Mr. & Mrs. R. Bendicksen
9049 Loyal Avenue N. W.
Seattle, WA 98117

A History of Engineering and Science in the Bell System

". . . Not only so, but, I believe, in the future, wires will unite the head offices of the Telephone Company in different cities and a man in one part of the country may communicate by word of mouth with another in a distant place."

Alexander Graham Bell

A History of Engineering and Science in the Bell System

The Early Years (1875–1925)

Prepared by Members of the Technical Staff,
Bell Telephone Laboratories.
M. D. Fagen, Editor.

Bell Telephone Laboratories, Incorporated

Credits for figures taken from
other than Bell System sources
appear on pp. 1043–1044.

Frontispiece: Alexander Graham Bell
in 1876. Photograph by E. A. Holton,
The Bell Family © National Geographic Society.

First Printing, 1975

Library of Congress Catalog Card Number: 75-31499

Printed in the United States of America

Contents

Foreword

The years 1974 to 1976 mark the first century of Alexander Graham Bell's invention and development of the telephone. The basic principle on which the telephone operates—the idea of an undulating current, the analog of a sound wave—was conceived in the summer of 1874 but not until a year later were sounds of a speech-like character heard over wires. On February 14, 1876, Bell applied for his first patent, which was granted on March 7. On the evening of March 10, he transmitted the first intelligible sentence. Many demonstrations of the new invention were conducted, probably none with so great an impact on the public feelings as that held in Philadelphia during the summer of 1876 at the exposition celebrating the centennial of American independence. This bicentennial year of our nation's independence therefore seems particularly appropriate for publishing a history of telephone communication, the growth of which has been so closely associated with the second century of our country's development.

This first volume of a series on the science and technology of telephony covers the half-century following Bell's invention. By the end of that time a great new industry had been developed. There were nearly seventeen million telephones in the United States, almost twelve million of them in the Bell System. And in perhaps no other field had the force of scientific research in support of engineering development been so effectively demonstrated.

The year 1975 marks another anniversary, the fiftieth year of the establishment of Bell Laboratories as the research and development unit for the Bell System. In 1925, it became a corporate entity, sharing responsibility with the Western Electric Company, the American Telephone and Telegraph Company, and the System's 24 Operating Telephone Companies for providing nationwide communications services and for planning, engineering, building, and operating the nationwide network.

The formal incorporation of Bell Laboratories was not a beginning of scientific research and engineering in the Bell System, but rather a stage in its growth in a line going back to Alexander Graham Bell's original laboratory in Boston. In 1907, a consolidation of the engineering forces took place in the Western Electric Company and in AT&TCo. As stated in 1925 by Mr. H. B. Thayer, then President of AT&TCo:

> The reorganization in 1907 consisted of a consolidation [whose] purpose was to avoid duplication of facilities as well as to get the greater efficiency coming from a closer contact between the staff of the Western Electric Company and our own. [He meant AT&TCo.] It brought to one point scientific study and research, manufacturing experience and operating ex-

perience . . . It simplified and expedited the work of the operating [telephone] companies in that it established one point where all statements of requirements, suggestions of improvements, or criticisms arising out of their operating experience could be considered and discussed from all points of view . . . It was helpful in the standardization of apparatus . . .

We had at the West Street laboratory (headquarters of the Western Electric Engineering Department) the scientists whose work involved laboratory facilities, the men conducting experiments, the shop design workers and the inspectors with suitable equipment of laboratories and model shops available for all. The ideas of our research and development scientists and engineers, worked out on paper or in rough mechanical form, were there developed into a finished piece for shop manufacture and after manufacture the product was then subjected to all the tests necessary to satisfy our engineers that it was worthy of introduction into or continuation in the plant of the Bell System.

This indicates clearly the important evolution in research and development management taking place in the Bell System early in the twentieth century. But as Thayer explains subsequently in the same article, it became evident by the early 1920s that even greater benefits could be expected by further centralizing research and development in a new organization, Bell Laboratories, working closely with the producers and users of communication systems:

Now the time has come when, it seems, we can take another step forward with advantage. What was contemplated in the reorganization of 1907 has been entirely accomplished, in that the development, research, and experimental work of the entire Bell System has been coordinated and has been concentrated as far as is desirable. The standardization of material is in effective operation. The different organizations composing the Bell System are working efficiently and harmoniously as parts of the greater organization, but this seems to be the time to get still more of the advantages in efficiency and economy which the consolidated organization now proposed makes possible.

This statement is an expression of what we have come to recognize as enduring themes in the maturation of technology in our business. First, there is reference to the economy and efficiency which comes through centralization of engineering and the standardization of systems which are meant to work together. Technical advance must be planned and orderly.

Second, there is the emphasis on technical integration, the need for intimate contact among the engineers and scientists working on all phases of a problem, extending from initial studies to the complexities of manufacture, installation, and use. And the latter specifically calls for close association of planners, designers, and the manufacturer with the operating entities of the organization.

Third, there is the emphasis on "scientific study and research," an idea which was quite new to industry in the early years of this century. The Bell System was one of a very few industrial organizations in which professionally trained scientists were doing basic research on fundamental problems related to company objectives. The experience of the intervening years has affirmed this belief in the values of research many times over, as, indeed, this history of communications will attest.

Finally, there is the important concept that our engineering and science through design, development, and manufacture be responsive to the needs of the final users—the Operating Telephone Companies who provide communication services to the public—and that innovations be "worthy of introduction into . . . the Bell System." The integrating influence of this ever-present goal—providing good telephone service to the ultimate consumer—has in a very real way tied together all the elements of the creative process throughout the Bell System over the 100 years of our history. We have no need to cite examples in this Foreword; the contents of this volume will do that in what I believe to be a most convincing fashion—by narration of a succession of technical triumphs unique in the history of industrial technology—producing what is generally acknowledged to be the most capable communications system in the entire world.

There is much to be learned here, as lessons from the past, true today as they were in the period 1875–1925, about the process of innovation as it really is. It was characterized then, as it is now, by continuity of technical activity from basic discoveries and inventions to direct operation in the communications network. We learned then and we know now, with a conviction born of experience, that conversion of new ideas, devices, and systems into something actually usable is a subtly demanding, personalized undertaking. The new concepts must be technically feasible and economically sound, they must fit into an existing operating plant, they must satisfy a real service need, they must be reliable and maintainable over a useful life of decades. And our long experience has demonstrated that this difficult task is unquestionably accomplished most effectively by the integration of technology through design, manufacture, and operation. In our semicentennial year, our intimate, daily associations with AT&TCo and Western Electric are at new, unsurpassed strengths.

Imbedded in the reality of technological innovation, but perhaps not as readily apparent as some of its other features, is the ever-present competition of ideas and approaches which exists in a large integrated structure such as ours. And it is the competition of ideas, rather than the competition of an undefined and arbitrary market place, that is the spur which really leads to technological progress. The reader of this volume will see it illustrated in the search by George Campbell for a better transmission line. He will see it a little later in the search by H. D. Arnold for an amplifier to implement the drive of T. N. Vail and John Carty for transcontinental telephone service; in the search by G. W. Elmen for a superior magnetic material for coils and transformers; in the quest by W. G. Houskeeper for a better glass-to-metal seal for high-power radio transmitting tubes that could generate the waves to carry the voice across the ocean.

It was the competition of ideas that stimulated their colleagues to press ahead, inventing oscillators, modulators, and wave filters that would permit the multiplexing of many conversations over a single circuit. And it was the same motivation that led to transmission of television signals by wire and radio over long distances, and to explanations of phenomena as fundamental as the complementary behavior of electrons and waves—a study which resulted in the award of the Nobel prize to C. J. Davisson.

There was a strong motivation then, as there is to this day, toward a well-defined goal—defined by the nature of the environment in which the work was done and by the searching spirit in a talented assemblage of technical experts who knew that what they were doing was relevant and what they produced would be useful.

In this, the hundredth year of the Bell System and the fiftieth year of Bell Laboratories, our modern world of communication is a world made up increasingly of digital signals and computers, of microwave radio and coaxial cable transmission systems, of satellites and broadband transoceanic cables, of electronic switching systems of unprecedented speed and versatility employing millions of transistor-like devices of microscopic dimensions. In the times ahead, we see millimeter waveguides, and optical fibers with capacities for conveying information at rates thousands of times greater than we know today, and especially automata enabling efficiencies and services making telecommunications a still greater frontier of human progress.

Those whose accomplishments are described in this volume could not foresee these things in detail, but they laid down the principles. It is clearly evident to us that the enduring themes in communication technology have not changed. Our incentives for acquiring new knowledge and our techniques for applying that knowledge to practical ends to satisfy human needs were right for our first fifty years and offer yet stronger opportunities for our latter fifty. Thus, this history is more than a mere record of past events. It provides an insight into the process of innovation and effective application of technology for beneficial purposes. Finally, it clearly demonstrates the intimate connection between the successes achieved by Bell System Operating Companies and the integrated structure they support for providing technological innovation, manufacture, and field application.

W. O. Baker
President,
Bell Telephone Laboratories

Acknowledgments

This volume is the work of many minds and hands, members of the Technical Staff of Bell Laboratories whose experience in communications research and engineering, going back more than 50 years, included personal acquaintance with outstanding individuals who went before them, as well as thorough knowledge of their published records.

The material in these pages was written by experts in their fields, individuals of proven technical competence who were, at the same time, in positions of administrative responsibility for planning and completion of technical projects. Thus it has been possible to achieve an account that goes beyond the simple narration of events and to deal with the "how" and "why," searching out the motivations and evaluating the long-range importance of the contributions to communication technology which are the substance of this history.

This is primarily an account of early Bell System achievements, but the authors have not been restricted by geographic or corporate boundaries and have recognized fundamental contributions originating outside the System. Nor has the treatment been constrained by strict time boundaries, since many programs initiated before 1925 were continued in the years following. Moreover, from our present position in the 1970s we can see, in retrospect, the formative stages of new ideas that were to have tremendous impact over long periods after 1925, some being even today the very essence of telecommunications.

Where outstanding developments and discoveries are clearly attributable to an individual, we have tried to give appropriate credit. The selection of a group of names is done with reluctance since others may well deserve equal mention. It is understandably impossible to list the hundreds who participated in the team effort responsible for so much of the technical advance made during these 50 years. We acknowledge the great debt owed to those dedicated workers whose creativeness, enthusiasm, and unselfish exchange of ideas with their fellows contributed so much to building up a completely new field of technology.

This history was initiated at the suggestion of James B. Fisk when he was President of Bell Laboratories. Most of the writing is the work of W. H. Doherty (Chapters 1, 2, and 10), and J. W. Emling (Chapters 3, 4, 5, and 6), both of whom also collaborated in the planning of the volume. The Non-Voice Communications chapter was prepared by F. J. Singer, the Materials and Components chapter is largely based on material submitted by A. G. Ganz and M. C. Wooley, and the Quality

Assurance chapter was primarily the work of L. N. St. James with some later assistance from G. R. Gause and G. J. Levenbach. Charles Breen supplied extensive sections on toll switching and PBX developments for the Switching chapter. To F. L. Crutchfield we owe thanks for his contributions on acoustics research and telephone instruments, material which was drawn upon in Chapters 3 and 10. F. A. Polkinghorn prepared an account of the history of radio communications in the Bell System which was helpful in preparing Chapter 5. Doren Mitchell provided helpful source material on speech privacy and radio control terminals for the same chapter. A. C. Keller wrote for us an authoritative account of sound recording and reproducing which was the basis for that portion of Chapter 4. S. O. Morgan's review of chemistry research was a valuable source document, as were supplementary writings by G. T. Kohman in this field. We are indebted also to G. E. Moore who provided us with much background material on physics, electronics, and materials research; to K. G. Van Wynen who did a preliminary paper on telegraph transmission systems; and to H. S. Renne who participated in the early stages of this work.

And, in the manner of all Bell Laboratories technical writing, the material, after preparation by experts, was reviewed and criticized by other experts who had established eminent reputations in their specific fields. In this respected list are E. I. Green, Ralph Bown, Lloyd Espenschied, W. H. Martin, R. D. Parker, H. T. Friis, A. E. Joel, B. D. Holbrook, G. D. Edwards, H. F. Dodge, E. G. D. Paterson, C. E. Fisher, and W. S. Hayward.

In addition, there were many who provided us with personal recollections in interviews and by correspondence, generously giving of their time in discussing and commenting on this history, and giving us access to the knowledge and experience of those who had been closer to the events described. Thanks are therefore due to M. J. Kelly, R. V. L. Hartley, H. W. Bode, W. C. Jones, A. W. Kishpaugh, M. B. McDavitt, A. C. Dickieson, J. C. Schelleng, S. P. Shackleton, E. L. Norton, P. W. Blye, R. M. Bozorth, H. A. Burgess, P. H. Evans, A. G. Hall, J. R. Hefele, J. G. Knapp, J. M. Duguid, W. Keister, J. W. Smith, E. A. Veazie, G. T. Ford, L. S. Inskip, J. F. Wentz, L. S. Hulin, L. J. Cobb, K. E. Hammer, R. R. Stevens, F. F. Romanow, H. E. Curtis, W. C. Babcock, H. C. Hart, J. B. Johnson, F. A. Korn, J. O. McNally, R. I. Wilkinson, A. A. Oswald, A. C. Peterson, J. R. Pierce, C. F. Wiebusch, J. R. Wilson, R. M. Burns, and Pierre Mertz. Regrettably, many in this list are no longer here to note our appreciation, but we ask the reader to think of this book as a memorial to their considerable achievements.

There are those who should be recognized for their part in producing the printed work: Miss R. L. Stumm, who worked tirelessly

on the typed script and searched the files for bibliographic material and illustrations; Mrs. Y. H. Quick of the Western Electric Company Historical Library, who provided us with reference documents from her archives; Mrs. V. B. Graeper of the AT&TCo Editorial Research Center, who was equally helpful with historical references from the Headquarters collection; B. A. Stevens, who compiled the Index; and R. E. Gillis and his associates in the Bell Laboratories Technical Publication Department, who took a manuscript and made a history book.

M. D. Fagen
Editor

Chapter 1

Bell's Telephone

The invention of the telephone was not inspired by a pre-existent popular demand. Rather, it came about largely through the ingenuity and vision of one man—Alexander Graham Bell. His belief that there was a great potential need for two-way voice communication over a distance, a need of which few men had been conscious, was confirmed by its immediate success and spectacular growth in spite of early technical limitations.

Bell's concepts were unusual in other ways also. Even before a means had been found for accomplishing his objective, he had a clear picture of the theoretical requirement, namely, a device for translating speech waves into analogous electric waves. As with many inventions, luck played a part—in this case in the form of a malfunctioning telegraph device with which he was experimenting. As a result of his clear understanding of the fundamentals of telephony, Bell recognized immediately the significance of this fortuitous event and, ably assisted by Thomas Watson, constructed within 24 hours the first device to convey speech-like sounds by electrical means.

Improvement in this first wave converter came rapidly and intelligible speech was transmitted about nine months later; the first commercial application occurred less than two years after the accident which led to the invention of the telephone instrument. But Bell's vision far transcended these early uses and even as these simple steps were being taken, he outlined his "Grand System" for a nationwide network for interconnecting any two users for voice communication wherever they might be. That plan has served as the broad telephonic objective ever since.

The events leading up to this "Grand System" are related in this chapter; subsequent chapters deal with developments in specific areas that led toward its realization.

I. INTRODUCTION

During the three decades preceding the invention of the telephone, telegraphy had made great strides in most countries of the world and had established a new tempo in business, in the conduct of government, in the functioning of railways and in the collection and dissemination of news. Continents had been interconnected, after great initial difficulties, by undersea cables. The telegrapher with his key and sounder was the respected operative of a miraculous vehicle for speeding important messages, both social and business; and the magnanimity of nature in permitting intelligence to be propagated with the speed of light had been gratefully accepted, if only superficially understood.

To an inquiring mind like Bell's, however, even an already established facility held intriguing scientific aspects to challenge the imagination, leading him as a very young man to experiment with telegraph apparatus. But more than this, his primary dedication to the study of speech and hearing made him particularly conscious of the basic shortcoming of the telegraph: it did not lend itself to man's natural way of communicating, the undelayed two-way vocal exchange. Indeed, to his observant mind, the cumbersome and time-consuming operations at the sending and receiving terminals of the telegraph undoubtedly were made more conspicuous by the instantaneity of actual transmission over the wires.

Thus was Bell, through his sensitivity to the human factor as much as his scientific insight, marked by destiny to conceive the instrumentality that would leap across this operational barrier and quickly outstrip the telegraph in popularity and usefulness.

History shows the conceptual steps beginning in the town of Brantford in Ontario, Canada, to which Bell had moved from England with his parents in 1870 when he was 23 years old. He had shortly thereafter taken up residence in Boston and had become Professor of Vocal Physiology at Boston University, but was spending his summer vacation in 1874 in Brantford at his father's home.

II. HARMONIC TELEGRAPH AND "ELECTRIC SPEECH" EXPERIMENTS

For several years Bell had been interested in the multiple telegraph (sometimes referred to as the harmonic telegraph), a scheme which hopefully would permit a number of telegraph messages to be sent simultaneously over a single wire by means of interrupted tones of different frequencies. Much of his spare time was devoted to working on this device. However, this activity was somewhat of a sideline, for he was primarily interested in developing techniques for teaching the deaf to speak and to read lips.

While engaged in these activities concerning the deaf, Bell had become familiar with a device known as a manometric capsule, invented by Koenig, in which a gas flame was made to vibrate by the action of the voice (Fig. 1-1). The voice, impinging on a membrane, produced a flickering which could be viewed as a continuous, wavering band of light in a revolving mirror. Bell felt that if he could reproduce this wavering band of light in some manner, he could teach his deaf students to speak by comparing their attempts at speech with a previously recorded pattern. However, the difficulties of photographing the band of light prevented the manometric capsule from becoming useful for this purpose.

Another voice-actuated instrument in which Bell had become interested was the phonautograph of Scott. This device consisted of a conical mouthpiece; a stretched membrane; a long, light lever of wood attached to the membrane; and a bristle or stylus on the end of the lever. A plain sheet of glass covered with lampblack was so arranged that when a sound was uttered into the mouthpiece and its vibrations transmitted by the stretched membrane to the wooden lever, the stylus wobbled up and down, tracing its motion on the lampblack. The sheet of glass was moved along at a uniform rate, recording the vibrations thus produced. Typical tracings of this kind are shown in Fig. 1-2.

In the course of his work Bell compared the output of the phonautograph with that of the manometric capsule for the same vowel sound and found that the two outputs did not match at all. This caused him to start searching in other directions for a possible teaching aid for the deaf. However, while conducting these experiments, Bell was struck by the similarity of the phonautograph to the human ear. He felt that a phonautograph modeled after the structure of the human ear probably would produce more accurate tracings of speech vibrations than the imperfect instrument with which he had been working.

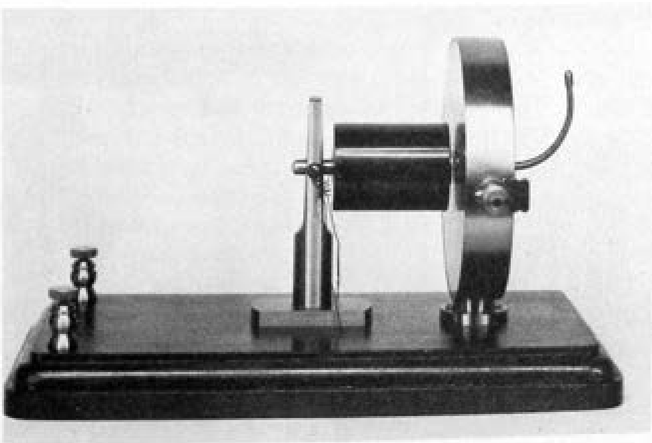
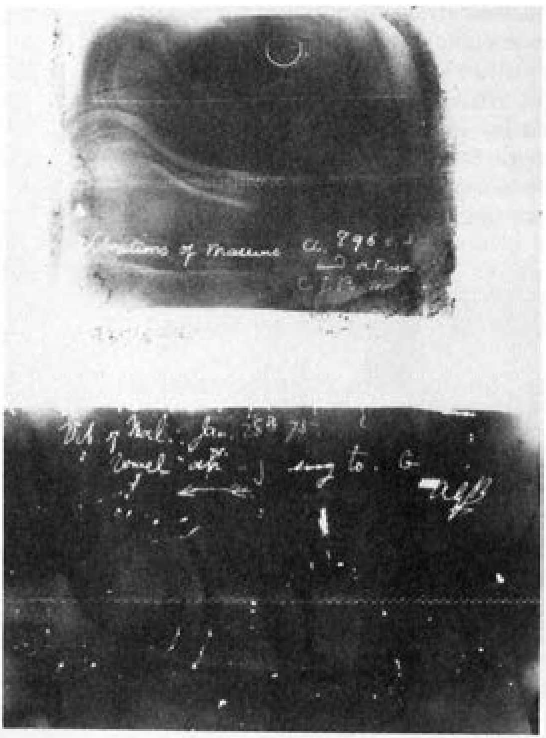


Fig. 1-1. Model of Koenig manometric capsule which Bell displayed at the Centennial Exposition in 1876. In this version, the diaphragm is actuated by an electromagnetic coil through which voice currents flow. In the unit used by Bell in 1874, the voice was directed to the diaphragm by a speaking tube.

To obtain expert advice on reproducing the structure of the human ear, Bell had called on an old acquaintance, Dr. Clarence J. Blake, who was later to become famous among American otologists. Blake had suggested, "Why not take an ear from a dead man and get tracings from the little bones of the ear?" He had also offered to obtain such an ear from the Harvard Medical School, and Bell had accepted his offer.

Bell carried this human ear with him to Brantford and during the summer of 1874 fashioned a phonautograph from it (Fig. 1-3). He moistened the eardrum with glycerin and water to make it flexible and used a small piece of hay for the bristle of the phonautograph. When he spoke into the ear, the piece of hay vibrated in accordance with the impinging sound. He then arranged to move a piece of smoked glass in such a manner that the piece of hay made tracings of the vibrations of the eardrum. Tracings of the type shown in Fig. 1-2 were obtained.



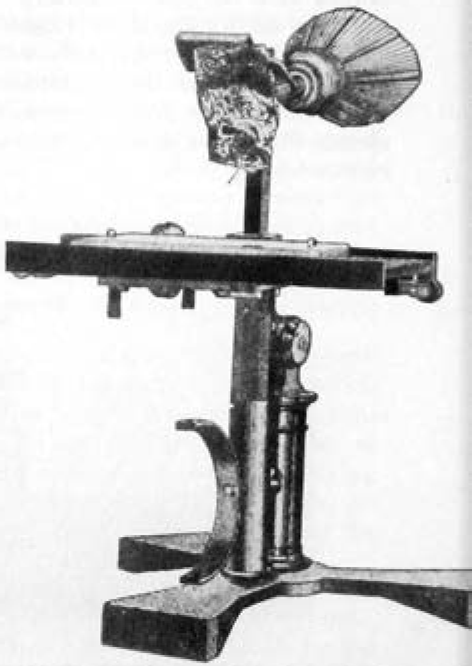


Fig. 1-2. Smoked-glass tracings of vowel sounds obtained from a phonautograph.
Fig. 1-3. Bell's human-ear phonautograph with which he experimented in Brantford, Ontario, in 1874.

These early experiments were of great significance to the future development of the telephone. Later, in 1916, when referring to this early work, Bell made the following remarks:

> Now it so happened that while I was experimenting with this human ear, I was at work on a very different problem. I was at work on a problem of transmitting musical sounds by a telegraphic instrument, by an intermittent current of electricity, and I had dreams that we might transmit the quality of a sound if we could find in the electrical current any undulations of form like these undulations we observe in the air.
>
> I had gradually come to the conclusion that it would be possible to transmit sounds of any sort if we could only occasion a variation in the intensity of the current exactly like that occurring in [the] density of the air while a given sound is made . . .
>
> I had obtained the idea that theoretically you might, by magneto electricity, create such a current. If you could only take a piece of steel, a good chunk of magnetized steel, and vibrate it in front of the pole of an electromagnet, you would get the kind of current we wanted . . . It struck me that the bones of the human ear were very massive, indeed, as compared with the delicate thin membrane that operated them, and the thought occurred that if a membrane so delicate could move bones relatively so massive, why should not a thicker and stouter piece of membrane move my piece of steel. And the telephone was conceived.

Bell then added the following statement:

> The conception of the telephone took place during that summer visit to my father's residence in Brantford, in the summer of 1874, and the apparatus was just as it was subsequently made, a one-membrane telephone on either end.

Bell returned to Boston in the fall of 1874 with his human-ear phonautograph and with many new ideas on the possibilities of "electric speech" as well as for his multiple telegraph. He communicated these ideas to a friend, Thomas Sanders. Sanders was somewhat startled by the concept of sending speech over a wire but shared Bell's enthusiasm about the possibility of the multiple telegraph. He offered to support Bell's telegraph work but suggested that the work on the transmission of speech might well be delayed since the multiple telegraph seemed to him to offer much greater opportunity for financial gain.

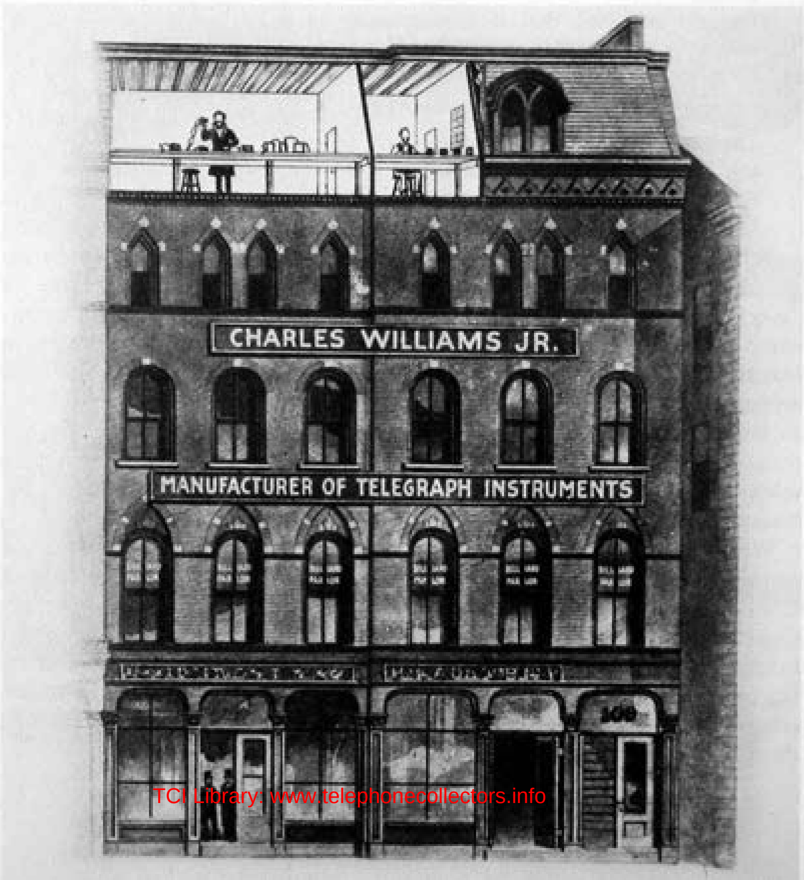
While experimenting with his multiple telegraph, Bell had become acquainted with a young machinist, Thomas A. Watson, at the shop of Charles Williams, Jr., a manufacturer of telegraph apparatus at 109 Court Street, Boston. This young man had been assigned to make much of the equipment requested by Bell for his experiments. During the late fall and winter of 1874, Bell spent more and more time on his telegraph experiments as well as more and more time with Watson designing and building new apparatus.

One evening, early in 1875, after a particularly discouraging session with the multiple telegraph, Bell said to Watson, "Watson, I have another idea I haven't told you about that I think will surprise you. If I can get a mechanism which will make a current of electricity vary in its intensity as the air varies in density when a sound is passing through it, I can telegraph any sound, even the sound of speech."

Later, in his autobiography, Thomas A. Watson made the following comments about this early conception of Bell:

> History gives us many illustrations of the transforming power of an idea but Bell's conception of a speech-shaped electric current ranks among the most notable of them. The conception itself was the great thing and any mechanism embodying it, even the very first form that was discovered, is of minor importance. If Bell had never found the apparatus for which he was searching to produce his ideal current, his name should have been immortalized.

Fig. 1-4. Birthplace of the telephone (109 Court Street, Boston) where Bell carried out his early experiments.



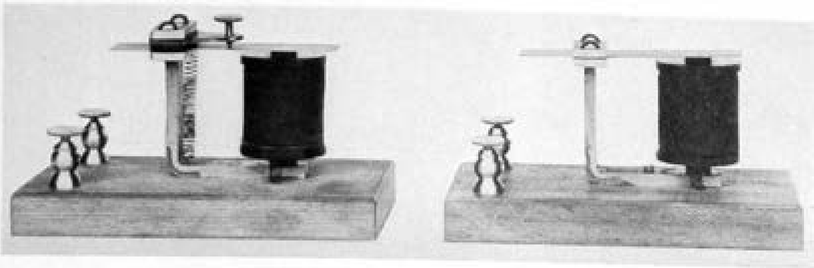


Fig. 1-5. Bell's multiple or harmonic telegraph apparatus. A reed transmitter is at the left and a reed receiver at the right.

III. FIRST SPEECH SOUNDS

During the early months of 1875, Bell and Watson spent long hours trying to perfect their multiple telegraph system. Experiments were conducted in the top story of the Charles Williams building, where Bell had obtained the use of two rooms with a wire line connecting them (Fig. 1-4). The apparatus consisted of tuned vibrating reeds at the transmitting end and similar tuned reeds at the receiving end (Fig. 1-5). Theoretically, a signal sent by means of one vibrating reed should be received only by the reed tuned to the corresponding frequency. By using different frequencies, it should be possible to send several telegraph messages simultaneously over a single wire. The long, arduous hours which Bell and Watson spent with this system were consumed primarily in attempts to tune the transmitting and receiving reeds to the same frequency. The transmitting reeds operated on the interrupted-current principle and both the transmitting and receiving reeds were tuned by clamping them at different points along their length. In the tuning process Watson would adjust one of the transmitting reeds while Bell would hold the receiving reed to his ear.

There are a number of accounts as to exactly what happened on the memorable day of June 2, 1875, but Watson's own words tell the story dramatically (Fig. 1-6):

> On that hot June day we were in the attic, hard at work experimenting with renewed enthusiasm over some improved piece of the apparatus. About the middle of the afternoon, we were retuning the receiver reeds, Bell in one room pressing the reeds against his ear one by one as I sent him the intermittent current of the transmitters from the other room. One of my transmitter reeds stopped vibrating. I plucked it with my fingers to start it going. The contact point was evidently screwed too hard against the reed and I began to readjust the screw while continuing to pluck the reed when I was startled by a loud shout from Bell and out he rushed in great excitement to see what I was doing. What had happened was obvious. The too-closely adjusted contact screw had prevented the battery

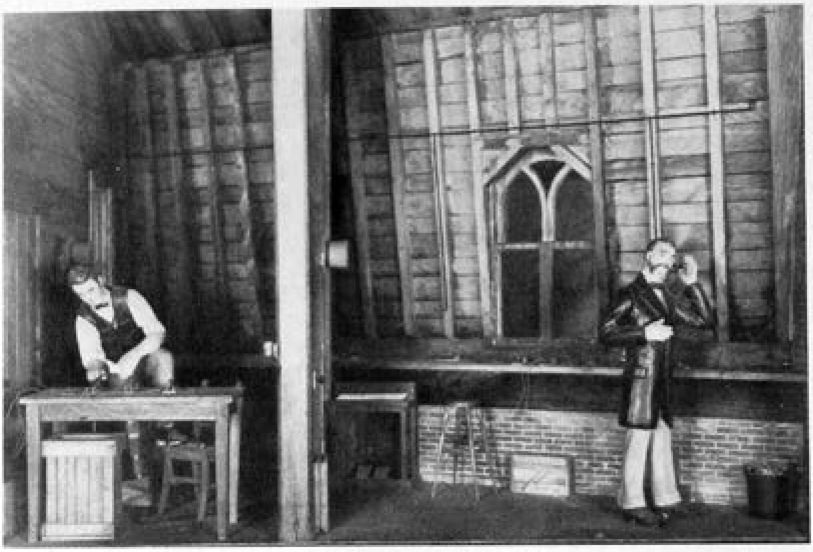


Fig. 1-6. Diorama of the "attic" scene of June 2, 1875, when the first speech sounds were heard.

> current from being interrupted as the reed vibrated and, for that reason, the noisy whine of the intermittent current was not sent over the wire into the next room, but that little strip of magnetized steel I was plucking was generating by its vibration over the electromagnet, that splendid conception of Bell's, a sound-shaped electric current.
>
> We spent the rest of the afternoon and evening repeating the discovery with all the steel reeds and tuning forks we could find and before we parted, late that night, Bell sketched for me the first electric speaking telephone, beseeching me to do my utmost to have it ready to try the next evening. And, as I studied the sketch on my way to Salem on the midnight train, I felt sure I could do so.

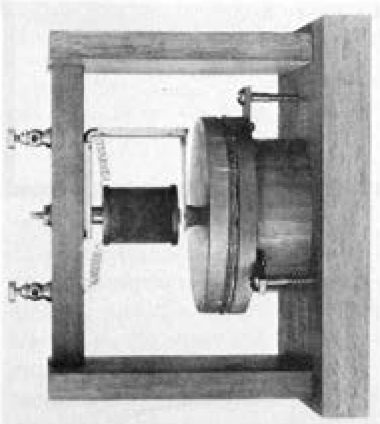
That first attempt at a telephone was a simple mechanism and has become known as the "gallows" telephone (Figs. 1-7 and 1-8). It consisted of a wooden frame on which was mounted one of Bell's harmonic receivers, a tightly stretched parchment drumhead to the center of which the free end of the receiver reed was fastened, and a mouthpiece arranged to direct the voice against the other side of the drumhead. It was designed to force the reed to follow the vibrations of the voice and so generate voice-shaped electric undulations.

This device was tested on June 3, 1875, using one of the harmonic telegraph receivers for listening, and although no intelligible words were transmitted, nevertheless speech sounds were heard. This indicated that Bell was on the right track. The sounds transmitted by means of this device have been considered the first voice waves ever transmitted by electricity.

IV. FIRST ELECTRIC SPEAKING TELEPHONE

The significant events of June 2 and 3, 1875, were followed by long months of tedious experimenting in an attempt to improve the device to the point where intelligible speech could be transmitted, using a second stretched-membrane instrument as a receiver. Bell was sick for several weeks during this period and also spent a great deal of time in preparing the patent application for the electric speaking telephone. The application was completed late in 1875 but filing was delayed because Bell wanted to file in England and the United States simultaneously. A friend of his from Brantford had promised to file the application in England but had not done so because of fear of ridicule. Finally, in desperation, Gardiner Greene Hubbard, one of Bell's financial supporters and his future father-in-law, filed the U.S. patent application on February 14, 1876. The application was allowed on Bell's birthday, March 3, and U.S. Patent No. 174,465 was issued to him on March 7, 1876, only about three weeks after it was applied for. The famous Fig. 7 of this patent is reproduced in Fig. 1-9 along with the cover sheet for the patent.

In January 1876, Bell set up shop at 5 Exeter Place, Boston, about a half-mile from the Williams shop where Watson was employed. He



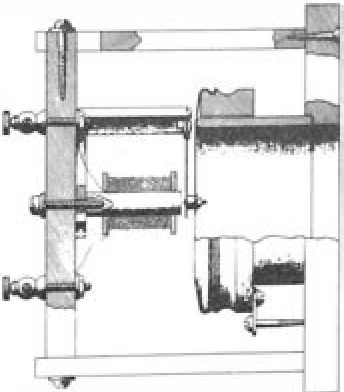


Fig. 1-7. Model of Bell's first telephone—the "gallows" telephone, so called because of the shape of its frame. Fig. 1-8. Cross section of the "gallows" telephone.


No. 174,465

The United States of America

TO ALL TO WHOM THESE PRESENTS SHALL COME:

Whereas Alexander Graham Bell, of Salem, Massachusetts

has presented to the Commissioner of Patents a petition praying for the grant of LETTERS PATENT for an alleged new and useful

Improvement in Telegraphy

a description of which invention is contained in the Specification of which a copy is hereunto annexed and made a part hereof, and has complied with the various requirements of Law in such cases made and provided, and

Whereas upon due examination made the said Claimant is adjudged to be justly entitled to a Patent under the Law,

Now therefore these LETTERS PATENT are to grant unto the said Alexander Graham Bell, his heirs or assigns for the term of Seventeen years from the seventh day of March one thousand eight hundred and seventy six the exclusive right to make, use and vend the said invention throughout the United States and the Territories thereof.

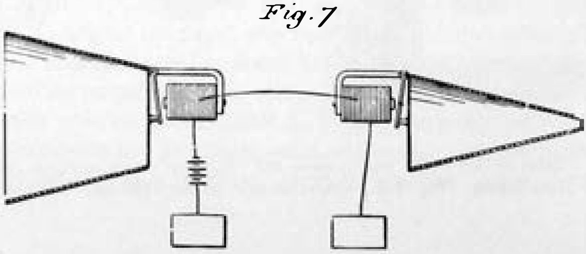
In testimony whereof I have hereunto set my hand and caused the seal of the Patent Office to be affixed at the City of Washington this seventh day of March in the year of our Lord one thousand eight hundred and seventy six and of the Independence of the United States of America the One hundredth.

Secretary of the Interior

Commissioner of Patents

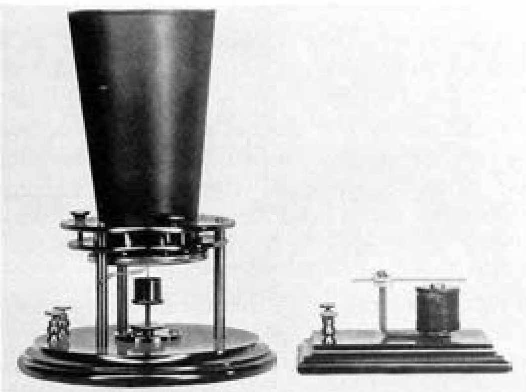
Fig. 1-9. Cover sheet for Bell's historic patent No. 174,465, which was allowed on his twenty-ninth birthday, March 3, 1876, and issued on March 7. A reproduction of Fig. 7 of Bell's first patent, showing his conception of the electric speaking telephone.

Fig. 7



had rented two rooms in the attic of a boarding house at that address for four dollars a week. Bell slept in the front room and fitted up the back room as a laboratory, with a wire running between the two rooms. Much of Bell's early experimental work was carried on in these two rooms. The move was made because of rumors he had heard of strangers visiting the Williams shop and examining his apparatus with curious eyes.

Soon after this transfer, Bell found time to continue some of his telegraph experiments, especially those on means for quenching the sparks at the contacts of the telegraph key. For this purpose he devised a variable water resistance to bridge the contact points. It was this work that undoubtedly suggested the first form of variable-resistance telephone transmitter, the so-called "liquid" transmitter.



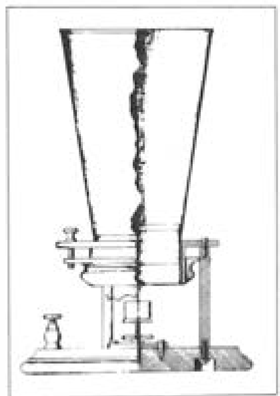


Fig. 1-10. Liquid transmitter and tuned-reed receiver used in the experiments of March 10, 1876. Fig. 1-11. Cross section of the liquid transmitter exhibited by Bell at the Philadelphia Centennial Exposition in 1876.

In the early spring of 1876, Bell had designed and Watson had built this new type of transmitter (Figs. 1-10 and 1-11) in which a wire attached to a diaphragm was inserted into acidulated water contained in a metal cup, both of which were included in a circuit through the battery and the receiving telephone. The resistance of this circuit was varied as the voice made the diaphragm vibrate and caused the wire to move up and down in the acid. Thus, the battery current was forced to undulate in speech form. On the evening of March 10, 1876, the cup of the liquid transmitter was filled with diluted sulphuric acid and it was connected to the battery and, via the wire running between the two rooms at No. 5 Exeter Place, to a tuned-reed receiver. Again, Watson's own words serve to tell the story dramatically:

> When all was ready I went into Bell's bedroom and stood by the bureau with my ear at the receiving telephone. Almost at once I was astonished to hear Bell's voice coming from it distinctly, saying, "Mr. Watson, come here. I want you." We had no receiving telephone at his end of the wire so I couldn't answer him, but as the tone of his voice indicated he needed help, I rushed down the hall into his room and found he had upset the acid of a battery over his clothes.

On reaching Bell's room, Watson exclaimed, "Mr. Bell, I heard every word you said distinctly." Elated over the success of this test, the accident of the spilled acid was forgotten. Both men immediately recognized the historic significance of this first complete intelligible sentence to be transmitted by electricity, for it represented the birth of the electric speaking telephone.

V. EARLY DEVELOPMENT

Soon after his initial success with the liquid transmitter, Bell replaced the tuned-reed receiver with one consisting of a cylindrical iron box with a central core around which a coil of wire was placed. The central core constituted one pole of the magnet and the rim of the iron cylinder the other pole. A sheet of iron formed the lid on one end of this magnet and served as the diaphragm. This design worked much better as a receiver than the tuned-reed arrangement, and was even used to some extent as a transmitter. It was called the iron-box receiver (Fig. 1-12).

Also developed at about the same time was a membrane transmitter (Figs. 1-12 and 1-13) using a non-magnetic diaphragm to which a piece of iron was attached. Motion of this membrane in the magnetic field produced by a coil caused the iron to induce currents in the coil in

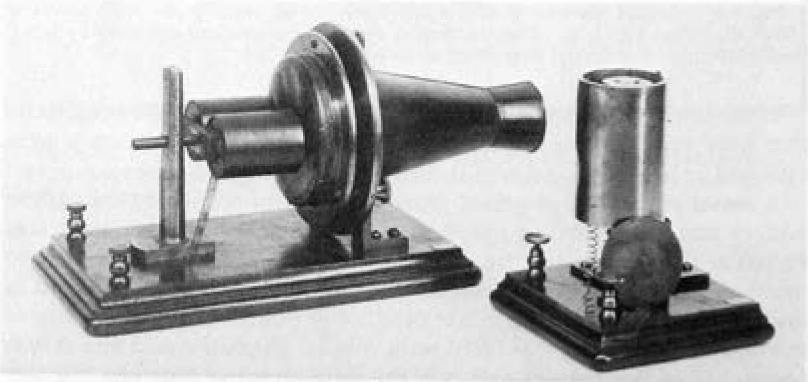


Fig. 1-12. Membrane transmitter and iron-box receiver demonstrated at the Centennial Exposition.

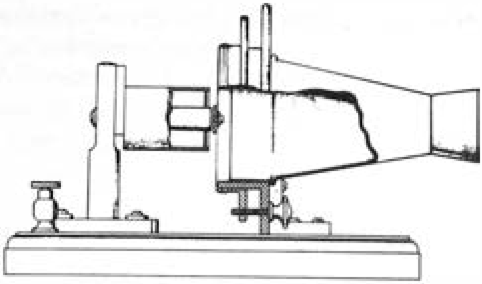


Fig. 1-13. Cross section of the Centennial membrane transmitter.

accordance with the membrane's motion. These currents then produced a corresponding motion in the diaphragm of the receiver.

This telephone system was described and demonstrated before the American Academy of Arts and Sciences in Boston on May 10, 1876, and about two weeks later before the Society of Arts at the Massachusetts Institute of Technology.

In the summer of 1876 the Philadelphia Centennial Exposition was held, commemorating the one-hundredth anniversary of American Independence. One of Bell's backers, Gardiner Hubbard, was a Centennial Commissioner and obtained a small table in the Department of Education portion of the Exposition, on which Bell displayed some of his apparatus. Included were two membrane transmitters, a liquid transmitter, and an iron-box receiver (Figs. 1-11 thru 1-13). In the latter part of June, Bell went to Philadelphia to demonstrate this apparatus to the Exposition judges.

On June 25, 1876, using the membrane transmitters and the iron-box receiver, Bell showed that intelligible speech could be transmitted over wires. This feat made a tremendous impression on those present, among whom were a number of famous scientists, including Elisha Gray who was later to enter into patent litigation with Bell, and Sir William Thomson, the British scientist later known as Lord Kelvin. The success of this demonstration gave a great boost to the telephone, and provided the stimulus for renewed development efforts.

The liquid transmitter was not used in the Centennial demonstrations and was apparently abandoned from this time on, as no further references to it have been discovered.

VI. EARLY DEMONSTRATIONS AND TESTS

During the summer of 1876, Bell again spent some time in Brantford, Ontario, at his father's home. While there he continued his telephone experiments and staged some significant demonstrations. In all of these, the membrane instrument was used as a transmitter and the iron-box apparatus as a receiver. This equipment was not designed to transmit and receive conversations simultaneously, so in these early demonstrations messages went in only one direction.

In some of the early demonstrations a triple mouthpiece, which Bell had designed for the membrane transmitter, was used. He felt that it was much more effective to show that two or three voices could be heard distinctly at the same time than to have it inferred that the electric current could carry only one voice distinctly. For these demonstrations a wire was run from the veranda of the house in Brantford to one of the outbuildings. Everyone took a turn, first at the mouthpiece and then at the receiver. When no one was available to share the trials, Bell strung a wire around the eaves of the house and sat telephoning to himself in his own room.

Another significant test which was carried out in early August was described by Bell, "Articulate speech was, for the first time, transmitted and received between places that were separated by miles of space." This test took place between Brantford and the little town of Paris about 8 miles away. Arrangements were made to use the telegraph wires of the Dominion Telegraph Company between the Brantford office and the Paris office. The triple-mouthpiece membrane transmitter was connected at Brantford and the iron-box receiver at Paris.

When the equipment was first connected there was a storm of bubbling and crackling sounds, but through these sounds, in a faint, faraway manner, voices could be heard from the transmitter at Brantford. In this first trial, low-resistance voice coils were used in both the transmitter and receiver. Bell, who was at Paris, telegraphed Brantford to substitute high-resistance coils and Bell did the same at his location. When this was done, the voices, and even singing, came through very clearly. The instruments used for these trials were very similar to those used at the Centennial Exposition.

In the Brantford–Paris test, electromagnets were used in both the transmitter and receiver, thus requiring a source of battery power. The

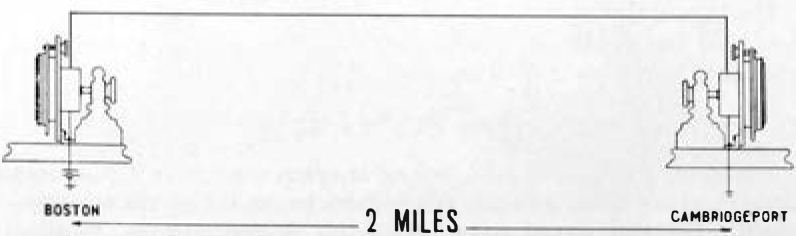


Fig. 1-14. Circuit used in Boston–Cambridgeport test.

battery for these tests was in Toronto, some 68 miles distant, and was connected to the instruments by means of telegraph wires.

Following this significant test there was a public reception and demonstration at the Bell home outside Brantford. For this demonstration, ordinary iron stovepipe wire was run from the Bell home to the telegraph office in Brantford, a distance of about one-quarter mile. From there the circuit was completed to Mount Pleasant, about 5 miles away, then back to Brantford, and to one of the outbuildings at the Bell home. Successful demonstrations of speech and singing were carried out, but only a brief note on these tests appeared in the *Toronto Globe* for August 11, 1876. However, the success of the experiments provided added encouragement to Watson and Bell for continuing their task of improving the telephone.

It was about this time that Hubbard offered Watson a one-tenth interest in the Bell patent if he would give up his job at the Williams shop and devote all of his time to making apparatus for Bell. Although pleased at this recognition, Watson was somewhat reluctant to accept since he had a well-paying job with Williams and the future of the telephone at that time was very uncertain. However, he finally accepted and moved into an attic room adjoining Bell's in the lodging house at 5 Exeter Place.

VII. CONTINUED TESTS

The tests and demonstrations at Brantford involved telephone transmission in one direction only. The next major forward step, two-way communication, was taken on October 9, 1876, after Bell had returned to Boston. The telegraph line used for this two-way test was owned by the Walworth Manufacturing Company and ran from their office in Boston to their factory in Cambridgeport, a distance of about 2 miles (Fig. 1-14). The results were very satisfactory, sustained two-way conversation between persons miles apart being carried on upon the same line using the same instruments alternately for talking and listening. These instruments were membrane telephones similar to the Centennial transmitters except that upon each membrane was glued a thin, circular sheet of iron almost as large in diameter as the membrane itself (Fig. 1-15). Comparisons by Bell and Watson the previous July had shown that the larger the sheet of iron the louder and more distinct was the tone.

To verify the accuracy of transmission in this significant test Bell and Watson each kept a record of the words transmitted and received. The *Boston Daily Advertiser* in its issue of October 19, 1876, published these two records in parallel columns to show that most of the conversation was transmitted with substantial accuracy. Two days after performing this test Bell communicated the results to the American Academy of

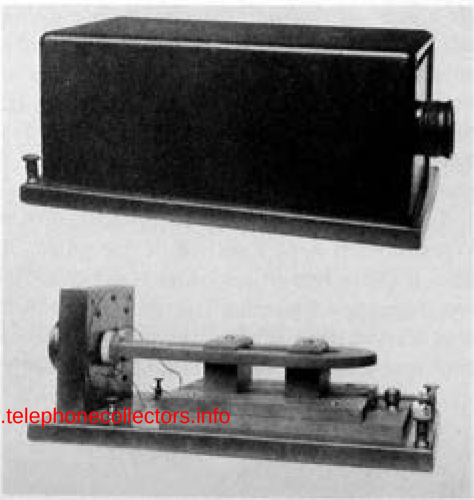
Arts and Sciences and exhibited in operation the telephones that had been employed.

The next step in improving the telephone was to dispense with the membrane diaphragms entirely and use diaphragms consisting solely of thin iron. These worked well and the construction was incorporated in the so-called "box" telephones (Fig. 1-16), the first of which was made in the latter part of October. The circuit employed with these telephones was a single wire with a ground return.

Progress in improving the telephone had been slow and had involved a great amount of hard work. Both Bell and Watson were alert to any new developments which might be adapted to improve its operation. Watson in particular had been spending a great deal of time in the local library reading everything he could lay his hands on about electricity in an attempt to find ideas that would be useful.

In the course of his library research he came across some information about a quick-acting magnet used by the Hughes Printing Telegraph Company. Watson dashed back to the laboratory and within an hour or two had constructed a magnet similar to the one about which he had been reading. This was a permanent magnet made up of four hardened-steel horseshoe-shaped plates bolted together, with a small soft-iron core, carrying coils for the voice currents, clamped to each pole. When mounted with the diaphragm and mouthpiece already developed, it became an improved instrument that functioned much better than any other telephone tried up to that time with either electromagnets or permanent magnets. From then on, all telephones requiring electromagnets went into the discard.

Fig. 1-15. Membrane telephone used for transmitting and receiving in Boston–Cambridgeport test. Fig. 1-16. Exterior and interior views of early box telephone.



A major test of the improved instrument (the so-called "magneto" telephone) was then made over a line of greater length than any previously employed. This test was conducted on Sunday, November 26, 1876, over a telegraph line of the Eastern Railroad running between Boston and Salem, a distance of about 16 miles. The test was highly successful, even a whisper or a loud breath being heard distinctly at the other end of the line.

Immediately thereafter an attempt was made to converse over a circuit about 200 miles in length. The voice could be heard "with considerable clearness," but sufficient "distinctness" was not attained to permit a normal two-way conversation to be carried on. However, on the Sunday following the Boston–Salem experiment, successful tests were conducted over a railroad wire between Boston and North Conway, New Hampshire. This was the first instance of a human voice being carried between points physically separated by more than 100 miles.

Early in 1877, a second patent was issued to Bell, covering the structural aspects of the magneto telephone in some detail.[1] This, together with his first patent in 1876, formed the basis for the Bell System monopoly in the telephone field which prevailed until the patents expired in 1893 and 1894.

VIII. LECTURES

Following these successful tests, Bell was invited to deliver a series of lectures explaining the telephone and demonstrating it in operation. The first such lecture was given in Lyceum Hall, Salem, on February 12, 1877 (Fig. 1-17) and, for the occasion, Bell's laboratory in Boston was connected with Lyceum Hall by a wire of the Atlantic and Pacific Telegraph Company. Watson spoke in Boston and his speech was audible to the audience in Salem. For this demonstration, permanent-magnet telephones with metallic diaphragms (Fig. 1-16) were employed and there was no battery in the line.

Reporting the event in its issue of February 13, 1877, the *Boston Globe* stated:

> This special by telephone to the Globe has been transmitted in the presence of about twenty who have thus been witnesses to a feat never before attempted—that is, the sending of a newspaper despatch over the space of eighteen miles by the human voice—and all this wonder being accomplished in a time not much longer than would be consumed in an ordinary conversation between two people in the same room.

This item served to wake newspaper editors with a jolt as they suddenly realized the tremendous possibilities inherent in transmitting newspaper "despatches" by voice.

[1] A. G. Bell; U.S. Patent No. 186,787; filed January 15, 1877; issued January 30, 1877.

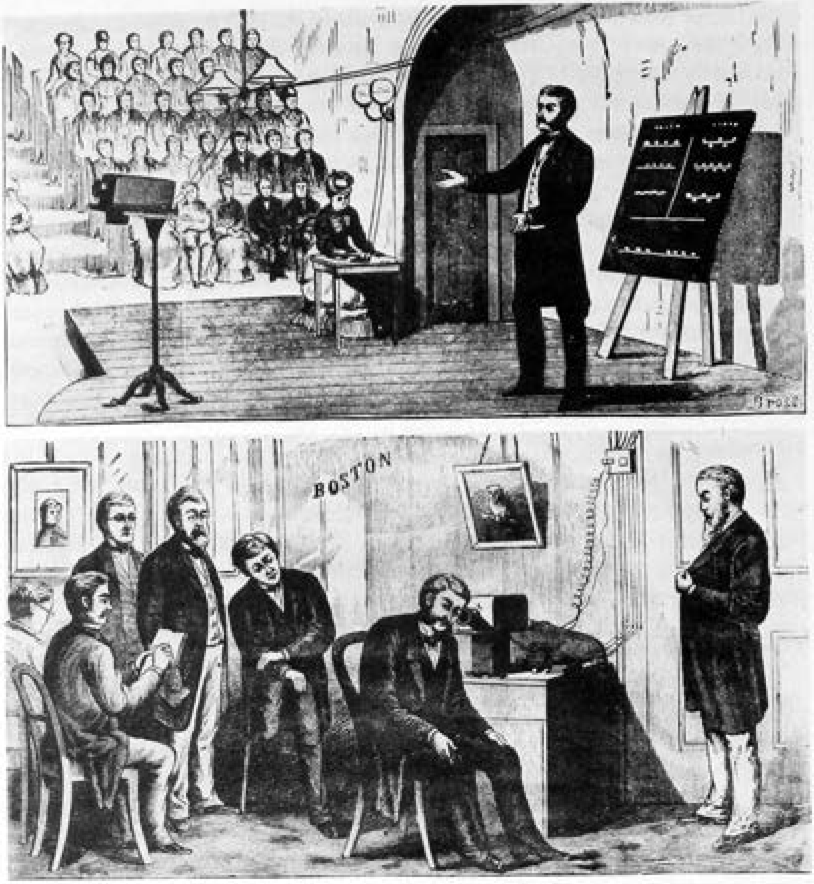


Fig. 1-17. Artist's sketches of lecture of February 12, 1877: (top) Lyceum Hall, Salem; (bottom) Bell's Boston laboratory.

Shortly afterward the first wire line constructed for regular telephone use was installed between Charles Williams' shop and his house in Somerville, a distance of about 3 miles. The inauguration of this line, a milestone in telephone history, took place on April 4, 1877.

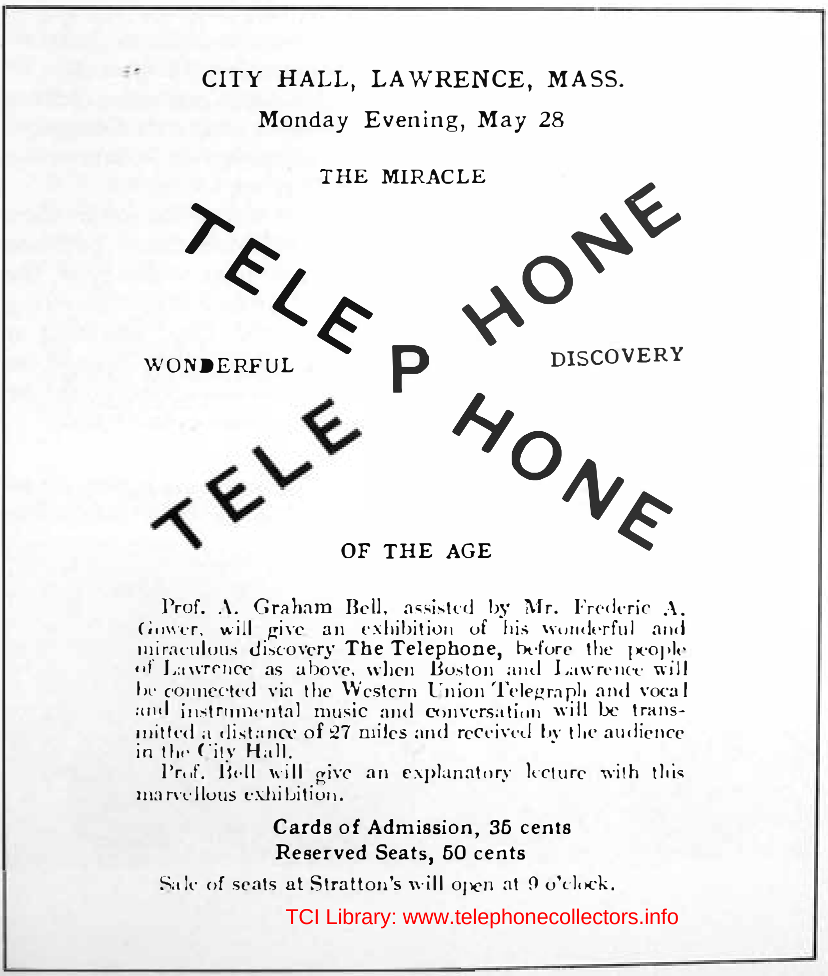
On April 5, another lecture was given in Music Hall, Providence, Rhode Island, before an audience of about 2,000 people. Telephones were arranged in different parts of the hall so that Bell, with the aid of a switch on the platform, could connect any one of them as desired. Again, the demonstration was highly successful.

In some of these early lectures, Watson would use the laboratory at 5 Exeter Place as the sending station and transmit voice and singing to the lecture hall where Bell was appearing. Bell, a master showman, cleverly arranged for Watson to sing as well as shout, for no matter how poor the telephone quality might be, the pitch was faithfully retained, and in recognizing the tune, people would think they had understood the words. Figure 1-18 shows the publicity for one of Bell's demonstrations.

Early in May, Bell delivered a series of three lectures in Chickering Hall in New York City. For these lectures, Watson's voice came over a telegraph line from New Jersey. According to Watson's memoirs, he and Bell had been discouraged by earlier tests from trying to stage a performance in New York City with Watson talking from Boston.

These lectures gave Bell his first monetary return from the telephone. He was very badly in need of the money and this income gave

Fig. 1-18. Facsimile of the publicity for one of Bell's demonstrations.



CITY HALL, LAWRENCE, MASS.

Monday Evening, May 28

THE MIRACLE

TELEPHONE

WONDERFUL P DISCOVERY

TELEPHONE

OF THE AGE

Prof. A. Graham Bell, assisted by Mr. Frederic A. Gower, will give an exhibition of his wonderful and miraculous discovery **The Telephone**, before the people of Lawrence as above, when Boston and Lawrence will be connected via the Western Union Telegraph and vocal and instrumental music and conversation will be transmitted a distance of 27 miles and received by the audience in the City Hall.

Prof. Bell will give an explanatory lecture with this marvellous exhibition.

Cards of Admission, 35 cents
Reserved Seats, 50 cents

Sale of seats at Stratton's will open at 9 o'clock.

him further encouragement to continue his development work. For example, during May he developed his receiving instrument into a form that began to resemble the modern receiver in external appearance. One early form of hand-held receiver was called the "butterstamp" due to its resemblance to that implement of the dairy (Fig. 1-19).

IX. FIRST COMMERCIAL APPLICATIONS

Also during May 1877, E. T. Holmes, who operated a burglar-alarm system in Boston, became interested in the telephone after seeing a demonstration. His interest was natural since he already had customers connected to his place of business by wires not generally in use in the daytime. Holmes ordered several instruments from Bell, the first three of which were of the box type (Figs. 1-20 and 1-21), numbered 6, 7, and 8. These were put in use late in May (less than 15 months after the first transmission of intelligible speech) for sending messages, requesting messenger or express service, between Holmes' burglar-alarm customers and his office on Washington Street (Fig. 1-22). To distinguish this business from his burglar-alarm system, Holmes conducted it under the name of the Telephone Despatch Company. Although this was a very restricted use of telephony, it is fair to say that it represented the first commercial application.

Holmes later broadened his use of telephony to include the modern concept of interconnecting lines for direct communication between telephone users. While he had made some early tests of this type, the credit for the first commercial application of the basic idea underlying present-day telephony belongs to George W. Coy. Coy, operating as the District Telephone Company of New Haven, formally opened on January 28, 1878, an "Exchange" with 21 customers who could be interconnected as they desired by means of a central switchboard.

Fig. 1-19. First four commercial hand-held receivers. All were made in 1877, and are arranged in chronological order from left to right. The butterstamp receiver is at the left.

X. THE "GRAND SYSTEM"

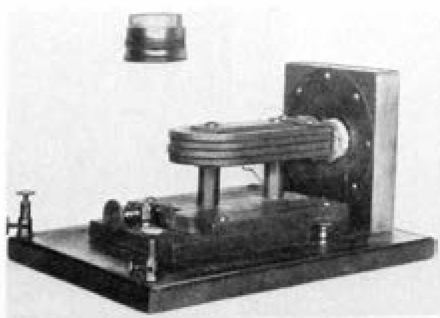
These first faltering steps in applying telephony showed a recognition of its value, but fell far short of Bell's ideas.

On July 11, 1877, Alexander Graham Bell married Mabel, the daughter of Gardiner Greene Hubbard. They sailed for Europe early in August and remained abroad for more than a year. Even though far removed from telephone activities, Bell's ingenious mind was actively at work. While he was in England, Bell prepared a prospectus designed to awaken the interest of a group of businessmen in promoting the use of the telephone. This document is quoted here substantially in its entirety since it indicates the remarkable prophetic vision of Bell at a time when telephony as a business was as yet undeveloped:

> The telephone may be briefly described as an electrical contrivance for reproducing in distant places the tones and articulations of a speaker's voice, so that conversation can be carried on by word of mouth between persons in different rooms, in different streets, or in different towns.
>
> The great advantage it possesses over every other form of electrical apparatus consists in the fact that it requires no skill to operate the instrument. All other telegraph machines produce signals which require to be translated by experts, and such instruments are therefore extremely limited in their application, but the telephone actually speaks, and for this reason it can be utilized for nearly every purpose for which speech is employed . . .
>
> At the present time we have a perfect network of gas-pipes and water-pipes throughout our large cities. We have main pipes laid under the streets communicating by side pipes with the various dwellings, enabling the members to draw their supplies of gas and water from a common source.

Fig. 1-20. First commercial box telephone. This is type used by E. T. Holmes in his early telephone "despatch" service.

Fig. 1-21. Early box telephone with cover removed.



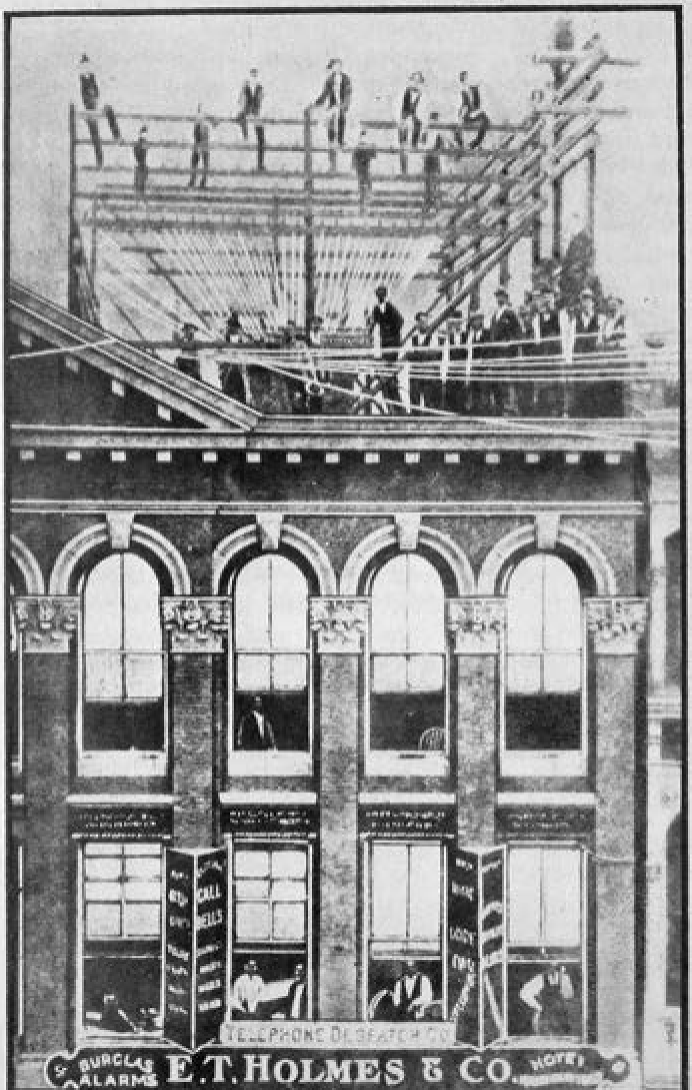


Fig. 1-22. Home of the first telephone company, the Telephone Despatch Company, at 342 Washington Street, Boston.

In a similar manner, it is conceivable that cables of telephone wires could be laid underground, or suspended overhead, communicating by branch wires with private dwellings, country houses, shops, manufactories, etc., etc., uniting them through the main cable with a central office where the wires could be connected as desired, establishing direct communication between any two places in the city. Such a plan as this, though impracticable at the present moment, will, I firmly believe, be the outcome of the introduction of the telephone to the public. Not only so, but I believe, in the future, wires will unite the head offices of the Telephone Company in different cities, and a man in one part of the country may communicate by word of mouth with another in a distant place.

I am aware that such ideas may appear to you Utopian and out of place, for we are met together for the purpose of discussing not the future of the telephone, but its present.

Believing, however, as I do, that such a scheme will be the ultimate result of introducing the telephone to the public, I will impress upon you all the advisability of keeping this end in view, that all present arrangements of the telephone may be eventually realized in this grand system . . .

In conclusion, I would say that it seems to me that the telephone should immediately be brought prominently before the public, as a means of communication between bankers, merchants, manufacturers, wholesale and retail dealers, dock companies, water companies, police offices, fire stations, newspaper offices, hospitals and public buildings, and for use in railway offices, in mines and [diving] operations.

Agreements should also be speedily concluded for the use of the telephone in the Army and Navy and by the Postal Telegraph Department. Although there is a great field for the telephone in the immediate present, I believe there is still greater in the future.

By bearing in mind the great object to be ultimately achieved, I believe that the Telephone Company can not only secure for itself a business of the most remunerative kind, but also benefit the public in a way that has never previously been attempted.

XI. CONCLUSION

Sensing that they were on the threshold of a new art of breathtaking possibilities, Alexander Graham Bell and Thomas A. Watson experienced alternate periods of high excitement and deep despair. As we have seen, there were discouragements, hard work, elation at successes (particularly those of June 2, 1875, and March 10, 1876), and then much more hard work and many obstacles to overcome before the telephone could be made practical. For the first few years the major technical load was carried by Watson, who took the telephone as Bell had invented it and toughened it into a rugged, usable instrument, acquiring in the process some 60 patents on the improvements he devised.

During this initial period the need for technological innovation became apparent and the power of organized scientific effort was sensed. The remaining chapters of our history, with one exception, will be devoted to relating the technical activities of the first 50 years of telephony and describing the many contributions made to this new art during the period. But before starting this major portion of our history, we shall cover briefly, in Chapter 2, the early corporate history of the Bell System and the part played by organizational structures in achieving Bell's "Grand System" through raising the necessary capital, promoting technological development, and integrating development, manufacture, and system operation into an effective organization with centralized direction, yet retaining a high degree of autonomy within the component elements.

Chapter 2

Early Corporate History

The immediate success and spectacular growth of the telephone as a local facility quickly led its backers to concur with the concept of a "universal service" as already envisioned by its inventor.

The technological barriers to this ultimate aim were formidable indeed and it became clear to the early managers that their science-based enterprise could not rely on fortuitous advances, coming from outside the industry, as the chief engine of progress. Accordingly they adopted, as an industrial innovation, a deliberate policy of supporting the pursuit of scientific knowledge to speed advancement in telephone technology.

As the technical imperatives of the expanding telephone network clarified, this farsighted policy became implemented by specific organizational arrangements, insuring a continued synergy of business planning and advancing technological skills.

This chapter presents a brief sketch of these organizational structures as they developed; for, as later chapters proceed to describe scientific and technological achievements, some of the more triumphant of these (such as the first transcontinental line) must be seen also as triumphs in organization and management.

I. CORPORATE ORGANIZATIONS

1.1 Bell Patent Association

Alexander Graham Bell, though primarily a teacher of the deaf, was a man of extraordinary scientific curiosity and insight. As noted in Chapter 1, during the time that his early ideas on telephony were developing, he was also experimenting with a type of multichannel telegraph. Thomas Sanders, a leather merchant of Salem, Massachusetts, who was a friend of Bell's and the father of one of his former pupils, had been following Bell's work and felt that the telegraph experiments promised more immediate prospect of success than the then somewhat vague ideas on telephony. Accordingly, in the fall of 1874 he made a verbal offer to Bell to help finance the telegraph work in return for a share in whatever patent rights might result. A short time later, another friend, Gardiner G. Hubbard, a Boston attorney and, as it developed, Bell's future father-in-law, made a similar offer. These informal offers were accepted by Bell, and later a written agreement was drawn up and dated February 27, 1875. This was the first, rudimentary corporate agreement, and, although the organization had no official name, it has been called the Bell Patent Association.

The interests of Sanders and Hubbard being confined to Bell's telegraph work, the written agreement of this Association did not speak of the telephone or refer to it in a distinctive way. However, Bell later took the position that the telephone was to be included despite Hubbard's offer to relinquish all right and title to that invention even after it had been successfully demonstrated. "Truly," writes telephone historian Frederick Leland Rhodes, "this original partnership was a gentlemen's agreement!" Thus, the Bell Patent Association became the first legal instrument of corporate telephone ownership and organization. (Figure 2-1 is a historical chart of Bell System corporate structures from 1875 to 1935.)

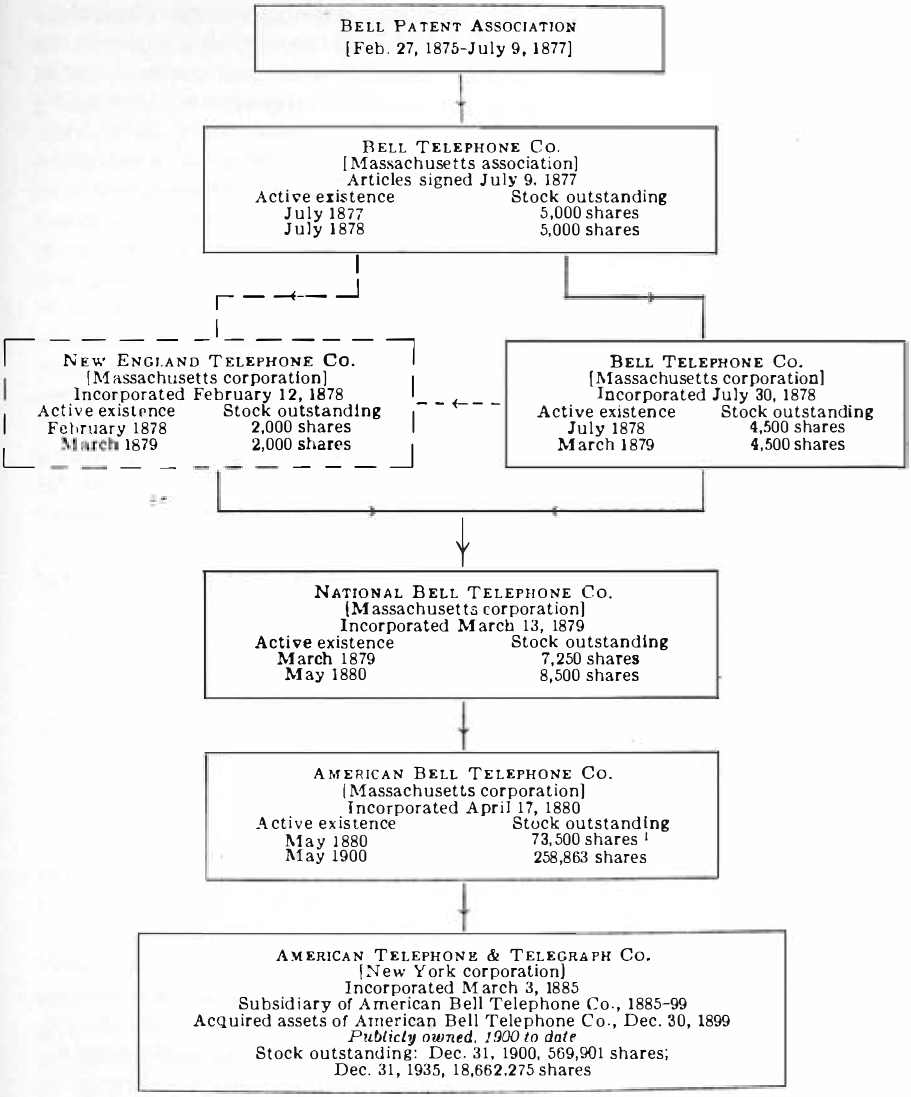
1.2 Bell Telephone Company Trusteeship

The original agreement provided that if any of the inventions of the Association proved to be of value, a company should be organized to manage and control the patents involved, and each of the three Association members should own one-third of the stock of that company. By the summer of 1877 it was recognized that Bell's basic telephone patent was very valuable, and steps were taken to organize a company in accordance with the agreement.

To provide the original Association with necessary technical skills, Thomas A. Watson, Bell's assistant, had been invited to become associated with it in a formal way, and to receive a one-tenth interest in all patents when the joint-stock company was organized. Watson had agreed, and a formal contract to this effect had been drawn up and dated September 1, 1876. This step had been taken by the three Association members "to associate with them a practical mechanician of sufficient skill and ability, under Mr. Bell, to make these inventions pecuniarily successful."

The new organization which came into existence on July 9, 1877, was given the name "Bell Telephone Company, Gardiner G. Hubbard, Trustee," and superseded the original Bell Patent Association. Sanders was Treasurer of this trusteeship, Bell was Electrician, and Watson was Superintendent.

Thomas Watson made many distinctive contributions to the enterprise. He was in effect the only member of the first research and



Legend: ———► Transfer of assets.
— —→— — - Licensee company.

NOTE.—All stock of corporations is $100 par.

[1] Includes 14,000 shares of trustee stock held by National Bell Telephone Co.

Fig. 2-1. Historical chart of the parent organizations of the Bell System.

development organization in the telephone industry. Watson had little to do with purely business affairs, but until his resignation in 1881 he carried the burden of the necessary research work and superintended manufacturing operations, thus playing a very large role in developing Bell's invention into a practical device with commercial possibilities. Although he had no technical training, he had to try to solve all of the baffling technical problems that arose in telephone construction and operation. This responsibility rested on Watson because Bell, having married Hubbard's daughter in July 1877, had sailed for Europe in August and had remained abroad for more than a year lecturing and promoting the telephone in England. Consequently, in four overworked years, during which he was called upon for many nontechnical tasks as well, Watson was the first consistent contributor to a research and development program that has continued for a century to solve the problems of telephony and advance communications technology.

Bell and Watson were both restless men with eager, probing minds, always ready to explore new horizons. By 1881, the success of the telephone had made both of them financially secure, and they both left the telephone business to embark on new ventures.

One of the legacies left by this early trusteeship was the leasing and licensing system, inaugurated by Hubbard and stubbornly maintained by him in the face of strong opposition from his associates. The Declaration of Trust in the original plan of organization provided for this policy as follows:

> The business of manufacturing telephones and licensing parties to use the same for a royalty, shall be carried on and managed by the Trustee, under the name of the Bell Telephone Company, under and in accord with such general directions, rules and regulations as may be made for that purpose by the Board of Managers.

This policy has been an essential element in maintaining the nationwide unity and efficiency of the Bell System.

1.3 New England Telephone Company

During these early days, money for the Bell Trusteeship was very scarce, and most of it was provided by Thomas Sanders. Having almost reached the limit of his resources, Sanders now cast about for new sources of capital and succeeded in interesting a group of Massachusetts and Rhode Island businessmen in the telephone. This group agreed to put up some money in return for permission to start a telephone business, but wished to confine its interests and responsibility to the New England area. Naturally, these men wished exclusive rights within that territory, and the members of the Bell Trusteeship agreed to this arrangement.

Accordingly, the New England Telephone Company came into existence on February 12, 1878. (This company, however, had no direct connection with the present New England Telephone and Telegraph Company, which was not formed until October 1, 1882, nearly five years later.) The New England Telephone Company held an assignment of rights to the Bell patents for the New England states, and its Articles of Incorporation committed it to the policy of leasing and not selling telephones. The Articles further stated that the corporation was formed "for the purpose of carrying on the business of manufacturing and renting telephones and constructing lines of telegraph therefor, in the New England States."

Earlier organizations operating in this area under Bell License, such as the Telephone Despatch Company in Boston and the District Telephone Company of New Haven (see Section IX of Chapter 1), were ultimately absorbed into the New England Telephone Company.

1.4 Bell Telephone Company

The success of the New England Telephone Company suggested establishment of a similar organization for the rest of the country. Money was needed and presumably could be obtained by such a corporation. Therefore, an Agreement of Association for a new company to be organized along these lines was signed on June 29, 1878, and a Certificate of Incorporation filed on July 30, 1878. This was an entirely new corporate organization, called simply the Bell Telephone Company. To transfer control of the telephone to the new company, Gardiner Hubbard assigned "all the patents, patent rights and interest in any and all contracts relating to any patents or future inventions owned or held by him as Trustee" to the Bell Telephone Company. Thus came to an end the organization known as "Bell Telephone Company, Gardiner G. Hubbard, Trustee," its place being taken by the new Bell Telephone Company.

Another significant step was taken by Hubbard at about this time. While serving a term in the United States Congress, he became acquainted with Theodore N. Vail, who at the time was Superintendent of the Railway Mail Service, and prevailed upon him to be General Manager of the Bell Telephone Company. Vail proceeded to bring order out of a considerable amount of organizational chaos, and encouraged Bell agents throughout the country to persevere in the face of severe competition from Western Union. Vail was to become one of the great figures in the Bell System—a master of organization, a leader of limitless courage and resources, unselfish and utterly honest in his management of the company's affairs. He contributed mightily to the System's success.

1.5 National Bell Telephone Company

During this early phase of corporate organization, William H. Forbes, a prominent Boston financier and son-in-law of Ralph Waldo Emerson, joined the group of Boston investors who had combined to provide financial support for the Bell interests. On December 31, 1878, Forbes was elected a Director of the Bell Telephone Company. He immediately demonstrated exceptional business ability and, with his associates, in effect took over control of the Company's affairs.

On January 29, 1879, Forbes took steps toward uniting all of the Bell interests in one company, to be called the National Bell Telephone Company. At a meeting of the Board of Directors of this new company on March 11, 1879, he was elected President. Two days later, the Certificate of Incorporation for the National Bell Telephone Company was filed. Then, on March 20 of the same year, both the New England Telephone Company and the Bell Telephone Company assigned their rights under the two basic Bell patents to this new company, and these two earlier companies were consolidated.

On the technical side, Thomas D. Lockwood joined the National Bell Telephone Company on July 26, 1879, as Assistant Inspector and Electrician. He functioned as an assistant to Thomas Watson in exercising general supervision, planning exchange apparatus, and writing pamphlets. He was later to assume responsibility for patent matters.

In 1879, the following force was at work looking after the technical problems of the telephone business: Thomas A. Watson, General Inspector; Lockwood, just mentioned; George L. Anders, an inventor; and Emile Berliner, another inventor, who had two assistants, Joseph H. Cheever and W. L. Richards (later head of the Bell System Historical Museum at 463 West Street in New York City). This last trio worked on inspection and adjustment of telephone instruments.

1.6 American Bell Telephone Company

In the late 1870s and early 1880s, the spread of the telephone over the United States was spectacular. The method of business organization at this time usually involved local companies which leased their instruments from the parent corporation. However, it soon became evident that mere leasing was not sufficient for effective coordination with the parent unit. Competition was becoming very strong at about this time, with many rival companies springing up. Thus it was necessary for the parent corporation to license local companies on an exclusive basis and, more than that, to gain control of these companies by buying a sufficient portion of their stock. This required far more capital than that for which the National Bell Telephone Company had been authorized ($850,000).

The most serious competitor at this time was the Western Union Telegraph Company. In 1876, this company had declined to purchase the basic Bell patent, offered for the sum of $100,000, the view being that the telephone was a toy and would never be of practical use. However, after witnessing the rapid growth of interest in the telephone, Western Union decided to take it more seriously.

In December 1877, the American Speaking Telephone Company was formed with Western Union owning two-thirds of the stock. By this time, the Bell Trusteeship had over 3,000 telephones in use. The Western Union affiliate immediately commenced making telephones in violation of the Bell patents and leasing these instruments in competition with the Bell Trusteeship and its assignees. This caused consternation in the Bell organization, and, shortly after the formation of the Bell Telephone Company in the summer of 1878, suit was brought against the Western Union Telegraph Company for infringing Bell patents. The actual suit was instituted against Peter A. Dowd, an agent of Western Union engaged in supplying telephones in Massachusetts. The Western Union subsidiaries involved in this action were the Gold and Stock Telegraph Company, the American Speaking Telephone Company, and the Harmonic Telegraph Company.

After the facts had been fully brought out, George Gifford, then counsel for Western Union, became convinced that the Bell patents were valid and advised a settlement. Negotiations were begun, and a settlement was reached which became effective on November 10, 1879, and covered a period of 17 years. Western Union agreed to withdraw from the telephone business and the National Bell Telephone Company (successor to the Bell company that instituted the suit) was granted a license to the telephonic inventions that Western Union had acquired or might acquire during the term of the agreement. The National Bell Telephone Company was to buy the telephones which Western Union or its subsidiaries had made, already totalling over 50,000, and the telephone exchanges they had established. In return, the Bell company agreed not to compete with Western Union in the public message-telegraph field. The final decree was approved by Judge Lowell on April 4, 1881. This marked the end of the Western Union dispute, but did not mark the end of the lawsuits engaged in by the Bell System to protect its patent rights against others.

The tremendous increase in business resulting from the settlement of the Western Union case, and the huge amounts of capital which were necessary to buy the Western Union telephone equipment, made it necessary to consider another reorganization. To permit this reorganization on an adequate scale and to permit the new corporation to hold stock in other corporations, the Massachusetts Legislature passed a special Act which was signed by Governor Long on March

19, 1880. The purpose as stated in this Act was to permit the "manufacturing, owning, selling, using and licensing others to use electric speaking telephones and other apparatus and appliances pertaining to the transmission of intelligence by electricity, and for that purpose constructing and maintaining by itself and its licensees public and private lines and district exchange." Capitalization was limited to $10,000,000.

In accordance with this Act, the American Bell Telephone Company was formed on March 20, 1880, and its Certificate of Incorporation was filed on April 17, 1880. To put into the hands of the new corporation all patent rights belonging to the Bell interests, assignments were made and licenses issued to the American Bell Telephone Company by all parties having any shadow of right or ownership.

On December 8, 1880, the American Bell Telephone Company declared its first dividend (3%), and on March 29, 1881, issued its first annual report.

1.7 Western Electric Company

Prior to 1878, all of the telephone equipment for the Bell organization was built by Thomas Watson in the Boston shop of Charles Williams, Jr., a maker of telegraph instruments. Watson's duties soon increased to the point where he could no longer build all of the equipment needed so he engaged a number of manufacturers, including Williams, to assist him. While this was going on, Western Union was having some of its telephone equipment manufactured in its own shops, and some by the Western Electric Manufacturing Company of Chicago.

With the November 1879 agreement by Western Union to relinquish its telephone interests, the Western Electric Manufacturing Company was licensed to manufacture telephones for the Bell organization. Soon thereafter, the American Bell Telephone Company purchased a controlling interest in Western Electric, and on November 26, 1881, Western Electric was reorganized under the laws of the state of Illinois with a new name, the Western Electric Company. On February 6, 1882, the Western Electric Company officially became the manufacturing unit of the Bell System, thus providing a dependable source of instruments and apparatus which would be interchangeable and would be of the desired quality.

An item in the Annual Report of the American Bell Telephone Company for 1881 describes this significant event in Bell System history:

> To obtain a permanent interest in the manufacture of telephones and apparatus, as well as to ensure the highest standards in the same, we have bought the plant and business of Charles Williams, Jr., of Boston, and an

> interest in the Western Electric Manufacturing Company, of Illinois, and propose to merge the two in a consolidated company, which will avail of the goodwill, business, and patents owned by that company, as well as our own, and secure an economical management for the whole of our manufacturing interests. We expect to make this an important and valuable part of our business.

The first part of the building at 463 West Street, New York City, was erected in 1897, and served as Western Electric's New York headquarters. This building was later (1907) to become the headquarters for the Western Electric Engineering Department. While Western Electric remained an Illinois corporation for many years, it was reincorporated in 1915 under the laws of the state of New York and has been a New York corporation ever since.

Western Electric had become interested in the sale and manufacture of communications equipment abroad a number of years before it became the manufacturing arm of the Bell System, and its foreign business continued for many years thereafter. A factory was built in Antwerp in 1882, and branches of Western Electric existed throughout the world by 1918. At that time, foreign business was so extensive that it was concentrated in a new subsidiary, the International Western Electric Company. This organization grew rapidly, but at the same time enormous growth was taking place in the United States and Bell management felt that it would be desirable to avoid any possibility that these foreign activities would interfere with the domestic business. Accordingly, in 1925, the International Western Electric Company was sold to the International Telephone and Telegraph Company and its name changed to International Standard Electric Corporation (neither of these companies had any further corporate connection with the Bell System). As a result of this policy decision, subsequent Bell System international business activities have been limited to those required for forming and operating a worldwide communications network. This international cooperation has been established largely by means of bilateral agreements with the connecting national telecommunications administrations. Wherever possible, construction of plant has been conducted as a joint enterprise with each administration providing its proper share of the equipment.

1.8 Long Lines System

The system of local exchanges, coordinating any number of telephone subscribers' lines and referred to as the "exchange system," had become well established by 1880. The problem now arose as to how to interconnect these various exchanges. The first major step in this direction was taken on June 2, 1880, when a telephone line from Boston to New York was authorized (this line was put into service on

March 27, 1884). On May 9, 1883, lines were authorized between Boston, New York, Philadelphia, and Washington, and between New York and Albany. These events marked the beginning of the Long Lines System.

The maximum capitalization allowed the American Bell Telephone Company was $10,000,000, which was quite inadequate for constructing lines on the large scale rapidly becoming necessary for interconnecting the exchanges of various cities. The Massachusetts Legislature refused to increase the authorized capitalization, making it necessary to organize a new company for building long lines.

1.9 American Telephone and Telegraph Company

The new company organized to provide these interconnections was named the American Telephone and Telegraph Company, and was chartered under the laws of the State of New York. Its purpose was to construct and operate lines all over the North American continent, not only in the United States but in Canada and Mexico, "and also by cable and other appropriate means with the rest of the known world." Thus, Bell's concept of a "Grand System" was expanded and formally stated as the goal of the Bell System. The word "Telegraph" was incorporated into the name by an Act of the New York Legislature. At that time the word "Telegraph" referred to the leased-wire business of the Company. In the early days, telegraphy and telephony were not dissociated in quite the same way that they are today; in fact, Bell's first patent (see Fig. 1-9) had been described by the inventor as an "Improvement in Telegraphy."

Signing of the Articles of Association for the new company took place on February 28, 1885. The company headquarters was in New York City, and it immediately started to build up its own Long Lines Engineering Department there to assist in carrying out various projects. Theodore N. Vail, who previously had been General Manager of the National Bell Telephone Company, was President of this new organization until he resigned in 1887. He was to serve another term as President from 1907 to 1919.

The company grew rapidly, and by 1899 it was deemed best that it should be made the central organization of the Bell System, primarily for financial reasons. (At the time the financial climate in New York State was much more favorable than in Massachusetts.) Accordingly, on December 30, 1899, the Directors of the American Bell Telephone Company conveyed all their assets, both in stock and property, to the American Telephone and Telegraph Company. Thus was formed the corporation which today is the parent company of the Bell System.

To handle the Company's long-distance lines, a Department of

Long-Distance Lines was established to start functioning on October 1, 1900. Then on November 14, 1917, its name was changed to the Long Lines Department.

A statement of the relationship of AT&TCo to the Bell System as a whole was contained in the AT&TCo Annual Report for 1908. Despite its length, substantial parts of the statement are quoted below because of their clarity and their pertinence to the "Grand System" concept:

> The relations of the American Telephone and Telegraph Company and the associated companies are not generally understood. The American Telephone and Telegraph Company is primarily a holding company, holding stocks of the associated operating and manufacturing companies. As an operating company it owns and operates the long-distance lines, the lines that connect all the systems of the associated operating companies with each other.
>
> In addition to these two functions it assumes what might be termed the centralized general administrative functions of all the associated companies . . .
>
> In the telephone business development is continuous. As conditions enlarge and change, new methods develop. The whole business suggests changes and stimulates inventions, and opportunities for improvements are frequent.
>
> If each separate exchange or group of exchanges had not been assisted and directed in the development and introduction of these new ideas, methods and inventions, there would now be as many systems, as many methods of operating as there are separate companies. This would have made impossible the organization which now gives the Bell system that universality and preponderance on account of which no matter how many systems may exist, every one of any commercial or social importance must have connection with the Bell system . . .
>
> The American Telephone and Telegraph Company owns and maintains all telephones. It also owns either directly or through the Western Electric Company all patents.
>
> It has a department which was organized at the very beginning of the business and has continued since, where is to be found practically everything known about inventions pertaining to the telephone or kindred subjects. Every new idea is there examined, and its value determined so far as the patent features are concerned.
>
> The Engineering Department takes all new ideas, suggestions and inventions, and studies, develops, and passes upon them.
>
> It has under continuous observation and study all traffic methods and troubles, improving or remedying them.
>
> It studies all construction, present and future development or extension schemes, makes plans and specifications for the same, and gives when desired general supervision and advice. It has a corps of experts which, in addition to the above work, is at all times at the service of any or all of the separate companies.

> When it is considered that some of these questions involve the permanency, duration and usefulness of a telephone plant costing millions of dollars, and changes costing hundreds of thousands, some idea of its importance can be formed. To give an illustration: one group of patents covering inventions which seemed likely to be useful and economical in the service was purchased by the company. These inventions were developed into operating apparatus and put into use. While this cost hundreds of thousands of dollars, placing it beyond the scope of one operating company, the saving already accomplished to the associated companies runs into the millions.
>
> A large staff has been and is continuously engaged in the consideration of disturbances arising from transmission and other lines carrying heavy currents, and in many cases that any telephone system can even exist in the vicinity of such lines is due to the constant and continued attention given this subject.
>
> Every new trouble, and there are many, comes before this department. When settled there, it is settled for all. This has established a commercial, operating and plant practice not only for our own associated companies, but for others of high standing throughout the world.
>
> All devices or inventions submitted receive the most thorough and painstaking investigation, and it is safe to say that there has as yet been no instance where any invention, system or method, rejected by the Patent and Engineering Departments of the American Telephone and Telegraph Company has ever had any permanent success when used elsewhere.
>
> The Manufacturing Department creates and builds the equipment and apparatus which have been adopted. In this way throughout the whole grand system will be found standardization and uniformity. This is not any handicap on improvement or development of the art, for, on the contrary, every suggestion or idea, and there are many, has abundant opportunity to be tested, which would not be possible otherwise. No one of the companies could by itself maintain such an organization, and it would be fatal to any service to introduce or try out undeveloped ideas in actual service.

In 1909, during Theodore N. Vail's second term as President, AT&TCo acquired a substantial interest in the Western Union Telegraph Company by purchase of stock, and Vail was elected President of Western Union. The purpose of this acquisition was to combine the advantages of the telephone and the telegraph so that each would find its level of use, on the basis of being complementary rather than rival services. This viewpoint had been taken by Vail as early as 1902, although he was not at that time connected with either company.

Much progress was made in implementing the concept of complementary services, but the United States Department of Justice soon decided that the association of these two companies violated the antitrust laws and required their separation. Accordingly, in 1913, AT&TCo disposed of its Western Union stock and Vail resigned the

Western Union presidency, remaining President of AT&TCo until 1919. Meanwhile, friendly business relations were maintained between the two companies, to the great advantage of the public at large.

II. RESEARCH AND DEVELOPMENT ORGANIZATIONS

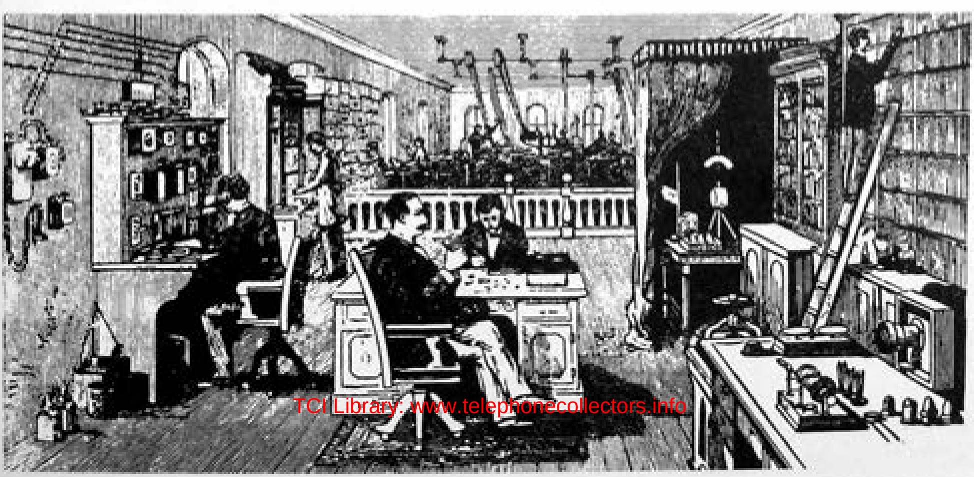
While the various organizational changes just related were largely made in response to a continuing need for the great amount of capital necessary to finance the "Grand System," the parallel necessity for promoting the required technical developments was at no time overlooked. From the beginning it was Bell System policy not to leave technological growth to chance but to foster it within the organization.

2.1 Mechanical Department of American Bell

Until 1883, the technical work of the American Bell Telephone Company had been carried on by two groups: the Electrical and Patent Department, which studied available patents and new apparatus and performed all of the engineering functions, and the Stock Testing Department. In 1883, an Experimental Shop was organized to supplement the activities of the Electrical and Patent Department. In June 1884, the Shop's name was changed to the Mechanical Department, although a similar activity today might be called development and research. The Mechanical Department was located at 101 Milk Street, Boston, and its appearance in 1884 is indicated in Fig. 2-2. A little about the activities of this new department appeared in the American Bell Telephone Company Annual Report for 1884:

> During the past year the Electrical Department has been reorganized and its scope somewhat extended . . .
>
> In the Mechanical Division, established this year, good and efficient service has been done.

Fig. 2-2. The laboratory of the Mechanical Department at 101 Milk Street, Boston, in 1884, from an old wood engraving published in *Scientific American*.



> The principal work which has shown results has been: A series of experiments on long-line apparatus; the protection of our system and apparatus from strong currents, either lightning or electric-light currents; perfecting the central office apparatus to be used in small exchanges; simplifying and perfecting the apparatus in general use; testing wire line and establishing a standard for the same.

Dr. Hammond V. Hayes, trained in physics and engineering at Harvard and the Massachusetts Institute of Technology, took charge of the Mechanical Department in November 1885. This was, in effect, the first formal organization in the continuous chain of research and development organizations leading to the present Bell Telephone Laboratories. An organization chart of the American Bell Headquarters technical staff as of December 31, 1885, is given in Fig. 2-3, and the Mechanical Department laboratory in 1886 is shown in Fig. 2-4. Some of the personnel of the Department as of about 1889 are shown in Fig. 2-5. A view of the laboratory where one of these engineers, E. H. Lyon, carried out a great deal of his work is shown in Fig. 2-6 as it appeared about 1890.

Fig. 2-3. Organization chart of the American Bell Headquarters technical staff as of December 31, 1885.

HEADQUARTERS TECHNICAL STAFF
AMERICAN BELL TELEPHONE COMPANY
DECEMBER 31, 1885
29 EMPLOYEES

ELECTRICAL AND PATENT DEPT.	MECHANICAL AND TESTING DEPT. (H.V. HAYES IN CHARGE)	STOCK AND PACKING DEPT.
SPECIAL DEVELOPMENT WORK: E. BERLINER W. W. JACQUES J. H. FLANNIGAN G. H. E. TROUVELOT PATENTS: T. D. LOCKWOOD V. M. BERTHOLD G. W. PIERCE D. E. RICHARDS F. J. SCHWARTZ W. B. VANSIZE SPECIAL WORK W. A. HOVEY JOSEPH LYONS	GENERAL EXPERIMENTAL WORK: MORGAN BROOKS E. H. LYON INSTRUMENT INSPECTION: W. L. RICHARDS THOS. F. MAGUIRE G. K. THOMPSON J. B. WILKINS SHOP: H. R. MASON W. P. RICHARDS G. W. THOMPSON CLERICAL: F. E. DONOHOE OFFICE BOY: F. C. BROWN	J. R. QUICK A. F. HALL T. FRANCIS MAGUIRE F. MARTIN H. G. MCKENNEY

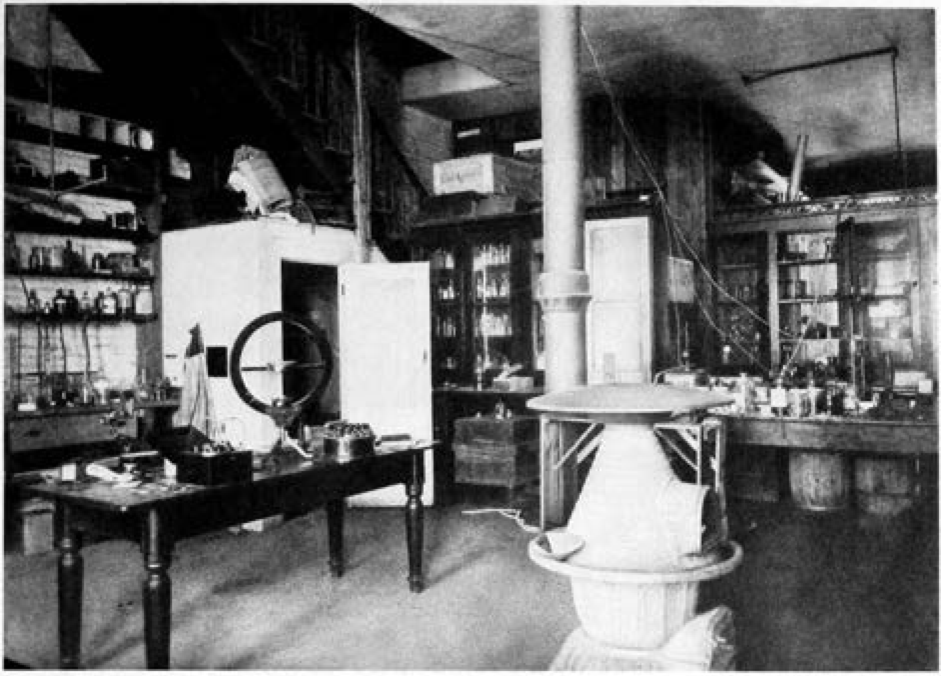


Fig. 2-4. The laboratory of the Mechanical Department at 141 Pearl Street, Boston, in 1886.

Fig. 2-5. Some of the personnel of the Mechanical Department about 1889. Seated, left to right, W. L. Richards, Harry Sears, Chauncey Smith, and W. J. Hopkins. Standing, left to right, E. H. Lyon and Anthony C. White.

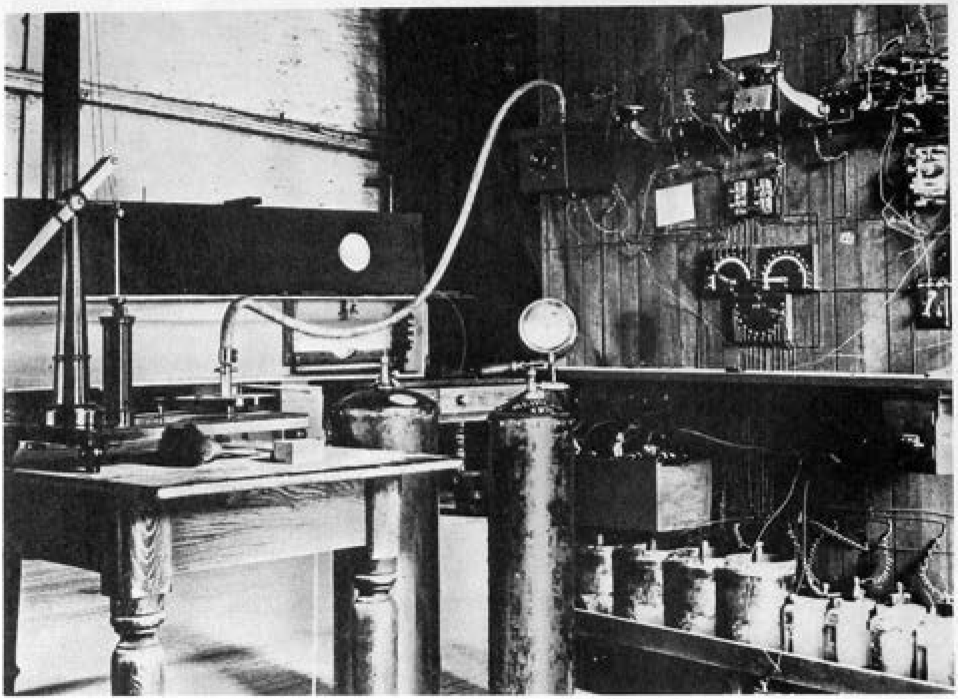


Fig. 2-6. The laboratory of E. H. Lyon at 127 Purchase Street, Boston, as it appeared about 1890.

2.2 Engineers Department of American Bell

In 1891, an Engineers Department was formed under American Bell's Chief Engineer Joseph P. Davis to standardize plant construction and operating methods. Although Davis was nominally in charge of both the Mechanical and Engineers Departments, Hammond Hayes continued to guide the Mechanical Department independently. The American Bell Telephone Company Annual Report for 1893 discussed the organization and functioning of the Department under Hayes:

> During the year a reorganization of the engineers' department has been determined upon and partly carried out. It has been decided to unite with it the mechanical department, thereby much enlarging the scope of the work of both departments. Mr. Hayes has been appointed electrical engineer, under Mr. Davis, our chief engineer. The work of the department has been very considerable in examining and advising upon projects for altering, extending, or otherwise improving the buildings, the equipment, and the plants of the company's licensees, and in giving personal assistance when desired while such work was in progress. Plans,

> specifications, and estimates have also been made in new underground systems where none before existed, and in extending the underground systems already existing in other cities.
>
> The advantage of uniformity in the form of apparatus used, in the mode of working, the details of construction, and the various methods of operation that are thereby introduced, are highly and extremely beneficial.

Figure 2-7 is an organization chart for the American Bell Headquarters technical staff, in Boston, as of December 31, 1895, showing that the staff had grown to 81 employees.

HEADQUARTERS TECHNICAL STAFF AMERICAN BELL TELEPHONE COMPANY DECEMBER 31, 1895 81 EMPLOYEES		
ENGINEERS DEPT. JOSEPH P. DAVIS, CHIEF ENGINEER	MECHANICAL DEPT. H. V. HAYES. ELECTRICAL ENGINEER	EXPERIMENTAL DEPT.
CONDUIT PLANS AND DEVELOPMENT STUDIES: W. S. FORD J. A. HIGHLANDS B. W. TRAFFORD J. WYMAN TOLL TRAFFIC STUDIES: T. B. DOOLITTLE G. T. BLOOD T. COTTER A. J. DELANO J. A. MCCABE R. A. NICHOLS BUILDINGS: L. F. RICE INSURANCE: C. J. H. WOODBURY INSPECTION: C. H. CUTLER DRAFTING: A. V. EDWARDS CLERICAL: S. H. MILDRAM M. J. HEANEY O. W. HOLLIS	TELEPHONE TRANSMISSION: W. L. RICHARDS J. S. STONE SWITCHBOARD ENGINEERING: L. S. GREENLEAF E. SLADE T. C. WALES, JR. W. R. WESCOTT A. S. WILLIAMS MECHANICAL CONSTRUCTION AND DESIGN: C. H. ARNOLD J. S. CODMAN A. DEKHOTINSKY F. C. MOODY E. C. ROBES G. K. THOMPSON GENERAL ENGINEERING: F. L. RHODES CHEMICAL LABORATORY: J. C. LEE G.D. BASSETT WIRE INSPECTION: 3 INSPECTORS SHOP: 7 MACHINISTS DRAFTING: 7 DRAFTSMEN CLERICAL AND STUDENTS: 8 EMPLOYEES INSTRUMENT TESTING AND PACKING: 18 EMPLOYEES	W. W. JACQUES J. H. FLANNIGAN H. R. MASON

Fig. 2-7. Organization chart of the American Bell Headquarters technical staff as of December 31, 1895.

2.3 Engineering Department of AT&TCo

At the time that the American Telephone and Telegraph Company became the parent company of the Bell System in December 1899, no changes were made in research and development organizations, the parent company's Headquarters technical staff remaining in Boston and the Long Lines Engineering Department in New York City.

In 1902, because of the poor health of Davis, the Mechanical and Engineers Departments were merged into a single organization called the Engineering Department and placed under a three-man committee of which Hayes was a member. On January 1, 1905, Davis resigned from AT&TCo and Hayes was named Chief Engineer and placed in charge of the combined Engineering Department. An organization chart of this department, now consisting of 195 employees in Boston, is shown in Fig. 2-8.

2.3.1 Traffic Division

It was in about 1899 when Hayes became concerned about the divergence of opinions among the Operating Telephone Companies as to permissible circuit and operator loads and other traffic matters, and the need for the Headquarters technical staff to keep itself so informed as to be able to speak authoritatively on all points relating to operating practices.

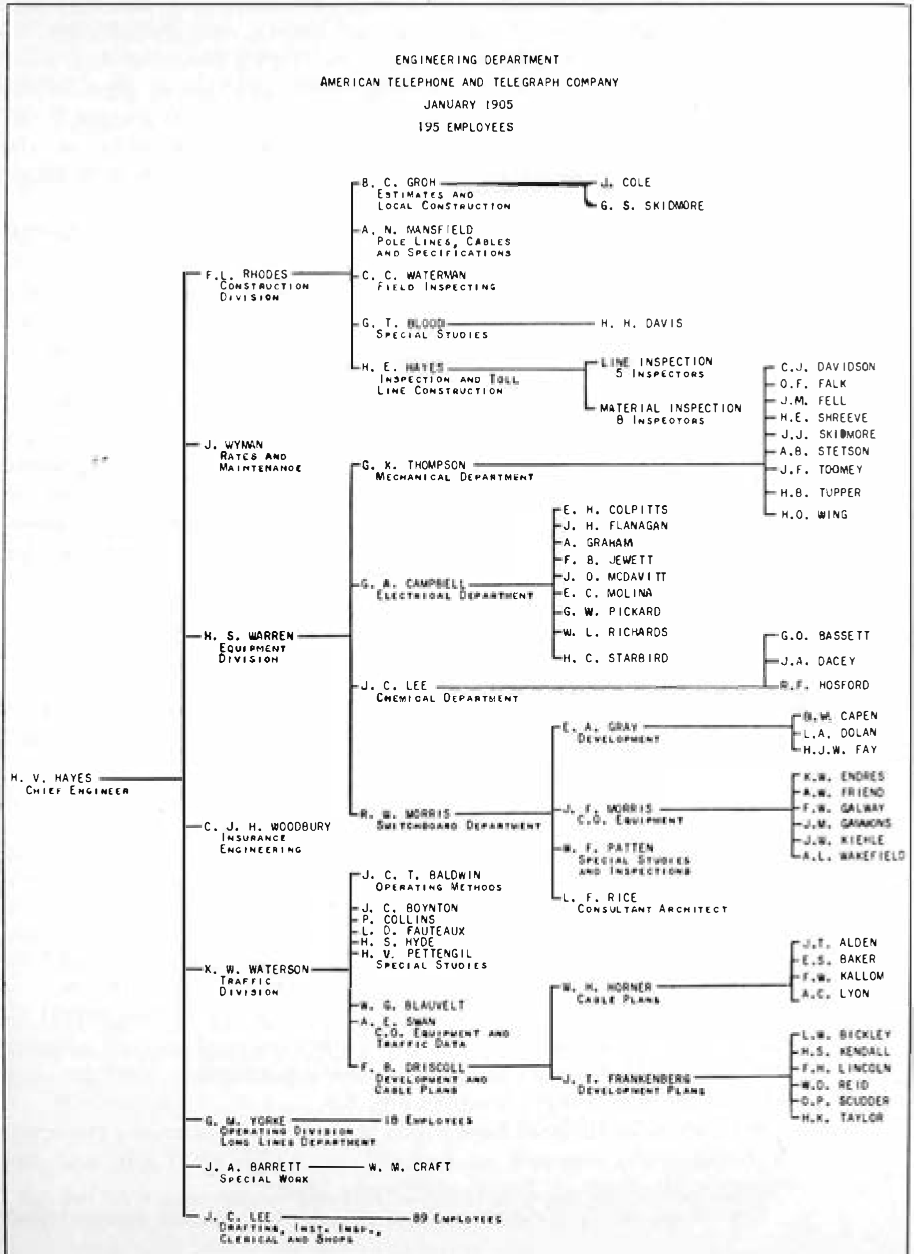
Again in 1905, emphasizing that there was no field of development work requiring more earnest attention than the improvement in toll operating methods, Hayes set up a separate Traffic Division in the AT&TCo Engineering Department, placing K. W. Waterson in charge. With Hayes' support, Waterson pioneered in the development of the traffic unit system that provided a useful tool for calculating force requirements and the loads which could be carried while giving satisfactory service.

2.4 Western Electric Engineering Department

During Western Electric's early history, a group of development and design engineers was established to: (*i*) study the functional or service requirements laid down by the parent company's engineers; (*ii*) develop instruments, switchboards, station apparatus, and wire and cable to meet these requirements; and (*iii*) prepare specific designs with specifications and drawings for use in manufacturing the products in quantity. This organization became known as the Engineering Department of the Western Electric Company. Most of its personnel were located at Western Electric's New York City headquarters at 463 West Street; another group was located in Chicago.

In 1907, Theodore N. Vail, returning to the presidency of AT&TCo,

Fig. 2-8. Organization chart of the AT&TCo Engineering Department as of January 1905.



took steps to bring together a number of Bell System research and development organizations. The Boston laboratory was discontinued, the key people (with the exception of Hayes, who terminated his connection with the System at that time) being transferred to New York City. Some of these people continued to function as the AT&TCo Engineering Department in New York City. Others, including E. H. Colpitts, were transferred to the Engineering Department of the Western Electric Company. The inspection function was also transferred to Western Electric.

A substantial number of engineers from the Western Electric Company's Clinton Street engineering unit in Chicago were moved to 463 West Street, thus completing the first sizable consolidation of the technical laboratory personnel in the Bell System. Western Electric's Chief Engineer was Charles E. Scribner, who during his career was to have a total of 441 patents issued to him.

John J. Carty, who supervised this consolidation, had been one of the original boy operators in Boston, and after holding increasingly responsible posts in the operating companies had been appointed Chief Engineer of AT&TCo by Vail. Carty was in a sense the creator of the profession of telephone engineer. He laid down the responsibilities of the profession on the widest and most comprehensive lines, and fought down flimsy, clumsy methods wherever he found them in the still-adolescent industry.

2.4.1 Research Branch

John J. Carty, with an uncanny ability to sense the lines of possible progress, strongly supported the buildup in research effort in the Bell System. He reported in 1911:

> To make adequate progress in this work [fundamental research], it was decided to organize a branch of the engineering department which should include in its personnel the best talent available and in its equipment the best facilities possible for the highest grade research laboratory work . . . A number of highly trained and experienced physicists have been employed . . . Another man in charge of an important investigation has had a considerable amount of post-graduate work and was for a time a professor of electrical engineering at a well-known institution.

The Research Branch referred to by Carty was organized in 1911 as part of the Engineering Department of the Western Electric Company and with E. H. Colpitts as its head. An organization chart showing this new branch is reproduced in Fig. 2-9.

In Chapter 10 we shall have much more to say concerning the spirit of fundamental research as fostered by AT&TCo's Carty and his colleague Dr. Frank B. Jewett of Western Electric.

Something of the makeup and activities of Bell System research and

TELEPHONE ENGINEERING DEPARTMENT N°600

LIST OF REFERENCE NUMBERS

Reference No.		Under Direct Supervision of	Reporting to
Administration			
601	Chief Engineer	C.E.Scribner	Vice-Pres.
602	Assistant Chief Engineer	J.I.McQuarrie	601
605	Assistant Chief Engineer	F.B.Jewett	601
Information Branch			
603		R. Raymond	602
Clerical Branch			
611	Chief Clerk	R.E.Williams	602
612	Records Division	R. Green	611
613	Expense and Order Division	H.C.Wetzelberger	611
614	Stenographic Division	E. Clinton	611
615	Correspondence Clerk	M.B. Walsh	611
Automatic Development Branch			
671	Supervision Section	C.F.Baldwin	602
672	Apparatus Design Section	J.N.Reynolds	671
673	Equipment Section	W.M.Smith	671
674	Circuit Section	F.N.Reeves	671
675	Laboratory Division	F.N.Reeves	671
676	Special Investigation Division	S.B.Williams	671
Special Studies			
640	Special Automatic Studies	A.H. Dyson	602
661	Special Telegraph Studies	J. H. Bell	602
		P.M. Rainey	602
Telegraph Development Branch			
660		A.F. Dixon	602
Inspection Branch			
648	Supervision	A.H.Verum	602
655	Apparatus Division	F.D.Thompson	648
656	Methods Division	F.D.Thompson	648
658	Central Office Equipment Division	W.C. Adams	648
Line Material Inspection Branch			
650	Supervision	C.R. Myer	602
651	General Supervising Inspector	C.A. Davis	650
652	Mfg. Cost Study Division	F.P. Moore	650
653	Special Studies Division	E.R. Scudder	650
Research Branch			
604	Supervision Division	E.H.Colpitts	605
606	Special Research Division	E.H.Colpitts	604
607	Repeater Division	H.E.Shreeve	604
Development Branch			
620	Supervision Division	E.B.Craft	605
625	Physical Laboratory Division	O.E.Stevens	620
633	Power Room Section	E.Kelber	625
Apparatus Design Division			
630	Supervision Section	J.J.Lyng	620
634	Design Section	J.J.Lyng	630
637	Drafting Section	R.C.Winckel	630
647	Model Shop Section	J.W.Upton	630
Circuit Laboratory Division			
623		H.L.Darrah	620
Transmission Branch			
621	Supervision Division	J.C.R.Palmer	603
677	Design Division	H.B.Wier	621
678	Laboratory Division	H.C.Benson	621
679	Investigation Division	H.W.Purcell	621
Chemical Branch			
624		J.W.Harris	605

Fig. 2-9. A Western Electric organization chart of June 1912 showing the Research Branch, under E. H. Colpitts, formed in 1911.

development organizations is contained in the AT&TCo Annual Report for 1913:

> At the beginning of the telephone industry there was no art of electrical engineering nor was there any school or university conferring the degree of electrical engineer. Notwithstanding this, the general engineering staff was soon organized, calling to their aid some of the most distinguished professors of science in our universities.
>
> As problems became more formidable and increased in number and complexity, the engineering and scientific staff was increased in size and in its specialization, so that we now have working at headquarters on the problems of the associated companies 550 engineers and scientists carefully selected with due regard to the practical as well as the scientific nature of the problems encountered.
>
> Among them are former professors and instructors of our universities, post graduate students and other graduates holding various engineering and scientific degrees from 70 different scientific schools and universities, 60 American and 10 foreign institutions of learning being represented.
>
> No other telephone company, no government telephone administration in the world, has a staff and scientific equipment such as this.
>
> The Bell Company, recognizing at the outset that the problems of telephony would require for their solution the highest degree of scientific and engineering skill, has been foremost in the development of telephone engineering and in the encouragement of scientific research.
>
> It can be said that this company has created the entire art of telephony and that almost without exception none of the important contributions to the art has been made by any government telephone administration or by any other telephone company either in this country or abroad.

The organization of the AT&TCo Engineering Department as of March 20, 1915, is shown in Fig. 2-10, and the Western Electric Engineering Department as of April 1, 1915, in Fig. 2-11. While these are in some detail, they merit this permanent recording because of the large number of names, even in lower echelons, which will appear in later chapters as notable contributors to technology.

2.5 Bell System Patents

From the time that the telephone was invented, patent protection has always been vital. The importance of patents to the Bell System and the necessity for an active patent staff are emphasized in these paragraphs from the AT&TCo Annual Report for 1918:

> At the end of 1918, the Bell System either owned, controlled or was licensed under 3,211 United States patents and 1,213 applications for such patents, making a total of 4,424. These patents and applications cover the entire known telephone field and may be roughly grouped into (1) a large number covering innumerable improvements in details, which collectively render it possible to vastly improve the character of

> the telephone service given, and to give this service at a minimum cost; and (2) those more or less fundamental to the rendition of a particular kind of service.
>
> It is not an exaggeration to assert that there is not a practical detail of telephone development work which is not represented by some patent under which the Bell System is free to operate. By reason of this favorable patent situation, the Bell System has the right to use all of the best and most efficient apparatus that has already been developed, and looking to the future there is no field of development from which it is excluded.

And again in the 1920 Annual Report:

> During the past year, the patent holdings of the Bell System have been increased by rights under more than a thousand patents and applications for patents, an increase of approximately 20 percent. It now owns or controls, or is licensed under, more than six thousand letters patent of the United States and applications therefor. The larger proportion of this increase is represented by inventions made by our own engineers. Rights acquired during the year under other inventions have been acquired largely in exchange for rights granted under our own inventions.

2.6 D&R and O&E Departments of AT&TCo

In 1919, the AT&TCo Engineering Department was reorganized into two groups. One was called the Department of Development and Research, under Vice-President J. J. Carty, and the other the Department of Operation and Engineering under N. C. Kingsbury. The Development and Research Department, commonly referred to as "D&R," was later (1934) consolidated with Bell Telephone Laboratories, while the Operation and Engineering Department (O&E) remained a part of AT&TCo.

A brief report of the activities of the D&R Department and the Western Electric laboratories appears in the AT&TCo Annual Report for 1920:

> The year just closed has been one of remarkable activity in the Department of Development and Research. In this department, including the laboratories at the Western Electric Company, 2,800 employees are engaged exclusively in research and the development and improvement of telephone and telegraph apparatus and materials and methods. Of these, 1,100 are engineers, chemists, physicists, and other scientists, among whom are graduates of more than 100 American colleges and universities. The remainder are laboratory assistants, draftsmen, stenographers, clerks, model makers, and administrative personnel.
>
> At the close of the year, upwards of 2,500 research and development projects were in hand, all these calculated to improve the service which the associated companies are rendering to the public or to make it more economical . . .

AMERICAN TELEPHONE & TELEGRAPH CO.

ORGANIZATION CHART

ENGINEERING DEPARTMENT

252 Employees – March 20, 1915

J.J. Carty
Chief Engineer

- **B. Gherardi** — Engineer of Plant
  - **Outside Plant Engineer** — F.L. Rhodes
    - Plant Methods and Costs: G.T. Blood; W.H. Horner, W.A. Taylor, R.A. Haislip, R.B. Hill, K.L. Wilkinson, C.C. Ayres, S.C. Carr
    - Materials: R.F. Hosford; L.B. Fish, E.B. Griffen, H.F. Stover, W.M. Rile, C.C. Fritz, H.D. Cutler, J.W. Hines
    - Exchange Aerial Plant, Toll Lines: C.C. Waterman; E.M. Matthews, R.C. Darrow
  - **Equipment Engineer** — L.F. Morehouse
    - Central Office Development: R.W. Morris; J.F. Toomey, H.M. Bascom, H.H. Davis, R.G. Ramsdell, L.H. Darrow, C.A. Wright, C.S. Demarest, J. Davidson, Jr., G.M. McCarty, G.W. Kuhn, W.A. Rhodes, O.M. Arnold, W.G. Freeman, E.F. Watson, 2 Clerks
    - Central Office Engineering: C.A. Berry; J.F. Morris, P. Woollcott, N.F. Roberts, H.N. Tyler, W.A. Hedrick, R.L. Young, O.A. Freund, E.C. Balch, E.N. Renshow, W.L. Steckel, L.W. Cooper, L.R. Mapes, F.K. Rowe
    - Special Switchboard Systems: W.G. Blauvelt; A.H. Moroweck, E.C. Molina
    - Substation Equipment: G.K. Thompson; C.J. Davidson, E.W. Niles, L.D. Barrows, A.O. Hoefer
    - Telegraph Equipment: J.M. Fell, R.D. Parker; R.E. Pierce, A.S. Littel, G.S. Vernam, 2 Boys
- **K.W. Waterson** — Engineer of Traffic
  - **Local Traffic Engineer** — W.E. Farnham
    - Traffic Engineering: H.B. Stimson; F.W. Galway, H.L. Brown, H.F. Parker
    - Operating: W.D. Sargent, E.B. Fairman
  - **Toll Traffic Engineer** — S.H. Browne
    - Traffic Engineering: J.W. Kiehle; F.P. Davis
    - Operating Practices: F.M. Bronson; A.A. Denico
    - Operating Methods: H.P. Charlesworth; W.C. Oakes
    - Toll Line Efficiency: G.W. McRae
    - Rate Quoting and Route Systems: A.C. Blood; P.R. VanMater, 5 Clerks
    - Toll Fundamental Plans: E.S. Baker
- **J.A. Barrett, H.S. Warren** — Engineering Work for Legal Cases
  - E. Miller, W.M. Craft
    - P.W. Spence, 1 Clerk

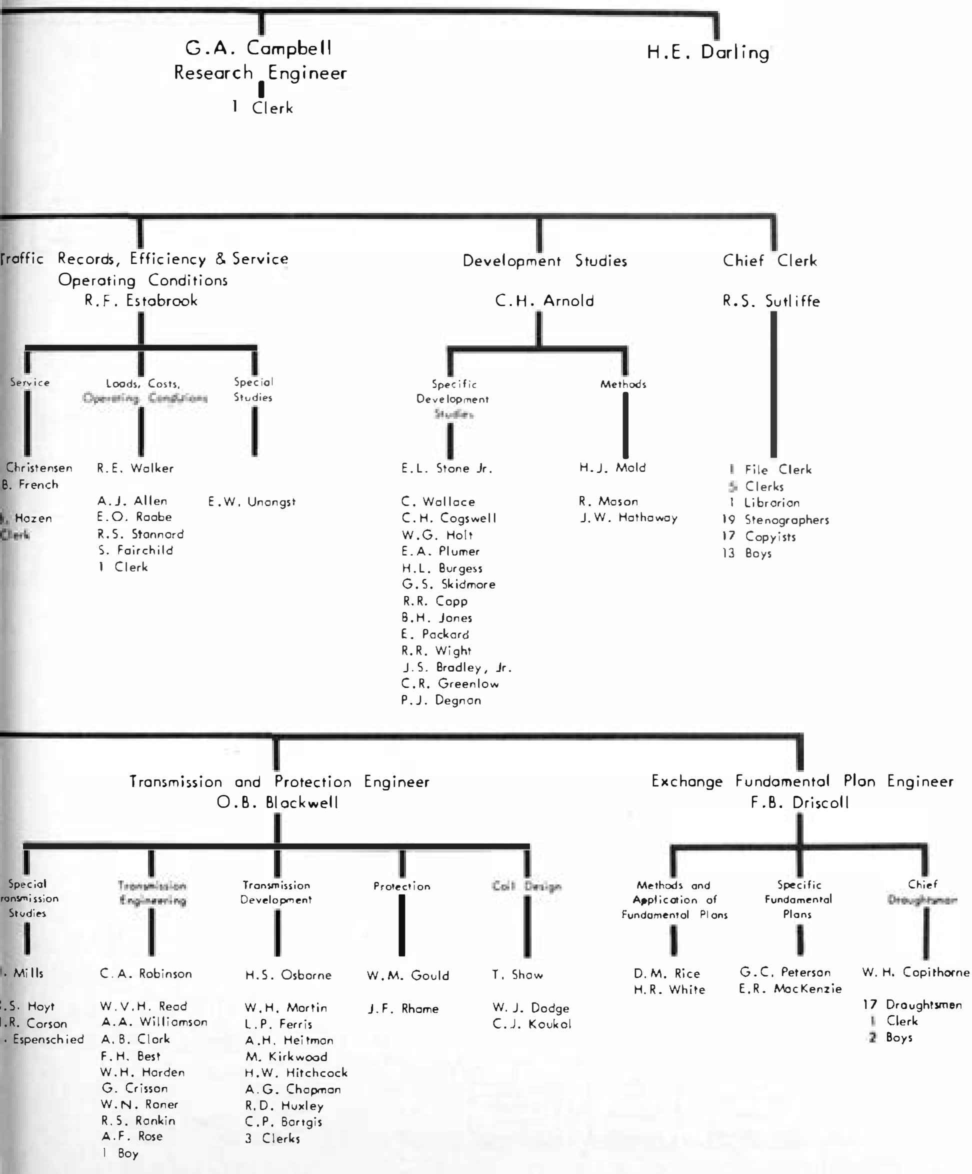
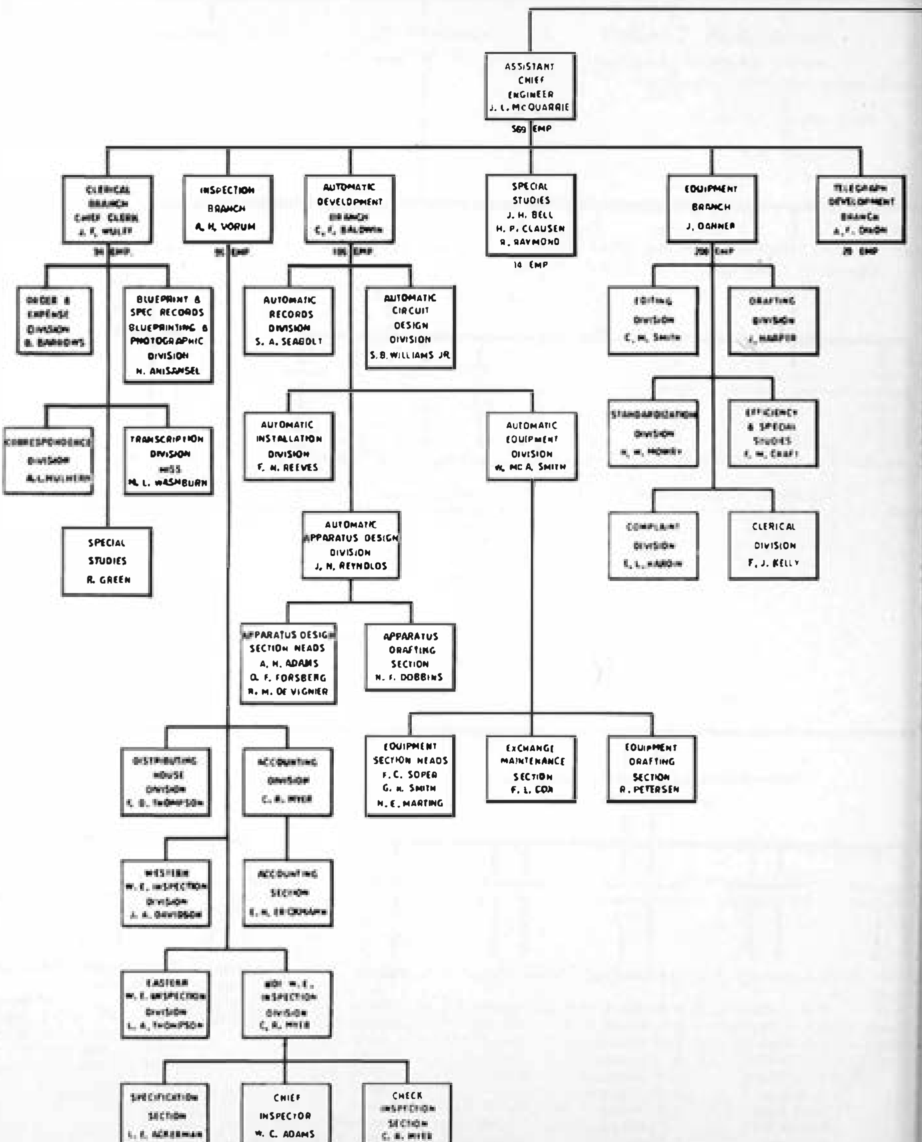


Fig. 2-10. AT&TCo Engineering Department as of March 20, 1915.

CHART F

ENGINEERING DEPARTMENT



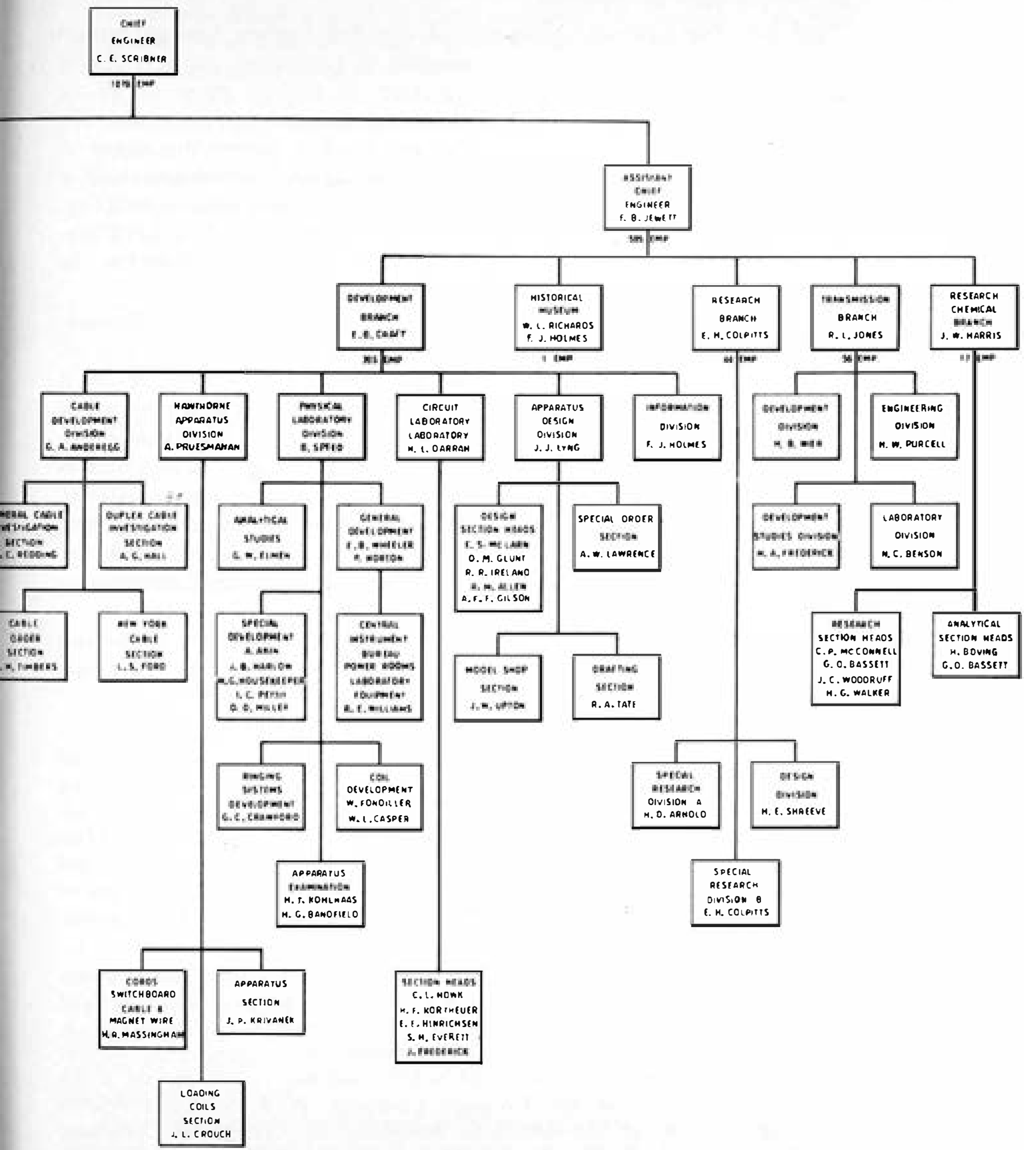


Fig. 2-11. Western Electric Engineering Department as of April 1, 1915.

2.7 Bell Telephone Laboratories

By 1924 the technical programs of the Bell System had so grown in range and intensity, and in number of personnel, as to suggest formation of a single new organization to handle most or all of these activities. Such an organization was formed on December 27, 1924, and started operations on January 1, 1925, under the name of Bell Telephone Laboratories, Incorporated. This corporation had a dual responsibility—to the American Telephone and Telegraph Company for fundamental researches and to the Western Electric Company for the embodiment of the results of these researches in designs suitable for manufacture.

In early discussions, it was felt that the AT&TCo D&R Department should be included in this new organization. However, not all of the research problems assigned to this department were of a laboratory character. The D&R Department had responsibilities in the development and establishment of proper standards for transmission and for other fundamental operations of communication. It also had to envisage the communication art as a whole and to foresee needs and trends, and to focus attention on desirable lines of development. Accordingly, this department was retained as a part of AT&TCo for the time being. It was not consolidated with Bell Telephone Laboratories until 1934.

One of the better accounts, for our purpose, of the formation of the new company appeared as a news item in the February 1925 issue of *The Telephone Engineer*:

> New York City—Extensions of laboratory facilities for the scientists and engineers of the new Bell Telephone Laboratories, Inc., are already under way. Laboratory space in the form of a new building covering almost a quarter of a city block will be added to the 400,000 square feet at present in service in the group of buildings at 463 West Street, New York City. At the date of incorporation, the personnel numbered approximately 3,600, of whom about 2,000 are members of the technical staff, made up of engineers, physicists, chemists, metallurgists and experts in various fields of technical endeavor . . .
>
> The chairman of the board of directors of the Bell Telephone Laboratories, Inc., is General J. J. Carty, Vice-President of the American Telephone and Telegraph Company. Other members are: Dr. F. B. Jewett, formerly Vice-President of the Western Electric Company, President of the new corporation, and also recently elected Vice-President of the American Telephone and Telegraph Company; W. S. Gifford, executive Vice-President of the American Telephone and Telegraph Company; Bancroft Gherardi, Vice-President of the same company; C. G. DuBois, President, and J. B. Odell, assistant to the President of the Western Electric Company.

> The operations of the Bell Telephone Laboratories, Incorporated, are under the direction of E. B. Craft, executive Vice-President, who was formerly chief engineer of the Western Electric Company.
>
> In the functional division of the research, development and engineering work of the laboratories, physical and chemical research is organized under Dr. H. D. Arnold, director of research; development of apparatus under J. J. Lyng, apparatus development engineer; and development of communication systems under A. F. Dixon, systems development engineer, all formerly concerned with similar activities in the engineering department of the Western Electric Company. Dr. R. L. Jones, inspection manager, continues his former responsibilities in engineering inspection, and S. P. Grace, commercial development engineer, those of commercial development.
>
> The patent work of the Laboratories is organized under J. G. Roberts, General Patent Attorney, formerly Assistant General Patent Attorney of the Western Electric Company.
>
> The corporate and commercial relations of the laboratories are under the direction of Vice-President E. P. Clifford, who was formerly Commercial Manager of the Engineering Department of the Western Electric Company. John Mills continues as personnel director, responsible for personnel activities, and educational and college relations.
>
> The formation of Bell Telephone Laboratories, Incorporated, provides an individual organization, the whole activities of which may be more efficiently devoted to the furtherance of research, development and engineering investigations along the line in which the parent companies have already made such remarkable progress. Its formation is an indication of the estimate which these companies place upon the importance of properly organized research and is a promise of continuous service to the public, to the communication art and to the progress of science.

Charts showing the organizational framework before and after the formation of Bell Telephone Laboratories appear in Figs. 2-12 and 2-13. The similarity of these two charts, and the carry-over of names, point up the smooth transition from the old organization to the new.

The building at 463 West Street, New York City, headquarters for this new organization, is shown in Fig. 2-14 as it appeared in 1923 when it was still the headquarters of the Western Electric Engineering Department. The same building was to remain one of the major locations of Bell Laboratories for more than 40 years. Today, although no longer associated with the Bell System, it still serves as a center for creativity, having been converted into studio-residences for artists.

Likewise, after an additional half-century, the place of Bell Laboratories in the organizational structure of one of the world's great corporate enterprises remains essentially as shown in the diagram of Fig. 2-15, dated 1925. The role of science and technology in the growth

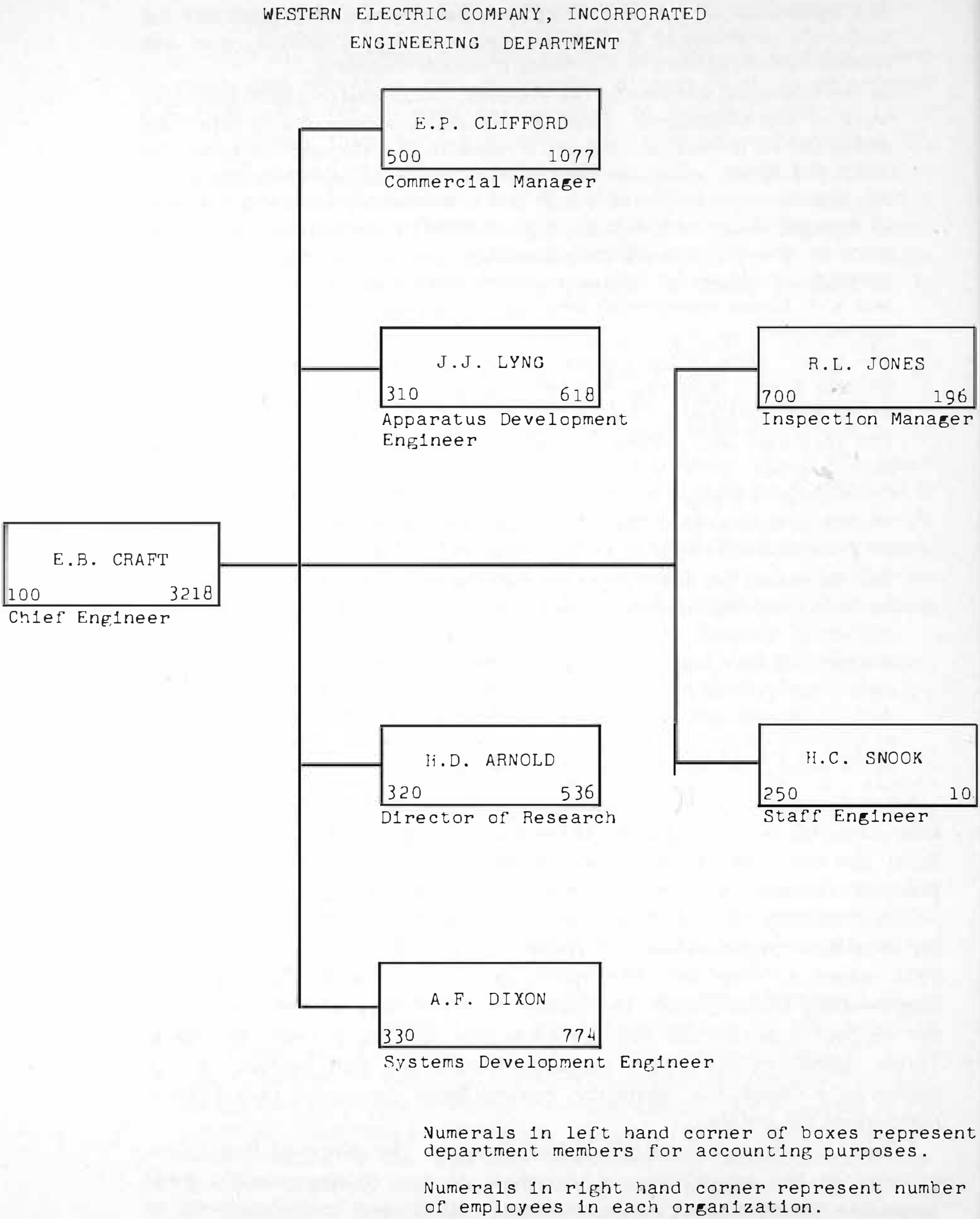


Fig. 2-12. Western Electric Engineering Department as of March 25, 1924.

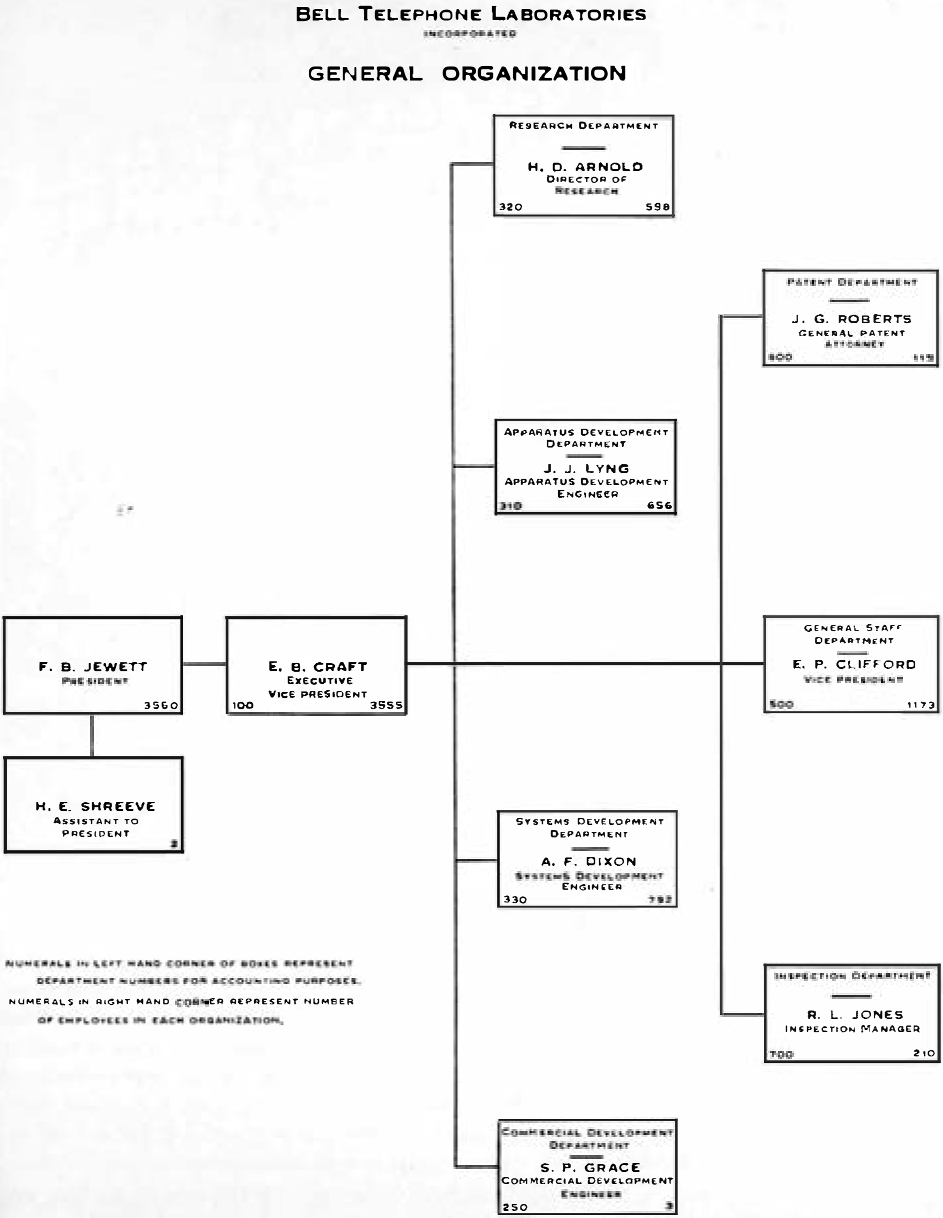


Fig. 2-13. General organization of Bell Telephone Laboratories, June 1, 1925.

Fig. 2-14. Western Electric's building at 463 West Street, New York City, shown in 1923. In 1925 this building became the headquarters of Bell Telephone Laboratories.

of the enterprise could have been seen only dimly by the small group of Boston merchants who laid the foundations of the business. Yet, with vision and almost prophetic foresight, they and their successors gave unwavering support to the pursuit of new scientific knowledge and its vigorous application, the subject of our remaining chapters and of subsequent volumes.

III. SUMMARY

The first 50 years following the invention of the telephone saw the solution of critical problems of organization that vitally affected the growth of the "Grand System" outlined by Alexander Graham Bell.

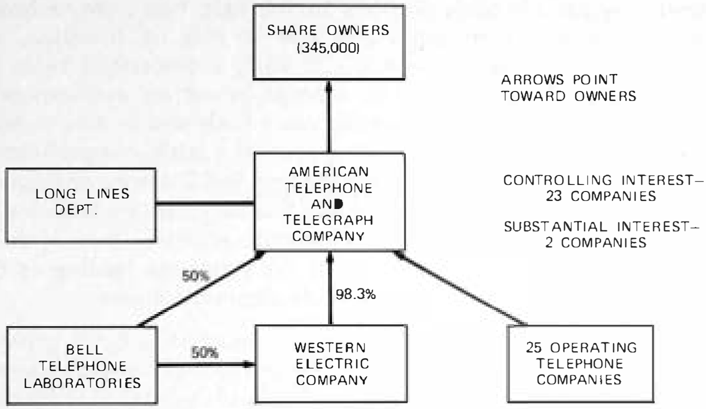


Fig. 2-15. Corporate structure of the Bell System in 1925, showing the new company, Bell Telephone Laboratories, owned jointly by AT&TCo and Western Electric.

Financing the rapid growth of this new venture started with a simple "gentlemen's agreement" among a few friends of Bell. When these resources proved inadequate, the transition to the corporate form was made smoothly under wise management without the stock-jobbing and bankruptcies that were so common during the early industrial era. By about 1900, the pattern had been set for an integrated system consisting of regional operating companies owned by a parent company furnishing the interregional connections and responsible for overall engineering, development, research, and manufacture. This relatively simple organizational pattern has continued to serve the needs of the Bell System with little change up to the present.

During all stages of the organizational evolution, development and research were not left to chance but as a matter of basic policy were promoted within the System. At first this work was carried out by engineering organizations within the parent and manufacturing companies but with the formation of Bell Telephone Laboratories it was consolidated in a separate corporate body owned jointly by the parent company and the manufacturing company.

The policy of promoting internal technological development was started when organized research and development was essentially unknown but has proved so successful that it is now commonplace among progressive industrial companies. However, few modern research and development organizations enjoy the autonomy provided by a separate corporate entity such as Bell Laboratories.

In the chapters which follow we shall examine in detail the technical advances which occurred during the first 50 years of telephony.

We shall see that the early workers in the field had little technical background on which to build and had to rely on intuition, ingenuity, and experiment. However, by 1900, a theoretical basis for electrical communication began to emerge based on application of the fundamental research then taking place both within and outside of the Bell System. By 1925 this had grown to a fairly comprehensive background of basic knowledge, and the new Bell Laboratories proved to be an ideal vehicle for expanding and building on this knowledge. Bell Laboratories' research and development activities then were to provide the starting point for many of the programs leading to the subsequent great expansion of worldwide communications.

Chapter 3

Station Apparatus

During the first half-century of telephony, the development of station apparatus, as related in this chapter, involved a great expansion in technology. The basic principles of the electromagnetic receiver and the variable-resistance transmitter were covered by the initial Bell patent and within a few years had become accepted throughout the industry. By the middle 1890s these principles had been implemented in a form which outwardly was not greatly changed during the first quarter of the twentieth century. A considerable number of internal changes were made during this period, however, largely on an empirical basis, to adapt station apparatus to the needs of the expanding and changing transmission and switching plant. The same period also was marked by the gradual evolution of the theory and the measuring techniques which were necessary to remove design from the realm of empiricism so that telephone stations could be developed systematically to meet preselected requirements. The fundamental design techniques worked out during the first quarter of the twentieth century were being applied before the end of the 1920s and provided much of the groundwork on which the later enormous advances in efficiency, quality, and utility could be based.

I. INTRODUCTION

Alexander Graham Bell should be credited with two basic communication inventions. The first was a device, the telephone instrument. The second was what today we would call a "system concept."

The essence of Bell's invention of the telephone instrument resided in his perception of an "analog"[1] relationship between sound pressure and electromotive force. What he demonstrated was a physical means for converting sound waves into what his associate Watson termed "sound-shaped" waves of electric current, followed by reconversion at the receiving end. Success on the first attempt, with the primitive

[1] In the language of today's communication art, the term "analog" signifies representation of one variable by another when the relationship is continuous, rather than coded or discontinuous.

and inefficient instruments of Bell and Watson, can be attributed to the extraordinary sensitivity with which nature has endowed the human ear; for no man-made instrument of that day could have detected and recognized the first faint words which electrified Watson on March 10, 1876.

The concept, or "Grand System," sketched by Bell in the 1877 prospectus quoted in Section X of Chapter 1, proposed a system for two-way voice communication between an individual in his home or place of business and any other chosen individual wherever located. It was essential in this concept that no special skill on the part of the user should be required in order to be connected, as desired, to anyone at any place.

The broad system concept was far in advance of existing techniques and its realization required much original work in many fields. The evolution of the organizational structure required to promote the rapid development of this concept has already been outlined and our remaining chapters can be devoted to the evolution of the necessary technology.

The complete system, as conceived by Bell, involved a number of elements, or subsystems. First came the end instruments for converting between sound and electrical waves. Next, means were needed for conveying the electrical waves over great distances. Providing such means involves the art of *transmission*. Third, the transmission paths had to be interconnected in such a way as to provide communication between the desired end instruments; this is the art of *switching*. And finally, means had to be found for directing the switching operation, alerting the person called, and later terminating the call and clearing the circuits; these are signaling techniques. In 1877, when the concept was outlined, none of these elements was available in usable form and there was little theory to guide their development.

The evolution of these developments will be the subject of the chapters which follow. For convenience they will be categorized by the major subsystem discussed: end instruments, transmission, switching, etc. The reader should, however, understand that development occurred in parallel along all lines and a major part of the technical effort was necessarily devoted to designing the elements so that they would work together as a coordinated system.

This chapter will be devoted largely to the terminal equipment commonly designated as "station apparatus." It also will cover similar apparatus used by telephone operators and for the other special purposes, as well as supplements to the more common customer apparatus such as coin telephones, telephone booths, and the like.

Broadly, station apparatus consists of the terminal equipment which provides the interface between users and the transmission and switching portions of the system. This equipment is somewhat unique in that it is multifunctional. It includes the end instruments which convert speech waves into their electrical analog and reconvert the electrical waves into the acoustical form needed by the user at the distant end. Thus they prepare the sound waves for electrical transmission and are closely allied to the transmission subsystem. The station apparatus also is part of the switching subsystem in that it initiates, under control of the user, the signals required to direct the routing set up by the switching mechanism to the called station. In some cases, rudimentary switching operations are also performed by the station apparatus to establish local, usually on-premise, connections to other stations or to establish connections to one of several outside lines. The station apparatus also includes ringers or other alerting mechanisms, such as lights, to signal that calls are waiting or to indicate the status of calls in progress. Finally, since the station apparatus is on the users' premises, it constitutes the part of the system visible to the users and, to a considerable extent, is representative of the entire telephone system. This is particularly true in these days of automation when personal contacts within the telephone system are so few. Thus, station apparatus must be designed to be compatible with both system and human factors. It must not only function properly with the transmission and switching elements of the system but also serve as a connecting link between the electrical and human elements which enter into a telephone call.

Because of the multifunctional character of telephone stations and their dependence on developments in other areas, it will be helpful to anticipate some of the material in later chapters, particularly the evolution of the problems faced in developing the wire transmission plant. Broadly, this plant is made up of the lines which carry the electrical waves between the stations. Originally, each line consisted of a bare iron wire supported by an overhead structure, with the ground used as the return portion of the circuit. Later, copper replaced iron as the conductor and the "metallic" circuit using a pair of wires replaced the ground-return circuit. This "open wire" structure was supplemented (and untimately replaced) by bundles of insulated copper wires in a waterproof sheath. These "cables" provided a much more compact structure than open wire and could be used either underground or suspended from poles. All of these structures presented transmission problems which for 40 years not only affected the growth of the network but also greatly influenced the development of station apparatus.

The main problem was "attenuation" or the reduction in strength suffered by the electrical waves as they traversed the transmission lines. The electric currents produced by the telephone from the voice waves were feeble at best and large amounts of attenuation could reduce them to a point where their reconversion to voice waves no longer produced audible speech. Consequently, for many years a major objective in designing wire plant was to find ways of reducing or compensating for attenuation. Considerable success was achieved in reducing attenuation, first, by the use of copper in place of iron wire, and later by series inductance, or "loading coils," inserted at regular intervals. But until the 1907–1912 period the only way found to compensate for attenuation was by designing the station instruments for the highest practical efficiency at the frequencies most important for speech transmission.

By applying all the ingenuity of transmission and station designers it became possible to telephone over roughly half the width of the continent by the early part of the twentieth century. For greater distances the need was obvious: a booster "amplifier" which could be inserted in the line to generate, from the attenuated waves, analogous but enhanced waves. We shall shortly see that one form of telephone instrument was an amplifying device which converted a weak voice wave to a higher-powered electrical wave, and it was largely through its use that transmission became possible over hundreds of miles. However, the development of a line amplifier eluded all efforts until the electron tube was invented by de Forest in 1907 and turned into a practical amplifier by Arnold and others in the next five to ten years. This great invention at last gave the transmission engineer the scope he needed and freed the station developer from the rigorous constraints under which he previously operated. Further freedom was given the transmission and station designer by the development, beginning about 1915, of a new transmission medium, a sort of "wired wireless," which came to be known as "carrier" transmission.

With this background, the reasons behind the evolution of station apparatus should be more understandable to the reader. However, one more matter should be mentioned for the benefit of those with little previous knowledge of telephone transmission. As the electric currents flow through successive line sections, they are attenuated exponentially. If the output/input ratio is x for one section of line, it will be x^2 for two, x^3 for three, and so on. In order to simplify the treatment of attenuation (and amplification) it has become the custom to describe it in terms of the logarithm of a power ratio, instead of using a simple ratio, in order to obtain units which are additive. The unit currently used almost universally is the decibel (dB) and the reader will find it used frequently in this and succeeding chapters. Its evolution is described in some detail in Section 5.1.1 of Chapter 4, but for the present it will be sufficient to say that mathematically it is defined as 10 log P_1/P_2, where P_1 and P_2 represent input and output powers respectively. This formula gives an answer having a plus sign for attenuation and a minus sign for amplification. Frequently the terms "loss" and "gain" are used to respectively represent a decrease and an increase of power, in which case the signs are implicit. Finally, those unfamiliar with the decibel may feel the need for an illustration of the magnitude of the unit. While this is covered in Chapter 4 also, the following table may meet the immediate need:

dB	Power Ratio (Approximate)
1	1.25
3	2
6	4
10	10
20	100

We now discuss the evolution of station apparatus during the first 50 years of telephony with the hope that our readers will have some feeling for the interaction between the growing technologies in the station and transmission areas.

II. ELECTROACOUSTIC CONVERTERS

In telephone terminology the device that converts acoustic waves into analogous electric waves is referred to as a "transmitter" and the one that performs the reverse conversion is designated a "receiver." Bell's electromagnetic converter[2] operated reciprocally and, as indicated in Chapter 1, was often used in the early days of telephony to perform both conversions. Sometimes the same device was used alternately as transmitter and receiver but more commonly separate devices, both using the electromagnetic principle, were provided.

While the use of the same basic device for both transmitting and receiving functions had the advantage of simplicity, it was highly restrictive since converters of the reciprocal type are inefficient, particularly when they must meet some of the difficult size and technical requirements imposed by telephony. Using nineteenth-century techniques, an electromagnetic receiver would roughly convert one thousandth of the applied electrical power into acoustic power in the ear canal. The efficiency of a similar device used as a transmitter

[2] The basic principle was that a diaphragm, moved by acoustic waves, caused an iron element to vibrate in a magnetic field. The resultant disturbances in the field induced an electric wave in a coil of wire also placed in the field. Conversely, electric waves in the coil disturbed the field causing the iron element (and diaphragm) to vibrate, thus generating acoustic waves. See Chapter 1 and Section 2.2 of the present chapter for illustrations of early electromagnetic converters.

would not be identical but would be of the same order of magnitude. Thus the dual conversion from acoustic to electric waves and back to acoustic would provide the listener with about one millionth of the voice power at the transmitting end, i.e., it would suffer a "loss" of 60 dB. With this large loss in the converters little margin is available for line attenuation and transmission distance is severely limited.

Fortunately, there was an alternative to the electromagnetic converter for performing the transmitter function. This was the variable-resistance converter in which the acoustic waves impinging on a diaphragm caused a resistance to vary in an analogous manner. By placing this variable-resistance element in a circuit with a direct-current source, the relatively low-powered sounds could cause considerably larger electrical waves. Such a device was a true amplifier using a low-power acoustic input, not as the source of the electric output, but to control a high-power output derived from the direct-current source and providing a gain of as much as 30 dB. The possibility of using such a device was understood from the beginning; it was one of the converting techniques mentioned in Bell's first patent,[3] and the principle was embodied in the electrolytic transmitter which Bell used in his first transmission of intelligible speech. After the first few years of telephony almost all effort on the improvement of transmitters went into the development of practical and efficient transmitters of the variable-resistance type and, as noted in Section 4.2.1 of Chapter 4, these devices were the only practical source of amplification available during the first 35 years of telephony. It should be stressed that without the development of the variable-resistance transmitter, telephony would not have become a commercial reality in the nineteenth century.

Unfortunately, variable-resistance converters were not reciprocal; they could not perform the electric to acoustic conversion and no practical source of amplification for the receiver was available. The receiver, therefore, developed on the basis of the electromagnetic principles used in Bell's original device and even now these basic principles are used in the vastly improved telephone receivers of today and, with slight modification, in the loudspeakers of the public-address and entertainment industry.

2.1 Transmitters

The evolution of the basic principles used in telephone transmitters went through several stages but they were reasonably well established by about 1890. After this time major effort was devoted

[3] A. G. Bell; U. S. Patent No. 174,465; filed February 14, 1876; issued March 7, 1876.

to improving performance and meeting the many evolving requirements of the user and the growing telephone plant. However, before describing these events we should consider a few related matters, some largely of historical interest and others providing substantial technical foundations for events to follow.

2.1.1 Early Background

Bell's liquid transmitter, described in Section IV of Chapter 1, was the first application of the variable-resistance principle. This device, which transmitted the first intelligible speech sounds, used a fine wire, attached to a diaphragm, partially immersed in a metal cup of acidulated water. The resistance varied as the speech sounds caused the diaphragm to move the wire up and down in the acid. This transmitter was satisfactory for demonstrating the principles of telephony but was obviously not practical for commercial usage. For this purpose Bell reverted to the electromagnetic principle and improved the electromagnetic converter to the point where commercial telephony over limited distances could be accomplished. Since this type of transmitter was essentially identical to the contemporary receiver, they will be described together in Section 2.2. However, growth of telephony was dependent on the improvement of the variable-resistance transmitter and the microphonic contact was to prove the key to this development. Before describing this principle it will be interesting to step back a few years before Bell began his work and examine another attempt at achieving telephony.

In 1861, well before Bell's work, Johann Philipp Reis, a professor of Natural Philosophy at Garnier's Institute in Frankfurt-am-Main, Germany, constructed an apparatus which he specifically called a telephone. This apparatus was demonstrated by him before the Physical Society in Frankfurt, and he authorized a Mr. Ladd to exhibit it to the British Association in 1863. The Reis telephone included a transmitter and a receiver (Fig. 3-1). The transmitter comprised a membrane with an attached electrode, and a second, spring-supported electrode, delicately adjusted so that at each vibrating excursion of the membrane an electrical contact between the two electrodes was made and broken. The receiver consisted of a steel rod attached to a sounding board and surrounded by a coil of wire; it operated on the magnetostriction principle that was first observed by Dr. C. G. Page of Salem, Massachusetts, in July 1837. This apparatus, when operating on the make-and-break principle proposed by Reis, could transmit and reproduce the pitch or frequency of sounds but not their variable intensity or amplitude as required by telephony. We can now see that with a different adjustment of the electrodes (in

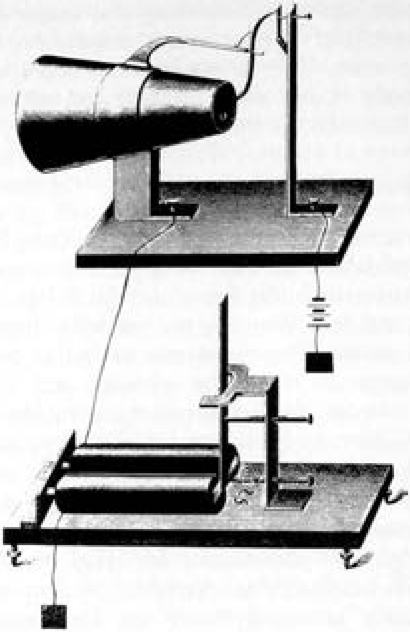


Fig. 3-1. Reis telephone. (Prescott 1879, Fig. 4)

a manner not contemplated by Reis) a microphonic contact might have been obtained that would have produced some amplitude change. Thus it is quite possible that a determined and perceptive experimenter, working with this device, might have discovered the microphonic principle.[4] But the Reis work was not pursued and instead the discovery and application of this important principle grew out of independent work largely carried out in the United States.

Credit for the disclosure of the microphonic principle usually goes to David E. Hughes but, as we shall see, both Edison and Berliner deserve consideration. The basis for the Hughes claim is a paper that he

[4] It will be recalled that Bell's telephone invention was the outgrowth of his work on an entirely different device, the multiplex telegraph. When his instrument produced unexpected results, under maladjustment, he was quick to interpret the reason and had the vision to see the possibility of furthering his concept of telephony. It is interesting to note that A. E. Dolbear, writing in *Scientific American* of June 1881, pointed out that the Reis transmitter did not make and break the circuit at low speech inputs and would, at these levels, act much like a Blake transmitter. However, it is clear from Dolbear's remarks that even as late as 1881, 20 years after the transmitter was proposed by Reis, most technicians still referred to it as a make-and-break device.

had written and sent to Professor Huxley for presentation and publication. Accordingly, this paper, describing experiments with loose contacts between various materials, was read before the Royal Society in London by Professor Huxley on May 8, 1878. Hughes, born in London, was at the time of his experiments Professor of Natural Philosophy at the College of Bairdstown, Kentucky. He had noted that a variation in electrical current followed from varying degrees of intimacy of contact between two conductors. He described how an assembly of three nails, one resting on the other two (Fig. 3-2), was highly sensitive to vibrations of either the supporting platform or the air. A second form, comprising sharpened pencils of carbon resting loosely in conical depressions in carbon blocks and mounted on a sounding board, as shown on Fig. 3-3, was found to be extremely sensitive. Hughes explained that "these effects are due to a difference of pressure at the different points of contact, and are dependent for the perfection of action upon the number of these points of contact."[5] To identify the phenomenon thus exhibited, Hughes revived the term "microphone," coined and first used by Wheatstone in 1827 for a purely acoustic device (a sort of stethoscope) which he had devised "for hearing sounds when it is in immediate contact with sonorous bodies." The term "microphone" was used for many years to refer

Fig. 3-2. Hughes loose-contact microphone.

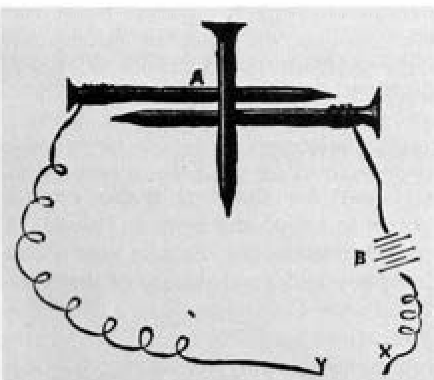
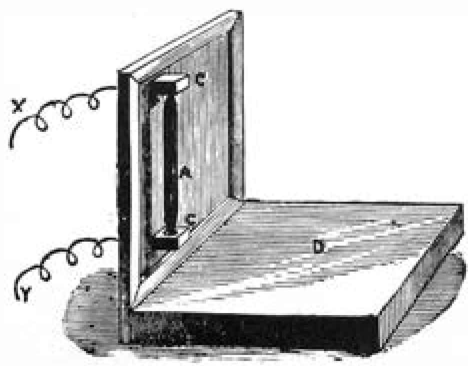


Fig. 3-3. Hughes loose-carbon-contact microphone.



[5] Perhaps the principle discovered by Hughes will be more understandable if restated in modern terms. He had found that an electrical resistance existed at the point of contact between two conductors. With low contact pressures and small contact areas, this resistance varied greatly with the amount of pressure applied. Thus the very small pressure variations in speech waves could produce the relatively large variations in resistance required to produce current changes of a magnitude sufficient for a useful telephone transmitter.

to sensitive loose-contact instruments. Later this term was used as a specific designation for the high-quality acoustoelectric transducers used in public-address, sound picture, radio, and television work. Curiously, it is still so used, even though these devices have not employed the Hughes microphonic principle for over 40 years, but the term is essentially unused in connection with the telephone transmitter, a device which still uses the principle.

The Hughes carbon-pencil microphone was further developed in France by Ader and his contemporaries. For commercial use, they constructed models having 6 to 12 pencils connected in series-multiple arrangements. An account, published in 1884, described a new acoustic effect discovered by Ader and demonstrated for the first time at the 1884 Paris Electrical Exposition in transmitting musical programs. He installed his microphones near the stage of the Grand Opera and telephones (receivers) at the Exposition Hall for listening. Each listener was provided with two Bell receivers (one for each ear) separately connected to two microphones that were situated some distance apart. As a result, the listener was enabled to picture in his mind the position on the stage of the performers and to follow their movements to the left or the right. This was the first demonstration of true binaural transmission.

Hughes and his contemporaries thus not only confirmed Bell's concept of a variable-resistance converter, as exemplified by his liquid transmitter, but had found a way to achieve large resistance variations with small acoustic inputs through the use of loose contacts between suitable materials. Neither the Hughes microphone nor the liquid transmitter was commercially practical for telephony but they had set the stage for developing such devices.

2.1.2 Learning to Apply the Microphonic Principle

2.1.2.1 Early Edison Transmitter. Credit for the first major application of the variable-resistance contact to telephony goes to Thomas A. Edison for his compressed-lampblack transmitter. Edison was a consultant for the Western Union Company and a subsidiary of this company, the American Speaking Telephone Company, was a Bell competitor. In 1877 he filed patent applications covering transmitters using variable-pressure contacts between metal and plumbago (or plumbago-coated) electrodes. Interference proceedings held up the granting of these patents for 15 years and there is no evidence that these devices were ever used. In the meantime he had found that highly compressed lampblack (carbon) in contact with metal provided a better variable resistance and he obtained a patent on this material.[6] The device

[6] T. A. Edison; U. S. Patent No. 203,016; filed March 7, 1878; issued April 30, 1878.

was used into the early 1880s and was found to be rugged and to provide a higher output than the magneto transmitters used by the National Bell Telephone Company at the time.

The lampblack transmitter, illustrated in Figs. 3-4 and 3-5, used a single contact between a metal diaphragm and the carbon disk so arranged that pressure on the carbon, and thus the resistance of the assembly, varied in accordance with the sound vibrations striking the diaphragm. The adoption of carbon for acoustic modulation of a direct current was the result of an extensive survey by Edison of available materials. The search included "hyperoxide of lead, iodide of copper, black oxide of manganese, graphite, gas carbon, platinum black, finely divided metals including osmium, ruthenium, silicon, boron, iridium and platinum, in fact, all conducting oxides, sulphides, iodides, fibers coated with metals by chemical means and pressed into buttons, liquids in porous buttons of finely divided nonconducting materials." The best material was found to be lampblack obtained from the combustion of light hydrocarbons, such as gasoline or naphtha, and compressed by the application of several thousand pounds of pressure into a solid disk.

2.1.2.2 Berliner–Blake Transmitter. Because of the superiority of the Edison instrument, the American Speaking Telephone Company made severe inroads into National Bell operations. Therefore, the Bell management instituted a patent infringement suit in 1878 on the basis that Edison's device operated on the variable-resistance principle covered by Bell's original patent. The suit was settled out of court in 1879. However, because of the early competition, National Bell intensified its efforts to produce a better transmitter.

These efforts resulted in the development of a superior transmitter, incorporating inventions of Emile Berliner and Francis Blake, which was named the Blake transmitter.

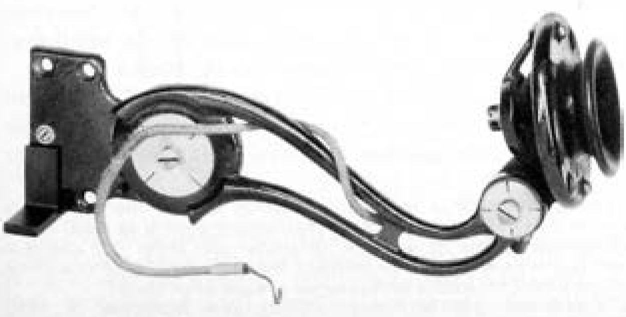


Fig. 3-4. Early Edison transmitter (1879).

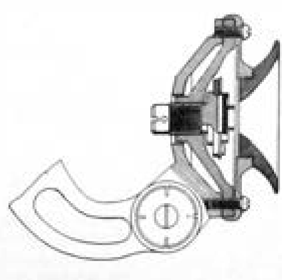


Fig. 3-5. Cross section of early Edison transmitter.

Emile Berliner, a native of Hanover, Prussia, was employed in Washington at the time of the Centennial Exposition in Philadelphia, and so learned of Bell's exhibit. He had also learned from a friend who was the chief operator of a telegraph office that a firm contact of the telegraph key was required to insure proper action of the sounder, and that few women were used as telegraph operators because of the lack of suitable strength in their fingers. He correctly deduced that a light contact might be used to provide the variable resistance needed for a telephone transmitter, constructed a device (Fig. 3-6) in which a steel ball pressed against an iron diaphragm, and found that the contact resistance did indeed vary as the diaphragm vibrated. Berliner filed a caveat for this invention on April 4, 1877; the U. S. Patent, No. 463,569, was not issued until November 17, 1891, but the National Bell Telephone Company purchased the right to use the invention and engaged Berliner in their service.[7]

The Berliner device was not commercially practical without further development since it was highly unstable, critical in adjustment, and unusable over a wide range of speech volumes without excessive distortion. The improvement of the device was undertaken by Francis Blake who joined the Bell Company in the summer of 1878. He had studied the reports of Hughes' work, and designed a transmitter on the microphonic principle that also included Berliner's concept. In this instrument one electrode was a platinum bead and the other was a hard carbon disk. Both electrodes were spring-mounted; the springs held them in contact with each other and also pressed the assembly against the diaphragm (which did not need to be a conductor) so as to be driven by its motion (Fig. 3-7). The edge of the diaphragm was encased in a single rubber band, stretched over the edge and extending about one-quarter of an inch inwards on both faces, an idea contributed by E. P. Wilson of Boston. This band served to hold the diaphragm firmly but not rigidly in place, and reduced the effects of mechanical vibration of the transmitter mounting. This transmitter was superior to the Edison instrument in clearness of articulation, in reliability, and in durability, but not in loudness. In 1881, Blake received four patents[8] covering this instrument.

After 1878, the Blake transmitter rapidly became the Bell System standard and remained so for a number of years even though the settlement of the patent suit against Western Union in November

[7] Curiously, the descriptive material in Berliner's patent implies the use of his device as a receiver as well as a transmitter and the illustrations show it so used. However, Patent No. 463,569 claims only its use as a transmitter and states that use as a receiver is the subject of a separate claim. Presumably the impracticality of this claim was recognized since no patent for such a device seems to have been issued.

[8] F. Blake; U. A. Patent Nos. 250,126 through 250,129; filed September 15, 1881; issued November 29, 1881.

1879 gave National Bell (soon to become the American Bell Telephone Company) access to the Edison transmitter patents. Aside from the transmitters acquired by the Bell Company from Western Union, the Edison lampblack transmitter received little use, but subsequent work by Edison on granular-carbon transmitters had a far-reaching effect on Bell transmitter development.

As a matter of interest, the frequency response characteristic of an 1878 model of the Blake transmitter was measured many years later, after equipment for making such measurements became available. The response curve is shown in Fig. 3-8, and was characterized by extreme peaks and valleys, the total variation being as much as 65 dB

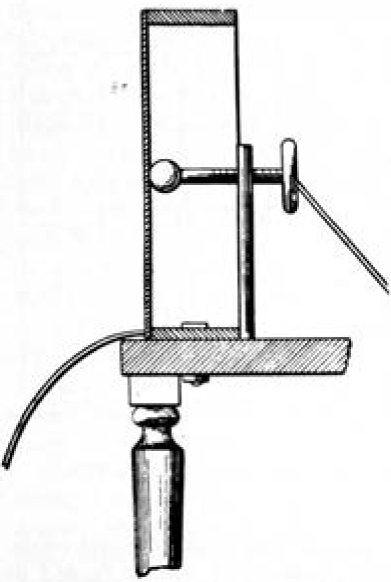


Fig. 3-6. Berliner microphone.

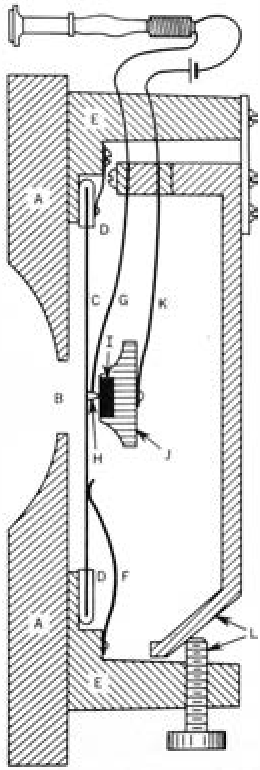


Fig. 3-7. Blake transmitter: cross section of working parts. A, door; B, orifice for sound waves; C, iron diaphragm; D, soft rubber ring covering diaphragm rim; E, iron mounting frame; F, damping spring; G, light spring holding platinum bead H; I, hard carbon disk set in metal backing J borne by heavy spring K; L, means for adjusting pressure at microphonic contact between carbon disk and platinum bead. (Redrawn from Rhodes 1929, p. 80)

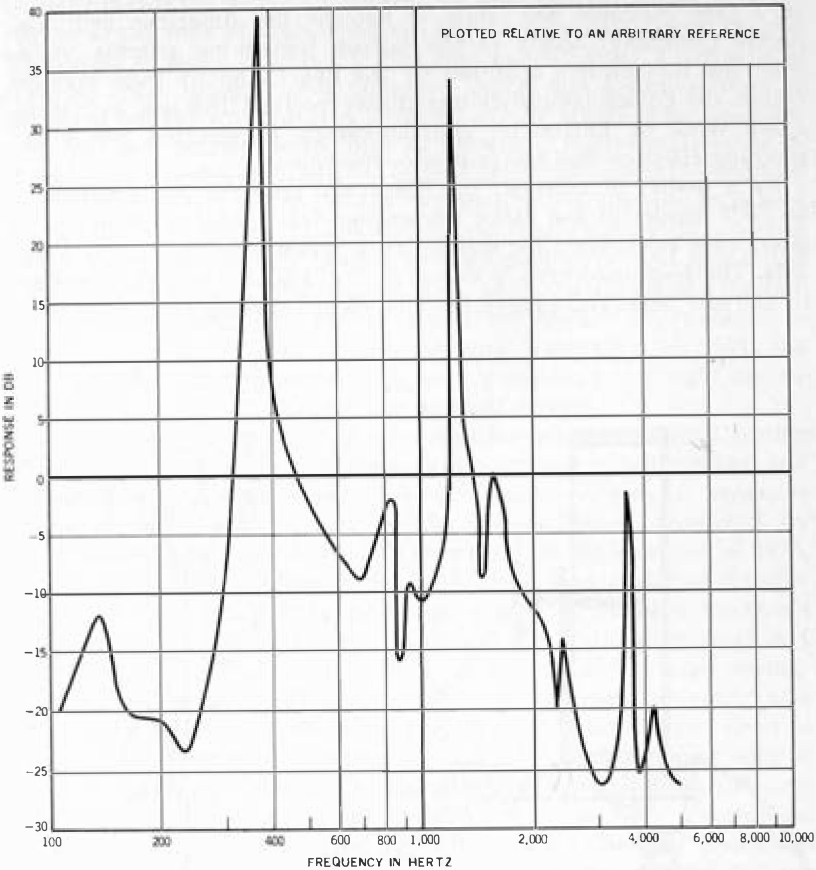


Fig. 3-8. Frequency response of a Blake single-contact transmitter of 1878.

from the peak at 380 hertz to the dip at 3,000 hertz. But despite peaks and valleys, useful output of the Blake transmitter was estimated to be nearly 20 dB higher than a contemporary magneto transmitter.

The Blake transmitter represented the peak development of single-contact transmitters, but at the very time when it was being introduced into commercial telephony the groundwork for its replacement was being laid by Henry Hunnings, an English clergyman. But before we discuss the important contribution of Hunnings, we should review briefly the inventions which carried the art to this point.

It should be clear to the reader that both Edison and Berliner had put to use the microphonic principle before it was disclosed by the

Hughes paper of May 1878. In the case of Edison there is some question as to whether or not he had appreciated as fully as Hughes the significance and generality of the microphonic principle. In his patent application (possibly to avoid conflict with Bell's patent) the claims emphasized the importance of the material rather than the principle. In the case of Berliner it seems clear that his objective was to take advantage of the resistance variation resulting from pressure variation on metals in loose contact. Today, some 90 years after the event, it appears that he clearly anticipated Hughes but probably did not have as broad an understanding of the principle. His transmitter performed rather poorly and was difficult to adjust and it was only the modifications made by Blake (after Hughes' disclosure) that made it a commercially usable device.

2.1.2.3 Hunnings Transmitter. On September 16, 1878, Henry Hunnings of Bothwell in Yorkshire received a British patent on a telephone transmitter using multiple contacts derived by partially filling a cavity between two electrodes with finely divided carbon ("engine coke") in a loose or free state.[9] An electrically conducting diaphragm of suitable metal—platinum, silver, ferrotype, or tinned iron—formed one electrode and a brass disk formed the other (Fig. 3-9). Because of the multiple parallel contacts, this instrument could carry higher currents than the single contacts of the Berliner, Edison, and Blake transmitters, and because of the multiple series contacts, its resistance could be made higher. Consequently, the amount of direct-current power that could be modulated by the diaphragm motion was substantially increased by this design, resulting in about a 9-dB increase in efficiency. Although the loose carbon particles had a marked tendency to pack together and become insensitive, Hunnings' United States patents[10] were purchased by the American Bell Telephone Company.

The packing characteristic of the original Hunnings design rendered it unsatisfactory for commercial use but it was developed into a more acceptable form by the engineers of the American Bell Company who redesigned it in 1885 to use a horizontal diaphragm and hard carbon granules. They also gold-plated the electrodes to reduce contamination of the electrode surfaces. This design was called the "long distance" transmitter because of its high output. Over ten thousand units of this type (Fig. 3-10) were used by American Bell in the following years.

[9] A modern transmitter may contain between 3,000 and 50,000 carbon granules, depending on design characteristics.

[10] H. Hunnings; U. S. Patent Nos. 246,512 and 250,250; filed May 14 and September 30, 1881; issued August 8, 1881, and November 29, 1881.

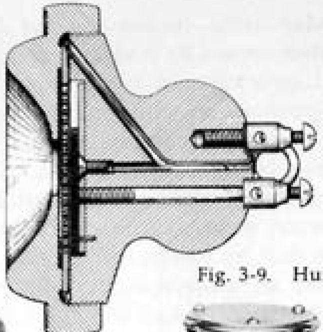


Fig. 3-9. Hunnings transmitter (1882).

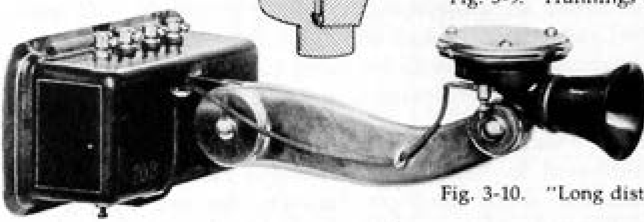


Fig. 3-10. "Long distance" transmitter.

Blake also attempted to adapt the Hunnings concept to his structure by replacing the single contact with a chamber containing granular carbon. Three designs were coded in the 1888–1890 period and a total of about 13,000 units manufactured. A typical design is shown in Fig. 3-11.

2.1.2.4 Edison Carbon-Granule Transmitter. A further improvement in the multicontact transmitter occurred in 1886, when Edison introduced the use of granules of anthracite coal that had been carbonized by roasting.[11] When properly processed, this carbon was much superior to previous forms of carbon in hardness, uniformity, and durability. It was quickly adopted as the standard transmitter material and has remained so for 80 years.

Commercial carbon had been available for many years as charcoal, obtained from wood or vegetable matter, and as coke, obtained from certain types of bituminous coal. It had also been available in molded forms such as rods, plates, and crucibles. The molded forms were produced by pulverizing the charcoal or coke and mixing this dust with syrup or molasses. This mixture was then compressed in prepared molds and raised to white heat for an hour or more; the compactness and hardness were increased by repeated soaking in syrup and heating. The finished product contained all of the non-volatile impurities of the raw material. As better-quality carbons were needed (specifically for arc light pencils), lampblack, retort

[11] T. A. Edison: U. S. Patent No. 406,567; filed February 19, 1886; issued July 9, 1889.

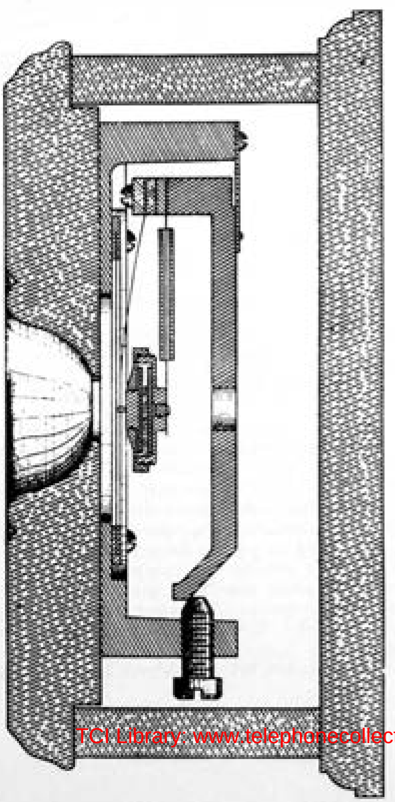
carbon, and gas carbon were substituted for coke and charcoal as the raw material; as noted previously, Edison used a selected lamp-black for preparing his early transmitter buttons.

The established superiority, for telephone transmitter usage, of roasted granular carbon made from anthracite coal relative to other known materials is due to its unique combination of properties which include: great hardness, elasticity, surface geometry (roughness), heat conductivity, electrical conductivity, infusibility, and the fact that its oxides are gases.

2.1.3 Large-Scale Commercial Production

The Edison granular carbon provided the last building block needed for developing variable-resistance transmitters which would be efficient, stable, and capable of large-scale production. Even though the basic principles and a suitable granular material had been developed, a practical design still had to be devised. A main problem remaining was the mechanical and electrical "packing" of the carbon

Fig. 3-11. Blake granular-carbon transmitter (1888). (Frederick 1931, Fig. 17)



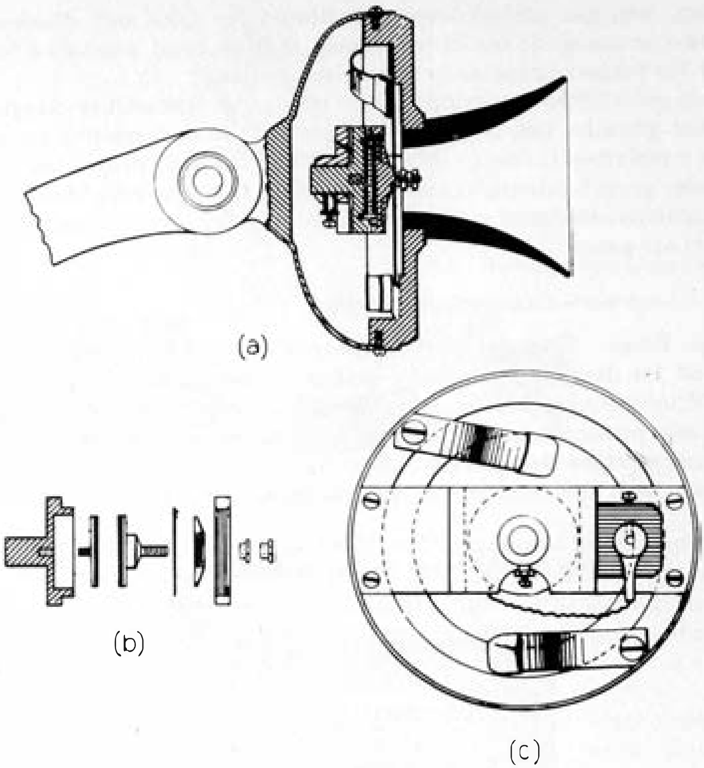


Fig. 3-12. White's solid-back transmitter (1890).

granules which greatly reduced the output.[12] It was frequently necessary to jar or shake the transmitter before normal operation could be obtained. A major step in overcoming this obstacle was made by Anthony C. White, an American Bell engineer, with his invention of the "solid-back transmitter" in 1890.[13]

[12] "Packing" is of two types, electrical and mechanical. Electrical packing is a form of the "cohering" of conducting particles that later was to prove so useful in early wireless. When this occurs, the carbon granules cling together, the resistance is greatly reduced, and the sensitivity to sound waves may be reduced to a few percent of the normal state. Fortunately, roasted carbon does not cohere readily, requiring a voltage of 1½ volts or more between contacts as compared to about 1/10 volt for metals, and can be reasonably well controlled by proper design and selection of operating voltages. Mechanical packing is due to a settling and compressing of the carbon mechanically, and in early transmitters was often associated with heating effects in the transmitter itself. The understanding of the causes of packing and design techniques to minimize it which were developed in the 1890s represented major advances in transmitter design.

[13] A. C. White; U. S. Patent No. 485,311; filed March 24, 1892; issued November 1, 1892.

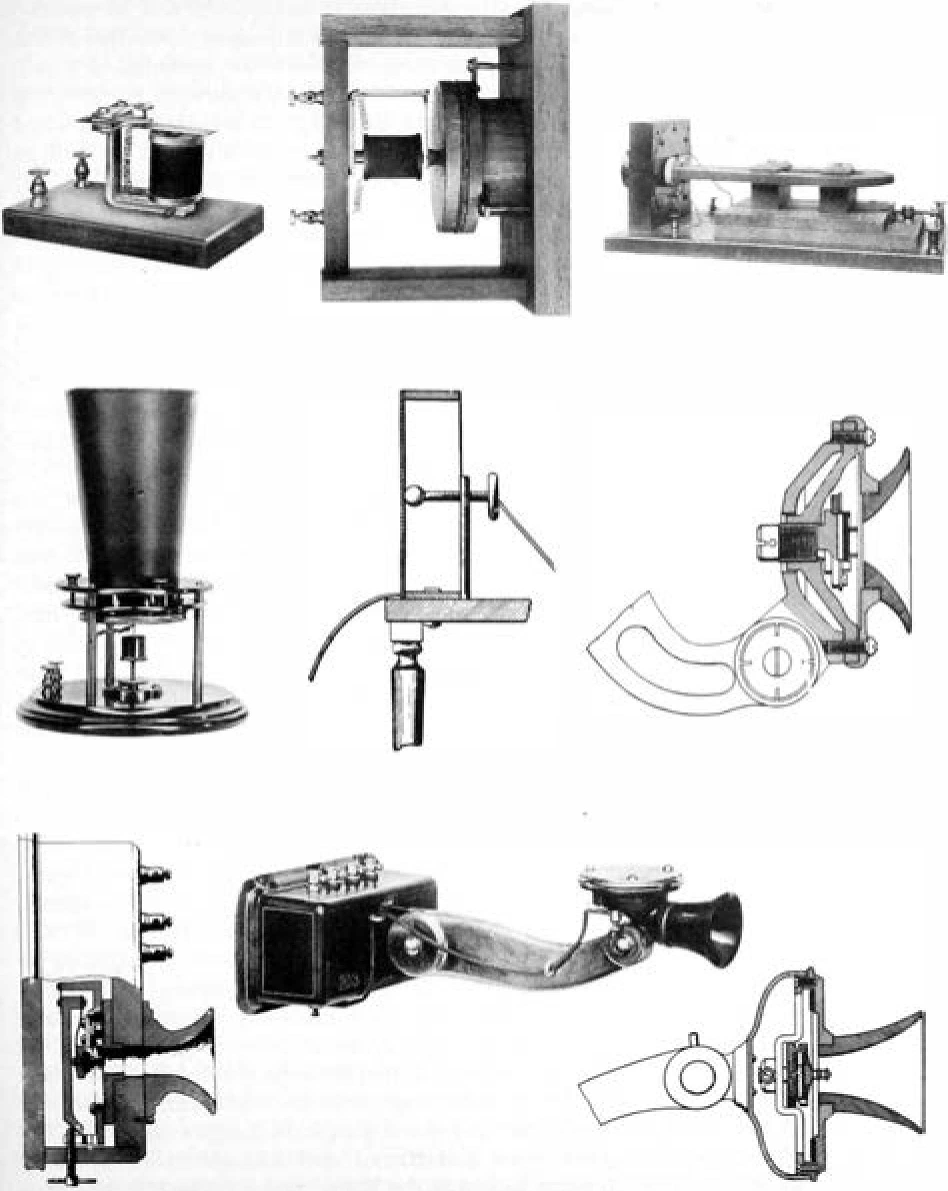


Fig. 3-13. Devices preceding the standardized solid-back-transmitter design of 1895 (No. 229 transmitter).

The structure and essential elements of the solid-back transmitter are illustrated in the three parts of Fig. 3-12, taken from the White patent. Part (a) is a cross section showing how the essential element, the so-called button, is placed between the diaphragm to which it is attached at the front and the metal bridge to which it is fastened at the back. This heavy metal bridge, which solidly held the button in place and gave rise to the "solid back" name, is shown more clearly in the rear view of the transmitter given in part (c). The latter also shows quite clearly the two rubber-covered springs which held the diaphragm against its support and provided mechanical damping to prevent unwanted spurious responses and thus control the electrical characteristics of the device. Following earlier practice the periphery of the diaphragm was encased in a rubber band. The button, which was the uniquely important element, is illustrated in part (b) of the figure. This button contained a chamber partially filled with roasted carbon granules providing multiple contacts between two carbon-disk electrodes. The rear electrode was attached to the bridge and the front electrode was free to move under the influence of the diaphragm (to which it was attached by a screw). The front enclosure of the button was an annular ring of thin mica which permitted motion of the electrode in a piston-like manner and insulated the two electrodes so that the current path was through the carbon granules. In addition to being more reliable and less susceptible to packing, the output of the solid-back transmitter was about 2 dB higher than the Hunnings long-distance transmitter.

This transmitter proved to be the prototype for nearly all the customer-used transmitters produced during the 35 years following White's invention. Before relating the evolution of the solid-back transmitter into high-production designs, it may be interesting to recall the devices which preceded it by reviewing Fig. 3-13 which shows the significant development steps in roughly chronological order. The three devices at the top are early forms of the magnetic transmitter (mentioned in Chapter 1 and discussed further in Section 2.2 of this chapter). The second line shows the early steps in developing the variable-resistance principle beginning with Bell's liquid transmitter and continuing with the early efforts of Berliner and Edison in applying the microphonic principle. All of these devices were produced within roughly the first two years after Bell's invention and received at most very limited commercial application. During the next few years the more practical designs shown at the bottom of the figure were introduced and manufactured in some quantity. These designs included the Blake device, which culminated the application of the single-contact principle, and the long-distance transmitter which embodied the first practical application of the

Hunnings multicontact invention. Finally, by 1890, 15 years after Bell's original work, the solid-back transmitter had been developed and was to serve as the basis for transmitter design until the mid-twenties of the following century.

The evolution of the Bell System solid-back design is illustrated in Fig. 3-14. By 1895 the design had been considerably refined and standardized as the No. 229. In this device the diaphragm was made of aluminum instead of the iron or steel used prior to White's solid-back design. The lighter diaphragm responded more readily to air motion and was retained in the later models.

Beginning in 1906 a number of special transmitters were built in which the case (and other parts coming in contact with the user) was insulated from the parts connected to the line. This desirable safety feature was standardized in the 329 transmitter introduced in 1913. Several minor changes were also made at this time including the use of steel instead of brass for the bridge (a change retained in later designs). Electrically the performance was unchanged.

In the Nos. 229 and 329 transmitters, the simple disk diaphragm was attached to the front button-electrode by a threaded nut, but this arrangement was changed to a pressure contact in the Nos. 323 and 337 (introduced in 1917 and 1919 respectively). The No. 323 transmitter employed the same button structure as the 229 and 329 but the vibratory system was changed to improve efficiency. A new form of diaphragm was used, the edge being turned over to provide a narrow contact with its seat at the outer edge. This edge rested on an insulating ring of varnished cambric (the long-used rubber band being omitted from the periphery). The dual damping springs were also omitted, a single spring near the center of the diaphragm being substituted. This spring not only provided damping but also served both to conduct current to the front electrode and to hold the electrode against the diaphragm. The diaphragm was thus less loaded, both at the edge and at the center, and so had increased freedom of vibration. The 323 structure was also arranged for bracket mounting and coded the 353.

The No. 337 transmitter of 1919 used a new form of button with a conical back and small back-electrode. This redesign increased the transmitter resistance with accompanying improved efficiency. However, since the high efficiency was undesirable on short loops and the transmitter tended to be noisy under these conditions, its use was limited to long loops, originally those over 300 ohms in resistance. Later this limit was reduced to 150 ohms.

During the period we have been discussing, the telephone plant had been undergoing a major change. Originally, the power for the transmitter was supplied at the user's premises by one to three

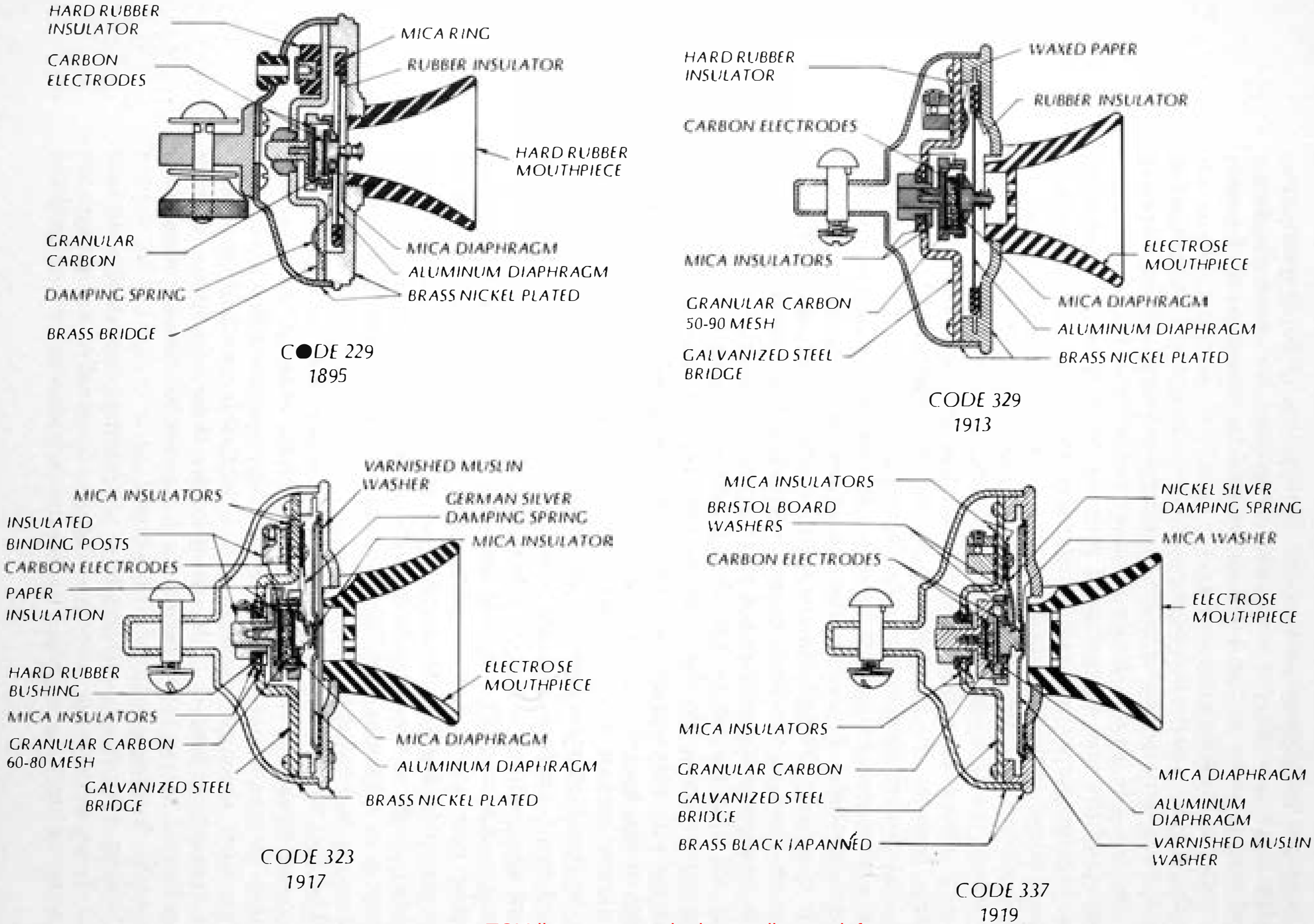


Fig. 3-14. Sectional views of transmitter showing development steps from 1895 to 1919.

cells of battery, each with an output of 1.5 volts. It was customary to design transmitters for this "local battery" service with a low resistance in order to maximize the amount of direct-current power available from the low-voltage source. Usually a resistance of from 5 to 10 ohms was achieved. The requirements changed, however, as "common battery" offices were introduced beginning late in 1893. In these offices a battery at the central office with an output of 16 or more volts was used to supply direct-current transmitter power over the telephone lines. This greatly reduced the cost and nuisance associated with local supply and the practice grew rapidly. However, with common-battery supply the resistance of the supply circuit was large and in spite of the higher-voltage source the transmitter current was usually smaller than that obtainable with local batteries.[14] In order to compensate for the loss in efficiency which this would cause, transmitters for common-battery use were designed for a higher resistance (usually 25 to 40 ohms). The original solid-back transmitter used a low-resistance button (No. 3 type) but a high-resistance button (No. 7 type) was developed for common-battery service. For some time both types remained in use, according to the battery supply employed, but before long the high-resistance type was used universally.

While the basic electrical design was not greatly changed over a long period from its 1890 prototype, very significant changes were made in manufacturing methods and particularly in the manufacture and control of the carbon granules. Exacting procedures were required in all stages of transmitter production. No contamination of the carbon electrodes or granules was permissible; the quantity of carbon granules used in the buttons was closely controlled. The tension of the diaphragm and button assembly required accurate adjustment and the completed instrument was voice-tested to assure proper performance. For a considerable period of time this work was done by the Western Electric Company at 463 West Street in New York City. A view of the transmitter assembly operations about 1906 is shown in Fig. 3-15.

The external appearance of Bell System transmitters varied widely over the years depending both on the electrical design and the manner in which they were used. During the first 50 years of telephony roughly a hundred different transmitters had been designed. Figure 3-16 shows in chronological order some of the variations. By 1925 all the designs shown had been discontinued and the Types 323,

[14] The current also varied with loop length, unfortunately being lowest on long loops where line attenuation was highest. Despite this problem, the advantages of common-battery operation were so great that it was rapidly adopted in new offices.

Fig. 3-15. Transmitter assembly on the 9th floor of Section C at 463 West Street, New York City about 1906.

353, and 337 were the basic standard models for customer use and were designed to fit in various mountings. Many of the models shown in Fig. 3-16 represent minor variations in mounting or design for limited special usage (police, military, etc.) and will not be individually identified. A few of the types illustrated were intended for operator use and will be mentioned later.[15] Some of the more important types designed for customer use are listed below, together with their production dates and quantities.

Identification No. in Fig. 3-16	General Type	Date First Produced	Quantity Manufactured (Bell System)
1–11	Magneto box telephones	1877	6,000
12–16, 19	Blake transmitters (16 and 19 most used)	1878	340,000
20, 21	Edison transmitter	1881	2,000
25–28	Long-distance (modified Hunnings) transmitter	1886	12,000
30–32	Blake granular-carbon transmitter	1888	13,000

[15] Telephone transmitter Type 234 (No. 77 in Fig. 3-16), introduced in 1900, is most significant since it introduced a pattern used for many years.

38–42, 54, 58, 69, 80, 85, 88	Early solid-back transmitter for local-battery use	1891	930,000
55–57, 64	Early solid-back transmitter for common-battery use	1894	590,000
65, 66, 81, 87	Solid-back transmitter of type 229 and related types	1895	4,600,000
94	Type 329 transmitter	1912	3,500,000
(Not shown)	Types 323 and 353 transmitters	1917	7,000,000*
(Not shown)	Type 337 transmitter	1919	700,000*

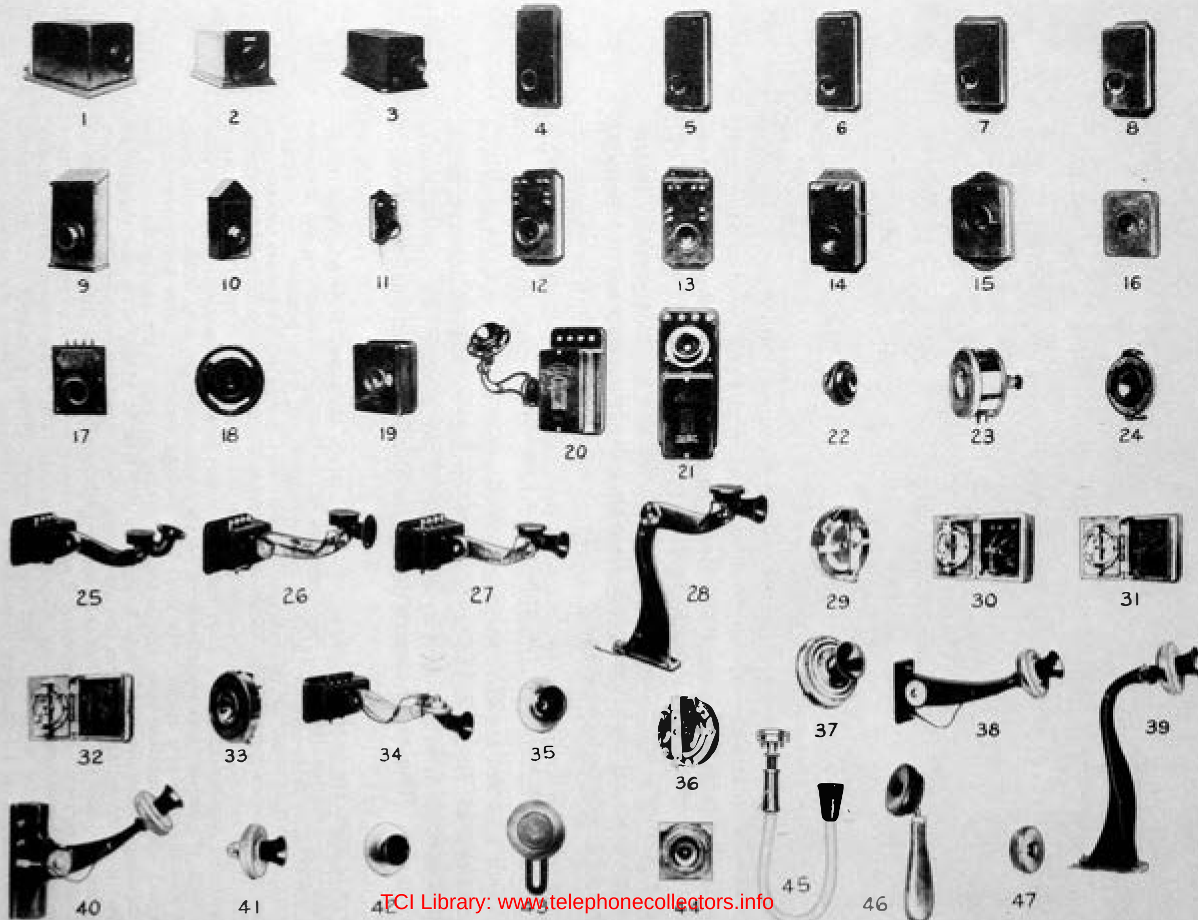
* Estimated production through 1925.

It will be noted that, despite its shortcomings, the Blake transmitter was the mainstay of the business in the early years. The Blake granular-carbon transmitter, first produced in 1888, presented so many problems due to packing that it did not receive large commercial use. White's solid-back design rendered the granular-carbon instrument practical and it was only then that the true potential of telephony could begin to be realized.[16]

2.2 Receivers and Other Electromagnetic Converters

Bell was not the first to construct a device to convert electric waves into sound. In 1837, Dr. C. G. Page, in Salem, Massachusetts, noted that sounds were emitted by a magnet if the magnetism was suddenly changed. (The magnetostriction principle later was used by Reis in

[16] In 1908, Hammond V. Hayes, who was closely associated with telephone developments between 1885 and 1907, testified in a court case concerning the performance of early transmitters. Some of his testimony will be of interest: "In the early days, 1879 and 1880 or thereabouts, we had two forms of telephone transmitters, the Blake transmitter and the Edison transmitter. These two were both objectionable. The Blake transmitter was unsatisfactory, for the reason that if you tried to make it powerful and spoke closely to it the instrument broke, rattled, was indistinct. The Edison transmitter was objectionable, because you could get very little volume from it except when you spoke with your lips pressed directly against the mouthpiece, which was a condition very hard indeed to get subscribers in the field to do. . . . the Hunnings transmitter . . . was bought by the Bell Company with the expectation that it would relieve us from the objections which were inherent in the two earlier forms . . . But we found that the Hunnings transmitter had a difficulty inherent to it, which rendered it uncommercial (i.e., packing)." Several men competed in an effort to produce a workable design. Hayes continued, "The other experimenters' results did not compare with those that Mr. White got. White produced this instrument, which we called a solid-back instrument to distinguish it from the pivoted instrument, which had a loose back, which the other men had in competition with White. White's instrument proved to be all right. We made the instrument up in the model form, and it has been in use from that day until the present time with practically no change whatever in design or proportions."


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

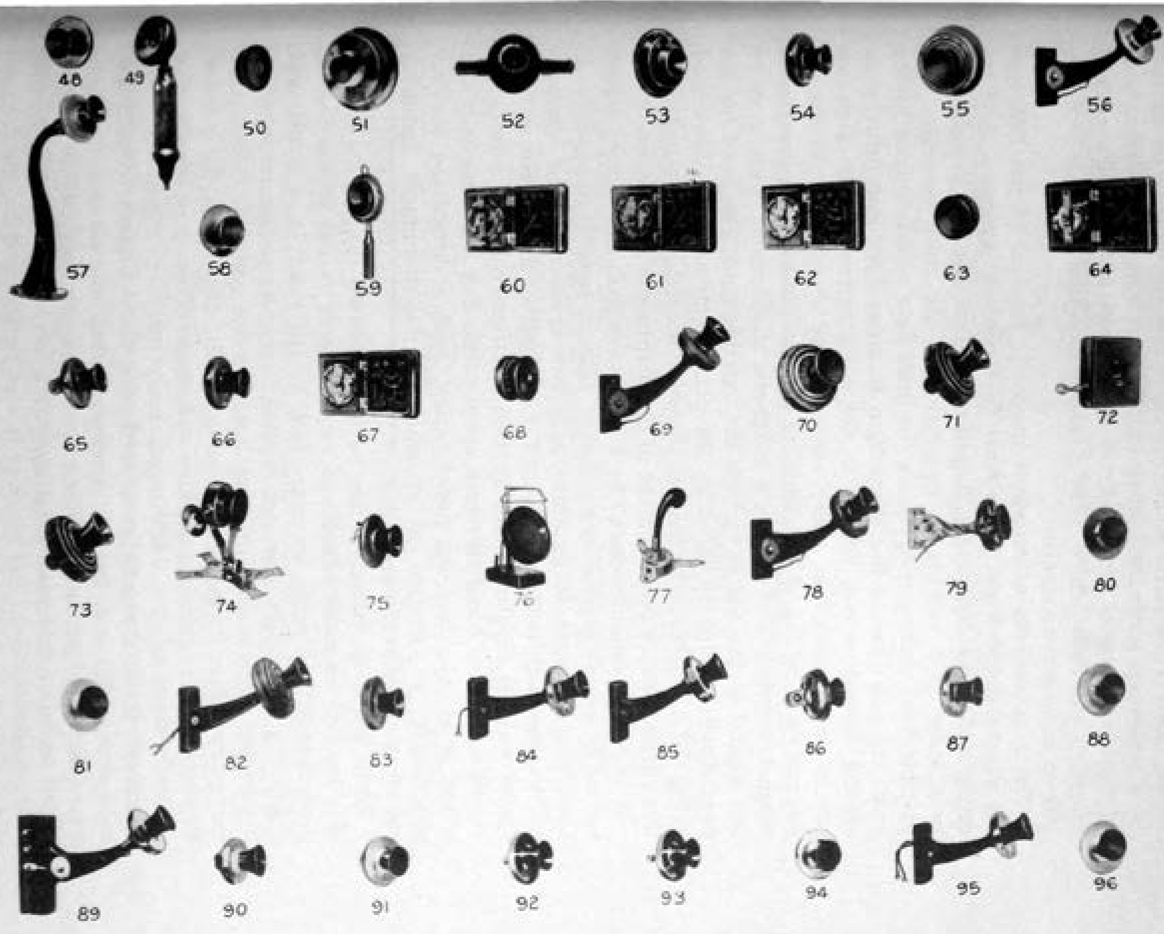


Fig. 3-16. Telephone transmitters which were once standard but were no longer being manufactured in 1925.

his experiments.) In 1870, Varley discovered another principle (put to use many years later) when he found that sound may be emitted by a condenser (capacitor) when its charge is varied. However, none of these discoveries led to direct application as did Bell's electromagnetic principle which, after nearly a hundred years, is still employed in telephone receivers.

The requirements of an electromagnetic receiver of the Bell type are simply stated: First, there must be a constant magnetic field exerting a pull on a diaphragm. Superimposed on this field there should be a lesser field varying in proportion to the voice currents. This variable field, aiding or opposing the fixed field as determined by the voice currents, causes the diaphragm to move back and forth, generating acoustic waves similar to the voice currents. Such a device can be implemented in many ways: a very simple one, illustrated schematically by Fig. 3-17a, consists of a simple bar-type permanent magnet exerting a force on a magnetic diaphragm, with a voice coil wound on the bar to provide the variable field. The need for the fixed magnetic bias provided by the permanent magnet is illustrated in Fig. 3-17b. Without the bias, as shown at the left, positive and negative currents cause displacement of the diaphragm in the same direction and the sounds produced by a sine wave with frequency f would be made up of several components with a tone at frequency $2f$ being preponderant. With a bias, as shown at the right, the displacement varies both with polarity and magnitude of the current and the sound wave is similar to the applied electrical wave. A number of variations in this basic structure are possible but fundamentally the action is as described. As noted earlier, such a device can also act reciprocally, a variation in pressure on the diaphragm inducing electric currents in the voice coil.

While the basic principle can be implemented simply, the design of a receiver to meet telephonic requirements is not so easy since performance is affected by many factors. For example, a diaphragm clamped at the edges does not vibrate freely. At certain frequencies it vibrates more readily than at others, these resonant frequencies being determined by the size and weight of the diaphragm, the manner in which it is supported, the size and shape of the acoustic cavities on both sides of the diaphragm, temperature changes, etc. The effects of these characteristics were not fully subject to analysis until about 1920 and, for the years preceding, the relation between mechanical design and acoustic performance was largely determined on an experimental basis.

The performance of the receiver is also determined by the magnitude of the fixed and variable magnetic fields. This was appreciated

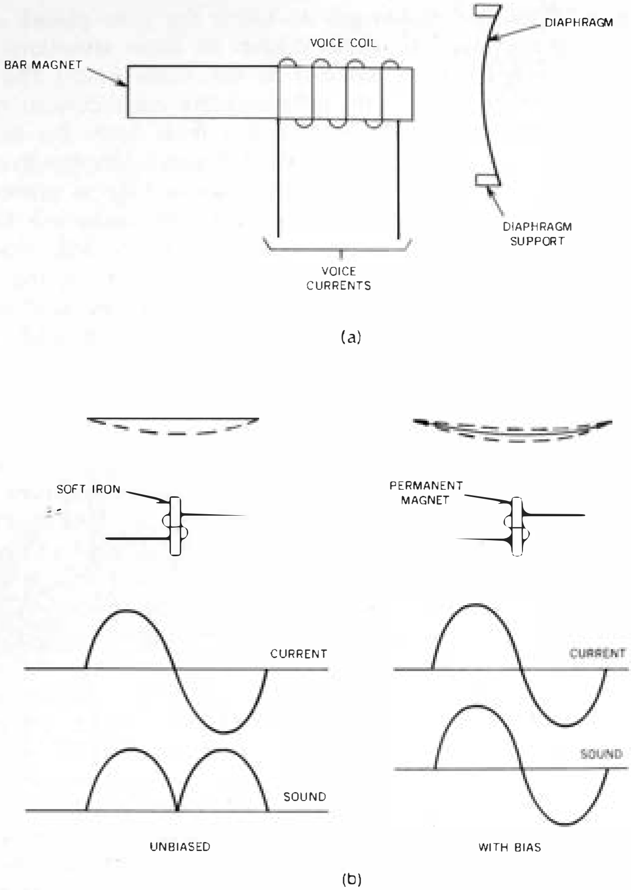


Fig. 3-17. (a) Schematic illustration of Bell-type electromagnetic receiver. (b) Effect of magnetic bias on reproduction of sound waves from a sine-wave current applied to Bell receiver.

in a general way very early in the development of the receiver,[17] but the complete mathematical analysis did not come for a number of years. In a somewhat simplified way it can be said that the force on the diaphragm can be broken down into three components. One is a steady force determined by the fixed magnetic field and does not directly affect the receiver performance so long as it is not large

[17] Maxwell indicated the effects of the fixed and variable fields in work published in the late 1880s.

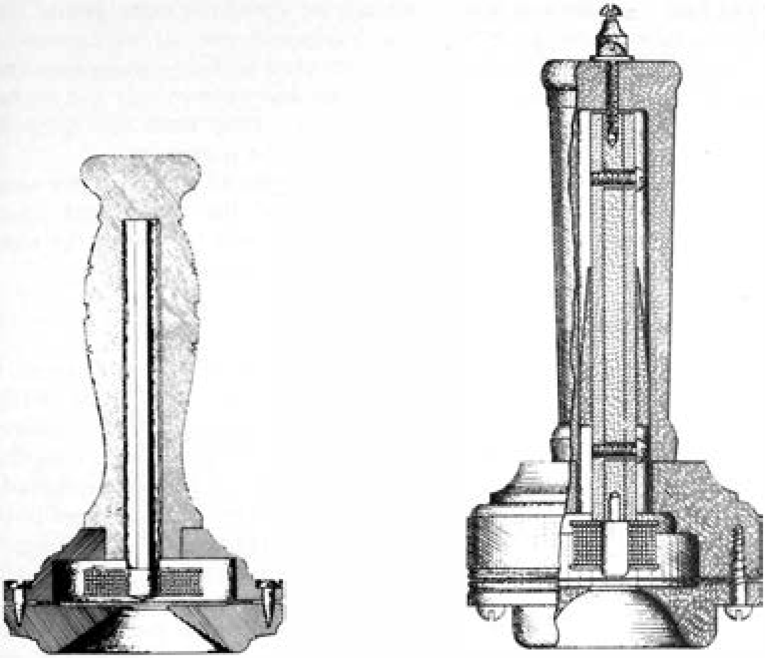
enough to cause the diaphragm to touch the pole pieces or cause saturation of the magnetic field. (Either of these situations would, of course, prevent a variation due to the voice field.) The second component is the variable force following the voice-current variation and its magnitude depends on the flux from both the fixed and variable fields. Because this component depends on the fixed field, it is desirable to have the fixed magnetic flux as high as practical consistent with the other limitations. The third component is a distortion component that is manifested principally as a sound with double the applied frequency. This component is dependent on the relative strength of the fixed and variable fields and can be kept small by having the fixed field large relative to the variable field, another reason for having a strong fixed field.[18]

2.2.1 Early Magneto Converters

The electroacoustic converters of the magneto type, used in Bell's early work, are covered in Chapter 1. The earliest ones used electromagnets which served both as the voice coil and, powered by a battery, as the source of the fixed magnetic field. Bell's reed receiver (see Fig. 1-5) had only a rudimentary diaphragm but a true diaphragm was incorporated in the "gallows" telephone (see Figs. 1-7 and 1-8) and all subsequent models. At first the diaphragm was non-metallic with an iron plate or slug serving as a driving armature, but the iron diaphragm was used in the Centennial Exposition demonstration of 1876 and thereafter was a permanent feature of the design. The permanent magnet was introduced by Watson in 1876 and thereafter completely superseded the electromagnet except for some experimental designs investigated in the 1919–1922 period.

Except for a few demonstrations with the electrolytic transmitter, the electromagnetic converter was used as transmitter and receiver both in early demonstrations and service installations. Bell's "box" transmitter was one of the earliest used for customer service and some 6,000 of the various models were constructed. Initially, with this device, the user alternately talked into or listened to the sound from the diaphragm behind a shallow opening. The active mechanism consisted of a large U-shaped permanent assembly of two bar magnets with soft-iron pole pieces on which the voice coils were wound. The diaphragm in this and all subsequent magneto converters was a circular piece of thin iron supported at the periphery. A typical telephone of this time is shown in Fig. 1-16 and several others are shown in Fig. 3-16 (item 1 through item 11).

[18] The extreme case, for which the fixed field is zero, is illustrated at the left in Fig. 3-17b.



(a) (b)

Fig. 3-18. Butterstamp receivers with wooden cases: (a) early model; (b) later model with compound magnet.

This awkward device was replaced by a hand-held converter (see Fig. 1-19) which came to be known as the "butterstamp" receiver because of its overall shape.[19] The fixed magnetic field was provided by a bar magnet along the axis of the case with a single voice coil at the diaphragm end, usually wound on a soft-iron pole piece. Two early types are shown in cross section in Fig. 3-18, one with a single bar magnet, the other with a two-bar assembly. In these models the case was made of wood but this material changed in size with moisture content giving rise to variation in the spacing between the pole pieces and diaphragm. Before the end of 1877, hard-rubber was being used for the case and this dimensionally stable material continued in use for roughly 50 years. The general shape of the

[19] It may be necessary to remind present-day readers that in the nineteenth century butter was sold in bulk. The butter stamp was a wooden device used by the housewife for molding butter into round pats, with a design on top, to make it more attractive for table use.

hand-held receiver was also continued for about the same period but with numerous changes in external details and internal mechanism.

The early butterstamp converters were used both as transmitters and receivers; some installations used a single unit alternatively and others provided two units, one for each function. They were also used as receivers with the box telephone serving as the transmitter.

But the use of the electromagnetic converter as a transmitter soon came to an end with the introduction of the Blake and other variable-resistance transmitters around 1878, and thereafter the electromagnetic device was developed solely as a receiver.

2.2.2 High-Production Hand-Held Receivers

The original butterstamp receiver evolved rapidly into a design suitable for commercial use and was coded the No. 101 in 1877 (Fig. 3-19). In many ways this was a remarkably advanced design incorporating features that were to remain satisfactory for many years such as the high-quality hard-rubber case and cap, the latter with internal threads for screwing to the case. The No. 101 receiver was redesigned, without change in code, in 1884 and again in 1898. In each case the changes involved mechanical details not affecting electrical performance. In all, over a half-million of the 101-type receivers were manufactured.

The magnetic circuit of the unipolar receiver was inefficient since the flux had a long return path through the air. Better magnetic circuits had, in fact, been used in Bell's Centennial Exposition instruments of 1876 (Fig. 1-12) and in the box telephone, but these designs did not readily lend themselves to the hand-held receiver which was proving so convenient.

By 1890, continued experiments by Bell and Watson had established that considerable additional improvement could be obtained with a

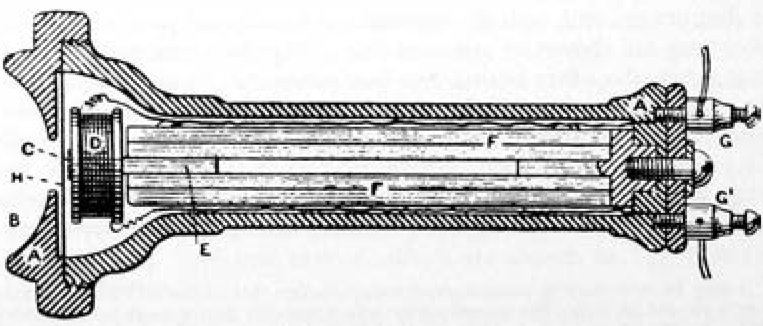


Fig. 3-19. Cross section of large hard-rubber hand-held receiver (code No. 101) of 1877.

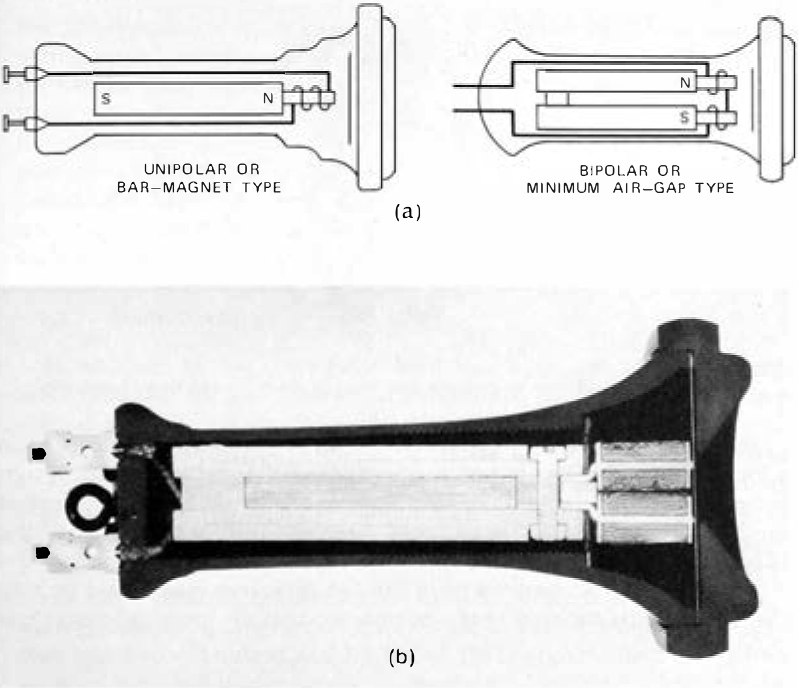


Fig. 3-20. (a) Two basic types of hand-held receiver. (b) Early bipolar receiver, 1895 version.

U-shaped magnet fitted with two pole pieces and two voice coils connected in series, and means were soon devised for fabricating such a structure in a hand-held receiver. In this bipolar structure the main flux path was short, comprising the two pole pieces, the diaphragm, and the two air gaps between the pole pieces and the diaphragm. The improved magnetic circuit of the bipolar receiver is clearly shown in Fig. 3-20a which compares it with the unipolar type. Figure 3-20b shows an 1895 design implementing the bipolar principle.

The amount of increase in receiver efficiency obtainable by improving the magnetic bias was limited by magnetic saturation in the diaphragm. This had been discovered as early as 1879 when the use of multiple magnets (Fig. 3-21) was tried with disappointing results. Even though the gain in efficiency was restricted by saturation, sufficient benefits were achieved to warrant the use of the bipolar design and such a structure has been used almost universally since about 1895.

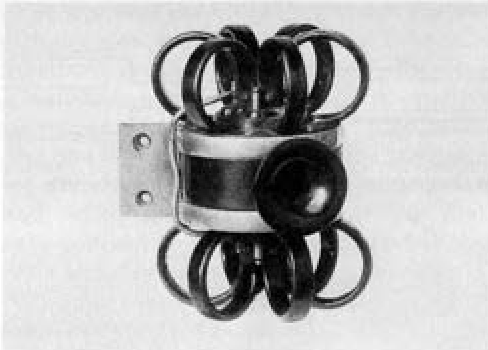


Fig. 3-21. Early attempt to increase receiver efficiency by addition of magnets.

The first bipolar receivers were made in 1890 (coded the No. 108) in a lot of about 500, but it was not until 1894 that an improved design became available. Several changes were made in 1895 and 1896 (coded the Nos. 111 and 112). These types were used quite widely until the 122-type receiver (Fig. 3-22) was introduced in 1902. This remained the standard hand-held receiver until the 144-type was introduced in 1912 (Fig. 3-23). This receiver was not only several dB more efficient than

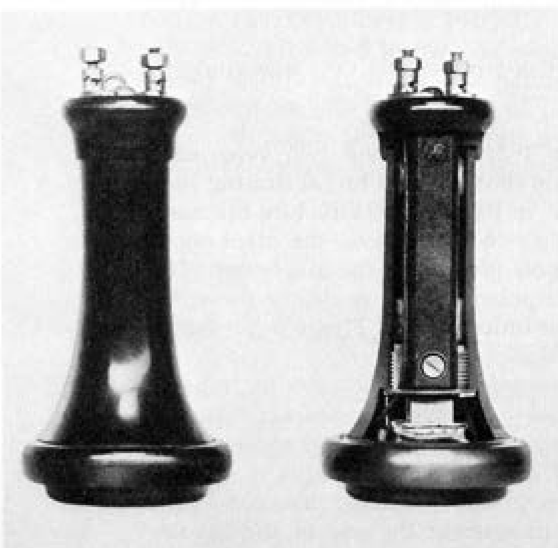


Fig. 3-22. The 122-type receiver of 1902.

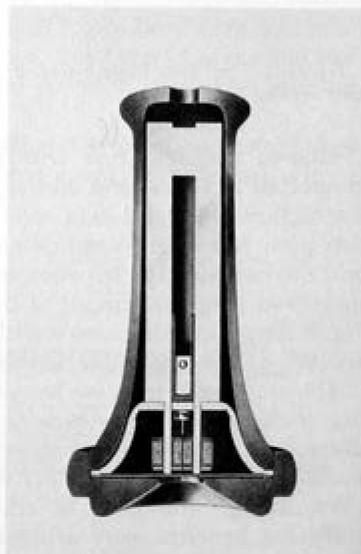


Fig. 3-23. The 144-type receiver of 1912.

the 122-type but also included a number of mechanical and manufacturing improvements. The cord terminals were placed inside the case so that the user could not come in contact with the electrical wiring. A brass cup welded to the magnet structure provided a rigid support for the diaphragm and reduced the variability in spacing between the pole pieces and the diaphragm. The cup also provided a closed space behind the diaphragm and kept the coils and air gaps free from dust and dirt. Typical manufacturing improvements were the use of form-wound coils and spot welding to join the pole pieces to the magnets.

2.2.3 *Small Receivers*

In addition to the large hand-held receivers just described, which were intended for the commercial subscriber, there was also a need for a small receiver for special applications, particularly one which could be mounted on a headband for operators and central-office workers who required the use of both hands in their work. The major problem was the production of a permanent magnet that was small, light, and sufficiently powerful to meet the requirements for efficiency and freedom from distortion. Some of the small instruments also had to meet other special requirements to permit their use outdoors, in diving helmets, and so forth. As a result, of the 64 designs of receivers that had been standardized by about 1912, approximately two-thirds were of the small types. Most of them were used with headbands to support them in position on the ear but one at least was attached to a silk cap worn by the operator.

The first small receivers made in any quantity were designed in 1884 and 1885 by Richards. These were single-pole receivers using a long bar-magnet external to the receiver case (Fig. 3-24). Originally these receivers used some of the parts of the 101-type hand-held receiver, but later there was some reduction in size. Several thousand were made. Berliner proposed the use of a magnetic headband in place of the external bar magnet used by Richards but the idea, while ingenious, seems to have had little more in its favor. In 1885, Gilliland introduced an interesting design with a cast-iron magnetized case supporting the diaphragm and serving as one pole piece. The voice coil was wound on a soft-iron pole piece in the center. Thus this receiver had a magnetic circuit, closed except for the gap between the central pole and diaphragm.

The bipolar principle, with its improved magnetic circuit, provided the most practical way to achieve sufficient magnetic bias with a magnet small enough to fit inside the case. The first small receiver incorporating this design was the "watch case" receiver (Fig. 3-25) which was introduced in 1894 and became the prototype for future

Fig. 3-24. Richards head telephone, old style of 1884.

small receivers. After some minor modification in the bipolar design, the No. 128 receiver was introduced in 1900 and it continued as the standard operator's receiver until 1920 when it was replaced by the more efficient No. 528. Well over 400,000 of the 128-type receiver were manufactured. This receiver was internally similar to the watch-case receiver, containing a small three-leaf magnet within the case. It was originally used with the rather heavy two-piece headband with which it is pictured in Fig. 3-26a. Simpler headbands were introduced later

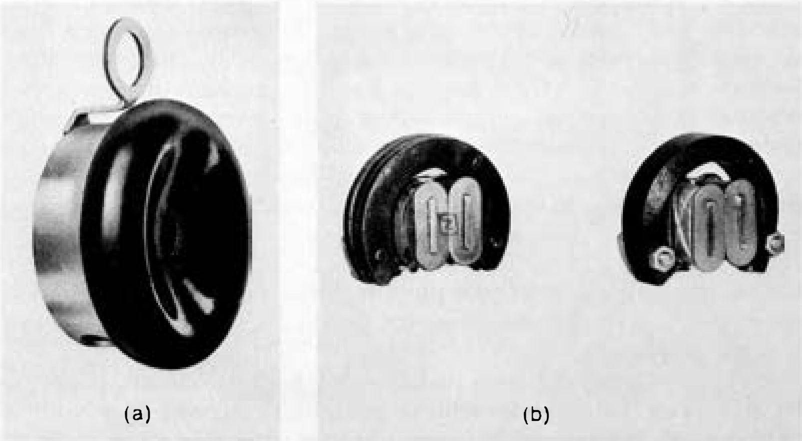


Fig. 3-25. First bipolar head receiver: (a) watch-case receiver (code No. 121) of 1894; (b) magnet and coil assembly.

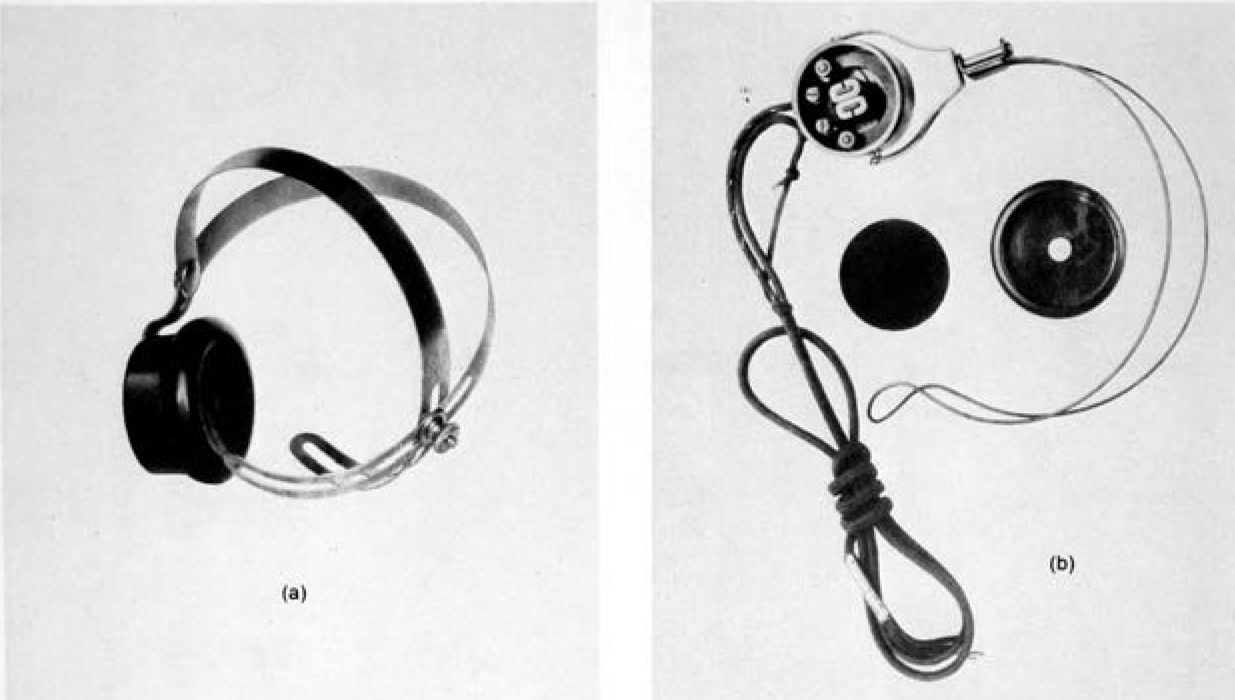
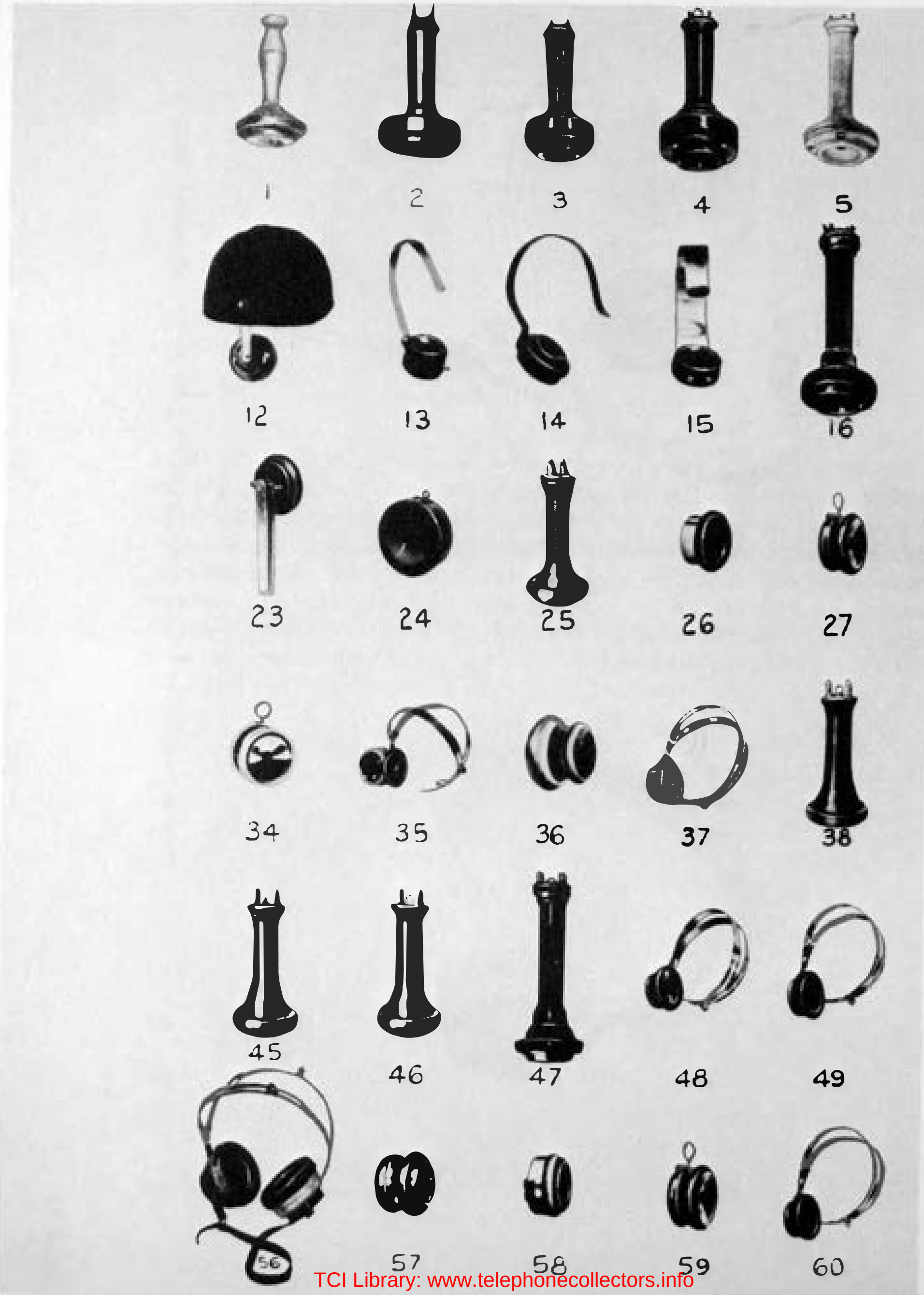
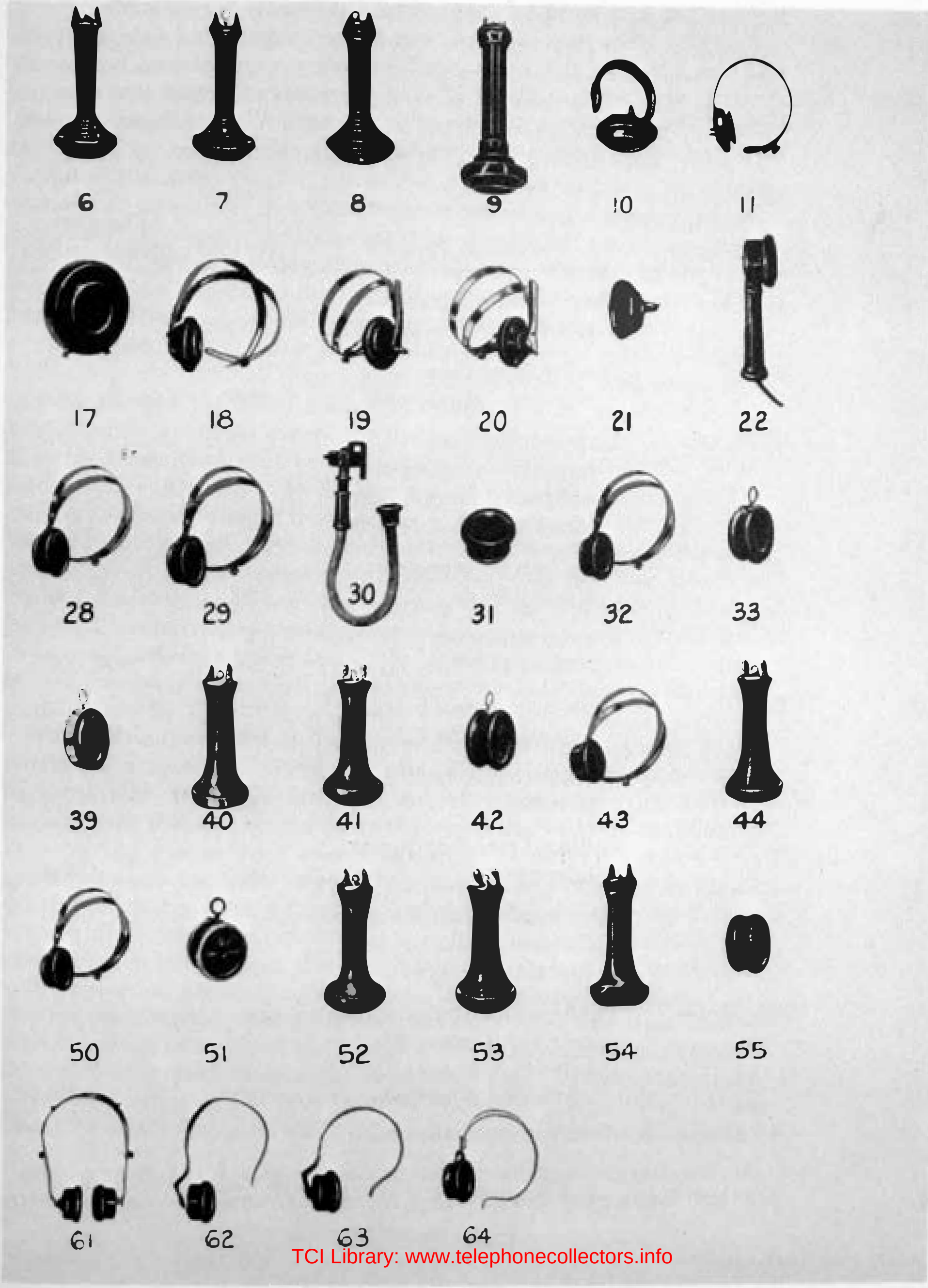


Fig. 3-26. Bipolar operator's receivers: (a) No. 128 receiver (circa 1900); (b) No. 528 receiver (circa 1920).

Fig. 3-27. Telephone receivers which were once standard but were no longer being



manufactured in 1925.



and when the smaller and lighter 528-type receiver was developed, a lightweight wire band became practical as shown in Fig. 3-26b.

By 1925, when the 144-type was the principal hand-held receiver and the 528-type the main head receiver in production, some 60 designs, illustrated in Fig. 3-27, had been manufactured and discontinued. The more important types are identified below, together with their production dates and quantities.

Identification No. in Fig. 3-27	General Type	Date First Produced	Quantity Manufactured
1–4	Butterstamp receiver—wooden case, single pole	1877	3,500
5–8	Small rubber-case receiver—single pole	1877	290
9, 16, 47	Large rubber-case receiver—single-pole compound magnet (three forms all coded 101)	1877	570,000
25	Early bipolar receiver (Code 108)	1890	530
38, 44, 45	Improved bipolar receivers (Codes 110, 111, 112)	1894	860,000
54	High-production bipolar receiver (Code 122)	1902	5,500,000
18–20	Early Richards head telephone	1884	4,000
21	Gilliland cast-iron head telephone	1885	1,700
39, 42	Early type of watch-case receiver—bipolar, self-contained magnet	1894	23,000
49, 62–64	High-production operator receiver (Code 128)—bipolar with self-contained 3-leaf magnet	1900	400,000

2.3 Design Objectives and Performance

At the beginning, the urgent necessity was for a *commercially* practical instrument design. This meant instruments with a usable quality of speech, with enough efficiency to permit transmission over a useful distance, simple enough to be used without special training, stable enough to be used on the user's premises with little adjustment or maintenance, and capable of being manufactured by the thousands at small unit cost. This was a large order. A reasonably good receiver design was achieved in a few years but the transmitters in the 1880s left much to be desired and it is rather surprising that, with the facilities available, the Bell System was to have as many as 200,000 stations in operation in 1890, the year when basic designs became available that were potentially capable of meeting the design objectives listed above. During the next 25 years the designers could concentrate on improving the performance and the manufacture of the bipolar receiver and the solid-back, granular-carbon transmitter.

Throughout this period an overriding consideration was that of increasing the distance over which communication was possible. It will be recalled that until about 1915 there was no source of speech amplification available except for the carbon transmitter. Once hard-drawn copper wire had been adopted for line construction in the middle 1880s, little could be done to reduce line attenuation (except for the limited benefits from increased wire diameter) until the introduction of loading around the turn of the century. Even this development was only a partial solution incapable of meeting the needs of transcontinental telephony. So, during this long period, the main potential for extending communication distance lay in improving the instrument efficiency. Consequently, designers struggled to achieve even small gains since each improvement by a decibel would extend this distance by 30 to 70 miles.[20]

It might be thought that the proper approach to instrument design would have been to concentrate on fidelity of reproduction. It so happens that this approach was not effective in maximizing the transmission distance. Instead, it was found desirable to compromise by accepting a measure of speech distortion in order to achieve large gains in speech loudness (the distance-conquering factor). The reason for this is that a large degree of speech intelligibility is carried by a band of frequencies centered at roughly 800–1,000 hertz. By adjusting the resonances in the mechanical systems of the transmitter and receiver to about this frequency, it is possible to enhance the diaphragm vibration over a narrow frequency band and thus achieve high efficiency in a narrow but highly effective speech band. The basis for optimizing performance (in the sense used here) was not fully understood until much later but by the early 1890s (long before objective measuring gear was available) a frequency response charac-

[20] The distances are based on the use of non-loaded and loaded 165-mil copper wire respectively.

teristic had been worked out empirically which continued to be used effectively for some 30 years. Figure 3-28 shows the characteristics of the transmitter and receiver used from the mid-1890s until the changing conditions of the late 1920s brought about modifications to be discussed later.[21] The Blake transmitter, it will be recalled by referring back to Fig. 3-8, had a highly irregular frequency response characteristic with multiple resonances which gave considerably less satisfactory performance.[22]

While the frequency response characteristics of the instruments had been pretty well established by 1895, when the 229-type transmitter and 122-type receiver were standard, there was still some possibility for increased loudness efficiency and this was accomplished with succeeding designs.

It is difficult to specify the relative efficiency of the instruments used during the first 50 years of telephony since during much of this period there was no means for making objective measurements and in the early days (when some of the changes occurred) there was literally no way to make quantitative measurements of any sort. However, various people have attempted to estimate performance on the basis of contemporary reports, tests on museum samples, and so forth. Figure 3-29 shows a set of such estimates which is probably reasonably accurate. The 229-type transmitter and the 122-type receiver operating in a common-battery circuit on zero loop is taken as the reference condition (except where noted) since, as discussed in Section 5.1.1. of Chapter 4, this condition was so used for many years in transmission rating. Over a period of 40 years the total increase in loudness efficiency (transmitter plus receiver) was about 50 dB of which 40 dB was in the transmitter, indicative of the important part played by the microphone principle.[23]

The availability of electronic amplifiers in 1914 did not immediately remove the pressure on designers for more efficient instruments and,

[21] The reason for the changes at that time and their nature are discussed in Section 6.4.

[22] The modern telephone user, conditioned as he is to the very much improved quality of present telephony, would find the peaked characteristics used in the early twentieth century highly unsatisfactory if not almost unusable. The fact remains that many millions of people learned to use this highly characteristic "telephone speech." Perhaps this is a tribute to the great value of telephony, as compared to other forms of communication then available, or maybe it is merely a reflection of the change in our standards for material comforts and conveniences.

[23] Figure 3-29 is based largely on a set of estimates made by K. S. Johnson in 1920 which he derived from information available at that time. Some later measurements on museum models have shown much poorer performance for early transmitters than indicated in the figure. There is always uncertainty as to the condition of equipment of this type 50 to 70 years after its manufacture and it is likely that Johnson's data are representative of instruments in good operating condition.

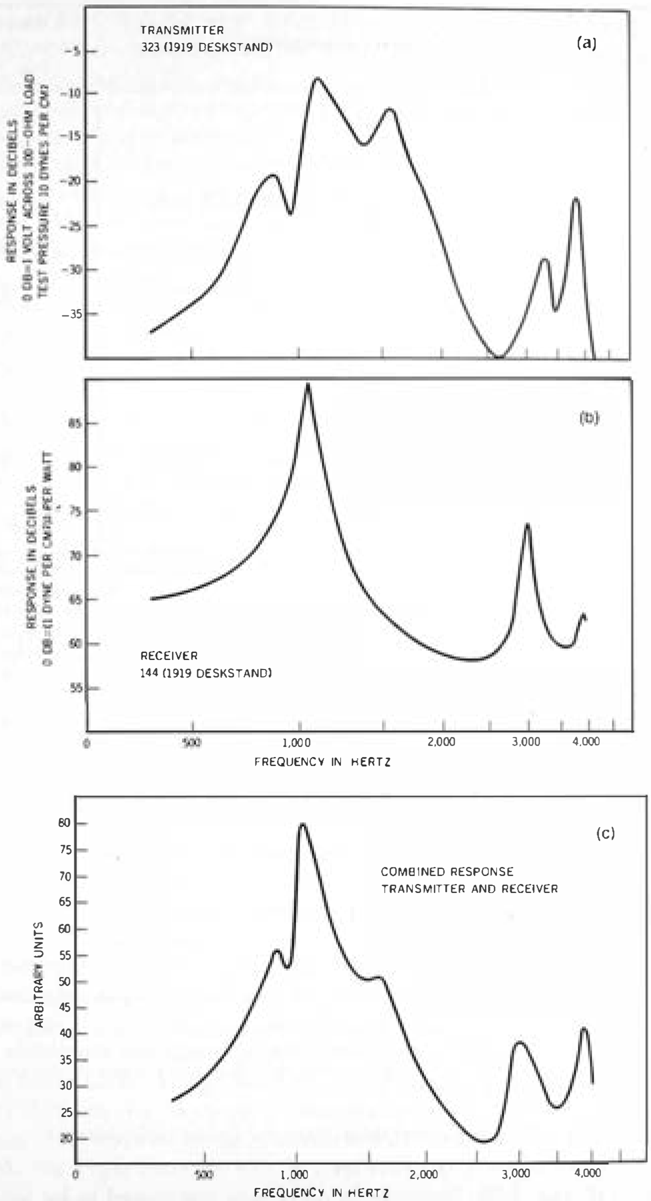


Fig. 3-28. Characteristics of deskstand transmitter and receiver, 1919.

First Introduced	Type of Transmitter	Efficiency — DB vs Reference (No. 229)
1877	Magneto type (butterstamp)	–35
1878	Blake — 1-cell local battery	–13
1886	Long distance (Hunnings) — 2-cell local battery	–4
1888	Blake granular carbon	–6
1890	Solid back — No. 3 Button 2-cell local battery	–2
1895	No. 229 common-battery zero loop — reference	0
1913	No. 329 common-battery zero loop	0
1917	No. 323 common-battery zero loop	+3
1919	No. 323 common battery (loops over 150 ohms — efficiency relative to No.229 on same loop)	+5
	Total increase in efficiency over 1877	+40

First Introduced	Type of Receiver	Efficiency — DB vs Reference (No. 122)
1877	Butterstamp receiver	–9
1878	No. 101 — unipolar hand receiver	–4
1890	No. 108, etc. — early bipolar hand receivers	–2
1895	No. 111, etc. — improved bipolar receivers	–1
1902	No. 122 — reference — high-production bipolar receiver	0
1912	No. 144 — improved high-production bipolar receiver	+2
	Total increase in efficiency over 1877	+11

Note: The efficiencies are specified relative to the No. 229 transmitter and No. 122 receiver since these instruments were used for many years as reference conditions for specifying station and loop transmitting losses as discussed in Section 5.1.1 of Chapter 4.

Fig. 3-29. Estimated transmitter and receiver efficiencies.

as shown in Fig. 3-29, transmitter efficiency continued to be increased until 1919 when the 337-type transmitter was introduced. The reasons are rather obvious. The first amplifiers were used to extend telephone service to transcontinental distances, a goal unachievable otherwise.

The next step was to use amplifiers to improve the quality and reliability of wire transmission by reducing the use of loading on open wire and increasing the use of cable circuits. As amplification became more reliable and cheaper (particularly with carrier transmission), its use became more widespread as a means for reducing the loss of telephone lines. With increasing use of inexpensive line amplification, it became possible to change the design objectives for transmitters and receivers. Up to this time the emphasis had been to maximize transmission distance by increasing loudness efficiency by all possible means. With loudness controllable by line amplification, the design objectives could include improved fidelity and ease of use. Concurrently, studies of speech and hearing were producing quantitative information useful in attaining these objectives. More will be said about this later but for the moment it is sufficient to note that the designers could plan to work along these lines as soon as electronic amplification was an accomplished fact, but it was not until the 1930s that amplification was used widely enough to permit the general use of the higher-fidelity but lower-loudness instruments.

III. SPEECH CIRCUITS

The speech circuit in the telephone station provides means for connecting the transmitter and receiver to the telephone line and to each other.

3.1 Circuit Configuration

With magneto converters the telephone instrument (used singly) was connected directly to the line wires as shown in (a) of Fig. 3-30.[24] A series battery, as shown, was required when the converters used an electromagnet but could be omitted with permanent magnet devices. When separate magneto devices were used for transmitting and receiving, they were merely connected in series as shown in (b) of the figure. This added convenience, incidentally, reduced the transmission distance capability quite seriously since the added converter absorbed power and a loss of 3 dB occurred at each end of the line.

When the carbon microphone was introduced as a transmitter, the same series arrangement (with battery) could have been used but would have been unsatisfactory for several reasons. First, direct current through the receiver, if incorrectly poled, would have reduced the receiver efficiency by opposing the flux of the permanent magnet.[25] Second, the high resistance of the receiver as well as the line would

[24] For convenience, Fig. 3-30 shows metallic circuits throughout but it will be appreciated that initially a ground return replaced one of the line wires.

[25] Since the receivers were fully magnetized in manufacture, a current that aided the magnet produced little effect but a current that opposed the magnet degraded the receiver's efficiency.

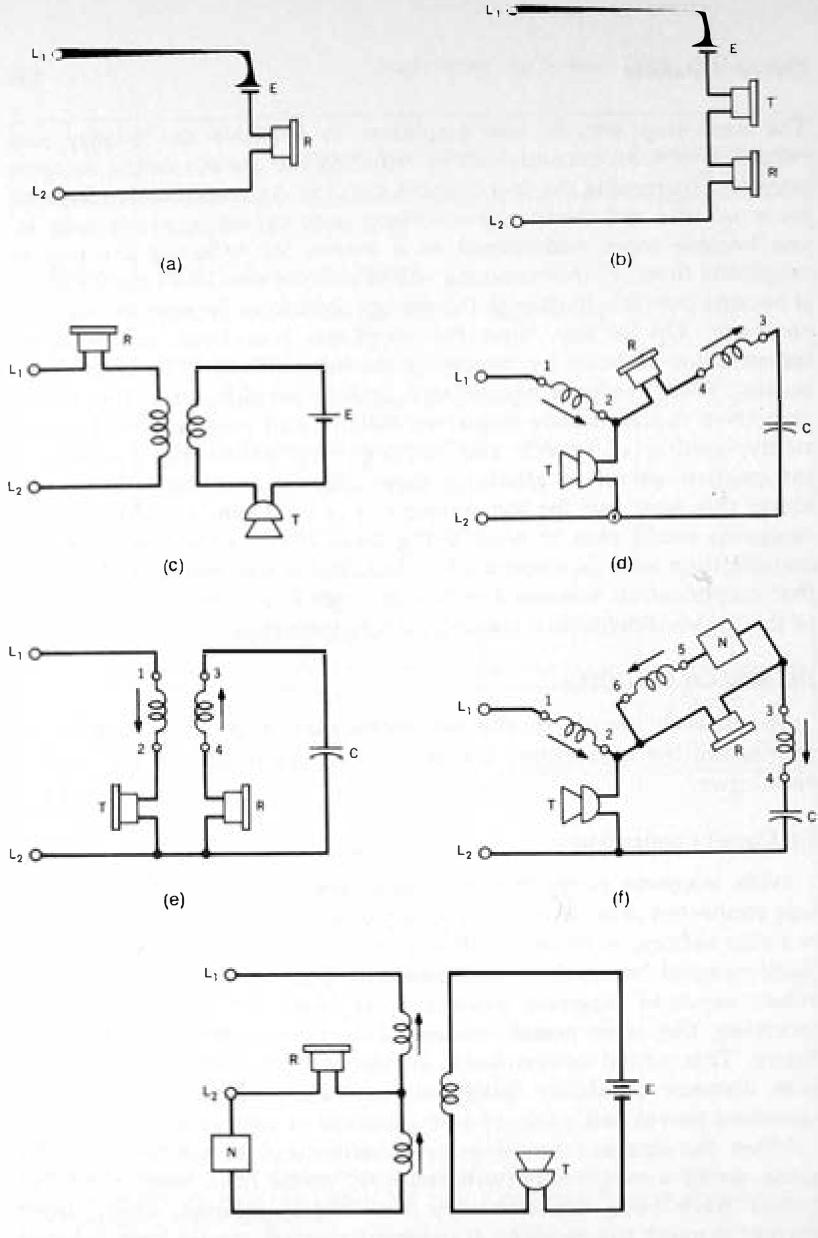


Fig. 3-30. Telephone speech circuits: (a) First telephone station using a single instrument. (b) Telephone station using two identical instruments. (c) Local-battery sidetone circuit. (d) Common-battery sidetone circuit. (e) Sidetone-reduction circuit. (f) Common-battery anti-sidetone circuit. (g) Local-battery anti-sidetone circuit. (In these diagrams, all induction-coil windings are coupled, i.e., wound on a common core. N is a sidetone balancing network which, with the associated circuitry, is in essence a Wheatstone bridge. Ideally, there is no current in the receiver when the user is transmitting.)

have been in the direct-current supply circuit and would have greatly reduced the current (and transmitter efficiency) unless high-voltage batteries were used. Third, and in many ways more importantly, the initial carbon transmitter had a very low resistance, of the order of 4 to 5 ohms. A low-impedance transmitter, in series with a high-impedance receiver working into a matching line, would be very inefficient. For optimum transfer of energy the transmitter and receiver should be of equal impedance and the two together should match line impedance, a matter of hundreds of ohms. This problem was solved by the use of a transformer and the circuit is shown in (c) of Fig. 3-30. The transformer was initially called a repeating coil since it transferred or repeated energy by induction from one coil, the primary, to another, the secondary. Later the term "repeating coil" was used solely for describing other telephone applications and the transformer used in the station set came to be called an induction coil. The latter designation will be used here. Both Berliner and Edison received patents in 1878 on transmitter circuitry in which this device was used. Its functioning is very simple and effective. As shown in the figure, the transmitter was connected in series with the primary of the coil and a local battery while the secondary was in series with the receiver, the combination being connected through the line to the distant station. The primary had relatively few turns and a low direct-current resistance. Thus a single-cell battery could be used to provide rather high direct current to the transmitter (200 to 250 milliamperes with a 4- to 5-ohm transmitter). The secondary had a large number of turns (some five or ten times the number of the primary) increasing the ac voltage from the transmitter accordingly and stepping up the transmitter impedance by the square of the turns ratio. Thus the transmitter impedance could be made effectively as large as that of the receiver and an efficient circuit, matching line impedance, achieved.

The common-battery system using eight or more storage cells was introduced in 1894. In this system, a single battery of large current-capacity was located at the central office and used to supply operating current to all the stations served by the office. With this system a new station circuit was required which would provide impedance matching and a direct-current path for the transmitter current and at the same time block the direct current from the receiver. The arrangement illustrated in (d) of Fig. 3-30 was used.

In all the station circuits which have been described so far, the receiver was driven by its associated transmitter; each customer, while talking, hearing his own voice in his receiver. He also could hear any background noise or other sounds picked up by the transmitter and fed to the associated receiver. This effect, called sidetone, was a disadvantage because the customer tended to speak softly when he heard his speech at high volume in his own receiver and background noise tended to mask incoming speech. Consequently, a large number of arrangements were developed to reduce sidetone. The simplest scheme was a rearrangement of the wiring of the normal common-battery set, as shown in (e) of Fig. 3-30, to give what was known as the STR (sidetone reduction) connection. This circuit had been authorized in 1903 for optional use when high sidetone was causing problems and was standardized about 1916 or 1918 for very short loops of 90 ohms or less resistance since the higher-efficiency instruments introduced at that time (323-type transmitters, 144-type receivers, and new induction coils) led to unpleasantly high sidetone on the short loops. This connection reduced sidetone by about 9 dB but was accompanied by a reduction in transmitting efficiency of 5 to 6 dB and an increase in receiving efficiency of 1 to 2 dB. This circuit was not a true AST (anti-sidetone) circuit in its commonly used telephone sense which K. S. Johnson has defined as "one in which there is in the ideal case, no power dissipated in the receiver when an EMF is generated in the transmitter."[26] The circuit most commonly used for common-battery subscriber station sets is shown in (f) of Fig. 3-30 and was introduced about 1930. The reduction in sidetone obtained varied considerably with the length and type of subscriber loop but was generally in the range of 10 to 15 dB. There was also an accompanying loss in transmitting and receiving efficiency totaling about 2 dB.

A local-battery version of the AST set was also introduced in the early thirties. The circuit is shown in (g) of Fig. 3-30. Local-battery sets were used on a variety of loops including non-loaded and loaded cable and open-wire extensions on cable. Consequently, the impedances faced by the set covered a wide range and the adjustable network (N) was included so that it could be selected to give minimum sidetone. While the theoretical benefits from such an arrangement were large, it proved difficult to administer in practice.

The AST set was to play an important part in station design during the 1920s and has contributed significantly to the big improvement in station transmission which has occurred in the last 40 years. It is, therefore, worth reviewing briefly the origin of this principle even though its application to subscriber stations occurred mostly beyond the period covered by this chapter.

The problems introduced by high sidetone had been appreciated for some time and techniques for its reduction dated from the mid-nineties, C. E. Scribner in Western Electric's Engineering Department being granted three anti-sidetone patents in the 1894–96 period. Because

[26] When the reader comes to Section 4.2.4 of Chapter 4, he will note that this is essentially the definition of a hybrid coil translated into the terms of the subscriber station set.

operators were subjected to long hours of high-level speech and worked under high levels of noise, the anti-sidetone principle was applied to their speech circuits as early as 1900 and soon became a standard feature of operator telephone sets. However, the circuit in wide use for common-battery operator circuits in the early 1900s was too elaborate and expensive for economical application to subscriber use.

The anti-sidetone subscriber circuit using a single induction coil was invented by G. A. Campbell, Research Engineer at AT&TCo. He showed in an article published in 1920 that the number of possible forms of anti-sidetone circuits is extremely large, depending on the assumptions made regarding the configuration of the elements, number of transformers (induction coils), etc. In fact, he indicated that the possible number was almost a half-million with a two-transformer circuit. Only a few of the simpler, more economical circuits were patented, nine patents being issued to Campbell in 1918.[27] Collectively these patents covered 37 circuits. It was deemed unnecessary to patent more of the circuits since it was practically impossible to determine which of the large number of possible circuits might prove important in the future. Instead it was decided to obtain protection by disclosing all possible variants by means of the 1920 Campbell paper. This not only made all these arrangements freely available to anyone who wished to use them but also prevented anyone from obtaining a basic patent on any or all and thus avoided endless future litigation.

Thus the technical situation by 1920 was such that the anti-sidetone set could be introduced at any time. This was not done for another ten years in part because practical designs were several dB less efficient than the No. 46 coil then in use and in part because of extra cost arising from the need for an extra winding on the induction coil and an extra conductor in the telephone cord. Towards the middle and late 1920s, tests made on working circuits showed that the circuit loss would be more than overcome by the benefits of sidetone reduction, i.e., by increased talker volume and reduced room noise transmitted to the ear. With this factual data in hand, the way was open to the new anti-sidetone designs which were introduced in the form of the No. 101 induction coil in 1930.[28] This coil was based on a new design

[27] G. A. Campbell; U. S. Patent Nos. 1,254,116–118 and 1,254,471–476; filed September 9, 1916 and March 15, May 18, and August 18, 1917; all issued January 22, 1918.

[28] It may well be surprising to the reader that the anti-sidetone set did not come into use for the subscriber until 25 to 30 years after it was used in operator circuits. Part of the reason was, as noted, that the circuit arrangements used by operators were large and costly and it was some 20 years before simplified ways were disclosed by Campbell. Perhaps an even greater reason was the doubt that the benefits to the subscriber would be very significant. Unlike the operator, the subscriber could always partially remove the receiver from his ear if talking sidetone was unpleasantly high and consequently many believed that the talking volume increase from anti-sidetone

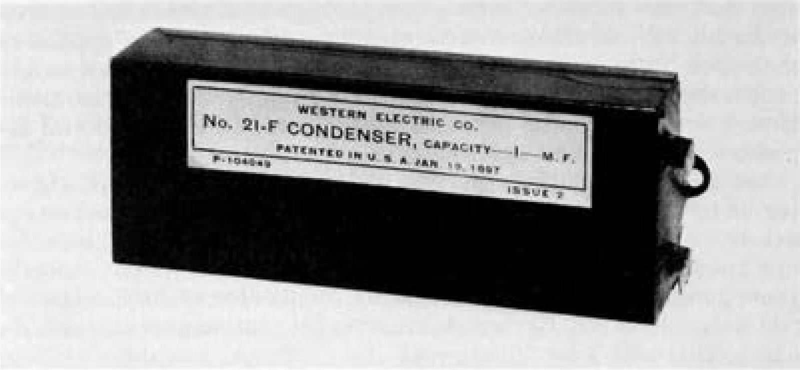


Fig. 3-31. No. 21 blocking capacitor.

adapted to more mechanized production and ultimately proved less expensive than the earlier two-winding coils. Provisions were also made for converting the older, sidetone coils to the AST type when they were returned to the shop for overhaul but the process was not used extensively.

3.2 Apparatus Components

The principal apparatus components of the subscriber circuit were the capacitor used to block direct current from the receiver and the induction coil for impedance matching of the transmitter. When not served on a common-battery basis, a local battery was also an essential part of the apparatus on the customer's premises.

Until about 1916 the capacitor was usually 2 microfarads in size, but it was changed to 1 microfarad in the war years as a measure for saving material. It was continued at this size after the war since the effect of the change on transmission proved minor. These capacitors were manufactured with metal foil plates separated by paper. The assembly was wrapped into a flattened roll and potted in a molten wax compound in japanned "tin" cases (Fig. 3-31).

The induction coil, until about 1930, was a simple solenoid with an open-ended core made initially of a bundle of parallel soft-iron wires.

---

would be negligible. Also the lower room noise levels at most customer locations led to the belief that the effects of noise reduction would be small. It required quantitative tests on working telephone circuits (mostly with AT&TCo and Bell Laboratories employees) to remove the question of anti-sidetone benefits from the realm of speculation and demonstrate factually the significant value.

Since these coils carried rather large direct currents (up to 0.25 ampere) the open magnetic circuit was used deliberately not only because it was simple to manufacture but to avoid the magnetic saturation that would occur with a closed-core structure. In 1916 a bundle of silicon-steel strips began to be used in place of the soft-iron wire giving about 0.5-dB transmission improvement with a few cents increase in cost. Originally the coils were wound individually on a layer of insulating paper or cambric surrounding the core but later they were machine wound, several at a time, on a long tube which was later cut up into individual coil forms. The core ends were of wood and served as terminal and mounting blocks as well as to hold the coils in place on the core. A typical coil, the No. 20, is shown in Fig. 3-32a. This was a common-battery coil introduced in 1897 and continued in manufacture until about 1918 when it was replaced by the somewhat cheaper and more efficient No. 46 coil. The predecessor of the No. 20 was the more expensive No. 18 that was used in the early common-battery installations beginning about 1894. Both the No. 18 and the No. 46 were structurally similar to the No. 20.

As we have noted, a new induction coil design was introduced in 1930 when the AST set was standardized. This coil, coded 101 and illustrated in Fig. 3-32b, presents a marked contrast to all the earlier types not only in size but also in configuration. The core was made up of L-shaped silicon-steel laminations which were assembled into a closed structure with small air gaps. The earlier Nos. 20 and 46 coils when converted to the AST type were coded 120 and 146 respectively. Their appearance, except for an added terminal, was essentially unchanged. Production was small because of the low cost of the 101-type.

The most commonly used local-battery induction coil was the No. 13 developed in 1893 as a replacement for the more costly No. 6 and, except for a change in core material in 1916, continued in manufacture without major modification in electrical characteristics through the 1920s. Some of the many types of induction coils that have been manufactured over the years are illustrated in Fig. 3-33. Some of the very large coils shown here were those used for the operator anti-sidetone circuits and their size indicates clearly why they were not introduced in subscriber station sets. Numbers 10, 23, and 24 are typical of such induction coils.

Until 1894, batteries were required on the customer's premises at all stations and their use was continued at a number of stations, particularly in rural areas, long thereafter. Up to 1900, batteries of the "wet" type were most commonly used. These came in several varieties with differing characteristics but all had in common the use of a more

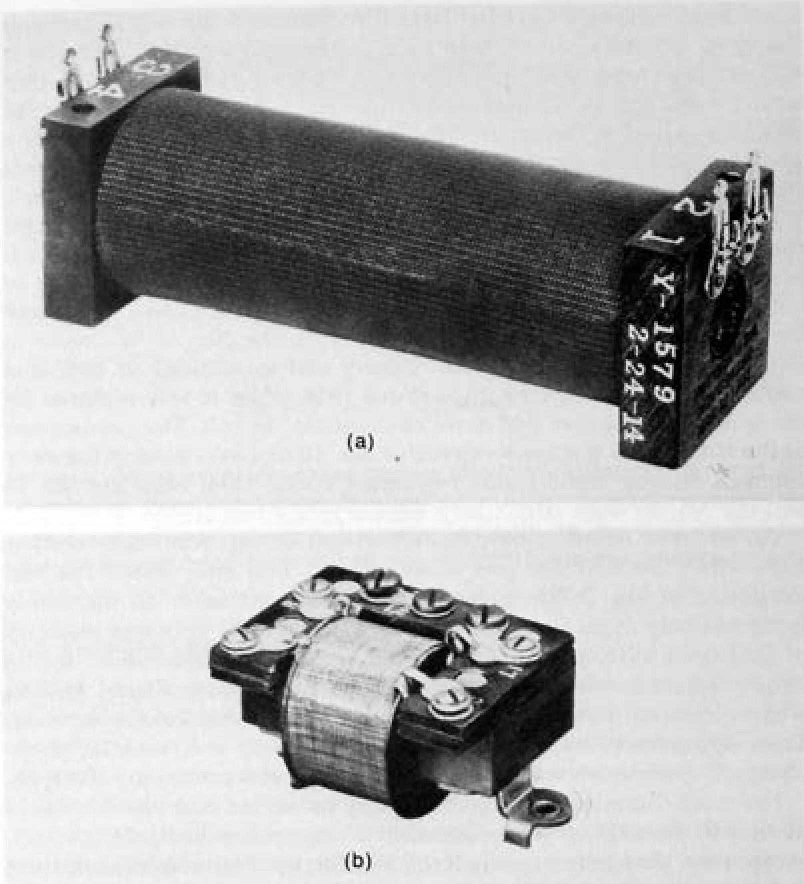


Fig. 3-32. (a) No. 20 induction coil. (b) No. 101-A induction coil.

or less corrosive liquid electrolyte that no one, if there were a choice, would care to have in his home.[29] Various types were used. In the early days a favorite was the LeClanche, a zinc-carbon-sal ammoniac cell (with

[29] The modern reader, accustomed to the use of present-day "leakproof" dry cells, will read with astonishment the following instructions for setting up a Fuller cell, as printed in an 1896 Western Electric catalog:

"Make a paste by mixing up pulverized bichromate of potash with strong sulphuric acid in about equal parts by weight. Put about ten ounces of this paste into the outside jar, pour over it two or three ounces of sulphuric acid and fill up with water. Into the porous cell pour a teaspoonful of mercury, put the zinc in place and fill up with water. The zinc should be lifted out occasionally and the sulphate washed off. Keep a supply of mercury in the porous cell, so as to have the zinc always well amalgamated."

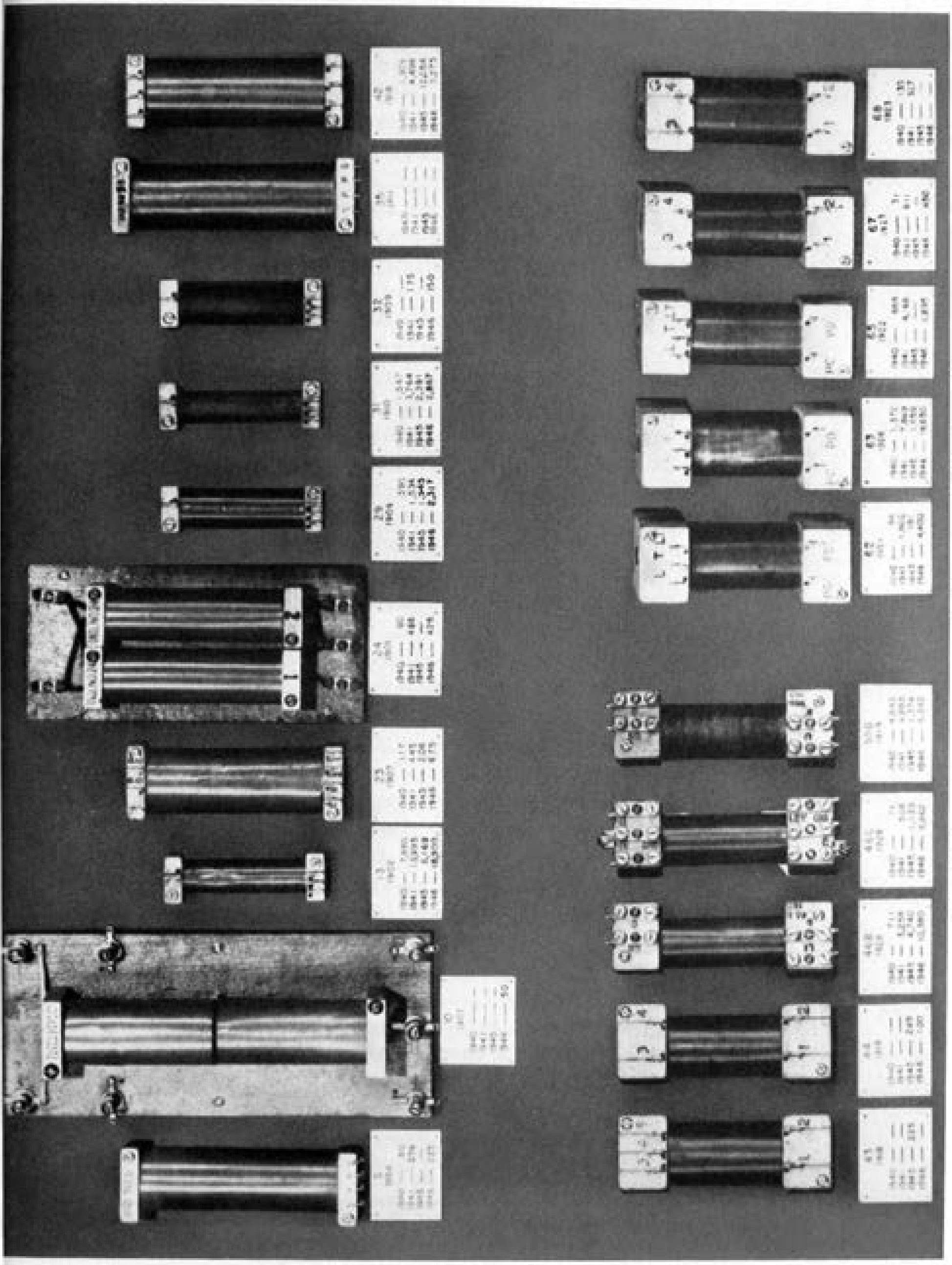


Fig. 3-33. Panel display tracing development of induction coils.

manganese dioxide depolarizer[30]) which was chemically the equivalent of the present-day dry cell. The electrolyte was somewhat less corrosive than used in some cells but had a tendency to "creep" over the edge of the jar and the cell had a lower voltage and polarized more readily than other types used for telephony. During the late 1880s and early 1890s the Fuller cell was the type most commonly used by the Bell Company. It used carbon and zinc electrodes with sulphuric acid electrolyte and potassium bichromate as a depolarizer around the carbon electrode. The zinc electrode was amalgamated to reduce local action and the production of zinc sulphate. The zinc was placed in a porous cup filled with water (or a salt solution) which separated this portion of the cell from the bichromate solution. The Fuller cell was favored because of its high electromotive force (about 2.1 volts), low internal resistance, and long life. Under ordinary circumstances with a low-resistance (5-ohm) transmitter, one cell was adequate but as many as three were used in some of the so-called "long distance" stations.

Commercial production of dry cells began in this country around 1890 and they were listed in Western Electric catalogs as early as 1891. By 1896 cells similar in size to the present No. 6 cell were offered for sale at a price of 80 cents. The dry cell is essentially a LeClanche type with a zinc electrode used as the container, a central carbon electrode, and the space between filled with a porous material saturated with a sal ammoniac solution. Manganese dioxide is used around the carbon as a depolarizer, preventing the formation of hydrogen gas at the electrode which would increase the internal resistance of the cell.

While the dry cell is obviously preferable to the wet type for home use, it was some years before the wet cells were fully replaced, particularly for heavy-duty service. The early dry cells were not completely free from corrosion problems since the zinc container would develop holes through which the sal ammoniac exuded. The life was short and varied greatly among the various brands manufactured. Some of the poorer brands were unreliable and gave the dry cell a bad reputation among telephone technicians. In 1901, typical cells lost about 35 percent of their capacity after six months of shelf life and new cells had a useful life of only about five to six months under light, intermittent service. Kempster Miller, in a book published in 1905,[31] pointed out that dry cells were at last being widely used but some cells still had a life of only a half a year even though better brands were becoming available with a life approximating two years. It is obvious

[30] In use, hydrogen formed around the positively poled element in all primary cells and quickly reduced the cell's capability. This reduction, known as polarization, was due both to chemical action and effective reduction in the working surface of the plate. It was compensated for by the use of a chemical oxidizing agent known as a depolarizer.
[31] *American Telephone Practice*, 4th ed., published by McGraw-Hill Book Co.

that even after the advent of the dry cell, the incentive to eliminate local-battery service was very great.

IV. SIGNALING ARRANGEMENTS

The electroacoustic converters, together with the speech circuit, supplied the means for conversing between two people but additional signaling equipment was necessary to initiate a conversation.

Before 1878, there was no central-office switching and telephony was confined to private lines with a telephone at each end. The signaling problem was simple, requiring only some means to alert the called subscriber to the fact that a conversation was desired. Originally the caller simply shouted into the mouthpiece using words with long, loud vowel sounds such as "ahoy" and "hello." This was not very satisfactory and various mechanical arrangments were developed, which we shall describe. Ultimately, schemes using special bells, or "ringers," were devised and the alerting process came to be known as "ringing the customer."

The need for establishing communication between any two of a large number of users was inherent in Bell's "Grand System" and means for "switching" the line from a caller's station to that of the desired called station were developed at an early date. Switching systems development is covered in Chapter 6 and for the present it will be sufficient to point out that basically each telephone station was connected by a transmission line or "loop" to a central office at which point it was "switched" to the called station either directly, if the two stations were served by the same central office, or via "trunk" lines to the serving central office. The connections at the central office were originally made manually by operators but later mechanized equipment was used for some forms of switching.

When switching was introduced, the ringing process became the responsibility of the operator (or the machine switching equipment) but the same ringing mechanism was used at the customer's premises since he still needed a ringer to announce an incoming call and a calling mechanism to indicate the need for an operator. Initially, the mechanism developed for ringing a customer was used to alert the operator but with the introduction of common-battery offices it became possible to call the operator merely by lifting the customer's receiver from its standby position.

With manual switching, information on the desired called customer was passed verbally to the operator who established the necessary connection through the switching system.[32] With the introduction of automatic switching in the 1890s it became necessary to add another

[32] In long connections the call often passed through a number of offices in tandem, each receiving verbal information from the initiating customer or preceding operator.

signaling arrangement to the telephone station, namely, means for directing and controlling the switching mechanism, as discussed further in Sections 4.2 and 4.3.

Thus, by the middle nineties the station set could include arrangements for performing three signaling functions: (*i*) ringing the called customer, (*ii*) alerting the operator (or the automatic switching mechanism), and (*iii*) controlling the switching mechanism (if it was automatic). These various arrangements will be described shortly but before doing so we shall mention a very simple invention that was to play an important part in the evolution of many signaling arrangements. This was the device known as the switchhook, invented by H. L. Roosevelt,[33] a prominent organ builder who was also a founder of the first telephone company in New York City. With certain ringing systems and, more importantly, with the introduction of the local-battery carbon transmitter, the station circuitry was different during the talking and the standby condition. For example, during standby, the transmitter battery was not needed and the circuit was opened to avoid unnecessary battery drain. Similarly, it was sometimes desirable to remove the ringer from the line during talking and reconnect it on standby. These functions were originally carried out by switches manually operated by the customer. Naturally, the operation was all too often overlooked leaving the set in the wrong state with unfortunate consequences. Roosevelt solved the problem by providing a hook for holding the receiver when not in use. The weight of the receiver operated the switch to the standby condition and a restoring spring transferred the switch to the talking condition when the receiver was lifted for use. An early station set with this feature is shown in Fig. 3-34. This obvious, but fundamental, idea of an automatic switch to transfer between the talking and standby conditions was ultimately used for a number of signaling functions such as control of talking battery, alerting the operator, tripping ringing, terminating the call, etc.[34]

4.1 Calling the User

Since oral calling was unsatisfactory with the very inefficient instruments available, various mechanical arrangements were soon tried. One expedient was to tap the diaphragm of the transmitter. This

[33] H. L. Roosevelt; U. S. Patent No. 215,837; filed October 3, 1877; issued May 27, 1879.

[34] Edwin T. Holmes, in his autobiography, *A Wonderful Fifty Years*, published in 1917, states that he invented the switchhook in 1877 to eliminate the bridging loss of the many telephone instruments he was installing on party lines. He shows a picture of a telephone station using such a device on page 72 of his book. He did not apply for a patent.

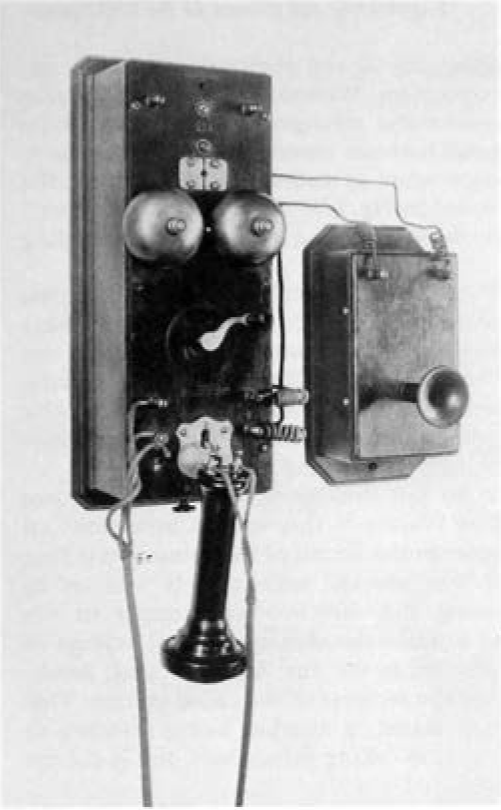


Fig. 3-34. Early station set with receiver switchhook.

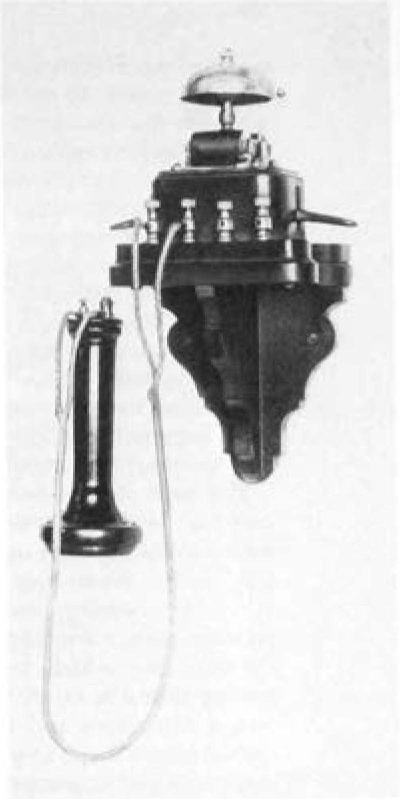
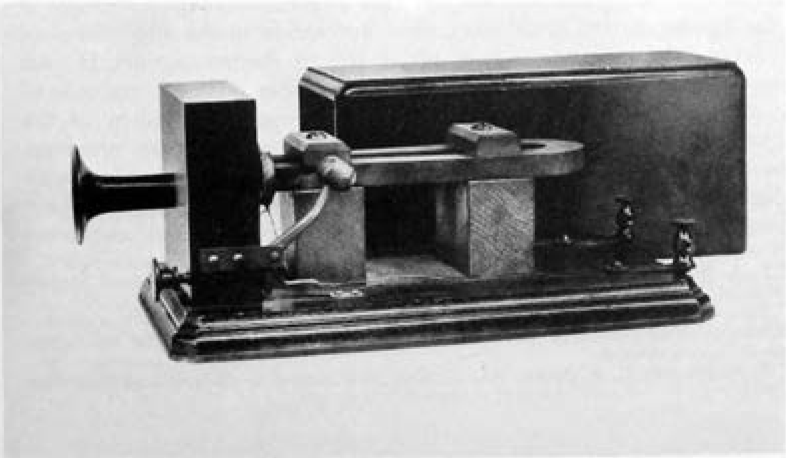


Fig. 3-36. Electric "tap bell."

Fig. 3-35. Watson's "thumper."



was not too effective as a calling signal and obviously could damage the instrument. To avoid this problem, Watson, in June of 1877, came up with the first built-in mechanical arrangement for his box-type telephone. It consisted of a small hammer mounted in the box so as to strike the edge of the diaphragm when operated by a knob outside the box. This arrangement, illustrated in Fig. 3-35 and called a "thumper," was purely a stop-gap measure, the tapping sound produced being weak and not distinctive.

This was followed by electric "tap bells" (Fig. 3-36), electric vibrating bells, and, in some cases, telegraph sounders. All of these were operated by power from batteries having limited voltage and power capacity which restricted the practical length of the serving lines. Also, these devices, particularly the interrupters on the vibrating bells, required frequent servicing. So the search continued for a better way to attract the attention of the person being called.

The next calling device to be put into service was a mechanism called a "buzzer" developed by Watson.[35] This was an induction coil with a vibrating-reed interrupter in the circuit of the primary winding (Fig. 3-37). When this reed was plucked manually, it was set in vibration, opening and closing the direct-current supply to the primary circuit and inducing a relatively high alternating voltage in the secondary which, when applied to the line, caused a loud, harsh, grating sound to be emitted by the receiver of the called station. This was a distinctive and effective signal, a number being installed in central offices for use by operators in calling subscribers, but it did not meet with public acceptance.[36]

The long-range solution to the ringer problem came with the invention of the polarized call bell, pictured in Fig. 3-38.[37] Its operation depended on the use of two magnetic forces: a steady force derived from a permanent magnet (the horseshoe-shaped element B in the figure) and an alternating force derived from the application of an alternating current to two series-aiding electromagnets, H. An annealed-iron armature, C, pivoted at the center, formed one pole of the permanent magnet and was located opposite the cores of the electromagnets which formed the opposite poles. When the armature moved on its pivot, a hammer, G, attached to it struck suitably placed gongs. In his patent application, Watson stated, "The object of my invention is to secure greater simplicity of construction and more

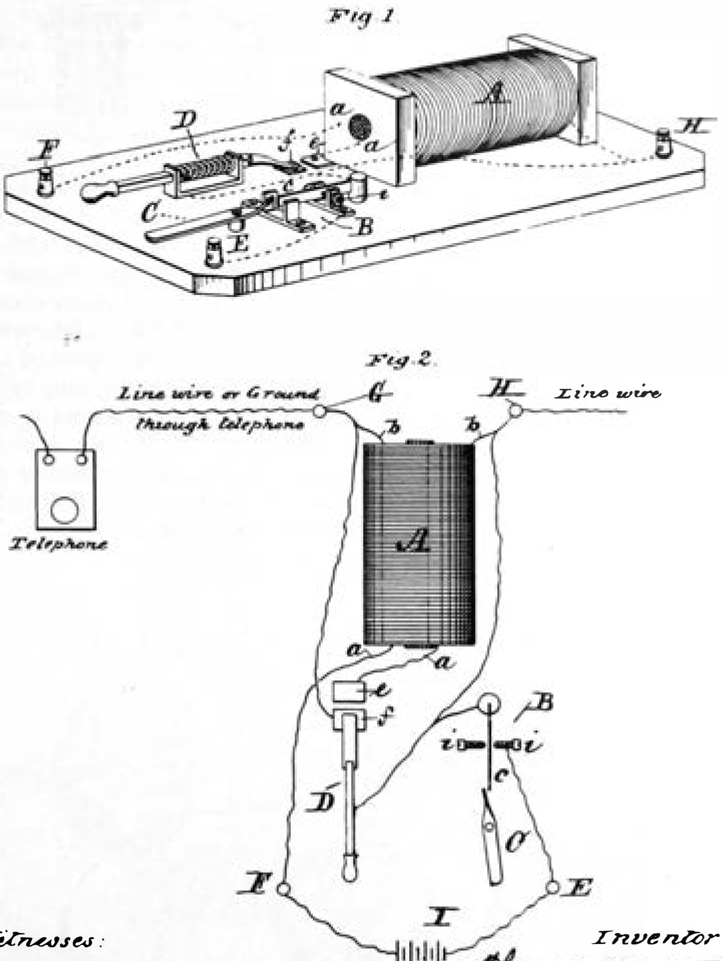
[35] T. A. Watson; U. S. Patent No. 199,007; filed December 5, 1877; issued January 8, 1878.

[36] The manager of the New Haven, Connecticut, office where this buzzer was first installed was George W. Coy, and the signal was promptly given the descriptive name of "Coy's chicken."

[37] T. A. Watson; U. S. Patent No. 210,886; filed August 1, 1878; issued December 17, 1878.

T. A. WATSON.
Telephone.

No. 199,007. Patented Jan. 8, 1878.



Witnesses:
E.E. Masson.
E. Asick

Inventor:
Thomas A. Watson by
A Pollok his attorney

Fig. 3-37. Watson's "buzzer."

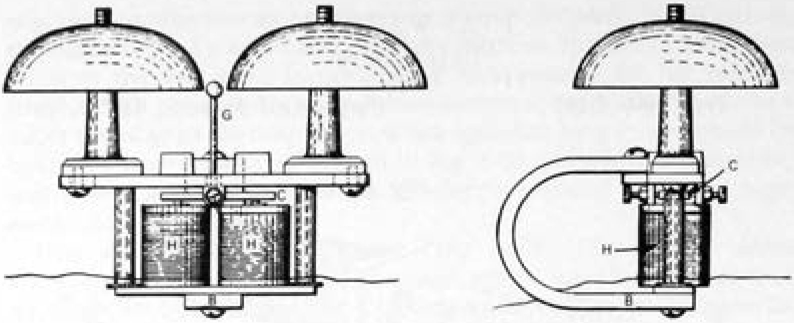


Fig. 3-38. Front and end views of Watson's polarized call bell or "ringer."

powerful operation." He obtained this "powerful operation" as follows: "When a current passes in one direction through the coils of the electromagnet, the armature is attracted at one end and repelled at the other. When a current passes in the opposite direction, this action is reversed." Watson's object in this device was so well achieved that the basic principle, a pivoted armature following an alternating magnetic field superimposed on a steady magnetic field is still in use; Fig. 3-39 shows the design that was standard in 1936. In the 1970s the same principles are being used but improved magnetic materials and better knowledge of acoustic theory have made possible both smaller devices and more pleasing and effective ringing sounds.[38]

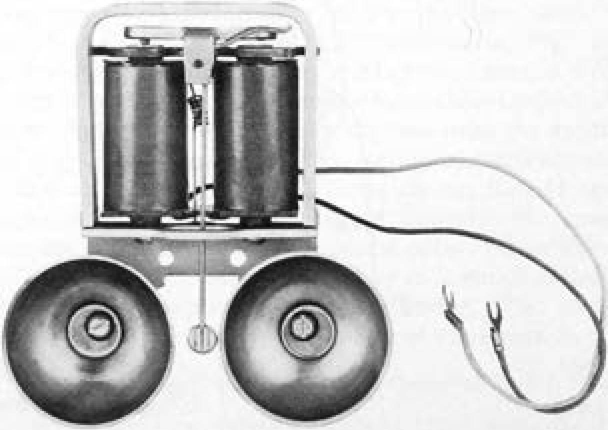
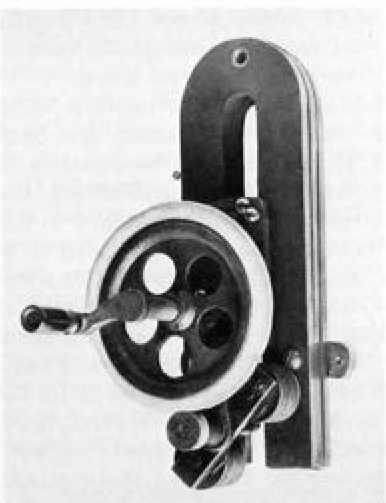


Fig. 3-39. Polarized ringer used in the 1930s.

[38] The balanced armature principle was fundamental to other communication devices, being used later in such diverse applications as coin collectors and loudspeakers.

To provide the necessary alternating field, Watson, after consulting Davis' *Manual of Magnetism*, published in 1847, constructed a small alternating-current generator (usually called a "magneto") which was mounted in the customer's telephone and operated with a hand crank.[39] Watson's magneto (Fig. 3-40) consisted of a substantial U-shaped permanent magnet and a bar armature, with a coil of wire at each end. The center of the armature was attached to a shaft driven by the crank through a pair of friction wheels. As the armature was rotated in the field of the permanent magnet, an alternating voltage was induced in the coils, the frequency being determined by the speed of rotation and the voltage by the speed, strength of the magnet, and the number of turns of wire in the coils.

This generator and ringer system was so successful that several manufacturers were licensed to manufacture the components. At a meeting of the National Telephone Exchange Association in 1881, it was stated that the magneto system was rapidly superseding all other forms of signaling. It was at about this time that Siemens developed the so-called "H-shaped" armature. A Bell System magneto employing this type of armature, a gear drive, and an automatic switch to connect the generator to the circuit only when cranked is shown in Fig. 3-41. A more powerful machine using five bar-magnets was also developed for



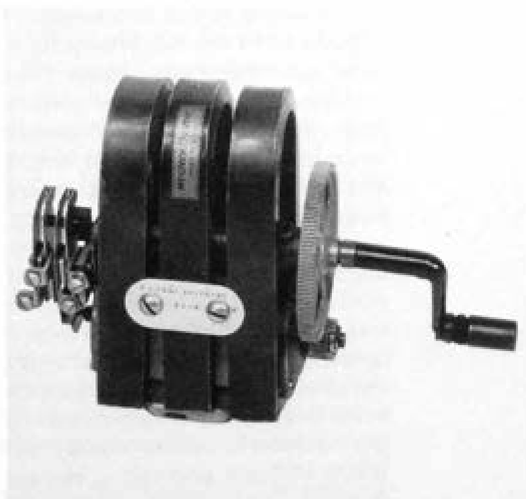


Fig. 3-40. Hand-operated alternating-current generator or "magneto," designed by Watson. Fig. 3-41. Later type of magneto with automatic cutout (at left).

[39] T. A. Watson; U. S. Patent No. 202,495; filed October 11, 1877; issued April 16, 1878.

use on very long circuits. This type of device was adopted by the Bell System in 1881 and used in local-battery subscriber stations, with minor modifications, for over 50 years.

As switching was introduced, a ringing generator at the central office became necessary. The hand-cranked machine required considerable operator effort and other arrangements, described in Section VII of Chapter 6, were soon substituted.

When the polarized ringing system was introduced, both the ringer and the generator were wired in series with the line, and were of relatively low resistance, about 80 ohms. It was soon found that the resulting loss in transmission was intolerable and means to reduce the loss were introduced. One of these was a short circuit applied across the coil of the generator while at rest; a mechanical linkage removed the short circuit when the crank was turned. This solved the problem related to the generator but not that of the call bell.

The ringer difficulty was resolved in 1890 by John J. Carty, at that time with the Metropolitan (New York) Telephone and Telegraph Company. He developed a high-impedance, 1,000-ohm ringer and connected it across the line, bridging the speech circuit.[40] The impedance of this ringer was very high at voice frequencies and thus resulted in little or no transmission loss. The associated generator was arranged to be bridged on the line during its operation, but to be disconnected, by a mechanical linkage, when at rest. The higher impedance of the ringer required a relatively high voltage, 75 volts or more, and such high-voltage ringing has been used ever since.

When common-battery switchhook signaling was introduced in 1897, the hand-cranked generator became obsolescent for both central-office and customer use, but its influence continued since it, and its associated ringer, were instrumental in establishing the frequency and voltage of ringing current that is still standard in the Bell System. The speed of the hand-cranked generator was such as to produce an average rate of about 17 hertz in normal usage, and the natural period of the ringer armature and clapper had been adjusted to correspond to this speed. When central-office ringing machines came into use, they were matched to the ringers already in service and were designed to a standard frequency of 16⅔ hertz supplied at 75 to 110 volts. With two-pole generators this corresponded to 1,000 revolutions per minute, a speed readily obtained with direct-current motors. When 60-hertz alternating current became the standard of the power industry, alternating-current motors were adopted for driving the ringing machines but belt drives were required since a speed of 1,000 rpm could not be obtained directly with existing motors designed for

[40] J. J. Carty; U. S. Patent No. 449,106; filed August 16, 1890; issued March 31, 1891.

60-hertz power. In order to convert to 60-hertz direct-drive generators, the ringing frequency was changed to 20 hertz in 1917, a change which did not greatly affect the capabilities of existing ringers.

Because of the high cost of long subscriber lines, it was often desirable to connect a number of users to the same line. On these "party lines," ringing current applied to the line by customers or operators rang all of the bells on that line. Customers identified their individual signals by a code of long and short rings. The provision of selective signaling to avoid this situation occupied the talents of many inventors for a number of years. In the period from 1879 to 1891, a total of 161 United States patents on selective ringing were issued. These included devices and systems based on ringing currents of different strengths, currents of different polarities, currents of different frequencies, electromagnetic and clockwork stepping arrangements for closing local circuits to the ringer, and multiple circuits requiring concurrently acting relays at each station.

The two major selective ringing schemes which ultimately became standard in the industry may be classified as harmonic and polarity arrangements. In the former, different frequencies of ringing current are made available at the central office for ringing selections and tuned ringers responding to these frequencies are used in the subscriber stations. Such a system was patented by J. B. Currier[41] and first used commercially by the New England Telephone and Telegraph Company in 1882. The technology available for implementing this scheme was not adequate and performance was originally unsatisfactory. Later the technique was used in England and, after improvements made in 1895 by J. A. Lighthipe of Pacific Coast Bell, the system began to be employed by non-Bell companies and ultimately was used widely by them.

The harmonic system, even though it originated in the Bell System, was never widely used by Bell companies, probably because other arrangements were available that were more practical at the time. One of these was a biased ringer scheme invented by G. S. Anders of the National Bell Telephone Company.[42] The basic idea was to employ two polarities of pulsating current together with a Watson-type polarized ringer with an added spring which "biased" the armature to one side or the other, as desired. Thus, when the armature rested against one pole piece, a current poled so as to move it in that direction produced no effect but an oppositely poled current was fully effective. By combining

[41] Three patents were granted to him in 1881 (240,010; 246,374; and 251,097) covering both ringers and ringing generators. He relied heavily on pendulum or modified pendulum arrangements for controlling the generated frequencies.

[42] G. S. Anders; U. S. Patent No. 218,153; filed July 7, 1879; issued August 5, 1879. F. A. Pinkernell; U. S. Patent No. 511,276; filed August 7, 1893; issued December 19, 1893. A. S. Hibbard; U. S. Patent No. 555,725; filed September 16, 1895; issued March 3, 1896.

polarization and bias, two selections of ringing were available on one line. This was one basic element of the Bell System selective ringing arrangement which ultimately developed. The other was a scheme invented by F. A. Pinkernell of the American Telephone and Telegraph Company.[42] In this system, one station ringer was connected to ground from each wire of the metallic circuit; the ringing current was selectively applied to either side of the line and ground at the central office. This also provided two ringing selections.

In 1896, A. S. Hibbard of the Chicago Telephone Company combined Anders' polarity system and Pinkernell's divided ringing system to provide the first commercial four-party selective ringing system.[42] As shown in Fig. 3-42, he installed two oppositely biased polarized bells from each side of the metallic circuit to ground. Ringing current from the central office consisted of pulsating current, made up of either positive or negative half-waves of alternating current, applied selectively to either side of the metallic circuit and ground. Thus, four unique ringing signals were made available and full four-party selective ringing was achieved.[43] This principle formed the basis for nearly all selective ringing arrangements used in the Bell System up to the present.

Over the years the four-party scheme was improved in various ways, most of which are beyond the scope of this chapter. One exception is a

[43] Coded signals were used in addition when the number of parties exceeded four.

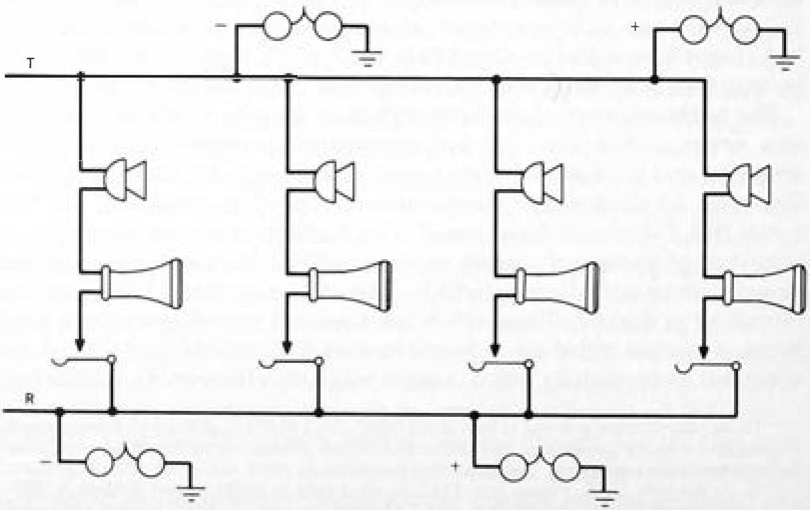


Fig. 3-42. A. S. Hibbard's circuit for full selective ringing on four-party lines.

patent by G. K. Thompson and E. C. Robes[44] which outlined a scheme whereby the ringers, with their ground connections, were normally detached from the line, but were brought into the circuit during the application of ringing current from the central office by means of alternating-current relays in series with capacitors bridged across this line. (Much later, electronic devices were used to replace the relays.) One of the several improvements from this arrangement was a reduction in the induced line noise which resulted from the line unbalance caused by grounded ringers.

4.2 Calling the Operator

As we have noted, the introduction of switching made it necessary to provide a means for the calling customer to alert the operator. To a considerable extent this replaced the need for the calling customer to ring the called user directly, but with party lines it was convenient to ring directly anyone on the caller's line and this practice was continued for some time. With this background, it is easy to see why it was convenient to use the same magneto generator for calling the operator that had previously been developed for subscriber ringing. At the central office it was impractical to use audible ringing and an "annunciator drop" associated with each line was used to alert the operator instead. In its simplest form this was a leaf hinged at the bottom and held in a vertical position by a latch which could be released by a magnet operated by ringing current. Upon release, the leaf fell under the influence of gravity indicating not only a waiting call but also indicating the particular line requiring the operator's attention. After establishing the desired connection, the operator would manually restore the drop to the latched condition and on completing the call the user would "ring-off" by turning his magneto generator, thus notifying the operator that the lines could be disconnected.[45]

This or very similar arrangements continued in use as long as local-battery offices were employed. When common-battery offices were introduced in the mid-nineties, the battery supply at the office provided enough power for signaling, and calling the operator was done automatically when the telephone receiver was lifted from its hook. The switchhook closed the dc path through the station circuit and the flow of current actuated relays at the central office which controlled actuated relays at the central office which controlled small lamps indicating the need for operator assistance and the status of the call. At this time all ringing, including party-line ringing, was done by the operator and the

[44] G. K. Thompson and E. C. Robes; U. S. Patent No. 644,647; filed July 18, 1899; issued March 6, 1900.

