# THE BELL SYSTEM TECHNICAL JOURNAL

# THE BELL SYSTEM

# TECHNICAL JOURNAL

## Telephone

### By E. I. GREEN

*This article, which appeared in the 1957 printing of the Encyclopedia Britannica,\* has been reprinted by special permission for readers of the Bell System Technical Journal. All statistics have been corrected to the latest available figures.*

*An objective account of the invention of the telephone itself is given, and the subsequent development and growth of telephony are described. A statistical summary is presented of the intensity of development in different parts of the world. This is followed by a comprehensive review of technical developments, including progress in station instrumentalities, and transmission and switching principles and methods. The article concludes with a prediction of trends to be anticipated in telephony as a result of recent technical advances.*

The term "telephone" (from the Greek roots τῆλε, far, and φωνή, sound) was formerly used to describe any apparatus for conveying sounds to a distant point. Specifically, the word was applied as early as 1796 to a megaphone, and not long afterward to a speaking tube. Subsequently the name "string telephone" was given to the device invented long before by Robert Hooke (1667), in which vibrations in a diaphragm caused by voice or sound waves are transmitted mechanically along a string or wire to a similar diaphragm which reproduces the sound. Still later, devices employing electric currents to reproduce at a distance the mere pitch of musical sounds were called telephones. Nowadays, however, this name is assigned almost exclusively to apparatus for reproduc-

---

\* Copyright 1957 by Encyclopedia Britannica

ing articulate speech and other sounds at a distance through the medium of electric waves. The term "telephony" covers the entire art and practice of electrical speech transmission, including the many systems, accessories and operating methods used for this purpose.

INVENTION

Like most inventions, the telephone drew heavily upon previous work and had scarcely appeared before notable improvements were made. Among the pioneer contributors in this field, the most outstanding was Alexander Graham Bell, who invented, and patented in 1876, the first telephone capable of practical use. As early as 1874 he had conceived the correct principle of telephone transmission, which he later stated as follows: "If I could make a current of electricity vary in intensity precisely as the air varies in density during the production of sound, I should be able to transmit speech telegraphically."

This conception of an undulatory current corresponding to a speech wave formed the foundation for the entire telephonic art. Earlier workers, notably C. G. Page in the United States (1837) and Charles Bourseul in France (1854), had devised methods employing the make-and-break principle of the telegraph for transmitting the pitch of sounds, but not articulate speech, to a distant point.

Philipp Reis came closer. Working at Frankfurt, Germany, in 1860 and subsequent years, he devised an apparatus using at the transmitting end a diaphragm structure something like that of the human ear to
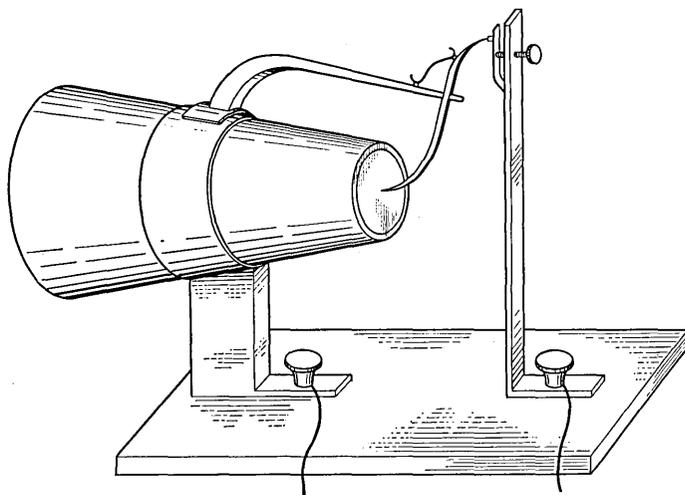


Fig. 1 — Transmitter of Philipp Reis, 1861 or 1862.

control an electric current (Fig. 1). At the other end of the circuit this produced audible tones by controlling the magnetization of a needle whose changes of length, in accordance with the magnetostrictive effect discovered by Page, vibrated a sounding board. In later apparatus an electromagnet was substituted for the needle. Reis's mechanism, like others before and after, included the important feature of deriving the electrical power for transmitting sound by making the sound power control the current from a battery.

Reis had some comprehension of the requirement for electrical transmission of speech, for he noted in his memoir on telephony the need "to set up vibrations whose curves are like those of any given tone or combination of tones" (S. P. Thompson translation). Though his apparatus served primarily to reproduce tones by the make-and-break scheme, he did, by extremely delicate adjustment, succeed in reproducing articulate sounds quite imperfectly. Reis seems not to have realized, however, that this success resulted because his apparatus could, over a very narrow range of speech volume, operate on the principle of changing an electric current in accordance with the voice wave by varying a loose contact, in this case a spring contact between platinum electrodes. His understanding of his own apparatus is indicated by his statement that "each sound vibration effects an opening and a closing of the current." About 20 years later, the German patent office after careful investigation decided that Reis's instrument was not a "speaking telephone."

In the following years attempts were made by other workers (e.g., the Italians A. Meucci and I. Manzetti), but without full realization of the requirements for articulate speech.

Bell's approach was different. In the summer of 1874 the idea of the "electric speaking telephone" became complete in his mind. He described to his father a form of apparatus consisting of a strip of iron attached to a membrane which, when actuated by the voice, would vibrate in front of an electromagnet, thus inducing an undulatory electric current theoretically capable of transmitting speech. At the receiving end a similar device could be used in reverse to reproduce the voice. But Bell doubted that the current generated by the voice would be strong enough to be useful, and for almost a year he made no attempt to construct the apparatus.

On June 2, 1875, while working in Boston on multiplex telegraph apparatus, Bell heard over an electric wire a sound corresponding to the twang of a steel spring at the other end. Recognizing this as a manifestation of the undulatory current principle, he gave his assistant, Thomas A. Watson, instructions for embodying it in a model (Fig. 2)
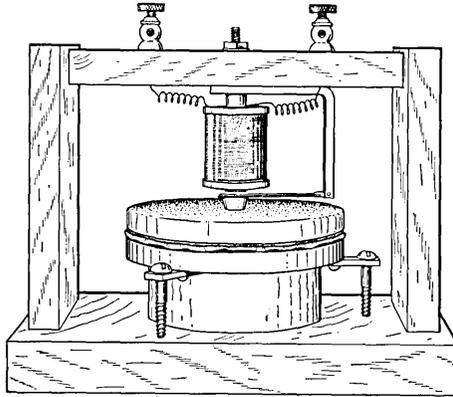
Fig. 2 — "Gallows frame" transmitter of Alexander Graham Bell, June 3, 1875.

of a telephone. The transmitter and receiver were of the electromagnetic type described a year before, and in between was a circuit which included an electric battery. This apparatus transmitted speech sounds the next day, June 3. Bell filed his application for a U.S. patent on Feb. 14, 1876 (Fig. 3). Further experiments produced an instrument (Fig. 4) which on March 10, 1876, transmitted the first complete sentence: "Mr. Watson, come here; I want you."

A few hours after Bell filed his application for patent, Elisha Gray filed a caveat (i.e., a notice of intent to perfect his ideas and file a patent application within three months) for an electric telephone. Gray described a "liquid transmitter," somewhat similar to one patented by Thomas A. Edison for telegraphy in 1873. In Gray's transmitter a voice-actuated diaphragm varied the electrical resistance, and hence the current, by changing the depth of immersion of a rod in water. Bell in
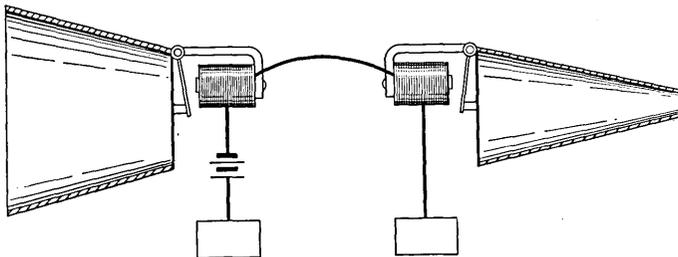


Fig. 3 — Electromagnetic transmitter (left) and receiver (right) illustrated in Bell's first telephone patent, filed Feb. 14, 1876.

his application mentioned the possibility of a similar liquid transmitter, and later used it for his historic summons to Watson. The variable-resistance principle, which subsequently, in the form of a variable carbon contact, proved of vital importance to telephone transmission, makes it possible to obtain an electric wave which is an amplified copy of the sound wave. Like Bell's receiver, Gray's was of the electromagnetic type, similar to one he had patented in Great Britain in July 1874 and in the U.S. in July 1875.

In view of Bell's prior filing of a patent application, the patent for the telephone was issued to him on March 7, 1876. Gray's status as to the invention of the telephone is best set forth in his own words, written to Bell on March 5, 1877: "I do not claim even the credit of inventing it." The claims of Gray, Daniel Drawbaugh and others were subsequently threshed out in prolonged litigation, involving about 600 separate suits, which finally resulted in Bell's patent being upheld in a divided vote by the Supreme Court of the United States.

Bell's first transmitter employing electromagnetic induction, while sound enough in theory, delivered such feeble electrical currents as to be inadequate for general application. The liquid transmitter afforded considerable improvement, but this too had drawbacks. The final essential element for a satisfactory working telephone was the variable-
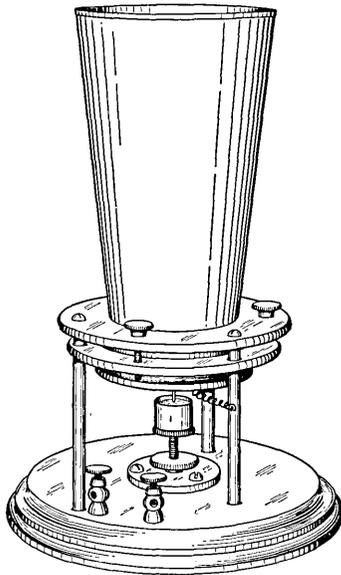


Fig. 4 — Liquid transmitter used by Bell on March 10, 1876.

contact carbon transmitter, due in large measure to Thomas A. Edison, as discussed below.

## INTRODUCTION AND GROWTH

### United States

The development of the telephone business in the United States was undertaken by a group of Bell's backers, under the leadership of Thomas Sanders and Gardiner G. Hubbard. They began by renting or lending telephones in pairs to individuals for local communication. The instruments were extremely crude. Connection between them was made by a circuit consisting of a single iron wire with ground return, and transmission, which was uncertain and poor at best, was possible for only a few miles. Initially there were no switchboards to interconnect a number of users. These came into being in 1877 and 1878.

It was on March 25, 1878, that Bell made a bold prediction that became a charter for the founders of the telephone business:

"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. . . .Believing, however, as I do that such a scheme will be the ultimate result of 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. . . ."

The owners of the telephone patent early incorporated their business, and funds were raised for its progressive development, under the leadership of Theodore N. Vail, who became general manager in 1878. It was recognized that telephone performance is a matter of mutual concern to users, and the practice was established of leasing telephones instead of selling them.

Within ten years after the issuance of the Bell patent, the organization of the Bell Telephone system had assumed something close to its

present form. The local systems were gradually brought together into regional companies operating throughout a state or several states. The systems of these regional companies were linked together by long-distance circuits operated by the American Telephone and Telegraph Company. This company, through ownership of stock in the regional companies, became the parent company of the Bell System.

It soon became evident that standardization of equipment and centralized research on improvements are essential to telephone progress. Accordingly, the Western Electric Company was acquired in 1882 as chief manufacturer and supplier for the Bell System, as well as to conduct research and development. In 1925 the Bell Telephone Laboratories was organized to take over the expanding research and development activities for the system. The 1,600,000 owners of shares in the American Telephone and Telegraph Company in the mid-1950s constituted the largest number of public owners for any one corporation in the world.

After the expiration of the basic telephone patents, many independent telephone companies, not affiliated with the Bell System, sprang up all over the country. Competition became so intense that in many localities there were two companies sharing the business. As this became an increasing source of inconvenience and expense to the public, the service was unified through acquisitions and mergers so as to leave a single company, either Bell or non-Bell, operating in each area. Provision was made for interconnecting the facilities of non-Bell companies with those of the Bell System, thus making possible the interconnection of nearly all telephones in the United States, as well as connections to the rest of the world.

Service and rates are regulated by state utility commisions and the Federal Communications Commission.

There were about 60,000,000 telephones in the U.S. early in 1957 or approximately 1 telephone for every 3 persons in the country. About 82 per cent of these telephones were served by the Bell System, comprising the American Telephone and Telegraph Company and its 20 operating subsidiaries, and 2 associated but noncontrolled companies (the Southern New England Telephone Company and the Cincinnati & Suburban Bell Telephone Company). The remaining telephones were served by about 4,200 independent companies and additional thousands of rural or farmer lines, virtually all of which connect with the Bell System. The total investment in telephone plant and equipment was nearly $20,000,000,000, of which 87 per cent belonged to the Bell System. Telephone traffic averaged over 216,000,000 conversations daily in 1956. Over 800,000 persons were employed by the telephone industry,

TABLE I—TELEPHONE DEVELOPMENT IN THE U. S.

| End of year | Number of telephones | Telephones per 100 population |
|---|---|---|
| 1880................... | 47,900 | 0.09 |
| 1890................... | 227,900 | 0.36 |
| 1900................... | 1,355,900 | 1.76 |
| 1910................... | 7,635,400 | 8.20 |
| 1920................... | 13,329,400 | 12.39 |
| 1930................... | 20,202,000 | 16.34 |
| 1940................... | 21,928,000 | 16.52 |
| 1950................... | 43,004,000 | 28.09 |
| 1954................... | 52,806,000 | 32.21 |
| 1956................... | 60,190,000 | 35.45 |

including more than 100,000 employed in the manufacture of telephone equipment and about 10,000 by the Bell Telephone Laboratories.

The growth of the telephone industry in the United States is shown in Table I.

*Great Britain*

Bell visited England and Scotland on his wedding trip in 1878 in the hope of developing a demand for the telephone. Despite the able support of Lord Kelvin, Sir William Preece and others, he aroused little public interest. He did, however, demonstrate his invention to Queen Victoria, who asked to purchase a pair of telephones, and instead was presented with two instruments done in ivory. Stimulated by this royal recognition, the first telephone exchange was opened in London in 1879 with seven or eight subscribers. Several telephone companies were organized in various parts of Great Britain, but in 1880 the British courts held that the telephone system was legally a telegraph system under an antecedent law which made the telegraph a government monopoly under the postmaster general. The government officials, reluctant to assume the risks involved in developing this new form of communication, issued licenses on a royalty basis to several private companies, which were later consolidated into a single company. A few municipal telephone systems were also established under license.

As the potentialities of telephone communication began to be appreciated, the government gradually took over the service. In 1896 the post office purchased the long-distance lines, and in 1902 it began establishing in London its own local telephone exchanges which were interconnected with those of the privately owned company. Finally, on Jan. 1, 1912, the post office acquired all the private telephone properties. Since then the post office has operated practically all telephones

in Great Britain and Ireland. When the Irish Free State was established, the British post office transferred to the Free State government its telephone system in southern Ireland.

At the end of 1956 there were about 7,214,000 telephones in the United Kingdom, which was more than in any other European country, and represented 1 telephone per 7 persons. The British post office maintains a large research and development organization. Telephone apparatus is supplied by private companies that manufacture to post office specifications.

### Germany

In Germany the telephone was a government monopoly from the beginning. Heinrich von Stephan, postamaster general and manager of the imperial telegraphs, when he learned about Bell's invention from an article in the *Scientific American,* ordered models for trial. After successful experiments at distances up to 90 mi. (Berlin-Magdeburg), he suggested on Nov. 9, 1877, to Prince Otto von Bismarck, the imperial chancellor, that the telephone be used as an adjunct to telegraph service in rural post offices where there was not sufficient traffic to justify a trained telegraph operator. Within two years 800 villages were thus connected.

So it came about that in Germany the first public use of the telephone was for long-distance communication, telephones being found only in government post offices. The situation was therefore the reverse of that in the United States and Britain, where local exchanges came first. Later, as public demand forced the issue, exchange service was gradually introduced, starting in the larger German cities. In the German states of Bavaria and Württemberg, the telephone systems were operated by the state administrations until 1920, when they were transferred to the German post office.

At the beginning of 1957 the Federal Republic of Germany, commonly known as West Germany, had the second largest telephone system in Europe, with 4,323,000 telephones or about 1 telephone per 12 persons. The German Democratic Republic, commonly known as the eastern or Soviet zone, had 1,067,000 telephones or about 1 telephone per 17 persons.

### France

In France the telephone was first exhibited at the Paris world's fair in 1878, where it attracted little interest. The next year, however, the French telegraph officials, unwilling themselves to pioneer in this new

field, granted concessions to several private companies which later consolidated into the Societé Générale des Téléphones. This company initiated public telephone service in Paris in 1881 and subsequently in other cities. Starting in 1883, the French government established exchanges in various centers. In 1889 the government took over the entire private system, and after that time operated the telephone service. With 3,313,000 telephones (over one-fourth of these in Paris) at the beginning of 1957, the French telephone system was third in size among those of Europe.

### Switzerland

In Switzerland a privately owned telephone system was established under government concession in Zürich in 1880, while the government itself opened exchanges in Berne and Basel in 1881. Thereafter the government proceeded rapidly to establish new exchanges, and in 1886 purchased the Zürich system. After that time the Swiss government operated all telephones. At the beginning of 1957 there were about 1,294,000 telephones, amounting to almost 1 telephone for every 4 persons.

### Scandinavia

The Scandinavian countries have achieved a high degree of telephone development. In Sweden, the International Bell Telephone Company of New York opened exchanges in Stockholm and Göteborg in 1880, and not long after in Malmö and elsewhere. In 1883 a competing telephone company was set up in Stockholm under the enterprising leadership of H. T. Cedergren. Co-operative telephone associations were established in Göteborg and in many rural communities. After a strongly competitive phase, in which the government participated, virtually all telephones were taken over by the state during the period 1890–1923. At the beginning of 1957 the Swedish Telecommunication administration operated about 2,312,000 telephones. The telephone density, approximately 1 telephone for every 3.2 persons, was higher than anywhere except in the United States.

The Swedish administration not only maintains a development organization, as do most other countries in Europe, but it is unique among European telephone administrations in possessing factories for production of telephone equipment.

In Norway the International Bell Telephone Company in 1880 secured franchises for Oslo and Drammen. The next year local companies

established exchanges in several cities, and a competing system in Oslo. Within ten years, largely through local enterprise, the telephone came to hold much the same place in Norwegian rural life that it does in the sparsely settled districts of the U.S. Ultimately the state acquired the more important local systems, and at the end of 1956 about 92 per cent of the country's 615,000 telephones were government-operated.

In Denmark, likewise, the telephone was introduced by private enterprise, the government interposing no serious difficulties. Gradually numerous small systems were consolidated into several relatively large organizations. The principal one is the Copenhangen Telephone Company, a stock company in which the government owns a controlling interest. Nearly half of Denmark's 923,000 telephones at the beginning of 1957 were in the city of Copenhagen. Denmark had at that time about one telephone for every 5 persons.

*Belgium and the Netherlands*

In Belgium telephone exchanges were first established in various cities by private concessionaires. In 1896, however, the telephone system became a complete government monopoly, and has remained so. At the beginning of 1957 there were 931,000 telephones. Private grants formed the initial pattern in the Netherlands also. As competition with the government telegraphs became apparent, the private telephone systems were integrated into a single government system, which at the end of 1956 comprised about 1,229,000 telephones.

*Austria, Italy and Spain*

Austria followed the typical European course, starting with private companies which later were bought in so as to create a government monopoly. There were 540,000 telephones at the end of 1956 or about 1 telephone for every 14 persons.

In Italy the story is somewhat different. To begin with, concessions were granted to private companies and in a number of important cities competititive situations arose. As the disadvantages of competition manifested themselves, consolidation was effected either voluntarily or by the dictates of local authorities. Partly as a result of onerous regulations, telephone development failed to keep pace with that in other European countries. In 1925 the structure of Italian telephone service assumed essentially the form that obtains today. Five concessionary companies operate local and toll service, each within one of the five zones into which Italy has been divided for this purpose. These com-

panies are privately operated, although the government participates in
the ownership. The government operates long-distance interconnecting
service as well as the landline portion of international service. Growth
in recent years has been rapid, and at the end of 1956 Italy had 2,609,000
telephones.

In 1924 Spain granted a concession to a subsidiary of the International
Telephone and Telegraph Company of New York to provide a nation-
wide telephone system to supersede the previous government- and
privately-owned systems which had attained only a limited development.
There were operating difficulties during the Spanish civil war, and in
1945 the government purchased the International Telephone and Tele-
graph operating interest. At the end of 1956, 1,199,000 telephones were
operated by a private company in which the government held a con-
trolling interest.

## U.S.S.R. and Satellites

In the mid-1950s no recent official statistics were available on the
telephone systems of the U.S.S.R. and satellite countries, but develop-
ment there had lagged far behind that in western Europe.

## Canada

Canada, divided by natural barriers and with most of its population
concentrated near the long southern border, nevertheless at the end of
1956 ranked next after the United States and Sweden in density of
development, having 1 telephone for every 4 persons and a total of
4,502,000 telephones. Moreover, Canadians led the world in use of the
telephone, making 481 calls per person in 1956, compared with 455 for
Sweden and 426 for the U.S. Most of the telephones in Canada were at
that time administered by seven private companies, which were banded
together as the Trans-Canada system to furnish through service.

## Latin America

Early development of telephony in Latin America was slow, for various
reasons. Colonization by different European countries, superposed on
the native background, led to a wide variety of languages and cultures.
Tropical areas, mountainous regions and other natural barriers tended to
limit development to the more populous isolated areas. In many cases
foreign capital was obtained through private concessions, but certain
countries where this was done followed the European plan of subsequent
transfer to government ownership.

At the end of 1956 Argentina, with 1,155,000 telephones, led the Latin-American countries in telephone development and Brazil, with 843,000, came next. Telephone systems of some size in Latin America were operated by subsidiaries of certain companies located in foreign countries, as follows: International Telephone and Telegraph Corporation, in Brazil, Chile, Peru, Cuba and Puerto Rico; the L. M. Ericsson Company, in Argentina and Peru; Cables and Wireless, Ltd., in Peru. Mexico's largest system was owned jointly by the I. T. & T. Corporation and the L. M. Ericsson Company. During the five-year period ending with 1956, the percentage growth of telephones in South America exceeded that in North America.

*Elsewhere*

Telephone development in the rest of the world may be summarized with the statement that, at the end of 1956, while some sort of telephone service is to be found almost everywhere, systems of substantial size are limited to Japan (3,487,000 telephones), Australia (1,762,000), South Africa (766,000), New Zealand (568,000) and Finland (486,000). Asia, with the lowest average level of telephone development, exceeded all the other continents in percentage growth during the five-year period ending with 1956.

*Growth Factors*

The rapidity of telephone growth and the intensity of development in different parts of the world were affected by a variety of factors, among which may be noted scientific advance, economic status, culture, degree of industrialization, size of market, linguistic and dialectic diversities, physiography, type of ownership, managerial enterprise and political and military considerations. Although the qualitative influence of these factors upon telephone development can frequently be observed, it is impracticable in any particular situation to assign specific weights to them. Sometimes cause and effect are inseparable. Civilization begets telephones, and telephones beget civilization.

*Telephone Statistics*

At the end of 1956 there were approximately 110,000,000 telephones in the world, 55 per cent of these in the United States. Any one of 106,000,000 telephones, scattered throughout the world, could be connected to a telephone in the U.S. Almost 70 per cent of the world's telephones were privately owned, the remainder government-owned.

Fig. 5 — U. S. overseas telephone circuits in the mid-1950s.

TABLE II—COUNTRIES WITH MORE THAN 1,000,000 TELEPHONES IN
SERVICE ON JAN. 1, 1957

| Country | Number | Country | Number |
|---|---|---|---|
| United States......... | 60,190,000 | Sweden............. | 2,312,000 |
| United Kingdom...... | 7,219,000 | Australia............ | 1,762,000 |
| Canada............... | 4,502,000 | Switzerland.......... | 1,294,000 |
| Germany, Federal Republic.............. | 4,323,000 | Netherlands......... | 1,229,000 |
| Japan................. | 3,487,000 | Spain............... | 1,199,000 |
| France................ | 3,313,000 | Argentina........... | 1,155,000 |
| Italy................. | 2,609,000 | Germany, Democratic Republic.... | 1,067,000 |

The distribution of telephones by principal countries is shown in Table II. Major U.S. overseas telephone connections are shown in Fig. 5.

Among the world's cities, Washington, D.C., led in density of telephone development, with 65.3 telephones per 100 population in 1956. Others ranking high among cities of more than 100,000 population included: Los Angeles, 64.9; San Francisco, 58.2; Stockholm, 54.8; Berne, Switz., 54.1; Basel, Switz., 55.9; Geneva, Switz., 53.5; Hartford, Conn., 53.2; Pasadena, Calif., 52.2; Wilmington, Del., 51.8; and Denver, Colo., 48.5.

## ITU and CCITT

Much of the progress in international telephone communication may be attributed to the work of international co-ordinating bodies. This work now centers in the Union Internationale des Télécommunications (UIT), called in English the International Telecommunication Union, abbreviated ITU. This organization, now an agency of the United Nations, was originally founded in 1865 as the Union Télégraphique Internationale. The purposes of this organization as set forth in its charter are "To maintain and extend international co-operation for the improvement and rational use of telecommunication; to promote the development of technical facilities and their most efficient operation, in order to improve the efficiency of telecommunication services, increase their usefulness, and make them, as far as possible, generally available; to harmonize the actions of nations in the attainment of those common ends." On July 1, 1954, the ITU had a membership of 90 countries or territories. The activities of the ITU and its constituent bodies are financed by contributions from member governments.

Of special interest and importance are three permanent and essentially autonomous organs of the ITU: (1) the Comité Consultatif International Téléphonique (International Telephone Consultative Committee),

abbreviated CCIF, founded in 1924; (2) the Comité Consultatif International Télégraphique (International Telegraph Consultative Committee), abbreviated CCIT, founded in 1925; and (3) the Comité Consultatif International des Radio Communications (International Radio Consultative Committee), abbreviated CCIR, founded in 1927.

For implementing its work, each of these three international consultative committees has organized a number of study groups covering different phases of activity. A plenary assembly distributes questions to be investigated by the study groups. These work by correspondence or by meeting and submit their findings to the succeeding plenary assembly which studies and discusses them and makes recommendations. Ordinarily the recommendations take the form of directives or rules which, without being mandatory, are usually observed by the technical and operating services of the government administrations and private operating agencies of the countries belonging to the union. The director of the consultative committee co-ordinates the work of the study groups, of the plenary assembly and of the committee as a whole.

The work of these three committees has emphasized the importance of good engineering and transmission standards to world-wide communication service and has gone far in bringing about the introduction of such standards. In particular, the CCIF has given exhaustive consideration to problems of transmission and interference, frequency bands and allocations, standards of telephone performance and questions of suitable rates and classes of service, all of which has contributed to effective international long-distance telephony. On Jan. 1, 1957 the CCIF and CCIT merged to form a new organization known as the CCITT (Comité Consultatif International Téléphonique et Télégraphique.

TECHNICAL DEVELOPMENT

The first telephones were extremely crude; it was barely possible to talk over them. The problems of transmitting speech over substantial distances were not understood, or even visualized. Nor was there the slightest comprehension of the complications involved in interconnecting large numbers of telephones. Far-reaching developments on many fronts were needed to bring the art of telephony to its present state of perfection.

One indispensable phase was the progressive improvement of telephone station instruments as to quality and loudness of speech, as well as convenience and cost. Concurrently, it was necessary to develop telephone lines and circuits that could transmit speech currents reliably, without appreciable impairment or interference, and economically, for

short or long distances. These two advances are sometimes grouped under the single term, transmission development. Through such development, telephone conversation has, except under adverse conditions, become practically effortless.

Another essential was the development of switching mechanisms whereby any two telephone instruments in a large system could be connected together certainly, rapidly and economically. This part of the advance is referred to as switching development. Progress in these areas came about gradually in different countries, along generally similar but not identical lines, and with, in many instances, substantial time differences.

*Station Instrumentalities*

The station set at the premises of a telephone customer normally consists of a transmitter, which converts the speech waves in the air into their electrical replicas; a receiver, which performs the reverse operation of converting the incoming electrical waves into sound waves; a transformer (formerly called an induction coil) designed to increase the effectiveness of the transmitter and to permit full duplex operation, *i.e.*, use of a single pair of wires for speech in both directions, without requiring a switch to change from talking to listening or vice versa; a bell or equivalent summoning device; a switch hook to control connection of the set to the customer's line; and various associated items. In automatic switching systems the station set also includes a device, such as a dial, for generating signals to actuate the switching mechanism.

Bell's original electromagnetic transmitter served also as the receiver, the same instrument being held alternately to the mouth and the ear. An important forward step was the invention by Émile Berliner in 1877, and by Thomas A. Edison later in the same year, of different forms of transmitters utilizing variable-contact resistance between two solid electrodes connected in a battery circuit. Berliner employed metallic electrodes, while Edison used semiconducting materials, particularly plumbago (*i.e.*, graphite). Berliner also introduced the induction coil into the transmitter circuit in 1877.

The principle of the microphone contact, which underlay both the Berliner and Edison transmitters, was elucidated by David E. Hughes in England in 1878. By virtue of this principle, energy from a battery is controlled by the voice waves so that the telephone transmitter acts as an amplifier as well as a converter of sound waves. The first recorded recognition of amplification through a microphonic contact seems to be found in a German patent of Robert Lüdtge, issued Jan. 12, 1878.

A transmitter in which granular carbon was used for the variable contact was invented in 1878 by Henry Hunnings, an English clergyman. Edison invented a transmitter whose carbon granules were obtained from anthracite coal, and Anthony C. White in 1890 invented the solid back transmitter, using a "button" of granular carbon placed between a fixed electrode and a diaphragm-actuated movable one. White's transmitter incorporated all the basic features of the modern telephone transmitter, though great improvements in efficiency (up to a thousand-fold amplification of speech sounds), naturalness, resistance to aging and other features have resulted from continuing research and development. Extensive studies of the properties of speech and hearing have contributed to this result.

Modern telephone receivers utilize the same basic principles found in Bell's original instrument. Fundamentally, the receiver consists of a permanent magnet having pole pieces wound with coils of insulated fine wire, and a diaphragm driven by magnetic material which is supported near the pole pieces. Speech currents passing through the coils vary the attraction of the permanent magnet for the diaphragm, causing it to vibrate and produce sound waves. Through the years the design of the electromagnetic system has been continuously improved to provide better talking qualities. In a new type of receiver, introduced in the
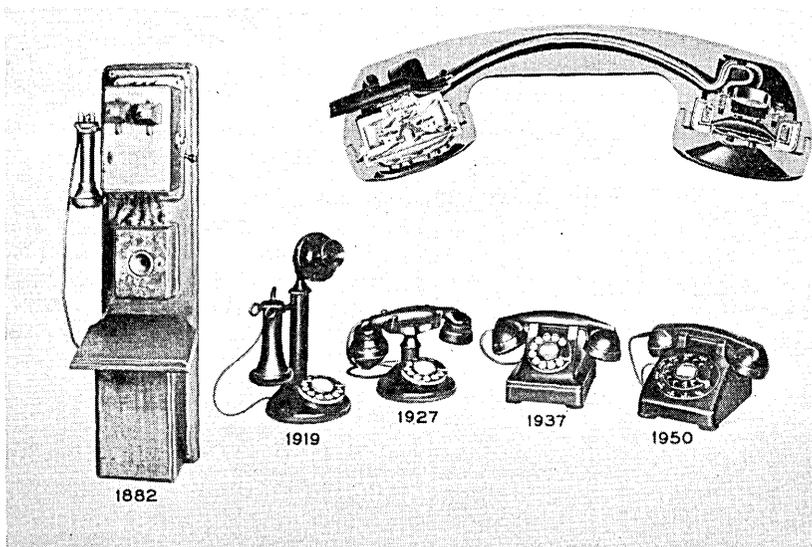


Fig. 6 — The evolution of telephone set design in the Bell System, including the approximate date when each design was adopted. *Inset:* Cross section of a modern handset.

United States in 1951, the diaphragm, consisting of a central cone and ring-shaped armature, is driven as a piston to obtain' efficient response over a wide frequency range.

In early telephone sets a hand-cranked generator, or so-called magneto, was used for signaling the operator, and local dry batteries supplied power to the transmitter. A centralized battery arrangment for signaling was invented in 1880 and one for both talking and signaling in 1886. With the gradual introduction of this common battery system, magneto telephone sets were largely superseded.

The first suggestions for mounting a telephone transmitter and receiver on a common handle, thus forming what is now known as a handset, were made by two Englishmen, Charles A. McEvoy and G. E. Pritchett, in 1877. Starting in 1878, a type of handset devised by Robert G. Brown was used by boy operators in the Gold and Stock telephone exchange in New York city. Handsets for customers' use were introduced in France about 1882 and spread quickly in Europe. Initial difficulties in meeting satisfactory transmission standards with the handset arrangement were gradually overcome and most modern instruments are of this type.

Methods of connecting the transmitter and receiver to the line have likewise been gradually improved. Starting with work by George A. Campbell, arrangements known as anti-sidetone circuits have been developed whereby most of the electrical energy generated in the transmitter is directed toward the distant station, with a minimum entering the speaker's receiver.

Apart from the basic telephone instrument, special customer equipment and arrangements are provided to meet individual requirements of various kinds. Telephone sets are available which permit the user to hold a conversation without lifting the receiver, a small microphone being used to pick up the voice and a small loud-speaker to reproduce the incoming speech. There are loud gongs or other signaling devices for noisy locations, amplifiers for persons with subnormal hearing, automatic telephone-answering devices, arrangements for recording conversations, loud-speakers for paging service, etc.

Telephone pay stations installed in public or semipublic locations provide a coin telephone set which includes, in addition to the usual station equipment, a coin collector having one or more slots designed to accept legitimate coins and reject slugs or spurious coins. The coins, in passing down their respective slots, strike distinctive gongs whose tones permit an operator to supervise the deposits. Many coin collectors are arranged to hold the coins in suspension in a hopper, with means provided whereby the suspended coins can be collected or refunded, depending on whether or not the desired connection is completed.
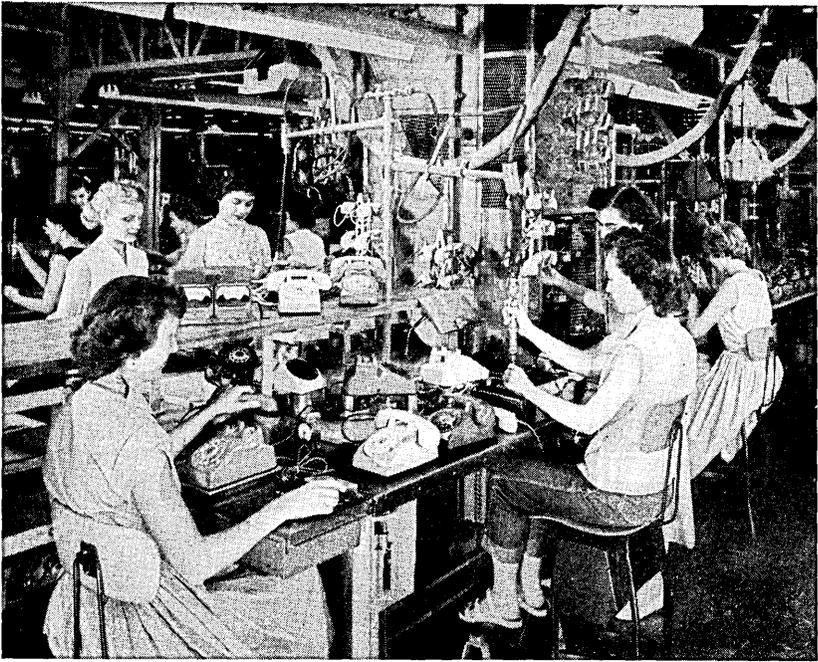
Fig. 7 — Assembly and testing of telephone sets.

Key telephone systems, affording greater flexibility of telephone usage than a single station set, are frequently used in business offices and residences. Combinations of push-button, turn-button or lever-type keys, installed on a desk or table or mounted integrally in the base of a telephone set, can be arranged to perform a variety of functions, particularly where two or more lines and a number of stations are involved. Thus a telephone may be connected to any of several lines, a call may be held on one line while conversation proceeds on another, etc.

*Telephone Lines*

Telephone circuits are furnished principally by wire lines, although radio is used to a moderate extent. Telephone lines comprise a network of wires which interconnect individual telephone stations, central offices and communities. These wire lines are of two forms — cable and open wire. The usual type of cable consists of insulated copper wires, twisted together and usually covered with a protecting sheath. Open-wire lines consist of bare wires, generally of copper, fastened to insulators which

are supported on poles at some distance above ground, commonly on pins in crossarms.

Following previous telegraph practice, the first telephone circuits utilized a single overhead wire, usually iron, with ground return. An important early improvement was the development by Thomas B. Doolittle in 1877 of hard-drawn copper wire, giving good tensile strength together with improved electrical conductivity. The advantages of a two-wire or so-called metallic circuit in reducing noise and interference were soon realized. It was found also that by transposing the wires, i.e., interchanging their positions, a number of circuits could be carried on a single pole line without excessive interaction which would permit the conversation on one circuit to be heard on another. These several features made possible a successful open-wire telephone line between Boston and New York in 1885.

Early in telephone history, as the number of open-wire circuits strung on poles and rooftops in the large cities began to reach the point of impracticability, methods of compacting the lines in overhead or under-ground cable were tried. At first the wires were placed in pipes and sealed against moisture with oil, paraffin or asphaltum. Means were soon developed whereby lead, heated to plasticity, could be extruded over a core of conductors. The introduction of dry paper as insulation for the conductors completed the foundation for the modern telephone cable.

Research and development covering many materials and processes made it possible to increase the number of pairs in a full-sized local cable from a maximum of 50 in the year 1888 to 2,121 in 1955. This increase was accomplished largely by reducing the size of the copper wires, from 25 lb. per wire mile in 1888 to 4 lb. in 1955. Following World War II, new types of cable sheath were developed in order to reduce cost and lessen dependence on lead supply. One form uses thin layers of aluminum and steel covered with polyethylene.

In a typical urban or suburban installation, insulated wires from each customer's premises extend to a distribution cable which in turn connects to a feeder cable leading to the central office. In densely built areas the feeders are generally placed in underground ducts; elsewhere, overhead. Circuits between central offices, known as trunks, are pro-vided in trunk cables which usually employ larger wires than those leading to customer stations.

The network of lines which interconnect approximately 73,000 com-munities in the U.S. included in the mid-1950s approximately 31,000,000 miles of wire on 230,000 miles of route. About 91 per cent of the wires were in cable, the remainder open wire. For the rest of the world, the per-
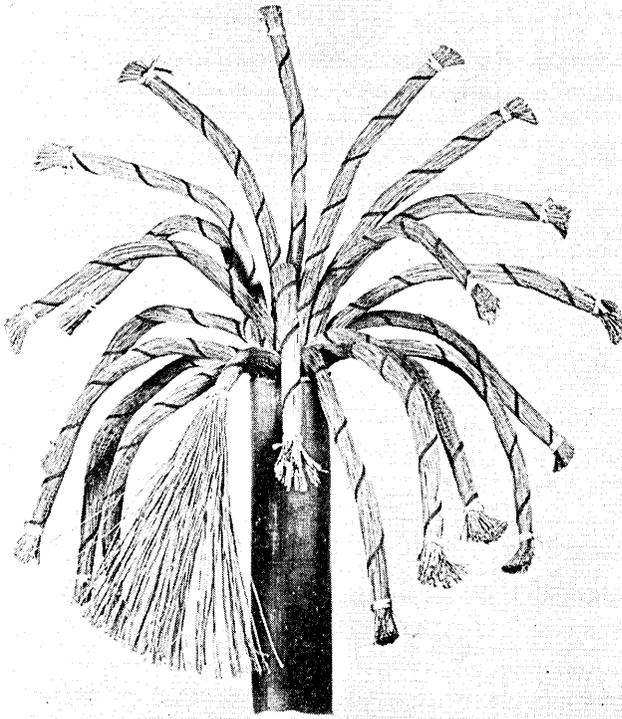
Fig. 8 — Multiple-unit telephone cable, consisting of 21 groups, each comprising 101 pairs of 26-gauge wires.

centage of long-distance facilities furnished by cable ranges from nearly 100 per cent in a few countries to a very low figure in the less populated regions. Since cable is virtually storm-proof, it affords much greater reliability than open wire.

A type of long-distance cable common in the U.S. employs conductors 36 mils in diameter, weighing 20 lb. to the mile (no. 19 AWG gauge). In making this cable, two paper-insulated wires are first twisted together to form a pair; then a quad is formed by twisting two pairs together. This twisting aids in the prevention of crosstalk (*i.e.*, overhearing) between different pairs. In European toll cables, four conductors are usually twisted together to form a spiral four or star quad. In either case the quads are grouped together and enclosed by the sheath. Such a

cable may be supported aerially or placed in an underground duct, or, with a suitable protective covering of jute or steel tape, may be buried in the ground. A special type of cable used for coaxial systems is discussed under *Carrier Systems* below.

*Phantoms*

Around 1900 the principle of "phantoming" two pairs of wires was introduced. The original idea of phantoming was devised by Frank Jacob in 1882, and the present method by John J. Carty in 1886. The phantom arrangement makes it possible to derive from two pairs of wires a total of three telephone circuits. The additional circuit, called the phantom, is obtained by using the two wires of each pair in parallel as a side of the phantom. To render this scheme practical, efficient balanced transformers were needed for the end connections, and transposition arrangements to keep the crosstalk between the three circuits within tolerable bounds. The phantom principle came to play an important part in both open-wire and cable plant. Because the presence of phantoms makes high-frequency transmission quite difficult, the advent of carrier systems has greatly curtailed the use of phantoms for long-distance circuits.

*Loading*

As speech currents pass along a line their strength decreases (a process referred to as attenuation), so that after some distance they become too weak to actuate a receiver properly. Studies by A. Vaschy (1889) and Oliver Heaviside (1893) developed the theoretical possibility of improving the transmission efficiency of telephone lines by artificially increasing their inductance. Various investigators speculated on the practicability of approximating the beneficial effect of an increase in uniformly distributed inductance by introducing in the line concentrated or lumped inductance in the form of low-resistance loading coils. Finally, in 1899, Michael I. Pupin and George A. Campbell (working independently, Pupin having a slight priority) discovered that the key to the problem was the spacing of the loading coils. By providing at least pi (*i.e.*, about 3.14) loading coils per wave length at the highest frequency to be transmitted, a substantial reduction in the attenuation of the speech waves is obtained. Thus a cable circuit may, by means of coil loading, be made to transmit telephonic currents as efficiently as a non-loaded one whose conductors weigh many times as much.

Another way of adding inductance to a cable circuit was proposed by Carl Emil Krarup of Copenhagen, his idea being to wind helically

along the copper conductor a fine wire of soft iron. A cable with this continuous loading was placed between Elsinore and Helsingborg in 1902. While continuously loaded cables have proved of some importance for submarine applications, coil loading has been almost universally preferred for loaded land cables.

By 1913 coil loading made it possible to extend the useful range of open-wire circuits to approximately 2,000 miles and to employ underground cable to connect Washington, D.C., and Boston via New York City. After electron tube amplifiers became available, open-wire loading was largely abandoned, but loading is still extensively applied on cable circuits used for voice frequencies, and especially on local trunk circuits.

Satisfactory loading requires that the loading coils have very low energy losses in their magnetic cores and copper windings. The toroidal-shaped cores of modern loading coils make use of a powdered magnetic alloy whose particles are individually insulated, compressed under high pressure and then usually heat treated to develop optimum magnetic properties. Progressive improvement of magnetic core materials has brought about large reductions in the size and cost of loading coils, as well as improved performance. A core material commonly used in the U. S. consists of an alloy of nickel, iron and molybdenum known as molybdenum permalloy, discovered by Gustav W. Elmen. In Europe powdered iron cores of somewhat larger size are employed.

*Electron Tubes and Repeaters*

For years the range of telephony was severely limited by loss of energy due to dissipation along wire lines, or to spreading in the case of radio waves. The idea of inserting one or more repeaters in a telephone line for the purpose of reinforcing or amplifying the telephonic currents from some local source of energy is almost as old as the telephone itself, but many years elapsed before the quest for a satisfactory repeater achieved success. In a so-called mechanical repeater, tried in 1904, inertia of the moving parts was found to present inherent limitations. Subsequently H. E. Shreeve developed a repeater employing carbon-contact amplification which was capable of practical use. The real solution to the amplification problem, however, was found in the device invented by Lee De Forest in 1906, which he called the audion, and which is now known as a three-electrode vacuum tube or electron tube, or in England as a valve. In its original form as used in radio telegraphy this tube was unsuited for telephone purposes. Research by Harold D. Arnold, Irving Langmuir and others showed that a major requirement for adequate performance in an amplifier was the creation of a high

degree of vacuum inside the tube envelope. By 1914 satisfactory high vacuum tubes were produced. Using telephone repeaters with amplification supplied by vacuum tubes, telephone service between New York and San Francisco was inaugurated in 1915. So telephony at last found the means for conquering distance, and it was this application of the vacuum tube in the telephone business that ushered in the electronic age.

Continuing development work brought about many improvements in vacuum tube repeaters. Efficient circuits and auxiliary equipment were devised for utilizing the amplifier element and associating it with the line circuits. The life of the repeater tube most commonly used in 1917 was about 1,000 hours, whereas standard tubes of a type introduced in 1935 have a life of about 90,000 hours, equivalent to more than ten years of continuous operation. In addition, the power required to heat the tube filaments was reduced to one-tenth that required in 1917.

New types of tubes with four or five elements, especially adapted for use in carrier and radio systems, were developed also. The role of vacuum tube repeaters in modern telephony is evidenced by the fact that in 1955 there were in the telephone plant in the U. S. about 480,000 voice and carrier repeaters using about 5,000,000 vacuum tubes.

With repeaters available, transmission defects of different kinds became increasingly apparent. The lines, particularly if loaded, were found to introduce severe distortion by reason of differences in the transmission efficiency and transmission velocity at different frequencies. Furthermore, large variations in transmission loss resulted from changes in the electrical resistance of cable conductors with temperature, or from changes in the leakage of open-wire conductors. Extensive developments in the field of network theory made it possible to design equalizers with frequency characteristics that compensate accurately for the line distortion. Variations in line loss are automatically counteracted by transmission regulators. Devices known as compandors, in which the amplitude of speech syllables serves to compress the range of speech volumes at the transmitting end and to introduce a corresponding expansion at the receiving end, have proved beneficial in reducing the effect of line noise on both wire and radio circuits.

Another unwanted effect on telephone circuits is the presence of electrical echoes due to irregularities in the line. These, when converted into sound, may disturb both talker and listener. The echoes become more annoying the longer they are delayed, and hence are of greatest concern on long circuits. They are controlled by restricting their occurrence, by reducing the line delay or by applying devices called echo suppressors to prevent their reaching the telephone users.

The vacuum tube provided for the first time the means for precise measurement of these and other transmission effects, and thus made it possible to establish a firm foundation for the transmission art. Vacuum tubes connected as generators of electric oscillations supply the range of frequencies needed for communication measurement. Other vacuum tubes amplify the extremely weak currents involved in telephone transmission to a level at which they can be conveniently measured.

*Carrier Systems*

It is now rather general practice on long-distance routes to multiply the circuit capacity by the application of carrier telephone systems. These are called carrier systems because of the underlying principle, known as modulation, whereby the voice frequencies modulate a higher frequency current which "carries" the voice currents. A voice wave is composed of undulations with a frequency of occurrence ranging from about 200 to 3,000 per second or, as it is commonly expressed, from 200 to 3,000 cycles per second. In the modulation process these frequencies are transposed to a higher frequency range, *e.g.*, 10,200 to 13,000 cycles per second, for transmission over the line. By using different carrier frequency bands a number of conversations can be sent simultaneously over one transmission path. At the receiving end the carrier frequency bands are separated by electric networks called filters (invented by G. A. Campbell in 1915), and the original voice frequencies recovered by an inverse process known as demodulation. Carrier systems make extensive use of vacuum tubes—as amplifiers, as oscillators for generating carrier frequencies, and sometimes as modulators for shifting bands of frequencies from one range to another.

The art of multiplex carrier telephony grew out of the harmonic telegraph systems associated with the names of Gray, Bell, E. Mercadier and others. The extension of the carrier principle to telephony, together with the use of electrical resonance instead of mechanical resonance for selecting the carrier frequencies, was invented in 1891 by two Frenchmen, Maurice Hutin and Maurice Leblanc. During the period 1908–11, demonstrations of carrier techniques were conducted by Ernst Ruhmer in Germany and Maj. Gen. George O. Squier in the U. S.

Following experiments between Toledo, O., and South Bend, Ind., in 1917, the first commercial application of the carrier principle was made in 1918 on an open-wire line between Baltimore, Md., and Pittsburgh, Pa., giving four additional telephone circuits on a pair of wires. With modern carrier techniques, as many as 16 telephone circuits are derived from a single open-wire pair. Beginning in 1934 in Germany, a

system which yielded one extra telephone channel on a lightly loaded cable pair was extensively applied in Europe.

Application of substantial numbers of carrier channels to cables necessitates many more repeaters because of the higher attenuation of the cable pairs, so that minor imperfections in each amplifier become quite serious. The solution to this problem was provided by Harold S. Black when he invented the negative feed-back amplifier in 1927. The complete development of the underlying theory came in subsequent work by Harry Nyquist. Though difficult in theory, the idea itself is quite simple. A part of the amplifier output is fed back to the input in such a way as to give practically distortionless amplification, together with almost complete absence of variation in amplification as a result of power supply variation and tube aging. The improvement thus obtained may be a thousandfold or more.

The negative feedback principle is now applied almost universally to amplifiers used for any purpose. This principle formed the basis for the introduction in 1937, between Toledo and South Bend, of a 12-channel carrier system operating on nonloaded cable pairs. Today carrier systems are quite generally used on long-distance cable routes, yielding circuits with excellent transmission properties for distances up to thousands of miles. Important in this connection has been continued progress in the understanding and control of crosstalk and noise. Cable carrier systems in the U. S. are arranged to derive 12 to 24 telephone circuits from two nonloaded 19-gauge pairs. Similar systems, some of them providing much larger numbers of channels, are used in Europe and elsewhere.

With the aid of a special type of cable conductor, called a coaxial unit or merely a coaxial, carrier techniques have been greatly expanded. A coaxial consists essentially of a copper tube, commonly about the size of a lead pencil, with a wire centrally supported inside. A full-size cable may contain as many as eight such coaxials, plus a number of conventional pairs of wire. By applying amplifiers at close intervals, four to eight miles, a very wide band of frequencies can be transmitted over a single coaxial.

The transmission properties of a coaxial circuit were considered by various 19th-century workers, especially Lord Kelvin and Alexander Russell. But it was a far journey from these studies to a wide-band transmission system suitable for long-distance multichannel telephony and television. The first such coaxial systems were applied in 1936 between New York City and Philadelphia and between Berlin and Leipzig. More highly developed systems are now widely used in both the U. S. and Europe.
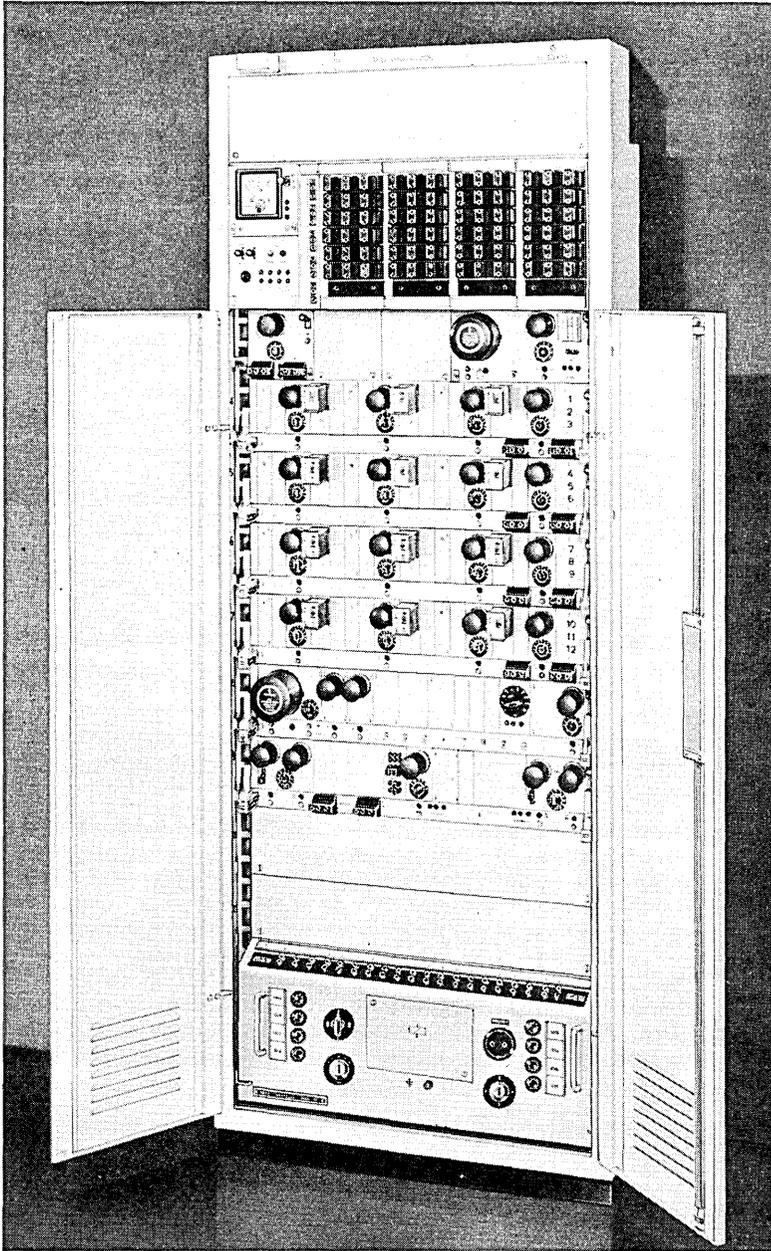
Fig. 9 — Twelve-channel carrier telephone equipment for use on balanced cable pairs. (By courtesy of Siemens & Halske, Ag.)
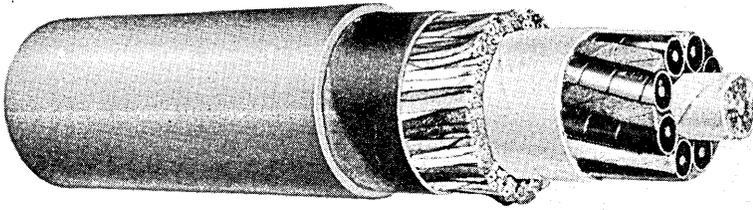
Fig. 10 — Coaxial cable, with eight coaxial units and a number of paper-insulated pairs.

A type of coaxial system introduced between Minneapolis, Minn., and Stevens Point, Wis., in 1939 utilizes a frequency band nearly 3 mc (*i.e.*, 3,000,000 cycles) wide, with two coaxials giving either 600 two-way telephone channels or a television circuit in each direction. Techniques were further augmented in a system, first applied in 1952, which provides a frequency band nearly 8 mc wide, so that two coaxials can provide either (1) 1,800 telephone channels; or (2) 600 telephone channels plus a 4.2-mc television circuit in each direction. Since a long coaxial system may have 1,000 or more repeaters connected in tandem, each one providing a ten-thousandfold amplification, the utmost perfection is necessary in the performance of each repeater.

The minimum distance for which carrier systems prove economical depends on the cost of the terminal apparatus. In recent types of systems the economical distance has been greatly reduced, so that carrier can be used for short-haul toll circuits, interoffice trunks and rural circuits.

## Radio

Radio, originally employed for telegraphy, has become an important instrument for telephone purposes. Both the theoretical and the utilitarian aspects of radio transmission are treated elsewhere.

The same basic principle of modulation used in carrier systems is required to shift the telephone signals to the desired radio frequency. There are several different kinds of modulation that may be used in either radio or wire systems. Simplest of these is amplitude modulation, in which the amplitude of a modulating wave (*e.g.*, a speech wave) controls the amplitude of a sine wave carrier. The first complete analysis of amplitude modulation was made by John R. Carson, who showed that in order to convey the intelligence only one of the bands of fre-
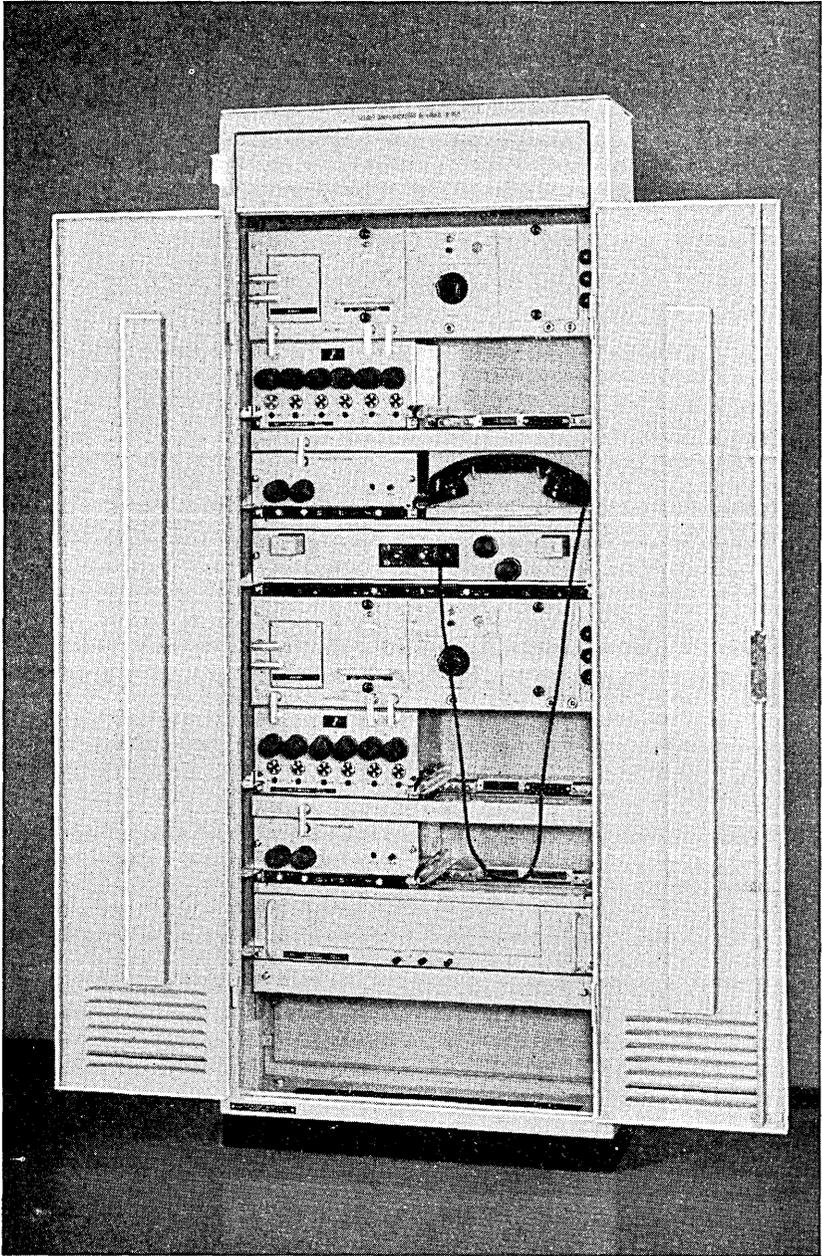
Fig. 11 — Carrier telephone equipment for coaxial cables. (By courtesy of Siemens & Halske, Ag.)

quencies generated in the process need be transmitted. Because of its efficient use of frequency space, Carson's single-sideband method is widely employed.

Another common form is frequency modulation, wherein the instantaneous frequency of a sine wave carrier is varied in proportion to the amplitude of the modulating wave. This method makes it possible by sacrificing frequency space to gain an advantage in the ratio of signal to noise.

Still another form is pulse modulation, in which the carrier consists, not of a sine wave, but of a series of pulses whose amplitude, duration, position or mere presence may be controlled so as to convey the message. In pulse code modulation, originated by A. H. Reeves in 1939, successive quantized samples of the modulating wave produce corresponding code patterns of pulses. This provides, at the expense of band width, low vulnerability to noise and interference and adaptability to repeated regeneration of the signals without distortion. For multi-channel transmission, the pulses corresponding to different channels are interleaved.

The advantages of being able to telephone without a wire connection were obvious from the beginnings of radio. Many early radio experimenters succeeded in transmitting speech over distances of a few miles, notably R. A. Fessenden and De Forest in America and Quirino Majorana, Giuseppe Vanni and V. Poulsen in Europe. Several essentials for practical application were lacking, however; to wit, a practicable generator of continuous high-frequency waves, a means for modulating these waves in accordance with speech and a receiving amplifier for revivifying the waves after enfeeblement in transit. In the main, it was the vacuum tube that provided the solution for all these problems.

In the same year, 1915, when telephone service across the U. S. was begun, intelligible speech was experimentally transmitted by radio from Arlington, Va., to Hawaii and to Paris. In 1927 the first commercial overseas radiotelephone circuit was opened between the U. S. and England. Service from Berlin to Buenos Aires, Arg., was begun in 1928, from England to South America and Australia in 1930, and service between the U. S. and South America, Central America, the Hawaiian Islands, the Philippine Islands, the Netherlands Indies and Japan during the years 1930 to 1934. At the end of 1956 radiotelephone service was available between all principal countries not connected by wire and in addition was used in many instances to supplement wire facilities. There were more than 1,600 calls a day between North America and Europe including radio and submarine cable facilities. Fig. 5 shows the principal overseas telephone connections. Most of the overseas radio-

Fig. 12 — Overseas telephone switchboard room, London, England. (By courtesy of Her Majesty's Postmaster General.)

telephone circuits were in the shortwave range from 3 to 30 mc., where long-distance transmission is accomplished by reflecting the waves from the ionosphere. The performance of the radio circuits, although fairly satisfactory, was subject to some difficulty due to atmospheric disturbances.

Radio systems that make use of scattering from the tropospheric layer in the atmosphere to transmit wide frequency bands, capable of handling a number of telephone channels, to distances "beyond the horizon" were introduced in the mid-1950s. A Miami-Havana system went into service in 1957.

Starting with service to the steamship "Leviathan" in 1929, radiotelephone service has been extended to many large ocean-going ships. Thousands of smaller ships in coastal waters, on lakes, in harbors and on rivers are equipped for connection to shore telephone stations by radio. Commercial radiotelephone service between the public telephone system and motor vehicles in cities and on highways began in the U. S. in 1946. In October 1957 there were about 18,000 such mobile stations.

In addition, about 500,000 other vehicles were provided with radiotelephone sets for private services such as those of airlines, police, taxicabs, etc., and there were 16 mobile radio stations for public service on trains. Radio systems were used also for bridging short water gaps and crossing other difficult terrain. Portable radiotelephone equipment
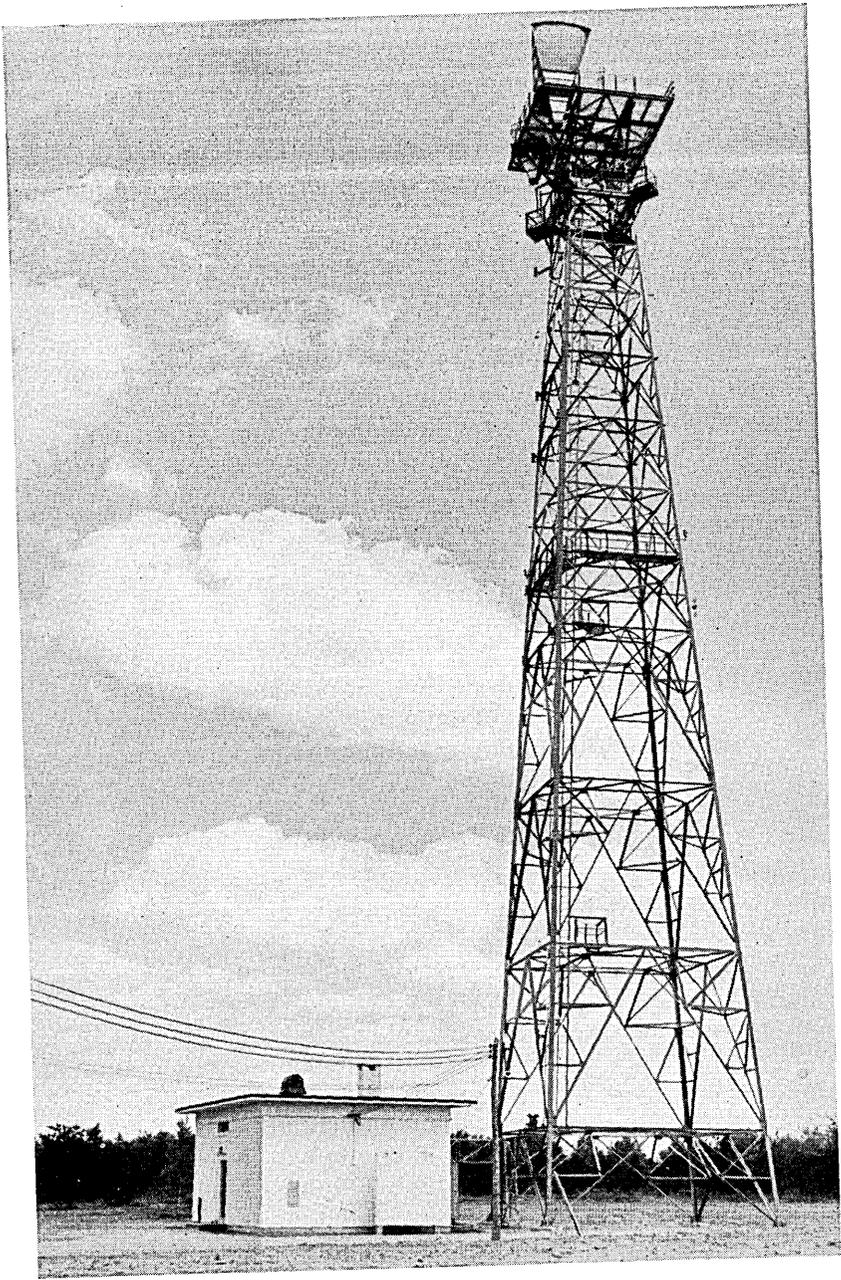
Fig. 13 — Microwave radio relay tower, Dawson, N.S.

was employed in emergencies such as floods or hurricanes to bridge gaps in the regular toll circuits until normal service could be restored. Radio paging systems were available in which a small receiving set carried on the person was used to summon an individual to the nearest telephone in order to place a call to his home or office.

Repeatered radio systems for obtaining substantial numbers of long-distance telephone circuits, as well as broad frequency bands for television transmission, are used in many parts of the world. These commonly transmit radio waves of thousands of megacycles, known as microwaves, over line-of-sight paths. As early as 1934 a microwave radio system developed by A. G. Clavier and others was employed for transmission across the English channel. The large-scale development during World War II of microwave techniques for radar use provided a powerful stimulus in the postwar development of microwave communication systems. A microwave radio-relay (i.e., repeatered) system was placed in service between New York City and Boston in 1947, and one across the United States in 1951. This latter system provides six radio channels in each direction, and each radio channel affords either 600 telephone circuits or a single television circuit.

Concentration of large numbers of circuits on a single route by means of carrier systems, coaxial systems or radio-relay systems yields large economic advantages through sharing common elements of cost including right of way, line conductors, installation, line maintenance, radio towers, power supply, etc. Combining the requirements for both telephony and television on a common route affords further economy.