[45] Annunciators, relays, and related signaling apparatus are discussed further and illustrated in Section 3.4 of Chapter 6.

magneto generator was no longer needed, thus greatly reducing subscriber station size by removing its bulkiest elements, the batteries and the magneto.[46]

4.3 Controlling Machine Systems

The evolution of mechanical, or "machine," switching is covered in Section IV of Chapter 6 and will not be discussed further here except to point out that its introduction in the mid-nineties added an extra function to the station set. Not only was it necessary to indicate the desire to make a call (and its termination) as required with manual switching, but means had to be provided for conveying the information on the desired connection which previously had been transmitted orally to the operator. Several schemes were devised for doing this but the most practical one was to use a series of dc pulses to indicate the successive digits. Early mechanisms used pushbuttons, one for each digit, which the user operated as many times in succession as necessary to indicate the desired digit (e.g., three pulses for the digit 3).[47] By 1896 the pushbutton controls were replaced by the fingerwheel shown in Fig. 3-43.[48] In dialing a number with this device, the fingerhold corresponding to the desired number was engaged and the rotary member was pulled down to the stop, S^2, and released. The pulling action wound up a spring whose tension caused the dial to return to its normal position. The return rotation was limited to a moderate speed by an escapement mechanism and, during the return, the required number of circuit interruptions took place to control the central-office switches. Shortly thereafter a nearly full-circle dial using finger-holes was developed (Fig. 3-44). This dial had a larger travel and produced longer pulse intervals which resulted in better control of the central-office switches. This remained the sole way to control machine switching until *Touch-Tone*® dialing was introduced commercially in 1963.[49]

Originally, machine switching was used in small cities where four- or

[46] Although complete common-battery offices were not introduced until the mid-nineties, a common-battery signaling system using annunciator drops at the office and a switchhook closure to initiate the signal was used by J. J. Carty as early as 1880 in connection with a special service for an express wagon company.

[47] This system, invented by Almon B. Strowger, was put into operation in La Porte, Indiana, in 1892 and was known as the step-by-step system. Strowger formed the Strowger Automatic Telephone Exchange in 1891 and A. E. Keith and the Erickson brothers, John and Charles, were among the first employees of the company.

[48] A. E. Keith and J. and C. J. Erickson; U. S. Patent No. 597,062; filed August 20, 1896; issued January 11, 1898.

[49] An interesting change was made in the rotary dial by C. F. Mattke in 1963 (Patent No. 3,108,159 filed May 9, 1960) when he invented the "Space-Saver Dial" with movable stop. In this device the stop was placed directly below the finger hole for the numeral 1 but moved a sufficient distance, when the finger was pushed against the stop, to send a pulse when released. This eliminated the blank space between the 1 and the 0 finger holes and reduced the dial diameter accordingly.

(No Model.) 2 Sheets—Sheet 1.

A. E. KEITH & J. & C. J. ERICKSON.

CALLING DEVICE FOR TELEPHONE EXCHANGES

No. 597,062. Patented Jan. 11, 1898.

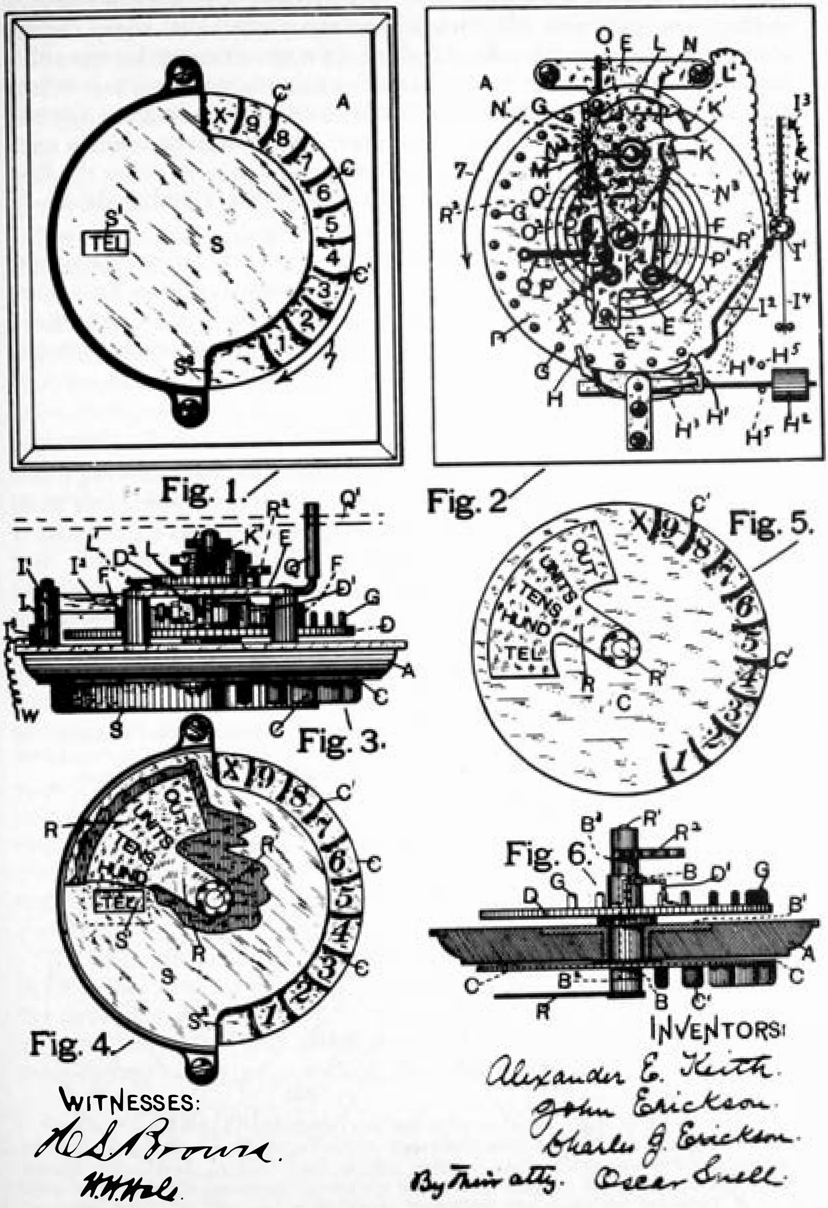


Fig. 3-43. First form of Strowger fingerwheel substation dial.

five-digit numbers were sufficient for the entire area and all-number dialing was employed (Fig. 3-44a). In large cities with many central offices the custom had developed of using a name to designate the office and a four-digit number to designate the line within the office. When machine switching had developed to the point where it could be applied in large cities, it was decided to continue the name-number system and, in 1917, W. G. Blauvelt of AT&TCo developed a system using the first three letters of the name together with four digits. The dial shown in Fig. 3-44b was devised to implement the system.[50]

The requirements on speed and duration of pulsing are quite rigorous and over the years many mechanical improvements in dials have been introduced to minimize the maintenance effort required to meet these requirements, reduce dialing noise, reduce dialing error, improve finger comfort, and so forth.

V. PROTECTION

The station apparatus presented the first need for measures to protect the user from electric shock and the delicate telephone apparatus from possibly damaging electric surges. At first, protection was accomplished by means of spark gaps incorporated in the station apparatus, but before long, fuses and other protective measures were incorporated and these devices were installed at the point where the telephone circuit entered the user's premises. The full story of protective devices is related in section 5.3.3 of Chapter 4.

[50] This system left many of the seven-digit number combinations unusable for phonetic and other reasons. Later the original system was changed to two letters plus five digits and currently an all-number system is being used in most places.

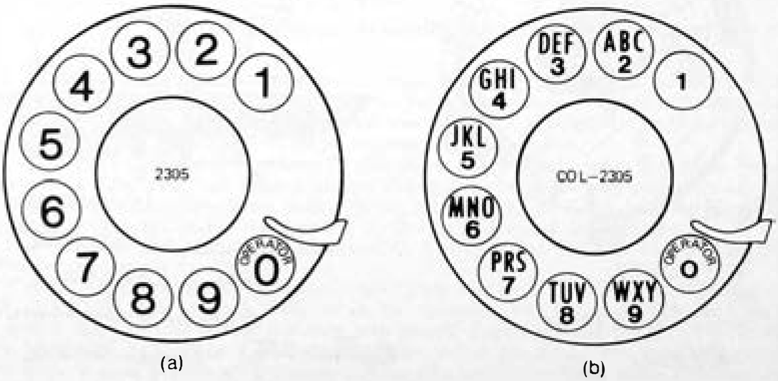


Fig. 3-44. Later types of rotary dials: (a) number-only dial; (b) number-and-letter dial.

Initially, some of the metallic parts of the telephone instruments connected to the line wires were accessible to the user. Beginning about 1906, special instruments were built in which these parts were either enclosed in an insulating housing or insulated from the accessible metallic parts of the station set. After about 1913 all standard designs were of the insulated type.

VI. CUSTOMER TELEPHONES

The circuit elements we have just discussed make up the equipment commonly found on the customer's premises. These elements have been assembled in many different ways to provide customer telephones but most of these fall into three basic categories. First, historically, was the wall telephone, a box fastened to the wall with the transmitter and receiver mounted on the outside, the inside containing most of the other elements. This was a "combined" type of station set in which essentially all station apparatus was contained in one package.[51] It was succeeded (but not fully supplanted) by the deskstand telephone which consisted of a small base (on which the dial was mounted) with a pedestal supporting the transmitter and a pair of prongs or other arrangement to hold the receiver and operate the switchhook. In order to keep this device small enough for practical desk use, the speech circuit and ringer were contained in a separate box (called a subscriber set) mounted on a nearby wall or some inconspicuous part of a desk. When the third basic type, the handset, was introduced, the practice of using two-part customer equipment was continued, the handset support or cradle containing only the dial and switchhook with the other elements mounted in a separate subscriber set. This was done partly because the ringer and speech circuit continued large in size and partly to facilitate the substitution of a handset for a deskstand without other costly changes. By 1937, improvements in materials and design techniques made it possible to include all the elements in a package small enough to use on a table or desk and the "combined set" again became practical and has continued in use ever since.

The first high-production handsets were introduced by the Bell System in 1927, just after the nominal period covered by this volume. However, the development is covered in considerable detail in Section 6.4 since most of the studies and development effort leading to the 1927 model were carried out in the preceding 5 to 10 years. In addition, the handset

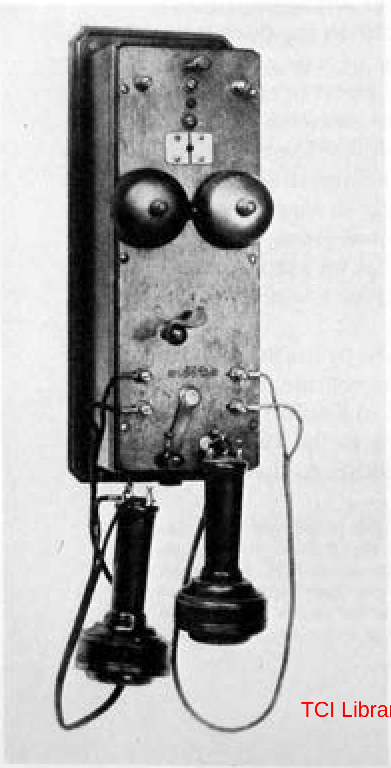
[51] After a few years the protective devices were removed to the point where the line entered the premises and this practice has continued. As a result these devices are usually considered more a part of the outside plant than as station apparatus. Batteries were often mounted separately when they were of the "wet" variety but as dry cells came into use, in the nineties, provision was made for including them in wall-set boxes.

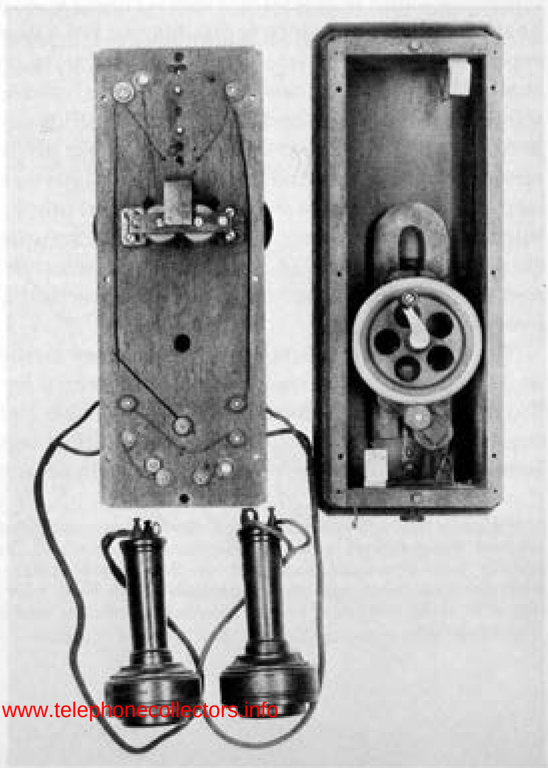
development illustrates an important change in design technique which began roughly 40 years after the telephone invention and has continued to influence station development ever since.

6.1 Wall Sets

While many of the early wall sets were manufactured only in small numbers, they are of considerable historical interest in illustrating the ingenuity of early technicians and the speed with which basic principles developed. The early telephone using a single hand-held butterstamp instrument for both transmitting and receiving was inconvenient and occasionally confusing, because it forced the customer to keep changing the receiver back and forth between his mouth and ear. To eliminate this situation, early in 1878 a wall set was introduced with two hand-held instruments, one for transmitting and one for listening. Such a telephone, shown in Figs. 3-45 and 3-46, was manufactured in the shop of Charles Williams, Jr., in Boston and was sometimes referred to as "Williams' Coffin" because of the shape and construction of the wall box.

Fig. 3-45. An early wall-mounted subscriber set, introduced in 1878, made by Charles Williams, Jr. Fig. 3-46. Interior view of Williams Magneto Set, showing hand-cranked generator and coils for the call bells (1878).





This set used the Watson polarized ringing system with the call bells and magneto crank on the outside of the box and the mechanism inside. Just above the call bells is the sawtooth spark gap or "lightning arrester" used for electrical protection. The switch below the magneto was used to transfer between the talking and calling condition. It was the inconvenience of this operation that led to Roosevelt's invention of the switchhook, illustrated previously in Fig. 3-34. The latter figure shows a wall-mounted transmitter, probably an early Blake type, in place of the second butterstamp telephone. Thus, from about 1879–80 the basic elements of the wall telephone were available and the many designs which followed differed from each other largely in physical arrangement. Some of the many varieties developed are shown in Figs. 3-47 and 3-48.[52]

Referring to the first figure, the 1877 box telephone was used only for private lines, the 1878 set shown as item (b) being the earliest form used with a commercial central office. This set, used in New Haven, Connecticut, employed a pushbutton to signal the operator by opening battery current from the central office. This released a numbered electromagnetic annunciator at the switchboard. Item (c), already discussed, incorporated magneto ringing but still used an electromagnetic transmitter.

The 1879 set shows the first use of the Blake transmitter with a local battery (of the wet-cell type) and also incorporates the switchhook and magneto ringing. The 1882 set was the first type manufactured by the Western Electric Company. The 1887 set used the long-distance transmitter illustrated previously in Fig. 3-10. The 1891 models were equipped with early types of the solid-back transmitter. The "cabinet" type set, item (h), was developed for use at prestigious public locations where considerable long-distance telephoning was to be expected. It contained space for three wet-cell batteries (in place of the single battery commonly used) in order to provide high transmitter output. This "desk type" set was in use as early as 1887 but at that time it was equipped with the Hunnings type of "long-distance transmitter." Items (i), (j), and (k) show early common-battery sets, local battery being used in all prior wall telephones.

By 1900 two distinct styles had developed, as shown more fully on Fig. 3-48. The local-battery, magneto sets [items (b), (d), (e), (f), and (k)] continued the use of wooden and usually large cases[53] whereas the common-battery sets were more compact and, after 1912, used metal cases.

[52] Some of the sets in Fig. 3-47 [e.g., items (a) and (h)] are not strictly wall sets but are included here since they were intended to be used in a fixed position with constraints on the user similar to wall sets.

[53] The wooden case was probably continued because these sets were manufactured in small quantity and did not justify the use of the expensive dies required to form large metal cases. They were large since they had to contain not only the common-battery apparatus but also the magneto and the batteries. The small magneto set of 1905 [item (d)] used a separate battery box.

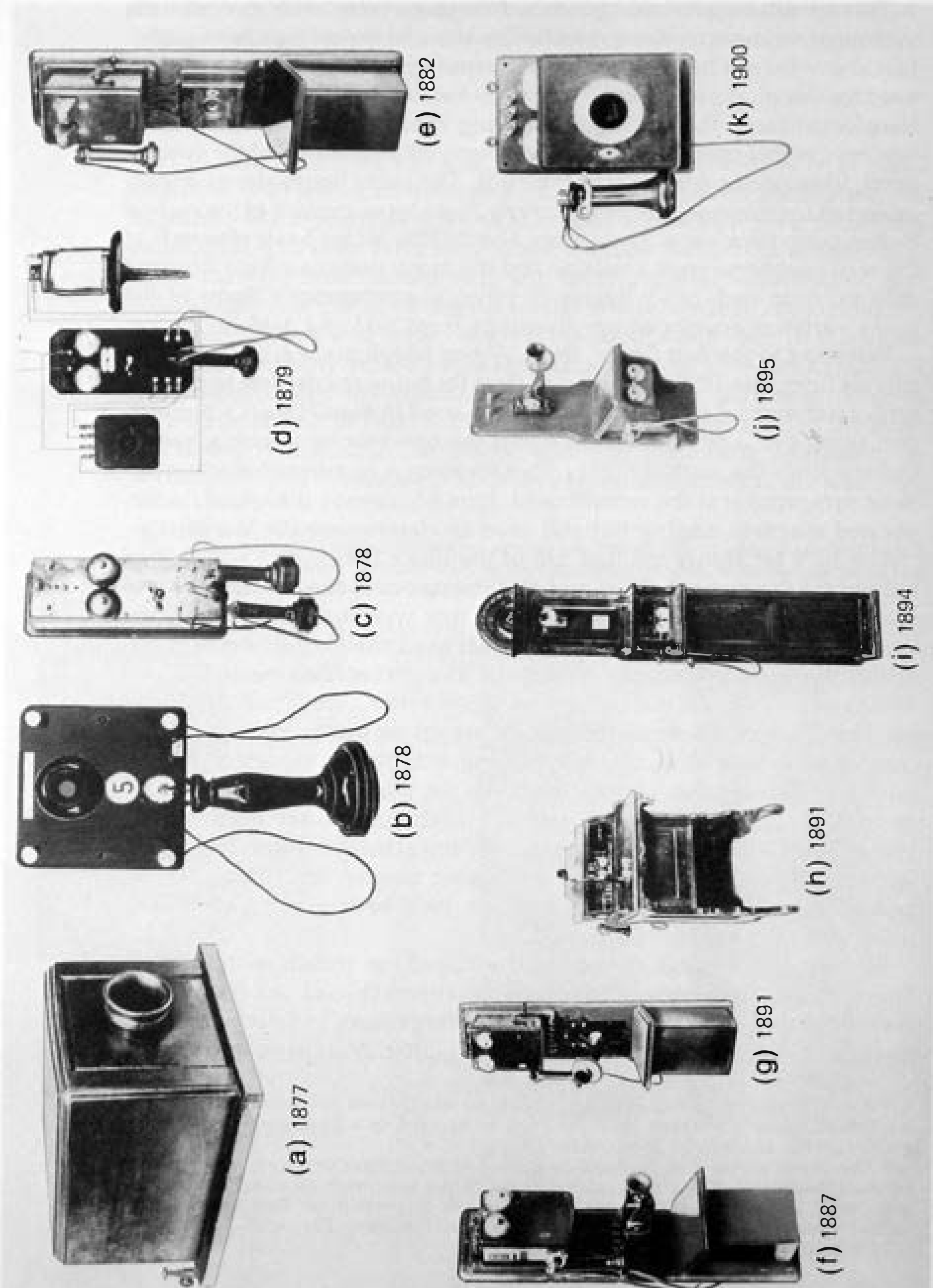


Fig. 3-47. Evolution of wall-type subscriber sets, 1877–1900.

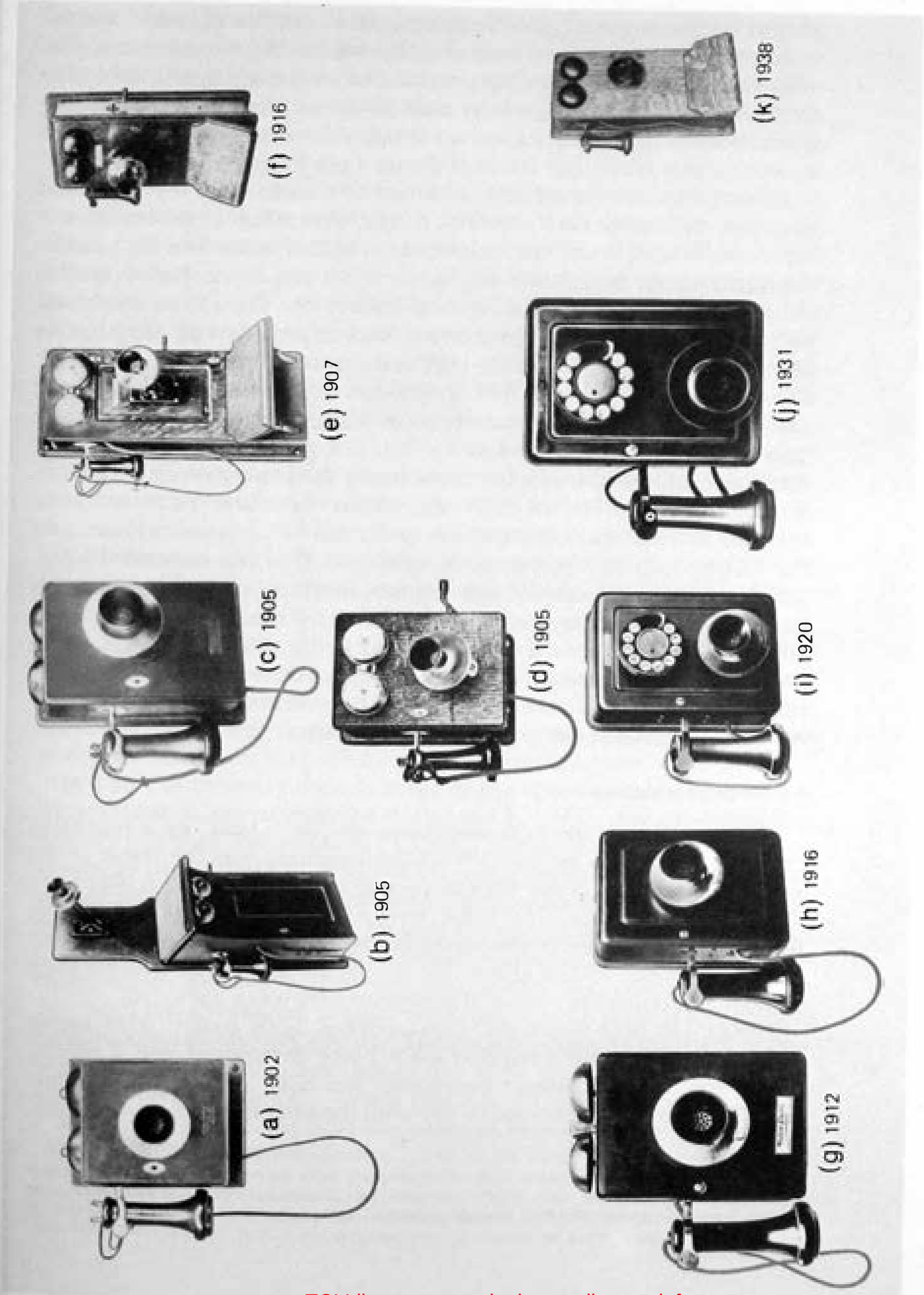


Fig. 3-48. Evolution of wall-type subscriber sets, 1900–1938.

Until 1916 the ringer gongs were external on all sets but after this date the common-battery sets were built with the whole ringer mechanism protected by the case, slots being provided to emit the ringer sound. The 1920 and 1931 models show how dials for operating machine switching systems were included; the former set illustrates the use of a dial with numerals only, the latter has a dial with both numbers and letters.

As noted later, when common-battery operation made the deskstand practical, the use of wall sets fell off. However, they never completely went out of favor, even in common-battery offices, and when the handset was introduced, they slowly began to stage a comeback. However, this is a story for a future volume. For local-battery use, the wall set continued to be the favored type of instrument since it provided all components in a single package.[54] The 1317-type set, shown as item (e) in Fig. 3-48, first standardized around 1907, continued in manufacture until about 1940. It was updated from time to time, a 1916 version being shown in item (f). This particular set was the first using two dry cells, instead of three, and improved instruments including the 144-type receiver and an insulated transmitter. In 1938 the transmitters and receivers were changed to make them comparable to current handset instruments and the AST induction coil was made available. This set, recoded the No. 1417, is shown as item (k). Some years later wall sets equipped with handsets were made available for local-battery service.

Some of our readers may be interested in Fig. 3-49 taken from a 1916 Western Electric catalog. It shows the construction of the No. 1317C telephone in detail and the notes at the side point out the design features which were considered important at the time.

6.2 Desk Telephones

The evolution of the desk telephone, or "deskstand" as it was often called, is illustrated by Fig. 3-50. This telephone was the result of the natural desire for greater convenience than possible with a telephone firmly fixed in place. The latter was not only inconvenient for the businessman who wanted access to his desk material while telephoning but also difficult to use effectively by people of various heights even though an adjustable transmitter support was often provided.

The earliest desk telephone [item (a) of Fig. 3-50], built in 1879, seems to have been merely a support for a Blake transmitter and a similar device for the long-distance transmitter was built in 1886. These primitive arrangements continued in use until the early nineties when a basic

[54] Common-battery operation was unsatisfactory both from the signaling and transmission standpoints on very high-resistance loops without complex and expensive supplemental circuitry. For this reason, local-battery, magneto sets continued in use on long "farmer's lines" even in nominally common-battery areas.

format began to evolve which became familiar to several generations of telephone users. One of the first features of this set, an angular adjustment for the transmitter, was introduced in 1891 and became universally used with the introduction of the solid-back transmitter as shown in item (e) and subsequent parts of the figure. The early deskstands used rather ornate curved supporting pillars with external switchhooks, but in 1892 a simple tubular support was introduced and the following year the switchhook contacts were enclosed [item (g)]. These two features were not universally adopted until 1900, the more ornate designs being slow to die. By 1900 [item (o)] the long-term format had been pretty well established and included the angular transmitter adjustment, tubular support with enclosed switchhook contacts, felt-covered bottom (introduced in 1894), and pronged switchhook with ends formed in a ring. This last feature was introduced in 1897 to avoid the possibility of damaging the receiver diaphragm by inadvertently striking it against the switchhook ends during the hang-up process. This design continued in use throughout the life of the deskstand. Another long-lived feature was incorporated in the 1900 design, namely, the use of an assembly within the tubular support, removable as a unit, which incorporated both the switchhook mechanism and the binding posts for connecting external and internal wiring. The ultimate design is illustrated in Fig. 3-51, and this highly functional arrangement continued in use for the next 40 to 50 years.[55] During this period the change in appearance was small but many changes in material and manufacturing processes were introduced to reduce cost and to improve reliability. The most important of the latter was insulation from line wires introduced in 1912–13 for all types manufactured.

In a few of the early models some of the speech circuit components were placed in the base of the deskstand but with one exception this practice had been discontinued by 1900. The exception is the 1904 model, shown as item (q) of Fig. 3-50, which was a portable set (probably one of the first of its kind) designed particularly for restaurant use. The more common arrangement was to provide the speech circuit and ringers in a separate box for wall or underdesk mounting. These boxes were called subscriber sets, one of which is shown with cover removed in Fig. 3-52. Subscriber sets are also included in several parts of Fig. 3-50 [specifically items (p), (r), and (t)]. Their omission from other parts of this figure does not imply their lack of use since they were required with all but a very few models. When machine switching was introduced, a dial was added on the base as shown in item (u) of Fig. 3-50.

While the need for deskstands arose at an early date, they could not

[55] Manufacture of deskstands was discontinued in 1940, but they remained in use for many years thereafter.

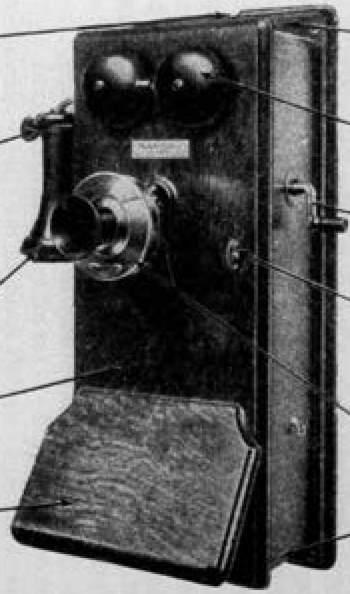
Backboard grooved for entrance of line wires at top or bottom.

Switch hook is compact, strong and durable. Finished in durable black.

Receiving efficiency is unsurpassed. Receiver strong and durable. Fits the ear.

Cabinet of solid quarter sawed oak, substantially made. Attractive design.

Writing-shelf placed at convenient and comfortable angle, securely fastened and supported.



No outside binding posts.

The 2½ inch brass gongs give a loud, clear tone. Finished in black.

One-piece generator crank. Finished in black.

Special self-adjusting lock.

High efficiency transmitter with low battery consumption. Mounted on substantial black enameled bracket, securely fastened to cabinet.

All corners dovetailed and glued.

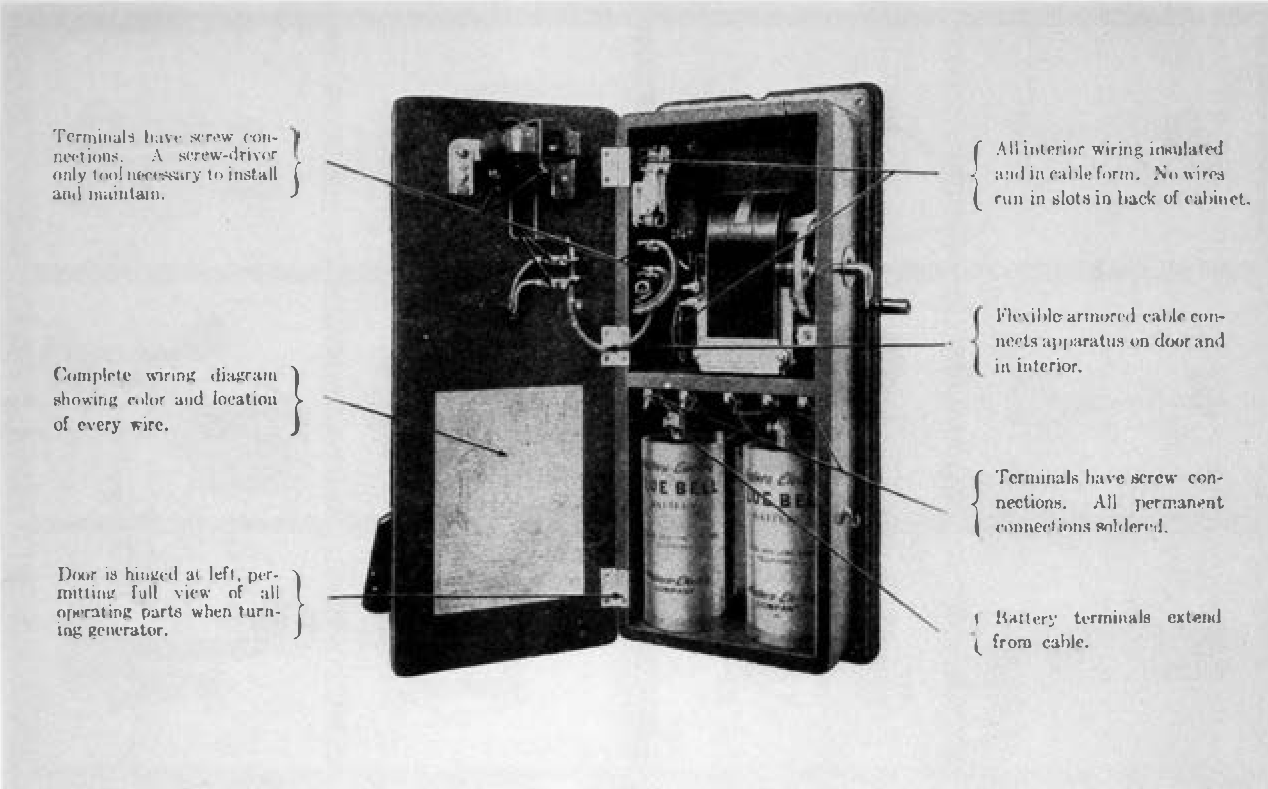


Fig. 3-49. No. 1317C magneto telephone.

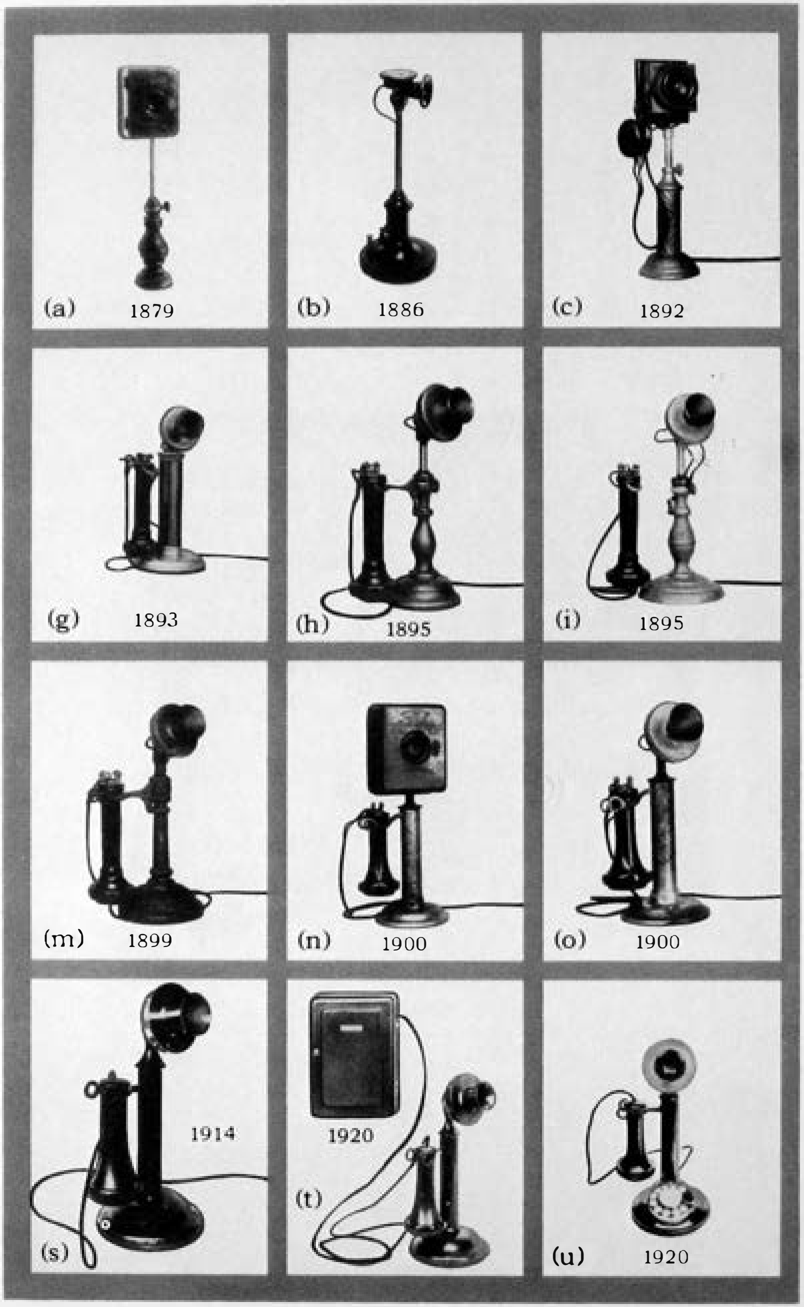
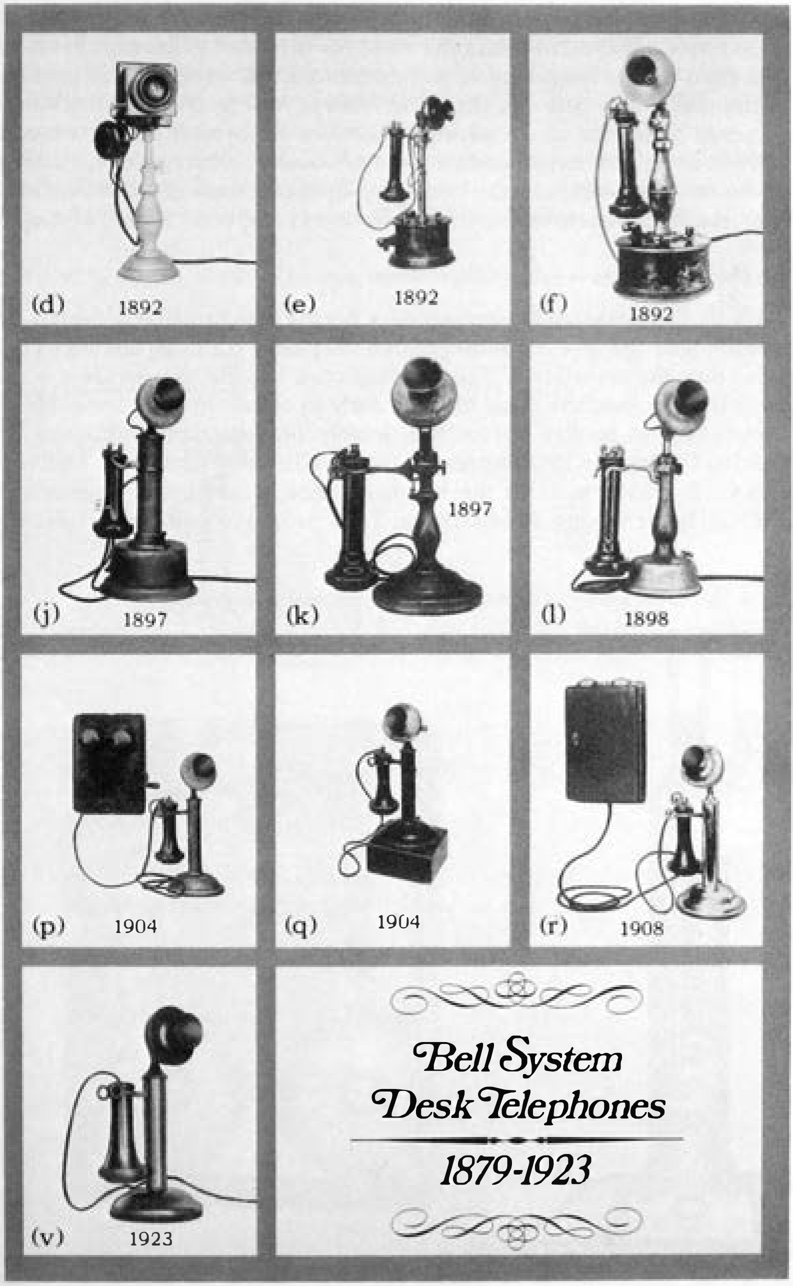


Fig. 3-50. Evolution of desk telephones, 1879–1923.


(d) 1892
(e) 1892
(f) 1892
(j) 1897
(k) 1897
(l) 1898
(p) 1904
(q) 1904
(r) 1908
(v) 1923
Bell System
Desk Telephones
1879-1923

realize their potential convenience until common-battery central offices were introduced, eliminating the need for the hand generator. Prior to this time it was necessary to either use a small generator of limited usefulness in the base [as shown in item (e) of Fig. 3-50] or include a standard generator in a separate subscriber set, which nullified much of the convenience associated with the deskstand. With common-battery offices the deskstand rapidly became the popular form of telephone and wall telephones became less commonly used except on "farmer's lines."

6.3 Hand Telephones—Early Experience

While the deskstand represented a highly functionalized design, it was obvious that the convenience of a telephone could be enhanced by mounting the transmitter and receiver on a handle to provide a one-piece instrument. The basic idea is nearly as old as the telephone. Bell's demonstration lecture before the Society of Telegraph Engineers in London in October 1877 impressed two Englishmen, Charles E. McEvoy and G. E. Pritchett, with the inconvenience of using one instrument alternately for talking and listening. They promptly and independently

Fig. 3-51. Cross section of deskstand showing internal arrangements.

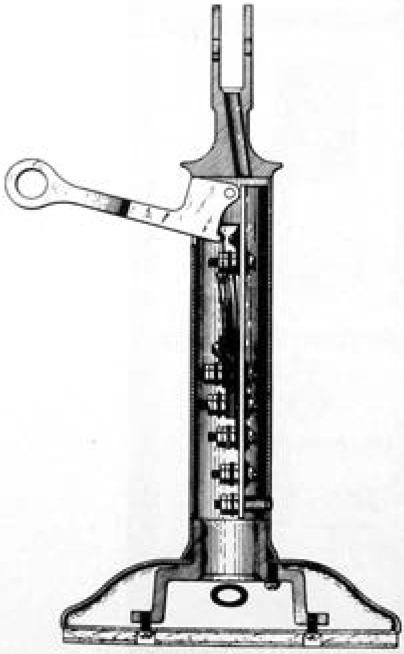
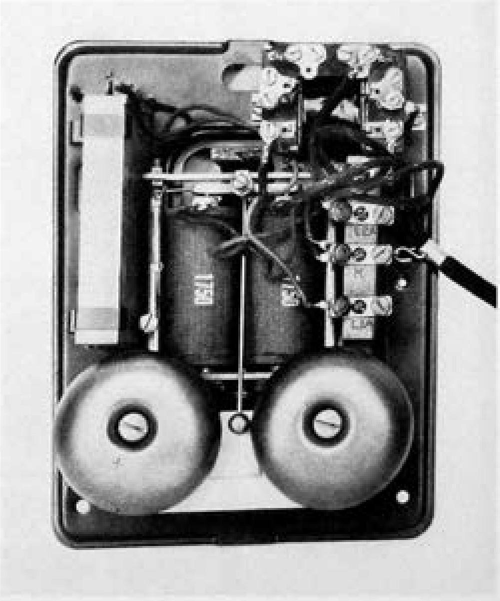


Fig. 3-52. Subscriber set.



obtained British patents on arrangements of instruments and speaking tubes that could be held with one hand or supported on the user's head or shoulder. These arrangements, in which the instruments are not specified, are illustrated by sketches appearing in the patents (Fig. 3-53). They were not developed for commercial use.

In the next year, 1878, R. G. Brown, the head operator of the Gold and Stock Exchange in New York, designed the handset illustrated in Fig. 3-54.[56] This handset comprised an Edison transmitter of the single-

[56] R. G. Brown; U. S. Patent No. 224,138; filed September 29, 1879; issued February 3, 1880.

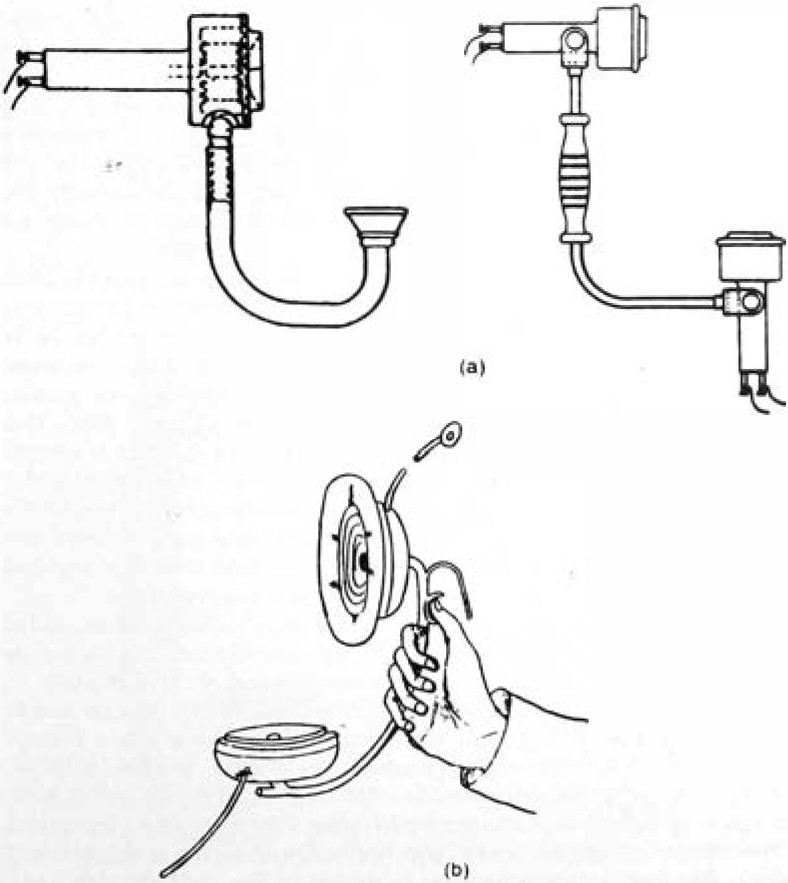


Fig. 3-53. Early handset designs: (a) by C. E. McEvoy; (b) by G. E. Pritchett.

contact type and a small single-pole receiver. The handle was a relatively heavy steel bar which served as the permanent magnet for the receiver. Handsets based on this design were used by the Gold and Stock Exchange operators. Brown later went to France where he accepted the post of Electrical Engineer of La Société Générale in Paris. Handsets based on his design were widely used in Europe, especially in France, and became known as French phones. They were not used in the Bell System because of their low efficiency and unstable performance relative to the wall and deskstand telephones available in the United States.

In the 1890–1902 period, three handset designs were produced in limited quantities by the Bell System. One of these was constructed to military specifications for use by the United States Army and Navy (Fig. 3-55). The second was an experimental model using a hollow aluminum handle (Fig. 3-56). Both of these were equipped with commercial carbon transmitters and watch-case-type receivers.

In 1902 a handset was produced for use by linemen in areas having common-battery service. This handset, called the model 1001 lineman's set, was equipped with a No. 244 carbon transmitter and a No. 131 watch-case receiver mounted on a hollow metal handle; initially the handle was formed of sheet steel but later it was changed to aluminum (Fig. 3-57).

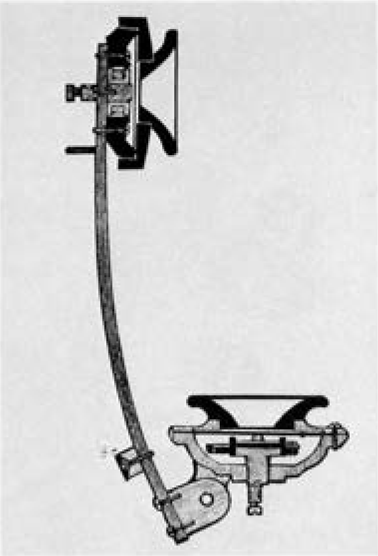
The Model 1002A handset (Fig. 3-58a) was originally designed in 1902; the Western Electric Company was authorized to produce 100 units in 1904 for trial. It comprised a No. 267W carbon transmitter and a No. 141W permanent-magnet watch-case receiver mounted on a hollow metal handle. In 1904 the transmitter was redesigned to place the carbon button at an angle of 45 degrees to the diaphragm (Fig. 3-58b). This change was expected to reduce trouble due to positional effects encountered in testing. The modified design was coded the Model 2 handset and the production of 3,000 units was authorized for use by the Associated Companies. All of these handsets were recalled from Bell System service in 1907, but a number were sold to non-associated telephone companies and apparently met their requirements.

The Western Electric Company produced an intercom handset, coded the Model 1003 in 1910 (Fig. 3-59), which received limited use for its special purpose. It used a No. 320 transmitter and No. 183 receiver.

During the 1900–1921 period the demand for handset telephones in Europe increased and designs were produced by the Western Electric Company to serve the European market. One of these, the Model 1005B, using the No. 324 transmitter and No. 188 receiver, is shown in Fig. 3-60. It was originally manufactured by Western Electric at the Hawthorne Works near Chicago but later production was transferred to the Antwerp shop. Another Antwerp handset is shown in Fig. 3-61 (Model 2266). Both of these were tested for possible Bell System use but were found

Fig. 3-54. Patent drawing of the handset designed by R. G. Brown in 1878.

Fig. 3-55. An Army-Navy handset of 1890.



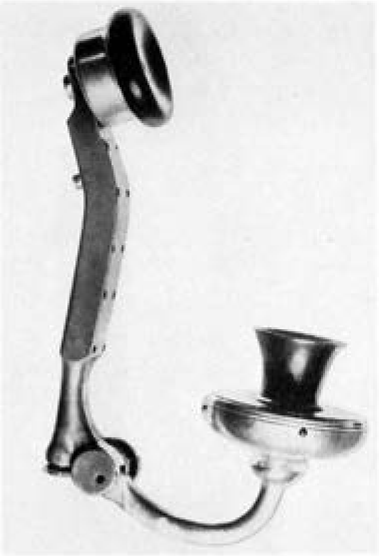


Fig. 3-56. Aluminum-handle handset of 1895.

Fig. 3-57. Lineman's handset for use in common-battery system.

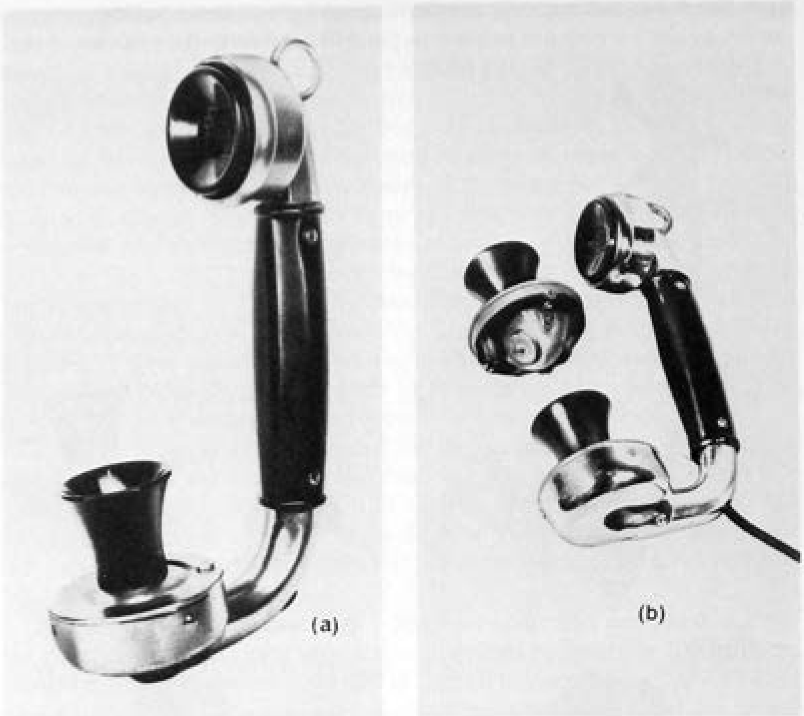


Fig. 3-58. Handset designs of 1902–1905: (a) Model 1002A (1902); (b) Model 2 (1905).

unsuitable because of the inherent limitations in early handsets, discussed in the following section.

6.4 Hand Telephones for Bell System Use

It is natural to wonder why it was, with all this background of experience, that practically no handsets were used in the Bell System until the late twenties. The answer, in simple terms, was that this experience clearly showed that the handsets available up to this time could not meet the rigorous requirements of Bell System service which, unlike the smaller European administrations, required transmission over distances of 3,000 or more miles.

The handsets produced up to about 1915 had two basic faults which had to be overcome before this type of instrument could receive widespread Bell System use. One of these was the variation in efficiency of the transmitter with position. The solid-back transmitter which had provided the efficiency and reliability necessary for the spread of tele-

Fig. 3-59. Western Electric interphone of 1910 with Model 1003 handset.

Fig. 3-60. Model 1005B handset (1915).

Fig. 3-61. Model 2266 handset (1921).

phone service throughout the United States was stable and uniform in performance when its position varied only within the limited range permitted by the adjustments available on deskstands or wall sets. It was found through experience that a handset was used in a large variety of positions in some of which the efficiency was greatly impaired. Some variation in efficiency was acceptable where the line losses were small, as on short-haul circuits, but on routes a thousand miles or more in length, maximum efficiency was necessary until such time as line losses could be reduced by amplification. Additionally, the constantly changing position associated with handset use caused abrasion of the carbon, increasing its resistance and generating noise.

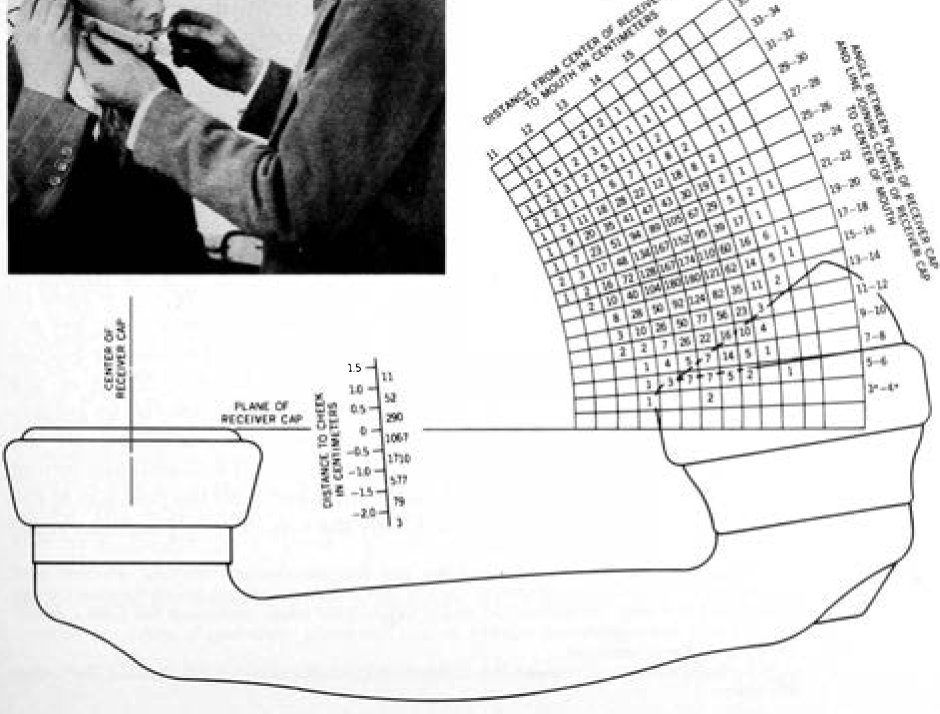
The second problem arose from the acoustic feedback between receiver and transmitter. Sounds from the receiver were conducted through the handle (or other coupling paths) to the transmitter and thence via the sidetone path of the speech circuit back to the receiver. Oscillation, or singing, occurred when there was a net gain (amplification) around this path. With the deskstand, this was not a serious problem since the receiver was ordinarily poorly coupled to the transmitter, being either on the ear (and shielded from the transmitter) or at some distance from the transmitter.[57] Several means for curing this problem existed with varying degrees of practicality. Employing a handle with poor sound-conducting properties was very helpful while the handset was being used, but with the resonant, high-efficiency instruments then in use singing would still occur when the handset was placed on a desk or other hard surface that provided good acoustic coupling. Lowering the peak instrument efficiency was also effective but undesirable for reasons previously noted. Finally, the use of a good anti-sidetone circuit would also reduce the potential for singing but simple circuits of this type did not become available until Campbell's work around 1918.

The faults of the handset were appreciated at an early date and were outlined in a 1907 letter written by J. J. Carty, the Chief Engineer of Western Electric. However, means for accomplishing a cure were not available at the time. By the end of World War I, the underlying causes were quite well understood and the availability of the anti-sidetone circuit and amplification for reducing line losses (and potentially reducing the need for extremely high instrument efficiency) offered some hope that the inherent faults of the early handsets could be corrected if the positional problem could be reduced. Accordingly, a comprehensive program for the development of a handset was begun. The objectives outlined by G. K. Thompson on May 18, 1918, were:

[57] However, older readers may recall that they or their childhood friends learned at an early age the possibility of annoying users by bringing the receiver and transmitter together until singing occurred.

(*i*) Dimensions must be based on head measurements.
(*ii*) The handset must work properly in any position.
(*iii*) Howling must be avoided.
(*iv*) The handle and all exposed parts must be made of insulating material.
(*v*) Carbon noise must be avoided.
(*vi*) The transmission performance must be equal to that of the most

Fig. 3-62. A gauge being used for making head measurements to determine the proper dimensions of the handset. Fig. 3-63. Distribution of head measurement data. (Redrawn from Jones and Inglis 1932, Fig. 11)



efficient transmitter and receiver in current use so that the handset could be interchangeable with existing station sets.

Optimum dimensions, including the spacing of the receiver and transmitter and their angular positions relative to the handle and to each other, were an important matter. A handset configuration had to be comfortable for most users while keeping the transmitter close to the lips since an unnecessarily large distance would reduce efficiency.[58] In order to resolve this problem, 4,000 measurements of head dimensions were made in 1919 on a sample of adults carefully selected to approximate the population as indicated by the most recent census. The gauge and method employed are illustrated in Fig. 3-62 and the results, together with an outline of the chosen handset drawn to the same scale, are shown in Fig. 3-63. With the chosen configuration the average distance between lips and transmitter mouthpiece was reasonably small and all but 3 percent of the adult population could use the handset with the receiver held to the ear in a normal manner. Only minor adjustments were required by the remaining users. This survey of the population was one of the early applications of anthropologic measurements in industry and one of the first applications of "human factors" studies in the Bell System. Succeeding studies in which the talker volume from the transmitter was measured during actual telephone calls and compared with similar information obtained with other transmitters extended the human factors approach, a technique much used throughout industry today.

Solving the transmitter positional problem required an entirely new approach to transmitter design. The solid-back transmitter had optimum efficiency in the usual vertical position of use. With handsets, the diaphragm was seldom vertical, more often being used in a nearly horizontal position in which case the carbon fell away from the uppermost electrode, giving rise to poor quality of speech and high resistance, the latter resulting in excessive heating and noise. Many shapes of carbon button were tested, but the solution finally adopted was the so-called barrier button, invented by C. R. Moore.[59] In the barrier-button transmitter (Fig. 3-64) the carbon container, placed in front of the diaphragm, comprised a hemispherical electrode separated by an insulating barrier from a ring electrode; the carbon mass was driven by the diaphragm but the electrical path through the mass was independent of the diaphragm and remained relatively unaffected by the position of the instrument. Performance was practically the same in any position except

[58] Increasing the distance between the lips and the transmitter from one-half inch to one inch reduced the transmitter output (for the same speech level) by about 3 dB. With the deskstand, the user could easily talk closely when necessary but with a handset, moving the transmitter relative to the lips could unfavorably affect the receiver position and performance.

[59] C. R. Moore; U. S. Patent No. 1,565,581; filed October 3, 1921; issued December 15, 1925.

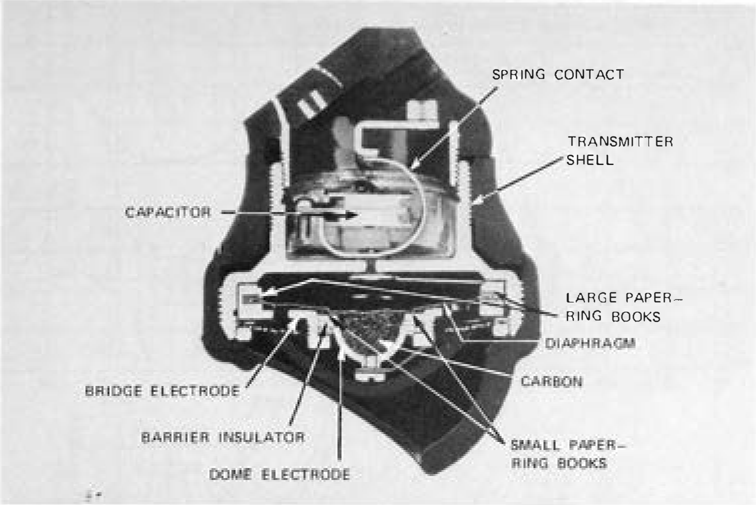


Fig. 3-64. Cross-section view of barrier-button transmitter (1922).

for some distortion when the domed portion of the button was underneath the diaphragm. To avoid this problem the button was located on the front of the diaphragm instead of the back as had been customary in the past. With this arrangement, the carbon granules are held in intimate contact with the diaphragm in all positions except when the diaphragm is face down, a position almost impossible to achieve in use. The extent of improvement in noise and performance achieved with this design is illustrated in Fig. 3-65.

Several measures were used to solve the howling problem. The effect of sound transmission through the handle was minimized by using a handle[60] made of material that dissipated sound energy and was designed to have a resonant frequency above the voice range (where the instrument efficiencies were low). Further control was obtained by a radical departure from previous transmitter design practices. Instead of using a resonant system peaked at the same frequency as the receiver, a highly damped system was used with a broad response peak in the upper voice range. This was obtained by using a very light but stiff diaphragm supported at the rim by a "book" of annular paper rings in a rather loose assembly. These rings damped out some of the resonances and permitted the diaphragm to vibrate as a unit. This resulted in a transmitter

[60] W. C. Kiesel; U. S. Patent No. 1,435,977; filed December 8, 1919; issued August 15, 1922.

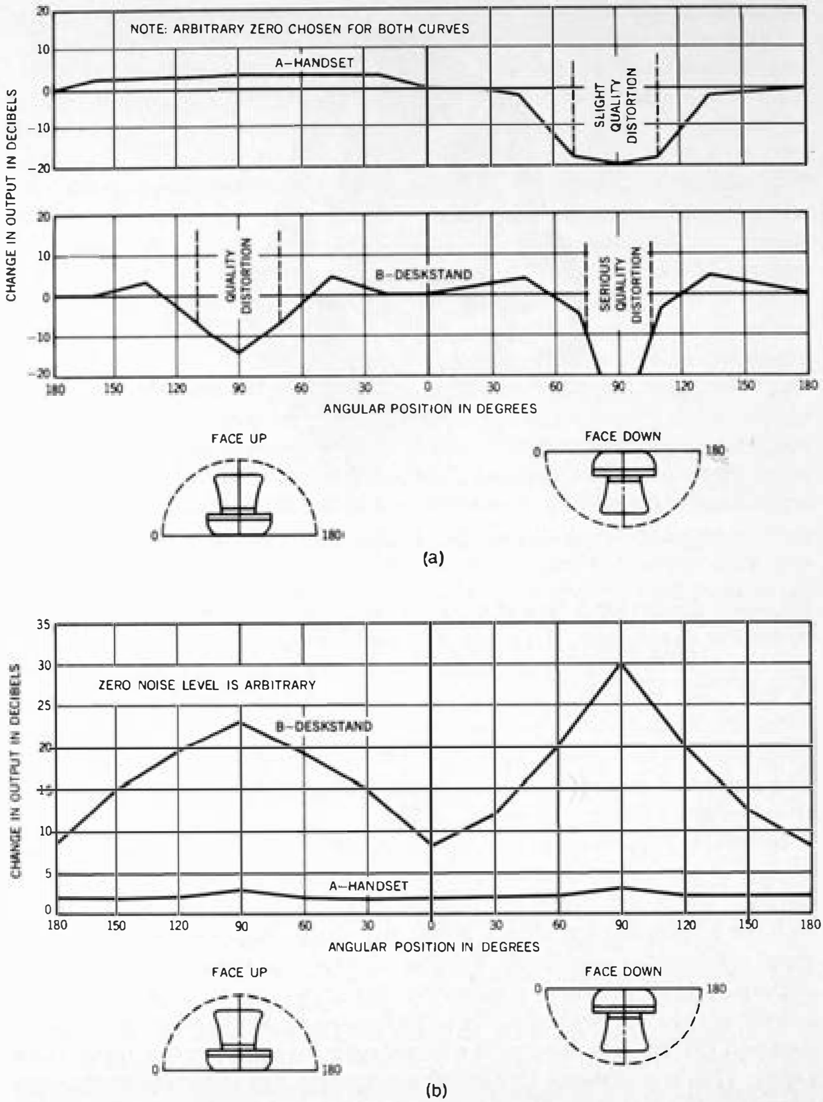


Fig. 3-65. Position effects with barrier-button transmitter: (a) effect of position on transmitter output; (b) effect of position on carbon noise. (Redrawn from Jones and Inglis 1932, Figs. 7 and 8)

responding to a wide range of frequencies but with a greatly reduced peak efficiency as shown in Fig. 3-66a (with the similar deskstand characteristics shown for comparison). The receiver characteristic was essentially unchanged as shown in Fig. 3-66b, but because of the transmitter

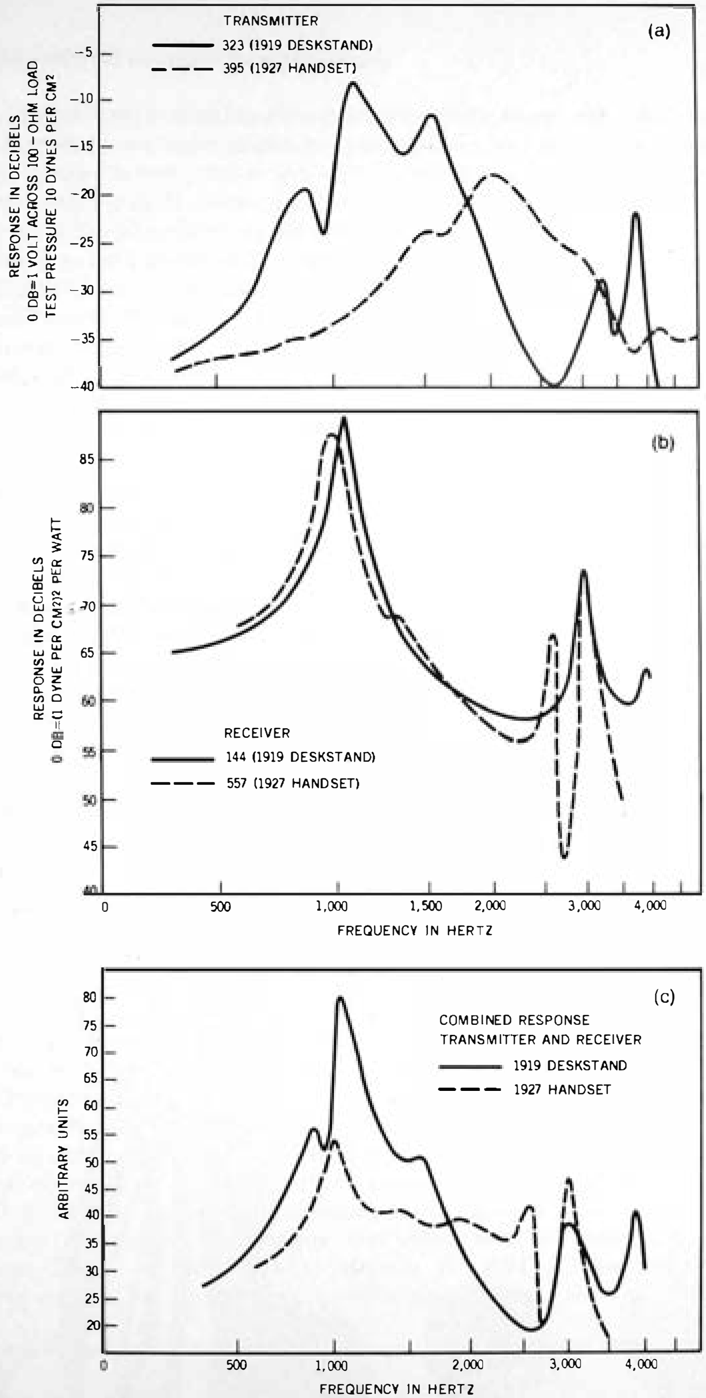


Fig. 3-66. Transmitter and receiver characteristics (1919 deskstand and 1927 handset).

modification, the peak efficiency of the combination (the significant factor in the singing path) was greatly reduced and the overall frequency response was flatter and broader. (See Fig. 3-66c.) For a fixed speech input, the output was, in terms of loudness, lower than for the highly resonant deskstand and would have been quite unacceptable on the basis of design criteria used ten years earlier. However, by the mid-1920s the use of field measurements to determine talking volume and articulation tests to evaluate overall quality changed the picture. It was found that the much reduced sidetone that accompanied the change in transmitter characteristic caused a considerable increase in the acoustic speech level used by the talker with a net decrease in electric volume on the line very much smaller than the apparent reduction in instrument efficiency. Articulation tests suggested that the improved response at the higher frequencies would easily compensate for the volume loss, and this was later confirmed by the repetition-counting technique.[61]

These were the major changes which made the handset commercially practical but numerous other improvements in the art were also incorporated in the final design. For example, methods for manufacturing carbon granules were revised, including the selection of raw material and techniques for grinding, roasting, and cleaning. Iron particles derived from iron pyrites occurring in the coal were removed by a magnetic separator. Flakes and slivers were removed from the final product by elutriation in an air stream. The material so processed had increased uniformity and stability.

The receiver ultimately used was basically similar to previous small receivers. It used improved magnetic materials that would have made an efficiency increase possible. However, the net benefit would have been small because of the accompanying increase in noise and sidetone. Instead, the improvements resulting from the new materials were used for two other design changes. One was increased separation between pole pieces and diaphragm to minimize the chance of contact (freezing) which sometimes occurs when clamped diaphragms are subject to temperature changes. The other was a reduction in receiver impedance which reacted with the speech circuit so as to give the same net receiving efficiency but with reduced sidetone and some increase in transmitting efficiency, thus partially compensating for some transmitter changes.

These design changes were accomplished over a considerable period of time during which numerous minor changes were made and tested. Tests were conducted on preliminary handset designs in the 1922–24 period; the first major test, under field conditions, started on September 14, 1924. A sample of 490 units of the Type A design (Fig. 3-67) was manufactured and placed in service with

[61] Section V of Chapter 4 covers these new testing techniques more comprehensively.

Bell System employees at the American Telephone and Telegraph and Western Electric locations in New York City. Several deficiencies were found in this trial and redesign continued, passing through Types B, C, and D. The major change during this period was in the receiver. The Type A set had used a split-winding receiver, serving as part of an anti-sidetone speech circuit. This required electromagnetic bias and proved to have unsatisfactory characteristics, and the more conventional permanent magnet using a new cobalt steel was substituted. A new form of receiver cap was also introduced which was easy to center on the ear and provided a good acoustic seal. Other, internal alterations included changes in dimensions of the carbon chamber, modification of the oiled-silk moisture barrier, and addition of a small capacitor across the transmitter terminal to reduce electrical packing found to be caused by radio signals and by transient currents due to switching operations. Finally, in late 1926, production was begun on the E1A handset (Fig. 3-68a) by the Western Electric Company in the Hawthorne Plant, and it was made available for general use early in 1927. This handset, which had been found to be "free of gross fault" in articulation,

Fig. 3-67. Tool-made sample of the developmental handset of 1923.

Fig. 3-68. (a) Cross section of early E1A handset (1926). (b) E1A handset and A1 handset mounting (1927).

efficiency, and durability during a two-year period of exhaustive tests, comprised the 395B barrier-button transmitter, the 557E receiver with cobalt-steel magnets and ferrotype diaphragm, and the E1 handle of phenol plastic material, equipped with a 4-foot 83B cord. When not in use, the handset was "hung-up" by placing it in a horizontal position on a cradle actuating the switchhook as shown in Fig. 3-68b. It was found to be practicable to use this handset interchangeably with the existing deskstand in the telephone plant. The existing speech circuitry was usable since the handset transmitter and handle designs eliminated the need for sidetone reduction except as already provided on very short loops by the STR connection (refer to Section 3.1). However, as noted previously, the new field-testing technique demonstrated the other advantages of anti-sidetone circuits and such arrangements became available a few years after the E1A handset.

The development of the handset has been dealt with at length because it marks a major milestone in the evolution of the Bell System and telephone technology. The handset of 1927 introduced a new form of telephone instrument which has become the universal type of customer instrument of today. But more importantly, it marked major changes in design techniques. Prior to World War I, the telephone instrument was largely the result of inspired invention and empirical testing, mostly subjective in nature, without the benefit of basic theory and quantitative testing techniques. By the 1920s transmission theory had developed rapidly and Bell engineers had adapted the advanced electrical theory to the design of electromechanical systems and the handset transmitter was one of the first telephone applications of this work.[62] At the same time both laboratory and field measurement techniques had developed to the point where the theoretically inspired designs could be evaluated quantitatively. It was only through such techniques that it became possible to demonstrate that optimum design, in the new age of amplification, lay in the direction of improved naturalness and quality of speech even though this might involve large reductions in loudness efficiency.

6.5 Public Telephones—Coin-Operated Sets

6.5.1 Background

The telephone sets just discussed were designed for installation in homes and business locations of the general telephone user. These users "subscribed" for telephone service and paid for it by means of a fixed monthly charge plus additional fees for long-distance or other special services, according to use. Generally speaking, usage was restricted by the subscriber to his family, friends, and business associates. There was, additionally, a need for a service, preferably on a pay-as-you-call basis, for those who were unable to subscribe for full service. This was a prime motivation for "public

[62] As discussed later, broad research into the speech and hearing process was started about 1915. In the twenties, instrument designs began to be greatly influenced by this work not only through the application of its results but also by the use of the analytic and measurement techniques devised in support of the research study.

telephone service"[63] in the early days of telephony, but another reason soon arose and has become more important as time has passed. This is the provision of readily available service to a subscriber or anyone else whenever the need arises and wherever he may be. Thus, in the early days, the greatest need was for well-distributed public stations in residential areas where subscriber service was sparse. Today, the use of public stations has greatly expanded to include locations along highways in lightly populated areas, as well as on busy streets in central city areas.

The simplest way to meet the initial need was to install a conventional telephone station in a public location with an attendant to collect the fee and set up the call or instruct the user as required. Because of the high expense involved in paying an attendant, this procedure is practical only where the calling rate is high and is still used at such locations. Since true convenience to the user required widespread stations, even where the usage was low, it was necessary to find a less costly scheme for such locations. A telephone automatically controlled by a coin deposit seemed to be the answer and such devices were invented at an early date.[64] The basic principle underlying most of these contrivances, which came to be known as "coin boxes," was the use of a coin to unlock a basic element in the station without which service would be impossible.[65] Originally they were adjuncts to conventional subscriber telephones but it was soon found preferable to design a special instrument or "coin telephone" in which the telephone and the coin-collecting functions were combined in a single housing.

The basic coin box scheme had two main variants: in one, known as the prepayment system, a coin deposit was required before the user could make contact with the operator (or dial a call in machine switching areas). With this arrangement a coin return mechanism was required so that the user could be reimbursed if for any

[63] Strictly speaking, two categories of public telephones are recognized today. The term, without qualification, refers to a telephone installed by the telephone company on its initiative in rented space or outdoors. The rent, or commission, is often based on the use of the phone. Such phones are usually not listed in the directory. "Semipublic" telephones are installed for a combination of subscriber and public use at locations such as filling stations, clubs, apartment houses, etc., where public usage is small or restricted. Such phones, often used mostly by the subscriber, are listed in the directory and involve a minimum charge which is made up by the subscriber if the collections fall short of the necessary amount. In the text, the term "public telephone" is used in a comprehensive sense since, from a technical standpoint, the distinction between categories is unnecessary.

[64] Coin-operated devices did not originate with the telephone but were reasonably well known before the 1880s.

[65] While the provision of the public telephones was the main incentive for developing coin boxes, they were also employed to some extent in residences, as covered in Section 7.3.

reason a charge was not required. In the second, or "postpayment" system, the call was set up in the usual manner without a coin deposit, but the connection to the called party was not completed until payment was made. The former scheme required more complex equipment to provide coin return but required less operator time. The latter scheme, while basically simpler, not only required more operator time but also frequently left the called subscriber waiting on the line while the caller fumbled for the necessary coins. The most significant advantage of postpay was that it permitted the caller to reach an operator without a coin deposit, a distinct advantage in an emergency (but also giving rise occasionally to endless arguments with the operator under non-emergency situations). Both systems have been and are used in the Bell System, but the prepayment plan was the preferred arrangement after about 1906 and became used almost universally because of traffic holding-time savings.[66]

With this as background, it is possible to summarize briefly the evolution of public telephones, sometimes referred to as "pay stations."

6.5.2 *Early Pay Stations*

Attended pay stations, located in public buildings, were operated as early as 1878 by Thomas B. Doolittle on a private intertown connection between Bridgeport and Black Rock, Connecticut, as a service provided by the Social Telegraph Association. A fixed rate of 15 cents for each call was charged.

Reference to a "pay station," which was probably the first public telephone in the world, appeared in the *New Haven Register* of May 25, 1880. The article announced details of the formation of the Connecticut Telephone Company and stated that "a 'pay station' will be established there, where for 10¢ or a similar sum, anyone can talk with any telephone owner." Other attended public pay stations were opened in New York City in 1880 in certain offices of the American District Telegraph Company.

The collection of the charge by the telephone company attendant

[66] The need for a coin deposit on emergency calls was recognized as a distinct disadvantage and was debated extensively before prepayment was adopted. As time passed, and the value of operator time increased, the economic benefit from prepayment grew accordingly. At the same time, the density of residential phones grew rapidly and the use of coin stations for meeting domestic emergencies practically disappeared. However, by the 1960s the situation had begun to change significantly as the accident rate on highways and the incidence of street crime increased and brought greater demands for emergency service in these areas. The expansion of coin service in such areas met much of this need but in a few cases the necessity for an initial coin deposit nullified their potential value. As a result, extensive central-office equipment changes were initiated in the 1960s to make emergency service available without a coin deposit while retaining the prepayment requirement on other calls.

or agents (such as hotel or drug store employee) was the principal method used for payment at public stations during the next 10 or 20 years but ideas for automatic collection began to develop during the 1880s. One of the first patents for a coin box was issued to Edmunds and Howard in 1885.[67] It provided for a prepayment box in which the coin could be collected or refunded by the central-office operator. This was a basic concept that was widely adopted later but the Edmunds and Howard scheme seems not to have been implemented.[68] The first public coin station (manufactured by William Gray and illustrated in Fig. 3-69) was installed at Hartford, Connecticut, in 1889. The following year the New York Telephone Company placed in service ten coin boxes manufactured by J. H. Bunnell and Company under a patent granted to H. C. Root. The Root patent covered a postpay box capable of accepting coins of various sizes and used a vibrating bell or buzzer signal to indicate coin deposits to the operator handling the call. For reasons lost in history, no more boxes of this type were installed during the next four years.

By the turn of the century more practical coin boxes became available. A box produced by the Baird Manufacturing Company about this time is shown in Fig. 3-70. This was a postpay box which would accept three sizes of coins. A lever had to be pulled down after each coin deposit to signal the operator that a deposit had been made. It can be noted that the instrument had a carrying handle and may well have been one of several portable devices, connected through jacks installed at each table, which were used in fashionable restaurants of the period. About 1902, coin-operated telephones began to be widely used, particularly in the Chicago area where large numbers were installed in residences. Figures 3-71 and 3-72 illustrate typical single-coin boxes of the period. Until about 1900 most development and manufacture of coin telephones was carried on outside the Bell System. By the turn of the century some 25 companies were manufacturing coin boxes in the United States. One of the earliest and most inventive people in this field was William Gray, who became interested in public coin boxes in the late 1880s[69] and received his first patent in 1889. Later George A. Long joined forces with him

[67] H. Edmunds and C. T. Howard; U. S. Patent No. 327,073; filed May 13, 1885; issued September 29, 1885.

[68] Many of the early coin box schemes were very ingenious but completely impractical. The Edmunds and Howard patent was no exception. It showed an excellent understanding of the functional requirements but the proposed mechanical arrangement was complicated and probably not very practical.

[69] Legend has it that he was inspired by difficulty he experienced in 1888 when he attempted to obtain the use of a nearby subscriber telephone for summoning a doctor during a family emergency. He manufactured the first public coin station (Fig. 3-69) as mentioned earlier.

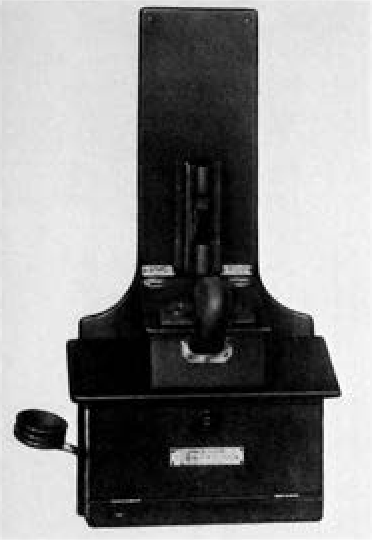


Fig. 3-69. First pay telephone (1889).

Fig. 3-70. Baird pay telephone (circa 1900).

in the Gray Telephone Pay Station Company. Together they were granted a large number of patents. The Gray and Long combination became dominant in the field as they developed a strong patent position covering basic and practical arrangements for handling coins of various denominations so that the boxes could be used for the more expensive toll calls as well as local "nickel" calls.

6.5.3 Basic Coin Box Problems

Of the many problems involved in coin box development two were outstanding. First was the need to inform the operator of the number and denomination of the coin deposited. This was the problem solved so well by the Gray Company. The second was the need for returning coins to the caller if for any reason the call was not completed. This was a less critical problem, since it could be avoided by postpayment, but was essential to the prepay system. Western Electric was instrumental in providing a practical solution to this problem in the early years of the twentieth century.

Gray and other coin box inventors used electrical signals, sounds from buzzers, plucked reeds, and similar arrangements for indicating a coin deposit but when more than one coin was involved, a complicated lever mechanism, operated by the caller, was usually employed.

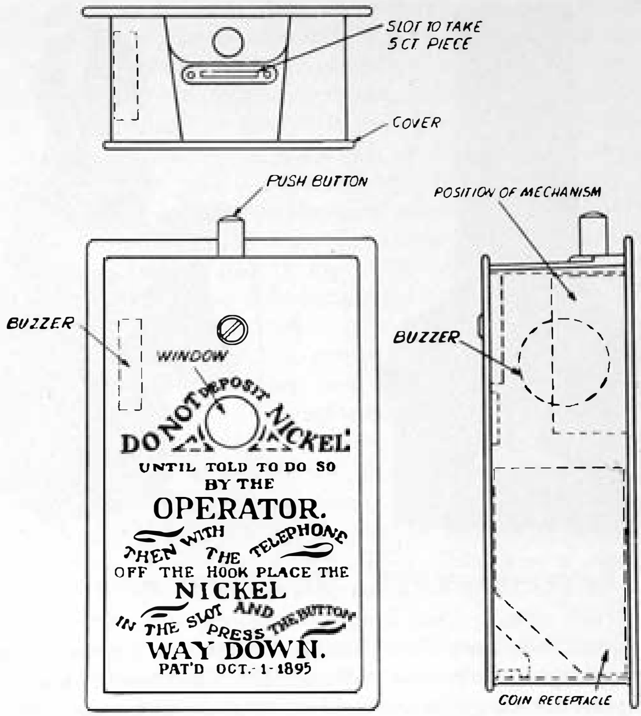


Fig. 3-71. No 1A coin collector.

A typical implementation by Baird is shown in Fig. 3-70, mentioned previously. About 1890, Gray accidentally dropped a coin on a bell and realized that the resultant sound, picked up by the telephone transmitter, could be used as a coin signal using no energy other than that of the falling coin. The basic idea was further developed into a multicoin arrangement which was covered by patents granted to Gray and to C. W. Holbrook in 1892. As implemented, this scheme used a separate slot and coin chute for each denomination. The coins, in passing through the chute, struck one or more gongs generating distinctive signals for each coin deposited. Boxes with as many as five slots were constructed, but the most common arrangement, still in use today, accepted three sizes of coins. The nickel, on its way to the cash box, struck a bell once; a dime hit the bell twice. The quarter followed an entirely different path, striking a so-called "cathedral gong."

Gray's first attempts to apply this scheme were not successful because the bell sounds were not efficiently conducted to the transmitter.

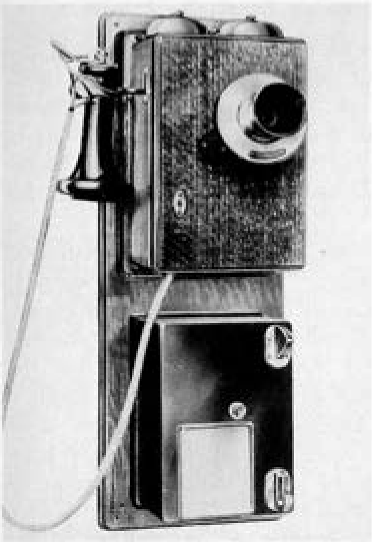


Fig. 3-72. Coin-operated telephone of the early 1900s with No. 7 collector. Fig. 3-73. Early No. 50A Gray/Western Electric coin box. (From *Western Electric News* of February 1916)

After numerous futile experiments involving horns and other means for conveying the sound through an air path, it was found that a rigid connection between the bell mechanism and the transmitter mounting (provided by a common plate or backboard) gave adequate sound transmission.[70] This arrangement, which was referred to as a "resonant connection," was patented in 1897.[71] This was such a simple and obvious way to couple the sound producer to the transmitter that it was widely adopted by other manufacturers and in 1905 the Gray Company started an infringement suit against Baird. The U. S. Circuit Court of Appeals, in a decision handed down in 1909, sustained the Gray patent and gave them control over the most satisfactory way then available for signaling the operator.[72]

[70] One of the schemes tried by Gray was the use of a second transmitter close to the signal bell for the sole purpose of picking up the bell sounds. The scheme was not used at this time (probably because of its complexity) but about 1930 when coin boxes were adapted for handset use, it was revived. With these instruments the coupling between the signal bell and the speech transmitter was very poor, and the supplemental signal was essential and is still used quite successfully.

[71] W. Gray; U. S. Patent No. 593,720; filed November 17, 1893; issued November 16, 1897.

[72] The term "resonant connection" seems to have been used rather loosely to mean a solid (rather than an air) path for the transmission of the bell sound to the trans-

6.5.4 Bell System Developments

The Bell people devoted little effort to coin telephones during the first 15 to 20 years of telephony[73] but beginning in the early 1890s the Boston staff conducted detailed tests on various designs offered by Gray, Baird, and others. A few found application in various parts of the system (in the absence of anything better) but, generally speaking, the attendant or agency schemes served most needs until near the end of the nineteenth century.

The first Bell coin box patent seems to have been granted to G. K. Thompson of the Boston laboratory.[74] This covered a postpay, single-coin box adaptable to either local- or common-battery stations. After depositing a coin, the user operated a lever which collected the coin and signaled the operator, by means of a buzzer, that the deposit had been made. C. E. Scribner of Western Electric applied for a patent on a prepay box in 1896. It was not issued until 1899. Several other prepay patents were issued about this time to Scribner and McBerty, but all covered rather crude arrangements which required action by the user for coin return. The first significant step forward in prepayment boxes came from A. M. Bullard of the Chicago Telephone Company who in 1898 invented a scheme for using a polarized magnet to control the disposition of the coin. A signal of one polarity originated by the operator would return the coin to the user if the call was free or incomplete. The opposite polarity would collect the coin in the cash box. Bullard applied for a patent on this arrangement about a year later. It was issued in 1901[75] and was assigned to the Western Electric Company. Scribner, also in 1901, adapted the principle of the polarized magnet to an improved collect-return mechanism.[76] Both patents were put to use in coin boxes manufactured by Western Electric in the early 1900s and coded Nos. 1 through 7.

---

mitter. There is no evidence that coin box manufacturers attempted to tune the connection so as to enhance the transmission of the bell frequencies. The judicial decision gives Gray credit for being the first to devise a "resonant connection" but seems to interpret this as "the idea of vibratory conduction of the sound signal along the solid wood or metal located in the path of vibration between the signal and the transmitter." This broad interpretation gave Gray an impregnable position.

[73] This was the period when major advances were being made in the new art. The companies were converting their lines to a metallic system using hard-drawn copper wire, common-battery exchanges were being introduced, and the reliable, solid-back transmitter had just become available and was being installed widely. In view of the large amount of engineering effort required by these activities, it is not surprising that the priority placed on coin box development was not high.

[74] G. K. Thompson; U. S. Patent No. 547,405; filed May 13, 1895; issued October 1, 1895.

[75] A. M. Bullard; U. S. Patent No. 665,874; filed June 7, 1899; issued January 15, 1901.

[76] C. E. Scribner; U. S. Patent No. 728,309; filed July 29, 1901; issued May 19, 1903.

The Nos. 1, 2, and 3 boxes were postpay, single-coin (optionally nickel or dime) devices made under the G. K. Thompson patent and were used for a number of years. The No. 1A box is illustrated in Fig. 3-71. Similar collectors were manufactured in the early 1900s by Baird, Gray, and others and used by a number of the Bell Companies. The No. 4 box was referred to as the "Esterbrook Box" and seems to have played an unimportant part in the coin box evolution.

The No. 5 box incorporated the Bullard patent. It was listed in a December 1902 Western Electric catalog. It was superseded almost immediately by the No. 7. However, the Chicago Company used the No. 5 box (probably manufactured locally) as early as 1898 and it seems to have been the first fully automatic prepay box in use in this country.

The No. 7 box, which is shown in Fig. 3-72, used the Bullard principle with the improved Scribner collect mechanism. It was a nickel-only box in which the deposit of a coin lighted a lamp in front of the operator. The coin was held in suspension until the end of the call, at which time it could be collected or refunded by the electromagnet in the box under control of the operator. Some eight or nine versions of the No. 7 box were ultimately produced and the basic design was not discontinued from manufacture until 1931.

In the latter part of 1903 a multicoin box was developed for toll service which used the No. 7 mechanism to initiate calls, further collections being made when the call was established by means of deposits through a second slot leading directly to the cash box. Around 1906 and 1908, modified versions were introduced which kept all coins in suspension until they were collected or refunded by the magnetic mechanism. In some of these multicoin boxes the user operated a lever, after each coin deposit, which struck a gong mechanically to inform the operator of the deposit and the size of the coin; thus these boxes involved a return to a mechanical principle requiring activation by the user.

About 1908, Forsberg of Western Electric began the development of an improved multicoin mechanism which was an extension of the Bullard-Scribner principles. The first model, known as Y-485, was produced in January 1909. In this device all coins were held in a bucket which could be tripped by a polarized magnet, controlled by the operator, so as to collect or return them as required.[77] By this time the Gray mechanism for signaling coin deposits had proven its worth and it was obvious that if this mechanism were combined with

[77] Further improvements were made in successive models and a patent application was made in 1911 and issued (No. 1,043,219) on November 5, 1912, in the form used for many years in the No. 50 box.

Forsberg's, the result would be a truly automatic, multicoin, prepay box. Consideration was given to developing such a box by Western Electric, but the 1909 court decision made it clear that it would not be possible without the use of the Gray patent. It was decided, therefore, to join forces with Gray and on October 18, 1910, an agreement was signed under which Gray would manufacture coin boxes for the Bell System using the Gray chute and box system but with a collection-magnet system supplied by Western Electric. Western Electric also supplied and installed the telephone instrument and station circuitry. This box, shown in Fig. 3-73, was coded the No. 50A in December 1911, and the basic principle involved continued in use for over 50 years. The original Gray agreement ran for about four years but was renewed, with changes as necessary, until the mid-1930s when Western Electric took over manufacture of the entire box.

During the long period in which the No. 50 coin box was manufactured, many changes were introduced. The boxes were adapted to dial operation in 1920. The anti-sidetone (AST) coil and the handset were introduced in the early 1930s. The original cast-lead chutes were replaced with more durable brass and many detailed changes were made to prevent vandalism and fraud from the use of slugs, misappropriation of collections, and so forth.[78] The locks were improved in various ways and in 1916 the box was strengthened by using a pressed-steel lower housing in place of the earlier cast iron. An important change was the introduction of a self-locking sealed coin receptacle. With this arrangement the receptacle was removed from the coin box by the collector and replaced with an empty one. Access to the coins required a key to the receptacle which was available only to the accounting office which supervised collections. In the 1920s this sealed box superseded the older arrangement in which the collector's key gave direct access to the coins.

New code numbers were assigned to boxes incorporating the changes and improvements mentioned above. Many of these were accomplished by changing the letter suffix in the No. 50 series. In other cases the number series was changed. For example, the 150 series was wired for the AST station set, the 160 designation was assigned to replace the 150 for boxes manufactured entirely by Western Electric (about 1935), and the 170 series covered coin boxes using handsets (1940).

[78] It is interesting to note that fraud is not a recent problem. In 1900, Scribner filed for a patent covering a device for detecting iron slugs and sending a warning signal to the operator.

VII. CUSTOMER TELEPHONE ADJUNCTS

Today we have become accustomed to supplementing the telephone station with many devices which increase the usefulness of the service. Message recorders, automatic dialers, complex switching mechanisms, and many other accessories to the telephone station are commonly used on the customer's premises. But during the first 50 years emphasis was on expanding basic telephone service and only a few adjuncts to the customer telephone set were used, notably the telephone booth and some rudimentary switching and metering arrangements.

7.1 The Telephone Booth

The primary purpose of the telephone booth is to provide a degree of privacy to the user of telephones at public locations. However, in the very early days the need to shout, particularly on long-distance calls, caused the user to be rather a nuisance to those around him and booths were sometimes used in business offices to reduce the disturbance as well as to provide privacy.[79]

The pay telephones used on Doolittle's Bridgeport–Black Rock circuit of 1878 were installed in enclosures and this was probably the world's first use of a telephone "booth." But the first booth patent was not filed until 1883. It described a booth 4 to 5 feet square with domed roof and a ventilator. It was mounted on wheels, presumably on the basis that it was a piece of furniture and was sure to be moved about.

By about 1890 a standardized series of five booths became available of which the No. 3 and No. 5 shown in Fig. 3-74 and Fig. 3-75, respectively, were typical. These booths, which were described in an April 1891 AT&TCo brochure, were available either in oak or cherry at a cost between $112 and $225, without "fixtures,"[80] depending on size and number of walls with windows. The Nos. 1, 2, and 3 booths differed only in the window arrangement. They required floor space of about 4 by 5 feet, and were intended for use of the long-distance cabinet

[79] Since Watson's landlady objected to the disturbance he caused with his early telephone experiments, he created a form of "booth" by using a roll of bedclothes to muffle the sound.

[80] It was considered appropriate to provide not only an ornate booth structure but also a decorated interior. Available fixtures were:

Wilton rug	$3.50 and $6.50
Revolving stools with russet leather tops	$2.00
Yellow silk window draperies	$3.00 per pair

By 1891 booths were already in use at such exclusive places as the Astor House in New York, the Grand Union Hotel at Saratoga, and the Union League Club of Chicago.

Fig. 3-74. No. 3 long-distance telephone booth.

Fig. 3-75. No. 5 long-distance telephone booth.

Fig. 3-76. Long-distance cabinet set (1887).

telephone illustrated in Fig. 3-76.[81] The No. 4 and No. 5 type (also differing only in window arrangement) were intended for use with a standard wall set and required only a 4-by-2½ foot floor space. All types had double walls and doors to make them soundproof and also included domed roofs. This type of booth was used extensively between 1890 and 1900 but some booths of simplified construction were also introduced during this period. About 1900 the Bell engineers drafted specifications for a much smaller and simpler booth which was built by outside suppliers on a competitive bid basis. These booths were about 34 inches square, had single walls (with some sound-deadening material added), and were lined with embossed sheet metal. They were the first of the unit construction type which could be assembled into groups. As in previous models, the door was hinged to swing outward and the floor (to provide structural strength) was raised about 4 inches off the building floor. Both of these construction details proved undesirable and were ultimately eliminated. By 1910 the door had been altered so that it could swing inward but it was not until 1930 that the elevated booth floor was replaced by a strong but thin steel plate.

The door problem arose about 1904 when booths were being installed in narrow passageways and odd corners where an outward-swinging door was awkward to use and at times hazardous. Maurice Turner, then associated with the New York Telephone Company, proposed the use of a door, without hinges, running on a curved track that swung the door inward and parallel to the side of the booth as it was opened. After several years of experimentation, the first booths of this type were installed. A considerable number were built in the next few years but the floor track tended to clog with dirt and some users never did master the technique for opening and closing. In 1910, Turner filed a patent for a double-hinged folding door with a simple guide at the top. This door operated easily and with motions which seemed logical to the user. It became the standard door mechanism and has continued in use pretty much ever since. At about the same time the floor space was reduced to 30 by 30 inches which is a common size for indoor booths today. A rather simplified construction gradually evolved so that the booths could easily be assembled into groups. Fronts and exposed end panels were well-finished hardwood but backs and sides used lower-grade material to minimize cost.

About 1912 these features were incorporated in a new standard booth listed as the No. 1 type in the Western Electric catalog and

[81] This was the telephone first used with these booths. In 1891 it was replaced by a similar cabinet phone [item (h), Fig. 3-47] using a solid-back transmitter in place of the "long distance" transmitter used earlier.

illustrated in Fig. 3-77. It was generally used throughout the Bell System for many years but receding-door and swinging-door booths were still available for those wanting lower-cost arrangements.[82] In 1930, the No. 5 and No. 6 type booths were introduced (identical except that a seat was supplied in the latter). These booths introduced the flush floor, improved lighting and ventilation, and numerous other features and accessories (Fig. 3-78).

Most booths were made by outside suppliers until the mid-1920s when Western Electric, in order to provide better and lower cost booths, purchased the Queensboro Shop of the Turner-Armour Company which had been a booth supplier since 1915. This shop became the principal Western Electric plant for the manufacture of booths and continued in operation until 1965.

Although standardization of booths has long been emphasized in order to minimize cost, there have always been a few places which warrant special designs. One of the most famous is the "Pagoda" booth used in San Francisco's Chinatown. The idea proved very popular and was adopted elsewhere. Figure 3-79 shows the manner in which the pagoda theme was used in New York's Chinatown to modify the austerity of an otherwise typical outdoor booth. Other special booths were developed for the rapid transit lines. Those used in the subways of New York were of all-metal construction and were made without doors for sanitary and other reasons. Probably one of the first outdoor installations was used on the elevated railroad platforms where waterproof booths were required. About 75 outdoor booths were used on a special basis in 1939 at the San Francisco Golden Gate Exposition, but widespread usage of the type did not come until later.

7.2 Switching—Key Telephones and Wiring Plans

During the first 50 years of telephony, the switching functions performed by telephone stations (or closely associated equipment) were primarily of two types which can be described broadly as intercommunication and line selection.

The intercommunication or "intercom" system essentially provides a small private telephone system, usually confined to a single building or small area, controlled by a set of keys or switches associated with the station set. A wall set for connection to ten lines, dating from about 1904, is shown in Fig. 3-80. In such a system each station

[82] The earlier use of Codes 1-5 for the booths illustrated in Figs. 3-74 and 3-75 seems to have been overlooked (probably because they were not marketed through the Western Electric Company). At any rate, the numbers were reused beginning about 1912, No. 1 designating the folding-door booth which could be assembled into groups. No. 2 was similar but a single unit with finished wood on all sides, and Nos. 3 and 4 were less-expensive types using receding and swinging doors respectively.

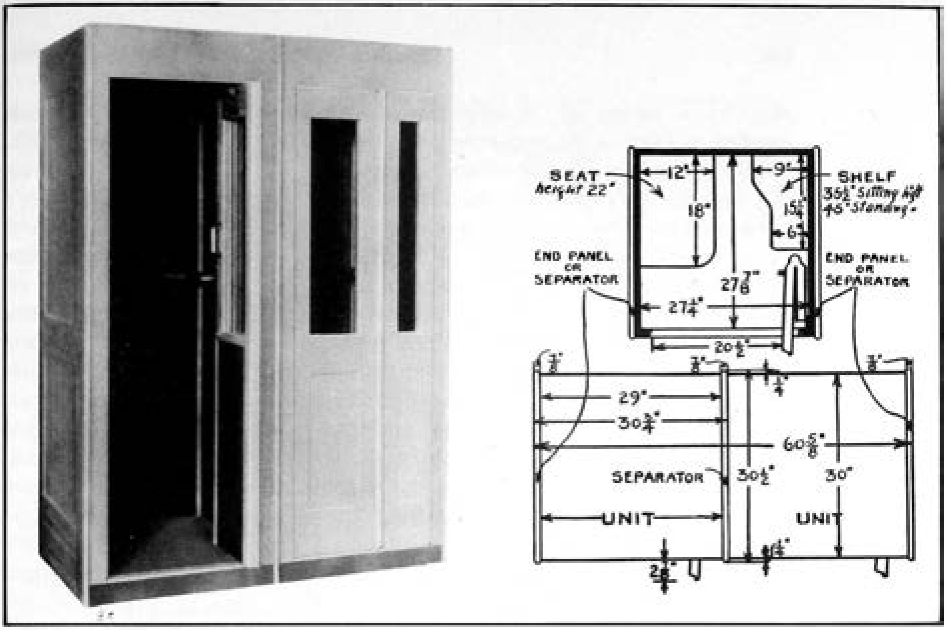


Fig. 3-77. No. 1 folding-door telephone booth (circa 1912).

was normally associated with a specific local cable pair. The pairs appeared at all stations in the system and access to them was by means of the pushkeys shown. To initiate a call the user pushed the key required to connect him to the desired line after which he sent a ringing signal by depressing the ring button. The called party actuated the answer key to connect his station to the line. In such arrangements one of the numbered lines could have been an outgoing line to a central office but in the early days it was common to use separate stations for intercom and network use since the latter required higher-grade station equipment.

The line selection function was used where a single user desired access to several lines to the central office. Commonly, this was accomplished by providing separate station sets for each line[83] but as early as 1892 the station pictured in Fig. 3-81 was provided with a key (directly under the receiver) to transfer the set between two lines. This arrangement seems to have been little used, probably because of the added space required in the base and the necessity for a special station design to meet a limited need. However, the need remained for switching a station among lines and later for other functions such as temporarily "holding" one line while using another, transferring calls from one station to another, and so forth. These needs were originally

[83] Older readers will recall that the movies in the first quarter of the twentieth century frequently pictured the affluent businessman with a battery of telephone stations on his desk.

met by a series of "wiring plans" introduced about 1921 which employed keys of the toggle or pushbutton type as shown in Fig. 3-82. These keys were usually mounted on the side of a desk and wired by the installer according to the standard wiring plan fitting the user's needs. These arrangements became quite common in the 1920s and by about 1930 the demand grew to the point where the more complex No. 100 key equipment for multiple lines was provided. This was an assembly of keys, relays, and signal lamps which had some of the characteristics of a small PBX[84] and provided more complex functions than the simple wiring plan. About 1940, when the combined telephone handset was introduced, keys and pushbuttons were again introduced in the base of the telephone set to control simple, on-premise switching functions. However, during the period covered by this portion of our history, local switching arrangements were largely adjuncts to the telephone station sets rather than an integral part and were controlled by separate key mechanisms wired to meet the user's needs. Mostly a standard wiring plan was used but custom designs were sometimes employed also.

7.3 Metering Arrangements

Around the turn of the century there was much discussion of the most equitable way to charge customers for telephone service. It was

[84] Private branch exchange, discussed in Section VI of Chapter 6.

Fig. 3-78. No. 6 type telephone booth (circa 1930).

Fig. 3-79. Pagoda booth.

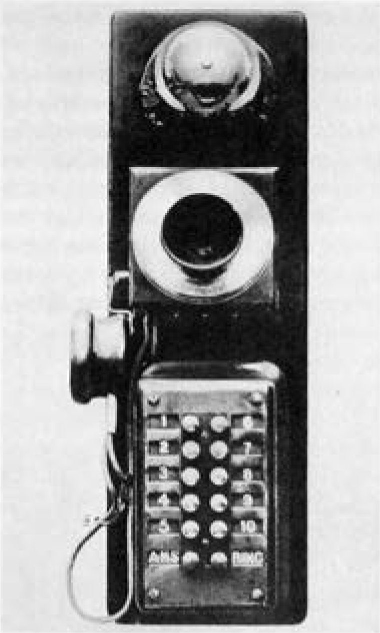


Fig. 3-80. Ten-line intercom set (circa 1904).

Fig. 3-81. Deskstand with line selection key (circa 1892).

generally agreed that long-distance calls should be billed individually according to distance and duration, but local service was ordinarily covered by a subscription fee that was independent of the number of calls made.[85] Many people felt that this was inequitable. Some customers who seldom used the phone objected to paying as much as heavy users did, and many telephone managers felt that as a result of flat-rate service, customers made their phones available to friends and neighbors who would otherwise become subscribers. There was, therefore, a considerable body of opinion in favor of charging, even for local calls, on the basis of the number of calls originated. There was general agreement that charges for incoming calls would be unfair since the recipient had no control over them. Two schemes for measuring service were introduced, both depending on apparatus added to the regular customer's telephone. One called for the use of

[85] This is known as flat-rate service. While the fee was independent of the number of calls per month, it differed with the type of service furnished, i.e., business and residence charges usually differed and the rates were less for party than for individual line service.

meters (message registers) to count the calls, and the other used the coin box.

The original idea was to place the meter on the customer's premises. Some of these devices required the user to press a button or operate a lever to initiate and register a call. They were fundamentally the same as coin boxes, registering a call on a counter instead of collecting a coin. They had the same disadvantages of requiring extra operator time and a visit to the premises to read the meter but the customer was billed monthly and spared the annoyance of having a supply of coins on hand. Hammond Hayes and others came up with counters operated electrically which were partly automatic but either

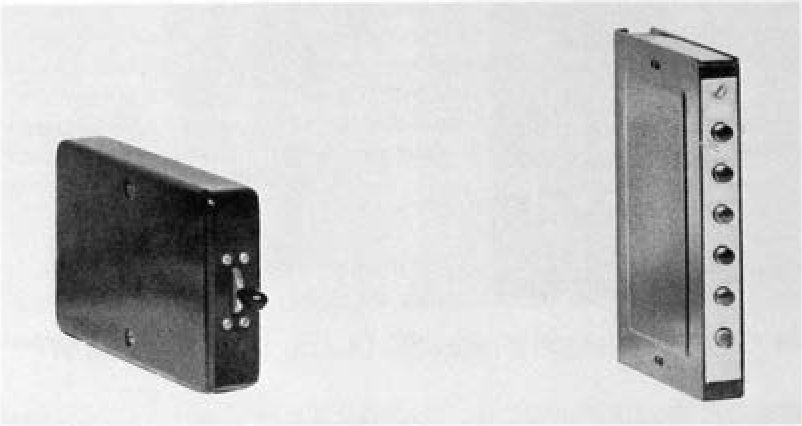


Fig. 3-82. Keys used with wiring plans introduced in 1921.

required complicated mechanisms or operator time to handle free and incomplete calls. This and the need for meter reading made the use of the meter on the customer's premises unsatisfactory and the scheme was little used.

A more satisfactory arrangement was devised by Scribner which used meters on a one-per-customer basis at the central office. This scheme required little extra operator time (it ultimately became completely automatic) and greatly reduced the cost of meter reading. It was the only meter scheme used extensively in the Bell System and has been used mostly in large cities. In such locations measured service is still employed but in small cities flat-rate service has always been preponderant.

The Chicago Telephone Company was the greatest user of coin boxes for providing measured service in the home, but they were also employed in a few other cities, of which Cincinnati was one. At first the idea was outlawed by the Chicago City Council but in 1900 the City Legal Counsel ruled that coin boxes in residences were permitted under the company charter.

Originally the idea was that they would be used more or less as semipublic stations, but the pay-as-you-call idea gained favor and by 1902 they were used widely in residences, 20,000 having been installed during the previous year. This one year's growth was just short of the total stations (25,000) in service during the previous year and explains why the Chicago Company developed great enthusiasm for this type of measured service.

The explosive growth did not last, but the use of residential coin boxes (at user's option) continued in Chicago into the fifties, various kinds of coin boxes being used during this period. Ultimately the collection cost became excessive and the pay-as-you-call system gave way to the monthly billing system in use elsewhere.[86]

Thus the idea of charging on the basis of originated calls started with the use of a station adjunct, but such arrangements were not used extensively. The basic idea was survived but metering has become one of the functions performed at the central office under control of the switching mechanism.

VIII. TELEPHONES FOR OTHER USES

In addition to the customer telephones that have been described, a number of other types evolved. Some of these, such as operator's and lineman's telephones, were used directly in the operation and maintenance of the national telephone network. Others, such as the high-quality instruments used for laboratory and comparison purposes, indirectly influenced the evolution of the network. Still others, while

[86] One of the unusual features of the Chicago coin service was that the use of slugs in home phones became quite common at one time. The collector on his monthly visit would separate out the slugs from the nickels and present them to the phone user who would redeem them in legal tender. The wide use of nickel slugs on a private basis obviously presented complications in the use of public telephones and other coin-operated devices and resulted in the use of special tokens of complicated design in place of nickels. These complications provided added reasons for discontinuing the home coin box but many users were reluctant to give them up, especially when use was shared with neighbors. In Chicago, residential coin boxes were discontinued as a new service offering at the end of 1937, but existing subscribers to this service were allowed to continue. Even though equivalent message rate service at the same cost had been offered as a replacement, it was not until 1958 that the number of coin boxes declined to about 1,000 and the service could be discontinued. Cincinnati had discontinued residential coin service about ten years earlier.

by-products of major telephone development programs, had special or limited applications with little relation to the network.

8.1 Operator's Telephone Sets

The need for an operator's set arose with the beginning of switching. Electrically the requirements were essentially the same as for customer sets except for the greater need for low sidetone brought about by the almost continuous usage, often in rather noisy surroundings.[87] Physically, the set had to be designed so that it was comfortable to use for long periods, left the operator free to move to a limited extent, and also gave her the use of both hands for manipulating the plugs and keys used in switching. This was a rather large order and was not completely achieved for a number of years.

The first operator's telephone sets were hand-held leaving only a single hand for switching. The first switchboard installed at New Haven, Connecticut, in 1878 provided a butterstamp instrument for the operator which served as both transmitter and receiver. The Gold and Stock Exchange in 1879 used handsets of the type designed by Brown which have been mentioned previously and illustrated in Fig. 3-54. Another operator's set introduced about this time used a Blake transmitter suspended by an arm in front of the operator and a hand receiver of the butterstamp type held by the operator.

In order to free both hands of the operator, E. T. Gilliland in 1881 devised a 6-pound arrangement to support a Blake transmitter and hand receiver on the operator's shoulder as shown in Fig. 3-83a. This equipment, because of the great weight, was unsuitable for use by the young women who were replacing the men operators used previously and was not put into widespread use. Instead, an arrangement was tried in which both the transmitter and the receiver were suspended by an arm as shown in Fig. 3-83b. This was more comfortable but was difficult to use since the operator was obliged to remain close to the instruments if the set was to perform efficiently. Obviously, none of these arrangements was satisfactory but with the development of the small receiver by Richards in 1884 a practical scheme began to evolve. As discussed in Section 2.2.3, this receiver was light enough to be suspended by a headband which held it tightly to the ear. The headband, originally of the double type, was simplified as time went on and by 1920 a very lightweight wire device was employed.

At first the Richards head receiver was used with a suspended or bracket-mounted transmitter as shown in Figs. 3-83c and 3-83d, but in

[87] As discussed in Section 3.1.

1900 the No. 234 breastplate transmitter (Fig. 3-83e) was introduced and, combined with the new No. 128 receiver standardized the same year, produced for the first time an operator set meeting the basic requirements. It continued in use with some modifications and improvements for the next 40 years. Its 1900 form is illustrated in Fig. 3-83f, and a simplified form introduced about 1920 (with the No. 528 receiver) is shown in Fig. 3-83g. This set had many features which contributed to its long use. From a technical standpoint it was highly satisfactory since it was based primarily on instrument designs similar to those manufactured in large numbers for subscribers. It was easily adjusted by means of a mouthpiece with ball and socket joint and a neckband of adjustable length. It was well liked by the operators since it was light in weight (less than 1 pound total) and easy to keep hygienic since the neck straps were readily and inexpensively replaced and the mouthpiece was removable for washing.[88]

8.2 Sets for Linemen and Craftsmen

During the construction and maintenance of telephone plant, craftsmen frequently had need to talk to, and conduct tests with, the wire chief or other supervisory people in the central office.

To meet this need, particularly in the case of outside-plant workers, two general types of lineman's sets were designed in the 1889 to 1910 period. For use on lines with local-battery service, the sets illustrated in Fig. 3-84 were developed. Each was a complete telephone with receiver, transmitter, hand generator, and a buzzer that responded to incoming calls. A small local battery with a push-to-talk button was provided. For use on common-battery lines, a handset, mentioned earlier and illustrated in Fig. 3-57, was designed. Since neither battery nor the ringing generator was needed, this set comprised only a carbon transmitter and a receiver mounted on an aluminum handle and wired in series. Because its use was restricted to trained telephone company personnel and to relatively short connections, it did not need to meet the requirements of commercial customer equipment.

The wire chief used standard instruments of the type designed for operators. Since he had to be free to leave his desk, the suspended transmitter of the type shown in Fig. 3-83c was used at the wire chief's position long after it was superseded by the breastplate transmitter for operators. This arrangement, particularly with the receiver and headband of Fig. 3-83g, provided a flexible setup since the small

[88] To further promote hygiene, each operator was assigned a telephone set which she alone used.

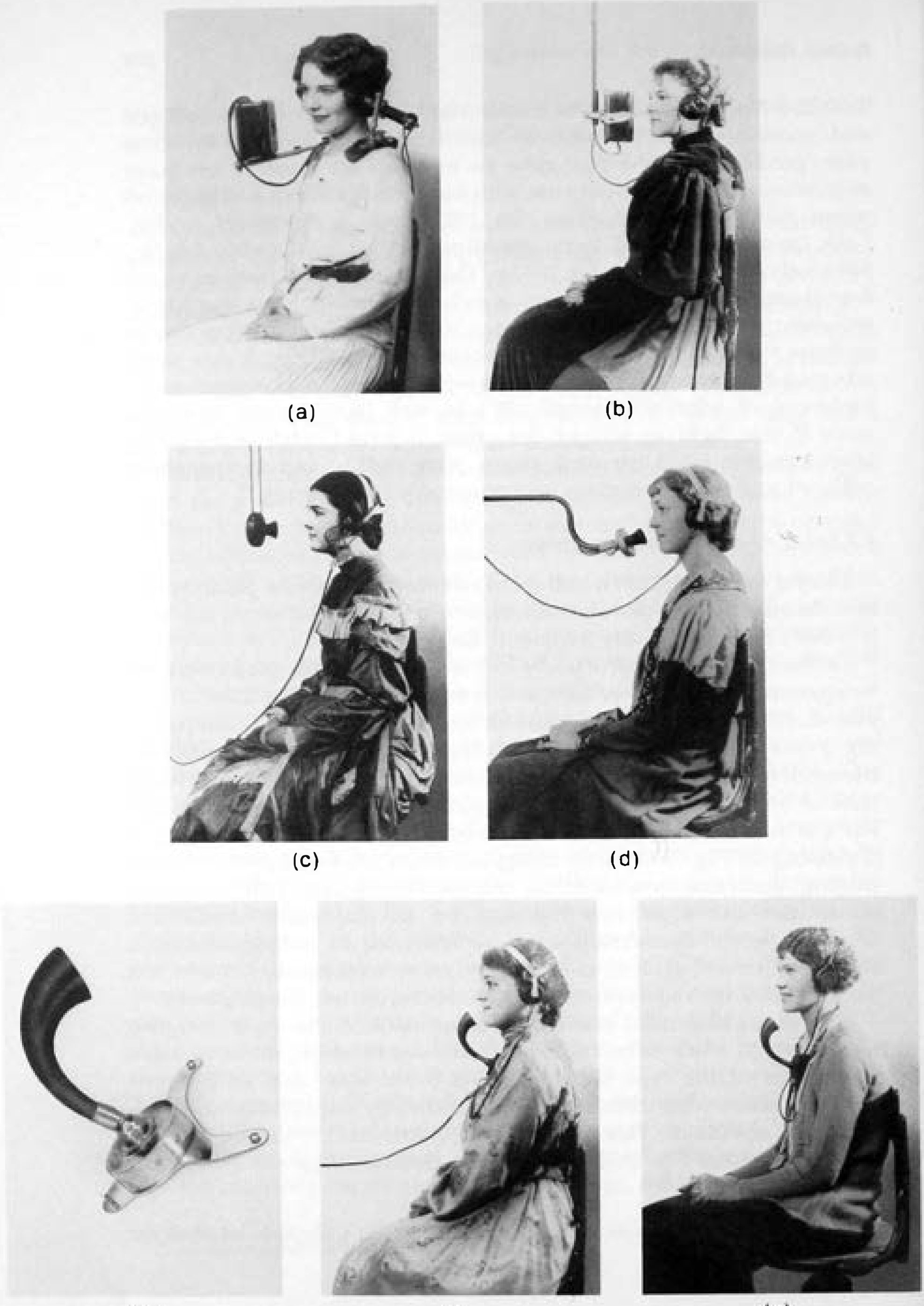


Fig. 3-83. Operator's telephones of 1881–1925.

receiver could be hand-held for short conversations or head-supported for a long series of tests.

The central-office craftsman originally used much the same equipment as the lineman but by 1912 the frame room had become quite large in some buildings. In order to provide greater flexibility for craftsmen, the loudspeaker systems mentioned in Chapter 4 (Section 4.3.1) were developed for communication with the wire chief.

8.3 Special-Purpose Telephones

There was an obvious need for various kinds of telephone communication not provided by the commercial network. A number of these needs were met by developing station sets to meet special service requirements. Intercom sets, already mentioned, are an example of a service originally separated from but later integrated with the commercial network. Other services with special requirements remained independent of the network but were able to take advantage of the technology developed for commercial telephony.

Telephone sets encased in iron boxes, designed to be mounted in outdoor locations on poles and buildings, were provided early in the twentieth century. Two such designs are shown in Fig. 3-85. Item (a) was produced in 1903 for street railway systems and item (b) was a policeman's call set of the same period.

Special telephones were also produced for use in mines where dampness and explosive gases were prevalent. An early design, of about 1892, is shown in Fig. 3-86a. The sealed box contained both the transmitter and the receiver with voice access through speaking tubes of rubber. All metal parts were lead plated to prevent rusting. In the 1910 version shown in Fig. 3-86b, all parts with electrical contacts were enclosed to isolate sparks and all metal parts were finished with japan or lacquer to protect against moisture damage.

Equipment to provide communication between divers and their tenders was supplied as early as 1892.

Telephone sets to meet the special needs of the military were produced beginning about 1900, including the handsets mentioned in Section 6.3. During World War I other special communication systems, including the necessary station apparatus, were designed for the military. Some of the more interesting were shipboard systems and means for communicating with captive balloons and airplanes. The U. S. Navy began the extensive use of shipboard telephones about 1909 when the "Great White Fleet" made its round-the-world cruise. By 1916, large ships were equipped with two extensive telephone systems, one for operational use and the other for fire control (see Fig. 3-87). Both of these systems involved unusual problems in protecting the equipment from moisture and physical damage during battle conditions. The development of ship-to-ship and other wireless communication systems is covered in Section 4.1 of Chapter 5. One of the interesting problems of airborne systems involved the development of transmitters and receivers (the latter built into aviators' helmets) for use in the presence of the airplane's high ambient noise.

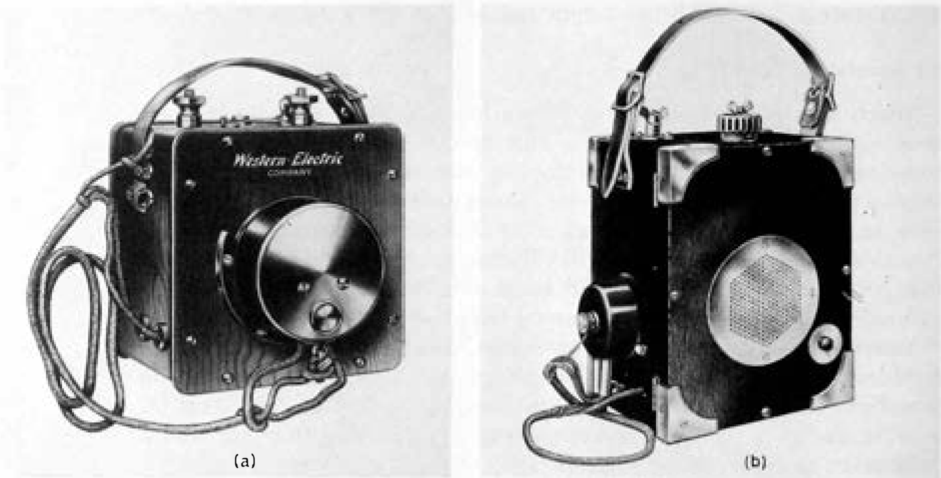


Fig. 3-84. Lineman's test sets: (a) 1889; (b) 1905.

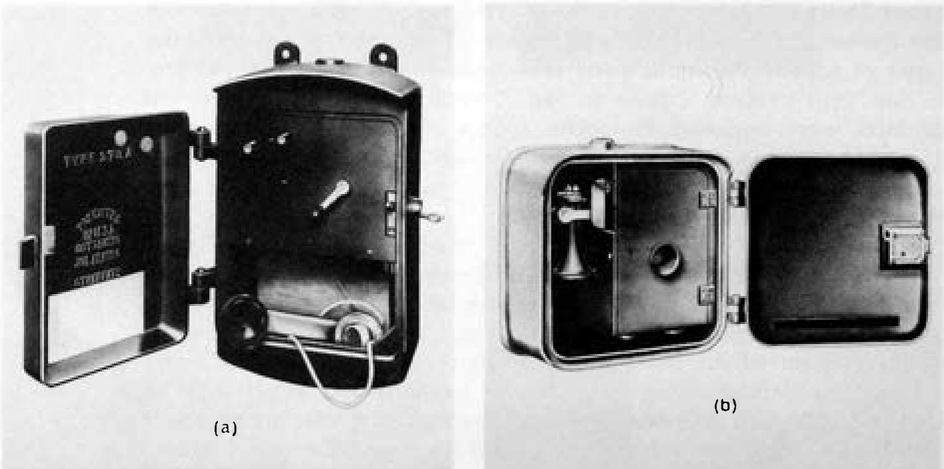


Fig. 3-85. Telephone sets encased in iron boxes for outdoor use. (a) Pole-mounted telephone set for street railway maintenance personnel. (b) Pole-mounted telephone set for police.

A by-product of the airplane system was telephonic communication between the pilot and his passengers. Previously communication had been accomplished by hand signals.

Thus, by the end of World War I, telephone sets had been designed for indoor and outdoor use on the earth, for use below the earth's surface, and in the air above.

8.4 High-Quality Electroacoustic Converters

The development of high-quality telephone instruments is a particularly interesting illustration of the research and development process. These instruments did not evolve directly, as might be expected, from the flow of technology which produced commercial telephone transmitters and receivers. Instead, their development drew largely on rather remote supporting technology which had been developed for quite different purposes. To complete the illustration of the serendipitous growth of technology, it turned out that these instruments, designed to meet a particular objective, found a much broader application in fields unthought of when their development was begun.

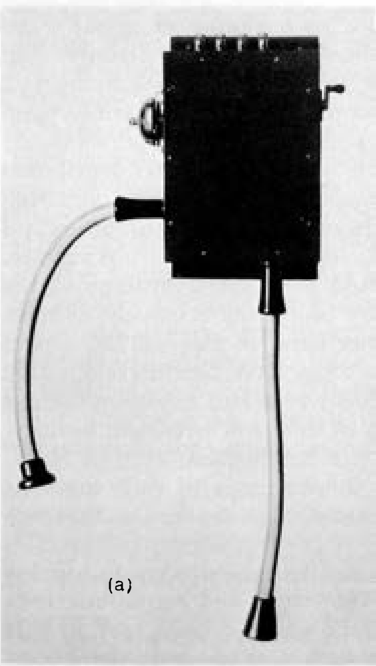
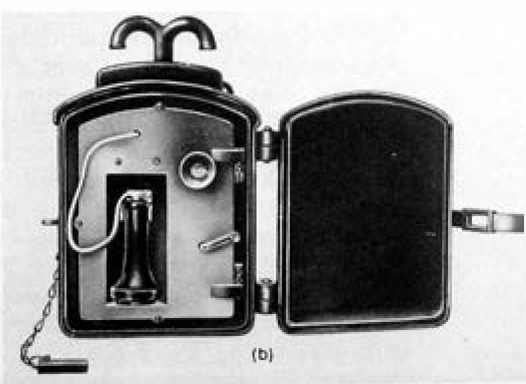


Fig. 3-86. Telephone sets for use in mines: (a) 1892 design; (b) 1910 design.



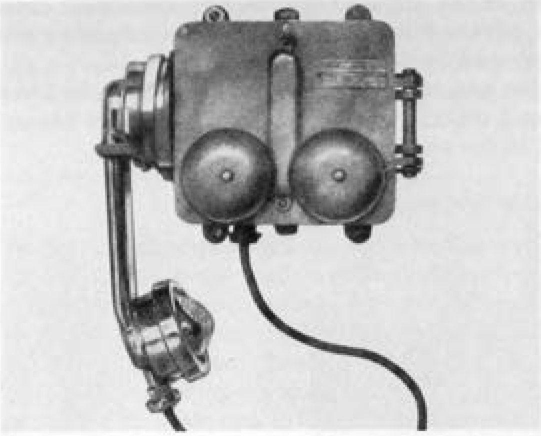


Fig. 3-87. Ship service telephone on board U.S.S. *Arizona*.

It will be recalled that in the early years of telephony, when the telephone transmitter was the only source of amplification, the most effective telephone instrument was one which provided high efficiency for a rather narrow band of frequencies centered at about 1,000 hertz. The speech from such instruments was highly distorted but still reasonably intelligible and the high efficiency attained by using a resonant structure which concentrated the energy in a narrow band was responsible for achieving transmission over distances which would have been completely unattainable if the speech energy was distributed over a wider frequency range. It was obvious that this situation would change as line amplification became available. Transmitters and receivers would no longer have to be designed to achieve maximum transmission distance. Instead, it would be possible to improve the ease and quality of communication by producing relatively uniform electroacoustic conversion over a wider band of frequencies. Useful designs of higher-quality instruments could have been developed by an extension of the experimental process which had already produced the reliable transmitters and receivers of the early twentieth century. But some of the technical leaders in the Bell System, such as H. D. Arnold and I. B. Crandall, saw the shortcomings of such methods and the advantage of developing a broadly based theoretical approach to acoustical design. Accordingly, they originated, shortly before 1915, a comprehensive study of the speech and hearing process to provide the necessary technical foundation. This work and its unique contribution to the art of acoustics will be covered elsewhere in this history. Of significance here is that such a study required highly sophisticated instrumentation including transmitters and receivers of great stability having a frequency response uniform over a band of frequencies far beyond anything needed for telephone speech. With such instruments, and appropriate equipment for measuring and generating electrical tones, it would be possible to explore the amplitude-frequency characteristics of speech and acoustic noise and also determine the sensitivity of the hearing process. The instruments, together with suitable connecting circuitry, would provide an optimum telephone transmission circuit against which other transmission paths could be compared and, by inserting transmission degradation in this path, the effects of noise and distortion on speech transmission could be evaluated.[89] Obviously, the key to this whole plan was the development of suitable high-quality transmitters and receivers.

8.4.1 Transmitters

The microphonic principle which formed the basis for all commercial transmitters was completely unsuitable for constructing the high-quality instruments needed for basic acoustic studies. These instruments lacked the necessary stability and also produced more harmonic distortion than permissible. In addition, the carbon button affected the motion of the diaphragm in a way that was difficult to predict with exactitude because of carbon packing. Instead, E. C. Wente, to whom the development problem was assigned, selected the condenser (capacitor) principle which had been discovered as early as 1870. At that time Varley had noted that a condenser could emit sound (due to movement of the plates) when an electromotive force was applied. Proposals for using this principle in the construction of a "telephone" were made shortly after Bell's invention by Edison, Du Marcel, Dolbear, and others.[90] These proposals covered primarily the receiver application but, since it was a reciprocal device, it could also be used as a transmitter because sound impinging on a condenser plate would vary capacitance and this variation could be used to generate an electromotive force.

While the simplicity of the concept was attractive, the condenser transmitter was so high in impedance and so low in efficiency, compared to the carbon microphone, that it was of no practical use

[89] In essence, this was the philosophical basis for the Master Transmission Reference System developed about 1928 and discussed in Section 5.1.1 of Chapter 4.

[90] Edison seems to have priority with two designs for condenser-type telephones which he tried in 1877. Apparently he did not follow up on these proposals. About 1881–82, Professor A. E. Dolbear strongly advocated the use of a condenser-type receiver and while it appeared to have certain advantages it could not, in fact, compete with electromagnetic devices, which were more efficient and more suitable because of their lower impedance, for use with available transmission systems.

during the 45 years after Varley discovered the principle. By 1915, Wente had two new tools which he could use. First was the de Forest audion. Although originally used as a radio detector, the improvements made by Arnold and others had turned it into a practical amplifier capable of solving the line repeater problem. Advances in tube construction and circuitry had come along so rapidly that Wente was convinced that electronic amplifiers could be used to convert the condenser-microphone voltage and impedance to the extent required for practical application.[91] The other tool was the new theory growing out of G. A. Campbell's work on loading, wave filters, and other electrical networks. Crandall, Wegel, Wente, and others began applying this theory to electroacoustic systems and showed how electrical networks could be devised that were the exact analog of vibrating systems and thus electroacoustic converters could be represented as electrical systems the performance of which could be computed by means of electrical network theory. Thus the characteristics of the transmitter could be computed in advance and the pertinent parameters varied to give the desired overall frequency response.[92]

The Wente condenser microphone of 1917 is shown in cross section in Fig. 3-88. It consisted essentially of two parallel plates insulated

[91] The technically inclined reader may be interested in a more detailed discussion of the problems involved in the practical application of the condenser principle. The device was not only far less sensitive than the familiar carbon transmitter but was also somewhat less sensitive than electromagnetic devices. However, for the acoustic levels of interest the output could be amplified to adequate levels by electronic amplifiers if the input could be kept reasonably free from noise arising either in the amplifier circuitry or through coupling to outside sources. Perhaps more serious than low sensitivity was the high and frequency-dependent impedance of the device. Since the transmitter was nothing but a small capacitor, the impedance was reactive, being very high at low frequencies and decreasing in an inverse manner with frequency (a typical commercial transmitter, the 394, varied from about 20 megohms at 25 hertz to 0.1 megohm at 5,000 hertz). The high impedance meant that even short lengths of cabling would introduce high loss and noise from crosstalk. The variable impedance meant that the transmitter voltage, which was reasonably independent of frequency, would be badly distorted unless it worked into an impedance many times higher than the microphone impedance. These problems would have been insoluble without the vacuum tube amplifier. The input impedance of these devices was inherently high and by placing a shielded first stage (preamplifier) close to the microphone the high-impedance wiring could be kept short and noice pick-up made negligible. This all sounds easy today but in the twenties there was little experience in producing quiet amplifiers with high and stable input impedances. Even the production of reliable high-value resistors was without precedent and the two 50-megohm resistors required in each preamplifier were, after much investigation, made by filling glass tubes with a mixture of Xylene and alcohol, contact with the mixture being made by platinum wires sealed into the ends of the containers. It was also necessary to use batteries to supply the transmitter bias and the plate and filament supplies of the preamplifier since other practical, low-noise power sources were not available.

[92] These two tools, the electron tube and electrical network theory, mentioned briefly above, are discussed in more depth elsewhere in this history in connection with subjects to which they are more directly related. See particularly Chapters 4 and 5 which not only provide more background on their origin and development but also cover their revolutionary effect on the evolution of wire and radio transmission.

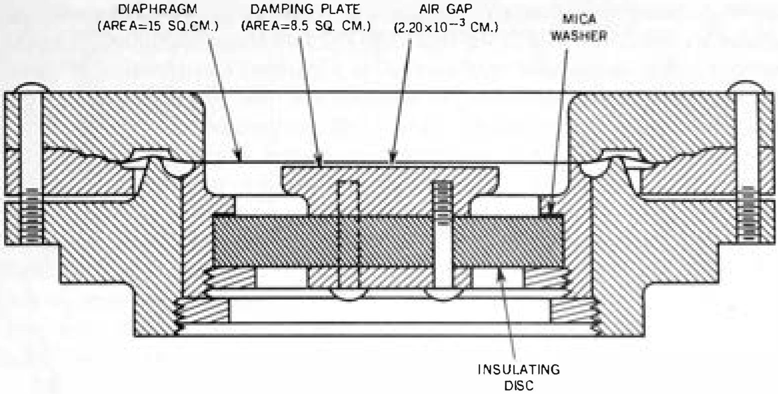


Fig. 3-88. The Wente condenser microphone of 1917.

from each other. One plate was fixed and the other, a diaphragm of thin sheet steel, was movable under the action of the alternating pressure of a sound wave. These plates were connected in series with a resistor and a battery; as the capacitance varied, the current flowing in the circuit also varied so that a varying potential was produced across the resistor which was proportional to the undulations of pressure in the sound wave. Distortion in the frequency range of interest was avoided by stretching the diaphragm so that its resonant period was at about 17,000 hertz. Such distortion also was avoided by the damping action of the thin film of air between the two plates. The design was improved by Crandall in 1918 by using a slotted back plate which increased the damping at the lower frequencies permitting the reduction of the resonant frequency and consequently increasing efficiency below this point. Research instruments built by Wente using this principle had a resonant period of about 10,000 hertz and an approximately uniform response with frequency from 20 hertz to 8,000 hertz.[93] They produced an open-circuit voltage of about 0.3 millivolt for a pressure of 1 dyne per square centimeter.[94] Commercial transmitters developed in the late twenties were 20 dB more sensitive.

In order to provide a useful output, a multistage amplifier was associated with the condenser microphone, the potential variation across the resistor in the microphone circuit being applied to the grid of the first stage. In one interesting application, where the transmitter was used in making oscillographic recordings of speech sounds, the

[93] The shape of the response in practical applications was, at the extreme frequencies, considerably influenced by the amplifiers used with the microphone.

[94] In the circuit used, the open-circuit voltage was essentially the operating voltage.

amplifier consisted of seven resistance-coupled stages having a voltage gain of 40,000 (over 90 dB). In this case the last stage consisted of eight vacuum tubes in parallel and served as a current transformer to drive the low-impedence oscillograph without the use of coil transformers which would have introduced a low-frequency cutoff.

Another interesting research application was the use of a probe, or search, tube, about six inches long and of one-quarter-inch or less diameter open at one end and coupled to the condenser transmitter at the other. The open end of the tube could be used to explore pressures in small chambers at frequencies up to about 25 kHz since the small size caused negligible wave distortion. The transfer of pressure to the microphone was affected by longitudinal resonances in the tube and consequently the response was not flat with frequency. As a result, the system had to be calibrated and used on a single-frequency basis.

Although developed primarily for research, the condenser transmitter found many other applications. Its development was completed just as public-address and radio broadcasting systems began to evolve, as will be related in Chapters 4 and 5, and for a number of years it was widely used commercially where the very best quality of speech transmission was required. The microphone developed for this purpose about 1928 was coded No. 394 and was associated with the No. 47A preamplifier, as shown in Fig. 3-89a. The amplifier was a single-tube device, shown schematically in Fig. 3-89b, which provided the 200-volt bias for the microphone and converted the 50-megohm grid input impedance to a low output impedance (50 or 200 ohms) for use with local cabling without undue loss or distortion. Later a small-diameter microphone and reduced-size amplifier were developed.

8.4.2 *Receivers*

Initially, high-quality receivers were obtained by modifying the bipolar telephone structure to provide greater stability under temperature variations and to damp out the resonant peaks. The former was accomplished by means of a diaphragm clamping ring having the same temperature coefficient of expansion as the diaphragm. Damping was provided by a loose pile of thin paper rings, treated to exclude moisture, behind the diaphragm. The air layers between the paper leaves and between the topmost leaf and diaphragm damped out most of the resonant peak.[95]

The response of these receivers was substantially uniform with frequency over a range of 20 to 4,000 hertz. These instruments, because of the damping, were far less efficient than commercial receivers but with

[95] A pile-up of paper rings was later used in handset transmitters to provide a loose but damped diaphragm support.

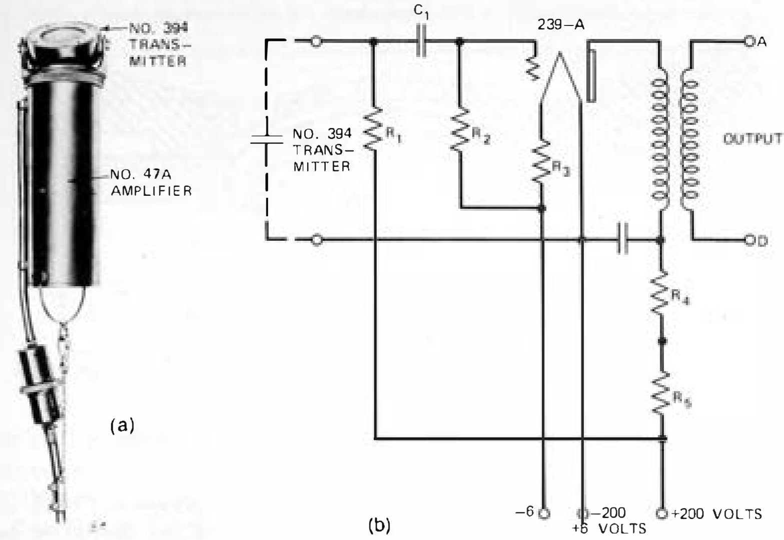


Fig. 3-89. Commercial condenser microphone (circa 1928). (a) Transmitter/amplifier assembly. (b) Schematic diagram of the No. 47A amplifier. (Diagram redrawn from Curl 1928, Fig. 1)

amplification they were suitable for limited speech testing in laboratories. When calibrated in terms of pressure output per unit current and driven by a suitable oscillator, they also could be used to supply known pressure levels to the ear and thus explore ear sensitivity.

These early receivers were too limited in high frequency response to cover the full range of speech sounds and fell far short of the spectral capabilities of the normal ear. This situation was corrected by an improved receiver developed by A. L. Thuras and E. C. Wente in the late 1920s. Again, the principle used was an old one, proposed as early as December 14, 1877, by Siemens and Halske.[96] But, as with the condenser transmitter, the old idea was made practical by modern techniques. The basic scheme is illustrated by the cross-sectional view of the receiver shown in Fig. 3-90. The diaphragm is dome-shaped to provide rigidity and is flexibly supported at the edges so that it can move more or less as a piston. Attached to this is a self-supporting coil, as shown in detail in Fig. 3-91, wound of aluminum ribbon held together and insulated by a film of lacquer. The coil is centered in the annular space between the pole pieces of the strong permanent magnet forming the shell of the receiver. In contrast to the bipolar structure

[96] German Patent No. 2355.

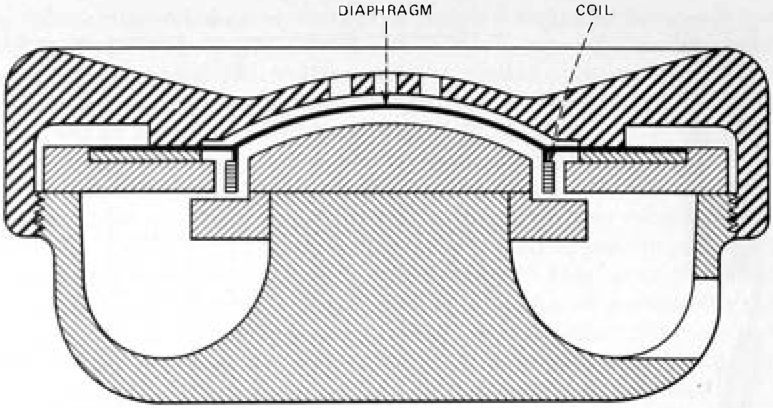


Fig. 3-90. Moving-coil head receiver. (Redrawn from Wente and Thuras 1931, Fig. 4)

used in commercial receivers, no biasing force is employed and an alternating voltage applied to the coil causes it to be displaced in either direction from the at-rest position depending on polarity and magnitude. Due to the absence of bias, the second harmonic of the applied voltage is not present; thus the main distortion present in commercial receivers is eliminated. Using network analysis techniques, the mass and stiffness of the diaphragm, together with the acoustical impedances of the various air chambers and passages, were proportioned to give a reasonably flat frequency response up to about 9 kHz as shown in Fig. 3-92. These responses were measured on a closed coupler and are a measure of the capabilities of the mechanism.

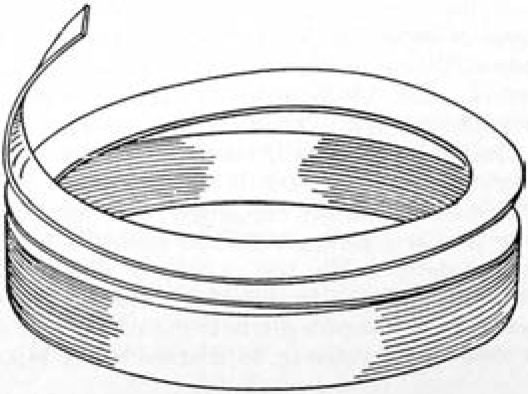


Fig. 3-91. Receiver driving coil. (Redrawn from Wente and Thuras 1928, Fig. 4)

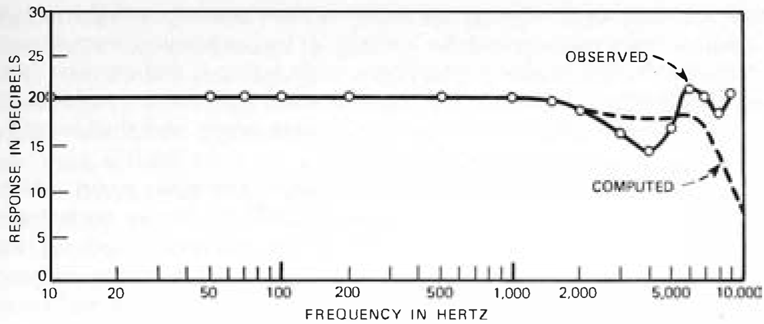


Fig. 3-92. Response of moving-coil receiver. (Redrawn from Wente and Thuras 1931, Fig. 6)

When used on the human ear, the lower-frequency response is affected by ear characteristics and particularly the leakage path between the ear and the receiver cap. The difference between the observed and computed curves shown in the figure is probably due to the fact that some of the quantities used in the calculations are not strictly constant up to the higher frequencies where the diameter of the diaphragm becomes comparable to the wavelength of sound.

One of the first telephone applications of these receivers was in the 1928 Master Transmission Reference System discussed in Chapter 4 (Section 5.1.1) but the moving-coil principle had been used in loudspeakers designed by Wente a few years earlier.[97]

8.4.3 Non-Telephone Transducers

By the late 1920s a substantial demand was developing for high-quality instruments for non-telephone applications such as broadcasting and talking movies. These outgrowths of the telephone are discussed in considerable detail in other parts of this history but a brief review of the instrument development will be appropriate to show the manner in which it evolved from telephone and speech research techniques.

Broadcasting to large audiences began with the public-address systems of the 1910–1920 period. At the beginning, vacuum tube amplification was not available and transmitter and loudspeaker developments were aimed at achieving the maximum practical sound output. Transmitters were of the carbon type designed to modulate

[97] Wente had applied for a moving-coil loudspeaker patent in August 1926 and No. 1,707,545 was issued April 2, 1929. The receiver patent, No. 1,766,473, was filed in 1928 and issued June 24, 1930.

large amounts of direct-current power. Water cooling and the use of double buttons were two of the techniques employed and were logical extensions of the existing telephonic transmitter. Loudspeakers also used basically telephone-type transducers with horns added to provide a better impedance match with free space and thus achieve more efficient electroacoustic conversion.

About 1917–18 the electronic amplifier brought about major changes. The development of the high-quality condenser transmitter has already been related and highly practical models such as the 394-type were available when radio broadcasting and talking pictures brought the need for stable and reliable high-quality microphones. Although these instruments were widely used wherever the highest quality of reproduction was required, there was also a need for a simpler, more efficient device for those applications where some reduction in frequency band and stability could be tolerated. This need was filled in 1921 by a carbon transmitter developed by W. C. Jones. This transmitter, shown in Fig. 3-93, used a stretched, air-damped diaphragm with a carbon button on each side, the two buttons being connected in a push-pull arrangement with an associated split-winding transformer so as to reduce the production of

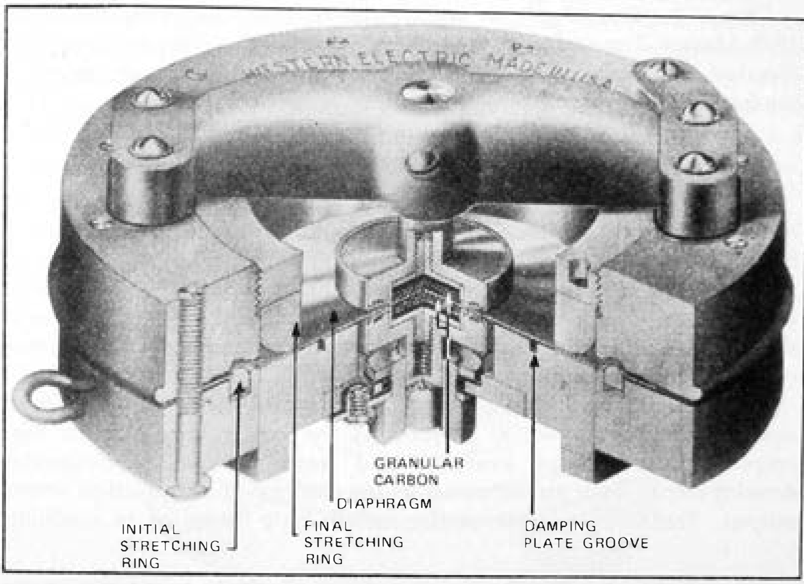


Fig. 3-93. No. 387 double-button, high-quality carbon microphone.

even-order harmonics.[98] Thus, this transmitter combined the techniques of the commercial telephone transmitter and the theoretical design methods underlying the condenser transmitter. This carbon transmitter, which was coded the No. 387 type, was suitable for many radio broadcast applications and efficient enough to reduce the required amplification by some three or four stages.

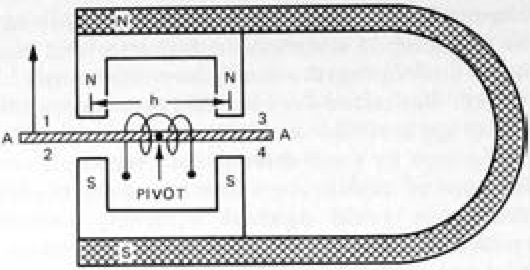
The development of loudspeakers also was influenced by electronic amplifiers since they could produce sufficient power to drive a broadband speaker at adequate acoustic levels. In 1918, as high-power amplifiers became available, a new type of converter using the balanced-armature principle was developed. This device, illustrated schematically in Fig. 3-94a, had a pivoted armature, AA, within a driving coil, h, mounted between the poles of a permanent magnet in such a way that the magnetic forces on the armature were in balance until the coil was energized. Since the armature was not permanently biased, it was free from second-order distortion. This structure could also be built to produce large excursions of the diaphragm and thus handle several watts of acoustic power in loudspeakers designed for reasonably flat frequency responses. The basic idea was not new. Thomas Watson had applied for a patent on a balanced-armature telephone in 1880[99] and Siemens, among others, had previously used pivoted-armature devices. During the early part of the twentieth century the balanced-armature principle received limited use in the construction of very efficient head receivers much favored by radio amateurs. In 1918, Henry Egerton of Western Electric devised a highly rugged balanced-armature unit capable of handling enough power for driving a loudspeaker.[100] The first application of this unit was in a horn-type speaker but in 1924 it was used in a new, cone-type structure based on developments by Ricker and Wegel.[101] This device, known as the No. 540AW loud-speaking telephone and illustrated in Fig. 3-94b, was widely used with home broadcast receivers and as a monitor in broadcast control rooms. Because of the 18-inch conical

[98] A double-button microphone had been constructed in 1910 in connection with early work on public-address systems. At that time the objective was to obtain increased power output, rather than high quality, since electron tube amplification was not yet available.

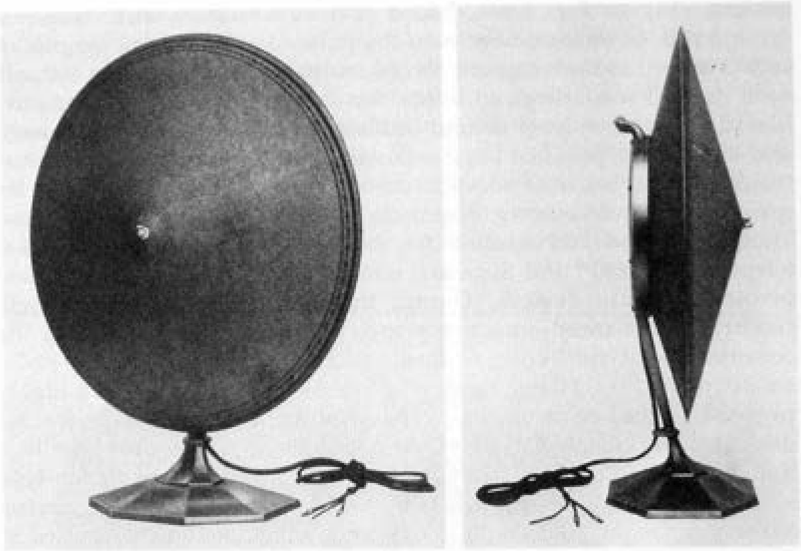
[99] U. S. Patent No. 266,567 was granted to him on October 24, 1882. His device differed from that shown in Fig. 3-94a in that the driving coil was wound on the armature instead of having a light armature moving within a fixed coil. This latter scheme was invested by Frank Capps in 1890.

[100] H. Egerton; U. S. Patent No. 1,365,898; filed January 8, 1918; issued January 18, 1921.

[101] N. H. Ricker; U. S. Patent No. 1,859,892; filed October 6, 1922; issued May 24, 1932. R. L. Wegel; U. S. Patent No. 1,926,882; filed October 17, 1929; issued September 12, 1933.



(a)



(b)

Fig. 3-94. (a) Schematic diagram of balanced-armature driving unit used in early horn and cone loudspeakers. (b) Front and side views of No. 540AW loud-speaking telephone. (Diagram redrawn from Hunt 1954, Fig. 7.1c; photos from Western Electric Bulletin T-750 of 1924)

diaphragm, matched to a similar conical support, this loudspeaker radiated frequencies much lower than obtainable with horns of reasonable size. In addition, its compactness and freedom from directivity made it ideal for home use and it was largely responsible for educating the broadcast audience to the benefits of higher-quality broadcasting. A few years later a 36-inch speaker known as the 548-type was introduced to provide even better low-frequency reproduction.

The cones and the horns using the balanced-armature driver met the needs of the public-address and broadcast audience but with the introduction of talking motion pictures in 1926 it became apparent that it would be necessary to fill large auditoriums with sound covering an extremely large range of frequencies.

It was this need that started Wente on a new approach which ushered in the "moving coil" era of electroacoustic converters. The moving-coil principle has already been described in connection with high-quality receivers (Section 8.4.2) but it was first employed in 1926 in a high-power speaker, the 555-type, which had a flat response from about 60 hertz to well over 5,000 hertz (the response was down only 10 dB at about 7,500 hertz). The network analysis technique that had proved so useful in designing condenser transmitters was again used in designing this speaker.

About 1930 an electromagnetic transmitter, coded the No. 618 and illustrated in Fig. 3-95, was developed using the same moving-coil principle employed in the receiver and loudspeaker discussed previously. This device was more stable than the carbon transmitter, had a wider frequency range than the No. 394 condenser transmitter, and was some 10 dB more efficient. Its output impedance was about 10 ohms.

Thus, by 1930 the moving-coil principle had been applied to microphones, loudspeakers, and research-type telephone receivers. The basic idea, together with a piston-type diaphragm with tangential corrugation invented by H. C. Harrison in 1926,[102] are still widely used today in electroacoustic converters. While the same basic mechanism

[102] H. C. Harrison; U. S. Patent No. 1,734,624; filed April 16, 1926; issued November 5, 1929.

Fig. 3-95. No. 618A moving-coil microphone.

is used for receivers, microphones, and loudspeakers, it should be appreciated that the design requirements and the acoustic constraints varied greatly with application. As a consequence, each proved to be a different design problem but all responded to the newly devised technique of applying network theory to electrical analogs of electroacoustic systems. Thus we see that this new design technique, together with the use of the vacuum tube, not only produced research tools essential to the better understanding of the telephone problems but also provided devices needed for applications not dreamed of when the research was started.

IX. SUMMARY

During the first 50 years of telephony, station apparatus had evolved from a crude assembly of available components, capable of use over only short distances, to highly developed instruments meeting the needs of transcontinental telephony. Even though large advances had occurred during this period, the year 1925 was less significant as concluding a period of station improvement than it was as marking a major transition point in station design. Up to this time, telephone instruments still consisted of deskstands and wall sets designed largely on the basis of invention and experiment supplemented to some extent by the laboratory measuring techniques which had come into common use in the preceding ten years. But large changes were about to take place and telephone engineers had been preparing for them for a number of years. The design of the stations in use in 1925 had been greatly influenced by constraints imposed by the state of the art in the transmission, switching, and signaling fields. For many years transmitter and receiver design had been directed almost entirely toward achieving the highly efficient but distorted form of speech transmission that was most effective in counteracting the inevitable attenuation of electric waves which occurred as the lines were extended. But during this period important developments were occurring in other parts of the telephone system, as will be related in subsequent chapters. These developments were gradually removing the constraints on station design, and this freedom made it possible to apply in 1925 the advances in theory and field evaluation which had taken place in the preceding ten years. As the first 50 years of telephone technology concluded, these methods were already being used to design a handset meeting performance requirements specified in advance. This handset, introduced in 1927, and the anti-sidetone set introduced shortly thereafter, marked the beginning of the application of theoretical and field evaluation techniques. Such work has been the basis for the many telephone station improvements which have played such a large part in achieving the ease of communication characteristic of today's telephony.

As a conclusion to this chapter and a preliminary to those which follow, we might point out that there was a pattern in technological evolution that was followed to a considerable extent in the early development of all parts of the telephone system. The development of station apparatus provides a particularly good illustration of this pattern but the perceptive reader also will see analogies in the evolution of transmission, switching, and signaling systems.

There are five rather clearly marked phases in this pattern. First, of course, comes invention, the discovery through intuition, observation, and alert perception of significant new principles. The first ten-year period of station development was a very fruitful time for invention. Bell propounded the idea of analog conversion between acoustic and electric waves and had envisioned both the electromagnetic and the powerful variable-resistance conversion principles. He had also devised means for implementing the principles. Watson in 1878 had devised the polarized ringer and adapted the magneto to actuate it. As early as 1877 the receiver-operated switchhook was invented and the use of an induction coil in the station set was proposed. All this early work stimulated other inventions and there were soon many proposed designs for microphonic transmitters, culminating in 1886 in Edison's proposal to use granules of carbonized anthracite coal.

The original implementation of most inventions is usually rather crude and is followed by the second stage in technological evolution, an era of experimentation during which improvements conceived on an intuitive basis are gradually evaluated and improved through a trail-and-error process. Naturally the age of improvement overlaps the period of basic invention but it would be fair to say that the second ten years of telephony was particularly productive in this field since during this time the technique was used for selecting the most effective frequency response, for devising a practical (solid-back) granular-carbon transmitter, and for finding various ways to improve the efficiency of receivers and station circuitry. Experiment continued to play a leading role in achieving improved performance and better manufacturing techniques until about 1910 or 1915.

The development of laboratory measuring methods was the third stage in evolving technology. Until then the effectiveness of experiments had been severely limited by the difficulty of measuring the results of change. Much depended upon judgment and personal opinion. About 1910, or shortly thereafter, a scheme was introduced for subjectively measuring performance. But the application of the vacuum tube, beginning about 1915, revolutionized the art of design by providing a means for objectively measuring detailed performance characteristics of instruments such as the input–output relation, noise pickup of transmitters, frequency response, and so forth. From this time on it became possible to use a scientific approach to the design, based on objective measures rather than personal opinion. Experiments could at last be evaluated on a quantitative basis.

This set the stage for the fourth phase, the theoretical approach to design. This involved two matters: (*i*) theory for predicting performance from a knowledge of physical characteristics, and (*ii*) a knowledge of the requirements (electrical characteristics, etc.) to give different grades of performance. Together, these two kinds of knowledge changed the whole design approach. Without theoretical methods, design is a matter of invention and intuition leading to a multitude of experiments aimed at empirically selecting the best design that has been derived by this largely chance procedure. With sound theoretical knowledge, instruments could, within limits, be designed to meet specific performance requirements. Search for this type of knowledge was begun on a broad scale about 1915. The development of analytic methods was greatly facilitated by prior developments in the transmission art by G. A. Campbell[103] and others. Johnson and Shea adapted this work to the design of station circuitry and other networks, and it was further extended to electroacoustic systems, as discussed in Section 8.4.1, by Crandall and others.[104] Understanding of design requirements came from research into speech and hearing conducted by Fletcher and his associates.[105]

With this material available, design for performance became possible. The basic work was accomplished largely in the second and third decades of the twentieth century and much of it was ready by 1925 and was put to use in the design of high-quality instruments, the first practical handset, and other large improvements in telephone instruments which followed.

A final important step in the development of telephone instruments came when techniques were devised for their field evaluation. The theoretical approach provided a basis for designing telephone instruments meeting specified physical characteristics and laboratory measuring techniques could be used to confirm theoretical predictions. But the overall performance of a telephone system depends on complex human reactions which are not easily predicted from theory.

[103] For example, his articles on "Cissoidal Oscillations," In *Proceedings of AIEE*, April 25, 1911, and "Physical Theory of Electric Wave Filters," in *Bell System Technical Journal*, November 1922.

[104] *Transmission Circuits for Telephonic Communication*, by. K. S. Johnson (1924); *Transmission Networks and Wave Filters*, by T. E. Shea (1929); *Theory of Vibrating Systems and Sound*, by I. B. Crandall (1926).

[105] *Speech and Hearing*, by H. Fletcher (1929).

As an example, there was no theoretical basis for estimating the extent to which a reduction in loudness, resulting from a flatter response characteristic, would be compensated for by the more natural-sounding speech. The articulation tests proposed by Campbell in 1910[106] and developed by Fletcher gave one means for deriving an overall performance figure but left unanswered questions of customer usage such as the effect of sidetone on talker volume. In the 1920s, W. H. Martin and others had evolved techniques for field testing on telephone users (employees of the American Telephone and Telegraph Company and Bell Telephone Laboratories) involving a vacuum tube device for measuring speech levels and repetition counts as an overall means for evaluating transmission. Shortly thereafter, McKown and Emling evolved a system for rating transmission of overall connections and their component parts based on the field testing concept supplemented by articulation and other laboratory techniques. This rating technique provided the telephone engineer with quantitative measures of the many changes being made in the telephone station and did much to encourage their adoption by operating companies.

Invention is, of course, the starting point for most great technical advances and it continued to play an important role throughout the history of telephony. However, after the first part of the twentieth century, the pattern of its reduction to practice was altered. Once scientific methods had become available in the field of telephony, and their power demonstrated, they played an increasing part in communication development. The use of objective measurement, analysis, and field evaluation reduced the reliance on empiricism, stimulated invention, and expedited the application of new concepts.

Thus by 1925 there had evolved a sophisticated approach to the development and application of telephone technology.

[106] See Chapter 4, Section 5.1.2.

Chapter 4

Telephone Transmission—The Wire Plant

In this chapter we chronicle the advances made in transmission of speech by wire during a period covering five decades, a period during which the barriers of cost and distance yielded to new insights, inventions, and organized engineering development. For more than 20 years practitioners of the telephone art, laboring with the earthiest problems of line construction and maintenance, understood little of the fundamental character of the signals passing over their wires or the mechanism of propagation. But by the turn of the century standardized types of wire facilities for telephony had evolved and analysis and measurement were penetrating much of the earlier obscurity. Thus the foundations were laid for the vast improvements in transmission quality and efficiency needed for the countrywide growth of telephony. By 1925 the distance barrier had been surmounted and the basic techniques were ready for developing the greatly improved quality of transmission and the enormous economies of scale which were to characterize the later years.

I. INTRODUCTION

This is the first of two chapters dealing with the transmission facilities used to convey the electric analogs of speech between telephone stations. It deals with wire transmission plant, i.e., those circuits employing physical conductors. These were the only practical kinds of transmission paths available for communication until the twentieth century at which time wireless (radio) means slowly became available and were used to supplement the wire plant. The evolution of these systems is covered in Chapter 5.

Wire communication circuits had been in use for transmitting telegraph and other signals for many years before the invention of the telephone and similar facilities were at first used for telephony, the pioneers of telephony concentrating most of their effort on improving the telephone instrument. This was the logical approach for the first few years when lines were short and the shortcomings of the instruments tended to obscure the impairments of transmission caused by the lines. However, it was soon discovered that the transmission of weak telephone waves was quite a different matter from the transmission of telegraph signals, and the manner in which these new requirements were met is the main subject of this chapter.

II. EVOLUTION OF THE WIRE TRANSMISSION PLANT

From the beginning, the developers of telephony faced very formidable transmission problems, both electrical and physical in nature. Often the two types of problem appeared to require incompatible solutions. As a consequence, much of the work in transmission has been, and continues to be, concerned with the development of workable compromises between conflicting electrical and physical requirements.

2.1 Basic Problems

Electrically, it was necessary to transmit the speech waves to a distance with adequate magnitude to be heard by the listener, sufficiently free from distortion to be understandable, and with unobjectionable amounts of interference from extraneous sources of sound. These were the early, basic problems; but as they were solved, others came to the forefront such as stability and speed of propagation.

Physically, it was necessary to provide conductors and structures strong enough to stand up against wind and sleet. As the demand for circuits increased, it became necessary to decrease the bulk by consolidating the conductors into cables and ultimately to place them underground to avoid the unsightliness and vulnerability of overhead construction. Throughout, it had been necessary to meet these requirements at a cost low enough to make telephony attractive to a potentially increasing market.

The physical problems were undoubtedly quite obvious from the outset and were the first solved, largely during the first 25 years and with little theoretical background. The electrical problems, while clear in a general way, required considerable development of theory because the complexities of telephone transmission first had to be understood. By the beginning of the twentieth century much of the theoretical groundwork had been laid and solutions to the electrical problems came along rapidly thereafter.

The evolution of transmission theory and the development of the transmission plant, which went hand in hand over a period of many years, will be more clearly understood if we first briefly review, from our

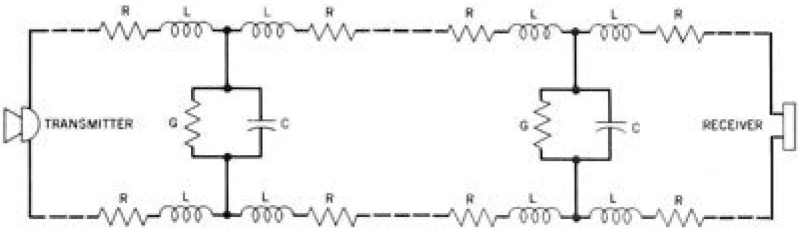


Fig. 4-1. Electrical equivalent of telephone line.

privileged position in the present, some of the characteristics of telephone transmission. With some simplification, a telephone line can be considered as a sequence of repetitive "H" structures, as shown in Fig. 4-1, with series elements consisting of resistance R and inductance L, and shunt elements consisting of capacitance C and leakage G.

When an alternating voltage is impressed at one end of the line, part of the power is dissipated in heat by R and G, part is stored in L and C, and part is transmitted to the load at the far end. For each unit length of the line, the power transmitted to the load beyond is the same fraction of the power introduced. Thus power is attenuated exponentially in successive sections. As mentioned in Section I of Chapter 3, this type of attenuation relation led to the development of logarithmic attenuation units so that total attenuation is additive and directly proportional to line length or number of line segments for any given type of line. Several of these units of attenuation have been used over the years and their evolution will be covered in some detail in Section 5.1.1. However, in this history, unless noted to the contrary, we shall use the decibel (or dB) to describe attenuation, amplification, or any ratio of two powers.[1]

When telephony began, there were a few theoreticians who had a rudimentary knowledge of the fundamentals of transmission as applied to telegraph practice. The exponential nature of attenuation was known or at least suspected by analogy with heat transmission, and the influence of R, L, C, and G was recognized even though the precise relationships would not be fully understood until Heaviside and others began developing the basic transmission equations ten years later. But it was even longer before this information was expressed in terms useful to the telephone practitioner.

Telegraphy involved the transmission of a series of impulses, and theoretical studies accordingly had been concerned largely with the tran-

[1] The dB difference between two powers, P_1 and P_2, is defined as $10 \log_{10} P_1/P_2$. The reader is referred to Section I of Chapter 3 for a brief discussion of some of the implications of this formula and its use for expressing attenuation (loss) and amplification (gain).

sient response of the system, that is, the retention in recognizable form of this series of pulses whose component frequencies were largely below about 30 hertz. Telephony, on the other hand, was characterized by a highly complex waveshape involving frequencies up to several thousand hertz. Transient analysis was not easy to apply to this type of wave, and it proved more satisfactory to develop new theoretical methods employing the "steady state" approach which considered the wave as a band of frequencies.

The commonly used transmission medium for telegraphy was iron wire, suspended on poles, with ground return. For this structure, at such low frequencies, the attenuation was very small and transmission was possible over considerable distance even in the early days of the art. The concept of relaying (or repeating) was introduced at an early stage so that when in practice it was found that telegraph currents decayed below a useful magnitude, a relay was introduced along the path, at a point where the currents were still usable, to control a new, locally supplied, high current which could carry for a further distance until it too attenuated to the point where relaying was required. Thus, long-distance telegraphy was accomplished at a very early date.

Telephony, on the other hand, required the transmission of frequencies 50 to 100 times as high as those used in telegraphy, and, as a consequence, the attenuation was many times greater. In addition, the problem of relaying or amplifying weak voice signals was not

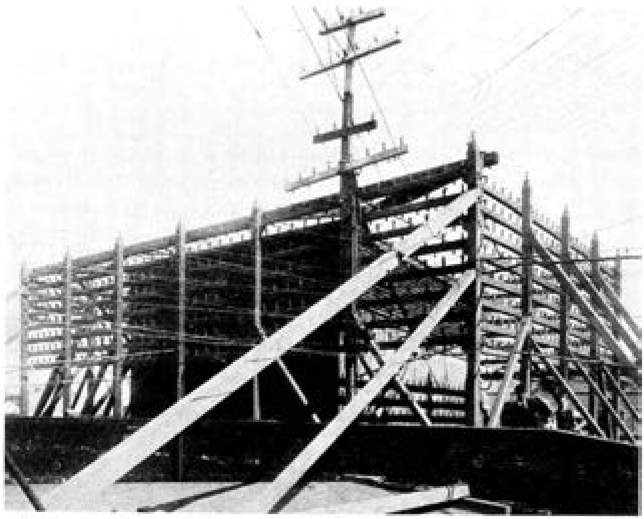


Fig. 4-2. Rooftop wire lines at Telephone Despatch Company, Boston.

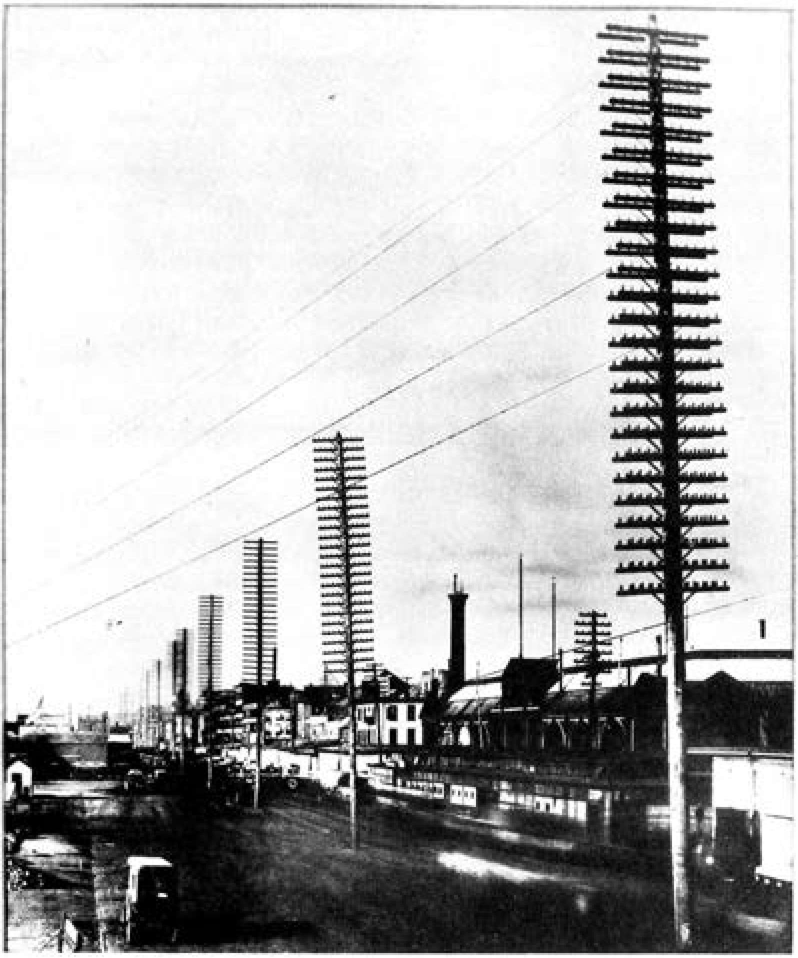


Fig. 4-3. West Street pole line in New York City, erected in 1887, with 90-foot poles, each carrying 25 crossarms and 250 wires. At some points the poles carried 30 crossarms and 300 wires. Picture taken at foot of Cortlandt Street.

solved for nearly 40 years. These were the two overriding facts that retarded the evolution of long-distance telephone transmission up to about 1915.

2.2 Open Wire Lines

In the early days, following telegraph practice, telephone transmission employed open wire lines, i.e., spaced wires with sufficient strength to be self-sustaining between infrequent supports. Initially, rooftop fixtures (Fig. 4-2) were used to support the wires, but as congestion increased it became the practice to use lines of poles. It was not until

February 25, 1891, however, that the first *standard* specification for pole lines was issued, calling for 40-foot poles of live cedar or chestnut. The minimum diameter at the butt was 17 inches for cedar and 12 inches for chestnut. Specifications in 1892 and early 1893 called for cedar only, since it was readily available, but by late 1893 the specifications again included the stronger but much heavier chestnut poles. Poles, however, could not always be limited in height to 40 feet. Figure 4-3 shows a pole line erected in New York City in 1887 using 90-foot poles. The early specifications called for pole spacing of 130 feet (40 per mile), which is still standard today for toll lines.

At about the same time, 10-foot crossarms spaced 2 feet apart, carrying ten wires each, became standard. The wires on the crossarms were spaced 12 inches apart (except for those adjacent to the pole) as indicated at (a) of Fig. 4-4. In the 1930s it became desirable to use closer spacing between pairs of wires and the arrangements shown on the other arms were introduced. The only one which came into common use was the 8-wire scheme illustrated in (c) of the figure. Even with these new arrangements, the crossarm size and spacing established in the early 1890s was continued.

Practically from the beginning, the wires were supported by, and secured to, glass insulators screwed onto threaded wooden pins projecting above the roof brackets or crossarms. This was a technique developed by the telegraph industry over a 40-year period of experimentation and was one of the few basic telegraph practices carried over into telephone line construction.

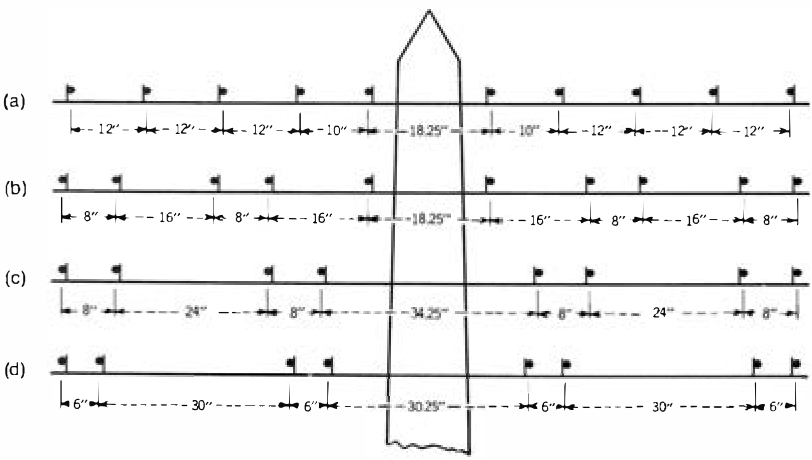


Fig. 4-4. Horizontal wire spacings used in Bell System at various times. (Babcock, Rentrop, and Thaeler 1955, Fig. 7)

Insulator development had been a long and somewhat controversial matter. In the early days of telegraphy various means had been used to support the wires, such as felt-lined grooves in the crossarm, grooved porcelain blocks, porcelain knobs projecting above or below the arms, etc. Some arrangements were roofed over to reduce the leakage from rain and snow. An almost endless variety had been invented[2] and in the process the requirements for insulator design had been clarified. Ultimately the glass knob above the pole proved preferable since it was reasonable in cost and had high insulation; though even with this type it was apparent from experience that there was a certain amount of conduction over the surface of the glass when it was wet. To render these insulators more effective, the lower part was molded in the shape of a bell or skirt for protection against dirt and moisture and to provide a longer leakage path. Later an internal skirt or petticoat was added to provide a still longer path.

With this long background of experience, it was natural to adopt the glass, telegraph-type, insulators for telephony, though a few technicians held out for other configurations. By 1890 a few types of glass insulators had been pretty well standardized (Fig. 4-5) and the same general types are still in use today.[3]

Following the pattern established for telegraph lines, wires were of iron or steel, sometimes galvanized and sometimes untreated. The iron was most often No. 12 gauge (109 mils) and the steel No. 14 gauge (83 mils), but other sizes were occasionally used.[4]

As can well be imagined, these lines were plagued with problems resulting from corrosion and rust, such as loose and noisy connections, and also with high attenuation, inherent in the high resistivity of iron and steel. Many different materials and techniques were tried in attempting to overcome these difficulties. One technique was the development of a compound wire in which a steel core was wrapped with a copper ribbon and the assembly tinned. However, it was practically impossible to eliminate tiny holes in the covering, and moisture entering these holes caused excessive corrosion. Another type of construction, somewhat better, was the application of an electrolytic coating of copper on a steel core. Although this construction materially reduced the corro-

[2] A writer of the time said, "The invention of insulators is a kind of scientific measles through which all telegraph managers pass with more or less danger."

[3] While the basic design of insulators was established at an early date, development continued over a long period to determine the most suitable material and the most effective configuration. For a brief period, porcelain was used. Ultimately, improved types of glass were used on very long lines employing the carrier systems introduced in the 1930s.

[4] Over the years various systems of gauges have been used to define wire size. In the earlier literature the particular system is not always stated. It was often, as above, in terms of the Birmingham Wire Gauge (BWG) but later the American Wire Gauge (AWG), or B and S system, became common.

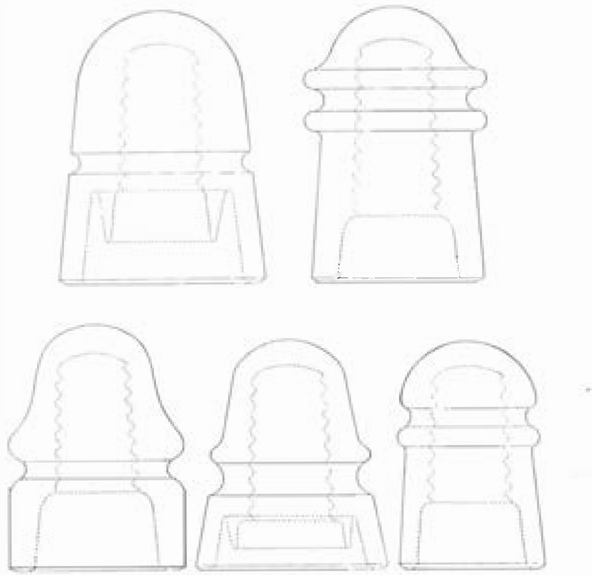


Fig. 4-5. Outline drawings of early glass insulators from an 1890 Western Electric catalog. Clockwise from lower left: Number 3, Double Petticoat, Western Union, Pony, Double-Petticoat Pony.

sion, it did not solve the basic problem of attenuation. Because of the very much higher frequencies used in telephony, the attenuation was an order of magnitude greater than at telegraph frequencies. Within a relatively short distance the feeble telephone currents decayed to the point where they were no longer usable. Until a way could be found to repeat (amplify) these signals, as had long been done in telegraphy, the only approach to extending telephony was by reducing attenuation.

A few individuals recognized that solid copper, with a resistance about 10 percent that of steel, would provide a large improvement in this respect, but methods of fabrication that were current did not produce a suitable wire. Wire then, as now, was produced by drawing strips or bars on copper through successively smaller dies until the desired size was achieved. Copper became brittle as it was drawn down to size and had to be annealed frequently during the drawing process to prevent breakage during succeeding passes. The final product was also annealed to improve its handling properties, and this incidentally improved its conductivity by a few percent. The annealing was desirable as long as the wire was to be used for winding coils, but unfortunately it also reduced the strength to the point where the wire could not support itself over an adequate span. The problem was solved by Thomas B. Doolittle, who was acquainted with wire manufacturing processes and believed that proper control of the size reduction made after annealing would produce a wire which would be hard and strong without being brittle. Experiments in the factory showed that such a "hard-drawn" wire could be obtained by decreasing the amount of area reduction produced by a die (i.e., using more passes for a given reduction) and by making a total size reduction of 11 numbers after the last annealing (e.g., from No. 1 gauge to No. 12).

This work of Doolittle's was of great significance. Commercial copper in its soft state has a tensile strength of about 28,000 pounds per square inch with an elongation of 36 percent. It was completely unsuitable for aerial construction not only because it lacked strength but because of the excessive and permanent sag taking place under load. The drawing process developed by Doolittle increased the strength to about 65,000 pounds per square inch with an elongation of only 1 percent.

Doolittle first used hard-drawn wire for telephone service in 1877 and in the following year he achieved a process that gave reproducible wire. But copper had a bad name among the practitioners and many would not be convinced that hard-drawn copper was not only suitable mechanically but also cheaper, for a given resistance, than steel (even though more expensive per pound).

Theodore N. Vail, then General Manager of the American Bell Telephone Company, was one who did listen to Doolittle. As a result, in 1881, extensive tests were conducted on hard-drawn copper wire under the supervision of Thomas A. Watson before his resignation that year as General Inspector of the New England Telephone Company. The results of these tests were very encouraging, and in 1883 an order was placed for 500 miles of such wire. On March 27, 1884, the first conversation was carried over this wire on a line between New York and Boston built largely by the Southern New England Telephone Company.

To explore further the properties and desirability of hard-drawn copper wire, a line was built between New York and Philadelphia in 1885 made up of wires of various sizes. Tests on this line provided data for determining the best size of wire to use for various installations and, from this time on, practically all overhead long-distance lines were made of hard-drawn copper.

In the meantime, other problems were arising from adherence to telegraph practices. It had long been the custom in telegraphy to use a single wire with ground return, not only to save the cost of the return wire but also because the ground return, on long circuits, offered a lower resistance path than a wire. Such a circuit is coupled conductively, inductively, and capacitively to other nearby electrical circuits and picks up unwanted currents from them. The interference of such currents with the telegraph signal was not serious, but with telephony even small interfering signals were annoying and the higher frequencies employed increased the inductive and capacitive coupling to the point where they were intolerable. It was soon learned that a full wire (metallic) circuit would reduce the interference from external sources, particularly those at some distance, since the induced currents in the two wires would be roughly of the same magnitude and tend to cancel each other.

It is generally agreed that the first major demonstration of a 2-wire metallic telephone circuit took place on January 10, 1881, on a commercial line between Boston and Providence. The demonstration was carried out by J. J. Carty, who is therefore credited with the introduction of the metallic circuit into commercial use. However, because of the expense of the additional wire and the need for new switchboards, several years elapsed before conversion to metallic circuits became general.[5]

While the metallic circuit greatly reduced the interference from sources at some distance from the telephone pair, it did not completely solve the problem of interference or "crosstalk" from a closely adjacent telephone pair on the same crossarm. The reason was that the disturbing wires were not both at the same distance from the two disturbed wires and consequently the induced currents did not completely cancel.

In 1886 the General Manager of the American Bell Telephone Company made the following statement:

> We found that between metallic circuits with wire connected straight through, there was a very considerable amount of crosstalk, not quite as much, perhaps, as there would have been on grounded circuits, but not materially less. We have entirely removed that by securing a balance between circuits.

This so-called balance was achieved by a system of transposing the wires of each pair in such a manner that each side of the circuit was balanced not only with respect to ground, but also relative to the wires of other pairs on the line.

[5] The basic idea of a metallic circuit was not new. The original concept of electrical transmission of signals involved the use of a metallic-return circuit. In 1838 Professor Steinheil of Munich discovered that the ground could be used as a return path, thus providing a large saving in conductors. Telegraphers hailed this as an epochal invention and quickly put it to use. Many years later, Werner Siemens reinvented the metallic circuit and described its use for telegraphy in British patent No. 2,089 issued in 1855. Bell, in a patent applied for in 1878 and issued in 1881, covering twisted pairs, clearly showed that he understood the technical benefits. The great reluctance to forego the savings and simplicity of ground-return circuits even as late as 1891 is illustrated by comments made by Thomas Lockwood at an AIEE meeting in March of that year: "Some of the considerations which tend to delay the universal employment of metallic circuits in telephony are, however, that there are an immense number of wires even when but one is used to a circuit; that twice the number of wires means twice the number to get out of order, to cross, break and leak, and consequently twice the number to take care of, and handle. I hardly think that sufficient thought is given to these matters by many who talk glibly about the necessity and essentiality of using metallic circuits only, for all classes of telephone work."

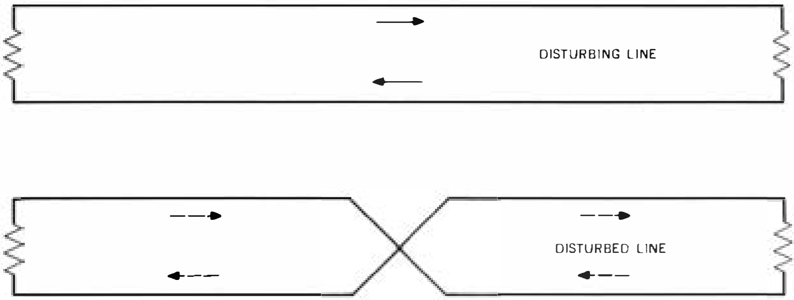


Fig. 4-6. Cancellation of crosstalk by transposing line. The disturbing line induces larger currents in the near wire of the disturbed line. By transposing, each wire of the disturbed line is, on the average, at the same distance from the disturber.

In its simplest concept, for a short, two-pair line, it merely required a "turn-over" of one pair so that residual currents on the west side of the turn-over were canceled by residual currents on the east side as shown in Fig. 4-6.

Obviously the problem was much more complicated when there were many pairs on a line since each needed to be balanced against all others. Moreover, the coupling between circuits proved to be much more complex than implied above, and additional problems arose as the transmitted frequency was increased by the use of carrier and as the length of long-distance lines was extended.

In 1885, John A. Barrett was given the task of balancing the line between New York and Philadelphia. He worked out an elementary scheme for such balancing which was later patented.[6,7] Then in March 1891, J. J. Carty outlined a basic theory of transposition which was to be used for a number of years. More elaborate and more effective systems for transposition were developed after E. H. Colpitts in 1904 conducted an elaborate series of measurements of direct capacitance using equipment devised by Campbell. His work led to modern transposition design. Two transposition techniques are illustrated in Fig. 4-7. The drop-bracket arrangement was most widely used, because of simplicity, until the advent of carrier transmission at high frequencies, at which time the point-type transposition began to receive wide use because of the more precise balance which could be achieved.

Thus by the last decade of the nineteenth century a distinctly "telephone" type of open wire line had evolved and pole line standards had

[6] J. A. Barrett; U. S. Patent No. 392,775; filed May 9, 1888; issued November 13, 1888.
[7] A more complete discussion of crosstalk and of the methods developed for its control will be found in Section 5.3.1. Figure 4-91 illustrates Barrett's first transposition scheme.

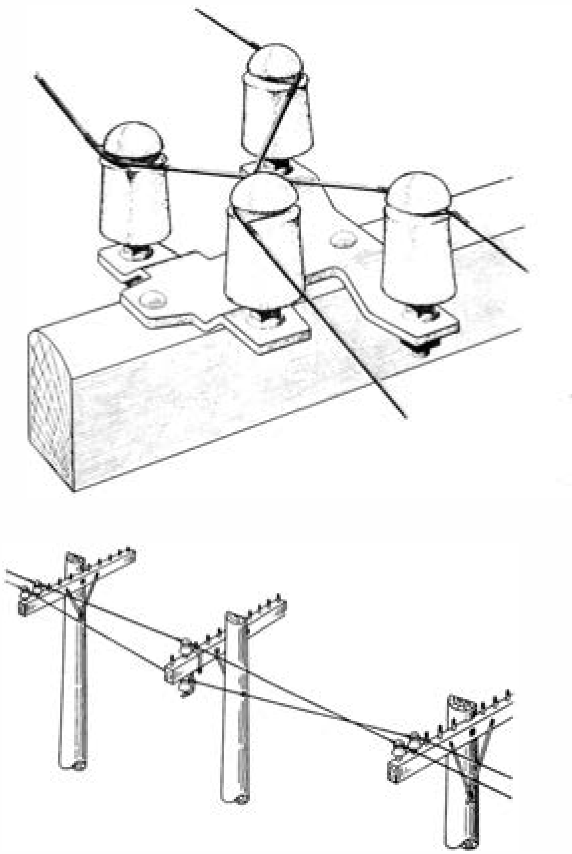


Fig. 4-7. Two standard methods for transposition on pole lines: top, point type, and bottom, drop-bracket type.

been established. These lines, using transposed metallic circuits of hard-drawn copper, were to form the backbone of the long-distance plant for many years and even today are the economic and accepted medium for wire telephone circuits in sparsely populated regions where only a few circuits are required.

The improvement in lines, together with improvements in telephone instruments, gradually extended the distance for usable telephone transmission until by 1892 a line was in operation between New York and Chicago, a distance of about 900 miles. In 1897 service was begun between New York and Omaha, over 1,300 miles. At this distance the attenuation of the line was so great, even with very large conductors, that service was just barely acceptable.[8] Thus it appeared that the limit of long-distance telephony had been reached unless a major advance of some kind could be achieved.

These advances were forthcoming after 1900, but before discussing them we should perhaps have a better understanding of the strengths and weaknesses of open wire lines so that the further evolution of transmission systems can be better appreciated.

The open wire line was in many ways not a bad medium for telephone transmission. The approximate attenuation for such a line at telephone frequencies[9] is given by

$$A \doteq \frac{R}{2}\sqrt{\frac{C}{L}} + \frac{G}{2}\sqrt{\frac{L}{C}}.$$

This relation between the primary constants and attenuation was not fully understood at the time, but with its help we can now see the basis for the pragmatic approach adopted of minimizing R through the use of copper (about 10 percent the resistance of steel) and minimizing G through the proper design of insulators. With a well-insulated line the part of the equation involving R was somewhat more significant than the G component. Intuitively one can sense the need for minimizing C, as a shunt across the line; but it was some time before Heaviside, Campbell, and Pupin convinced the practitioners that there could be a benefit from adding to the series inductance. In fact, there is for any line an optimum relation between C and L since a high C/L ratio increases the effects of R and a low ratio increases the effects of G. For open wire, it turns out, the ratio is not too far from optimum, though some improvement can be obtained by increasing L beyond the inherent inductance of spaced wires. The expression for A also shows that it is independent of frequency, and thus the transmission of voice frequencies is, in simple theory, without distortion. In actuality this is not quite true, since neither R nor G is itself independent of frequency. As the frequency is increased, R in particular is severely modified since the internal inductance of the conductor forces the current to flow near the surface. As a result of this "skin effect," the effective resistance ultimately gets higher and higher, increasing in proportion to the square root of frequency so

[8] For these long circuits it was necessary to use No. 8 BWG hard-drawn copper wire 165 mils (about 1/6 inch) in diameter. This was the largest size found practical. It weighed 435 pounds to the mile and corresponded roughly to No. 6 gauge in the B and S system usually used to describe cable conductors.

[9] This approximation does not apply at the low frequencies of telegraph where the series impedance is predominantly resistive rather than reactive, and is completely inapplicable to cable circuits. The quantity A is in nepers per unit length of line, the neper being a unit of attenuation equal to the natural or Napierian logarithm of the current or voltage ratio, or ½ $\log_e$ of the power ratio. The other terms, all on a unit length basis, are ohms, henries, farads, and mhos.

that ultimately attenuation is no longer constant but also increases as the square root of frequency. Fortunately, for copper wire, this effect is small at voice frequencies. With iron, internal inductance is higher because of its high permeability and the skin effect becomes noticeable even in the voice range. This causes an additional penalty for the use of iron and a discrepancy from expectations which proved puzzling to the practical communication man until skin effect was understood.

Thus, for voice-frequency currents, the metallic copper open wire line proved to be a good means for meeting the electrical requirements, since resistance and capacitance were low and the inductance nearly large enough to be optimum and in any event sufficient to offset some of the effects of capacitance and provide a nearly distortionless line. One undesirable aspect was that leakage tended to increase during rainy periods with consequent increase in attenuation. This effect was not too serious on early telephone lines but became more important as the line impedance was raised by the use of loading in the early 1900s.[10] Later, when frequencies well above the voice band were used for carrier transmission, weather which caused the formation of sleet or hoar frost became a serious problem. Under these conditions not only the leakage but also the effective resistance and capacitance were radically altered.

Physically, the open wire line was far from desirable, since it was bulky and unsightly and required a large amount of expensive copper and a complex supporting structure. The solution obviously was the use of small insulated conductors in close proximity; in other words, a cable. But the electrical effects were highly discouraging; not only did the use of smaller conductors increase the resistance but the close proximity between wires increased the capacitance by a factor of about 10 and decreased the inductance to a negligible amount. The reduction in conductance G, resulting from the improved environment, was only a small compensation. Even for the same size conductors the attenuation of cable at 1,000 hertz was roughly ten times as great as that of open wire, and for practical sizes of cable conductors the attenuation might be 30 times that of heavy open wire. Further, since the inductive reactance of cable was negligibly small at voice frequencies, the effect of the capacitance was controlling and the attenuation increased as the square root of frequency. Transmission was no longer distortionless, the quality of speech deteriorating as the circuit was lengthened because the high frequencies were attenuated more rapidly than the low.

As the effects of capacitance and inductance were better appreciated,

[10] Leakage was a particular problem in "alkali" country where the dust deposited on the insulators contained a large quantity of conducting salts. At one time the Mountain States Company, operating in the Salt Lake region, used steam from the boiler of a Stanley Steamer automobile to clean insulators.

the possibility was explored of reducing the high attenuation by adding inductance in some manner. The practical answer to the question came after 1900 and will be covered shortly as we describe the evolution of loading. At this point it suffices to say that this was a partial and highly valuable solution which considerably ameliorated the distortion and reduced attenuation by a factor of 3 or 4; but the final solution to the attenuation problem did not come until amplification became practical.

2.3 Cable

2.3.1 Structural Design

The first need for cable arose, as it did earlier in telegraphy, where it was necessary to go through tunnels or to cross small bodies of water, either above the surface or on the bottom. Telegraph practice was followed using cables made up of insulated wires drawn into metal pipes or sometimes merely protected by canvas jackets. Gutta percha, rubber, or a rubber compound (such as Kerite) was used as insulation. Single-wire, ground-return circuits were employed. The close proximity of the wires greatly increased interference between circuits as compared to the wider-spaced open wire lines. This interference, or some components thereof, was reduced by shielding, i.e., covering each insulated wire with a conductor at ground potential. Lead, tinfoil, or spiraled copper wires were used for the shield. By 1880 four cables, each with seven conductors, had been installed on the Brooklyn Bridge.

These cables had many faults. The insulation had a high dielectric constant, causing high attenuation; the insulating materials tended to deteriorate with age; and the shielding was only partially effective. Thus the use of any but short lengths was impractical. But the need for cables was increasing rapidly as telephony grew, and even early in the 1880s the congestion on open wire lines in the cities was so serious (Figs. 4-2 and 4-8) that some way to reduce the bulk was essential. The hazards of wind and ice storms had always been a serious problem with open wire. When these were combined with the great numbers of wires required as telephone customers multiplied, chaotic breakdowns in service could occur (Fig. 4-9) that might require weeks to restore. It was therefore of critical importance for telephone engineers to solve the cable problem.

David Brooks, an independent engineer, and later his son and namesake, were to take important steps in this direction. In 1878 Brooks had discovered that cotton and other fibrous materials were good insulators if kept dry. He made cables in which the individual conductors were insulated with a wrapping of cotton and a smaller, similarly insulated, copper wire was spiraled around the conductor

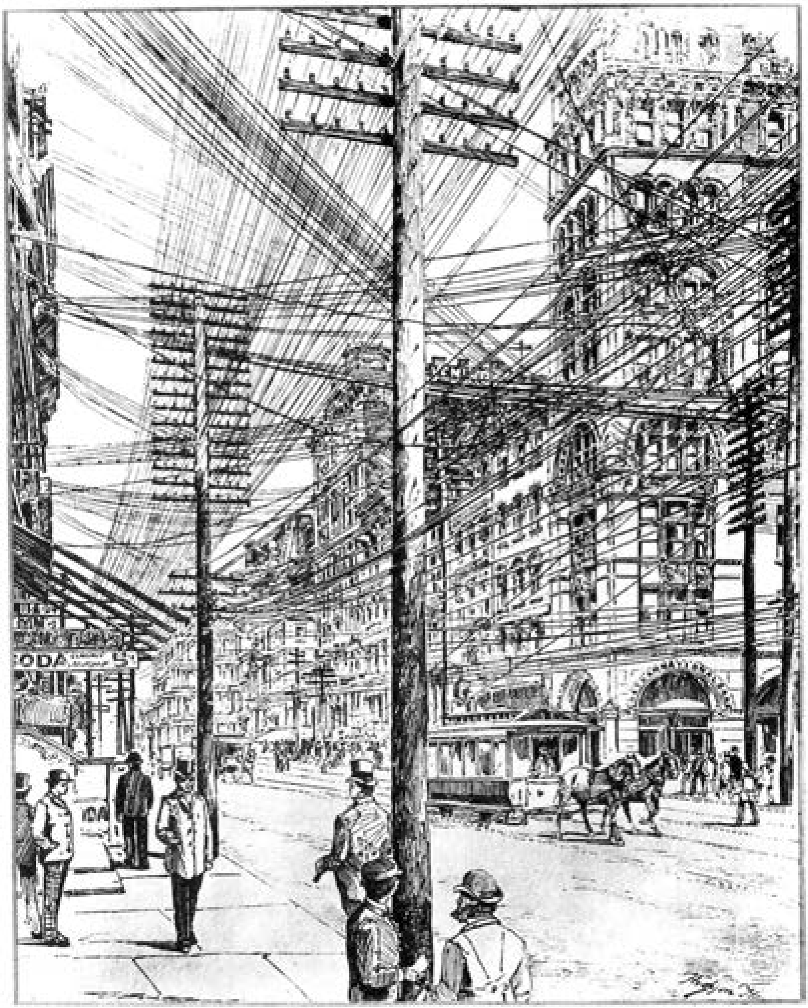


Fig. 4-8. Typical pole-line construction in the early 1880s. (Emling 1962, p. 8)

as a shield. A number of these assemblies were pulled into iron gas pipe, with special recessed couplings, which acted as a protective sheath. Brooks dried the cable core by heating it in oil, and kept it dry by drawing it directly from the oil into the iron pipe, which was then filled with oil. A reservoir of oil above the highest part of the piping system served to maintain the system under pressure, the theory

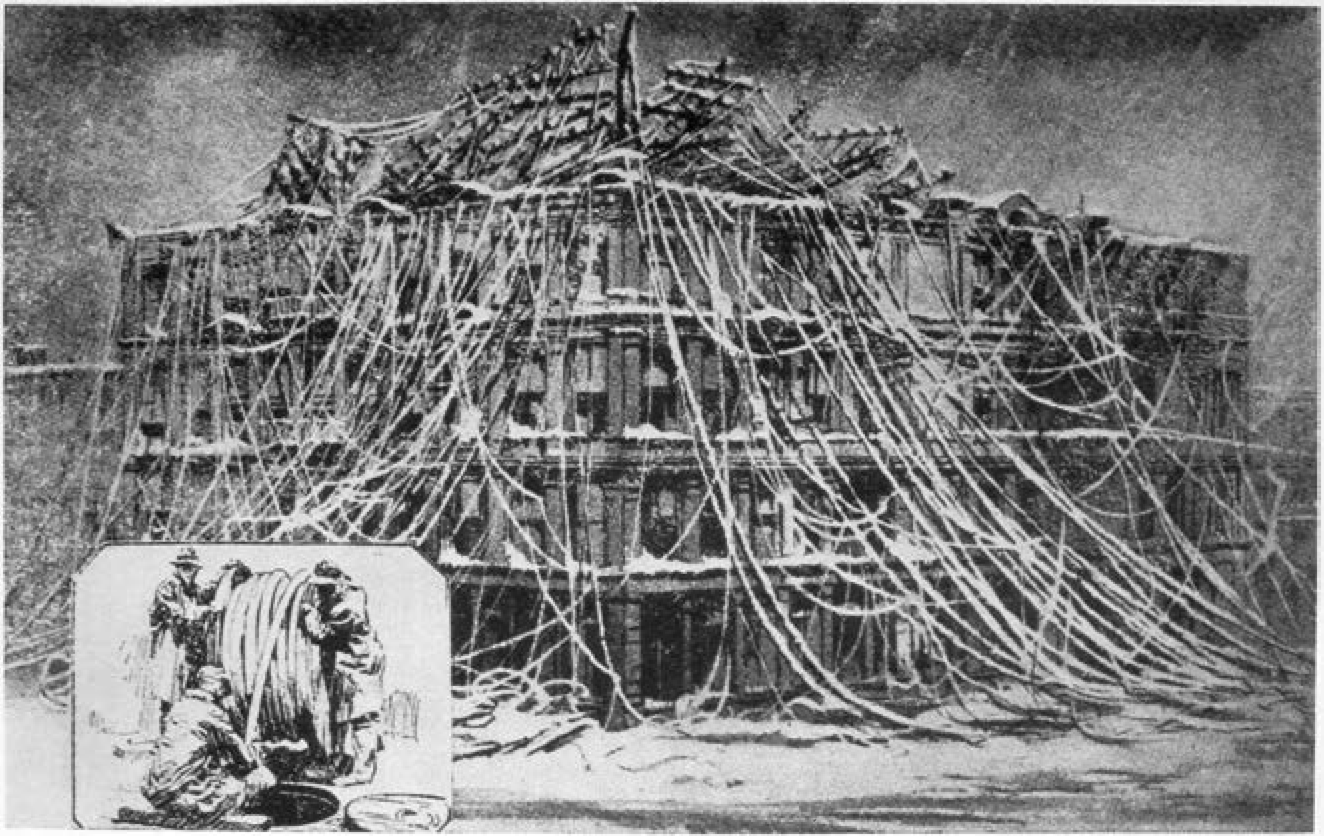


Fig. 4-9. Boston central station after sleetstorm of 1881. Inset shows installation of underground cables. (From contemporary account in *Telephone Topics*.)

being that in case of a small leak, the oil would seep out and prevent the entry of moisture. The Brooks system did not work too well in practice and was not widely used, but it did establish the possibility of textile insulation, and the use of oil under pressure was the precursor of gas pressure techniques introduced much later. Brooks also discovered that a metallic circuit with twisted pairs would solve the crosstalk problem, but in a patent interference suit it was found that Bell had come up with this idea as early as 1876 and he was judged to have priority. Bell was granted a patent in 1881 in which he stated with his usual clarity and comprehension of basic principles, "With twisted wires the relative distance of the wires is of little or no consequence so far as obviating inductive disturbance is concerned, since by the twist the wires of each pair are brought alternately to the same position relative to the other wires."

In 1880 the Western Electric Manufacturing Company (soon to become affiliated with the Bell System as the Western Electric Company) became interested in the Brooks cable, and obtained the rights to manufacture it. A young Western Electric man, W. R. Patterson, was assigned the task of working with Brooks and learning how the cable should be manufactured and installed. While engaged in this task, Patterson arrived at two very significant conclusions. First, if the pipe could be maintained free of leaks and if all moisture had been removed from the core, the oil filling was unnecessary. Second, a flexible pipe of lead, if it could be maintained leakproof, would offer important advantages over the rigid iron pipe.

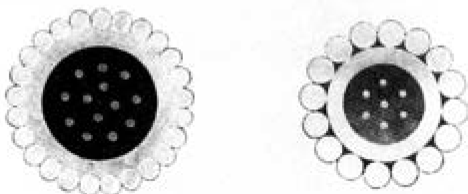
Patterson experimented with various techniques for implementing his conclusions. He developed alloys of lead that would provide pipe with adequate strength and resistance to corrosion. At first he was tempted to produce cable without any filler, but apparently his cable customers were not convinced that the filler was unnecessary and he cast about for a substitute for the oil filler.[11]

One of the materials which had engaged Patterson's attention was paraffin wax. He noted that when melted paraffin cools, it hardens gradually, shrinking as it solidifies. He tried charging melted paraffin with gas and allowing the mixture to cool under pressure. With this treatment, the material did not shrink but assumed a white, spongy appearance. When used as a filler in lead-sheathed cable, this spongy paraffin not only kept out moisture, but greatly improved the transmission characteristics (see Fig. 4-10). It was determined later that the improved transmission characteristics of this cable over previous cables with solid dielectrics were due to the fact that the gas in the

[11] Patterson has been quoted as saying, "It would have been as easy to sell a cable filled with air as to sell a sausage filled with air. So it was necessary to find a filler."

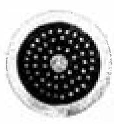
PATTERSON CABLE.

The "Patterson Cable" is a group of conductors, each covered with two or more windings of cotton or jute, or both, saturated with paraffine and protected by a lead pipe. The space between the group of conductors (or core) and the pipe is filled with aerated paraffine.



The core may be made in continuous lengths of 1,500 feet, and the protecting pipe is jointed over it in lengths of 75 to 100 feet. Any flaw which may exist in the pipe and any leaky joint is detected by the process of filling.

The almost invisible globules of gas, scattered uniformly through the mass of the paraffine filling, render it elastic, so that the natural shrinkage of the paraffine in cooling is compensated for by the expansion of these globules, preventing the formation of cracks and longitudinal fissures, through which water would penetrate indefinitely in case of a break in the protecting pipe.



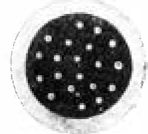


Fig. 4-10. Description of the Patterson cable, taken from the 1883 Western Electric catalog.

paraffin reduced the dielectric constant and thus reduced the capacitance between conductors. Over 75 years later, in the age of plastics, this principle was to be revived when expanded polyethylene and similar substances were studied as potential wire insulation.

Although difficult to manufacture, the Patterson cable, as it was called, became an important item in the early 1880s and provided the telephone industry with its first dependable multiconductor cable.[12] In 1883 the cable was being made available with either

[12] The status of cable development at the time of Patterson's work is well illustrated by a quotation from the American Bell Telephone Company Annual Report for 1882: "There is no strictly underground telephone system in any part of the world, the Paris wires being

single-line conductors or twisted pairs, but by 1887 only balanced, twisted pairs were being supplied.

The Patterson cable was not without competitors, a few of which should be mentioned to illustrate the ingenuity of the American engineer. One example was the Delaney Button cable, in which four No. 20 cotton-covered wires were run through a long string of ordinary shirt buttons. The buttons were spaced close enough to keep the wires from touching but far enough apart to form a flexible cable that could be wound on a reel and drawn through the desired size of pipe or conduit. Many years later, the coaxial cable was to use a not unsimilar scheme to position the center conductor. Another unique idea, which neither at the time nor later enjoyed much success, was the cable made up of about 20 bare wires, each drawn into a small glass tube. The group of tubes was then drawn into a lead pipe and a solution of some kind of grease or rosin forced in under pressure. When the pipe was bent, the tubes were, of course, broken, but the desired separation of the wires was supposedly maintained by the injected material holding the glass in place.

Because of the many different types of cables in use and under development in the mid-1880s, the need for standardization became more and more apparent. A Cable Conference was called in September of 1887 at which leading telephone executives and engineers gathered to discuss their experiences with cables and to establish the qualities needed for good performance.

Specifications for telephone cables were issued in 1888, based on the conclusions reached at this Cable Conference. These specifications called for the use of No. 18 B and S gauge copper wire covered to an outside diameter of ⅛ inch with not less than two wrappings of cotton. The conductors were to be twisted in pairs with a 3-inch twist and formed into a core made up in spiraling layers with alternating directions of lay. The spaces in the core and between the core and the pipe were to be filled with an insulating material. The sheath was to be an alloy of 97 percent lead and 3 percent tin about ⅛ inch thick. Electrostatic capacitance was limited to a maximum of 0.20 microfarad per mile, each wire tested against all others grounded.[13]

---

placed in huge sewers, practically large galleries, entirely altering the ordinary underground electrical conditions. Foreign companies are apparently waiting to see what is done in America; we have, therefore, nothing to guide us but our own experiments."

[13] This was the common way to express capacitance in the early days of cable development and was logical with ground-return circuits. When metallic circuits became common, it was realized that capacitance between wires was the significant characteristic and this was accordingly specified after about 1890. There is not a unique relation between the two figures but the capacitance between wires usually runs lower than the capacitance of a single wire to ground by a small amount (roughly 10 percent).

With these requirements, 52 pairs of wires could be placed in a cable having an outside diameter of 2 inches.

The next important step in the development of cables was the transition from cotton-insulated, filled cable to paper-insulated, dry-core cable. This type of cable served as a prototype for essentially all paired and quadded cables for the next 50 years, at which time the advent of reliable plastics began to influence design significantly. However, even in the late 1960s, 80 years later, over half of the cable conductors manufactured utilized some form of paper insulation.

The major advantage of these materials was a significant decrease in the electrostatic capacitance, due to the fact that the wire insulation was mainly air rather than solid dielectric. It became possible not as the result of a single invention but more as a matter of gradual evolution resulting from a number of developments. It will be remembered that Patterson had suggested early in his work that a filler would be unnecessary if the sheath could be maintained watertight. Improvement in manufacturing technique and the development of extruded lead sheathing were to solve this problem.

John A. Barrett, working on Bell System cable problems, was instrumental in first introducing the lead extrusion technique into the manufacture of telephone cable. He became acquainted with John Robertson, an inventor, who had developed a process for extruding a thin lead sheath on single electric-light wires. Barrett worked with Robertson on extending the principle to multiple-conductor telephone cables, and in 1886 a press capable of extruding a pipe 2 inches in diameter was built and limited production begun. The lead was alloyed with about 3 percent tin to improve mechanical strength and increase resistance to corrosion. This alloy had been developed earlier by Patterson and provided considerably higher breaking strength and resistance to corrosion than unalloyed lead. Figure 4-11 pictures an early sample of cable with extruded lead sheath.

About 1907 work was started to develop a material that would be at least as satisfactory as the lead-tin alloy but less expensive. Extensive laboratory tests and field trials covering a wide range of alloys were conducted over a period of four or five years. This work resulted in the development of a new cable sheath material consisting of about 1 percent antimony alloyed with the lead. This material was adopted for underground cables early in 1912 and for aerial cables about a year later. It was used almost universally until 1948, when the use of plastic-sheathed cable in aerial installations was started.[14]

[14] It is of interest that in England, in 1845, William Young and Archibald McNair received a very broad patent (No. 10,799) on the manufacture of cable. The patent

Fig. 4-11. Early cable with extruded lead sheath (1888).

These developments made the dry core possible, but several other developments were required before the modern type of structure was evolved. The use of cotton insulation, which required winding the conductors with a layer or two of fine yarn, was slow and expensive and experiments with the use of a strip of paper in place of yarn had begun at an early date.

described "Certain Improvements in the Construction and Means of Manufacturing Apparatus for Conducting Electricity" and covered the manufacture of a cable made up of wires surrounded by nonconducting substances enclosed in tubes. The use of thread in plaited or braided form as insulation was described, and means were outlined for impregnating the wire assembly with an insulating compound, the whole being enclosed in a lead pipe extruded around the core by means of hydraulic pressure.

This patent certainly anticipated much of the work on cable which we have described, but seems to have been premature since there was little follow-up and many of the ideas bore no fruit until they were independently resurrected 35 to 40 years later. It illustrates the fact that an invention, to be useful, must come at a time when there is both a need and a sufficiently developed technology to make it feasible.

At the 1889 Cable Conference, W. R. Patterson stated:

> . . . it is the sense of this conference that it is very important to follow up the experiments in the use of paper or other materials as a substitute for cotton and that it is desirable to authorize such substitution wherever the manufacturers will guarantee a substantial reduction in static capacity therefrom.

Early work with paper involved the use of sealing compounds, but experience with a cable in which the penetration of the compound was imperfect demonstrated the benefits that might arise from its omission. Toward the latter part of 1889 it was clearly recognized that the sealing mixture in a cable detracted from the desired electrical properties, and by the same time it had become apparent that trustworthy cable sheaths could be produced in large quantities. Some important installations of dry-core cable were made at about this time, and experience indicated that they were very satisfactory. By 1891, the cable specifications called for paper insulation as a standard. The size of the conductor was reduced to No. 19 B and S gauge and the electrostatic capacitance requirement was made 0.085 microfarad per mile (between the wires of a pair), very close to the present-day standard for exchange cable.

From this time on, the development of cables using the paper-insulated dry-core principle was rapid, with most of the effort going into the production of cables with finer gauges so that more and more conductors could be placed in a sheath. In 1901, for example, cables were placed in use containing 400 pairs of No. 22 gauge wire, followed the next year by cables containing 600 pairs of the same. These and other steps in the evolution of cable are shown in Figs. 4-12 and 4-13.[15]

Until the late twenties the layer type of structure, illustrated by the 24-gauge cable in Fig. 4-14, was common. This, it will be noted, is the type of structure, with reversed layers, described in the 1888 specifications. Originally all pairs had a 3-inch non-staggered twist, but in 1920 staggered-twist cables were introduced that used two different lengths of twist to reduce crosstalk in adjacent pairs. Later, as many as five twists were used in a layer. These measures greatly improved crosstalk suppression by reducing the chance of coincident twists in adjacent pairs.

In the late twenties the unit type structure was introduced with the manufacture of 26-gauge cable. This structure, clearly illustrated at the right of Fig. 4-12 and at the bottom of Fig. 4-13, was made up of a number of bundles of conductors (or small cables) which had been

[15] It is to be understood that the use of fine-gauge cables became practical only after the introduction of measures for improving transmission which will be described shortly.

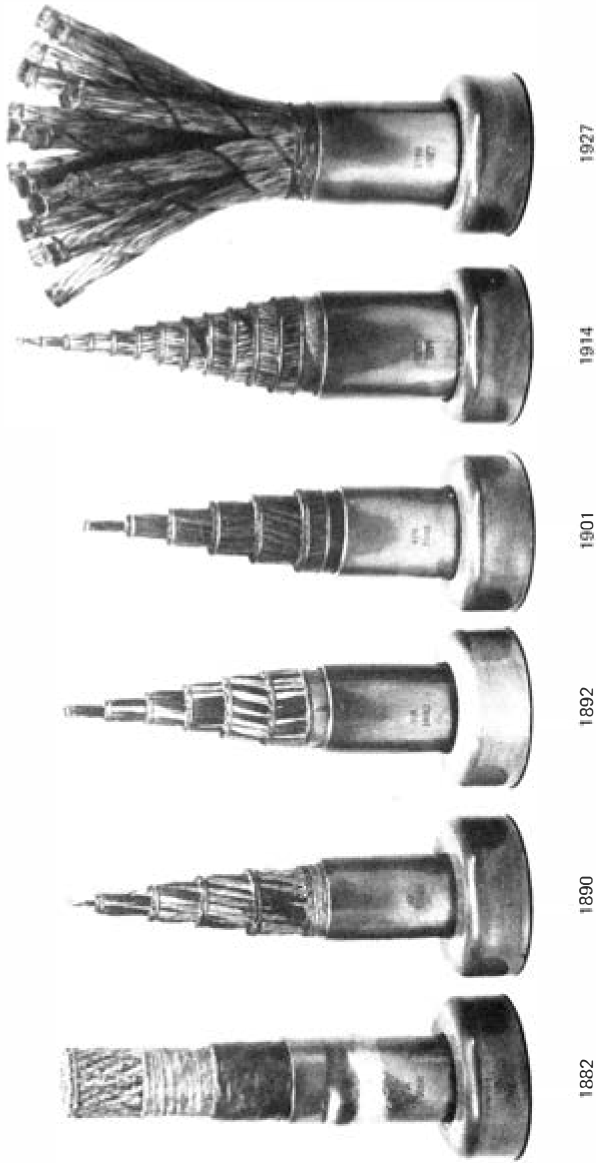
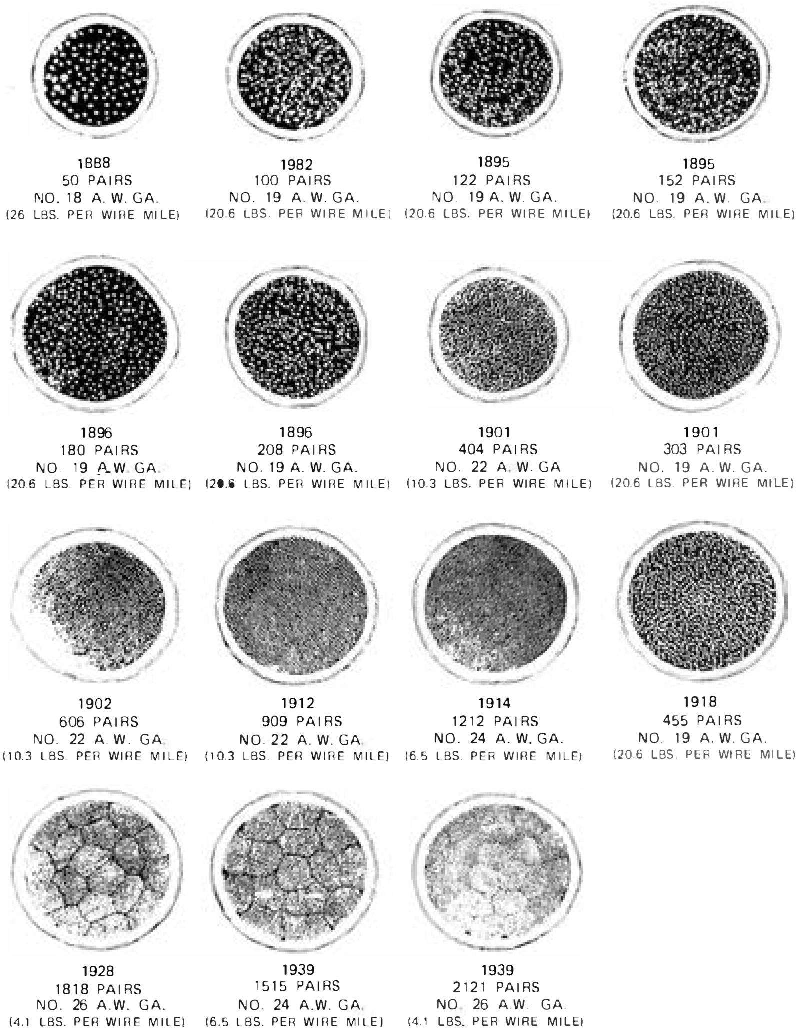


Fig. 4-12. Genesis of cable construction.

Fig. 4-13. Principal stages in the development of paper-insulated exchange cables.

CABLES SHOWN BELOW WERE FULL–SIZED CABLES AS OF THE DATES INDICATED. PHOTOGRAPHIC REPRODUCTIONS ARE REDUCED IN SIZE (ACTUAL OUTSIDE DIAMETER OF CABLES IN THE LAST TWO ROWS 2 5/8 INCHES)



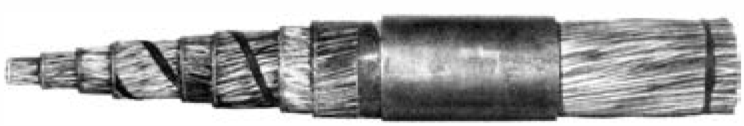


Fig. 4-14. A 1212-pair, 24-gauge cable showing how pairs are cabled together (outside diameter of 2⅝ inches).

formed in an earlier step in the manufacturing process. The use of a large number of twists was important in controlling crosstalk within a unit. Current cables are made up of 25-pair units, each pair of which has a different twist.

Another important change in cable manufacture began in the twenties, the use of a homogeneous paper covering formed directly on the wire in place of a helically wound paper tape. This "pulp-insulated" cable eliminated several steps in the fabrication process and led to considerable cost reduction. Work on the process began in 1921 but it was not until about 1925 that the process began to appear feasible. At this time a 10-wire machine was used to supply pairs for a number of experimental cables, several of which were installed in the plant. The work was so promising that the pulp-insulating machine was expanded to a 50-wire size and began operation in March 1928. Expansion after this was rapid; by 1932 all 24- and 26-gauge exchange area cables were being manufactured with pulp insulation.

2.3.2 Cable Splicing

The adoption of the air-core, lead-sheath cable introduced several problems in joining cable sections together.

During the splicing operation it was necessary to avoid the entrance of moisture through the open end. Hence a first step in splicing, after making the conductors accessible by removing the end of the sheath, was to pour molten paraffin over the conductors until it penetrated back into the sheathed portion far enough to block the penetration of moisture. This "boiling out" process was much later supplemented by the addition of desiccants (such as silica gel) before resealing.

The wires themselves were ordinarily joined by simply twisting the two ends together after removing the paper insulation. The twisted joints were often solder-dipped when the wires were large or in toll cables. The twisted joint was insulated by slipping over it a woven cotton tube. A large lead collar covered the spliced wires and was connected to the lead sheath by the standard plumbing practice of "wiping." This consisted of manipulating hot lead alloy around the junction of the collar and sheath until they were integrally joined.

2.3.3 Cable Installation Techniques

Use of cable was stimulated by a need to reduce the space occupied by the wires carried on overhead structures and, beyond this, by the desire to replace these structures entirely with underground installations.

Originally, overhead installations of cable were used in place of open wire on bridges, where space was limited, and in the neighborhood of central offices, where congestion was particularly severe. These installations were often little more than bundles of wires insulated with gutta percha or rubber, banded together with a wrapping or protective jacket, and usually suspended between rather closely spaced supports. As lead-sheathed cable became available, largely in response to the need for underground installation, it was also used overhead, and a need arose for supporting this heavy but inherently weak structure on pole lines having the wide spacings developed for open wire. Several attempts were made to manufacture self-supporting cable by incorporating steel wires during the manufacturing process. An interesting approach was the use of two longitudinal steel wires outside the sheath, held in position by circumferential wires spaced 3 inches apart and soldered to the longitudinal. This proved difficult to handle and tended to cause sheath breaks at the circular wire supports. The most practical approach, adopted at an early date, was the use of a supporting strand or messenger with the cable supported directly below it by means of metal or rope hangers at frequent intervals, as shown in Fig. 4-15. This technique continued until the advent of the lashing process, about 1939. In this process the cable is temporarily suspended from the messenger and a small reel of binding wire is pulled along the assembly in such a way that the wire is wrapped around the cable and messenger in long spirals and firmly lashes the two together (see Figs. 4-15 and 4-16). It is interesting to note that a similar process using tarred rope cord for lashing was used about 1881 or 1882, and while described as "very original, practical and effectual," it apparently did not prove as satisfactory at this time as the hanger technique.

After the initial need for short cables for river crossings and the reduction of congestion near offices, the greatest need was for underground installations to avoid the potential damage from storms and fire and to meet the public demand for eliminating unsightly overhead structures. In these cases, greater lengths were required and the inherently high attenuation and distortion of cable assumed importance. These effects were small for a short entrance cable or river crossing but became serious when long cables were tried. Often the reason for the poor performance was not clearly understood. The discouragement prevailing at the time was eloquently expressed later by Thomas Watson:

> My recollection of all this work is chiefly characterized by a feeling of hopelessness I then had. I felt as if I were gazing into an abyss that must be bridged by lives of work and experiment before we could get to the other end.

A major concern, when attempts were first made to place cables underground, was how to protect them from moisture, corrosion, and physical damage. An 1881 installation in Pittsburgh indicates one ap-

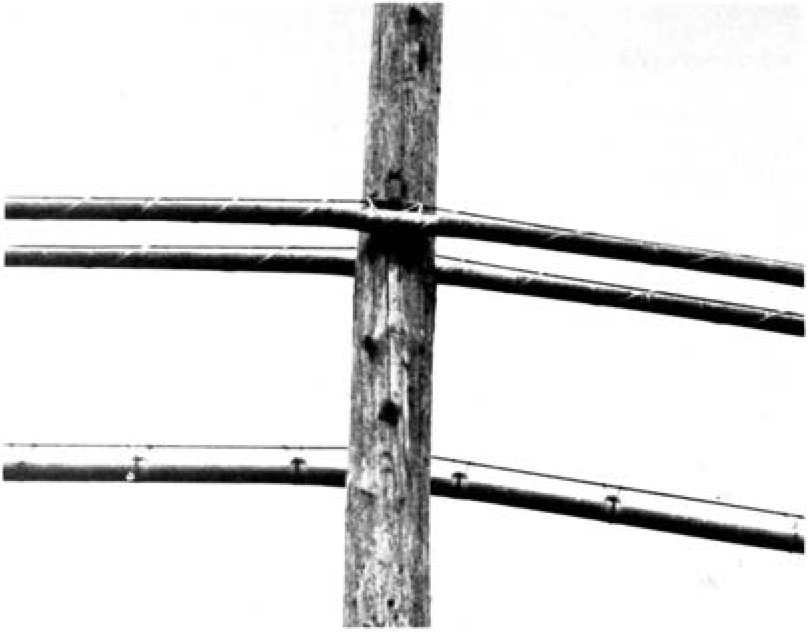


Fig. 4-15. Cable supported by wire "messenger." Wire hangers support lower cable; lashing is used on upper cables.

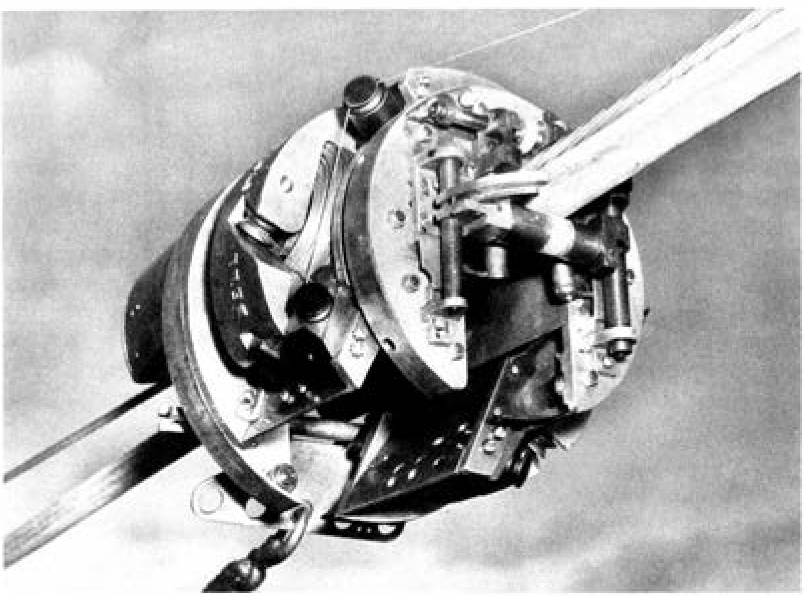


Fig. 4-16. Method for lashing supporting wire to cable.

proach. Three Patterson cables having 50 conductors each of No. 26 copper were laid in a wooden box and covered with pitch or asphalt. The following year eight more cables were installed. These experiments did not prove very successful, possibly because the very small wire made the attenuation quite high. In addition, the installation techniques proved laborious and not very practical.

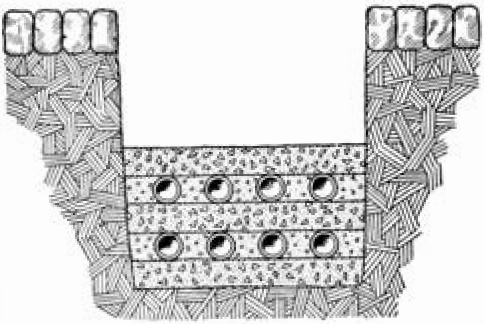
Over a hundred schemes had been proposed for getting wires underground, and an experimental evaluation of some of them was obviously needed. This was undertaken about 1882 in a 6-mile installation along the right-of-way of the Boston and Providence Railroad. The trench for the cables was opened by a plow drawn by the train employed in the experiment. After the cable was placed and joined and the jointing boxes made watertight, the trench was backfilled by an improvised scraper pulled by the train. Various types of cable using rubber-insulated wire were laid for comparison, the whole weighing nearly 30 tons and filling three freight cars. This ambitious experiment was not too encouraging. None of the cables gave very good transmission, presumably because of a high attenuation.

One thing became clear at this time. Considering the imperfections, both mechanical and electrical, of the types of cable then current, it was not wise to bury the structure permanently. Rather, it was concluded, a system of ducts should be provided into which the cable could be drawn, and from which it could be withdrawn for salvage or replacement if there was failure or if improved cable became available.

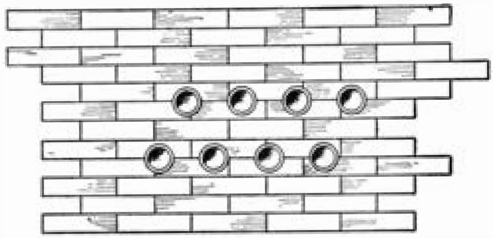
The first such duct system was installed in Boston. The ducts consisted of buried wrought-iron boiler pipes about 3 inches in diameter. A flaring bell attached to each pipe at a manhole prevented injury to the cable as it was drawn into the duct (Fig. 4-17). The manholes were constructed very carefully in an attempt to exclude water.[16] Several different types of cable were installed in such a conduit system early in 1883. Even though ground-return circuits were still commonly used, the conductors in these cables were twisted together in pairs. Thus it was possible to use them later for metallic circuits.

As time went on, other types of conduit were developed for use with underground cable. For example, in the latter part of 1885 trials were made of Dorsett conduit composed of a mixture of asphalt and sand. This material was molded into cylindrical sections. However, it was not very satisfactory, since its interior surface was rough and the sealing material squeezed through the cracks at the joints, making projections which broke off and then had to be cleaned out.

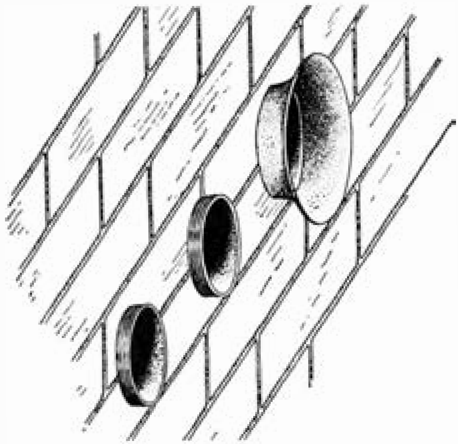
[16] The idea of building a watertight conduit system was soon abandoned in favor of a watertight cable system.



(a)



(b)



(c)

Fig. 4-17. First telephone "drawing-in" underground conduit system—Boston, Mass., 1882. (a) Cross section in street. (b) Arrangement of ducts entering manhole. (c) Brass bell attachment to ducts to prevent injury to cable when drawing in. (Rhodes 1929, Figs. 14 and 15)

The American Bell Telephone Company Annual Report for 1887 indicates that progress, though slow, was significant:

> It will be noted that the number of miles of wire underground has materially increased, and we expect further large additions to the underground mileage in the current year, especially in New York, where the Subway Commission has made arrangements for the working of an underground system, which will meet the requirements of the wire companies. . . . The Metropolitan Telephone and Telegraph Company is prepared to put a large part of its wires into subway conduits as soon as the obstacles referred to are removed. (The obstacles in this case were legal.)

Two different types of wooden conduit were developed and used to a limited extent. First there was the Macdonald conduit (Fig. 4-18) consisting of blocks of wood so grooved that when they were assembled a series of holes 2½ inches in diameter was available for drawing in the cables. Next came a creosoted wooden conduit called a "pump log." Each log was several feet long and about 4½ inches square, pierced longitudinally with a 3-inch round hole (Fig. 4-19). A much less sophisticated approach was represented by the very early creosoted wooden conduit shown in Fig. 4-20.

Around 1890, vitrified clay or glazed tile was tried out as a conduit material. Trouble was encountered with breakage and with cables becoming jammed so that they could not be removed. Shortly thereafter production was started on conduits of another form of vitrified clay known as hollow brick. These hollow bricks were laid in cement mortar and the entire structure then embedded in concrete. Following this, a vitrified clay conduit called the "multiple duct" came into use.

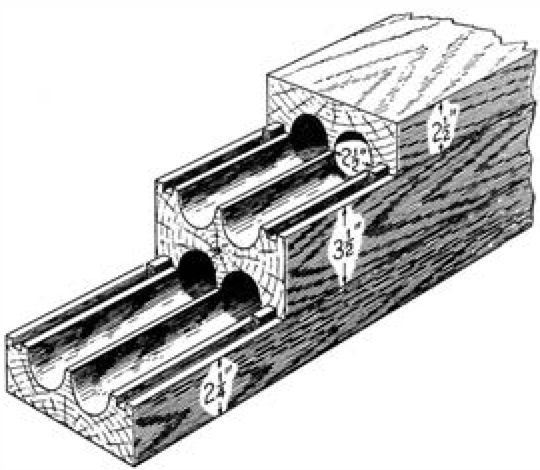


Fig. 4-18. Macdonald wooden conduit. (Rhodes 1929, Fig. 18)

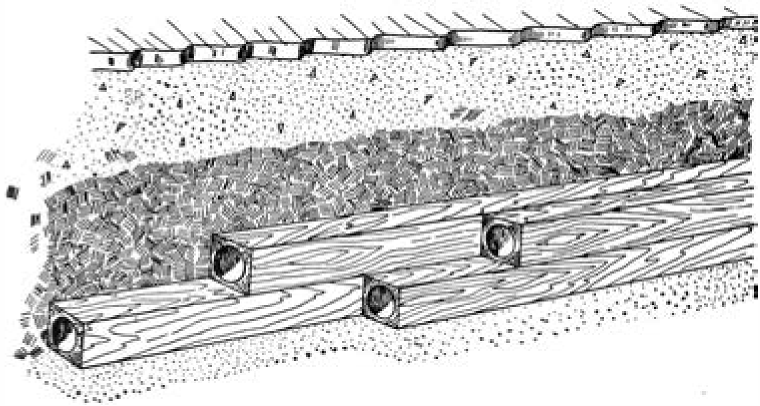


Fig. 4-19. "Pump log" conduit. (Rhodes 1929, Fig. 19)

In this construction a separate square hole was provided for each cable, the sections being generally made with four, six, or nine holes. This type of conduit was used very widely, and is similar to that employed in modern construction (Fig. 4-21).

It would be fair to say that by the nineties the general pattern had been set for underground conduit systems as we know them currently. Much later the plowing technique was to be revived when high-grade cables and tractor-drawn plows came into existence, but for about 50 years the system of manholes spaced to fit cable

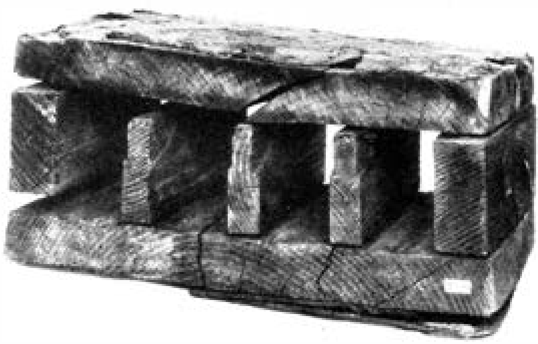


Fig. 4-20. Early creosoted wooden conduit.

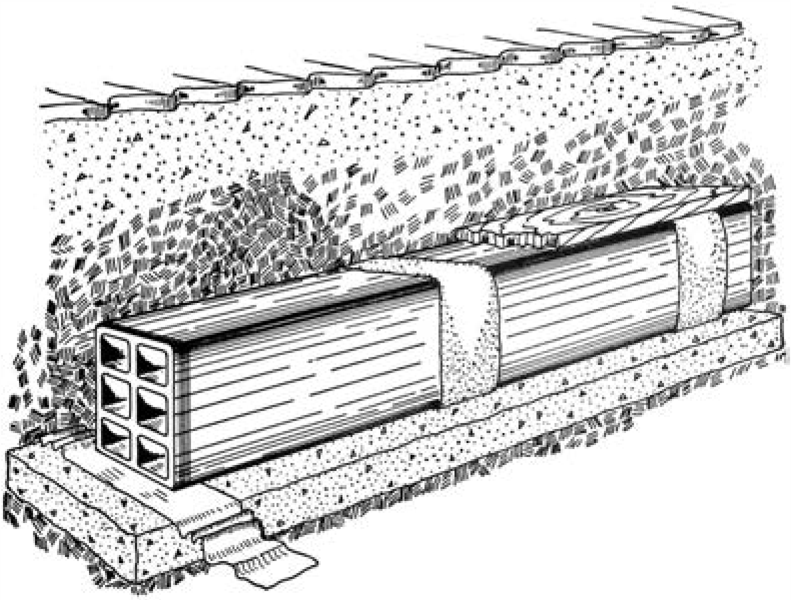


Fig. 4-21. Multiple-duct vitrified clay conduit. (Rhodes 1929, Fig. 20)

reel lengths and connected by vitrified clay conduit characterized the majority of underground installations.

2.4 Special-Purpose Wire and Cable

The facilities just discussed were the basic transmission media used in the early years of telephony. It was not long, however, before the need was recognized for specialized forms of wire and cable. For example, the conductors used for connection into the actual customer's premises needed to be resistant to the abrasion of tree limbs, and the connections into a central office were subject to space limitations because of the hundreds of pairs of wire. Submarine cable obviously had to be protected from a quite different natural environment, whereas toll cable had special characteristics dictated by the transmission requirements of long, balanced circuits.

2.4.1 Drop-Wire

At the user's end of the transmission facility it was impractical to continue the main line facilities all the way to the customer's premises. With rooftop distribution the customer's line was "dropped" to his premises below using the available type of insulated conductors. As pole distribution was introduced, these connections retained their name of "drop-wire," but a special twisted pair was developed with each wire insulated by a rubber compound. The rubber was covered with a braid saturated with asphaltum which protected it from abrasion and the deteriorating effects of light. This type of drop-wire, with No. 14 copper conductors, was adopted about 1900. Ten years later, higher-strength conductors of bronze or copper-steel were used to permit the use of longer spans and smaller wires. Since the total length of drop-wire used was commonly small, the increased attenuation was not important. In the late 1920s a parallel pair was used with insulation extruded over the two wires simultaneously. Initially a braid cover was also used but as the more durable synthetics, such as neoprene, superseded rubber it was omitted. This parallel construction not only was cheaper, because it required fewer manufacturing operations, but avoided degradation of transmission during wet weather by eliminating the crevice between wires. Moisture had tended to accumulate in the crevice, increasing attenuation because of the high dielectric constant of water.

2.4.2 *Terminating and Office Cables*

Special facilities were also required at the central office end of telephone lines. Originally drop-wire was used, but as the number of wires increased, cable was introduced to decrease the bulk. With the introduction of dry-core cables, it became necessary to seal them at the ends to prevent the entrance of moisture. To some extent plugs of paraffin or similar substances were used, but it became more common to splice on a "tip cable," often with textile insulation filled with wax or other impermeable substance. At the outer end, the tip cables terminated in screw terminals to which the open wire was connected. At the office end, these cables were connected to distribution frames[17] for interconnection with the office equipment. Within the office, "office cabling" was used which employed textile insulation saturated with wax, the whole being enclosed in a textile jacket treated with fire-resistant compounds.

2.4.3 *Toll Cable*

The early forms of cable were used universally, wherever needed. As the intercity or "toll" plant grew, pressure developed for underground installation as protection against storms. Because of the very high

[17] See Chapter 6, Section 3.4.6.

attenuation of cable relative to open wire, it was necessary to use large conductors and a loose "pack" of the wires to minimize capacitance. Thus, low resistance and low capacitance[18] were the main distinguishing characteristics of toll cable until the introduction about 1910 of "quadded" cable which was required to implement the phantom principle. Briefly, quadded cable was made up of two twisted pairs which were both twisted together. After about 1911, essentially all toll cables manufactured were of the quadded type. The phantom technique and the requirements it placed on cable will be discussed in Section 4.1.1.

2.4.4 Submarine Cable

As early as 1880 an attempt was made to telephone over the Atlantic telegraph cable, but without success. The reasons for failure were not clear at the time but it is now obvious that the attenuation and distortion of the medium were far too great. Three years later, success was achieved in a more modest trial when five cables crossing the North River between New York and New Jersey were connected in tandem, totaling 5 miles. Blake transmitters were used with Bell receivers, and it was reported that "conversation was carried on with the greatest of ease." In 1884 a 4½ mile, 7-conductor, gutta-percha-insulated cable was laid across the bay between San Francisco and Oakland.

As lead-covered cables were developed for underground use they were also used for short water crossings, but the cables were usually filled with bitumen to prevent the penetration of water in the event of sheath injury. As experience demonstrated the reliability of lead sheath, dry-core cable came into use about 1899. Protection and strength were provided by the use of extra thick sheaths covered by jute, the whole being surrounded by spiraling galvanized-steel armor wires. Hemp, saturated in a preservative compound, was wrapped around the armor. About 1900 a double lead sheath without armor was used in a few deep-water crossings where damage from dragging anchors was unlikely. The reasoning behind the double sheath was that defects in the two sheaths would probably not occur near enough together to cause serious damage.

These cables were all essentially the same as underground cables except for mechanical protection and since the pressure of cost dictated the use of small conductors, their use was limited to short distances.

As inductive loading became available for reducing attenuation, as described later, it was also applied to submarine cables. Ex-

[18] Wires as large as No. 10 B and S gauge (0.1-inch diameter) and cables with capacitance as low as 0.04 mF per mile were used.

amples were the Chesapeake Bay crossing in 1910,[19] the Tarrytown to Nyack crossing in 1916, and the Raritan Bay crossing in 1917. All of these were relatively short cables in shallow water, using specially encased loading coils.

About 1920, long-distance types of submarine cable were developed through adaptations of the single-conductor telegraph-type cable, and systems over 100 miles long were placed in service. One of these, the Key West-Havana system (Fig. 4-22), also revived a very early form of loading, in which inductance was added in a continuous manner along the line.

III. THE SITUATION IN THE NINETIES

Though the telephone system after 25 years was still somewhat short of attaining its inventor's goal of nationwide interconnection, progress had nevertheless been quite remarkable considering the technical obstacles, and it can fairly be said that by the end of the nineteenth century the service had come of age and established itself as a vital social and economic necessity. Thus the late 1890s offered in retrospect a view of vast local and regional expansion, outwardly apparent in lengthening mileage of pole lines and multiple crossarms, but with the revolutionary inventions still not at hand for solving the problem of transmission over long distances in an economical way.

3.1 Growth

By the turn of the century almost 2,000,000 miles of wire were devoted to telephone transmission in the Bell System, about six times the amount in use ten years earlier. Bell telephones, numbering over 800,000, were to be found in every major city and every part of the country; indeed, even as early as 1881 there had been only nine cities of more than 10,000 population without a telephone exchange. Some interesting statistics on this rapid growth appear in Fig. 4-23, taken from the Annual Reports to Stockholders of the American Bell Telephone Company (1880–1899) and the American Telephone and Telegraph Company (1900).

In 1885, as already recorded in Section 1.9 of Chapter 2, the American Telephone and Telegraph Company had been formed to build and operate lines interconnecting cities not only in North America but "also by cable and other appropriate means with the rest of the known world." For reasons that we have touched upon but shall examine in more depth, progress in interconnection was more difficult than

[19] This was the first Bell System underwater cable to use loading and it pioneered practices that were followed in later projects. It contained 17 pairs of 13-gauge paper-insulated conductors and was loaded at two underwater points.

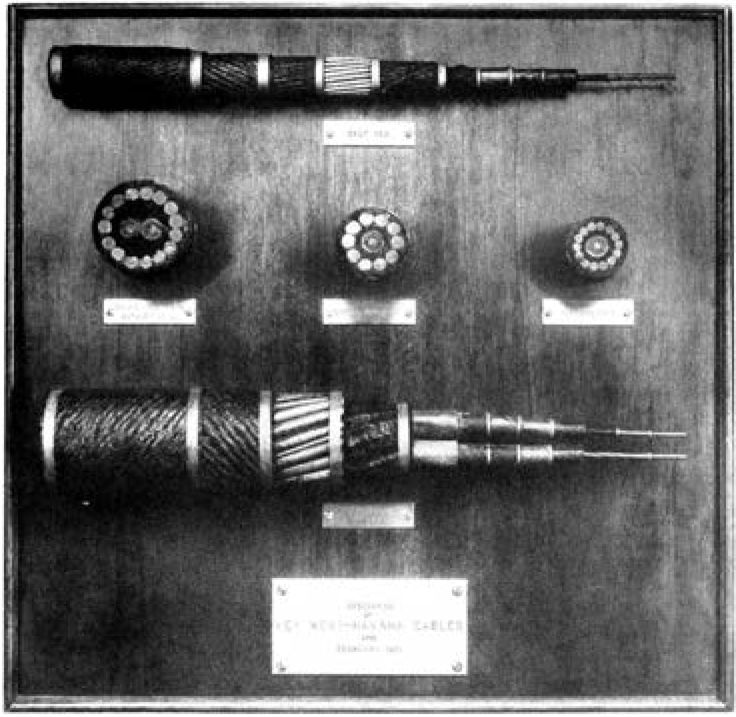


Fig. 4-22. Key West–Havana submarine cable construction.

had been originally expected. The first truly long-distance circuit was provided by the experimental copper line between New York and Boston in 1885. Two years later commercial service was opened from New York to Philadelphia. Lines spread out rapidly thereafter and by 1893 service from New York had extended as far south as Washington, as far north as Augusta, Maine, and as far west as Chicago. By the turn of the century western expansion had reached Omaha. While this was still far from the ultimate goal, it was sufficient to demonstrate the potential of long-distance telephony to the using public and to show the determination of the System to overcome technical limitations.

3.2 The Physical Problem

In the beginning, as we have noted, it had been necessary to use the highly unsuitable transmission media that had evolved for teleg-

Year	Stations*		Exchange Lines*			Toll Lines*	
	Total	On Metallic Circuits	Miles of Pole Lines	Total Miles of Wire	Miles of Under-ground Wire	Miles of Pole Lines	Total Miles of Wire
1880	48	NA†	NA	28	NA	0.7	1
1882	98	NA	NA	69	NA	6	14
1885	138	NA	10	114	3	31	42
1887	159	NA	17	146	8	32	56
1890	203	NA	24	240	55	36	91
1892	232	23	23	308	90	43	133
1895	282	95	25	460	185	53	215
1897	384	NA	23	626	283	68	325
1900	801	538	30	1,354	705	101	608

* All figures in thousands.

† Data not available.

Fig. 4-23. Bell System telephone growth, 1880–1900.

raphy, but within 20 years new media had been developed which were better adapted to the specific needs of telephony. In spite of the inadequacies in theory under which early technicians worked, these media were fundamentally so sound as to require no basic changes for many years.

The open wire line had been standardized using transposed balanced pairs of hard-drawn copper, with insulators, crossarms, and poles basically the same as used 50 years later. As time went on, improved transposition systems were to be introduced to care for larger numbers of conductors and later to permit the transmission of frequencies far beyond the voice range. These changes were to be accompanied by use of improved insulators and means for tensioning wire to give uniform sag. The supply of long-life cedar and chestnut poles was to become exhausted and means were found to protect the more abundant and cheaper woods from decay and make them suitable for long-life plant. Yet, important as these developments were, they represented no basic change in the open wire line available in the 1890s.

Similarly, the fundamental cable systems using dry-core construction with sealed lead (or lead alloy) sheath had been developed. These and the suspension and conduit systems were to remain basically unchanged for many years. This in no way implies that progress in developing cable systems stopped in the late nineties. Indeed, the great bulk of cable development lay ahead but emphasis changed from developing a basic structure to developing this structure into a practical and economic reality. Cable manufacturing processes were improved to speed fabrication and reduce cost. Crosstalk was reduced by the use of staggered-twist pairs, and ultimately the spiraled paper tape used on many forms of cable was superseded by pulp insulation. In the twenties the earlier idea of using dry air (or nitrogen) under positive pressure for testing and for avoiding the entrance of moisture became practical and its application has since become widespread. These were highly valuable developments and essential to the provision of telephony of the high quality and low cost which we have today but did not fundamentally change the basic structure, designed by Patterson and his co-workers, which has continued in use until the present. A new structure, the coaxial, was introduced for long-haul systems in the thirties, but it was not until the forties and later that plastics began to effect fairly fundamental changes in paired cable construction. Even today the underlying principle in most balanced-cable design is the use of twisted pairs, with thin insulation, loosely packed in a sheath so that air (or other dry gas) forms the significant part of the dielectric.

It is fair to say, therefore, that the basic physical problems of providing transmission media suitable for telephony over short to medium distances, strong enough to withstand the elements and capable of being compressed into a small space, had been solved by the middle nineties; but progress on the problems of longer-distance transmission and the problems of cost offered no cause for satisfaction.

3.3 The Distance Problem

The distance over which conversations could be conducted was severely limited by the attenuation of the lines. As we have noted, the problem differed from that encountered by telegraphers in two ways. First, the attenuation was much greater at telephone than at telegraph frequencies, and somewhat more for the metallic circuit required for telephony than for the ground-return telegraph circuit. Second, it was relatively easy to "repeat" the telegraph signal (i.e., send out a reasonably good facsimile of the original) after attenuation or distortion had reduced it to the point where it was just recognizable. While a similar need to repeat or amplify the telephone signal was recognized in the early years of telephony, the problem was not to be solved successfully until the next century.

During the first 25 years of telephony the major contribution to reduced attenuation was the use of copper in place of steel. When, for a given size wire, the line had been extended to the point where total attenuation was sufficient to render the speech signals unrecognizable, the only remedy was to use more copper, i.e., larger wire. Wires as large as one-sixth inch in diameter were used for the longest distances.

For open wire, the attenuation per mile was roughly proportional to the resistance, so that doubling the limiting transmission distance required twice the cross-sectional area of copper. This meant that the cost of long-distance lines did not merely increase in direct proportion to length; rather, the copper costs tended to increase as the square of the distance. This explains to a large extent why growth in long distance was slow and limited, as Fig. 4-23 indicates. With the 165-mil copper which was considered to be about the largest size economically and physically practical, the length over which conversation could be carried, even with the greatest effort, was not much over 1,000 miles.

Cable made the situation much worse because of high capacitance between wires. For practical cables the attenuation per mile in the middle of the voice-frequency range was 10 to 25 times as great as for common types of open wire, and the limiting length was reduced accordingly. Moreover, the attenuation of cable, unlike that of open wire, was affected by frequency f and, because of the absence of appreciable inductance, depended almost entirely on resistance and capacitance as follows:

$$A \doteq \sqrt{RCf}.$$

This was a very unfortunate situation since it meant that attenuation was greater at high frequencies than at low and telephone technicians, long before they understood the cause, learned that increasing the length of cable circuits not only reduced the volume of the speech but distorted it severely. In addition, because of the square-root relationship, the old remedy of "more copper" was less effective than with open wire lines since it required four times the cross-sectional area to cut the attenuation in half. Thus, for a given limiting total attenuation the copper costs increased as the cube of distance. Obviously cables were admirable for solving the problem of bulk in the cities, where distances were short, but would only have aggravated the difficulty of conquering distances.

A partial solution to the attenuation problem came in a few years with the advent of loading, but the final solution, amplification, came only during telephony's second quarter-century.

3.4 The Cost Problem

It is apparent by now that the first two decades, while bringing solutions to the physical problems of line configuration and construction, appeared to bring only retrogression from the standpoint of cost. The dismay of the early telephone manager can easily be imagined when he learned that the single-wire circuit must be replaced by a pair of wires and the relatively inexpensive steel wire by costly copper. Then there was further disappointment when it was found that much larger wire sizes were needed as the distance was extended. And finally, cable, which made it physically possible to use very small wires, proved usable only for very short distances.

Fortunately the technician, even though he did not have the answers, had begun to understand the problem. First, there was needed a means for reducing attenuation that would be cheaper than adding copper. Second, a way must be found to use the circuits more efficiently by multiplexing several conversations on a single pair.

IV. BEGINNING THE TWENTIETH CENTURY

Although practical solutions for the distance and cost problems had not yet been found, the outlook at the turn of the century was far from bleak. The preceding ten to fifteen years had seen a marked increase in interest in the theoretical aspects of telephone transmission, coinciding with a period of healthy reform in the relations between theorists and the so-called practical men. If the latter could justly complain of the aloofness of the theoreticians and their unwillingness to express their results in usable form, it was equally appropriate for Lord Rayleigh in 1884 to chide the practical electricians "whose ideas [as he put it] do not easily rise above ohms and volts." Thus, under the pressure of the telephone transmission problem, there was a gradual rapproachement as the work of Heaviside, Pupin, Campbell, and others developed a theoretical basis for solving these critical problems. In particular, Heaviside in Britain demolished the simplistic view of electric currents that had been prevalent even into the telegraph age, and showed that the propagation of speech currents over wires must be understood on the basis of waves—the electrical analogues of mechanical waves, as elucidated by the mathematician Lagrange a century before, and waves in fluids as studied by Newton a century earlier; with the inherent requirement—the *sine qua non*—for efficient transport of any kind of wave: a continuously moving cross-shuffle of potential and kinetic energy. It was the kinetic energy, lodged in the magnetic field, that had been appropriately ignored for the low frequencies of telegraphy, where the counterelectromotive force induced by the field was small compared with the potential drop in the series resistance.

As the revelations of Heaviside got across in their full portent, there was a diligent search in the telephone laboratories for ways of implementation. Thus, during the first quarter of the twentieth century an interim solution, in the form of inductive loading, was widely implemented, and in addition rudimentary multiplexing schemes evolved. Later the development of a practical amplifier lifted the barrier on distance, and the amplifier, together with the wave filter, opened the way to the use of carrier-current systems. These systems greatly extended the use of the multiplex principle and were to prove the most significant single factor in the evolution of long-haul and low-cost transmission systems during the middle portion of the twentieth century.[20]

4.1 Interim Solutions to the Cost and Distance Problems

4.1.1 Multiplexing—Phantom Circuits

As noted earlier, the cure of the interference problem carried with it a bitter pill indeed, the use of two wires per telephone circuit instead of the single wire which had been adequate for telegraphy. The notion of countering this setback by using the same wires for more than one communication channel must have been in the minds of telephone engineers practically from the beginning.

The first proposal of a telephone multiplex is attributed to Frank Jacob in England. Jacob, in 1882, suggested superimposing, by means of a Wheatstone Bridge arrangement, additional telephone conversations on circuits already carrying a conversation.[21] Several schemes for implementing this were proposed, the most significant one being shown in a simplified manner in Fig. 4-24a. The superposed circuit, for telephone C, soon to be called a "phantom circuit," was carried on two metallic circuits already in use for telephones A and B. In essence, the current from C was divided into two equal parts by means of a resistance bridge and each part transmitted over one wire of the pair used for A. Similar arrangements served to combine the components at the far end and to provide a return path over the pair used for B. Since the phantom currents flowed in the same direction over the two wires of the pair, there would be no mutual interference between the "phantom" circuit and the "side" circuit

[20] While most of the basic inventions and developments which made carrier transmission possible came during that first quarter of the twentieth century, one was not to materialize until just afterward. This was the invention of negative feedback by H. S. Black in 1927.

[21] F. Jacob; U. S. Patent No. 287,288; filed May 2, 1883; issued October 23, 1883.

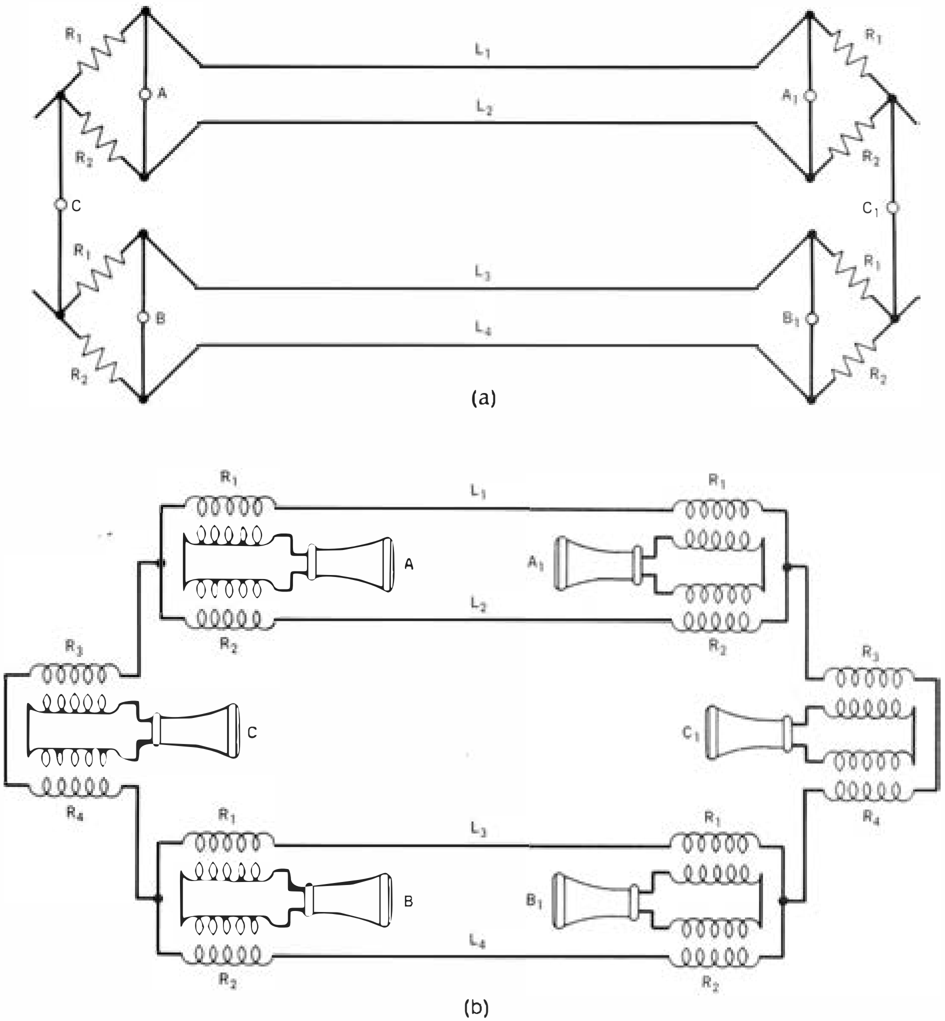


Fig. 4-24. Development of the phantom circuit: (a) phantom circuit, based on concepts of F. Jacob, intended for the simultaneous operation of three telephone circuits over two metallic circuits; (b) phantom circuit (scheme of J. J. Carty) using induction or repeating coils instead of resistances. A, B, and C are telephone sets. (Redrawn from Rhodes 1929, Figs. 43 and 54)

carried metallically on the pair so long as the phantom currents were equal in the two wires.

The resistances proposed by Jacob for deriving the phantom circuit provided an easy way to divide the current into two equal parts, i.e., to "balance" the currents between the two wires of a pair, but introduced a considerable attenuation to the voice currents. In 1886, J. J. Carty proposed a scheme for avoiding the major part of this loss by using transformers (repeating coils) in place of resistances. The Carty equivalent of the Jacob proposal is shown in Fig. 4-24b.

In theory, the phantom circuit was very simple, but the basic requirement of precise division of the current into equal parts was not easy to achieve at all frequencies within the voice range. Difficulties were encountered in attempts to make repeating coils with satisfactory balance, and the state of the electrical art at that time was not such as to enable even a skillful professional to comprehend all the steps involved. Thus for many years the phantom circuit remained scarcely more than an interesting scientific curiosity.

In 1894 and 1895 tests were conducted on phantom circuits on interoffice trunk lines in underground cable, and in a few cases fairly favorable results were obtained. Then in 1899 tests were made on open wire lines with the following results, as reported by C. H. Arnold:

> The trunks between Gloversville and Johnstown in their present condition cannot be commercially duplexed because it would be impossible to ring on the duplexed trunks because the phantom would be too noisy and because there would be objectionable crosstalk on the phantoms and on certain trunks. Before further duplex tests on aerial lines are made, systems of transpositions for duplexed lines should be worked out.

In the latter part of 1902 a phantom circuit was put into commercial operation between Lewiston, Maine, and Berlin, New Hampshire, over a transposed line. Although the results were fairly satisfactory, it was not until 1903 that really good performance was obtained by replacing the coils originally used in this installation with new repeat coils of the toroidal type.

It was the impetus given to coil design by the exhaustive early work on loading (described subsequently) that finally resulted in the production and use of phantom deriving coils of the toroidal type which, in connection with the working out of suitable plans for transposing the wires, put the use of the phantom circuit for open wire lines on a commercial basis in the 1904–05 period.[22]

The first transposition system for preventing crosstalk between phantom and other circuits was proposed by Carty in a patent issued in 1889. As soon as well-balanced coils became available, improvement followed improvement in rapid sequence, no less than eight phantom transposition specifications being issued by Bell engineers in a period of about four years, from 1904 to 1908.

[22] Full commercial use of the phantom principle, however, did not come until the development, about 1910, of phantom loading and quadded cable.

The design of cables for phantom circuit operation was considered as early as 1894. In April 1895, C. H. Arnold stated in a report that the simplest arrangement of wires in a cable for multiplexing would seem to be twisted pairs twisted. This arrangement had previously been proposed by J. J. Carty. First tests with such a cable were unsuccessful in keeping the crosstalk between the phantom and side circuits within commercial limits. The basic approach was sound, however, and continued development produced in 1910 a successful cable made up of "quads," consisting of twisted pairs twisted together.[23] The important refinement was to use manufacturing methods giving the greatest possible degree of symmetry in the two conductors forming a pair and of the two pairs making up the quad. Different pair and phantom twists were used on quads in close proximity. Cable cores were built up in layers with successive layers spiraled in opposite directions. In installation, the capacitance unbalance between pairs was measured and successive lengths spliced so as to minimize the accumulation of unbalance.

These measures made phantoming practical, and after about 1911 all toll cable manufactured for the Bell System was of the quadded type. The use of phantoming on both open wire and cable was to grow rapidly on the longer circuits until about 1930 when it became apparent that phantoming tended to complicate the application of carrier techniques, which in the long run offered much more promise as a method of multiplexing.

The phantom technique could in theory be pushed beyond the gain of one circuit for each two pairs. It was, in principle, possible to superimpose a "ghost" circuit on two phantoms and an additional "wraith" on two "ghosts." These were of more theoretical than practical interest, since the gain was small and the difficulty of maintaining balance was formidable. "Ghosts" were, however, occasionally used where the value of an extra circuit was very great, as on some submarine cables. Even phantom circuits were limited largely to long circuits, since on short circuits the cost of achieving good balance could be greater than the copper saving.

Cost reduction through an increase in the number of circuits per pair was the principal benefit from phantoming, but there was an additional small bonus applicable to the distance problem. Since the

[23] The multiple twin-quad, adopted by the Bell System, was not the only cable structure that could be used for phantom circuits. In Europe, the star-quad has been extensively used. In this structure the plane of the wires in one pair is at right angles to the plane of the other pair. If this configuration is rigorously maintained, there is theoretically no coupling between the pairs. The choice of the multiple twin structure for the Bell System was based on a number of practical factors involving both manufacture and application which are too complex to discuss here. An important consideration was the belief that if phantoming proved to have limited use, the twisted side circuits would prove superior to the side circuits of a star-quad.

phantom used two conductors in parallel, the resistance was cut in half, and since the capacitance was increased by only 50 percent, the net result was about a 20 percent reduction in attenuation. As a consequence, phantom circuits were for some time preferred over side circuits for very long hauls.

4.1.2 Multiplexing—Composite Circuits

The phantom principle could also be used for transmitting telegraph signals over pairs in use for telephony. The basic concept is illustrated in Fig. 4-25, which shows how a telephone pair can be used to provide a ground-return telegraph circuit. This is basically the "simplex" circuit, the evolution of which will be described elsewhere.

From the standpoint of cost reduction and the solution of basic transmission problems, the simplex concept was far less significant than the composite principle which was first used in 1892 and developed into its present form within the next few years. This too was a scheme for transmitting telegraph signals over working telephone lines. Whereas the simplex telegraph circuit merely replaced a phantom telephone path, the composite scheme employed a different principle which today we would call "frequency-division" multiplex.

As we have noted, telegraph signals require only a path for low frequencies, roughly under 50 to 100 hertz. Telephony, on the other hand, involves much higher frequencies and indeed is quite satisfactory in quality if nothing is transmitted below about 300 hertz. When this was understood, it became obvious that the two types of signals could be transmitted simultaneously over the same pair of wires, without mutual interference, if the two frequency bands could be kept separate. In the 1880s the Belgian, van Rysselberghe, had obtained patents on schemes for doing this.[24] These ultimately evolved into the "composite" circuit, illustrated in Fig. 4-26. In essence, this was a combination of high-pass and low-pass filters separating the telegraph and telephone frequencies. It permitted the addition of two ground-return telegraph circuits to a metallic telephone circuit.

Prior to the development of composite telegraphy, signals in the telegraph range (roughly about 20 hertz) were used to control circuit switching. When this band was used for telegraph, it became necessary to develop new signaling techniques. This was done by using a frequency of 135 hertz, which fell between the telegraph and telephone bands.

[24] F. van Rysselberghe: U. S. Patents Nos. 306,665 (October 14, 1884); 320,987 (June 10, 1885); 321,404 (June 30, 1885); 322,333 (July 14, 1885); 323,239 (July 28, 1885); 361,734 (April 26, 1887). In a sense, the basic idea of frequency multiplex had been anticipated by Cromwell Fleetwood Varley in British Patent No. 1,044 (April 8, 1870) covering the superposition of "tone" signals on dc telegraph by means of rudimentary filters.

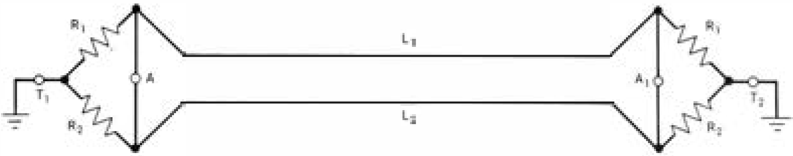


Fig. 4-25. Telegraph simplex based on the Jacob concept. A and A_1 are telephone sets, T_1 and T_2 are telegraph sets. (Redrawn from Rhodes 1929, Fig. 52)

Obviously, compositing was an important contribution to cost reduction since the telegraph circuits could bear part of the cost of a metallic pair. In retrospect, compositing was perhaps even more important as the first application of the frequency-division principle which was later to be exploited so extensively in carrier-current telephone transmission.

4.1.3 Inductive Loading

Well before the end of the nineteenth century the importance of minimizing resistance and capacitance to control attenuation was appreciated by telephone technicians. In 1854, before the invention of the telephone, Lord Kelvin had propounded the so-called *KR* law in connection with telegraph transmission. In essence this "law" stated that the maximum speed (characters or impulses per second) of telegraph transmission possible over a line was inversely proportional to the product of the total capacitance *K* and total resistance *R* of the line. It is obvious that Kelvin was thinking in terms of signal

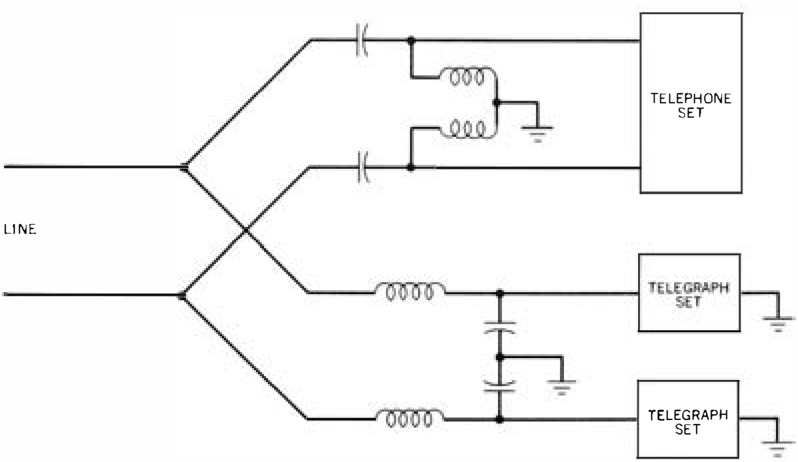


Fig. 4-26. Basic "composite" circuit for operating telephone and telegraph systems on the same line without interfering with each other. (Redrawn from Jewett 1923, Fig. 12)

distortion (intersymbol interference), not attenuation, since he specifically stated that the number of current reversals (i.e., signaling speed) was independent of the applied emf.

The telephone problem was different. Here the initially applied emf was very small because of the weak transmitting devices available, and distance was limited not so much by distortion as by attenuation, which reduced the signal below the sensitivity of the receiving device (or below the interfering noise on the line). Nevertheless, prominent telephone technicians agreed that some form of *KR* law was applicable, since telephony required the transmission of specific frequencies which imposed an upper limit on the *KR* product. Hence it was felt that the limiting length of transmission could be specified in terms of the resistance and capacitance per unit length.

There was a fair amount of justification for this argument, particularly for cable circuits where the attenuation is indeed determined by the capacitance-resistance product. There was, perhaps, just enough justification to delay a full understanding of the effects of the other primary constants.

A number of people challenged the *KR* concept, but Oliver Heaviside in England probably deserves the credit for expounding most completely the theory of electrical transmission. His results appeared in a series of papers presented in *The Electrician* and in the *Philosophical Magazine* beginning in 1873 and continuing until about 1901.

An outstanding disclosure in Heaviside's work was that series inductance could be beneficial at voice frequencies. Today we can illustrate this by means of a simplified formula such as the one referred to in Section 2.2:

$$A \doteq \frac{R}{2}\sqrt{\frac{C}{L}} + \frac{G}{2}\sqrt{\frac{L}{C}}\,.$$

What occurs, when inductance is added, is that the line impedance is increased, and a given amount of power can be transferred with less current (but at a higher potential). The result is a reduction in the series losses (first term of the equation) and an increase in the shunt losses (second term). Since the former are usually preponderant, there is a net reduction in attenuation until the inductance is increased to the point where series and shunt losses are equal. The process is somewhat analogous to the use of high-voltage lines for long-distance power transmission.

While these relationships are easy enough to comprehend after many years of experience with the transmission equations, it was not nearly so easy to appreciate Heaviside's work when it appeared, particularly when the *KR* path had been followed so long. It is likewise understandable why the practical telephone man was reluctant to accept the idea that insertion of an additional series impedance in the line could improve transmission. There were, however, a number of men capable of following the theory and comprehending its significance; two outstanding examples in the United States were Professor Michael I. Pupin of Columbia University and Dr. George A. Campbell of the Boston laboratory of the American Telephone and Telegraph Company. Along with these should be mentioned Dr. Hammond V. Hayes, Chief Engineer of AT&TCo, and John Stone Stone of his staff. Stone was probably the first member of the American telephone community whose interests were almost entirely in the theory of transmission. He, as well as Hayes, had been following the theory developed by Maxwell, Kelvin, Heaviside, and others and they were convinced that the analysis of Heaviside, recognizing the role of the magnetic field in propagation of the signal, offered the hope of solving some of their problems. Stone, as early as 1894, had suggested the use of cable conductors with continuously distributed inductance as a means for reducing attenuation, and increasing the impedance, of entrance cable in order to provide a better terminating impedance or "match" for long-haul open wire lines. In 1897 he received a patent on a bimetallic wire to accomplish this. Hayes, impressed with the potential of the theoretical work, obtained appropriations to continue it and also, in 1897, recruited Campbell to join the small laboratory force. Campbell, a graduate of M.I.T. with five years of additional training at Harvard, Paris, Vienna, and Göttingen, was the logical person to continue Stone's work when the latter resigned in 1899.

Campbell took over this work in February of 1899 and was soon convinced that it would be far more practical to concentrate the inductance in "loading" coils introduced at discrete intervals if the effects would be equivalent to distributed inductance. He promptly developed the necessary theory and in September 1899 experimentally verified it in the Boston laboratory using reels of actual cable. Experimental loading coils were used on two 24-mile cable circuits in Boston in May 1900 and on a 670-mile open wire line in July of that year. Both installations were later used commercially.

Shortly before the installation of the experimental cable loading, the American Telephone and Telegraph Company learned that Professor Pupin of Columbia University had independently worked out a theoretical solution of the coil-loading problem and had filed a patent application in December 1899. On June 19, 1900, two United States patents were issued to Pupin. The conflicting claims of the Pupin and Campbell applications resulted in extended interference proceedings which ended, on April 6, 1904, in an award to Pupin on the basis of two weeks' priority in disclosure. Before the interference action had gone far, Pupin's rights in the invention were purchased to protect the AT&TCo interest whichever way the case might be decided.

While Pupin has commonly received credit for the invention of the coil-loaded line (outside of the United States it is referred to as a "Pupinized" line), it is only fair to note that he and Campbell were working in an era when there was a great expansion in theoretical effort, many contributions being made in both Europe and the United States. Heaviside had outlined the basic principles in June of 1887 and suggested bimetallic wires in which iron was to be one of the metals to provide additional inductance, and had also suggested the use of a wire insulation impregnated with iron dust.

In the July 2, 1887, issue of *La Lumière Electrique*, the Frenchman, A. Vaschy, likewise suggested that uniformly distributed inductance would improve transmission. Again in 1889 he proposed to improve transmission "without modifying the state of the line, by introducing at certain distances . . . coils offering sufficient self-induction and insignificant resistance." But he gave no indication of the construction of such coils, or of how they should be spaced.

In 1893, Professor S. P. Thompson in England suggested techniques using inductive shunts and transformers, but these ideas did not prove of much practical value. Also in 1893 Heaviside mentioned the possible use of series coils and suggested a spacing of 1 kilometer. He specified the inductance to be used and recommended that the inductor make use of a closed magnetic circuit using finely divided iron.

Another way of adding inductance to a cable circuit was proposed by Carl Emil Krarup of Copenhagen, his idea being to wind a fine wire of soft iron around the copper conductor. A cable with this continuous loading was placed between Elsinore and Helsingborg in 1902. This method of loading was also used on three submarine cables installed between Key West and Havana, Cuba, in 1921. These cables, each about 115 miles long, were the longest and most deeply submerged undersea telephone cables in the world at the time. While continuously loaded cables have proved of some importance for submarine applications, the use of discrete coils has been almost universally preferred for loaded land cables.

Amongst these and other contributors, Pupin and Campbell deserve special recognition since they not only recognized the benefits of coil loading but also established the rules under which these benefits could be realized.[25] They showed that a continuously

[25] As might be expected, the long litigation, which finally awarded the patent to Pupin on the basis of two weeks' prior disclosure, did not end discussion of the merits of the case. An interesting and very thorough review of the history by Dr. James E. Brittain of the Georgia Institute of Technology appeared in *Technology and Culture* as recently as January 1970 and will be found very rewarding by those who desire details of the

loaded line could be simulated by a coil-loaded line when the coils were closely spaced, but that the simulation failed with increased spacing. The important factor was the length of the wave corresponding to the frequency being transmitted. With ten coils per wavelength, for example, the equivalence is quite close. Expressing this somewhat differently, for a given spacing between coils the simulation is good at low frequencies (long wavelengths) and becomes poorer with increasing frequency (decreasing wavelength) until a critical or cutoff frequency is reached where there are π (that is, 3.14) coils per wavelength.[26] Beyond this point, transmission essentially ceases. Thus the cutoff frequency is given by:

$$f_c \doteq \frac{1}{\pi\sqrt{LSC}},$$

where

L = coil inductance
S = coil spacing in miles
C = capacitance per mile.

The benefits ultimately to be realized through the introduction of loading on cable are clearly illustrated by the curves of Fig. 4-27, drawn many years later. Obviously a large transmission improvement was possible with the line when loading was designed to place the critical frequency above the important speech components. Not only is the attenuation itself reduced, but transmission becomes much more uniform over the major part of the passband, thus overcoming the undesirable distortion of non-loaded cable.

To achieve these results, however, much more was required than devising rules for coil inductance and spacing. There was at first no experience to assist telephone engineers in determining the permissible cutoff frequency, or to guide them in the actual fabrication of good coils for this application.

Campbell's 1899 tests carried the cutoff up to 11,000 hertz, but for the early commercial tests it was reduced to 4,100 hertz. Further testing resulted in 1904 in the adoption of 2,300 hertz, this remaining the standard until 1918. At that time there commenced a trend

---

controversy. Dr. Brittain strongly favors Campbell for reasons which require discussion beyond the scope of our history. Regardless of the detailed argument, there is no doubt that Campbell played the major part in developing loading in the Bell System (and the many administrations that followed Bell practices). Campbell was not only readily available as a consultant to the Bell engineers who reduced it to practice but also had provided a mathematical analysis more exact and more useful to designers.

[26] The wavelengths referred to here are those of an equivalent loaded line with inductance uniformly distributed. Near cutoff the wavelengths on a coil-loaded line are shorter than on a uniform line and the actual wavelength at the critical frequency is twice the coil spacing.

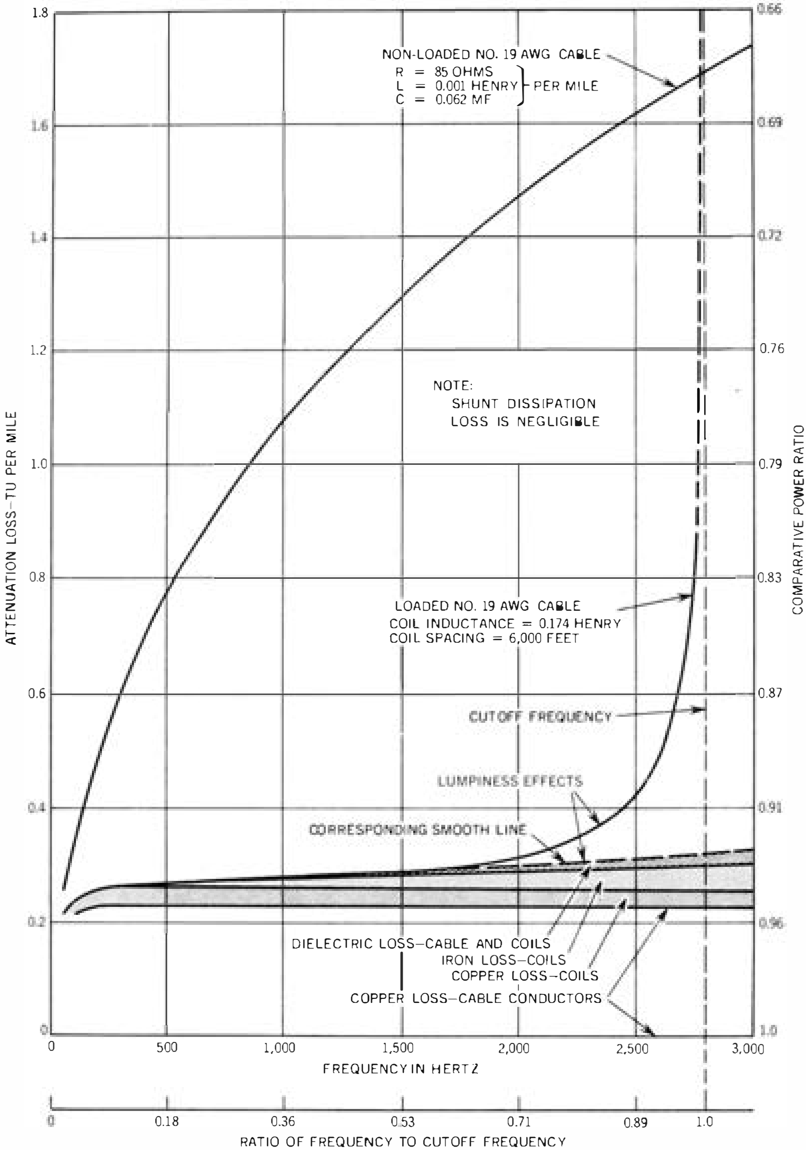


Fig. 4-27. Attenuation-frequency characteristics of loaded and non-loaded No. 19 AWG cable. [A 1926 curve. Note that attenuation is given in Transmission Units (TU), renamed decibels in 1929.] (Redrawn from Shaw and Fondiller 1926, Fig. 1)

toward higher cutoffs in order to meet the more stringent requirements of repeatered transmission and to take advantage of improved telephone instruments which responded to higher frequencies. Cutoff standards have gone through several stages which cannot be covered in a brief history, but currently all transmission systems, including those employing loading, are designed to transmit frequencies from user to user from about 300 up to 3,000–3,300 hertz with only minor distortion.

The coils used in the experiments of 1899–1900 were of the air-core solenoidal type. These were unsatisfactory for commercial use. For one thing, their resistance was not at all negligible but rather reminiscent of Heaviside's speculation that "inductance coils have resistance as well, and if this be too great the remedy is worse than the disease." Moreover, being large in size, the coils created a considerable external magnetic field which could interfere with other loaded circuits. By April 1901 a toroidal or doughnut-shaped coil had been developed using a core made up of about 10 miles of very fine (4-mil diameter) lacquer-insulated iron wire. The toroidal form confined most of the magnetic field within the core, the high permeability of iron as compared to air reduced the number of turns (and resistance) of the inductive winding, and the use of fine insulated iron wire for the core reduced eddy current losses. The inductive winding was divided into two halves, one half being introduced in series with each line wire (see Fig. 4-28). By using a single core with two windings, the mutual inductance was added to the self-inductance, further reducing the required number of turns (and the resistance).

Many improvements in loading coils have been made since 1901 (see Figs. 4-29 and 4-30), but the basic principles of the toroidal coil with split windings on an iron core designed to minimize eddy currents have continued in use. About 1916 the core was changed by using powdered iron, with the granules insulated and compressed into rings, in place of the iron wire. Later, special heat-treated iron alloys (the permalloys), having higher permeability, were introduced, further reducing the core size.

As noted previously, it was the work on loading coils that made possible the development, in 1904, of a repeating coil (transformer) which met the requirements for deriving phantom circuits. This coil opened the way to commercial application of phantoming, but its widespread use did not come about until coils could be developed for loading the phantom circuits as well as the side circuits. This was accomplished in 1910 by using a set of three coils (Fig. 4-31),

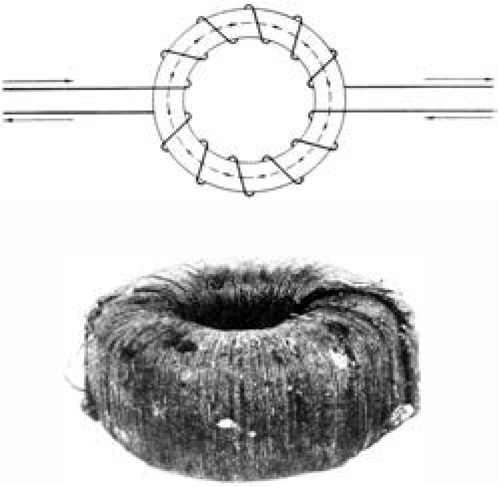


Fig. 4-28. Early loading coil and circuit.

one for each side circuit and one for the phantom. While the basic design principles were the same, refinements were needed to achieve a very high degree of balance and to make the side circuit coils essentially non-inductive to the phantom circuit and vice versa. This required close magnetic coupling between the line windings in each coil, achieved by careful interleaving of the two halves of the winding.

Early priority in the development of loading was given to open wire systems in order to extend the maximum usable communication distance. By means of loading, the attenuation of such lines could be approximately halved. The open wire application turned out to be particularly difficult, requiring protection against lightning surges, the maintenance of stable magnetic properties despite these surges, the coordination of coil spacing and transpositions, and the maintenance of low line leakage.[27] Final standards were not developed for a number of years (1905 for 104-mil lines and 1910 for 165-mil lines), but sufficient initial progress was made on the manu-

[27] Leakage was a particularly difficult problem in the case of 165-mil lines, for which the leakage component was an important part of the total attenuation, even without loading, and was aggravated by loading. This accounts for the late development of commercial loading for the large-size conductors. The use of suitable double-petticoat insulators played a large part in controlling leakage. Originally, molded porcelain was used, but the opacity of this material seemed to encourage the nesting of insects under the petticoats with consequent impairment of insulation. In a few years porcelain was superseded by glass.

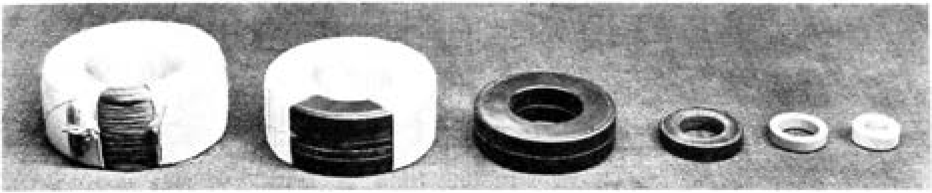


Fig. 4-29. Six stages in the development of cable loading-coil cores.

Fig. 4-30. Six stages in the development of cable loading coils.

facturing and other problems to permit the introduction of improved service on an existing 165-mil line between New York and Chicago in November 1901. In May 1902 commercial service was inaugurated on three 104-mil loaded open wire lines between Philadelphia and Chicago. By 1911 the problems of open wire loading combined with phantoming had been solved and a 165-mil phantom group between New York and Denver was put into commercial service in May. This represented about the maximum practical distance for non-repeatered telephone circuits, leaving coast-to-coast service still unrealized.

Experience gained in developing loading coils for open wire lines assisted in the development of satisfactory toroidal-type coils for the first commercial loaded cable project, a 19-gauge underground cable between New York City and Newark, New Jersey, completed in August 1902. Prior to the invention of loading, a relatively expensive 13-gauge non-loaded cable had been planned for this route. Thus from the very beginning the use of loading yielded large economies in the cable plant of the metropolitan areas.

The commercial exploitation of loaded cables went forward rapidly

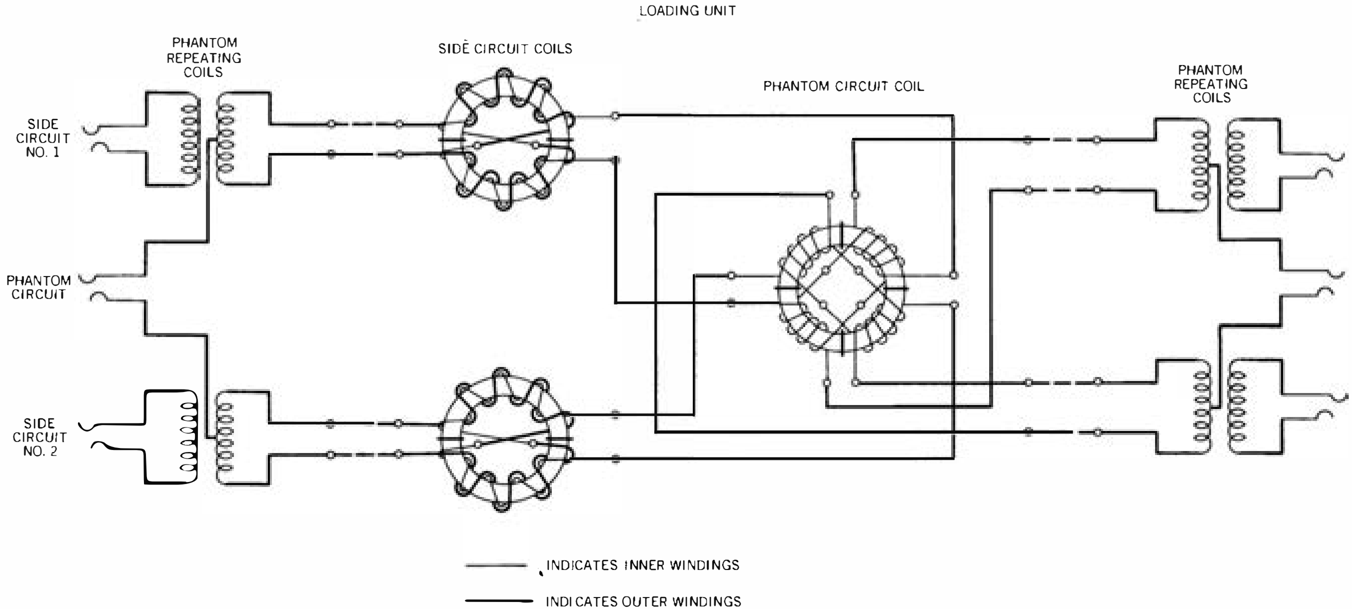


Fig. 4-31. Bell System standard method of loading phantom circuits and their side circuits.

after the New York-Newark installation, resulting by 1904 in the adoption of loading system standards that prevailed for over a decade. Figure 4-32 lists characteristics of the first standard loading systems.

A loaded cable was installed between Boston and Worcester in 1904, and cables between New York and Philadelphia and between New York and New Haven in 1906.

A sleet storm in 1909, just before the inauguration of President Taft, disrupted communication with Washington and painfully emphasized the need for storm-free facilities along the eastern seaboard. Plans for a loaded cable using No. 10 and No. 13 B and S gauge conductors began the following year. The Philadelphia-Washington section was placed in service in 1912 and the complete Boston-Washington service opened in 1913. (The latter circuit employed an early repeater of the mechanical type.)

It was soon appreciated that cable transmission was the field in which maximum benefits would be realized from the application of loading, and by the end of 1907 about 60,000 loading coils had been installed on some 86,000 miles of cable circuit. It has been estimated that by the end of 1925 about 1,250,000 coils were in use in the Bell System to load about 1,600,000 miles of cable circuits and 250,000 miles of open wire. Since the loading coil spacing is

Loading Designation	Coil Induct-ance (henrys)	Coil Spacing (miles)	Approx Nominal Imped-ance (ohms)	Attenuation at 1,000 hertz (dB per mile)		
Copper Open Wire				104-mil dia		165-mil dia
Non-loaded			650	0.075		0.033
Heavy	0.265	8	2,100	0.031		0.014
Cable				19 ga AWG (B&S)	16 ga AWG (B&S)	14 ga AWG (B&S)
Non-loaded				1.05	0.74	0.59
Light	0.135	2.5	900	0.51	0.27	0.17
Medium	0.175	1.75	1,300	0.39	0.21	0.14
Heavy	0.250	1.25	1,800	0.28	0.16	0.11

Nominal cutoff of all loaded circuits is about 2,300 hertz.

Open wire attenuation assumes dry line.

Cable attenuation assumes capacitance of 0.070 mF/mile.

Later toll cables had capacitances of 0.062 mF/mile and gauges as large as No. 10 were used giving attenuation of 0.050 dB per mile on side circuits and 0.042 dB per mile on phantoms with heavy loading.

Fig. 4-32. Characteristics of first standard loading systems.

much smaller for cable (about 1 mile) than for open wire (about 8 miles), this meant that about 95 percent of the coils were on cable. Since 1925, open wire loading has been discontinued, while cable loading has continued on the increase.

The reason for this rapid growth of cable loading is obvious. Loading greatly reduces the distortion inherent in cable and reduces the attenuation by a factor of 3 or 4 at 1,000 hertz and much more at higher frequencies. To achieve the same reduction in attenuation would require roughly a ten-fold increase in copper at an enormously greater cost.

Before the advent of repeaters, loading was essential on long open wire circuits since no other practical way existed to extend the limiting communication distance. The benefits, however, were not as great as on cable. The attenuation was decreased only by a factor of about 2, and this was accompanied by some increase in distortion. Without loading, open wire was a reasonably distortionless medium, and loading tended to introduce certain undesirable transmission effects associated with the sharp rise in attenuation near cutoff. In addition, by raising the line impedance, loading increased the problems of line leakage and crosstalk. Thus, as the repeater art developed, it was found that open wire circuits with improved quality and lower cost could be obtained by using repeaters, instead of loading, to reduce overall attenuation. Years later, beyond the time frame of this chapter, the introduction of carrier on cables provided an economic alternative to long-haul loaded cable, but even as this is written (the mid 1970s), loading continues to be the economic means for reducing attenuation on cable circuits only a few miles in length.

4.2 Longer-Term Solutions to the Transmission Problems

With the vastly superior facilities at our disposal today, it may be hard to appreciate the importance of phantoming and loading to the growth of telephony in the early twentieth century. These techniques came at a time when cost and attenuation were seriously impeding the development of a nationwide system. Within a few years phantoming reduced the cost of long circuits by roughly a third, while loading not only doubled the maximum communication distance but also greatly extended the usefulness of cable circuits, previously limited to very short distances. Perhaps the best indicator of the importance of these measures is their effective life. Phantoming remained an important technique for over a third of a century, and loaded circuits are still being added to the telephone plant in large numbers after some 65 years.[28]

[28] The production for 1974 was 14,500,000 loading coils or well over ten times the total number in the plant in 1925.

Nevertheless, these measures provided only limited solutions to the cost and attenuation problems. The phantom principle, pushed to its extreme theoretical capability, could have only doubled the number of circuits per pair of wires, and a 50-percent increase proved to be the practical limit. Loading could never, even theoretically, overcome fully the effects of attenuation; it could only reduce them. In practice, reduction by a factor of 3 or 4 on cable and 2 on open wire was about all that could be achieved.

By contrast, the two techniques of amplification and carrier transmission, which were soon to follow, provided almost unlimited potential for solving the transmission problems when the necessary technology for their application had been developed. Today, by the use of carrier methods, the capacity of a pair of wires, originally designed for a single voice channel, has been increased by a factor of 12 to 16 (much more with special structures), while amplification permits the use of conductor attenuations three or more orders of magnitude greater than those which limited transmission in the early years of the century; and it is still not evident that a limit has been reached.

We now turn to the development of these powerful techniques which originated during the second quarter century of telephony and have since come to dominate evolution of telephone transmission.[29]

4.2.1 Amplification by Mechanical Means

As we have noted earlier, it had long been recognized that if electrical speech waves, having been attenuated by passage over a long line, could be faithfully restored to their original magnitude, they could then be transmitted over an extended line. By repetition of this process, transmission might then be possible over almost any distance. What was needed, therefore, in the parlance of telephone engineers, was a source of *gain* to compensate for the attenuation or *loss* introduced by the line.

One source of gain for speech signals had long been available. The carbon transmitter was not only a converter but also an amplifier, supplying electric power some 100 or more times greater than the acoustic power impinging on its diaphragm. Thus, a logical approach to the amplification problem was to use the attenuated speech signals to move the diaphragm of a carbon transmitter, which would then yield a much greater energy output by varying the resistance of a transmitter controlling the current flow from a local source of dc power. Within a few years after the Bell invention, patents had been filed by Edison, Thompson, Houston, Hughes, Ludtge, and

[29] Only part of the story is told here since, as related in Chapter 5, these developments in wire telephony were closely related to parallel work in radio, effort in each area benefiting the other.

others for the amplification of speech by means of the receiver-transmitter combination. Thomas Lockwood, in reviewing the repeater situation in 1896,[30] noted that some 27 U.S. patents on repeaters had been issued. Most of the early patents covered minor variations of the same basic principle of using a telephone receiver or its equivalent to actuate a carbon transmitter but Edison made an important step forward by devising circuitry for using the mechanical amplifier as a true telephone repeater providing bilateral gain.[31] This was the 21-type repeater circuit which is discussed more fully in Section 4.2.4. Lockwood in his review stated that by 1896 the art had advanced to the point where technical success had been achieved, but doubted that the principles so far developed would lead to a repeater that would be of practical value in extending the limiting length of transmission. In his opinion the feebleness of the currents available at the end of a long line and the distortion introduced by the mechanical structure would restrict the output so much that only a negligible amount of line could be added.

Regardless of the opinion of the experts, the fact remained that in the early years of the twentieth century the simple approach of using the vibrations of a telephone receiver to actuate a transmitter was still the only one that appeared possible. In 1903, H. E. Shreeve was assigned the problem of developing a mechanical amplifier,[32] using the receiver-transmitter principle. It was recognized by Shreeve and his superiors that the requirements for such a device would be very severe, since it would have to be sensitive to small inputs, produce an adequate output, faithfully reproduce the input wave, and, finally, remain stable over long periods of operation. Shreeve first dispensed with the diaphragms of the transmitter and receiver. Their normal functions relate to the collecting and emitting of actual acoustic waves, and he recognized that these were unnecessary features in an amplifier for electric waves only. He then developed a mechanical linkage, so arranged that the voice currents in the electromagnetic receiving coil drove a piston-like rod which was connected directly to a plunger acting on the carbon granules of a transmitter.

A major difficulty encountered in earlier mechanical amplifier designs had been packing of the carbon granules in the transmitter over a period of time. This reduced the transmitter resistance and seriously reduced the output. Shreeve found that this packing was due to expansion caused by heat generated in the carbon chamber. In his

[30] In a series of three articles appearing in the November 14, 21, and 28 issues of the *Electrical World*.

[31] T. A. Edison; U. S. Patent No. 340,707; filed December 15, 1884; issued April 27, 1886.

[32] These devices are often referred to as "repeaters." It seems somewhat more desirable to use the term "amplifier" and retain "repeater" to describe the assembly of amplifiers and other equipment required for the practical introduction of gain in a telephone circuit.

first successful laboratory amplifier he used a stretched steel strip as a connecting link between the receiver and the transmitter, the transmitter being designed so that the expansion of its parts under the influence of heat did not subject the granular carbon between the electrodes to increased pressure. This device was successfully tested in 1904 on a circuit between Amesbury, Massachusetts, and Boston. A schematic diagram of the basic mechanical element is shown in Fig. 4-33.

It was recognized that the large inertia of the moving parts of this device impaired the quality of reproduction. The moving parts were therefore made lighter and the natural or resonant frequency of the moving system increased, resulting in an improved model which was commercially operated on a circuit between New York and Chicago from August 1904 to February 1905 (Fig. 4-34).

In spite of the success of this first commercial installation the amplifier was not standardized, since its performance was extremely difficult to reproduce. As a partial remedy the sensitivity of the transmitter was regulated by a thermostat in the form of a zinc strip, the strip being heated by a coil connected in series with the transmitter so that whenever the current through the transmitter increased as a result of incipient packing, it served to withdraw the rear electrode and thus decreased the tendency to pack—an interesting early example of feedback control. This device (Fig. 4-35) was coded the 1A "repeater"[33] in 1906.

[33] Although called a repeater at the time it was developed, it was an amplifier under the definitions used here.

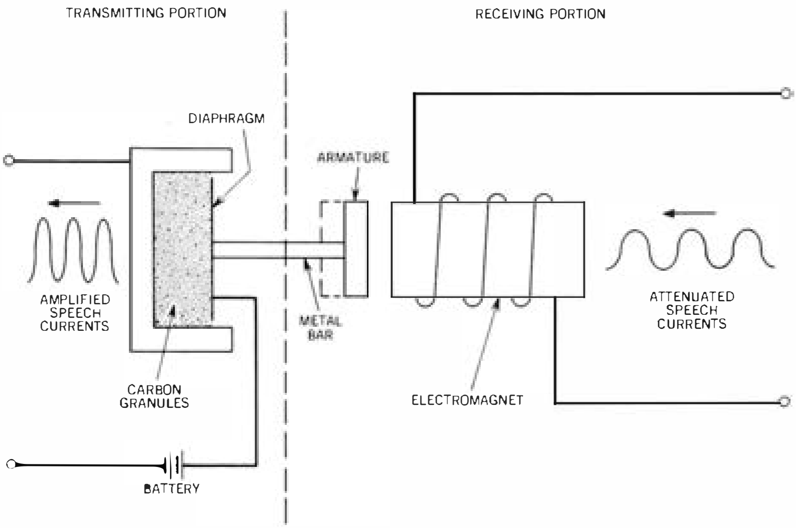


Fig. 4-33. Schematic diagram of a basic mechanical amplifier.

The 1A gave fair satisfaction and was used to some extent, but there was sufficient distortion of the speech to preclude the use of a number of them in tandem on the same line. Since three or more would have been required to achieve transcontinental distance, the search for improved characteristics was continued, resulting in the 3A (Figs. 4-36 and 4-37). This amplifier, developed in 1912 and standardized about 1914, had a higher natural frequency and incorporated additional compensating features. Particularly novel was the use of a plug-in cartridge containing those working parts which might become defective and therefore require replacement.

The 3A represented the ultimate in mechanical amplifiers and was standard for Bell System use until the vacuum tube amplifier was introduced. Like all mechanical repeaters, its performance was limited by distortion, for it was not possible at the time to build a device with the necessary efficiency that would also amplify all frequencies equally. In addition, the output did not vary linearly with the input, the sensitivity falling off rapidly at low inputs. As a consequence the signal was not faithfully reproduced and the speech quality was further degraded with each successive amplifier. Even the best mechanical amplifiers were thus limited in use to about three in tandem, so that their employment in the telephone plant never became widespread.[34] Even so, the mechanical amplifier played an important role in the evolution of telephony, for the effort going into its development did much to clarify the requirements for a more nearly ideal amplifier and stimulate the invention of circuitry for adapting the amplifier to the specific needs of telephone communication. Largely because of this early work, it was possible to move forward rapidly when a far better amplifying device became available a few years later.

4.2.2 *Electronic Amplifiers—The Vacuum Tube*

The search for amplification was not confined to the mechanical principle.[35] About 1910, Dr. F. B. Jewett, Transmission and Protection Engineer in the Engineering Department of AT&TCo, alive to the need for a new approach, told his friend Dr. Robert A. Millikan of the University of Chicago about the repeater problem and asked him to recommend a man trained in the "new physics" who could conduct research

[34] The principle was revived about 1940 and a considerable number of mechanical amplifiers were used in telephone stations for the hard of hearing (No. 332- and 324-type telephone sets). A few years earlier similar amplifiers were used in hearing aids.

[35] As early as 1907, G. A. Campbell, then at the Boston laboratory, appears to have recognized the need for investigating the use of the high-speed electron stream as a means for solving the amplification problem. The depression of that year and the subsequent reorganization which brought about Campbell's move to New York seems to have prevented the active pursuit of these ideas.

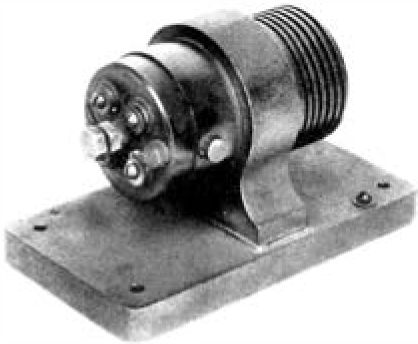


Fig. 4-34. The Shreeve mechanical amplifier (circa 1904).

Fig. 4-35. The early Shreeve 1A repeater.

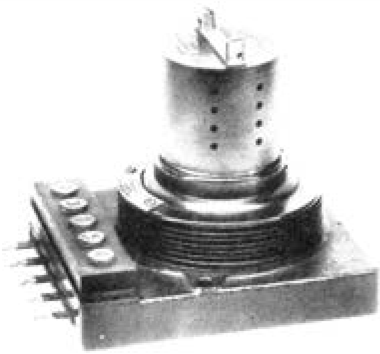


Fig. 4-36. Cartridge-type mechanical repeater. (Model 3A cartridge repeater, which became the standard type.)

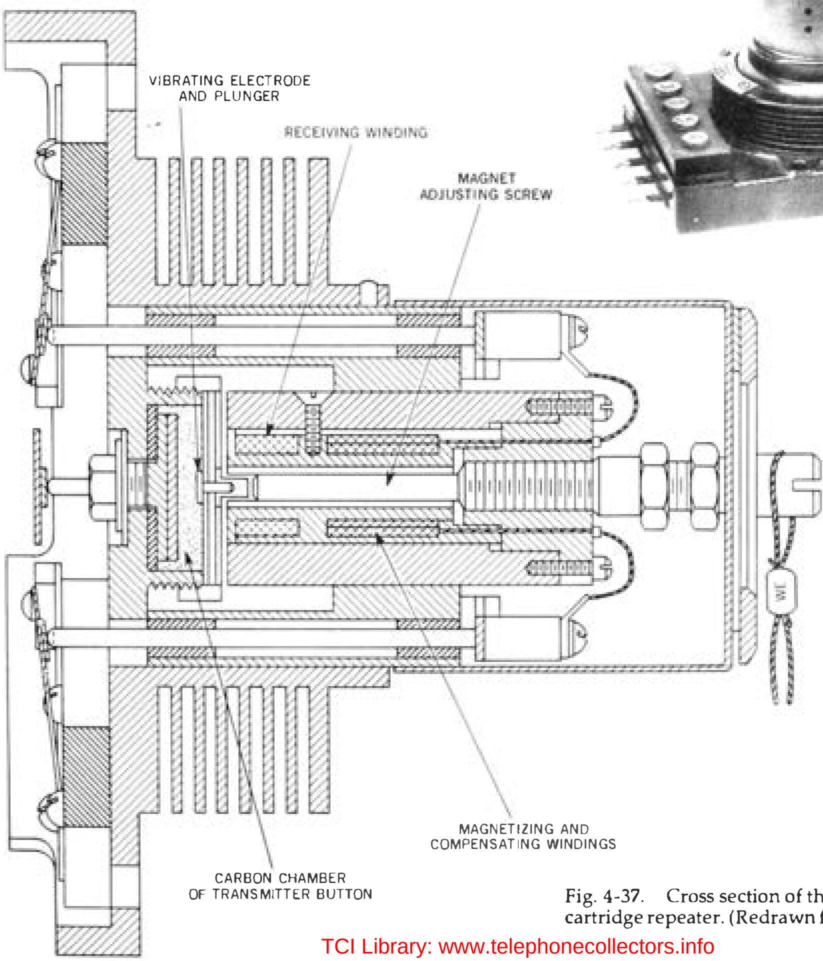


Fig. 4-37. Cross section of the Model 3A cartridge repeater. (Redrawn from Stecker 1959)

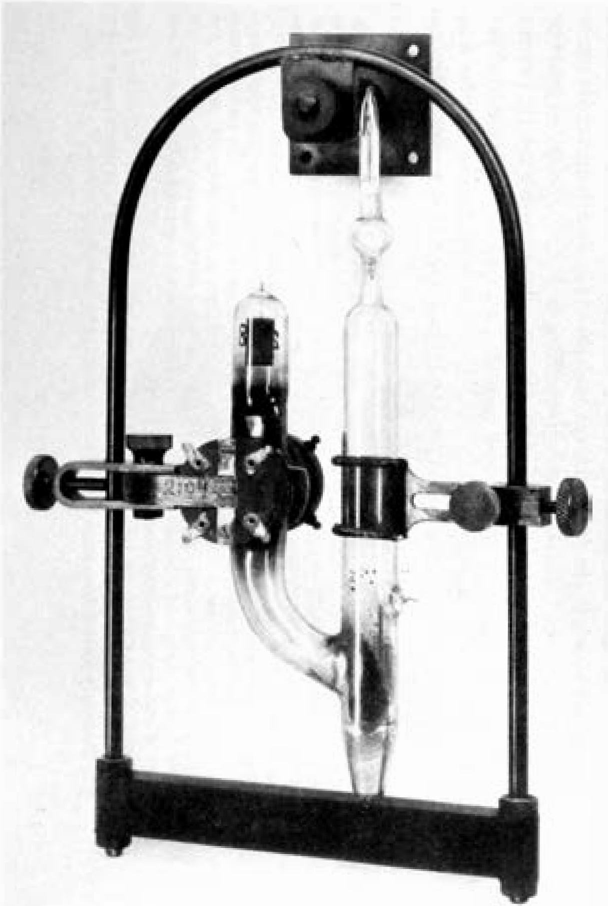
in the area. Jewett's statement of the problem, quoted many years later by Millikan, shows unusual insight:

> . . . Such a device, in order to follow all of the minute modulations of the human voice, must obviously be practically inertialess, and I don't see that we are likely to get such an inertialess moving part except by utilizing somehow these electron streams which you have been playing with here in your research work in physics for the past ten years . . .

In due course Millikan recommended Dr. H. D. Arnold, who joined the newly formed Research Branch of the Engineering Department of Western Electric in January 1911. By 1912, Arnold had developed an amplifier based on earlier work by Peter Cooper Hewitt on mercury-vapor-discharge tubes. In this device a stream of ionized molecules of mercury was controlled in deflection by the magnetic field from telephone currents flowing through a coil transverse to the discharge (Figs. 4-38 and 4-39). Though capable of amplification and fairly free from distortion, it had problems of starting and maintenance such that special engineering supervision was necessary in its use, and though units were installed experimentally on several lines, including the first transcontinental line, such amplifiers were never used commercially.

Other devices, such as the von Lieben cathode ray type of amplifier worked on in Germany, were also examined. The invention from which the solution ultimately evolved was disclosed as early as January 15, 1907, in a U.S. patent granted to Lee de Forest for a "Device for Amplifying Feeble Electric Currents."[36] Lee de Forest's invention consisted first of a heated filament (cathode) emitting electrons which were drawn to a plate (anode) charged positively with respect to the filament—as already done in the "Fleming Valve," a thermionic type of rectifier of 1904 based on the "Edison effect" known since 1883. De Forest's very essential contribution was to introduce a third or control electrode so placed as to influence the flow of electrons to the plate. In the original patent the control electrode was shown near the cathode but not in the main electron stream. Thus its influence on the electron flow was not great; but in a second patent—"Space Telegraphy"—the added electrode was shown as a grid or screen through which the electrons would pass.[37] A potential applied to this grid would, under proper conditions, markedly affect the flow of electrons (plate current) while causing but little current flow to the grid itself. Thus a small amount of input energy to the grid circuit could control a large amount of output power in the plate circuit, and since an almost "inertialess" stream of electrons was involved, postulated by Jewett as a necessity for a successful amplifier, the frequency response was no longer subject to mechanical limitations.

[36] L. de Forest; U.S. Patent No. 841,387; filed October 25, 1906; issued January 15, 1907.
[37] L. de Forest; U.S. Patent No. 879,532; filed January 29, 1907; issued February 18, 1908.



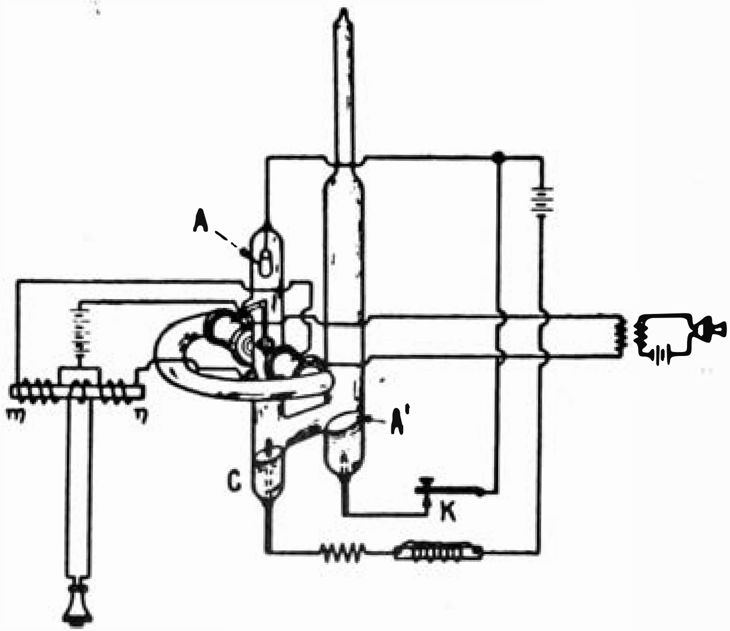


Fig. 4-39. Drawing from patent of Arnold's mercury-arc amplifier.

Fig. 4-38. Laboratory form of the mercury-arc amplifier.

For nearly five years this potential application for de Forest's device was essentially unrecognized by anyone, even the inventor. For one thing, it had been found almost immediately to be a very sensitive detector of modulated radio waves, far surpassing any previous means. We now know that this was largely attributable to its amplifying capabilities but at the time the device, called the "audion" by de Forest, was merely looked upon as an improved detector and was sold almost exclusively to enterprising wireless amateurs for this purpose, only a few going to experimental laboratories.[38]

In January 1912 Fritz Lowenstein demonstrated to Bell officials an amplifier in a sealed box using a device that later was proved to be an audion. The equipment amplified erratically and uncertainly but since the amplifying device and circuitry were not disclosed at the time, there was no basis for judging its potential on that occasion.[39] In the meantime John Stone Stone, who had been in the Boston laboratory of American Bell from 1890 to 1899 but was now on his own, learned of experiments, made by de Forest in 1912, aimed at using the audion as an audio amplifier. Stone appreciated its amplifying capabilities and, because of his background, had recognized its potential value in telephony. He and de Forest demonstrated the device as an amplifier to Bell officials on October 30 and 31, 1912, with full disclosure of the circuit used. Again the performance was erratic and the output too low for a telephone repeater, but the telephone people were impressed with its possibilities and organized a project under H. D. Arnold to study the device and the circuits to see whether its shortcomings could be overcome.

Arnold noticed immediately that a blue glow, indicative of ionized gas, occurred when the plate voltage was raised in an attempt to increase the output. Arnold, unlike others (including de Forest), believed that the presence of gas was unnecessary and that its elimination would permit the plate voltage to be raised and a larger output power obtained. Preliminary steps during the next two weeks convinced Arnold that he was on the right track, but it was not until April 1913 that a pump giving

[38] L. de Forest was an inventor almost solely interested at this time in wireless. There is rather good evidence that, in spite of the broad title of one of the patents, he was thinking almost entirely in terms of very weak wireless currents when he invented his audion. Though this preoccupation of his is understandable, it is surprising that someone in the field of telephony did not appreciate the potential of the audion before 1912. A partial explanation may lie in the state of wireless development at this time which was, to a considerable extent, dominated by promoters and stock jobbers—an environment not conducive to dissemination of technical knowledge. Robert Marriott, one of the founders of the Institute of Radio Engineers, relates in his autobiography the relief he felt in 1912 when he joined a Federal regulatory organization and entered into "a job where none of my associates sold stock."

[39] Several years later, an incidental feature in the box, namely, the "C" battery for rendering the grid negative, proved to be vital and the right to use was purchased by the Bell Company from Lowenstein.

the necessary high vacuum[40] had been obtained from Germany. By October 18 of that year, less than 12 months after the Stone-de Forest demonstration, tubes had been built and tested on commercial circuits between New York and Baltimore. In July 1914 vacuum tube amplifiers (Fig. 4-40) were used on the transcontinental line to go into commercial use six months later. The mechanical amplifiers which had been built for this service as well as the Arnold mercury-arc amplifiers were installed on a standby basis but never used commercially.

Thus the "audion" or electron tube, within a year after being brought to Arnold's attention, was developed into the long-sought telephone amplifier; but its usefulness was to extend far beyond amplification of the voice. The amplifying property made it easy to generate oscillations or electric waves of any desired frequency by coupling output to input through suitable control circuits. Already used as a "detector" or demodulator of electric signals, the device was soon found capable of serving also as a modulator.[41] Circuitry for all of these applications

[40] Irving Langmuir of General Electric also pursued the high-vacuum path in improving the vacuum tube, and applied for a patent in 1913, which was issued in 1925. After long litigation the U. S. Supreme Court declared the patent invalid on the basis that the use of a high vacuum was "a natural development of the art," citing the prior work of Arnold and others. A more complete treatment of early vacuum tube development is given in Chapter 8, Section VII.

[41] Modulation and demodulation are explained briefly in Section 4.2.5.

Fig. 4-40. Vacuum tube used in repeaters on the first transcontinental telephone line.

was rapidly developed by Arnold, van der Bijl, R. V. L. Hartley, E. H. Colpitts, R. A. Heising, and many other Bell engineers.

Chapter 5 will discuss some of this circuitry and its application to the evolution of radio telephony. Elsewhere, more will be said about the development of the crude audion into the sophisticated and reliable device that has played such a large part in the growth of electrical communication, but before we leave this subject we should mention the important contribution of H. J. van der Bijl.[42] He was largely responsible for analyzing tube performance in terms of simple parameters and establishing techniques for designing circuits with performance predictable from these parameters. This work was reported in a number of publications, beginning about 1913, and culminated in the 1920 book, *The Thermionic Vacuum Tube and Its Applications*, which guided designers for many years.

4.2.3 The Transcontinental Line

At this point we digress from our discussion of specific technical developments to note a highly important event in the evolution of telephony—the opening in January 1915 of telephone service between New York and San Francisco (Fig. 4-41), a major step in the conquest of distance. If any doubt still lingered that the Bell System concept of worldwide telephony would be achieved, it was surely dispelled later in the same year when wireless waves carrying speech, sent out from Arlington, Virginia, were received in both Paris and Honolulu. By the end of 1915, fundamental technical barriers to the problem of distance had been largely overcome.

The wireless story will be told in Chapter 5. Here we should note that the establishment of transcontinental service was a crowning achievement of 40 years of development in transmission over wires.

Most of this New York–San Francisco line was built of No. 8 (BWG) copper (diameter about ⅙ inch), weighing 870 pounds per loop mile. Supported on 130,000 poles, the weight of copper in the line totaled 2,500 tons. Loading coils were placed every 8 miles, and the line was provided with three 2-way vacuum tube repeaters. Early in 1915 three more repeater points were added,[43] and in 1918 two more for a total of eight.

The application of loading to this line reduced the usable bandwidth to about 900 hertz, giving poor intelligibility to the transmitted speech.[44]

[42] Van der Bijl returned to his native South Africa in 1920 and took a leading part in the scientific and industrial development of his country, becoming Director-General of Supplies during World War II.

[43] The original line with three repeaters was, in a sense, an experimental installation to prove the feasibility of transcontinental telephony and the vacuum tube repeater. There was, however, some commercial use.

[44] As noted previously, the theoretical cutoff frequency of the loading systems in use at the time was about 2,300 hertz, but the attenuation of the line increased rapidly with frequency and little useful power was transmitted above 1,250 hertz. Repeating coils,

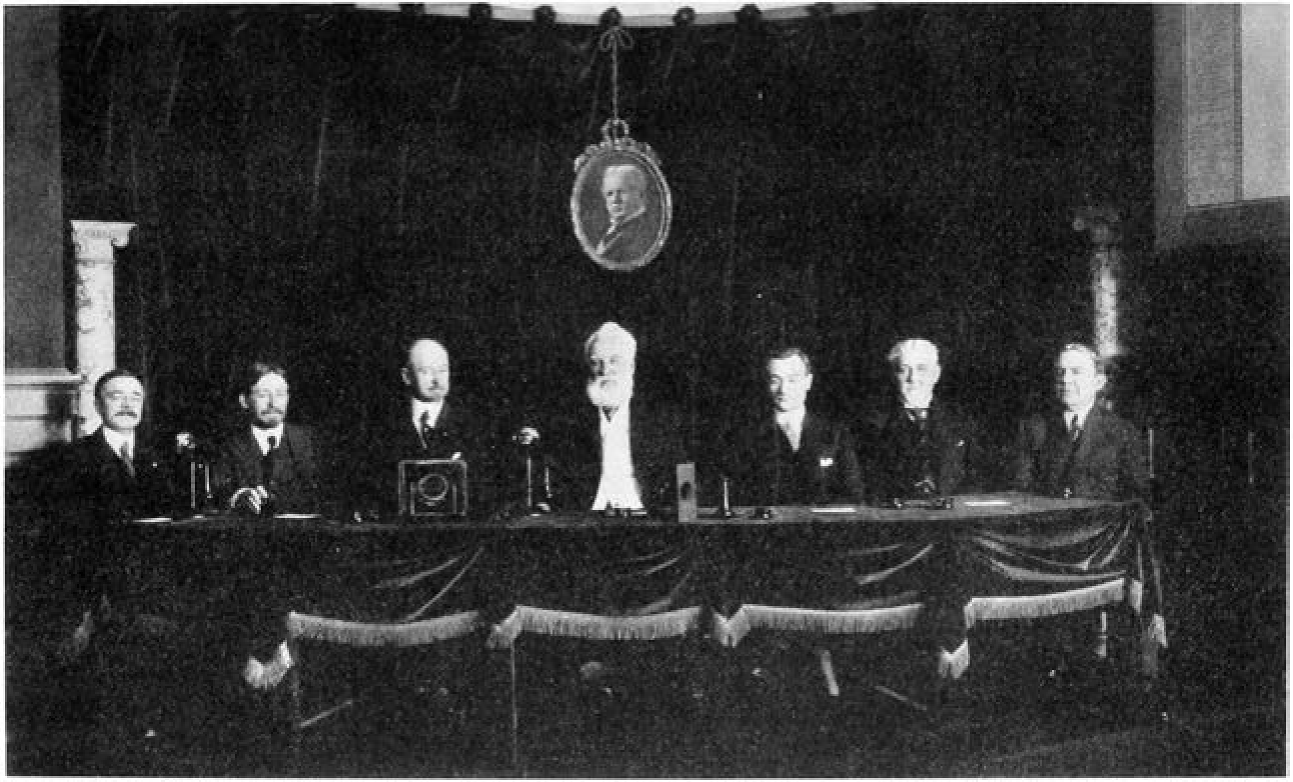


Fig. 4-41. Opening of the first transcontinental telephone line on January 25, 1915. Seated at the table in the New York City offices of the American Telephone and Telegraph Company are, left to right: J. J. Carty, AT&TCo Chief Engineer; G. McAneny, President of the New York City Board of Alderman; U. N. Bethell, AT&TCo Senior Vice President; Alexander Graham Bell; Mayor J. P. Mitchell of New York City; C. E. Yost, President of the Nebraska Telephone Company; and Comptroller W. A. Pendergast of New York City. A similar group, including Thomas A. Watson, was in the San Francisco offices of the Pacific Telephone and Telegraph Company (Mills et al. 1940, p. 30)

As repeaters were improved, it became possible to space them more closely and thus provide enough gain without the use of loading. In 1920 the transcontinental line was unloaded and equipped with 12 repeaters using 3,000-hertz filters. As a result, the transmission band was doubled, the loss cut about in half, and the speed of propagation increased by a factor of 3.5. This last factor greatly reduced the effect of echoes, which, we shall see, can seriously impair transmission on long loaded circuits.

The problem of cost was still not solved. The charge for a 3-minute call in 1915 was $20.70. Means for reducing the cost of the service, as well as improving the quality, have been the goals of transmission development ever since this epochal triumph over distance in 1915. Accordingly, after this brief recognition of a milestone in telephone history, we return to the development of technology.

4.2.4 Adapting Amplifiers to the Telephone Line

While a suitable amplifying device was the key to solving the attenuation problem, its successful application to telephony required much ingenuity. Fortunately some of the basic needs had been recognized early in the search for amplification when mechanical amplifiers offered the only prospect for introducing gain. The development of circuitry for introduction of amplification accordingly proceeded in parallel with development of the amplifying devices themselves.

One complication arose from the fact that telephone pairs were normally used for transmission in both directions at once. Thus a telephone "repeater," to be used in a conventional two-way telephone circuit, had to provide gain in each direction, while amplifiers are normally one-way devices with separate input and output terminals. The solution of this problem was complicated by the fact that if amplified currents found their way back to the input they would again be amplified, with the consequent possibility of oscillation or "singing" at a frequency determined by the circuit constants.

Several solutions to this seemingly difficult problem had been proposed. One of the first was the ingenious circuitry shown in Fig. 4-42. This was referred to as the 21-type repeater since it provided two-way transmission with one amplifying device.[45] With this circuitry the output currents were divided between the eastbound and the westbound lines so that if these were precisely alike, none of the output was returned to

---

composite circuits, etc., cut off the lower end of the speech band and limited effective transmission to frequencies above 350 hertz; hence the bandwidth figure of 900 hertz given above.

[45] The principle is essentially the same as disclosed in the Edison patent of 1886 (No. 340,707) mentioned earlier.

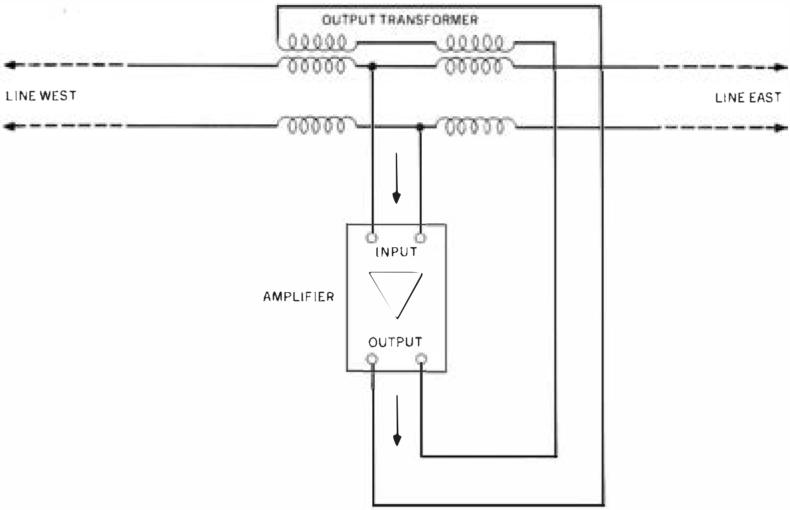


Fig. 4-42. Block diagram of the 21-type repeater circuit.

the amplifier input. If, however, the two lines were not precisely alike, some of the output was returned to the input, depending upon the degree of symmetry or "balance," and singing would occur unless the amplifier gain was reduced to the point where it was less than the loss in the path coupling output to input. In actual practice it was never possible to make two lines precisely alike and the amount of gain achievable with a 21-type repeater was severely limited.

George Campbell had studied this problem in 1912 (when mechanical amplifiers were still the only source of gain) and had provided an analysis of the balance requirements. He concluded that the 22-type[46] of repeater, shown in Fig. 4-43, would be greatly superior to the apparently simpler 21-type. This arrangement required two amplifiers for two-way transmission but there were important compensations. Each line was connected to the input of one amplifier and the output of another by means of a so-called "hybrid coil," a transformer arrangement with four points of access. When the line was balanced by an identical artificial line, or "network," connected as shown, there was equal division of the amplifier output between the real line and the network and no transfer of energy across the hybrid to the input of the other amplifier. One advantage of this arrangement was that since each line was balanced by

[46] The 22-type repeater was not new. It had been invented by W. L. Richards in 1895, long before practical amplifiers were available. At first it was believed that the balancing networks would have to be complex multisection lines, but Campbell found a simple solution using a basic 3-element network plus a few "building out" components.

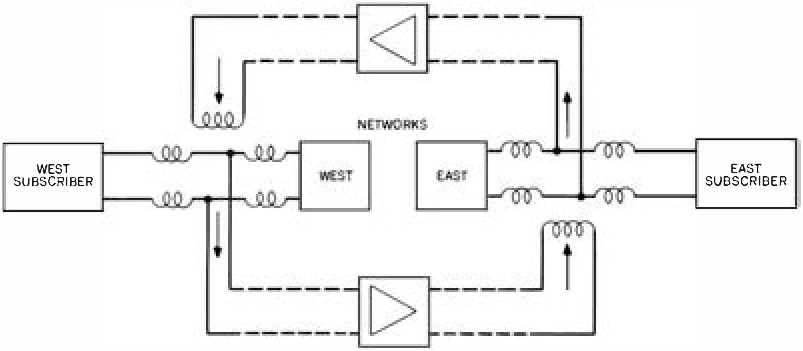


Fig. 4-43. The 22-type repeater. (Redrawn from Bouton 1938, p. 116)

its own network, the two lines did not have to be alike (e.g., one could be cable and the other open wire). It also turned out that the allowable unbalance of each line was about double that with the 21-type repeater, and singing did not occur with even very bad unbalance on one line if balance was adequate on the other.

Even so, practical requirements limited the balance (and gain) rather severely. Campbell made numerous suggestions for improving repeater operation, including one which must have appeared unbelievably radical. He pointed out that the path of the 22-type repeater circuit shown dashed in Fig. 4-43 could be stretched out and even include additional line and one-way amplifiers, as indicated in Fig. 4-44. A two-path voice-frequency telephone circuit, such as this, is usually spoken of as a 4-wire circuit. The apparent extravagance in the use of line conductors by 4-wire circuits is offset by the fact that much higher repeater gains are allowable than in 2-wire circuits, and accordingly smaller wires may be employed or the amplifiers may be spaced farther apart. This idea was tested with mechanical amplifiers in 1913 and was used commercially in April 1915 with electron tube amplifiers on a 450-mile cable circuit between Boston and Washington. Because of the high gains possible, and the consequent large saving in copper, this 4-wire technique was to prove the answer for many years to the economic use of essentially storm-proof cable for long-distance service. By 1925, 4-wire cable circuits were introduced between New York and Chicago and by 1933 New York–Dallas cables were in service. By this time the technical problems had essentially been solved for transcontinental-length circuits, but before these could be introduced a better solution was provided by other techniques using carrier multiplex.

Although 4-wire transmission was the preferred technique for long-distance cables, the use of 2-wire with 22-type repeaters was preferable

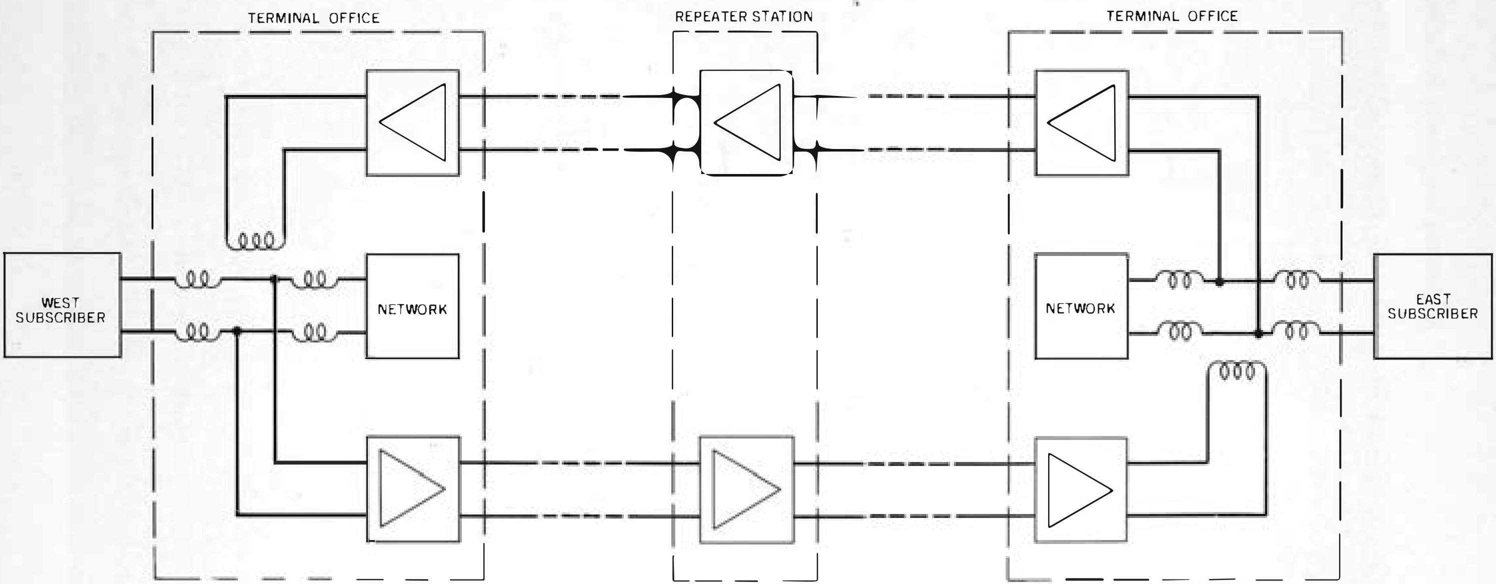


Fig. 4-44. Four-wire circuit for a two-path voice-frequency circuit. (Redrawn from Bouton 1938, p. 117)

for open wire since the conductor size, largely dictated by physical conditions, could not be reduced sufficiently to pay for the extra pair of wires.

Present-day readers familiar with the extreme miniaturization of circuitry achieved in recent years will be interested to see the amount and kinds of apparatus required for a repeater with the technology available in the first quarter of the twentieth century. The wall-mounted repeaters used in 1915 on the transcontinental line are pictured in Figs. 4-45 and 4-46. In 1917 these were replaced by the floor-type repeater of Fig. 4-47, which shows the beginning of the relay rack type of mounting. In all of these it is difficult at first glance to find the electron tube. While this was the vital element without which a successful repeater could not be built, it was physically a very small part of the complex apparatus required to make it function in a working telephone circuit. For many years the physical bulk and the cost of a repeater resided largely in the devices for deriving one-way circuits, for controlling gain, for protection against lightning, for bypassing signaling and telegraph, etc. Some reduction in size was achieved by 1920 when the Reading-type repeater of Figs. 4-48 and 4-49 was introduced, and in 1923 the Type 22A-1 vacuum tube repeater (Fig. 4-50) was put in production. This repeater, which could be used with either open wire or cable, remained the standard for many years. It should be noted that, at about this time, it became the practice to mount apparatus required for different functions on separate racks in order to achieve greater flexibility in installation and modification with plant growth. Accordingly, the photograph of the 22A-1 repeater does not include all the types of equipment (e.g., telegraph bypass, etc.) shown in the earlier photographs.

While the adaptation of a one-way amplifier to the two-way telephone line was the outstanding initial problem in the use of gain, others were soon to become apparent. Some of the more outstanding were:

(*i*) Equalization of frequency response
(*ii*) Regulation of gain
(*iii*) Control of echo.

As we have noted, the increase in attenuation with frequency which is characteristic of cable circuits produced highly objectionable distortion of the voice sounds, an effect greatly reduced by the use of loading; yet even small variations in the attenuation-frequency characteristics could become important in the case of very long circuits in which undistorted or "flat" amplification was employed to reduce the overall loss. On open wire lines the solution was easy. The distortion was introduced largely by the loading itself and the return to non-loaded lines, when repeaters became available, provided not only less distortion but also a less expensive line, since repeaters were cheaper than loading. This solution was not possible with cable because of its inherent distortion. Two means were resorted to: one, an increase in the cutoff frequency of the loading

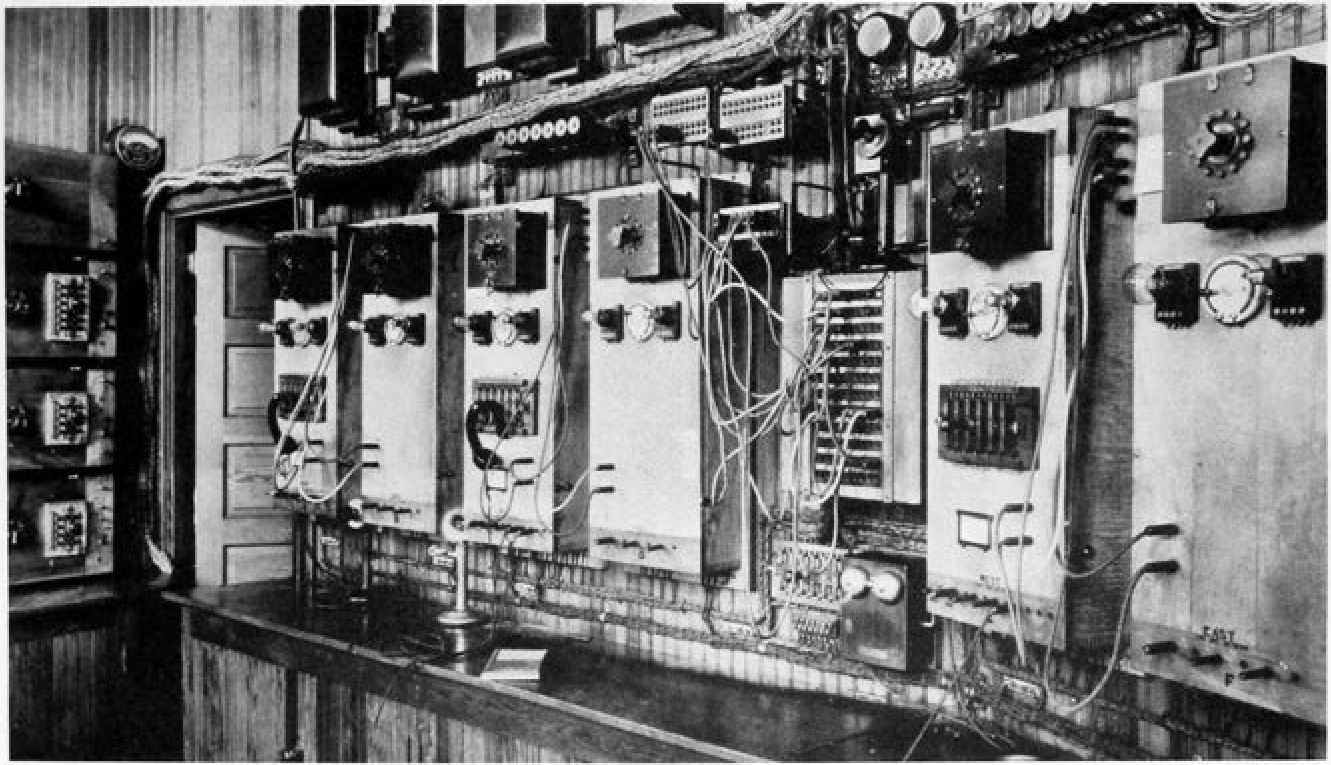


Fig. 4-45. A group of the first vacuum tube repeaters installed at Brushton, Pa., on the transcontinental line.

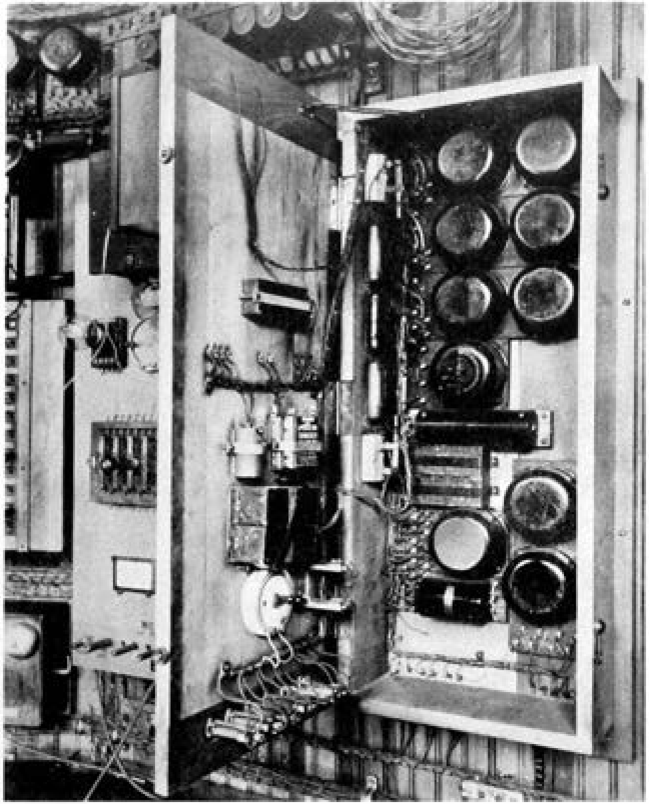
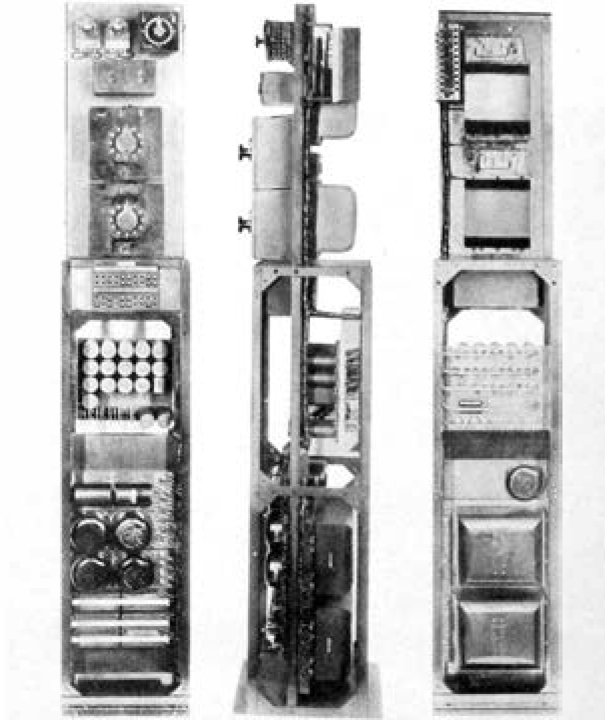


Fig. 4-46. One of the repeaters of Fig. 4-45 with the cover removed.



(a) (b) (c)

Fig. 4-47. Floor-type repeater which replaced Fig. 4-45 units in 1917. Equipment for one 2-wire circuit of loaded or non-loaded open wire line (22-type repeater circuit). (a) Front view. (b) Side view. (c) Rear view.

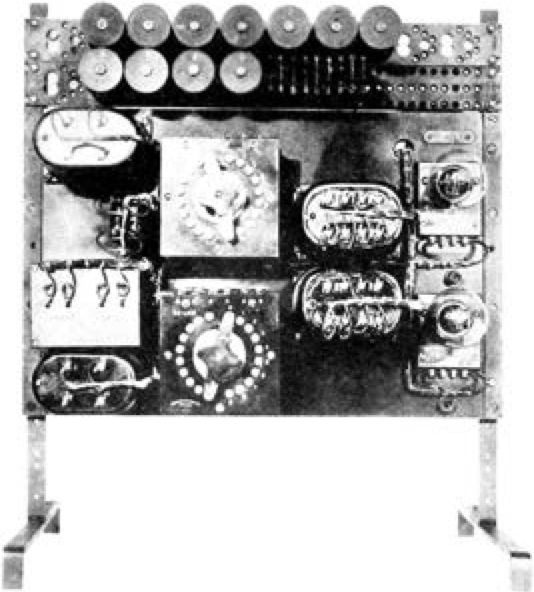


Fig. 4-48. Front view of Reading-type repeater. Equipment for one 2-wire circuit of loaded cable (22-type repeater circuit).

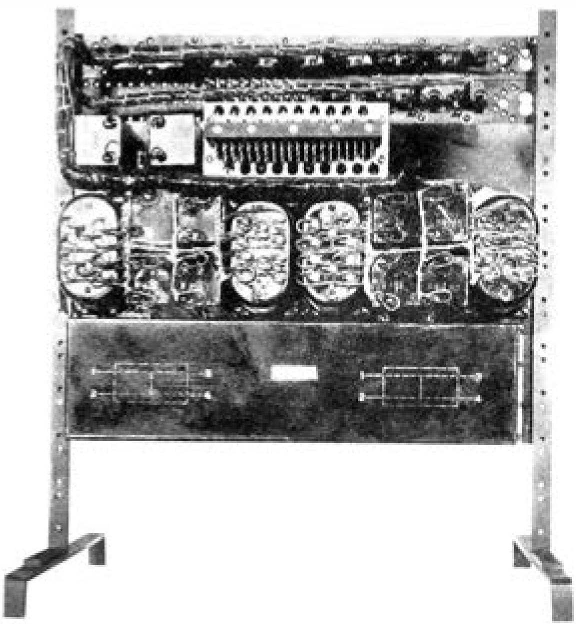


Fig. 4-49. Rear view of Reading-type repeater.

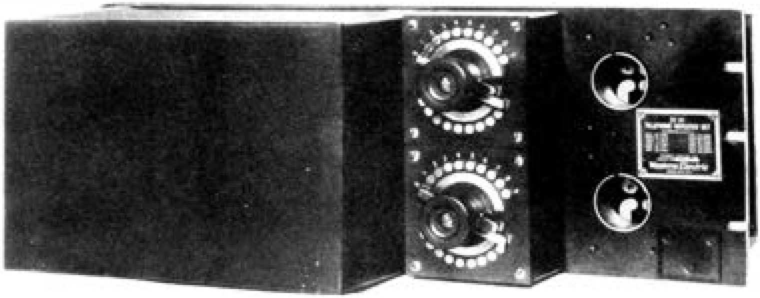


Fig. 4-50. Type 22A-1 vacuum tube repeater.

system, giving a flatter attenuation characteristic; the other, introduction (in the repeater) of equalizers having attenuation characteristics inverse to those of the line. In this way the repeater gain could be made equal at all frequencies to the line loss. The design of these equalizing networks, including filters to prevent amplification at frequencies where line balance was difficult to achieve, was one of the outgrowths of Campbell's work on the mathematical theory of the loaded line.

The attenuation of any line is dependent on temperature, rather large relative changes (5 to 10 percent) being possible from winter to summer or even between day and night. For lines with small attenuation these changes are unimportant, but as amplifiers permitted the use of very large attenuations, a 5- or 10-percent change with temperature became prohibitive on an absolute basis. The solution was the development of automatic regulation, i.e., devices to adjust the repeater gain to compensate for the effects of temperature. An early solution, introduced in 1923, was the pilot wire regulator. This device measured the resistance of a "pilot" wire in the cable and by means of a simple "computer" determined the gain required and adjusted all the repeaters using the cable accordingly.

The problem of echo was the most complicated of all. At any point in a circuit where an electric wave meets a discontinuity, a portion of the wave is reflected back to the sending end and is heard by the talker as an echo. These discontinuities may occur anywhere along the line due to irregular spacing of loading coils, mismatch between lines and balancing networks, etc., but the most common source is the impedance mismatch between the line and the terminating telephone station. The echo is not troublesome if it is small or is not greatly delayed in its return. For practical conditions any echo which returns to the talker in less than 40 or 50 milliseconds is not objectionable. For circuits which propagate waves near the speed of light (186,000 miles per second) echoes are not troublesome for circuit lengths of 3,000 to 4,000 miles. However, when a circuit is loaded to reduce attenuation, one of the attendant penalties is a reduction in speed of propagation to 10,000–20,000 miles per second. Thus in a circuit 1,000 miles long having a 10,000-mile speed, 100 milliseconds are required for the wave to travel to the end and another 100 milliseconds for the echo to return. Such a delayed echo, if strong, can make conversations almost impossible. This problem was studied extensively as long-distance circuits were developed, and a practical solution was achieved in the early 1920s with the development of the echo suppressor. Later, as we shall see, the echo difficulty was greatly reduced with the use of high-speed carrier circuits and did not again become serious until recent years when submarine cable and satellite circuits extended communication to distances of many thousands of miles.

The action of an echo suppressor can best be described in connection with a 4-wire circuit. In an oversimplified form the arrangement consists of relays so connected that the return path of the 4-wire circuit is short-circuited whenever the forward path is in use (Fig. 4-51). This was accomplished by bridging two similar high-impedance amplifier-detectors across the two sides of the 4-wire circuit, each amplifier-detector having an associated relay which operates whenever an alternating voltage of sufficient strength is present (Fig. 4-52). The operation of

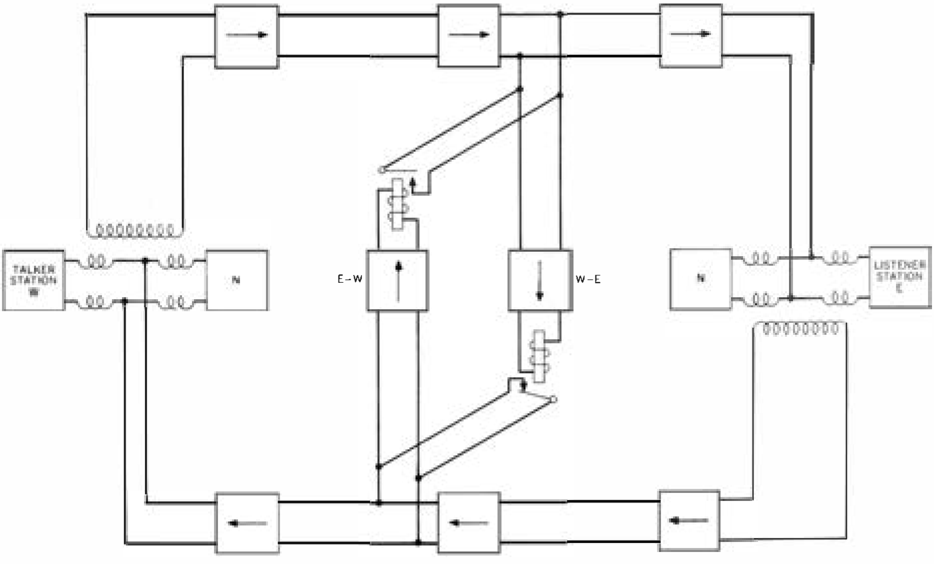


Fig. 4-51. Basic operation of echo suppressor in a 4-wire circuit.

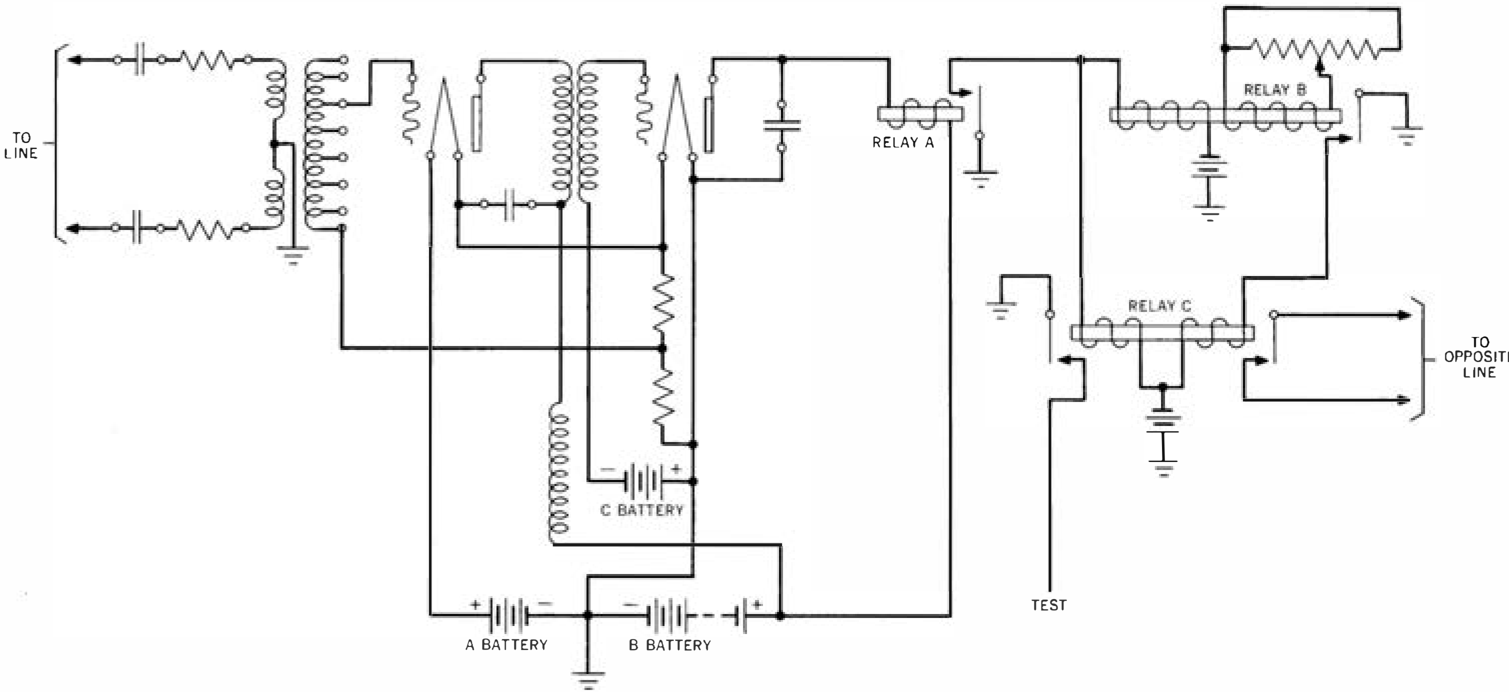


Fig. 4-52. Simplified circuit diagram of one-half of a 4-wire echo suppressor, corresponding to either block E-W or W-E of Fig. 4-51.

either relay, due to strong speech on one side of the 4-wire circuit, places a short circuit across the opposite side, thereby not only blocking passage of any signal on that side but also insuring that the other relay will not be actuated and short-circuit the first side. A similar effect could be achieved by opening instead of closing appropriate circuits.

A number of echo suppressors were placed on 4-wire lines in Harrisburg, Pennsylvania, during 1924 (Fig. 4-53). Techniques were also developed about this time for applying echo suppressors to 2-wire lines.

The basic method of echo suppression outlined above continues to be used and has been successful in almost completely eliminating echoes, but does introduce some transmission impairments of its own which tend to be characteristic of voice-operated devices. One shortcoming of the echo suppressor is the limiting of transmission to only one direction at a time. Thus there is difficulty in breaking into a flow of speech once started, so that the natural two-way flow of conversation is impeded. In more recent years some rather sophisticated techniques have been introduced so that the modern echo suppressor has improved break-in characteristics and better behavior in several other respects.

With echoes under reasonable control, and with means available for amplification, equalization, and regulation, it would have been technically possible to provide cable circuits over transcontinental distances—a desirable objective in the interest of maintaining trouble-free service. The problem of costs, however, had not been fully solved in spite of the many economies made possible by the electron tube amplifier.

It will be recalled that the initial impact of the vacuum tube telephone repeater had been on open wire transmission, where it made transcontinental telephony possible without using conductors so large as to be completely uneconomic. The repeater was now proving its effectiveness in the economics of cable transmission; by the 1920s, even the longest cable circuits employed 19-gauge (36-mil diameter) conductor in place of the earlier 13-gauge (72-mil diameter) conductors. But this saving was offset to some degree as 4-wire transmission began to replace 2-wire for greater stability on the longer cable routes, so that even with phantoming, only three-fourths of a two-way telephone circuit was available per pair of wires. As repeaters were reduced in cost, some further economies could be made by trading off additional repeaters for reduction in wire diameter, but potential for such improvements was definitely limited by physical restrictions on wire size. A satisfactory solution of the cost problem required transmission facilities much less costly than promised by this approach; facilities that would permit rates so low that long-distance calls would become a commonplace, not a rarity.

What was needed was a fundamentally new transmission technique

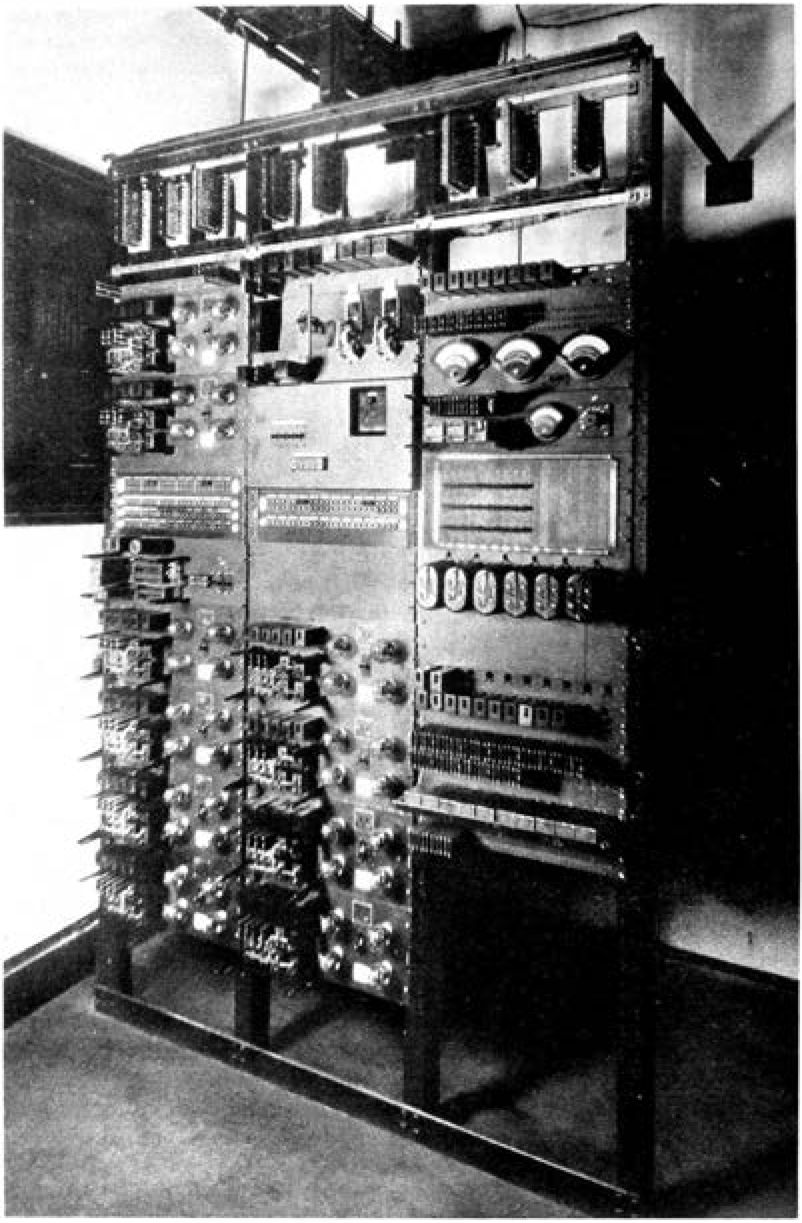


Fig. 4-53. An early echo suppressor.

that would permit a more efficient use of the conductors. This advance came in the form of *carrier multiplexing*, applied first to open wire and later to cable. This development, now to be described, permitted the use of a single set of conductors for a number of telephone conversations, ultimately a large number. It proved to be a technical innovation of far-reaching consequence in the reduction of transmission costs and the resulting stimulation of toll usage.

4.2.5 *Carrier Multiplexing*

The basic concept of increasing line capacity by carrier multiplexing is older than telephony. It will be recalled from Section II of Chapter 1 that Bell was experimenting with telegraph multiplex (which he called the "harmonic" telegraph) when he recognized the possibility of transmitting the voice.

The primitive form of carrier multiplexing involved in the telegraph application consisted merely in permitting the telegraph key to break up the flow of an intermittent or alternating current, instead of steady or direct current, into the appropriate coded sequence of impulses. The alternating current or tone is referred to as the "carrier" of information, and the key or other device is said to "modulate" this carrier. By using different tones for different messages, they can all be sent over the same wires (i.e., multiplexed) and decoded at the distant end if means exist for filtering or separating the tones.

This was all pretty well appreciated in the 1870s and 1880s and various schemes for telegraph multiplexing were proposed by Gray, Bell, Van Rysselberghe, Edison, Mercadier, and others. These early schemes depended on the use of mechanical resonance[47] for generating and separating the tones, but in the early 1890s Pupin at Columbia, Stone at American Bell, and Leblanc in France had made somewhat more sophisticated proposals employing electrical resonance or tuned circuits. Pupin was adjudged the earliest inventor in the United States. His patent included not only the use of electrical resonance for discriminating between the received tones but also the use of a "detector" to convert the received pulses of tone into dc pulses.

While the early work on multiplexing was concerned with telegraphy, the Frenchmen Hutin and Leblanc pointed out as early as 1891 the possibility of using the same basic idea for telephony. In the United States, John Stone Stone was working along the same lines and

[47] The mathematician Maxwell had perceived, as early as 1868, the analogy between electrical resonance in a circuit containing capacitance and inductance and the already well-known mechanical resonance of a system involving a mass and a spring; but at that early date there was no profession of electrical engineering and, as we have already observed in connection with Heaviside and his work in transmission theory, there were few channels of communication between the analysts and the practitioners.

tested a system in the laboratory of the American Bell Telephone Company in 1894. Thus, by the middle 1890s the basic requirements for carrier multiplex telephony were known, but the devices to make it practical were unavailable.

The devices needed were: first, a source of continuous tone (or carrier) at a frequency well above the maximum voice frequency conveyed; second, a "modulator" to impress the voice currents on the carrier so that it could convey the intelligence; third, a means for separating the various modulated carriers at the receiving end; and finally, a means for demodulating the received carrier, i.e., deriving the voice wave originally impressed on it. In addition, it was necessary to adapt telephone lines for carrying the high frequencies involved so as to overcome the effects of attenuation and interference (or crosstalk) between pairs. Both of these effects were more serious at carrier frequencies than at voice frequencies.

In the experiments carried out in the years up to about 1912, carrier-frequency currents were mainly generated by electric arcs or high-frequency alternators, these currents then being modulated by a carbon transmitter and demodulated or converted to audio by a crystal or other type of rectifier. Frequency separation was accomplished by tuning resonant circuits. These, it will be recognized, were the techniques of early wireless telephony, and carrier telephony as contemplated at this time was often called wired wireless, much of the work being done by those primarily concerned with wireless communication such as Ernst Ruhmer in Europe (1908) and Major George O. Squier in the United States (1910–11).

None of these components could be called really practical. Arc oscillators were not very stable and alternators were ponderous. The carbon transmitter was a very inefficient modulator for high-frequency currents, and crystal detectors were highly unstable. Resonant circuits could be built for selecting the narrow bands of frequencies required for manually keyed telegraphy; but although modulation theory had not been developed, there was general awareness that a carrier modulated by the voice could no longer be considered as a single frequency but involved a band of frequencies related to the frequency spectrum of the voice itself. This meant that conventional tuned circuits sufficiently selective to prevent interference between the several speech channels would inevitably introduce large amounts of distortion, due to discrimination within a desired band, unless the carrier frequencies were widely separated. Finally, the attenuation of telephone lines at the higher frequencies prevented use over practical distances without amplification.

The state of the carrier art in 1912 is clearly brought out in a contemporary paper by John Stone Stone.[48] Stone, who earlier had worked at the Boston laboratory of American Bell, was at this time an independent worker with a broad theoretical background and long communication experience. He was an early enthusiast for carrier multiplex, his paper opening with the statement, "A new art has been born to us. The infant art of high-frequency multiplex telephony and telegraphy is the latest addition to our brood of young electric arts." His paper showed a good understanding of the promise of this new art but reading it in retrospect we can also see quite clearly the limitations of the technology proposed for implementing the principle. Most of these limitations have been mentioned in the previous paragraph but the paper gives emphasis to the importance of the Campbell wave filter, the modulation theory developed by Carson, and the critical problem of crosstalk between pairs.

By about 1914, concerted attacks on these problems in the Bell System were commencing to bear fruit. The electron tube quickly extended its versatility and became a stable device for tone generation (oscillation), for amplification, and for modulation and demodulation.[49] Van der Bijl and Heising in the Western Electric Engineering Department showed how the non-linear portion of an electron tube characteristic could be utilized to modulate and demodulate. The problem of selective transmission of frequency bands was solved by G. A. Campbell's invention of the wave filter. Continuing on in the Engineering Department of AT&TCo after his work on loading, Campbell by 1910 had devised various types of selective circuits which could limit the transmitted frequencies to those below a specified frequency, or above a specified frequency, or within a particular range of frequencies. These circuits, made up of inductances and capacitances, were referred to respectively as low-pass, high-pass, and band-bass filters. A typical band-pass filter used in an early carrier system is shown schematically in Fig. 4-54, together with its frequency transmission characteristics. This particular filter was designed to transmit a band of frequencies about 2,000 hertz wide with little distortion or non-uniformity and to offer high attenuation at any frequencies 1,000 hertz or more outside of this band. It was thus possible, by using a set of such filters, to have carrier channels close together with little mutual interference, a result quite unattainable with circuits previously known. It would thus be no exaggeration to

[48] "The Practical Aspects of the Propagation of High-Frequency Electric Waves Along Wires," published in the *Journal of the Franklin Institute* for October 1912.

[49] Later, copper-oxide rectifiers were to replace the electron tube in the modulation and demodulation functions in carrier systems, but the electron tube was used in the first commercially successful systems.

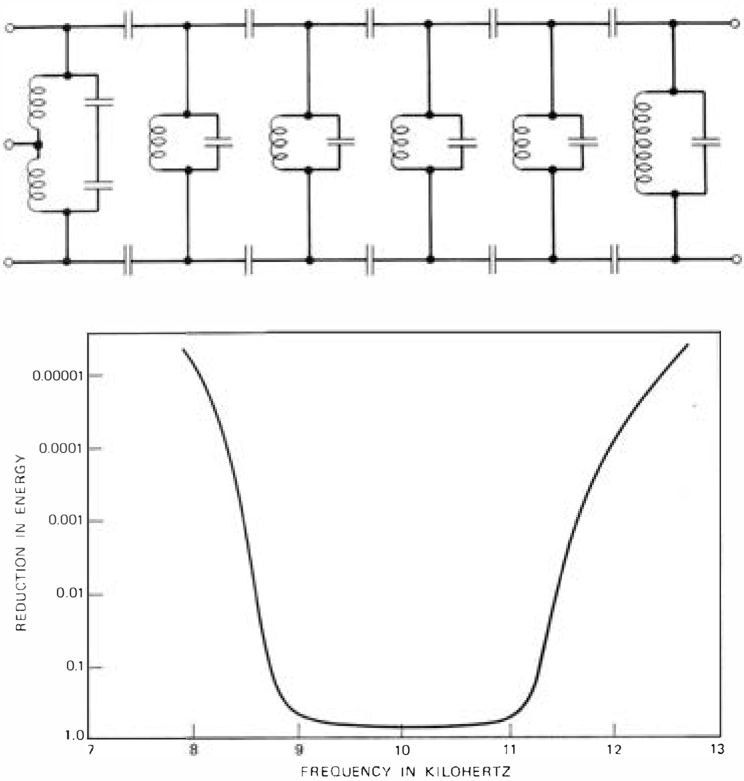


Fig. 4-54. Band-pass filter. (Redrawn from Colpitts and Blackwell 1921, Fig. 14)

state that the wave filter was fully as essential as the electron tube to the successful development of carrier transmission.[50]

Concurrently with these achievements, a better understanding of modulation and the modulated-wave frequency spectrum was developed. The laboratory notebook of a young Western Electric engineer, C. R. Englund, for August of 1914 shows the geometrical relationships of carrier and sideband waves as worked out by him, and in 1915 J. R. Carson published a mathematical analysis of the modulation and demodulation process, showing that when a carrier c and a voice wave v were passed through a non-linear device, the output contained not only the input frequencies of c and v but also

[50] As this is written, in the mid 1970s, the electron tube has been entirely superseded by electronic solid-state devices, but the wave filter in highly developed form remains an essential component of all frequency-division carrier systems.

"sidebands" of frequencies on each side of the carrier. As illustrated in Fig. 4-55, the lower sideband contained the frequencies $c - v$ and the upper sideband $c + v$. Additional frequencies might also be inadvertently generated but could be eliminated by wave filters, leaving only the desired carrier and its sidebands. Carson showed that when the carrier and its sidebands were again passed through a non-linear device (detector or demodulator) the original voice frequencies were included in the output along with other, undesired, frequencies which could be filtered out. Carson further showed that it was not necessary to transmit the carrier and both sidebands, one sideband being sufficient to convey the full information. At the receiving end, this single sideband could be demodulated to produce the original voice wave by supplying a local carrier of the same frequency as the original.[51] This type of carrier transmission (known as single sideband, carrier suppressed) is highly advantageous, requiring only half the frequency space of double sideband and simplifying the design of repeaters, since suppression of the carrier greatly reduces the electrical power on the line. When this was understood, modulator circuits were devised for balancing out the carrier, the ever-useful wave filter then being employed to suppress the unwanted sideband. It can thus be seen that with single-sideband, carrier-suppressed transmission, the modulation and demodulation processes are, in effect, means for frequency shifting. Referring to Fig. 4-55, it can be seen that the upper sideband is the same as the original voice wave but with each frequency component shifted upward in frequency by the frequency of the carrier. Thus it was possible to take a number of voice channels, each necessarily arriving over its own pair of wires, and shift each one by a different amount so that they were spread out

[51] Actually, with single-sideband transmission, the original voice is reproduced satisfactorily if the resupplied carrier is within a few hertz of the original carrier.

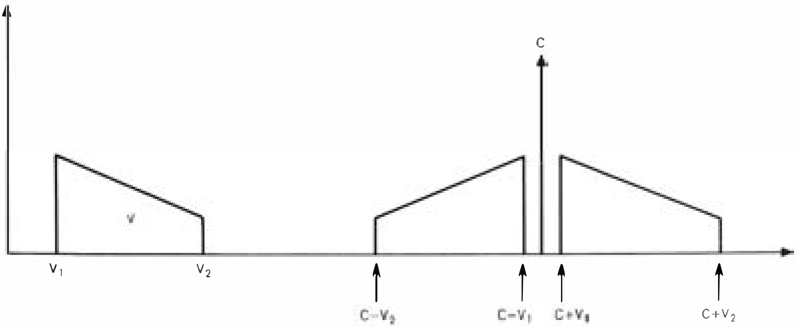


Fig. 4-55. The modulation process.

in frequency with no overlap as illustrated in Fig. 4-56. They could then all be transmitted over a single circuit and amplified together in common repeaters, to be eventually separated out by wave filters and the original waves derived through demodulation. This is the power of carrier multiplex.

While these basic contributions were being made, the development of working systems was being actively pursued. By 1914, simple carrier circuits, of which Fig. 4-57 is representative, were successfully tested in the laboratory, after which a complete multiplex telephone system was built for operation over an artificial line. This laboratory apparatus (Fig. 4-58) was then installed in 1917[52] on an experimental basis at Maumee, Ohio (near Toledo). This equipment was used with highly successful results over a specially transposed commercial line to South Bend, Indiana, at which point a repeater was installed and the circuit looped back to Maumee.

The first commercial carrier system, installed between Baltimore and Pittsburgh in 1918, provided four two-way carrier channels above the voice channel on open wire pairs in the frequency range between 5 and 25 kHz. This installation used the Maumee equipment at Pittsburgh, but new and more compact apparatus was built for Baltimore and was subsequently standardized as Type A carrier. One of the longest Type A installations was between Chicago and Harrisburg. The terminal installation at the former location is shown in Fig. 4-59.

In the Type A system, each channel used the same frequency for both directions of transmission, the necessary directional discrimination being secured by hybrid-coil balance at terminals and repeater stations. The lower sidebands of carrier frequencies at 10, 15, 20, and

[52] Somewhat earlier, in January of 1917, a carrier multiplex telephone system operating over radio channels was installed by the Bell System on several U. S. Navy ships.

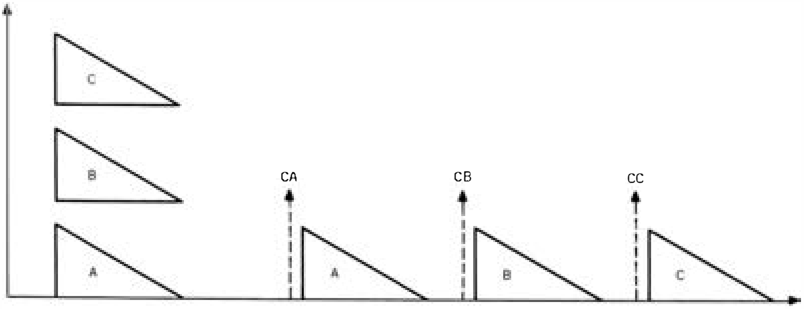


Fig. 4-56. Frequency translation.

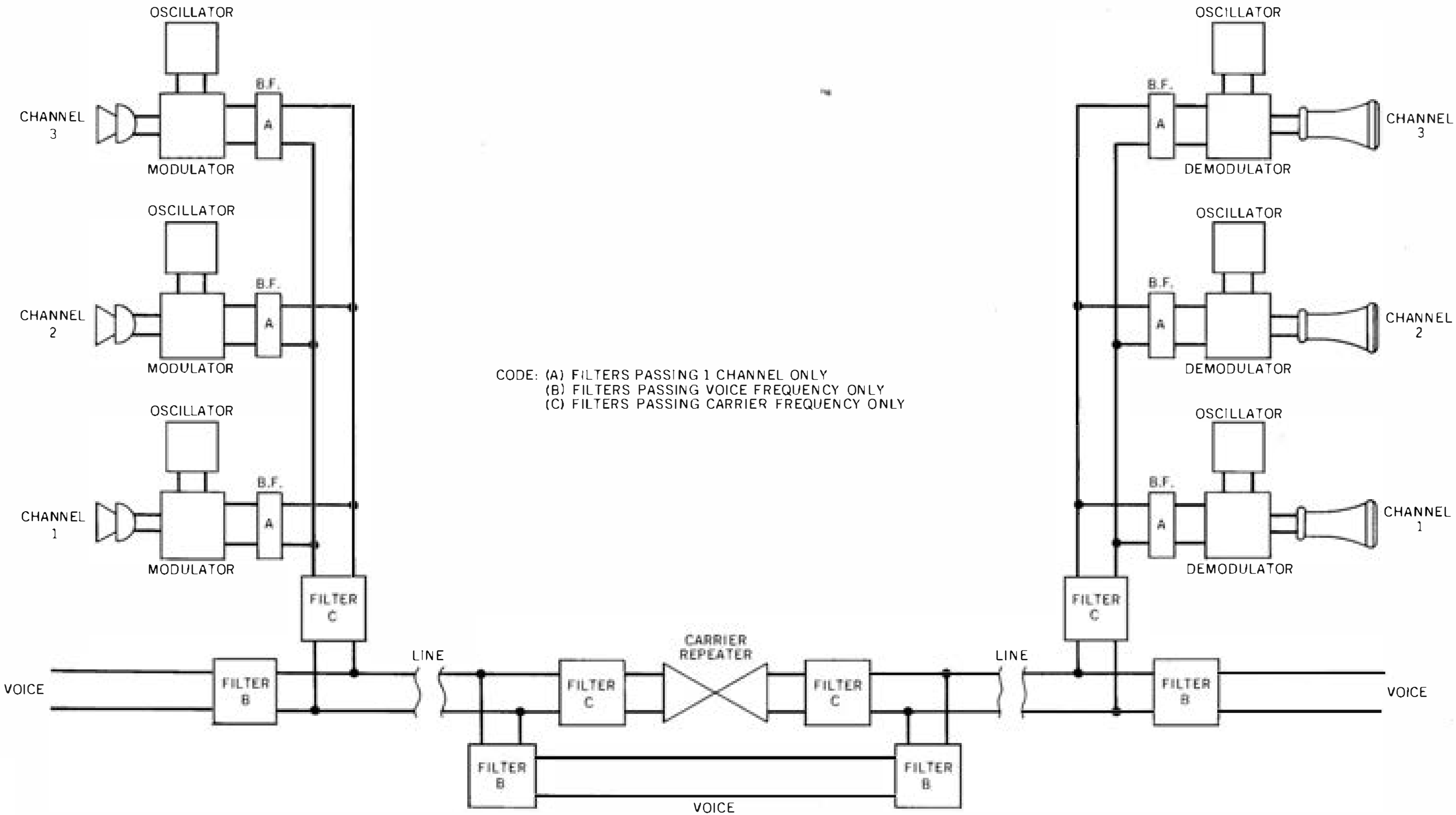


Fig. 4-57. Simplified block diagram of a 3-channel carrier system.

Fig. 4-58. Type A carrier telephone system—experimental equipment used in original carrier transmission tests at Maumee, Ohio, in 1917.

25 kHz were employed, with the carrier suppressed. Seven Type A systems were ultimately installed, the last one remaining in service until the 1940s.

The Type A system, by employing the same frequencies for both directions of transmission (i.e., 2-wire transmission), as was common for open wire voice transmission, provided efficient use of frequency space, and reasonably good service was given, but the number of systems that could be placed on a pole line was limited by near-end crosstalk, that is, undesired coupling between the outgoing carrier circuits and other incoming circuits on the pole.[53] Also, repeater gains were limited by the degree of balance obtainable from the hybrid coil, so that the utility of the system for very long circuits was limited.

[53] Crosstalk problems are discussed in more detail in Section 5.3.1.

In the next system developed, Type B (Fig. 4-60), a slightly different approach was used. It was decided to reduce the number of channels to three above the voice band and to use different frequencies for transmission in the two directions. In this way wave filters could be used to separate the two directions and it would not be necessary to rely upon line balance to prevent singing. This is very much like the 4-wire voice transmission discussed previously and is usually referred to as *equivalent* 4-wire transmission. This system also used single sideband, but in an attempt to avoid the carrier resupply problem the carriers were transmitted. The lower sidebands on carriers at 6, 9, and 12 kHz were used east-to-west and the upper sidebands on carriers at 15, 18, and 21 kHz were used for the opposite direction.

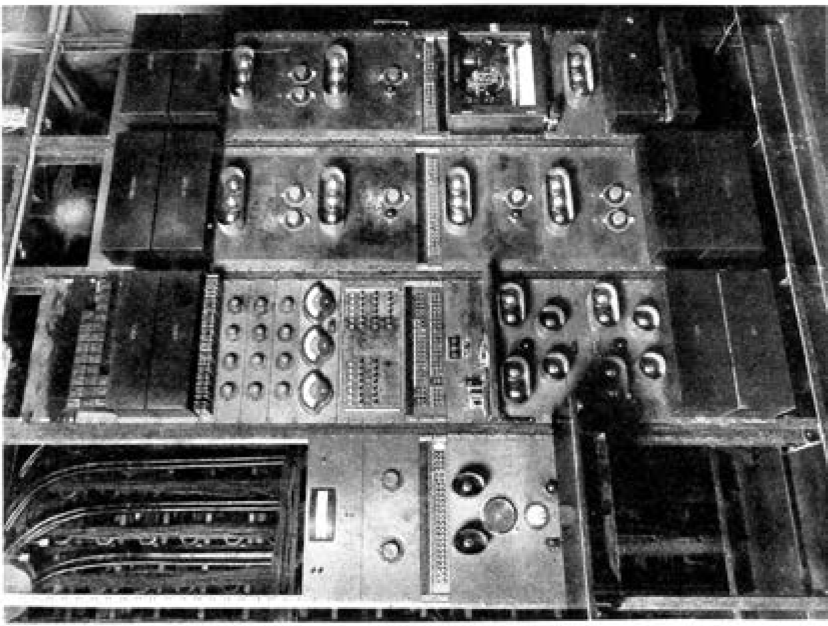
This system demonstrated the great benefits of *equivalent* 4-wire transmission from the standpoint of reduced crosstalk and increased repeater gain, but the transmission of the carrier proved restrictive and in the attempt to conserve frequency space the speech bands had been narrowed undesirably. About a score of these systems were built, beginning in 1920, and some continued in use until about 1940.

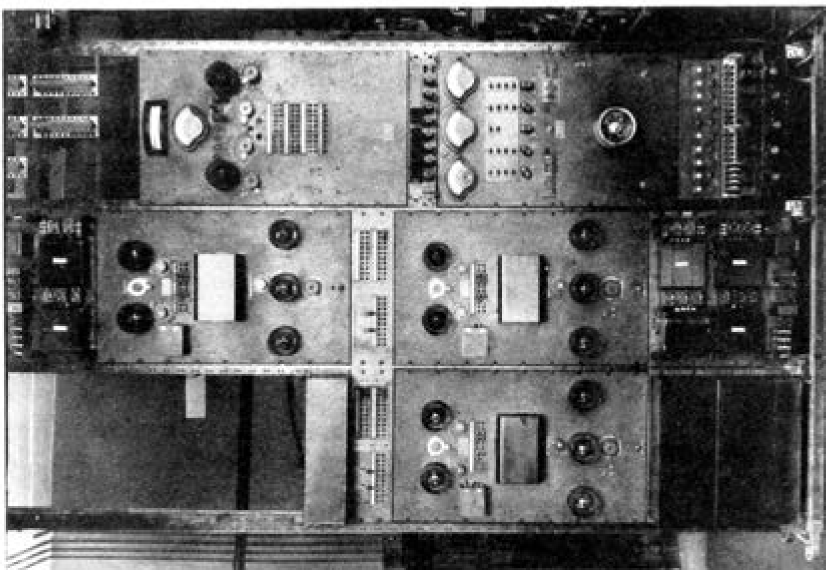
These two early systems tested out the basic principles and set the pattern for the future. Subsequent to the development of Types A and B, almost all[54] carrier systems have been of the single-sideband, carrier-suppressed type and have used 4-wire transmission (*equivalent* 4-wire transmission has commonly been used on open wire and *physical* 4-wire transmission on cable).

Type C, the next to be developed (Fig. 4-61), followed this new pattern but employed better selectivity to separate the two directions of transmission and the top frequency was increased to about 30 kHz to provide wider bands.

Where a number of carrier systems were employed on the same pole line, special transpositions had to be used to reduce carrier crosstalk. In addition, it was found advantageous to shift slightly the frequencies used on carrier systems being transmitted over adjacent pairs to obtain a so-called "staggering advantage," and thus reduce crosstalk impairment. Thus, Type C systems having three slightly different frequency allocations were developed, designated CN, CS, and CT. As improvements were made in the Type C system, successive models were designated C2, C3, and C4, with letters added to indicate the frequency allocations (CN4, CS4, CT4). The Type C system was again redesigned (C5) in the late thirties. Type C is now largely being replaced by more modern systems, but at its peak in the early 1950s

[54] An exception is the Type N system developed around 1950, which returned to the double-sideband principle. It employed both different pairs and different frequencies for the two directions of transmission.





about 1.5 million circuit miles of C carrier were in use in the Bell System. Repeaters (Fig. 4-62) were provided for the Type C system to be spaced at about 125-mile intervals, and by their use the systems could be used for total distances of 1,000–2,000 miles.

Although the Types A, B, and C systems had found ready acceptance for application to long circuits, there were many places in the toll plant where a less expensive type of carrier could be used to advantage on open wire lines for shorter circuits and in areas of slow growth. To fill this need, the Type D system was developed. Type D provided one two-way telephone circuit on a pair of wires in addition to the voice-frequency circuit already in use. Like the Type C system, the Type D employed single-sideband transmission with the carrier suppressed, and used different frequencies for opposite directions of transmission, employing the lower sidebands of carriers at 10.3 and 6.87 kHz. Originally the system was used without amplification, but later a repeater was supplied to extend the range of the system to about 200 miles. A later and similar single-channel carrier system, coded Type H, differed in plan from Type D in using the same carrier frequency, 7.15 kHz, for both directions of transmission, the upper sideband used for one direction and the lower sideband for the other.

In the early 1920s an investigation was undertaken to determine whether telephone communication could be furnished over high-voltage power transmission lines by means of carrier currents. The answer was in the affirmative, and a Type E system was developed for the purpose. Previous development work on multiplex carrier telephone systems on telephone lines was drawn on heavily, but many of the problems were unique, such as the line irregularities, effects of power equipment, coupling to the line, etc. (Fig. 4-63). While these problems were technically solvable,[55] power lines did not provide, on an overall basis, a practical medium for commercial telephony. A few Type E systems were built to provide the communication needs of power companies but the production of this type of carrier was not continued by the Bell System.

One more "first-generation" carrier system was to be designed, just beyond the time frame of this volume, in the early 1930s. This unique and simple single-channel system, designated Type G, was a first attempt to apply carrier to the very short haul field. In order to do this, it was necessary to simplify greatly and reduce the cost of the carrier terminal. This was done by using double sideband with carrier transmitted, operating on a 2-wire basis. This made possible the

[55] Figure 4-63, which pictures the capacitors used to couple a carrier system to a high-voltage line, illustrates some of the unusual problems involved in this method of communication.

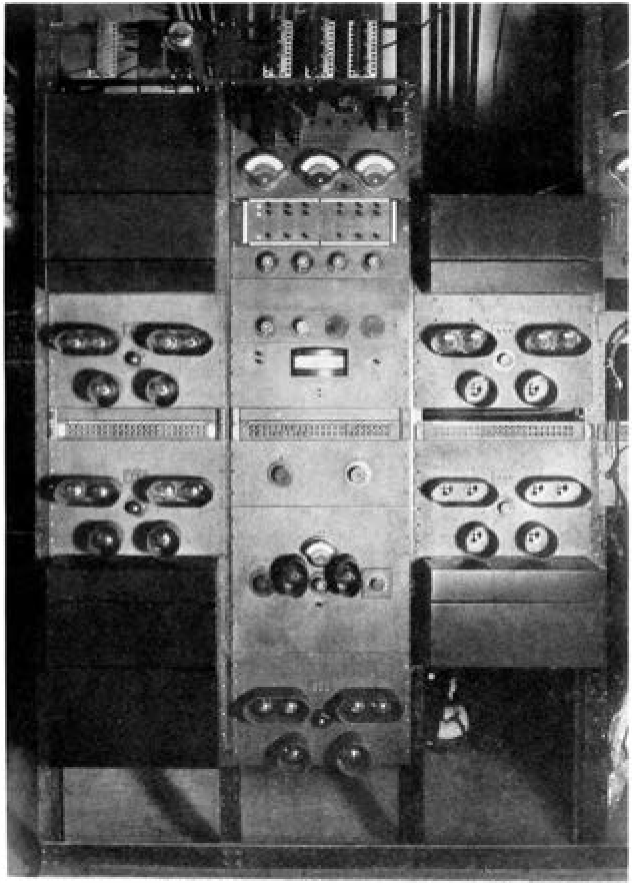


Fig. 4-62. A typical repeater installation for the earlier carrier systems.

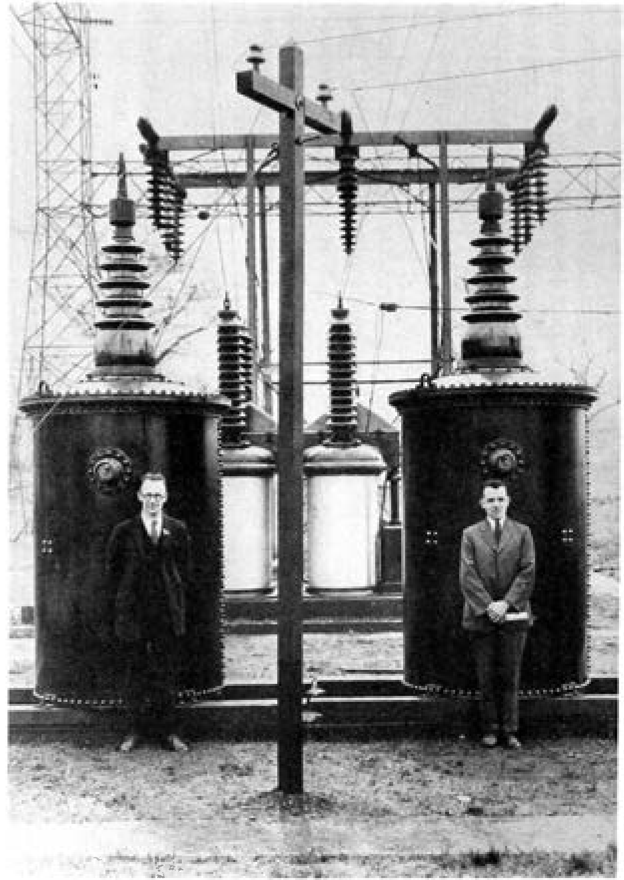


Fig. 4-63. High-voltage (120-kV) coupling capacitors for power-line carrier telephone system.

generation of the carrier at one end only, the other end using an "inert" terminal, i.e., one using no vacuum tubes or other source of amplification and hence requiring no power source. While this greatly reduced cost, the limitations imposed by attenuation and crosstalk were so severe that the field of use was greatly restricted and very few systems were built.

A summary of the characteristics of the early carrier systems just described is given in Fig. 4-64.

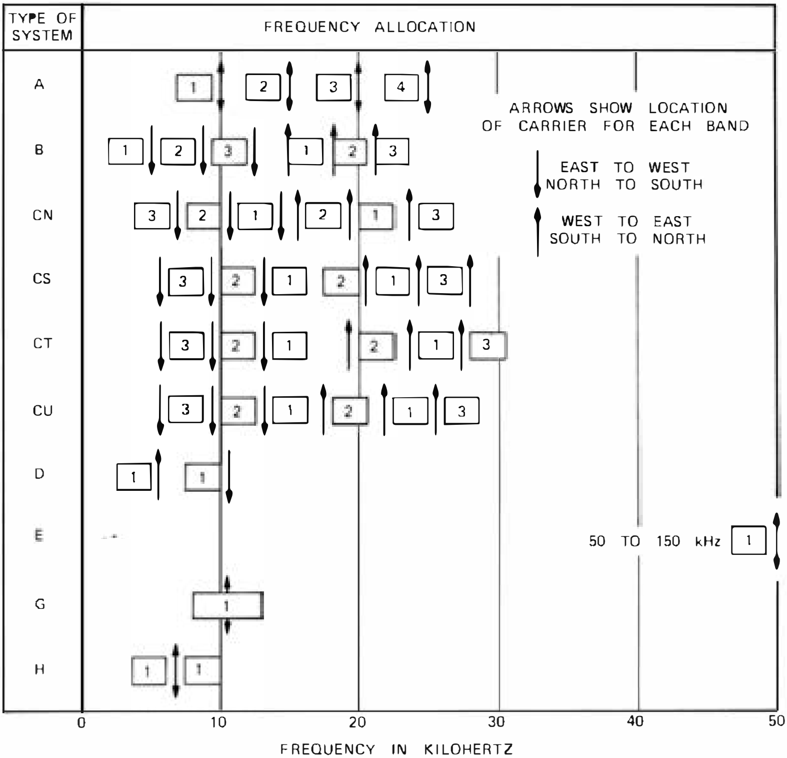
Systems for carrier telegraph transmission were developed at about the same time as the telephone carrier developments just related. These systems, which are described in Section 9.5.5 of Chapter 7, provided for transmitting a large number of telegraph channels in the voice-frequency range and in the frequencies above the voice channel. Initially, carrier frequencies as high as 11 kHz were used, but later it became customary to use only frequencies below those employed in the Type C telephone carrier system.

While a major part of the early effort on carrier went into the development of the carrier equipment itself, a large amount of supplemental work was required to adapt the telephone wire plant for carrier transmission. Crosstalk was a major limitation and the solution of crosstalk problems will be discussed in Section 5.3.1. There also were other kinds of problems. As an example, signals for controlling the switching of circuits were, previous to the development of carrier, transmitted at frequencies below those required for voice transmission, that is, around 20 or 135 hertz. Since these frequencies were not transmitted over carrier channels, being suppressed by the channel filters, it became necessary to develop completely new signaling techniques. A system employing 1 KHz modulated at a 20-hertz rate was used for many years.

4.3 Supplemental Transmission Systems

Although the major effort of Bell System technology has always been directed toward two-way communication between individuals connected through the switched network, some development of diverse but related forms of communication was inevitable. These developments, brought about both by market demand and through the natural evolution of telephone technology, were usually pursued by the Bell System far enough to understand their long-range implications. Many of these projects appeared to lead beyond the mainstream of Bell System communication and work on them was soon discontinued after making pertinent technology available to those in a position to put it to use.[56] However, some of the "spin-offs" from telephone develop-

[56] Examples, to name only a few, were hearing aids, an electric stethoscope, the artificial larynx, etc.



Desig–nation*	Date of First Use	Channels	Treatment of Carrier	Sidebands	Method of Two-Way Operation
A	1918	4	Suppressed	Single	Balanced 2-Wire
B	1920	3	Transmitted	Single	Equivalent 4-Wire †
C	1924	3	Suppressed	Single	Equivalent 4-Wire
D	1926	1	Suppressed	Single	Equivalent 4-Wire
E	1928	1	Suppressed	Single	VF switching ‡
G §	1936	1	Transmitted**	Double	2-Wire
H	1937	1	Suppressed	Single	Equivalent 4-Wire

* The types A, B, C, D, G, and H systems were used on open wire lines; the Type E system was designed for use on power transmission lines.

† Transmission at different frequencies in opposite directions makes each direction of transmission independent of the other to a degree which is equivalent to 4-wire operation.

‡ The same frequency was used for the two directions but a voice-operated switch enabled the circuit in one direction at a time.

§ The F designation was never used for a commercial system.

** Carrier was generated at one end only.

Fig. 4-64. First-generation carrier systems. (Adapted from Bower 1938) ment were closely related to the main Bell System interests and required extensive effort before a decision could be made as to their ultimate disposition.

We do not have space to pursue the many offshoots of telephone technology which took place but one, the public address system, provides a particularly interesting example of the complex and somewhat unpredictable manner in which technology can evolve and it will be covered both here and in Section V of Chapter 5.

Originally, public address systems were aimed at transmitting speech and music by acoustic means to a group of people from one or at most a few loud-speaking telephones instead of using the individual receiver employed in telephony. But they soon grew into a form of broadcasting by wire involving remote audiences. This application received only limited usage but further work led into developments of long-range technical and economic significance both within and outside the Bell System. For example, the public address system and related developments provided basic technology for the high-power, high-fidelity audio systems needed for talking moving pictures and for high-fidelity music reproduction. While neither of these developments was of great value to the Bell System, their economic impact on outside industry has been very large indeed.

Other offshoots of public address system development held more promise of lasting benefit to Bell communications. Quite fortunately the high-quality audio equipment of public address systems became available just as Bell work on radio transmission was maturing and the two techniques were combined to open up the new field of radio broadcasting, as discussed in Chapter 5. But here again the greatest impact of this development was outside the Bell System since it was ultimately decided that the broadcast type of transmission fell outside the System's main field of interest.

But this was not yet the end of the influence of the public address project. The application of public address work had led to the design of special transmission systems, or "program circuits," for carrying high-quality speech and music over Bell System lines. The availability of these techniques was an important factor in the development of network radio broadcasting which in turn stimulated further development of program circuits. This was a natural field for the use of the Bell network, and audio program circuits and their video counterparts became an important segment of Bell enterprise. At last, the wisdom of pursuing the development of a public address system became apparent.

4.3.1 Public Address Systems

Public address work started in 1907 when the development of a loudspeaker was begun in order to meet the requests of a few customers who wished to use this device instead of the telephone receiver. A few loudspeakers were produced and used on a trial basis in the following year. In 1910, a double-button microphone was developed to provide a source of electric waves some 13 to 15 dB more efficient than the regular telephone transmitter. In 1912, a high-efficiency transmitter and a loudspeaker were standardized and used for communicating between the wire chief's desk and craftsmen at work in the frame room. Originally this was a one-way system from wire chief to craftsmen, but later a return path was added which employed a distant-talking transmitter in the frame room.

The first long-distance public address broadcast occurred in 1913 when the Governor of Oklahoma, speaking at Oklahoma City, addressed an audience of over 300 persons in Tulsa, 122 miles away, the speech being transmitted over No. 12 gauge open wire pairs. No amplification, except from the microphone, was available at this time and the necessary energy was obtained from a water-cooled microphone supplied with a 3-ampere direct current (about 50 times the current used in an ordinary telephone transmitter). This device provided enough energy, not only to overcome the wire line attenuation, but also to drive the loudspeakers. The demonstration apparently was highly successful but the use of high-powered transmitters as a source of amplification had limited application because of the potential for crosstalking into the much-lower-level telephone circuits.

What was needed, as in the telephone field, was a source of amplification that could be introduced along the line and also be used for producing high-power electrical waves for driving loudspeakers. The vacuum tube made this possible and this type of amplifier was first used in a public address demonstration in 1916. After this, it became possible to concentrate effort on developing microphones and loudspeakers with high-fidelity reproduction instead of emphasizing high efficiency alone.

In 1918, a new type of loudspeaker using the balanced-armature principle was developed and used on submarine chasers. Later this principle was applied in a driving unit, shown in Fig. 4-65, which was used as a sound source with various types of horns (Fig. 4-66). As narrated in Section 8.4.3 of Chapter 3, in 1924 the same general scheme was incorporated in the 540-AW cone-type loudspeaker that was widely used with early broadcast receivers.

By 1919 the art had developed to the point where an installation could be made, for a Victory Loan Drive, which covered three city blocks along Park Avenue in New York City. The installation involved over 100 loudspeakers driven by a number of high-power amplifiers. The first widespread testing of public address system characteristics

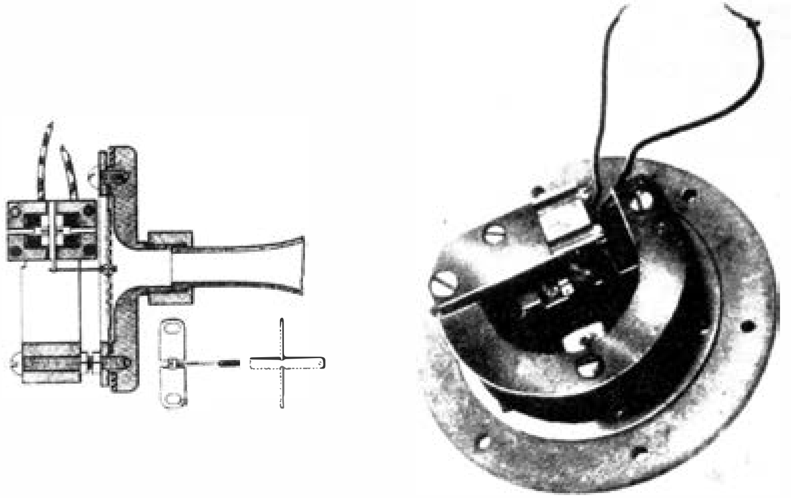
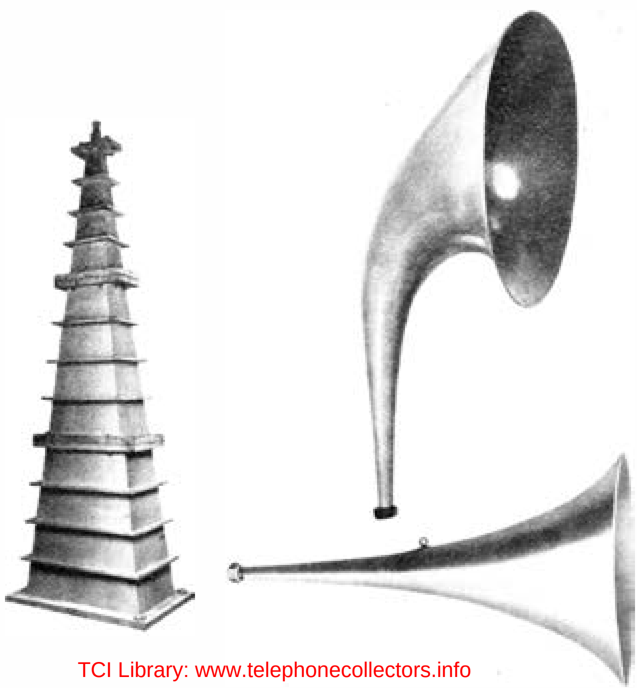


Fig. 4-65. Receiver designed to carry several watts with small distortion. (Green and Maxfield 1923, Fig. 11)

Fig. 4-66. Three types of horns used. (Green and Maxfield 1923, Fig. 12)



was made in connection with this demonstration. Measurements included tests of intelligibility, sound distribution, frequency response, and noise. In 1921, a high-quality, push-pull, double-button microphone and a volume indicator (see Section 5.1.4.3) were used for the first time. Both received much use in public address and early broadcasting systems. In fact, the volume indicator, in modified form, is still being used widely for measurement and control of speech volume. However, within a few years the carbon microphone was superseded by the condenser microphone which was originally developed as a byproduct of Bell System acoustical research.

On Armistice Day, 1921, one of the most elaborate wire broadcast and public address installations ever made was set up to cover the burial of the Unknown Soldier at Arlington Cemetery. The ceremonies were heard by 30,000 people in New York City and 20,000 in San Francisco in addition to 100,000 at Arlington. In all, the demonstration involved 3,900 miles of telephone lines, three transmitter locations, and some 80 loudspeakers. It not only did much to demonstrate the high level of technology available in the audio field but also showed the tremendous interest in public events that would soon provide an audience for radio broadcasting. The first standardized public address system was produced in 1922 and a number more were developed in the next ten years. With the standardization of a complete system, application became a fairly routine matter and Bell System participation in installation fell off; however, system manufacture by Western Electric continued until the middle 1930s.

An exception to this general trend away from the public address field occurred in 1932–33 when Bell Telephone Laboratories and the American Telephone and Telegraph Company, in cooperation with the Philadelphia Orchestra, conducted experiments in very high power, wideband stereophonic transmission of music. The ultimate result of this work was a three-channel system with an 80-dB volume range capable of producing sound power about ten times that of an unamplified orchestra and covering a 15-kHz frequency range. Loudspeakers set up in Constitution Hall in Washington reproduced the playing of the Philadelphia Orchestra in the Philadelphia Academy of Music. For this demonstration, Leopold Stokowski, the conductor, turned over the baton to an assistant and manned the system control.

Again, in the Second World War, Bell Laboratories reentered the public address field and developed a number of systems for use on naval ships which met requirements for high reliability under the shock of gunfire and high intelligibility under extreme and rather unusual types of noise.

4.3.2 Program Transmission Systems

The technical requirements underlying public address transmission were considerably different from and mostly much more rigorous than those applying to telephone transmission. The early realization of these requirements and developments made to meet them did much to supply the acoustical needs of radio broadcasting when it began its remarkable growth in the early 1920s (as related in Section V of Chapter 5).

In a sense, program transmission was simpler than telephony since it was a one-way service and the complications of bilateral transmission were avoided, but this was more than offset by the need for transmitting a broader frequency band and providing a wider range of sound level and greater freedom from distortion and noise. Telephone speech was quite intelligible when limited to a band of 300–3,200 hertz and was even acceptable when the frequencies around 1,000 hertz were greatly emphasized to achieve high efficiency in acoustic-electric conversion. On the other hand, speech so distorted and limited in frequency sounded unnatural when heard over a loudspeaker and music was completely unrealistic. In addition, music covered an extremely wide range of power levels, some 50 dB or more. Systems for its transmission had to be able to reproduce the lowest of these sounds well above the noise on the system and also handle the highest without distorting the signal by "overloading" the equipment. By the early 1920s reasonably good short-haul audio transmission had been achieved. Low-noise circuitry had been devised for the low-level portion of the equipment (often this required use of batteries for plate and grid supply of vacuum tube amplifiers) and the frequency band had been extended to a range of 50–5,000 hertz. The full range of sound level was not transmitted but an acceptable range was provided, without overloading, by some compression of the volume range by means of manual control.

Much of the early effort devoted to meeting these requirements went into microphone and loudspeaker developments. A major problem was concerned with extending the lower frequency limit since musical sounds, particularly those from percussion instruments, were unrealistic without full reproduction of the low frequencies. As the acoustic converters were improved, parallel work was devoted to improving the wire transmission lines. Here major changes were required. Loading, which restricted transmission to a top frequency of about 3 kHz, had to be removed or replaced by a system which extended the cutoff well above 5 kHz.[57] In order to utilize the low frequencies, it

[57] In later years carrier systems presented serious problems and it was necessary to develop special terminals so that the wide frequency band required for program transmission could be obtained by utilizing the space normally occupied by several telephone channels.

was necessary to eliminate the composite equipment used for providing telegraph and signaling channels. To compensate for line loss, it was necessary to develop special amplifiers meeting the necessary requirements for bandwidth, noise, and volume range. And finally, equalizers were required to compensate for the variations in attenuation with frequency which were introduced by the line facilities, particularly the non-loaded cables.

The development and construction of line facilities meeting program requirements took a number of years. For the 1921 Arlington demonstration the local circuits, using short non-loaded cable, were equalized reasonably well over the 50–5,000-hertz range as illustrated by Fig. 4-67. However, the long-distance transmission used telephone-grade circuits of the best type then available, consisting of repeaters and non-loaded No. 8 BWG open wire lines carefully aligned and maintained for optimum performance. The transmission characteristics of the New York to San Francisco circuit, which included 12 repeaters, is shown in Fig. 4-68. While this represented excellent performance for the time, it was far from the desired objective for music. However, with good-quality microphones and loudspeakers,

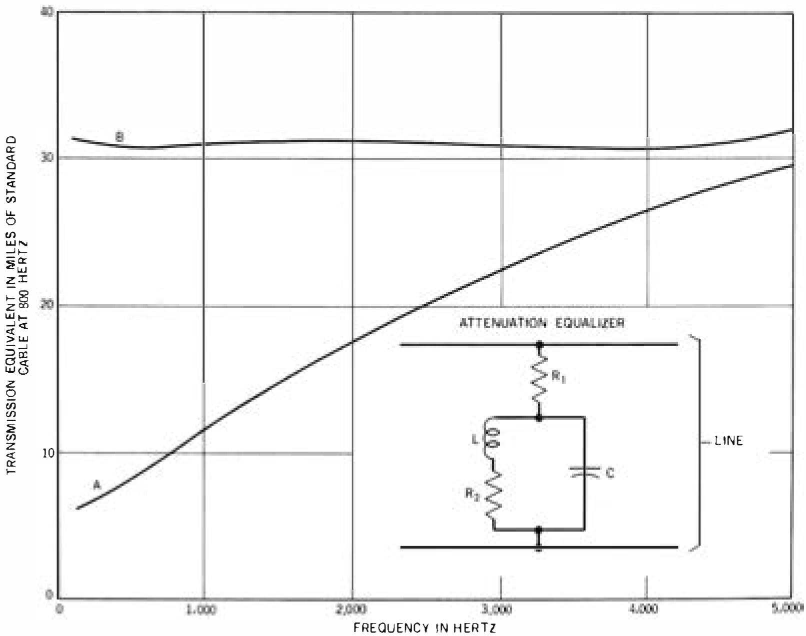


Fig. 4-67. Transmission characteristic of a 10-mile, 19-gauge, non-loaded cable circuit. Curve A—without attenuation equalizer. Curve B—with attenuation equalizer. (Redrawn from Martin and Clark 1923, Fig. 2)

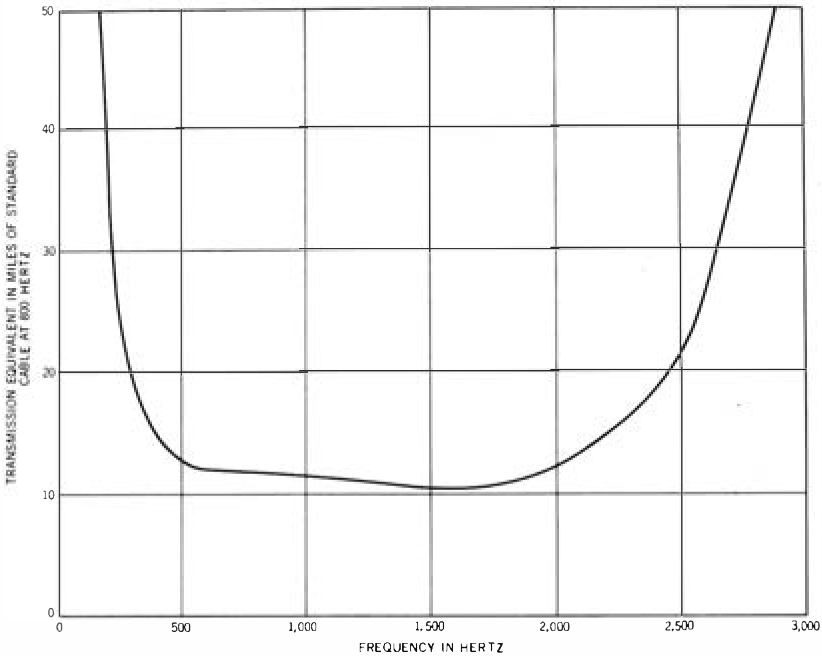


Fig. 4-68. Transmission characteristic of transcontinental circuit, New York to San Francisco. (Redrawn from Martin and Clark 1923, Fig. 1)

these circuits were reasonably satisfactory for the outdoor broadcast of speech, the main objective of the 1921 demonstration. The frequency band could have been extended at the lower end by removing composite circuits and at the upper end by removing carrier telegraph. However, at this time transcontinental circuits were in short supply and it was decided to use standard telephone circuits, temporarily diverted for the demonstration, since completely modifying them would have had broader reaction on telephone service.

A similar approach was used in some of the earlier radio network broadcasting, namely using well-equalized non-loaded cable for the short and more or less permanent circuits between studio and radio transmitters with the best available grade of telephone circuits, carefully aligned to meet program requirements, for the temporary long-distance lengths. Here again this was reasonably satisfactory for things like news and sports broadcasting when these lines were used with high-performance microphones and good-quality broadcasting stations. It should be recalled that the broadcast listener at this time usually used headphone receivers which had a highly peaked frequency response transmitting little energy above 2,500–3,000 hertz.

As network broadcasting spread, and began to be used for high-grade music programs, the need for better-quality transmission arose and was met by removing composite and carrier telegraph circuits and developing wideband repeaters capable of handling a wide range of volume levels without noticeable non-linear distortion. Originally these circuits were set up on a temporary basis but as network broadcasting stabilized, circuits were set apart especially for the purpose. Even in the early days, when telephone circuits were used for programs with little modification, a very considerable amount of time and effort was required to prepare them for program use. First, of course, they had to be disconnected from switchboards and patched through to provide the desired route. Next, they had to be carefully aligned for optimum performance and measured carefully to provide assurance that lineup was being maintained. Finally, after use, they had to be returned to message service and again patched. Usually three circuits were assigned for handling program transmission, one acting as an order wire so that technicians along the line could be constantly in communication, and two assigned for program use, one being designated for normal use and the other held in reserve as a spare. As program circuits were improved, by removing telegraph, adding special amplifiers, and so forth, the work required to prepare them for broadcasting and subsequent restoration to message use increased greatly and as much as forty man-days might be required for a 1,000-mile circuit. To avoid some of this work, beginning about 1923, certain circuits were designated for broadcast use and special arrangements were made to facilitate their transfer between message and program service. These were marked with a red pencil and referred to as the "red layout." By 1926 the demand for broadcasting had become quite large and long-distance-circuit availability had grown to the point where these circuits could be assigned exclusively for broadcasting, and they became known as the "red network."

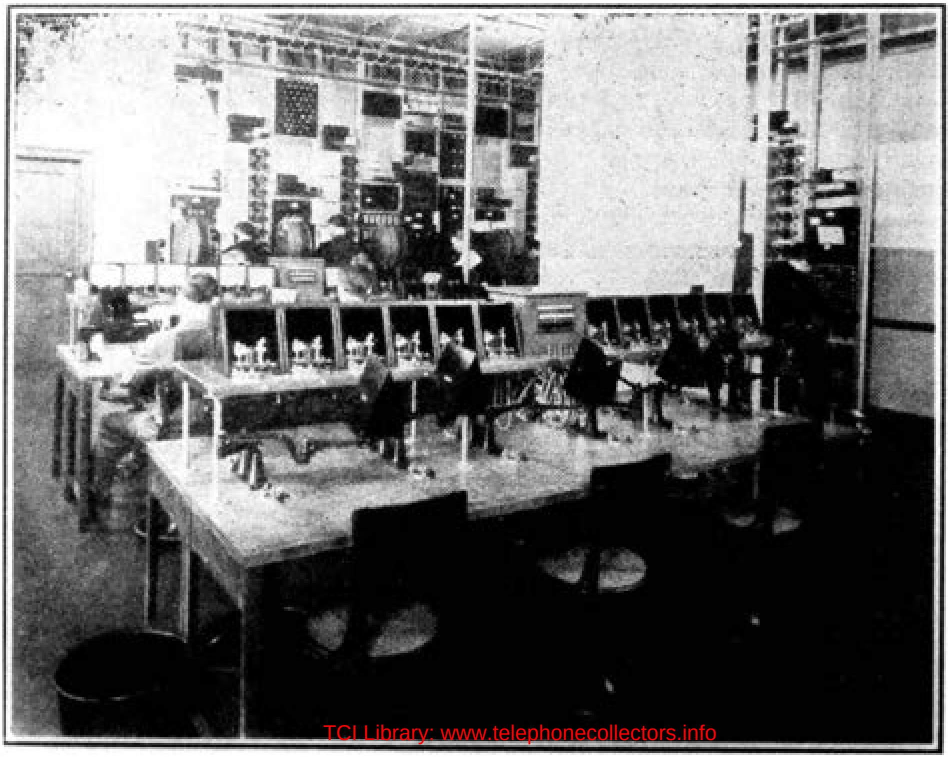
Growth was rapid and in 1930 about 35,000 miles of circuits were being regularly used for supplying about 150 radio stations. These circuits were organized into four regional areas, each with a control center. A typical center, located in Chicago, is shown in Fig. 4-69. These centers were connected to each other and to radio stations and line repeater points by an elaborate order wire system using about 43,000 miles of telegraph circuits. At this time, most of the program circuits used No. 8 BWG (165-mil), non-loaded, copper open wire with amplifiers about 150 miles apart. The amplifiers included adjustable equalizers to compensate for line attenuation. Incidental cables used No. 10 or No. 11 B and S pairs with special loading to match the open wire impedance and provide a cutoff well above 5 kHz. The transmission band of the entire system extended from 100 to 5,000 hertz but plans were being made for widening it significantly.

By 1930 the Bell System telephone cable network extended from the East to well beyond Chicago and a countrywide system was being planned. It was obvious that provision should be made to use these cables for program circuits. In the late 1920s development of such a system was begun and in 1930 a trial installation was made on pairs looped back and forth in the New York-Pittsburgh cable for a total length of 2,200 miles. The system provided for the transmission of a frequency band extending from 50 to 8,000 hertz and for handling a volume range of about 40 dB without material noise interference.

The cable system used No. 16 B and S gauge conductors loaded with 22-mH inductance coils spaced 3,000 feet apart, giving a nominal cutoff frequency of 11 kHz. A special cable was designed to distribute the 16-gauge pairs among the 19-gauge message pairs in a manner that would minimize crosstalk into and from the program circuits. Special amplifiers were designed to handle the wide range of volume and included gain regulation and equalization to provide not only constant amplitude but also reasonably constant delay over the transmitted band.[58] These amplifiers (repeaters) are pictured in Figs. 4-70 and 4-71.

[58] When waves are transmitted for long distances over loaded circuits (or circuits containing wave filters), the high frequencies tend to be delayed in arrival as compared to the midband frequencies. This can be quite disturbing on some program material.

Fig. 4-69. Program transmission equipment in Chicago office. (Cowan 1929, Fig. 6)



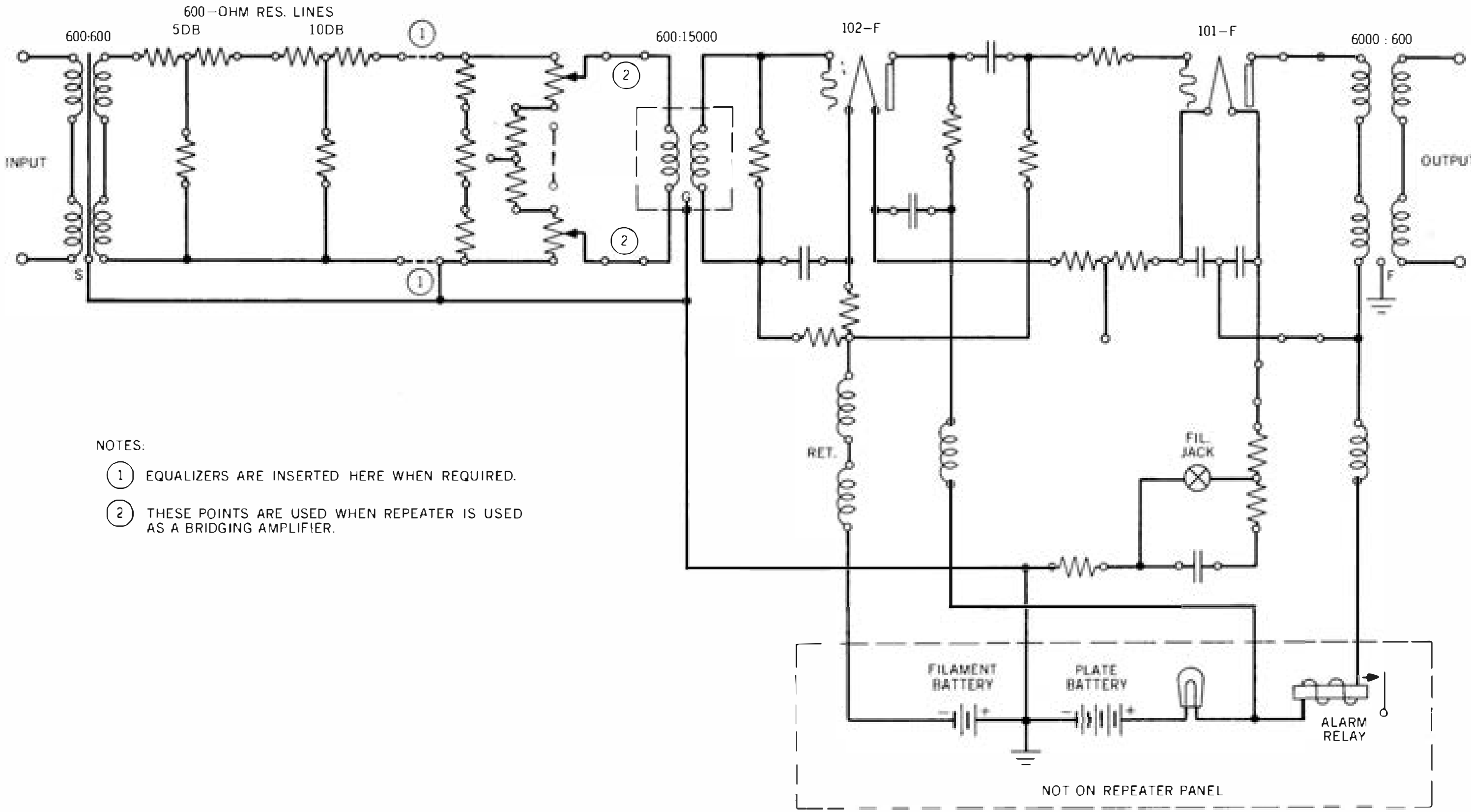


Fig. 4-70. Schematic of non-regulating repeater for cable program system. (Redrawn from Clark and Green 1930, Fig. 10)

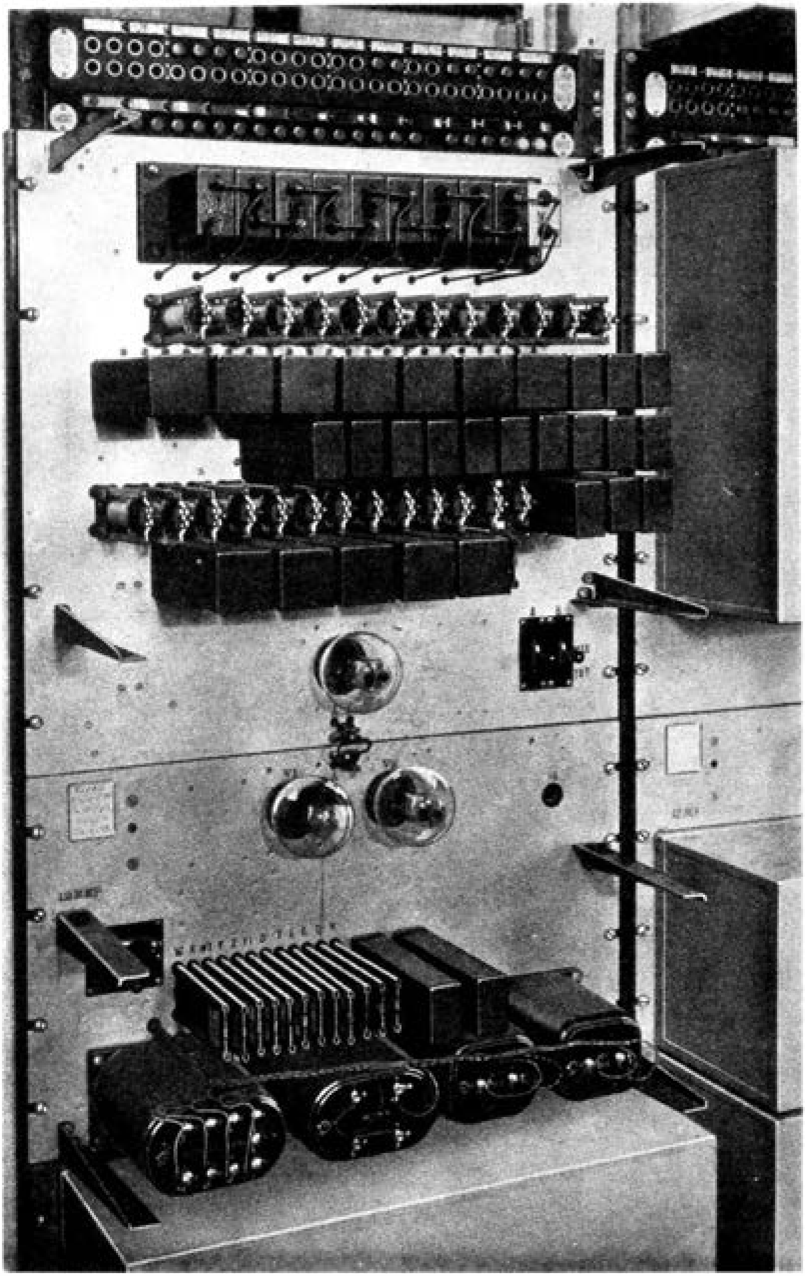


Fig. 4-71. Front view of program repeater and associated regulating system. (Clark and Green 1930, Fig. 11)

Performance characteristics of the complete system operating over 2,200 miles of cable are shown in Figs. 4-72 and 4-73.

V. ANCILLARY TRANSMISSION PROBLEMS

So far, we have concentrated our attention on the evolution of transmission media (and the underlying theory), on the devices used to reduce or compensate for attenuation, and on systems for multiplexing a number of circuits on a single physical channel. Needless to say, these steps in solving the transmission problem would have been impossible without supporting effort in related fields. Chief of these were the provision of suitable measuring instruments, the reduction of interference from external sources, and the development of techniques for arriving at the most economical design of plant meeting transmission requirements.

5.1 Measuring Instruments

Today, when practically any type of measuring device can be ordered from a catalog, it is difficult to appreciate the situation prevailing in the early days of telephony. At that time the applicable measuring equipment consisted largely of the d'Arsonval galvanometer (the predecessor of the dc voltmeter and ammeter) and the Wheatstone bridge. Alternating current measurements, particularly at low amplitudes and telephone frequencies, were practically impossible until the telephone receiver provided an audible means for detection. This device, plus buzzers or other primitive sources of tones, extended the use of the Wheatstone bridge and provided means for measuring capacitance, inductance, and impedance in the voice range.

Direct-current meters proved very useful in checking the early telephone lines for leakage, opens, grounds, etc. These instruments and variants of the Wheatstone bridge (Murray or Varley loop circuits) were important devices for assuring correct performance and still are used for maintenance of physical circuits such as those connecting the telephone user to the telephone switching office.

However, means for measuring and comparing overall telephone performance were badly needed. For many years subjective tests provided the only way to do this.

5.1.1 Subjective Tests—Loudness Balancing

Obviously, the simplest way to compare two circuits was to talk over them alternately and make a subjective judgment as to which was better. This was the basis for all early comparisons of copper and iron wire, open wire and cable, etc. However, this technique was ill-adapted to rating a large number of conditions or to relating tests

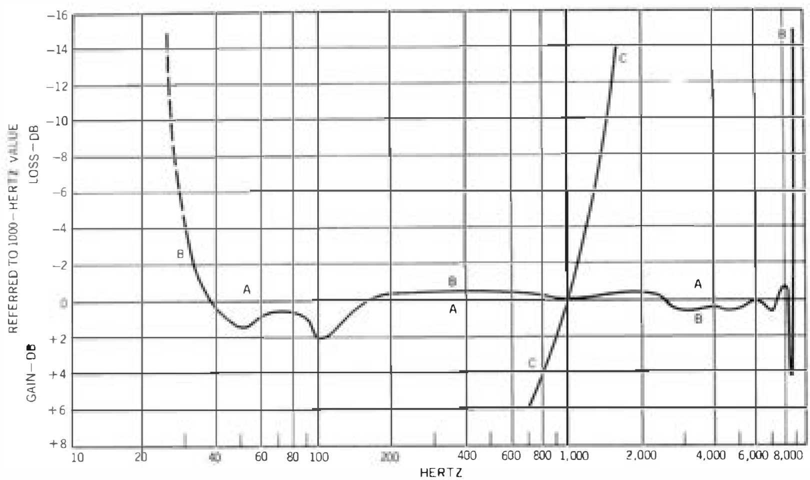


Fig. 4-72. Transmission frequency characteristics of 2,200 miles of 16-gauge B-22 cable program transmission circuit. Curve A—ideal characteristic. Curve B—measured characteristic. Curve C—line without equalizers. (Redrawn from Clark and Green 1930, Fig. 21)

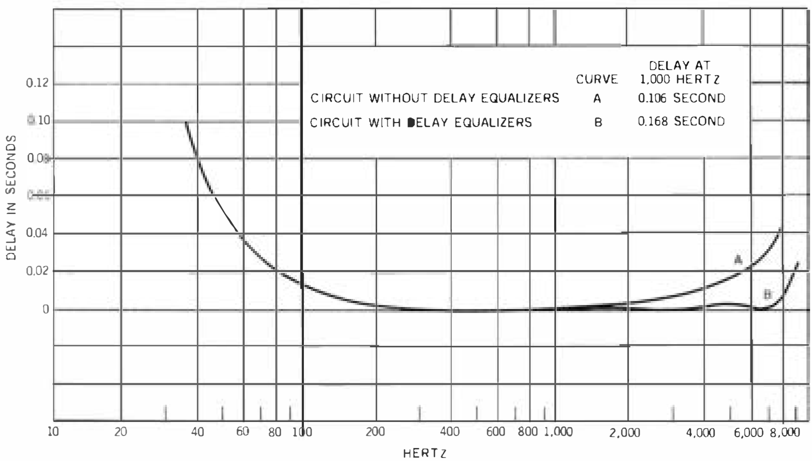


Fig. 4-73. Delay characteristics of 2,200 miles of 16-gauge B-22 cable program transmission circuit. (Redrawn from Clark and Green 1930, Fig. 22)

made at different times. What was needed was a more quantitative technique. Two requirements were basic. First, there must be a point of reference or zero, and, second, there must be a scale or unit of measurement.[59]

Such a technique was developed around the turn of the century. The

[59] The measurement of temperature provides a useful analogy. On the Centigrade scale, for example, the temperature of melting ice is taken as the reference, or zero, point. The scale is the degree and measures temperature relative to this zero.

best telephone connection possible at the time was selected for the reference point. This consisted of a telephone station connected with minimum-length line through a simulated central office to a similar far-end equipment. This connection had a rating of zero. An artificial line, which simulated a typical cable used in the plant, was connected between the two central-office circuits and provided the scale. The whole, shown in Fig. 4-74, constituted a reference system against which telephone connections could be compared. The unknown and the reference systems were talked over alternately and artificial cable was added to the reference system until the received speech was judged equally loud. The rating of the unknown circuit was taken to be the number of miles of cable in the reference system. The artificial cable was standardized at a resistance of 88 ohms per loop mile and a capacitance of 0.054 mF per mile, values representative of cable then in general use, and the circuit rating was accordingly referred to as the "loss in miles of standard cable." These ratings were also referred to as "transmission equivalents." A rather elaborate system of losses was ultimately worked out to provide ratings of telephone sets, customers' loops, central-office equipment, trunks, etc., so that design objectives could be established for each component of the overall circuit and the overall loss, or equivalent, determined by adding component losses.

Several problems soon developed. When improved instruments were designed, the reference set was no longer the best telephone connection available and balances often required the insertion of cable in the unknown circuit in order to equal the zero connection. Thus such circuits had negative transmission equivalents, or gain, relative to the reference system. While this was somewhat awkward, the situation arose only after the reference system had been long in use, and it was considered to be less confusing than abruptly changing the reference point.[60] Another problem arose from the distortion introduced by cable. As artificial line was added, it attenuated higher frequencies more than the lower and speech was highly distorted. This made the rating of relatively distortionless circuits, such as open wire and loaded cable, very difficult. To correct this difficulty, it became common around 1920 to use a distortionless line having a constant loss and impedance (600 ohms non-reactive) throughout the voice range of frequencies, in place of the artificial cable. In order to alter the significance of measurements as little as possible, the unit of attenuation used for this line was the attenuation of a mile of standard cable at 800 cycles per second (hertz) and was accordingly referred to as "an 800-cycle mile."[61]

[60] This problem, it will be noted, has long been faced by users of thermometers, and the use of "negative" temperatures is commonplace.

[61] Actually, the frequency seems to have been arrived at by selecting the product of $2\pi f$ to be equal to 5,000 (presumably because this was a nice round number). On this basis, f worked out to be 794 hertz. This, not being a nice round number, was soon rounded to 800 and the origin largely forgotten. The effect on speech loudness of an 800-cycle mile was roughly the same as the effect of an actual mile of cable.

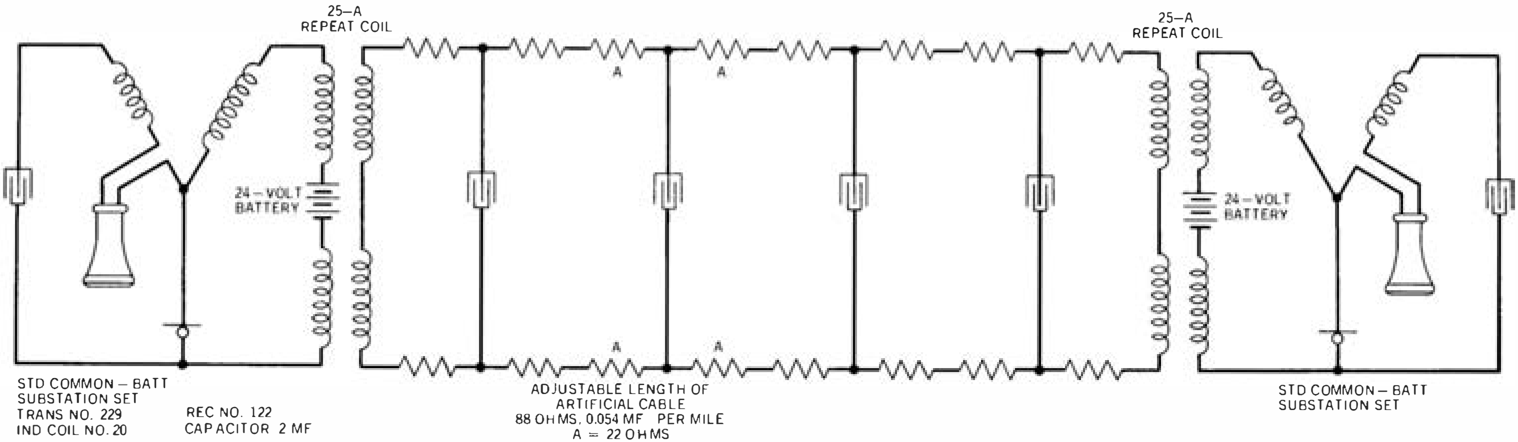


Fig. 4-74. Bell System Reference Circuit.

The distortionless attenuator proved highly successful, but the 800-cycle mile was an awkward unit to use and describe. Its great virtues were its similarity in magnitude to the familiar mile of cable and the fact that, like miles of cable, it was a logarithmic unit. Attenuations, which represent power ratios, are therefore easy to handle numerically when expressed in this form. The overall effect of a series of attenuations can be obtained by adding the logarithmic units instead of multiplying the power ratios.

It was soon noted that the 800-cycle mile was mathematically equal to 10.56 $\log_{10} P_1/P_2$, where P_1 and P_2 are the input and output powers, respectively. This coefficient was so close to 10 that the next step was obvious. In 1924 the Bell System adopted a new unit of attenuation, called the "Transmission Unit," or TU, which was defined as 10 $\log_{10} P_1/P_2$. This was close enough to the 800-cycle mile to be used almost interchangeably therewith, and at the same time had a definition that was more satisfying to those who like simple relations using easily remembered round numbers.

It was proposed to the International Advisory Committee on Telephony that this unit be standardized on a worldwide basis. The proposal met with the approval of many telephone administrations since the mile of cable had received fairly widespread use.[62] However, a considerable number of countries had been using a different unit which was nearly ten times the size of the TU. This unit, called the napier after the inventor of logarithms, was defined in terms of natural logarithms, i.e., using the base e (2.718) instead of the base 10. It also was defined in terms of current ratios instead of power ratios. The upshot was a compromise, in 1928, which standardized several units. One of these was that based on natural logarithms and called the neper. The other (of approximately the same magnitude) was based on common logarithms (base 10) and called the bel. One tenth of a bel, or a decibel, was exactly equal to the TU, and the latter term was accordingly replaced in the Bell System by the decibel, or dB.

Over the years the term bel has fallen into disuse and the neper, while still in use, is far from common. Thus the decibel or dB is fast becoming the accepted unit of attenuation worldwide. The following table provides a comparison of the various units:

Various Ways to Define Attenuation

$$\begin{aligned} \text{bel} &= \log_{10} P_1/P_2 \\ \text{TU} = \text{dB} &= 10 \log_{10} P_1/P_2 \\ \text{800-cycle mile} &= 10.56 \log_{10} P_1/P_2 \\ \text{neper} &= 1.151 \log_{10} P_1/P_2 \\ &= 0.5 \log_e P_1/P_2. \end{aligned}$$

[62] Such a system was adopted in the United Kingdom as early as 1904.

Examples of these relations are shown below:

Power Difference in Logarithmic Units

Power Ratio P_1/P_2	neper	bel	TU or dB	800-cycle mile
1	0	0	0	0
2	0.35	0.30	3.01	3.18
4	0.69	0.60	6.02	6.36
10	1.15	1.00	10.00	10.56
100	2.30	2.00	20.00	21.12
1,000	3.45	3.00	30.00	31.68

While the mile of standard cable was originally introduced as a means for rating the loudness of telephone circuits, relative to reference, the dB into which it evolved has become a far more versatile unit. Since it corresponds to a ratio of two powers, it is used today in a host of electrical and acoustical measurements such as noise, crosstalk, frequency response characteristics, echoes, etc. While the dB is strictly a measure of power ratio, it is commonly used as an absolute measure of power level where the power reference is either expressed or implied.

At about the same time that the unit of attenuation was being changed, the transmission reference system was also undergoing review. The use of a system defined in terms of instrumentalities, i.e., specific types of transmitters and receivers, was obviously undesirable since such devices could not be manufactured and maintained within close limits. In addition, the instruments introduced a large amount of distortion. This distortion caused no problems as long as it was similar to that of commercial instruments, but in the early twenties it was evident that it would soon be possible to develop instruments with much less distortion so that the continued use of a highly distorted reference system might become undesirable. Accordingly, there was developed a Master Transmission Reference System which used instruments that were practically distortionless within the telephone band, stable in performance, and defined in purely physical terms (i.e., in terms of acoustic and electric power). Means for measuring the instruments and adjusting them to specified standards were a part of the system. By 1925 a preliminary model of such a system had been built. In 1926 the International Consulting Committee on Telephony (CCITT) recommended that such a system, with minor modifications, be adopted internationally. By the summer of 1929 two identical systems had been built by the Bell System, one installed at Bell Telephone Laboratories and the other at the CCITT laboratory in Paris.

The place of this system in the evolution of telephony and the story of its successors goes well beyond the scope of this portion of our history. It is sufficient to say here that it represented an important step in the evolution of transmission rating systems and in the development of techniques for measuring telephone systems in physical terms.

Paralleling the development of these subjective techniques for measuring transmission equivalents, similar methods were developed for measuring crosstalk and noise. In each case an ear balance was made between an unknown and a reference speech (or noise) which could be attenuated until the two were judged equal. Later these ear balance techniques were superseded by meter measurements, but for the most part the latter were not widely adopted until after 1925.

5.1.2 Subjective Tests—Articulation Testing

The ear balance tests we have been discussing were intended to give a measure of the loudness of speech transmitted over a telephone circuit. It had long been appreciated that this was only one of the factors affecting the quality of speech; it was also affected by distortion, noise on the line, and external factors such as room noise and user characteristics. In 1910 the ingenious and versatile George Campbell described, in the *Philosophical Magazine*, transmission tests he had made in an attempt to better evaluate the many factors influencing transmission. Campbell's technique involved a talker at the sending end of the circuit to be rated and a group of listeners at the other end. The talker transmitted over the circuit a list of meaningless monosyllables made up by combining the sounds of speech in a systematic way. The listeners recorded the syllables they heard, and the percent received correctly (called percent articulation) was taken as a measure of the transmission capability of the circuit. This "articulation" testing technique was later developed extensively by Harvey Fletcher and others and has led to much insight into the requirements for good speech transmission.

Later, in the mid-twenties, W. H. Martin proposed a count of requests for repetitions on working telephone circuits as an overall measure of transmission. About 1930 a technique was developed for specifying "effective transmission losses" based on this suggestion.

5.1.3 Objective Tests—Attenuation Measurement

Subjective tests, such as ear balances between sounds or articulation tests, were at best difficult to make even under laboratory conditions and were entirely unsatisfactory for maintaining the specified performances of actual telephone lines. Fortunately, until the advent of the telephone repeater and carrier systems, the adequate performance of

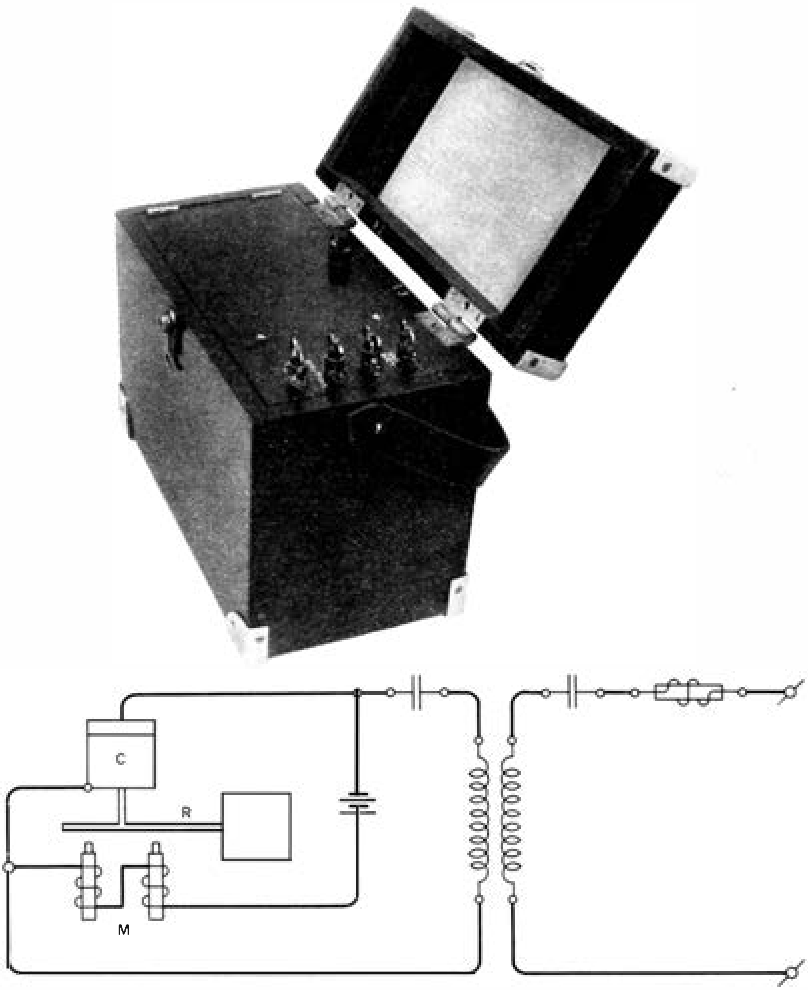
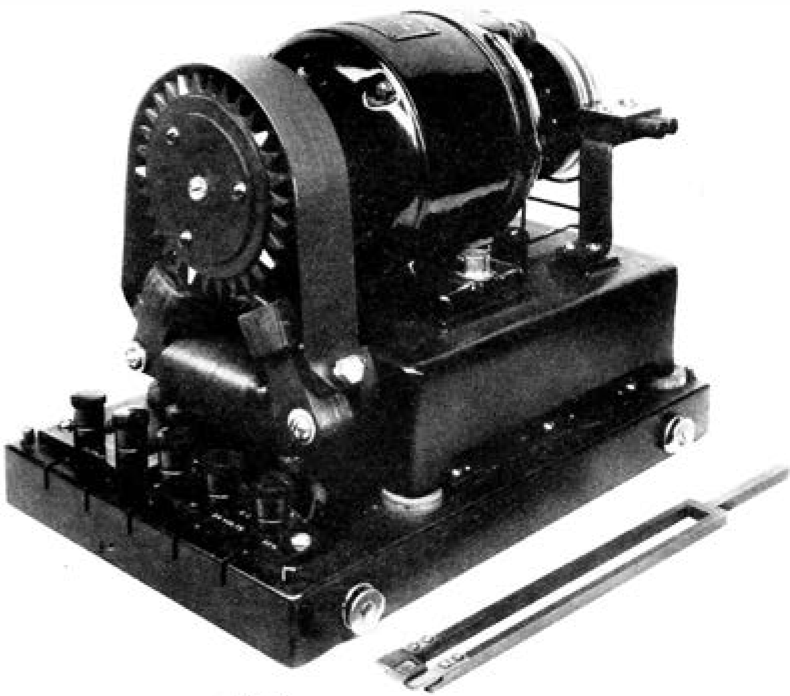


Fig. 4-75. The 2A oscillator, a vibrating-reed unit with a fixed frequency output of 800 hertz.

telephone lines was reasonably assured if they met specified requirements on resistance, leakage, and capacitance. These were characteristics that could be measured by the available dc meter and bridge techniques.

As repeaters and carrier systems were developed, a need arose, first in the laboratory and later in the field, for far more sophisticated techniques that could measure quickly and with a high degree of precision the attenuation of lines and filters as a function of frequency, the characteristics of crosstalk, noise, etc. Fortunately, the very devices (vacuum



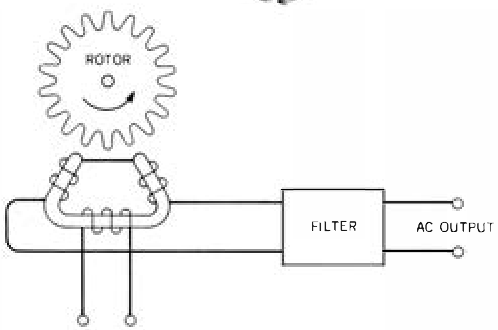


Fig. 4-76. The 3A oscillator, a motor-driven alternator with an output of 1,000 hertz.

tubes and filters) which brought about the need for objective measurements also provided the means for making them.[63]

Two basic devices were required: first, a source of ac tone (an oscil-

[63] Prior to the development of vacuum tubes, means for alternating-current generation and measurement were rudimentary and because of their size or lack of stability were confined largely to the laboratory. In 1903, for example, a 10-kHz generator was so uncommon that G. A. Campbell visited MIT to inspect such a recently acquired device. In 1906 the test gear available in the telephone engineering department laboratories consisted of several single-frequency tone-generators ranging in frequency from 200 to 5,000 hertz,

lator) and, second, an indicating meter for measuring the amplitude of the tone. The latter was usually called a transmission measuring set.

An early oscillator, introduced into the plant in 1921, was the 2A (Fig. 4-75), which was a type of buzzer with a vibrating reed tuned to 800 hertz. This was followed in about 1925 by the 3A (Fig. 4-76), a simple inductor alternator with the motor driven from the 24-volt office battery. By this time, 1,000 hertz had been adopted as a standard test frequency, in place of 800 hertz, and the governor-regulated motor was designed to operate at a speed which would provide that frequency.

The first electron tube oscillator for field use was the 4B (Fig. 4-77). It could be adjusted to generate frequencies between 100 and 3,000 hertz, and was the first source of test tones covering the voice-frequency range.

The 5A oscillator (Fig. 4-78) was designed for routine measurements on toll circuits where it was not necessary to make measurements over a wide range of frequencies. It generated an output which varied continuously and periodically in frequency over a range of 900 to 1,100 hertz, somewhat similar to a siren. Frequency variation was brought about by changing the inductance of an inductometer in the oscillator circuit, the moving coil of the inductometer being rotated by a motor. This "warble tone" technique was particularly useful in measuring circuits for which the attenuation varied rapidly with frequency due to loading or other irregularities. It was later adapted to crosstalk measurements.

---

two shielded bridges for ac measurements together with the necessary standards for inductance and capacitance, and facilities for subjective tests. Some progress had been made in the use of thermocouples for measurement of low-amplitude alternating currents but none of the equipment was suitable for use in the field. When development of the transcontinental line started in the 1911–12 period, impedance measurements on the actual circuits were needed. The Vreeland mercury-arc variable-frequency oscillator had recently become available and was used in obtaining these data.

Fig. 4-77. The 4B oscillator could be set to deliver any frequency between 100 and 3,000 hertz.

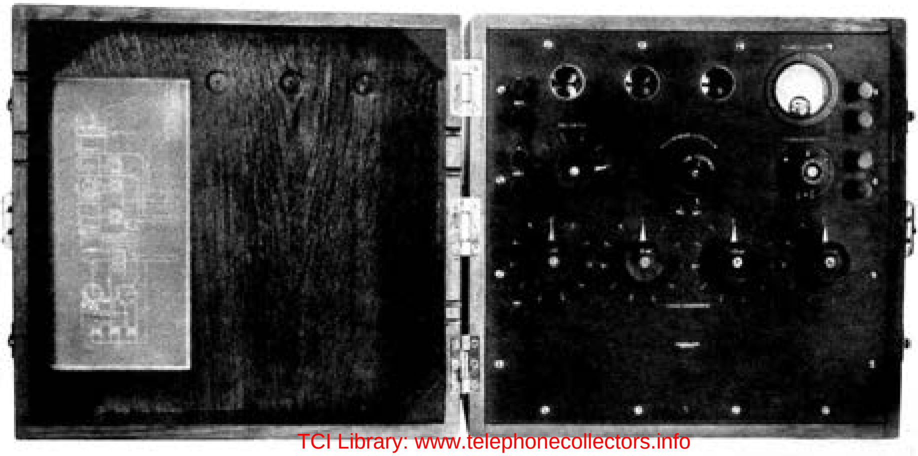


Fig. 4-78. The 5A oscillator provided an output signal that varied continuously and periodically over a frequency range of 900 to 1,100 hertz.

Transmission measuring sets to work with these oscillators began with the 1A, designed in the early 1900s. This set (Fig. 4-79) involved ear balancing techniques for comparing an adjustable known line against the unknown, but used an 800-hertz tone from the 2A oscillator as a source in place of speech. It also employed resistance terminations which could be adjusted to 600, 1300, or 2200 ohms to match the various types of line employed in the plant. Operation, illustrated schematically on Fig. 4-79, was simple. The known and unknown branches were fed in parallel and at the receiving end a telephone receiver was switched from one branch to the other, the known network being adjusted until the sounds were equal. Losses of up to thirty 800-cycle miles could be measured. The 1A set was not provided with a voltmeter to measure oscillator output. Instead, the two branches were bridged together to assure equal voltages on the known and unknown circuit. Thus, both the input and the output of the circuit under test had to be available at the test set terminals. This prevented the measurements of circuits between offices unless two circuits were connected together at the far end to form a loop back to the set.

In situations where the attenuation was very high, the listening technique for obtaining balance was unsatisfactory because of the low vol-

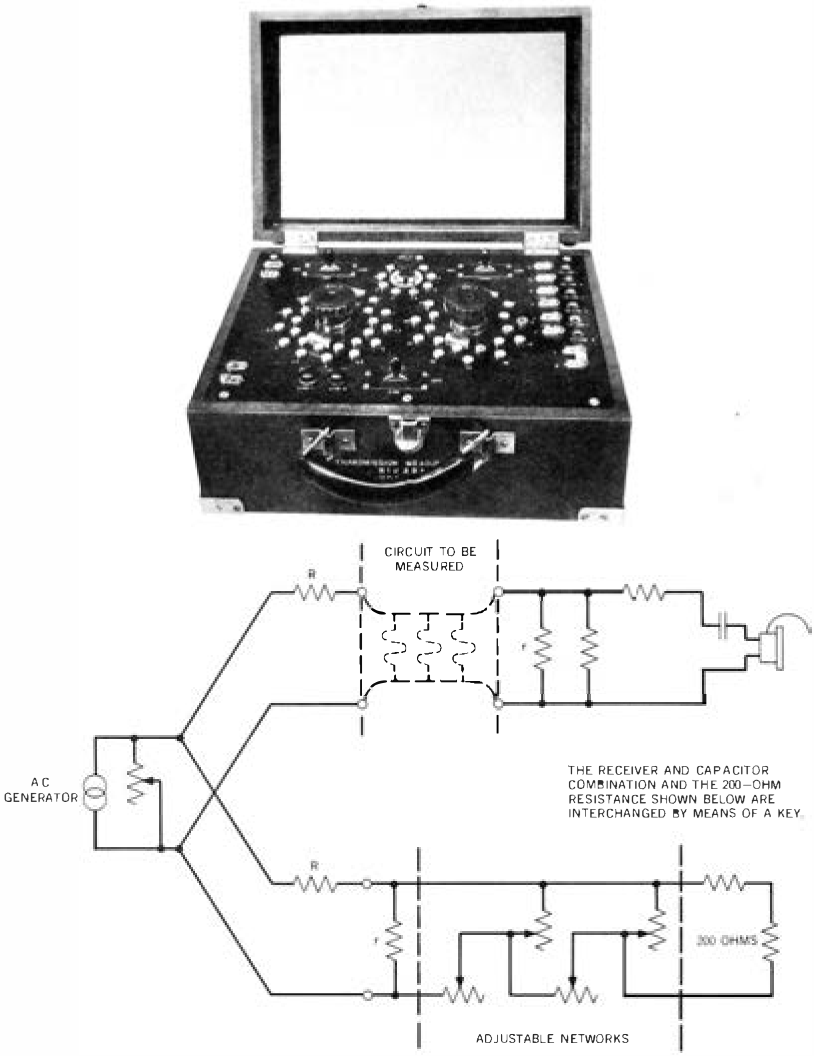


Fig. 4-79. The 1A transmission measuring set.

ume from the receiver. As the vacuum tube amplifier came into use, it was applied as a detector along with a meter to determine balance with very small signals. These principles were incorporated in the 3A transmission measuring set (Fig. 4-80), which also could make "straightaway" tests as well as "loop" tests. The straightaway tests were made possible

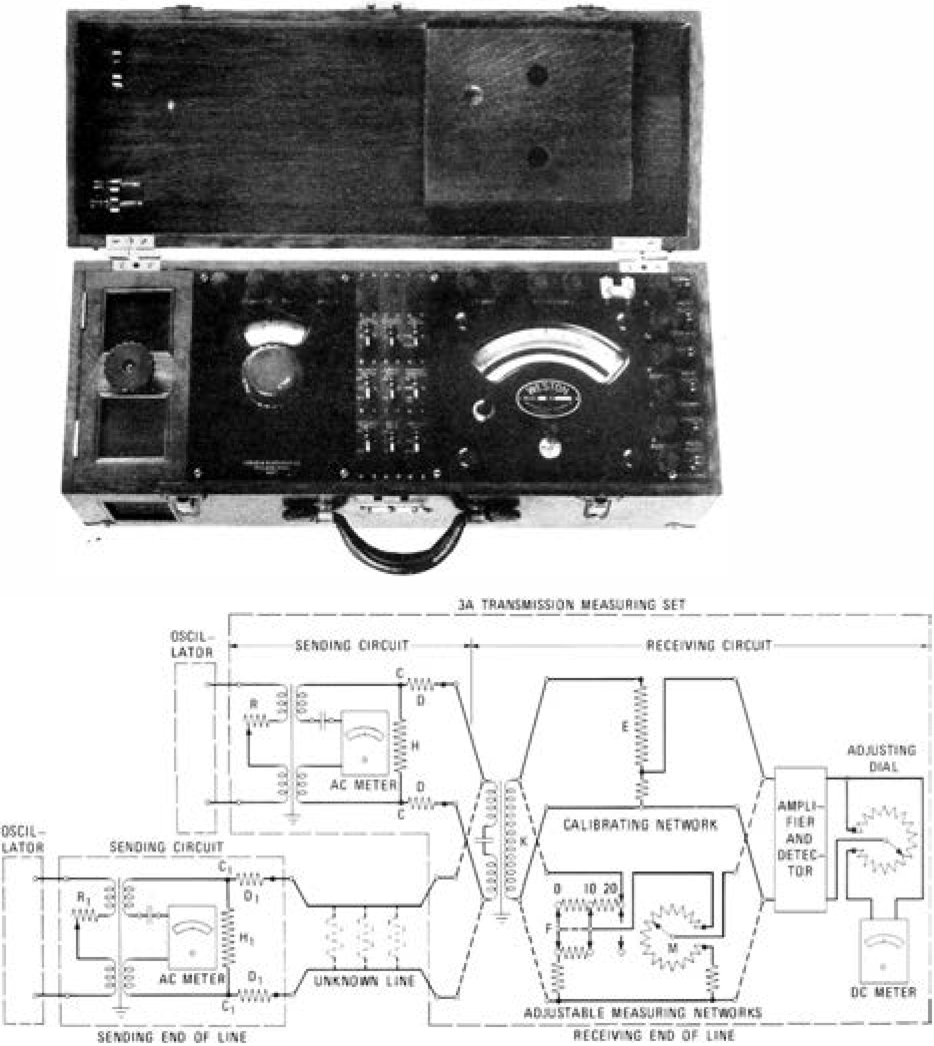


Fig. 4-80. The 3A transmission measuring set.

by the inclusion of a voltmeter for measuring the input to the set from the signal generator. The same voltage was then impressed at the distant end of the circuit under test, and the attenuator dials adjusted for balance. The loss in the circuit was then read directly from the dials. This set was mostly used for 1,000-hertz measurements with the 3A oscillator as a source.

The next set to be developed was the 4A (Fig. 4-81) designed to measure up to sixty 800-cycle miles of standard cable, and also to measure the gain of vacuum tube repeaters. Electrically it was very similar to the 3A except for the increased sensitivity. Because it was so sensitive, extraneous noises on the line under test could lead to incorrect measurements. A filter was therefore provided to eliminate all signals except the signal from the oscillator being used for the tests.

The 4A transmission measuring set with its companion 4B oscillator was intended for permanent installation in the large toll centers. By 1925 about 50 of these sets had been installed throughout the country for measuring and diagnosing troubles on the bulk of the 20,000 long-distance circuits then in service. The 3A set was used at the smaller offices where the more expensive and complicated 4A could not be justified. Exchange area circuits used mostly passive components and did not require frequent testing by sophisticated equipment. In 1925 it was the custom to have all offices tested periodically by men equipped with portable testing sets, such as the 3A. A typical test team with theirequipment is shown in Fig. 4-82.

The 3A and 4A sets established a basic pattern that was used for some time. They employed a precise but not very sensitive meter, usually of the thermocouple type, to adjust the oscillator output to a standard value, usually 1 milliwatt. Amplification was used to provide for measurement of low loss, and filters excluded unwanted frequencies. A high-precision attenuator was used as the measuring standard. The detector and receiving amplifier required only short-time stability since they were switched between the unknown line and the high-precision line, the latter being adjusted to give equal meter reading with the unknown. This was, in fact, only an extension of the basic principle developed many years before in voice testing.

Along with the development of these field testing sets, there was needed even more sophisticated equipment for designing and building the carrier systems and other electronic equipment.[64] Carrier systems required stable oscillators for frequencies of up to about 50 kHz, and meters, attenuators, impedance bridges, etc., all capable of working at these high frequencies. This involved the design of low-inductance resistors and shielded systems to minimize external interference as well as the solution of many other problems.

[64] As an example of the state of the art in 1925, specifications for a "universal transmission measuring set" being designed at the time were as follows:

"Capable of measuring transmission gains from 0 to 111 TU and transmission losses from 0 to 81 TU in 0.2 TU steps between 30 and 10,000 CPS; transmission levels from 0 to +30 TU (without amplifier) in steps of 2 TU; balanced and unbalanced apparatus; and arranged to match impedances from 0 to 26,000 ohms."

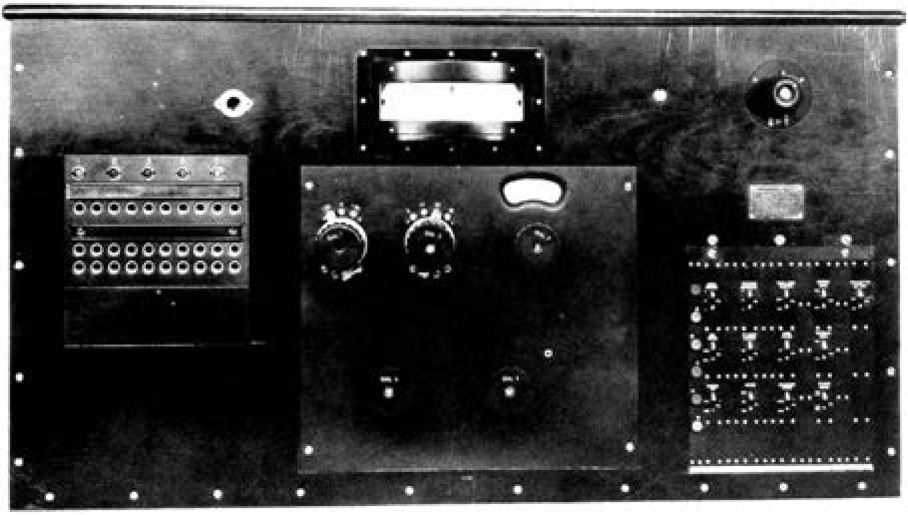


Fig. 4-81. The 4A transmission measuring set.

Fig. 4-82. Typical transmission testing team. (Harden 1925, p. 41)

5.1.4 Supplementary Test Gear

While the measurement of attenuation was of first importance, the need for many other kinds of measurement and related devices arose and was met by the use of vacuum tubes, filters, and other transmission system components. Space limitations permit the description of only a few of the many interesting devices and techniques developed in the fruitful years following the development of workable electron tubes.

5.1.4.1 Frequency Control. With the advent of carrier-current telephony and radio broadcasting (to be discussed in Chapter 5), it became necessary to generate alternating currents with a high degree of frequency stability. In the early 1920s it appeared that the most promising approach would be to use a 100-cycle tuning fork driven by a vacuum tube circuit. Such a frequency standard was built in 1923. Careful control of the many factors which could affect frequency, such as the temperature of the tuning fork, loading of the oscillator circuits, and the like, resulted in a generator that could be relied upon to 1 part in 100,000. At the same time techniques were devised for producing harmonics of the basic frequency so that standard frequencies were available in steps of 100 hertz as high as desirable.

This unit was used as a primary frequency standard for a number of years. It was frequently checked with the primary standard of frequency at the National Bureau of Standards in Washington, and these checks indicated that the desired standard of precision could be maintained.

About 1917, A. M. Nicolson, working in the Engineering Department of Western Electric Company, became interested in the piezoelectric effect of crystals and discovered that the mechanical vibration of a crystal could energize an electrical circuit by coupling through suitable electrodes and that the resulting electromechanical vibratory system performed exactly like an electrically tuned circuit. He built a crystal-controlled oscillator which was operated successfully in 1917 (Fig. 4-83). Nicolson applied for a patent in 1918 but because of various interferences the patent was not issued until 1940.[65]

Nicolson's early work had been with both quartz and Rochelle salt crystals but most of his effort had been devoted to the latter because of its very strong piezoelectric effects even though it required rigorous control of moisture content for stable operation. Later, Dr. Walter Cady of Wesleyan University pointed out the superior stability of the quartz-controlled oscillator and devised circuitry for achieving crystal control that was slightly different from Nicolson's. Subsequently G. W. Pierce of Harvard showed that quartz could be used to essentially duplicate oscillators of the type devised by Nicolson. Ultimately, quartz crystals

[65] A. M. Nicolson; U.S. Patent No. 2,212,845; filed April 10, 1918; issued August 27, 1940.

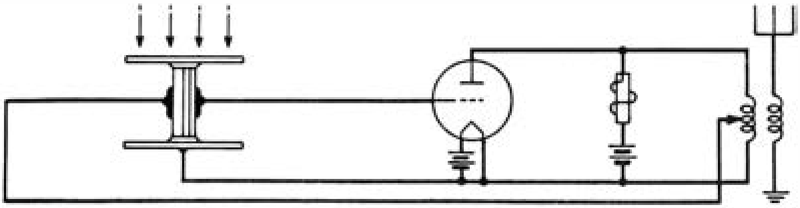


Fig. 4-83. The piezoelectric oscillator circuit shown in A. M. Nicolson's patent.

became widely used elements in both oscillators and wave filters. One of the earliest applications was the control of the frequency of AT&TCo's broadcast station WEAF at 610 kHz. W. A. Marrison ground a crystal to operate at this frequency (Fig. 4-84) and built a simple oscillator circuit which was installed in breadboard form on June 19, 1924. It was used until 1925 when it was replaced by a basically similar but more permanent arrangement.

One of the disadvantages of the crystal-controlled oscillator was the relatively high frequency produced. This made it difficult to operate a clock, as could easily be done with the tuning-fork oscillator. In November 1924, W. A. Marrison suggested a means for accomplishing the reduction of the output frequency of the crystal oscillator, one in which a synchronous clock motor was driven by a second oscillator operating on a controlled subharmonic of the first.

5.1.4.2 Vacuum Tube Voltmeter. Another device of great utility in electronic development was the vacuum tube voltmeter. About 1915, Heising

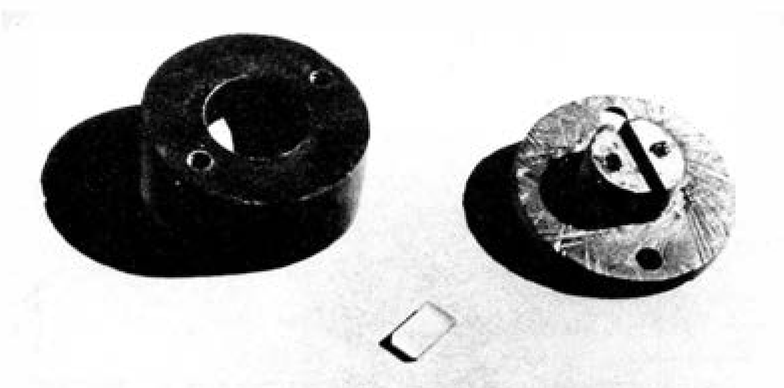


Fig. 4-84. Original crystal and its holder used for the first crystal-controlled broadcast transmission (station WEAF).

noted that a properly biased electron tube, offering an extremely high input impedance over a wide frequency range, could be operated so that its plate current, measurable with a simple dc meter, was related to the ac voltage applied to the input. He used a known dc voltage as a reference standard, with a voltage divider adjusted to give a plate current equal to that with the unknown ac voltage (Fig. 4-85).

5.1.4.3 Volume Indicator. When electron tubes were introduced into telephony, particularly radio telephony, it became desirable to adjust the input level to the maximum the tubes could carry without distorting speech. To accomplish this, E. L. Nelson devised in 1922 the volume indicator (Fig. 4-86), essentially a simple form of damped vacuum tube voltmeter in which the dc meter in the plate circuit followed the envelope of the speech wave. In use, the speech applied to an amplifier was adjusted so that the swings of the volume indicator seldom exceeded a prearranged amount.

With some modification this has become the device now known as the VU meter, commonly used today for electrical measurement of speech levels in volume units.

5.1.4.4 Cathode Ray Oscilloscope. A device which could give a visual trace of the shape of speech waves, or other alternating-current waves, was obviously much needed in pursuing transmission studies. Oscillographs, using mechanical elements such as galvanometer movements, were unsatisfactory because their inertia was too great to follow the frequencies of speech. The cathode ray tube, invented in 1897 by the German, Karl Ferdinand Braun, appeared to be ideal for this purpose since it employed an electron beam focused on a fluorescent screen.

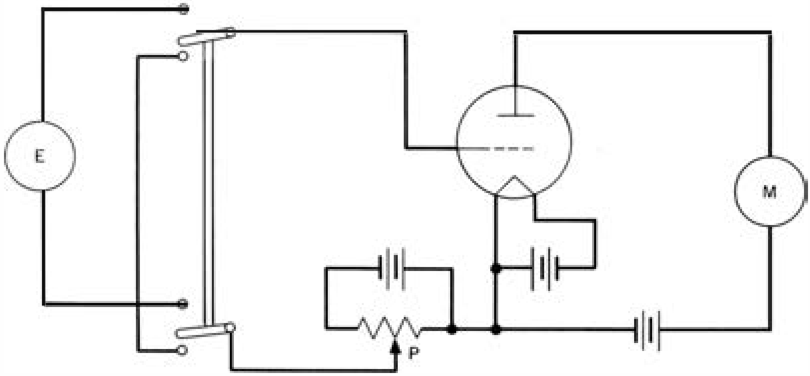


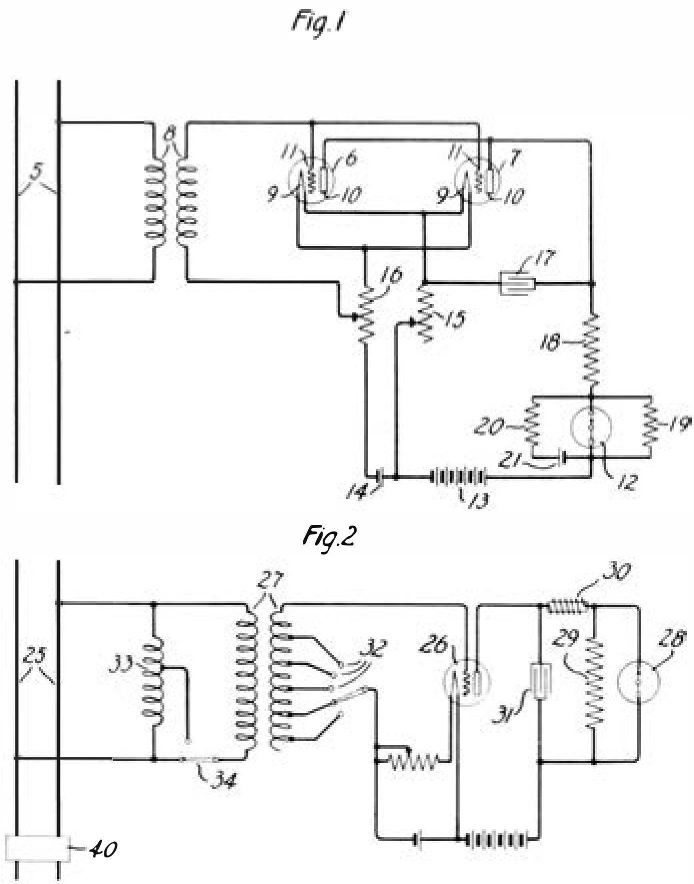
Fig. 4-85. Basic circuit of the Heising vacuum tube voltmeter.

Jan. 20, 1925. **1,523,827**

E. L. NELSON

TRANSMISSION CIRCUITS

Filed Aug. 31, 1922



Inventor
Edward L. Nelson
by Att'y.

Fig. 4-86. Circuits from E. L. Nelson's volume indicator patent.

In Braun's original tube the beam was electromagnetically deflected; subsequent developments made use of electrostatic deflection by means of internal electrodes.

Unfortunately, in its original form the Braun tube required a steady potential of 10,000 to 50,000 volts to accelerate the electron beam. Such a voltage supply was not only dangerous but would, in the twenties, have been bulky and expensive. J. B. Johnson, working in the Engineering Department of Western Electric in the early twenties, solved the problem with a cathode ray tube requiring only 300–400 volts. He accomplished this by using a Wehnelt cathode consisting of an oxide-coated platinum ribbon of the same type that had been developed for long-life repeater tubes. With a copious supply of electrons thus available, the necessary operating voltage was determined primarily by the speed with which the electrons must bombard the fluorescent material on the screen. With the low-voltage electron beam, focusing became a problem, but Johnson's co-worker H. J. van der Bijl solved this ingeniously by introducing a small amount of gas in the tube. Ions from the gas, interacting with stray electrons, set up a field surrounding the beam which tended to pull the electrons inward and so counteract the spreading due to mutual repulsion within the beam. These techniques made the cathode ray oscilloscope a practical laboratory tool and Western Electric furnished many cathode ray tubes to experimenters and to oscilloscope manufacturers.

5.2 Transmission Objectives and Standards

It is difficult to describe the grade of transmission provided during the first 50 years of telephony in terms that have significance today. As noted previously, transmission performance and standards were specified on a volume basis in terms of the standard cable reference system. The telephone instruments used commercially, and in the reference system, were designed for maximum efficiency at about 1,000 hertz, the response falling off rapidly above and below that frequency. With such instruments it was possible to achieve maximum speech loudness and consequently maximum transmission distance at a time when the conquest of distance was a major objective of transmission engineers. With a little practice, speech from these instruments proved intelligible enough to be usable but was highly distorted and obviously very much poorer than today's reader would infer from the attenuation of the reference system. In some cases (the first transcontinental line, for example) the performance was quite poor by present standards but was considered commercially usable under the conditions then current.

Another problem in describing early transmission performance arises from the fact that the standards were usually specified in terms of limiting loss, i.e., the loss when all components of the plant (the subscribers' loops and the trunks) were at their limit. Obviously, many loops and trunks were much better than the limiting value, so that the average performance was considerably better than the limit, but it is difficult to say, at this late date, how much better or to estimate the probability of obtaining a limited connection. On the other hand, limiting losses were occasionally exceeded during bad weather, or under trouble conditions, or on particularly long switched connections.

A final problem is the lack of specific and detailed information on standards in the early years. Much was left to the judgment of individual managers, and the major criterion was often pragmatic. A circuit that could be used, even if many repetitions were required, was obviously better than no circuit at all, regardless of what headquarters said about standards. The availability of loading and repeaters made it much more feasible to work to a systematic set of standards and, generally speaking, limiting losses of about 20 miles of standard cable for local calls and 30 miles for long-distance calls were fairly common in the period just following World War I and in the early 1920s. By the late 1920s and early 1930s limiting objectives of 15 to 18 dB (the numbers are approximately the same as numbers of miles of standard cable) for local calls and about 25 dB for toll calls were more common.

It is interesting to look at the attenuations characterizing some of the early, very long haul circuits. Prior to loading, 165-mil copper wire circuits were the best available, and this facility was used at the turn of the century for distances of over 1,000 miles. With an attenuation of 0.033 dB per mile this meant a toll line loss of about 35 dB in dry weather. To this must be added the loss of the customer's loop and the trunk from his office to the toll office, which together certainly added another 5 to 7 dB at each end. The first transcontinental line was only a little better. The loss of the line and associated apparatus in dry weather was about 60 dB, this being reduced to 20 dB by the use of repeaters. However, the band transmitted was only about 900 hertz wide and this narrow band introduced a significant impairment as compared to the early circuits using nonloaded facilities. Under bad weather conditions the line loss practically doubled, and while the repeater gain was increased by manual adjustment to keep the overall loss within reasonable bounds, it was not practical to compensate fully for the losses introduced by rain, frost, and sleet. Fortunately, improvements in transcontinental telephony came along rapidly. By 1920, despite the delays resulting from engagement in World War I, the number of repeaters that could be used successfully on the line had been increased to 12, which permitted the unloading of the entire line. These two changes, together with improved repeater balancing, permitted the reduction of the net loss on the transcontinental line to 11 dB and permitted doubling the bandwidth. The elimination of loading lowered the impedance by a factor of 3.5, thus reducing the effect of leakage during bad weather, and, even more importantly, increased the speed of transmission by the same factor of 3.5, thereby reducing the effects of echo to where they were insignificant.

5.3 Control of Interference

Several kinds of external electrical waves, picked up by telephone lines, can interfere with proper reception of the wanted telephone signal.

One of the most troublesome in the early days, and later when carrier techniques were introduced, was unwanted speech from nearby telephone circuits. Generically, such interference is known as "crosstalk," but often the term is used only to designate intelligible speech interference. When crosstalk comes from many circuits, the various disturbers may so interfere with each other that the result is unintelligible and it is referred to as "babble" or "crosstalk noise." Babble is undesirable, as are other forms of noise, because it provides an annoying background of sound and may interfere with reception by covering up, or masking, the weaker speech signals. Intelligible crosstalk is particularly undesirable since it represents a violation of the privacy of communication. Carrier transmission can give rise to both intelligible crosstalk and babble similar to that arising in voice-frequency circuits. In addition, the frequency translation process may produce a form of crosstalk that has the speech frequencies inverted or displaced as compared to the original. Such speech sounds are much less intelligible than simple crosstalk, but the speech origin is recognizable from the rhythm of the sound. As noted in the discussion of Type C carrier, systems are often designed to produce inversion or displacement by proper choice of carrier frequency and thereby attain a so-called "staggering advantage" which permits crosstalk levels several dB higher than with intelligible speech. This technique is advantageous when there are only a few sources. When there are many staggered sources there is a babble similar to that from unstaggered sources.

Atmospheric static, another source of interference of importance in the very early days, was brought under control rather easily but again became of considerable importance with carrier. The level of atmospheric noise is greatest at low frequencies and falls off rather rapidly as frequencies increase. Ground-return telephone lines acted as long radio antennas and on such lines this type of interference was audible even with relatively inefficient telephone receivers, particularly a randomly occurring swishing or whistling sound. Long after atmospheric noise on voice-frequency circuits was eliminated by the use of metallic lines, these "whistlers" were identified as waves, generated by lightning discharges in the atmosphere, that had traveled many times around the world and had been spread out in time because the frequency components had been differently diffracted and hence traveled paths of different lengths in their long journey. The lower-magnitude, high-frequency sounds only became important in the 1920s and later when the high amplification used with carrier raised them to levels which could interfere with speech. These crackling-type noises were produced only by rather nearby thunderstorms and occurred very infrequently. During local thunderstorms, however, they could be quite high in level and were important factors in determining carrier system power level and repeater gain.

Electric railways and power lines can also be a serious source of noise. Neither of these was of much importance when the telephone was invented, but within ten or fifteen years arc lighting, electric railways, and alternating-current power transmission were introduced and commenced a rapid growth. It might appear that little interference would result from these applications of electricity because they employ frequencies well below the telephone range. In practice, however, the waveform is far from "pure" and contains many high frequencies superposed on the basic direct-current or 60-hertz power frequency. The use of metallic lines did much to bring this type of interference under control, but as time went on and communication and power transmission developed, literally side by side, further measures were required.

In addition to being a cause of annoyance by inducing noise, both atmospheric and power currents could be a source of danger. Nearby thunderstorms could induce high potential on telephone lines and direct lightning strikes could cause personal injury to a telephone user or seriously damage telephone equipment. Very long and close exposures to power lines could result in the buildup of undesirably high voltages on telephone lines and, even where such voltages were normally not excessive, the heavy power currents during power line faults could cause a dangerous situation. Obviously, a direct "cross" between telephone and high-voltage lines was a matter of serious concern. From the beginning of telephony the necessity for protecting the telephone user and telephone plant from these potential dangers was recognized and much effort has been devoted to the development of adequate protective measures ever since the very early days.

The following sections discuss briefly some of the work, during the first 50 years of the telephone, that was devoted to control of voltages from external sources so as to reduce crosstalk and noise and protect the user against physical harm.

5.3.1 Crosstalk Control

Practically from the beginning, the control of crosstalk was an important factor influencing the design of the transmission medium itself. In the very early days, attenuation was the controlling factor in determining limiting transmission distance, but once amplification became available, crosstalk often determined ultimate system performance in terms of distance and in other ways. This was particularly the case with carrier systems.

In the early years, as we have noted, metallic circuits were adopted in place of ground return in order to reduce crosstalk and noise. They proved rather disappointing in reducing crosstalk, particularly on open wire lines, but the metallic principle was indirectly responsible for interference control since it permitted the use of measures which ultimately proved the basis for success in this field.

At this point we should refer back to Fig. 4-6 and note that the current carried by one pair on a line is induced unequally in the wires of adjacent pairs giving rise to a residual, unwanted, "crosstalk current." The crosstalk current may be transferred from the disturbing circuit by inductive coupling (i.e., by magnetic lines of force) or by capacitive coupling (i.e., by the capacitances between the various wires).[66]

Crosstalk can be reduced by any means that tends to equalize the current induced in the two wires. One way is to have the wires in the pair close together relative to the distance from the disturbing pair, and this principle has been used to some extent. Another approach is to use configurations such that the electric and magnetic field from a pair has the same effect on each wire of the disturbed pair. Such configurations are shown in Fig. 4-87. (The left one represents the "star-quad," used extensively in European cable construction.) Neither of these configurations has proved practical for open wire construction. The most effective measure for open wire by far has been the transposition illustrated in Fig. 4-6.

The transposition principle appears very simple since, offhand, it seems only necessary to balance out the undesired induced currents by "turning over" or transposing the relative positions of the wires of the disturbing and disturbed circuits. Physically, this can be accomplished very easily on cable pairs by giving them a twist every few inches. In practice, this by itself was not sufficient since it was found that irregularities or differences in the resistance or capacitance

[66] The possibility of transferring crosstalk through capacitive coupling seems to have been questioned by some practitioners. J. J. Carty, by means of a series of ingenious experiments, reported in 1891, clearly demonstrated the possibility but he seems to have overestimated somewhat the importance of this kind of coupling on typical open wire lines.

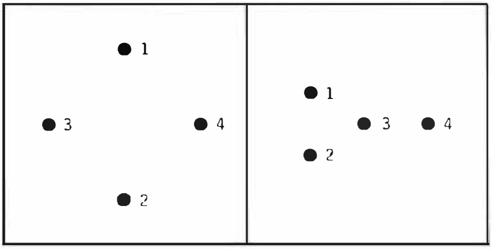


Fig. 4-87. Crosstalk-free wire configurations.

of the wires in a pair would prevent complete cancellation of the crosstalk current. This problem was solved, as we have noted earlier, by developing cables with wires having a high degree of uniformity and in some cases by correcting for the residual unbalances by compensating splices. Residual balances were also corrected by the use of corrective capacitors or inductors. The latter technique was to be used extensively with the introduction of long-haul cable carrier (Type K) in the late 1930s, and this and other ingenious techniques were used in applying carrier to cable in the second half-century of telephony. Until that time, adequately low crosstalk was obtained on cables by using careful control in manufacture, compensating splices, and an adequate number of twist lengths to provide frequent, relative turnovers on adjacent or nearby pairs.

The physical problems were greater with open wire since it was desirable to minimize the number of transpositions because they complicated line construction and were costly to make. In addition, the wires of the pair were far enough apart so that some crosstalk was caused by nearly every pair on a pole line. Crosstalk on open wire, as in cable, was influenced by unbalanced capacitance or resistance and, in addition, rather small irregularities in construction were important. For example, as carrier techniques were introduced, it was found that small differences in wire sag caused significant unbalances.

As a consequence, much work was devoted to the design of transposition systems that would adequately meet current needs with minimum transposition cost. This work, which began in the 1890s, continued for 40 or 50 years as new transmission techniques were introduced which complicated the crosstalk problem. Some appreciation of how these problems were solved is essential to understanding the growth of transmission, but first we should lay some further groundwork.

The discussion and illustration of crosstalk so far (Fig. 4-6) has been rather elementary and has not recognized a number of complicating factors. Consider, for example, the very simple situation with two pairs,

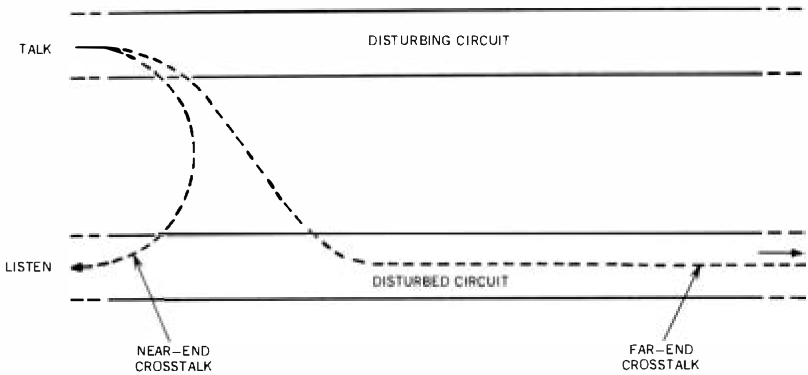


Fig. 4-88. Near-end and far-end crosstalk. (Redrawn from Babcock, Rentrop, and Thaeler 1955, Fig. 17)

one acting as a disturber and the other as the disturbed circuit (Fig. 4-88). Crosstalk induced in a short length of exposure travels in both directions from the point at which the induction occurs. That part going toward the same end of the line as the disturbing source is called "near-end crosstalk" and is designated NEXT. That part going to the distant end is "far-end crosstalk" and is designated FEXT. In each case crosstalk results from both capacitive and inductive coupling. For near-end crosstalk the two effects tend to add, whereas for far-end crosstalk the effects are in opposition. Thus, on short circuits NEXT may be rather more severe than FEXT, but this is not necessarily true for long circuits since the near-end effect is determined very largely by coupling at the sending end[67] whereas FEXT adds up throughout the line. In the days before the use of repeaters, far-end crosstalk was often the controlling factor on long open wire lines.

With the advent of repeaters the importance of NEXT was greatly increased. Consider the situation illustrated in Fig. 4-89. At the output of the west amplifier of the disturbing pair, the voice currents are at their highest level because they have just been amplified and *the NEXT currents are correspondingly high* on the disturbed pair. On the disturbed pair at the same point the wanted voice currents coming from the opposite direction are at low level because of the attenuation involved between the east and west repeaters

[67] The technical reader will note that the speech on the disturbing circuit is attenuated by the line as the distance from the sending end increases. With NEXT the crosstalk current is also attenuated on its return over the disturbed circuit. Hence the effect of coupling at a distance is small. With far-end crosstalk the wanted and unwanted (crosstalk) sounds are attenuated together and consequently the ratio of the unwanted to wanted sounds increases with distance.

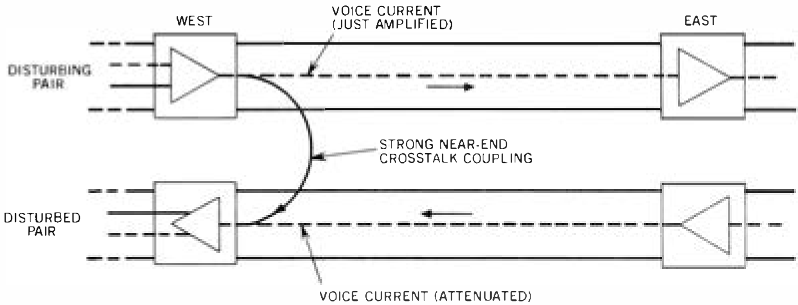


Fig. 4-89. Near-end crosstalk on repeatered circuits. (Redrawn from Babcock, Rentrop, and Thaeler 1955, Fig. 18)

and both the wanted and unwanted currents are amplified as they propagate to the west. Thus, crosstalk current generated through the near-end coupling at the west repeater is amplified by the west amplifier on both the disturbing and the disturbed circuits. A similar analysis of FEXT will show that it is (for this simple example) unaffected by repeaters since they amplify the wanted and unwanted currents equally. Thus, when repeaters were introduced, NEXT became of very great importance. At carrier frequencies, where the coupling between circuits and the amplifier gain are both very high, it becomes impractical to control NEXT by means of transpositions, so that other techniques must be used. In cable carrier a common method is to use different cables (or the equivalent shielded groups) for the different directions of transmission so that there is essentially no NEXT exposure. On open wire lines the same thing is accomplished by using different frequencies for the two directions of transmission.[68]

So far, we have been discussing direct crosstalk, that is, crosstalk resulting from direct coupling between two circuits. Indirect crosstalk also occurs when there are more than two circuits on a line. The disturbing circuit crosstalks into a tertiary circuit which in turn crosstalks into the disturbed circuit. In fact, some form of indirect crosstalk is unavoidable, since even with only two pairs on a line, there are always tertiary circuits in existence such as the simplex circuit with ground return or the phantom circuit. In actuality, therefore, the crosstalk on a line is always the sum of that received over the direct and all of the indirect paths. In many cases indirect crosstalk is a second-order effect of not too great significance, but it can play an important part in the design of transpositions since it is obvious that the tertiary modifies their effect.

[68] The Type N cable carrier system introduced in 1950 also used different frequency bands for the two directions.

A particularly important type of indirect crosstalk can occur on repeatered lines. One such is illustrated in Fig. 4-90, which shows how crosstalk through two near-end paths can be amplified twice and can end up as high-level far-end crosstalk. Such crosstalk, usually referred to as near-end-near-end interaction crosstalk, can be quite important when only a few carrier systems are installed on a large pole line or cable since the only cure for this type of crosstalk requires the blocking of carrier signals in the many unequipped tertiary circuits.

Many other aspects of crosstalk had to be analyzed and their effects determined before adequate transposition systems could be designed for carrier systems. Without going into more of these situations, it should be apparent that the art of crosstalk reduction, particularly at carrier frequencies, required a highly sophisticated approach involving both theoretical analysis and a large amount of experimental effort. Much of the latter work was carried out on an abandoned open wire line at Phoenixville, Pennsylvania, which was also used for testing insulators and determining the attenuation at carrier frequencies of open wire lines having different configurations and operating under

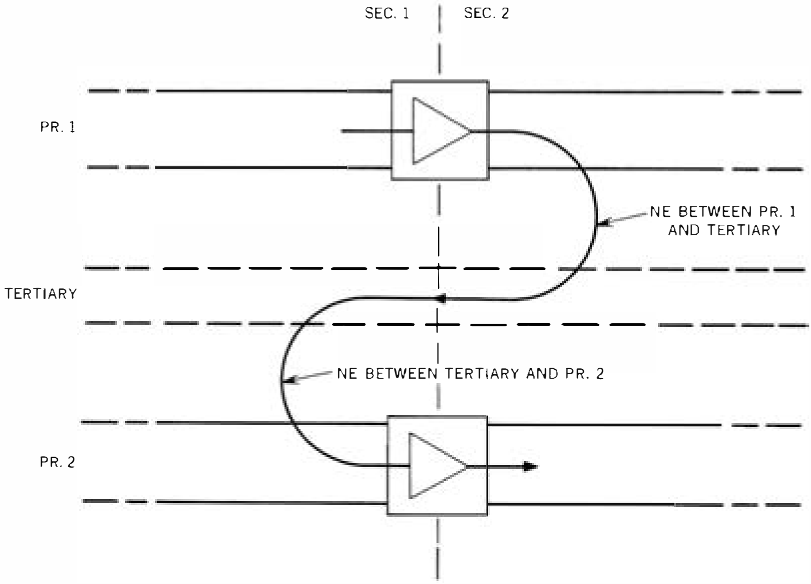


Fig. 4-90. Near-end-near-end interaction crosstalk. (Redrawn from Babcock, Rentrop, and Thaeler 1955, Fig. 22)

various weather conditions.[69] By the middle of the 1930s the theory of crosstalk had been well developed and had been tested to the point where transposition could be carried out on a more or less routine basis. This work was largely done by A. G. Chapman, building on theoretical contributions by G. A. Campbell and R. S. Hoyt.

Before concluding the discussion of crosstalk we should review the physical development of transposition systems and see how the evolving theoretical crosstalk studies were reduced to practice.

The beginning of the long period of transposition design effort dates from 1885–1886. At that time the first line of the newly formed long-distance company (AT&TCo) was being completed between New York and Philadelphia. It consisted of 12 metallic circuits, the first instance of such a large number on the same long-distance pole line. Contrary to previous experience with lines having fewer circuits, the crosstalk performance was nearly as bad as on ground-return circuits. As was stated in Section 2.2, dealing with open wire lines, John A. Barrett was engaged to study the problem and work out a solution. The general concept of reducing crosstalk by twisting wires was not new at this time, but there was no prior experience in applying the technique to open wire lines. Despite the lack of background, Barrett devised within a few months the so-called ABC transposition scheme shown in Fig. 4-91. With this system successive transposition poles, spaced about 1,300 feet or ten poles apart, were lettered A, B, C, B, A, B, C, B, A, and so on. On poles designated by the same letter, the same circuits were transposed as shown in the figure.

This system was easy to install because of its repetitive pattern, and worked reasonably well when the number of circuits was small, but later was found to be inadequate with large numbers of circuits. Some of the reasons for this now appear rather obvious. Since all alternate crossarms were transposed alike, there was considerable coupling between arms 1 and 3, and also between 2 and 4. In addition, the outermost pairs on a given arm were transposed at the same point and hence were not transposed relative to each other. The ABC system was therefore superseded in 1898 by the so-called standard system which remained the basic pattern for many years. This new system

[69] One of the interesting outputs from this work was the suggestion by L. T. Wilson to replace the previously used wooden insulator pins with metal pins and a low-resistance electrical connection between the pins associated with the two wires of a pair. At high carrier frequencies, the current flowing from wire to wire through the insulators was largely determined by the insulator capacitance. Reducing the resistance path between insulators had negligible effect on this current but reduced the power loss (I^2R) through the path with corresponding decrease in line attenuation. This technique also reduced the variation in loss between wet and dry weather.

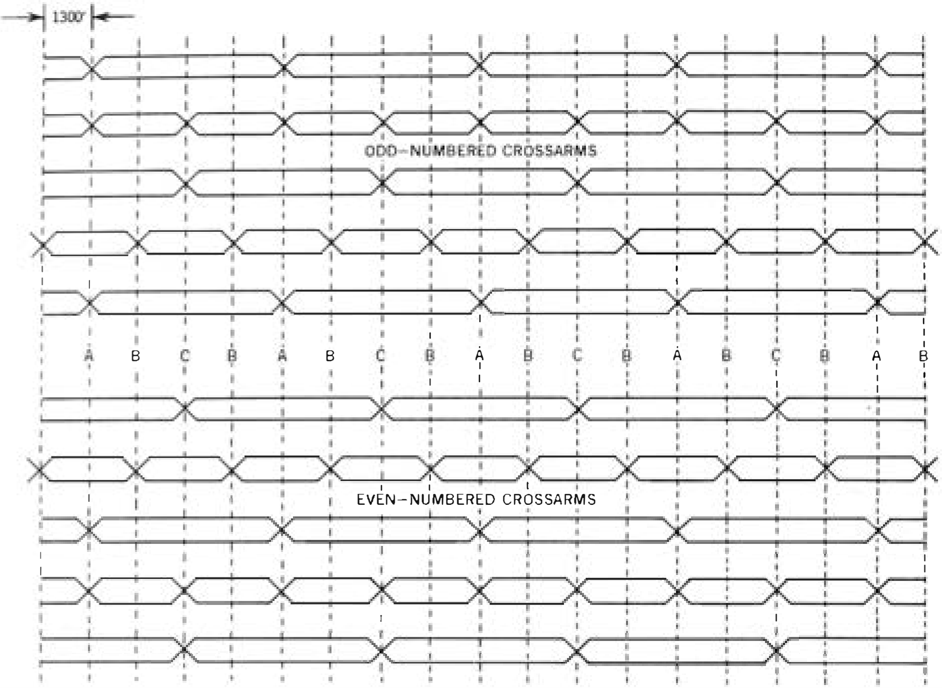


Fig. 4-91. The ABC transposition system, shown above, was the first to result commercially from Barrett's invention. (Redrawn from Hill 1949, Fig. 1)

and modifications made in it much later were all based, to a large extent, on principles outlined in the Barrett patent of 1888.[70]

The standard system was based on the use of repetitive 8-mile sections consisting of 320 spans (i.e., spaces between poles) each 130 feet long. This basic section was divided into 32 transposition intervals of ten spans each. For the basic section Barrett devised a system which contained 32 different transposition types, as shown in Fig. 4-92, which varied all the way from 0 to 31 transpositions in the section. With such a scheme 32 pairs could be used, each of which would have one or more transpositions relative to every other pair. For example, a Type O pair had 1 transposition relative to the un-

[70] This patent outlined the need for relative transpositions and showed how a scheme could be built up on a given basic section of line using 0, 1, 2, 3, etc., transposition points on different pairs. It is apparent that Barrett had a remarkably sound understanding of transposition design, but in terms of present knowledge he somewhat overestimated the importance of diagonally adjacent pairs and also overestimated the shielding effect of nearby pairs.

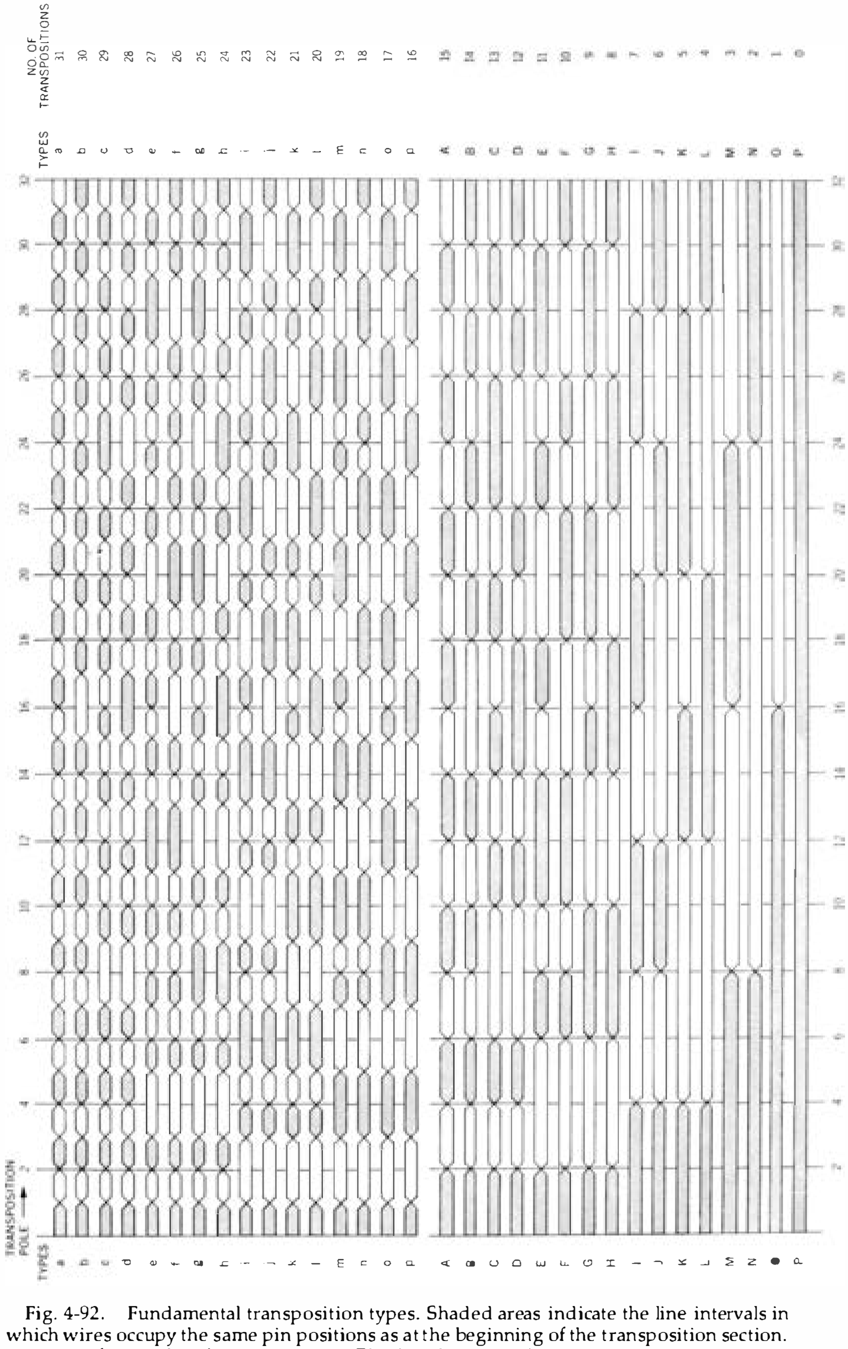


Fig. 4-92. Fundamental transposition types. Shaded areas indicate the line intervals in which wires occupy the same pin positions as at the beginning of the transposition section. (Redrawn from Babcock, Rentrop, and Thaeler 1955, Fig. 1)

transposed Type P pair (or to the 2-transposition Type N pair). Similarly, Type a had 1 relative transposition as compared to Type b, but it had 31 relative to P. This scheme provided great flexibility since the types could be arranged to provide a large number of relative transpositions between closely coupled pairs and fewer between remote pairs. Even a full six-arm line, with 30 pairs, could be transposed without repeating types within the transposition section.

This basic system worked very well for voice-frequency circuits but was not suitable for the higher-frequency carrier currents, and a number of line and transposition modifications were introduced to meet carrier needs. One of the first requirements was the use of shorter distances between transposition points. The reason for this is that a transposition interval, to be effective, must have very little attenuation and phase shift. Otherwise the induced currents on the two sides of the transposition may not be sufficiently alike to cancel out. This requirement was met satisfactorily with the basic system in the case of voice frequencies but not with carrier. For the latter, a scheme was devised whereby a transposition would be placed in the middle of each interval of the basic scheme. This resulted in 32 extra transpositions for each letter type in an 8-mile section and, in effect, gave 32 additional transposition types to use along with the original 32 fundamental types. They were called "single extras." Later, as frequencies were increased and shorter intervals were required, double and higher-order extras were added. Throughout this evolution the basic types devised by Barrett were retained but the use of ten-span intervals was not satisfactory for the higher-order extras since they sometimes resulted in unequal numbers of spans between transpositions. To reduce this problem the basic interval was reduced to eight, and the basic section to 256 spans (6.4 miles).

These changes took care of the attenuation and phase shift problems, but carrier also increased the effect of minor physical differences between pairs and transposition intervals. Improvement in these factors was required in the middle 1930s when long-haul carrier (Type J) using frequencies up to 143 kHz was introduced. These improvements were accomplished by reducing the spacing between the wires in each pair to 8 inches (with an increase in the distance between pairs) and by eliminating both phantom and pole-pair circuits which introduced hard-to-control tertiary circuits.[71] In addition, point-type transpositions were used in place of drop brackets (Fig. 4-7) and the regularity of sag and pole spacing were improved. Where uniform spacing was not practical because of natural obstacles,

[71] It was at this time that the crossarm configuration shown in (c) of Fig. 4-4 was adopted.

transpositions were made within the span by means of a bracket (Fig. 4-93) supported by the wires between poles.

In order to simplify the discussion of transposition, it has been carried out in terms of means for handling pairs. Obviously, phantom groups also had to be transposed and suitable schemes were provided for voice-frequency circuits. However, they greatly complicated the use of carrier and when 12-channel carrier became available, the use of phantoms was gradually eliminated since a carrier circuit proved less costly than the transposition and other line rearrangements required to provide the phantoms.

5.3.2 Noise Control

A major step in noise control was taken with the introduction of metallic circuits. For more-or-less-distant sources of power or railway noise, this measure balanced to a very large extent the induced noise in the two wires of the pair. When the source of interference was close and the exposures were long, the unbalance was imperfect, but the transpositions introduced to control crosstalk were also highly effective in reducing induced noise.

The control of crosstalk did not automatically control noise. Crosstalk coupling tended to increase with frequency and hence crosstalk balancing was most important at the middle and upper voice

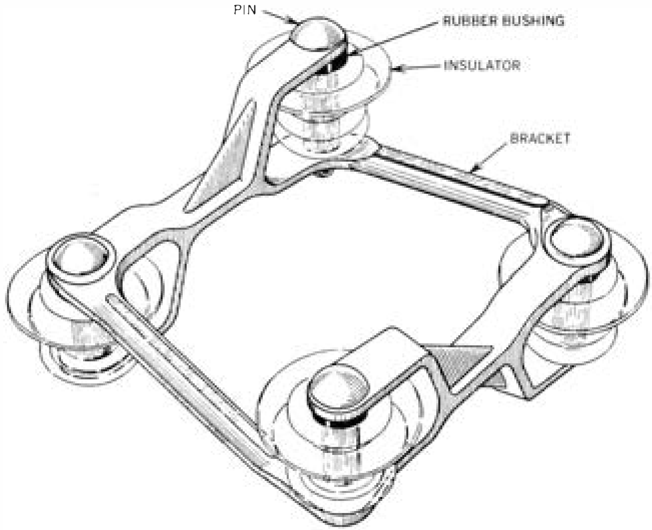


Fig. 4-93. Span transposition bracket. (Babcock, Rentrop, and Thaeler 1955, Fig. 16)

frequencies. Noise from power and railway sources was largely concentrated in the lower voice range and its successful reduction often required specific measures to achieve good balance at these frequencies. For example, it was common practice to call a telephone user to his phone by means of a 16-hertz ringer connected between one side of his loop and ground. Such ringers had a reasonably high impedance in the upper voice region and hence the unbalance to ground was not very important from a crosstalk standpoint. However, the ringers, of necessity, had to have low impedance at the ringing frequency and consequently the unbalance tended to be significant at 180 and 300 hertz which were important noise frequencies. Consequently, arrangements had to be devised that were adequate for ringing but with sufficiently high impedance at low frequencies to keep noise at acceptable levels. Similarly, central-office circuits had to be designed so that they performed their major functions properly and still were acceptable from the standpoint of balance at the important noise and crosstalk frequencies. Thus, over the years, a high degree of cooperation was required between the developers of telephone station apparatus, switching equipment, and transmission systems.

Prior to 1900 the telephone and power industries developed more or less independently, adequate noise control being obtained through the relatively simple means mentioned. However, interference problems increased rapidly in the first quarter of the twentieth century as both industries not only grew rapidly but also strove to better meet their customer needs. The communication industry raised its standards of service by placing lower limits on acceptable interference and by introducing amplification and higher efficiency receivers. The power companies spread their lines to give wide geographic coverage and delivered more and more electrical energy to their customers. Separation between the two types of lines was decreased, exposures became longer, and interference problems multiplied. A joint committee on inductive interference made up of telephone and power engineers conducted studies and tests in California in 1913–1917 in an attempt to resolve this growing problem, but all too often litigation was becoming the accepted approach with each side attempting to force the other to solve the problem.

In the latter part of 1920 it was evident that the situation was menacing both industries, and in early 1921 a group of power and telephone men met under the chairmanship of Owen D. Young of the General Electric Company to form a permanent committee to work out jointly a solution to the interference problem. This group, which later became known as the Joint General Committee of the National Electric Light Association (NELA) and the Bell Telephone System, established a subcommittee made up of power and telephone engineers headed jointly by R. F. Pack from the NELA interests and Bancroft Gherardi of AT&TCo. This subcommittee was charged with identifying the situations giving rise to the interference problem and with determining the best technical solution regardless of the division of cost.

During the next year or so, progress appeared to be slow, but in retrospect it is clear that the major step in solving the interference problem had already been taken when the decision was made to find solutions on an engineering basis instead of through litigation. This was easier said than done, since it was soon found that the technical information on which to base corrective measures in the power and telephone field was often not available. It was necessary, therefore, to organize theoretical and experimental programs to provide the necessary data. This was done by telephone and power groups operating in various parts of the country where suitable facilities were available, and a long series of reports on these investigations was issued by the Joint Committee beginning about 1925 and extending over a period of years. These reports covered the characteristics of power lines and loads causing harmonic generation, the telephone line and equipment factors affecting coupling, the effects of noise on the transmission of speech, means for measuring telephone-line noise and power-line interference factors, and so on. Although some of this work was done prior to 1925, the major part of the effort was reported on and applied later. It is sufficient to point out here that the basic groundwork was laid by the formation of the 1922 Joint Committee and that nine years later Mr. Pack, in opening a symposium, was able to say, "Today inductive coordination as between the Bell Telephone System and the power companies is no longer a problem but only a routine day-to-day job of cooperatively continuing research work and developing the art of both systems to eliminate as far as possible causes for inductive interference."

5.3.3 Electrical Protection

At the beginning, the problems of electrical protection were basically the same as those faced by the telegraph industry and were solved in the same manner. Practically the only danger came from lightning surges, and protection was provided by a small spark gap between the wires and ground. The very simplest form consisted of two pieces of silk-insulated wire twisted together. A lightning surge would puncture the insulation and provide a path between the line wire and the one connected to ground. A more highly developed arrangement was made up of two small metal plates, the edges of which were separated by an air gap. Often the edge of one plate was serrated to promote more uniform breakdown voltages. Such a protector is shown just above the ringer gongs on the telephone pictured in Fig. 4-94. This telephone was one of the first in commercial use (circa 1878), and illustrates the very early recognition of the telephone protection problem. The line could be grounded as further protection during a storm by inserting a brass plug in the circular opening between plates.

The telephone problem soon required special techniques, as telephone lines were extended into the same areas where the growing electric power and transit industries were spreading. With the increasing proximity to these systems, the dangers of high voltages and currents from induction or accidental contact increased. Furthermore, telephone instruments were small and particularly sensitive to high voltages which could break down the wire insulation and to heavy currents which could overheat the wires and present a fire hazard. Finally, since telephones were installed in homes where technical skills were unavailable, very safe and maintenance-free protective devices were required. These factors set the basic requirements for protective devices, namely, they must protect the user from physical harm, avoid the buildup of voltages and currents that would damage equipment, and operate with minimum interruption to service.

Various devices using the fuse principle were introduced to provide the needed protection but were not entirely satisfactory. In 1890, Hammond V. Hayes made a comprehensive study of the protection problem. He concluded that a protection system should contain three elements:

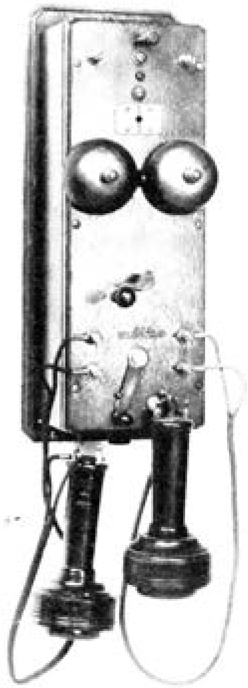
(*i*) A *spark gap* connected between each wire and ground which would break down at voltages exceeding about 400 volts.

(*ii*) A *fuse* in each wire (on the line side of the gap) which would interrupt the circuit when the current continued after the gap breakdown and exceeded 8 to 10 amperes.

(*iii*) A *heat coil*, or sneak-current arrester, which would operate on a persistent voltage too low to operate the gap but high enough to cause currents damaging to the telephone equipment. Such currents would gradually increase the temperature of the heat coil until a fusible plug melted, releasing a plunger which grounded the circuit.

These three elements have formed the basis of telephone protection practically to the present day, even though many improvements and modifications have been made as time passed.[72]

[72] For example, most of these devices are now omitted on completely underground installations since the chances of excessive currents and voltages are essentially nil. In the 1950s, fuses on the customer premises were found to be no longer necessary since currents sufficient to operate them were practically non-existent and fine-gauge conductors served the same purpose as fuses on the few occasions where needed. Still later, as transistorized apparatus was introduced, the protection given by spark gaps was inadequate to prevent breakdown of the low-voltage apparatus used in these circuits and solid-state protectors were used to supplement the gap.



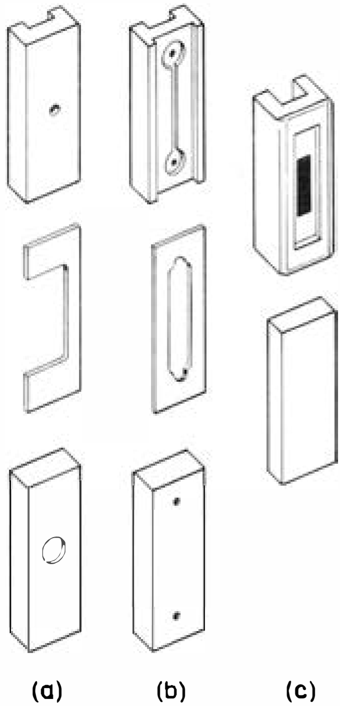


Fig. 4-94. Airgap protector on early telephone (circa 1878).

Fig. 4-95. Steps in the development of the protector block. (a) The formerly standard combination of the No. 1 and No. 2 protector blocks and the No. 3 mica. This was used for central-office and substation protection. (b) The formerly standard copper block protector, consisting of the No. 19 and No. 20 protector blocks and the No. 10 mica. (c) The final step in development, wherein the separator is omitted and proper separation is secured by depressing the carbon insert. (Redrawn from D. T. May 1932, Figs. 1 and 2)

All three of these elements were available in one form or another before the Hayes study, but considerable development effort was required to produce a reliable system.

A major problem with the spark gap (often called a "protector") was to produce a device that would not permanently short-circuit after a few discharges and thus require an expensive visit by a maintenance man to replace the protector at the customer's premises. It was found that carbon was a more suitable material for the gap than metal, since a wider spacing met the same breakdown requirements and the gap was more nearly self-cleaning because the carbon tended to burn away at the sparking point whereas metal could melt and form a permanent bridge. Such a protector made up of two parallel carbon plates with a mica separator was standardized about

1890 and used for both central office and substation protection for many years (Fig. 4-95a). A similar copper block was used where a higher breakdown voltage could be tolerated (Fig. 4-95b), and was sometimes used in parallel with a carbon block to reduce the number of discharges of the latter.

By 1914 the telephone system had expanded to the point where the high maintenance cost of protectors had become a serious concern. A vacuum arrester was tried but was found to be impractical because of breakage and loss of vacuum and the delay in operation as compared to airgap devices.[73] Intensive work was therefore begun on the design of a block which would be inexpensive to manufacture and maintain. Many designs were tried, the final product being the No. 26/27-type block (and other members of the same generic type meeting somewhat different physical and electrical requirements). This protector, shown in Fig. 4-95c, consisted of a hard carbon block (earlier blocks used a soft carbon) and a porcelain block in which was mounted a small carbon block, the latter being ground 0.0028 inch[74] below the procelain. This two-piece construction eliminated the need for the separator, which had presented serious problems in manufacture and handling.

The need for short-circuiting protector blocks when an arc presisted had long been recognized since overheated blocks and mountings would constitute a fire hazard. In the protector formerly standard, with the No. 1 and No. 2 carbon blocks, this need was filled by inserting in one of the blocks a plug of metal with a low melting point which would melt to bridge the gap between the blocks. In the new block it was achieved by using a lead-borate glass to mount the small carbon block in the porcelain block. This material does not cold-flow, but softens at a comparatively low temperature and allows the pressure of the clamping spring in the protector mounting to move the carbon insert into contact with the block connected to ground.

Fuses had been used as early as 1884, when Theodore Vail designed one consisting of a tinfoil conductor, laid in a channel cut in the side of a non-conducting rod. He was granted a patent in the following year. Various modifications of this basic idea were invented by Hibbard, White, and Hayes. A fundamental problem in the design of fuses is that they must open reliably and safely on reasonably low current (usually continuous currents of 5 amperes or more and short-

[73] The search for a practical arrester with hermetically sealed container was periodically revived because of its great potential for maintenance-free operation. By 1964 the 460A gas tube had met the necessary requirements and was being used where the economic climate was favorable.

[74] A tolerance of only ±0.0004 inch is permitted.

duration currents of about 8 amperes), but should not operate on lightning surges. One of the most practical arrangements was to place the fuse wire in an insulating, heat-resistant tube. Usually the tube was filled with asbestos or similar material to help quench the arc resulting when the wire is vaporized. Figure 4-96 illustrates the general configuration and also shows the means for electrical connection and the vents which exhaust the gases generated when the fuse blows.

A primitive sneak-current arrester had been patented by I. H. Farnham in 1887. He used a ball of wax or pitch between a stiff spring, connected to the line, and a ground electrode. This device was associated with the equipment to be protected and when the latter overheated, the wax melted and the spring provided a bypass for the current. A more reliable device was the heat coil, developed by Hammond Hayes and W. L. Richards in the early 1890s. It introduced into the line circuit a small coil of insulated wire wound on a fine copper tube in which a pin was soldered with an alloy that melted at a low temperature. When sufficient sneak current flowed through the coil, the heat which was generated melted the alloy and allowed the pin to be forced into contact with a ground plate (Fig. 4-97).

During the early 1900s, as the power companies increased power-line potential, there was much effort devoted to meeting the needs for high-voltage protection. Considerable improvement was made in fuses and protector blocks (the 26/27-type block was one of the products of this work), but up to 1925 reliable protection was available only against lines having voltages of up to about 3,000 volts to ground (about 5,000 volts between phase wires). Exposure to greater voltages were avoided or, where physical crossovers were unavoidable, the telephone lines were buried or otherwise insulated against contact. After 1925,

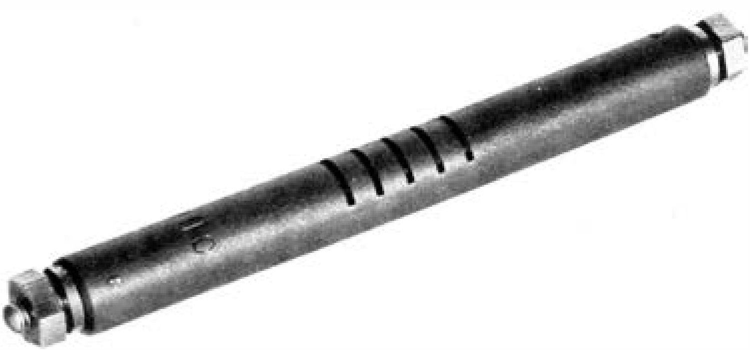


Fig. 4-96. No. 11 tubular fuse. (Savage 1933)

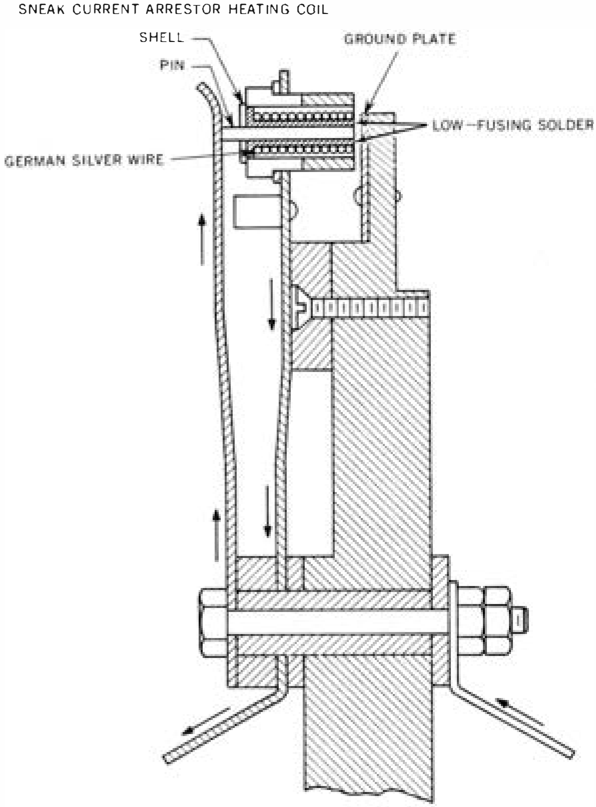


Fig. 4-97. Early form of heat coil and mounting for central-office use (1899).

the joint work (already mentioned) by the Edison Institute and the Bell System provided techniques for joint use of pole lines at somewhat higher voltages.

As cables came into use for entrances to cities, it became necessary to protect them from the high voltages induced on the long open wire lines to which they were connected. Much work was done on this matter, beginning about 1905. The problem was solved by using a number of devices at the junction between the wire and cable, such as protector blocks, fuses, insulating repeating coils, and so on.

5.4 Cost Studies

Once telephony changed from a scientific miracle to an accomplished fact, concern for minimizing cost developed rapidly. It was apparent that Bell's goal of a nationwide communication system could be achieved only if the benefits were worth the charges. In providing local service, much of the cost was in the telephone instruments and means for signaling and switching, while on long distance the transmission facilities accounted for the bulk of the cost and were a major limitation on the growth of this service prior to the advent of loading, repeaters, and carrier techniques. Until these cost-reducing measures were introduced, beginning in the early 1900s, there was little possibility for reduction in toll line costs; but even before this time, techniques were developed for minimizing the overall system cost.

As will be explained more completely in Chapter 6, the early introduction of switching greatly reduced outside plant costs, since instead of the direct interconnection of all telephones, each user was given access (by means of a "loop") to a central switching office at which point he was connected by a "switch" to another loop, in the case of a nearby user, or to a "trunk" connecting to a distant office where he could in turn be connected to the loop desired. This not only minimized the amount of wire required but made it possible to utilize the trunks more efficiently since they could be kept more constantly in use. In the case of long-distance trunks, their use was increased further by the employment of delayed service so that they were less subject to the whims of the user and could be kept in operation almost continuously during the waking hours of the day.

It is apparent that there are basic differences between loops and trunks. The former are short and large in number, whereas the latter are longer and relatively few as compared to loops (about one-tenth in the case of local trunks and far fewer for long distance). As a consequence, it was decided at an early date to use minimum-cost facilities for loops and reserve the high-quality, high-cost facilities for trunks, particularly those going long distances. Thus, loops employed inexpensive insulators and often used galvanized iron wire, since the total reduction in attenuation from better facilities would have been very small for the distances of up to a half-dozen miles that were common. Short trunks used 104-mil copper wire with high-grade insulators and the very heavy 165-mil open wire was reserved for the longest distances. In the cities, where congestion was an important factor, it was practical to use cable for short distances, the finer gauges being used for loops and the heavier for trunks.

It is not clear just when it was realized that these general principles of selecting facilities could be developed into a systematic technique for design of minimum-cost plant, but it was certainly before 1906, when loop and trunk studies, aimed at achieving the best economic balance between these facilities, became part of the responsibility of the AT&TCo Construction Department headed by F. L. Rhodes. It was also about this time that organizations were established for "fundamental planning," i.e., planning the growth of plant on the most economical basis to meet traffic needs in both the near and the more distant future. By 1925, the need for fundamental planning of central-office locations and subway, cable, and wire routes and economic assignment of type and size of facility had been long accepted and was an important part of the Engineering Department's responsiblity.

The advent of loading, phantoms, and repeaters stimulated the development of other types of cost studies aimed at determining the most economic way to use these facilities for reducing wire size. Some of these studies were rather straightforward since it was relatively easy to balance the cost of loading coils or repeaters against the cost of copper. However, developing technology presented the designer with many choices affecting economy. For example, the resistance and core losses of loading coils reduced their effectiveness and hence required some offsetting increase in the wire size. For a price, these losses could be reduced, and it became important to determine how much should be spent on loading coils in order to save line wire. Similar questions arose endlessly in connection with the design of phantom coils, composite sets, repeater balancing networks, and so forth. Many of these problems justified specific, detailed cost studies, but others required a more general approach and for these an interesting concept was developed known as the W.A.C. (Warranted Annual Charge) of transmission. This was the annual cost of providing a transmission improvement of 1 mile of standard cable (or dB) in the most economic manner. Originally, this concept seems to have begun at a time when the use of additional copper was the only way to improve transmission, and it provided a useful means for determining the justifiable cost for improving repeating coils, central-office equipment, and so forth. For example, if the W.A.C. of transmission (based on copper) was at the rate of $50 per dB, then improvement in a repeating coil was justified up to the point where a small improvement just equaled this rate, at which point it would be in economic balance with wire plant (i.e., each 1/10 dB improvement in a coil was justified if it could be obtained for $5 a year or less; at the $5 rate the coil design was in balance with the remainder of the plant).

Later the W.A.C. of transmission was derived in the course of loop and trunk studies for a plant with loops and trunks in economic balance. The process is too complicated to be described here. It is only necessary to note that by 1925 the process was highly sophisticated and formed the basis for designing a very complex telephone transmission plant made up of many elements each of which was designed to minimize the total plant cost.

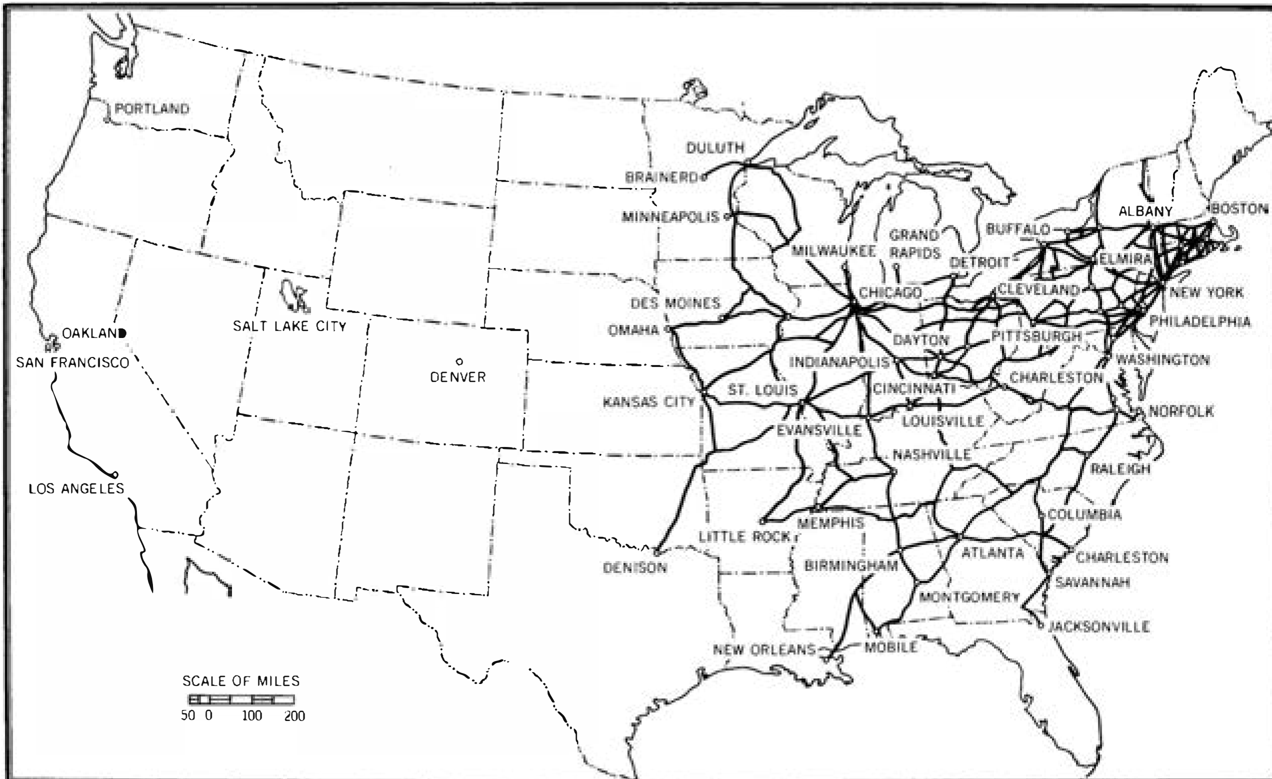


Fig. 4-98. Long-distance lines of the American Telephone and Telegraph Company in 1906. (Redrawn from Shaw 1944)

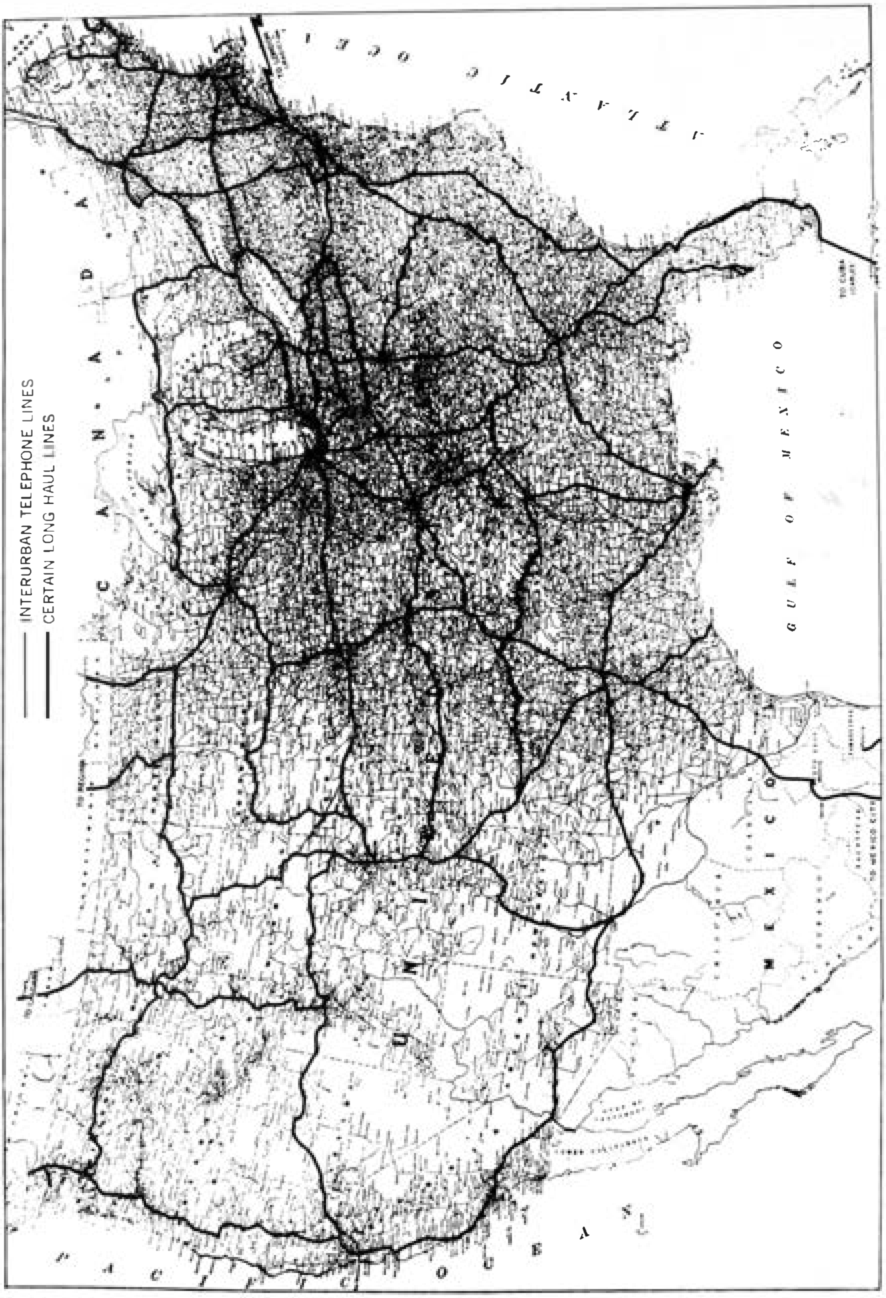


Fig. 4-99. Toll lines of the Bell Telephone System in 1929. (Gherardi and Jewett 1930, p. 100)

VI. SUMMARY

The simplest way to summarize the wire transmission situation in the middle 1920s is to say that it had reached technical maturity.

The first 25 years of telephony was largely the period of empiricism. Theoretical studies of wave propagation had started but their use by the practitioner was not great before the end of this period. However, in this new art of telephone transmission the empirical approach was highly productive and resulted in basic wire media, both open wire and cable, which were to continue in use for many years.

The second 25 years of telephony were characterized by the growing knowledge and application of wire transmission theory. Analytic methods were increasingly applied not only to the solution of wire transmission problems but also were adapted for use in station apparatus design. The value of the application of transmission theory was manifested in many ways as indicated below.

Growth of long-haul transmission facilities was extremely rapid. In 1900, toll lines totaled 600,000 miles of wire or approximately 300,000 miles of circuits. By the end of 1926, toll circuit miles had increased by a factor of ten. Even as late as 1906, the long-distance network was confined to the eastern part of the United States (Fig. 4-98). Twenty years later all parts of the country had been opened to telephony (Fig. 4-99).

The use of the newer types of transmission media were greatly stimulated. In 1900, almost all toll circuits were open wire, but by 1926 about 45 percent of the mileage was in cable. Far more significant than indicated by the numbers, about 2 percent was in carrier.[75]

Many of the problems that restricted telephony by wire at the turn of the century had been solved 25 years later. Distance had been conquered by the use of loading and amplification. Major steps had been taken toward the solution of the cost problem. Most significant in this connection, the use of carrier multiplex had started and groundwork had been laid for the vast expansion and cost-reduction potential of this technique. Also in this period a new form of transmission began which freed telephony from the constraint of wire connection. This wireless or radio transmission, which will be covered in Chapter 5, not only opened new fields for telephone communication but also opened new fields of technology. In doing so it borrowed heavily on the technology of wire transmission but it also contributed much in return. It led the way to the use of the higher frequencies and originated many of the techniques and devices used in the carrier transmission over wires. These two developments, carried on in parallel, provided an

[75] Twenty-five years later, carrier provided over 65 percent of the Bell System toll circuit mileage and by 1976, the year of the Telephone Centennial, voice-frequency circuits will have practically disappeared from toll use, amounting to only about 2 percent of the total.

interesting example of the way in which technology can be used to produce an integrated system employing two media differing quite basically in their characteristics. The pattern set at this time continued during the succeeding years as each technology evolved, and by the late 1960s resulted in a long-distance plant having roughly equal mileage served by the wire and wireless (radio) media.

In brief, by 1925, solid groundwork had been laid for developing the greatly improved transmission and the enormous economies of scale which would take place in the following years.

Chapter 5

Telephone Transmission—The Advent Of Radio

This chapter reviews the early development of telephone communication without wires and shows how, like wire telephony, it advanced to a high degree of technical maturity by the late twenties. Wireless transmission, beginning about ten years after Bell's telephone invention, evolved quite independently during its first two decades with major emphasis on the telegraph application. De Forest's invention of the audion brought together workers in the wire telephone and wireless fields; and thereafter, despite major differences in the media, wire and wireless transmission techniques were closely linked in their evolution. Engineers in the two fields were able to use many of the same devices, such as oscillators, amplifiers, modulators, and wave filters, and to exploit much of the same new technology, such as the carrier multiplex and single-sideband principles. As a result the wireless medium, soon referred to as radio, became usable for voice communication and came into service as a highly useful supplement to wire telephony.

I. HISTORICAL BACKGROUND

As with many great inventions, one cannot assign credit to a single individual for the invention of wireless communication. Much is owed to contributions by many workers over a long period of time, but of all of the early contributors the German physicist Heinrich Hertz is probably most deserving. Hertz, working between 1884 and 1888, not only demonstrated that high-frequency oscillations could produce electrical phenomena at a distance but also recognized that the effect was due to electromagnetic waves conforming to the laws of geometric optics and traveling at the speed of light. Hertz thus experimentally substantiated the predictions of James Clerk Maxwell, whose electro-magnetic-wave equations had been published 20 years earlier.

Hertz also invented the method that was used for many years for generating and launching these waves, namely, the discharge of a capacitor through a spark gap connected to short lengths of heavy wire. (Today, we would call it a dipole antenna.) This generator of waves was somewhat more advanced in concept than his detector (see Fig. 5-1), which consisted of a large, nearly closed loop of wire with the ends brought close together so as to form a very closely spaced spark gap which was observed visually. This means of detection was highly insensitive, requiring even for the smallest gap a potential of about 300 volts to produce a visible spark. His transmissions were thus limited to very short distances and it is remarkable that he could achieve so much with such primitive equipment. Perhaps for this reason Hertz did not pursue his discoveries to the point of applying them to communication, but it is more likely that, having investigated the intellectual problem to his satisfaction, he was anxious to move on to other fields.[1]

Probably the earliest recognition of the possible use of Hertzian waves for communication came from William Crookes who wrote[2] in 1892:

> Here is unfolded to us a new and astonishing world, one which is hard to conceive should contain no possibilities of transmitting and receiving intelligence.
>
> Rays of light will not pierce through a wall, nor as we know only too well, through a London fog. But the electrical vibrations of a yard or more in wavelength . . . will easily pierce such mediums, which to them will be transparent. Here, then, is revealed the bewildering possibility of telegraphy without wires, posts, cables or any of our present costly appliances.

He then proceeded to enumerate some of the developments required to

[1] An 1889 letter of Hertz, replying to an inquiry concerning the practical value of his discoveries for telegraphy, is often quoted as evidence that Hertz missed the significance of his work. What he said, in essence, was that it would be possible but impractical to radiate waves at the extremely low frequencies of telegraphy. While this showed a profound understanding of wave propagation, it is an interesting example of how technical comprehension of the difficulties sometimes obscures a simple solution, in this case the use of high frequencies interrupted (modulated) to form telegraph signals.
[2] In the non-technical *Fortnightly Review*.

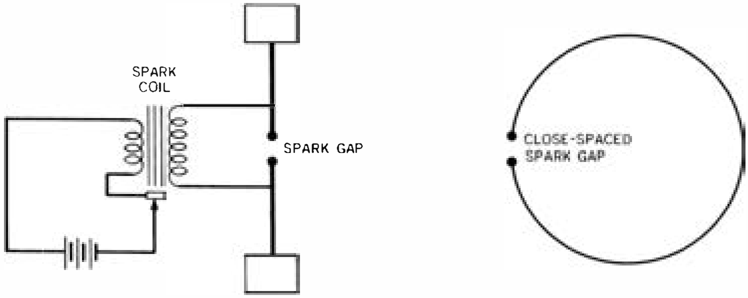


Fig. 5-1. Wireless transmitter (left) and receiver (right) of Heinrich Hertz.

bring about practical communication, namely, reliable transmitters for various wavelengths, sensitive tunable receivers, and directional antennas.

Several experimenters, their interest aroused by the work of Hertz, soon produced a more sensitive detector than Hertz's spark gap by using a phenomenon observed many years before. As early as 1835, it had been noticed that the conductivity of a mixture of metal filings, normally poor, was increased when subjected to the discharge of a Leyden jar and remained so until physically disturbed. A number of people explored this effect but the most significant work was done by Professor E. Branly in Paris in 1890–91, using an insulating tube containing metal filings between two conducting plugs, an arrangement later to be known as a coherer. Branly showed that an electric spark, *at a distance*, could change the resistance of this device.[3] He seems not to have appreciated that the cohering of the filings was caused by Hertzian waves, but within a year or two a number of people, including G. M. Minchin and Oliver J. Lodge in England, recognized that Branly's device was responding to waves created by the spark discharge, and Lodge then proceeded to apply the coherer principle in several demonstrations made in 1894. (He also contributed the name.)

In these demonstrations, Lodge used an oscillatory spark generator and several types of coherers. The lowered resistance was demonstrated by the deflection of a galvanometer in series with the coherer and a battery. In one demonstration he substituted an electric bell for the galvanometer and set it to ringing by an electric discharge. He also showed that the vibration of this bell, if mounted on the same base as the coherer, was sufficient to "de-cohere" the device and make it ready for the next signal. Thus Lodge approached very closely to the signaling arrangement illustrated in a general way by Fig. 5-2, which was used for a number of years in early wireless telegraphy. The use of Hertzian waves for telegraphy seems obvious but, to quote Lodge, he "did not realize that there would be any particular practical advantage in thus with difficulty telegraphing across space instead of with ease by the highly developed and simple telegraphic and telephonic methods rendered possible by the use of a connecting wire."[4] While Lodge missed some of the significance of his 1894 work, he does deserve credit for appreciating and demonstrating the importance of tuning the transmitter and receiver to the same frequency. Hertz had employed a similar principle but his circuits were highly damped and

[3] As noted in Section 2.1.3 of Chapter 3, Hertzian waves, or other electrical impulses, could also cause carbon granules to cohere and this electrical "packing" of the telephone transmitter was one of the problems faced by the designers of these instruments. Fortunately, carbon required a much higher voltage than metal filings before cohering occurred.

[4] This rather involved statement is from the third (1900) edition of *Signalling Through Space Without Wires* by Lodge.

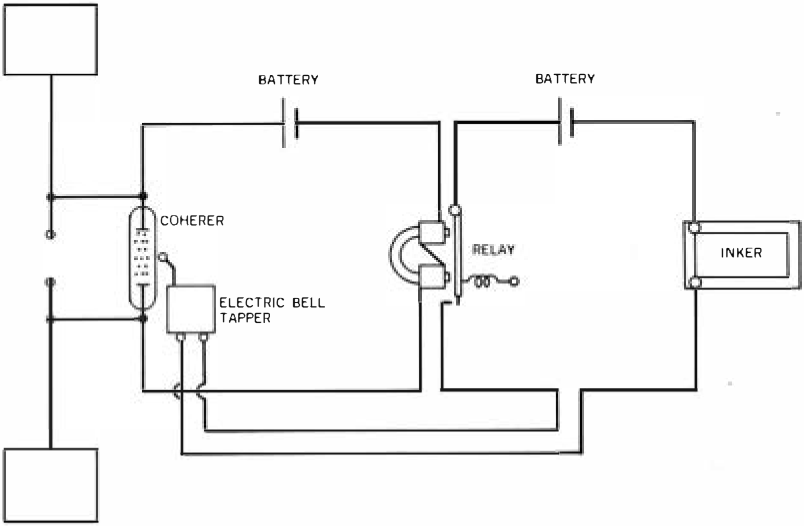


Fig. 5-2. Generalized schematic of early wireless receiver.

did not emphasize the significance to the extent of the work by Lodge and the latter was ultimately awarded basic tuning patents.[5]

After seeing an account of Lodge's lectures, Professor Alexander Popov, in Russia, conducted experiments with the coherer in 1895–96. His setup was almost identical to that shown in Fig. 5-2. He used the bell tapper in contact with the coherer tube to de-cohere the filings and also provided means for protecting the coherer from local sparks. He also departed from Lodge's example by using a large vertical antenna. Popov seems to have been aware that this equipment had a potential for signal transmission "as soon as a source of such oscillation possessing sufficient energy will be discovered" but his main interest presumably was in the study of atmospheric electric discharges.

It will be noted that most of the early work on Hertzian waves had been done by scientists more interested in exploring basic principles than in commercial application.[6] What was needed was a technically-minded entrepreneur who could put these basic discoveries to work. The man to do this was Guglielmo Marconi of Bologna, Italy.

[5] Marconi also appreciated the significance of tuning and protested the Lodge patent. When this was upheld, Marconi quickly purchased the right to use.

[6] An interesting contrast between wire telephony and wireless can be mentioned at this point. By 1895, wire telephony had been developed very extensively on an empirical basis, with theoretical understanding of transmission still limited, as we have noted in Chapter 4. In wireless, at this time, the main interest was theoretical with little interest in its application.

Marconi had an absorbing interest in physics and chemistry but, in contrast with his predecessors in the wireless field, was not a highly trained scientist. To quote his later collaborator, Fleming, "Marconi was eminently utilitarian. His predominant interest was not in purely scientific knowledge per se, but in its practical application for useful purposes."

In 1894, at the age of 20, Marconi had heard of the work of Hertz and Lodge, and after that time he was constantly concerned with wireless experimentation and the practical development of the medium. His early work was very similar to that of Popov but done independently.[7] He improved the coherer, developed the method of application illustrated in Fig. 5-2, and adopted an elevated antenna instead of Hertz's dipole. By 1896 he was transmitting Morse code messages nearly 2 miles. Moving to England, he demonstrated that messages could be transmitted as much as 8 miles, and by 1897 had formed the British Marconi Company, the first in the wireless field. Immediately afterward he began experimenting with long-distance transmission, but progress was slow and it was not until December 12, 1901, that he was able, by using a kite-supported antenna, to hear faintly the Morse letter "S" transmitted over the 1,700 miles from Cornwall to Newfoundland.

Promising as this was, and despite the opening of commercial service in 1907, Marconi realized that the existence of undersea telegraph cables, together with the requirements for high power and costly antennas for bridging the ocean, would severely limit the commercial application of transatlantic wireless telegraphy. Accordingly, he entered the field of ship communication, an area where there was no practical alternative to wireless. Interest in this field was grealy stimulated in 1909 when wireless messages from S.S *Republic* were responsible for prompt rescue action when that ship was hit and sunk at sea, and as early as 1910 laws were passed in the United States requiring wireless equipment and a skilled operator on vessels carrying 50 or more persons between ports 200 miles or more apart.[8]

During the early 1900s, though there were many competitors, Marconi's company occupied a prominent position in marine telegraphy, with the right to use important basic patents on vertical antennas, methods of selective tuning, the coherer, the magnetic detector, and the Fleming valve. While most of the competitors proved less astute businessmen than Marconi, two basic contributions were to come from

[7] The relative merits of the Marconi and Popov claims to the invention of radiotelegraphy have been argued at length. It seems likely that Marconi has priority both on publication and demonstration. Perhaps of more importance is the fact that Marconi foresaw the potentialities of Hertzian waves as soon as he heard about them and never ceased in his efforts to utilize them for telegraphy.

[8] The *Titanic* disaster of 1912 reemphasized the importance of wireless at sea and pointed out the need for 24-hour coverage and for emergency power. The United States Radio Act was subsequently amended to include these important provisions.

these sources that profoundly affected the course of radio development. One was the use of continuous waves instead of the intermittent and non-coherent oscillations of the spark discharge; the other was the de Forest audion. Fessenden was probably the most active advocate of the continuous-wave technique in the United States, but it was the audion as a "CW" generator that ultimately established the continuous coherent type of oscillation as the base for efficient radio communication.

Early in Chapter 4 (Sections 2.1 and 2.2) some of the fundamentals of voice-current propagation over wires were discussed as a basis for the more detailed treatment of transmission development which followed. Similarly, having now sketched the first primitive steps in establishing a new transmission medium not requiring wires, we shall examine more closely those technical characteristics of the radio medium which were to determine the special roles it would play in domestic as well as overseas communication.

II. TECHNICAL BACKGROUND

An important requisite to progress in radio was the development of good methods for generating and detecting electromagnetic waves. Once this was done, its field of application was greatly influenced by the manner in which these waves were propagated and the history of radio is largely the story of how technology has been developed to take optimum advantage of the characteristics of the radio medium.

2.1 Generation and Detection of Radio Waves

The equipment most commonly used for wireless telegraphy in the years before the vacuum tube was basically very simple (Fig. 5-3). The transmitter was fundamentally similar to that used by Hertz, consisting of a high-voltage source[9] which charged a capacitor. When the voltage reached the breakdown point of the spark gap across the capacitor, the latter discharged through the gap and set up a fast-decaying or "damped" oscillatory current in the antenna circuit. The electric wave thus generated was highly irregular in magnitude, the high degree of damping or "decrement" leading to a frequency spectrum rather broadly dispersed on each side of the nominal frequency determined by the inductance and capacitance of the oscillatory circuit (including the antenna). In practice it was the type and size of the

[9] The Ruhmkorff induction (or "spark") coil, using an interrupted direct current as primary power, was used for many years, but alternating-current sources feeding step-up transformers were more common for high-power commercial installations. Also, while the simple open spark gap of Hertz was often used, various other forms, including rotary gaps, quenched gaps, etc., were used before the more sophisticated continuous-wave system came into use.

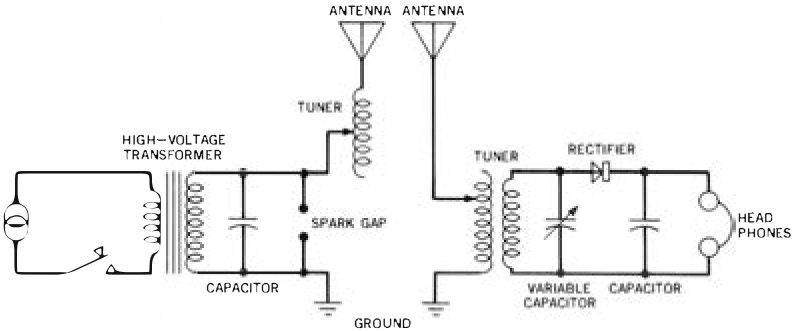


Fig. 5-3. Wireless transmitter (left) and receiver with diode detector (right) of the early twentieth century.

antenna and the adjustment of the inductance in the antenna circuit which established this nominal or mean frequency.

Hertz in his experiments deliberately made the frequency of oscillation high (as much as 80 to 100 MHz) so that he would have waves short enough to follow the laws of optics, a major objective of his research. He did this by using small antennas, essentially simple dipoles. Commercial wireless installations, on the other hand, grounded one side of the gap and connected the other through a tuning inductance to a long and high antenna system. This gave them very low nominal frequencies, hence waves of greater length, which they found advantageous for long-distance communication because of the tendency of longer waves to bend around the earth.

Commercial receivers employed a similar antenna system together with coupled air-core tuning coils of variable inductance which allowed the user to select the frequency to be received. The detector in the early days was a coherer but it was far from satisfactory. Not only was it unstable and insensitive but its response to bursts of static was not readily distinguished from bona fide signals.[10]

About 1904, Professor J. A. Fleming of University College, London, working as a consultant for Marconi, devised the two-element vacuum tube which came to be called the Fleming valve. This device, consisting of a hot cathode and a nearby cold anode in an evacuated glass bulb,

[10] The magnetic detector was another interesting device for producing audible signals. This device, which had some similarity to a tape recorder, had a long history of indifferent success, but Marconi reduced it to practice and used it extensively for about ten years. It consisted of a continuous belt of soft-iron wires traveling past permanent magnets and along the axis of two solenoidal coils. The magnets placed a magnetic bias on the belt which was disturbed by radio signals passing through one of the coils. This belt, with the magnetic domain reordered by the radio signals, generated an emf in the second coil which produced audible signals in a telephone receiver connected to it.

conducted in only one direction and thereby acted as a rectifier of the received oscillations so that the envelope of the wave could be heard in a head receiver.[11] Shortly thereafter the rectifying properties of certain crystals, such as carborundum, galena (lead sulphide), and silicon, were put to use as detectors that were so simple and efficient that the Fleming valve received little practical use. Listening to the rectified signal in a head receiver proved to be an important step forward since a skilled operator could readily discriminate between the wanted signal and static or other interfering sounds.

De Forest devised the three-element audion (Chapter 4, Section 4.2.2) in an attempt to produce a detector with the stability of the Fleming valve but with sensitivity comparable to or better than the somewhat unstable crystals. In this he was so successful that for a number of years he concentrated on this application of the audion so that it remained for others to discover many of its even more important applications.

The great simplicity of the spark system attracted many amateur experimenters. Almost anyone with a little manual skill and a few dollars could build a wireless transmitter and receiver. All that was required was a spark coil (a common part of the Ford automobile ignition system), a telephone receiver, a piece of galena, and a miscellaneous collection of tinfoil, wax paper, glass plates, copper wire, and cereal boxes for the construction of capacitors and tuning coils. As a consequence, a large group of amateurs provided a readymade audience after World War I when radio broadcasting made its first faltering transmissions.

The simplicity and ease of obtaining fairly powerful oscillations made the spark system very attractive but it had some basic disadvantages. The waves set up by the spark discharge were a series of highly damped irregular oscillations, each of which began with a high amplitude and rapidly decayed as the capacitor discharged. (See Fig. 5-4a.)[12]

Such waves had two basic faults. One, they used the oscillatory energy very inefficiently since the average power, which determined communication range, was very much less than the peak power, which was limited by the antenna and associated apparatus. Second, the energy was not concentrated in a single frequency but was spread out on either side of the nominal frequency determined by the transmitter constants. As a result, sharp tuning was not possible and there was much interference between stations unless they were widely spaced in frequency. More importantly, it was not possible to modulate these waves effectively

[11] Wehnelt, in Germany, had produced a thermionic rectifier in 1903 but his interest at the time was in converting ac power into dc power and it was not until several years later that he pointed out that the principle could also be used for the high frequencies of radio.

[12] It was the rectified envelope of this irregular wave train that was heard in the headphone of the receiver using diode detection.

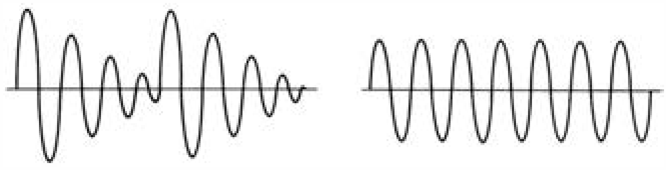


Fig. 5-4. (a) Damped-wave oscillations. (b) Continuous-wave oscillations.

with speech since demodulation would reproduce not only the speech but all the irregular sounds in the original damped oscillation.

A number of people[13] realized that both of these disadvantages could be overcome by the use of continuous waves, i.e., waves generated by pure tones (Fig. 5-4b). Such waves used the energy efficiently (concentrated it at one frequency and with a high ratio of average to peak power) and could also be modulated and demodulated for the transmission of speech as outlined in Chapter 4 (Section 4.2.5). They had a minor disadvantage for telegraphy since the output of diode detectors was above audible frequency, but this could be corrected by chopping (interrupting) the signals or by other means. A major problem was the generation of these continuous waves at frequencies high enough for radio application and with enough power to be useful. Prior to the vacuum tube there were two main means: one, promoted by Tesla, Fessenden, and Alexanderson, was the rotary alternator, and the other, invented by Poulsen, was the electric arc. Either of these devices could produce waves suitable for telegraphy but the problem of modulating them efficiently kept them from successful commercial application for radio telephony. The three-element vacuum tube ultimately provided the solution for both generating and modulating continuous waves and completely changed the wireless art as developed by Marconi. But more of this later.

2.2 Propagation of Radio Waves

Hertz had demonstrated that radio waves followed the laws of geometric optics. For the very high frequencies he used, they followed straight lines, could be reflected from surfaces, and with suitable reflectors could be concentrated in beams much as searchlights concentrate light. Ultimately (after World War II) the major development of two-way radio communication was to occur at and above the very high-frequency range in which Hertz experimented. But for many

[13] In the United States, John Stone Stone (of the Boston laboratory) and Reginald Fessenden were early proponents of continuous waves.

years commercial communication was practical only at lower frequencies. We can now see rather clearly why this occurred, but at the time the choice of frequency involved much experimental effort devoted to determining propagation characteristics. Rather than guide the reader through the devious path followed by the radio practitioner, we shall briefly review some of the salient points in the light of current knowledge.

It is convenient to consider radio propagation in three categories, determined by frequency,[14] having distinctly different characteristics. The dividing line between these categories is not sharp. Instead, between them are broad transitional regions showing some of the characteristics of the area on either side.

Above about 1,000 MHz the waves behave much like light through a clear, reasonably uniform atmosphere. When originated from a point source, they propagate in all directions and, since the area of the wave-front spreads out spherically, the energy in a given area decreases roughly as the square of the distance from the source. Variations around this

[14] The literature describing the radio spectrum is about equally divided between the use of wavelengths and frequencies. We shall endeavor to use the frequency description as much as possible since it will be more consistent with our telephone material. For those who are interested, the relation between wavelength and frequency is shown in Fig. 5-5, which also lists a series of nominal descriptions for the spectral regions now rather widely used. In the 1920s and earlier the designations used did not conform to this classification.

Wavelength vs Frequency

$$\lambda \text{ (wavelength in meters)} = \frac{300{,}000{,}000}{f \text{ (frequency in hertz)}}$$

Frequency Classification		
Nominal Classification	Frequency	Wavelength
VLF—Very Low Frequency	3—30 kHz	100,000—10,000 meters
LF—Low Frequency	30—300 kHz	10,000—1,000 meters
MF—Medium Frequency	300—3,000 kHz	1,000—100 meters
HF—High Frequency	3—30 MHz	100—10 meters
VHF—Very High Frequency	30—300 MHz	10—1 meter
UHF—Ultrahigh Frequency	300—3,000 MHz	1—0.1 meter
SHF—Superhigh Frequency	3—30 GHz	10—1 centimeter
EHF—Extremely High Frequency	30—300 GHz	10—1 millimeter

Fig. 5-5. Frequency classification.

inverse-square rule can result from interaction between the wave in free space and the "echo" reflected from the earth, but unlike guided waves there is essentially no energy absorbed in the medium and there is little or no attenuation in the same sense as the power loss of a transmission line.[15] The dispersion, or spreading out of radio waves, has been an important factor in radio application from the beginning. It means, of necessity, that the amount of energy available for the receiver decreases rapidly as the distance from the transmitter increases (at least a 6-dB decrease each time the distance is doubled) and explains why rather sizeable amounts of power are required to cover even moderate distances.[16] The major problem in using these frequencies is that, because they propagate like light, their transmission tends to be limited by the horizon. Thus, until it became possible many years later to transmit hundreds of telephone channels on a beam, there was little interest in radio transmission over 20- to 30-mile, horizon-limited distances since, for this range, wire transmission was simple, inexpensive, and avoided the interference problem of radio.

At the other end of the spectrum, below about 500 kHz, radio waves tend to follow the curvature of the earth, being guided between the earth and the ionized layer of the upper atmosphere (ionosphere). Early experimenters soon discovered this advantage of the lower frequencies and exploited it for overseas and ship-to-shore transmission. In this low-frequency range there is some dissipation of energy, or attenuation, in both the earth and the ionosphere. But in the very low part of this band (below 50 to 100 kHz) the added loss is small and not important compared to the elimination of the distance barrier.[17]

The use of low frequencies was not an unmixed blessing. Since the waves tended to spread out in all directions, a transmitter capable of covering the transatlantic distance of 3,000 miles interfered with reception on that frequency at all receivers within a 3,000-mile radius (i.e., within an area of 30 million square miles). Thus each transmitter tended to preempt both physical space and frequency space.[18] Obviously, the number of stations that can operate 24 hours a day throughout the world is very limited at the low frequencies favorable

[15] Exceptions occur above 10,000 MHz but these were not to prove important for another 50 years.

[16] Concentrating the radiation in a beam reduces the power required to reach a given distance but does not modify the inverse-square law since the beam constantly spreads out as distance is increased.

[17] The empirical propagation formula proposed in 1911 by L. W. Austin and L. Cohen of the National Bureau of Standards provided a basis for estimating propagation loss and showed clearly the advantage of the lower frequencies.

[18] The amount of frequency space required for communication depends on the state of the art and the kind of signal transmitted. For telephony it cannot be less than 3 kHz and with allowance for practical matters a figure of 5 kHz is about the minimum for very low-frequency transmission and becomes more like 25 to 50 kHz at high frequencies, when frequency-modulated.

for overseas transmission. The concentration of the energy in beams reduces the area of interference and a modest amount of directivity has been achieved at low frequencies, but high directivity requires antenna structures many times the transmitted wavelength. Such antennas are almost prohibitively large at frequencies around 50 to 100 kHz which correspond to waves several miles long. This frequency-space limitation has been a basic problem in radio from the beginning and explains the emphasis, over the years, that has been placed on developing the economic use of higher and higher frequencies. Such frequencies provide the frequency space for more stations and also tend to favor the use of directivity and thus reduce area interference.

In the 1920s, about the time when the limitations on the use of the lower frequencies were beginning to be felt, it was discovered that the high-frequency region between about 3 and 30 MHz was also usable for long-distance communication. These frequencies were too high for effective transmission along the earth's surface to distances much beyond line of sight but they had the peculiar property of penetrating the lower boundary of the ionosphere and being reflected (sometimes several times) within upper layers of the ionosphere, ultimately being returned to earth. Thus, these frequencies were usable as "surface waves" close to the transmitter and again as so-called "sky waves" at long distances, with an unusable skip-distance between. This discovery opened an enormous frequency space for long-distance transmission (some 25 million hertz as compared to the 100 or so thousand hertz previously available). Again, this was not an unmixed blessing. Propagation varied greatly with the season of the year and time of day, and over the seven-year cycle of sun-spot activity. For a given route of a few thousand miles, a specific frequency provides good transmission for only part of a day, and other frequencies must be used for the remainder of the time. Thus, each 24-hour communication circuit may require the assignment of three or four frequencies. Even so, the gain in available communication circuits resulting from opening up the high-frequency band was enormous compared to the band below 100 or 200 kHz.

Propagation in the three frequency bands just discussed can be described in rather simple terms as surface waves, sky waves, and "optical" or line-of-sight waves. In the transition bands between, the propagation characteristics are more complex. The region between 500 kHz and 3 MHz is one of these transition bands. At these frequencies the waves are not greatly limited by the horizon and they propagate along the earth's surface during the daytime to distances several times the line-of-sight distance. At night, much greater distances are achieved by reflection from the ionosphere, the maximum ratio of night-to-day transmission occurring in the 1- to 2-MHz band. The lower part of this band was initially used extensively for ship-to-shore communication but later was assigned to AM radio broadcasting.

The frequencies between 30 and 1,000 MHz constitute the other transition region. In this band, ionospheric transmission does not apply, the surface wave is weak, and the wave reflected from the earth tends to cancel the direct wave. The horizon effect is noticeable but not serious at 30 MHz but it may be controlling at 1,000 MHz. This frequency range is used primarily for mobile radio communication and television broadcasting. As the frequency increases above 1,000 MHz (microwave range), the horizon effect becomes more and more controlling and the shorter wavelengths make it easier to provide both a sufficient antenna height (as measured in wavelengths) and a significant amount of antenna directivity for point-to-point free-space transmission with optical-like waves.

As we have already implied, radio waves do not always follow a single path. With the exception of the very longest waves, there are almost always multiple paths of transmission, some of which involve reflections from the ground or from irregularities in either the atmosphere or the ionosphere. Since these paths are of different lengths, the waves can arrive in opposition to those following the direct path or in phase with the direct wave, giving rise either to cancellation or reinforcement. Since the various paths are constantly changing, the strength of the composite arriving signal is subject to constant variation. This so-called "fading" is characteristic of all long-distance transmission at frequencies of several MHz or more in the daytime and of even lower frequencies at night. At microwave frequencies, rapid fading of large magnitude may be experienced over short, line-of-sight distances as the result of interference between the direct wave and reflections from both the ground and atmospheric irregularities.

Before we complete this brief technical review, we should mention the problem of noise interference. Every source of electric discharge is a potential source of radio noise. Probably the most important, since not under man's control, are the electrical disturbances from lightning discharges which reverberate around the world between earth and ionosphere. The noise, called static, increases in magnitude at the lower frequencies and proved to be an important limitation in the development of early radio communication. Static from distant sources is unimportant at frequencies above the sky-wave band at which point man-made sources and local thunderstorms become controlling.[19]

[19] In 1927 J. B. Johnson of Bell Laboratories showed that the agitation of the electrons in any conductor was a source of noise which was proportional to the temperature of the conductor. This type of noise became controlling at the very high frequencies. Later, Karl Jansky discovered that a similar type of noise came from distant stars and nebulae and this discovery originated radio astronomy. Still later, in the satellite era, the noise from our own sun became a factor of importance in limiting the pointing of the very narrow beam receiving antennas used with this type of transmission.

III. THE BEGINNINGS OF RADIO IN THE BELL SYSTEM

3.1 Early Attempts at Telephony

Even before the work of Hertz, Bell had transmitted speech over a beam of light and was granted a patent in 1880 for a device which he called the "photophone." An improvement was worked out in the Boston laboratory of the American Bell Telephone Company in 1897 on which a fundamental patent was issued,[20] but the work was not pursued beyond the experimental stage. It is interesting to note that E. J. P. Mercadier re-christened the device "radiophone," the first use of the word "radio" in the sense we use it today. Optical systems, with their weather dependence, could not compete with the much lower-frequency Hertzian waves which were essentially independent of weather. In 1892, following Crookes' comments on the potentialities of Hertzian waves, Hammond V. Hayes, head of the Boston laboratory, suggested to John Stone Stone that Hertzian radiations might be used to signal vessels at sea. Stone investigated the possibility but was unsuccessful; the required instrumentalities did not exist at this early date (as Crookes had noted). In 1902, Hayes, still interested in wireless, engaged G. W. Pickard to investigate the possibility of radiotelephony. During the following six months, while working in the Boston laboratory of what was now AT&TCo, Pickard succeeded in achieving a degree of voice communication by using a transmitter diaphragm to mechanically vary the size of a spark gap. This work was not continued, in part because it became clear that the road to success was long and uncertain and in part because of an apparently fundamental patent granted to Reginald Fessenden about this time.

As we have noted, a basic requirement for telephony was a source of continuous waves. This seems to have been recognized by Stone as early as 1892. In reporting on his work in July of that year he stated, "The high frequency transmission has reached a state where a high frequency alternator is essential for satisfactory work." Both the alternator and the oscillating arc provide sources of continuous waves and were used in the 1906–1912 period for experimental radiotelephone transmission by a number of people outside the Bell System, including Poulsen, Fessenden, de Forest, and Vanni.[21] In early 1907, Hammond Hayes, on the basis of a report by E. H. Colpitts, stated to President Fish of the Bell System:

> I feel that there is such a reasonable probability of wireless telegraphy and telephony being of commercial value to our company that I would

[20] Lloyd Espenschied, in an unpublished memorandum, has pointed out that the first claim of this patent covers, for the first time, the modulation of radiant energy by a voice current and, in a sense, is the basis for all amplitude-modulated radio and carrier telephone systems.

[21] In 1912, Vanni succeeded in telephoning a distance of about 600 miles.

advise taking steps to associate ourselves with Mr. Fessenden if some satisfactory arrangement can be made.

The Bell System was close to completing negotiations with Fessenden when the financial panic of 1907 hit the country. As a result, there was a change in Bell System management, headquarters were shifted to New York, and all expenditures were examined very critically. After a reappraisal of the Fessenden patents, it was decided not to buy. Thomas Lockwood, the Bell patent counsel, summarized his conclusions as follows:

> If an individual or a corporation is desirous of becoming interested in wireless telephony for the sake of keeping in touch with progress in electrical transmission work, based on recent scientific research, this would seem to be an excellent opportunity; but for a telephone company, the possibility of substituting a wireless system for a system of toll lines is the most attractive feature of the proposition, and I have a strong conviction that this feature cannot and will not reach any practical realization within the term of years yet remaining to Fessenden's fundamental patents.

The fact was that prior to the audion invention in 1907 the devices needed to make radiotelephony commercially successful were not clearly in view. In addition to a simple, stable source of high-frequency power (which conceivably could have been supplied by alternators or arcs), three additional elements were needed: first, means for controlling these currents in accordance with the weak speech currents; second, means to receive and amplify the waves greatly weakened in transit; and, third, a means for deriving the original speech signal from the received wave which was better than the crystal detector then in use. These elements were those also required for carrier telephony on wires and the problems were to be solved in the Bell System more or less simultaneously. The key elements, in each case, proved to be the vacuum tube and the wave filter.

To appreciate the significance of the vacuum tube in radiotelephony, we should note that long-distance radio requires enormous power as compared to wire transmission, hundreds of kilowatts as compared to a few milliwatts. A major problem in radiotelephony was, therefore, that of efficiently modulating these enormous powers with the minute telephone currents available. No very satisfactory way to do this had been devised for arc or alternator sources of continuous waves. But with the vacuum tube it proved possible to generate the carrier and modulate it at lower power levels, not greatly different from those used for wire communication. The modulated waves could then be amplified to whatever power level was required to achieve the distance objective.[22]

[22] A simpler but less satisfactory approach, used in the early days of radiotelephony, was to modulate the signal at high level in the output amplifier using a voice signal amplified by electron tubes to a power level comparable with the radio signal.

3.2 Learning to Use the Vacuum Tube

As recounted in Chapter 4, de Forest invented the vacuum tube in 1907 but for the next five years he was concerned mostly with its use as a highly sensitive radio detector. It was brought to the attention of the Bell System in 1912 by John Stone Stone as a potential voice-frequency amplifier. In its original form it was not suitable for this purpose, but Arnold recognized and corrected the problem by evacuating the tube to a high vacuum. The tube was further improved by Western Electric engineers in 1913 by applying Wehnelt's oxide-coated cathode. This cathode emitted an adequate stream of electrons at relatively low temperatures and provided the means for achieving long life. This form of vacuum tube, stable and long-lived, was the basis not only for the voice-frequency repeater but also for the practical development of carrier and radiotelephony.

In the Bell System, understanding of the physics of the tube, its characteristics, and its equivalent circuit developed in parallel with the improvement of the tube itself, enabling it to be applied in various roles. On November 4, 1912, E. H. Colpitts conceived the push-pull amplifier. The Lowenstein amplifier circuit, which employed negative grid bias, was disclosed to the Bell System in January 1913 and the patent later acquired. Following work by Arnold on the "feedback" audion, Colpitts, in February 1914, devised a circuit for producing and modulating high-frequency oscillations. On February 10, 1915, R. V. L. Hartley disclosed vacuum tube oscillator circuits in which the plate was connected to one end of the tuning coil, the grid to the other end, and the filament to some intermediate point. This became known as the "Hartley oscillator." A month later Colpitts showed the capacitative equivalent, later called the "Colpitts oscillator."

In the same period there was also activity outside the Bell System in developing vacuum tube applications. Armstrong and de Forest in the United States, Meissner in Germany (using the von Lieben tube), and Franklin and Round of the British Marconi Company all developed oscillators, but de Forest, after prolonged litigation with Arnold, won patent priority on the basis of his 1912 experiments. Alexanderson of General Electric worked on modulators and produced a circuit at about the same time as Colpitts.

3.2.1 Short-Haul Experiments

In 1914 the Bell management decided that it was time to bring the vacuum tube out of the laboratory and attempt a series of radiotelephone field tests in the low-frequency part of the spectrum (mostly between 50 and 100 kHz).[23] In May a development program was begun to learn how to generate and control continuous-wave powers as high as 100 kW at low frequencies. Poulsen arcs and radio-frequency alternators were investigated, but in late 1914 it was decided that the state of the art was sufficiently advanced to warrant the construction of an experimental transmitter using vacuum tubes and the design of suitable equipment was undertaken. It was thought that the available power would be sufficient for at least a 300-mile transmission over water and, accordingly, a transmitting station was set up at Montauk, Long Island, and a receiving station on the roof of the du Pont Building at Wilmington, Delaware, 250 miles away. Successful one-way transmission was achieved on April 4, 1915, and then on May 18 speech was transmitted to St. Simon's Island, Georgia, a distance of about 900 miles.

Difficult problems had been faced in achieving these successful tests but many more had to be solved before commercial telephony became practical. Vacuum tubes had to be designed that would handle the large amounts of power required. Studies had to be made of the modulating, detecting, and amplifying properties of the tubes. Circuits had to be devised for carrying out the basic functions. In particular, oscillators were required for a wide range of frequencies, amplifiers were needed not only for the very high power signals radiated but also for the extremely weak received signals, and suitable modulation systems had to be devised. Finally, circuits had to be designed for interconnecting radiotelephone and wire telephone systems so that the speech signals from commercial telephone lines could be used to modulate the radio transmitter and the speech signals received by radio could be furnished to wire connections with characteristics suitable for wire transmission.[24]

3.2.2 *Early Overseas Tests*

Immediately following the successful Montauk–Wilmington tests, AT&TCo authorized an attempt to demonstrate radiotelephony across the Atlantic Ocean. In order to save time and minimize the cost of

[23] The work described in this and the following sections was carried out by a team of skillful and highly trained engineers and scientists organized to investigate and solve the many complex technical problems involved. While this was not the first use of this approach (the transcontinental wire transmission problem was handled in the same way), it certainly represented one of the most complex technical problems, up to this time, handled by coordinated team effort. It demonstrated the effectiveness of a technique which was to become the pattern for modern industrial research and development.

[24] The interconnection requirement ruled out the use of special types of microphones and head receivers.

the experiment, the U. S. Navy was asked for permission to use the antenna of its Arlington, Virginia, radio station. The Bell System would, of course, provide the radio transmitter, since such facilities were nowhere in existence, but use of the Arlington antenna, then one of the largest in the country, would avoid unnecessary expense and greatly reduce the lead time required. Permission was granted and every facility given to aid the experiment. A special type of vacuum tube was developed for the transmitter, many times larger than those used in wire telephony, with an output of 25 to 50 watts (Fig. 5-6). A transmitter was built, the final stage of which contained some 300 to 550 of these tubes connected in parallel (Type "W" in Fig. 5-6). When installed at Arlington, this transmitter gave an average power of about 2 or 3 kilowatts at the antenna at a frequency of around 50 kHz.

In this transmitter, a modulation technique known as the van der Bijl System was employed.[25] This system, classified as modulation by amplification control, was based on using the audion with suitable bias voltages so that the curve relating output current to grid voltage was parabolic. In effect, the audio- and radio-frequency signals were connected in series with the grid circuit of the modulating tube. The output contained the modulation products (see Chapter 4, Section 4.2.5) amplified to a level well above the input. The modulated signal was picked off the modulator tube by means of a tuned circuit and then fed to two stages of linear amplification. From 2 to 12 audions were connected in parallel for the first stage, depending on the frequency and number of power tubes being used. The second stage was the final amplifier mentioned above. The receivers used in these tests also depended heavily on new vacuum tube techniques. They employed a stage of radio-frequency amplification and two audio-frequency stages.

When the transmitter had been installed and tested, observers were sent to remote points to listen for the signals. R. H. Wilson went to Panama; H. D. Arnold and R. V. L. Hartley to Mare Island, California; Lloyd Espenschied to Hawaii; and Herbert E. Shreeve and A. M. Curtis to Paris. The Navy provided useful facilities at naval installations at all points except Paris, where the French Government permitted the use of its Eiffel Tower station for a short time each day.[26] The observers were supplied with their own radio receiving equipment.

On August 27, 1915, music from a phonograph and live speech

[25] H. J. van der Bijl; U. S. Patent No. 1,350,752; filed August 21, 1915; issued August 24, 1920.

[26] Even this small use of the Eiffel Tower station represented a remarkable degree of cooperation on the part of the French authorities. The war was still in a highly critical phase. Less than a year before, in September 1914, the Germans had been near the outskirts of Paris and were stopped only by rushing reinforcements in a fleet of Paris taxicabs. While this crisis had been averted, the German forces were still uncomfortably close at the time of the Paris tests and warfare at sea was being intensified as evidenced by the sinking of the *Lusitania* in May 1915.

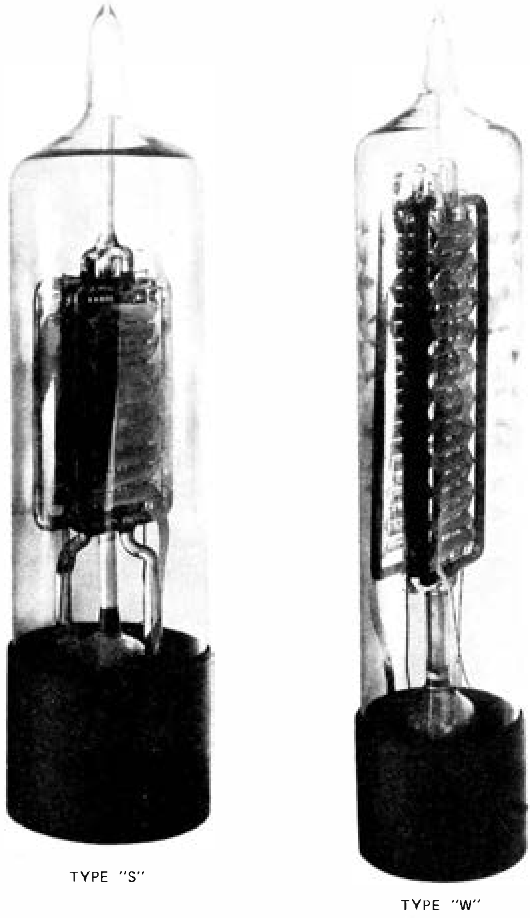


Fig. 5-6. Power tubes, designed by H. D. Arnold. The Type "W" was used in the first transatlantic radiotelephone transmission.

were successfully received at Darien, Panama. On September 29 speech was transmitted from New York to Arlington by wire and thence by radio to Mare Island and on the following day speech from Arlington was received at Pearl Harbor, Hawaii. The time available for reception at Paris was severely limited since the Eiffel Tower station was almost constantly in use by the French military. Scraps of speech were heard in the October 12–21 period, but it was not until the last day that connected, verifiable speech was heard. On October 23 a formal demonstration was arranged, and a number of observers at Paris listened to telephone transmissions from Arlington, 3,600 miles away (Figs. 5-7 and 5-8). Many of the transmissions to Paris were also heard by Espenschied at Pearl Harbor.

Following this demonstration, a Poulsen arc was set up at Montauk, Long Island, and on October 27, 1916, a one-way transmission took place between Montauk and New York. This was expanded into a two-way demonstration for officials of the Western Electric Company on December 17, the transmission from Montauk being by radio and from New York by wire. At about this time, the Bell System engineers became so heavily engaged in war developments that work on a transatlantic circuit was temporarily suspended.

IV. TWO-WAY RADIOTELEPHONE APPLICATIONS

The work just described was purely experimental. It was aimed at demonstrating feasibility of long-distance radiotelephone transmission and at evaluating some of the techniques required. It clearly showed a possibility for this type of telephone communication and also demonstrated the capabilities of the vacuum tube, but the work did not progress to the point of developing a working system for two-way communication. The first effort in this direction occurred in connection with Bell System work for the military.

4.1 Military Applications

The war came to Europe before the vacuum tube art had been applied to telephony and as a consequence the vacuum tube received only limited use by the European nations and only for radiotelegraphy. In the United States it was possible to continue the peacetime development for several years after the war began and as a consequence this country's technicians were in a much better position to apply radiotelephony to the war effort.

4.1.1 Ship-to-Ship and Ship-to-Shore Telephony

In 1916, about a year before the United States entered World War I, the U. S. Navy became interested in the possibility of quick voice communication between ships at sea and between ships and headquarters on land. Part of the latter concept involved the use of the wire telephone network to provide connections to appropriate land-based radio transmitting and receiving stations. In view of the successful overseas tests in which the Navy had cooperated, it was natural for them to request assistance from the Bell System in this project.

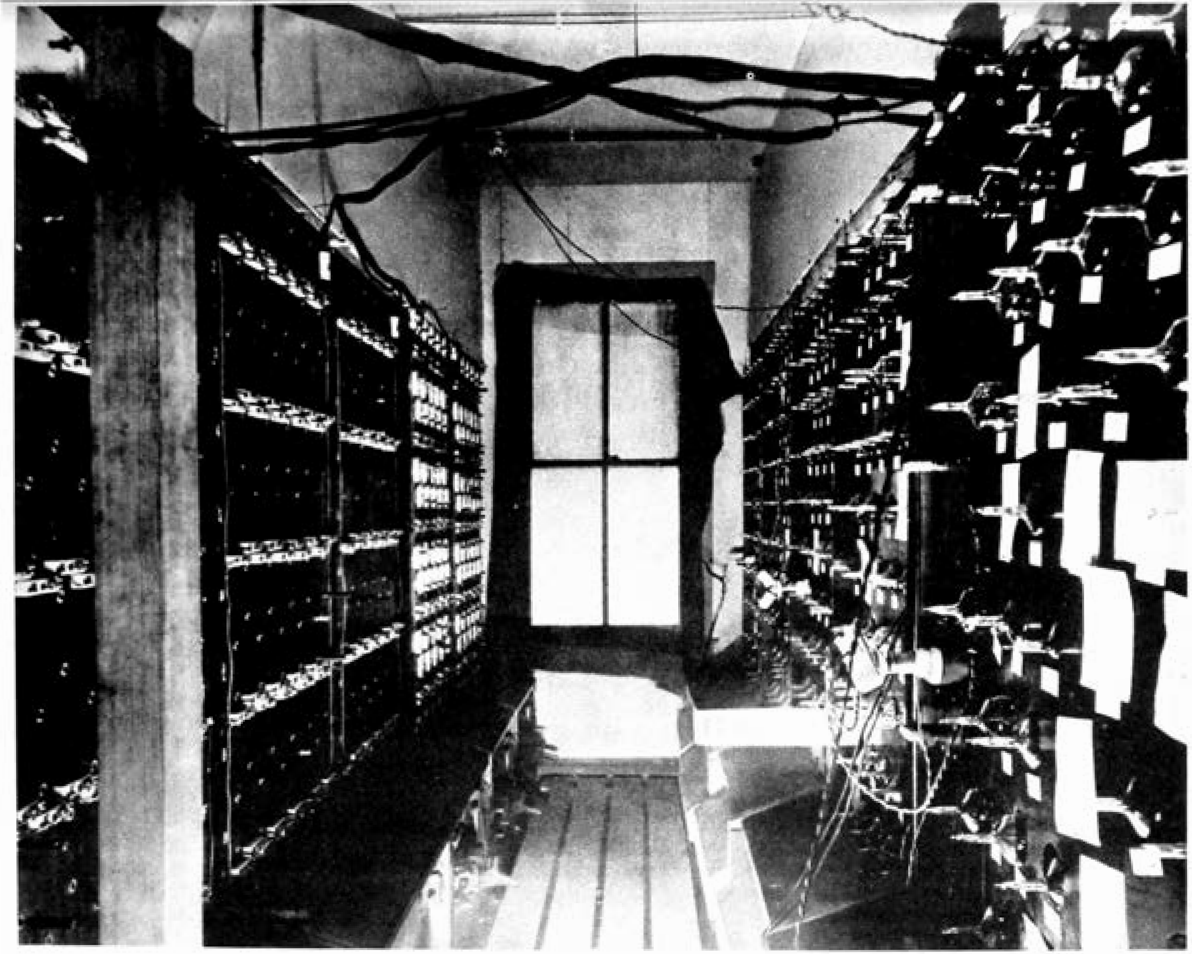


Fig. 5-7. Arlington long-wave transmitter of 1915. At right are some of the 550 tubes used in the final stage.

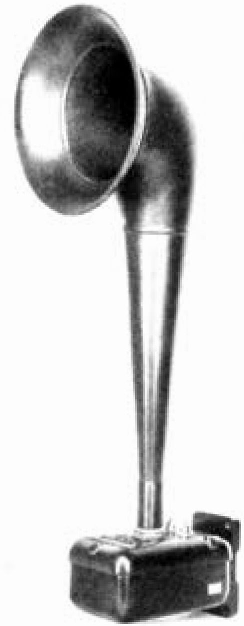


Fig. 5-8. Paris transatlantic loud-speaking receiver of 1915.

In April 1916 the Bell System began the installation of a special wireless telephone set on U.S.S. *New Hampshire* and on May 7 this equipment, together with the transmitting installation at Arlington and a receiving installation at Norfolk, Virginia, was utilized for a demonstration of long-distance radiotelephony for the Navy Department. Captain Chandler of the *New Hamsphire* talked to Secretary of the Navy and Mrs. Josephus Daniels and Mrs. Chandler in Washington, as well as with naval officers at Mare Island, California. This was the first time that two-way telephony through the wire telephone network had been extended to a vessel at sea.

Even earlier, the need for ship-to-ship communication was foreseen and the Naval Bureau of Steam Engineering requested assistance in its development. The construction of two sets was begun in December 1915 and they were completed in January and installed on the battleships *Arkansas* and *Florida* in February 1916. Conversations were carried out between ships up to 30 miles apart and signals were overheard as much as 175 miles away. This system also was similar to that at Arlington but much smaller in size and power output. It used roughly the same frequencies employed for radiotelegraph (600 to 1,200 meters or 500 to 250 kHz) and the two systems could not be used simultaneously without undesirable interference.

In order to avoid interference between telegraph and telephone and also to provide frequency space for a number of telephone channels, it was decided to exploit shorter wavelengths. In July of 1916, tests were carried on between U.S.S. *Arkansas* and U.S.S. *Florida* on 150 to 238 meters (2,000 to 1,200 kHz) to investigate the operation of multiplex radiotelephone equipment at the same time that the ship's radiotelegraph equipment was in use. Satisfactory performance was obtained at a distance of about 30 miles. These tests resulted in the design and installation of complete multiplex telephone and telegraph equipment on U.S.S. *Pennsylvania*, U.S.S. *Seattle*, and U.S.S *Wyoming* in January 1917. This equipment, designed by R. A. Heising, had subcarriers at about 25,000, 35,000, and 45,000 hertz which were modulated by voice and then used simultaneously to modulate the transmitter at any of the three wavelengths (150, 189, or 238 meters). Thus, a complete installation covering the three wavelengths could handle nine conversations at the same time. This appears to have been the first practical use of the "carrier" principle.

Upon declaration of war by the United States on April 5, 1917, effort in radiotelephony was changed from general, experimental projects to specific applications. In the previous month, the Navy Department had requested the construction of 15 experimental sets for use on submarine chasers where rapid short-range communications were important for coordinating their movements. These sets were primarily intended for continuous-wave telegraphy, but were also equipped with a telephone-modulating attachment. They were almost immediately superseded by an improved design based on developments in aircraft communication which will be described shortly. On November 10, 1917, a demonstration of experimental gear was held at New London, Connecticut, in which communication between submarine chasers was established at a distance of 4½ miles. At a second demonstration a loudspeaker was used so that a number of people could participate in the communication. The installation on the first group of 12 boats was completed in February 1918.

The standard equipment was designated CW-936 and operated on five frequencies in the 500- to 1,500-kHz range with a power of about 5 watts. This was the first successful radiotelephone equipment standardized by the Navy. It was not only responsible for popularizing voice operation in the Navy, but served as a prototype for following generations of radiotelephone equipment. The CW-936 equipment quickly came into demand for all kinds of uses and over 2,000 units were installed on ships of both the United States and British Navies. As late as 1930 they were still in wide use. In the early 1920s they were a popular item on the surplus market and found their way into college laboratories as well as private hands.

4.1.2 Aircraft Communication

Both the U. S. Army Signal Corps, which operated the Army Aviation Service, and the U. S. Navy saw the need for radio communication between airplanes and ground and between the planes themselves. Both agencies called on Western Electric for assistance. As early as September 1916 an experimental seaplane radiotelephone set capable of operating in the 300- to 1,000-meter (1,000- to 300-kHz) range was developed for the U. S. Navy. It was tested in Washington by the Navy but was never installed on a plane because of weight limitations.

Early in May 1917 radio work within the Western Electric Company was redirected towards areas that might be useful to the military and the organization was well prepared when on May 22 a formal request was received from General Squier, the Chief Signal Officer, for the development of an airplane set with 2,000-yard range. The speed of response seems marvelous today when we have become accustomed to long lead times. A few dates will illustrate:

May 22, 1917—Authorization for development received
June 5—Requirements review in Washington

July 1—Field test started on experimental equipment
July 2—First air-to-ground transmission (about 2 miles)
July 4—First ground-to-air transmission (about 2 miles)
August 20—Two-way communication between planes in flight (up to 2 miles)
October 14—Standardized sets tested
October 16—Distance tests successful to 23 miles (ended by break in plane cooling system)
October 24—Sets boxed for shipment abroad for field evaluation
December 2—Final tests at Dayton, Ohio; quantity orders for manufacture of set coded SCR-68 placed immediately after
November 2—Request for adaptation of airplane set to submarine chaser
November 10—First tests at sea
December 27—First installation of new set designed for chasers
January 11, 1918—Successful field test
February 1918—Installation on 12 boats completed.

All of the above work was done with transmitters employing the Colpitts oscillator and the Heising "constant current" modulator. Receivers used a "gridleak" detector circuit followed by two stages of audio amplification. Two types of vacuum tube using the long-life Wehnelt cathode were employed, the VT-1 for receiving and the VT-2 for transmitting, the latter delivering about 3 to 5 watts of high-frequency power. Transmission was in the 200- to 500-meter range (1,500 to 600 kHz) using a trailing-wire antenna on the plane. Power on the submarine chasers was obtained from the ship's battery but on planes no power was available, and attachments to the engine were not permitted. Therefore a wind-driven generator placed in the propeller's slipstream was developed to meet the communication needs. Available regulators were not adequate to maintain constant voltage over the wide range of generator speeds which resulted, and a unique vacuum tube regulator was developed by H. M. Stoller that maintained voltage within close limits over a three-to-one speed range.[27]

Experimental equipment used in the early airplane tests is pictured in Fig. 5-9. Somewhat later equipment is shown in Fig. 5-10, together with participants in a demonstration. The standardized SCR-68 transceiver is shown in Fig. 5-11 and the complete equipment for a two-place airplane is shown in Fig. 5-12. Included were means for voice communication between the two occupants which

[27] The generator had a main field with a vacuum tube filament in series and an opposing field in series with the plate circuit. Increasing speed increased the main field and the filament current which in turn caused the plate current and opposing field to increase and thus maintain the voltage approximately constant.

Fig. 5-10. Early aircraft-radio tests. E. B. Craft is at the left, and Col. N. H. Slaughter at the right. The two officers behind the table are First Lt. Ralph Bown and Major Nathan Levinson.

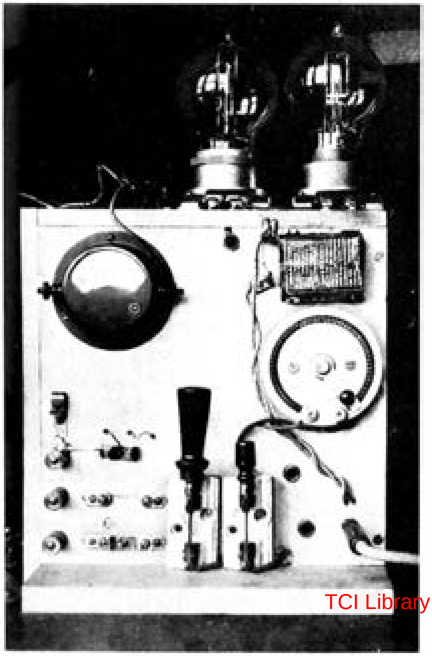
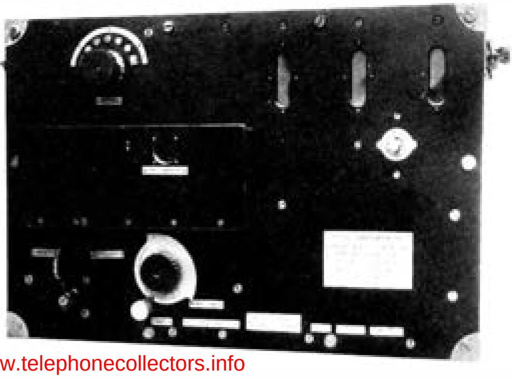


Fig. 5-9. Experimental equipment used in early airplane radio tests.

Fig. 5-11. Signal Corps Standard SCR-68 radiotelephone transmitting and receiving set—transceiver. (Craft and Colpitts 1919, Fig. 32)



previously, because of the high noise level, had only been possible by hand signals. The wind-driven generator is shown in Fig. 5-13. A circuit schematic for the transceiver is shown in Fig. 5-14. The weight of the complete equipment was about 58 pounds, including the generator and the two operator's sets.

Probably the most important requirement in the work just described was rapid development. The systems resulting were not what would have been developed under normal conditions with time available for study of technical problems and optimization of design. Instead, the design represented a compromise using available circuitry and parts, so far as practical, supported by brief experimental demonstrations of feasibility.

One of the important limitations on the use of the SCR-68 airplane equipment was imposed by the long trailing-wire antenna. This was satisfactory for training and some other applications but would be hazardous in tactical flights requiring high maneuverability. A smaller radiating system was obviously desirable and this in turn would require the use of higher frequencies. In October 1917 some laboratory sets were made for wavelengths of 70 to 150 meters (4 to 2 MHz). Later the development of such a set was requested by the Signal Corps and it was built in April 1918 and tested thereafter at Camp Alfred Vail in New Jersey. Although the design was finalized (Figs. 5-15 and 5-16), it came near the end of the war and apparently there was little, if any, production. Electrically these sets were similar to the SCR-68 with changes to minimize the effects of wiring capacitance and inductance which became much more important at the higher frequencies. Because of the short wavelength, smaller antenna systems were possible. Originally, a short structure on the top of the airplane was used with two wires extending to the tail. Later studies of antenna radiation resistance and pattern showed that a superior arrangement consisted of two short unweighted wires, one from each wing tip, joined by a wire above the fuselage, the whole working against the conducting portion of the plane as a counterpoise.

4.1.3 Postwar Effort

Some of the work described above naturally carried over after the armistice of November 11, 1918, to bring going projects to completion. In addition, development of a line of standardized radiotelephone and radiotelegraph sets was undertaken. It was planned that the smallest of these would have a power of 5 watts and a nominal range of 10 miles, and the largest would have a power of 500 watts and a nominal range of 100 miles. Demonstration models of the 500-watt set were built and in 1922 used to provide a radiotelephone circuit between Peking and Tientsin, China. In addition, a duplex

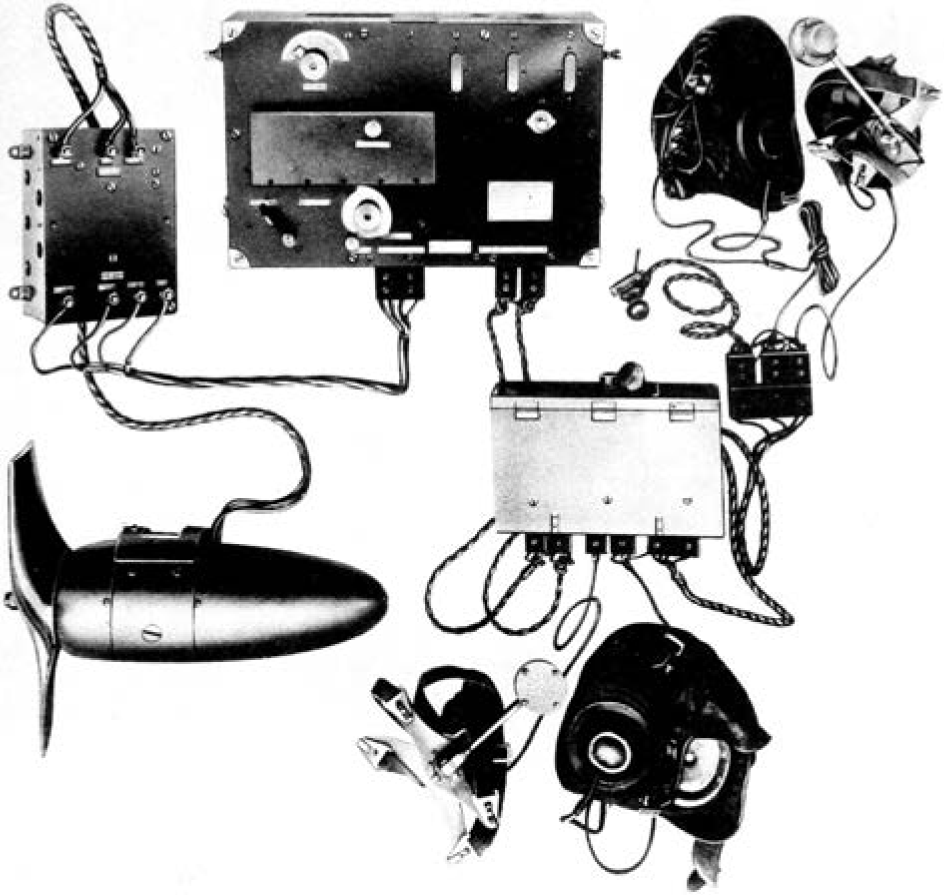


Fig. 5-12. Signal Corps Standard SCR-68 radiotelephone transmitting and receiving set—complete equipment for two-place airplane. (Craft and Colpitts 1919, Fig. 31)

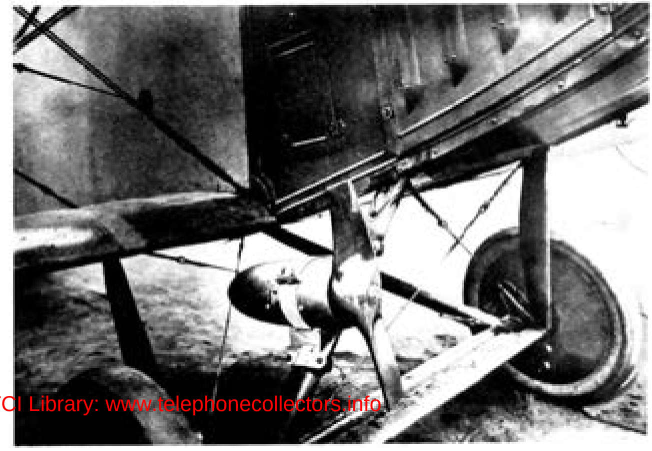


Fig. 5-13. Method of mounting wind-driven generator. (Craft and Colpitts 1919, Fig. 33)

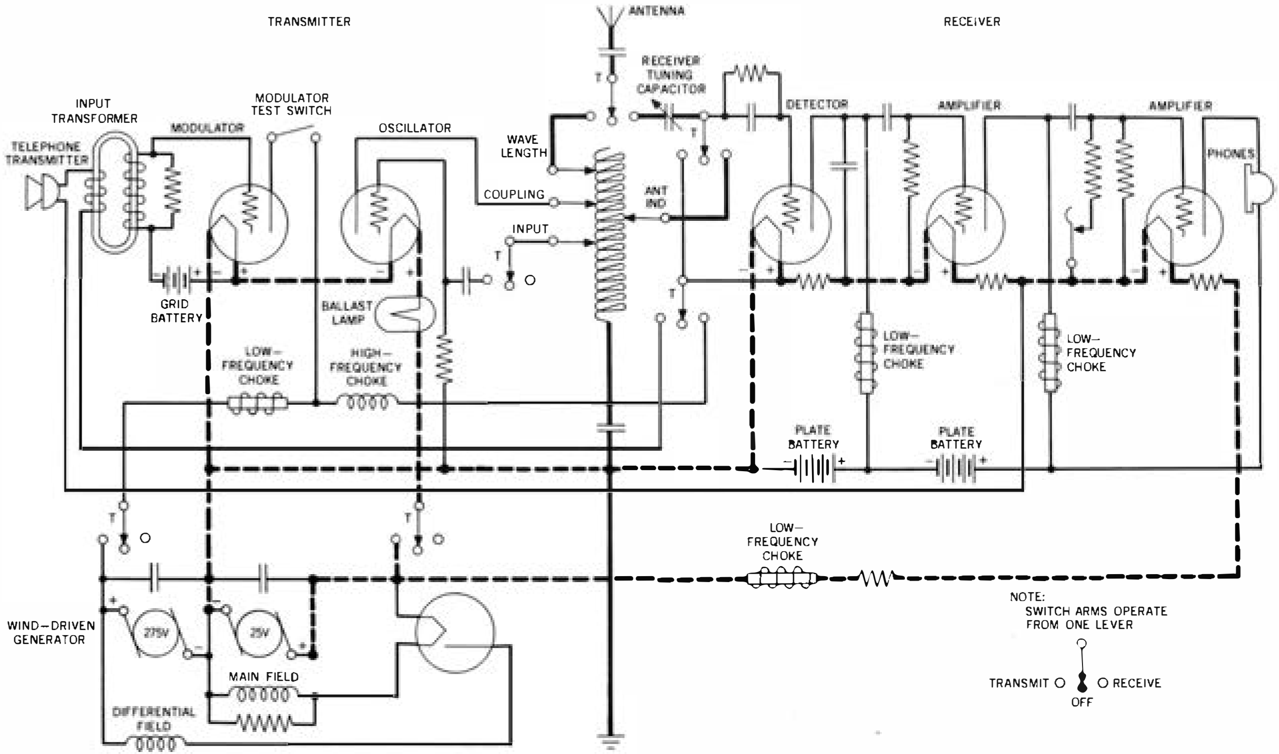


Fig. 5-14. Signal Corps Standard SCR-68 radiotelephone transmitting and receiving set—schematic diagram. (Redrawn from Craft and Colpitts 1919, Fig. 35)

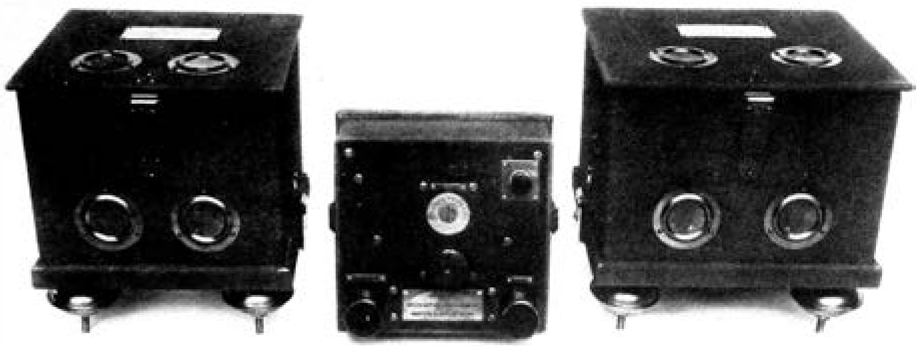


Fig. 5-15. Short-wave-type airplane radiotelephone set. (Craft and Colpitts 1919, Fig. 36)

radiotelephone set for fire control use on all types of naval vessels was developed for the Navy in 1919.

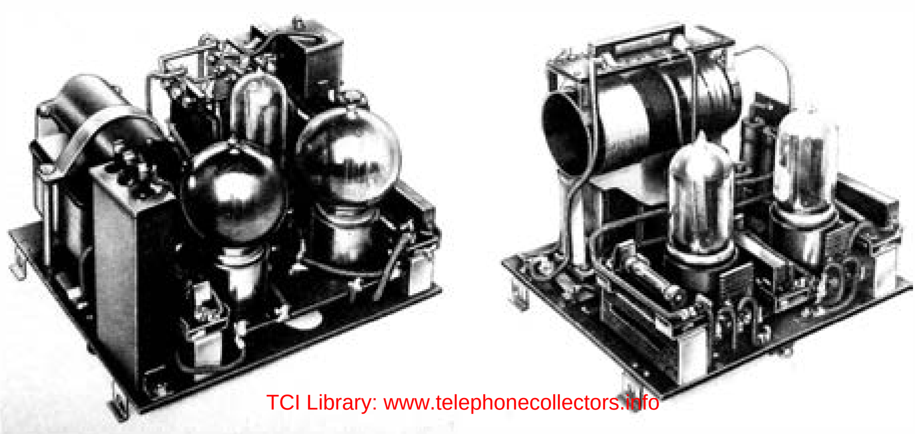
However, it soon became apparent that Bell System needs would place heavy demands on the available development manpower and Western Electric largely withdrew from the military communication field[28] until the late 1930s when the technical skills of the Bell System were again required in preparing for and prosecuting World War II.

4.1.4 The Military Effort in Retrospect

The developments in radio communication just described had little direct effect on the outcome of World War I since the 19 months

[28] The patent agreements reached in 1920 (see Section 4.2) also must have played a part in reaching this decision.

Fig. 5-16. Interior construction of short-wave transmitting and receiving units (Craft and Colpitts 1919, Fig. 39)



between our entry and the war's termination provided little opportunity for the application of new technology. Of the many radio sets for airplanes, submarine chasers, etc., in production, few actually saw service. However, the impact of the wartime effort on military thinking was considerable. The feasibility of rapid voice communication between all military vehicles in the air or on the sea had been effectively demonstrated and there would be no return to the primitive means for communication in use prior to 1917.

The impact of the war on technology was also very substantial. Because of the need for haste, there had been little time or effort devoted to fundamental analytic studies. But the pressures of necessity greatly stimulated invention and ingenious empirical solutions to a wide variety of technical problems. In addition, this work provided useful experience in the standardization and quantity production of apparatus and components, particularly vacuum tubes.[29]

The problems involved in airplane communication illustrate the effect of war requirements on technology. In June 1917, when the first meeting was held in Washington to establish technical requirements, it was found that practically nothing was known about the wavelengths desired, antenna types available, antenna radiation characteristics, power supplies, and similar essentials. Many of the decisions were made by judgment backed by a few field tests. It was on this basis that the original wavelengths were selected. The trailing-wire antenna, adopted for early systems, had been used abroad for telegraph transmission. The solution of the power problem has been described. Circuitry for detectors, amplifiers, and oscillators was available as the result of prewar effort. In airplane systems low weight was essential, and this required efficient modulation of the carrier to minimize the amount of carrier and prime power required. The Heising constant current modulator proved most suitable and, with the Colpitts oscillator, formed the basis for all airplane sets.

However, there was no experience available for the solution of many problems. Acoustic noise in the plane, in which the occupants sat in open cockpits, was enormous. Helmets containing built-in telephone receivers were developed to reduce the noise reaching the ear (Figs. 5-12 and 5-17) and microphones that minimized noise pickup were developed largely on a trial and error basis. Electrical interference from the ignition systems of airplanes and submarine chasers was at a very high level. On planes no alteration of the plane equipment was permitted, but shielding of the ignition wires was found to be acceptable. Small choke coils in the ignition leads solved

[29] The total output of commercial vacuum tubes was less than 200 per week in August 1917, but by the end of the war in November 1918 deliveries were at the rate of 25,000 per week.

the problem in sumarine chasers. Weight and space problems responded to new packaging approaches, including the use of a multiple-unit plan of assembly. Both circuit and equipment design were aimed at eliminating minor adjustments. Vibration was far beyond anything experienced on the ground and was particularly critical in the design of vacuum tubes. Figures 5-18 and 5-19 show the tube structures developed to meet the mechanical requirements of airplane service.

Somewhat hidden in the war effort was a lesson that was soon to be recognized and has had a continuing impact on Bell System research and development. So much progress could be made in such a short time in an area where there were so many unknowns only because there was available a sound background of basic knowledge on which to build. Previous work on acoustics, interference prevention, power supplies, and design of telephone instruments, when added to theoretical and practical knowledge of the vacuum tube, provided the background for quickly putting together a complex system meeting completely new requirements. The large element of judgment and empiricism which was required to meet the needs of the moment was successfully applied only because it was backed by broad-based knowledge and experience.

Probably the greatest impact of the war on radio technology came from the demonstration of feasibility. At the end of the war many theoretical problems, bypassed through empirical solutions, would have to be solved to put radio communication on a sound basis, but these could be approached with the knowledge that "it worked."

4.2 Postwar Bell System Experiments

Work during World War I had eliminated doubts as to the technical feasibility of radiotelephony but left many questions concerning its commercial future. In reviewing the situation Bell System managers and scientists found three questions of considerable concern:

(*i*) What is the most likely field of use for radio in the Bell System?

(*ii*) How could progress in applying radio be made in the face of the complex patent situation?

(*iii*) Is commercial development consistent with the availability of frequency space?

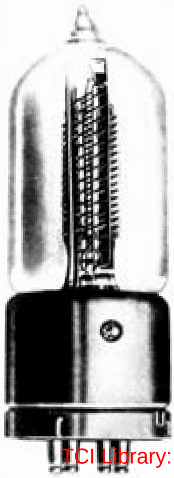
Before relating the technical evolution of commercial radiotelephony, we should examine these questions a little more closely to see how they affected the program which was ultimately followed.

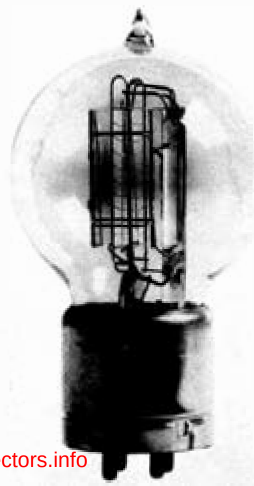
The justification for the wartime program had been obvious. Rapid communication between bases and vehicles without the use of skilled operators was highly desirable and, in the face of this military need, the cost of providing the service was not of major concern. In com-

Fig. 5-17. Telephone receiving helmet for airplane radio. (Craft and Colpitts 1919, Fig. 40)

Fig. 5-18. Western Electric receiving-type vacuum tube—Army and Navy Standard. (Craft and Colpitts 1919, Fig. 17)

Fig. 5-19. Western Electric transmitting-type vacuum tube—Army and Navy Standard. (Craft and Colpitts 1919, Fig. 25)





mercial service, on the other hand, the use of radio must be justified economically as compared to other media and, even where there is no alternative, a commercial service is warranted only if, in the long run, the users are willing to pay the cost.[30] As early as February 1919, E. B. Craft and E. H. Colpitts reviewed this matter in a paper presented to the American Institute of Electrical Engineers. Their conclusion was that with the conditions prevailing then, and in the foreseeable future, radio transmission would not be a competitor for wire transmission in densely populated areas because of cost, lack of privacy, and problems of selectivity and interference. On the other hand, there were areas where radio was feasible and wire telephony was a physical impossibility. Their final conclusion is best presented by quoting their summary:

> Fortunately, however, the connection of a wire system to a radio system is [not] complicated[31] . . . Leaving aside for the moment the particular methods by which radio telephone communications is to be carried on, it is clear that the establishment of communication between two individuals will be most efficiently realized through the use of a combination of wire transmission on a network extending over perhaps 99% of the stations and radio transmission to those relatively few stations to which it is either impossible or impracticable to build lines. These stations will be of two kinds:
>
> 1. *Moving*, such as ships, airplanes, trains, trucks;
> 2. *Fixed but inaccessible*, such as on islands, in deserts and very sparsely settled regions.
>
> A third class of service is that which is concerned, not with single individuals, but with groups; such service as the broadcasting of news, time and weather signals and warnings. In some cases one objection to radio telephony (i.e., omni-directional radiation) would be an advantage in this class of service.

This situation existed for the next 20 to 30 years after which the evolution of microwave and carrier telephone techniques began to eliminate some of the factors previously limiting radio applications and opened the field to economic competition with wire lines.[32] Before passing on to our next problem we should note that even with the limited outlook for radio application that existed at the end of the war there was never any doubt that the technological development should be pursued. The prewar experiments had clearly demonstrated the

[30] An exception might be a supplementary service which, while not paying its way originally, provides intangible benefits by greatly increasing the usefulness of the telephone network as a whole. Such arguments could be advanced on behalf of overseas and mobile services.

[31] This statement is slightly optimistic as the reader will find when he comes to Section 4.3.5 which covers the interconnection problem.

[32] At the time this is being written, in the early 1970s, microwave radio has become an important competitor for wire, particularly over long distances, but wire is still the preferred medium for circuits in densely populated areas.

value of developing technology so that it would be available when need for it arose.

Before the United States entered World War I, the patent situation had become complex since work in radio, particularly that involving the new audion, had become widespread. Inventors all over the world had applied for or received patents, many involving conflicting or overlapping claims, that were fundamental to the development of the art. During the war the urgency was great and work proceeded regardless of the rights and claims of individual inventors, all companies making their inventions freely available to the government. The stimulus to invention provided by this work produced even more patent applications. As a result the patent situation at the end of the war approached chaos. There were a large number of conflicting claims concerning fundamental principles based on inventions made almost simultaneously. No one could be sure which inventor held priority until this was decided through litigation which, under prevailing conditions, might take many years to resolve. Some orderly basis for proceeding without the possibility of future penalties for patent violation was necessary if the art was to be reduced to practice.

Where it was practical the Bell System purchased the right to use fundamental patents, such as that for the Lowenstein negative C battery (grid bias), but the vacuum tube patents and applications held by the large manufacturing organizations were so many, complex, and conflicting that resolution of the patent problem through bilateral agreements was unlikely. Instead, it appeared more practical to pool these patents and define their field of use by the interested concerns. Accordingly, at the request of the United States Government, the General Electric Company and AT&TCo entered into a cross-license patent agreement, effective as of July 1, 1920. Following the original agreement, extension agreements were made with the Westinghouse Electric and Manufacturing Company and the newly formed Radio Corporation of America (which had taken over the interests of the Marconi Company in the United States). As a result of these agreements, the situation in the early 1920s was roughly as follows: Radiotelephone equipment for commercial and public service uses was to be provided by AT&TCo or through its manufacturer, the Western Electric Company. Amateur radiotelephone equipment (later interpreted to include radiotelephone broadcast receiving sets) and radiotelegraph equipment were to be manufactured by the General Electric Company and the Westinghouse Company and sold through the Radio Corporation. The underlying principle throughout this cross-licensing agreement was to free radio development from the possibility of future disastrous litigation and thus assure availability to the public of the best technical methods at an early time. It also played a large part in defining the areas in which the Bell System was to apply radio technology.

The problem of frequency allocation had been of importance practically from the beginning of wireless. By the very nature of radio waves, which could travel great distances without regard for physical boundaries, it was an international problem and was so recognized as early as 1903 when a preliminary conference was held in Berlin. Three years later the first truly international conference (attended by 29 nations) met, again at Berlin, to consider commercial and technical problems, the major one being frequency allocation. At this time the frequencies below 188 kHz were assigned to long-distance coastal (point-to-point) communication, the frequencies from 188 to 500 kHz were designated for military use, and two frequencies (500 and 1,000 kHz) were set aside for maritime (ship-to-ship and ship-to-shore) use. It is interesting to note that the last two frequencies were shared by all maritime operations, the problem at this time being more that of establishing a common working frequency where all could come together, rather than finding a way to spread them out over a band; congestion was obviously not yet a problem. These allocations were confirmed at the London Conference of 1912, the last before the first World War, and a few more allocations were made for new services such as radio beacons, weather reports, and time signals. However, telegraphy was still the service of main concern and the only technically feasible type of two-way communication existing at the time.

After World War I the situation was completely changed; two-way radiotelephony was a technical reality and by 1920 broadcasting was clearly a possibility. The question of where to place these services in the frequency band was fundamental to their development and was further made critical by the fact that each channel would require some one hundred times as much frequency space as a telegraph channel. Clearly the problem of congestion was now paramount.

Some of the frequency allocation problems were entirely domestic and all had at least some effect on domestic developments. Therefore, as a first step in bringing order to this matter, the U.S. Secretary of Commerce appointed a committee to study radio allocations in the light of the changed situation. This committee, meeting first in February 1922, produced recommendations on which subsequent legislation was based. They proposed tentative frequency assignments which were consistent with their recommendation that a large part of the available frequency space be set aside for various kinds of broadcasting with a small amount reserved for ship-to-shore, transoceanic, and fixed stations.

After this, and similar preliminary work elsewhere, the problem was taken up on an international basis at the Third International

Conference in Washington in 1927 with U.S. Secretary of Commerce, Herbert Hoover, as Chairman. This meeting undertook to allocate all frequencies between 10 hertz and 60 MHz.[33] By this time broadcasting had become a major user of frequency space. The full solution of the broadcast allocation problem did not come for some time and involved complications we need not go into here. Of most significance to the United States was the confirmation of the allocation, previously agreed to domestically, of 500 to 1,500 kHz for United States broadcasting. The use of this wide frequency band for broadcasting involved clearing the space of other users and had a considerable effect on Bell System work which had begun shortly after the war.

The uncertainty concerning frequency allocation was the last of the major postwar problems to be cleared up, but by the latter half of the twenties it had been clarified to a point where commerical planning and development work could proceed.

Before leaving the subject of frequency allocation, we should mention a rather special situation that existed in the United States from an early date and later developed international implications. There had grown up in this country a large body of amateur wireless experimenters. It has been estimated that as early as 1910 about 4,000 amateur stations existed, some four or five times the total number of commercial and military stations. These amateurs operated without regulation, being free to use any power and work at any frequency they saw fit. Being non-commercial, they felt free to ignore patent licensing regulations and were quick to adopt new and promising techniques; as a consequence they sometimes had equipment superior to commercial stations and by sheer numbers practically dominated the air. Several attempts had been made to control (or prohibit) their operation but these were not successful until 1912 when a regulation covering amateurs was included in the Alexander Bill which was intended primarily to implement the 1912 international agreements. The regulation covering amateurs (which had no international counterpart at that time) provided for licensing of stations, prohibited the use of wavelengths over 200 meters (frequencies under 1,500 kHz), and limited power to a maximum of 1 kW. At that time the frequencies above 1,500 kHz were considered largely useless, particularly with low radiated power, since surface-wave losses were very high. The regulation was intended to eliminate interference from amateurs at commercial frequencies and hopefully to eliminate amateur stations completely. Fortunately the latter objective was not accomplished, the

[33] A start was also made on limiting the inefficient use of frequency space. The spark station, which had so long been a wasteful user of frequencies, was outlawed after 1930 (except for very low power stations) and no new stations of this type were permitted after 1927.

amateur group continued to grow, and the amateurs provided a large number of experienced operators when they were badly needed during World War I and in the postwar years. Amateurs also were to play an important part in the rapid development of broadcasting. Even more important, they discovered that the frequencies allocated to them were usable and their experiments, carried out in the early twenties, had a significant effect on the development of overseas communication.[34] By 1927, the worth of amateur radio had been so well demonstrated that specific frequencies in six bands were assigned on an international basis in spite of initial opposition from almost every government except that of the United States.

4.2.1 Ship-to-Ship Experiments

Although the commercial outlook for radiotelephony was uncertain, it was agreed by Bell management that the radio art could best be developed by selecting a specific goal rather than attempting broad and undirected experimentation. The first work undertaken, in 1919, was to develop a radiotelephone system for vessels at sea that could be connected to Bell System telephone customers through the wire network.

The experiments were carried out with two coastwise ships, the *Ontario* and the *Gloucester*, which sailed between Boston, Philadelphia, and Baltimore and thus were nearly always within a few hundred miles of the land stations located near Green Harbor, Massachusetts, and Deal, New Jersey.[35] Transmission was by double sideband with carrier transmitted, using separate transmitters and transmitting antennas for each channel at the shore stations. On shipboard, the same antenna was used for transmitting and receiving. Different frequencies were used for the two directions of transmission, and further decoupling of the two directions (necessary to avoid singing, as discussed in Section 4.2.4 of Chapter 4) was obtained by using space separation between the transmit and receive antennas at the shore station, together with hybrid balance at the connection to the 2-wire telephone line. The technical advantages of single-sideband, suppressed-carrier transmission were recognized at the time of these experiments but the technique was not adopted because of practical problems. It was highly desirable to minimize the amount of special equipment required on shipboard and at the time most ship receivers were not suitable for single-sideband reception.

[34] Some of the contributions of amateurs will be covered in more detail later in this chapter.
[35] In 1922, in cooperation with the Radio Corporation of America, S. S. *America* was also equipped and tests were made on transatlantic runs, including the measurement of field strength over the entire distance.