*Submarine Telephony*

Submarine cables for transmitting telegraph signals antedated the invention of the telephone. Many years went by, however, before long submarine cables suitable for telephony became practical. As early as 1891 a cable containing four wires was laid under the English channel, between St. Margaret's Bay, Eng., and Sangatte, Fr. During the next five years several cables of similar construction, suitable only for shallow water and relatively short distances, were placed beneath the channel.

When cable is laid in deep water, the high pressure necessitates a different type of cable, the usual construction for deepwater telephone cables consisting of a central conductor surrounded by insulation around which is a return conductor. The first such cables were laid in 1921 between Key West, Fla., and Havana, Cuba, a distance of more than 100 miles. Each of these cables provided a single voice circuit and four telegraph circuits. Two years later the first long submarine cables

adapted for carrier telephone transmission were placed across the 40-mile stretch between the California coast and Santa Catalina Island.

For spanning greater distances with the wide frequency band required to accommodate a number of telephone channels, intermediate repeaters are needed. The first submerged telephone repeaters were applied between Anglesey, Wales, and the Isle of Man in 1943 and between Lowestoft, Eng., and Borkum, Ger., in 1946. These repeaters were adapted only for shallow-water operation. Repeaters designed for deep-water use were included in a pair of cables placed in 1950 between Key West and Havana, the two cables giving a total of 24 telephone circuits.

To handle the expanding requirements for transatlantic telephone service, installation of the first transoceanic telephone cable system, between Clarenville, Nfd., and Oban, Scot., with an extension from Newfoundland to Nova Scotia and a radio-relay link to Portland, Me., was begun in 1955 and the system was put in service in 1956. This was a joint project of the American Telephone and Telegraph Company, the British post office and the Canadian Overseas Telecommunication Corporation. The two cables on the main crossing were designed to have 52 submerged repeaters each, spaced at approximately 40-mile intervals, and to provide a total of 36 telephone circuits. The vacuum tube repeaters were designed to operate continuously and flawlessly, with no attention for at least 20 years, at depths up to 2,000 fathoms.

By the end of 1957 similar submarine systems between Port Angeles, Wash., and Ketchikan, Alsk., and between California and Hawaii were in service.

*Information Theory*

Progress in telephony and allied arts has been greatly furthered by the formulation of a comprehensive theory underlying the communication of information. This theory, variously referred to as information theory or communication theory, has resulted from studies by Claude E. Shannon, Norbert Wiener and others, following earlier work by R. V. L. Hartley. It states in essence that information of the kind contained in messages transmitted over communication systems is measurable; as a consequence, the loss of information caused by unpredictable perturbations (noise) introduced during transmission can be evaluated. The basic unit of measure is a "yes-or-no" choice, which is called a bit (short for binary digit). All information can for communication purposes be expressed as, or encoded into, sequences of on-or-off (*i.e.*, binary) pulses. The development of the basic mathematical theory has made it
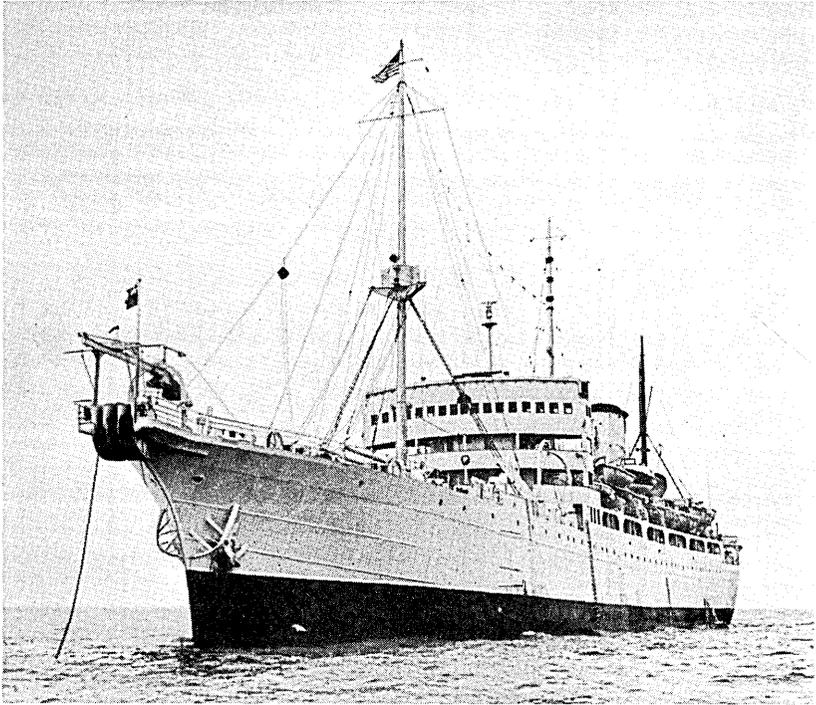
Fig. 14 — H.M.T.S. *Monarch*, owned and operated by the British post office, used in laying the first transatlantic telephone cable.

possible to treat transmission factors such as band width, noise, distortion and the like, and the relations among them, in quantitative terms. This facilitates the comparison of different transmission systems—for example, systems using different forms of modulation, different coding, etc. The theory also establishes absolute upper limits for the rate at which information can be transmitted over systems of different kinds.

*Switching*

Establishing a connection between any two telephones out of a large group is a complicated process. Even in the simplest situation, where both telephones are served by the same switchboard, it is necessary to: (1) observe that a customer wishes to make a call; (2) connect the operator (or switching mechanism) to his line; (3) determine what telephone he wishes to be connected with; (4) select a speech path between them which is not already in use; (5) determine whether the wanted telephone is idle or busy; (6) if idle, ring the bell, or if busy,

inform the calling customer; (7) determine when the call has ended; and (8) restore all equipment to its quiescent state, in readiness for other calls. If the telephones are served by different switchboards, perhaps in widely separated cites, it is also necessary to determine what switchboard serves the wanted telephone and how it may be reached.

From the beginning, the importance of the switching function in telephony was recognized, but the means initially employed were primitive. The first telephone switching arrangement was provided by the Holmes Electric Company at Boston in 1877, by using for telephone connections in the daytime the line wires and plug-and-block connectors employed for a burglar alarm system at night. A commercial telephone switchboard placed in operation at New Haven, Conn., in 1878 served 21 stations on 8 grounded lines. In contrast, a modern central office may serve as many as 100,000 telephones on 50,000 lines.

The New Haven exchange, and other primitive ones, provided a separate switch for each combination of lines that might have to be connected together. For 8 lines this required only 28 switches; but the number of switches increased much more rapidly than the number of customers and quickly became impracticable. With 100 lines, for example, 4,950 switches would have been required. To get around this difficulty, the cord circuit was introduced in 1880. Each line was terminated on the switchboard in a socket (called a jack), and a number of short flexible circuits (called cords) with a plug on each end were also provided. Two lines could thus be interconnected by inserting the two ends of a cord in the appropriate jacks. This was an efficient system as long as the number of calls passing through a switchboard could be handled by a single operator.

Another early invention, which permitted much larger volumes of traffic to be handled efficiently on a manual basis, was the multiple system devised by Leroy B. Firman in 1878. With this, each customer's line is connected to a number of jacks, placed at suitable intervals along the switchboard, so that one is within the reach of every operator.

With the introduction of the multiple system, it became necessary to develop a busy test to determine whether or not a line with which connection was desired was already in use through a connection made at some other part of the switchboard. In a manual switchboard the operator performs this test by touching the tip of a connecting plug to a conducting sleeve forming part of the jack of the desired line. If the line is busy, a click is heard in the operator's receiver.

In 1895 the magnetically operated drop signals which informed the operator of a customer's desire for service were first replaced by incandescent lamps. These were more reliable and smaller than the drop

signals, and permitted a much more compact arrangement of the sub-
scribers' jacks.

Since then, there has been continuous improvement to adapt manual
switchboards to the growing volume and complexity of telephone traffic
and to modern standards of workers' comfort and of customers' con-
venience. There have also been improvements in the associated circuits.
But the design of all subsequent manual switchboards has been based
upon the plug-ended cord, the multiple, the busy test and lamp signals.

The idea of fully automatic switching also appeared quite early.
In fact the first patent for an automatic switching system was issued to
Daniel Connolly, T. A. Connolly and T. J. McTighe in 1879, only two
years after the first primitive switchboard in Boston. This system never
achieved commercial success. In 1889 Almon B. Strowger invented an
automatic system which was installed at La Porte, Ind., in 1892, the
first commercial automatic exchange in the world. The system was
subsequently developed by Alexander E. Keith and other engineers
of the Automatic Electric Company into a form which is still extensively
used throughout the world under the names step-by-step or Strowger
system. In this system the switching mechanisms are operated directly
by pulses generated at a customer's instrument. Originally a cus-
tomer operated a push button to produce the switching pulses, but this
and other types of calling devices gradually gave way to the dial mech-
anism now in general use. This was invented in 1896 by A. E. Keith,
C. J. Erickson and John Erickson of the Automatic Electric Company.

The basic mechanism around which the Strowger system is built
is the Strowger switch, which is sometimes called a connector and some-
times a selector according to the use to which it is put. In complete
form this consists essentially of two parts: a ten-by-ten array of ter-
minals (the bank) arranged in a cylindrical arc; and a movable switch
(the brush) which is translated along the axis of the cylinder by one
ratchet mechanism and rotated about it by another, so that it can be
brought to the position of any one of the 100 terminals. Each ratchet
mechanism is driven by an electromagnet, which can respond to the
pulses produced by a telephone dial.

Many other switching systems have been invented, notably in the
United States, Germany and Sweden, and a number of them have
been put into successful commercial operation. Some of these have
operated on the step-by-step principle but have used different types
of apparatus. Others have operated on the principle of common control,
in which pulses are stored for a short time in a device which then controls
the switches either directly or through some intermediate mechanism.

The distinguishing feature of such systems is that the common-control
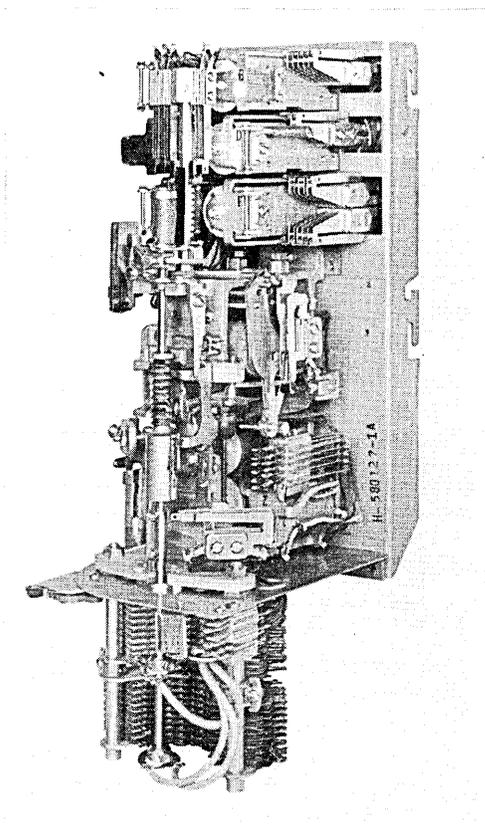
Fig. 15 — Step-by-step selector (United States). (By courtesy of Automatic Electric Co.)

mechanisms (known as registers, senders, translators, directors, markers, etc.) are not assigned to the customer for the duration of his call, but are used only as long as they are needed and are then free to serve other customers. Thus each serves many calls per hour, and few are required; it is therefore practicable to provide complicated devices which can perform a variety of useful functions, but which would be too expensive to assign for the duration of the call. Systems using common control have great flexibility and efficiency in the use of trunk groups, and are especially advantageous for large exchanges and for auto-matic routing of long-distance calls.

The earliest common-control systems were developed by the Western Electric Company, primarily to meet the needs of large metropolitan centers. In a program beginning in 1906, two systems were developed
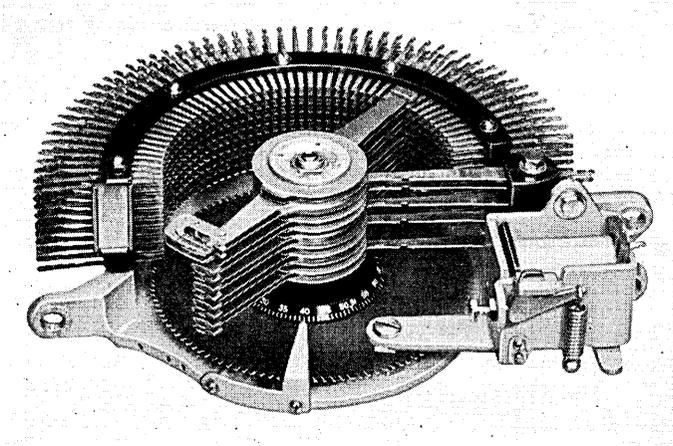
Fig. 16 — A rotary selector (Belgium). (By courtesy of International Telephone and Telegraph Co.)

alongside one another, and both proved to be successful. One, called the rotary system, was first put into commercial service in England in 1914. The other, called the panel system, went into commercial use at Newark, N. J., in January of the following year. Common-control principles were also applied to the step-by-step switch by the Automatic Electric Company, and the resulting system was adopted by the British post office in 1922 for use in London. The most recent common-control systems have been crossbar systems, of which several commercial types have been developed by the Bell System in the United States, by the Swedish Telecommunication admistration and the L. M. Ericsson Company in Sweden and by subsidiary companies of the International Telephone and Telegraph Company in Belgium and Germany.

The most important concepts in the evolution of the modern types of crossbar exchange were probably translation, the sender, the marker, the crossbar switch and the principle of call-back operation.

Translation, invented by E. C. Molina of the American Telephone and Telegraph Company in 1906, makes it possible to convert incoming dial pulses from decimal to nondecimal form and thus affords flexibility and efficiency in the use of trunk groups.

The sender, first used in the rotary system, is essentially an automatic mechanism which generates new dialing signals, either in the code given by the translator or in other appropriate codes.

The basic function of the marker is to make a preliminary test of several alternative paths to a wanted destination through an array of switches, before any of the switches is closed, so as to avoid the pos-

sibility of encountering a busy switch after part of the switching operation has been performed. It was invented by N. G. Palmgren and G. A. Betulander of Sweden in 1912, and was first used commercially in a system manufactured by the Relay Automatic Company of England from 1915 to 1920. As used in crossbar systems of the Bell System, it has been developed into a complex assemblage of electromagnetic relays which serves as the basic control element for the entire switching operation. Among other things, it tests the circuits before connections are established; it seeks out alternate paths when needed; and it reports trouble conditions which may be encountered in its preliminary tests.
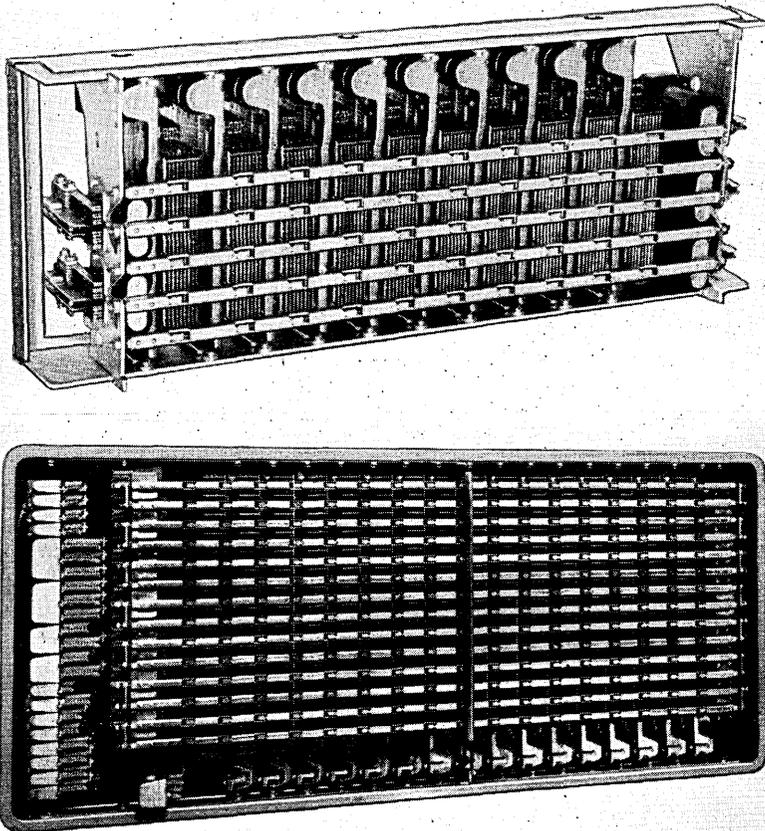
Fig. 17 — *Top:* Crossbar switch, present date design by the Swedish Telecommunications Administration. (By courtesy of Swedish Board of Telecommunications.) *Bottom:* "Pentaconta" crossbar (France). (By courtesy of International Telephone and Telegraph Co.)
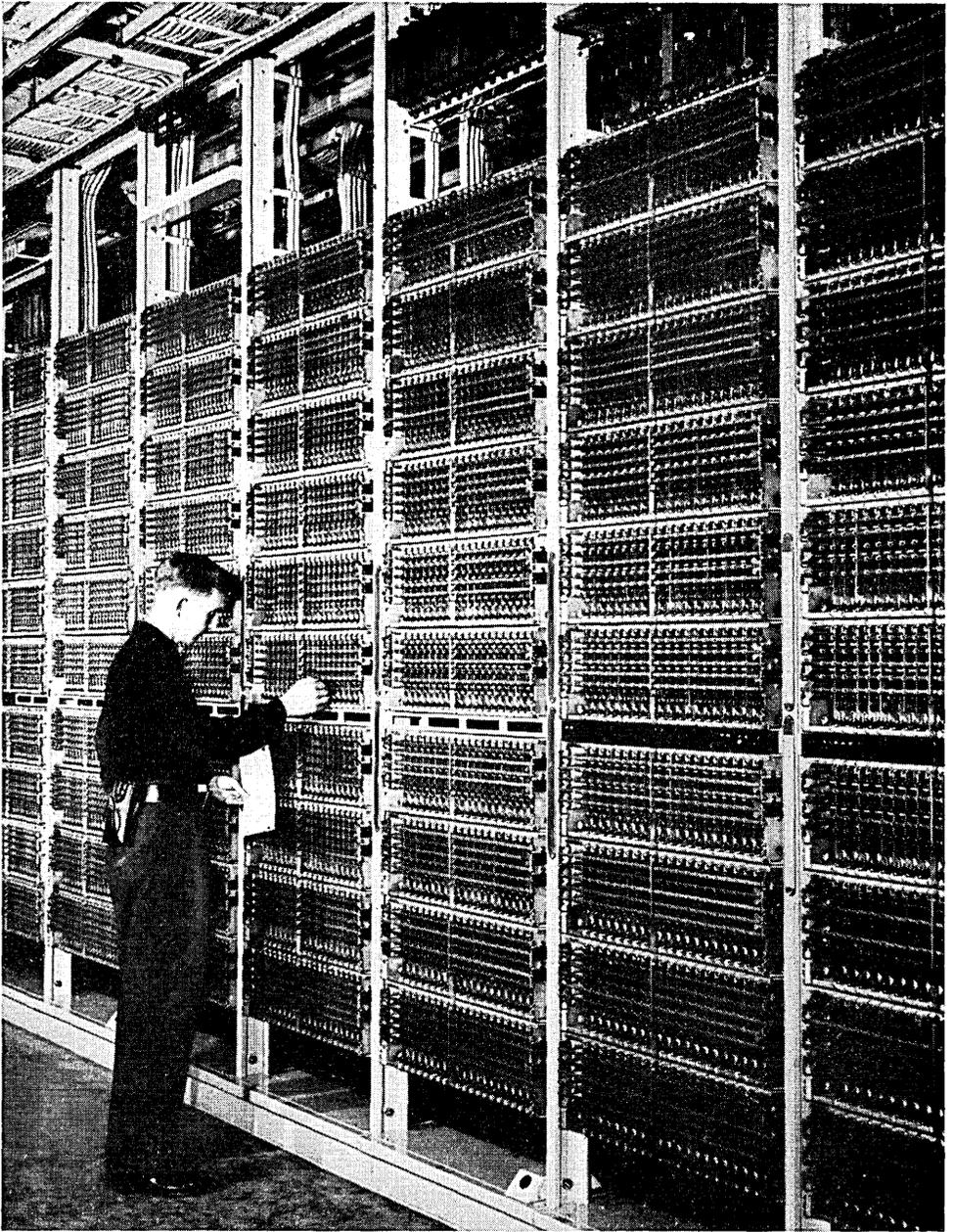
Fig. 18 — Part of a crossbar switching system.

Since it can examine a large number of trunk circuits practically simultaneously, it uses them with great efficiency.

The crossbar switch is essentially a multiple relay structure affording fast operation and reliable contacts of precious metal. Unlike other switches mentioned above, the moving parts have little inertia and move through small distances. It was first conceived in America by Homer J. Roberts of Automatic Electric Company in 1901, and later patented in separate forms by John G. Roberts and John N. Reynolds, both of Western Electric Company. The first satisfactory mechanical design, worked out by Palmgren and Betulander in Sweden in 1919, was used by the Swedish telephone authorities in a commercial exchange in 1926. Aside from a small-scale trial installation in Stockholm in 1919, which was subsequently abandoned, this was the first public use of the crossbar switch. The system in which it was used, however, and all other crossbar exchanges in Sweden until very recent years, operated on step-by-step principles. Most of the crossbar switches used in America and other countries today follow rather closely the Swedish design.

Call back is a principle of operation, invented by Edson L. Erwin in 1938, which has been effectively used in crossbar systems. When a customer originates a call, the register stores not only the wanted number, but also the identity of the calling telephone which is determined automatically. The connection with the calling subscriber is then disconnected, and an entirely new connection established, from a favourable point within the exchange, to both the calling and wanted telephones.

As automatic switching systems were improved, their application was extended until, in the mid-1950s, 77 per cent of the world's telephones were automatic, as compared with 15 per cent only three decades earlier. This metamorphosis occurred not only because automatic operation is faster, more accurate and more economical than manual service, but more basically because in many areas the enormous number of operators necessary to support the rapid telephone growth would have far exceeded the possible supply.

In America the crossbar system, designed originally for metropolitan areas, presented so many advantages as compared with earlier automatic systems that its basic principles were soon applied in other fields. Crossbar tandem equipment, for example, is the modern version of the tandem principle developed originally for manual systems. Through the tandem scheme, traffic between offices on opposite sides of a large metropolitan area is handled through one or more intermediate (tandem) offices which act as clearinghouses for these relatively small amounts of traffic, handling them more efficiently than if direct paths were provided and thus reducing the number of trunk circuits required.

Another type of crossbar system is specially adapted to the automatic switching of long-distance circuits. With this, connections are made on a two-path or four-wire basis, a separate path being used for each direction of transmission so as to obtain superior transmission performance on connections comprising a number of circuit links in tandem.

The first crossbar switching system in the long-distance field, introduced in Philadelphia in 1943, enabled operators but not customers to dial long-distance calls. This advance was followed by another which permitted direct dialing by customers without the aid of an operator. The switching mechanism selects the route to the distant telephone. Should it find all direct circuits busy, it explores in succession as many as five alternate routes, and establishes the connection along the least circuitous one which is available. All this is done automatically. Necessary information regarding direct and alternate routes to the destination is permanently available in a translator, which supplies it to the control circuits as required.

Arrangements for customer dialing of at least some intercity calls were provided in all the principal countries of the world by the mid-1950s. In North America a complete program was worked out for eventually handling substantially all calls in this way. The initial installation of equipment to permit nation-wide dialing by customers was placed in service at Englewood, N. J., in 1951. With this equipment, customers were able to dial directly about 11,000,000 other customers in selected areas as far away as San Francisco. They dial a three-digit area code followed by the digits of the directory listing of the called number. At the end of 1956 this type of service had been extended to 253 originating locations reaching about 30,000,000 customers.

Closely related to automatic switching is the automatic recording of data for preparing a customer's bill. The simplest means for this purpose, used in either manual or automatic offices, is an electromechanical counter, known as a message register, which records the number of calls made by a customer. In a more elaborate arrangement, referred to as multi-unit registration, the register can be operated more than once for a single call, the number of operations depending on the distance and duration of the call. Each customer is billed on a bulk basis for all of the "message units" totaled on his register. This plan is widely used in America, and almost universally elsewhere.

Another method of charging is "automatic ticketing," first introduced in Belgium. With this method, automatic equipment prints for each call a ticket similar to one that might be prepared by an operator.

In America, where distances are great and the tariff structure complicated, a highly versatile type of message registration system, called

automatic message accounting (usually abbreviated to AMA), was developed by the mid-1950s. In this system the information needed for billing calls is recorded in the form of coded holes in paper tape. Automatic processing of the tapes in an accounting center yields bills with any desired amount of detail. The AMA system is extensively used in association with direct distance dialing.

A big step in the improvement of telephone service in outlying areas resulted from automatizing the telephone switching in very small communities. Community dial offices, either of the step-by-step or all-relay type, were provided in many localities. These function with no attendance except for occasional maintenance visits.

Another phase of telephone switching is found in switchboards, known as private branch exchanges, which are located on the customer's
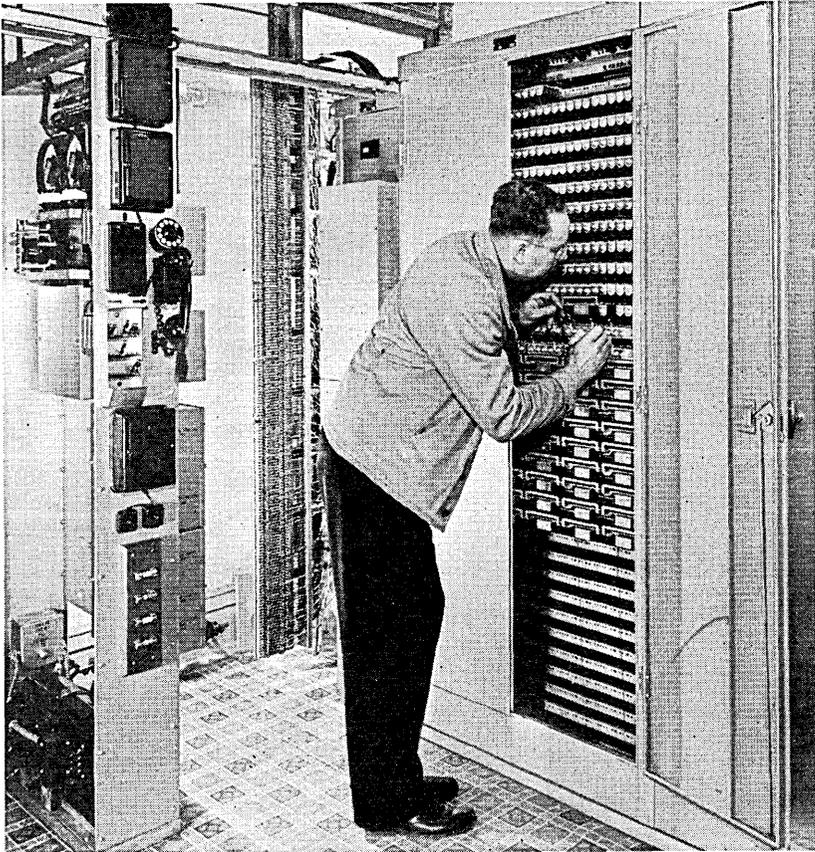


Fig. 19 — All-relay community dial office for telephone service in outlying areas. (By courtesy of North Electric Co.)
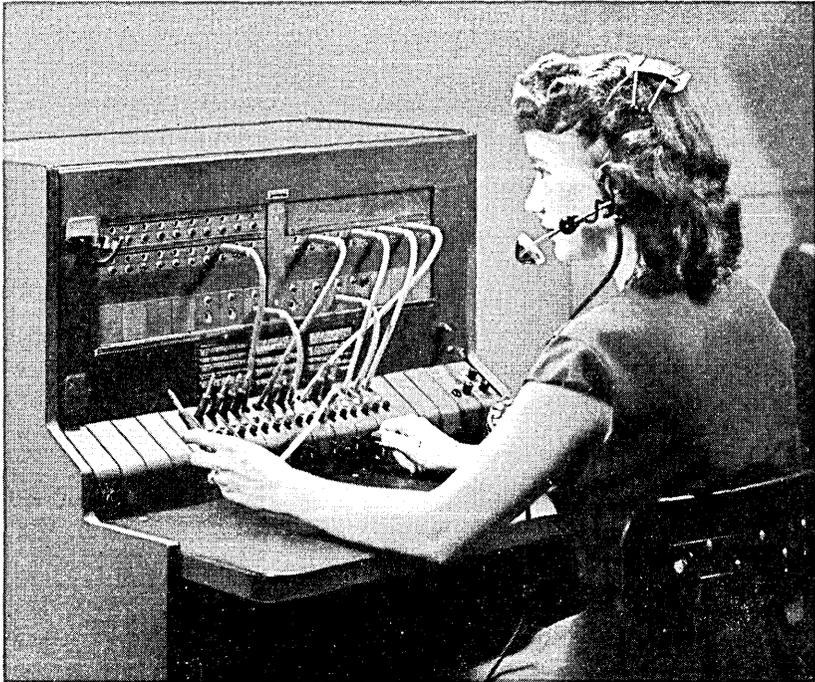
Fig. 20 — Private branch exchange switchboard, capable of handling both internal and central office calls. Used in business offices.

premises. The private branch exchange (abbreviated PBX) serves both for interconnecting the station sets of the customer's establishment and for connection over trunk lines to the central office. The smaller PBX's can be placed on a desk and operated by a person who may perform other duties as well. Large PBX's resemble central office switchboards and may be of either the manual or automatic type.

By the mid-1950s exploratory developments were under way in the principal communications laboratories of many countries with the ultimate objective of replacing electromechanical switching systems by electronic ones. A small beginning had already been made in some units that are partly electronic and partly electromechanical. Among these may be mentioned the line concentrator of the Bell System, an unattended switching unit which permits a small number of lines to be shared by a substantially larger number of customers without the disadvantages of party-line service; and the mechanoelectric system developed by the International Telephone and Telegraph Company in Belgium and first installed at Ski, Nor., in 1954.

The Bell System announced also that a fully electronic system was

under development, and was expected to be placed in commercial service in Morris, Ill.

*Associated Services*

Much of the plant required for telephone service is well adapted to the provision of other types of communication service, so that economies can usually be effected by designing the plant to handle the combined requirements of different services. The earliest of such associations was with the telegraph. Direct current telegraph channels superposed on the telephone wires were leased to private customers and to the companies furnishing message telegraph service, and were employed also for telephone line maintenance. The introduction of carrier transmission methods made it possible to derive as many as 18 telegraph channels from one telephone circuit. Such channels serve both for teletypewriter exchange service, which is a switched service analogous to telephone service, and also for private-line teletypewriter service.

Satisfactory picture transmission requires a frequency band of about the same width as that for telephone conversation. Special networks are provided for picture service, and occasional use is made of regular telephone circuits for this purpose. Telephone circuits are employed also to render facsimile and other services. High-speed transmission of information in the form of pulses is being increasingly applied for both commercial and military purposes, one common application being to supply data to electronic computers.

Generally speaking, the chain networks used for the radio broadcasting of either sound or television programs are provided in conjunction with the telephone plant. For good quality of music reproduction, a frequency band somewhat wider than that adequate for speech is required. Television broadcasting necessitates a very wide frequency band, equivalent to that required for about 1,000 telephone circuits.

Television circuits are derived, commonly in association with large numbers of telephone channels, from coaxial or microwave systems. Wire broadcasting systems, in which sound programs are conveyed by wire directly to the customers' premises, are in limited use.

TRENDS

Technical trends likely to be present in telephony over the years to come were foreshadowed by various developments in progress by the mid-1950s. Foremost in the area of new art was the transistor, invented by W. H. Brattain and John Bardeen in 1948. This is a three-electrode amplifying device employing a semiconductor such as germanium or silicon. A preferred form, known as the junction transistor, was invented

later by William Shockley. The transistor can perform many of the functions previously assigned to vacuum tubes, but is far more efficient because it requires no power to heat a cathode. Additionally, it operates with much lower electrode voltages and is much smaller than a vacuum tube (Fig. 21). Because of the transistor's low power dissipation and low voltage requirements, other electronic components associated with it can be miniaturized. The over-all result is that the size, weight and power consumption of apparatus employing transistors can usually be reduced to a small fraction of that for equivalent apparatus using vacuum tubes.

Transistors therefore held prospect of finding large and varied applications in almost every field of communication, and particularly in telephony and associated services, including not only areas where vacuum tubes were previously used, but many new areas as well. Thus transistors promised economies through more extensive use of amplification and carrier techniques in the local telephone plant.

All of the automatic switching systems described have been built around the electromagnetic relay and other electromagnetically operated devices. The relay dates back to the early part of the 19th century, and in a form much cruder than the present was the basis for the development of telegraphy. It employs magnetic attraction produced by an electric current to move an armature and thus open or close electric contacts so as to perform switching or other operations. Devices of this kind require a minimum of several thousandths of a second to operate. In contrast, electronic devices such as the transistor may be used to perform switching operations with a speed of the order of a few millionths of a second. Not only so, but a more complex operation involving choices based on conditions existing at the instant can be performed in a similar time interval.
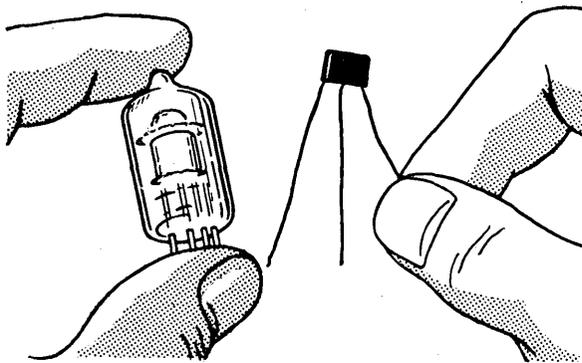


Fig. 21 — Comparative sizes of vacuum tube (left) and transistor.

This means that no longer must a large complement of apparatus be set aside to serve a customer during the entire period of his call, or even a large part of it. The extremely high speed was expected to afford substantial economies in new switching systems through centralizing of functions and time sharing of the apparatus used for establishing connections. The common apparatus would serve a large number of customers in such rapid succession that each one receives the equivalent of continuous and exclusive service. Reduction of space and power requirements would yield further economy. The same speed advantage could be realized with vacuum tube electronics, but with attendant disadvantages in size, power consumption, life and reliability. Altogether, the new art of solid state electronics provided the basis for a revolution in both switching and transmission technology.

The introduction of intercontinental submarine cables, not only of the type being installed in the mid-1950s but also with new apparatus capable of handling larger numbers of telephone channels, was expected to be of particular importance to world communication. Television transmission across ocean barriers was likewise a possibility. Research under way on the use of hollow wave-guide conductors as transmission lines held in store new long-distance systems that would transmit extremely wide bands of frequencies, of the order of thousands of megacycles, able to provide a great multiplicity of communication channels.

As to telephone service in general, there were, in the mid-1950s, no indications of saturation in demand, even in highly developed areas. Substantial further growth in number of telephones and their utilization was therefore in prospect. Large expansion in the services associated with telephony could likewise be envisioned. More fundamentally, it was anticipated that improved communications would contribute greatly to the development of culture and understanding among peoples.

BIBLIOGRAPHY

1. L. W. Taylor, The Untold Story of the Telephone, Am. Phys. Teacher, **5**, pp. 243–251 (1937). This contains an account of Elisha Gray's invention.
2. D. E. Hughes, On the Action of Sonorous Vibrations in Varying the Force of an Electric Current, Telegraphic J., **6**, p. 201, May 15, 1878.
3. Circuit Court of the U. S., District of Massachusetts, 1879 (Dowd Case Record).
4. Count T. A. L. Du Moncel, *The Telephone, Microphone and Phonograph*, New York, 1879.
5. S. P. Thompson, *Philipp Reis: Inventor of the Telephone*, London, 1883.
6. T. Schwartz, *Telephon, Mikrophon und Radiphon*, Vienna, 1883.
7. G. B. Prescott, *Bell's Electric Speaking Telephone*, New York, 1884.
8. Supreme Court of the U. S., October Term, 1887, Book 31, **126**.
9. H. N. Casson, *The History of the Telephone*, Chicago, 1910.
10. A. N. Holcombe, *Public Ownership of Telephones on the Continent of Europe*, Cambridge, Mass., 1911.
11. A. B. Smith, *Telephony Including Automatic Switching*, Chicago, 1924.

12. F. L. Rhodes, *The Beginnings of Telephony*, New York, 1929.
13. H. A. Frederick, The Development of the Microphone, Bell Tel. Quart., **10**, p. 164, April, 1931.
14. W. Aitken, *Who Invented the Telephone?*, London, 1939.
15. A. W. Page, *The Bell Telephone System*, New York, 1941.
16. A. L. Albert, *Electrical Communication*, 3rd ed., New York, London, 1950.
17. E. Horstmann, *75 Jahre Fernsprecher in Deutschland*, Frankfurt, 1952.
18. A. B. Clark, The Development of Telephony in the United States, Trans. A.I.E.E., **71**, p. 348, 1952.
19. O. Myers, Common Control Switching Systems, B.S.T.J., **31**, p. 1085, Nov., 1952.
20. H. Sterky, The First Century of Swedish Telecommunications and What We Can Learn From It, Tele (English ed.), no. **1**, p. 1, 1954.
21. F. E. Terman, *Electronic and Radio Engineering*, 4th ed., New York, 1955.
22. M. J. Kelly, Sir G. Radley, G. W. Gilman and R. J. Halsey, A Transatlantic Telephone Cable, Communication and Electronics, **17**, p. 124, March, 1955.

# Tone Ringing and Pushbutton Calling

## Two Integrated Exploratory Developments

### By L. A. MEACHAM, J. R. POWER and F. WEST

*Forward-looking work on the telephone set has yielded attractive new solutions for the problem of signaling between the customer station and the central office. This work takes advantage of new electronic devices such as the silicon junction diode, the transistor and the ferrite coil. Two exploratory developments are described: a "tone ringer" and a "pushbutton caller," both of which employ signals within the telephone speech band. Although usable separately, as justified by economic and other considerations in particular telephone systems, the developments have been conceived primarily as an integrated and moderately long range attack on the signaling problem, with special regard for the field of electronic switching.*

I. GENERAL

1.1 *Introduction*

The rapid advances of recent years in the development of electronic techniques and components have brought about both a need and an unusual opportunity for new station apparatus.

The need has appeared unmistakably in connection with the development of electronic switching systems,[1] themselves outgrowths of the new technology. To such systems, the relatively large currents and voltages that are associated with signaling functions in existing telephones would represent formidable design difficulties. For example, the conventional ringer draws nearly a watt from a 90-volt, 20-cycle generator, whereas the design of electronic switching networks is properly focused upon handling speech signals that are far smaller, of the order of a volt or a milliwatt. Roughly parallel remarks might be made with regard to dialing, to signals for controlling coin telephones, and even to the direct currents sent over the lines to power speech transmitters. Electronic switching clearly calls for a fresh attack upon station problems.

In a preliminary broad study of this need and opportunity, it was

recognized that exploratory development should strive toward certain long-term goals. For instance, it would be desirable to use the same frequency band as speech for signaling functions, thereby avoiding the need to transmit more than the one band; to provide party lines with full-selective ringing and with automatic party identification to facilitate automatic message accounting;[2] to eliminate the use of a "station ground" for purposes other than lightning protection, thus doing away with a troublesome source of transmission noise; and to replace the rotary dial with pushbuttons for the sake of speed, convenience and other benefits. Of course the actual adoption of any such changes must always depend upon a vast number of detailed considerations, human, technical, administrative, and economic, which are beyond the scope of this paper.

This paper describes two integrated developments that have grown out of the study — new arrangements for signaling to and from the customer's station. The "tone ringer" and "pushbutton caller" are closely related in that they both employ multi-frequency pulsing,[3] and in fact share a common series of signaling frequencies in the speech band chosen for their mutual benefit. They also share circuit components that can be transferred by switchhook contacts from one function to the other, and in numerous design details they profit from their interrelationship. An experimental telephone set of the 500 type, modified to include both the pushbutton caller and the tone ringer, is shown in Fig. 1.



Fig. 1 — Modified 500-type telephone set equipped experimentally with tone ringer and pushbutton caller.

1.2 *Signaling Plan*

One of the first problems to arise in this exploration was the choice between frequency division and time division or pulse coding of some sort. From the viewpoint of information theory, either ringing or calling involves only a few bits of information, and could be accomplished very rapidly and efficiently with time-division pulses in the speech band.[4] But on the local telephone line as it exists and will probably remain for a substantial period, information-carrying capacity is not at a premium during either ringing or calling; in fact this capacity may be spent rather liberally to minimize the cost and complexity of terminal apparatus. This is the basic argument behind the selection of multi-frequency signaling for the developments here described.

The tone ringer performs full-selective ringing of up to eight parties on the basis of a "1-out-of-8" code. Specifically, each ringer is arranged to be tuned upon installation to an assigned frequency; it will then respond to one and only one of eight alternative ringing tones sent from the central office. Each tone is a simple sine wave, interrupted about twelve times per second to give the resulting sound a distinctive character, and turned on and off at intervals of several seconds as in conventional ringing. The signal selected by the tuned circuit might conceivably be arranged to turn on an oscillator or other local sound generator in the station set, but here it is used directly, being amplified by a transistor and applied to a small sound radiator mounted in the base of the set. With the transistor biased for "Class C" operation, selectivity requirements can be met (with the moderate values of Q afforded by practicable coils in the audio range) at frequency spacings in the neighborhood of 10 per cent.

In the pushbutton calling arrangement under consideration, number information is sent to the central office in the form of oscillatory pulses, using a 1-out-of-10 code. As will be seen later, party identification is also provided. In earlier studies using reeds, a 2-out-of-5 code was employed; that is, five reeds with different frequencies were provided and two of them were plucked for each digit of the called telephone number. As compared with using a 1-out-of-10 code, this saved five reeds. But in the case of the tuned circuit, a single tapped coil can generate any reasonable number of frequencies, one at a time. The latter code therefore becomes attractive, because 2-out-of-5 would involve either a pair of tuned circuits (exclusive of party identification) or the difficulties of pulsing twice in sequence at different frequencies for each pushbutton

operation. Chiefly for this reason, ten frequencies are employed for the ten decimal digits.

In addition, party identification is desired. The resonant circuit of the ringer, tuned to a different frequency for each party, is fortunately well-suited for this purpose. It is therefore switched from the ringing to the calling function when the handset is lifted, with no change in its party tuning. Accordingly, when any pushbutton is pressed two simultaneous pulses of different frequency are generated. One, the "party pulse," is produced in the tuned circuit transferred from the ringer; the other, the "digit pulse," comes from a separate coil, differently tuned for each of ten pushbuttons.

1.3 *Frequency Plan*

From the foregoing it is clear that eighteen frequencies are required, eight for ringing and party identification and ten for digit transmission. In urban systems, only four ringing frequencies are likely to be needed but in this case the others may be found useful for special signaling functions such as those associated with pay stations.

In existing operator multi-frequency key-pulsing systems a set of six frequencies is used, spread between 700 and 1,700 cps at uniform intervals of 200 cps. If this series were to be extended to as many as 18 frequencies, it would cover, say, from 300 to 3,700 cps, and the steps would range from 66.7 to only 5.4 per cent. With such a frequency allocation the total bandwidth is excessive, the percentage spacing is extravagantly wide at low frequencies, and the spacing is too close at the high end to be compatible with inexpensive electrical tuning elements.

Geometric spacing was selected as the logical alternative, the set of frequencies being given by

$$f_n = a^{n-1} f_1 . \qquad\qquad (1)$$

Here the constant percentage spacing is simply $100 \, (a-1)$. This series fits in naturally with the use of tuned circuits having uniform values of $Q$ and uniform frequency stabilities — properties that are easy to attain. The transient rise and decay times associated with constant-$Q$ tuned circuits vary inversely with their resonant frequencies, but these time differences have not led to serious difficulties.

The geometric series has an important advantage over other possible frequency distributions, in that by carefully choosing the factor $a$ it is possible to make both the second and third harmonic of each frequency

fall approximately midway between two adjacent higher frequencies in the same series. This interlace is helpful in reducing the possibility of interference from non-linear distortion.

A method of choosing an optimum value for $a$ is illustrated in Fig. 2. This shows the relative frequency of each term in a geometric series $f_1$, $f_2 \cdots f_{18}$ as a function of the spacing. Harmonics of $f_1$ up to the sixth are also displayed on the logarithmic frequency scale. Points where a harmonic coincides with any frequency $f_n$ are marked by circles, and intervening points where a harmonic falls geometrically midway between $f_n$ and $f_{n+1}$ are shown as dots. An ideal spacing factor $a$, if one existed, would be represented as a vertical straight line passing through a dot for every harmonic. Actually, the dot alignment is not so fortunate, but two spacings do exist (near 8.5 per cent and 11.1 per cent) for which the
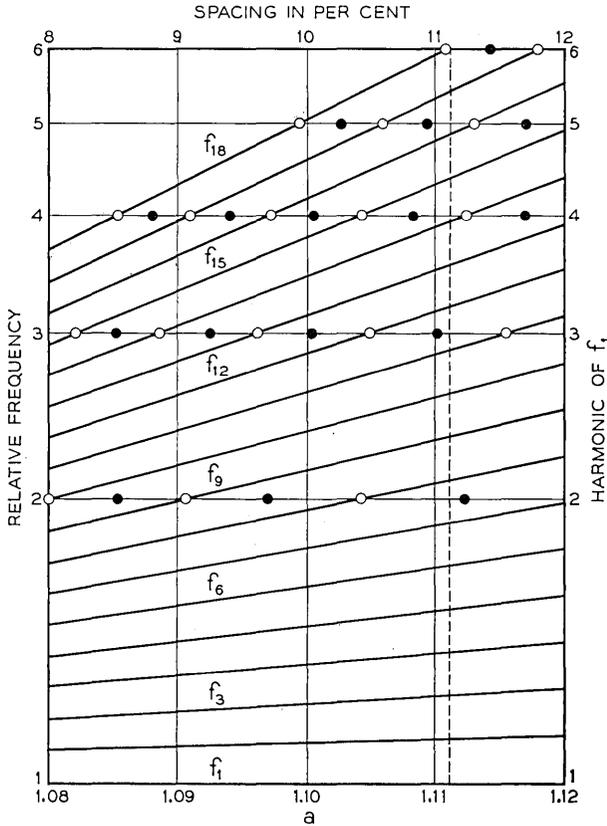
Fig. 2 — Interlacing of harmonics in geometric series of frequencies.

second and third harmonics are both fairly well interlaced. These two harmonics are of primary interest because in the system as planned all higher ones are not only weaker, but also fall above the frequency range of critical interest (the portion of the band devoted to the same function — party or digit). Since a larger spacing was preferred for the sake of selectivity, a figure of 11.11··· per cent was finally chosen, corresponding to a ratio of 10/9 between adjacent frequencies. This happens to give a range of almost exactly 6 to 1 for the whole set of eighteen. The spacing is indicated by a dashed line in Fig. 2.

Among many considerations affecting the choice of specific frequencies may be mentioned transmission, tuning elements, the acoustic radiator, and the sound of the ringing tone. The series used in the development extends from 478 to 1,000 cycles for ringing (also party identification) and from 1,111 to 2,868 cycles for calling. The lower frequencies were assigned to ringing chiefly because the resulting sounds were found to be more pleasing. As will be seen, strong harmonics of the fundamental ringing frequency are generated in the tone ringer, and such of these as fall within the pass band of the acoustic radiator form the actual sound output.

## II. TONE RINGER

### 2.1 Basic Considerations

In addition to the fundamental choices affecting the tone ringer that have been presented in foregoing sections, a few others should be mentioned. A primary requirement of the sound of a ringer is that it must be effective (not readily confused with other common sounds) as well as pleasant. It must attract attention when the customer is in the far reaches of house or garden but must not be annoyingly loud or harsh when he is nearby. High pitched tones are not suitable for this purpose because they are shrill and because they are rapidly attenuated as they travel through passageways and around corners. On the other hand, low pitched tones are pleasant and carry well but require a large sound radiator and are likely to be masked by background noise. The chosen compromise is a complex tone in the region of 900 to 2,400 cycles per second. This tone is composed almost entirely of harmonics of the selected ringing frequency; in general, the fundamental lies below the band efficiently radiated. The result is a distinctive richness of quality, with the ear effectively supplying the missing fundamental. The tone is made still more distinctive by interrupting it about 12 times per second. A secondary requirement of the tone is that its loudness should

be reasonably independent of circuit and environmental conditions but that volume control with about 15-db range should be available to the customer.

A problem that arises when ringers are operated in the frequency and power range of speech is that of "talk-off" protection; i.e., the ringer must be prevented from responding to speech or other in-band voltages that appear on the line during conversation. This problem also is believed to have been effectively solved, in a manner to be explained.

### 2.2 Tone Ringer Circuit

The tone ringer circuit consists of four main elements — an input limiter, a frequency-selective network, an amplitude-selective output amplifier and a sound radiator — as indicated in the simplified schematic of Fig. 3. A volume control and a dc power filter are also included. Many variations in the details of circuitry have been explored; the example shown is one developed for the initial trial of electronic switching.

The voice frequency signal voltage at the customer's end of the line may have any value between 0.5 and 2.5 volts. It is the primary function of the input limiter to convert this variable voltage to a constant-current drive for the frequency-selective elements. Some amplification is also provided. The limiter draws a constant direct current of 100 microamperes plus the leakage current, $I_{co}$, of the transistor. The ac output is a symmetrically clipped sine wave that approaches a square wave, its peak-to-peak value being limited to the 100 microamperes drawn from the line.
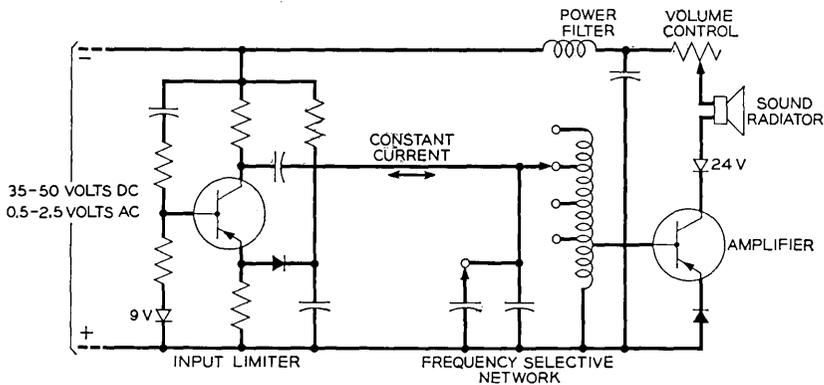


Fig. 3 — Simplified circuit schematic of the tone ringer.

The frequency selective network is simply a parallel resonant circuit consisting of one coil with four taps and either one or two capacitors. Eight resonant frequencies in the range of 478 to 1,000 cycles per second are available by various tap-capacitor combinations, an adjustment readily made by the installer. The inductor has a ferrite core consisting of two molded cups with a cylindrical coil nested inside them. This design is simple and provides stability, small size and reasonably high $Q$. In addition, the inductance can be adjusted during assembly of the set to any value within ±10 per cent of the nominal value by rotating one half of the core; thus, accuracy requirements for the tuning capacitors are reduced.

The constant current from the input limiter flows through the resonant circuit. At or near resonance the circuit impedance is high and a correspondingly high voltage is developed, which drives the power amplifier. Away from resonance this voltage is small.

The output amplifier consists of a transistor in a common emitter Class C circuit, the bias being provided by the forward voltage drop of a silicon junction diode in series with the emitter. Such a diode changes quite sharply from a very high impedance to a low one at a forward potential of about 0.6 volt. This circuit has many advantages. A principal one is that it does not respond until the input voltage from the resonant circuit exceeds the bias voltage and a small increase in voltage above this threshold value drives the transistor to its maximum output, i.e., to overload. This effect not only sharpens the selectivity of the tone ringer but also holds the output constant over a frequency band of several percent centered about resonance, thus providing a margin for small inaccuracies in central office and resonant frequencies. Another advantage is that the input impedance of this amplifier is high, up to the overload point, and does not seriously lower the $Q$ of the resonant circuit. Still another, the amplifier output consists of current pulses which are rich in harmonics of the driving frequency. It is these harmonics that are reproduced by the sound radiator. A final advantage is that, when not operating, the amplifier does not draw any direct current except the $I_{co}$ of the transistor — a few microamperes.

A second silicon diode, having an avalanche breakdown at about 24 volts, is connected in series with the collector, which it thus holds at a value some 24 volts less than the dc line voltage; that is, at not more than 26 volts. This is within the permissible range of presently available transistors. The diode also provides talk-off protection as will be described later.

The sound radiator, shown in Fig. 4, consists of a ring-armature
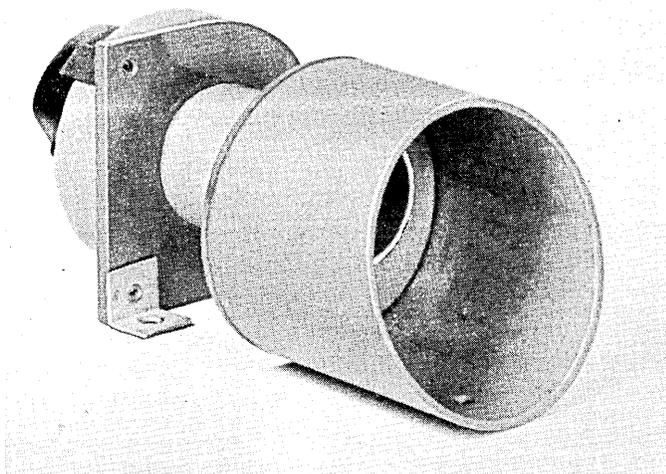
Fig. 4 — Sound radiator of the tone ringer.

telephone receiver[5] and double-tube resonator. The usual acoustic damping of the receiver is omitted, giving the response a resonant peak at 1,500 cycles per second, and the resonator introduces other peaks at 950 and 2,200 cycles per second. This results in efficient electro-acoustic transformation in the frequency band from 900 to 2,400 cycles per second and in very little response outside that band. The characteristic curve is shown in Fig. 5. The figure also shows the frequencies present in the pulses supplied to the radiator when the tone ringer is driven at its lowest frequency (478 cps). This illustrates the dominance of harmonics in the sound output. Volume control is provided by means of an adjustable resistance in series with the sound radiator.

A low-pass filter must be interposed between the line and the output circuit. This has the dual function of maintaining a high ac input imped- ance and of minimizing feedback from the output circuit to the line. A series inductor and shunt capacitor are adequate for these purposes.

One of the previously mentioned basic requirements of the tone ringer is that it must not operate on speech voltages. One way of assuring this is to disable all ringers on a line by means of the reduction in dc line voltage that occurs whenever any associated handset is removed from its cradle. This drop is from nearly 50 volts to less than 24 volts. Since the latter value is not sufficient to break down the diode in series with the collector of the amplifier transistor, the diode becomes essentially

"open circuit". This prevents the flow of any current in the sound radiator, and so the ringer remains silent.

### 2.3 *Performance*

The tone ringer will operate over as much as five miles of 26-gauge cable (2,200 ohms loop resistance) when the voice frequency ringing power is +6 dbm at the central office end of the line. It requires less than 1.5 ma of direct current per operating ringer from a 50-volt office battery.

About 0.5 milliwatt of sound power is generated by a tone ringer operating at full volume on a short loop, which is roughly the same as for a conventional ringer in a 500-type set under the same conditions. When several ringers are operated simultaneously at the end of a long loop, their individual outputs may be reduced as much as 4 db. Tests conducted in the laboratory have indicated that the carrying power and attention attracting qualities of the ringer tone are actually better than those of the conventional bell and that the tones are considered pleasant by most people and acceptable by nearly all the others. These results have been substantiated by a field trial of a few tone ringers on rural lines at Americus, Georgia, and by a more extensive trial at Crystal Lake, Illinois.

The frequency selectivity and sensitivity of the ringer are illustrated
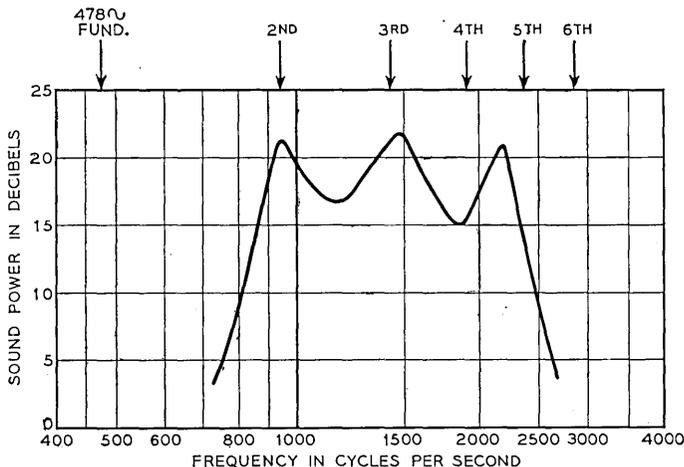


Fig. 5 — Frequency response of the sound radiator, and harmonic spectrum of a typical party frequency. Sound power is given in decibels relative to 1 microwatt for 1 milliwatt of available electrical input.
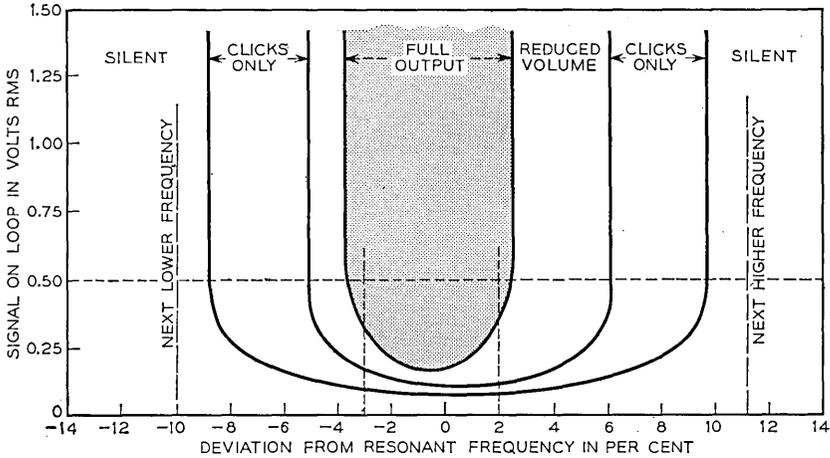
Fig. 6 — Selectivity and sensitivity of the tone ringer.

in Fig. 6. The coordinates of this chart are the amplitude of the ringing voltage on the line at the telephone set and the deviation of its frequency from the resonant value. The curves define areas within which various output levels are obtained. Full sound output is delivered whenever the intersection of frequency and voltage falls within the central shaded area, which extends to beyond +2 per cent and −3 per cent from resonance at all voltages above 0.5 volt. This range is wide enough to tolerate the expected frequency deviations due to possible manufacturing variations and to changes in ambient temperature from −20°F to 130°F. In a band just outside the full output area the tone ringer operates with a reduced volume. Surrounding this is a band in which no sustained tones are produced but a transient click is generated each time the ringing voltage is turned on. As the ringing frequency departs further from the resonant frequency of the ringer, the clicks become fainter and finally no sound at all is produced. As this silent condition prevails at the adjacent frequencies in the ringer series, complete selectivity is obtained. The dissymmetry which may be noted between the left and right portions of the "reduced volume" area is a result of feedback from the output of the Class C amplifier to the input of the tone ringer.

The voice-frequency input impedance of a tone ringer that is not operating is between 20,000 and 25,000 ohms. Its phase angle is a function of frequency, varying from 40° lagging (inductive) at 350 cycles per second, to 15° leading at 3,000 cycles per second. When the ringer is operating, its impedance is lowered by feedback associated with the

ripple voltage across the filter capacitor, the reduction being a compli-
cated function of frequency, length of loop and number of ringers.
Under extreme conditions the impedance may be as low as 9,000 ohms.
However, as this value is still much higher than the characteristic
impedance of the loop, the termination represents almost an open circuit,
even with several extensions sounding simultaneously. During ringing,
this mismatch aids in providing ample signal voltage at the set. During
talking, the still higher standby impedance avoids any appreciable
effect on speech transmission.

The direct current standby drain of this ringer consists of the 100
microamperes drawn by the input limiter and the $I_{co}$ of the two tran-
sistors. For the units presently used, typical values of $I_{co}$ are 5 micro-
amperes at 75°F and 40 microamperes at 130°F. All sets on an eight-
party line are unlikely to be at 130°F simultaneously, but on a hot day
their combined standby drain may be slightly over 1.0 milliampere.
This drain is trivial if the ringers are associated with existing systems hav-
ing conventional talking and supervisory arrangements, but somewhat
objectionable in connection with projected electronic switching systems
in which the talking and supervisory currents are materially reduced.
The continuing development of transistors which have values of $I_{co}$
several orders of magnitude smaller promises to permit use of simpler
circuitry, in which the limiter draws no direct current and the total
standby drain of the ringer is negligible.

III. PUSHBUTTON CALLER

3.1 *Basic Considerations*

To generate pushbutton calling pulses, magnetic energy derived from
the dc loop current is stored in the coils of the "party" and "digit" tuned
circuits while no pushbutton is pressed, and released in the form of
damped oscillatory transients upon operation of a button. It is found that
ample signal energy can be obtained in this way without the use of ampli-
fication, either to prolong the oscillations or to increase the useful output.
With its energy thus derived, the design must depend to some extent
upon the associated talking arrangements, particularly with regard to
whether they are "high-current" as in existing systems, or "low-current"
as they may possibly be for electronic switching. In laboratory studies
of the two cases, however, it has been found that the circuitry to be
described is applicable to both with only quantitative variations (as-
suming that the low-current talking arrangements draw on the order
of 10 milliamperes). Little if any generality is lost, therefore, in con-

fining our attention to a design for the pushbutton set and associated central office receiver which is based on conventional speech networks.

### 3.2 *Station Set Circuit*

In the station circuit shown in Fig. 7 the pushbutton caller is associated with a 500-type telephone set, the talking portions of which are only slightly modified. Most of the calling components are connected in series with the speech network on one side of the line.

As indicated in Fig. 7, which shows the state of the circuit when no pushbutton is operated, substantially all of the loop current flows through the low resistance of the two tuning inductors, storing energy in their magnetic fields. One tuned circuit, which generates the digit signals, employs the tapped inductor associated with capacitors C1, C2 and C3. The other, which generates the party-identification signal, employs a duplicate tapped inductor and capacitors C4 and C5. If one imagines points A and B to be strapped together, two parallel resonant circuits become apparent. Further, if the closed contact of K1 were suddenly opened, the energy stored in the fields of the two inductors would be dissipated in two independent oscillatory discharges, both of which would pass through the A-B strap. Actually, the assumed strap is replaced by a low-impedance winding added to the induction coil, T1, of the conventional speech network, as a means of coupling signal energy from both tuned circuits to the line. Since the telephone receiver
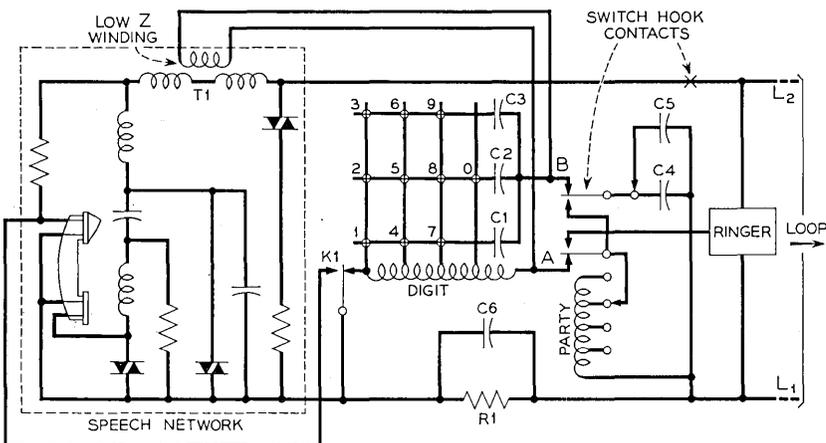


Fig. 7. — Circuit schematic of a pushbutton telephone set. Switchhook contact arrangements are indicated for sharing use of the party tuning elements with a tone ringer.

is also coupled to the induction coil, the pulse may be heard at a moderate level, being attenuated to some extent by sidetone balance. The rate of decay of the oscillations as well as the coupling between the tuned circuits is affected by the impedance of the added winding. Since a minimum of damping and inter-coupling is desirable, this impedance is made as low as is consistent with a satisfactory output amplitude. When switch K1 interrupts the current in the inductors, the dc loop current would drop to zero except for the presence of the resistor R1. This resistor maintains the current at a value sufficient to hold the central office supervisory relay operated and thus prevents loss of the call. Capacitor C6 shunts R1 to bypass voice frequencies. The abrupt drop in line current produced by switch K1 is used in "talk-off" protection as described in the following section.

The damped oscillations are frequency coded to convey the party and digit information. As noted earlier, the party frequency is preset at the time of installation by use of combinations of C4 and C5 and the four taps on the inductor. The ten digit frequencies are associated respectively with the ten pushbuttons by means of a crosspoint switching device.

This switch, as shown diagrammatically in Fig. 7, is associated with the digit inductor and capacitors C1, C2, and C3. Four parallel wires are connected to the inductor taps and three others perpendicular to these are connected to the capacitors. Thus 12 crosspoints exist in the switching mesh any one of which can associate a coil tap with a capacitor and establish a resonant frequency. Since only ten frequencies are
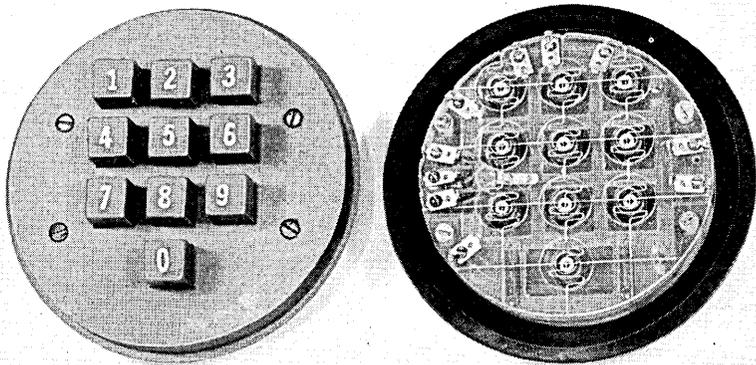


Fig. 8 — Front and rear views of a pushbutton crosspoint switch.

required for the digits, two of the crosspoints are not implemented. The ones used are shown circled and labeled with digit numbers in Fig. 7.

The crosspoint switch is part of the pushbutton assembly, a photograph of which appears in Fig. 8. The back view shows the crosspoint wires supported in clear plastic details. One set of wires is straight and the other provided with "S"-like bends at each crosspoint to give it the proper elasticity. Each pushbutton is aligned axially with a crosspoint, and when depressed closes the corresponding contact through a spiral spring which permits overtravel of the pushbutton. This overtravel allows crosspoint closure to occur early in the travel, thereby establishing the tuned circuit before interruption of the direct current by the switch K1. The actuation of K1 is accomplished by means of a slide-plate free to move transversely to the direction of motion of the pushbutton and driven during the overtravel period by a wedge-shaped portion of the pushbutton body. In order to prevent arcing, the break contact of this switch is arranged to open rapidly, using a toggle action, regardless of the speed of the customer's finger. Fig. 9 illustrates the slide-plate mechanism and the switch action.

An additional function of switch K1 is to short-circuit the transmitter in the telephone set as shown in Fig. 7. This provides protection against "talk-off", i.e., against false operation due to party or digit frequencies
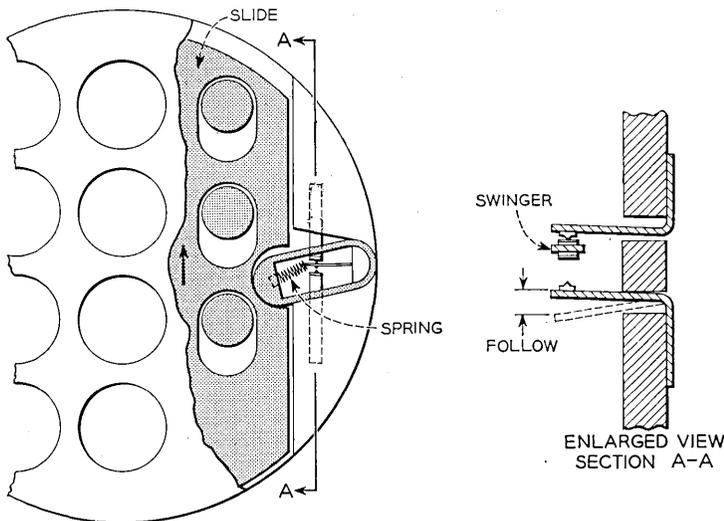


Fig. 9 — Switch detail, showing the slide plate and follow arrangement for toggle contact K1.

that might be present in the transmitter output. Since K1 is actuated only while a pushbutton is depressed, however, it affords no protection during idle periods before and between operations. Protection required during these idle periods is provided by letting the central office receiving equipment be continuously disabled except for a brief enablement triggered by the aforesaid sudden reduction in dc loop current. This method, combined with the transmitter shorting function of K1, affords substantially absolute protection.

Fig. 10 is a representation of the wave forms appearing at the line terminals of the station set as it is taken off hook and a button is pushed
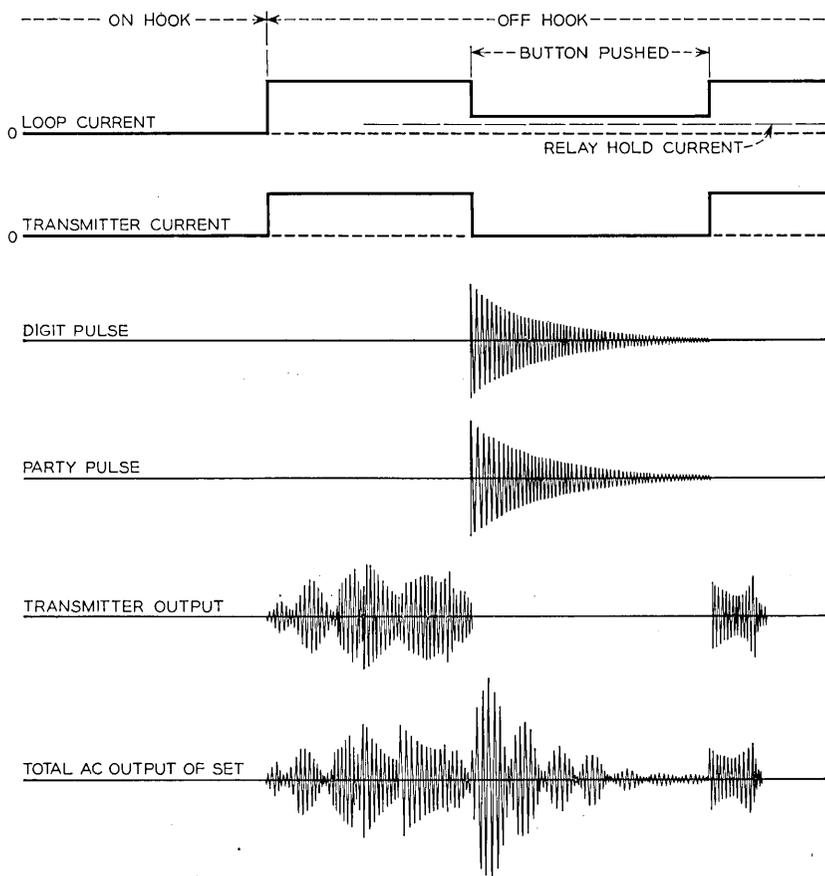


Fig. 10 — Waveforms illustrating operations at a station set in pushbutton calling.

and released. This figure summarizes the processes described in foregoing paragraphs.

### 3.3 *Central Office Receiver*

Concurrently with work on the station set, it was necessary to provide some suitable receiving equipment for the calling signals, to permit over-all tests in the laboratory. A schematic diagram of an experimental receiver is shown in Fig. 11. Voice frequency signals from the station set are transmitted through a line circuit to a "cutapart" filter, where the party identifying frequencies (478 cps to 1,000 cps) are separated from the digit frequencies (1,111 cps to 2,868 cps). Each group is then sent through one of two similar analyzing circuits, each consisting of an amplifier-limiter, selective circuits, a translator, and trigger circuits. A common enabling circuit is provided for talk-off protection. Both analyzing circuits terminate in a register where the information is available to subsequent elements of the switching system. In the laboratory model the received information is merely displayed on indicating lamps.

The line circuit is conventional except for the pair of diodes connected across one of the repeat coil windings. These serve to absorb large voltage surges such as those due to switchhook operation, which if not suppressed, might excite troublesome resonant frequencies in the cutapart filter.

In the enabling circuit, the drop in line current produced whenever a pushbutton is depressed is used to control a relay closure which briefly provides for recognition of the ac signals. When the line current decreases, it produces a rise in voltage across the blocking capacitor of the line circuit. This rise is differentiated and the resulting pulse of limited duration is applied to the base of a transistor, which is thereby energized to operate the relay for a period of about 40 milliseconds. This operation is delayed a few milliseconds by the capacitor shunted across the relay winding in order that any prior excitement of the tuned circuits may have time to decay before enablement. Enablement is thus made to occur only at a time when the station set transmitter is short-circuited, a proper signal is being sent, and the selective circuits are free of extraneous energy. It follows that all unwanted signals, except possible rare accidental bursts, are precluded and talk-off protection is made substantially absolute. One exceptional condition is the sending of an enabling pulse (dc step) by an accidental switchhook operation having too short a duration to lose the call. Such enablement is not troublesome as the
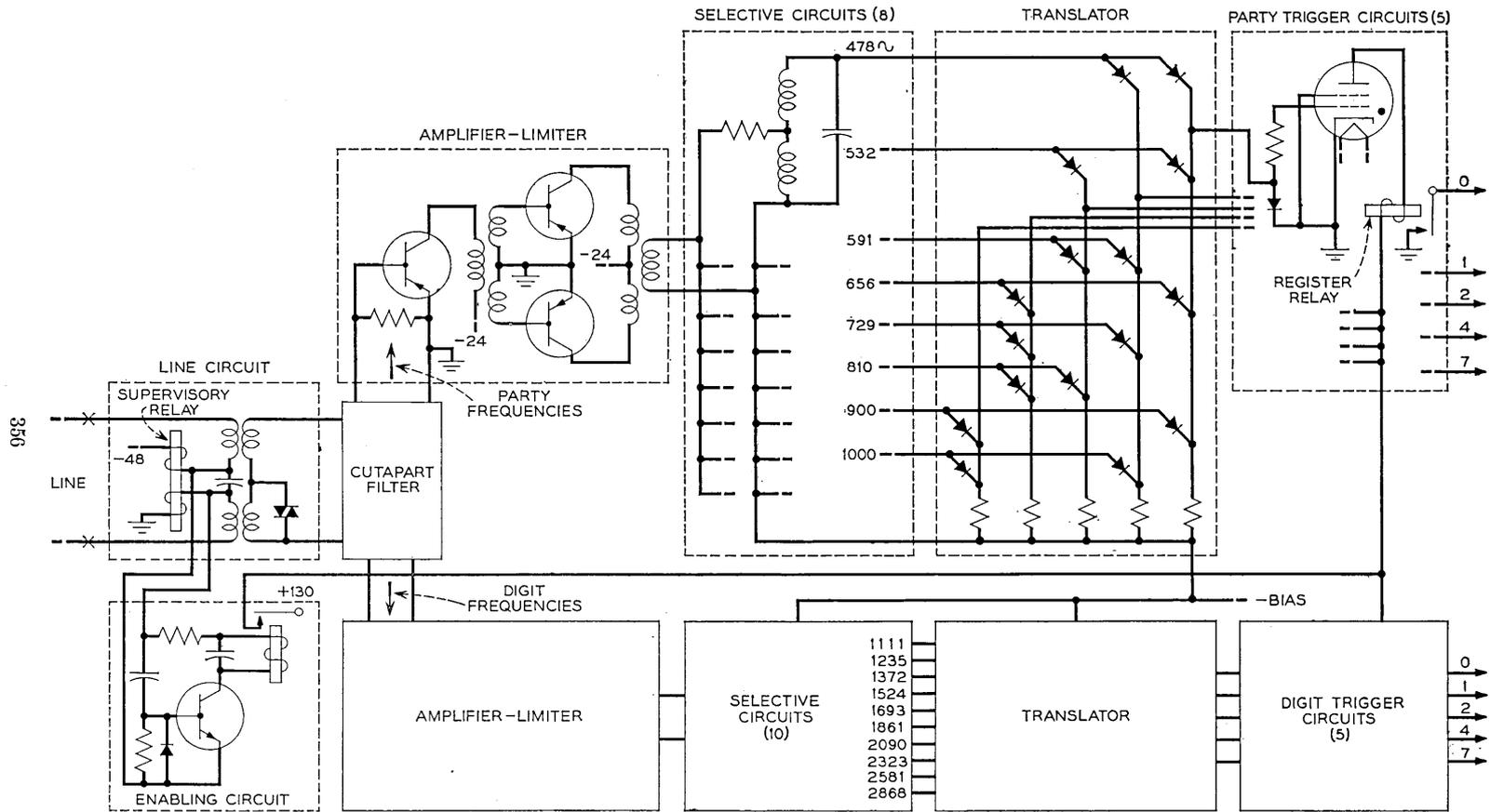
Fig. 11 — Schematic diagram of an experimental pushbutton calling pulse receiver.

set, disconnected from the loop by the same action, cannot send an interfering signal.

Each of the two channels emerging from the cutapart filter is applied to a symmetrical amplifier-limiter which provides, for a certain time, a substantially square wave of constant amplitude having the fundamental frequency of the received signal. This square wave affords a constant-amplitude drive for the tuned circuits, a technique employed also in the limiting amplifier of the tone ringer. An important feature of the limiter is that, whenever the wanted signal is at least slightly larger than any existing interference, the well-known effect of limiter capture tends to discriminate in favor of the wanted signal and thus aids in selectivity.
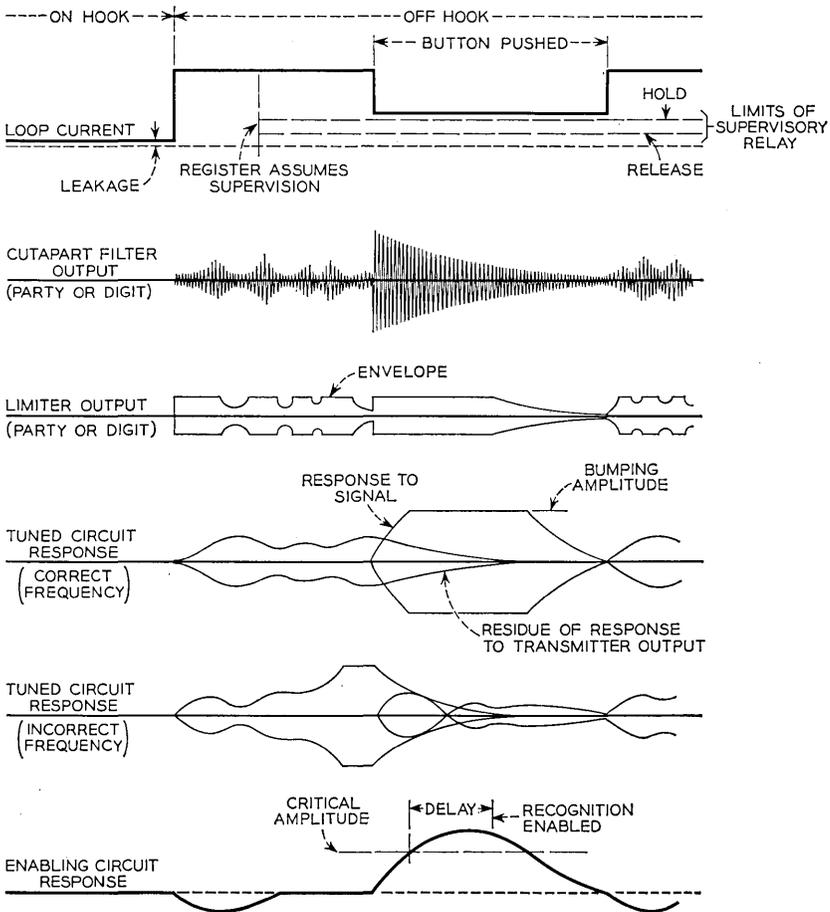


Fig. 12 — Waveforms illustrating operations at a pulse receiver in pushbutton calling.

The group of eight (or ten) tuned circuits may employ simple adjustable coils similar to those used in the station set. The outputs of these circuits are connected to a translator comprising a diode matrix which transmits the output of any one tuned circuit to two of five trigger circuits. In this manner the one-out-of-eight party (or one-out-of-ten digit) indication is translated into a two-out-of-five code with a resultant reduction in the required number of triggering devices. Other codes could be used but this one is commonly employed in existing dial pulse registers.

Each trigger circuit consists simply of a suitably biased thyratron, with its anode voltage supplied through a contact of the relay in the enabling circuit. When fired, the thyratron operates a register relay. It is extinguished when voltage is removed by the enabling circuit.

An essential feature of this trigger device, or of possible alternative circuitry, is that it must provide a fairly sharp drop in input impedance at or slightly above the triggering threshold, even when actual triggering is not enabled. This change in impedance (afforded by the grid-cathode diode of the thyratron) "bumps" the tuned circuit, restricting its amplitude to little more than that required for triggering. As a result, the amount of extraneous energy that can be stored in the tuned circuit
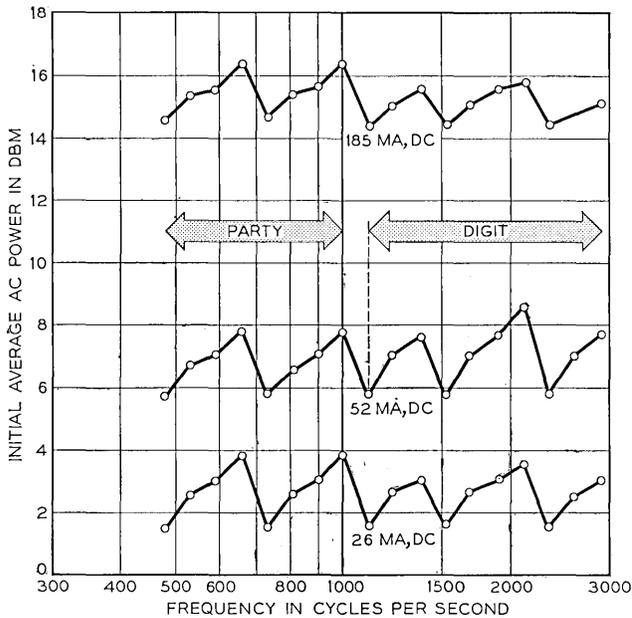
Fig. 13 — Power delivered to the line at the beginning of the damped oscillatory party or digit pulse is shown as a function of frequency and loop current.

is regulated, and the period of its decay to a negligible value, after transmitter disablement, is shortened.

Fig. 12 shows the waveforms at the central office corresponding to those at the station set as shown in Fig. 10. This figure also summarizes the action of the receiver in registering a signal. The loop current, limiter, and enabling diagrams are self-explanatory. The diagram for the cutapart filter shows random output in the first portion of the off-hook period resulting from speech or noise components in its pass band. When a button is pushed the transmitter output ceases and is replaced by the damped digit or party pulse. Release of the button restores the random output. The tuned circuit diagrams show representative types of response to the "correct" or resonant frequency and to the "incorrect" or non-resonant frequency. It is assumed for illustration that the "incorrect" circuit happened to be excited to the "bumping" amplitude by a component of the transmitter output just before the transmitter was disabled and the dial signals began. An appreciable time is seen to exist between the fall of the "incorrect" signal from, and the rise of the "correct" signal to, the bumping amplitude. It may also be noted that enablement does not occur before the correct signal has grown to the bumping amplitude.

### 3.4 *Performance*

The system as described was assembled for study and demonstration of the voice frequency signaling scheme. It has served this purpose well
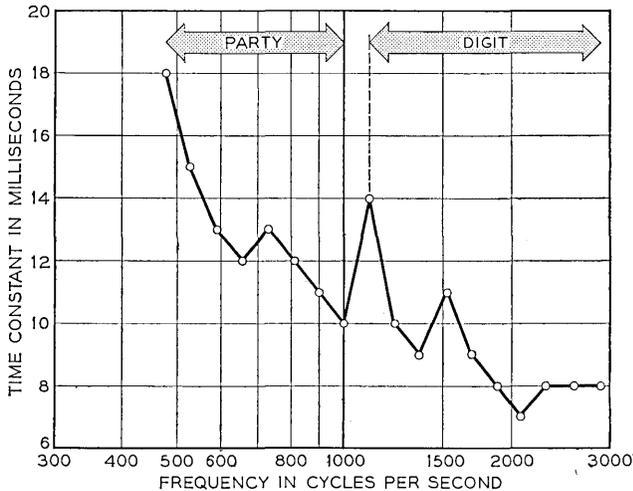


Fig. 14 — Time constant of the damped oscillatory party or digit pulse is shown as a function of frequency.

and in fact has operated dependably in the laboratory over a period of many months and over wide ranges of simulated operating conditions. However, the arrangement is only exploratory and numerous simplifications and improvements are known to be possible. Therefore, inclusion here of extensive performance data would be unwarranted. A few significant items are given below.

Fig. 13 shows the average ac power in dbm at the beginning of the oscillatory transient delivered at the station set terminals to a 900-ohm line for the 18 calling frequencies. Three curves are shown for dc loop currents of 185, 52, and 26 milliamperes representing loops of 0, 2, and 4 miles of 26-gauge cable, respectively. The saw-toothed appearance of Figs. 13 and 14 is related to the use of taps on the inductors of the signal generating circuits for tuning purposes. This results in progressively more rapid energy dissipation as capacitor connections are made to lower coil taps. Fig. 14 shows the time constants representing the decay rate of the voice frequency signal voltages across the 900-ohm line for a dc loop current of 26 milliamperes. For the least available power in a 4-mile loop and the shortest time constant, the duration of limiting at the receiver is about 16 milliseconds which is adequate for proper signal registry.

IV. ACKNOWLEDGEMENTS

The exploration described has depended upon the cooperation of many people. It represents contributions of information, judgment, and effort from members of many departments in the Laboratories, and of other companies in the Bell System. The writers wish to acknowledge indebtedness to all of them for this cooperation, and particularly for encouragement and constructive interest.

REFERENCES

1. A. E. Joel, Electronics in Telephone Switching Systems, B.S.T.J., **35**, pp. 991–1018, Sept., 1956.
2. G. V. King, Centralized Automatic Message Accounting System, B.S.T.J., **33**, pp. 1331–1342, Nov., 1954.
3. C. A. Dahlbom, A. W. Horton, Jr., and D. L. Moody, Application of Multifrequency Pulsing in Switching, A.I.E.E. Transactions, **68**, pp. 392–396, 1949.
4. A. Weaver and N. A. Newell, In-Band Single-Frequency Signaling, B.S.T.J., **33**, pp. 1309–1330, Nov., 1954.
5. E. E. Mott and R. C. Miner, The Ring Armature Telephone Receiver, B.S.T.J., **30**, pp. 110–140, Jan., 1951.

# Attenuation in Continuously Loaded Coaxial Cables

### By GORDON RAISBECK

*The formula for the attenuation of a coaxial line loaded continuously with magnetic materials involves, after a change of variables, only three parameters, even when the effect of dielectric losses is included. If the dimensions of the line have their optimum values, the attenuation is a function of only two parameters. The relation is exhibited both graphically and analytically, in forms which can be applied to practical problems. A simple numerical example is given to illustrate the use of the formula.*

## INTRODUCTION

The development of ferrite materials having high permeability and low loss at high frequencies has raised again the question of continuous loading of coaxial cables. The question is one of many treated in a recent thorough paper on loading by P. M. Prache.[1, 2] He considers, among other things, cables loaded with coaxial cylinders of magnetic material, and attacks the problem of minimizing the attenuation through the regulation of various free parameters.

The present paper is another attempt to find the conditions under which attenuation in a loaded coaxial cable is a minimum.[3] It differs from the previous attempts in several respects. First, a drastic set of changes of variable makes the problem much easier to manage mathematically. Second, the resulting simplification makes it possible to include the effect of dielectric loss in the loading material. Third, a detailed analysis of magnetic losses, such as found in Prache's paper, is ignored and losses are described in terms of $Q$, the ratio of peak energy stored to energy lost per radian. Special assumptions about the joint restrictions on $Q$ and the permeability (or dielectric constant, as the case may be) are reserved for later steps in the analysis.

This paper is entirely theoretical. The results indicate that substantial reductions in attenuation are possible at frequencies of a few megacycles with magnetic materials having a $pQ$ product ($p$ = relative per-

meability) of $10^4$. Whether this is worth while practically is another matter. For example, ferrites having a $pQ$ product greater than $10^4$ and $p$ less than 100 have been made,[4] but these are brittle and rigid. Stringing beads of ferrite or assembling split cylinders about a center wire are quite conceivable, but the mechanical difficulties are enough to dampen the enthusiasm of anyone designing for manufacture. Nevertheless, the analysis in this paper provides a sound basis for comparing the electrical characteristics of various loading methods as they became available, and provides quantitative information about dimensions and electrical properties required if new loading materials are sought.
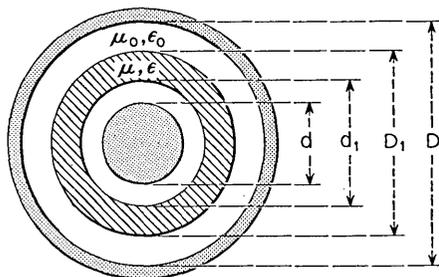


Fig. 1 — Schematic cross section of a coaxial transmission line loaded with magnetic material. The diameters of the inner and outer conductors are $d$ and $D$, respectively. The inner and outer diameters of the cylinder of loading material are $d_1$ and $D_1$, respectively.

## CABLE STRUCTURE AND EQUIVALENT CIRCUIT

A cross section of the cable is shown in Fig. 1. The diameter of the inner conductor is $d$, and the (inside) diameter of the outer conductor is $D$. The inside and outside diameters of the loading cylinder are $d_1$ and $D_1$. The main dielectric has dielectric constant $\epsilon_0$ and permeability $\mu_0$, and the loading material has dielectric constant $\epsilon$ in the radial direction and permeability $\mu$ in the tangential direction. (Other components of $\mu$ and $\epsilon$ are not important to the problem; the material may as well be assumed isotropic. However, if the shell has air-gaps, the effective dielectric constant and permeability are not in fact isotropic.) The permeability of the loading material relative to that of the main dielectric is defined as*

$$p = \frac{\mu}{\mu_0},$$

and the reciprocal of the dielectric constant of the loading material rela-

---

* The notation used in this paper is due mostly to Prache, Reference 2.

tive to that of the main dielectric is defined as

$$q = \frac{\epsilon_0}{\epsilon}.$$

Notice that $\epsilon_0$ and $\mu_0$ need not be the dielectric constant and permeability of free space. It will turn out that the various diameters $d$, $d_1$, $D_1$ and $D$ are not individually important, but their ratios are. Accordingly we adopt the notation

$$N = \frac{D}{d},$$

$$n = \frac{D_1}{d_1}.$$

A convenient equivalent circuit of a unit length of line is represented
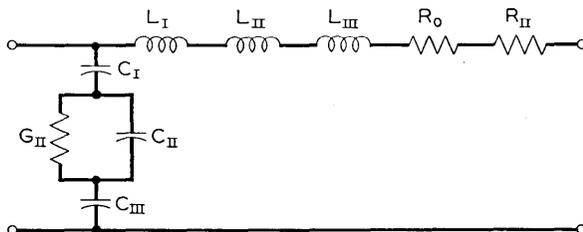


Fig. 2 — Equivalent circuit of a unit length of magnetically loaded line.

in Fig. 2. The series elements are $L_I$, the inductance due to magnetic fields between $d$ and $d_1$ ; $L_{II}$, the inductance due to fields between $d_1$ and $D_1$ ; $R_{II}$, the loss associated with $L_{II}$ ; $L_{III}$, the inductance due to fields between $D_1$ and $D$; and $R_0$, the series conductor resistance. The shunt elements are $C_I$, the capacitance due to electric fields between $d$ and $d_1$ ; $C_{II}$, the capacitance due to electric fields between $d_1$ and $D_1$ ; $G_{II}$, the loss conductance associated with $C_{II}$ ; and $C_{III}$, the capacitance due to fields between $D_1$ and $D$. The main dielectric is assumed to be free of loss, and the frequency is assumed high enough so that all current flow is at the surface of the conductors. The resulting expressions are:

$$L_I = \frac{\mu_0}{2\pi} \log \frac{d_1}{d},$$

$$L_{II} = \frac{\mu}{2\pi} \log \frac{D_1}{d_1} = \frac{p\mu_0}{2\pi} \log \frac{D_1}{d_1},$$

$$L_{III} = \frac{\mu_0}{2\pi} \log \frac{D}{D_1},$$

$$\frac{1}{C_\text{I}} = \frac{1}{2\pi\epsilon_0} \log \frac{d_1}{d},$$

$$\frac{1}{C_\text{II}} = \frac{1}{2\pi\epsilon} \log \frac{D_1}{d_1} = \frac{q}{2\pi\epsilon_0} \log \frac{D_1}{d_1},$$

and

$$\frac{1}{C_\text{III}} = \frac{1}{2\pi\epsilon_0} \log \frac{D}{D_1}.$$

The losses are evaluated through the relations

$$R_\text{II} = \frac{2\pi f L_\text{II}}{Q'},$$

and

$$G_\text{II} = \frac{2\pi f C_\text{II}}{Q''}.$$

It is convenient to define the "$Q$" of the unloaded line as

$$Q = \frac{\mu_0 f \, \log \, (D/d)}{R_0}$$

so that

$$R_0 = \frac{2\pi f L_0}{Q}.$$

The inductance $L_0$ is simply the inductance per unit length which the line would have if it contained no loading material.

The various components of series impedance and shunt admittance are easily added to make an equivalent circuit with one series inductance $L$, one series resistance $R$, one shunt capacitance $C$, and one shunt conductance $G$, as in Fig. 3. Here

$$L = L_\text{I} + L_\text{II} + L_\text{III},$$

$$= \frac{\mu_0}{2\pi} \left[\log \, (N/n) + p \log n\right],$$

and

$$R = R_0 + R_\text{II},$$

$$= \frac{2\pi f}{Q} \frac{\mu_0}{2\pi} \log N + \frac{2\pi f}{Q'} \frac{p\mu_0}{2\pi} \log n,$$

$$= f\mu_0 \left[\frac{1}{Q} \log N + \frac{p}{Q'} \log n\right].$$
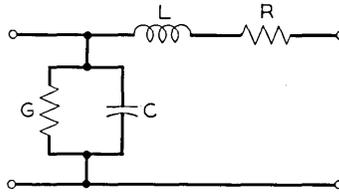
Fig. 3 — Simplified equivalent circuit of a unit length of magnetically loaded line.

If $Q''$ is assumed to be large compared to unity, then

$$\frac{1}{C} = \frac{1}{C_{\mathrm{I}}} + \frac{1}{C_{\mathrm{II}}} + \frac{1}{C_{\mathrm{III}}},$$

$$= \frac{1}{2\pi\epsilon_0} [\log(N/n) + q \log n],$$

and

$$G = \frac{2\pi f}{Q''} \frac{\dfrac{1}{C_{\mathrm{II}}}}{\left[\dfrac{1}{C_{\mathrm{I}}} + \dfrac{1}{C_{\mathrm{II}}} + \dfrac{1}{C_{\mathrm{III}}}\right]^2},$$

$$= \frac{2\pi f}{Q''} \frac{2\pi\epsilon_0 q \log n}{[\log (N/n) + q \log n]^2}.$$

TRANSMISSION CHARACTERISTICS

The propagation constant and characteristic impedance of the line are easily described in terms of $R$, $L$, $C$ and $G$. In fact, to a first approximation

$$Z_0 = \text{characteristic impedance}$$

$$= \sqrt{\frac{L}{C}}$$

$$\beta = \text{phase constant}$$

$$= \omega\sqrt{LC}$$

$$\alpha = \text{attenuation per unit length}$$

$$= \frac{R}{2}\sqrt{\frac{C}{L}} + \frac{G}{2}\sqrt{\frac{L}{C}}.$$

It is easy to show that

$$Z_0 = \frac{1}{2\pi} \sqrt{\frac{\mu_0}{\epsilon_0}} \sqrt{[\log (N/n) + p \log n][\log (N/n) + q \log n]},$$

and

$$\beta = \omega\sqrt{\mu_0\epsilon_0}\ \sqrt{\frac{\log\ (N/n)\ +\ p\ \log\ n}{\log\ (N/n)\ +\ q\ \log\ n}}.$$

We shall devote more attention to the attenuation-per-unit-length $\alpha$. Substituting values for $R$, $L$, $C$ and $G$, one finds

$$\alpha = \frac{f\mu_0}{2}\left[\frac{1}{Q}\ \log\ N + \frac{p}{Q'}\ \log\ n\right]$$

$$\frac{2\pi\sqrt{\epsilon_0/\mu_0}}{\sqrt{[\log(N/n)\ +\ p\ \log\ n][\log(N/n)\ +\ q\ \log\ n]}}$$

$$+ \frac{2\pi f}{2Q''}\ \frac{2\pi\epsilon_0\ q\ \log\ n}{[\log\ (N/n)\ +\ q\ \log\ n]^2}\ \frac{1}{2\pi}\ \sqrt{\frac{\mu_0}{\epsilon_0}}$$

$$\sqrt{[\log\ (N/n)\ +\ p\ \log\ n][\log\ (N/n)\ +\ q\ \log\ n]}.$$

The factor $\pi f\sqrt{\mu_0\epsilon_0}$ is common to both terms. In fact, a little reflection about the unloaded line (set $Q' = Q'' = \infty$, $p = q = 1$) shows that for the unloaded line

$$\alpha_0 = \frac{\pi f\ \sqrt{\mu_0\epsilon_0}}{Q}.$$

Now make a change of variable

$$\frac{\log\ N/\log\ n + q - 1}{p - q} = v.$$

Then

$$\frac{\alpha}{\alpha_0} = kv^{-1/2}(v\ +\ 1)^{-1/2} + v^{1/2}(v\ +\ 1)^{-1/2} + mv^{-3/2}(v\ +\ 1)^{1/2},$$

where

$$k = \frac{1\ -\ q}{p\ -\ q} + (Q/Q')\ \frac{p}{p\ -\ q},$$

and

$$m = (Q/Q'')\ \frac{q}{p\ -\ q}.$$

Note that the effect of magnetic loss is included in $k$, and the effect of dielectric loss in $m$.

OPTIMUM DIMENSIONS FOR MINIMUM ATTENUATION

In the ordinary case where the conductor resistance $R_0$ is simply ohmic skin resistance, $Q$ is a well-known function of $D$ and $N$:

$$Q = \frac{D \log N}{1 + N} \sqrt{\pi\sigma\mu_0 f},$$

where $\sigma$ is the conductivity of the conducting material. For fixed $D$, the maximum value of $Q$, and hence the minimum value of $\alpha_0$ occurs when

$$1 + N = N \log N,$$

or

$$N = 3.59.$$

If $v$ is kept constant (e.g., by keeping $\log N/\log n$ constant) this also yields the lowest value of $\alpha$. Hence, unless some mechanical or dimensional restraint limits $n$, the value $N = 3.59$ is the optimum for attenuation reduction whether loading is used or not.

If we now regard $Q$, $Q'$, $Q''$, $p$, $q$, $d$, $d_1$ and $D_1$ as independent parameters, we find that all of the new variables are independent of the dimensions save $v$. Hence it is possible to find the optimum dimensions of the line by differentiating with respect to $v$, and setting the derivative equal to zero. After solving for $v$, the result is

$$v = \frac{k + 5m + \sqrt{k^2 - 14mk + m^2 + 12m}}{2(1 - 2k - 2m)}.$$

It is necessary to reject the root with the negative radical, because $v$ is always a positive number. Unfortunately, the expression for $\alpha$ becomes rather complicated if this value is substituted for $v$. The result is

$$\frac{\alpha^2}{\alpha_0^2} = \frac{16}{9}(1 - k)(k - m) + \frac{8}{9}(1 - k)(k - m + \sqrt{\phantom{x}})$$

$$+ \frac{8(1 - k)(k - m)^2}{9(k - m + \sqrt{\phantom{x}})},$$

$$= \frac{8(1 - k)(2k - 2m + \sqrt{\phantom{x}})^2}{9(k - m + \sqrt{\phantom{x}})},$$

$$= \frac{8}{3}(1 - k)(k - m) - \frac{2(k - m)^3}{27m} + \left[\frac{8}{9}(1 - k) + \frac{2(k - m)^2}{27m}\right]\sqrt{\phantom{x}},$$

where

$$\sqrt{\phantom{x}} = \sqrt{k^2 - 14mk + m^2 + 12m}.$$

Of these, the first seems to be suited to general computation, and the second to slide rule computation. The third form is indeterminate when $m = 0$, which is just the region which it is desirable to investigate. However, if special relations between $k$ and $m$ are assumed, then some easy results are available. For example, suppose dielectric loss can be neglected, i.e., assume $m = 0$. Then

$$\frac{\alpha}{\alpha_0} = kv^{-1/2}(v + 1)^{-1/2} + v^{1/2}(v + 1)^{-1/2}.$$

A family of these curves for various values of $k$ is plotted in Fig. 4. The minimum occurs when

$$v = \frac{k}{1 - 2k}$$

and has the value

$$\frac{\alpha}{\alpha_0} = 2\sqrt{k(1 - k)}.$$

This curve is the lowest curve in Fig. 5.

To include the effect of dielectric loss, one can assume a fixed value of
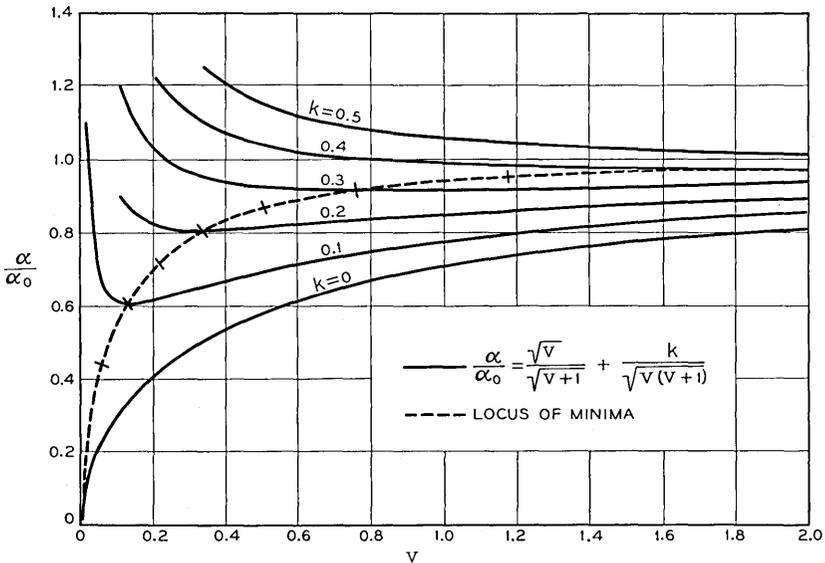


Fig. 4 — Reduction of attenuation in a magnetically loaded line in terms of the normalized dimension parameter $v$ and the normalized magnetic loss parameter $k$. For each fixed $k$, the value of $v$ yielding minimum attenuation is shown.
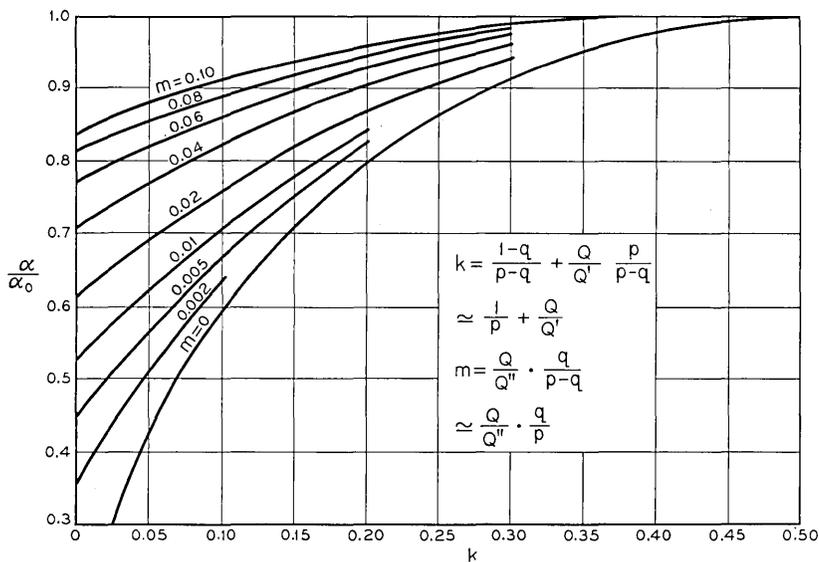
Fig. 5 — Maximum reduction of attenuation in a magnetically loaded line in terms of the magnetic loss parameter $k$ and the dielectric loss parameter $m$ of the magnetic material.

$m$ and proceed as before. The result is a family of curves of $\alpha/\alpha_0$ versus $k$ for fixed values of $m$. Such a family is plotted on Fig. 5 for various values of $m$. These data may be replotted as in Fig. 6 to show $m$ versus $k$ for constant $\alpha/\alpha_0$. This family of curves is useful to show the extent to which one kind of loss can be traded for another. As $\alpha/\alpha_0$ decreases, the slope of the curves decreases. This means that the role of $m$, and hence of dielectric loss, is becoming more and more important compared to that of $k$, which includes magnetic loss.

REDUCTION IN ATTENUATION WITH OPTIMUM DESIGN

If all dimensions have their optimum values, then the reduction in attenuation can be described in terms of two variables only, $k$ and $m$. If, furthermore, dielectric loss is known to be negligible, for each value of $\alpha$ there is a unique value of $k$, and

$$k = \frac{1 - q}{p - q} + \frac{Q}{Q'} \frac{p}{p - q} = f\left(\frac{\alpha}{\alpha_0}\right) = \frac{1 - \sqrt{1 - (\alpha/\alpha_0)^2}}{2}.$$

If we assume that $p$ is large compared to $q$, this reduces to

$$\frac{1}{p} + \frac{Q}{Q'} = \frac{1 - \sqrt{1 - (\alpha/\alpha_0)^2}}{2}.$$
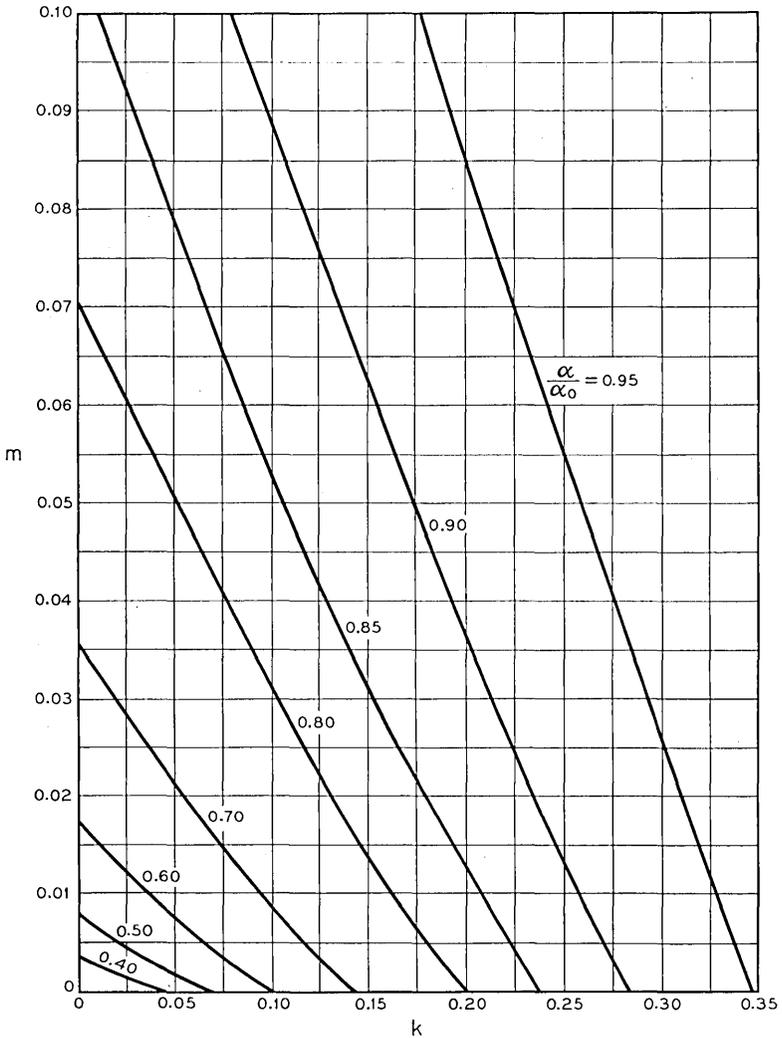
Fig. 6 — The relation between dielectric loss parameter $m$ and magnetic loss parameter $k$ for which the maximum possible reduction of attenuation is the same, plotted for several different values of attenuation reduction.

If we assume further the old rule of thumb that $pQ'$ equals a constant for different types of cores made from the same type of material, then the lowest value of $\alpha$ is found when

$$p = \sqrt{\frac{pQ'}{Q}}.$$

Then

$$k = 2 \sqrt{\frac{Q}{pQ'}},$$

and

$$\frac{\log N}{\log n} = 1 + \frac{2}{1 - 2 \sqrt{Q/pQ'}} \cong 3$$

if $pQ'$ is large compared to $Q$. These results, and their counterparts involving $q$ explicitly, were found by Prache.[2] They have been criticized on the ground that $pQ'$ is not a constant for all types of materials. It is clear, however, that the minimum value of $\alpha$ arrived at from the above formula is nearly correct, even if $p$ deviates rather widely from $\sqrt{pQ'/Q}$, whereas the value of $\log N/\log n$ required to realize the minimum attenuation will be altered. In fact, if $p > \sqrt{pQ'/Q}$, $\log N/\log n > 3$, i.e., the shell of loading material is thinner than before. It is unlikely in practical cases that we should want to use $p < \sqrt{pQ'/Q}$, but if we did, a thicker shell of loading material would be required.

The analogous approximation when dielectric loss is not neglected is:

$$k = \frac{1}{p} + \frac{Q}{Q'},$$

$$m = \frac{Q}{Q''} \frac{q}{p}.$$

It is clear from the graphs that if the loading material has dielectric losses, then one should use a thinner shell of material having higher $p$. The relative variation of $p$, $Q'$ and $Q''$, the relative permeability and the magnetic and dielectric "$Q$" respectively, in one general type of material has not been described in detail, and probably no concise approximation like ($pQ'$ = constant) exists which involves all three. It must be borne in mind furthermore that $k$ and $m$ involve $Q$ and hence depend on the dimensions of the line. For a given line size one could compute $k$ and $m$ for known magnetic materials and deduce from Fig. 5 or Fig. 6 whether it is worth while to load the line with any particular material.

All of the reasoning on previous pages applies, with appropriate modification, to the continuous loading of any transmission link whose cross-section is a conformal transformation of a circular annulus. This broad class includes any uniform transmission line having just two conductors. Some examples are given in the author's patent.[3]

NUMERICAL EXAMPLE

Suppose for the sake of a numerical example that dielectric losses are negligible, that $pQ' = 10^4$ for the loading material to be used, that the

conductors are made of copper, and that the dielectric constant of the loading material is high. At frequencies such that the skin depth of currents is small compared to conductor thickness,

$$Q = (D/3.59)\sqrt{\pi\sigma\mu_0 f},$$

where $\sigma$ is conductivity and all quantities are in $MKS$ units. For copper

$$\sigma = 58.10^6 \text{ mho/meter},$$

and for most dielectrics

$$\mu_0 = 12.57 \times 10^{-7} \text{ henry/meter.}$$

Then

$$Q = 4.23 \, D\sqrt{f}$$

when $D$ is in meters and $f$ in cycles per second. If $D$ is measured in inches and $f$ in megacycles per second, then

$$Q = 107 D_{\text{in}}\sqrt{f_{\text{mcps}}}.$$

The optimum value of $p$ is

$$p = \frac{10^4}{\sqrt{107 D_{\text{in}} f_{\text{mcps}}^{1/2}}} = 5.44 D_{\text{in}}^{-1/2} f_{\text{mcps}}^{-1/4},$$

and $k$ has the value

$$k = \frac{2}{p} = 0.368 D_{\text{in}}^{1/2} f_{\text{mcps}}^{1/4}.$$

Then

$$\frac{\alpha}{\alpha_0} = 2 \sqrt{0.368 D_{\text{in}}^{1/2} f_{\text{mcps}}^{1/4} (1 - 0.368 D_{\text{in}}^{1/2} f_{\text{mcps}}^{1/4})}.$$

The resulting curves are plotted on Fig. 7 for $D = 1''$, $\frac{1}{2}''$, $\frac{1}{4}''$, and $\frac{1}{10}''$, for values of $f$ between 100 kilocycles per second and 10 megacycles per second. It is clear from these curves that the improvement attainable in a one-inch line is modest over this frequency range, but that worthwhile improvements can be attained in a $\frac{1}{10}''$ cable even at many megacycles.

The effect of dielectric loss might be included thus: Take the 0.25'' line at 1 megacycle per second. Imagine that $Q''$ is constant, say, and that, e.g.

$$Q'' = 200,$$

$$q = 0.10.$$
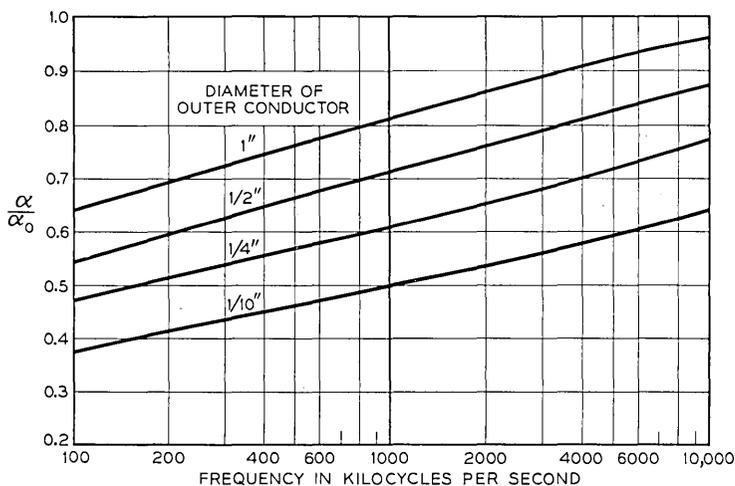
Fig. 7 — Maximum possible reduction in attenuation by magnetic loading with material having a $pQ'$ product of $10^4$ for copper coaxial lines of various diameters, plotted as a function of frequency.

Then

$$\frac{Q}{Q''} = \frac{85}{200} = 0.425,$$

$$m = \frac{Q}{Q'} \frac{q}{p}$$

$$= \frac{0.0425}{p}.$$

Neglecting the dielectric loss we find

$$p = 5.44\ D_{\text{in}}^{-1/2} f_{\text{mcps}}^{-1/4} = 10.9.$$

Hence

$$m \leqq 0.0040.$$

From Fig. 5 it is easy to see that $\alpha/\alpha_0$ is increased from 0.78 to about 0.81 by the dielectric loss. A little experimentation with different values of $p$ shows that this is very nearly the optimum. If, on the other hand, e.g.,

$$Q'' = 20,$$

then

$$m = \frac{0.425}{p},$$

and $\alpha/\alpha_0$ is increased from 0.78 to 0.90. This can be reduced to 0.88 by choosing $p \cong 20$.

REFERENCES

1. P. M. Prache, Noyaux et Coquilles dans le Domaine des Tele-Communications, Cables et Transmission, 6, 1 and 2, Jan. and April 1952.
2. U. S. Patent 2,727,945, High Frequency Magnetic Elements and Telecommunication Circuits, issued December 20, 1955 to M. P. Prache.
3. Certain aspects of this analysis are treated slightly differently in U. S. Patent 2,787,656, Magnetically Loaded Conductors, issued April 2, 1957, to the present author.
4. Private communication from F. J. Schnettler. There appears to be no reference to such materials in publications, where emphasis has been placed on high relative permeability $p$.

# New Developments in Military Switching

By A. C. GILMORE, P. R. GRAY and W. S. IRVINE

*Plans for a new communication network for a military theater of operation were recently formulated. These plans include not only a number of presently available new facilities such as carrier and radio systems, but also a general switching plan together with means for its implementation. It is with these switching arrangements that this article is concerned.*

*A manual switchboard of advanced design is described in which many formerly manual procedures are mechanized. An orderly arrangement of switching centers and trunk groups provides for efficient usage of trunk facilities with alternate routes engineered as needed to minimize call delays.*

*The equipment design features considerable flexibility so that rearrangements to decrease or increase the size of a switching center may be accomplished with a minimum of effort. Rapid installation or dismantling is facilitated by limiting individual equipment packages to 250 pounds and by the use of patch cables which eliminate the need for soldering when an office is installed.*

## I. INTRODUCTION

Some years ago, following examination of the effectiveness of communication techniques and facilities employed in prosecuting World War II, plans were formulated for a new communication network for a military theater of operation. These plans comprise not only a number of new facilities, such as new carrier and radio systems now available and described elsewhere,* but also a general switching plan for a military theater together with means for its implementation. It is with these switching arrangements that this article is concerned.

---

* Bell Laboratories Record: G. Rodwin and G. H. Huber, New Military Carrier Telephone System, pp. 274–275, July, 1955; G. H. Huber, A Military Communication Network Using Wire and Radio, pp. 290–293, Aug. 1955; A. L. Durkee, New Military VHF-UHF Radio Set, pp. 382–384, Oct., 1955; and W. G. Hensel and Reinschmidt, AN/TRC-24 Radio Transmitter, pp. 428–431, Nov., 1955.

The problems involved in military switching, where mobility and adaptability are prerequisite, differ radically from those encountered in a commercial system. In a commercial system both central offices and stations remain comparatively fixed, thus permitting the use of a pre-engineered numbering pattern. A study of these problems and of available art indicated that automatic switching for a military theater would be difficult to realize at this time. Accordingly, development work was initiated on a manual switching system of advanced design which would to a considerable extent mechanize former manual proce- dures and provide many of the benefits of the automatic switching. At the same time the flexibility necessary of the shifting demands of military situations will be provided. Salient features of this switching system include:

(a) Adaptability to an orderly arrangement of switching centers and trunk groups so as to provide efficient usage of facilities with alter- nate routes as needed to minimize calling delays.

(b) Complete control and supervision of calls by the originating operator with automatic differentiation between the call status when the called party has not answered and when he has hung up. Through these and other features, the new switching equipment reduces operator time to about half that previously required.

This newly developed straightforward military switchboard has been designated "Manual Telephone Central Office". The switchboard, together with associated equipment developed on this and previous contracts, provides the necessary facilities for an integrated switching network for military theater communications.

GENERAL

II. FEATURES OF MANUAL TELEPHONE CENTRAL OFFICE

The central office is a universal, common battery switchboard devel- oped for military use at both toll and local switching centers. It is arranged for use with a nominal 48-volt battery power plant. It is designed primarily for straightforward type trunking. However, special trunks are provided to connect to existing military switchboards. In addition, a combination line and trunk is provided to connect on a ring- down basis to magneto switchboards or lines. A special trunk is furnished which permits connection to a civilian local central office on a dial or manual basis.

The development model of this switchboard was subjected to military environmental and operational tests and was approved. While the development model was a three position, two hundred line board, the

flexibility of the unitized equipment design permits a progressive growth to a maximum of twenty positions and fourteen hundred lines and/or trunks when using a four panel multiple.

III. SWITCHING CONSIDERATIONS

Civilian communications networks such as those used with the nationwide dialing switching plan of the Bell System are engineered with the assurance that the various centers comprising the network will be fixed as to location. The connecting trunk routes and the numbering plan will also remain comparatively fixed for some years.

A military theater of operations switching plan is difficult to establish with any assurance that the various centers comprising the network will remain fixed for any length of time. Accordingly the concept of a dial switching system for a military theater of operations was temporarily deferred in favor of an interim plan employing a manual system of advanced design.

The design requirements envisioned a manual switching system in which the straightforward mode of operation would provide the most desirable features of a dial system but which would be tailored to meet the needs of a fluid military situation.

The basic Military Straightforward Switching Plan is patterned after the Bell System Plan for nationwide dialing. The two plans are illustrated in Fig. 1 for ready comparison.

The basic switching plan for a military theater contains a complex arrangement of interconnected switching centers serving a variety of functions but integrated in an orderly sequence. Some of these switching centers serve important large areas and have trunk groups to all parts of the theater. Others cover more restricted and somewhat less important areas and depend upon centers higher in the chain of authority to reach points outside their own sectors. In diminishing order of rank or importance, the trunk switching centers in the military long distance network are zone centers, primary centers and secondary centers. Next below these are the local or end centers which are at the bottom of the long distance chain.

Magneto or tributary exchanges at which station loops terminate are next in order of rank below the local or end centers. They connect one loop to another or connect a loop to a tributary trunk which leads to the switching center for the area, or to a terminal trunk leading to another tributary exchange in the same local area.

From a systems engineering standpoint it is necessary to limit the number of switching points or links in a connection involving toll

facilities. Reference to Fig. 1(b) shows that a call from the local center LC1 to the local center LC2 could involve six intermediate switching points and seven links. Such a call would be switched successively through secondary, primary and zone center and thence successively through a zone, primary and secondary center of the second area. This number of switches on a large number of calls would cause adverse reactions on service for two reasons. First, the number of switching points involving operators would increase the time required to establish the connection and, second, the over-all transmission loss which is the sum of the individual link losses would be too great for satisfactory transmission. Therefore, additional routes are provided for most calls of this type which bypass certain switching centers. The high-usage trunk group from secondary center SC1 to primary center PC2 in Fig.



(a)
BELL SYSTEM SWITCHING PLAN
FOR NATIONWIDE DIALING

REGIONAL CENTER

SECTIONAL CENTER

PRIMARY CENTER

TOLL CENTER

●END OFFICE

—— FINAL TRUNK GROUP
━ ━ ━ HIGH USAGE GROUPS ARE PROVIDED AS JUSTIFIED BY TRAFFIC AND COST RATIOS
TRANSMISSION OBJECTIVES FOR TRUNKS: VIA TRUNK NET LOSS 0.4 – 4.5 DB DEPENDING ON RANK OF OFFICES AT EACH END
ADDITIONAL LOSS FOR END OFFICE LINKS – 2 DB
MAXIMUM TOTAL LOSS FOR END OFFICE LINKS – 4 DB

(b)
MILITARY STRAIGHTFORWARD
SWITCHING PLAN

ZONE CENTER  Z ZC1        Z ZC2

PRIMARY CENTER  PC1        PC2

SECONDARY CENTER  SC1        SC2

LOCAL CENTER  LC1        LC2

●MAGNETO SWITCHBOARDS

—— FINAL TRUNK GROUP
━ ━ ━ HIGH USAGE GROUPS ARE PRO- VIDED AS JUSTIFIED BY TRAFFIC
TRANSMISSION OBJECTIVES FOR TRUNKS: VIA TRUNK NET LOSS 3 DB
NO ADDITIONAL LOSSES AT THE TERMINALS (NO PADS)

Fig. 1 — Comparison of Bell System Switching Plan for Nationwide Dialing and a new Military Straightforward Switching Plan designed to facilitate long-distance military traffic on an essentially "no delay" basis.

1(b) illustrates this type of trunk. Now a call from local center LC1 to local center LC2 would be routed successively through secondary center SC1, primary center PC2 and secondary center SC2. This routing over a high-usage trunk group reduces the number of switching points from six to three and the number of links from seven to four. This represents a 50 per cent decrease in operating time and almost a 50 per cent improvement in over-all transmission.

The example given above typifies the manner in which high-usage trunk groups are utilized. These trunk groups are provided on a tightly engineered basis. That is, they are engineered so that, during the busy hour, a considerable portion of the calls offered to them cannot be accommodated and must be diverted to different, alternate routes. The alternate route employs final trunk groups, and this may introduce one or more extra switches as compared with the more direct first choice routes. This method of trunk engineering results in a very efficient use of trunks in the high usage groups.

Final trunk groups shown in Figure 1(b) are engineered for low delay. That is, each group is provided with enough trunks to carry not only all of the first choice traffic offered to it, but also any overflow from high usage groups which may use it as part of an alternate route. The engineering objective is that, during the busy hour, an average of not more than three calls in 100 will find all final trunks occupied in any given group. With this low delay, it follows that the final trunk groups are used relatively inefficiently in the sense that many of the trunks usually will not be in use continuously during the busy hour. However, a combination of highly efficient high usage groups and relatively inefficient final groups results in an over-all plant with trunk efficiency comparable to that of a ringdown plant. While the number of switching terminations is increased, total network trunk mileage will be less than for ringdown plant giving comparable speed of service.

The type of long distance network described above must be evaluated from an over-all performance standpoint on the basis of the number of delayed calls during the busy hour. First choice routes for a large proportion of the total calls originated will consist largely of high-usage trunk groups. During the busy hour, these groups will be able to carry (with no delays) most of the the calls offered to them (in the order of 70 per cent). The remainder must be completed over alternate routes (final groups). Of these alternate routed calls, and those calls whose first choice route is over final groups, slightly more than three in 100 may be expected to encounter delays in obtaining trunks. Therefore, with a majority of the calls encountering no delays and only slightly more than three in 100

of the remainder encountering delays, the net result will be that, during the busy hour, throughout the network, less than three calls in 100 may be expected to be delayed because of all trunks being occupied.

## IV. RANK AND FUNCTIONS OF SWITCHING CENTERS

The order of importance of a switching center tends to follow the rank of the unit or group with which it is associated. This is so, because the volume of traffic at the more important centers is roughly proportional in quantity and order of priority to the rank of the military unit with which the center is identified. The following paragraphs describe switching centers in descending order of importance.

### 4.1 *Zone Center*

A zone center is established in a military theater as soon as it becomes impractical to provide direct trunk groups between all long distance switching centers. Theater expansion may make it impractical to connect all long distance trunk switching centers to a single zone center. It is then necessary to establish forward and rear zone centers along the axis of communication. As the theater continues to grow the number of zone centers will increase but in any event each zone center in the theater must be connected by direct trunk groups with every other zone center if delays are to be minimized. While a zone center's primary function is to switch toll trunks, it may also take on some of the functions of a lower ranking center.

In Fig. 1(b), the zone centers correspond in importance and usage to the regional centers shown in Fig. 1(a).

### 4.2 *Primary Centers*

All local centers that are so situated strategically that they may be useful as through switching centers for other local centers are logical candidates for consideration as primary centers. To qualify they must either be fairly close to the zone center or have such a large volume of long-distance traffic to the zone center that a direct group of trunks can be justified. The function of a primary center is to serve as a concentration point for long-distance traffic of associated centers of lower rank.

As shown in Fig. 1, the primary center is the military equivalent of the Bell System sectional center.

### 4.3 *Secondary Center*

A secondary center functions to group a large volume of long distance traffic originating from many local centers geographically located around

it. Such centers are established to economize on trunk plant which would be necessary between local centers and primary centers.

As shown in Fig. 1, the secondary center corresponds to what is known as a primary center in the Bell System plan.

### 4.4 *Local Centers*

Local centers serve as access points for calls between local networks and long-distance networks. Similar centers are known as toll centers in the Bell System toll switching plan. They are distinguished from other switching centers in the long-distance network in that they do not connect long-distance trunks together. A local center usually homes on a secondary center but in some instances, when strategically situated, it may home on a primary center or even a zone center.

Local centers are also known as end centers since they are located at the end of the toll axis of communication. As such they must be arranged to connect to a variety of lower echelon switchboards on a ringdown basis, the lower echelon switchboards usually being of the magneto type.

The plan, then, is arranged to provide an integrated long-distance switching network which allows for theater-wide toll switching operations with essentially no call delay. The layout of switching centers is so arranged that a minimum of toll links and toll switching points is used on most toll calls. This arrangement decreases the time necessary to set up a call and provides high grade transmission circuits.

Prior to the development of the plan herein described, most military toll traffic was engineered on a "ringdown" basis. Such a method of operation is time consuming in that on a multi-link connection each operator contacted must in turn ring forward to the next office in line. On disconnect the same procedure must be followed to release the multi-link connection. Also, any recall signals must be repeated by every intermediate operator in the connection.

The new plan speeds up the network operation by reducing the operator work time. The originating operator controls the setting up of a call until the called station is reached. An intermediate operator receives a verbal order from the originating operator, plugs into the desired trunk and has no further work to do on the call until a disconnect is received. Since automatic ringing is provided, the terminating operator also has no further work to do on the call until a disconnect is received.

Meanwhile, the originating operator has a positive indication that the called party has not answered since a "ringing" or "R" lamp is provided in the cord circuit which remains lit until the called party has

answered. Since the originating operator will leave a call as soon as the terminating operator has plugged into the desired line, this "R" lamp will allow him to determine later whether the called party has answered. This avoids operator work time by eliminating the necessity for the originating operator to be associated with a connection or to monitor the connection at intervals to determine its status.

Furthermore, once a call has been established the originating operator receives switchhook supervision from both the calling and called party without any assistance from intermediate operators. If it becomes necessary, the originating operator may recall the terminating operator by actuating a ring-forward key. This recall does not involve any intermediate operators as it would in a "ringdown" plan. In fact the intermediate operators receive no indication that a recall has taken place.

When the calling and called parties both disconnect, the originating operator receives disconnect supervision. Pulling down both cords automatically sends disconnect supervisory signals to the next office in line until the terminating office is reached. Since the intermediate and terminating operators can pull down these connections without challenging on the connection and without ringing forward, the time required to take down a built-up connection is greatly reduced as compared to a "ringdown" system. Hence trunks can be engineered more tightly since the time to build up and release a connection has been reduced.

V. BLOCK DIAGRAM

Although many circuits were of necessity developed to realize the design requirements, only a few basic circuits are involved once an inter-office connection is established. These circuits are illustrated in block diagram form in Fig. 2 in single line notation.

Using Fig. 2 and starting in the upper left-hand corner a multi-link call can be traced through three toll switching offices equipped with switchboards. A common battery line is used at both the originating and terminating end of the call. The cord circuits provide interconnecting facilities at each of the switchboards with an operator telephone circuit bridged to each cord circuit. Interoffice channels are provided by the two-way straightforward trunks interconnected by carrier channels and associated signal converters. In Fig. 2, an alternate inter-office channel is shown using signal-extension circuits instead of a carrier channel.

It will be observed that the systems control board is shown between carrier facilities and the signal converters. This unit consists of a patching jack field and testing equipment. It provides access for testing and patching but is not actively involved in the connection.
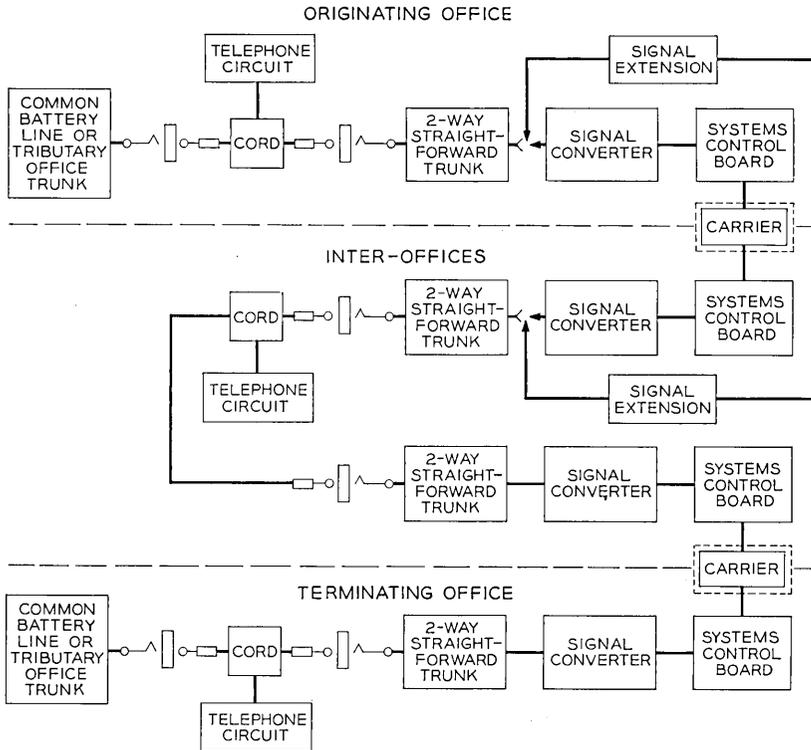
Fig. 2 — Block diagram showing inter-office connections.

VI. BASIC FEATURES OF THE CENTRAL OFFICE

The circuitry for the office may be divided into four general categories according to the functions performed. In the first category is the positional and face equipment which is used by the operator. This includes cord circuits, operator telephone circuits, holding and bridging circuits, local conference circuits, busy back circuits, and trunk and line jack appearances.

The trunk and line relay circuits comprise a second category. Although these circuits are associated electrically with the face equipment, the units themselves may be located as much as 25 feet away from the switchboard positions.

The power circuits include the 48-volt batteries, the battery charger and the power distribution circuits. This is the third category.

The fourth category covers the test and miscellaneous circuits. It includes the local test desk position, the systems control board used for

toll test work, the portable relay test set, the trunk group register circuit, and the signal converters.

## VII. POSITIONAL CIRCUITS

Each position consists of two half-position units mechanically and electrically tied together. A full position contains twenty cord circuits and two operator telephone circuits. Provision is made to split a full position so that two operators working side by side may handle heavy position traffic.

Each position is also arranged so that it may be used as a teletypewriter switching position when voice frequency teletype service is used. With this mode of operation the operator telephone circuit is replaced by a teletypewriter through appropriate switching and the regular straight-forward trunks are used to transmit VF teletypewriter signals. When a trunk signal appears, the operator plugs a cord circuit into the associated trunk jack. The order is then typed on the position teletypewriter instead of being passed verbally. After reading the order, the position operator plugs into the desired trunk to complete the connection.

This type of operation was a design requirement for the over-all system. Therefore, the straightforward trunk and cord circuits were designed to limit distortion in the voice frequency band, so that VF teletypewriter and facsimile signals may be passed without undue distortion.

## VIII. CORD CIRCUIT FEATURES

The universal-type cord circuit is a ten relay circuit arranged for automatic ringing, supervisory signals on both front and rear cords, visual indication that called station has answered, locked-in automatic flashing recall on rear cords and manual flashing recall on front cords, manual re-ring on called lines and trunks, and tripping of ringing during the silent interval after the called party has answered.

This cord circuit is used for station-to-station, station-to-trunk, trunk-to-trunk, and trunk-to-station calls. A low loss permalloy repeat coil is provided in the cord circuit. However, since the repeat coil must be bypassed on a trunk-to-trunk call to reduce transmission losses, means are provided to bypass automatically the repeat coil on this type of call.

The cord sleeve circuit consists of a marginal relay and a sensitive relay in series so that a low resistance trunk or station sleeve condition will operate both the marginal and sensitive relays whereas a high resist-

ance sleeve condition is recognized by the operation of the cord sensitive relay only. In this switchboard all station and trunk sleeve conditions are low resistance with the exception of the straightforward trunk which has a high resistance sleeve condition. This circuitry, therefore, allows the cord repeat coil to be switched out on trunk-to-trunk connections involving straightforward operation. All other trunks are arranged to switch in the repeat coil since these other trunks are used to connect to tributary or terminating offices only.