By the fall of 1920 a demonstration of simultaneous three-channel operation from a shore station had been carried out. The wavelength was in the neighborhood of 400 meters (750 kHz) and the transmitters were capable of putting out about 1 kW of modulated radio-frequency power. This output was provided by six power tubes connected in parallel, each with a rating of 250 watts. The shore station for this operation is shown in Fig. 5-20 and an interior view of the operating room in Fig. 5-21. An inside view of one of the transmitters is presented in Fig. 5-22. The shielded oscillator is at the upper right and the other two shielded compartments contain amplifiers. The large unshielded compartment contains six radio-frequency power amplifiers.

On shipboard three power tubes rated at 250 watts were employed, one being a master oscillator, one a power amplifier, and one a modulator. A front view of an early experimental shipboard transmitter showing the three tubes is presented in Fig. 5-23.

The objective in all of this work was to provide a radio circuit having a net loss, that is, a loss from the standpoint of voice transmission, of 6 Transmission Units (dB). This was achieved by manually adjusting the gain, or amplification, of the receiver to compensate for the many variations that occurred in the transmission path. A range of 40 dB was required in order to handle all of the variations resulting from distance and diurnal effects.

This work was brought to a conclusion in the early 1920s by the postwar depression in shipping and by the allocation of the most desirable frequency space to broadcasting. As we shall see shortly, it was soon found possible to use high frequencies for long-distance transmission and work on ship-to-shore service was again undertaken in the late 1920s, largely as an outgrowth of the transoceanic developments discussed in Section 4.3. Service was initiated on the North Atlantic in 1929 using shore facilities similar to those used for transatlantic service. The ship installation included a 500-watt transmitter employing a screen-grid power tube. This service has been continuously extended ever since, not only to large passenger ships but also to small boats on coastal and inland waters. In a sense, all of today's service to mobile units, which now includes automobiles, trains, and airplanes, has grown out of the original concept of broadening communication coverage by using radio to supplement the fixed-wire network.

Although the work of the early 1920s did not immediately lead to a commercial service, it had an important impact on technology and provided much background information that was required when mobile service was established at a later date. Some of the technical achievements were:

(*i*) Establishment of the requirements for, and development of, a workable duplex system
(*ii*) Development of highly selective, superheterodyne receivers which permitted close spacing of channels
(*iii*) Progress in placing radio transmission on a quantitative basis by the measurement of field strength and circuit net loss over long periods of time, thus demonstrating not only the variation with distance but also the large variation between day and night transmission.
(*iv*) Demonstration of the importance of path variability and talker volume on system design and operation. This pointed up the ultimate need for automatic control of radio circuit gain and voice level.
(*v*) Demonstration of the serious nature of interference from spark-type transmitters and the need for their elimination.

4.2.2 First Commercial Application

In the meantime, work was going on that was to lead to the first use of radiotelephony for public service. Early in 1920, a need developed for telephone service to Catalina Island, off the coast of California. A submarine cable would have been the natural choice for this service since the distance was short and ultimately more circuits would be required than could readily be supplied by radio. However, conditions growing out of the war were such that cable could not be supplied in the time desired and the Pacific Telephone Company requested the American Telephone and Telegraph Company to develop a radio link to the mainland. The first conference on this project was held on April 30, 1920, and just 77 days later, on July 16, the first commercial radio and connecting land toll line was opened for public service between Avalon, Catalina Island, and Long Beach on the mainland (about 30 miles distant) using equipment shown in Figs. 5-24 and 5-25.

Each transmitter in this installation had a power output of about 100 watts, and 4-wire communication was obtained by using two frequencies—638 kHz from the mainland to Catalina Island and 750 kHz in the opposite direction. Hybrid balance was used at the junction with the 2-wire land network. A telegraph channel was multiplexed with the telephone channel and used independently for telegraph service to the Island. Another interesting feature was the use of through voice-frequency ringing, the same technique that was being applied to carrier-current transmission over wire.

This system provided good service, but the frequencies employed were in the middle of the broadcast band, and the government soon

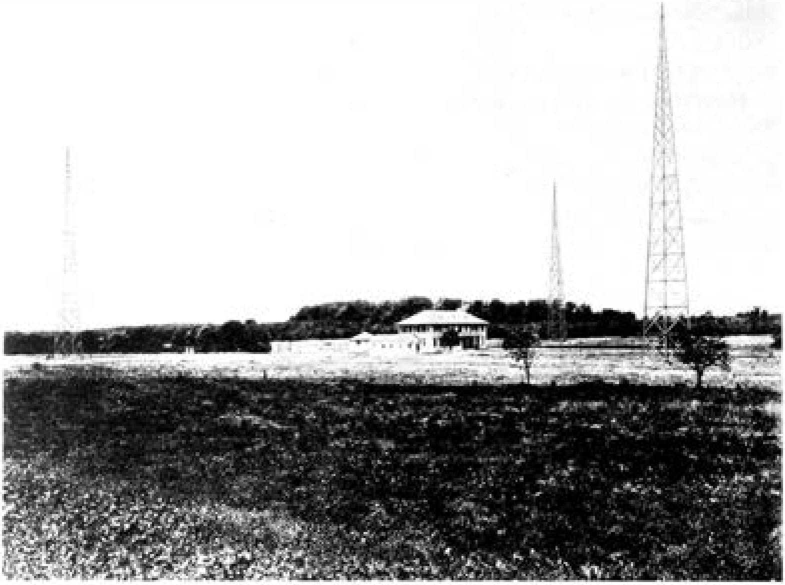
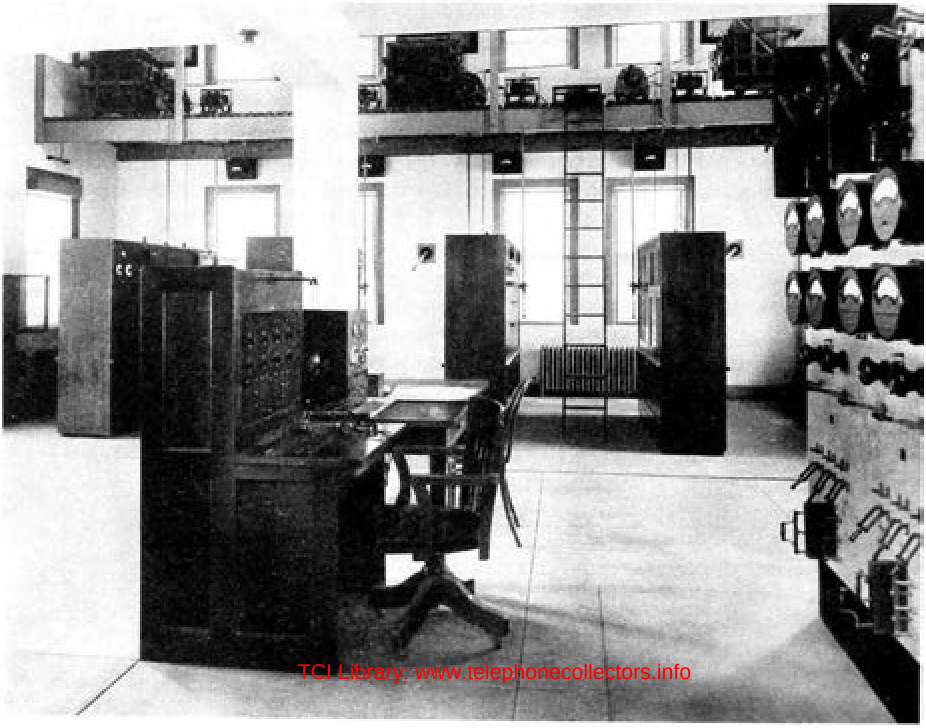


Fig. 5-20. Short-wave radio station at Deal, New Jersey. (Craft and Colpitts 1919, Fig. 27)

Fig. 5-21. Operating room of Deal Beach Station. Operator's desk is in the foreground, transmitters in the background, and antenna loading and coupling coils on the balcony.



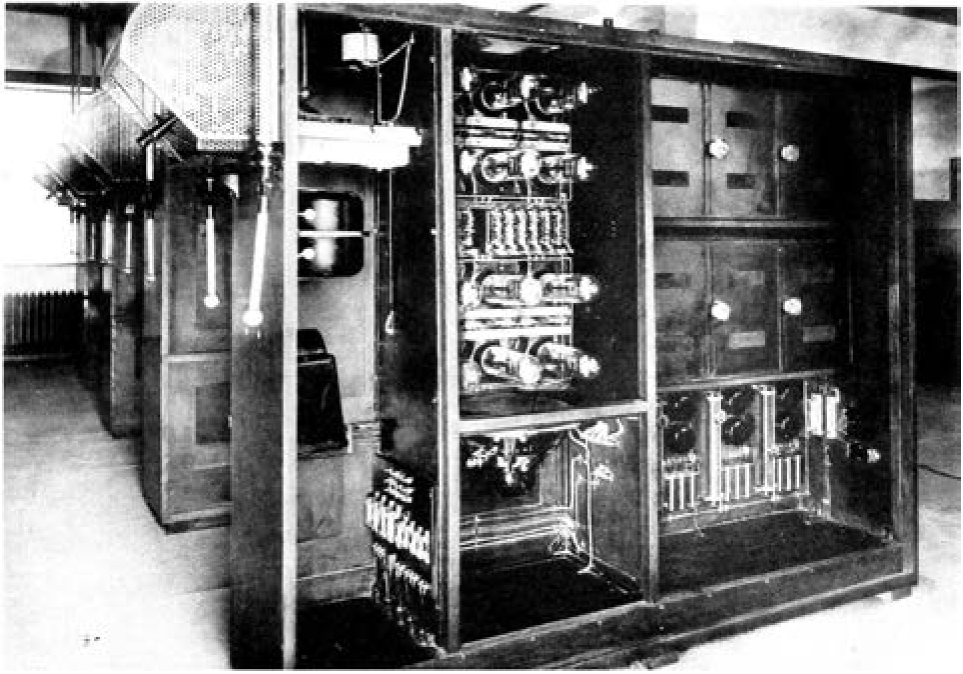


Fig. 5-22. Interior view of one of the four transmitters at Deal Beach Station. These transmitters can be seen in the background in Fig. 5-21.

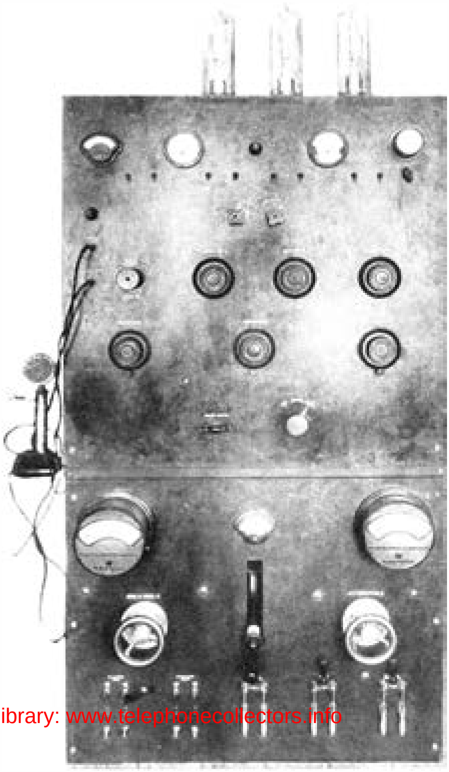


Fig. 5-23. Front view of experimental shipboard transmitter.

Fig. 5-24. Radiotelephone station at Pebbly Beach, Catalina Island.

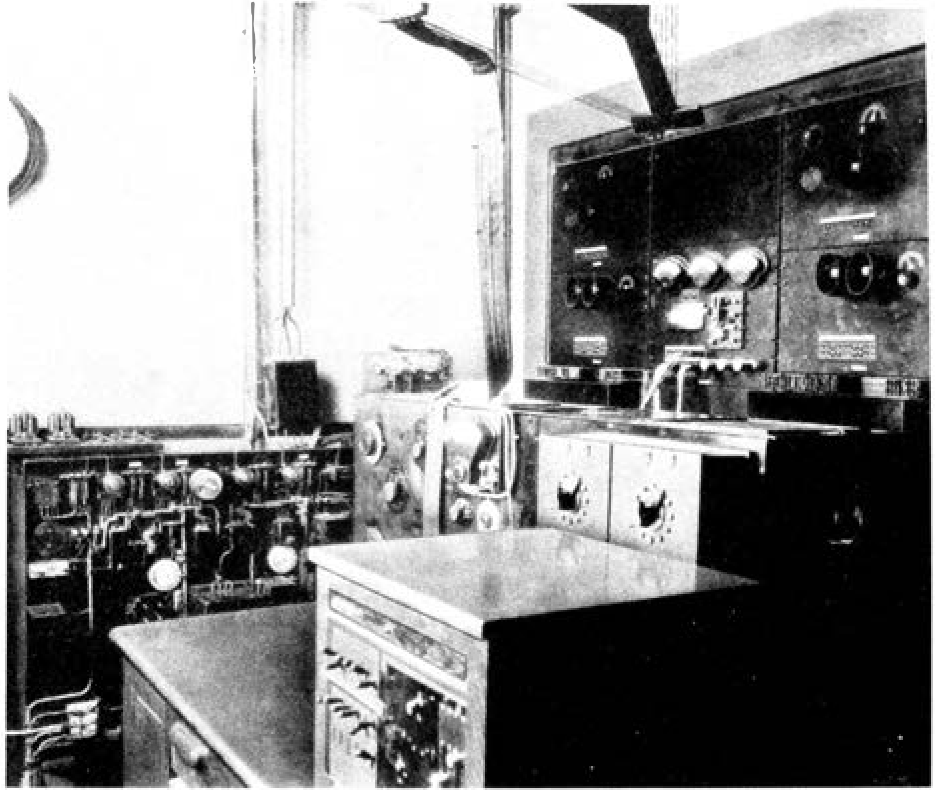


Fig. 5-25. Receiver at Long Beach, California, for the California–Catalina Island radiotelephone link.

requested that service be terminated so that the frequencies could be used for the rapidly expanding broadcast services. Thus the system was shut down on August 1, 1923, and telephone service was provided thereafter by two submarine cables.

One disadvantage of the radio system was that anyone with a suitable receiver could listen to the conversations (and nearly everyone did). To solve this problem a "privacy" system was added some six weeks before the radio setup was taken out of service.[36]

Work on the Catalina system not only provided additional technical background in the use of full-duplex systems, multiplexing, signaling, and privacy but also demonstrated the possibility of using radio for temporary circuits when a quick response was required.

Present-day workers in the field of radio will be interested in Fig. 5-26 which shows the radio development laboratory at 463 West Street (then the headquarters for the Engineering Department of the Western Electric Company) about 1922, i.e., towards the latter part of the period we have been discussing.

4.3 Transatlantic Telephony

To many in the telephone industry the postwar events just described must have seemed peripheral to the most challenging problem. Since 1915, when transcontinental telephony became a reality, and transmission of speech across the oceans was a demonstrated possibility, the next major goal in telephony had been the extension of commercial communication overseas.[37] Only thus could the Bell System objective of a universal worldwide service be realized. But the momentary transmission to the Eiffel Tower in 1915, and the subsequent commercial applications of radio at comparatively short distances, left a long way to go. The ship-to-shore and similar two-way communication systems went far to define and point up the "system" problems, which we shall return to later, but did little to solve the problem of furnishing a reliable telephone circuit over transatlantic distances.

When the transatlantic experiments were resumed in the early 1920s, there appeared to be only one region in the frequency spectrum capable of spanning the distance, namely, the region below about 100 kHz, and the initial work was carried out at these frequencies. Very soon the high frequencies (5 to 25 MHz) began to show promise and were studied extensively, but the main effort continued at the low frequencies, since they offered the best chance of early success, and was carried through

[36] See Section 4.3.5 for a description of the privacy problem and how it was handled.
[37] There was considerable doubt that enough traffic could be generated at rates adequate to support this obviously expensive system. However, Bell management, and in particular J. J. Carty, felt that the ultimate advantage of worldwide service would be very great and urged development if there was a remote chance of economic feasibility.

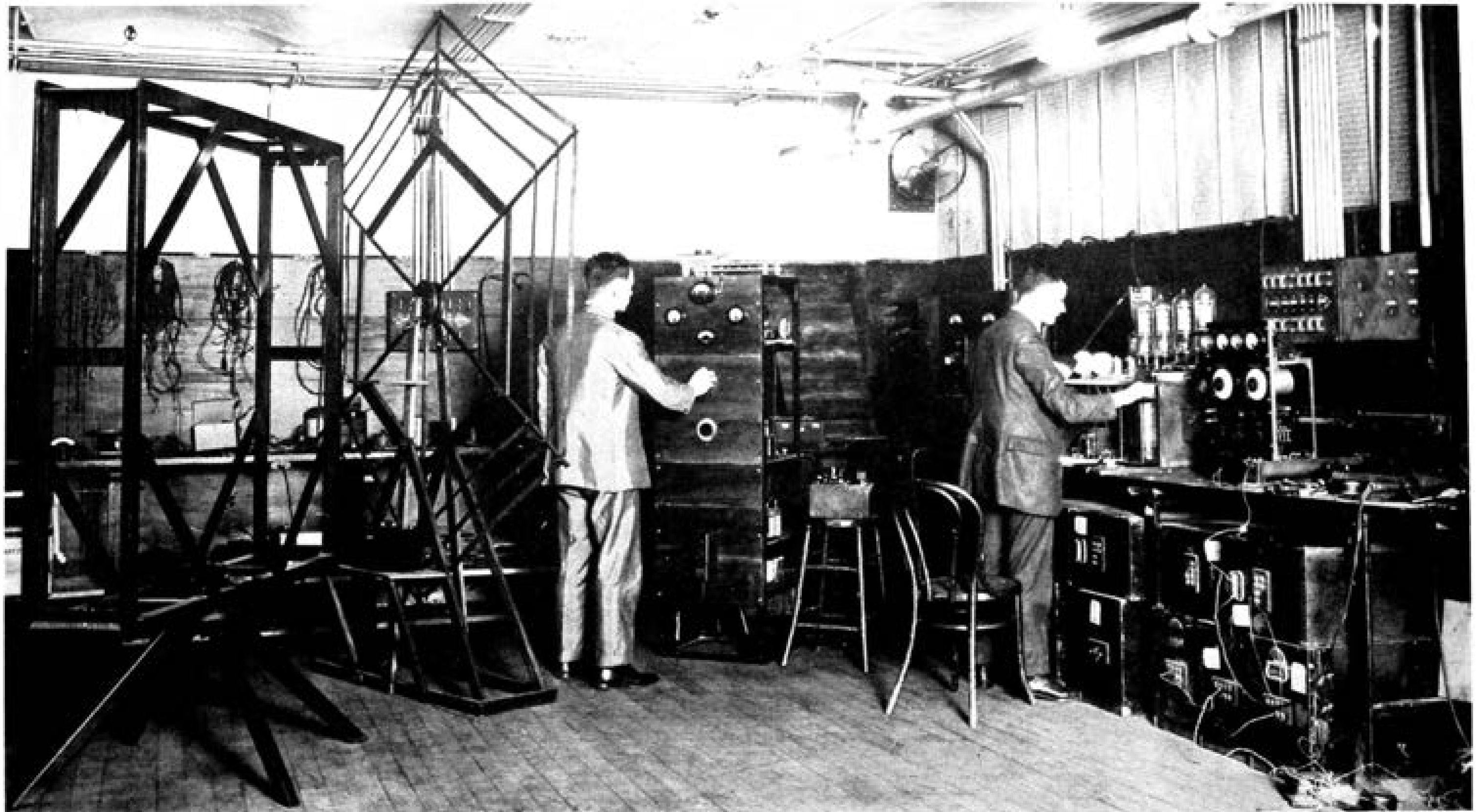


Fig. 5-26. Radio development laboratory at 463 West Street, New York City (headquarters for Engineering Department of Western Electric Company) about 1922.

to commercial application in early 1927. The parallel effort at high frequencies advanced very rapidly and by the middle of 1928 a high-frequency circuit was occasionally used for commercial service. Work in both areas continued for many years. Originally the high-frequency circuits were used only to supplement the low, but as techniques were refined, they proved capable of furnishing year-round service and ultimately provided the basis for the extensive overseas network built up in the 1930s and 1940s.

4.3.1 Low-Frequency Transmission—Technical Considerations

When the low-frequency work began in the early twenties, the most pressing problems centered on power. It was known that very large amounts of radiated power would be required, much more than had been used in 1915, but it was not certain precisely how much was necessary for year-round, 24-hour service. Thus, there were two separate power problems:

(i) Determining the amount of radiated power required for reliable telephony, and

(ii) Finding practical means for supplying this power.

Power requirements were difficult to predict because the radio medium, unlike wire (particularly cable), provided an unstable transmission path. Noise, commonly called static, varied greatly on both a long-term and short-term basis and was also greatly affected by the location of the receiving station. The wanted radio signal, which must at all times be large enough to override noise, was in itself subject to large diurnal and seasonal variations as the characteristics of the propagation path changed.

This problem existed in spite of years of transatlantic radiotelegraph service. Radiotelegraphy, as carried out at the time, was a much simpler operation than radiotelephony and required much less knowledge of the characteristics of the medium. Short-term variations in noise were not too important since some deferment of transmission, during periods of high noise, was tolerable and repetition of the message could be made whenever reception was doubtful because of momentary bursts of static. In addition, the experienced operators employed were skillful in adapting to the vagaries of the medium. Communication was carried out mostly at frequencies below 30 kHz and as low as 10 to 15 kHz, a range where there was reasonably constant wave propagation with only modest departures from the inverse-square law. Thus for telegraphy, the amount of power required was reasonably predictable and service could be adapted to meet conditions if the available power occasionally proved to be inadequate.

Commercial radiotelephony, if it was to be comparable to wire communication, had to provide usable communication between unskilled users, essentially on demand, and be reliable enough to provide a signal, for a large percentage of the time, that was sufficiently above noise to convey the message with few repetitions. The amount of radiated power required for such service could be determined only after a long and practically continuous series of observations of noise and propagation over a period of at least a year. (In actuality, these tests were continued for nearly ten years.) The problem was complicated by the fact that the more stable transmission frequencies, below 30 kHz, had already been preempted for telegraphy and could not be used for telephony.[38]

Some propagation tests had been made, at 800 kHz, in connection with the ship-to-shore experiments discussed previously. These had shown extreme variability between day and night and the necessity for using a much lower frequency to avoid excessive daytime attenuation. Accordingly, a much larger program was begun towards the end of 1922 covering both noise and propagation at about 60 kHz on a transatlantic path. More will be said about this work later when the Bell System pioneering measuring program is described. Suffice to say that transmission was found to be highly variable, as indicated in Fig. 5-27 which shows the range and general pattern of the first two months of 1923. The factor of importance is, of course, the ratio of signal to noise. This is shown on an average basis by month in Fig. 5-28. It will be noted that on the average there is a large diurnal variation and also a considerable seasonal variation. There are also large day-to-day variations around the monthly average curves. Unfortunately, nature in its perversity has chosen to provide some of the best transmission during the hours of darkness when telephone communication is at its ebb. Poorest transmission tends to occur during afternoon hours in the United States (early evening in Great Britain).

While preparations were being made for the long transmission testing program required, means were being devised for solving the power problem. It was already apparent that for commercial applications signal-to-noise ratios would have to be improved by many orders of magnitude as compared to the 1915 experimental setup. To accomplish this solely through the use of more vacuum tubes would have been completely impractical. The "tyranny of numbers," which we hear so much about in discussion of present-day complex systems, was even

[38] Even if not already in use, these frequencies would have been of limited value for telephony since this service required a transmission band of 3 to 5 kHz for each channel. Thus the favorable frequency band was only wide enough to accommodate at best a few telephone channels and the technical problem of devising antennas and other equipment providing the relatively wide telephone bands at these frequencies was very formidable at best.

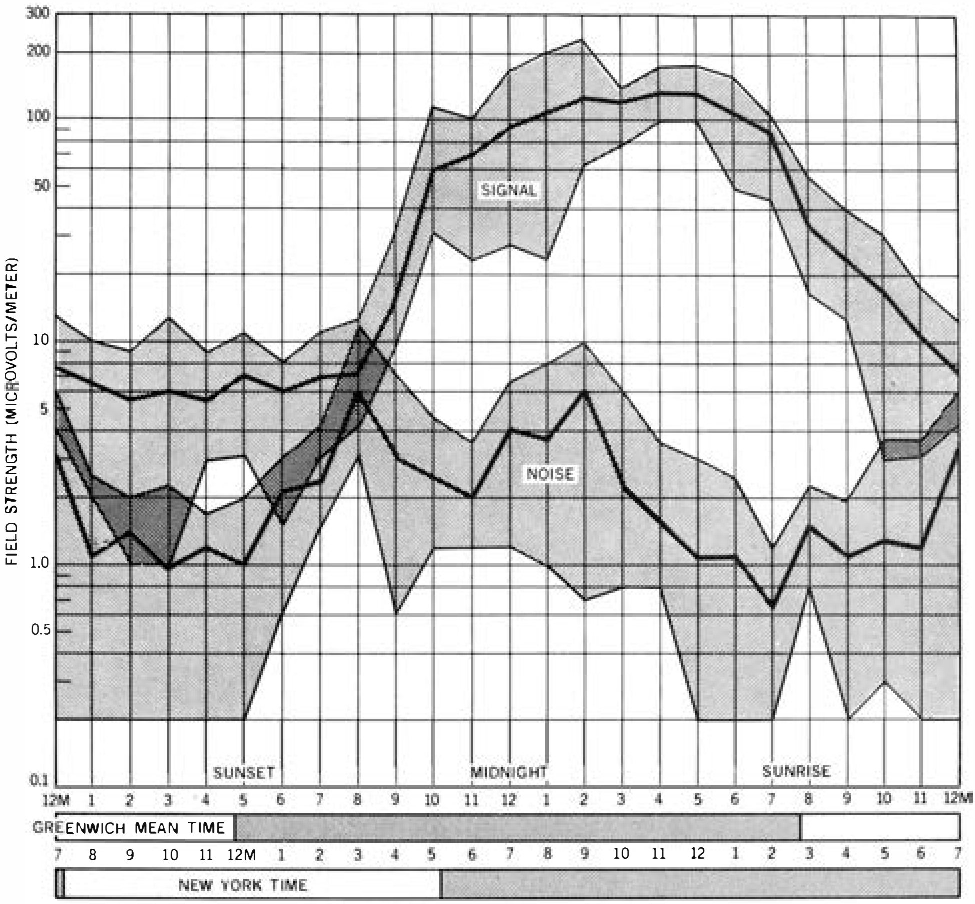


Fig. 5-27. Transatlantic radio transmission measurements: diurnal signal and noise variation, January 1 to February 23, 1923. Shaded area shows range of variation; heavy lines are average. Signal field strengths corrected to 300 amperes antenna current. (Redrawn from Arnold and Espenschied 1923, Fig. 9)

then an important factor. With the 500 tubes used in 1915, the expected failure rate was probably several per day and this would have increased manyfold with the greater numbers required for commercial service. Instead, three major steps were taken:[39]

(*i*) Development of very high power tubes so as to provide greater radiated power with only a modest number in parallel

[39] Numerous other improvements and changes were made which collectively were important but individually were far less significant than the major steps enumerated.

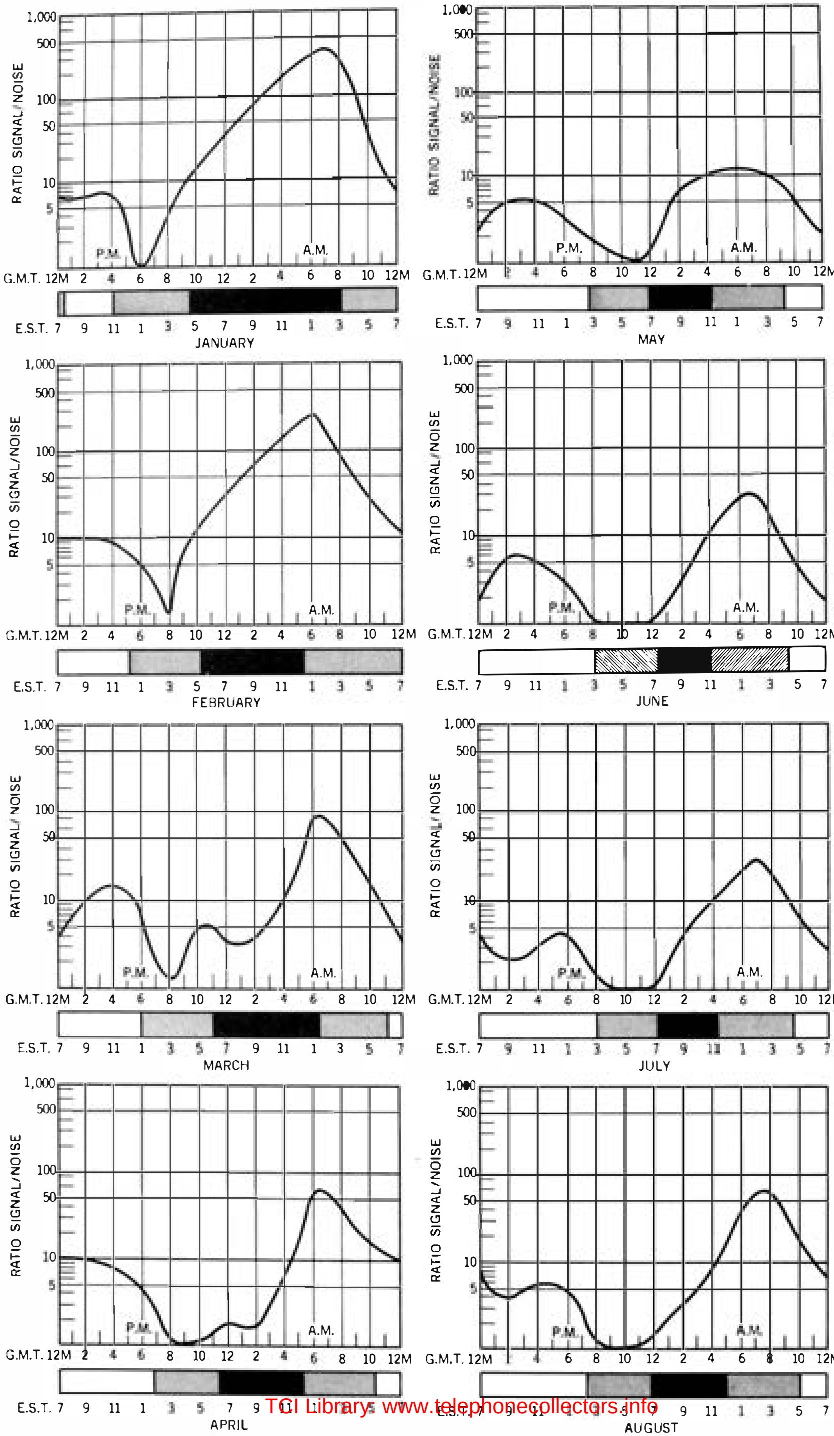


Fig. 5-28. Transatlantic radio transmission measurements: monthly averages of diurnal variations in signal-to-noise ratio for 1923. Transmission on 57 kHz. Curves corrected to 300 amperes antenna current. (Redrawn from Arnold and Espenschied 1923, p. 820)

(*ii*) Use of more efficient modulation schemes so that the radiated power could be used to greater advantage

(*iii*) Reduction in the noise picked up by the receiver.

A major problem in building high power tubes is the dissipation of heat. Even with well-designed circuits the amount of heat energy which must be disposed of at the anode (plate) of the tube is about the same as the useful power that is delivered to the antenna. During the postwar years, considerable advance had been made in increasing the power of conventional air-cooled tubes, as noted in a previous section (Section 4.2.1) discussing ship-to-shore transmission, but only a limited amount of heat could be carried away from a glass-walled tube by a stream of air.[40]

A much more efficient thermal system was required for very large power production and was provided in the early 1920s by the water-cooled tube shown in Fig. 5-29. In this tube the metal anode forms a container for the active portion of the tube and in operation is surrounded by circulating water. The upper portion of the tube is glass and serves to support and insulate the grid and filament structures. By using a cooling medium with high heat capacity in direct contact with the source of heat, a highly efficient system was obtained capable of dissipating over 10 kW of heat. Later, this same principle was used for larger tubes rated at 100 kW.

The three main difficulties which had to be overcome in constructing these tubes were, first, making a vacuum-tight seal between the copper anode and the glass; second, the provision of adequate means for conducting the large filament currents through the glass wall; and, third, obtaining the necessary vacuum for high power operation.

The first of these problems was solved in 1919 by the development by W. G. Houskeeper of a new metal-to-glass seal. In making this seal, the glass and metal parts are brought into contact while hot, the temperature being high enough for the glass to wet the metal. The part of the metal in contact with the glass is made so thin that the stresses which are set up when the seal cools are not great enough to fracture the glass or to break it away from the metal at the surface of the contact. Seals made in this way are sufficiently rugged to stand repeated heating and cooling from the temperature of liquid air to that of molten glass without deterioration.

A seal employing the same principle but different in form was also used at the point where the leads, carrying the large filament current, pass through the glass walls of the tube.

In exhausting the tubes, it was found necessary to subject all the metal parts to a preliminary heat treatment in a vacuum furnace

[40] The tubes used for the ship-to-shore work were rated at 250 watts whereas the 1915 tubes had a rating of 25–50 watts.

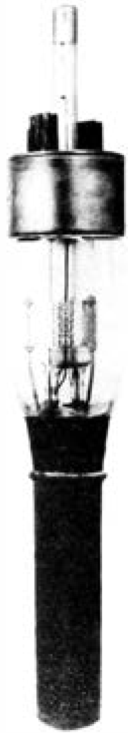


Fig. 5-29. Water-cooled vacuum tube 220B. (Arnold and Espenschied 1923, Fig. 6)

during which the great bulk of the occluded gases was removed. By this process the time of exhaust was considerably reduced but the vacuum conditions to be met were so stringent that the final process of evacuation often occupied as much as 12 hours.

The resultant tubes were operated at a plate voltage of 10,000 volts and were capable of delivering 10 kW at this voltage in a suitable oscillator circuit. When tubes are used to amplify modulated currents with the large peak values characteristic of telephone signals, the maximum electron current is several times the average operating current and therefore, to insure the necessary high quality of transmission, these tubes were operated for telephone purposes with an average output of about 5 kW. The nominal rating was usually referred to as 10 kW which corresponds to the maximum permissible plate power dissipation.

In practice, about 20 to 30 of these tubes were used in parallel giving a fiftyfold increase in power as compared to the hundreds of tubes used in the 1915 tests.

This large increase in power did much to meet the requirements for commercial service but further measures were required. One of these was to improve the efficiency of modulation by using the single-sideband, carrier-suppressed technique which has been described in detail in a previous chapter.[41] With the double-sideband, carrier-transmitted method used in previous radio experiments, about two-thirds of the total power radiated is concentrated in the carrier which conveys no intelligence. Thus a large increase in efficiency can be obtained by eliminating the carrier and concentrating the power in the intelligence-carrying sidebands. The transmission of both sidebands is unnecessary and undesirable since this use complicates the carrier resupply at the receiver and also requires twice the frequency band of a single-sideband system. The narrower single-sideband system practically doubled the number of channels possible in the very limited frequency band available.[42] In addition, the receiver band could also be narrowed with a considerable reduction in noise pickup. The combined benefit from single sideband was close to an equivalent power increase of tenfold as compared to earlier systems. The modulation efficiency was further improved by adjusting the speech from the wire lines to a constant volume, as discussed in Section 4.3.5.

Perhaps the largest improvement grew out of studies of noise sources and characteristics. At the low frequencies first used for transatlantic telephony, the principal noise source lies in a belt near the equator. Thus, locating the receiver as far north as practical considerably reduced the received noise (and also shortened the distance to England). Receiving at Houlton, Maine, for example, provided a combined improvement, as compared to New York, equivalent to a fiftyfold power increase, or roughly equal to the benefit from the high-power tube development. Furthermore, on the important transatlantic path the signal tended to arrive from a near northerly direction or almost opposite to the direction of noise arrival. Because of this situation, it was possible to utilize the Beverage wave antenna to discriminate between the arrival directions of the signal and noise. The benefit on the average from such an antenna was found to be equivalent to a power increase of about 100. Such antennas, because of the very long waves involved, are far from simple. The one at

[41] See Chapter 4, Section 4.2.5. The single-sideband technique was invented by J. R. Carson in 1915 and is widely used in wire transmission. Its application to radio is one of the many illustrations of the manner in which the two media benefited from common developments.

[42] Another important reason for using the narrower band lies in the difficulty of constructing a low-frequency antenna system which transmits a uniform band wide enough for double-sideband operation.

Houlton consisted of four parallel pole lines some distance apart, each about 3 miles long, the outputs being connected to the receiver by transmission lines and combined so as to yield an optimum signal from the overseas direction. (One of the pole lines is shown in Fig. 5-30.) Originally a two-element wave antenna was used for receiving in the United Kingdom but it was later replaced by an array of loop antennas when it was found that the ground conditions were not very favorable for wave antennas.

With these techniques available, which gave a combined improvement of some six orders of magnitude as compared to the 1915 experiment, commercial service appeared feasible.

4.3.2 Implementing the Low-Frequency System

Although much thought had been given to transatlantic radiotelephony for a number of years, the development of an overseas system did not get truly under way until early 1921. At this time a decision was made to develop the necessary high-power vacuum tube and studies were also started on the characteristics and benefits of wave antennas. In 1922 construction was begun of apparatus for measuring noise and signal strength at low frequencies and at the same time the use of the large RCA antenna at Rocky Point, Long

Fig. 5-30. Part of transatlantic receiving antenna at Houlton, Maine. (Page et al. 1932, Fig. 15)

Island, was being investigated as a possibility for transmitting to England. This antenna was a large, multiple-tuned "flat top," 400 feet high, consisting of 12 parallel wires each 1-½ miles long. It was supported on six steel towers at each of which leads were brought down through inductance coils to ground. This provided a sharply resonant radiation system with very little directivity but of high efficiency. It had been built by the Radio Corporation of America for transatlantic telegraphy, but the frequency band over which it radiated efficiently was made sufficiently wide for telephony by appropriate tuning and by equalizing the input circuits. It was ultimately used for all commercial low-frequency telephone service to Great Britain.

All of this preparatory work culminated in a demonstration of one-way transmission between New York and London on January 14, 1923. The radio circuit employed the high-power tube described previously in a transmitter which was a close prototype of that used later in commercial service. The Rocky Point antenna was used for transmitting and a large loop antenna for reception. Officials of the Bell System talked to British Post Office officals using land line extensions at each end. Speech was considered to be of good commercial quality, and the demonstration generated great enthusiasm. By March of the same year a decision was made to build a complete two-way system using Western Electric transmitting equipment both in England and in the U.S.A. The transmitter for the London end was completed and shipped in 1926, and by the end of that year the entire system was completed and public telephone service between the two countries was opened on January 7, 1927, with a charge of $75 for a 3-minute call.

The commercial service used the same frequency band, centered at 60 kHz, for both directions of transmission. The transmitting stations were at Rocky Point, Long Island, and Rugby, England, and the receiving stations at Houlton, Maine, and Cupar, Scotland. The transmitting antennas were nearly omnidirectional but tuned to give efficient radiation over the required frequency band. Initially the receiving antennas were of the Beverage type at both ends, as mentioned previously.

The transmitters were almost identical to the one used in the 1923 demonstration except for a slight increase in power. They consisted of three main parts as shown in Fig. 5-31:

(*i*) a low-power modulation-amplification system
(*ii*) a high-power two-stage amplifier
(*iii*) a three-phase power supply and rectifier.

Two steps of modulation were employed, both at low-power levels, and both using balanced modulators to suppress the carrier in the output. The voice band (200–3,000 hertz) first modulated a 33,200-hertz carrier and the lower sideband, of 33,000–30,200 hertz, was

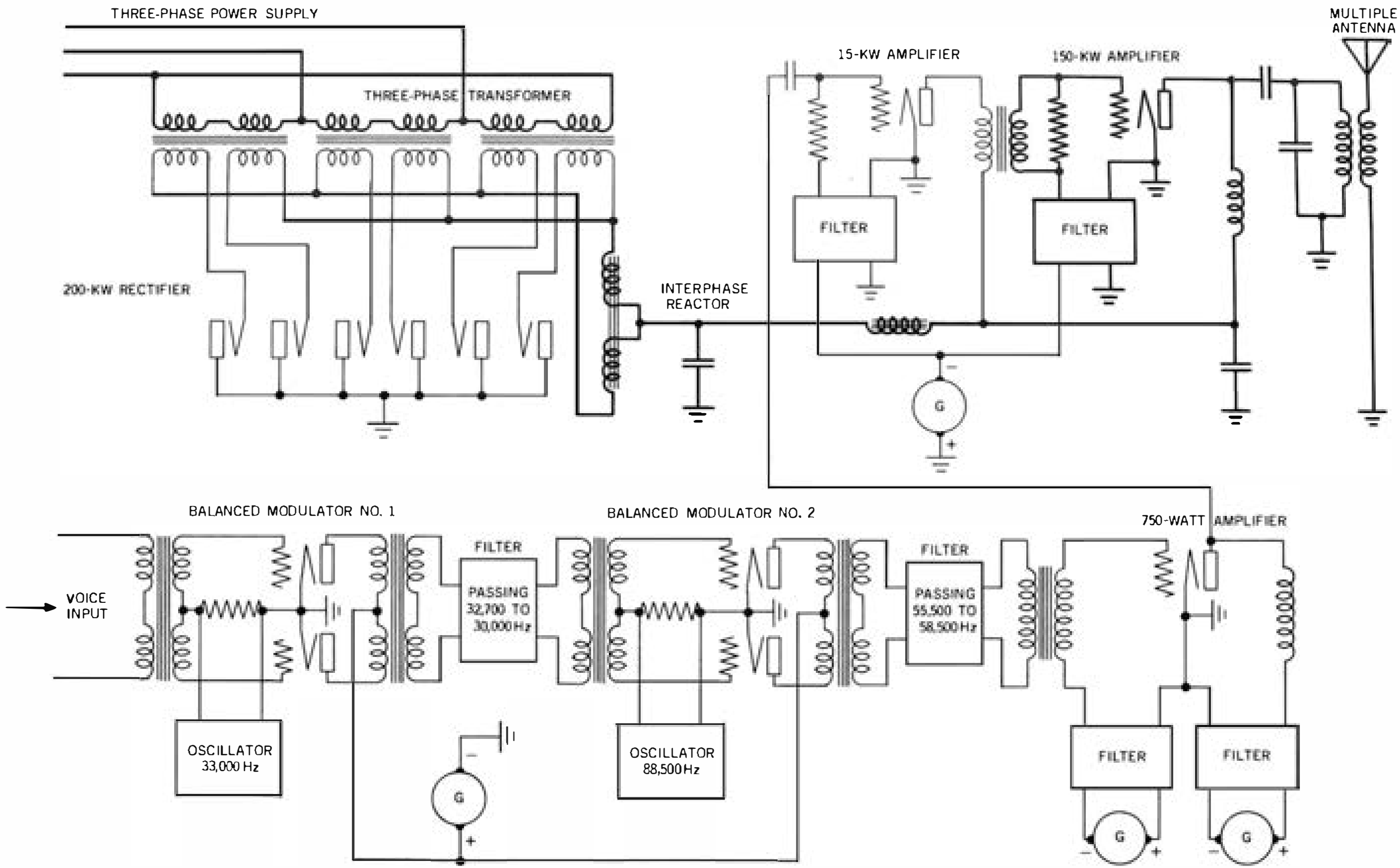


Fig. 5-31. Single-side band, carrier-eliminated transmitter for commercial transatlantic telephone service.

selected by a band-pass filter for modulating the second carrier of 91,700 hertz. The lower sideband gave the signal desired for transmission (61,500–58,700 hertz), the upper sideband and the unwanted carrier being easily filtered out since they were far displaced from the desired band.[43] The latter was amplified to about a 750-watt level by means of three parallel, air-cooled, glass-envelope tubes.

Power amplification involved two stages, each employing the water-cooled tubes mentioned in the previous section. The first stage used two tubes in parallel, and the final stage used as many as 35.[44] (The 1923 demonstration used only 20.) A multitube output amplifier (several were used in parallel) is illustrated in Fig. 5-32 which clearly shows the cooling-water duct coiled at the base. This cooling-water tube was made long to give a high resistance to ground and minimize the leakage of current from the anodes which operated at 10,000 volts. The complete amplifier assembly is shown in Fig. 5-33.

The power rectifier used water-cooled diode vacuum tubes otherwise similer to the three-element amplifiers. Full-wave rectification of a three-phase supply gave an output from which the ripple currents were easily removed by filters.

The receiver also used a double-modulation scheme, the now familiar superheterodyne principle, with intermediate-frequency-band and beat-oscillator frequencies the same as used in the transmitter.

This equipment performed well during many months of the year, but during the summer the level of static noise frequency became too high for commercial daytime service. There seemed to be two ways in which this could be corrected; one was to increase the signal-to-noise ratio and the other was to move into the high-frequency band which usually performed well during this season. The latter approach ultimately proved the practical one for transoceanic telephony, as we shall shortly see. However, in the late 1920s the characteristics of the high-frequency band were not clearly understood and some years of laborious measurement were required before the capabilities of this band could be defined. It was decided, therefore, that further development of transatlantic services would follow a two-pronged approach. One was the development of a new and more adequate low-frequency system and the other was the continued study of the capabilities of the high frequencies. At the time, it was believed that

[43] The double-modulation scheme not only facilitated the removal of unwanted frequencies but made it easy to shift the transmitted band by changing the frequency of the second carrier. Because of the wide separation of the wanted (lower-sideband) signal and the unwanted frequencies, moderate frequency changes could be made without changing the filters. The 1923 demonstration used essentially the same transmitter but an 88,500-hertz second carrier giving a transmitted band about 3,000 hertz lower.

[44] This was commonly referred to as a 150-kW amplifier but the peak envelope power was about 200 kW.

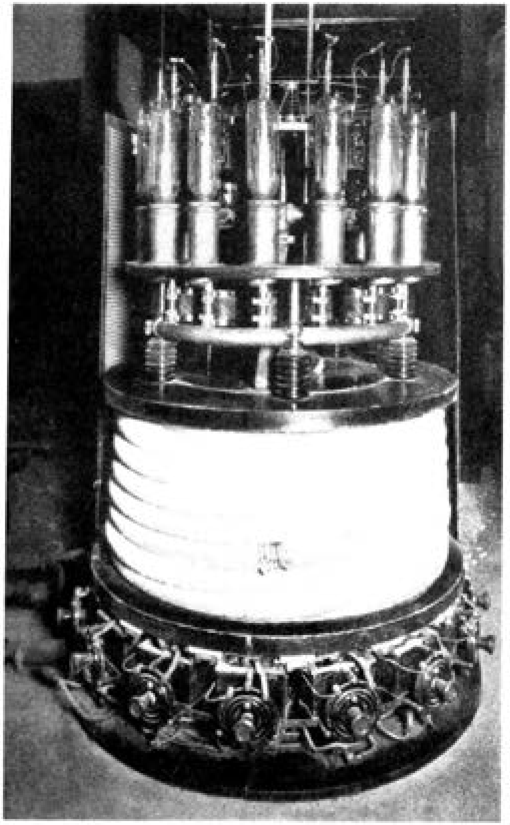


Fig. 5-32. Large vacuum tubes used in the last stage of the Rocky Point long-wave radio transmitter.

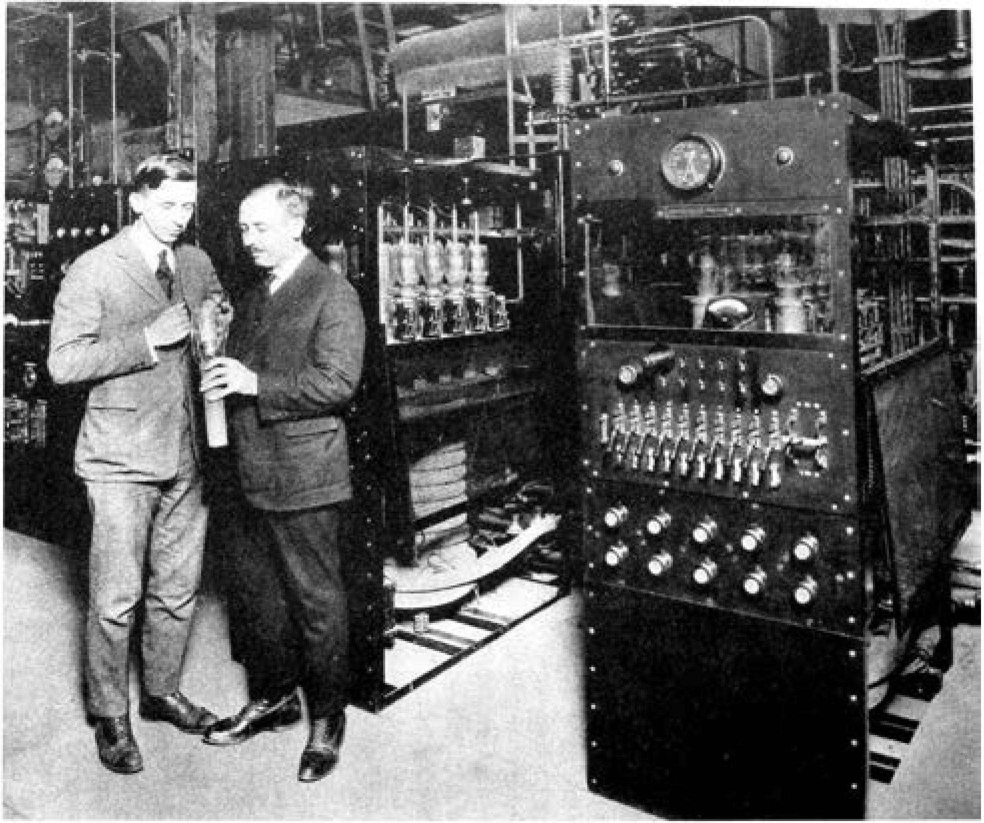


Fig. 5-33. Final amplifier of 150-kW capacity. Pictured are J. C. Schelleng and A. A. Oswald, authors of the classic paper, "Power Amplifiers in Trans-Atlantic Radio Telephony," *Proc. IRE*, Vol. 13, June 1925, pp. 313–361.

the latter work would at least provide a useful supplement to the low frequencies and might in the long run have broader applications, as it eventually did.

Planning of the new low-frequency system started in early 1929 and by June 4 the outlook appeared sufficiently promising to authorize development. The main changes planned were in the transmitting system which would use a final amplifier stage employing six 100-kW tubes in push-pull parallel feeding a broadband wave antenna which could handle several channels simultaneously. The transmitting station would be located in Maine, thus bringing it several hundred miles closer to Scotland and reducing the path loss accordingly. Two channels, both working from the same antenna, were planned. The first to be constructed would be centered at about 68 kHz and supplement the existing channel. The second, at 60 kHz, would ultimately replace the Rocky Point channel. Work on the amplifier and power supply progressed rapidly with a full-scale experimental model operating at the Whippany Laboratory by the middle of 1930. Also, by that time, the practicability of the transmitting wave antenna had been demonstrated through tests made in Maine with a low-power transmitter. Unfortunately the stock market crash of 1929 and the subsequent business depression raised serious question as to the immediate future of transatlantic telephony. Accordingly, work was discontinued in November 1930 except for completion of certain tests in progress which continued until March 1931, at which time the completed work was documented[45] and further effort discontinued until conditions were more favorable. This time was never to arrive since the shortwave work covered in the next section had, in the meantime, demonstrated that this band could not only supplement the long waves during periods of adverse static but could provide the basis for a complete worldwide transoceanic system.

However, the low-frequency system continued to be of value as a backup during periods of magnetic storms and in 1939 a second channel operating at 68 kHz was added as an "applique" to the Rocky Point system. This duplicated the low-power portions of the 60-kHz system and the two channels were multiplexed and used the high-power amplifiers in common. Appropriate modifications in antenna tuning were made so the two channels could be handled. Two-channel operation was obtained with some loss in power on each channel. This system was continued in use or on standby until 1957 when it was dismantled.

[45] Internal report, June 11, 1931, R. A. Cushman and W. H. Doherty, "Engineering Report of the Development of the High Power Amplifier and Antenna Coupling Circuits for the Two-Channel Long-Wave Transatlantic Radio Telephone System." This report includes the first known theoretical treatment of intermodulation distortion in single-sideband transmitters.

4.3.3 High-Frequency Transmission—Technical Considerations

We shall now cover the short-wave work that so greatly changed the outlook for overseas communication, but before doing so we should remind the reader that the work described both here and in the preceding section covers only the provision of one-way point-to-point radio paths. For commercial telephony it is necessary to combine two of these unidirectional channels into a two-way system capable of working with switched, wire telephone extensions on each end of the radio path. The means for doing this is very much the same regardless of the radio band employed and will be covered after discussing the development of high-frequency radio channels.

As we have noted, until the early 1920s it appeared that the only frequencies which could possibly be used for transatlantic service were those below 100 kHz (3,000 meters wavelength). Above this frequency the attenuation of surface radio waves was so great that only relatively short distances could be covered with practical amounts of power. As frequencies were reduced below this point, the attenuation was reduced and in the 15- to 30-kHz region used for long-distance telegraphy the propagation approached the inverse-distance law even for transatlantic distances.

In 1912, when the radio amateurs were assigned to the frequencies above 1,500 kHz (200 meters), it was believed by many that the band was so useless, because of the limited power and distance capabilities, that amateur radio would soon disappear entirely. In spite of these difficulties, and the complete ban on amateur radio during the war years, amateur interest refused to go away. By the early 1920s a few determined amateurs were working not only at 1,500 kHz but at twice this frequency. To the surprise of everyone, transmission was occasionally obtained to great distance and in November 1923 transatlantic transmission was accomplished on three successive nights. Within a year, signals from Australia were heard in California over a distance of about 7,000 miles.

During this same period the English Marconi Company was investigating the capabilities of the high-frequency band for commercial service. In 1921 they had achieved duplex telephony at 3 MHz between England and Norway. Two years later they built an experimental short-wave station in Poldhu, England, which included a directional "beam" antenna with parabolic reflector. In 1924 this station operated just above 3 MHz. Later that year, experiments at frequencies up to about 10 MHz showed improved daytime range. The following year, frequencies as high as 150 MHz were investigated and by 1926 Marconi opened commercial telegraph service to Canada using frequencies of about 10 and 20 MHz.

In the middle twenties it was apparent that there were some interesting possibilities in this frequency region. The important question concerned the reliability of such transmissions. Were they reproducible and dependable enough for commercial service, or were they unpredictable and only suitable for communication on an occasional, unscheduled basis? In 1923 and 1924 when the sporadic long-distance, high-frequency transmissions were first obtained, the Bell System had already demonstrated good low-frequency transmission and, through a comprehensive and highly successful measurement program, had provided the data on noise and propagation needed for engineering commercial service. Undoubtedly the main effort had to be devoted to the long-wave development if the earliest possible service was to be achieved. But it was also wise to see what the high frequencies offered. This meant not only development of equipment and antennas for this previously unused band but a measurement program even more extensive than that carried out at the low frequencies.

During the war years and in the early twenties, the Bell System had developed short-haul airplane and shipboard equipment operating at frequencies as high as 3 to 4 MHz. While this was far from meeting the needs of transatlantic service, it did provide some insight into the problems of high-frequency equipment and inspired some confidence that they could be solved. By the middle of 1925 a transmitter using water-cooled tubes had been built at the Deal Beach laboratory, the output consisting of two 5-kW tubes in push-pull configuration. At first this was only capable of operating at the lower end of the high-frequency band but the operating frequency was gradually pushed upward until 20 MHz was achieved about the middle of 1926.

No adequate quantitative information was available on radio propagation at these frequencies. Nichols and Schelleng of Western Electric had made some preliminary studies in 1924, but no comprehensive work could be done without a high-power transmitter. Such a program was started as soon as the Deal transmitter was available. Initial tests were over land from Deal to various points to the north, south, and west, and later they were extended to a ship traveling between New York and Bermuda. In 1925 weekly measurements began on transmission to England. The initial work covered a period of nearly two years, but similar work went on for many more years and ultimately covered not only North Atlantic paths but also paths to other parts of the world.

This work showed that as the distance was increased the high-frequency signals disappeared rapidly in accordance with previous experience, but they reappeared at high intensity at distances of several hundred miles and continued at a high level with little attenuation for considerable distance thereafter. The performance depended greatly on frequency, time of day, and season of the year.

Shortly after low-frequency service was inaugurated, enough data was available to show that the high frequencies would be useful, at least as a supplement to the low frequencies, and development of commercial equipment was started. Before we outline this work, it will be worth comparing the capabilities of the two frequency bands.

The low frequencies had four major disadvantages:

(*i*) Only a narrow frequency band was available and consequently it would be impossible to implement more than a few telephone channels.

(*ii*) Static noise was severe, particularly during summer on the North Atlantic path, and was present essentially at all times on paths crossing the equator.

(*iii*) The high noise and large path attenuation required very high power. Power levels available at that time practically prohibited transmission to the Southern Hemisphere or for distances much beyond three to four thousand miles in the Northern Hemisphere.

(*iv*) Antenna directivity, which would ameliorate the noise and attenuation problem, was attainable only with very large and expensive structures.

The high frequencies overcame most of these basic faults. A band about 20 MHz wide was available instead of the 40 to 50 KHz in the low-frequency region. Static noise was low throughout the band and negligible at the upper end. Power requirements were modest and, finally, and perhaps most importantly, reasonably small directive antennas could lower interference and greatly reduce the transmitting power requirements by concentrating the power on the desired path. While the high frequencies seemed to solve the main problems being experienced at low frequencies, they were not without their own peculiar disadvantages. New forms of interference such as automobile ignition and noise generated within the receiver were important and had to be guarded against. Not all frequencies were usable at all times. Typically, a given frequency would be usable for a few hours and then have to be replaced by another, as shown by Fig. 5-34 which also provides a low-frequency comparison. Similar changes occurred seasonally. The signal was not constant but tended to fade rapidly due to multipath transmission. When the paths differed greatly in length, the fading was selective, that is, it occurred in narrow bands that swept through the voice band, causing a peculiarly unpleasant distortion. Finally, service could be disrupted completely for days at a time as a result of changes in the earth's magnetic field (magnetic storms) which were related to sunspot activity as shown in

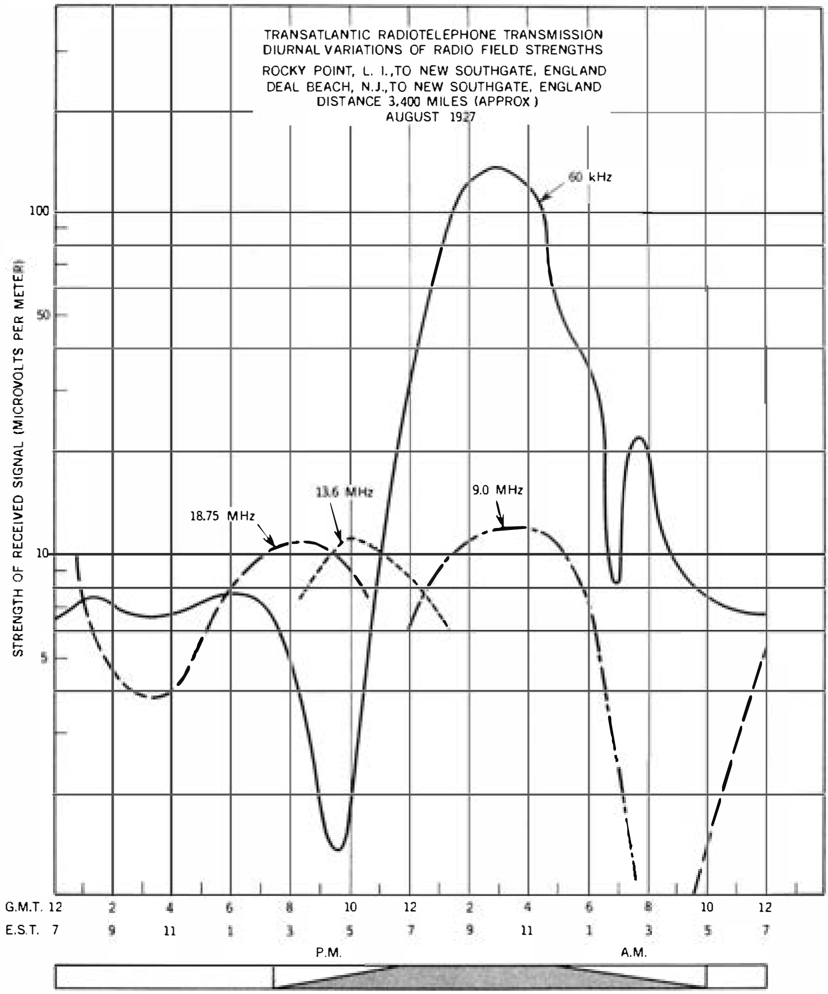


Fig. 5-34. Comparison of transmission at low frequency and various high frequencies.

Fig. 5-35.[46] As indicated in the figure, long-wave transmission was entirely unaffected by such occurrences.

4.3.4 Implementing the High-Frequency System

Early in 1927 an experimental one-way high-frequency channel was set up with the transmitter at Deal, New Jersey, and the receiver at

[46] Transmission tended to be particularly erratic when the path passed through the auroral belt in the Far North.

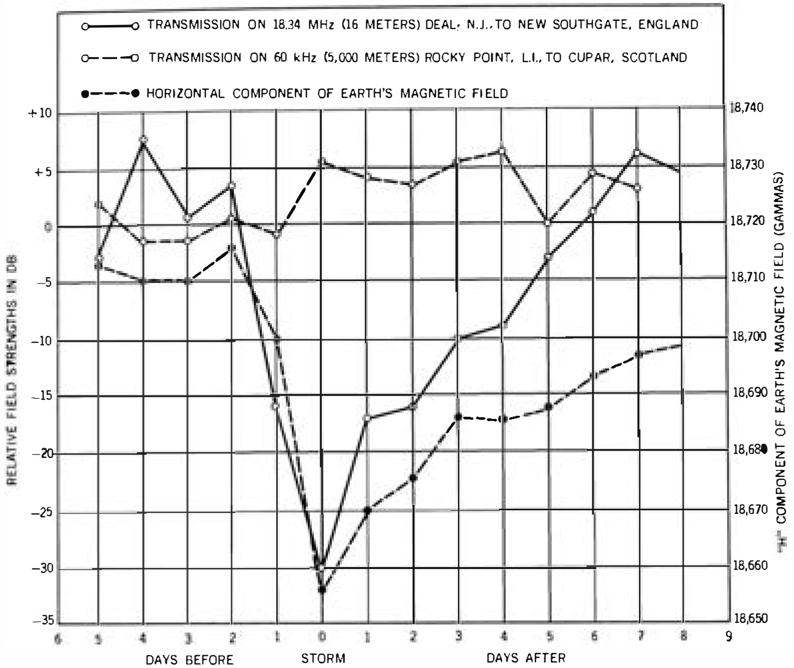


Fig. 5-35. Effect of magnetic disturbances on radio transmission.

New Southgate, London. During the summer of that year, when the static difficulties were severe on the long-wave system, this channel was used as an emergency commercial channel to England with the long-wave Rugby-Houlton channel used for the return direction. Early in 1928 an east-to-west short-wave channel began experimental transmission from Rugby to Cliffwood, New Jersey, and by the middle of the year the experiments had progressed to the point where a two-way circuit could be set up and used occasionally for commercial service when long-wave transmission was impractical. In June 1929 two commercial two-way circuits were put into operation between England and the United States. In the United States the transmitting station was set up at Lawrenceville, New Jersey, and the receiving station at Netcong, New Jersey. These sites were selected to minimize man-made noise and interference from transmitting stations of their own and other communication agencies. In addition, the locations would permit their ultimate use for service to South America as well as Europe; such service was inaugurated to Buenos Aires in 1930. In England the transmitters and receivers were located at Rugby and Baldock, Hertfordshire, respectively.

The equipment at Lawrenceville and Netcong in 1930 was typical of that used in the early years of high-frequency commercial service. At this time the stations were equipped for four circuits, three for New York–London service and the remaining one for New York–Buenos Aires service. The low-frequency system to London was also in use but, since the performance was about equal, the high-frequency systems were favored where a choice was possible because the cost of operation tended to be lower.

Each of the four channels had its own transmitter, for which the frequency could be changed to meet diurnal and seasonal requirements, and its own power plant and antenna system. The transmitter schematic is shown in Fig. 5-36. The telephone speech, suitably amplified and monitored at a control desk, was used to modulate the plate voltage of an oscillator consisting of two 250-watt air-cooled tubes in push-pull. This oscillator operated at the carrier frequency to be radiated, the frequency being closely controlled by quartz crystals in temperature-regulated ovens. To facilitate manufacture, low-frequency crystals were used to operate at a submultiple of the main oscillator frequency. Harmonic generators multiplied this basic frequency as required. Each transmitter was capable of working over the range of 9 to 21 MHz, the operating frequency being determined by an appropriate selection of inductances, capacitors, and crystals. From three to five preselected frequencies were assigned to each transmitter and frequency changes could be made in a matter of minutes. The modulated signal was fed to a two-stage power amplifier using water-cooled tubes with 10,000-volt plate supply. The first stage contained two tubes and the second six, the carrier output from the latter being 15 kW (60-kW peak power for 100-percent modulation).

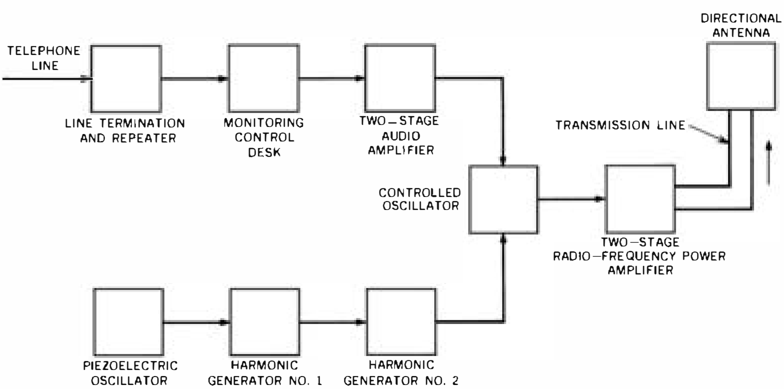


Fig. 5-36. Block schematic of transmitting system.

This system, unlike the long-wave system, employed two sidebands with transmitted carrier. The benefits of single sideband were appreciated, but the power and bandwidth problems were far less severe in the high-frequency region and the double-sideband approach was adopted to minimize other problems. However, with the low-level modulation scheme employed, it was possible to convert the transmitters to single sideband with no appreciable modification of the power amplifiers which represented the major portion of transmitter cost. In 1933 the experimental use of single sideband was begun and provided not only a roughly tenfold improvement in signal-to-noise ratio but also a reduction in distortion from selective fading. Commercial application began about two years later and it is used almost universally today.

As we have noted, the short wavelengths of the 10- to 20-MHz band made the use of highly directive antennas practical. A very interesting type of vertical screen antenna, illustrated in Fig. 5-37, was used at Lawrenceville.[47] It consisted of a single conductor folded so as to form several loops in the same plane. The antenna currents, shown by the arrows, were at any instant in phase in the vertical members whereas they tended to cancel in the horizontal. This arrangement multiplied the radiated power at right angles to the plane of the antenna by a factor of four as compared to a single-element radiator. A similar system parallel to and ¼ wavelength behind the main radiator was used as a reflector to reduce backward radiation and double the power in the desired direction. The basic plan was further extended by adding vertical and horizontal elements and by so doing power radiation was achieved at Lawrenceville that was as much as 50 times that of a single half-wave radiator. Sleet and ice greatly affected antenna performance and provision was made for melting it by the use of 60-hertz power applied through a sort of composite arrangement which permitted simultaneous thawing and antenna operation. These antennas, while highly directive, were effective for only a very narrow frequency band so that it was necessary to use separate structures for each transmitted frequency. Later, the non-periodic rhombic antenna developed by E. Bruce and H. T. Friis in 1930 was adopted since it gave adequate gain over a wide range of frequencies.

The receivers were of the conventional superheterodyne type with a 400-kHz intermediate frequency. Automatic gain control, actuated by currents from the intermediate-frequency output, was used to maintain a reasonably constant output signal despite the large variation in input caused by fading. Each receiver operated at three fixed frequencies

[47] Sometimes called the Grecian Key or Bruce Array type. It was developed at the Cliffwood laboratory.

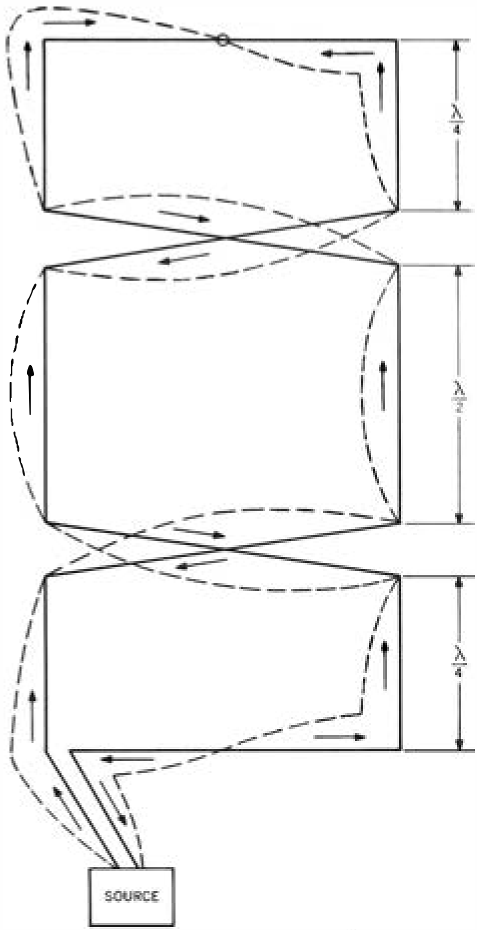


Fig. 5-37. Conductor bent to form one section of simple directive antenna. This was the type used for transmitting at Lawrenceville.

between 9 and 21 MHz from three separate antennas each connected through its own transmission line and permanently tuned preamplifier. The preamplifier outputs were switched to a tunable radio-frequency input amplifier as required. Final adjustment of the beat oscillator completed set tuning.

The receiving antenna was similar to that used for transmitting but had only a single row of vertical elements since the vertical angle of arrival of the incoming wave varied throughout the day and a high degree of vertical directivity was not desirable unless means were provided to adjust for the angle of arrival. (This was accomplished by the Multiple Unit Steerable Antenna developed in 1935.) An interesting feature of the Netcong receiving installation was the use of air-insulated coaxial-type transmission lines connecting the antennas and receivers. These lines, using 3/16 and 5/8 inch diameter conductors, not only provided low attenuation but also effective shielding against local interference.

4.3.5 Solving the System Problems

Transoceanic radiotelephony required the solution of two quite different technical problems. The first was providing dependable radio channels across the ocean, and the second was making these channels function as a link between the extensive wire telephone networks at each end. The first of these, the radio link problem, was successfully solved in the manner just described. To many readers this may seem to be the more glamorous of the two problems since it represented such a large quantitative advance by extending the limits of useful transmission from distances of tens of miles (as achieved in the war years) to thousands of miles. However, to the telephone technician the second problem of incorporating isolated radio links into a working telephone system was no less challenging and required no less ingenuity in its solution. These are the problems which we shall now discuss.

Radio channels are unilateral circuits transmitting in only one direction, whereas telephony requires to and fro communication. In the twenties much telephone communication was on a 2-wire basis, i.e., both directions being carried on the same pair of wires, and all switching and interconnection of circuits was at 2-wire points in the network. Therefore, a major system problem required the interconnection of the basically 2-wire telephone circuits with the basically 4-wire radio channels. This was precisely the problem faced when amplification and, later, carrier transmission was introduced into the wire plant. The analogy is very close since the radio channels were a combination of amplifying and frequency-translation equipment fundamentally the same as used for carrier-current transmission on wire. To some extent the same technique involving the hybrid coil[48] could be used for 2-wire to 4-wire conversion, but radio transmission was sufficiently different from wire transmission to introduce additional problems (Fig. 5-38).

A principal problem in using a circuit such as that shown is the prevention of singing, or oscillation resulting from the feedback of high-level power from the transmitter output into the low-level receiver input. With wire circuits the main cause of singing is high gain

[48] See Chapter 4, Section 4.2.4, for a discussion of the 2-wire to 4-wire conversion technique used in the wire plant.

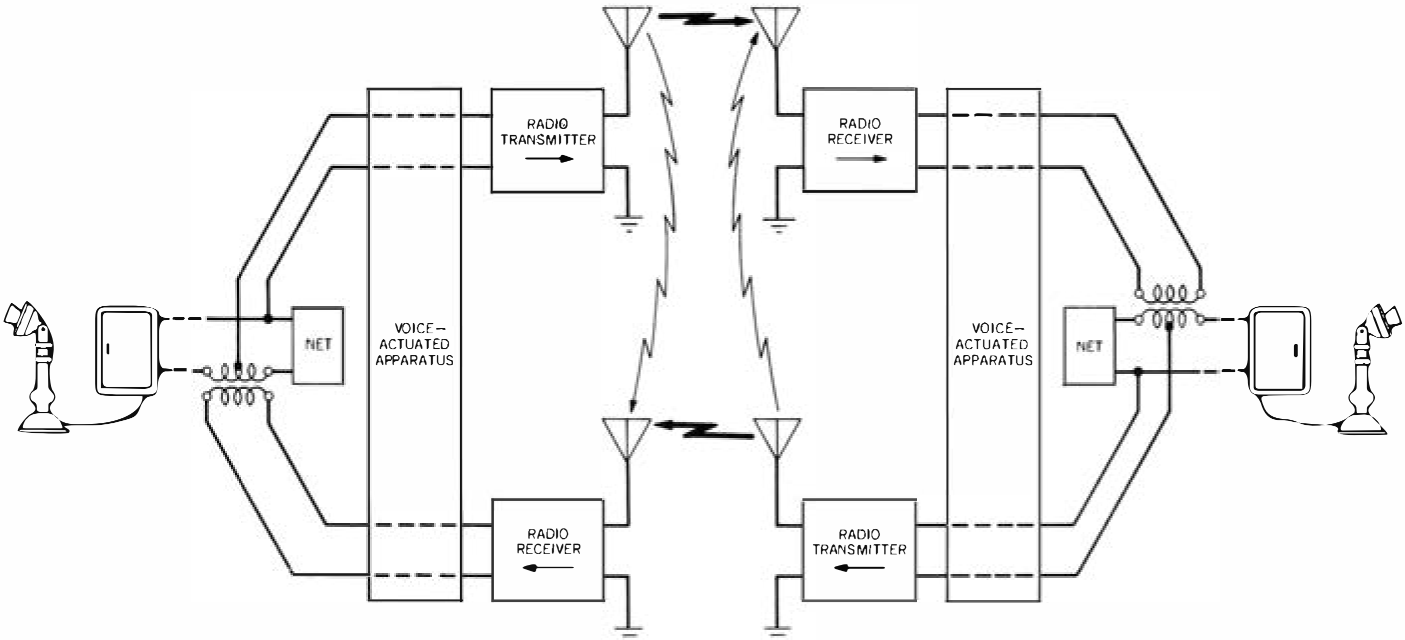


Fig. 5-38. Block schematic of transoceanic circuit. (Redrawn from Blackwell 1928, Fig. 13)

around the loop (as shown by the heavy arrows in the figure), i.e., energy from the west, after amplification, is transferred across the east hybrid coil owing to poor balance and returned to the sending end where it finds its way into the transmitter input by way of imperfect balance in the west hybrid. With wire circuits such singing is easily controlled by keeping the loop circuit gain well below the losses across the hybrid coils. (This is also necessary for echo control.) This is not readily done with radio since the transmission path between transmitter and receiver is highly unstable, particularly at high frequencies. A circuit adjusted for stability under average conditions may, therefore, easily sing when the path loss is momentarily reduced during the upswing of a fade.

Another possibility for singing occurs as a result of transfer of energy from the transmitter output to the input of its own receiver (as shown by the light arrows in Fig. 5-38). This type of coupling is ordinarily unimportant on wire circuits but can be of considerable importance with the high amplification used on radio. The transfer can be greatly reduced by using different frequencies in the two directions, by locating the transmitting and receiving stations at a distance, and by using directional antennas located so as to give poor transmission between them. These measures were used successfully in the ship-to-shore experiments and on the Catalina circuit, but the problem is more severe in the case of transatlantic service. The path loss across the Atlantic is so great that power amplifications of about 100 million (80 dB) are required in the transmitter and the receiver (160 dB total). With these large amplifications it is impossible to reduce the energy received from the local transmitter to the point where hybrid balance is sufficient to control singing. This was particularly true in the case of the low-frequency system which used the same frequencies for both directions of transmission and had only moderate antenna directivity.

The singing problem was solved by using voice-actuated apparatus (located as shown in Fig. 5-38). This was arranged so that when there was no speech present, the transmitter input at each end was short-circuited, making both transmitters inactive and opening all singing paths. As soon as speech was detected at one end, the transmitter at this end was enabled and the receiver disabled, thus providing a through path to the far end but keeping the singing path open. The essential features of the voice-controlled equipment, known as a VODAS (*V*oice *O*perated *D*evice *A*nti *S*inging[49]), are shown in Fig. 5-39.

[49] This seems to be one of the first outbreaks of the acronym disease which has since grown to pestilential proportions. The apparatus can take several forms and over the years numerous refinements were added, but Fig. 5-39 illustrates the basic principles. The resemblance to the echo suppressor, discussed in Chapter 4, Section 4.2.4, will be noted. It was indeed a rather natural outgrowth of this device and illustrates again the close relation between developments in the wire and radio fields.

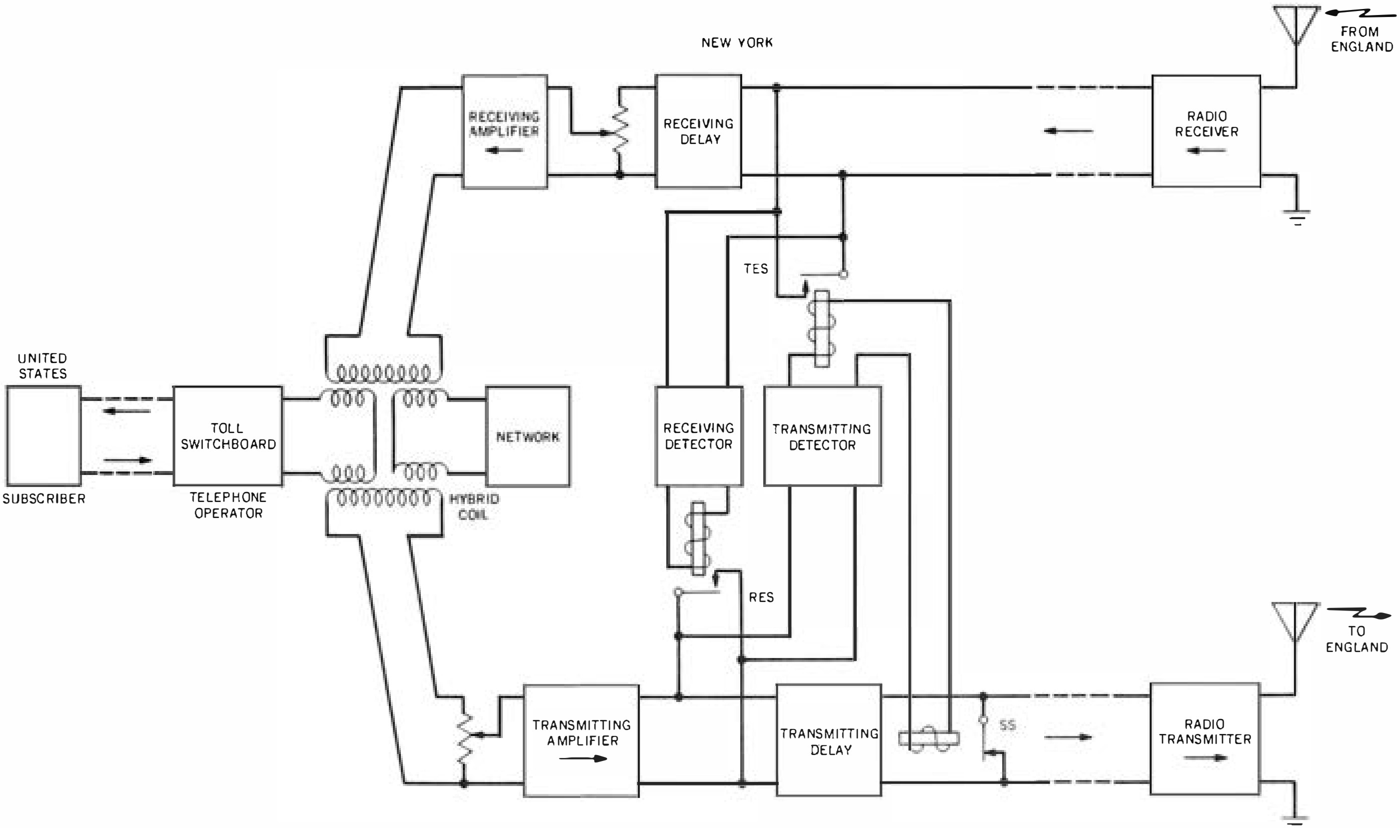


Fig. 5-39. Circuit diagram illustrating operation of voice-operated switching device. (Redrawn from Bown 1930, Fig. 1)

The functioning of the apparatus is briefly as follows: The relay TES is normally open so that received signals pass through to the subscriber. The relay SS is normally closed to short circuit the transmitting line. When the United States subscriber speaks, his voice currents go into both the transmitting detector and the transmitting delay circuit. The transmitting detector is a device which amplifies and rectifies the voice currents to produce currents suitable for operating the relays TES and SS which thereupon short circuit the receiving line and clear the short circuit from the transmitting line, respectively. The voice currents are delayed a few hundredths of a second in the delay circuit so that, when they emerge, relay SS has had time to operate and clear the path ahead. (Without this feature, initial speech sounds would be "clipped" because of the slow response time of the relay.) When the subscriber has ceased speaking, the relays drop back to normal.

The function of the receiving delay circuit, the receiving detector, and the relay RES is to protect the transmitting detector and relays against operation by echoes of received speech currents. Such echoes arise at irregularities in the 2-wire portion of the connection and are reflected back to the input of the transmitting detector where they are blocked by the relay RES which is closed and which hangs on for a brief interval to allow for echoes which may be considerably delayed. In this way, transmission around the loop is prevented and singing due to momentary low path losses is avoided.

The VODAS solved the major problem of adapting radio channels to the telephone plant but there were other differences between wire and radio transmission that required special treatment. One of these was the variability of the radio path loss. On the low-frequency circuit these variations were relatively slow and were taken care of by monitoring the incoming signal and adjusting received amplification as necessary to maintain a reasonably constant voice-frequency loss between the transmitter input and receiver output. At high frequencies, these variations were very rapid and, as indicated in the previous section, were taken care of automatically by means of an automatic gain control circuit. In this way the vagaries of the transmission medium did not result in enormous variations in received speech level. (But the variation in receiver gain was manifested to some extent as a change in received noise.)

Variability was not confined to the radio link alone. The speech volume presented to the radio system by the wire end-links varied greatly from call to call. (The total range in speech power was as much as several hundred to one, i.e., 20–25 dB.) Some of this variation came from the difference in length and loss of the wire circuit, but most of it resulted from different talker characteristics. Some people had loud, powerful voices and others were low-level talkers. This variation complicated the design of voice-operated devices such as the VODAS and, more importantly, if not corrected, would have greatly lowered the transmitter efficiency. Enough transmitter power capacity had to be supplied to handle the very loudest talker but with normal speech the transmitter would be operating at only about one-tenth of the maximum capability. This situation was corrected by adjusting all incoming speech levels to about the same level (within a few dB) as that of the loudest talker. In this way the transmitter modulation was maximized at all times. Initially, these adjustments (as well as the control of VODAS sensitivity) were done by a technical operator who monitored the circuit with the aid of a volume indicator (see Chapter 4, Section 5.1.4.3). Later (about 1938) the adjustment was made automatically with a VOGAD (*V*oice *O*perated *G*ain *A*djusting *D*evice). With this change and with other improvements in control equipment, it became unnecessary to have a technical operator on each circuit.

The adjustment of talker volume to give essentially constant (and optimum) modulation of the transmitter resulted in a circuit with variable net loss.[50] For a typical circuit the net loss might be about 6 dB for a loud talker whereas a weak talker would result in a circuit adjusted for 20-dB higher amplification, i.e., for a net *gain* of 14 dB. Such a high gain would cause very bad echoes and possibly singing around the loop on a conventional wire circuit but was made possible on radio circuits by the protection afforded by the VODAS.

The manual adjustment of speech volume and the VOGAD tended to equalize talker volume on a long average basis but large variations still occurred at a syllabic rate as the speech sounds varied from weak consonants to strong vowel sounds. To partially compensate for this variation, the compandor (contraction for compressor-expandor) was invented by R. C. Mathes and used in the low-frequency circuit in 1932. This device added amplification at the transmitting end on low speech sounds (compressed the range of speech levels usually to one-half of the input range) and at the same time added compensating loss at the receiving end (i.e., expanded the range of speech levels). Thus the compressor raises the level of the weak sounds so they more readily override the noise on the radio link. The expandor, with its compensating loss, restores the range in speech sound level and reduces the noise level during the transmission of the weak sounds and during silent intervals. These effects are accomplished with no change in circuit net loss.

The compandor was not used extensively on radio circuits because it exaggerated the effects of fading. However, the compandor received extensive use on wire circuits, particularly after the late forties. In the late sixties the compandor principle was revived and applied successfully

[50] Net loss is a measure of the speech level at the receiving hybrid output as compared to the level at the transmitting hybrid input.

to high-frequency radio in a device called the Lincomplex (linked compressor-expandor) based on British Post Office and Bell System developments.

Another major problem arising out of the use of radio was the matter of privacy. Almost anyone with a typical amateur-radio receiver could monitor the high-frequency circuits (although they would commonly hear only one side of the conversation). The low-frequency circuits, using single sideband without carrier, were not picked up by the average receiver but modifications to do so were easily made and would then detect both sides of the conversation since the same frequency was used for both directions.

The problem of privacy was recognized and worked on at an early date. The 1920 AT&TCo Annual Report outlined the then current situation:

> The problem in attaining privacy in radio telephone transmission is peculiar and difficult. Nevertheless, in the solution of this problem we have also made important progress. Our engineers have carried on conversations by radio telephony according to a method which they devised whereby ordinary receiving stations can hear nothing but unintelligible sounds; yet at all stations equipped with the necessary special apparatus, and in possession of the requisite operating information, the spoken works can be heard and understood. Our development in this direction is being continued.

The system referred to, and used on the Catalina circuit just before it was discontinued, was a simple translator that inverted the sidebands so that the lower voice frequencies were at the edges and the higher next to the carrier. This technique was rather easy to apply and thoroughly effective on double-sideband systems since demodulation by a normal receiver gave an inverted voice band (i.e., the low end being translated to the upper end and vice versa). This peculiar inverted speech had the rhythm of normal speech but was unintelligible to the casual listener. However, it proved possible for some people, with practice, to develop a fair capability for understanding it. In addition, it was fairly easy for a knowledgeable and determined eavesdropper to build equipment which would restore the previously inverted speech to normal. As a partial deterrent, there was later added means for "wobbling" the transmitted carrier and sidebands a few hundred hertz.

The inversion techniques were ineffective on single-sideband, carrier-suppressed systems since the carrier that was resupplied at the receiver could easily be placed to give right-side-up speech. The so-called split-band techniques were developed to meet privacy requirements on such systems. With these techniques the speech was divided into small bands (originally four, and later five). These bands were reordered before transmission and restored to normal order after reception. The individual bands could be transmitted right side up or inverted. Figure 5-40 illustrates the process. Originally the frequency arrangement was

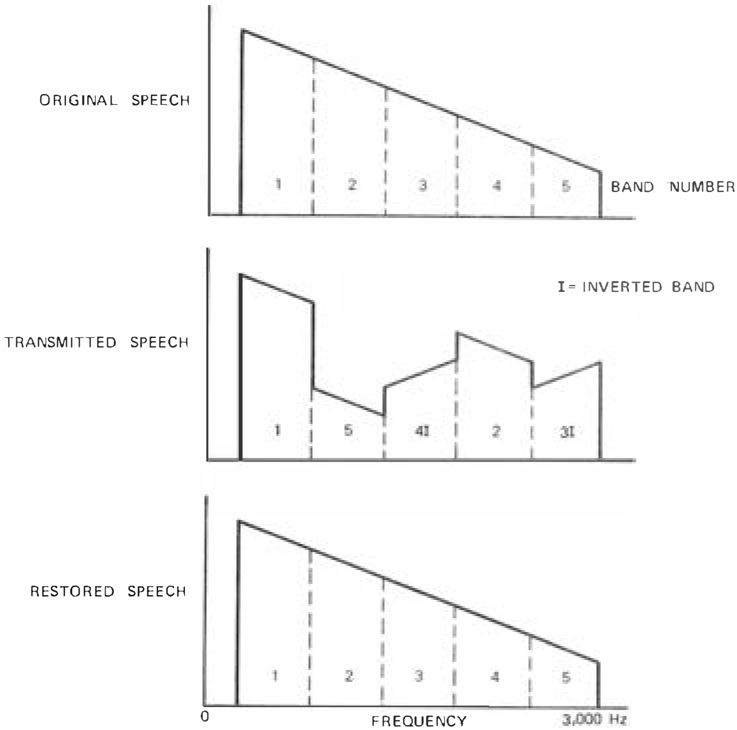


Fig. 5-40. Illustration of split-band privacy techniques.

changed manually and left constant for fairly long periods, but later the order could be changed automatically a few times a minute, the switching at the two ends being done in synchronism. Split-band techniques were used rather widely on both the low- and high-frequency systems (after the latter were changed to single sideband) and were an effective deterrent to the casual listener. However, they did not provide complete security against a determined and technically equipped intelligence expert.[51]

4.3.6 Further Evolution of Transoceanic Telephony

The events described so far carry us well beyond the time frame for this portion of our history. This has been done deliberately since otherwise the continuity of the overseas development story would have

[51] The analogy to the door lock is obvious. It discourages the casual lawbreaker but cannot prevent the knowledgeable and determined housebreaker.

been broken in midstream. By 1930 overseas service was a reality and many of the basic principles for its further evolution had been developed. This is therefore a logical place to stop our detailed story, but a few of our readers may appreciate a glimpse of the years ahead.

After a year or two of commercial service to Europe, it was found that short-wave circuits were available for service about 80 percent of the time (although not all of this circuit time provided high-grade service). This was approximately equivalent to the performance of the long-wave system, and the two systems together gave a fairly large percentage of availability since the high frequencies tended to perform well during the summer months when low-frequency static was severe.

Therefore, in the late twenties it appeared that the main use of the high frequencies might be as a supplement to the low frequencies in North Atlantic service and for circuits requiring very long hops (as in the Pacific) or a crossing of the high-static equatorial belt. However, the demand for overseas circuits continued to grow rapidly and it became clear that the scarcity of long-wave frequency space would prevent the necessary growth. There was no alternative to developing high frequencies for universal use by improving transmission as much as possible. This work was actively promoted during the next ten years and to some extent continues to this day since, in spite of undersea cables and communications satellites, high-frequency radio is still the simplest and cheapest way to establish very long overseas circuits between points with traffic low enough to require at most one or two circuits.

The high-frequency work proved highly successful despite the propagation problems at these frequencies. Exploitation of transoceanic telephony not only provided a useful service but demonstrated the great need for international communications. Evidence of this need was the enormous growth during the ten-year period following its introduction which encompassed one of the worst periods of depression ever experienced by this and the European countries. During this ten-year period, an extensive worldwide short-wave system was set up, as shown in Fig. 5-41, which permitted the interconnection of about 93 percent of the world's telephones and enabled United States telephone users to reach 68 other countries. These circuits developed a market that could only be filled by the submarine cable and satellite technology of the 1950s and 1960s.

Throughout this period of high-frequency development the low-frequency systems were maintained as a backup for transatlantic service and were used when high-frequency transmission was completely interrupted by magnetic storms.

As time went on, numerous improvements were made in the high-frequency system. As noted, single-sideband operation was introduced in 1933 and this was followed by multiplex operation. (Currently,

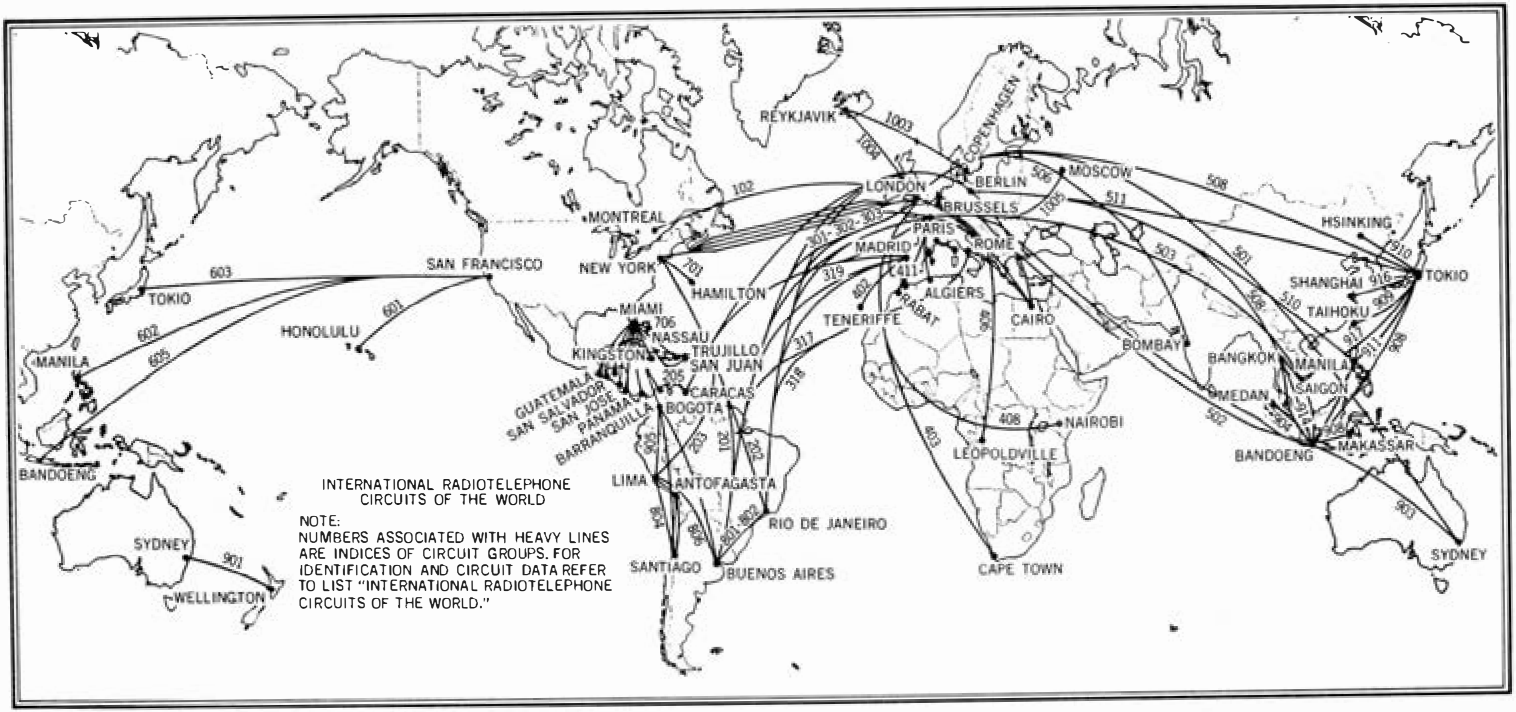


Fig. 5-41. Circuits in operation on the tenth anniversary of the initiation of transoceanic radiotelephone service. The first circuit, New York–London, long wave, was opened on January 7, 1927.

where traffic is sufficient, it is common to use three single-sideband channels, multiplexed on a single transmitter, in a single 10-kHz frequency assignment.[52]) Non-periodic antennas of the rhombic type were commonly used, both for transmitting and receiving. In 1935 the MUSA (Multiple Unit Steerable Antenna) was developed which could be steered electronically to maximize the incoming signal in both the vertical and horizontal directions. This antenna received considerable use at the time, but its benefit became less valuable as increased transmitting power became cheaper and it became less competitive with the simpler non-periodic rhombics. Power amplifiers were increased in size and now output powers as great as 80 kW (peak envelope power) are used.

In 1931 short-wave transmitter and receiver stations were established at Dixon and Point Reyes, California, to handle transpacific service, and Florida stations were set up in 1932 for Caribbean and Central American service. By the early 1960s when high-frequency radio had achieved peak use, about 200 circuits were in use to 58 countries.

V. RADIO BROADCASTING

In retrospect, the evolution of radio broadcasting seems inevitable once the technological problem of speech modulation had been solved. Radio waves tended to propagate in all directions from their point of emission; as noted in the Craft–Colpitts AIEE paper of 1919 (see Section 4.2), broadcasting could therefore take advantage of a characteristic of the medium which was a handicap to the development of point-to-point service. But, while the technical path seemed clear, the commercial evolution was far from certain. There were no precedents to guide the development of this new service, and even today, 50 years after the first broadcast, the commercial approach differs widely among the countries of the world and the objectives that should guide broadcast programming are still debated. For this reason the early history of broadcasting in the Bell System was not concerned only with technical developments, which to a large extent flowed directly from the development of two-way point-to-point telephony. It also involved the evolution of a commercial approach to broadcasting and decisions regarding the Bell System position within this framework.

5.1 Early History

The broadcast concept was not new. It will be recalled that Bell, as early as 1877, had transmitted music by wire and had promoted his invention by demonstrating such transmissions to an audience

[52] In some cases, four channels are multiplexed in a 12-kHz assignment.

assembled in an auditorium. There was no dearth of imaginative prophets who foretold the time when large audiences would be entertained by opera and orchestral programs transmitted by wire. The fact was, however, that these early demonstrations were in the nature of stunts and the prophecies were premature since the technology to support commercial broadcasts of any kind was not available until high-power amplifiers became practical after World War I. Nevertheless a number of important preliminary steps had been taken at an early date.

The Bell System had developed and used a loud-speaking telephone as early as 1908 and a few years later added high-power microphones to form the basis for a primitive public address system. In 1913 such equipment was used for a long-distance broadcast over an open wire circuit at which the Governor of Oklahoma, speaking in Okalahoma City, addressed a group of over 300 people in Tulsa some 120 miles away. Since this was before the day of successful vacuum tube amplifiers, only the amplification provided by the carbon transmitter was available and a very powerful, water-cooled device was employed. By 1919 the vacuum tube amplifier was a reality and a public address installation covering three city blocks and using over 100 loudspeakers was employed during a Victory Loan Drive in New York City. Probably the most elaborate wire broadcast demonstration to date was held on November 11, 1921, when the services held at the burial of the Unknown Soldier in Arlington, Virginia, were broadcast by wire and loudspeakers to an audience of about 150,000 people in Arlington, New York City, and San Francisco. This represented the high point for this type of broadcast by audio means to large numbers of people assembled in a public place since by this time radio had developed to the point where it was apparent that it could be readily used for broadcasting to individuals in their homes. These early experiments with the public address type of broadcast did much, however, to prepare the way for radio broadcasting since they provided the necessary high-quality audio equipment (microphones, amplifiers, loudspeakers) which were required for the subsequent development of radio broadcasts. In addition, they provided experience in wire transmission of high-quality music and speech (subsequently required for networking) and demonstrated that there was an enormous potential audience for the broadcast of important events.

The first radio broadcast was probably that made by Lee de Forest in New York City in 1907.[53] Again, this was in the nature of a stunt

[53] His "fanmail" consisted of three notes from Brooklyn amateurs who heard the broadcast.

since commercial broadcasting could not become a reality until the postwar development of vacuum tube radiotelephone techniques. By 1920–21 radiotelephone technology was becoming available in the relatively unused frequency range centered at about a megahertz and a sizable and growing audience existed in the radio amateurs spread through the country. Many of these operated only receiving stations and were highly receptive to an opportunity for using them for something other than listening to distant telegraph stations. Moreover, they were anxious to demonstrate their technical skill to friends and neighbors when speech or music transmission provided an interesting opportunity.

Broadcasting on an informal basis had occurred sporadically in the years after World War I but Frank Conrad, an employee of the Westinghouse Electric and Manufacturing Company in Pittsburgh, seems to have been the first individual to begin broadcasting on a continuing basis. He started transmitting speech and music in 1919 from his amateur station 8XK located in his home. The Westinghouse Company at this time had begun manufacturing radio receivers for sale to amateurs and, seeing this as an opportunity for promoting sales, took over Conrad's broadcasting activities using a commercial station licensed with the call letters KDKA. This station inaugurated service on November 2, 1920, by broadcasting returns of the Harding–Cox presidential election.[54] This created a local sensation and soon amateurs all over the country were adapting their transmitters for speech and music broadcasts and commercial interests were not far behind in applying for licenses. At this time a license involved little more than the registration of a station with the Secretary of Commerce and could not be denied. The Commerce Department had set aside two frequencies (360 and 400 meters) for broadcasting, but these were obviously insufficient for the enormous growth in stations in the middle twenties and conditions soon became chaotic as broadcasters attempted to find a frequency free from local interference. This condition continued until the Radio Act of 1927 set up the Federal Radio Commission as a regulatory body and made a wide band of frequencies available for broadcasting.

5.2 The Bell System Position Evolves

Some involvement of the Bell System in broadcasting was certain

[54] Although KDKA is usually referred to as the first commercial broadcasting station, the National Association of Broadcasters gives this credit to a station operated by the *Detroit News*. This station, originally operated under an amateur license, later became WWJ. It began operation on a fairly regular basis some ten weeks before KDKA and included among its broadcasts election returns from the primary election of September 1920.

since its technical qualifications in this field were unique. Research in the field of acoustics and experience with public address systems had provided a background in the audio field that was unmatched. The early work on radiotelephony had provided basic vacuum tube circuitry and the Bell System's own strong patent position was enhanced by the cross-licensing agreement reached in 1920. The ship-to-shore and Catalina experiments had provided experience in the frequency band open to broadcasting and as early as 1918 development had started on the high-power vacuum tubes which would be needed for metropolitan service. Finally, it became clear that wire connections would often be needed between studios or pickup points and the radio transmitters. The Bell System not only had the physical facilities available for this type of work over much of the country but also had experience in adapting the circuits for the high-fidelity transmission required for broadcasting.

Thus, in the early 1920s it was clear that the Bell System was in a unique position for supplying either broadcasting systems or services. But the manner in which broadcasting would develop and the role to be played by Bell was far from clear.

Westinghouse began broadcasting on the assumption that it would promote the sale of receivers manufactured by the company. The motivation of other early broadcasters was not always so certain. The newspapers undoubtedly saw radio as a logical extension of the news media which they could not ignore. Large department stores built stations as a form of "institutional" advertising that would publicize their name and enhance their image.[55] Soon broadcast stations were being built all over, many by people who had no clear idea as to their objectives. It should be remembered that at this time a capital investment of as little as $10,000 to $15,000 would build a complete station. Many station owners ignored the fact that annual operating costs would be several times this amount and no one foresaw the ultimately enormous cost of programming with paid performers.

The Bell System could enter the broadcasting field either as a manufacturer supplying equipment or directly as a station operator. Neither field was entirely satisfactory. There was, as noted, a great demand for transmitters and for the moment the equipment field looked very attractive but, with only one or two wavelengths available, it was apparent that either the market would soon be limited or the interference from hundreds of stations would be so intolerable that each would have to be limited to short, preassigned broadcast periods. This same factor also made the direct entry into broadcasting unattrac-

[55] It is interesting to note that the John Wanamaker organization, an early broadcaster, had for a long time featured daily organ concerts and similar public entertainment in their stores which logically led into similar radio broadcasts.

tive since it was apparent that even the modest costs of that time could not be supported with only short periods of use.

It was decided to enter the equipment manufacturing field with some caution and to explore ways for providing broadcast service which would meet current needs and have a possibility of being economically sound. The development of equipment will be covered later. At this point we need only point out that the first station, a 500-watt transmitter, was sold to the *Detroit News* and put into service in January 1922. By April 1922 sixteen stations had been manufactured and quotations were outstanding on some 50 more.

In 1921 the outlook for providing broadcast service was far from attractive. Millions were anxious to listen, hundreds were setting up stations, but no one was sure how broadcasting could be supported economically or how the many stations could operate without intolerable mutual interference. In October 1921 an internal AT&TCo memorandum suggested a possible way to handle these problems by means of a wire network, with alternate routing to provide reliability, linking public address locations and radio stations in all parts of the country so that major events could be broadcast to a countrywide audience. By the end of 1921 this concept of a nationwide broadcasting service had developed further. The idea of public address broadcasting had been de-emphasized (but not completely dropped) in favor of more extensive radio broadcasting from 38 stations located strategically on main Bell System long-distance routes so as to essentially cover the country. The stations could be used individually for local programs or connected together through the wire plant for events of widespread interest. In early January 1922 these plans were refined and a decision made to build an initial station so that the basic concepts, both engineering and commercial, could be tested in practice.

The unique commercial concept was that of "toll broadcasting" which can best be explained by a quotation from the public announcement of this service made February 11, 1922. After announcing that a new station, with some unique technical features, would be constructed in the long-distance building at 24 Walker Street, New York City, the announcement continued:

> It will be unique in another respect because it will be the first radio station for telephone broadcasting which will provide a means of distribution and will handle the distribution of news, music or other programs on a commercial basis for such people as contract for this service.
>
> The American Telephone and Telegraph Company will provide no program of its own, but provide the channel through which anyone with whom it makes a contract can send out their own programs. Just as the Company leases its long distance wire facilities for the use of newspapers, banks and other concerns, so it will lease its radio telephone

> facilities and will not provide the matter which is sent out from this station. . . .
>
> This is a new undertaking in the commercial use of radio telephony and if there appears a real field for such service and it can be furnished sufficiently free from interference in the ether from other radio services, it will be followed as circumstances warrant by similar stations erected at important centers throughout the United States by the American Telephone and Telegraph Company. As these additional stations are erected, they can be connected by the toll and long distance wires of the Bell System so that from any central point, the same news, music or other programs can be sent out simultaneously through all these stations by wire and wireless with the greatest possible economy and without interference.

The new station, operating with the call WBAY, began operation on July 25, 1922, on a frequency of 830 kHz (360 meters) and sharing time with ten other stations in the New York area operating on the same frequency.[56] Service was offered to anyone wishing to broadcast at the rate of $50 for a 15-minute period in the evening and $40 for a similar afternoon period. It was a month before anyone engaged the facilities[57] and it soon became evident that "sustaining" programs would have to be provided in order to gain and hold the interest of the audience between commercial programs. This was contrary to the basic premise that the Company would not do programming but the alternative was to abandon the experiment. So started a whole new endeavor, namely, the development of program policies and means for implementing them. Soon this was to become a large part of the commercial development, leading to numerous innovations which are beyond the scope of this history.[58]

The technical performance of WBAY was disappointing owing to radiation problems at the Walker Street location, and on August 16 operation was switched to a similar experimental station, with call letters WEAF, set up at the Western Electric Engineering Laboratory at 463 West Street. The performance was greatly improved and the station continued at this location until after WEAF was sold by the

[56] At this time, 360 meters was the only wavelength used for commercial broadcasting. An assignment of 400 meters (750 kHz) had been requested by the Company, but it was not until September that this wavelength was authorized for "Class B" stations such as WBAY and WEAF.

[57] The first commercial broadcast was by the Queensboro Corporation on the afternoon of August 28, 1922. It was discovered that prospective users were deterred by the low charges, believing they must be indicative of little value. Later the rates were increased to $100 for a 10-minute period in the evening and $50 for the afternoon. Another deterrent was the prohibition, placed by the Company, on "direct" advertising mentioning prices or describing packages.

[58] A very complete history of early Bell System broadcasting is given in *Commercial Broadcasting Pioneer: The WEAF Experiment*, by William P. Banning, published in 1946 by the Harvard University Press.

Company in 1926. The studios were continued at the Walker Street building until new and larger quarters were made available at the AT&TCo headquarters building at 195 Broadway in April 1923.

During the next few years the broadcasting pattern evolved. Many new types of programs were developed, including orchestras, glee clubs, musical shows, religious services, prize fights, baseball games, etc. A second Bell station, WCAP, was opened in Washington, D. C., on July 4, 1923, with a program transmitted from WEAF over wire lines. But well before this, network broadcasting began, the first experiment being on January 4, 1923, when WEAF and WNAC in Boston were interconnected by wire and broadcast the same program. Throughout the summer of 1923, WEAF and WMAF in Round Hill, Massachusetts, were regularly broadcasting the same program, and on June 7 a chain of four stations, WEAF in New York, KYW in Chicago, KDKA in Pittsburgh, and WGY in Schenectady, were joined by wire for a broadcast from New York. From this time on network broadcasting grew rapidly and within a few years networks connecting 20 or more stations became common.

Technical changes also occurred. In May 1923, WEAF began broadcasting on 492 meters (610 kHz), a frequency which was used for many years. A crystal-controlled master oscillator was introduced in 1924 to reduce frequency modulation and improve fidelity in areas where multipath transmission caused signals of poor quality (as discussed subsequently). Beginning early in 1924, steps were taken to increase power beyond the 500 watts of the initial installation. This was opposed by many city listeners who, using primitive sets with little selectivity, had difficulty in tuning out WEAF when listening to other stations. The move was applauded by those at a distance, particularly those in poor coverage areas. Opposition decreased as the problem became better understood and the value of highly selective receivers was recognized by the listener, but it was not until September 1925 that the full transmitter power of 5 kW was used following small incremental increases during the preceding year and a half.[59]

By this time the WEAF experiment was drawing to a close, but before concluding this part of our narrative, the reader may be interested in seeing some of the physical aspects of broadcasting at the time. Figure 5-42 shows the transmitter used when the AT&TCo station was opened in 1922 as WBAY. The equipment used by WEAF (Fig. 5-43) was essentially identical but it included additional experimental and control gear. The antennas were simple flat-tops with little directivity, WBAY's antenna being on the roof of 24 Walker

[59] Actually the first experimental broadcast at 5 kW was made on the last day of 1923 but, as related above, it was some time before such power was approved for routine use.

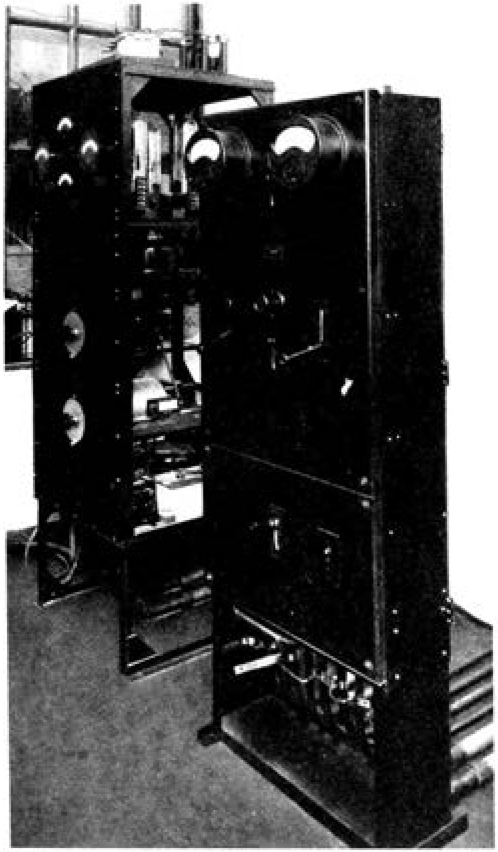


Fig. 5-42. WBAY's radio broadcasting transmitter, 1922, at the American Telephone and Telegraph Company Long Lines Building, Walker Street, New York City. (Banning 1946, opp. p. 79)

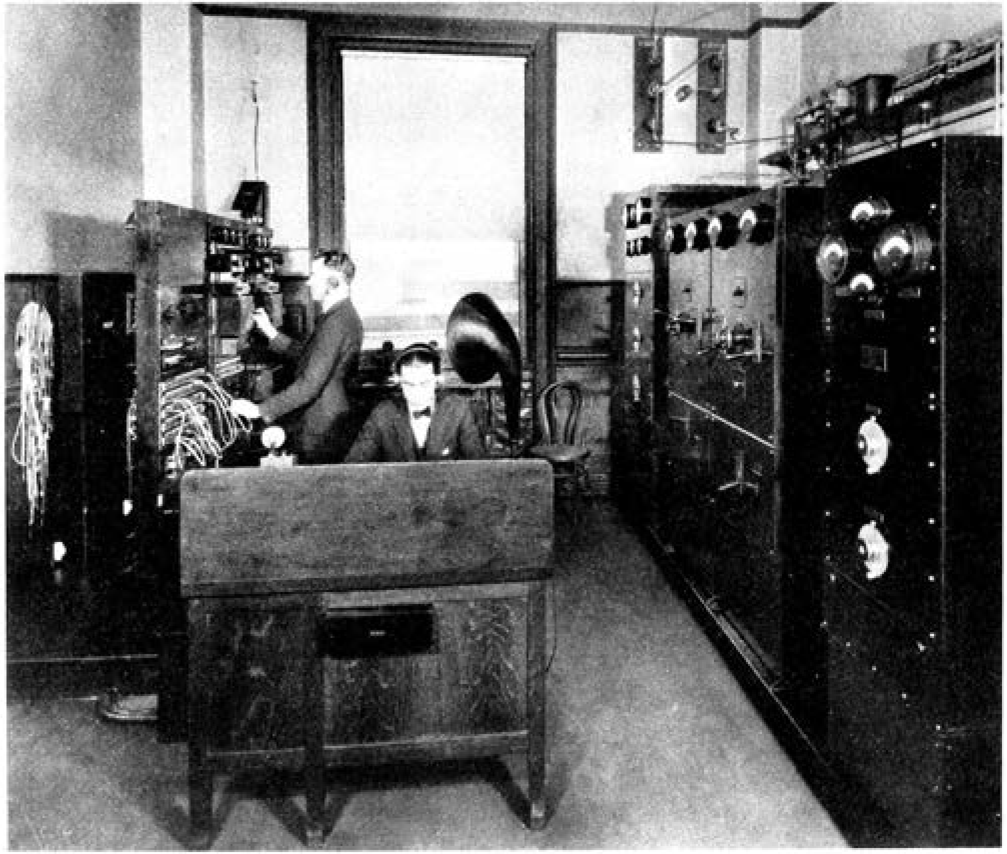
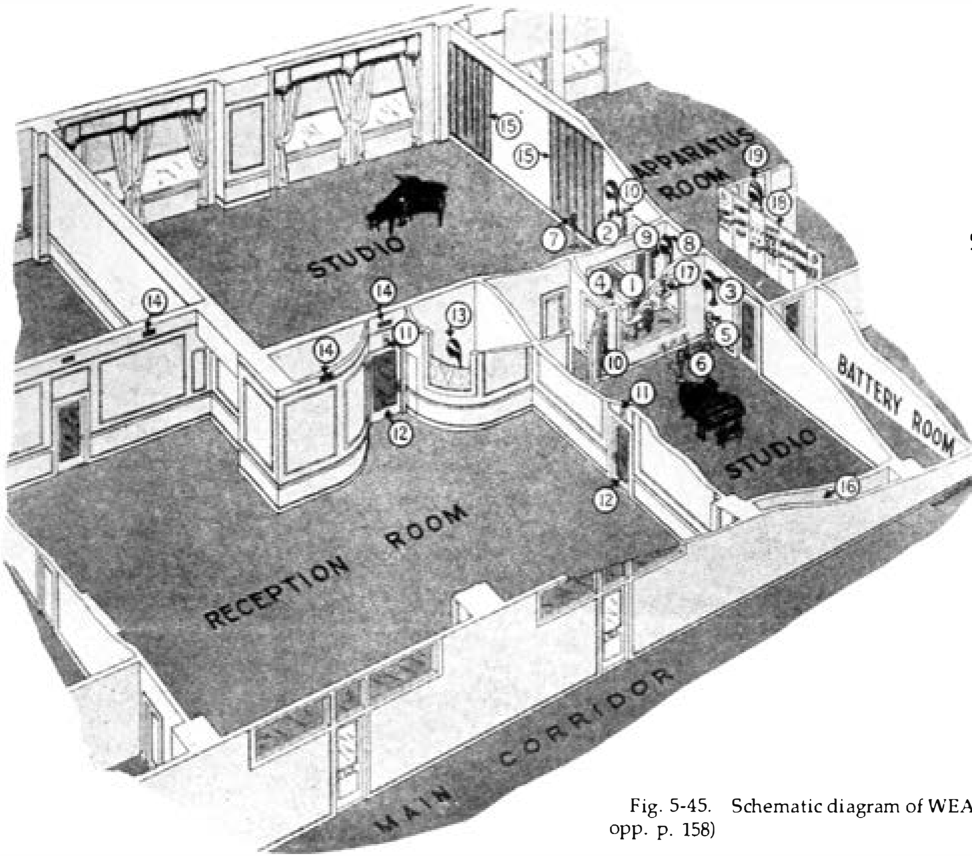


Fig. 5-43. Experimental station 2XB at the West Street laboratory, which in 1922 became WEAF. (Banning 1946, opp. p. 95)

Fig. 5-44. Miss Helen Hann, of the American Telephone and Telegraph Company Long Lines Department, was WEAF's first studio hostess and an announcer. (Banning 1946, opp. p. 110)

Street and WEAF on the roof at 463 West Street. The studio, Fig. 5-44, was looked upon as little more than an enlarged and upgraded telephone booth with facilities the sponsor could use for speech and simple musical programs (including a commercial phonograph). The move to 195 Broadway in 1923 provided a studio setup (Fig. 5-45) prophetic of the future. By 1924 the transmitter had grown to a large size (Fig. 5-46) and the input equipment (Fig. 5-47) was growing in complexity.

By 1925 a change in the Bell System position with regard to broadcasting became imperative. There were many contributing causes of which only a few can be mentioned briefly here. The patent position for some time had been troublesome. The agreements made in 1920 preceded the advent of broadcasting and their interpretation in regard to this new form of communication was in dispute and would not easily be resolved so long as Bell continued to operate a broadcasting station. Bell was primarily interested in preserving its freedom to develop and manufacture tubes and equipment for telephone purposes, and was willing to make some compromise in the broadcast field. The reason for doing so is a curious one: the WEAF experiment had become, in a sense, too successful.



1. Announcer's microphone.
2, 3. Loud-speakers in the studios.
4. Announcer's control panel.
5, 6, 7. Microphones in studios.
8. Announcer's loud-speaker.
9, 10. Signal lights.
11. Door signal lights showing when studio was on the air.
12. Special door knobs to prevent interruption.
13. Loud-speaker for reception room.
14. Ventilation system.
15. Curtains for acoustic effects.
16. Sound-proof wall.
17. Announcer's telephone for communication with engineers.
18. Equipment panels.
19. Loud-speaker for monitoring engineers.

Fig. 5-45. Schematic diagram of WEAF studios at 195 Broadway, 1923. (Banning 1946, opp. p. 158)

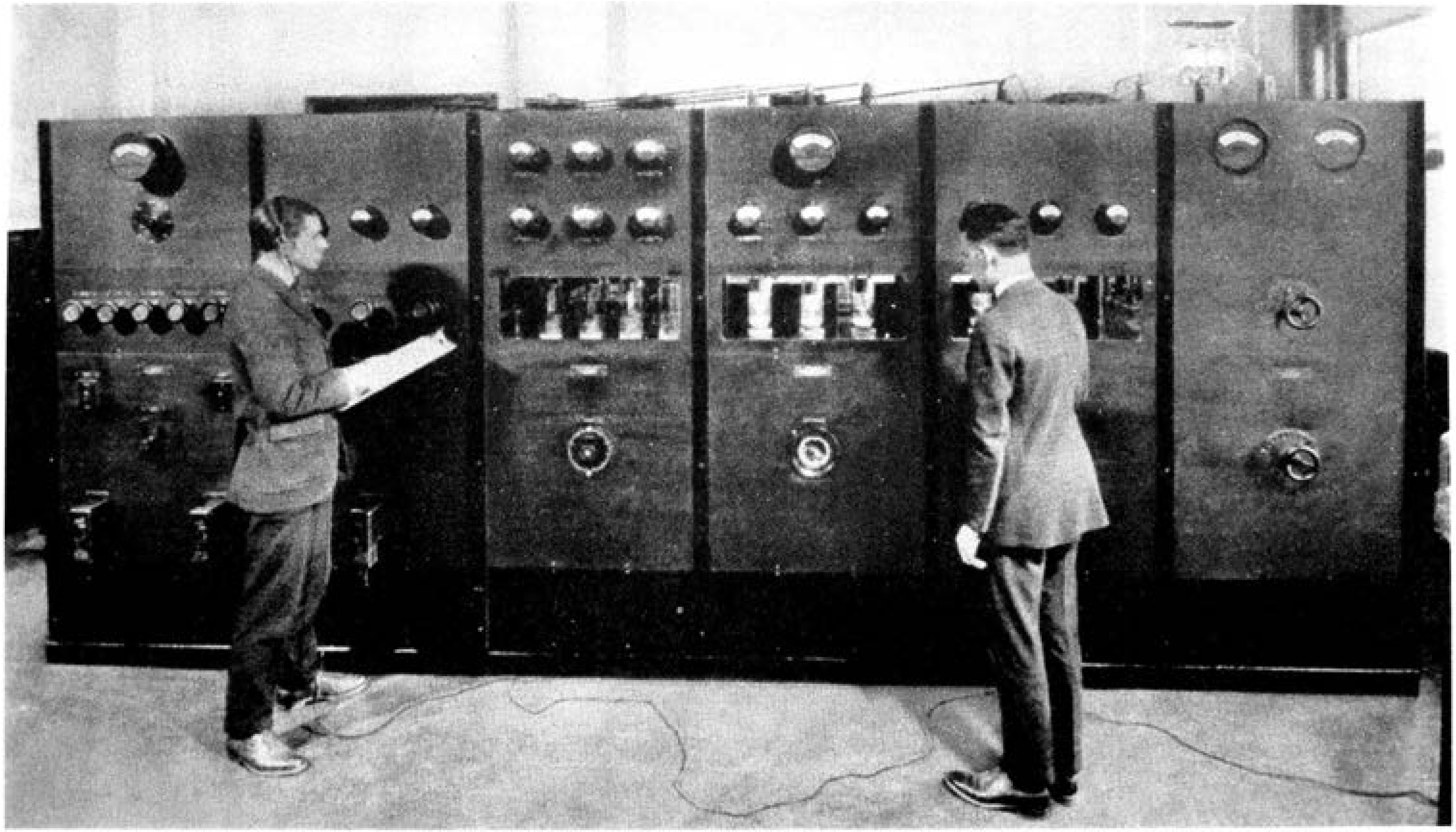


Fig. 5-46. The world's first 5-kilowatt broadcast transmitter, put in service by WEAF on the eve of January 1, 1924. In June 1924, crystal control was added, making it the first broadcast transmitter with this feature. (Banning 1946, opp. p. 191)

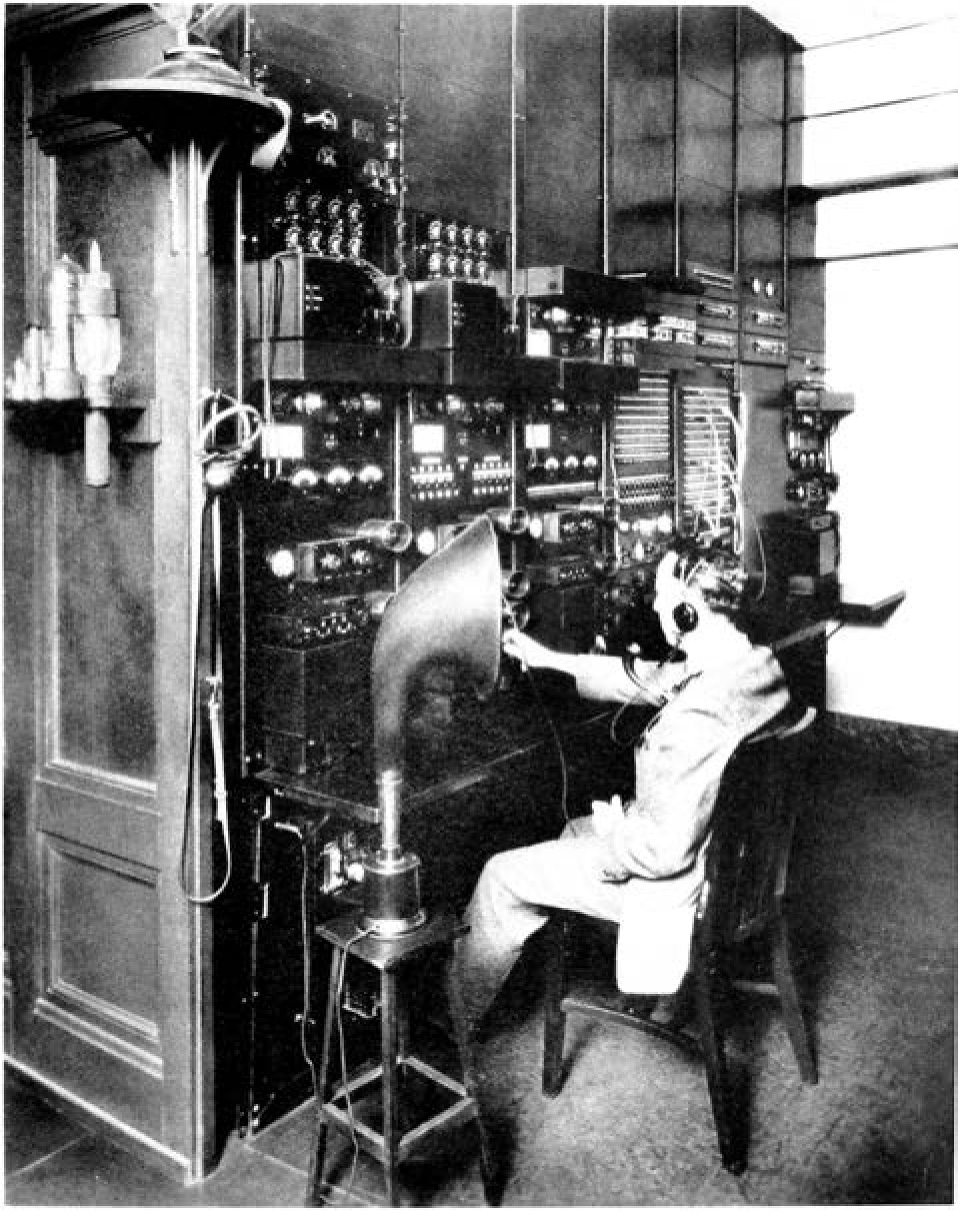


Fig. 5-47. WEAF's speech-input equipment at American Telephone and Telegraph Company Headquarters, 195 Broadway, New York City, 1924. (Banning 1946, opp. p. 190)

The original intent of the toll broadcasting proposal was, it will be remembered, to provide a service which could be used by commercial users to reach audiences in all parts of the country without going to the expense of establishing individual stations for each user. In this, the experiment was highly successful.[60] Regularly scheduled network programs identified with commercial sponsors had become commonplace and an accepted part of the lives of millions of listeners. Programs which initially had been simple announcements and largely amateur musical performances had grown to the point where the Atwater–Kent program, inaugurated in October 1925, brought to the listeners of 11 stations performances by the world's most eminent musicians at an annual cost of over $100,000. Not only had the operation achieved a considerable degree of artistic success but by 1925 it was also a commercial success, the operating profit for that year being about $150,000 compared to losses of over $100,000 in each of the preceding two years. But this success was achieved only by assuming responsibilities not contemplated when the experiment was initiated and remote from the normal Bell System functions. As noted earlier, a large and receptive audience could be assured only by operating many hours of the day when commercial sponsors were not available. This could be done only by providing station- or network-sponsored "sustaining" programs. This in turn led to assistance in programming for others and ultimately to setting up an artists' bureau through which performers who had achieved popularity by radio could be engaged for non-radio events. It was evident that radio was not only a communication enterprise but was also largely an entertainment business. What is more, it did not seem possible to separate this part of the business from the commercial operation of a radio station, and Bell management began to question the appropriateness of their continuing the entertainment aspects of the project.

While success of sponsored network broadcasting was being demonstrated by WEAF, stations being operated by individuals found expenses rising and began to question the benefits accruing. It was apparent to many that sponsored network broadcasting, as developed during the WEAF experiment, was the answer to achieving economic viability. It was with this background that negotiations were begun to settle the patent situation and resolve the future of broadcasting in the Bell System. The upshot was a decision to sell broadcasting station WEAF but to remain in the business of furnishing wire communication between stations and between stations and remote studios or pickup

[60] This may be disputed by modern listeners, jaded by years of commercial listening, but the unbiased historian judging the situation in the light of the time will probably agree as to the quality of much of the WEAF programming and the restraint exercised in commercial announcements.

points. Thus Bell would provide the physical facilities for networking and continue the strictly communication business but retire from station operation and the associated entertainment business. This division of effort not only resolved the business conflict but also removed most of the disputatious points in the patent situation so that an agreement protecting Bell interests could easily be reached.

The sale of WEAF to the Radio Corporation of America was announced on July 22, 1926, and in October this company formally transferred the station to the newly formed National Broadcasting Company (NBC). The studios continued at 195 Broadway until August 1927 when they moved to the uptown headquarters of NBC.

5.3 Technical Developments

The technology of broadcasting developed in parallel with and in support of the commercial evolution just recounted. In the Bell System these developments followed three main paths:

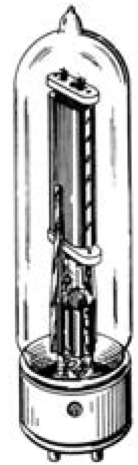
(*i*) Provision of audio input and output equipment, together with means for transmitting audio program material by wire lines
(*ii*) Production of radio broadcasting transmitters and associated gear
(*iii*) Investigation of solution of wave propagation problems characteristic of the broadcast frequencies.

The development of audio equipment and program transmission systems had begun earlier under the stimulus of public address broadcasting but achieved technical maturity only as the concept of network radio broadcasting evolved. This work has been covered fully in the preceding chapter on wire transmission and needs no further exposition here. The present portion of the broadcast history will therefore be devoted to the solution of the technical problems uniquely concerned with radio equipment and propagation.

5.3.1 Radio Broadcast Equipment

At the time when experimental broadcasting was undertaken by Frank Conrad in Pittsburgh and by the *Detroit News*, the Bell System had built relatively high-power stations, operating in the 1-MHz band, for experimental ship-to-shore and point-to-point communication. The further development of this equipment to meet the requirements for broadcasting (i.e., for a wider audio band and wider volume range) was a logical step and was undertaken in 1920 with testing being done by experimental station 2XB at the West Street laboratory. There were many favorable comments on the high technical quality of this station, leading to the first sale of transmitting equipment, rated at 500 watts, to the *Detroit News* station WBL (later WWJ) which went into service in January 1922. Very similar equipment

NO. 212A VACUUM TUBE



NO. 211A VACUUM TUBE

Fig. 5-48. Western Electric Company 250-watt and 50-watt tubes. (From WECo Bulletin T-670)

was used by WBAY and WEAF, the Bell System stations, when they began operation in mid-1922.

In May 1923 a Western Electric Company bulletin announced the availability of two standardized broadcast equipments, the 101A[61] rated at 500 watts and the 102A with a 100-watt rating. In 1924 a 50-watt equipment designated Type 103 was added to the line to fill the need of stations with very small coverage area. All of these were basically similar, using glass-envelope, air-cooled tubes of the 212 and 211 type (rated respectively at 250 and 50 watts) in the high-power stages. Both of these tubes, shown in Fig. 5-48, used long-life Wehnelt-type oxide-coated filaments operating at a dull red heat. The

[61] Later, basically similar equipments used the same code number but different letter suffixes. The 101B was one of the most commonly used early systems.

manufacturing dates of these early systems and information on the high-power tube complement are shown below:

System Code	Transmitter Code	Power Rating	Mfg. Date*	High-Power Tube Complement
101A	1A	500 watts	1-10-23	Four Type 212 One Type 211
102A	2A	100 watts	6-10-23	Five Type 211
103A	3C	50 watts	5-20-24	Two Type 211

* Dates on which final manufacturing information was given to the manufacturing organization by the development engineer. Manufacture is often begun well before this date on the basis of preliminary information.

All three systems used the so-called "constant current" modulating scheme devised by Heising. In this scheme, also called "plate current" modulation, the audio signals are amplified to a high power level and superposed on the direct voltage impressed on the plate of the oscillator tube. This type of modulation proved practical and simple to apply. It had been used in the military applications of World War I and the ship-to-shore experiments and was widely applied in early broadcasting. Later, more sophisticated schemes were employed using low-level modulation with linear amplification to the desired output power. The initial 101- and 102-type systems were designed to operate at frequencies between 500 and 1,000 kHz but later systems had their range extended to 1,500 kHz.

A unique feature of these equipments was that they were offered as complete systems incorporating essentially all the components needed for a broadcast station except the prime power, antenna, and building. Figure 5-49, from the bulletin covering the 103C system, illustrates the completeness of the equipment package. In all of these systems a radio receiver was an essential part of the package since government regulations required monitoring at intervals to determine if distress signals were being sent (in which case broadcasting had to be discontinued) or if there were indications that broadcasting was interfering with other radio communication.

Typical broadcasting equipment of this early period is shown in Fig. 5-42 (radio transmitter and power board), Fig. 5-50 (radio receiver), and Fig. 5-51 (microphone). The last figure shows the commonly used carbon-type transmitter consisting of a tightly stretched diaphragm with two carbon buttons, one on each side of the diaphragm, used in a push-pull arrangement which greatly eliminated production of even-order harmonics. For the very highest quality audio reproduction a condenser (capacitor) transmitter was available, but since it required a very quiet

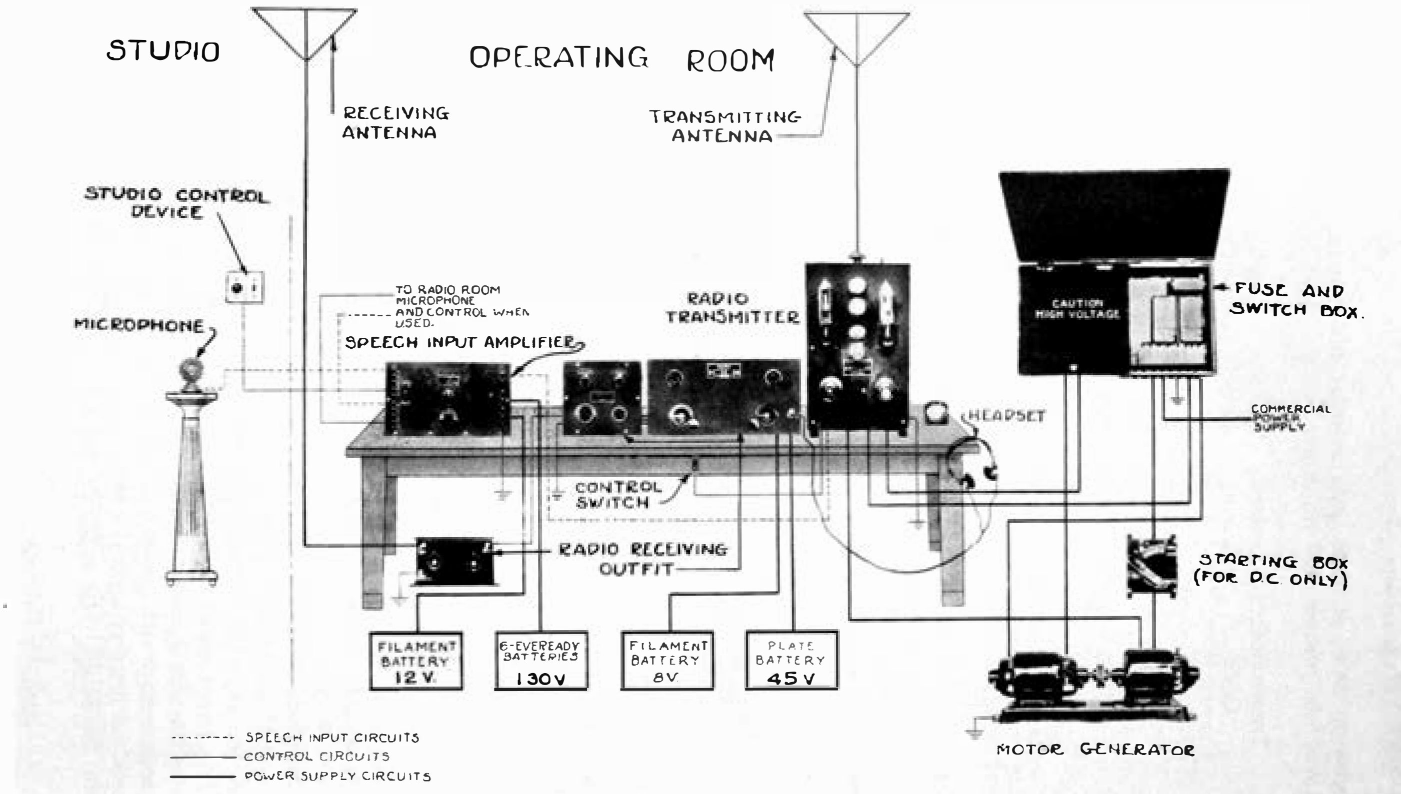


Fig. 5-49. No. 103C radiotelephone broadcasting equipment.

Fig. 5-50. No. 2C radio receiver. (From WECo Bulletin T-670)

Fig. 5-51. Carbon microphone. (From WECo Bulletin T-670)

multistage battery-operated amplifier, in close proximity to the microphone, this device originally was suitable only for studio use. An early form of this equipment is shown in Fig. 5-52. Somewhat later, more compact and rugged condenser-transmitter equipment was developed, as related in Section 8.4.1 of Chapter 3.

Even with the carbon transmitter, the frequency response of this early broadcasting equipment was quite good as shown by Fig. 5-53 which applies for the 101B system of 1923.

Growth in broadcasting systems manufactured by Western Electric was rapid in the early years, averaging about 30 per year. Most of these were of 500-watt rating or less, but a few towards the end of 1925 were of higher power.

The need for higher power to provide adequate coverage of large metropolitan areas was recognized at an early date and a 5-kW transmitter was installed at WEAF in late 1923 and used for the first time on the last day of that year (see Fig. 5-46). The high power output was achieved by adding a 5-kW amplifier, using water-cooled tubes, to the existing 500-watt Type 1 transmitter. This combination of the 500-watt transmitter as a driver for a 5-kW output amplifier was later coded as the Type 4 transmitter and used in the 104 system by a number of stations.

At WEAF another important advance was made with the addition of a crystal-controlled oscillator to stabilize the carrier frequency, a "breadboard" installation being made on a trial basis in June 1924. In 1925 a more permanent arrangement was installed and WEAF remained under crystal control thereafter. This feature was not made a standard part of the

Fig. 5-52. Condenser-transmitter equipment. (Nelson 1924, Fig. 2)

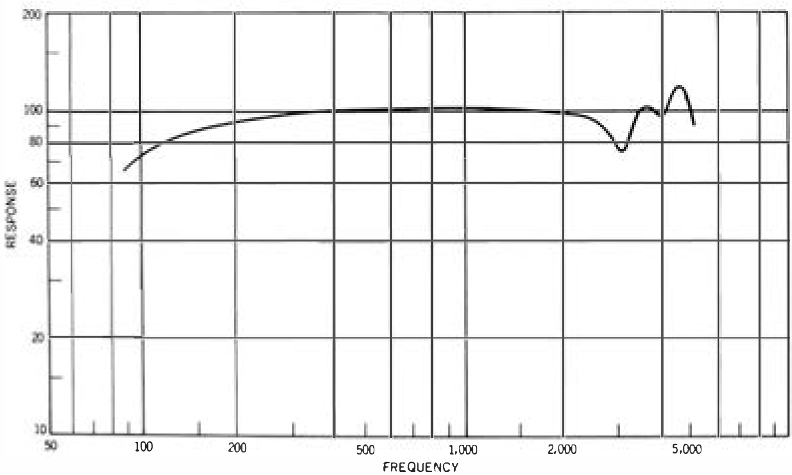


Fig. 5-53. Frequency response characteristic with carbon transmitter. (Redrawn from Nelson 1924, Fig. 13)

104 system since this system was only an interim step toward a new series of high-power systems, all using water-cooled tubes, which were introduced in the late twenties, as indicated below:

System Code*	Transmitter Code	Power Rating	Mfg. Date	Water-Cooled Tubes in Final Amplifier Stage
105B	5C	5 kW	6–28	Two Type 220 (10 kW)†
106B	6B	1 kW	10–28	One Type 228 (4 kW)
107A	7A	50 kW	6–28	Six Type 232 (40 kW)

* These are the codes of the earliest "second generation" systems described in the text. In a few cases the same code numbers had been used earlier, with other letter designations, for interim systems.

† These are the peak power ratings of individual tubes.

The No. 5 type transmitter is illustrated in Figs. 5-54 and 5-55 and its circuitry is shown schematically in Fig. 5-56. The other transmitters in the series used the same unit type of construction but employed different numbers of bays. The small, 1-kW Type 6B transmitter is illustrated in Fig. 5-57 and the 50-kW Type 7A in Fig. 5-58. The latter is shown in the Whippany, New Jersey laboratory where it was developed. In addition to the ten bays of equipment shown in the figure, a 5-bay structure, at the base of the antenna, was required for coupling and tuning.

These were truly "second generation" systems differing from the early systems in many ways besides physical size and high output power. Major performance changes were: greatly improved frequency response, reduction in audio- and radio-frequency harmonic distortion, and improved precision and stability of carrier frequency. The improved frequency response and reduction in audio-frequency harmonic production was needed to take advantage of the growing improvement in the fidelity of receivers and loudspeakers available to the public. Radio-frequency harmonics had to be kept low in order to minimize interference with other stations operating at frequencies which were an integral multiple of that used by the radiating station. Precise carrier frequency was necessary to avoid unpleasant "beat frequency" interference from other stations at a distance operating on the same nominal frequency. By reducing this type of interference it became possible to reduce the spacing between stations on the same frequency and hence achieve a better utilization of the available frequency band. By improving carrier stability (i.e., reducing short-term variations in frequency) the distortion resulting from frequency modulation, discussed in Section 5.3.2, was greatly reduced.

Two major design changes were introduced in achieving these improvements, low-level modulation and crystal control of frequency. Low-level modulation, it will be remembered, was the technique used for transatlantic telephony. It was technically more difficult to implement

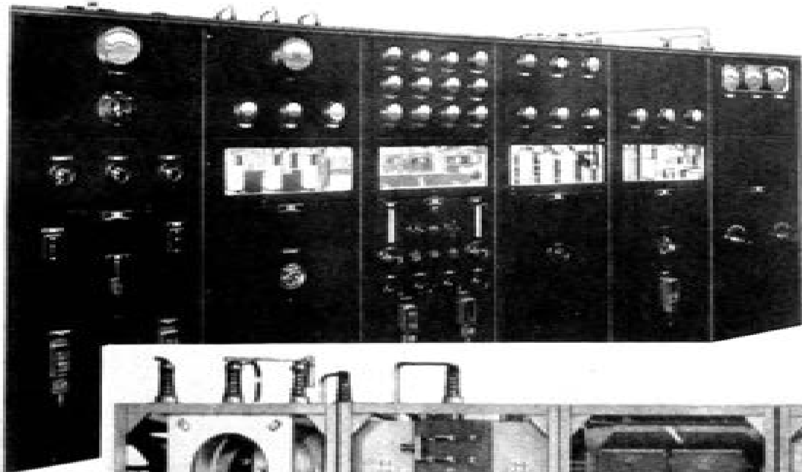
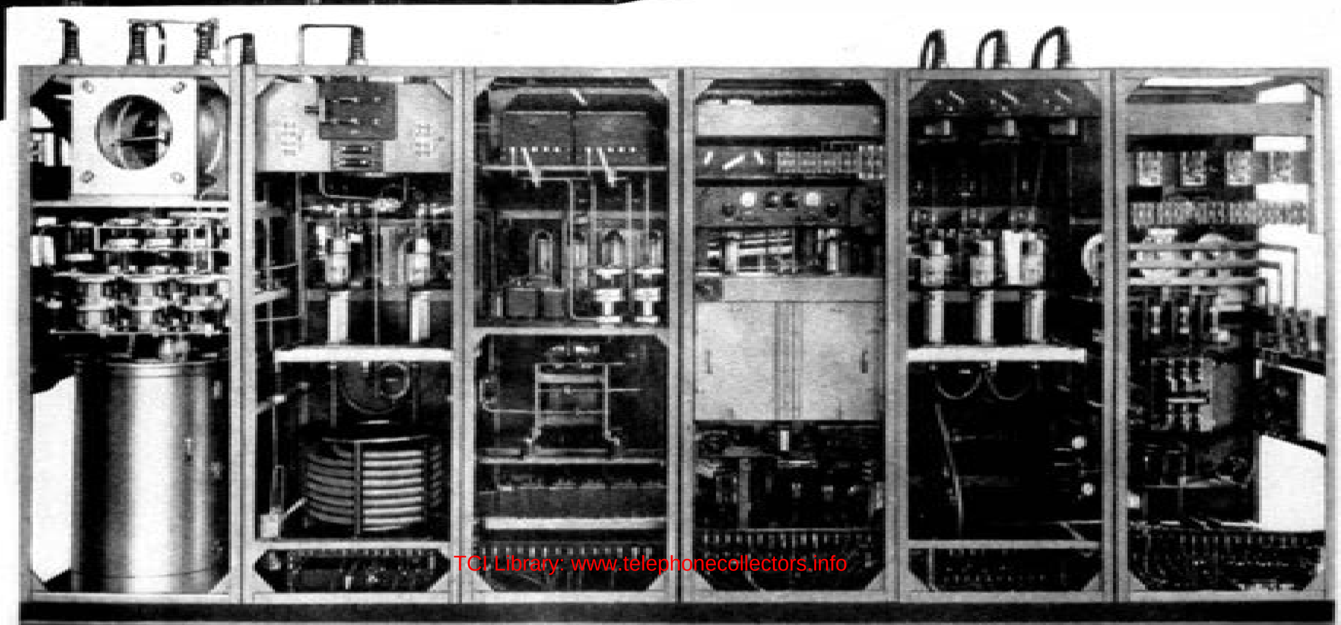


Fig. 5-54. No. 5C radio transmitter—front view. (From WECo Bulletin 410)

Fig. 5-55. No. 5C radio transmitter—rear view of panel units. (From WECo Bulletin 410)



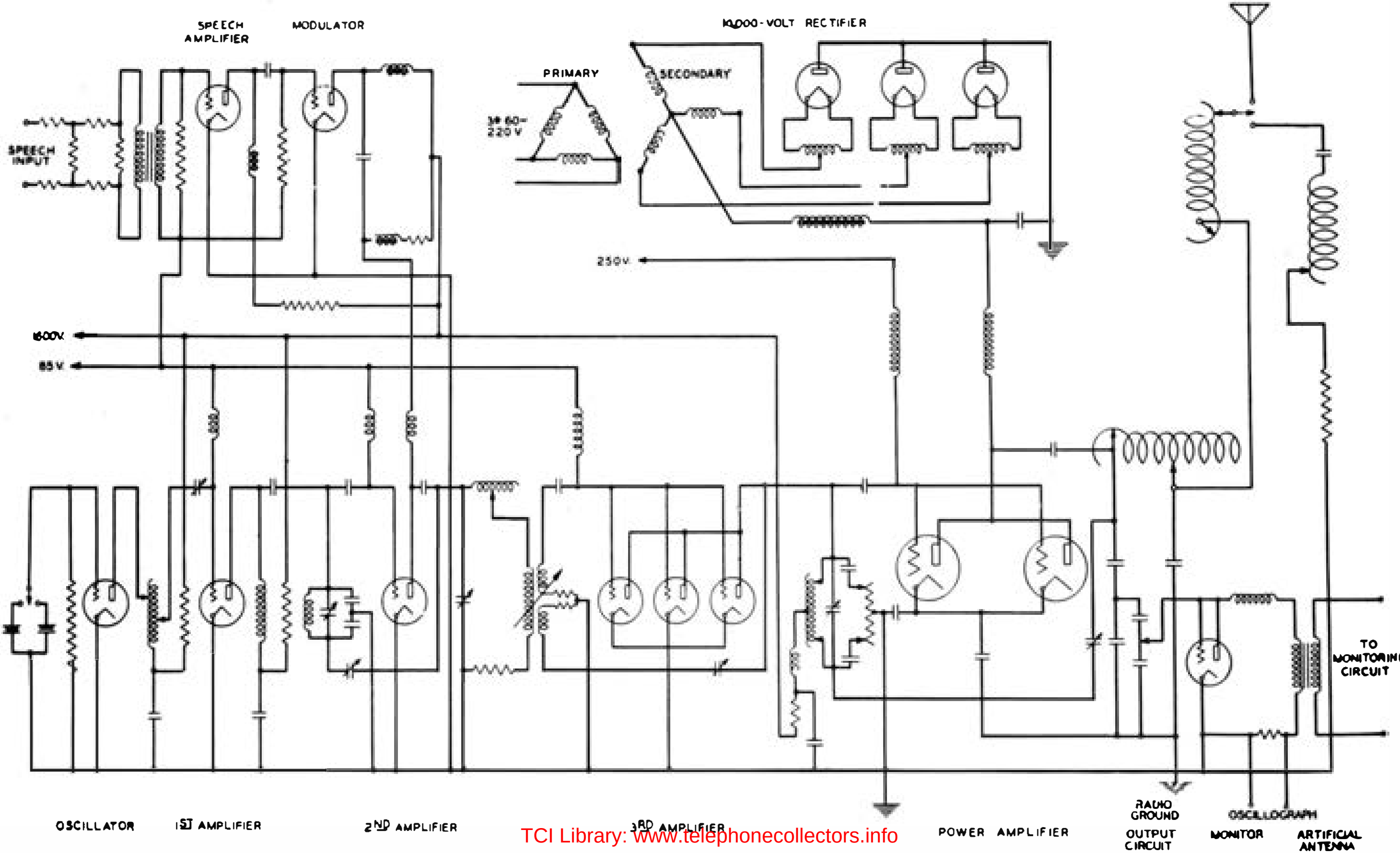


Fig. 5-56. No. 5C radio transmitter—simplified schematic. (From WECo Bulletin 410)

since it required highly linear high-power amplifiers, but it made possible better efficiency by achieving 100 percent modulation without overloading whereas the early transmitters with plate modulation produced noticeable overload distortion above 50 percent modulation. The low-level modulation technique also avoided the necessity for amplifying the audio signal to approximately the level of the radiated signal. As implemented in the 50-kW transmitter, the modulating amplifier was a 50-watt tube with audio signals supplied by a 250-watt tube which provided ample level for 100 percent modulation without distortion. The modulated wave was amplified by three push-pull power stages in tandem. The first power stage used two 250-watt air-cooled tubes, the second stage used two water-cooled tubes, and the last stage used six

Fig. 5-57. No. 6B radio transmitter.

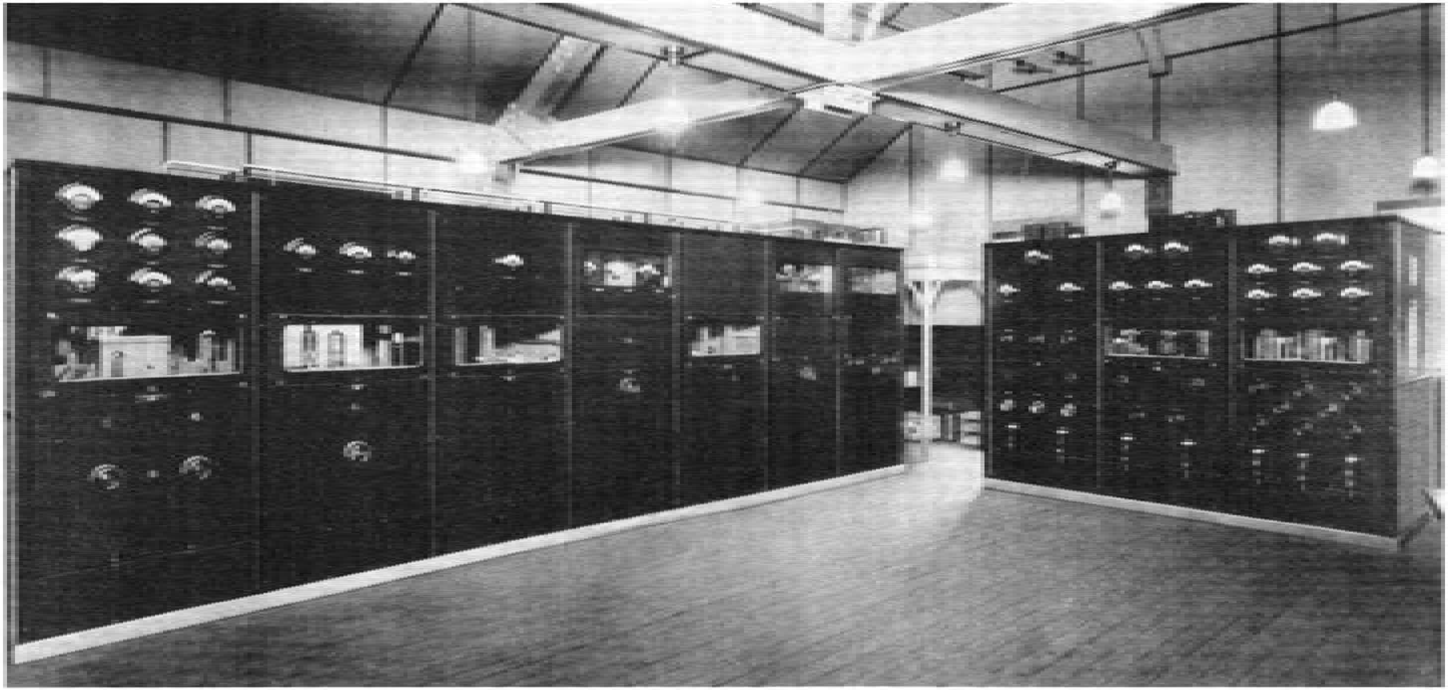


Fig. 5-58. No. 7A radio transmitter.

water-cooled tubes with 40-kW peak ratings. This technique, coupled with careful design, gave an audio output with harmonic content of less than 5 percent and carrier-frequency harmonics of about 0.03 percent.

Frequency control, using an oven-mounted quartz crystal, achieved a highly stable carrier practically free from short-term frequency modulation and having a long-term stability with ±30 hertz, well over an order of magnitude better than the then-current legal requirement of ±500 hertz.

Improved frequency response was achieved by careful control of audio components, particularly those in the input and coupling circuits. But this control was greatly facilitated by the use of low-level modulation which avoided the necessity for building very high power audio amplifiers with flat response at low frequencies. The net result of these measures was a frequency response flat within ±1 dB from 30 hertz to 10 kHz.

Some years later, and well beyond the time frame of this chapter in our history, a new form of linear amplifier developed by W. H. Doherty had a lasting effect on broadcast transmitter design. The Doherty amplifier, by using a unique method for adapting load impedance to fit the momentary output requirement, essentially doubled the operating efficiency with consequent large savings in power consumption, an important part of the operating expense of a high-power broadcast station.[62] The Doherty amplifier was first applied commercially in the transmitter built for station WHAS in Louisville, Kentucky. This station, built in 1938, had an installed power rating of 50 kW but was designed to drive a 500-kW amplifier and permit this expansion with minimum expense. The basic amplifier principle received wide application after this time.

Western Electric continued the manufacture of radio transmitters until 1948 at which time it was decided to concentrate Western Electric production facilities on meeting the heavy postwar demand for telephones.

5.3.2 Radiation and Propagation Problems

As we have noted, station WBAY with antenna atop the 24-story Walker Street building was not a technical success and within a month operation was transferred to a much lower antenna on the roof of the Western Electric building at 463 West Street. Here, considerably improved performance was achieved with an identical transmitter. This station, operating as WEAF, continued in successful operation for many years.

One of the problems with WBAY was a much lower signal than anticipated from the amount of power delivered to the antenna. The reason was that the steel frame of the Walker Street building provided an electrical structure which, unfortunately and quite by chance, resonated at the frequency used for broadcasting. Thus, much of the energy delivered

[62] The power saving on a 50-kW station amounted to about $6,000 per year.

to the antenna was absorbed by the building framework instead of being radiated into space. A substantial change in frequency would have solved this problem but this was not practical and the change to West Street was made since earlier experience had shown adequate radiation from its antenna.[63]

Another problem, emphasized by the poor radiation of WBAY, was the existence of very uneven propagation, with areas of low field strength within a few miles of the transmitter. Very soon it also became apparent that certain areas were marred by fading and sometimes by signal distortion even though these locations were no more distant from the transmitter than others entirely free from such transmission defects. These situations were not fundamentally improved by the shift in station location but the improved radiation and subsequent increases in power made the signal impairment more tolerable.

It was apparent that many defects in the received audio signal were not directly due to the equipment but resulted from signal impairment during transit through space. In these cases a better understanding and utilization of the transmission medium offered the only path to improvement.

Some of the deficiencies in the transmission medium were not unexpected by radio engineers since the propagation of radio waves at the broadcast frequencies was known to be affected by the electrical constants of the ground over which they passed and by obstructions such as hills or man-made structures which could attenuate waves and cast "shadows" behind them. There was, however, no theoretical basis for quantitatively evaluating the combined effects of these factors, and their magnitude and certain side effects were surprising.

Obviously, the first step toward a better understanding lay in an experimental program to determine the magnitude (field strength) and quality of the signal as a function of geographical location. Fortunately, the quantitative measurement of signal strength had been recognized long before as an essential part of developing the radio medium and suitable measuring techniques were available when needed.

Early in 1923, field strength measurements were made in the New York suburban area of signals from the West Street and Walker Street stations and also from a proposed site on the New York Municipal Building. This last, very tall, building proved to be an even poorer site than Walker Street. All locations showed heavy shadow areas to the northeast, the

[63] The WBAY location had been selected since previous experience had shown it desirable to use the highest antenna site available. The Walker Street building seemed ideal since it towered far above anything nearby. Unfortunately, the height of the building was close to one-quarter of the wavelengths then available for broadcasting (360 and 400 meters), about the worst possible height that could be used. Subsequent measurements showed large improvements possible with a change in wavelength (something over 20 dB for doubling) but such changes were not possible at the time and the only solution available was the use of a building considerably under one-quarter-wavelength high.

southeast, and the south due to the cluster of large steel buildings in the Grand Central and Wall Street areas. Later measurements around Washington, D.C., showed a much more uniform distribution around this city free from tall steel-frame buildings and also provided a measure of the shadows cast by the 1,000- to 2,000-foot mountains to the northwest. In 1925 a much more detailed survey was made of both areas. From these measurements it was possible to draw contour maps showing typical field strengths to be expected throughout the coverage areas as shown in Figs. 5-59 and 5-60.[64] After this work a field strength survey was recognized as a necessary preliminary to locating a broadcast station except in areas of highly uniform terrain.

In the surveys around New York it was found that the shadows cast by the buildings were limited in extent. At distances well beyond the shadowing structures, energy tended to be fed into the shadow zone from the sides and from above and on the average the field strength tended to return toward the amount expected for the distance from the transmitter. It was soon found, however, that this recovery in transmission was not entirely beneficial. The fine-grain exploration of field strength made in 1925[65] showed areas with the highly variable field strength contours, illustrated by Fig. 5-61, which are characteristic of wave interference. These contours, which showed wide field strength differences between points only short distances apart, were quite reproducible during the daytime. It seemed likely, therefore, that energy bypassing the Grand Central area and traveling up the Hudson Valley and Long Island Sound fed into the shadow zone in the Westchester area to the north of the city. These waves tended to reinforce or cancel each other, depending on the relative distance traveled. The fact that the patterns were stable suggested that the paths followed by the waves were determined by fixed features of the terrain.

At night, however, the situation was quite different. In the shadow zones, fading was extremely bad, even at short distances which normally would have reasonably stable transmission. What is more, it was found that fading was selective (i.e., frequency sensitive). Within the narrow limits of the audio band the field strength was highly variable both with frequency and with time; narrow bands of frequency would show near cancellation and these bands would shift about in the audio band, giving rise to a highly unpleasant distortion of the program. It was apparent that propagation at night, particularly in the shadow zones, was a highly complex matter with waves following several paths which differed

[64] As noted later, there was often a fine-grain variation superimposed on these "average" curves.

[65] In this survey (which appears never to have been completely published), measurements were made, so far as geographical features permitted, at 1-mile intervals around concentric circles centered on the transmitter. The radii of the circles increased in steps of 5 miles.

greatly in distance. Some waves undoubtedly traveled close to the earth but others must have followed a longer path reaching higher altitudes before returning to the point of reception. This longer path was probably constantly changing in length, thus accounting in part for the selective fading. Another contributor to selective fading was frequency instability of the transmitter. The carrier frequency tended to vary from instant to instant. There were rather large, slow drifts in carrier frequency and also extremely fast variations occurring within the carrier cycle. This latter effect was a form of frequency modulation superimposed on the amplitude-modulated signal. The net effect was a constantly changing wavelength which acted very much the same as a change in path length. As soon as this was appreciated, crystal control of frequency was added to WEAF with significant reduction in distortion in the shadow areas.

The very fundamental investigation of propagation, just discussed, was carried out in a very short time period. The first hint that radio was not the perfect transmission medium presumed by early experimenters was obtained in the early 1920s and more particularly about 1922. By late 1925 the Bell System had built highly sensitive and precise measuring gear and had made thousands of measurements which formed the basis for two milestone papers on transmission in the broadcast band.[66]

As a result of this work, it was apparent that broadcast transmission was greatly affected by terrain features. In a broad way these effects of terrain were predictable on theoretical grounds, but accurate, fine-grain predictions of performance required empirical data. (This is still pretty much true today.) To some extent the problems of radiation and propagation could be avoided or ameliorated by appropriate antenna siting. However, even with the most favorable location some problems continued to exist in physically complex areas. In such problem areas the distortion due to selective fading was greatly reduced by the introduction of precise frequency control but the final solution (so far as one exists) was the radiation of an adequate amount of power to give a field strength well above the existing noise level together with the use of receivers having adequate selectivity.

VI. THE ART OF MEASUREMENT

Measurements have always played an important part in science and its applications. They provide the means for confirming a hypothesis and also, when adequate theory has become available, they provide the primary data on which estimates of performance are based. In addition,

[66] "Some Studies in Radio Broadcast Transmission," by Ralph Bown, DeLoss K. Martin, and Ralph K. Potter, presented before the Institute of Radio Engineers, New York, November 4, 1925. Also published in *Bell System Technical Journal*, Vol. 5, No. 1 (January 1926). "Radio Broadcast Coverage of City Areas," by Lloyd Espenschied, *Bell System Technical Journal*, Vol. 6, No. 1 (January 1927).

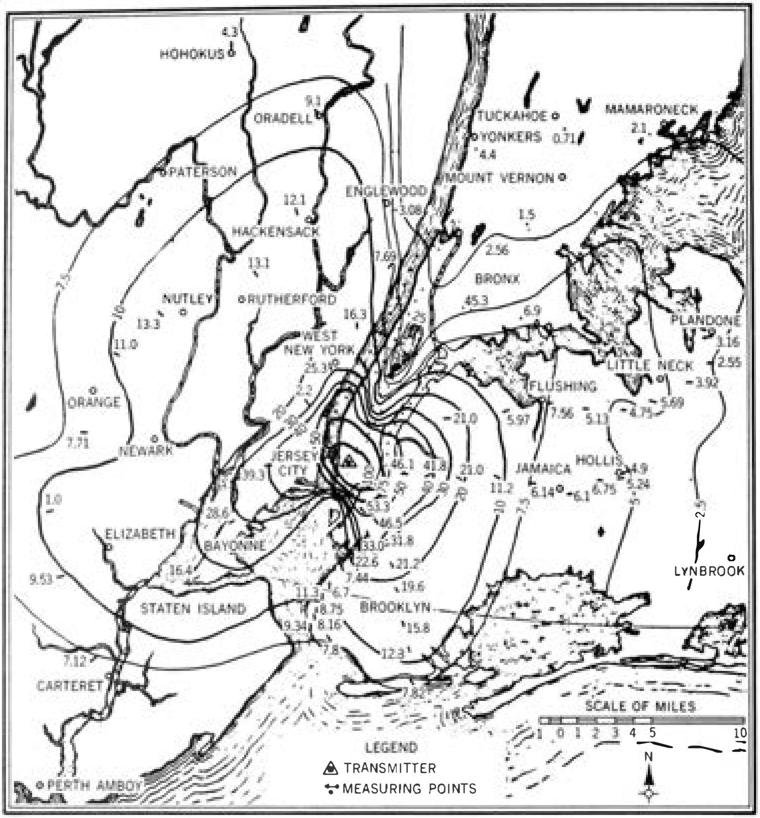


Fig. 5-59.

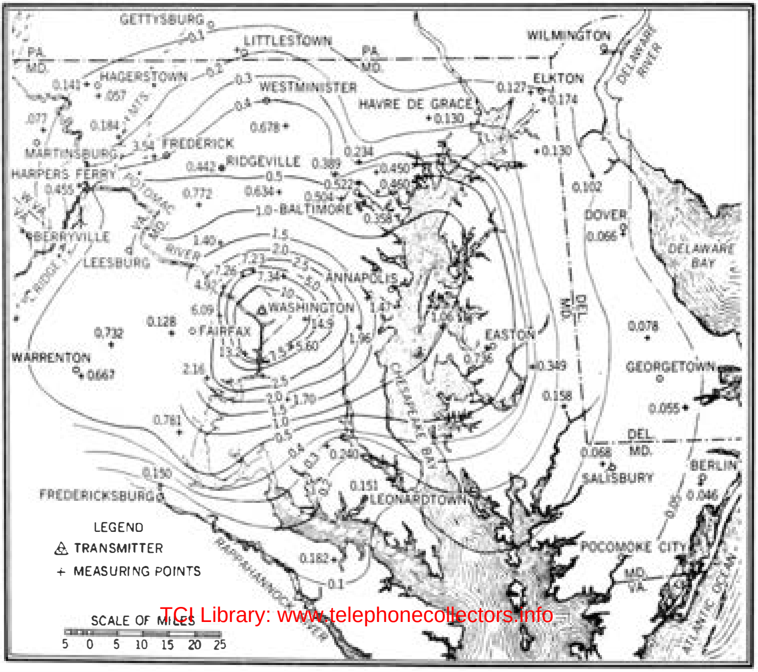


Fig. 5-60.

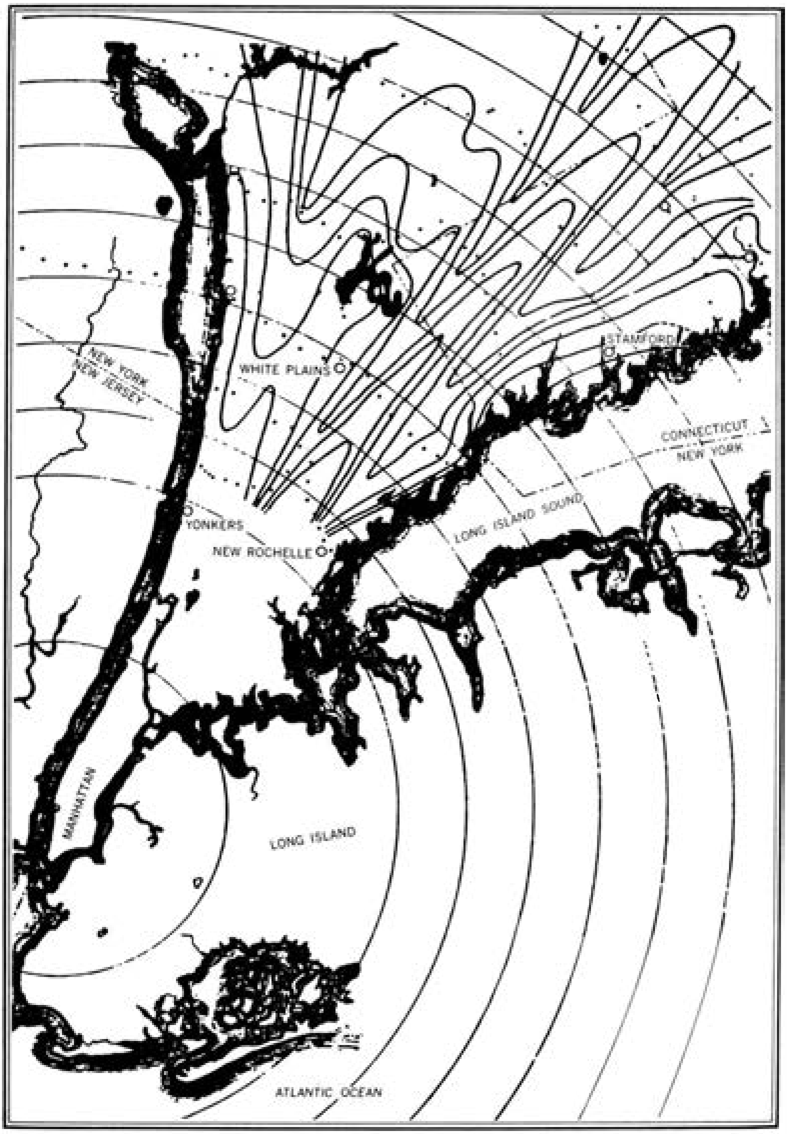


Fig. 5-61.

Fig. 5-59. Radio field strength contour map of New York metropolitan area (field strengths in millivolts per meter). (Bown and Gillett 1924, Fig. 8)

Fig. 5-60. Radio field strength contour map of Washington, D.C. and vicinity (field strengths in millivolts per meter). (Bown and Gillett 1924, Fig. 9)

Fig. 5-61. Radio contour map showing wave interference pattern to the north of New York City. (Bown et al. 1926, Fig. 1) in the area of application, they provide a means for monitoring performance to assure the maintenance of adequate service. In radio communication, measurements play an additional role in providing information on propagation needed for designing adequate radio circuits.

The need for empirical data on propagation is particularly necessary in radio because of the highly variable nature of the medium. This was also the case in the early days of wire transmission when the somewhat unstable open wire line was the principal transmission facility employed. With this medium, information on attenuation, noise pickup, and crosstalk had to be obtained largely by empirical means, but as the more stable cable medium was developed, and the theory underlying its performance became understood, performance could be accurately predicted from relatively simple measurements covering the primary constants of the line. In the case of radio, on the other hand, the prediction of performance in sufficient detail for economical engineering is often not possible on strictly theoretical grounds because of the extreme complexity of the medium. As we have noted, performance varies greatly with the frequency employed and even when satisfactory information has been derived for one frequency region, it has not been possible to extrapolate performance to adjacent regions since the mechanism of propagation may be quite different. Even when the mechanism is well understood, the theory developed to quantify it is usually adequate only for "average" or "typical" situations, and the large and often rapid departures from such conditions require a statistical description of performance which is only partly predictable by theoretical means. Thus, measurement of performance has played an important part not only in understanding the mechanism of propagation but also in providing statistical expressions of performance needed for designing radio circuits which will meet overall requirements in an economical manner. This need was particularly great at the time that the extremely complex high-frequency band was being developed for overseas communication but the need still exists today when the line-of-sight, microwave portion of the spectrum is the principal medium for radiotelephone communication.

The importance of measurements in the application of radio was recognized by the Bell System at an early date. As each frequency region was developed for use, Bell engineers played a leading role both in developing accurate measuring equipment and in using it to explore the propagation mechanism and performance. Some of this work has been touched on in previous sections but its importance warrants further coverage.

6.1 Developing Measuring Gear

Almost from the beginning, radio engineers recognized the need for quantitative measurement of the strength of the received signal and noise, but until the early 1920s adequate experimental tools were lacking. In the very early days of radio this lack did not greatly handicap application since performance of radio systems was largely limited by available technology. Systems could not be designed to meet specific performance requirements; instead they were built to utilize the best available equipment and the performance was defined after the fact by subjective judgments based on listening tests. The war years, with the intense pressure on practical results, made this the only feasible approach. As communication became less limited by technology and as commercial application brought cost factors to the fore, the need for economical design to meet a specified performance became possible and desirable. Fortunately the technological advances that brought about this change also made possible the development of suitable measuring gear.

The basic problem was to determine the most economical distribution of the gains and losses in a radio circuit among the various elements of the circuit, namely, the transmitting terminal (and its antenna), the receiving terminal (and its antenna), and the radio path between the terminals. The key to this problem was a means for measuring on an absolute basis (microvolts per meter) the strength of field from radio signal and noise. The development of equipment for making such measurements was begun in the Bell System by Englund and Friis about 1920, the work being carried out mostly at the Cliffwood, New Jersey, laboratory.

The basic technique for field strength measurement is simple in theory since it only requires the measurement of received current in an antenna of known constants. In some cases an ordinary open "flat top" antenna had been used. This made current measurements easy because of the relatively large amount of power collected by these large antennas but Englund and Friis felt that current measurements at low levels could be accomplished readily with electron tube techniques and they chose the relatively less efficient loop antenna because it was stable and portable and could be easily and accurately calibrated. This decision contributed largely to the success of their work. The measurement techniques were initially developed for long waves (first at 23.5 kHz and later at 57 kHz) but were soon extended into the higher frequencies.

The measurement of antenna current was based on the substitution scheme illustrated in Fig. 5-62. A known resistance, R, connected to the input of a radio receiver was inserted in series with a loop antenna (connections x, x) and the receiver tuned to the desired signal and adjusted for a convenient output level. The current in R, producing this signal, was then determined by disabling the distant (unknown) signal and adjusting a local oscillator connected to R (connection y, y) until it produced a signal equal to the unknown both in frequency and magnitude. The field strength of the unknown could then be calculated from a knowledge of the loop characteristics and the amount of current flowing

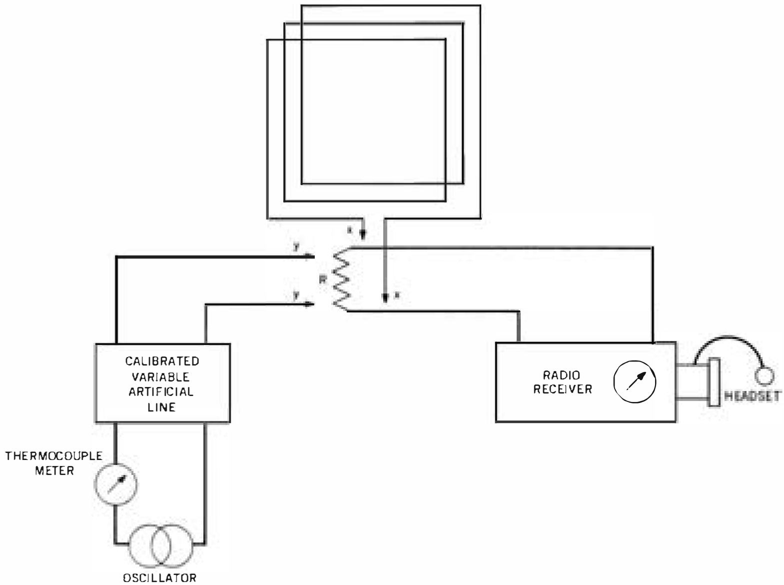


Fig. 5-62. Basic scheme for measurement of field strength.

in the resistance from the local oscillator. The loop current was very small and its accurate determination presented one of the major problems of field strength measurement. This was solved by adjusting the local-oscillator output to a fixed amount at a level high enough to measure with a thermocouple, after which the tone was passed through an adjustable, calibrated artificial line to achieve a comparison signal equal in magnitude to the unknown. Obviously the success of this method depended not only on an accurately calibrated artificial line[67] but also on assurance that all the current from the oscillator passed through this line and none was introduced into the calibrating resistance by crosstalk paths bypassing the line.

Equality of the unknown and comparison signals was initially determined by listening tests using a telephone headset fed by the final detector but later a meter in the detector plate circuit was used. At first the radio receiver was of the single-detection, tuned-radio-frequency type, but it was soon replaced by the superheterodyne which provided much more sensitivity and discrimination against unwanted signals.

[67] The production of an accurate artificial line usable over a considerable range of radio frequencies was a challenging problem since at these frequencies even minute inductances and capacitances associated with the components of the line (or their wiring) were significant.

All sets were battery operated, usually with separate supplies for the oscillator and the receiver.

In the transatlantic and similar tests, where the transmitter was under control of the experimenter, the comparison of known and unknown signals was accomplished rather easily by having the transmitter send alternate dashes and spaces. The known signal was then introduced during the spaces and adjusted to equal magnitude.

As the measuring range was extended to higher frequencies, the measurement problem became more difficult (particularly for weak signals) since it was difficult to prevent some of the local-oscillator current from entering the loop through crosstalk paths and thus introducing an error by bypassing the calibrating device. In the first broadcast measuring set this problem was solved by disconnecting the loop from the calibrating resistance, R, during the adjustment of the comparison tone from the local oscillator. This was satisfactory for laboratory measurements since the loop characteristics would remain constant for very long periods. In field measurements, however, the technique was less satisfactory since weather conditions affected the loop resistance. These resistance changes affected the current from the desired signal but not the comparison signal and thus introduced errors unless the loop constants were frequently determined by a rather inconvenient and time-consuming process. The problem was solved by building a very carefully shielded local-oscillator circuit to assure that all current was introduced in the loop through the calibrating circuit with negligible pickup through stray fields. With this well-shielded oscillator, the loop could remain connected during adjustment of the comparison tone and any change in loop resistance affected the unknown and comparison signals equally. Another minor problem in the measurement of broadcast fields was that the carrier was no longer under control of the experimenter but radiated continuously. With a loop antenna it proved easy to reduce the distant signal, while adjusting the comparison tone, by turning the loop to its null position at right angles to the position of maximum response used for measurement.

As might be expected, there is a considerable similarity in the appearance of the various field strength measuring sets developed in the 1920s. One of the earliest long-wave equipments is shown in Fig. 5-63. The sets for the shorter wavelengths were smaller and those used for broadcast surveys were made more suitable for transporting in motor vehicles. One of these, a second-generation equipment which became available about 1925, is shown in Fig. 5-64. In this figure, the superheterodyne receiver is at the top and the measuring equipment at the bottom, the variable artificial line for introducing known currents in the loop being at the left and the oscillator at the right. The latter is in a double-shielded compartment, with the outer shield enclosing all high-level equipment

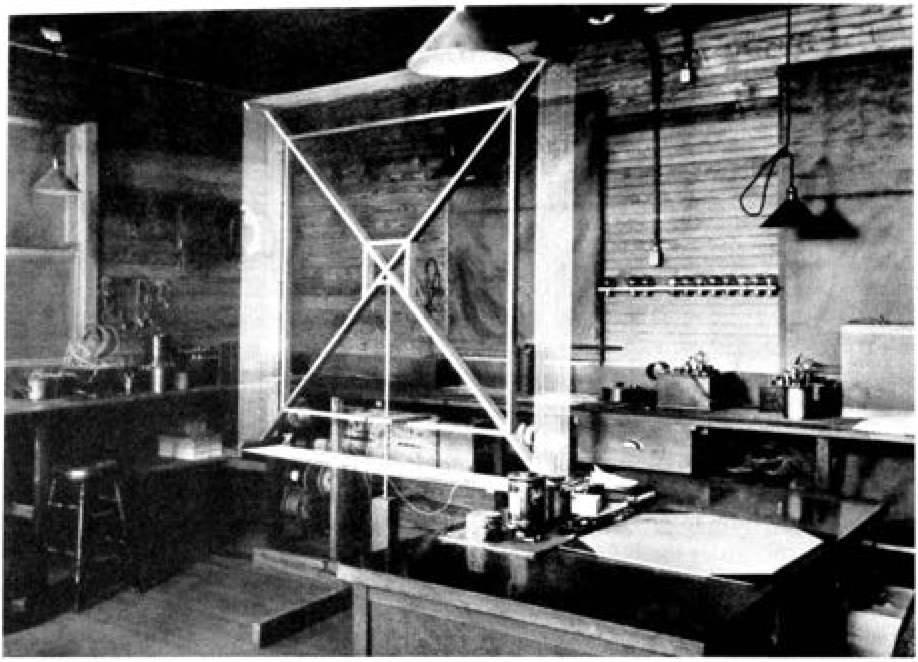
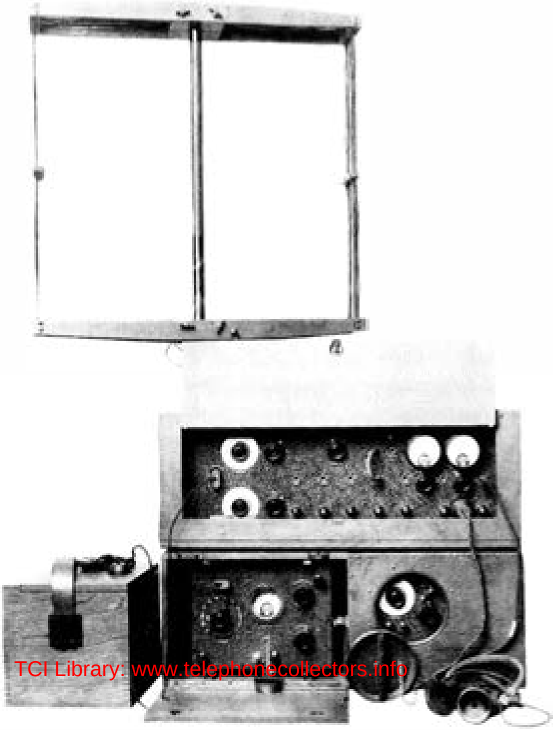


Fig. 5-63. Long-wave field strength measuring set showing shielding of local signal input system. (Bown et al. 1923, Fig. 3)

Fig. 5-64. Portable field strength measuring set of 1925. (Jensen 1926, Fig. 3)



during measurement. Access to the calibrating meter and frequency control was by means of a round port which is shown in the figure with its cover removed. This equipment is also shown in Fig. 5-65 mounted on a Ford roadster as used in the 1925 broadcast survey.

In addition to these early field strength measuring sets, arrangements were developed for recording signals over long periods of time and for other special applications. Other types of measuring equipment were also developed as they became necessary, such as vacuum tube voltmeters, oscilloscopes, and frequency measuring and monitoring gear. Some of this equipment is described elsewhere in this history (Chapter 4, Section 5.1.4) and had an important impact on the development of electronics. But important as these devices were, they are perhaps of less interest historically than the development and application of field strength equipment since the latter so clearly marked the end of the pragmatic approach to radio communication and the beginning of economical engineering design to meet preselected standards of performance.

6.2 Measurement Programs

Most of the early measurement programs have been mentioned in previous sections of this chapter which point out the part they played in

Fig. 5-65. Portable equipment mounted on a Ford roadster as used in 1925 broadcast survey. (Gillett, *Proc. IRE,* 1926, Fig. 2)

the specific developments covered. Accordingly, only a brief summary of this work will be given here.

The first comprehensive Bell System program covering the propagation of radio waves as affected by time and distance was carried out in connection with the ship-to-shore experiments described in Section 4.2.1. Much of the early work consisted of net loss measurements made by using a 1,000-hertz tone of known magnitude to modulate the transmitter. In so far as the gain of the transmitter and receiver remained constant, the 1,000-hertz tone at the receiver output gave a measure of the net loss; its variation with time and distance gave an indication of the variation in the propagation path. By July 1921 the development of field strength measuring equipment had advanced to the point where absolute measurements of field could be made. Some of the first data of this type, obtained during a southbound trip of S.S. *Gloucester*, are shown on Fig. 5-66. In this figure the dashed curve represents the prediction of the Austin–Cohen formula. The experimental results to the northeast, which involved a path largely over water, are in good agreement, but the path to the south, which skirted the shore, shows greater loss and indicates the greater absorption of such paths. In early 1922, field strength measurements were made over long distances as the signal from S.S. *America* was measured during its crossing of the Atlantic. The results, shown on Fig. 5-67, show clearly the very large difference between the day and night

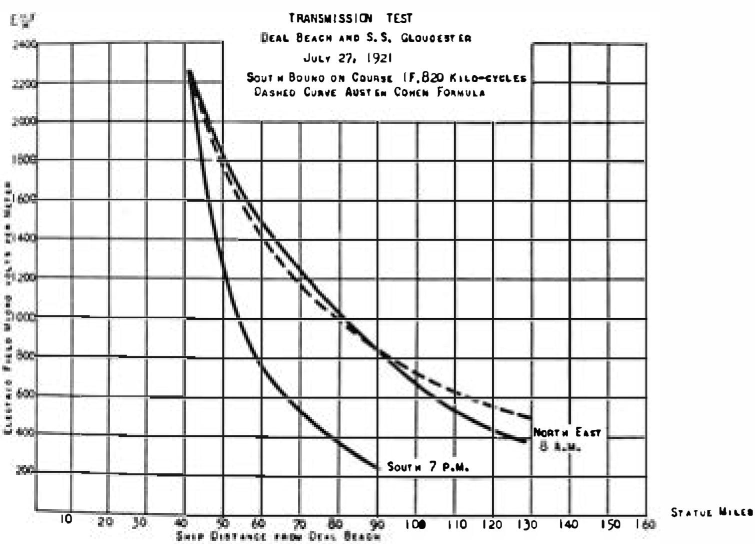


Fig. 5-66. Some of the first data obtained during a south-bound trip of S.S. *Gloucester*. (Nichols and Espenschied 1923, Fig. 18)

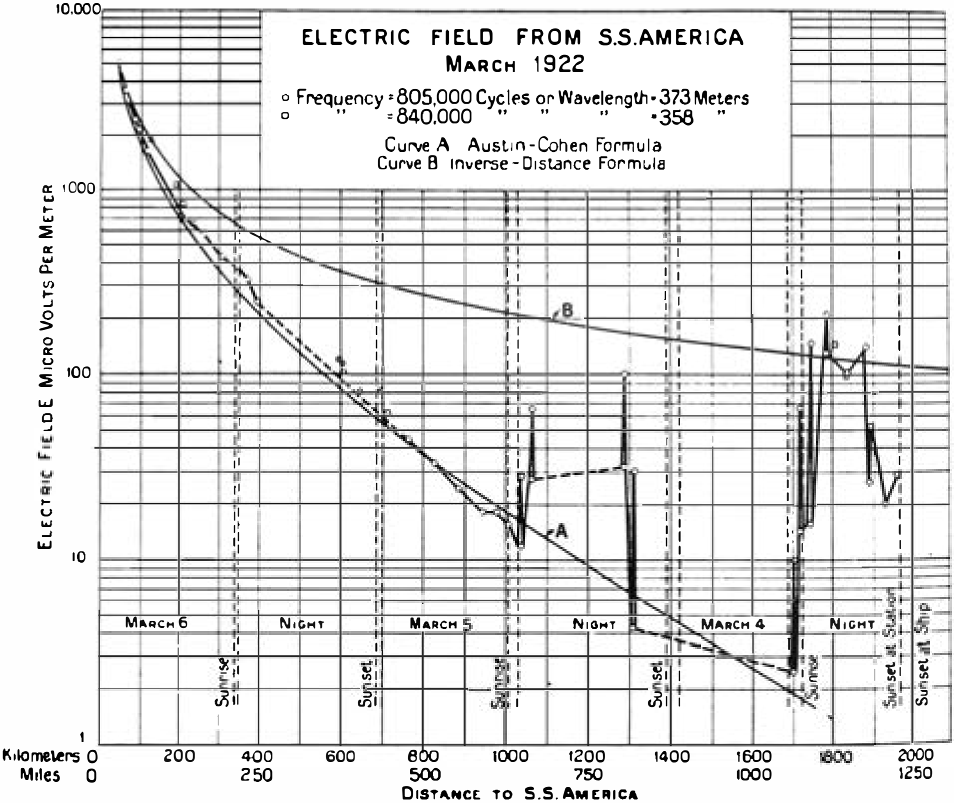


Fig. 5-67. Results of field strength measurements made of signal from S.S. *America* during its crossing of the Atlantic in 1922. (Nichols and Espenschied 1923, Fig. 20)

transmission at long distances with daytime results in good agreement with the Austin–Cohen formula and the nighttime data showing a highly variable signal with a field at times close to the prediction of the inverse-square law. These early tests had a profound effect on the radio engineers' views concerning radio propagation and demonstrated the great need for comprehensive surveys of radio propagation and noise measurement prior to establishing commercial radio systems.

Surveys of the transatlantic path at the low frequencies were begun in 1922 and those at high frequencies in 1925. These surveys have been discussed and some of their results illustrated in Sections 4.3.1 and 4.3.3.

At about the same time, the anomalous propagation of broadcast signals had been noted, and in 1923 the field strength measurements around New York and Washington, discussed in Section 5.3.2, were made. More comprehensive surveys were conducted in 1925 and about this time a sophisticated set of measurements involving the use of the recording oscillograph was made to investigate the mechanism of the irregular propagation and the several types of fading experienced in the New York City area. These tests and their analysis by Espenschied, Bown, Martin, and Potter showed quite clearly that the radio medium was far from ideal and was capable of introducing serious degradation and distortion of the signal. They also pointed out some techniques which could reduce this degradation.

It was also in 1922 that the poor radiation from the WBAY antenna on the Walker Street building was noted. Subsequent field strength measurements of signals from this and other sites quickly brought to an end the earlier belief that increasing antenna height was always beneficial and pointed out the important absorption effects resulting from resonance in steel buildings about a quarter-wavelength high.

The short period of 1922 to 1925 was in a sense the golden era of measurement. For the first time, quantitative information on the propagation of signals and the performance of antenna systems was being obtained at frequencies ranging from about 25 kHz to 25 MHz. These measurement programs were highly productive in explaining some of the important factors underlying the radiation and propagation of radio signals. As this understanding came about the need for the work lessened but it continued of great importance for a matter of ten years or more and even today much engineering data used in designing microwave systems is based on statistical data obtained from field strength measurement programs carried out since World War II.

VII. SUMMARY

By 1930, some 55 years after the Bell invention and 45 years after the work of Hertz, radiotelephony had achieved a small but important position in the Bell System. It was providing the only means then available for overseas telephony and would continue to do so for the next 25 years. A beginning had been made in providing telephone service to mobile stations, again a field where radio was (and continues to be) the only suitable medium. Point-to-point radiotelephony over land was on the whole not competitive with wire transmission and would not be for roughly another 20 years when the microwave field opened up. The radio use of most concern to the public, and the one with the greatest immediate potential, was that of broadcasting. The importance of broadcasting had been recognized by the Bell System in the early 1920s and over a period of some five years much had been done to develop the field both technically and commerically. However, the operation of commercial broadcasting stations (and particularly the proportionately large effort required for producing programs) had appeared inconsistent with Bell System communication objectives and it had been decided to withdraw from this field, continuing only as a supplier of broadcast transmitting equipment and program circuits for connecting stations into a network. Today, program transmission remains the only major broadcast endeavor in which the Bell System is engaged. It is interesting to note that, while for many years program circuits were provided by wire, the majority, particularly the television circuits, are presently carried by microwave radio.

The small amount of telephone communication carried by radio in 1930 was in no sense a measure of its importance to the Bell System. Through the preceding 15 years radio as well as wire technology had achieved a remarkable degree of maturity. The electron tube, the wave filter, and the carrier principle all had been developed to the point where their impact on communications would be felt for many years. In the evolution of this technology, developments in radio and wire transmission had gone hand in hand, work on each medium contributing not only to its own capabilities but also to those of the other. All through this period the problems faced by radio proved a great stimulant in developing the new electronic techniques which were proving so valuable in expanding wire transmission. It was this technology which would provide much of the basis for the extensive communication growth during the second half-century of the telephone. It is difficult to assign the precise contribution made to the field of electronics by the work on wire telephony and on radiotelephony but there can be no doubt that the work in the two fields together led to far greater progress than would have resulted from a development program restricted to one alone. Moreover, the work on wire and radio during this period had demonstrated the importance of broad technological experience in solving new and often unexpected communication problems when they arose.

In writing a brief history covering 50 years of rapidly developing technology, it is inevitable that the emphasis may not always be placed where it should be, or where the reader believes it should be. In concluding these two chapters on the evolution of the transmission media, it will perhaps be permissible to point out a few implications which may not have been given the emphasis that they deserve.

It is natural to emphasize physical things, the inventions and devices which have had far-reaching effect on the evolution of telephony, and the individuals primarily responsible. The years have produced many such, as for example Bell's magnetic telephone, the carbon microphone of Edison and Hunnings, the loading coil of Pupin and Campbell, de Forest's audion, and numerous others. Important as the invention of these devices was, the inventors must share some part of their credit with a large number of technicians who labored to reduce them to practice. To name only a few: The success of the loading coil was due not only to its inventors but also to the numerous engineers who over the years developed a practical configuration, appropriate magnetic materials, and economical manufacturing processes. The far-reaching effect of the audion arose not solely from its invention but also from the recognition by Arnold and Langmuir that it was an electronic and not ionic device, the application of the earlier discovery by Wehnelt of the low-temperature cathode, and the work of numerous technicians who found ways to manufacture this delicate device cheaply and in quantity.

Even when these important devices had been made practical, their usefulness was limited until means were devised for their application. The application of the audion, for example, required the development of circuits in which it could be used effectively to generate signals, modulate and demodulate them, and amplify them by useful amounts. And finally, all of these devices had to be integrated into a system to provide a complete communications channel, and these channels in turn had to be fitted into the complex system of telephony already existing and in a way that would not inhibit future growth. Back of this work were teams of technicians who unfortunately cannot be individually recognized but who in the aggregate were largely responsible for expanding communications capability.

In addition to the physical things, we should recognize what Lloyd Espenschied[68] calls the "leaven of the art," the "ideas which gave rise to [the art of communication]; the analyses and reductions to measurement which enabled results to be obtained by design." One of the many examples mentioned by Espenschied is van der Bijl's early study of the vacuum tube which derived expressions describing tube performances and ultimately led to his comprehensive 1920 book which was for many years the authoritative text used by designers of tube circuitry. Again, it was the analysis of the modulated wave by Carson that led to the extremely useful concept of a carrier with sidebands that underlay much of the development of radio and carrier systems.

Going beyond these technical contributions, we should not overlook the part played by the general public and the technically enlightened entrepreneurs who could see a future in this complicated and initially little-understood field of communications.

Finally, we should note that some readers may find undue emphasis on the smooth and continuing progress of technology. It may all appear too simple and easy. Let us hope that the understanding reader will appreciate that this work was carried on by people who shared the faults as well as the virtues of mankind. In real life, progress never fol-

[68] The quotations are from a paper presented to the Institute of Radio Engineers on May 10, 1937.

lows a direct line to its objective. There are frequently excursions down blind alleys that must be retraced. Often each dozen steps forward are accompanied by several backward. For every dozen good judgments, there are a number of poor ones. All these frailties were to be found in those participating in these events. Perhaps our narration would have been more spritely (and it certainly would have been much longer) if we had chosen to follow each step and presented the conflicts and differing points of view as they arose during the 50 years covered. Instead, we have chosen to look only at the net result and attempt to identify some of the important milestones along the path. The fact that progress during this period was very considerable implies in no way that the road was simple and straightforward or that the end was achieved without error or conflict. Instead, it would be fairer to say that the contributors to the evolution of communication showed a remarkable ability to rise above the normal impediments to progress. In reading the literature of the period, and in talking to those who lived through the latter part of it, the impression is inescapable that this ability was based on the confident belief in the effectiveness of technology in overcoming seemingly impossible obstacles and the equal confidence that the ultimate aim was a worthy one.

Chapter 6

Switching and Signaling Systems

Switching and signaling systems have been developed to establish direct communication between telephone users. In the first century of telephony, these systems have progressed from almost totally manual to almost totally automatic. The first 50 years of this development are related in this chapter.

Early communication experience provided little background on which to base the development of switching and signaling systems, but gradually the concept evolved of a number of distributed central offices, each containing switching devices and supporting equipment, for economically interconnecting telephone lines. This plan was first used for telephone users in a local area, or exchange, and was later extended to the interconnection of these exchanges over long-distance lines to form the nationwide network foreseen by Alexander Graham Bell in his "Grand System." The exchange systems also were extended to include private branch exchange (PBX) systems, usually located on a customer's premises, which not only provided internal communication but also gave access to the common-carrier network. At first the switching mechanisms were operated manually and required complicated actions on the part of both the customer and the central-office operator. Gradually the mechanisms were simplified so that many of these operations were carried out automatically even in those offices employing an operator for establishing connections. While this progress was being made in the "manual" offices, much effort went into the development of mechanisms for automatically interconnecting lines under the control of the user. The field for such "machine switching" was at first greatly restricted by both economic and technical factors but by the end of telephony's first 50 years many of these problems had been resolved for exchange plant applications. By 1925, although toll service was still handled manually, about 12 percent of Bell lines were switched by machine in the exchange plant. More importantly, the first steps had been taken toward the sophisticated switching methods, involving the common-control principle, which laid the groundwork for the almost complete mechanization of switching in the next 50 years.

I. INTRODUCTION

The transmission systems described in the two preceding chapters provide the paths over which electric analogs of speech waves are carried. This chapter describes the switching and related arrangements used to connect individual paths together into a network so that anyone, anywhere, can talk to anyone else, anywhere else in the network, at the desire of the calling party.

To understand the evolution of switching systems it will be helpful to start by examining the network configurations capable of providing this service and the requirements placed on the switching and signaling arrangements used for interconnecting the component circuits.

II. BASIC CONCEPTS

The very earliest telephone communications were carried out over what today we would call party-line or "full-period" circuits. As illustrated in (a) of Fig. 6-1, two or more telephone instruments were connected to a single telephone line and anyone on this line was able to converse with anyone else (and everyone else) so connected. Such arrangements are still used for some data-transmission or telephone circuits where the community of interest is limited and communication between specific users is maintained over a long period. But, obviously, the usefulness of this configuration is limited. It suffers from lack of privacy since any station can monitor all communications carried by the line. In addition, the total number of stations that can be served is severely restricted since, for its duration, a call between any two users denies service to all others connected to the line.

It was apparent from the beginning that there was need for a service that went far beyond the capabilities of a single, party-line arrangement. Bell saw immediately that telephony, unlike telegraphy, required no special skills on the part of the user and thus offered the potential for a widespread communication service. To accomplish this, it was necessary to select a specific path connecting the desired users, preferably with complete or nearly complete privacy, in a way that would have little effect on other users. One possibility, illustrated in (b) of Fig. 6-1, would be to provide a line from each telephone station to every other station, with means at each location whereby the user could connect, or switch, his instrument to the desired line and exclude all others. This "station" switching overcame two of the objections of the single-line system by providing privacy and leaving the unused part of the network available for other communications when portions were in use. The price paid for the improvement was high because a completely separate line was required

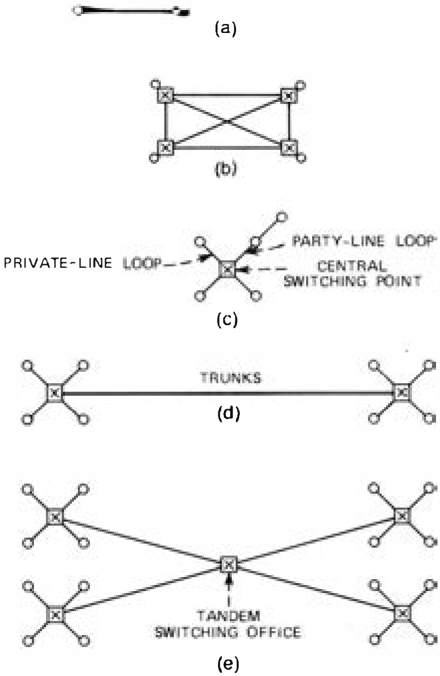


Fig. 6-1. Methods for interconnecting telephone users: (a) party-line or "full-period" circuit; (b) station switching; (c) centralized switching; (d) centralized switching with trunking; (e) centralized switching of loops and trunks—tandem trunk switching.

from every station to every other. Thus, except for very small networks, the arrangement was impractical since the number of lines increased roughly in proportion to the square of the number of stations. However, station switching has received some use and is still employed for intercommunication systems, where the distance between stations is small, and for some special-purpose data networks.

A much more practical arrangement, which came into use only a few years after Bell's invention, was the centralized switching scheme illustrated in (c) of Fig. 6-1.[1] With this arrangement each station (or a small number of party-line stations) had a direct line, called a "loop," to a central switching point at which the necessary connections or switches were made to set up a path between those stations desiring to communicate. The connection was established only for the duration

[1] As related in Section 3.1, this arrangement had been patented as early as 1851 for telegraph interconnection but had been forgotten and the concept evolved anew in the 1870s and 1880s.

of the call and, so long as adequate interconnection facilities were available, any of the stations not already carrying on communication were available for interconnection. The number of lines required was at most one per station and the length of these lines could be minimized by selecting a switching point near the center of the network. The switching point was commonly called a "central office." Initially, it was sometimes called an "exchange," but later this term was used in a broader context.

Centralized switching was an excellent means for handling interconnection in small geographic areas but left something to be desired in covering large areas since each station required its own line to the central switching office. For large cities, these lines could be long and costly, and basically uneconomical, since most of them were used for only a few calls a day. The party-line concept, in which a number of stations were assigned to the loop, reduced the cost per station but with some loss of privacy and availability. The solution, illustrated in (d) of Fig. 6-1, was the use of small central-office areas interconnected by lines called "trunks." All intercommunication within a given central-office area was set up by means of a switch at the office. Communication to a different office was accomplished by switching a loop at the initiating office to a trunk extending to the terminating office, at which point it was connected to the loop of the called station. This arrangement, by slightly complicating the switching procedure, kept the loops short and provided an economical means for utilizing the longer trunk circuits between offices. Obviously, even in the smallest network employing only two offices, the number of trunks could be far fewer than the number of loops. One reason was that many calls were confined to the originating office and, even on interoffice calls, the probability was small that a large number would be made simultaneously. Thus, due to the randomness of traffic, it was possible to use trunk lines more economically than loops because they could handle many more calls in the course of a day. (Roughly, the trunks in a large system numbered about one-tenth the number of loops.)

Once the trunking principle evolved, it was natural to extend it. In large cities, "tandem offices" were introduced, as shown in (e) of Fig. 6-1. These offices switched only trunks, thus reducing the number of trunks required for direct interconnection of all offices. The tandem scheme was used particularly for connecting rather remote portions of the city having only a small community of interest while direct trunking, shown in (d), continued to be used for connecting the nearer offices and those with high traffic density. Therefore the trunking pattern in a large city tended to be a combination of the configurations shown in (d) and (e).

The centralized switching schemes just described were those used largely for communication within a city or its environs. This type of area is usually referred to as providing "exchange" service as distinguished from intercommunication between cities commonly called long-distance, or "toll" service.[2] The long-distance network is largely based on the extension of the tandem principle used in large exchange areas. In each city there is usually a "toll office"[3] having long-distance lines or trunks extending to other cities. All intercity calls are carried over toll connecting trunks from the local office to the toll office which provides means for determining the special charge or toll made for the long-haul call and also provides a switching mechanism for connection to the appropriate intercity line either directly or through a tandem switch.

The toll network is illustrated in a highly generalized manner in Fig. 6-2. In the particular configuration shown, City A has direct toll trunks to Cities B and C, but if communication is desired to City D, toll trunks must be switched together in tandem at either B or C. In theory at least, B could be reached from A not only by the direct trunk linking the two cities but also by tandem switches at C and D. In practice, this use of alternate routing between cities is a valuable means for avoiding breaks in the line or congestion in the network but must be used under rigorously prescribed conditions to avoid an excessive number of trunks in tandem, each of which can add its quota of attenuation or other transmission impairment. The art of designing a complex network to provide the desired degree of circuit availability at minimum cost with adequate protection against path failure is a complicated matter beyond the scope of this history. Suffice to say that it has been a long and continuing study which has contributed greatly to the rapid and flexible system of communication available today.

In addition to the exchange and toll switching systems, a third type known as the private branch exchange, or PBX, formed a part of the comprehensive telephone network. The PBX was much like a small exchange office but was located on customer rather than telephone company premises. The customer was usually a business organization and the PBX switched two kinds of traffic. Most of the traffic was internal, between customer stations, and was handled locally

[2] The distinction between exchange and toll switching cannot be defined very rigorously but generally speaking the former covers calls for which the charge on all or at least a minimum number is governed by a base charge for service. It sometimes includes calls on which a small charge is made, usually in the form of metered or bulk billing. Toll service involves calls which are billed on an individual basis, the charge being determined by duration and distance.

[3] In a city with only a single central office, the toll office might be merely a specially equipped portion of the switchboard.

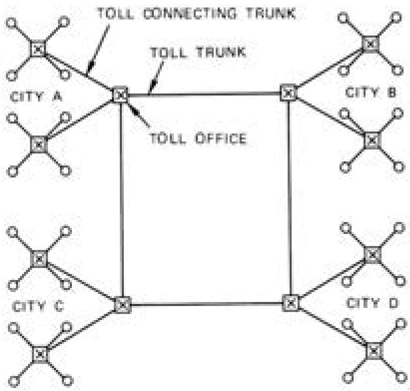


Fig. 6-2. Generalized toll network.

without entering the common-carrier network. However, the PBX also connected the customer stations to the exchange and toll networks by switching them over loops (commonly called PBX trunks) to a nearby telephone company central office.

The various arrangements for centralized switching, just discussed, formed the basis for the expanding telephone network and continue in use today. Over the years, the means for making the interconnection has varied greatly. In the beginning it was accomplished entirely by the manual manipulation of interconnecting devices. Originally, these manipulations were complex and laborious and much effort was devoted to reducing complexity by minimizing and automating the required operation. As early as 1879, proposals were made for automating central-office operation under control of the telephone user. These machine arrangements evolved more slowly than manual switching but with some few exceptions have gradually superseded manual central-office operations. However, the problems of complete automation were formidable and took many years to solve. The steps toward automation taken during the first 50 years of telephony are discussed in Section IV.

Regardless of the manner in which the interconnection is carried out, there are a number of basic requirements which all switching systems must meet. These are:

(*i*) Means for notifying the central office (either the operator or the machine) that a subscriber wishes to make a call.

(*ii*) Means for indicating the station with which communication is desired.

(*iii*) Mechanisms for making the desired interconnection between the loops of the calling and the called subscriber, either directly or via a trunk, for the duration of the call.

(*iv*) Means for indicating to the called subscriber (if not already engaged) that a call is waiting.

(*v*) Means for indicating the status of the connection to the calling party (i.e., ready to proceed, or unavailable for use due to absence of the called party, busy line or station, and so forth).

(*vi*) Means for indicating the end of the call so that the interconnection can be terminated.

These are the minimum requirements for any switching system and were soon incorporated in all successful systems. Some were accomplished rather readily. Others, originally complex, were gradually simplified as the art evolved. It is interesting to note that of the six items listed above, only (*iii*) is strictly an interconnection, or switching, function. All the others involve signals to convey information on the destination and status of the call. Thus the art of signaling is a large and essential contributor to the development of any switching system. But signaling is so closely related to interconnection equipment and methods that it is not always practical to discuss separately these apparently different functions of a switching system.

As telephone communication developed, other requirements evolved. For example, charges for some services were on the basis of number of calls made or their duration. For such situations, means were needed for counting or timing calls, and these and other special requirements will be covered as telephone services are discussed in detail.

III. MANUAL EXCHANGE SWITCHING

3.1 Contribution of Prior Art

Prior to the invention of the telephone, telegraph technicians had developed a transmission system consisting of spaced iron wires supported on insulators mounted on poles or brackets 100 or so feet apart. As related in Section II of Chapter 4, this system was not optimum for telephony and was soon greatly modified and ultimately supplanted by other systems. However, it did serve as a useful transmission medium for the early years.

In the case of switching, the prior telegraph and related signaling art had somewhat less to offer. Certain apparatus items (such as vibrating bells and buzzers, simple switches, and annunciator drops and relays) provided components for early switchboards and signaling systems, but there was little experience to assist the development of basic concepts. The reason is that telegraphy had developed as a specialized, rather than a universal, communication medium. The manual sending and reception of telegraph signals required skill and long training and the average person had found the reward not worth the effort; mechanical sending and print-out devices were also available but were crude and expensive. However, the press, military, and railways found high-speed communication of sufficient value to employ skilled telegraphers and built networks fitting their needs. On the whole, these were simple full-period circuits, often with many stations on a line. Switching was not required or, if employed, was used to set up circuits for rather long periods, much as a patch board might be used today.

The needs of the general public were met by the message telegraph service. The sender delivered his written message to a nearby telegraph office and a written copy was delivered by messenger to the recipient. Since the messages were usually short and could often be sent in a fraction of a minute, there was no need for an elaborate switching system for setting up a direct path from the originating to the terminating office. Instead of switching circuits, message switching was employed. That is, the message was transmitted toward its destination as far as a direct line was available, at which point it was translated back into alpha-numeric form. When a circuit became available, it was retransmitted further, the process being repeated as many times as necessary to reach the final office. This was a highly efficient process since by introducing only minor delays at points of retransmission, it was possible to keep the circuits in operation during a large part of the day and still provide message transmission over the lines in a very short time compared to that required for pickup and delivery at the terminals. Even with moderately long delays in transmission, messages traveled rapidly as compared to the physical transportation required by the mails.

Obviously, the various types of switching as developed for telegraphy could contribute little to meeting the very different requirements of telephone switching, although, as with all generalities, there were exceptions. Shortly before the invention of the telephone, a switched telegraph service was offered to and used by a number of law firms in New York City. These organizations had enough need to communicate among themselves to justify the use of trained telegraphers. The addition of the telephone to this network was a natural step and a considerable amount of early telephone service grew out of the so-called "Law System." Even the written message service contributed to the growth of telephone switching in an indirect way. The need to convey these messages to a telegraph office was naturally a nuisance, particularly to business firms with an appreciable amount of service but not enough to employ a full-time messenger. To meet such needs, the so-called district telegraph service was established as a separate organization or, in some cases, as part of the message telegraph company. In providing this service, wires were run from a central point to the subscribers to the service, largely on a party-line basis. Each subscriber was provided with an automatic transmitter (or "call box") operated by a spring motor. When the spring was wound, it actuated a transmitter which sent a coded message to the central point, designating the point of origin, and indicating the need for a messenger who was subsequently dispatched from the central office. Variants of this scheme included means for sending fire and burglar alarms and a choice of coded messages to indicate not only the need for a messenger but the desire for package pickup or other service. Still other arrangements used a central office with wire networks, again largely on a party-line basis, to broadcast information such as stock prices and financial clearing house reports. Many people operating services employing a network of wires saw the telephone as an opportunity for expansion and did indeed provide a basis for the early growth of telephony. Their contribution to the art of switching is, however, at least debatable. There was a strong tendency to look on the telephone as a useful adjunct to the existing service, and many of the telephone applications were based on fitting it into an existing service pattern rather than developing a new service to take advantage of a unique and different means of communication. This Procrustean approach may well have restricted the evolution of the switching art even though the services provided undoubtedly did much to increase public exposure to the telephone.

Finally, we should not close this review of the early art without mention of Francois DuMont. In 1851, when the electric telegraph was little more than 15 years old, DuMont took out a comprehensive British patent (No. 13,497) covering a system of telegraph intercommunication with centralized switching of both loops and trunks much like the system that evolved years later for telephony. DuMont's proposal must have been known to the telegraph community of the day but seems to have been promptly and completely forgotten, probably because it was so unrealistic as compared to current needs and technology. Thus this inventor had great foresight but little influence on the telephone network that gradually evolved much later under the logic of need and with little influence attributable to any single individual.

3.2 Early Networks and Switchboards

Much of the early history of the application of telephony was not documented at the time and our present knowledge depends to a considerable extent on recollections, often conflicting, recorded many years later. We shall do our best to present here the probable course of events but will not pretend to be definitive. At this late date it is probably impossible, even with the best of intentions, to resolve contradictory statements made by those writing much closer to the event.

Several telephone networks came into existence as early as 1877. One of the earliest was set up by E. T. Holmes in Boston during May of that year by connecting Bell telephones to wires already in use for a centralized burglar-alarm system. A small plug-type switchboard, shown in Fig. 6-3, was installed at the central office. Holmes, in an undated but signed manuscript which probably was written in 1883, states that on May 17, 1877, two lines were connected through the switchboard and conversations carried out. Similar interconnections were made later and although these communications were of historical interest, undoubtedly being the first switched telephone calls, they added little to the art of switching since they appeared to be viewed more as an experiment than as the beginning of a new service. This is evident from the conclusion of the Holmes document that states, "It was however obvious that an express telephone business substantially as now conducted (in which subscribers requiring an express [man] telephone to the central station from whence the order is telephoned to the General Express office) would prove of more immediate pecuniary value, and that system was therefore immediately adopted as the first field of the telephone in Boston." Thus it seems that the first switchboard was used only incidentally for interconnection of subscribers and primarily as a means for transferring a line from the burglar-alarm mechanism to a telephone at the central office.[4] In the following year a new switchboard for interconnection seems to have been added by Holmes and for a number of years

[4] In his autobiography, *A Wonderful Fifty Years*, published in 1917, Holmes states that he was the first to suggest the use of direct connection of users through a central switchboard. Letters reproduced in the book show that in July 1877 he was negotiating a license agreement with Gardiner Hubbard (of the Bell Telephone Company Trusteeship) which included this as one of the permissible uses of the telephone. However, a number of other people seem to have had the same idea at about this time and, of course, the concept was basic to Bell's "Grand System" mentioned in Section X of Chapter 1. Holmes does, however, deserve full credit for the first commercial use of the telephone in his despatch service.

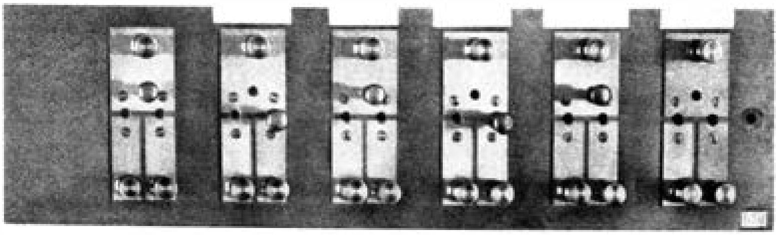


Fig. 6-3. E. T. Holmes switchboard of 1877. (Rhodes 1929, Fig. 30)

he conducted, as separate entities, both a switched exchange service and his express despatch service, the latter using lines installed primarily to furnish burglar-alarm service after the close of the business day.

In July of 1877, Isaac Smith, a Hartford druggist, established a telephone network for connecting his store with several doctors and livery stable operators. At first this network relied on the use of an operator to relay messages verbally, but later some primitive arrangements seem to have been used for interconnection. There is evidence that this scheme relied heavily on a party-line arrangement.

Both of these networks were originally intended for private rather than commercial use. A third such system was put in use in 1877 by the ingenious Thomas Doolittle, whose name appears frequently in early telephone history. In 1874, he had initiated an organization called the "Bridgeport Social Telegraph Association" made up of friends and business associates in the area. Members were connected by wire to each other and to the local commercial telegraph company, the manager of which taught them the Morse code. In the summer of 1877, Doolittle secured a subagency for the Bell telephone but it was fall before he obtained instruments. He then began to install telephones on the lines of the Association, and limited service started with a simple interconnecting switchboard. Doolittle seems to have had line interconnection as an objective from the beginning, but service was confined to the private network until July 1878 at which time a switchboard designed by Doolittle and built by Charles Williams, Jr., of Boston became available and commercial service was offered.

The first truly commercial exchange was established in New Haven, Connecticut, in 1878, by George W. Coy, a licensee under the Bell patents. His company, the District Telephone Company of New Haven, was formed on January 15, and almost immediately distributed 1,000 copies of a prospectus describing a public switched service which would be made available at a cost of $4.50 per quarter. This resulted in only one application for service, made by the Reverend John Todd. With the help of direct solicitation, other customers were obtained and the exchange was formally opened with 21 subscribers on January 28, 1878. On February 21, the company issued its first list of subscribers, the number at that time having grown to 50.[5] This, the world's first telephone directory, reproduced in Fig. 6-4, was of the classified type giving no addresses or telephone numbers, the latter not coming into use until about ten years later.[6]

[5] Appropriately enough, the first name listed was that of the Reverend John Todd. It is interesting to note that 70 years later, 14 of the original subscribers (or their corporate successors) were still listed in the Southern New England Telephone Company directory.

[6] As late as 1880, the list of New York City subscribers did not show telephone numbers, even though it consisted of about 1,500 names distributed among six exchanges.

LIST OF SUBSCRIBERS

New Haven District Telephone Company

OFFICE 219 CHAPEL STREET

February 21, 1878.

Residences.

Rev. JOHN E. TODD.
J. B. CARRINGTON.
H. B. BIGELOW.
C. W. SCRANTON.
GEORGE W. COY.
G. L. FERRIS.
H. P. FROST.
M. F. TYLER.
I. H. BROMLEY.
GEO. E. THOMPSON.
WALTER LEWIS.

Physicians.

Dr. E. L. R. THOMPSON.
Dr. A. E. WINCHELL.
Dr. C. S. THOMSON, Fair Haven.

Dentists.

Dr. E. S. GAYLORD.
Dr. R. F. BURWELL.

Miscellaneous.

REGISTER PUBLISIHNG CO.
POLICE OFFICE.
POST OFFICE.
MERCANTILE CLUB.
QUINNIPIAC CLUB.
F. V. McDONALD, Yale News
SMEDLEY BROS & CO.
M. F. TYLER, Law Chambers.

Stores, Factories, &c.

O. A. DORMAN.
STONE & CHIDSEY.
NEW HAVEN FLOUR CO. State St.
" " " " Cong. ave.
" " " " Grand St.
" " " Fair Haven.
ENGLISH & MERSICK.
New Haven FOLDING CHAIR CO.
H. HOOKER & CO.
W. A. ENSIGN & SON.
H. B. BIGELOW & CO.
C. COWLES & CO.
C. S. MERSICK & CO.
SPENCER & MATTHEWS.
PAUL ROESSLER.
E. S. WHEELER & CO.
ROLLING MILL CO.
APOTHECARIES HALL.
E. A. GESSNER.
AMERICAN TEA CO.

Meat & Fish Markets.

W. H. HITCHINGS, City Market.
GEO. E. LUM.
A. FOOTE & CO.
STRONG, HART & CO.

Hack and Boarding Stables.

CRUTTENDEN & CARTER.
BARKER & RANSOM.

Office open from 6 A. M. to 2 A. M.
After March 1st, this Office will be open all night.

Fig. 6-4. Facsimile of the first telephone directory.

This information on the New Haven exchange is well documented, but the physical characteristics of the switchboard are less well established since the board formed part of the partition between the office and battery room and did not survive subsequent building alterations. Available information depends on a sketch made from memory by Coy in 1905 and his statement that it was similar to a board, first used in Meriden, Connecticut, in February 1878, which has survived. On this basis the reconstructed model shown in Fig. 6-5 and the circuit diagram of Fig. 6-6 are usually taken to be representative of the equipment first used for commercial switched service in January of 1878.[7]

The many interesting features incorporated in this board warrant a close examination of Fig. 6-6. For illustrative purposes, we shall assume that a caller on Line 1 wished to call a party on Line 4. Normally, all switches designated L and C were open and all those marked K were closed. In this situation, battery current from B-1 was sent over each line to a ground at each station. When the sub-

[7] J. E. Kingsbury in his 1915 book, *The Telephone and Telephone Exchanges*, says categorically that a somewhat smaller (four-line) board with different operating features was used and he illustrates the circuit with a drawing he obtained from T. D. Lockwood. He states that the original board was replaced in about two months by one similar to the Meriden board which also was designed by Coy. The precise sequence of events is of little importance since, in either case, Coy deserves credit for introducing many fundamental principles of telephone switching.

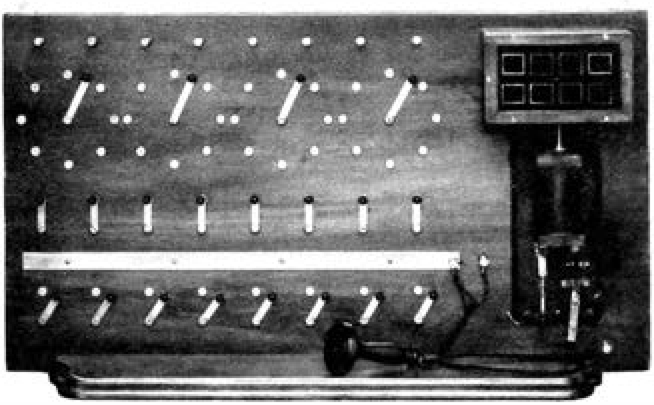


Fig. 6-5. Model of the first commercial telephone switchboard, installed at New Haven, Connecticut, in 1878. (Rhodes 1929, Fig. 31)

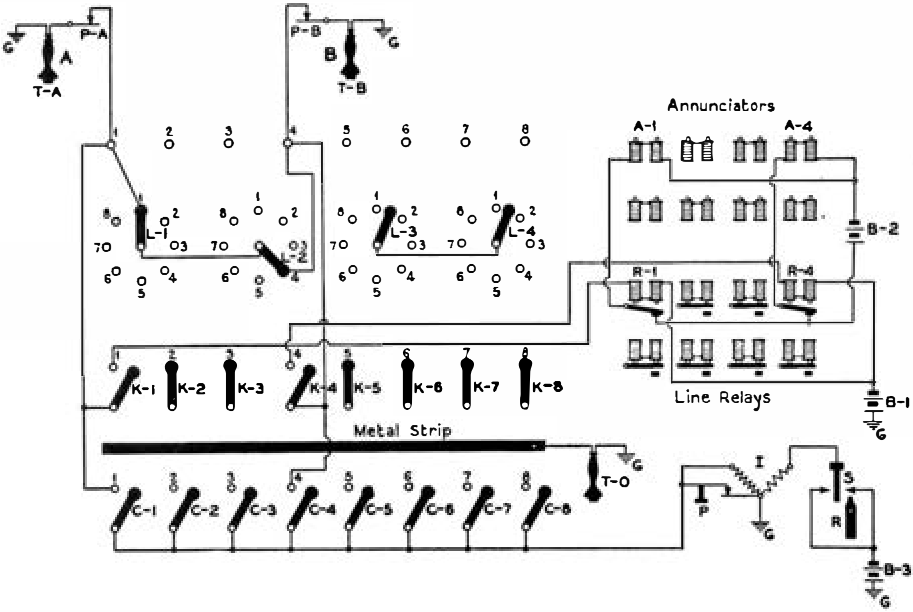


Fig. 6-6. Circuit diagram for the New Haven switchboard. (Rhodes 1929, p. 167)

scriber wished to make his call, he momentarily opened the circuit by pressing a button at his station (shown at the upper left of Fig. 6-6). This operation opened the circuit through line relay R-1 and actuated annunciator A-1 which remained operated until manually restored. (At a later date a bell was also added in order to awaken the night operator.) The operator responded by restoring the annunciator and turning switch K-1 to the metal "listening" strip which connected the caller to the operator's telephone. He then obtained the name of the desired called party, which he recognized as being on Line 4. Thus, the board fulfilled the first two requirements of any switching system (listed at the end of Section II) using, incidentally, a common-battery signaling arrangement which was soon to be discarded but would be reintroduced about 20 years later and become a standard feature of telephone systems.[8] The operator, knowing that the called party was on Line 4, opened both K-1 and K-4 and closed C-4, after

[8] Later systems supplied battery only during a call, not on a continuous basis, including idle periods, as used by Coy. Continuous battery was probably a hangover from telegraph service where it was common practice. With the poorly constructed lines of the time, it was advantageous since the central office was immediately notified of a line break.

which the ringing mechanism[9] (at the lower right of Fig. 6-6) was operated to send out a sequence of buzzes corresponding to the code of the called party. After ringing, the operator opened C-4 and connected the called line to the operator's telephone by turning K-4 to the listening strip. If the phone was answered, he connected the two parties by turning switches L-1 and L-2 as indicated in Fig. 6-6. If there was no answer, the calling party could be notified verbally by switching K-1 to the listening strip. Up to this point, the first five functions of a switching system had been fulfilled. The call was terminated by restoring all switches to normal but no means was provided for notifying the operator that the call was completed except by his occasionally swinging the K switches to the listening strip, an operation which was often carried out too infrequently (or sometimes too frequently) to be satisfactory. Thus, the final or "recall" function was not adequately handled by this board but was added by other manufacturers within the year.

This rather lengthy explanation has been given not only to point out how farsighted Coy had been in recognizing the functional requirements of a switching system but also to point out the complexity of the board and the large number of manual operations required. The simplification of these operations was one of the main objectives of switchboard development in the next 10 to 15 years. A complementary objective was to find much more compact arrangements so that large numbers of lines could be handled. It was obvious that expanding Coy's arrangement of switches to handle more than a few dozen lines would provide a very unwieldy system, and for large expansion a radically new approach would be needed.

Before examining other solutions to the switching problem, we should, however, point out one technique used by Coy which in principle has continued to play an important role in switching. This is the arrangement of switches designated L on Fig. 6-6. It will be noted that these connecting devices were used in pairs with a wire connection between. Interconnection of lines was accomplished by connecting one of the pair to the calling party and the other to the party called. In essence, this combination was a trunk with a connecting mechanism at each end. While the rotary switch was soon abandoned as the connecting device, the basic idea of a trunk with a connection at each end has been used repeatedly in subsequent switching systems under various names such as transfer trunk, cord circuit, link circuit, junctor, etc., and is a basic means for providing a simple, flexible

[9] This was Watson's buzzer, also known as "Coy's chicken" which is described with other ringing arrangements in Section 4.1 of Chapter 3. It produced a high-voltage alternating current that caused a loud buzz in the distant telephone.

means for interconnecting a large number of terminals.[10] Further, it will be noted that Coy used these trunks with economy. With eight lines available, there was a possible need to establish as many as four calls at once. However, he recognized that the probability of this occurrence was small and on a strictly instinctive basis he reduced the size of his board by providing switches for only two simultaneous conversations. Much later, mathematical analysis to determine the precise number of trunks required to provide a desired grade of service was to become an important milestone in the evolution of switching systems.

Telephone exchange service captured public imagination far beyond anyone's expectation. To quote testimony by C. E. Scribner[11] in an 1896 law suit: "The value of the telephone exchange came to be appreciated by the public very soon and the growth of the telephone exchange gave to its promoters surprise following surprise. A switchboard designed and well calculated to provide for the estimated growth of years would be found of too small capacity for the needs of the time when completed, the growth of the business between date of order and that of completion being sufficiently great for this result."

Naturally, such a situation attracted many technicians to the field. At the time, the Bell interests were concerned with the many attempts being made to infringe the important basic patents on telephone instruments and their policy on switchboards was that, in the absence of a reliable central manufacturing company, the local company should adopt the best apparatus readily available to them and make improvements as they came along. As a consequence, every little shop with electrical experience entered the field and applied the inventive genius of its staff to switchboard development. In fact, inventions were made in so many places and came along so fast that it is practically impossible to produce a reliable chronological account of this period. Dozens of different types of boards were produced and we shall not attempt to describe more than a few.

The Coy board, it will be recalled, used the same electrical circuits for communication and signaling. A competing concept, originating

[10] It is not intended to imply that Coy was the inventor of all the principles employed in his board. To a considerable extent he merely put to work the material and ideas available to anyone versed in the art of the time. However, it is to his credit that they were put together in a way to meet the needs of telephony instead of adapting the telephone service to fit existing switchboards.

[11] Charles Scribner was one of the most prolific early telephone technicians, with over 400 inventions to his credit. He joined the Western Electric Manufacturing Company in 1877, at the age of 18, as a designer of telegraph instruments. Shortly thereafter, when the company began the manufacture of telephone apparatus, the engineering activity in this field was assigned to Scribner. He literally grew up with the telephone, spent practically his entire active life in this field, and served for many years as Chief Engineer of Western Electric after it became the manufacturing branch of the Bell System.

largely in the district telegraph community, was the use of separate circuits for communication and the major signaling functions. Probably the first such exchange was installed in 1878 by the American District Telegraph Company of Chicago. This company operated a call-box system used by subscribers for summoning messengers, expressmen, etc. Leroy Firman, the General Manager, seems to have been a highly ingenious individual and decided to supplement the call-box service with a telephone exchange. Stations using Edison transmitters, manufactured under Western-Union-held patents, were installed on individual lines run to subscribers of the service. The call-box service was supplemented by adding two codes to designate "Telephone Use" and "Telephone Through." A customer wishing to make a call sent the former signal, and the call-box recording operator made a note of the request and name of the caller and passed the information to the switchboard operator who obtained by voice the name of the called party, called him, and set up the connection through the telephone switchboard. At the end of the call, the "Telephone Through" signal was sent and the connection was taken down by the switchboard operator after notification from the recording operator. Although the procedure was slow, since it involved two operators, the switchboard itself was simple and compact, consisting of rows of metal terminals, one per subscriber, with holes in which metal connecting pegs could be inserted. The operator's set was grounded at one side (through a battery during signaling) and the other side was connected through a flexible cord to a peg which could be plugged into the terminal of the customer desiring service. After the operator learned the call destination, he disconnected his set and connected the appropriate line terminals with a flexible wire or "cord" having suitable pegs at each end, these cords being carried in a loose bundle in his pocket. Electric tap bells or one of the other simple arrangements described in Section 4.1 of Chapter 3 was employed for ringing.[12]

The Chicago system was a highly primitive arrangement and would have been of little historical interest had it not stimulated Firman, Scribner, and other inventive visionaries into new approaches to switching from which much of the manual switching art evolved. All this will be related later. For the present it is enough to note two improvements on Coy's board. One was the introduction of an "end of call" or "recall" signal.[13] The other was the use of a double-ended

[12] Firman's U.S. Patent No. 328,305, filed January 16, 1880, and issued October 13, 1885, seems to give a fairly good description of his early system but does not agree in all respects with contemporary accounts. Undoubtedly, changes were many and frequent.

[13] The price was high, however, since it was obtained as a result of a separate signaling circuit. On June 9, 1880, H. H. Eldred applied for a U.S. patent (No. 303,714, granted August 19, 1884) which included recall, or "clearing-out" drops, for a system using the talking line

cord with plug-in pegs for connecting lines. This was not an original idea but was a great improvement in the connecting function as compared to Coy's rotary switches which required two switchpoints per line for each connecting circuit and became unwieldy with more than a few dozen lines.

The use of a separate signaling circuit (or "order wire") was favored by a number of early telephone technicians because it greatly simplified the switchboard circuitry. Since this circuit was only used infrequently, it could be run to many subscribers who used it in common. This "call circuit" principle, as it was referred to, was abandoned for subscriber signaling after a few years but was used for trunk signaling until about 1925. Before it was abandoned for subscriber use, rather extensive use of call circuits was employed in the so-called "Law" telephone system. This system was the outgrowth of the one begun in 1874 by W. E. Childs for exchange telegraph service between law firms. It was converted to telephone around 1880 and was introduced into a number of eastern cities. The calling procedure differed from that used in Chicago in that the subscriber connected his station to the order wire when he wished to make a call and forwarded his request orally to the operator who continuously monitored the call circuit. The telephone line of each subscriber terminated in a cord and plug and, after ringing the called subscriber, the operator made the desired connection by inserting the appropriate two plugs into one of a number of connecting strips provided for the purpose. A Law system board used in Richmond, Virginia, in 1882 is shown in Fig. 6-7. Earlier boards used boys as operators but by this time women were beginning to be used and more decorum was introduced into the switchroom.[14]

Although the Law system had little lasting impact on the evolution of telephone switching, a patent assigned to the Law Company by Frank Shaw[15] is worth noting. This patent seems to be the first recorded proposal to assign numbers to subscribers to avoid ambiguity on calls to customers with similar names. It also outlines the possible use of multiple central offices connected by trunks.

Another early telephone system receiving considerable use was that of the Gold and Stock Company, originally formed to broadcast

---

for signaling. This seems to be the origin of the scheme used in the Western Electric "Standard" board. A slightly different but related arrangement was used in the Western Electric "Universal" board and Gold and Stock Company boards of 1879. (All of these switchboards are discussed in Section 3.3.)

[14] A Mr. Eckert, of Cincinnati, speaking in 1881 before the National Telephone Exchange Association, said, "I would like to say right here, I've been asked by Mr. Sabin what our experience has been with young ladies' help; the service is very much superior to that of boys and men. They are steadier, do not drink beer and are always on hand."

[15] F. Shaw; U.S. Patent No. 220,874; filed August 11, 1879; issued October 21, 1879.

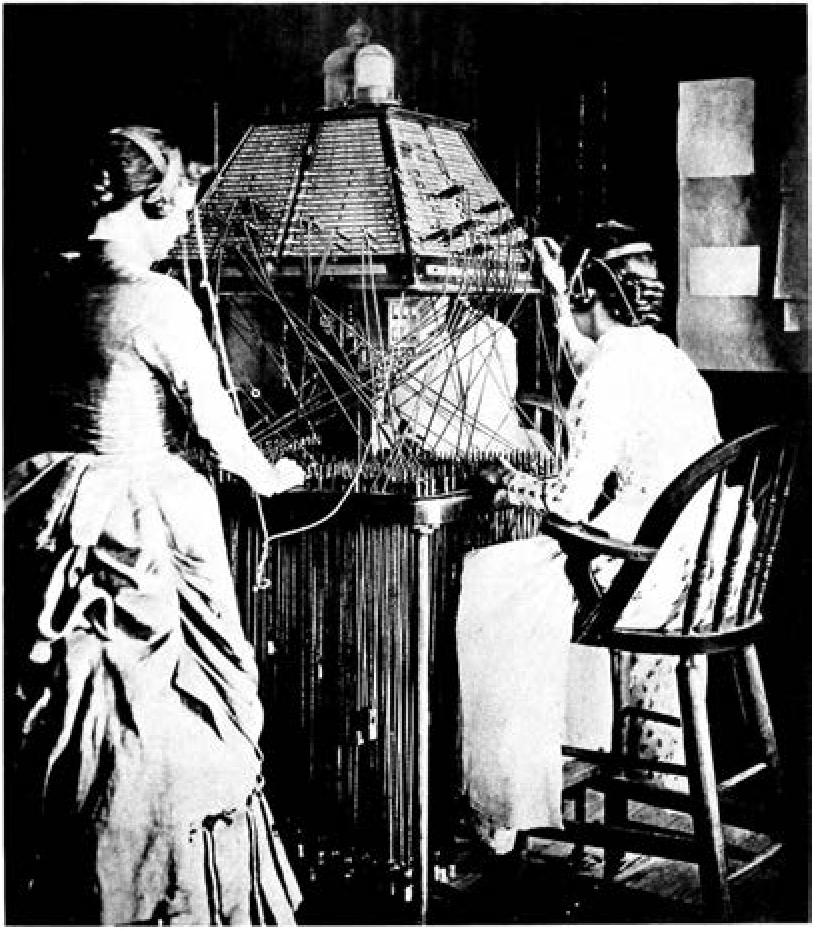


Fig. 6-7. Law system board at Richmond, Virginia, 1882. (Rhodes 1929, Fig. 39)

stock market and banking reports by means of telegraph printers. A large switchboard used by this company in 1879 is illustrated in Fig. 6-8; Fig. 6-9 shows a single position. This was a single-circuit board, i.e., using the same wire for signaling and communication. Drops designating lines were at the top of the board with connection terminals arranged in the same order directly below. This part of the board was essentially the Western Electric "universal switch," described in Section 3.3.1, but some of the additional features were unique. As shown in Fig. 6-9, the operator could communicate directly with

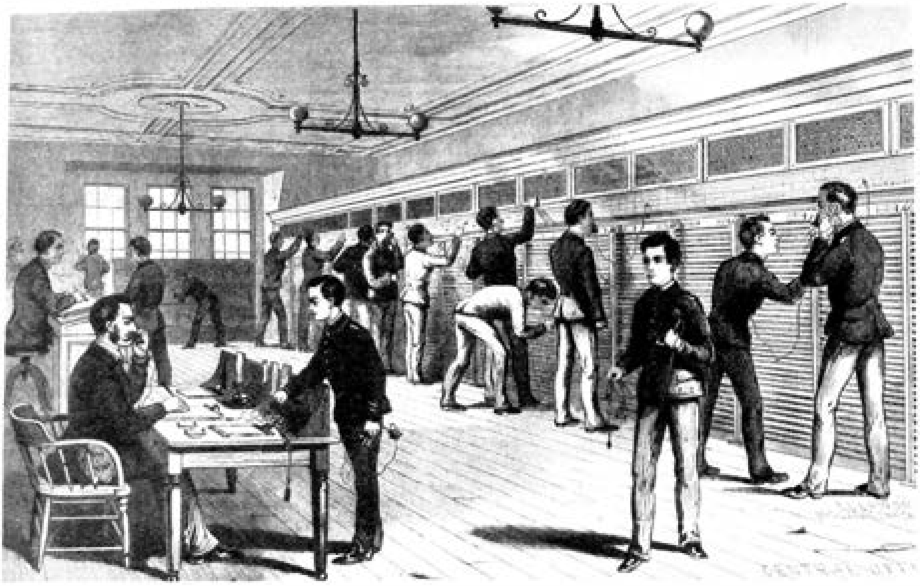


Fig. 6-8. General view of the "Gold and Stock" switchboard of 1879. (Rhodes 1929, Fig. 34)

Fig. 6-9. An operator's position of the "Gold and Stock" board. (Rhodes 1929, Fig. 35)

the subscriber by plugging one cord of his telephone into a ground plate and the other into the line terminal. Communication between subscribers was established by connecting the appropriate line terminals through plugs and cords. However, since this was a large switchboard, the desired subscriber line was likely to be terminated in a jack distant from the operator who answered the call. To avoid long cords, a series of metal bars running the length of the board was installed below the line terminals. These bars were so mounted that they could be turned on their long axis. An operator, wishing to connect two remotely located lines, connected one to a bar through a cord using pins or clips for connection, and rotated the bar slightly to indicate to the other operators that it was in use. At the same time he called (through the air) to the operator at the distant position to connect the desired subscriber to the same bar.[16] This practice, of course, resulted in much clamor and confusion. Another feature was the use of "clearing-out" drops associated with each bar which would operate when the calling subscriber "rang off" on completing a call by operating a signaling device at his station. All these drops appeared in a group where they could be observed by a special operator. When a drop fell, indicating call completion, the operator shouted to the line operators to take down the connection to the bar (designated by number), thus adding further noise and confusion.

The preceding pages have discussed only a few of the many switchboards that were used in the early days of telephony. It is tempting to explore further the design of early boards, since they provide a fascinating example of the technical ingenuity of the era, but space limitations prevent this indulgence with one exception. A number of manufacturers such as Williams of Boston and Gilliland of Indianapolis used a principle employed in the Western Union Peg Switch and known in Europe as the "Swiss commutator." The basic plan was to have a series of metal bars crossing the board horizontally. Above them, but not in contact, a series of bars was arranged vertically, each of which was connected to a line. Holes were drilled through the horizontal and vertical bars at each point of intersection so that a vertical bar (line) could be connected to a horizontal bar by means of a metal plug passing through the bars at the intersection. Another plug could similarly connect the horizontal bar to a different vertical bar, thus interconnecting the lines. For many reasons this scheme did not continue in use very long for manual boards. However, many years later the idea was incorporated in crossbar switching systems, in which the vertical and horizontal "bars" operated switch contacts at the intersections.

[16] At this time all telephone communication employed ground-return circuits and switching in this and other contemporary boards involved only a single conductor.

After this final example of the diversity of ideas explored by early telephone technicians, we shall turn to the main stream of development effort, which ultimately resulted in standard types of switchboards which served the growing needs of telephony for many years.

3.3 Meeting Telephone Requirements

As just noted, the functions of the telephone switching system were met in the early years by a number of technicians using diverse methods based on available apparatus and to some extent by adaptation of telegraph practices. While this approach met immediate needs and gave the public a taste of the potential value of telephony, it led to dead ends when expansion was required. To meet this need, a distinctive approach had to be developed that not only met the functional requirements of switching but did so with the minimum of equipment and operating effort. This development began only a few years after the telephone invention. The path followed was in part determined by developments closely related to the switching process but was also greatly influenced by progress in the transmission and station areas, covered in other chapters, and these influences should be kept in mind as switching development is discussed.

To review briefly, the main transmission objective was the conquest of distance but the use of the best line materials available came far short of providing even tolerable conversation over United States continental distances. The advent of loading early in the twentieth century helped to some extent but the problem was not fully solved until the introduction of the vacuum tube repeater about 1914. Thus, for some 40 years there was great pressure put upon the development of high-efficiency station equipment since gains in this area were the most effective means available for achieving increased distance.

The electromagnetic transmitter used by Doolittle, Coy, and other early entrepreneurs was soon replaced by battery-powered microphonic devices which were true amplifiers and had a potential for increasing the electric power on the line by roughly a thousandfold.[17] However, the early instruments of this type were far from reaching the ultimate potential of the principle, and it was necessary to achieve the greatest practical efficiency by using a high battery current which, with the high-resistance iron-wire loops then in use, was more readily obtained from a local battery consisting of a few cells at the station than from a large battery at the central office.

Signaling the operator by means of a common battery, as used by Coy and others, also was a problem with high-resistance loops,

[17] The evolution of station apparatus is covered in detail by Chapter 3.

particularly when the line, as often the case, was not well-insulated and therefore subject to leakage to ground on wet days. A satisfactory scheme for signaling the operator, as well as for calling the subscriber to the phone, was found in Watson's ringing system. This consisted of a high-voltage, low-frequency (about 17 hertz) alternator as a source of power and a balanced-armature ringer which followed alternations in the current.[18] The alternator, in the form of a hand-cranked "magneto," was installed as part of the telephone station. It was used by the subscriber not only to "ring" parties on his own line but also to "ring down" an annunciator or drop at the central office in order to signal the operator that assistance was needed in setting up a call or for disconnecting one at its termination. Similar generators, originally driven by hand or foot, but soon electrically powered, were used at the central office to ring the called party.

As a result of these circumstances, the telephone system for about 15 years following 1880 consisted of station sets using locally supplied power for the transmitter and for signaling the operator. By about 1895, line construction had improved, loops had been shortened, and stable transmitters had been developed that would operate efficiently on battery current as low as 30 to 50 milliamperes. So, at that time it became possible to consider once more the use of a central rather than a local battery both for transmitter supply and for signaling the operator by means of sensitive relays that were then available. Thus the trend toward common-battery offices began, and with it the evolution of more sophisticated signaling arrangements that reduced the burden on both subscriber and operator.

3.3.1 Local-Battery Systems and the Multiple Board

With this background in mind, we can now return to the development of a series of switchboards which successively met the growing needs of telephony and gradually evolved into standardized boards meeting Bell requirements.

The framework for this evolution grew out of Firman's Chicago exchange and was largely the result of work by Scribner and his associates at Western Electric as they strove to improve the original, rather crude, arrangements provided.

The Western Electric Manufacturing Company had for many years provided telegraph equipment for the Western Union Company[19]

[18] Such equipment constituted a large part of the early station sets and is described in Section IV of Chapter 3.

[19] A brief review of the relations existing among the Bell interests, Western Union, and Western Electric at this time will provide useful background for the events related here. The Western Electric Manufacturing Company of Chicago was an independent manufacturer of electrical equipment with considerable experience in the production of telegraph

and it was natural that the latter organization should turn to them for technical and manufacturing help when Western Union decided to purchase a license for the Edison transmitter and challenge the Bell patents by entering the telephone business. The first fruit of this combination was the so-called "Edison" exchange opened in Chicago in 1878. This, as related previously, used separate call-circuit signaling and would have been unimportant if its growth had not inspired Firman to make a major switching invention.

The problem Firman faced was that the number of customers had grown to the point where the traffic between them could not be handled by the two or three operators who could be crowded in front of a single switchboard on which the lines were terminated. (With call-circuit signaling, only simple connectors were required at the telephone board and a very large number could be crowded in a small area, but the space problem became more acute as this operating procedure was abandoned.) The solution, worked out in late 1878 and early 1879, was to provide several duplicate switchboards, each of which contained terminations for every customer in the exchange; thus a connection between any two customers could be made at any board. These duplicate boards were referred to as "multiple" boards and this term was later used to designate any arrangement which provided more than one jack appearance per line. Unfortunately, since there were several terminations for each customer, it was quite possible for a connection to be made at one board to a line already in use at another. What was needed was a means for designating a busy line and this was accomplished by using an additional "target" board, visible to all operators, on which tags were posted by a separate operator.[20] Crude as this basic scheme was, it soon

---

equipment. In 1876, Bell's telephone patent rights had been offered to Western Union for the modest sum of $100,000. The offer was turned down on the basis that the invention would be of little use to Western Union, but within a year they decided that there was indeed a future in telephony and entered the field. Equipment was based on patents acquired from Gray, Edison, Dolbear, and others. Both the Western Union shops and the Western Electric Manufacturing Company produced equipment. Following court action started by the Bell interests, Western Union in November 1879 settled out of court, agreeing to retire from the telephone business, grant rights to Bell on all patents they had acquired, and sell to Bell all of the telephone systems that Western Union and their subsidiaries had established.

Experience had shown the difficulties of building a system from elements designed and produced by many manufacturers, and Bell, in a step toward standardization, purchased the Western Electric Manufacturing Company in November 1881, after which time it operated under the name of the Western Electric Company. At about the same time, Western Electric acquired the licenses of Williams and Gilliland which, with the Western Union and Bell patents, gave them manufacturing rights under all the extant patents of any importance. In February of 1882, agreements were reached whereby Western Electric became the main supplier of Bell for major telephone system components.

[20] Firman, at the time, did not consider that his multiple board and target indicator constituted an invention but, under the urging of E. M. Barton, he filed for a U.S. patent on January 7, 1881, and No. 252,576 was granted January 17, 1882, and sold to the Western Electric Company.

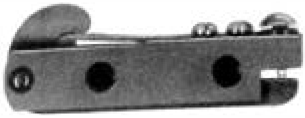
evolved into a highly workable arrangement with the help of Scribner and others. But before we discuss the practical development of the multiple switchboard, we need to consider other problems of the era.

The Edison exchange had started with call-circuit signaling but soon there were demands for service from areas not covered by the call-box system. To handle these lines, special arrangements were provided, incorporated first in the Western Electric "Universal" switchboard and later in an improved board designated as the "Standard." The basic idea was that current from a local source at the customer's station was used to operate an annunciator, or drop, at the switchboard to indicate the desire to make a call. The customer's line terminated in a "jackknife" switch, invented by Scribner, which served two purposes. As a plug was inserted, it made connection to the customer's line and at the same time actuated a movable member, hinged like a jackknife blade, which opened two contacts disconnecting the annunciator circuitry. This device, shown in Fig. 6-10, was the first piece of telephone apparatus in which the performance of a primary switching function (connecting) automatically carried out a secondary (signaling-related) function. It was the first of many steps taken to reduce the number of actions required by the operator.

The "universal switch" was used in the Edison exchange for stations without call boxes and also employed by the Gold and Stock and other companies. Its 1879 implementation is shown in Fig. 6-11 with a call set up between Lines 1 and 9 (corresponding to the actuated drops at the top of the board) by cords and plugs connected through a "call plate" at the bottom of the board. This board will also be recognized in the background of Fig. 6-9.

In 1880, the improved "Standard" switchboard, Fig. 6-12, was introduced. This board incorporated several new features that were to be a part of future switchboards. One was the use of a keyshelf and flexible, double-ended cords arranged with weights to restore them to an orderly at-rest position, thus doing away with the loose cords used previously. Another feature was the use of keys, associated with the cords and mounted in front of them on the shelf, to connect the operator's telephone and to send out ringing current. Finally, each cord included a "clearing-out drop" which could be operated at the end of the call by a spurt of current sent out by the customer. As originally conceived, these boards used the local talking-battery at the customer's station as a source of current to operate the drops but as Watson's magneto ringing system was introduced in subscriber stations,[21] the magneto was adopted as a source of signaling current and the switchboard drops modified, where necessary, to operate on low-frequency alternating current.

[21] As described in Section 4.1 of Chapter 3.



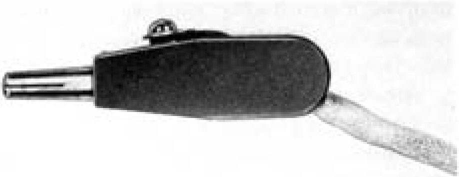


Fig. 6-10. Scribner's "jackknife" switch and the plug used with it. (Rhodes 1929, Fig. 32)

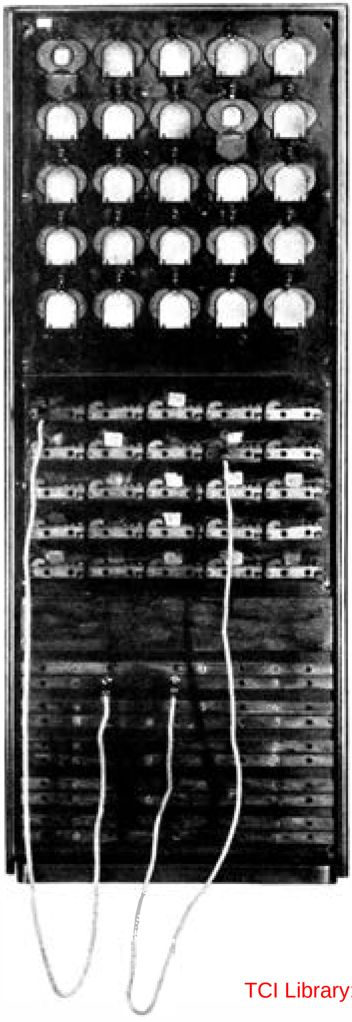


Fig. 6-11. The "Universal" switchboard of 1879. (Rhodes 1929, Fig. 33)

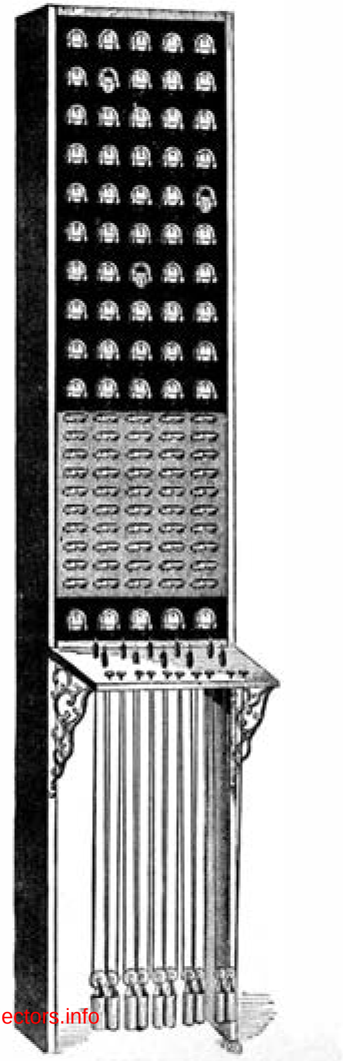


Fig. 6-12. The "Standard" switchboard of 1880. (Rhodes 1929, Fig. 37)

Similarly, a magneto or equivalent ac generator was used at the central office in place of the direct current originally used for ringing. These measures greatly extended the distance over which loop signals could be sent and by 1881 or 1882 magneto signaling and ringing had become almost universally used.

The "Standard board" of 1880 approximated the characteristic shape of later manual switchboards but lacked much of the substance. The major unsolved problem was that of size. A unit, or "position" as it came to be known, could take care of only 50 lines.[22] When more lines were required, additional units were added side by side. This was satisfactory for lineups of about three units since during light traffic a single operator could be seated at the center position and reach lines in both her own and the adjacent units. During busy periods, operators could be added at the end positions and complete calls to all lines in the group of three positions with occasional assistance from the center operator in connecting to jacks beyond reach. For more than three positions it was necessary to use transfer trunks between positions. With this arrangement, the answering operator did not complete the call herself but connected her cord to a trunk appearing before an operator having the called number within reach. It was this operator who made the final connection. With such a scheme, it was necessary to pass the completing information between operators. To avoid other problems, this was done orally, usually over the trunk but sometimes through the air as in the Gold and Stock Company exchange, with obvious noise and confusion.

Several improvements were necessary and soon introduced in Western Electric boards. An important step was the reduction in size of the drop and jack apparatus resulting from inventions by E. P. Warner and J. C. Warner respectively, made about 1882 while they were employed by Western Electric. The Warner jack, the prototype for the present-day plug and jack, will be discussed later. For the present it is sufficient to note that, like the Scribner jackknife switch, it not only performed a connecting but also a signaling function. However, it occupied less space on the face of the board, the size ultimately being reduced to the point where over 10,000 jack openings could be placed within reach of a single operator. These were important steps in solving the space problem but not nearly so significant as modifications and improvements made in the application of the multiple jack field proposed by Firman some years before.

Firman's scheme was originally proposed in connection with call-box

[22] The large amount of space required by the jackknife switches and the line drops was one of the penalties paid for eliminating call-circuit signaling and its associated complicated operating practices.

signaling under circumstances requiring only a few duplicate jack fields. A target board to indicate busy signals, while awkward, was at least usable under these conditions. Subsequently, call-circuit techniques had been abandoned and signaling had become associated with each line so that a single operator could perform both the answering and connecting function. The increased space required by this operating change and the rapid growth of telephony clearly indicated that a large number of positions with many multiple appearances would soon be required. To have signaling drops associated with each appearance would be wasteful of space and equipment and the whole scheme would soon become unworkable unless a scheme was devised for indicating busy lines which was more suitable than the target board. The steps in achieving the necessary improvements were many, including a number which proved somewhat impractical and were abandoned in favor of later developments. For our purpose it is sufficient to indicate the ultimate outcome. About 1883, Scribner devised a circuit which made it possible to test a line by touching the plug tip of the operator's answering cord to the shell (sleeve) of the jack. A click indicated a busy line and the calling party was so informed.[23] With this arrangement the need for a target board was eliminated, operating procedure was simplified, and there was no limit on the size or layout of the multiple imposed by line-of-sight requirements.

Another and equally important change occurred in the layout of the vertical face of the board. The basic idea involved a separation between jacks used for answering (including the drop or other associated signals) and those used for completing. The completing jacks were placed in the upper part of the board, a full field providing jack terminations for every line in the office and occupying the space covered by three positions. This field was repeated (multipled) in the adjacent section of three positions, one section following another until an adequate number of appearances were achieved. Since the reach of an operator covered the board face constituting three positions, each operator, regardless of where she sat, could complete calls to every line in the office without assistance from another operator.[24] It was appreciated that an operator could handle only a limited number of calls during the busy hour and did not need access to all lines for answering purposes. Instead, the answering

[23] Originally the click could be heard by both the operator and the calling party, but in 1885 J. J. Carty devised an improved circuit (not immediately adopted) which prevented transmission of the click to the customer.

[24] In the Bell System the width of an operator's position became standardized at about 2 feet. Thus a three-position section of multiple was 6 feet wide. The height varied with the size of office and type of board but was limited to about 3 feet.

jacks were distributed over the entire switchboard with only the small number which could be handled in a busy hour placed in each position. The drops associated with each line to signal a request for service and those associated with cords to indicate completion were located below the answering jacks or on the keyshelf. This shelf, which had been introduced by the 1880 "Standard" board, continued to be a feature of all later boards and provided the mounting for double-ended cords and plugs retracted by pulley weights, ringing and listening keys, and other auxiliary apparatus. With this arrangement of keyshelf and backboard with separate answering jacks and completing multiple, the operations on each call were carried out by a single operator who had all necessary controls and connections within reach.

The multiple board just described evolved in several stages and was the product of a number of people working over a period of years. In the limited space available only the major contributors and a few of the steps in the development can be covered in summary form below:

(*i*) *Basic idea for use of duplicate switchboards*—Invented by Firman about the end of 1878 and patented in 1882.

(*ii*) *Means for handling customers not on call-box lines*—Circuitry and jackknife switch invented by Scribner and used in "Universal" switchboard of 1879.

(*iii*) *Improved jack capable of more compact mounting*—J. C. Warner (1882).[25] Later improvements by others to reduce size and cost. Compact strip mounting contributed by Richard Freeman in 1884 and H. B. Thayer in 1887.

(*iv*) *Sensitive and compact drop for switchboard mounting*—E. P. Warner (1880).[26] Later improvements by others to reduce size and increase stability.

(*v*) *Click test for busy line*—Scribner (1883)[27] with improvement by J. J. Carty in 1885.

(*vi*) *Continuously repeating multiple spread over section made up of three positions*—Scribner (1883).[28] In this proposal the multiple was used for both answering and completing, thus requiring a drop per jack. The basic idea of a continuously repeating multiple was a permanent feature of future boards, but Scribner's proposed layout of drops and cords was impractical.

(*vii*) *Separate answering jacks distributed over the entire board with only*

[25] J. C. Warner; U.S. Patent No. 281,741; filed June 8, 1882; issued July 24, 1883.
[26] E. P. Warner; U.S. Patent No. 232,093; filed April 14, 1880; issued September 7, 1880.
[27] C. E. Scribner; U.S. Patent No. 330,060; filed October 22, 1883; issued November 10, 1885.
[28] C. E. Scribner; U.S. Patent No. 330,062; filed November 10, 1883; issued November 10, 1885.

a portion before each operator—J. A. Seely (1885).[29] This proposal more than doubled the number of completing jacks that could be placed within any operator's reach.

Most of these improvements were incorporated in the 1888 Cortlandt switchboard in New York City, the largest board built up to that time. The planned capacity of this board, pictured in Fig. 6-13, was over 10,000 lines. An interesting aspect of this board was that it was designed originally for ground-return circuits but during construction it was decided to change it for use with metallic lines. It was accordingly modified and rewired to provide the necessary extra contacts and circuitry. This contribution by C. E. Scribner and F. P. and J. A. Cook resulted in the first major switchboard for metallic circuits.[30]

With the Cortlandt board, the basic design of multiple magneto boards was nearly completed. The final steps were taken about 1890 to 1892 and most were incorporated in the Albany, New York, board of 1892. As the number of multiple sections were increased, difficulty was experienced with the many series contacts required by previous designs. Scribner, with the assistance of J. A. Steiner and O. A. Bell, eliminated this problem by designing a "branching" system using bridged multiples with high-impedance line signals.[31] At the same time drops were introduced which were restored to normal by insertion of a plug in the answering jack. These "self-restoring" drops eliminated the separate hand operation previously required. At about the same time (1890) E. M. Barton proposed an intermediate distributing frame to provide the easy rearrangement of answering jacks to provide for uniform distribution of traffic to all operators.[32]

[29] J.A. Seely; U.S. Patent No. 330,067; filed March 28, 1885; issued November 10, 1885.

[30] There are some minor discrepancies in the records of this board. Scribner says it was built in 1887 but it was not put in use until the following year. While it was undoubtedly the largest board built at the time, the full design objective of 10,000 lines was not achieved for some years. In 1890 there were about 5,000 lines in use and by 1900, after conversion to common battery, the number had increased to 9,000.

[31] Originally, in boards with a single line jack, the line signal was a low-impedance drop, normally connected between line and ground, the drop being cut off by spring contacts in the jack upon insertion of a plug. The same general scheme was continued when multiple jacks were introduced by wiring them in series and using spring contacts to give a normal connection through each jack. Inserting a plug opened the contacts and cut off both the line drop and all jacks beyond the one in use. The series arrangement became particularly complex with metallic boards since excessive crosstalk could occur unless both sides of the line were opened when a plug was inserted. The parallel arrangement, with a high-impedance line signal that could remain bridged during conversation, simplified the boards greatly by eliminating the jack-spring contacts.

[32] Originally, the answering jacks were assigned to the operators' positions in numerical order, but it was soon found that the traffic pattern was highly variable with some operators under this system being assigned highly active lines and others being lightly loaded. The use and structural features of the distributing frame are discussed in Section 3.4.6.

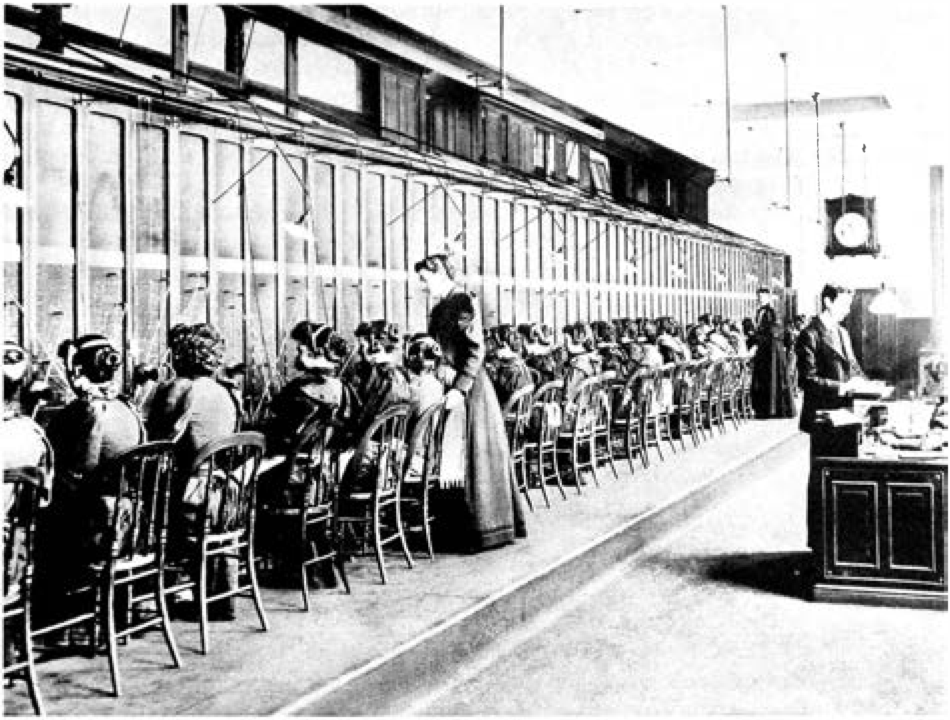


Fig. 6-13. Cortlandt board in New York City, 1888. (Rhodes 1929, Fig. 38)

3.3.2 Common-Battery Systems

The local-battery, magneto system served well for the first 20 years of telephony. The high transmitter current provided the loud output required to converse over high-attenuation transmission systems and the signaling system was simple with the high-voltage, low-frequency magneto source of current being well adapted to the poorly constructed, high-resistance loops commonly used. Without these features, the growth of telephony might well have been slower and less widespread. But this system had important disadvantages from the standpoint of both user and the telephone administration. The batteries on the customer's premises were, up to the late 1890s, of the wet type and were not only bulky but used fluids which were damaging if spilled. Dry cells, being introduced at this time, were an improvement from the user's standpoint but were of uncertain life and still bulky. With either type of cell, replacement was costly and unpredictable. Battery failures caused special trips for replacement or long service outage. The magneto was a reliable device but not particularly satisfactory from the user's standpoint since it constituted a bulky part of the station set and prevented the use of compact, movable stations of the desk type. Its use was relatively simple but added specific operations to initiate and terminate a call. The latter operation was often forgotten, resulting in long and unnecessary tieup of equipment.

By the middle 1890s, the situation had become more favorable for eliminating these nuisances by means of a common-battery system that used power sources at the central office for both talking and signaling current. Some of the major factors in this change were transmitters of greater stability capable of efficient operation on low currents, sensitive dc signaling devices operating over a wide range of currents, and, finally, well-insulated loops with lower resistance. This last item resulted from improved outside plant and also from the dispersion of switching offices connected by exchange trunks, with consequent reduction in the average length of loop. This was a highly significant factor since throughout the manual switching era the loop resistance, with normal central-office equipment, was limited to a maximum of about 635 ohms. For higher-resistance loops it was necessary to use special "long-lines" equipment and for the very longest loops there was no substitute for the local-battery, magneto system. Thus the transition to common battery was necessarily carried out over a long period. In the large cities, with short loops, it was relatively fast, depending largely on growth and practical replacement rates. In the smaller towns and rural areas, complete changeover to common battery was not practical and combination switchboards that could handle either system were commonly used well beyond the 1920s in order to service rural and other stations beyond the range of common-battery operation. It is interesting to note that the Western Electric Company Bulletin 1017 (of about 1911) states: "In the small plant of 100 lines or less, with a prospect of a small future growth, there is no question but that the local battery system should be used. It is also just as certain that in the larger plant of 500 lines or more, central battery equipment will be more profitable." For situations with uncertainties about future growth they suggested the use of a convertible board which could start out as a local-battery system and be changed to common battery when conditions were more suitable.

The development of the high-capacity common-battery switchboard occurred over a period of several years. A few early experiments were made in the late 1880s but significant commercial installations were not made until the 1890s, some of the important early steps being listed below:

Date	Location	Type of Board	Type of Signal	Signaling Current	Talking Current
1893	Lexington, Mass.	Non-multiple	Drops	Common Battery	Common Battery
1896	Worcester, Mass.	Multiple	Lamps	Common Battery	Local Battery
1897	Louisville, Ky.	Multiple	Lamps	Common Battery	Common Battery

One of the great improvements accompanying the common-battery board was the use of miniature incandescent lamps in place of drops for giving both line and clearing-out signals. These were more readily seen by the operator, occupied less space on the board, and avoided the noise and confusion due to the falling and restoring of drops. They were used with jack control circuits which eliminated the need for any operator action for restoration to normal. The use of lamps was first proposed by J. J. O'Connell of the Chicago Telephone Company and received limited use on a local-battery board in that city in 1894. But the Worcester and Louisville boards marked their first major application. Their use, it should be noted, was dependent on the availability of suitable control relays operable over a wide range of line current. This was necessary since the lamps available at the time operated over only a relatively small range of currents and could not be used with any degree of satisfaction in series with a line.

A major problem in the use of a common battery was to devise circuits for supplying a large number of lines without annoying crosstalk between them. Figure 6-14 shows schematically two schemes which were devised about 1896 for this purpose. The one shown in Fig. 6-14a, invented by Hammond B. Hayes of the American Telephone and Telegraph Company, used a repeating coil (transformer) and introduced a relatively small series resistance in the battery supply circuit. The large mutual impedance of the repeating coil was in series with the battery and minimized the transmission of crosstalk.[33] This circuit was used in the Lexington board and the basic principle continued in use in Bell manual switchboards and in the panel-type machine switching system. The other scheme (Fig. 6-14b), using bridged impedance coils, was patented by Scribner in 1896. It required relatively high-resistance coils since a high impedance was needed to avoid transmission loss and reduce crosstalk, there being no mutual coupling to assist in this process as in the case of the repeating

[33] The cord circuit used at the far end of the trunk (the B board, discussed later) did not utilize the mutual impedance of the repeating coil and other measures had to be adopted to control crosstalk.

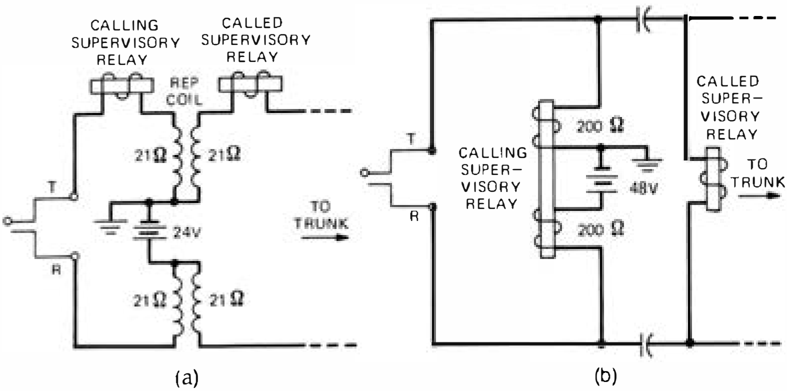


Fig. 6-14. Battery supply circuits for common-battery offices: (a) repeating coil circuit; (b) bridged impedance circuit.

coil. The bridged retard system was used by Bell in its step-by-step systems and in all crossbar systems. An important consideration in adopting the retard system was that often it was possible to use relays, required for switching operations, as battery supply inductors, thus achieving a very economical system. A 24-volt battery was commonly used with repeating-coil cord circuits except on long-distance calls where the voltage was increased to 48 volts where conditions permitted. With bridged-retard cord circuits the normal voltage was 48 volts in order to compensate for the high resistance of these coils. Even where the bridged retard system was normally used, special 48-volt, repeating-coil battery supply arrangements were often used on long-distance calls in order to achieve a small transmission advantage by means of increased talking-battery current.

Various improvements and changes were made in the early common-battery boards, but they had become pretty well standardized by 1900. A typical large multiple board of this period is illustrated by Fig. 6-15, and a simplified circuit schematic is shown in Fig. 6-16.

Operation of the manual common-battery board is roughly as follows (see Fig. 6-16): A subscriber desiring to make a call merely lifts his receiver (A) from the switchhook. This closes the battery circuit from the central office through line relay C and lights lamp D placed just above answering jack F, thus notifying the operator that a call is waiting. The operator then inserts her back plug (E) into the answering jack. This completes the sleeve circuit through G and operates cutoff relay H, extinguishing lamp D and supplying talking battery through the repeat coil in the cord circuit. She connects her telephone set (K) through a key (not shown), determines the desired number, makes a busy test, and

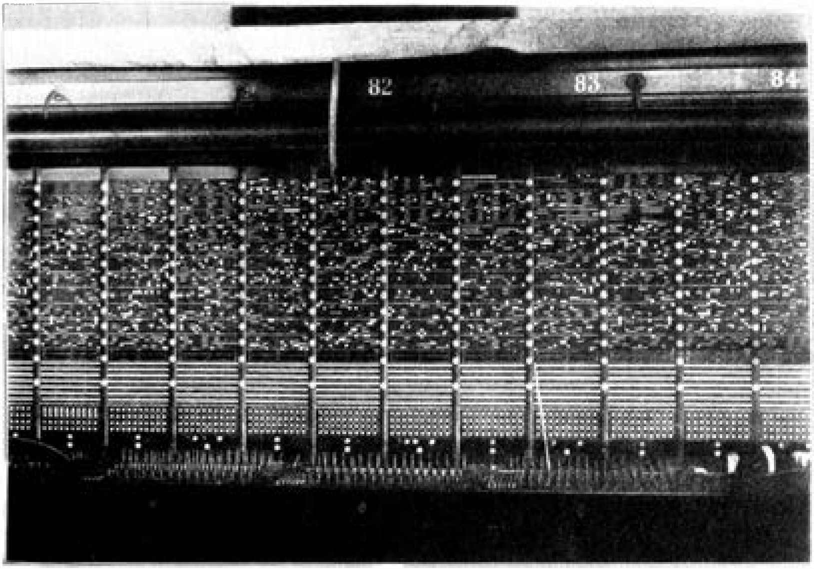


Fig. 6-15. Overall view of large common-battery switchboard, circa 1900.

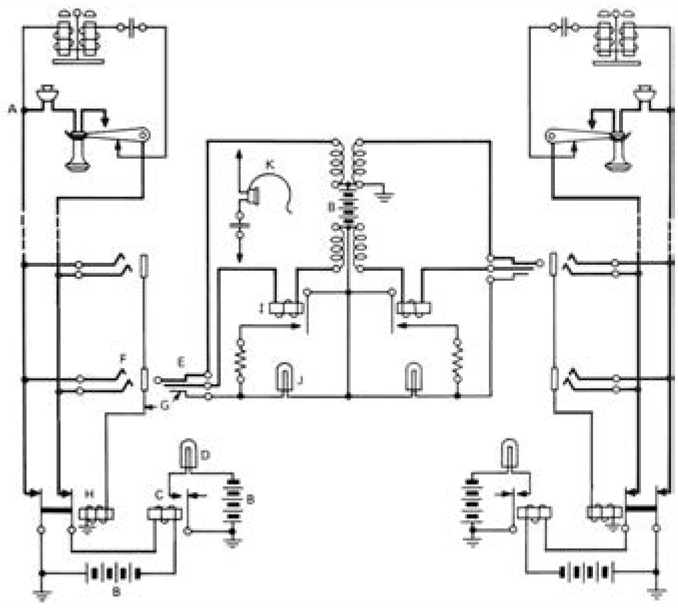


Fig. 6-16. Simplified circuit schematic of common-battery board. (Rhodes 1929, Fig. 44)

connects to the called line through the multiple by means of her front plug. Ringing current is sent out by depressing a ringing key (not shown). The cord circuit includes two supervisory lamps (J) associated with the two plugs. Each lamp is connected to battery through the cord circuit but controlled by supervisory relay I so that it is extinguished when the associated plug is connected to an off-hook subscriber and illuminated when the subscriber is on-hook.[34] Thus the operator can determine through the lamp signals when the called party answers and when either party hangs up. Repeated switchboard operation causes a flashing supervisory lamp, thus indicating the need for assistance.

Long before this point many readers may have decided that the above description far exceeds their interest in manual boards. For such readers let us merely point out that by 1900 complex operating procedures (as required with early boards) had been reduced to a few simple manual operations controlled by lamp signals. Each call was completed by inserting two plugs in appropriate jacks and operating a listening and ringing key. Complex circuitry performing a number of automatic operations had reduced the complexity of the operator's procedures. We should also note that the user's operation had been reduced to the minimum effort ever achieved. He merely lifted his receiver and verbally informed the operator of his wishes. When switching was automated, the need for adding dialing to the user's effort was a matter of considerable concern to some telephone administrators. Fortunately the user recognized the need as valid and other benefits from automation helped to compensate for the nostalgic days when the operator gave personal service in response to verbal requests.

3.3.3 Large-Exchange Systems

So far we have discussed the boards applicable to a single-office exchange. In a common-battery board of this type, such as the No. 1A introduced in 1897, the operator answered calls through answering jacks placed within her reach near the bottom of the board[35] and completed them through the subscriber's multiple placed above the answering jack. In multioffice exchanges, completion to distant offices was over trunks with access by means of jacks placed above the answering

[34] Relay I, when operated, does not cut off all current from lamp J but reduces it by a shunting resistance to the point where it is not visible. By keeping the lamp warm during the "off" condition, it responds more rapidly to intermittent signals.

[35] Initially there was only one answering-jack appearance for each subscriber line. Later (about 1907) they were multipled so each appeared in several sections of the board, but one appearance was always designated as first priority for answering. With this arrangement a momentarily overloaded operator could be relieved by a less busy one at another section. In addition, in a light-traffic period one operator could handle a large number of answering jacks without leaving her position.

jacks and below the multiple. The trunk jacks were multipled through the various sections of the board so as to be available to all operators. In such offices it became desirable to concentrate the incoming trunks from all offices in special operating positions used only to complete calls from the distant offices. These positions were designated the No. 1B board and were equipped with subscriber multiple jacks only, the incoming trunks being terminated by cords and plugs on the keyshelf. These two types of boards came to be known by the shortened designation of A and B boards and the same designations were used for operators, i.e., an A operator was one who answered calls, and a B operator completed calls from an incoming trunk. A and B board positions can be compared by referring to Figs. 6-17 and 6-18.

The arrangement just described (using an A board having answering trunk and completing jacks) was usually equipped with a subscriber multiple for up to about 8,000 lines. However, as the number of offices increased, more space was required for trunk jacks and less space was left for the subscriber multiple. If this trend had been continued without limit, the line capacity of offices would have been reduced, thus adding to the number of trunks required. To avoid this situation it became the practice in large cities to eliminate the subscriber multiple in the A board and trunk all calls to a B board, even those originating in the same office.[36] This scheme provided space for a very large number of trunk appearances and at the same time made it possible to increase the size of the subscriber multiple to 10,500 lines since completing jacks appeared only at the B board where all the space on the face of the board was available for this purpose. This arrangement became standard in large cities. The small exchanges continued to use A boards with subscriber multiples, the choice between the types of boards being determined largely by economic factors.

Communication between A and B operators was initially by means of call circuits. With this scheme, a separate talking circuit was permanently connected to a B operator's telephone and extended to all A operator positions at all offices having access to the trunks terminating in front of the B operator. The A end of the call circuit was normally open but could be connected to the A operator's telephone by operating a push-button key. Each A position had access to a call circuit to each office for which it had trunks. The call-circuit buttons were grouped together at the left of the position, as shown in Fig. 6-17. Operation was as follows: The A operator upon answering a call determined the office and the line number desired. She then depressed the appropriate call-circuit button,

[36] In a metropolitan area, such as New York City, as many as 90 percent of the calls originating in an office were to subscribers in other offices. Obviously, the retention of the A-board multiple in order to complete 10 percent of the calls would provide only small benefit at a high cost.

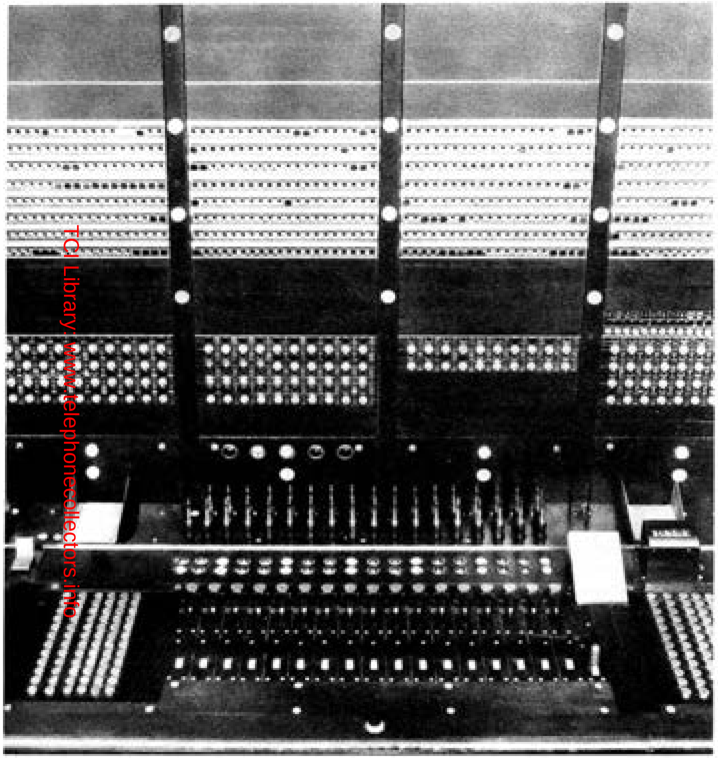


Fig. 6-17. No. 1 switchboard—"A" position.

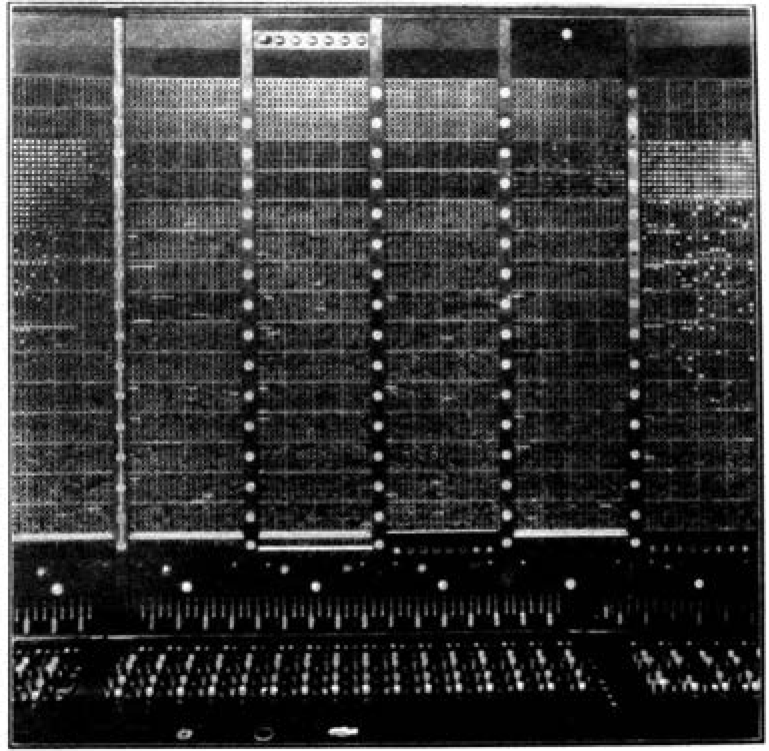


Fig. 6-18. No. 1 switchboard—"B" position or trunk section.

waited until the circuit was idle, and then by voice passed the name of the calling office and the number in the called office to the B operator. The latter observed her trunk cords, selected an idle one, and gave the trunk number to the A operator. While the A board connection was being made, the B operator tested the called number for busy and, if free, completed the call by inserting the trunk plug into the appropriate line jack, following which she rang the called party.[37] Status of the call was indicated by lamps so that verbal communication was kept to a minimum. Supervision remained the responsibility of the A operator, her cord signals being controlled by the connected subscribers. When she took down her cord, the B operator received a signal and disconnected the trunk from the called line.

About 1924, a number of cities had grown to the point where call circuits between A and B operators became unwieldly and the "straightforward" system, which used the trunk itself for operator communication, was introduced. Additional circuitry was added so that the A operator could determine which trunks were idle. She then selected one of these and at the far end the B operator's set was connected (either automatically or in response to a light signal). As soon as this occurred, the A operator received a tone signal and passed the called number orally. It was not necessary to pass the calling office name with this system since the A operator had already selected the proper trunk. Completion of the call and supervision were the same as with call circuits.

Direct trunking between A and B boards was satisfactory for most exchanges but trunks to distant offices, in large metropolitan areas, were often used inefficiently because of the light traffic between remote areas.[38] To solve this problem, the tandem principle (discussed in Section II) was introduced about 1908. This required the use of a third type of board, to switch incoming to outgoing trunks, interposed between the A and B boards. Functionally, these "tandem" boards were very simple since they were only required to establish connection between a relatively small number of incoming and outgoing trunks. However, the early tandem boards, using call-circuit trunking, introduced operating complications and other problems because the most practical operating method involved the A operator passing the called

[37] Originally, manual ringing was employed at the B board as well as the A board but about 1915 the B cords were equipped for automatic ringing. With this arrangement, ringing automatically started when the plug was inserted in the jack and stopped when the customer answered. On party lines the proper ringing selections were made by manually operating appropriate keys prior to making the jack connection. If the party was busy, the trunk plug was inserted in a "busy" jack which sent a busy signal to the calling subscriber and the A operator.

[38] Section 3.6 discusses the manner in which the amount of traffic affects the number of trunks required between two points and the efficiency with which the trunks can be used. As indicated in that section, considerable penalties result from the use of small trunk groups.

number to the tandem operator after being assigned a tandem trunk. The latter operator, in turn, obtained a completing trunk assignment from the B operator and passed the number to the latter who completed the connection. This tandem operation involved a considerable amount of additional operating time and the second repetition of the number was a potential source of error.

Tandem operation was made much more practical by two developments in the early twenties. One was the introduction of semimechanical tandem offices and the other was straightforward trunking. The former will be discussed in Section 4.3.5.3 but the benefits of straightforward trunking are properly noted here. With its use, the operating functions at the tandem board became very simple in the ultimate state of development. Incoming trunks terminated in cords and were selected by the A operator who received a signal when the tandem operator was connected by automatic circuitry. The A operator passed only the name of the called office and the tandem operator made the connection, her telephone being automatically disconnected when she did so. The call was then handled by the A and B operators in the same manner as with direct trunks. This simplified procedure reduced errors by eliminating the need for number repetition, increased greatly the number of calls per hour handled by the tandem operator, and resulted in a tandem board relatively uncomplicated in appearance (Fig. 6-19).

Tandem trunking, although requiring a third operator, brought about a number of benefits in large exchanges. Trunks were used more efficiently between offices with low traffic and even though direct trunking was continued on high-traffic routes, tandem trunks provided a means for handling overflow traffic on these routes and were often used for light-period traffic in order to concentrate operation at a few positions. Tandem operation also was used to replace toll board operation on short hauls, the A operator, in this case, doing the timing and ticketing. By 1930, some 20 tandem boards had been installed in the Bell System, of which five were in New York City (two of these were mechanical and three manual). In a few cases, two tandem boards were interposed between the A and B board but this "double tandem" operation was not a completely satisfactory scheme with the art of the 1920s and 1930s because of the additional call setup time, and greater chance for error. In addition, very low-loss trunks were required to provide satisfactory transmission on double tandem calls.

The preceding material has described the basic features of the common-battery manual boards introduced in standardized form beginning about 1897. The No. 1 board, produced at that time, and its successors continued in wide use for well over 40 years. While the basic principles remained much the same over this long period, many features were introduced to

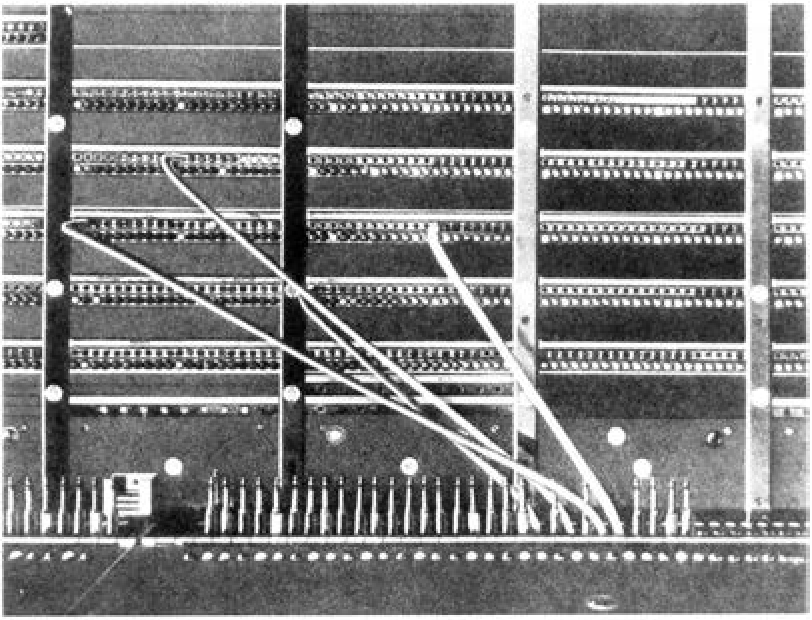


Fig. 6-19. Manual tandem board.

reduce cost by simplifying apparatus components, improving manufacturing methods, and simplifying operating procedures. Some of the last, such as straightforward trunking, have already been mentioned. Other important steps were the introduction of machine ringing, automatic connection of the operator telephone, and audible "ring-back" to inform the calling subscriber that ringing was in process. For those interested in further details of Bell System switchboards during this period, their characteristics are summarized in Fig. 6-20.

In closing our discussion of boards used for exchange switching we should mention that unlike toll boards, discussed in Section 5.4, they commonly used one-way trunks. It is obvious that trunks between offices A and B are required for traffic originating in each office and in theory the same physical facilities could be used for calls originating at either office. This "two-way" operation uses the circuits very efficiently (particularly on light-traffic routes) and was quite commonly used on long-distance circuits during the first 50 years of telephony. However, signaling is either complicated or slow and on local calls the use of one-way trunks has been the usual practice, particularly with common-battery systems. With this scheme, physically separate trunk-groups

<table>
<tr><th>Code</th><th>Year</th><th>Capacity (Lines)</th><th>Volt–age</th><th>Ringing</th><th>Battery Feed</th><th>Line Signal</th><th>Cord Signal</th></tr>
<tr><td>1</td><td>1897</td><td>8,000 – 10,500</td><td>24</td><td>Manual</td><td>Repeat Coil</td><td>Lamp</td><td>Lamp</td></tr>
<tr><td>8</td><td>1902</td><td>1,500</td><td>36*</td><td>Manual</td><td>Capacitor Line Ckt.</td><td>Lamp Series Jacks</td><td>Lamp</td></tr>
<tr><td>9D</td><td>1903</td><td>800</td><td>24</td><td>Manual</td><td>Bridged Imped.</td><td>Magnetic Series Jacks</td><td>Magnetic</td></tr>
<tr><td>9C</td><td>1905</td><td>800</td><td>38</td><td>Manual</td><td>Bridged Imped.</td><td>Magnetic Series Jacks</td><td>Magnetic (In Key–Shelf)</td></tr>
<tr><td>10</td><td>1906</td><td>1,600</td><td>24</td><td>Manual</td><td>Repeat Coil</td><td>Lamp Series Jacks</td><td>Lamp</td></tr>
<tr><td>1C</td><td>1909</td><td>8,000– 10,500</td><td>24*</td><td>Machine</td><td>Repeat Coil</td><td>Lamp</td><td>Lamp</td></tr>
<tr><td colspan="8">Automatic Listening, Audible Ringing</td></tr>
<tr><td>1D</td><td>1918</td><td>3,000</td><td>24*</td><td>Manual</td><td>Repeat Coil</td><td>Lamp</td><td>Lamp</td></tr>
<tr><td>11</td><td>1924</td><td>A = 4,000†
B = 10,500</td><td>24</td><td>Machine</td><td>Repeat Coil</td><td>Lamp</td><td>Lamp</td></tr>
<tr><td colspan="8">Prepay coin, MR, Service Observing, Call Indicator (Type E) Automatic Listening, Manual Ringing Start and Key Listening (Type G), False Recall, Automatic Peg Count, Call Distribution (8).</td></tr>
<tr><td>12</td><td>1932</td><td>640‡</td><td>48</td><td>Manual</td><td>Bridged Imped.</td><td>Lamp § Series Jacks</td><td>Lamp</td></tr>
</table>

* 48V on toll cords

† 7,200 (1927), 10,500 (1929) – A board

‡ 1,400 (1935), 2,000 (1938)

§ Lamps (1,000 ohms) in line – no line relays

Fig. 6-20. Characteristics of Bell System common-battery manual switchboards.

are used for traffic originated at the two ends. With high-density traffic, common in exchange service, there is little loss in efficiency and a considerable simplification in signaling. The use of tandem service tends to eliminate the few cases where light traffic would make two-way trunks preferable.

3.3.4 Supplementary Switchboards

The switchboards just described provided the backbone network for exchange communication. In addition to these boards, a communication system requires supplementary facilities to provide functions related to but not always involved in establishing a communication path. These special-purpose switchboards, commonly referred to as desks, are used to provide information and intercept service, supervision of performance, and maintenance of lines and service.

The information desk is used when the subscriber needs a number not listed in the directory available to him. In the early days the switching operator provided this information but since about 1910 it has been handled by special operators to whom the subscriber is switched by the A operator. The No. 1 and No. 2 information desks are shown in Figs. 6-21 and 6-22, respectively. It will be noted that the latter uses rotary card index files. Such files provide a quick and more readily updated source of information than the books used at the No. 1 desk but require much more space and are impractical for covering very large cities.

Intercept desks, as illustrated in Fig. 6-23, are used to provide information to the user on changes in service made subsequent to the latest directory issued. When a line has been temporarily or permanently disconnected, the number changed, or if for any reason the directory information is not valid, the line jack is marked in the multiple (by a closure plug or similar arrangement) and the A operator switches the call to the intercept operator. By inquiring the number called, she can determine, from her file of all lines so marked, the status and inform the caller accordingly.

The chief operator's desk (Fig. 6-24) provides a turret by means of which she can talk to supervisors over her telephone or can monitor the work performed by operators and supervisors. Supplementing this form of supervision is the use of routine service-observing carried out by means of either permanent or portable desks. These desks provide facilities for an observer to monitor preselected lines in the office to determine performance such as time for operator to answer, time to perform switching operations, ringing, etc.

Maintenance is carried out with the use of the two desks shown in Fig. 6-25. All calls reporting trouble are routed through the desk of the repair clerk who records the complaint, selects the maintenance card for the line, and passes the material to the test man at the local test desk. This desk is equipped with a volt-milliammeter and other facilities to determine the state of the line. If more detailed tests are required or if circuit changes are to be made, these are carried out by a craftsman in the frame room under direction of the test desk man. Communication between them is by telephone set or loudspeaker system. The test desk contains the necessary keys and jacks for communicating with the frame room, various supervisory personnel, local and toll switchboards, and other test desks.

3.4 Apparatus Components

So far we have been discussing switching principles and, very broadly, the switching and signaling arrangements required for their implementa-

Fig. 6-21. No. 1 information desk.

Fig. 6-22. No. 2 information desk.

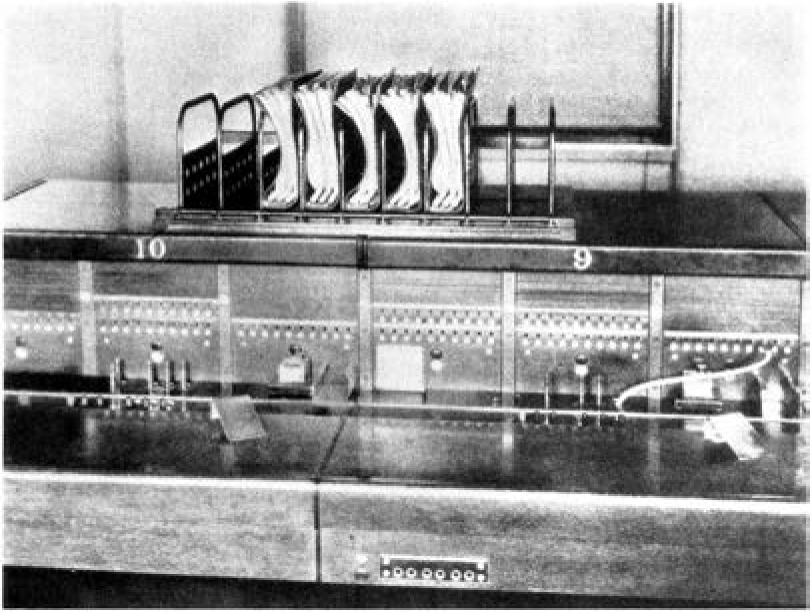


Fig. 6-23. Intercept desk.

Fig. 6-24. Chief operator's desk.

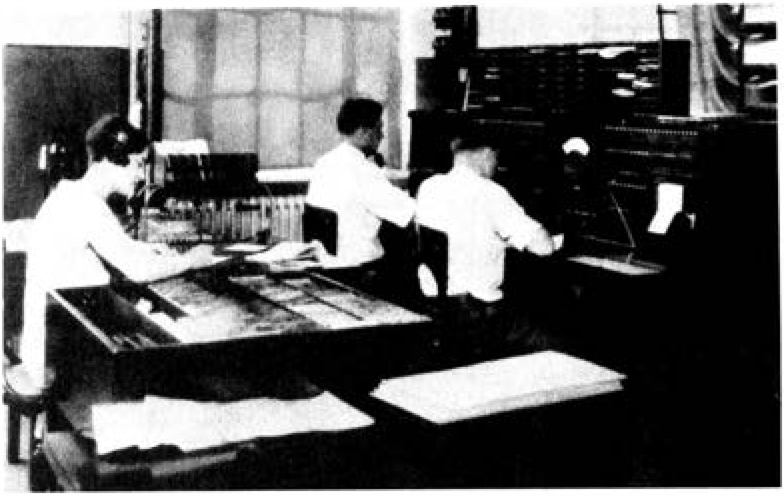


Fig. 6-25. Repair clerk's and local test desks.

tion. From this discussion it will be apparent that the evolution of these systems from crude arrangements requiring complex operating methods to the compact and efficient switching systems of the early 1900s was highly dependent on the development of new and improved apparatus components. While some of these have been mentioned in general terms during the preceding discussion, further details may be of interest to many readers and will be covered here.

At the time of Bell's invention, electrical communication apparatus consisted largely of that required for telegraph operation, and for the simple signaling systems used in homes and public buildings to announce visitors, summon assistance, or report emergencies. The apparatus used, other than the telegraph transducers, consisted largely of pushbutton switches, vibrating bells, annunciators to record signaling operations, and relays to extend communication range by providing means whereby weak currents attenuated by passage over long lines could control the high current (supplied locally) required for operating audible or visual signaling devices.[39] Current for operating these devices was furnished by many forms of primary batteries.

Connections between apparatus components were commonly made on a more or less permanent basis by means of soldering or screw

[39] Relays were, as used at this time, a form of regenerating amplifier which was practical for signaling purposes long before means became available for the amplification of the weak and complex voice currents of telephony.

terminals, although simple switching arrangements were occasionally used, both in telegraph systems and laboratory equipment, employing rotary switches or metal-peg connectors that could be plugged into holes in heavy brass blocks.

All of these devices were employed in early switching systems but, as in the field of transmission, it was soon found necessary to modify them to meet telephone demands and ultimately to invent completely new apparatus. The new requirements arose because of the need for providing new operating functions, as well as achieving greatly reduced size and cost and a degree of reliability far beyond that of pretelephone equipment. The evolution of some of the components which entered into the sophisticated manual switching systems used after about 1890 is described below.

3.4.1 Plugs and Jacks

Although metal pegs had been used for some time in making temporary connections, the true prototype of the telephone plug and jack combination was Scribner's jackknife switch. A photograph of this apparatus is shown in Fig. 6-10 but the cross section, shown in (a) of Fig. 6-26, will perhaps clarify the mechanism. The basic principle was that in the normal state a circuit was closed through the contacts shown. Inserting a plug opened this circuit by moving the jackknife blade, and established a connection between the frame and the plug. Thus this arrangement went beyond the earlier peg connectors and opened a simple, normally-closed switch when the plug or peg was inserted. Considerable improvement over Scribner's device resulted from J. C. Warner's 1882 invention of the plug and jack illustrated in (b) of Fig. 6-26.[40] At first this provided only a space reduction, the function performed being much the same as the jackknife switch, i.e., making a simple connection and performing a signal-related switching operation. But the basic design was capable of modification so as to make multiple connections and operate several switching contacts, as shown in (c) of Fig. 6-26. In addition, by reducing the diameter of the plug, rearranging the terminals, and mounting the jacks in strips [as shown in (d) of Fig. 6-26], it became possible to reduce the required switchboard space. Ultimately, two sizes of jacks became standard, each capable of making three connections designated tip, ring, and sleeve. The first two carried the metallic transmission circuit and the last the signaling circuit used within the central office. Some characteristics of these jacks and their associated plugs are given below:

[40] J. C. Warner; U.S. Patent No. 281,741; filed June 8, 1882; issued July 24, 1883.

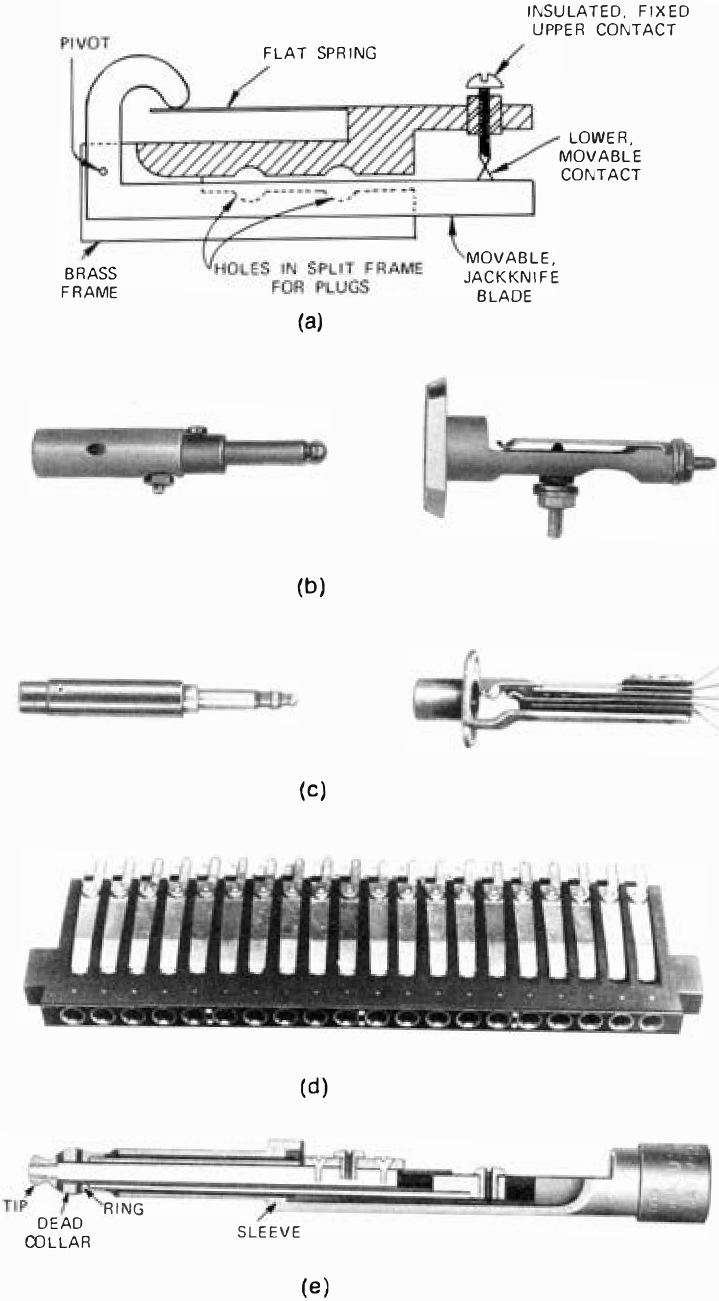


Fig. 6-26. Evolution of the telephone plug and jack. (a) Cross section of jackknife switch. (b) Warner plug and jack. (c) Later plug and jack designed to make multiple connections. (d) Strip-mounted jacks. (e) Cross section of 109-type plug.

Jack Spacing (In.)*	Jack No.	Date Intro	Plug Diam (In.)	Plug No.	Date Intro	Space Between Tip and Ring
7/16	49	1897	0.25	64	1897	Hard-rubber ring
7/16	49	1897	0.25	104	1900	Dead collar
7/16	49	1897	0.25	110	1905	Improved dead collar
3/8	92	1901	0.20	101	1901	Dead collar
3/8	92	1901	0.20	109	1904	Improved dead collar

* Minimum horizontal and vertical distance between axes of the jacks.

It was the 92-type jack that made it possible to build a switchboard multiple with 10,500 lines within the reach of an operator.[41] When the three-conductor jack was introduced about 1891 to provide for 2-wire (metallic) transmission, there was some difficulty in providing satisfactory insulation and avoiding clicks as the plug was inserted in the jack since there was a tendency to bridge over between the various jack contacts. The plug profile finally adopted is shown in (e) of Fig. 6-26. Illustrated is the relative diameter employed for the various contacts and the method of insulation. This plug used an insulated metallic band, called a "dead collar," between tip and ring, that acted as a spacer to prevent simultaneous contact between jack springs as the plug was inserted. Use of the metallic band also avoided the excessive wear which occurred on previous plugs using a wide collar of hard rubber.

At one time, jacks were one of the highest production items of the Western Electric Company. In 1917, for example, about 14,000,000 jacks were manufactured, about three-quarters of them being of the 92-type mounted in strips of 10 or 20.

3.4.2 Keys and Switches

Although some switching operations were performed by plugs and jacks, it was necessary to use manually operated switches to perform a number of short-term operations. The apparatus used usually was referred to as a key and was operated either by a lever or a push-button. The former provides great flexibility since it can be arranged to have three positions, e.g., center (neutral), forward, and back. It also can be arranged to be restored to neutral when pressure is removed or to remain in an operated state until manually restored.

[41] The 0.20-inch-diameter plug was the smallest used in Bell switchboards. Smaller diameters, permitting larger-capacity boards, have been used abroad but were not favored by the Bell System because of increased wiring and maintenance problems.

The first lever key was introduced in the Bell System about 1888 and is illustrated in (a) of Fig. 6-27. Operated in one direction this key connected the operator's set to the cord circuit and in the other direction it sent out ringing current. Later, when selective ringing was introduced, additional types of keys were needed and a smaller and simpler version, the 102-type shown in (b) of the figure, was introduced so that two could be installed one behind the other on the keyshelf. Still later, vertical springs were introduced, as shown in (c) of Fig. 6-27, to give more efficient utilization of keyshelf space. Space requirements and flexibility of use were further improved with the A-type key, introduced in 1912 and illustrated in (d) of Fig. 6-27. This was known as the universal-type key since a relatively small number of piece parts could be assembled so as to give a wide variety of switchleaf pileups of either the locking or non-locking type. A pushbutton variety also was introduced and the individual keys of both lever and pushbutton type could be assembled in a mounting plate in any desired combination, as illustrated in (e) of the figure.

The universal keys proved so useful that they were used in large quantities for many years and this general type still is being manufactured in small quantities. Before leaving this subject, we should note that the structure, consisting of a pileup of appropriately shaped flat springs with suitable contacts, was quite similar to, but more complex than, the jack structures.

3.4.3 Annunciators or Drops

The main purpose of plugs, jacks, and keys in the early switchboards was to provide the mechanism for establishing talking paths between calling and called subscribers and between subscribers and operators. However, they also formed part of the complex system which conveyed the various signals required for controlling the switching functions. Perhaps the most important single device used in early boards for signaling was the annunciator or drop.[42] These devices came into existence long before telephony, being used in large call-bell systems to indicate the particular source of the call. They were used without change in early switchboards but were soon modified to meet the special requirements of telephony for small size, operation on low current, and freedom from mutual interference when mounted in close proximity.