## IX. LINE AND TRUNK CIRCUIT FEATURES

The line circuit is a conventional common-battery, two-relay circuit.

The two-way straightforward trunk designed for this switchboard provides means for two-way service on a straightforward basis between switchboards. The circuitry is arranged to receive dc signals from the answer cord circuit over the "tip" lead and to transmit dc signals over the "ring" lead to the calling cord circuit.

On through connections when the repeat coil is switched out of the transmission path, a reversal is switched in between the "tip" and "ring" leads so that dc signals applied on the "ring" lead are transposed onto the "tip" lead at the opposite end of the cord.

Other features such as ac idle indicating lamps, locked-in disconnect, and capacitor-resistance terminations on the talking conductors in the idle or disconnect condition, are provided in the straightforward trunk design.

Trunk grouping is simplified by providing a grouping key with each trunk. A three-position toggle switch allows each trunk to be removed from a group of trunks, to be the first or intermediate trunk of a group, or the last trunk of a group. A "make busy" toggle switch is also associated with each trunk so that the switchboard appearance of a trunk may be made busy for maintenance purposes.

The straightforward trunk circuit has been designed as a two-wire trunk; that is, voice transmission in both directions takes place over the same pair of conductors. When used with carrier facilities which, for transmission reasons, are usually four-wire circuits, a signal converter is necessary to provide the hybrid coil used to derive a four-wire circuit from the two-wire trunk circuit. This mode of operation is more fully described in Section X.

As previously mentioned, it is possible to have two toll switching offices within a short geographic distance of each other. Provision of expensive carrier facilities between two such centers would not be economically sound in a civilian toll plant and would certainly be

unsound for a military theater. Accordingly, a unit was designed which within certain limitations of distance would provide for the interconnection of straightforward trunks without using carrier facilities and associated signal converters. This unit is the signal-extension circuit.

The signal-extension circuit is the military equivalent of the Bell System signal lead extension circuit. Its use permits the interconnection of two toll offices on a loop or two-wire basis. The main component of the signal-extension circuit is a polar relay with four balanced windings. The windings of this relay are duplexed on the trunk conductor loop at each end in such a way that in the idle condition, battery and ground applied through the relay windings at each end does not allow the polar relays to operate at either end. When the trunk is seized at one end the battery and ground is reversed through the windings of the local polar relay. The local polar relay does not operate but at the distant end a similar polar relay does operate. This operation causes trunk seizure to take place. When the trunk signal at the distant end has been answered, the battery and ground through the distant polar relay windings is also reversed causing the local polar relay to operate as an off-hook indication. Since the polar relay windings are balanced, the tip and ring leads of the trunk loop remained balanced so that voice transmission over the trunk loop is not impaired by the imposition of the signaling voltages applied to the trunk loop through the polar relay windings. This particular type of signal-extension circuit is classified as a double-pole changer since the voltage applied to both tip and ring of the loop is changed to transmit signals over the loop.

Of course signal-extension circuits must be provided at both ends of the trunk loop with this arrangement. For short loops the saving in equipment over carrier-type operation is a very significant factor.

## X. SIGNAL CONVERTER

When trunk loops between central-office installations exceed the design limits, loop signaling is no longer feasible using signal-extension circuits. Carrier facilities must be utilized for this long haul traffic. Since the carrier facilities in the military plan are four-wire facilities employing voice-frequency signaling and since the straightforward trunks are two-wire trunks employing dc signaling, it is necessary to provide a device which will make these two different types of facilities compatible on a system basis.

The signal converter was designed to perform this translation function. It is similar to the Bell System single-frequency signaling circuit and is an "in band" voice-frequency signaling device. On the four-

wire side, signaling is accomplished by using two different frequencies, one for each direction. This is done to minimize the effect of signal echo currents. An "on-off" signaling arrangement is used, as in the Bell System counterpart, with the signaling tones "on" in both directions during the idle period and "off" on an established connection.

The converter includes a hybrid coil whereby the four-wire circuit for connection to carrier facilities is derived from the two-wire straight-forward trunk appearance. Signaling tones are supplied by two vacuum tube oscillator circuits. The individual channel units, four per cabinet, are combination transmitter and receiver units in that the signal tone is applied to the transmitting side of the four-wire facility under control of the connected trunk circuit. Incoming signal tone or the absence of tone is recognized by the receiver side of the unit and is translated into dc supervisory signals to the connecting trunk. These units are not arranged to transmit dial pulses.

The circuit design of the converter is such that the nominal net trunk transmission loss between switchboards is limited to 3 db. This design permits the use of as many as seven links on a toll connection before the over-all transmission is impaired to the point where it is unacceptable to the military.

## XI. TESTING AND MAINTENANCE FEATURES OF THE SWITCHBOARD

Basic design requirements for the switchboard and associated equipment stressed ease of maintenance. This requirement was met by utilizing an equipment design featuring readily removable units containing one or two circuits only. With this type of design, provision of replacement units makes it possible to relegate repair maintenance to a rear echelon maintenance depot. Such a maintenance plan can be successful however only if sufficient test features are provided at the central office to allow operating personnel to determine readily when an individual replaceable unit is defective.

Test and maintenance components provided are as follows:

1. A test desk position similar to an operator position, for making operational tests on lines, trunks and cord circuits.

2. A system control board for patching toll circuits (4-wire) and for making sectionalized tests and transmission measurements on toll facilities.

3. A portable relay test set for use in testing line and cord circuits in small offices where a test desk position is not justified.

4. A trunk group register unit which records trunk usage data for traffic study purposes.

In addition to the above mentioned test components, it is expected that standard test equipment now in general use will be available.

## XII. POWER CONSIDERATIONS

Two main power sources are required to operate the central office:

1. A 115- or 230-volt ac source, and

2. A 48-volt dc battery source with the necessary ac rectifier to float the 48-volt dc battery supply.

The first power source may be derived from commercial power lines or from any one of the many available standard alternator sets.

The 48-volt dc power supply normally consists of four 12-volt storage batteries and a battery charger, or chargers, floated across the nominal 48-volt battery source.

Distribution of all the various power supplies to the relay units is accomplished through the equivalent of a power distribution board. This unit is known as the power supply control unit. In addition to providing the necessary outlets for distribution of ac and dc power, this unit contains a dc operated rotating ringing machine which provides ringing current and the ringing interruptions required.

## XIII. BATTERY CHARGER

Three 10-ampere rectifier units are provided for each three switchboard positions and associated relay equipment.

The switchboard may include a maximum of 20 positions and provision is, therefore, made to multiple together both the 48-volt storage batteries and the battery chargers.

The battery chargers are so designed that one or more chargers may be used during light load periods to charge the spare set of storage batteries which will be provided.

Each battery charger provides a closely regulated dc supply for the floating battery at 48–52 volts dc or for battery charging at 57.5–62.5 volts dc. It operates on 115/230 volts ac, 50- or 60-cycle, single-phase power.

## XIV. POWER SUPPLY CONTROL UNIT

One control unit per three positions is provided. This unit furnishes the necessary power connections for all types of power used in the central office. Facilities are included for multipling to similar power plants. It is at this point that the prime source, 115 or 230 volts ac, is connected for system wide distribution.

## XV. SWITCHBOARD POSITIONAL EQUIPMENT

The basic switchboard position unit is packaged in a half-section sheet-metal container. Each container has rain proof, removable front and back covers. Half-section units are utilized to keep package weights below 250 pounds.

A full position comprises two identical half-section units mounted side by side and held together by the keyshelf extension, the multiple rack, and the position power adapter supply. The first line up of equipment shown in Fig. 3 shows how the positions are assembled. In this illustration the three positions to the left are operator positions while the extreme right hand position is a test desk position. Fig. 4 shows a rear view of the same equipment. Rear covers are removed on the right-hand position to show the manner in which cord packs are mounted in the half-section cases.

Each half-section case contains five cord packs or ten cord circuits. Shock-mounted plates, top and bottom, protect the relay equipment from shock and vibration in transit. An individual cord pack with the gasketed sheet-metal cover removed is illustrated in Fig. 5. Each vertical row of 10 relays is associated with one cord circuit. When the cover is in place it protects the relays from dust and mechanical damage.

The keyshelf extension shown in Fig. 3 provides a working surface for the operator and contains two operator telephone circuits. These units are of the plug-in type for ease of maintenance. The top of the keyshelf is 30 inches above the floor level thereby providing a comfortable desk height working level for the operator. Special operator's chairs are not necessary since any desk chair available will be satisfactory.

The position dial and dial cord unit is mounted in the middle of the keyshelf between the two half-section cases for use with civilian office dial trunks.

## XVI. MULTIPLE RACKS

All line and trunk circuits must have lamp and jack appearances in the face of the switchboard to provide operator access. The lamps and jacks are packaged in groups of twenty appearances. Such units are known as line or trunk packs. Since each trunk pack appearance has an idle-indicating lamp as well as a trunk answer lamp, the over-all height of the trunk packs is greater than the line pack which provides only answer lamps. This difference in height necessitates two different multiple racks in which the line and trunk packs are mounted. In the illustration, the left position is equipped with only five line packs, and spacers fill up the five unused pack positions.
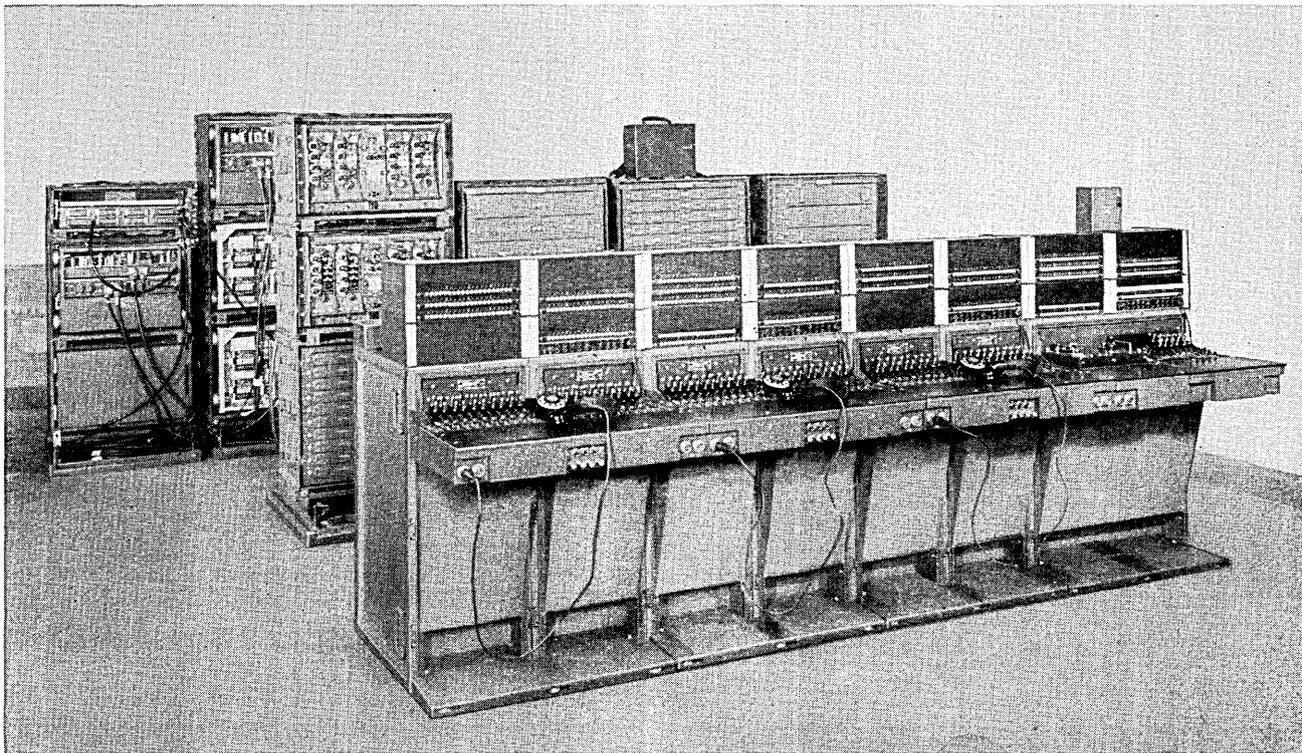
Fig. 3 — Typical three-position installation showing the preferred arrangement of switchboard and associated equipment.
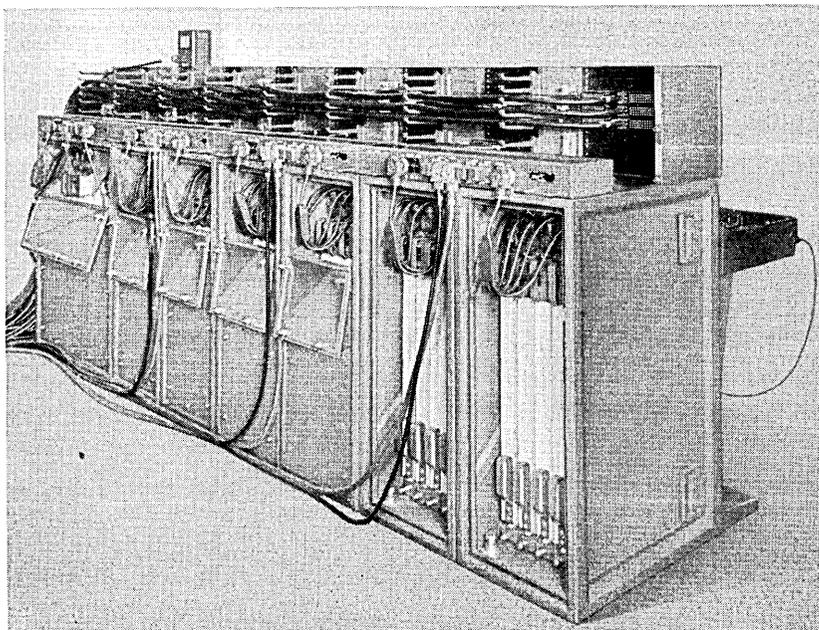
Fig. 4 — Rear view of switchboard position line-up. The right-hand position has both half-section rear covers removed to display the manner in which cord packs are mounted in position cases. The line and trunk multiple mounted on the top front of the position cases shows the orderly manner in which multiple cables are racked.

The rear of the multiple racks is equipped with removable cable racks which hold the multiple patch cables in an orderly fashion. Fig. 4 shows the rear of the position line up with the switchboard multiple on the top of the positions. In this instance a four-panel multiple was used. That is, a particular line or trunk appeared in every fourth panel, there being two panels in front of each full position. Multiple cables have been designed for a four panel multiple but if it is more desirable to confine all trunk and line appearances to one position, a two panel multiple may be installed. Such an arrangement would be suitable for a small office but a large office would undoubtedly require more lines and trunks and hence a four panel multiple.

XVII. GENERAL PURPOSE CASES

Portability and flexibility were the determining factors in the design of the carrying cases for the central office. To meet the desired objectives, five different size carrying cases are provided.
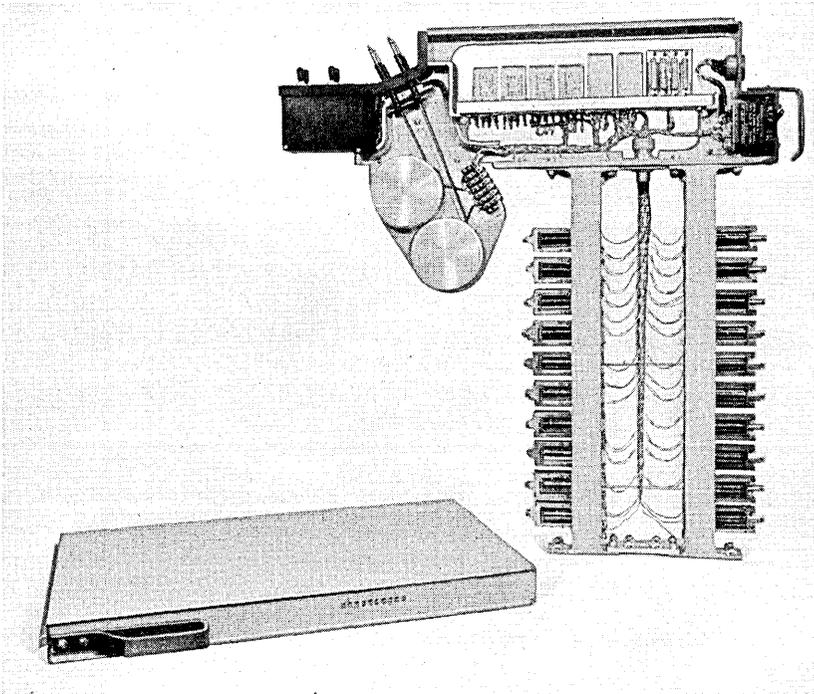
Fig. 5 — Switchboard cord pack containing two cord circuits with the dust cover removed.

Both multiple racks span a full switchboard position. Protruding "V" blocks on the base of each rack recess within indented slots in the position cases. Thumb screws provide means for fastening the multiple rack to the position. Successive multiple racks are fastened by similar means to the lower racks.

The trunk multiple rack is arranged to hold six trunk packs or 120 trunk appearances. The line multiple racks have a greater height and hold ten line packs or 200 line appearances.

The height to which the switchboard multiple may be raised is limited by the height at which an operator may readily see a lighted lamp. Because of this, the multiple is limited to 440 trunks and 960 lines. In an emergency, as many as 40 additional lines could be added with only a slight degradation in service.

Fig. 3 shows the trunk and line multiple fastened on the top of each position in the first line of equipment. In this illustration the multiple rack resting on the position cases is a trunk multiple rack in the position at the left but it contains only one trunk pack instead of the full complement of five. Spacers are provided to fill up the unused pack positions.

The top multiple rack is a line multiple rack.

Since the general-purpose cases used for mounting relay equipment were required to stack one upon another as shown in the rear lines of equipment of Fig. 3, it was necessary to make the base dimensions uniform and to vary the height to accommodate the varying amounts of equipment. Also to facilitate stacking, interlocking cleats or registering plates were riveted to both bottom and top surfaces of these general purpose cases to prevent slippage when stacked. Separate clamps were also designed which clip to the edges of adjacent cases when stacked to hold the stack in both a horizontal and a vertical direction.

Stiffening channels on the insides of the sheet-metal cases are provided. This is necessary to insure a rigid case, since it is necessary to remove both front and back covers before stacking. These channels also serve as points of attachment for shock mounts which hold all interior frameworks upon which relay equipment is mounted.

All cases feature an arrow-head type construction on all edges. These arrow heads absorb sharp blows on the case edges to prevent deformation of the case proper in transit. Four of the five types of cases are shown in Fig. 6 with both covers in place.

One of the five types of cases mentioned is provided to store the numerous quick-disconnect patching cables used throughout this system. These cases are equipped with means for holding down cables and components during shipment so that no damage occurs. These cases are also equipped with a special cover which when mounted flat provides a mounting base for the relay unit stacks. These bases may be observed in the two rear line ups of relay units shown in Fig. 3. Since these bases
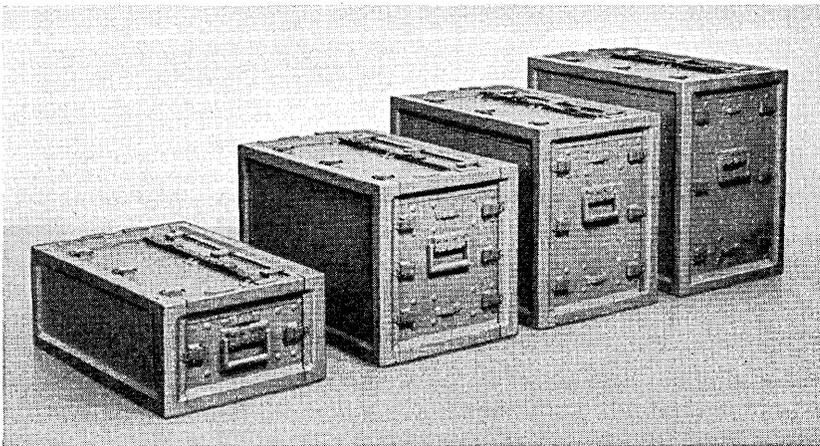


Fig. 6 — Typical general purpose cabinets. Front and rear covers are both removed when the office is installed.

provide about six inches of space under the first unit of a stack, spare cable may be dressed into them both for appearance and safety.

## XVIII. UNITIZED RELAY MOUNTINGS

A unitized relay mounting is provided which permits rapid replacement of a defective unit without the necessity of unsoldering wired connections.

Three types of plug-in units were designed to package the eight different circuits. These units feature a multi-contact plug which mates with multicontact jacks mounted in a multi-unit framework. This multi-unit framework is shock mounted in the appropriate size of general purpose carrying case. Typical examples of these units are shown in Fig. 7.

Each plug-in unit is covered by a gasketed sheet-metal cover which protects the relays and other apparatus from moisture, dirt and dust. The plug-in units are mounted on stainless steel slides riveted to the framework and are held in place in the framework by means of quick-disconnect fasteners. Although three basic designs are used, the individual apparatus mounting plates may be drilled in a variety of ways to accommodate the various types of apparatus employed.

A typical trunk framework is shown in Fig. 8. This framework is shock mounted in a general purpose carrying case. Twenty trunk units
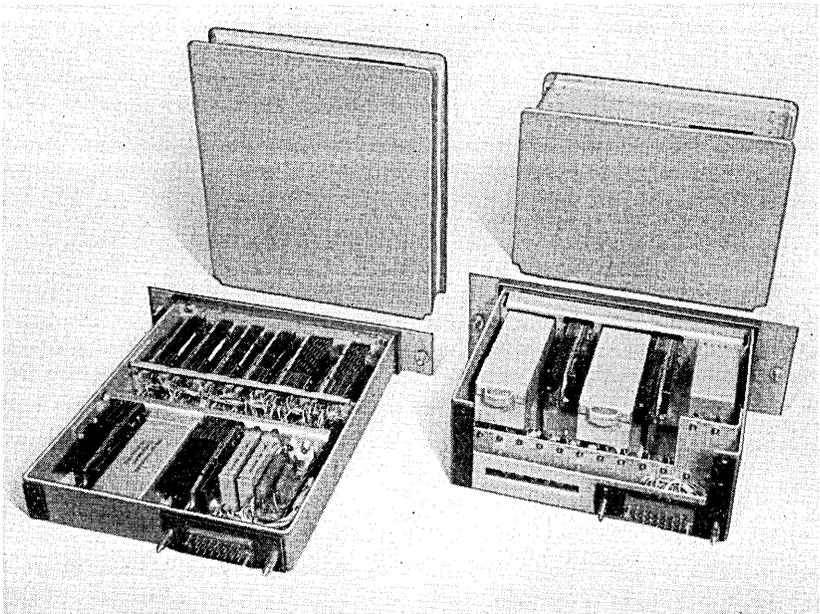


Fig. 7 — Typical plug-in type trunk units with dust covers removed showing in detail the manner in which apparatus is mounted.
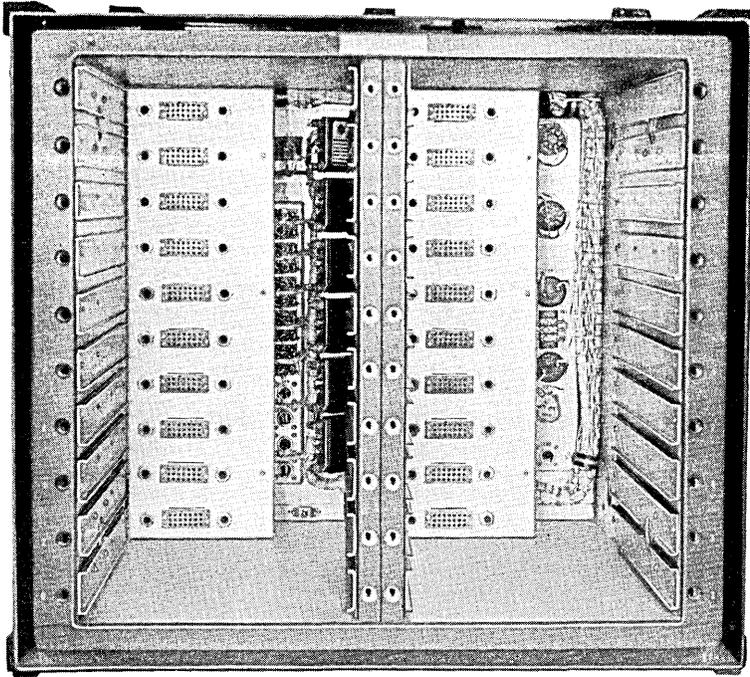
Fig. 8 — Trunk cabinet with the twenty plug-in trunk units removed to show interior construction of cabinet.

mount from the front of this cabinet as shown in Fig. 9. All power distribution and interconnecting leads are brought out on a hinged door on the rear of the unit. Therefore with this type of design, all cable connections are made at the rear of the cabinet but all plug-in units are removed from the front of the unit.

## XIX. FIXED INSTALLATIONS

The central office meets military conditions where mobility and adaptability are prerequisite. Individual units are designed in such a manner that many physical arrangements are possible to meet the variety of conditions that may be encountered. In general, the tactical situation will be such that a permanent installation will be feasible. If such is the case, the floor plan arrangement shown in Fig. 3 is applicable. In this illustration a three position office together with the necessary power and relay equipment is located in an 18 by 21 ft. area. The typical layout is so arranged that all cable patching can be accomplished using a maximum cable length of 25 ft.

Previously, reference has been made to Fig. 3 to clarify the multiple-
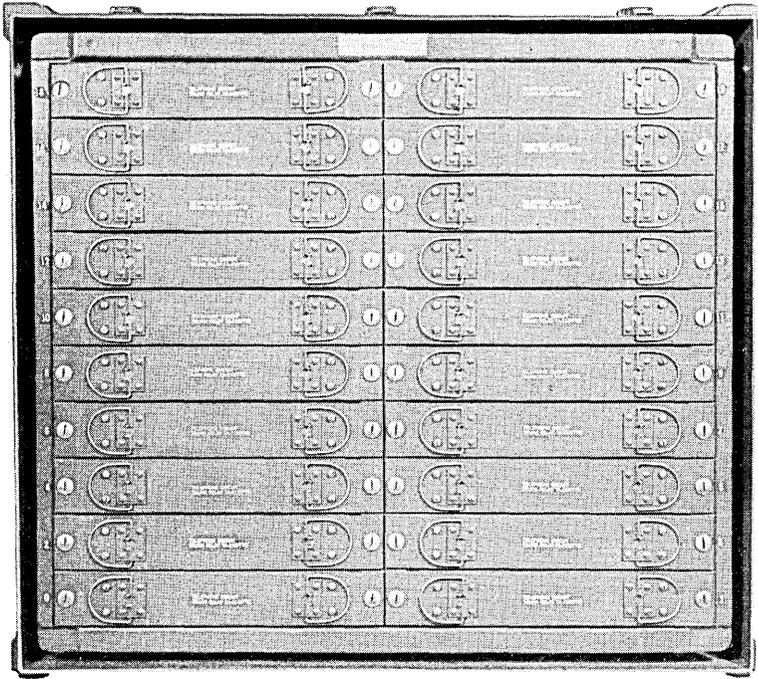
Fig. 9 — Two way straightforward trunk cabinet equipped with twenty trunk units. Individual trunk units may be removed from this side by releasing two fasteners.

rack arrangement. This figure also illustrates the manner in which a fixed installation may be set up. The first line of equipment from left to right shows three identical operator positions. The position at the extreme right is a test desk position used for maintenance purposes.

The second line of equipment in Fig. 3 on the extreme left shows partially two units of signal converters stacked upon a unit containing 20 trunk circuits. The remaining units in the second line-up partially display other trunk units and at the extreme right part of a unit containing 100 line circuits may be seen.

The third line of equipment does not show well except at the extreme left end. The first stack of units shows the cabling side of the systems control board used in making transmission measurements on toll facilities. Not visible in the third line-up are the main distributing frame and the various power units.

XX. MOBILE INSTALLATIONS

Details have been designed which would permit a complete central office to be mounted in a truck and a semi-trailer.

XXI. MAINTENANCE

Ultimately, all detailed maintenance on units of the central office will be handled at a repair depot except for some small maintenance jobs such as lamp or vacuum-tube replacement or the replacement of an equipment unit. To facilitate this type of maintenance all units are readily removable in groups of one or two circuits. The line unit is an exception in that the smallest plug-in assembly houses twenty circuits.

XXII. CORD REEL

Conventional switchboard designs in the Bell System employ 6-foot switchboard cords which hang down from the keyshelf in a cord pit. The cords are kept in position by the use of pulley weights on each cord. Such a design is not acceptable for military use because the space is not available for cord pits and frequent cord rebutting is not tolerable.

A new method of storing the switchboard cordage was developed as shown in Fig. 10.

The design of the cord pack on which this cord reel is used permits cord reels to be assembled on a mounting plate in groups of four, thereby providing the necessary switchboard cords for the two cord circuits housed in one cord pack. These assemblies include reels, cords, plugs, and connecting blocks which permit complete shop assembly and pre-tensioning of the reels prior to final assembly. Fig. 5 shows the manner in which cord-reel assemblies are fastened to cord packs. This method also permits rapid replacement of defective cord reel assemblies without removing the dust covers on the cord pack.

XXIII. 52-CONTACT CONNECTOR

A 52-conductor connector, shown in Fig. 11, is used to interconnect the various units. The receptacle, which mounts in the various units, has stationary contact pins. The cable connector, however, employs a spring-loaded solid pin to which the cable leads are connected. When the cable connector is tightened down on the receptacle, shown in the lower left-hand corner of Fig. 11, the spring-loaded terminals maintain adequate pressure at the point of contact. An additional feature is that a second connector may be attached to the first connector. Fig. 12 illustrates this connection as employed at the line and trunk multiple.

Fig. 10 — Cord reel assembly showing the switchboard cords in the retracted position. This assembly provides access to two cord circuits.
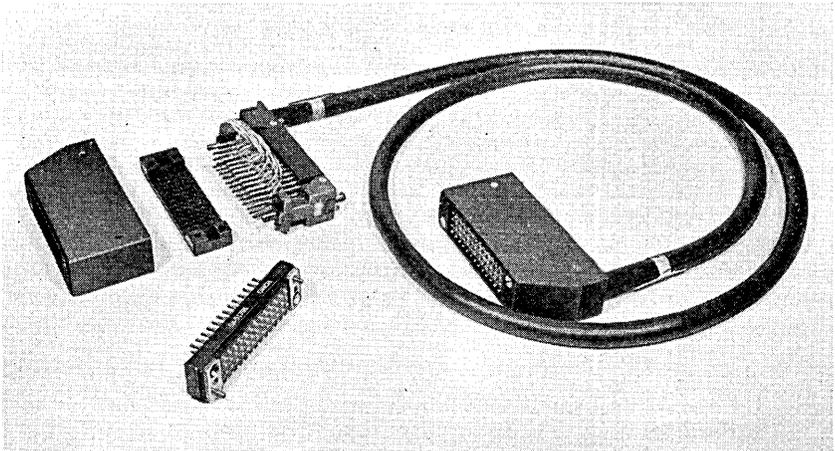


Fig. 11 — Patching cable assembly showing the method of terminating the cable on the 52-pin connectors. The connector in the lower left-hand corner is the mating connector which mounts in each unit and to which the patching cable connects.
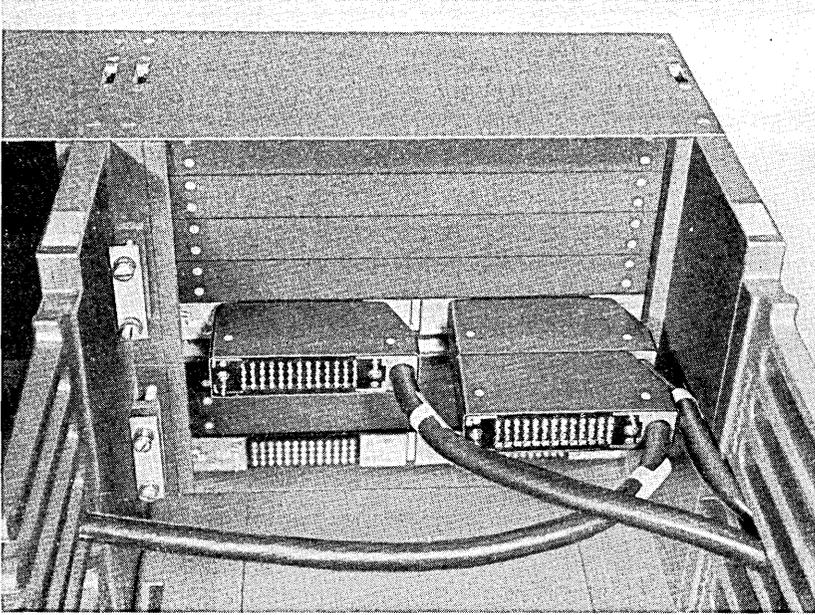
Fig. 12 — Rear view of a line and trunk multiple showing on the right-hand side the manner in which 52-contact connectors are used in establishing multiple trunk and line appearances in the switchboard.
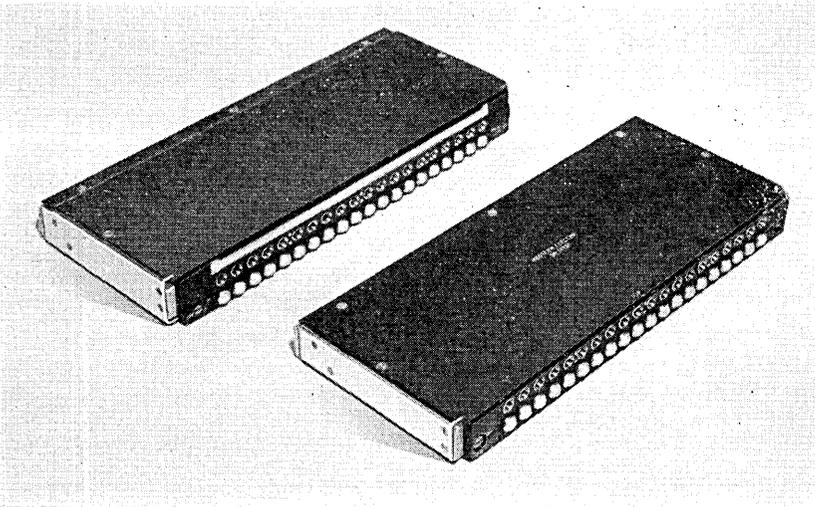


Fig. 13 — Trunk and line packs each containing twenty jack and lamp appearances. The pack on the left is a trunk pack, that on the right a line pack. Connection is made via two 52-pin connectors not shown here but mounted on the rear of each pack.

Soldered wire-wrapped connections are used for connecting wires to both the cable plug and the unit terminal receptacle.

## XXIV. LINE AND TRUNK PACKS

Although the relay equipment for line and trunk circuits is located in carrying cases stacked near the switchboard line-up, provision must be made for the jack and lamp appearance of each line and trunk circuit in the switchboard multiple. The switchboard features self-contained line and trunk appearances in which each pack contains twenty line or trunk appearances. These packs are each equipped with two 52-contact connectors to provide for patching to the relay equipment and to the next multiple appearance. These packs are front mounted in multiple racks designed to accommodate them, and are held in place by removable number plates.

The answer lamps in each line or trunk pack are arranged to feed through a switch mounted in the face of each pack in groups of ten lamp appearances. This permits traffic-load control by providing a means of opening answer-lamp circuits in groups of ten lamps by operating a slotted switch from the face of the multiple. This switch may be operated either by means of a screw driver or a coin.

Fig. 13 shows the two types of pack provided. The pack at the left of the illustration is a trunk pack. The idle-indicating lamps are located under the top horizontal strip on the face of the pack. This strip is translucent allowing a lighted idle-indicating lamp to be seen by the operator. The second horizontal strip shows the 20 lamp caps for the 20 trunk answer lamps. In line with the lamp caps at each extreme end may be seen the slotted switch which actuates switch contacts to open up ten lamp circuits at a time for load control.

The pack shown at the right in Fig. 13 is a line pack. No idle-indicating lamps are provided, thereby reducing the thickness of the line pack over that of the trunk pack. The lamp switch arrangement, however, is the same as that used in the trunk pack.

The design of the line and trunk packs also provides ease of maintenance in that a defective pack may be removed from the front of the switchboard without disturbing adjacent line and trunk packs.

## XXV. ACKNOWLEDGEMENT

The development of the switchboard and associated equipment involved numerous departments both at Bell Telephone Laboratories and at the Kearny works of the Western Electric Company. It was the teamwork, knowledge and ability of all which produced the final design.

# A Proposed High-Frequency, Negative-Resistance Diode

## By W. T. READ, JR.

(Manuscript Received March 2, 1957)

*This paper describes and analyzes a proposed semiconductor diode de-signed to operate as an oscillator when mounted in a suitable microwave cavity. The frequency would be in the range extending from 1 to 50 kmc. The negative Q may be as low as 10 and the efficiency as high as 30 per cent.*

*The diode is biased in reverse so as to establish a depletion, or space-charge, layer of fixed width in a relatively high resistance region, bounded by very low resistance end regions. The electric field has a maximum at one edge of the space-charge region, where hole-electron pairs are generated by internal secondary emission, or avalanche. The holes (or electrons) travel across the space-charge layer with constant velocity, thus producing a cur-rent through the diode. Because of the build-up time of the avalanche, and the transit time of the holes across the depletion layer, the alternating current is delayed by approximately one-half cycle relative to the ac voltage. Thus, power is delivered to the ac signal. When the diode is mounted in an induc-tive microwave cavity tuned to the capacity of the diode, an oscillation will build up. It appears possible to obtain over 20 watts of ac power in con-tinuous operation at 5 kmc.*

## I. DESCRIPTION

This paper discusses a proposed oscillator consisting of a semicon-ductor diode biased in reverse and mounted in a microwave cavity. The impedance of the cavity is mainly inductive and is matched to the mainly capacitative impedance of the diode so as to form a resonant system. We shall show that the diode can have a negative ac resistance so that it delivers power from the dc bias to the oscillation. The negative $Q$ may be as low as 10 and the efficiency as high as 30 per cent.

The principle of operation is as follows: a reverse bias is applied to establish a space-charge, or depletion, layer of fixed width in a relatively weakly doped region bounded by highly doped end regions. A possible

structure is the $n^+$-p-i-$p^+$ structure, where the $+$ denotes high doping and i means intrinsic. This structure is shown in Fig. 1(a). The field distribution under reverse bias is shown in Fig. 1(b). The voltage is always well above the punch-through voltage, so that the space-charge region always extends from the $n^+$-p junction through the p and i (intrinsic) regions to the i-$p^+$ junction. The fixed charges in the various regions are shown in Fig. 1(b). A positive charge gives a rising field in going from left to right. A positive field makes holes move to the right. The maximum field, which occurs at the $n^+$-p junction, is of the order of several hundred kilovolts per cm, so that hole-electron pairs are generated by internal secondary emission (also called multiplication or avalanche). The electrons go immediately into the $n^+$ region. The holes move to the right across the space-charge region. The field throughout the space-charge region is in the range (above about 5 kilovolts per cm) where carriers move with constant velocity independent of the field. For practical purposes we can forget about the electrons in the space-charge
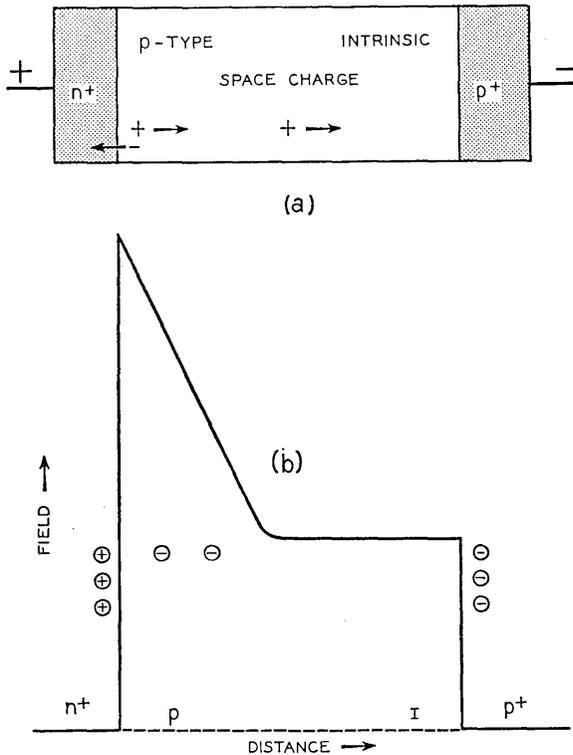


Fig. 1 — The structure (a) and field distribution (b) under reverse bias.

region and consider that current is generated at the $n^+$-p junction, so that only the holes move in the space-charge region. Thus, the physical picture is as follows: a current of holes $I_0(t)$ is generated at the $n^+$-p junction. The holes move to the right and traverse the space-charge region, moving with constant velocity, $v$. In silicon $v$ is about $10^7$ cm sec$^{-1}$. The transit time of a hole across the space-charge region is $\tau = W/v$, where $W$ is the width of the space-charge region. Throughout the discussion we shall illustrate various quantities by giving numerical values for the case $W = 10^{-3}$ cm. Thus $\tau$ would be $10^{-10}$ sec.

The holes moving across the space-charge region produce a current $I_e(t)$ in the external circuit. It can be shown that $I_e$ is equal to the average current in the space-charge region. Since velocity is constant, $I_e$ is simply $v/W = 1/\tau$ times the total charge of the moving holes. (We shall assume throughout that all quantities refer to unit area of junction.) Suppose, for example, that a pulse of holes of charge $\delta Q$ is suddenly generated at the $n^+$-p junction. Immediately, a constant current $I_e = \delta Q/\tau$ begins to flow in the external circuit, and continues to flow during the time, $\tau$, that the holes are moving across the space-charge region from the $n^+$-p junction to the $p^+$ region. Thus, on the average, the external current $I_e(t)$ due to the moving holes is delayed by $\tau/2$ relative to the current $I_0(t)$ generated at the $n^+$-p junction. We shall show (in the discussion of multiplication) that the current $I_0(t)$ is delayed by one-quarter of a cycle relative to the ac voltage. Thus, to get a total delay of one-half cycle, we want the delay $\tau/2$ due to transit time to be one-quarter of a cycle. The cavity should therefore be tuned to give a resonant frequency $(2\tau)^{-1}$.

The conductive current $I_e(t)$, which arises from carriers moving through the space-charge layer, should be distinguished from the displacement, or capacitative, current $I_c$, which charges and discharges the diode regarded as a capacitor. This current, $I_c$, supplies the variation in charge at the edges of the space-charge region, where the field changes abruptly and, for practical purposes, can be considered discontinuous. Since $I_c$ is $90°$ out of phase with the voltage it contributes nothing to the power.

## 1.1 Multiplication

Carriers moving in the high field near the $n^+$-p junction acquire enough energy to knock valence electrons into the conduction band, thus producing hole-electron pairs. The rate of pair production, or multiplication, is a sensitive nonlinear function of the field. By proper doping, the field can be given a relatively sharp peak so that multiplication is confined to a very narrow region at the $n^+$-p junction. The multiplication

rate can be regarded as a function of the peak field, $E_0$. At a critical field, $E_c$, breakdown will occur; that is, any current will be self-sustaining; every pair produced will on the average produce one other pair. We shall neglect the thermally generated reverse saturation current, which will be small compared to the generated current $I_0(t)$. Then, when the field is above $E_c$, the current $I_0(t)$ will be more than self-sustaining and will build up. When the field is below the critical breakdown field $E_c$, the current is less than self-sustaining and will die down.

In operation, the diode is biased so that the peak field is above $E_c$ during the positive half of the voltage cycle and below $E_c$ during the negative half. Hence the current $I(t)$ builds up during the positive half and dies down during the negative half. Therefore $I_0(t)$ reaches its maximum in the middle of the voltage cycle, or one-quarter of a cycle later than the voltage.

We have assumed so far that the field varies in phase with the voltage. This will be a good approximation if the carrier space-charge can be neglected. This can be seen from Fig. 1(b). If the voltage is above the punch-through voltage, an increase in voltage simply raises the whole field distribution throughout the depletion layer. That is, additional charges simply appear at the edges of the space-charge layer. Hence the field at each point varies in phase with the voltage. The space charge of the carriers will, however, affect the shape of field distribution in the space-charge region. When the current becomes too large, the carrier space-charge cannot be neglected. As we shall see, the effect of carrier space-charge is to reduce the delay between the voltage and the current generated, $I_0$. In practice this limits the dc bias current.

When the dc bias is small enough so that carrier space charge can be neglected, the operation can be summarized as follows: the peak field varies in phase with the voltage and generates (at one edge of the depletion layer) a current delayed by 90°. This current gives rise to a current through the diode delayed by 90° relative to the current generated and, therefore, by 180° relative to the voltage.

## 1.2 *Diode and Cavity*

The capacity of the diode and the inductance of the cavity determine the frequency, $f$, of oscillation. We have seen that the optimum frequency, $f$, is such that the transit time, $\tau$, is one-half cycle. Hence the angular frequency, $\omega$, should be

$$\omega = 2\pi f = \frac{\pi}{\tau} = \frac{\pi v}{W}. \tag{1}$$

For $W = 10^{-3}$ cm, the frequency, $f$, would be 5,000 mc.

We have seen that the current into the diode is $I_e + I_c$, where $I_e$ is a conductive current due to holes moving in the space-charge region and $I_c$ is a displacement current required to charge the diode regarded as a capacitor. Thus, the diode acts like a capacity and negative resistance in parallel.

The capacity is related to the width, $W$, of the depletion layer by

$$C = \frac{\kappa}{4\pi W},$$
$$= \frac{1.06 \times 10^{-12}}{W} \quad \text{for Si,} \tag{2}$$

where $C$ is in farads per cm$^2$ and $W$ in cm. The negative conductance of the diode will be small compared to $\omega C$, so the admittance is mainly that of the capacity. Thus for $W = 10^{-3}$ cm the admittance will be about 30 mhos per cm$^2$.

To make a resonant system we mount the diode in a microwave cavity having a mainly inductive impedance. The inductance, $L$, of the cavity is chosen so that the cavity and diode together have a resonant frequency $f = \frac{1}{2}\tau$.

Fig. 2 shows a possible cavity with the diode mounted in it. The diode should have narrow end regions so that the depletion layer is as close as possible to the metal base and center post of the cavity. Then the heat generated will flow away rapidly. The cavity is tuned by adjusting the vertical position of the plunger forming the top of the cavity. There must be some break in the cavity so that a dc bias can be applied. This is accomplished in Fig. 2 by insulating the center lead. The break could also be made by insulating the walls from the base. On the right is an outlet through which the ac power is tapped.
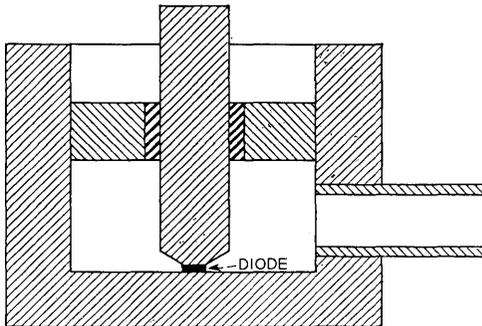


Fig. 2 — Diode mounted in microwave cavity.

The impedance, $\omega L$, of the cavity must be equal to the impedance of the diode. If the diode is circular with radius $R$, then the junction area is $\pi R^2$ and the impedance, $(\pi R^2 \omega C)^{-1}$. Thus the area of the diode is limited by how small the impedance of the cavity can be made. Setting $\omega L = (\pi R^2 \omega C)^{-1}$ and using (1) and (2) we have

$$\left(\frac{R}{W}\right)^2 = \frac{4}{\pi \kappa v \omega L},$$

$$= \frac{9.5 \times 10^3}{\omega L} \quad \text{for Si,} \tag{3}$$

where $\omega L$ is in ohms. It would be difficult to make $\omega L$ less than about 10 ohms. Hence $R$ could not be more than about $30W$.

## 1.3 *Stable Operation*

We have shown that the diode will deliver power, so an oscillation will build up. In Section III we shall calculate exactly the $Q$ for the small signal case where everything is linear. For reasonable values of the direct current bias, $I_d$, the $Q$ can be made negative and as low in magnitude as desired. As the oscillation builds up, the behavior becomes highly nonlinear. In the practical range of operation the field will vary by a fraction — probably less than 20 per cent — of its average value but the current will vary by orders of magnitude. The current $I_e$ approaches a square wave, being negligibly small during the positive half of the ac voltage cycle and almost constant during the negative half cycle. Since the direct current $I_d$ is the average conductive current, it follows that the amplitude of variation of $I_e$ is approximately equal to $I_d$. If $V_a$ is the ac voltage amplitude, the ac power delivered $P_r$ is found to be

$$P_r = \frac{2}{\pi} I_d V_a \tag{4}$$

per unit area of junction. Thus, if the dc bias is applied by a constant direct current generator, the power delivered is proportional to $V_a$. The energy stored in the capacity is proportional to $V_a^2$. Hence, $-Q$ is proportional to $V_a$. In other words, if the amplitude increases, the stored energy, or energy of oscillation, increases faster than the energy delivered per cycle. This is the condition required for a stable oscillation to be possible. At the stable operating point, $-Q$ of the diode is equal to the effective $Q$ of the cavity. If the amplitude increases, the energy delivered by the diode increases less than the energy lost to the cavity. Thus, the amplitude decreases. In the same way a decrease in amplitude is also self-correcting.

The total power output is $P_r$ times the area $\pi R^2$ and is therefore pro-

portional to $R^2 I_d V_a$. The maximum $R$ is given by (3). The maximum allowable $I_d$ is limited by the fact that carrier space-charge cannot be too large, as we now show.

### 1.4 Space Charge of the Carriers

The holes moving across the space-charge region will affect the field distribution. As shown in Fig. 1(b) a positive charge gives a rise in field (in going from left to right) and a negative charge, a decrease in field. Thus, the effect of the holes will be to oppose the fixed negative charge in the depletion layer, and hence to flatten the field distribution shown in Fig. 1(b). Therefore, for a given voltage, the effect of the holes is to reduce the peak field at the $n^+$-p junction. The multiplication rate increases with the peak field. Thus, the current generated has a space charge that tends to reduce the rate of current generation. That is, the current tends to shut itself off. Instead of building up throughout the positive half of the voltage cycle, $I_0(t)$ builds up until the carrier space-charge has reduced the peak field below the sustaining field $E_c$. Then the current decreases. Thus, $I_0$ reaches its peak *before* the middle of the cycle. This reduces the delay and hence the power. Increasing the current, therefore, increases the ac power only up to a certain point where carrier space-charge begins to spoil the phase relations.

We have seen that in the practical range of operation the current $I_e$ varies by approximately its dc value $I_d$. A current $I_d$ corresponds to a total carrier charge $\tau I_d$ in the space-charge layer (since $I_e$ is the average current in the space-charge layer and all carriers travel with velocity $v = W/\tau$). If the carrier space-charge is to be neglected it must be small compared to the charge $CV_a$ that produces the voltage variation. Thus we want $\tau I_d \ll CV_a$. Actually, as we shall show in Section III, it is sufficient to take $I_d \tau = CV_a/2$. Since the period of oscillation is $2\tau$, this becomes

$$I_d = \frac{\omega}{2\pi} CV_a. \tag{5}$$

If $I_d$ is no greater than this the current and voltage will be roughly 180° out of phase and we can use (4) for the power per unit area. For larger $I_d$ the increase in power would be more than offset by the effect of carrier space-charge on the phase relations.

### 1.5 Power Output

Combining (3) with (5) we have, for the total ac power output

$$\pi R^2 P_r = \frac{V_a^2}{\pi^2 \omega L}. \tag{6}$$

As we have seen, the impedance $\omega L$ of the cavity could be as low as 10 ohms. Now consider the maximum $V_a$. This is limited by the dc voltage $V_d$. It is seen from Fig. 1(b) that if $V_a$ is too large the field in the intrinsic region will be reduced to zero during the negative voltage cycle. By careful doping, the voltage at punch-through can be made small compared to the dc bias at breakdown. Even allowing for the reduction in field due to the carrier space-charge, we will be safe with $V_a = V_d/2$. This gives an efficiency of about 30 per cent. From (6) the power becomes

$$\text{Power Output} = \pi R^2 P_r = \left(\frac{V_d}{2\pi}\right)^2 \frac{1}{\omega L}. \tag{7}$$

The dc voltage is limited by the following considerations: (a) In the negative half cycle the field must not fall below about 5 kilovolts per cm. Otherwise carrier velocity will depend on field; so the carriers will slow down during the negative half of the ac voltage cycle and thus reduce the power. This can be avoided if $V_d$ is at least $10^4 W$ volts, where $W$ is in cm. (b) To localize the multiplication the field must remain well below $E_c$ except near the $n^+$-p junction. This will be so if $V_d = 2V_a$ is less than about $(2/5)E_c W$. Then the maximum field in the intrinsic region will always be less than $0.6\ E_c$. Thus we have $10^4 W < V_d < 0.4\ E_c W$.

For $W$ as large as $10^{-3}$ cm the critical field will be about 350 kilovolts per cm. For $W$ as low as $10^{-4}$ cm the multiplication would have to be confined to a region no more than $10^{-5}$ cm in width. This would require a somewhat steeper field gradient and a higher critical field — say about 650 kilovolts per cm. Thus for $W = 10^{-4}$ cm, or 50 kmc, $V_d$ would have to be less than 26 volts. For $\omega L = 10$ ohms, the maximum power output would then be less than 2 watts. At $W = 10^{-3}$ cm, or 5 kmc, the maximum $V_d$ would be 140 volts, which gives a power of 50 watts. At $W = 10^{-2}$, or 500 mc, the maximum voltage would be 1400 volts, and the minimum 100 volts. The power would then be between 25 watts and 5 kw.

1.6 *Zener Current*

Chynoweth and McKay[1] have shown that in sufficiently narrow junctions, where breakdown occurs at around 10 volts or less, the current is generated not by multiplication but by internal field emission. That is, electrons tunnel from the valence to conduction bands. This is known as Zener current. We shall show in Section IV and Appendix E that the device will operate on Zener current, but much less effectively. This,

even more than the limited power, may limit the minimum $W$ and hence the maximum frequency that would be practical.

## Experiments

An experimental program has been undertaken to construct and test a device operating by avalanche.

## II. CURRENT AND SPACE CHARGE

In this section we consider the physics of multiplication, space-charge and carrier flow in more detail and obtain the equations that govern the behavior of the diode.

### 2.1 Multiplication

Electrons moving with velocity $v$ generate hole-electron pairs at a rate of $\alpha v n$, where $n$ is the electron concentration and $\alpha$ is the ionization rate. By definition $\alpha$ is the number of pairs produced on the average by an electron in moving unit distance. Thus $\alpha^{-1}$ is the average distance between ionizations. We shall be dealing with the case where each pair produces roughly one other pair; thus $\alpha^{-1}$ is of the order of the width of the narrow region near the $n^+$-p junction where the multiplication is localized. For fields below about 600 kilovolts per cm $\alpha$ can be considered a function of field, $\alpha = \alpha(E)$. For larger fields the electrons do not have time to get into equilibrium with the field in the average distance $\alpha^{-1}$ between collisions. We shall be dealing with fields where $\alpha = \alpha(E)$.

McKay[2] has measured $\alpha$ as a function of $E$ for electrons in silicon. Fig. 3 is a plot of the best fit curve to McKay's data, together with the theoretical curve calculated by P. Wolff[3]. A straight line of slope 6 is also shown and is seen to be a good fit to the data below about 500 kilovolts per cm. In the formulas we shall use the general relation $\alpha \approx E^m$ and take $m = 6$ for numerical calculations.

## Assumptions

We shall make the simplifying assumption that $\alpha = \alpha(E)$ is the same for both holes and electrons. Actually this is not so. However, $\alpha$ is so sensitive to field that the difference in field for the same $\alpha$ is small. Any difference in the $\alpha$'s will alter both the carrier and field distributions so as to favor the carrier with the lower $\alpha$. Consequently, it is believed that this assumption will not give misleading results.

We also assume that holes and electrons travel with the same velocity in the high fields involved. This will not lead to serious error since the motion of one type of carrier always plays a minor role.

## 2.2 *DC Case*

We take the $x$ axis normal to the junctions with $x = 0$ at the $n^+$-p junction and $x = W$ at the i-$p^+$ junction, as shown in Fig. 4. It can be shown that the direct current $I$ is related to the thermally generated re-
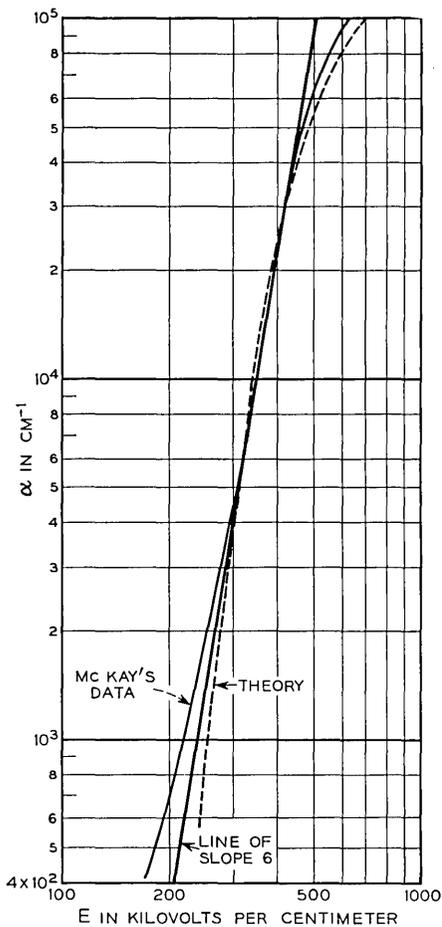


Fig. 3 — Ionization rate vs electric field.

verse saturation current $I_s$ by

$$\frac{I_s}{I} = 1 - \int \alpha(E) \ dx \qquad (8)$$

where the integral is taken over the space-charge region. See, for example, McKay[2]. We will be dealing with space-charge regions where most of the multiplication occurs in a narrow region, which we shall call the *multiplication region*, near $x = 0$. Thus we need take the integral in (8) only from $x = 0$ to $x = x_1$, where $x_1 \ll W$ is the width of the narrow multiplication region.

Breakdown occurs, that is, the direct current becomes infinite, when $\int \alpha \ dx = 1$. Physically this means that each hole-electron pair generated in the multiplication regions will, on the average, generate one other pair. We now consider how $\int \alpha \ dx$ depends on the peak field $E_0$ for various field distributions including the one shown in Fig. 4.

(a) The simplest case is that where the field is a constant $E_0$ in the narrow multiplication region. This would correspond to a short region of high constant field followed by an abrupt drop to a much lower field.
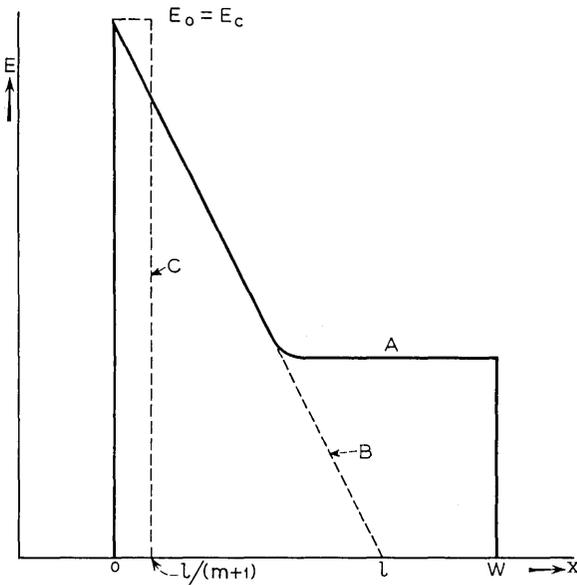


Fig. 4 — Field distribution at breakdown.

Then

$$\int \alpha \, dx = \int_0^{x_1} \alpha(E_0) \, dx = x_1 \alpha(E_0).$$

So if $\alpha$ is proportional to $E^m$

$$\int \alpha \, dx = \left(\frac{E_0}{E_c}\right)^m,$$

where $E_c$ the critical field for breakdown is given by $\alpha(E_0) = 1/x_1$.

(b) Next consider the linear field distribution shown in Fig. 4. Curve A is the field distribution at breakdown. We can take $E = E_0 - kx$ in the multiplication region. The slope $k$ is proportional to the charge in the p-region. We shall neglect the effect of the carriers on the space charge. Then $k$ will be proportional to the fixed charge in the p-region and will be constant independent of current. In practice $k$ would be well above $E_c/W$. In order for the carriers to produce a field gradient comparable to this, the current would have to be of the order of $(\kappa v/4\pi)$ $(E_c/W)$, or several thousand amperes/cm². Actual currents will be much smaller than this.

In Fig. 4 the flat section of field will contribute negligibly to the multiplication, so we can replace the actual field distribution $A$ by the tangent curve $B$. Then

$$\int \alpha \, dx = \left(\frac{E_0}{E_c}\right)^{m+1}, \tag{9}$$

where $E_c$ is given by $\alpha(E_c) = (m + 1)/l$ and $l = E_c/k$ is the zero intercept of the tangent curve $B$ at breakdown. Thus the peak field $E_c$ at breakdown is equal to the breakdown field for a constant field region of width $l/(m + 1)$, as shown by curve $C$ in Fig. 4.

(c) For a linear gradient junction the field varies parabolically with $x$ and

$$\int \alpha \, dx = \left(\frac{E_0}{E_c}\right)^{m+1/2}.$$

In what follows we shall use (9), which applies to a linear field, or abrupt junction. This will be perfectly general if $m$ is left arbitrary. The results obtained for $m = 6$ will be a good approximation for a linear gradient junction, since the difference between 7 and 6.5 is within the experimental uncertainty in $m$.

As an example, consider a step junction with $m = 6$ and $W = 10^{-3}$ and let $l = 0.7W$. Then the critical field $E_c$ is given by $\alpha(E_c) = 7/l = 10^{-4}$ cm. From Fig. 3, $E_c = 350$ kilovolts/cm.

## 2.3 *Time Dependence*

In the dc case $\int \alpha \, dx$ cannot be greater than unity. That is, we cannot get above breakdown. This is not necessarily so for rapidly varying fields. We now derive a differential equation for the current as a function of time. This will reduce to (8) in the dc case.

Let $p$ and $n$ be the hole and electron densities respectively. The corresponding currents are $I_p = qvp$ and $I_n = -qvn$, where $q$ is the electronic charge. Hole-electron pairs are being generated at a rate $\alpha v(n + p)$. This is so large compared to the rate of thermal generation that the latter can be neglected. The equations of continuity then become

$$\frac{\partial p}{\partial t} = -\frac{1}{q}\frac{\partial I_p}{\partial x} + \alpha v(n + p), \tag{10}$$

$$\frac{\partial n}{\partial t} = \frac{1}{q}\frac{\partial I_n}{\partial x} + \alpha v(n + p), \tag{11}$$

where $\alpha = \alpha(E)$. The three variables, $p$, $n$ and $E$ are determined by (10) (11) and by Poisson's equation and the boundary conditions at the junctions. An exact solution is impractical. Instead, we make an approximation based on the fact that most of the generation is confined to the narrow multiplication region near $x = 0$. This is shown in Fig. 5. The multiplication region extends from $x = 0$ to $x = x_1$. The width $x_1$ is a small fraction of the width $W$ of the space-charge region.

## 2.4 *Assumptions*

We shall assume that (a) *all* of the multiplication occurs in the relatively narrow multiplication region and (b) the total current $I(x, t) = I_p(x, t) + I_n(x, t)$ in the multiplication region is a function $I(x, t) = I_0(t)$
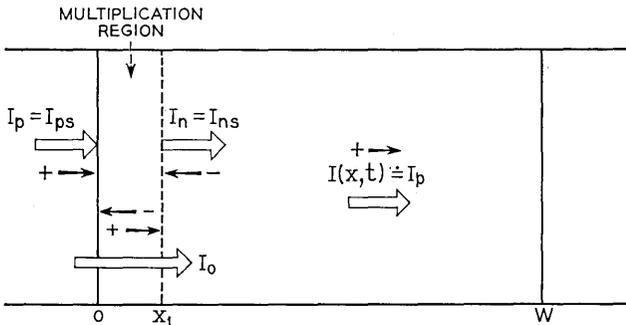


Fig. 5 — Boundary conditions on the multiplication region.

of time only. This, of course, will not be true of $I_p$ and $I_n$ individually in either the ac or dc cases. Assumption (a) will be valid if the net acceptor concentration in the p-region is sufficient to give a sharp drop in field and hence to localize the multiplication. The second assumption will be a good approximation provided the current does not vary by a large percentage in the time $\tau_1 = x_1/v$ required for carriers to cross the multiplication region. This will not be so at large amplitudes. However, the errors for rising and falling currents tend to cancel, so the equation is right on the average. As we shall see, at large amplitudes only the average is involved.

### 2.5 Differential Equation for $I_0(t)$

Adding (10) and (11), using the assumption $I_p + I_n = I_0(t)$ and integrating from $x = 0$ to $x = x_1$ gives

$$\tau_1 \frac{dI_0}{dt} = - [I_p - I_n]_0^{x_1} + 2I_0 \int_0^{x_1} \alpha \, dx, \tag{12}$$

where $\tau_1 = x_1/v$ is the transit time across the multiplication region. The boundary conditions are shown in Fig. 5. The hole current at $x = 0$ consists entirely of the reverse saturation current $I_{ps}$ of holes thermally generated in the $n^+$-region; these have moved to the $n^+p$-junction by diffusion. Thus at $x = 0$, $I_p - I_n = 2I_p - I_0 = 2I_{ps} - I_0$. At $x = x_1$ the electron current consists of the reverse saturation current $I_{ns}$ of electrons thermally generated both in the space-charge region and in the $p^+$-region, so $I_p - I_n = -2I_{ns} + I_0$. With these boundary conditions, (12) becomes

$$\frac{\tau_1}{2} \frac{dI_0}{dt} = I_0 \left( \int_0^{x_1} \alpha \, dx - 1 \right) + I_s. \tag{13}$$

In the dc case $I_0$ is the direct current $I$, so this reduces to (8).

The condition for breakdown is $\int_0^{x_1} \alpha \, dx = 1$. If a field that satisfies this is suddenly applied, $I_0$ will increase linearly at a rate of $2I_s/\tau_1$ and become infinite. If a larger field is applied, $I_0$ will approach infinity exponentially. For a smaller field $I_0$ will approach a finite value.

The integral $\int \alpha \, dx$ over the multiplication region depends very little on carrier space-charge. Hence it will be negligibly affected by the small differences in carrier distribution between the ac and dc cases. We can therefore combine (9) and (13) to obtain a differential equation

$$\frac{\tau_1}{2} \frac{d}{dt} \ln I_0 = (E_0/E_c)^{m+1} - 1 + \frac{I_s}{I_0} \tag{14}$$

relating the current $I_0 = I_0(t)$ in the multiplication region to the peak field $E_0 = E_0(t)$.

In most practical cases the current $I_0$ will be so large compared to $I_s$ that the effect of $I_s$ can be neglected. The correction due to $I_s$ and the detailed formulas for evaluating the effect are given in Appendix D.