The basic operating scheme is illustrated in Fig. 6-28. A small plate, hinged at the bottom, was normally held in a vertical position by means of a latch. A current through the electromagnet pulled up an armature

[42] These devices were originally referred to as "annunciators" in the United States and "indicators" abroad. For reasons that will become apparent, they were also referred to as "shutter drops" and soon the short term "drops" became universally used.

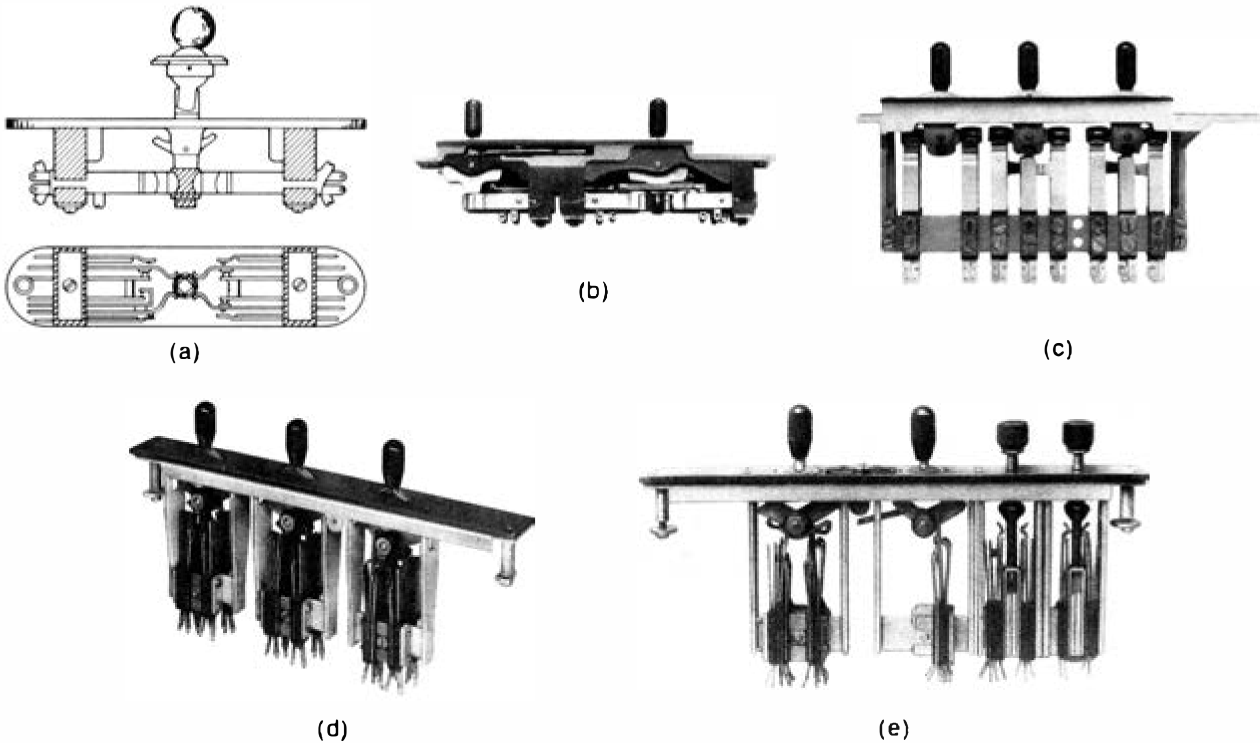


Fig. 6-27. Evolution of the telephone key switch. (a) 52-type key, the first lever key containing two pileups. (b) 102-type key mounted two in-line on keyshelf. (c) 463-type key with vertical springs to provide better keyshelf utilization. (d) A1, universal-type key. (e) Universal keys of both lever and pushbutton type.

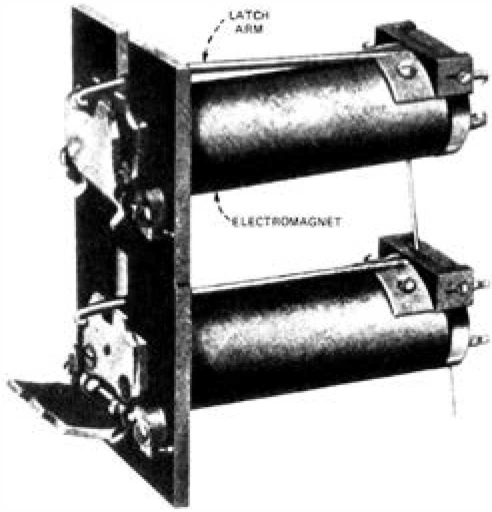


Fig. 6-28. Typical latch-type drop.

at the rear of the drop, lifting the latch arm and allowing the hinged plate to fall forward. This notified the operator that her assistance was required in establishing a call or in disconnecting it, depending on the drop location in the line or recall position. After taking the necessary action, the operator reset the drop manually to the normal position. Many variations were introduced. E. P. Warner in 1880[43] invented a drop without a latch which depended on the drop being pivoted so that the normal position was a few degrees back of vertical. The electromagnet pulled the drop plate forward past the vertical, causing it to fall under the force of gravity. This was a highly sensitive drop but subject to false operation due to air movement or slight mechanical blows. In 1889, J. C. Warner[44] invented a latch mechanism which required only small magnetic forces to actuate and he also provided a soft-iron shell around the electromagnet. The latter acted as a shield, thus preventing false operation through inductive coupling between adjacent magnets. The shell also improved the magnetic circuit and, together with the improved latch, provided a very sensitive and stable device.

Later, electromagnetic signals were used in which current in the magnet lifted a small target which remained visible as long as the cur-

[43] E. P. Warner; U.S. Patent No. 232,093; filed April 14, 1880; issued September 7, 1880.
[44] J. C. Warner; U.S. Patent No. 477,616; filed June 17, 1889; issued June 21, 1892.

rent flowed. There were several types of these signals but for convenience we shall refer to all of them as "self-restoring drops." Various forms of control were used. In some cases they were actuated directly by line current. In other applications the drop was actuated by a spurt of ringing current and locked itself in place by closing contacts when the target was pulled up. A locally supplied current held the target in place until operation of a jack or key by the operator opened the local circuit. Such a drop is shown in Fig. 6-29. Self-restoring drops were not used extensively since the switchboard lamp, to be described later, proved a much more satisfactory indicator.

3.4.4 *Relays*

The relay was originally very similar structurally to the drop but functionally very different. It was invented by Morse and covered by his original telegraph patent. In his conception it provided a means for extending the range of telegraphy. He realized that telegraph currents were attenuated by the resistance of the lines as they were extended, and as the current approached the minimum actuating point a relay was introduced. This device, activated by the weak incoming signal, closed a contact which sent out a new high-level current from a local source. This same principle of using a small current to control one or more larger ones was the main purpose for which the relay was used in telephone signaling for many years.[45]

[45] The relay proved a very versatile device and later became a part of "counting" and other "logic" circuits which are an essential part of machine switching and computer systems. In these latter applications, electronic counterparts have now superseded the mechanical relay to a very large extent.

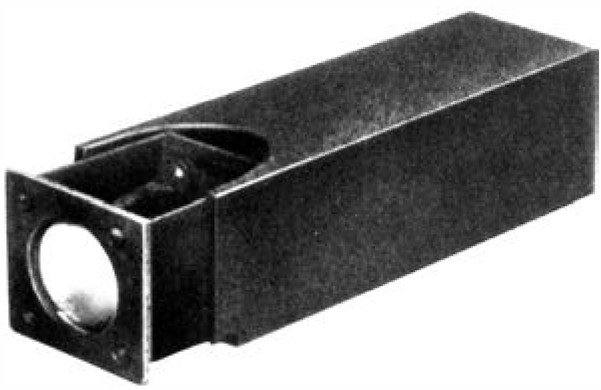


Fig. 6-29. Self-restoring drop.

Morse's original relay was a ponderous device weighing several hundred pounds and, although much simplified by Alfred Vail and others, it was still a large piece of apparatus weighing 3 to 5 pounds in the early days of telephony. With the advent of common-battery service and machine switching, large numbers were required for telephony[46] and obviously devices of small size, low weight, and high efficiency were needed. By the 1920s, the weight of a telephone relay had been reduced to about 3½ ounces with corresponding improvements in size and efficiency.

One of the earliest forms of relay, used in the Worcester, Massachusetts, board of 1896, is shown in Fig. 6-30a. This device, employing a reed-type hinge for the armature, was very rugged but rather insensitive and slow to operate and was, accordingly, unsuitable for line and supervisory relays that needed to follow the switchhook rapidly. To meet the special requirements for these relays, one using a knife-edge armature support was developed. The schematic of this relay, shown in Fig. 6-30b, also serves to illustrate the basic principles of relay construction. In this figure, T_3 and T_4 are the terminals for the control current which through winding W on core C provides the magnetic force to pull up armature A. This, in turn, closes contacts C_1 and C_2 and completes the circuit through terminals T_1 and T_2. Photographs of such relays are shown in Fig. 6-30c. The lower photograph shows a cover used to protect the contacts in some circumstances. Fig 6-31a shows a more advanced form of relay in which the armature actuates multiple contacts. This particular relay is of the "transfer" type, using three layers of springs. One set of such spring-contacts is shown schematically in Fig. 6-31b. When the armature is pulled up, an insulated stud pushes the middle spring away from the bottom spring and brings it in contact with the top spring, transferring the current path.

Many kinds of relays were ultimately required, involving fast action, slow action, high and low impedance, and many contact arrangements. Space will not permit further discussion of these many types except to note that 50 years after the telephone invention Western Electric was manufacturing about 100 basic types of relays which, with various contact arrangements, amounted to about 3,500 different kinds.

Obviously, the mechanical structures illustrated in Figs. 6-30 and 6-31 did not lend themselves to economical large-scale production to meet a variety of needs. Knife-edge supports required much maintenance and

[46] In the 1920s, a typical manual connection involved about 20 relays and close to 150 were required in a machine-switched call. Some long-haul calls with several offices in tandem involved as many as 300 relays. A typical 10,000-line office required the installation of 50,000 relays if it were of the manual type and roughly three times as many for machine switching. The Western Electric Company production was about 5,000,000 relays per year in the 1920s.

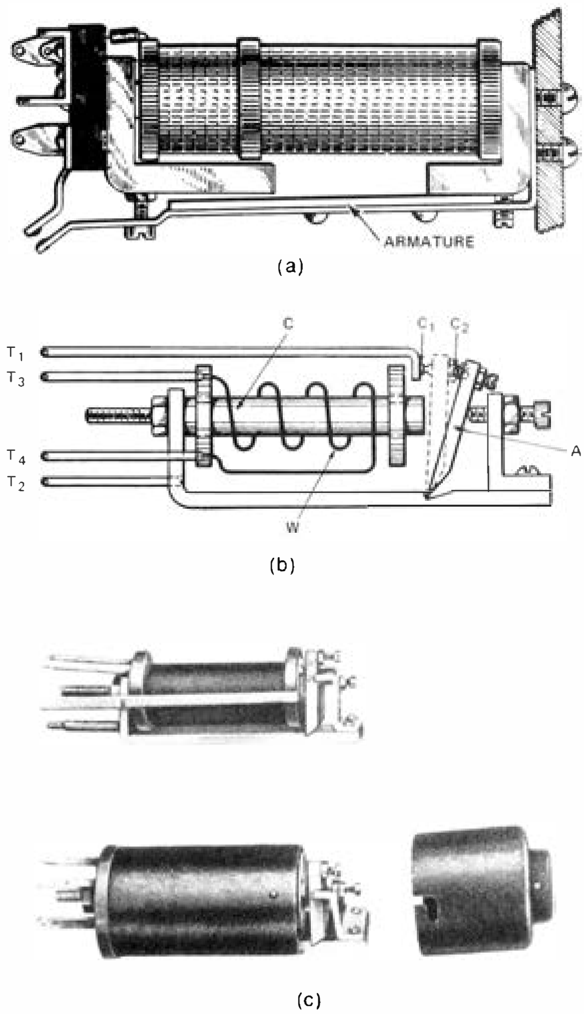


Fig. 6-30. Early telephone relays: (a) reed-armature relay; (b) schematic of single-contact relay with knife-edge armature support; (c) implementation of single-contact relay. (Shackleton and Purcell 1924, p. 10)

manufacturing costs were high because of the numerous hand and screw-machine operations required. In 1910, E. B. Craft proposed the use of a so-called "flat-type" relay largely made up from piece parts produced at high speed by means of punch presses. By this time, analytic techniques were being used in relay design and had shown that less

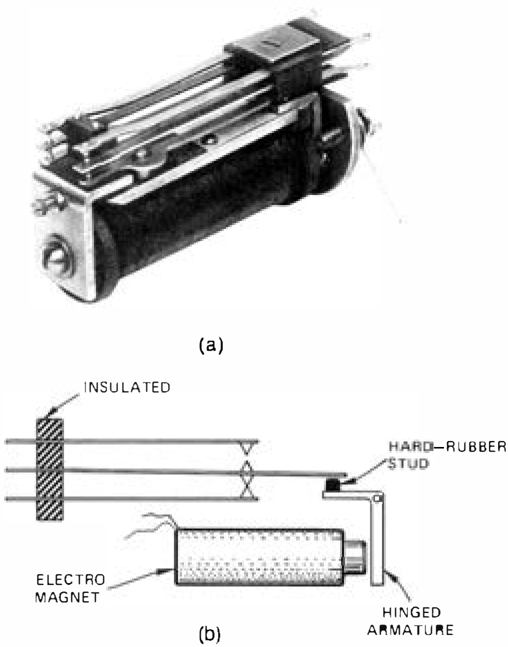


Fig. 6-31. (a) Early multicontact telephone relay. (b) Schematic of transfer-type relay.

iron could be used, particularly with new materials such as silicon steel. Craft's original sketch proposing the flat-type relay, shown in (a) of Fig. 6-32, proved to be remarkably prophetic of the ultimate design. It involved a return to the reed-hinged armature but by using the new materials and analytic design techniques it was possible to meet the stringent requirements for line relays. The same structure was used for cutoff relays and later a similar arrangement, the D type, was used for supervisory relays. The last type required individual covers to serve as electromagnetic shields to prevent crosstalk between closely mounted relays. For other functions it was possible to use a common cover for a multitude of relays. Some of these early punch-press relays and their mounting arrangements are shown in (b), (c), and (d) of Fig. 6-32.

These early flat-type relays were highly successful and about 1915 the E-type relay was developed using the same general principle but modified to serve as a general utility relay which could use a variety of windings and have sufficient magnetic force to operate numerous pairs of springs and thus permit an almost unlimited number of contact combinations. This relay, shown in Fig. 6-33, was manufactured with about 3,000 varieties of windings and spring arrangements. Later the relay was improved by compressing the core portion of the punching into a roughly round cross section, thus providing a more efficient winding requiring less wire for a given performance. Such relays, introduced in 1923, were designated the R type. They were very widely used but did not fully supplant the E type.

3.4.5 Switchboard Lamps

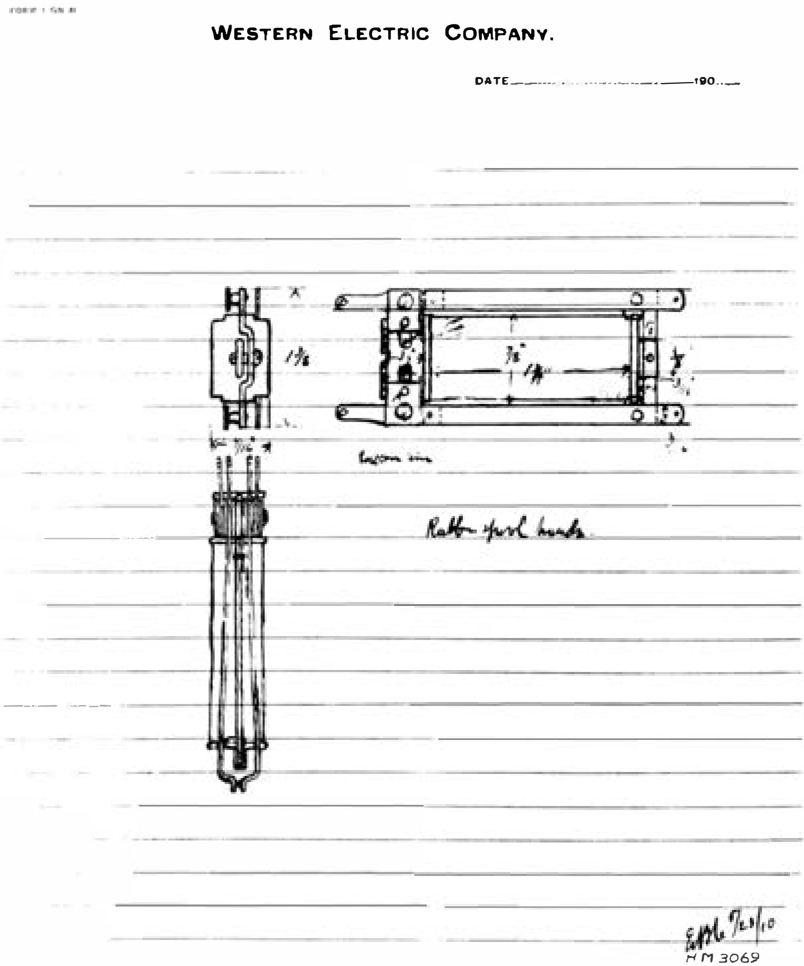
As indicated previously, the annunciator which antedated telephony was modified to meet the requirements for a telephone signal indicator. Such telephone "drops" proved very satisfactory in the early days of telephony since they were responsive to magneto signals over long loops. However, by the middle 1890s, telephony had advanced to the point where a new device was needed. The growth of switchboards placed a premium on space and a smaller indicator was clearly desirable and, as indicator size was decreased, it became important to increase the contrast between on and off signals. At the same time the need increased for an indicator that was restored automatically as a part of the plug-jack connecting operation. This need arose not only to save operator time and effort but also because of the difficulty of using manual restoration on the compact indicators required on boards serving large numbers of subscribers. A miniature electric lamp seemed to be the ideal replacement for the drop. However, it had several inherent disadvantages. It could not be substituted directly for the drop on a magneto line since the short spurts of current from the magneto would give at best only a flash of light that would not readily be detected by the operator.[47] In addition, the light intensity varied greatly with the current through the lamp and hence was highly dependent on the resistance of the loop and the intensity of the magneto operation. Fortunately, these problems were readily solved in the common-battery offices that came into use in the late 1890s. These offices required sizeable power plants to supply the talking and signaling power and the burden of powering small lamps was not too severe. The loop signaling current varied so greatly with loop resistance that it could not be used directly for operating lamps[48] but, fortunately, relays were available that operated over a wide range of current and these could be used to control lamp signaling circuits as described in Section 3.3.2.

Lamps were used as switchboard line signals as early as 1894.[49] In

[47] Lock-up circuits could have been used, as they were with self-restoring drops, but this required relay arrangements which were more compatible with common-battery than with local-battery offices.

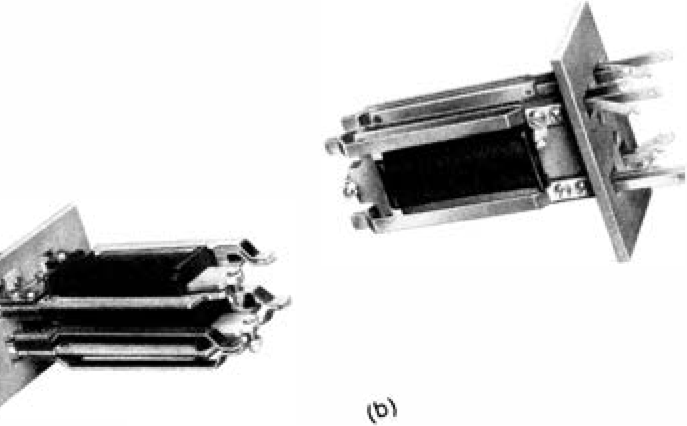
[48] In some cases it was possible to use lamps as line signals by placing them in series with a loop but, generally speaking, a relay-type line signal was usually required.

[49] The use of lamps was first proposed by J. J. O'Connell of the Chicago Telephone Company. The original suggestion, in 1888, covered a burglar-alarm application. About 1890, he used a lamp in a trunk-signaling circuit.

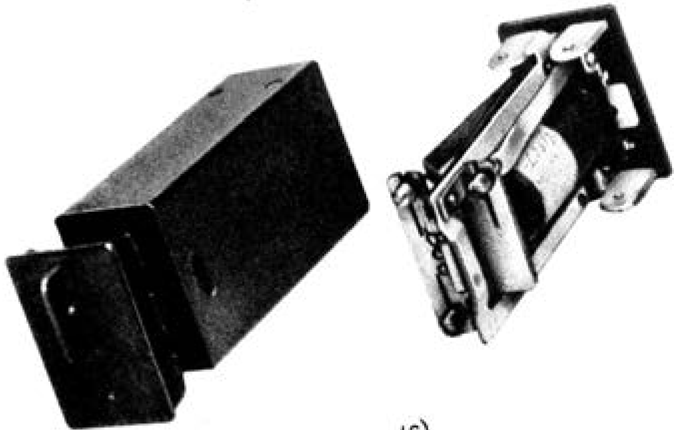


(a)

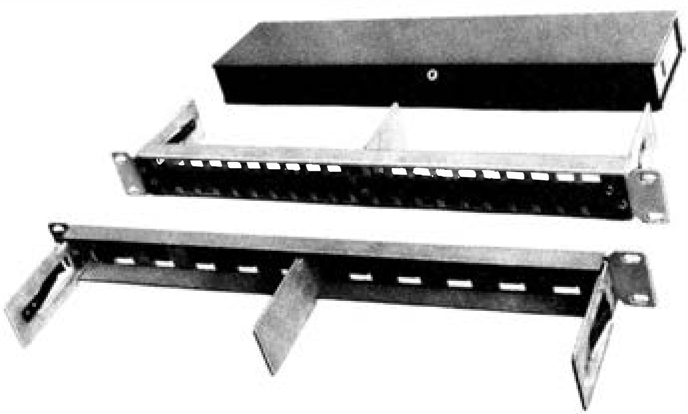
Fig. 6-32. Early flat-type relays. (a) E. B. Craft's original sketch embodying his invention of the flat-type relay. (b) First punched-frame (flat-type) relay. (c) B-type relay. (d) Mounting plates for strip of flat-type relays. (Shackleton and Purcell 1924, pp. 16–18; Mead 1926, p. 84)



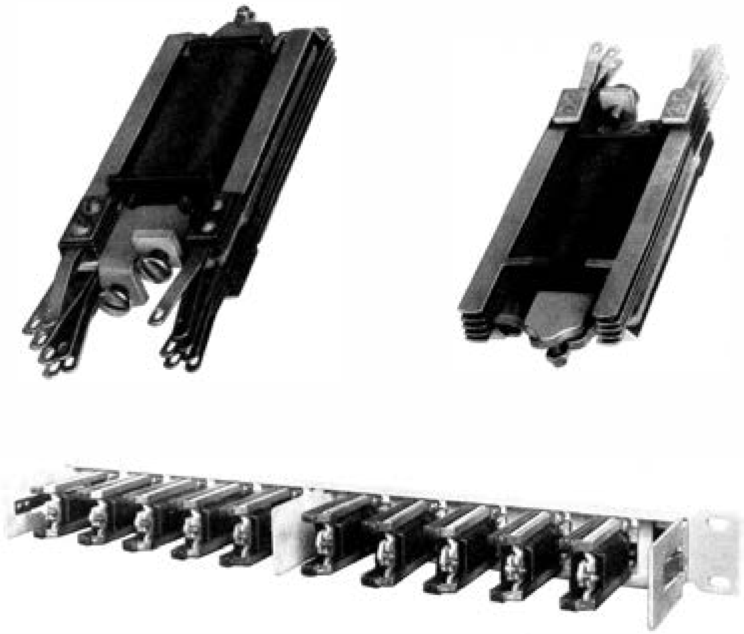
(b)



(c)



(d)



10 E–TYPE RELAYS ON 737 B MOUNTING PLATE

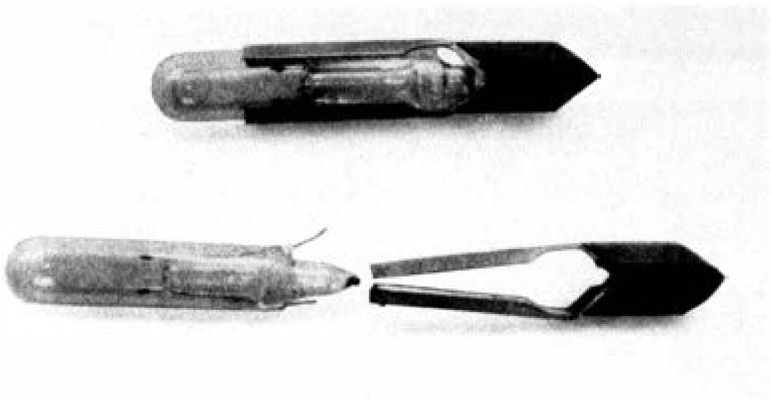
Fig. 6-33. The E-type (universal) flat relay.

addition to designing a rugged, low-powered lamp of small size, it was necessary to develop a means for mounting them so that they occupied only a small area on the face of the board. Over a period of years several designs were developed, as shown in Fig. 6-34. Around 1900 the scheme that was to endure for well over 50 years was introduced.[50] This lamp, known as the 2 type, is illustrated in Fig. 6-35. As shown, it consisted of two main parts, a bulb containing the filament with appropriate lead-out wires, and a base which supported the bulb and provided a means for making contact with the socket. The bulb itself was about ¼ inch in diameter and about 1¼ inches long. The base consisted of a small, V-shaped, wooden block with two tinned-brass terminals. The latter were cemented to the bulb to support it and were soldered to the lead-out wires. They also made contact with the socket. The overall size of the assembly was about 1¾ inches long by 5⁄16 inch in diameter.

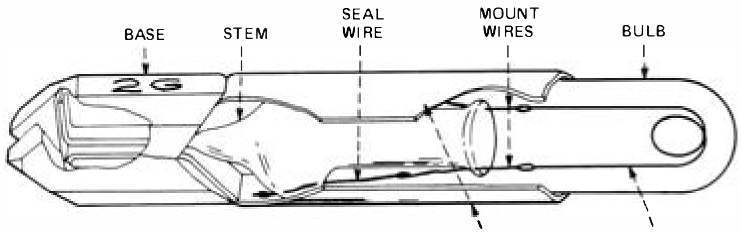
Originally, the filaments were made of carbonized cellulose. This

[50] This type of lamp is still widely used in the 1970s, but much smaller modern devices are used where space is at a premium.

Fig. 6-34. Various changes in switchboard lamps. The original type employed is shown in upper left-hand corner. (A. V. Abbott, *AIEE Trans.* 1898, Fig. 7)

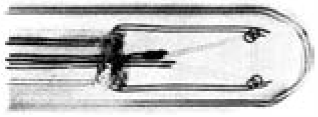


(a)



(b)

(c)



(d)

Fig. 6-35. Construction details of the 2-type switchboard lamp: (a) Complete lamp assembly (above), base and bulb (below), showing brass terminals; (b) constituent parts of lamp assembly; (c) carbon-filament lamp; (d) enlarged section of tungsten-filament lamp showing method of supporting filament. (Adapted from illustrations in *Western Electric News* of October 1928, article by J. C. Wright in *Bell Laboratories Record* of March 1932, and unpublished manuscript by W. Willenbruch of March 1938.)

material has a high resistivity and thus it was possible to use a relatively heavy, rugged, self-supporting filament. About 1927, tungsten filaments began to be used. Lower resistivity of this material necessitated the use of very fine wire which required auxiliary support. The two filament materials have very different characteristics. Tungsten produces a visual signal over a wider range of applied voltages, is brighter, and withstands overload voltages better. It responds more quickly to the application of current and hence has better flashing characteristics. Carbon, on the other hand, performs better in some circuits which rely on a reduction in current for extinction since it will become invisible on current reductions which would cause a dim glow with tungsten.

An important requirement of all switchboard lamps is long life since their operation is essential to many switching functions. Lamps are designed therefore to be practically a permanent piece of equipment.

In 1925, about 30 different types of lamps were in regular manufacture, designed for nominal voltages from 4 to 48 volts and currents from 0.003 to 0.03 ampere.

3.4.6 Terminals and Distributing Frames

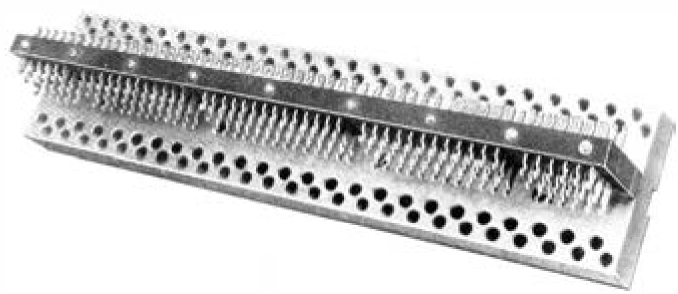
It is apparent that a switching system consists of a large assembly of apparatus components. The plugs and jacks of a manual system (and the switches of the machine system) are used to make the temporary connections required to set up the desired talking path, but the remaining elements are more or less permanently wired together. In some cases (the telephone plug is an example) screw terminals are used for connecting the attached wires but most connections are of the soldered type.[51] Many of these remain undisturbed during the working life of the component but some flexibility is introduced at various points in the system by using terminal strips which, although using soldered connections, are designed so that connections can be rearranged conveniently. Several strips of the type used in the early 1920s are shown in Fig. 6-36. Each of these strips has a maple base perforated with "fanning holes" through which the wires to each set of terminals are fed. This provides an orderly distribution of wires and facilitates tracing them when wiring changes are required. On top of the base is a pileup of hard-rubber strips slotted to carry the tin-plated brass punchings which form the terminals to which the wires are soldered. It will be noted that these punchings go through the rubber strips and thus provide a means for connecting wires on one side of the pileup to those on the other side.

[51] Beginning about 1952, a solderless-type connection was developed that had low contact resistance and reliability equal to or better than a soldered connection and this type is superseding the earlier arrangements used.

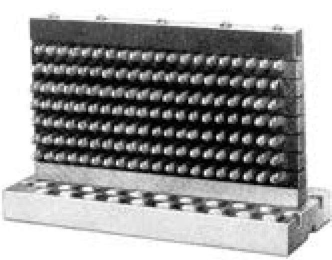
Terminal strips are used in many places in a telephone office, wherever wiring rearrangements are likely to be needed. Probably the most common use is on the distributing frames that provide flexible wiring arrangements between the outside cable plant and the switchboard jacks (or switches in the case of a machine system). Up until about 1890, lines entering an office had been distributed to the switchboards by means of connecting wires run in a direct but disorderly way between the boards and outside conductors (Fig. 6-37). The mass of wires became more and more confused as offices grew and, in 1891, Angus S. Hibbard invented an arrangement known as the Main Distributing Frame, or MDF, to bring order to this chaos. In the meantime, E. M. Barton suggested a similar arrangement, which has become known as the Intermediate Distributing Frame (IDF), to provide flexibility in distributing subscriber lines to the answering jacks at operator positions.[52] In 1892, W. S. Ford and E. A. Lenford invented a

[52] The Hibbard patent, U.S. No. 453,863, was filed February 28, 1891, and issued June 9, 1891. The Barton application was made March 8, 1890, and U.S. Patent No. 431,962 was issued July 8, 1890.

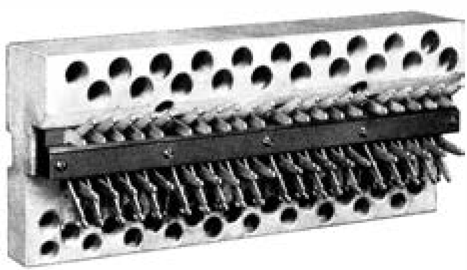
Fig. 6-36. Various types of terminal strips. (a) No. 183 terminal strip. (b) Strip with a large number of terminals. (c) No. 65 terminal strip; staggering of the punchings permits close spacing.



(a)



(b)



(c)

framework made up of metal rods and pipes to implement the ideas of Hibbard and Barton. The original proposals have been subject to refinement over the years but the basic ideas remained unchanged for a long time and are still widely used. A typical distributing frame consists of a long open structure several feet wide and extending over nearly the entire height of the office. A portion of such a frame is shown in Fig. 6-38. On one side of the frame, terminal strips of the type shown in Fig. 6-36 are mounted in horizontal rows spaced about a foot apart. On the other side, similar strips are mounted in a series of vertical rows. Horizontal racks run through the center of the frame to support connecting wires. Incoming cables are wired permanently to one side of the frame and outgoing to the other side. Jumper wires, which run vertically and horizontally through the frame as required, are used to make the desired connections between incoming and outgoing cables. If change in the interconnection is desired, it is accomplished by rewiring the jumpers.

The use of distributing frames can be clarified by considering a typical application to the manual switchboard used in the early part of the twentieth century (a somewhat simplified schematic of the entire arrangement is shown in Fig. 6-39). The outside-plant (subscriber) lines commonly were brought into the office by underground cables which terminated in a vault in the basement, as shown in Fig. 6-40. Here the paper-insulated cables from outside were spliced to the textile-insulated cables used inside the office. The latter were terminated on connector strips on the vertical side of the MDF.[53] The horizontal side was permanently connected to switchboard multiple jacks.[54] Jumper wires were used to connect the two sides, thus making possible the connection of any outside pair to any line number. The multiple (including the sleeve connection) was also wired to the horizontal side of the IDF and then jumpered to the vertical side of the same frame where the terminals were permanently connected to the answering jacks and associated relay equipment. Thus, the IDF provided a means for distributing line numbers to the answering-jack locations so as to provide each operator with a roughly equal traffic load.

3.4.7 Other Apparatus Components

The preceding sections have described briefly some of the components used in quantity in switching systems, particularly those of the manual

[53] Line protection equipment, described in Section 5.3.3 of Chapter 4, was mounted on the vertical side where it was permanently wired to the cable pairs.

[54] Current practice is described. Early practice was the reverse, i.e., with line wires terminated on the horizontal side. Since protectors were on the vertical side, the early practice provided some economy in protection, because there were fewer switchboard lines than cable pairs, but left unused cable pairs unprotected.

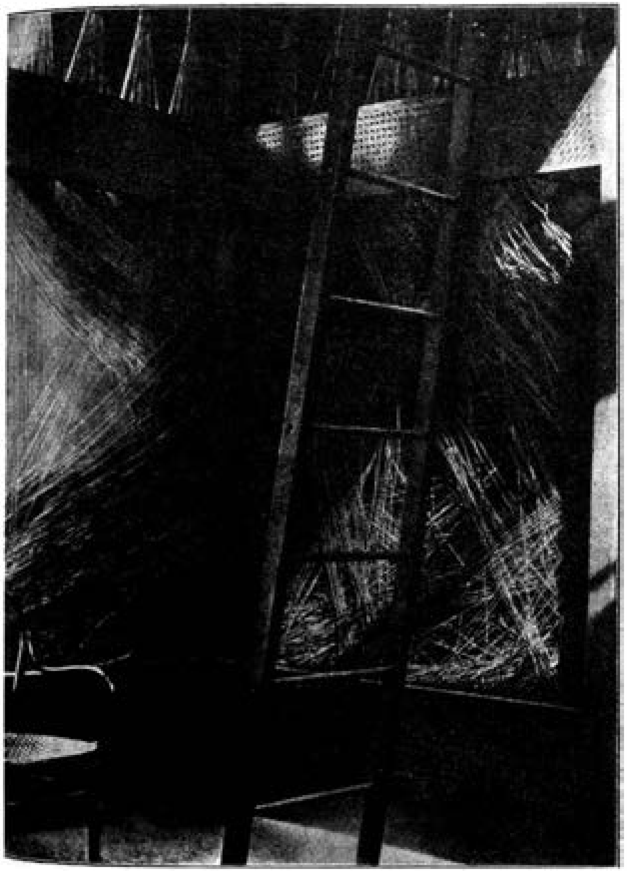


Fig. 6-37. Early distributing frame.

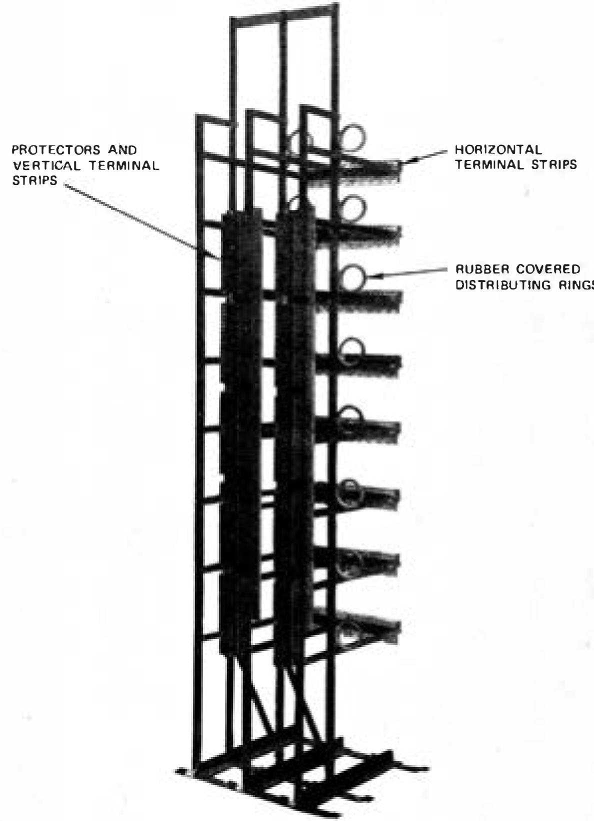


Fig. 6-38. Ford and Lenford type distributing frame of early 1900s.

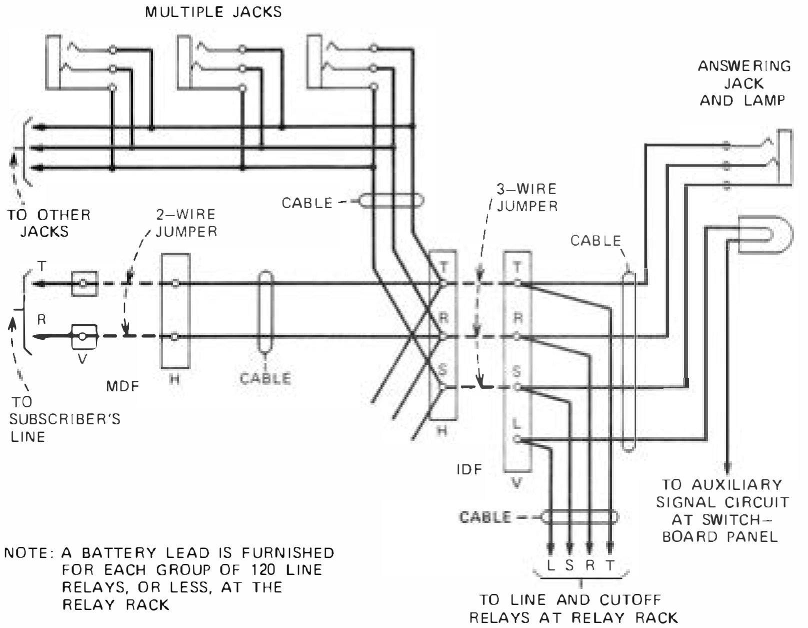


Fig. 6-39. Connections through distributing frames. (Redrawn from Miller 1905, p. 321)

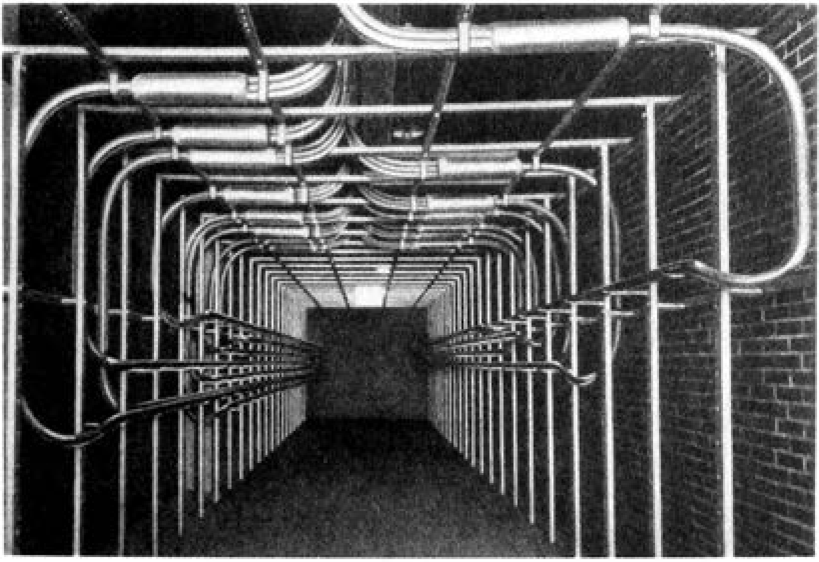


Fig. 6-40. Cable vault. (Gherardi and Jewett 1930, p. 20)

type. In addition, there are many other components entering into the complete telephone system, some of which are located on the customer's premises, others in the outside plant, and still others in the central offices. Many of these are described in other chapters where such a description seems to best fit into the narrative structure we have adopted.[55] Some items such as resistors, capacitors, transformers, and so forth have found wide application and, so far as space permits, their development will be covered in Chapter 8 (Materials and Components).

3.5 Standardization

As we have noted, the first ten years of telephony saw the use of a great variety of switching systems. This was undoubtedly very conducive to the development of original approaches to switching in a new and rapidly developing art. However, it also was accompanied by serious disadvantages when establishing an interconnected but geographically widespread communication network. In recognition of these problems, the National Telephone Exchange Association began, in 1880, to hold annual conventions for comparing the experience of the many telephone organizations involved and exchanging information on their systems and practices. Thus there began at an early date the cooperation among telephone workers of the United States that still continues.[56]

A significant step towards selecting the most desirable systems and standardizing operations was taken when the American Bell Telephone Company called a nationwide Switchboard Conference for December 1897. Two earlier and highly productive Cable Conferences (mentioned in Section 2.3.1 of Chapter 4) gave impetus to this meeting. Naturally, the Switchboard Conference did not attempt to establish a standardized switching system but it did resolve a number of questions and clarified many others. The technique used was to establish the basic requirements of a switching system and then to examine the many extant systems and evaluate the manner in which they met these requirements with respect to efficiency, economy, and permanency. Some of the Conference results were:

(*i*) Recognition of the need for educating the public in the use of the telephone and suggestion of techniques for doing so.

[55] For example, components making up the telephone station set are described in Chapter 3 even though some of these such as the magneto and ringer are closely associated with the signaling. Protection devices, whether located on the customer's premises or in the central office, are covered in Chapter 4 since they are so closely related to the wire plant. Likewise, much information on the electron tube is given in Chapters 4 and 5 in order to clarify the development of transmission systems.

[56] J. E. Kingsbury, writing in England in 1915, points out that there was no cooperation among the companies and administrations operating in Europe. The implication is clear that he believed the cooperative approach in the United States greatly advanced the art.

(*ii*) Determination of the amount of effort which could reasonably be expected from the customer in establishing a call. (It was agreed that switching and signaling should be a responsibility of the telephone company and that the work required of the customer be kept to a minimum.)

(*iii*) Agreement on the need for uniformity of operating procedures carried out in a disciplined manner with courtesy. (The Conference favored women as operators and the establishment of training schools.)

(*iv*) Decision that the multiple board be given first preference but that other types be considered acceptable.

(*v*) Proposals that all new boards be adaptable to metallic circuits and that crosstalk in the new boards be reduced by improved circuitry.

(*vi*) Suggestion of separate answering jacks and distributing lines to give uniform operator load.

(*vii*) Recognition of the disadvantages of party lines. (The Conference suggested lower rates for such service but failed to find a feasible economic substitute.)

It is rather clear that this Conference, which had no authority for implementing its suggestions, nevertheless set the pattern which was to be followed for many years to come. It is an interesting example of how problems of standardization can be handled in a flexible manner through the cooperative effort of people motivated to find a preferred solution based on a technical approach.

A number of subsequent Switchboard Conferences were held during the 1890s and helped to guide the evolution of the common-battery system and other improvements mentioned in previous sections.

3.6 Traffic Engineering

Our discussion up to this point has been largely concerned with the development of switching "hardware" and with practical ways for combining and operating it to meet the needs of a universal telephone system. Much of this work was done by a number of highly ingenious practitioners of a completely new art relying heavily on invention and intuition. But, as in other aspects of telephone communication, science began to supplement art in the early years of the twentieth century. In switching, science and its handmaiden mathematics first made their appearance in the area which can be designated broadly as "traffic engineering."

To most readers the term "traffic" will bring to mind the interminable flow of automobiles on the highway. To the telephone engineer the word has a somewhat similar connotation; but in this technical usage it refers to the flow of telephone messages over the many paths they traverse between users.

Basically, the problem of telephone traffic engineering involves the question of how many calls are to be expected over a given pathway at a given time and what quantity of telephone facilities is required to handle these calls in a reasonable manner. In this definition we use the term pathway very broadly to mean a group of trunks between two points, an outside-plant route which may in the aggregate handle many trunk groups, a central-office switchboard, or any other part of the telephone highway which, if inadequately equipped, could act as a bottleneck to slow or interrupt the orderly flow of telephone calls.

The problem arises because telephone calls are originated at the whim of the user and there is no way to predict when a particular caller "A" will want to talk to another user "B" and precisely how long this call will last. Fortunately, typical calling patterns for a group of users can be forecast in a general way even though the precise prediction of an individual's performance is not possible. Intuitively, we feel sure that there will be fewer calls in some parts of the day than in others and by supplementing our intuition with traffic counts we can determine typically the number of calls through any pathway at any given time of day and "on the average" how long calls will last.

If all calling followed precisely the typical pattern, it would be rather simple to determine how large to make our pathways, i.e., how many trunks to provide between any two switching points, how many switchboard positions to install in our office, and how many operators to have on duty during each hour of the day. Unfortunately, things are not so easy because the element of chance (or probability) enters to a large extent. Not all calls, for example, last the same length of time. Some are very short, others last a matter of minutes, and a few will be prolonged for an hour or more. Similarly, the number of calls made in a given hour is not the same from day to day. It will depend on the individual needs of customers, the weather, the state of business, the occurrence of unusual events, etc. The best we can do is to obtain by observation of "statistical" description of the flow of traffic; that is, not only a description of the average number of calls to be expected and the average duration but also some measure of the variability of the calls which flow over our pathways. With this information and the use of the mathematics of probability and statistics, we can obtain information which provides a basis for deciding (with the help of good judgment) the amount of equipment required to provide a reasonable grade of service. The art of making these decisions is the province of the traffic engineer and is based on a knowledge of the pattern of traffic flowing over the pathway in question and on the application of a traffic mathematics that is much the same as the mathematics of the statistician.

Statistical mathematics is widely used by life insurance actuaries, agronomists, gamblers, and many other people dealing with matters governed largely by chance. Its use in telephone communication was an early application in the business field and represents one with enormous cost benefits since an error in engineering telephone equipment can result in enormous excess costs if supplied too generously or in great dissatisfaction to the user if too sparsely furnished.[57]

We shall shortly outline the manner in which traffic mathematics evolved; but before doing so it is important to point out the special characteristics of the answer provided by the statistician. It is highly precise in the sense that it is mathematically correct, but it is always presented with qualifications. Since we are dealing with matters influenced by chance, the statistical answer can never tell us the precise number of trunks (or other pieces of equipment) required to handle a given quantity (or pattern) of traffic. It can only tell us the chance, or probability, that a given number will be enough to meet our objective. As an example, a statistical answer might tell us that, for a given pattern of traffic, a specific number of trunks will in the long run be adequate to handle 99 out of 100 calls without delay but the implication is inherent that, on the average, one call will be delayed. There is a further qualification of the answer that should be noted. In our example, one in one-hundred calls will encounter delay *on the average*. This does not mean that one *will* be delayed in every successive group of one-hundred calls but that this is the average rate of occurrence in a sample involving many hundreds of calls. Any block of one hundred picked at random may have no delay or well over the single delay which occurs on the average. Refinements of statistical methods provide a means for computing the probability of such "atypical" performance but always the answer is given not in terms of the specific number of occurrences but as the average rate at which an event will occur in the long run.

By the use of these methods, the traffic engineer can determine the grade of service provided by various amounts of a critical facility,[58] but there is nothing in the statistical method which tells what grade of service is adequate. It is for this reason that we have included the exercise of judgment among the functions of the traffic engineer. He

[57] The use of the mathematics or probability and statistics in the Bell System was not limited to the field of traffic engineering. The power of this tool in solving industrial and engineering problems was widely recognized at an early date. In 1928, Thornton Fry published his book, *Probability and Its Engineering Uses*, which remained the classic text in its field for many years. In the early twenties, W. A. Shewhart used statistical methods for setting up sampling techniques for the inspection of manufactured product which served as the genesis of Quality Assurance. (This subject is discussed at greater length in Chapter 9.)

[58] As an example, a given traffic pattern which requires ten trunks to give a one in a hundred chance of delay will require about 20 percent more trunks for a chance of one in a thousand and 40 percent more for one in ten thousand.

must choose a performance standard that provides a suitable compromise between the high cost of trunks, which goes with excessively high standards, and the high chance of busy circuits, which goes with low. The former will inhibit usage because of its impact on cost of service and the latter because of dissatisfaction and inadequacies of service. In the early days of telephony, when telephone circuits were very costly, delay was unavoidable and almost all very long-distance calls were handled on a delayed basis during the busy time of day. But as circuit cost has been reduced, good judgment has dictated a reduction in delay until today it is negligible under normal conditions.

3.6.1 Evolution of Traffic Mathematics

The first traffic problem was encountered in the design of Coy's New Haven switchboard in 1878. As will be recalled, his office had eight lines and calls from an incoming line were connected to the desired outgoing line by means of switches connected together by what later came to be known as transfer trunks. Theoretically, it would have been possible to carry out four calls simultaneously if each of the four calling lines desired connection to the appropriate one of the remaining four lines available in the office. Coy seems to have recognized the very small chance that this would occur and, on an intuitive basis, provided only two transfer trunks and sets of associated switches. This approach was satisfactory for the conditions existing, but as the offices increased in size, and the use of interoffice trunks began, traffic engineering problems increased and the need grew for deciding on the number of switchboard positions to install, the number of operators required at various times of the day, and the number of intraoffice and interoffice trunks required. In the early years of telephony, most of these decisions seem to have been made on intuitive or empirical bases but, since the methods used by various telephone engineers were usually considered to be proprietary information, there is little detailed material available on specific techniques used prior to the turn of the century.

Up to this time, much work was certainly done to determine the rate at which calls were made as a function of time of day, type of service, month of year, and so forth. The concept of the "busy hour" was adopted very early since observations soon showed that traffic, although highly variable, tended to follow a reasonably consistent pattern for any given class of call (e.g., business or residential) and such calls reached a peak for a short period during the day. Obviously, if satisfactory service was provided for the busiest hour of the day, there would be no problem during the remaining period except for unpredictable catastrophic events completely destroying the pattern.

In 1898, G. T. Blood of AT&TCo, who for a number of years had been engaged in traffic studies, made the first recorded attempt to apply mathematical methods to the solution of traffic problems. He noted a close agreement between observations on the distribution of busy trunks in a group and the terms of an algebraic expression which mathematicians had shown to be related to the probability of occurrence of certain random events. No records of Blood's studies are available and the extent of his work is not known, but in the following year M. C. Rorty, then in the Traffic Department of AT&TCo in Boston, became interested in this work and in October 1903 prepared an internal paper describing mathematical means for preparing curves that could be used for solving some of the common traffic problems. In his own words, " . . . an attempt has been made to calculate the distribution of calls over a short period, such as the busy hour, to calculate trunk efficiencies, etc., by means of the mathematical theory of probability." Rorty's paper was widely circulated in the Bell System and the substance was presented before a meeting of the Chicago Section of the AIEE in April 1905. This seems to be the first published description of the application of statistical methods to telephone traffic problems.

October 1903, when Rorty's paper was prepared, also marked another important advance in traffic engineering. At this time, AT&TCo issued the first comprehensive traffic engineering practice, aimed at outlining "Methods to be Followed in General in Preparing a Traffic Study on Which to Base the Plans for a Central Office Equipment." These notes included advice on measuring traffic characteristics and predicting traffic expansion as well as data on the periods to be used for initial and long-term engineering studies and some dozen curves for determining the size of equipment groups. These curves were probably empirical since the use of mathematical analysis was still very new; but as the work of Rorty and others progressed, the traffic practices were expanded and ultimately became a highly comprehensive, mathematically based treatise on traffic engineering.

Returning to Rorty's work, his mathematical derivation was based on a number of assumptions, namely, that a fixed number of calls were placed at random within the busy hour and that all calls were of the same length. These assumptions greatly simplified the problem. Since calls were all of the same length and occurred at random, it was only necessary to look at a segment of the busy hour equal to the length of a call. Furthermore, the probability of occurrence of any given call in this interval was equal to the call length expressed in hours, i.e., if the call length was 3 minutes, or $1/20$ hour, the probability of a particular call being made in this interval was one-twentieth. Since the number of calls offered to a group of equipment in the busy hour was assumed equal to the average, it became possible to compute, by using the binomial expansion, the probability of exceeding any number of calls during an interval. Or expressed differently, it was possible to compute for a given size of trunk group (or amount of equipment) the probability that there would be a delay or blocking of calls.[59]

Although his assumptions seem somewhat artificial, the agreement between computed blocking probabilities and those determined empirically was quite good, and as a result of Rorty's work there was increasing recognition of the possibility of predicting aggregate human performance in statistical terms even though individual performance was random.

The complexity of the computations required by Rorty's method, and the assumptions necessary, stimulated further search for suitable techniques for probability computations, and E. C. Molina[60] became the main contributor. His first contribution was to restate the problem in terms of the user's point of view, namely, "given a certain number of calls in the busy hour and a certain number of trunks, what is the probability that all trunks will be busy when a particular call is made?" He found that Rorty's use of the binomial could, with slight modification, be used to answer this question provided that a further assumption was made that calls which found all paths busy were delayed and ultimately were completed with their normal holding time reduced by the amount of the delay. The other assumptions made by Rorty about holding time and uniform loading within the busy hour were continued

[59] Only a very superficial discussion of the mathematics of traffic engineering is possible without getting into complexities beyond the scope of this book. A few readers may find our approach unsatisfactory and for these perhaps some further details of Rorty's mathematical approach should be given. In essence, he assumed that the probability of a specific call being initiated during a given period was p and the probability of a non-occurrence at this time must then be q since $p + q = 1$. If the number of calls offered during the busy hour is N, the probability of occurrence of precisely 0, 1, 2, etc., calls is obtained by substituting the values for p and q in the successive terms of the algebraic expansion of the binomial $(p + q)^N$.

[60] Molina was one of the extraordinary individuals who left a lasting mark on the evolution of telephony. He joined the Western Electric Company in 1898 at the age of 21 with no formal education beyond high school and proved to be a versatile and prolific inventor, making important contributions to the development of machine switching, some of which will be described later. Molina joined the Research Department of AT&TCo in 1901 and, during his long career in this company and with Bell Telephone Laboratories, contributed extensively to the development and application of traffic mathematics. This was made possible by a disciplined self-study of mathematics that led him to become a recognized expert in certain fields, particularly in the work of Laplace. In recognition of this work and the encouragement and example it offered to young people for self-study, he received, in 1952, the Honorary Degree of Doctor of Science from Newark College of Engineering, an institution at which he taught mathematics until 1964, some 20 years after his "retirement" from the Bell System.

since they appeared to have little effect on the result; but Molina noted that the normal approximation for the binomial summation (as used by Rorty) was not reliable when the probability of finding all trunks busy was less than one in ten. Since good service required much less chance of delay, Molina devised a new formula, easy to compute, that provided an excellent approximation for the delay probabilities of interest, namely those between 1 in 30 and 1 in 100. Computation of new traffic engineering tables began in earnest and early in this work it was observed that when the product of the number of callers and their probability of calling was constant, the number of trunks required was almost independent of the number of callers so long as this number was large. Based on this observation, Molina devised a new approximation which greatly simplified computation; traffic capacity tables based on this method have been in continuous use since 1908. The approximation was called the Molina formula until several years after his work when Molina discovered that it had been derived earlier by the French mathematician Poisson, whose name has since then been associated with it.

Traffic engineering also had been the subject of work in Europe, chiefly by A. K. Erlang in Denmark. Unfortunately, his work did not receive attention beyond his own country until its publication in English in *Journal of the Post Office Electrical Engineers* (London, 1918). This opened a completely different approach to traffic theory, an approach based on an exponentially distributed holding time instead of the constant holding time used previously. While this had little effect on most of the traffic problems of importance during the first 50 years of telephony, it was much more realistic and permitted analytic approaches to problems otherwise not readily solvable. It provided formulas for treating different types of behavior when the customer found all trunks busy. For example, the capacity of trunk groups could be computed for the case where the customer abandoned his call and also for that in which he waited for an available trunk. In the latter situation, either the shortened talking time, assumed by Molina, or normal holding time could be assumed. Erlang's work also was the forerunner of much of the "queuing theory" work which has found wide usage in recent times.

3.6.2 Application of Traffic Mathematics

The reasons for the development of traffic mathematics have already been covered, namely, the determination of the amounts of equipment required to furnish a good grade of service at a reasonable cost. Commonly, the basis for making such decisions is a series of curves, such as shown in Fig. 6-41a, which indicate the average number of calls that can be handled by various numbers of trunks (or pieces of switching equipment) for several probabilities (p) of blocking, i.e., finding all trunks busy.[61]

Perhaps Fig. 6-41b gives a way of presenting the same data which is more impressive to the average reader. This shows, for two blocking probabilities, how trunk efficiency[62] depends on groups size. It clearly brings out the benefits resulting from the use of large trunk groups, a matter which has been of great significance in designing switching systems, particularly the automatic systems described in Section IV.

The original traffic engineering problems, exemplified by Fig. 6-41, have expanded to those of much more complexity involving the entire communication network. Full treatment of such problems is beyond the scope of this history but some of the potentialities can be indicated. As an example, it is apparent that in large cities the concentration of a number of central offices in the same building not only leads to economies in the craft forces but also permits the use of larger trunk groups with accordingly improved efficiency in traffic load. However, this may lead to somewhat longer and more costly loops and trunks. Traffic mathematics is obviously one of the components of the studies which determine the most economical arrangement. Similarly, with tandem switching, choices are available between direct trunking, which often provides low-cost trunks but in groups so small as to be inefficient, and more costly tandem trunks, which can be used more efficiently because the groups are large.

3.7 Summary

By the early 1900s, manual switching in the exchange plant had become a well-developed art based on principles and components devised to particularly fit the needs of telephony. Beginning with large and crude devices and arrangements involving complex operations, both on the part of the telephone user and the switching operator, switchboards had developed to the point where manual effort by the user and oper-

[61] These curves, which have so much significance for the traffic engineer, are not too easy to interpret by the uninitiated. Perhaps an example will help. Let us assume that we have a group of 45 trunks and wish to know what average traffic load can be carried during a given period (e.g., the busy hour) without blocking, on the average, more than 1 call out of every 100 submitted. The upper curve indicates that we could handle 30 calls on the average. This does not mean that only 30 trunks would be in use at all times. On the contrary, this is an average for a varying traffic pattern and if during the busy hour we examined the trunk group at regular intervals, we would find usage varying. Sometimes there would be less than 30 busy trunks and sometimes more. Very occasionally all 45 would be in use and a submitted call would be blocked (or delayed). However, the average of all of these examinations would show 30 trunks in use and the blocking situation would be found only once in 100 times.

[62] Efficiency can be defined as the average proportion of time a trunk is in use during a specified period (such as the busy hour) or, expressed differently, it represents the average proportion of the trunk group in use during the period.

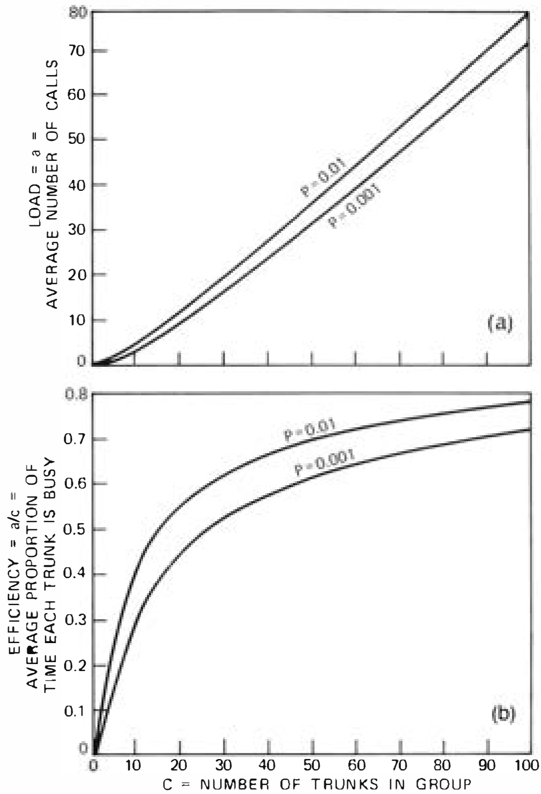


Fig. 6-41. (a) Load carried by a trunk group without exceeding the indicated proportion (p) of blocked calls. (b) Trunk efficiency for indicated proportion of blocking.

ator had become minimal. The size of components had been greatly reduced so that as many as 10,000 lines could be handled as one central-office unit and means had been devised for economically switching calls in exchange areas of roughly any size from a small village to a large metropolitan area.

Intimately associated with the evolution of switching systems was the concurrent development of signaling arrangements for control of the interconnecting network. While much of the control information could be conveyed by voice, it was soon found that supplementary electrical signals could be used with much saving in operator and user effort. Typically, such signals were used to alert operators and customers and for conveying information on progress and status of the call. In the early days of telephony, this type of signaling was done by means of

16–20-hertz alternating current generated either by a hand-cranked "magneto" in the station set or machine-driven generators at the central office. Beginning in the late nineties, direct current from a common battery in the central office began to be used both for signaling and for supplying the talking battery that previously had been furnished by primary batteries on the customer's premises. These common-battery systems, by eliminating the need for magneto and battery supply at the customer's station, simplified both station apparatus and the user's effort, greatly stimulating the growth of telephone usage. Alternating-current signals were not completely eliminated, however. The low-frequency ringing system was retained since it was highly effective and simple to use once the central office provided the power source. Certain audible, alternating-current signals also were used to indicate progress of the call to the customer (i.e., ringback, busy, etc.)

While many changes were introduced to simplify procedures, the basic interconnecting operation on most calls was performed manually by the central-office operator. There was a natural desire to find a way to automate the interconnecting function and as early as about 1879 technical means for doing this had been proposed and ten years later active development of automatic (machine) switching started. This work, described in our next section, proceeded simultaneously with the development of manual systems, which has just been described. Despite the large amount of effort expended on automatic systems, manual systems dominated the field during the span of this history. By 1925, only about 12 percent of the exchange telephone lines in the Bell System operated on an automatic basis. But, as we shall see, developments of lasting importance had taken place which were already eroding the dominance of manual systems.

Another development during these years was the use of mathematical methods for engineering plant to provide the most economical means for handling telephone traffic. These traffic engineering techniques not only minimized the amount of plant expenditure required but also had an important influence on the evolution of automatic switching.

IV. AUTOMATIC EXCHANGE SWITCHING

4.1 Origin

The first proposals for "automatic" switchboards[63] came only a year or two after the first commercial manual boards were put into service. Probably the earliest was a patent application by M. D. and T. A.

[63] The systems covered in this section originally were referred to as "automatic switching." Later the terms dial switching, mechanical switching, and machine switching were used. We have used the original term generally throughout this section, but the synonomous terms are used to some extent in the later part where they are more in keeping with current usage.

Connolly and Thomas McTighe.[64] This system, based on rotating a connecting switch under control of a series of dc impulses sent out by the subscriber, was very similar to a system for sending telegraph signals proposed by Froment (in France) almost 30 years before. Whether or not these two inventions were independently conceived is of no moment since neither proposal had much application.

Numerous other early inventors were active in this field, both in the United States and abroad, but it is generally considered that the evolution of the first widely used automatic systems stems from the proposals of A. B. Strowger.[65] This, too, was a system in which the lines were connected at the central office by means of a switch directly controlled by pulses sent out by the subscriber. While the pulsing mechanism originally proposed was both crude and impractical (see Section 4.2.2), the switch was unique in that it employed two directions of motion and with further development provided a key element in one of the first practical automatic systems.

A system based on the Strowger patent was installed at La Porte, Indiana, in November 1892. This installation is commonly credited with being the world's first commercial automatic central office. It was, however, a very crude affair that demonstrated a concept but was far from being a prototype for a practical system.[66]

Strowger, an undertaker from Kansas City, Missouri, did not have the technical background for developing his original concept into a practical system, but this was accomplished by ingenious partners with whom he became associated. In 1892, A. E. Keith joined Strowger and about two years later the Erickson brothers, John and Charles J., became associated with the organization and they, with Keith, made up a technical group largely responsible for converting the Strowger concept into a working system. Frank A. Lundquist also was an important member of this group for a period of about two years beginning in 1894. Strowger withdrew from the telephone business in a half-dozen years and died in 1902, but the work he started was continued by his associates in the Automatic Electric Company, which had been formed to develop and manufacture automatic systems.[67] But more about this later.

[64] M. D. and T. A. Connolly and T. McTighe; U.S. Patent No. 222,458; filed September 10, 1879; issued December 9, 1879.

[65] A. B. Strowger; U.S. Patent No. 447,918; filed March 12, 1889; issued May 10, 1891.

[66] While this system utilized a number of the proposals in the Strowger patent, they mostly involved ideas that were soon abandoned as impractical (the most important Strowger invention, the switch with two directions of motion, was not used at first in La Porte). The original La Porte installation was largely replaced by another experimental arrangement, using the two-direction switch, in the fall of 1894. The following year, the switches were again replaced, this time with a design closely resembling the type which was to be the basis for the commercial Strowger system.

[67] The first company formed to exploit the Strowger patent was incorporated October 30, 1891, as the "Strowger Automatic Telephone Exchange." The Automatic Electric Company was organized in 1901 and acquired the U.S. rights to manufacture and sell Strowger equipment. The original Strowger Automatic Telephone Exchange remained in existence as a patent-holding company until 1908 when its stock was acquired by Automatic Electric.

The Bell System became interested in automatic switching as early as 1884, some four years before the Strowger invention. This early Bell work, and that of the next 10 to 20 years, was aimed largely at supplying service to small offices where there was not a full-time load for even one operator. At the time, this was a very logical decision since during the later years of the nineteenth century many great improvements were being made in the large manual systems, which reduced both operator and customer effort, but the very small offices continued to be a serious problem because the high operating costs per station were not affected by these improvements, being inherent in the size of the office.[68]

The first automatic system tried by Bell was the "Village System" invented by Gilliland in 1884 and on which he received some half-dozen patents. The first installation was opened in Leicester, Massachusetts, on September 21, 1885, and some 50 systems were ultimately put in service. This system was a combination of the station-switching and party-line concepts described in Section II. No central office was employed, the various lines in the exchange all being looped through each customer's station, which included keys for connecting his telephone to the desired lines. Means were also provided for connection to a nearby toll operator. The field of use for such a system was severely limited since outside plant costs rose rapidly as the number of customers and the distance between them increased. The intended field of use, 20 to 40 customers on four lines, was attractive in 1885 but the popularity of telephony increased rapidly and most of the "Village System" installations had been replaced by higher-capacity common-battery manual switchboards by 1902.[69]

During the early 1900s, further attempts were made by the Bell System to solve the very small office problem by means of automation. Because of the outside plant problems, station switching was abandoned and central-office switching, using single-level selectors similar to the Connolly–McTighe proposal, was introduced. These selectors were controlled by a dial, having one hole per line, on each customer's station. An arm was set to the desired line position and upon release

[68] In cost studies made during the early 1900s the wages for an operator were figured at $300 per year for 12-hour service, six days a week. Full 24-hour, seven-day service would require three operators at $800 per year. Thus, the traffic costs per line for a 25-line office would be $1 per month for 12 hours limited service or $3 for full 24-hour coverage. These traffic costs may not seem extraordinary by present standards but were prohibitively high at the time. Service in small towns had limited value in the early days of telephony because of the small number of phones which could be reached by a customer. As a consequence, the charges for phone service had to be kept low and might not greatly exceed the traffic costs if full-time service (potentially available with automatic switching) were provided.

[69] While most of the work on the Village System was carried out before the Strowger installation at La Porte, its effect on the development of automatic telephony was small since the basic principle of switching at the station received little further use except for systems using key telephone sets for local intercommunication and primitive PBX systems (see Section 6.1 in this chapter and also Section 7.2, Chapter 3).

sent out pulses to operate the selectors. A 100-line system of this type was tested in Queens, New York, in 1902 and, two years later, systems for 20 and 100 lines were made available. The size of these systems was limited by the use of single-level selectors and station-set dials with one position per line (station sets with dials for these systems are shown in Fig. 6-42). In spite of these limitations, the need for inexpensive small systems was so great that some 40 were sold within a few years after they were placed on sale. However, as a result of technical problems and growth beyond the system capacity, ultimately these installations also were changed to manual operation which, as developed by Bell in the early 1900s, gave better and more economical service for all exchanges except the very smallest.

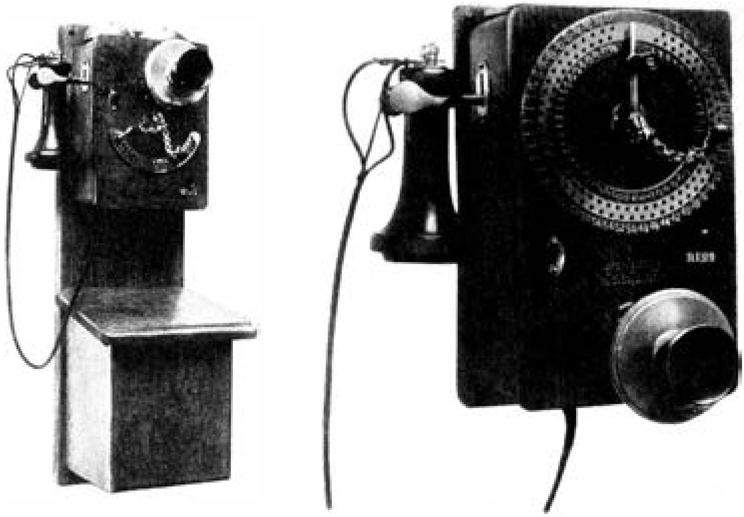
In 1902, the American Telephone and Telegraph Company authorized Western Electric to develop a 10,000-line system. Work continued through 1903 but produced no evidence that such a system would provide material savings if designed to meet the same service standards as current manual systems. In 1903, Bell also purchased patent rights to the Lorimer system, the first of the automatic systems to use motor-drive instead of magnetic stepping mechanisms. Several of these systems had been built by Lorimer but they were not found suitable for Bell use. Probably because of these experiences, Hammond V. Hayes, toward the end of 1904, reported that the future of automatic exchanges in the Bell System, particularly in sizes over 100 lines, was uncertain.

During the early 1900s, the Automatic Electric Company had been pressing development of the Strowger system (as covered in Section 4.2) and had installed systems with capacities of as many as 6,000 lines in Grand Rapids, Michigan; Dayton, Ohio; and several other cities. Kempster B. Miller, writing in 1905 about these systems, said:[70]

> Summing up, therefore, the statements already made, the automatic system is not only a possibility, but is actually here. With the interjection of human intelligence to supplement it in performing certain functions, it seems to be as flexible as the manual. Party line, common battery and measured service working, while not yet achieved commercially, so far as I am aware, seem to be well within the grasp of those who are doing the development work. The public seems to like it, and we do not know whether it is too complex or not.

Thus, we find a rather anomolous situation some 15 years after the Strowger automatic concept had been introduced. On the one hand, some half-dozen cities had installed automatic systems with capacities up to about 6,000 lines. At the other extreme, the Bell System had built 50–100 small systems and found them not satisfactory, and its experience with larger systems had led Bell to believe that the outlook

[70] See 4th Edition of *American Telephone Practice.*



(a) (b)

Fig. 6-42. Station sets for Western Electric Automatic System, circa 1904: (a) set for 20-line system; (b) set for 100-line system.

for these systems also was uncertain. In between these views were those of an independent telephone expert, Kempster Miller, who seems to have had considerable enthusiasm for the Strowger type of system, not as it then existed, but as it might develop in the near future.

With our advantage of looking backwards we can speculate as to the reasons which might have led to this situation. Probably there were a number of factors responsible for these divergent views including service standards, technology, economics, the available market, and unknowns concerning customer reaction.

By 1904, manual-switching technology in the Bell System was highly developed. Common-battery boards using indicating lamps had been in use for over ten years and were widespread throughout the system. In large cities, 10,000-line offices were common and trunking arrangements had been developed (using separate A and B boards) to interconnect essentially any number of such offices that might be required. Many of the complex manual operations originally required of both subscriber and operator had been replaced by automation. The subscriber's operation had been reduced to a minimum; he had merely to lift his telephone from a hook to signal his desire for communication and then he was able to pass his orders verbally to the operator. The operator carried out these orders by making connections through easily operated and conveniently located plugs, jacks, and keys. She also was available to assist the subscriber in any way desired, and provided supplementary operations going beyond simple local connections.[71]

Automatic telephony had received a considerable amount of attention over a period of 20 years with indifferent results. It was not a great success in the very small exchange, the field where operating savings first appeared so attractive, largely because these exchanges grew rapidly to a size where operators were used more efficiently with consequent reduction in operating cost per line. Strowger automatic systems had been developed for offices with several thousand lines but from a service standpoint they were a definite step backward. Much of the burden of making connections was shifted from the operator to the subscriber. He not only had to operate a dial mechanism to indicate the party desired, but he also had to initiate ringing by operating a button on the station set. Local batteries were used at the station, again increasing the bulk of station equipment and reintroducing the replacement problem. Finally, some forms of service were unavailable with automatic systems. Obviously, as of 1904, there was a real question as to whether or not these retrogressions would not cause an unfavorable subscriber reaction and thus inhibit the rapid growth then being experienced. As time went on, the requirements for local batteries and customer ringing could be eliminated but improvements in manual boards to reduce operator effort also were to be expected. Therefore, it was not completely certain that automatic service could be expected to be superior, from the subscriber's viewpoint, to good-quality manual service since, at best, an automatic system would require the extra operation of dialing on the part of the customer. In all Bell System work a basic requirement was that service provided by an automatic system should not be inferior to that of a manual system of current design. Until wide use of the dial proved its acceptability to the user, many telephone engineers were reluctant to place this burden on the user. On the other hand, ardent proponents of automation believed that this disadvantage was more than offset by privacy, speedier service, and operator savings.

The Bell view in the early years was that the main justification for automatic service would have to be economic.[72] The reduction in

[71] The value of an operator in maintaining good service should not be underrated. The answering operator supervised the connection and could, in response to a switchhook flash, take care of any difficulties encountered on the circuit. She also talked over the calling loop and after the introduction of straightforward trunking, she also talked over the trunk. Thus she was in a position to rapidly detect poor transmission and often detected other problems as they arose. With automatic systems, the loss of these operator functions placed additional burdens on the maintenance staff to detect and correct faults promptly before serious situations developed.

[72] In later years the availability of a suitable labor supply became an important factor.

operator expense and secondary costs (rest rooms, restaurants, etc.) would have to compensate for the higher cost of automatic equipment and the added burden of dialing. The extent to which this would become possible was far from clear. For example, in a small city with modern manual equipment, the cost of growth by adding more manual positions was low as compared to automation, which would require a mix of two systems (with impractical complexities) or a replacement of existing manual equipment. A large multioffice city, where growth was obtained by adding completely new offices, obviously favored automatic service, providing manual-automatic interface problems could be solved.

Another economic factor was the extent to which operators could be eliminated. No one in the early twentieth century foresaw the extensive automation of toll calls and other services that we have today. It seemed obvious that many operators would always be required and consequently the secondary savings could be less than one might expect. Thus, the operator-cost savings in a particular situation were dependent on the services provided, e.g., toll operating, information, etc.[73]

With these facts in mind we can see shaping up two different markets for automatic switching equipment. One is the rural area where new service is being developed, or the small town where existing manual equipment requires replacement, particularly one in which there is little need for toll or other supplementary operating. The second market is the large multioffice city where the growth is met by establishing new central-office units and where operating personnel with suitable skills will probably be in short supply in the future. This latter market would, of course, only materialize if the problems that arose in interconnecting the two types of systems in the large cities could be solved (see Section 4.3.2.2).

During the period between 1904 and about 1914, while the Strowger system was evolving from the rudimentary, local-battery arrangements of Grand Rapids into a common-battery system with the features available in modern manual offices, the number of independent telephone

[73] It is interesting to note the extent to which operator requirements have changed over the years. In 1902, as common-battery switchboards were being introduced widely in Bell offices, the Bell System employed roughly 30,000 operators or about 22 per 1,000 telephones. At the time of World War I, when only about 1 percent of Bell telephones were automatic, the System employed about 100,000 operators or roughly 15 per 1,000 telephones. This reduction in operators per telephones was very largely the result of the many automatic features and improvements in manual boards which were then commonplace throughout the system. Thirty years later at the time of World War II, the number of dial phones had increased to 45 percent and the number of operators amounted to about 6 or 7 per 1,000 telephones but totaled about 140,000. By 1970, when the Bell System was on a full dial basis with automation covering toll calling, timing, message accounting, and other services, there was still an operating force of about 166,000 or 1.7 per 1,000 telephones. This is a rather interesting example of the manner in which the introduction of automation can be keyed to growth so as to minimize adverse effects on the labor force.

companies in the United States was growing rapidly as a result of the expiration of some of the earlier Bell patents.[74] Certain of these companies were of the so-called mutual type, owned largely by the subscribers, and were formed to serve some of the sparsely settled areas not covered by Bell. The toll lines connecting these companies, and very often the toll boards also, were supplied by Bell; hence automatic exchange service could greatly reduce the need for operators in the mutual company. In addition, the fact that the subscribers owned the companies often made them tolerant of service standards that might otherwise have been less acceptable.

The Bell System, on the other hand, had a large proportion of its customers in the major cities where Bell companies were experiencing extraordinary growth requiring the frequent addition of central offices.[75] Many of the smaller and medium-size cities had relatively modern manual offices where growth could be met quite readily by adding similar manual positions.[76] Hence, the Bell interest in automatic switching was directed more and more towards solving the problems of the larger cities.

But let us forego our speculations and return to the chronological narrative. By 1910 the Bell position was beginning to clarify and was expounded well by J. J. Carty, the AT&TCo Chief Engineer.[77] His paper has sometimes been cited as evidence that the Bell System was opposed to automation. However, a careful reading will show that Carty was opposed not to automation *per se*, but to any system that would not provide service equivalent to existing systems, would not offer the potential for meeting the rapid growth expected in large cities, and would not achieve the broad objective of rapid, comprehensive service

[74] The basic Bell patents expired in 1893–94. Prior to 1892, about 60 independent companies had been formed, mostly of small size. Fifteen years later (1907) there were over 5,000 independent companies owning about 45 percent of U.S. telephones. Most of these companies were small, the average number of stations being under 5,000.

[75] New York City presented an extreme example of the growth problems that the Bell System was facing. J. J. Carty in his 1910 paper, discussed in the next paragraph, gave the following data for the area within the greater New York municipal limits:

Year	Population (millions)	Telephone Stations (thousands)	Central Offices
1900	3.4	51	43
1910	4.8	376	52
1930 (projected)	8.8	2,142	109

[76] Wherever possible, new Bell equipment was designed so that it was compatible with older designs and could be added for growth with a minimum of office rearrangement.

[77] In a paper presented at the International Conference of European Telephone and Telegraph Administrations, held in Paris September 4–11, 1910, Carty was careful to point out that in accord with the paper's title, "Telephone Service in America," he was confining his comments on automation to the conditions applicable in America.

covering the country. The basic argument presented in the paper was that all modern systems (dial or manual) were really semiautomatic and, as far as could be foreseen, some manual auxiliary operation or supervision of equipment would continue to be needed for a long time to come. To Carty, the basic question was not manual versus automatic, but which semiautomatic system should be used, what functions should be automated, and who should carry out the manual operations, the traffic operator or the customer. It was his view, based on observation of the existing technical situation, that the fully automatic system then available required an extensive maintenance force, not only to keep the equipment running but to detect failures and generally monitor service, a duty carried out with little effort by the A operator as a normal part of her job. This led him to conclude that full automation would reduce the manual effort less than many people expected since much of it would be transferred to the customer and the maintenance forces. This, together with other limitations, led him to believe that the automation of the B board was the first logical step (since such operators performed an almost solely mechanical job). A further step would be automatic distribution of incoming calls to A operators, thus eliminating the necessity for manual connection to answering jacks and the waiting time between calls. The basic concept was that the A operator, after being connected automatically to a calling subscriber, would obtain the necessary call information by voice, "dial" the desired number, and continue to supervise calls as required to assure satisfactory completion and to carry out other duties which automation could not accomplish (timing of AB toll calls for example). One of the features of operator dialing would be the use of a keyboard instead of a rotary dial, which would be faster and also could be checked before sending out pulses. Such a "semimechanical system" was receiving the major Bell development effort at the time and was subjected to large-scale field trial in the 1913–1917 period.[78]

These trials and the development work preceding and accompanying them were so successful that a decision was reached in 1917 to develop mechanical systems for large, multioffice cities. Because of problems associated with manual-dial interconnection in such cities, the initial plan called for semimechanical operation in large metropolitan areas with customer dialing elsewhere. However, in 1918 the metropolitan area problems had been solved (as described in Section 4.3.2.2) and it was decided to use full mechanical systems, with customer dialing, wherever automation proved economical. The rapid rise in operator

[78] About 10,000 lines, in two Newark, New Jersey, offices, were put in commercial service in 1913 to provide a field trial of operator "dialing." Similar large-scale trials of automatic distribution of calls to operators were conducted in Newark and Wilmington, Delaware, in 1917.

wages at this time greatly emphasized the benefits of automation and expanded the potential field of use. Unfortunately, the onset of World War I (which diverted much Bell manpower) delayed this work and the first fully mechanical office was not put into service until December 1921 (in Omaha, Nebraska) and the first fully automatic prototype of the metropolitan system (the Pennsylvania exchange in New York City) was not cut over until October 1922.

In concluding this summary of the evolution of mechanical switching in the Bell System, two excerpts from AT&TCo Annual Reports will be of interest. In 1913, it was stated:

> We have designed, and manufactured, and installed all kinds of switchboards—automatic, semiautomatic, and manual—and we have exhaustively studied the practical workings of every type of switchboard in use. . . . As yet it has not been demonstrated that the automatic system would give as good and dependable service as we now render to the public when used in connection with the extensive and comprehensive suburban and inter-urban telephone system of Bell.

By 1919, the doubts which had existed for so many years had been resolved and the Annual Report said:

> During the past year the Engineering Department has been engaged in planning and directing the introduction of machine switching or automatic switchboards into the Bell System. . . . Such studies show that in the large cities machine switching equipment should be employed for extensions necessary to provide for growth and for reconstruction to replace wornout equipment. Our experience has shown that by this procedure we are enabled constantly to change to new types of apparatus as they are developed, with the least amount of disturbance to the service, in the minimum time and without disturbing effects upon the employees or on the financial situation.

The technical developments that led to this great change in outlook with regard to the use of full automatic switching (i.e., customer dialing) for large-city offices will be covered in Section 4.3.

During the period when Bell was concentrating on large-city problems, the Automatic Electric Company was continuing to develop the Strowger approach. The many faults of the 1904 system gradually were eliminated. Common-battery operation was introduced in 1905, the equipment was made more reliable and used more efficiently, and features were added that made automatic service more comparable to good manual service. By 1914, the Strowger system was much more attractive than ten years earlier, there had been many sales to independent companies, and to the surprise of many telephone people it was found that relatively few customers objected to the added work of dialing. Although Bell System development effort had been devoted largely to the American "big-city" problem, for which

Strowger equipment was not well adapted, there was no lack of experience with Strowger systems. By 1915, Bell was operating equipment manufactured by Automatic Electric that served some 90,000 stations, mostly acquired through purchase of independent properties. This first-hand experience and further studies led to a patent agreement with the Automatic Electric Company in May of 1916 giving the Western Electric Company manufacturing rights for Strowger-type equipment. In 1917, an order was placed on Automatic Electric for equipping the Bell office in Norfolk, Virginia. This was the beginning of a long period during which Bell used Strowger-type equipment in smaller cities and in some PBX systems. From the start, this equipment was made by Automatic Electric to Bell System specifications and over the years many changes were introduced by Bell to improve performance and flexibility and to add operating features. Beginning about 1926, Western Electric began manufacturing this type of equipment but Automatic Electric remained an important supplier through 1936.

After this somewhat lengthy and broad outline of the evolution of automatic switching in the United States, a discussion of the technology will be appropriate. It is already clear that two basic systems had evolved, one for the small cities, another for the large.[79] These will be covered separately. The first, using the Strowger principle, came to be known as the step-by-step system. This system is a form of direct dial control in that the switches at each stage in the dialing process are directly responsive to the digits being dialed. The second is of the common-control type. In such systems the dialed information is stored in centralized equipment before being used to control switching operation.

4.2 Direct Dial Control—The Step-by-Step System

Strowger's direct control system, as described in his patent, was in part impractical, but his concept for the connecting switch was not only original but so sound that it evolved into the basic piece of apparatus used in the step-by-step type of automatic system.

4.2.1 Basic Connecting Switch

The underlying principle was a switch that could be moved readily in two directions, vertically to give access to different levels of contact points, and horizontally to give access at any desired level to

[79] While only two automatic systems were extensively used in the United States, several other systems received limited use. Probably the most important was that developed by Clement and manufactured by North Electric Company. Western obtained manufacturing rights to this system in April 1916.

contact points on segments of a circle. As it finally evolved, it was a decimal system with ten vertical levels, each of which carried ten sets of contact points in the horizontal plane, thus giving 100 connecting positions. The general form it assumed after a number of years of development is shown in Fig. 6-43. The assembled switching mechanism consisted of a "bank of contacts" on the arc of a circle with a "wiper" mounted on a shaft rotating on the axis of the circle. The bank assembly is shown in (a) of Fig. 6-43 and (b) is a top view showing some of the positions which could be assumed by the wiper. The wiper is shown in greater detail in (c). As ultimately developed, each level of the bank had two sets of contact points, one above and one below an insulated mounting plate. The wiper had two arms also insulated from each other, one making contact with the points above the insulating plate and the other with those below. Thus, each bank contained 100 pairs of contact points, sufficient for making connections between the wiper and any 1 of 100 pairs of talking or signaling conductors. Several banks could be mounted one above the other with the wipers all on a common shaft, thus giving an assembly that could handle not only the connections required for the talking path but also those for control of switching (or signaling).[80]

The bank arrangement for the T and R conductors is shown schematically in Fig. 6-44a. The normal (inactive) position of the wiper was below row 1 and to the left of column 1. In operation (Fig. 6-44b), under control of signals from the subscriber, it was first moved vertically one level at a time to the desired level, after which it was turned on its shaft to the desired pair of contact points.

Figure 6-45a not only shows the way in which wipers on several banks could be controlled by one shaft but also illustrates the manner in which the vertical and horizontal (rotary) motions of the wipers were accomplished. The drive mechanism used two kinds of ratchets mounted on the shaft above the banks, both actuated by magnetically operated pawls. The upper ratchet provided the vertical motion and the lower one the rotary motion. To avoid unnecessary complications, the details of the mechanism are not shown. However, some idea of both the complexity of the switch and the variety of assemblies obtainable is provided by (b) and (c) of Fig. 6-45, which show a set of modern step-by-step switches, all using the basic bank mechanism evolving out of the original Strowger idea of a switch with two types of motion. The large number of relays associated with each switch should be noted. These relays, mounted above the banks and occupying

[80] As with manual boards, the connections within an office were usually on a 3-wire basis. The talking pair (designated T and R by analogy with the manual counterpart) was handled by one bank of a switch and a single-conductor local-signaling circuit was carried by the second bank of the switch.

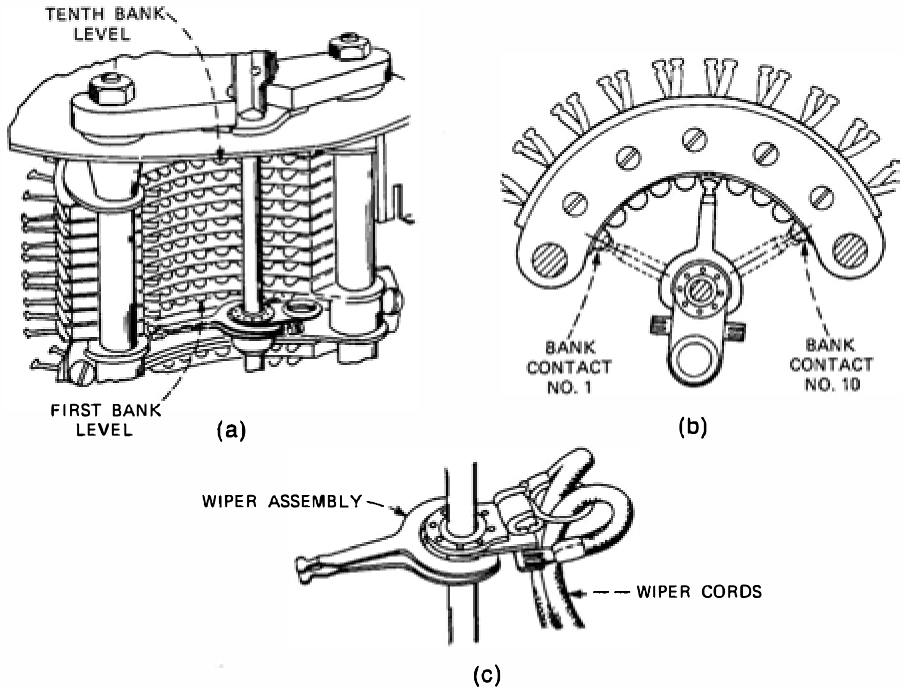


Fig. 6-43. Details of the step-by-step switch: (a) bank assembly; (b) top view of switch bank; (c) wiper assembly.

the major portion of the space, provided the control circuitry that made it possible for the switching mechanism to follow the dial pulses and stay in place not only between pulse trains but until released by switchhook action at the end of the call.

4.2.2 Early Step-by-Step Development

After this explanation of the principles involved in the step-by-step switch, we can return to the beginnings of automatic telephony and see how the switch evolved from the Strowger concept and the manner in which it was incorporated, with the help of other inventions, into a practical system.

The original Strowger concept involved a ten-level switch such as we have described, but each level consisted of 100 contact points mounted in a circle. Each line would be connected to the wiper of its own switch and have access to 1,000 contact points. These contact

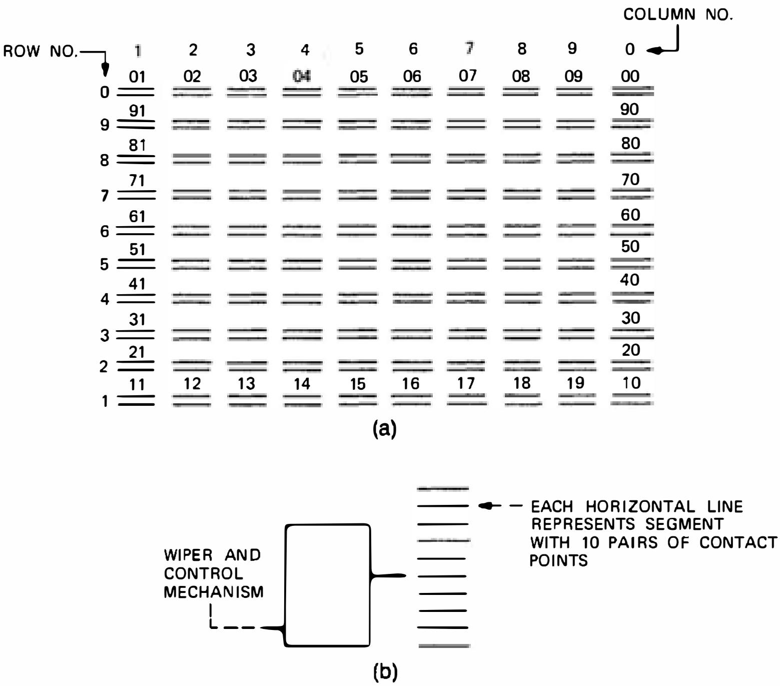
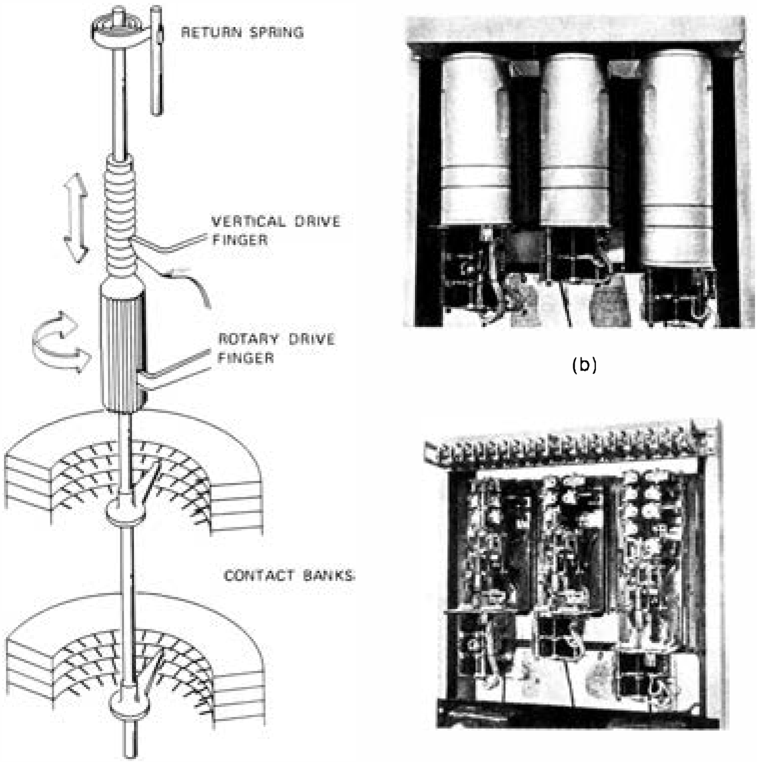


Fig. 6-44. (a) Schematic of switch bank showing numbering system (as viewed from contact side). (b) Simplified representation of bank and wiper.

points would be connected to the lines entering the office and be multipled to identical points on all other switches. Thus, a complete system could handle up to 1,000 lines, each of which would have a switch giving direct access to each of the other lines in the office.[81]

Strowger proposed a magnetic ratchet and pawl system for operating the wiper under control of four pushbuttons in the subscriber's station set. Each of these buttons was connected to the central office by its own wire and sent out a direct-current pulse each time the button was pushed. The first button to be operated moved the wiper vertically, one step for each closure, thus selecting the level corresponding to the hundreds digit. The second button, indicating the tens digit, moved the wiper arm horizontally ten steps

[81] The basic idea, as in all early automatic central offices, was to use the switch as a mechanical substitute for the switchboard cord in making a direct connection between calling and called lines. This concept imposed severe limitations on switch design and until the use of switches in tandem, covered later, a 1,000-line office was about the maximum size practical.



(a) (b) (c)

Fig. 6-45. Step-by-step switches: (a) drive mechanism; (b) modern switches, with equipment covers; (c) modern switches, covers removed. (Diagram redrawn from Feder and Spencer 1962, p. 137)

for each closure, and finally each closure of the units button moved the arm one additional step horizontally. The fourth button was pushed at the end of the call and restored the mechanism to normal. Ringing was by the magneto arrangement customarily used with local-battery operation. It will be noted that this was an automatic system in the sense that the functions of the operator were performed mechanically but was anything but automatic from the customer's viewpoint since he was required to perform many precise operations. For example, connection to Line 999 required nine properly timed closures of each of the three digit-buttons, manual ringing, and finally an end-of-call signal. Besides these complex customer operations, the telephone plant was complicated by the use of four wires in addition to the voice transmission circuit. Furthermore, the wiring involved in connecting 1,000 step-by-step switches in multiple was not only complex but added considerable transmission loss. And, finally, the size of an office was limited to the capacity of the switch.

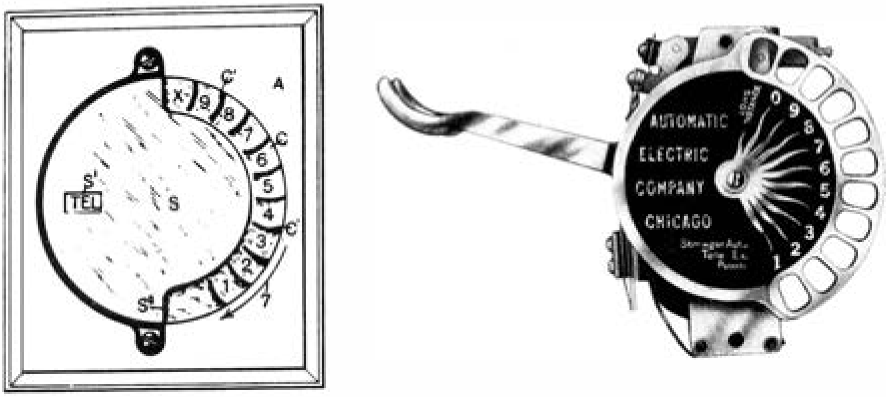
The problems inherent in Strowger's original proposal were solved in various ways. One of the first was the invention, in 1896, of the dial as a replacement for the three pulsing pushbuttons.[82] Its original external form is shown in Fig. 6-46a. This fingerwheel dial consisted of a disk rotating on a center spindle. The ten digits were marked on a portion of the periphery and a ridge, or fingerhold (C^1 in figure), was adjacent to each digit. The dial was operated, in a manner familiar to everyone today, by placing the finger on the desired digit and rotating the dial until the finger reached the stop position (S^2 in figure). Upon release, a spring and governor mechanism returned the dial to normal at a controlled speed and during the return sent out a series of pulses corresponding to the number above the fingerhold.[83] The fingerhold, however, proved awkward to use and was superseded by a hole for the finger. A dial of this type used in the early 1900s is shown in Fig. 6-46b.[84] Over the years, dials with further improvements were developed, as covered in Section 4.3 of Chapter 3, but the basic principle remained the same, namely, setting the dial for the digit desired and sending out pulses during the dial's return to normal. This, incidentally, is a procedure which dates back to the call boxes used by companies such as the American District Telegraph Company to provide messenger and express service (refer to Sections 3.1 and 3.2).

The dial was a great improvement over manually operated pushbuttons since the pulses it sent out were uniform, having a duration and spacing designed to best operate the central-office mechanism.

[82] A. E. Keith and J. and C. J. Erickson; U.S. Patent No. 597,062; filed August 20, 1896; issued January 11, 1898.

[83] From the beginning, the decimal system formed the basis of dialing. In American practice, digit 1 sent one pulse on the dial return, 2 sent two pulses, and so forth. The tenth position, which sent out ten pulses, was originally marked X but was soon changed to 0. Some administrations have used different arrangements. Sweden, for example, uses a dial numbered from 0 to 9 in a counterclockwise direction so that dialing 0 sends out one pulse and 9 sends ten pulses. New Zealand also numbers from 0 to 9 but in a clockwise direction so that the first position, marked 9, sends a single pulse and decreasing dial numbers send increasing numbers of pulses. The problems arising many years later when international dialing was introduced can be imagined.

[84] At this time, a mechanical interlock between the dial and switchhook was employed to provide a signal for releasing the switches upon hang-up, thus eliminating the end-of-call pushbutton. The dial illustrated also had 11 finger positions, the last two being marked 0 and "long distance," respectively. Both sent ten pulses and brought in an operator when dialed as the initial digit. Later the use of the ten-position dial was resumed, the last position being marked both 0 and "operator" as we do at present.



(a) (b)

Fig. 6-46. Early forms of Automatic Electric Company dials. (a) First form of station-set dial, which replaced the pushbuttons previously used. (b) Dial with fingerholes and 11 positions, circa 1900–1905. [Part (a) from U.S. Patent No. 597,062; part (b) from Miller 1905, Fig. 508]

It also eliminated the need for a separate wire to the central office for each digit. One of the first schemes for using the dial employed signals to ground over each side of the line. The dial pulses for operating the central-office switches were sent over one side and, when the dial had sent all of these and returned to normal, a pulse was sent over the other side to indicate the end of digit pulsing. Such a scheme was used until about 1905 after which timing arrangements were introduced to indicate the end of the digits. This avoided the necessity for two types of signals and opened the way for loop (metallic) signaling and the provision of common-battery talking current.

About 1896, Keith and the Erickson brothers started work on means to eliminate the necessity for multipling all the subscriber lines to each switch. They worked on a 1,000-line system using two stages of switching. Instead of a single switch having 1,000 contact positions (which previously had proved impractical), they used two smaller switches connected in tandem by means of transfer trunks.[85] Specifically, their approach was to make connection to the called subscriber by means of "connectors" that were 100-point switches (ten vertical levels, ten horizontal positions) much like the ultimate design

[85] The idea was not new. The use of transfer trunks between groups of selectors had been proposed several years earlier for a telegraph application by J. G. Smith who had been granted U.S. Patent No. 481,247 on August 23, 1892. Keith and the Ericksons were, however, responsible for the commercial application and their ideas were covered by U.S. Patent No. 672,942, applied for June 23, 1897, and issued April 30, 1901.

discussed earlier. The switchpoints were connected to the lines, one of these connectors (or a multiple group of them) being required for each 100 lines in the office. The calling subscriber had to have access to the input, or wiper, of the appropriate connector, and this was accomplished by having each incoming line connected to the wiper of a switch that selected a transfer trunk connected to the wiper of the required connector. These earlier switches in the train, known as "selectors," responded to the first digit dialed while the 100-point connector responded to the last two digits. The original 1,000-line system was installed in Augusta, Georgia, in March of 1897 to serve some 400 lines. In this system, there was a simple 10-point switch serving as a selector, each switchpoint providing a trunk to a connector giving access to 100 called subscribers.[86] The scheme is diagrammed in Fig. 6-47.

As shown in this figure, the output of several selectors was multipled so that several subscribers had access to a given set of connectors. This was necessary to reduce the number of switches but also led to the failure of the system using 10-point selectors. Without multipling, a 1,000-line office would require not only 1,000 selectors but also 10,000 connectors (ten for each line). By multipling the selector outputs, the number of connectors was reduced accordingly, but as soon as one connector was in use for a call from a line, the remaining 99 lines wired to it were denied service from any line connected to the same selector multiple. Obviously, service would frequently be blocked if many selectors were in multiple. The trouble was that only one transfer trunk was provided between any connector and a group of subscribers (multiple selectors). As trunking studies had shown (see Section 3.6), this was a very inefficient way to operate. What was needed was a much larger group of trunks between switching stages.

This needed improvement was accomplished by using selectors with more than one contact on each level so that a number of trunks (and connectors) could be made available to each line. The concept as it developed, in somewhat simplified form, was to divide the 1,000 incoming lines into ten groups of 100, each line in the group having access to ten connectors for each 100s group. Thus, with ten choices available to each 100 lines, there was little chance of having the call blocked by finding all ten connectors in use.

It is both difficult and unnecessary for the purpose of this discussion to describe the process in detail, but the principle was fundamental

[86] An interesting feature was that the 10-point selector used essentially the same mechanism as the connector but with only a single switchpoint at each level. The first digit dialed moved the wiper vertically to the proper level with the usual pawl and ratchet arrangement and at the end of the pulse sequence the wiper was rotated one step to engage the switchpoint.

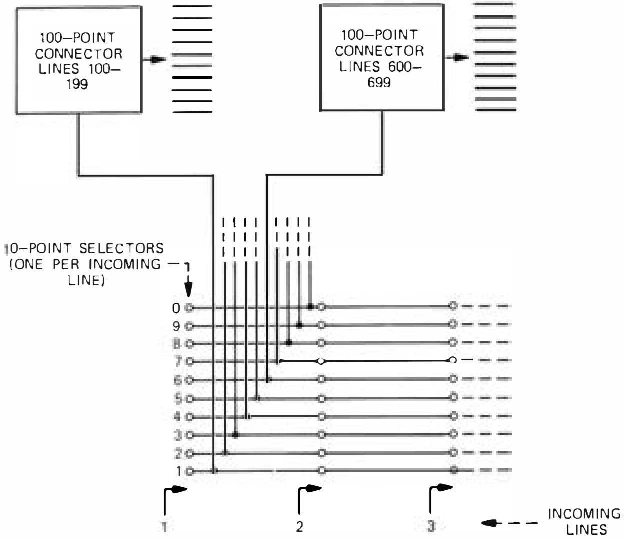


Fig. 6-47. Early 1,000-line system with transfer trunks and 10-point selectors.

to the practical application of step-by-step switching and will perhaps be made clear by Fig. 6-48. As shown, each line was connected to the wiper of an individual 100-point selector with banks arranged in the same manner as a connector, namely, ten levels with ten contacts each. In a 1,000-line office, the first digit dialed elevated the wiper to the level corresponding to the hundreds-group dialed (the 600s-group only is illustrated in detail). Each contact on this level was connected by a trunk to one of the ten connectors serving the appropriate hundreds-group. After the first (hundreds) digit had been dialed, the selector mechanism automatically and rapidly moved horizontally until an idle connector trunk was found, at which point it stopped so that the next two digits dialed could move the connector brush vertically for the tens digit and horizontally to the units digit to make connection with the desired line. Provision was made for a busy signal if the line was in use, but the means for accomplishing this, the automatic "hunt" for an idle connection, and other details are unnecessary to the understanding of the basic principle, namely, that by the use of two levels of switching, it was possible to use 100-point switches to interconnect 1,000 lines.

The application of multilevel switches connected by transfer trunks did not stop here. It was possible to use a three-stage system to reach 10,000 lines by having the first-stage selector (connected to the calling

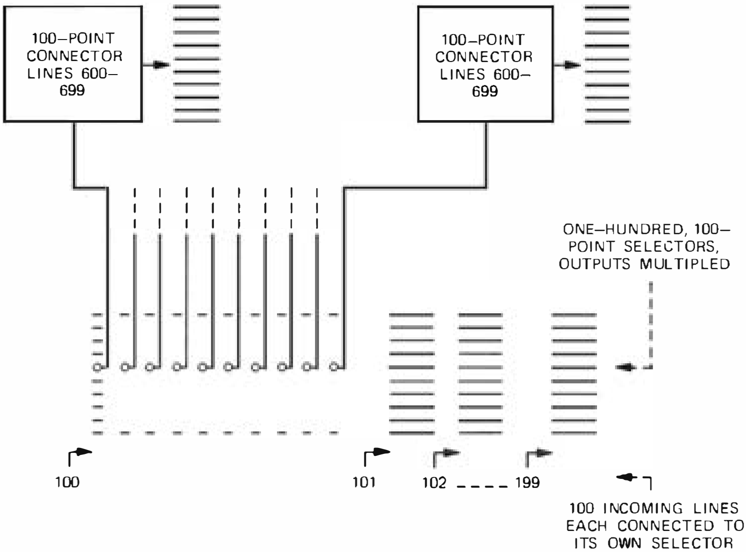


Fig. 6-48. Switching in a 1,000-line office using 100-point selectors and connectors that are corrected by transfer trunks. Illustration covers only one 100-line group of calling lines (100–199) and one 100-line group of called lines (600–699).

line) respond to the thousands-digit and select a second-stage selector which was actuated by the hundreds-digit and selected a connector (controlled by the last two digits) that established a circuit through to the called line. Obviously, the principle could be extended and was indeed used to handle more than four digits, some of which were used to select trunks to selectors in other central offices, thus accomplishing step-by-step switching in a multioffice city. This idea of an expanding, tree-like system of trunks and 100-point switches was basic to the practical evolution of the step-by-step system.[87] It was one of the major inventions underlying the type of direct-control switching which has received the most use up to the time when this history is being written (the 1970s).

Most of the developments just described had been reduced to practice by the early 1900s and were implemented in the major step-by-step installations made at Grand Rapids, Dayton, etc. Specifically, they included the use of 100-point banks connected in tandem by transfer trunks, the invention of the dial as a replacement for push-

[87] This is often described as a "Progressive Switching Network" since the path is established progressively, link-by-link, from the input toward the output. The idle or busy status of each successive link or output is not determined until an actual connection is attempted.

button pulsing, and the use of talking pairs for carrying signaling pulses from the subscriber to the central office.[88] A number of other refinements also were introduced but we need not discuss them in detail. Probably the most important was the use of a rather elaborate switchhook mechanism to replace the end-of-call button originally proposed by Strowger.

After examining these improvements in detail, it is easy to understand some of Kempster Miller's enthusiasm in 1905 for the future of automatic telephony. It is true, as pointed out previously, that as of that time the service provided was far inferior to that provided by a good manual system of the current type. However, the enormous improvements made in the 15 years after Strowger's original, somewhat crude, proposal certainly gave hope that the remaining problems also could be solved so that ultimately service, functionally at least, could be comparable to good manual service. Most of the problems were indeed solved in the next ten years, as covered by the following section, and a field of use for step-by-step systems became well-defined.

4.2.3 Evolution of the Practical Step-by-Step System

By 1904, after 15 years of development, the step-by-step system had received a modest amount of use in offices of various sizes, some with as many as 6,000 lines; but, as we have noted, the service was inferior to that of contemporary manual systems. A considerable amount of work was still needed to develop a system competitive with manual boards from the standpoint of both service and economy. Some of the necessary developments had already been started and most were completed during the following five to eight years. We cannot take the space to give full details, but a few rather fundamental inventions deserve discussion.

4.2.3.1 Fundamental System Concepts. Even though the systems of 1904 were deficient from the service standpoint, they incorporated most of the basic concepts that were to make step-by-step a viable switching system. Of most importance were: (*i*) the dial, which generated uniform control signals as a result of a simple customer operation; (*ii*) the two-direction, 100-point switch bank, which was adaptable to many types of connecting functions; and (*iii*) the concept of successive switching stages, connected by transfer trunks, which used 100-point switches to interconnect lines in central offices of large capacity

[88] While signaling was carried out over the talking conductors, it was accomplished by means of two types of signals, one using the tip conductor and the other the ring conductor with ground return.

and even to interconnect central offices. Only a few basic inventions were still needed to meet practical requirements. Perhaps the most important of these were the line-switch and line-finder concepts.

It will be recalled that, in the system of 1904, each line required a 100-point selector switch to pick a transfer trunk to the next higher switching stage. These relatively expensive switches made up a large percentage of the switches in the office but received little use, each one being actuated only during the small part of the day when the subscriber's line associated with it was being used on an outward call. In many cases this was not much more than a half-hour total for the day. This costly use of first selectors was obviated by using either line switches or line finders. These two devices were functionally similar in that they provided means for using less than one selector per line but acted in a different manner. The principles are shown in Fig. 6-49. The basic line switch, illustrated by (a) of the figure, is a simple switch, with its wiper connected to the line, which hunts an idle trunk to a first selector when the subscriber line is ready to place a call. One of these switches was still required for each line but they were relatively inexpensive and their use reduced the number of much more complicated selectors by a factor of ten or so. The line-finder objective was the same but operated in the reverse manner

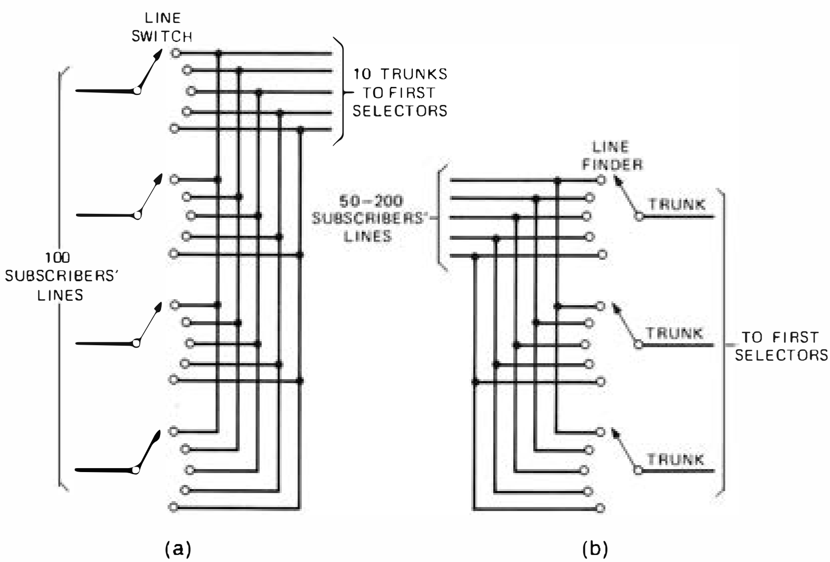


Fig. 6-49. Line-switch and line-finder concepts: (a) basic line switch; (b) basic line finder. (Redrawn from Smith and Campbell 1914, Figs. 18 and 25)

[see (b) of the figure], the wiper being connected by trunks to the selectors and the switchpoints to lines. When a call was initiated, the wiper from an available trunk hunted for the line requiring service.

In practice, both line finders and switches were developed beyond the simple rotary switches illustrated to provide more economical and compact apparatus. Generally speaking, the Automatic Electric Company favored line switches devised by Keith and the Bell System preferred line finders, particularly after Bell had developed, in the twenties, a 200-point device (using two 100-point banks per finder) with a special commutator invented by W. W. Carpenter to provide vertical as well as horizontal hunting.

The Automatic Electric line switch developed by Keith[89] was a rather ingenious multiple switch with a common mechanical control. In a sense, it was a form of coordinate switch array providing connections between lines and trunks as shown schematically in (a) of Fig. 6-50. The mechanical device for accomplishing these connections is illustrated in (b) of the figure. It consisted of a number of plungers, on different levels, moved around an arc by common vertical shaft S. Actuating a plunger magnetically closed contacts between line and trunk conductors and also disengaged the plunger from the vertical shaft. The plunger remained actuated as long as the line was in operation and the shaft was free to move other plungers about as required to effect other line-trunk connections.

It will be observed that the basic purpose of line finders and switches is to provide transfer trunks between the subscriber's originating line and the first level of selectors. By so doing, the selectors can be reduced in number and used more efficiently. Commonly, the number of selectors required is one for every ten lines but better efficiencies are possible by using two levels of line switches or with 100- or 200-point line finders as used in Bell System step-by-step offices after about 1927.

Two other inventions improved the efficiency of the switching network within an office. One of these was the use of "bank slip"[90] which, in addition to other benefits, reduced the time required for selector operation. It will be recalled that in order to obtain trunk efficiency, a group of selectors was given access to ten transfer trunks which were multipled at the selector output. During the periods of heavy traffic, many of these trunks would be used simultaneously and, if all trunks appeared on the same bank number, a selector in need of a trunk would have to "hunt" over a number of contacts to find an idle one. To reduce the hunting time, the wiring

[89] U.S. Patent No. 1,304,324; filed April 29, 1905; issued May 20, 1919. It was first put in use in 1907.

[90] A. E. Keith; U.S. Patent No. 831,876; filed March 9, 1905; issued September 25, 1906.

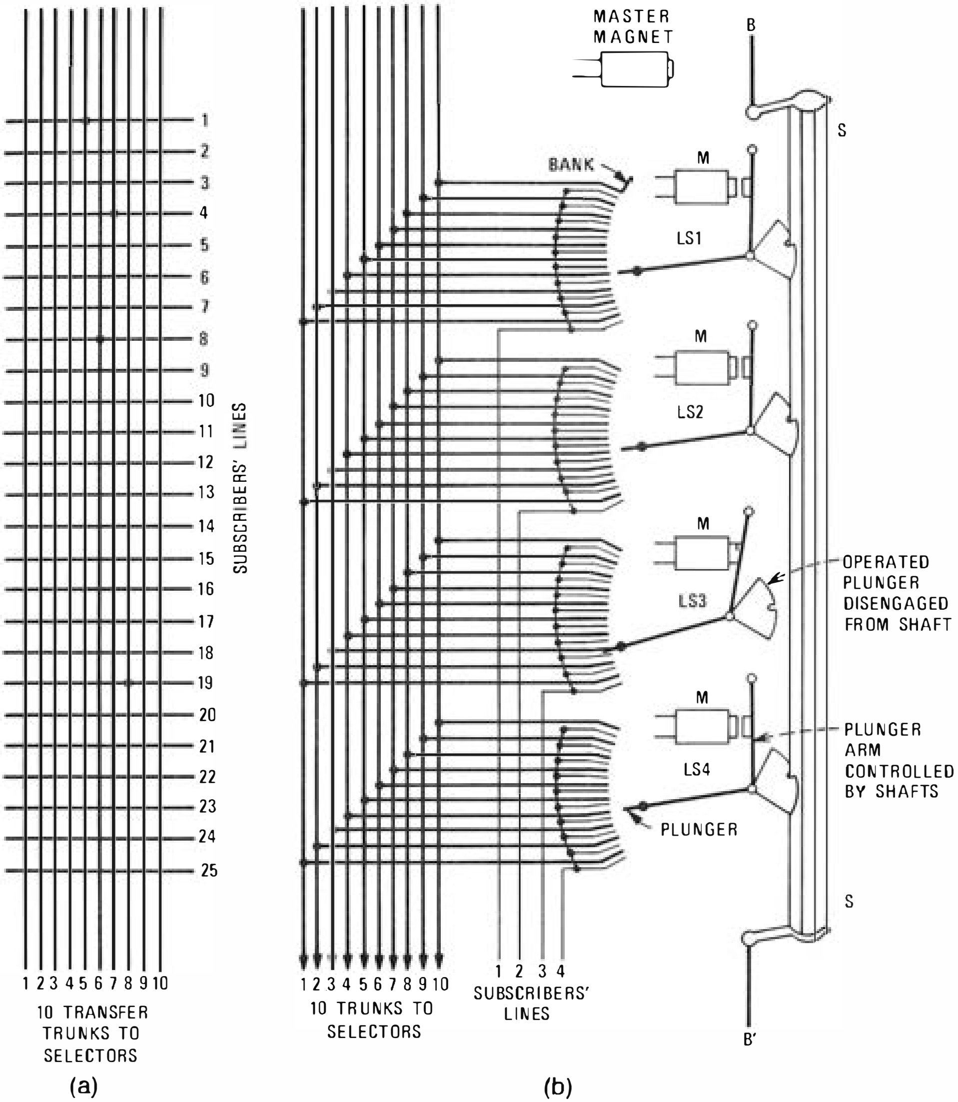


Fig. 6-50. Keith line switch: (a) relation of subscriber's lines to transfer trunks; (b) mechanical arrangement. (Redrawn from Smith and Campbell 1914, Figs. 19 and 20)

was "slipped" so that a different trunk number appeared as first choice on each selector in a group of ten. A common scheme for accomplishing this is shown in Fig. 6-51. With this arrangement there is a large probability that the first contact will be idle and thus hunting over a large part of the bank will be avoided.

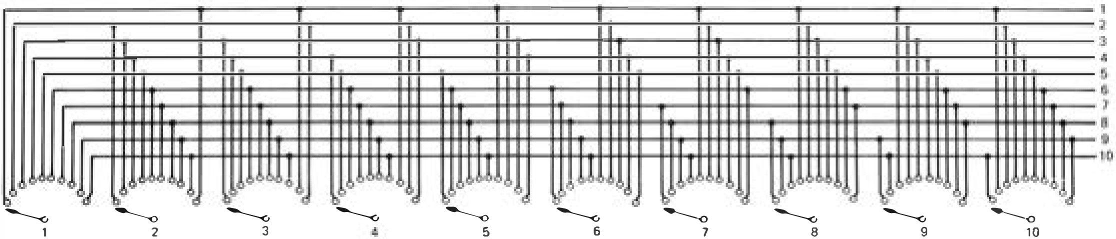


Fig. 6-51. Illustration of one scheme for "bank slip." (Redrawn from Smith and Campbell 1914, Fig. 12)

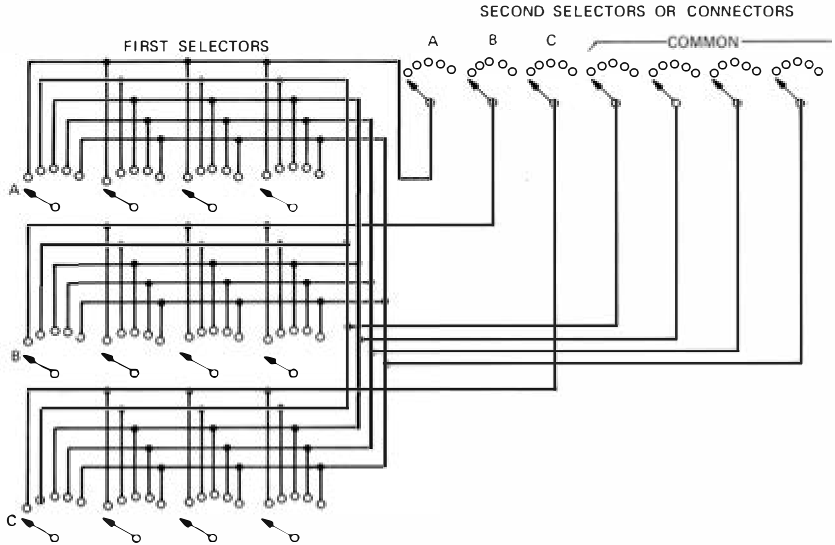


Fig. 6-52. Simplified schematic illustrating "graded multiple" combining individual trunks (A, B, C) with common trunks. (Redrawn from Smith and Campbell 1914, Fig. 16)

Another bank wiring scheme, perhaps more fundamental in importance, was that known as "graded multiple." It will be recognized that the strictly decimal organization of the network of switches could be highly restrictive since the traffic was not always such that a selector needed outputs to precisely ten switches in a given digit group at the higher level. If traffic was light, the operating and control mechanism of high-stage switches could be omitted. (The banks were usually wired into the circuit, however.) But the need for more than ten higher-stage switches was not so easily handled. To meet this need, E. A. Gray of AT&TCo invented the graded multiple.[91] The idea was to have one or more individual trunks (and corresponding switches) wired as first choices to a limited number of selector banks. The remaining selector contacts were connected to common trunks multipled to the banks of all selectors and were used when the individual trunks were busy. The simplified diagram of Fig. 6-52 shows how five first-selector trunks can be used to give access to seven second selectors. The extension to ten-trunk switching is obvious. The principle could also be extended further by having several groups of common trunks, each of which was wired to the contacts of a portion of the incoming selectors. By methods such as these, it was possible to have a group of 100 selectors feed more than the ten trunks that normally would be associated with the ten contacts on a given bank level.

The techniques just described provided the necessary means for organizing a switching network which could meet load requirements in a reasonably efficient manner. These techniques, and manufacturing improvements which became available in the 1904–1914 period, did much to make step-by-step switching economically more competitive with manual. By 1914, almost all of the basic principles that were to be used in this type of switching during the next 50 years had been developed.[92] Their durability is a tribute to the great ingenuity of the early workers in this field.

4.2.3.2 Service Improvements. While improved network organization did much to make step-by-step economically competitive with

[91] This resulted from the effort being devoted by Bell to the development of "large switch" systems, as described in Section 4.3. The concept originated about 1905 and a patent application was filed July 30, 1907. U.S. Patent No. 1,002,388 was issued September 5, 1911.

[92] During the long life of step-by-step, many apparatus and manufacturing improvements were devised and incorporated in production. However, with the exception of the addition of a degree of indirect control in certain special situations, the basic network organization for the small and medium-sized cities was not significantly changed. The main impact of the common-control system described in Section 4.3 was on large cities and for many years step-by-step remained a system functionally and economically attractive only for small and medium-sized exchanges.

manual switching, automatic switching could never compete seriously with manual systems until it provided all the services offered by the latter. Providing these services automatically was beyond the technology available in the first part of the twentieth century and a compromise approach was used. Instead of attempting full automation, development was confined largely to automatic exchange (and some PBX) switching systems which could be combined with manual operation in those areas where technology could not provide suitable automatic replacement.[93]

Even with this limited goal, the service provided by early step-by-step systems lagged behind that furnished by manual systems, but the capabilities of the two types of systems began to be equalized after 1905. It was in this year that common-battery talking was first used (at South Bend, Indiana), but it was not until 1907 that the metallic circuit was used, as in manual boards, for carrying speech, battery supply, and control signals. At this time, the use of slow-release relays made it possible to eliminate signaling over ground-return circuits as used previously. Typical battery supply arrangements used on exchange calls are shown in Fig. 6-53. On a non-trunked call, talking battery was supplied to both parties through a bridged retard scheme using holding magnets in the connector as retardation coils. On trunked calls, the connector furnished battery to the called party but the calling party was supplied from its own office through a "repeater," a device which provided an interface between trunk and interoffice signaling.

Four-party, full-selective ringing service was introduced about 1905. The Automatic Electric Company used a "harmonic" ringing scheme to distinguish among the parties on the line. Each part was assigned a specific one of four ringing frequencies and also was furnished with a ringer tuned to respond only to that frequency. The frequencies commonly used were 16.6 hertz and its second, third, and fourth harmonics. Each party had its own line number; usually these were alike except for the third digit from the end. This digit served to trunk the call to a group of connectors which supplied ringing current at the appropriate frequency. Step-by-step systems manufactured for the Bell System were designed for the biased ringing that the System had long used (see Section IV of Chapter 3).

Ringing initially had required specific action by the calling party but in the 1905–07 period automatic ringing was introduced, the ringing current being sent out through the connector as soon as the

[93] For example, manual operation, with a few exceptions, was used for handling and charging toll calls until after World War II and even now some services such as directory assistance are still handled more effectively on a manual basis.

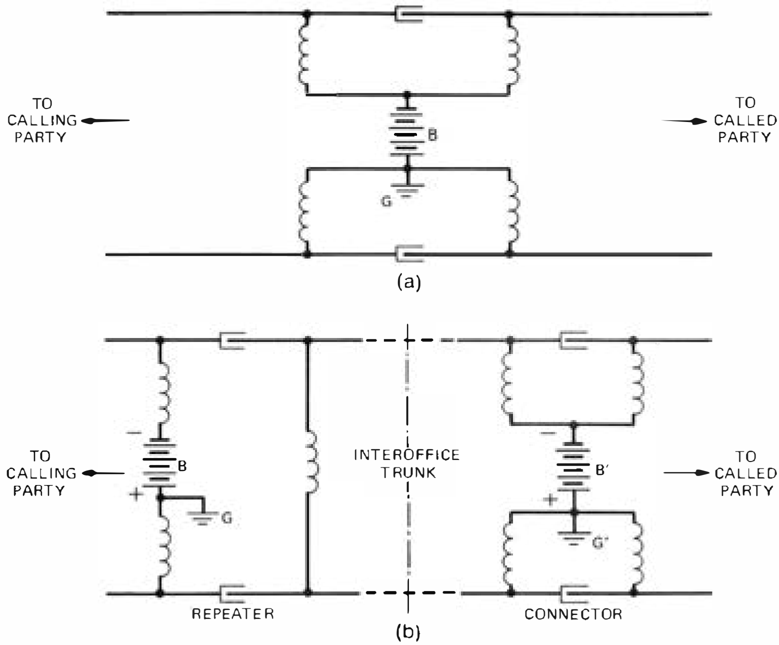


Fig. 6-53. Battery supply circuits for step-by-step offices: (a) talking circuit of single-office connection; (b) talking circuit of interoffice connection. (Redrawn from Smith and Campbell 1914, Figs. 51 and 56)

wiper made contact with the desired line. The calling party received an appropriate signal to indicate that ringing was in progress.

Two other signals also were supplied automatically to replace operations normally performed by the operator. The first of these was dial tone, which indicated that equipment was available for service and thus replaced the operator's request for a number.[94] The second signal indicated unavailable (busy) equipment. If the connector, upon reaching its dialed position, found the line in use, it continued stepping to a final segment where the brush was connected to a busy tone transmitted to the calling party. Later, similar arrangements were used to indicate busy intraoffice or interoffice trunks.

4.2.4 Manual and Automatic Interconnection

Originally, the interconnection of manual and step-by-step offices presented no problem since conversion to automatic usually was

[94] This was required when line finders and switches were introduced since they were not always immediately available to the calling subscriber. Prior to this time, dial tone was not needed since each subscriber's line was permanently connected to its own first selector.

done only in small cities where it was carried out for an entire exchange. Calls to areas beyond the automatic exchange were handled through the toll operator. Later, new automatic offices were introduced on a growth basis in cities where the existing manual offices were retained. Interface requirements were an important part of the "Big City Problem" faced by the Bell System and will be discussed in the following section. Generally speaking, the same general principles were followed in step-by-step areas.

4.3 Common-Control Systems

The unique feature of the system just described was a piece of apparatus of great versatility. It was a switch, directly controlled by the caller, with two directions of motion. This basic configuration, with appropriate associated control relays, could be used to perform the various functions required to set up progressively a network for interconnecting subscriber lines. The unique feature of the system we are about to describe was not a piece of apparatus but a system concept, common control,[95] which could be implemented with a great variety of connecting switches. Over the years the basic concept has indeed been used with two-dimensional switches of the Strowger type, other types of rotary switches, linear switches of the panel type, crosspoint switches, relays, electronic devices, and even with switching on a time-division basis. While this history will be devoted largely to progressive systems using switches of the panel, or linear, type, it is important to remember that the common-control principle, which (in simplified form) made its start with this system, was one of great potential power. It has expanded in scope and been adapted to the application of new technology up to the present day and in all likelihood this process will continue well into the future.

4.3.1 The Basic Concept

In a simplified way, the common-control principle can be stated rather briefly. Instead of using the dial pulses directly for the operation of connecting switches, the pulses are processed in a manner that operates the interconnecting mechanism and performs other switching functions in the most effective way. A common-control switching system, particularly in its later and more sophisticated form, has been likened to a highly complex computer with many input and

[95] Common control is a generic term used widely in the Bell System as a broad description of any system in which the dial pulses from the user do not directly control the connecting switches but are first processed by equipment used in common by many switches while the call is being established. The simpler forms of common-control systems, such as those described in this chapter, are sometimes referred to as Indirectly Controlled Systems.

output terminals (subscriber lines and trunks) which can be interconnected at will as directed by dial pulses sent from any of the input stations. The common control accepts the dial pulses, puts them in a temporary store, and translates them into the machine language that is best adapted to controlling the interconnecting switches. In this converted form the stored signals are used, under the direction of a built-in program, to control the switch operation required to make the desired connection, alert the output station, and send information to the input station on the status of the operation (e.g., in progress, complete, or unavailable).[96]

4.3.2 Genesis of Common Control—The "Big City Problem"

In retrospect we can see that the common-control concept evolved from Bell work on what may be called the "Big City Problem." At the time, bits and pieces of the problem were solved by ingenious ideas, probably with no broad philosophical unifying principle, but as these parts were developed, it became clear that collectively they constituted a new approach to machine switching, giving efficiency, flexibility, and potential for new functions that were impossible with direct control of each switch by pulses from the subscriber. Let us go back and review how this came about.

Before the turn of the century, offices of 100 lines or less appeared to be the most favorable market for automation because of the high cost of operators in such locations. This proved not to be the case since growth was rapid and the simple schemes favorable for very small offices were not well adapted to growth into the larger sizes soon needed. Early in the twentieth century it was apparent that the bulk of the cities would fall into the single-office (100–10,000 lines) category, a field actively being developed by the Automatic Electric Company, using the step-by-step approach. In 1905, Bell explored the small-city area very carefully. It was found that the systems then available could not provide service equivalent to manual. Even after service deficiencies were removed (as they were in a few years), they could not at that time compete economically with modern manual systems. Such existing systems could expand very inexpensively by adding positions whereas automation usually required replacement of the entire office.

Thus, about 1906, both experience and logic suggested a look at automation in large cities. This was reinforced by the enormous

[96] This is oversimplified and describes in a general way the original concept of a switching system. Today, systems are far more complex and designed to perform many functions besides interconnection of telephone subscribers. But this is beyond the scope of this portion of our history.

telephone growth in these cities (mostly served by Bell). Such growth required the frequent addition of entire offices (economically the most favorable field for machine switching) and also was causing difficulties in obtaining and retaining adequate operator personnel. A review showed that this was indeed the field in which Bell effort should be concentrated even though this was, unfortunately, the area where technical problems were most severe and new approaches most needed. It was a fortunate decision since the solution of these problems led to the common-control concept.

Two of the technical problems were outstanding. One was the inflexibility and poor economy resulting from the size limitation of the available switches. (This type of switch provided access to only ten transfer trunks when used as a selector and only 100 subscribers as a connector.) The other problem arose in adding machine offices to a large network of manual offices.

4.3.2.1 Switch Limitations. As we have noted in Section 3.6, theoretical studies of the relation between size of trunk group and its efficiency of use had been started as early as 1898 and ten years later the mathematics underlying this relation was well understood and emphasized the restrictions of the 10 × 10 switch. As an example, if it is tolerable to have all trunks busy 1 percent of the time, a group of ten trunks can carry a load averaging slightly over four calls, but with a 20-trunk group the average permissible load becomes 11 as compared with the eight which would be obtained with two ten-trunk groups.[97] Obviously there was a great potential economic advantage in a system organized about large-access switches.

Gray's invention of the graded multiple went far toward overcoming the capacity limitations of the small-access 10 × 10 switch but was both less efficient in the use of trunks (and associated equipment) and less flexible than a larger switch would have been. The design of a larger switch was not the major problem;[98] the main difficulty lay in finding means by which such switches could be controlled by the dialing system with which the subscriber was familiar. The problem was solved in 1905 by E. C. Molina who proposed translation at the central office of the incoming decimal system of pulses into a non-decimal control system adaptable to large switches and groups of equipment.

As a result, AT&TCo early in the following year proposed a scheme based on the idea of registration and translation of signals by means

[97] Greater improvements result from even larger trunk groups or from more stringent limitations on the permissible percentage of busy trunks.
[98] A number of these large switches are described in Section 4.3.4.

of a "sender" which acted as intermediary between the incoming signals from the subscriber and the switches. Registration made it possible to record, without delay, the desired number in decimal form as received from the customer and allowed the necessary time for processing these signals and for hunting over large groups of lines and trunks. Translation converted the decimal signals into the format best suited to the switch mechanism and trunking plan required for economical operation.

The sender obviously would be a complicated and expensive mechanism but, since it was required only during that part of the call during which connection was being established, it was possible to have it completely separated from the switching mechanism proper and connected only as needed. Thus a small number of these complex senders could be used in common for controlling a large number of the rather simple switches that established the talking path.

This was the beginning of the common-control concept. This concept not only solved the problem of using large-capacity switches but also removed much of the control equipment from individual switches, where it was used inefficiently,[99] and concentrated it in common equipment where it received high usage. In addition, when translation was made changeable by some means such as cross-connection, common control provided flexibility both for adapting machine switching to the trunking plans required in large cities and for meeting the requirements which arose when an increasing number of machine offices were added to a large-city network of manual offices.

Before discussing the problem of manual-machine interconnection, we should note another aspect of common control that was invented during the first 50 years of telephony even though it was not applied commercially until many years later.

Both in direct control and in early common-control systems (such as those described in this portion of our history) the switching network was of the "progressive" type. Selectors "hunted" for an idle transfer trunk, testing each contact position until one not in use was found. In the period between 1921 and 1925, Western Electric engineers developed a coordinate machine-switching system and built a complete laboratory model. In many ways, the system concepts were highly advanced and too far ahead of available technology for economic implementation and, for this reason, development did not go beyond the laboratory at the time although most of the basic ideas were

[99] As noted previously, each step-by-step switching unit consisted of two parts, first, banks of contacts and a wiper arm for establishing connection and, second, some half-dozen relays to control the wiper arm under guidance of pulses from the subscriber and other, intraoffice signals. These relays constituted a significant part of the switch unit but many received very little usage, most being idle during much of the day.

applied later. Of these we have space to discuss only one, the marker principle.[100] In simplified terms, this was an extension of the common-control function to include the directing of connecting switches to idle terminals instead of hunting until one was found.

4.3.2.2 Addition and Interconnect Problems. If machine switching were to be used efficiently for meeting growth requirements in large cities, it meant adding machine offices of 10,000-line ultimate capacity to an existing large network of manual offices of a similar size. Since machine offices would initially be in the minority, it obviously was important to find ways to make these additions which would not require major changes in the manual offices, the trunking plant, and the calling methods used by the customer. Furthermore, any solution adopted would have to use a switching plan that would be flexible enough to provide not only for the predictably large average growth but also for unpredictable variations which were certain to occur. These problems were not only technical in nature but also psychological and while the outcome of the solutions devised for the former were predictable, the solutions of those involving customer reaction were not. As we shall see, the uncertainty about subscriber acceptance of changes in telephone procedures was one reason for delaying the solution of problems involving manual office and machine office interconnection.[101]

The problem of calling from a manual to a machine office was simply solved. The caller could give the desired number by voice to his A operator and she would either dial it over a trunk to the machine office or pass it by voice in the usual manner to a B operator in the machine office who would do the dialing. The latter procedure involved no changes in existing offices and gave minor operator savings. The former procedure eliminated the B operator at the machine office, but required the installation of dials (or equivalent keysets) at all manual A positions. Economic considerations could determine the procedure followed since it was quite practical to start with B-operator dialing when the city had only a small percentage of machine

[100] In addition to the marker, other innovations in the coordinate system which were used some years later in other systems were the crossbar switch (first proposed by J. N. Reynolds of the Bell System in 1913), large multicontact relays, a trouble indicator which used both indicator lamps and a teletypewriter printout, self-checking circuits, and the ability for a "second trial" if a first attempt failed to establish a connection through the switches. Although several of these innovations involved new types of apparatus, other features were based on extensions of the common-control principles already described.

[101] We have space to outline only a few aspects of the "Big City Problem." An excellent discussion of this problem is presented in considerable detail in Chapter VIII of the book by J. H. Robertson, *The Story of the Telephone*, published in London in 1947 by Sir Isaac Pitman and Sons, Ltd., and shows that other administrations also were seeking solutions to the problem.

offices and convert to A-operator dialing when the percentage increased enough to make the procedure economical.

The reverse interconnection, where a call to a manual office originated in a machine office, presented more difficulties. One possibility was to have the machine subscriber dial an operator, ask for the call by voice, and have it handled from there on by manual methods. There was general agreement among telephone administrators that this procedure would be acceptable for occasional calls and it was used for many years to handle toll calls which required operator assistance for other reasons. But it was generally agreed that this solution would be unacceptable on exchange calls. Not only would operator savings be small (being restricted to full-machine calls), but the burden on the customer would be too great. He would be asked to use different procedures for reaching two groups of people and, since machine offices would be added frequently, the people in the groups would be constantly changing. There was a technical solution to this problem too since it was quite possible (particularly with common control) to have the signals from the machine office go directly to the manual office B board where they operated a unit (known as a "call indicator") that visually displayed the desired number.[102]

The most difficult problem came from the need for a different numbering plan if machine offices were to dial calls into either a manual B board or into another machine office. As the manual system evolved to the point where a number of central offices were required, it had become the custom to assign a name, usually with local geographic significance, to each office. Thus each customer's number consisted of a name to designate his office and four digits to indicate one of the roughly 10,000 jack positions, or lines, in the office to which he was connected. The problem in calling from a machine office was that dials were intended to handle digits and had no provisions for handling names. A few small cities using step-by-step systems employed five digits, one of which was for all practical purposes an office designation. However, it was easy to determine that the growth expected in large cities would ultimately require a seven-digit number. This had two disadvantages. Some tests made in the early 1900s indicated that this would be unsatisfactory because the short-term memory span of many people could not handle seven digits and many dialing errors due to memory lapse might occur.[103] This, in itself, was con-

[102] As discussed later, this scheme was adopted and used for many years. As technology became more sophisticated, "call announcer" machines were devised which translated the signals into the spoken word, but this goes well beyond the scope of this chapter.

[103] Unfortunately, the documentation of these tests seems to be unavailable today and it is not possible to state definitely why it appeared so certain that seven-digit numbers would be impossibly long while present-day use is commonplace. It is possible that the interpreters of the tests placed too much emphasis on the difficulties involved in going from

sidered a sufficient reason for not using all-number dialing and, in addition, if dialing from machine to manual office was to be accomplished without passing the call by voice at the machine office, it would be necessary to simultaneously change all numbers within a city to the digital form as soon as one machine office was cut into the exchange network. This was an additional powerful deterrent.

The problems just dealt with in detail were some of the reasons why J. J. Carty, at the time of his 1910 paper, favored only partial automation since the interconnection between dial and manual offices could be handled with fewer complications if the A operator instead of the customer did the dialing.

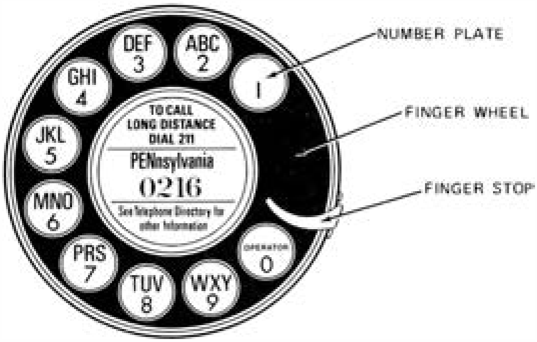
However, the solution to the seven-digit problem proved very simple when it was discovered in 1917 by W. G. Blauvelt of the American Telephone and Telegraph Company.[104] He proposed that positions on the dial be used to designate letters as well as numbers. The first or 1 position on the dial and the last or 0 position would remain unchanged. Each of the other eight positions would be imprinted with three letters as well as the number, ABC being associated with 2, DEF with 3, etc., as shown in (a) of Fig. 6-54. The letters Q and Z were omitted because of their little use. Each phone number would be made up of the first three letters of the office name followed by four digits. This meant that the information to be remembered in transferring from the telephone directory to the dial was very nearly the same as for manual calling. No change in existing numbers in manual offices was required except in those few cases where the first three letters or two or more office names translated into the same digits. Aside from these cases, no directory changes were required, but as an aid to dialing, the first three letters in an office name were ultimately capitalized as shown in (b) of Fig. 6-54.

Common control did little to solve the interconnecting problems just discussed except in the case of the call indicator. In this development, registration in the sender provided the necessary time for transmitting the signals and setting up a display; translation made possible

---

five-digit to seven-digit dialing. This would not be surprising since "human factors" testing of this sort is full of pitfalls and in the early 1900s was in its infancy. In a typical memory test, the digits are presented by voice, one per second, and there is no opportunity for review or "rehearsal" at the end of the presentation. This is a much more severe condition than is encountered in the usual telephone dialing situation where numbers are displayed visually and the dialer can rehearse the number until he is confident that he has memorized it. He also has an opportunity to refer back to the display during dialing if the need arises.

[104] After some 50 years' experience with this scheme, it seems so obvious that it is unbelievable that it took so long to invent and it is difficult to realize the tremendous significance of the proposal when it was made. Perhaps this is merely one more indication that the genius behind invention often lies in the ability to see the simple and obvious solution that others have overlooked.



(a)

Argent Co. 1400 Bway.............GRE eley 5513
Argentina Brazil & Chile Shipping Co
70 Wall.HAN over 0307
Argentine Genl Consulate. 17 Batry pl..REC tor 6948
Argentine Impt & Expt Corp. Prod Ex...BRO ad 1768
Argentine Mercantile Corp. 42 Bway.....BRO ad 5066
Argentine Naval Commission. 2 W 67..COL mbus 5623
Argentine Quebracho Co. 80 Maiden la....JOH n 1652
Argentine Railway Co. 25 Broad.......BRO ad 1383
Argentine Trading Co. 1164 Bway......MAD Sq 1871
Argeres Bros. Restrnt. 86 6th av......SPR ing 5337
Argero A. Grocer. 119 9th av........CHE lsea 6255
Arghis A. Tobacco. 74 Wall.........HAN over 6311
Argirople Theodore. Jwlr. 406 8th av..FAR ragut 9772
Argo Packing Corpn. 705 Greenwich...FAR ragut 4505
Argon Dress Co. 24 E 12..........STU yvsnt 2011
Argonaut Supply Corp. 50 Union sq..STU yvsnt 7476
Argonne Steamship Co. 17 Battery pl...REC tor 2493
Argos Ad-Art Co. 1133 Bway........FAR ragut 5986
Argosy The (A Pub). 280 Bway......WOR th 8800

(b)

Fig. 6-54. (a) Dial for letter-number dialing. (b) Sample of telephone directory set up for three-letter, four-number dialing. (Craft, Morehouse, and Charlesworth 1923, Fig. 9)

an efficient signal format for rapid transmission over trunks and actuating the display. However, the flexibility of common control was to prove of great importance in implementing a highly efficient trunking plan for large cities. The factors involved are on the whole too complex for detailed exposition in this history but one example will serve as an illustration. As important factor in trunk economy is the ability to use tandem systems for connecting offices with light traffic between them and direct trunking to connect offices with heavy traffic. With direct-control systems this capability, generally speaking, does not exist.[105] The basic reason is that one digit is required for operating each selector switch and a tandem routing will consequently require an extra digit as compared to a direct trunk route. With the indirect control used with the panel system, the digits are stored until switching is completed and translation can be arranged to set up whichever route is programmed.

While much more could be said about common control, there is no need to belabor its widespread benefits. However, we should not close this section without pointing out that there are some disadvantages also. The system is basically complex and generally requires somewhat higher levels of skill on the part of maintenance personnel although in newer systems this is more than offset by built-in testing and diagnostic features. Because common-control systems use trunks more efficiently they tend to be more susceptible to trouble under conditions of severe overload, but these problems can be mitigated

[105] Rather cumbersome expedients sometimes can be used to achieve some of the benefits but the schemes are likely to place severe restrictions on the trunking plan and result in limited economies.

by load-control features. Finally, and perhaps of most importance, common-control equipment is expensive. A substantial amount of this expensive equipment is required for even a small system; there is a sizable getting-started cost and the cost-per-line of the common equipment may be high in small, slow-growing offices. Once full size is approached, however, the cost-per-line of the central equipment becomes much less important. Thus the common-control approach is less attractive for the smaller, isolated systems and this is an important reason why step-by-step has remained the system of choice for such locations for such a long period.

4.3.3 Evolution of the "Panel" System

The need for a big-city system using mechanical switching and the probable economies resulting from its use were well known by the middle of 1905. The problems also were fully appreciated and appeared almost insurmountable. When Molina's translation and registration scheme promised to solve one of the major problems, it was decided early in 1906 to go ahead with development. Since no other way for handling the manual-machine interconnecting problem existed, the basic plan at this time, and for several years to come, was to concentrate on a semimechanical scheme using automatic distribution of calls to A operators who would set up the calls on keysets for mechanical completion through the switches that replaced the B operators. This would eliminate B operators and hopefully would give faster service than the best manual system.

The first effort went into developing a large switch since this probably would be responsible for a major part of the system cost. By March 1906, Western Electric engineers had designed two rotary switches, one a 200-point (20 around and 10 up) selector and the other a 306-point (34 around and 9 up) selector. In examining the manner in which they would be used in a system, it was apparent that conventional means for multipling these switches would involve large amounts of cabling at a prohibitive cost. It was at this time that the unique approach was adopted of setting up two separate design groups for developing competing ways of using the registration and translation system for large-city systems.[106] One of these groups continued with the rotary-type

[106] This was probably the first large use of competitive development groups within an industry to stimulate invention and improve the product. A similar competitive approach to transmitter development had been used on a small scale in the early nineties and had resulted in the solid-back transmitter (see Section 2.1.3 of Chapter 3). Perhaps this stimulated the approach used for switching development. At any rate the competition was highly successful in that the resultant system was a composite of the best features devised by each group. Competition among development organizations within the Bell System is commonplace today as witnessed by concurrent work on wire, microwave, waveguide, and satellite transmission systems as well as the intensely competitive work on analog and digital transmission and on space-division versus time-division techniques in switching. These examples represent only a few of the cases where parallel work is carried on, aimed at developing the best technology for handling all aspects of communication.

switch and devised a special flat-woven cable to provide simpler and more economical multipling. The other devised a linear switch (later known as the "panel bank") that employed punched metallic strips to serve both as contact points and connections between points, thus obviating much of the need for cable multiples.

In June of 1907 the two schemes were reviewed in an effort to decide on a system incorporating the best features of each. The time was not ripe for such a decision and parallel development of the two systems was continued during the next few years with frequent review and interchange of ideas. Desirable features of each system were incorporated in the other. From the beginning, registration and translation in a sender had been a basic part of each system and each also employed power drive for the switches, but there were many differences. The fundamental trunking plans differed and since that devised for the panel system seemed superior, it was applied to the rotary system. The group developing the rotary system had devised revertive control, multiple brushes with selective tripping, and the rotary sequence switch, all of which were adapted for use in the panel system.[107] Thus, within a short time the two systems were fundamentally the same, differing only in those particulars governed by the differences in the switching mechanism.

By the middle of 1908, developments had advanced far enough to permit the field trial of a semimechanical system in the form of a PBX. This trial system was cut into service at the Western Electric Company building at 463 West Street, New York City, on November 29, 1910, and operated with satisfactory results. At this time the development of the rotary switch mechanism was further advanced than the panel switch and it was used in the trial PBX. It also was used in the system manufactured by Western Electric in Europe about 1911 in answer to the strong competition of other mechanical systems. Largely as a result of this offering, the rotary system has received large use in Europe.[108] However, in 1911, there was still doubt that the rotary system was the best answer for the large-city conditions existing in the United States even though it appeared adequate for the smaller installations required abroad. Further studies were made that, by 1912, showed the panel system had a potential for better serving the large trunk groups associated with the rapid telephone development being experienced in the big cities of the U.S.A. Consequently, all further development effort was concentrated on this system for Bell use.

After a trial of the panel switch in the West Street PBX, it was decided in 1913 that the panel system was ready for large-scale field trials. Although full mechanical switching employing panel selectors was operating in the laboratory at the time, semiautomatic operation was selected for field trial since it could be used for a large-scale trial of the

[107] These techniques and devices are discussed in Section 4.3.4.
[108] As indicated in Chapter 2, Western Electric was active in the foreign field until 1925.

essential elements of the system in a large city without the complication and ambiguities involved in simultaneously introducing dialing by the subscriber. It would also provide more flexibility in modifying or adding to the equipment as the trial progressed. A further decision was to separate the trials of the mechanical switching and automatic call distribution in order to expedite the trials and better evaluate the benefits of the two functions.

Equipment to implement the field trials was designed and built by Western Electric to provide semimechanical switching in two Newark, New Jersey, offices. The Mulberry office with 3,640 equipped lines was placed in commercial service on January 16, 1915, and the Waverly office with 6,480 lines was cut over on June 12 of the same year. These offices worked in association with manual offices in the area and two tandem offices. Beginning in April 1917, automatic call distribution was tested in another semimechanical office in Newark, Branch Brook, which was equipped for 7,400 lines. Two months earlier, a manual office with automatic call distribution had been cut into commercial service in Wilmington, Delaware.

While the mechanical switching trials were being conducted, operation and maintenance were extensively studied. Operating loads were not entirely up to expectation but the trials in other respects were highly successful and demonstrated that panel equipment could be depended upon to give high-quality, reliable service in an area where the requirements were as severe as any in the world. At the same time as these field trials were made, work was being carried out in the laboratory on adapting the equipment for customer dialing (i.e., full mechanical switching) and extensive cost and service studies were made of manual, semimechanical, and full mechanical systems when used in cities of various sizes. The mechanical systems studied included not only panel but also the step-by-step system of the Automatic Electric Company and the Clement System manufactured by North Electric. As previously noted, patent agreements had been made during 1916 with both of these companies.

As a result of this work, much of the material was at hand in early 1917 for making definite recommendations relative to the use of machine switching in the Bell System. In some areas the labor situation was becoming desperate because the vast expansion of industry, resulting from World War I, was not only increasing labor costs but greatly reducing labor availability. As a result, a memorandum for J. J. Carty, then AT&TCo Chief Engineer, was prepared under the date of July 24, 1917, making recommendations for all-size offices as follows:[109]

[109] As we shall see, this memorandum was premature by a matter of several months since a number of problems, including the important interconnection problem in metropolitan areas, were still unresolved, but the enormous traffic and labor problems disclosed at a Bell technical conference in 1916 made immediate decisions on mechanization imperative.

Single-Office Cities—Continue with manual switchboards in most such cities, deciding after the completion of trials in progress whether or not automatic call distribution to the answering operator should be used. The chief reasons for the recommendation were that speed of service and accuracy were essentially the same for the two systems and cost favored manual at 1916 salary levels and would break even with machine at a 25-percent increase. Growth could be cheaper for manual, particularly if machine replacements elsewhere made excess manual positions available. The availability of labor was not controlling in most small cities and machine switching systems could be considered where labor was an important consideration.

Multioffice Cities Not in Large Metropolitan Areas—Adopt a full mechanical system for all growth and convert existing equipment as rapidly as practical (probably in about seven years on the average). Principal reasons for the recommendation were the improved speed of service on trunked calls, reduction in errors, improved economy, and difficulty in obtaining adequate labor supply for full manual operation. At the rate of conversion recommended, no need for dismissing employees would arise. Means for handling the addition and conversion of offices were available but some problems had to be solved before handling message-rate service on four-party lines.

Metropolitan Areas Such as New York, Philadelphia, Chicago, etc.—Use semiautomatic system for all growth and for conversion of existing manual systems, but with a 15- to 20-year period for replacement rather than the shorter period recommended for the smaller offices. A full mechanical system was considered to be more desirable for all reasons except those arising from the need for seven-digit dialing. The semimechanical system would avoid these problems and most of the central-office equipment could be utilized later for full mechanical switching if and when the customer dialing problem was solved. In the meantime, semimechanical systems would give speed of service equal to manual, improve accuracy, reduce cost, and ameliorate the labor problem to some extent by reducing the need for operators by about 30 to 40 percent and eliminating some of the more boring and burdensome parts of their work. It was believed that the latter aspects would reduce the turnover of operators.

PBX Service—Make available semimechanical private branch exchanges on a provisional basis until experience with this equipment in the hands of the public demonstrated more clearly the advisability of full mechanization.

An interesting theme running through this memorandum is the question of customer acceptability. Not so many years before, many telephone people were uncertain that the public would accept the added burden of dialing. However, experience had shown that customer sentiment ran strongly in favor of dialing, one reason being that they believed it provided much faster service since, being occupied during dialing, they tended to ignore the time required for dialing. Thus, in 1917, there was no longer a question about the acceptability of dialing but, on the contrary, there was considerable concern over the adverse reaction of those customers who, of necessity, would have to continue with what they might consider the less-modern manual system even though the actual service in many cases would be quite comparable.

The recommendations of the July memorandum were approved by top management of AT&TCo in a letter to Carty on September 13, 1917. By a strange quirk of fate, about the time that this letter was being written, Blauvelt proposed his three-letter, four-number system which would eliminate the reservations on full mechanical systems in metropolitan areas. As a consequence of further work on his proposal, plans were approved in November 1918 for designing the first metropolitan machine office for customer dialing and all subsequent work was on full mechanical systems with one exception. As early as 1916, the New York Telephone Company was in need of an additional tandem office to supplement the three manual tandems then in use. The need was acute because of the large growth of suburban traffic requiring expensive and inefficiently used trunks if handled by means of direct trunking. Tandem trunking greatly reduced the cost of outside plant but increased the error rate since the operating practice in use at the time required a repetition of the number at the tandem office. For this reason, and because of congestion in the trunk multiple of the tandem boards, it was decided in September 1917 to make the new tandem office a semimechanical installation. Work on this New York City office, known as Metropolitan Tandem and located in the Walker–Lispenard Building, was started in early 1918 and it was placed in service in the middle of 1920. The use of the semimechanical system reduced the errors, as compared to manual tandem, because it eliminated one repetition of the called number. In addition, it provided an important step toward further mechanization by furnishing a highly useful link in the interconnection of manual and machine offices. Both of these advantages will become apparent subsequently (Section 4.3.5.3).

Unfortunately, the ambitious program for introducing mechanical systems planned in 1917–18 was negated by United States entry into World War I on April 6, 1917. By the time the decision to proceed had been made, employees were beginning to join the military services and large parts of the development and manufacturing organizations were beginning to concentrate effort on war projects. On August 15, 1918, the Postmaster General (who was given control of all communication during the war and remained in charge until July 31, 1919) ordered all telephone companies "to confine extensions and betterments to imperative and unavoidable work to meet war requirements and the vital commercial needs of the country." Following the end of the war there was an acute shortage of labor, an unprecedented increase in cost, and a great deficiency of all kinds of telephone plant, which had to be supplied in the quickest way possible. It was not a favorable atmosphere for introducing new types of equipment and, as a result, the first full-mechanical panel office (the Atlantic office in Omaha, Nebraska) was not cut over to service until December 10, 1921, and the first metropolitan area office (Pennsylvania in New York City) was put in service October 14, 1922.[110] Growth of panel continued at the rate of about 100,000 lines a year through 1926, after which it increased more rapidly, reaching a rate of nearly 400,000 lines a year by 1931.

4.3.4 Implementation of the Panel System

4.3.4.1 Basic Plan. The basic plan of the panel system underwent some changes in the early years but by 1923 the basic scheme used for a full machine connection was as diagrammed in Fig. 6-55.[111] This bears some resemblance to the step-by-step system in that both are of the progressive type using a number of switches in tandem. However, they differ basically in that the panel system uses larger switches controlled by a sender. Because these switches give access to many outputs, fewer are required in tandem.

These switches will be described shortly. For an understanding of the manner in which a machine call is set up it is only necessary to know that the switches consist of hundreds of contact points in vertical columns and that the switch arm (commonly called a selector) is a rod which moves vertically and has access to any of the points along the line of motion. Progress of the call on a full machine connection is as follows:

Subscriber lines are connected to switchpoints in a line finder frame and when the receiver is removed from the switchhook, preparatory to dialing, a switch arm is actuated on this frame and moves vertically to select the caller's line. At the same time the switch arm is connected to a sender through an associated sender selector. Upon completion of these operations, which take only a short time, a "dial tone" is sent to the caller indicating that the machinery is ready to accept dial pulses. When the subscriber dials, his decimal-type pulses are registered in the sender and translated into the signals required for controlling the subsequent selectors, which operate on a non-decimal basis. As a first step in the switching process, the sender causes the particular "district selector"

[110] The Omaha office used six-digit dialing. The New York City and all subsequent offices were designed for seven-digit. In both cases the subscriber numbers used a combination of letters and numerals.

[111] Arrangements for handling calls between machine and manual offices are covered in Section 4.3.5.2.

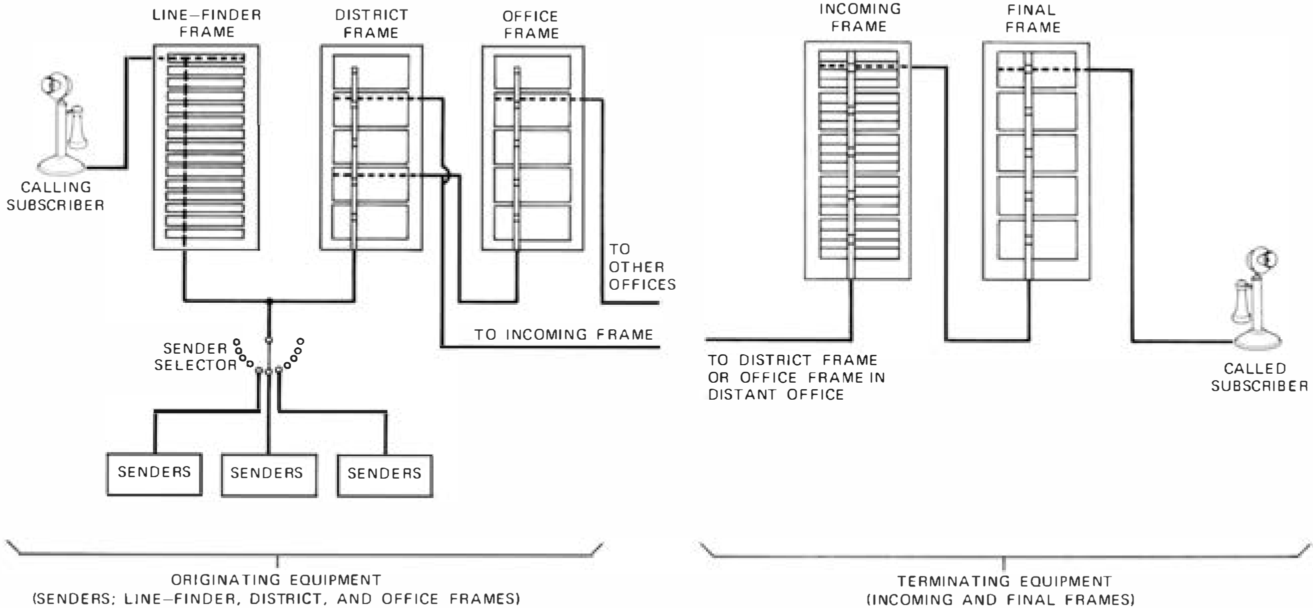


Fig. 6-55. Diagram of panel machine-switching system for full machine call. (Redrawn from Craft, Morehouse, and Charlesworth 1923, Figs. 18 and 20)

(which is permanently associated with the line-finder selector used) to select a trunk to the office desired.

If the call is for a subscriber in the same office, the trunk chosen will be an internal one terminating at an "incoming selector" frame and the sender will cause the call to be routed through the incoming selector to a final selector and thence to the particular line desired. When the connection is thus completed, audible signals will be sent back to the calling subscriber to indicate that the station is being rung or that the line is busy. In either case the sender is released from the call and becomes available for use on another.

For calls to a subscriber in another machine office, two alternatives exist for completing the call over direct trunks.[112] If the trunking network is small, the district selector will connect to an external trunk to the distant office. If, as more likely, the network is large, with many interoffice trunks, the district frame, despite its large access, will not have enough contacts to reach all offices directly and the trunk will be routed to an additional switching stage through the office frames that are provided in sufficient numbers to give access to all the outgoing trunks. The outgoing interoffice trunks (regardless of routing in the originating office) terminate at incoming frames at the distant office and connection is made to the called subscriber through these and the final frames as described previously.

Connections between the machine and manual network will be described later. For the present it suffices to say that it is unnecessary for either the caller or the manual A operator to know the type of office in which the call terminates. Regardless of the type of office in which a caller is located, the procedure is the same for calls to either type of called office. Manual B operating procedure is basically the same on calls incoming from manual and machine offices with the exception that the desired number is received aurally in the former case and by means of a visual display if from a machine office.

Machine switching does not require radical changes in private branch exchanges. The PBX is provided with dials, and calls to the central office are dialed by the attendant or by the extension user in the same way as the ordinary subscriber dials. No change is required for handling incoming calls, the PBX trunk in a machine office being treated very much in the same way as an individual subscriber line except that the final selector automatically passes over busy PBX trunks and connects to the first one found idle. At the PBX the incoming call from a machine office reaches the attendant in the usual manner and is handled by her in the same way as one from a manual office.[113]

[112] As discussed in Section 4.3.5.3, tandem trunking also could be used with the panel system.

[113] It was not until the 1960s that it became possible for callers to dial PBX extensions directly, except for panel offices.

After this broad overview of the way in which panel machine-switching functions, we can examine the equipment used and better understand the design features developed to implement the system plan.

4.3.4.2 Switching Equipment. The switching frames make up the bulk of the equipment in a panel office and represent the most obvious difference from the step-by-step system. These frames are double-sided, with the two sides essentially the same and generally much as pictured in Fig. 6-56. The five identical sections in the center are the panel banks that contain the contact points, roughly 90,000 in a frame. Below these banks is the machinery for vertically moving the selector rods, shown in front of the banks, which provide the means for making electrical connection to the bank contacts. The motion of these rods is controlled by the topmost (commutator) panel working in conjunction with the equipment to the right and left of the switch banks under direction of the sender.

The construction of a panel bank is illustrated in Fig. 6-57a. It consists of flat brass punchings about 3½ feet long with an overall width of about 1 inch. Three of these are shown in (b) of the figure, and a punching is shown in detail in (c). Each punching has lugs on both sides which form contact points. The central portion of the punching not only provides the means for holding the contacts in position but also connects or multiples them together, thus avoiding the need for the soldered connections and wiring used for this purpose in a step-by-step switch. In this particular bank, three types of punching are used, with the contact lugs slightly shifted in position to provide tip, ring, and sleeve contacts in offset vertical columns. The bank is assembled by first piling the three types of punching one above another, separated by insulating strips. Then, one hundred of these groups of three, also separated by insulation, are in turn piled above each other and bolted together to form a panel about 15 inches high. This panel provides a multiple consisting of tip, ring, and sleeve contacts, for each of 100 lines, which appear 60 times (30 on each side of the bank) in a horizontal row. Thus the five banks in a frame provide contacts for 500 lines (or trunks) each of which is accessible at 60 horizontal positions on the frame.

Access to these contacts is by means of 30 selectors on each side of the bank. Each such selector consists of a hollow metal tube carrying five sets of brushes, each set arranged for making contact with the lugs in one of the five banks. Similar contact members of the five brushes are multipled together. The brushes are normally free from contact with the bank terminals but any set can be tripped mechanically so that contact can be made. Thus, by tripping an appropriate set of brushes and moving the selector rod vertically by a distance no greater than the

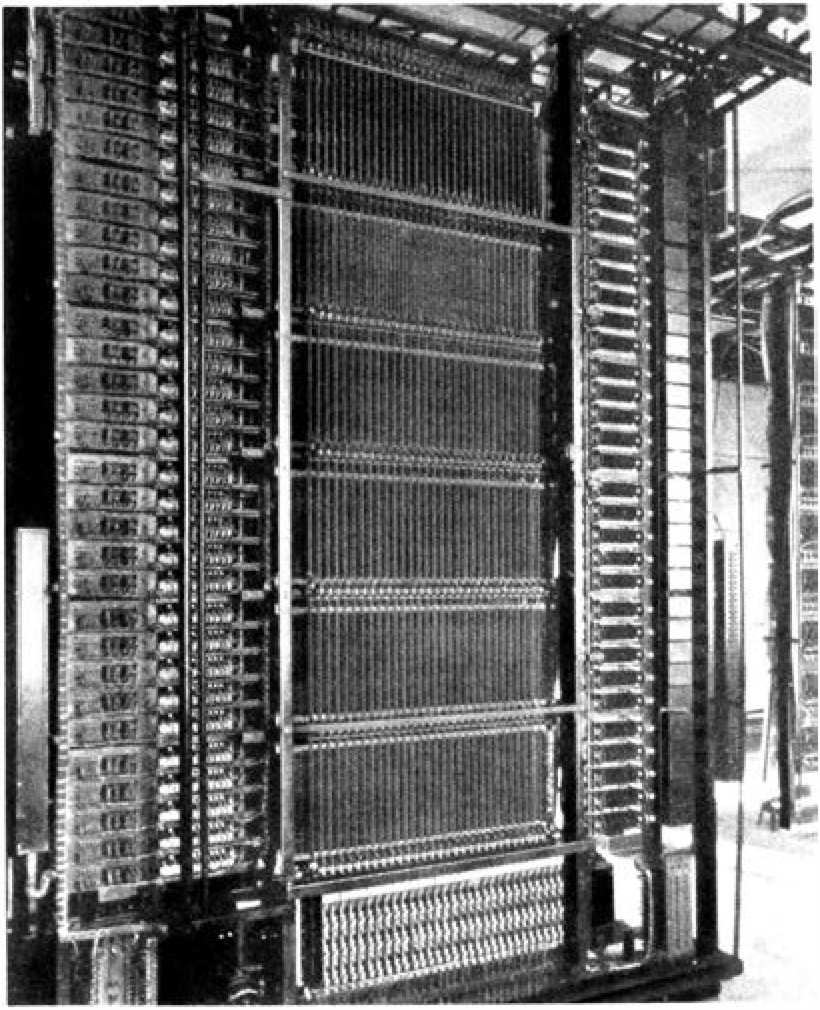


Fig. 6-56. Overall view of typical panel switching frame (opposite side of frame is identical). (Craft, Morehouse, and Charlesworth 1923, Fig. 13)

height of one bank, any one of the 500 sets of contacts can be connected to a selector.

The selectors are moved up and down by motor-driven cork rollers which can be made to engage a flat, slotted strip at the bottom of the selector rod by operating a clutch mechanism. When the appropriate

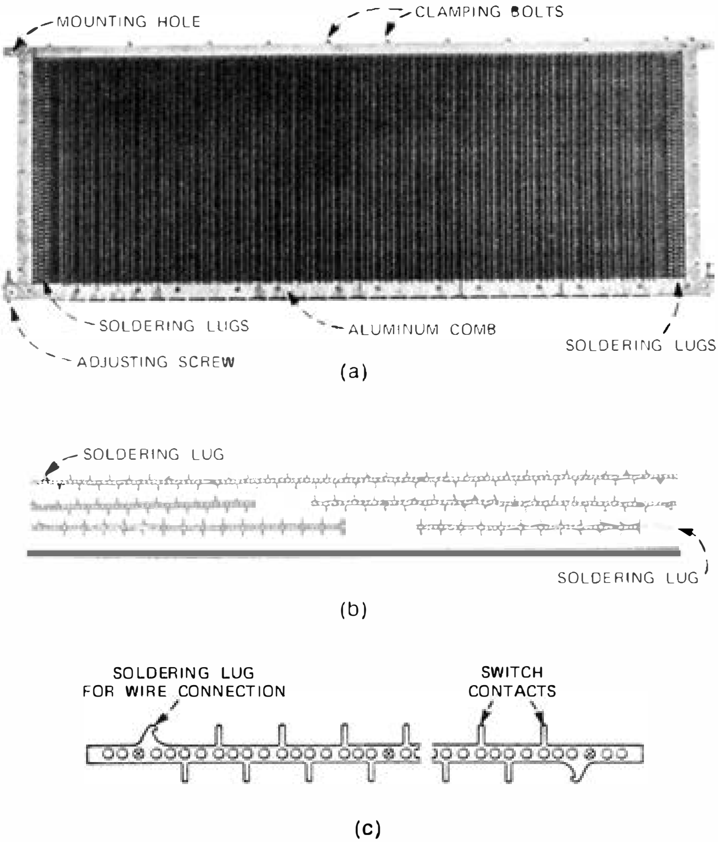


Fig. 6-57. Panel bank details: (a) assembled bank; (b) set of punchings; (c) detail of a punching.

position has been reached, a pawl engages in one of the slots and holds the selector in position until the call is completed. Figure 6-58 shows the lower bank on a frame with the selectors in different positions. Some of the latter show the slotted strip used for moving and holding them in position. The pawl and clutch mechanism is directly below the horizontal bar across the lower part of the photograph and the roller drive is out of view behind this mechanism.

Contact to the brushes is made by means of commutators, one for each selector, at the top of the frame. The commutator, shown on Fig. 6-59, consists of strips of insulating material in which are imbedded strips or segments of brass. Brushes at the top of the selector rod are connected to the bank brushes and complete connection to external

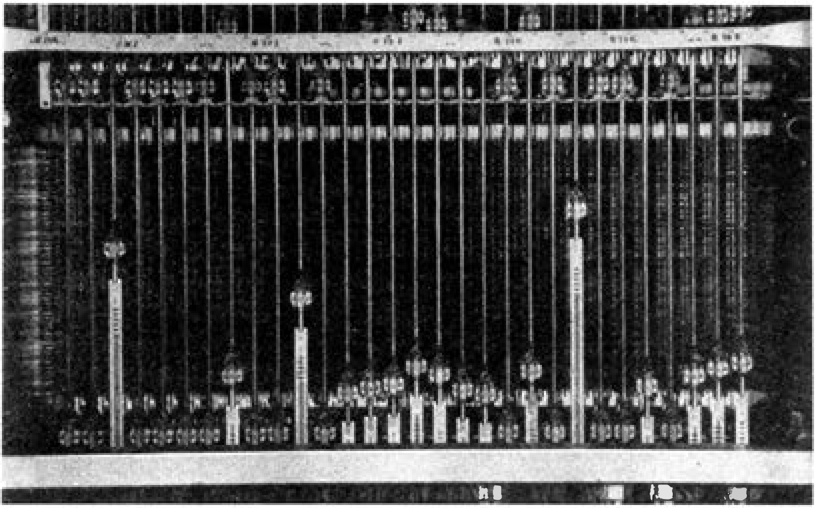


Fig. 6-58. Typical lower bank on panel switching frame. (Feder and Spencer 1962, p. 132)

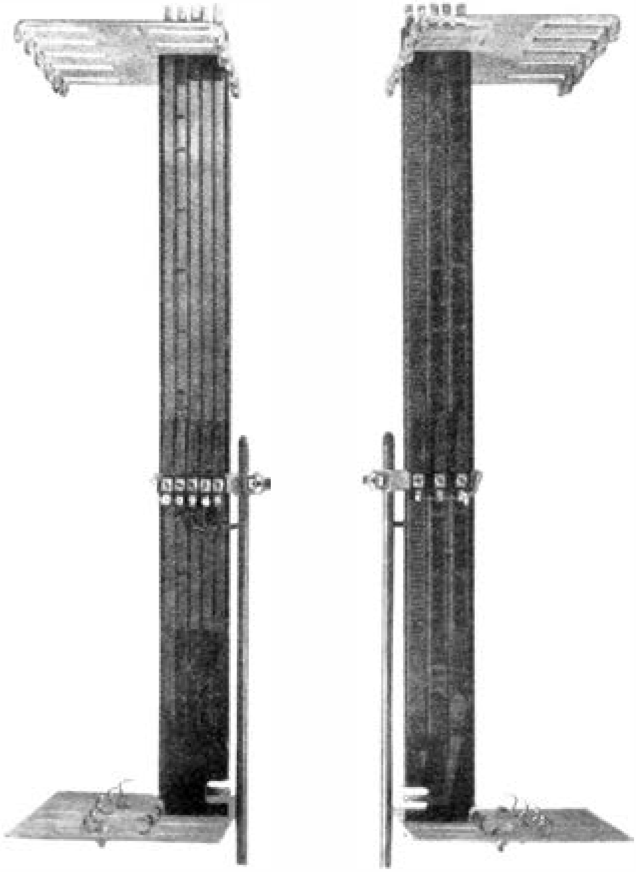


Fig. 6-59. Panel-switch commutator for connecting to and controlling vertical movement of selecting mechanism, side views. (Craft, Morehouse, and Charlesworth 1923, Fig. 11)

circuitry through the brass strips, thus avoiding long flexible connections. Perhaps a more important function is performed by similar brushes which, while the selector is in motion, pass over the small segments generating impulses which, when sent back to the sender, indicate the position of the selector so that it can be stopped at the position necessary for making the desired bank contact.

Obviously, the control of the many operations involved in operating a panel bank is highly complicated. While the sender is the master control, much of the routine operation is directed by the relays and sequence switches mounted to the right of the panel banks. The sequence switch, illustrated in Fig. 6-60, is both a unique device and an important element of the control system. It is a rotary switch made up of circular disks called cams mounted on a shaft driven by the same motor[114] used to move the selector rods, the rotation of each switch being controlled by a magnetically operated clutch. The switch is rotated through 18 discrete positions and can be stopped on any one for as long as necessary. The plates of the cams are cut to provide contact between brushes in the sequence required to carry out the functions necessary to control the panel switch during all the operations involved in setting up, carrying out, and terminating a telephone call. In modern terms, it is the per-

[114] Obviously, operation of the drive shafts for the selectors and sequence switches must be highly reliable. This is achieved by using two drive-shaft assemblies, each of which operates half of the mechanism in a frame. In addition, each drive shaft has two motors, one driven by commercial 60-Hz power for normal service and a second driven by the office storage batteries during failure of the commercial power.

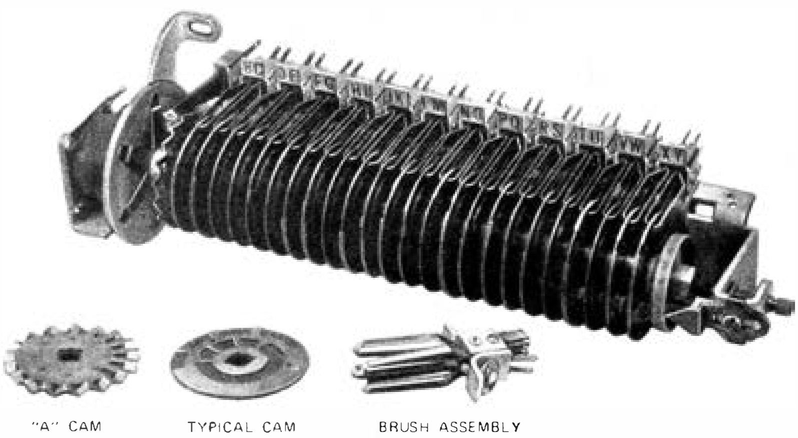


Fig. 6-60. Sequence switch. (Craft, Morehouse, and Charlesworth 1923, Fig. 12)

manent memory containing the program for controlling the routine operations of the panel switch.

The mechanisms just described, namely the panel banks, commutators, and sequence switches, formed the basis for all major selector switches in the panel office.[115] The different types of frames were much alike in appearance but obviously differed in control circuitry and minor mechanical arrangements. The major difference was in the line finder, which originally used 15 banks with 20 contact levels, giving a total capacity of 300 lines. (Later the capacity was increased to 400 lines by using ten banks of 40 levels each.) The bank connections also were split in the center horizontally, with the wiring order reversed in the two halves, so that the lines appearing at the bottom of the left side also were at the top of the right side. Both the reversal in wiring and the smaller bank were used in the line finder to reduce the amount of travel (and time) required to connect with lines. With the arrangement adopted, the maximum travel required, in most cases, was 20 levels and most connections were made with much less travel.

The other differences among the various kinds of selectors are of importance to the efficient functioning of the system but are mostly of interest only to the student of machine switching. A possible exception is in the district and incoming selectors which in addition to performing their switching function also supply battery current for the customer's transmitter and, in the case of the incoming selector, current for signaling over the trunk. Battery-supply arrangements were, until the late 1930s, very much like those used in manual offices, employing a repeating-coil supply with 24 volts on exchange calls and 48 volts on toll calls.

4.3.4.3 The Sender. While the complex switching equipment forms the largest and most obvious part of a panel office, the most significant part, as noted previously, is the common-control equipment, or sender. The sender is not of particular interest from the apparatus standpoint since it mostly used relays, sequence switches, and selectors already described[116] but it is notable for the many and complex functions which it carried out. Some of these are mentioned below and, of course, the sender controlled many additional operations too numerous to mention but necessary for the systematic switch operations required for establishing a call. The sender was required to:

[115] The device which connected the sender to the switching circuitry while the call was being established was an exception. It required very few switchpoints and consisted of a simple rotary switch.

[116] Early types of senders used a pulse-generating machine not required elsewhere in the system. It consisted of a number of rotary commutators continuously generating various pulse sequences of the type required for controlling the translation to office codes. The sender selected the pulse train needed at each stage of the operation. A single pulse machine could supply as many senders as required in a panel office.

(*i*) Accept the decimal signals from the subscriber, store them, and translate them to the non-decimal basis required for the particular path the call was to follow.
(*ii*) Control the selecting mechanism in the required time, independent of the rate at which information was received from the dial.
(*iii*) Provide flexibility in relating office designations to trunk-group locations on the frame, thus not only providing for efficient use of equipment but means for trunk rearrangement and growth.
(*iv*) Identify the type of terminating office (manual or dial) and type of trunk (direct or tandem) and arrange for sending out the appropriate type of signal.
(*v*) On calls from coin-box lines, control the call as required to assure coin deposit or transfer of the call to an operator.
(*vi*) Identify calls and establish proper connection where operator assistance was required.

Details of the means by which these functions were carried out are beyond the scope of this history but some further discussion of the special requirements resulting from integration into the manual telephone system will be given later.

We should note one change in sender design which occurred during the first decade of the panel system. Originally, a translator connected to a common pulse generator was part of the sender during the entire time required to receive pulses from the customer (or operator) and complete the connection to the point of starting the ringing or busy signals. During the latter half of the twenties, some of the translating functions of the senders were incorporated in a separate piece of equipment known as a decoder. This device performed the translation after the first three digits received from the dial (the office code), or after the zero dialed to call an operator. Like the translator it provided the coded information needed to designate the office connection in the switching equipment. Since this operation could be performed very rapidly during the early part of the call it was unnecessary to have this complex decoder device connected during the full time required to establish a call and consequently a few decoders could serve the larger number of senders required in an office. Relays (including some with as many as 50 contacts) were used for connecting decoders to the senders. The decoder arrangement eliminated the need for the pulse machine and also effected economies in both space and cost.

4.3.4.4 Signaling. Exchange plant, common-battery systems prior to panel mostly used a simple open-and-closed type of direct-current signaling over the metallic pair. In the step-by-step system a similar arrangement, using pulses of current from the customer's dial (and repeated over trunks where necessary), was used to direct the switching mechanism. In the panel system, as we have noted, the dial pulses were recorded in a sender and then were translated into a non-decimal form suitable for use with the new signaling arrangements that were introduced to better fit the panel equipment. Two types were employed; the most extensively used was known as revertive pulsing, but a second type known as PCI (Panel Call Indicator) was used on calls from a panel to a manual office.

Revertive pulsing was in a sense the reverse of the direct-control pulsing of step-by-step and was brought about by the use of power-driven, large-access switches. It was impractical to design a power drive for a high-access switch which would accurately follow a string of control pulses, without placing excessively severe requirements on synchronization between the pulsing and drive mechanisms.[117] Instead, the drive was set in motion by a signal from the sender and pulses were then sent back to the sender by the selector as it rose vertically and moved a brush over commutator segments at the top of the frame. These pulses were counted by the sender and when the selector arrived at the desired level a signal from the sender disengaged the drive and set the locking pawl to hold the selector in place. Thus, this pulsing in the backward direction achieved precise location of the selector without placing undue restrictions on the drive mechanism. It should be appreciated that this is a highly simplified description of the revertive pulsing system and obviously its implementation involved complex consideration of the character and timing of the pulses in order to meet the requirements of transmission over metallic trunks (and later over carrier telephone systems).

As noted previously, calls from a machine office to a manual office required the transmission of information over a trunk to the manual office where it could be used to provide a visual display of the called number at the manual B board. In the early trials of this display system, revertive signaling was used for controlling displays, but a considerable amount of information was required for this purpose and in order to accomplish signaling accurately and rapidly a coded signal was later devised. The code, transmitted by a sequence switch in the sender, is essentially a binary system employing four bits per decimal digit. Four signaling states are used for designating digits: current off, light positive current, and light and heavy negative current. Combinations of these pulses indicate the digits to be displayed (and the party letter if required) by actuating a rather simple relay arrangement at the manual office. A fifth state using a pulse of heavy positive current is transmitted to

[117] The problem of obtaining equal drive speed for the various pieces of equipment was not too difficult but in addition the clutches would have to be free from slipping and the time for their engagement and release would have to be closely controlled.

indicate the end of digit transmittal. Later, this coded signaling system, or PCI, also was used for other signaling applications.

4.3.4.5 Maintenance. The maintenance of a manual office was a relatively simple matter. An operator was connected to the circuit used by the customer, while the call was being established. She could easily detect many central-office troubles or incipient troubles and report them to the repair service and, if a malfunction was encountered, she could substitute an alternate circuit so that the customer's service was not seriously affected. The test-desk man worked with the craftsman in correcting troubles within the office and also were concerned with locating outside plant troubles. This work could be done with relatively simple direct-current equipment such as a high-impedance voltmeter or possibly a Wheatstone bridge.

It was obvious that the highly complicated machine office required a completely different approach. An operator was no longer available for detecting problems on calls and other means would have to be used to detect troubles, or preferably incipient troubles, before they became serious. In addition, the complicated circuitry and many components provided a much greater potential for trouble than the simpler manual office. The solution for this problem was based largely on the use of preventive measures and several approaches were followed.

As a preliminary step in carrying out this philosophy, all new pieces of apparatus designed for the machine system were rigorously tested during development to assure long life and operation over a wide range of conditions.

Office test gear included the test-board equipment used for maintaining the outside plant; in addition, means were provided for timing the dial pulses from the subscriber's station to detect a faulty dial and assist a repairman in adjusting it where necessary.

Furthermore, a large amount of special test gear was developed to provide arrangements for automatically subjecting the various office circuits to routine tests designed to assure that each was capable of operating under more severe conditions than those to be met in practice. In cases where gradual deterioration occurred, the severe testing conditions would detect incipient failure long before service was affected. A test frame for controlling this routine procedure is shown in Fig. 6-61. When a circuit failed to meet test requirements, an alarm was sounded and the circuit in trouble was indicated by an array of lamps on the test frame. The craftsman in charge could then correct the trouble or if necessary remove the circuit from service until repair had been made. Portable test gear, shown in Fig. 6-62, was provided for on-the-spot observing and checking of frames during operation.

The sender, being both the most critical element in the office and the most complicated, received special attention. In addition to an automatic

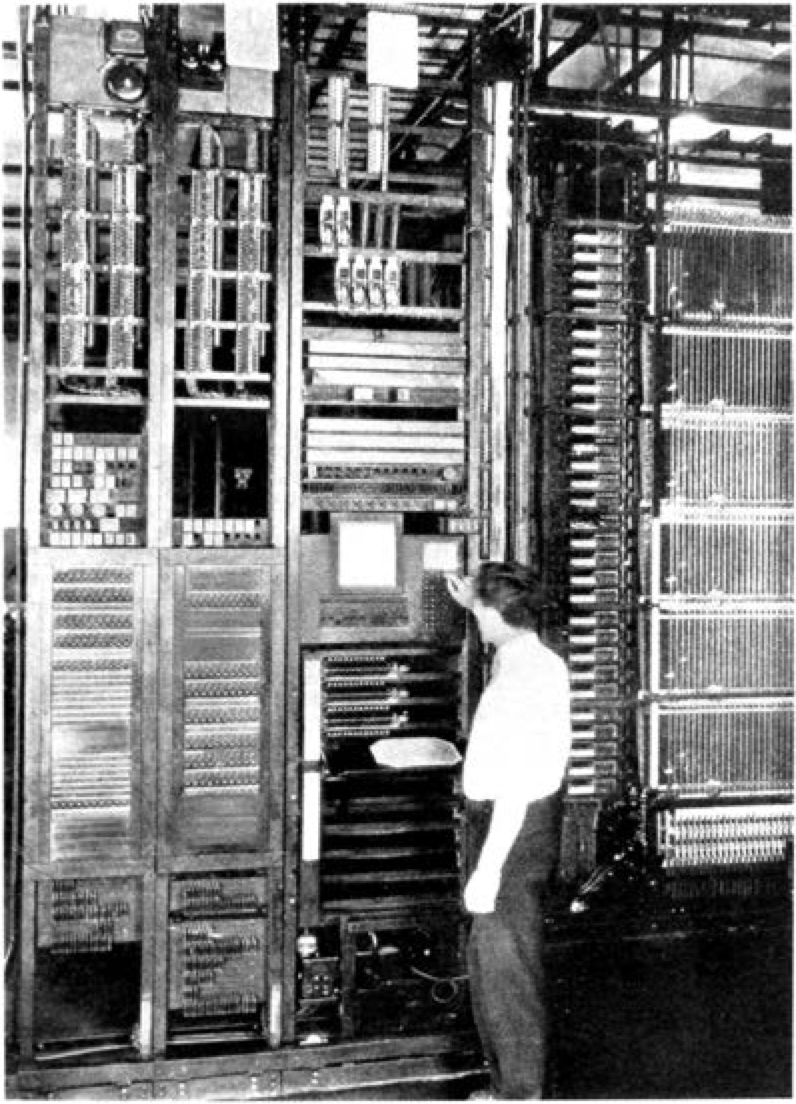


Fig. 6-61. Incoming-selector routine test frame and floor alarm board.

test circuit, a sender monitor was provided to time critical operations involved in establishing a call. If unreasonable sender holding times occurred, an alarm sounded and the monitor operator made connection to the talking circuit in order to determine the cause. If the fault was due to the subscriber (through dialing too few digits or similar irregularity)

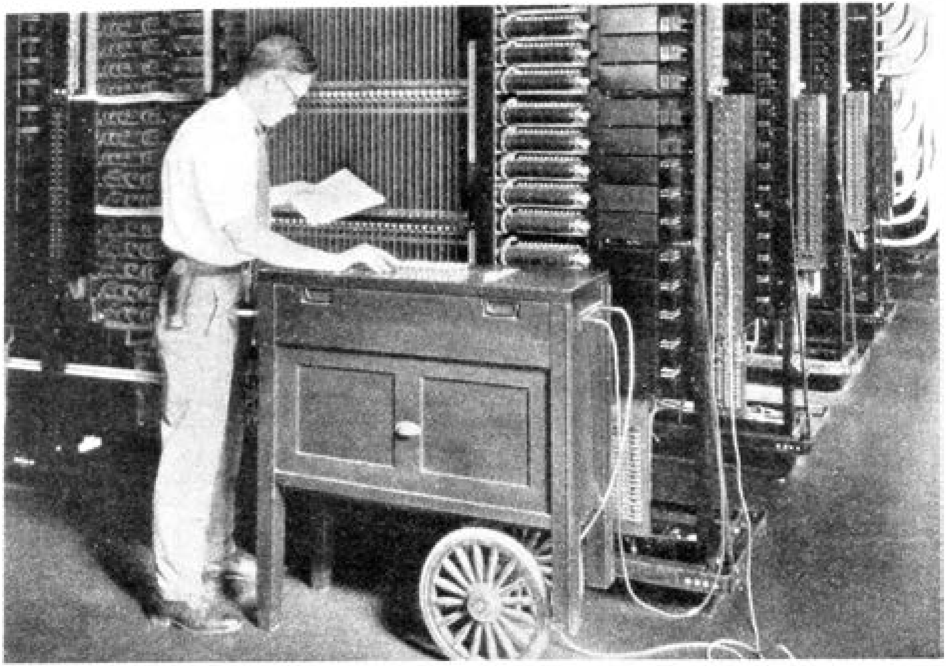


Fig. 6-62. Portable selector test set.

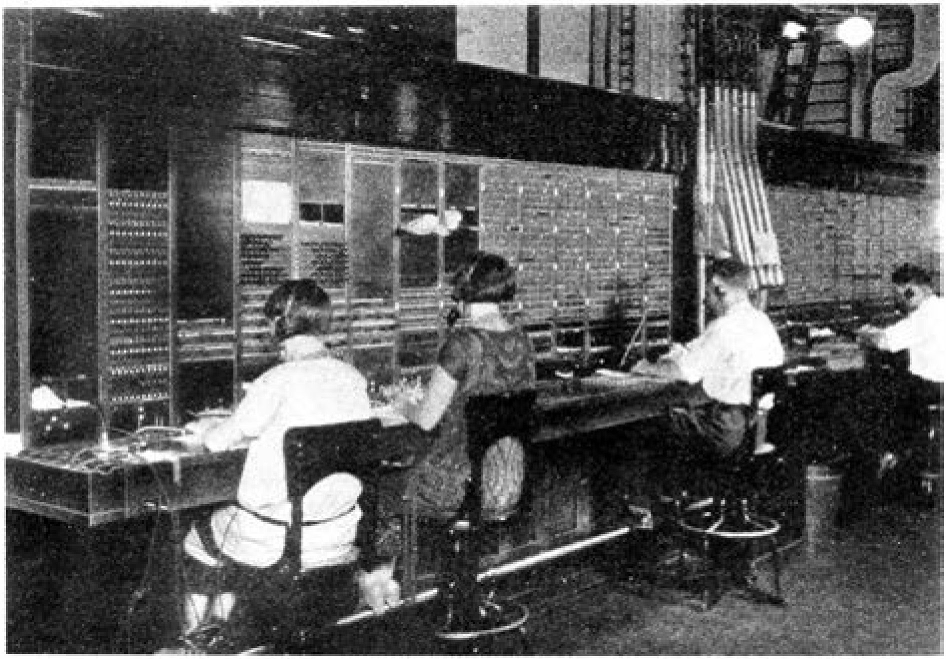


Fig. 6-63. The sender monitor.

appropriate instruction was provided. If the monitor operator found a sender with a trouble condition, she released it from the line and removed it from service until repaired. Some common subscriber troubles, such as receiver off-hook without dialing, caused a special signal after a reasonable time and the sender was automatically removed from the line and made available for other calls. The sender-monitor positions that controlled these operations are shown in Fig. 6-63.

4.3.5 Integration Into Existing Plant

We have already pointed out many of the difficulties involved in developing a machine switching system that would supply all the required services and work smoothly with a large and sophisticated manual system that had evolved over a period of some 40 years. In this section we shall merely review briefly some of the solutions developed for integrating machine and manual systems with minimum effect on the people using them.

The reader should bear in mind that, during the initial years, the broad Bell System plan for adopting machine switching involved gradual introduction by using it primarily for growth and for manual-office replacement made necessary by obsolescence. In all cases it was to be introduced in units capable of expansion to full-size offices handling 10,000 lines, although not necessarily equipped for the ultimate capacity from the beginning. In large cities and metropolitan areas with rapid growth, panel-type switching was employed to add a series of offices, each working at close to full capacity from its cut-over. Often a number of these offices would, in the long run, be located in a single building. For the smaller cities, machine switching also was installed as distinct office units which either replaced obsolete manual offices or provided new office growth. Here, step-by-step was the more commonly used system (except in the fringe of metropolitan areas) since it was more economical than panel in small offices, particularly where the growth rate was slow.[118] By following this general plan of addition or conversion by

[118] There were a few exceptions to this general plan of using panel for large cities and step-by-step in small, the most notable being Los Angeles. This area had a very complex growth involving many companies. For many years the city had dual services but by 1916 the fallacy of this form of competition had been demonstrated and the public decided (by popular vote) in favor of consolidation. The consolidated company, the Southern California Telephone Company, a subsidiary of the Pacific Telephone and Telegraph Company, began operation of unified plan on January 1, 1918. Included were some 35,000 dial stations served by step-by-step offices, one of the largest installations in the country at that time. Because of this large number in the city, and also the use of step-by-step systems in the many independent companies outside the city limits, it was decided to use step-by-step for further expansion of automatic plant in Los Angeles. The many problems of developing a large metropolitan area without common control will not be described here but furnish an excellent example of the wisdom of the engineers who, working in the early part of the twentieth century, decided that direct control was an unacceptable basis for a large machine system.

distinct office units, the cost was minimized since little replacement of modern equipment was required and the problems of training employees and subscribers involved in a changeover were minimized.

4.3.5.1 Supplementary Manual Services. A number of services furnished in manual offices could not be provided on a full machine basis and it was decided at an early date that a large number of operators would have to be retained to supplement the automatic equipment. The basic procedure for providing many services, such as directory assistance, remained largely unchanged by machine switching. The subscriber was given access to the operator providing the service by dialing a special code (usually consisting of three digits) instead of making an oral request but otherwise there were no changes. Similarly, long-distance telephone calls which required special routing information, timing, and charging (see Section 5.2) were handled by an operator also reached by special dialing codes.

From the subscriber's standpoint the most important operator contact was with the "special service operator." This operator was also known as the "zero operator" since this was the code dialed by the subscriber in reaching her. An important part of this operator's duties was to handle emergency calls and to assist the subscriber whenever dialing did not or could not provide the desired result. She occupied what was essentially an A operator's position with the additional equipment necessary to carry out her many duties. Some of these were:

Subscriber Assistance

Dialing zero connected the subscriber to a jack, with lamp signal, in front of the operator who answered with a plug-in cord. She either gave proper instructions to the subscriber or completed the call as required. The assistance function was of particular value in providing fast service in case of fire or medical emergency since this was one of the calls previously handled directly by the operator. Her assistance was also valuable to subscribers who through physical handicaps were unable to dial complete seven-digit calls. During the period after a machine cut-over the zero operator served a valuable educational function by instructing those who had difficulty in mastering the dialing operation.

Unused Office Codes and Out-of-Service Numbers

If a subscriber, through an error, dialed a code not assigned to a working office, a special tone was returned as an indication and the zero operator was not involved unless the subscriber had continued trouble of this type. However, lines in the office not in use (either because they were unassigned or temporarily out of service) were connected to a series of jacks which could be answered by the operator, who then gave the correct information for redialing or completed the call as appropriate.

Toll Calls

Long-haul toll calls were handled through the toll board operator as described in Section 5.2.2 and she was ordinarily reached by dialing a special code. Short-haul (A-B) toll calls were, however, handled by the zero operator. Ordinarily, she was reached by dialing zero but if the subscriber was unaware that he was making a toll call and dialed the listed number, the mechanism directed the call to the assistance operator. In either case, this operator established the call, did the timing, and carried out all the duties of an A-B toll operator.

A typical position for an assistance operator (or panel A board) is shown in Fig. 6-64. This board includes a multiple for checking the originating number given by the caller on a toll call, and also a keyset for rapid dialing. Special numbers for emergencies and routing instructions were displayed on cards carried in the bulletin holder on the keyshelf and sometimes on special racks above the checking multiple. A rotary dial was used in place of the keyset when only a small percentage of the calls answered were completed in the local area. Such a board is shown in Fig. 6-65.

4.3.5.2 Machine-Manual Interconnection. The basic objective underlying the means for handling calls between machine and manual offices was that the subscriber would not have to know the type of office he was calling and could follow the same procedure for making a call regardless of office destination. In addition, operators, so far as practical, carried out their procedures in the same manner regardless of the type of office that the subscriber called.

Calls from machine to manual offices were handled by equipping the latter with special B positions which differed from the usual B board in having a "call indicator" which visually displayed the number being called. These devices received field trials as early as 1910 (in Cortlandt office, New York City) and underwent several evolutionary stages. A typical call indicator used with panel systems is shown in Fig. 6-66. Some manual offices were capable of handling 10,500 lines and four party letters. Provision was made therefore for displaying a fifth digit (0 or 1) and the party letter but on most calls only four digits were required. Otherwise, operation of the B board was much the same as on a manual call using straightforward trunking. The machine equipment connected the call to an idle plug-ended trunk and the B operator was notified by

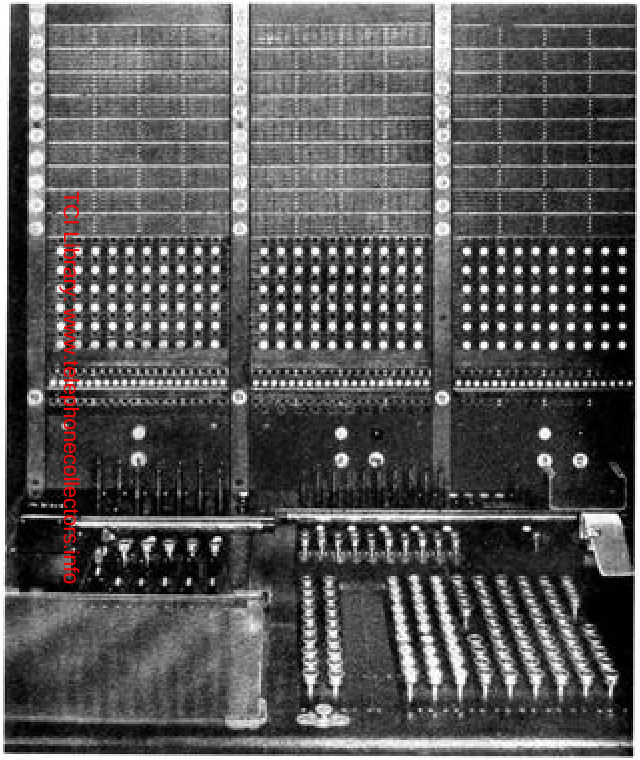


Fig. 6-64. Semimechanical A switchboard, combination position.

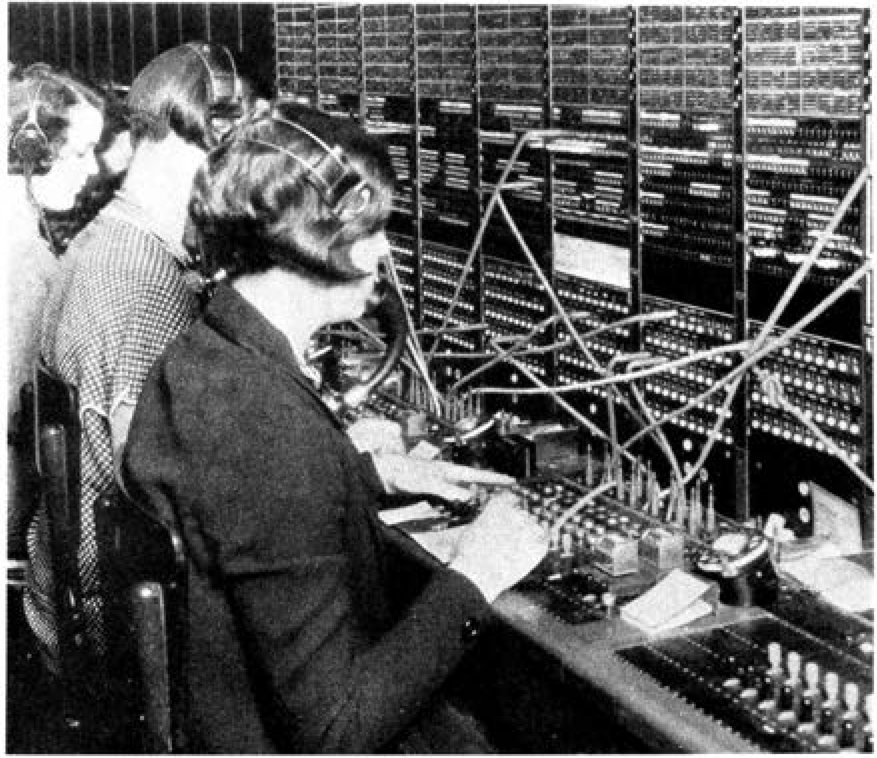


Fig. 6-65. Dialing A board.

lighting the associated lamp. She then pressed a key to actuate the display and completed the call by inserting the plug in the proper jack of the multiple after making a busy test. The plug would, of course, be inserted in a busy-back jack if the line was busy. The equipment used is diagramed in Fig. 6-67. All of the mechanical functions were under control of the sender including the trunk signaling and call indicator operation. Signaling in panel areas employed the PCI code system mentioned previously, except during the early trial period.[119]

Calls from manual to machine offices could have been handled in either of two ways. Under some conditions (particularly after the manual offices were in the minority) it would have been economical to equip the manual A board with a dial and have this operator make the conversion between spoken and electrical representation of digits. However, this would have required a procedure on the part of the A operator which depended on the call destination and it was decided to avoid this complication by making the conversion at the machine office.[120] This was done by the equipment shown in Fig. 6-68 and involved establishing a B board at the machine office with an operator communicating vocally with the manual A-operator in the same manner as during a call between manual offices, call-circuit signaling being used initially and straightforward trunking later as that scheme was introduced. The machine B-operator, instead of completing the call by means of cords plugged into the multiple, established connection by means of keys for assigning trunks and a keyset for "dialing" the number. The machine operations carrying out the connection were under control of the office sender and completed through the machine equipment. The so-called "cordless B-boards" initially used are shown in Fig. 6-69.

In the latter part of the twenties, a "call distributing B-board" was introduced in panel offices along with straightforward trunking. With this board, incoming calls were distributed and connected automatically to operators in rotation. If all operators were busy, the call was held momentarily until one was free. Connection to an idle operator sent a signal to pass the number orally. The B operator then punched the number on a keyset, the connection was set up automatically, and the operator was released to handle another call. This resulted in much faster handling of calls and the simplified B-board shown in Fig. 6-70.

*4.3.5.3 **Tandem Systems.*** Tandem switching had been introduced in the metropolitan areas as a means for providing more economical trunk

[119] Step-by-step call indicators used the slower dial-pulsing.

[120] For reasons of economy, some non-Bell administrations, particularly outside the U.S., used A-board dialing. This was not customary in the Bell System even when the manual offices were a minority since at this time means had been devised for achieving economical interconnection by simplified, centralized conversion between voice and machine signals.

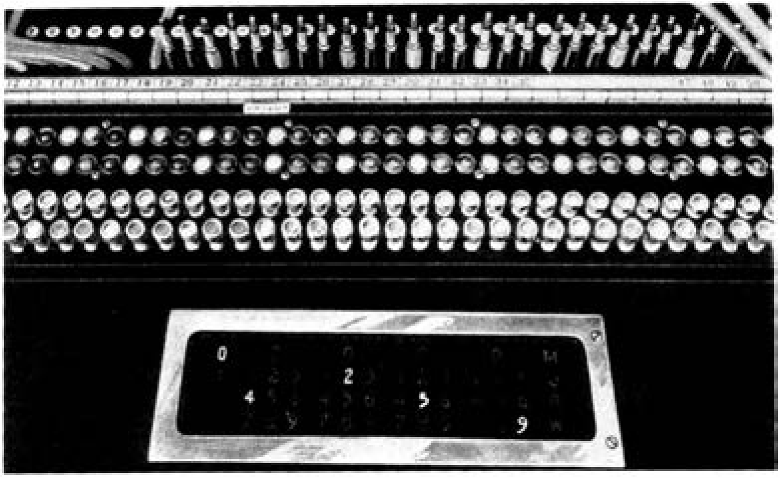


Fig. 6-66. Call indicator at an incoming-trunk position in a manual office. (Craft, Morehouse, and Charlesworth 1923, Fig. 23)

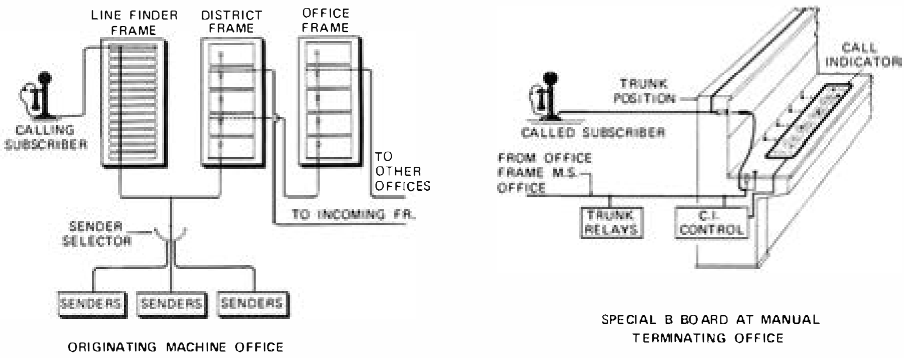


Fig. 6-67. Diagram of a connection from a machine to a manual office. (Redrawn from Craft, Morehouse, and Charlesworth 1923, Figs. 18 and 21)

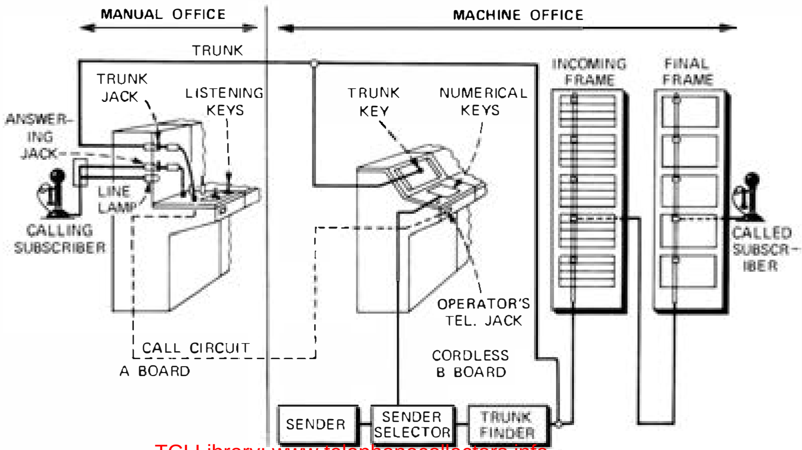


Fig. 6-68. Diagram of a connection from a manual to a machine office. (Redrawn from Craft, Morehouse, and Charlesworth 1923, Fig. 24)

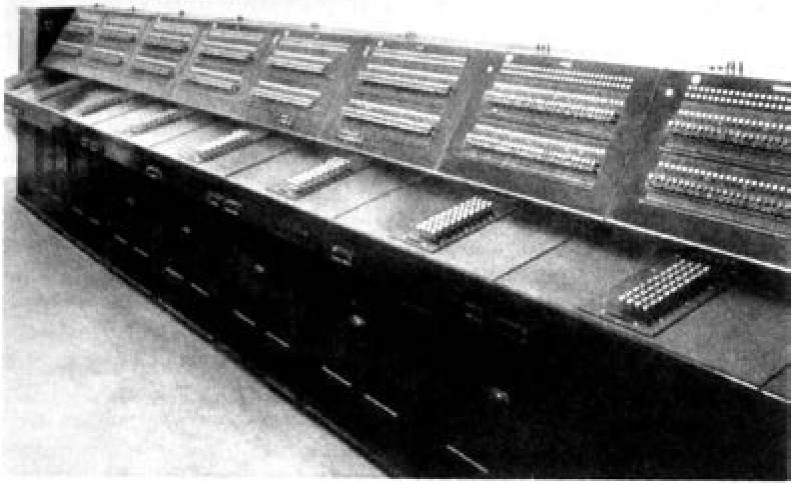


Fig. 6-69. Cordless B-positions in machine switching office. (Craft, Morehouse, and Charlesworth 1923, Fig. 25)

plant before machine systems were introduced. Obviously, machine systems had to be devised to take advantage of this important trunking principle. Up to about 1930, there were several kinds of tandem offices using panel switching which functioned roughly as follows:

Semimechanical Metropolitan Tandem

As indicated previously (Section 4.3.3), this system, put into service in 1920, was the first commercial application of the panel system and accordingly it was designed primarily to interconnect local manual

Fig. 6-70. Call distributing B-board.

offices but also had the capability of working with machine offices when they were introduced. When all local offices were of the manual type, the tandem operator received the called number from the A operator orally over a call circuit. She entered the equivalent of the office code and subscriber's number on a keyset which operated machine equipment (district and office frames under control of the tandem office sender) to extend the call from the incoming tandem trunk to the appropriate office where the number was displayed at a call indicator B-board. At this point, the operator completed the call through the manual office multiple.

Obviously, rather simple changes adapted this type of office for use with panel-type local offices. When a machine office originated a call, seven-digit signals from the originating office were passed on to a different tandem sender where they were re-registered and connections were made, under control of the tandem sender, to the called number in the usual way. If the terminating office was manual, the scheme described above (using call indicators) was used. If the call number was in a machine office, completion was on a mechanical basis through incoming and final frames.

Office Sender Tandem (2-Wire-Office Tandem Centers)

As machine offices were introduced, it became possible to use a simplified tandem office under some circumstances when the call originated in a panel office. Basically, this consisted of installing office frames in a tandem office which were connected by 2-wire trunks to the originating office. Calls for such a tandem office were registered and translated in the sender of the originating office which directed the connection through the local district frame where the trunk to the tandem office was selected. Signals from this trunk, of the revertive type (see Section 4.3.4.4), directed the call through the tandem office-frame and selected a trunk to the terminating office where it was completed in the usual way, i.e., either by panel-office machine methods or through the multiple, under direction of a call indicator, if the office was manual.

This scheme had certain limitations since the office sender could control only one stage of switching at the tandem office and the size of the system was severely limited.

Panel Sender Tandem

This system was adopted in early 1930 to eliminate the limitations of the office sender tandem. It was, in a sense, a refinement of the metropolitan tandem described previously. It was a self-contained switching system complete with senders capable of recording and translating seven-digit numbers. It could receive either oral signals from a manual office (recording them through a keyset) or the called office code and number from a machine system in the form of PCI signals (described in Section 4.3.4.4). All control beyond the tandem office was exercised by the tandem sender and it could, if desirable, direct the call to a second tandem office as well as completing directly through a local office in any of the standard methods.

It is interesting to note that the tandem office used, to a considerable extent, equipment such as office selectors which otherwise would have been installed at local offices. Since tandem offices were concentrating points where equipment received high usage, a degree of economy was achieved by this concentration and furthered the use of the tandem principle. This was particularly true in the case of office sender tandems. These required no tandem senders and used office frames and the trunks associated therewith which otherwise would have been required, in greater number, at the widely dispersed local offices feeding into the tandem center. The economy so achieved made them highly desirable in spite of the reduced flexibility as compared to panel sender tandem.

4.4 Developments Abroad

While this history is largely concerned with the evolution of communication technology in the Bell System, completeness requires the recognition of interaction with technological growth outside the Bell System. During the 1910–25 period there was much interest in automation of telephony in a number of European countries and the influence of American technology was strong. The Automatic Electric Company actively marketed step-by-step systems abroad and they were well received since they met the needs of European administrations with low telephone development. As noted previously, Western Electric, which at this time was very active abroad, felt the need for an automatic system for sale in Europe and about 1911 introduced the rotary system which, up to that time, had been developed by the Bell System in parallel with the panel system and differed basically from the latter in using selecting switches employing rotary rather than linear motion. The rotary system, like panel, differed basically from step-by-step in that it used power-driven sequence and selecting switches controlled by revertive pulsing and also employed registration and translation. The rotary system, which had been developed by a group led by F. R. McBerty, used somewhat smaller access switches than panel and, while they were judged to be less favorable for Bell System use, they were fully adequate for European service and in 1911 had received more testing than panel switches. Accordingly, McBerty was sent to Europe to establish manufacturing facilities and carry on further development work there.

The first rotary exchange in Europe was a small (800 lines) fully automatic exchange in Darlington, England, put in service during October 1914. The larger Dudley central office was opened in Birmingham in September 1915. In the meantime semiautomatic installations were made (1915) in Sweden and France. Prior to World War I, the Western Electric works at North Woolwich, England, manufactured these systems, but after the war the rehabilitated works at Antwerp, Belgium, became the main center for development and production of automatic systems with Woolwich concentrating on manual apparatus and cable. Rotary systems manufactured in Antwerp were installed throughout the world, and amounted to about nine million lines in 1973.

Although early rotary systems were installed in England, there was little subsequent use there. The story is interesting. After World War I, the British Post Office[121] was greatly concerned with automation since they had many outmoded manual offices needing replacement. They believed that automation was the way of the future and were reluctant to defer it by using long-life manual replacements. At the same time they, like the Bell System, were much concerned with the "Big City Problem." Although many ideas were advanced, none satisfactorily met the needs of London, which required great flexibility, since physical-plant limitations seemed to dictate the use of a large number of central offices of various sizes ranging from a few thousand to many thousands of lines. In 1919, Mr. McQuarrie of the Western Electric Company visited Colonel Purves, then the Engineer-in-Chief of the Post Office, and outlined in great detail the manner in which the Bell System planned to meet a similar problem through the use of the panel system, incorporating common control, together with Blauvelt's numbering plan. Purves saw at once that this approach would solve all of London's problems and preliminary arrangements were made for its adoption. However, Western's proposal to manufacture the equipment in Antwerp had political repercussions that could not be ignored. When the Automatic Telephone Manufacturing Company offered to manufacture a system in Britain that would meet the Post Office's needs, their proposal was examined carefully and adopted in 1922, but it was not until 1927 that the first London installation took place. The proposal adopted by the Post Office was that the Automatic Telephone Manufacturing Company would develop a "director," incorporating the registration and translation principles developed by Bell, which could be used with Strowger switches and the Blauvelt numbering system. The lack of large-access switches was to be overcome as far as practical by the use of the graded

[121] The telephone and telegraph branches of the Post Office were basically operating and engineering organizations with essentially all the development and manufacturing being carried out by private industry. The technical staff, headed by the Engineer-in-Chief, consisted largely of civil servants but overall administration of the Post Office was in the hands of political appointees.

multiple which Gray of AT&TCo had proposed in 1905. Thus, while the panel system was not adopted, as originally proposed, its basic concepts were used with step-by-step switches in the solution to the "Big City Problem" ultimately adopted in Great Britain.[122]

Another early Bell invention also had an interesting international history. As noted previously, J. N. Reynolds invented a crossbar switch in 1913 as part of the so-called "coordinate system" which did not go beyond the experimental stage at that time. In 1919, Betulander in Sweden invented a small crossbar switch which was the predecessor of larger switches of that type put into production in early 1920. These switches were used for some 20 years in Sweden in a direct-control system much like step-by-step. In 1930, when the Bell System was considering the design of a second-generation common-control system, W. H. Mathies of Bell Telephone Laboratories visited Sweden and was impressed with the potentialities of the crossbar switches he saw there. Samples were ordered, and after improvement and adaptation for common control, this switch was used as the selector mechanism for the No. 1 crossbar system that became the large-city successor to panel.

4.5 Summary

4.5.1 Fifty Years of Development

It may come as a surprise to some of our readers that work on mechanical switching started, and was pursued diligently, shortly after the invention of the telephone. This may seem particularly unusual because manual switching appears to be a relatively simple process and also offers opportunities for personal services not possible with a machine. It is always hazardous to offer explanations for historical events but perhaps we will not be far wrong if two factors of possible significance are suggested.

A major factor was probably the status of technology. The great inventions which led to the Industrial Revolution occurred roughly in the century preceding the telephone and, when the invention of the latter

[122] The account is given in considerable detail in the book, *The Story of the Telephone* (1947), by J. H. Robertson. Robertson was a professional journalist engaged by the Telecommunication Engineering and Manufacturing Association to prepare this excellent history of the telecommunications industry of Britain. There is no doubt that the Post Office technical staff believed that the "director" system provided the best available solution to the problems existing in Britain where the telephone development was much smaller than in the U.S. and was based on the use of smaller central offices. These considerations changed the emphasis on economic factors and also placed a premium on the use of a system employing the same type of switches throughout the country. However, the admiration of Purves for the panel system concepts is evidenced by his July 1925 I.E.E. paper in which he says: "The panel system in its entirety includes a multitude of electrical circuits of great complexity. It is the product of many wonderful brains and its rapid development and installation are the outcome of probably the greatest engineering effort so far made in any field of industrial endeavor."

occurred, the weaving of cloth, the refining of iron, the generation of mechanical energy, and the transportation of people and goods had all been transformed by the invention of fundamental mechanical devices. The revolution of industry was in full swing and there was boundless confidence in the power of machines. It was an article of faith that mechanization, which had brought such miracles in the past, would continue to do so in new fields in the future. To a considerable extent it must have been some such faith that drove people in the pursuit of switching mechanization even when it was difficult to justify on the basis of economies and services.

Another factor was that manual switching in the early days of telephony was not the simple operation that it appears in retrospect to those few readers who have used it. Early manual switching was a highly complicated affair involving much work on the part of both the operator and user. As technology advanced, manual switching procedures were simplified and many eliminated by performing the operations automatically. Thus, manual switching was slowly automated until only a few functions were being performed manually. The technological advances that made these changes possible also facilitated the improvement of full mechanical systems. For a long time, manual systems were always a few steps in advance but the gap was small enough to encourage the mechanical enthusiasts to press forward.

After a number of early, abortive attempts at mechanization, two main lines of development evolved, one sponsored by the Automatic Electric Company and the other by the Bell System. The Automatic Electric Company, using the direct-control principle and a modification of the Strowger switch, developed what is now known as the step-by-step system. In its earlier form this system did not provide many of the service features of contemporary manual systems but by 1910, or shortly before, the services provided by manual systems in the exchange plant were largely duplicated by the step-by-step system. One exception was that machine systems placed the extra burden of dialing on the customer. To the surprise of many, this burden was accepted willingly by the public[123] and after about 1910 the choice between manual and step-by-step systems in the smaller cities was largely a matter of relative cost, i.e., balancing the saving in operator salaries against the added cost of machine switching equipment. An exception was that the step-

[123] Even today the widespread acceptance of dialing seems somewhat surprising. Rather careful tests made about 1960 showed that most people would choose a system in which the number was passed orally in preference to one requiring dialing. Obviously, the very fortunate acceptance of dialing (which is basic to the modern telephone) is a very complicated matter. Some people undoubtedly were influenced by the feeling that there was greater privacy in a machine system and others felt that it was more rapid and more free from errors. Still others were influenced by the novelty and were anxious to adopt dialing because it represented the "modern" approach. Perhaps we might summarize the situation by saying that the general public too was imbued with the faith in machinery and the belief in the fallibility of man that spurred much of the Industrial Revolution.

by-step systems at this time were not very suitable for meeting the requirements of large metropolitan areas. Thus it turned out that technically the most favorable field for the step-by-step system was exchange switching in the small city with at most a few office units (each handling 10,000 or fewer lines). Economically, step-by-step could best compete with manual in those areas where the addition or replacement of complete office units was required by growth or obsolesence of manual equipment.

In the years before World War I, there were not many Bell System areas in which step-by-step was the system of choice. Small cities were mostly equipped with relatively modern, common-battery, manual systems and growth could be handled more economically by adding manual positions than by intermixing automatic with the existing manual. Large metropolitan areas were another matter. Growth here would be rapid and require the frequent addition of full office units. Furthermore, a shortage of suitable operators could be foreseen in these locations. Accordingly, after some early and unsuccessful attempts to employ machine switching in very small towns, the Bell System concentrated development effort on automation toward the solution of large city and metropolitan area problems.

The step-by-step system with its small access switch was inefficient for large areas and the Bell System developed two large access switches, the rotary and the panel types. Bell also devised means for controlling these switches based on signal registration and translation in common-control equipment. These concepts led to a system, both efficient and highly flexible, which solved the unusual problems associated with service in big cities.

At first, the problems of interconnection with manual offices, and the belief that the required seven-digit dialing would be unacceptable to users, suggested the use of a semimechanical system in which the A operator, instead of the customer, did the dialing. However, these problems were solved and plans were approved in November 1918 for designing a metropolitan area machine office for customer dialing. World War I deferred realization of these plans and the first common-control office using the large access, panel-type switch was not put into service until 1921. As soon as adequate production facilities could be built up, this "panel system" provided most of the growth and replacement offices in large cities and metropolitan areas under Bell operation until the introduction of crossbar systems in the mid-thirties. Panel equipment was manufactured for additions for some time after, and at its peak use in 1956 there were about seven-million Bell System stations served by panel systems.[124]

[124] In terms of usage, relative to other types of switching, the peak development was achieved in 1940 when 28 percent of the Bell System telephones, or about five million, used panel.

Although panel was the preferred system in large cities, it was not as economical as step-by-step in small cities, particularly those with low growth. For these, Bell adopted the step-by-step system, the first use being in cities where the plant was acquired by purchase of independent properties. However, in 1917, step-by-step equipment was purchased for installation in Norfolk, Virginia, and subsequently this type of system became widely used in the smaller Bell cities. All of this equipment was manufactured by the Automatic Electric Company under Bell specifications until 1926 at which time Western Electric began manufacture also; but Automatic Electric continued as a major supplier until about 1936.

The step-by-step system served small cities economically for many years until new developments after World War II provided an economic competitor in the form of the Number 5 crossbar system. This system, using the new crossbar technology introduced in 1937 and an advanced type of common control, was able to avoid the rigidity inherent in step-by-step and to provide many new service features. However, there is still a large amount of step-by-step equipment in the Bell plant and at its peak in 1972 there were about 42 million telephones being served.[125]

Equipment designs of both the step-by-step and the panel systems were highly ingenious mechanically. The step-by-step switch bank was relatively simple in design using inexpensive parts manufactured by means of punch presses. Since the same basic switch bank served many functions, the economy of large-scale production was achieved. However, the actuating mechanism of the switch (with associated relay controls) was relatively expensive and, since it reached only 100 contacts, the cost was fairly high per contact. The panel bank had many of the manufacturing advantages of step-by-step. It, too, used punch-press parts of only a few basic designs which could be manufactured inexpensively in large numbers. In addition, its design eliminated much expensive wiring required at the time in step-by-step systems to connect contacts in multiple. Since each selector arm could reach 500 contacts, the cost per contact was modest even with a drive mechanism of some complexity.

Both types of equipment had an inherent electrical fault in that their switch contacts used base metals such as brass and bronze. These tended to develop high contact resistance under some atmospheric conditions and this, in turn, produced a scratchy type of noise requiring routine contact cleaning for its control. Since the mid-thirties, all new systems developed by Bell have used switch contacts which retain a low resistance over the life of the switch.

[125] In terms of relative use, the peak was in 1960, with 30 million telephones, or 49 percent of the total.

4.5.2 Retrospective View

Although this history is concerned primarily with the first 50 years of telephony, the work on machine switching during this period can best be evaluated when viewed from our position in the seventies. In 1925, the impact on the plant was small since less than 1.5 million or about 12 percent of the Bell system telephones were on a dial basis. But this was only a first step toward the evolution of a new and completely different switching plant that 40 years later had grown to over 75 million stations of which all but 0.2 percent were dial.

Remarkably, the number of operators in 1965 was about 150,000, or roughly the same as in 1925, and at no time during the intervening years had the number been less than about two-thirds of this amount. Thus, by carefully coordinating the introduction of automation with plant growth, a complete changeover was accomplished without the necessity for large reductions in the operating staff. In most cases the normal staff turnover was more than adequate for handling displacements resulting from mechanization. It has been estimated that without automation the amount of telephone business actually handled in 1965 would have required some seven times as many operators or roughly one million. In a sense these are not very meaningful figures since the population could not have supported such a work force (and the large number of new hires required to maintain it) and the cost would have been so great that the economy could not have supported 75 million telephones. However, the figures do provide some feeling for the size of the mechanization task growing out of the pioneering effort prior to 1925 and the enormous growth in communication that it made possible.

Another rather remarkable aspect of this great mechanization program is the fact that it was accomplished with rather small changes in the central-office maintenance staff (relative to the number of telephones). In 1920, with a purely manual plant, one maintenance man was required for each 1,600 telephones. In the early days, mechanization, as might be expected, placed a heavier burden on maintenance personnel. In the late thirties, when half the stations were dial, the average number of stations handled by a maintenance man dropped to 1,250 but the remarkable improvements in systems made after World War II reversed this trend and in 1965, with an all-dial plant, the number of stations per maintenance man averaged over 1,800, or about 15 percent better than the manual plant of 1920.[126]

[126] The figures quoted are Bell System averages based on the total central-office craftsmen employed by the Operating Telephone Companies and the Long Lines Dept. of AT&TCo. The extraordinary improvement in systems reliability that brought about this reduction in maintenance effort will be appreciated when it is realized that, by 1965, plant complexity had been greatly increased not only by the mechanization of the exchange and toll plants, but by the widespread use of sophisticated electronic equipment.

As we have implied, most of the credit for the amount of growth and service improvement in machine switching since 1925 is the result of the remarkable technical advances which took place during this period. Much as these advances have contributed, it would be unfair to overlook the valuable groundwork laid before 1925.

The step-by-step system, while presently obsolescent, was for many years the most economical machine system for small cities, received widespread use throughout the country, and in the early seventies still served about 40 percent of Bell System telephones. It would be fair to say that it was largely responsible for familiarizing the public with automated switching. The dial, invented in 1896 to control step-by-step switches, was one of the fundamental pieces of machine switching equipment. Its basic principles are still being used in most station sets today in spite of the rapid growth of pushbuttons, introduced in 1961 under the name of *Touch-Tone*® dialing.

The panel system too has had a long service life, a few offices still remaining in use in the early seventies. To this system goes the credit for introducing new concepts which met the needs of large metropolitan areas. It not only made the first application of common control but also introduced the numbering plan that has been used, with some supplementary changes, to meet growth far beyond anything anticipated in the 1920s including expansion to worldwide long-distance calls. The replacing crossbar systems, introduced in the late thirties, used improved switch banks (with rare-metal contacts) but originally retained much of the basic panel-system concept. While present-day systems, both of the crossbar and electronic type, use much more sophisticated controls than employed by panel, it would be fair to say that the concepts of registration and translation centralized in a common-control unit, first introduced by panel, are still basic to present-day systems. In addition, a number of the principles of present systems that contribute to their speed, flexibility, and reliability are traceable to the coordinate system, using the Reynolds switch, that was developed by the Bell System in 1913 but not adopted at that time very largely because of the lack of suitable production technology.

In summary, it is proper to say that by 1925 switching developments in the Bell System had provided much of the groundwork for the enormous expansion of mechanical switching that was to follow in the next half-century. The few cases where a progressive direct-control system was chosen to serve large cities have emphasized the far-reaching effect of many of the inventions incorporated in the panel system and have shown the wisdom of the decisions made early in the twentieth century to adopt highly novel concepts in the system design.

This concludes our discussion on the switching systems used in the exchange plant. This plant provides the main body of traffic for the telephone system since the major interests of most people are limited to rather small geographic areas. But the value of such systems is limited unless there are means for extending the range of communication far beyond local exchanges. Alexander Graham Bell had recognized from the start that one of the most valuable characteristics of telephony was its potential capability for supplying rapid two-way communication over great distances without the need for any special training on the part of the user. For this reason the conquest of distance was a major goal among transmission engineers and as early as 1885 AT&TCo was incorporated with the major objective of establishing long-distance lines for interconnecting exchange areas. The successful outcome required not only the extensive transmission developments related in preceding chapters but also new switching techniques, which will be covered in our next section.

While the exchange and long-distance systems provide the necessities for common-carrier systems, there is another type of switching which supplements the regular exchange system by providing conveniences not so readily obtained from a large central office.[127] These are the private branch exchange, or PBX, systems usually located on the customer's premises. These systems are devoted largely to the customer's own internal communications needs but also provide connection to the central office and thus form an important part of the overall switching network. These switching arrangements also will be discussed shortly in order to provide an overall picture of the major elements of a complete switching system capable of covering unlimited geographical areas and meeting the needs of a wide variety of users.

V. TOLL SWITCHING

The provision of communication over long distances was part of Bell's "Grand System" concept of 1878 and within a few years after Bell's prospectus, lines between cities were built in several parts of the country.[128] It was apparent from the beginning that the cost of the lines increased rapidly with distance and special charges, or "tolls," were introduced for calls between cities to supplement the charge (usually on a monthly basis) for local or exchange service. These tolls, based on distance and time of usage, were, from the customer's standpoint, the significant characteristic of long-distance service and the term "toll" was used as a designation for the service, for calls, and for the lines involved.

[127] This situation held until the 1960s when some PBX-type service began to be supplied by central-office switching.

[128] By 1880, in Massachusetts, Springfield had been connected to Lowell and Boston to Holyoke. In the West, lines had been run between Champaign and Urbana, Illinois, and several mining communities of California were connected. In a primitive way, long-distance communication had begun.

We shall continue that usage here even though present-day nomenclature is slightly different.[129]

By 1885, toll service had made a significant step forward when the New York-Boston line, using metallic transmission, was placed in service. In the same year AT&TCo was incorporated to help bring about a complete and integrated service and specifically to provide connecting lines between cities. It is interesting to note that the Certificate of Incorporation, in describing the lines to be supplied by the Company says, in part, that they:

> . . . will connect one or more points in each and every city, town or place in the state of New York with one or more points in each and every other city, town or place in said state, and in each and every other of the United States, and in Canada and Mexico; and each and every other of said cities, towns and places is to be connected with each and every other city, town or place in said states and countries, and also by cable and other appropriate means with the rest of the known world, as may hereafter become necessary or desirable in conducting the business of this association.

Obviously, this was a very long range goal since technical obstacles of great difficulty stood in the path, but its existence did much to define the route to be followed and guide the technical effort required to solve the problems of communication over great distances.[130] At the time, even the problem of interconnecting exchange and toll lines had not been solved. Long-distance telephony required special, high-efficiency terminal equipment which was impractical for local service. Hence it was necessary for the user to go to a special telephone when such service was desired.

Although universal toll service was recognized as a formal goal in 1885, it remained in a preliminary, or exploratory, phase for about ten years. By the mid-1890s, transmission, station, and switching development all had advanced to the point where the requirements for a unified system were understood and the means for implementing such a system over moderate distances were available and coming into rapid use. The solid-back telephone transmitter (see Section 2.1.3 of Chapter 3) provided a stable and high-efficiency instrument suitable for both toll and local usage. The theory of transmission was becoming understood,

[129] "Intertoll trunks" is the accepted name for toll lines; however, it is interesting to note that interstate trunks are furnished by the "Long Lines Department" of AT&TCo.

[130] As related in previous chapters, it was 30 years before transcontinental telephony was technically possible and another ten years before means were found for spanning the oceans. In both of these cases, a new invention, the electron tube, not even dreamed of in 1885, played a dominant role. The original overseas service was by radio, a technique unknown in 1885, and the use of undersea cable for long distances did not come about until the mid-1950s. The present-day microwave and cable systems which provide long-distance circuits several orders of magnitude cheaper than the early transcontinental systems owe much of their effectiveness to still another far-reaching invention, the transistor, which didn't come along until more than 60 years after the formation of AT&TCo.

metallic transmission (with low noise and reduced crosstalk) had become standard for both local and long-distance lines, and inductive loading (by reducing the attenuation of lines by a factor of two or three) not only had extended the range of open wire lines but also made practical the use of compact, fine-wire cables in congested areas and on heavy traffic routes. The common-battery, 2-wire multiple switchboard with simplified operating methods also had been introduced into the exchange plant about 1895, and by the end of the decade much of the plant had been changed or was being changed to incorporate these new devices and techniques. By the turn of the century, the Bell System was ready to begin universal service and meet the demands for extremely rapid growth, a phenomenon which occurred in three dimensions—numbers of telephones, calling range, and volume of traffic. It was in this period that the development of true long-distance switching systems began.

5.1 The Special Problems of Toll Switching

To a considerable extent the toll interconnection problem was similar to that in the exchange plant. The same basic mechanism (consisting of jacks and indicators mounted in a vertical backboard with a keyshelf containing the interconnecting plugs, cords, and supplementary keys and signals) was applicable to both. To this extent, toll switching systems relied heavily on the development of exchange switchboards. But the greater length of toll lines introduced two basic problems which affected the development of other aspects of toll switching. First, the transmission of switching signals was much more difficult[131] and, second, the line cost was many times greater. Thus, toll service required both special signaling systems and also special arrangements for minimizing the cost to the user. The cost problem ultimately was solved by using carrier and other transmission techniques described in Chapters 4 and 5. These techniques reduced the cost by several orders of magnitude but for many years the only way in which long-distance service could be brought into the consumer's price range was by employing traffic and pricing methods which provided high usage of toll lines. Such methods led to operating schemes differing greatly from those appropriate for the exchange plant.

In the exchange plant the basic concept was of a service furnished on demand and as nearly instantaneous as possible. As a result, the lines were used very inefficiently; loops, for example, might handle only a half-dozen calls a day and trunks perhaps five times as many. During

[131] The voice-transmission problem is covered in previous chapters and was solved by loading, amplification, and carrier systems. These techniques did not necessarily improve the transmission of switching signals. They often complicated the problem by requiring the development of new signaling techniques.

the first 50 years of telephony it was unthinkable to provide enough lines for this type of service over long distances since, if the needs were met for the peak demand period, large numbers of costly facilities would be idle for most of the time. It was unavoidable, therefore, that some means be found for using the lines during a large part of the day. Two schemes were devised. One was to provide delayed service so that the calls could be queued up in the order of receipt and completed over the lines as they became free. In this way, lines could be kept busy over long periods with only short idle intervals between calls. Another approach was to extend the hours of usage by charging less for calls during the night periods when traffic would otherwise be negligible.

The cost of toll service was made equitable by charging on the basis of distance and duration of call.[132] Thus the high cost of toll lines brought about basic differences in the operating methods used for exchange and toll service, the latter requiring more complicated procedures in order to handle delayed calls, to take care of accurate timing and charging, and to handle these operations with the maximum practical satisfaction to the user.

In summary, toll switching systems could use much of the art developed for exchange switching but also required additional arrangements to provide the best balance between operating methods and the cost of switching, signaling, and circuits. Economic factors required delayed, rather than instantaneous, service on many calls but throughout the evolution of toll switching an overriding objective has been to minimize the interval between placing and establishing the call to the greatest extent consistent with the cost of providing service.

5.2 Operating Methods

From an early date it was recognized that different operating methods were appropriate for short- and long-haul calls. Short-haul toll service (for distances up to about 30 to 50 miles) included the traffic in extensive suburban areas around large cities and constituted the great majority of toll calls. On these calls, since the traffic was fairly heavy, there was little need for delayed service. The trunks were too expensive for inclusion in the exchange calling area, with bulk billing, but, being short, were relatively inexpensive as compared to the very long haul variety.[133] While tolls on such calls were necessary, they were modest

[132] Originally, the charge was on the basis of a 5-minute initial period with overtime charges for each additional minute, but in 1896 the 3-minute initial period was introduced and remained in use until very recent times when a 1-minute initial period was introduced on a limited basis. The 3-minute period provided a reasonably low rate for the economically minded but did not inhibit usage to any great extent since toll calls averaged about 6 minutes in length for many years.

[133] Although the length of these circuits was short, the fact that they traversed city areas often called for the use of cable which, being an inherently high-loss facility, often required special arrangements to provide adequate transmission and, until the use of loading became common, necessitated the use of very heavy conductors.

in amount and did not encourage the high use of person-to-person calls that were common on the more expensive long-haul circuits. Thus, relatively simple operating procedures were possible.

5.2.1 Short-Haul Methods—A-B Toll Calls

Most short-haul traffic, except for charging, was handled very much like local traffic with respect to switchboard equipment, operating methods, and interarea signaling. As in exchange service, a short-haul toll call appearing at a local switchboard was answered by the A operator and connected by her to a trunk to a distant B-board at which point the B operator completed the connection to the proper subscriber while the calling party remained on the line. This type of service became known as A-B toll and evolved in step with the local service, being implemented with similar switching and signaling facilities. Charging was done by the A operator, using simplified charge tickets[134] on which the starting and completing time was obtained from the standard switchboard clock and entered manually on the ticket by the operator. A-B toll was used largely for distances under about 30 miles and was limited to calls for which the distant number was known to the caller.[135] Time and charge information could be furnished the caller but most calls involving operating complexities were handled through the toll board. Later, non-delayed service was extended to longer distances by toll board methods which now will be described.

5.2.2 Long-Haul Calls—Toll Board Methods

Long-haul methods differed from short-haul methods for many reasons. The high cost of lines required high toll charges and this encouraged the use of person-to-person calling with a considerable number of callbacks resulting when the desired party was not reached on the first attempt. The high cost, together with the long distances involved, resulted in low traffic density which could only be handled efficiently by the introduction of queuing and delay. All of these factors tended to reduce the number of routes justifying direct trunks. Consequently, the routing between cities often was complicated by the necessity for a number of tandem switches. Finally, since long-haul calls were infrequently made, the caller seldom knew the called number and obtaining it involved widespread use of information facilities.

To meet these situations it became customary, early in the history

[134] The toll ticket, which receives so much mention in this section, was a key element in the toll telephone process. It was a small slip of paper made out for each call and served as a working memorandum containing the information required by the operators involved in the call, it provided the time and distance information (i.e., terminal points) from which the charges were computed, and it provided a record which could be consulted in case of subsequent queries concerning the call. It will be described more completely subsequently.

[135] Information service was available to subscribers for offices in the A-B toll area.

of toll service, for the local operator to route long-haul calls to operators at separate toll switchboards.[136] Once the toll office was reached, operators there took full charge of the call, handling completion, supervision, timing, and so forth. Several toll operators often were involved in processing the call, sometimes as many as four or five being required. Very large cities might have one or two centralized toll boards; medium-sized cities would have one board at most and, if they were not sizable, the board often consisted of only a few special positions at an exchange office. In very small communities, and in offices suburban to large cities, toll switchboards were usually located in nearby cities known as toll centers and the local operators switched the call to these centers. Although the basic methods for completing toll calls were essentially the same, there were many modifications in particular cases, depending on the type of equipment, size of the office, and cost of switching.

5.2.2.1 Callback Methods. For many years the standard method for handling long-haul toll calls required the calling party to give information on the call to a toll operator and then hang up while the call was being processed. When the desired line (or party) had been reached, the originating party was called to the phone by the toll operator and the call proceeded. If there was any delay due to a lack of answer, party unavailable, and so forth, the originating party was called and given an appropriate report.

The earliest operating method of the callback type was referred to as the two-ticket process since tickets were made out by operators at both ends of the toll circuit. Under this method of operation, a customer wishing to place a toll call asked his local operator to connect him with long distance. This connection was made over a toll connecting trunk of the "recording" type. Such trunks only could be answered by one of a special group of operators trained specifically for recording the customer's order. The customer told the recording operator whom he wished to reach, where he might be found, and the type of call (person-to-person or station-to-station) desired. He then was told that the operator would call him back and he hung up his receiver.

After recording the information supplied by the customer on the "outward" ticket form, the ticket was sent to other operators for further handling. If the customer had not supplied the number of the called telephone, the ticket was sent to a directory operator who looked it up in the directory of the called place. Then, since each toll office did not have direct circuits to all other toll offices, it often was neces-

[136] The connection to the toll board was made over circuits generically called "toll switching trunks." As related later, there were several types of such trunks.

sary to determine the route to the called place. In this case, the ticket next went to a routing operator who indicated on the ticket the path to be followed in reaching the called place. The ticket finally was sent to the particular outward-line operator having access to toll lines which would reach the desired terminal office.[137]

Once this stage was reached, the outward-line operator was virtually in charge of the call but, since local operating procedures were not completely standardized, she needed some assistance from the inward toll operator at the distant office. The outward operator reached the inward operator either over a direct trunk or by proceeding in the direction of the called office over trunks which were connected in tandem by intermediate operators. Once the inward operator was reached, the details of the call were passed to her and recorded on an "inward" ticket which she used as a memorandum during the time required to obtain connection with the called telephone (or party) or to find out where or when the called person might be reached. Once the called station or party had been reached, she notified the originating operator, who then rang the calling party over a toll switching trunk of the "connecting" type. When both the calling and called parties answered their telephones, the originating operator entered the starting time on the ticket. When the calling party hung up his receiver at the end of the call, the originating operator received a disconnect signal whereupon she entered the call duration on the ticket and took down the connection to the calling party. The distant inward operator took down the connection at her end upon the receipt of the signal indicating that the called party had hung up his receiver. In some cases, the inward operator also provided a duration check by recording the starting and completion times on her ticket.

Once signaling, accounting, and operating procedures were standardized throughout the Bell System, it became possible for the outward operator to take over all the responsibility for handling the call at the far end except for the mechanical process of making physical connections to the called telephone. By making this change it was possible to reduce the time during which two operators were required on the line and also eliminate the inward ticket. Under this single-ticket operating method the early stages of call processing were the same as with the two-ticket scheme but the line operator, upon receiving the ticket, obtained a connection to the inward operator and asked for a connection

[137] Obviously, all these operations involved a large amount of physical transportation of toll tickets between the various operators involved. In a few very large offices, conveyor systems were used. In smaller offices, the tickets were transported by messengers. (In a few cases the messengers were furnished with roller skates to speed them on their way.) Ultimately, the transportation of tickets was eliminated by using telephone calling for obtaining number and routing information and by more sophisticated operating methods such as the CLR scheme described later.

to the called number. At the start of ringing, the inward operator dropped off and the outward operator continued all further negotiations with the called and calling parties. From the customer's standpoint the procedure was virtually unchanged but it did result in some reduction in operating time.[138]

As noted, two types of toll switching trunks were used to provide connection between toll and exchange offices. A "toll recording trunk" was used in making the initial connection when the customer asked for long distance and passed his request to the recording operator. Since these trunks were not part of the toll connection and were used only to pass information as far as the first toll office, they could be inexpensive, relatively high-loss trunks. However, when the customer was called back, his toll switching trunk was a link in the toll connection and, in view of the problems involved in transmission over long distance, it was desirable to use a low-loss circuit in order to reduce the burden on toll line facilities. Therefore, these "toll connecting trunks" used high-grade facilities and, in addition, terminated at the exchange B-board in a special toll-grade cord circuit which provided higher battery current to the transmitter with consequent improvement in transmission from the station.[139] Toll switching trunks with high battery supply were also used at the distant end of the toll circuit. From a transmission standpoint these were equivalent to toll connecting trunks but were often referred to as "completing" trunks.

5.2.2.2 Methods Not Requiring Callback. The callback method was obviously an awkward procedure from the customer's standpoint and was justified only because it was a practical way to queue up calls so as to

[138] The single-ticket method appears so simple and obvious it seems surprising, in retrospect, that the two-ticket system ever received widespread use. While documentation on the subject is not very complete, there seems to have been a number of reasons. The lack of standardization, mentioned above, was undoubtedly an important one since the outward operator could not be expected to know all of the intricacies of the methods used in the distant cities to which she had access. Another reason for the second ticket seems to have been to provide a check on timing and charges. Timing differences were not uncommon since, with the primitive arrangements used for supervision, a busy operator might not always be aware of the precise beginning and end of a call. As a consequence, it was customary to reconcile the timing data obtained by the two operators and make adjustments in the division of revenue about once a month. Finally, transmission from operators' sets often left something to be desired and there was a considerable advantage in having negotiation with the distant party carried out by the closest operator since person-to-person and reverse-charge calls often involved extensive and complicated discussions.

[139] It might seem that it would have been advantageous to use high battery current with its improved transmission on all calls. This was not the case partly because of the added cost on exchange calls but more particularly because the added transmission efficiency would have been unpleasant on non-trunked calls which made up a large part of the traffic. In addition, while high battery current could be tolerated occasionally, its use on every call could have made the transmitter age more rapidly, leading to noise and unsatisfactory performance.

provide heavy usage of very expensive toll lines. From the customer's standpoint, A-B toll was ideal since he could call by number and stay on the line until the call was completed. However, it was limited by economic and transmission considerations to short, heavy traffic routes. As long-distance traffic increased, it became possible to introduce toll board methods on some of the heavy routes (such as New York to Philadelphia or Boston) that allowed completion while the customer remained connected to the office.

The first approach to non-callback toll board operation was the two-number system introduced in large metropolitan areas where separate trunk plant and a separate "two-number" office could be justified for handling traffic to points beyond the range of A-B toll. The origin of the name for this type of operation will be clear from the description of the procedure. As with A-B toll, the customer gave his A operator the city and number of the party desired. The A operator then called the two-number toll operator over a toll-recording-type trunk and passed the number of both the called and the calling parties. The two-number operator then ordered up a connection to the calling party over a high-grade toll connecting trunk to the B-board, the connection being made at her request without busy test. Once the B-board connection was made, the two-number operator disconnected the recording trunk, giving a signal to the A operator to release her original connection. The two-number operator then assumed full charge of the call and established connection with the called party by an appropriate toll line. She also made out a ticket covering the call and recorded the necessary timing. On a simple station-to-station call the caller remained on the line until the call was completed (or terminated by a "busy" or "don't answer").

About 1925, another method was introduced for handling toll calls while the called party remained on the line. This method, called CLR (for combined line and recording), was based on using a single operator to handle a call instead of dividing the work among specialized operators.[140]

Under the CLR method of operating, the customer dialed or asked for long distance in the usual way. The signal at the long-distance board appeared before the line operator who answered with the words "Long Distance." The line operator recorded the call in the usual way except that when the customer gave the name of the called place and the number of the called telephone, she selected a circuit to the called place and recorded the information on the ticket while waiting for the inward operator at the called place to answer. After passing the called number to the inward operator and while waiting for the called telephone

[140] By the end of 1926, about 150 of the larger cities were using this method.

to answer, she asked the calling party for his telephone number. Conversation was timed and the ticket disposed of in the usual way.

Under the single-ticket method, calls to or via a city were always handled by the same group of operators. Under the CLR method, any line operator could handle a call to any place in the toll system. If the call was not completed on the first attempt, the ticket was sent to the "point-to-point" position designated to handle calls to the particular city being called. This assured prompt and careful handling of those calls which encountered delay and the handling of such calls did not interfere with work on new calls at the CLR position.

In summary, it will be noted that this method of operation differed very little from short-haul methods such as the A-B and two-number schemes except that the desired number was given by the calling party to a toll operator rather than to a local operator. Otherwise the procedure from the customer's standpoint was very much the same and on a normal call the customer was able to remain on the connection until the called party answered.

The CLR method differed somewhat from earlier methods in the toll switching trunks employed. At the originating end the connection between the local and toll offices was made over high-grade "recording-completing" trunks, the higher-loss, less-expensive recording trunk no longer being employed. This was economically practical because of the highly improved and lower-cost facilities which had become available with fine-gauge, loaded cables. The use of high grade completing trunks with high current battery was continued at the distant end but the high battery supply was not always used at the originating end. When the calls originated at a dial office, it was relatively easy to program the equipment to provide such a supply when the long-distance code was dialed. In manual offices, however, high battery supply involved a complex process and ordinarily was not used since the low-loss transmission facilities available when the CLR method was introduced made the battery-supply improvement less necessary.

5.2.3 Ticketing, Timing, and Charging

A basic difference between exchange and toll operating was the necessity, with the latter, for providing accurate records for each individual call, which could be used by the Accounting Department for billing the customer for the appropriate toll charge. The "ticket," frequently referred to above, provided this information. While the format of the ticket varied in the early days of telephony, standardized ticket formats were soon introduced and continued in use for many years.

The ticket used by outward toll operators was made of fairly heavy paper, about 3 by 5 inches in size. As illustrated in Fig. 6-71, the front of the ticket had designated spaces for all the information required for establishing the call and billing the customer. Included were names, numbers, terminals, routing information, operators involved, type of call, (collect, person-to-person, etc.), charges, etc. The back was blank and used for recording the time the call began and the duration.

Calls handled at the toll board often involved fairly high charges and accurate time records were required. These were obtained with a timing device, called a calculagraph,[141] mounted on the keyshelf. The calculagraph (Fig. 6-72) had a clock face and a slot for inserting the ticket. With the ticket in place, front-side up, the right lever was pushed backward and then pulled forward at the start of the call. These operations recorded (on the back of the ticket) the calculagraph number, the time of day at which the call started, and a pair of dial imprints to be used later for indicating the elapsed time. At the end of the call, the ticket was again inserted in the same calculagraph and the left lever pulled forward. This operation again recorded the calculagraph number and also pointers which indicated the elapsed time on the dials previously printed. The various recording steps are shown on Fig. 6-73. The left side of the figure shows the type of recording obtained with calculagraphs made until about 1930. These devices had spring-driven, 8-day-clock mechanisms and recorded the elapsed time by means of the two left-hand dials, one being marked in 5-minute intervals up to 60 minutes and the other in quarter-minute steps up to 5 minutes. In the 1930s, regulated 60-hertz power became widely available and a synchronous electric motor was used as the drive mechanism. With the extra power available, it became possible to record time to the nearest second. Recordings made with an electric-drive calculagraph are shown at the right of Fig. 6-73.[142] Regardless of the drive, each operation recorded the calculagraph number to assure that the same instrument

[141] The calculagraph was invented by Henry Abbott, a 38-year-old New York clockmaker, in 1888. Tradition has it that the invention was made on the March day that brought the famous blizzard to New York City. Abbott did not have the timing of toll calls specifically in mind. Instead, the invention came about in response to a vague feeling that there must be a considerable market for a device that would not only measure elapsed time (the stop watch, already in existence, did this) but also provide a permanent record. The device was not used in telephony until 1894. Before this time, other means for timing telephone calls appeared adequate for the small amount of toll business carried on. One telephone manager, on being solicited by Mr. Abbott, said he had no use for such a device since his company's toll lines received so little use he would be glad to have his customers talk over them without limit. However, the situation changed after the turn of the century and Mr. Abbott found himself with a very profitable business which he headed in person until he reached an age well into his nineties.

[142] Readers interested in mechanical devices will undoubtedly recognize that each pointer always pointed at 0 on its scale and both the pointers and dials rotated together at appropriate rates. Since the first operation recorded the position of the dials only, and the second recorded the pointers only, the two operations provided a record of elapsed time, the necessary subtractions being made automatically.

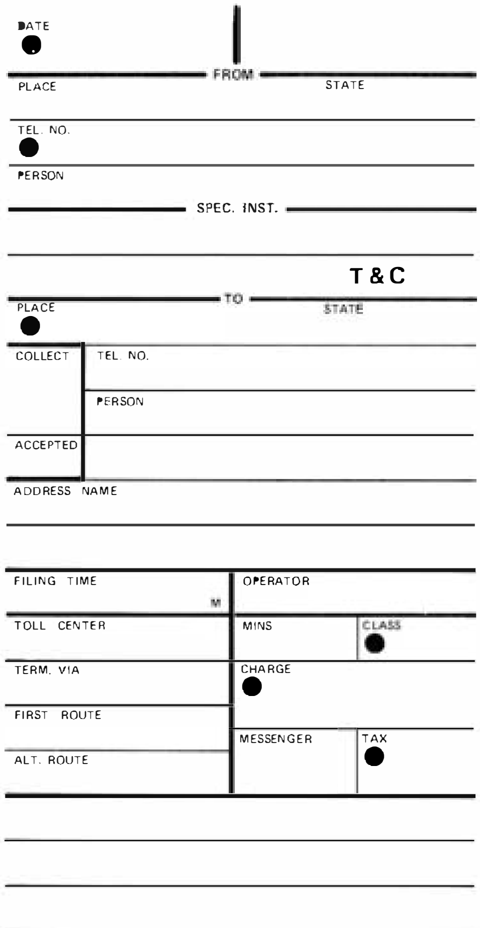


Fig. 6-71. Outward toll ticket.

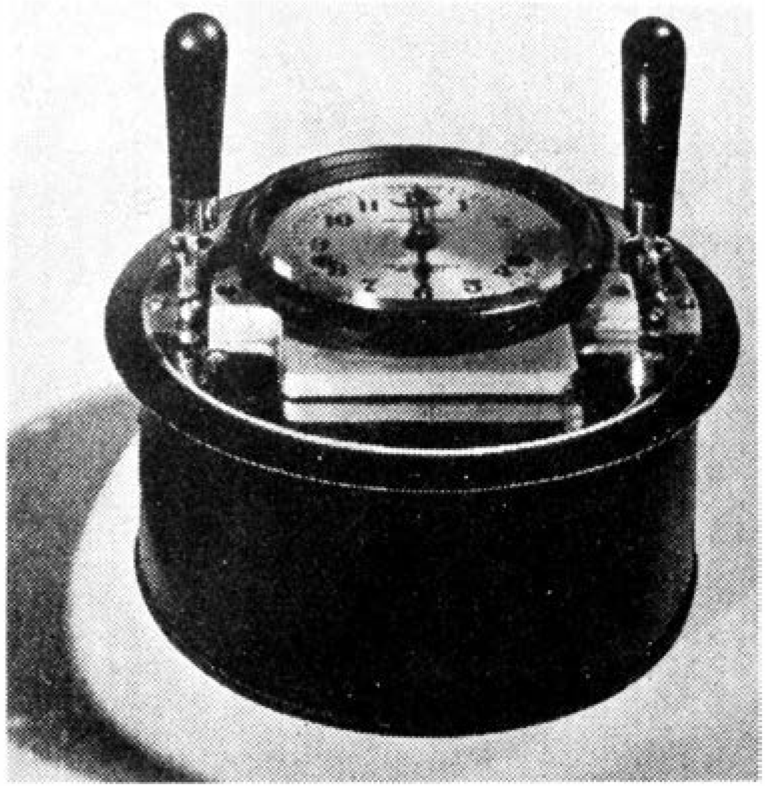


Fig. 6-72. Abbott's calculagraph.

had been used for the two recordings. An error would have resulted otherwise since the dials of the various instruments were not positioned identically.

Since toll tickets provided the basic information on time and charges, it was necessary to file them for a period long enough to meet legal requirements. This ultimately involved large amounts of storage space and was one of the many reasons why automatic message accounting was developed after World War II.

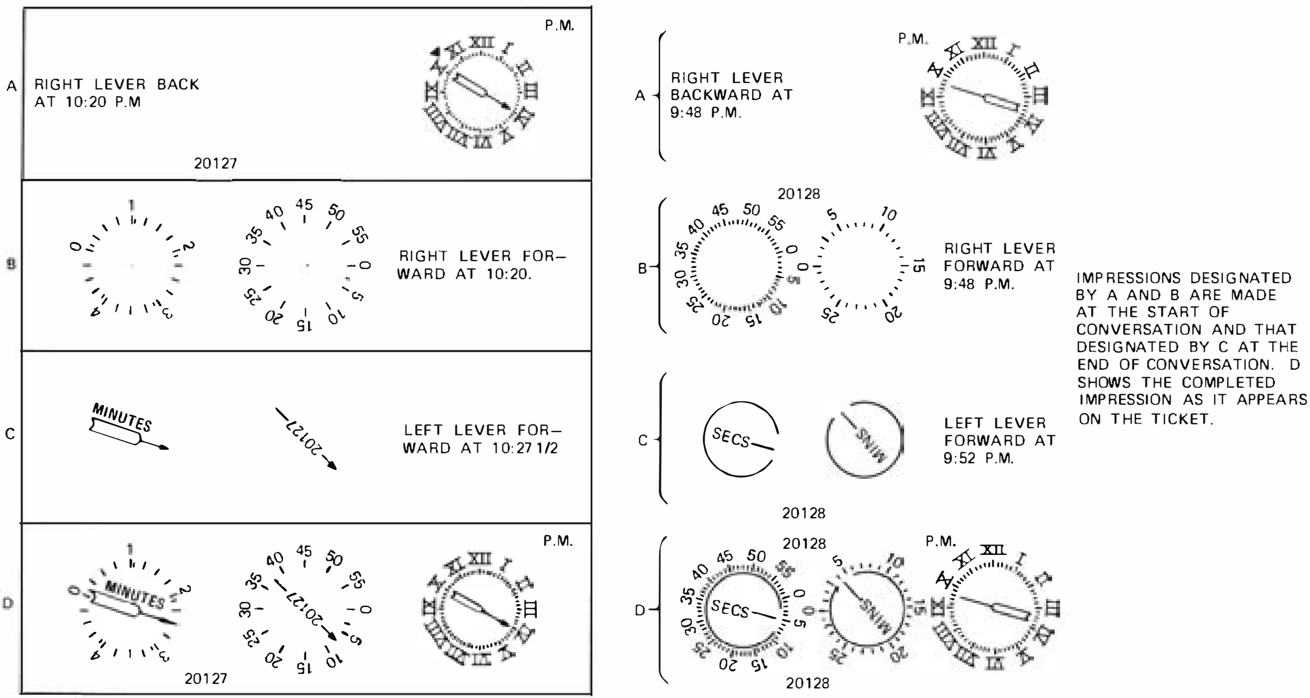
A-B toll calls were mostly handled on a station-to-station basis with numbers supplied by the caller. Hence, they could be processed with less information, and a smaller and simplified ticket was used. Since the calls were relatively inexpensive, less precise recording was allowable and was obtained from the standard clock on the A operator's keyshelf. Entries for starting and completing time were recorded manually on the ticket by the operator who also made the subtraction to determine elapsed time. The keyshelf clocks were of the 12-hour digital type and were usually installed on the left side of alternate positions so each operator would have one clearly in view. Figure 6-17 shows one of these clocks immediately behind the array of call-circuit buttons at the extreme right of the picture. (To the left of the clock is a holder for a pad of tickets.) The keyshelf clocks initially were very much like the message register used to record calls in large cities where measured service was used (see Section 7.3 of Chapter 3). The clocks were actuated by pulses from a master clock in the office every 6 seconds and recorded time in hours, minutes, and tenths of minutes. Later these clocks too were driven by 60-hertz synchronous motors and the dials arranged to show seconds instead of tenths of minutes.

The charges for toll calls changed greatly over the years. Typical charges at the end of the twenties are shown on Fig. 6-74.[143]

5.3 Intertoll Signaling Systems

The preceding sections have outlined the manner in which operators, in establishing connections between customers, met the special requirements of long-distance service, particularly the need for using circuits intensively and charging for the service equitably. The methods used by the operator were carried out largely by two subsystems, the signaling systems described in this section and the switchboard systems covered in Section 5.4. Before proceeding with the discussion of signaling,

[143] As a matter of interest, in 1972 the maximum 3-minute daytime (8:00 A.M.–5:00 P.M.) station-to-station rate for calls anywhere in the continental United States was $1.35. The maximum weekday evening rate was 85¢ and the weekend daytime rate was 70¢. Every night between 11:00 P.M. and 8:00 A.M. the initial period was reduced to 1 minute and cost 35¢ with additional minutes adding 20¢. Person-to-person calls, reflecting the high cost of operating expense, cost $3.55 at any time of day.



SPRING DRIVE
ELAPSED TIME = 7–1/2 MINUTES

ELECTRIC DRIVE
ELAPSED TIME = 4 MINUTES, 5 SECONDS

Fig. 6-73. Examples of calculagraph impressions.

Type of Call	3 – Minute Charge for Airline Miles Shown				
	10	100	500	1,000	3,000
Station-to-Station:					
4:30 A.M to 7:00 P.M.	$.10	.70	2.05	3.75	10.25
7:00 P.M to 8:30 P.M.	.10	.55	1.75	3.00	8.25
8:30 P.M. to 4:30 A.M.	.10	.35	1.15	2.00	6.25
Person-to-Person:					
All Hours	.20	.90	2.55	4.75	12.75

Fig. 6-74. Typical interstate toll charges as of February 1929.

perhaps it will not be out of place to emphasize once again the interdependence of the operating, signaling, and switching systems; the division we have made in this chapter has been adopted as a matter of descriptive convenience only. At any stage in the evolution of toll switching the particular solution adopted for an operating, signaling, or connecting problem was determined not only by the technology available to the particular subsystem involved but by the interaction among the various subsystems and equipments used throughout the toll switching system. And perhaps more importantly, we should also recall the important part played by the technology currently available for designing transmission and station systems.

Signals of various kinds were required on a toll call. Reaching the toll operator and conveying the customer's desires involved electrical and voice signals already described in Sections III and IV. The toll operator, in establishing a connection, required signals to alert the distant and intermediate operators when she was setting up the call or when she needed to recall them for subsequent action. Signals also were needed in the reverse direction to indicate actions by the distant party or the operator. For the period we are discussing (through the mid-1920s) much of the information was conveyed to operators by speech but later signaling systems became more sophisticated and electrical signals were used for conveying information previously transmitted by voice.

In the exchange plant, it was common to simplify signaling systems by using trunks designed so that calls could be originated at one end only. Thus each trunk group consisted of two parts, one capable of originating at one end, the other functioning in a reverse manner. Where large groups were involved, the number was not significantly increased by the use of trunks with "one-way" signaling. A similar scheme was used for short-haul toll trunks or on heavy routes, but for many years the number of trunks in a long-haul group was very small and used "two-way" signaling so that calls could be originated at either end of a given trunk. The signaling complications so introduced were more than offset by the reduced number of trunks required.

5.3.1 *Signaling Methods*

Generally speaking, the methods used for toll signaling were, from the operator's viewpoint, very similar to those used in the exchange plant but, as will be covered shortly, often employed very different technological means. Three basic signaling modes[144] were employed:

Ringdown—Broadly, this referred to the case where electrical signals were sent over the toll circuit that was to carry the conversation and, more particularly, where initiation of the signals was controlled by the operator at either end. This was the method first used on long-haul toll lines and remained the predominant method throughout the manual toll era. Circuits were of the two-way type.

Call Circuit—This method used a separate "order wire" circuit for conveying signals by voice means. It functioned in the same manner as the exchange call-circuit trunking described earlier. It used one-way circuits and was employed only with large trunk groups where the saving in signaling equipment and operating expense justified the additional circuits required for the order wire and for one-way operation. Signaling over the talking circuit was not completely eliminated since it was necessary to provide switchhook supervision to the originating operator, i.e., let her know whether the called subscriber's receiver was on or off hook.

Straightforward Operation—When this form of trunking was introduced in the exchange plant in the mid-1920s, it was also made available for toll systems as a replacement for the call-circuit method. Functionally it was the same as described under exchange trunking. Its use was somewhat limited since it required one-way trunks and caused complications when CLR operation was introduced since all outward CLR positions had to be equipped to handle all signaling modes. Although not extensively used, the technology developed for straightforward trunking was directly applicable to the subsequent evolution of nationwide operator dialing started in the 1940s.

5.3.2 *Signaling Technology*

Regardless of the method employed, signaling systems are alike functionally in that they must accept information generated at an originating point, convert it to a form suitable for transmission over telephone lines, and reconvert it to appropriate format for use at the

[144] These were also referred to as "trunking methods" since they involved rather distinctive ways of setting up trunks between offices.

destination. Criteria must be met with respect to speed and accuracy, efficient use of the line, and signal level and sensitivity of detection. Level and sensitivity present particular difficulties since they have to be adequate to avoid error from signal mutilation caused by line distortion or from unwanted signals such as speech or noise. At the same time, the signals must not be of a type or magnitude that would interfere with speech or other communication carried by the line. Meeting these complex requirements in an economical manner constitutes the technological problem of signal system design.

Although the requirements are basically the same for toll and exchange systems and the signaling methods employed are superficially alike, the technology required for toll signaling is often quite different since the high attenuation, distortion, and cost of toll lines prevent the use of many of the simple techniques applicable to the exchange plant.

5.3.2.1 Ringdown Systems. The information requirements for ringdown signaling are relatively simple since the only need is to notify an operator at one end of a trunk that voice communication is desired (or no longer required) by the operator at the other end. In all ringdown systems this was accomplished by applying a high-voltage alternating current at one end that operated a drop or other indicator at the distant end. Similar arrangements were provided at each end so that the trunks operated in the two-way mode.

The signal originally used was about 100-volt, 16-hertz current, roughly the same as used for the customer ringing generated by a hand-cranked magneto in the station set. Consequently, ringdown toll lines were often referred to as "magneto toll lines" even though the toll signaling currents were derived from power-driven generators at the toll office. As covered in Section 4.1 of Chapter 3, the ringing frequency was changed to 20 hertz about 1917 when commercial 60-hertz current became the commonly used prime power source. Except for this minor change, the low-frequency system, using 16- to 20-hertz signals,[145] had a long history of use throughout the manual toll era, but deficiencies showed up first on very long open wire lines and later when loaded cable began to be used extensively.

The problem with open wire was to a large extent economic. As the length of these lines increased, they became very costly and in order to use them more efficiently, means were devised in the early 1890s to use the lower end of the frequency band, not required for speech transmission, for transmitting telegraph signals from each

[145] As a matter of convenience, we shall refer to the signaling tone as 16 hertz in this chapter unless the precise frequency used is critical.

wire to ground.[146] Thus, each pair provided two telegraph channels to help carry the cost of the line. However, there was a serious conflict with 16-hertz signals since they fell in the same frequency band used for telegraph. If metallic signaling was continued, both telegraph channels were lost. As an alternative, it was possible to use one of the telegraph channels for signaling but this still represented a large economic loss. To avoid this penalty, a signaling system using 135 hertz, transmitted on a metallic basis, was developed about 1902. This frequency was above the band devoted to telegraph but was lower than necessary for telephone communication. Thus long-haul circuits were, in effect, divided into three frequency bands: one below 100 hertz for telegraph, a second around 135 hertz for signaling, and the third above about 200 hertz for voice transmission. Short lines, on which the cost of composite telegraph was not justified, continued to use 16-hertz signaling applied on a metallic basis (i.e., using a balanced rather than a ground-return circuit).

Since both the 16-hertz and 135-hertz systems were to be used for many years, the latter system was introduced without changing the operating method or switchboard design. The operator continued to ring the distant end by operating the same switchboard key regardless of the signaling system employed and the output was still 16 hertz. For short lines this output was used directly for signaling but for longer lines, carrying telegraph, the output operated a relay arrangement, called a "composite ringer," which responded to the 16-hertz current from the switchboard key and applied the 135-hertz current to the toll line. At the far end an ac relay, mechanically tuned to 135 hertz, responded to the signal and connected 16-hertz ringing current toward the toll switchboard to operate the drop or line-lamp relay.

The choice of 135 hertz for the new signaling system was wisely made since the frequency not only fell between the voice and telegraph bands but also differed from the fundamental and the harmonics of all other alternating-current signals used in the telephone or related plants. This avoidance of other tones and their harmonics was an early recognition of a signaling objective, which continues today even despite the great number of signals now in the plant. The 135-hertz signaling system, with detailed improvements made along the way, was unchallenged as the major technology for ringdown operation on the longer circuits for the next 15 years.

The problems introduced by cable were more technical than economic. On cable it was customary to use the simplex circuit [147] for telegraph

[146] See Chapter 4, Section 4.1.2. This was done by means of a "composite set" which was a combination of high- and low-pass filters dividing the open wire circuit into two frequency bands, the dividing point being roughly 100 hertz, the lower band being used for telegraph and the upper for voice transmission.

[147] See Section 4.1.2 of Chapter 4 for a description of simplex transmission.

and the metallic circuit for signaling. These two modes of transmission were compatible and, except for some minor problems, could use the same frequencies without loss of telegraph channels. The major problems with 16-hertz signaling arose with the introduction of loading, which extended the length of cable that could be used and also encouraged the use of fine, high-resistance conductors. This made signal transmission more difficult because of the high attenuation and distortion with consequent signal mutilation. Perhaps more important was the necessity for using repeating coils for impedance matching between cable and open wire when the facilities were used in tandem. Such transformers badly attenuated and distorted low frequencies and seriously limited 16-hertz signaling. This situation was handled for a time by the use of improved detecting relays and repeat coils with improved ringthrough characteristics or bypass circuitry. These measures, together with extensive tests of the ringing characteristics of lines and equipment, kept 16-hertz signaling in use on short cable systems for a number of years, but the 135-hertz system was adopted for the longer systems since it proved to be easier to transmit through repeating coils.

By 1917, it was apparent that the use of telephone repeaters would grow rapidly and introduce serious signaling complications in the near future. On September 28 of that year a meeting was held among the foremost experts of the time on signaling, switching, transmission, and telegraph technology to take a system view of the situation as it existed and to evaluate the future outlook. Some of their more important conclusions were:

(*i*) Cable would gradually take the place of open wire for toll transmission. This would be accomplished by the use of large numbers of repeaters which would require means for passing signaling through or around them.

(*ii*) Where telegraph was to be provided over long loaded-cable circuits, it would be necessary to develop a metallic transmission system using very small currents.[148]

(*iii*) Metallic telegraph transmission would be incompatible with 16-hertz signaling and might well be difficult to coordinate with the existing 135-hertz metallic signaling system.

(*iv*) Circuit groups were becoming larger because of both normal traffic demand and the stimulation provided by the economy of

[148] The reasons were rather complex but in a rather simplified form were as follows: The rather high telegraph currents used on open wire would have reacted on telephone transmission by saturating the magnetic material used in the loading coils. Lowering the currents, while retaining ground return, would have led to impairment of the telegraph signal by variable ground potential. In addition, the wires in a cable were so closely coupled that interference between grounded telegraph circuits would have been severe. The use of metallic telegraph operating on small current solved all these problems.

cable circuits and as a result there would be a demand for automatic signaling (as required for switchhook supervision with call-circuit operation) in place of ringdown.

This appraisal stimulated effort along a number of lines. Included were the following possibilities: improved telephone repeater transmission at 135 hertz; improved 135-hertz signaling operating at lower, non-interfering currents; use of vacuum tubes in signaling circuits to obtain greater sensitivity; higher-frequency signaling; and common-channel multiplex signaling to perform the function for a group of circuits.

The final answer to the three-cornered problem caused by the expansion of repeatered cable operation, the need for metallic telegraph operation, and the use of ringdown operation with 135-hertz signaling did not come for several years. The initial solution was a new 135-hertz signaling system using low enough currents to avoid interfering with the metallic telegraph system, tuned sharply enough to reject telegraph interference, and sufficiently sensitive to operate on all except a few of the longest toll circuits.[149] On long circuits, the 135-hertz ring was relayed by an intermediate ringer at alternate repeaters. Other physical and maintenance improvements permitted successful use for many years thereafter. Much of the sensitivity and selectivity of the new system was due to the design of a new special relay called the 218B. This relay was used in an anti-resonant circuit sharply tuned to 135 hertz. When a signal was received, the relay armature, which was only lightly damped, vibrated and caused a secondary circuit to be continuously closed through appropriate contacts. In the development of this ringer, vacuum tubes were tried for the first time in signaling apparatus. While they proved unnecessary for this particular system, they formed an essential part of the system which soon followed.

It was appreciated, even while the improved 135-hertz system was being developed, that a preferable arrangement would be one operating in the band of frequencies used for voice transmission. With such a system, a signaling path would exist whenever it was possible to talk and repeaters, instead of being an obstruction to be bypassed, would aid signaling by amplifying the ringing as well as the voice currents. Furthermore, it would avoid foreseeable incompatibilities with future developments in telegraph and telephone technology and, in particular, would be a highly acceptable means for signaling over carrier channels.

The requirements for such a system were not easy to meet. The signal currents had to be much lower than previously used and of the same magnitude as voice currents since otherwise they would overload the voice system and cause objectionable crosstalk into voice circuits.

[149] The complete system was standardized by D&R Bulletin 133, dated March 21, 1924, but various features undoubtedly were used much earlier as required by the spreading cable network.

In addition, the signal detector could not be falsely activated by voice currents (a condition referred to as "talkoff").

A number of Bell System inventors contributed to the solution of this problem, and in 1921 patents were filed on signaling systems operating in the voice range by A. B. Clark, D. K. Gannett, and H. Nyquist[150] at AT&TCo, and by P. B. Murphy[151] at Western Electric.

By 1922, experimental operation was under way of a signaling system in which the basic signaling frequency was 1,000 hertz, just about at the optimum transmission point. To provide adequate sensitivity, a vacuum tube was employed to detect this current—the first commercial use of vacuum tubes in telephone signaling equipment. A sensitive, telegraph-type polar relay, the 215-type, operated from the output of the vacuum tube.

To avoid "talkoff" by voice harmonics or from 1,000-hertz test tones or harmonics, the 1,000-hertz signal tone was interrupted 20 times per second and the resulting signal was detected by relay apparatus tuned to 20 hertz. To further reduce the possibility of false operation, the circuit logic was designed so that the signal needed to be present for a short time before response took place.[152] Thus, while the system could not be said to have complete freedom from talkoff from all possible sources, it was reasonably improbable that false operation would occur. A simplified schematic is shown in Fig. 6-75.[153]

In 1923, the 1,000/20 system was installed on some of the longer circuits, such as Chicago to Los Angeles, and New York to New Orleans. In 1924, the southern transcontinental route was equipped.

A few years later a number of changes were made. The one-tube signal receiving circuit was further improved by the addition of a second vacuum tube and directionally selective arrangements were incorporated to reduce the effects of echo currents on re-rings with built-up connections, which otherwise would cause signal cancellation. The interrupted 1,000-hertz signaling supply was replaced by a 1,000-hertz carrier modulated by a 20-hertz tone. Intermediate 1,000/20 ringers also were developed so that they could interconnect at an intermediate repeater point with 20- or 135-hertz signals.

[150] A. B. Clark, et al.; U.S. Patent No. 1,519,573; filed November 17, 1921; issued December 16, 1924.

[151] P. B. Murphy; U.S. Patent No. 1,537,653; filed November 10, 1921; issued May 12, 1925.

[152] Originally it was referred to as the 1,000/cycle system but later it became known more accurately as the 1,000/20 system.

[153] These terminals, following the pattern established when the 135-hertz system was introduced, wereessentially two-way converters. A 1,000/20 signal applied to the left-hand side of the converter would be detected and send out corresponding 20-hertz signals applied to the right-hand side. Conversely, 20-hertz signals applied to the right-hand side would produce 1,000/20 signals at the left-hand. While usually employed to connect the 1,000/20 toll line system to the 20-hertz system used in the toll board, they could, in fact, be used wherever a conversion between the two systems was required.

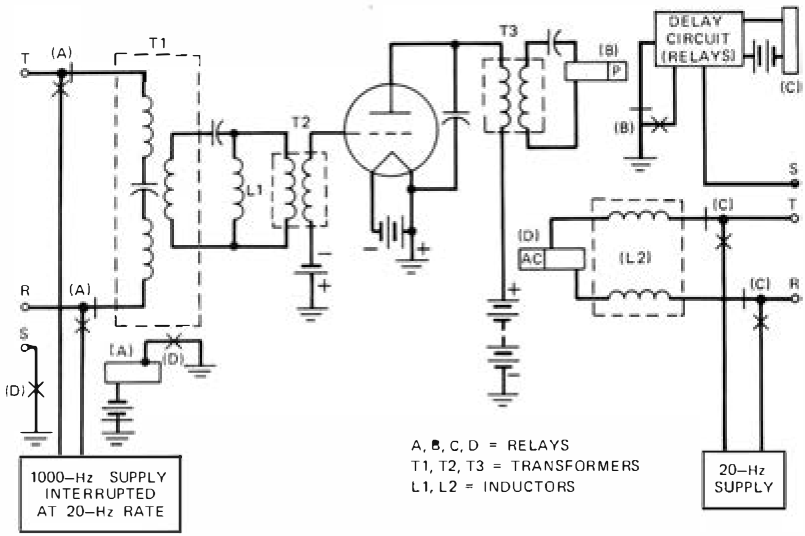


Fig. 6-75. Initial 1,000/20-hertz terminal ringer circuit, simplified schematic.

The 1,000/20 signaling system thereafter became the predominant standard for medium and long toll circuits, its use being found more economical than 135-hertz signaling when the toll lines had more than one intermediate repeater. Even on single-repeater lines, the use of the 1,000-20 system was encouraged for reasons of uniformity and it operated satisfactorily for many years.[154]

The observant reader will have noted that all the signaling arrangements we have discussed used alternating current. In the exchange plant, 16-hertz ac signaling techniques were being rapidly supplanted by dc methods at the turn of the century and about 1903 a somewhat comparable complicated dc technique was devised for toll signaling. However, direct-current signaling was abandoned for toll board calls about a year later (except for some use in call-circuit trunking) as its limited possibilities became apparent. Instead, the various ac methods we have outlined, using 16, 135, and 1,000 hertz, were chosen for toll service. In retrospect, we can see that this was a wise and foresighted decision since future developments, which made long-distance transmission possible, virtually eliminated the possiblity of maintaining a direct-current path from end to end and, where it was possible, it could be done only at the cost of losing a valuable telegraph channel.

[154] It is interesting to note that European administrations adopted a similar scheme for signaling over the voice channel but used 500 hertz as the carrier frequency while retaining 20 hertz as the modulating tone.

5.3.2.2 Call-Circuit and Straightforward Systems. Technically, these two systems had much in common in that they required an automatic type of signaling to minimize operating time for successful and economic application. Straightforward operation was little used during the period covered by this history but some of the technology developed during the period for other systems was useful for its later application and also was of significance in more recent systems.

The signaling requirements for call-circuit operation were very simple. Most information was transmitted orally and only the status of the called subscriber's switchhook had to be conveyed by electrical signals. Call-circuit operation was found economical for large circuit groups as early as 1907 and special arrangements were developed for this type of trunking even though large groups were uncommon until much later. The early systems used a dc arrangement (colloquially known as "wet-dry" signaling[155]) over the metallic circuit, or an equivalent scheme using a telegraph channel, to obtain switchhook supervision.

The use of either of these schemes caused the loss of telegraph channels. Originally this was of small moment since the large trunk groups necessary to justify call-circuit signaling tended to be so short that the cost of deriving telegraph channels made their value small. However, as circuit groups increased in size and call-circuit trunking was applied to longer circuits, the disadvantage of using telegraph channels increased, particularly on cable circuits which required metallic transmission of telegraph. The 1917 meeting to review signal systems recognized this situation and, as noted previously, pointed out the advantage of multiplexing signaling systems by using a single circuit to perform the signaling function for a group of voice channels, thus releasing the telegraph dc path otherwise required. Work soon began on developing this so-called "common-channel signaling system"[156] using basic concepts patented by Bancroft Gherardi[157] of AT&TCo and H. D. McPherson[158] of the Western Electric Company. The resultant common-channel signaling system (known as the distributor system for toll

[155] "Wet-dry" signaling was a two-state system, one state (wet) being indicated by the application of a dc potential to the line, the accompanying flow of current operating a relay. The absence of current indicated the second (or dry) state.

[156] This was not the first proposal for common-channel signaling. Before the development of the 135-hertz system, a scheme known as the Morse ticket system received some use. The basic idea was to use one telegraph channel as a Morse order wire to provide signaling information for a number of circuits. This had the advantage of saving telegraph channels and some telephone circuit time but obviously complicated operating methods very severely and its general use was abandoned when 135-hertz signaling became available.

[157] U.S. Patent No. 1,251,363; filed August 31, 1915; issued December 25, 1917.

[158] Nine consecutively numbered patents starting with U.S. Patent No. 1,387,284; filed July 31, 1919; issued August 9, 1921.

lines) was used in a commercial trial in 1922 and standardized by D&R Bulletin No. 162, dated August 15, 1924.[159]

The distributor system was originally used with call-circuit trunking but it soon became apparent that straightforward operation required little or no additional cost to implement since the capability was present in the signaling system. Accordingly, the full group of 240 circuits between New York and Philadelphia was converted to straightforward operation. Eight distributor signaling systems were required for this and they remained in service through the 1940s. Similar systems were installed in a few other cities but wide use did not come about because of engineering uncertainties of the time.

One of the uncertainties was the size of the circuit group needed to justify the expense of a common system—a problem frequently encountered in the Bell System as it explored the use of common-system technology for groups of circuits. Costs were such that economies for the common signaling system over 135-hertz or 1,000/20 ringdown signaling appeared possible only when circuit groups were as large as 25, 30, or even 50 circuits each, depending upon particular configurations. Since only about 1 percent of the AT&TCo Long Lines circuit groups were more than 30 circuits per group, and since a far smaller percentage prevailed for Associated Company circuit groups, widespread use of the system was not economical. Other considerations inhibiting the use of the common signaling system involved differences between the supervisory features of this system as compared with those that had been developed for straightforward operation with individual signaling per circuit.

The straightforward signaling technology developed for individual circuit operation stemmed from the techniques used for composited grounded telegraph systems. A number of applications based on these techniques were made but the technology was in conflict, as previously noted, with the objective of providing telegraph service over toll lines. Accordingly the need became apparent in the latter part of the 1920s for some other signaling technology for straightforward operation. Considerable thought was given to adapting the 1,000/20 system to this use, but it was not for many years that a suitable in-band signaling system for straightforward, and later dialing, purposes was developed.

Even though the distributor system was used only to a limited extent, its description is of interest since it was a precursor of concepts used much later, including modern electronic systems which

[159] At this time, a system became a Bell System Standard when a descriptive bulletin was issued by the Development and Research Department of AT&TCo. These "D&R Bulletins" were usually issued when the system had been put into final form after extensive field trials. Frequently, nearly identical systems had been in commercial use for a number of years on an experimental or provisional basis before standardization.

transmit signaling information for a plurality of lines over a common data link. The concept also is being put to use in other modern signaling systems as a replacement for the voice-frequency systems that have been used so extensively since World War II.

The distributor system developed in 1922 compressed the supervisory signals for 30 toll lines into one telegraph channel through the use of time-division multiplexing techniques, thus freeing all other low-frequency channels associated with the 30 voice circuits for telegraph use. Such compression was possible because the system effectively transmitted only two status conditions per circuit for each direction of traffic, i.e., off-hook and on-hook; and changes of state occurred infrequently compared to the holding time of each circuit. The techniques used were those of the start-stop telegraph-printer multiplex system described in Section IX of Chapter 7.

Fundamentally, the distributor system consisted of two commutators at each end of the telegraph channel, as shown in Fig. 6-76. The sending commutator at one end was paired with the receiving commutator at the far end to provide signal transmission in one direction and a similar set of commutators were paired for signal transmission in the reverse direction. A single, full-duplex telegraph circuit connected the two ends and carried the signals in both directions. Each commutator consisted of a flat disk with a circular contact plate broken up into a number of segments. A motor-driven metallic brush rotated so that it passed over the segments successively, contact being made between the brush and a particular segment simultaneously[160] at the two ends of the circuit. The commutator had eight segments, six of which were used for signaling information and two to maintain synchronism by means of the start-stop principle. In modern terms, the six commutator segments used for signaling carried a six-bit code, one bit indicating the supervisory status (switchhook on or off) and the other five identifying the particular one of the 30 lines for which information was being transmitted.

Some readers may desire more technical detail. For those it should be noted that the basic idea of the distribution system was that the sending and receiving commutators were adjusted to approximate synchronism but one would always run at a slightly different speed than the other and if allowed to run freely would sooner or later get out of step. To avoid this, the brushes were stopped each revolution, after they had transmitted the six signaling bits and reached the stop segment of the commutator, and remained so until further transmission

[160] This was effectively but not literally true. Synchronization of the two ends was not exact and the receiving commutator, of necessity, had to run behind the sending by the propagation time of the circuit. The start-stop technique, described in the next paragraph, took care of both of these factors.

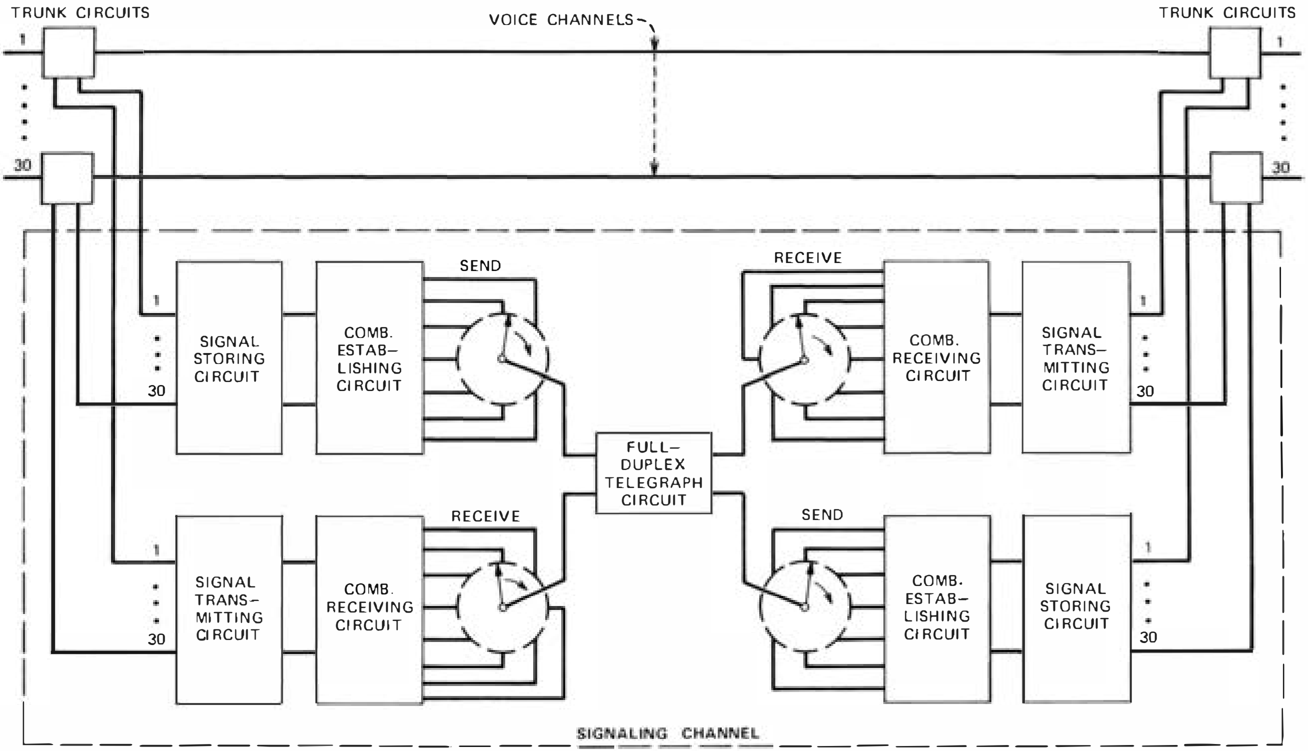


Fig. 6-76. Distributor signaling system for call-circuit toll lines, simplified circuit.

of information was required. At that time, the control mechanism initiated signals which started another revolution of the brushes. Thus, synchronism had to be maintained only during a single commutator revolution, a requirement that was easily met. It was not necessary for the commutators to rotate unless they had information to transmit, i.e., when some change in signal status occurred in one of the 30 telephone circuits connected to the system. Change of status information was provided by relay circuitry ahead of the sending commutator which stored the information until time for transmittal had arrived. Other relay circuits were used to accept the input from the store and encode the status and identity of the line. Complementary circuitry was used at the receiving end. For simplicity, the relay control and encoding circuitry is shown in Fig. 6-76 as symbolic boxes preceding and following the commutators.

5.4 Toll Switchboards

As we have noted, the methods used by the toll operator were implemented by the signaling systems just described and the switchboard systems that provided the physical means for interconnection. Toll switchboards, to a large extent, used principles and apparatus developed for the exchange plant but required additional and more sophisticated arrangements to meet the special needs of toll service. Thus, while toll switchboard development could start with basic ideas growing out of exchange plant needs, it was strongly influenced by the technology currently available for signaling and the special requirements for handling calls involving timing, delay, and other complex operating methods. As we have noted, the operating methods used for meeting toll switching requirements were gradually simplified. This simplification was brought about partly by adopting new trunking and signaling methods and partly through switchboard development which gradually produced equipment best fitting the capabilities of the operator, a process which today we call "human engineering."

A switchboard has four basic components: a set of inward jacks; a set of outward jacks; a group of plug-terminated cord circuits (with associated keys and switches) to interconnect the inward and outward jacks; and, finally, an appropriate set of indicators to show call status. The board also includes supplementary equipment such as clocks, call-circuit keys, jacks for operator telephone sets, and so forth as required to carry out the operating methods applicable. In toll switching, one set of jacks is normally[161] connected to exchange plant circuits and the other to toll trunks. Thus, the toll board provided the interface

[161] An exception was the "through" board which interconnected toll trunks. It functioned much like the tandem board used in the exchange plant which has been described earlier. These were much simpler boards than those serving as the exchange-toll interface and were used in much smaller numbers.

between these two types of plant and the cord circuits (with associated equipment) had to be designed to operate with exchange signaling systems on the one side and toll signaling systems on the other. As in exchange switchboards, the jacks in a toll board were located in a vertical backboard, the cord circuits and associated keys in a horizontal keyshelf in front of the operator, and the indicators (lamps or drops) adjacent to the jacks or plug sockets as appropriate for providing ready identification.

It was recognized early in the century that two basic types of toll switchboards were required—one for use in small, single-office cities where the volume of traffic and number of toll circuits were relatively small, and the other for use in large multioffice cities or where it was desired to handle toll traffic for a group of small cities at a single point. In the first situation, i.e., the single-office city, the toll switchboard was installed as part of the local-office switchboard lineup. This switchboard had the entire subscribers multiple appearing on the jacks in front of the operator, and she could use her cord circuits to connect the subscriber directly to a toll line. Where the separate toll board was used, toll switching trunks were run between the local and toll switchboards and used as described previously.

Early boards of the type used in the local lineup included the No. 1B, No. 2, No. 9C, and others. Since common-battery (direct-current) supervision was used toward the subscriber and alternating-current signaling toward the toll line, the cord circuit had to be designed to be responsive to both of these signals. Where toll-to-toll (through) connections were to be made at the same board, separate cord circuits were supplied for this use that were responsive to alternating-current signals from each direction. These switchboards also were adapted for short-haul A-B toll traffic. The general principles used in these boards were much like those employed in the separate toll boards such as the No. 1 type described below.

5.4.1 The No. 1 Toll Switchboard

This board was developed soon after the turn of the century and continued in use for about 25 years during which time many changes were made to adapt it to the evolving requirements. Photographs of several types of the No. 1 board are shown in Figs. 6-77, 6-78, and 6-79. The first of these is a rather sophisticated version of the board equipped for CLR operation (note the calculagraphs on the keyshelf). The second is an inward board equipped for ringdown operation, and the third is a tandem version. Neither of the latter two, of course, required calculagraphs since no timing was done at these boards.

An example of the circuitry associated with trunks and cords is shown schematically in Fig. 6-80. In ringdown operation, for which this circuit was used, the operator at the far end of the toll line called the toll operator by ringing over the line with a 16-hertz current or other appropriate signal. The toll-switchboard end of the line was equipped with a relay to respond to this ringing current and operate a line signal such as the lamps shown in the schematic. After the toll operator answered this signal by plugging a cord into the line jack, the line relay was disconnected, and the line cut through to the apparatus in the cord circuit. The cord circuit of the No. 1 switchboard had two supervisory relays. The relay on the toll-line side responded to the 16-hertz current and gave the toll operator an indication that the distant operator was recalling. The second supervisory relay responded to direct current received from the switching trunk to give the toll operator switchhook supervision of the subscriber. Other relays (not shown) permitted the toll operator to make a busy test or to use the cord for a terminating or a through connection. Keys (not shown) were included in the circuitry to permit the toll operator to connect her telephone set to any cord circuit from time to time in order to talk and listen. Keys also were included to permit her to ring on any front or back cord and to split any cord circuit, i.e., cut off either front or back cord from her own telephone set and from the other cord.

Over the 25 years during which the No. 1 toll switchboard remained the main Bell System toll board, many changes and additions were made to handle the various operating methods introduced in this period, including the single-ticket, all-circuit, and straightforward techniques. As dialing was introduced in the exchange plant (see Section IV), circuits and equipment were added so that the toll operator could receive calls from and complete calls to dial offices. However, all switching to and from toll lines was handled on a manual basis through the first 50 years of telephony.

The transmission characteristics of the toll switchboard also were improved. Special cord circuits containing telephone repeaters were developed and used to some extent to improve transmission on through connections, but were later abandoned due to the difficulty of maintaining transmission balance with 2-wire through connections. Transmission on through switches was improved by developing means for cutting off the outward and inward position toll-line multiple appearances on such connections.

Modifications were made to improve the speed of service and to adapt the switchboard to changing modes of operation. For example, there was developed the capability to operate groups of positions on a teamwork basis, whereby operators could handle traffic on adjacent positions by grouping those cords to the home position. Circuits were developed

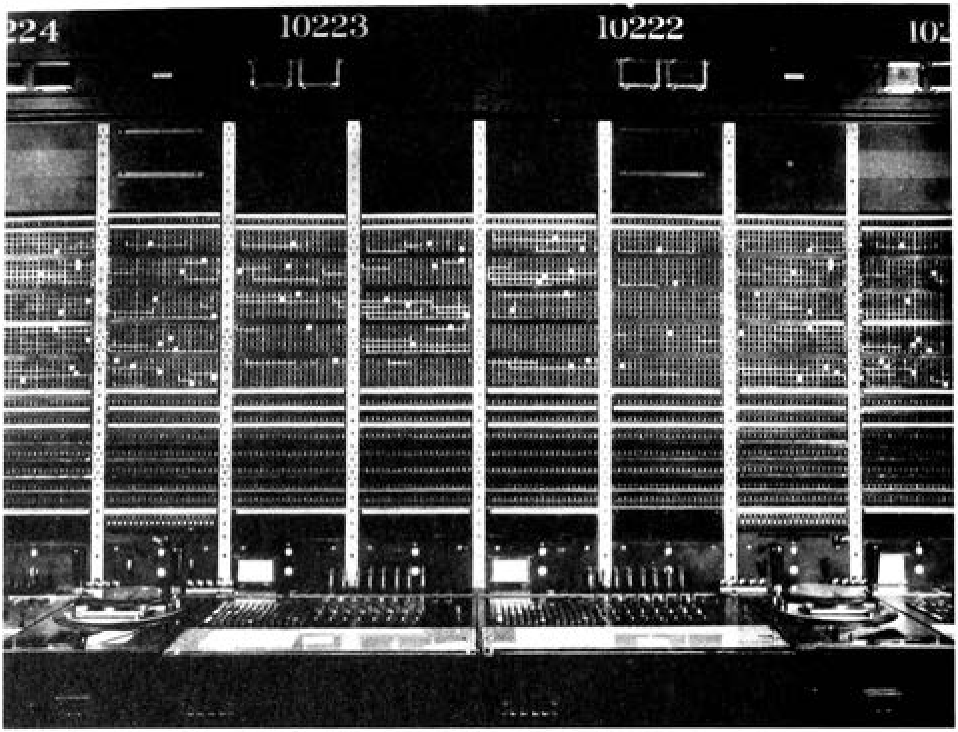


Fig. 6-77. No. 1 toll switchboard equipped for combined line and recording operation.

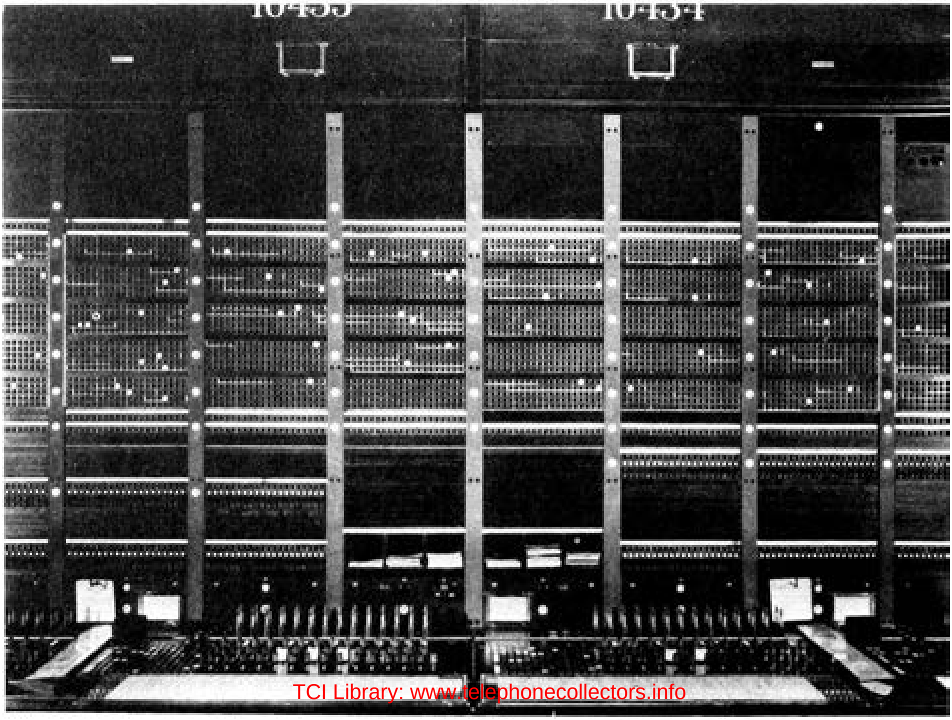


Fig. 6-78. No. 1 toll switchboard, inward board equipped for ringdown operation.

to permit the rapid transfer of inward calls, resulting from delayed operation at the far end, to outward positions for completing.[162] Busy signals were developed which automatically indicated whether the toll line was busy or not without making a busy test with the cord. These signals consisted of magnetically operated shutters which were visible to the operator in a small window associated with the line jack.

5.4.2 The No. 3 Toll Switchboard

In the early 1920s it became apparent that the continued adaptation of the No. 1 toll switchboard to changing technology had proceeded far enough, and that a new switchboard was needed not only to best accommodate the technology as it then existed but also to be designed with a view toward the future. Accordingly, such a board, designated the No. 3 toll switchboard, was developed and made available about the mid-1920s. The design of this board was aimed at a high degree of uniformity in the interface between the cord circuits and the line and

[162] When a toll call could not be completed immediately at the distant end, an operator at that point might later call back the originating office. This call, answered at the inward board, would be transferred to the outward board where the original ticket was being held. For this purpose, there was provided for each line at the inward board a pushbutton which lighted a line lamp at the outward board.

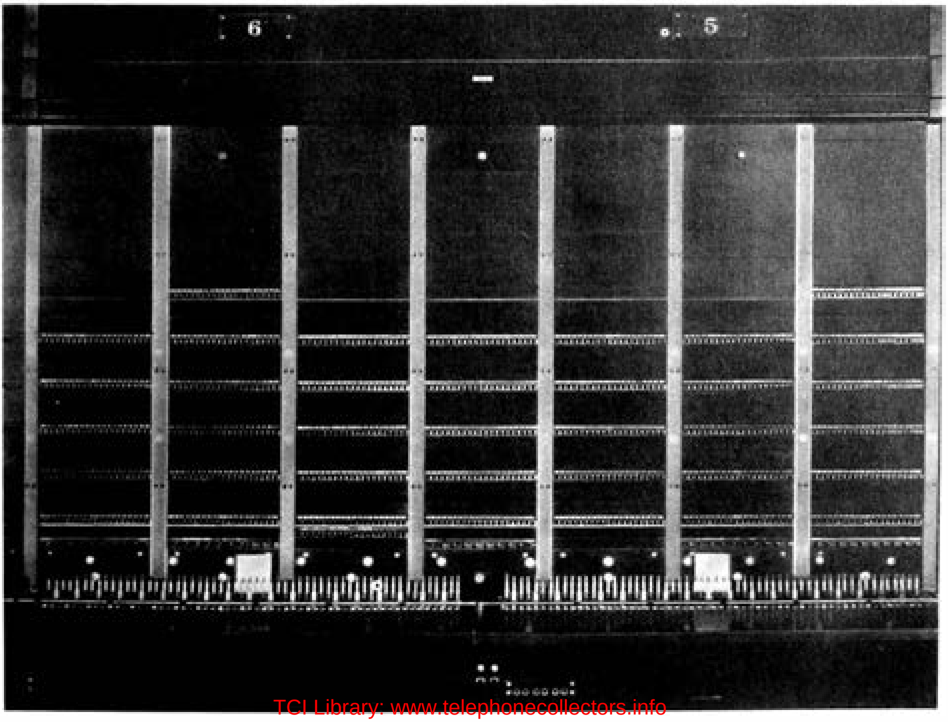


Fig. 6-79. No. 1 toll switchboard, tandem version.

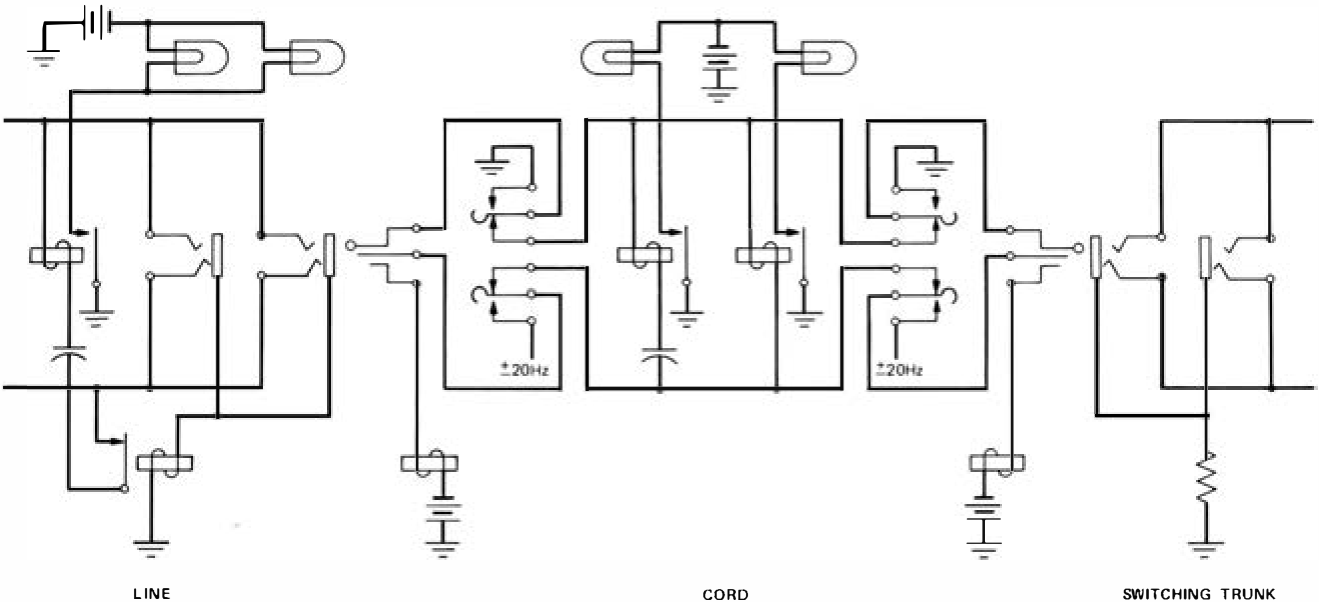


Fig. 6-80. No. 1 toll switchboard, trunk and cord circuits for ringdown operation. (Redrawn from Davidson 1927, Fig. 3)

trunk circuits. This was achieved by redistributing and altering the signaling arrangements between the cord and trunk circuits. The new board also achieved a uniform interface between the line-circuit relays and the intertoll signaling equipment, and this permitted the intertoll signaling equipment to be considerably simplified. A schematic of the No. 3 board and its connecting circuits is shown on Fig. 6-81.

L. F. Porter of Bell Telephone Laboratories patented the basic system.[163] In essence, the signaling rearrangements consisted of moving the cord-circuit ringing relays and dc supervisory relays from the cord to the line and switching trunk, and signaling between the cord and its connecting circuits with direct current over the sleeves. In addition, ringing from the cord toward the line-trunk circuits was accomplished by direct current rather than the 16-hertz ringing that had been retained in the No. 1 board, thus providing an interface between the line and the intertoll signaling circuits. In addition to the cord circuit changes, the position circuit was greatly simplified and arranged to accomplish the delayed call transfer function from the position rather than from a key per line circuit. Improvements were also made in handling calls to exchange dial offices.

The transfer of apparatus from the cord to the line and switching trunk and the use of common positional equipment made possible a relatively simple toll board suitable for inward, outward, or through operation, and easily adapted to future trends in dialing and signaling. The transfer of equipment from cord to line circuit also was of interest from the economic standpoint because in toll switching at the time there were 60 percent more cords than lines and 25 percent more cords than switching trunks.

For many years after its introduction, the No. 3 board, a photograph of which is shown on Fig. 6-82, was standard for new installations made to accommodate growth in traffic. In some cases it replaced the No. 1 toll switchboard where substantial traffic rearrangements were being made. At about the same time, another new board (designated No. 11) was developed to perform more efficiently the functions of the No. 1D, No. 2, and No. 9C boards where the toll board was part of the local lineup. Figure 6-83 shows a photograph of a No. 11 toll switchboard taken during installation.

5.5 General Toll Switching Plan

In the very early days of telephony, toll service was provided between only a few selected locations using special station equipment, connected by direct circuits of the highest grade technically feasible at the time. As technology developed the concept of a universal service came closer to realization. High-efficiency, stable telephones,

[163] L. F. Porter; U.S. Patent No. 1,608,524; filed July 24, 1924; issued November 30, 1926.

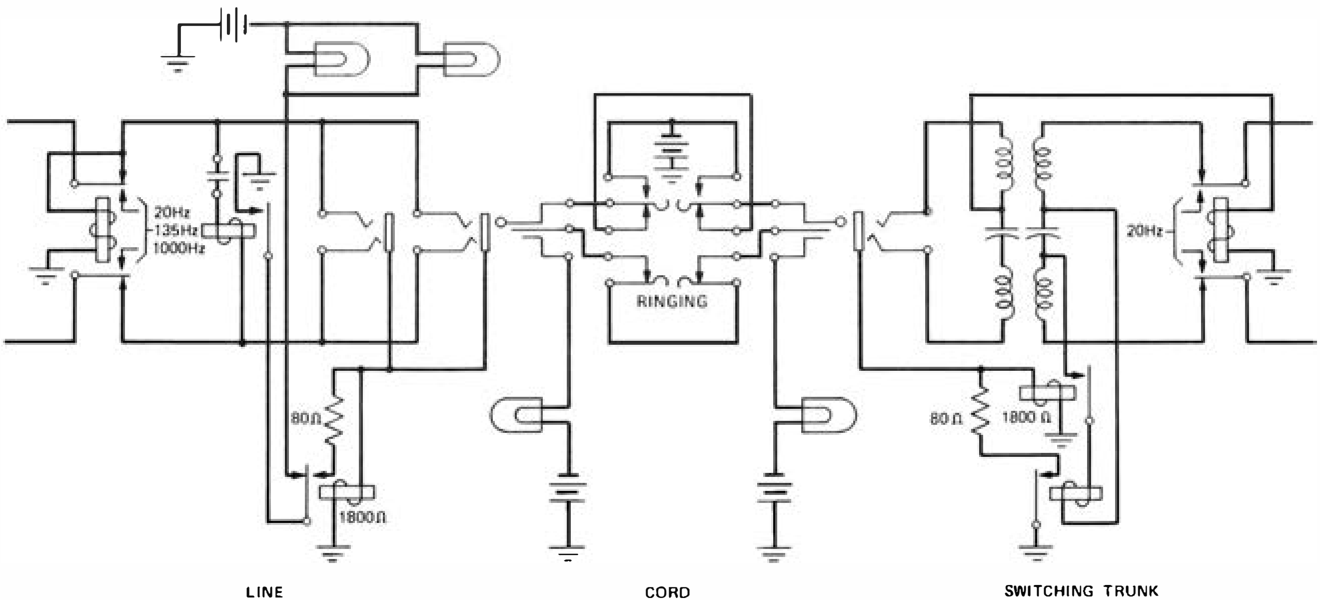


Fig. 6-81. No. 3 toll switchboard, trunk and cord circuits. (Redrawn from Davidson 1927, Fig. 4)

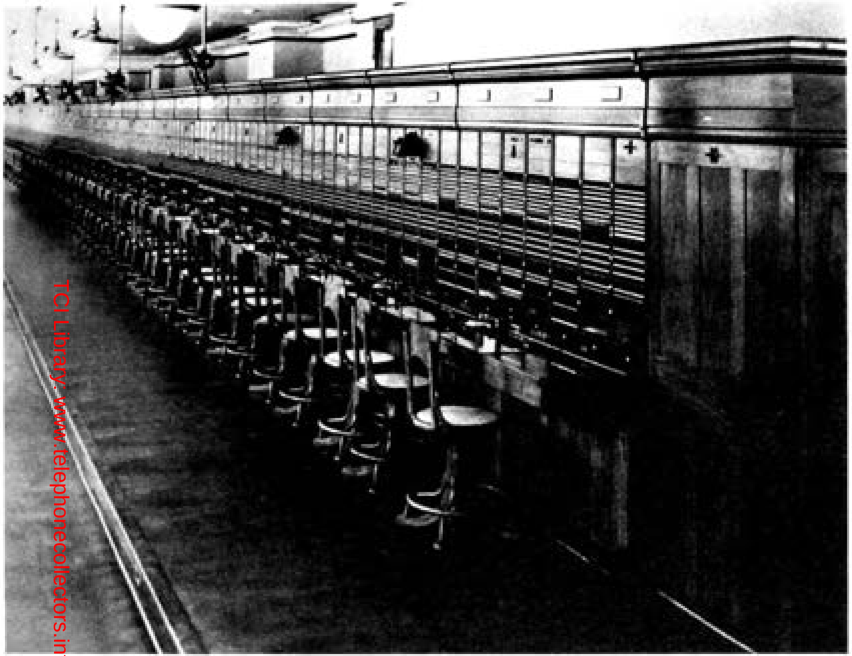


Fig. 6-82. No. 3 toll switchboard, Chicago, 3,400 lines.

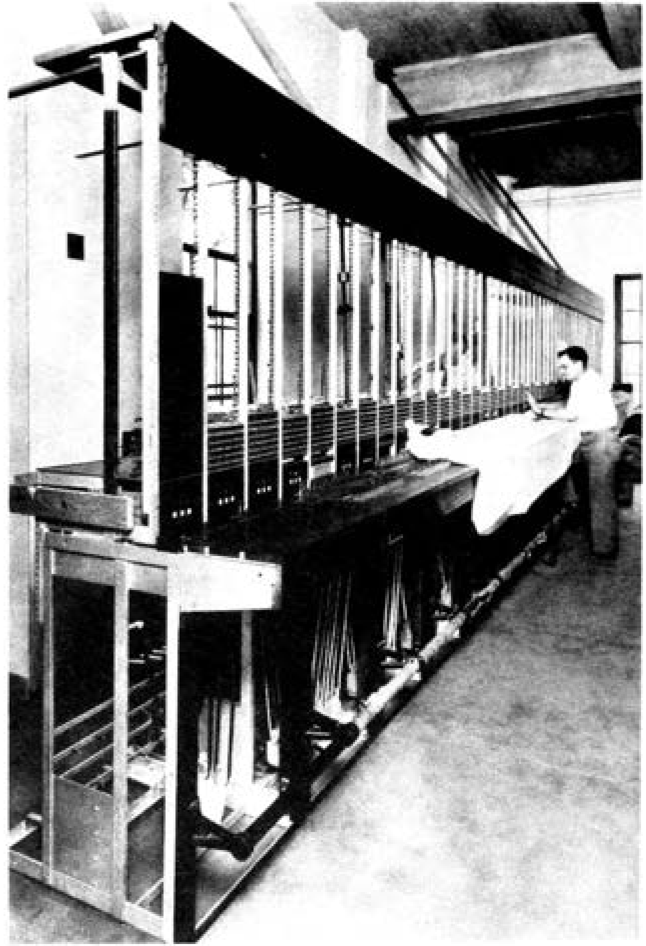


Fig. 6-83. No. 11 toll switchboard.

which could be installed at any customer location, and low-loss lines extended the area over which long-distance service could be provided. It was apparent that a need for a fantastic number of interconnections would soon develop if any customer, wherever he might be, was to be capable of talking to any other customer in the network. We have seen how this need was met in the exchange plant by the use of central offices for grouping together subscriber lines so they could be efficiently interconnected by a system of trunks. A somewhat similar solution was used in the toll plant. A major difference was that only a few pairs of cities had sufficient traffic between them to justify the direct interconnection of their toll boards. In a large majority of cases it was necessary to adopt tandem switching concepts. In these cases a toll call involved a number of lines connected in tandem at intermediate points. In establishing such built-up connections it was important to keep the number of tandem links to a minimum to avoid excessive switching time and transmission loss. It was also desirable that toll circuits be consolidated into large routes to provide economies of scale when this could be done without transmission degradation.

For these reasons the routing of a call between two cities was not a haphazard matter but had to follow a plan carefully worked out to meet the many operating, cost, and transmission requirements. As we have noted, an important part of toll operation was consulting a routing operator or some form of written routing schedule to determine the most suitable tandem switching pattern. For a considerable period the routing pattern changed frequently as toll facilities were added and experience in routing techniques accumulated. By the late 1920s, the situation had reached the point where the techniques could be formalized into a "General Toll Switching Plan," which was adopted in 1928–29. While the greatest impact of this plan was on the period beyond the first 50 years of telephony, it is nonetheless of considerable interest to this history since it represents the codification of techniques which gradually evolved during the first quarter of the twentieth century.

The general features of this plan are illustrated by Figs. 6-84 and 6-85. Figure 6-84 shows how the plan applied in a limited area such as an Operating Telephone Company or large state. A few important switching points were selected in each such area and designated as "primary outlets" (PO) which served as tandem switching points. Each toll center (city or area with a toll board) was connected directly to at least one primary outlet and each primary outlet was connected directly to every other primary outlet in the area. In addition, a few toll centers with heavy traffic between them were connected by direct circuit groups. Thus, any two toll centers in the area could be connected together with no more than three toll trunks in tandem (two PO switches) and those with high community of interest would involve only one toll trunk.

The nationwide application of the plan is illustrated in Fig. 6-85.

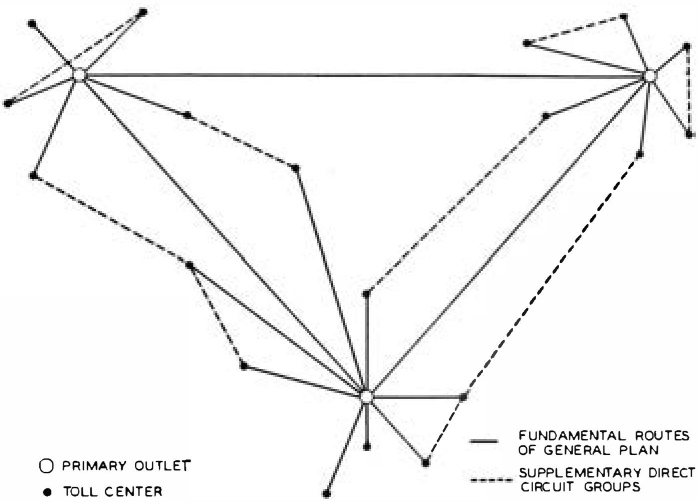


Fig. 6-84. Application of the toll switching plan to an operating area. (Osborne 1936, Fig. 22)

In order to tie together the interconnected groups of primary outlets with a minimum number of switches, a higher level was added to the hierarchy of switching points and served as the center for a large region. Each primary outlet had direct connection to one or more of these regional centers and each of the latter was connected to all others so that complete interconnection of all toll centers was possible with anywhere from one to five toll trunks in between. By 1935, the system had grown to the eight regional centers and 143 primary outlets shown on Fig. 6-86.

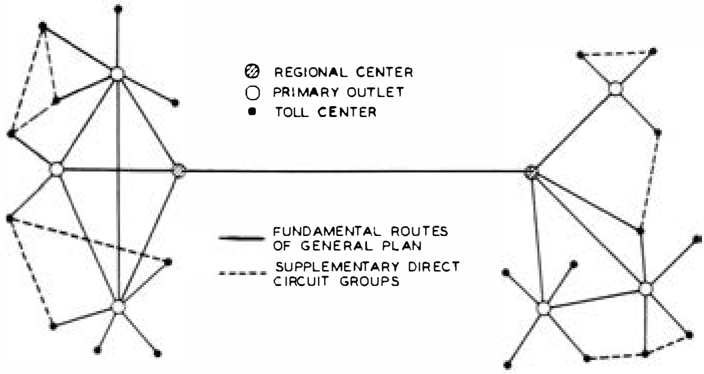


Fig. 6-85. Application of the toll switching plan to the country as a whole. (Osborne 1936, Fig. 23)

Fig. 6-86. Locations of primary outlets and regional centers. (Osborne 1936, Fig. 25)

The selection of the location for a switching center was a complicated process aimed at best meeting the requirements of good service with minimum cost and shortest practical delay due to switching and busy circuits. In most cases, principal cities were the logical choice for switching centers but other main switching points were chosen because of advantageous location such as the intersection of important toll routes. Transmission design was also a complicated matter since the trunks making up the network had to be designed for satisfactory attenuation and echo characteristics over nationwide connections involving as many as five trunks in tandem. But these matters are well beyond the scope of this history.[164] While the use of multilink connections was unavoidable for connecting small, remote toll centers with little community of interest, the steady growth of long-distance service made possible increased use of direct trunks. Figure 6-87 shows, for example, the direct circuits from Chicago to primary outlets and regional centers in 1935.

5.6 Summary

Because of the serious technical problems involved in transmitting speech over long distances, the development of a switching system in the toll plant started much later than in the exchange plant. Although AT&TCo had been incorporated in 1885 to establish long-distance service, and a few toll lines existed even earlier, it was not until the 1890s that the evolution of the toll network truly began.

This delay made it possible to use, for toll switching, many of the interconnection techniques devised earlier for the exchange plant and, as a consequence, toll switching development could concentrate largely on the problems peculiar to long-haul communication, namely, means for signaling over lines extending over considerable distance and techniques for timing, charging, and handling calls in ways that used the expensive lines economically.

Signaling methods originally employed the 16-hertz alternating current that also was used extensively in the early exchange plant both for ringing and signaling. When practical means were found for reducing the cost of long telephone calls by using the lines simultaneously for telegraph and telephone service, the 16-hertz system was incompatible

[164] For those interested in some detail, it should be pointed out that attenuation requirements were met initially by using repeaters in the cord circuits at through (tandem) switchboards. With this scheme it proved difficult to achieve the impedance balance required to give adequate repeater gain and it was superseded by an improved method using repeaters permanently inserted in the toll lines at their terminals. The amplification or gain of these repeaters was adjusted automatically as required when the lines were connected in tandem at a through switchboard. A system of attenuators or pads was used for making the necessary gain adjustments.

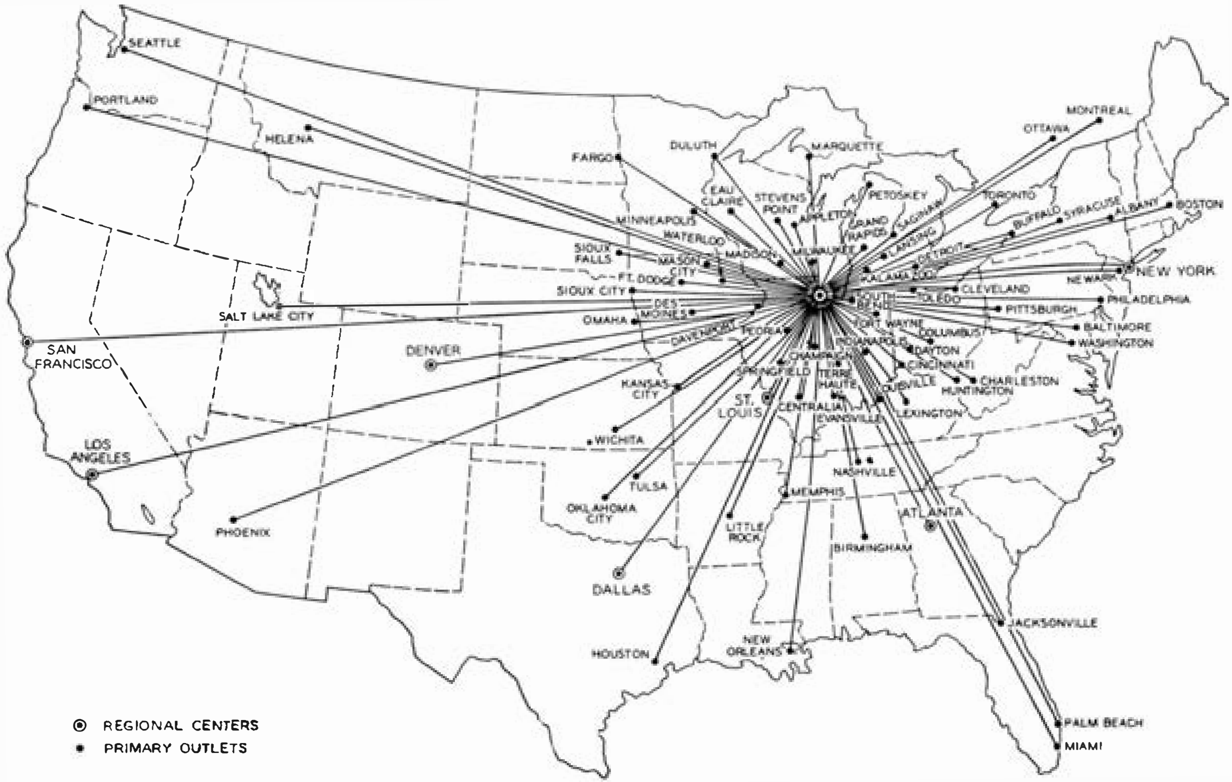


Fig. 6-87. Direct circuits from Chicago to primary outlets and regional centers. (Osborne 1936, Fig. 26)

with telegraph and was replaced by 135-hertz signaling which did not conflict with either of the communication services. Later, transmission developments such as electron tube repeaters and carrier systems were devised to achieve nationwide communications at low cost. With these facilities, low-frequency signaling paths were difficult to achieve and in the 1920s, when it was clear that future long-haul transmission systems would follow the electronic route, signaling again was changed to fit the new circumstances. One change, which was to last many years, was to signal over the voice band so that any system capable of carrying voice transmission also would be suitable for signaling. Another development of this period was to transmit signals for numerous voice channels over a single common signaling channel. This system, which employed time-division multiplexing, was not extensively used at the time but the technique had a resurgence some 40 to 50 years later and was used in the then-evolving electronic switching systems.

Toll line interconnection during the first 50 years of telephony was accomplished solely by manual techniques, the use of automatic switching methods not yet having developed to the point where they were competitive for the many complex operations associated with toll switching. At first, the operating methods employed were highly involved because of the primitive equipment available, the lack of uniform operating procedures throughout the country, and the necessity for using expensive toll lines for the greatest number of calls possible. Meeting this last objective required many operators, tandem toll lines, delayed service, and high usage of person-to-person calls.

Gradually, standardization of methods reduced the need for double ticketing, and the reduction of toll line costs encouraged the use of toll service. These factors led to expanded routes and to operating methods that more and more resembled the techniques of the manual local plant in which an operator made the physical interconnection by means of plugs and jacks. However, many of the other associated functions, such as supervision, ringing, etc., were largely carried out by automatic arrangements employing relays, self-restoring drops, switchboard lamps, and so forth.

It is important to realize that the evolution of transmission, switching, signaling, and station systems was a highly interdependent process that resulted in a spiraling (and related) reduction in cost and increase in traffic. Each improvement in one aspect of the plant usually resulted in potential for improving the others. These improvements gave better service and lower cost and stimulated more usage, which in turn led to the demand for more circuits, giving economies of scale which further stimulated the cycle.

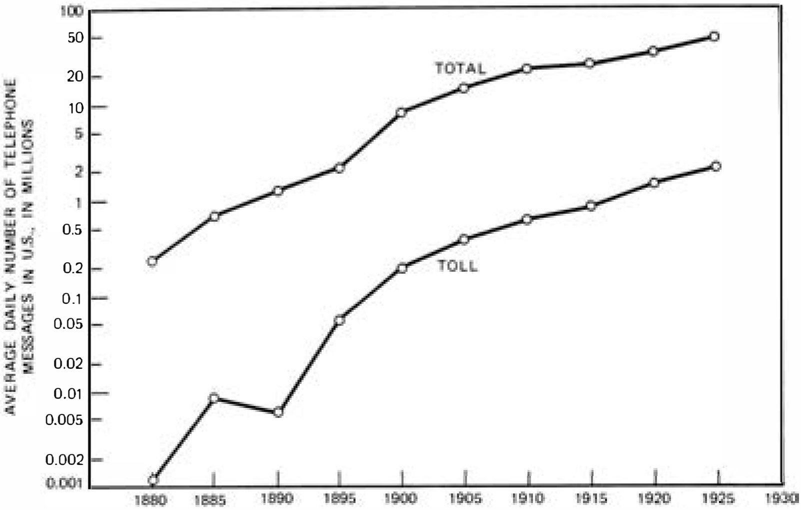


Fig. 6-88. Toll and total daily messages, 1880–1925.

Figure 6-88, which shows the growth in toll calls as well as in total telephone calls during the first part of the twentieth century, illustrates the extent to which efforts in the various lines of endeavor achieved success. Growth in all calls was extremely rapid in the years up to about 1910 but flattened off to a considerable extent after that date. In the toll area, however, the rapid expansion continued for many years later. This rapid expansion resulted from many factors but chief among them were technical advances in transmission systems, which conquered the distance and economic problems of transmission (as outlined in Chapters 4 and 5), and switching system improvements, which led to better and more convenient toll services. Before going into these matters, it may be interesting to consider the growth data shown on Fig. 6-89 for three selected routes. By 1923, the beginning of the period considered, toll service between New York and Boston had been in existence for over 25 years. Its use had become commonplace and growth had settled down to about 5 to 6 percent a year. New York–Chicago service did not become commercially practical until the early 1900s with the advent of loading and its cost remained high until electronic repeaters began to be introduced in quantity about 1920. Hence, the improvement in quality of transmission and cost had a highly stimulating effect on the growth of traffic in the mid-1920s. The effect is even more noticeable on transcontinental traffic which, while possible as early as 1914, was difficult to use and limited in quantity until transmission improvements

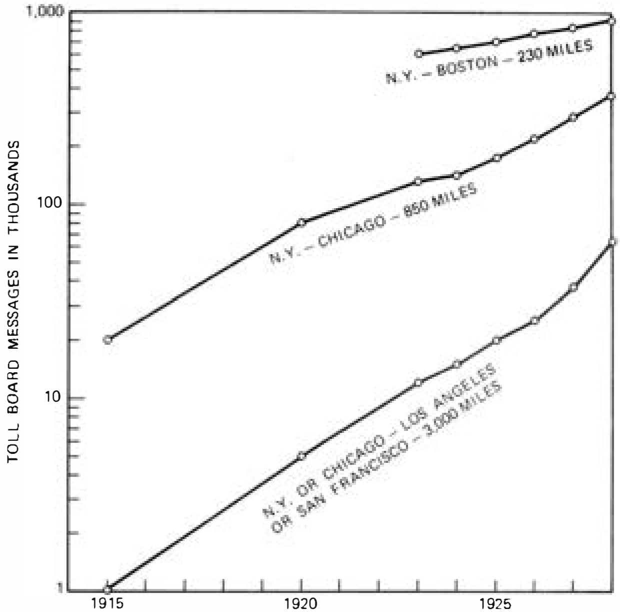


Fig. 6-89. Toll board messages per year on selected routes.

introduced in the 1920s made line construction more economical and brought about a quality of transmission more comparable to the shorter lines to which the user had become accustomed.

Throughout the entire period of toll switching development there were two outstanding objectives. One was the reduction of the time required to establish a call and the second was simplification of operating methods so as to minimize the burden not only on the operator but more particularly on the user. Figure 6-90 shows the extent to which the first objective had been achieved by the late 1920s. Early long-distance calls often required long delays before completion since the circuits were expensive and few in number and the user had to await his turn in obtaining access to those available. By 1920, technical developments reduced circuit costs to the point where large numbers of circuits could be supplied economically, and rapid reduction in service time became possible. One factor of importance was that more of the short-haul circuits could be supplied in sufficient number to be handled as A-B toll with consequent reduction in average speed of service (the measure used in Fig. 6-90). Changes in operating methods also played an important part in improving speed of service. Figure 6-91 shows quantitatively the way in which operating methods were changed. Originally, the complex

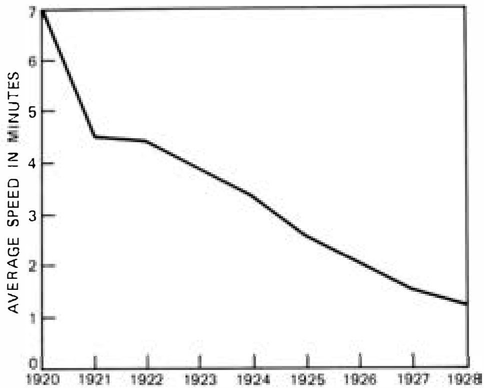


Fig. 6-90. Improvement in speed of service (all toll calls averaged), with speed of service defined as the average time required from the placing of a toll call to the response of the called party, or until a definite report is made by the operator. (Redrawn from Gherardi and Jewett 1930, Fig. 26)

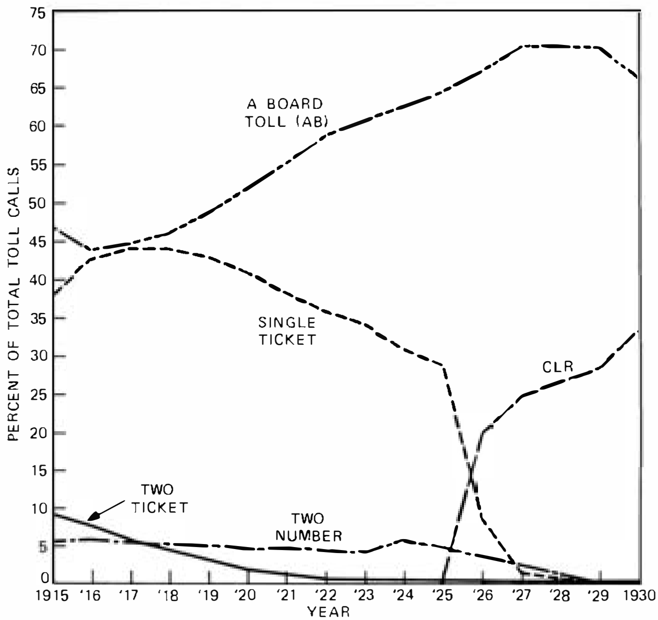


Fig. 6-91. Distribution of toll calls by operating methods at Bell-operated offices in the United States, 1915–1930. (Redrawn from Osborne 1936, Fig. 20)

two-ticket arrangement was largely used for toll board calls but by 1915 it was mostly replaced by the single-ticket method and it was phased out almost completely by 1930. Where it was impractical to use A-B toll, because of the scarcity of traffic, the two-number method had been introduced to give the user some of the benefits of A-B service by completing calls while he remained on the line. With the introduction of the CLR (combined line and recording) method in 1925 the number of calls handled while the user stayed on the line increased rapidly and by 1930 essentially all calls were handled either on an A-B or CLR basis.

The many changes in the outside plant and in operating methods had large effects on the speed with which the long-haul toll board calls were handled. In 1920, the average speed of service on these calls was 13 minutes; by 1926, this had been reduced to about 7 minutes, a large part of this time being used in transferring tickets within the office and waiting for availability of operators. Three years later, largely through the use of CLR operating, it had been reduced to 2.6 minutes.[165] It is interesting to note that the limitation of call length has never been used in the Bell System to provide increased circuit availability with consequent improved speed of service. The basic philosophy has been that, except for emergency situations, the customer should use the circuit as long as necessary to meet his needs and that reductions in delay should be accomplished as far as economically practical by providing enough circuits to meet the customer's requirements.

In conclusion, it should be noted that by 1925, when the telephone ended its first 50 years of use, the long-distance switching field was in the middle of a period of great development and growth. This was producing technology that in a few years was to mature into a solid base for a long period of toll service expansion during which the ultimate goal was achieved of low-cost service handled on much the same basis (from the subscriber's standpoint) as exchange service.

VI. PRIVATE BRANCH EXCHANGE (PBX) SWITCHING

The private branch exchange (PBX) systems described in this section were aptly named and highly useful supplements to the common-carrier networks discussed previously.

Many business and other organizations had a need for internal telephone communication and this was supplied through a small switching system, usually located on the organization's premises, and operated by one of their employees. In this sense it was a private

[165] These times are not to be compared with those shown in Fig. 6-90, which gives average time for all toll calls, many of which were handled very rapidly by A-B method.

switching system even though the equipment involved was leased from the telephone company. These same users also had need for connection to the common-carrier network and this, too, was accomplished by means of the PBX, which could connect their local extension lines to the telephone central office over PBX trunks.[166] In this usage, the PBX served as a branch switching point (or "exchange," to employ the original usage of the word). Thus, the PBX switchboard made possible both on-premise communication and also communication via the common-carrier network by means of a single telephone station set for each user location.

The PBX concept had many other advantages besides eliminating the need for separate telephone stations to serve the dual services. Since only a small part of the traffic went through the PBX to the central office, a relatively small number of PBX trunks could serve in place of the large number of equivalent loops that would have been required if each station had been individually connected to the central office. Thus, there was a considerable saving in outside plant that helped pay the cost of the PBX equipment and operator. Very large organizations, operating at several locations, often had a PBX installation at each location, all being connected together by "tie lines" which made possible interconnection by PBX switching without going through the common-carrier network, thus achieving more economical use of outside plant.

There also were many service conveniences associated with the PBX. PBX users did not need to be listed individually in the telephone directory since a single listing of the organization would be sufficient, the PBX operator serving to direct the incoming calls to the desired individual; if the particular desired person were not known by name, the operator, being an employee, could usually direct the caller to the appropriate individual.[167] On outgoing calls the operator also could be of help in locating the number and location of the person being called and establishing the call with a minimum use of the caller's time. Thus, the operator frequently served as a telephone secretary, and in recognition of the fact that her duties went beyond switchboard operating, she came to be known as a "PBX attendant." Finally, the PBX provided an opportunity for conveniently switching various kinds of special services, and considerable technological effort has been expended in effectively accomplishing this function.

[166] PBX terminology has varied somewhat over the years. The line from the PBX to the user was often referred to as a "line" with no further qualification but it was also called an extension, an extension line, or a station line. The user's station was referred to as a PBX station, an extension station, or sometimes without qualification. The line between the PBX and the central office has some of the characteristics of a subscriber loop but was usually called a PBX "trunk" since it connected two switching points. In this discussion, we have attempted to use the simplest, most unambiguous terminology.

[167] If desired, an individual's name could be listed in the appropriate alphabetic position in the directory together with the PBX number.

Although the PBX, to a large extent, is a private communications system, it must meet standards of performance that are compatible with other telephone system equipment so that the PBX, and its connected lines and stations, will in no way degrade performance when used as part of the overall network. Therefore, the Bell System approach from the beginning was that PBX communication was properly a part of the overall communications system and should be treated as an extension of the telephone network. It followed that, throughout their evolution, PBX systems were closely integrated with contemporary Bell System technology in the fields of transmission, switching, traffic engineering, and the design of station instrumentalities.

This approach benefited the PBX and other users as well as the telephone companies. Not only did both types of users obtain uniform service but PBX development could take advantage of the technological developments taking place elsewhere in the System and the development of PBX technology also could benefit the overall network. The special nature of PBX service has resulted in some unique problems and solutions with resultant technical advances of broad benefit. Among the problems was the need for optimizing designs to meet the special size and service requirements of the PBX, including installation and maintenance on customer premises. Among the technological benefits from PBX development was the opportunity the PBX offered for testing new technology on a small scale in a controlled environment. As an example, the PBX in the Western Electric Engineering Department was an important vehicle in the development and testing of early automatic switching systems. Much later (in the 1960s), PBX development pioneered time-division switching and spearheaded the development and testing of new service features of broad impact.

As with other systems, PBX switching was initially carried out manually, and the operator (or attendant) played an important role for a long time. However, automation also played a significant role when it became practical. The step-by-step and relay systems proved particularly well suited to this service. The evolution of both manual and automatic PBXs will be covered in the sections which follow.

6.1 The First 25 Years—The Formative Period

The very earliest applications of telephony, such as the pioneering work of Holmes, Doolittle, Isaac Smith, etc., were private systems but not PBXs since they did not connect with other networks. The first PBX in accord with present concepts seems to have been installed by the Dayton, Ohio, Telephone Company in 1879 at the local Soldier's Home. It had seven extension stations and at least one trunk to the local exchange.

In the same year, several private intercommunication systems were installed which were not initially connected to central offices. One of these was installed for the Northern New Hampshire Railroad and another at Columbus, Ohio. The latter employed 42 telephones and in 1880 connection to the local exchange was provided.

In 1887, the Western Electric Company had in manufacture a telephone system designed for use in large buildings which, if desired, could be connected also to the central exchange. This system had been used since about 1885 in the Western Electric Company itself, in the City Hall in Chicago, and in the offices of the Union Pacific Railroad in Omaha. The system was not a prototype of future PBX systems but contained some features of interest.

The general plan of the Western Electric system is shown in Fig. 6-92. The station set, shown in Fig. 6-93, consisted of a transmitter, receiver, and vibrating dc bell; a switchhook connected the line to the transmitter-receiver combination when off-hook, or to the bell when on-hook. Interconnection was accomplished by means of one transmission wire per station, which had an appearance at all other stations. Two additional wires were multipled to each station, one for common ringing battery and the other for common ground return. The system was claimed to be good for any number of stations, but 12 appeared to be the usual limit.

Each station had a plug connected to the telephone set. The plug could be inserted into any one of a circular array of jacks, each of which was connected to one of the other station's line wires. Each of these jacks, or plug sockets as they were called, also had a separate spring which could be pressed to connect the ringing battery to the desired line in order to ring the associated station. The manufacturer's directions to the users stated: "If you want to call another party, you take the plug that hangs at the right-hand side and insert it in the socket of the number you want, and then push the spring of the same socket, which will ring the bell of the party wanted. Before pushing the spring, however, it will be well for you to listen in your telephone to learn whether that party is using his telephone or not." At the end of the call, the plug was removed from its jack and the receiver returned to the switchhook.

An interesting feature of this system was a consulting arrangement which is still a highly stressed attribute of modern systems. The directions covering this arrangement stated: "Sometimes two persons may be talking and you may want to consult a third who is in some other part of the building. They can call him up and the three can talk together." It was claimed that the system could be connected to the city exchange if one of its instruments was in the building, "so that either of the rooms could be connected with one of the city subscribers, or a city subscriber could call and talk to one of the telephones in the building."

This 1887 arrangement probably did not exist very long. As noted in

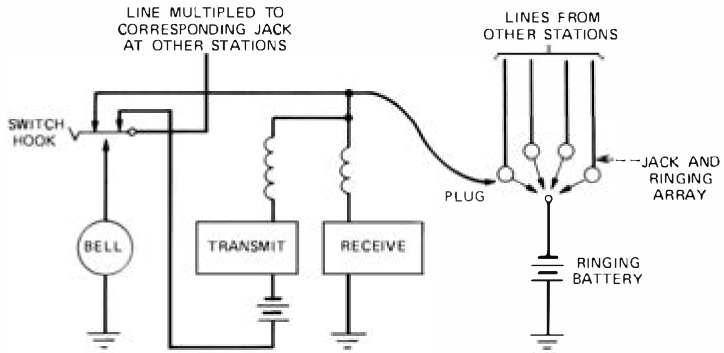


Fig. 6-92. Western Electric telephone system for large buildings (1887), simplified schematic.

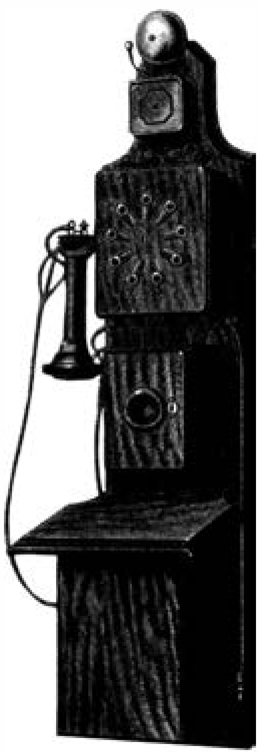


Fig. 6-93. Western Electric system for large buildings, telephone station set. (From *Western Electrician*, July 2, 1887)

the local-exchange portion of this history, about this time a strong and rapid movement toward 2-wire transmission evolved, which was highlighted in 1888 by the conversion during installation of the large Cortlandt exchange from 1-wire to 2-wire "metallic" operation. The transmission advantages of metallic operation spread rapidly, not only to the central-exchange systems but also to PBX operation.

By the end of the century, firm concepts and objectives for PBX systems were reached, in concert with the system as a whole. An idea of the state of the art can be gained from Fig. 6-94, which shows the private branch exchange network of the metropolitan newspaper *The New York World* as it existed in 1897. It will be noted that the PBX relationships were much as we know them today, consisting of a switchboard, having a number of cords, at which were terminated 27 extension telephones scattered about the various offices of the newspaper building, and a few off-premise telephones. The switchboard gave access to ten trunks to the Cortlandt central-office telephone exchange (the same one whose rapid metamorphosis had taken place less than ten years before). It was felt that the trunks were engineered liberally to reduce "busy" reports.

In 1897, there were in New York City alone, 150 such PBXs in operation, some large, as in the case of the *World*, others small, having ten or fewer telephones. Rate structures were worked out which had parameters generally similar to some of today's. Illustrations of large and small switchboards are shown in Figs. 6-95 and 6-96.

So ended the first quarter-century, with fundamental concepts well established, and with engineers working hard to codify standards and develop new and better systems for use in the twentieth century.

6.2 Standard Systems Evolve (1900–1910)

Prior to the turn of the century, the Associated Telephone Companies had designed their own equipment or modified central-office equipment to meet PBX requirements. The result was that a number of different types of switchboards were manufactured at a relatively high cost. In the century's first decade, AT&TCo began the process of standardization so that the Telephone Companies could have the benefit of soundly engineered systems at lower costs by concentrating manufacture on a smaller number of designs. Standardization involved coordinating the PBX equipment designs with the improvements achieved in central-exchange switchboards, while keeping in mind the service and economic requirements unique to private branch exchange operation. Stressed also was coordination of transmission technology and PBX traffic operation with that of the overall system.

The need for both large and small PBXs and the inherently different requirements for different sizes were recognized. Throughout this period the primary emphasis was on manual switching systems,

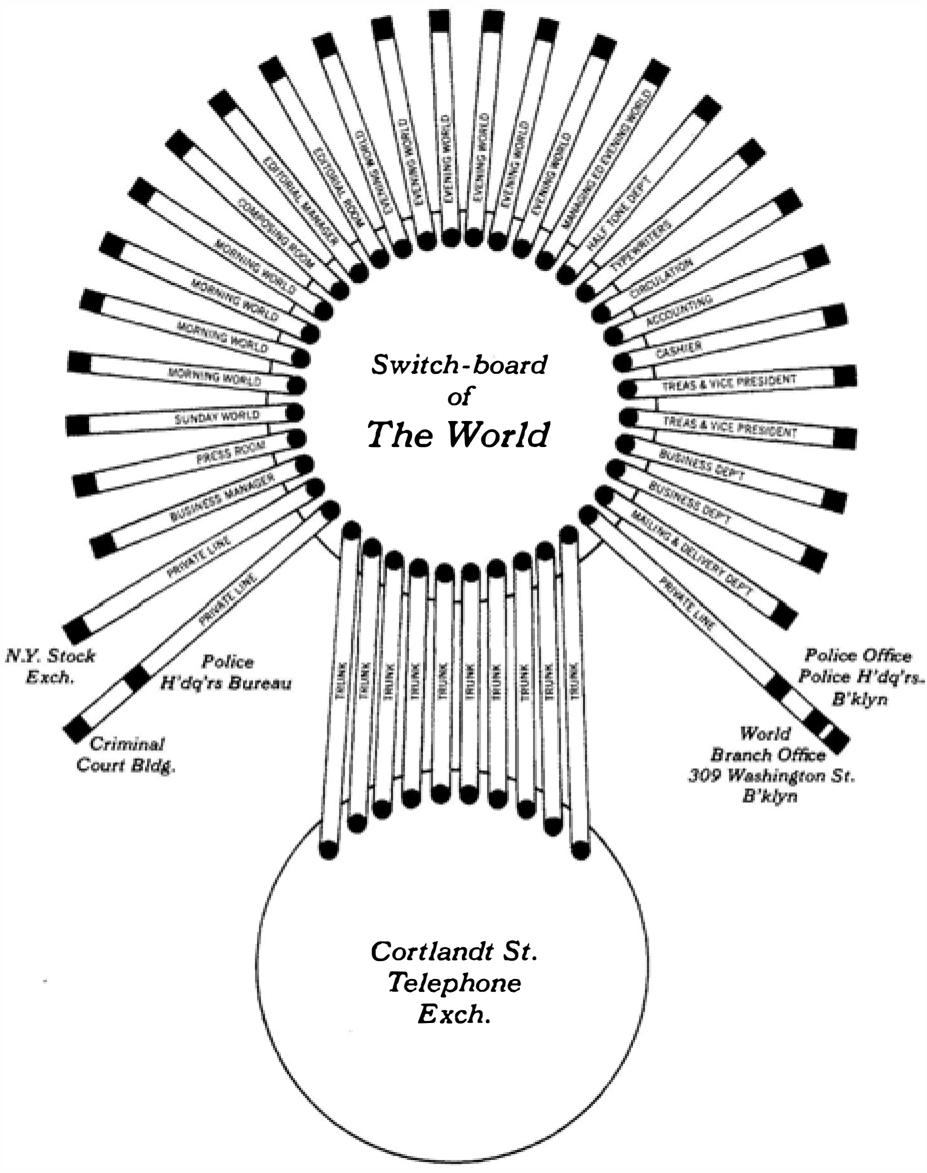


Fig. 6-94. Private branch exchange of *The New York World* and its connection into the system of the New York Telephone Company, 1897. (From *Electrical Engineer*, November 25, 1897)

Fig. 6-95. Large PBX switchboard at the H. B. Claflin Company, New York City. (From *Electrical Engineer*, November 25, 1897)

Fig. 6-96. PBX switchboard for ten or less telephones. (From *Electrical Engineer*, November 25, 1897)

consistent with the emphasis in the central-exchange area, but toward the end of the decade there began some investigation of the possibilities of dial operation.

In July 1902, AT&TCo announced its first standard PBX. Known later as the No. 1 PBX, this cord switchboard also had the distinction of being described in Bulletin No. 1, the first to be issued by the recently formed Engineering Department of AT&TCo.

The No. 1 PBX switchboard was made in two standard sizes: one for customers having up to 30 station lines, the other for use with up to 80 station lines. Both sizes had a capacity of ten trunks to the central office and ten cord circuits.

Because of rapid improvements in the art, the No. 1 PBX remained the standard for just a few years, being followed during the next 25 years by a number of boards that reflected changes in technology and customer requirements. While we cannot cover all of these boards in detail, it perhaps will be of interest to examine the features of their antecedent, the first standard PBX of 1902.

The No. 1 PBX was designed to work with standard station sets as they prevailed at the time. Station lines were terminated in cutoff jacks which, when a plug was inserted, disassociated an electromechanical line signal from the talking circuit.[168] Electromechanical signals also were associated with the trunk jacks and with the cord circuits. The trunk signal was permanently bridged so that the central-office operator could always call the PBX operator.

So-called "universal" cord circuits were employed that could be used for either station-to-station calls or station-to-trunk calls. The distinction between the circuitry involved in the two types of calls lay in the source of talking battery—the cord itself when connecting station-to-station, and the central office when connecting the station to the trunk. In the latter case, the station was on the same transmission and signaling basis as main stations with respect to the central office. The electrical state of the trunk sleeve caused the cord to determine which battery feed should be used. A signal associated with the cord circuit provided called-party supervision to the PBX operator on station-to-station calls and station supervision on trunk calls. Listening keys in each cord circuit permitted the PBX attendant to associate her telephone set with any cord. Ringing keys permitted ringing on either end of the cord; alternating current for ringing could be supplied by a hand generator or from the central office. DC power for the PBX was supplied from the central office, and a technique was developed for engineering the number of cable pairs required to furnish adequate current as a function of conductor gauge, distance to the central office (up to 10,000 feet), and number of simultaneous cord-circuit connections. A simplified schematic of the PBX is shown on Fig. 6-97.

Very soon it was recognized that there was a need for standardizing an economical PBX for situations where the traffic was not enough to justify a full-time operator. Accordingly, the No. 2 PBX was announced by AT&TCo in October 1903. It, too, warrants detailed description since it incorporated new principles, some of which have had a long life.

This PBX, unlike the No. 1, did not use cords for making connections at a central switchboard. Instead, it used keys, mounted on a wall or desk near each station, which could select any other station or make connection to one of the trunks to the central office. As such, the PBX was the forerunner of our present key telephone systems, which today have the selection keys built into the telephone set.

The Engineering Bulletin covering the No. 2 PBX pointed out that poor service often resulted on small cord-type switchboards because of lack of attention on the part of an attendant who also had other duties. The Bulletin continued by stating that:

> Private branch exchange No. 2 furnishes a system of intercommunication whereby a subscriber at any station may call any other station of the branch

[168] Several types of electromechanical signals, sometimes called "drops," are described in Section 3.4.3. Lamp signals were not used in PBXs until about 1904.

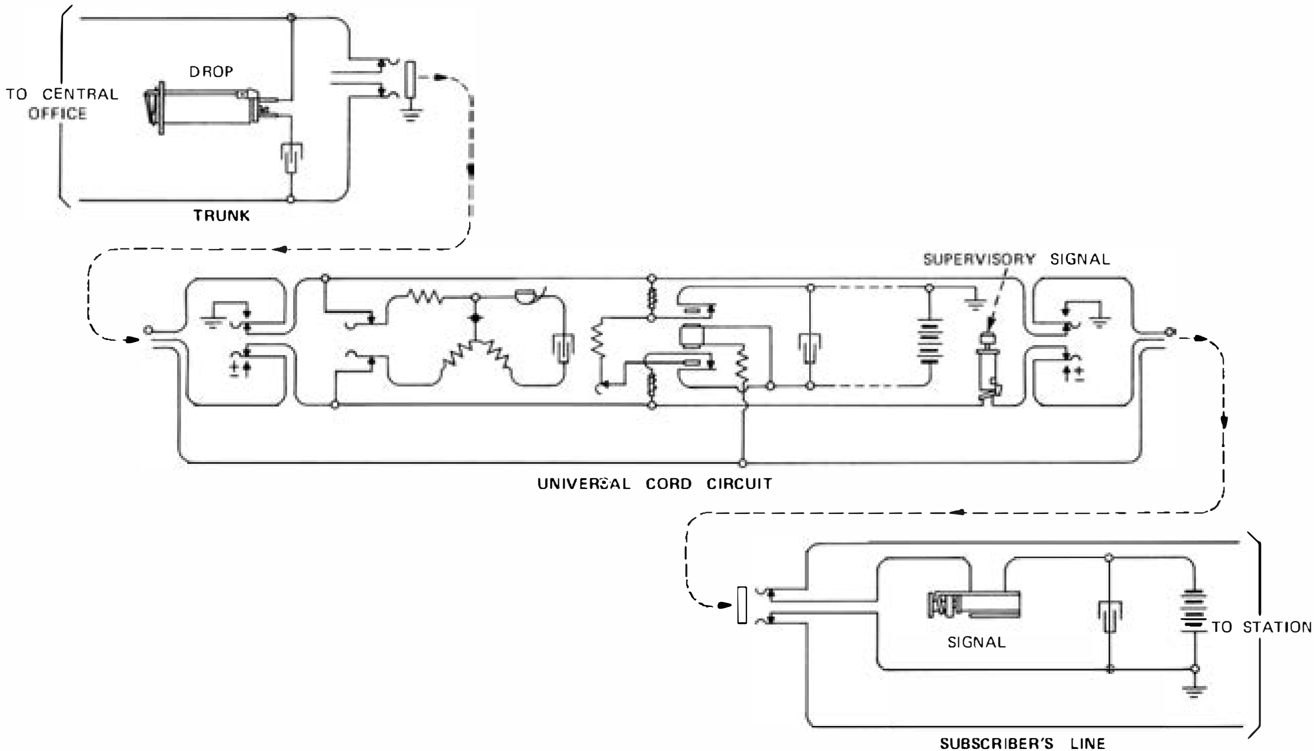


Fig. 6-97. No. 1 PBX, simplified schematic.

> exchange without the assistance of an operator. Facilities can be provided for central office connections with either common battery or magneto switchboards. In such cases no private branch exchange operator is needed for out-going calls from the private branch exchange, but for incoming calls from the central office, services of an attendant or clerk at one of the private branch exchange stations, arranged as a receiving station, will be required to distribute such calls to the proper stations. There is no delay from slow disconnections at the private branch exchange, the supervisory signal at the central exchange being under control of the private branch exchange subscriber who is talking.

One implementation of the No. 2 PBX is shown in Fig. 6-98 and a simplified schematic of the system is shown in Fig. 6-99. (In the latter figure, each of the horizontal and vertical buses represents a pair of wires.) The functions of the system were basically as follows: To make an intercom call, station 3, for example, would call station 1 by depressing key S1 (which locked into the connect position) and operating the ringing key. This rang the bell at station 1, who answered by depressing the answer key. When the person at station 1 took the receiver off the switchhook, the bell was disconnected. Other stations could intercommunicate in a similar manner. Any station could make an outgoing trunk call by going offhook and depressing a trunk key. An incoming trunk call was brought to the attention of one of the stations, who acted as an attendant, by ringers (not shown) associated with the trunks. The attendant would answer the call by depressing the trunk key, and then would determine the station desired and signal it; that station then would connect to the trunk designated by the attendant.

It will be noticed that this PBX resembled the Western Electric system of 1887 (shown in Figs. 6-92 and 6-93) not only in the use of station switching but also in the signaling plan. Although the earlier system was a 1-wire transmission system, whereas the No. 2 PBX was 2-wire, the intrastation signaling in each system was accomplished with direct current over one of the wires to operate a vibrating dc bell. Connection from one station to another was made in the Western Electric system by inserting a station plug into a jack (connected to the other station's line) at the originating station; the No. 2 PBX accomplished this by depressing a locking key that made similar connection.

The type of switching matrix used in the No. 2 PBX usually is attractive when the number of interconnected lines is small but it becomes cumbersome from the standpoint of wiring and key arrays and also becomes uneconomical as the switching complex becomes larger. This occurs because, in the extreme case where all stations have full access to each other, the number of key switching points is the square of the number of stations, whereas with simple cord-switchboards the number of jack switching points is equal to the number of stations. The cord-switchboard advantage is diminished, of course, by the expense of the operator and the limiting effect of the switching links (cord circuits).

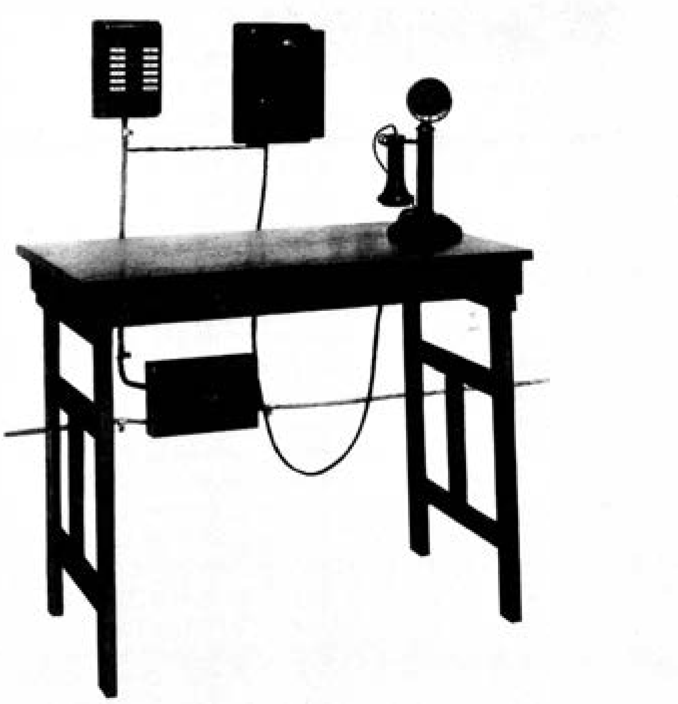


Fig. 6-98. A No. 2 PBX station equipped with a No. 472A key assembly mounted on the wall and a No. 20CN deskset on the table.

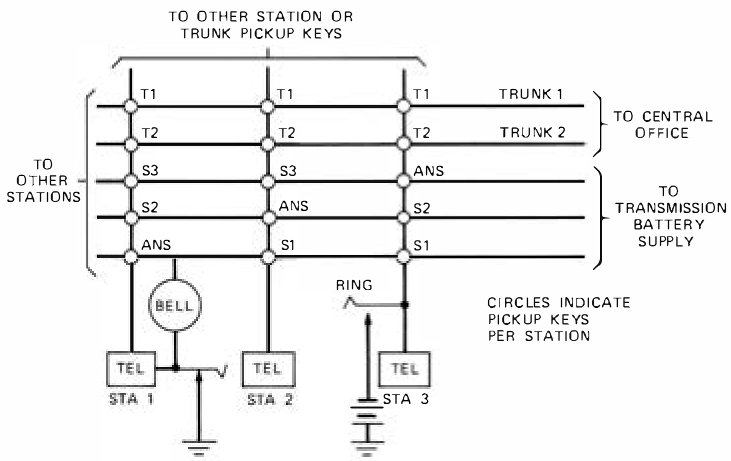


Fig. 6-99. No. 2 PBX, simplified schematic.

By 1904, the Bell System was stressing the importance of integrating PBX service with the System as a whole. In this year AT&TCo, through Engineering Bulletin No. 11, outlined to the Associated Telephone Companies practices which AT&TCo felt should be observed at the central office and at the private branch exchange in order to secure satisfactory operation of the PBX. The parent company pointed out: "Poor private branch exchange operating has a detrimental effect upon exchange service as a whole. For this reason it is absolutely necessary that the Telephone Company should maintain a general supervision over all private branch exchanges. Much improvement can be accomplished in private branch exchange operating if the Telephone Company will encourage private branch exchange subscribers to apply to it for experienced operators for the more important private branch exchanges, or better in some cases, to hand over to it the entire charge of the operating of such exchanges. The Telephone Company could give careful instruction in operating methods to such inexperienced persons as may be put in charge of private branch exchange switchboards." The Bulletin then proceeded to describe the purpose and advantages of PBXs in essentially the same terms that we see them described today, and as already outlined in the introduction to this section of our history. It also pointed out the desirability of avoiding apparent advantages which were really pitfalls: "One of the principal advantages the private branch exchange system offers to the businessman is that feature of operating which permits the subscriber to give a call to his private branch exchange operator and then to hang up the receiver until the called subscriber is at his instrument ready to talk. This feature, however, interferes most seriously with the character of the telephone service as a whole. If an operator is obliged to handle many calls of this kind, it is impossible for her always to remain on a connection until the called subscriber answers, and in consequence that subscriber, upon answering his call, will find no one on the line to speak to him. Such a system of operating can lead to confusion when an attempt is made to establish telephone connection between two subscribers, neither of whom will have come to the telephone until the other party is on the line. Such a system of operating is to be strongly advised against by the Telephone Company." Apparently, some of our less-desirable telephone customs are of long standing.

Outlining some positive actions which should be taken to ensure good coordination between the PBX and the central office, AT&TCo said: "At the central office the auxiliary trunks to the private branch exchange should terminate in consecutive jacks . . . when the private branch exchange is one of considerable size, some economy can be effected by providing separate trunks for incoming and outgoing calls . . . the method of night service which has proven the most satisfactory in practice is to connect to the trunk lines those stations at which it is desired to receive or originate calls during such hours." Operating objectives such as indicated above helped to stimulate technological advances in the early years and for that matter many of the objectives are still valid today and continue to stimulate the art.

In the same year, 1904, a new cord-type switchboard was developed and standardized to replace the No. 1 PBX. The new PBX was known as the No. 4[169] (a No. 3 PBX had been designed in the meantime to meet the special traffic requirements of hotels and similar large installations). The general appearance of the No. 4 switchboard is shown in Fig. 6-100.

Taking advantage of device and circuit improvements made in the general switchboard art, the No. 4 board provided lamp signals on extension lines, trunks, and cords. These signals replaced the electromechanical devices used on the No. 1 board. Signals, which

[169] Some PBX historians have stated that the No. 4 was the first standard PBX. This impression possibly was gained because the No. 1 was quite short-lived, particularly when compared with the longer life of the No. 4, and quite probably had very little publicity beyond the early engineering bulletins issued by AT&TCo.

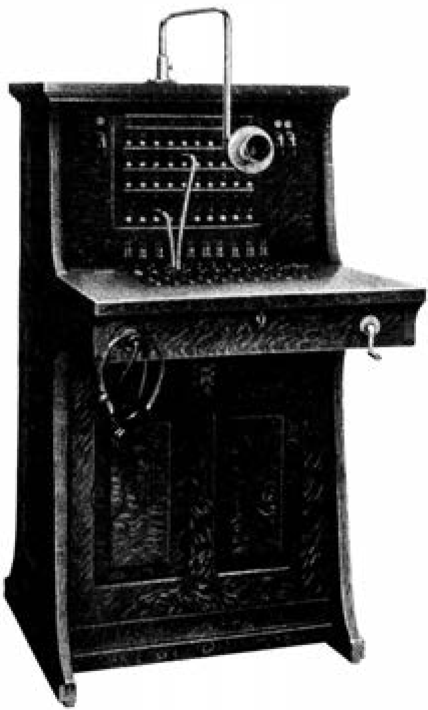


Fig. 6-100. No. 4 PBX switchboard (30 lines).

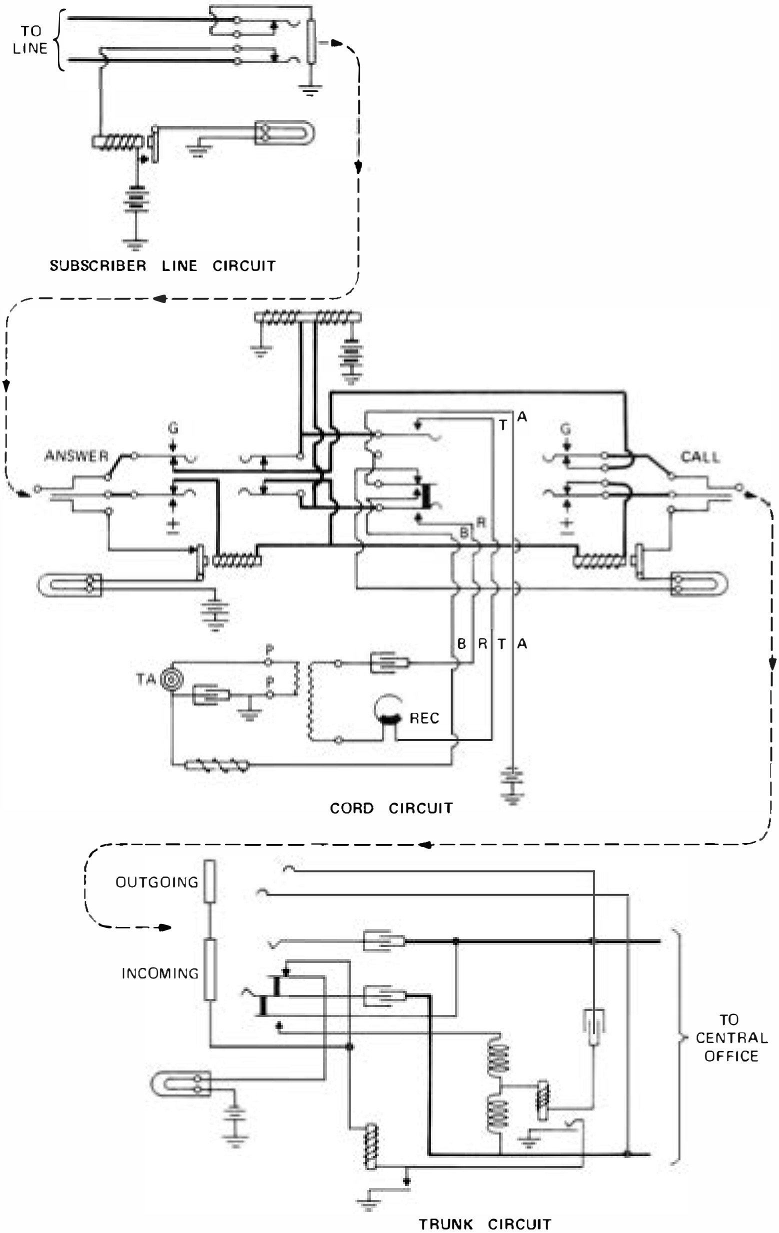


Fig. 6-101. No. 4 PBX circuits.

separately indicated the supervisory status at each end of the cord, replaced the single signal of the No. 1 cord, thus providing more position supervision. In addition, the transmission circuit was improved to provide a higher degree of transmission efficiency. By this time, PBX circuitry, following that of local-exchange boards, had become somewhat complex, largely as the result of simplifying operating procedures. For the reader who is interested in the circuitry employed, Fig. 6-101 shows some simplified schematics.

After the introduction of the No. 4 PBX, a number of changes were made in its circuits and equipment to keep it up-to-date with other branches of the art, and its capabilities were increased so that it could cover efficiently a wide range of customer sizes. By the end of the decade, a line of PBX switchboards was being manufactured by the Western Electric Company which provided capacities from 30 to 1,500 extension lines. These boards were grouped into two types. One was a single-position (non-multiple) type for situations where the traffic could be handled adequately by one operator. This board was engineered for 30, 80, or 320 extension lines. A second type, which provided for multipling the circuits before more than one operator (the multipling technique being adapted from the general switchboard art), took care of situations where more than one operator was required to handle the traffic. Two versions of the multiple board were made available, one with a capacity for 640 lines and the other with a capacity for 1,500 lines; these are illustrated in Figs. 6-102 and 6-103. The resemblance to a small exchange switchboard is obvious.

For much the same reason as the No. 1 PBX was supplemented by the No. 2, the initial No. 4 cord switchboards were supplemented by a PBX switching arrangement designed to meet the needs of customers having very few lines and whose traffic was insufficient to justify a full-time operator. This was a key-controlled switching system that did not require plugs and jacks. Unlike the No. 2 PBX, it did not employ switching at the stations but was a true centralized switching arrangement with all of the equipment contained in a small cabinet. The cabinet could be mounted on a desk at which the attendant could do other work as well as perform the minimal switching functions. This PBX, a photograph of which is shown in Fig. 6-104, was introduced in 1907. Called the No. 505, it was the prototype for a succession of cordless PBXs, which with improvements were manufactured until the mid-sixties.[170] Indeed, the objective of cordless switching control was carried over to the auxiliary

[170] Lest the reader feel that suddenly a whole host of PBXs between the No. 4 and the No. 505 have slipped into technical oblivion, we hasten to point out that a three-digit identification system beginning with the 505 was the start of a systematic numbering plan which would help to keep track of the PBX genealogy for generations to come. The chart shown on Fig. 6-114 outlines this plan.

Fig. 6-102. No. 4 PBX switchboard (640 lines). (From WECo Bulletin No. 1018)

Fig. 6-103. No. 4 PBX switchboard (1,500 lines). This eight-position multiple board, installed in the Boston City Hall Annex, had 80 trunks to the central office. (From WECo Bulletin No. 1018)

control of small dial switching PBXs and persists to this day in trim, exotic desk-top "consoles" associated with modern electronic systems.

The No. 505 PBX was designed to serve up to seven stations and three trunks to the central office. As a "first" of the cordless type, it seems worthy of some detailed description. As shown in the switching plan, Fig. 6-105, all connections between stations or between stations and trunks were made by the attendant via links which served much the same purpose as cord circuits. The attendant connected a station to a link by operating a key, and then with another key connected the same link to another station or trunk. Although this was a key system like the No. 2 PBX, the use of a centralized link-switching arrangement controlled by an attendant reduced the number of switching points in the No. 505 PBX to the product of links and stations (plus trunks) as contrasted to the larger number of switching points required by the more or less stations-squared array of the No. 2 system.

The family of No. 4 PBX cord switchboards, and the No. 505 cordless switchboard just described, set a lasting pattern for efficiently and economically handling the traffic of small, intermediate, and large customer groups. The No. 4 PBX and the original version of the No. 505 remained the standard until 1916 but basic technical approaches and the size pattern prevailed, with some modifications, well beyond that year. Naturally, many of these boards remained in service long after 1916 since, as with other telephone equipment, newly standardized designs represented merely a change in manufacture and were compatible with earlier designs, which remained in plant for their useful life.

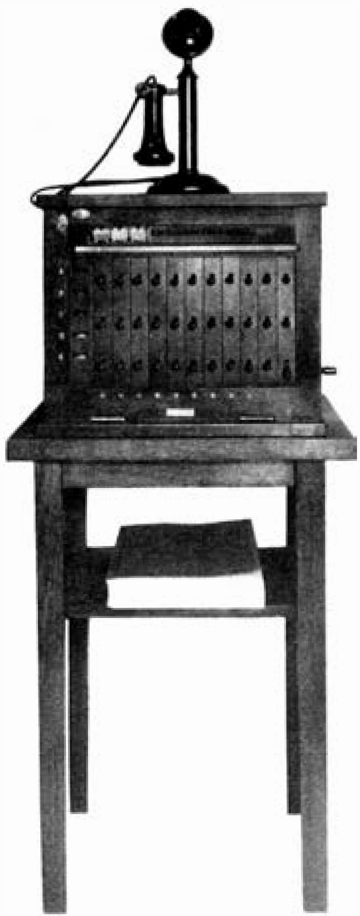
6.3 Completing the First Half-Century

By 1910, the basic concepts underlying PBX service had solidified, and a line of manual systems had been developed that met the needs of the time, ranging from the small cordless board with a maximum of seven extension lines up to multiposition cord-type boards handling up to 1,500 lines.

The principal development work remaining for the next 15 to 20 years was concerned with improving and extending the capacity of manual boards through the application of evolving technology, adapting manual boards for use with the dial exchange system then being introduced, and finally applying automation to the PBX itself. This work is covered in the following sections, which carry this history a few years beyond our nominal objective of 1925 to describe developments actually under way as the first 50 years of telephony ended.

6.3.1 Manual Systems

Replacement of the No. 4 PBX, which had been the mainstay for a dozen years, was begun about 1915 with code series No. 550. These



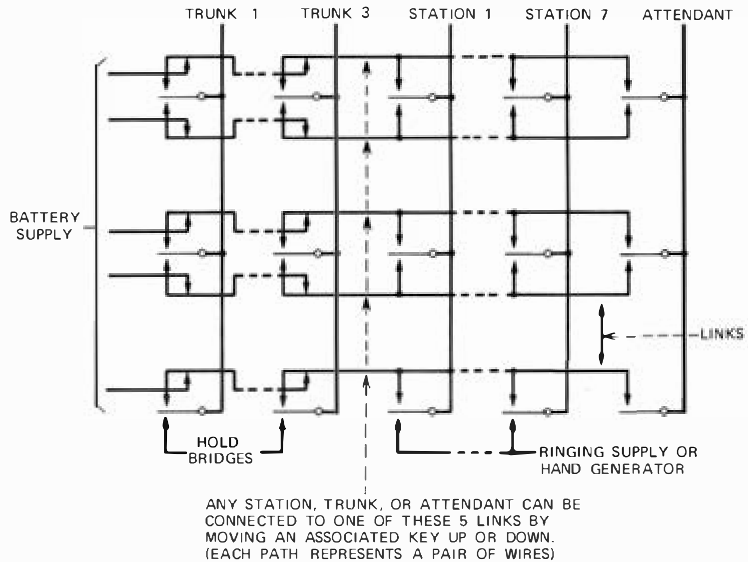


Fig. 6-105. No. 505 PBX switching plan.

Fig. 6-104. No. 505B PBX switchboard. This cordless board, installed in the Boston offices of Ferguson, Loud and Ferguson, had seven station lines and three trunks to the central office.

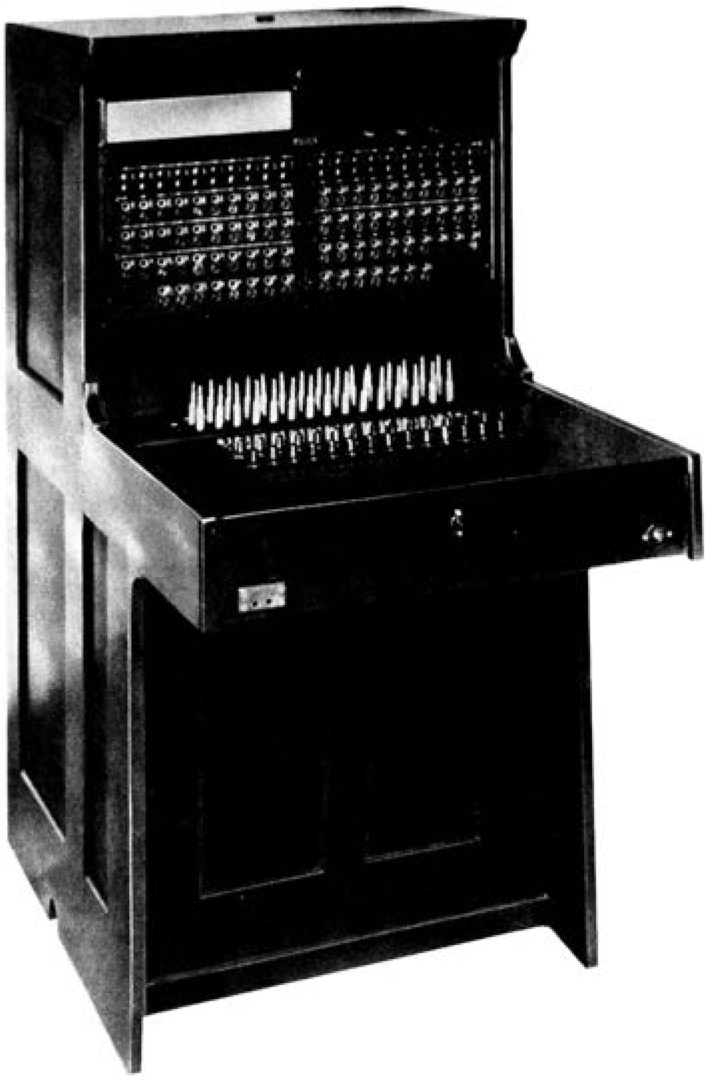


Fig. 6-106. Typical 550-type PBX switchboard of 1916. (Abbott 1968, p. 10)

boards were similar functionally to the non-multiple versions of the No. 4 board but improved in construction and appearance. A typical code-550 board is shown in Fig. 6-106. The complete code-550 series consisted of the following:

Code No.	Introduced About	Size and Features
550A	1915	30 and 80 lines
550B	1916	30, 80, 320 lines
550C	1920	30, 80, 320 lines. Adapted for use with dial offices (attendant and station dialing)
551A&B	1927	40, 80, 320 lines. Replaced the manufacture of the 550 boards with improved equipment arrangements

The No. 505, seven-line cordless board also was improved in 1916 and adapted for use with dial offices about 1920, receiving the codes 505B and 505C respectively. In 1928, the manufacture of the 505 board was supplanted by the 506A and 506B for 7 and 12 lines respectively. The circuit and operating features were essentially the same as the 505C but incorporated improved equipment arrangements.

The No. 4 board continued to be used for installations of over 320 lines until 1922 when the 600C board was introduced for installations up to 640 lines. This switchboard was more economical than the No. 4 due to substantial reductions in power made possible by the design of more efficient circuits and the use of new, more sensitive B- and E-type relays. It was the first of a new family of boards in the 600-code series. Since it was apparent by 1922 that dial exchange offices would soon be used widely, PBX boards designed at this time and subsequently were designed to work with both manual and dial central offices. In the latter case, the board was designed so that those stations that were equipped with dials could dial their central-office numbers through the cord-circuit connections established by the attendant (through-dialing). The board also was designed so that the release of the central-office connection on outward calls could be controlled optionally by the originating station (through-supervision) or by the attendant (nonthrough-supervision). Both the through-dialing and the through-supervision features were intended to reduce the work of the attendant (they also speeded up the service and reduced unproductive central-office holding time) but required that the customer relieve the attendant of the responsibility of supervising these calls.

The complete line of multiple boards in the code-600 series, developed in the 1920s, is described briefly below:

Code No.	Introduced About	Maximum Size and Features
600C	1922	640 lines
604C	1925	2,000 lines

605A	1928	Replaced the manufacture of all 600C, and 604 boards in sizes up to 1,500 lines
606A	1929	5,000 lines—used small (No. 92) jacks

A 600C board is illustrated in Fig. 6-107 and its successor the 605A in Fig. 6-108. This latter board provided a wide field of use with a single type of switchboard since it could be provided with any number of positions required to meet service requirements up to a maximum of 1,500 lines, thus accomplishing economies in manufacture, stocking, installation, operation, and maintenance. With the completion of the 5,000-line 606A board the requirements of practically all customers were met by a family of PBXs starting with sizes as small as seven lines.[171]

Superficially, there may seem to be a lack of order in the plan followed in PBX development. The No. 4 board was a rather successful attempt at providing a "universal" board meeting maximum requirements of 30, 80, 320, 640, or 1,500 lines. This was followed by a seemingly piecemeal attack resulting in boards covering only a limited range of sizes and mostly of small capacities. However, there were good reasons for this approach stemming from a combination of factors. First, the PBX had to be designed for installation and maintenance on a customer's premises with installation completed as rapidly as possible and with minimum interference with the customer environment. Further, the PBX equipment had to be flexible enough to accommodate frequent moves in and out of a customer's premises—particularly with the small and intermediate sizes. These restrictive requirements favored the tailoring of the PBX to particular size limits in order to make the equipment as economical as possible. Why then was not a family of switchboards developed all at once to meet each segment of requirements? The answer here probably lies in the need to place priority on high-demand areas which required the small- and intermediate-size PBXs. For example, during the period 1910 to 1926 more than 100,000 PBXs were manufactured by the Western Electric Company. Of these, 87 percent were about equally divided between the No. 505 (capacity of seven station lines) and the No. 550 (capacity limits of 30, 80, or 320 lines), 9 percent were No. 2 and No. 4 types under 300-line size, and only 4 percent of the demand was for the larger PBXs of the types developed in the 1920s. It was essential, of course, to develop these large PBXs, even though they were statistically few, because of the importance of this field and the large revenues involved; but the pressure of high demand emphasized the need for maintaining high manufacturing efficiency by frequently up-dating the smaller systems.

[171] Occasions could arise where a 5,000-line capacity was not adequate. These would be few and would often involve special requirements so that a custom design was a logical solution in such cases.

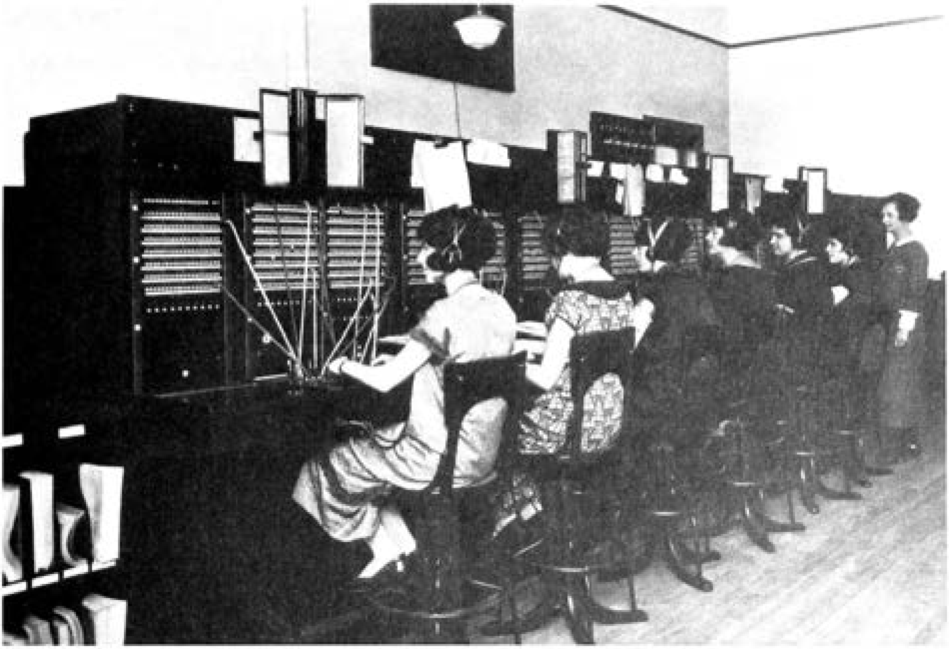


Fig. 6-107. A six-position No. 600C PBX switchboard, equipped with 480 station lines and 40 trunks, installed in Copley Plaza Hotel, Boston.

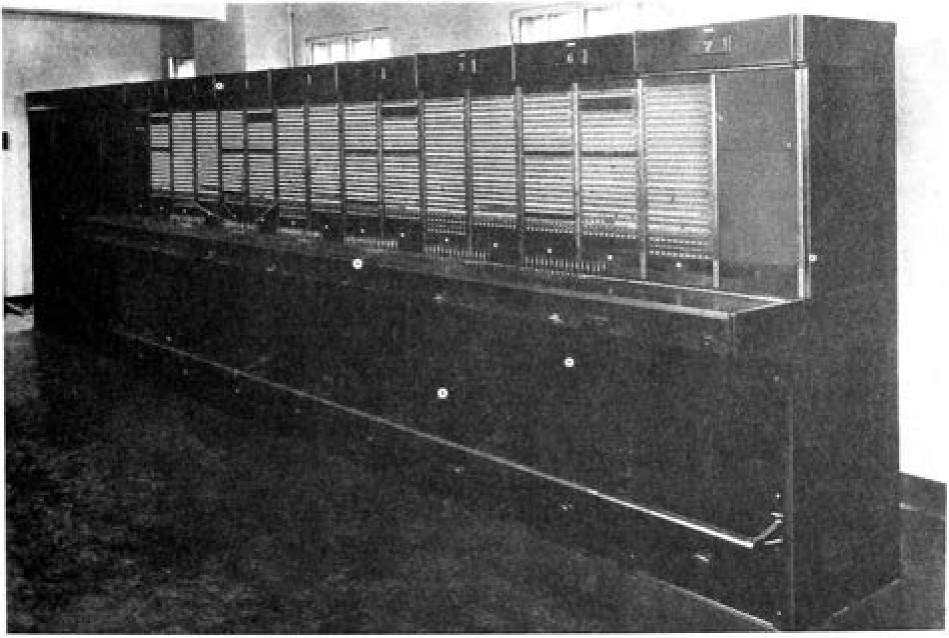


Fig. 6-108. No. 605A PBX switchboard in American Women's Clubhouse, New York City. (R. W. Harper, *Bell Laboratories Record* 1929, Fig. 1)

6.3.2 Automatic Systems

Since PBX systems were based very largely on the same technology as exchange systems, it is only natural that automation also should be considered at an early date. For example, the 1904 "Queens System" described in Section 4.1 was used as an intercommunication or PBX system at several locations including Dartmouth College and the U.S. Artillery School. While customer reaction was good initially, these systems apparently could not show significant benefits and their life as PBXs was short.

During the next dozen years after the Queens System trial, the major applications of automation to Bell System PBXs were as trials of technology being developed for exchange plant switching, commercial service during this period being furnished by manual boards.

As related in Section 4.3.3, the common-control systems being developed by Bell were put on trial in the Western Electric PBX at 463 West Street in New York City (later the home of Bell Telephone Laboratories). Initially, this installation used rotary switches, and it was placed in service on November 29, 1910. This was primarily a laboratory experiment, but since it involved bonafide service and the gathering of customer reaction data as well as technical data, it could also be thought of as a PBX "first."

In the West Street rotary trial, about 400 extension telephones in the Engineering Department were served on a semimechanical basis; that is, calls placed by the stations were answered at a switchboard, whereupon the operator advanced the calls, either to another station or over a trunk to a central-office operator, through the sender and selector-controlled switching mechanism. Incoming calls from outside were also answered at a switchboard (made necessary because these calls were all from manual offices), whereupon the inward operator advanced the call to the desired station via the automatic mechanism. In September 1912, panel-type switches were added to this trial to provide more facilities and to make a practical trial of the new equipment which by that time appeared to be economically more favorable than rotary switches. In this addition, all calls from one position were handled by panel-type equipment consisting of district, office, incoming, and final selectors. In order to provide a direct comparison of the systems, both rotary and panel connectors had access to the same 100 lines. It was estimated that the rotary equipment handled 7,000 calls per day and the smaller amount of panel equipment 2,000 calls per day.

During this trial, the design of the apparatus was carefully studied to detect defects, and modifications were made to insure improved operation and provide measures to compensate for variations in manufacture and to facilitate maintenance. The satisfactory results of the PBX trial, together with the results of economic studies, solidified the preliminary plans to proceed with large-scale central-office field trials of the panel system on a semimechanical basis. It had little effect on the automation of the PBX system since the panel system was designed to meet the complex problems of metropolitan areas, whereas the much simpler requirements of the PBX could be met with systems no more complicated than those required for a small town. Such systems using the step-by-step (S × S) switch were available by 1915 but prudence dictated one more try at developing a more efficient switch, particularly for use with large PBXs. Accordingly, in 1915, development was resumed on a hybrid S × S-panel system that had been started in 1913 for central offices but subsequently discontinued.

The objective of the S × S-panel system was to provide a switching multiple as efficient as the panel bank, but without the power-driven controls of the panel system. The mechanism consisted of a vertical rod with ten brushes located at different points on the rod in a spiral relationship and mounted in front of a panel-type bank consisting of ten groups of contacts with ten sets of contacts in each group. A brush was located opposite each group of bank contacts. Any one of the ten groups of bank contacts was selected by rotating the rod through the proper angle to place a brush in contact with the proper bank and then the particular set of bank contacts was selected by vertical movement of the rod. In November 1916, a PBX embodying these mechanisms was cut over at 463 West Street. After operating successfully, it was removed in 1917 and reinstalled to serve the Post Office Department in Washington, D.C. It remained in service there until mid-1934. Another system was installed at the Hotel Commodore in New York City in 1919, where it was used until 1929. No further applications of the system were made since, while the term of service indicated this system was technically satisfactory, the cost could not be justified.

As this work was going on, the advantages of dial operation, particularly for large PBXs, became increasingly apparent because of rising labor costs and the increased speed and privacy of intercom service. In 1916, some privately owned S × S dial intercom systems, whose customers wished to coordinate them with Bell System central offices, were purchased by the telephone companies and development work was carried on to arrange these systems to operate satisfactorily with the Bell System plant.

By the middle of the following year, ideas on the mechanization of all forms of switching had solidified and were covered in a letter of July 24 from Bancroft Gherardi to J. J. Carty who was then Chief Engineer of AT&TCo. The recommendations for exchange switching in cities of various size already have been covered in Section 4.3.3. The plan for PBXs covered the use of a semimechanical arrangement which is of particular interest since it established the operating plan that persisted for many years. This plan was: "Local calls, that is, calls within a private branch exchange, are completed automatically. Calls to the central office are made by connecting with a trunk from the PBX to the central office by means of a single movement of the dial,[172] the calling party then giving his call directly to the central office operator; or else any extension may call the attendant by one movement of the dial[173] and have the attendant make the outgoing call as at present. As an alternative, any extension station may be connected to a line lamp on the PBX and make all calls as at present on manual PBXs. Incoming calls will be received by the attendant at the PBX and distributed by her." In forwarding the recommendation to top management, Carty added these points, which illustrate some of the thinking of the day: "It is thought that most of the private branch exchanges will be better served by manual switchboards, but wherever all of the conditions indicate a mechanical system, we will be prepared to provide one. The most favorable conditions for the mechanical PBX are found where a 24-hour service is given, where much of it is within the PBX itself and where the PBX is relatively large." Quite obviously, these conditions were the most favorable since they permitted the largest operating savings with which to offset the dial equipment cost. Only small operating savings were realized on traffic external to the PBX since, at the time, most of these calls required the use of an attendant.

As noted previously, Western Electric obtained patent rights to the S × S switching system in 1916 and placed their first order for a small-city exchange system in 1917. In this same year, Bell also adopted the S × S equipment for PBXs. As with exchange systems, the first equipment was obtained from the Automatic Electric Company under Western Electric specification. Early installations were for the DL&W Railroad at Hoboken, New Jersey, the Emergency Fleet Corporation in Philadelphia, and the War Department in Washington, D.C.; these probably were stimulated by the increasing communication needs of the country as it became involved in World War I.

The step-by-step PBXs followed more or less conventional lines. Calls originated by stations were detected by line switches, and selectors routed these calls to the switchboard on the first pull of the dial if it were an outgoing central-office call. Station-to-station calls were routed via selectors and connectors to the stations. Incoming calls were answered at the switchboard and the attendant completed the call directly to station jacks appearing in the switchboard.

During the start-up period of mechanizing PBXs in the Bell System several versions of the step-by-step system were included in the "700"

[172] The numeral 9 was adopted and is still used in most PBX services to obtain an outward connection to the exchange plant.

[173] The numeral 0 (Operator) was adopted for this purpose.

series of codes. These systems bore various subcodes to designate the special purposes for which they were employed. However, they used practically the same fundamental technology, with some alterations in the basic switching plan. A recitation of their characteristics will be of little interest except possibly to note that the 700A code was applied to the previously mentioned private systems that were taken over by Bell and adapted to PBX service in 1916, the 700B utilized the 550B switchboard as an attendant's position, the 710A was used in 1918 for several railroad installations that did not connect to central offices, and the 720A was for the use of railroads requiring long ringdown toll lines.

Thus, during the second decade of the twentieth century the first steps toward PBX automation had been taken but this service was still supplied preponderantly by manual systems. However, customer interest had been aroused in dial PBXs and the next ten years would bring a complete line of standardized dial systems aimed at making this service more useful and economical over the entire range of customer sizes and to function expeditiously not only with the manual exchange plant but with the increasing number of machine offices.

The mounting customer interest in dial PBX service was stimulated by the growing awareness of the properties of dial systems: the improved speed, accuracy, and availability of intercom calling among the PBX customer's employees and also the expansion of outward dialing capability as dial central offices became more numerous. Thus, with dial PBX operation, the PBX attendant was not only dispensed with on intercom calls but it was also possible to bypass her on calls to other customers outside the PBX reached through the dial central office. But the case was not all one-sided. First, the increased cost of the dial equipment and its installation required higher rates to be paid by the customer. As far as he was concerned this had to be offset by savings in PBX attendant costs (the customer paid these) and by the service and intangible advantages of dial service. Second, the PBX attendant was necessary for distributing calls from outside to the PBX station; her knowledge and judgment of what organizations or people should be connected to the inward call was considered a cardinal PBX asset. Indeed, some people felt that this was valuable even on internal calls. Finally, the PBX customer was not always willing to give all of his station users unrestricted access to the central office. To the extent that this unwillingness prevailed, the outward dial advantage was diminished until systems were developed which permitted the automatic application of calling restrictions at the customer's option. Under the scheme ultimately developed, some stations could be given free access to the central office, whereas attempted outward calls from restricted stations would be diverted to the attendant who would then handle them in accordance with the customer's company policy. For example, she could keep track of the charges, and had the options of denying or forwarding the call. To meet these requirements the PBX had to be designed so that the attendant could dial the call for the station user, or alternatively he could dial through her cord connection, a feature known as "through dialing."

With these broad objectives in view, the first standard dial PBX system, the No. 700C, was developed about 1922. It used step-by-step selectors and connectors together with plunger-type line switches similar to those used in central-office systems at the time. It was the first Bell System standard dial PBX arranged to operate with manual and dial central offices, and it replaced the 700A used in the latter part of the previous decade, which interfaced with manual offices only.

Similar to the 700A, the 700C PBX served intercommunicating calls between stations by means of step-by-step switches. It also permitted establishing outgoing calls to the central office either by dialing the central office directly or by dialing the PBX attendant and having her establish the connection. In the latter case, calls could be dialed by the attendant or she could operate a "through-dialing" key and the call could then be dialed from the station. All incoming local and toll calls from the central office and all outgoing toll calls were handled by the attendant. Outgoing calls to other PBXs reached over tie lines could be established either through the step-by-step switches without the aid of an attendant at the originating point or they could be routed via the switches to the attendant who then would complete them with her cord equipment via the tie line. The manual switchboard associated with this system had jack appearances for the extension lines, tie lines, and PBX trunks to the central office. The attendant used her cord circuits for interconnecting the jacks as required by the appropriate procedure for handling the call. The attendant's switchboard was a modification of the 550-type PBX for installations less than 300 lines and of the 600-type PBX for larger installations up to 1,000 lines.

The switches of the 700C could be engineered for two-digit, three-digit, or combined three- and four-digit operations in accordance with the number of digits needed to identify the station on intercommunicating calls. This permitted the system to serve small customers (under 80 lines) or customers with up to 1,000 lines or more, with a flexible array of switches. A number of features which today are considered desirable also were provided with this system in the year 1922—hunting, night service, restricted service, intercept on vacant connector terminals, machine ringing, and conference and code calling. Dial-repeating tie lines (which permitted stations to directly dial stations in distant PBXs) were under development when this PBX was first introduced. According to one authority, the first two 700C PBXs were installed for the Ford Motor Company in Detroit and in Windsor, Canada, and were connected together with tie lines. Approximately 100 of these PBXs were manufactured.

After obtaining experience with the 700C, the development of an improved step-by-step system was started in 1926. In 1929 this system, coded the 701A, superseded the 700C. It used the same basic traffic plan as the 700C but had a capacity for 3,200 station lines. Line finders of the 200-point type replaced the plunger-type line switches of the 700C and the system contained improved switches and equipment arrangements which reduced installation effort and facilitated stocking of the equipment by the Western Electric Company, factors which in turn reduced the overall time between customer service request and actual service. The 605A switchboard, previously described as a manual-only switchboard, was adapted to function with the 701A, broadening its field of use.

This system remained as a standard step-by-step PBX for 30 years and continued thereafter, with equipment improvements, as the 701B until the present day.

At about the time the 700C became the first standard Bell System dial PBX, a method of operation known today as "Main-Satellite" was being evolved for the Bell System. This type of operation was desirable, and still is, for those customers who have two or more locations, such as a general office location and a remote plant, each of which may "home" on a different central office. Under these conditions, the customer may wish the two units to function as an integrated PBX. In those cases where it is not economical to wire all of the station lines to a single switching system, a separate switching system is provided at each location and these systems are connected together by tie lines. Intercom calling between locations was dialed directly by the stations over the tie lines (after dial-repeating tie lines became available) and outgoing traffic was completed generally by direct dialing to the associated central office. The principal technical difference in Main-Satellite operation arises from the handling of inward traffic from the central office. With this type of operation, only the main location receives inward central-office calls; the attendants at this location distribute them either to stations homing on the main, or via low-loss tie lines to stations on the satellite. For this purpose and other reasons the 710C was developed for use as a satellite, replacing the previous 710A and 730A which had served manual central-office areas only. Later, at the time the 701A was standardized to replace the 700C, the 711A replaced the 710C and used the same technology as the 701A, i.e., line finders instead of line switches and improved selector and connector switches. Like the 701A, the 711A remained the standard for many years thereafter.

Although the 700C and 701A step-by-step PBXs were technically capable of serving customers of all sizes, they were not economically attractive for small installations because the saving in attendant costs could not pay for the higher equipment costs relative to a manual PBX. Since an attendant was required to handle inward calls on a dial PBX and a full-time attendant could take care of a manual PBX with about 100 lines, the greatest operating saving to be expected with a small dial system was the part-time effort of a single attendant. The service advantages of dial operation still remained, of course.

Therefore, development work was started to provide small dial systems that would be more competitive with manual systems than would result from a trimmed-down 700C or 701A. Two systems were the initial result—the 740A, a small system which permitted partial attendant savings, and the 750A, a very small system requiring no special attendant at all.

The development of the 740A, started in 1924, was completed in 1926. Its target was the customer size defined by two-digit intercom dialing. Numerically, two decimal digits consisting of a tens and units digit would provide for 100 numbers, but the capacity of the 740A was a little less—88 stations in fact—because it was necessary to set aside numbers to reach the central office and the attendant and for test purposes. Consisting of conventional step-by-step switches, with 100-point line finders and combination selector-connectors, the 740A design resulted in lower cost per line because, by restricting its range, a more or less pre-engineered factory-wired switching frame as well as a simpler power plant could be produced. These features led to a system requiring less floor space and promoting rapid installation.

The capabilities and service features of the 740A were largely the same as the 700C and 701A with the exception of the switching equipment for the attendant. This, a new dial PBX concept for the Bell System, was a cordless-type control cabinet similar in general appearance to earlier manual cordless switchboards. It was used instead of the conventional cord switchboard to distribute inward calls and to handle the occasional outward calls. The trunk circuits were terminated on keys and the attendant established connection to the station by dialing the extension number, thus obviating the need for extension-line jack appearances and cord circuits as was usually required at the conventional cord switchboard. The attendant's services were not needed for disconnect as the release of the PBX switches was controlled jointly by the restoration of the switchhook at the called extension and disconnection at the central office. These features resulted in a system whereby the attendant could perform other duties, thus resulting in partial savings. The attendant cabinet used with the 740A is shown in Fig. 6-109. Some readers may be interested in the methods used for switching various types of calls at this stage in PBX development. For these, the routing plan of the 740A is shown in Fig. 6-110.

The success of the 740A led to the development, in 1928, of two similar but smaller systems with a maximum capacity of 38 lines. Fifty-point line finders and 50-point selector-connectors were used in both systems

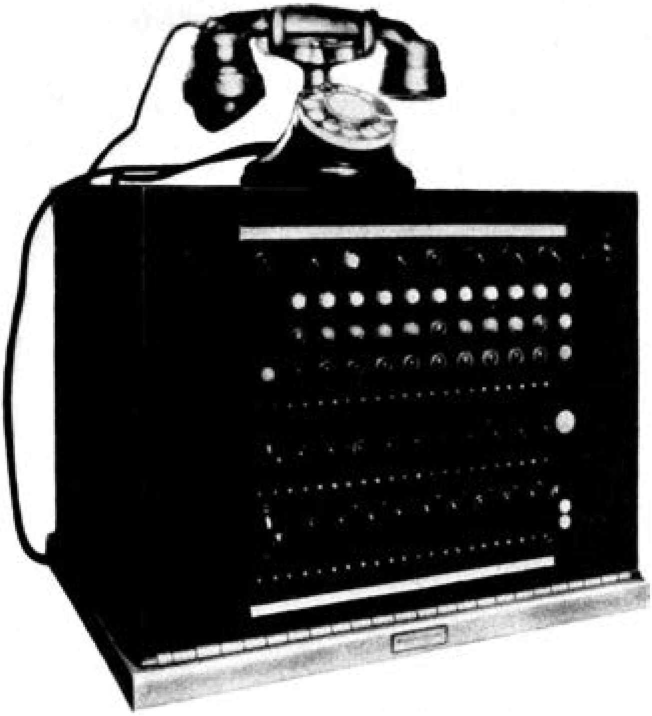


Fig. 6-109. No. 740A PBX attendant cabinet. (Abbott 1968, p. 12)

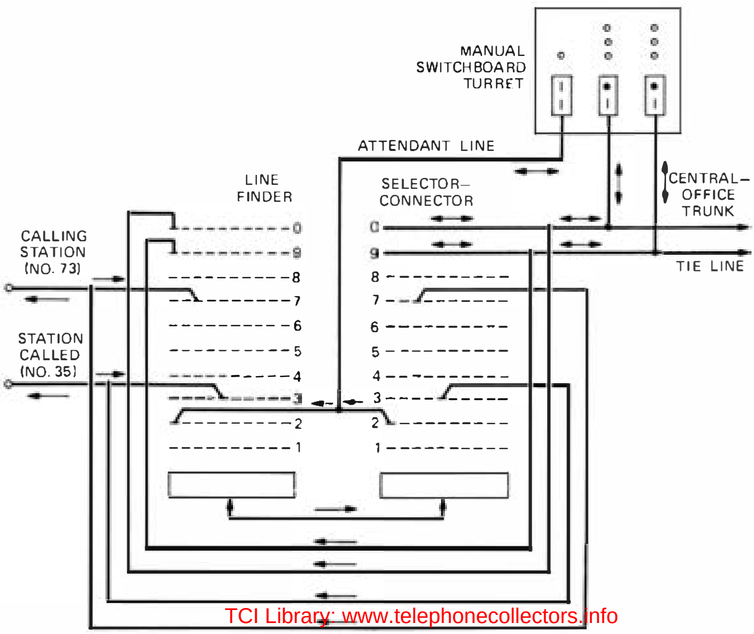


Fig. 6-110. Routing of calls in No. 740A PBX.

which, together with improved equipment arrangements, decreased the cost per line for this size and also the floor space required. The first of these systems, the 740B, was engineered for the traffic requirements of business customers, while the second, the 740C, was engineered for the somewhat lesser traffic but more esthetic requirements of residence use. Toward this end, an attendant's cabinet was designed for the 740C that fitted into the decor of the times, as shown in Fig. 6-111.

The final thrust toward the target was made with the development of the 750A—the smallest dial PBX produced up to that time. The system was designed for residence use, but also found application in small businesses. Made in two sizes, 15 lines and 8 lines maximum, the system solved the attendant problem by not requiring a specific attendant at all. To accomplish this, the system had several interesting innovations. Using relays only to accomplish switching, station telephones were equipped with an appliqué base containing five switching-control pushbuttons (a forerunner of key telephone sets) as shown in Fig. 6-112. One of these buttons gave access to a dial intercom circuit through

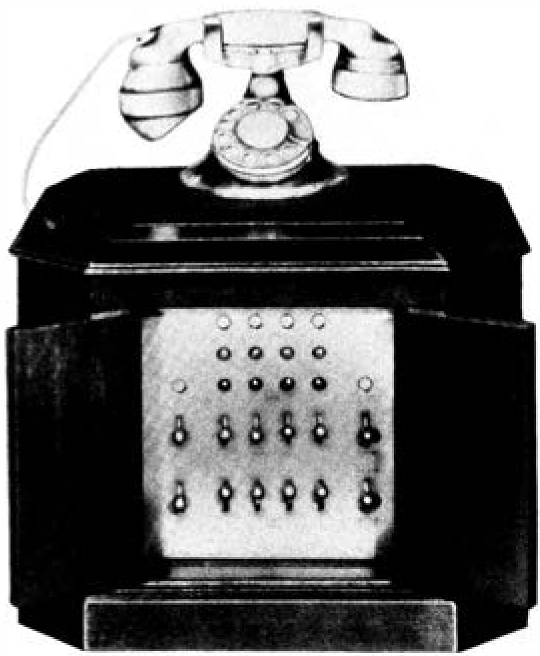


Fig. 6-111. No. 740C PBX attendant cabinet. (Abbott 1968, p. 12)

Fig. 6-112. No. 750A PBX station telephone. (R. W. Harper, *Bell Laboratories Record* 1930, Fig. 1)

which any station could dial any other (a rotary selector followed the pulsing and operated appropriate relays to connect the stations together). Other buttons on the telephone base provided direct access to central-office lines, over which calls could be originated or answered. In the latter case, a call answered by one station could be transferred to another station, if desired, simply by dialing that station through the intercom and asking the person to pick up the proper central-office line. The number of wires per station was minimized to six and the system contained several interesting features for signal identification, lockout and privacy, holding, restriction, and other functions. Equipment design innovations are shown in Fig. 6-113.

The development of the small relay-type 750 PBX supplemented a flexible array of step-by-step dial PBXs. These, together with the manual PBXs mentioned earlier, provided at this time a complete line of manual and dial PBX systems, for all customer sizes, which made efficient use of the technology of the day.

6.4 Summary

The private branch exchange mode of operation had its roots in the beginnings of telephony, and PBX technology was nurtured by the widespread research and development of other switching systems. Within the span of a half-century, PBX systems became an integral and substantial part of the overall Bell System service. By 1929,

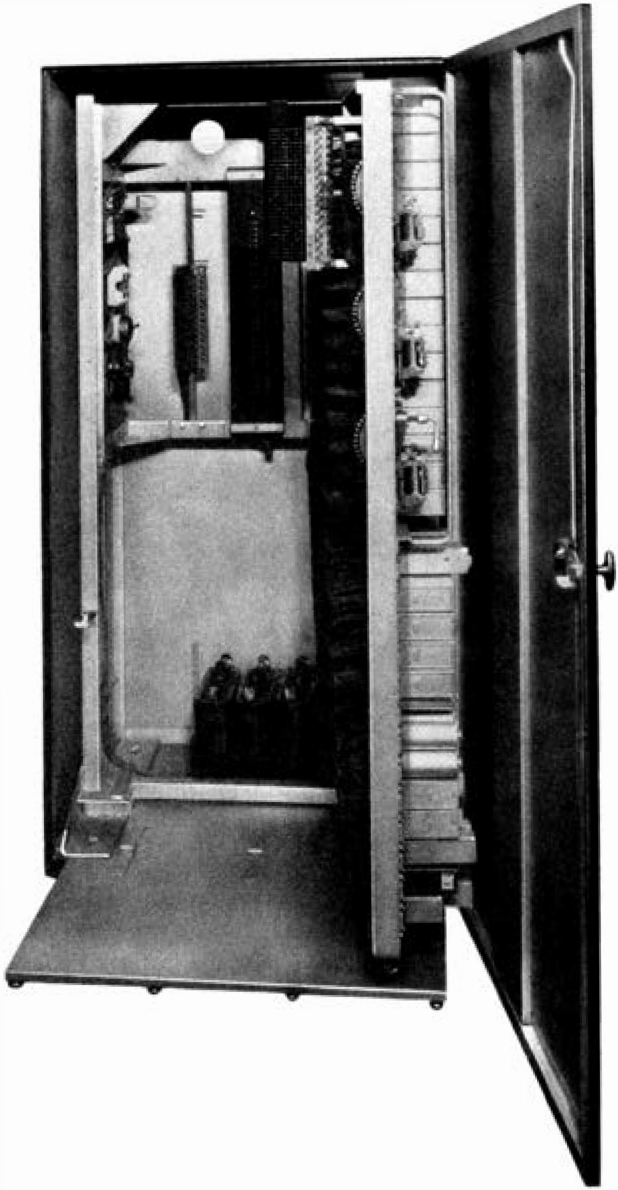


Fig. 6-113. No. 750A PBX equipment cabinet. (R. W. Harper, *Bell Laboratories Record* 1930, Fig. 4)

there existed 130,000 PBXs serving customers whose requirements ranged from less than seven station lines up to many thousands. The number of stations served by these PBXs was almost one-third of the main stations then in existence. While the trunks to these PBXs accounted for perhaps less than 10 percent of the total central-office line terminals, in some business districts such as the downtown New York City area, the PBX trunks comprised 75 percent of the line terminals. In some of the large cities the number of PBX attendants outnumbered the central-office operators; for example, in Manhattan there were 20,000 PBX attendants compared to 9,500 central-office operators.

In view of the vast extent of this service, it is not surprising that the PBX appeared at many times and in many forms during the technological evolution of the Bell System. It employed some of the earliest forms of manual and dial switching technology, and it was the vehicle for some of the fundamental exploratory work on rotary, panel, and Bell System step-by-step technology.

After a little more than half a century the PBX position appeared as follows: Firm operating procedures had evolved which were to continue to the present day. These included the operating interfaces between the attendant, central office, and stations; the coordination of the attendant procedures with dial PBXs; the methods of supervision and privacy; and, very importantly, the numbering plans. It is of interest that one of the most important influences on the technology of switching of all types has been the problem of numbering. In the PBX case, the isolation from the exchange numbering plans permitted by the "dial 9" procedure (adopted by the System in 1917) allowed complete freedom in numbering the stations of large and small PBXs and making provisions for special services. It also allowed freedom in the choice of technology.

Further, the first half-century brought about the evolution of service features which have persisted to the present day—some in their original form, others modified or augmented as the technology advanced. Among these, perhaps those that permitted such operations as station hunting, call transfer, restrictions, night service, conferencing, and interconnection of PBXs via tie lines might be considered fundamental. These served as a solid base for the more sophisticated features that were to come later.

Finally, during the first 50 years, there occurred the hardware implementation of principles, procedures, and features that resulted in a full line of manual and dial systems covering customer requirements of all magnitudes. Figure 6-114 shows the genealogy of these PBX systems. The fundamentals of these systems remained, with improvements, for many years. Some remain to the present day; others yielded after a while to the powerful modern technology of common-control crossbar and solid-state switching. But all provided a firm technology upon which to build in the future.

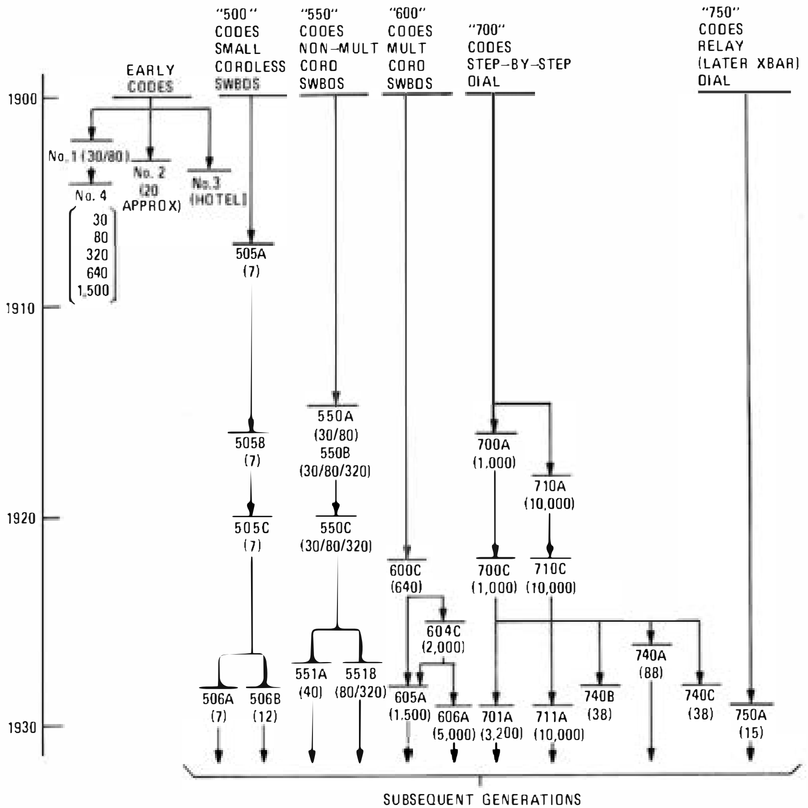


Fig. 6-114. PBX genealogy (maximum number of station lines in parentheses).

VII. CENTRAL-OFFICE POWER SYSTEMS

Except for a few of the very first applications, power supplies for early telephone systems were widely dispersed. On each customer's premises, primary batteries supplied direct current for the telephone transmitter and a hand-cranked magneto furnished low-frequency alternating current for signaling and ringing (as described in Sections III and IV of Chapter 3). The central office also used primary batteries for the operator's telephone and a magneto generator for ringing. The latter was, in the early offices, either hand-cranked or operated by a foot treadle, as shown in Fig. 6-115. Later, various forms of vibrating interrupters or motor-driven alternating-current generators were used. An early form of vibrating interrupter from about 1900 is shown in Fig. 6-116.[174]

[174] This particular device was a reversing switch which sent out alternately positive and negative pulses from a battery supply. Other vibrators supplied on-off pulses which were converted to alternating current by passing them through a transformer.

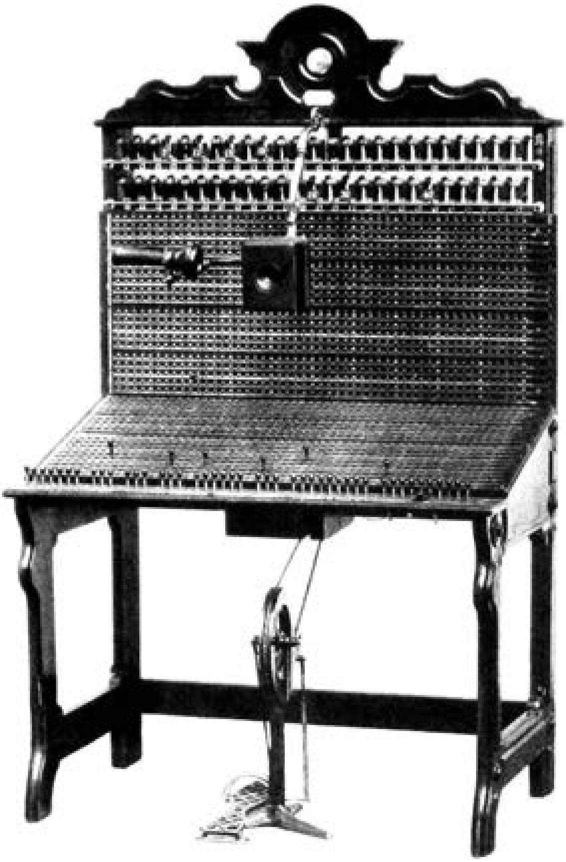


Fig. 6-115. Post 50-line switchboard, with magneto ringing generator operated by a foot treadle, used in Washington, D.C., in 1878.

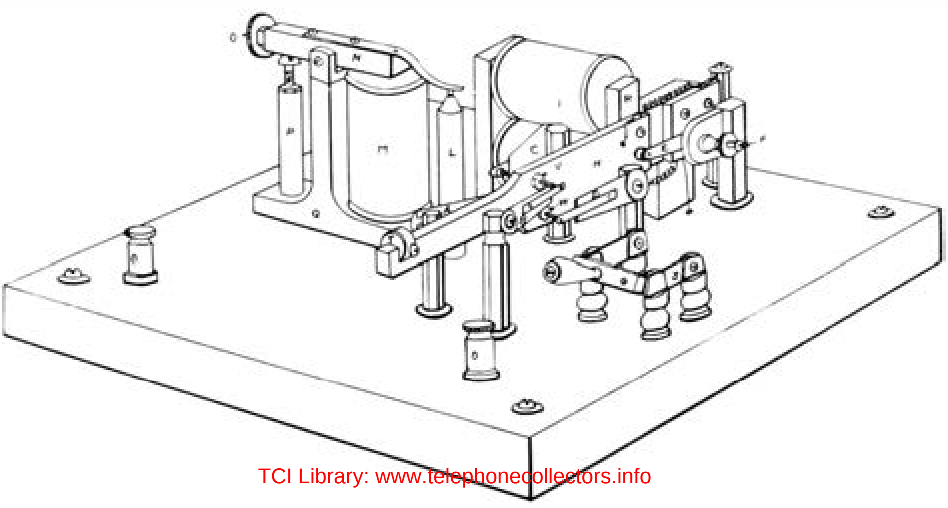


Fig. 6-116. Ringing generator of the pole-changer type.

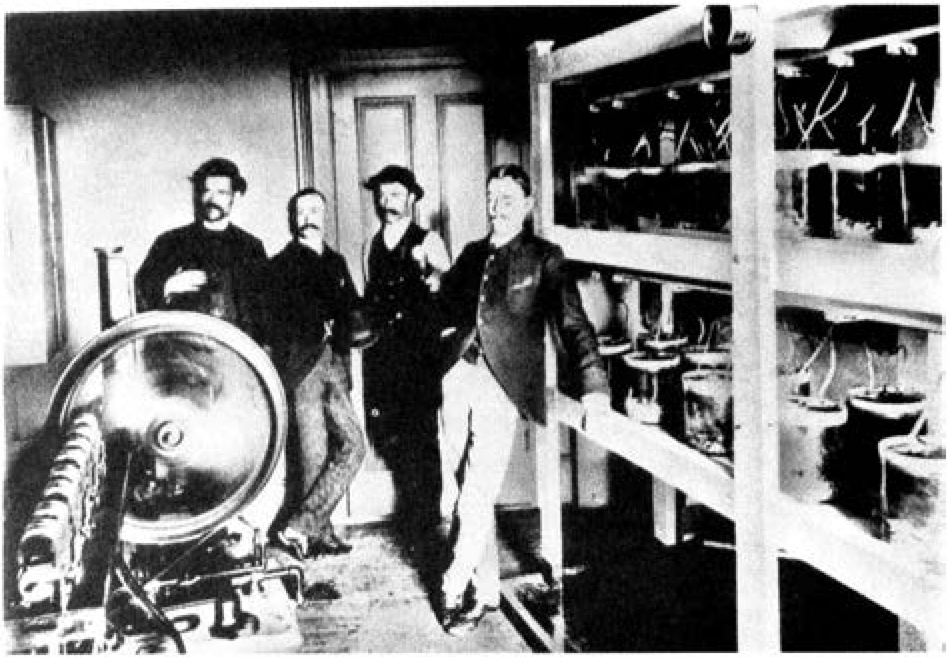


Fig. 6-117. Power plant in Montreal local-battery office, 1886. (From *The Blue Bell*, Bell Telephone Company of Canada magazine, February 1931)

Even with local-battery systems, the need for central-office ringing and talking-battery power increased rapidly as the office became larger and some sizeable installations were required, as indicated by Fig. 6-117 which shows the power plant in Montreal as of 1886. At the left is a lineup of motor-driven magnetos for ringing. Operator talking battery was furnished by the primary batteries at the right. These batteries, which generated electricity by chemical decomposition, were large and required frequent replacement of the active chemical elements as they were consumed by conversion to electrical energy. Serious maintenance and space problems arose as the offices grew in size and would have been greatly increased by the introduction of common-battery systems, which not only concentrated all power supply at the central office but also greatly enlarged the amount of power required for each operating line.[175]

Fortunately, a better alternative became available just as the need

[175] As related elsewhere, the central-office power problems associated with common-battery systems were more than offset by the system's many advantages, not the least of which was the elimination of battery maintenance at the widely dispersed customer locations.

became urgent. Commercial power generation for lighting and traction had been growing rapidly in the last quarter of the nineteenth century and could be used in the large cities, where it was commonly available, as the prime source of electric power for telephony. Commercial power could not be used directly since it did not have the reliability and other characteristics required for telephone applications but, after suitable conversion, it could be used in combination with secondary (storage) batteries to meet telephone needs while reducing battery size and eliminating the costly and messy replacement of the electrodes and electrolyte as required with primary batteries.

The secondary battery, of which there were several types, also produced electricity by chemical action but the process was reversible and once a battery was partially depleted (discharged) it could be renewed (charged) by the simple process of applying a suitable source of electric power, hence the common name of "storage battery." Thus this combination could meet telephone needs (as it still does today), the storage battery supplying the necessary continuity of service and commercial power renewing (or maintaining) the battery charge without replacement of battery elements.

The lead-acid battery, which is the type commonly used for telephony, had been invented by Planté in 1859. However, it was not until the 1880s that it became commercially practical, largely as a result of European developments. The American power industry imported some of these improved batteries as a reserve lighting source and manufacture was started in this country in 1888.

In the early nineties, these batteries were introduced into large local-battery telephone offices and were an essential part of all common-battery offices beginning with the early installations of 1893. The power plant of an 1893 local-battery office in Jersey City is shown in Fig. 6-118. The benchtop contained charging equipment and ringing generators and the storage batteries were placed below it; the control apparatus was mounted on the backboard. The much larger power plant required for a common-battery office is illustrated by Fig. 6-119 which pictures an 1896 installation at Worcester, Massachusetts. A reserve engine-driven generator also is part of this plant.

The use of common-battery telephones required the development of not only large central-office power systems but also those having great reliability. With local-battery telephones, failure of customer power supply affected only a single station and complete failure seldom occurred since it was usually forewarned by a gradual deterioration of transmission as the battery voltage dropped with age. The power supply for the local-battery central office was not much of a problem either so long as it was based on primary cells. It usually consisted of a number of battery units, each of which furnished current to only a portion of the

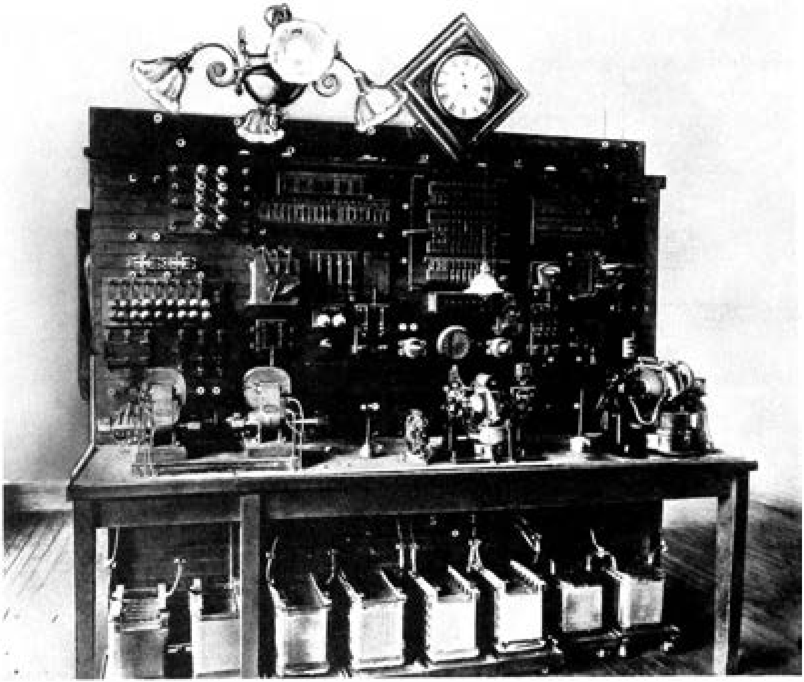


Fig. 6-118. Power plant with storage batteries, in Jersey City local-battery office, 1893.

Fig. 6-119. Power plant in Worcester, Massachusetts, common-battery office, 1896.

office, and cells could be replaced a few at a time as they deteriorated. In large common-battery offices a reliable office supply was essential since failure would have catastrophic effects. The introduction of machine switching increased the amount of power required by the office and emphasized the need for reliable power since even short interruptions could have serious repercussions. Storage batteries charged by commercial power met the basic needs so long as prime power failures were of short duration but other measures were required to protect against interruptions lasting longer than the battery reserve capacity. Battery capacity was designed to carry the system power load anywhere from a few hours to several days depending on the size of the office, the reliability of the commercial system, and the charging procedure. Frequently, two commercial lines supplied from separate power stations were connected to an office for additional reliability but even this was often inadequate. More commonly, in very large cities where uneconomically large batteries would be required for extensive reserve periods, a generator powered by an internal-combustion engine was furnished to provide a local substitute for commercial power during an outage. The engines were often designed for normal use on illuminating gas with gasoline as an alternative.

Safeguards also were provided when installing equipment for converting prime power to the various kinds used in the telephone system. Power-conversion units were supplied in numbers greater than required for a normal load so that enough of such equipment always was available when units were shut down for routine maintenance, or failed. A common arrangement in large offices was to furnish three units, two of which could supply the peak load leaving the third available as a spare. Due to the reduction in power requirements in an off-peak period, even one power supply unit could carry an office over a fairly long period because of the battery reserve available for peak hours. As a result of all these precautions, telephone system failure due to loss of power was almost unheard of, short of a natural disaster such as a fire, flood, or earthquake.[176]

In describing more specifically the evolution of power supplies, it will be convenient to consider them in several categories. The first, and largest, was the basic direct-current central-office supply. This was used to power the talker's transmitter, operate relays and signal lamps, and serve as a basic reserve. As we have noted, this supply evolved into a system using storage batteries charged from commercial power together with appropriate standby generators. In addition, each office required a

[176] A study made in the mid-twenties covering a 20-year period showed that the average outage of the main direct-current supply was 2 seconds per office per year. The less-disastrous loss of ringing current, including both complete and partial failures, amounted to 2 minutes per office per year.

number of supplementary supplies. Some furnished small amounts of direct current at special voltages and others supplied alternating current of various frequencies, both continuous and interrupted, for ringing and signaling. Finally, there were a considerable number of locations, mostly at PBXs, which required many of the power sources used in a central office, and much of the reliability, but where it was impractical to install the full array of equipment used in a large office.

7.1 Basic Direct-Current Supply

The early common-battery offices used a number of different voltages to meet the needs of the various pieces of apparatus employed and power plants were essentially tailormade to meet local conditions, but in the early 1900s power requirements were increasingly standardized. By 1925, the basic direct-current supply provided only two voltages, 24 volts and 48 volts. The former was used in manual offices for exchange-area transmitter supply and most lamp and direct-current signaling operations. In manual offices, 48 volts were used as talking battery on toll calls and for some trunk signaling. Panel machine switching used roughly the same arrangement but step-by-step offices used 48 volts for both talking and signaling purposes.

In the first common-battery offices, two sets of storage batteries were used, one being charged while the other was supplying the office power. This arrangement was particularly necessary in supplying transmitter current since the then-available dynamos used for battery charging produced a large ac "ripple" superposed on the charging current. Since the ripple frequency was in the audible range, an excessive amount of noise resulted.

In 1897, the first steps toward simplification were taken by standardizing on a plant using a single ten-cell battery, all the office apparatus being designed to operate on the voltage supplied. (This battery was charged by a nominally 24-volt dynamo driven by an induction motor from commercial power, with an engine-driven generator as backup). One regular and one reserve ringing machine completed the plant. The battery had sufficient capacity to carry the load for 17 to 20 hours and customary procedure was to rely on the battery alone overnight and to operate the charging generator during the day at a rate high enough to carry the office load and recharge the battery. An attendant regulated the charging voltage as required during the day. The electrical noise caused by the charging generator was reduced by using large choke coils in the charging leads between the generator and battery.

The noise reduction was not completely successful and in 1898 the "M-type" generator, designed specifically for charging telephone batteries, was introduced. This type of generator, which remained in use for nearly 30 years, produced a very quiet output with smaller ripple at a less audible frequency. This was accomplished by using a smooth-surfaced armature with many coils in place of the commercial design, which used fewer coils in a deeply slotted armature. The M-type machine also used copper-gauze brushes instead of carbon. These proved less noisy but unfortunately required frequent lubrication. A typical M-type machine is pictured in Fig. 6-120. By 1927, efficient filters employing the newly available electrolytic capacitors in a size of 1,000 microfarads or more became available and the return to commercial-type charging machines became possible. At the same time, synchronous driving motors were adopted and their capability for power-factor correction reduced the cost of prime power.

The practice, introduced in 1897, of charging the batteries during the day and allowing them to carry the load without charge during the night was continued with some modification until about 1920. The alternate charge and discharge resulted in voltage variations amounting to as much as 25 percent. At an early date the standard battery was increased to 12 cells for 24-volt nominal supply and 23 or 24 cells for 48 volts. With this larger battery, excessively low voltages were avoided and some reduction in the total variation could be achieved by switching the number of cells manually in large offices. In 1919, after considerable testing, the "continuous float" method of charging was introduced in large offices. This involved operation of the charging generator for 24 hours a day at a rate sufficient to supply the office load and maintain the battery at a nearly constant voltage. Originally, the adjustment of the charging generator to meet load requirements was made manually, and

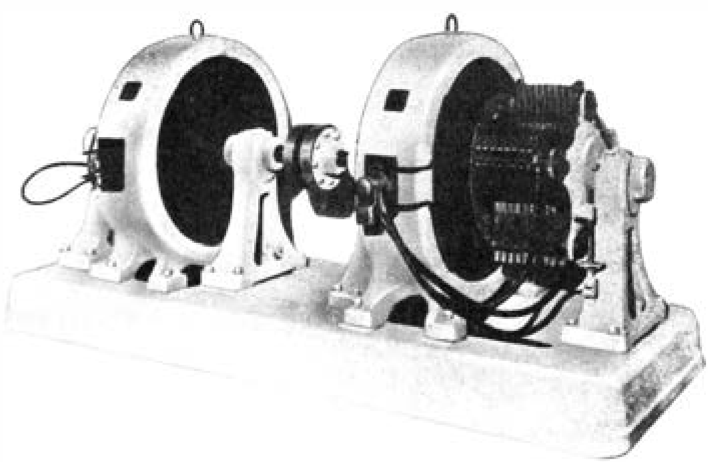


Fig. 6-120. M-type battery-charging machine.

by this means the voltage variation could be held within a range of about 2 volts. This obviously required fairly close supervision by the attending craftsman and, by the middle of the 1920s, techniques for automatic regulation were devised which maintained voltage regulation within about a ½-volt range.

Since many of the devices, such as relays, operating from the standard direct-current central-office voltages, generate a transient voltage which could appear as noise in transmission circuits, it was usual to have two battery outputs. One was the so-called "quiet battery" used to supply current to the customer's loop and transmitter or to other transmission circuitry. The other, or "signaling battery," was used where there were no severe noise limitations. In some cases, separate batteries and chargers were used for the two purposes, but in others the two outputs were supplied by the same battery, the "quiet" output being isolated from the "signaling" output with its high-frequency transients by means of choke coils or other filtering arrangements.

At first, the battery-charging machines were manufactured by the Western Electric Company, but in 1909 Western Electric sold its interest in this type of business to the General Electric Company and the latter manufactured subsequent M-type machines with capacities of as much as 1,500 amperes until the use of commercial types was resumed.

The batteries used in the central office were the product of commercial battery manufacturers and were made to Western Electric specification using the best techniques of the time. Since they were very large in size, they usually were contained in glass jars with open tops. Top enclosures with vents for the generated gases were introduced on small batteries and were gradually extended to larger sizes. These closed cells were commonly used for PBX or other remote supplies but the open cells continued in use past 1925 in large offices. The smaller closed batteries were made up in units of three cells but the individual cell was the basis for all large batteries.

The internal-combustion-engine-generators used for standby power also were produced by commercial manufacturers and modified where necessary to meet Bell System needs.

7.2 Supplementary Power Supplies

The supplementary power supplies furnished a number of different voltages of both alternating and direct current. Most of these supplies were of relatively small output but were subject to rather exacting performance requirements.

One of the main supplementary supplies was required for current to operate the customer's ringer. Originally, this current was provided by manually operated devices with voltages and frequencies that were highly variable. As power drive was introduced, the frequency was standardized at 16⅔ hertz since this was the approximate frequency of the manual generators used and that at which the ringers were tuned. The ringing frequency later was changed to 20 hertz since this was more readily produced by motor-generator sets operated off 60 hertz, the frequency which was gradually standardized for commercial power in the United States. By 1925, the use of various signaling arrangements required alternating currents of approximate sine wave form at several voltages ranging between 105 to about 75 volts. In addition, four-party ringing required superposed positive or negative direct current of comparable voltage and when machine ringing was employed the ringing current had to be made available in 1- or 2-second intervals separated by silent intervals during which direct current replaced the ringing current in order to provide a means for stopping ringing as soon as the call was answered.

Toll-line ringing initially required a source of 135-hertz current and later a 1,000-hertz source also was required.

A number of direct-current sources also were needed. Message registers operated from their own 39-volt supply and 110 volts (poled both positively and negatively relative to ground) was required for coin-box collect and refund. As electronic equipment was introduced, special supplies with stringent requirements had to be provided for the grid, filament, and plate circuits of vacuum tubes.

Various tones approximating 160 and 480 pulsations per second were needed for audible signals to the subscriber and operator and some of these were interrupted to convey "busy" and other information. A siren-like tone, or "howler," also was used to alert a subscriber on an unused station when his receiver was inadvertently left off-hook.

The need for various types of supplementary power with more stringent requirements increased with the evolution of an ever-more-complex telephone plant and led to the development of power supplies of increasing variety and complexity. We have not the space to cover these developments in detail. Suffice to say that originally the demands for customer ringing and signaling tone were supplied by rather simple motor-generator sets driven by commercial power, with a similar set having a motor designed for operation off the 24- or 48-volt battery as a reserve. Direct current for coin-box operation or other small-demand requirements often was supplied by dry batteries with a suitable set held in reserve.

As the needs for supplementary power increased both in amount and variety, standardized units were developed for meeting all the common needs. By the middle twenties, arrangements similar to the ones shown in Fig. 6-121 were being introduced in all but the smallest offices. These units had two drive motors; one operating off commercial power was

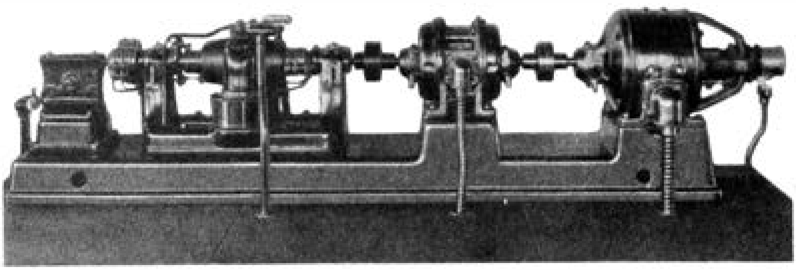


Fig. 6-121. One type of standardized supplementary-power unit, supplying 20-hertz ac for ringing and ±110 volts dc for coin-box operation, and with reduction gearing for low-speed signaling. This unit was driven by an ac-line motor with a reserve battery motor automatically energized upon power failure. A speed controller on the battery motor maintained rotational speed within close limits. (Young 1927, Fig. 13)

normally in use and the other, designed for battery operation, was rotating but disconnected from its power source until failure of commercial power at which time it was automatically connected to the battery, thus preventing any delay in activating the reserve. Under battery operation, automatic control maintained the rotational speed within close limits. Generators produced both direct current at plus or minus 110 volts for coin-box operation and also 20-hertz ac for ringing. The latter was used with a tapped transformer to supply the various voltages required for subscriber ringing. Commutators on the drive shaft interrupted the basic battery current to provide pulsating audible signals at 4, 80, and 160 interruptions per second. Commutating arrangements, driven through a reduction gear, provided the low-rate interruptions that were required for "flashing" signals and the control of automatic ringing. While this "universal" machine was more complex than arrangements required in some offices, the simpler installation and great flexibility made it economically more attractive in the twenties than an assembly of simpler arrangements in a plant that included a considerable variety of manual and machine offices, particularly in those cases where several offices were installed in one building, all being supplied by one power room.

Obviously, not all offices required such elaborate ringing machines and, for these, appropriate smaller equipment was used, some of which continued to employ permanent-magnet, bipolar generators driven by commercially powered motors. In 1914, the 84-type interrupter was introduced. This was a simplified version of the "pole changer" of 1900 illustrated in Fig. 6-116. It interrupted the office battery at an appropriate rate and converted dc to roughly 20-hertz ac of the proper voltage by means of filters or a transformer. Beginning in the 1930s, a static device known as the subcycle generator (developed by an outside supplier) began to be used widely where only a small amount of ringing current was required.

7.3 Power Plants for Toll Offices

The power plants described above were developed specifically for the large exchange-plant offices but the same general principles were applied in toll offices. The basic direct-current power plant was the same but the supplementary plant, while much the same in toll and exchange, required some additional supplies for the former. In particular, 135-hertz and 1,000-hertz generators were needed for toll signaling and these were usually individual motor-generator sets developed for the specific purpose and operating either from the main office battery or from commercial power with battery reserve. A source of telegraph battery also was needed to supply plus or minus 130 volts direct current. Storage batteries, charged from a motor-generator set operating from commercial power, were commonly used for this purpose.

As vacuum tubes were introduced into the telephone plant, additional power sources were required. The grid-bias or "C" battery was commonly a small three-cell dry battery, a separate one being used for each piece of equipment to avoid crosstalk between circuits. The filaments, where possible, operated from the basic office battery with choke coils in the individual circuit to provide electrical isolation. The plate supply ("B" battery) of about 130 volts was often a set of small dry cells built into blocks of 15 or 30, each block supplying a nominal voltage of 22.5 or 45 volts. Taps were provided on both the C- and B-type dry batteries to provide voltages less than the maximum. Later, as the power needs increased, storage batteries were provided to supply plate voltage, with choke coils in the individual leads to avoid coupling between equipment. These batteries were usually charged by means of hot-cathode rectifiers although other arrangements also were used.

7.4 Power for Small and Remote Locations and Special Applications

The systems we have been describing were intended primarily for providing highly reliable power sources for the larger offices. Many of the locations requiring power were remote from the central office or were of such small size that fulltime attendance by craftsmen was impractical. For such situations special arrangements were required to provide both economical and reliable power. Although such systems provided only a small portion of the total telephone power requirements, they were of particular importance since they stimulated the application of new techniques and devices such as unit-type construction, sealed batteries, static rectifiers (of the hot-cathode and copper-oxide type), and automatic control; all of these techniques being particularly adaptable to plants of small size. Much of this work was being done in the twenties and early thirties and resulted in a considerable number of standardized plants for special applications. We have space to mention only a few of these here. Some of these plants were not put into use until after the nominal time frame of this history but are significant for our purpose since they illustrate the important changes in power plant development, largely stimulated by developments started before 1925, which were to have lasting effects on telephone power supply systems.

7.4.1 Plants for Remote Locations

The most numerous remote locations were the private branch exchanges installed on customer premises. Early PBXs obtained both their direct-current and alternating-current (ringing, etc.) supplies over special cable pairs from their serving central offices, thus obtaining the required reliability without complicated equipment on the premises. With this arrangement, voltage varied greatly with load and small storage batteries were soon added to reduce the variation. These were charged on a constant-current basis by means of the current supplied over the central-office pairs, but sometimes local rectifiers were used. Maintenance men visited the PBX periodically to check the electrolyte level and battery charge, adjusting the latter as required. Owing to variations in the power requirements of the PBX, this arrangement was not too satisfactory and would not meet the requirements of the dial PBX with its need for a larger number of tones and close voltage regulation. To meet these needs the 102 series of power plants (illustrated by Fig. 6-122) was introduced. These plants (except for the batteries) were preassembled in the factory and required little installation effort other than the addition of the batteries, which were shipped separately, ready for use. The plant incorporated automatic control, hot-cathode rectifiers of the "Tungar" type, and in some cases ringing machines supplying all the necessary tones.

Occasionally, electronic repeaters were installed at locations where 24-hour attendance was not available, and even when they were located in large toll offices they required small power supplies additional to those commonly available. In larger offices, filaments were often supplied from the quiet 24-volt battery but where the repeater installation was small or remote from the main power supply a rack-mounted plant of the 602 type was used. Plate supply was furnished by similar rack-mounted supplies of the 604 or 603 type. The former used automatic control of a storage battery charged by a Tungar rectifier but the latter, intended for very small installations, used a copper-oxide rectifier with dry cells as a reserve.

7.4.2 Small-Office Plants

Small offices were frequently not manned by craftsmen during the night hours or over the weekends and batteries were usually provided of sufficient size to carry the load during the period without charging. In moderate-size offices a battery plant large enough for the weekend load was expensive and the alternative of overtime attendance for Saturday afternoon and Sunday charging also was costly. This situation stimulated the Pennsylvania Bell Company in 1917 to produce plants charged by means of mercury-arc rectifiers which were automatically turned on

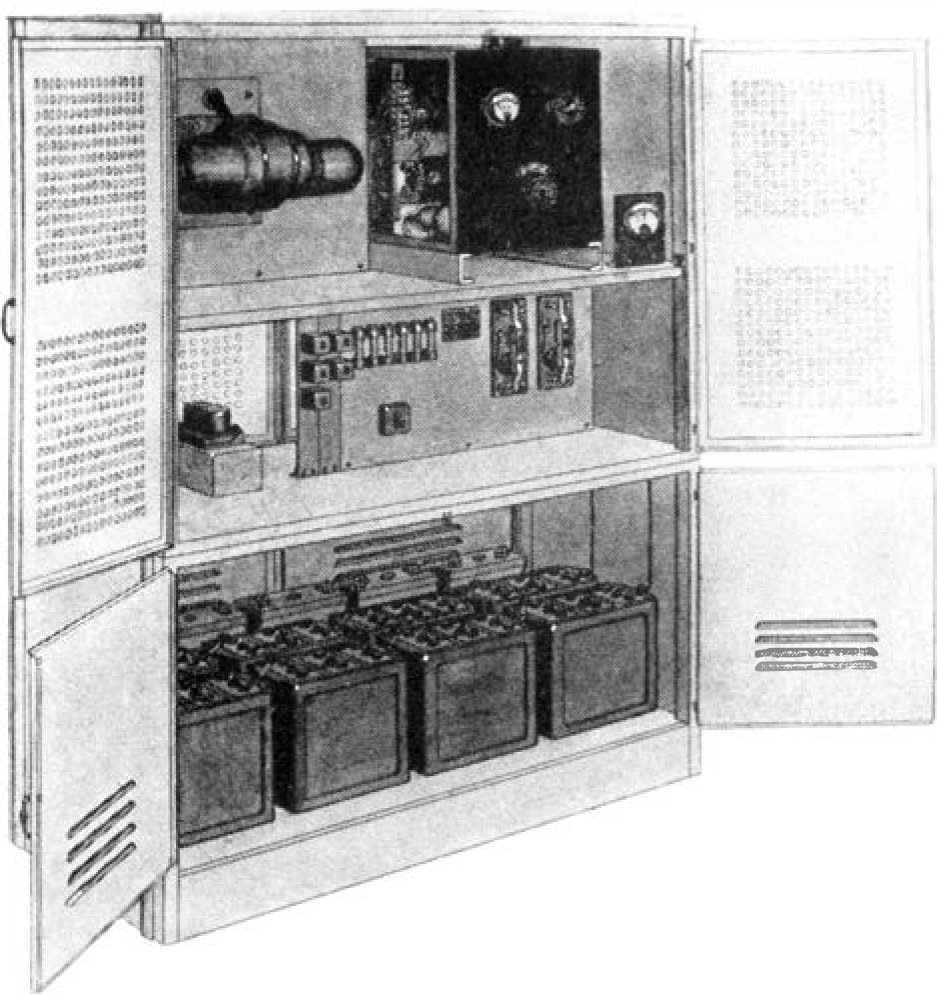


Fig. 6-122. No. 102B power plant for small, dial-type PBX (44 to 50 volts, 2.5 amperes).

whenever the specific gravity of the battery electrolyte fell below a specified figure. Three of these plants were installed and served satisfactorily for many years. However, they were somewhat complicated and better techniques soon became available and were incorporated in small, standardized, Bell System plants of the automatic type. Typical of these was the 200-type design in sizes of 5- and 100-ampere capacity for supplying nominal voltages of 24, 34, and 48 volts. They also usually included the necessary ringing equipment. The 203A plant illustrated in Fig. 6-123 is representative. Charging of these plants was

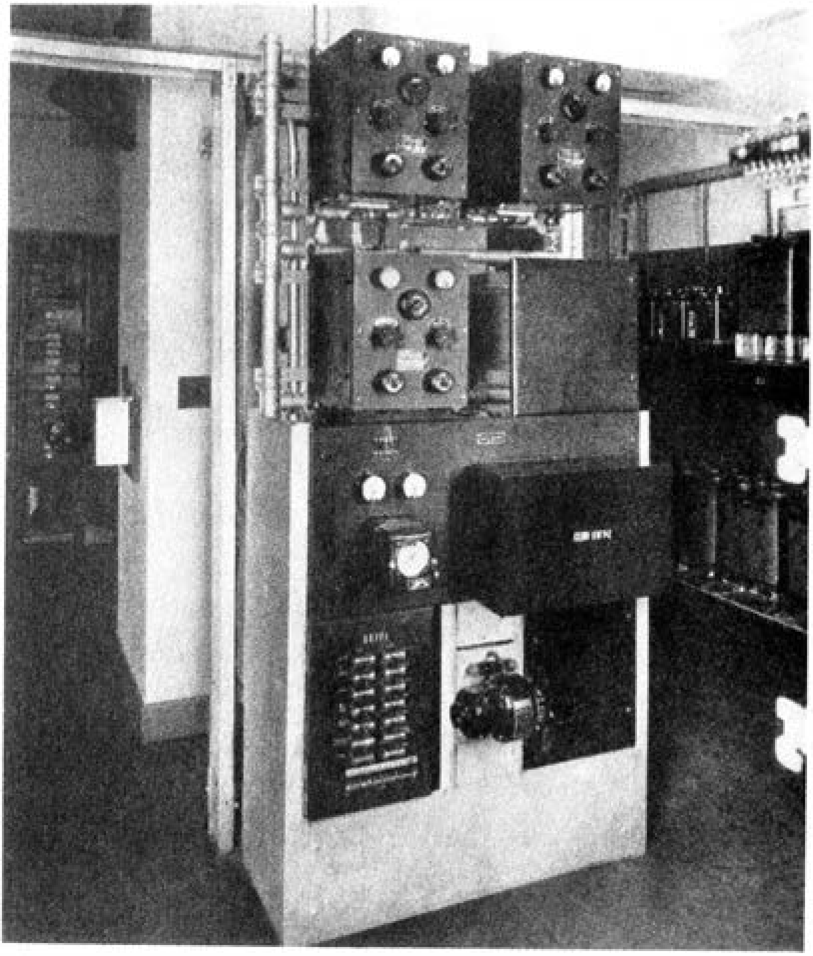


Fig. 6-123. No. 203A automatic-type, low-voltage power plant for small central offices (44 to 50 volts; 15, 30, or 50 amperes).

controlled by an ampere-hour meter which turned on a Tungar-type rectifier when a given amount of discharge had occurred and disconnected it after the charge was restored. Alarms were provided to alert craftsmen in nearby attended offices if the system did not function correctly or if other irregularities such as blown fuses, ringing failure, etc., occurred.

For even smaller offices, still using local-battery operation, the 801C plant furnished both talking battery for the operator and ringing current. The former was supplied by an 8-volt storage battery under continuous charge by means of a copper-oxide rectifier. Ringing was provided by a motor-driven magneto with a vibrating interrupter operating from dry cells as a reserve source.

7.4.3 Special-Purpose Plants

In addition to the plants already mentioned, a number of other standardized special-purpose plants were developed for the office with modest power requirements. These included not only supplies for customer and trunk ringing at 20 hertz but also arrangements for 135- and 1,000-hertz toll line ringing. Direct-current supplies also were provided for coin-box control and for telegraph service. Included in the latter category was an interesting system for compensating for ground potential which, at a time when dc street railway systems were common, could often be quite large and highly variable. We shall not discuss these arrangements since the material already presented illustrates our main objective of showing the many facets of the telephone power problem, the need for great reliability, and the manner in which the techniques which became available in the early part of the twentieth century led to a large effort in the 1920s aimed at providing standardized power plants, including the use of automatic control, which met the increasingly more rigorous and complex requirements arising from the new telephone systems coming into use. However, some information on these systems is included in Fig. 6-124, which summarizes some of the small and specialized plants available on a more or less "package" basis by the end of the 1920s.

7.5 Summary—The Situation in the Twenties

Since the mid-nineties, commercial systems were the prime source of power for telephone operation. Power as received from these systems could not be used directly as supplied but required equipment of special design to convert it to the needs of telephone systems. In order to meet the need for reliability, storage batteries, charged from the prime source, were the major direct source of telephone power, and provided a continuous supply regardless of interruption in the prime source. As an

Type No.	Service	Output	Power Converter and Reserve
102A 1028	PBX — Basic Power	44-50 V dc, 2.5 A 44-50 V dc, 2.5 A & 5 A	Tungar rectifier and enclosed storage batteries
200-Code Series	Small Central Office — Basic Power — Various Sizes	24, 34, 48 V dc, 5 A to 100 A	12-Ampere Tungar rectifier (used in parallel as required); 3-cell (105-Ampere-hour) or single-cell (480-Ampere—hour) enclosed storage batteries as required.
602A	Vacuum Tube Filament Supply	20-28 V dc, 10 A	Tungar rectifier and enclosed storage batteries; supplementary (close voltage) regulation available
603A 604A 6048	Vacuum Tube Plate Supply	130 V dc, 50-80 mA 130 V dc, 50-200 mA 130 V dc, 200-800 mA	Copper-oxide rectifier with dry—cell reserve. Tungar rectifier and enclosed storage batteries
801A	Ringing	75-110 V 20 Hz, 0.19 A	Motor-driven magneto with dry-cell reserve operating vibrating interrupter
801C	Ringing Operator's Transmitter	75-110 V 20 Hz, 0.19 A 8 V dc, 20 or 40 Ah	Same as 801A Copper-oxide rectifier and enclosed storage battery
802A 802B	Toll Signaling	35 V 135 Hz, 40 W 6 V 1,000 Hz, 0.015 A 4.25 V 1,000 Hz, 0.025 A	24-volt basic supply

Fig. 6-124. Typical "package" power plants available in the late 1920s.

additional precaution, most large offices had local engine-generators which could replace the commercial power in case of extended failure.

Originally, power plants were more or less individually designed for each office, but standardization was started in the late nineties to provide a small number of basic plants to meet the special requirements of telephony, and this continued to be a major objective in telephone power development in the years that followed. One of the important needs was for sources nearly free from electrical noise when the power was used for transmission equipment. Special charging generators were developed for the purpose so that charging could take place while the batteries were in service.

As part of the standardization effort, two direct-current voltages, 24 and 48 volts, were selected as major sources for direct operation of telephone equipment. However, increasing numbers of other sources, both dc and ac, were required in small sizes for special purposes, and converters were developed to meet these needs as they arose. In all cases these sources were provided with the necessary reliability, usually by using the main battery as a backup source. Often the auxiliary power supplies were rather complex assemblies of special generators of the rotary type but static converters using mercury-arc rectifiers and hot-cathode tubes were coming into fairly common use in the early twenties for supplying small direct-current needs. They were used not only for the many special voltages required but also for the 24 and 48 volts supplied at remote locations where only small amounts of power were needed and attendance was part-time.

All of these requirements added up to large and complex installations for supplying central-office power. One of the large installations required for a two-unit machine switching office is shown in Fig. 6-125. The cost of providing and operating this equipment and of compensating for losses in the power conversion was sizable and resulted in a cost of telephone power at point of use which was as much as ten times the cost of the prime power.

The twenties marked a period of great change in telephone power systems. Some of these changes were already in use by 1925 and for others the basic technology had just become available and would soon be put to use. Around 1920, the office voltage regulation in large offices was greatly improved by the "continuous float" method of charging using manual adjustments. By 1925, techniques for automatic regulation became available and their adoption followed rapidly thereafter. Means for reducing the cost and doubling the life of batteries were devised and put into production in 1927. Static rectifiers using hot-cathode vacuum tubes of the Tungar type had been adopted for small power plants, particularly for furnishing vacuum tube voltages, before 1920; but during the early twenties technology for using metallic rectifiers of the copper-oxide type was evolving and these devices were put to use in 1927. They received large usage in small plants, ultimately replacing many of the small Tungar type and being themselves replaced by selenium rectifiers beginning about 1938. However, the Tungar rectifiers continued in use in medium-sized unattended locations until they were replaced by the thyratron type about 1935. Thus, as in the other parts of the plant, the end of the first 50 years of telephony found new technology in the field of power supply developing rapidly and laying the groundwork for significant changes to be made in the years to follow.

VIII. SWITCHING SUMMARY

The first half-century of switching development can be characterized briefly as a period of establishing basic technical foundations on which the future could be built.

Prior telecommunication art had little to contribute to telephone switching and the early years were devoted to finding techniques for interconnecting stations in a rapid and economical manner. Accomplishing this required not only new principles but new hardware for implementing them. All had to be done within the constraints imposed by the primitive state of the station and transmission components of the telephone system. By the early 1900s, developments throughout the Bell

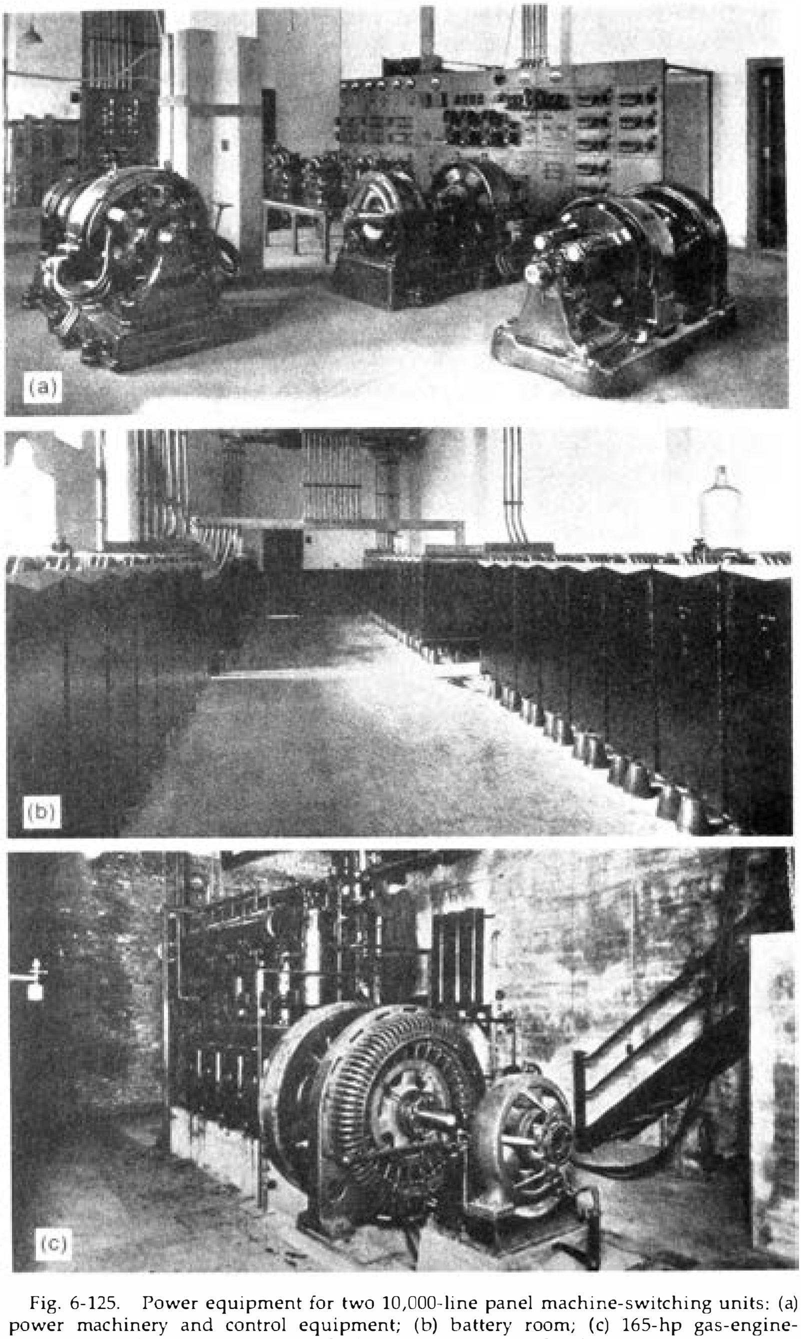


Fig. 6-125. Power equipment for two 10,000-line panel machine-switching units: (a) power machinery and control equipment; (b) battery room; (c) 165-hp gas-engine-generator set for emergency use. (Craft, Morehouse, and Charlesworth 1923, Figs. 15, 16, and 17)

System brought a degree of sophistication to all parts of the telephone plant and gave confidence that Bell's "Grand System" for universal telephone communication ultimately could be achieved even though many problems had yet to be solved. In the switching area, a basic interconnection plan had evolved employing a system of switching points, or central offices, connected directly to telephone stations by wire lines (loops) and in turn connected to each other by higher-grade lines referred to as trunks, sometimes directly but often through intermediate switching points. During this first 25 years, switching was handled almost entirely on a manual basis and by the turn of the century the originally large and complex switching arrangements had been simplified to the point where they were compact and required minimal effort on the part of both the telephone user and the switching operator. Metallic (2-wire) transmission and the common-battery principle had been adopted and means had been devised for economically switching calls in exchange areas of roughly any size from a small village to the largest metropolitan area.

During the first 25 years of the twentieth century most switching continued to be handled on a manual basis with considerable improvement being made in economy, service features, and operating methods. Particular emphasis was placed on meeting the special problems of toll switching and, while some switching of toll calls occurred before 1900, it was the developments of the second 25 years of telephony that demonstrated the practicality of Bell's "Grand System." It was also this second 25 years which saw great progress in the development of PBX systems as a supplement to the common-carrier network, establishing an integrated system serving all telephone needs from on-premise intercommunication to intercontinental long-distance calling.

This second 25 years also brought developments that demonstrated the practicality of automatic (or machine) switching and laid the groundwork for its ultimate widespread usage. Proposals for automated switching were made as early as 1879 and an automatic exchange was installed in 1892; but it is fair to say that automation did not move much beyond the experimental stage until the 1900s. By the 1920s, two basically different systems had come into use in the Bell System. For small exchange areas (up to several central-office installations) the Strowger or step-by-step system usually was preferred. This was a progressive, direct-control system in which pulses from a customer's dial operated directly a series of switches (with two directions of motion) arranged in a tree-like configuration. This type of system involved many serious problems when attempts were made to apply it to large metropolitan areas. Such areas formed an important part of the Bell System and to meet their special needs Bell developed a completely different system using large linear (panel) switches controlled indirectly by the customer's dial through common equipment which registered dial pulses and translated them to signals suitable for operating the switches. This indirect, common-control approach proved to be a powerful tool and became basic to all modern large-size switching systems.

In spite of the large amount of development effort during this period, the number of Bell stations switched on a machine basis was only about 12 percent in 1925. Thus the major effect of the work on automation conducted during the first quarter of the century was in establishing a foundation for the future. In this respect the situation was very similar to that we have found in examining the development of stations and transmission systems where the late twenties found telephony with the technical background needed for great future expansion.

Chapter 7

Non-Voice Communications

The use of coded light signals to transmit information over considerable distances extends into the dim past of recorded history. However, the electromagnetic mode of transmission over wire-line media did not come into use until the nineteenth century. The notable experiments made by Samuel F. B. Morse and the collaborating work by Alfred Vail in evolving signal codes to identify letters and figures provided the necessary means of transmitting a coded message over a wire line for considerable distance and its instant recording at the receiving station. Morse demonstrated the feasibility of this mode when he transmitted a message over an open-wire grounded line extending from Baltimore to Washington, D.C., during 1844.

This chapter highlights the early history of telegraphic communications prior to the time of the Morse demonstration, traces the growth and evolution of wire-line telegraph communications within the United States prior to formation of the Bell System, and finally describes the means employed by the Bell System to provide telegraph and other non-voice services to the public during the 50-year period beginning within the decade of the 1880s and extending through the 1920s.

I. INTRODUCTION

The primary form of communication between people throughout the ages has been a procedure known as talking. This acoustic art is a mouth-to-ear sort of process which is severely limited by increasing the distance between the individuals involved. Human innovation overcame this handicap by making use of the eye which could recognize changes in a pattern of lights or objects at a considerable distance. Clearly this process would have a limited vocabulary and take a long time to transmit a given message, compared to the verbal method. Nonetheless, ancient peoples found visual signaling a high-speed method compared to sending a messenger on foot to complete the verbal process. We are told by Aeschylus that, long before

Agamemnon returned home, Clytemnestra learned of the fall of Troy by means of a sequence of fire signals flashed from one mountain top to another.

Communication over a distance resembles transportation. The commodity to be transferred from one place to another is information, and the system has attributes of speed and capacity or quantity of information per unit time. Speed is a relative sort of descriptive term in that a proposed system will be called "high speed" if it is faster than the system in present use. At some future time the term "high speed" will be dropped from the then-existing system and transferred forward when a new, faster generation is proposed. Thus we see that Clytemnestra's signal fires were high speed compared to sailing across the Aegean Sea, but in terms of today's technology would seem incredibly slow. It is probable too that her system transmitted only one item of information, i.e., "Troy has fallen and the Greeks will be returning."

With the passage of time the users of signaling systems exhibited the natural desire of wanting something better, faster, and more flexible. It became apparent that better performance could be obtained if written messages were sent letter-by-letter using some sort of coding system in which each letter or number was represented by a unique pattern that was discernible to the eye. Two hundred years after the fall of Troy, Greek military leaders displayed sets of torches in patterns which signified the letters of the alphabet.

The period before the year 1800 was one of intensive development of semaphore transmission. It was sponsored and used mainly by governments. In France a message could travel a distance of nearly 500 miles in about 12 minutes. This means that a question could be asked and the answer received in half an hour. Again compared to use of messengers on foot or horseback, this was indeed high speed. Such systems were in common use in Europe and ruins of the signal towers along the Georgian military road across the Caucasus mountains are still very much in evidence today. The success of the semaphore system depended on the use of telescopes which enabled the relay points to be stretched out; but weather in the form of fog, clouds, or precipitation was an adverse factor. A locally heavy storm could block transmission for the duration of the storm.

II. EARLY ELECTROMAGNETIC EXPERIMENTS

Experiments with electricity in the eighteenth century were directed toward the possibility of developing a new kind of signaling system. These experimental trials were the forebear of modern telegraph systems. The word "telegraph" derives from the Greek, "tele" meaning far-off and "graph" meaning something written, i.e., a symbol of some sort; hence a telegraph system was one which could reproduce symbols at a distance. Although static electricity, conductors, and insulators had been recognized since the days of the Greeks, efforts to make a working system based on known principles showed that such a system was impractical. Experimentation with electricity continued and it was learned that a magnet could be constructed by causing an electric current to flow through a coil of insulated wire wrapped around a piece of soft iron. Work by Arago and Faraday in 1820 and by Henry in 1830 was largely responsible for providing the electromagnet which in turn provided the basic technologic instrumentality needed for the development of an electromagnetic telegraph system. Literally dozens upon dozens of experimental systems had preceded it and proved impractical.

III. THE MORSE TELEGRAPH

Shortly after 1830, Samuel F. B. Morse began experiments and the system which he devised became widely accepted. Morse had been educated at Yale, graduating in 1810. He went to England to study art and became a moderately successful portrait painter after returning to America. He had been interested in electrical phenomena in his college days but apparently that interest had remained dormant until middle age. There were many electrical experiments being reported and lectures given in the period before 1830, which reawakened Morse's interest in electricity and he began experimenting with the electromagnetic telegraph. In his first system Morse used a pencil fastened to the movable armature of an electromagnet. It left a trace on a slowly moving roll of paper. At some distant point the electric circuit through the electromagnet was momentarily closed by operating a telegraph key. The resultant current flowing through the magnet caused the armature and pencil to be displaced and then return to the normal position. The result was that a V or notch would occur in the steady trace made by the pencil, every time the key was operated at the distant point. Morse's initial designs would probably be called somewhat cumbersome but they were sufficient to demonstrate the potential worth of his proposed system.

Discussions with the faculty at the University of New York stimulated the active participation of Alfred Vail, an inventor, 16 years younger than Morse, who in turn sought financial support from his father, Judge Stephen Vail of Morristown, New Jersey. But the Judge remained skeptical until a message devised by Alfred Vail was successfully transmitted over 3 miles of wire that had been strung around the barn on the Vail estate. This occurred on January 6, 1838, and the favorable impression secured the further financial support which enabled the development of a working telegraph instrument.

Morse originally thought he could construct a dictionary of numbered words and phrases corresponding to the notches on the tape and spent several months at the endeavor. It became clear that it would be better to code letters by a sequence of short dots and long dots, i.e., by dots and dashes. It is said that Vail went to the newspaper office in Morristown and counted the number of pieces of type in each letter box of the alphabet. He concluded that telegraph code should use the simplest telegraph signal, a dot, for the most frequently used letter of the alphabet, e, and gradually increase the complexity of the signal as the frequency of usage of a letter decreased. This concept had already been adopted by European experimenters. It was expected that the dots and dashes on the paper would be read and converted to a hand-written message. Some years later the operators who had the task of making the conversion realized that the auditory sense was superior for the task at hand and that merely by listening to the clicks of the electromagnet they could write down the message as rapidly as it arrived. It proved easy for the ear to recognize the sounds made when the magnet operated and when it released. In order to make the task easier an improved listening device known as a "sounder" was added so that both the sender and the receiver could hear the code being transmitted.

It has been claimed that Morse did not invent the telegraph. It is certainly true that the concepts are lost in the mists of history, and there certainly was much prior and concurrent experimentation from which there were conflicting claims. It does appear clear, however, that Morse and Vail designed, built, and demonstrated the first practical and successful electromagnetic telegraph system in the United States. It is a historical fact that in periods of scientific ferment the same discovery may be made independently at the same time in widely separated locations, and that, further, the workers in a particular field "feed" on each other's discoveries, stimulating still greater advances.

IV. CODES USED IN TELEGRAPH COMMUNICATIONS

It was mentioned earlier that codes for letters of the alphabet had existed since the time of the Greeks. With the passage of time, coded signals became more efficient in transmitting a given amount of information in the minimum time. Figure 7-1 shows a number of codes, all but one of which were developed in the first half of the nineteenth century. It will be noted that the concept of two elements in varying combinations was well understood. The codes of Bacon in 1605 and Rees in 1809 made use of a binary code of five elements which give 2^5 or 32 different combinations or characters. Swaim's code

	Bacon	Rees	Swaim	Schilling	Gauss & Weber	Steinheil	Original Morse	Later Morse (American)	European Morse (International)
	1605	1809	1829	---	1833	1836	1838	1844	1851
A	aaaaa	11111	t	rl	r	lrl	sss	sl	sl
B	aaaab	11112	tt	rrr	ll	lrrl	ss ss	lsss	lsss
C	aaaba	11121	ttt	rll	rrr	llr	s ss	ss s	lsls
D	aaabb	11122	tttt	rrl	rrl	rl	sss s	lss	lss
E	aabaa	11211	s	r	l	l	s	s	s
F	aabab	11212	ss	rrrr	rlr	lrr	s sss	sls	ssls
G	aabba	11221	sss	llll	lrr	rrl	ss s	lls	lls
H	acbbb	11222	ssss	rlll	lll	rrrr	ssss	ssss	ssss
I	abaca	12111	ts	rr	rr	r	sl	ss	ss
J	---	12112	t tts	rrll	—	r	ss s	lsls	slll
K	abaab	12122	t t	rrrl	rrr	llr	lsl	lsl	lsl
L	ababa	12211	t tt	lrrr	llr	rll	l'	l'	slss
M	ababb	12212	t ttt	lrl	lrl	rrr	lss	ll	ll
N	abbaa	12221	t tttt	lr	rll	rr	ls	ls	ls
O	abbab	12222	t s	rlr	rl	lll	ss	s s	lll
P	abbba	21111	t ss	llrr	rrrr	rllr	sssss	sssss	slls
Q	abbbb	21112	t sss	lllr	——	——	ssls	ssls	llsl
R	baaaa	21121	t ssss	lrr	rrrl	ll	ss	s ss	sls
S	bcacb	21122	t ts	ll	rrlr	llrr	sls	sss	sss
T	baaba	21211	tt tts	l	rlrr	lr	lls	l	l
U	----	21212	tt t	llr	lr	rlr	sll	ssl	ssl
V	bacbb	21221	tt tt	lll	rlr	rlr	l	sssl	sssl
W	babac	12121	tt ttt	rlrl	lrrr	rlrl	ssl	sll	sll
X	babab	22212	tt tttt	lrlr	——	——	ll	slss	lssl
Y	babba	22221	tt s	rllr	——	——	sl	ss ss	lsll
Z	babbb	22122	tt ss	rlrr	rrll	rrll	sls	sss s	llss

Fig. 7-1. Comparison of early alphabetic codes with the various Morse codes.

is said to have been used for communication between prison cells, where "t" stands for "tap" and "s" for "scratch." The other alphabetic codes were used in experimental setups, such as left or right on a galvanometer, black or white, short or long pulse. The Morse codes of 1838 and 1844 very largely reflect the inventive genius of Morse's assistant Alfred Vail. It will be seen that a few letters are omitted from some of the codes and there are some duplications which the receiving telegraph operator was expected to properly interpret.

V. GENESIS AND EARLY HISTORY OF TELEGRAPH OVER WIRE LINES

The ferment and excitement of instant communication had permeated political realms in the United States and in 1843 the Congress passed the Telegraph Bill which provided $30,000 to build a telegraph line from Baltimore to Washington, D.C. On May 27, 1844, a demonstration took place which assured success to Morse's system. The historic message "What hath God wrought?" was transmitted a distance of 40 miles! The message was received on a paper tape known as a recording telegraph, which is shown in Fig. 7-2. Early systems were operated over a single wire and the apparatus at each end was connected between the wire and ground.

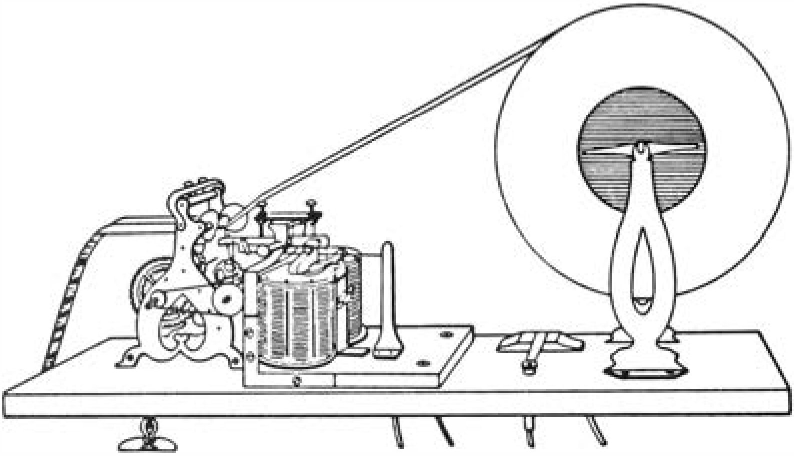


Fig. 7-2. Morse's recording electric telegraph as used on the experimental line between Baltimore and Washington in 1844. (Still 1946)

This single-wire system reflected the desire to make the operation as economical as possible by using the earth in place of a second return wire. As the separation between the two terminals increased, technical problems arose. The potential of the ground at the two locations might be different; hence the ground behaved as a giant battery which caused an unwanted current to flow in the wire when the circuit was closed, interfering with and altering the wanted current. Additional components were needed to compensate for this effect, reducing the economic advantage. Furthermore, the ground potentials changed with time and distance, thus requiring human intervention and causing further expense. It is a fact that nature behaves differently from the basic concepts of the textbook, and much of the work of the engineer is concerned with compensating for these differences. The harder he pushes a system, i.e., the more he attempts to maximize the efficiency and reliability, the greater are the technical additions he must make to the original system to assure success.

It is true that as soon as a new system proves its worth, expanded services are suggested. Thus by 1849 a telegraph service was provided between New York and Philadelphia in which each message was recorded in printed form for delivery to the recipient. This service grew rapidly since the printed message was very useful to commercial businesses and press organizations receiving private messages. The concept of a typewriter was in existence but a practical machine as it is known today did not come along until 1875, three decades after the inception of the manual-type telegraph system.

Telegraph service and the physical plant providing it expanded rapidly. Quite naturally it grew where the service would be valuable. It interconnected cities and followed main railroad lines, having service points, called drops, at railroad stations. Many small companies sprang up, sometimes operating only a single-wire-line system between two relatively nearby places. The transmission of military messages during the Civil War provided a strong stimulus to the expansion of telegraph service. Finally, the Western Union Telegraph Company began absorbing the smaller operating companies and emerged as one of the largest and most important companies in the history of the country up to that time. Telegraph became a universal service within the country. It was used to transmit information in the fields of banking, business, government, railroads, and news gathering, and in the sector of the private citizen.

Since telegraph service originated with a piece of paper bearing a message and terminated with a piece of paper having a reproduction of that message, it was considered a competitor of the U.S. mail service. The advantage of telegraph obviously was its high speed, i.e., reduction in the time needed to transmit, record, and deliver the message and to get a reply if one were requested. This feature was evidently worth the expense involved for news associations, businesses, or any organization where prompt action was synonymous with the improvement of operations.

The telegraph plant consisted largely of single-wire, ground-return circuits. These could be operated in only one direction at a time, i.e., an incoming message would be interrupted to send an urgent outgoing message.

The early systems involved "neutral operation" in which there was a flow of current for the "marking," i.e., closed, position and zero current for the "spacing" (open) position. Figure 7-3 shows the fundamental principle of a "single Morse" circuit. At each station there is a key, relay, sounder, and local battery. The stations are connected together by the single-wire line which is grounded at B and connected to battery at A. Ordinarily the contacts of each telegrapher's key were shorted so that current flowed continuously. If the operator at B wished to send a message, he opened the shorting switch on his key, thereby putting both relays and their sounders in a condition to respond to operations of the key. A short make and break of the key (followed by the sounders) was interpreted as a "dot"; a make lasting a bit over one-tenth of a second followed by a break was recognized as a "dash." Letters and figures were sent according to the American Morse code, each followed by a space. This code is tabulated below for easy reference, although it is included in the comprehensive table (Fig. 7-1).

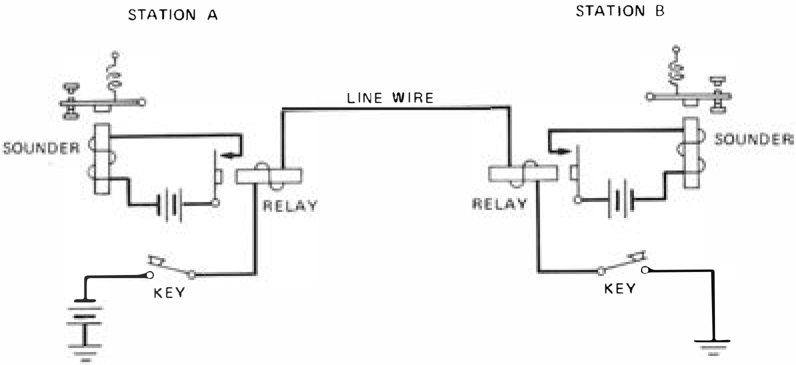


Fig. 7-3. Elementary neutral telegraph circuit.

A	·—	H	····	O	· ·	V	···—	3	···—·
B	—···	I	··	P	·····	W	·——	4	····—
C	·· ·	J	—·—·	Q	··—·	X	·—··	5	———
D	—··	K	—·—	R	· ··	Y	·· ··	6	······
E	·	L	⸺	S	···	Z	··· ·	7	——··
F	·—·	M	——	T	–	1	·——·	8	—····
G	——·	N	—·	U	··—	2	··—··	9	—··—
								0	⸺

If the operator at A wished to interrupt, for instance to ask for a repeat, he simply opened the shorting switch of his key, thereby opening the circuit. B upon hearing no response of the sounder to his key would realize that A wished to send a message. He would then throw his shorting switch to receive from A.

Operation of the neutral system was generally satisfactory for short distances, but as the distance was increased, the corresponding increase of the resistance lowered the current and ultimately affected the response of the relay. In addition there was a tendency for the current to leak to ground along the line, particularly under wet weather conditions. Operation was also affected by earth potentials and capacitive effects from the line and equipment. It will be clear from the schematic that the circuit impedance changed by a finite value as the key went from a marking to a spacing condition. The waveshapes of the current buildup and decay were different, which was a disadvantage. All in all the neutral operation had very definite limitations and something better was urgently needed.

"Polar" operation provided a considerable improvement. Whereas the neutral relay operated on current of either direction, the polar relay was designed to operate when current of the right polarity flowed through the windings but to have no response when current of the opposite polarity flowed. In the polar arrangement the telegrapher's key connected positive battery to the line on marking and negative battery on spacing. The transition from marking to spacing therefore involved a symmetrical current reversal in a circuit of constant impedance. The fact that there was current flow for both marking and spacing improved the operating margin over the neutral mode without increasing the magnitude of the steady state current in any part of the circuit and without increasing the voltage to ground.

The need for two-way service, called duplex service, brought about the development of new equipment at the telegraph stations in 1872. In one form of duplex equipment a Wheatstone-bridge configuration was used. (See Fig. 7-4a.) The two windings of a bridging coil formed two arms of the bridge; the line to the distant equipment and its

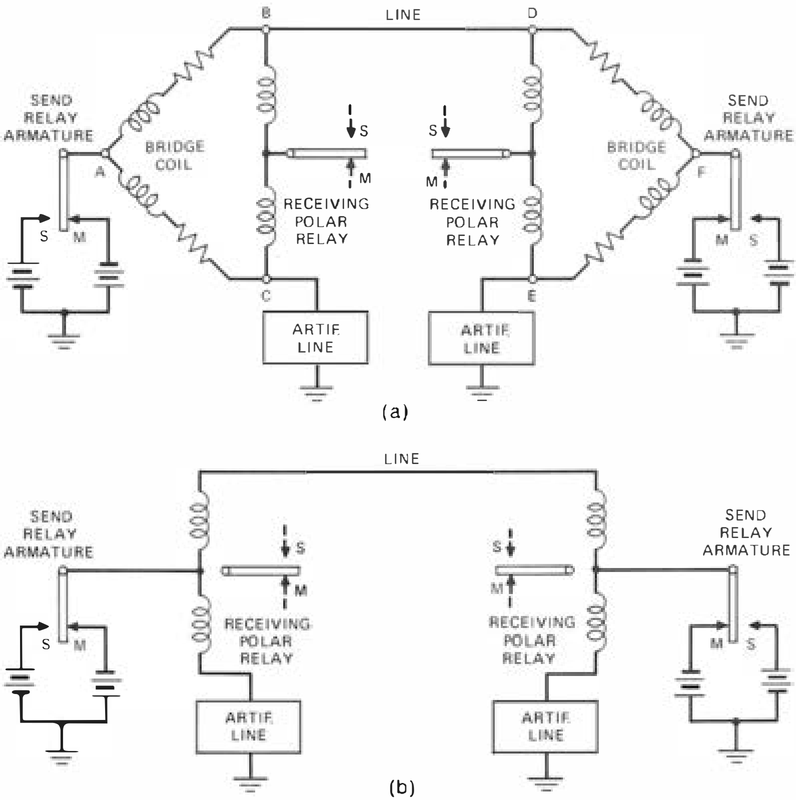


Fig. 7-4. Polar transmission uses positive and negative pulses passed through a bridge circuit, as at (a), or a differential relay circuit, as at (b). (Redrawn from Hamilton 1946, Fig. 1)

simulating network formed the other two arms. The receiving polar relay was placed in the galvanometer position and showed no response to local keying when the bridge was balanced, but it responded to sending from the distant key. In a second form, differential duplex, the receiving relay was provided with two windings. (See Fig. 7-4b.) These were connected in such a way that the armature responded when the distant telegraph key was operated, but the armature was not affected when the local key was operated. It was necessary to connect one of the two windings of the receiving relay to an artificial line which, in both arrangements, was composed of resistors and capacitors, so arranged that it was a good impedance simulation of the actual line and terminating equipment connected to the other winding of the relay or coil. Operation of the local key would cause the current to divide equally in the two windings which were arranged to oppose each other so that there was no effect on the armature. Again there would be need for adjustment of the network according to the length of the line and if there were changes in leakage.

Edison developed a quadruplex system in 1874 combining polar and neutral differential duplex systems so that two two-way channels could be operated over a single line simultaneously.

In a brief review such as this, it is possible to touch on only a few of the highlights. There were many practical problems associated with daily operation, especially those having to do with maintenance. Variations in battery voltage, insulation of the line, adjustment of the relays, etc., only added to the dedication of the operators who realized that they were in the center of a new endeavor which recognized the importance of communications in the mercantile and social life of the country.

5.1 Printing-Telegraph Technology During the Nineteenth Century

A short historical review of the development of telegraph printers provides a bit of perspective in seeing how pressures of the marketplace combine with existing technology and stimulate further developments toward improved service. It will also be pointed out that there was a widespread interest in printing telegraph with developments showing up at widely separated locations in North America and Europe in the early days.

In 1848, only four years after the Baltimore–Washington demonstration of telegraphy, a Vermont inventor, Royal E. House, recieved a patent for a printing-telegraph system using a piano-type keyboard which actuated a type wheel by compressed air. In 1855, David Hughes, of Kentucky, developed a greatly improved machine along similar lines, which was used commercially. At about the same time a Connecticut man, named John P. Humaston, patented a keyboard-and-treadle scheme whose perforations of a paper tape represented the telegraph message in Morse code signals. In Germany, Dr. Werner Siemens patented in 1868 the first completely automatic keyboard perforator for use with Morse code. Improvements were made in England by Creed and his apparatus was generally adopted there around 1900.

At about the time Siemens was at work in Germany, Elisha Gray, of the firm of Gray and Barton in Boston, Massachusetts, was busy developing a private-line printer for use on short lines, such as that connecting a downtown office with a factory some little distance away. Gray's machine was evidently superior to competing machines and came to have fairly wide usage in the larger cities of New York, Pittsburgh, Detroit, Chicago, and St. Louis.

Emile Baudot, working in France in 1874, developed a practical system of printing telegraphy employing paper strips and a five-unit selecting code which has been in common use since that time. His type-printing multiplex system, which employed time division, was the basis for the Murray and Western Electric multiplex systems which came into commercial use after the turn of the twentieth century. (The multiplex system is described in Section 9.3.)

The need for records in the form of the printed page continued. The typewriter, invented in 1875, provided such a record, and it was logical to unite this device with the telegraph system. The first page printer was an Essick machine which is believed to have been put in operation over telephone lines in 1891. Donald Murray was also active in this field and in 1901 he introduced a high-speed automatic system, and a multiplex system which was used by Western Union.

Despite these early attempts to introduce suitable printing-telegraph machines for general public use in transmitting messages over wire lines, they failed to come into more than limited use. The lack of appeal was due in part to limitations of the machines themselves and in part to the inadequate quality of transmission over the wire-line media.

VI. TELEGRAPH OVER SUBMARINE AND OTHER CABLES

The expansion of telegraphy naturally brought with it the problem of finding the best way to cross bodies of water. In 1847, gutta percha, an inelastic latex, was discovered and found to be a very reliable insulator when submerged in water. In 1851, England was linked to the Continent when a telegraph cable was successfully installed across the English Channel, using gutta-percha insulation. In 1854, Cyrus Field, a retired American merchant, only 35 years old, conceived the idea that a submarine cable under the Atlantic Ocean would be a good business venture. He organized American and English companies and after a variety of failures succeeded in establishing a cable connection in 1858. The cable failed after three weeks of operation and he had to reorganize the company to raise the funds necessary for further work. In 1866, cables were permanently established. The application of regular telegraph equipment proved rather disappointing in that not more than two words could be transmitted per minute.

The submarine telegraph cable can be electrically simulated to a good approximation by considering the series resistance of the conducting wire and the shunt capacitance of the insulation. A consideration of the time interval needed to establish a steady value of current shows that the signaling speed should have been very slow.

In 1855, Lord Kelvin of England studied the speed at which telegraph signals could be sent and showed that, for a given type of cable, doubling the length resulted in slowing the speed to one-fourth; in other words, the signaling speed was inversely proportional to the square of the length. This forecast the slow speed found on the Atlantic cable. Stimulated by this adverse performance, Lord Kelvin devised a mirror galvanometer which raised the speed to seven or eight words a minute, and in 1870 his siphon receiving mechanism raised the speed to about 20 words per minute. From a modern vantage point it is known that the signaling speed depends on the frequency bandwidth which in the first submarine cable would have been extremely narrow indeed. This understanding explains why it has been economical to abandon the early submarine cables even though they may be in working order, and represent millions of dollars of investment. It is not the physical state of the cable but its ability to transmit the maximum information per unit time that is important. But transferring the point of view back to the epoch of 1870, the alternative to the cable was to send a message by steamboat which required ten days or more for a passage in one direction. Hence at the time the cable was a high-speed system.

Later, in the 1880s, Oliver Heaviside, father of the operational calculus, made a rigorous analysis of the transmission of telegraph signals and showed that the uniform addition of inductance along the line could produce a "distortionless" line with unlimited speed. Once again there was a difference between the theoretical ideal and that which could be practically achieved. Four decades later the spiral winding of a ribbon of a new magnetic material (permalloy) around the submarine-cable conductor made possible a speed of 400 words per minute.[1] In telephone cables it was found advantageous to add the inductance in a lump at uniform intervals by inserting wound

[1] See "References" at the back of this volume for journal articles describing advances made in submarine-cable technology during the early twentieth century.

inductances, called loading coils. This process was known as "loading" and is discussed in Section 4.1.3 of Chapter 4 and in Chapter 10.

VII. EARLY HISTORY OF TELEGRAPH SWITCHED SERVICES

The history of the telegraph exchange is older than that of the telephone exchange. In fact, some of the fundamental ideas which have lead up to the establishment of telephone exchange systems were developed for application with the telegraph.

The early history of the telegraph exchange is described in fairly complete form in an article by Thomas D. Lockwood, entitled "The Forerunners and Genesis of the Telephone Exchange." This article appears as an Addendum to the Proceedings of the Fifth Annual Convention of the Telephone Pioneers of America, which was held in San Francisco during 1915. The interesting history contained in Mr. Lockwood's article is quoted at length below:

> The idea of an exchange system for the electrical intercommunication of intelligence was an old and familiar one long before the advent of the telephone; and it is clear that a telegraph exchange is as natural an outcome of the electric telegraph, as the telephone exchange was and is of the electric telephone.
>
> Several such telegraph exchanges were proposed, and some were constructed and permanently operated. The first practical electric telegraph was that of Cooke and Wheatstone, in England, which was operated as early as 1837. The first telegraph exchange proposition was fourteen years later, and is presented in the form of a British patent No. 13497 granted to a French inventor, F. M. A. Dumont, February 7, 1851. The plans of this patent were comprehensive and included a number of central stations in the same city united by trunk circuits. That is, each central station had lines converging to it from the substations of its own group; and each was also connected by trunks with the principal central station. Several forms of switchboard were shown and described, but in the arrangement evidently preferred by the inventor, each line entering the central station passed first to a two-point switch. The normal point of this switch connected through a call instrument to earth; and the other or operating point connected through a Dial or A. B. C. telegraph instrument to the switchboard proper, which was shown as comprising a number of circular boards having contact points arranged near the edge, and two switch levers pivoted at the center and in contact with each other. There were as many of these circumferential contact points as there were lines in the system, and the corresponding numbers on all of the circular boards were united by wires at the back of the boards. To unite any two lines for through transmission, the two-point switches were first moved from the call to the switchboard branches, and then on the circular switch of either the calling or called line the two central switch levers were turned to the contacts corresponding respectively to the two lines concerned.

This system was afterwards described in Blavier's *Treatise on the Electric Telegraph*, published in 1867, as being actually in use in France.

Also at Colchester in the south of England and at Newcastle on Tyne, in the north, the British Post Office for many years prior to the telephone operated small telegraphic exchanges, employing Morse instruments in the former, and the Wheatstone A. B. C. Dial telegraph instruments in the latter system.

Several systems of this kind were also in regular use within the United States. Probably the earliest of these was one operated in New York by the Private Line Department of the Gold and Stock Telegraph Company. This was organized in 1869, and survived until 1880 when its supersession by the telephone exchange, which had long been inevitable, came about. Its history was outlined to me in the autumn of 1881 by Mr. G. L. Wiley, at that time Assistant Superintendent of the Metropolitan Telephone and Telegraph Company, but who had previously held a similar position with the Gold and Stock Telegraph Company.

At the time of its inception some twenty-five banks were connected by direct wires (one wire for each) with a central office of the Gold and Stock Company, at the Clearing House. The general business of that central office at first was to report to each bank in the system by means of printing telegraph instruments, its daily debit or credit balance, and to repeat to any bank, any message received for it from any other bank. But it also was the occasional practice to connect by means of a switchboard, (the Jones Lock Switch) the wires of any two banks desiring direct communication.

In 1871 this central office was moved to the General Office of the company at 61 Broadway, and from that time until the system was superseded by the telephone exchange, by means of a pin switchboard, a suitable battery on a loop with terminal plugs, and a supervisory central printer, any bank was on request connected with any other bank, or with the Clearing House. The system was both useful and popular. All the wires were permanently connected each through its own relay to the switchboard and the relays during a through connection were left in circuit, and of course rattled and clicked, all the time that the two banks concerned were working together by their printing telegraph instruments.

When the subscribers were finished and were ready for disconnection, they were instructed to make three long pulsations as a signal. But they were very much like the early telephone subscribers, and usually forgot it. However, when they failed to give the desired signal, and the relays were silent for some time they were disconnected without signal.

When in 1881 I was compiling this information, I received many letters in relation to it, and among others one from Edward A. Calahan, the practical organizer of the "ticker" system; who said,

> "I find by referring to my diary of 1869 that on Saturday, April 10, 1869, the first bank instrument was placed in the Clearing House and twenty-seven banks were connected between that date and September

> 7th of the same year. . . . At the time referred to, you may remember that in principle we were doing precisely what the telephone exchanges are now doing."

During the later years of this particular telegraph exchange an offshoot of it, similarly constructed and arranged, was installed and operated at Pittsburgh by T. B. A. David, the agent there of the Gold and Stock Telegraph Company. But here as elsewhere a printing telegraph could not satisfactorily compete with the telephone, and in due season the system was superseded by and became the nucleus of a telephone exchange established in the interest and under the instructions of the Western Union Telegraph Company.

During these years also there was a telegraph exchange employing Morse telegraph instruments, operated in Philadelphia. It was organized by the late Henry Bentley of that city as a personal business of his own; formed the nucleus of an early telephone exchange established at Philadelphia in the Western Union interest, which exchange later was merged with the Bell Exchange to constitute the Bell Telephone Company of Philadelphia. The switching apparatus used to connect these Philadelphia telegraph lines was for a number of years very crude; but in 1876 the Western Electric Manufacturing Company made for this system a standard Western Union pin switchboard. In this board each upright conductor was connected at its lower end through a springjack to the line leading to the subscribers' station. An adjustable resistance was connected in many of the lines to equalize their resistance; this being quite desirable when several lines were supplied from a single main battery. The horizontal switchboard conductors were grouped in pairs, with a battery in circuit permanently with each pair.

The lowest cross conductor was not paired but was used as a ground plate to which all lines when at rest were plugged, a common battery being connected between this crossbar and the earth connection.

Whenever two lines were united in this system for through communication, their plugs were transferred from the ground plate to the two cross conductors of any pair, so that the circuit extended from one of the lines to one crossbar of the pair, then through the battery of that pair, and from thence to the other crossbar of the same pair, and out to the second line. To establish such a connection, a call coming in on the sounder of any line for any other line was answered as usual by the central office operator, and upon the calling subscriber stating his desire, his line was immediately switched to that for which he had called, and he was instructed himself to give the call of the wanted station. The subscriber, with whom communication was desired, hearing his call would respond to it, and the two would at once converse with each other telegraphically.

But perhaps the best known of these pretelephone exchange systems was that of the Law Telegraph Company, New York.

This was planned and put into operation either late in 1874 or early in 1875, by William A. Childs, a gentleman of much enterprise and ability, who became manager of the company, and who is at the present time a member of the Pioneers. His idea was at first merely to extend telegraph lines equipped with alphabetical Dial instruments from the New York Courthouse to the offices of lawyers, and other subscribers; but this plan was soon abandoned on account of the number of wires centering at the Courthouse that would be required .

A new plan that then formed itself in the mind of Mr. Childs was to connect the offices of lawyers, and others, who would consent to join the system, by special and private lines to a central office, where there should be a switchboard, and auxiliary apparatus, to be manipulated by a force of trained operators in uniting any two of the lines on call. As in the first plan, the communicating instruments were to be of the alphabetical Dial type, the transmission being effected by keys corresponding to letters or other characters, and the reception of the message by a needle pointer, centrally poised and arranged to move within and around a dial indicator marked in correspondence with the keys. The characteristic feature of this system was the "call wire". In addition to the special telegraph wire of each substation, there was also this call or signal wire extending from the central station to a considerable number of subscribers' offices, looping into them one after another, and furnished at each with a Morse key and sounder, or a key alone. The key was a strap key with a back or resting contact only, and connected through this back contact with the call line in one direction, and through its own substance with the same line toward the other direction. The key when depressed acted to open the circuit of the "call wire" which was normally closed by being grounded both at the central office end and the outer terminal. At the central office the call circuit passed to ground through a key similar to those at the substations and also through a relay controlling a local circuit including a register and bell; the relay being arranged to close the register circuit through its back contact, so as to be responsive to the opening of the circuit by the substation keys. This "call wire" was used only for ordering connections and disconnections.

The direct intercommunicating subscribers' line circuits were at the substation end closed to ground after passing through the magnet of a call-bell, but at the central station end were open when not in use, and terminated each in a flexible cord and plug which hung down loose over the face of the switchboard.

The switchboard was a simple affair, consisting merely of a board with a number of metal connecting bars extended across its face, and painted with several different colors so as to be readily distinguishable. A considerable number of plug holes were also made in every bar. Each bar was divided at the middle of the board so that each was really a pair of bars in line with one another. Each pair had its own battery, furnishing working current for the Dial instruments, and the several batteries were permanently looped between the two bars of each pair. The direct Dial lines to the substations were not ordinarily connected with any battery, so that the object of the looped battery was to vitalize the Dial instruments.

The subscribers' stations were designated by numbers, and the calls and disconnections were always ordered by number; not by name.

In the operation of this system when any subscriber desired to communicate with another, for example, 21 with 369, No. 21 would begin by pressing his key to transmit over the call wire the desired number 369, followed by his own number 21. This operation broke the circuit in conformity with the numbers transmitted, permitting the armature of the call wire relay at the central station to fall back and close the local circuit a corresponding number of times, repeating the call number upon the register or sounder. An operator at the central office, stationed always at the register, at once observed the recording of the called and calling numbers, and repeated them aloud, saying "369 for 21"; whereupon a separate operator, whose business it was to do the actual switching, would signal the wanted line by taking up the plug of 369 and bringing it into contact with a metal call current bar connected with one pole of the calling battery, which was grounded at its other pole. The plugs of the two lines concerned were then placed in holes of the two connection bars respectively of one of the pairs. A complete telegraphic circuit extending between substations 369 and 21, through the common central station battery, was thus established. 369 and 21 would then communicate by means of the Dial instruments as long as they pleased, and when finished would order disconnection by means of the call wire, in the same way that the connection order was transmitted.

This system is of considerable historic interest. Its essential principles were retained when, after the advent of the telephone, the "Law" company acquired the right from the Bell Telephone Company of New York to sublease a limited number of telephones, so that it might transform its telegraph exchange into a telephone exchange. This it speedily did, and as a telephone exchange it was operated successfully for a number of years. When thus operated the characteristic combination of the call and direct wires was still employed, but the Dial instruments being superseded by telephones, each substation was provided with a switch to transfer the substation telephones between the call and direct wires. The central operators at the call wire were also equipped with telephones, and listened for call and disconnect orders continuously; and later, a form of multiple switchboard was incorporated in the system, materially enhancing its efficiency. The "Law" system found favor with a number of the "Bell" operating companies; was for a long period of time installed in the Philadelphia Exchange; was extensively employed by the Southern Bell Telephone Company; was highly thought of by the late George F. Durant of the Bell Telephone Company of Missouri, being installed at St. Louis for many years; and a closely similar system was during 1883, and for a time before and after, also to be found in Oswego, N. Y., under the supervision of our fellow member, Mr. Martin J. Joyce, whom I there and then met for the first time.

Another interesting feature of the Law system was its use from the first of numbers to designate its subscribers,so that in operation, names were

> never mentioned. This practice prevailed as previously noted, even in the telegraphic days of the system, and was continued when telephones were substituted for the dials, thus constituting a telephone exchange. The Law system was the first to thus employ numbers, and from it, the practice gradually spread to other exchanges; until its advantages being universally recognized, it is now adopted in telephone service all over the world.

It is evident from the above references which Mr. Lockwood made to the early history of the telegraph that telegraph exchanges consisting of a number of lines convergent to a central office and arranged for interconnection through a switchboard were not a new thing when the telephone was generally introduced commercially in 1877. It is also interesting to note that these telegraph exchanges formed a background from which the telephone exchange was developed.

Telegraph exchange service, however, at the time of the introduction of the telephone, soon ceased to appeal to the subscribers since they favored the telephone. This was probably due to the fact that in order to use the manual "Morse" code of communicating the customers were required to spend considerable time and effort in learning the code. In the case of the A-B-C and most other systems, the communication was too slow. Although in the case of the printing-telegraph systems the written record was provided, the telephone was preferred because the early printers had not been developed sufficiently to be acceptable instruments to the subscriber. Commercially, the early telegraph exchange systems, therefore, passed into the background.

VIII. TELEGRAPH WITHIN THE BELL SYSTEM, 1876–1910

8.1 Experiments of Alexander Graham Bell

The equipments used in Alexander Graham Bell's experiments of 1875 which resulted in the transmission of speech sounds were referred to as a harmonic telegraph transmitter and a harmonic telegraph receiver. These names reflected the fact that the telegraph had a well-established position as a means of rapid communication. What Bell was attempting to do was to establish a multiplicity of telegraph channels over a single wire. He proposed the use of vibrating tuned reeds, each resonating at a different pitch and producing undulating currents, as the transmitting devices, and similar resonant reeds (or tuning forks) to respond at the receiving end. Bell saw that such a system sharing a single line wire would greatly reduce the cost of telegraph service. In recognizing the possibility of sending a number of different frequencies over a single wire, Bell foreshadowed by over 40 years the development of carrier telegraph systems.

By a happy accident, Bell discovered that a reed (or diaphragm) that was not resonant (or was clamped) could respond to a range of frequencies and thereby transmit the human voice when driven by sound waves. Nevertheless, Bell's financial backers, Gardiner G. Hubbard and Thomas Sanders, urged him to concentrate on the telegraph experiments as having more promise for financial returns. For a time, work was continued on both telephone and telegraph fronts, and U.S. Patent No. 174,465 covering both aspects was filed on February 14, 1876, and issued on March 7 of the same year. Anxious to get some return on the money advanced, Bell's backers later offered this and other key patents to the Western Union Company for $100,000, but were refused. It has been reported that two years later Western Union would have been willing to pay $25 million.

It was soon realized that a new communication technique had been discovered and the term "telephone" was applied, based on the concept that sound was reproduced at a far-off location. After development of a satisfactory variable-resistance transmitter (microphone) the telephone became an instant success. It provided the equivalent of face-to-face conversation by extending each participant's mouth to the other's ear through telephone instrumentalities. Following the prior practice of the telegraph, a single wire using ground return was employed. In fact, the first demonstrations of telephony were performed by using existing single-wire, ground-return telegraph circuits. This usage was satisfactory for the earliest customers who typically had a permanent connection from a merchant's office to his home or warehouse.

8.2 Bell System Policy—Dual Use of the Wire-Line Plant

The Bell patent granted in 1876 disclosed the means of providing either telegraph or telephone communications over a wire line. Consequently the innovators and entrepreneurs of the period appreciated the economic and service advantages of utilizing wire-line outside plant dually for both services. This philosophy has been the guiding principle of scientists and engineers throughout the history of the Bell System. Plant arrangements have been designed and provided to enable telephone (voice) and telegraph (non-voice) communication services to be accommodated either dually or on an either/or basis over the extensive common wire-line and radio networks.

8.3 Transmission Interference

With expansion of telephone plant and the interconnection of cities it was necessary to mount a number of the single-wire, ground-return circuits on each crossarm of a telephone pole. It was then found that a person talking on one connection could be heard by a person listening on the adjacent telephone connection. This phenomenon is known as "crosstalk" and is a typical "systems" problem. A single circuit was satisfactory but a complete system, involving a number of such single circuits in close proximity to each other, i.e., no longer in isolation, becomes subject to other natural laws. Crosstalk was greatly mitigated by using a second wire for the return current instead of the earth. This use of two wires for a single conversation was referred to as a metallic circuit, and since the very early days have been almost universally used for telephone communications in the Bell System.

The problem of crosstalk continued to be acute as telephone lines were extended further and further. The solution involved an arrangement in which each wire of the second metallic pair was periodically reversed, i.e., transposed, in its position with respect to the first pair. This caused the crosstalk induced in the section following the reversal to oppose the crosstalk induced in the section before the reversal, thus canceling completely in the ideal case. The third metallic pair required transpositions with respect to both the first and second pairs. As more pairs were added the scheme became increasingly complicated. It is discussed in more detail in Section 5.3.1 of Chapter 4.

8.4 Bell System Entry into Telegraph Private-Line Service

The first ten years of expansion in the telephone business were confined to urban areas where the demand for service was greatest. Communication between cities was provided solely by means of telegraph. Most of the telegraph service within the Bell System was on a message basis, but about 1879 private-line service was provided between New York and Philadelphia and also between New York and Boston. The telegraph lines terminated on the customer's premises and were available for his operators to use between specified hours of each working day.

In 1886 the Bell System completed a long-distance telephone line between New York and Philadelphia, again reflecting the mercantile community of interest between those two cities. Because of contract licenses and technical difficulties this initial telephone service was provided between booths set up in public places. Some slight skepticism greeted the proposal by the Bell System that customers lease a private line and use it for either telephone or telegraph service. L. H. Taylor and Company, a brokerage firm, was the first customer. They expressed the belief that telegraph would be needed to insure reliability to the operation. Service started on January 15, 1887, and it is interesting to note that 75 of the first 76 messages transmitted over the private line were telephone calls.

Expansion of private-line telegraph service continued; next to Boston on January 2, 1888, for the Globe Newspaper Company; and to Chicago on December 8, 1892, for the brokerage firm of Hubbard, Price and Company. In some cases the private line provided only telephone service, but in others arrangements were made for either telephone or telegraph, but not both simultaneously.

Service on these early private telegraph lines was by Morse operation over a ground-return circuit, a procedure which was to be followed until private-line service using printer operation was first introduced early in the twentieth century.

8.5 Means of Providing Separate Telephone and Telegraph Circuits Over a Two-Wire Line

The rapid and extensive growth of telephone service into all sorts of business offices and between many cities resulted in the building of a large wire plant. In order to provide private-line telegraph service over telephone wire plant used simultaneously for telephone, two means were implemented.

The simplex network, first used on a regular basis in 1888, is analogous to the Wheatstone bridge. The transmitted telegraph signal is applied to the middle point of a large inductance called a retardation coil the two outer terminals of which are connected, one to each side, to the metallic telephone circuit. At the distant location the telegraph apparatus is connected to a similarly arranged retardation coil. At both ends of the telegraph circuit the apparatus was connected to ground, a procedure which was generally followed until 1919. These arrangements are shown schematically in Fig. 7-5a.

The successful application of the simplex method depends on the assumption that the two halves of the retardation coil are identical and that the two telephone wires are identical. In that case the telegraph current is split exactly in half and these two halves oppose each other in their effects on the telephone apparatus, thereby canceling and having no net effect. In the real world it is not possible to produce identical halves and so it was necessary to establish tolerable limits or standards concerning the inequality, i.e., the amount of unbalance. In the early days these were pragmatically reached by judgments and experimentation.

About 1900 a simplex circuit was developed which employed a 4-winding transformer called a repeating coil. Two of the windings were connected in series, the outer points being connected to the telephone line, and the inner points being connected together provided the simplex connection for the telegraph apparatus. The remaining two windings were connected together and inductively coupled the voice currents of the line and telephone apparatus, and excluded the conductive current of the telegraph apparatus. Figure 7-5b shows this arrangement.

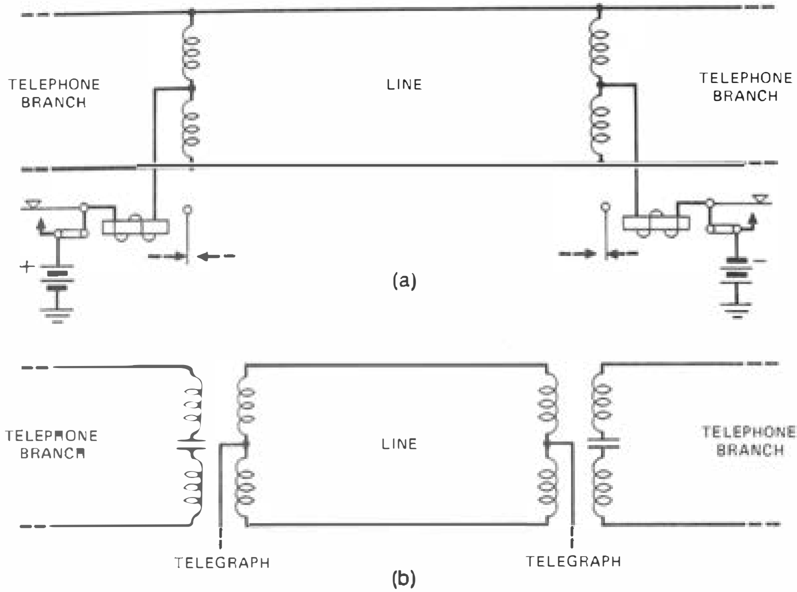
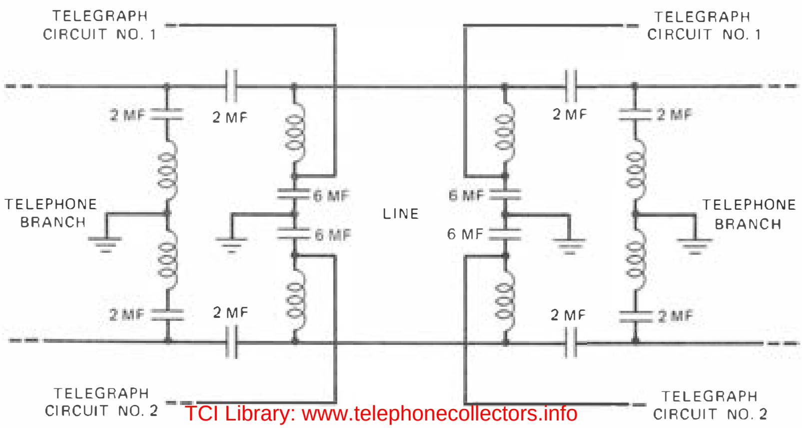


Fig. 7-5. A telephone line may be adapted for telegraph transmission without mutual interference by the use of a center-tapped induction coil, as at (a), or a transformer with a center-tapped secondary, as at (b). (Redrawn from Hamilton 1945, Fig. 2)

The simplex circuit provided one ground return telegraph channel for each metallic telephone circuit.

The composite set was first used in 1892. It is an assembly of apparatus which enables each wire of the telephone pair to be used as a telegraph channel. Telegraph was considered to be a dc operation and voice was recognized as alternating currents. Therefore, a conductive path was provided for the direct current of the telegraph, but it was excluded from the telephone. Figure 7-6 is a circuit schematic of a composite set.

Fig. 7-6. By using composite coils, a single telephone circuit can provide two telegraph channels with no mutual interference. (Redrawn from Hamilton 1945, Fig. 4)



By 1910 the application of Fourier analysis revealed that frequencies up to 50 or 60 hertz made important contributions to digital-type telegraph signals. After Campbell's invention of the filter in 1915 it soon became common to think of telegraph signals as occupying a low band of frequencies ranging from 0 to 80 hertz and the composite set was recognized as a low-pass filter. The analog-type voice currents occupied a higher band which was subsequently established as 300 to 3,200 hertz for good-quality telephone transmission. The telegraph apparatus was connected between the composite set and ground return. Since two telegraph channels were derived from one metallic telephone circuit, the composite had twice the efficiency of the simplex arrangement in so far as the number of telegraph channels was concerned.

8.6 Service Problems

As the length of the open-wire telephone lines increased, the technical problems already mentioned became increasingly important and lines were often described in terms of being short or long. Neutral operation, in which current/no-current corresponded to marking/spacing, proved satisfactory up to about 1904 but was then replaced by bridge polar duplex on composited circuits. A ground-potential compensator was also developed to overcome differences in earth potential, which were sometimes caused by dc railways. This work was done in 1913. Another of the improvements made to the dc ground-return circuits during this period was the addition of a neutralizing arrangement to cancel the mutal interference, called crossfire, between the circuits of a group.

As the length of a line was extended, a point was reached such that the addition of another section of line would cause errors due to the reduced current. To overcome this difficulty, a device, known as a single-line telegraph repeater, was inserted between the two sections of line. The repeater was first employed in 1892 and, initially, consisted of two simple relays. Each line section was connected through one relay winding and the front contacts of the other relay. A local battery was connected through additional windings and contacts so that transmission could always take place from an open-circuited line to a closed-circuited line. The circuit schematic of the repeater is shown in Fig. 7-7. This repeater was a simple device that served very well for Morse operation with signaling speeds of about 10 to 12 dots per second. Although it would send forward a full value of the signal current, it could make no correction for changes in the duration of marks and spaces. (These changes are called timing errors and result from distortion of the waveshape of the modulated dc signal.) As a result, the employment of more than a few repeaters in tandem would cause transmission errors due to accumulated timing errors.

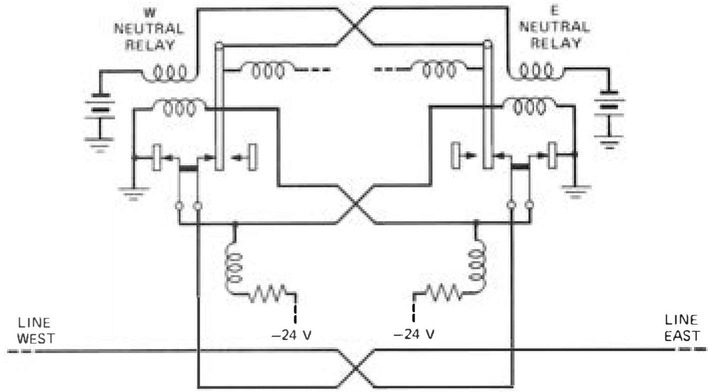


Fig. 7-7. Simplified schematic of a single-line telegraph repeater. (Redrawn from Hamilton 1945, Fig. 6)

Repeaters were initially used to extend the range of operation and were standardized in 1893, a year after initial trial. They were later employed for connecting together dissimilar facilities, especially in setting up private-line telegraph networks, and for connecting the customer's loop to the trunk line. This use of the repeater permitted instant reversal of the direction of transmission in a half-duplex operation, i.e., one where transmission is in one direction at a time. This application of the repeater gave the Bell System a substantial advantage over its competition in private-line service.

In 1905 a New York–Pittsburgh–Chicago–St. Louis circuit was considered to be a very large network and was in fact difficult to keep in satisfactory operation. Maintenance was a constant problem, requiring the attention of skilled men. With the passage of time, semiautomatic keys raised the signaling speed, until by 1905 it had reached 16 to 18 dots per second.

8.7 Establishing and Maintaining Private-Line Networks

As covered earlier, from the beginning the Bell System recognized the value of utilizing its local and intercity wire-line plant for various forms of public telephone and telegraph services. However, from the time of Alexander Graham Bell's experiments which demonstrated the feasibility of voice telephone service, public acceptance and demand for telephone service far exceeded that for telegraph. This was mainly due to the fact that suitable telegraph instruments (printers) were not to become available for general public use until about 1910. Communication over telegraph networks was by Morse telegraph, which required special training of operators. Consequently, until around 1910 the Bell System restricted its development of Morse telegraph service to that involving private-line networks in contrast to the telephone service which from the start became a public universal switched service.

In telegraph service the Bell System followed the precedent of telephone service by owning and maintaining the equipment at the customer's premises. It was necessary therefore to have a general plan which would facilitate access to the lines and loops which terminated in the telegraph sets. This was accomplished by having a common location, generally in the telephone office, where the lines, subscriber loops, repeaters, batteries, and testing apparatus terminated on jacks and plugs. Various testing arrangements could then be set up by plug-ended cords. The telegraph equipment and lines were tested at the telegraph test board. In cases where lines were found faulty the troubles were reported to the maintenance forces who used telephone test boards to locate and clear the troubles.

In the early days the testing arrangements were handled on a job basis and consequently varied from one location to another. Often they were associated with the arrangements for setting up and changing the private lines. Later they were often close to or part of the telephone test board. The continuing growth of Morse telegraph private-line service reached such proportions shortly after the turn of the century that it was deemed desirable to standardize a test board which could serve all the larger cities. It would then be possible to manufacture and stock the boards, and then install them on order, with a minimum amount of engineering.

The first of these standardized test boards was called the No. 4 Morse Board. It came into service about 1910 and was of the single-position-switchboard type. Subscriber loops and test equipment were terminated on plug-ended cords in the keyshelf. Jacks in the upper part of the vertical panel gave access to lines, positive and negative battery, and repeaters, and were connected to jacks in the lower part of the panel into which the subscriber loops and testing equipment could be plugged. Several subscribers could be connected if required. Double-ended patching cords were used to set up desired testing arrangements. A simplified circuit schematic of the No. 4 board is shown in Fig. 7-8.

8.8 Status of the Telegraph Business in the United States in 1910

Growth of the telegraph business in the United States from its introduction by Morse during 1844 and on through into the twentieth century has been highlighted in Section V. During the late nine-

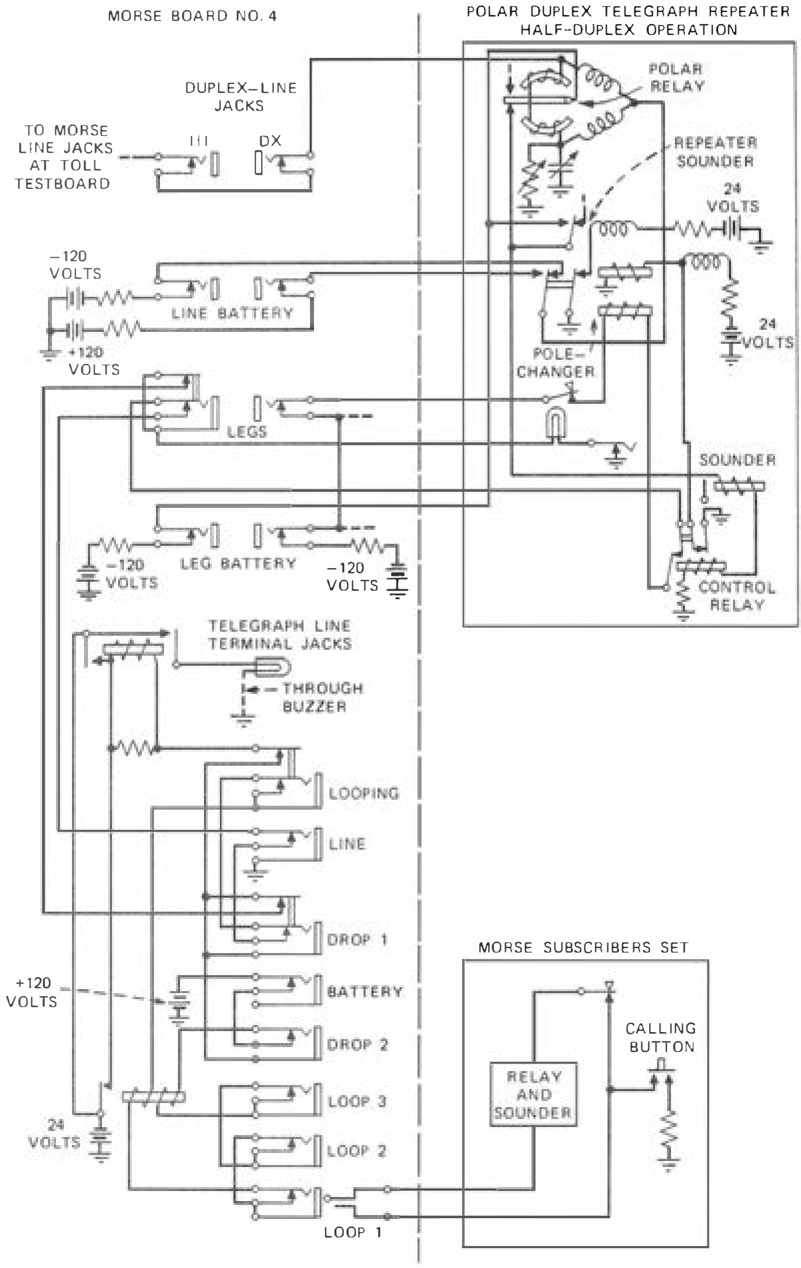


Fig. 7-8. Simplified schematic of No. 4 Morse Board. (Redrawn from Wier 1935, Fig. 1)

teenth and early twentieth centuries, telegraph communication services had become a most important means of transmitting information in fields of banking, business, government, railroads, and news gathering, and in the sector of the private citizen when expedited record communications were desired. It became a preferred mode over mail services. During this adolescent period in the growth of telegraph in the United States most of the services were furnished by the Postal and Western Union Telegraph Companies and by the Bell System. The telegraph companies were competing strongly for public "message telegram" business and for the private-line business of heavy users of the service. Meanwhile, the Bell System restricted its service offerings to private-line customers. Although there continued to be a limited number of small companies who were competing for both types of business, their survival was to be short-lived because they could not afford to build wire-line plant and operating and test centers in duplication of those owned by the large telegraph companies and by the Bell System.

As will be covered in the next section, the telegraph printer had not as yet been adopted for use by Bell System customers. Nonetheless, where large amounts of information needed to be disseminated in an expedited manner, such as press association news and the like, the professional telegraphers in employ of the users were capable of transmitting and receiving information, and in turn recording it on a typewriter, at a continuous rate of as much as 35 words per minute. These expert telegraphers became capable of this high attainment by using semiautomatic key senders ("Bug Keys") to transmit the information in a code known as the Phillips code (named after its compiler). Abbreviations for common words and groups of words within the basic Morse code were used (example: POTUS for President of the United States). During reception the receiving operator transcribed the Phillips-code aural signals into full English text on a typewriter.

At the turn of the twentieth century the classical method of furnishing public message-telegraph service, i.e., the telegram, had become ubiquitous within at least the large communities throughout the United States. The Postal and Western Union Telegraph Companies were formidable competitors in promoting the service particularly within and between metropolitan areas. Both companies had established operating and test centers at which their respective open-wire interconnecting plant was terminated. The telegraph companies employed boy messengers to pick up and deliver telegrams offered by the public. Alternately, whenever the customer preferred to dictate his message to the telegraph business office via telephone, the delay of messenger pickup was avoided.

IX. ADVANCES IN NON-VOICE SERVICES WITHIN THE BELL SYSTEM, 1910–1925

During the late nineteenth and early twentieth centuries, the Bell System concentrated on development of its telephone business. Meanwhile there was relatively limited activity on development of its Morse telegraph private-line business. Many scientific and engineering advances were made to improve telephone services. They were applied to the open-wire and cable plants, to customer station arrangements, to switchboards and systems located in central offices, and to facilities located in test and maintenance centers. The technical advances made and applied to the telegraph plant were limited to those described in previous sections of this chapter.

At the turn of the century a majority of the local telephone operating companies that comprise the present Bell System had come under general management of the American Telephone and Telegraph Company and the Western Electric Company had been acquired. The Long Lines Department of AT&TCo had the responsibility of building, maintaining, and operating interregional plant and of providing long-distance telephone and private-line Morse telegraph services. Headquarters Engineering within AT&TCo performed the function of coordinating Bell System construction programs, conducted communication research, and specified characteristics of new products and changes in existing products needed by the operating companies. Headquarters Engineering within the Western Electric Company was responsible for research and development programs including design of products for manufacture by Western Electric for subsequent use by the Bell System.

During 1910, under the able guidance of Theodore N. Vail, a grandson of Alfred Vail, the American Company acquired sufficient stock in the Western Union Telegraph Company to have a controlling interest. Western Union had at that time become the largest of the competing telegraph companies engaged in furnishing public telegraph (telegram) service and in the provision of private-line circuits for Morse telegraph use. Evidently higher management of both companies viewed the acquisition as being of advantage to both companies and to the public. The extensive wire plants of the two companies became available for either telephone or telegraph services and the scientific and engineering effort and advances being made by both companies could be brought to bear toward development and introduction of new telegraph services.

This acquisition of Western Union by AT&TCo and knowledge that there was a sizable public demand for a printing-telegraph communication service may be regarded as a turning point in Bell

System policy as regards development and promotion of non-voice communication services over its wire-line network. The impetus was provided for development and introduction of a variety of products and systems for application within the plants of the Bell System and for use by its customers on their premises. There were indeed many advances made and applied during the two decades beginning around 1910. These will be described in the remainder of this chapter.

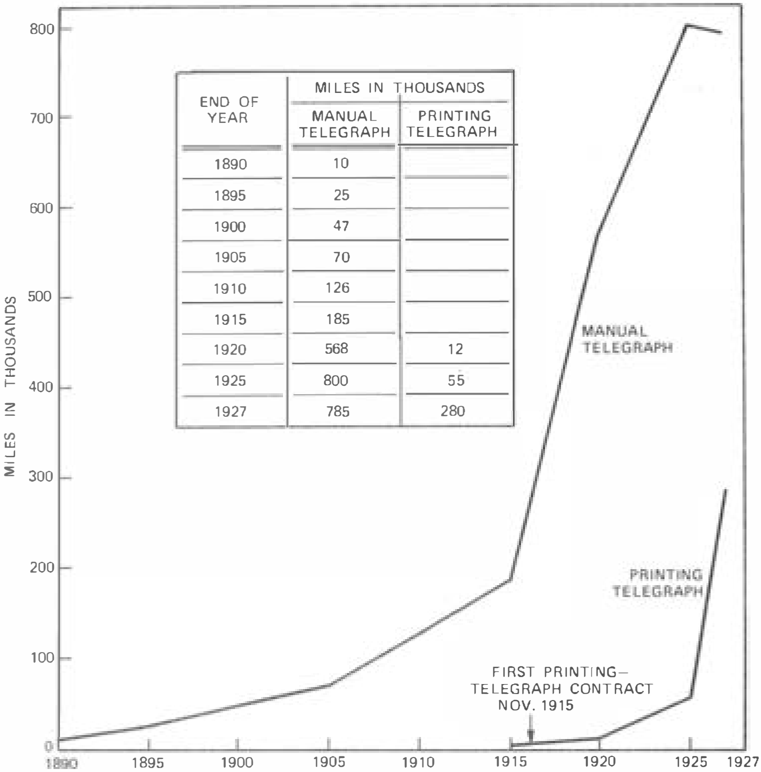
It is of interest that, due to legal problems, the intercorporate ownership and management relationships between Western Union and AT&TCo were terminated about 1914. However, by that time the development programs that were to lead to provision of multiplex telegraph and start-stop private-line printing-telegraph services were well along.

Figure 7-9 shows the growth of Bell System private-line telegraph contract mileages from 1890 through 1927. The growth of manual telegraph milage until 1910, although steady, was restricted to the use of the relatively unstable dc grounded telegraph facilities. As mentioned above, 1910 marked the turning point in Bell System policy regarding development and promotion of non-voice communication services over its wire-line networks. The effects of this new policy on growth are clearly shown by the data in Fig. 7-9.

9.1 Start-Stop Printers and the Multiplex Printing-Telegraph System

Following acquisition of Western Union by AT&TCo, Bell System management directed attention to modernization and expansion of its telegraph business. It was decided to follow the same procedures that had been successful in development of the telephone portion of its business. It was clear that a family of printing-telegraph machines having the same convenience in handling record communications as telephone sets provided for telephone communications were needed. It was also evident that the standards of performance of such machines in customer private-line networks would need to be high and the overall cost of the service to the users should be consistent with that of the existing Morse telegraph private-line services.

Investigations and studies that followed led to the conclusion that none of the printing-telegraph machines in use or being offered by suppliers would be suitable. Accordingly it was decided that the Western Electric Company should undertake development of such machines. The American Telephone and Telegraph Company initiated the first of these development programs in the form of a work order dated August 1, 1909, assigning the Western Electric Company to "the development of a page printing telegraph system for use on cable circuits of moderate length." Following soon thereafter, Western Electric was authorized to develop a multiplex printing-telegraph system.



END OF YEAR	MILES IN THOUSANDS	
	MANUAL TELEGRAPH	PRINTING TELEGRAPH
1890	10	
1895	25	
1900	47	
1905	70	
1910	126	
1915	185	
1920	568	12
1925	800	55
1927	785	280

Fig. 7-9. Bell System private-line telegraph contract mileages, 1890–1927.

Sections 9.2 and 9.3 describe the principles of start-stop printer and multiplex printing system operations and highlight the programs engaged in by AT&TCo and Western Electric over the period beginning during 1909 and extending into the latter part of the decade of the 1920s.

9.2 Start-Stop Page Printer Developments

The first design resulting from the authorization of August 1, 1909, was a type-wheel page printer using ink rolls. The type wheel was actuated by a spring which in turn was wound by an electric motor. Two models were built, but tests made in 1910 showed that the performance was unsatisfactory.

At about this point it was realized that a page-type telegraph printer was a complicated piece of machinery that would require serious design effort if it were to meet rigid standards of performance and reliability. The first step in meeting the needs was to establish a clear definition of the requirements which the machine should meet. In 1911, engineers at AT&TCo and Western Electric agreed to the following five items:

(*i*) Would operate over 10 miles of cable
(*ii*) Must use standard-width paper (presumably 8½ inches wide)
(*iii*) Would always be attended by an operator
(*iv*) Must be highly accurate ("accuracy is vital")
(*v*) Would operate at a speed of above 20 words per minute.

Looking backward in time, the stated requirements seem lenient, but when issued they were realistic and quite beyond the existing state of the art. The American Company and Western Electric were now committed to an earnest effort to provide reliable page-printers for use by customers of Bell System private-line telegraph service. Later, after improved machines were available, it was natural to extend the range of service. This need focused attention on the characteristics of the lines and resulted in objective methods for rating the telegraph lines and establishing satisfactory performance. It also brought about the regenerative repeater. These items will be discussed in subsequent sections.

The period from 1910 onward was one of rapid change as well as great difficulty in deciding the optimum procedure for handling the telegraph business. There were several companies in the business of producing machines, each with its own strengths and points of view. There was rapid growth and extension of the telephone plant with attendant effects on the performance of the printing machines, and there also was new technology entering the telephone plant. Insofar as the page printers themselves, there were questions of printing wheel versus type bars, moving versus stationary paper-carriage or its opposite: stationary versus moving type, and inking roller versus ribbon. There also were problems of paper feed, carbon copies, simple copy or billing forms and invoices, and a whole host of items important to the several types of users of telegraph private-line systems.

In such a rapidly changing scene it was difficult to foresee the ultimate objective. The expense of making radical changes with possible mistakes was too large to gamble. Consequently, improvements were incremental in nature, followed by new demands and more changes, all being modified by advances in telephone technology.

9.2.1 Short-Line Page Printer

Late in 1911 an improved short-line page printer, known as the cross magnetic printer, was produced. This machine used electromagnets and relays, and required two wires for operation. In 1912 the Western Electric Company produced 12 of these machines at an estimated cost of $200 each. Ten were installed by Western Union between main and branch offices. The other two were put on test between two offices in the Bell System. Experience with these machines focused attention on the need for a number of improvements in page printer design.

Work was started on two models of a short-line page printer in 1913 using the multiplex machine, which will be described later, and a so-called stepping distributor. Tests, which continued until mid-1915, showed performance of the short-line printer to be inferior to the two-wire cross magnetic printer which had earlier been placed in service by Western Union and the Bell System.

Once again customer demand ran ahead of available service and prompted further development. Three press associations requested the following short-line-printer services for distribution of news in the New York City area:

(*i*) United Press wanted one-way service from its New York Bureau Office to six receiving stations.

(*ii*) International News Service requested one-way service to five newspapers, including New Brunswick and Asbury Park, New Jersey.

(*iii*) Associated Press desired a circuit from its Bureau Office in New York to two newspapers in Newark, New Jersey. This service was to have both perforated-tape and direct-keyboard sending at the Bureau Office with a telephone circuit on the same pair of wires for return communications.

There was no inventory of equipment at the time of the first request so work was pushed to produce the needed machines, including improvements which could be made based on the experience to date. The United Press service was hooked up to New York Telephone Company circuits in November 1915. This date, then, marks the beginning of *private-line page-printer service*.

9.2.2 Start-Stop Printer and Regenerative Repeater Principles

At this point a review will be made of some of the principles and factors that related to the operations of start-stop printers and regenerative repeaters of that period.

A five-unit binary code was employed. There were two and only two possible values, such as current/no current, or black/white. If a pack of five cards, each of which is black on one side and white on the other, is available, then it is readily shown that there are 32, i.e., 2^5, patterns in which the five cards can be placed from left to right. In the case of the telegraph printer, the five units were short time intervals whose duration (in milliseconds) depended on the speed of sending (in words per minute); thus the interval was 33 ms for 40 wpm, 22 ms for 60 wpm, and

18 ms for 75 wpm. In each of the five time elements of the code there were only two possible conditions, and one of these was chosen: either the presence or absence of a current impulse.

Typical telegraph-printer copy contained the 26 letters of the alphabet, ten digits, and a variety of punctuation marks and other characters, all of which added up to considerably more than the 32 patterns in the five-unit code. The apparent discrepancy was resolved through use of the shift key. The alphabet was printed only in capital letters in the usual lower-case or "LTRS" position of the keyboard. The upper-case or shift position was called "FIGS" and had all the remaining characters. The six codes available after coding the 26 letters of the alphabet were used for instructions to the typewriter. They were (*i*) line feed, (*ii*) space, (*iii*) carriage return, (*iv*) letters, (*v*) figures, and (*vi*) blank.

Operation of a key on the keyboard caused five contacts to close (or open) according to the code of the depressed key, thus connecting (or not connecting) battery to the segments of the sending distributor. The signal could be sent directly into the telegraph line or it could be sent to a tape perforator, which punched holes in a paper tape for later transmission to the line.

Figure 7-10 shows the code, in the form of a perforated tape, for each letter, digit, fraction, punctuation mark, and other character. Figure 7-11

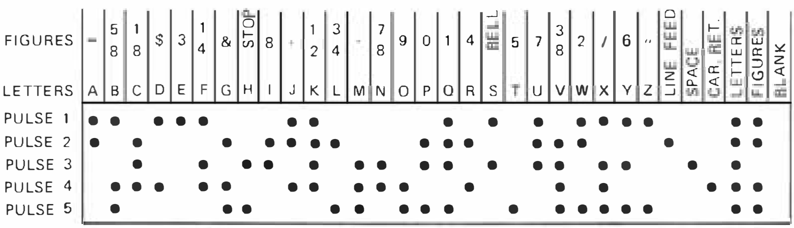


Fig. 7-10. Five-unit start-stop printer code. (Redrawn from Watson 1938, Fig. 1)

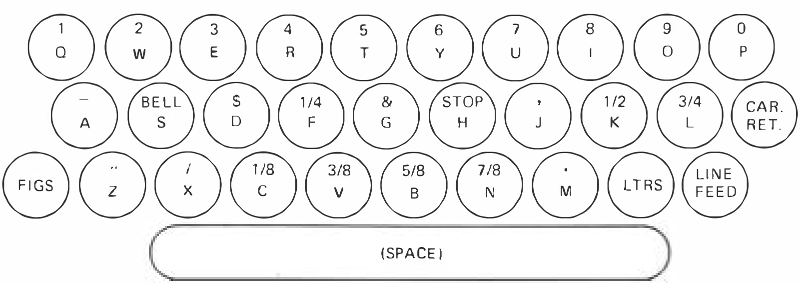


Fig. 7-11. Start-stop printer keyboard. (Redrawn from Watson 1938, Fig. 2)

shows the keyboard and the symbols corresponding to the LTRS and FIGS positions. Unlike the usual office typewriter, there were only three rows of keys instead of four.

In the start-stop scheme the five-unit code was preceded by a start pulse of the same unit length and followed by a stop pulse which, when a sending machine was running at a uniform rate, was 1.42 times that of the others. Today the stop pulse is generally 1.50 times the unit length.[2] These seven pulses were represented by circular segments on a sending distributor (see Fig. 7-12) and they were connected by a rotating brush to an inner ring contact which was connected to the line. In a synchronous system only five segments were needed, and so the two segments represented a loss in transmitting efficiency which was the price paid for the start-stop method of maintaining accurate response of the receiving printers. The rotating brush was normally held stationary by a latch. When the latch was released, the brush was driven by a shaft through a friction clutch, making one rotation and again becoming latched by the stop pulse. The depression of a key set up the codes on the five contacts shown at the bottom of the sending distributor of Fig. 7-12, thereby putting the proper condition on the segments over which the brush traveled. Figure 7-13 shows the current condition on the line when the letter A was transmitted. At each receiving station a similar latched brush was released by the arrival of the start pulse (an open from the line) and rotated in approximate synchronism with the transmitting brush for one turn; hence each transmitted pulse was distributed to the proper receiving magnet of the printer.

To insure constant speed for the driving motors, use was made of centrifugal governors in the earlier models. Later models took advantage of the constant frequency of the 60-hertz power systems and used synchronous motors. Figure 7-12 shows sending and receiving stations connected together by a telegraph line. A given private-line network might have a dozen or more receiving stations, of which the most remote might have been as much as several thousand miles away. The time of transmission depended on the combination of facilities used; and the speed of these could vary from 20 to 120 miles per millisecond. The great advantage of start-stop operation was that no matter how long it took the start pulse to arrive, the five units of the code would follow sequentially and the brush would stop before arrival of the next start pulse. Start-stop operation thereby brought flexibility to network operation, making it possible to add or subtract stations from the network without requiring adjustments. Small differences in motor speed did not seriously affect operation, but it was usual to keep variations within ± 0.75 percent.

[2] The lengthened stop pulse enabled bringing a system back into proper phase in case of a loss.

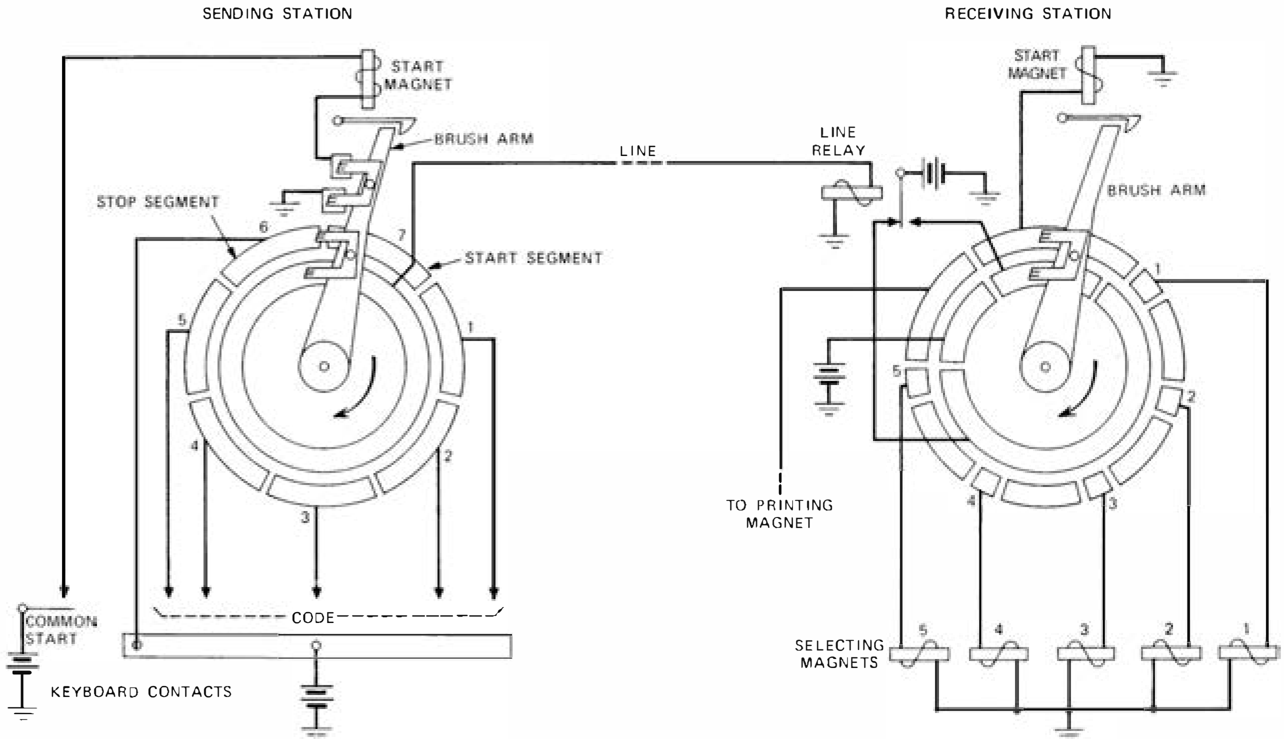


Fig. 7-12. Simplified diagram of start-stop system. (Redrawn from Watson 1938, Fig. 3)

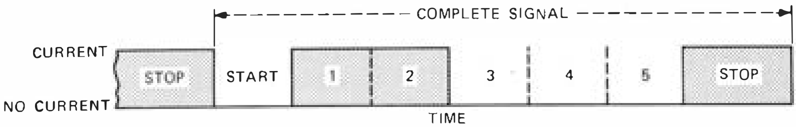


Fig. 7-13. Line signal for letter A. (Redrawn from Watson 1938, Fig. 4)

The selecting magnets of the receiving distributor were connected to rather small segments in the outer ring and there were large blank segments. The pulses of the code were generally distorted, sometimes on the leading side and sometimes on the trailing edge. However, the center of the time interval provided ample margin to give a true indication of the presence or absence of a pulse. Accordingly, a short sample was taken at the middle of each time interval (short segment) and the remainder of the interval (blank segment) was ignored. After the five elements of the code had been received, the printing magnet was energized by passage of the rotating brush, and the typebar made its imprint on the paper.

The commutator nature of start-stop operation shown in Fig. 7-12 did, of course, have the problems associated with wear. In later machines the brush and segmented-distributor were replaced by simple mechanical devices, sending contacts and a single receiving magnet. They operated in essentially the same manner, but were cheaper to manufacture and, most important, they improved the reliability and reduced the maintenance.

The nature of the five-unit-code regenerative repeater will be clear from the above discussion of start-stop operation. The receiving distributor identified the incoming code and delivered the information to a transmitting distributor which rotated in synchronism to send out pulses of the proper duration. Since the incoming pulse was not identified until the midpoint of its time interval, each regenerated outgoing pulse was delayed by approximately half the duration of a pulse in passing through the repeater.

Whenever a dc pulse was transmitted over a telegraph line its shape would be altered compared to the idealized shape that would be expected from the simple make and break of the transmitted signal. The change from the ideal shape is called distortion, and it depended on the equipment, the line, and also interfering effects from sources outside the telegraph system.

In Morse operation the effects were subjective and varied from one operator to another. The skill of the man receiving could compensate for a considerable amount of distortion, and he was aided by his knowledge of the English language in forecasting the next letter of ordinary copy. In printer operation the machine did not have any such innate ability; but its performance could be fairly well predicted from its design and adjustment. The effects of distortion could be considered in terms of the shortening or lengthening of the signal pulse. If it approached 50 percent of the pulse duration, the sample taken at the midpoint of the pulse interval would be in error and hence the code would be misinterpreted by the printer.

9.2.3 The All-Purpose Page Printer

Inquiries about service possibilities focused attention on the fact that there was need for a long-line duplex page printer. Requirements set forth in 1915 stated that the system should operate over long-distance telegraph circuits of the Bell System with all the flexibility obtained by Morse operation in private-line service, i.e., operate on a network with a large number of stations in such a way that any station could break and take control for sending. In addition, there was to be provision for a spare machine at each location, a home copy of messages sent, and provision for multiple received copies.

It was a period of considerable tribulation. The stepping distributor was proving unsatisfactory and it was decided to develop a start-stop distributor and to make the short-line printer and the long-line duplex printer identical. There were a variety of mechanical problems of noise and wear due to the violent action of the electromagnets in the printer as well as troubles relating to paper feed, inking of characters, and provision of satisfactory printed records. There were also electrical circuit problems including that of fire hazard. It seemed best to undertake a completely new design effort in the middle of 1916. The new printer would have a stationary paper-carriage, use type bars with an inking ribbon, make multiple copies, use cam-controlled mechanical action rather than electromagnet action, and have reduced noise, good appearance, and a speed of 60 words per minute. The objective was to drastically reduce maintenance to the point where a routine visit required for lubrication would be sufficient. Realization of these goals was still a long way off.

In 1917 a pioneer long-line duplex network operating at 42 words per minute was installed for the Associated Press between New York and Boston, with intermediate stations. On the opening day the AP established a Morse wire network in parallel with the printer network to guarantee delivery of the news in case the printer failed to operate satisfactorily. The fastest telegraph operator in the New York office was put on the Morse wire. He soon was ahead of the printer operator, which was a great disappointment to the printer people observing the operation. However, it was later admitted that the receiving operators at the newspapers on the network were unable to copy the high-speed

Phillips code being sent from New York, and quit trying. Consequently the trial ended with nothing available but copies of the news taken from the printer.

Operation of the AP network provided experience leading to improvements for testing and maintaining the lines and equipment and for handling emergency situations. It also provided those engaged in the design of the all-purpose page printer with information regarding performance of the early model in service.

It is of interest at this point to mention that by 1919 sufficient progress had been made in transmission technology to insure that an all-purpose page printer operating at 60 words per minute could be used in either local or intercity networks with equal assurance that the quality of transmission would be adequate. The dc metallic-telegraph system, introduced during 1918, came into use to provide reliable transmission of printer signals over a composite circuit derived from two pairs within moderately long intercity cables. The open-wire carrier telegraph system, introduced during 1919, provided ten suitable telegraph circuits over a 2-wire open-wire line.

During the 1918–1923 period, Western Electric incorporated improvements in the electrical and mechanical units making up the overall design. Flexibility for association of the various apparatus units comprising each complete installation was also provided in order that each installation could be tailored to meet the specific service needs of the user. Two categories of sets were found desirable; these became known as the 12-type set and the 13-type set. Subsequently, both types served well in meeting various customer demands for page-printer services. They were used extensively well into the decade of the 1930s, when improved designs came into use. These designs will be described later in this chapter, and in a subsequent volume of this history.

The 12-type set (Fig. 7-14) was usually used to furnish combined keyboard-sending and receiving service. For this service each installation required the following units:

12A printer
12A or 12B printer base
12A printer stand
12AL or 12AH keyboard distributor
12A, 12S, or 12S (modified) printer cover
215A relay
71510 or 71508 motor generator (for ac sets—to provide direct current for magnets of the printer).

If the installation was to be used for receiving-only service, a 12AL or 12 AH receiving distributor would be provided as a substitute for the 12AL or AH keyboard distributor.

Fig. 7-14. The 12-type printer set.

The 13-type set was designed to furnish both automatic-sending (perforated tape) service and keyboard-sending service as well as receiving service. Although the 13-type set could be arranged to provide keyboard-sending-only service or receiving-only service, the 12-type set was preferable because of its lower cost.

Each installation of the 13-type set required the following units:

12A printer
13A operating table
5A or 5B distributor
10A relay box
5A perforator transmitter
1B telegraph transmitter
10A distributor cover
13A printer cover
215A relay
5A copy holder.

9.2.4 The Western Electric Tape Printer

As discussed in some detail on the preceding pages, beginning during 1909, AT&TCo and Western Electric placed great emphasis on the design and supply of various models of page printers for use by customers of its private-line services. These machines fulfilled the communication needs of those customers whose message traffic was amenable to recording in page form. However, others, such as stockbrokers, whose communications were inherently short, the record of which needed to be acted upon expeditiously, felt the need for the message to be recorded in tape rather than page form. Consequently, during 1920, Western Electric produced a model of a simple start-stop tape printer the keyboard of which contained features to permit the printer to be used in networks containing start-stop page printers. Its design also permitted interfacing with the multiplex printing-telegraph system, the description of which is included in Section 9.3.

9.2.5 The Regenerative Repeater

The growing demands for extended private-line networks comprised of several stations equipped with start-stop printers raised requirements for a repeater, requirements which could not be met by the single-line telegraph repeater in private-line Morse telegraph networks. A mechanical start-stop regenerative repeater, the principles of which are described in Section 9.2.2, was developed in time to permit introduction into the telegraph plant during 1925.

Whenever a telegraph line was extended almost to the point where errors would result from any additional line, the start-stop repeater would be inserted. It would correctly interpret the distorted incoming pulses and send forward full-current-value pulses with correct timing, thereby completely removing any distortion introduced by the previous section of line. This concept of using a regenerative repeater at a point in the line where the detected signal was substantially error-free removed the limitations of length and complexity in establishing private-line networks for start-stop printer service.

9.2.6 The Cipher Printing-Telegraph System

During World War I, agencies of the U.S. Government, including its military forces, became quite alarmed because of the inadequacies of the "code book" method of enciphering Morse-telegraph and printing-telegraph communications. Regardless of the care used in developing the codes and the frequency of practicable change, expert cryptologists were able to decipher "secret" messages within a period of a few hours.

Gilbert S. Vernam, an AT&TCo engineer who was skilled in telegraph-printer technology, conceived of a means for automatically enciphering printing-telegraph messages on a character-by-character basis and, in turn, automatically deciphering them at the receiving end of a printing-telegraph circuit. The means of enciphering utilized the methodology of automatically mixing the five-unit code of each plain text character with a five-unit character code derived from a *random non-repetitive* source. The methodology provided a major breakthrough in the science of cryptology.

During May 1918, AT&TCo and Western Electric engineers started development of a printing-telegraph system having capabilities of automatically enciphering and, in turn, deciphering messages on a random non-repetitive basis. The system included an automatic machine perforator that was developed for the purpose and other Western Electric printer apparatus that was available at the time. The first systems came into use during 1918 in furnishing Government services between New York, Washington, D.C., and Newport News, Virginia. A large number were ordered for use by the American Expeditionary forces in France but were not delivered because of the Armistice.

Following the war, work was continued in a move to improve the system. Models of a simplified printing-telegraph cipher set utilizing a tape printer and a combination keyboard perforator and machine perforator were made. However, the apparatus never reached the manufacturing stage because of the lack of interest on the part of Government and public for hardware capable of handling secret communications during peacetime. Nevertheless, Gilbert Vernam and R. D. Parker conceived of a number of novel ways of implementing the random non-repetitive mode of encrypting and decrypting and obtained patent coverage of their ideas.

The disclosures in the inventions revolutionized the science of cryptology not only in the United States but throughout the world. During World War II the Western Electric Company and the Teletype Corporation became suppliers of a number of types of "secret" telephone and telegraph systems and of printing-telegraph apparatus for use by the U. S. Government and its allies in handling global communications. Readers are referred to a later volume of this history for a review of the systems that were provided during the World War II period.

9.3 The Multiplex Printing-Telegraph System

Development work on a multiplex telegraph system was started by Western Electric during 1912 primarily to meet the needs of the Western Union Telegraph Company. The system was intended to permit most efficient use of a telegraph line facility in transmission of a large volume of public "message telegram" traffic. The large amount of traffic passed over such lines between Western Union message centers accentuated the importance of obtaining a highly efficient use of the lines.

The multiplex system developed by Western Electric employed the Murray form of five-unit Baudot selecting code in such an efficient manner that the 20 selecting-code impulses required within the four-channel system obviated the need for additional impulses to maintain synchronism between the signal distributors located at the two ends of each line.

The general principles of operation of the system are shown on Fig. 7-15. The figure shows the four-channel system with associated sending and receiving printing apparatus at each end of a line. Assume that transmitter A at station X is sending to printer A at station Y. The transmitting-distributor brush arm at station X is rotating in synchronism with the receiving-distributor brush arm at station Y and the relative position of the two brush arms is so maintained that the pulse sent when the transmitting brush at X is on segment 1 arrives at station Y just as the receiving brush is passing over segment 1. In this way all five pulses necessary to select a character are transmitted to printer A. The brushes then pass to transmitter and printer B and a character is transmitted in the same way to B and to C and D in turn. The mechanical operations of the printing are carried out during the interval when the circuit is being used by other channels.

The segments of the receiving distributor are narrower than those of the transmitting distributor so that the pulse which operates the receiving mechanism is taken from the center of the received pulse, thus minimizing effects of distortion.

The complete system, including page printers, keyboard, paper-tape perforators, tape transmitters, and receiving perforators, in addition to the multiplex distributors, was delivered to Western Union late in 1912. The page printers, like the multiplex distributors, employed the Murray form of five-unit Baudot code, with selecting levers dropped into slots in disks. The printer, coded 1B, was designed to operate in a speed range of up to 60–65 words per minute.

To insure accurate and efficient use of the complete system, messages were typed out in advance and recorded in perforated-tape form for later feeding into the tape transmitters associated with the respective sending channels of the multiplex distributors.

In 1913, five sets of multiplex equipment were produced and delivered to Western Union for use on its New York–Boston route. Demand continued and by 1916 Western Union had 26 multiplex circuits in operation and orders for 25 more. Since some of the multiplex circuits had way stations between repeater points, a development known as "forking" was undertaken in 1916 to serve them.

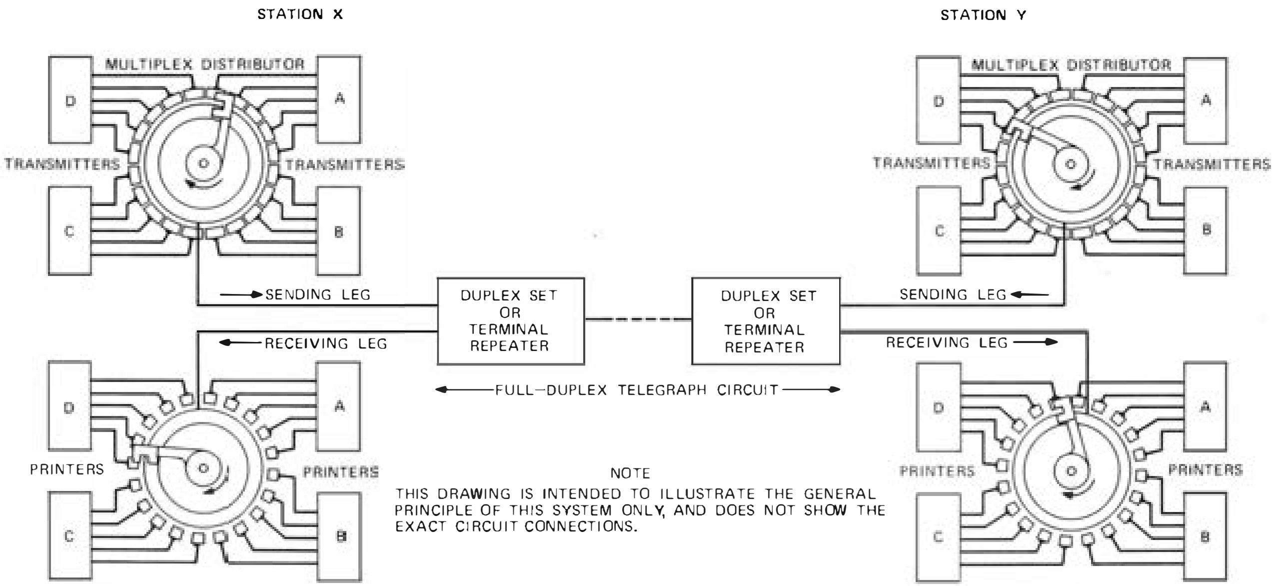


Fig. 7-15. Simplified diagram of multiplex system. (Redrawn from *Notes on Telegraph Engineering*, AT&TCo 1929, Fig. 3)

The American Company and the Western Electric Company established a trial of this equipment, between New York and Chicago, which was discontinued in 1918. In 1920, however, regular service was established, each company using one channel for its own business.

Increased operating experience indicated areas of possible improvement and these were taken care of in 1922. A cost study, made in 1923, showed that the multiplex system provided the most economical means of transmitting a large amount of traffic.

Beginning in 1926, a review of this system brought out that there was some difficulty with maintaining synchronism, maintenance was greater than desired, and there was a lack of flexibility in adapting the system to network service. Consideration was given to simplifications in the forking arrangement, the possible application of start-stop techniques, and other improvements. To achieve the desired improvements would have required a major development with a high cost in money and manpower. In view of the lack of definitive demand for a radically new system, no further active work was undertaken. A factor in this decision was that the rapid development of the telephone plant created an abundance of telegraph facilities, and since the speeds were comparable there was no real pressure for a new multiplex system. In Europe, on the other hand, there was a greater development of multiplexing because of the more limited telephone plant and the need to use available facilities more efficiently.

9.4 Non-Bell Designs of Start-Stop Printers

As mentioned earlier, AT&TCo and Western Electric obtained information on the characteristics of printing-telegraph products being offered by various suppliers prior to the time that Western Electric undertook the design of such products. Although no printers being offered were suitable, the surveys uncovered the fact that a number of competing manufacturers, both in the United States and Europe, were undertaking design programs pointing toward later production. The suppliers were familiar with the demands of telegraph companies and others who utilized private-line circuits for Morse-telegraph-type communications.

Later, around 1915, a number of manufacturers, notably the Cardwell, Kleinschmidt, and Morkrum companies in the United States and Creed in Europe, began selling their printing machines and associated apparatus to Western Union and other telegraph companies. The Bell System became involved because a growing number of its private-line customers became interested in the operating and service features provided in the outside suppliers' machines. Models of the earlier designs of the Western Electric all-purpose page printer had not been supplied to many of these customers and much work remained to be done to prepare for standardization and quantity production.

In line with the established policy that the Bell System wanted the best machines available at the lowest cost commensurate with meeting required standards of performance, some of the outside suppliers' products were examined. In the interval 1916–1919, Cardwell printers were studied and tested but found to be unsuitable. A printer design originated by Potts was not sufficiently developed to be serviceable but it contained a number of fundamentally good features of value to the Bell System. The patents covering the features were purchased from Potts. Also during this period, printers supplied by the Kleinschmidt Electric Company of Long Island City were tested. Performance was considered to be unsatisfactory compared to the Western Electric all-purpose page printer.

During 1923, the Creed organization, which had come to the fore in England, submitted their printing equipment for possible application to press association services furnished by Western Union and the Bell System. Completion of an investigation covering a two-year period led to the conclusion that the all-purpose Western Electric page-printing equipment, which had at that time been standardized, was more suitable for press association services than the Creed equipment.

Meanwhile, the Morkrum Company of Chicago offered, at least in model form, a rather comprehensive line of start-stop page and tape printers and associated apparatus that contained features of considerable interest to the Bell System. As a prelude to describing the relationships between the Morkrum Company and the Bell System that resulted from this review of Morkrum's product line, a brief resume will be made of its early history.

During 1901, Joy Morton, the wealthy owner of the Morton Salt Company, became interested in telegraph-printer technology and its potentialities as a new business. He engaged Ray Perne to convert typewriter mechanisms into printing-telegraph mechanisms. C. L. Krum soon replaced Ray Perne and built a model of a page printer around the Blickensderfer electric typewriter. Later, during 1901, a type-bar printer was developed, based on the Oliver typewriter. At about that time Krum's son Howard entered the field. He designed a printer using a moving type-wheel and a stationary paper-carriage. A five-unit selecting code, which is still in use, was employed. A considerable number of these machines were placed in commercial use. Encouraged by that and other early success it was decided to form the Morkrum Company, the name of which was a contraction of those of Joy Morton, its financial backer, and Howard Krum, who by that time had proved to be an expert machine designer and highly skilled engineer. Over the next few years, under the able guidance of Howard Krum, the Morkrum Company became a recognized leader in development and production of telegraph printers and associated mechanisms.

During 1923, patent and service agreements were reached between AT&TCo and Western Electric on the one hand and the Morkrum Company on the other.[3] The agreements included a number of terms of overall benefit to the three companies. Morkrum operated a specialty shop capable of manufacturing teletypewriters in small quantities at a considerably lower cost than Western Electric. Western Electric had over the years developed efficient means of manufacturing and supplying telephone sets, and related station apparatus, and central-office and outside-plant products in large quantities for Bell System use and was not accustomed to the "model shop" mode of manufacture. It was agreed that Morkrum would undertake manufacture of start-stop teletypewriter products of Western Electric design and in accordance with Western Electric specifications and be the future supplier of such products to the Bell System. It was also agreed that Morkrum would undertake design of a new generation of start-stop page teletypewriter sets containing features covered by Morkrum patents that were considered to be superior to those provided in the Western Electric 12- and 13-type sets. Finally, it was agreed that the models of tape teletypewriters of Western Electric and Morkrum design would be compared leading to a decision as to which design to adapt for manufacture and supply to the Bell System.

Detailed design of the next generation of page teletypewriter sets was left in the hands of the Morkrum engineers with the understanding that AT&TCo Headquarters engineers would specify service features and requirements. Western Electric Headquarters engineers were delegated the responsibility of insuring that all requirements, particularly those relating to convenience in operating, maintenance, and reliability, were adequately met.

Immediately following the agreements, Morkrum undertook production of the 13-type sets which had at that time reached the stage

[3] Following agreements reached with the Bell System during 1923, the Morkrum Company consolidated its operations with those of the Kleinschmidt Electric Company. The consolidated firm adopted the name *Teletype®* Corporation (during 1929 the *Teletype* Corporation was acquired by the Bell System). It also began to use the term "teletypewriter" as a designation on printers sold to the Bell System since the Bell System had adopted the term as a standard for use in defining the class of service which it offered, using its start-stop printer sets as instruments on customer premises. Therefore, throughout the remainder of this chapter, the term "teletypewriter service" will be used in place of "printing-telegraph service" and the term "teletypewriter" (TTY) will be used in place of "printer."

of manufacture by Western Electric to meet Bell System demand. Thirteen of these sets were produced during 1924 and placed in service on a trial basis. The results being satisfactory, Morkrum undertook production of additional sets to fill the original order for 200.

Morkrum engineers also began development work on the next generation of start-stop teletypewriter sets to replace the 12- and 13-type sets. After several model stages and tool-made samples which were thoroughly tested by Western Electric engineers and placed on trial in service the sets were standardized. The No. 15 teletypewriter, like the 12-type set, provided for keyboard-sending and receiving or receiving-only services. The No. 19 teletypewriter set included the No. 15 teletypewriter and other auxiliary operating units to provide either automatic-sending (perforated tape) service or keyboard-sending service as well as receiving service.

The No. 15 teletypewriter used a typewriter ribbon for inking, had a moving basket of type, employed a paper roll inside the machine cover, and made very satisfactory carbon copies with various types of paper. It was not subject to earlier problems of paper feed, inking, and partial printing of adjacent letters. Other features included sprocket feed for commercial forms with perforated edges, thereby giving perfect registration for all copies. Tabulating features were also provided.

The No. 15 teletypewriter is shown in Fig. 7-16 as used when the service was limited to keyboard-sending and receiving. The No. 19 set including the No. 15 teletypewriter and auxiliary apparatus required to provide for either automatic or keyboard sending is shown in Fig. 7-17.

During 1923–1924, thorough comparisons were made of the models of the Western Electric and Morkrum *tape-type teletypewriters*. During 1924 it was decided to make the Morkrum teletypewriter standard even though its performance was not quite as good as the Western Electric model. (The higher cost of the latter did not justify its selection.) The Morkrum model, coded the No. 14 teletypewriter, was standardized during 1927. It is shown in Fig. 7-18.

In summary, during the 1910–1930 period, the Bell System succeeded in making a family of reliable and convenient teletypewriters and associated apparatus available for use by its customers within their leased private-line networks. During this period notable improvements were also made to enable the Murray form of five-unit Baudot code signals used to define the various letters, numerals, and controls of the teletypewriter language to be transmitted satisfactorily over the private-line networks. Toward the end of the period reliable transmission

Fig. 7-16. The No. 15 teletypewriter.

Fig. 7-17. The No. 19 teletypewriter set. (Watson 1938, Fig. 10)

Fig. 7-18. The No. 14 teletypewriter. (Watson 1938, Fig. 8)

of such signals at the rate of 60 words per minute became an accomplished fact.

9.5 Transmission Improvements

Early in the twentieth century, the management of the Bell System decided to undertake programs that would lead to conversion of the then-existing manual Morse telegraph service to the more convenient automated forms that would be possible as soon as suitable teletypewriter mechanisms were made available for customer use. At the time of decision the wire-line networks that were used to provide the manual Morse service required adjustments to maintain adequate quality of transmission of the telegraph signals. It was quite evident that substantial improvements would need to be made to meet the exacting standards for transmission of automated signals originated by teletypewriter machines.

The decision proved to be the stimulus to undertake scientific investigations, engineering studies, and the development of transmission systems for introduction into the Bell System telegraph plant. In keeping with the established long-range policy of the Bell System that its transmission plant should be capable of dual use for both analog (voice) and digital (teletypewriter) communications, the telegraph engineers and scientists coordinated their activities with those of their telephone counterparts.

Sections 9.5.1 through 9.5.6 to follow highlight the important advances that were made and applied to the Bell System transmission plant, and at its operating and maintenance centers, to make the plant suitable for handling communications between teletypewriter machines. At the end of the period reliable teletypewriter service at a speed of 60 words per minute was available throughout the Bell System.

9.5.1 Status of Wire-Line Plant in 1905

The growing demand for telephone and telegraph services during the early period of Bell System history showed up in requests for circuits extending over longer and longer distances and also in greater concentration of services within urban areas. This latter problem of many open-wire circuits in a limited area produced something like a "copper-wire haystack" which was difficult to service and maintain. To overcome these handicaps, lead-covered cable was developed. Initially, cable was employed where distances were relatively short and, to minimize its size, fine-gauge copper wires were used. Cable was introduced into the local plant about 1888. By 1904 it had largely replaced open wire plant in urban areas.

Cables were not affected by changes in weather to more than a minor degree as compared to open wire lines. This was an advantage, hence there was a natural tendency to extend the distance over which cables might be used. An aerial cable was in operation between Boston and Worcester during 1904. Later, in 1906, an underground cable was placed in service between New York and Philadelphia. These cables were used to provide voice-telephone and grounded-telegraph transmission circuits, with less than satisfactory results. It became evident that technical effort was needed on several fronts to overcome the obstacles that limited the length of lead-covered, loaded, fine-gauge cable. It was not until many years later that the problems were considered to be under sufficient control to justify application of the cable wherever needed within the Bell System.

Open wire lines were the sole means of providing voice-telephone and telegraph circuits over considerable distance. When used for manual Morse telegraph service single-line repeaters were inserted in the circuit at intermediate points along the right of way. Close attention by skilled maintenance forces was required in order to make the service acceptable. The low resistance of the open-wire-line conductors provided an advantage for transmission of both voice-telephone and Morse currents as compared to the fine-gauge wires within cable. Under favorable weather conditions, open wire circuits provided relatively good transmission of voice currents for several hundred miles. But due to the instability of the telephone circuits there were frequent periods when they were considered unsuitable for commercial use.

9.5.2 Scientific Advances and Inventions

The limitations of the transmission plant within the Bell System during the 1905 period are highlighted briefly above. Evidently this knowledge served the purpose of proving once again that "necessity is the mother of invention." Several notable advances in transmission technology by J. R. Carson and others came to the fore during the several years that followed. History reveals that these were to have a profound impact on the composition of the Bell System transmission plant that came into use during the 1915–1925 period. There were many contributions by many contributors but, at the risk of omitting some worthy of mention, the following have been chosen as of prime importance.

The audion was invented by Lee de Forest and patented during 1907. In 1912, H. D. Arnold undertook research and device development work in the Engineering Department of the Western Electric Company. Arnold determined the methodology that enabled reliable, long-life vacuum tubes to be manufactured. Western Electric established a "tube shop" soon thereafter to produce high-quality vacuum tubes and over the next 40-plus years served as a supplier to the Bell System. Over this period vacuum tubes came into use in a variety of transmission systems. In fact, the vacuum tube became as versatile a device in such systems as the relay in switching systems. Neither was displaced by solid-state devices until well into the 1950s.

The first commercial use of vacuum tubes within the Bell System was in the repeaters connected into the transcontinental open-wire line during 1915. An extensive account of the application of electronic amplifiers to early transmission systems is given in Chapter 4, Section 4.2; the story of the underlying research is presented in Chapter 10.

Concurrent improvements in a variety of materials, notably magnetic alloys, facilitated a major reduction in the size and cost of apparatus as well as in the power required to operate line repeaters. These developments are discussed in Chapter 8. Improved repeaters for insertion in 2-wire open-wire lines and 2- and 4-wire circuits within fine-gauge cable were designed and made available. The 2-wire repeater, known as the

22-type, was a two-stage amplifier containing the 40,000-hour-life-expectancy 101- and 102-type vacuum tubes. The 4-wire repeater was composed of apparatus used in a pair of 2-wire repeaters but connected to provide amplification in each of the two directions of transmission. The 2- and 4-wire repeaters came into extensive use throughout the long-distance wire plant in the Bell System. Their use in cables marked the beginning of the long-delayed program of extending lead-covered, fine-gauge, loaded cable over great distance. During 1925 such a cable was placed in operation between New York and Chicago.

By 1910 the application of Fourier analysis revealed that digital telegraph signals passed over the dc grounded circuit desired by the compositing method contained frequencies of up to about 60 hertz. The analysis also revealed that speech currents occupied a higher band which was subsequently established as 300 to 3,200 hertz for good-quality telephone transmission. These theoretical studies together with information gained from measuring the loss characteristics of the composite set and of the typical telephone line provided information for detailed studies of the characteristics of wire lines and transmission networks.

As described in Chapter 4, George Campbell invented the ladder-type filter during 1915. The filter concept provided an economical means of subdividing available frequency spectrum over a wire line in an efficient manner to provide dual use of the line for both telephone and telegraph communications.

The Campbell invention unleashed a great deal of activity in the design of filter and other networks including those used in combination with vacuum tubes to provide oscillators, modulators, demodulators, and rectifiers. Thus the basic technology and building-block units for composition of transmission systems using the frequency-division method of multiplexing became available.

During this period Harry Nyquist of AT&TCo Headquarters Engineering undertook a series of comprehensive theoretical studies relating to the behavior of digital and analog signals within typical closely coupled electrical circuits where the phenomenon of mutual signal interference needed to be fully understood. He reported on his findings in articles that appeared in *The Bell System Technical Journal* and in other technical publications. The work of Harry Nyquist has since been regarded as a major contribution to the advancement of transmission technology.[4] Application of this knowledge is reviewed in Section 9.5.6.

[4] Journal articles describing advances in transmission theory by J. R. Carson, Harry Nyquist, and others are listed in "References" at the back of this volume.

9.5.3 *The Metallic Telegraph System*

The metallic telegraph system was developed to meet the need for a stable direct-current telegraph system with suitable quality of transmission to permit it to be used in handling teletypewriter communications over long, fine-gauge, loaded cables equipped with repeaters. It was designed to operate on a 4-wire basis using two direct-current metallic circuits derived by compositing two pairs of wires in a fine-gauge cable. It could also be used with spare phantom circuits or two spare pairs when available.

The system was comprised of two types of telegraph repeaters. The terminal type was used to associate the 4-wire line circuit with the loop (half duplex) or loops (full duplex) extended to the customer's station over a pair, or two pairs, of wires in the local cable. The through type was used to repeat teletypewriter signals between line sections. Through repeaters were spaced along the line at about 100-mile intervals. The telegraph repeaters were located at telephone repeater points in order that the number of arrangements for bypassing the telegraph circuits around telephone repeaters would be minimized. Telephone cables having repeater spacings of about 50 miles made it necessary to bypass the telegraph circuit around the telephone repeater.

The choice of metallic rather than ground-return line circuits made it possible to apply a low voltage (34 volts from a non-grounded source) to each of the two lines and to reduce the line currents to about 5 milliamperes, (±130-volt batteries were used to supply line currents of about 60 milliamperes to the grounded telegraph line circuits used with the family of telegraph repeaters described in earlier sections of this chapter). The use of metallic line circuits, and low voltages and currents, permitted reduction of interference from power circuits, effects of ground potential differences, and interference from other telegraph circuits in the same cable to very small amounts. The low voltages and currents also minimized the effects of telegraph thump and flutter on the telephone circuits in the cable. To further reduce the high-frequency components of the telegraph currents, a series-inductance-and-shunt-capacitance network, known as a "noise killer," was connected between the line-relay windings and the loop (sending) relay armature. (See Fig. 7-19.)

In addition to adoption of a metallic line circuit, the repeaters of the system contained a number of improvements as compared to the earlier family of telegraph repeaters used to furnish service over grounded lines (Fig. 7-7).[5] The most important improvement came in

[5] See "References" at the back of this volume for journal articles describing later improvements that were made in dc telegraph repeaters and the modes used to improve their operation over grounded open-wire circuits.

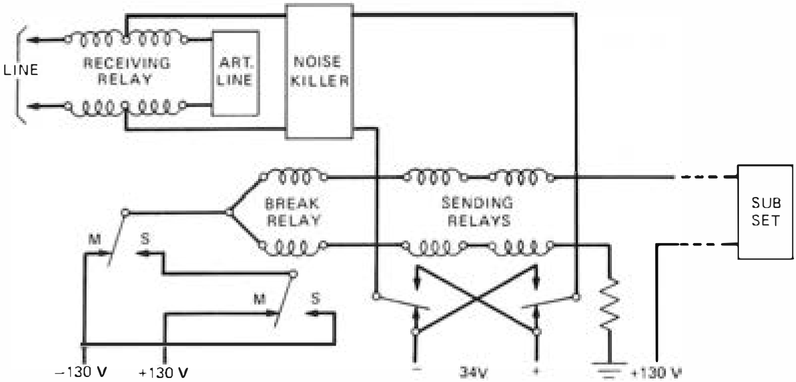


Fig. 7-19. Simplified schematic of the metallic telegraph system for cables. (Redrawn from *Notes on Telegraph Engineering*, AT&TCo 1929, Fig. 14)

the form of a sensitive line relay, coded 209FA, capable of operating on very low values of current. The 209FA relay contains four equally balanced main windings and two auxiliary windings. The latter were connected across the contacts of the line relay thus providing energy induced into the auxiliary windings from the main windings of the relay to increase its sensitivity. As a by-product of the apparatus development work that was carried out to produce the 209FA polar relay, a less-sensitive type of polar relay, which was coded 215A, was made available for use in the loop circuit of the terminal repeater. Figure 7-20 shows the 209FA and 215A relays.

The basic circuit arrangements of a terminal repeater for half-duplex service are shown in schematic form on Fig. 7-19. Note that two 215A sending relays are required to provide a means of reversing the voltage applied to the line. Also note that a 215A break relay is provided to permit an operator at the local station to readily stop an operator at a distant station on the repeatered line circuit from sending and thereby gain control of the circuit.

The metallic telegraph system was first applied on a trial basis on a cable circuit between Kingston, New York, and New York City. The first standard systems came into use between New York and Philadelphia during 1919. Many systems were subsequently installed throughout the Bell System where cable facilities became available and demand for service became sufficient to justify their installation.

9.5.4 The Open-Wire Carrier Telegraph System

The metallic telegraph system did indeed provide high-quality telegraph circuits for teletypewriter communications over routes where

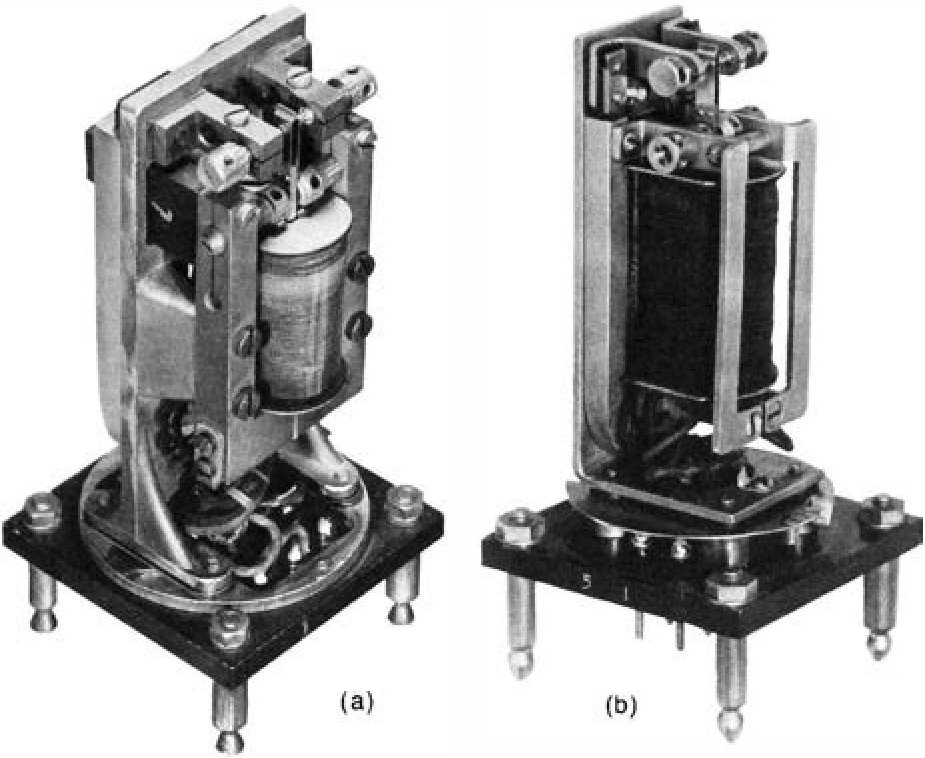


Fig. 7-20. Polar telegraph relays: (a) Code 209FA; (b) Code 215A.

fine-gauge, repeatered cable circuits were available for voice-telephone use. However, extension of these long cable circuits outside of the northeastern area of the United States was several years off.

It was evident that the open wire plant would have to be utilized for both telephone and telegraph purposes for an indefinite period into the future. This led to the decision to intensify technical effort on improvement in performance of the telephone and telegraph circuits provided within the open wire plant and on innovations to increase its circuit capacity. Several creditable modifications were subsequently made to the open-wire outside plant and repeaters which made a frequency spectrum available for carrier circuits extending from about 300 hertz to several kilohertz.

The planners and designers of the open-wire carrier telephone and telegraph system decided not to attempt to derive both telephone and telegraph circuits within a common system but to design two systems, one for telephone and the other for telegraph. Both systems made use of the newly developed technology of vacuum tubes and associated networks and of tuned circuits and wave filters.

The features of the open-wire carrier telephone system that first came into commercial use during 1918 are described in Section 4.2.5 of Chapter 4. The characteristics and capabilities of the open-wire carrier telegraph system are reviewed below.

Figure 7-21 shows the schematic circuit of one terminal of the system including the manner of associating it with an open wire line. The figure also shows the manner of connecting one of the available telegraph line circuits to a loop extending to a subscriber's office.

The system occupied a frequency band, above the voice-telephone circuit, in the range from 3,300 hertz to 10,000 hertz. The following table shows the frequencies used in the two directions:

Circuit Number	Frequency in Hertz	
	West to East	East to West
1	5,500	6,500
2	5,250	6,800
3	5,010	7,110
4	4,770	7,440
5	4,530	7,800
6	4,290	8,180
7	4,050	8,590
8	3,810	9,030
9	3,570	9,500
10	3,330	10,000

The 20 one-way channels were paired in the manner indicated above to provide the ten two-way circuits. The lowest and the highest frequencies were paired together. This was done to allow for emergency line conditions where, for example, the system could only be made to operate by raising the outgoing currents to the maximum and increasing the sensitivity of the receiving circuits to a maximum. The first one-way channels that were likely to fail were the 3,300-hertz channel because of interference from the telephone circuit, the 10,000-hertz channel because of line attenuation, and the 5,500-hertz and 6,500-hertz channels because of mutual interference. Under such marginal conditions only two of the ten circuits would be out of service.

Variable air capacitors were used in the oscillator and sending and receiving tuned circuits to provide means of adjusting the frequency of the oscillator and the midpoint of each tuned circuit. Means of adjusting these circuits were provided because, at the time that the system was being developed, stable band-pass filters and constant-frequency oscillators were not available.

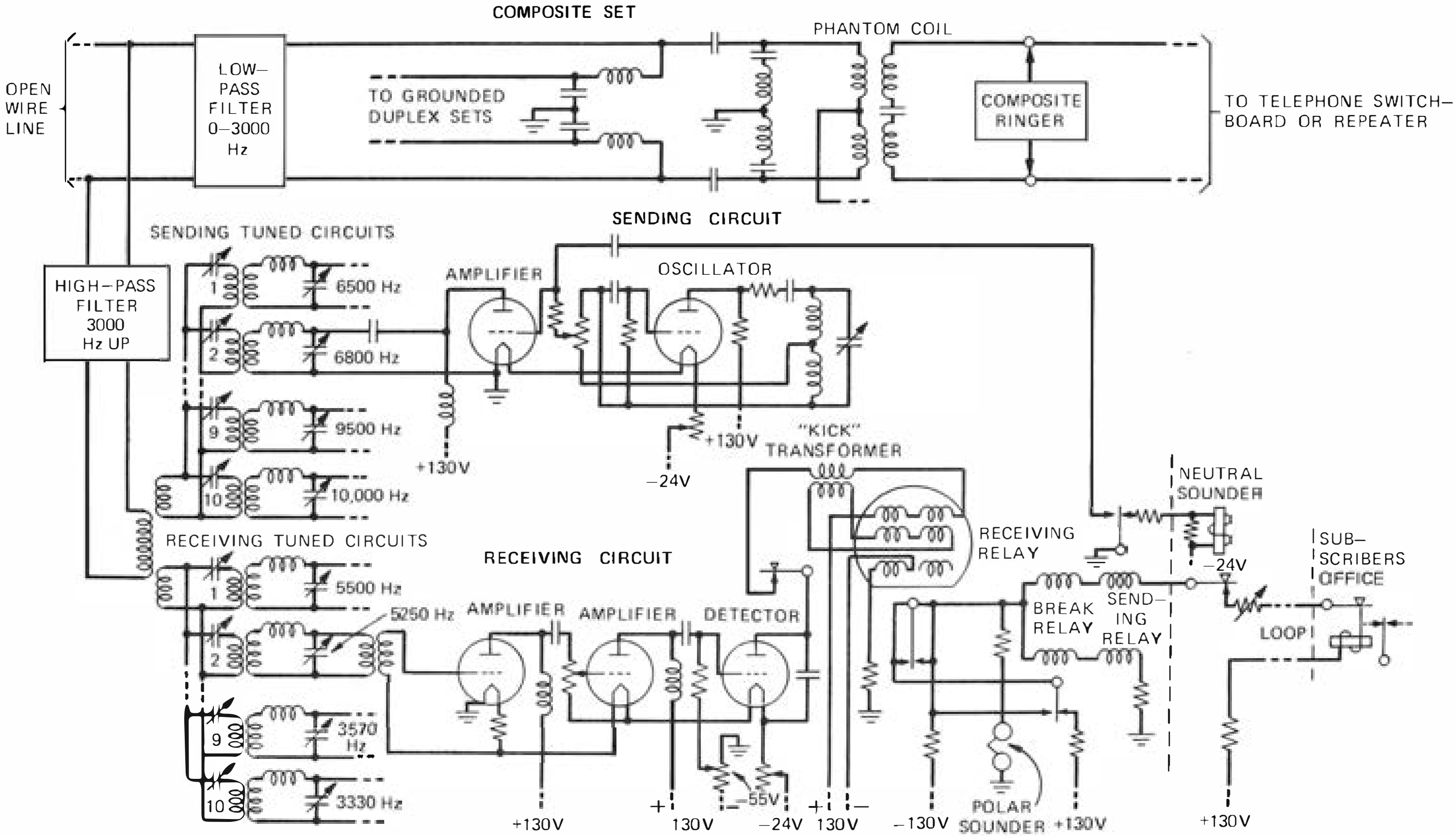


Fig. 7-21. Simplified schematic of the first carrier telegraph system. (Redrawn from Hamilton 1948, Fig. 1)

Models of the system were installed at Pittsburgh and Harrisburg, Pennsylvania, and connected to an open wire line between these two cities during 1919. Later, during 1920, the trial system was extended to Chicago. These first installations were used to check performance under field conditions and to determine the nature of changes that were needed prior to standardization.

The first standard system, which became known as the Type B open-wire carrier telegraph system, was installed at Chicago, Illinois, and Oakland, California, and connected to an open wire line containing repeaters spaced at intervals of 250 to 300 miles. After cutover its capabilities were verified by looping the ten circuit terminals at Chicago and Oakland together thereby setting up a circuit whose length of about 30,000 miles was the longest that had ever been tested up to that time.

During the following period of many years Type B systems were added to the open wire plant throughout the Bell System to provide needed transmission circuits suitable for teletypewriter communications. They were replaced gradually by the voice-frequency carrier telegraph system for cables over routes where the cables replaced the open-wire-line facilities.

9.5.5 The Voice-Frequency Carrier Telegraph System for Cables

As stated earlier, the metallic telegraph system, which became available during 1919, provided a superior transmission circuit for teletypewriter use as compared to the earlier dc grounded telegraph systems over routes where repeatered, fine-gauge, loaded cables were in use. However, the number of circuits that could be derived within the cable by compositing means for use by the metallic system was reduced by a factor of 4/1 as compared to the dc grounded systems. Although it was possible to use spare pairs within the cable for metallic telegraph rather than voice telephone, their use was severely limited because it was uneconomical to provide a high percentage of spares within any cable.

Voice-telephone circuits within the cables were operated on either a 2-wire or 4-wire basis depending upon the plant situation at the time. But the superior quality of transmission over the 4-wire circuits was convincing proof that 2-wire operation should gradually be eliminated, particularly over the longer cable routes.

The bandwidth for transmission of voice-telephone currents over the long 4-wire repeatered cable circuits was well beyond that required for speech (300–3,200 hertz). However, the bandwidth that could be made available for telegraph use was considered to be insufficient to justify degradation of the high-quality voice circuits within the 4-wire cables.

Comprehensive engineering studies which included all of the above factors and others led to the conclusion that a voice-frequency carrier telegraph system should be designed for 4-wire operation over the 16- and 19-gauge medium-heavy-loaded repeatered cable circuits. These cables had been adopted as the most suitable for additions to the long-distance cable plant.

Figure 7-22 shows one terminal of the voice-frequency carrier telegraph system in schematic form including the manner of connecting it to the 4-wire cable circuit and to one of the local cable circuits extending to the premises of a subscriber of the service. The schematic discloses the composition of the finalized design of the system, the first installation of which was made in New York and Pittsburgh during 1923. A 4-wire line circuit within the 19-gauge medium-heavy-loaded cable with repeaters located at 50-mile intervals was used to interconnect the two terminals. Figure 7-23 is a photograph of one terminal of the system.

It is of interest that the voice-frequency carrier telegraph system came into use four years later than the open-wire carrier telegraph system. During that interval there were several advances in technology which enabled designers of the voice-frequency system to choose components having a high degree of stability that were not available for use in the open-wire carrier system. Included were the band-pass channel selection filters and the multifrequency generator of carrier frequencies.

The multifrequency generator provided a source of 12 carrier frequencies in 170-hertz steps. All of these frequencies were generated within a machine which consisted of 12 rotors on a single drive shaft. The reason for this design was that the carrier frequencies could not drift with respect to each other. However, a drift in speed of the drive shaft affected all channels adversely. This had serious implications because the band-pass filters were designed to allow a space of 170 hertz between channels with the midband of each filter occurring at the carrier frequency supplied by the generator. It was necessary to maintain this frequency within close limits to minimize interference between channels due to modulation in the repeaters along the cable circuit. The carrier frequencies were chosen at odd multiples of 85 hertz, this value being half of the frequency separation (170 hertz) between the channels. By this scheme the sum or difference of any two carrier frequencies lay just midway between carrier frequencies and the chance for interference from these components was minimized by the band-pass filters.

Drift in speed of the generator of the carrier frequencies was maintained within limits of ± 1 percent. But the centrifugal governor on the shaft of the dc motor which drove the generator required frequent adjustments to maintain the speed within these limits. Once

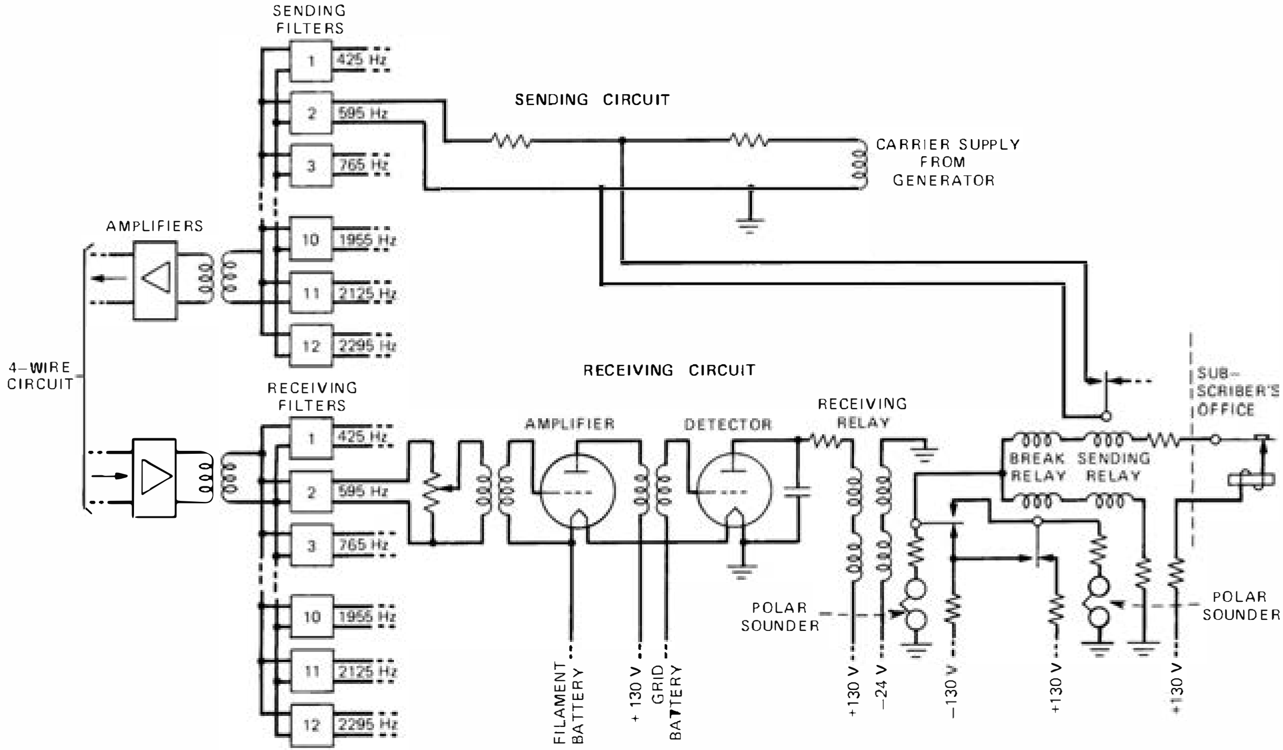


Fig. 7-22. Simplified schematic of the 12-channel voice-frequency carrier telegraph system that replaced the open-wire carrier system shown in Fig. 7-21. (Redrawn from Hamilton 1948, Fig. 2)

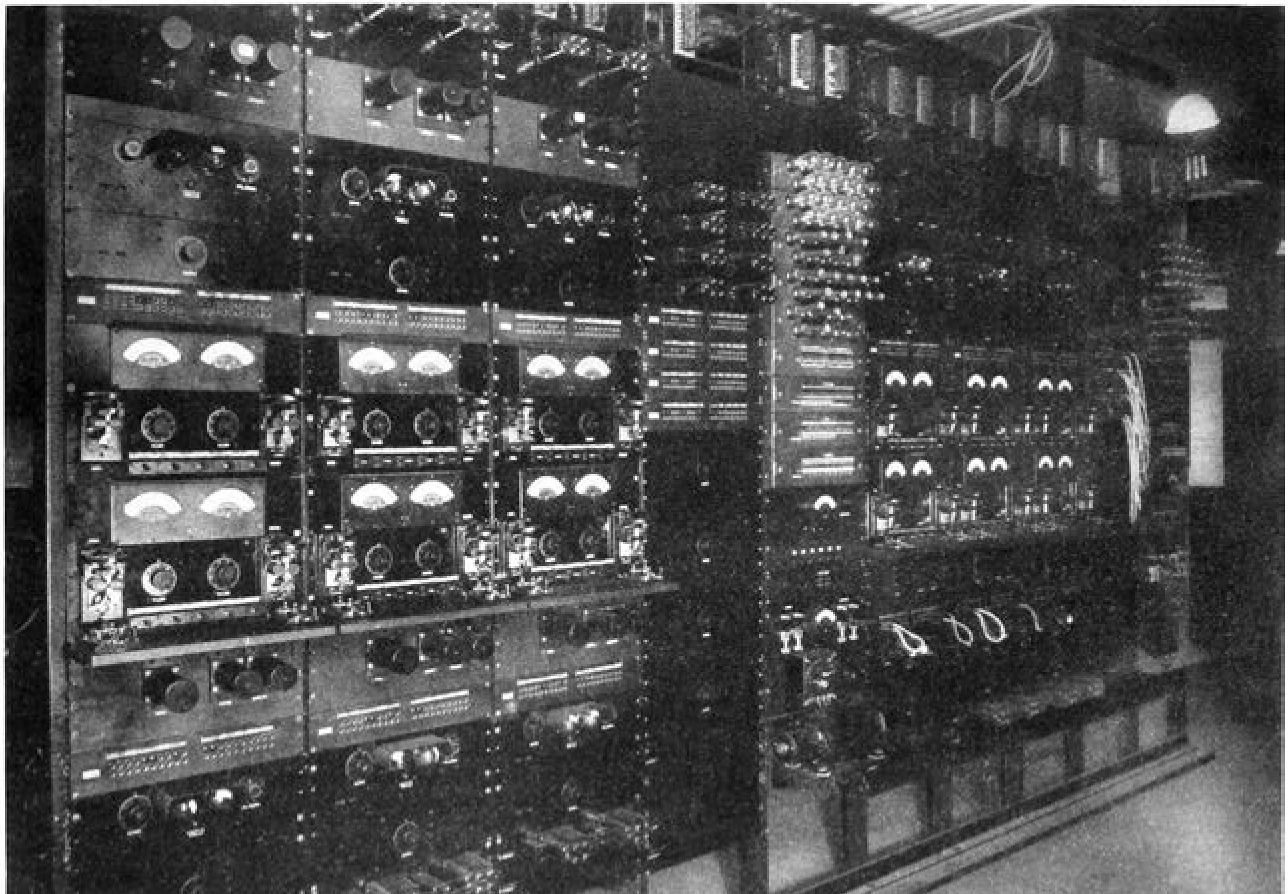


Fig. 7-23. Terminal of a 12-channel voice-frequency carrier telegraph system for cables.

again, advances in technology were available. The stability of vacuum tube oscillators had been greatly improved and circuitry had been developed for locking oscillators in step with harmonic frequencies from a control oscillator. So within a few years the motor-generators were replaced by vacuum tube oscillators.

During the period when the voice-frequency carrier system was being designed, the technology of filter design had reached the stage where ladder-type filters having very stable characteristics could be produced. Inductors, capacitors, and resistors for use in composition of the filters and having very close tolerances were being manufactured and assembled in encapsulated containers to insure a high degree of stability of the filters in service environments.

The 102- and 101-type (40,000-hour life expectancy) vacuum tubes were adopted for use within the receiving circuit of each channel of the carrier system. These tubes had at that time come into extensive use within the 22-type and 44-type repeaters where their performance in service was considered to be of the highest order.

The 209FA and 215A polar relays used in the balanced loop circuits of the system had been in use in similar circuits of the metallic and open-wire carrier telegraph systems. Experience with these systems showed that the relays were capable of operating for many hours of continuous service and of transmitting many millions of digital pulses before requiring cleaning and burnishing. The relays were plug-mounted to enable removal for test and adjustment at a relay test table.

Thus the configuration and content of the voice-frequency carrier telegraph system was a blending of knowledge drawn from several technical disciplines and combined suitably to meet the applications in service for which it was intended. History tells us that the many systems that were placed in service over the following decades continued to meet the exacting service requirements and required a very low degree of maintenance to keep them in operation.

9.5.6 Distortion and Measurement of Digital Signals

The telegraphic mode of communicating over distances makes use of groups of discrete digital pulses arranged in suitable time sequence and duration to enable trains of such pulses to define suitably each letter, figure, or other symbol within the message being transmitted.

Prior to 1910 the private-line telegraph services offered by the Bell System were operated on a manual Morse basis. The dot and dash signals used to define each character were sent in time sequence at speeds up to the equivalent of about 18 dots per second, depending upon the quality of the transmission medium, the capabilities of the telegraph sending device, and the skill of the sending and receiving operators. The time separation between dot and dash pulses within a character was usually equal to the dot length and the dash pulse was normally three times the length of the dot.

The sending keys and receiving sounders within each manual Morse circuit provided the means to enable skilled operators to make adjustments to meet their individual idiosyncrasies so as to allow for adequate identification of the various coded characters. Prior to turning over a private-line circuit to customer use in service or after receiving a customer complaint the test room attendants "lined up" the relatively unstable circuit. The line-up procedure usually consisted of measurement of the ability of the circuit to transmit a succession of dot pulses originated by an automatic vibrating-key sender or equivalent device. Circuit response was observed at various points by means of a zero-center meter. If the meter reading deviated from zero, the signals were considered to be biased to "light" or "heavy" depending upon the direction of the deviation. To compensate for this bias in received signals the attendants at the various repeater points along the way made adjustments of the repeaters. Other than this elementary mode of measuring bias of repeated dot signals no distortion measuring had been considered necessary.

At the time of acquisition of the Western Union Telegraph Company, Bell System management decided to modernize its telegraph business by providing its customers with convenient teletypewriter instruments in place of Morse telegraph sets. This decision provided the impetus for a systematic investigation of the principles of digital transmission over the wire-line media comprising Bell System plant and of the mode of generating and receiving digital signals by teletypewriters and regenerative repeaters. Considerable knowledge was needed to determine the causes of distortion of such signals at their source, during transmission over various media, and at their destination. It was obvious that without this knowledge it would have been difficult, if not impossible, to design suitable signal generators and receivers and to incorporate features in transmission media to enable each network to provide reliable service at signaling speeds far in excess of the 18 dots per second that had up to that time been possible with the manual Morse mode of communicating.

During the period covering development and application of the start-stop teletypewriter, the multiplex telegraph system, and the metallic and carrier telegraph systems, Harry Nyquist and his transmission-engineering associates in AT&TCo Headquarters Engineering conducted theoretical studies, made laboratory experiments and measurements in Bell System plant, and designed telegraph distortion-measuring and relay-test sets for use by plant engineers and maintenance forces.

The three types of distortion which affect the length of a signal pulse were defined in 1928 by Nyquist, Shanck, and Cory. They were called (*i*) *bias*, which is a uniform lengthening of the marks (called positive bias) or a uniform lengthening of the spaces (called negative bias); (*ii*) *characteristic bias*, which is a function of the signal combination as well as the electrical and mechanical characteristics of the particular circuit under consideration; and (*iii*) *fortuitous distortion*, which is the random superposition of extraneous interfering currents upon the signal pulse. The principal factors causing these distortions are as follows:

Bias	—variations in circuit resistance or loss
	—variations in battery voltage at the terminals, or variations in the oscillator voltage
	—high-resistance sending-relay contacts
	—non-symmetrical relay adjustments.
Characteristic Bias	—characteristics of relay windings, coils, filters and other equipment
	—wave shaping of detectors.
Fortuitous Distortion	—noise
	—lightning
	—ground-potential variations
	—functional switching operations
	—relay-contact troubles
	—variation in repeater gain with load.

As a general rule, distortion is not allowed to exceed 35 percent in teletypewriter operation. If distortion were to reach 55 percent then every symbol would be in error and total failure would result.

The effects of distortion on Morse telegraph are more subtle. Tests reported in 1929 by J. Herman noted that telegraph operators described circuits with bad distortion as "heavy," "light," "hangs," or "no good," and that the different types of distortion had different results. Judged by the opinions of the operators, negative bias became objectionable at 30 percent, positive bias at 40 percent, and all other types of distortion at about 50 percent. On the other hand, judged by the accuracy of the operators' performance all distortion, with the possible exception of positive bias, only became serious at values above 50 percent, ranging up to 85 percent for fortuitous distortion.

The design of a satisfactory teletypewriter network depended on a knowledge of the distortion in each circuit component. The distortion-measuring set was used to determine the total distortion, as well as its components. It will be clear that a signal composed of ideal alternate marks and spaces of equal duration would have zero-percent distortion. A lengthening of the marks (and corresponding shortening of the spaces) would read positive bias, and a shortening of the marks would read negative bias; hence a center-reading meter can be used for measuring bias. In an actual signal, the indication would jump about because of dependence on the transmitted-signal transitions from mark to space. The total-distortion-measuring set was designed to give a relatively steady reading, but variations with time would be expected.

The term "telegraph coefficient" was adopted for use in circuit layout. It was related to total distortion and the probabilities inherent in the causes. The limit of momentary total distortion of 35 percent or less in teletypewriter circuits was based on the normal statistical relation which states that a circuit whose measured distortion had a root-mean-square value of 8 percent would nonetheless reach or exceed 35 percent, and thereby incorrectly print a character 1 time in 44,000. This was the same as saying that there would be one character error in two hours of 60-word-per-minute teletypewriter service, on the average. This might have seemed to be a high accuracy requirement for simple messages but it reflected the Bell System objective of providing the best service that is economically sensible to customers who require that kind of accuracy, such as financial institutions. It could have been made more or less accurate at a corresponding economic offset.

A circuit whose root-mean-square total distortion was 8 percent, i.e., a mean-square value of distortion of 64, was arbitrarily assigned a coefficient of 10. Therefore, a circuit having a mean-square distortion of 6.4 had a coefficient of 1.0. Coefficients assigned to the different segments were added arithmetically to get an overall coefficient which could not be more than 10.0. Mathematically, this statement is equivalent to saying that the distortions contributed by the different segments of the overall circuit accumulated in random fashion.

Short subscriber lines of less than 5 miles were considered to contribute nothing to the overall distortion. Loops of up to about 35 miles had coefficients of 1.0 to 1.5, but those of up to 50 to 60 miles had coefficients of 3.5 to 4.0. For longer circuits, used to provide 60-speed TTY service, typical values were as follows:

	Coeff/Sect	Max Length/Sect
DC grounded on open wire	2.5–4.0	300 miles
DC metallic on cable	2.0–3.0	150 miles
High-freq carrier on OW	2.6	1,150 miles
Voice-freq on cable or OW	2.0–2.2	3,500 miles

If it was required to set up a circuit such that the coefficients of the segments added up to more than a total of 10, it was necessary to insert a regenerative repeater in such a position that the sum of the distortions on either side would be less than 10. The repeater would then wipe out the distortion in the preceding section. Figure 7-24 shows a circuit layout in which the coefficients add up to 13.2 between the West and East stations. This exceeds the objective coefficient of 10, and so a repeater is inserted at a point such that the coefficient is 7.0 on the West side and 6.2 on the East side. This arrangement would be expected to give better performance than that stated above for a coefficient of 10.

9.6 Special-Purpose Teletypewriter Private-Line Switching Systems

During 1916, when the Bell System began leasing teletypewriter service to press associations, a teletypewriter switching system was developed for use by the United Press Association. This system interconnected six newspaper offices in greater New York City with the central transmitting station in the World Building. By means of it the central bureau could connect one or more of these six lines to the sending instrument and send information to one or more of the newspaper offices. The teletypewriter circuits to each of the newspaper offices operated on a simplex basis. During 1917 this system was modified and enlarged to include eight newspaper offices and at that time a telephone circuit was provided between each newspaper office and the central bureau. By means of a hand generator, the attendant at the newspaper office could signal the central bureau. The attendant at the newspaper office would then report to the attendant at the central bureau by telephone the trouble or other information deemed necessary regarding the teletypewriter or teletypewriter circuit.

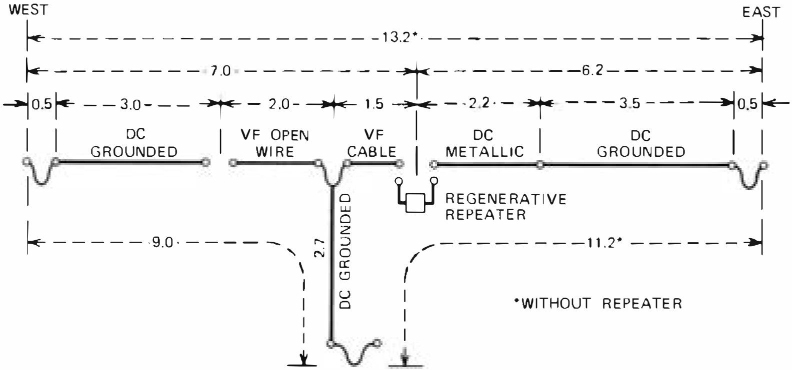


Fig. 7-24. Circuit layout illustrating telegraph coefficients.

Another one of the earliest teletypewriter switching systems developed and used in Bell System subscriber networks was one provided for the Otis Elevator Company. In December 1916, a three-station network was installed and arranged to permit the attendant at the New York City station, by means of switching keys, to connect his teletypewriter to either the line to Harrison, New Jersey, or to the line to Yonkers, New York. Two-way communication was provided for in either case. Later, in September 1917, a more elaborate switching arrangement permitted any one of the stations to signal and become connected to either of the other two for two-way teletypewriter communication.

These two examples of teletypewriter switching applications are believed to be of interest as they illustrate that both press associations and commercial customers demanded and required the development of teletypewriter switching arrangements as soon as the Bell System started furnishing teletypewriter service to private-line customers.

Following the development and application of these early types of teletypewriter private-line switching systems there was a continuous demand on the part of the AT&TCo Long Lines Department and practically all of the Associated Companies of the Bell System for the development of special teletypewriter switching systems to meet specific requirements of particular private-line installations of customers to whom they had leased the service. Systems were developed for use by banks, brokers, hotels, a large variety of manufacturing concerns, merchandising organizations, credit bureaus, press associations, public utilities, state and municipal police systems, and government agencies such as the Bureau of Aeronautics of the Department of Commerce. These systems were designed for use in local networks, for use over toll wire-telegraph lines, and in networks involving both. They also included both manual and automatic systems as in certain cases the demands of the customer were such that a manual system seemed proper whereas in other cases automatic arrangements were provided to meet his requirements.

With the commercial introduction of the more compact tape- and page-type teletypewriters, beginning around 1920, commercial private-line teletypewriter service had a large growth and naturally with this growth there was a considerable expansion in the number and varieties of switching arrangements. This was naturally to be expected because the modern teletypewriter became a business tool that had decided advantages over the telephone and Morse telegraph instrument. Probably the most important of these was its ability to provide an accurate written record at both the sending and receiving ends of the circuit and its ability to record a message at the receiving end even though the party was not present at the time. New service features not heretofore provided in telephone networks were immediately apparent to the customer and were developed and incorporated in the switching arrangements. Such features included broadcasting to a group of stations and remote control of power to permit a teletypewriter motor to be started and stopped remotely for unattended service. Other features, which had also been used in telephone networks, were conference connections and selective calling by which a number of stations could be associated in the same network and arranged so that if the network were idle any one station could automatically select one or more other stations to the exclusion of the rest of the network.

We will not describe all of the different special-purpose types of private-line teletypewriter switching systems that were developed for use by Bell System subscribers. A few of these special-purpose systems, which were not necessarily the first of their kind but are considered of particular interest, however, have been selected for description to illustrate some of the development work that was done in this field.

9.6.1 System Developed for Use by New York City Police

This system was developed for New York Bell Telephone Company to meet the specific needs of the New York City police. It was cut into service during 1928, and included five switchboards and over 100 teletypewriters.

The city of New York is divided into five boroughs—Manhattan, Brooklyn, Bronx, Queens, and Richmond. At the time, each borough had a headquarters office and the General Police Headquarters was located in the Manhattan office. The system provided two-way teletypewriter communication between General Police Headquarters and each borough headquarters, and one-way communication from each borough headquarters to associated stations.

At General Police Headquarters there were four receiving-only machines for handling messages from other headquarters, and one receiving-only machine at each of the four borough headquarters for messages from General Police Headquarters.

There was a switchboard at each borough headquarters for two-way service between Manhattan and the other four headquarters and for one-way service from each borough headquarters to the associated receiving-only stations in that respective borough. Each switchboard had associated with it two sending-receiving teletypewriters, one normally used for two-way service and the other for one-way service, with flexibility of interchange of machines or use of either for both purposes. The switchboard operators could select and send messages over any single line, to a group of lines, or to all lines associated with the switchboard. Receiving-only stations could light a lamp at the board to acknowledge receipt of a message by operating a key at the station.

9.6.2 System Developed for Use by Hudson Motor Car Company

This switching system was developed for the Michigan Bell Telephone Company for use by the Hudson Motor Car Company of Detroit to coordinate the flow of parts in its assembly lines. The system was installed in 1928. It consisted of a cordless key-type switchboard (having a capacity of 12 lines) and two teletypewriters at the control point with lines radiating to various points throughout the factory. The main features of this system were as follows:

(*i*) Means were provided for two-way communication between the control station and the branch stations or between the various branch stations.

(*ii*) Either teletypewriter at the switchboard could be associated with any line for two-way communication.

(*iii*) One teletypewriter at the switchboard could "broadcast" to one or more stations by operating a broadcast key associated with each line.

(*iv*) Two stations could be connected together for two-way communication and either switchboard teletypewriter could be connected in for monitoring.

(*v*) A station could call the switchboard by operating a key. A line lamp was lighted at the switchboard and extinguished when the operator answered the call.

(*vi*) During broadcasting, a station could signal the switchboard, lighting a line lamp by "flashing" a key at the station.

(*vii*) Lamp signals received during a broadcast could be answered by the second teletypewriter at the switchboard.

(*viii*) Disconnect on a through connection was provided by releasing a key at the station.

(*ix*) The teletypewriter motor at each station was started automatically when a connection was established to that line at the switchboard and was stopped when the connection was broken.

This system was restricted to use over comparatively short lines without repeaters. Two-wire lines were used, one wire for printing and motor control and the other wire for signaling. Neutral or "open and close" transmission was employed.

9.6.3 System Developed for Use by Pennsylvania State Police

The Pennsylvania State Police switching system is of particular interest because it was the first teletypewriter switching system to operate over such a large territory and because it included a scheme of remote control which permitted the State Headquarters at Harrisburg to automatically seize all lines at the three Zone Headquarters for a statewide broadcast.

This system was developed for the Bell Telephone Company of Pennsylvania and service was started during December 1929. The network included cordless switchboards at Harrisburg, Wyoming, Philadelphia, and Pittsburgh, and over 100 receiving-only stations scattered throughout the state.

Each of the three zone switchboards and the Harrisburg board were arranged to permit sending to one or more stations associated with that switchboard either singly or on a broadcast basis. When broadcasting took place from any of the three zone switchboards, the information was also sent to a receiving-only teletypewriter at Harrisburg. The teletypewriters at the various receiving-only stations in the network were started and stopped automatically when a connection was established. Two-way teletypewriter communications could be established between each zone switchboard and Harrisburg.

In case an operator at Harrisburg wished to broadcast to every station in the state, a means was provided to automatically connect all stations at every switchboard into the broadcast network even though connections might at that time have been established at any of the zone boards.

As covered above, during the 1916–1925 period, the demand for teletypewriter private-line switching systems for use by Bell System customers was met by development and application of a number of distinctive special-purpose switching systems. During this period the planners and developers of such systems gained knowledge of customer needs and desires which led to establishment of service requirements for a family of standard private-line teletypewriter switching systems and of the service requirements for Teletypewriter Exchange (TWX) Service. The standard private-line systems and the switching arrangements for TWX Service came into use after 1925.

9.7 Telephotography

Dating back to the earliest days of modern civilization, transmission of visual information continued to be an intriguing challenge to innovators of improved means of communication over distances beyond the range of audibility. Examples of early forms of communication by coded visual signals are included in an early section of this chapter.

Bell System interest in telephotography dates back to the nineteenth century. Evidence of this interest was recorded in a comprehensive report prepared by T. D. Lockwood's organization at the American Telephone and Telegraph Company. The report, entitled "Transference of Images by Electricity," was completed during 1903. It dealt with two classes of images that might be transmitted over long wire-line media, viz., those received with no permanent record and those received and recorded. It was recognized at the time that a number of scientific and engineering advances beyond the then-existing state of communication technology offered formidable challenges.

Early within the twentieth century the burgeoning motion picture industry created an insatiable public interest and demand for "moving pictures" for entertainment and educational purposes. Responding to this stimulated public interest in pictorial displays of current news items, the newscasters increased the number of graphics and photographs appearing in newspapers and periodicals. Nationwide press associations that served competing newspaper publishers in various metropolitan areas throughout the country found it necessary to prepare and mail film negatives of photographs covering important news events to their syndicate publishers. The unpredictable time delays in transporting this pictorial information through the mail service convinced the press associations that there must be a better way. Accordingly, they turned to the telegraph companies and the Bell System for assistance.

During World War I, and in response to requests by the press associations, engineers at AT&TCo Headquarters made a new appraisal of the practicability of making a picture transmission system available for use in providing a private-line service to the press associations. Voluminous reports were prepared covering inventions submitted by outside inventors of picture systems who thought they had solved the problem. In every case it was found that the innovations were not suitable for use with Bell System long-distance wire plants. The common deficiency was a lack of appreciation of the *system problem* involving transmission of pictorial information over long Bell System wire-line circuits, particularly the effects of attenuation, delay, and noise on picture quality.