### 2.6 *Example*

At low enough amplitudes of oscillation we can expand $E_0/E_c$ in powers of $E_0/E_c - 1$ and retain only the linear term. Then, neglecting $I_s$, equation (14) becomes

$$\frac{d}{dt} \ln I_0 = \frac{2(m+1)}{\tau_1} \left( \frac{E_0}{E_c} - 1 \right). \tag{15}$$

If $I_0$ is to be periodic, then $E_0$ must be periodic and the dc bias must be such that the average $E_0$ is $E_c$. Suppose we apply a periodic voltage with the proper bias so that $E_0 = E_c + E_a \sin \omega t$, where $\omega$ is the optimum frequency $\pi/\tau$. Then

$$\ln \frac{I_0(t)}{I_0(0)} = \frac{2(m+1)}{\pi} \frac{\tau}{\tau_1} \frac{E_a}{E_c} \left( 1 - \cos \frac{\pi t}{\tau} \right). \tag{16}$$

Fig. 6 shows the field and $lnI_0$ as functions of time. Suppose $\tau/\tau_1 = W/x_1 = 20$ and $m = 6$. Then even if the amplitude $E_a$ of variation of $E_0$ is as small as 1.3 per cent of $E_c$, $I_0$ will vary by a factor of 10 over a cycle. Thus we can have small signals in field and voltage but large signals in current. We shall call this the intermediate range of amplitudes.

From (16) the maximum value of $I_0$ is seen to occur in the middle of the cycle, where $t = \tau$. Thus, if $I_0$ varies by a large factor, the current is generated mainly in a pulse in the middle of the ac voltage cycle, as shown in Fig. 6.

Actually the space charge of the current will affect the $E_0(t)$ curve, which will not be exactly sinusoidal for a sinusoidal voltage. However, this does not affect the conclusion that a small field and voltage signal can give a large current signal and that the current $I_0$ approaches a pulse as the ac amplitude increases.

### 2.7 *Carrier Space-Charge*

We have now dealt with the multiplication region, and obtained an equation relating $I_0$ and $E_0$. It remains to consider the rest of the space-charge region, where current generation is negligible. We shall also neglect the reverse saturation current $I_{ns}$ of electrons, which will be negligible compared to the total current $I(x, t)$. From Poisson's equation

we shall derive another relation between $I_0$ and $E_0$. This will involve also the voltage $V(t)$ and together with (14) will uniquely determine $I_0(t)$ and $E_0(t)$ for any $V(t)$.

### 2.8 *Physical Picture*

The physical picture is shown in Fig. 7(a). The width of the narrow multiplication region is small compared to the total width $W$ of the space-charge region. Therefore, in treating the current, space-charge and field distributions throughout the whole space-charge region, we shall assume that the multiplication region has zero width so that all current is generated at $x = 0$. Then $I_0(t)$ is a current of holes flowing out of $x = 0$, and the only carriers in the space-charge region are the holes.

The current $I(x, t)$ at any point $x$ and time $t$ is

$$I(x, t) = I(0, t - x/v) = I_0(t - x/v). \tag{17}$$

Thus the entire hole and current distributions are given in terms of $I_0(t)$.
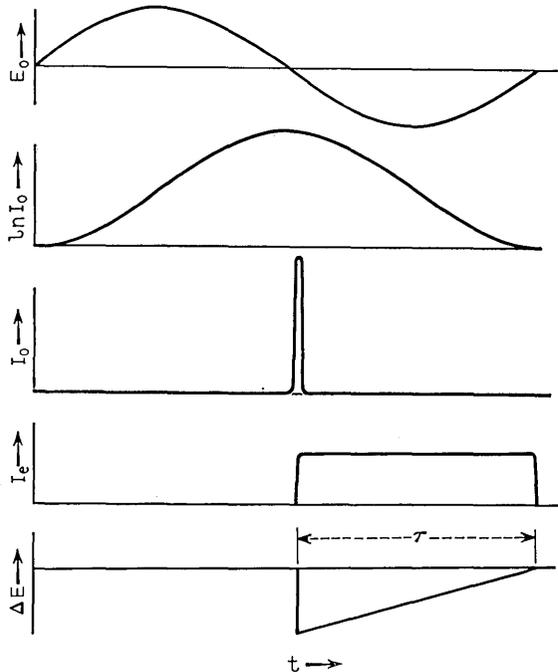


Fig. 6 — Case of a sharp current pulse.

Next consider the effect of the holes on the space charge and the field. The field distribution $E_p(x)$ at punch-through is shown by the dashed curve in Fig. 7(b). If no current flowed the field distribution for any voltage above the punch-through voltage would be simply $E_p(x)$ plus a constant determined by the voltage. The fixed negative charge in the p-material gives a drop in field across the space-charge, or depletion, layer. The positive space charge of the holes opposes that of the acceptors and hence reduces the drop in field. So, for a given voltage, $E_0$ will decrease as the current flowing in the space-charge region increases and flattens the field distribution.

2.9 *External Current*

The holes traversing the space-charge region give rise to a current $I_e = I_e(t)$ in the external circuit. $I_e$ is equal to the average current flow-
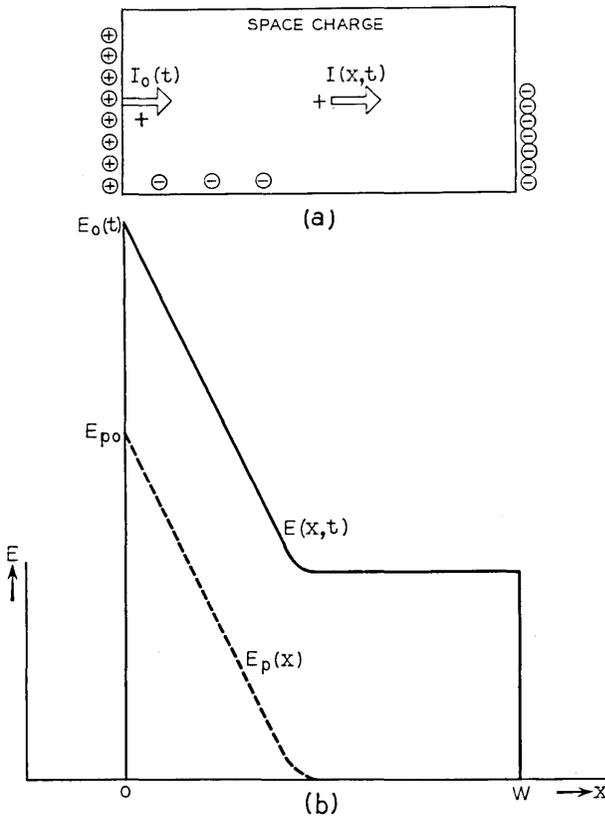
Fig. 7 — Current (a) and field (b) in the space-charge region.

ing in the space-charge region. (A rigorous and simple proof of this for
the present case of plane parallel geometry is given in Appendix A.) Thus

$$I_e(t) = \frac{1}{W} \int_0^W I(x, t) \, dx$$

$$= \frac{1}{W} \int_0^W I_0(t - x/v) \, dx \tag{18}$$

$$= \frac{1}{\tau} \int_{t-\tau}^t I_0(t') \, dt'.$$

In other words, the current $I_e$ in the external circuit is the total charge in
the space-charge region divided by the transit time $\tau$. Fig. 6 shows $I_e$ for
the case discussed at the end of Section III, where $I_0$ was a sharp pulse
in the middle of the cycle. From (18) it follows that the average value of
$I_e(t)$ is equal to the average of $I_0(t)$.

In addition to $I_e$, which arises from carriers moving through the space-
charge region, there is a capacitative current

$$I_c = C \frac{dV}{dT} \tag{19}$$

flowing in the external circuit. When the voltage $V(t)$ across the diode is
above the punch-through voltage, $V_p$, the capacity is a constant given
by (2). As discussed earlier, $I_c$ is the current required to charge and dis-
charge the diode regarded as a capacitor. It furnishes the variation in
charge at the edges of the space-charge region.

2.10 *Effect of Current on Field*

We now show how the space charge of the holes reduces the peak field
for a given voltage. The stability of the device comes from the fact that
current multiplication increases as $E_0$ increases but the current carriers
give a space charge that reduces $E_0$.

If there were no current flowing, any increase in $V$ above $V_p$ would
simply raise the entire field distribution by an amount $(V - V_p)/W$.
Fig. 8(a) is a plot of the difference $E'(x, t) = E(x, t) - E_p(x)$ at a given
time. The slope of the $E'(x, t)$ curve is determined entirely by the space
charge of the holes; the effect of the fixed charge is already included in
$E_p(x)$. The holes give a charge density $I(x, t)/v$. Thus, Poisson's equa-
tion is

$$\frac{\partial E'}{\partial x}(x, t) = \frac{4\pi}{\kappa v} I(x, t) = \frac{4\pi}{\kappa v} I_0(t - x/v). \tag{20}$$

The excess of $V$ over $V_p$ is equal to the area under the curve in Fig. 8(a). This is equal to $WE'(0, t) = W[E_0(t) - E_{p0}]$ plus the sum of the areas of a number of horizontal strips like the one shown. The area of such a strip is $(W - x) \, dE'$. So

$$V(t) - V_p = W[E_0(t) - E_{p0}] + \int_0^W (W - x) \frac{\partial E'}{\partial x} \, dx.$$

Substituting (20) into this and setting $t' = t - x/v$ gives

$$E_0(t) = E_{p0} + \frac{V(t) - V_p}{W} + \Delta E(t),$$

$$\Delta E(t) = -\frac{4\pi}{\kappa \tau} \int_{t-\tau}^t I_0(t')[\tau - t + t'] \, dt'. \tag{21}$$

Here $\Delta E(t)$ is the effect of current on the field. The quantity $\Delta E$ is always negative.
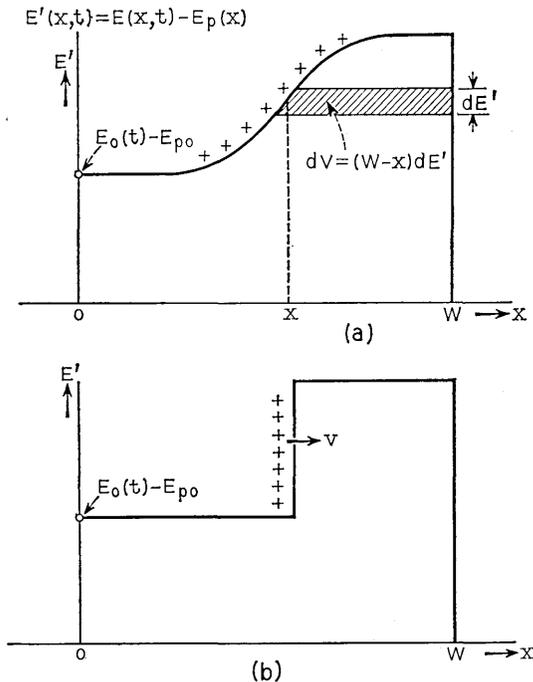


Fig. 8 — Effect of carrier space-charge on field.

2.11 *Effect of a Current Pulse*

To illustrate the physical meaning of (21) we return to the example illustrated in Fig. 6. A sharp pulse of current was generated near the middle of the cycle. The reduction $-\Delta E$ in $E_0$ due to this pulse is also shown in Fig. 6. If the pulse is instantaneous, $-\Delta E$ jumps at once to its maximum value and then declines to zero linearly in the time $\tau$ required for the pulse to cross the space-charge region. The physical reason for this is easily seen by reference to Fig. 8(b), which shows the same thing as Fig. 8(a) except that the carrier space-charge is concentrated at one point, that is, in a pulse. For a given voltage (area under the curve), the reduction in $E_0$ will decrease from its maximum value to zero as the pulse moves to the right across the space-charge region. From the effect of a single instantaneous pulse of current on $E_0$, the effect of any arbitrary current pulse can be found by resolving it into a series of instantaneous pulses and superposing the effects. An instantaneous pulse of current of charge $\delta Q$ gives an instantaneous drop in field of $4\pi\delta Q/\kappa$.

III. ANALYSIS

We now have two equations, (14) and (21), relating the current $I_0(t)$, the field $E_0(t)$ and the voltage $V(t)$. Thus in principle the current can be found for any applied voltage. Actually the exact solution is impractical except in limiting cases. In this section we present (a) an exact solution for the linear small-signal case, (b) an approximate analysis for large amplitudes and (c) a rapidly converging iteration method that yields solutions of any desired accuracy.

3.1 *Voltage*

We shall assume that the voltage varies sinusoidally, $V(t) = V_d + V_a \sin \omega t$. This will certainly be a good assumption in the small-signal range, where the diode is linear. The cavity is linear at all amplitudes. At large amplitudes of oscillation a sinusoidal voltage gives a sinusoidal capacitative current $I_c$ plus a conductive current $I_e$ which approaches a square wave as the amplitude increases. Thus we are assuming that the voltage across the cavity is sinusoidal while the current contains a distribution of higher frequencies. For a cavity like that shown in Fig. 2, this assumption may be a relatively good approximation. If the cavity is tuned to the fundamental it may be almost a short circuit for the higher frequencies.

3.2 *Dimensionless Variables*

It will simplify the discussion to express everything in terms of dimensionless variables, taking as units parameters that characterize the de-

vice. For example, $W$ would be unit length, $\tau$, unit time, and $E_c$, unit field. Then the carriers would travel with unit velocity $v = W/\tau = 1$. As can be seen in Fig. 1, the voltage will be of the order of, but less than, the unit voltage $WE_c$. It is convenient to choose the unit charge so that $4\pi/\kappa = 1$. Then unit charge produces a unit gradient of field. Since $v = 1$, a unit charge moving in the space-charge region gives unit current. Hence unit hole current in the space-charge region produces unit field gradient, and the average current $I_e$ in the space-charge region is equal to the total drop in field due to the carrier space-charge. The actual current will be small compared to unit current since the curent produces a drop in field that is smaller than the ac field variation, which in turn is small compared to $E_c$.

From the choice of units it follows that the diode has unit capacity and that optimum frequency is $\omega = \pi$. The following table summarizes the units and gives typical values for a silicon diode with $W = 10^{-3}$ cm. McKay's data,[2] plotted in Fig. 3, gives $E_c = 350$ kilovolts per cm; the effective width of the multiplication region at breakdown is taken to be $10^{-4}$ cm.

| Quantity | Unit | Example |
|---|---|---|
| length | $W$ | $10^{-3}$ cm |
| time | $\tau$ | $10^{-10}$ sec |
| field | $E_c$ | $3.5 \times 10^5$ volts/cm |
| voltage | $WE_c$ | 350 volts |
| current density | $\dfrac{\kappa E_c}{4\pi\tau} = vCE_c$ | $3.7 \times 10^3$ amps/cm$^2$ |
| power density | $\dfrac{\kappa v}{4\pi} E_c{}^2$ | $1.3 \times 10^6$ watts/cm$^2$ |

The unit of power is seen to be relatively independent of $W$, since unit voltage goes as $W$ and unit current as $1/W$. (A given current causes a greater drop in field across a wider space-charge region.)

### 3.3 Governing Equations

Since the peak field $E_0(t)$ varies around $E_c$ it is convenient to define a dimensionless field

$$E(t) = \frac{E_0(t) - E_c}{E_c}. \tag{22}$$

Then, for a sinusoidal voltage variation, equation (21) becomes in dimensionless terms

$$E(t) = E_b + V_a \sin \omega t + \Delta E,$$

$$\Delta E = -\int_{t-1}^{t} I_0(t')[1 - t + t']\, dt'. \tag{23}$$

Here $\Delta E$ is the effect of current on $E$ and $E_b$ the effect of the dc voltage $V_d$. Let $E_{p0}$ be the peak field at punch-through. Then, at punch-through, $E$ is $(E_{p0} - E_c)/E_c$. If no current flowed, any increase in dc voltage above the punch-through voltage $V_p$ would simply raise the whole field distribution by $(V_d - V_p)/W$ (see Fig. 7). Thus the dimensionless parameter $E_b$ is given in terms of dimensional quantities by

$$E_b = \frac{E_{p0} - E_c}{E_c} + \frac{V_d - V_p}{WE_c}.$$  (24)

The value of $E_b$ will be very small compared to unity.

In equation (14) we neglect $I_s$ and expand the right-hand side in terms of powers of $E = E(t)$. It will be sufficient to stop the expansion at $E^2$. Then

$$\frac{d}{dt} \ln I_0 = \frac{2(m + 1)}{\tau_1} \left( E + \frac{m}{2} E^2 \right).$$  (25)

The current through the diode is the capacitative current

$$I_c = C \, dV/dt = \omega V_a \cos \omega t$$  (26)

plus the conductive current $I_e$, where, from (18),

$$I_e(t) = \int_{t-1}^{t} I_0(t') \, dt'.$$  (27)

*Average, or dc, Values*

We shall let $\langle \ \rangle$ denote time averages and define $I_d$ as the direct current $I_d = \langle I_e + I_c \rangle = \langle I_e \rangle = \langle I_0 \rangle$ where the last step follows from (27). Averaging (23) gives

$$\langle E \rangle = E_b - \frac{I_d}{2}$$  (28)

or $\langle \Delta E \rangle = -I_d/2$. Since $I_0(t)$ is periodic, we have, from (25)

$$\langle E \rangle + \frac{m}{2} \langle E^2 \rangle = 0.$$  (29)

Thus, in the small-signal limit the average field $\langle E \rangle$ vanishes, and $I_d = 2E_b$, where $E_b$ is given by (24).

### 3.4 *Linear, Small-Signal Case*

When all quantities vary by small fractions of their average values, then the equations separate into a dc part and a linear ac part which is

easily solved. We now derive the impedance as a function of the direct current $I_d$ for the optimum frequency $\omega = \pi$. The ac voltage is $V_a \sin \pi t$ or, in complex form, $V_a e^{i\pi t}$. We write $I_0 = I_d + I_{0a} e^{i\pi t}$ where $I_{0a}$ is complex. Similar expressions can be written for $I_e$ and $E$. Equations (23), (25) and (27) give the ac relations

$$i\pi I_{0a} = \frac{2(m + 1)}{\tau_1} I_d E_a ,$$

$$I_{0a} = \frac{i\pi}{2} I_{ea}, \qquad (30)$$

$$i\pi(V_a - E_a) = I_{0a} - I_{ca} .$$

From equations (30), $V_a = Z I_{ea}$ where

$$Z = \frac{1}{2} \left[ 1 - \frac{\pi^2 \tau_1}{2(m + 1)I_d} \right] + \frac{i}{\pi} . \qquad (31)$$

The current through the diode has two parts, $I_e$ and $I_c$, where $I_c$ is a pure capacitative current. Thus the equivalent circuit consists of a unit capacity in parallel with an element of impedance $Z$, where $I_e$ goes through $Z$. The impedance $Z$ consists of a fixed reactance and a resistance that varies with the dc bias. A simpler equivalent circuit emerges from considering the admittance $Y$ of the diode. Since the capacity and the impedance $Z$ are in parallel, $Y = i\pi + 1/Z$. Here the $Q$ of the diode is the ratio of the imaginary part $Y_i$ of $Y$ to the real part $Y_r$. From (31)

$$Q = \frac{\pi}{2} \left[ 1 - \frac{\pi^2 \tau_1}{2(m + 1)I_d} \right] ,$$

$$Y_r = \frac{\pi Q}{1 + Q^2} , \qquad Y_i = \frac{\pi}{1 + Q^{-2}} . \qquad (32)$$

Thus $Q$ varies linearly with $\tau_1/I_d$ and is negative for $I_d$ less than $[\pi^2 \tau_1/2(m + 1)]$. When $I_d$ is equal to this critical value the diode becomes an open circuit for this frequency. This means that none of the alternating current generated in the multiplication region flows out of the diode. Rather it flows into the edges of the space-charge region and provides the current that charges and discharges the diode regarded as a capacitor. In other words, the unit acts as a capacitor generating its own charge internally. Hence voltage can vary with no external alternating current.

### 3.5 *Equivalent Circuit*

Equations (32) describe a simple equivalent circuit consisting of a fixed unit capacity in series with a conductance $\pi Q$, where $Q$ depends on

the dc bias. The equivalent circuit suggests the following practical conclusion: If the cavity is designed to act like an inductance in series with a variable load resistance, then the load resistance can be made equal to the negative resistance of the diode, so that the two resistances cancel, and the equivalent circuit consists of the fixed capacity of the diode and the inductance of the cavity. Hence at small amplitudes the resonant frequency will be independent of the dc bias.

### 3.6 Sharp Pulse Approximation

As the oscillation builds up, the behavior rapidly ceases to be linear, and it is impractical to solve the equations exactly. However, as we have seen, when the amplitude increases the current $I_0(t)$ approaches a sharp pulse and $I_e(t)$ approaches a square wave as shown in Fig. 6 for the optimum frequency $\omega = \pi$. The average current $I_d$ is half the maximum. In the limit of a sharp pulse the problem again becomes simple. We now derive some approximate relations for this case, and show how the oscillation can be stabilized.

The power delivered $P_r$ is

$$P_r = -\frac{V_a}{2} \int_0^2 I_e(t) \sin \pi t \, dt.$$

We can substitute (27) for $I_e(t)$ into this and reduce the double integral to a single integral by integrating by parts. The result is

$$P_r = -\frac{2V_a}{\pi} \langle I_0(t) \cos \pi t \rangle \tag{33}$$

where again the brackets denote the average over a cycle.

Thus, if the current $I_0$ is generated in a pulse near the middle of the cycle, where $\cos \pi t$ is negative, $P_r$ will be negative and power will be delivered to the ac signal.

We define the $Q$ of the diode as $2\pi$ times the ratio of the energy stored in the capacity to the energy lost in a cycle. The stored energy for unit capacity is $V_a^2/2$. The energy lost is the negative of the power delivered, $P_r$, times the period $2\pi/\omega = 2$. So

$$\frac{1}{Q} = \frac{4}{\pi^2} \frac{\langle I_0(t) \cos \pi t \rangle}{V_a}. \tag{34}$$

Let the pulse of current occur at a time $t_1$. Then in the limit of an instantaneous pulse (34) becomes

$$\frac{1}{Q} = \frac{4}{\pi^2} \frac{I_d}{V_a} \cos \pi t_1. \tag{35}$$

The phase relations will be ideal for $t_1 = 1$; that is, when the pulse occurs in the middle of the cycle. In this case $-Q$ increases with $V_a$ for constant direct current, so the oscillation is stable.*

We now consider how $t_1$ depends on $V_a$ and the dc bias. The current pulse becomes sharper as $V_a$ increases and $\tau_1$ decreases, as can be seen from (16). In the limit of vanishing $\tau_1$ the pulse becomes instantaneous and the problem can be solved exactly. This is done in Appendix B. Here we give a simple physical argument that will be a good approximation so long as the duration of the pulse is small compared to a period, as will be the case in the range of practical operation.

Fig. 9 shows $E(t) = E_b + V_a \sin \pi t + \Delta E$ for the case where the current pulse occurs at an arbitrary time $t_1$. As illustrated in Fig. 6, a
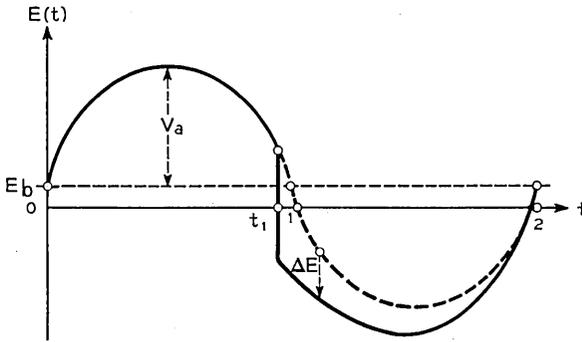


Fig. 9 — Variation of peak field with time for sharp current pulse.

current pulse causes $\Delta E$ to drop abruptly and then rise linearly to zero in a transit time. Let $E_\delta$ be the abrupt drop in $E$. We have already seen that the average $\Delta E$ is $-I_d/2$. Since $\Delta E$ is a triangular wave, lasting half a cycle, it follows that the maximum is four times the average; so $E_\delta = 2I_d$.

Let us consider in some detail the relatively short interval during which current is being generated. The holes can move only a short distance during this time. Hence, as seen from Fig. 8, the field will drop by an amount roughly proportional to the amount of charge generated. It will, therefore, have dropped by $E_\delta/2$ when half the charge has been generated. The pulse will be roughly symmetrical and will have reached its peak when half the charge has been generated. The current is a maximum at $E(t) = 0$ since it builds up for positive $E$ and decreases for negative $E$. Thus, during the first half of the pulse $E$ drops to zero, the

---

* I am indebted to J. L. Moll for pointing out the advantage of applying the bias with a constant direct current generator.

drop being $E_\delta/2$. During the second half of the pulse the field continues to drop to $-E_\delta/2$. This is shown in Fig. 9. Before the pulse occurs the field is $E(t) = E_b + \sin \pi t$. Hence in the limit of an instantaneous pulse $E_b + V_a \sin \pi t_1 = E_\delta/2$. Since $E_\delta = 2I_d$

$$I_d = E_b + V_a \sin \pi t_1 . \tag{36}$$

Eliminating $\langle E \rangle$ from equations (28) and (29) gives another relation between $I_d$ and $E_b$ :

$$I_d = 2E_b + \left(\frac{m}{2}\right) \langle E^2 \rangle .$$

Since $E_b$, $V_a$ and $-\Delta E \leq 2I_d$ are all small compared to unity, the only contribution to $\langle E^2 \rangle$ that cannot be neglected in comparison with $I_d$ and $E_b$ is $\langle V_a^2 \sin^2 \pi t \rangle = V_a^2/2$. So we have

$$I_d = 2E_b + \left(\frac{m}{2}\right) V_a^2. \tag{37}$$

Eliminating $E_b$ between (36) and (37) gives

$$\sin \pi t_1 = \frac{I_d}{2V_a} + \frac{m}{4} V_a . \tag{38}$$

When the right-hand side is larger than unity the current cannot be an instantaneous pulse. What happens then is that the current varies almost sinusoidally and produces a space charge which keeps $E$ small at all times. In this case the phase relation between the voltage and current makes $Q$ positive.

In practice, we begin with a small enough bias current so that $I_d/2V_a$ has become small before the oscillation has built up to the range where the sharp-pulse approximation is valid. The effective $Q$ of the cavity will be chosen so that the oscillation will be stabilized before $(m/4) V_a$ becomes comparable to unity. For example, suppose we take $I_d = V_a/2$ as discussed in Section I. Then $-\cos \pi t_1$ will have dropped only to 0.85, even for as large an amplitude as $V_a = 0.2$. For $V_a = 0.1$, $\cos \pi t_1$ would be $-0.92$.

### 3.7 *Constant dc Voltage Bias*

If a constant dc voltage bias $V_d$ is applied then $E_b$ is constant. As the oscillation builds up $I_d$ increases in accord with (37).

In terms of $E_b$ and $V_a$, equations (35) and (38) become

$$\frac{1}{Q} = \frac{8}{\pi^2} \left(\frac{E_b}{V_a} + \frac{m}{4} V_a\right) \cos \pi t_1 , \tag{39}$$

and

$$\sin \pi t_1 = \frac{E_b}{V_a} + \frac{m}{2} V_a. \tag{40}$$

The solid curves in Fig. 10 are plots of $-1/Q$ vs $V_a$ for two values of $E_b$. In the range shown, $\cos \pi t_1$ is approximately $-1$. The curves, therefore, have minima at $V_a = 2\sqrt{E_b/m}$. As $V_a$ increases, $-1/Q$ will eventually reach a maximum and begin to decrease because the term $(m/2)V_a$ in (40) will become important, and the pulse will occur too soon in the cycle. This range lies beyond that shown in Fig. 10.

At the amplitudes shown in Fig. 10, and especially in the stable range, the main error in the above approximations comes from the fact that the pulse is not sharp. We now turn to a method of obtaining more accurate results at low amplitudes.

### 3.8 Iteration Method

Equation (25) gives $I_0(t)$ when $E(t)$ is known and (23) gives $E(t)$ if $I_0(t)$ is known. Thus we can guess at $E(t)$, find the corresponding $I_0(t)$ from (25) and use it to determine a new $E(t)$ from (23), and so on.
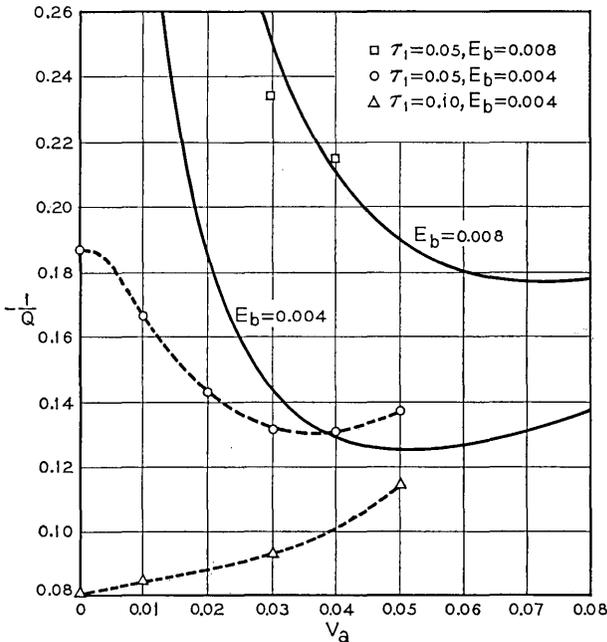


Fig. 10 — Variation of $Q$ with amplitude for several dc voltage biases.
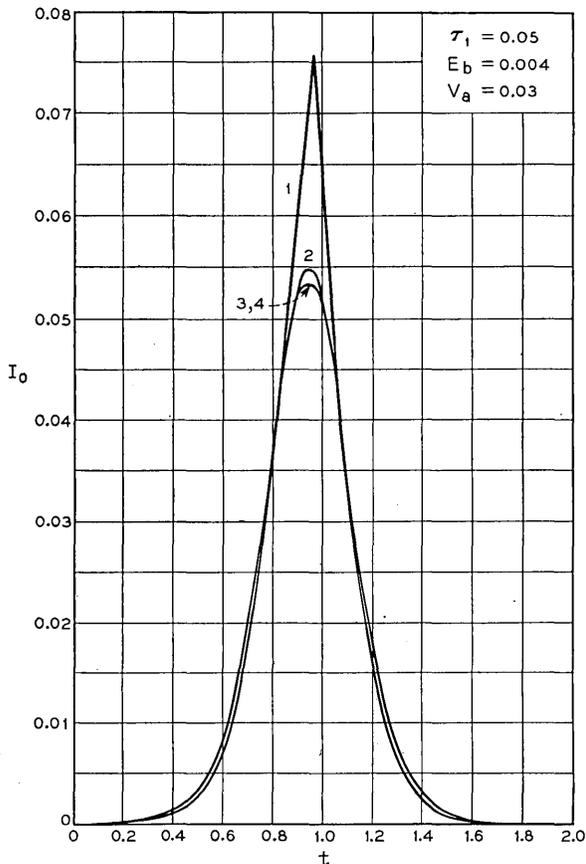
Fig. 11 — Converging solutions for $I_0(t)$.

For the initial $E(t)$ we take the instantaneous pulse solution. The procedure converges rapidly and results in solutions of any desired accuracy. In finding $I_0(t)$ from $E(t)$ in (25) a constant of integration, $I_0(0)$, is involved. What is found directly from $E(t)$ is $I_0(t)/I_0(0)$; $I_0(0)$ is so chosen that the next $E(t)$ will satisfy the condition (29) that the following $I_0(t)$ be periodic. The procedure is discussed in more detail in Appendix C.

Fig. 11 shows plots of $I_0(t)$ for $\tau_1 = 0.05$, $E_b = 0.004$ and $V_a = 0.03$. The various iterations are numbered. The values of $-1/Q$ in successive iterations were (to four places): 0.1416, 0.1330, 0.1321 and 0.1321. Fig. 12 shows $-\Delta E$ vs $t$ for $\tau_1 = 0.05$, $E_b = 0.004$ and various values of

$V_a$. The corresponding $I_0(t)$ curves are shown in Fig. 13, and $I_e(t)$ curves are shown in Fig. 14.

Fig. 10 is a plot of $-1/Q$ vs $V_a$ for various values of $E_b$ and $\tau_1$. The points were obtained from the iteration procedure and the solid curves from the sharp-pulse approximation. As expected, this approximation improves as $V_a$, $E_b$ and $1/\tau_1$ increase. With $E_b = 0.004$ there is a minimum for $\tau_1 = 0.05$ but none for $\tau_1 = 0.1$. In the former case stable oscillations would be possible at amplitudes below about 0.035.

### 3.9 *Intermediate Amplitudes*

We now have solutions for the linear small-signal range and for the large amplitude range, where the sharp-pulse approximation is good. However, there remains an intermediate range where the current variation is too large for the small-signal analysis to apply but not large enough
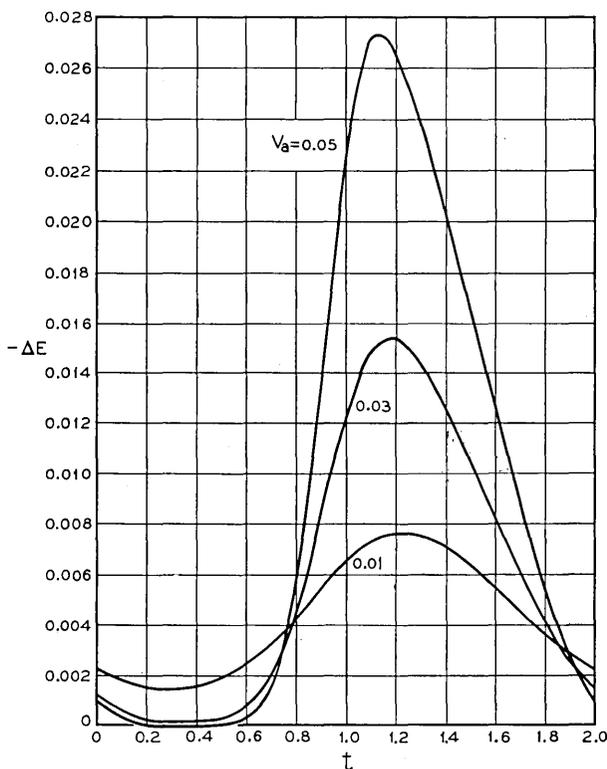


Fig. 12 — Variation of $-\Delta E$ with time for several amplitudes.
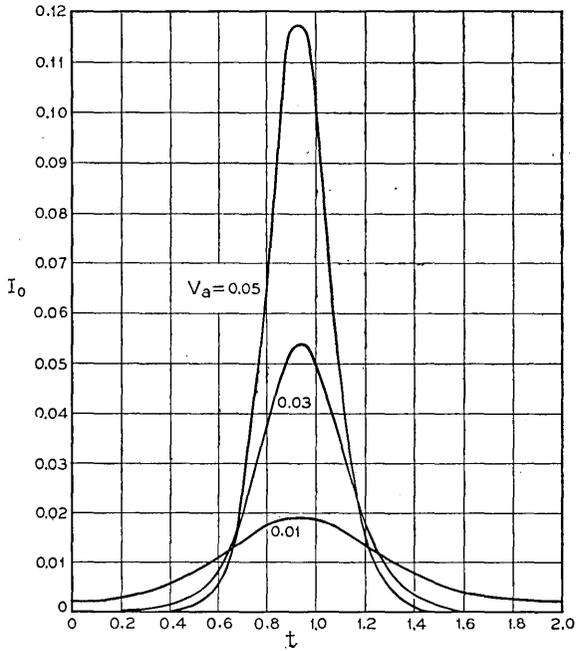
Fig. 13 — Current generated approaches a sharp pulse as amplitude increases.
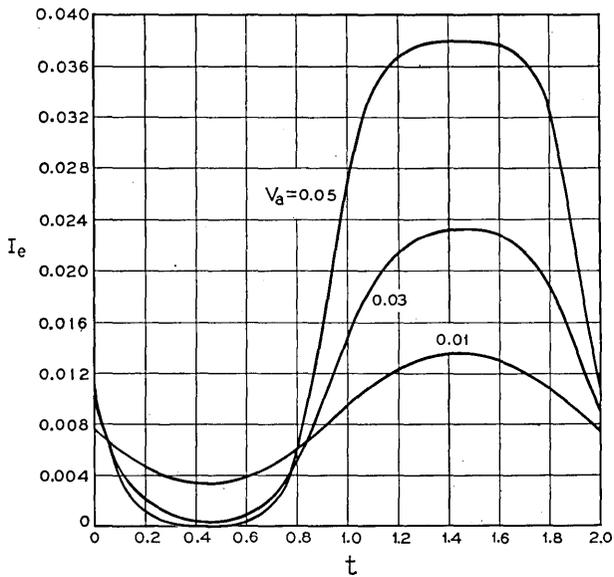


Fig. 14 — Conductive current through diode.

for the sharp-pulse approximation to be valid or even practical as a start for the iteration procedure. This range can be dealt with by an approximation that becomes more accurate as $-Q$ increases and can be within 10 per cent even for $-Q$ as low as 7. The results show how $I_0(t)$ changes from a cosine wave to an increasingly sharp pulse as $V_a$ increases.

If $Q^{-2}$ is small compared to unity then, in the small-signal case, $-Q = \pi V_a/I_{ea}$, where $I_{ea}$ is the amplitude of variation of $I_e$. The same is true at large amplitudes if $I_e(t)$ is analyzed into a Fourier series and $I_{ea}$ taken as the amplitude of the fundamental. In the linear small-signal range, $I_{ea}$ is proportional to $V_a$ so $-Q$ is constant. If $I_d$ is kept constant as the amplitude builds up, then $I_{ea}$ increases less rapidly than $V_a$. This can be seen from the fact that $V_a$ can increase without limit while $I_e(t)$ approaches a square wave of amplitude $I_d$, for which $I_{ea} = (4/\pi)I_d$. Thus $Q$ approaches $-(\pi^2/4)(V_a/I_d)$ as given by (35), for $t_1 = 1$. The phase shift, $\pi t_1$, between voltage and $I_e$ is $\cot^{-1} Q$. We have seen that $I_e$ is always 180° out of phase with the peak field. Therefore, if $-Q$ is large enough so that $I_e$ and voltage are approximately 180° out of phase, the peak field is in phase with the voltage and we can use equation (16) with $E_a = V_a$. The current $I_0(t)$ is then

$$I_0(t) = I_0(0)e^{-x(1-\cos\pi t)},$$

$$x = \frac{2(m + 1)V_a}{\pi\tau_1}. \tag{41}$$

This is seen to approach a pulse of increasing sharpness as $V_a$ increases, and to reduce to the small-signal results as $V_a$ approaches zero.

The $Q$ can be found from (35) and (41). From (35) $1/Q = (4/\pi^2)(I_d/V_a)f(x)$, where

$$f(x) = -\frac{\langle I_0(t) \cos \pi t \rangle}{\langle I_0(t) \rangle} = -\frac{d}{dx}\int_0^\pi e^{-x\cos\theta} d\theta. \tag{42}$$

The function $f(x)$ is the ratio of the first to the zero order Bessels functions of pure imaginary argument. As $x$ decreases, $f(x)$ approaches $x/2$, which means that $Q$ is constant. As $x$ increases, $f(x)$ bends over and approaches unity asymptotically. Values are $f(1) = 0.45$, $f(2) = 0.70$, $f(4) = 0.87$, $f(8) = 0.94$ and $f(16) = 0.97$. When $I_d$ is constant, $Q_0/Q = 2f(x)/x$, where $Q_0$ is the small-signal $Q$. Thus the curve of $-1/Q$ vs $V_a$ starts out flat and then decreases and approaches the form $1/V_a$ as $V_a$ increases. This is illustrated by curves 1 and 2 in Fig. 15.

If the $-1/Q$ vs $V_a$ curves for constant $I_d$ are calculated from the sharp-pulse approximation [equations (35) and (38)], they have the form of the solid curves 3 and 4 in Fig. 15. As $V_a$ decreases, the calculated $-1/Q$

goes through a maximum. For $I_d$ below about $4/m$, the maximum would occur at about $V_a = I_d/\sqrt{2}$. However, as $V_a$ decreases, the sharp-pulse approximation will break down. If $I_d$ is small enough so that the small signal $Q$ is negative, then the sharp-pulse approximation must break down before the maximum is reached, since, as we have seen, $-1/Q$ increases monotonically as $V_a$ decreases. So the curves have the form of curves 1 and 2 in Fig. 15. However, if the small-signal $Q$ is negative, then the curves have the form of curves 3 and 4 and the sharp-pulse approximation breaks down in the range of positive $Q$ (shown dotted in Fig. 15). In this range the current varies roughly sinusoidally and produces a space charge that keeps the field variation small.

IV. OPERATION

In this section we consider in more detail some of the practical questions about the design and operation of the diode. In particular, we discuss the stability for both constant current and constant voltage bias, the limitations on both dc bias and ac amplitude, the effects of heating and finally the frequency dependence of the effective admittance.

4.1 *Stability at Constant Direct Current*

Fig. 15 shows the form of the variation of $-1/Q$ with voltage amplitude $V_a$ when the bias is applied with a constant direct current generator. The oscillation will be stable at any point where $-1/Q$ is decreasing with $V_a$ and the $Q$'s of cavity and diode are equal in magnitude. The horizontal dashed line in Fig. 15 represents the effective $1/Q$ of the cavity. Thus the oscillation can be stabilized at the points A, B and C. The curves are numbered in order of increasing direct current $I_d$. If the direct current is turned on slowly compared to the response time $Q\tau$ of the diode, the oscillation will begin when $I_d$ is slightly above the value for curve 1. Thereafter the oscillation will build up and $V_a$ will increase as $I_d$ increases. By slowly varying $I_d$ we can establish a stable oscillation at any desired amplitude.

By raising $I_d$ rapidly enough we could get onto curves 3 or 4 while the amplitude was so low as to be in the unstable range. The oscillation would then start but quickly die out.

Because of the nonlinearity of the diode, the dc voltage will vary as $V_a$ increases at constant $I_d$. Eliminating $E_b$ between (24) and (37) gives

$$I_d = 2(V_d - V_p + E_{p0} - 1) + \frac{m}{2} V_a^2, \qquad (43)$$
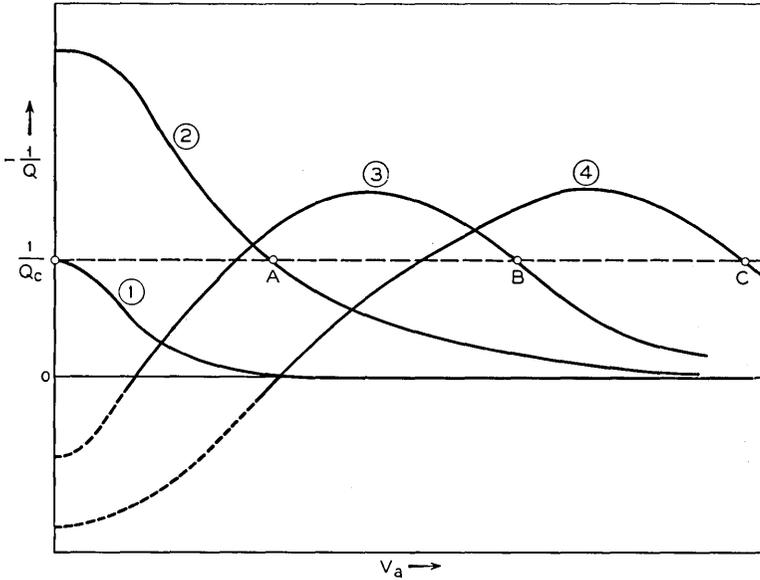
Fig. 15 — Variation of $Q$ with amplitude (schematic) for constant direct current.

where $E_{p0}$ is the peak field at punch through (taking $E_c$ as unit field) and like $V_p$ is a constant of the diode.

### 4.2 Stability at Constant dc Voltage

When the bias is applied by a constant voltage, $E_b$ will be constant as seen from equation (24). Curves for several values of $E_b$ and $\tau_1$ were shown in Fig. 10. The small-signal $Q$ is given by (32) with $I_d = 2E_b$. As $E_b$ is raised, the small-signal $Q$ will become negative and $-1/Q$ will rise. If $Q_c$, the effective $Q$ of the cavity, is small enough the oscillation will initially be stable. For example, in Fig. 10, if $\tau_1 = 0.05$ and $1/Q_c = 0.16$, the oscillation will begin when $E_b$ is raised to about 0.004. As $E_b$ is further increased, the amplitude $V_a$ will also increase. However, as seen from the figure, the oscillation will not remain stable as $E_b$ increases unless $1/Q_c$ also increases. For example, at $E_b = 0.008$ the minimum $-1/Q$ is almost 0.18. We have seen that at large enough amplitudes $-1/Q$ reaches a maximum and decreases to zero. Thus, if $E_b$ is increased with constant $Q_c$, the amplitude will suddenly jump from the range shown in Fig. 10 to a much larger value. In the range where the sharp-pulse approximation is good, $-1/Q$ is given by (37) and (40). For example, when $V_a^2$ is large com-

pared to $(4/m)E_b$, we have, approximately,

$$\frac{1}{Q} = \frac{4m}{\pi^2} V_a \sin 2\pi\tau_1,$$

$$\sin \pi t_1 = \frac{m}{2} V_a. \tag{44}$$

So there is a maximum in $-1/Q$ at $V_a = \sqrt{2}/m$. The maximum will be at lower amplitudes for larger $E_b$.

It is seen from (39) and (40) that the oscillation will always be stable when $V_a^2$ is between two and four times $E_b/m$, provided $E_b$ is large enough that the sharp-pulse approximation is good in that range. Thus, with constant dc voltage, the oscillation can be stabilized at any amplitude, but to reach the operting point it may be necessary to vary the $Q$ of the cavity as the bias voltage is increased.

In the remainder of this section we shall consider several effects, such as heating, which may cause the power and $-1/Q$ to decrease with increasing amplitude. These may limit the maximum power output (especially at low frequencies) and may also be used to stabilize the oscillation in the case of constant dc voltage.

### 4.3 *Efficiency*

We define the efficiency, $\epsilon$, as the ratio of the ac power, $P_r$, to the power $P_d$ delivered by the dc voltage or current source; therefore,

$$P_d = V_d I_d. \tag{45}$$

The power $P_h$ that goes into heat is the difference $P_h = P_d - P_r$. Therefore,

$$P_d = \frac{P_r}{\epsilon} = \frac{P_h}{1 - \epsilon}. \tag{46}$$

Later in this section we consider the temperature rise caused by $P_h$.

At the optimum frequency, $\omega = \pi$,

$$\epsilon = \frac{P_r}{P_d} = -\frac{2}{\pi} \frac{V_a}{V_d} \frac{\langle I_0(t) \cos \pi t \rangle}{I_d}. \tag{47}$$

At small and intermediate amplitudes this can be evaluated by (41) and (42), which hold when $-Q$ is well above unity. In the sharp-pulse approximation

$$\epsilon = -\frac{2}{\pi} \frac{V_a}{V_d} \cos \pi t_1. \tag{48}$$

Thus, for $V_a = V_d/2$ and $t_1 = 1$, the efficiency would be over 30 per cent.

### 4.4 *Limitations on Bias and Amplitude*

For a desired operating amplitude we want $V_d$ to be as small as possible to maximize the efficiency. However, as discussed in Section I, the minimum $V_d$ is limited by the requirement that the field in the intrinsic region must not become negative during the negative half of the voltage cycle, and the maximum $V_a$ is limited by the necessity of keeping the multiplication localized. We now consider these requirements in more detail.

In the analysis, we have assumed that the field in the intrinsic region is always high enough so that the carrier velocity remains constant, independent of field. However this will not be so at large enough amplitudes. The field may, in fact, momentarily become negative. This would reduce the negative resistance and eventually destroy it. The effect of field on velocity gives both an upper limit on the allowable amplitude for a given bias and a method of stabilizing the oscillation at any desired amplitude.

The minimum field will occur in the intrinsic region and at the trailing edge of the pulse of holes advancing to the right, as illustrated in Fig. 8. The minimum field will fall below the constant-velocity range only at high enough amplitudes so that we can use the instantaneous pulse approximation. The holes will then be closely bunched as in Fig. 8(b). The pulse causes a drop in field equal to the current $I_e$, which is $2I_d$ during the half cycle that the pulse is moving. The field $E_i$ immediately behind the pulse is, therefore,

$$E_i(t) = V_d - V_p + V_a \sin \pi t - 2I_d(1 - t + t_1). \qquad (49)$$

This has a minimum at a time $t_2$ where

$$\cos \pi t_2 = -\frac{2}{\pi}\frac{I_d}{V_a}, \qquad (50)$$

$$\pi < \omega t_2 < 3\pi/2.$$

Equation (38) gives $t_1$, and (43) gives $I_d$ in terms of $V_d - V_p$ and $V_a$. Thus for a given diode the minimum $E_i = E_i(t_2)$ is determined by the voltage bias and amplitude.

In general, the average field will be well above the range where velocity depends on field. Hence we can take $E_i(t_2) = 0$ as the critical condition where the drop in velocity in the negative half of the cycle begins to reduce the power appreciably. As $V_a$ increases, the minimum $E_i$ decreases. If the direct current is held constant, so that the oscillation is stable, then we will want $V_d - V_p$ to be large enough to build up $V_a$ to the desired amplitude without reducing $E_i$ below zero and thereby

losing power. However, $V_d$ should be made as low as possible to maximize the efficiency. Thus $E_i(t_2) = 0$ is an optimum condition for operation at constant direct current. On the other hand, if we wish to operate with constant dc voltage and in the range which would be unstable for constant velocity, then $V_d$ should be low enough so that the velocity variation will come in at the desired amplitude and stabilize the oscillation; so the condition $E_i(t_2) = 0$ would not only be optimum but necessary.

We have seen that $E_i(t_2)$ depends on $V_d - V_p$, $V_a$ and the peak field $E_{p0}$ at punch-through. Thus for a given diode $E_i(t_2) = 0$ gives a relation between $V_d - V_p$ and $V_a$. The quantity $V_a$ and the corresponding $V_d - V_p$ could then be chosen to maximize the power. However, an upper limit on the voltages is set by the necessity of localizing the multiplication. The field throughout the intrinsic region must be well below the breakdown field, especially when the current is flowing. From Fig. 8(b) the field at the leading edge of the current pulse is $V_d - V_p + V_a \sin \pi t + 2I_d(t_1 - t)$. This determines how many hole-electron pairs the pulse of holes will produce in moving across the space-charge region. The holes produced join the pulse and add to the power. However, the electrons moving in the opposite direction will disturb the phase relations.

In the arguments above $V_d$ occurs only in the combination $V_d - V_p$. It is therefore desirable to make $V_p$ as small as possible so that $V_d$ is small and $\epsilon$ large.

Both the voltage $V_p$ and the peak field $E_{p0}$ at punch-through are determined by the impurity distribution in the p-region. The amplitude at which the diode is designed to oscillate will fix the choice of $E_{p0}$. The larger the desired amplitude, the smaller $E_{p0}$. However, if $E_{p0}$ is too small the multiplication will not be localized.

### 4.5 *Example*

We may illustrate the above discussion by applying it to an actual design. In the discussion at the end of Section I we took $V_d = 2V_a = 4I_d$ as a reasonable operating condition. This would give an efficiency of about 30 per cent. To operate at an amplitude of $V_a = 0.2$ would then require that $V_d - V_p$ be no less than 0.275. For $W = 10^{-3}$ and a unit voltage of 350 volts this would mean $V_a = 70$ volts, $V_d = 140$ volts and, for optimum operation, $V_p = 44$ volts. These parameters are reasonable. In fact, both $V_d$ and $V_p$ could probably be slightly lower so that the efficiency would be higher.

### 4.6 *Heating*

The power $P_h$ that goes into heat will cause the temperature of the diode to rise above that of the surroundings. We shall assume that the surface of the cavity is kept at constant temperature. If the radius $R$ of the diode is small compared to the radius of the center post of the cavity, as in Fig. 2, then heat can flow away from the diode in almost all directions. The diode can be made thin enough and mounted sufficiently close to the metal so that the temperature drop in the silicon is small compared to that in the metal. Let $\Delta T$ be the difference between the diode temperature and the temperature of the surface of the cavity. Then $\Delta T$ is related to $P_h$ and $R$ by the formula for spreading resistance

$$\Delta T = \frac{P_h R}{4K},\tag{51}$$

where $K$ is the thermal conductivity of the metal. For copper, $4K = 16.7$ watts per cm per °C.

The temperature will rise to its equilibrium value in a time of about $R^2/D$, where $D$ is the coefficient of thermal diffusion, which is about unity for copper.

We may now apply these results to the examples discussed at the end of Section I assuming an efficiency of $\frac{1}{3}$ so that $P_h = 2P_r$. The 50 watts of power output at 5,000 megacycles and $R = 0.03$ cm would produce a temperature rise of about 60°C. At 500 megacycles and $R = 0.3$, however, the 5 kw maximum power would raise the temperature by about 600°C. Thus at low frequencies the maximum power output in continuous operation would be limited by heating rather than by how small the impedance $\omega L$ of the cavity can be made. However, the time constant of the temperature rise for $R = 0.3$ cm would be almost a tenth of a second, so the temperature rise in pulse operation would not be serious.

### 4.7 *Effect of Temperature on Critical Field*

McKay[2] has found that the critical field increases with temperature. For critical fields between 250 and 500 kilovolts per cm in silicon, a change in temperature changes the critical field by 0.05 per cent per °C. Increasing the critical field effectively decreases the dc voltage and current. Thus when the diode begins to get hot it effectively reduces its bias. The heating will therefore stabilize the oscillation.

### 4.8 Effect of Reverse Saturation Current

In silicon the reverse saturation current will become important only at large amplitudes. In the small-signal range it is easily shown that $I_s$ can be neglected if it is small compared to $\pi[\pi\tau_1/2(m+1)]^2$ which will be of the order of $10^{-3}$, or a few amperes per cm$^2$ for $W = 10^{-3}$ cm.

Even at large amplitudes $I_s$ will be small compared to the average current and will have a negligible effect on the space charge. However, as $V_a$ increases, the ratio of the maximum and minimum values of $I_0$ (which varies exponentially with $V_a$) will increase much faster than the maximum $I_0$. Thus the minimum $I_0$ becomes very small at high amplitudes. It cannot, however, fall below $I_s$.

If the effect of $I_s$ is included, the equation (25) for current generation becomes

$$\frac{\tau_1}{2}\frac{d}{dt}\ln I_0 = (m+1)\left(E + \frac{m}{2}E^2\right) + \frac{I_s}{I_0}. \tag{52}$$

As $V_a$ increases and the minimum $I_0$ decreases, the term $I_s/I_0$ becomes important near the current minimum, and prevents the current from becoming too low. Thus $I_s$ will be important in the equation only near the current minimum. However, by increasing the minimum $I_0$, $I_s$ also increases subsequent values of $I_0$ and hence increases the associated carrier space-charge. The carrier space-charge will therefore shut the current off (by reducing the field) earlier in the cycle. This will reduce the delay between voltage and current and so reduce the power.

The amplitude at which $I_s$ becomes important can be roughly estimated as follows: From (41) the ratio of maximum and minimum values of $I_0$ is related to $V_a$ by

$$V_a = \frac{\pi\tau_1}{4(m+1)}\ln\frac{I_{\max}}{I_{\min}}.$$

A formula for $I_{\max}$ is derived from the instantaneous-pulse approximation in Appendix B [equation (63)] 11. In practical cases, where $I_d$ will be of the order of $10^{-1}$, $I_{\max}$ is seen to be of the order of unity. At room temperature the reverse saturation current in silicon can be made less than a microampere per cm$^2$, which is a dimensionless current of around $10^{-10}$ in the kmc range. Thus the minimum $I_0$ will be greater than $I_s$ if $V_a$ is no larger than about 2.6 $\tau_1$. To operate at an amplitude $V_a = 0.2$ we therefore want $\tau_1$ to be no smaller than 0.07.

The effect of $I_s$ is treated in detail in Appendix D and more general relations are found to replace (37) and the equations derived from it.

### 4.9 *Frequency*

So far we have considered only the optimum frequency $\omega = \pi/\tau = \pi$. The small-signal analysis is easily carried out for arbitrary frequency. The admittance $Y$ of the diode is found to be

$$Y(\omega) = \frac{4\omega y[(1 - \cos\omega) + i(2y - \sin\omega)]}{2y^2 - 2y\sin\omega + 2(1 - \cos\omega)},$$

$$y = \frac{\omega}{2}\left[1 - \frac{\omega^2\tau_1^2}{2(m + 1)I_d}\right].$$

(53)

This reduces to (32) for $\omega = \pi$. When $-y$ is well above unity, it can be held constant by varying $I_d$ in proportion to $\omega^3$. Then $-1/Q$ will vary only with the phase factor $1 - \cos\omega$. Thus by varying $I_d$ the device can be tuned mechanically over a frequency range extending from $\frac{1}{2}$ to $\frac{3}{2}$ of the optimum frequency.

At large amplitudes the current no longer varies sinusoidally. However, as mentioned in Section III, the cavity may be almost a short circuit for frequencies higher than the fundamental. Hence, to evaluate the diode as an element in the oscillator, we can analyze the current $I_e + I_c$ into a Fourier series and retain only the fundamental. The relation of this to the ac voltage then defines the conductance, $G$, and capacity, $C$, of the diode. The results are

$$G = \frac{2}{\omega V_a}\left[\langle I_0(t)\sin\omega t\rangle\sin\omega + \langle I_0(t)\cos\omega t\rangle(1 - \cos\omega)\right],$$

(54)

$$C = 1 + \frac{2}{\omega^2 V_a}\left[\langle I_0(t)\sin\omega t\rangle(\cos\omega - 1) + \langle I_0(t)\cos\omega t\rangle\sin\omega\right]$$

where again the brackets denote time averages. When current and voltage are approximately 180° out of phase, we can use (42) to evaluate the time averages in (54).

In the sharp-pulse approximation the averages become $I_d\sin\omega t_1$ and $I_d\cos\omega t_1$ respectively. Equation (37), which is independent of frequency, remains valid. Since the drop $E_\delta$ in field is $(4\pi/\omega)$ times the average $\langle -\Delta E\rangle = I_d/2$, the general form of (36) becomes

$$\frac{\pi}{\omega}I_d = E_b + V_a\sin\omega t_1.$$

(55)

Above the small-signal range the admittance $Y = G + i\omega C$ of the diode depends on voltage amplitude as well as on frequency and direct current. For stable oscillation the admittance $Y = Y(\omega, I_d, V_a)$ of the diode is equal to $-Y_c$, where $Y_c$ is the admittance of the cavity. For a

given setting of the plunger (Fig. 2), $Y_c = Y_c(\omega)$ depends only on the frequency. Thus the frequency and amplitude of oscillation are determined in terms of the direct current by $Y(\omega, I_d, V_a) + Y_c(\omega) = 0$. When $I_d$ is varied both $V_a$ and $\omega$ will vary. However, it is not certain whether the variation of frequency with $I_d$ could be made large enough for a practical frequency modulation device. Near the optimum frequency and bias the frequency will remain relatively constant as $I_d$ and $V_a$ vary. Thus the frequency is primarily determined mechanically by adjusting the height of the cavity (Fig. 2) and the amplitude of oscillation is determined electrically by the bias.

### 4.10 *Zener Current*

In sufficiently narrow junctions, where breakdown occurs, the rate of generation is an extremely sensitive function of field. So, as in secondary emission, the generation can be highly localized. The diode could be made to operate by Zener current rather than multiplication if the $p$ region is sufficiently narrow. However, the conditions for negative resistance would be less favorable. The current $I_0$ generated by field emission is a function of the peak field $E_0$. Hence $I_0$ and $E_0$ are in phase, so all of the delay has to come from the transit time. For ideal phase relations at large amplitudes the bias should be such that the current is generated mainly in a short burst near the voltage peak. Then, if the transit time is $\frac{3}{4}$ of a cycle, the current, $I_e$, will flow during the last three quarters of the cycle and power will be delivered to the ac signal. However, the $Q$ and efficiency are considerably lower than for secondary emission. The small-signal $Q$ can be varied by varying the bias. A small change in bias will change $dI_0/dE_0$ drastically (since we are on the knee of the current-voltage curve). For the frequency $0.75/\tau$, the minimum $-Q$ is 25 at small signals and over 100 at large signals. The latter might be improved somewhat by increasing the frequency in relation to $1/\tau$. The efficiency could probably not be raised above about 5 per cent without ruining the $Q$.

An analysis of the diode operation on Zener current is discussed in Appendix E.

### V. ACKNOWLEDGEMENTS

APPENDIX A

*Derivation of $I_e$*

We first give a simple analytical derivation and then a physical argument. A more general proof is given by Shockley (1938).

Subtract (4) from (3) and use Poisson's equation

$$\frac{\partial E}{\partial x} = \frac{4\pi q}{\kappa}(p - n + N)$$

for $p - n$. The result is*

$$\frac{\partial}{\partial x}\left[\frac{\kappa}{4\pi}\frac{\partial E}{\partial t} + I\right] = 0.$$

In other words the quantity in brackets is a function of time only and is the same, at a given time, throughout the length of the diode — not only in the space-charge region, but also in the ends. The current in the ends, or leads, is $I_c + I_e$. Let the ends be of sufficiently low resistance so that the field can be assumed to vanish there. Then

$$\frac{\kappa}{4\pi}\frac{\partial E}{\partial t} + I = I_c + I_e.$$

Averaging over the length $W$ of the space-charge region and using $C = \kappa/4\pi W$ gives

$$C\frac{dV}{dt} + \frac{1}{W}\int_0^W I\, dx = I_c + I_e.$$

The physical argument can be illustrated by Fig. 8(b). Let a small pulse of charge $\delta Q$ be generated at $x = 0$. It causes a drop in field $\delta E = (4\pi/\kappa)\delta Q$. If the field at the edges remains constant as the pulse moves, the voltage will drop at a rate $v\delta E = (4\pi/\kappa)v\delta Q$. Therefore, if the voltage is to remain constant, a current $\delta I_e$ must flow in the

---

* G. Weinreich (private communication) has pointed out that this is a special case of the general three-dimensional result

$$\text{div}\left[\frac{\kappa}{4\pi}\frac{\partial \vec{E}}{\partial t} + \vec{I}\right] = 0$$

which follows from

$$\text{curl } \vec{H} = \frac{\kappa}{c}\frac{\partial \vec{E}}{\partial t} + \frac{4\pi}{c}\vec{I}$$

and div curl = 0.

external circuit and increase the voltage by adding charges to the edges of the space-charge region. The rate of voltage increase will be $\delta I_e/C$. Setting this equal to $v\delta E$ gives $\delta I_e = \delta Q/\tau = v\delta Q/W$.

APPENDIX B

*Instantaneous Pulse*

The equations can be solved exactly for any amplitude and bias in the limiting case of $\tau_1 = 0$. As we have seen, the current $I_0$ approaches an instantaneous pulse. This solution will give reasonable approximations to the $Q$, efficiency and power for actual cases if the effective duration of the current pulse is a small fraction of a cycle. We let the frequency $\omega$ be arbitrary but less than $2\pi/\omega$, so that only one current pulse is flowing at a time. The solution is completely specified by the time $t_1$ at which the pulse and the discontinuity in field occur and the magnitude $E_\delta$ of the discontinuity. From $\langle \Delta E \rangle = -I_d/2$ and the form of $\Delta E(t)$,

$$E_\delta = \frac{4\pi}{\omega} \langle -\Delta E \rangle = \frac{2\pi}{\omega} I_d. \tag{56}$$

The condition that $I_d$ be periodic is

$$\langle E \rangle + \frac{m}{2} \langle E^2 \rangle = 0 \tag{57}$$

where $E = E_b + V_a \sin \omega t - E_\delta(t - t_1)$ in the internal from $t_1$ to $t_1 + 1$ and $E = E_b + V_a \sin \omega t$ at all other times. Thus (57) is a quadratic equation for $E_\delta$ as a function of $E_b$, $V_a$ and $t_1$. Only the smaller of the two roots is meaningful. When $E_b$ and $V_a$ are small compared to unity (57) reduces to (37) and is independent of $t_1$.

To obtain another relation between $t_1$ and $E_\delta$ we solve (23) and (25) and find the relation between $I_0$ and $E$ during the pulse. In the limit of an instantaneous pulse, $t$ in equation (23) does not vary during the pulse, so $dE = -I_0 \, dt$. We can use this to eliminate $dt$ in (25). Then integrating we have the relation

$$I_{\max} - I_0 = \frac{(m + 1)}{\tau_1} E^2 \left[ 1 + \frac{m}{3} E \right] \tag{58}$$

for $I_0(t)$ as a function of $E(t)$ during the pulse. The values of $E$ at the beginning and end of the pulse are found by setting $I_0 = 0$ in (58). Let $E_1$ be the field at the beginning of the pulse. Then

$$E_1 = E_b + V_a \sin \omega t_1, \tag{59}$$

and the maximum $I_0$ is

$$I_{\max} = \frac{(m+1)E_1^2}{\tau_1}\left(1 + \frac{m}{3}E_1\right). \tag{60}$$

Neglecting terms in $E_1^2$ in comparison with unity, (58) gives

$$E_\delta = 2E_1\left(1 + \frac{mE_1}{6}\right). \tag{61}$$

Combining this with (57) and (60) we have

$$E_b + V_a \sin \omega t_1 = \frac{\pi}{\omega}I_d\left(1 - \frac{m\pi}{6\omega}I_d\right). \tag{62}$$

In practical cases, the term $(m\pi/6\omega)I_d \approx I_d$ can be neglected in comparison with unity. Then (63) becomes (55). This, together with the quadratic (57), determines $E_\delta$ and $t_1$ for any $E_b$ and $V_a$.

From (60) and (61),

$$I_{\max} = \frac{(m+1)E_\delta^2}{4\tau_1} = \frac{\pi^2(m+1)}{\omega^2\tau_1}I_d^2. \tag{63}$$

Thus the effective duration of the pulse $\Delta t = I_d/I_{\max}$ is

$$\Delta t = \frac{\omega^2\tau_1}{\pi^2(m+1)I_d}. \tag{64}$$

So the pulse becomes sharper as $\tau_1$ decreases and $I_d$ increases. We can estimate the accuracy of the instantaneous pulse approximation by comparing $\Delta t$ with the period $2\pi/\omega$.

APPENDIX C

*Iteration Method*

The procedure in detail is the following: Each iteration goes from an $E(t)$ through an $I_0(t)$ to a new $E(t)$. In the first iteration we have to guess at $E(t)$. The procedure has been to begin with the $E(t)$ corresponding to the instantaneous current pulse and illustrated in Figs. 6 and 9. The magnitude $E_\delta$ of the discontinuity is chosen to satisfy

$$\langle E + (m/2)E^2 \rangle = 0.$$

The time $t_1$ when the discontinuity occurs is found from (40). Putting this $E(t)$ into (25) and integrating gives a periodic $I_0(t)/I_0(0)$ from which, using (23), we find a function $F(t) = -\Delta E(t)/I_0(0)$. Putting $E = E_b + V_a \sin \pi t - I_0(0)F(t)$ into $\langle E + (m/2)E^2 \rangle = 0$ gives a quadratic equation

for $I_0(0)$. Only the smaller of the two roots is meaningful. From the known $F(t)$ and $I_0(0)$ we have a new $E(t)$. Since it satisfies $\langle E + (m/2)E^2 \rangle = 0$, the $I_0(t)/I_0(0)$ calculated from it in the following iteration will be periodic.

To plot $I_0(t)$ in Fig. 11 we have determined both $I_0(t)/I_0(0)$ and $I_0(0)$ from the same $E(t)$. The quantity $I_0(t)/I_0(0)$ is found from (25) and $I_0(0)$ is chosen to satisfy $\langle I_0 \rangle = -2\langle \Delta E \rangle$. This relation is automatically satisfied if $\Delta E(t)$ has been determined from $I_0(t)$ using (23). However, only in the exact solution is it satisfied if $I_0(t)$ was determined from $\Delta E(t)$. To determine $I_0(0)$ from $\Delta E(t)$ we have

$$I_0(0) = \frac{-2\langle \Delta E(t) \rangle}{\left\langle \dfrac{I_0(t)}{I_0(0)} \right\rangle}$$

where $I_0(t)/I_0(0)$ is found from $\Delta E(t)$ using (25) with $E = E_b + V_a \sin \pi t + \Delta E$. The value of $I_0(0)$ found in this way, in each iteration, was compared with the value that makes the following $I_0(t)$ periodic. When the two values agreed within a specified amount, usually taken to be about one per cent, the iteration procedure was terminated. The procedure was programmed on an I.B.M. 650 Magnetic Drum Calculator. The machine would give a solution of the required accuracy in about four iterations, or forty minutes, on the average.

In one case, where the current peak occurred at $t_1 = 0.92$, the iteration procedure was repeated starting with $t_1 = 0.85$. The peak in successive iterations moved from 0.85 to 0.92. In other words, the final result was independent of the initial $t_1$. However, more iterations were required for a poor initial choice.

APPENDIX D

*Effect of $I_s$*

It is convenient to consider the solution for a single cycle extending between two current minima. The effect of $I_s$ will be important only near the ends of the cycle where $I_0$ is small and the term $I_s/I_0$ cannot be neglected in (52). Let the solution be $f(t)I_0(t)$ where $I_0(t)$ is a solution for $I_s = 0$ and is the correct solution during most of the cycle. Then $f(t)$ will be unity during most of the cycle and will rise sharply at the beginning and end of the cycle. Since the current is always positive, each rise in $f(t)$ must be less than unity. The condition that the current be periodic is $\Delta \ln f + \Delta \ln I_0 = 0$, where $\Delta$ denotes the change during a cycle. The change in $\ln I_0$ is found by the same procedure that led to

(37). If $\Delta f$ is the total change in $f$ in a cycle we have, instead of (37),

$$\ln\left(\frac{2 + \Delta f}{2 - \Delta f}\right) = \frac{4\pi(m + 1)}{\omega\tau_1}\left[E_b - \frac{I_d}{2} + \frac{m}{4}V_a{}^2\right].\qquad(65)$$

Replacing $I_0(t)$ by the correct solution $f(t)I_0(t)$ in (52) gives

$$\frac{df}{dt} = \frac{2}{\tau_1}\frac{I_s}{I_0}.\qquad(66)$$

To determine $\Delta f$ we need to solve this only in the short interval near the current minima where $f$ is changing. In practical cases the current will have stopped flowing slightly before the end of the negative half of the cycle. Hence when $f$ is changing, $E + (m/2)E^2$ will vary as $V_a\omega t$ and

$$\ln\left[\frac{I_0(t)}{I_{\min}}\right] = \frac{(m + 1)V_a\omega t^2}{\tau_1}\qquad(67)$$

where $I_{\min}$ is the minimum of $I_0(t)$ and $t = 0$ is taken at the minimum. Substituting (67) into (66) and integrating over the short interval where $f$ is changing gives

$$\Delta f = \frac{2I_s}{I_{\min}}\sqrt{\frac{\pi}{(m + 1)V_a\omega\tau_1}}.\qquad(68)$$

The ratio of $I_{\min}$ to $I_{\max}$ is found by integrating (25) from the current minima at $E = 0$ to the maximum at $t = t_1$. At the amplitudes where the effect of $I_s$ is important we can use the sharp-pulse approximation, so $E = E_b + V_a\sin\omega t$ during this interval. The quantity $I_{\max}$ is given by (63). Thus $\Delta f$ is found in terms of $E_b$, $V_a$ and $t_1$. Equation (65) now replaces (37) and together with (55) determines $t_1$ and $I_d$ for a given voltage bias and amplitude.

APPENDIX E

*Zener Current*

The small-signal case is easily solved. The ac variations in $I_0$ and $E_0$ are proportional. At large signals it is probably a good approximation to say that (a) no current is generated until $E_0$ has risen to a critical value, which we shall define as $E_c$, and (b) thereafter $I_0$ will be such as to keep $E_0$ from rising above $E_c$. Thus (25) is replaced by $I_0 = 0$ for $E < 0$ and $E = 0$ for $I_0 > 0$. Equations (23) and (26) remain unchanged. The current $I_e(t)$ will flow during the following three intervals:

(A) Beginning at the time $t_0$ when $E_0 = E_c$, current begins to be generated and continues to be generated until a time $t_1$ when no more carrier space-charge is required to keep $E_0$ from rising above $E_c$, that

is, to keep $E$ from becoming positive. The last holes to be generated will have been generated before the first have crossed the space-charge region. During this interval, $I_e$ increases from $I_e(t_0) = 0$ to its maximum $I_e(t_1)$, while $I_0$ jumps at once to its maximum at $t_0$ and thereafter declines, reaching zero at $t_1$. The governing equation is

$$\tau \frac{dI_e}{dt} - I_e = C \frac{dV}{dt} = CV_a \omega \cos \omega t. \tag{69}$$

This can be obtained by physical reasoning from Fig. 8 or by differentiating (23) and (26). Equation (69) is easily solved, subject to $I_e(t_0) = 0$. The time $t_1$ and the maximum current $I_e(t_1)$ are found from $dI_e/dt = 0$ at $t = t_1$.

(B) From the time, $t_1$, when $I_e$ reaches its maximum and $I_0$ drops to zero, until the time $t_0 + \tau$, when the first holes to be generated reach the i-p$^+$ junction, $I_e$ remains constant and equal to $I_e(t_1)$. A constant number of holes are moving with constant velocity across the space-charge region.

(C) Between $t_0 + \tau$ and $t_1 + \tau$ the holes generated in the first interval are flowing out at the i-p$^+$ junction. Since the first holes to be generated are the first to flow out, $I_e(t) = I_e(t_1) - I_e(t - \tau)$ during this interval.

Thus, by solving (69) for $I_e(t)$ in the interval $t_0 \leq t \leq t_1$ we know $I_e(t)$ throughout the cycle. Except in the first interval, $E \leq 0$ and $I_0 = 0$.

The equations have been solved for the frequency $0.75/\tau$ or $\omega = 3\pi/2\tau$. The $Q$ is a function of $\omega t_0$, which is determined by the voltage bias and amplitude. The $Q$ is negative for $13° < \omega t_0 < 90°$. The maximum $-1/Q$ is about 0.0095 and occurs at about $\omega t_0 = 40°$. The quantity $[\epsilon/(1 + \epsilon)]$ $(V_d/V_a)$ rises almost linearly with $\omega t_0$ from zero at $13°$ to 0.21 at $90°$. The limitations on $V_d$ and $V_a$ would be about the same as for a diode operating by multiplication.

REFERENCES

1. A. G. Chynoweth and K. G. McKay, Phys. Rev., **106**, (1957), 418.
2. K. G. McKay, Phys. Rev., **94**, (1954), 877.
3. P. Wolff, Phys. Rev., **95**, (1954), 1415.
4. W. Shockley, J. Appl. Phys., **9**, (1938), 635

# Broadband Oscilloscope Tube

By D. J. BRANGACCIO, A. F. DIETRICH and J. W. SULLIVAN

*By applying traveling wave tube principles to the design of a helix type vertical deflection system it has become possible to build an oscilloscope tube whose bandwidth characteristic is flat over 600 megacycles. The trace on the fluorescent screen is readable without other optical means. The tube construction is similar to commercial cathode ray tubes, the only exception being that the alignment of the various tube elements has a closer tolerance. In actual use this tube allows one to view directly repetitive pulses a few millimicroseconds in width.*

## I. INTRODUCTION

New transmission systems which employ binary pulses and regeneration[1] require pulses which are a few millimicroseconds in length. One of the primary characteristics of short pulses is the wide bandwidths required for any system which is to handle them. The advantages of using a wideband system have been discussed by others[2,3] and in this paper we will describe an oscilloscope which is capable of presenting a visual display for these millimicrosecond pulses.

In a conventional oscilloscope which uses a pair of plates to deflect the electron beam the frequency range is limited to the extent that a typical tube might have a response which is down by 3 db at 60 mc. This bandwidth is limited by the transit time of the electrons through the deflection system. This time interval should be short as compared to the period of the RF voltage on the plates. We can decrease the time with a resulting increase in bandwidth by shortening the axial length of the plates and increasing the speed of the electrons. Both of these factors, however, tend to decrease the sensitivity which we define as the deflection of the beam measured in units of spot diameters for a unit input voltage. A much improved method of overcoming the transit-time problem has been suggested by Pierce[4] where he uses a slow traveling RF wave to deflect the beam. The deflecting voltage travels at the same speed as the electrons and consequently a given electron sees

447

the same phase of the wave throughout the deflection system. The bandwidth in such a system is now limited by the frequency range over which the RF structure is capable of propagating a wave at a constant velocity. This extends the frequency limit well into the microwave region. This principle has recently been used by Germeshausen[5] in a tube which displays single pulses or transient responses of very short duration.

The oscilloscope discussed here, shown in Figs. 1 and 2, has a flat
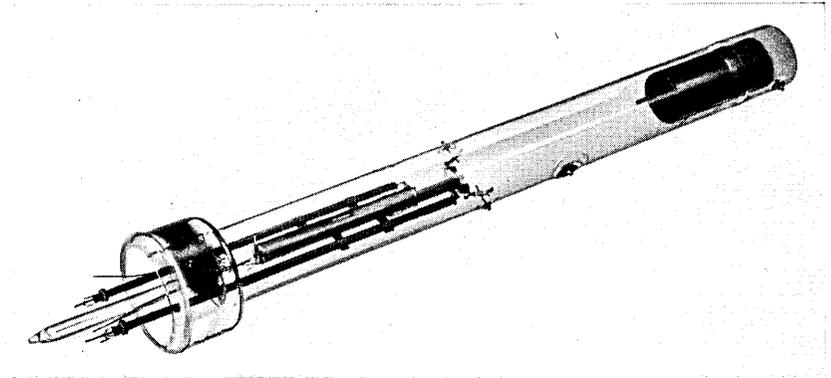


Fig. 1 — Photograph of broadband oscilloscope tube — Model III.
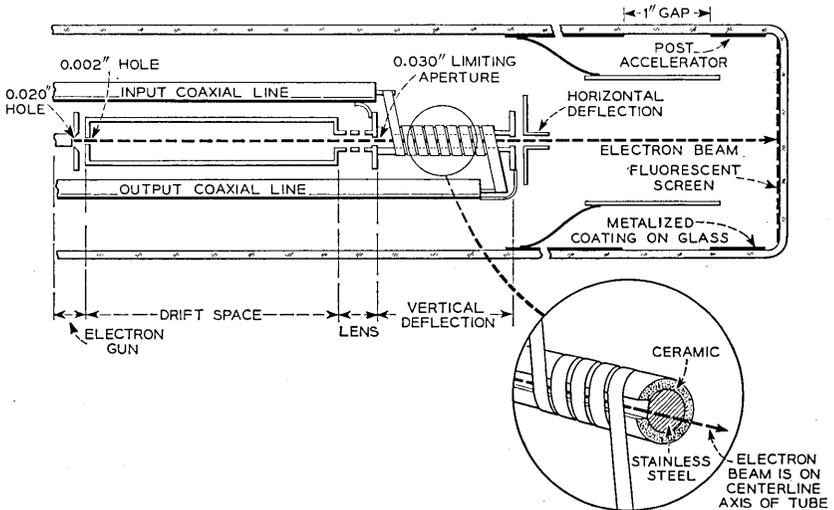


Fig. 2 — Cross section of broadband oscilloscope tube — Model III.

response from 0 to 600 mc and is useful for viewing repetitive pulses within this bandwidth. This limit is determined by the particular transition from the 76-ohm coaxial input line to the slow wave structure and is not inherent in the traveling wave deflection system. The viewing field is approximately one inch square when no post acceleration is applied. This increase in picture size is obtained by increasing the tube length and therefore the optical system of Pierce's previous model has been eliminated. With a beam voltage of 1,000 volts the sensitivity is 0.02 trace widths per millivolt which tells us that 0.41 milliwatts of power into the 76-ohm input lead is required for a peak-to-peak deflection of 10 trace widths.

## II. GENERAL DESCRIPTION OF THE TUBE

The broadband oscilloscope tube of Fig. 1 consists of an electron gun which forms a small beam (0.002″ diameter) that is focused on the screen by the lens which precedes the slow wave structure as illustrated in Fig. 2. The electron gun employs a type "B" Philips impregnated cathode 0.025″ in diameter and it is capable of giving current densities in excess of one ampere per square centimeter. The gun design is simple in that we aperture the beam with a small anode hole rather than making use of a sharp crossover. The ion bombardment of the cathode was considerably reduced by aperturing the beam with a 0.020″ hole in a low voltage electrode near the cathode surface. This electrode also serves as a current control which is very useful for blanking the return trace. The 0.030″ aperture at the input of the deflection system eliminates those electrons which do not pass near the center of the focusing lens and insures that the vertical dimension of the beam in the traveling wave deflection will be small. This system of deflection is the component which is new in this tube and it will be discussed in detail in Section III. The horizontal sweep for this tube is a pure sine wave of a frequency much lower than the repetition rate of the incoming pulses. Since a single frequency is involved, the horizontal sweep circuit can be resonant. Therefore, large sweep voltages are easy to obtain. Two parallel plates $\frac{1}{4}$″ long spaced 0.060″ apart are used for the horizontal deflection system. The picture size increases with distance beyond the deflection circuit and as a compromise between over-all tube length and picture size, a drift distance of 8″ was chosen. There is a simple post acceleration lens near the screen for the purpose of increasing the intensity. Since it is a converging lens it has the disadvantage of decreasing the picture size, as a result the accelerating voltage is limited to about 3,500 volts.

The use of a strong post-acceleration lens wherein the beam actually

crosses the axis before reaching the screen has been suggested by a number of people as a method of avoiding the reduction in picture size with increasing screen voltage. Such a system has been built (described in Section VI) which operated at a post-accelerating voltage of 10 kv. However, the system does suffer from increased astigmatism and more work is required before it can be used.

### III. VERTICAL DEFLECTION

The vertical deflection system is a coaxial tape helix made of gold-plated molybdenum tape 0.115″ wide wound at eight turns per inch over a stainless steel arbor as shown in the enlarged sketch of Fig. 2. The rectangular opening for the beam is 0.044″ high and 0.108″ wide. At either end there is a simple transition from the 76-ohm coaxial line to the helix as shown. We see that the coaxial line is essentially turned inside out with the outer coaxial conductor becoming the inner portion of the deflecting system. This transition from coaxial line to the deflecting system represents a discontinuity which is mainly due to a radical change in geometry in a short distance. This is true even though impedance is preserved. At about 650 mc the input and output transitions are spaced about a half wavelength apart and, consequently, a large reflection occurs. Actual measurements on the tube show that the reflections from this tube are 25 db down from 0 to 600 mc (vswr < 1.13). The bandwidth could be increased by an improved coaxial to circuit transition. The high frequency limit to this helix as a propagating structure will occur when the circumference becomes equal to one-half wavelength. At this point the discontinuities due to the slot in the ceramic spacer are one-half wavelength apart and a stop band in the propagation characteristic will appear. This will occur at about 3,000 mc for the dimensions of this tube.

The maximum angle of deflection is limited by the ratio of the length to vertical separation of tape helix and the center conductor. This angle is 3.7°.

When the beam voltage is properly adjusted, the wave and the beam travel along together and transit time effects are minimized. If the beam velocity differs from the synchronous value, the amplitude of a sinusoidal deflection will be reduced by[6]

$$\frac{\sin \frac{\omega l}{2}\left(\frac{l}{v_s} - \frac{l}{v}\right)}{\frac{\omega l}{2}\left(\frac{l}{v_s} - \frac{l}{v}\right)}, \tag{1}$$
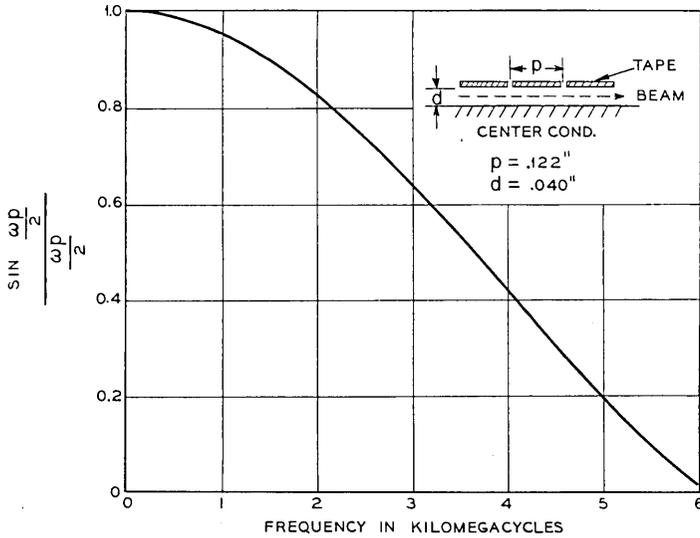
Fig. 3 — Transit time versus frequency for a beam in the presence of parallel plates.

where $\omega$ is the radian frequency, $l$ is the length of the deflecting system and $v_s$ and $v$ are, respectively, the synchronous and the actual beam velocity. Equation (1) shows that, in any case, this deflection system gives better response than would a system using parallel plates. For parallel plates the reduction due to transit time effects is

$$\frac{\sin \frac{\omega l}{2v}}{\frac{\omega l}{2v}}, \tag{2}$$

where the symbols have the meanings given above.

There is another point to consider and that is the transit time effects due to the finite width of tape. Fig. 3 shows the magnitude of this effect. Sinusoidal signals suffer only a loss in amplitude because of transit time effects; however, any other waveform will be distorted. This can best be seen by considering what happens when a square wave with a zero rise time amplitude $V_m$ is applied to this deflecting system.

The vertical deflection on the fluorescent screen will be proportional to the transverse velocity gained by the beam in passing through the deflecting system. This is given by

$$v_y = \frac{e}{m} \int_0^t \frac{V}{d} \, dt. \tag{3}$$

When the voltage goes from zero to $V_{max}$, some electrons will just be entering under a tape. These electrons will be influenced by $V_{max}$ for all the time that they are under the tape. Electrons which have further progressed under the tape will be influenced for a lesser period of time. When the proper limits of integration are placed on (3), it reads

$$v_y = \frac{e}{md} \int_0^{x/u_0} V_m dt,$$

$$= \frac{e}{md} V_m \left(\frac{x}{u_0}\right) \quad 0 \leq x \leq p, \tag{4}$$

$$= \frac{e}{md} V_m \tau \quad p \leq x,$$

where $p$ is the pitch of the tape helix and $\tau$ is the time taken by the beam in traveling a distance of $p \left(\tau = \dfrac{p}{u_0}\right)$. Equation (4) shows that a square wave with a zero rise time will appear on the screen to have a rise time $\tau$. For the dimensions of this tube $\tau$ is $1.65 \times 10^{-10}$ seconds.

## IV. SPOT SIZE AND SENSITIVITY

The elements in this tube which determine spot size are the 0.002″ hole and the subsequent lens system. This hole is imaged on the fluorescent screen and its size at the screen is determined primarily by the intervening lens system. Theoretically the spot size should be 0.007″- however, in most cases the measured size was approximately 0.010″; Part of this difference is due to the fact that the beam boundaries are not clearly defined, consequently measurements may be somewhat inaccurate.

Sensitivity is best expressed as trace width of deflection per volt of applied signal. This is particularly significant since the amount of information which can be obtained from a picture depends not on the absolute height of the picture but rather on the number of elements (trace widths) which make up the picture. Thus it is desirable to have the spot size as small as possible. The vertical sensitivity of this tube is 50 millivolts per trace width and the field is about 100 trace widths both horizontally and vertically. A 25 milliwatt sinusoidal signal will give maximum vertical deflection. The horizontal sensitivity is 0.42 volts per trace width or 42 volts per inch.

## V. WRITING SPEED

One of the most important considerations in high speed oscilloscopes is the writing speed of the instrument. This factor tells how fast the

sweep can travel and still give recordable information. This is usually expressed in trace widths per second. To measure the writing speed of the broad-band oscilloscope a square wave with a 50 millimicrosecond rise time was applied to the tube and the repetition rate of the square wave was lowered until the rise time portion could no longer be observed. With 2000 volts of post acceleration and a square wave with a height of 100 trace widths, rise time was still visible at a repetition rate of 50 times a second. This means that the writing speed was $2 \times 10^9$ trace widths per second when repeated 50 times a second. To compare this with the conventional method of stating writing speeds, the number of "frames" instrumental in forming an image must be known. For a human, the persistence of vision is approximately $\frac{1}{20}$ second. This means that the number of events occuring in this period aid in forming an image; any greater number is superfluous. Using the above numbers, our writing speed becomes $0.8 \times 10^9$ trace widths per second. This would also be the maximum writing speed detectable on film with the same sensitivity as the human eye. The writing speed of this tube should be compared to that of $1 \times 10^{11}$ trace widths per second available in special high speed oscilloscopes. The relatively low writing speed of this tube makes it more useful in observing recurrent phenomena rather than transients (or single pulses). The most effective way of obtaining a greater writing speed is to increase the beam voltage. This results in a beam with a greater power density since not only do the electrons strike the screen with a greater velocity, but also the current density in the beam may be increased. However, as previously mentioned, this results in a loss in sensitivity. This can be overcome to some extent by using a strong post acceleration lens and high voltages.

## VI. ALTERNATE LENS SYSTEM FOR POST ACCELERATION

In order to obtain higher writing speed in the broadband oscilloscope tube, a cross-over lens was built to compare this type of system with "ordinary" post acceleration. An operating voltage of 10,000 was chosen. A photograph and cross-section of the tube is shown in Figs. 4 and 5. The lens consists of two cylinders whose diameters are in the ratio of 1.5 to 1. The small cylinder must be long enough to prevent fields from the lens from penetrating into the horizontal deflection region. This essentially determined the position of the lens center. The diameter of the small cylinder was determined by setting a limit to the amount of aberration which can be tolerated. It was decided that the small cylinder diameter should be three times the maximum deflection at the lens center. The beam with its increased velocity travels about ten inches before it strikes the fluorescent screen.
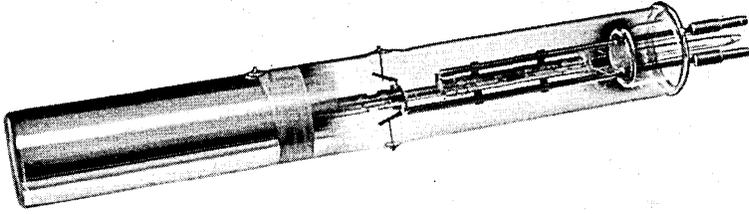
Fig. 4 — Photograph of broadband oscilloscope tube — Model IV.
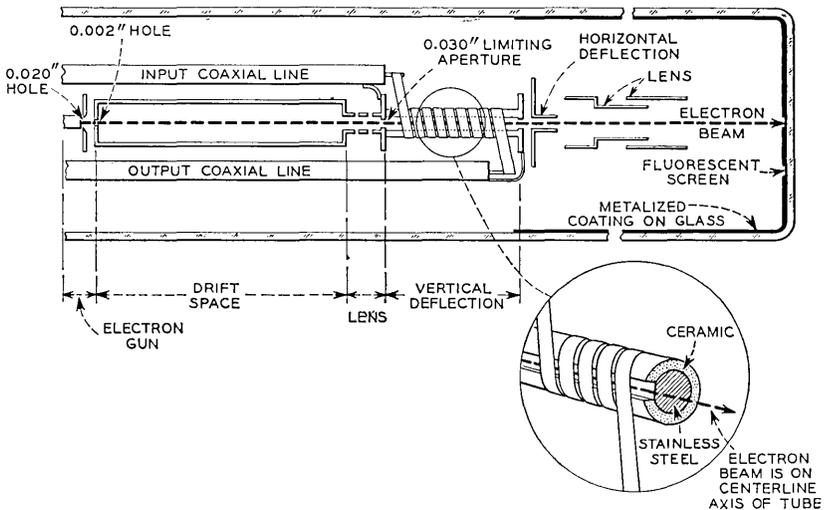


Fig. 5 — Cross section of broadband oscilloscope tube — Model IV.

For this simple lens system illustrated in Fig. 6 the image distance is fixed by the distance from the lens center to the screen. However, the object distance as defined by distance between the lens center and the center of deflection is different for vertical and horizontal deflections. This results from the axial displacement of the deflection regions and causes a difference in the magnification of the horizontal deflection system as compared to the vertical. With the use of Spangenberg's[7] $P$-$Q$ curves one can find the crossover point (2) for each value of $V_2/V_1$. This allows us to calculate the magnification and the results are plotted in Fig. 7, together with the observed magnification. The agreement is fairly good until high values of $V_2/V_1$ are reached — here the observed curve seems to "saturate".

This effect was corrected and the brightness was improved by using an "aluminized" screen. The screen is made by evaporating a thin (approx. 5000 Å) film of aluminum on the back of the phosphor.[8] This is necessary since at high velocities the electrons penetrate so deeply into the phosphor that secondaries do not readily escape; consequently the screen charges to a potential lower than that applied to the lens and the beam is slowed down at the screen. A further advantage of aluminizing is the fact that the aluminum film acts as a mirror and reflects forward part of the light that was originally lost at the back of the screen. Fig. 8 shows the magnification of an aluminized tube as a function of the lens voltages. The "astigmatism" of this particular design is worse than the conventional "post acceleration" systems and more effort is needed to improve this feature.

## VII. MECHANICAL DESIGN

The mechanical design of the broadband oscilloscope tube is similar to the design of a conventional cathode ray tube. That is, the various electrodes which make up the tube are mounted in cylinders which are supported from ceramic rods. The materials used in the tube are type 302 and 304 stainless steel, 326 Monel and molybdenum. The type of glass used is 7052 Pyrex. The conductive coating inside the envelope is tin oxide. The screen is a blue sulphide phosphor, seven to ten microns particle size.

In constructing the tube, the various sub-assemblies consisting of the gun, focusing lens and RF structure are properly aligned and mounted as a unit from the coaxial lines. Before the tube is sealed in its envelope, the RF match of the helix circuit is measured. When the electrical
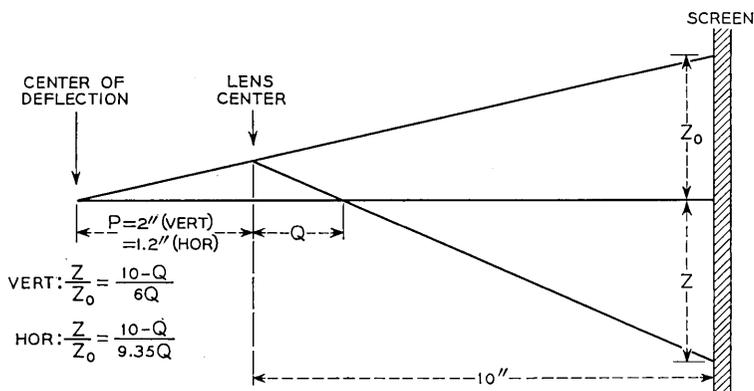


Fig. 6 — Method for determining magnification.
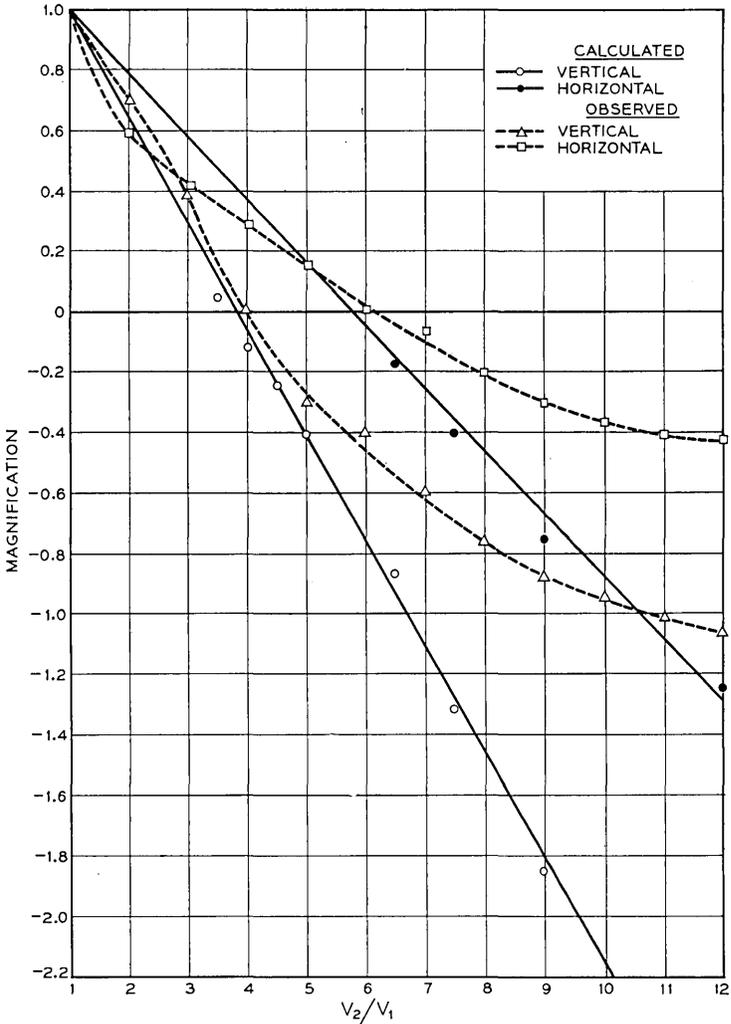
Fig. 7 — Magnification versus $V_2/V_1$, for cross-over lens with non-aluminized screen.

tests have been completed, the alignment of the entire structure is checked using an optical comparator. The alignment of the drift tube, focusing electrodes and RF structure is acceptable when the axial deviation is less than one part in five thousand. The cathode assembly must be within three thousandths of an inch axially.

Although the alignment of the electrodes for the broadband oscilloscope tube appears to be rigid, it is not a difficult tube to build.

VIII. APPLICATIONS

Currently, studies are being made on microwave regenerative repeaters with pulse repetition rates of one hundred and sixty million pulses per second. Timing pulses needed for this system are only 2 millimicroseconds wide at the base. Bandwidths of 500 megacycles or more are required for satisfactory reproduction of these short pulses.

Fig. 9 is an enlarged photograph of timing pulses occurring at the
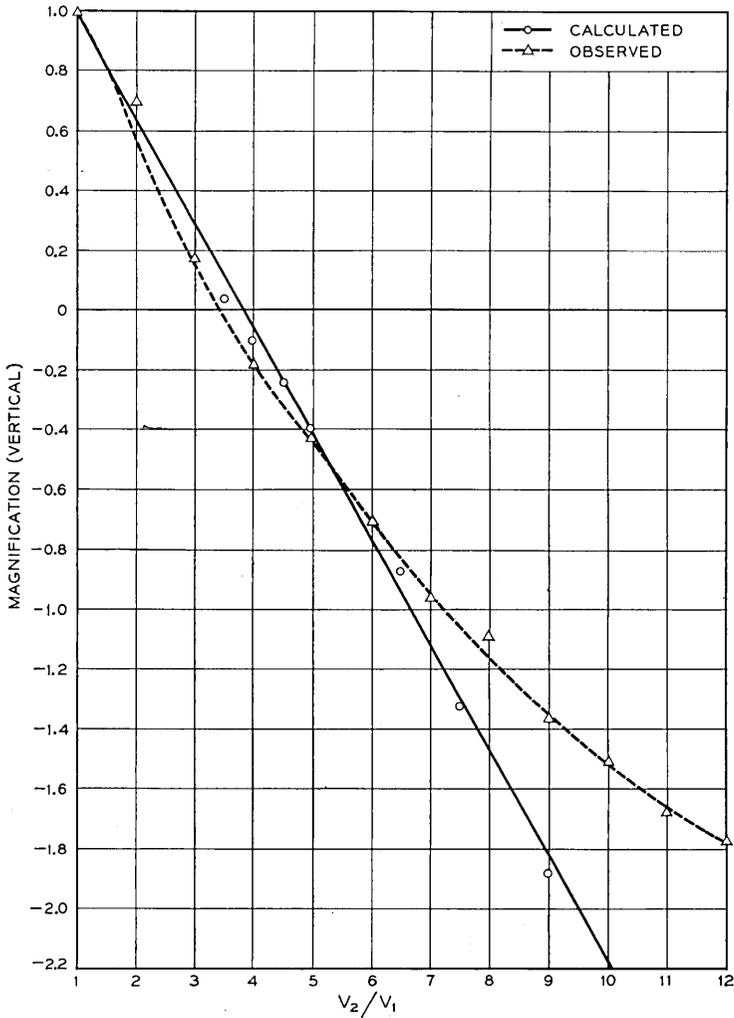


Fig. 8 — Magnification versus $V_2/V_1$, for cross-over lens with aluminized screen.

rate of one hundred and sixty million pulses per second as seen on the face of the broadband oscilloscope tube. It will be noted that the slight imperfections in the wave caused by leakage of pulses at half this rate can be clearly seen. Sine wave frequencies considerably higher than 160 megacycles have been observed on the broadband oscilloscope tube. Fig. 10 shows an enlarged photograph of an 800 megacycle sine wave as seen on the face of the broadband oscilloscope tube.

By using television techniques and televising the image on the face of the broadband oscilloscope tube, a much larger image can be displayed on a television monitor. The waveform on such a monitor is enlarged by a factor of 10, with an over-all vertical deflection sensitivity of one volt per inch. The picture size compares favorably with the size of trace one normally views on a 5″ cathode ray tube. Successful operation of a system of this type requires that the image be a steadily recurring display. In this application the broadband oscilloscope tube might be regarded as a storage device where information is written in at the rate of ten million pulses per second and read out by television scanning at a 60 cycle rate.

Fig. 11 shows a typical photographic reproduction of the face of the television monitor picture tube. Fig. 11(a) is the detected envelope of an 11 kmc pulse group at the output of the microwave modulator. Fig.
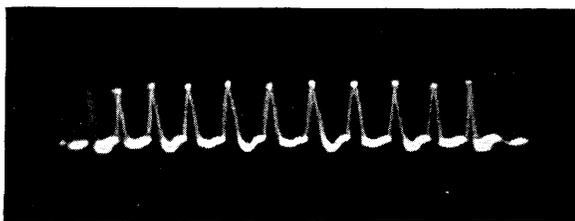


Fig. 9 — Photograph of timing pulses.
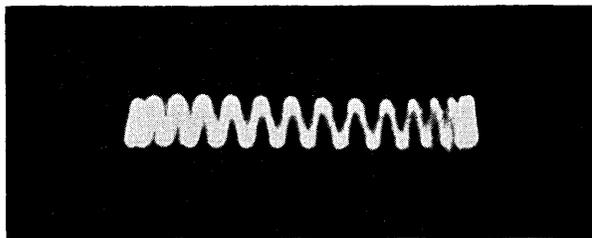


Fig. 10 — Photograph of 800 mc sine wave.
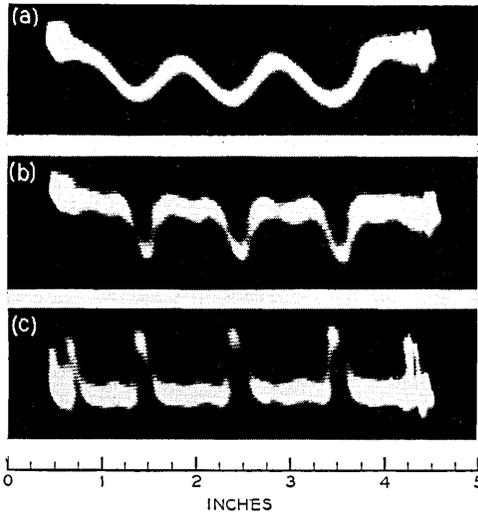
Fig. 11 — Photograph of pulses on the television monitor.

11(b) is the detected envelope of the same pulse group after it has been retimed in an 11 kmc microwave gate by the 2 millimicrosecond baseband timing pulses shown in Fig. 11(c).

IX. CONCLUSION

The broadband oscilloscope tube with its inherent bandwidth has made it possible to pursue studies on systems using millimicrosecond pulses. The present writing speed of this tube restricts its use to high speed repetitive pulses. The writing speed could be increased without a reduction in deflection sensitivity by increasing the cathode current density. The bandwidth could be increased with an improved transition from the coaxial line to the vertical deflection system. In principle the bandwidth could be extended to 2500 megacycles without appreciable loss in sensitivity.

A combined unit employing both the oscilloscope tube and the television system results in a broadband oscilloscope which is capable of resolving one millimicrosecond pulses. The reproduced enlarged images compare favorably in size with presentations on a conventional 5″ oscilloscope tube. The availability of the unit greatly facilitates progress towards a short pulse microwave regenerative repeater which would be capable of transmitting more than 100 million pulses per second.

## X. ACKNOWLEDGEMENTS

REFERENCES

1. O. E. Delange, Experiments on the Regeneration of Binary Microwave Pulses, B.S.T.J., **35,** pp. 67–90, Jan., 1956.
2. S. E. Miller, Waveguide as a Communication Medium, B.S.T.J., **33,** Nov., 1954.
3. W. M. Goodall Television by Pulse Code Modulation, B.S.T.J., **30,** Jan., 1951.
4. J. R. Pierce Traveling Wave Oscilloscope, Electronics, Nov., 1949.
5. K. J. Germeshausen, S. Goldberg and D. F. McDonald, A High Sensibility Cathode Ray Tube for Millimicrosecond Transients, I.R.E. Transactions on Electron Devices, P.G.E.D., p. 152, April 1957.
6. K. R. Spangenberg, *Vacuum Tubes*, Chap. 15, p. 468, McGraw Hill, 1943.
7. K. R. Spangenberg, *Vacuum Tubes*, Chap. 13, p. 379. McGraw Hill, 1948.
8. I. A. D. Lewis and F. H. Wells, *Millimicrosecond Pulse Techniques*, pp. 192–193 McGraw Hill, 1954.

# Optimum Tolerance Assignment to Yield Minimum Manufacturing Cost

By DAVID H. EVANS

(Manuscript received September 9, 1957)

*The problem considered is that of choosing approximate component tolerances in order to minimize mass production costs. The basic item considered is a unit with a single nominal design response. This unit has several components with given nominal design values such that the unit nominal response is as required. We assume that the components are in statistical control and that we can compute the statistical behavior of the response as a function of the assignment of component tolerances. Further, we assume that the cost and salvage value of a unit are known as a function of the assignment of component tolerances. We impose the restriction that the sum of the responses of n identical units in combination must be within a prescribed tolerance with probability 1 − ε. We can then find a relation involving the tolerance limits on the sum of the responses, the rejection limits on the response of a single unit, the variance of the response of a single unit, and the probability ε. Using this relation, which effectively introduces the rejection rate as an additional variable, we then show how to assign component tolerances to minimize production costs. As an illustrative example we consider the design for production of an idealized lumped-constant delay line.*

## I. INTRODUCTION

A valid area of investigation for the cutting of manufacturing costs in the mass production process lies in the assignment of tolerances. In this paper we examine a problem in that area. Consider the following fairly typical sequence of events: A piece of equipment is to be designed with a specified nominal response, for example, an amplifier with a specified nominal gain, or a logic gate with a specified nominal time delay. The circuit is designed and nominal values are assigned to the components of this piece of equipment so that it has the required nominal response. Next, this piece of equipment is to be mass produced, and

461

mass produced economically. One of the manifold problems which arises at this point is the assignment of tolerances to the various components of the piece of equipment. It is at this point in the design for production that the considerations in this paper enter.

The effect of component tolerances is to cause the response to deviate from the nominal in a statistical manner. A common approach to component tolerance assignment ignores the statistical behavior of the response deviation and bases the tolerance assignment on the "worst case" approach, i.e., the deviations from nominal for all components are assumed to act in concert to maximize the deviation from nominal of the response. This criterion corresponds to a very pessimistic viewpoint because, usually, the probability of such a simultaneous occurrence of worst values is extremely small. In fact, it is often so small that in a very practical sense it is zero. Within the past several years the statistical approach to assigning component tolerances, which makes use of the statistical nature of the response deviation, has been gaining in popularity. J. M. Juran[1] gives examples and several references to uses of statistical tolerancing. A fine case history of a statistical tolerancing approach is that of the design for production of the repeaters used in the Bell Systems L3 coaxial system.[2] In References 3 and 4 the particular problems of statistical distribution requirements and quality control requirements for the components of the L3 system are considered. In statistical tolerancing, in order that the deviation of the response be in control, it is necessary that the component manufacturing processes either be in control or sufficiently compensated so that they are effectively in control at all times. We will assume statistical tolerancing in this paper; thus, we are also forced to assume the restrictive implication that the component manufacturing processes are in control.

In any kind of tolerancing there are many possible component tolerance assignments for which the response tolerances are identical or reasonably so. The costs associated with the different component tolerance assignments, however, will not in general be the same.

*For example, consider an R-C circuit.* Assuming for the sake of this example that both resistors and capacitors come in truly uniform distributions it is obvious that the statistical behavior of the time constant ($\tau = RC$) will be the same if the resistor is from a 5 per cent distribution and the capacitor is from a 10 per cent distribution or if the resistor is from a 10 per cent distribution and the capacitor is from a 5 per cent distribution. However, the costs will generally be different.

The desired tolerance assignment is the least expensive tolerance assignment (of those tolerance assignments which engender identical

response tolerances). Pike and Silverberg[5] have considered this problem for linear, or approximately linear, (mechanical) systems using statistical tolerancing. They show how to adjust the component tolerances (actually the variances) to get maximum value for minimum cost.

Next let us discuss some characteristics of the particular type problem we wish to consider:

## 1.1 *Response Tolerances*

It very often happens that the deviation from nominal of the response of an individual piece of equipment — or unit as we shall call it henceforth for brevity — is relatively unimportant; the quantity which is important is the algebraic sum of the deviations of the responses from nominal of a combination of several units.* For example, the repeaters in the L3 system are in series and the primary requirement is that the sum of the gains compensate for the line loss plus or minus a small tolerance. Another example of this type is a string of several logic gates for which the total time delay must be less than some prescribed value. The sum requirement gives us considerably more latitude in the assignment of the response tolerances for the individual units because of the nature of a sum of random variables — for indeed, the deviation from nominal of the response is a random variable under statistical tolerancing.

## 1.2 *Rejection Rate*

A criterion which is often used to measure the efficiency and economy of a production process is the rejection rate. Completed or uncompleted units may be rejected for any number of reasons, but here we confine our attention to those units which are rejected solely because the deviation of their components from nominal is such that their responses are out of tolerance. That is, we ignore those units which must be rejected because of cold solder joints, flaws, broken leads, and a multitude of similar causes. Hereinafter, we shall use the term rejection rate to mean the fraction of completed units which have a response which is outside of tolerance but which are otherwise acceptable. The usual assumption is that the rejection rate must be small for an economical production process, however, we take the viewpoint that the rejection rate is an-

---

* We have chosen to use the hierarchy: components, units, combinations of units. This triad may be thought of as corresponding to any similarly ordered threesome in any hierarchy which may be more familiar to the engineer, e.g., raw materials, piece parts, subassemblies, assemblies, units of product, subsystems, and systems.

other variable which may be introduced in order to minimize production costs. Note that this implies 100 per cent testing on finished units and the consequent added cost thereof.

But then, what of the rejected units? The rejected units will have some salvage value. The salvage value for a rejected unit may range from a positive value which is a fairly large percentage of the cost of manufacturing a unit (such as would be the case if only a small additional charge were necessary to bring the unit into tolerance or if out of tolerance units could be selectively assembled), to a negative value (such as would be the case if the unit were a total loss and there was an additional charge to dispose of it). In the most general case the salvage value is a statistical quantity since its value might depend, for example, on how far out of tolerance the response is or what component or combination of components is the essential cause of the response being out of tolerance.

### 1.3   *Aim*

Before going to the analysis, I would like to indicate the tenor of this work. Certainly we are trying to decrease production cost by an intelligent assignment of tolerances. However, it is important to note that this assignment is made at a point in the production process immediately after the final circuit design and specification of nominal component values have been completed. At this stage only the rudiments of the projected manufacturing process are known since many final answers must await the assignment of component tolerances. Hence, the figures for the production costs and the salvage values are not known precisely and may be in fact only educated guesses; in addition, the distributions for some of the components may not be known precisely. And further, it would be uneconomical to get precise estimates of the figures for each and every possible combination of component tolerances which could reasonably be used in the production model since the number of such combinations can easily be enormous. Thus, since the cost figures are not known precisely, it would be so much wasted effort to make the rest of the analysis exact. The principal advantage to be gained from the following analysis is to eliminate all except, say, two or three possible combinations of component tolerances for ultimate consideration for the production model.

### II.   GENERAL STATEMENT OF THE PROBLEM

Let us denote the deviation from nominal of the response of a unit by $x$; $x$ is a random variable. Let there be $n$ units in combination;*

---

* Combination, as we use it here, implies only that (1) holds.

$n$ is a fixed but arbitrary number. Let $x_i$ be the deviation from nominal of the response of the $i$th unit, $i = 1, 2, \cdots, n$, and let the $x_i$'s be independent. Let it be required that, for proper over-all operation, the algebraic sum of the random deviations of the $n$ units in combination be constrained to lie between $\pm B$; i.e.,

$$| x_1 + x_2 + \cdots + x_n | \leqq B. \tag{1}$$

That is, $\pm B$ are the tolerance limits on the deviation from nominal of the response for the combination of $n$ units.

As is usual in statistical tolerancing let us be willing to assume a small risk $\epsilon$ that the combination will not operate properly, i.e., that the sum (1) will exceed $B$. Thus,

$$\Pr(| x_1 + x_2 \cdots + x_n | > B) = \epsilon. \tag{2}$$

Next let us look at the assignment of component tolerances. The assignment variable is really the independent variable in a tolerancing problem. That is, let the unit which is to be manufactured have $k$ components; number these components arbitrarily, $1, 2, \cdots, k$. Let component $\#j$ be available in $r_j$ different tolerance distributions which are to be considered as candidates for possible use in the production model of the unit; number these tolerance distributions arbitrarily, $1, 2, \cdots, r_j$. Do the same for all components, $j = 1, 2, \cdots, k$.

*For example, suppose that component $\#j$ is a resistor.* Let the available tolerance distributions considered be 5 per cent resistors and 10 per cent resistors. Thus $r_j = 2$ and we can arbitrarily number the 5 per cent resistors as distribution $\#1$ and the 10 per cent resistors as distribution $\#2$.

A particular assignment of component tolerances can thus be characterized by the ordered set of numbers

$$(i_1, i_2, \cdots i_k) \tag{3}$$

which is to be understood to mean that component $\#1$ comes from the $i_1$th distribution from the set of available distributions for component $\#1$, component $\#2$ comes from the $i_2$th distribution from the set of available distributions for component $\#2$, and so forth. The number of different tolerance assignments can be very large since the assignments range over all possible combinations of available distributions.* Although in the above we have only considered a finite number of tolerance distributions for each component there is no reason, in principle, why one

* The number is $\prod\limits_{j=1}^{k} r_j$.

or more of the components cannot come from a continuum of possible tolerance distributions.

The component tolerance assignment variable, (3), is unwieldy; let us replace it by a more manageable independent variable. To do so we argue as follows: Since the deviation from nominal of the response of a unit, $x$, is a random variable, as such, it is characterized by a probability distribution, say $D(x)$. But $D(x)$ also depends on which set of component tolerances is used since different assignments of component tolerances will, in general, manifest themselves in different statistical behaviors for the response. A measure of the distribution $D(x)$ is the variance of $x$ — var $x = \sigma^2$. The quantity $\sigma^2$, or $\sigma$, is an excellent measure if all distributions $D(x)$ are normally distributed with mean zero, as we shall shortly assume is the case in our problem; otherwise, the aptness of $\sigma$ diminishes as $D(x)$ departs from normal with mean zero. Thus we can make the new independent variable $\sigma$ (or $\sigma^2$, whichever is more convenient) instead of (3). The range of $\sigma$ is determined by considering all possible combinations of component tolerances and $\sigma$ can only take on the discrete values determined by the possible combinations of component tolerances (if only a finite number of distributions is considered for each component). Note that, at this point, the correspondence between $\sigma$ and the particular assignment of component tolerances is not necessarily one to one, (see example in Section I). A unique (or effectively unique) correspondence will come about naturally when we consider the costs, below.

Before considering the costs we must consider the rejection rate. In the introduction we defined the rejection rate as the fraction of the completed units which have responses outside of tolerance but which are otherwise acceptable. That is, if the tolerance limits on the unit response are $\pm b$,* then every unit which has a response deviating from the nominal response by more than $\pm b$ is to be rejected, i.e., reject all units such that $| x | > b$. Since the rejection rate is a variable, $b$ is a variable which must be determined. The tolerance limit $b$ is a function of three variables $\sigma$, $B$, and $\epsilon$, and must be chosen to satisfy (2). Qualitatively, for fixed $B$ and $\epsilon$ it is obvious that, in order to satisfy (2), as $\sigma$ increases $b$ must decrease and vice versa. We will obtain a quantitative relation later.

Finally we consider the manufacturing costs per unit. We distinguish two types, the *raw cost* and the *real cost*. The raw cost per unit is the amount of money which must be spent to manufacture one unit regardless of whether it has a response which is or is not within the tolerance

---

* We are only going to consider symmetrical distributions about the nominal, hence $b$ is sufficient.
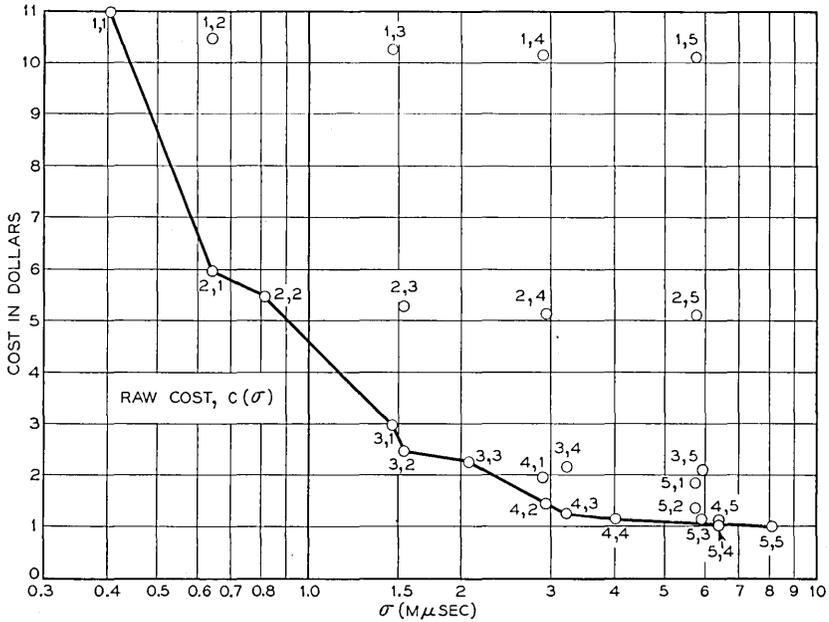
Fig. 1 — Raw cost for a single $L$-$C$ section of a lumped constant delay line as a function of the standard deviation of the delay per section. (See Table II-8 for explanation of the code $(i,j)$).

limits determined by $\pm b$ and independent of any salvage value a unit may have. The real cost per unit is the raw cost plus the unsalvageable raw cost per rejected unit prorated among the units within tolerance. It is obvious that for a well behaved manufacturing process the raw cost should be a monotone decreasing function of $\sigma$. Furthermore, if two or more different component tolerance assignments give the same — or approximately the same — variance, $\sigma^2$, the assignment which should be chosen to correspond to that $\sigma$ is the one which minimizes the raw cost. A better statement of the criterion for choosing the component tolerance assignments which make up the raw cost curve as a function of $\sigma$ is that an assignment lies on the raw cost curve if there is no other assignment which has *both* a smaller (or as small) $\sigma$ *and* a lower (or as low) cost.* A raw cost curve, $C(\sigma)$, is illustrated in Fig. 1, i.e., if

---

* Stated precisely, the points of the raw cost function as a function of $\sigma$, $C(\sigma)$, are determined as follows: Let us denote the component tolerance assignment variable, (3), by $\beta$; let the raw cost for each $\beta$ be $C(\beta)$; let the variance of the response for each $\beta$ be $[\sigma(\beta)]^2$; then the points of the raw cost curve are given by

$$C(\sigma) = \min C(\beta) \qquad (4)$$

where (a) the minimum is taken over all $\beta$ such that $\sigma(\beta) \leq \sigma$, and (b) the only allowed values of $\sigma$ are those such that there exists a corresponding $\beta$ *and* $C(\beta) = C(\sigma)$, i.e., the "corners" of (4).

one plots the point $(\sigma, C)$ — or equivalently $(\sigma^2, C)$ — for each assignment of component tolerances then $C(\sigma)$ is the set of points which determine the polygonal curve which is the lower envelope of the set of points for all possible combinations of component tolerances. The set of points $C(\sigma)$ are connected for illustrative purposes only; $C(\sigma)$ exists only as a pointwise function (for component tolerances which do not come from a continuum of allowed tolerances).

Finally there is the salvage value for the rejected units. We denote the salvage value by $\alpha(\sigma)C(\sigma)$, i.e., $\alpha$ is the ratio of the salvage value to the raw cost. In the general case the salvage will be a function of $\sigma$, i.e., of the particular set of component tolerances, and it will also be a random variable which depends on $x$, the deviation of the response from nominal. We will retain the dependence of $\alpha$ on $\sigma$; but we will ignore the fact that it may be a random variable and take $\alpha$ as a constant for each $\sigma$. This constant value may be an expected value. The assumptions set forth on $\alpha$ are in accord with the aim set forth in the introduction, for, if the cost figures are not precise estimates, then certainly the salvage value as a distribution function cannot be known precisely. If we took $\alpha$ in all its generality, we would only succeed in cluttering up the analysis with functions and figures for which we could not possibly get realistic estimates. We can, however, reasonably expect to get a realistic estimate for the expected value of the salvage as a function of the component tolerance assignment, or equivalently $\sigma$. Along this same line of reasoning, in connection with the salvage value, we note that we assumed that if two different assignments of component tolerances give the same $\sigma$ then the possible difference in their salvage values was to be ignored in choosing the assignment which determines the raw cost curve $C(\sigma)$. This assumption could possibly lead to a real cost which is higher than necessary since the salvage value is inherent in the real cost. However, the possibility of such an occurrence is doubtful, and if such an occurrence were suspected it could always be calculated as a special case.

The problem we want to solve is:

    (a)  given $n$, $B$, $\epsilon$, as defined in the first two paragraphs of this section,

    (b)  given the raw cost as a function of $\sigma$, $C(\sigma)$,

    (c)  given the salvage as a function of $\sigma$, $\alpha(\sigma)$,

    find the value of $\sigma$ such that the real cost, $C^*(\sigma)$, is a minimum and find the tolerance limits on the unit response, $\pm b$.

We are able to solve this problem under restrictive but widely applicable conditions.

III.   RELATIONSHIP BETWEEN $b$ AND $\sigma$

We assume that the deviation of the response from nominal of a unit, $x$, is normally distributed with standard deviation $\sigma$. This is a realistic assumption if the components used in manufacturing the unit have independent random variations (not necessarily normally distributed) which in turn influence the response additively. We assume, further, that the mean of $x$ is always zero which in turn implies that the mean does not shift significantly with change in $\sigma$, i.e., with change in the assignment of component tolerances, and further that the component manufacturing processes are in control.

Let $F(y)$ be the cumulative normal distribution function and $\varphi(y)$ the normal probability density function:

$$F(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-t^2/2} \, dt,$$
$$\varphi(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}. \tag{5}$$

If the tolerance limits on $x$ are $\pm b$ then the probability that an individual unit will be rejected, i.e., the rejection rate, is

$$\text{Probability of rejection} = \text{rejection rate} = 2[1 - F(b/\sigma)]. \tag{6}$$

Since only the units which fall within the rejection limits $\pm b$ are to be used in the combination of $n$ units, the probability density function for the acceptable units is

$$\psi(x) = \begin{cases} \dfrac{\varphi(x/\sigma)}{\sigma[2F(b/\sigma) - 1]}, & |x| < b \\ 0, & |x| > b \end{cases}. \tag{7}$$

We want next to find the distribution function for the random variable

$$\xi = x_1 + x_2 + \cdots + x_n, \tag{8}$$

where the $x_i$ are independent and distributed according to (7). We assume that $n$ is sufficiently large to apply the central limit theorem. Performing the necessary integration to find the variance of $x$ distributed according to (7), we have that $\xi$ is normally distributed with mean zero and variance

$$\sigma_\xi^2 = n\sigma^2 u(b/\sigma), \tag{9}$$

where

$$u(t) = \left[ 1 - \frac{2t\varphi(t)}{2F(t) - 1} \right], \quad t > 0. \tag{10}$$

Note that $u(t) \leqq 1$.

Because of the above assumptions we can rewrite (2) as

$$F(B/\sigma_\xi) = 1 - \epsilon/2. \tag{11}$$

From tables for $F(y)$ we can find the standard normal deviate $r = B/\sigma_\xi$ for a given risk $\epsilon$. Introducing $r$ in (9) to eliminate $\sigma_\xi$ we find

$$(B/\sigma)^2 = r^2 n u(b/\sigma). \tag{12}$$

Equation (12) gives the desired relationship among $B$, $r$ (or $\epsilon$), $\sigma$ and $b$ in order to satisfy (2) for an arbitrary value of $\sigma$. Note that if, in trying to satisfy (12), $u$ turns out to be greater than one this simply means that although all units are accepted the probability that $|\sum x_i|$ exceeds $B$ is still less than $\epsilon$.

## IV.   REAL COST PER UNIT

The raw cost per unit, as we have defined it, does not include the penalty that must be paid for producing units which are outside of tolerance and therefore must be sent to salvage, nor does it include any salvage value the rejected units may have. Call the raw cost per unit

$$C = C(\sigma). \tag{13}$$

In addition to knowing the raw cost we must also know the salvage value of a rejected unit, i.e., a unit such that $|x| > b$. As before, we define the salvage value per unit to be

$$S = \alpha(\sigma)C(\sigma). \tag{14}$$

Here, $\alpha$ is a proportionality factor which will, in general, depend on $\sigma$. Obviously $\alpha$ is less than 1; on the other hand it may range downward thru negative values, e.g., if it costs additional money to dispose of a rejected unit.

Define the real cost per unit to be, as before,

$$C^* = C^*(\sigma). \tag{15}$$

The real cost is related to the raw cost and the salvage value as follows: If $M$ units are produced in all and $m$ of these $M$ units must be rejected and sent to salvage because their responses are out of tolerance, then

$$C^*(\sigma) = \frac{1}{M - m} [MC(\sigma) - m\alpha(\sigma)C(\sigma)] \tag{16}$$

is the real cost per unit, i.e., $C^*$ is the total raw cost for all units produced minus the salvage value of the unacceptable units all prorated among the acceptable units. For large $M$, $m/M$ is the probability that a unit will fall outside of tolerance, i.e., it is the rejection rate, (6); hence

$$C^*(\sigma) = \left[ \frac{1 - \alpha}{2F(b/\sigma) - 1} + \alpha \right] C(\sigma). \tag{17}$$

By proper choice of $\sigma$, we want to minimize the function $C^*$.

## V.   MINIMIZATION OF THE REAL COST

In principle we could give functional forms for $C(\sigma)$ and $\alpha(\sigma)$ and then minimize $C^*(\sigma)$ by the usual analytical methods. However, one would rarely, if ever, know the functional form for either. Hence we go to a graphical method.

So that the necessary calculations may be carried out expeditiously we redefine some of the previously formulated functions. First, however, let us see exactly what is desired. We are given $B$, $\epsilon$(or $r$) and $n$. We want to calculate $C^*(\sigma)$ throughout the range of interest of $\sigma$ (or specifically, for a set of values of $\sigma$ in the range of interest). After plotting $C^*(\sigma)$ we can pick off the minimum, or minimums, of $C^*$; we then need to calculate $b$ for the minimum, or minimums. The calculation of $C^*(\sigma)$ and $b$ for given $B$, $r$, $n$, $\sigma$, can be done stepwise:

1. From (12) calculate $u$.

2. From (10) calculate the implicitly defined variable $t = b/\sigma$ for $u$ from step 1. (The correspondence between $t$ and $u$ is one-to-one since $u$ is a strictly monotone* function of $t$). This step essentially gives us $b$.

3. From (17), using $t = b/\sigma$ from step 2, calculate $C^*(\sigma)$.

Now that we know exactly what is desired we can expedite the calculations. A convenient combination of the variables is

$$q = \frac{r\sqrt{n}}{B} \sigma = u^{-1/2}. \tag{18}$$

Since $u(t)$, defined by (10), is strictly monotone it can be inverted (numerically) to get

$$t = w(q). \tag{19}$$

Also, define the function

$$H(q) = \frac{1}{2F(w(q)) - 1} = \frac{1}{2F(t) - 1} \tag{20}$$

---

* The fact that $u$ is strictly monotone will become obvious from the graph of the related function (19), Fig. 2.

## TABLE I

### $H(q)$ and $w(q)$ as functions of the argument $q$

| $q$ | $H$ | $w$ |
|---|---|---|
| 1.01 | 1.002 | 3.11 |
| 1.02 | 1.004 | 2.86 |
| 1.03 | 1.007 | 2.69 |
| 1.04 | 1.010 | 2.57 |
| 1.05 | 1.013 | 2.48 |
| 1.06 | 1.017 | 2.40 |
| 1.07 | 1.020 | 2.33 |
| 1.08 | 1.024 | 2.26 |
| 1.09 | 1.028 | 2.21 |
| *1.10 | 1.032 | 2.16 |
| 1.12 | 1.041 | 2.05 |
| 1.14 | 1.049 | 1.99 |
| 1.16 | 1.058 | 1.92 |
| 1.18 | 1.069 | 1.85 |
| *1.20 | 1.077 | 1.80 |
| 1.25 | 1.10 | 1.68 |
| 1.30 | 1.13 | 1.58 |
| 1.35 | 1.16 | 1.49 |
| 1.40 | 1.19 | 1.42 |
| 1.45 | 1.22 | 1.35 |
| *1.50 | 1.25 | 1.29 |
| 1.60 | 1.31 | 1.19 |
| 1.70 | 1.37 | 1.11 |
| 1.80 | 1.43 | 1.03 |
| 1.90 | 1.49 | 0.98 |
| *2.00 | 1.56 | 0.92 |
| 2.50 | 1.90 | 0.72 |
| 3.00 | 2.25 | 0.59 |
| 3.50 | 2.60 | 0.50 |
| 4.00 | 2.95 | 0.44 |
| 4.50 | 3.30 | 0.39 |
| *5.00 | 3.65 | 0.35 |
| 6.00 | 4.37 | 0.29 |
| 7.00 | 5.09 | 0.25 |
| 8.00 | 5.82 | 0.22 |
| 9.00 | 6.54 | 0.19 |
| 10.00 | 7.27 | 0.17 |

\* Indicates a change in increment of the argument $q$

using $t$ and $w$ as defined by (19). Both of the functions $H$ and $w$ have the same independent variable, $q$. In terms of the functions $w$ and $H$ we have

$$C^*(\sigma) = \left[ (1 - \alpha)H\left(\frac{r\sqrt{n}}{B}\sigma\right) + \alpha \right] C(\sigma), \qquad (21)$$

and

$$b = \sigma \cdot w\left(\frac{r\sqrt{n}}{B}\sigma\right) \qquad (22)$$

as the desired formulas. The functions $w(q)$ and $H(q)$ are tabulated in Table I and plotted in Fig. 2 for convenient use. We note for $q < 1$ that $t = w(q)$ is infinite; this simply expresses the fact that for $\sigma$ sufficiently small the probability that $| \sum x_i |$ will exceed $B$ is less than $\epsilon$, and, hence, that the rejection limits are $\pm b = \pm \infty$, cf., remark about $u$, following (12).

In terms of $H$, the rejection rate (6) is

$$\text{Rejection rate} = 1 - 1/H. \tag{23}$$
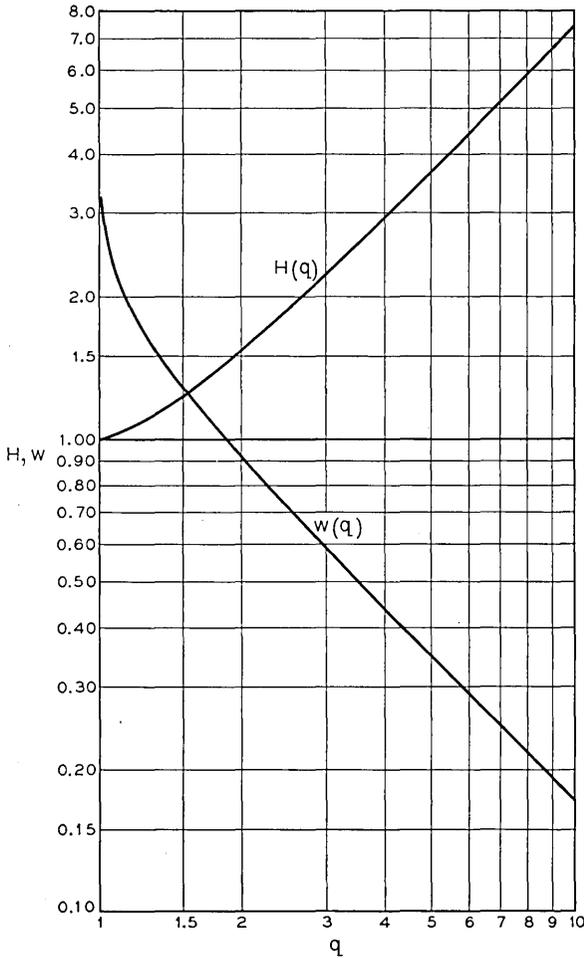


Fig. 2 — $H(q)$ and $w(q)$ as functions of the argument $q$.

## VI. COMPARISON WITH OTHER CRITERIA

In the above we have tacitly assumed that each unit would be tested. Let us now consider the case in which this test is omitted, that is, at least as a test on 100 per cent of the units. One can still satisfy (2) by choosing the proper value for $\sigma$. Since no units are rejected the distribution for $\xi = \sum x_i$ is normal with mean zero and variance $\sigma_\xi^2 = n\sigma^2$. Hence (2) becomes

$$F(B/\sigma\sqrt{n}) = 1 - \epsilon/2. \tag{24}$$

Letting $r$ be the standard normal deviate which satisfies (24) one has

$$\sigma = B/r\sqrt{n}, \qquad (b = \infty). \tag{25}$$

This is, of course, well known and is in use. In comparing the cost by this last method with the cost by the previous method, one must remember to take into account the cost of 100 per cent testing of units. The testing cost could easily swing the balance in favor of the no-test method.

Another criterion to consider is the zero risk case. Here, $\epsilon$ equals zero and the rejection limits are then given by $b = B/n$. It still remains to choose the optimum $\sigma$ for the manufacturing process. Proceeding in the same manner as previously, one finds that the real cost is given by

$$C^*(\sigma) = \left[ \frac{1 - \alpha}{2F(B/n\sigma) - 1} + \alpha \right] C(\sigma), \tag{26}$$

This can obviously be plotted as a function of $\sigma$ and the minimum for $C^*$ obtained graphically. Note that in this case the component distributions do not have to be in control to satisfy the tolerance limits $\pm B$; however, they must be reasonably in control to make (26) true.

## VII. EXAMPLE

As an idealized example of the method described in the foregoing we consider the design for production of a lumped constant delay line. This example is meant to be strictly illustrative since we wish to concentrate on explaining the technique. We are going to ignore some factors which must be taken into account in practical applications but which, in the present example, would only serve to clutter up the explanation. For example, the costs would be influenced by whether we use printed wiring, just how and what we are salvaging, whether special care should be taken with certain close tolerance components, the testing cost as a function of the limits, and so on.

With the above reservations in mind we make the following specifications for the example:

1. An $L$-$C$ section will be the unit, the delay will be the response.

2. The raw cost of the unit will be the sum of the costs for the inductor and the capacitor plus a fixed cost $K$ independent of the component tolerances. We will use $K$ as a parameter.

3. The salvage value of a rejected unit will be one-half of the component costs.

4. All component distributions will be normal distributions about the nominal and will be in control.

We will consider two examples which differ from one another only in the tolerance on the over-all delay $B$, and for each example we will consider several different values for the fixed (i.e., independent of $\sigma$) cost $K$ to be added to the component cost to get the total raw cost. We introduce these variations to give the reader a quantitative idea of the trends they induce. We use the values given in Table II.