Beginning around 1918, engineers of the newly formed AT&TCo Department of Development and Research conducted exploratory work to determine the composition of a suitable picture transmission system for use with the high-quality long-distance wire-line circuits normally used for Bell System telephone or telegraph communications. Early in 1920 this exploratory work showed sufficient promise to justify intensive applied research and development effort and construction of a laboratory model of a complete system for demonstrations.

During the 1920–1923 period, engineers and scientists of the AT&TCo Department of Development and Research and at Western Electric Headquarters Engineering collaborated closely in the design and application of building-block units comprising a laboratory assembly of the system. This work was completed in time to make a private demonstration of capabilities at the Bell System President's Conference held at Yama Farms, New York, during May 1923. Pictures were sent from the Western Electric Engineering Laboratory in New York City to Yama Farms over a telephone circuit of about 60 miles. The receiving unit at Yama Farms was demonstrated as the pictures were being received. The enthusiastic response to this demonstration led to authorization for development of a commercial system for transmission of pictures over Bell System private-line telephone circuits (July 25, 1923).

A public demonstration of the preliminary design of the commercial system was made on May 8, 1924, between Cleveland and New York and subsequently used for sending pictures from Cleveland to New York covering highlights of the 1924 Republican National Convention. A second system was also set up between Chicago and New York during 1924 and used to send pictures to New York of important views of activities during the Democratic National Convention held in Chicago. The 5″ × 7″ pictures derived from these experiments were published in newspapers within these cities. The prompt publication on the same day that the pictures were taken provided excellent publicity to demonstrate the time advantage of transmitting pictorial information electrically over long wire-line media as compared to sending pictures by mail.

Encouraged by the general public interest in this "rapid" electronic means of transmitting pictorial information over long distances, Bell System management decided to have additional telephotograph machines installed in other cities. The plan would provide means of sending a picture from a specific point of origin to one or more designated receiving points, thus allowing study of the scope and character of nationwide market demand.

To implement the plan it was decided to modify the telephotograph machines to improve the fineness of grain of the received picture. The scanning rate was increased from 65 to 100 lines per inch and recording on the receiving film was changed from the variable-line to the variable-density method.

During the latter part of 1926 the modified machines at New York, Chicago, and Cleveland, and machines installed in Boston, Atlanta, Cleveland, St. Louis, and Los Angeles, provided the means of testing the market demand for a transcontinental commercial service involving the transmission and recording of pictorial information electronically.

The composition of the telephotograph system that came into commercial use between the pairs of cities mentioned above is shown schematically on Fig. 7-25. The diagram shows only half of an installation at each station, i.e., the means used to transmit a picture from Station A to Station B for reception and recording. Transmission of the visual information from A to B requires only one of the two available paths in a 4-wire line circuit. The other half is thus available for transmitting information from B to A.

The building-block composition of the sending apparatus includes:

(*i*) A motor-driven carriage containing a cylinder upon which the positive 5″ × 7″ transparent print is mounted. The film cylinder is moved along its horizontal axis by means of a micrometer screw at the same time that it is rotated. Thus, the mechanism provides the means of causing a small spot of light to traverse the entire film area in a long spiral.

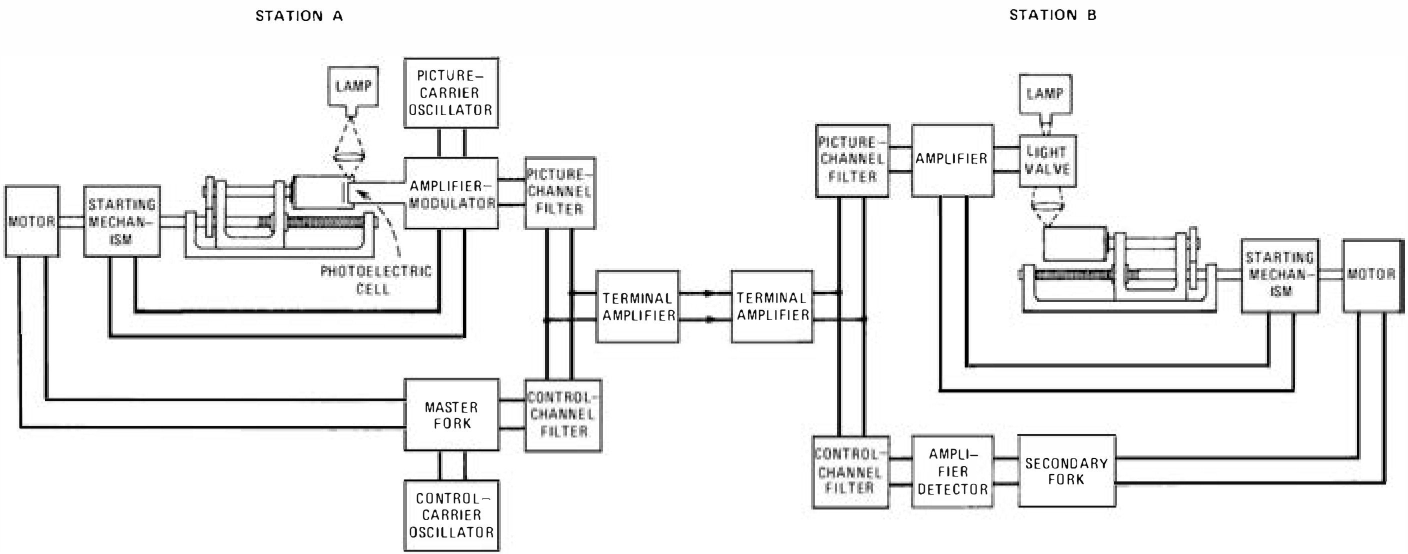


Fig. 7-25. Simplified schematic of telephotograph system. (Redrawn from Ives et al. 1925, Figs. 9 and 10)

(*ii*) An optical system, shown on Fig. 7-26, which includes a photoelectric cell, shown in Fig. 7-27. The photoelectric cell, a vacuum tube device containing an alkali-metal cathode, is capable of emitting electrical currents of varying values that are directly proportional to the varying amounts of light that pass through the positive film transparency onto the photoelectric cell during the point-by-point scanning operation.

(*iii*) A picture-carrier oscillator to furnish a frequency of 1,300 hertz.

(*iv*) An amplifier-modulator which amplifies the currents of varying value that are emitted from the photoelectric cell, and in turn modulates the 1,300-hertz carrier frequency from the oscillator to provide pulses of varying amplitude in proportion to those emitted by the photoelectric cell.

(*v*) A band-pass filter having a "flat in-band" characteristic within the range of frequencies from about 800 to 2,000 hertz.

(*vi*) Apparatus to provide accurate synchronization of the rotating cylinders at the sending and receiving stations. The apparatus at the sending station includes a phonic-wheel motor controlled by an electrically operated tuning fork, a 400-hertz control-carrier oscillator, and a band-pass filter having a passband of approximately 300–500 hertz. Synchronization of the cylinders at the sending and receiving stations is obtained by sending pulses from the 400-hertz source to the receiving station. The pulses are controlled by the electrically operated tuning fork at the sending station. The received pulses are used to control the speed of the electrically operated tuning fork and phonic-wheel motor at that station.

The apparatus at the receiving station is comprised of:

(*i*) A motor-driven carriage assembly similar to that provided at the sending station, including a cylinder for mounting an unexposed film negative.

(*ii*) An optical system, shown on Fig. 7-28, which includes a light valve, shown on Fig. 7-29. The system provides the means of projecting a beam of light having an intensity corresponding to that absorbed by the photoelectric cell at the sending station for each small area of the film exposure during movement of the synchronized cylinders.

(*iii*) A band-pass filter and amplifier to provide input of the 1,300-hertz pulses received over the picture channel.

(*iv*) A band-pass filter, amplifier-detector, secondary tuning fork, and phonic-wheel motor to supply means of driving the cylinder and maintaining it in synchronism with the cylinder at the sending station.

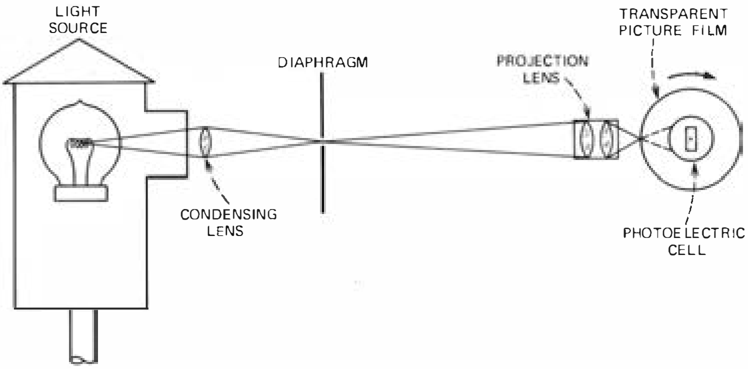


Fig. 7-26. Cross section of sending-end optical system. (Redrawn from Ives et al. 1925, Fig. 1)

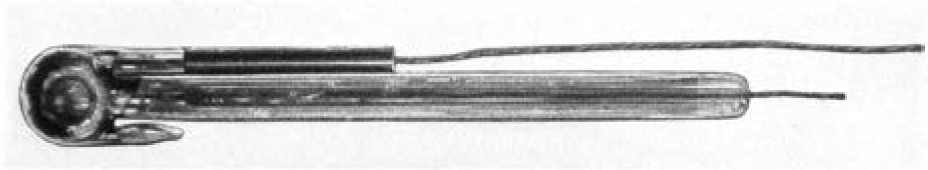


Fig. 7-27. Photoelectric cell of type used in picture transmission. (Ives et al. 1925, Fig. 2)

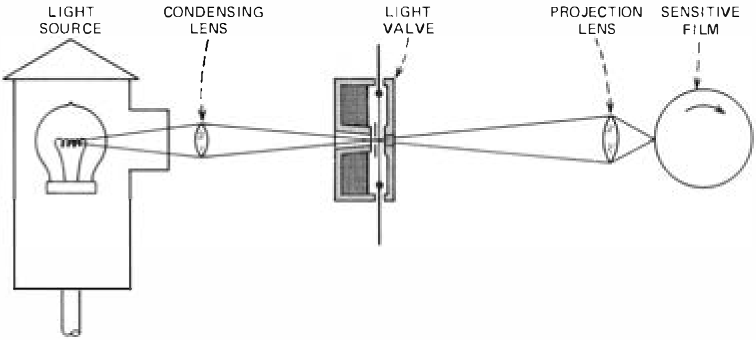


Fig. 7-28. Cross section of receiving-end optical system. (Redrawn from Ives et al. 1925, Fig. 4)

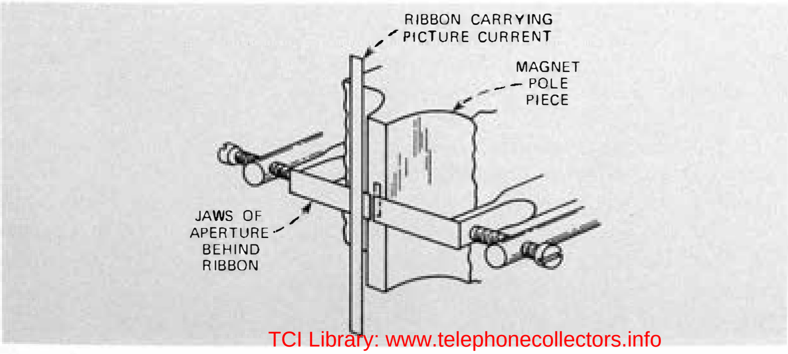


Fig. 7-29. Details of light valve. (Redrawn from Ives et al. 1925, Fig. 3)

Figure 7-30 shows the apparatus at the sending end and Fig. 7-31 that at the receiving end of the system.

The time of transmission of each 5″ × 7″ picture was, for a 100-line-to-the-inch picture, about 7 minutes. This time was a relatively small part of the total time required from the taking of a picture until it was delivered to the user in the form of a print. If the transparent positive at the sending end was loaded on the cylinder while still wet and a print was made at the receiving end from the exposed negative before it was dry, the overall time was in the order of 45 minutes. Figure 7-32 is a 5″ × 7″ print of a picture taken during the inauguration of President Coolidge on March 4, 1925, and sent from Washington simultaneously to New York, Chicago, and San Francisco.

Beginning during 1927, the AT&TCo Long Lines Department made use of the eight installations of the telephotograph system described above, which was coded Type A, to furnish coast-to-coast commercial service to its customers. Interest in this service, particularly by press associations, led to the decision by Bell System management that an orderly research and development program leading to an improved system should be undertaken by Bell Telephone Laboratories (formed during 1925). Accordingly, in a letter dated July 27, 1927, E. H. Colpitts (AT&TCo) authorized E. B. Craft (BTL) to carry out research and development effort on the Type B system. The Type B system is covered in a later volume of this history.

9.8 Television

A demonstration of the experimental telephotograph system during 1923 (see Section 9.7) provided convincing evidence that transmission of pictorial images over long wire-line circuits used for telephone communications had reached the stage of commercial practicality. Consequently, the teams of researchers and technical planners in AT&TCo and Western Electric Engineering who had combined their skills to reach that plateau of accomplishment were eager to explore the technology of television, i.e., transmission of information defining images in motion over long distances.

Herbert E. Ives,who was a major contributor in the technical aspects of the work on the telephotograph system, was asked by H. D. Arnold to draw up a research program on television, starting with a modest attack on the most fundamental problems, with provision for expansion to assembly of an experimental system.

Fundamentals of television, in the broadest sense, include reception of a stream of varying-intensity-light signal elements reflected from a moving image being scanned, conversion of these elements into electrical signals, transmission of the electrical signals to a distant

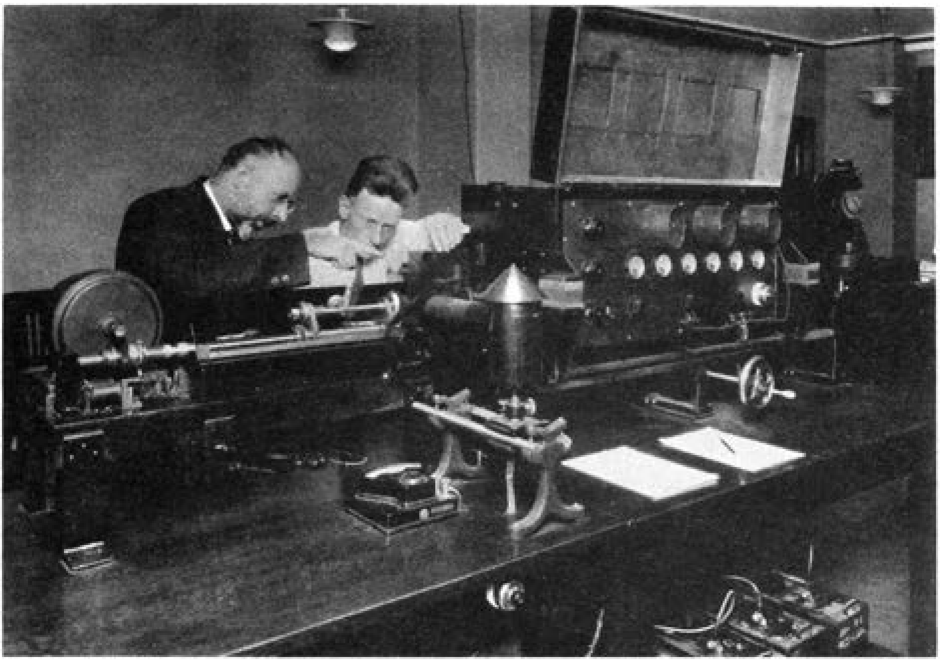


Fig. 7-30. Sending-end apparatus showing motor, film carriage, optical system, and amplifier-modulator. (Ives et al. 1925, Fig. 7)

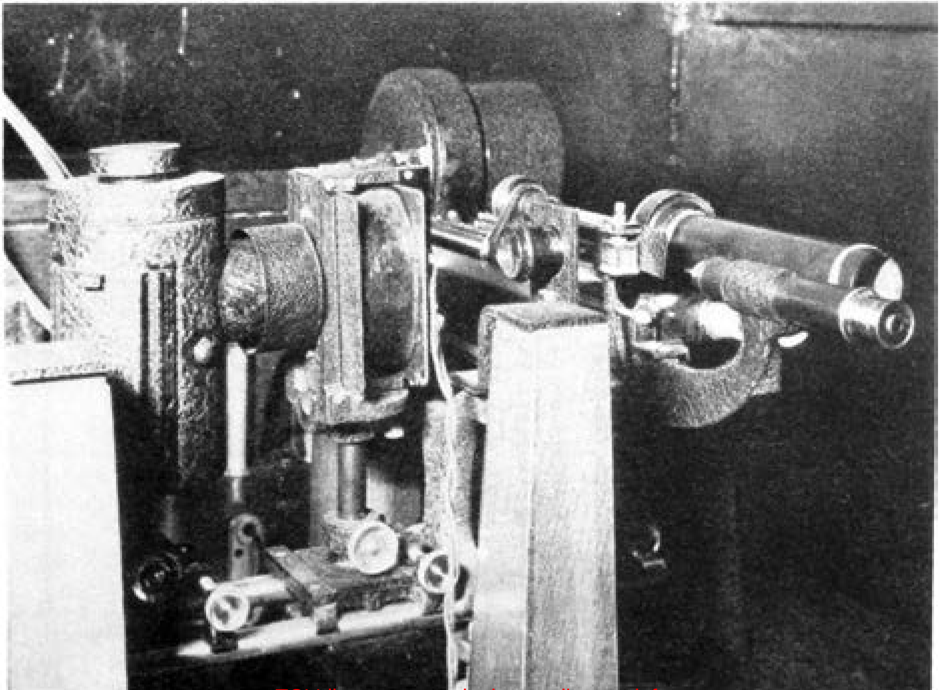


Fig. 7-31. Receiving-end apparatus showing light valve and observation microscope. (Ives et al. 1925, Fig. 8)

Fig. 7-32. Telephotograph sent, using a 100-line-per-inch scan, over a transcontinental circuit.

location, and then conversion of the electrical signals into light signal elements which by means of a scanner are used to recreate the moving image.

The researchers and planners were confronted with the problem of formulating means of performing these sequential functions to a requisite degree of sensitivity, speed, efficiency, and accuracy in order to recreate a changing scene without appreciable lapse of time and in a form satisfactory to the eye.

Analyses led to the conclusion that the limited capabilities of optical and transmission systems at that time precluded the possibility of televising extended scenes of fast-moving objects. The available optical systems lacked sensitivity in conversion of the low levels of light intensity desired from elements of an extended image while it was being scanned. The available transmission systems lacked the capability of transmitting trains of electrical pulses at microsecond speeds.

To enable the research program to reach an experimental stage, the participants proceeded with design and assembly of functional building-block units to make up a system having the capability of televising individuals engaged in a two-way telephone conversation.

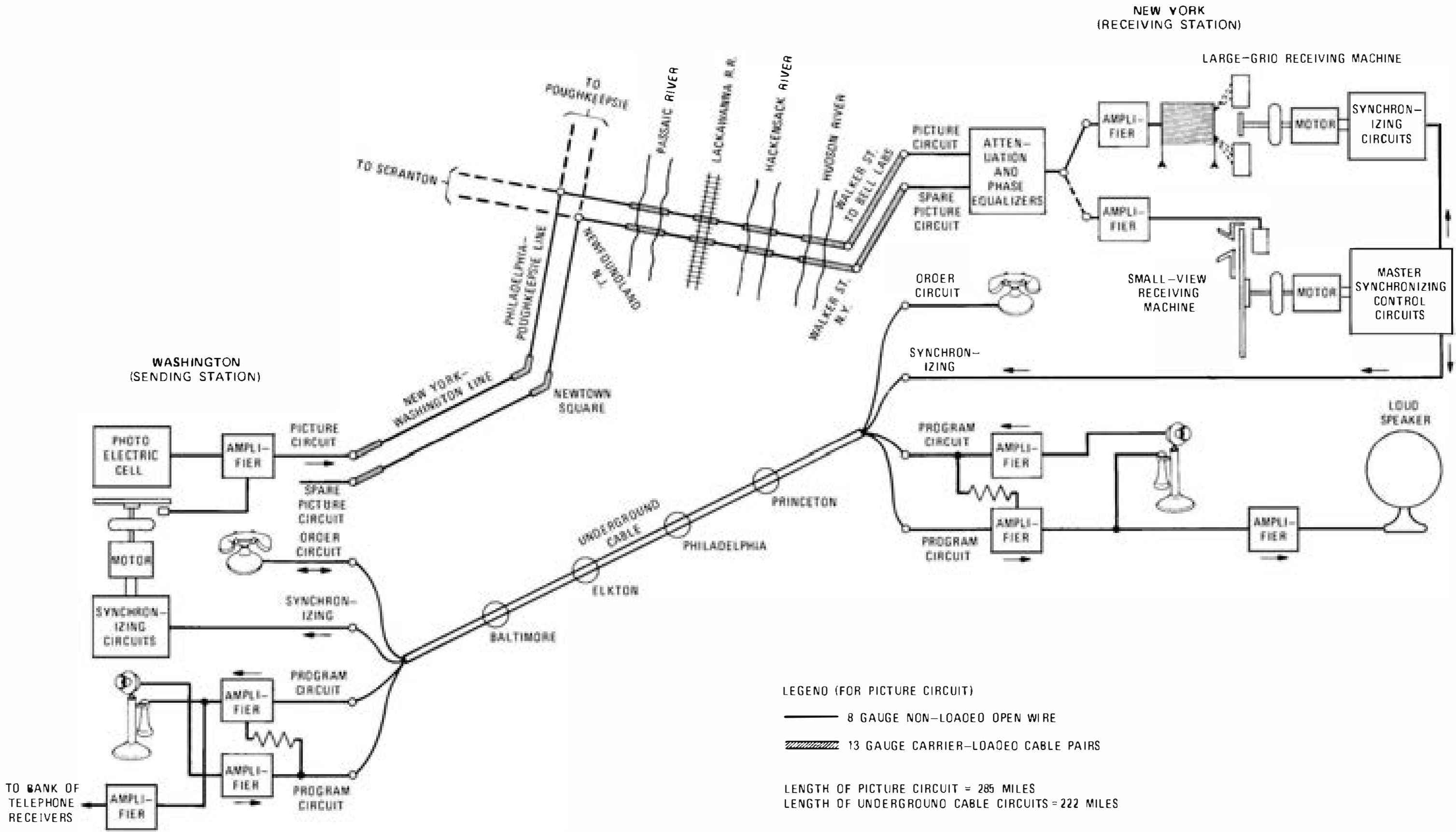
For the experiment it was decided that an adequate definition of the images would be obtained on a 50-line viewing screen comprised of 2,500 visual signal elements. The rate of displaying each complete image at least 16 times per second to extinguish flicker was used to determine the rate of scanning, i.e., 40,000 elements per second. This very modest rate was adopted to fall within the limits of capability of the existing wire-line and radio channels.

The program culminated in private and public demonstrations during April and May 1927 of a series of two-way telephone conversations between individuals located in Washington, D.C., and New York City and between others located in New York City and Whippany, New Jersey. The terminal apparatus in New York City was located in the Bell Laboratories auditorium and that in Whippany in the Bell Laboratories radio laboratory. The apparatus assembly at each of the three locations permitted a two-way telephone conversation to be conducted by individuals between each pair of terminals. Each participant observed the image of the other on a 2-inch-by-2½-inch display located near the telephone instrument.

A public address system, and a visual adjunct in the form of a 2-foot-by-2½-foot-grid luminous screen comprised of 2,500 neon lamps, were installed at the New York terminal. Thus, a performing speaker or singer, located in either Washington or Whippany, could be seen and heard by the large audience in the New York auditorium.

Figure 7-33 shows the composition of the Washington-New York system in schematic form including the wire-line facilities between the

Fig. 7-33. Schematic diagram of circuits for 1927 television demonstration. (Redrawn from Gannett and Green 1927, Fig. 2)



two locations. Radio was used as the transmission link between the New York and Whippany terminals, a distance of about 22 miles.

All participants in the demonstrations and the audiences that witnessed them were thrilled to know that at long last electrical transmission of views of images in motion over long distances had become an accomplished fact. However, the opinions of the researchers and system planners who were looking to the future were well expressed by Dr. Ives whose appraisal is quoted below:

It is not easy at this early date to predict with any confidence what will be the first or the chief uses for television, or the exact lines that future development may take. It must be clearly understood that television will always be a more expensive service than telephony, for the fundamental reason that it demands many times the transmission channel capacity necessary for voice transmission. This expense will inevitably increase in proportion to the size and quality of the transmitted image.

The kinds of service which are naturally thought of upon consideration of the services now rendered in connection with sound transmission are: first, service from individual to individual, parallel in character to telephone service, and as an adjunct thereto; second, public address service, by which the face of a speaker at a distant point could be viewed by an audience while his voice was transmitted by loud speaker; third, the broadcasting of scenic events of public interest, such as athletic contests, theatrical performances and the like.

The first two types of service just mentioned lie within the range of physical practicability, with apparatus of the general type already developed. The third type, because of the uncontrolled conditions of illumination, and the much finer picture structure which would be necessary for satisfactory results, will require a very considerable advance in the sensitiveness and the efficiency of the apparatus, to say nothing of the greatly increased transmission facilities. For all three types of service, wire or radio transmission channels could be utilized, for while the problems incident to securing distortionless transmission over wide frequency bands, or multiple transmission channels, are different in detail in the two cases, they appear to be equally capable of solution by either means. However, the very serious degradation of image quality produced by the fading phenomena characteristic of radio indicates the practical restriction of radio television to fields where the much more reliable wire facilities are not available.

The above coverage highlights the activities and accomplishments of the research and engineering teams that brought the "television" system experiment to fruition during the 1920s. Those interested in details relating to this extraordinary accomplishment are referred to the following technical articles published in the October 1927 issue of *The Bell System Technical Journal*:

"Television," by H. E. Ives

"The Production and Utilization of Television Signals," by F. Gray, J. W. Horton, and R. G. Mathes

"Synchronization of Television," by H. M. Stoller and E. R. Morton

"Wire Transmission System for Television," by D. K. Gannett and E. I. Green

"Radio Transmission System for Television," by E. L. Nelson.

The knowledge gained from this combined research and system-engineering experiment provided the base for further attack on all of the technical problems relative to televising extended scenes containing images in motion and televising images of individuals conducting a two-way telephone conversation. For details covering these activities the reader is referred to a later volume of this history. Meanwhile, to illustrate how industries are born and thrive, as a result of an important breakthrough in science, the evolution of the television industry will be highlighted below.

Motion picture producers and entrepreneurs in the budding broadcasting and electronics manufacturing industries realized the potential of televised broadcasting of athletic events, of studio and theatrical performances, and of a gamut of other events of public interest. Scientists and engineers in these industries were motivated to direct effort on television instrumentation problems.

Toward the end of the decade of the 1930s, cameras and optical systems having capabilities of televising extensive scenes had reached the stage of commercial practicality. Concurrently, black-and-white television receivers became available for public use.

Following the end of World War II, commercial television broadcasting became a flourishing business. The early success of this mode of broadcasting created public demand for national coverage of important events. The Bell System was ready to meet the demands of the television broadcasters for interconnection of studios. Carrier-derived transmission channels available as spares within the long repeatered coaxial cables having capabilities for telephone and telegraph use could readily be combined to provide transmission of televised picture-element pulses utilizing bandwidths of 3 MHz or more depending upon the needs of the broadcasters. By 1951, highly reliable microwave radio facilities were augmenting the coaxial cable to provide transcontinental service for television networks. The Bell System again demonstrated its ability to meet public demand for high-quality communication facilities as the need arose.

9.9 Summary of Accomplishments from 1910 to 1925

Founders of the Bell System and particularly Theodore N. Vail foresaw the importance of a unified approach to the development and expansion of its telephone and telegraph businesses. Nevertheless, during the early years of Bell System history, public acceptance of the more-convenient mode of communicating by telephone set the style of its research and development programs. Prior to 1910, many scientific and engineering advances were made to improve telephone service. They were applied to the open-wire and cable plants, to customer station apparatus, to switchboards and systems located in central offices, and to facilities located in test and maintenance centers. The technical advances made and applied to the telegraph plant were limited to those described in Section VIII of this chapter. At the end of the period the telegraph business of the Bell System was restricted to the funishing of Morse telegraph private-line services.

During 1910, Bell System management decided that the time had come to stimulate research and development effort on telegraph and other non-voice modes of communicating. Emphasis was placed on development of printing-telegraph instrumentalities and on telegraph plant arrangements to anticipate furnishing convenient telegraphic communication services comparable to those being furnished by the telephone.

The management decision provided the impetus for research, engineering, and development groups at AT&TCo Headquarters and Western Electric Engineering to undertake coordinated programs that led to introduction of a succession of new products into the Bell System plant. The programs that were activated during the 1910–1925 period have been described in Sections 9.1–9.8 of this chapter. The classes of products that became available for commercial use during the period are identified in the following listings:

Start-stop printing-telegraph instrumentalities:
- Nos. 12 and 13 page-type printer sets (*Teletype* Corp.) and also earlier replaceable designs
- Nos. 15 and 19 page-type teletypewriters (*Teletype* Corp.)
- Tape-type printer and associated apparatus
- No. 14 tape-type teletypewriter (*Teletype* Corp.)
- Regenerative repeater
- Four-channel multiplex printing-telegraph system with start-stop printer extensions
- Cipher printing-telegraph system.

Miscellaneous telegraph apparatus:
- 209- and 215-type polar relays
- Capacitors, inductors, and resistors used exclusively in telegraph applications
- Noise killers and telegraph-thump suppressors
- Ground-potential compensators
- Hit suppressors and lightning drainage networks
- Cross-fire neutralizers.

Telegraph transmission systems:
- Improved designs of neutral and polar telegraph repeaters having capabilities of furnishing printing-telegraph communications over

dc grounded simplex and composite "legs" derived from 2-wire open-wire line circuits

DC telegraph equipment for Key West–Havana submarine cables

DC metallic telegraph system for use with composited circuit derived from a 4-wire fine-gauge cable circuit

High-frequency carrier telegraph system for use above a voice-telephone circuit over a 2-wire open-wire line (capacity of ten telegraph circuits)

High-frequency carrier telegraph system for use with the Key West–Havana cable

Voice-frequency carrier telegraph system for use with a 4-wire fine-gauge cable circuit exclusive of voice telephone (capacity of 12 telegraph circuits).

Teletypewriter switching systems:

Loop switchboards for use on private-line-customer premises

Teletypewriter switchboards of special design to meet service needs of each particular private-line customer.

Telephotograph system:

This is the system that was installed in eight metropolitan areas of the United States to provide commercial service from coast to coast.

Testing and maintenance apparatus and equipments:

Tables and panels for testing 209- and 215-type polar relays

Automatic means of sending teletypewriter test signals over telegraph lines

Telegraph-stability test set using reversals as a source of supply

Telegraph-transmission test sets for measuring distortion of start-stop teletypewriter signals at receiving terminals of various transmission media.

No. 5 testboard (telegraph positions of this board contained improvements in features over those contained in the earlier No. 4 Morse Board).

Thus, at the end of its first 50 years, the Bell System had successfully demonstrated the unique value of systematic and coordinated research and development effort in the advancement of technology, not only in telephonic but also in telegraphic and other non-voice modes of communicating. At the end of the period (1925) the Bell System was positioned to furnish nationwide public telephone and non-voice services at a technical level superior to any other throughout the world. In the television field the successful research experiment and related published papers (1927) provided technical guidance to the radio-television manufacturing and broadcasting industries during the development of television broadcasting equipment and television receivers which came into commercial use following World War II.

Chapter 8

Materials and Components

A great deal has been said about the components used in transmission, switching, and station systems in Chapters 3, 4, 5, and 6. As new communication systems were developed, there was an increased need for new kinds of components having highly specialized characteristics in the areas of electrical performance and physical configuration. Rarely was it possible to use the ordinary commercial product of the time—developed mainly for the electrical power and later for the radio entertainment industry—to meet the requirements posed by carrier transmission systems, high-quality audio systems, and radio communications. Substantial research and development effort went into the design of vacuum tubes, transformers, loading coils, relays, resistors, and capacitors to ensure long-lived, trouble-free performance. This chapter supplements what has been said about apparatus and components in the earlier portions of this history and complements what will be covered in Chapter 10 relating to physical phenomena and the properties of materials.

I. INTRODUCTION

There are many components used in a telephone system, most of which have required continuing development to meet the changing requirements of new communications technology. As one examines the technical history in some detail, it becomes apparent that advances in components and innovations in design have, in fact, been a necessary condition for progress in systems development, as for example, the vacuum tube for long-distance transmission, and the wave filter for multichannel carrier telephony.

Only a few of the devices will be described in this chapter, as illustrations, to supplement what has been said in preceding chapters in relation to station apparatus, transmission, and switching.

Magnetic materials, inductors, transformers, resistors, capacitors, and vacuum tubes will be discussed briefly as examples of the interrelatedness of systems and their components.

Resistors and capacitors were initially basic laboratory devices used in connection with instrumentation for electrical measurements. They had been traditionally manufactured in small quantities with few restrictions as to shape, size, and interchangeability. Their application to the new and special needs of telephony required that there be a carefully written specification covering element values, a stated tolerance that defined the limits of manufactured product, a full description of the materials to be used, their purity, treatment, and inspection attributes, and other specified qualities that would ensure long life and stable performance. And of equal importance, the design of the component had to be such that very large quantities of these specialized products could be made at a reasonable cost. Thus, as early as 1901, the flat 18- and 19-type resistors were designed specifically for mounting in relay racks, in a flattened form that resulted in low inherent inductance and minimum space requirement, readily manufacturable in quantity, inexpensive, and with very long service life. The application of capacitors to telephone needs took a similar path, evolving from a laboratory standard made in small numbers, to a mass-produced, long-life device with controllable characteristics. It was not a simple transition, for it required years of study of the properties of insulating paper used as a dielectric, the conducting foil, and the impregnants before the desired stability and cost criteria could be attained.

The design and use of inductors and transformers specifically for communication circuits is another illustration of sharp departure from the engineering practices of the early 1900s. In contrast to electrical power applications, the devices used for telephony were required to cover a wide frequency range and to operate with very high efficiency at low flux density. Accordingly, the development of suitable components was early directed to a search for magnetic materials having a high permeability at low magnetizing force, low hysteresis and eddy-current losses, and to designs that minimized distributed capacitance and leakage inductance. The problems of impedance matching and of operation with superimposed direct current in the windings were further complications for the communications transformer engineer. These special requirements had not existed in the prior art, nor were the right materials available. The solutions were found in a program of research on new magnetic materials, notably with the invention of permalloy and other high-permeability alloys, and in the analytic approach to transformer design based on network theory.

The vacuum tube illustrates another way in which new ideas were applied to communications technology. What was needed was a device with consistent performance, utmost reliability over a long life, predictable characteristics that could be controlled in manufacture, and the ability to be produced in large quantities at reasonable cost. As a component with a high degree of complexity, it required scientific studies that would ensure a basic understanding of the physical phenomena involved. With this background it became possible to provide the advanced technology needed by the system designer as he moved to higher and higher frequencies and faced more stringent limitations on noise and distortion.

The problems of design were thus somewhat similar for the relatively complex vacuum tubes and communication transformers. They were solved largely by development of classes of general purpose apparatus and, in part, by special devices for specific systems. Advances in capability followed from fundamental studies of the properties of materials and of the physical principles involved. That new devices opened the way to new kinds of systems is a theme that recurs in the history of communications technology, from the loading coil to the transistor and beyond.

II. MAGNETIC MATERIALS

2.1 Introduction

Magnetic materials have played a vital role in the Bell System almost from its beginning, and research on these materials was actively pursued by the Western Electric Company and its successor, Bell Telephone Laboratories. In particular, 1913 to 1935 was a period of extraordinary creativity in the field of magnetic materials, unmatched anywhere else in the world.

There were two principal areas of activity. One was primarily a search for better understanding of the fundamental physics of magnetism. Typical interests were in crystal structure and magnetostriction. Work in the second area was directed to the discovery of new magnetic materials, mostly under the direction of G. W. Elmen, a man well-trained in physics and electrical engineering, and an inspired, persevering experimentalist. The two areas of research complemented each other, as will be noted in Chapter 10, Section IV, "Research in the Sciences of Materials," where Elmen's work on permalloy and L. W. McKeehan's work on the crystal structure of iron-nickel alloys are described.

2.2 Permalloy—The Binary Alloys

In 1913, when Elmen first undertook his investigation of nickel-iron alloys, the magnetic materials used in transformers and coils were principally iron wire and some silicon-steel laminations, both obtained from commercial sources. The iron wire, in annealed or in partly hard-drawn form, had been used in coils in the Bell System since before 1900. However, the permeability of both the iron wire and silicon steel used had an upper limit in the order of only 300 at the weak fields corresponding to voice currents.

Elmen was also concerned with the study of continuous loading of Bell System submarine cables, where higher permeability in the magnetic materials to be used for loading was of pressing importance. This was in addition to the need for higher-permeability materials for higher efficiency in apparatus. Accordingly, he tried to find a material that, at magnetizing forces of less than 0.2 oersted, would have higher permeability and also lower losses than either silicon steel or iron wire. By the end of 1915, he had obtained initial permeabilities[1] of as high as 2,500 for a 70-percent nickel, 30-percent iron alloy.

By 1921, it had been found that an alloy of 78.5-percent nickel and 21.5-percent iron, specially heat-treated, gave initial permeabilities of upwards of 9,000, and not long thereafter, a value of 12,000 was obtained.

At that time, the invention of permalloy was perhaps the outstanding engineering contribution in magnetics. The realization of permeabilities one or two orders of magnitude greater than previous materials, with correspondingly lower hysteresis losses, was of great scientific interest. The engineering value of permalloy was attested to by its immediate use in transformers in which a frequency band could be transmitted that was four or more times wider than was otherwise possible, in transformers having less than one third of their previous volume; and by its use in the continuously loaded transatlantic telegraph cable laid in 1924, where the loading increased the capacity of the cable from 250 to 1,900 letters per minute.

An indication of the interest in permalloy was the exhibit at the Chemical Industries Exposition in New York City in 1925, shown in Fig. 8-1. This exhibit, devised by Bell Telephone Laboratories, attracted a great deal of attention. By turning the knob at the top of the exhibit, visitors could complete a magnetic circuit by successively introducing rods of brass, iron, and permalloy. A current was passed through a winding placed around one of the legs of the magnetic circuit. When this current was reversed, the change in magnetic flux induced a voltage in a second coil connected to a meter. With the brass rod in position and the current reversed, the meter showed no deflection, and with iron only a slight deflection. With the permalloy rod in position, the meter went off scale, giving dramatic evidence of its high permeability.

By 1925, 78½ permalloy and 45 permalloy were well-established materials for design. The 78½ permalloy was preferred where the

[1] Initial permeability is the permeability when the flux density becomes vanishingly small.

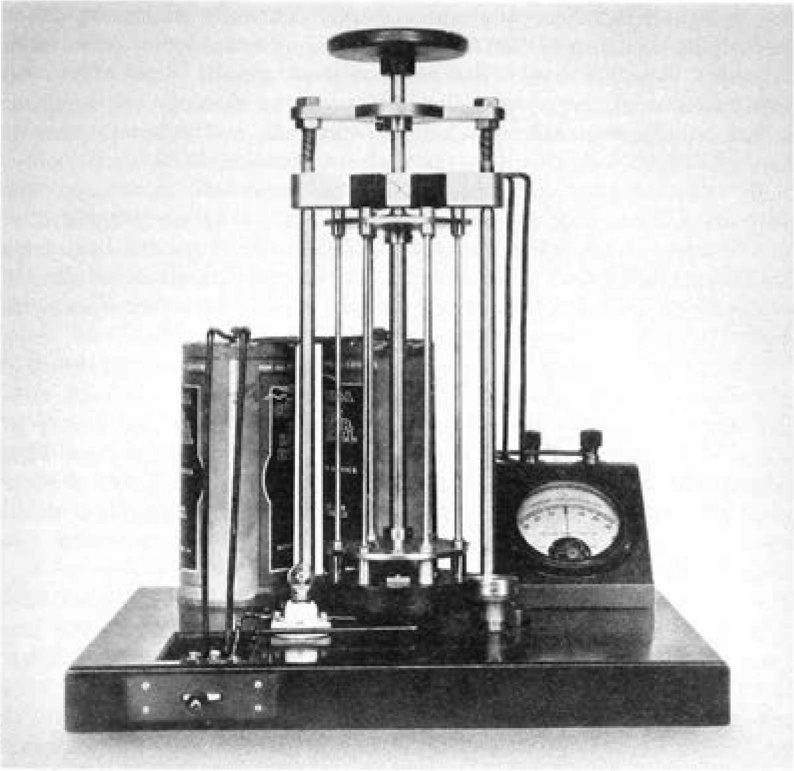


Fig. 8-1. Permalloy exhibit at the Chemical Industries Exposition of 1925. This exhibit permitted the visitor to test for himself the permeability of brass, iron, and permalloy.

highest permeability was needed, as in high-quality audio transformers working into grid circuits of vacuum tubes; and 45 permalloy was used where its lesser sensitivity to direct current was important, such as in transformers working out of the plate circuits of vacuum tubes. The continuing work on permalloy involved the ternary alloys described in the next section. The development of high-nickel permalloy as compressed powder for use in magnetic cores for loading coils is described in Section 2.4.3 of this chapter.

2.3 Permalloy—The Ternary Alloys

At audio frequencies and higher, eddy currents usually account for a large part of the losses in magnetic materials. The continuing development to improve permalloy was therefore directed to increasing its resistivity, thereby minimizing eddy currents. Since it was well known in metallurgy that the addition of certain elements in small amounts would increase the resistivity of metals, a program was initiated to study the effect of the addition of the elements silicon, chromium, manganese, antimony, copper, tungsten, and others to permalloy, while hopefully retaining the high permeability of 78½ permalloy.

By 1921, an alloy of 55-percent nickel, 11-percent chromium, and 34-percent iron was developed, with a resistivity of 100 microhm centimeters, compared to about 17 for 78½ permalloy, and having an initial permeability of 1,000 or more. This was followed by an alloy of 78.5-percent nickel, 8-percent chromium, and 13.5-percent iron, with a resistivity of 90 and an initial permeability above 2,000.

At first, the addition of the third element increased the resistivity at the expense of permeability, but further work showed that with chromium permalloy it was possible to obtain higher resistivity as well as initial permeability comparable to or greater than 78½ permalloy. By early 1925, an eminently useful material with a permeability of 13,500 and a resistivity of 45 was obtained with an alloy of 78-percent nickel and 1.9-percent chromium, with the remainder consisting of iron.

In the team environment of Bell Telephone Laboratories, the apparatus designers were aware of the development of and the superior characteristics of the new chromium permalloys. To facilitate the introduction of new alloys in designs, a magnetics expert was placed in the design organization, with the specific responsibility of evaluating new materials and establishing requirements for manufacture. The designers were eager to test sample laminations of the new materials. The new alloys, however, could not be incorporated in designs until it was established that they could be manufactured to reasonable magnetic performance tolerances, and that the Hawthorne Works of the Western Electric Company could make laminations on a dependable production basis. By 1929, the supply of laminations was so assured, and 3.8 chromium permalloy with a permeability of 10,000 and a resistivity of about 60 was the alloy chosen for transformers.

Development work on permalloy containing molybdenum followed closely after the work with chromium. In 1925, a permeability of over 20,000, higher than that previously attained with any magnetic material, was realized in an alloy of 78.5-percent nickel, 3.7-percent molybdenum, and 17.8-percent iron, using a special heat treatment with a precisely controlled cooling rate. (This alloy had a resistivity of 55.) The characteristics of the molybdenum permalloy and the earlier chromium permalloy were sufficiently similar to make them rival materials. The earlier alloy, 3.8 chromium permalloy, was being used in audio-frequency transformer designs with magnetic characteristics highly satisfactory in 14-mil lamination form and much superior to 78½ permalloy. Since it would have been uneconomical to standardize both materials, a choice between them had to be made.

One of the determining factors in the choice between the two alloys was surface effects in chromium permalloy. The permeability of the chromium permalloy laminations was not as high as the intrinsic permeability of the alloy itself. Studies revealed that there was a surface layer in the laminations whose permeability and losses were markedly inferior to the material in the interior of the laminations. When this surface layer was etched off with acid, the lamination, though thinner, exhibited the high permeability characteristic of the alloy. The "layer effect" was not pronounced in laminations 14 mils and thicker at low audio frequencies, the first and most important use. At higher frequencies, where eddy currents cause a higher concentration of magnetic flux near the surface, and in thinner laminations, where the surface layer is a larger proportion of the cross section, the effect is serious. The chromium permalloy was also found to have a harder surface, causing greater wear on punching dies than did molybdenum permalloy.

Since the 3.7 molybdenum permalloy did not show the surface effects, and the trend in development work was toward higher frequencies where surface effects were most deleterious, a decision was made to standardize the molybdenum alloy. This decision, made around 1930, has stood the test of time, and this alloy is still the principal permalloy, both in laminar and in powder form.

2.4 Magnetic Materials for Loading Coils

In passive components, the largest and most important use of magnetic materials was the loading of telephone lines (see Chapter 4, Section IV). The demand for loading coils in the Bell System has always been very large, with an accumulated total production through 1915 of 530,000, continuing at the rate of 90,000 per year. The average annual production rose to 600,000 coils in the 1920s. Since the need for loading is a unique feature of a telephone system, and the demand for telephony was great, it was necessary to have a continuing research and development program on magnetic materials specifically for loading.

By far the largest factor in the efficiency of loading coils is the magnetic material used in them. The development of loading coils, accordingly, has been linked closely with the development of magnetic materials specifically for such use, with the closest cooperation between research and loading-coil development people. In the years up through the 1920s, loading was of the highest technical and economic importance to the Bell System; and despite the introduction of carrier systems on a large scale, loading was still important in the 1970s, with annual production of about 10 million coils.

2.4.1 *Wire Cores*

From the early 1900s, loading coils were of toroidal-type construction, consisting of hard-drawn insulated steel wire 0.004 inch in diameter, obtained from commercial sources, the fineness of which assured low eddy-current losses. The hard-drawn wire was preferred over soft iron wire even though the permeability was lower, because it was magnetically more stable and had lower hysteresis, which meant lower losses in the magnetic core. The wire was produced in permeabilities of 95 and 65, the lower but more stable permeability being attained by less frequent annealing between draws. Loading coils with such cores were used in large numbers but were not on the whole as stable as desired. For example, a superposed dc magnetizing force of $H = 45$ would reduce the inductance of a 65 permeability iron-wire core to 35 percent of its original value, recovering to only 62 percent upon removal of the current. In composite systems transmitting telegraph and voice currents simultaneously, the variation in inductance of the loading coils with large telegraph currents caused a corresponding variation in the voice transmission, a highly objectionable effect known as "flutter." The addition of air gaps in 1914 by the method of cutting cores in half and inserting non-magnetic spacers was only partially effective in reducing inductance instability.

2.4.2 *Iron-Powder Cores*

Research had been pressed to develop improved magnetic materials for loading. The concept of using divided iron to minimize eddy-current losses had appeared in the earliest literature, and around 1913, J. B. Speed, who was "rather inventive and something of a genius," according to one of his contemporaries, conceived of a core of insulated iron powder compacted under high pressure. The fineness of the powder would ensure low eddy-current losses, and the myriad air gaps due to the insulation between the particles would provide more than adequate inductance stability. Despite intensive and varied experimentation, a problem persisted: obtaining an insulation thin enough to minimize the separation of the particles, yet tough enough to withstand the 100,000 or 200,000 pounds per square inch pressure required to compact the cores without having insulation breakdown. Both high pressure and thinness of insulation are essential to reduce the effective air gaps and provide a dense core, with adequate permeability upwards of 50. The advent of World War I increased the need for a solution to the problem, since it prevented the importation from Europe of the diamond dies for drawing the wire used in the wire type of cores.

As so often happens in intensive research, a fortuitous circumstance caused the breakthrough. It was observed that cores consisting of powder milled in one drum, which had originally been used for another purpose, were superior to those from powder milled in other drums. In searching for reasons for the superior performance, an examination of the particular drum revealed that it was zinc-lined, and chemical tests quickly disclosed that there was zinc in the coating over the iron particles. Further work confirmed that this coating, plus shellac, was tough enough to withstand the high compacting pressures. A commercial process for their manufacture now was immediately available, and iron-powder cores[2] were placed in production in 1916.

The process of manufacturing the cores, in brief, was to ball-mill electrolytically deposited iron into powder and to anneal the powder in cast-iron boxes at 850°C. The iron powder, mixed with zinc flakes, then was rolled in a drum, the flakes removed, and shellac applied. Cores were pressed in steel dies at 200,000 pounds per square inch, which gave them a high tensile strength as well as good permeability. The cores were made in Grades "A," "B," and "C," with permeabilities of 55, 30, and 26, respectively, and with the lowest permeability having the lowest losses. Grade "A" consisted of 80-mesh annealed iron powder and Grade "B" of 90-percent annealed and 10-percent unannealed powder. Grade "C" differed from "B" in fineness of the powder (200 mesh) and in the use of somewhat heavier insulation. The choice of the grade in any coil application depended on the importance of high permeability in relation to low losses in the specific design.

The development of iron-powder cores was a major accomplishment in the field of magnetics. The powder cores were half the volume, cheaper, and more stable, with lower losses than the iron-wire cores with air gaps (see Fig. 8-2). Loading coils using these cores were admirably suited to the needs of the telephone plant because of their stable inductance and low losses, and for these same reasons, the cores also found important uses in retardation coils for filters and networks.

2.4.3 Permalloy-Powder Cores

Despite the success of iron-powder cores, magnetic materials research—the search for even better and cheaper materials—continued. Permalloy, with its characteristically high permeability and very low hysteresis, was a logical material to use for powder cores.

[2] The success of the development was largely due to the direction of W. Fondiller.

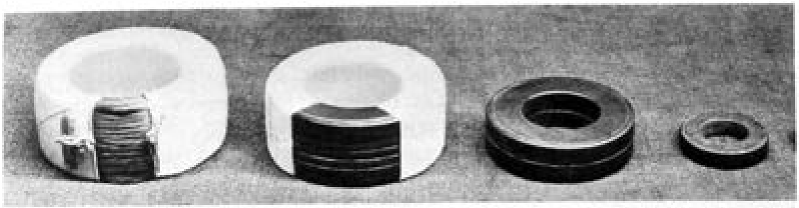


Fig. 8-2. Iron-wire and iron-powder cores.

The first step in the long development of permalloy-powder cores was taken in 1920. A mixture of fine powders of 78-percent nickel and 22-percent iron was heated to cause alloying and, at the same time, to form a sintered mass. This mass was crushed in a ball mill to form powder, which then was placed in dies and molded under pressure into cores. Although the results were not significantly superior to the iron-dust cores, the experiment did demonstrate that permalloy-powder cores could be made, and this discovery provided the basis for broad patent coverage for nickel-iron powders.

The development of permalloy-powder cores was a far more difficult undertaking than the work on earlier iron-powder cores. Permalloy, which was characterized by such high permeability and low losses, was also physically soft and malleable. The latter property made reduction to powder difficult. Permalloy was also highly strain-sensitive, so that the deformation of the particles at the high pressures up to 200,000 pounds needed to compact the powder would in turn reduce the permeability to values comparable to ordinary iron. Also, to capitalize on the high permeability of the permalloy, it was essential that the air gaps due to the insulation spaces between the particles be at an absolute minimum. There was some question in the early days as to whether these problems might be insurmountable barriers to the development of a successful permalloy-powder core.

Nevertheless, the investigation continued. In the manufacturing department at Hawthorne, research work on manufacturing methods was expedited. Early in 1925, these efforts produced the first modest successes. The engineers at Hawthorne had devised methods of melting and subsequent working of the cast alloy ingot into strips rolled to ¼-inch thickness and thereby made so brittle that the metal could be reduced to powder by standard methods. After the powder was pressed into cores, permeabilities that were significantly higher than those with the iron-powder cores were achieved after moderate heat treatment.

The development of permalloy-powder cores proceeded on a broad front, with the specialists in metallurgy being responsible for the work on powders, and the magnetics research and development people working on the powder cores. It was evident that for high permeability the density of the cores would have to approach that of the metal itself, calling for high core-forming pressures, unavoidable deformation of the particles, and corresponding loss of intrinsic permeability.

The solution appeared to be heat treatment of the cores after their formation to restore the high permeability and the low losses of the permalloy particles. Temperatures required for such treatment were in the order of 500°C, so that insulation that would give only minute separation of the particles and would stand up under such high temperatures had to be found for the particles. Progress came stepwise, and by 1926 a successful process was achieved. As described in 1927, the development required "continuous experimental work over a period of several years before it was brought to a successful issue. It represents the result of highly cooperative efforts embracing research and developments in chemical, metallurgical and electrical fields."

Permalloy-powder cores went into production in 1926, superseding the iron-powder cores. The insulation was sodium silicate, powdered talc, and chromic acid, the last being later replaced, because it was somewhat toxic, by tartaric acid. The powder was 120-mesh, 81 permalloy. Forming pressures were 200,000 pounds per square inch and the cores were annealed in air at 500°C. The permeability of the cores so produced was 75.

A measure of the improvement obtained over the existent Grade "B" iron-powder cores, which are most nearly comparable, is given by the following values: Residual coefficient 37×10^{-6}, hysteresis coefficient 5.5×10^{-6}, and eddy-current coefficient 51×10^{-9}. The corresponding coefficients for the 35 permeability Grade "B" powder cores are 109×10^{-6}, 49×10^{-6}, and 88×10^{-9}, respectively. (The coefficients are defined in the next paragraph.) A lower permeability (26) permalloy-powder core, which was more stable, later followed in production. The superior characteristics of the new cores made possible a reduction in core volume of 5 to 1.

2.4.4 Core Loss Analysis

A major contribution that evolved from the development of powder cores was the precise determination of core losses. Studies had disclosed that the conventional eddy-current and hysteresis concepts were not adequate to describe the losses in the cores. An additional loss, referred to as the residual loss, had to be taken into account. An exact formulation of core losses requiring three terms was developed in 1927 or earlier, as follows:

$$\frac{R}{\mu L f} = c + aB + ef,$$

where R and L are the effective resistance, in ohms, and the inductance, in henries, of the coil (with wire resistance subtracted); μ is the core permeability relative to air, f the frequency in hertz, and B the maximum flux density in gauss; and c, a, and e are the residual, hysteresis, and eddy-current coefficients, respectively. By taking measurements at different frequencies and flux densities, the coefficients can be determined by straightforward analytical or convenient graphical methods.

This breakdown of the loss into its components greatly assisted the course of the development. The eddy-current coefficients could be related to the sizes and shapes of the particles, the resistivity of the alloys, and the effectiveness of the electrical insulation between the particles. The two other coefficients could be related to the magnetic characteristics of the alloys and the heat treatments of the particles and the cores. These coefficients later became the industry standard for characterizing compressed powder cores with precision.

2.4.5 Molybdenum-Permalloy-Powder Cores

While the permalloy-powder development was still under way, work was started on powder cores of permalloy in which a third metal such as chromium was added. The purpose was to take advantage of the higher resistivity of the ternary alloys to lower the eddy-current coefficient. By 1926, molybdenum permalloy (in laminar form) had been developed, having the expected high resistivity and low losses with a permeability of 20,000. Development work on powder cores was therefore channeled toward utilizing these superior properties.

The higher permeability of molybdenum permalloy did not in itself increase the core permeability very much because the myriad air gaps in the core largely controlled its permeability. The higher resistivity and lower magnetic losses characteristic of the material, however, offered significant advantages. The development proceeded with the same thoroughness and in much the same way as that of the 81-permalloy-powder cores. The latter had represented so large an advance over the earlier iron-powder cores that the price of further improvement was a great deal of effort, much of which was in the nature of refinement of methods and processes. In this development, many problems of alloy embrittlement, pulverization, insulation, and heat treatment had to be solved, on both a laboratory and a factory scale. In the course of the work it was found that, by heat treating the cores in a hydrogen atmosphere at temperatures considerably higher than that used for the permalloy-powder cores, much lower hysteresis losses could be realized. At the same time, a new type of ceramic insulation was introduced that was more inert than that used previously. Refinements in insulation methods were made to effect a high degree of compaction and correspondingly higher permeability than before.

The later part of the development work was slowed by the onset of the Depression, and the new molybdenum-permalloy-powder cores, which superseded the permalloy-powder cores, did not go into production until 1934. The composition finally chosen was 2-percent molybdenum, 81-percent nickel, and 17-percent iron, designated 2-81 molybdenum permalloy, used as 120-mesh powder. The cores were heat-treated in hydrogen at 650°C, with a resulting core permeability of 125. The residual, hysteresis, and eddy-current coefficients were 30×10^{-6}, 1.6×10^{-6}, and 19×10^{-9}, respectively, in the order of a two-fold improvement over the previous permalloy-powder cores. The new cores[3] made possible a 2:1 reduction in core volume for loading coils, compared with the earlier permalloy-powder cores, while at the same time giving appreciably better coil performance in stability modulation and "flutter."

2.5 Other Applications

Although the new cores were primarily intended for loading coils, they found immediate and equally important use in coils for oscillators, filters, and networks for carrier systems because of their low losses and low modulation. The introduction of these new cores was very timely, since it coincided with the rapid expansion of carrier system development for open wire lines, cables, and the newly introduced coaxial cables. For carrier filters and network coils, cores with permeabilities of 26 and 14 were introduced. These were similar to the 125-permeability cores except for a somewhat smaller particle size and the dilution of the powder with non-magnetic powder, yielding from 25- to 50-percent-higher coil Qs (Q is the ratio of reactance to resistance and is a measure of coil efficiency).

Later, a series of stabilized cores were made available, in which the permeability changed less than 0.03 percent from 45°F to 108°F, which was necessary in some cases to prevent the shifting of transmission bands of filters due to changes in the inductances of component coils with temperature. The cores were stabilized by adding a small amount

[3] V. E. Legg was one of the principal contributors.

of 12-percent molybdenum permalloy powder, which has a very high negative permeability-temperature coefficient, to offset the small positive coefficient of the regular powder cores. By adjusting the proportion of additive powder, a wide range of temperature coefficients could be provided. The other properties of the cores were not significantly affected.

Inductance coils using the new cores came into large-scale general use in the Bell System and were one of the underpinnings of the carrier technology. In the succeeding 30 years or more, the molybdenum-permalloy-powder cores—virtually unchanged, and widely used in World War II—became the standard of the industry. Successful production of the cores is a highly detailed art, and this art and its patent rights were given to industry. As of the early 1970s, some 20 million cores, nearly half for loading coils, were manufactured by the Western Electric Company, and nearly as many by companies outside the Bell System.

2.6 Magnetic Materials for Continuous Loading

For circuits on land, the loading by use of coils at intervals along the lines is more economical than continuous loading. Coils also are more economical for cables in relatively shallow water, like those crossing wide rivers. On the other hand, the loading of deep-sea cables by the addition of coils presents formidable problems in fabrication, in the laying of the cable with the coils, and in protecting the coils from ocean-bottom pressures. For such cables, continuous loading not only presents a better technical solution but also is more economical.

Theoretical studies of continuous loading were undertaken in the Bell System as early as 1897. Iron-wire-type loading was used, but to a very limited extent, in Europe in the early 1900s, and in the Vancouver cable in 1913. In the Bell System, loading developed along the lines of using spaced coils as the more economical method for land lines. The first application of continuous loading in the Bell System was for the submarine-telephone coaxial cable from Key West, Florida, to Havana, Cuba, placed in service in 1921.

Initially, because of its superior properties high-permeability permalloy tape was to be used for loading this cable. These plans, however, were slowed up by American entry into the war in 1917. After the war, the demand for telephone service to Cuba had become urgent. At that time, the use of permalloy tape for the cables would have entailed the development of cable manufacturing methods, which would have taken considerable time; and, in addition, delays were probable if new materials and new methods were used. Therefore, the time element compelled the use of iron wire for loading.

A single close layer of 0.008-inch iron wire was applied directly over the central copper conductor, resulting in a total inductance of 4.35 mH per nautical mile. The iron wire was covered by insulation, over which concentric copper tapes were wrapped to provide the return circuit. In addition to two-way voice circuits, each cable was intended to carry as many as three carrier-telegraph circuits above the voice range, as well as dc telegraph.

Since the effective permeability of the iron wire was only about 115, permalloy was chosen for the continuously loaded telegraph cable laid between New York and the Azores in 1924 and placed in service in 1925. The method used was to spirally wrap the central conductor with 6-mil, 0.125-inch-wide, 78½ permalloy tape; to heat-treat the wrapped conductor to restore the magnetic properties of the permalloy without damaging the copper conductor; and then to insulate and complete the fabrication of the cable in the usual manner (see Fig. 8-3). The effective permeability of the permalloy spiral was 2,300 and the resulting total inductance of the loaded cable was 54 mH per nautical mile. The capacity of the cable was 1,900 letters per minute, compared with 250 for an unloaded cable. It should be noted that the New York–Azores cable was a Western Union cable, but it was designed and its construction in England was supervised by Western Electric development engineers, and the Western Electric Company furnished the permalloy tape used in it. This cable served as a prototype for many similar transoceanic telegraph cables around the world.

Effort was next directed to the continuous loading of a transatlantic cable for telephone transmission, a formidable engineering project. The technology of the day had not advanced sufficiently to make deep-sea repeaters practicable. For natural speech, the transmission of

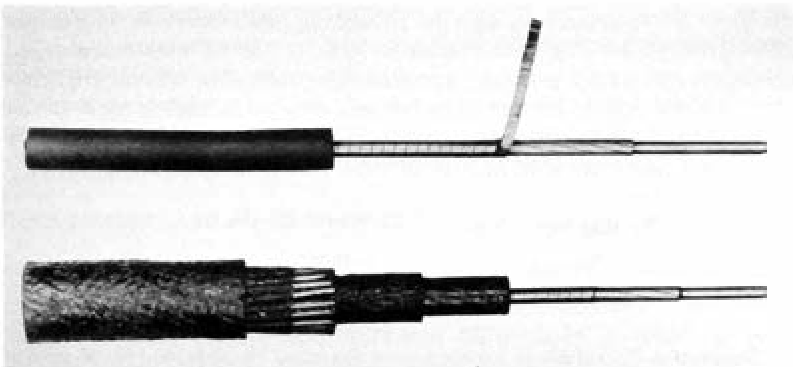


Fig. 8-3. Deep-sea permalloy loaded cable.

a frequency range up to 3,000 hertz is necessary, nearly two orders of magnitude higher than that required for telegraph. Since the attenuation of a cable increases rapidly with frequency, the loss at the highest frequency is extremely high. The cable was planned for use between Newfoundland and Ireland, some 1,800 nautical miles, over which distance the attenuation at 3,000 hertz is 165 dB. It was essential, therefore, that the attenuation of the cable be reduced to an absolute minimum through inductive loading, and that the loading be as loss-free as possible if the weak signal received at the far end of the cable were to be detectable. For this application, the unique properties of perminvar were essential prerequisites, that is, extremely low hysteresis loss and constant permeability.[4] The eddy currents were to be minimized by using the perminvar in the form of very thin tape. Voice switching was to be used, which meant that at both ends the receiving apparatus was to be normally connected to the line; then each speaker's voice currents would switch his end of the line to the transmitting apparatus.

The design was completed in the late 1920s. The loading consisted of four spiral layers of 1-mil-thick perminvar tape and the return circuit consisted of copper tape, as in the Havana–Key West cables. Because at that time the only cable manufacturer was in Europe, a task force of Bell Laboratories people was stationed in Germany for a period of two to three years, beginning about 1930, to assist in the manufacturing development, oversee the production, and test the completed cable.

From an engineering standpoint the project was a success. A 20-mile section of the manufactured cable was laid in the deep-water part of the Bay of Biscay, and tests made on this sample section were satisfactory in all respects. Unfortunately, at about that time the Depression began. The work was at first slowed down and then the laying of the entire cable was finally abandoned for lack of funds. Work on the loaded transatlantic telephone cable was never revived. In the late 1930s, the possibility of a carrier system on cable with submerged repeaters was being seriously considered, promising several channels at a cost not much greater than for the single-channel loaded cable. After World War II, the repeatered transatlantic cable was successfully developed and went into service in 1956.

2.7 Magnetic Testing Equipment

The extensive programs for developing magnetic materials were facilitated by measuring equipments specific for that use. Main

[4] The perminvar composition considered was 47-percent Ni, 25-percent Co, 20-percent Fe, 7.5-percent Mo, and 0.5-percent Mn. The addition of the molybdenum increased the resistivity.

dependence originally was on the standard ballistic methods of measurement. Such methods, however, were not sufficiently sensitive for measurements at very low magnetizing forces, corresponding to those in most communication circuits. For such measurements, the inductance bridges developed in the 1900s in the Engineering Department of the Western Electric Company proved to be a valuable tool. They provided an easy and far more precise method than the ballistic galvanometer for measuring permeabilities at low flux densities.

In 1911, W. J. Shackelton, then a young engineer, was given the task of improving these bridges, which had lacked precision and often gave erratic results, to meet the need for increased accuracy for the advancing communication art. Through study of the circuitry, Shackelton discovered the sources of instability and inaccuracy and redesigned the bridges, using skillfully disposed electrostatic shielding of the bridge arms. Work was continued over the years to refine and extend the frequency and inductance ranges of these bridges to keep pace with the magnetic material, coil, and transformer development work.

An important instrument specifically developed in 1920 for magnetic tests was the Kelsall[5] furnace permeameter. With this instrument, samples of magnetic materials could be tested at temperatures as high as 900°C. The furnace permeameter depends on the transformer principle for its operation (see Fig. 8-4). The furnace itself is a thermally insulated torus within which the toroidal sample is placed. The sample is heated by wires adjacent to it. A high-permeability toroidal core wound with many turns of fine wire serving as a primary is located coaxially with the furnace and close to it. A hollow copper rod extends through this coil and the furnace, with the rod, the top and bottom copper covers, and the copper outside shell functioning as a single turn linking the specimen and the wound core. The magnetic properties of the specimen, by transformer action, can be determined by measurements made on the primary winding described above.

The furnace permeameter, in addition to determining the magnetic properties of alloys at different temperatures, made it possible to measure these properties accurately as the temperature of the sample was raised or lowered through the transition region, giving important insight into the metallurgical structure of magnetic alloys.

III. COMMUNICATIONS TRANSFORMERS

3.1 The Early Years

The first transmission of speech by telephone, the famous words of Bell to Watson, took place in March 1876. Even in that earliest

[5] Named after its inventor, G. A. Kelsall, a Bell Laboratories engineer.

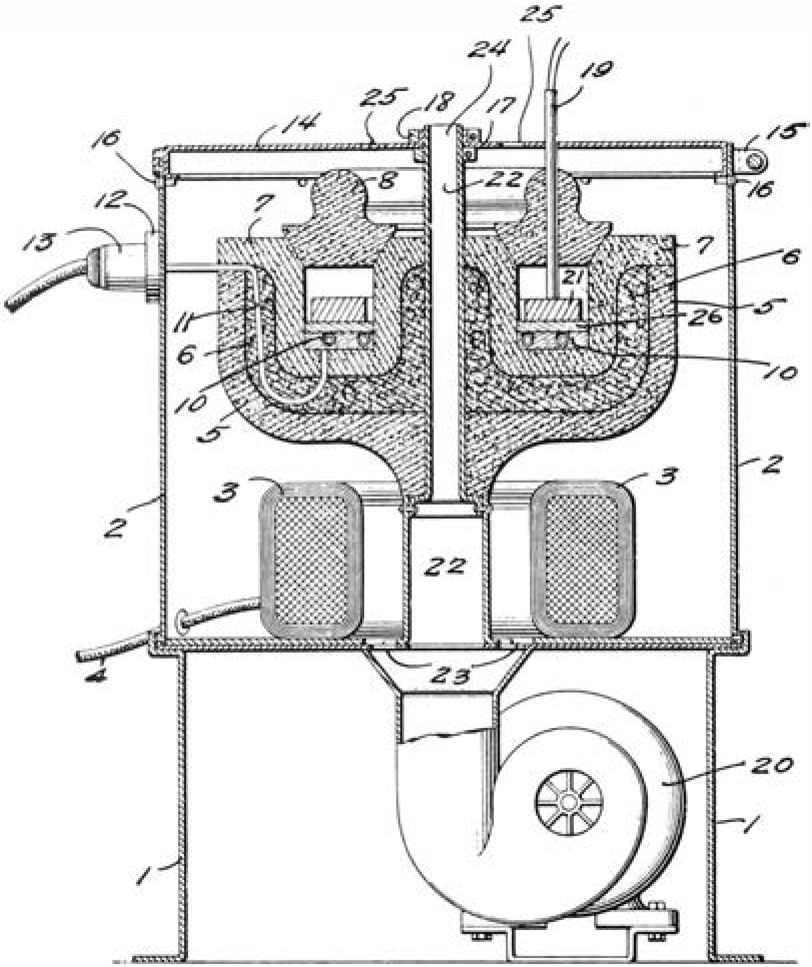


Fig. 8-4. Kelsall furnace permeameter. (From U. S. Patent No. 1,475,438)

telephone circuit, the communications transformer played an integral part. Only about a year later, in October 1877, a patent application was filed on the use of a transformer to interconnect the battery, transmitter, receiver, and line. In 1879, the Western Electric Company marketed the "Universal Switch," an elementary form of switchboard employing a transformer in the operator's telephone circuit. A year later, in 1880, a patent was issued to Watson[6] covering the use of a

[6] T. A. Watson; U.S. Patent No. 232,788; filed July 10, 1880; issued September 28, 1880.

transformer to connect a two-wire metallic line circuit to a subscriber's grounded circuit to minimize the effect of noise picked up by the line.

In those days transformers were called induction coils. Some years later the term repeating coil came into use; the coil was said to repeat the signal by induction from one circuit to the next. In the course of time, the term induction coil became restricted and applied only to transformers used in telephone sets; the term repeating coil was applied to transformers used to interconnect lines and, until about 1913, for all other purposes.

3.1.1 Early Induction Coils

It is not surprising that the transformer found such early use in the subscriber's telephone set (see Chapter 3, Section 3.1). In the first telephone set, the battery, transmitter, receiver, and line were all connected in series. Since the resistance of the transmitter was very small, the variations in this resistance with the talking sound waves resulted in only minute changes in current in the circuit. By connecting the battery and transmitter to a primary winding of an induction coil, and the receiver and line to a secondary winding having a much larger number of turns, two advantages accrue: a larger direct current flows through the transmitter, giving higher current changes with impressed sound waves; and, by transformer action, relatively high voltages are impressed on the line, resulting in a far more efficient set (see Fig. 8-5). The early induction coils consisted of a bundle of straight iron wires in a wooden spool wound with a primary and a secondary winding. This form of construction was quite inexpensive, gave adequate performance, and therefore remained unchanged in essential features well into the 1900s (see Fig. 3-33).

3.1.2 Early Repeating Coils

Repeating coils with center taps found early use in circuits for simultaneous telephone and telegraph operation, called "simplexing." They were also used in a common-battery exchange installed in Lexington, Massachusetts, in December 1893. There the coils[7] kept the various voice circuits separated while simultaneously supplying direct current from a common-battery source to all the subscribers. Repeating coils later found a very important use in phantom circuits, where they made possible three telephone conversations over two pairs of wires on an efficient basis. Figure 8-6 shows the arrangement of the windings of such repeating coils, as covered in a patent granted to J. J. Carty in 1886.[8] However, the design of repeating coils for adequate

[7] Later called "battery-supply repeating coils."
[8] J. J. Carty; U. S. Patent No. 348,512; filed February 2, 1886; issued August 31, 1886.

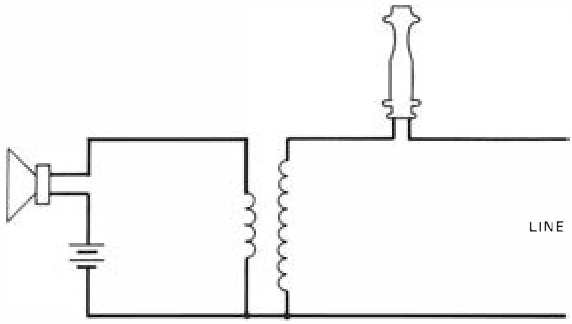


Fig. 8-5. Telephone circuit of 1878.

balance was not understood at that time, and phantom circuits remained something of strictly scientific interest for many years.

3.1.3 Early Phantom Repeating Coils

It was not until 1902–1903 that a phantom circuit was put into commercial operation between Lewiston, Maine, and Berlin, New Hampshire, initially using solenoidal repeating coils. H. S. Warren had established the design of loading coils on a more scientific basis in 1901, with their construction consisting of copper wire wound on toroidal cores made up of fine insulated iron wire. The design

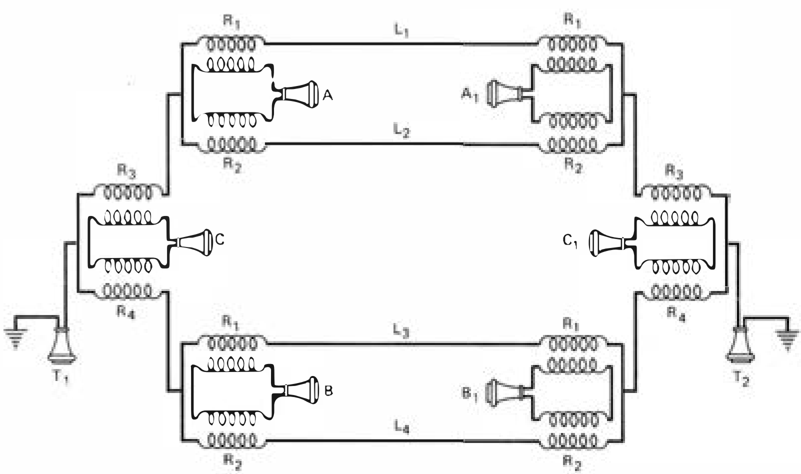


Fig. 8-6. Phantom circuit showing use of repeating coils. (Redrawn from U. S. Patent No. 348, 512)

concepts developed by Warren for loading coils were useful also for repeating coils. Late in 1903, repeating coils of the new toroidal type replaced the coils initially used in the first phantom circuits, and they "showed great improvement" in circuit performance. Shortly thereafter, in 1904, the 37A repeating coil was designed, characterized as the "first satisfactory" phantom repeating coil (see Fig. 8-7).

3.1.4 Hybrid Transformers

One form of the communications transformer that early performed a unique and very valuable function in telephony is the hybrid repeating coil or transformer, in later years called the bridge transformer. The concept of the hybrid transformer is old, antedating 1890, and the use of such a transformer in repeaters was covered by a Bell System patent filed in 1902. The first successful electromagnetic repeater, called the Shreeve "telephone relay," came into use in 1904 and depended on hybrid transformers for its operation. The hybrid transformer construction that evolved for this purpose was similar to

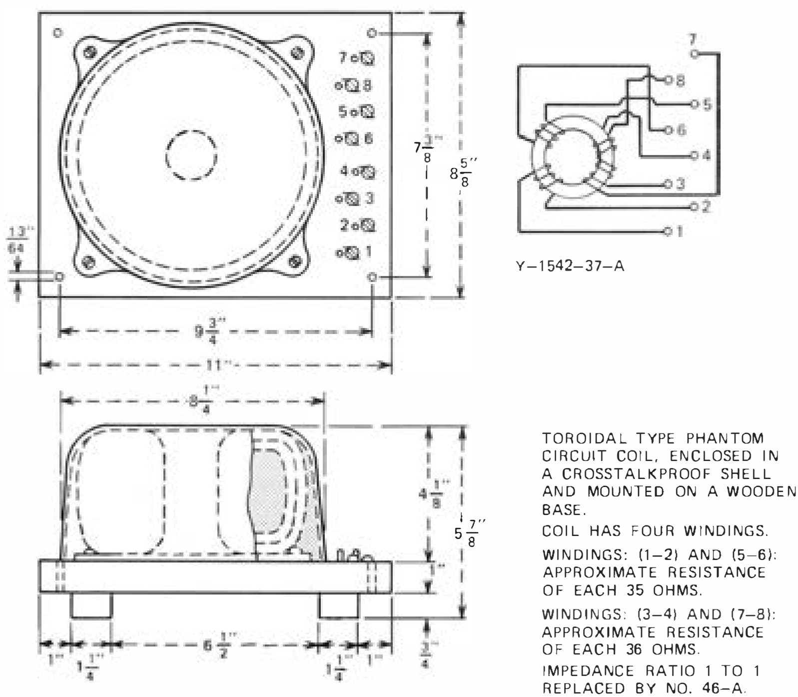


Fig. 8-7. No. 37A repeating coil. (Redrawn from Western Electric Apparatus Catalog)

the repeating coils of that period. An amplifier is inherently a one-way device, receiving weak currents at the input terminals and transmitting amplified currents at the output. The hybrid transformer makes it possible to use such a one-way amplifying element for two-way transmission, as in the 21-type repeater circuit,[9] which came into use before 1911, and the 22-type repeater circuit[9] in 1914.

3.2 Distinctive Features

Even in the early years the communications transformer had little design resemblance to its progenitor, the power transformer. The communications transformer usually operates at milliwatts where heating is not a factor, instead of kilowatts as in the power variety; millivolts instead of kilovolts; the magnetic materials at a flux density of a few gauss instead of kilogauss; and over a band of frequencies instead of at a single frequency.

A property of the transformer that accounts for much of its use in telephony is that of providing an impedance match. A transformer may be used, for example, to connect a cable having a characteristic impedance of 130 ohms to an open wire line of 600 ohms. If the proper turns ratio is used in the transformer, the cable and open wire line impedances are matched, which means that signals are transmitted smoothly in either direction, without reflection, just as if the circuit consisted entirely of a cable or an open wire line.

3.3 Transformer Equivalent Circuit

At this point in the discussion, a brief reference to the equivalent circuit of the communications transformer is made to help the reader understand the description of the various designs that follow.

From the transmission standpoint, the communications transformer may be represented exactly by the equivalent circuit of Fig. 8-8. In the circuit shown, the T network is the familiar unity ratio T of a transformer, with Z_A and Z_B related to the primary and secondary leakages and dc resistances respectively, and with Z_M the impedance due to the magnetic core. C_1 and C_2 are the primary and secondary winding capacitances, and the dashed lines enclose an idealized transformer whose ratio is equal to the turns ratio of the actual transformer. Thus, the various elements of the equivalent circuit can be related to the various components of the physical transformer and its performance calculated in advance. Z_M is usually relatively large and influences only the low end of the band, and Z_A, Z_B, C_1, and C_2 are all relatively small and influence primarily the upper end of the band.

[9] See Section 4.2.4, Chapter 4.

3.4 Input Transformers—Early Designs

The development of vacuum tube amplifiers around 1912 called for a new and different kind of transformer. The input or grid circuit of the vacuum tube has an exceedingly high impedance; consequently, higher amplification could be gained in a repeater simply by increasing the number of secondary turns, as may be seen from Fig. 8-8. Such transformers working into a grid circuit are called "input transformers." Where the application requires a resistive impedance viewed at the primary winding of the transformer, a very high resistance is usually placed across the secondary winding, and the "matched impedance" principle described previously applies.

In the toroidal construction used at the time for repeating coils, only No. 31 gauge wire or heavier could be wound with the existing winding machines, severely limiting the number of secondary turns

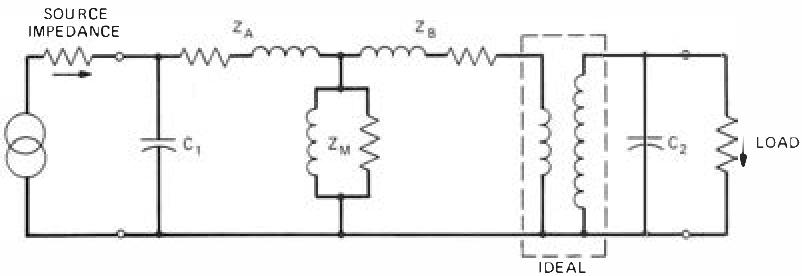


Fig. 8-8. Exact equivalent circuit of communications transformer.

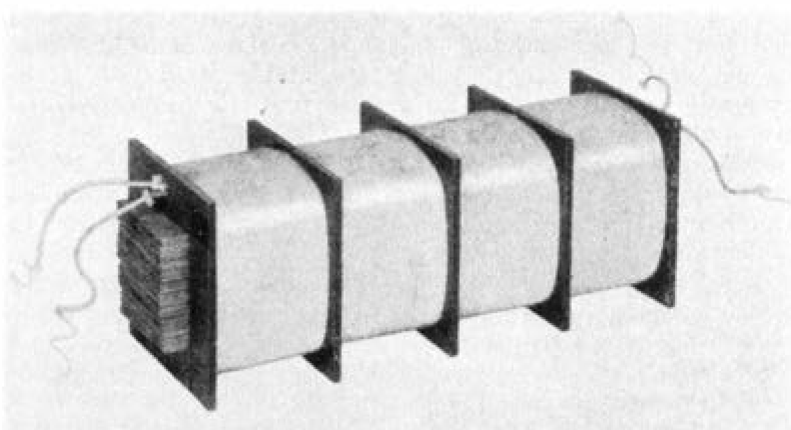


Fig. 8-9. Straight silicon-steel transformer core design of 1913; its size is 9 by 3 inches.

that could be applied, and placing a ceiling on the step-up and consequent amplifier gain that could be obtained. This limitation led in 1913 to a new type of design, whose construction is illustrated in Fig. 8-9. The magnetic core consisted of 14-mil silicon-steel strips. The spool-type windings permitted very fine wire to be used so that a large number of secondary turns giving high step-up and high amplifier gain could be obtained. The spool was arranged in sections to minimize the secondary winding capacitance which, referring again to Fig. 8-8, might otherwise have caused attenuation of the higher frequencies in the voice transmission band with consequent loss of distinctiveness in the transmitted speech.

The external field of the 1913 transformer proved troublesome and led to the closed magnetic circuit designs of 1914 and 1915, shown in Figs. 8-10 and 8-11. These also used 14-mil laminations and, as may be seen in the illustrations, borrowed construction details from the power transformer art. The advanced state of the technology at that time is illustrated by laboratory reports written in 1915 and early 1916. These reports give a sophisticated discussion of an experimental design having the very high impedance ratio of 2,000 to 600,000 ohms. The loss and reflection coefficient at 400 hertz were 0.5 dB and 12 percent, and at 2,000 hertz, they were 0.25 dB and 3.6 percent, respectively, which is good performance even by modern standards. The desire for a smaller and cheaper construction led briefly to a reversion to the toroidal type in 1916, using, however, annular laminations of 14-mil silicon steel, as shown in Fig. 8-12.

As a result of continuing development, the toroidal design was replaced in 1918 by the two types shown in Fig. 8-13. These designs were much smaller and cheaper than the toroidal and at the same time had higher step-up. The 201-type was introduced in 1919 in the transcontinental lines because it was small, relatively inexpensive, and gave superior performance.

The course of input transformer development culminated in 1923 in the 234-type shown in Fig. 8-14. This design was adopted to attain low manufacturing cost, while retaining the performance properties of its predecessors. The core was designed for automated production; windings could be applied very rapidly, and the assembly of the transformer was a simple operation. The continuing improvement in performance of the described designs is shown in Fig. 8-15, each design being identified by the year of its introduction. In all cases, the windings were vacuum-dried and impregnated in rosin-rosin oil; whenever a high degree of moisture protection and long life were needed, they were potted in rosin in cast-iron cases.

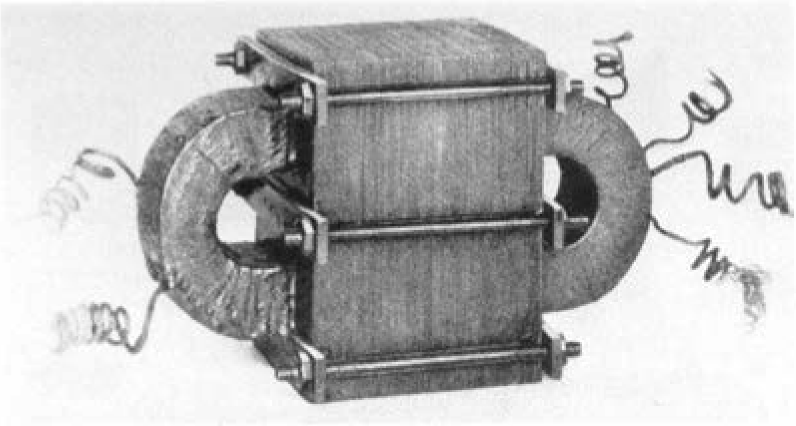


Fig. 8-10. Shell-type design of 1914; its size is 6⅝ by 4 inches.

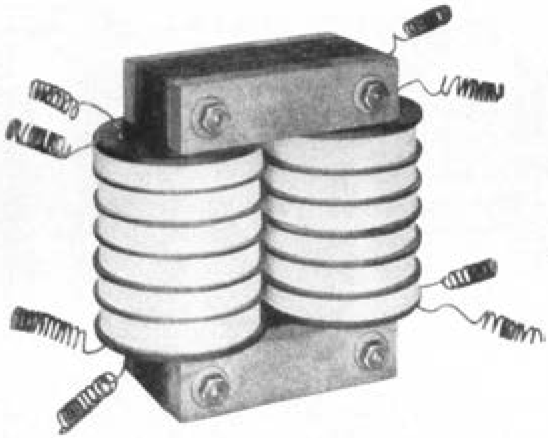


Fig. 8-11. The 1915 transformer core design; its size is 6¾ by 4⅝ inches.

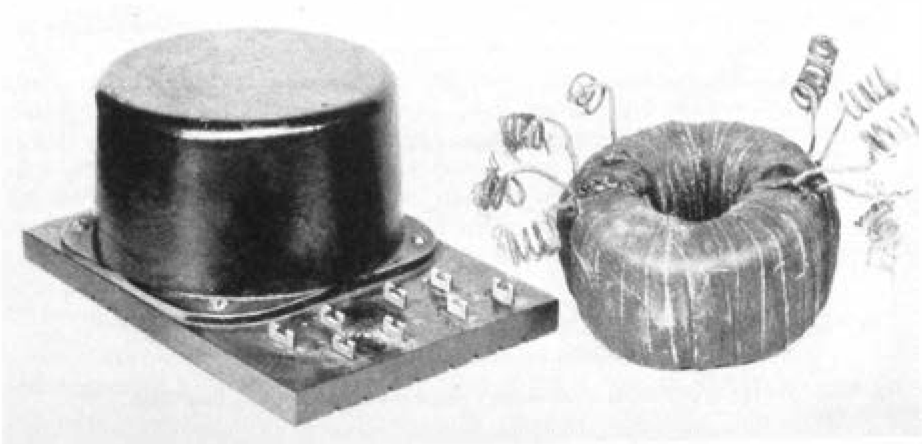


Fig. 8-12. Toroidal transformer, mounted on a block; its size is 6 by 4⅝ inches.

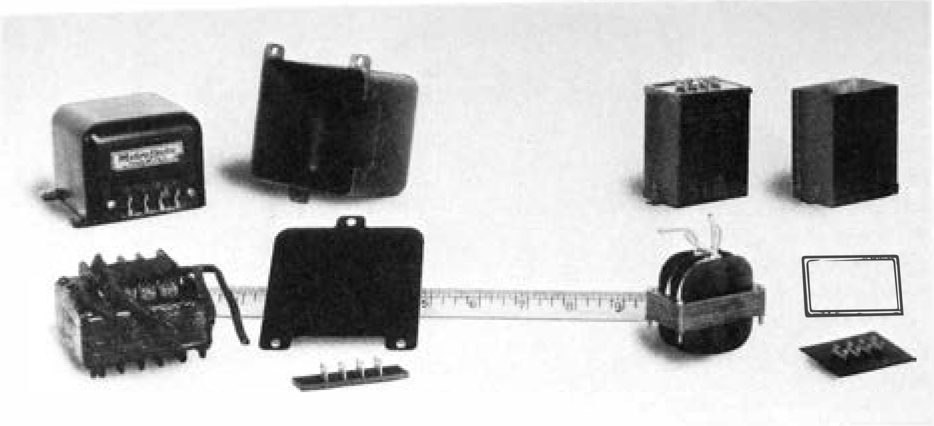


Fig. 8-13. 201- and 213-type transformers.

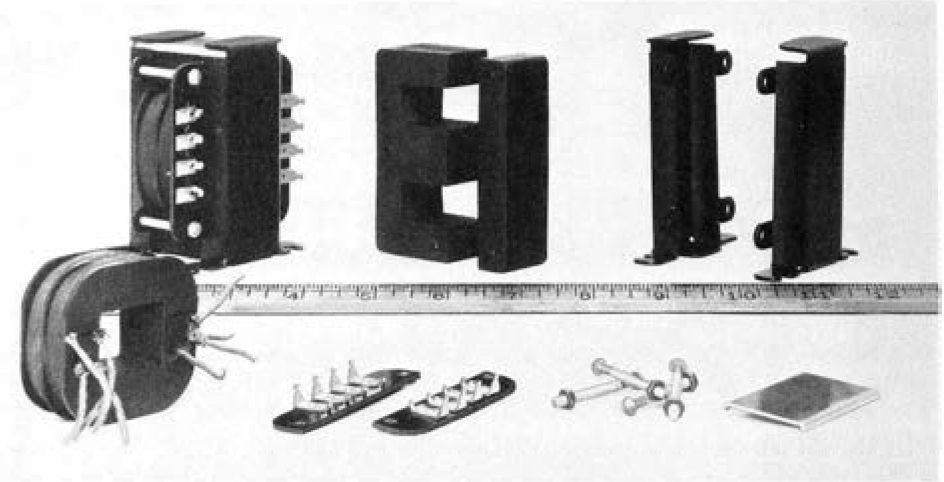


Fig. 8-14. 234-type input transformer (1923).

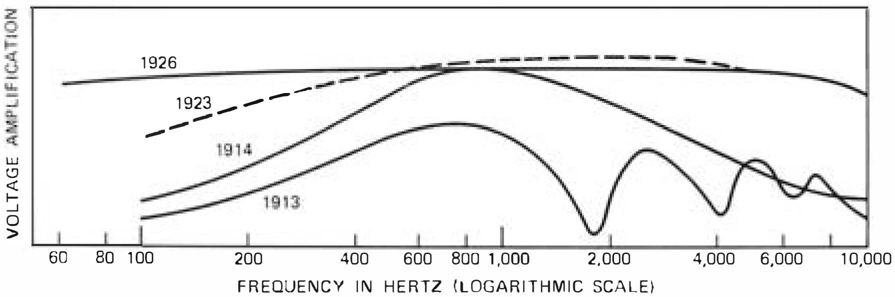


Fig. 8-15. Voltage amplification characteristics of transformers. (Redrawn from Field 1926, p. 34)

3.5 State of the Art in the Early 1920s

A review of the state of the art as it had developed up to the early 1920s reveals that a number of factors, principally work on the new carrier systems, were coming into play at about that time. These would greatly affect the course of development work on communications transformers.

3.6 Types of Construction

For input transformers, where minimizing secondary capacitances and maximizing step-up were important, the 234-type transformer was the common type of construction used, with modifications consisting of longer or shorter legs for the "E"-shaped cores, according to the needs of the design. Where small size was at a premium, the 201- and 213-types shown in Fig. 8-13 were used. Similar constructions were used for interstage transformers, which operated between the plates and grids of vacuum tubes in amplifiers. For output transformers (i.e., transformers working out of the plate circuits of vacuum tubes), the same construction was used as for input transformers, since the plate impedances, while not so high as the grid circuit impedances, were still many thousands of ohms, and minimizing high side-winding capacitances was still a factor in the design. Air-gap spacers were used in the cores of the output transformers to minimize the effect of the plate direct currents which tended to bias the magnetic material in the cores and seriously reduce the inductance.

For low-impedance applications, as in connecting lines and cables, the toroidal type of construction (shown in Fig. 8-12) was used. At the low-impedance levels of the circuits, the inherently high capacitances of the windings could be tolerated. Furthermore, the inherently close coupling of the windings and correspondingly low leakage inductances permitted designs with wide frequency bands and minimum transmission distortion.

In the early 1920s, the invention of 78½ and 45 permalloy had a direct and important influence on the design and construction of transformers. They alloys contained 78.5-percent and 45-percent nickel, respectively, the balance being iron. The application of 78½ permalloy, which had an initial permeability of 10,000 compared to 400 for silicon steel, made possible designs of 10 to 25 times the previous frequency ranges, as well as other improvements. The 45 permalloy was superior where there was superimposed direct current.

3.7 Status of Transformer Theory

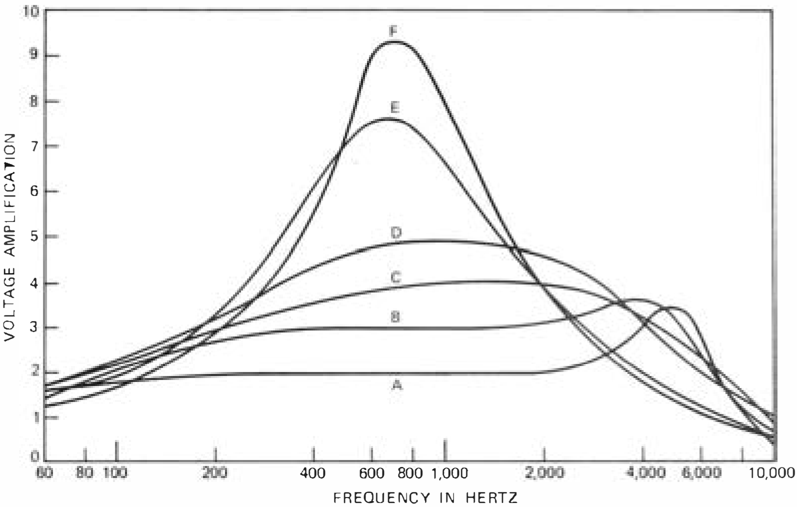
The evolving communications transformer art was accompanied by and supported by an evolving theory. Out-of-hours courses in which the theory was developed and enlarged had been in progress in the Western Electric Company Engineering Department since 1919. Much of the new material was converted into a text by K. S. Johnson during the years 1920–1924, and it was included in his landmark published book, *Transmission Circuits for Telephonic Communication.* At about that time, H. Whittle, whose background experience was in filters and networks, was transferred to the transformer area. He brought the network approach to transformer design, where it proved to be a very effective tool.

By the early 1920s, the principles of transformer design were well established. Specifically, the roles of transformer leakage inductance and of the distributed capacitances of the windings in shaping the transmission characteristics of the transformer were established art (see Fig. 8-16). A body of information concerning the characteristics of typical windings—the permeability and losses of the various magnetic materials as used in the different core structures, as a function of superimposed direct current, air gaps, and frequency—had been made available to the designer. The technology of the use of twisted pairs, copper and tinfoil electrostatic shields, and balanced winding methods were all familiar (see Fig. 8-17) and transformer performance could be calculated in advance, instead of relying on "cut and try." Nearly all designs were in the voice-frequency range, but a few had already been designed in the carrier range. Reliability was a way of life, so transformers rarely failed.

The communication principles and theories, of which transformer theory was a part, so highly developed in the Engineering Departments of the Western Electric Company and AT&TCo, had little counterpart outside the Bell System. Courses on the subject were not taught in the universities. The unique contribution of the Bell System in this field is attested to by the series of lectures given on communication circuits, including transformer theory, at Harvard University during the winter of 1921–1922 at the invitation of the Department of Electrical Engineering.

3.8 New Requirements for High-Frequency and High-Quality Audio Circuits

Around 1922, a number of factors came into play that had a large bearing on the course of transformer development in the succeeding years. The most direct and important influence on design was the application of 78½ permalloy, a newly invented magnetic material with an initial permeability of 10,000, compared to the 400 permeability of silicon steel. Permalloy made possible transformer designs of 10 to 25 times the previous frequency ranges, as well as other improvements. Another major factor was the rapid increase in the develop-



VOLTAGE AMPLIFICATION CHARACTERISTICS OF INPUT TRANSFORMERS OF VARIOUS TURNS RATIOS OPERATING FROM 20,000–OHMS RESISTANCE INTO A 216A VACUUM TUBE

COILS	TURNS RATIO
A	1:2
B	1:3
C	1:4
D	1:5
E	1:8
F	1:10

Fig. 8-16. Voltage amplification characteristics of input transformers of various turn ratios.

ment of new carrier systems, which had its beginnings before 1920. Carrier systems began to be proliferated with transformers of all kinds and varieties, covering new frequency ranges and meeting new kinds of requirements.

The most valuable contribution of the transformer development work was the support given to the systems, amplifier, and other circuit development activities. A great deal of skill was directed to tailoring transformer designs for best circuit performance. The measure of worth of a design was the new kind of, or the improved performance of, the circuit that the transformer made possible. Only a modest proportion of the effort was devoted specifically to developing new types of transformers; it was the challenge of new and more difficult design requirements that provided the main impetus for new transformer developments. Literally thousands of new designs were originated in the next decade.

The 1920s saw a large upsurge of work on "by-products" or "specialty products" in Bell Telephone Laboratories. Under this

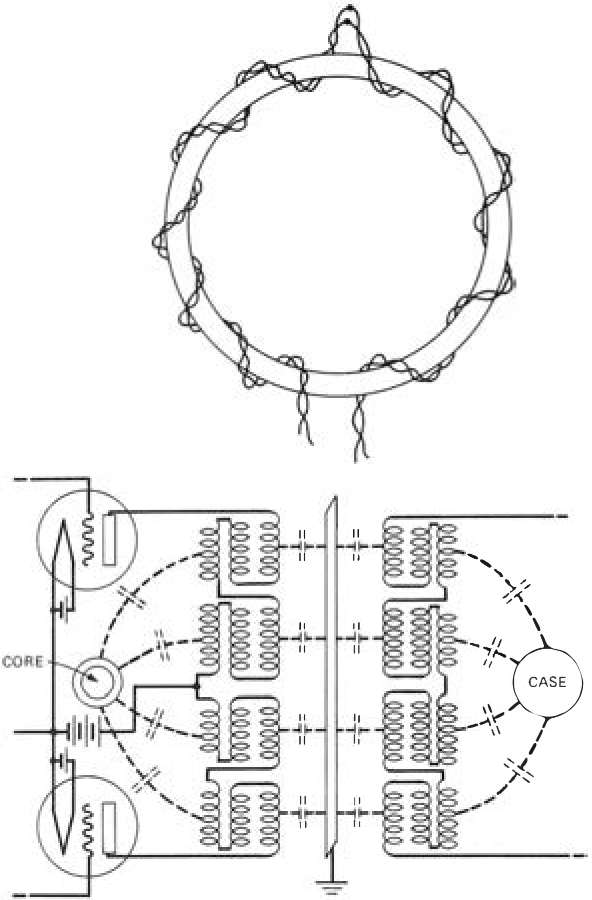


Fig. 8-17. Winding arrangements for inductive and capacitive balance.

heading came radio receivers for home and aircraft, broadcasting equipment, public address systems, and recording and reproducing systems for sound movies. The transformers for such applications were predominantly audio-transformers. This design work depended largely on the materials, construction, and technology that evolved in the development of transformers for the Bell System. Some of the designs for specialty products were unique for that period. Radio broadcasting required the transmission of wide frequency bands, 30 to 7,000 hertz. By taking advantage of permalloy and interleaving the windings, transformers that provided such a wide range were designed in 1924.

The 1920s saw radio broadcast programs transmitted over long-distance telephone circuits, the so-called radio networks. The high-fidelity requirements for music called for the transmission of much wider audio-frequency bands than those for voice. It became necessary to control the phase shift in transformers to minimize phase delay distortion (i.e., the propagation of some frequencies at a faster rate than others).[10] Delay distortion of the transformer was usually greatest at the low end of the band. Here also, introduction of the high-permeability 78½ permalloy assured high transformer inductance, which reduced the delay and made it feasible to meet the requirements.

Many transformers in the 1920s were designed for carrier systems at frequencies upwards of 100 kHz; and for other purposes, transformers were being designed for operation at increasingly higher frequencies.

IV. INDUCTORS (RETARDATION COILS)

4.1 Introduction

The terms "loading coil," "choke coil," and "retardation coil" all were applied to coils whose primary function was to introduce inductance into a circuit. Loading coils were regarded as a separate class because of their unique use and their economic importance. The term "choke coil" in Bell System terminology was used for coils in dc power circuits, where the inductance served to keep noise out of the telephone circuits. The design and construction employed in such power chokes resembled power transformers rather than the communications transformers previously described.

The retardation coils discussed in this section, called inductors in present-day terminology, found their principal use as components of filters and networks where the requirements are for very precise inductance values and high *Q*s. They had earlier and equally important uses in telephone station sets, switchboard circuits, battery supply, and repeater circuits. With the expansion of carrier systems from 1920 on, filter and network activity was greatly stimulated, and the design of the component inductors began to be recognized as a field distinct from loading coil work.

4.2 Retardation Coil Development

Inductor technology as it evolved around 1920 was represented by four basic types of designs. The wood-core toroid was commonly used. It was wound with silk-covered wire consisting of strands individually insulated with enamel to reduce eddy-current losses. The

[10] Phase delay equals $\frac{d\phi}{d\omega}$ in seconds, where ϕ is phase shift in radians, ω equals $2\pi f$.

wire was manufactured by the Western Electric Company to exacting specifications. The calculation of such losses for simple shapes had been mathematically established, and the adaptation of the formulas to the toroidal shape, together with data accumulated experimentally, was worked out to provide a good basis for design.

In the toroidal construction, the magnetic flux is ideally confined within the windings and core, and it is true in practice that the external field is quite small. This is a distinct advantage in the assembly of the components in a filter. The winding capacitances were reduced by applying the windings in two or four sections, and "bank windings" were used to further minimize capacitances. A typical coil is described as being wound on a toroidal wood core, 3 inches in outside diameter, having inductance of 0.054 henry and a Q of 14.5 at 3,000 hertz. Such inductors had Qs of 50 or higher at 30 kHz, the highest frequency then in use. The wound coil, covered with cotton tape, was impregnated in rosin-rosin oil, dipped in Western Electric Superior compound (asphaltic), but not potted. The coil was provided with leads of flexible conductor. Wood-core toroids were principally used at the higher voice frequencies and in the carrier frequency range. They were largely displaced after 1925 by solenoidal air-core inductors and permalloy-powder toroids, which had higher Qs.

Toroidal inductors with iron-powder cores were extensively used in audio-frequency filters and networks. The most common type of core was made of type "C" iron powder, with a permeability of only 26 but the advantage of low losses. The iron-powder cores were replaced in the mid-1920s with permalloy-powder cores, which had much lower losses. The toroidal type of construction placed a practical limit on the inductances obtainable because wire finer than 31 gauge could not be used in the existing machines. For higher inductance values, the practice was to break the powder cores in half, insert the halves in one or two wound spools, and then cement the butting surfaces of the core halves together to reform the core. This construction, developed in 1922 or earlier and employed, for example, in the 94-type retardation coils (see Fig. 8-18a), permitted the finest wire to be used. The air gap in the core could be varied during manufacturing for small inductance adjustment. The stray magnetic field was greater than in the toroidal construction and the coils were therefore often potted in, and "shielded" by, copper cans. In the early 1930s, the permalloy-powder cores were replaced by those of molybdenum permalloy, having lower losses and yielding inductors with even higher Qs.

In 1924, the 119-type retardation coil was just beginning to come into use. It consisted of a wound thermoplastic spool potted in a thermoplastic case (see Fig. 8-18b without the square loop), and it used the stranded wire previously described. This type of construction gave values of Q considerably higher at carrier frequencies than did the toroidal construction. However, the stray magnetic field was high, requiring care in the mounting arrangement of the coils in the filter, to be done at right angles or with axes parallel but displaced to avoid mutual coupling. The 119-type was manufactured in large quantities for many years.

4.3 Adjustable Precision Inductors

In carrier communication circuits, bandpass filters were employed at both ends of the line to separate the channels. Since these channels were placed within narrow limits in the frequency scale, the cutoff frequencies of the filters that pass these channels had to be precisely determined. The cutoff frequencies depended on the values of the capacitances and the inductances comprising the filters, requiring that these values be held to very close limits.

The capacitors could be held to about 0.3-percent capacitance, and the inductors were adjusted by taking off single turns after winding. Subsequent potting and the location of the inductors with reference to the shielding used in the filters added to the variables, so the possible error in inductance when the inductor was mounted in place could be 1.8 percent. This variation could cause frequency shifts of the filter that would be far greater than permissible. To solve this problem, an adjustable inductor was invented around 1927.

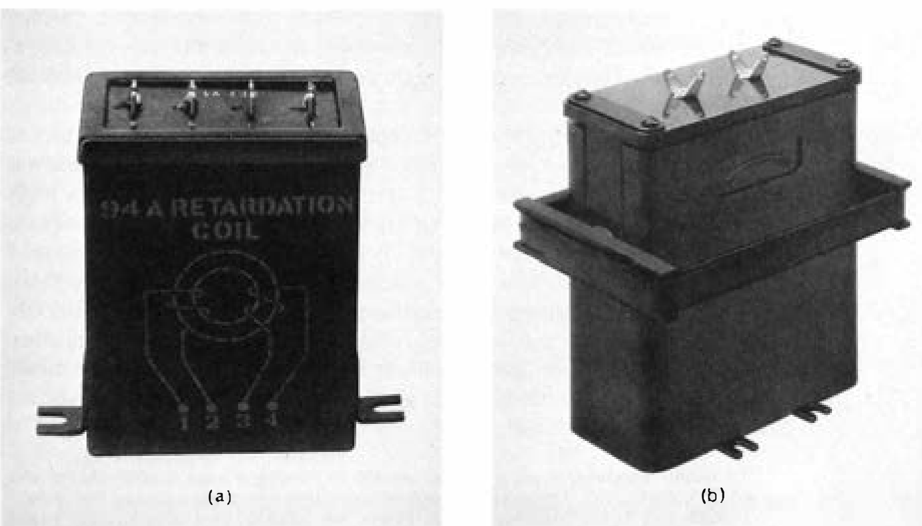


Fig. 8-18. (a) No. 94A retardation coil. (b) Adjustable precision inductor.

Figure 8-18b shows a solenoidal inductor with a square loop, which contains a short-circuited winding, around the case. In the horizontal position, the loop cuts no lines of magnetic force and so does not affect the inductance of the inductor. As the loop is turned near end up, the eddy currents in it reduce the inductance. A 4-percent reduction in inductance was obtained by turning the loop up to 40 degrees. The loop could be set to within one-half degree and the corresponding precision of inductance adjustment was 0.012 percent near the horizontal position and 0.1 percent in the upper position. The adjustment of the inductor was made in the filter so that the variation in the associated capacitor value, the effects of shields in the filter on the inductor, and the variations in the inductor itself could all be corrected by a single adjustment. In this way, the filters could be held to ±0.2 percent in frequency, which was a five-fold improvement over the values obtainable previously.

V. CAPACITORS

5.1 Early Applications

The use of capacitors[11] in the telephone system was started soon after the invention of the telephone. In 1878, Prescott mentions the use of "condensers consisting of alternate sheets of tinfoil and paraffined paper" to pass voice currents around the coils of the ringer that was in series with the line. Another early application was for separating "galvanic" and "derived" circuits on a common line so that, for example, it could be used for both telephony and telegraphy.

The number of capacitors increased greatly with the advent of the common-battery system, in spite of the fact that the units available at the time were bulky and expensive.

Black and Rosebrugh's patent (1879)[12] describes the capacitor as tinfoil layers separated with layers of mica, gutta percha, or paraffined paper. These layers required hand stacking, sometimes done at a high temperature to maintain the paraffin in a molten state, which was a slow and tedious process for capacitors of the order of 0.1 microfarad or more. Furthermore, since the insulating layers were relatively thick and some form of clamp was necessary to hold the foils and dielectric in intimate contact, the assembly was large in size. This situation existed until 1897 when Lee, Westcott, and Robes of the American Bell Telephone Company developed and patented an improved wound

[11] The name "condenser" for a device capable of storing a large electric charge was applied by Volta in 1782. The name "capacitor" came into general use during the 1940s.

[12] G. Black and A. M. Rosebrugh; U. S. Patent No. 212,433; filed June 4, 1878; issued February 18, 1879.

capacitor in essentially its present pressed and elliptical form.[13] Kingsbury states that the production of the foil in continuous rolls, which was a practical necessity for wound capacitors, was a considerable part of this development. He also indicates that this method of construction reduced capacitor costs by from 10 to 20 times. The rapid increase in the need for capacitors is attested to by Mansbridge's statement that the total output of European and American manufacturers was 5 farads in 1907. He attributes this demand to their use in common-battery telephone systems.

5.2 Paper Capacitors

The construction of paper capacitors underwent only minor changes from the time of the Lee, et al., patent until about 1930. The Western Electric 5-type capacitor followed the patent description, but when the 21-type was first made, the flat, wound unit was folded into a closed U. These two types are shown in Fig. 8-19.

The paper used in the early capacitors was a rag stock that was 0.001-inch thick. In 1905, the 21-type was made with one layer of 0.001 and one layer of "thin" paper (probably about 0.00065-inch thick). A purchase specification dated 1910 specified a thickness of 0.00065 inch. The thinner paper initially became available in Europe and there can be little doubt that the telephone industry was responsible for the pressure on domestic suppliers to make a similar product. Subsequently, 0.0005-inch material became available, but it was not until after 1925 that adequate quality was obtainable in 0.0004-inch and thinner papers. Throughout this period, capacitor paper was the subject of many laboratory studies that had as their object the attainment of higher dielectric strength with a reduction of capacitor volume. Large quantities of paper were purchased in Europe, but during World War I, supply was cut off and competition for rags increased so that rag stock paper was difficult to obtain. This led to a more extensive investigation of papers, including one made of a mixture of rag stock and wood pulp. The results were, in most cases, disappointing, and it was not until much later that wood pulp (Kraft) papers with adequate properties were produced and used.

The foil used in the early capacitors was commonly referred to as "tin" but actually it was a high-lead and tin alloy. Initially, the thickness was of the order of 0.0006 inch. A foil approximately 0.0003-inch thick became available in 1904, with a higher proportion of tin in its composition; it was more costly, so the thicker foil continued in use for a number of years.

[13] J. C. Lee, W. R. Westcott, and E. C. Robes; U. S. Patent No. 575,653; filed July 8, 1896; issued January 19, 1897.

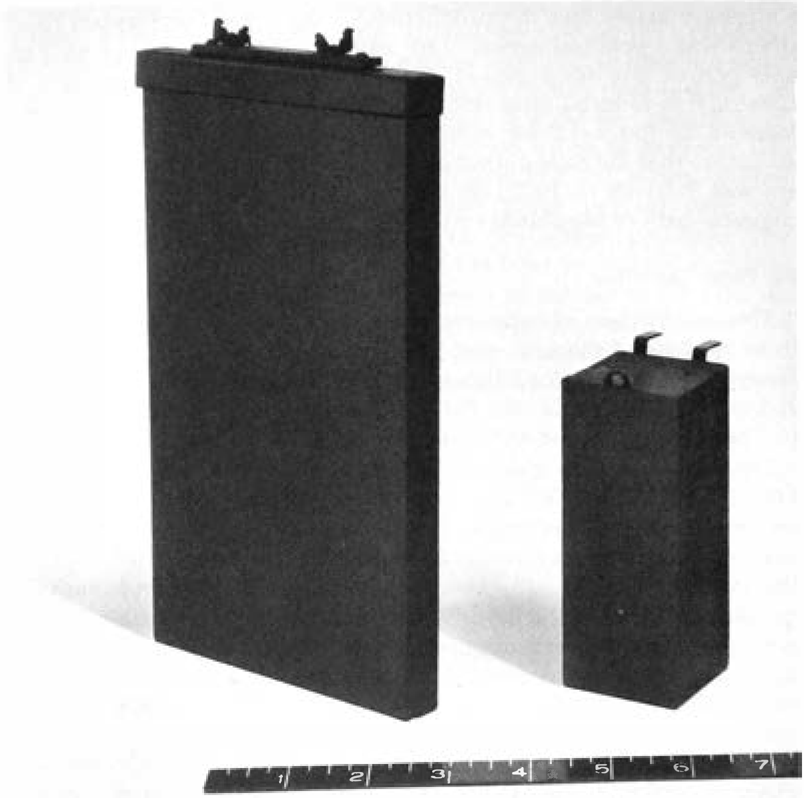


Fig. 8-19. Western Electric 2-microfarad, 200-volt capacitors. At left, the 5-type, standard until 1905; and, at right, the 21-type, standard from 1905 to 1929.

Paraffin was used exclusively as the impregnant in capacitors until after 1925. The low melting point was of some concern because, in warm locations, it melted and leaked out of the container. This problem was solved by the use of higher-melting-point paraffins and additives such as beeswax and carnauba wax, a mixture that was standardized as Western Electric No. 9 Compound.

The early capacitors used paper that had been soaked in paraffin. This process was replaced in 1907 by vacuum impregnation of the wound units. A Western Electric processing specification dated November 1914 (No. 50216, Issue 1) called for oven drying followed by 20 minutes under vacuum prior to the admission of the wax for impregnation. Initially, vacuum was used for economic reasons; it reduced the amount of paper and foil required for a given capacitance by more effectively eliminating trapped air. It was known that adequate drying was essential for a high insulation resistance, but the level of insulation resistance achieved by oven drying alone was considered adequate. Consequently, vacuum drying was not believed to be necessary from that standpoint. The 20-type and a few other capacitors were oddities in this respect. At one time they were intentionally made with a high moisture content to obtain a higher dc capacitance. Capacitance was adjusted by baking, to decrease capacitance, or by steaming, to increase it. The moisture degraded other properties and caused capacitance to decrease drastically as frequency increased. The process was reported to be "difficult to control."

The metallized-paper capacitor developed by Mansbridge was never used in the Bell System, although at least two investigations of its properties were carried out. The Mansbridge capacitors were found to be electrically inferior (lower insulation resistance and higher power factor) and more costly than the conventional metalfoil-paper types.

The following is a chronology of events relating to the paper capacitors that were widely used in the telephone system prior to 1925:

Code[14]	Event and Date
5A	Standard prior to 1905
18B	Standardized in 1901
20C	Standardized in 1901
21A	Standardized in 1902
21B	Standardized in 1904
21D	Baked and pressed—Standardized in 1905
21D	Vacuum treated—1907
21D	Elliptical units—1909
21D	Potted in paraffin—1910
21D	Dipped and potted in asphalt (WE compound)—1912

Prior to 1903, code marking was not applied to capacitors, the only identification being the capacitance marking. Due to the proliferation of types, and the availability of the same capacitance value in different types, it was decided in 1903 that the marking should include the code number. At about the same time, alternating-current measuring equipment was introduced in the Western Electric capacitor shops. Prior to that, ac measurements were made only in the laboratory and dc was used in the shop for capacitance measurements.

[14] Codes 18B, 20C, 21A, and 21B were the wet capacitors mentioned above. They were abandoned in 1904 because of their inferior properties.

The earliest reference to life testing of capacitors on continuous voltage as a means of measuring their capability for satisfactory service is contained in a memorandum from C. N. Frazee to J. A. Davison dated November 26, 1917. The short-time breakdown voltage was then the normal criterion for this property, but Frazee proposed conducting long-time tests.

5.3 Mica Capacitors

Although mica was one of the earliest dielectric materials used for capacitors (capacitors with tinfoil interleaved between sheets of mica were described in 1845), it was not employed extensively in telephone systems until carrier systems were introduced. Most early applications for capacitors required relatively large values of capacitance, which could not readily be produced with small sheets of mica. However, with the advent of carrier telephone systems, small, stable capacitors were required for use in filters, oscillators, and other equipment. Mica capacitors met this need.

Earlier, the main applications for mica capacitors had been as capacitance standards or for use in test sets. The first Western Electric coded mica capacitor was the 38A, a 1-microfarad unit brought out in 1914 as a standard. Mica capacitors for experimental and developmental systems were, however, produced much earlier.

It was recognized at an early date that an intimate contact between the dielectric and electrodes was essential to attaining stable capacitance. This led to the development, prior to 1904, of a capacitor in which silver electrodes were deposited chemically on the mica. Details of the manufacturing process for these capacitors were contained in a memorandum from E. B. Craft to W. S. Fulton dated October 23, 1907. A single capacitor unit of this type, which was part of a shop capacitance standard made prior to 1919, is in the Bell System museum.

The production of silvered mica capacitors by this process was slow and tedious, so their use was limited to applications requiring the highest stability. In the more common construction, foil electrodes were interleaved between mica sheets and the assembly held between rigid clamping plates. This method of fabrication continued in use until the 1930s.

5.4 Electrolytic Capacitors

The unusual properties of electrolytically formed films on aluminum were discovered in 1855, and capacitors based on this principle were used in Germany for early common-battery telephone systems. By 1908, they had been replaced by paper capacitors. The unpredictable (at that time) electrical performance and the problems associated with the use of an aqueous liquid electrolyte discouraged wide use of this type of capacitor.

However, there was a great need for compact, high-capacitance units in the telephone system. With small systems and light (talking) loads on the common battery, intermittent charging of the battery in the central office at "off" hours was adequate. However, as systems grew and other direct-current loads, notably switching apparatus, were added, it was necessary to keep the dc generator operating continuously. The ripple voltage in the output of conventionally wound and commutated generators produced objectionable noise in the system, so specially designed low-ripple generators were required.

Several workers, both in the United States and Europe, made studies of the electrolytic capacitor around 1900, and about 1902, studies were begun in the Engineering Department of the Western Electric Company. This work contributed to improved performance and life of the capacitors through better understanding of the effects of contaminants, electrolytes, and physical factors. These innovations were covered by a number of patents relating to aluminum and tantalum electrolytic capacitors (see Fig. 8-20). The improved performance and longer life stimulated large-scale use in the Bell System in filters to minimize the ripple voltage in the output of conventional direct-current generators.

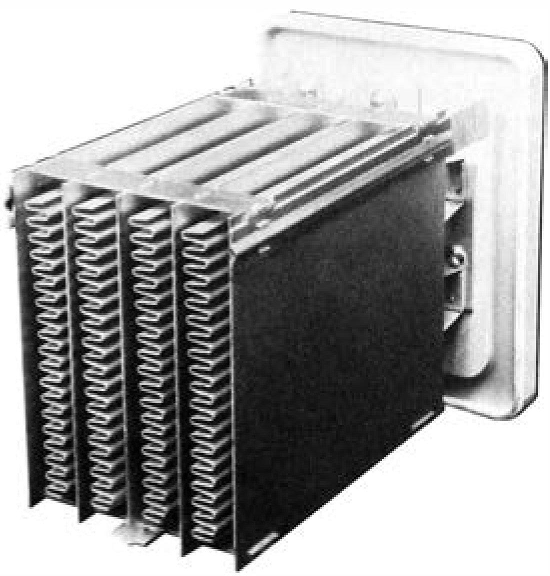


Fig. 8-20. Interior view, showing construction, of electrolytic capacitor. The corrugated plates are the positive electrode.

5.5 Miscellaneous Capacitors

Glass was the dielectric in the first capacitor, the Leyden jar, and it has continued in use to the present. Leyden jars were being used for high-voltage work in the laboratory until 1905 or later, but there was little, if any, significant use of glass capacitors in telephone systems until the 1930s.

Advances were being made by other workers and a brief account of one of these will set the stage for developments described in a later volume of this history. The early experimenters recognized that one of the causes of failure of the Leyden jar at high voltage was the "brush discharge," or corona, that occurred at the edges of the electrodes or where the electrodes did not contact the glass. Their solutions to the problem included silvering of the glass by a chemical process used for mirrors, followed in some instances by copper plating over the silver. With copper plating to protect the thin and fragile silver coating, relatively thick metal foil electrodes were no longer required. One example was the Moscicki capacitor patented in Great Britain in 1904. This capacitor and its variants included several means for reducing the voltage stress at the edges of the electrodes, thereby avoiding corona at this point. These were the forerunners of modern tubular glass capacitors, which had to await improved and more practical materials.

Ceramics were not used as capacitor dielectrics until the 1930s. Coursey does include porcelain in a table of dielectric properties of materials, but it was used only for insulators and bushings.

Air capacitors, particularly the variable variety, have been used since about 1900. Their early use in the telephone system was in test sets, but with the advent of radio and carrier, they found wider use as tuning elements. The early need for and interest in variable air capacitors are indicated by the fact that, prior to 1925, more patents relating to them were granted to the Western Electric Company than for any other type of capacitor.

VI. RESISTORS

With a few exceptions, resistors developed for Bell System use in its early years were of the wire-wound type. Most of these were designed for manufacture by Western Electric, but a substantial number of the wire-wound power types and fixed-composition resistors were obtained from outside suppliers under Bell System specifications.

For installation in telephone exchanges, equipment bays were developed and standardized, using relay racks with uniform hole spacings for mounting components and related apparatus. Two examples of the specialized forms of resistors developed as early as

1901 are the 18- and 19-types, illustrated in Fig. 8-21. These were flat units that could be mounted interchangeably with other kinds of apparatus, such as relays and capacitors. Various kinds of windings were used: single layer, multilayer, bifilar, and reverse layer, in accordance with intended circuit requirements. The windings were applied to a phenolized asbestos core and terminated in sideposts with terminals, two for the single-winding 18-type and three for the double-winding 19-type. The units were conservatively rated at 5 watts dissipation and made in resistance values up to 10,000 ohms, with tolerances of from 5 percent to 0.1 percent, as required. They were extensively used, from the time of their initial manufacture in 1901 by the Western Electric Company, in quantities which in subsequent years were as high as 6 million per year. Though improvements were made and costs reduced over the years in the 18- and 19-types, the basic structure remains to this day as an excellent example of design for efficient production and economical use.

Another widely used standard type of resistor (also shown in Fig. 8-21) was the No. 1 type, small and compact, having one winding on a brass core, with a brass cylindrical shell over the outside. These were made in inductive and non-inductive designs. (The No. 5 type, wound on a wooden spool, dates from the years before 1905, as does the No. 1.)

Resistors of the wire-wound type are restricted in the maximum values obtainable by physical limits to the wire size that can be handled in production. One of the earliest designs using carbon as a resistive element was the 38-type, first made in 1910. This unit consisted of a single carbon filament winding placed in a spiral groove on a cylindrical lavite core. Insulating and moistureproofing compounds were applied to the spool after winding, for protection from

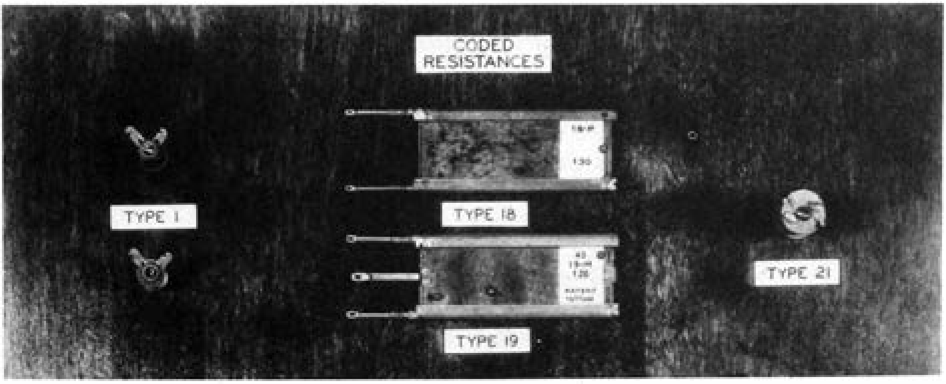


Fig. 8-21. A display board of coded resistances showing early designs.

damage. Resistance values in the range of several hundred thousand ohms were obtainable in the 38-type design. In later years, the use of carbon resistors expanded greatly as electronic circuits were applied to telephone technology, and sophisticated designs of great precision and matched temperature coefficient were developed by Bell Telephone Laboratories for use in the highly complex radar computers and gun directors of World War II.

VII. VACUUM TUBES

7.1 The Goal—Reliability, Uniformity, Long Life, and Low Power Consumption

The advent of the vacuum tube and its earliest application to long-distance wire and radio communication have previously been covered in Chapters 4 and 5. Chapter 10 contains information on the Bell System's fundamental research on vacuum tubes and the results of basic studies of thermionic emission and tube design begun in 1912.

The properties of vacuum tubes were intensively explored, and improved designs of the tubes and the circuits in which they were used came out of the Engineering Department of the Western Electric Company and the Research and Development Department of AT&TCo, working in close collaboration at an innovative rate never before known in industrial research and development. The capability of the vacuum tube for almost distortionless amplification and its ability to function as an oscillator, detector, rectifier, modulator, or demodulator were so useful that it became the essential element in long-distance communication by wire, cable, and radio systems, in radio broadcasting, in telephotography and television, and in a variety of means for mobile, maritime, and military communication for the next 50 years.

Special requirements on vacuum tubes for telephone service had to be met to assure long life, uniformity, stability, and freedom from distortion of the transmitted signals. In long telephone circuits, over 100 tubes might be included in the circuit in each direction of transmission. Since they were operated in cascade, imperfections that might cause noise and distortion were additive. Uniformity and stability of operating characteristics were thus imperative. Furthermore, the failure of a single tube could cause interruption of service, so design for long life was another controlling objective. It is not surprising, then, that a continuing strong effort for new and improved vacuum tubes was maintained in the Bell System laboratories.

As was noted earlier (see Section 4.2.2 of Chapter 4), the original vacuum tube, invented by de Forest, was unsuitable for use as an amplifier because it depended for its action, in part, on the ionization of residual gas within the tube. Discovering the value of, and developing the means for, obtaining a very high vacuum made possible, for the first time, a true vacuum tube—having qualities of uniformity and reproducibility, and the possibility for practically distortionless amplification necessary in the telephone "repeater."

The work done by Arnold and his associates, beginning in 1912, was the start of a comprehensive program of research and development in the laboratories of the Bell System. An important part of this program was the investigation of the physics and chemistry of the cathode, or filament—the electron-emitting element of the vacuum tube. This research, described in Section IV of Chapter 10, was directed toward obtaining an understanding of: the factors that affected electron emission from cathode surfaces; the more promising emitting materials; practical methods for manufacturing tubes having cathodes of high efficiency; the best metals for plates and grids; the best varieties of glass for the envelopes; and the best methods of exhausting the tubes to remove residual gases from constituent parts.

By 1925, tube life expectancy, as compared to the early tubes of 1913 and 1914, had been increased by 50 times and more, from a few weeks to several years of continuous operations, and the filament power reduced by a factor of 5, from 10 watts to 2. The number of vacuum tubes in use in the Bell System in the late 1920s was over 100,000—having more than doubled every two years since their first use in the telephone plant early in 1914. Over half of this number were used in telephone "repeaters," or what we now call amplifiers. About 10 percent were used in station ringers, 20 percent in other applications in carrier and telegraph systems, and the remainder in amplifiers for broadcast program circuits, test equipment, telephotograph apparatus, radio communications, and public address systems. The tubes ranged in size from the miniature "Peanut" tube with a power output rating of 1/100 of a watt, used in radio receivers, audiometers, and amplifiers for the hard-of-hearing, to the large water-cooled tubes used for radio broadcasting and transoceanic telephone service, with power ratings up to 100 kilowatts.

The tube types in most common use were the 101D, earlier known as the "L" tube, and the 102D, originally known as the "V" tube. These tubes, standardized in 1922, used platinum-alloy filaments, operating at a current of 1 ampere. By the end of 1923, these two types accounted for about half of the total number in use. In 1927, improved models using 0.5-ampere filaments, the 101F and 102F, were put into service, and they soon replaced their less economical predecessors. Estimates were made in the early 1930s showed that the improved tubes were saving the Bell System about 10 million dollars per year through lower cost of manufacture, longer life, and smaller power consumption.

7.2 The 101F Vacuum Tube and Its Predecessors

In a 1926 paper in the *Bell Laboratories Record*, M. J. Kelly, reviewing the manufacture of vacuum tubes as of that time, wrote about the 101D tube: "Historically, this is the oldest high vacuum thermionic tube used in commercial communication (wired or radio) in the world. In the thermionic tube family it stands supreme, for to our knowledge there is no other standard high vacuum tube of as long life and of as high quality. It has been used for over twelve years in the telephone repeaters of our Bell System, in fact, ever since its proportions were established by H. D. Arnold."

In an early form as applied to the telephone repeater, the vacuum tube developed by Arnold is shown in Fig. 8-22. This tube, with the designation "Type A," was first used in service in the Philadelphia repeater of the circuit between New York and Washington in October 1913. It used a platinum filament coated with a barium compound, a nickel plate and grid, and a rugged internal supporting structure for the elements.

Shortly thereafter, an improved version, Type B (Fig. 8-23b), with a longer filament and an improved grid structure was developed. Provided with a base for mounting, it was known as Type M (Fig. 8-23c) and, in this form, it was used as the essential element in the transcontinental telephone line, opened to commercial service on January 15, 1915 (see Section 4.2.3 of Chapter 4). It is interesting to note that the amplification factor of this early tube was 5; it operated at a filament current of 1.5 amperes, a plate potential of 100 volts, and a plate current of 10 to 15 milliamperes; and it had a useful life of approximately 400 hours.

Before the end of that year, an improved tube was available with a life expectancy of over 4,000 hours. Designated Type L (Fig. 8-23d), this tube incorporated a filament that could operate at a lower temperature with increased emitting area, thus reducing the filament current requirement to 1.3 amperes.

A word about nomenclature is in order at this point. About the middle of 1916, the letter designations used by the laboratory design group were replaced by Western Electric code numbers similar to those assigned to identify other types of telephone apparatus. The Type M became the 101A and the Type L, the 101B.

Further improvements in internal mechanical design and in lowering filament current resulted in a new design, the 101D, first produced in 1921. This tube was characterized by an amplification factor of 6, a plate resistance of 5,700 ohms, and a plate current of 7.0 milliamperes. Filament current required now was 1 ampere, a one-third reduction from its predecessor. It was fabricated from a new design of a carefully

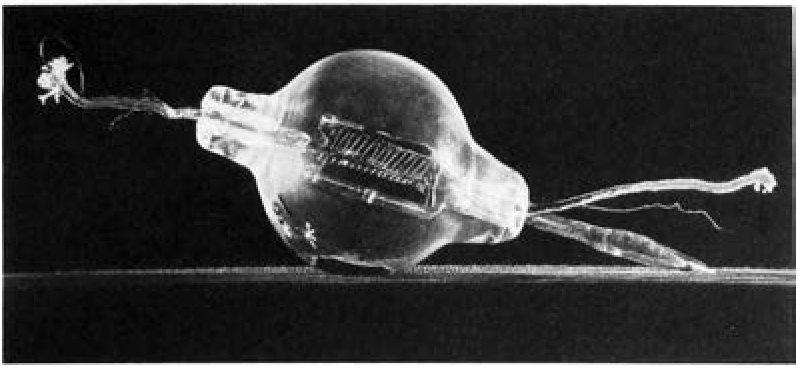


Fig. 8-22. Arnold's high-vacuum tube, used as a telephone repeater at Philadelphia on a New York–Washington cable circuit in October 1913.

controlled platinum-nickel alloy, which, like the earlier platinum-iridium filament, was coated with a mixture of barium and strontium oxides applied in a number of successive layers. Since the coating literally combined chemically with the filament metal, it was referred to as a "combined" coating. In a few years, by 1927, studies of filament characteristics and materials had progressed to the point where satisfactory operation could be obtained with a current of 0.5 ampere, or only 2 watts of power, just half of that previously required. This tube was designated 101F. By the end of 1928, this tube was in widespread use throughout the Bell System. The successive developments leading to the 101F are shown in Fig. 8-24; its component parts and assembly are shown in Fig. 8-25.

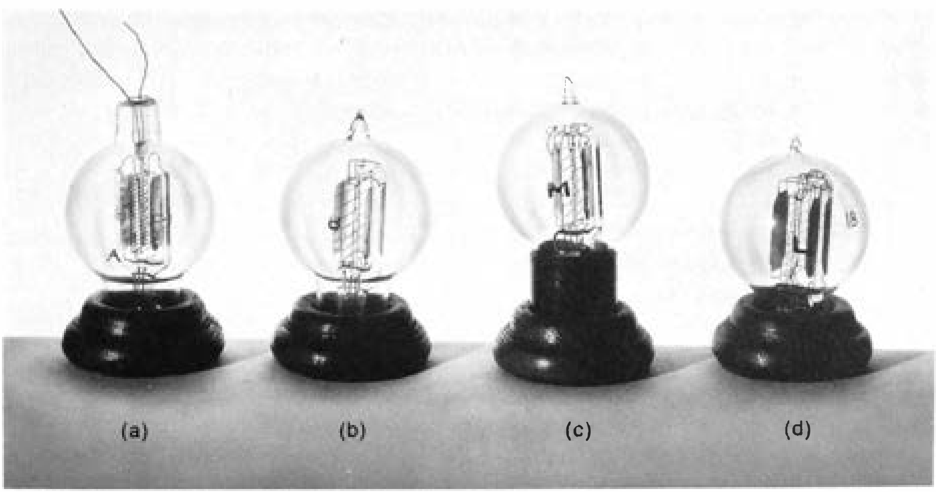


Fig. 8-23. Four early vacuum tubes: (a) Type A of 1913; (b) Type B of 1914; (c) Type M of 1914; (d) Type L of 1915.

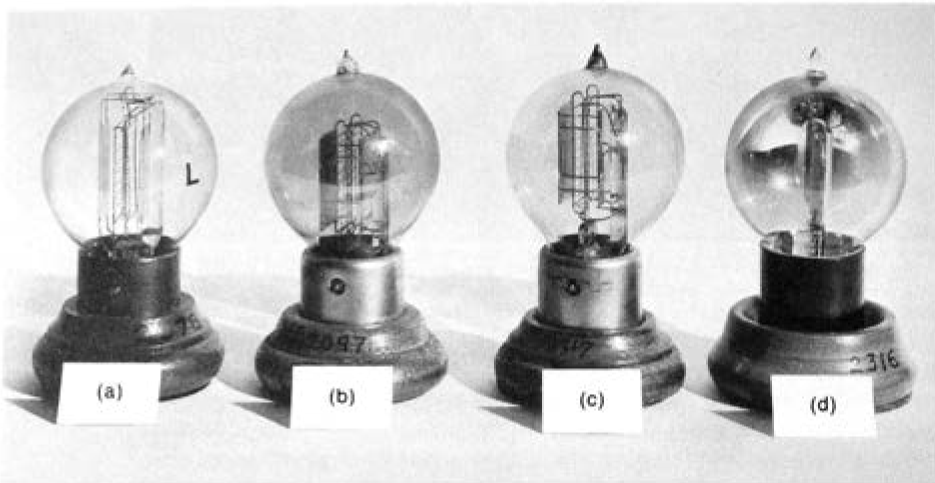


Fig. 8-24. Evolution of the Type L tube: (a) Code 101B of 1915; (b) Code 101B of 1919; (c) Code 101D of 1920; (d) Code 101F of 1927.

7.3 From the "V" Tube to the 102F

The earliest amplifiers in Bell System use were single-stage, utilizing the successive forms of the "L" tube. By 1915, in response to the need for higher amplification, two-stage repeaters were developed and a

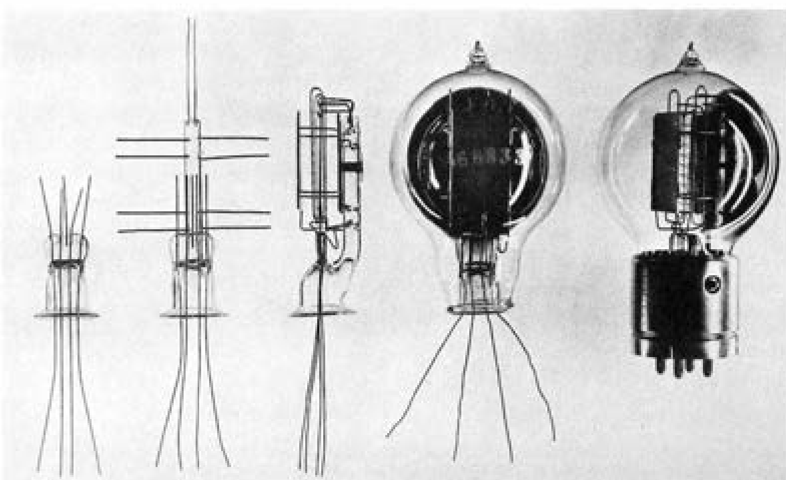


Fig. 8-25. Steps in the manufacture of the Code 101D tube.

tube was specifically designed for high amplification in the input stage. Appropriate changes were made in the grid and plate structures, in the filament, and in the spacing between the elements for this application as a voltage amplifier. In 1916, the designation was changed to 102A; in 1922 successive improvements led to the 102D, and later improvements paralleled the design changes for longer life and reduced filament current as described for the 101 series. The characteristics of the 102D tube were: an amplification factor of 30, a plate resistance of 50,000 ohms, and a plate current of 1 milliampere. The filament current was 1 ampere, with a voltage of 2.1 volts. The 102F had similar characteristics except that the rated filament current was 0.5 ampere. Like the 101 tube, it had been carefully designed and manufactured to exacting standards to ensure long, trouble-free performance. The average lifetime of these tubes, when used in the equipment for which they were designed, was 70,000 hours or about eight years of continuous operation.

7.4 The VT-1 and VT-2—Landmarks in Design for Military Service

Some highlights in the development of specialized systems and apparatus for military use in World War I by the Western Electric Company are covered in Section 4.1 of Chapter 5, where Figs. 5-18 and 5-19 illustrate the VT-1 and VT-2 tubes, designed to meet critical needs in aircraft and naval service. They were the first of what we now term "ruggedized" structures built to successfully withstand unusually high degrees of vibration and shock. VT-1 was the Signal Corps designation for this type, which had originally been identified as the "J" tube and later coded in the Western Electric 203 series. The VT-1 was a general-purpose tube and was used as a detector, amplifier, or low-power oscillator, with a rated plate current of 2 milliamperes, an amplification factor of about 6, and an internal plate impedance of about 20,000 ohms. The VT-2 was a small transmitting tube rated at 5 watts output and operated at approximately 300 volts and 40 milliamperes plate current. Earlier designations for this design were, first, Type E and, later, Western Electric codes in the 205 series. Together, the VT-1 and the VT-2 accounted for most of the Western Electric production of vacuum tubes for military communications in World War I, an output which reached 25,000 units per week by the end of hostilities in November 1918.

7.5 The "N" Tube—The First Miniature Vacuum Tube

In the last year of the war, the Western Electric Company developed a special tube for the Signal Corps for operation in portable equipment, where low filament and plate power are important requirements. This need was successfully met by the "N" tube, later coded 215A, which operated with a filament current of 0.25 ampere supplied by a single dry cell of 1.5 volts. Utilizing a concentric form of assembly and with a cylindrical plate, a spiral wire grid, and a very fine oxide-coated filament, the tube, shown in Fig. 8-26, was a little over 2 inches high and ⅝ inch in diameter. Because of its appearance, it was called the "Peanut" tube, an apt designation by which it was known to the postwar generation of radio experimenters and enthusiasts. The "N" tube was used in large quantities for many years in radio receivers and in instruments like the audiphone (an early hearing aid) and audiometers, where small size and low current drain were important.

7.6 The Equipotential Cathode

In a notebook entry dated November 20, 1913, A. M. Nicolson described a vacuum tube that he had constructed, which utilized a "unipotential" cathode. This design, the first of its kind, had its cathode surface independent of and insulated from its heating mechanism. This meant that all parts of the cathode surface could be maintained at the same potential with respect to the other elements, a feature which proved to be of value in many specialized amplifier and detector circuits.

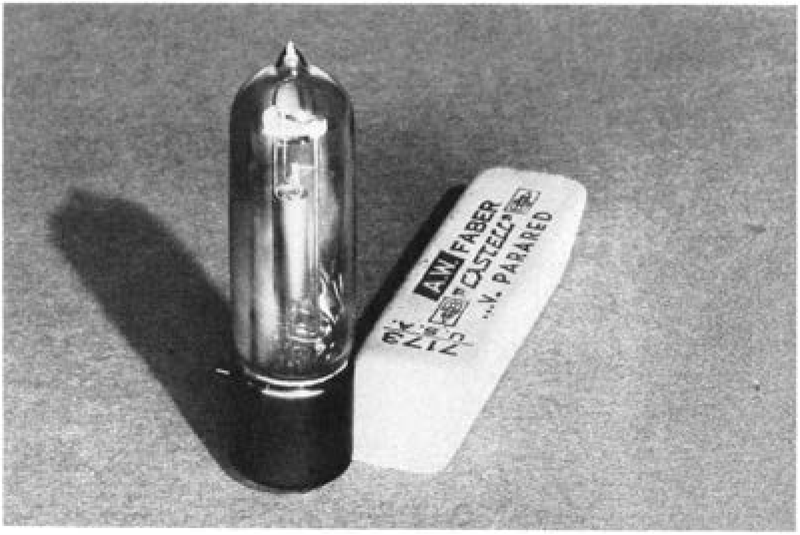


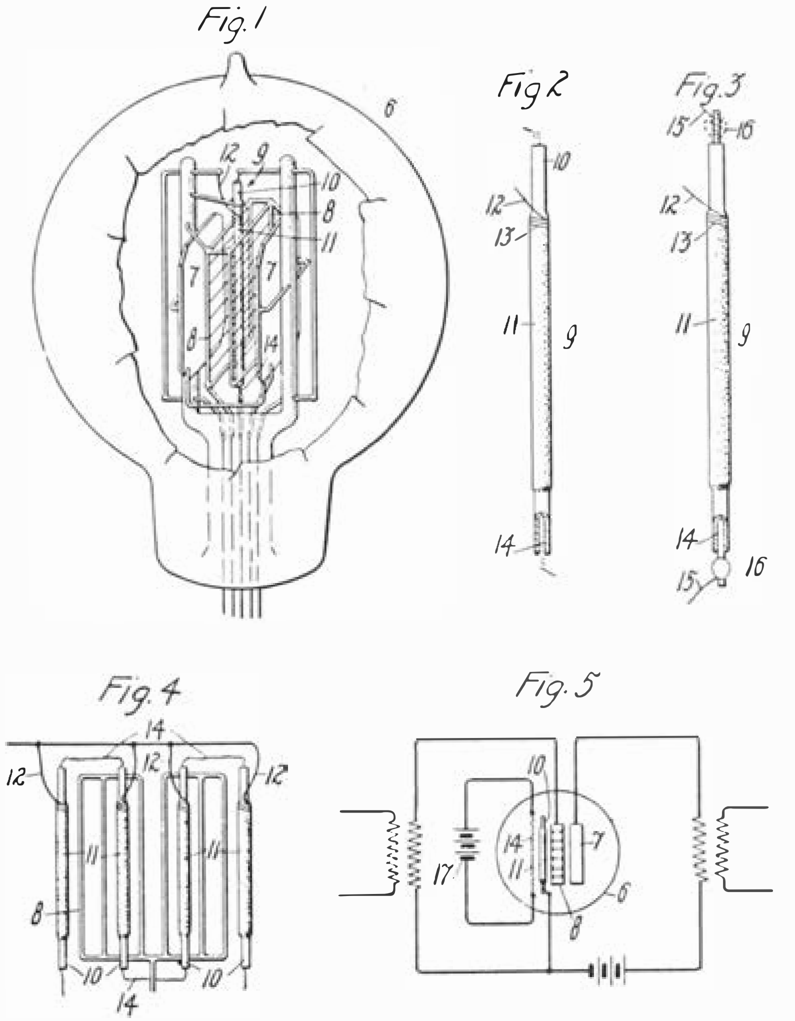
Fig. 8-26. The Peanut tube, patented by H. W. Weinhart.

June 19, 1923. 1,459,412

A. McL. NICOLSON - 27

THERMIONIC TRANSLATING DEVICE

Original Filed April 16, 1915



Inventor
Alexander McLean Nicolson
by W. E. Beatty. Att'y

Fig. 8-27. A page from Nicolson's patent on the equipotential cathode.

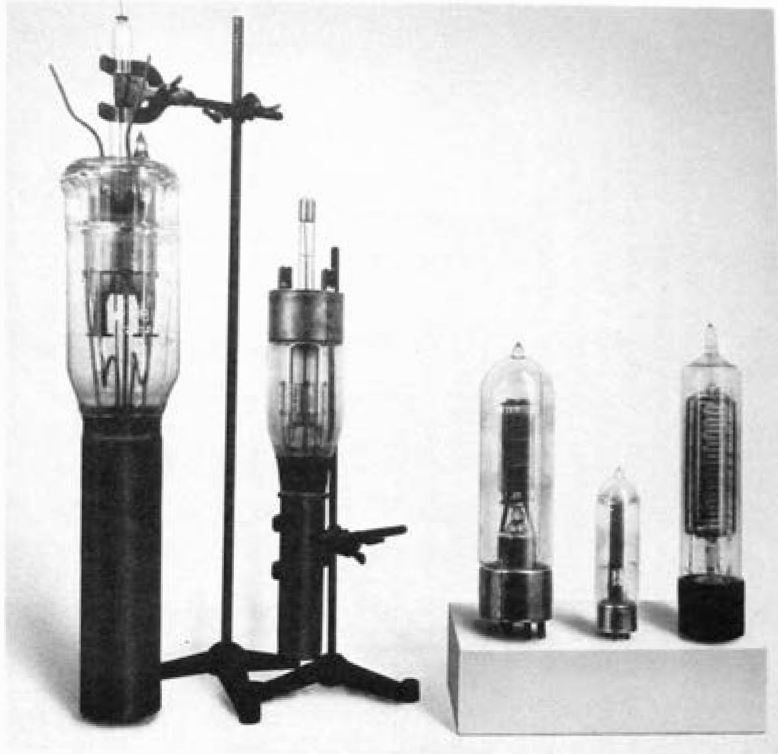


Fig. 8-29. Steps in the development of the high-power vacuum tube: from right, Type W (50 watts) of 1915, Type G (50 watts) of 1919, Type I (250 watts) of 1920, Type T (10 kilowatts) of 1921, and Type U (100 kilowatts) of 1922.

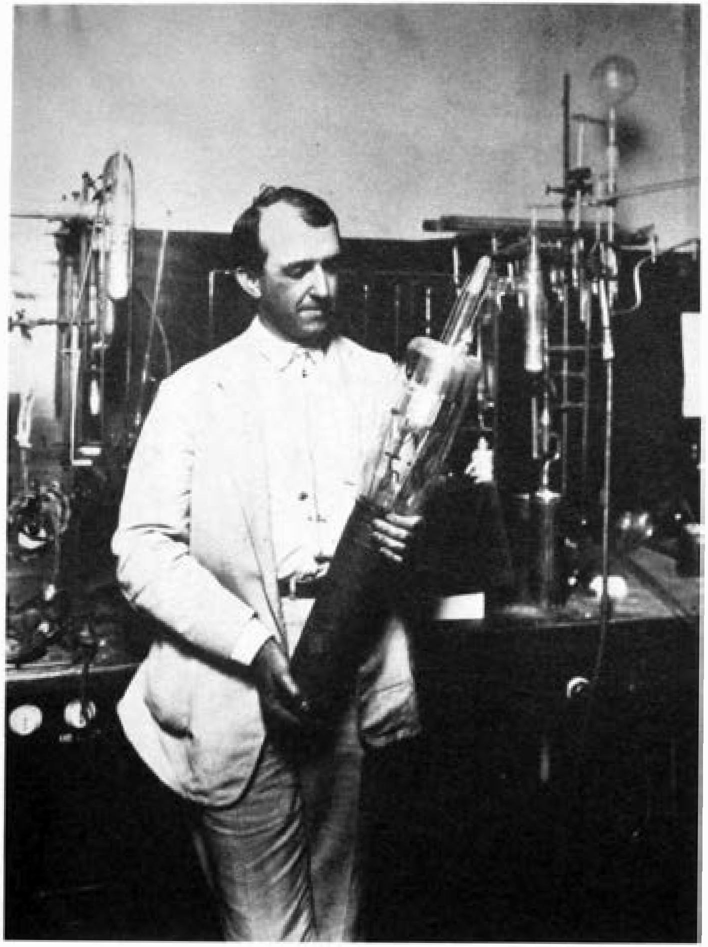


Fig. 8-28. William G. Houskeeper with a 100-kW high-power, water-cooled vacuum tube incorporating the Houskeeper seal.

Of even greater future importance was the fact that, by separating the heating agent from the element that was the thermionic source of electrons, the cathode could be heated by alternating current. In the 1920s, when radio broadcast receivers were finding their way into homes all over the world, the "heater-type" tube made it possible to move the storage battery out of the living room, replacing it with a small, simple transformer hidden in the receiver or part of a separate ac power supply. The application was essentially universal since the indirectly heated structure was adopted by all the manufacturers of radio tubes and receivers.

Figure 8-27 is a page from Nicolson's patent,[15] the original of which was filed on April 16, 1915. Though at that time it was too early for him to anticipate the widespread application of his idea to home radio use, his basic concept of "divorcing the heating agent from that which produces the thermionic activity" was an outstanding innovation in vacuum tube technology.

7.7 High-Power Vacuum Tubes

The development of high-power vacuum tubes in the Bell System was stimulated, first, by the need in the transatlantic radio experiments of 1915 and, later, by the requirements for other long-distance point-to-point communications radio broadcast transmitters, and large public-address systems. Reference should be made to Chapters 4 and 5 for descriptions of the systems and the devices that were developed for specific areas of application. The external-anode, water-cooled tube using the Houskeeper copper-to-glass seal was an innovation that permitted the design of tubes with power output ratings up to 100 kilowatts, one of which is shown in Fig. 8-28, along with its inventor. Other early high-power tubes, air- and water-cooled, are shown in Fig. 8-29. For many years after the time period covered by this history (until the late 1940s), research and development continued in Bell Laboratories, with manufacture by Western Electric, on broadcasting equipment that earned a worldwide reputation for excellence in performance and reliability. All such equipment utilized Western Electric tubes specifically designed for this application. They were used also in ship-to-shore, air-to-ground, and land mobile communication systems, and in the high-power transoceanic shortwave transmitters that provided worldwide telephone and telegraph communications.

[15] A. M. Nicolson; U. S. Patent No. 1,459,412; filed April 16, 1915; issued June 19, 1923.

Chapter 9

Quality Assurance

Functions, like apparatus, devices, and equipment, appear to develop from a recognized need. Implementation sometimes closely follows recognition of a need; more often it eludes fulfillment and is a continuous process. Quality Assurance is such a function.

I. INTRODUCTION

The term is relatively new—some 30 years old—but the need for some function to generate confidence in the quality of "instruments" and "telephonic appliances"[1] supplied the Licensees was recognized at least as early as the "1882 Contract," an agreement between the Western Electric Company and the American Bell Telephone Company. In regard to "instruments" the agreement states:

> At all times during their manufacture and upon their completion the instruments and the materials employed shall be subject to the inspection and acceptance . . . by said first party. (American Bell Telephone Company)

With reference to "telephonic appliances" the agreement provides:

> . . . apparatus manufactured under this license . . . shall be subject to inspection of the party of the first part. (American Bell Telephone Company)

These simple statements suggest a basic concept of Quality Assurance: one organization operates as an agent for assuring the quality of apparatus supplied to another. Some rather rigorous inspection applying to "instruments" was almost mandatory since the American Bell Telephone Company retained ownership of all "instruments" and, accordingly, assumed responsibility for their serviceability in leasing them to the Licensees (Operating Telephone Companies). The somewhat looser inspection provision applying to "telephonic appliances" was not

[1] "Proprietary instruments" were transmitters, receivers, and induction coils. Telephonic appliances included calls (signals), switches, switchboards, annunciators, exchange furniture, and other apparatus adapted for use on telephone lines.

applied to other apparatus, cable, wire, etc., although both categories, "telephonic appliances" and other apparatus, were sold to the Licensees by the Western Electric Company. The Operating Telephone Companies accordingly found it necessary to inspect much of the material and apparatus supplied by the Western Electric Company and, as the System grew, this inspection was extended to essentially everything except proprietary instruments.

The situation was reviewed as part of an extensive restudy of the functions and operations of the Engineering Department of the American Telephone and Telegraph Company, apparently instigated by E. J. Hall, Vice-President of the Company. The following excerpts from a letter dated July 17, 1907, from J. J. Carty, Chief Engineer of AT&TCo, to Hall, pertain specifically to the inspection operations:

> Pursuant to your instructions, I have made a plan for the reorganization of this department so that it may be operated with greater economy and efficiency . . .
>
> * * *
>
> With respect to the inspection and testing of this apparatus, (telephones, transmitters, relays), it should be conducted by the Western Electric Company, following methods and standards which have been approved by this Department, and which would be checked by us from time to time so as to assure ourselves of their efficiency.
>
> * * *
>
> The inspection and testing of material and supplies for the various Licensee Companies.
>
> At the present time there is maintained by this Department a staff of inspectors for this purpose. The Western Electric Company and many of the Licensee Companies also have a separate staff engaged in this work. It would make for economy and efficiency if this testing and inspection work for the Licensee Companies, for the Western Electric Company and for this Company were done by one central organization. In view of the Western Electric Company's relation as Supply Agent and Purchasing Agent to so many of the telephone companies and in view of their close touch with the suppliers, they are in a position to conduct this inspection more economically and efficiently than could be done elsewhere.
>
> Any objection which might be raised against this plan, because the Western Electric Company is the seller, is met by having this Department (The Engineering Department of the AT&TCo) from time to time check the inspection and inspection methods of the Western Electric Company in such a manner that the efficiency of their inspection may be determined.

It is evident from this that AT&TCo, while divesting itself of inspection responsibilities, was not eliminating its obligations to the Licensees for the quality of product supplied.

Hall endorsed the program proposed by Carty and authorized him, "to go ahead along the lines proposed," in a letter also of July 17, 1907.

The Western Electric Company's phase of the program was implemented during the next few months. The following is an excerpt from the Western Electric Engineering Department Report of Work to May 1, 1908:

> During the period covered by this report our responsibility with respect to the inspection of material which the Company supplies has been enlarged. The Department has for a long time maintained a force of inspectors at the Chicago and New York houses for the purpose of checking the Shop products; we have also checked the work of the Installing Department by inspecting a limited number of new central office equipments and we have been inspecting the line material purchased by most of the associated companies. It has been the practice in the past for some of the associate companies to supplement our work by an additional inspection performed by men in their employ and inasmuch as this seemed to involve unnecessary expense, a plan has been established by which the inspection of the associate companies will be eliminated and the responsibility for the quality of our apparatus and supplies has been delegated entirely to this department.

Thus, by 1908 what might be called the delegated-agency phase of the Quality Assurance function had begun to take shape in broad outline under the driving force of economy and efficiency. Briefly, it may be described as follows: First, AT&TCo continued its overall responsibility for making available a source of product adequate for the needs of the Operating Companies. Second, AT&TCo delegated to the Western Electric Engineering Department the responsibility for conducting inspections and tests of product to assure its quality. Third, AT&TCo reserved to itself the approval of methods and standards for test and inspection and their periodic checking to assure efficiency.

Means available to the Western Electric Company in 1908 for discharging the responsibilities delegated to it were inadequate. Almost the only tool available was inspection in various forms. Both sampling and detailed inspection were employed extensively throughout manufacturing processes and after all manufacturing operations had been completed. Perhaps the most unique method was the "Engineering Inspection" conducted by the Engineering Department discussed at some length in the next section. However, viewing the Quality Assurance function chiefly as an inspection operation had serious limitations. Although the Western Electric Company did in fact assume a greater responsibility for quality, inspection of central offices, PBXs, poles, and some supplies continued to be performed by the Operating Companies well into the 1920s.

II. THE QUALITY ASSURANCE FUNCTION AS IMPLEMENTED IN 1906–1908

To place these developments in proper perspective, reveal their limitations, and enable us to anticipate future developments, we should investigate the implementation of the Quality Assurance function at the time of, and immediately following, Carty's letter. The drastic change in procedures had developed gradually over the preceding years. Although the driving force was economy and efficiency, quality was not to be impaired.

A document of the Engineering Inspection Department of the Western Electric Company dated March 7, 1906, and titled "Routine,"[2] detailed the procedures used to verify and, to some degree, control the quality of apparatus supplied the Operating Companies. This Inspection Department was part of the Engineering Department of the Western Electric Company located at 463 West Street in New York City. The procedures embraced the following:

(*i*) Complaints from customers or distributing houses
(*ii*) Complaints from other departments of the Western Electric Company
(*iii*) Inspection of apparatus in stock
(*iv*) Inspection of toolmade samples of new apparatus
(*v*) Inspection of first apparatus embodying changes.

In connection with customer or distributing house complaints, it is significant that New York investigated all complaints except those involving apparatus manufactured at the Hawthorne plant in Illinois (the Engineering Inspection Department maintained a small unit at Hawthorne for this purpose). The prescribed procedure provided for: determining if the complaint was justified, determining responsibility for trouble, notifying the head of the responsible organization, and informing the originator of the complaint of the action taken.

The actual inspections conducted by the Engineering Inspection Department, although mentioned under three categories, were actually of two types. The "Routine" of March 7, 1906, describes the toolmade-sample inspections as follows:

> Samples are to be inspected first to determine whether they are made in accordance with the prints. The design should then be investigated to ascertain whether it will fulfill the requirements. A life test or other tests of durability to be made when the nature of the apparatus warrants it.

[2] The "Routine," dated 1906 and prepared by the Engineering Inspection Department, listed 75 types of apparatus that had been inspected during 1905 along with the output, sample size, and frequency of sampling for each. It also listed other obligations of the department, such as complaint handling and toolmade-sample approval. As such it represents the situation as it had been during the year before its preparation. It would appear that this information had been compiled into a single document with the possible intent of issuing it as a GEI (General Engineering Instruction). However, this cannot be verified since no files of these old GEIs issued by the Western Electric Company are in existence at this time and only an occasional GEI issued before 1925 is encountered in Bell Laboratories files.

This statement implies a fundamental concept of Quality Assurance, which undoubtedly had been in the minds of many of the people involved, but was nowhere more clearly stated than in this "Routine." The concept is that the manufactured product must not only be in accordance with the specifications and drawings but also that the design itself must fully express the objective of the designer. In other words, the design and the resulting product must be capable of meeting service "requirements" with adequate "durability."

The inspection of apparatus in stock was called, "Periodical, Routine Engineering Inspection," and was regarded as quite specialized and well beyond ordinary inspection for conformance to drawings. This is indicated by the following excerpt from the 1906 "Routine":

> . . . each line of work to be specialized, that is, each kind of apparatus to be in charge of a man who will make a special study of the design, manufacture, packing, installing and maintaining of the apparatus under his charge. A catalogue list should be kept of all apparatus, specifying their vital points, and giving a list and the nature of all complaints received. This list is always consulted before starting a routine inspection.

Concerning the results of these inspections, the "Routine" continues,

> A curve should be made, showing the results of these inspections, for use of the engineering executives. Detail reports showing the results of each inspection are to be furnished to the departments interested.

Inspection of apparatus in stock was of necessity done on a sampling basis in addition to being periodic. The "Routine" gives no rationale for setting either the period or the sample size but it is reasonable to assume that both derived from engineering judgment based upon the number manufactured, complexity of the apparatus, difficulty of manufacture, complaints received, cost of inspection, etc. The rather meager statistical techniques available in 1906 which could be applied directly to sampling problems and the almost complete lack of knowledge of these techniques in the engineering field would preclude any statistical approach. Terms such as process average, allowable percent defective, consumer's risk, controlled product, and sampling error—the current terminology of Quality Control and Quality Assurance—did not appear for nearly 20 years.

The 1906 "Routine" did list some 75 types of apparatus, each with the 1905 output, the inspection sample size, and the number of times per year each sample should be inspected. The samples ranged from a 1,500-sample size three times a year for plugs with an annual output of 500,000, down to a sample size of five once a year for telegraph arrestors with an annual output of 100. The largest production item was protector carbons with an output of 7,000,000 and a sample size of 600, three times a year. The predominant sample size—100—

occurred 13 times; the next most frequent—25—occurred 11 times. The predominate sampling period was once a year. This statement accompanied the list:

> This list should not be strictly followed, but should be changed as the nature and quantity of the apparatus shipped changed.

This quote, labeled a "thinking clause" many years later by Major General Leslie E. Simon, verifies that "Engineering Inspection" differed from the usual inspection and was, in fact, looked upon as an engineering function with all the latitude and responsibility usually delegated to an engineering operation. It also was vested with substantial power as indicated in the following quotations from the 1906 "Routine":

> If the apparatus is not in accordance with the drawings and the trouble is vital to its operations, a memo is to be sent to the Store requesting them to issue an order on the shop to reinspect the stock.
>
> If the defect is due to the design, a memo is to be sent the designing department, requesting them to issue a change order. If the change is vital, the change ticket reads "including finished apparatus in stock."

It is quite evident why Carty thought the Engineering Department of the Western Electric Company was the logical organization to assume the responsibility for assuring (under the direction of AT&TCo) the quality of the apparatus and equipment purchased by the Operating Companies. Western already had in being many of the elements of a comprehensive mechanism for providing such assurance to Western management. Extension of this assurance to the Operating Telephone Companies appeared direct and straightforward if the periodic engineering inspections were made in accordance with methods and standards approved by and subject to review by AT&TCo as proposed in the Carty 1907 letter. Again we see the emphasis on inspection as the major, if not the sole, tool of the Quality Assurance function.

It would be well at this point to examine the effectiveness of the Engineering Inspection routines employed by the Western Electric Company at the time of Carty's letter from the viewpoint of the Operating Telephone Companies. A rather complete analysis is embodied in a letter from R. M. Ferris, Chief Engineer of the New York Telephone Company, dated September 14, 1907, to H. F. Thurber, General Manager of the New York Company. Ferris divided the products purchased from the Western Electric Company into three classes:

(*i*) Material manufactured by the Western Electric Company

(*ii*) Material obtained through but not manufactured or inspected by the Western Electric Company

(*iii*) Material obtained through but not manufactured by the Western Electric Company for which Western would make a fixed charge for inspection.

In describing inspections made by Western on the first and third classes of material, Ferris mentions the regular shop inspection on Western-manufactured product and the inspections on suppliers' product, but omits reference to the "Periodical, Routine Engineering Inspection." Ferris then goes on to detail the results obtained by the New York Telephone Company in its reinspection of the first class of material, items such as heat coils, switchboard lamps, and cable. He lists rejection rates ranging from 3 to 10 percent.

One would be inclined to conclude from these results that prior Western Electric inspection had not been overly effective in view of the rather large percentage of defects found by the New York Telephone Company. Of particular significance is Ferris' discussion of cable. He says that Western samples only about 1½ percent of the cable reels for opens, crosses, and capacitance when the local Telephone Companies make an acceptance inspection themselves, but that they test each cable reel when no subsequent Telephone Company inspection is made.

Only four types of material are listed in the second class. (It will be recalled that this apparatus was subject to the "Periodical, Routine Engineering Inspection.") New York Telephone Company inspections of the two largest runners—glass insulators and copper sleeves—resulted in zero rejections.

The third class of material includes items such as pole-line hardware, crossarms, wire, and strand. For this class, Ferris concentrates on a comparison between the cost of inspection to the New York Telephone Company and the prices quoted by the Western Electric Company. In those instances where costs differ widely (occasionally as much as four to one in either direction), Ferris develops some explanation. The explanation never involves sampling, but rather describes differences in methods of conducting the actual inspection and test operations. It is probable that 100 percent inspection is implied by both companies.

It is evident from Ferris' study that he did not attach much significance to "Routine Engineering Inspection" as the major implementation of the Quality Assurance function. He chose to place his trust in shop inspections which were of an entirely different sort, not under direct control of the Engineering Department of the Western Electric Company, and only incidental to the plan proposed by Carty. This misunderstanding, however, did not put any major obstacles in the way of implementing Carty's proposal and a new Supply Contract was instrumented to take effect January 1, 1908, with the New York Telephone Company and the New York and New Jersey Telephone Company. However, Thurber, General Manager of the New York Telephone Company, in his letter of March 23, 1907, to H. B. Thayer, Vice-President of the Western Electric Company, approving the new Supply Contract had the following to say:

> The contract is necessarily very general in its character: in view of this it seems necessary to outline certain matters of policy and certain detailed arrangements in connection therewith . . .

Of the numerous matters of policy and detailed arrangements Thurber discussed, only two are significant to the developing Quality Assurance concept. The first concerns cost:

> That under this form of contract the Telephone Company will effect as large savings as it would under any form of contract now in effect between the Western Electric Company and an operating telephone company.

The second concerns quality:

> That the Western Electric Company recognizes that in undertaking the work of inspecting material it becomes absolutely responsible for the character of material in the operating telephone plant. . . . that such responsibility rests with the Engineering Department of the Western Electric Company. The Engineering Department of the Western Electric Company will accordingly undertake the supervision of all material inspection and will carry out the work under the methods approved by the Chief Engineer of the American Telephone and Telegraph Company; further, the inspectors employed will in no way report either to the purchasing or shop branches of the Western Electric Company organization.

This essentially completes the picture of the early vision of an economical and efficient method of assuring the quality of product supplied the Operating Telephone Companies. With the eventual approval and signing of the new Supply Contract by all of the Operating Telephone Companies, it might be assumed that implementation of the Quality Assurance function was underway along the broad outlines proposed by Carty. However, if we scrutinize the details of implementation, particularly the area called "Periodical, Routine Engineering Inspection," it becomes evident that difficulties could develop. This inspection was similar to that defined later as check inspection and currently defined as Quality Assurance inspection. It also embodied some of the concepts of the Quality Survey which was not to evolve for many years. It was not realized in 1908 that to be effective in appraising quality, such an inspection must have certain prior assurances. These assurances involve knowledge that the product is regularly controlled at some satisfactory level and that adequate prior inspections have been made. Furthermore, rational means for arriving at the frequency and sample sizes to be used in such inspections were not available. What may even be more important, statistical means for analyzing the inspection data and criteria for drawing valid conclusions had yet to be developed.

By looking into the future, we are led to the conclusion that more than a basically sound philosophy would be required to secure the inspection economies contemplated under the 1907 Supply Contract.

The Quality Assurance function, as it is known today, was not yet born, tools for its successful implementation had not been developed, and most likely the need for such tools was not even recognized. Hindsight enables us to understand why many of the Telephone Companies chose not to recognize this phase of the Supply Contract. Many Operating Companies, among them the Pacific, Mountain States, Southwestern Bell, Southern Bell, Cincinnati and Suburban, New York, and New England, were still inspecting central offices, PBXs, some cable, poles, and some supplies well into the 1920s.

III. THE INSPECTION ENGINEERING CONCEPT

The 1922 Annual Report of the Engineering Department of the Western Electric Company includes an unusually extensive discussion of the purpose and work of the Inspection Department. The discussion opens with the following statement:

> The Inspection Department operates as a check on the Quality of the engineering development work and safeguards by inspection telephone apparatus manufactured or furnished by the Company, and materials of other manufacture furnished to the Telephone Companies, . . .

While the concept of an inspection department functioning as a check on the quality of engineering development work was still unique in 1922, it was definitely in line with the ideas expressed in the 1906 "Routine." (See Section II.) It is evident from the following, however, that the amount of inspection effort had been expanded considerably beyond that covered by the "Periodical, Routine Engineering Inspection" described in the "Routine":

> The work of inspection of telephone products normally carried on at Hawthorne, is now extended to the Philadelphia factory and the factories of the Automatic Electric Company and the Stromberg-Carlson Company.
>
> Periodic inspections by visiting inspectors are made of material repaired at the twenty-seven distributing house shops scattered throughout the country. The inspection of non-Western Electric material, such as . . . line material and construction tools for the use of the Telephone Companies, necessitates continual supervision at the manufacturers' plants as well as study and revision of specifications to insure the proper quality of product.
>
> The timber products business is a very large one and of great importance to the Associate Companies. As a result of studies in both the laboratory and the field in cooperation with the American Telephone and Telegraph Company, the method of handling the inspection of timber products has been put on a very much improved basis.

The report closes with the following paragraph which categorizes all these inspections as "routine check inspections"—a term which is used less broadly today:

> The entire subject of engineering inspection is being studied with a view to bringing our methods and practices more nearly in harmony with present day requirements. There is a feeling that less time should be devoted to actual routine check inspections and more time devoted to a scrutiny of inspection methods and results, and the performance of the apparatus and equipment in the field. A comprehensive survey of the possibilities along this line is underway and it is hoped will be completed before the end of the year.

As one would expect, the Annual Report did not go into detail concerning the reason for the extensive review of established procedures in the engineering inspection area. The Bell System had grown in magnitude and complexity to the point where it was essential for all units, not just AT&TCo, but all Operating Companies and the Western Electric Company, to assume a system viewpoint of much greater breadth than before. Then too, the rapid installation of the newly designed machine-switching central offices placed a greater burden on the Western Electric Company to assure quality in all phases: design, manufacture, and installation. There is some evidence that increased inspection effort was tried as a stopgap measure. For instance: G. D. Edwards, who later headed the Quality Assurance Department, quotes E. B. Craft, Chief Engineer of the Engineering Department, as stating that there were more Western inspectors than there were installers on the "State-Central" panel job in Chicago. A December 18, 1934, letter from H. S. Sheppard, Commercial Engineer at Bell Telephone Laboratories, to C. Uhrig, Assistant Comptroller of AT&TCo, states that as late as 1925, the New York Telephone Company had 30 inspectors in its Engineering Department. The increase in number had followed the introduction of panel central offices.

The opposite approach was also tried. J. L. Kilpatrick, General Installation Manager of the Western Electric Company ordered that all Western inspectors be removed from selected panel offices and that each installer be told individually to do his work right the first time and inspect it himself to insure that it was done right. The invention of the command, "do it right the first time" bears an interesting relation to the "Zero Defects" slogan which became popular during the middle 1960s among military contractors. Edwards' comment then, "Only the Deity had so far been found competent to inspect His own work and find it good," revealed the depth of his insight that made him so well suited to direct the Quality Assurance Department some years later.

In 1922, C. G. Stoll, Assistant General Superintendent in the Manufacturing Department of the Western Electric Company, among other responsibilities, had general responsibility for installation of central offices while E. B. Craft, Chief Engineer, had charge of the

Engineering Department. These men, together with W. H. Harrison, Building and Equipment Engineer of AT&TCo, led the new attack on the quality problem referred to in the 1922 Annual Report. The original ideas underlying Inspection Engineering must be attributed largely to Stoll and Craft. The 1923 Annual Report of the Western Electric Company announced the formation of the Inspection Engineering Department charged with the responsibility of taking a "Company-wide view" including a critical analysis of design work. To accomplish this, the new department was to:

(*i*) Arrange for more contact with the field and do part of the system job

(*ii*) Arrange for more effective contact with the Engineering and Installation Departments

(*iii*) Provide for closer cooperation between Distributing Houses and the Engineering Department

(*iv*) Coordinate inspection work with the Laboratory Staff of the Engineering Department.

A memorandum of November 2, 1923, by R. L. Jones, head of the newly formed Inspection Engineering Department, stated:

> A type of criticism of engineering design work never previously attempted will be provided. Engineering inspection will take a Company-wide view of the technical quality of the systems and apparatus being produced, analytical in character, but broad as the Company's activities in its scope. Thus it is planned to supplement the work of the designer in a most important way.

The March 7, 1906 "Routine" did specifically refer to testing during toolmade-sample inspection to confirm the adequacy of the design to meet service needs. However, this test was made on only the first few articles manufactured, with no provision for follow-up. Also, the intent of the testing was directed more toward establishing the *extent* to which the apparatus reflected the design concepts rather than the *adequacy* of the design concepts themselves.

Jones' concept of the intent of Engineering Inspection was indeed new and far-reaching. Its significance went far beyond the interchange of words in the title of Inspection Engineering. The formation of the new Inspection Engineering Department was announced to the Operating Telephone Companies in Bancroft Gherardi's letter of April 19, 1924. It indicated the principal aims of the new plan as follows:

> (a) More intimate contact with field conditions and equipment results in order that the Electric Company may do its part of the System job with the fullest understanding.
>
> (b) More effective contacts between the Electric Company's Engineering and Installation Departments in various parts of the country, promoting mutual cooperation.

(c) Closer cooperation between the Electric Company's Engineering Department and its Distributing Houses, expediting the investigation and settlement of questions or complaints relating to quality of materials and equipment.
(d) Coordination of their inspection activities with the work of the laboratory staff.

Gherardi then comments,

> Broadly speaking, therefore, the underlying motive of the new plan is to insure that, with the constantly increasing growth of the system, the Engineering Department of the Western Electric Company shall be kept continuously advised as to whether the responsibility of the Company for furnishing equipment in accordance with the requirements of the System is being satisfactorily fulfilled. . . . nothing in the plan is designed to alter the established practices for handling matters pertaining to the operation or adequacy of the service rendered by apparatus, material and equipment turned over to the Telephone Company in accordance with established standards and specifications which matters, as heretofore, are to be taken up directly by the Telephone Company with the American Telephone and Telegraph Company.

It is quite evident from Gherardi's announcement to the Operating Telephone Companies that the organization of the Inspection Engineering Department within the Engineering Department of the Western Electric Company was most significant. We can conclude from Jones' November 2, 1923, memorandum that the criticism of engineering design work in terms of "a Company-wide view of the technical quality of the systems and apparatus being produced" implies the provision for a direct feedback of the operating and field use experience to the engineering design organization. Some feedback existed even prior to 1907, by virtue of the setup for handling complaints, but the implementation of Jones' concept went far beyond this.

The organization structure, as set up early in 1924 and shown in Fig. 9-1, was transmitted with Gherardi's letter. It shows R. L. Jones as Inspection Manager reporting directly to E. B. Craft, Chief Engineer. The organization is divided into two functions: first, Inspection Operations involving primarily the actual inspection of material not of Western Electric Company manufacture and, second, Inspection Engineering embracing the Quality Assurance function.

Although the 1924 organization chart was transmitted with Gherardi's letter, it should be thought of as a single frame in a rapidly evolving panorama, more transient than static. The extent of the changes made up to this time can be seen by referring to a prior chart (Fig. 9-2) issued in August 1922. This shows F. D. Thompson heading the Inspection Department and reporting to W. F. Hendry, Assistant Chief Engineer, who reported in turn to E. B. Craft, Chief

Engineer in charge of the Engineering Department. First, Inspection Operations under Thompson no longer was responsible for Check Inspection and Merchandise Inspection, both under J. A. Davidson at Hawthorne. Second, a new organization called Inspection Engineering was set up under Jones. The remaining Inspection Operations under Thompson shown on Fig. 9-1, were transferred to Western Electric with the incorporation of Bell Telephone Laboratories in January 1925.

The organization chart shown in Fig. 9-1 was perhaps thought to be most significant to the Operating Companies because of the inclusion of ten field engineering locations under H. G. Eddy (some of which had not as yet been manned). It is also of interest in that it shows D. A. Quarles heading Apparatus Inspection and W. A. Shewhart, Theory and Special Studies, both of whom will make themselves heard later. It does not include, however, G. D. Edwards, the long-time Director of Quality Assurance who was assigned to head the Methods and Results organization shortly after this chart was issued. A few months later H. F. Dodge was appointed to fill the Apparatus Inspection Methods and Results block under Edwards.

With the incorporation of Bell Telephone Laboratories in 1925, E. B. Craft became Executive Vice-President of the new company. Jones continued to report to Craft as Inspection Engineer along with the Inspection Engineering organization, but as has already been stated, the Inspection Operations division remained with the Western Electric Company. Figure 9-3 shows the functional organization as it was developed in September 1926. This chart is also a snapshot of a transient situation but is significant in that it shows Edwards heading Systems Inspection along with the Complaint Bureau and both Dodge and Shewhart reporting directly to Jones. Actually, many of these changes had been under way even before the incorporation of Bell Telephone Laboratories. In an article, "The Viewpoint of Inspection Engineering," published in the *Bell Laboratories Record* for August 1926, Jones states:

> The duties of the department are closely interrelated with those of the development departments and certain technical groups in the Western Electric Company. With proper interpretations they are described as follows:
>
> 1. To develop the theory of inspection: putting existing mathematical knowledge into available form for use in laboratory and factory, and developing new principles where existing knowledge is inadequate.
> 2. To develop methods of stating the quality of various types of apparatus and switchboards, and methods of applying these concepts in everyday work; to develop economic standards of quality which telephone materials should meet.
> 3. To maintain oversight of the quality of apparatus, supplies and systems currently being furnished for communication service; to make regular reports on the current quality of these materials.

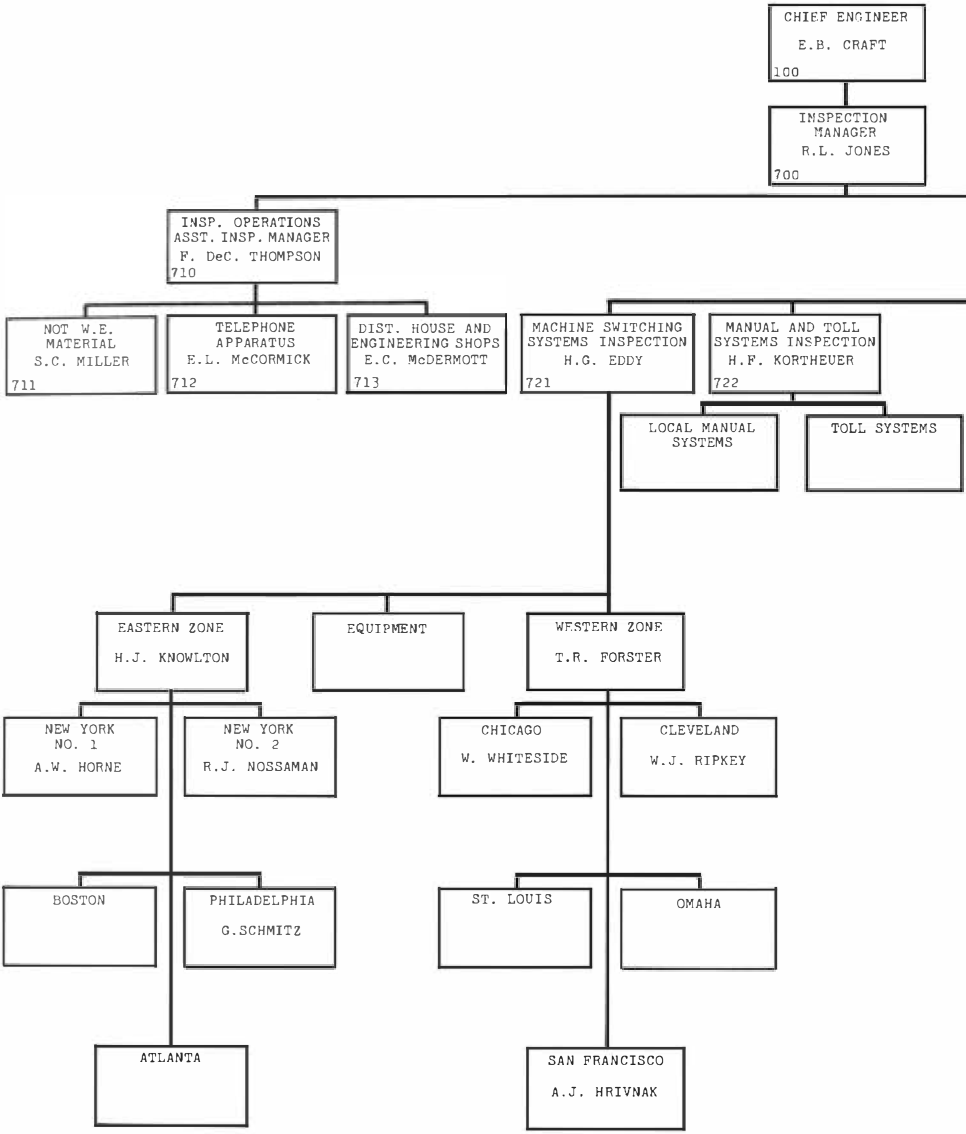
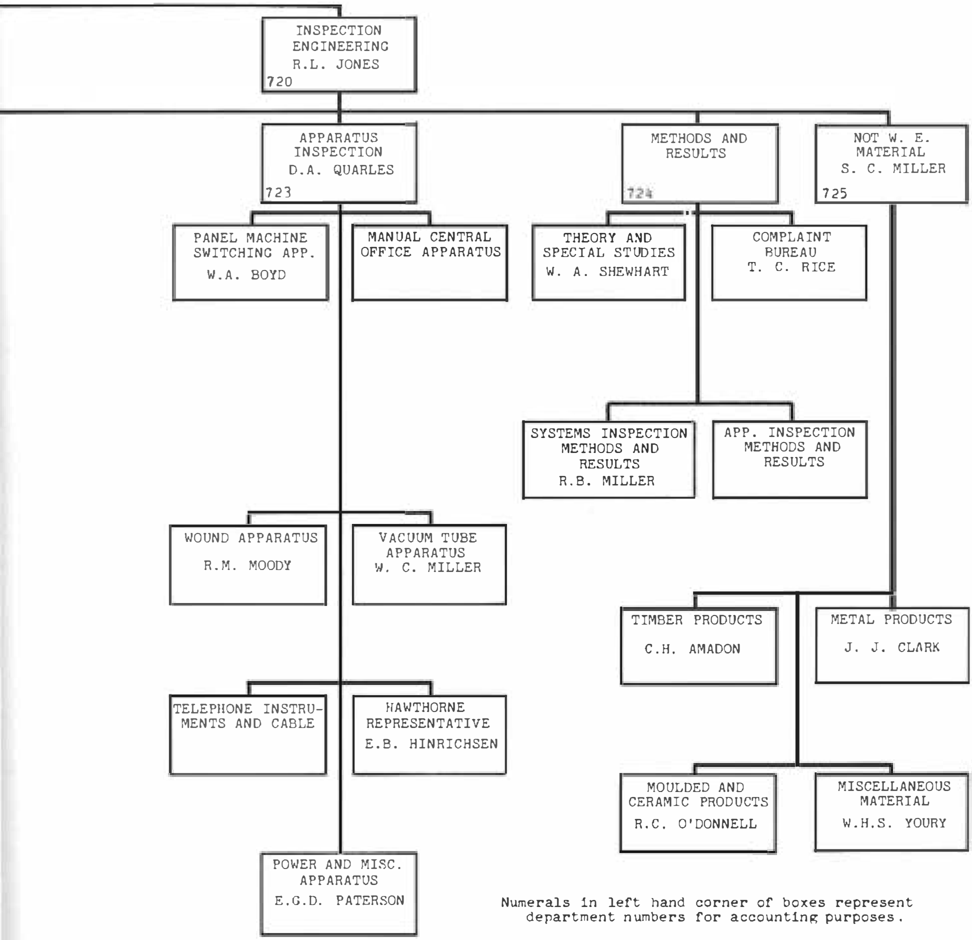


Fig. 9-1. Organization chart of the Engineering Inspection Department within the Engineering Department of the Western Electric Company, as of early 1924.



Numerals in left hand corner of boxes represent department numbers for accounting purposes.

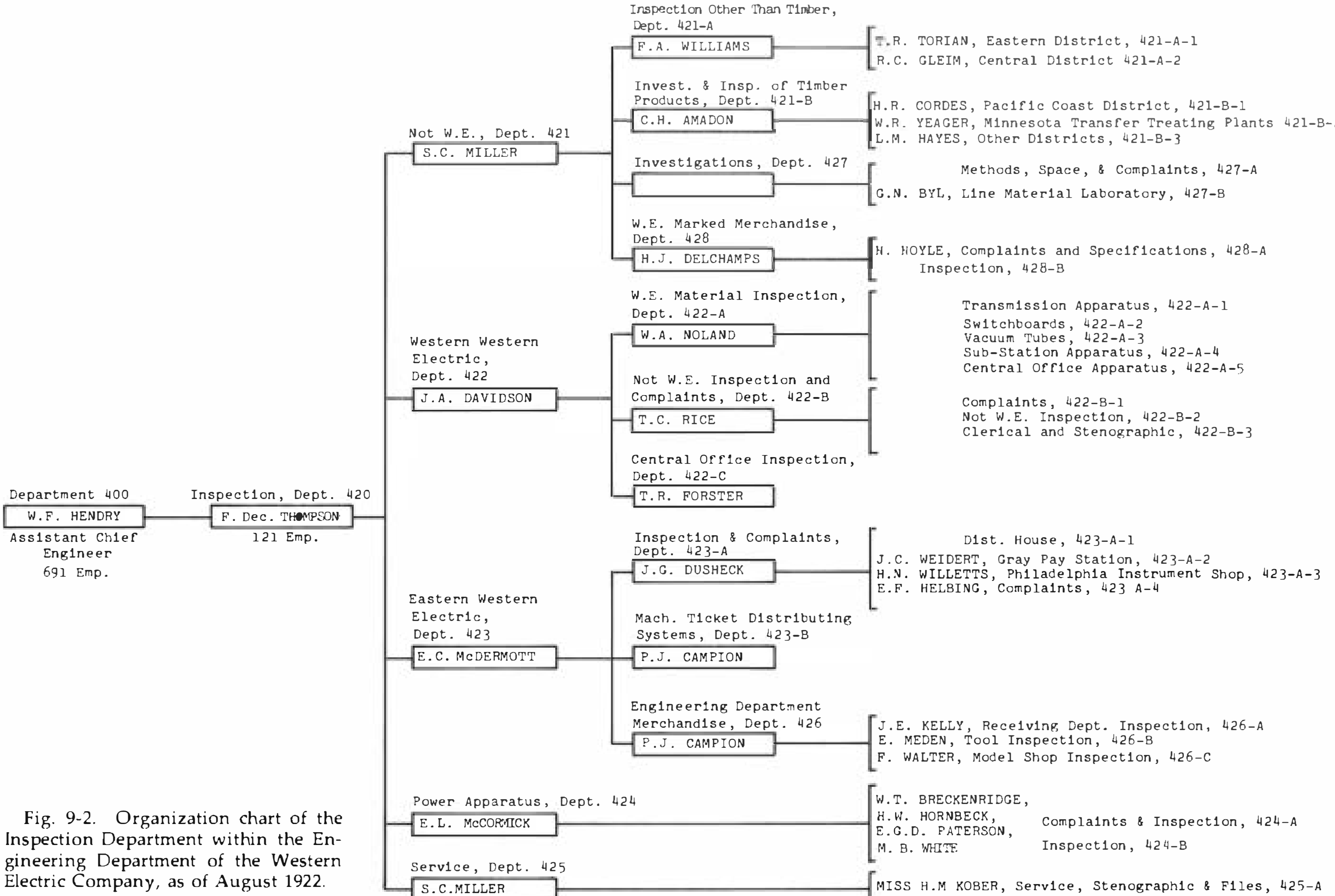


Fig. 9-2. Organization chart of the Inspection Department within the Engineering Department of the Western Electric Company, as of August 1922.

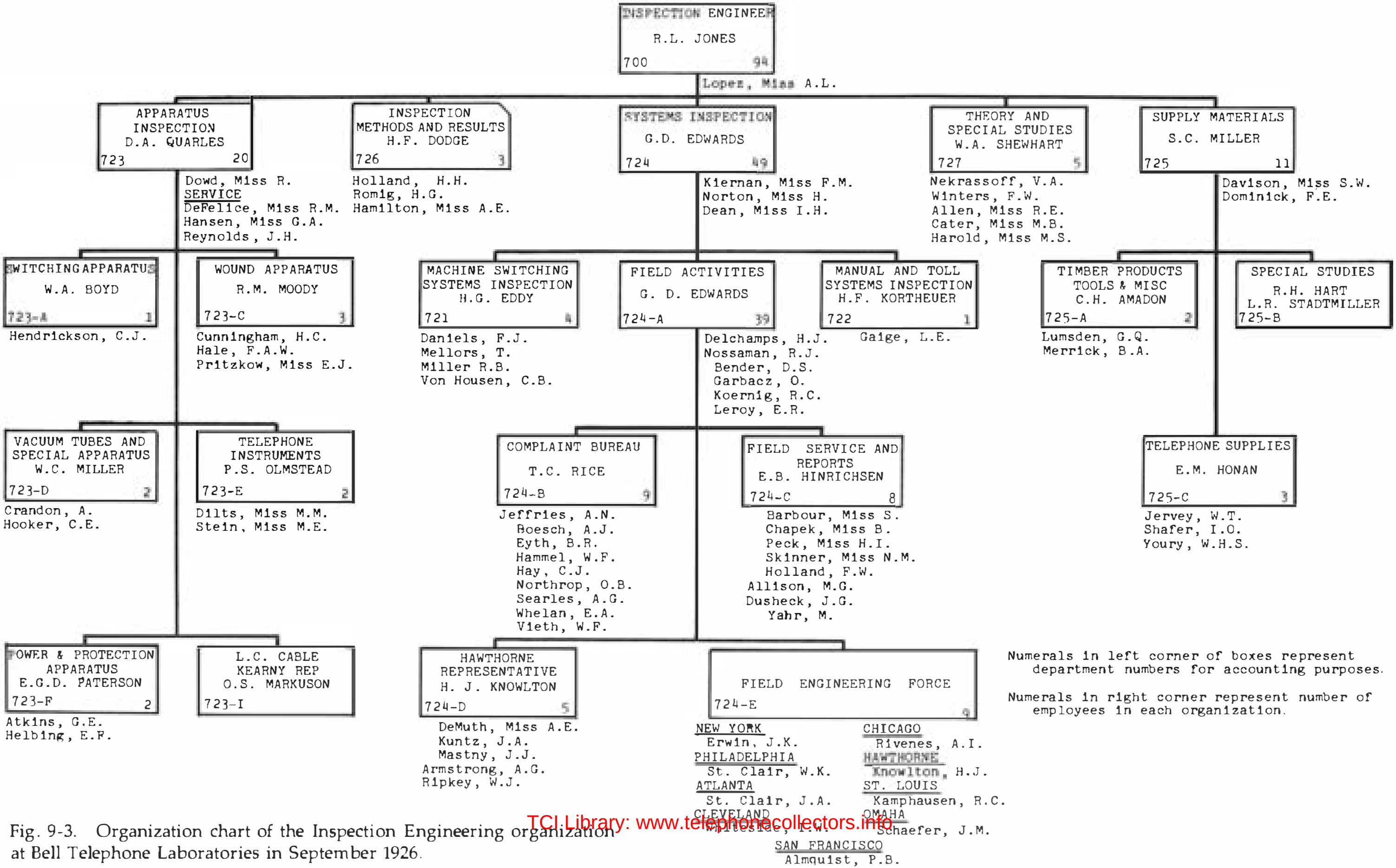


Fig. 9-3. Organization chart of the Inspection Engineering organization at Bell Telephone Laboratories in September 1926.

> 4. To study the quality and performance of telephone plant in service as an aid to further and improved developments; to give special study to instances where unsatisfactory quality is reported and to guide the steps taken to prevent recurrences of such cases.
>
> The fulfillment of these duties requires men of many talents: physicists, mathematicians; experts in glass, metals, wood, leather, paper, textiles; engineers experienced in design, manufacture, and operation.
>
> In all this work of the Inspection Engineering Department, consulting and cooperating with many departments and companies, the activities of its members must be those of team play—they have no separate laboratories, no individual problems apart from the broad problem of contributing, through cooperation, towards the best telephone instrumentalities which present day science and economies admit.

Jones' description of the function of the Inspection Engineering Department gives prominence to the development of "the theory of inspection," "methods of stating quality," and "economic standards of quality"—all mathematically oriented. Responsibility for the development in these areas was shared by W. A. Shewhart and H. F. Dodge. In describing the caliber of men required, it was made evident that the organizations headed by D. A. Quarles, G. D. Edwards, and S. C. Miller were groups of specialists in the fields of their responsibility. They were not inspectors in any sense of the word and they performed no actual inspection. Inspections were performed by the Western Electric Company and in the case of check inspection by an engineering organization separate from and not reporting to the Production Organization. The Laboratories' specialists were responsible for the development of inspection procedures defining such things as sample size, frequency of sampling, and methods of reporting data compatible with the statistical theories and techniques developed by Shewhart and Dodge.

Also, the article was more a discussion of the means for developing the Inspection Engineering function rather than a definition or description of the function itself. Jones was fully aware of this, as is indicated in the title of the article "The Viewpoint of Inspection Engineering," and confirmed by the following quotation which was italicized.

> In order that engineering inspection shall be of greatest value there must be an art and science of inspection engineering.

Undoubtedly, Jones looked upon the development of the "art and science of inspection engineering" as his major responsibility; its definition could, and must, wait for a later date. As the art and science evolved, the name was changed, appropriately, to Quality Assurance, but there is still no concise and completely descriptive definition of the function involved.

Figure 9-4 shows an organization chart dated August 1928, which continues the evolution of the Inspection Engineering Department under

Jones. It shows Edwards as Assistant Inspection Engineer and four functional departments: two concerned with the development of theory (Shewhart) and methods (Dodge), and two concerned with physical product under Edwards and Quarles. A year later (Fig. 9-5) Edwards became head of the entire organization, a position he was to hold, first as Inspection Engineer, then as Director of Quality Assurance (1940) until his retirement in 1955.

While it cannot be stated that the August 1929 chart had true long-term stability—the organization was still evolving—the four basic functions had been crystallized and were destined to continue for many years.

IV. THE FUNCTION OF THE INSPECTION ENGINEERING DEPARTMENT

We have seen how, under the general guidance of Stoll and Craft, and the direct leadership of Jones, the Inspection Engineering concept was born. The latter's statement in 1923, "A type of criticism of engineering design work never previously attempted . . . to supplement the work of the designer in a most important way," was truly a landmark. Inspection had not previously been used for such a purpose and indeed it was not Jones' intent to use inspection, as then understood, as the only tool at his disposal. In 1926, the Inspection Engineering Department employed no inspectors nor did it conduct regular inspections of apparatus, equipment, or central-office installations. Actual inspection was delegated to inspection organizations which in a broad sense were part of the various operating divisions of the Western Electric Company. These inspections were to be performed under the general direction of, and in accordance with Inspection Procedures, prepared by the Inspection Engineering Department. Furthermore, the organizations doing the actual inspection work were intended to be independent and well removed from the responsibility of meeting delivery schedules and cut-over dates. In other words, the independence they enjoyed earlier as part of the Engineering Department was to be preserved. This was, and still is, a vital point if the inspection results are to be objective.

The Inspection Engineering Department appears as a new organization of less than 100 people superimposed upon an existing setup. Its purpose was to criticize design, to subject the inspection function to a critical engineering analysis, to evaluate quality from a consumer's viewpoint, and to save money. One can easily imagine the contemporary comments: "We got along without it for 20 years, what will it do for us now?" "How does one save money by spending more?" "We do not have to be told how to inspect, we have been doing it for years." "What are all these people going to do?" "What is Inspection Engineering anyway?"

As is so frequently true, changes of this magnitude are not made

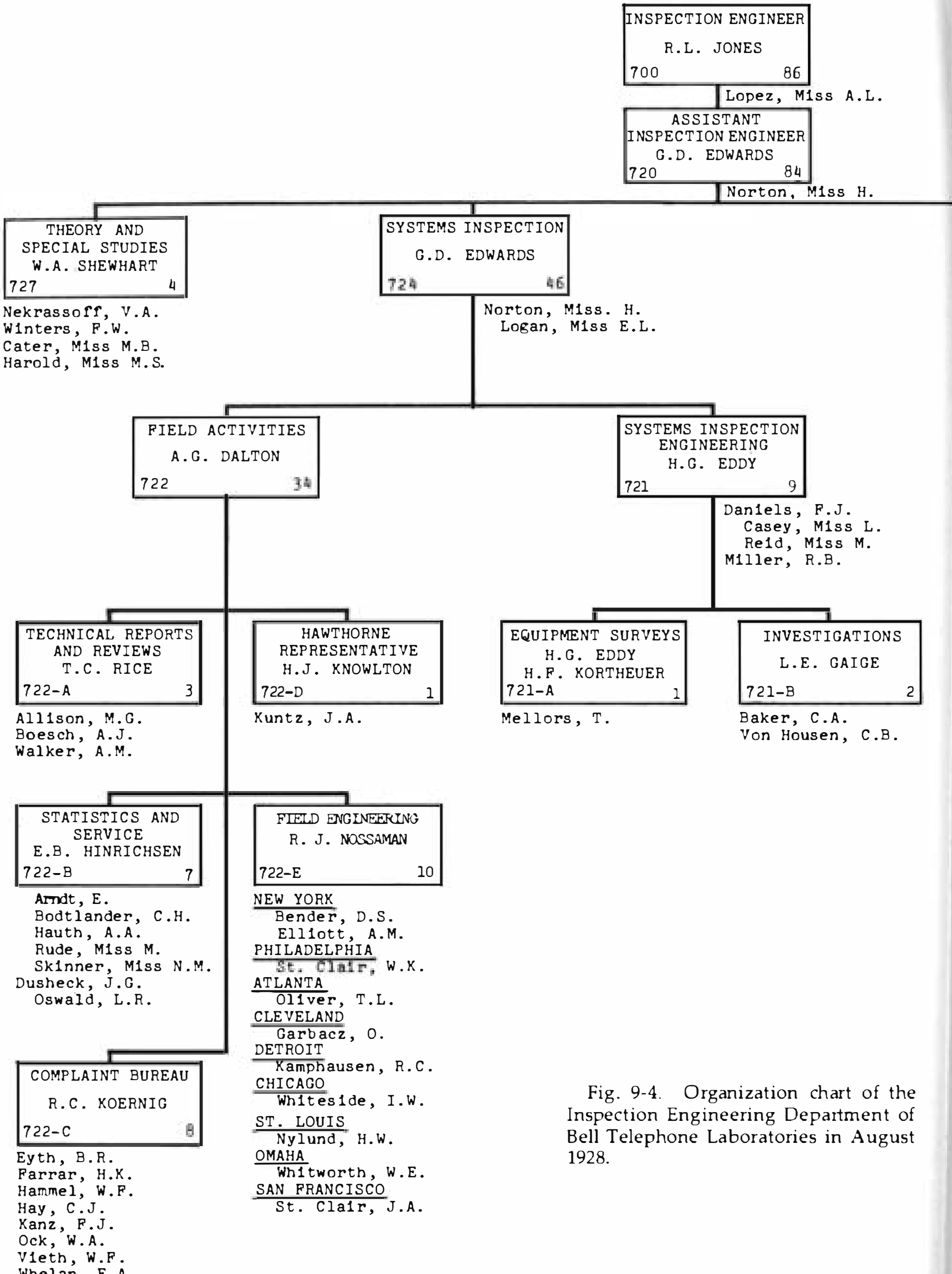


Fig. 9-4. Organization chart of the Inspection Engineering Department of Bell Telephone Laboratories in August 1928.

- **INSPECTION METHODS AND RESULTS** — H.F. DODGE — 726 — 4
  - Holland, H.H.
  - Romig, H.G.
  - Hamilton, Miss A.E.
  - Dieneman, Miss K.D.
- **APPARATUS INSPECTION ENGINEERING** — D.A. QUARLES — 723 — 26
  - Allen, Miss R.E.
  - SERVICE
    - DeFelice, Miss R.M.
    - Hansen, Miss G.A.
    - Cook, R.A.
  - **DIAL AND MANUAL APPARATUS** — W.A. BOYD — 723-A — 2
    - Hendrickson, C.J.
    - Leroy, E.R.
  - **WOUND AND SUBSTATION APPARATUS** — R. M. MOODY — 723-C — 5
    - Cunningham, H.C.
    - Grendon, A.
    - Hale, F.A.W.
    - St. James, L.N.
    - Pritzkow, Miss E.J.
  - **VACUUM TUBES AND SPECIAL PRODUCTS** — A. N. JEFFRIES — 723-D — 2
    - Hooker, C.E.
    - Yahr, M.
  - **TELEPHONE INSTRUMENTS** — P.S. OLMSTEAD — 723-E — 2
    - Dilts, Miss M.M.
    - Stein, Miss M.E.
  - **POWER & PROTECTION APPARATUS** — S.H. ANDERSON — 723-F — 2
    - Atkins, G.E.
    - Helbing, E.F.
  - **L.C. CABLE KEARNY REP.** — O.S. MARKUSON — 723-I
  - **OUTSIDE PLANT APPARATUS** — E.G.D. PATERSON — 725 — 2
    - Smith, F.I
    - Smith, R.D.

Numerals in left corner of boxes represent department numbers for accounting purposes.

Numerals in right corner represent number of employees in each organization.

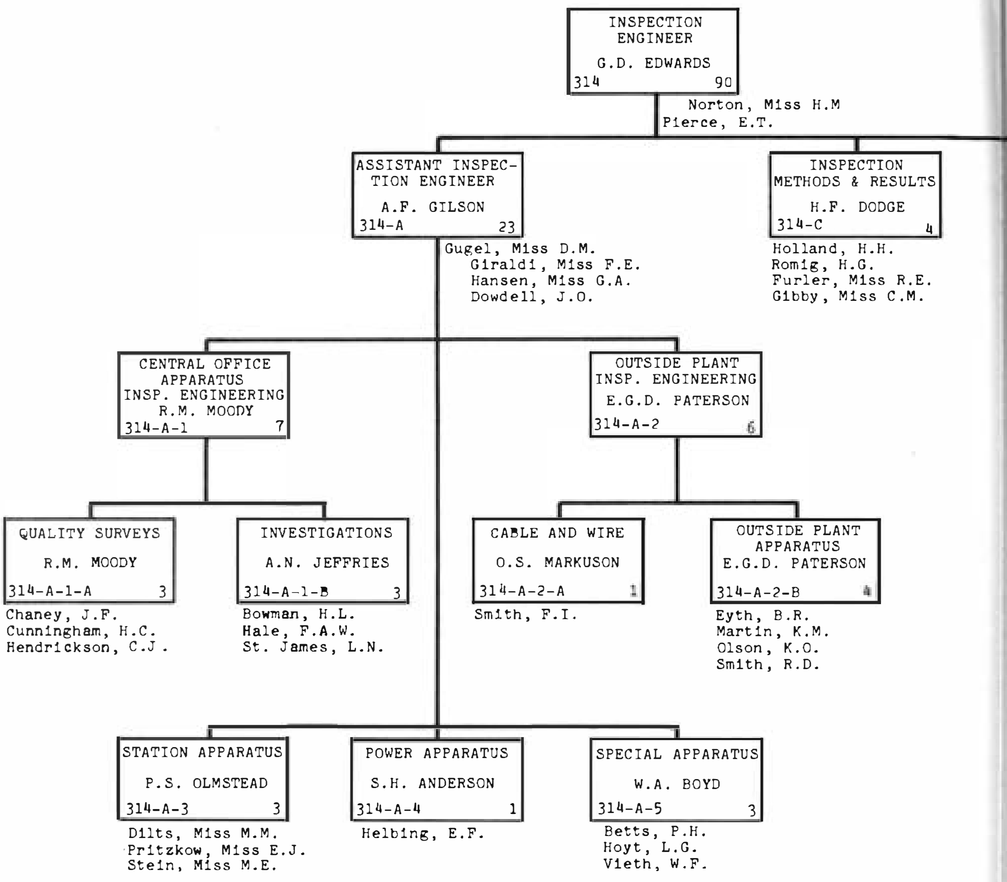


Fig. 9-5. Organization chart of the Inspection Engineering Department of Bell Telephone Laboratories in August 1929.

INSPECTION THEORY & SPECIAL STUDIES
W. A. SHEWHART
314-D 4

Winters, F.W.
Cater, Miss M.B.
Harold, Miss M.S.
Rodgers, Mrs. A.R.

SYSTEMS INSPECTION ENGINEER
A.G. DALTON
314-B 47

Newman, Miss H.M.
Logan, Miss E.L.

FIELD ACTIVITIES
A.G. DALTON
314-B-1 36

SYSTEMS INSPECTION ENGINEERING
H.G. EDDY
314-B-2 8

Daniels, F.J.
Lane, Miss E.D.
Miller, R.B.

FIELD ENGINEERING
R.J. NOSSAMAN
314-B-1-A 16

Bateman, I.L.
Bonnesen, E.J.
Craig, H.M.
Farrar, H.K.
Knapp, F.H.
Shepard, J.H. Jr.

NEW YORK NO. 1
Bender, D.S.

NEW YORK NO. 2
Elliott, A.M.

PHILADELPHIA
Whiteside, I.W.

ATLANTA
Oliver, T.L.

CLEVELAND
Garbacz, O.

DETROIT
Kamphausen, R.C.

CHICAGO
Johnson, C.A.

ST. LOUIS
Nylund, H.W.

OMAHA
Whitworth, W.E.

SAN FRANCISCO
St. Clair, J.A.

TECHNICAL REPORTS AND REVIEWS
T.C. RICE
314-B-1-C 2

Allison, M.G.
Wean, M.J.

COMPLAINT BUREAU
R.C. KOERNIG
314-B-1-B 15

TELEPHONE PRODUCTS
Boesch, A.J.
Baker, C.A.
Gallup, W.S.
Walker, A.M.
Whelan, E.A.

SPECIAL PRODUCTS
Hay, C.J.
Hauth, A.A.
Kanz, F.J.

STATISTICS AND FILES
Daly, W.L.
Bodtlander, C.H.
Pitt, H.E.
Rude, Miss M.A.
Smith, Miss N.C.
Staehler, G.J.
Dusheck, J.C.

QUALITY SURVEYS
H. G. EDDY
314-B-2-A 2

Hagenbuck, R.O.
Kortheuer, H.F.

INVESTIGATIONS
L.E. GAIGE
314-B-2-B 2

Hammel, W.F.
Von Housen, C.B.

Numerals in left corner of boxes represent department numbers for accounting purposes.

Numerals in right corner represent number of employees in each organization.

from the bottom up, but from the top down. If the new concept had not been instigated by Stoll and Craft with the endorsement of Harrison of AT&TCo, the function of Inspection Engineering might not have seen the light of day for several years. Recall that no one had yet proposed a concise and unambiguous definition or even a complete description of the function. The intent, of course, was that the new organization would develop the science and art of a new discipline —Inspection Engineering—and apply it to so control and assure the quality of product delivered that Operating Company inspection would be unnecessary. Earlier attempts to eliminate Operating Company inspections had been unsuccessful.

The earliest attempt to orient the Inspection Engineering function within Bell Telephone Laboratories and to establish the Laboratories' responsibilities with respect to the Western Electric Company and the Operating Telephone Companies was covered in General Executive Instruction (GEI) 1.3 issued over the name of E. B. Craft, Executive Vice-President, on December 31, 1925. This document was titled: RESPONSIBILITY FOR QUALITY OF PRODUCTS FURNISHED BY WESTERN ELECTRIC COMPANY, INCORPORATED. In order to fully convey its intent and implications, it is necessary to quote rather extensively:

> 1.1 In accordance with understandings between the Bell Telephone Laboratories, Incorporated (hereinafter referred to as the Laboratories), and the Western Electric Company, Incorporated (hereinafter referred to as Western Company), the Laboratories are responsible for defining the standards of quality for materials, apparatus, equipment and systems (hereinafter referred to as products) standardized for purchase, manufacture and/or installation and furnished by the Western Company to the Bell System.
>
> . . .
>
> 1.3 This instruction outlines the division of functions and responsibilities between the Laboratories and the Western Company in respect to the quality of products.
>
> 2. RESPONSIBILITY OF LABORATORIES
>
> 2.1 This Company shall be responsible for:
>
> 2.11 Development and standardization of products for purchase, manufacture and/or installation by the Western Company;
>
> 2.12 Prescription of the standards of quality by:
>
> 2.121 Preparation of specifications for products covering design and essential requirements;
>
> 2.122 Examination and approval of tool-made samples of new products;
>
> 2.123 Examination and approval of initial installations of new products.

One might infer from paragraph 2.12 and subparagraph 2.121 that "standards of quality" referred exclusively to the engineering requirements included in design specifications. Such an inference would not be in accord with the terms "quality standards" and "standard quality" used a short time later but, at this point, the meaning is not too clear. GEI 1.3 continues:

> 2.13 Development of manufacturing or inspection methods when necessary, coincident to the development of new or improved products or as an aid in expressing design requirements . . .;
>
> 2.14 Systematic review of the characteristics of products in relation to service requirements, . . .
>
> As a means to this end it shall:
>
> 2.141 Make studies and analyses of reports of check inspections made by the Western Company, showing the quality of the current product;
>
> 2.142 Make studies reviewing with the Western Company, as by co-operative surveys, design and performance requirements and their interpretation, methods of inspection being employed, and quality results obtained.
>
> 2.15 Making available mathematical principles which will be of assistance in establishing design requirements, in determining the extent and character of inspections necessary to insure that inspection results are indicative of the quality of the whole, and in arriving at the conclusions to be drawn from the results of inspections;
>
> 2.16 Study of the performance of products in service, as an aid to development;
>
> 2.17 Investigation of all customer's complaints on quality of products except as provided in paragraph 2.172;
>
> . . .
>
> [The exception concerns customer's inspections of new installations which complaints are to be referred directly to the Installation Branch.]
>
> 2.18 Advising the Western Company whether or not products should be shipped in cases where they fail in some respect to meet prescribed requirements or are questionable in a characteristic not covered by the specifications, but where special considerations may make their use desirable or permissible;
>
> 2.19 Furnishing such reports as are made to the American Telephone and Telegraph Company on the quality of products.

It appears that in GEI 1.3, the responsibilities and obligations of the new Inspection Engineering Department were intermingled with the responsibilities of other departments in Bell Telephone Laboratories. However, if one looks more closely, only paragraph 2.11 and (with the possible reservation already noted concerning the intended meaning of "standards of quality") paragraphs 2.12 and 2.121 apply specifically to departments other than the Inspection Engineering Department. The remaining paragraphs of Section 2 can be looked upon, and indeed were construed at the time, as a charter for the Inspection Engineering Department. The conclusion is inescapable that, at least from this point on, the Laboratories shared heavily in the responsibility for assuring the quality of products supplied to the Operating Companies by Western, and that discharge of this responsibility was vested in the new Inspection Engineering Department under R. L. Jones.

With this responsibility charged to the Inspection Engineering Department in 1925, the definition of quality should be considered. It has already been indicated that the term "standards of quality" seemed to refer to the engineering requirements in the design specifications. In a paper prepared for use in the department, "The Control of Quality of Manufactured Product," dated March 1929, Shewhart states:

> Dating at least from the time of Aristotle, there has been some tendency to conceive of quality as indicating the *goodness* of an object. . . . advertisers of the present day appeal to the public upon the basis of the quality of product. . . . they implicitly assume that there is a measure of goodness which can be applied to all kinds of products . . .

Thus Shewhart looks upon even the traditional or popular concept of quality as implying a measurement and therefore, the existence somewhere of a means of measurement. This thought appears to underlie the use of the word quality in the title of GEI 1.3.

Shewhart then discusses the "Quality of a Thing as a Variable":

> The quality of a thing is that which is inherent in it and we cannot alter the quality without altering the thing.
>
> Going a little deeper we see that possibly without exception every conceptual something is really a group of conceptions more elementary in form. The minimum number of conceptions required to define an object may be called the qualities thereof and this definition is consistent with that given by JEVONS [W. S. Jevons, *The Principles of Science*]. The mind learns to regard each object as an aggregate of qualities and acquires the power of dwelling at will upon one or other of those qualities at the exclusion of the rest.
>
> For our purposes we shall assume that, had we but the ability to see, we would find a very large number of different characteristics required to define what even the simplest thing really is.

Within this framework we can think of "standards of quality" (paragraph 2.12 of GEI 1.3) as being limited to an engineering definition of the attributes of an article that must be measured to determine its quality—for example, capacitance, insulation resistance, and physical dimensions of a capacitor. How collective measurements of such quality characteristics were to be interpreted in evaluating the quality of a product was, in 1925, a process yet to be developed. Further discussion along these lines would be out of place at this point. The brief discussion of quality given is intended to indicate the many sources of confusion and misunderstanding in this new field caused by the use of inadequate, ordinary terms.

V. THE CONTROL CHART

"The object of this note is to emphasize what appears to be a comparatively new field of applications of statistical methods." With this opening sentence, Shewhart began his paper, "The Application of Statistics as an Aid in Maintaining Quality of a Manufactured Product," published in the December 1925 *Journal of the American Statistical Association*. In it, he discussed the question of limits of variation and the mathematical methods for determining when a product was being manufactured under a constant system of causes and how significant deviations in the quality of product could be detected and brought under control. The tool which he had conceived for the first scientific approach to maintaining uniform quality in manufacture was the Control Chart. It was one of the most valuable and versatile devices that emanated from the new art and science of Inspection Engineering. The concept was both simple and elegant, perhaps even elegant in its simplicity, while being most remote from contemporary thinking. Within 25 years or so, it had become so commonplace, so inseparable a part of a good economical manufacturing operation, particularly where precision products were involved, that it was difficult to realize it had not always been around.

Shewhart's first Control Chart was attached to a memorandum dated May 16, 1924, addressed to Jones. The chart, together with Shewhart's memorandum, is shown here as Fig. 9-6. (The same items were reproduced in the Shewhart Memorial Issue, of August 1967, of *Industrial Quality Control*.) The chart gives upper and lower limits, now generally called upper control limit (UCL) and lower control limit (LCL). The brief memorandum gives no clue as to the mathematical determination of these limits or their philosophical justification. The memorandum does state, however, that when a point goes outside the limits, the product is not satisfactory. This is almost too strong a statement until one considers its framework. At this point, it should be remembered that Shewhart chose his words carefully; he did not say unserviceable. In a Memorandum for File of the same date, Shewhart wrote:

> The observed values of the fractional number p having a given characteristic may, however, be expected to vary within certain limits, as a

> result of uncontrolled random variations in the manufacturing process. The report should indicate, therefore, if the observed fluctuations of p are within these limits . . . The usefulness of the chart depends upon the significance of the lines drawn thereon (upper limit and lower limit) and not upon the method of drawing the lines.

In an article a few months earlier, "Some Applications of Statistical Methods to the Analysis of Physical and Engineering Data," published in the January 1924 *Bell System Technical Journal*, Shewhart states:

> Statistical methods alone do not answer all of the questions that are raised in this problem nor do they answer them in many others. There is almost always room for judgment to enter.

It is evident that the physicist Shewhart did not consider statistics to be a substitute for thought but rather an adjunct which could assist in directing thought along fruitful lines.

In the immediately ensuing years, Shewhart wrote many memoranda covering his work and published several papers. The first paper specifically titled "Quality Control Charts" was published in the October 1926 *Bell System Technical Journal*. In this paper the philosophical basis for the Control Chart is made clear, sometimes with quite homely examples. Liberal quotation is justified since many users of the technique have lost sight of the original intent.

First, Shewhart draws a nice distinction between "assignable" and "nonassignable" causes, a concept which is vital:

> To make clear the significance of the terms "assignable causes" and "nonassignable causes," we may make use of the following illustration. Suppose a person were to fire one hundred rounds at a target. We know what probably would happen—the individual would not hit the bull's-eye every time. Possibly some of the shots would fall within the first ring, others within the second ring, and, in general, the shots would be distributed somewhat uniformly about the center of the target. We have a more or less definite picture of some of the possible reasons why the individual would not hit the bull's-eye every time, but we probably cannot assign the reasons or causes for his missing the bull's-eye in any particular instance—the causes of missing are nonassignable. Suppose, however, that the individual tended to shoot to the right of the bull's-eye. Naturally we would conclude that there was some discoverable cause for this general tendency, i.e., we would feel that the observed effect could be assigned to some particular cause.

He then presents the essence of the problem of controlling the quality of a product, as he sees it, within this framework:

> Here then is the practical commercial problem—When do the observed differences between the product for one period and that for another indicate lack of control due to assignable causes, and when, on the other hand, do the differences in quality of manufactured product observed from one

Case 18013

WAS-724-5-16-24-FQ

Mr. R. L. JONES:-

A few days ago, you mentioned some of the problems connected with the development of an acceptable form of inspection report which might be modified from time to time, in order to give at a glance the greatest amount of accurate information.

The attached form of report is designed to indicate whether or not the observed variations in the percent of defective apparatus of a given type are significant; that is, to indicate whether or not the product is satisfactory. The theory underlying the method of determing the significance of the variations in the value of p is somewhat involved when considered in such a form as to cover practically all types of problems. I have already started the preparation of a series of memoranda covering these points in detail. Should it be found desirable, however, to make use of this form of chart in any of the studies now being conducted within the Inspection Department, it will be possible to indicate the method to be followed in the particular examples.

W. A. SHEWHART.

Enc.:
Form of Report.

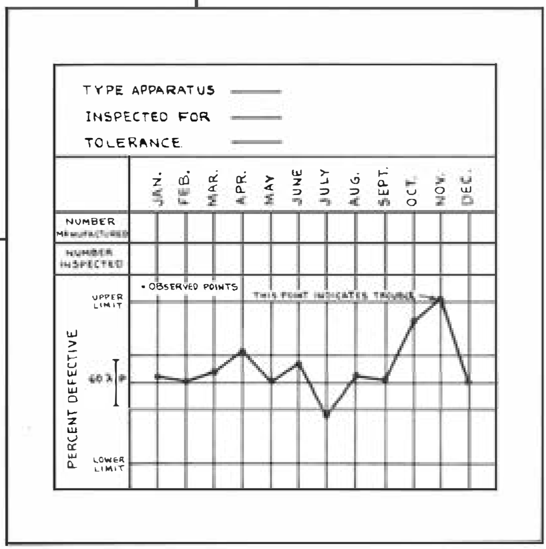


Fig. 9-6. W. A. Shewhart's first Control Chart and accompanying memorandum of May 16, 1924.

period to another indicate only fortuitous, chance or random effects which we cannot reasonably hope to control without radically changing the whole manufacturing process?

The reader is then led, almost literally by the hand, through a process for determining a frequency distribution from actual observed data for one characteristic of a telephone transmitter. After obtaining this distribution, Control Charts are set up for such statistics as average, standard deviation, skewness and kurtosis, using three-sigma limits and the values of each of these statistics, calculated from each month's data. It is evident that Shewhart was thinking of very large samples taken over a substantial period of time which required a search for relatively small changes in the distribution from month to month. The application of the Control Chart concept to very small samples, four or five, was not proposed until some time later.

VI. A PRELIMINARY NOTE ON SAMPLING INSPECTION

Inspection by attributes, the examination and test of individual items of production to classify them as either good or bad, was the usual type of inspection employed by the Western Electric Company. As has been mentioned earlier, sampling was resorted to frequently to reduce inspection cost. What was needed was some program of sampling inspection which would minimize inspection cost and guarantee product of a quality acceptable to the user.

Perhaps the earliest documentation of a suggested plan for attacking the sampling inspection problem is found in Memorandum for File I.M. 118, dated July 22, 1925, by H. F. Dodge, who directed the group responsible in this area. This memorandum opens by stating that the consumer specifies the requirements needed for service and translates these requirements into specifications. The manufacturer endeavors to meet these requirements at the lowest possible cost. One item of this cost is assuring, by inspection, that the product meets the consumer's requirements. Sampling inspection, therefore, is dictated by economy. Dodge then discusses the use of *a posteriori* as opposed to *a priori* probabilities in solving this problem and develops reasons for proposing the latter. The Shewhart influence in setting up the basic structure is evident in the following, quoted from I. M. 118:

> With a controlled process, the average composition of the product is known. Individual lots produced by this controlled process may then be considered as samples from a single universe, and as such their compositions will deviate from the "process average" in accordance with the fluctuations of sampling.
>
> In sampling inspection, we inspect only a portion of each lot. These inspection samples can be considered as smaller samples from the same universe. The composition of the inspection samples will likewise deviate from the "process average" in accordance with the fluctuations of sampling. This provides a basis for drawing conclusions as to the probability of occurrence of samples having any stated composition, and permits us to set up criteria for accepting individual lots.

Within this framework, the memorandum develops and defines the basic concepts of consumer's and producer's risks. (These definitions have remained essentially unaltered except for the more or less general dropping of the "A" and the frequent substitution of "product" or "process" for "lot" in the second definition.)

> The Consumer's "A" Risk is the probability of finding not more than the acceptance number of defects in a sample drawn from a lot containing the tolerance fraction defective.

> The Producer's "A" Risk is the probability of finding more than the acceptance number of defects in a sample drawn from a lot containing the process average fraction defective.

In essence, then, the consumer's risk is the probability of acceptance for some stated value of "poor" quality, while the producer's risk is the probability of rejection for a stated value of "good" quality. Incidentally, Samuel S. Wilks in his book *Mathematical Statistics*, page 397, states, "The concepts of producer's and consumer's risks were, in fact, the forerunners, respectively, of risks of Type I and Type II errors," used in the testing of statistical hypotheses.[3]

Subsequent accomplishments over the next few years are well described by H. F. Dodge and H. G. Romig in the Introduction to their classic book *Sampling Inspection Tables*:

> The use of these two risks alone for obtaining a sampling plan was insufficient to the needs of the problem. They merely provide the equivalent of a statistical test for a unique sample from a unique lot. What was required was the inclusion of some of the ideas contained in the theory of "quality control" then under development in the Bell Telephone System. These ideas involved the treatment of lots not individually but as a series. Such a treatment provided a basis for securing certain fundamental and over-all advantages by an adjustment of inspection procedure to the actual quality conditions existing at any time. One of the more important practical advantages is "a minimum amount of inspection for a given inspection procedure," which is one of the salient features of the tables in this book.
>
> Shortly thereafter, in 1926, in a program of active cooperation between the Laboratories and the Western Electric Company,the first set of sampling inspection tables was prepared for shop use. These tables provided a complete set of "minimum inspection" solutions for several selected levels of process average quality using the lot tolerance concept with a 10 percent Consumer's Risk. Studies had shown, however, that a double sampling procedure offered both theoretical and practical advantages. These tables were accordingly prepared for double sampling as well as single sampling covering ranges of conditions dictated by the needs of actual shop practice.

Thus began a period of great creativity at Bell Telephone Laboratories with the development of the basic concepts of statistical quality control and its application to specific fields such as sampling inspection theory. As will be shown in a later volume of this history, with the application of the principles developed by Shewhart, Dodge, Romig, and their associates in later years, a new technology was created which rapidly spread through the industrial community in the United States and many other countries. A new technical society was later formed,

[3] In statistics, a Type I error is the probability of rejecting the null hypothesis when it is true. A Type II error is the probability of accepting the null hypothesis when it is false.

the American Society for Quality Control, and the field grew in scope to become a powerful management tool to assure efficient production of high-quality products.

VII. CONCLUDING REMARKS

The need for sound economic means for insuring that the Operating Companies were supplied with product adequate to meet service demands was recognized at the earliest inception of the Bell System. The foregoing highlights major innovations and developments of the then unnamed Quality Assurance function and shows that the goals of the function were well understood by 1925 when Bell Telephone Laboratories was incorporated.

At that time the department responsible for the function was still called "Inspection Engineering Department" and it is shown how the organization structure evolved during the first few years following incorporation. However, the required engineering and mathematical developments had just started, so that a more complete discussion of the art and science of Inspection Engineering will be reserved for a later volume of this history. In that volume we will also trace all other developments up to the present in the Quality Assurance area, mainly from the vantage point of Bell Telephone Laboratories, of which the Inspection Engineering Department formed a part since 1925.

Chapter 10

The Spirit of Research

"Research is the effort of the mind to comprehend relationships which no one has previously known. And in its finest exemplifications it is practical as well as theoretical; trending always toward worthwhile relationships; demanding common sense as well as uncommon ability." H. D. Arnold's definition of Research characterized a movement initiated even before his time, when a few farsighted executives sensed the potency of the scientific method and commenced enlisting trained scientists to work on fundamental problems. Thus there emerged a generation of men for whom an industrial setting, far from impeding scientific inquiry or diluting its fundamental quality, proved a stimulant to prolific contributions to science.

How quickly these contributions came to advance the telephone art, beginning about the turn of the century, we have described in previous chapters. In this concluding chapter, we discuss Bell System scientific research in a more specific way: that is, who the individuals were, what were their backgrounds, and how they were motivated, still working on the frontiers of science, to collaborate and direct their insights against challenges posed by the new art of telephony.

I. INTRODUCTION

Alexander Graham Bell was not a trained scientist. His discoveries resulted more from insight, persistence, and a keen sensitivity to human needs. Actual commencement of scientific research in the telephone industry is more appropriately dated from late 1885 when Hammond V. Hayes joined the staff of the American Bell Telephone Company in Boston. Hayes, a graduate of Harvard, had studied electrical engineering at the Massachusetts Institute of Technology and had then received one of Harvard's earliest doctorates in physics. Hayes was quick to sense the ultimate technical complexity of telephony and to perceive that its scientific roots must extend into deeper soil than could be cultivated with the primitive tools his telephone associates were using.

The period immediately preceding had produced many notable engineering advances—the first skyscraper, in Chicago; the telephone, in Boston; the electrified subway, in London; the electric trolley car, in Baltimore; the electric elevator, in New York; the incandescent bulb, in Menlo Park, New Jersey—all within a space of a dozen years. Yet, with the universities considered to be the primary sources of new scientific knowledge, there prevailed a science-technology gap that demanded reform in the relations between theorists and the so-called practical men. Thomas Edison's laboratory, organized in 1876 and responsible for many extraordinary inventions, had been characterized more by the try-and-try-again (so-called Edisonian) technique. In the General Electric laboratory at Schenectady, established in 1892 and stemming from Edison's work, we see some of the earliest signs of the scientific approach. Likewise, in 1889 the United States petroleum industry employed its first professional chemist, William Burton, who devised a once-important oil cracking process; and in 1902, E. I. du Pont de Nemours instituted organized chemical research, the first company in its field to do so.

These beginnings, however, were not viewed everywhere without some skepticism and suspicion, and the general extension of scientific methods into arts and industries was a slow process. As late as 1916, J. J. Carty, Chief Engineer of the American Telephone and Telegraph Company, in his presidential address at the 33rd Annual Convention of the American Institute of Electrical Engineers, remarked that:

> While many concerns in America now have well-organized industrial research laboratories, particularly those engaged in metallurgy and dependent upon chemical processes, the manufacturers of our country as a whole have not yet learned of the benefits of industrial scientific research and how to avail themselves of it.

Colonel (later Brigadier General) Carty went on with the following statement concerning research in his own industry:

> Industrial research, conducted in accordance with the principles of science, is no new thing in America. The department which is under my charge, founded nearly forty years ago to develop, with the aid of scientific men, the telephone art, has grown from small beginnings with but a few workers to a great institution employing hundreds of scientists and engineers; and it is generally acknowledged that it is largely owing to the industrial research thus conducted that the telephone achievements and development in America have so greatly exceeded those of other countries.

Carty additionally remarked that though the development of electric lighting, electric power, and electric traction had come later than the invention of the telephone, nevertheless some of the larger electrical manufacturing concerns had founded industrial scientific research laboratories which had already attained worldwide reputation and had brought returns to their sponsors, in the form of profitable improvements in the art, worth many times their cost.

We have cited the enlistment of Dr. Hammond V. Hayes in 1885 as representing the initiation of genuinely scientific approaches to the problems of telephony. Indeed, looking back to the earliest research efforts within industry, we find it difficult to cite a more clearly illustrative case of sheer intellectual power, brought to bear on an urgent and baffling problem, than the fundamental approach to telephone transmission encouraged and sponsored by Hayes.

Accordingly, our discussion of telephone research begins with early research in *electrical communication theory* and the supporting mathematics associated with electrical transport and distribution of voice signals over wires, in networks, and in space. We then consider *speech, hearing, and sound* and the penetrating researches in these fields—including their subjective aspects, which were vital in telephony and yet unapproachable for investigation until the telephone itself became available as an instrument of measurement. Finally, recognizing the critical importance of *materials* and their properties in both the physical makeup and the performance of the vast telephone plant, we trace the stepping-stones in chemistry and metallurgy and the advances in physics, up through the mid-1920s, which were fundamental to the understanding and control of materials behavior and the creation of new and better materials.

II. RESEARCH IN ELECTRICAL COMMUNICATION THEORY

It has been obvious, in our discussions in earlier chapters on the work of Bell and his immediate successors, that while they knew something about mechanics and acoustics and about resonance phenomena associated therewith there was scarcely a suspicion that these phenomena had precise electrical analogues that would have to be understood before telephone transmission could advance beyond its primitive stages.

The British physicist Sir William Thomson, afterwards Lord Kelvin, whose early studies in transmission we have already noted in Section 4.1.3 of Chapter 4, had published in 1853 a mathematical analysis of the discharge of a condenser (capacitor) through a circuit containing resistance and inductance. His solution to the differential equation had shown that the discharge could be oscillatory, as had been suggested earlier by Helmholtz (Berlin, 1847). Such oscillations had been observed experimentally by others, but the reasons for them had been obscure.

Fifteen years passed, however, before Maxwell (in 1868) performed a similar analysis in which he assumed, instead of a condenser discharge, the application of a sinusoidal electromotive force, and showed the conditions for resonance or maximum current. Maxwell, in this exposition, provided the tie-in with Newton's mechanics of the previous century by likening the resistance of the circuit to the effect of a viscous fluid in which a body is made to move backwards and forwards; by likening the inductance of the electromagnetic coil to the mass of a large boat resisting efforts to accelerate it alternately in opposite directions; and by comparing the condenser to a railway buffer resisting the impingement of the "carriage" against a fixed obstacle. This analysis by Maxwell enhanced the reputation already achieved through his electromagnetic field equations of 1864, establishing him as the Newton of nineteenth-century physics.

Thus the knowledge underlying alternating-current electric-circuit theory was coming to light nearly 20 years before the arrival of the telephone, yet possessed by scarcely anyone except the advanced physicist; and even in 1884, the year of the founding of the American Institute of Electrical Engineers,[1] the inductance coil remained a "choak" to alternating currents while the condenser or "Leyden" was mainly used to store and discharge static electricity. About that time, however, it had been noted that the output of an alternating-current generator could be increased by connecting a condenser in series; and the British physicist Hopkinson, in explaining this observation on the basis of Maxwell's circuit analysis of 1868, pointed out that under resonance conditions the potential across the condenser might be many times the generated voltage. Startling as this must have appeared, it failed to attract much attention from practicing engineers, many of whom were preoccupied with direct-current problems and satisfied to let such an intricate and apparently abstruse matter lie at rest until its consideration might be forced on them. But it was soon afterwards that Hertz, in his high-frequency experiments already described in Section I of Chapter 5, made practical use of the resonance phenomenon (Fig. 10-1) in adjusting his circuits to draw the maximum length of spark in his wave-detecting loop. And by about 1890, what appeared a few years earlier to have been merely an interesting laboratory experiment, having no further application than perhaps the breaking down of a condenser, was developing some practical aspects to which the advent of alternating-current power systems was forcing engineers to give attention. For example, the

[1] AIEE was founded in that year "to promote the Arts and Sciences connected with the production and utilization of electricity, and the welfare of those employed in these industries; by means of meeting for social intercourse, the reading and discussion of professional papers, and the circulation, by means of publications among its members and associates, of the information thus obtained." Among its charter members, in addition to power and telegraph engineers, were some dozen telephone men, including Alexander Graham Bell, who was president of the Institute in 1891–92.

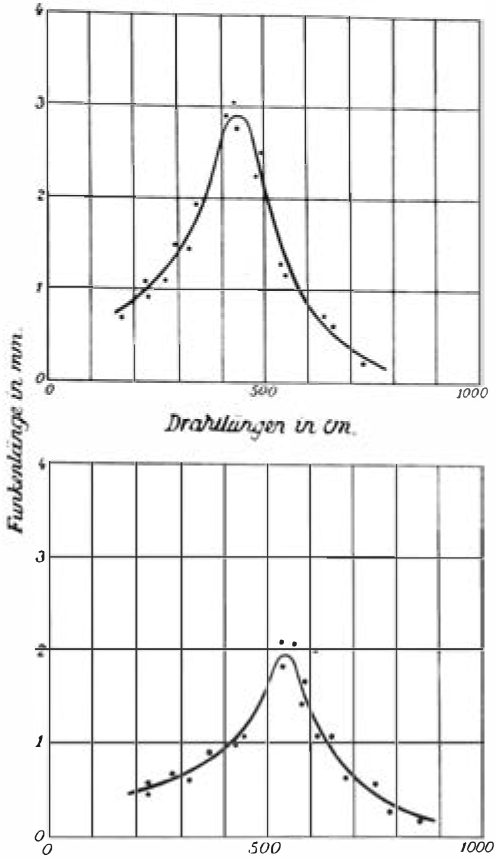


Fig. 10-1. Earliest known electrical resonance curves (from Hertz's experiments in 1887).

French power engineer Boucherot brought to light a highly interesting property of a circuit (Fig. 10-2) composed of a coil and a condenser, each of X ohms reactance at the frequency employed, whereby an alternating voltage E applied at the input terminal will produce an alternating current of E/X amperes at the output terminal, regardless of the impedance connected thereto; and this property found important

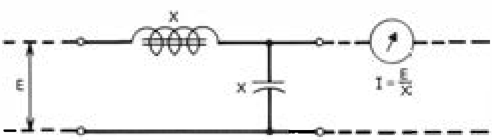


Fig. 10-2. An early example of resonant circuit theory: one of the Boucherot "constant current" networks delivering a current that is independent of the terminating impedance.

uses, for example, in the series-connected arcs that were used in early street lighting.

Apart from such special applications, resonance was generally something to be avoided under the constant-frequency, constant-potential conditions that were sought in power engineering and electric lighting. But for scientists and engineers whose interest was in electrical communication—and more specifically, telephony, where response to a wide spectrum of frequency and range of amplitude is a *sine qua non*—the thought was dawning that the resonant network with its dynamical character was the electrical analogue of their acoustomechanical contrivances, and far more adaptable. It was apparent that the attempts of Bell, Gray, and others to develop the "harmonic telegraph" might have gotten further had they appreciated the equivalence of coils and condensers to the masses and springs of their vibrating instruments. Likewise, in limited circles there was an appreciation of the more generalized form taken by Ohm's law and Kirchoff's laws, and by such structures as Wheatstone's bridge, with the introduction of reactive circuit elements and the inclusion of frequency as a parameter.

It was into this livening atmosphere of electrical theory that Hammond Hayes found himself projected in 1885 as he took charge of the experimental work at the American Bell Telephone Company in Boston; and he must have made the observation that much of the advance on the theoretical side was being made in foreign countries, despite the slow start which actual telephone usage was making in these countries because of the entrenched position of telegraphy in the postal administrations. We cite two further examples of advancement in electrical theory abroad which undoubtedly came to Hayes' attention:

(*i*) a network theorem (1883) of L. Thévenin,[2] which was to become fundamental in developing transmission network theory, and

(*ii*) the first known arithmetical calculation (1884), by Lord Rayleigh, of the distance over which speech might be transmitted in a cable.

These two disclosures deserve more than passing notice, one typifying the penetrating acuity of French thought in theoretical matters, the other illustrating a British capacity for supplementing theory with practical approximations to cut through to a needed conclusion.

Thévenin's theorem—which, like Ohm's law, can be generalized to cover the alternating-current case, i.e., to include reactances as well as resistance—as illustrated in Fig. 10-3, and states that the current I_T that will flow in an impedance Z_T terminating a network N can be

[2] "Sur un nouveau théorème d'électricité dynamique," *Comptes Rendus*, 1883.

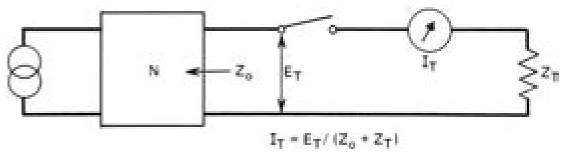


Fig. 10-3. Thévenin's theorem, central to circuit theory, postulated that the internal details of a network need not be known if certain external measurements can be made.

determined by (*i*) observing the voltage E_T at the terminals with Z_T open-circuited, and (*ii*) measuring the impedance Z_O looking back into the network with the original source (or sources) still connected but generating no voltage. The current I_T will then be $E_T/(Z_O + Z_T)$. Thus, if the required measurement of Z_O can be made, the complete performance of the network can be predicted with no knowledge of its internal makeup; or alternatively, if the details of the network are known, the performance can be predicted by calculating Z_O with the simplification that the original source is removed and replaced with its internal passive impedance.

Thévenin proved his theorem by a clever application of the Superposition Principle, already well established, whereby the currents in a linear passive network due to a plurality of voltages in different branches can be determined by simple addition of the currents that would be produced by the individual voltage sources acting separately. His theorem,[3] a brilliant conceptual achievement for his time and a perfect illustration of "research" as defined by Arnold, became of immense practical value in subsequent decades in the daily network problems handled by circuit engineers.

In the entirely different contribution by Rayleigh, which he presented at a Montreal meeting of the British Association, we find what is thought to be the first instance of an expression for *exponential* decay being applied to a sinusoidal voltage in the voice-frequency range impressed on a cable. Choosing "a note rather more than an octave above middle C"—specifically, with $2\pi f$ equal to 3,600 radians per second, that is, a frequency of roughly 600 hertz—and using resistance and capacity figures representative of the Atlantic telegraph cables of that time, Rayleigh deduced that a distance of 20 statute miles would reduce the "intensity" (i.e., ultimately the sound-power

[3] Both Thévenin's theorem and a companion or "dual" thereof stated by E. L. Norton (Bell Telephone Laboratories, 1926) had been anticipated by the German physicist Helmholtz in 1853. Norton's theorem states that the current I_T can be determined by noting the short-circuit *current* at the terminals and assuming that this same current will flow when Z_T is connected but will be partially diverted into an impedance Z_O shunting Z_T.

level) "to about a tenth" (more exactly, by a factor of 0.135). Rayleigh then demonstrated how clearly he appreciated the *iterative* nature of this decay by observing that such a reduction in intensity could not often be repeated without rendering speech inaudible; and that consequently "with such a cable the practical limit would not be likely to exceed fifty miles, more especially as the easy intelligibility of speech required the presence of tones still higher than is supposed in the above numerical example."

The intricacy of the transmission problem was pointed up still further the following year (1885) in a treatise on "Alternating Currents of Electricity," by Blakesley of the Royal Naval College, Greenwich, in which the application of hyperbolic functions to calculations of alternating-current transmission was discussed and numerical values were given for relative attenuation of different harmonics of the fundamental note. The treatise stated: "Thus, at the end of a cable of any considerable length and capacity the various tones of the voice would be received in a state of degradation depending upon their pitch . . . the ear has not the synthetic power of reconstructing a composite tone from the wreck of variously degraded components. In this consideration reside the limits of telephony . . ."

Such was the intellectual ferment stirred up by the invention of the telephone, highlighting as it did the ignorance prevailing in the world of engineering and physics concerning electric waves and oscillations. The excitement of this new area of investigation and analysis, for those equipped with the intellect and training to explore it, was reaching a high point in the mid-1880s, as demonstrated by the examples just given, and was sustained by the further revelations of Heaviside, beginning in 1886, concerning wave propagation on wires and the role of the magnetic field therein.

One can imagine the challenge felt by a keen and well-trained mind given the responsibility of making this new knowledge useful in an infant industry, an industry already outpacing, in its almost wild growth, its capacity to adapt; for hardly any but empirical and cut-and-try methods had characterized telephone development up to that time. Because of the host of practical problems facing Hammond Hayes as he took over his assignment in December of 1885, he could never spend more than a small fraction of his time on the strictly scientific problems that intrigued him. Like others of similar background who came in later decades, he made up for this with his keen scent for promising trails to be followed and his brilliant direction of the minds that would explore them. In his quest for such talent his first discovery was John Stone Stone, recruited from Johns Hopkins University in 1890 through the recommendations of the renowned physicist Rowland, then on the Hopkins faculty.

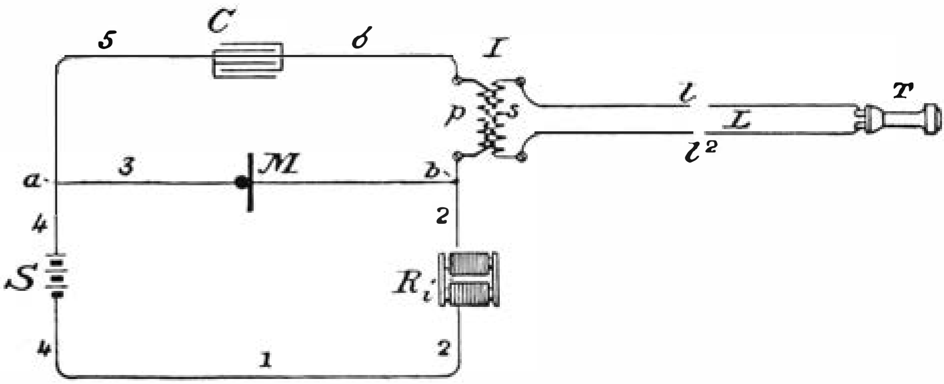


Fig. 10-4. An early concept of John Stone Stone: use of a choke coil R_i and a condenser C to separate the direct-current supply and the alternating-current product of modulation. An advantage cited was that a larger direct current could be fed through the microphone M than would otherwise be possible with the same size of primary wire on the line transformer I. (Redrawn from J. S. Stone, 1892, U. S. Patent No. 487,102, Fig. 2)

Well-equipped mathematically, already fascinated by the development of telephony, and impressed with the scientific beauty of electric-wave propagation as demonstrated in the experiments of Hertz, Stone embarked on his telephone career with enthusiasm. Hayes could scarcely have found a man better prepared, for that stage in the art, to make original contributions requiring both an understanding of existing theory and an imaginative approach to practical utilization, however far in the future. Stone could well, therefore, be considered the progenitor and exemplar of the communications research engineer of today, as Hayes was the pioneer administrator.

Some 20 patents granted in Stone's name during his ten years in Hayes' Boston laboratory—one, concerning cables, a joint patent with Hayes—attest to the breadth of Stone's perspective. In Stone's first patent application[4] we find him drawing a distinction between the direct-current supply of a microphone telephone transmitter and the alternating-current products of modulation. Figure 10-4 corresponds to the second figure of the patent. Lloyd Espenschied, engineer-historian who was to know Stone in a later decade of his career, sees in this disclosure early evidence of Stone's analytical ability and originality, his separation of the two components presaging the later separation of carrier and sidebands in high-frequency telephony.

[4] J. S. Stone; U.S. Patent No. 487,102; filed October 10, 1891; issued November 29, 1892.

In the latter part of 1891, after some brief and unsuccessful attempts at wireless communication but with Hertz's demonstrations of resonance still fresh in his mind, Stone began to entertain the notion of using resonant circuits in telephony—first as they might apply to selective signaling for an automatic telephone exchange, but soon afterward (May 1892) for the purpose of multiplex telephony. The following testimony, offered several years later in some patent suit proceedings, outlines Stone's ideas as he laid them before Hayes:

> I described to him how it could be possible to develop the high frequency currents by means of the Tesla oscillator which I had used in my experiments on signaling through space; how these currents could be varied in amplitude so as to correspond to the sound variations of the voice; and how, at the other end of the telephone circuit resonant electric circuits, one attuned to each of the frequencies in the high frequency currents impressed upon the line, could be employed to selectively receive these high frequency telephone currents. I pointed out to him that the frequencies of these high frequency currents could be made so great that except when their amplitude was varied they would produce inaudible effects in telephone receivers through which they pass, and that when their amplitudes were varied so as to correspond to the vibrations of the voice, the receivers through which they passed would reproduce the vibrations in the voice.

Encouraged by Hayes, Stone and an assistant worked for a year to try to implement these ideas; but, as with many concepts that are far ahead of their time, the notion of transporting speech waves by superposition or modulation upon a wave of higher frequency serving as a "carrier" had to await developments that were still in the future. One of these, it was fairly self-evident, was a generator of clean, continuous carrier waves. Interestingly, as we have already noted in Section 4.2.2 of Chapter 4, it was to be Stone himself who would introduce to Bell engineers two decades later the revolutionary device of de Forest that would supply this need. More obscure but even more fundamental, and likewise not to be understood for 20 years, was the process of modulation itself and the stringent requirements it would impose upon the selective circuits that Stone had tried to imagine. Nevertheless, this pioneering effort toward multiplex telephony played its part in sustaining Hammond Hayes' conviction that fundamental science, vigorously pursued, would hold the key to telephony's future.

Early in 1896, when the publicity on X rays and the discovery of radioactivity were reaching a peak, the alert mind of Stone speculated on the possibility that X-ray or other discharge tubes might be used for amplifying purposes, i.e., as "telephone relays." On the theory that the effect of radiations in such tubes would be "to temporarily convert dielectrics through which they pass into conductors of electricity, i.e.,

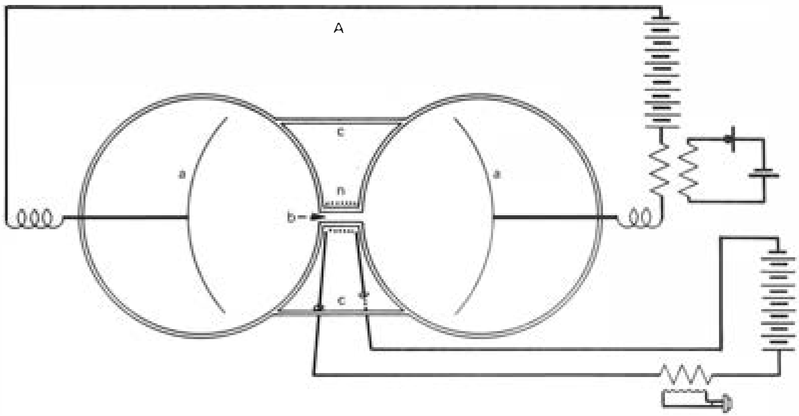


Fig. 10-5. In J. S. Stone's sketch of a possible speech amplifier or telephone relay, A is a "vacuum tube," a-a are electrodes of large surface, b is a very narrow passage with thin walls, c-c is a thick tube to strengthen an otherwise weak construction, and n is two parallel isulated wires coiled around the passage b. It appears that the potential developed across the two wires n by the telephone transmitter is intended to interact with the discharge through passage b and thus to appear in amplified form in the output circuit involving electrodes a-a. (Redrawn from J. S. Stone's sketch of March 6, 1896)

to enormously enhance the conductivity of the dielectric *pro tempore*," Stone suggested the arrangement shown in Fig. 10-5.

Indicative of Stone's clear appreciation of what was required for a successful speech amplifier was his specification that the conductivity be "approximately a direct linear function of the potential difference," and that there be "no appreciable time lag" in the changes in conductivity.

Other ideas[5] of a similar nature, involving applications of Roentgen rays or other radiations to telephone relaying, were proposed by Stone's colleague F. L. Rhodes. Yet this at a time when the electron itself had not yet been identified; and eventually it was not to be X rays, but rather the electrons that create them, that would become useful for amplification. What is remarkable in retrospect is the clarity with which these young engineers saw what was needed, so far in advance of its possible achievement. Even earlier (1890–92), when Stone's notes indicate that the subject of telephone "relays" was under discussion in the department, he had entertained the idea of using a dynamo as an amplifier of telephone sounds, feeding the microphone into the field winding and taking the amplified output from the armature.

[5] L. Espenschied, "Early Dreams of Electron Relays," *Bell Laboratories Record*, October 1947.

It was in 1897 that Hayes, in his continued objective of establishing a modest corps of highly trained scientists in the Headquarters technical staff of American Bell, brought Dr. George A. Campbell into what was still called the Mechanical Department. Hayes already had Stone working on implementing the mathematical discoveries of Heaviside, particularly relating to loading of cables, and in Section 4.1.3 of Chapter 4 we noted Stone's work on continuous loading and the use of inductance to match the impedance of "entrance" cables to the impedance of open wire lines. With the attenuation problem on long lines becoming so serious as to demand the highest level of attention, Hayes directed the efforts of Campbell thereto. Since the results of this concentrated effort by Campbell have been given in detail in the same section of Chapter 4, our present purpose will be served by citing the comments of Vannevar Bush,[6] many years later, in appraising Campbell's contribution:

> . . . In Heaviside's day the time was not ripe for great communication systems, nor were men of his time receptive to what appeared to be radical departures from convention. Unless there had been prophets to see the light and to carry it over into the new day, the great work of Heaviside would have been lost. These men who carried the torch can be numbered on the fingers of one hand, and of them Campbell stands out as the predominant figure, not alone because his grasp was sound and persistent, but also because he was associated with those who by transportation and communication developments were beginning the great work of joining America into a single nation.
>
> Into this early advance, to meet a commercial exigency, came the loading of telephone lines. The vision was Heaviside's. The design formulae were Campbell's. Out of a mathematical argument came coils of the right size and in the proper places to enable wires to go underground and to carry speech to long distance.

In Campbell's time the problem of propagation over wires was largely abstract, and it is accordingly difficult in retrospect to appreciate its intricacies. And though electrical configurations were more adaptable to mathematical treatment than were mechanical ones, it was a long road from the bare idea of continuous loading, or even the suggestion that the loads might be discrete, to an actual working cable or open wire line with the proper loads at the correct spacings.

> Empiricism here [continued Bush] would have been utterly hopeless and unjustifiably expensive. One cannot temporize with a long transmission line running across actual country by trying this and replacing that. Haphazard work of that nature could have but resulted in discouragement and failure. Procedure by thoroughly rigorous mathematical treatment to the condition where operation is understood, is the necessary step if the result

[6] Foreword to *The Collected Papers of George Ashley Campbell*, AT&TCo, 1937.

is to be practically accomplished. For this, Campbell, and Campbell almost alone, deserves the credit. He gave us the loaded line.

In telephony there is a subtlety in the end product, human communication. This subtlety, and the silent instantaneity of the process of telephoning, have supplied a glamor of an intellectual sort, which has always aided in the enlistment of superior minds. Doubtless this appeal was felt by Campbell when he signed the roster in 1897, and by others who joined the slowly growing force of talented young men built up by Hayes. John Stone Stone left in 1899 to work on his own, being especially intrigued by the possibilities of wireless, a medium he had tried while under Hayes but found not yet ready for exploitation in telephony. Into his place Hayes brought E. H. Colpitts from the Harvard faculty, who worked with Campbell to "prove in" and develop the early loading systems, later to become famous for his inventions in wire and radio telephony and as a research administrator. Also brought in by Hayes was G. W. Pickard, trained at Harvard and M.I.T., who carried Stone's early attempts at wireless to a somewhat more advanced state in 1902. Though still lacking the necessary source of continuous oscillations, Pickard succeeded in modulating a high-frequency spark discharge in a crude manner with the voice. Then in 1904, through the personal efforts of Campbell, Dr. Frank B. Jewett, a University of Chicago physicist, was recruited from the M.I.T. faculty. Jewett was destined two decades later to become the first president of Bell Telephone Laboratories, and was one of the world's most articulate spokesmen for the pursuit of scientific research in industrial laboratories.

With these and other capable recruits, some of them brought in by Jewett, the small coterie of men with advanced training grew[7] in size, and with its growth there developed the essential ingredient of large-scale success: the cross-pollination of ideas that comes from intimate association of highly trained minds exposed to a common broad problem. Hammond Hayes had established the nucleus and provided the early environment. Jewett in turn saw in the pooled contributions of Campbell, Colpitts, and their associates the supporting evidence for his convictions. Continued advancement in telephony, he could see, would require assimilation of exceptionally talented men into a cohesive force, capable of a penetrating attack on fundamental problems.

Heavily occupied though it was with the great number of engineering and operational problems posed by a complex young industry, the Boston laboratory under Hayes left a record of keen sensitivity to

[7] The reader is referred to Fig. 2-8 of Chapter 2, which indicates the place of the "Electrical Department" under Campbell in Hayes' overall organization as of January 1905 when Hayes became Chief Engineer. The following year, Jewett took over responsibility for this group from Campbell, who preferred to have fewer distractions from theoretical matters.

scientific advance. For example, in the Boston files we find records of additional attempts to carry on voice communication over light beams, as had been tried in the late 1880s in the Photophone experiments of Bell; in particular, an improvement[8] was worked out in 1897 using an arc as a light source and modulating it with the voice electrically. The files also contain records of diligent pursuit of methods for amplifying voice signals over wires, despite the spectacular advance made possible by Campbell's invention of loading, which for many years held center stage in the effort to extend the length of lines. Campbell himself considered negative resistances, such as might be afforded by arcs, for offsetting the effect of line resistance. Observing the fundamental weakness in the acoustomechanical approach that produced the earliest repeaters or "telephone relays," he concluded[9] that "the high speed of electrons in cathode tubes seemed . . . to be the solution for high fidelity amplification. I felt the need of experimental trial of such tubes. A laboratory room was assigned to me in the remodeled building at the back of 125 Milk Street, Boston. Before anything could be done there, we were transferred to New York, and I was among the dozen men who were crowded into Mr. Carty's old office on Dey Street. This delayed the work on the electronic repeater."

It was not to telephone transmission alone that theoreticians were being asked to make their contributions. Problems of administering the plant and the growing forces of operators began to engage the attention of the mathematically trained at an early date. Prominent among these were the relationships between average telephone traffic loads during the busy hour and those peak loads that could occur over short intervals during the hour. It was soon apparent, as the business grew, that empirically determined relationships were not good enough to form a basis for assigning operators, or for design of the "trunk" plant; deeper understanding of the nature of call fluctuations would be necessary. As early as 1898, G. T. Blood, whose name appears in connection with "Toll Traffic Studies" on the 1895 organization chart of American Bell (see Fig. 2-7 of Chapter 2), found a close agreement between the terms of a binomial expansion and the observations made on distribution of busy calls. The following year, M. C. Rorty, a young Cornell engineering graduate who had been brought in from the New York Telephone Company through the personal intervention of Hammond Hayes, became interested in the possibility of describing call fluctuations by means of probability theory. In 1903, Rorty composed a paper entitled "Application of the Theory of Probability to Traffic

[8] H. V. Hayes and E. R. Cram; U.S. Patent No. 654,630; filed June 7, 1897; issued July 31, 1900.

[9] From attachment to a letter written by Campbell, after his retirement, to L. Espenschied.

Problems"—a paper that he later presented before the AIEE and that generated wide interest throughout the industry.

There had been skepticism, understandable for that time, that people's social habits could be predicted by any kind of theory. "The belief is common," wrote Rorty, "that, at the point where the human element enters a proposition, all recognized rules cease to hold." It was much to the credit of Rorty, and of the stimulating environment provided by the Hayes organization, that where humanity in the aggregate is involved, the applicability of a mathematical approach could indeed be demonstrated. But it was a self-educated young circuit engineer and mathematics student in the Campbell group, E. C. Molina, who became most interested in the telephone traffic problem upon learning of Rorty's work. Molina's contributions to the switching and trunking art over many years have been detailed in Chapter 6. It was the stimulating milieu of the Boston laboratory that kindled the interest of this major contributor in fields in which he possessed extraordinary talent.

Several years later, in 1907 (the year that Hayes retired), general financial panic descended, in which the Bell System inevitably shared in general business retrenchment. Since such crises are generally characterized by cutbacks in forward-looking industrial developments, the atmosphere was a discouraging one for scientists and engineers; and for the Boston staff of AT&TCo it was especially so, for Theodore Vail, returning to the presidency of AT&TCo, decreed as a part of his drastic reorganization plans that the Engineering Department should move to New York City. Even such a stalwart as Jewett, with his bright prospects for the future, briefly felt the temptation to resume the academic career broken three years before. Nevertheless, the outstanding men with fundamental research capabilities, with the exception of Hayes,[10] made the move from the Boston establishment and were ready to resume work as business confidence improved and prospects brightened for fruitful application of their ideas.

Such applications were not long in coming. The new Chief Engineer of AT&TCo appointed by Vail was J. J. Carty, a man whose vision and appreciation for scientific research in industry we noted earlier. Carty and Jewett, with the encouragement of Vail, saw the long-distance transmission problem as Hayes had seen it, namely, the system's most pressing technical challenge. It was clear, looking ahead to the westward extension of the New York–Chicago circuit as already contemplated, that Denver would be the end of the line, for at any greater distance the inexorable toll taken by ohmic attenuation would be far too great for

[10] The reader is referred to *Proc. IRE*, April 1948, where Edward Lindley Bowles, in a tribute to Hayes, discusses the conflict of views between Hayes and Vail that led to the termination of the former's service.

any practical size of conductors to overcome, even with loading pushed to the limit. Thus, beyond the necessity for continual engineering improvement, there was no question about the need for a scientific breakthrough, and the need to marshal the necessary talent for its accomplishment.

We have already seen (in Section 4.2.2 of Chapter 4 and Section VII of Chapter 8) how the enlistment of physicist H. D. Arnold—with the timely demonstration of de Forest's three-element tube the following year to Arnold and his superiors—unleashed a concerted drive in the Bell quarters to capitalize on the extraordinary versatility of this new component. Early in 1911, the formal establishment of a Research Branch, headed by Colpitts, in the Engineering Department of Western Electric had ensured that work of a fundamental character would receive the management's full support. Thus, 50 years later[11] it could be said that "The preeminent discovery of the twentieth century is the power of organized scientific research."

The invention and perfection of the vacuum tube must be granted undisputed claim as the key to unlimited long-distance telephony. Yet major problems of the Bell System would have been left unsolved but for new fundamental understandings in communication theory reached in those same early years of the century. We will discuss those that were most important.

Of great conceptual and practical significance was the understanding of the "band" nature of signals, whereby any translation or displacement of telephone or other signals in *frequency*, for multiplexing, radio transmission, or other purposes, must necessarily provide a certain width of band for their accommodation. People had not been thinking explicitly in such terms. Indeed, if one were to define the "frequency" of an oscillation, e.g., a carrier wave, as the number of regularly-spaced axis crossings per second, it would not be expected (nor was it supposed by many) that modulation of the amplitude alone would be reflected in any way in the frequency spectrum. Even John Stone Stone (then no longer with the Bell System) held this view as late as 1912.[12]

In the minds of Campbell and Colpitts, however, there was no such naive assumption. An old phenomenon in acoustical physics was the combining of two waves in a non-linear element to produce sum and difference waves, and there appears to have been a certain carry-over of

[11] "The World's Greatest Industrial Laboratory," *Fortune*, November–December 1958.

[12] Stone's Franklin Institute paper in October of that year, entitled "The Practical Aspects of the Propagation of High-Frequency Electric Waves Along Wires," advocating carrier multiplex, made the claim that while attenuation, because of the higher frequencies required, would necessarily be greater than with straight voice transmission, the usual distortion would be eliminated. "There is, in fact," the paper reads, "in the transmission of a given message, but a single frequency of current involved, and therefore no unequal attenuation of components of different frequencies and no distortion."

that knowledge as they contemplated electric wave modulation, tacitly assuming the existence of a band of frequencies even though the spectral structure was not completely visualized.

Further evidence of Campbell's awareness of the "band" nature of signals is seen in the wave filter, invented by him in 1909, stemming from his studies on loaded lines and constituting an achievement of greatest fundamental importance to the art. Campbell's report[13] to Carty, dated March 7, 1910, gave drawings and characteristic curves not only for low-pass and high-pass filters, but for band-pass filters as well. Thus, from the earliest work in the Western Electric Engineering Department and AT&TCo on the twin developments of radiotelephony and high-frequency wire multiplex, the maturing theory of the transmission band played a dominant role. By the summer of 1915, both H. D. Arnold and a young mathematician from Princeton, J. R. Carson, recruited in 1914, had come to recognize that one sideband alone sufficed to convey the information in the signal and had thought, independently, of eliminating the other sideband. Carson proposed suppressing the carrier as well. Thus was born single sideband, the priority for which went to Carson.

The penetrating power of Carson's and Campbell's mathematical approach illuminated many investigations[14] during the early decades of this century. It was a time when the intuition of the electrical engineer needed support from the rigorous analysis of the professional mathematician; and the collaborative setup between Colpitts' Research Branch and the analytical group under Campbell at AT&TCo provided this support. In retrospect, one can see, as a product of that period, the development of basic principles and a method of analysis that would be universally adopted for communications research.

Heaviside's studies in the 1880s had been carried out against a background of 30 years of telegraphy, during which a fundamental objective was the preservation of the identity of telegraphic impulses as they traversed the medium. It was natural, therefore, that his approach should include a special concern for the transient, or starting-and-stopping, aspect of the received impulse or signal. Moreover, experimental work of that time, such as the investigations by Professor David Hughes in England on "skin effect" in ferrous and non-ferrous conductors, made use of an interrupted or "square wave" current, and the Heaviside

[13] Cited in a communication from L. Espenschied to *IEEE SPECTRUM* on the electric wave filter: August 1966, p. 162.

[14] An example we shall not be pursuing further, but outstandingly clear-cut, was Carson's analysis of modulation theory, published in 1922, which exposed the fundamental fallacy in certain proposals, ingenious and superficially plausible, for further economizing in band requirements by modulating the *frequency* (instead of the amplitude) by small amounts.

analysis had to seek reconciliation with the hard-to-interpret findings. Thus was Heaviside led to the concept of "indicial admittance," i.e., the response (as a function of time) at any point in a circuit or transmission line, previously in equilibrium, when a sudden unit increment in potential is applied at the sending point. Determining the indicial admittance from the differential equations was the essence of the problem; for, with this factor known through integration, the response to any arbitrary signal could be determined by considering the signal as a series of such increments, of appropriate amplitudes, closely spaced in time. In the limit, then, the response would be found from the solution to an integral equation.

Heaviside, however, had broken with classical methods and chosen a shorthand or symbolic procedure in which the differential operator d/dt was replaced by the symbol p and the operation $\int dt$ by $1/p$. This reduced the differential equations to a form that was algebraic and therefore superficially simple, known as the Heaviside *operational equations*. Heaviside's methods, though potentially of wide applicability in electric circuit theory, were used by only a few specialists; the neglect, according to Carson when he analyzed[15] the Operational Calculus many years later, being due "less to the intrinsic difficulties of the subject than to unfortunate obscurities in Heaviside's own presentation."

Even the classical approach, however, involving the solution of integral equations as indicated above, could have presented formidable difficulties in the solution of communication engineering problems for those who were not professional mathematicians. But fortunately, beginning in the 1890s, an alternative to this "transient" method of attack, known as the steady-state approach, was promulgated in industrial and academic circles and became part of the training of all electrical engineers. The field of alternating-current power, which was advancing rapidly at that time, was inherently characterized by a constant-amplitude, constant-frequency view of electrical oscillations. Yet it could be seen that even the complex and transitory sounds of speech, in the light of Fourier analysis, were representable to a high degree of accuracy by including a sufficiently large number of sinusoidal components. Thus the mathematical treatment of telephone transmission networks lent itself to the same methods that were being popularized at that time, primarily for use in the power field, by C. P. Steinmetz of the General Electric Company and by A. E. Kennelly of Harvard University. Kennelly in particular, many of whose papers appeared in the *Transactions of the AIEE* and in the *Harvard Engineering Journal* in the mid-1890s, focused much light on the evolving concepts of reactance and im-

[15] J. R. Carson, *Electric Circuit Theory and the Operational Calculus,* McGraw-Hill, 1926.

pedance; on the use of complex algebra to represent orthogonal quantities in a single two-dimensional equation; on the use of the exponential function e^{jx} as the equivalent of the complex quantity ($\cos x + j \sin x$); and on the applicability of hyperbolic functions to network and transmission line problems, especially where waves were reflected at a terminal.[16]

To mathematicians endowed with the insights of Campbell and Carson, the transient and steady-state approaches to circuit theory were merely different aspects of the same thing; and to Bell System engineers concerned with telegraphy and later telephotography, where preservation of wave *shape* was important, an understanding of solutions based on the Fourier Integral, as well as the Fourier Series, was necessary. Still more relevant and useful was the transient approach to come with the development of television and digital or pulse transmission. But in the telephone transmission research of the 1910s and 1920s, the steady-state method held dominance and was the foundation for most of the development of network theory in that period.

Especially noteworthy, and introducing a powerful approach to circuits generally, was the Campbell AIEE paper of 1911, entitled "Cisoidal Oscillations," in which the author employed the expression cis pt to represent ($\cos pt + i \sin pt$), or alternately e^{ipt}, and argued for the representation of signals by *pairs* of cisoidal oscillations having positive and negative frequencies, i.e., having the time factors $+pt$ and $-pt$. Characterizing all of Campbell's work was an adherence to mathematical rigor and a disparagement of what he considered loose practices, such as the use by engineers of j rather than i as the imaginary coefficient, and the use of the term "vector" for quantities inherently scalar but represented on a two-dimensional graph. If Campbell's uncompromising rigor gave his papers an *hauteur* or loftiness possibly discouraging to those brought up on more rudimentary concepts, his clear mastery of the subject inspired many to work more diligently on deepening their own understanding. It is the observation of the present authors, historically, that those with extraordinary genius and insight must operate at their own levels, while colleagues assume responsibility for interpretation and dissemination when necessary. It therefore becomes doubly laudable when outstanding theoreticians are led by practical urgencies to engage, for a time, in the less esoteric phases of a project, as did Campbell in the early proving-in period of inductive

[16] Hyperbolic functions are characterized by the combination of a term of the form e^{-ax} (where a is a complex attenuation factor and x is a distance), plus a term of the form e^{ax}, the latter representing a backward-moving wave, similarly attenuated, from a reflecting terminal. Similar expressions had appeared in the equations of Fourier (1822), describing the flow of heat along a bar of finite length, and in the equations of Lagrange (1759), describing motion of elastic waves along a loaded string.

loading. Yet this is a process that has repeated itself continually in science and technology, and it is a part of the genius of enlightened research management that it can bring this about.

Before the advent of telephone amplifiers, there was little interest in devising electrical networks for reducing distortion in the transmission of speech. The major problem in transmission was to conserve speech power. The mere correction of defects by circuits that could only *absorb* power was not an objective of high priority, much as these defects were deplored.[17]

Once the telephone circuit was freed from energy limitations by the availability of the vacuum tube repeater, it was immediately possible to invoke techniques for equalization, whereby versatile combinations of resistances and reactances could be devised for making the transmission uniform over the frequency range. An even more urgent need for frequency-dependent networks, however, arose in connection with the repeater itself, particularly in the case of two-way repeaters; for, as has been discussed in Section 4.2.4 of Chapter 4, it was essential to balance impedances quite accurately to have stability and adequate repeater gain. To simulate or balance the impedance of a loaded line over the voice-frequency range was a challenging problem, especially since the line could commence at an arbitrary point with respect to the loading coil locations. This problem was solved by R. S. Hoyt, who had entered the Bell System in 1912 from Princeton, as did J. R. Carson two years later. Hoyt's important patent,[18] applied for in 1913, demonstrates the proficiency in mathematical treatment of electrical circuits which would soon have to become a part of the analytical equipment of all communication engineers.

Starting with a formula Campbell had given[19] ten years earlier for the "mid-load" impedance of a long periodically loaded line, Hoyt derived in his patent, using complex quantities, an expression for the resistance and reactance of such a line, as functions of frequency, when entered at any point in a section. Figure 10-6 is a graph from his patent, showing how the *resistive* component of this impedance would vary with frequency for various points of connection. Connection in the vicinity of the point 0.2 (one-fifth of a loading section) results in a nearly constant resistance up to 85 percent of the critical frequency. Hoyt was then able to show that the complex impedance of the line at this point could be

[17] The Campbell paper, "Telephonic Intelligibility" (*Philosophical Magazine*, January 1910), describes elaborate statistical tests made to determine the degree to which familiarity with the speaker's voice, and with the context, aided in distinguishing sounds that would otherwise be confused. More will be said on this subject in Section III of this chapter.

[18] R. S. Hoyt; U.S. Patent No. 1,124,904; filed April 30, 1913; issued January 12, 1915.

[19] "Loaded Lines in Telephone Transmission," *Philosophical Magazine*, March 1903, Equation 19.

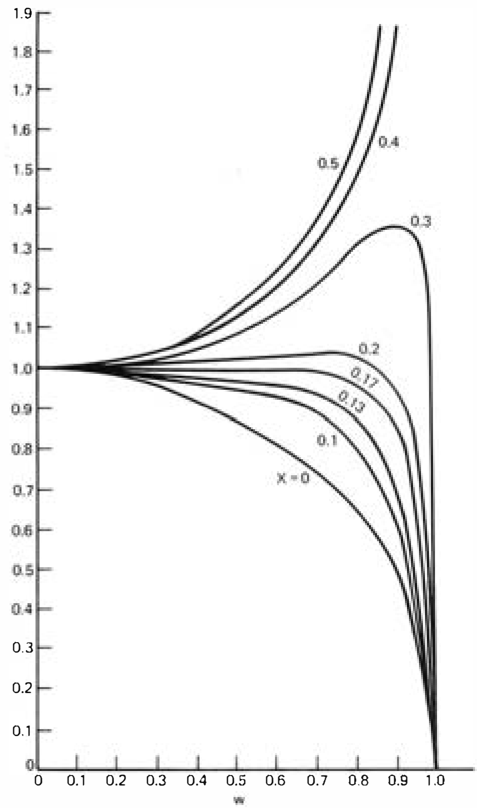


Fig. 10-6. A graph from R. S. Hoyt's 1915 Artificial Line patent, showing the variation in the resistive component of the impedance of a loaded transmission line (w being the ratio of frequency to the critical frequency of the loaded line) for various points of entrance, i.e., fractions of a loading section. (Redrawn from R. S. Hoyt, 1915, U. S. Patent No. 1,124,904)

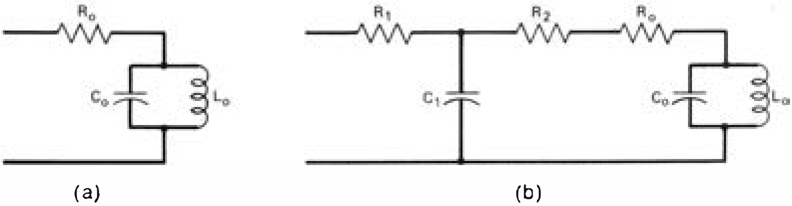


Fig. 10-7. Simulation of the loaded line, (a) for connection at 0.2-section, and (b) for connection at any other point. (Redrawn from R. S. Hoyt, 1915, U. S. Patent No. 1,124,904)

simulated (i.e., balanced) with fair precision by means of the simple circuit shown in Fig. 10-7a, in which an anti-resonant circuit is added in series with the resistance to give the required reactance; and that by the addition of a simple three-element artificial line section (as shown in Fig. 10-7b) to "build out" the network, the impedance for fractional lengths other than one-fifth of a loading section could also be simulated.

Hoyt's balancing networks, together with simple types of Campbell wave filters for excluding frequencies outside the range of balance, were essential to the successful operation of the first transcontinental line in 1915, as related in Section 4.2.3 of Chapter 4. Three years later, when the first carrier system was installed, band-pass filters were employed to separate the four channels; and at about the same time, the first attenuation equalizer was used, consisting of a resistance-reactance network across the line. Thus the frequency-dependent network in various forms was on its way to very widespread use, and the time had come for a thorough theoretical study of its properties and applications. To this task O. J. Zobel was assigned when he entered the Engineering Department of AT&TCo in 1916 with a doctorate from the University of Wisconsin.

Zobel's achievements in establishing an art of filter design demonstrated an apt combination of mathematical skill and engineering insight. The Campbell approach had treated filter sections as elements in a series of the same type, terminated in what Campbell called their "iterative"[20] impedance, the attenuation of each section being readily computed as a function of the ratio of its series to shunt impedances. But since the iterative impedance varied with frequency, typically (in the absence of dissipation) going to zero or becoming infinite at the cutoff frequency, the actual termination was ordinarily much different, and a "reflection factor" and "interaction factor" had to be introduced to permit accurate computation of actual loss and phase shift through the filter. Another innovation, with the introduction by Zobel of "composite" filters containing unsymmetrical sections or half-sections of various types, was the term "image impedance" where "iterative impedance" would not be properly descriptive because of dissymmetry between the two ends. Accompanying this new term was the term "image transfer constant" defining (by its real and imaginary parts) the loss and phase shift of the network when connected between its image impedances.

One of Zobel's most notable achievements was his invention in 1920 of the "*m*-derived" type of filter. Most of Campbell's filters had been of the type known as "constant *k*," i.e., the product of the impedances

[20] "Iterative impedance," "characteristic impedance," and "surge impedance" were interchangeable terms applied by early investigators to the input impedance of an infinitely long and therefore reflectionless transmission line or series of artificial line sections.

of the series and shunt elements of a section (e.g., the series inductances and the shunt condenser of a low-pass T section) was a constant, k^2, independent of frequency. Similarly, a constant k band-pass filter might include resonant circuits in the series arms and anti-resonant circuits in the shunt arms. Such filters, considered as "prototypes" from which more complex structures were to be "derived," left much to be desired in steepness of cutoff, especially looking to multichannel systems where conservation of frequency space was important. Moreover, the image or iterative impedance of a constant k section was far from constant over the pass band, leading to difficulty in properly terminating the section. Zobel, striving to extend network theory in directions that would aid filter designers, made the remarkable discovery that certain modifications of the prototype, including a resonant or anti-resonant circuit, introduced to give a peak of attenuation at a frequency outside the pass band and thus substantially increasing the steepness of cutoff, could be so chosen with respect to a parameter m —less than unity but otherwise arbitrary—as to leave the image impedance of a section the same as that of the prototype at all frequencies. Figure 10-8 shows how the parameter m is used in deriving such a section from its prototype in the case of a T-section low-pass filter. The relationship between the parameter m and the ratio of the frequency of the attenuation peak to the cutoff frequency is given by:

$$f_\infty/f_c = 1/\sqrt{1 - m^2}.$$

The identity of image impedances just mentioned permits use of a number of sections or half-sections in cascade, having different values of m, and including sections or half-sections of the prototype. These constitute a "composite" filter with matched (reflectionless) junctions throughout, so that the image transfer constants are directly additive to give the required total attenuation over the band to be rejected.

Zobel made the further highly valuable discovery that by terminating such a filter in an appropriate m-derived *half*-section, using a value for m of 0.6 (giving, in the low-pass case, an f_∞/f_c of 1.25), its image impedance at either or both ends could be made remarkably constant over

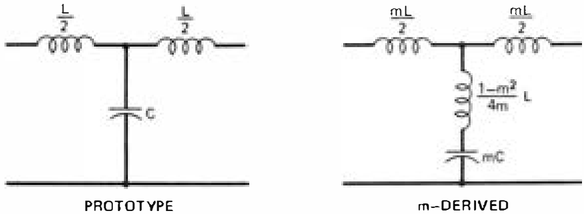


Fig. 10-8. O. J. Zobel's invention (1920): the m-derived wave filter.

most of the pass band. This minimized reflection losses when introduced between constant resistance terminations and greatly increased the confidence with which the filter's performance could be predicted when inserted in a system.

It was characteristic of Zobel, in these pioneering developments in transmission network theory, that he paid meticulous attention to the preservation of matched image impedances so that reflection and interaction effects could be minimized, and he extended these principles to equalizer theory as the art advanced and transmission requirements became more severe.[21] The early development of equalizer theory was largely the work of Zobel and R. S. Hoyt, whose work on balancing networks we have already discussed. Hoyt's equalizers were in the form of two-terminal networks connected either in series or, more often, in shunt with the line, and because of their simplicity were widely used. Hoyt's 1918 patent application[22] included the mathematical theory of both series and shunt two-terminal networks as loss equalizers.

The Zobel constant-resistance network approach, on the other hand, is exemplified in Fig. 10-9 where, if Z_1 and Z_2 are two "inverse" impedances, i.e., impedances having a constant product equal to R^2, and the network is terminated in a load resistance equal to R, the input impedance will likewise be R, independent of frequency. The desired variation of loss with frequency is obtained simply by relating Z_1 (and thus necessarily Z_2) appropriately to R, the loss curves appearing as circles with these coordinates. Thus, as would not be the case with one of Hoyt's two-terminal equalizers, large variations in loss could be introduced without reflections, and any desired number of sections could be cascaded with results that were simply additive.

Proceeding in the late 1920s with further refinements in filter and equalizer theory, Zobel (still at AT&TCo after the formation of Bell Telephone Laboratories) applied himself likewise to phase equalizers. Interest in the phase characteristic of filters and networks came from

[21] In the early 1920s, with the advent of radio broadcasting, the demand for maximum flatness of frequency response in cable circuits was an added incentive.
[22] R. S. Hoyt; U.S. Patent No. 1,453,980; filed June 29, 1918; issued May 1, 1923.

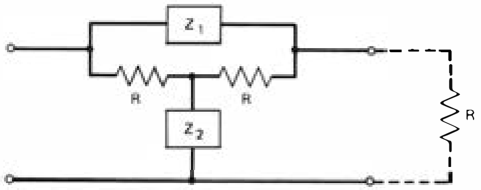
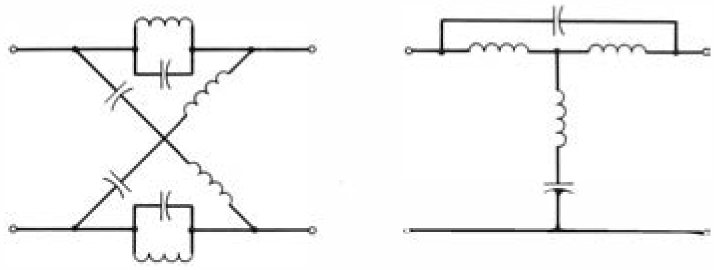


Fig. 10-9. Example of constant-resistance type of loss equalizer, where $Z_1Z_2 = R^2$. The attenuation is expressible very simply in terms of Z_1/R or R/Z_2.



(a) (b)

Fig. 10-10. Example of (a) lattice-type phase equalizer, and (b) an equivalent network of the unbalanced or bridged-T type. The latter is sometimes not physically realizable (because of negative values) without using mutual inductance.

increasing appreciation of its effect on the transient response, as in telegraph transmission over cables and the transmission of visual-type signals. A phase curve increasing as a linear function of frequency had no effect on the transient except to delay it by a time equal to $dB/d\omega$, where B was the phase in radians and ω was 2π times the frequency characterizing the "bundle" or spurt of waves composing the transient. Deviations of $dB/d\omega$ from a reference delay were referred to as delay distortion.

It is intuitively apparent that an all-pass lattice-type[23] network, exemplified by Fig. 10-10, would be a basic network for introducing delay. With two of the arms arranged to be the inverse of the other two, and with the product a constant equal to the square of the terminating resistance R, the lattice exhibits a constant image impedance R and a phase shift which varies from 0 to 360 degrees, i.e., the network changes gradually with frequency from a direct connection to (in effect) a reversing switch, then continues to shift phase and ultimately becomes a direct connection again.

Many design charts were plotted for the use of circuit engineers, showing delay or phase shift in terms of other parameters, both for lattice networks and for other types. Preparation of such charts and methods of computation for filters and other networks by Zobel and his colleagues had been a contribution of the utmost importance. It illustrates one of the responsibilities of fundamental research beyond the discovery and exploration of new ground, i.e., the systematization of knowledge in a form that is quickly usable by those responsible for

[23] The lattice is the most *general* of the network types because of the additional "handle" on the signal afforded by the diagonal branches; but, as a filter, the lattice was not widely used (until the advent of the balanced piezoelectric crystal filter in the 1930s) because a higher degree of precision was required in the element values.

its repeated application in design. More than this, the leaders in these advances not only published their findings in the technical and scientific journals but also lectured at the universities[24] to stimulate interest in the field; for it was a time when, as expressed by T. E. Shea, "the various dynamic sciences, which had rapidly grown in scope under the influence of Maxwell, Helmholtz and others, having been for a time pent up, were fairly bursting in their eagerness to spill over into the communication field innumerable facts capable of being put to great practical use."

Among the contributors to this systematized knowledge was E. L. Norton, whose network theorem has been mentioned earlier. Norton entered Western's Engineering Department from M.I.T. in 1922. A major specific contribution of his was the invention of impedance transformation in band-pass filters (patent applications dated 1924 and 1925), used in some cases to eliminate the need for an actual transformer but more often as an *internal* impedance change within a network to give more practicable values for some of the elements or even to eliminate them entirely.

The few examples we have given of developments in network theory illustrate how the urgent needs of a technologically based industry can spur rapid scientific advance when trained minds are enlisted and provided with appropriate inspiration. And any who in the mid-1920s may have felt that all of the facts concerning networks were known were to find new challenges on the immediate horizon; for the invention of negative feedback by H. S. Black in 1927 was to impose drastically more severe requirements on the phase and loss characteristics of networks, and would require new understandings of the stability of dynamic systems. To these understandings, H. Nyquist and H. W. Bode, stimulated by the atmosphere we have described, would be the outstanding contributors. Of Nyquist, a native of Sweden, whose career at AT&TCo had begun in 1917 when he received his doctorate from Yale, we shall say more in the next few pages.

Also participating in the lively extension of knowledge in signal and circuit theory was Pierre Mertz, a Cornell graduate and an associate of Nyquist in the theory of telegraphic types of signals. It was Mertz whose quantitative analysis of visual signal transmission in 1921 gave

[24] The classic texts, *Transmission Circuits for Telephonic Communication* (K. S. Johnson, 1924) and *Transmission Networks and Wave Filters* (T. E. Shea, 1929), were prepared from notes of lectures given not only within the Bell System, but at Harvard and M.I.T., of which the respective authors were graduates. These texts have been used worldwide. Similarly, J. R. Carson's book (1926) *Electric Circuit Theory and the Operational Calculus* embodies lecture material delivered at the Moore School of Electrical Engineering, University of Pennsylvania. These texts include references to other papers on circuit theory by authors in the Bell System and elsewhere, including important contributions by the German electrical engineer K. W. Wagner.

the first indication of the very wide frequency bands that would be required to accommodate motion. In a less inspiring environment, the disclosure could have discouraged Bell System entry into television transmission; yet by 1925, the first year of operation of the new Bell Telephone Laboratories, research on television was in full swing, leading to successful transmissions by wire and radio soon afterward, as described in Section 9.8 of Chapter 7. Toward the end of the present section, we shall touch again upon the research aspects of television.

The problem of minimizing bandwidth requirements is one we have been touching upon repeatedly, since frequency is one of the most fundamental parameters in communication; and we have cited particularly the contributions of Campbell and Carson to frequency conservation. In the mid-1920s, R. V. L. Hartley, a University of Utah graduate and a Rhodes scholar, already well known for his vacuum tube circuit inventions in the research group of Western Electric's Engineering Department (particularly for the oscillator bearing his name), studied the basic relationship between width of frequency band and the capacity of a system to transmit "information." His work followed mathematical studies by Nyquist at AT&TCo concerning the maximum number of symbols that could be transmitted in a given frequency band. Hartley's results, published[25] before the end of this fast-moving decade, were to be guiding rules for transmission engineers for nearly 20 years, giving way to a more comprehensive theory only when C. E. Shannon (Bell Telephone Laboratories, 1948) would carefully examine the effect of *noise* on the ability of a system to preserve faithful indications of amplitude.

Thus, Nyquist's and Hartley's studies related to a hypothetical noise-free medium, though electrical noise and interference from diverse causes had plagued telephone development from the beginning. In Chapter 4 we recorded the long and difficult program of interference prevention in telephone transmission, from the earliest steps of employing balanced pairs of wires and transpositions to the systemwide collaboration with other electrical industries. The vacuum tube repeater, boon though it was in periodically boosting the signal level along the line to preserve necessary margins against interference, generated noises of its own with which the early vacuum tube workers at Western Electric had to struggle.

Similar problems, including cathode instability, inadequate pumping, mechanical resonances, and poor welds, must have beset the tube developers in other countries in the years of World War I; but in a few

[25] R. V. L. Hartley, "Transmission of Information," *Bell System Technical J.*, July 1928. Note also Nyquist's papers, "Certain Factors Affecting Telegraph Speed," *Bell System Technical J.*, April 1924, and "Certain Topics in Telegraph Transmission Theory," *Trans. AIEE*, 1928.

laboratories, these difficulties did not deter scientists who wanted to know what the ultimate electrical noise level in an amplifier might be when the grosser current fluctuations from faulty tube construction could be eliminated. It was J. B. Johnson, a native of Sweden as was Nyquist, and with a doctorate in physics from Yale, who sought these answers in the Engineering Department of Western Electric.

The classical paper on noise in amplifiers had been written in 1918 by W. Schottky of the Siemens and Halske firm in Germany. More than 50 years later[26] Johnson was to observe, in retrospect, how remarkable it was that out of the Germany of those years, faced with military defeat and economic collapse, there could come a scientific paper of the quality and technical importance of that paper of Schottky's.

Schottky concluded that an amplifier would have two sources of noise of a fundamental nature, perhaps irreducible. One, which he called the *Wärmeeffekt*, or thermal noise, would arise in the input circuit and would be produced by the random flow of charges in the conductors in response to the heat motion of their molecules. It would involve the Boltzmann constant k, a universal constant of physics, times the absolute temperature of the system; and the fluctuating voltage corresponding to this noise, being applied to the input electrode of the tube, would cause the noise to appear at the output terminal in amplified form. The other fundamental source of noise, called *Schroteffekt* or shot effect, would be internal to the tube and would result from the randomness of electron emission, and emission velocity, from the cathode. This effect was calculable by statistical theory, and it was first experimentally identified and measured in Schottky's laboratory.

It was Schottky's view that shot noise would be much stronger than thermal noise, and the efforts of Johnson and colleagues at Western Electric, as well as investigations at the General Electric Company and elsewhere, were concentrated for some time on the former effect. An important 1925 paper[27] by Johnson, however, included a set of observations that we here reproduce as Fig. 10-11. As Johnson studied these data afterward, his special interest was aroused by the fact that the quietest tubes (those represented along the sloping line at the left) showed a noise level proportional to their amplification, suggesting that the effect of shot noise in those samples was being masked by the thermal noise in the input circuit, the latter being then amplified by the tube. Thus, the *Wärmeeffekt* had been tentatively identified, and its identity was verified by subsequent tests in which the temperature, value, composition, etc., of the input resistor were varied. Johnson's

[26] "Electronic Noise: The First Two Decades," *IEEE Spectrum*, February 1971.
[27] *Phys. Rev.*, 1925.

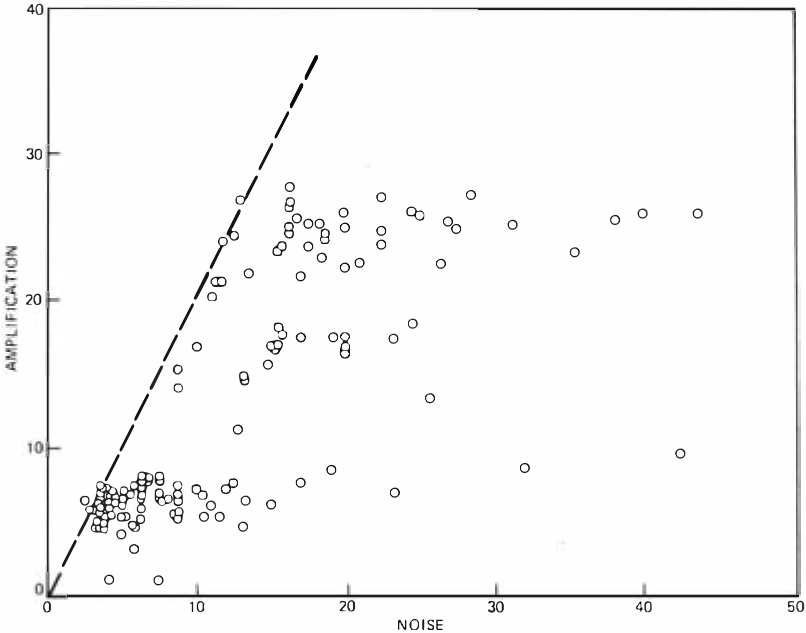


Fig. 10-11. The first experimental evidence of thermal or "Johnson" noise. Each point represents a vacuum tube under test. The sloping line at left shows that the "quietest" tubes had a noise output proportional to amplification and therefore presumably arising in the input circuit. (Redrawn from J. B. Johnson, 1925)

associate, Nyquist, then proceeded to analyze the effect mathematically from thermodynamic considerations, coming up with a formulation for thermal noise *per unit of bandwidth* of 4 kT watts, where k is Boltzmann's constant and T the absolute temperature.

Johnson's discovery and Nyquist's formulation, published in companion papers,[28] are an outstanding example of the scientific method: painstaking, orderly experimentation joined with insight and careful theoretical analysis. Coming during the formative years of Bell Laboratories, the achievement reinforced industry's growing appreciation of the role of research—for here was being uncovered a fundamental physical phenomenon whose application would prove to be an engineering milestone as well.

The study of electronic noise was to continue throughout subsequent decades. At Bell Laboratories, F. B. Llewellyn and others were

[28] *Phys. Rev.*, 1928, pp. 97–109 (Johnson) and 110–113 (Nyquist). Johnson had also published (in 1927) short notes in *Nature* (January) and in *Phys. Rev.* (February).

to be major contributors in this field, both in experiment and in mathematical theory. Johnson had, himself, uncovered several sources of noise within the tube. Shot noise, in particular, was greatly dependent on such factors as cathode design and material and the nature of the electron flow, whether emission-limited or space-charge-limited; and various strategems were to be devised, beyond our present scope, to bring circuit noise as close as possible to thermal or "Johnson" noise—even, 30 years later, to circumvent this seemingly ultimate limitation. The latter accomplishment was to depend on new concepts of the 1930s and 1940s whereby economy in width of frequency band could be sacrificed for noise immunity—an exchange whose ultimate profitability could scarcely be envisioned in the frequency-starved years that had preceded.

Accordingly, before concluding our present section on research related to transmission, we shall devote a few pages to the new vistas in transmission that unfolded as the frequency spectrum widened. In particular, we emphasize radio; for, as the growing research staff that included Colpitts, Hartley, R. A. Heising (an electrical engineer coming in 1914 from the Universities of North Dakota and Wisconsin), and H. J. van der Bijl (coming in 1913 from the University of Chicago, a native of South Africa but with a doctorate from Leipzig) created the first oscillators and modulators for wire carrier systems, they were envisioning transmission in free space as well. It was apparent to them that any extensions they could make in the frequency range of their circuits would give not only more radiotelephone channel capacity but also more efficient "coupling" (by way of the antennas) to the new medium. It was even possible, looking ahead, to envision wavelengths sufficiently short so that the dimensions of an antenna, or an array[29] of antennas, might encompass a number of wavelengths and, by the laws of geometrical optics, focus the radiation in chosen directions (as had already been demonstrated by Hertz) or might similarly receive radiation from selected directions. It was, therefore, not beyond the imagination of these pioneers that the human voice, even though unguided by wires or other physical boundaries, might still be projected and received on radio waves in a directional manner, thereby conserving both power and frequency space and suppressing interference.

Thus, research in radiotelephony was particularly rich in the leaven of ideas which engendered it, and likewise was to prove rich in the analyses and reductions to measurement that would open the way to

[29] G. A. Campbell's 1919 patent application (U.S. Patent No. 1,738,522 issued in 1929) contains a theoretical discussion of antenna arrays, illustrating how far advanced was the mathematical thinking on this possibility. Later, Campbell's associate, R. M. Foster, collaborated in the mathematical aspects of an exhaustive theoretical study of three-dimensional arrays by G. C. Southworth.

design and planning of systems. These relationships were well brought out, in retrospect, in a review paper[30] by Espenschied in 1937, from which we reproduce here, as Fig. 10-12, a flow-of-the-art diagram for radiotelephony.

Most difficult of all radio problems to understand and deal with, however, was wave propagation over the earth's surface and through that complex medium already encountered by the wireless telegraphers, i.e., the combination of the atmosphere itself and the reflecting-absorbing ionic layer suggested independently by Kennelly and Heaviside in 1902.

The equations of Maxwell had lent themselves readily to solution under simple boundary conditions where the geometry and electrical properties of the boundaries were well defined, as with wires or parallel plates of metal or known dielectric. But in the application to

[30] "The Origin and Development of Radiotelephony," presented before the Institute of Radio Engineers on May 10, 1937. Published in *Proc. IRE* (September 1937).

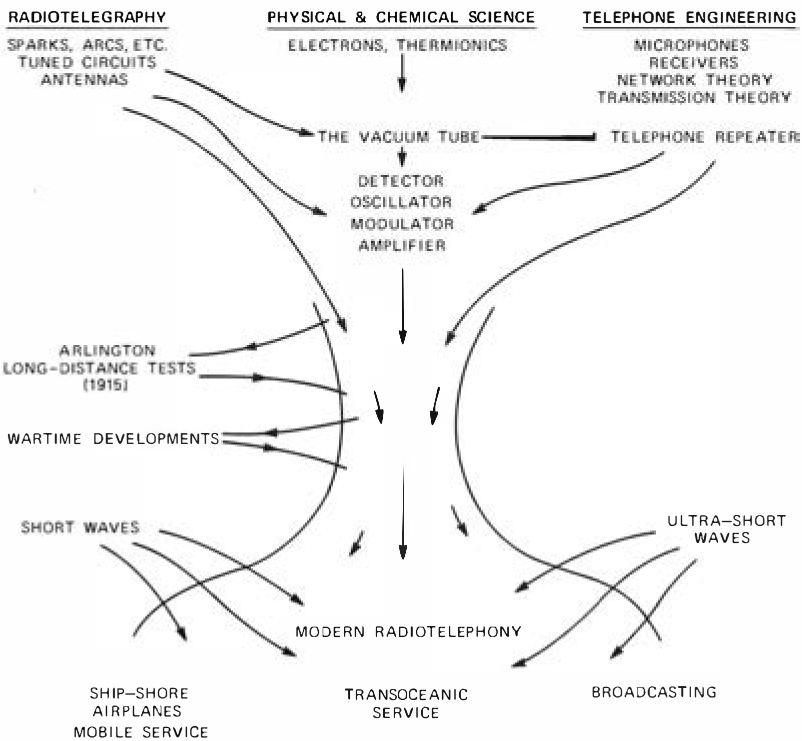


Fig. 10-12. A diagram of the flow of the art in radiotelephony, many of the elements being beneficiaries of progress in fundamental science. (Redrawn from L. Espenschied, 1937)

communication, diffraction over a curved earth of highly variable material had to be dealt with first. Then there was the variation in air density (and hence dielectric constant) with altitude, which tended to bend the waves[31] somewhat around the earth, but was found to be so changeable in character, due to temperature inversions, that waves could on occasion be bent *away* from the earth, or down *toward* the earth, or along the earth's surface in a "duct"-type trajectory.

Finally, and by far the most frustrating, there was the highly complex character of the ionosphere itself. It was not until many years later (the late 1920s and early 1930s) that techniques for exploring this ionized region began to strip the mystery from its seemingly erratic behavior, demonstrating the "layer" to be in reality three or more distinct layers, in the sense that a curve of ion density versus height displayed several peaks whose location shifted from hour to hour and with the seasons, with additional variations from year to year. It was already appreciated, from the work of Eccles (1912) and later Larmor (1924) in England and other investigators, that the effective dielectric constant in an ionized medium could be *less* than unity because the motion of the charged particles—particularly electrons, because of their high mobility—would lag the impressed electric field by 90 degrees, while the normal "displacement" current leads the field by 90 degrees; thus, a positive or negative gradient in the degree of ionization with height would result in additional downward or upward bending of the ray.

Dating from the time of Marconi's triumphant transmission across the ocean (1901), a vast amount of experiment and theory was devoted, in both Europe and America, to radio wave propagation. In Chapter 5 we have recorded many of these findings. To a large extent, as pointed out in that chapter, they were more applicable to the needs of narrow-band radiotelegraphy than to the far more exacting requirements of radiotelephony, with its need for instant two-way uniform transmission over bands that were some kilocycles in width. Thus, the Bell System investigators were confronted with radio transmission problems of vastly greater difficulty and had to devise methods of measurement and analysis[32] all their own.

[31] Much later (1933) a paper by Schelleng, Burrows, and Ferrell of Bell Laboratories showed that this refractive effect, *on the average*, could be taken into account by assuming a fictitious radius of the earth equal to about 4/3 times the actual radius. Alternatively stated, the refraction produces in the wave path, on the average, a radius of curvature four times that of the earth.

[32] The records of radio wave propagation research contain many examples of the interaction between theory and experiment as these proceed cooperatively, often separated in time and distance, each supplying to the other the needed checks or, just as often, the challenges requiring a new look. One good example of this was the theoretical study in 1909, by Sommerfeld in Germany, of the propagation of a

In conformity with the research philosophy preached by Jewett, the approach to these complex problems was characterized by a step-by-step advance in which, at each stage, the number of variable factors was minimized in the interest of reaching a reliable conclusion. An essential in such efforts is precision of measurement. One of the earliest workers in the carrier-frequency and radio-frequency range, C. R. Englund, who was trained at the University of South Dakota and the University of Chicago, moved from a teaching career to enter Western Electric's Engineering Department in 1914. Englund and his associates made many contributions in the high-frequency measurement field, particularly in establishing the shielded bridge, conceived by Campbell[33] ten years earlier, as being far superior to the voltmeter-ammeter type of measuring technique used elsewhere. In Fig. 10-13 we reproduce one of Campbell's diagrams of a shielded bridge—at that time (1904), of course, for audio frequencies only—to illustrate the meticulous attention Campbell had paid at that early date to the factors concerned with precision of measurement. Campbell had gone farther and had shown[34] how distributed capacity and inductance could affect the measured value of a component, such as a resistor, particularly as the frequency increased, and how these effects could be compensated for. There was thus in the air around the telephone laboratory a passion for accurate measurement and elimination of disturbing effects, and as the complex phenomena of radio wave propagation came under study, this meticulousness was to serve well in the sifting of experimental data for significant conclusions.

Englund was also an early advocate (1917) of the method of measuring the strength of received radio waves wherein an adjustable local input is compared with the signal to be measured. He used this method, employing a Colpitts-type oscillator as a local signal source, on an extended series of observations, beginning in 1921, on the signals coming from POZ, the powerful radiotelegraph transmitter at

---

wave along the curved interface of two media, such as the imperfectly conducting ground and the atmosphere above. A similar study, ten years later (1919), by Weyl, also in Germany, led to a different formulation that did not explicitly include a type of "surface wave" appearing in Sommerfeld's result. Many years later (1937), C. R. Burrows at Bell Laboratories discovered that these two formulations differed exactly by the surface wave of Sommerfeld, and in collaboration with his associates L. E. Hunt and A. Decino, conducted the crucial experiment that showed that Weyl's formulation was correct and that the surface wave of Sommerfeld did not exist. Soon afterward, S. O. Rice at Bell Laboratories and K. F. Niessen at Philips (Eindhoven) independently found the source of the error in Sommerfeld's work—the incorrect choice of the square root of a complex quantity in an intricate mathematical derivation.

[33] "The Shielded Balance," *Elec. World and Eng.*, April 2, 1904.

[34] "Resistance Boxes for Use in Precise Alternating Current Measurements," *Elec. World and Eng.*, October 29, 1904.

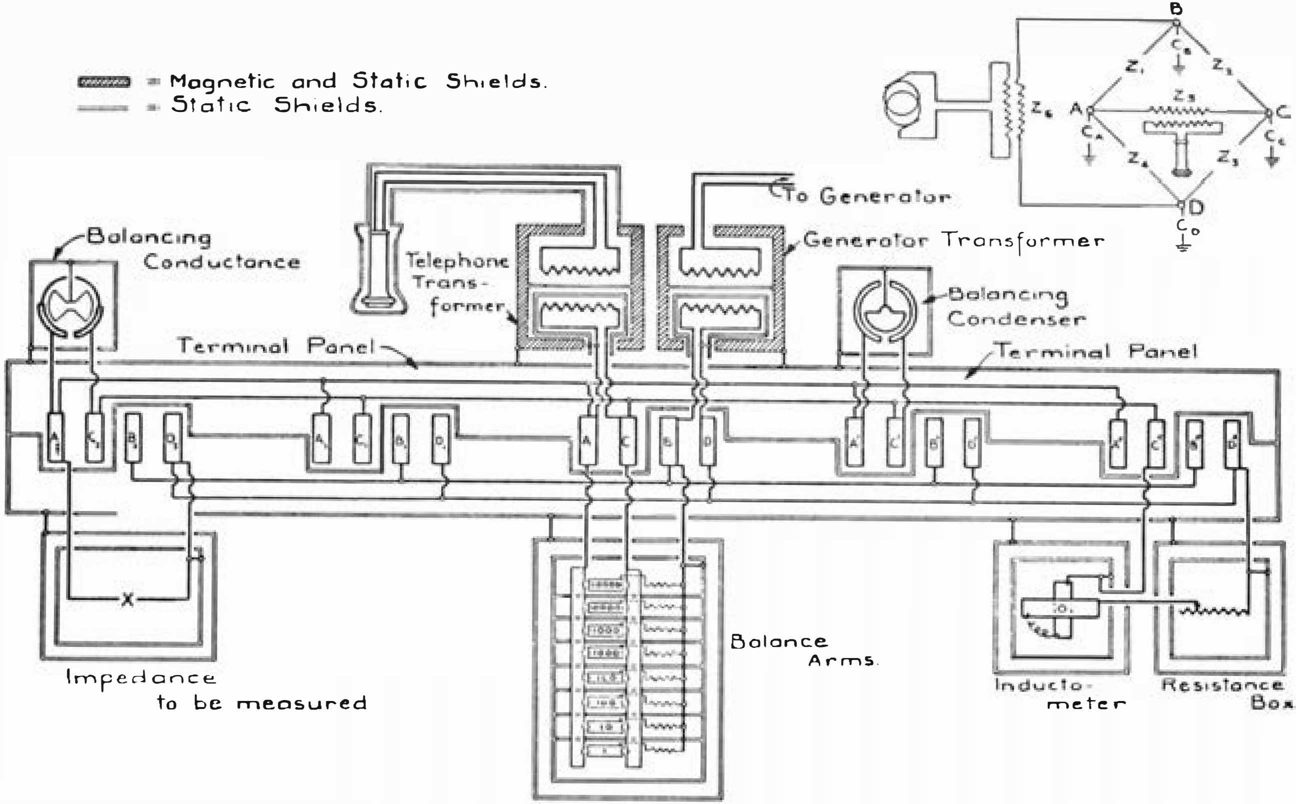


Fig. 10-13. One of G. A. Campbell's bridges or "balances" (at that time for voice frequencies only) showing disposition of both inductive and electrostatic shielding, sometimes involving shields within shields. These principles were carried along into the carrier- and radio-frequency ranges, where accurate measurement was equally important. (G. A. Campbell, 1904)

Nauen, Germany. Englund's paper[35] on this work graphically points up the difficulties in such measurements at that early date due to stray fields, requiring heavy shielding; frequency instability at the transmitting station; strong interference from other stations because of close frequency spacing; and numerous other obstacles to accurate measurement. But there was no doubt that very extensive tests of this type, covering not months but years, would have to be carried out before transatlantic telephony could move ahead from the realm of the possible, as demonstrated in 1915, to a commercially practicable service.

In 1919, World War I having come to a close, two outstanding newcomers, both destined to become directors of radio research at Bell Telephone Laboratories decades later, added their talents to the growing research activity in radio. They were Ralph Bown, who had been a Signal Corps officer and had received his doctorate at Cornell, now joining the Development and Research group at AT&TCo; and H. T. Friis, an electrical engineering graduate of the Royal Technical College of Denmark (1916), joining Colpitts' Research Branch in Western Electric's Engineering Department after studying radio under Morecroft at Columbia University. Friis was hired by H. W. Nichols, a physical scientist of exceptional analytical and inventive ability, who had taken charge of Englund's work, as well as that of Heising, already mentioned. Nichols, a graduate of Armour Institute (now the Illinois Institute of Technology) had come to Western Electric in 1914 from the University of Chicago, which later (1918) awarded him a doctorate in physics. He would undoubtedly have risen high in the ranks but for an untimely death in late 1925.[36]

In their long careers Bown and Friis personified the spirit of research, particularly research in the industrial environment where an awareness of practical goals is in the atmosphere but does not limit the depth to which relevant knowledge can be pursued. Thus the vagaries of wave transmission, as observed in the empirical data being acquired by Englund, were seen as offering a challenge that had to be met. It was apparent that good statistics had to be obtained, as to both signal strength variations and radio noise level; but it was also apparent that explanations had to be sought for these phenomena, in the interest of circumventing the limitations where possible.

[35] "Note on the Measurement of Radio Signals," *Proc. IRE*, February 1923.

[36] One of Nichols' outstanding papers, "Propagation of Electric Waves Over the Earth," co-authored by J. C. Schelleng (who had come in 1919 from the physics faculty at Cornell) and published in the *Bell System Technical J.* in April 1925, developed a mathematical theory of transmission phenomena in a medium containing free electrons and ions and simultaneously subject to the earth's magnetic field, whereby the wave is split into two parts by magnetic double refraction, one of these components being then subject to an absorption that is a maximum at a wavelength of about 200 meters. Appleton in England independently published on this anisotropic effect, by which the behavior of waves in this turbulent region of space was still further complicated.