We must first examine the distribution of the delay (response) of the individual $L$-$C$ sections as a function of the component tolerance distributions. Normalizing the formula for the delay so that $L$ is in $\mu h$, $C$ is in $\mu\mu f$, and $\Delta$, $x$ are in $m\mu sec$,

$$\Delta = 100 + x = \sqrt{CL} \qquad (27)$$

(where $x$ is the deviation of the response from nominal). Linearizing (27) and using the ordinary linear propagation of error formula one finds that the variance of $x$ is

$$\sigma^2 = \tfrac{1}{4} [\sigma_L{}^2 + \sigma_C{}^2], \qquad (28)$$

where $\sigma_L$ (in $\mu h$) is the standard deviation of $L$, $\sigma_C$ (in $\mu\mu f$) is the standard deviation of $C$, and $\sigma$ (in $m\mu sec$) is the standard deviation of $x$. One should satisfy himself that $x$ is normally, or approximately normally, distributed with mean zero, or approximately zero, and variance as in (28), or approximately as in (28). We can do so by using the nonlinear propagation of error approximate formulas[6] for the range of combinations of distributions for the components, Table II-8. It turns out that (see Table III)

(a) the variance as given by (28) is negligibly different from the true variance ($\leq 1$ per cent, see column 5, Table III),

(b) the coefficient of skewness $\beta_1 = \mu_3{}^2/\sigma^6$ is small ($\leq 3$ per cent for all combinations, $<1$ per cent for most combinations, column 6, Table III) and

(c) the coefficient of excess $\beta_2 = \mu_4/\sigma^4 \approx 3$, the standard for normal

TABLE II — VALUES FOR EXAMPLE

| | |
|---|---|
| 1. Nominal inductance: | $L_0 = 100\ \mu h$ |
| 2. Nominal capacitance: | $C_0 = 100\ \mu\mu f$ |
| 3. Nominal delay per unit: | $\Delta_0 = \sqrt{LC} = 0.1\ \mu sec$ |
| 4. Number of units per delay line: | $n = 10$ |
| 5. Nominal delay of delay line: | $n\Delta_0 = 1\ \mu sec$ |
| 6. Tolerance on delay line, | |
|     First example: | $B = 5\ m\mu sec\ (0.5\ per\ cent)$ |
|     Second example: | $B = 15\ m\mu sec\ (1.5\ per\ cent)$ |
| 7. Assumed risk of out of tolerance delay line: | $\epsilon = 0.01$ per cent |
| | $(r = 3.89)$ |

8. Cost versus tolerance of Components:

| Code† | Inductors | | Capacitors | |
|---|---|---|---|---|
| | $\sigma_L$ in $\mu h$ | Cost | $\sigma_C$ in $\mu\mu f$ | Cost |
| 1 | 0.577   (1%)* | $10.00 | 0.577   (1%)* | $1.00 |
| 2 | 1.155   (2%) | 5.00 | 1.155   (2%) | 0.50 |
| 3 | 2.887   (5%) | 2.00 | 2.887   (5%) | 0.30 |
| 4 | 5.774  (10%) | 1.00 | 5.774  (10%) | 0.20 |
| 5 | 11.55  (20%) | 0.90 | 11.55  (20%) | 0.15 |

| | | |
|---|---|---|
| 9.  Fixed cost parameter to be added to compo- | $K =$ | $0.30 |
|     nent costs to get total raw cost, both exam- | | 1.00 |
|        ples: | | 3.00 |
| | | 5.00 |
| | | 10.00 |
| | | 15.00 |
| 10. Salvage value: | $\alpha C = \frac{1}{2}$ component cost for | |
| | all $\sigma$. | |

---

\* The figures in parenthesis give the tolerances for uniform distributions which have the same standard deviations as the normal distributions.

† We will use the code number to refer to these distributions. The same code number is used for both inductors and capacitors. For a pair of components we will use the code pair $(i, j)$ where $i$, the first entry, is the code for the inductor distribution, and $j$, the second entry, is the code for the capacitor. For example the code pair $(2, 1)$ means that $\sigma_L = 1.155$ at a cost of \$5.00 per inductor and $\sigma_C = 0.577$ at a cost of \$1.00 per capacitor.

distributions, (column 7, Table III) for the whole range of combinations of distributions. However, one does get into some small difficulty with the mean. The non-linear formula for the average, after dropping terms which turn out to be negligible in this case, is

$$\text{ave } x = \tfrac{1}{2}\ \Delta_{CC}\sigma_C^2 + \tfrac{1}{2}\ \Delta_{LL}\sigma_L^2,$$
$$= -\tfrac{1}{8}\ 10^{-2}(\sigma_C^2 + \sigma_L^2), \tag{29}$$

where the partial derivatives are evaluated at $L = 100$, $C = 100$. One finds that

$$\frac{\text{ave } x}{\sigma} = \frac{\bar{x}}{\sigma} = -\frac{1}{4}\cdot 10^{-2}\sqrt{\sigma_L^2 + \sigma_C^2} \tag{30}$$

TABLE III

Various Indices of the Distribution of $x$ as a Function of the Component Distributions. (For a normal distribution with mean zero and variance $\sigma^2 = \frac{1}{4}(\sigma_L{}^2 + \sigma_C{}^2)$ these indices have the values shown in the first row.)

| Code | $-\dfrac{\bar{x}\,\P}{\sigma}$ | $-\dfrac{10\bar{x}\,\P}{5}$ | $-\dfrac{10\,\bar{x}\,\P}{15}$ | $\dfrac{\delta\sigma^2}{\sigma^2}$ | $\beta_1$† | $\beta_2$ | Minimums of C*‡ at (√) for | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | $B = 5$ | $B = 15$ |
| $N(0,\sigma)$ | 0 | 0 | 0 | 0 | 0 | 3 | | |
| 1,1 | 0.002 | 0.002 | 0.001 | 0.000+ | 0.000+ | 3 | √ | |
| 1,2 | 0.003 | 0.004 | 0.001 | 0.000+ | 0.000+ | 3 | √ | |
| 1,3 | 0.007 | 0.022 | 0.007 | 0.001 | 0.001 | 3 | | |
| 1,4 | 0.015 | 0.084 | 0.028 | 0.003 | 0.007 | 3 | | |
| 1,5 | 0.028 | 0.326 | 0.108 | 0.012 | 0.030 | 3 | | |
| 2,2 | 0.004 | 0.007 | 0.002 | 0.000+ | 0.000+ | 3 | | |
| 2,3 | 0.008 | 0.024 | 0.008 | 0.001 | 0.000+ | 3 | √ | √ |
| 2,4 | 0.015 | 0.088 | 0.029 | 0.003 | 0.006 | 3 | | |
| 2,5 | 0.029 | 0.339 | 0.113 | 0.012 | 0.028 | 3 | | |
| 3,3 | 0.010 | 0.042 | 0.014 | 0.001 | 0.000+ | 3 | | |
| 3,4 | 0.016 | 0.104 | 0.034 | 0.002 | 0.001 | 3 | | √ |
| 3,5 | 0.031 | 0.357 | 0.119 | 0.011 | 0.020 | 3 | | |
| 4,4 | 0.020 | 0.167 | 0.056 | 0.003 | 0.000+ | 3 | | |
| 4,5 | 0.032 | 0.419 | 0.140 | 0.009 | 0.005 | 3 | | |
| 5,5 | 0.041 | 0.671 | 0.224 | 0.010 | 0.000+ | 3 | | |

¶ $\bar{x} = $ ave $x$.
† $\sqrt{\beta_1}$ is negative, i.e., negative skewness.
‡ Checks (√) in the last two columns indicate combinations which are used in final solutions for $C^*$. The left column is for $B = 5$, the right for $B = 15$.

is not significant ($\leq 4$ per cent for all combinations of $\sigma_L$, $\sigma_C$, and $<2$ per cent for all combinations not involving any code 5 element). However, now compute the ratio

$$\frac{n \text{ ave } x}{B} = \frac{10 \text{ ave } x}{B} = \frac{10\bar{x}}{B}, \tag{31}$$

which is a measure of the shift in the average of the delay for the complete delay line compared to the tolerance on the delay line. One finds this ratio is appreciable (i.e., about 10 per cent or more) for some of the combinations of component tolerances; namely, for $B = 5$, all combinations involving code 5 elements and the one other combination (4,4); for $B = 15$ only the combination (5,5). Thus, in general, the assumptions on normality and mean zero are approximately fulfilled; however, one should view with suspicion any solution we may get which involves one of the above mentioned combinations.*

* Note that these same combinations include all the larger values for $\beta_1$.

TABLE IV

Calculation of $C^*(\sigma)$ for $K = \$0.30$, $B = 5$ m$\mu$sec

| Code | $\sigma$ | $c(\sigma)$ | $q$ | $H(q)$ | $\alpha$ | $C^*(\sigma)$ |
|------|------|------|------|------|------|------|
| 1,1 | 0.408 | $11.00 | 1.004 | 1.001± | 0.49 | $11.30 |
| 2,1 | 0.646 | 6.00 | 1.59 | 1.30 | 0.48 | 7.30 |
| 2,2 | 0.817 | 5.50 | 2.01 | 1.56 | 0.47 | 7.50 |
| 3,1 | 1.47 | 3.00 | 3.62 | 2.68 | 0.45 | 6.35 |
| 3,2 | 1.55 | 2.50 | 3.81 | 2.82 | 0.45 | 5.60 |
| 3,3 | 2.04 | 2.30 | 5.02 | 3.64 | 0.44 | 6.45 |
| 4,2 | 2.94 | 1.50 | 7.24 | 5.23 | 0.42 | 6.20 |
| 4,3 | 3.23 | 1.30 | 7.95 | 5.70 | 0.41 | 6.05 |
| 4,4 | 4.08 | 1.20 | 10.04 | 7.5± | 0.40 | 7.35 |
| 5,4 | 6.46 | 1.10 | 15.88 | — | 0.39 | — |
| 5,5 | 8.17 | 1.05 | 20.09 | — | 0.39 | — |

The next item we want is the raw cost function, $C(\sigma)$. From the formula for $\sigma$, (28), and from the component cost versus tolerance functions, Table II-8, we can compute the cost and the variance for every combination of tolerance distributions for $L$ and $C$. These points are shown in Fig. 1; the numbers near each point are the code pairs which indicate the particular combination of distributions used to calculate each point. The points which are connected together, call them $c(\sigma)$, give the raw cost of the components (only) since this is the lower envelope of the set of all points. To get the raw cost we must add the fixed cost parameter $K$, i.e.,

$$C(\sigma) = K + c(\sigma). \tag{32}$$

The points which make up $c(\sigma)$ are tabulated in Table IV in the first three columns; the first column shows the particular combination of component distributions, the second column, $\sigma$, and the third column, $c$.

All that remains is to perform the calculations set forth in the text. We will perform these calculations carefully for one set of data in order to show the method. We will take the case $B = 5$ and $K = 0.30$; therefore,

$$q = \frac{3.89\sqrt{10}}{5} \sigma = 2.46\sigma, \tag{33}$$

and

$$C(\sigma) = 0.30 + c(\sigma). \tag{34}$$

The value of $q$ for each $\sigma$ is shown in the fourth column, Table IV. For each $q$ we find the corresponding $H(q)$ from either Table I or Fig. 2, whichever is more convenient; $H(q)$ is entered in the fifth column,

Table IV. Since the salvage value is one-half the component cost

$$\alpha = \frac{c(\sigma)/2}{C(\sigma)} = \frac{c(\sigma)/2}{0.30 + c(\sigma)}; \tag{35}$$

$\alpha$ is entered in the sixth column, Table IV. The real cost, $C^*$, is to be computed from this data. Modifying (21) to fit the form of this data, $C^*$ is given by

$$C^*(\sigma) = [(1 - \alpha)H(q) + \alpha][30 + c(\sigma)]; \tag{36}$$

$C^*$ is entered in the last column, Table IV, and is plotted in Fig. 3 as the curve marked $K = \$0.30$. From either the tabular form of $C^*$ or the graphical form of $C^*$ it is easy to see that the minimum real cost is $C^* = \$5.60$ per unit. Checking back through the calculations we see



Fig. 3 — Real cost per section of a ten-section, one-microsecond, lumped-constant delay line as a function of the standard deviation of the delay per section and for various values of the fixed cost $K$. The delay tolerances are $\pm B = \pm 5$ m$\mu$sec.

that the minimum real cost is realized for the code pair $(3,2)$, i.e., $\sigma_L = 2.89$ and $\sigma_C = 1.16$. The rejection limits, $\pm b$, are calculated from (22) and for this case

$$b = \sigma w(q) = 1.55\, w(q) = 0.72 \text{ m}\mu\text{sec} \tag{37}$$

where $w(q)$ is obtained from Table I or Fig. 2. Finally, the rejection rate, (23), is 65 per cent.

In addition to the case detailed above different values of the fixed cost were considered, as listed in Table II-9. The next variation was to change the over-all tolerance to $B = 15$ and consider the same range of values for $K$ again. In all cases the salvage value was taken as one-half of the material cost, for simplicity.

Furthermore, all of the above cases were considered using the other



Fig. 4 — Real cost per section of a ten-section, one-microsecond, lumped-constant delay line as a function of the standard deviation of the delay per section and for various values of the fixed cost $K$. The delay tolerances are $\pm B = \pm15$ m$\mu$sec.

TABLE V — RESULTS OF EXAMPLE

| Line | Risk, $\epsilon$, per cent | Fixed Cost, $K$ | Min $C^*$ | Code | Rejection Rate, per cent |
|------|------|------|------|------|------|
| | | $B = 5$ | | | |
| 1 | 0.01 | $0.30 | $5.60 | 3,2 | 65 |
| 2 | 0.01 | 1.00 | 7.60 | 3,2 | 65 |
| 3 | 0.01 | 3.00 | 10.80 | 2,1 | 23 |
| 4 | 0.01 | 5.00 | 13.40 | 2,1 | 23 |
| 5 | 0.01 | 10.00 | 19.90 | 2,1 | 23 |
| 6 | 0.01 | 15.00 | 26.02 | 1,1 | 0.1 |
| 7 | 0.0 | 0.30 | 7.45 | 3,2 | 75 |
| 8 | 0.0 | 1.00 | 10.15 | 2,1 | 44 |
| 9 | 0.0 | 3.00 | 13.70 | 2,1 | 44 |
| 10 | 0.0 | 5.00 | 17.30 | 2,1 | 44 |
| 11 | 0.0 | 10.00 | 25.45 | 1,1 | 22 |
| 12 | 0.0 | 15.00 | 31.85 | 1,1 | 22 |
| 13 | 0.01+ | min $C^* = 11.00 + K$ | | 1,1 | no-test |
| | | $B = 15$ | | | |
| 14 | 0.01 | $0.30 | $2.55 | 4,3 | 50 |
| 15 | 0.01 | 1.00 | 3.75 | 3,2 | 10 |
| 16 | 0.01 | 3.00 | 5.95 | 3,2 | 10 |
| 17 | 0.01 | 5.00 | 8.20 | 3,2 | 10 |
| 18 | 0.01 | 10.00 | 13.75 | 3,2 | 10 |
| 19 | 0.01 | 15.00 | 19.30 | 3,2 | 10 |
| 20 | 0.0 | 0.30 | 3.30 | 4,3 | 64 |
| 21 | 0.0 | 1.00 | 4.65 | 3,2 | 34 |
| 22 | 0.0 | 3.00 | 7.65 | 3,2 | 34 |
| 23 | 0.0 | 5.00 | 10.05 | 2,2 | 2 |
| 24 | 0.0 | 10.00 | 16.25 | 2,1 | 2 |
| 25 | 0.0 | 15.00 | 21.35 | 2,1 | 2 |
| 26 | $<10^{-4}$ | min $C^* = 5.50 + K$ | | 2,2 | no-test |
| 27 | 0.1 | min $C^* = 3.00 + K$ | | 3,1 | no-test |

criteria discussed previously. That is, the no-test method, (25), for $\epsilon = 0.01$ per cent, and the zero risk method, (26).

The results of the above calculations are shown in Figs. 3 and 4 and in Table V. The results of the zero-risk method were not plotted; the curves are similar to the ones shown but are shifted in a manner indicated by the shift of the minimums as recorded in Table V. Obviously, no curves of this type can be plotted for the no-test method.

VIII.  DISCUSSION OF EXAMPLE

The author must admit that this is not the best of all possible examples. Better examples of units would be the amplifiers in a long transmission line, or individual logic packages in a logic network, or gas tube crosspoints in a switching network, and so forth. However, there is a very real difficulty involved in constructing such examples; namely, it is extremely difficult, but possible, to obtain the raw cost as a function

of the variance for any but the simplest kind of unit. Therefore, in order not to get side-tracked the author has chosen a simple unit, a delay network, and asks the reader to use his imagination. In addition to asking the reader to accept some of the simplifications and their likes previously noted, and to accept the possibility of making some difficult tests (e.g., rejecting on ±0.7 mμsec, which, however, is certainly no more difficult than rejecting on ±0.5 mμsec as would be necessary in the zero-risk case), he also asks the reader to imagine some good economic reason why the completed delay lines cannot be tested to determine whether they are within the tolerance limits. For, if the completed delay lines could be tested we would have to consider another possible production process.

It is to be emphasized, however, that all the shortcomings of this example can be overcome because most of the extensions we require are not out of line with usual practices, cf., parenthetical statement in the preceding paragraph, for example. If it would make the reader any happier he can relabel the scales in Fig. 1, the raw cost curve, and pretend that he has an amplifier, for instance. The calculations will be the same from there on except for any adjustments in the salvage values that the reader cares to make.

Probably the most striking result is the size of the rejection rate. For instance, in Table V, lines 1, 2 the rejection rate is 65 per cent which is much larger than the rejection rates for production processes which are usually considered as satisfactory. The important point to realize, however, is that *under the assumptions* considered and in order to produce delay lines at the *minimum cost per usable delay line* this is the rejection rate. All rejection rates are not this high. For, as the fixed cost $K$ is increased (lines 3, 4, 5) the rejection rate decreases to 23 per cent. Of course, a rejection rate of 23 per cent is also rather high compared to the usual. It is not until the fixed cost is increased to $15 (line 6) that the rejection rate is of a usual size, about 0.1 per cent.

For the set of entries in Table III which correspond to $B = 15$ the rejection rates are usual (10 per cent for lines 2, 3, 4, 5, 6) except for line 1 (50 per cent).

For comparison we have also tabulated the minimum real costs for the zero-risk case and for the no-test case in Table V. These values were obtained under the same assumptions as were used above. But, in order to compare the real costs on a fair basis one must attach a cost figure to the risk of having one out of ten thousand delay lines out of tolerance, or to the cost of 100 per cent testing of the individual units, respectively. Note that for $B = 15$, the no-test case (lines 26, 27), we run into trouble because of component availability. The only compo-

nents which are available either make the risk much lower than desired — at a high cost — or they make the risk too high.

It is worthwhile making a point here. We assumed that we knew the raw cost at one stage in the procedure. On that assumption we obtained a minimum, or a set of minimums and near minimums, for the real cost. Also, we could obtain all the associated rejection limits. We now have some new knowledge which is important if the cost of testing is not insignificant compared to the raw cost. For, notice, it is certainly more expensive to accept or reject on, say, $0.72\sigma$ than on $2.0\sigma$. And this is a type of information which we could not have used intelligently initially but which we can use now. Hence, we can now readjust the raw cost curve and reperform the calculations.

The above idea illustrates a general principle. One need not think of this procedure, as a whole, as necessarily leading to the answer in one stroke. Rather, one should think of it as a procedure which can be applied again and again in order to converge on the answer. For, after any one application the number of combinations of component tolerances which are candidates for the production model has been reduced, and one can then afford to get more precise information on fewer possibilities in order to reapply the method.

IX.   CONCLUSIONS

We have given a method for finding the optimum tolerance assignment from the viewpoint of giving the lowest cost per acceptable unit. We have compared it with two other criteria in an idealized example and have shown that the method described is usually the best — for this example.

As was noted in the discussion of the example, the most startling result has been the generally large sizes of the rejection rates. In order to investigate this, one must remember that the rejection rate has been thrown in as another variable instead of allowing the rejection rate to be a measure of the optimization of the production method. Now a low rejection rate is not a bad measure for many production processes; the thing which these processes have in common is that the ratio of the material cost to the labor cost and other fixed costs is very small, assuming that these costs are unsalvageable.* This was the condition for some of the cases considered. And, in general, if this be the case the method we have given will also predict a low rejection rate since the raw cost will be almost a constant. However, the advantage our method has is

---

* These costs are not necessarily unsalvageable. One might often be able to salvage a good share of the labor cost by selectively assembling the rejected units.

that it will handle the other cases, too. These cases are not unusual today and with increasing automation they will become more and more the order of the day.

A word of caution is important. We have given a method of optimizing tolerances to reduce manufacturing costs under a special set of assumptions. *Our analysis and conclusions are valid only for manufacturing processes in which these assumptions prevail under actual manufacturing conditions.*

In obtaining our formulas we have considered only the special case of symmetric tolerance limits. It is obvious that the method can easily be extended to cover the case of one-sided tolerance limits and the case of unsymmetric tolerance limits. Further, we have only considered the case in which all the distributions are normal; it is not so obvious how the method can practicably be extended to the non-normal case. However, for distributions which are not violently non-normal it is not clear that what we have done is not of sufficient accuracy. Indeed, in practice, both the raw cost and salvage rate as functions of the variance of the response of the unit would be given only approximately. Further, it appears from the example that the minimum is not critical. Thus, the principal benefit of the method would be to get an approximation to the optimum. Then, after finding the correct neighborhood one could make an exhaustive cost and statistical study to determine the optimum production process.

Also, we have dealt only with the case where there is only one response per unit to be considered. The practicable extension of this work to multiple responses is not obvious. However, if one can single out one response which is more critical than the others remarks similar to the ones in the preceding paragraph about using this method to get in the correct neighborhood are in order.

X.   ACKNOWLEDGMENT

XI.   REFERENCES

1. J. M. Juran, Quality Control Handbook, McGraw-Hill, New York, 1951, pp. 75–79.
2. B.S.T.J., **32,** July, 1953. The entire issue is devoted to the L3 coaxial system.
3. H. F. Dodge, B. J. Kinsburg and M. K. Kruger, Quality Control Requirements of the L3 Coaxial Cable, B.S.T.J., **32,** pp. 943–967, July, 1953.
4. R. F. Garrett, T. L. Tuffnell and R. A. Waddell, Application of Quality Control Requirements in the Manufacture of Components of the L3 Coaxial System, B.S.T.J., **32,** pp. 969–1005, July, 1953.
5. E. W. Pike and T. R. Silverberg, Assigning Tolerances for Maximum Economy, Mach. Design, **25,** p. 139, Sept., 1953.
6. J. W. Tukey (unpublished manuscript).

# The Measurement of Power Spectra from the Point of View of Communications Engineering — Part II

### By R. B. BLACKMAN and J. W. TUKEY

(Manuscript received August 28, 1957)

*The measurement of power spectra is a problem of steadily increasing importance which appears to some to be primarily a problem in statistical estimation. Others may see it as a problem of instrumentation, recording and analysis which vitally involves the ideas of transmission theory. Actually, ideas and techniques from both fields are needed. When they are combined, they provide a basis for developing the insight necessary (i) to plan both the acquisition of adequate data and sound procedures for its reduction to meaningful estimates and (ii) to interpret these estimates correctly and usefully. This account attempts to provide and relate the necessary ideas and techniques in reasonable detail. Part I of this article appeared in the January, 1958 issue of THE BELL SYSTEM TECHNICAL JOURNAL.*

## DETAILS AND DERIVATIONS

In this part we will reconsider some of the earlier analysis, either in greater detail, or from alternative points of view. We shall assume familiarity with the material on fundamental Fourier techniques presented in Appendix A of Part I.

The sections of this part will be numbered in exact correspondence with the sections of the general account. Thus, for example, Section B.7, below, presents the details and sidelights related to Section 7, of Part I. (Certain sections will be omitted.)

### B.1 *Gaussian Processes and Moments*

There are two common modes of description of a random process, intuitively quite different. One uses the idea of an ensemble, the other a function of infinite extent. The first is undoubtedly more flexible, as it can describe processes, even non-stationary (e.g. evolving) ones, which cannot be described by any single function, even one of infinite extent. The first is also, at least in the eyes of the statistician, more fundamental, since uncertainty, which he regards as a central concept, enters directly and explicitly. It is possible to regard the single-function approach as an attempt to minimize recognition of the statistical aspects of the situation. Once, such minimization may have been of some value, but today the essentially statistical nature of communication, be it of symbols, voice, picture or feedback information, is well established. The communication engineer is aware that he must have designed not only for the message which was sent, but also for the one which might have been sent — moreover that his design demanded consideration of the relative probabilities of various messages that might have been sent (and were not).

Such a statistical view of message or noise confronts us with the need, not only of picking out what functions might arise, but also of attaching probabilities to functions (at least to sets of functions). To do this directly and completely requires much careful mathematics. As far as questions associated with observations and data are concerned, there is,

fortunately, no need for such care and complexity. We know that any empirical time function can be adequately represented by some finite number, large or small, of ordinates. Thus, for practical purposes, it suffices to be able to assign probabilities to sets of $n$ ordinates — for $n$ finite but possibly quite large. It is for this reason that we went directly to probability distributions of such $n$-dimensional sections in the general account.

This replacement of a continuous record by discrete ordinates is related to the sampling theorem of information theory. The relationship is, regrettably, not quite simple. Given a band-limited "signal" defined for all time from $-\infty$ to $+\infty$, and moderately well-behaved otherwise, the sampling theorem (Nyquist[23]), which is also known as the Cardinal Theorem of Interpolation Theory (Whittaker[24]), states that equi-spaced ordinates, if close enough together, extending from $-\infty$ to $+\infty$ will precisely determine the function. Given a band-limited function over a finite interval, the corresponding result is almost practically true. It is not true in a precise impractical sense, since every band-limited function can be obtained from an entire function of exponential type (of a complex variable) by considering only the values taken on along the real axis. Consequently, if we know a band-limited function precisely in an interval, its values are determined everywhere. Theoretically determined, but not practically so, since the kernels expressing this determination behave like hyperdirective antennas. Since the values at equi-spaced points in the interval do not determine the values at equi-spaced points outside the interval (which would be determined by precise knowledge throughout the interval) the latter cannot be obtained from the equi-spaced values in the interval. In practice, however, functions are not quite band-limited, and measurements always involve measurement noise. When these two facts are considered, a sufficiently closely spaced set of equi-spaced ordinates extracts all the practically useful information in the continuous record.

The fact that averages, variances and covariances completely characterize any $n$-dimensional Gaussian distribution is common statistical knowledge, and follows by inspection of the conventional general form of Gaussian probability density function, in which these moments appear as the only parameters.

B.2 *Autocovariance Functions and Power Spectra for Perfect Information*

If $X(t)$ is a function generated by a stationary Gaussian process, and if

$$\lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} X(t) \, dt = 0 \tag{B-2.1}$$

then the *autocovariance function* of the process is

$$C(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} X(t) \cdot X(t + \tau) \, dt. \qquad \text{(B-2.2)}$$

In particular, the *variance* of the process is

$$C(0) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} [X(t)]^2 \, dt. \qquad \text{(B-2.3)}$$

If the function $X(t)$ is passed through a fixed linear network whose impulse response (response to a unit impulse applied at $t = 0$) is $W(t)$, then the output of the network will be

$$X_{\text{out}}(t) = \int_{-\infty}^{\infty} W(\lambda) \cdot X(t - \lambda) \cdot d\lambda \,,$$

and the autocovariance of the output will be

$$C_{\text{out}}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(\lambda_1) \cdot W(\lambda_2) \cdot X(t - \lambda_1)$$

$$\cdot X(t + \tau - \lambda_2) \cdot d\lambda_1 \cdot d\lambda_2 \cdot dt$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(\lambda_1) \cdot W(\lambda_2) \cdot C(\tau + \lambda_1 - \lambda_2) \cdot d\lambda_1 \cdot d\lambda_2.$$

If we now let

$$C(\tau) = \int_{-\infty}^{\infty} P(f) e^{i\omega\tau} \, df \qquad (\omega = 2\pi f), \quad \text{(B-2.4)}$$

then

$$C_{\text{out}}(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(\lambda_1) \cdot W(\lambda_2) \cdot P(f) \cdot e^{i\omega(\tau + \lambda_1 - \lambda_2)} \cdot d\lambda_1 \cdot d\lambda_2 \cdot df.$$

But

$$\int_{-\infty}^{\infty} W(\lambda_2) \cdot e^{-i\omega\lambda_2} \cdot d\lambda_2 = Y(f)$$

is the transfer function (ratio of steady-state response to excitation, when the excitation is $e^{i\omega t}$) of the network, and

$$\int_{-\infty}^{\infty} W(\lambda_1) \cdot e^{i\omega\lambda_1} \cdot d\lambda_1 = Y(-f)$$

is the complex-conjugate of $Y(f)$. Hence,

$$C_{\text{out}}(\tau) = \int_{-\infty}^{\infty} |Y(f)|^2 \cdot P(f) \cdot e^{i\omega\tau} \cdot df. \qquad \text{(B-2.5)}$$

In particular, the variance of the output is

$$C_{\text{out}}(0) = \int_{-\infty}^{\infty} |Y(f)|^2 \cdot P(f) \cdot df. \qquad \text{(B-2.6)}$$

Since $|Y(f)|^2$ is the power transfer function of the network, it is natural to call $P(f)$ the *power spectrum* of the process. The power spectrum has the dimensions of variance per cycle per second. By equation (B-2.4) we have

$$P(f) = \int_{-\infty}^{\infty} C(\tau) \cdot e^{-i\omega\tau}\, d\tau\,, \quad (\omega = 2\pi f)\,, \quad \text{(B-2.7)}$$

and, by equation (B-2.5),

$$P_{\text{out}}(f) = |Y(f)|^2 \cdot P(f).$$

Autocovariance functions and power spectra are usually regarded as one-sided functions of lag and frequency, respectively, related by the formulae (Rice[3] p. 285),

$$P(f) = 4 \int_0^{\infty} C(\tau) \cdot \cos \omega\tau \cdot d\tau\,, \qquad \text{(not used here)},$$

$$C(\tau) = \int_0^{\infty} P(f) \cdot \cos \omega\tau \cdot df\,, \qquad \text{(not used here)}.$$

However, we will find it very convenient for analytical purposes, to continue to regard them as two-sided even functions related by equations (B-2.4) and (B-2.7). This will be evident in Section B.4 where spectral windows will be convolved with power spectra, and in Section B.6 where we would otherwise have to make use of rather complicated trigonometric identities.

In a few places where we contemplate computations, we may write

$$P(f) = 2 \int_0^{\infty} C(\tau) \cdot \cos \omega\tau \cdot d\tau\,, \qquad (\omega = 2\pi f),$$

instead of equation (B-2.7), but it is important to observe that the power spectrum $P(f)$ thus obtained must still be regarded as a two-sided even function which contains only a half of the total power or variance in the positive frequency range. If we prefer to think ultimately of a one-sided power spectrum (over only positive frequencies), in accordance with engineering practice, then we should take $2P(f)$ as the power density (per cycle per second) for positive frequencies.

The power spectrum $P(f)$ may also be expressed directly in terms of $X(t)$. The formal derivation of this expression on the basis of the formula (B-2.2) for the autocovariance function is complicated by the fact that the integral

$$\int_{-T/2}^{T/2} X(t) \cdot X(t + \tau) \, dt$$

actually depends upon two more-or-less distinct pieces of $X(t)$, one in the range $-T/2 < t < T/2$, the other in the range

$$-\frac{T}{2} - \tau < t < \frac{T}{2} - \tau.$$

We may avoid this complication by using the equivalent formula

$$C(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-\infty}^{\infty} G(t) \cdot G(t + \tau) \cdot dt, \qquad (B\text{-}2.8)$$

where

$$G(t) = X(t), \qquad |t| < \frac{T}{2},$$

$$= 0, \qquad |t| > \frac{T}{2}.$$

Let

$$S(f) = \int_{-\infty}^{\infty} G(t) \cdot e^{-i\omega t} \, dt,$$

so that

$$G(t) = \int_{-\infty}^{\infty} S(f) \cdot e^{i\omega t} \, df.$$

Then

$$C(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-\infty}^{\infty} G(t) \cdot \int_{-\infty}^{\infty} S(f) \cdot e^{i\omega(t+\tau)} \, df \cdot dt$$

$$= \lim_{T \to \infty} \frac{1}{T} \int_{-\infty}^{\infty} S(f) \cdot e^{i\omega \tau} \int_{-\infty}^{\infty} G(t) \cdot e^{i\omega t} \, dt \cdot df$$

$$= \lim_{T \to \infty} \frac{1}{T} \int_{-\infty}^{\infty} S(f) \cdot S(-f) \cdot e^{i\omega \tau} \, df.$$

Comparing with (B-2.4) we get, at least formally,

$$
\begin{aligned}
P(f) &= \lim_{T \to \infty} \frac{1}{T} \mid S(f) \mid^2 \\
&= \lim_{T \to \infty} \frac{1}{T} \left| \int_{-\infty}^{\infty} G(t) \cdot e^{-i\omega t} \, dt \right|^2 \\
&= \lim_{T \to \infty} \frac{1}{T} \left| \int_{-T/2}^{T/2} X(t) \cdot e^{-i\omega t} \, dt \right|^2 .
\end{aligned}
\tag{B-2.9}
$$

(See Rice,[3] p. 320, and Bennett,[25] p. 621.)

B.3 *The Practical Situation*

In the study of second moments of random processes, the balance between the approach through autocovariances and the approach through power spectra is, in at least one sense, a little closer than Section 3 would seem to imply. As a means for understanding, and as a guide for intelligent design, the power spectrum is without a peer. The autocovariance function is of little use except as a basis for estimating the power spectrum. This is fundamentally because, in most physical systems, power spectra have reasonable shapes, are relatively easily understandable, and often are quite directly influenced by the basic variables of the situation, whatever these may be. The process of using an empirically observed and analyzed power spectrum usually goes through some such chain of steps as this:

(i) Planning and design.

(ii) Observation and recording.

(iii) Analysis and preparation of estimates.

(iv) Comparison of estimates with existing and synthesizable theoretical structures and quantitative information.

(v) Selection of the best working version of a theoretically-guided approximation to the estimates.

(vi) Use of this working version.

Our theoretical understanding of the situation, and of the forces in it, play important roles, which we should never allow ourselves to forget, in steps (i) and (v). (We use this understanding to the utmost, in its proper place, but we do not, and should not, allow it to narrow down steps (ii) and (iii) to the point where we have little or no chance of discovering that it was incomplete or in error. Thus, we estimate a considerable number of smoothed spectral densities, and not merely a few constants of a suggested theoretical curve. After we have compared curve and points, we may then wish to estimate the constants.)

The simplest and most straightforward use we can make of the power spectrum is to predict the output spectrum, or perhaps only the output power, when the process studied provides the input to a linear device with a known power transfer function.

Except for purely descriptive uses, checking on performance for agreement with anticipation, and for predicting the behavior of already designed linear systems, the most elementary use of a power spectrum lies in optimizing the performance of some linear predictor or filter as measured á la least squares. The nature of this situation is not quite that one which most persons imagine.

If we really have no theoretical insight into the situation *at all*, we might as well (nay, perhaps, might better) stay in the time domain. We have autocovariances (say) for some limited range of lags. If the duration of the transient response of our filter or predictor is not going to exceed one-half this time limit, then we can write out the estimated variance of any predictor or filter directly in terms of our estimates of autocovariances and of the time description of the filter or predictor, and could then minimize this directly. With no theoretical insight this should work at least as well as any other way. With no theoretical insight, analysis in the time domain would be relatively good, perhaps even optimal—and probably absolutely poor.

But we do not, and almost always should not, optimize filters or predictors in this way. The reason is simple. Actual power spectra are often simple and understandable. Actual autocovariance functions are hardly, if ever, simple and understandable. The intervention of theoretical insight and human judgment at step (v) is crucial and valuable. This intervention is effective in the frequency domain, but not in the time domain. (Step (v) is likely to stand for some time, as a challenge to the ability of statisticians to wisely and effectively automatize inferential procedures.)

An additional advantage of power spectrum analysis over autocovariance analysis was pointed out in Section 3, namely the ease of compensation for (linear) modification before measurement. When the random function $X(t)$ is passed through a time-invariant linear transmission system whose impulse response is $W(t)$, the output random function, which may be the only function accessible for measurement, is

$$X_{\text{out}}(t) = \int_{-\infty}^{\infty} W(\tau_1) \cdot X(t - \tau_1) \, d\tau_1.$$

The relation between the autocovariance $C_{\text{out}}(\tau)$ of the modified process,

and the autocovariance $C(\tau)$ of the original process may be derived as follows:

$$
\begin{aligned}
C_{\text{out}}(\tau) &= \text{ave} \ \{X_{\text{out}}(t) \cdot X_{\text{out}}(t + \tau)\} \\
&= \text{ave} \left\{ \iint\limits_{-\infty}^{\infty} W(\tau_1) \cdot X(t - \tau_1) \cdot X(t + \tau - \tau_2) \cdot W(\tau_2) \ d\tau_1 \ d\tau_2 \right\} \\
&= \iint\limits_{-\infty}^{\infty} W(\tau_1) \cdot C(\tau + \tau_1 - \tau_2) \cdot W(\tau_2) \ d\tau_1 \ d\tau_2 .
\end{aligned}
$$

Putting $\tau_2 = \tau_1 + \lambda$, we get

$$
\begin{aligned}
C_{\text{out}}(\tau) &= \int_{-\infty}^{\infty} C(\tau - \lambda) \cdot \left[ \int_{-\infty}^{\infty} W(\tau_1) \cdot W(\tau_1 + \lambda) \ d\tau_1 \right] d\lambda \\
&= C(\tau) * W(\tau) * W(-\tau).
\end{aligned}
$$

Measurement of $X_{\text{out}}(t)$ can give estimates of $C_{\text{out}}(\tau)$ which must subsequently be converted into estimates of $C(\tau)$. The only practical way to make this conversion seems to be through Fourier transformation of the estimates of $C_{\text{out}}(\tau)$ into the frequency domain, compensation there, and Fourier retransformation. Such a procedure, in effect, invokes the relation between the power spectrum $P_{\text{out}}(f)$ of the modified process and the power spectrum $P(f)$ of the original process. This relation is

$$
P_{\text{out}}(f) = P(f) \cdot |\ Y(f)\ |^2,
$$

where $Y(f)$ is the transfer function corresponding to the impulse response $W(t)$.

## DETAILS FOR CONTINUOUS ANALYSIS

B.4 *Power Spectrum Estimation from a Continuous Record of Finite Length*

In the ideal case considered in Section B.2, which assumes that we have an infinite length of $X(t)$, we can calculate the power spectrum $P(f)$ in two ways — either directly from $X(t)$, or indirectly as the Fourier transform of the autocovariance function $C(\tau)$ which is calculable directly from $X(t)$. The basic choice is, leaving limiting problems aside, between squaring a Fourier transform, or Fourier transforming an average of products. In either case multiplication and Fourier transformation must enter.

From the point of view of the ensemble, as opposed to the single time function, we seek to estimate a particular basis for the second moments

of all linear combinations. We may estimate any convenient basis as a start, and then transform.

Clearly, from either point of view, any result obtained in one way can also be obtained in the other. Differences between a time approach and a frequency approach must be differences in (i) ease of understanding, (ii) ease of manipulation of formulas, (iii) ease of calculation with numbers, rather than in anything more essential.

Understanding and a simple description of the procedure which yields reasonably stable estimates, and which we have discussed in general, is more easily obtained by the indirect route, so we shall proceed accordingly, beginning with a general outline of a hypothetical procedure for power spectrum estimation from a continuous record of finite length.

A general outline of a hypothetical procedure for power spectrum estimation from a continuous record of finite length, specifically $X(t)$ for $-T_n/2 \leq t \leq T_n/2$, is as follows:

(1) *Calculate the apparent autocovariance function*

$$C_{00}(\tau) = \frac{1}{T_n - |\tau|} \int_{-(T_n-|\tau|)/2}^{(T_n-|\tau|)/2} X\left(t - \frac{\tau}{2}\right) \cdot X\left(t + \frac{\tau}{2}\right) dt \quad \text{(B-4.1)}$$

for $|\tau| \leq T_m < T_n$, where $T_n$ is the length of the record, and $T_m$ is the maximum lag to be used. We shall see in Section B.9 that the stability of our power spectrum estimates depends upon how small we take the ratio $T_m/T_n$.

(For the purpose of the theoretical analysis in this section we assume that the data contain no errors of measurement; in particular, no bias due to a displaced (perhaps drifting) zero. The effects of such errors are considered elsewhere.)

(2) *Calculate the modified apparent autocovariance function*

$$C_i(\tau) = D_i(\tau) \cdot C_{00}(\tau), \quad \text{(B-4.2)}$$

where $D_i(\tau)$ is a prescribed lag window, an even function such that

$$D_i(0) = 1,$$

and

$$D_i(\tau) = 0 \quad \text{for} \quad |\tau| > T_m.$$

Note that $C_i(\tau) = 0$ for $|\tau| > T_m$ although $C_{00}(\tau)$ is not available for $|\tau| > T_m$.

(3) *Calculate the estimated power spectrum*

$$P_i(f) = 2\int_0^\infty C_i(\tau) \cdot \cos \omega\tau \cdot d\tau. \quad \text{(B-4.3)}$$

Our object is now to determine the relation between ave $\{P_i(f)\}$ and $P(f)$, where "ave" denotes the ensemble average, that is, the average over all possible continuous pieces of $X(t)$ of length $T_n$. (The variability (specifically, the variance) of $P_i(f)$ will be examined in Section B.9.) Since $C_{00}(\tau)$ is not calculated for $|\tau| > T_m$, it is clear that

$$\text{ave } \{C_{00}(\tau)\} = C(\tau), \qquad \text{only for } |\tau| \le T_m .$$

However, because $D_i(\tau) = 0$ for $|\tau| > T_m$,

$$\text{ave } \{C_i(\tau)\} = D_i(\tau) \cdot C(\tau), \qquad \text{for any } \tau.$$

Hence,

$$\text{ave}\{P_i(f)\} = \int_{-\infty}^{\infty} D_i(\tau) \cdot C(\tau) \, e^{-i\omega\tau} \, d\tau.$$

Then, if $Q_i(f)$ is the Fourier transform of $D_i(\tau)$, the relation we seek is, symbolically,

$$\text{ave } \{P_i(f)\} = Q_i(f) * P(f),$$

or explicitly,

$$\text{ave } \{P_i(f_1)\} = \int_{-\infty}^{\infty} Q_i(f_1 - f) \cdot P(f) \, df . \tag{B-4.4}$$

This relation is in a form, (B-2.6), which is familiar to communications engineers except for the fact that $Q_i(f_1 - f)$ is not an even function of $f$, when $f_1 \ne 0$, and may be negative in some ranges of $f$. However, taking advantage of the fact that $P(f)$ is an even function of $f$, we may write

$$\text{ave}\{P_i(f_1)\} = \int_{0}^{\infty} H_i(f; f_1) \cdot P(f) \, df , \tag{B-4.5}$$

where

$$H_i(f; f_1) = Q_i(f + f_1) + Q_i(f - f_1). \tag{B-4.6}$$

The function $H_i(f; f_1)$ is an even function of $f$ as well as of $f_1$. Hence it satisfies one of the necessary conditions for a physically realizable *power transfer function*. However, inasmuch as it may be negative in some ranges of $f$ (actually an advantage as we will see), it may still not be physically realizable. Nevertheless, it is convenient to regard $H_i(f; f_1)$ as the power transfer function of a network, and to regard ave $\{P_i(f_1)\}$ as the long-time-average power output of the network when continuously driven by the random process.

Since $P_i(f_1)$, whether calculated from a single piece of $X(t)$ as outlined

above, or calculated as an average over a finite number of pieces of $X(t)$, is an estimate of ave $\{P_i(f_1)\}$ in the usual statistical sense, it is evident that the calculated power density $P_i(f_1)$ is an estimate of an average-over-frequency of the true power spectrum $P(f)$, and not an estimate of the local power density $P(f_1)$. The calculated power density $P_i(f_1)$ may be regarded as an estimate of the local power density $P(f_1)$ only to the degree to which $Q_i(f)$ approximates $\delta(f)$. However, under the restriction that $D_i(\tau) = 0$ for $|\tau| > T_m$, the degree to which $Q_i(f)$ approximates $\delta(f)$ depends chiefly on how large we take $T_m$. On the other hand, as we will find in Section B.9, the larger we take $T_m/T_n$ the less stable will the estimates be. Hence, in general, it will be wasteful to demand more frequency resolution than we actually need. In many cases we may even have to take less frequency resolution than we would like to have, in order to secure a reasonable stability of the estimates. Clearly, for any specific value of $T_n$, (and number of pieces of record), we can increase frequency resolution (or stability) only by sacrificing stability (or frequency resolution).

We have just examined a hypothetical method of power spectrum estimation, in which we compute an apparent autocovariance function, modify it, and take the cosine transform. We will now examine a method, also hypothetical, in which we modify the data, take the sine and cosine transforms, compute the sum of the squares at each frequency, and divide by the length of the record. If the data is $X(t)$ for $0 < t < T_n$, and the weighting function (data window) is $B_i(t)$, the estimated power spectrum is computed essentially according to the formula

$$P_{ei}(f) = \frac{1}{T_n} \left| \int_0^{T_n} B_i(t) \cdot X(t) \cdot e^{-i\omega\tau} \, dt \right|^2. \tag{B-4.7}$$

To determine the average it is convenient to assume that $X(t)$ is of unlimited extent, to specify $B_i(t)$ to be identically zero for $t < 0$ and $t > T_n$, and to allow the *data window* $B_i(t)$ to be located anywhere in time by substituting $B_i(t - \lambda)$ for $B_i(t)$. Then

$$P_{ei}(f;\lambda) = \frac{1}{T_n} \left| \int_{-\infty}^{\infty} B_i(t - \lambda) \cdot X(t) \cdot e^{-i\omega\tau} \, dt \right|^2. \tag{B-4.8}$$

This is the Fourier transform of

$$C_{ei}(\tau;\lambda) = \frac{1}{T_n} \int_{-\infty}^{\infty} B_i(t - \lambda) \cdot X(t) \cdot B_i(t - \lambda + \tau) \cdot X(t + \tau) \cdot dt. \tag{B-4.9}$$

[Compare with (B-2.8) and (B-2.9)]. Now, since the random process is

stationary, the ensemble average is equivalent to the average over $\lambda$; that is,

$$\text{ave } \{C_{ei}(\tau)\} = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} C_{ei}(\tau; \lambda) \cdot d\lambda.$$

Substituting (B-4.9) and $\lambda = t - \xi$ into the right-hand member, we get

$$\text{ave } \{C_{ei}(\tau)\} = \frac{1}{T_n} \cdot \lim_{T \to \infty} \frac{1}{T} \int_{-\infty}^{\infty} X(t) \cdot X(t + \tau)$$

$$\cdot \left[ \int_{t-(T/2)}^{t+(T/2)} B_i(\xi) \cdot B_i(\xi + \tau) \cdot d\xi \right] \cdot dt.$$

Reversing the order of integration we get

$$\text{ave } \{C_{ei}(\tau)\} = \frac{1}{T_n} \int_{-\infty}^{\infty} B_i(\xi) \cdot B_i(\xi + \tau)$$

$$\cdot \left[ \lim_{T \to \infty} \frac{1}{T} \int_{\xi-(T/2)}^{\xi+(T/2)} X(t) \cdot X(t + \tau) \cdot dt \right] d\xi.$$

The quantity in brackets is the true autocovariance function $C(\tau)$. Hence,

$$\text{ave } \{C_{ei}(\tau)\} = D_{ei}(\tau) \cdot C(\tau) \tag{B-4.10}$$

where

$$D_{ei}(\tau) = \frac{1}{T_n} \int_{-\infty}^{\infty} B_i(\xi) \cdot B_i(\xi + \tau) \cdot d\xi \tag{B-4.11}$$

is the *lag window equivalent* of the data window $B_i(t)$. Therefore,

$$\text{ave } \{P_{ei}(f)\} = Q_{ei}(f) * P(f) \tag{B-4.12}$$

where $Q_{ei}(f)$ is the *spectral window* corresponding to (i.e. the Fourier transform of) the lag window $D_{ei}(\tau)$.

If $J_i(f)$ is the *frequency window* corresponding to (i.e. the Fourier transform of) the data window $B_i(t)$, then

$$Q_{ei}(f) = \frac{1}{T_n} \mid J_i(f) \mid^2. \tag{B-4.13}$$

(It will be noted that (B-4.10) can be obtained more directly from (B-4.9), by taking the ensemble average of the right-hand member of (B-4.9). In this case, $\lambda$ is superfluous and need not have been introduced at the start.)

These formulas for the "direct" method, where Fourier transforma-

tion (here modified by the data window) precedes multiplication (here squaring) can be used in several ways. Let us consider three possibilities:

(a) we may choose $B_i(t)$ non-zero over as long a range as possible, so that $D_{ei}(\tau)$ will be broad and $Q_{ei}(f)$ narrow,

(b) we may choose $B_i(t)$ non-zero over a much shorter range, making $D_{ei}(\tau)$ not so broad and $Q_{ei}(f)$ not so narrow,

(c) we may choose a number of such short windows, use each for $B_i(t)$ and average the corresponding values of $P_{ei}(f)$ so obtained into a general average.

If we follow choice (a), our result will behave similarly to those obtained from the indirect method when we try to make $Q_i(f)$ like $\delta(f)$. We shall estimate an average over a very short frequency interval, and our estimate will be excessively variable.

If we follow choice (b), we shall estimate an average over a wider frequency interval, but our estimate will remain just as variable.

If we follow choice (c) our estimate will refer to the same sort of smoothing as in (b), but we shall gain increased stability for the estimate. The behavior of the estimate will resemble that of a reasonable estimate by the indirect route.

Finally, there is another way in which we can apply the formulas for the direct route. We may use a long data window, calculate many values of $P_{ei}(f)$, and then average these results over moderately wide frequency intervals. Again our estimates will be estimates of considerably smoothed spectral densities; again our estimates will be moderately stable.

Of all these, the simplest description of an estimate which is moderately stable, and must, consequently, be an estimate of an at least moderately smoothed spectral density, is the indirect route. We shall stick to the indirect route for the present.

B.5 *Particular Pairs of Windows*

In this section we will consider five pairs of windows. They are illustrated in Figs. 1, 14, and 15. We begin with the
*Zeroth Pair*

$$D_0(\tau) = 1, \qquad |\tau| < T_m,$$
$$= 0, \qquad |\tau| > T_m,$$

and

$$Q_0(f) = 2T_m \frac{\sin 2\pi f T_m}{2\pi f T_m} = 2T_m \text{ dif } 2fT_m.$$

Notice that $C_0(\tau) = D_0(\tau) \cdot C_{00}(\tau)$ coincides with $C_{00}(\tau)$ wherever $C_{00}(\tau)$
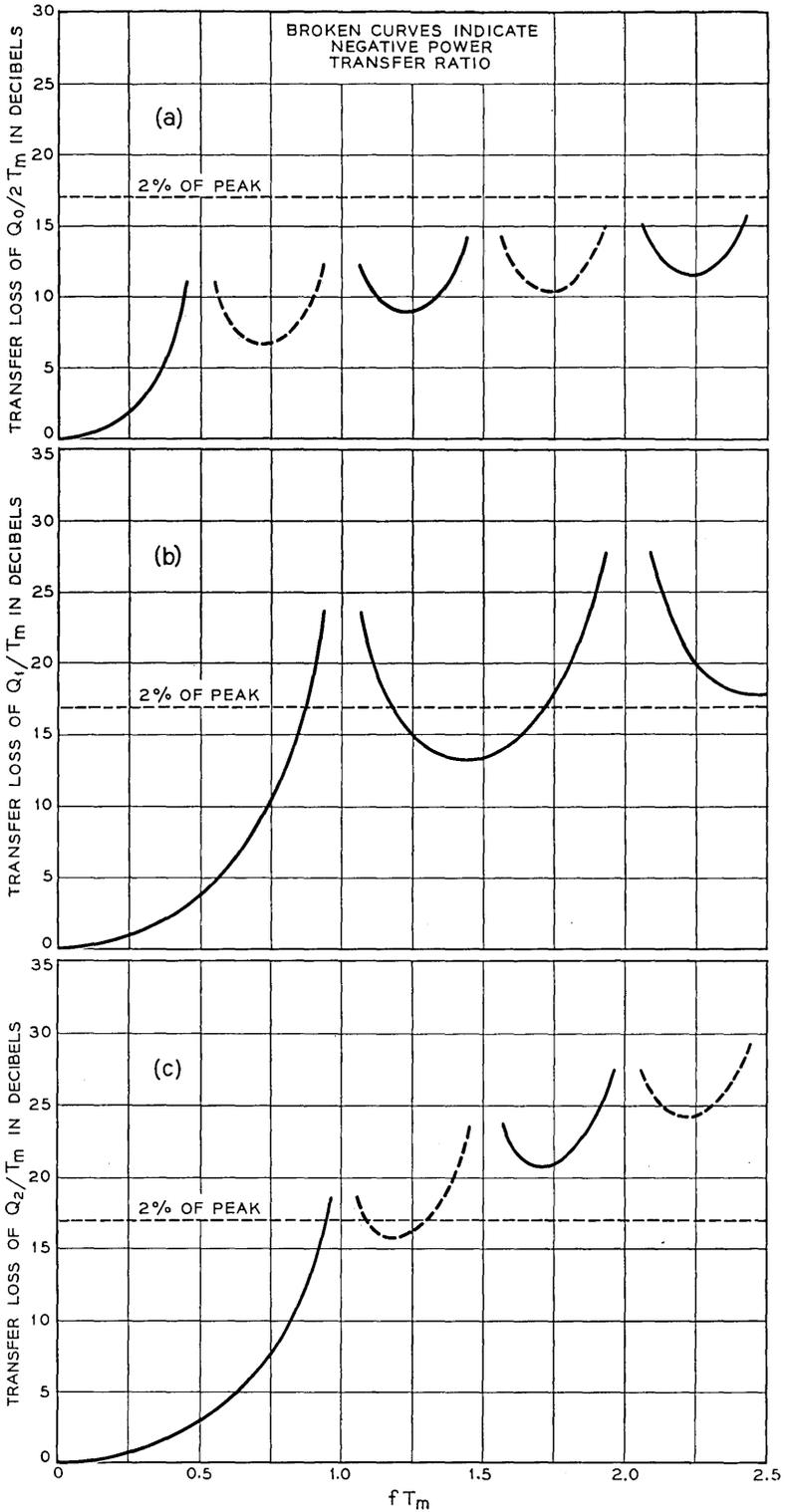
Fig. 15(a)(b)(c) — Transfer loss of spectral windows $Q_0$, $Q_1$, $Q_2$ in db.
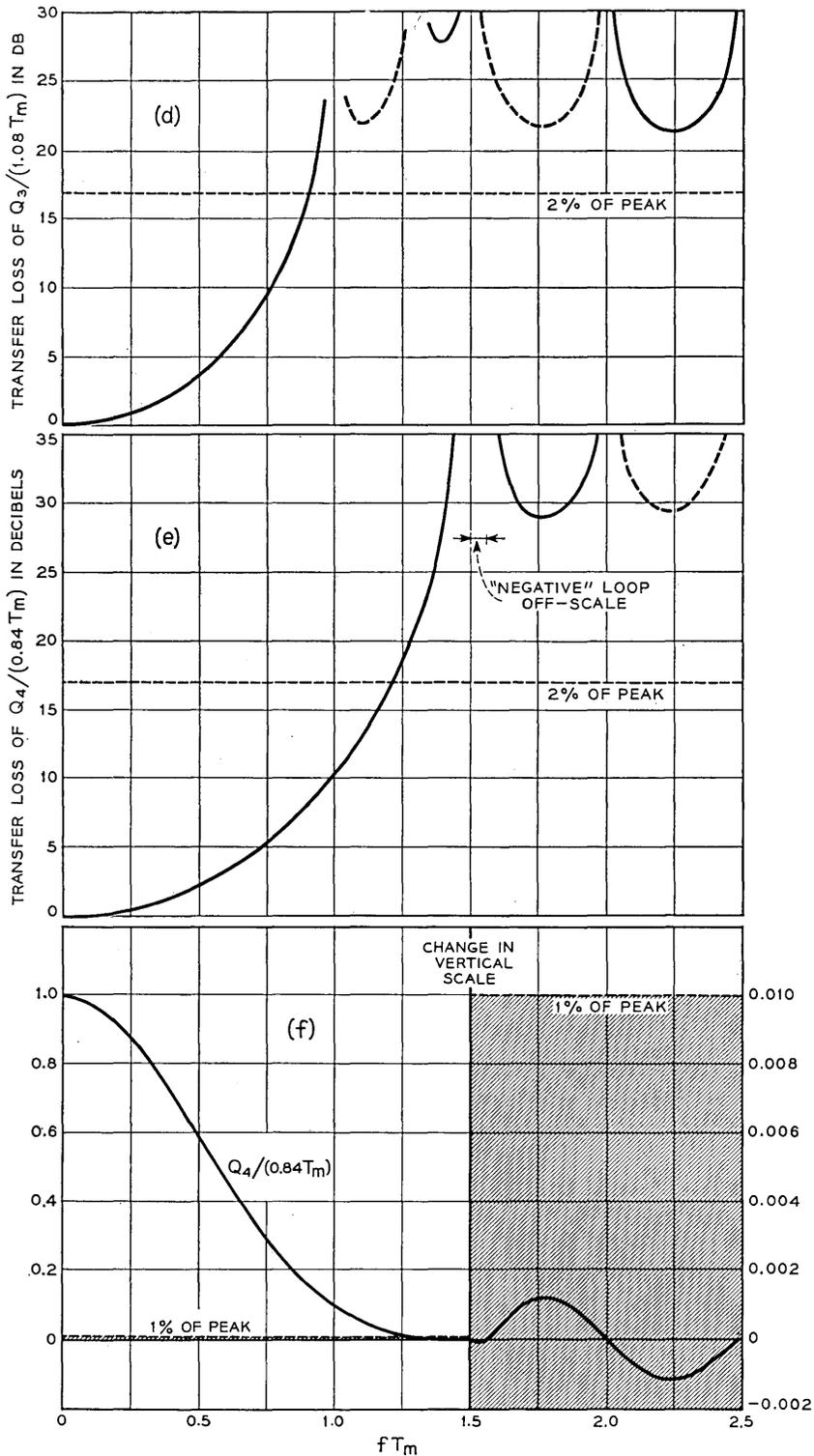
Fig. 15(d)(e)(f) — Transfer loss of spectral windows $Q_3$, $Q_4$ in db and power transfer ratio of spectral window $Q_4$.

is defined, and vanishes elsewhere. This is a "do-nothing" pair. However, $Q_0(f)$ is neither simple nor well-behaved, the first side lobe on each side of the main lobe being about $\frac{1}{5}$ the height of the main lobe (and negative).

The specifications of the other four pairs are:

*First Pair* (Bartlett[5])

$$D_1(\tau) = 1 - \frac{|\tau|}{T_m}, \qquad |\tau| < T_m,$$

$$= 0, \qquad\qquad |\tau| > T_m,$$

and

$$Q_1(f) = T_m \left(\frac{\sin \pi f T_m}{\pi f T_m}\right)^2.$$

*Second Pair* (sometimes called "hanning", after the Austrian meteorologist Julius von Hann)

$$D_2(\tau) = \frac{1}{2}\left(1 + \cos\frac{\pi\tau}{T_m}\right), \qquad |\tau| < T_m,$$

$$= 0, \qquad\qquad |\tau| > T_m,$$

and

$$Q_2(f) = \frac{1}{2} Q_0(f) + \frac{1}{4}\left[Q_0\left(f + \frac{1}{2T_m}\right) + Q_0\left(f - \frac{1}{2T_m}\right)\right].$$

*Third Pair* (sometimes called "hamming", after R. W. Hamming[26])

$$D_3(\tau) = 0.54 + 0.46 \cos\frac{\pi\tau}{T_m}, \qquad |\tau| < T_m,$$

$$= 0, \qquad\qquad |\tau| > T_m,$$

and

$$Q_3(f) = 0.54 Q_0(f) + 0.23 \left[Q_0\left(f + \frac{1}{2T_m}\right) + Q_0\left(f - \frac{1}{2T_m}\right)\right].$$

*Fourth Pair* (RBB's not very serious proposal)

$$D_4(\tau) = 0.42 + 0.50 \cos\frac{\pi\tau}{T_m} + 0.08 \cos\frac{2\pi\tau}{T_m}, \qquad |\tau| < T_m,$$

$$= 0, \qquad\qquad |\tau| > T_m$$

and

$$Q_4(f) = 0.42Q_0(f) + 0.25\left[Q_0\left(f + \frac{1}{2T_m}\right) + Q_0\left(f - \frac{1}{2T_m}\right)\right]$$
$$+ 0.04\left[Q_0\left(f + \frac{1}{T_m}\right) + Q_0\left(f - \frac{1}{T_m}\right)\right].$$

These specifications are all special cases of

$$D_1(\tau) = a_{i0} + 2\sum_{j=1}^{\infty} a_{ij}\cos\frac{j\pi\tau}{T_m}, \qquad |\tau| < T_m,$$
$$= 0, \qquad\qquad |\tau| > T_m,$$

with

$$a_{i0} + 2\sum_{j=1}^{\infty} a_{ij} = 1,$$

whence,

$$Q_i(f) = a_{i0}Q_0(f) + \sum_{j=1}^{\infty} a_{ij}\left[Q_0\left(f + \frac{j}{2T_m}\right) + Q_0\left(f - \frac{j}{2T_m}\right)\right].$$

The coefficients in $D_3(\tau)$ may be regarded as convenient approximations to

$$a_{30} = \frac{25}{46}, \qquad a_{31} = \frac{21}{92},$$

which would have produced a zero of $Q_3(f)$ at $|f| = 1.25/T_m$, with other zeros occurring at all integral multiples of $0.5/T_m$ except at 0 and $\pm 0.5/T_m$. (The zero which could have been produced at $|f| = 1.25/T_m$ actually occurs at approximately $|f| = 1.3/T_m$.) The coefficients in $D_3(\tau)$ were actually selected to minimize the height of the highest side lobe.

The coefficients in $D_4(\tau)$ are convenient approximations to

$$a_{40} = \frac{3969}{9304}, \qquad a_{41} = \frac{1155}{4652}, \qquad a_{42} = \frac{715}{18608},$$

which would have produced zeros of $Q_4(f)$ at $|f| = 1.75/T_m$ and $|f| = 2.25/T_m$, with other zeros occurring at all integral multiples of $0.5/T_m$ except at 0, $\pm 0.5/T_m$, and $\pm 1/T_m$.

In view of the fact that

$$D_i(\tau) \cdot D_0(\tau) = D_i(\tau),$$

we have

$$Q_i(f) * Q_0(f) = Q_i(f),$$

relations which may be of interest, as may the fact that using $D_i(\tau) \cdot D_j(\tau)$ for the lag window corresponds in general, to a spectral window $Q_i(f) * Q_j(f)$.

The writers have spent considerable time and effort inquiring into other possible window pairs. One of the most promising approaches was that of Čebyšëv or Chebyshëv polynomials (see Dolph[27]) to obtain side lobes of equal height. Their present conviction is that: (i) special windows *cannot* eliminate the need for prewhitening and rejection filtration and (ii) good prewhitening and rejection filtration *can* eliminate the need for special windows. Accordingly they do not recommend expending extensive effort on special windows.

Readers familiar with physical optics will recognize the close relation between the considerations of this section and diffraction by slits of uniform ($i = 0$) or varying ($i = 1, 2, 3, 4$) width. The literature on apodization (Jaquinot,[28] Boughon, Dossier, Jaquinot[29]) is relevant.

B.6 *Covariability of Power Density Estimates — Basic Result*

To derive a formula for the covariance of two power density estimates $P_i(f_1)$, $P_j(f_2)$ obtained from the same record, we will first derive a formula for the covariance of $M(t_1, \tau_1)$, $M(t_2, \tau_2)$ where

$$M(t, \tau) = X\left(t - \frac{\tau}{2}\right) \cdot X\left(t + \frac{\tau}{2}\right).$$

For this we will use a formula for

$$\text{cov } \{wx, yz\} \equiv \text{ave } \{wxyz\} - \text{ave } \{wx\} \text{ ave } \{yz\}$$

dating back to Isserlis,[30] and used by Hotelling[31] in a similar connection. If $w$, $x$, $y$, $z$ are joint Gaussian variates with zero averages, then

$$\text{cov } \{wx, yz\} = \text{ave } \{wy\} \cdot \text{ave } \{xz\} + \text{ave } \{wz\} \cdot \text{ave } \{xy\}.$$

(This result is easily derived from the "characteristic function".) If we take

$$w = X\left(t_1 - \frac{\tau_1}{2}\right), \qquad x = X\left(t_1 + \frac{\tau_1}{2}\right),$$

$$y = X\left(t_2 - \frac{\tau_2}{2}\right), \qquad z = X\left(t_2 + \frac{\tau_2}{2}\right),$$

and make use of

$$\text{ave} \left\{ X \left( t - \frac{\tau}{2} \right) \cdot X \left( t + \frac{\tau}{2} \right) \right\} = C(\tau) = \int_{-\infty}^{\infty} P(f) \cdot e^{i\omega\tau} \, df$$

we get

$$\text{cov} \left\{ M(t_1, \tau_1), M(t_2, \tau_2) \right\} =$$

$$\iint_{-\infty}^{\infty} [e^{i(\omega_1 - \omega_2)(\tau_1 - \tau_2)/2} + e^{i(\omega_1 - \omega_2)(\tau_1 + \tau_2)/2}] \qquad \text{(B-6.1)}$$

$$\cdot e^{-i(\omega_1 + \omega_2)(t_1 - t_2)} \cdot P(f_1) \cdot P(f_2) \cdot df_1 \cdot df_2.$$

It may be noted that

$$\text{var} \left\{ [X(t)]^2 \right\} = \text{cov} \left\{ M(t, 0), M(t, 0) \right\} = 2 \left[ \int_{-\infty}^{\infty} P(f) \cdot df \right]^2,$$

while

$$\text{ave} \left\{ [X(t)]^2 \right\} = \int_{-\infty}^{\infty} P(f) \cdot df.$$

Hence,

$$\frac{2 \cdot [\text{ave} \{ [X(t)]^2 \}]^2}{\text{var} \{ [X(t)]^2 \}} = 1,$$

in accordance with the fact that $[X(t)]^2$ is a constant multiple of a chi-square variate with one degree of freedom.