Through the efforts of Friis and others, the measurement tools employed by Englund were sharpened to provide greater accuracy, convenience, and portability, so that an extensive program of field strength and noise measurement could be followed, continuing on from the very long waves, to be employed in the first transatlantic radiotelephone system, to the shorter (now called "medium") waves where the phenomena of rapid fading and "skip distance" effects due to the ionosphere would dominate the coverage from a transmitter except at the shortest ranges. These studies were therefore of great importance to the infant broadcasting industry as well as to the planning of point-to-point, ship-to-shore, and mobile communication systems.

Looking ahead to the still shorter waves, when it became apparent later in the 1920s that these offered great possibilities for expansion of the transoceanic service, it could be seen that the multiple paths possible through the ionosphere presented a most serious transmission obstacle in the form of *frequency-selective* fading caused by interference between them. Friis and his associates accordingly launched an intensive study of the *angle of arrival* (in the vertical plane) of these waves so that the mechanism of interference could be understood and some form of vertical directivity could be employed in the hope of mitigating the effect of selective fading on transmission quality. Ideally, such studies would have employed test antennas of very high directivity and variable vertical angle so that the waves arriving at different angles could be separated. Lacking such facilities at this early stage, Friis and his colleagues utilized oscillographic observations of the phase difference at two points, and despite the difficulties introduced by multiple waves arriving at different angles, it was possible to report preliminary findings[37] as early as 1928. In the years immediately following, as more refined methods became possible, angle-of-arrival measurements continued to be vigorously pursued. Some of these were pulse methods that showed, as expected, longer propagation times (over the great circle path) for waves arriving at higher angles, further confirming the need for high discrimination in the vertical plane to avoid serious impairment of signal quality in a working system. A still further refinement to result from these studies would be the multiple unit steerable antenna[38] or MUSA, whereby the elevation angle could be varied electronically for best reception of the incoming signal.

In the absence of distortion, and provided the signal at the input terminals of the radio receiver has sufficient margin over "set

[37] H. T. Friis, "Oscillographic Observations on the Direction of Propagation and Fading of Short Waves," *Proc. IRE*, May 1928.

[38] H. T. Friis and C. B. Feldman, "A Multiple Unit Steerable Antenna for Short Wave Reception," *Proc. IRE*, July 1937.

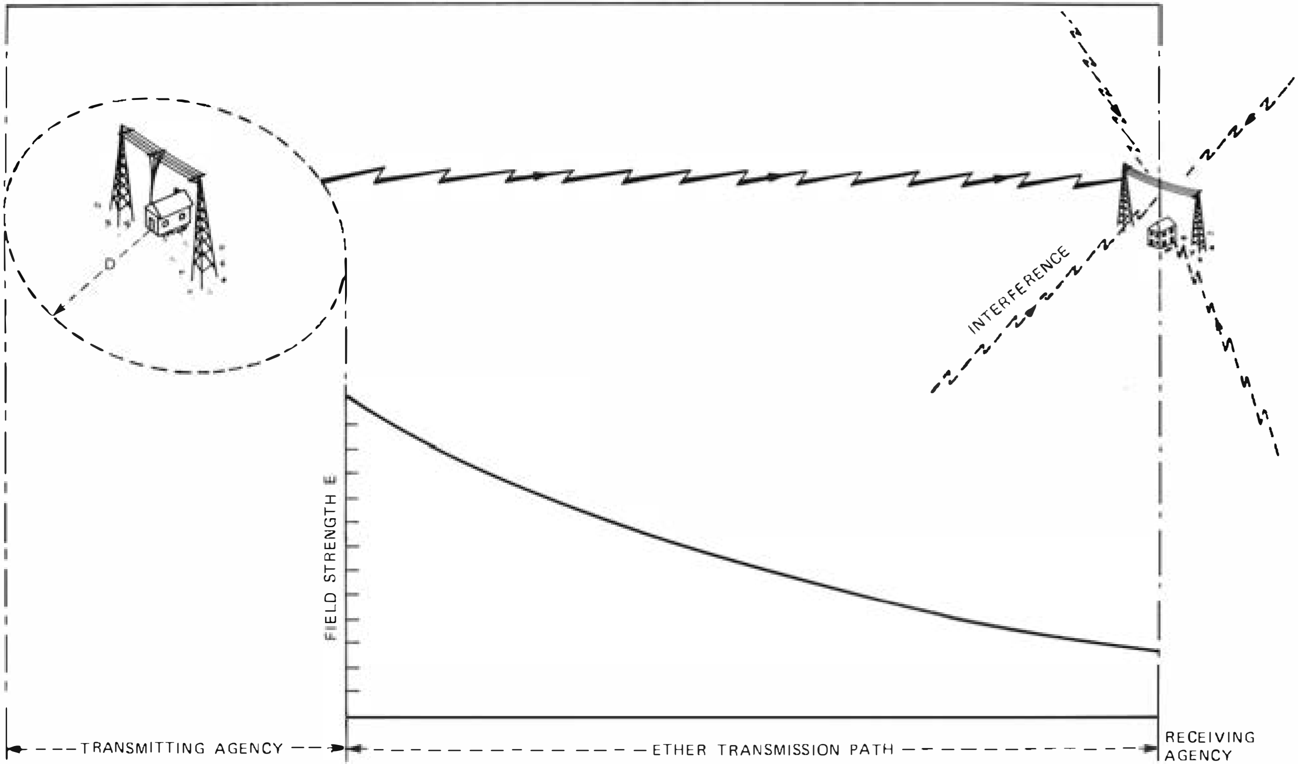


Fig. 10-14. Transmission analysis of radio circuit. (Redrawn from R. Bown, C. R. Englund, and H. T. Friis, 1923)

noise"—i.e., over Johnson or thermal noise (already discussed) and tube noise—the limiting factor in transmission then becomes the noise encountered in the medium itself, in the form of atmospheric static or other disturbances. It was apparent from the beginning that any strategems that could be devised to reduce such interference would make possible great economies in other parameters, of which transmitted power and antenna size were most obvious. Consequently the study of static, in relation to magnitude and direction of origin, was an essential part of the research program that looked ahead to commercial service. While the literature on manual measurements of static was already plentiful, there had been little continuous automatic recording.[39] Moreover, any meaningful application to telephony had to include specific consideration of the effect of the static on the intelligibility of speech.

These broad considerations were covered in 1923 in a classic paper[40] by Bown, Englund, and Friis, from which we show here, as Fig. 10-14, their initial diagram displaying the three distinct parts into which a radio system naturally splits from a transmission standpoint. It was shown how the gains in the transmitting and receiving portions would have to be apportioned, by the radio system designers of the future, so as to compensate for the loss in the transmission path, at minimum cost, while satisfying the necessary criteria of service quality. It was characteristic of these authors to explore carefully the subjective aspect of received noise and to consider adequate margins of signal over noise as one of the most essential service criteria. For this purpose they introduced the idea of a simple type of artificial speech whose amplitude could be adjusted so as to be just discernible through the noise. Thus, in the vast quantities of data taken over periods of many months, looking ahead to the first transoceanic radiotelephone service, it was the essential factor of speech-to-subjective-noise ratio rather than radio field strength alone that received prime attention.

An important AT&TCo contributor to these studies was R. K. Potter. Because of major contributions that Potter made later, related to analysis of speech, we shall have more to say about his career toward the end of Section III of this chapter.

Long after radio, employing both short and long waves, had become a well-established service in the Bell System, the measurement of radio noise was to continue. For the tools of measurement would become more refined and productive as time went on, and the character of noise itself would change as the frequency range continued to expand.

[39] H. T. Friis, "A Static Recorder," *Bell System Technical J.*, April 1926.
[40] R. Bown, C. R. Englund, and H. T. Friis, "Radio Transmission Measurements," *Proc. IRE*, April 1923.

"Set" noise, for example, became predominant beyond 200 MHz, rather than atmospheric static or man-made interference, or the extra-terrestrial type of noise[41] that was to be discovered at Bell Laboratories in the early 1930s.

Since radio was still considered, in the 1920s and 1930s, the only means by which multichannel telephony could economically bridge the ocean for some time to come—and indeed was to remain so for nearly three decades, including the years of World War II—these fundamental studies paid for themselves many times over in making it possible to live with the shortcomings of an erratic, defiant medium and create highly successful and useful systems. More than this, the men involved in these studies laid the groundwork and set up an organized, talented group for even more comprehensive and diversified approaches to the problems of the future—a future that would make radio a competitor to wire lines over land, even giving it superiority in some applications; a future, too, that would extend radio to frequency ranges far beyond the few hundred megahertz reached in the 1930s, even to the millimeter wavelength range where new mechanisms could be invoked for the guidance of waves by conductors.

Thus have Friis and his associates contributed to a vast accumulation of knowledge and art, extending for half a century. Here we cite only a few of their many publications not already mentioned, as representative of the character and broad scope of their work in the 1920s and the succeeding few years:

R. Bown, D. K. Martin and R. K. Potter, "Some Studies in Radio Broadcast Transmission," *Proc. IRE*, February 1926.

A. G. Jensen, "Portable Receiving Sets for Measuring Field Strengths at Broadcasting Frequencies," *Proc. IRE*, June 1926.

J. C. Schelleng, "Note on the Determination of the Ionization in the Upper Atmosphere," *Proc. IRE*, November 1928. (A further note by the author appeared in the same journal in August 1929.)

G. C. Southworth, "Certain Factors Affecting the Gain of Directive Antennas," *Proc. IRE*, September 1930.

C. B. Feldman, "The Optical Behavior of the Ground for Short Radio Waves," *Proc. IRE*, June 1933.

The numerous contributions of this same group in later decades, when they continued as leaders of a worldwide attack on radio propagation problems and their ramifications, are left for recording in a later volume of this history.

[41] Two papers by K. G. Jansky describe some of these investigations: "Directional Studies of Atmospherics at High Frequencies," *Proc. IRE*, December 1932, and "A Note on the Source of Interstellar Interference," *Proc. IRE*, October 1935.

While radio with its inherent novelty and new challenges was a focus for the interest of numerous applied scientists no longer pressed by the critical demands of World War I, other aspects of communication were likewise beckoning for their attention. One of these was the transmission of visual information, which had been a goal of research workers from the earliest days of electrical communication but had resisted any practical advance. The blossoming of electronics now revived hopes for the long-dreamed-of transmission of visual images, with television—images in motion—a more distant goal. Into this field came Herbert E. Ives, with a doctorate from Johns Hopkins, Signal Corps experience in aerial photography, and a lively interest in optical and photographic problems. Ives was the leader in the technical aspects of the work in Western Electric's Engineering Department leading to a demonstration of long-distance telephotography over telephone wires in 1924, and at H. D. Arnold's instigation drew up a proposed program of research in television, starting with a modest attack on the most fundamental problems, with provision for expansion as these might be solved and give way to problems of a developmental character.[42]

In this way, as in the Englund–Friis approach to radio, invasion of the new territory of television proceeded by thoughtful steps toward objectives only dimly seen at first, but becoming clearer as fundamental obstacles gave way to a heightening onslaught. The developments and early demonstrations, through the late 1920s, resulting from this program have been described in Section 9.8 of Chapter 7. Here we make note of just two examples of fundamental contributions that were to demonstrate their worth many years later:

(*i*) The suggestion by P. M. Rainey[43] that the elements of a picture be represented, for transmission purposes, by coded binary (two-valued) signals in order to circumvent the imperfections of the medium. More than a decade later, when a corresponding proposal was made for *speech* transmission by A. H. Reeves of ITT, the idea was still considered audacious because of the greatly increased signal frequency band required; yet by mid-century the development of "pulse code modulation" was given high priority, and soon afterward this technique figured prominently in the transmission plant of the Bell System.

(*ii*) The invention by Frank Gray[44] of a method of transmitting two television signals within a frequency range normally sufficient

[42] An example of Ives' early interest in signals for transmitting motion pictures was his paper "A Theory of Intermittent Vision," *J. Opt. Soc. Am. and Rev. Sci. Instr.*, June 1922.

[43] P. M. Rainey; U.S. Patent No. 1,608,527; filed July 20, 1921; issued November 30, 1926.

[44] Frank Gray; U.S. Patent No. 1,769,920; filed April 30, 1929; issued July 8, 1930.

for only one. The method was based on use of the *interstices* discovered by Gray (and independently by P. Mertz, whose work has already been cited) between the spectrum lines of a television signal, these lines being spaced at intervals equal to the line-scanning frequency. More than two decades later, the great value of this invention was appreciated when it became necessary to transmit, independently, a "luminance" signal and a "chrominance" signal to handle color television; and Gray received, in 1953, months after his retirement from Bell Laboratories, the Vladimir K. Zworykin Award of the Institute of Radio Engineers for this far-ahead-of-its-time contribution to electronic television.

Gray, a Ph.D. from Wisconsin, where he was on the physics faculty, had come to Western Electric in 1919, and under the inspiration of Ives and others he was to make numerous innovative contributions in a career of three decades.

Going back now to the earliest days of the telephone, it will be recalled that the switching of telephone calls did not at first seem to be a field requiring deep analysis, but it developed some aspects in the 1890s that called for mathematical talent of a high order; and, as we have seen earlier in this section, the names of M. C. Rorty and E. C. Molina became known worldwide for bold and original thinking in trunking theory and the application of probability theory to switching problems. Going now into the 1920s, as telephone growth continued[45] and the power of new methods introduced by Erlang of the Copenhagen Telephone Company was recognized, a train of theoretical studies was set in motion that continued strongly into later decades. Molina, himself a devotee of Laplace, with the additional inspiration of association with Campbell, led this effort in its theoretical aspects[46] and was instrumental in attracting the talent that would be needed in the future, when "common control" switching systems—discussed in Section 4.3 of Chapter 6—would introduce additional degrees of freedom into the switching plant and would require even more powerful analytical methods.

The design of circuits, the economic study of automated systems, and the theory of probability all seemed to complement one another in the person of Molina. His work even made a contribution to the planning of transmission systems by showing the combined effects of

[45] A comprehensive review of the development of trunking theory overseas, which paralleled the thinking in the U.S.A. in this field, is given in a classic publication by G. F. O'Dell, "An Outline of the Trunking Aspect of Automatic Telephones," *J. Inst. of Elec. Eng.* (London), 1927. O'Dell also was the author of an excellent bibliography appearing in the *Post Office Elec. Eng. J.*, October 1920.

[46] "The Theory of Probabilities Applied to Telephone Trunking Problems," *Bell System Technical J.*, November 1922.

a multitude of irregularities in loaded lines—a problem that had to be studied intensively as loading came into use.

One of Molina's admirers, himself a distinguished mathematician, was T. C. Fry, who came to Western Electric from the Wisconsin faculty in 1916. Fry, whose work at Western in the 1920s, largely concerned with probability theory, ran closely parallel to Molina's at AT&TCo, paid tribute[47] many years later to two sides of Molina's character, the scholarly side as revealed in his expert knowledge of Laplace, and the very practical side as displayed by his conception of the *translator*, a great basic invention (1905) that completely freed the automatic telephone switching mechanisms from slavish dependence on the decimal numbering system and was later embodied in all of the great computing systems.

Fry's own contributions demonstrated the power of mathematical methods in numerous applications, often in collaboration with physicists or engineers we have already mentioned; for example, with Hartley in 1922[48] on statistical aspects of studies on the binaural location of complex sounds; and with Ives, also in 1922[49] concerning phenomena in photoelectric cells. Coincidentally with J. B. Johnson's work on thermal noise, Fry published on the theory of shot noise[50] due to randomness in electron emission. Fry's celebrated book *Probability and Its Engineering Uses* (Van Nostrand, 1928) included basic contributions to the theory of stochastic (random) processes, with applications to telephone traffic, thus tying in with Molina's work and the accelerating programs in traffic and trunking theory.

We conclude this section with a brief mention of another field, related to transmission, whose flowering commenced in the 1920s as a natural area for the combined talents of telephone scientists and their mathematical collaborators—the field of electromechanical-electroacoustic systems. It was pointed out early in this section that frequency-dependent electrical elements and the reactance concept had originally become understood in the 1860–1890 period through their mechanical analogues. Now, a half-century later, with the maturing of electric circuit theory under Campbell, Carson, Zobel, Norton, and others we have mentioned, the rebound from their discoveries was felt in full impact, especially in electroacoustics. Thus it became possible to progress from the primitive methods devised by Alexander Graham Bell, through an evolution recorded in Chapter 3, to far more

[47] Internal Bell Laboratories memorandum, January 31, 1949, to R. I. Wilkinson from Fry, who was then Director of Switching Research and Engineering.

[48] R. V. L. Hartley and T. C. Fry, "Binaural Locations of Complex Sounds," *Bell System Technical J.*, November 1922.

[49] H. E. Ives and T. C. Fry, "Voltage-Current Relation in Central Anode Photo-Electric Cells," *Astrophys. J.*, July 1922.

[50] T. C. Fry, "Theory of the Schroteffekt," *J. Franklin Inst.*, February 1925.

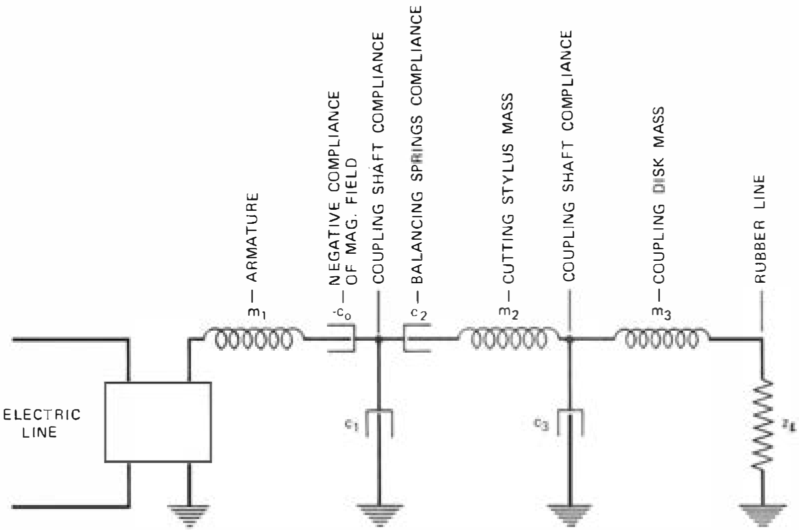


Fig. 10-15. Equivalent circuit of electromagnetic recorder. One of the earliest detailed expositions of the application of classical electrical network theory in electroacoustic systems, bringing out the various electrical analogs of mechanical elements. (Redrawn from J. P. Maxfield and H. C. Harrison, 1926)

sophisticated methods of electromechanical coupling that would provide higher efficiency of energy transfer and vastly improved frequency response. We reproduce here, as Fig. 10-15, one of the early diagrams from a classic 1926 paper by Maxfield and Harrison[51] wherein the electrical and acoustomechanical analogies are clearly brought out.

We leave further discussion of this field to the next section, where the context of speech and hearing phenomena will provide a more natural setting. But we point out that, as with radio, it was the spirit of uncommitted research, fostered by Jewett, Arnold, and their associates, which inspired much of this recorded advance. "Try as we may, and as we ought," wrote Arnold[52] in 1926, "to maintain an even and considered front in our attack on the boundaries of knowledge, there are always some salients which will yield only to siege or to extended flanking operations. So we find in the department men who are patiently and cunningly attacking old problems—problems which it might seem we had passed in our rapid progress, but which have

[51] J. P. Maxfield and H. C. Harrison, "Methods of High Quality Recording and Reproducing of Music and Speech Based on Telephone Research," *Bell System Technical J.*, July 1926.
[52] H. D. Arnold, "Organizing our Researches," *Bell Laboratories Record*, June 1926.

still remained unconquered, and are frequently key positions of the greatest value. Compensating for these long-established sieges are slender lines of adventure which have been thrust forward into the unknown far beyond any present hope of consolidation. In this virgin territory we find men, whose success must depend largely upon their own initiative and resourcefulness, striving for some point which may bring with its winning the conquest of new and broad regions."

In this section, the potency of the scientific method applied to fundamental problems, as voiced by Carty and quoted at the beginning of this chapter, has required only a modest number of examples to demonstrate its worth to industry itself, as well as to the scientific community. In the fields to be covered by Sections III and IV we shall similarly cite selected examples, rather than be exhaustive, and hope to depict further through these illustrative cases the intellectual climate in which research thrives and can lead, in Carty's words, to "profitable improvements in the art, worth many times their cost."

III. RESEARCH IN SPEECH, HEARING, AND SOUND

Having no records of the speech of primitive men, we could assume that it consisted of elementary sounds, perhaps in the nature of growls and grunts; and that these, supplemented by physical gestures, conveyed the simple ideas that had to be communicated. We can imagine that, with time, men learned to control the organs of speech with more skill and precision to impart shades of meaning, and to edge and emboss the vowel sounds with an increasing variety of cosonants and diphthongs, giving the ear and brain of the listener a more complex pattern of sounds to interpret. With the evolution of *language,* it then became possible to introduce reasoning, logic, and persuasion on a higher intellectual plane; so that now, for many thousands of years, the race has had the benefit of a medium of communication capable of accommodating even the most abstruse ideas and the most sensitive emotions, yet with apparatus that is already built-in at birth.

It is only when the *range* of voice communication has to be extended (beyond a few feet or a few hundred feet), or when a *record* of the voice must be preserved, to be listened to at a later time, that man-made devices, the telephone and the recorder-reproducer, must be brought into play. Until a century ago, hardly anyone thought of the voice as a thing that could start from here and be heard somewhere else, for the voice was a fleeting thing, not to be transported anywhere, not to be stored anywhere. That a sound might be uttered, but not immediately lost forever—except in people's memories—was the limit in fantasy:

> Antiphanes said merrily, that in a certain city the cold was so intense that words were congealed as soon as spoken, but that after some time they thawed and became audible; so that the words spoken in winter were articulated next summer. —Plutarch

The stimulus for Bell's invention of the telephone was his recognition of the value of the spoken word, in an instant two-way exchange, as compared with communication in any other form. It was a recognition that came from his dedication to human speech—to giving the power of speech to people who could not speak because they were deaf and could not hear anyone else speak. With this appreciation of the importance of vocal communication in the life of human beings, any possible extension in the *range* of communication was something to be ardently sought.

As compared to the discoveries in astronomy, chemistry, and electricity, there has been little in the histories of science concerning speech, hearing, and acoustics. It would seem that the marvelous flexibility of the speech process and the analytical ability of the hearing process were too commonplace to excite wonderment. Yet there are some names that stood out and a few discoveries that provided a background of knowledge not only for Alexander Graham Bell but for his father and grandfather, in whose steps he eagerly followed. For example, Dr. John Wilkins, an English clergyman–philosopher, had published a book in 1668 clearly showing many of the relations between voice sounds and the tongue and lip positions used in producing them, and had devised a series of schematic drawings of the mouth that could to some degree specify speech sounds. A more accurate representation was provided by an ingenious "visible speech" presentation devised in 1867 by Alexander Graham Bell's father, Alexander Melville Bell. From such a record made by the father or son, the other could accurately imitate even unalphabetical articulations such as a yawn, a hiccough, or a clearing of the throat. Thus the Bells understood perfectly the relations between the parts of the vocal tract and the sound produced.

Most notable of the works providing a basis for the studies of the Bells was that of the anatomist, physiologist, and physical scientist Helmholtz in Germany, especially his publication in 1862 of "Die Lehre von den Tonempfindungen," on sensations of tone as a physiological basis for the theory of music. We have also noted, in Section II of Chapter 1, the use of the "phonautograph" of Leon Scott, a French engineer–scientist, and the manometric capsule of Koenig, a contemporary of Helmholtz, in connection with Bell's early visions of how telephony might be carried out, as well as his attempts to portray speech waves to his deaf pupils. We also noted his use of the bone mechanism of an actual human ear in attempting to get a more realistic phonautograph record. These were surely research efforts in the true sense, the exploration of new territory; but once the breakthrough had been accomplished, and the telephone had become a commercial thing, the record shows little more of truly fundamental work in speech, hearing, and sound for several decades. Speech itself was so elusive in its content, and the hearing/understanding process so highly subjective, as to defy quantitative analysis until vastly improved instrumentation for measurement became available.

The vacuum tube amplifier, arriving on the scene in 1913, gave new life to studies of speech as it had done for telephony itself; for, with amplification, it was possible to make measurements at levels far below the sensitivity of the ear, marvelously sensitive though that organ was, and to do so with a quantitative precision not previously possible.

It had, of course, been appreciated that the intelligibility of human speech, when electrically transmitted and then reproduced, must depend on the overall responsiveness of the system in some regions of the voice spectrum more than in others. And, to the extent that "naturalness" or recognizability of individual voices was of concern at that time, it must have been suspected that the frequencies contributing thereto were not the frequencies that were most essential to good intelligibility. But all of this had remained qualitative only; for, as we noted in Section II, the objective had necessarily been to transmit and reproduce as much power as possible at every frequency.

Thus when Irving B. Crandall, a professor of physics and chemistry with degrees from Wisconsin and Princeton, came into the research ranks of Western Electric in 1913, a rich field for investigation was opening up and the new tools for its cultivation were being invented by his own colleagues in the form of oscillators and amplifiers, and in the form of filters and other frequency-dependent elements. The art of analyzing electrical currents into their component frequencies and accurately measuring them, and the related art in connection with mechanically vibrating systems, were rapidly advancing, and it was apparent that great advantages would come from similarly analyzing speech and hearing; for an accurate knowledge of every part of a system, from the voice through the telephone instruments to and including the ear, would permit more intelligent design of the parts under control.

A factor that had been especially difficult to analyze and measure was the ability of an individual to recognize small defects in familiar sounds. It is remarkable how quickly we note a small change in a friend's voice, and with what uncanny skill a musician will detect minute imperfections in complex sounds. Accordingly, it could be seen that a study of the speech and hearing process as related to telephony would require construction of devices so nearly perfect that the keenest ear could not find a flaw in the rendition, and then the step-by-step introduction of measured imperfections until an observer could detect a fault. Then, by collating the observations from a wide variety of subjects, the degree of perfection that might be demanded of the telephone instrument could be forecast in a reasonable way.

The philosophy, therefore, of the investigations of speech and hearing—captained by Crandall beginning in 1914—was, as later expressed[53] by H. D. Arnold, "to get an accurate physical description and a measure of the mechanical operation of human ears in such terms that we may relate them directly to our electrical and acoustical instruments; to test the keenness of the sound-discriminating sense and find what is the smallest distortion which the mind can perceive and how it reacts to somewhat larger distortions; and thus to reach a reasonable basis of design both for separate instruments and for systems, as a whole, to give a proper balance between cost and performance."

We sense again, in these words of Arnold's, the spirit of fundamental research in the industrial environment, where the long-range usefulness of new knowledge can be envisioned long before it can be commercially demonstrated. Arnold's was, in fact, the guiding spirit that conceived and initiated the plan for acoustics research that Crandall and his colleagues would carry out; for it was Arnold, more than anyone, who recognized the need for precise measurement, the need for a basic transmission reference system with components as nearly perfect as possible, and the desirability of an excursion into the psychological and physiological aspects of speech and hearing. Less than 15 years later, though in Arnold's words some parts of the work were "hardly more than started," yet the results had been so great, both for the original purpose and for the many issues that had since arisen, that they presented "a unique exemplification of the worth of systematic and sustained research."

The attack was first launched most vigorously on the constitution of speech in an effort to establish some kind of description of average speech, and to find to what extent small imperfections and variations affected intelligibility. This work obviously could not go far without including a study of the organs of speech and the organs and mechanism of hearing, and its scope came to be extended as well to some of the abnormalities in these faculties. As better and more precise instruments became necessary, some of the effort was devoted to securing devices that would convert sound waves into electrical form and reconvert them again to sound with the least possible distortion. Out of this were to come unexpected rewards to the

[53] From Arnold's introduction to *Speech and Hearing* by H. Fletcher, D. Van Nostrand Company, Inc., 1929.

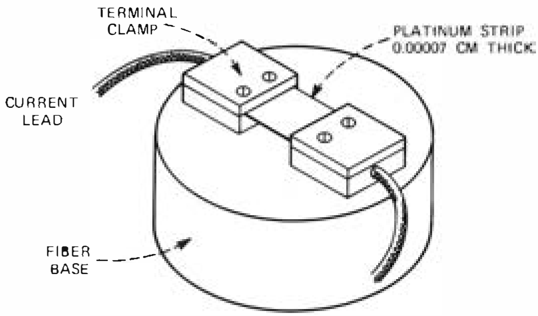


Fig. 10-16. A simple form of thermophone for precision generation of sound waves. (Redrawn from H. D. Arnold and I. B. Crandall, 1917)

telephone and phonographic arts, as well as a succession of important scientific and technological advances.

Together, Crandall and Arnold developed the thermophone,[54] a device based on earlier work abroad, particularly the discovery by F. Braun[55] that acoustic effects could be produced by passing alternating currents through a bolometer. Requiring a laboratory sound source that was simple and dependable, and capable of reproducible results, Crandall and Arnold started with a simple form of instrument as depicted in Fig. 10-16 (reproduced from their 1917 paper). The periodic heating of the very thin metallic strip, in response to an alternating current fed through the strip, set up temperature variations in the film of air immediately adjacent, which in turn created sound waves as a result of the thermal expansion and contraction. The instrument was therefore unique in that there was no reaction on the driving source from the medium as would be the case with a mechanically vibrating type of instrument.

In a mathematical analysis backed up by careful experimentation, the conditions were derived for maximum efficiency of sound generation consistent with purity of response, i.e., the absence of extraneous frequencies such as harmonics and sum-and-difference tones. As a source of single-frequency sound waves, the thermophone could be actuated by a sinusoidal current of half the desired frequency, since the instantaneous heating in the resistance R of the strip due to a current i, varying sinusoidally with time, would be

$$i^2R \sin^2 pt = \tfrac{1}{2}\, i^2R\,(1 - \cos 2pt).$$

This mode of operation presupposed an adequate supply of pure single-frequency alternating current to bring the average temperature

[54] "The Thermophone as a Precision Source of Sound," *Phys. Rev.*, July 1917.
[55] *Ann. Physik*, 1898.

of the strip up to a level where sound would be efficiently generated. In the case of a weaker source, the sound output could be increased by supplying direct current for heating the strip, in which case the signal input would be at the desired frequency instead of half the desired frequency. In this mode of operation, it was necessary to restrict the alternating-current amplitude to a value much smaller than the direct current so that the second harmonic generated would not be excessive.

Where the input to the thermophone was to be a combination of frequencies, as was more nearly representative of speech waves, this second mode was essential, the direct current again being several times the amplitude of the alternating components so that the temperature wave and consequent sound wave generated would faithfully reproduce the sinusoidal components and be sufficiently free of harmonics and sum-and-difference tones.

Since the thermophone was a basic instrument for the researches that were to follow, it was essential for its fundamental properties to be known quantitatively. In their analysis, Arnold and Crandall determined its theoretical frequency response under two conditions: first, with the sound wave permitted to travel freely away from the thermal strip as a plane or spherical wave; second, with the strip placed in a small cavity for the purpose of producing pressure changes that could actuate the ear or some mechanical "phonometer" or pressure-measuring instrument that would constitute one wall of the enclosure. The analysis showed the response to be greatly different for the two cases, being proportional to frequency in the first case and *inversely* proportional to the *cube* of the frequency in the second case.

Both of these relationships, with the unaided human ear used as a comparator, received surprisingly good experimental confirmation. In these tests, the thermophone and a telephone receiver with which it was compared as to acoustic output were looked upon as two pistons that communicated their amplitudes of motion to the adjacent medium, the strength of each source being proportional to the area of the imagined piston times the amplitude of its motion. The telephone receiver had been calibrated as a sound generator by measuring the motion of the diaphragm with a microscope, and it was known that for small vibrations the bowing of the diaphragm gave an integrated displacement equivalent to that of a piston having 0.306 times its area.

A much more accurate and satisfactory verification of the theory was made possible by the construction of a successful phonometer, in the form of a condenser-type microphone, by Crandall's colleague E. C. Wente, who came to Western Electric in 1914 with degrees from the University of Michigan and M.I.T., to which a doctorate from Yale would be added in 1918.

The principle of the condenser microphone, in which the spacing between plates of a condenser is varied by the pressure of a sound wave, was not new, having been suggested[56] as early as 1881; but in its early forms, and lacking an amplifier to make up for its feeble response as a transducer, it had not been a useful instrument. What

[56] "Condensateur Employé comme Transmitteur," *La Lumière Electrique,* 1881.

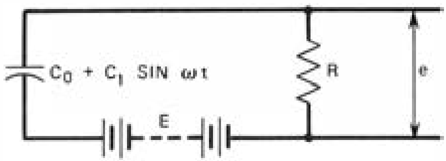


Fig. 10-17. Circuit used in E. C. Wente's analysis of the condenser microphone. (Redrawn from E. C. Wente, 1917)

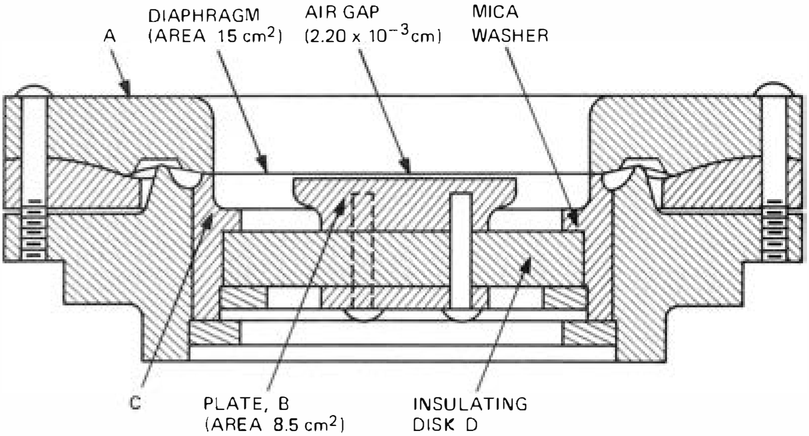


Fig. 10-18. Sectional drawing showing construction of condenser microphone. (Redrawn from E. C. Wente, 1917)

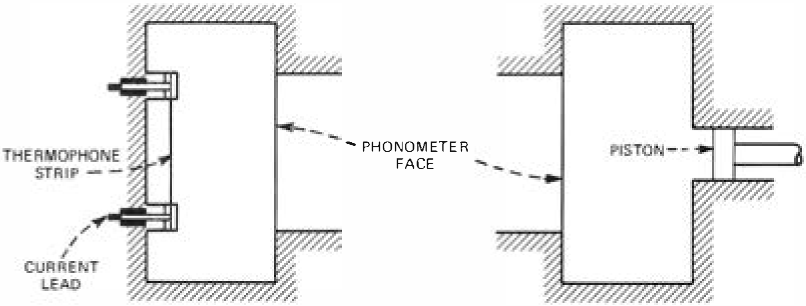


Fig. 10-19. The use of Wente's condenser microphone to verify Arnold's and Crandall's theory on the efficiency of the thermophone as a source of sound. (Redrawn from H. D. Arnold and I. B. Crandall, 1917)

appealed to Wente was the fundamental simplicity of the device and the possibility of extending its inherent mechanical resonance far beyond the frequency range of interest so that its response might be smooth and reproducible. In a 1917 companion paper[57] to the paper of Arnold and Crandall previously cited, Wente analyzed the device mathematically and gave details of construction of a successful instrument. Wente's paper shows the basic circuit considered in his derivation, here reproduced as Fig. 10-17, whereby a small sinusoidal variation of amplitude C_1 in the capacity of a condenser C_0, in series with a battery of voltage E, was shown to be equivalent to the insertion, in series with C_0, of a generator having a sinusoidal open-circuit voltage $E(C_1/C_0)$.

Also reproduced from Wente's paper is Fig. 10-18, which gives some details of construction, notable features being the use of a very thin steel diaphragm stretched nearly to its elastic limit, with a very small air gap for maximum sensitivity, the plate B being ground extremely flat so that several hundred volts could be used without leakage or flashover. With such construction, it is obvious that the thin film of air between plates would have very significant damping effects, and the analysis of these effects was an essential part of Wente's and Crandall's studies of the instrument.

Figure 10-19, reproduced from the Arnold–Crandall 1917 paper already referred to, shows how Wente's phonometer was used to corroborate their theory on the acoustical output of the thermophone. The thermophone element was placed in a small enclosure whose walls included the diaphragm of Wente's instrument. Then, to eliminate any absolute calibration of the phonometer, an identical one was incorporated in a second enclosure excited by a piston at the same frequency. The pressure variations caused by the piston were easily calculated from mechanical considerations, and the comparison was easily made by matching the electrical outputs of the phonometers. The experiment verified Arnold and Crandall's theory to within 5 percent, a remarkable result in that pioneering stage in electroacoustic science.

The field of electroacoustics in which Crandall and Wente and their colleagues were laboring was, of course, far from being an exclusive domain of Bell System workers; for the invention of the telephone by Bell had been quickly followed by the invention of the phonograph by Edison and the development of variations thereof, such as the graphophone and gramophone. Alexander Graham Bell himself had participated in these efforts. Other workers had devised oscillographic methods for portraying sound waves visually for study and analysis.

[57] "A Condenser Transmitter as a Uniformly Sensitive Instrument for the Absolute Measurement of Sound Intensity," *Phys. Rev.*, July 1917.

Among those responsible for these achievements the names of Hermann, Bevier, Blondel, and Duddell testify to the international character of the rising scientific interest in acoustical science, whereby the theoretical and largely mathematical expositions by Rayleigh—in his remarkable two volumes, *The Theory of Sound*, first printed in 1877—could now have greatly amplified opportunity for experimental demonstration.

One of the most useful instruments for study and demonstration of tone qualities in speech and music was the Phonodeik (1909) devised by Dayton C. Miller[58] of the Case School faculty. In the Phonodeik (meaning to show or exhibit sound), a diaphragm was so coupled to a tiny mirror as to cause rotation of the mirror in response to an impinging sound wave. A beam of light focused on the mirror could then be deflected to give a greatly magnified excursion on a moving photographic film or, for observation while experimenting, on a piece of ground glass. Since the parts of the device were extremely light, the Phonodeik could be used for portrayal of harmonic components in a sound wave as rapid as 10,000 vibrations per second, and it was therefore a widely used instrument in scientific laboratories.

The Crandall group at Western Electric participated in true "research" fashion in the worldwide dissemination of new knowledge in acoustics; and though their work was not aimed at satisfying the growing public mania for talking machines, inevitably some of their creations were adapted to these entertainment uses. Toward the end of Section II we saw (in Fig. 10-15) how the new understanding of electrical-mechanical analogues was making it possible to fashion transducers by employing electrical transmission theory, this work by Maxfield and Harrison being an indispensable contribution to the development of high-quality recording and reproduction of speech and music. Even before this, Wente's condenser microphone had been hailed as launching a new era in high fidelity;[59] and numerous later achievements of Wente's, aimed primarily though they were at solving fundamental problems, found wide adaptation in the livening atmosphere of sound entertainment and radio broadcasting. J. R. Pierce, many years later, referring to the solid theoretical foundations upon which Wente's ideas were always built, remarked[60] that "When Wente designed a horn, he understood its cutoff and other acoustical

[58] Miller's book, *The Science of Musical Sounds*, first published in 1916, includes a detailed account of various methods employed for sound portrayal in the late 1800s and early 1900s.

[59] Olson and Massa of RCA, in their book, *Applied Acoustics*, generously remarked that "Modern acoustics may be said to have begun with the development of the condenser microphone by Wente."

[60] Address at the presentation of the Gold Medal of the Acoustical Society of America to Wente, May 15, 1959.

properties. He knew why his tweeter had to be cellular and how cellular it had to be. When he designed a transducer to drive his horns, he understood its electrical and acoustical properties in a very fundamental way."

Pierce further noted that "Like many physicists, Wente possessed an excellent mechanical sense and great mechanical skill. Rather than airily waving his ideas into the hands of others, there to meet disaster, he could cleverly design and bring into reality the things he envisioned."

Several other youthful scientists of exceptional promise were recruited for the work on speech, hearing, and sound in that period before the needs of World War I began to draw heavily on their talents. Notable among these were R. L. Wegel, coming in 1914 (the same year as Wente) from the Edison Laboratory after specializing in theoretical physics at Wisconsin; C. F. Sacia, a graduate of the University of Michigan in electrical engineering, coming in 1916; and Harvey Fletcher, likewise arriving in 1916, a graduate of Brigham Young and head of the physics department there, with a doctorate from the University of Chicago. Fletcher was destined to assume the responsibilities of Crandall upon the latter's death in 1927 and was eventually (1933) to become director of physical research of Bell Telephone Laboratories.

The efforts of this group being largely redirected to investigation of devices for underwater sound transmission and detection, it was not until after the war ended—more specifically, 1922—that their previous and subsequent results on speech and hearing could begin to appear in *Physical Review* and in *The Bell System Technical Journal*, founded in July of that year. Some of the large number of fundamentally important papers, as well as papers from other laboratories, are referred to in a 1925 paper[61] by Crandall, which starts with the historical setting for the speech studies then in progress. A view of the main line of Western Electric research effort in this field is given by the following brief list of selected titles, supplementing the 1917 papers already cited.

(*i*) H. Fletcher and R. L. Wegel, "The Frequency-Sensitivity of Normal Ears," *Phys. Rev.*, June 1922.

(*ii*) C. E. Lane, "Minimum Sound Energy for Audition for Tones of High Frequency," *Phys. Rev.*, May 1922.

(*iii*) I. B. Crandall and D. Mackenzie, "Analysis of the Energy Distribution in Speech," *Phys. Rev.*, April 1922.

(*iv*) R. L. Wegel, "The Physical Characteristics of Audition and Dynamical Analysis of the External Ear," *Bell System Technical J.*, November 1922.

[61] I. B. Crandall, "The Sounds of Speech," *Bell System Technical J.*, October 1925.

(*v*) H. Fletcher, "The Nature of Speech and Its Interpretation," *J. Franklin Inst.*, June 1922.
(*vi*) H. Fletcher, "Physical Measurements of Audition and Their Bearing on the Theory of Hearing," *J. Franklin Inst.*, September 1923.
(*vii*) R. L. Wegel and C. E. Lane, "Auditory Masking and Dynamics of the Inner Ear," *Phys. Rev.*, February 1924.
(*viii*) I. B. Crandall and C. F. Sacia, "A Dynamical Study of the Vowel Sounds," *Bell System Technical J.*, April 1924.
(*ix*) H. Fletcher, "Useful Numerical Constants of Speech and Hearing," *Bell System Technical J.*, July 1925.
(*x*) C. F. Sacia, "Speech Power and Energy," *Bell System Technical J.*, October 1925.
(*xi*) J. C. Steinberg, "The Loudness of a Sound and Its Physical Stimulus," *Phys. Rev.*, October 1925.

Prior to the sensitivity studies reported by Fletcher and Wegel in reference (*i*), which they had described before the National Academy of Science in November 1921, there had been numerous investigations over a 50-year period, beginning with those of Toepler and Boltzmann in 1870 and Rayleigh in 1877, with results that were too widely disparate to be of practical use. For one thing, there had been no appreciation of the great variation—as much as a thousandfold in terms of acoustical energy—in the hearing ability of supposedly "normal" ears. Beyond this, there was the lack of any adequate apparatus for measurement, the sound sources being typically organ pipes in which light interference methods had to be used to determine the amplitude of vibration of the air particles, or a whistle whose energy output was calculated from the pressure used in blowing it, or a tuning fork mounted on a resonator whose energy output was calculated from the decrement that it imposed on the vibration of the fork.

Fletcher and Wegel showed their results, as compared with those of other investigators, in the first figure of their paper, which we here reproduce as Fig. 10-20.

The new tools making these more accurate measurements of audition possible were, of course, the vacuum tube oscillator and amplifier, the condenser transmitter, and the thermal receiver, together with a special attenuator whose range of variation in current output was more than three-millionfold. The assemblage as a whole was known as an audiometer and the curve relating sensititivity to frequency was called an audiogram.

From these and other measurements on a large number of subjects considered to have normal hearing, Fletcher produced a highly significant pair of curves published as his first figures in reference (*vi*) showing the *threshold of audibility* and the *threshold of feeling* for

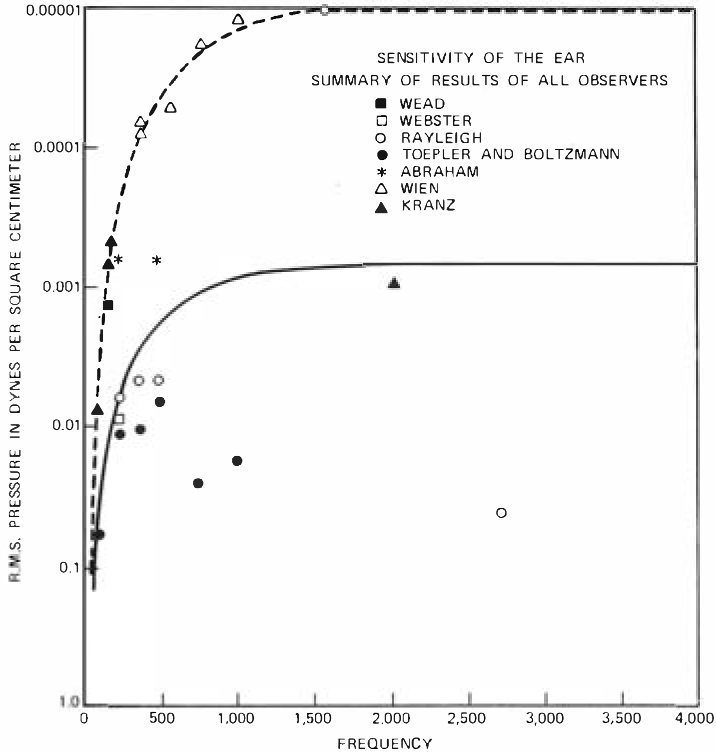


Fig. 10-20. Redrawn from a 1921–22 paper by H. Fletcher and R. L. Wegel, showing scattered results of earlier investigators on the frequency sensitivity of the ear. Their own measurements are shown by the solid line.

the average human ear, as related to frequency. Reproduced here as Fig. 10-21 is a pair of such curves, which thus become the boundaries of the area of auditory sensation. Sounds of lower intensity at the indicated frequency would not be detectable, even in the complete absence of interfering noises, while sounds exceeding the upper threshold would be detected first as a tickling sensation or pain. The curves bring out graphically, first, that the frequency range of maximum aural sensitivity is in the vicinity of a half-dozen octaves; and second, that over much of this range the ear has the remarkable capacity to accommodate a range or ratio of energy levels of the order of a million million.[62]

[62] For good telephonic speech, fortunately (for economic reasons), a frequency range of about four octaves and an energy range more like 100,000 to 1 are adequate; but, as demonstrated dramatically by Fletcher and associates a few years later, the full artistic effect of musical renditions is achievable only by accommodating most of the audible frequency range and most of the intensity range bounded by the curves.

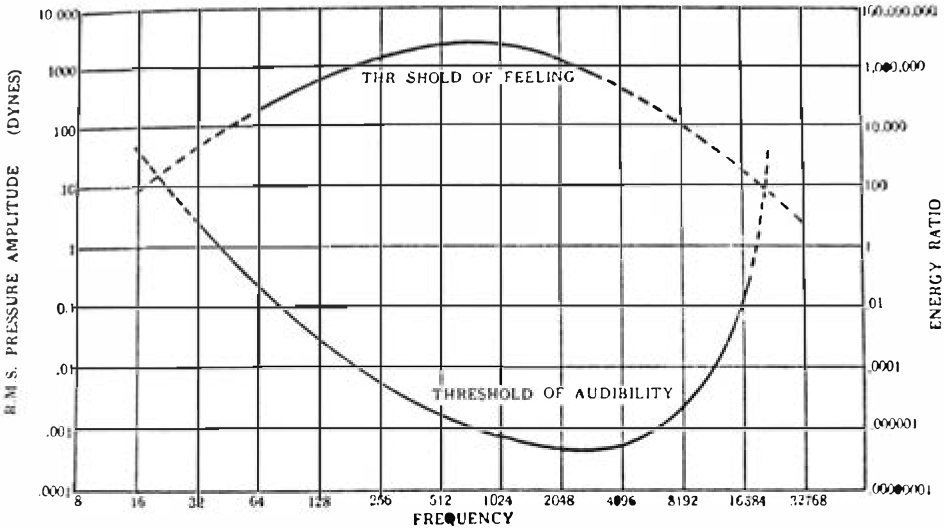


Fig. 10-21. Curves representing the "area of auditory sensation" as explored by Fletcher and Wegel. The above figure, essentially Fletcher's curves but with simplified coordinates, is reproduced from Fig. 1 in the volume *Transmission Circuits for Telephonic Communication* by K. S. Johnson, Bell Telephone Laboratories, 1927.

On these curves the ordinate is plotted logarithmically, in recognition of the psychological "law" established much earlier by the German physicists Weber (1834) and Fechner (1860), whereby the sensation caused by a stimulus is proportional to the logarithm of the stimulus. As applied here, the increment in sound pressure that is just observable to the ear bears a constant ratio to the original pressure, the ratio being found to be about 1.1 over a wide range of pressures and frequencies. Since a similar relationship holds, over a wide range, in respect to *pitch* sensibility—the ratio in this case being 1.003—the abscissa or frequency scale employed by Fletcher and Wegel was likewise logarithmic, as is the keyboard of a piano. It would thus be possible to subdivide the auditory sensation area between the two curves into a large number of small squares, each of which could be described as representing a frequency and amplitude just distinguishable from its neighbors.

From the number of differentiable values[63] of intensity, measured as approximately 270, and the number of differentiable frequencies,

[63] Fletcher obtained these data through the work of V. O. Knudsen, a graduate student at the Ryerson Physical Laboratory, University of Chicago. Knudsen's results, published in a paper "The Sensibility of the Ear to Small Differences of Intensity and Frequency," *Phys. Rev.*, January 1923, were more accurate than those of many

approximately 1,300, it could be said that there are about 300,000 pure tones that can be perceived as being different; and the extent to which an individual's area of auditory sensation fell short of this number could be taken as a measure of hearing deficiency. For specific purposes, however,—such as intelligibility in telephony—it is obvious that a criterion emphasizing a more restricted frequency and amplitude range would have more significance. Fletcher and his colleagues investigated a large number of cases of deficient hearing to correlate the audiograms with the subjective conversational experience of the individuals.

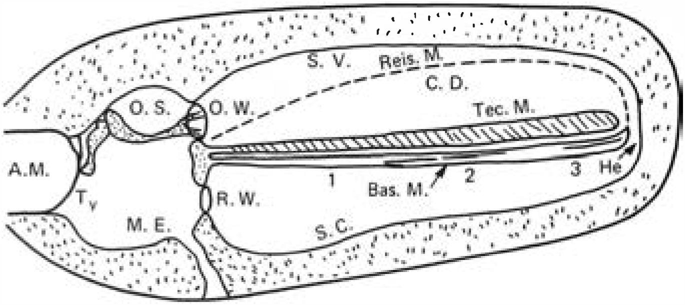
The very large number of distinguishable pure tones could lead one to expect that the number of distinguishable *complex* tones would be astronomical; but because of a phenomenon known as *masking*, to be discussed presently, there are limitations in the hearing mechanism which place bounds on this number.

From 1910 to the early 1920s, there were already many theories of hearing, and Fletcher and his colleagues sought to reconcile these with their own subjective observations, being made with electrical and acoustical measuring tools never before available, as well as with the new understandings of complex dynamical systems which they owed mainly to their electrical network theory confreres. Fletcher, in reviewing a few of these theories in reference (*vi*), published a diagrammatic representation of the auditory function, which we here reproduce as Fig. 10-22. The representation starts with the external ear canal or meatus (on the left), which terminates in the eardrum, whose minute vibrations (of the order of 10^{-8} centimeter) are mechanically transmitted by way of three tiny bones or ossicles, located in the air-filled middle ear, to the small and comparatively rigid oval window, which then excites the dense, incompressible fluid filling the inner ear.

The Bell family, and the early otologists like Clarence Blake (from whom Alexander Graham Bell obtained, for experiment, these mechanical parts from a cadaver, as related in Section II of Chapter 1) had been well aware of the functioning of these portions of the auditory system and of the important transforming or "impedance-matching" role of ossicles. What was now in question related more to the complex functioning of the inner ear, especially the mechanism responsible for pitch discrimination and for excitation of the neural pathways to the brain.

---

earlier investigators because of improved instrumentation and a quieter environment. Later work at Bell Telephone Laboratories, with further refinements in measurements and covering a greater range of frequencies and intensities, indicated a figure above 500,000 for the number of distinguishable tones. This work is described in the following papers: "Differential Pitch Sensitivity of the Ear," R. Biddulph and E. G. Shower, *J. Acoust. Soc. Am.*, October 1931; "Loudness and the Minimum Perceptible Increment of Intensity," R. R. Riesz, *J. Acoust. Soc. Am.*, January 1933.



A.M.	Auditory meatus	O. W.	Oval window
Bas. M.	Bas. mem. including organ of Corti	Reis. M.	Reissner's mem.
C. D.	Cochlear duct	R. W.	Rd. window
E.	Eustachian tube	S. C.	Scala cochlea
He.	Helicotrema	S. V.	Scala vestibuli
M. E.	Middle ear	Tec. M.	Tectorial membrane
O. S.	Ossicles (malleus, incus, stapes)	Ty.	Tympanic membrane

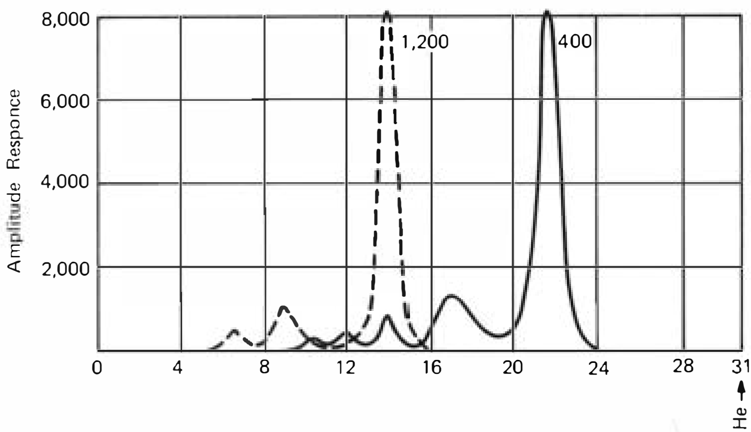


Fig. 10-22. Fletcher's diagrammatic representation of the auditory function. The cochlea (actually a snail-shaped organ) is shown "rolled out" for clarity. The lower graph, based on calculations by R. L. Wegel and C. E. Lane, shows relative vibration amplitudes and positions (in millimeters from the oval window) along the basilar membrane, for tones of two frequencies, each 80 units (minimum detectable increments) above the threshold of audibility. (Redrawn from H. Fletcher, 1923)

In the earliest formulation of the Helmholtz theory, it was supposed that the organ of Corti, located between the basilar membrane and the tectorial membrane, acted like a set of resonators, large in number and sharply tuned. Each tone would stimulate a single resonator, depending upon its pitch. Later this theory had been somewhat modified when it was thought that the resonant property might reside in one of the membranes in the cochlea.

Some theories bypassed the pitch question by placing the burden on the neural pathways to the brain and suggesting that pitch discrimination took place there in some manner.

A theory held by many, and seemingly consistent with the dynamical constants involved, suggested that the maximum amplitude of vibration of the basilar membrane would occur at a distance along the membrane determined by the pitch, the high tones stimulating regions near the base and the low tones regions near the apex of the cochlea; and that the pitch would be judged by this position of maximum response. It was an amplification of this theory that Fletcher arrived at, postulating that only those nerves are stimulated that are at particular parts of the membrane vibrating with more than a certain critical amplitude, and that the pitch is judged from the part of the membrane where the nerves are stimulated. For loud pure tones there would be several such regions of maximum amplitude, as shown in the lower graph of Fig. 10-22 (calculated by Wegel and Lane), corresponding to the tone and the harmonics introduced by the non-linear response of the middle and inner ear, the latter maxima increasing very rapidly with increased loudness.

Since, according to the anatomists, there were only a few thousand nerve cells in the basilar membrane, with four or five fiber hairs per cell, there would not be nearly enough discrete units (even if each hair fiber could act as a unit) to account for the 300,000 (or more) differentiable tones. It was thus necessary for Fletcher to postulate that a large number of units must act at one time and that the ear and brain must become educated to interpret the spatial and amplitude configuration, observing differences in the intensity of excitation of each nerve cell, as well as the position of each one excited.

It is the spatial or prism-like character of the basilar membrane's response that permits the hearer to sense two mixed tones as distinctly two tones, while the visual mechanism cannot sense individually two mixed colors, since the retina and the visual cortex of the brain are already burdened with a spatial function in two coordinates.

On the other hand, as indicated a few paragraphs earlier, the auditory system's response to complex tones becomes greatly complicated at higher intensities because of the "masking" effect of one tone upon another. It had been known for nearly 50 years that strong tones of low frequency had a powerful effect in partially or completely obliterating sensitivity to higher-frequency tones, and this observation was considered fatal, by many, to any theory of hearing that involved resonance. But when the phenomenon was studied in great detail by Wegel and C. E. Lane (who had joined him, from the University of Iowa, in 1921), as reported by them in their 1924 paper, reference

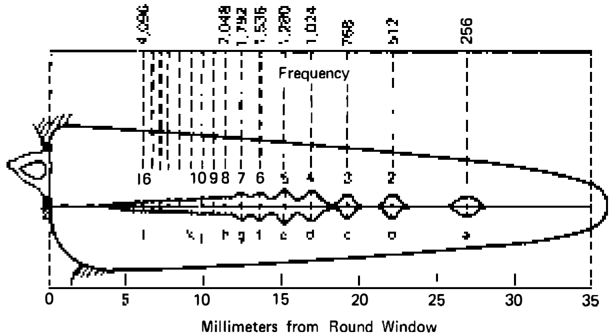


Fig. 10-23. Resonant response of the basilar membrane to a strong 256-hertz tone, illustrating the potential of such a tone to mask higher-frequency tones by overlapping of the vibratory regions. (Redrawn from a published note by R. L. Wegel, 1925)

(*vii*), it proved to be the very thing that would furnish experimental data from which the vibration characteristics of the inner ear could be calculated. The lower graph in Fig. 10-22, already noted, is an illustration of these calculations. We also reproduce here, as Fig. 10-23, a longitudinal section of the basilar membrane showing the pattern of deformation (with amplitudes greatly magnified) in response to a strong musical tone of pitch 256 (middle C). Taken from a subsequent note by Wegel (*Nature*, September 12, 1925), the pattern shows clearly why a strong low-frequency tone, through second-order distortions, can obliterate the response to a weaker high-frequency tone, while a high-frequency tone would have to be extremely powerful before the excited region could overlap that of a tone of substantially lower frequency.

It could have been suspected that masking effects would still be observable if two tones were introduced into opposite ears instead of the same ear, but this was found not to be the case. It was observed that only a minute portion of the sound energy impinging on one ear could affect the other, this being the portion transmitted by the bones of the head. Thus the masking phenomenon was clearly not a "central" or brain effect but was attributable to non-linearity in the individual ear, the creation of "subjective" harmonics of a lower-pitched tone; these harmonics excite nerve terminals that would otherwise be receptive to stimulation by a higher-pitched tone. The inference from this observation was that interference with telephone conversation by room noise is not principally due to noise impinging on the free ear, but to noise getting into the same ear to which the telephone receiver is being held, due mainly to leaks under the receiver cap. This inference was confirmed by direct experiments with telephone users.

The new concepts about the hearing process that were generated by Fletcher and his colleagues came at a time when otologists, physiologists, psychologists, and physicists had been at odds, each sponsoring a theory that accorded with his own training and point of view. The work in the Western Electric laboratories accordingly had a unifying effect that was universally applauded, manifesting itself in numerous papers by joint authors in cooperating fields. An outstanding example of such cooperation was a 1923 paper[64] by Wegel and E. P. Fowler, M. D., "Audiometric Methods and Their Applications." Fowler, an outstanding otolaryngologist, in a published tribute[65] to Fletcher many years later, wrote that "One of my treasured memories is my association with Harvey Fletcher and that wonderful group of men who worked with him in the Acoustical Research Laboratories of the Western Electric Company." Calling Fletcher the father of modern scientific acoustics, Fowler also praised Wegel for his part in devising practical audiometers for use in measuring deficiencies in hearing.

Going far beyond studies of the hearing process itself, exhaustive tests were made on various forms of distortion and their effect on syllabic articulation and overall intelligibility. The pioneering studies by Campbell[66] on intelligibility were thus greatly extended through the use of a laboratory telephone system that could reproduce speech with practically no distortion, and in which Campbell's filters were an essential component. By means of distortionless attenuators, equalizing networks, and adjustable low-pass and high-pass filters, a vast amount of information was obtained, which is best summarized in a book by Fletcher[67] published a few years later. We reproduce here, as Fig. 10-24, one type of result that would bear quite directly on the planning of telephone systems, namely, the effect upon articulation of eliminating frequencies below a certain value and above a certain value. The economic importance of such quantitative knowledge is of great significance in engineering the transmission plant and providing satisfactory telephone instruments, as has been brought out in Chapters 3 and 4. The results shown by Fig. 10-24 are quite striking. For example, a system which transmits only frequencies below 1,000 hertz exhibits a syllable articulation of only 40 percent, although 84 percent of the speech energy is present; while a system which rejects those frequencies and transmits only those above 1,000 hertz exhibits a syllable articulation of 86 percent, a degree of recognition that for many purposes might be satisfactory, although only 17 percent of the speech

[64] *Annals of the American Rhinological, Laryngological and Otological Society*, June 1923.
[65] *Archives of Otolaryngology*, January 1967.
[66] "Telephonic Intelligibility," 1910, *loc. cit.* (Section II of this chapter).
[67] H. Fletcher, *Speech and Hearing*, 1929.

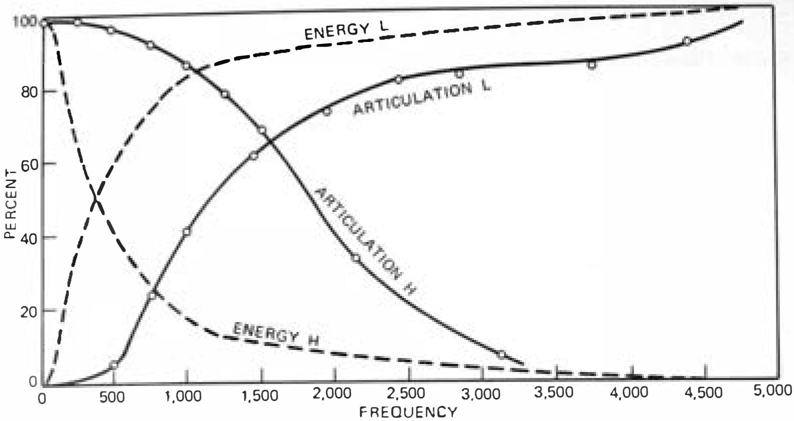


Fig. 10-24. The solid curves show the effect upon articulation of transmitting only frequencies *below* a certain value (Articulation L curve) and *above* a certain value (Articulation H curve). The dotted curves show the average speech energy transmitted in each case. (Redrawn from H. Fletcher, 1929)

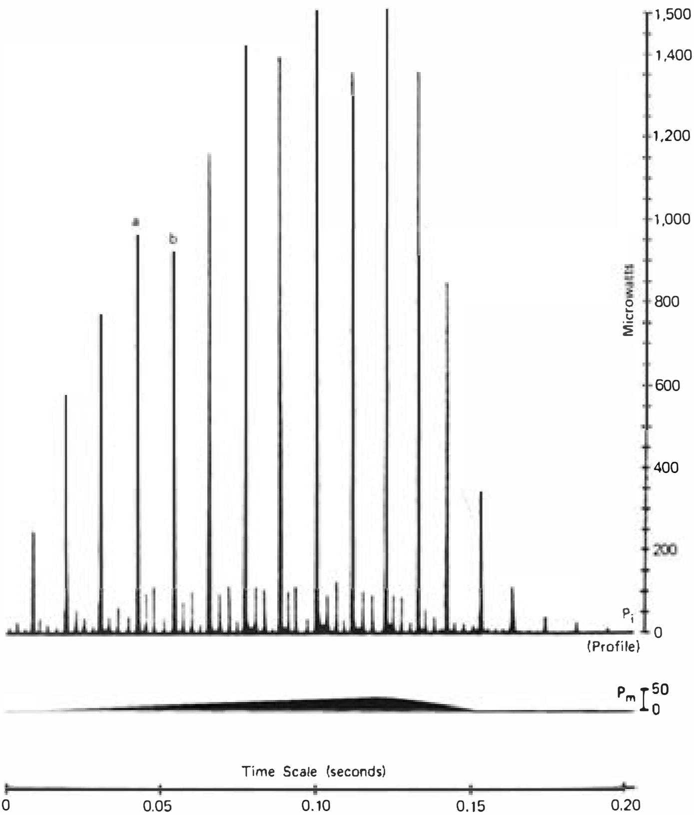


Fig. 10-25. Instantaneous and mean power. Enlarged copy of original oscillogram of the word "quite." (Redrawn from C. F. Sacia, 1925)

energy is present. In the latter system, however, the speech reproduced would sound very peculiar, its "naturalness" being destroyed; hence, both ends of the voice spectrum had to be respected in designing the telephone plant.

Since speech intelligibility varies somewhat with sound level, the intensity of the received speech, in obtaining these results, was adjusted by means of an attenuator to give maximum articulation for each setting of the filters.

While the frequency range necessary for good telephony was thus being studied, another investigation of importance was not concerned with frequency as the argument, but rather with speech amplitude or energy as a function of time—for a salient characteristic of speech waves was the generally high ratio of peak value to mean or mean-square value. Failure to take this into account could cause serious load distortion in speech-transmitting amplifiers by suppression of the peaks—an effect readily detected by an accustomed ear, provided that the rest of the system is relatively distortion-free. C. F. Sacia was the principal investigator of this aspect of speech waves, some of his major conclusions being published in reference (*x*).

Reproduced here as Fig. 10-25 is one of Sacia's diagrams illustrating, for the word "quite," both the "instantaneous" power P_i, which of course varies from zero to high values on each cycle[68] of the wave, and the "mean" power P_m, as measured over an interval of a hundredth of a second (i.e., approximately one cycle of the fundamental frequency). The latter is the power that would be recorded by a very fast-acting wattmeter.

The ordinate values (microwatts) indicated in the figures were determined by measuring the pressure on a condenser transmitter 9 centimeters from the speaker's lips, and translating this into terms of total emitted microwatts by using fairly well-established assumptions about energy distribution over the wavefront. The fluctuations in amplitude from one instantaneous power peak to the next were explained by Sacia as a modulation or tremolo in the voice which can appear even in so short a syllable as this word, and having in this case, a periodicity of around 0.02 second.

From such measurements Sacia determined the "peak factor," a term which he applied to the square root of the ratio of a peak value of P_i to the corresponding value of P_m. The peak factor was obviously a useful index of the waveform. In extended tests

[68] When a wave contains strong even harmonics, the instantaneous *power* includes a component at the fundamental frequency, while for a pure sine wave or a wave with only odd harmonics, the power wave fluctuates with double the frequency. Thus, in the figure, *a* to *b* is the actual vocal period, or reciprocal of the pitch.

with a variety of sounds and speakers, Sacia arrived at a peak factor of 5 as making a reasonable allowance for the worst cases. "Hence," he wrote, "the effective voltage should not exceed one-fifth the overload voltage of the system."

These findings of Sacia's, employing apparatus of careful design, were of fundamental importance and came at a time when demands in the entertainment field, as well as in telephony, were requiring better understanding of the nature of speech waves.

Also of much interest to acousticians was the highly subjective question of loudness of a sound as related to its physical stimulus. The simple logarithmic relationship of Weber and Fechner, already referred to, had obvious limitations. It could not, for example, be depended upon near the threshold of hearing, nor could it be expected to apply at high sound intensities because of non-linearities in the hearing mechanism. It was likewise suspect in the case of complex tones, of which speech and random noise were prime examples.

J. C. Steinberg, joining Western Electric in 1922 with a doctorate in physics from the University of Iowa, collaborated with Fletcher in investigations of loudness reported in 1924, and then, in the paper listed as reference (*xi*), published the derivation of a formula covering the general case of both voiced and unvoiced sounds over a wide range of intensities. The derivation had to take account not only of the energy frequency spectrum of the sound, but of the threshold peculiar to each component in the presence of all the others. The result, therefore, had to be consistent with the theories of masking that were being published at the same time.

Steinberg's formulation for the loudness of a voice-like sound—i.e., a sound characterized by discrete frequencies more than 60 hertz apart—was expressed in terms of the logarithm of a summation of an appropriate *root* of the weighted spectral components of pressure:

$$L = (10/3)r^2 \log \sum_{i=1}^{i=k} (W_i P_i)^{2/r},$$ [69]

where the weight factor W_i for the i^{th} component depended on its frequency and on the sensation level for the sound as a whole, while the root r, shown applied to each squared weighted sound

[69] Steinberg was amused, 50 years later, at the suggestion that he could have omitted the factor 10/3 in this expression if he had used the base-2 logarithm instead of base-10, since the ratio of these logarithms is 10/3. Few scientists in the 1920s has reason to foresee the importance which the binary notation would assume with the advent of information theory and digital computation. Moreover, the use of common or base-10 logarithms in communication was being stimulated by the Bell System's promotion of the Transmission Unit and later the decibel, as recounted in Section 5.1.1 of Chapter 4.

pressure, depended on the sensation level only, being 1 for sounds near the threshold of audibility but rapidly rising to 3 in the middle and high intensity range. *Sensation level*, as used here, was a term arising in earlier work based on root-mean-square pressure (with respect to threshold of audibility for a 700-hertz tone), but not on the weighted pressure spectrum.

Experiments showed that in such voice-like sounds, since there is no overlapping of excited regions along the basilar membrane at low intensities, components which would not be strong enough to be heard alone do not, in ensemble, contribute to the loudness. Hence the above summation was properly applied only to those components with pressures above their minimum audible pressures when heard alone. With other types of sounds, where the various components were not multiples of a fundamental voice pitch and were too close together to be resolved by the ear, this independence, of course, did not hold and it was necessary to modify the formulation by an artifice that considered the totality of energy in each of a plurality of narrow bands, of which each could be considered as a single frequency.

Like many of the investigations we are mentioning in speech and hearing, the study of subjective loudness entailed large amounts of experimental data painstakingly planned and recorded. Such data are of permanent usefulness because they relate to unchanging human characteristics, not to devices or materials that will become obsolete with advances in technology. The more immediate usefulness of Steinberg's results was in other investigations where it was necessary to calculate how much the sensation level of one sound had to be changed so that its subjective loudness would be the same as some other sound. Examples would be the case where a portion of the frequency range was to be removed from a complex sound, or the case where known amounts of distortion were to be introduced into a system and the effect on loudness would have to be predicted.

Equally fundamental in character, and likewise exhaustive and requiring new refinements in experimental gear, were the dynamical studies of the vowel sounds in speech, as reported in 1924 by Crandall and Sacia in the paper listed as reference (*viii*), supplemented by graphical records of semivowel and consonant sounds as reported in Crandall's 1925 paper (*loc. cit.*). The anatomy and physiology of the speech organs producing these sounds had, of course, been well known to the Bells, Alexander Graham Bell himself having been a teacher of vocal physiology; and the brilliant investigations of Helmholtz, cited early in this section, had provided a sound basis for the view that the oral cavities could be considered as resonators, responsive or unresponsive to the various Fourier components of the sound wave originating at the vocal cords.

The oral cavities constitute a vibrating system of two or three degrees of freedom, the theory of which had been developed by Rayleigh and others, and it could be expected that the vowel qualities would be well accounted for by postulating harmonic forced vibrations in these cavities. There had evolved, however, a supposedly competing theory whereby the "characteristic" frequencies of the vowel sounds were the natural vibrations or *transients* in the oral cavities when excited impulsively by the (more or less) periodic puffs of air from the glottis, the opening between the vocal cords. According to this "inharmonic" or transient theory, no harmonic relations need exist between the characteristic frequencies of the vowels and the fundamental or cord tone accompanying them, and the classical Fourier analysis would not be considered applicable in resolving the vowel sound into simpler components. A number of investigators had constructed machines on this principle which, like others based on the forced or harmonic theory, produced sounds resembling human speech.

This difference of views was closely analogous to the controversy concerning electrical network theory reported in Section II of this chapter, wherein the transient approach of Heaviside had to be reconciled with a "steady state" approach having more practical usefulness in telephone engineering. And, as in that case, it became appreciated that the two methods of analysis were fundamentally the same and would yield similar results, as would the experiments, when properly interpreted. Speech being a variable phenomenon, with the cord tones often unstable and with rapid elision from one sound to another as both the pitch and the cavity configurations are manipulated, a certain *lack* of periodicity was to be expected in any recorded phenomena. In speaking, as distinguished from singing, the way the sounds are started and ended has much to do with the ability to recognize them. The characteristics of the vocal cords determine the type of voice and identify the speaker, but it is the *modulation* of the cord tone that imparts the distinguishing characteristics that permit recognition of the many different sounds of speech.

The analysis and interpretation of both vowel and consonant sounds made possible by the highly accurate waveform records of Crandall and Sacia shed much light on this aspect of human speech and its production. These records were the realization of an objective sought by Alexander Graham Bell a half-century earlier, when the smoked-glass vowel tracings depicted in Fig. 1-2 of Chapter 1 were the best that his crude phonautograph would yield. The tests of Crandall and Sacia made use of a Wente condenser transmitter, a multistage amplifier, and an oscillograph, all specially designed and calibrated, the overall response being substantially free of amplitude and phase distortion up to a frequency of 5,000 hertz. Illustrative of the specialized character of this apparatus was the final stage of the amplifier, which employed eight vacuum tubes in parallel to act as a "current transformer" working into the low impedance of the oscillograph vibrating element. To assure themselves of the fidelity of the system, the experimenters generated square-topped acoustical waves up to 300 hertz in frequency and determined, from observation of the oscillogram under this severe test, that the highest harmonics they sought to observe were being faithfully registered.

In selecting the vowel sounds to be analyzed, Crandall and Sacia made use of the "vowel triangle" or diagram shown in Fig. 10-26, well known to phoneticians. In this diagram there are eleven standard "pure vowel" sounds from *oo* to long *e*, with two related vowel sounds *ar* and *er* interpolated in their proper places. The diagram also shows four "transitional" vowels, with arrows indicating direction of transition. The triangle is intended to represent the change in lip and tongue positions as different sounds are made, beginning with the *oo* sound, where the lips are rounded, the tongue raised, and there is a large resonating cavity in the front part of the mouth and a smaller and less important one in the throat. Passing down the left side of the triangle from *oo* to *a* (as in *father*), the mouth is gradually opened and the tongue lowered, the throat resonance

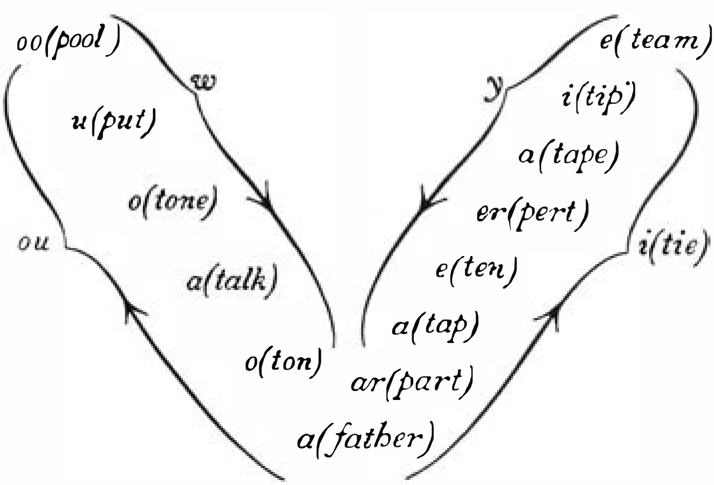


Fig. 10-26. The "vowel triangle" used by phoneticians to illustrate (left to right), in a rough manner, the lowering and subsequent raising of the tongue and the gradual opening and subsequent partial closing of the mouth as the indicated sounds are uttered.

playing only a minor part. The sound *a* (*father*) is, as it were, the center of gravity of the diagram and occupies the key position in the phonetics of most languages. Going up the right side of the triangle from *a* to long *e*, the tongue is gradually raised to the front part of the mouth, thus forming two major resonance chambers. It was therefore an object of the Crandall–Sacia study to correlate the waveform analysis with the resonances that might be expected from these configurations of the oral cavities, and also the nasal cavities, whose role becomes dominant in the case of sounds like *m*, *n*, and *ng*.

The Crandall–Sacia results included 160 graphical records of vowel and consonant sounds, of which 104 were vowels. All records were subjected to painstaking analysis by ruler and compass. The subjects were four men and four women, and in each case, the duration of the sound was less than a second, the average being less than 0.5 second. In each case, with the vowel sounds, there was observed (1) a building-up period in which the oscillations rose from zero to an amplitude that showed all the components clearly; (2) a middle period in which the general amplitude remained nearly constant; and (3) a period of decay in which the components disappeared and the oscillation gradually lost its characteristic waveform.

In a typical case, the sound *oo* (as in *pool*) pronounced by one of the male speakers, period (1) was 0.05 second, period (2) 0.20 second, and period (3) 0.06 second; total duration was 0.31 second. The fundamental or cord tone was observed to be 102 hertz at the start, rising to 108 in the middle, and rising to 120 at the end. The low-frequency "characteristic," i.e., first apparent cavity response, was measured as 400 at the start, 430 in the middle, and 440 at the end. Its amplitude was greater than that of the fundamental, and it was approximately a fourth harmonic, but its amplitude variation during the cycle suggested a transient. A high-frequency component of 3,300 to 3,600 hertz was observable throughout, with amplitude variation again suggesting a transient. It was not practically possible to relate these frequencies harmonically to the fundamental.

By taking special pains with the optical system to ensure fine definition and strong illumination, and developing the films for maximum contrast, it was possible to make the records sufficiently clear to permit reproduction by the line-engraving process so that wide distribution of the results was practical. To illustrate the clarity of the result and its usefulness to investigators, Fig. 10-27 shows two short sections (from Crandall's 1925 paper), each of 0.01-second duration, at instants spaced 0.1 second apart as a single sound was being pronounced. Both lower and higher frequencies are clearly reproduced.

The detailed inspection and analysis of all the oscillographic waveform records were formidable tasks, but they furnished new insights

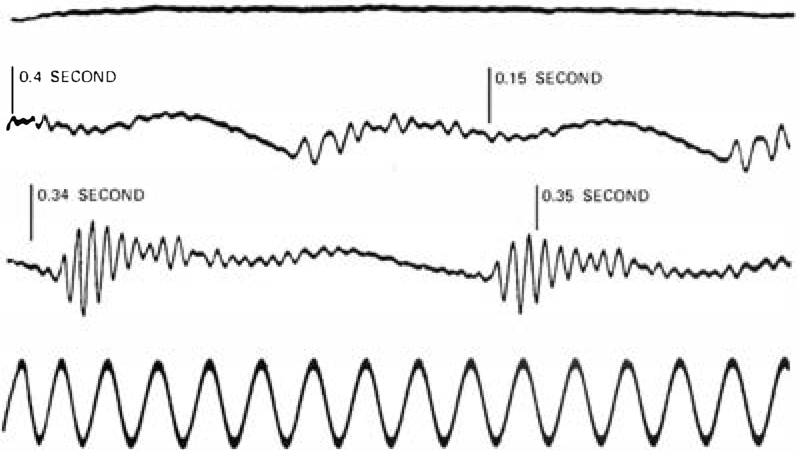


Fig. 10-27. Short sections (each 0.01 second) from the record of a single speech sound as finally reproduced for publication, including a 1,000-hertz timing wave. (Redrawn from I. B. Crandall, 1925)

into the comparative structure of the different vowel sounds, and the records themselves were basic material from which more extended studies could be made. In particular, to obtain quantitative information on spectral energy distribution, a machine method of waveform analysis devised by Sacia[70] was applied to all of the records. By the use of an endless belt, the records, of which a single one was, of course, a non-periodic function representable analytically by a Fourier integral, were passed repeatedly through the analyzer to build up periodic functions whose magnitudes were related to those of the infinitesimal components of the Fourier integral. Using averaging methods, the results could then be plotted as the sound spectra for these vowels for the four male and four female voices. To avoid undue emphasis on the low frequencies, whose energy content is large but to which the ear is less sensitive, the ordinates (proportional to acoustic pressure amplitudes) were corrected for relative ear sensitivity in accordance with Fletcher's 1923 curves, which we have already noted in Fig. 10-21.

Reproduced here as Fig. 10-28 are two such sound spectra, selected out of 13 published by Crandall and Sacia, corresponding to the 13 vowels noted on the triangle of Fig. 10-26. These represent the sound *u* (*put*), well up on the left side of the triangle, and the sound *i* (*tip*) well up on the right side. The difference in energy distribution is clearly depicted, corresponding to the single-cavity

[70] "Photomechanical Wave Analyzer Applied to Inharmonic Analysis," *J. Opt. Soc. Am. and Rev. Sci. Instr.*, October 1924.

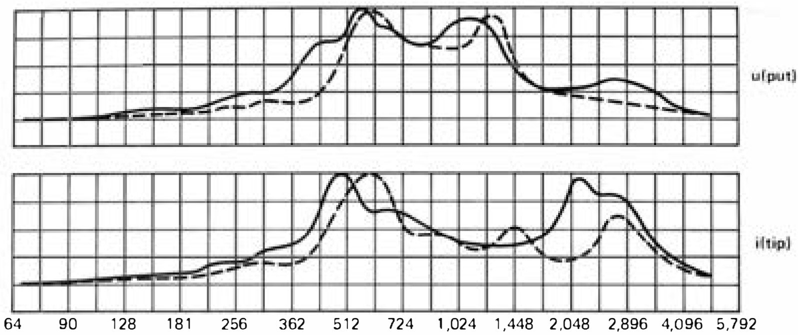


Fig. 10-28. Spectral energy distribution for two vowel sounds, one on each side of the vowel triangle, portraying the effects of single-cavity resonance (upper curves) and double-cavity resonance as determined by tongue and mouth positions. Ordinates, representing sound pressure, are arbitrary but are corrected for ear sensitivity. The full lines are for male voices, the dotted for female. (Redrawn from papers by Crandall and Sacia)

and double-cavity resonances characterizing these sounds. One of the important conclusions from study of the spectral diagrams was that the *ratio* of resonant frequencies is an important clue in the aural interpretation of vowel sounds, whether male or female. "It is only in this way," noted Crandall, "that we can account for what is a matter of universal experience in using the phonograph, namely, that moderate variations from normal speed in recording and reproducing speech leave the vowel sounds still intelligible."

Crandall and Sacia applied the same analysis to the semivowels *l*, *ng*, *n*, and *m*, and also studied the waveforms of the various consonants, voiced and unvoiced, seeking those subtle differences that constitute the niceties involved in the miracle of articulate speech. As these fine distinctions were traced, easily observable to the ear, it was found more and more that transients and minute modulations of higher frequencies during the fundamental cycle played an essential role. "For the present," wrote Crandall, "the record is the important thing, and we believe that a set of faithful records opens a new prospect in the field of speech investigation."

Other able scientists in America and abroad, too numerous to list here, were carrying out types of experiments in the 1920s that supplemented and reinforced the studies of Fletcher and Crandall and their colleagues. Not all of these depended on the new technologies associated with the telephone. Especially noteworthy were the achievements of Sir Richard Paget, who with great ingenuity devised experiments using comparatively simple equipment, supplemented by an exceptionally keen sense of hearing and pitch. For example, Paget trained himself with great patience to identify the individual overtones of speech corresponding to the primary resonances in the oral cavities. The observations he reported were referred to repeatedly by Crandall and Sacia in their own analysis of sound spectra.

In another related area, G. O. Russell,[71] director of the Phonetics Laboratories at Ohio State University, contributed greatly to the understanding of the vocal tract with his X-ray photographs in the median plane of the head, taken during speech, supplemented by cross measurements by means of "palatograms" and a "laryngo-periskop" that he had devised. Much of the theory concerning resonances had previously been based on dimensions that were conjectural.

H. K. Dunn, who joined the Fletcher–Crandall group in 1925 (the first year of Bell Telephone Laboratories) with a doctorate from the California Institute of Technology, was later to become a principal worker in vocal tract theory and experiments. Dunn's first studies[72] were in statistical analysis of spectra of speech and music. As knowledge of speech spectra accumulated, Dunn was to take special interest, beginning in the 1930s, in relating this knowledge to the physiology of the vocal tract in terms of equivalent mechanical cavities and constrictions, leading eventually (after the diversions of World War II) to an electrical analogue[73] of the tract wherein the masses and compliances, including the highly variable constrictions between cavities, could be simulated by adjustable electrical elements in accordance with transmission line theory.

Dunn's apparatus, excited by an electric wave generator comparable to the vocal cords in harmonic output, could produce the whole series of English vowels, not perfectly but distinctly better (because of the distributed constants) than was possible with the cavities represented by ordinary tuned circuits.

Thus the understanding of electrical-mechanical analogues we have mentioned earlier, supplemented by transmission line and network theory expertise, all born in a telephone laboratory, found another application more than a decade later that would be useful in checking vocal tract theory and investigating the phonetic effects of the independent movements of the articulating elements.

The findings in the mid-1920s on the nature of speech emphasized that its phonetic quality is associated not with the precise form of its pressure fluctuations but rather with its frequency content. The

[71] Russell's book, *The Vowel*, Ohio State University Press, 1928, includes a historical summary of vowel theories going back more than two centuries.

[72] "Absolute Amplitudes and Spectra of Certain Musical Instruments and Orchestras," *J. Acoust. Soc. Am.*, January 1931.

[73] "The Calculation of Vowel Resonances, and an Electrical Vocal Tract," *J. Acoust. Soc. Am.*, November 1950.

vowels and semivowels derive their color from vocal tract resonances that result in "groupings" of cord tone harmonics in the vicinity of the cavity resonances, the groupings being known as *formants*. Fricatives and plosives, on the other hand, are characterized by smudges of energy that are relatively broad in the frequency dimension. The flow of speech is a combination of these features into a dynamically varying whole. It was these findings, together with the developments of that period in electronics and circuit theory, that began to pave the way for synthetic generation of speech, a dream more than a century old but realizable only in cumbersome mechanical ways. H. W. Dudley, who joined Western Electric in 1921, became an active worker in this field. Dudley was an electrical engineering graduate of Penn State who later did postgraduate work at Columbia.

Starting with a device[74] that Bell people called the "voder" (voice demonstrator), employing a bank of ten electrical resonators excited selectively from an electrical buzz or hiss source by pressing keys, Dudley proposed (in the 1930s) to go a step further and control such a device directly and instantaneously from the voice of a speaker. The implications of this proposal were exciting, since only the control signals (which would be telegraph-like and narrow-band, corresponding to the slow maneuvering of the vocal tract muscles) would have to be transmitted over a line. Thus the vocoder (voice coder) principle had obvious possibilities for systems where security might be of overriding importance and the quality of speech a secondary matter. We have already seen that the infinite subtleties of human speech could scarcely encourage hope for a truly natural speech from any but a human source.

Of more immediate usefulness was a device called the artificial larynx, which actually used the cavities of the human vocal tract but supplied the air stream through a tube inserted in the speaker's mouth and including a buzzing reed, activated by the air flow (or turned off, for unvoiced sounds). Described[75] by R. R. Riesz, who came to Western Electric in 1925 from the physics faculty at Wisconsin, the device was later used by thousands of people who had lost their vocal cords through surgical procedures. In such persons, the windpipe terminated (Fig. 10-29) in a hole in the front of the neck, and the lungs were thus able to supply the necessary air stream by connecting the tube to this opening. In later years, refinements

[74] An early form of electrical speech synthesizer was designed by J. Q. Stewart, a Ph.D. in physics from Princeton who was with AT&TCo from 1919 to 1921 and then returned to Princeton. A somewhat similar one was demonstrated publicly by Fletcher in New York (1924).

[75] R. R. Riesz, "Restoring Speech," *Bell Laboratories Record,* September 1929.

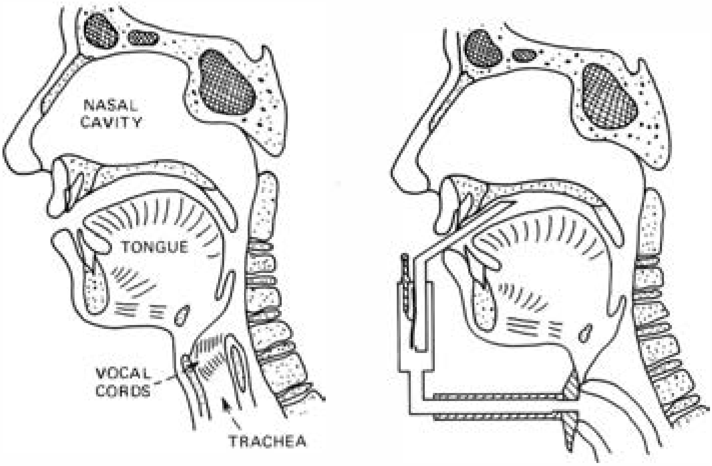


Fig. 10-29. In the artificial larynx, the normal passageway for air through the vocal cords (left) is replaced by tubing attached to a breathing hole in the neck (right). The airflow from the lungs then actuates a vibratory metal reed, injecting sound vibrations into the mouth, where they can be modulated by the same cavity resonances as in normal speech. (Redrawn from R. R. Riesz, 1929)

would become possible to improve the flexibility of the device (for example, by electronic generation and control of the "cord tone") and to increase the convenience of use.

While the artificial larynx, manufactured by Western Electric and distributed at cost, exemplified the kind of useful by-product[76] that can arise from a fundamental research program, it also deserved notice as a demonstration of the often-observed eagerness of supposedly "detached" scientists to make their findings useful when human needs can be served.

The new techniques for acoustical measurement that were evolving in the 1920s were naturally of great benefit in the developing field of architectural acoustics. Moreover, radio broadcasting and the increasing use of microphones and loudspeakers for entertainment imposed more stringent requirements on the acoustical properties of rooms and auditoriums as well as studios—requirements that must now reflect the new understandings of subjective effects,

[76] The artificial larynx is but one example of the many useful products that have resulted from Bell System research. Others from the field of speech and hearing and acoustics include the audiometer, hearing aids, high-fidelity recording and reproduction, and sound motion pictures. It is not the purpose of this volume to list the many hundreds of innovations in science and engineering that have helped create new industries (outside the Bell System) and new products. A Bell Laboratories 1971 publication, *Impact*, Library of Congress Catalog Card Number 79-157516, is a compilation of many of these.

including noise, reverberation, distortion, and frequency response. As a result, acoustical engineers in various laboratories who had been considering formation of a society to be concerned chiefly with architectural acoustics became interested in enlarging its scope to include workers in all branches of acoustics, an idea supported enthusiastically by Fletcher and his colleagues at the newly formed Bell Telephone Laboratories. Prominent among those promoting this concept were D. C. Miller of the Case School and V. O. Knudsen of the University of California, both of whom we have mentioned in this section as important contributors in acoustical theory.

This was the genesis of the Acoustical Society of America, formally founded in 1929 with Fletcher of Bell Laboratories serving as its first president. Its list of charter members included 89 from the Bell System. In the years following, the *Journal* of the Society has been one of the major channels through which Bell System scientists concerned with speech, hearing, and acoustics in general have published their contributions in these fields. Accordingly, in a subsequent volume where these contributions of the 1930s and later years can be described more fully, numerous references will be made to *J.A.S.A.* papers.

For the decade immediately following the mid and late 1920s, to which the present volume must be limited, we only indicate here the continuing scope of Bell System interest in the field pioneered by Helmholtz, Rayleigh, and the Bells. We do this by listing a few publications, selected from a large number:

L. J. Sivian, "Speech Power and Its Measurement," *Bell System Technical J.*, October 1929.

R. L. Wegel, "Viscosity in Solids," *Bell Laboratories Record*, November 1929.

W. A. MacNair, "Optimum Reverberation Time for Auditoriums," *J. Acoust. Soc. Am.*, January 1930.

R. L. Wegel, "Theory of Vibration of the Larynx," *Bell System Technical J.*, January 1930.

H. Fletcher, "Space-Time Pattern Theory of Hearing," *J. Acoust. Soc. Am.*, April 1930.

J. C. Steinberg, "Effects of Phase Distortion on Telephone Quality," *Bell System Technical J.*, July 1930.

E. H. Bedell and E. C. Wente, "Chronographic Method of Measuring Reverberation Time," *J. Acoust. Soc. Am.*, April 1930.

R. L. Wegel, "Study of Tinnitus," *Archives of Otolaryngology*, August 1931.

L. J. Sivian, "Acoustical Impedance of Small Orifices," *J. Acoust. Soc. Am.*, October 1935.

A. L. Thuras, R. T. Jenkins, and H. T. O'Neill, "Extraneous Frequencies Generated in Air Carrying Intense Sound Waves," *J. Acoust. Soc. Am.*, January 1935.

H. Fletcher, "Newer Concepts of the Pitch, Loudness, and Timbre of Musical Tones," *J. Franklin Inst.*, October 1935.

E. C. Wente, "Principles of Measurements of Room Acoustics," *Soc. Motion Picture Engineers J.*, February 1936.

W. B. Snow, "Change of Pitch with Loudness at Low Frequencies," *J. Acoust. Soc. Am.*, July 1936.

J. C. Steinberg, "Positions of Stimulation in the Cochlea by Pure Tones," *J. Acoust. Soc. Am.*, January 1937.

H. Fletcher and W. A. Munson, "Relation Between Loudness and Masking," *J. Acoust. Soc. Am.*, July 1937.

H. Fletcher, "Mechanism of Hearing as Revealed Through Experiment on the Masking Effect of Thermal Noise," *Proc. Nat. Acad. Sci. U. S.*, July 1938.

The papers listed above do not, by their titles, herald startling advances, and in fact represent for the 1930s mainly a consolidation and perfecting of the art that had blossomed with such rapidity, as we have seen, in the previous decade. More spectacular would be creations of the 1940s, including the sound spectrograph and the beginnings of digital representation of speech signals.

In these later achievements, R. K. Potter was to figure prominently as a leader and an innovator. The Bell System career of Potter, from its beginning in 1923, is a particularly interesting example of the research mind in an orderly attack on fundamental problems. With degrees from Whitman College and Columbia, he was first concerned, along with Bown and others at AT&TCo whom we have mentioned in Section II of this chapter, with problems of speech quality in broadcast and other radio systems where transmission irregularities could lead to serious impairments. Such problems were especially acute when voice signals were "scrambled" for purposes of privacy, using methods described in Section 4.3.5 of Chapter 5. Later, when Potter was Director of Transmission Research at Bell Laboratories, the observation and recording methods he had developed were to provide a basis for instant visible portrayal[77] of the energy-

[77] R. K. Potter, G. A. Kopp, and H. C. Green, *Visible Speech*, Van Nostrand, 1947.

frequency content of speech, whereby deaf subjects learning to speak could continuously observe an "on-line" record of their own voices. The detailed techniques involved in instantaneous sound spectrography, and the invaluable uses to which the sound spectrograph could be put in speech research, must be left for recording in another volume.

Early in Chapter 4 we indicated that conventional "analog" representation of speech waves would at some distant time give way, for some purposes, to the use of digital methods; i.e., the dissection of the voice wave by sampling methods for the purpose of either "time-division multiplexing" or coding for improved transmission. It was Potter who was to take the lead, in the 1940s, in exploring and promoting this radical departure, which ultimately would have profound consequences in the transmission plant of the Bell System.

It is universally appreciated that digital representation of speech and other signals, like many other advances of the last two decades (looking back from the 1970s), could scarcely have become practical had it not been for solid-state technology, manifesting itself most spectacularly in the transistor and in the microminiaturization that has made complex circuitry manageable and economical. Perhaps not so well appreciated are the many hundreds of earlier, less publicized advances in the chemistry, metallurgy, and processing of materials, and the advances in physics whereby the properties of materials of the period up to the 1930s could be understood and predicted. In the present section and in Section II we have said little about the physical materials of which communication apparatus is made and through which communication research is conducted. Thus we are led to Section IV, which deals specifically with research in these fields and records many contributions thereto by the Bell System.

IV. RESEARCH IN THE SCIENCES OF MATERIALS

American Bell's first physicist-engineer, Hammond Hayes, rightly saw as telephony's first intellectual problem a better understanding of electric wave propagation along wires. His associate Campbell was able to demonstrate, following the brilliant revelations of Heaviside, how such waves could be induced to travel with much less attrition by reducing the burden on the conductors. This was accomplished by depending on lower ohmic currents in the conductors while placing greater stress on the dielectric—which in the earliest days was mainly air, but also involved insulators mounted on poles and exposed to the weather. Thus the properties of materials were of fundamental importance from the outset: the resistivity of copper —which, thanks to the magnanimity of nature, was only a few percent higher than that of silver, the best conductor of all—and the dielectric constants and loss factors of those earliest insulating materials, glass and porcelain, and their counterparts for cables, such as rubber, cotton, and gutta percha.

Moreover, the loading coils through which Campbell brought about this transmission improvement turned out to be most successful when employing iron cores, and here the losses inherent in ferromagnetic materials (by way of hysteresis and eddy currents) at voice frequencies introduced a third problem of a strictly materials character, not faced in the telegraphic arts.

Thus telephony, because of the frequency range and the subtle frailty of its minute currents, made demands that had not surfaced before, quite apart from the intrinsic economic and structural importance of materials in an industry whose physical plant, even before the turn of the century, had become enormous by any standard.

Yet, at the time of the consolidation of Western Electric's laboratory facilities at West Street in New York City in 1907, as related in Section 2.4 of Chapter 2, the need for truly fundamental researches in the sciences of materials had not become widely appreciated. The engineers, chemists, and metallurgists responsible for selection and process control of the vast quantities of lead, copper, and iron, of paper and textiles, of rubber and asphaltic compounds used in the Bell System were therefore practitioners of an extensive accumulation of art dating from the beginnings of electrical science and benefiting from the experiences of the telegraph and power industries.

The chemical specialists in the original West Street laboratory numbered only five, under Jonathan W. Harris, a Michigan graduate in chemical engineering who had headed up similar work for Western Electric in Chicago. The group included a member from the AT&TCo Boston Laboratory, G. O. Bassett, an M.I.T. graduate and chemical engineer, whose name appears on the 1895 and 1905 rosters of that laboratory as shown on Figs. 2-7 and 2-8 in Chapter 2. One of the primary functions of these specialists was chemical analysis, whereby the degree of purity of a submitted material could be determined and the foreign substances identified. While this activity tended to identify the role of the chemists as largely that of a service group, the record[78] shows an aggressive assault on broader materials problems as instrumentation and methodology improved. Typical of these problems, in this case a problem of cost, was the search for a substitute for tin to be alloyed with lead in the fabrication of cable sheath. This search culminated in 1910, after four years of study and testing, in the adoption of a

[78] "Seeking the Better Material and the Better Method," *Western Electric News*, November 1916. (This article describes the work of the chemical branch of the Engineering Department.)

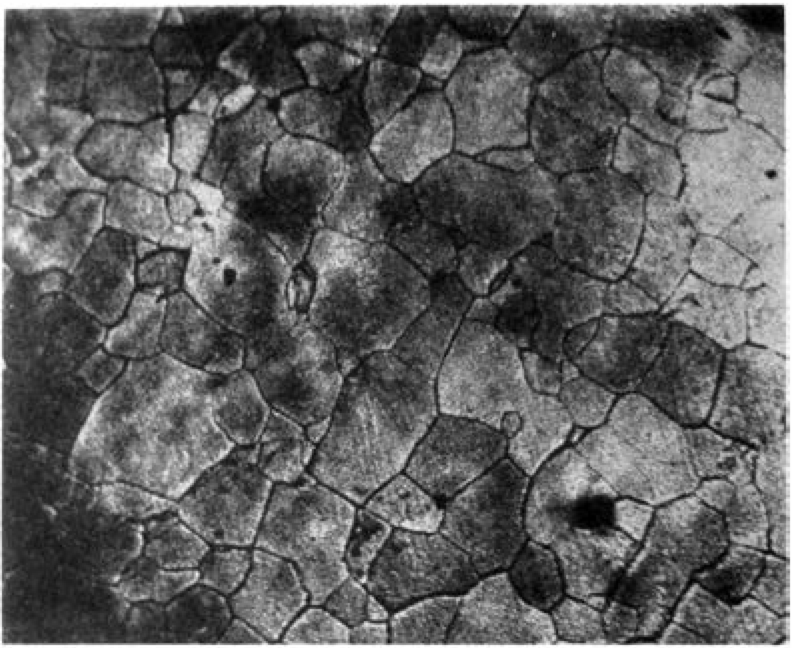
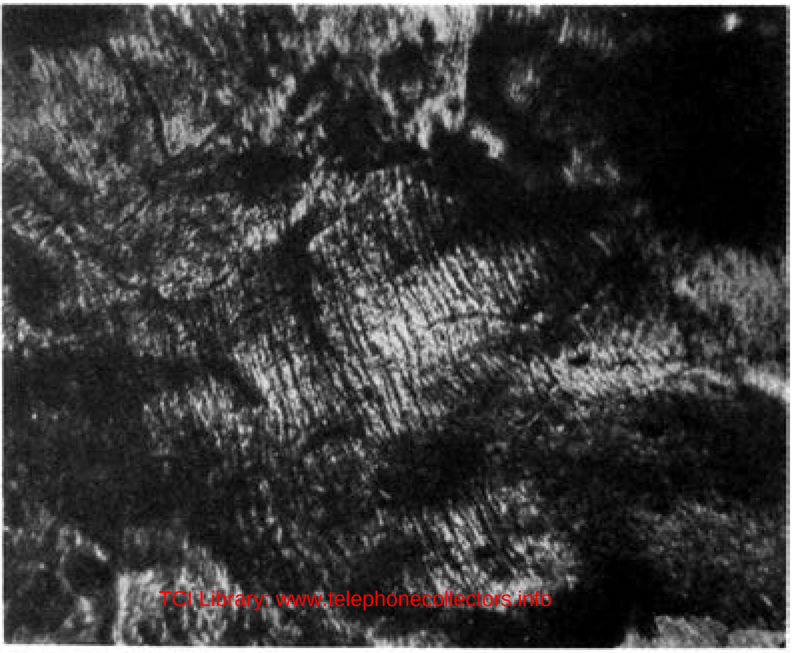


Fig. 10-30. A pair of photomicrographs taken in the Western Electric chemical laboratory. Above, a view of the normal crystal structure of lead-antimony cable sheath (Magnified 100 times). Below, the same type of sheath after severe physical strain (magnified 200 times). (*Western Electric News*, November 1916)



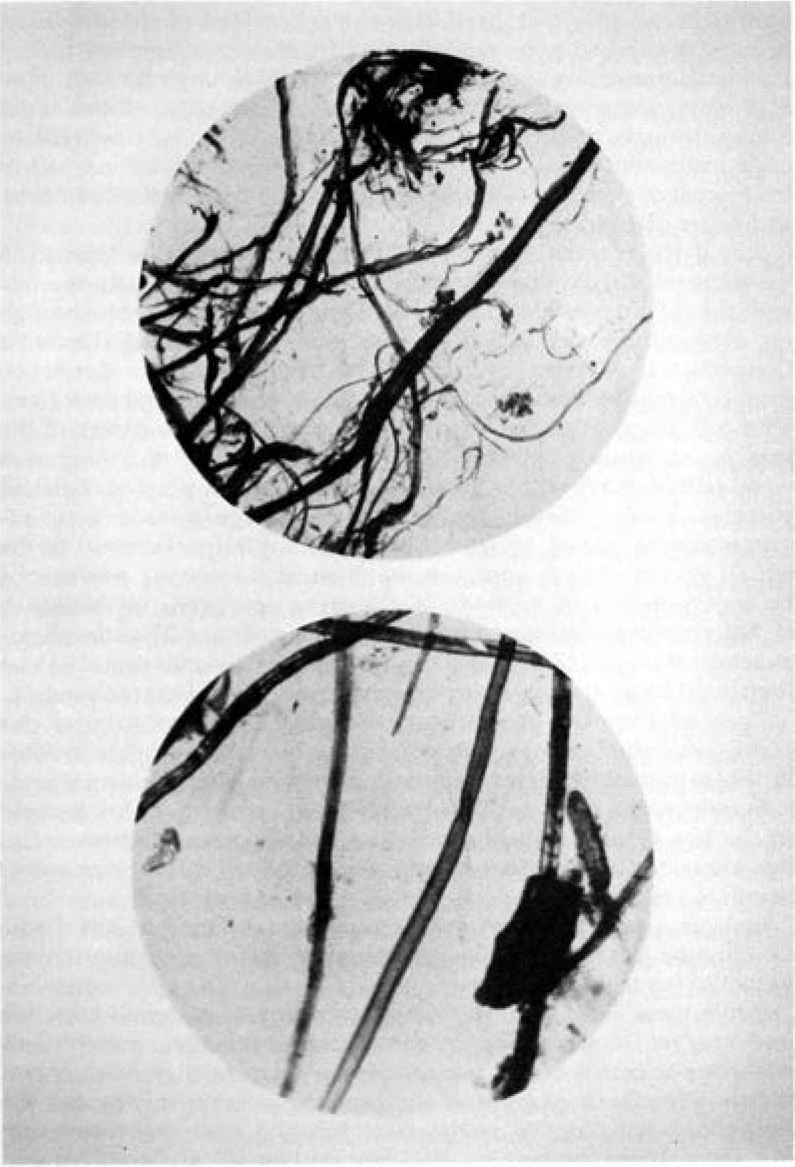


Fig. 10-31. Microscopic pictures, also taken in the Western Electric chemical laboratory, showing the difference between rope fibers from cable paper (at the top) and linen fibers from condenser paper (at the bottom), both magnified 150 times. (*Western Electric News*, November 1916)

lead-antimony alloy that saved large sums. Problems of physical strain in cable sheath had to be carefully studied, the crystalline structure of the metals portraying such deformations clearly through the technique of photomicrography (Fig. 10-30). Similarly, microscopic studies of the fibrous textures of dielectric materials (Fig. 10-31), as employed in cable insulation and in great quantities of condensers, were a part of the arsenal of analytical weapons leading to the development of newer and better dielectrics.

Fine detail at magnifications of 100 or 200 diameters, as typified in these figures, did not remotely approach what was theoretically possible with the apertures used and at the wavelengths of visible light. Through the outstanding work of F. F. Lucas, who came from the Operating Companies to Western Electric in 1910, techniques were developed for increasing the useful magnification to several thousand diameters, with full crispness of the image. Striving to reveal the nature of the granular structure of metals, to observe the survival of a fungus in wood cells, to bring within the range of vision the dispersion of colloid particles in other media, Lucas and his colleagues made their advances over a period of years by painstaking improvements in the optical system so as to approach the theoretical resolving power, and by improvements in methods of preparing specimens for brilliancy and contrast in observation, often with the help of color. Thus the microstructure of many elusive substances became better understood so that they could be used more advantageously in the manufactured product.

These achievements in microscopy brought many honors to Lucas, one of those scientists who, though lacking an advanced education, are able to discern profitable spheres of application for their gifted inquisitiveness.

In cooperation with the Zeiss staff in Jena, Lucas was to bring about in the late 1920s a further spectacular advance, the development of the ultraviolet microscope, taking advantage of the shorter wavelength to yield twice the detail as well as other advantages.

In later years, the electron microscope, developed at the Radio Corporation of America, was to increase useful magnification (in vacuo) to the tens of thousands.

While these new avenues to closer scrutiny of materials structure had not yet been opened in the decade 1910–1920, nevertheless telephone scientists of that period could see hopeful signs; for, immediately after the beginning of the century, electron theory and the Rutherford–Bohr atomic model were infusing new excitement into the speculations concerning the composition of matter. Chemical reactions, all thoroughly catalogued, were known to be characterized by a strict proportionality determined by the "valences" of the elements involved; but not until the atom was seen as a planetary system composed of a central nucleus and orbiting electrons, had it been suspected that chemical affinity or valence was a property controlled by the number and disposition of electrons in an outermost orbit or shell, these being the most distant from the nucleus and hence most easily detached, or added to, as the interaction with other atoms might require. There was therefore a hope that the making and breaking of chemical bonds might be analyzable through advances in electromagnetic theory, rather than being imputed to the action of some sort of "hooks" that could mysteriously engage and disengage.

It was one of the triumphs of chemistry to have posed the problem of chemical reactions in such challenging terms, with the phenomena catalogued in such detail as to force upon mathematical physicists the necessity of inventing a model of the atom that could provide some rational explanation for their observations. The concept of hard elastic balls, their collisions governed by statistical laws, had served usefully as a basis for the great advances in thermodynamics in the mid-nineteenth century; but it had not been chosen to explain even the simplest chemical reactions or to account for the radiation spectrum of hydrogen, the simplest of atoms. The Periodic Table of the elements, originating (1869) with the Russian chemist Mendeleev, proved (as later modified) to be the most useful method of classifying the elements and comparing their properties; but it only emphasized the challenge to physics, for it strongly hinted at an intricate inner dynamism in the atom that would prove even less tractable mathematically than the many-body problem of astronomy.

Moreover, the challenge was sharpened as chemists continued laboriously, around the turn of the century, to fill in the vacancies in the Periodic Table and thereby spotlight certain anomalies that required courageous interpretation. It was necessary, first, to reverse the accepted order of elements in three cases—putting argon before potassium, cobalt before nickel, and tellurium before iodine, thus testifying to a faith that there must be something more fundamental than "combining weights" governing the nature of chemical elements. Then, in several instances where the Periodic Law implied that there ought to be an additional element between two apparently consecutive ones, the chemists audaciously left a vacant space between the two for an element still unknown but presumed to exist; and some of these elements were soon afterward discovered, thus justifying the faith in a most impressive way.

But of the nature of the fundamental "something" leading to these anomalies, there was no inkling except for the suspicion that it involved the nucleus. As Bell chemists and others studied the elements of the Periodic Table, they noted that the "atomic" weights of the first ten, as observed in chemical combinations, were as follows:

H	He	Li	Be	B	C	N	O	F	Ne
1.008	4.00	6.94	9.02	10.83	12.00	14.01	16.00	19.0	20.20

Out of this group, six had "combining" weights that are integer multiples of one-sixteenth the combining weight of oxygen[79] (within observational error), while four (hydrogen, lithium, boron, and neon) certainly did not. Thus arose the most stimulating situation that can confront the scientist: a manifest rule restricted by undeniable exceptions.

The key to these anomalies did indeed reside in the nucleus; and it was to be found through careful spectroscopic experiments that there could be more than one "nucleus" of hydrogen, lithium, boron, and neon, and of many other elements; and that the observed "atomic" weight was merely a weighted average of these, the two or more varieties being dubbed *isotopes*. For each of the isotopes characterizing an element, the number n of orbiting electrons, each of negative charge e, was to be the same; the positive charge on the nucleus (ne) was to be the same; and the chemical properties of the atom consequently almost identical. But the *mass* of the nucleus, and hence the atomic weight, was to be greater than the mass of n protons by a factor that averaged around 2 over the range of approximately 90 elements. Thus atomic weight became dethroned, giving way to atomic *number* (the nuclear charge) as the primary attribute of an element for determining its place in the Periodic Table.

It was the additional chargeless mass in the nucleus, approximately doubling its weight (even for an isotope of hydrogen, known as deuterium), that was to confound theoreticians until the discovery of the neutron by Chadwick in England in 1932. For, even after the discovery of the electron[80] in 1897, the notion had prevailed of the atomic "nucleus" as a sphere or pudding of positive charge in which the electrons were somewhat passively embedded. Such a concept would have been difficult to reconcile with the elaborate systems of highly discrete spectrum lines observable in the radiations from even simple atoms when in excited states, such measurements being among the most precise in all experimental physics. Moreover, the "pudding" concept was shattered when Rutherford and co-workers at the University of Manchester found it necessary—after experiments performed about 1906 on the scattering of fast atomic projectiles impinging on thin metallic films—to consider the nucleus as being al-

[79] Oxygen had arbitrarily been chosen as the element whose atomic weight, assigned the integer value 16, would be the standard of comparison for all elements in the Periodic Table.

[80] The "discovery" cited was really the determination by Sir J. J. Thomson, at the Cavendish Laboratory at Cambridge, England, of the precise charge-to-mass ratio e/m of electrons. The actual charge on an individual electron was measured for the first time by Millikan at the University of Chicago twelve years later. Incidentally, the term *electron* itself had originated with the Irish physicist Stoney in 1891, but he was thinking of it as the natural (and indivisible) unit of electricity without reference to a discrete physical particle that would carry such a charge.

most incredibly small, with the electrons revolving at relatively large distances therefrom.

Such a system appeared superficially amenable to the equations of classical dynamics, with electrostatic attraction merely substituted for the gravitational attraction of Newton. But, to the discomfiture of physicists who found the planetary concept enormously appealing, the theory appeared to collide with the well-established Maxwellian principle whereby an electron undergoing acceleration inevitably radiated energy, and would therefore be expected in such an orbital system to spiral in toward the nucleus and to emit a continuous spectrum as it did so.

It was the Danish physicist Bohr who, in 1913, resolved this dilemma with two startling postulates: first, that an electron can revolve around a nucleus in certain specific orbits, designated by "quantum" numbers 1, 2, 3, . . . related to its angular momentum, without radiating, so that the atom may remain stable; and second, that an electron may spontaneously undergo transfer from one of these specified orbits to another one of lower energy (lower quantum number), this transfer being accompanied by the radiation of a "packet" of electromagnetic wave energy equal to the energy difference between the two orbits. The frequency ν of the radiation, in terms of the two orbital energies W_1 and W_2, would be given by

$$\nu = (W_1 - W_2)/h,$$

where h was a universal constant of "action" proposed by the German physicist Planck in 1900, the birthyear of quantum physics.

Planck's proposal, and an extension thereof by Einstein in 1905, had represented a new audacity that was to characterize twentieth-century science, an audacity that would challenge concepts considered to be established beyond possible doubt. Radiation was now to be considered as discontinuous, residing in minute wave packets or "photons," each possessing a "quantum" of energy $h\nu$—i.e., an amount strictly proportional to the frequency. Thus the photon was reminiscent of the "corpuscle" that had already been discarded in favor of the wave theory of light; for it could, by impact, wrench an electron out of one orbit and inject it into another of higher energy. Yet the photon could not be a corpuscle in the material sense, for its speed was the speed of light, and Einstein had already shown this speed to be unattainable by a material particle.

The new quantum physics, even in these earliest stages, explained puzzling effects observed in the spectrum of radiation from hot bodies and in the emission of electrons from surfaces exposed to light (the photoelectric effect first discovered by Hertz in 1887). The

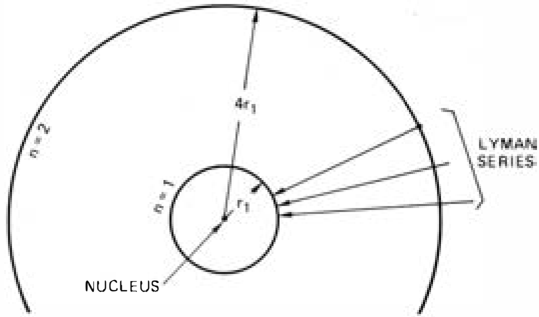


Fig. 10-32. Innermost circular orbits (quantum numbers 1 and 2) for hydrogen according to Bohr's atomic model. Electron transitions to $n = 1$ from other orbits give rise to the Lyman Series of spectrum lines (ultraviolet). Transitions to $n = 2$ from outer orbits would generate lines in the Balmer Series (visible).

Bohr atom model, devised to be compatible with Planck's theory, was phenomenally successful in explaining the radiation (as well as absorption) spectra of the two simplest atomic systems—hydrogen and ionized helium—both of these being characterized by a single orbiting electron. For example, Fig. 10-32 shows the two innermost orbits (quantum numbers 1 and 2) when they are circular; elliptical orbits are also possible. With hydrogen, transitions from $n = 2$ (or more) to $n = 1$ give rise to the Lyman series of spectrum lines, which are in the ultraviolet. Transitions from orbits of higher quantum number to orbits $n = 2$, $n = 3$, etc., give rise to other well-known spectral series in the visible and infrared. In all these cases, there is agreement between calculated and observed wavelengths.

But difficulties were to arise in applying the theory to elements of higher atomic number, even to normal helium whose atomic number was 2. There was something deeper than Bohr's orbital electron mechanics. This was *wave mechanics*, not to be conceived for another ten years, but destined to be the unifying principle of all atomic phenomena. And, as we shall see later, the laboratories of the Bell System were to be conspicuous in the resolution of a heightening paradox, the seemingly dualistic nature of electrons and waves.

The handful of theoreticians in the Boston laboratory of AT&TCo, under the guidance of Hammond Hayes, had taken more than passing notice of the revolution that was beginning around 1900. We have already noted, in Section II of this chapter, their alertness to the possibility of employing Roentgen rays for altering (at voice frequencies) the electrical properties of materials, in the hope of achieving amplification. We also noted the interest of Campbell, about the time of the move to New York (1907), in doing something with electrons, in view of their known lightness and consequent mobility. But the first forthright declaration of war by the communication industry on the redoubtable fortress of atomic science was the enlistment, in 1911, of physicist H. D. Arnold from the University of Chicago. The mission, as spelled out by Jewett, was clear and urgent: to find out how electrons could be used to amplify speech currents and to organize an attack at whatever depth might be necessary.

In Chapters 4 and 8, and in Section II of this final chapter, we have considered some of the practical problems of developing a useful and reliable electronic amplifier when the urgency was such that not all of the related phenomena could be fully explored. Here, we are concerned more with fundamental properties of materials that were intrinsic to the device chosen; most particularly with thermionic emission, for the phenomenon of electron emission from hot bodies was destined to be a subject of intensive study and experiment[81] in the Bell System for 40 years.

It was understandable that Arnold, already expert through his work under Millikan in the area of electric discharges in gases, should choose to get under way by invoking the ease and copiousness with which carriers could be produced through ionization of mercury vapor; and this approach, described in Section 4.2.2 of Chapter 4, was not unproductive, leading as it did to a usable electronic repeater. What is most notable, reflecting on the challenges that confronted Arnold, is the decisiveness with which he cut loose from this approach, at one stroke, to the relatively clean-cut notion of straight thermionic production of electrons, with their passage to an output electrode controlled in a relatively simple manner by a third (generally non-current-carrying) electrode, the grid of de Forest.

Up to this point in time, the academic world had been the primary source of knowledge concerning thermionic emission of electrons and the dynamics of electron and ion flow. The workers who laid the foundations in this field had for the most part been men who conducted their research without monetary motives and with little or no thought to any future commercial application. An industrial laboratory beginning to participate in this field was the beneficiary of their findings, of which three of the most notable within the decade had been Richardson's equation, the Wehnelt cathode, and Child's equation.

In 1901, the British physicist O. W. Richardson published a theory of emission based on thermodynamical reasoning, wherein the free electrons in random motion within the material of a heated cathode were assumed to have a velocity distribution postulated by Maxwell ("Maxwell–Boltzmann statistics"). That proportion of the electrons

[81] W. Wilson, "Reducing the Cost of Electrons," *Bell Laboratories Record*, November 1926.

whose kinetic energies exceeded a certain value characterizing the material, and known as its *work function* (expressible in electron volts), would escape therefrom and could be induced to move to another electrode. Richardson's expression for the electron current I per square centimeter of heated surface was

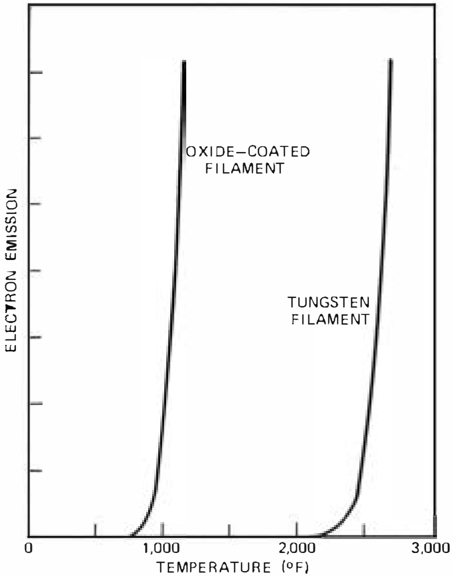
$$I = AT^{\frac{1}{2}}e^{-b/2T},$$

where the constant A depends on the units employed, T is the absolute temperature, and b is a constant depending on the emitting surface, relatively independent of temperature. Other and more accurate forms of the equation appeared later, some of them through further studies by Richardson himself and by Langmuir and Dushman of General Electric (U.S.A.). In these, other powers of the factor T appeared, as high as 2 or 3; but in general the expression and the experimental results were in close agreement, the shape of the emission versus temperature curve in the critical region being controlled primarily by the exponential term.

In 1904, Professor A. Wehnelt of Erlangen and Berlin disclosed that the electron emission from pure metals when heated was enormously increased by coating them with an oxide of one of the alkaline earth metals,[82] preferably barium or strontium or a combination thereof. His first such observations were made with calcium oxide. Thus, while the pure metals such as tungsten, platinum, tantalum, and molybdenum were good emitters because they could be operated at much higher temperatures due to their high melting points, there was the potential for large savings in heating power as well as possible longer life by taking advantage of the much smaller work function characterizing the oxide coatings. The effect of the oxide coating is brought out strikingly by the curves of Fig. 10-33, from data taken some years later when the techniques of coating Western Electric tube filaments had been standardized and more accurate and reproducible measurements could be made.

After the discovery by Wehnelt, many researchers in the universities were led to study emission from oxides under various conditions, generally under conditions of partial ionization because of the presence of one or more gases at low pressure. C. D. Child at Colgate University, studying the emission from hot calcium oxide, mathematically analyzed the situation arising when the ion density might be sufficient to place a top limit on the discharge current because of

[82] The term "alkaline earth metals" had been used by chemists and metallurgists in reference to the elements barium, calcium, and strontium (with magnesium sometimes included). These elements, in contrast with a univalent group known as "alkali metals," are bivalent and have an outermost orbital shell or subshell containing two electrons.



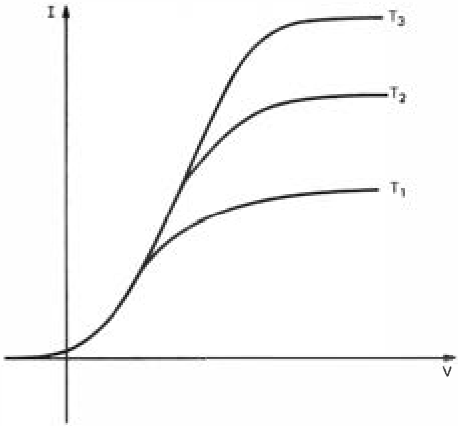


Fig. 10-33. Relative emission from oxide-coated filaments and tungsten filaments, demonstrating the greatly improved thermionic efficiency made possible through Wehnelt's discovery. (Redrawn from W. Wilson, 1926)

Fig. 10-34. The combined effect on space current of Child's 3/2-power law and Richardson's law of thermionic emission for different cathode temperatures.

mutual repulsion of the charges. Child found that this maximum current would be proportional to the 3/2 power of the accelerating potential applied to the anode.

Had it not been for this "space charge" effect, the de Forest three-element device would never have worked, for the function of de Forest's grid, even if imperfectly understood by the inventor, was purely and simply to control the electron flow by reducing or enhancing, as desired, the limiting effect of the space charge due to the cloud of electrons in the vicinity of the cathode.

Child's analysis was aimed at the case of a preponderance of positive ions over electrons; but the same approach was shown in 1913 by Langmuir of General Electric to be applicable also to a space charge of electrons only; and the 3/2-power relation is consequently often referred to as the Langmuir–Child law.

The combined effects of Child's and Richardson's equations are illustrated in Fig. 10-34, where Child's 3/2-power relation prevails up to a point where all of the emitted electrons are induced to move to the anode, whereupon (depending on the temperature of the emitting cathode) the curve necessarily flattens. In telephone repeaters, only the left-hand region would be of interest. It should be noted that even at zero anode potential there is a small flow of space current. This is because a fraction of the emitted electrons have sufficient velocity to reach the anode unaided.

Acquainted with the advances just listed, and confident of his conclusion that the three-element tube with highest possible vacuum was finally the right approach, Arnold plunged ahead in late 1912 with his program. Seeing far beyond the immediate necessity for a small number of amplifiers for a single coast-to-coast connection, Arnold's vision was of a long-distance transmission plant that would in a few years require many thousands of such devices, of uniform quality and of sufficiently long life and low operating cost to be economically feasible. At the same time, as a research physicist, he could not be satisfied with superficial answers to fundamental questions that presented themselves, even if some of these appeared to be of more scientific than commercial interest. The wisdom of this approach was to prove itself; for, while the oxide-coated cathode proved to be a complex chemical system with many baffling aspects, it was by far the most efficient of emitters, and over the next decade its superiority in terms of life, reliability, and emission efficiency were to be improved by large factors.

Working closely with Arnold on these problems was A. M. Nicolson, of British parentage and educated mainly in South America and England, recruited late in 1912 after graduate study at Harvard. With Arnold coordinating the project and devoting much of his thought to structural design and electrode geometry, Nicolson concentrated on the cathode, striving through exhaustive experimentation to bring coating methods under control so that the early repeater tubes, all necessarily handmade in the research laboratory, might last long enough, and manifest sufficient uniformity in characteristics, to serve their purpose while further knowledge could be brought to bear.

Another felicitous find, discovered by Colpitts in 1914 and temporarily teaching at the University of Toronto, was the Englishman William Wilson, an honors graduate from Cambridge with a doctorate in physics from Victoria University at Manchester. Wilson's record of fundamental researches since 1907 under Rutherford and J. J. Thomson, including a study of emission[83] from hot surfaces, presaged a distinguished career in science as he joined the growing phalanx of academically oriented men in the service of industry. We have already

[83] "Discharge of Positive Electricity from Hot Bodies," *Philosophical Magazine*, May 1911; "Quantum Theory and Emission of Electricity from Hot Bodies," *Proc. Phys. Soc.* (London), August 1913.

mentioned, in Section II, the enlistment of the South African van der Bijl in 1913, another of the protégés of Michelson and Millikan of the University of Chicago, where Jewett himself had been a graduate student. The contribution of van der Bijl[84] was broad indeed, covering not only fundamental physical phenomena such as electron emission, but those applications aspects that were the "coupling" between the device itself and the multitudinous uses (discussed in Section II) to which his "circuit" confreres Hartley, Colpitts, and Heising were eagerly applying it. Van der Bijl himself, taking advantage of the non-linear characteristics of a tube when operated over a wide amplitude range, was the originator of one of the first successful methods for generating voice-modulated radio waves.

Thus Wilson, Nicolson, and van der Bijl composed the powerful team, under Colpitts and Arnold, that would handle the most fundamental problems of electron emission and control in these earliest years of Western Electric experience with the vacuum tube. It can be seen that these few years, from 1911 to the beginning of World War I, were marked by the same new spirit of university/industry relations in the materials sciences that we have already noted (Sections II and III) in electric communication theory itself and in the field of acoustics, speech, and hearing. At one end, this new rapport owed itself to the vision of men already in the industrial ranks like Jewett, to their intellectual stature and their warm relationships in the academic community. At the other end, it was much to the credit of distinguished university scientists, typified by Millikan, that they were able to recognize the fertility of a virgin industrial field such as electric communication for cultivation even by the academic type of scientist in a sufficiently stimulating environment and under wise leadership.

We are not through with emission. The free and mobile electron, smallest and lightest of carriers, had established itself. No one could know that mastery of the science of materials and the physics of solids would one day reach a point where free electrons could be situated *by design* in the interstices of a crystal lattice. Accordingly, the literature of 1910–1930 is voluminous on thermionic work in many countries, where accurate measurements were ardently sought to verify or dispute Child's law and to determine the quantities A and b to be used in Richardson's equation.

[84] As early as 1913, just before coming to Western Electric, van der Bijl (having come from Leipzig) published in a German paper the fundamental equivalence (with respect to anode current) of grid potential variations and variations in anode potential when the latter were divided by a factor μ (to become known as the amplification factor of a tube) determined entirely by electrode geometry. Van der Bijl's book *The Thermionic Vacuum Tube and Its Applications* (McGraw–Hill 1920) was to be the classic in electronic engineering for many years.

Verification of Child's relation involved a complication, studied by Wilson at Western Electric in 1914, due to the potential drop along the cathode filament or ribbon itself, which tended to raise the apparent exponent to as much as 5/2, instead of 3/2, for low values of anode potential. Later the invention of the unipotential or indirectly heated cathode by his colleague Nicolson[85] was to circumvent this particular discrepancy. Other observations were complicated by the effects of small amounts of gas and other contaminants that could greatly affect the emission. These effects included bombardment of the cathode by positive ions created by collision of the electrons with gas molecules. Extreme measures had to be taken to eliminate occluded gases from the metallic and glass parts, which would be released at operating temperatures. Although it had already been established, for the case of pure metal emitters, that there was no basis for a "chemical" theory of emission, it was not until 1920 that Arnold[86] could lay to rest, once and for all, the suspicion in many quarters that emission from oxide-coated cathodes had to involve chemical and gas effects and could not be a purely thermionic effect in its own right.

For such fundamental studies, accurate and convenient measurement of gas pressure was essential. Until 1916, the only manometers available for measuring extreme vacua had been the Knudsen manometer and the Langmuir molecular gauge, both of which were of delicate construction and slow in action. A new manometer, using ionization of the gas by an electron discharge, was devised by O. E. Buckley[87] and proved to be free of these limitations and to handle a greater range of pressure.

A graduate of Grinnell College with a doctorate in physics in 1914 from Cornell, Buckley was to be responsible a few years later for pioneering work on high-speed telegraph cables, particularly the application of magnetic loading thereto; and, after being in charge of the Signal Corps laboratory in Paris in World War I, he was to become Assistant Director of Research under Arnold, and eventually (1940) President of Bell Telephone Laboratories.

Buckley's manometer, which was the best ionization gauge for many years, was in essence one of Arnold's triodes located within the evacuated space being studied. The grid, being biased negative, acted as a collector of positive ions. In another version, the roles of grid and anode were interchanged. In either form, the ratio of positive "collector" current to electron current was shown by Buckley to be

[85] A. McL. Nicolson; U.S. Patent No. 1,459,412; filed April 16, 1915; issued June 19, 1923.

[86] H. D. Arnold, "Phenomena in Oxide-Coated Filament Electron Tubes," *Phys. Rev.*, July 1920.

[87] O. E. Buckley, "An Ionization Manometer," *Proc. Nat. Acad. Sci. U.S.*, December 1916.

proportional to the pressure and could be so used over a pressure range from 10^{-3} to 4×10^{-6} mm of mercury. Simple in construction and exactly reproducible, the Buckley manometer also had the advantage of rapidity and ease of measurement of a varying pressure, since only the readings of a sensitive galvanometer (in the collector circuit) need be followed. These attributes were obviously of great value in a research program that of necessity involved measurements on thousands of filaments to find the best materials and processes for a demanding production program, and at the same time to unlock the secrets of Wehnelt oxide cathodes and remove some of the empiricism upon which further improvement seemed dependent.

Thus, when the advent of World War I brought a need for far greater numbers of vacuum tubes for military applications, it was recognized that fundamental scientists as well as engineers would have contributions to make; and the background of C. J. Davisson, then at the Carnegie Institute of Technology, led to his recruitment early in 1917 for what was expected to be a temporary assignment in Arnold's laboratory, from which he expected to return to his academic post. Of particular relevance[88] to the immediate objective was his doctoral work in 1911 at Princeton, where he had worked under Richardson, who had come from England to take a professorship there. The field, a long-time interest of Richardson's, was the emission of positive ions from heated salts of the alkali metals. At that time, scientists had not yet realized that emission of electrons would soon become far more important than emission of positive ions.

The atmosphere of urgency and intense activity of an industrial laboratory in wartime were in sharp contrast with the academic environment of Davisson's earlier years. The fact that he remained there for nearly 30 years would seem to be good evidence that Arnold and his associates knew how to stimulate Davisson's thoughtful kind of contributions, and that, as an individual, despite having little appetite for engineering problems, he found the same exhilaration as other pure scientists when they see their findings adapted to practical use. Davisson's office mate, less than a year later, was M. J. Kelly, who came with degrees from the Missouri School of Mines and Metallurgy and the University of Kentucky, and with his doctoral work in physics completed at Chicago. Kelly, an intense hard-driving scientist, impatient to get the job done, was of vastly different temperament from the quiet Davisson, but had the highest respect for his thoughtful and sound advice. The two maintained a warm friendship throughout

[88] Interestingly, earlier work by Davisson, likewise at Princeton and described in his 1909 maiden paper, had concerned the impingement of electron beams upon metal targets; and it was in studies of this kind that he was to achieve his greatest fame, many years later, at Bell Telephone Laboratories.

their careers. Kelly, who was to assume direction of all vacuum tube research and development in 1928, was another president-to-be of Bell Telephone Laboratories (following Buckley), and was to become celebrated as the country's leading scientist–industrialist.

The possibilities for important contributions to engineering problems by pure scientists are outstandingly illustrated by Davisson's invention of the "power emission chart," which became standard in the art. The chart is illustrated in Fig. 10-35. It consisted of a form of coordinate paper so designed that if the abscissas, which are curved, represent power supplied to the filament, and the ordinates represent emission (i.e., current to an anode), then points on the chart coordinating these two quantities will fall on a straight line if the emission satisfies Richardson's equation and if the thermal radiation from the filament is proportional to the fourth power of the absolute temperature in accordance with a well-known law due to Stefan and Boltzmann. For practical purposes in studying the thermionic efficiency of a large number of sample filaments, the advantage of using power rather than temperature as the independent variable is obvious.

The types of investigations in which Davisson was especially interested might have been considered somewhat far afield for an industrial laboratory, including, as they did, such abstruse theoretical studies as the nature of the surface forces that hold the electrons in the filamentary material, the reflection of electrons from solid surfaces, and the distribution of the velocities of the emitted electrons. These studies, however, besides shedding light on the general problems of contemporary physics, gradually led to a better understanding of the phenomena involved so that the best possible theoretical considerations could be brought to bear on the more practical problems at the other end of the scale. Of these, the preparation of filament materials and coating, and determination of the effects of impurities therein, were vital to success. Following on the work of Nicolson already cited, much of this experimentation was led by chemist James E. Harris. This Harris, younger than the Jonathan Harris who still headed the Chemical Laboratory but, like him, a University of Michigan graduate, had received his doctorate there as well, and was recruited from the Michigan chemistry faculty in 1917. The following year he was joined by E. E. Schumacher from the same faculty. This pair in collaboration made notable contributions to the improvement of cathode materials, coating, and processing, as well as the solution of other materials problems[89] of tube making. Schumacher was later to be named Chief Metallurgist in an expanded metallurgical department and eventually (1954) Metallurgical Director of Bell Telephone Laboratories.

[89] For example, "Measurements on the Gases Evolved from Glasses of Known Chemical Composition," *Ind. Eng. Chem.*, 1922, and *Bell System Technical J.*, January 1923.

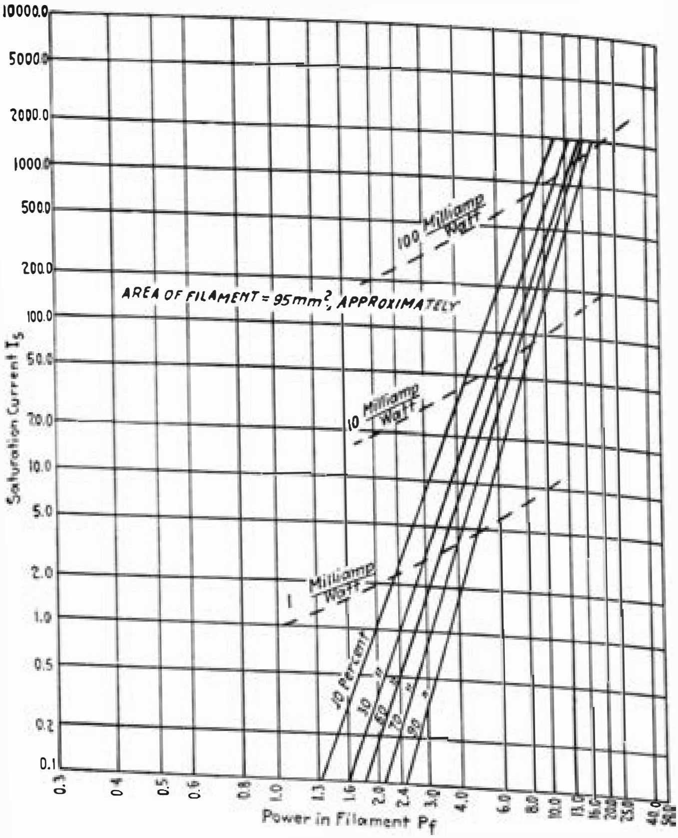


Fig. 10-35. Example of the power-emission chart devised by Davisson. Full lines represent average emission for a large number of different oxide-coated filaments, all of the same area. Dashed lines represent constant thermionic efficiency in milliamperes per watt. The constants of Richardson's equation were directly determinable from such a chart.

The more theoretical work of Davisson dovetailed so neatly with some of these materials problems that scientists as well as engineers could take much satisfaction when the emission efficiency of Bell System repeater tubes advanced by two successive factors of 2 in the 1920s, leading to great savings in power required for the repeaters, accompanied by enormous improvement in the reliability and life of the tubes.

We mentioned earlier the uncertainty in the exponent of the *A* term in Richardson's equation, the early theoretical exponent of ½ being based on Maxwell–Boltzmann statistics applied to the distribution of velocities of electrons within the emitting material. The later suggested power of 2, likewise supported by Richardson and derived from thermodynamic arguments, would have required, in order to fit the same data, a different value for the constant b, which was associated with the thermionic work function for the particular emitting material, i.e., the work of egress that an electron must do (at the expense of its kinetic energy) in order to go from the inside to the outside of the metal. Which theory was right? This was precisely the kind of question to which a scientist in a purely academic environment could happily devote several years. But with the wartime pressures removed, and with the continued encouragement of Arnold and Wilson, it was now possible for Davisson and his associate L. H. Germer (who had also come in 1917, with degrees from Cornell and Columbia, and in 1927 would receive a doctorate from Columbia) to proceed to find the answer to this question by determining the work function by another method, the "calorimetric" method. In this method, the "cooling effect" due to the latent heat of evaporation of the emitted electrons would measurably lower the resistance of the filament wire. It was an experiment that superficially would appear straightforward, but was vastly complicated—as others attempting it had found—by the "bleeding" of electrons, which caused the current to become smaller as it traversed the length of the wire. Davisson and Germer[90] found a way to circumvent this difficulty, permitting an accurate calculation of the work function that substantiated the newer value of b and thereby established A^2, rather than $A^{1/2}$, as the correct term in Richardson's equation.

In effect, Davisson and Germer had established that Maxwell–Boltzmann statistics were not applicable; that, rather, an electron velocity distribution conforming to newer statistics, to be proposed some years later by Fermi and Dirac to be compatible with the new quantum or wave mechanics, would be governing. Thus Davisson was in the position of confirming the Fermi–Dirac distribution law years before it had been stated.

These results were obtained with an emitter of pure tungsten. A few years later, the same workers[91] experimented with oxide-coated platinum wire, which had become standard[92] for electron tubes. These experiments were much more difficult because of the complex chemical system involved, but they continued to provide clues to the basic mechanisms of emission, which were to have puzzling aspects for yet another decade.

[90] C. J. Davisson and L. H. Germer, "The Thermionic Work Function of Tungsten," *Phys. Rev.*, October 1922.
[91] C. J. Davisson and L. H. Germer, "Thermionic Work Function of Oxide-Coated Platinum," *Phys. Rev.*, December 1924.
[92] The "core" material finally settled on was an alloy of platinum with 5 percent nickel.

Davisson and Germer, describing their work, paid glowing tribute to the findings of Harris on the chemical properties of the coated filament, which involved not only chemical combination of the core platinum with the oxides of the coating, but also diffusion of the coating into the core during use—either of these effects being potentially damaging to their sensitive measurements of temperature.

Thus was pointed up the powerful support given to experimentalists and theoreticians by the chemical and metallurgical specialists, whose number increased and whose advanced training and knowledge were more and more in demand as the scientific method more firmly established itself in industry. The chemist had traditionally been a living compendium of vast numbers of facts concerning a hundred thousand substances he had prepared, and their interactions. Much of what he contributed was based on experience more than theory; but, as Jewett remarked[93] in 1917, ". . . our so-called scientific methods would be of little avail were it not for the accumulation of facts laboriously dug out by empirical methods at a time when such methods were the only ones that could be applied." This was a period, however, when the revelations of atomic structure were giving the chemist and metallurgist new insights that would prepare them for a future role as molecular architects and engineers, able to design new material structures having revolutionary properties. The way to this goal was indeed to be long and hard; and even as late as 1927 Schumacher was to remark[94] that ". . . we cannot understand the behavior of metals until we know more of the properties of their structural elements, the atoms. For this the metallurgist must wait upon the physicist. If, on the other hand, the metallurgist can simplify his observations . . . he may furnish the physicist with knowledge as to the properties of atoms from which may be derived further knowledge as to their necessary structure."

This leads us to depart, for the present, from thermionics and take note of another aspect of materials research to which telephony was uniquely sensitive: the field of magnetic materials. Here too we shall observe how a far-seeing approach, though starting with empirical knowledge only, was to bear fruit not only in an immediate practical way, but also in furthering scientific understanding of a basic physical phenomenon.

The simple superficial facts of ferromagnetism as exhibited by the three elements iron, nickel, and cobalt had of course been exploited for many years in the electrical and related industries in motors,

[93] F. B. Jewett, "Industrial Research, with Some Notes Concerning its Scope in the Bell Telephone System," an address at a meeting of the American Institute of Electrical Engineers, Philadelphia, Pennsylvania, October 8, 1917.

[94] E. E. Schumacher, "The Hardening of Lead," *Bell Laboratories Record*, August 1927.

generators, transformers, choke coils, electromagnets, and permanent magnets. Improvements in magnetic properties were always to be welcomed. Telephone transformers, for instance, could have profited through reduced size and cost and better performance. But the critical requirements came with the invention of loading and the need for loading coils in the hundreds of thousands to extend the range of telephony on open wires and cables. Even more immediate, and more demanding technically, was the need for a magnetic tape-like material for continuous loading of submarine cables for both telephony and high-speed telegraphy. In submarine applications, discrete or periodic loading with coils was not practical.

The critical requirement in such applications was a high value of permeability extending down essentially to zero flux density so that the feeblest currents could be responded to. Such a requirement was unknown in the power field. Accompanying this high permeability there had to be low hysteresis; otherwise, the cyclic energy loss in the material, a magnetic "friction" proportional to frequency, would vitiate the benefits of the magnetic loading. Finally, it was necessary to minimize resistive losses due to ohmic or "eddy" currents in the material, transverse to the magnetic flux. Ordinary laminations as used at power frequencies would be relatively ineffective in the voice or high-speed telegraph range. Much better success with loading coils was obtained by winding the toroidal cores out of very fine lacquer-insulated iron wire, generally some miles in length.

Since electric communication was the only field making these stringent demands, it was in the Western Electric laboratories that powdered-iron technology was developed in the 1910–20 period for loading coils and telephone transformers, the finely ground and insulated particles being compressed into rings and offering substantially lower loss, as well as reduction in coil size.

These developments and their important applications have been detailed in Chapters 4 and 8. They did not invoke any new fundamental understandings of the phenomenon of ferromagnetism itself. Thus a mind such as Arnold's, aware that this elusive property must necessarily be rooted in the atomic structure, could not fail to speculate that deeper insights might lead to discoveries as revolutionary in impact as the Wehnelt cathode had been in thermionics.

Modern ferromagnetic theory has been made difficult by the fact that the magnetic forces between the electrons in an atom are small, by a factor of some ten thousand, compared to the electrostatic forces. Before the electron theory of the atom came into being, theorists beginning about 1900 had proposed atomic and molecular models to explain ferromagnetism in terms of the forces between atoms, assuming each atom to be a permanent magnet. Neither the distances between atoms nor the crystal structure of solid iron was known. With later knowledge of these, and of the magnetic moment of each iron atom, these early hypotheses were shown to be untenable; for even at extremely low temperatures, the energy of thermal agitation was far too great to allow formation of stable magnetic configurations. Nevertheless these early theories kept alive a healthy speculation and encouraged a variety of experiments to sharpen the theory or change its direction.

Also fueling speculation on the physical basis of ferromagnetism were photomicrographic studies of the crystal structure of metals—such as those shown earlier in this section (Fig. 10-30)—and the findings of von Laue in Germany in 1912, and the Braggs in England afterward, using X rays to reveal the internal order of the atoms in crystals. With ferromagnetic properties in the solid metal known to be sensitive to the rate of crystal formation as well as to composition, such findings were providing an accumulation of knowledge and technique giving further encouragement to a mission-oriented materials study.

Into this partial vacuum, with much at stake in the fast-expanding telephone network, moved G. W. Elmen, a native of Sweden and a University of Nebraska graduate who had joined Western Electric in 1906 and had, since his master's work at Nebraska, devoted all his time to magnetics studies. Elmen worked in the Physical Laboratory under B. Speed, who was said to be something of a genius; and a joint paper of theirs[95] describes the development, under wartime pressure, of the powdered-iron technology we have just mentioned.

Seeking a magnetic material that would combine high permeability at low flux densities with low hysteresis, Elmen in 1913 had become interested in alloys containing mostly iron and nickel. One of these, a commercial alloy of approximately 70-percent nickel and 30-percent iron used primarily for resistance wire, and in the normal hard-worked condition inferior to iron magnetically, turned out under heat treatment to be far superior to iron in permeability at low field strengths, and to have much lower hysteresis at both low and high flux densities. The heat treatment consisted of a rapid cooling during the transformation, indicating that a highly important factor in developing the magnetic properties of new alloys would be the determination of the cooling rate for best magnetic quality for each composition.

This was the lead that Elmen, an inspired and persevering experimentalist, had been looking for, and it triggered an investigation of the magnetic properties of the whole series of nickel-iron alloys, including not only permeability and hysteresis but also

[95] B. Speed and G. W. Elmen, "Magnetic Properties of Compressed Powdered Iron," *J. AIEE*, 1921.

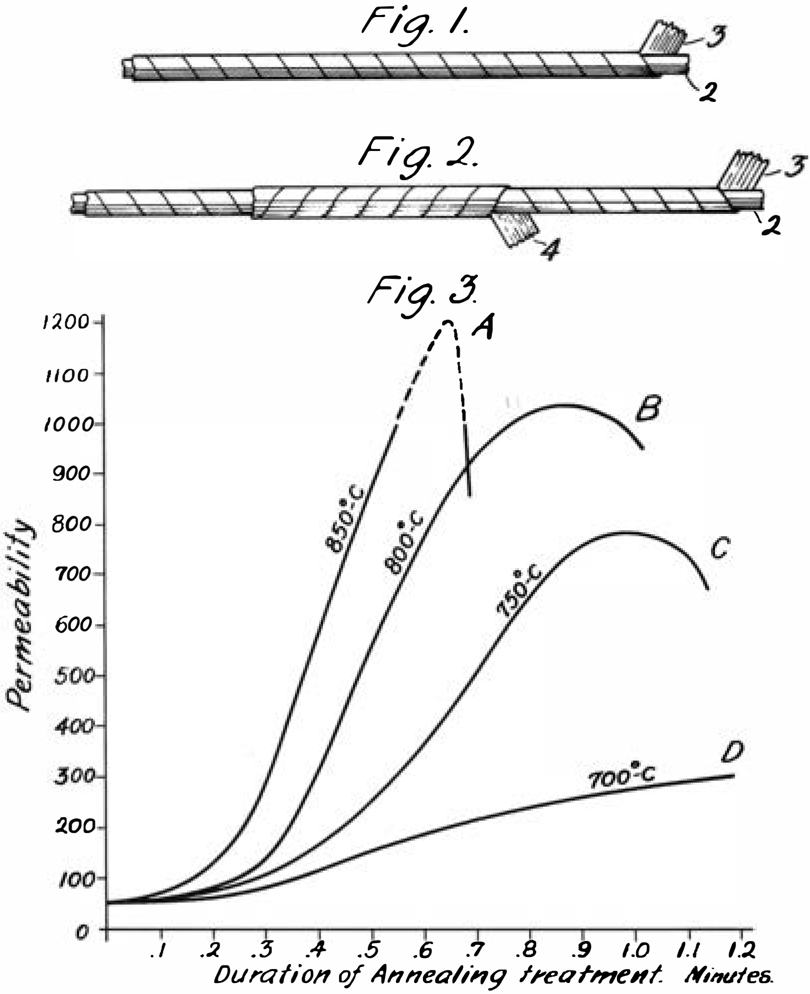


Fig. 10-36. The first three figures from G. W. Elmen's 1917 Canadian patent, depicting methods of applying nickel-iron tape to a conductor and showing the striking dependence of his high permeabilities upon heat treatment.

electrical resistance, in the interest of minimizing eddy-current losses. The vigor and thoroughness of these investigations are evidenced in Elmen's patent applications filed in 1916. The first three figures of his Canadian patent,[96] which was issued in November of 1917, are shown

[96] The opening paragraph of this patent, No. 180539 (Canada), is interesting in a personal way:

in Fig. 10-36. His Figs. 1 and 2 relate to the use of nickel-iron tape as a wrapping around the copper conductor of a submarine cable. His Fig. 3 demonstrates the critical dependence of the ultimate permeability on the temperature and duration of the annealing treatment, for the particular case of an alloy composed of 70-percent nickel and 30-percent iron. Compositions in this range, where the properties were most spectacular, were given the name "permalloy" to emphasize the high permeability displayed. Also emphasized in the patent application was the susceptibility of the material to reduction in permeability (after the heat treatment) due to mechanical strains imposed by subsequent manufacture and handling of an actual cable. It was because of this sensitivity to inevitable mechanical strains that Elmen and his collaborators chose not to overemphasize, at this stage, the importance of extreme chemical purity in ingredients, even though experiments showed that removal of the last vestiges of impurities, notably carbon, could even more strikingly enhance the magnetic properties.

Further evidence, from this early patent, of Elmen's enthusiastic pursuit of his objective is the inclusion of data concerning chromium as a constituent in these highly permeable alloys, though chromium by itself was not ferromagnetic. For that matter, nickel by itself is not nearly as magnetic as iron; it was the combination of their properties that was so remarkable. A "ternary" alloy particularly noted consisted of 55-percent nickel, 34-percent iron, and 11-percent chromium. The introduction of chromium increased the ohmic resistance of the alloy by a factor of five or more, thus substantially reducing eddy-current losses, and there were applications in which this advantage was more important than achieving the highest permeability.

The first scientific presentation on permalloy was an American Physical Society paper[97] presented in 1923. By this time, Arnold and Elmen were able to describe highly refined methods for making and testing the very large number of samples, in the form of thin tapes, covered by their exhaustive study of the nickel-iron series. Figure 1

"Be it known that I, GUSTAF W. ELMEN, originally a subject of the King of Sweden, who arrived in Philadelphia from Sweden on August 1, 1893 when sixteen years of age, declared my intention of becoming a citizen of the United States on October 5, 1900, and again on April 30, 1915; have been a constant resident of the United States since my arrival, but have not taken out my final naturalization papers; am now residing at Bogota, in the County of Bergen and State of New Jersey, and have invented certain new and useful improvements in THE LOADING OF TELEPHONE CONDUCTORS of which the following is a full, clear, concise and exact description."

Elmen's U.S. application, filed in July 1916, encountered questions of prior art, which were settled by limiting the claims to magnetizing forces below 0.2 oersted. The relevant patents were finally issued in June 1926 and refer back to the July 1916 application.

[97] H. D. Arnold and G. W. Elmen, "Permalloy, an Alloy of Remarkable Magnetic Properties," *J. Franklin Inst.*, May 1923, and *Bell System Technical J.*, July 1923.

from their paper (reproduced as our Fig. 10-37) shows strikingly the remarkable initial permeability observed in the region of about 80-percent nickel, 20-percent iron, after careful exploration to determine the best heat treatment (temperature range and rate of cooling) for that composition.

Initial permeability for each sample had to be determined through careful measurements at very weak field strengths. On the graph of permeability against magnetizing field strength, an example of which is shown in Fig. 10-38 in comparison with (but to a different scale) the best annealed iron, the straight line through the points for 0.002, 0.003, and 0.010 oersted was extended down to field strength zero to give the initial permeability, the property of most importance in communication engineering.

In the series of tests summarized by Arnold and Elmen at that time, the largest value of initial permeability found for permalloy at room temperature was 13,000, more than 30 times the value for the best soft iron. How extraordinary this was can be appreciated from the fact that this material, although it had a saturation value of magnetic induction comparable to that of iron, approached magnetic saturation in the earth's field. Thus, unusual caution had to be exercised in measuring its properties to protect the sample from the influence of stray fields.

One of the additional graphs presented by Arnold and Elmen gave the hysteresis curves shown in Fig. 10-39, where the area of the loop for permalloy (78.5-percent nickel, 21.5-percent iron) at moderate values of induction was only one-sixteenth of the area for soft iron. This great advantage prevailed at weak fields as well.

While development engineers moved fast to exploit the advantages of the new material in transmission circuits, Arnold saw in the behavior of permalloy new evidences that ferromagnetism must be associated with fundamental material structure in ways that were different from ordinary physical and chemical properties. In the scientific challenge thus presented, the extreme sensitivity of magnetic properties to strain and to heat treatment, without alteration of composition, appeared to offer new controls for use in magnetic investigations.

At the same time, L. W. McKeehan, who had come to Western Electric in 1921 from the physics faculty of the University of Minnesota, having received his doctorate there, was exploring through X-ray diffraction patterns the crystal structure (the arrangement of atoms) in iron-nickel alloys[98] and other alloys that were solid solutions of the components. The natural crystal structure of iron was body-centered cubic, i.e., the atoms were sited at the eight corners of a

[98] L. W. McKeehan, "The Crystal Structure of Iron-Nickel Alloys," *Phys. Rev.*, April 1923. Other alloys studied by McKeehan, prepared for him under chemist-metallurgist J. E. Harris, with whom we are already acquainted, included some having abnormally high resistivity despite the good conductivity of their constituents (e.g. silver and gold).

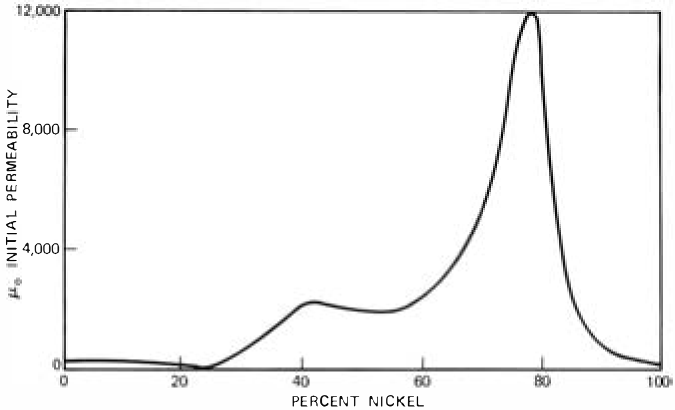


Fig. 10-37. Initial permeability in the iron-nickel series, showing remarkably high values in the vicinity of 78.5-percent nickel. (Redrawn from Arnold–Elmen, 1923)

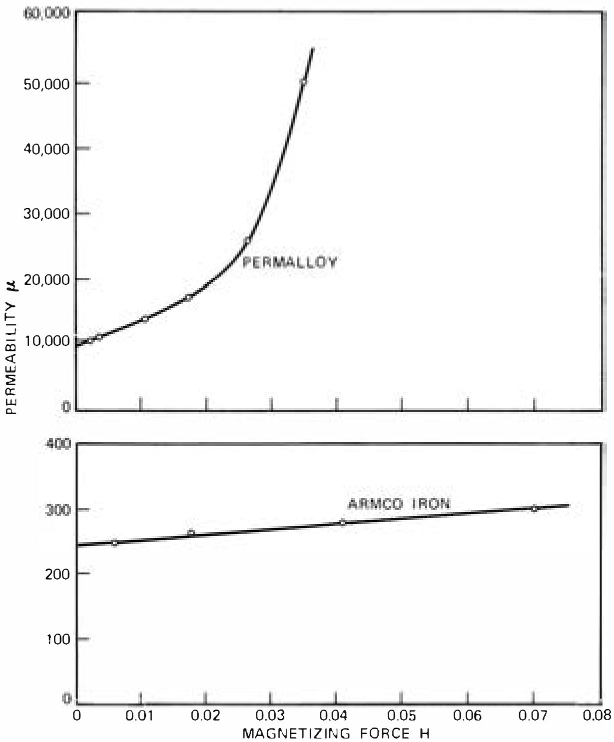


Fig. 10-38. Comparison of initial permeability (very low magnetizing forces) between 78.5 permalloy and (to a different scale) the best soft iron. (Redrawn from Arnold–Elmen, 1923)

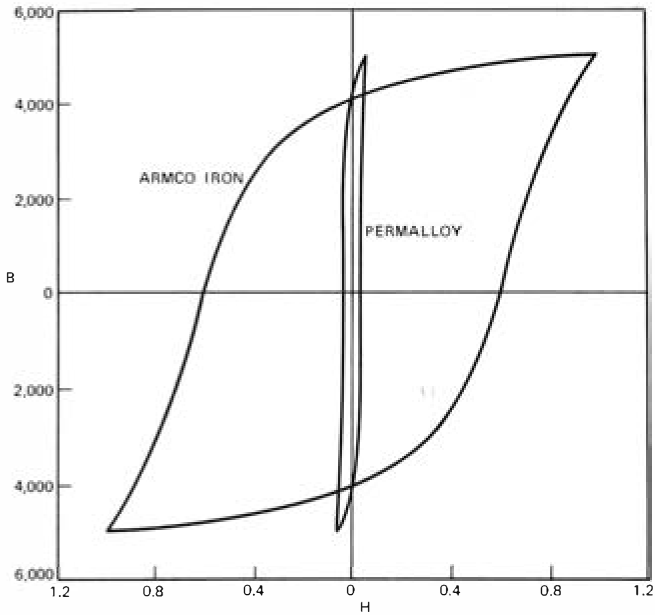


Fig. 10-39. Comparative hysteresis loops of 78.5 permalloy and iron. (Redrawn from Arnold–Elmen, 1923)

cube, with an additional atom at the center of the cube. A crystal of nickel, in contrast, was face-centered cubic, i.e., each square face of the cube had an atom at its center and there was no atom centered in the cubic body. In studying the space-lattice of the nickel-iron series, it was found that with higher than 30-percent nickel the alloy exhibited throughout, in general, the face-centered arrangement characteristic of nickel. "It is suggested," remarked McKeehan, "that the electrons responsible for the cohesion of metallic crystals are not the so-called valence electrons but those of a deeper layer relatively more firmly attached to the nucleus, the external electrons being left free in the sense required for electrical conductivity."

Additional observations from around the world deluged the scientific community. The Heusler alloys (Germany, 1922), notably an alloy of copper, manganese, and aluminum—none of them ferromagnetic by itself—displayed ferromagnetic properties related to structure when cooled at an appropriate rate producing a body-centered cubic with a face-centered "superlattice." The properties were not commercially useful, but the scientific interest was high. Concurrently, spectroscopists everywhere worked with furious energy to explore the "astronomy" of every kind of atom and provide maps of all the electronic orbitals. Iron, cobalt, and nickel, having atomic numbers 26, 27, and 28, respectively, each had a "filled" innermost or K shell of two electrons, a filled second or L shell of eight electrons, an unfilled third or M shell, and a fourth or N shell containing two "valence" electrons out of the 32 that (with an appropriate nucleus) would be its full complement. The anomaly responsible for ferromagnetism appeared to reside in the unfilled M shell (in particular, in a third "subshell" of the M shell), yet there were other elements with atomic numbers between 21 and 30 of which the same could be said, but which displayed no ferromagnetism. The suggestion that the ultimate magnetic entity might be a "spinning" electron—a notion discomforting to those who envisioned the electron as being dimensionally hardly more than a mathematical point—still did not explain why free atoms, coalescing into a solid, could leave only iron, cobalt, and nickel with their fundamental magnetic particles disposed in some manner that supported strong ferromagnetism.

We shall have more to say about magnetics when we describe the volcanic state of affairs in the middle and late 1920s, where this research story concludes. But having emphasized thermionics and ferromagnetism as particularly outstanding examples of concerted research effort that more than paid for itself, it is necessary to fill out the picture; for Jewett, Colpitts, and Arnold did many other things to foster the research spirit we are attempting to describe.

With their encouragement, and with wartime tensions alleviated, in 1919 a small group of able research men organized the Colloquium, a serious discussion forum for employees working on the frontiers of science. Membership required affiliation with one of the national scientific societies—the American Chemical Society, the American Physical Society, or the American Mathematical Society—and required being involved in or in active touch with research work. According[99] to Crandall, one of its charter members whose leadership in acoustical research we discussed in Section III, these requirements contributed much to stability and assured a certain standard in the quality of papers given, thus benefiting the scientific worker, whose very craft depends not only on knowing what has gone before, but what is actually in process at the moment.

The quality and reputation of the Colloquium's membership is evidenced by their success in having from time to time, as speakers, such distinguished guests as R. A. Millikan, A. Sommerfeld, F. W. Aston, and Ernest Rutherford.

The "fireworks" of the Colloquium, according to Crandall, was the small group of "electron mechanics"—arch-priests of the craft, who

[99] I. B. Crandall, "A History of the Colloquium," *Bell Laboratories Record*, November 1925.

set a rapid pace; while the steadiest influence and contributor of the best paper each season was Karl K. Darrow, brilliant master of exposition in mathematical physics, who had been attracted to Bell Telephone Laboratories in 1917 after graduate study at the Universities of Paris and Berlin and a doctorate from the University of Chicago.

Darrow's assignment at Western Electric was to study and correlate scientific information for his colleagues and keep them abreast of current advances by workers everywhere in fields related to their own activities. His "Some Contemporary Advances in Physics," a review of recently established knowledge in the physical sciences, appearing several times a year in *The Bell System Technical Journal* (beginning in 1923), received universal acclaim for its depth and clarity of presentation.[100]

Greatly strengthened though they had become through acquisitions we have already named, the Western Electric research ranks were not yet ready in the materials sciences, as the 1920s approached, for some of the demands that were coming as the communication industry advanced. Yet in some of these areas the best universities were well fortified with talent. On Jewett's recommendation (in 1919), Colpitts engaged the services of John Johnston, Sterling Professor of Chemistry and chairman of the department of chemistry at Yale, to assist in strengthening the organization to handle problems of a more fundamental character in organic chemistry. It could be seen, for example, that with the prospect of ocean cables of much higher capacity, due to the advent of permalloy tape loading, many new materials problems would arise; indeed, inductive loading itself would be a factor in these, since the higher impedance of a loaded cable would place new requirements on dielectric quality. Dielectric theory—in which one of the pioneers was the Dutch physical chemist P. Debye, who developed theories of rotation of polar molecules under the influence of electric fields—had been of major interest to Johnston in his researches at Yale. These researches had included studies of the chlorinated benzenes, some of which possessed large dipole moments and offered promise of much higher dielectric constants for use in capacitors.

Johnston continued as a consultant for 11 years and was largely responsible for the early scientific development of the department, especially for the selection of able men, including some of his own doctoral students, notably (in 1923) A. C. Walker and G. T. Kohman. One of Johnston's prime acquisitions (1919) was R. R. Williams, who later became Chemical Director of Bell Laboratories. A University of Chicago graduate and already a distinguished chemist in government

[100] Darrow was the author of *Introduction to Contemporary Physics*, D. Van Nostrand (1926), a recasting and unification of some of these "Contemporary Advances," and also the author of *Phenomena in Gases*, Williams and Wilkins (1932), and *The Renaissance of Physics*, MacMillan (1936).

service, Williams promptly set in motion a program for improvement in the quality and stability of cable dielectrics, having in mind the possibility of replacing gutta percha,[101] the prevailing choice at that time, which was expensive and of doubtful availability in the required quantities.

Williams was to achieve special fame for work he pursued as an avocation, seeking to identify a constituent of brown ricebran that promised relief from the scourge of beriberi, which was killing millions in the Far East. His independent search was to triumph in the mid-1930s with the discovery and synthesis of Vitamin B1, or thiamine chloride, a crystalline compound easily and cheaply added to milled rice and now a lifesaving part of the diet of Americans as well as Orientals.

The cable dielectric study, primarily aimed toward the substitution of rubber, was conducted in its practical aspects chiefly by A. R. Kemp, who had been recruited from the California Institute of Technology after graduate study there. Kemp was to become recognized as one of the country's leading rubber chemists. A joint paper[102] by Kemp and Williams, summarizing some of this work, was one of many that were appearing in this expanding field. Paralleling the engineering approach were more fundamental studies by Johnston's protégé Kohman and by H. H. Lowry, a Princeton Ph.D. recruited in 1920. Their interests[103] were in the basic mechanisms of water absorption by rubber, the removal of electrolytes to give stable dielectric properties when used under water, and those intrinsic features of the material in its purest state that governed electrical properties—the symmetry of the molecule as affecting dielectric constant, and the distribution of energy loss as between the electron, the atom, and the molecule. In these efforts their early training under distinguished faculties provided an intellectual background of excitement that projected itself over into the world of hard practical application. This could not have happened had they not considered themselves members of an already-respected scientific community.

Another problem pursued in the 1920s by Lowry, in what was one of the best pieces of chemical research[104] of the period, related to carbon as the active element in telephone transmitters. It had been the

[101] Gutta percha, a natural rubber-like product from a Malayan tree with a vast field of usefulness, had been suggested in the mid-1800s as a cable insulation by Faraday in England and promoted by Werner Siemens in Germany. It combined excellent electrical properties with plasticity when heated, facilitating continuous extrusion over the conductor.

[102] R. R. Williams and A. R. Kemp, "Submarine Insulation with Special Reference to the Use of Rubber," *J. Franklin Inst.*, 1927.

[103] H. H. Lowry and G. T. Kohman, "The Mechanism of the Absorption of Water by Rubber," *J. Phys. Chem.*, January 1927.

[104] H. H. Lowry, "The Hydrogen Content of Certain Charcoals," *J. Am. Chem. Soc.*, April 1924; and in the same issue (with S. O. Morgan) "The Rate of Oxidation of Certain Charcoals," and later papers.

practice since Edison's time to make transmitter carbon by roasting anthracite coal; but the wide range of characteristics, depending on the source and the processing technique, was quite incompatible with the tightening requirements of the growing telephone system. To characterize the different coals, Lowry developed a continuous roasting process employing a rotating furnace, and he found that with an atmosphere of hydrogen he could produce a surface of pyrolytic carbon on the granules, which resulted in both a higher microphonic sensitivity and a lower adsorption and absorption for gases—both of these being of top importance to development engineers in overcoming the practical problems of transmitter design, as described in more detail in Section II of Chapter 3. From these studies, the best sources of anthracite coal were selected, a roasting process standardized, and large supplies of coal stored for future telephone use.

Since there were many different coals roasted by different processes and being tested for different characteristics, one of the problems was a statistical analysis of the results. Here W. A. Shewhart, whose achievements in statistical quality control have been mentioned in Chapter 9, made one of his early and important contributions to telephone technology.

Other research chemists brought in, both of them in 1922, through the efforts of Johnston and Williams and destined to make notable contributions were R. M. Burns, a Ph.D. from Princeton, and S. O. Morgan, who was to receive the same degree from Princeton in 1928 after a period of postgraduate study there. The interests of Burns, who had also been a student and faculty member at the University of Colorado, were chiefly in electrochemistry and focused on corrosion prevention and the fundamental mechanisms of the corrosion process in metals. Burns was later (1945) to become Chemical Director of Bell Laboratories upon the retirement of Williams. Morgan's early work was on dielectrics as related to molecular structure; in collaboration with Lowry, Kohman, and E. J. Murphy (whom we shall mention again), Morgan later led studies in areas related to solid-state physics and chemistry, including important fundamental advances in ceramic materials in the mid-1930s, and became Burns' successor as Chemical Director in 1954.

In these ways the tapping of the great intellectual resources of the universities proceeded with accelerated speed. And the process was of mutual benefit; for, in the outstanding cases we are citing, the academic contacts were retained and often served to inspire new programs of research in the institutions.

We return to Nicolson, whom we left in 1917 working on chemical and metallurgical problems with vacuum tube cathodes. When this work was taken over by Harris and Schumacher, the fertile imagination of Nicolson shifted to the piezoelectric effect and its possible application in communications. Discovered by H. and P. Curie almost 40 years before, piezoelectricity was a natural bonanza awaiting its time; and though the effect was far greater in Rochelle salt (potassium sodium tartrate)[105] than in any other material—leading to growth of many Rochelle salt experimental crystals by the chemists for Nicolson's studies from nuclei or seed crystals—the practical choice for foreseeable communication uses was the relatively rugged and stable quartz crystal (silicon dioxide).

The characteristic phenomenon of piezoelectricity was a "coupling" between an electric and a mechanical axis whereby, with appropriate dimensions and a mechanically loss-free mounting or suspension, one could achieve the equivalent of an extremely low-loss (high "Q") electrical circuit. The internal mechanisms responsible for this action were the same as those responsible for the (much older) electro-optical effect or rotation of the plane of polarization of polarized light—namely, an asymmetry in molecular structure. In quartz, the asymmetry involved the isosceles triangle composed of a single silicon atom and two oxygen atoms, and the final symmetry in the space-lattice of the crystal, which was hexagonal, required that three such elbow-shaped SiO_2 molecules be properly disposed in each unit of structure. This disposition was explored by X-ray crystal analysis by McKeehan,[106] along with his study of metallic crystals already cited.

Nicolson's pioneering paper[107] included as Figs. 1, 2, and 3, here reproduced as Fig. 10-40, sketches of the shapes of three mineral crystals exhibiting pronounced piezoelectric effects, including an illustration of a flat "blade" excised from a crystal of quartz for measurement of these properties. Also included, surprisingly for that early date, were sketches of a phonograph transmitter (using torsional vibrations) and a crystal loudspeaker as examples of electromechanical transducers employing Rochelle salt crystals. Nicolson also pointed out that the electromechanical vibratory system could be used to introduce positive or negative reactance into an electrical circuit, depending on the frequency, thus paving the way for the crystal filters to be used in great numbers in the telephone carrier systems of the future, to which W. P. Mason was to make many notable contributions beginning in the 1930s.

[105] Morgan suspected that the remarkable properties of Rochelle salt were associated with hydroxyl (OH) groups attached to the molecule. In an interesting confirmation of this suspicion, crystals were grown from heavy water, wherein the normal hydrogen in these groups was replaced with deuterium, whose nucleus was twice as heavy. The very large change in piezoelectric properties demonstrated the correctness of the hypothesis.

[106] L. W. McKeehan, "The Crystal Structure of Quartz," *Phys. Rev.*, May 1923.

[107] A. M. Nicholson, "The Piezo-Electric Effect in the Composite Rochelle Salt Crystal," *Trans. AIEE*, October 1919.

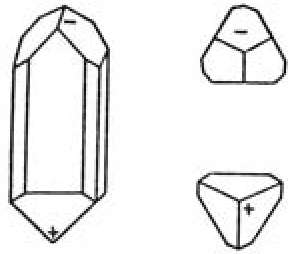


Fig 1—Tourmaline

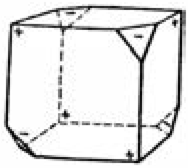


Fig 2—Boracite

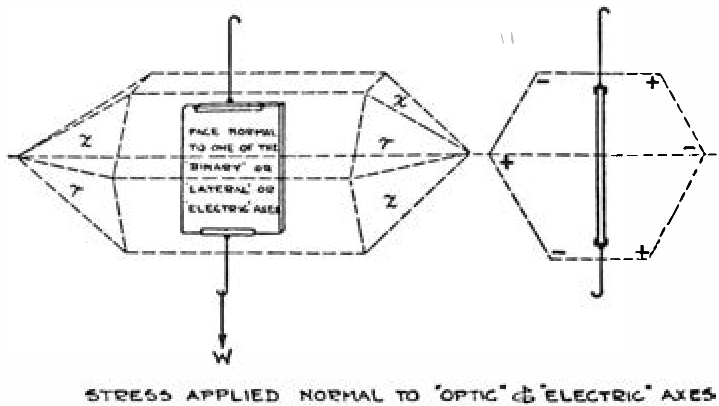


Fig. 3—Indicating Preparation of the Quartz "Piezo-Electrique"

Fig. 10-40. Three figures from Nicolson's 1919 AIEE paper, indicating shapes of three piezoelectric crystals and an appropriate "cut" from quartz for measurement of properties.

Probably the most valuable contribution coming from Nicolson's work was the crystal-controlled oscillator, on which he filed a patent application in April 1918, antedating all others in this field. The patent, not issued until 1940 because of numerous interferences,[108] stands as a monument to him in consideration of the millions of crystals that were to be used in the Bell System and elsewhere for oscillator frequency control.

The use of crystal-controlled oscillators in radio broadcasting and other services (beginning in the mid-1920s), and the use of temperature control therewith, were discussed in Section 5.3.1 of Chapter 5. A

[108] Among Americans most active in exploiting the piezoelectric effect in these earliest days were G. W. Pierce at Harvard and W. G. Cady and K. S. Van Dyke at Wesleyan University (Connecticut). Van Dyke was responsible for the derivation of the equivalent electrical circuit of a crystal (1925).

quartz-crystal-controlled clock[109] (also intended as a frequency standard for Bell System use), employing a specially cut and shaped quartz plate giving zero temperature coefficient of frequency in the vicinity of a chosen temperature, was a "research spectacular" of the late 1920s, the achievement of W. A. Marrison.

Later on, many piezoelectric crystals, eventually including quartz, were to be produced synthetically. In this field A. C. Walker, whom we have mentioned as one of Johnston's Yale protégés, was to take a leading part, in collaboration with Kohman, after spending his earlier

[109] W. A. Marrison, "The Crystal Clock," *Proc. Nat. Acad. Sci. U.S.*, July 1930.

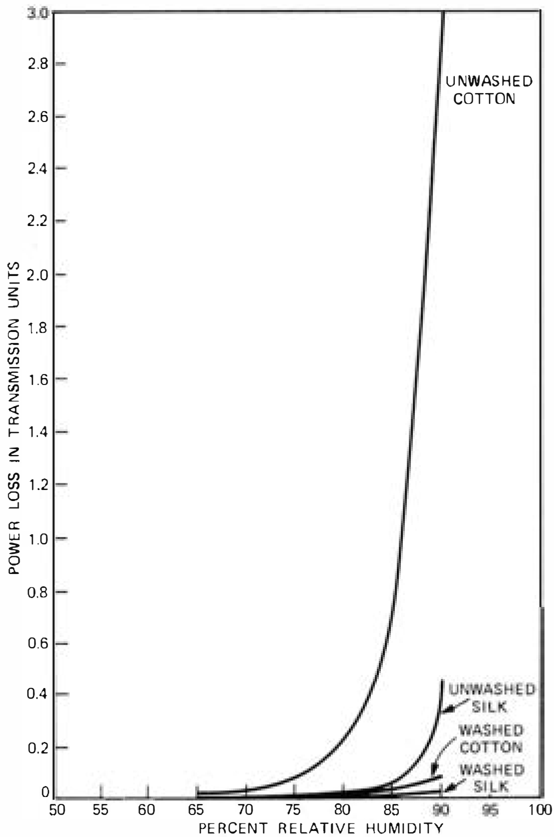


Fig. 10-41. Transmission loss at 1,000 hertz in 50 feet of central-office twisted pairs, as calculated from measurements of ac conductance and capacitance, showing benefits of "washing" to remove water-soluble salts. (Redrawn from paper by H. H. Glenn)

years on purification of textile materials for improving their insulating properties under conditions of high humidity. That earlier work of Walker, participated in by an associate E. J. Murphy who had had graduate training at McGill and Harvard, likewise had interesting theoretical[110] as well as practical aspects, an essential one being the discovery that removal of water-soluble impurities by a "washing" process could produce spectacular improvement. Figure 10-41, from a developmental paper by H. H. Glenn, depicts striking benefits to which Murphy's and Walker's fundamental studies led in connection with central-office wiring, a subject of great practical importance in the telephone plant.

Jewett and Carty continued to find their faith in industrial science vindicated as numerous other talented contributors made their appearances to extend the frontiers of knowledge. We mention a few more of these whose promise of achievement, as members of Western Electric's Engineering Department, was to be amply fulfilled in their Bell Laboratories years that followed.

J. A. Becker came from the faculties of Stanford and the California Institute of Technology in 1924. His Cornell doctoral thesis (1922) had been one of the better though unsuccessful efforts to identify the "ultimate magnetic particle." At the California Institute of Technology, as a National Research Fellow, he had studied the velocity distribution of electrons resulting from the bombardment of metallic surfaces. His long Bell System career, spanning the eras of both thermionics and solid-state electronics, included numerous contributions to the understanding of cathode behavior and of physical phenomena in semiconductors. Becker was especially interested in adsorption,[111] the process by which atoms (e.g., of thorium or caesium) are held on a surface instead of penetrating into the interior, and can greatly influence the emission of electrons. Becker's summaries[112] of the contemporary state of knowledge concerning coated filaments represented the results of long and careful experiments at a time when thermionic emission still had several useful decades to go.

P. P. Cioffi, initially involved (1917) with vacuum tubes and the study of gaseous effects therein, moved over to magnetic materials investigations as Buckley and McKeehan attempted to penetrate more deeply beyond the empirical knowledge being rapidly accumulated by Elmen. One of Cioffi's attributes was a mastery of meticulous measurement[113]—an essential in scientific investigations, as we have

[110] A. C. Walker, "Textiles as Insulators," *Bell Laboratories Record*, April 1929.

[111] J. A. Becker, "The Life History of an Adsorbed Atom," *Bell Laboratories Record*, September 1927.

[112] J. A. Becker, "Phenomena in Oxide Coated Filaments," *Phys. Rev.*, November 15, 1929 (Part I) and December 15, 1931 (Part II, co-authored by R. W. Sears.)

[113] P. P. Cioffi, "Measuring to Four Parts in a Billion," *Bell Laboratories Record*, February 1927.

noted in Sections II and III, but especially so in studies of effects as difficult to measure as *magnetostriction*, where magnetic fields produce dimensional changes of only a few parts in a hundred thousand, at the most. McKeehan had noted, in seeking an explanation for the behavior of permalloy, that iron and nickel have opposite coefficients of magnetostriction, and that, from data published by Japanese experimenters, the effects should neutralize in approximately the composition that gave permalloy its most strikingly abnormal behavior. This led McKeehan, making use of some highly precise measurements by Cioffi,[114] to hypothesize that low hysteresis and high initial permeability or ease of magnetization were directly linked, causally, with the absence of magnetostriction; and though this conjecture was later found to be incorrect, the findings (especially with respect to sensitivity to strain) were useful to the continuing accumulation of data, worldwide, that was aimed toward real understanding of these elusive phenomena.

Cioffi's ingenuity and skill in exact measurement were of increasing value as Elmen's investigations of great numbers of ternary alloys[115] responded to the varied needs for special magnetic materials in the Bell System. Cioffi was especially alert to evidences of degraded properties due to impurities, and he was to discover a few years later that extraordinary magnetic properties, superior even to those of permalloy, could be achieved[116] on an experimental basis by removing impurities through high-temperature hydrogen treatment.

Ferromagnetism was likewise the specialty of R. M. Bozorth for most of his nearly 40 years in the Bell System. Bozorth was a Reed College graduate who started with the Bell System in 1923 after holding a postdoctoral fellowship, as had Becker, at the California Institute of Technology. Through his lifelong collaboration with theorists, on the one hand, and with engineers always seeking new magnetic properties and more sophisticated ways to achieve them, Bozorth has been widely regarded as having "done more to promote the interaction between research and engineering in magnetism than any other person of our time."[117] Author of over a hundred papers, most of them combining both scientific and technological aspects of magnetism, many of them stressing the correlation of magnetic properties with crystal structure, Bozorth was to publish in 1951 a monumental treatise (Fig. 10-42) of nearly a thousand pages on ferromagnetism.[118]

[114] L. W. McKeehan and P. P. Cioffi, "Magnetostriction in Permalloy," *Phys. Rev.*, July 1926.

[115] G. W. Elmen, "Magnetic Alloys of Iron, Nickel, and Cobalt," *J. Franklin Inst.*, May 1929.

[116] P. P. Cioffi, "Hydrogenized Iron of High Permeability," *Bell Laboratories Record*, January 1932.

[117] Quoted from an editorial in *IEEE Trans. Magnetics*, December 1969.

[118] R. M. Bozorth, *Ferromagnetism*, D. Van Nostrand Co., 1951.

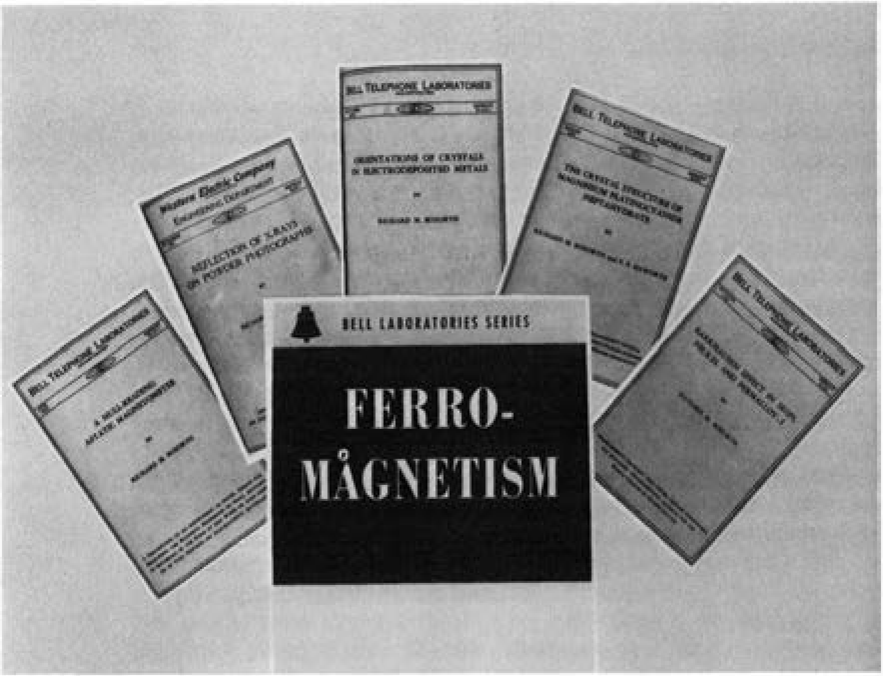


Fig. 10-42. R. M. Bozorth's classic text on ferromagnetism, published late in his career (1951), and several of his early papers of the 1920s.

About 1924, as final plans were being made to establish Bell Telephone Laboratories as a distinct corporate entity, the still unsolved mysteries of magnetism and of atomic structure, of waves and quanta, were reaching that stage where maximum frustration lived side by side with maximum determination to get through to the answers, if answers there were. In Darrow's view, what was probably required was "a modification, indeed a revolutionary extension in the art of thinking—such a revolution as took place among a few mathematicians when non-Euclidean geometry was established by the side of Euclidean, as is taking place today among the disciples of Einstein who are striving to unlearn the habitual distinctions between time and space . . ."[119]

Likewise, while spectroscopy and the Bohr atom and advances in thermionics cleared up much that had been murky concerning electrical conduction in gases, the interpretations of various experiments on conduction in *solids* (metals and non-metals) had become so conflicting as to be termed by Darrow "a hopeless entanglement of incongruous rules diversified by numberless exceptions."[120]

[119] "Contemporary Advances in Physics—VIII (Waves and Quanta)," *Bell System Technical J.*, April 1925.

[120] "Contemporary Advances in Physics—V (Electricity in Solids)," *Bell System Technical J.*, January 1925.

Yet events were conspiring that would lead, eventually, to a compatibility in the multitude of partial theories, each designed to explain a limited area in the patchwork, each trying to overlap and absorb the others; and each piece of evidence, whatever its source, was shared with Bell physical scientists by their fellow workers internationally.

A highly important experiment in 1924 by Stern and Gerlach in Germany showed that each electron had a fixed magnetic moment, and gave its magnitude. For any individual electron in a given situation, the magnetic moment could assume only certain discrete directions in space.

In 1925, Uhlenbeck and Goudsmit in Holland proposed that every electron had a mechanical angular momentum or spin equal to $h/4\pi$, here h is Planck's constant. This was likewise a vector property, and in any individual "system" (such as a suitable subgroup of electrons in a single atom), the spin for each electron must have one or the other of two allowed directions. Thus the spinning electron—at first a purely theoretical proposal that had been greeted with skepticism when suggested by A. H. Compton in 1921—was recognized as the ultimate source of all magnetism; and the "pairing" or cancellation of spins within an atom—rather, within individual shells or subshells—was seen as being implicated in the absence or presence of ferromagnetism, yet not in a simple way; for quantum mechanics or wave mechanics was to impose new rules, interpretable only through formidable mathematics, that would constitute the final criteria, and would postpone for more than another decade the arrival of any mature theory of ferromagnetism.

Quantum or wave mechanics had its origin in suggestions, in a 1924 doctoral thesis by de Broglie in France, that just as radiation must sometimes be considered as particles (photons), so might the electron be looked upon, for some purposes, as a wave or bundle of waves instead of as a discrete particle; and that the wavelength (non-relativistically) would be given by

$$\lambda = \frac{h}{mv},$$

where h is again Planck's constant, v is the velocity of the electron, and m is its mass at rest.

De Broglie's proposal, unsupported by experiment, required great boldness, as well as a philosophical insight needed in the confused situation where physics found itself.

Schrödinger in Germany adopted de Broglie's philosophical ideas and proposed a "wave mechanics" or quantum mechanics to which Heisenberg and Sommerfeld in Germany, Dirac in England, and others added ideas in a remarkable sequence of papers. Thus was initiated the "revolutionary extension in the art of thinking" that

Darrow saw as necessary, a way of thinking that would reconcile the comfortable continuum of classical wave theory with the impacts and instantaneous inter-orbital transitions that had so startlingly characterized the twentieth-century physics of Planck and Bohr.

No longer would it be meaningful to try to specify, simultaneously, the exact position and the velocity of an electron, or to describe an orbit precisely in astronomical terms. Atomic distances and spacings and the movement of charges therein had to be viewed in probabilistic terms, conforming to Schrödinger's "wave equation."

Those who could handle the abstruse mathematics of wave mechanics found that it gave answers that agreed with spectroscopy in cases where Bohr's theory completely failed, i.e., when there was more than a single orbiting electron, as (to take the simplest cases) in normal helium, or the common molecular hydrogen, H_2. More than this, there evolved from wave mechanics the concept of *exchange interaction*, which accounted for the strong forces (electrostatic, they turned out to be) involved in inter-atomic bonding, the formation of molecules. For example, wave mechanics could account quantitatively for the "covalent" type of bond, wherein atoms would begin to "share" electrons, as did two atoms of univalent hydrogen when they moved closer together, and thus developed an affinity that created the molecule; and the "ionic" type of bonding, wherein an atom such as hydrogen or sodium, having a single electron unaccompanied in its shell, would "donate" this electron to an unfilled shell such as the M shell of an atom of chlorine, thus establishing the electrostatic potential difference that would bind them together.

Thus were evolved the foundations of molecular theory, the much-needed basis on which the chemists of the following decades would establish an architecture for revolutionary new substances. Wave mechanics would also, ultimately, account for the elusive ferromagnetic properties of iron, cobalt, and nickel, for which the theories of paired and unpaired spins had been unable to give the complete answer; and would provide a quantum mechanical basis for seeking an understanding of conduction in solids, the great goal of solid-state physics in the decades that followed.

These remarkable advances initiated in the mid-1920s generated high excitement at the West Street laboratory in New York City, just at the time when scientists of the Arnold–Davisson–Darrow type were feeling gratified at the new organizational status of Bell Telephone Laboratories. Still more remarkably, experiments then in progress by Davisson and Germer were to provide the first clear-cut demonstration that electrons in motion did indeed display the wavelike qualities that de Broglie had audaciously suggested.

Beginning in 1920, in addition to his interests (with Germer) in thermionic emission as already described, Davisson was concerned with exploring the "scattering" characteristics of metals for beams of positive ions, and also of electrons, impinging upon them. The desired results could have yielded information on the outer or "orbital" regions of the atom, as distinguished from the regions near the nucleus that had been explored many years before by Rutherford with his massive and high-speed alpha particles. With a younger associate, C. H. Kunsman, a Ph.D. from Berkeley, Davisson was observing[121] the scattering of "secondary" electrons[122] produced by bombardment of polycrystalline metals by an electron beam, when it was noticed that, instead of being slower, some of the scattered electrons had the same speed as the primaries, and hence, Davisson suspected, were primary electrons that had, through some mechanism, been reflected without loss of energy.

These experiments were pursued vigorously by Davisson and Kunsman for several years, with Germer then becoming involved (Fig. 10-43). Because of a directional peculiarity in some of the results which depended on the speed of the primaries (as determined by the accelerating voltage), a young German physicist, W. Elsasser, suggested in 1925 a relationship between the electrons and "de Broglie waves." It appears that Davisson did not consider that anything in the polycrystalline scattering patterns could be taken as corroboration of such a hypothesis; but in that same year, when he and Germer were bombarding a target that included (by accident, in the first instance) a number of small single crystals of nickel instead of wholly polycrystalline material, they observed a startling difference in the angular distribution of the scattered electrons *after* a laboratory accident that had led to the partial crystallization. The new scattering curves (Fig. 10-44) exhibited much more "structure" than the originals, suggesting that an even more marked dependence on crystal direction might be observable with a single crystal. It was this conclusion that set in motion a new program to study this dependence carefully; and though it was expected that the strongest electron beams might issue from the crystal along its "transparent" directions—the directions in which the atoms in the lattice would present the least obstruction—and that the

[121] C. Davisson and C. H. Kunsman, "The Scattering of Electrons by Metals," *Phys. Rev.*, September 1923.

[122] Secondary emission (particularly when the number of secondaries exceeded the number of impinging primaries) was a useful, but sometimes exasperating, phenomenon in vacuum tube applications. The degree to which secondaries from a molybdenum grid, for example, in a very high-power (100 kW) amplifying tube, can affect circuit stability was later illustrated in a note by W. H. Doherty, *Proc. IRE*, February 1934.

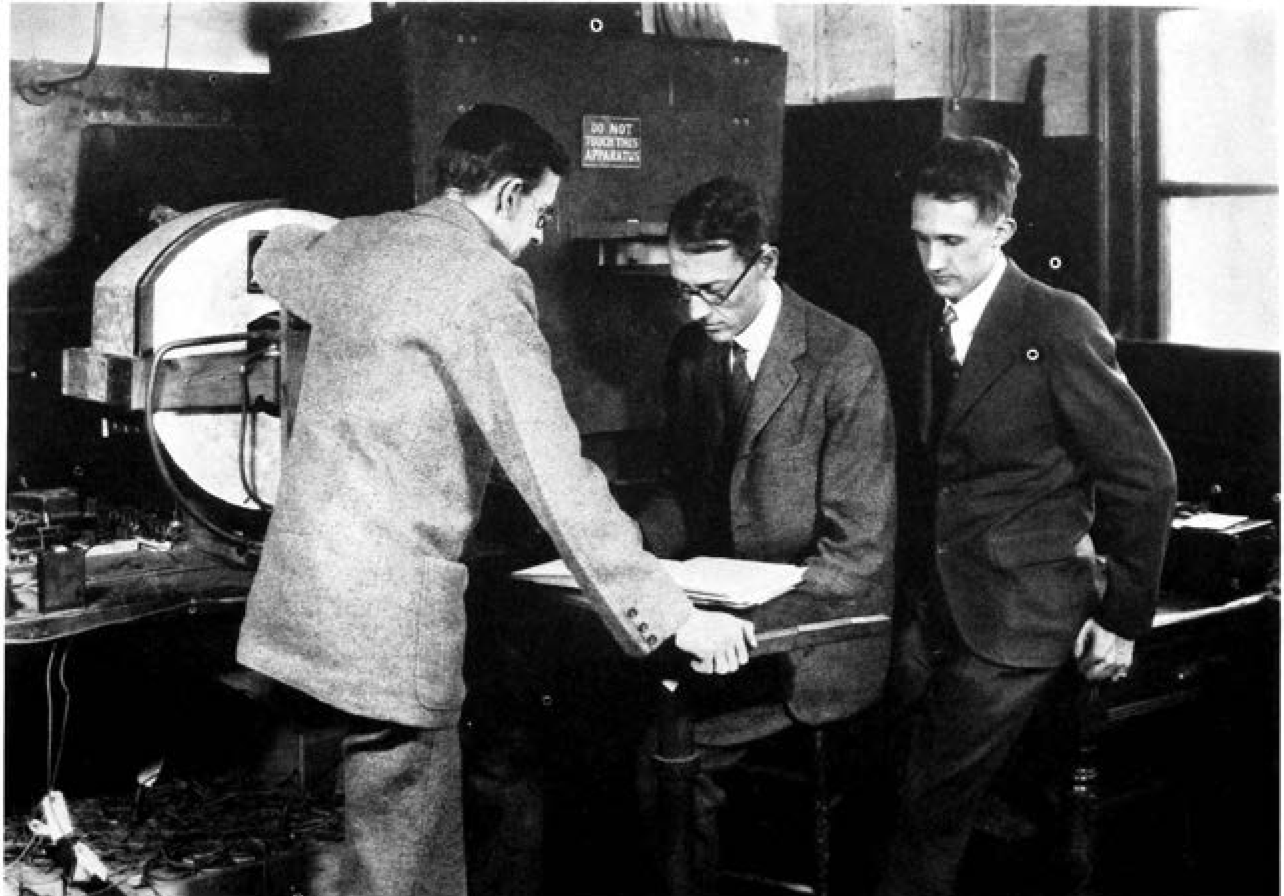


Fig. 10-43. C. J. Davisson (at left) and L. H. Germer (at center) with their apparatus for studying the scattering of bombarding electrons from nickel crystals. Looking on (at right) is their colleague C. J. Calbick.

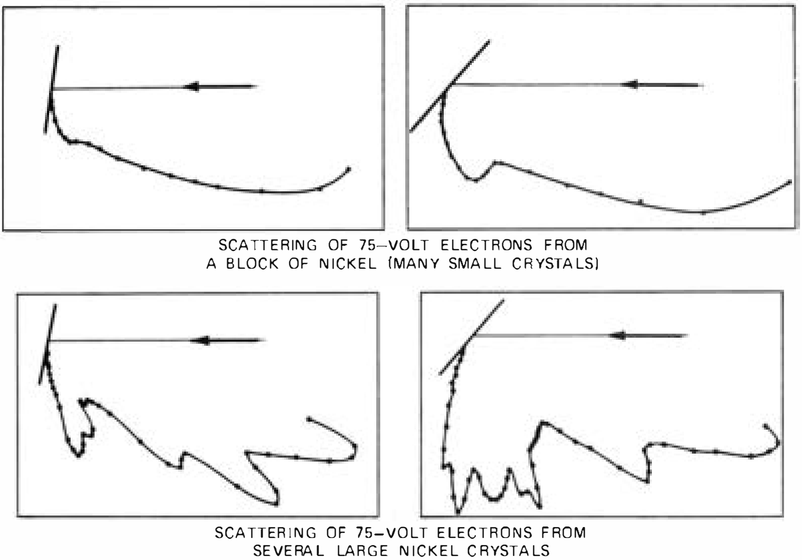


Fig. 10-44. Electron-scattering curves from nickel (1925). The upper curves are for two different angles of incidence of the 75-volt electron beam upon a polycrystalline surface. The lower curves show the change resulting when the surface included approximately ten "single" crystals.

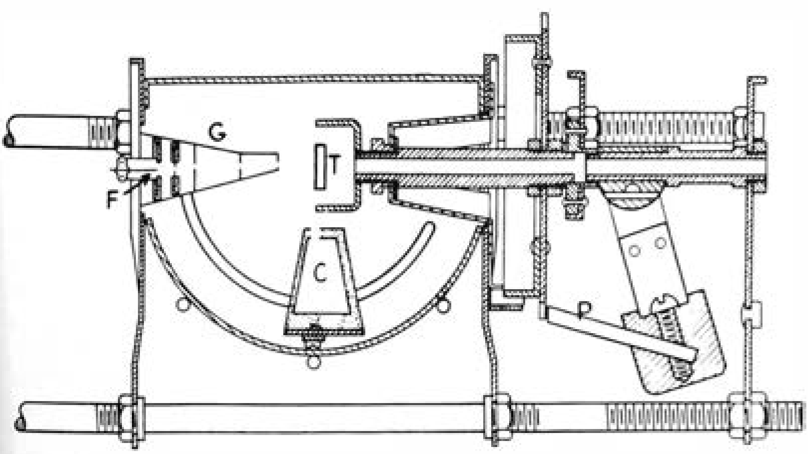


Fig. 10-45. Cross-sectional view of experimental apparatus (the highly evacuated glass bulb not shown). Essential elements are the tungsten-ribbon-filament emitter F, furnishing electrons to the "gun" G, which could project the electrons at any desired speed; the crystal target T; and the collector C, which could be adjusted to collect only those electrons having approximately the same velocity as those projected from the gun. The collector could be conveniently moved along two orthogonal axes for complete exploration of the beam pattern. (From Davisson's and Germer's December 1927 *Physical Review* paper)

scattering pattern would be unrelated to electron speed, nevertheless the experimental setup (Fig. 10-45) included facilities for varying the accelerating potential over a wide range. This was a wise provision, for though strong beams were indeed found, this was only when the speed of bombardment lay near one or another of a series of critical values, and then in directions quite unrelated to crystal transparency.

The most striking feature of these beams was a one-to-one correspondence that the strongest ones bore to the X-ray beams ("von Laue beams") that would have been found issuing from the same crystal if the incident beam had been a beam of monochromatic X rays of wavelength appropriately related to the atom spacings in the crystal. Moreover, the patterns displayed a selectivity similar to the wavelength selectivity for the X-ray case, when the accelerating potential for the electrons[123] (and hence presumably the "de Broglie wavelength") was varied. Figure 10-46, from later data obtained in a modified experiment involving "reflection" rather than scattering, brings out this wavelength dependence most strikingly.

It will be recalled that, around 1906, Rutherford had found the nucleus to be mainly surrounded by empty space, the individual orbiting electrons having dimensions of the order of 10^{-13} cm, about 1/100,000th the size of the atom as a whole. Hence there was no basis for imagining that the electrons from Davisson's "gun" could be deflected in appreciable numbers from their course, heading straight into the material, unless they possessed some attribute that would excite a resonance in the much larger mesh constituting the crystal structure; and the relationships uncovered by Davisson and Germer clinched the case. Electrons (and presumably other forms of matter as well) *had* to be looked upon also as waves.

Although the supreme recognition of Davisson's achievement, the Nobel prize in physics, was not awarded until 1937, the scientific importance of his discovery was acclaimed at once. The two papers involved, both co-authored by Germer, were "The Scattering of Electrons by a Single Crystal of Nickel," a preliminary announcement in "Letters to the Editor," *Nature*, April 16, 1927, and "Diffraction of Electrons by a Crystal of Nickel," *Physical Review*, December 1927. These were clearly the first publications offering experimental proof that individual electrons, unassociated with an atom, displayed wave characteristics in flight through space.

Very soon afterward, at the University of Aberdeen, Scotland, G. P. Thomson independently and in a somewhat different experiment

[123] A body of charge e and mass m accelerated (from rest) by a potential of V volts acquires (non-relativistically) a velocity v equal to $(2Ve/m)^{\frac{1}{2}}$, i.e., proportional to the square root of V; hence, according to de Broglie's relationship, the "electron wavelength" h/mv would be inversely proportional to $V^{\frac{1}{2}}$.

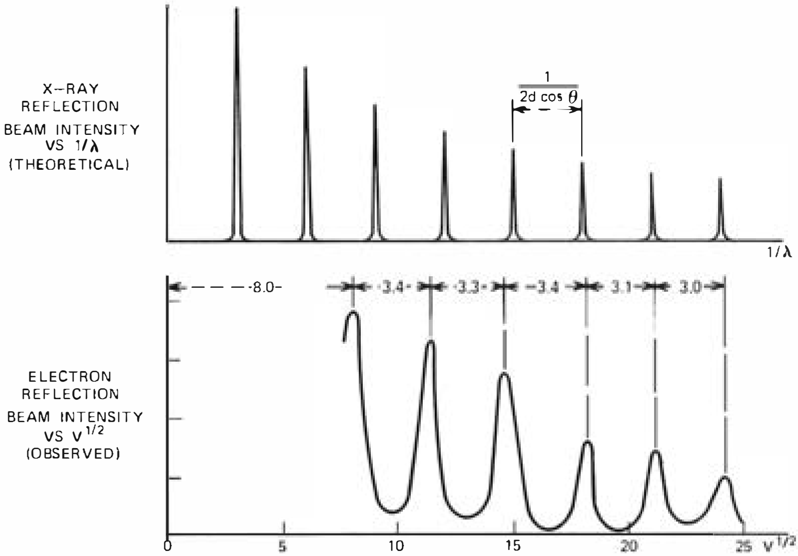


Fig. 10-46. Comparison of the (calculated) selectivity of X-ray reflection and the observed selectivity of electron reflection based on the de Broglie relationship (electron "wavelength" proportional to the square root of the accelerating potential). (Redrawn from L. H. Germer, "Optical Experiments with Electrons," *J. Chem. Education*, September 1928, Fig. 3)

arrived at a similar result, and consequently shared the Nobel award with Davisson.

In his accounts of these experiments Davisson characteristically acknowledged the assistance of his immediate co-workers, the contributions of the chemists in preparing crystals, of the mechanics in refinements of construction of his apparatus; also the valuable discussions with McKeehan, Darrow, and Bozorth; and finally, the encouragement given by Arnold and Wilson to pursue an investigation that was eminently "scientific."

Breakthroughs like that of Davisson and Germer were of great value in establishing Bell Laboratories as an institution where scientists could pursue their interests without domination by business or manufacturing pressures. Thus men like Davisson, McKeehan and Bozorth, and Johnston and Williams, represented an environment which, only a decade or two earlier, had not been known outside the universities, and were instrumental in attracting younger men of similar inclinations, typified (in 1929) by W. H. Brattain, the first of the trio (Brattain, Bardeen, and Shockley) whose work in the quantum physics of solids (culminating in the invention of the transistor in 1947) would bring to the Bell System family (1956) its second Nobel prize.

It would take far too many pages to cover even sketchily the many other areas of materials research that were being profitably pursued in the mid-1920s, or to pay deserved tribute to the contributors. Some of the scientific papers cited in our footnotes include bibliographies that list many other papers and the names of their authors. Here we have only hoped to show, through a few examples, the fruitfulness of materials science as a field for investigation in an industry based on high technology. With the electrical and acoustical arts continuously advancing in sophistication, new properties in materials always found immediate engineering use and, in turn, generated new requirements for quality in composition that would further enhance performance or contribute to long life and serviceability. An appreciation was developing during this period for such qualities as unprecedented chemical purity and crystalline perfection. With deepened understandings of chemical properties on an atomic and molecular level, there evolved a new day-to-day collaboration between physicists and chemists without which the achievements of the next two decades in the chemistry of synthetic polymers, in ceramics technology, and in the physics of semiconductors could scarcely have been initiated. These were fields to which Buckley and Kelly, successors to Jewett as presidents of Bell Laboratories, were to give strong encouragement and support, based on their own years under the leadership of Jewett, Colpitts, and Arnold.

It is therefore appropriate that this section, as well as the chapter, should reach its conclusion on a note of recognition of what the Spirit of Research is. It is, as we know it, a spirit generated in the deep recesses of the mind, seeking understanding of natural truths; yet, in Arnold's words, it is "practical as well as theoretical; trending always toward worthwhile relationships." This was the spirit that had motivated Hammond Hayes when the early sponsors of the telephone industry in Boston engaged him in 1885; and the torch has been carried by his successors, always with unstinting management support, to the great benefit of the business and the public, as well as the world of science.

Postscript: After 50 Years

Even a brief review of early telephony would suggest that the year 1925, the nominal terminus for this volume, represented a satisfying anniversary. It was the conclusion of the first half-century of a new form of human communication, ranking with the great developments of all time. Too, the mid-twenties can be looked upon as a time of transition in the evolution and application of technology. Technological innovation had formed the indispensable core for telephony's growth up to 1925, but was even more significant to the future because so much of it was fundamental: the way was being prepared for more powerful systems yet to come, which would be essential to the enormous expansion felt to be lying ahead. Perhaps more significantly, the application of scientific methods to solving the "system" problems of telephony set a pattern which influenced industrial research and development by demonstrating the power of these methods and developing techniques of management which encouraged their use.

We see the time around 1925 also as marking a changeover to a new generation in engineering leadership. The earliest fundamental contributions after the telephone itself—the balanced pair, the phantom circuit, transposition, the loaded line—had been the product of young minds; but while still active and dedicated, three decades later, to the technology of the new profession, Campbell and Carty and Colpitts were exerting their major influence by forming broad principles and by inspiring their successors to a dynamic appreciation of the opportunities and challenges of nationwide and worldwide telephony.

We proceed with this summary as we look back at those 50 years with personal recollections of the mature judgments and guidance which these early masters of the art, and their leader Jewett, were exercising as the first half-century of telephony drew to its close.

I. PHYSICAL GROWTH

By 1925, the telephone system in the United States had grown to a network covering the length and breadth of the country. It connected with a sizable network in Canada and modest but developing systems in other countries such as Cuba and Mexico, with the first step toward a worldwide system only two years away—the opening of long-wave radio service to Great Britain.

In the United States alone there were roughly 17 million telephone stations, of which 71 percent were owned by Bell companies and the remainder by independent companies, most of which connected with the Bell network (see Fig. P-1). Business probably represented the

Telephone Stations by end of 1925	
Total Stations :	
Bell System	11.9 million = 71%
Connecting	4.8 million = 28%
Non–Connecting	0.2 million = 1%
Total	16.9 million = 100%
Total per 100 population	14.5
Households with telephone service	39%
Bell System Stations :	
Residence	61%
Business	39%
Main	73%
PBX	19%
Extension	8%
Panel dial	7%
Step–by–step dial	6%
Manual – common battery	80%
Manual – local battery (magneto)	7%
Residence main stations (1929)*	
1–party	36%
2–party	34%
4–party	26%
More than 4–party	4%

*Data not available prior to 1929.

Telecommunications in the United States

	1909	1925
Number of telephone conversations in millions	12,617	22,400
Number of telegrams in millions	98	215
Total	12,715	22,615
Percent telephone	99.28	99.04

Fig. P-1. United States telephone statistics.

greatest density of use (reliable statistics are not available), but the telephone was rapidly becoming a part of the country's social structure with about 60 percent of the stations being installed in residences and affording service to about 40 percent of the country's households. However, close to two-thirds of the residential stations were on party lines and only 13 percent of the total number of stations were of the dial type. Comparison with other forms of communication may not be too significant but it is interesting at least to note that as early as 1909 over 99 percent of the United States message telecommunication was carried out by telephone.

It is difficult to appraise growth on an absolute basis, but the significant accomplishments of 50 years will be more apparent by comparison with growth elsewhere. As shown in Fig. P-2, the 17 million telephones in this country represented about 15 per 100 population, exceeding the development in any other country of the world. It was about five times that of Great Britain, and ten times the world average.

On any basis, the growth of telephony in the United States during the first 50 years can be viewed with much satisfaction, particularly considering the vast distances covered by the network (compared to other countries) and the inherent technical and economic problems associated therewith. Growth in the first years was almost explosive, as shown in Fig. P-3. Commercial, switched service was not offered until 1878, but only four years later roughly 100,000 Bell stations had been installed in some 1,000 exchanges. Nearly 140,000 stations were

	Telephones Per 100 Population	Millions of Telephones	Percent of World Telephones	Telephone Calls per Capita
United States	15	16.9	61	196
Canada	12	1.1	4.1	NA*
Denmark	9	0.3	1.1	135
New Zealand	9	0.1	0.5	NA
Sweden	7	0.4	1.6	106
Norway	6	0.2	0.6	107
Australia	6	0.4	1.3	50
Switzerland	5	0.2	0.7	39
Germany	4	2.6	9.3	33
Great Britain	3	1.4	5	25
Total World	1.5	27.8	100	NA

*Data not available.

Fig. P-2. Telephone statistics for ten most highly developed countries at end of 1925.

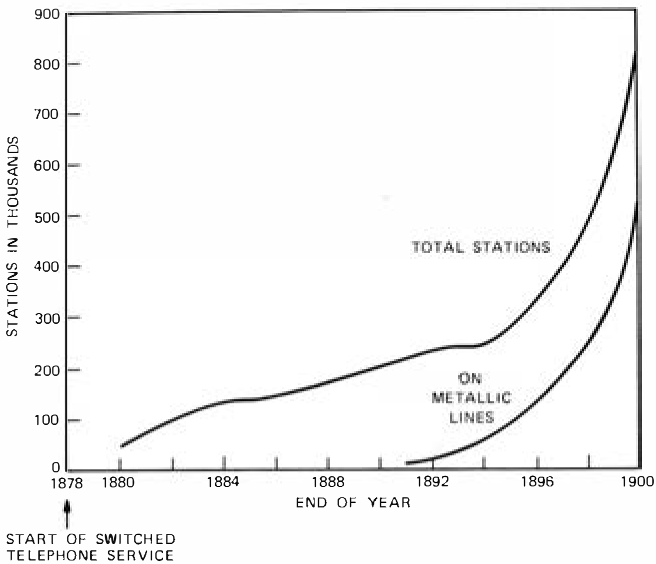


Fig. P-3. Bell System telephone stations, 1880–1900.

in use in 1885, only ten years after Bell's invention. Some of this rapid growth was probably due to novelty and the extraordinary interest in new things exhibited by people of the period. Part of the growth must also be credited to the telegraph which, starting about 40 years before the telephone, had demonstrated the value of rapid communication. Thus the public was ready to accept this new form of telecommunication, which added to the transmission speed of telegraphy a grand new dimension: the instant two-way vocal exchange between individuals with no need for extensive experience or complex prior training.

The great potential desire for this form of communication is evident when we recall how poor transmission quality was in 1885. Telephone communication at the time was carried out almost entirely by means of the unsuitable ground-return, iron-wire circuits previously developed for telegraphy. The unique electroacoustic converters which made the telephone possible were also in a very early stage of development. Receiver design had been pretty well stabilized in the 101 type of unipolar device with permanent-magnet bias, but transmitters were still of the single-contact carbon type, mostly of the Blake design. These left much to be desired from the standpoint of speech quality, efficiency, and stability, but provided proof to the public that the voice communication which they obviously desired was technically feasible. However, it was apparent to the technicians of the time that growth, par-

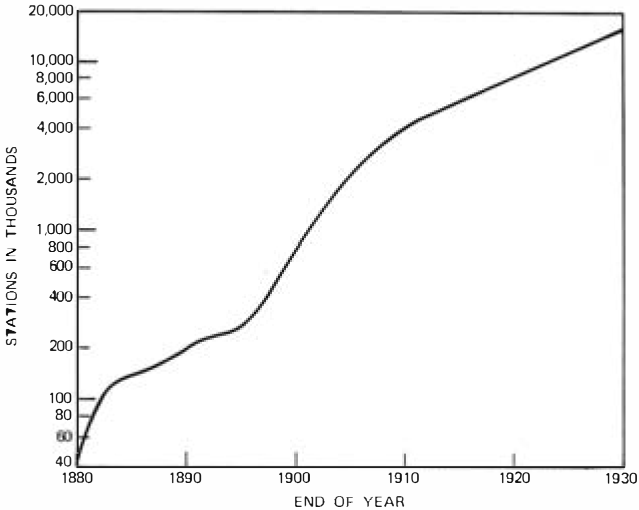


Fig. P-4. Bell-owned telephone stations, 1880–1930.

ticularly of long-distance telephony, was technologically limited unless important improvements could be made in transmission media and telephone instruments. Growth did indeed slow down for the ten years following 1884, but during this period telephone engineers were rapidly devising improvements, and these came along fast enough to support a linear increase of about 12,000 stations per year.

By the early 1890s, a completely new technological basis had been devised for a system of telephony capable of operating over distances of a thousand or so miles. New transmission media had been devised to replace grounded iron wire, the early telegraph-type facilities being rapidly replaced by balanced copper pairs in the form of open wire for long-distance circuits and paper-insulated, dry-core cable in the cities where both aerial and underground installations were being made. Telephone instruments were improved greatly by the use of the bipolar magnet in the receiver and the granular-carbon transmitter. The latter, which became practical with White's solid-back structure, provided the efficiency and stability that was needed for the network of long-distance lines which was the primary objective in the founding of the American Telephone and Telegraph Company in 1885. And, finally, common-battery switching had eliminated the twin nuisances of batteries and hand generators on the customer's premises.

Even though the goal of a nationwide network was still not in sight, system performance had been placed on a sound basis and, as shown in Figs. P-3 and P-4, a new spurt of growth began in the middle 1890s which continued on an exponential basis for many years.[1] While this growth was stimulated by the technical developments of the first 25 years of telephony, it continued at a high pace only because the second quarter-century brought completely new technology such as loaded lines, vacuum tube amplifiers, and carrier-current transmission systems. During most of this period of intense growth the telephone transmitter was the only source of amplification readily and cheaply available. Thus, a large part of the successful expansion of telephony over the vast area of the United States was owed to instruments designed to achieve the greatest possible distance with commercially practical transmission facilities.

Toward the end of our 50-year period, the development of line amplification began to free instrument design from the rigorous demands whereby transmission quality had been secondary to the achievement of maximum transmission distance. By the late 1920s, radically new instrument designs could be introduced into the plant and begin the vast improvement in ease and quality of transmission which featured the second 50 years of telephony and contributed largely to continued growth. However, the groundwork had been well laid by 1925.

II. TECHNOLOGICAL GROWTH

The growth of telephone technology has been the major subject of this volume, and we have just now mentioned briefly the more significant developments as background for interpreting plant growth. Thus there is no further need for a detailed summary of technical accomplishments in terms of hardware and systems; but a few words on the evolution of communication research and development techniques during these years will be pertinent.

It was during these years that a whole new field of technology evolved. Starting with Bell's telephone invention and essentially no practice or theory to guide them, a small group of ingenious, prolific inventors, with little knowledge of the fundamentals of electroacoustic conversion or the propagation of electric waves, set out to develop a communication system that could be used by anyone without previous experience. Later, some of these individuals were brought together in an integrated organization known today as the Bell System which also encompassed operation, engineering, and manufacture. This was to prove a wise and productive move, since it provided close interaction between user and producer and also provided means for stimulating

[1] Between 1895 and 1910 the growth was at an annual rate of about 20 percent. After 1910, growth slowed to 7–8 percent, but even including the seven depression years of the thirties, during which there was no growth, the annual increase in the 60 years since 1910 has averaged just under 6 percent.

invention and development through an interchange of ideas among the most knowledgeable practitioners of the art.

For a number of years, inventiveness and empiricism served well to provide the necessities for the fast-growing enterprise; but within the Bell System there was a visionary group of technical executives who foresaw the growth of a highly complex system and the need for a sounder theoretical approach based on analysis and measurement. To provide this base, the technical force was augmented by university graduates well trained in mathematics, physics, and the new art of electrical engineering. Their assignment was to establish for the developing Bell System the deeper theoretical foundations required in order that the fast-growing plant could benefit from the best of technology. In some cases this could be done by interpreting, for the practitioner, the research done by the "pure scientist" working far from the field of telephony. In other cases—and this became more common as telephonic science became more sophisticated—it became necessary to extend the theory provided by others or fill gaps in the field. By 1925, the value of this approach had been well demonstrated and the Bell R&D organization included many workers in pure and applied research as well as the engineers and technicians concerned with development and manufacture. Along with the theoretical work, measurement techniques, which are an essential in all true scientific work, were developed to fit the needs of the new field so that objective as well as subjective means became available for specifying and judging performance.

III. EVOLUTION OF TECHNICAL ORGANIZATION AND MANAGEMENT

The value of research, analysis, and measurement was not the only lesson learned in applying technology to the development of a nationwide telephone plant. From an early date it was found that certain techniques of management greatly enhanced the effectiveness of an R&D organization. As an example, it was soon found that purchasing equipment, such as switchboards, from a variety of manufacturers led to a lack of compatibility which degraded system performance. As a result, the Western Electric Company was purchased in the early 1880s and became the manufacturing unit of the Bell System. This made possible the needed close cooperation between the designer-producer of telephone equipment and the engineering force of the parent company which, as representative of the user, specified requirements.

This close liaison between user and manufacturer vastly increased the assurance that equipment designs would meet the user's needs and be compatible over the years and throughout the system. However, no mass production system is perfect, and techniques were needed to assure a satisfactory product without 100-percent inspection, which would have been prohibitively expensive. This problem was also solved by the application of scientific methods using the techniques of mathematical statistics. The methods used, later referred to as "Quality Assurance," were so successful that they have been widely adopted throughout industry. In an oversimplified form, Quality Assurance as developed in the Bell System involved both an organizational technique and technical methodology. The essence of the former was assigning responsibility for control of quality to a group remote from the production organization. The technical methods used the application of mathematical statistics to determine when a production method was "in control" (subject only to unavoidable random variation) and the assignment of appropriate performance limits when such production was achieved. With this information as a starting point, further statistical analysis was used to set up an inspection system based on relatively small samples of product, to provide assurance that the production standards were being met and that lack of production control would be promptly discovered.

Not all the problems of industrial research and development involve recourse to abstruse scientific approaches. Many involve, to a greater degree, the quality called "common sense." With the growth in size and complexity of the telephone plant, the solving of "system" type problems soon exceeded the capability of any one individual. Some form of group approach was required. It was soon found that a group working in concert was far more productive than the same number of people working as individuals, particularly when the group was composed of people trained in different but complementary disciplines. This "team" approach has been one of the most valuable characteristic techniques in the technical advancement of the Bell System.

As the System grew, remaining the dominant contributor to telephone technology, management recognized the potential dangers of depending too much on advances originating from within, and took steps to continually appraise the benefits which could arise from both external and internal competition. Some of the inventions which were to prove of great importance came from outside the Bell organization, such as the Edison transmitter, the de Forest audion, and the Strowger switch. As a result, much effort was devoted to following the art and testing new ideas so that valuable inventions such as these could be put to use. Very often these inventions were not of practical value as first conceived and needed extensive study and improvement by the Bell technical staff before their potential could be realized. The audion is an outstanding example of a device which originally had limited use because its functioning was not understood until subjected to intensive theoretical and experimental study by Arnold and his co-workers.

However, too much reliance could not be placed on external contributions, for relatively few outside workers had a broad knowledge of the communication field and it would be only fortuitous if the solution to an important current problem were to come from these sources. Primary reliance had to be placed, by necessity, on internal development effort; yet many solutions to a problem were often possible, with the optimum approach difficult to determine until development was well advanced. Thus there was instituted the unique approach of competing development teams, one of the first products of this approach being White's 1890 solid-back transmitter which successfully outperformed the instruments developed by several competing Bell designers. Probably the first wide-scale use of this technique was in the development of common-control switching. In this project, dual development teams not only provided competitive stimulation of ideas but, since they worked in the same general organization, could meet frequently for an interchange of ideas so that the general system plan could, at each development stage, incorporate the better features of the two approaches being followed. Since that period (the first decade of the twentieth century), internal intellectual competition has become the rule rather than the exception.

Towards the end of the nineteenth century, the addition of research scientists to the Bell staff brought new challenges to management. There had been a feeling in the predominantly academic research field that the industrial area, dominated by the goal of utility, was incompatible with the researcher's primary objective of a search for scientific knowledge without regard to immediate application. The manner in which these supposedly incompatible goals were melded into successful industrial research in the Bell System has been told very completely in Chapter 10, and only a few aspects will be briefly mentioned here to emphasize the necessity of meeting the special needs of the scientist.

The field of research, like the world in general, includes many kinds of people and it was of first importance to select those who had both the capability for outstanding scientific investigation and also the breadth of view to see that the application of new knowledge could be as exciting as its production. Fortunately, "detached" scientists having no desire, or even capability, for solving the day-to-day problems of applied technology, often display an eagerness to make their findings more useful when they discover that human needs are to be served. Such scientists, however, might not find industrial research permanently satisfying unless the environment is one which recognizes the long-range value of new knowledge long before it can be commercially demonstrated. There must be assurance that research can be pursued without dominance by business or manufacturing pressure. There must be encouragement and the physical means for free interchange of ideas both within the organization and in academic and professional circles. And finally, there must prevail an awareness of practical goals, a sense of purpose, without limiting the depth to which relevant knowledge can be pursued. It was an auspicious sequel to one of the greatest inventions in history, that the leaders of the enterprise had the vision to see these needs as they had not been seen before, and at an early date to adapt management style to their fulfillment.

The organizational milestone reached in 1925 by the formation of Bell Telephone Laboratories (as related in Chapter 2) was a formal recognition of the great value of the technological advances of the first 50 years. We have been pointing out that these advances included not only the creation of new technology but also major steps in the development of the art of managing R&D. This 50-year period had demonstrated the benefits of an integrated organization with close liaison between all elements from the producers of the system to the user of the services. It had demonstrated the value of pure and applied research, and finally it had shown that a large organization could maintain the highest degree of vitality by the use of internal competition and through guidance by an experienced and understanding management alert to changing times and ready to adapt to them.

IV. THE OUTLOOK

The management of the new Bell Telephone Laboratories was faced with many challenges, but obviously had good reason for facing them with confidence. As we look back from the seventies we can see that the second 50 years of telephony was a tremendously productive period that no one could have predicted in specific terms. Its history will be covered in other volumes, but it will not be out of place to speculate on what the problems looked like in 1925 and the paths toward their solution as seen by the telephone engineer of that day.

It must have been obvious that the telephone market was far from saturated. Only 40 percent of the United States households had telephones and only about one-third of these had single-party service. Service over transcontinental distances was still new and overseas service was not to commence for two years. Rapid growth in traffic had accompanied each extension of the service distance, and large growth in both long-distance and local service seemed inevitable if costs could be kept within the reach of the bulk of the populace. Transmission over the early telephones was tolerable, but long-distance transmission was often achieved only with considerable effort. By 1925, improvements were being made and accepted with enthusiasm. Early misgivings about machine switching had given way to enthusiastic acceptance. With only 13 percent of the Bell stations handled by machine methods, the potential for growth in this field was very large indeed.

The direction the future should take was also fairly obvious. Expansion of plant to provide more useful service, improvement in the speed of service and quality of transmission, and mechanization to avoid a potential shortage of operators were some obvious goals. And all of this had to be accomplished at reasonable and hopefully lowered cost to the user with appropriate return to the owners for the use of their capital. It was all summarized by Walter S. Gifford, then President of AT&TCo, in a speech made in Dallas on October 20, 1927, at a meeting of the National Railroad and Utility Commissioners. Two brief sentences quoted from his remarks will give the essence of his views:

> Our policy and purpose are the same as yours — the most telephone service and the best, at the least cost to the public . . . With your sympathetic understanding we shall continue to go forward, providing a telephone service for the nation more and more free from imperfections, errors or delays, and always at a cost as low as is consistent with financial safety.

Bell Laboratories could face this challenge with confidence.

The station designers were at last freed by recent transmission developments from the need to design for maximum volume efficiency. They could concentrate on improving the naturalness of speech and optimizing the transmission of intelligence within the constraints imposed by reasonable costs. And they had the theoretical and laboratory tools needed for designing to meet specifications and testing the efficiency of their designs.

The designers of the transmission network had physical and technical means in sight for carrying speech for almost any distance over land or sea. The use of radio (for overseas transmission) would be restricted by limitations in the available frequency spectrum, but the use of undersea cable over long distances was no longer inconceivable. Technically, the barrier of distance had been overcome and the main constraint was now economic. Here, carrier transmission had already demonstrated its potential for savings, though its ultimate power for economy of scale could not be imagined since it was to depend on inventions not yet conceived. (However, probably the most powerful tool, negative feedback, was only a few years away.) Improvement in the quality of speech transmitted was a certainty. Improved station sets were already being designed and further transmission improvement in telephone circuits would be the obvious result of the new technology made possible by the electron tube, carrier transmission, and the many offshoots of network theory.

The general pattern of the switching network had been set so that rapid establishment of calls from anyone, anywhere, to anyone else, anywhere else, was no longer a potential technical problem. Here again economics was the main constraint; the desired service was an objective to work towards with every hope for its ultimate achievement. Effective, easy-to-use, manual switching systems were available for switching the network and two highly usable machine systems were already in use. For the moment, growth in mechanization was largely a matter of cost reduction, orderly growth to avoid loss of jobs, and improvement in service features. All of these matters were well within the realm of practicality. There were surely some who had begun to think of possible new fields for machine switching (toll systems, for example) but there was probably no one who could conceive of all the wheels which would ultimately be set in motion by the invention some years before of the indirect, common-control type of system.

Backing up the work on systems, which had laid the groundwork for so much that was yet needed, were the successful management techniques which had been developed for conducting and applying research, the means for closely controlling the quality of manufactured product, and a type of organization providing close integration of the user, technical developer, and manufacturer.

One who lived through these days has an indelible memory of the feeling that pervaded the Bell R&D community, a feeling of confidence that a way would be found to meet their objective, and that the key to opening the closed doors would be the cooperative application of the best in science and technology while working closely with the manufacturer to fulfill the needs of the user.

References

Chapter 1. Bell's Telephone

Annual Reports. American Bell Telephone Co. and American Telephone and Telegraph Co.

Bell, A. G. *The Deposition of Alexander Graham Bell*. Boston: American Bell Telephone Co., 1908.

Bruce, R. V. *Bell: Alexander Graham Bell and the Conquest of Solitude*. Boston: Little, Brown & Co., 1973.

Carty, J. J. "Telephone Engineering." *AIEE Trans*. 25 (1906): 81.

Casson, H. N. *The History of the Telephone*. Chicago: A. C. McClurg & Co., 1910.

Coon, H. *American Tel & Tel, the Story of a Great Monopoly*. New York: Longmans, Green & Co., 1939.

Costain, T. B. *The Chord of Steel*. New York: Doubleday & Co., 1960.

Deland, F. "History of the Telephone." Unpublished manuscript.

Events in Telephone History. American Telephone and Telegraph Co., 1969.

Gray, Dr. E. *Experimental Researches in Electroharmonic Telegraphy and Telephony—1867–1878*. New York: Russell Bros., 1878.

Hanscom, C. D. *Dates in American Telephone Technology*. Preliminary ed. Bell Telephone Laboratories, 1961.

Holmes, E. T. *A Wonderful Fifty Years*. Privately published, 1917.

"Human Voice Transmitted by Telegraph." *Sci. Amer*. 11 (September 9, 1876): 163.

Kingsbury, J. E. *The Telephone and Telephone Exchanges*. New York: Longmans, Green & Co., 1915.

Mackenzie, C. *Alexander Graham Bell, the Man Who Contracted Space*. Boston: Houghton Mifflin Co., 1928.

Miller, K. B. *American Telephone Practice*. New York: American Electrician Co., 1899.

Miller, K. B. *American Telephone Practice*. 4th ed. New York: McGraw–Hill Book Co., 1905.

Page, A. W. *The Bell Telephone System*. New York: Harper & Bros., 1941.

Pound, A. *The Telephone Idea—Fifty Years After*. New York: Greenberg, 1926.

Proceedings of the National Telephone Exchange Association. 1880–1888.

Rhodes, F. L. *Beginnings of Telephony*. New York: Harper & Bros., 1929.

Richards, W. L. "Notes on Bell's Discovery of the Speaking Telephone." Unpublished manuscript.

Robertson, J. H. *The Story of the Telephone: History of the Telecommunications Industry of Britain*. London: Isaac Pitman & Sons, 1947.

Stehman, Dr. J. W. *Financial History of the AT&T Company*. Boston: Houghton Mifflin Co., 1925.

Still, A. *Communication Through the Ages, From Sign Language to Television*. New York: Holt, Rinehart & Winston, 1946.

Taylor, L. W. "The Untold Story of the Telephone." *The American Physics Teacher* 5 (1937):243.

Ten Men and the Telephone. American Telephone and Telegraph Co.

Things Worth Knowing About the Telephone. American Telephone and Telegraph Co., October 1, 1926.

Watson, T. A. *The Birth and Babyhood of the Telephone*. New York: American Telephone and Telegraph Co., 1931.

Watson, T. A. *Exploring Life*. New York: D. Appleton & Co., 1926.

Williams, A. *Telegraphy and Telephony*. Nashville, Tenn.: Thomas Nelson, Inc., 1928.

Chapter 2. Early Corporate History

Annual Reports. American Bell Telephone Co. and American Telephone and Telegraph Co.

Bruce, R. V. *Bell: Alexander Graham Bell and the Conquest of Solitude*. Boston: Little, Brown & Co., 1973.

Craft, E. B. "The Bell System Research Laboratories." *Electrical Commun.* 2 (January 1924):153.

Dixon, A. F. "Twenty Years at West Street." *Bell Laboratories Record* 5–6 (December 1927):138.

Doherty, W. H. "The Bell System and the People Who Built It: The Early Days." *Bell Laboratories Record* 46 (February 1968):38.

Doherty, W. H. "The Bell System and the People Who Built It: The Growing Years." *Bell Laboratories Record* 46 (March 1968):76.

Hill, R. B. and Shaw, T. "Hammond V. Hayes: 1860–1947." *Bell Telephone Mag.* 26 (Autumn 1947):150.

Langdon, W. C. "The Early Corporate Development of the Telephone." *Bell Telephone Quart.* 2 (July 1923):133.

Langdon, W. C. "Two Founders of the Bell System." *Bell Telephone Quart.* 2 (October 1923):213.

Mabon, P. C. *Mission: Communications*. Bell Telephone Laboratories, 1975.

Paine, A. B. *In One Man's Life*. New York: Harper & Bros., 1921.

U. S. Federal Communications Commission. "Investigation of the Telephone Industry in the United States." U. S. House of Representatives Document No. 340, June 14, 1939.

Chapter 3. Station Apparatus

General

American Telephone and Telegraph Co. *Telephone and Telegraph Statistics of the World*. Chief Statistician's Division, Annual Issues, 1912–1926.

Atkinson, E. *Ganot's Physics*. New York: William Ward & Co., 1893.

Campbell, G. A. "Cissoidal Oscillations." *AIEE Trans.* 30, Part 2 (April 1911):873.

Collected Papers of G. A. Campbell. New York: American Telephone and Telegraph Co., 1937.

Coolidge, O. H.; Gannett, D. K.; and Van Wynen, K. G. *Transmission Aspects of Wire Communication Systems*. Vol. 1, Chaps. 3 and 4. Bell Telephone Laboratories Communications Development Training Course, Issue 1.

Crandall, I. B. *Theory of Vibrating Systems and Sound*. New York: D. Van Nostrand Co., 1926.

DeLand, F. "Notes on the Development of Telephone Service." *Popular Science Monthly* 69 (November 1906):425; 69 (December 1906):490.

Dolbear, A. E. "A New System of Telephony." *Sci. Amer.* 44 (June 1881):388.

Du Moncel, Count. *The Telephone, the Microphone and the Phonograph*. New York: Harper & Bros., 1879, pp. 18–22.

Events in Telephone History. American Telephone and Telegraph Co., 1969.

Fletcher, H. *Speech and Hearing*. New York: D. Van Nostrand Co., 1929.

Gherardi, B. "Henry as an Electrical Pioneer." *Bell System Technical J.* 11 (July 1932):327.

Green, E. I. "Telephony." *Encyclopaedia Britannica* 21 (1957):895–906.

Hanscom, C. D. *Dates in American Telephone Technology*. Preliminary ed. Bell Telephone Laboratories, 1961.

Hunt, F. V. *Electroacoustics*. Cambridge, Mass.: Harvard University Press, 1954.

Inglis, A. H. Notes of interview by F. L. Crutchfield at Bell Laboratories (1968).

Johnson, K. S. "Discussion of the Various Developments Made in Substation Sets—With an Estimate of Their Economic Values." Internal memorandum, Bell Telephone Laboratories (March 25, 1920).

Johnson, K. S. *Transmission Circuits for Telephonic Communication*. Western Electric Co., 1924.

Kingsbury, J. E. *The Telephone and Telephone Exchanges*. London: Longmans, Green & Co., 1915.

Martin, W. H. "Seventy-Five Years of the Telephone: An Evolution in Technology." *Bell System Technical J.* 30 (April 1951):214.

Miller, K. B. *American Telephone Practice*. 4th ed. New York: McGraw–Hill Book Co., 1905.

Prescott, G. B. *Bell's Electric Speaking Telephone*. New York: D. Appleton & Co., 1884.

Prescott, G. B. *Speaking Telephone, Electric Light and Other Recent Electrical Inventions*. New York: D. Appleton & Co., 1879.

Rhodes, F. L. *Beginnings of Telephony*. New York: Harper & Bros., 1929.

Rhodes, F. L. *John J. Carty, An Appreciation*. New York: Privately printed, 1932.

"Unveiling of Memorial Tablet to Alexander Graham Bell at 16 South Charlotte Street, Edinburgh." *Scottish Electrical Engineer* 8 (December 1937).

Vinal, G. W. *Primary Batteries*. New York: Wiley, 1950.

Walsh, J. L. *Connecticut Pioneers in Telephony*. New Haven: Morris F. Tyler Chapter, Telephone Pioneers of America, 1950.

Western Electric Co. Catalogs, Bulletins, and Apparatus Card Files (various years, 1890–1930).

Transmitters and Receivers

Bell, A. G. *The Deposition of Alexander Graham Bell*. Boston: American Bell Telephone Co., 1908.

Frederick, H. A. "The Development of the Microphone." *Bell Telephone Quart*. 10 (July 1931):164.

Hammer, K. E. "Transmitters and Receivers Used as Wall Sets and Deskstands." Internal memorandum, Bell Telephone Laboratories (October 1949).

"List of Transmitter Types That Have Been Standard, Including Current Standard Types, in the Bell System." Unpublished document, Bell Telephone Laboratories.

Thompson, G. K. Internal memorandum for L. F. Morehouse, Equipment Development Engineer (January 18, 1924).

Thompson, G. K. "Notes on Telephone Transmitters." Internal memorandum, Bell Telephone Laboratories (November 6, 1924).

Customer Telephones

Campbell, G. A. and Foster, R. N. "Maximum Output Networks for Telephone Substations and Repeater Networks." *AIEE Trans.* 39, Part 1 (1920):231.

Foley, J. W. "The Anti-Sidetone Station Circuit." *Bell Laboratories Record* 17 (July 1939):347.

Johnson, K. S. "Chronological Development of Commercial Subscriber Circuits." Unpublished document, Bell Telephone Laboratories (March 8, 1920).

Johnson, K. S. "Discussion of the Various Developments Made in Substation Sets—With an Estimate of Their Economic Values." Internal memorandum, Bell Telephone Laboratories (March 25, 1920).

Thompson, G. K. "Development of the Wall Type of Substation Set." Internal memorandum, Bell Telephone Laboratories (April 10, 1930).

Wright, C. A. and Puchstein, A. F. *Telephone Communication*. New York: McGraw–Hill Book Co., 1925.

Ringing, Signaling, and Switching

Allen, L. H. "The 1A Key Telephone System." *Bell Laboratories Record* 18 (June 1940):315.

Cruser, V. I. "Multiplying the Subscriber's Line." *Bell Laboratories Record* 8 (July 1930):527.

"Early Developments in Telephone Signalling." *Bell Laboratories Record* 1 (January 1926):207.

Hill, R. B. "Early Work on Dial Telephone Systems." *Bell Laboratories Record* 31 (January 1953):22.

Stacy, L. J. "Calling Subscribers to the Telephone." *Bell Laboratories Record* 8 (November 1929):113.

Stacy, L. J. "The History of Party Line Ringing." Internal memorandum, Bell Telephone Laboratories (1935).

Stuart, W. M., Jr. "The Station Ringer." *Bell Laboratories Record* 10 (July 1932): 385.

"Telephone Signalling." *Bell Laboratories Record* 1 (February 1926):241.

Watson, T. A. *The Birth and Babyhood of the Telephone*. New York: American Telephone and Telegraph Co., 1931.

Power Supplies

Miller, K. B. *American Telephone Practice*. 4th ed. New York: McGraw–Hill Book Co., 1905.

Vinal, G. W. *Primary Batteries*. New York: Wiley, 1950, pp 22–23, 104–105.

Special-Purpose Telephones

Beardsley, H. I. "Operators' Transmitters and Receivers." *Bell Laboratories Record* 7 (January 1929):203.

"The Telephone's Use in Naval Warfare." *Western Electric News* 5 (May 1916):1.

Handsets

Frederick, H. A. "Early Handsets." *Bell Laboratories Record* 11 (July 1934):322.

Jones, W. C. and Inglis, A. H. "The Development of a Handset for Telephone Stations." *Bell System Technical J.* 11 (April 1932):245.

Coin Boxes

Agreement Between Gray Telephone Pay Station Company and the Western Electric Company (October 18, 1910).

"The First Telephone Pay Station." *The Telephone Bulletin*. Southern New England Telephone Co. (October 1923).

History of the Development of the Gray Pay Station. Hartford, Conn.: The Gray Telephone Pay Station Co., 1941.

Mabon, P. C. "Public Telephones." *Bell Telephone Quart.* 18 (October 1939):246.

Willis, W. L. "History of Coin Collectors." Internal memorandum, Bell Telephone Laboratories (January 18, 1924).

Booths

"A Not So Private Peek at Some Very Private Public Telephones." *The Telephone News* (July 1965):20.

Mabon, P. C. "Public Telephones." *Bell Telephone Quart.* 18 (October 1939):246.

High-Quality Instruments

Curl, H. C. "Amplifier for Condenser Transmitter." *Bell Laboratories Record* 6 (June 1928):329.

"For Better Radio Reception." *Bell Laboratories Record* 1–2 (December 1925):160.

Green, I. W. and Maxfield, J. P. "Public Address Systems." *Bell System Technical J.* 2 (April 1923):113.

"Historic Firsts—The Condenser Microphone." *Bell Laboratories Record* 21 (1943):394.

Jones, W. C. and Giles, L. W. "A Moving-Coil Microphone for High-Quality Sound Reproduction." *J. Soc. Motion Picture Engineers* 17 (December 1931):977.

Sivian, L. J. "A Telephone Transmission Reference System." *Electrical Commun.* 3 (October 1924):114.

Wente, E. C. "Absolute Measurement of Sound Intensity." *Phys. Rev.* 10 (July 1917):51.

Wente, E. C. "Electrostatic Transmitter." *Phys. Rev.* 19 (May 1922):498.

Wente, E. C. and Thuras, A. L. "A High Efficiency Receiver for a Horn-Type Loud Speaker of Large Power Capacity." *Bell System Technical J.* 7 (January 1928):140.

Wente, E. C. and Thuras, A. L. "Moving-Coil Telephone Receivers and Microphones." *Bell System Technical J.* 10 (October 1931):565.

Western Electric Co. "No. 540-AW Loud Speaking Telephone." Instruction Booklet (1924).

Summary—Technological Growth

Campbell, G. A. "Cissoidal Oscillations." *AIEE Trans.* 30, Part 2 (April 1911):873.

Campbell, G. A. "Physical Theory of the Electric Wave-Filter." *Bell System Technical J.* 1 (November 1922):1.

Crandall, I. B. *Theory of Vibrating Systems and Sound.* New York: D. Van Nostrand Co., 1926.

Hartley, R. V. L. "The Transmission Unit." *Electrical Commun.* 3 (July 1924):34.

Martin, W. H. "The Transmission Unit and Telephone Transmission Reference Systems." *Bell System Technical J.* 3 (July 1924):400.

Shea, T. E. *Transmission Networks and Wave Filters.* New York: D. Van Nostrand Co., 1929.

Chapter 4. Telephone Transmission—The Wire Plant

General Transmission Material

American Technical Society. *Encyclopedia of Telephony and Telegraphy.* 1919.

American Telephone and Telegraph Co. *Telephone Almanac.* New York: 1922–1925.

Capen, W. H. "Wire Telephone Communication in Theory and Practice." *Proc. Telephone and Telegraph Section*, American Railway Assn., October, 1925. [Also in *Electrical Commun.* 4 (January 1926):161].

Clark, A. B. "The Development of Telephony in the United States." *AIEE Trans.* 71, Part 1 (1952):348.

Emling, J. W. "The Conquest of Distance." *IRE Student Quart.* (May 1962):3.

Green, E. I. "Telephony." *Encyclopaedia Britannica* 21 (1957):895–906.

Historic Firsts. Pamphlet consisting of selected reprints from the *Bell Laboratories Record*. New York: Bell Telephone Laboratories, 1948.

Jewett, F. B. "Long Distance Telephony in America." Paper presented at the Congresso Internazionale delle Applicazioni Elettriche, 10–17 September 1911, at Torino.

Stecker, R. B. "A History of Transmission Systems." *Western Electric Engineer* 3 (October 1959):2.

Early Development of Open Wire and Cable

Awarding of the Edward Longstreth Medal to Thomas B. Doolittle, Abstract of proceedings of June 1, 1898 meeting of the Franklin Institute Committee on Science and the Arts. *J. Franklin Inst.* (August 1898):158.

"Cable Sheath—a Review." *Western Electric Engineer* 2 (October 1958):2.

Campbell, A. S. "We Had Little to Guide Us." *Long Lines* 4 (April 1925):17.

Doolittle, T. B. "Hard Drawn Copper Wire." *Telephone Topics* (December 1910).

Grace, S. P. "The Telephone Wire Plant." *AIEE Trans.* 26, Part 1 (1907):569.

Green, E. I. "The Transmission Characteristics of Open-Wire Telephone Lines." *Bell System Technical J.* 9 (October 1930):730.

Griswold, A. H. "Design of Telephone Pole Lines for Conditions West of the Rocky Mountains." *AIEE Trans.* 31 (May 1912):481.

Jewett, F. B. "The Modern Telephone Cable." *AIEE Trans.* 28, Part 2 (June–December, 1909):1079.

Lawson, C. C. "Neoprene Jacketed Telephone Wires." *Wire and Wire Products* 25 (October 1950):829.

Lockwood, T. D. "Earth Wires; or, the Earth as an Electric Circuit Completer." Paper read before the American Institute of Electrical Engineers, October 1881, at Philadelphia.

Lockwood, T. D. "Some Recent Advances in Telephony." *AIEE Trans.* 3 (1886):74.

Martin, W. H.; Anderegg, G. A.; and Kendall, B. W. "Key West–Havana Submarine Telephone Cable." *AIEE Trans.* 41 (March 1922):184.

Mills, J.; Nance, H. H.; Oram, R. M.; and Jewett, F. B. "A Quarter Century of Transcontinental Telephone Service." *Bell Telephone Quart.* 19 (January 1940):3.

Owens–Illinois Glass Co. *Hemingray Insulators*. Chicago, 1946.

Pilliod, J. J. "Philadelphia–Pittsburgh Section of the New York–Chicago Cable." *AIEE Trans.* 41 (June 1922):446.

Rhodes, F. L. "Cable Development and Research." *Bell Telephone Mag.* 3 (January 1924):30.

Rhodes, F. L. "How the Telephone Wires Were First Put Underground." *Bell Telephone Mag.* 2 (October 1923):240.

Rhodes, F. L. and Hosford, R. F. "Recent Results Obtained from the Preservative Treatment of Telephone Poles." *AIEE Trans.* 34, Part 2 (1915):2549.

Rhodes, F. L. "A Short History of the Development and Use of Cable in the Bell System." Unpublished manuscript.

Weaver, M. A. "The Long Struggle against Cable Crosstalk." *Bell Telephone Quart.* 14 (January 1935):3.

Western Electric Co. Catalogs (1908, 1911, 1916, 1929).

Transmission Theory

Appleyard, R. "The Origin and Development of the Transmission Equation." *Electrical Commun.* 2 (January 1924):183.

Johnson, K. S. *Transmission Circuits for Telephonic Communication.* Western Electric Co., 1924, pp. 143–151.

Loading, Phantom Circuits, and Related Matter

Brittain, J. E. "The Introduction of the Loading Coil: George A. Campbell and Michael I. Pupin." *Technology and Culture* (January 1970):36.

Campbell, G. A. "On Loaded Lines in Telephonic Transmission." *Philosophical Mag.* 5 (March 1903):313.

The Collected Papers of George Ashley Campbell. New York: American Telephone and Telegraph Co., 1937.

Fondiller, W. "Commercial Loading of Telephone Cable." *Electrical Commun.* 4 (July 1925):24.

Gherardi, B. "The Commercial Loading of Telephone Circuits in the Bell System." *AIEE Trans.* 30, Part 3 (June–December 1911):1743.

Hill, R. B. and Shaw, T. "Hammond V. Hayes: 1860–1947." *Bell Telephone Mag.* 26 (Autumn 1947):150.

Jewett, F. B. "Dr. George A. Campbell." *Bell System Technical J.* 14 (October 1935):553.

Shaw, T. "The Evolution of Inductive Loading for Bell System Telephone Facilities." *Bell System Technical J.* 30(1951):149.

Shaw, T. and Fondiller, W. "Development and Application of Loading for Telephone Circuits." *Bell System Technical J.* 5 (April 1926):221.

Repeaters, Their Application to the Line, and Other Voice-Frequency Problems Involved in Building Up Voice-Frequency Telephone Plant

Bouton, L. L. "Four-Wire Circuits in Retrospect." *Bell Laboratories Record* 17 (December 1938):116.

Clark, A. B. "Telephone Transmission over Long Cable Circuits." *Bell System Technical J.* 2 (January 1923):67.

Clark, A. B. and Mathes, R. C. "Echo Suppressors for Long Telephone Circuits." *AIEE Trans.* 44 (April 1925):481.

Demarest, C. S. "Telephone Equipment for Long Cable Circuits." *Bell System Technical J.* 2 (July 1923):112.

Gherardi, B. and Jewett, F. B. "Telephone Repeaters." *Proc. of AIEE* 38 (November 1919):1255.

Jewett, F. B. "Transcontinental Panorama." *Bell System Quart.* 19 (January 1940):46.

Lockwood, T. D. "Telephone Repeaters or Relays, and Repeating Systems." *The Electrical World* 28 (November 1896):593.

Osborne, H. S. "Telephone Transmission over Long Distances." *AIEE Trans.* 42 (October 1923):1051.

Shaw, T. "The Conquest of Distance by Wire Telephony." *Bell System Technical J.* 23 (October 1944):337.

Smith, C. A. "Fifty Years of Telephone Repeaters." *Bell Laboratories Record* 27 (January 1949):5.

Carrier, Public Address, and Program Transmission

Bower, M. M. "The Carrier Telephone Alphabet." *Bell Laboratories Record* 16 (February 1938):208.

Clark, A. B. and Green, C. W. "Long Distance Cable Circuit for Program Transmission." *Bell System Technical J.* 9 (July 1930):567.

Clark, A. B. and Osborne, H. S. "Long Distance Telephone Circuits in Cable." *Bell System Technical J.* 11 (October 1932):520.

Colpitts, E. H. and Blackwell, O. B. "Carrier Current Telephony and Telegraphy." *AIEE Trans.* 40 (1921):205.

Cowan, F. A. "Telephone Circuits for Program Transmission." *AIEE Trans.* 48 (July 1929):1045.

Green, I. W. and Maxfield, J. P. "Public Address Systems." *Bell System Technical J.* 2 (April 1923):113.

Hartley, R. V. L. "Transmission Limits of Telephone Lines." *Bell Laboratories Record* 1 (February 1926):225.

Horton, J. W. "Multiplex Transmission by Carrier Currents." *Bell Laboratories Record* 1 (December 1925):147.

Jewett, F. B. "Making the Most of the Line." Paper read before the American Institute of Electrical Engineers, 17 October 1923, at Philadelphia.

Kendall, B. W. "Carrier Current-Telephone Systems." *Bell Laboratories Record* 1 (December 1925):154.

Martin, W. H. and Clark, A. B. "Use of Public Address System with Telephone Lines." *Bell System Technical J.* 2 (April 1923):143.

Martin, W. H. and Fletcher, H. "High Quality Transmission and Reproduction of Speech and Music." *J. AIEE* 43 (March 1924):384.

Rose, A. F. "Practical Application of Carrier Telephone and Telegraph in the Bell System." *Electrical Commun.* 1 (April 1923):16.

Rose, A. F. "Twenty Years of Carrier Telephony." *Bell Telephone Quart.* 17 (October 1938):245.

Slaughter, N. H. "Carrier Telephone Equipment for High-Voltage Power Transmission Lines." *Electrical Commun.* 3 (October 1924):127.

Slaughter, N. H. and Wolfe, W. V. "Carrier Telephone on Power Lines." *AIEE Trans.* 43 (April 1924):620.

Stone, J. S. "The Practical Aspects of the Propagation of High Frequency Electric Waves along Wires." *J. Franklin Inst.* 174 (October 1912):353.

Wolfe, W. V. "Carrier Telephony on High Voltage Power Lines." *Bell System Technical J.* 4 (January 1925):152.

Vacuum Tubes

"Hendrick Johannes van der Bijl." *Proc. IRE* 33 (1945):352.

King, R. W. "Thermionic Vacuum Tubes and their Applications." *Bell System Technical J.* 2 (October 1923):31.

Llewellyn, F. B. "Birth of the Electron Tube Amplifier." *Radio and Television News* (March 1957):43.

Pratt, H. "Radio Ancestors—An Autobiography by Robert H. Marriott." *IEEE Spectrum* 5 (June 1968):52.

Van der Bijl, H. J. *The Thermionic Vacuum Tube and Its Applications.* New York: McGraw–Hill Book Co., 1920.

Measuring Gear and Maintenance

Alford, R. S. "Crosstalk Measurements." *Bell Laboratories Record* 15 (August 1937):377.

Augustadt, H. W. "Transmission Networks and their Measurement." *Bell Laboratories Record* 10 (April 1932):296.

Barstow, J. M. "A New Noise Meter." *Bell Laboratories Record* 15 (April 1937):252.

Bertels, A. S. "Testing Ringers and Dials at Subscribers' Stations." *Bell Laboratories Record* 8 (February 1930):263.

Best, F. H. "Decibel Meters." *Bell Laboratories Record* 15 (January 1937):167.

Best, F. H. "Measuring Methods for Maintaining Transmission Efficiency of Telephone Circuits." *J. AIEE* 43 (February 1924):136.

Castner, T. G. "A General-Purpose Frequency Analyzer." *Bell Laboratories Record* 13 (May 1935):267.

Douglass, H. T. "Transmission Testing of Central-Office Circuits." *Bell Laboratories Record* 7 (March 1929):283.

Ferguson, J. G. "A Clock-Controlled Tuning Fork as a Source of Constant Frequency." *Bell System Technical J.* 3 (January 1924):145.

Ford, G. T. "Standard Housings for Portable Test Sets." *Bell Laboratories Record* 8 (November 1929):120.

Gerks, I. H. "Extending the Usefulness of the Oscillograph in Circuit Testing." *Bell Laboratories Record* 7 (May 1929):352.

Harden, W. H. "Electrical Tests and their Applications in the Maintenance of Telephone Transmission." *Bell System Technical J.* 3 (July 1924):353.

Harden, W. H. "Practices in Telephone Transmission Maintenance Work." *Bell System Technical J.* 4 (January 1925):26.

Hartley, R. V. L. "The Transmission Unit." *Electrical Commun.* 3 (July 1924):34.

Hartley, R. V. L. " 'TU' Becomes 'Decibel.' " *Bell Laboratories Record* 7 (December 1928):137.

Heising, R. A. "Generating High Frequencies with Precision." *Bell Laboratories Record* 11 (December 1932):100.

Heising, R. A. *Quartz Crystals for Electrical Circuits; their Design and Manufacture.* New York: Van Nostrand, 1946.

Henneberger, T. C. and Edwards, P. G. "Bridge Methods for Locating Resistance Faults on Cable Wires." *Bell System Technical J.* 10 (July 1931):382.

Johnson, J. B. "A Low Voltage Cathode Ray Oscillograph." *Bell System Technical J.* 1 (November 1922):142.

Krazinski, L. C. "Primary Toll Test Boards." *Bell Laboratories Record* 13 (April 1935):245.

Kuebler, R. E. "An Improved Volume Indicator." *Bell Laboratories Record* 13 (December 1934):122.

Lundius, E. R. "Automatic Testing Equipment for Trunk Circuits." *Bell Laboratories Record* 9 (August 1931):589.

Lutomirski, K. "Toll Transmission Measuring System for the No. 8 Test and Control Board." *Bell Laboratories Record* 12 (August 1934):367.

Martin, W. H. "The Transmission Unit and Telephone Transmission Reference Systems." *J. AIEE* 43 (June 1924):504.

Martin, W. H. and Gray, C. H. G. "Master Reference System for Telephone Transmission." *Bell System Technical J.* 8 (July 1929):536.

Moore, C. R. and Curtis, A. S. "An Analyzer for the Voice Frequency Range." *Bell System Technical J.* 6 (April 1927):217.

Nielsen, J. E. "A Crosstalk Measuring Set of Improved Precision." *Bell Laboratories Record* 11 (October 1932):53.

Pascarella, A. J. "Locating Faults on Toll Lines." *Bell Laboratories Record* 7 December 1928):161.

Sivian, L. J. "Speech Power and Its Measurement." *Bell System Technical J.* 8 (October 1929):646.

Sivian, L. J. "A Telephone Transmission Reference System." *Electrical Commun.* 3 (October 1924):114.

Slonczewski, T. "Automatic Measurement of Transmission." *Bell Laboratories Record* 15 (October 1936):56.

Smith, C. W. "Practical Application of the Transmission Unit." *Bell System Technical J.* 3 (July 1924):409.

Zammataro, S. J. "Impedance Bridges." *Bell Laboratories Record* 8 (December 1929):167.

Zammataro, S. J. "A Universal Laboratory Transmission-Measuring Set." *Bell Laboratories Record* 2 (May 1926):98.

Growth and Plant Statistics

Annual Reports to Stockholders. American Bell Telephone Co. and American Telephone and Telegraph Co.

Baldwin, F. G. C. *History of the Telephone in the United Kingdom*. London: Chapman, 1925.

Gherardi, B. and Jewett, F. B. "Telephone Communication System of the United States." *Bell System Technical J.* 9 (January 1930): 1.

Mougey, W. E. "Principal Stages of Development in Toll Cables in the Bell System." From "Story of Telephone Cables," Internal memorandum, Bell Telephone Laboratories (July 5, 1945).

Rhodes, F. L. *The Beginnings of Telephony*. New York: Harper & Bros., 1929.

Royal, W. C. "A Brief History of Telephone Cable Development." Internal memorandum, Bell Telephone Laboratories.

Control of Interference

Babcock, W. C.; Rentrop, E.; and Thaeler, C. S. "Crosstalk on Open-Wire Lines." Bell System Monograph 2520 (1955).

Carty, J. J. "Inductive Disturbances in Telephone Circuits." *AIEE Trans.* 8 (1891):99.

Chapman, A. G. "Open-Wire Crosstalk." *Bell System Technical J.* 13 (January 1934):19; 13 (April 1934):195.

Corbett, L. J. *Inductive Coordination of Electric Power and Communication Circuits.* San Francisco: The J. H. Neblett Pressroom, Ltd., 1936, Introduction.

Friedlander, G. D. "Railroad Electrification: Past, Present, and Future," *IEEE Spectrum* 5 (September 1968):77.

Hill, R. B. "John A. Barrett, Pioneer in Telephone Research." *Bell Laboratories Record* 27 (July 1949):255.

Jansky, C. M., and Faber, D. C. *Principles of the Telephone, Part I.* New York: McGraw-Hill Book Co., 1916, pp. 117–131.

May, D. T. "The Development of the Protector Block." *Bell Laboratories Record* 11 (November 1932):80.

Pack, R. F. "Introductory Remarks—Symposium on Coordination of Power and Telephone Plant." *Bell System Technical J.* 10 (April 1931):155.

Savage, E. S. "Fuses." *Bell Laboratories Record* 12 (September 1933):27.

Telegraphy

Bell, J. H. "Composite Telegraphy." *Bell Laboratories Record* 7 (December 1928):140.

Chapter 5. Telephone Transmission—The Advent of Radio

General

Banning, W. P. *Commercial Broadcasting Pioneer: The WEAF Experiment, 1922–1926*. Cambridge, Mass.: Harvard University Press, 1946.

Blake, G. G. *History of Radio Telegraphy and Telephony*. London: Chapman and Hall, Ltd., 1928, pp. 65–67.

DeSoto, C. B. *200 Meters and Down*. West Hartford, Conn.: American Radio Relay League, 1936.

Espenschied, L. "The Origin and Development of Radiotelephony." *Proc. IRE* 25 (September 1937):1101.

Fleming, J. A. *Electric Wave Telegraphy and Telephony*. 3rd ed. London: Longmans, Green & Co., 1916, pp. 467–474.

Fleming, J. A. *Principles of Electric Wave Telegraphy and Telephony*. 4th ed. New York: Longmans, Green & Co., 1919, pp. 368–372.

Lodge, O. J. *Signalling through Space without Wires*. 3rd ed. New York: D. Van Nostrand Co., 1900.

Maclaurin, W. R. *Invention and Innovation in the Radio Industry*. New York: Macmillan Co., 1949.

Michaelis, A. R. *From Semaphore to Satellite*. Geneva: International Telecommunication Union, 1965.

Page, A. W., et al. *Modern Communication*. Boston: Houghton Mifflin Co., 1932.

Pratt, H. "Radio Ancestors—An Autobiography by Robert H. Marriott." *IEEE Spectrum* 5 (June 1968):52.

Shiers, G. "The First Electron Tube." *Sci. Amer.* 220 (March 1969):104.

Susskind, C. "The Early History of Electronics. II. The Experiments of Hertz." *IEEE Spectrum* 5 (December 1968):57; "III. Prehistory of Radiotelegraphy."

IEEE Spectrum 6 (April 1969):69; "IV. First Radiotelegraphy Experiments." *IEEE Spectrum* 6 (August 1969):66.

Wheeler, L. P. "Nikola Tesla's Achievements in the Electrical Arts; Part II. Tesla's Contribution to High Frequency." *Electrical Engineer* 62 (August 1943):355.

General Technical Background

Bullington, K. "Radio Propagation Fundamentals." *Bell System Technical J.* 36 (May 1957):593.

Everitt, W. L. *Communication Engineering*. New York: McGraw–Hill Book Co., 1937.

Hartley, R. V. L. "Relations of Carrier and Side-Bands in Radio Transmission." *Proc. IRE* 11 (February 1923):34.

Heising, R. A. "Modulation in Radio Telephony." *Proc. IRE* 9 (August 1921):305.

Heising, R. A. "Production of Single Sideband for Transatlantic Radio Telephony." *Proc. IRE* 13 (June 1925):291.

"Historic Firsts: The Heising Modulator." *Bell Laboratories Record* 23 (January 1945):6.

Polkinghorn, F. A. "Commercial Single Sideband Radiotelephone Systems." *Communications* 28 (December 1948):24.

Early Bell System Work Through World War I

Craft, E. B. and Colpitts, E. H. "Radio Telephony." *Proc. AIEE* 38, Part 1 (March 1919):337.

Heising, R. A. "History of Airplane Radio." *Bridge of Eta Kappa Nu* (Fall 1962):2.

"Pioneering in Radio Telephony." *Bell Telephone Mag.* 20 (February 1941):21.

Early Postwar (I) Era

Griswold, A. H. "The Radio Telephone Situation." *Bell Telephone Quart.* 1 (April 1922):2.

Nichols, H. W. and Espenschied, L. "Radio Extension of the Telephone System to Ships at Sea." *Proc. IRE* 11 (June 1923):193.

"The Place of Radio in the Communication Field." *Western Electric News* 10 (October 1921):16.

Transatlantic—General and Low-Frequency

Arnold, H. D. and Espenschied, L. "Transatlantic Radio Telephony." *J. AIEE* 42 (August 1923):815.

Bailey, A. and McCann, T. A. "Application of Printing Telegraph to Long-Wave Radio Circuits." *Bell System Technical J.* 10 (October 1931):601.

Blackwell, O. B. "Transatlantic Telephony—The Technical Problem." *Bell System Technical J.* 7 (April 1928):168.

Bown, R. "Transoceanic Radiotelephone Development." *Proc. IRE* 25 (September 1937):1124.

"Who Wants to Talk Across the Atlantic?" *195 Mag.* 38 (October 1965):24.

Transatlantic High-Frequency

Bown, R. "Transoceanic Telephone Service—Short-Wave Transmission." *Bell System Technical J.* 9 (April 1930):258.

"Historic Firsts: The Horizontal Rhombic Antenna." *Bell Laboratories Record* 22 (June 1944):440.

O'Dea, W. T. *Science Museum of the Handbook of the Collections Illustrating Electrical Engineering. II. Radio Communication. Part I. History and Development.* London: His Majesty's Stationery Office, 1934.

Oswald, A. A. "Transoceanic Telephone Service—Short-Wave Equipment. *Bell System Technical J.* 9 (April 1930):270.

Schelleng, J. C. "Some Problems in Short-Wave Telephone Transmission." *Proc. IRE* 18 (June 1930):913.

Broadcasting—General

Barrett, R. T. "Network Broadcasting." *Bell Telephone Quart.* 13 (April 1934):75.

Bown, R. and Gillett, G. D. "Distribution of Radio Waves from Broadcasting Stations over City Districts." *Proc. IRE* 12 (August 1924):395.

Bown, R.; Martin, D. K.; and Potter, R. K. "Some Studies in Radio Broadcast Transmission." *Bell System Technical J.* 5 (January 1926):143.

Clark, A. B. "Wire Line Systems for National Broadcasting." *Bell System Technical J.* 9 (January 1930):141.

Clark, A. B. and Green, C. W. "Long Distance Cable Circuit for Program Transmission." *Bell System Technical J.* 9 (July 1930):567.

Cowan, F. A. "Telephone Circuits for Program Transmission." *AIEE Trans.* 48 (July 1929):1045.

Espenschied, L. "Radio Broadcast Coverage of City Areas." *Bell System Technical J.* 6 (January 1927):117.

Events in Telephone History. New York: American Telephone and Telegraph Co., p. 28.

Foland, H. S. and Rose, A. F. "Telephone Circuits Used as an Adjunct to Radio Broadcasting." *Electrical Commun.* 3 (January 1925):194.

Martin, D. K., Gillett, G. D., and Bemis, I. S. "Common Frequency Broadcasting." *Proc. IRE* 15 (March 1927):213.

Thayer, H. B. "The Radio Broadcasting Situation." *Bell Telephone Quart.* 3 (April 1924):113.

Broadcast Equipment

Doherty, W. H. "A New High Efficiency Power Amplifier for Modulated Waves." *Proc. IRE* 24 (September 1936):1163.

Doherty, W. H. and Towner, O. W. "A 50-Kilowatt Broadcast Station Utilizing the Doherty Amplifier and Designed for Expansion to 500 Kilowatts." *Proc. IRE* 27 (September 1939):531.

"Historic Firsts: Crystal-Controlled Broadcasting." *Bell Laboratories Record* 26 (October 1948):420.

Logan, M. A. "A Recording Transmission Measuring Set." *Bell Laboratories Record* 13 (June 1935):295.

Nelson, E. L. "Broadcasting Transmitters and Related Phenomena." *Bell System Technical J.* 9 (January 1930):121.

Nelson, E. L. "Transmitting Equipment for Radio Telephone Broadcasting." *Proc. IRE* 12 (October 1924):553.

Slaughter, N. H. "Four Years' Progress in Radio Broadcasting." *Bell Laboratories Record* 3 (September 1926):1.

Western Electric Co. Bulletin "No. 103-C (50-watt) Radio Telephone Broadcasting Equipment." (Plus others covering equipment developed before 1926.)

Radio Measurements

Bown, R., Englund, C. R., and Friis, H. T. "Radio Transmission Measurements." *Proc. IRE* 11 (April 1923):115.

Coram, R. E. "A Frequency Monitoring Unit for Broadcast Stations." *Bell Laboratories Record* 11 (December 1932):113.

Jensen, A. G. "New Portable Sets Measure Radio Transmission." *Bell Laboratories Record* 2 (June 1926):177.

Jensen, A. G. "Portable Receiving Sets for Measuring Field Strengths at Broadcasting Frequencies." *Proc. IRE* 14 (June 1926):333.

Roetken, A. A. "Measuring the Frequencies of Radio Signals." *Bell Laboratories Record* 9 (August 1931):585.

Chapter 6. Switching and Signaling Systems

Introduction and General Concepts

Events in Telephone History, 1970 to Date. New York: American Telephone and Telegraph Co., 1975.

Green, E. I. "Telephony." *Encyclopaedia Britannica* 21 (1957):895–906.

Holmes, E. T. *A Wonderful Fifty Years.* Privately printed, 1917.

Kingsbury, J. E. *The Telephone and Telephone Exchanges.* London: Longmans, Green, & Co., 1915.

Miller, K. B. *American Telephone Practice.* 4th ed. New York: McGraw–Hill Book Co., 1905.

Preece, W. H. and Maier, J. *The Telephone.* New York: D. Van Nostrand Co., 1889.

Reports of the Meetings of the National Telephone Exchange Assoc. (1880–1890).

Rhodes, F. L. *Beginnings of Telephony.* New York: Harper & Bros., 1929.

Robertson, J. H. *The Story of the Telephone: History of the Telecommunications Industry of Britain.* London: Isaac Pitman & Sons, 1947.

Smith, A. B. and Campbell, W. L. *Automatic Telephony.* New York: McGraw–Hill Book Co., 1914.

Statistical Reports. American Telephone and Telegraph Co. (1913–1945).

Stehman, J. W. *The Financial History of the American Telephone and Telegraph Company.* Boston: Houghton Mifflin Co., 1925.

Telephone and Telegraph Statistics of the World. New York: American Telephone and Telegraph Co., Chief Statistician's Division, 1912–1926 (Annual Issues).

"Telephone Service." *Circular of the Bureau of Standards, No. 112*. Washington: Government Printing Office, 1921.

Western Electric Co. Catalogs, Bulletins, and Instruction Manuals (1890–1930).

Evolution of Manual Exchange Switching

Barrett, R. T. "The Changing Years as Seen from the Switchboard." *Bell Telephone Quart*. 10 (January–October, 1935).

Bennetto, J. "New Haven Pioneer Exchange, Marks 60th Birthday." *The Telephone Bulletin* 31 (January 1938).

Bouman, B. M. "Early Manual Switchboard Development." *Bell Laboratories Record* 9 (February 1931):265.

Bronson, F. M. "Tandem Operation in the Bell System." *Bell System Technical J*. 15 (July 1936):380.

"Central Battery versus Local Battery Operation." Western Electric Co. Bulletin No. 1017 (August 1911).

Collins, C. A. and Schreiner, L. R. *Telephone Practice*. Bell Laboratories Out-of-Hours text, Course 160 (1937–1938):Chap.9.

DeLand, F. "Notes on the Development of Telephone Service." *Popular Science Monthly* (November 1906):425; (December 1906):490; (January 1907):48; (March 1907):229; (April 1907):317; (May 1907):404; (June 1907):518; (July 1907):21; (August 1907):139; (September 1907):227; (October 1907):313; (November 1907):433.

Findley, P. B. "Our Historical Museum." *Bell Laboratories Record* 1 (December 1925):137.

Holbrook, B. D. "Early Switching History." Transcript of interview at Murray Hill, N. J. (September 9, 1971).

Holmes, E. T. "Notes on Early Telephone Work." Handwritten document prepared for T. D. Lockwood, General Inspector of the American Bell Telephone Co. (1883).

James, F. S. and Farnham, W. E. "Operating Features of the Straightforward Trunking Method." *Bell Telephone Quart*. 9 (January 1930):25.

Jewett, F. B. "The Telephone Switchboard—Fifty Years of History." *Bell Telephone Quart*. 7 (July 1928):149.

Kingsbury, J. E. "The Origin and Development of the Telephone Switchboard." *Proc. IEE* 24 (1895):36.

Lewis, F. P. "The Pioneer Exchange in New Haven." *The Telephone Bulletin* 4 (February 1911). (Manuscript in American Telephone and Telegraph Co. Historical Files.)

Lockwood, T. D. "The Evolution of the Telephone Switchboard." *AIEE Trans*. 21 (January–June, 1903):127.

Mead, E. D. "Saving By Swaging." *Bell Laboratories Record* 3 (October 1926):83.

Neill, P. "Fifty Years of Telephone Plugs." *Bell Laboratories Record* 5 (December 1927):104.

Richards, W. L. "Operation of New Haven Switchboard Model." Unpublished document, American Telephone and Telegraph Historical Files (1927).

Ritchie, G. A. "Lever-Type Keys." *Bell Laboratories Record* 7 (September 1928):56.

Scribner, C. E. "The Genesis of the Telephone Switchboard." *Proc. of the Sixth Annual Convention, Telephone Pioneers of America*, October 27–28, 1916, at Atlanta, Georgia.

Shackleton, S. P. and Purcell, H. W. "Relays in the Bell System." *Bell System Technical J.* 3 (January 1924):1.

Spencer, C. G. "The Manual Tandem Board." *Bell Laboratories Record* 8 (January 1930):233.

Thompson, G. K. and Hill, R. B. "The First Telephone Switchboard and Its Method of Operation." *Bell Telephone Quart.* 9 (July 1930):205.

"Untechnical Talks on Technical Topics." ("Keys.") *Western Electric News* 7 (August 1918):7.

"Untechnical Talks on Technical Topics." ("The Multiple Telephone Switchboard.") *Western Electric News* 1 (June 1912):20.

"Untechnical Talks on Technical Topics." ("Relays, Drops, and Signals.") *Western Electric News* 3 (December 1914):26.

Walsh, J. L. *Connecticut Pioneers in Telephony*. New Haven: Privately printed, 1950.

Willenbruch, W. "Switchboard Lamps." Unpublished manuscript, March 1938.

Trunking and Switching Theory

Crowley, T. H.; Harris, G. G.; Miller, S. E.; Pierce, J. R.; and Runyon, J. P. *Modern Communications*. New York: Columbia University Press, 1962, Chap. 10.

Hayward, W. S. "Traffic Engineering." Internal communication, Bell Laboratories (May 14, 1974).

Molina, E. C. "Application of the Theory of Probability to Telephone Trunking Problems." *Bell System Technical J.* 6 (July 1927):461.

Molina, E. C. "Theory of Probability and Some Applications to Engineering Problems." *AIEE Trans.* 44 (February 1925):294.

Rorty, M. C. "Application of the Theory of Probability to Traffic Problems." Unpublished manuscript, American Telephone and Telegraph Co. (October 22, 1903).

Wilkinson, R. I. "The Beginnings of Switching Theory in the United States." *Electrical Engineering* 75 (September 1956):796.

Automatic Switching

"Automatic Exchange Apparatus." Circular Letter No. 387, Western Electric Co. (May 23, 1904).

"Automatic Switchboards for Small Exchanges." Circular No. 25, American Telephone and Telegraph Co. (May 21, 1924).

Carty, J. J. "Telephone Service in America." Paper presented at the International Conference of European Telephone and Telegraph Administrations, September 4–11, 1910, at Paris.

Clark, E. H. "Evolution of the Call-Indicator System." *Bell Laboratories Record* 8 (December 1929):171.

Craft, E. B.; Morehouse, L. F.; and Charlesworth, H. P. "Machine Switching Telephone System for Large Metropolitan Areas." *Bell System Technical J.* 2 (April 1923):53.

Curran, S. T. "Key-Display Type Call Indicators." *Bell Laboratories Record* 8 (July 1930):515.

"Description of the Village System of 1895." Unpublished notes prepared from contemporary records of 1884 to 1902, Bell Telephone Laboratories (1947).

Engineering Considerations of the Bell System. Training material used by Bell Telephone Laboratories (1967): chaps. 5 and 6.

Feder, H. S. and Spencer, A. E. "Telephone Switching." *Sci. Amer.* 207 (July 1962):132.

Freeman, W. G. "History of the Development of the Panel Machine Switching System." Unpublished document, Bell Telephone Laboratories (1937).

Gherardi, B. and Charlesworth, H. P. "Machine Switching for the Bell System." Supplement to the *Telephone Review* (April 1920).

Goldsmith, E. H. "The Panel Type Dial Telephone System." Training material used by the New York Telephone Co. (December 1926).

Hill, R. B. "Early Work on Dial Telephone Systems." *Bell Laboratories Record* 31 (January 1953):22.

Hill, R. B. "The Early Years of the Strowger System." *Bell Laboratories Record* 31 (March 1953):95.

Hill, R. B. "History of Dial Switching in the Bell System." Unpublished notes, Bell Laboratories (1945 and 1947).

"Important Improvements in Automatic Exchange Circuits and Apparatus." Information Bulletin No. 30, Western Electric Engineering Dept. (December 1, 1905).

Joel, A. E., Jr. "A Chronology of Bell System Automatic Switching System Developments." Internal memorandum, Bell Laboratories (May 18, 1971).

Keister, W. "Switching History." Transcript of interview at Murray Hill, N. J. (April 1, 1969).

Laidlaw, E. A. "Facts or Opinions." *P. O. Elec. Eng. J.* 5 (July 1912):207.

Machine Switching Equipment Notes, Vol. 2. Training material, Western Electric Engineering Dept. (1920).

McDavitt, M. B. "History, Development, and Application of Machine Switching in the Bell System." Unpublished document (January 6, 1949).

Myers, O. "Common Control Telephone Switching Systems." *Bell System Technical J.* 31 (November 1952):1086.

Notes on Tandem Telephone Switching Systems. Training material, Bell Telephone Laboratories (1973).

"Survey of Telephone Switching." Training material prepared by the Pacific Telephone and Telegraph Co. (November 1956).

Telephone Switching Systems in the Bell System. Training material used by Bell Telephone Laboratories (August 1971).

Washburn, S. H. "An Instance of Automation: The Mechanization of Telephone Switching." *1967 IEEE Convention Record,* Part II:158.

Toll Switching

Breen, C. and Dahlbom, C. A. "Signaling Systems for Control of Telephone Switching." *Bell System Technical J.* 34 (November 1960):1381.

"Calculagraphs, Electrically Driven." Bell System Practices, Section 030-101-101, Issue 1, American Telephone and Telegraph Co. (November 1935):1.

"Calculagraphs, Spring-Driven." Bell System Practices, Section 030-100-101, Issue 1, American Telephone and Telegraph Co. (September 1926):1.

Campbell, W. L. "A Study of Multi-Office Automatic Switchboard Telephone Systems." *AIEE Trans.* 27 (1908):503.

Carty, J. J. "The Telephone's Development." *Bell Telephone Quart.* 1 (April 1922):23.

Clark, A. B. "Development of Telephony in the United States." *AIEE Trans.* 71 (November 1952):350.

Davidson, J. "Toll Switchboard No. 3." *Bell System Technical J.* 6 (January 1927):18.

"Distributor Signalling." Bulletin No. 162, American Tlephone and Telegraph Co. (December 19, 1924):25.

Fundamentals of Switching. BTL Training Course 330 (1948): chaps. 8 and 11.

Gherardi, B. and Jewett, F. B. "Telephone Communication System of the United States." *Bell System Technical J.* 9 (January 1930):1.

Hamilton, B. P. "Single-Line Telegraphy in Early Bell System Practice." *Bell Laboratories Record* 23 (November 1945):419.

Harrison, W. H. "Recent Developments in Toll Telephone Service." *Bell Telephone Quart.* 9 (April 1930):124.

"Henry Abbott—Father of the Calculagraph." *Telephony* 124 (May 22, 1943):24.

"Improved Telephone Signals and Telegraph Arrangements." Memorandum for Mr. Gherardi, American Telephone and Telegraph Co. (January 26, 1918).

James, F. S. and Farnham, W. E. "Operating Features of the Straightforward Trunking Method." *Bell Telephone Quart.* 9 (January 1930):25.

Jewett, F. B. "Electrical Communications: Past, Present, Future," *Bell Telephone Quart.* 14 (July 1935):167.

"KS-14156 Switchboard Position Clocks." Bell System Practices, Section 030-114-801, Issue 1, American Telephone and Telegraph Co. (February 1953):1.

"1000 Cycles Signaling System." Bulletin No. 170, American Telephone and Telegraph Co. (November 7, 1924).

Osborne, H. S. "A General Switching Plan for Toll Service." *Bell System Technical J.* 9 (July 1930):429.

Osborne, H. S. "Technical Developments Underlying the Toll Services of the Bell System." *Bell System Technical J.* 15, Supplement (January 1936):2.

Pilliod, J. J. "Engineering the Long Lines." *Bell Telephone Quart.* 2 (January 1923):22.

Porter, L. F. "A New Type of Toll Switchboard." *Bell Laboratories Record* 4 (June 1927):337.

"Switchboard Position Clocks." Bell System Practices, Section 030-114-811, Issue 1, American Telephone and Telegraph Co. (January 1953):1.

"Telephone Signalling." *Bell Laboratories Record* 1 (February 1926):241.

"Terminal Ringer." D&R Bulletin No. 133, American Telephone and Telegraph Co. (March 21, 1924):1.

Waterson, K. W. "Functions and Management Problems of the Traffic Department." *Bell Telephone Quart.* 5 (October 1926):203.

Waterson, K. W. "Improvement in Toll Service through Combined Line and Recording Method." *Bell Telephone Quart.* 5 (April 1926):71.

Private Branch Exchange (PBX) Switching

Abbott, H. H. "Development of Dial PBX's." Internal memorandum, Bell Laboratories (April 10, 1935).

Abbott, H. H. "60 Years of PBX Development." *Bell Laboratories Record* 46 (January 1968):8.

"Automatic Exchanges." General Engineering Circular No. 25, American Telephone and Telegraph Co. (May 21, 1904).

Circular Letter No. 511, Western Electric Co. (June 21, 1907).

Events in Telephone History, 1790 to Date. New York: American Telephone and Telegraph Co., 1975, p. 9.

Ferguson, J. G. "Recollections of Dial PBX and C. D. O. Developments." Internal memorandum, Bell Laboratories (November 21, 1960).

"551-A and 551-B Non-Multiple Private Branch Exchange." General Engineering Circular No. 982, American Telephone and Telegraph Co. (July 1, 1920).

Gherardi, B. "Machine Switching Systems." Internal memorandum, American Telephone and Telegraph Co. (July 24, 1917).

"The Growth of the Private Branch Telephone Exchange Service in New York City." *Electrical Engineer* 24 (November 25, 1897):497.

Harper, R. W. "Private Branch Exchanges." *Bell Laboratories Record* 7 (February 1929):226.

Lockwood, T. D. "Ten Years Progress in Telephony." *Western Electrician* 1 (October 1, 1887):165.

"Machine Switching Private Branch Exchange No. 700-C." General Engineering Circular No. 1302, American Telephone and Telegraph Co. (September 25, 1923).

"Multiple Private Branch Exchange No. 600-C." General Engineering Circular No. 1232, American Telephone and Telegraph Co. (January 23, 1923).

"No. 4 Private Branch Exchange Boards." Bulletin No. 1018, Western Electric Co. (1911).

"No 505-C Non-Multiple Cordless Private Branch Exchange for Operation with Machine-Switching Central Offices." General Engineering Circular No. 1081, American Telephone and Telegraph Co. (August 29, 1921).

"No. 604-C Private Branch Exchange." Letter to Chief Engineers, American Telephone and Telegraph Co. (November 29, 1926).

"No. 605-A Private Branch Exchange." Plant Engineering Circular No. 380, American Telephone and Telegraph Co. (July 5, 1929).

"Operating Methods." Engineering Department Bulletin No. 11, American Telephone and Telegraph Co. (February 1904).

"Private Branch Exchange No. 2." Engineering Department Bulletin No. 9, American Telephone and Telegraph Co. (October 1903).

"Private Branch Exchange Switchboard No. 550-B." General Engineering Circular No. 665, American Telephone and Telegraph Co. (September 18, 1916).

"Private Branch Exchanges." Engineering Department Bulletin No. 1, American Telephone and Telegraph Co. (July 1902).

"Re-Designed Cordless Private Branch Exchange." General Engineering Circular No. 577, American Telephone and Telegraph Co. (January 20, 1916).

Rhodes, F. L. *Beginnings of Telephony.* New York: Harper & Bros., 1929, p. 9.

"Sales of Automatic Exchanges and Notes on Recent Improvements." Information Bulletin No. 30, Western Electric Co. (December 1, 1905).

"Telephone Systems for Large Buildings." *Western Electrician* 1 (July 2, 1887):1.

Central-Office Power Systems

Callahan, V. T. "A New Emergency Power Supply Unit." *Bell Laboratories Record* 8 (December 1929):158.

Callahan, V. T. "New Standards in Emergency Power-Supply Units." *Bell Laboratories Record* 6 (June 1928):318.

DeKay, R. D. "Power for Magneto Offices. *Bell Laboratories Record* 9 (December 1930):180.

Duguid, J. M. "Historical Review of Bell System Power Development." Internal memorandum, Bell Telephone Laboratories (March 3, 1961).

Duguid, J. M. "Milestones in Telephone Power Plant Development." Internal memorandum, Bell Telephone Laboratories (February 28, 1957).

Martin, R. P., Jr. "Talking Battery." *Bell Laboratories Record* 15 (January 1937):138.

Van Duyne, C. W. "Insuring Central-Office Power Supply." *Bell Laboratories Record* 7 (September 1928):21.

Young, R. L. "Power Plants for Telephone Offices." *Bell System Technical J.* 6 (October 1927):702.

Young, R. L. and Lunsford, R. L. "Automatic Power Plant for Telephone Offices." *AIEE Trans.* 50 (September 1931):949.

Chapter 7. Non-Voice Communications

General

Bennett, W. R. and Davey, J. R. *Data Transmission.* New York: McGraw–Hill Book Co., 1965.

Brown, F. J. *Cable and Wireless Communications of the World*. 2nd ed. London: Isaac Pitman & Sons, 1930.

Crotel, A. *Elementary Telegraphy and Telephony*. London: E. and F. N. Spon, 1903.

Everitt, W. L. *Communication Engineering*. New York: McGraw–Hill Book Co., 1932.

Gherardi, B. and Jewett, F. B. "Telephone Communication System of the United States." *Bell System Technical J.* 9 (January 1930):1.

Harrison, H. H. *Printing Telegraph Systems and Mechanisms*. New York: Longmans, Green & Co., 1923.

Hausmann, E. *Telegraph Engineering and Manual for Practicing Engineers and Engineering Students*. 2nd ed. New York: D. Van Nostrand Co., 1922.

Herbert, T. E. *Telegraphy–A Detailed Exposition of the Telegraph System of the British Post Office*. 5th ed. London: Isaac Pitman & Sons, 1930.

Jewett, F. B. "Electrical Communications, Past, Present, Future." *Bell Telephone Quart.* 14 (July 1935):167.

Jewett, F. B. "Utilizing the Results of Fundamental Research in the Communication Field." *Bell Telephone Quart.* 11 (April 1932):143.

Maver, W., Jr. *American Telegraphy and Encyclopedia of the Telegraph*. New York: Maver Publishing Co., 1912.

Notes on Telegraph Engineering. New York: American Telephone and Telegraph Co., 1929.

Still, A. *Communication Through the Ages, From Sign Language to Television*. New York: Holt, Rinehart & Winston, 1946.

Stone, A. E. *Text Book of Telegraphy*. London: Macmillan & Co., 1928.

Williams, A. *Telegraphy and Telephony*. London: Thomas Nelson & Sons, 1928.

Telegraph Within the Bell System:

General

Bell, J. H. "Composite Telegraphy and Telephony." *J. Post Office Electrical Engineers* 15 (April 1922):1.

Bell, J. H. "The KR Law as Applied to Quadruplex Working." *Telegraph and Telephone Age* (May 16, 1912):313.

Bell, J. H. "Printing Telegraph Systems." *AIEE Trans.* 39 (1920):167.

Bell, J. H. "Western Electric Company Multiplex Telegraph System." *Telegraph and Telephone J.* 1 (March 1915):118.

Bell, J. H.; Shanck, R. B.; and Branson, D. E. "Metallic Polar-Duplex Telegraph System for Long Small-Gauge Cables." *AIEE Trans.* 44 (April 1925):378.

Colpitts, E. H. and Blackwell, O. B. "Carrier Current Telephony and Telegraphy." *AIEE Trans.* 40 (1921):205.

Duncan, J. A.; Parker, R. D.; and Pierce, R. E. "Telegraphy in the Bell System." *AIEE Trans.* 63 (June 1944):1032.

Hamilton, B. P. "Early Bell System Polar Telegraphy." *Bell Laboratories Record* 24 (June 1946):224.

Hamilton, B. P. "Single-Line Telegraphy in Early Bell System Practice." *Bell Laboratories Record* 23 (November 1945):419.

Jewett, F. B. "Enter Multiplex Telephony and Telegraphy." *Western Electric News* 7 (January 1919):1.

McMurry, F. R. "Telegraphy by Typewriter." *Bell Laboratories Record* 3 (September 1926):3.

Meyer, D. C. "New Telegraph Equipment." *Bell Laboratories Record* 2 (May 1926):103.

Parker, R. D. "Telegraphy." *Nelson's Encyclopedia* (October 1922).

Parker, R. D. "Telephone Typewriters and Auxiliary Arrangements." *Bell Telephone Quart.* (July 1929):181.

Pierce, R. E. "Communication by Police Teletypewriters." *Electrical Engineering* 51 (August 1932):550.

Pierce, R. E. "Modern Practices in Private Wire Telegraph Service." *AIEE Trans.* 50 (1931):426.

Rainey, P. M. "A New Printing Telegraph System." *The Electrical World* 65 (April 3, 1915):848.

Vernam, G. S. "Cipher Printing Telegraph Systems for Secret Wire and Radio Telegraphic Communications." *AIEE Trans.* 45 (February 1926):109.

Watson, E. F. "Fundamentals of Teletypewriters Used in the Bell System." *Bell System Technical J.* 17 (October 1938):620.

Transmission Theory

Carson, J. R. "Theory of the Transient Oscillations of Electrical Networks and Transmission Systems." *AIEE Trans.* 38 (1919):345.

Herman, J. "Effect of Signal Distortion on Morse Telegraph Transmission Quality." *Bell System Technical J.* 8 (April 1929):267.

Jewett, F. B. "Some Recent Developments in Telephony and Telegraphy." *Proc. Pan-American Scientific Congress*, 1915–1916.

Nyquist, H. "Certain Factors Affecting Telegraph Speed." *Bell System Technical J.* 3 (April 1924):324.

Nyquist, H. "Certain Topics in Telegraph Transmission." *AIEE Trans.* 47 (March 1928):214.

Nyquist, H. "Certain Topics in Telegraph Transmission Theory." *AIEE Trans.* 47 (February 1928):617.

Submarine-Cable and Radio Telegraphy

Bailey, A. and McCann, T. A. "Application of the Printing Telegraph to Long-Wave Radio Circuits." *Proc. IRE* 19 (December 1931):2177.

Buckley, O. E. "High-Speed Ocean Cable Telegraphy." *Bell System Technical J.* 7 (April 1928):225.

Buckley, O. E. "The Loaded Submarine Telegraph Cable." *Bell System Technical J.* 4 (July 1925):355.

Carson, J. R. and Gilbert, J. J. "The Sea Return Impedance of a Submarine Cable." *The Electrician* 88 (April 28, 1922):499.

Carson, J. R. and Gilbert, J. J. "Transmission Characteristics of the Submarine Cable." *J. of Franklin Institute* 192 (December 1921):705.

Clokey, A. A. "Automatic Printing Equipment for Long Loaded Submarine Telegraph Cables." *Bell System Technical J.* 6 (July 1927):402.

Curtis, A. M. "The Application of Vacuum Tube Amplifiers to Submarine Telegraph Cables." *Bell System Technical J.* 6 (July 1927):425.

Curtis, A. M. " 'Signal Shaping' for Submarine Cables." *Bell Laboratories Record* 6 (April 1928):237.

Gilbert, J. J. "Balanced Sea Earth for Cables." *The Electrician* 97 (August 6, 1926): 152.

Gilbert, J. J. "Determination of Electrical Characteristics of Loaded Telegraph Cables." *Bell System Technical J.* 6 (July 1927):387.

Gilbert, J. J. "Extraneous Interference on Submarine Telegraph Cables." *Bell System Technical J.* 5 (July 1926):404.

Heising, R. A. "Printing Telegraph by Radio." *J. of Franklin Institute* 193 (January 1922):97.

Repeaters

Bell, J. H. "Cross-Fire Neutralization of Telegraph Circuits." *Bell Laboratories Record* 9 (December 1930):171.

Bell, J. H. "A Regenerative Telegraph Repeater." *Bell Laboratories Record* 8 (August 1930):570.

Cummings, G. C. "The Single-Line Telegraph Repeater." *Bell Laboratories Record* 10 (October 1931):39.

Fry, J. R. and Gardner, L. A. "Polarized Telegraph Relays." *AIEE Trans.* 44 (March 1925):223.

Shanck, R. B. "Neutralization of Telegraph Crossfire." *Bell System Technical J.* 5 (July 1926):418.

Spenser, H. H. "Telegraph Ground-Potential Compensator." *Bell Laboratories Record* 9 (February 1931):271.

Wilburn, S. D. "Line Current Regulation in Bridge Polar Duplex Telegraph Circuits." *Bell System Technical J.* 5 (October 1926):625.

Voice-Frequency Carrier

Bell, J. H. "Carrier-Current Telegraphy." *Bell Laboratories Record* 1 (January 1926):187.

Hamilton, B. P. "Carrier Telegraphy in the Bell System." *Bell Laboratories Record* 26 (February 1948):58.

Hamilton, B. P.; Nyquist, H.; Long, M. B.; and Phelps, W. A. "Voice-Frequency Carrier Telegraph System for Cables." *AIEE Trans.* 44 (March 1925): 213.

Mahoney, J. A. "New Equipment for Voice-Frequency Telegraphy." *Bell Laboratories Record* 7 (February 1929):241.

Testing and Maintenance

Osborne, H. S. "Detection and Location of Trouble on Telephone and Telegraph Circuits." *Proc. American Railway Assoc.* (September 1924).

Wier, A. J. "Telegraph Testing Facilities." *Bell Laboratories Record* 14 (October 1935):45.

Telephotography and Television

Gannett, D. K. and Green, E. I. "Wire Transmission System for Television." *Bell System Technical J.* 6 (October 1927):616.

Gray, F.; Horton, J. W.; and Mathes, R. C. "The Production and Utilization of Television Signals." *Bell System Technical J.* 6 (October 1927):560.

Harkness, W. E. "Commercial Use of Telephotographs." *Bell Telephone Quart.* 6 (July 1927):171.

Ives, H. E. "Television." *Bell System Technical J.* 6 (October 1927):551.

Ives, H. E.; Horton, J. W.; Parker, R. D.; and Clark, A. B. "The Transmission of Pictures over Telephone Lines." *Bell System Technical J.* 4 (April 1925):187.

Nelson, E. L. "Radio Transmission System for Television." *Bell System Technical J.* 6 (October 1927):633.

O'Brien, J. B. "The Transmission of Pictures over Telephone Wires." *Bell Telephone Quart.* 3 (July 1924):186.

Stoller, H. M. "Synchronization System for Two-Way Television." *Bell System Technical J.* 9 (July 1930):470.

Stoller, H. M. and Morton, E. R. "Synchronization of Television." *Bell System Technical J.* 6 (October 1927):604.

Chapter 8. Materials and Components

Magnetic Materials Development

Arnold, H. D. and Elmen, G. W. "Permalloy, An Alloy of Remarkable Magnetic Properties." *J. of Franklin Institute* 195 (May 1923):621.

Buckley, O. E. "High Speed Ocean Cable Telegraphy." *Bell System Technical J.* 7 (April 1928):225.

Buckley, O. E. "The Loaded Submarine Telegraph Cable." *Bell System Technical J.* 4 (July 1925):355.

Buckley, O. E. "Transoceanic Telephone Cables." *Bell System Technical J.* 20 (June 1942):238.

Elmen, G. W. "New Permalloys." *Bell Laboratories Record* 10 (September 1931):2.

Fondiller, W. "Developments in Communication Materials." *Bell System Technical J.* 9 (April 1930):237.

Fondiller, W. "A New Era in Loading." *Bell Laboratories Record* 5 (September 1927):1.

Hill, R. B. "Birth of the Loading Coil." *Bell Laboratories Record* 30 (June 1952):263.

Legg, V. E. "Magnetic Measurements at Low Flux Densities Using the Alternating Current Bridge." *Bell System Technical J.* 15 (January 1936):39.

Legg, V. E. and Given, F. J. "Compressed Powdered Molybdenum Permalloy for High Quality Inductance Coils." *Bell System Technical J.* 19 (July 1940):385.

Martin, W. H.; Anderegg, G. A.; and Kendall, B. W. "Key West–Havana Submarine Telephone Cable System." *AIEE Trans.* 16 (February 1922):1.

"The New York–Azores Cable." *Bell Laboratories Record* 1 (November 1925): 116.

Peterson, E. and Wrathall, L. R. "Eddy Currents in Composite Laminations." *Proc. IRE* 24 (February 1936):275.

Schwartz, E. L. "Permalloy in Audio Transformers." *Bell Laboratories Record* 6 (April 1928):259.

Shackelton, W. J. and Barber, I. G. "Compressed Powdered Permalloy, Manufacture and Magnetic Properties." *AIEE Trans.* 47 (April 1928):429.

Shaw, T. "Inductive Loading for Bell System Telephone Circuits." *Bell System Technical J.* 30 (1951):149.

Shaw, T. and Fondiller, W. "Development and Application of Loading for Telephone Circuits." *Bell System Technical J.* 5 (April 1926):221.

Speed, B. and Elmen, G. W. "Magnetic Properties of Compressed Powdered Iron." *AIEE Trans.* 40 (June 1921):1321.

Communications Transformers

Casper, N. L. "Telephone Transformers." *AIEE Trans.* 43 (March 1924):197.

Colpitts, E. H. and Blackwell, O. B. "Carrier Current Telephony and Telegraphy." *AIEE Trans.* 40 (1921):205.

Field, F. E. "Evolution of the Input Transformer." *Bell Laboratories Record* 3 (October 1926):33.

Hanscom, C. D. *Dates in American Telephone Technology.* New York: Bell Telephone Laboratories, 1961, p. 25.

Hitchcock, H. W. "Carrier-Current Communication on Submarine Cables." *Bell System Technical J.* 5 (1926):636.

Johnson, K. S. *Transmission Circuits for Telephonic Communication.* New York: D. Van Nostrand Co., 1927, Chaps. 6, 7, and 8.

Miller, K. B. *American Telephone Practice.* 4th ed. New York: McGraw–Hill Book Co., 1905.

Rhodes, F. L. *Beginnings of Telephony.* New York: Harper & Bros., 1929, pp. 152, 166, 171, and 179.

Shaw, T. "The Conquest of Distance by Wire Telephony." *Bell System Technical J.* 23 (October 1944):337.

Smith, C. A. "Fifty Years of Telephone Repeaters." *Bell Laboratories Record* 27 (January 1949):5.

"Technical Developments Underlying the Toll Services of the Bell System." *Bell System Technical J.*, Supplement (July 1936):11.

Inductors (Retardation Coils)

Darnell, P. S. "Retardation Coils for Precision Filters." *Bell Laboratories Record* 12 (August 1934):375.

Kahl, W. E. "Adjusting Precision Filters." *Bell Laboratories Record* 12 (August 1934):379.

Capacitors

Coursey, P. R. *Electrical Condensers.* London: Isaac Pitman & Sons, 1927, pp. 412–471.

Froberg, M. A. "Commercial Generator for Central Office Power Plants." *Bell Laboratories Record* 5 (December 1927):113.

Kingsbury, J. E. *The Telephone and Telephone Exchanges.* London: Longmans, Green and Co., 1915, p. 457.

Mansbridge, G. F. "The Manufacture of Electrical Condensers." *J. of IEE* 41 (1908):535.

Prescott, G. B. *Bell's Electric Speaking Telephone*. New York: D. Appleton & Co., 1884, p. 23.

Siegmund, H. O. "The Aluminum Electrolytic Condenser." *Bell System Technical J.* 8 (January 1929):41.

Wheatstone, C. "On the Position of Aluminum in the Voltaic Series." *Philosophical Mag.* 10 (1855):143.

Vacuum Tubes

Arnold, H. D. "Phenomena in Oxide-Coated Filament Electron Tubes." *Phys. Rev.* 16 (July 1920):70.

Gannett, D. K. "The Vacuum Tube in Telephone Communications." *Bell Telephone Quart.* 12 (October 1933):240.

Heising, R. A. "The Audion Oscillator." *Phys. Rev.* 16 (September 1920):216.

Houskeeper, W. G. "The Art of Sealing Base Metals Through Glass." *AIEE Trans.* 42 (June 1923):870.

King, R. W. "Calculation of the Constants of the Three-Electrode Thermionic Vacuum Tube." *Phys. Rev.* 15 (April 1920):256.

King, R. W. "Thermionic Vacuum Tubes and Their Applications." *Bell System Technical J.* 2 (October 1923):31.

Llewellyn, F. B. "Birth of the Electron Tube Amplifier." *Radio and Television News* 57 (March 1957):43.

Nichols, H. W. "The Audion as a Circuit Element." *Phys. Rev.* 13 (June 1919):404.

Tyne, G. F. J. "The Saga of the Vacuum Tube." *Radio News* 29 (March 1943) through 35 (April 1946).

Van der Bijl, H. J. "Theory of the Thermionic Amplifier." *Phys. Rev.* 12 (September 1918):171.

Van der Bijl, H. J. *The Thermionic Vacuum Tube and Its Applications*. New York: McGraw–Hill Book Co., 1920.

Wilson, W. "A New Type of High Power Vacuum Tube." *Bell System Technical J.* 1 (July 1922):4.

Chapter 9. Quality Assurance

Annual Report. Western Electric Co., 1922.

Jevons, W. S. *The Principles of Science*. London: Macmillan & Co., 1907.

Jones, R. L. "The Viewpoint of Inspection Engineering." *Bell Laboratories Record* 2 (August 1926):241.

Olmstead, P. "Our Debt to Walter Shewhart." *Industrial Quality Control* (Shewhart Memorial Issue) 24 (August 1967):72. (In particular, note the illustration, "The First Control Chart," which precedes the article.)

Shewhart, W. A. "The Application of Statistics as an Aid in Maintaining Quality of a Manufactured Product." *J. of American Statistical Assn*. 20 (December 1925):546.

Shewhart, W. A. *The Economic Control of Quality of Manufactured Product*. New York: Van Nostrand & Co., 1931, Chap. 4.

Shewhart, W. A. "Quality Control Charts." *Bell System Technical J*. 5 (October 1926):593.

Shewhart, W. A. "Some Applications of Statistical Methods to the Analysis of Physical and Engineering Data." *Bell System Technical J*. 3 (January 1924):43.

Wilks, S. S. *Mathematical Statistics*. New York: John Wiley & Sons, 1962.

Chapter 10. The Spirit of Research

Arnold, H. D. "Systematized Research." *Bell Laboratories Record* 6 (June 1928):313.

Bello, F. "The World's Greatest Industrial Laboratory." *Fortune* 58 (November 1958):148; (December 1958):117.

Bown, R. "Vitality of a Research Institution and How to Maintain It." *Proc. Sixth Annual Conf. Admin. Res.* (1953):31.

Findley, P. B. "The Research Department." *Bell Laboratories Record* 2 (June 1926):164.

Fisk, J. B. "Strategy in Industrial Research." *Research Management* 6 (September 1963):325.

Jewett, F. B. "Elements of Industrial Research." *Chemical Age* 1 (December 25, 1919):286.

Kelly, M. J. "Bell Telephone Laboratories—An Example of an Institute of Creative Technology." *Proc. Royal Soc., A* 203 (October 10, 1950):287.

Williams, R. R. "Chemistry in the Telephone Industry." *Industrial and Engineering Chemistry* 22 (April 1930):316.

Credits

Acknowledgment is made to the following for their kind permission to reprint copyrighted material:

Figures 3-7, 4-17 through 4-19, 4-21, 4-24, 4-25, 6-3, 6-5 through 6-13, and 6-16 from Frederick Leland Rhodes, *The Beginnings of Telephony*. Copyright 1929 by Harper and Row, Publishers, Inc.; renewed 1957 by Eleanor A. Rhodes. Reprinted by permission of the publisher.

Figure 3-94a from Frederick V. Hunt, *Electroacoustics*. Copyright 1954 by Harvard University Press. Reprinted with permission.

Figure 4-8 from J. W. Emling, "The Conquest of Distance." *IRE Student Quarterly*, May 1962. Copyright 1962 by The Institute of Electrical and Electronic Engineers, Inc. Reprinted with permission.

Figure 4-54 from E. H. Colpitts and O. B. Blackwell, "Carrier Current Telephony and Telegraphy." *AIEE Transactions*, Vol. 40 (1921). Copyright 1921 by The Institute of Electrical and Electronic Engineers, Inc. Reprinted with permission.

Figure 4-69 from F. A. Cowan, "Telephone Circuits for Program Transmission." *AIEE Transactions*, July 1929. Copyright 1929 by The Institute of Electrical and Electronic Engineers, Inc. Reprinted with permission.

Figures 5-27 through 5-29 from H. D. Arnold and L. Espenschied, "Transatlantic Radio Telephony." *Journal of the AIEE*, August 1923. Figures 5-63 and 10-14 from R. Bown, et al., "Radio Transmission Measurements." *Proceedings of the IRE*, April 1923. Figures 5-66 and 5-67 from H. W. Nichols and L. Espenschied, "Radio Extension of the Telephone System to Ships at Sea." *Proceedings of the IRE*, June 1923. Copyright 1923 by The Institute of Electrical and Electronic Engineers, Inc. Reprinted with permission.

Figure 5-30 from Arthur W. Page, et al., *Modern Communication*. Copyright 1932 by Houghton Mifflin Company. Reprinted with permission.

Figures 5-52 and 5-53 from E. L. Nelson, "Transmitting Equipment for Radio Telephone Broadcasting." *Proceedings of the IRE*, October 1924. Figures 5-59 and 5-60 from R. Bown and G. D. Gillett, "Distribution of Radio Waves from Broadcasting Stations Over City Districts." *Proceedings of the IRE*, August 1924. Copyright 1924 by The Institute of Electrical and Electronic Engineers, Inc. Reprinted with permission.

Figure 5-64 from A. G. Jensen, "Portable Receiving Sets for Measuring Field Strengths at Broadcasting Frequencies." *Proceedings of the IRE*, June 1926. Figure 5-65 from G. D. Gillett, "Portable Field Strength Measuring Equipment." *Proceedings of the IRE*, May 1926. Copyright 1926 by The Institute of Electrical and Electronic Engineers, Inc. Reprinted with permission.

Figure 6-45a from H. S. Feder and A. E. Spencer, "Telephone Switching." Copyright © 1962 by Scientific American, Inc. All rights reserved.

Figure 6-58 from H. S. Feder and A. E. Spencer, "Telephone Switching." *Scientific American*, July 1962. Copyright 1962 by David Linton. Reprinted with permission.

Figure 7-2 from Alfred Still, *Communication Through the Ages*. Copyright 1946 by Holt, Rinehart and Winston, Inc. Reprinted with permission.

Figure 10-11 from J. B. Johnson, *Physical Review*, Vol. 26 (1925). Copyright 1925 by the American Institute of Physics. Reprinted with permission.

Figure 10-12 from L. Espenschied, "The Origin and Development of Radiotelephony." *Proceedings of the IRE*, September 1937. Copyright 1937 by The Institute of Electrical and Electronic Engineers, Inc. Reprinted with permission.

Figure 10-20 from H. Fletcher and R. L. Wegel, "The Frequency-Sensitivity of Normal Ears." *Physical Review*, June 1922. Copyright 1922 by the American Institute of Physics. Reprinted with permission.

Figure 10-22 from H. Fletcher, "Physical Measurements of Audition and Their Bearing on the Theory of Hearing." *Journal of the Franklin Institute*, September 1923. Figures 10-37 through 10-39 from H. D. Arnold and G. W. Elmen, "Permalloy, an Alloy of Remarkable Magnetic Properties." *Journal of the Franklin Institute*, May 1923. Copyright 1923 by the Franklin Institute. Reprinted with permission.

Figure 10-45 from C. J. Davisson and L. H. Germer, "Diffraction of Electrons by a Crystal of Nickel." *Physical Review*, December 1927. Copyright 1927 by the American Institute of Physics. Reprinted with permission.

Figure 10-46 from L. H. Germer, "Optical Experiments with Electrons." *Journal of Chemical Education*, September 1928. Copyright 1928 by the American Chemical Society. Reprinted with permission.

Index