Substituting $f_1 = f' + f$, $f_2 = f' - f$, and noting that $df_1 \cdot df_2 = 2df \cdot df'$, we get

$$\text{cov} \left\{ M(t_1, \tau_1), M(t_2, \tau_2) \right\} =$$

$$\frac{1}{2} \int_{-\infty}^{\infty} [e^{i\omega(\tau_1 - \tau_2)} + e^{i\omega(\tau_1 + \tau_2)}] \cdot \Phi(f, t_1 - t_2) \cdot df, \quad \text{(B-6.2)}$$

where

$$\Phi(f, \lambda) = 4 \int_{-\infty}^{\infty} P(f' + f) \cdot P(f' - f) \cdot e^{-i2\omega'\lambda} \, df'. \qquad \text{(B-6.3)}$$

Since $\Phi(f, \lambda)$ is an even function of $f$ we may replace $e^{i\omega(\tau_1 \pm \tau_2)}$ by $\cos \omega(\tau_1 \pm \tau_2)$, and since $P(f' + f) \cdot P(f' - f) = P(f + f') \cdot P(f - f')$ is an

even function of $f'$ we may replace $e^{-i2\omega'\lambda}$ by $\cos 2\omega'\lambda$. Hence, we have

$$\text{cov } \{M(t_1, \tau_1), M(t_2, \tau_2)\} =$$

$$\int_{-\infty}^{\infty} \cos \omega\tau_1 \cdot \cos \omega\tau_2 \cdot \Phi(f, t_1 - t_2) \cdot df, \quad \text{(B-6.4)}$$

where

$$\Phi(f, \lambda) = 4 \int_{-\infty}^{\infty} P(f + f') \cdot P(f - f') \cdot \cos 2\omega'\lambda \cdot df'. \quad \text{(B-6.5)}$$

The next step is to determine the covariance between $C_0(\tau_1)$ and $C_0(\tau_2)$. By definition, and for hypothetical computation,

$$C_0(\tau) = \frac{1}{T_n - |\tau|} \int_{-(T_n - |\tau|)/2}^{(T_n - |\tau|)/2} M(t, \tau) \cdot dt, \quad \text{(B-6.6)}$$

but this is inconvenient for present purposes on account of the dependence of both the limits of integration and the divisor upon $\tau$. A more convenient form would be

$$C_0'(\tau) = \frac{1}{T_n'} \int_{-T_n'/2}^{T_n'/2} M(t, \tau) \cdot dt, \quad \text{(B-6.7)}$$

where $T_n' \leqq T_n - T_m$. This form could actually be used for computation, but it would not make the maximum possible use of the data. The range of integration in $C_0'(\tau)$ is less than it is in $C_0(\tau)$ for any $\tau$ except possibly $|\tau| = T_m$. A good approximation to the use of $C_0(\tau)$ for computation is to regard this as (approximately) equivalent to the use of a hypothetical, modified $C_0'(\tau)$ with $T_n - T_m < T_n' < T_n$. We will take

$$T_n' = T_n - \alpha_i T_m, \quad \text{(B-6.8)}$$

where $\alpha_i$ depends upon the lag window to be used. Since, for each value of $\tau$, the range of integration in formula (B-6.6) suffers a loss of $\tau$ out of $T_n$, and since the seriousness of this loss depends on the value of $D_i(\tau)$, it seems to be a reasonable approximation to take

$$\alpha_i = \frac{1}{T_m} \frac{\int_0^{T_m} \tau \cdot D_i(\tau) \cdot d\tau}{\int_0^{T_m} D_i(\tau) \cdot d\tau}, \quad \text{(B-6.9)}$$

which yields, for the first four lag windows described in Section B.5, $\alpha_0 = 0.50$, $\alpha_1 = 0.33$, $\alpha_2 = 0.30$, $\alpha_3 = 0.33$.

Using the approximation described in the preceding paragraph, we find (omitting the prime in $C_0'$)

$$\text{cov}\,\{C_0(\tau_1),\,C_0(\tau_2)\} = \int_{-\infty}^{\infty} \cos\omega\tau_1 \cdot \cos\omega\tau_2 \cdot \Gamma(f)\;df \quad \text{(B-6.10)}$$

where

$$\Gamma(f) = \frac{1}{(T_n')^2} \iint\limits_{-T_n'/2}^{T_n'/2} \Phi(f,\,t_1 - t_2) \cdot dt_1 \cdot dt_2$$

$$= 4 \int_{-\infty}^{\infty} P(f + f') \cdot P(f - f) \cdot \left(\frac{\sin\omega' T_n'}{\omega' T_n'}\right)^2 \cdot df'. \tag{B-6.11}$$

The final step is to determine the covariance of $P_i(f_1)$ with $P_j(f_2)$, recalling that, for example,

$$P_i(f) = \int_{-\infty}^{\infty} C_i(\tau) \cdot \cos\omega\tau \cdot d\tau,$$

where $C_i(\tau) = D_i(\tau) \cdot C_0(\tau)$. We get

$$\text{cov}\,\{P_i(f_1),\,P_j(f_2)\} = \frac{1}{4} \int_{-\infty}^{\infty} H_i(f;\,f_1) \cdot H_j(f;\,f_2) \cdot \Gamma(f) \cdot df, \quad \text{(B-6.12)}$$

where

$$H_i(f;\,f_1) = 2 \int_{-\infty}^{\infty} D_i(\tau_1) \cdot \cos\omega\tau_1 \cdot \cos\omega_1\tau_1 \cdot d\tau_1$$

$$= Q_i(f + f_1) + Q_i(f - f_1) \tag{B-6.13}$$

with a similar formula for $H_j(f;\,f_2)$. In particular, of course,

$$\text{var}\,\{P_i(f_1)\} = \frac{1}{4} \int_{-\infty}^{\infty} [H_i(f;\,f_1)]^2 \cdot \Gamma(f) \cdot df. \tag{B-6.14}$$

The *power-variance spectrum* is given by (B-6.11), and involves the true power spectrum in an essential way, as we would expect. Together with (B-6.12) and (B-6.14), this is the basic result. It is exact for estimates based on $C_0'(\tau)$, approximate for estimates based on $C_0(\tau)$.

### B.7 *Covariability of Estimates — Various Approximations*

We now need to obtain a variety of approximate forms of the basic result, suitable for use under different conditions. We shall, in turn; (i) assume that the typical frequency scale of $H_i(f;\,f_1)$ and $H_j(f;\,f_2)$ is much

larger than $1/T'_n$, (ii) assume that $P(f)$ varies slowly enough for the *quadratic* term in its Taylor series to be neglected at distances up to, say, $3/T'_n$, (iii) assume both (i) and an assumption similar to (ii) about $P_{i1}(f)$ and $P_{j2}(f)$, and, finally, (iv) assume only that $P(f)$ is concentrated in a sharp peak, narrow in comparison with $1/T'_n$.

Combining formulas (B-6.11) and (B-6.12) just obtained, we find

$$\text{cov}\ \{P_i(f_1),\ P_j(f_2)\}\ =$$

$$\iint\limits_{-\infty}^{\infty} H_i(f;\ f_1)\cdot H_j(f;\ f_2)\cdot P(f + f')\cdot P(f - f') \tag{B-7.1}$$

$$\left(\frac{\sin \omega' T'_n}{\omega' T'_n}\right)^2\cdot df\cdot df'.$$

The last factor in the integrand becomes rapidly negligible for $|f'|$ greater than, say, $1/T'_n$. On the other hand, the typical frequency scale of $H_i(f; f_1)$ and $H_j(f; f_2)$ is $1/T_m$ which is usually much larger than $1/T'_n$. In most cases, then, we will find that the approximation

$$H_i(f; f_1)\cdot H_j(f; f_2) \approx H_i(f + f'; f_1)\cdot H_j(f - f'; f_2)$$

is a good approximation for the values of $f'$ making contributions of any importance to the integral. Under this approximation, and a change of variables of integration to

$$f'_1 = f + f', \qquad f'_2 = f - f',$$

we have, approximately,

$$\text{cov}\ \{P_i(f_1),\ P_j(f_2)\}\ \approx$$

$$\frac{1}{2}\int_{-\infty}^{\infty} H_i(f'_1;\ f_1)\cdot P(f'_1)\cdot H_j(f'_2;\ f_2)\cdot P(f'_2)$$

$$\cdot\left[\frac{\sin \pi(f'_1 - f'_2)T'_n}{\pi(f'_1 - f'_2)T'_n}\right]^2\cdot df'_1\cdot df'_2.$$

If we let

$$P_{i1}(f)\ =\ H_i(f; f_1)\cdot P(f), \tag{B-7.2}$$

and

$$P_{j2}(f)\ =\ H_j(f; f_2)\cdot P(f), \tag{B-7.3}$$

then, approximately,

cov $\{P_i(f_1), P_j(f_2)\}$

$$\approx 2 \iint\limits_0^\infty P_{i1}(f_1') \cdot P_{j2}(f_2') \cdot \left[\frac{\sin \pi(f_1' - f_2')T_n'}{\pi(f_1' - f_2')T_n'}\right]^2 \cdot df_1' \cdot df_2', \quad \text{(B-7.4)}$$

which is our approximation on assumption (i).

In terms of the same quantities we have, from (B-4.5),

$$\text{ave } \{P_i(f_1)\} = \int_0^\infty P_{i1}(f_1') \cdot df_1', \quad \text{(B-7.5)}$$

and

$$\text{ave } \{P_j(f_2)\} = \int_0^\infty P_{j2}(f_2') \cdot df_2'. \quad \text{(B-7.6)}$$

The following heuristic interpretation of the last three formulas is useful. Frequency $f_1'$ is involved in $P_i(f_1)$ with a net weight of $P_{i1}(f_1')$, while frequency $f_2'$ is involved in $P_j(f_2)$ with a net weight of $P_{j2}(f_2')$. These are net weights, and might represent partial cancellation. The covariance between $P_i(f_1)$ and $P_j(f_2)$ might involve additional sources of variability. To the extent that the approximation to $H_i(f; f_1) \cdot H_j(f; f_2)$ is valid, there are no additional sources of variability — the covariance involves the same net weights. The fact that we have only a record of equivalent length $T_n'$ is represented by a tendency of the frequency $f_1'$ to become entwined with the frequency $f_2'$, measured by the factor

$$\left[\frac{\sin \pi(f_1' - f_2')T_n'}{\pi(f_1' - f_2')T_n'}\right]^2.$$

The fact that there are no additional sources of variability as we pass from first to second moments (to the accuracy of the approximation) is good evidence that we are using the data in a relatively efficient way.

If we begin anew from (B-7.1) by expanding $P(f \pm f')$ in Taylor series around $f' = 0$, we find

$$P(f + f') \cdot P(f - f') = [P(f)]^2 + (f')^2\{P(f) \cdot P''(f) - [P'(f)]^2\} + \cdots\cdots$$

symmetry forcing the odd-order terms, including the first, to vanish. Since the trigonometric factor is very small for $|f'| > 2$ or 3 times $1/T_n'$, we may often replace $P(f + f') \cdot P(f - f')$ by $[P(f)]^2$ to a very good approximation. This yields, after integrating out $f'$,

$$\Gamma(f) \approx \frac{2}{T_n'} [P(f)]^2,$$

whence,

$$\text{cov}\ \{P_i(f_1),\ P_j(f_2)\} \approx \frac{1}{T'_n}\int_0^\infty H_i(f;f_1)\cdot H_j(f;f_2)[P(f)]^2\cdot df$$

$$\approx \frac{1}{T'_n}\int_0^\infty P_{i1}(f)P_{j2}(f)\ df,$$

our useful approximation on assumption (ii).

In carrying out this approximation we only needed smoothness of $P(f)$ for $f$'s which make a non-negligible contribution to the final result. If $H_i(f;f_1)P(f)$ and $H_j(f;f_2)P(f)$ are both smooth, as under assumption (iii), then $P(f)$ must be smooth except where both $H_i(f;f_1)$ and $H_j(f;f_2)$ are small, and the contribution from such regions can be neglected. Thus our useful approximation for the *covariance* under assumptions (iii) is the same as that under assumption (ii). (The approximation for $\Gamma(f)$ need not be so accurate.)

Finally, if $P(f)$ consists only of a very sharp peak (width $\ll 1/T'_n$) at $f = f_0$ (and, of course, at $f = -f_0$), with area $A$, then

$$P(f) \approx A\cdot[\delta(f+f_0) + \delta(f-f_0)],$$

whence,

$$\Gamma(f) \approx 2A^2\cdot\left[\delta(f+f_0) + \delta(f-f_0) + 2\left(\frac{\sin \omega_0 T'_n}{\omega_0 T'_n}\right)^2\cdot\delta(f)\right].$$

Hence,

$$\text{cov}\ \{P_i(f_1),\ P_j(f_2)\} \approx A^2\left\{H_i(f_0\ ;\ f_1)\cdot H_j(f_0\ ;\ f_2)\right.$$

$$\left. +\ H_i(0;\ f_1)\cdot H_j(0;\ f_2)\cdot\left(\frac{\sin \omega_0 T'_n}{\omega_0 T'_n}\right)^2\right\},$$

which is our useful approximation under assumption (iv). In case $f_0 \gg 1/T'_n$, the trigonometric factor may be neglected, and we may write

$$\text{cov}\ \{P_i(f_1),\ P_j(f_2)\} \approx \left(\int_0^\infty P_{i1}(f)\ df\right)\cdot\left(\int_0^\infty P_{j2}(f)\ df\right).$$

B.8 *Equivalent Widths*

Under assumption (ii) of Section B.7, i.e. $P(f)$ slowly varying, we have, for the dimensionless variability of $P_i(f_1)$,

$$\frac{\text{var}\ \{P_i(f_1)\}}{[\text{ave}\ \{P_i(f_1)\}]^2} = \frac{1}{T'_n W_e}, \tag{B-8.1}$$

where

$$W_e = \frac{\left[\displaystyle\int_0^\infty P_{i1}(f)\cdot df\right]^2}{\displaystyle\int_0^\infty [P_{i1}(f)]^2 \cdot df} \tag{B-8.2}$$

is the *equivalent width* of

$$P_{i1}(f) = H_i(f; f_1)\cdot P(f).$$

We will now determine the equivalent width of $P_{i1}(f)$ under the assumption that, for each value of $f_1$, $P_{i1}(f)$ is essentially a constant times $H_i(f; f_1)$. The value of the constant may depend upon the value of $f_1$, but it will not have any effect on the value of $W_e$. It is convenient to express (B-8.2) in terms of the normalized frequency

$$\varphi = 2\pi f T_m, \qquad (\varphi_1 = 2\pi f_1 T_m),$$

so that

$$W_e = \frac{1}{4\pi T_m}\frac{\left[\displaystyle\int_{-\infty}^\infty \tilde{P}_{i1}(\varphi)\cdot d\varphi\right]^2}{\displaystyle\int_{-\infty}^\infty [\tilde{P}_{i1}(\varphi)]^2 \cdot d\varphi}, \tag{B-8.3}$$

where

$$\left.\begin{aligned}
\tilde{P}_{i1}(\varphi) &= \tilde{Q}_i(\varphi + \varphi_1) + \tilde{Q}_i(\varphi - \varphi_1), \\[2mm]
\tilde{Q}_i(\varphi) &= (1 - a_i)\tilde{Q}_0(\varphi) + \frac{a_i}{2}\,[\tilde{Q}_0(\varphi + \pi) + \tilde{Q}_0(\varphi - \pi)], \\[2mm]
\tilde{Q}_0(\varphi) &= \frac{\sin\varphi}{\varphi}, \\[2mm]
a_0 &= 0, \qquad a_2 = 0.50, \qquad a_3 = 0.46.
\end{aligned}\right\} \tag{B-8.4}$$

Since

$$\int_{-\infty}^\infty \frac{\sin\varphi}{\varphi}\, d\varphi = \pi,$$

we get.

$$\int_{-\infty}^\infty \tilde{P}_{i1}(\varphi)\cdot d\varphi = 2\pi,$$

and, since

$$\int_{-\infty}^\infty \frac{\sin(\varphi + \alpha)}{\varphi + \alpha}\frac{\sin(\varphi + \beta)}{\varphi + \beta}\, d\varphi = \frac{\sin(\alpha - \beta)}{\alpha - \beta}\,\pi,$$

we get

$$\int_{-\infty}^{\infty} [\tilde{P}_{i1}(\varphi)]^2 \, d\varphi = 2\pi B \,,$$

where

$$B = \left(1 - 2a_i + \frac{3}{2} a_i^2\right)\left(1 + \frac{\sin 2\varphi_1}{2\varphi_1}\right)$$

$$+ a_i(1 - a_i)\left[\frac{\sin(2\varphi_1 - \pi)}{2\varphi_1 - \pi} + \frac{\sin(2\varphi_1 + \pi)}{2\varphi_1 + \pi}\right] \qquad \text{(B-8.5)}$$

$$+ \frac{a_i^2}{4}\left[\frac{\sin 2(\varphi_1 - \pi)}{2(\varphi_1 - \pi)} + \frac{\sin 2(\varphi_1 + \pi)}{2(\varphi_1 + \pi)}\right].$$

Hence, for these windows,

$$W_e = \frac{1}{2T_m B}. \qquad \text{(B-8.6)}$$

It will be noted that

$$B = 2\left(1 - 2a_i + \frac{3}{2} a_i^2\right), \qquad \text{for } f_1 = 0 \,,$$

$$= 1 - a_i + \frac{1}{2} a_i^2, \qquad \text{for } f_1 T_m = \frac{1}{4} \,,$$

$$= 1 - 2a_i + \frac{7}{4} a_i^2, \qquad \text{for } f_1 T_m = \frac{1}{2} \,,$$

$$= 1 - 2a_i + \frac{3}{2} a_i^2, \qquad \text{for } f_1 T_m = \frac{r}{4} \,,$$

$$(r = 3, 4, 5, \cdots).$$

To a close approximation,

$$B = 1 - 2a_i + \frac{3}{2} a_i^2, \qquad \text{for } f_1 T_m \geqq 1,$$

whence, for

$i = 0$, the main lobe is $1/T_m$ wide, and the equivalent width of $P_{01}(f)$ is
$1/(2T_m)$,

$i = 2$, the main lobe is $2/T_m$ wide, and the equivalent width of $P_{21}(f)$ is
$4/(3T_m)$,

$i = 3$, the main lobe is $2/T_m$ wide, and the equivalent width of $P_{31}(f)$ is
$1.258/T_m$ .
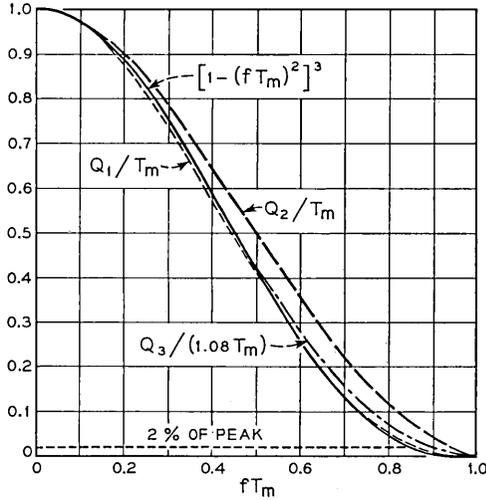
Fig. 16 — Approximation to standard spectral windows.

For $f_1 < 1/T_m$ these equivalent widths are to be reduced. For $f_1 = 0$, they are to be halved.

Assume, next, that

$$Q_i(f) = A [1 - (T_m f)^2]^3, \qquad |f| \leqq \frac{1}{T_m},$$

$$= 0, \qquad \text{otherwise,} \qquad \text{(B-8.7)}$$

which approximates $Q_i(f)$ quite well for $i = 1, 2$, and 3, as is shown in Fig. 16. Let us further assume that $f_1 \geqq 1/T_m$ and that, at least for $|f - f_1| \leqq 1/T_m$,

$$P(f) = P(f_1)[1 + \beta T_m(f - f_1)]. \qquad \text{(B-8.8)}$$

Then we can evaluate the equivalent width of

$$P_{i1}(f) = H_i(f; f_1) P(f)$$

by simple integrations, finding

$$W_e = \frac{429}{350 T_m \left(1 + \dfrac{\beta^2}{15}\right)}.$$

Since $\beta^2$ cannot exceed 1, we have

$$\frac{1.15}{T_m} < W_e < \frac{1.23}{T_m}, \qquad \text{(B-8.9)}$$

showing how little effect a linear slope of $P(f)$ across the main lobe of $H_i(f; f_1)$ has on $W_e$, even when $\beta$ is large enough for $P(f)$ to vanish at one edge.

If $P(f)$ has exactly the same form as the $H_i(f; f_1)$ corresponding to (B-8.7), then

$$W_e = \frac{0.94}{T_m},\tag{B-8.10}$$

a rather smaller value.

If, on the other hand, we take

$$Q_i(f) = A[1 - (T_m f)^2]^\ell, \qquad |f| \geqq \frac{1}{T_m},$$
$$= 0, \qquad \text{otherwise,}\tag{B-8.11}$$

then for $\ell = 2$, and $P(f)$ given by (B-8.8),

$$W_e = \frac{7}{5T_m\left(1 + \dfrac{\beta^2}{11}\right)},$$

so that

$$\frac{1.28}{T_m} < W_e \leqq \frac{1.40}{T_m},\tag{B-8.12}$$

while for $\ell = 1$,

$$W_e = \frac{5}{3T_m\left(1 + \dfrac{\beta^2}{7}\right)},$$

so that

$$\frac{1.46}{T_m} < W_e < \frac{1.67}{T_m}.\tag{B 8.13}$$

In general, we can use

$$W_e \approx \frac{1}{T_m}\tag{B-8.14}$$

as a conservative approximation which provides a factor of safety often near 1.15 or 1.20.

All this was for $f_1 \geqq 1/T_m$. As $f_1$ is reduced below $1/T_m$ there is overlapping between $Q_i(f - f_1)$ and $Q_i(f + f_1)$ in $H_i(f; f_1)$. As a consequence, the equivalent width decreases in a way which is not worth examining in

great detail. (At $f_1 = 1/(2T_m)$ the equivalent width has decreased about 15 per cent, and at $f_1 = 0$ it has fallen to just half its usual value.)

B.9 *Equivalent Degrees of Freedom*

Let us assume that $P(f)$ is flat — uniformly equal to $p_0$. Then the power variance spectrum $\Gamma(f)$ will also be flat, but with a value of $2p_0^2/T'_n$. Let us consider $H_i(f; f_1)$ to be the ideal bandpass power transfer function

$$H_i(f; f_1) = A, \quad f_1 - \frac{W}{2} < |f| < f_1 + \frac{W}{2},$$

$$= 0, \quad \text{otherwise},$$

where $f_1 > W/2$, although such a transfer function is not even approximately realizable under the requirement that $D_i(\tau) \equiv 0$ for $|\tau| > T_m$. Then

$$\text{ave } \{P_i(f_1)\} = AWp_0,$$

and

$$\text{var } \{P_i(f_1)\} = \frac{A^2 W p_0^2}{T'_n}.$$

If we equate these moments with the corresponding moments of a multiple of a chi-square variate with $k$ degrees of freedom, we get

$$k = \frac{2[\text{ave } \{P_i(f_1)\}]^2}{\text{var } \{P_i(f_1)\}} = 2WT'_n.$$

We get the same result if $H_i(f; f_1)$ is assumed to be the ideal lowpass power transfer function

$$H_i(f; f_1) = A, \quad |f| < W,$$

$$= 0, \quad \text{otherwise}.$$

In either case we get only one degree of freedom when $W = 1/(2T'_n)$.

This suggests that the frequency range $f > 0$ be divided into *elementary bands* of width

$$\Delta f = \frac{1}{2T'_n} \tag{B-9.1}$$

with *one degree of freedom in each*. (In the presence of very sharp peaks in the original spectrum it would be somewhat more accurate to divide the frequency range $-\infty < f < \infty$ into bands of width $1/T'_n$ with one

degree of freedom in each, and with one band centered at $f = 0$.) It follows from the results of the preceding section that, if $P(f)$ is reasonably smooth, we may regard the stability of the power density estimate $P_i(f_1)$, for $i = 1, 2,$ or 3, as approximately equal to that of a chi-square variate with $k$ degrees of freedom, where

$$k = \frac{2T'_n}{T_m} = 2\left(\frac{T_n}{T_m} - \frac{1}{3}\right) \tag{B-9.2}$$

for one piece (or record).

B.10 *Filtering and Analog Computation*

Power spectrum analysis from continuous data is frequently done by filtering techniques. In this section we will examine some of these techniques. In particular, we will try to express their results in such a form that we are led to express their reliability in equivalent numbers of degrees of freedom.

A common technique is to apply the signal to a narrow-band filter, allow some time for initial transients to become negligible, and then measure the average output power over the remaining time of the record. This corresponds to Mode I as described in Section 10 of Part I. In this technique it is clear that the estimate $P_Y(f_1)$ of the power density $P(f_1)$ in the power spectrum $P(f)$ of the original random process at the filter input, is, substantially,

$$P_Y(f_1) = \int_{-\infty}^{\infty} P_{\text{out}}(f; f_1) \cdot df$$

where

$$P_{\text{out}}(f; f_1) = |Y(f; f_1)|^2 \cdot P(f)$$

is the power spectrum of the modified random process at the filter output, and $Y(f; f_1)$ is the transfer function of the filter with a narrow passband around the frequency $f_1$. The record length is shorter for the modified process than for the original process, by the time allowed for the initial transients to become negligible (at least the reciprocal of the bandwidth in cycles). The effective record length for the modified process determines the width of the elementary bands, and the equivalent number of degrees of freedom is the number of elementary bands in the equivalent width of $P_{\text{out}}(f; f_1)$, or the equivalent width of $|Y(f; f_1)|^2$ if $P(f)$ is reasonably constant in the filter passband.

In the technique described as Mode II in Section 10 we integrate *all* of the power output of the filter, and divide by the length $T$ of the original record to obtain the estimate $P_Y(f_1)$. The result may clearly be ex-

pressed in the form

$$P_Y(f_1) = \frac{1}{T} \int_{-\infty}^{\infty} \{W(t; f_1) * [B(t) \cdot X(t)]\}^2 \, dt,$$

where $W(t; f_1)$ is the impulse response of the filter, and $B(t) = 0$ for $t < 0$ and $t > T$, but is otherwise arbitrary. To reduce this to a familiar form we first write it out in detail as

$$P_Y(f_1) = \frac{1}{T} \iiint\limits_{-\infty}^{\infty} W(\tau_1; f_1) \cdot W(\tau_2; f_1) \cdot B(t - \tau_1) \cdot B(t - \tau_2) \cdot X(t - \tau_1)$$

$$\cdot X(t - \tau_2) \cdot d\tau_1 \, d\tau_2 \, dt,$$

so that

$$\text{ave} \{P_Y(f_1)\} = \frac{1}{T} \iiint\limits_{-\infty}^{\infty} W(\tau_1; f_1) \cdot W(\tau_2; f_1) \cdot B(t - \tau_1) \cdot B(t - \tau_2)$$

$$\cdot C(\tau_1 - \tau_2) \cdot d\tau_1 \, d\tau_2 \, dt.$$

Now,

$$C(\tau_1 - \tau_2) = \int_{-\infty}^{\infty} P(f) \cdot e^{-i\omega(\tau_1 - \tau_2)} \, df,$$

while, if $J(f)$ is the Fourier transform of $B(t)$,

$$\int_{-\infty}^{\infty} B(t - \tau_1) \cdot B(t - \tau_2) \cdot dt$$

$$= \iiint\limits_{-\infty}^{\infty} J(f') \cdot J(f'') \cdot e^{i\omega'(t-\tau_1)+i\omega''(t-\tau_2)} \, df' \, df'' \, dt$$

$$= \iint\limits_{-\infty}^{\infty} J(f') \cdot J(f'') \cdot \delta(f' + f'') \cdot e^{-i(\omega'\tau_1+\omega''\tau_2)} \, df' \, df''$$

$$= \int_{-\infty}^{\infty} |J(f')|^2 \cdot e^{-i\omega'(\tau_1-\tau_2)} \, df'.$$

Hence,

$$\text{ave} \{P_Y(f_1)\} = \frac{1}{T} \iiiint\limits_{-\infty}^{\infty} |J(f')|^2 \cdot W(\tau_1; f_1) \cdot W(\tau_2; f_1) \cdot P(f)$$

$$\cdot e^{i(\omega-\omega')(\tau_1-\tau_2)} \, d\tau_1 \, d\tau_2 \, df' \, df.$$

Further, since $Y(f; f_1)$ is the Fourier transform of $W(t; f_1)$,

$$\text{ave}\{P_Y(f_1)\} = \frac{1}{T} \iint\limits_{-\infty}^{\infty} |J(f')|^2 \cdot |Y(f - f'; f_1)|^2 \cdot P(f) \cdot df' \, df.$$

Finally, therefore,

$$\text{ave } \{P_Y(f_1)\} = \int_0^\infty H_Y(f; f_1) \cdot P(f) \cdot df, \qquad \text{(B-10.1)}$$

where

$$H_Y(f; f_1) = \frac{2}{T} |J(f)|^2 * |Y(f; f_1)|^2. \qquad \text{(B-10.2)}$$

Since (B-10.1) is in the same form as (B-4.5), we may now apply the results of Sections B.8 and B.9. In particular, if $P(f)$ is reasonably smooth, we may regard the stability of the estimate $P_Y(f_1)$ as approximately equal to that of a chi-square variate with $k$ degrees of freedom, where

$$k = 2T \frac{\left[\int_0^\infty H_Y(f; f_1) \cdot df\right]^2}{\int_0^\infty [H_Y(f; f_1)]^2 \cdot df}. \qquad \text{(B-10.3)}$$

From the fact that $H_Y(f; f_1)$ is the convolution of the power transfer function of the filter with $(2/T)|J(f)|^2$ it is clear that the passband of $H_Y(f; f_1)$ is at least as wide as that of the wider of the filter and $(2/T)|J(f)|^2$. Hence, the resolving power of the filter method of power spectrum analysis is limited by the length $T$ of available data, just as is that of any other method. If the filter passband is made narrower than $1/T$, say, not only do we gain very little in resolving power, but the stability of our power density estimates is then, at best, approximately that of a chi-square variate with only one or two degrees of freedom. To obtain a reasonable degree of stability the filter passband should be several to many times $1/T$ wide. The resolving power then depends largely on this width. Under these circumstances it should not make much difference which of the four modes described in Section 10 is used.

In Mode III, the output power may clearly be expressed in the form

$$P_Y(f_1) = \frac{1}{T} \int_{-\infty}^{\infty} [\tilde{B}(t) \cdot \{W(t; f_1) * [B(t) \cdot X(t)]\}]^2 \, dt,$$

where $\tilde{B}(t) = 0$ for $t < 0$ and $t > T$, but is arbitrary otherwise. The re-

duction of this to a familiar form follows closely that of the preceding
case. The final result is that

$$\text{ave}\ \{P_Y(f_1)\}\ =\ \int_0^\infty\ H_Y(f;\,f_1)\cdot P(f)\cdot df, \qquad\text{(B-10.4)}$$

where

$$H_Y(f;\,f_1)\ =\ \frac{2}{T}\int_{-\infty}^\infty\ \left|\ \int_{-\infty}^\infty\ J(f'')\cdot Y(f''\ -\ f;\,f_1)\right.$$

$$\left.\cdot\tilde{J}(f'\ -\ f'')\cdot df''\ \right|^2\cdot df'. \qquad\text{(B-10.5)}$$

The analysis of Mode IV differs from that of Mode III only in the
specifications of $\tilde{J}(f)$. Indeed, the results for Modes I and II are also
special cases of the result for Mode III. Mode II corresponds to $\tilde{J}(f)\ =\ \delta(f)$, while Mode I corresponds to $J(f)\ =\ \delta(f)$ as well as $\tilde{J}(f)\ =\ \delta(f)$.

It was noted in Section 10 that the zero of the input noise may not be
quite at ground potential in Fig. 5. The discrepancy may be considered
to produce a spurious *line* or *delta* component at zero frequency in the
spectrum of the input noise. It will have no effect on the average output
power in Mode I if $f_1\ \neq\ 0$, because the spectral window, which is simply
$|\ Y(f;\,f_1)\ |^2$ is opaque at zero frequency. In Mode II, however, the spec-
tral window specified by (B-10.2) may not be sufficiently opaque at
zero frequency. The effect on the average output power depends upon
the value of $H_Y(0;\,f_1)$. If the spectral window $H_Y(f;\,f_1)$ were ideal in
the sense that it has no side lobes, there would be no effect on the
average output power for values of $f_1$ at least half of the width of the
spectral window. This indicates the desirability of using a graded data
window $B(t)$ in order to reduce the side lobes in the spectral window
$H_Y(f;\,f_1)$, that is, essentially the side lobes in $(2/T)\ |J(f)\ |^2$. Since the
latter window is necessarily positive for all values of $f$, reduction of side
lobes here is not quite as easy as it is in the indirect computation tech-
nique described in Section B.4, where selection can be exercised di-
rectly on the lag window, and the spectral window can be negative at
some frequencies.

Another filter technique frequently used in power spectrum analysis is
to apply the available record to the filter as a periodic function with a
period equal to the length of the record. If a data window is used, so
that the periodic function applied to the filter is of the form $B(t)\cdot X(t)$
in the interval $0\ <\ t\ <\ T$, then the power spectrum of the input function
is a *line spectrum* with power concentrated at integral multiples of $1/T$

cps. The total power in this spectrum, expressed in a form which displays the distribution in frequency, is

$$\text{input power} = \frac{1}{T} \sum_{q=-\infty}^{q=\infty} P_{\text{in}}\left(\frac{q}{T}\right),$$

where, essentially,

$$P_{\text{in}}(f) = \frac{1}{T} \left| \int_{-\infty}^{\infty} B(t) \cdot X(t) \cdot e^{-i\omega t} \cdot dt \right|^2.$$

Hence, the output power of the filter, is

$$P_Y(f_1) = \frac{1}{T} \sum_{q=-\infty}^{q=\infty} \left| Y\left(\frac{q}{T}; f_1\right) \right|^2 \cdot P_{\text{in}}\left(\frac{q}{T}\right).$$

Aside from the fact that we are now dealing with sums instead of integrals, this analysis parallels very closely our analysis of Mode II. It should be noted, however, that while our summands are even functions of $q$ we may not run our sums from $q = 0$ to $q = \infty$, and then double the result, unless we take only half of the first ($q = 0$) term. Hence, with this technique the equivalent number of degrees of freedom should be taken as

$$k = \frac{\left[ \sum_{q=-\infty}^{q=\infty} \left| Y\left(\frac{q}{T}; f_1\right) \right|^2 \right]^2}{\sum_{q=-\infty}^{q=\infty} \left| Y\left(\frac{q}{T}; f_1\right) \right|^4}$$

provided that this turns out to give $k > 3$ or 4, say. It is easily seen that this formula discourages the use of this technique with a filter whose passband is only a few times $1/T$ wide.

B.11 *Prewhitening*

The techniques required here are standard in communication engineering and do not require specific treatment. As these techniques are ordinarily used, the original stationary random process is in effect continuously acting on the input end of the prewhitening filter, so that the output of the filter may be regarded as another stationary random process for purposes of spectral analysis. If these techniques are used where it is practical to obtain only a finite length of the original process to apply to the input of the filter, we must discard an initial portion of the output, corresponding to the time required for the filter transients to die out, as well as all of the output after the input has ceased. These

considerations are taken up in greater detail in Sections B.15 and 16 in connection with the analysis of equi-spaced discrete data.

## DETAILS FOR EQUI-SPACED ANALYSIS

B.12 *Aliasing*

Let us consider a stationary random process whose autocovariance function $\tilde{C}(\tau)$, and power spectrum $\tilde{P}(f)$ are known exactly. If we now take only the values of $\tilde{C}(\tau)$ at uniformly spaced values of $\tau$, viz.

$$\tau = 0, \ \pm\Delta\tau, \ \pm 2\Delta\tau, \ \cdots$$

we can, in principle, calculate a corresponding (aliased) power spectrum $\tilde{P}_a(f)$, by using the formula

$$\tilde{P}_a(f) = \int_{-\infty}^{\infty} [\nabla(\tau; \Delta\tau)\cdot\tilde{C}(\tau)]\cdot e^{-i\omega\tau}\,d\tau,$$

where $\nabla(\tau; \Delta\tau)$ is an infinite Dirac comb as defined in Section A.2. Making use of the results of Sections A.2 and A.3 we have

$$\tilde{P}_a(f) = A\left(f; \frac{1}{\Delta\tau}\right) * \tilde{P}(f),$$

or, explicitly,

$$\tilde{P}_a(f) = \sum_{q=-\infty}^{q=\infty} \tilde{P}\left(f - \frac{q}{\Delta\tau}\right).$$

Thus, if it happens that $\tilde{P}(f)$ is zero for $|f| > 1/(2\,\Delta\tau)$, as illustrated in Fig. 17(a), there will be no overlapping of the individual terms in $\tilde{P}_a(f)$, and the result of the summation will be as illustrated in Fig. 17(b). In this case, we can restore the original spectrum by multiplying $\tilde{P}_a(f)$ by a rectangular function according to the formula

$$\tilde{P}(f) = \tilde{P}_A(f) = \tilde{S}(f)\cdot\tilde{P}_a(f),$$

where

$$\tilde{S}(f) = 1, \quad |f| < \frac{1}{2\Delta\tau},$$

$$= 0, \quad |f| > \frac{1}{2\Delta\tau},$$

for, in the absence of such overlapping, $\tilde{P}_a(f) = \tilde{P}(f)$ for $|f| < 1/(2\Delta\tau)$. Hence, we will recover the original autocovariance function $\tilde{C}(\tau)$ from

that for the discrete series by a convolution according to the formula

$$\tilde{C}(\tau) = \tilde{G}(\tau) * [\nabla(\tau; \Delta\tau) \cdot \tilde{C}(\tau)]$$

$$= \Delta\tau \cdot \sum_{q=-\infty}^{q=\infty} \tilde{G}(\tau - q\Delta\tau) \cdot \tilde{C}(q\Delta\tau),$$

where

$$\tilde{G}(\tau) = \frac{\sin\frac{\pi\tau}{\Delta\tau}}{\pi\tau} = \frac{1}{\Delta\tau} \cdot \text{dif}\,\frac{\tau}{\Delta\tau}.$$

Now let us consider a second stationary random process whose auto-covariance function $C(\tau)$ happens to be related to $\tilde{C}(\tau)$ by

$$C(\tau) = G(\tau) \cdot \tilde{C}(\tau),$$

where $G(\tau)$ is unity at $\tau = 0$, $\pm\Delta\tau$, $\pm2\Delta\tau$, etc. Since

$$\nabla(\tau; \Delta\tau) \cdot C(\tau) = \nabla(\tau; \Delta\tau) \cdot \tilde{C}(\tau), \quad .$$

it is clear that $P_a(f)$ will be quantitatively, and in the utmost detail
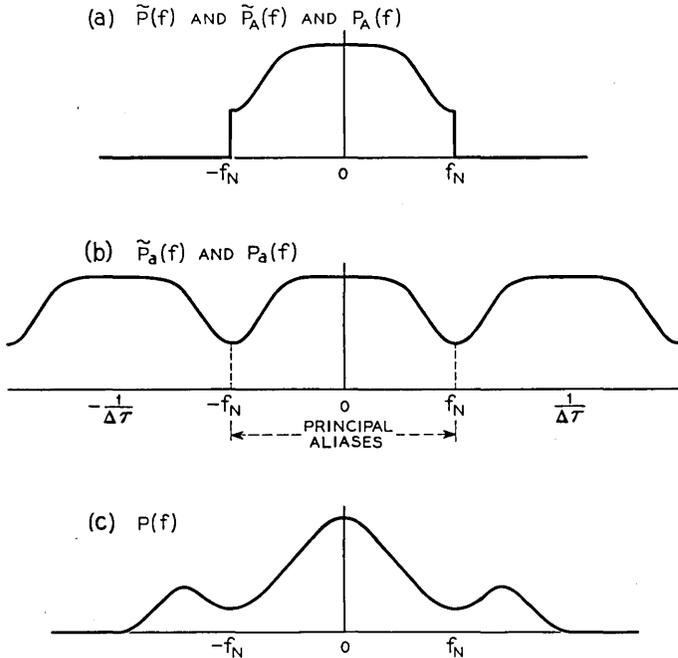


Fig. 17 — Effect of sampling on power spectra.

identical with $\tilde{P}_a(f)$. Hence, the *principal part* $\tilde{P}_A(f)$ of the aliased spectrum $\tilde{P}_a(f)$, that is, the part in the range $|f| < 1/(2\Delta\tau)$, will not in general be the power spectrum of the second process. In fact, if $S(f)$ is the Fourier transform of $G(\tau)$, the power spectrum of the second process is related to the power spectrum of the first process by

$$P(f) = S(f) * \tilde{P}(f).$$

In general, this spectrum covers an infinite range of frequencies, so that the aliased spectrum

$$P_a(f) = A\left(f; \frac{1}{\Delta\tau}\right) * P(f) = \sum_{q=-\infty}^{\infty} P\left(f - \frac{q}{\Delta\tau}\right)$$

will involve overlapping of the individual terms. This overlapping will account for the quantitative identity of $P_a(f)$ with $\tilde{P}_a(f)$, and the failure of its principal part $P_A(f)$ to represent $P(f)$ even in the range $|f| < 1/(2 \Delta\tau)$. Fig. 17(c) illustrates such a $P(f)$.

It may well have already occurred to readers familiar with amplitude modulation that using only uniformly spaced values of $C(\tau)$, viz., $C(r\Delta\tau)$ where $r = 0, \pm1, \pm2, \cdots$, has the same effect on the power spectrum as the simultaneous amplitude modulation of carrier waves with frequencies $q/\Delta\tau$ where $q = 0, \pm1, \pm2, \cdots$. If the two-sided power spectrum $P(f)$ corresponding to $C(\tau)$ is visualized as side-bands on a zero-frequency carrier, then the aliased spectrum $P_a(f)$ corresponding to $C(r\Delta\tau)$ will be naturally visualized as the same sidebands on carrier frequencies $q/\Delta\tau$, where $q = 0, \pm1, \pm2, \cdots$. If each sideband of $P(f)$ does not extend beyond the frequency $1/(2\Delta\tau)$, then there will be no overlapping of side-bands in $P_a(f)$, and the principal part $P_A(f)$ of $P_a(f)$ will be identical to $P(f)$. Contrariwise, if each sideband of $P(f)$ extends beyond the frequency $1/(2\Delta\tau)$, then there will be overlapping of side-bands in $P_a(f)$, and the principal part $P_A(f)$ of $P_a(f)$ will not be identical to $P(f)$. In any case, however, it is important to note that

$$\int_{-1/(2\Delta\tau)}^{1/(2\Delta\tau)} P_a(f)\ df = \int_{-1/(2\Delta\tau)}^{1/(2\Delta\tau)} P_A(f)\ df = \int_{-\infty}^{\infty} P_A(f)\ df = \int_{-\infty}^{\infty} P(f)\ df.$$

If we examine the relation of $P_A(f')$ to $P(f)$, where

$$0 \leq f' \leq \frac{1}{2\Delta\tau}, \quad \text{and} \quad f \geq 0,$$

we find that

$$P_A(f') = P(f') + \sum_{q=1}^{\infty}\left[P\left(\frac{q}{\Delta\tau} - f'\right) + P\left(\frac{q}{\Delta\tau} + f'\right)\right],$$
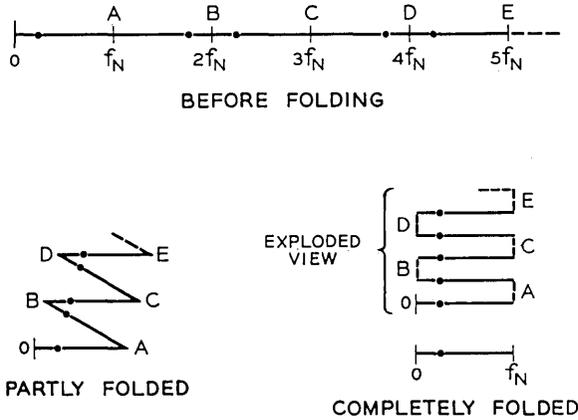
Fig. 18 — Spectrum folding.

(the right-hand side representing $P_a(f')$ for all $f'$). We say, therefore, that the power density at $f = (q/\Delta\tau) - f'$ and at $f = (q/\Delta\tau) + f'$ in the true spectrum $P(f)$, where $q = 1, 2, \cdots, \infty$, are *aliased* at $f'$ in the principal part $P_A(f')$ of the aliased spectrum. Clearly, the presence of aliases in power density estimates from time series is a matter of some concern.

Aliasing is sometimes called *spectrum folding*, because the pattern by which various frequencies are aliased with one another corresponds to the result of folding up the frequency axis, as illustrated in Fig. 18. The frequency of the first fold is called the *folding* or *Nyquist frequency*. In the discussion above this was $f_N^* = 1/(2\Delta\tau)$, which usually coincides with $f_N = 1/(2\Delta t)$.

B.13 *Transformation and Windows*

A general outline of a hypothetical procedure for power spectrum estimation from a uniformly spaced discrete time series of finite length, by the indirect route, is as follows:

(1) Let $X_0, X_1, \cdots, X_n$ be the time series and let $\Delta t$ be the time interval between adjacent values. Compute *mean lagged products*, with lag interval $\Delta\tau = h\Delta t$, according to the formula

$$C_r = \frac{1}{n - hr} \sum_{q=0}^{q=n-hr} X_q \cdot X_{q+hr}$$

$$\left( r = 0, 1, \cdots, m \quad \text{where} \quad m \leqq \frac{n}{h} \right).$$

(In a practical procedure, this formula will have to be modified to avoid difficulties with spurious low-frequency components.)

(2) Compute *raw spectral density estimates* according to the formula

$$V_r = \Delta\tau \cdot \left[ C_0 + 2 \sum_{q=1}^{q=m-1} C_q \cdot \cos\frac{qr\pi}{m} + C_m \cdot \cos r\pi \right].$$

Since $V_r$ is symmetric in $r$ about every integral multiple of $m$, it is necessary to compute it only for $r = 0, 1, \cdots, m$. The frequency corresponding to $r$ is $r/(2m\Delta\tau)$, as shown below.

(3) Compute *refined spectral density estimates* according to the formula

$$U_r = a_{i0}V_r + \sum_{j=1}^{\infty} a_{ij}[V_{r+j} + V_{r-j}],$$

where the $a$'s are the same as in Section B.5. In particular, for the third pair of lag and spectral windows described in Section B.5, we have $a_{30} = 0.54$ and $a_{31} = 0.23$, all others being zero, so that

$$U_r = 0.23\ V_{r-1} + 0.54\ V_r + 0.23\ V_{r+1}.$$

(These power density estimates should of course be doubled if they are to be referred to positive frequencies only. This doubling may in fact be introduced through the mean lagged products.)

Comparing this outline with the one for the continuous case (Section B.4), it will be noted that the window is introduced by different methods. In the continuous case the window is introduced as a lag window before cosine transformation, in order to avoid convolution after transformation. In the discrete series case, since convolution is not difficult, indeed is very simple, convolution after transformation is convenient, and the lag window is shaped after the cosine transformation.

To relate the outlined procedure for the discrete series case to that for the continuous case, we note that

$$\text{ave } \{C_r\} = C(r\Delta\tau),$$

where $C(\tau)$ is the true autocovariance function. Hence,

$$\text{ave } \{V_r\} = \Delta\tau \cdot \left[ C(0) + 2 \sum_{q=1}^{q=m-1} C(q\Delta\tau) \cos\frac{qr\pi}{m} + C(m\Delta\tau) \cos r\pi \right].$$

This may be expressed in the form of a Fourier transform, viz.,

$$\text{ave } \{V_r\} = \int_{-\infty}^{\infty} [\nabla_m(\tau; \Delta\tau) \cdot C(\tau)] \cdot e^{-i\omega\tau}\ d\tau,$$

where $\nabla_m(\tau; \Delta\tau)$ is the finite Dirac comb defined in Section A.2, and

$$f = \frac{\omega}{2\pi} = \frac{r}{2m\Delta\tau}.$$

Therefore,

$$\text{ave } \{V_r\} = \left[ Q_0(f; \Delta\tau) * P(f) \right]_{f=r/(2m\cdot\Delta\tau)},$$

where $P(f)$ is the true power spectrum. Hence, $V_r$ may be regarded as an estimate of an average-over-frequency of $P(f)$ in the *aliased* neighborhood of $f = r/(2m\Delta\tau)$, with the spectral window $Q_{0A}(f) = Q_0(f; \Delta\tau)$ illustrated in Fig. 9.

Three other views of the relation of $V_r$ to $P(f)$ may be developed from the fact that (from the end of Section A.2)

$$\nabla_m(\tau; \Delta\tau) = D_0(\tau) \cdot \nabla(\tau; \Delta\tau),$$

$$Q_0(f; \Delta\tau) = Q_0(f) * A\left(f; \frac{1}{\Delta\tau}\right),$$

whence,

$$\text{ave } \{V_r\} = \left[ Q_0(f) * A\left(f; \frac{1}{\Delta\tau}\right) * P(f) \right]_{f=r/(2m\cdot\Delta\tau)}.$$

The view taken in the preceding paragraph corresponds to the substitution

$$Q_0(f) * A\left(f; \frac{1}{\Delta\tau}\right) = Q_{0A}(f).$$

A second view corresponds to the substitution

$$A\left(f; \frac{1}{\Delta\tau}\right) * P(f) = P_a(f),$$

and a third to the identity

$$A\left(f; \frac{1}{\Delta\tau}\right) * P(f) = A\left(f; \frac{1}{\Delta\tau}\right) * P_A(f).$$

The latter two correspond to the results

$$\text{ave } \{V_r\} = \left[ Q_0(f) * P_a(f) \right]_{f=r/(2m\cdot\Delta\tau)},$$

$$\text{ave } \{V_r\} = \left[ Q_{0A}(f) * P_A(f) \right]_{f=r/(2m\cdot\Delta\tau)}.$$

Hence, $V_r$ may also be regarded as an estimate of an average-over-frequency of the entire *aliased power spectrum* $P_a(f)$ in the neighborhood of $f = r/(2m\Delta\tau)$, with the spectral window $Q_0(f)$ illustrated in Fig. 14, or, and usually most usefully, as an estimate of an average-over-frequency of the *principal part $P_A(f)$ of the aliased spectrum*, in the neighborhood of $f = r/(2m\Delta\tau)$, with the aliased window $Q_{0A}(f)$, which is illustrated in Fig. 9.

Finally, a fourth view corresponds to the substitution (from Section B.4)

$$Q_0(f) * P(f) = \text{ave } \{P_0(f)\}.$$

Hence, $V_r$ may also be regarded as an aliased version of the estimated power spectrum $P_0(f)$ which would have been obtained from the continuous data in accordance with the hypothetical procedure outlined in Section B.4.

Regarded from any one of the four points of view developed above, we see that the raw spectral density estimates $V_r$, (assuming $\Delta\tau$ sufficiently small, so that aliasing is negligible) are in the same position as the estimated power spectrum $P_0(f)$ in the continuous case; the spectral window $Q_{0A}(f)$ is essentially $Q_0(f)$ with the same undesirable side lobes. Hence, the raw power density estimates must be refined. As in the continuous case, the refinement can be done by using a graded lag window before transformation. In the discrete series case, it may also be done by convolving the raw estimates with the $a_{ij}$'s of Section B.5 to obtain the refined estimates, as described under (3) in the outline.

In any event, we may write, as we did at the end of Section 4, of Part I,

$$\text{ave } \{U_r\} = \int_0^\infty H_i\left(f; \frac{r}{2m\Delta\tau}\right) \cdot P_a(f) \cdot df,$$

where

$$H_i(f; f_1) = Q_i(f + f_1) + Q_i(f - f_1).$$

Alternatively, we may write

$$\text{ave } \{U_r\} = \int_0^\infty H_{iA}\left(f; \frac{r}{2m\Delta\tau}\right) \cdot P(f) \cdot df,$$

or, more usually,

$$\text{ave } \{U_r\} = \int_0^{1/(2\cdot\Delta\tau)} H_{iA}\left(f; \frac{r}{2m\Delta\tau}\right) \cdot P_a(f) \cdot df$$

$$= \int_0^\infty H_{iA}\left(f; \frac{r}{2m\Delta\tau}\right) \cdot P_A(f) \, df,$$
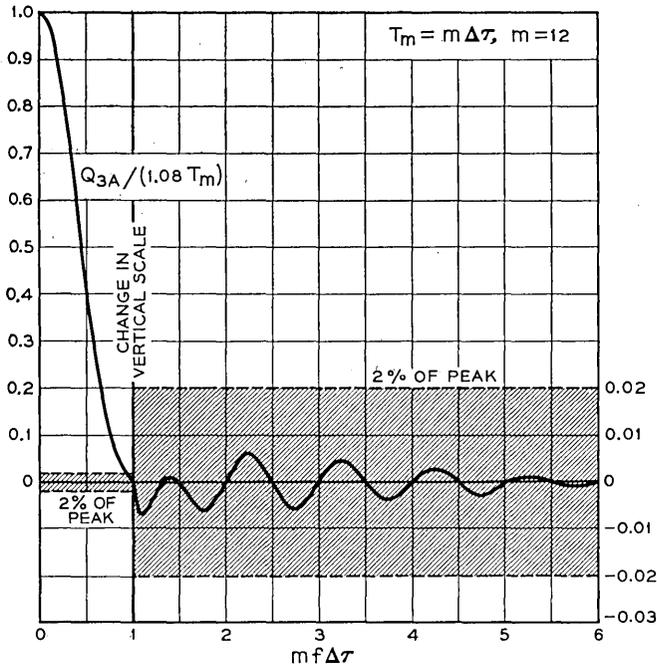
Fig. 19 — Aliased spectral window $Q_{3A}$ for $m = 12$.

where

$$H_{iA}(f; f_1) = Q_{iA}(f + f_1) + Q_{iA}(f - f_1),$$

and

$$Q_{iA}(f) = Q_i(f) * A\left(f; \frac{1}{\Delta\tau}\right) = \sum_{q=-\infty}^{+\infty} Q_i\left(f - \frac{q}{\Delta\tau}\right).$$

In particular, $Q_{2A}(f)$ and $Q_{3A}(f)$ are illustrated in Figs. 8 and 19.

B.14 *Variability and Covariability*

The analysis of variability of power density estimates from discrete time series is a repetition of Section B.6 up to and including (B-6.4) for cov $\{M(t_1, \tau_1), M(t_2, \tau_2)\}$. Beyond this point we now have to deal with summations rather than integrations with respect to $t_1$ and $t_2$.

If the range of summation in the formula for the mean lagged products, viz., $n$-$hr$, is replaced by an average range $n'$, which may be regarded as

the effective length of the series, and which may usually be taken as

$$n' = \text{largest integer less than} \left( n - \frac{hm}{3} \right)$$

then

$$\text{cov} \{C_r, C_s\} = \int_{-\infty}^{\infty} \cos \omega r \Delta \tau \cdot \cos \omega s \Delta \tau \cdot \Gamma_{\Delta t}(f) \cdot df,$$

where

$$\Gamma_{\Delta t}(f) = 4 \int_{-\infty}^{\infty} P(f + f') \cdot P(f - f') \cdot \left( \frac{\sin \omega' n' \Delta t}{n' \sin \omega' \Delta t} \right)^2 \cdot df'$$

(reducing to (B-6.10) and (B-6.11) as $\Delta t \to 0$ with $n' \Delta t = T'_n$). Finally, if $U_r$ and $U_s$ are refined power density estimates based on applying the spectral windows $Q_i(f)$ and $Q_j(f)$, respectively, to the aliased power spectrum $P_a(f)$, then

$$\text{cov} \{U_r, U_s\} = \frac{1}{4} \int_{-\infty}^{\infty} H_{iA} \left( f; \frac{r}{2m\Delta\tau} \right) \cdot H_{jA} \left( f; \frac{s}{2m\Delta\tau} \right) \cdot \Gamma_{\Delta t}(f) \cdot df.$$

In particular, of course

$$\text{var} \{U_r\} = \frac{1}{4} \int_{-\infty}^{\infty} \left[ H_{iA} \left( f; \frac{r}{2m\Delta\tau} \right) \right]^2 \cdot \Gamma_{\Delta t}(f) \cdot df.$$

We see now that there is essentially no difference between power density estimation from uniformly spaced time series and power density estimation from continuous data, except for aliasing and its secondary consequences. In particular, if $P(f)$ is reasonably smooth, and if aliasing is negligible, then we may judge the stability of the power density estimates $U_r$ for the uniformly spaced case by analogy with a chi-square variate with $k$ degrees of freedom, where

$$k = 2 \left( \frac{n\Delta t}{m\Delta\tau} - \frac{1}{3} \right), \quad r = 1, 2, \cdots, (m - 1),$$

$$= \text{half as much for } r = 0, m.$$

B.15 AND 16 *Transversal Filtering**

If the $Z$'s are moving linear combinations of the $X$'s, for example

$$Z_q = c_0 X_q + c_1 X_{q-1} + \cdots + c_k X_{q-k},$$

---

* See Kallmann[32] for the origin of this term.

and we take $\Delta t = 1$ for convenience, then the spectra are related by

$$P_Z(f) = P_X(f) \mid c_0 + c_1 e^{-i\omega} + c_2 e^{-2i\omega} + \cdots + c_k e^{-ki\omega} \mid^2$$

$$= P_X(f)[(c_0^2 + c_1^2 + c_2^2 + \cdots + c_k^2)$$

$$+ 2(c_0 c_1 + c_1 c_2 + \cdots + c_{k-1} c_k) \cos \omega$$

$$+ 2(c_0 c_2 + c_1 c_3 + \cdots + c_{k-2} c_k) \cos 2\omega + \cdots$$

$$+ 2(c_0 c_k) \cos k\omega],$$

where $\omega = 2\pi f$. The first equality arises by considering $X_q \equiv e^{iq\omega}$, and the second involves the sequence of coefficients

$$b_{-k} = c_0 c_k,$$

$$b_{-k+1} = c_0 c_{k-1} + c_1 c_k,$$

$$\cdot$$
$$\cdot$$
$$\cdot$$

$$b_{-1} = c_0 c_1 + c_1 c_2 + \cdots + c_{k-1} c_k,$$

$$b_0 = c_0^2 + c_1^2 + \cdots + c_k^2,$$

$$b_1 = c_1 c_0 + c_2 c_1 + \cdots + c_k c_{k-1},$$

$$b_2 = c_2 c_0 + c_3 c_1 + \cdots + c_k c_{k-2},$$

$$\cdot$$
$$\cdot$$
$$\cdot$$

$$b_{k-1} = c_{k-1} c_0 + c_k c_1,$$

$$b_k = c_k c_0,$$

which represent the convolution of the sequence $c_0, c_1, \cdots, c_k$ with itself.

As a filter, the moving linear combination is characterized by

$$\mid Y(f) \mid^2 = b_0 + 2 b_1 \cos \omega + 2 b_2 \cos 2\omega + \cdots + 2 b_k \cos k\omega,$$

which is never negative (as the square of an absolute value). Since we can write $\cos j\omega$ as a polynomial in $\cos \omega$ of degree $j$, we know that:

(1) $\mid Y(f) \mid^2$ can be written as a polynomial of degree $k$ in $\cos \omega$,

(2) for $-1 \leqq \cos \omega \leqq 1$, it is not negative.

Any such polynomial can be realized as a moving linear combination (in several ways, see Wold[33]). The simplest way to see this is to factor the given polynomial into linear and quadratic factors. By appropriate

choice of signs these factors will satisfy (2), and, if each corresponds to a real moving linear combination, the moving linear combination obtained by applying them successively will correspond to the given polynomial.

Any such linear factor takes the form

$$| Y_i(f) |^2 = A^2(1 + a \cos \omega)$$

with $| a | \leq 1$, which may be realized by $k = 1$ and

$$c_0, c_1 = \tfrac{1}{2}A(\sqrt{1 + a} \pm \sqrt{1 - a}).$$

Any such quadratic factor takes the form

$$| Y_j(f) |^2 = A^2(1 + a_1 \cos \omega + a_2 \cos^2 \omega).$$

The condition for this to be non-negative for $| \cos \omega | \leq 1$ is

$$| a_1 | \leq 1 + a_2,$$

and if $a_2 \geq 1$, so that the first condition forces an internal extremum on $[-1, +1]$, we must also have $a_1^2 \leq 4a_2$. We may write

$$| Y_j(f) |^2 = A^2 \left( 1 + \frac{a_2}{2} + a_1 \cos \omega + \frac{a_2}{2} \cos 2\omega \right),$$

which is realized by a three-point moving linear combination with

$$c_0, c_2 = \frac{A}{2\sqrt{2}} \left( \sqrt{1 + a_2 + \sqrt{\phantom{-}}} \pm \sqrt{1 - a_2 + \sqrt{\phantom{-}}} \right),$$

$$\sqrt{\phantom{--}} = \sqrt{(1 + a_2)^2 - a_1^2},$$

$$c_1 = \frac{A}{2} \left( \sqrt{1 + a_2 + a_1} - \sqrt{1 + a_2 - a_1} \right),$$

the conditions stated above ensuring that all radicals are real.

These two cases — linear and quadratic factors — not only prove that every polynomial non-negative on $(-1, +1)$ can be obtained, but provide at least one way to find a moving linear combination with an assigned polynomial.

One special case deserves record. If we require a simple moving linear combination with $Y(f_0) = 0$, we may use

$$c_0 = \frac{1}{2 \cos \omega_0},$$

$$c_1 = -1,$$

$$c_2 = \frac{1}{2 \cos \omega_0},$$

which has $| Y(f) |^2 = [1 - (\cos \omega/\cos \omega_0)]^2.$

Now that methods are available for finding moving linear combinations whose spectral windows are prescribed polynomials in cos $\omega$, some attention should be given to approximating an arbitrary desired response by polynomials. If a roughly "equal ripple" approximation (where the local maxima deviations of approximation from desired are roughly equal) is desired, then the techniques described in the next paragraph will be quite effective. If, as seems likely, however, we desire the *fractional* error

$$\frac{\text{approximation} - \text{desired}}{\text{desired}}$$

to have roughly equal ripples, then no specific method seems to be available. All we can suggest is the following procedure: (i) find a roughly equal ripple approximation, (ii) find the zeros of its error, (iii) squeeze these zeros together where smaller ripples are desired, and open them out where larger ripples can be permitted, keeping as much of the same general pattern of distances between zeros as possible, (iv) construct a new polynomial with these points for the zeros of its error, (v) adjust the result slightly, if necessary, to make it non-negative. (This procedure sounds quite plausible, but the reader should be warned that we do not know how to be more specific about "keeping as much of the same general pattern of distances as possible". However, we expect the procedure to work in many hands.)

The construction of the roughly "equal ripple" approximation can proceed in many ways. In almost every case, one should begin by calculating values of the desired response at values of cos $\omega$ equally spaced from cos $\omega = +1$ to cos $\omega = -1$. The semi-classical approach (DeLury,[34] Fisher-Yates,[35] or Milne[36]) would be to fit orthogonal polynomials to these equi-spaced points. The results would be least-square fits, but might be far from equal-ripple. The process of "economization" (Lanczos,[37] Lanczos,[38] streamlined by Minnick[39]) will allow us to take an over-fitted least-square fit and back up to an "equal-ripple" polynomial of lower order. However, the direct attack, based on central differences of the desired response (at equi-spaced values of cos $\omega$) as proposed by Miller[40] seems likely to be shorter, even allowing for the expansion of the Čebyšěv or Chebyshěv polynomial series into single polynomials.

One further set of considerations remains which is sometimes important. These relate to the starting up of a moving linear combination. If $Z_q$ involves $X_q$ back to $X_{q-k}$, then there will be $k$ less $Z$'s than $X$'s, and nothing can be done about this. Any transversal filtration causes the loss of some data, and if the filter characteristic is complicated (as a

polynomial in cos $\omega$) the loss will have to be correspondingly great. This is usually unimportant, but, with very short pieces of record, might become crucial.

In Section 15 we remarked that an autoregressive relation, e.g.,

$$X_q = c_0 Z_q + c_1 Z_{q-1} + \cdots + c_k Z_{q-k},$$

between $X$'s and $Z$'s enabled us to obtain the reciprocal of any suitably non-negative cosine polynomial as the ratio of the spectrum of $Z$ to the spectrum of $X$. There are different ways of looking at the situation which make this statement true, not true, or partly true. If we have all the $X$'s back to $q = -\infty$, we can calculate the corresponding $Z$'s and it is true. If we have only a finite number of $X$'s, as always in practice, then we have to start the calculation up in some other way, perhaps like this

$$X_0 = c_0 Z_0,$$

$$X_1 = c_0 Z_1 + c_1 Z_0,$$

$$\vdots$$

$$X_{k-1} = c_0 Z_{k-1} + c_1 Z_{k-2} + \cdots + c_{k-1} Z_0,$$

$$X_k = c_0 Z_k + c_1 Z_{k-1} + \cdots + c_k Z_0.$$

As a consequence, we will have introduced an initial transient into the form of the autoregressive transformation so that our $Z$'s are never related to the $X$'s in the way we supposed. In this sense the statement is untrue. In many cases, however, this initial transient dies out quite quickly, and if we discard enough initial $Z$'s, perhaps $2k$ to $4k$ of them, we can regard the reciprocal of the cosine polynomial as a satisfactory approximation. In this sense the statement is partly true.

In theory, an autoregressive scheme corresponds to an infinitely long moving linear combination. In practice it corresponds to a sequence of changing moving linear combinations of finite but increasing length which approximate the infinitely long one. Sometimes the approximation is quite good enough. (Perhaps the main advantage to the autoregressive scheme is its likely reduction in arithmetic — a few autoregressive coefficients corresponding to a long moving linear combination.)

B.17 *Smoothing and Decimation Procedures*

We now study the effects of applying, successively and in any order, simple smoothing and decimation. The basic operations are taking

equally weighted means of $\ell$ consecutive values, and discarding all but every $j$th value. Simple sums are usually more convenient than means, and, since the results differ only by a fixed constant factor, lead to the same spectra except for a constant. We define, then, as our basic operations, $S_\ell$ and $F_j$, where (for definiteness)

$$Y = S_\ell X$$

means

$$Y_i = X_i + X_{i+1} + \cdots + X_{i+\ell-1}, \qquad (\ell \text{ terms}),$$

while

$$Z = F_j X$$

means

$$Z_i = X_{1+(i-1)j}.$$

It will also be convenient in dealing with the algebra of these operations, to use an operator for simple summing at wider spacings. We therefore define $S_\ell^{(h)}$ by taking

$$W = S_\ell^{(h)} X$$

to mean

$$W_i = X_i + X_{i+h} + \cdots + X_{i+(\ell-1)h}, \qquad (\ell \text{ terms}).$$

It is immediately clear that

$$S_\ell^{(h)} S_h = S_{\ell h} \qquad (\text{B-17.1})$$

both sides corresponding to forming sums of $\ell h$ consecutive $X_i$'s.

Now, consider the equation

$$Z = S_\ell F_j X,$$

which means that if

$$Y = F_j X,$$

then

$$Z = S_\ell Y.$$

These, in turn, mean that

$$Y_i = X_{1+(i-1)j},$$

and

$$Z_i = Y_i + Y_{i+1} + \cdots + Y_{i+\ell-1}.$$

Hence,

$$Z_i = X_g + X_{g+j} + \cdots + X_{g+(\ell-1)j},$$

where $g = 1 + (i - 1)j$. Now, since

$$W = S_\ell^{(j)} X$$

means

$$W_g = X_g + X_{g+j} + \cdots + X_{g+(\ell-1)j},$$

we have

$$Z_i = W_{1+(i-1)j},$$

which corresponds to

$$Z = F_j W.$$

Thus, we have shown that

$$S_\ell F_j = F_j S_\ell^{(j)}. \tag{B-17.2}$$

Similarly, we find that

$$S_\ell^{(h)} F_j = F_j S_\ell^{(jh)}. \tag{B-17.3}$$

The order in which $F_j$ and $S_\ell$ or $S_\ell^{(h)}$ are applied is thus important. On the other hand, it is easy to show that

$$S_\ell S_h = S_h S_\ell \neq S_{\ell h}, \tag{B-17.4}$$

and

$$F_j F_h = F_h F_j = F_{jh}. \tag{B-17.5}$$

Thus, sample reductions are

$$S_2 F_4 S_4 = F_4 S_2^{(4)} S_4 = F_4 S_8,$$

and

$$F_2 S_4 F_2 S_2 = F_2 F_2 S_4^{(2)} S_2 = F_4 S_8.$$

It may be noted also that $F_\ell S_\ell$ corresponds to "summing in (barely) non-overlapping blocks of $\ell$ terms".

A reason for these differing relations is easily found. The changes in the spectrum due to $S_\ell$ and $F_j$ are quite different in character. $S_\ell$ multiplies spectra by the power transfer function

$$| 1 + e^{-i\omega\Delta t} + \cdots + e^{-i(\ell-1)\omega\Delta t} |^2 = \left( \frac{\sin \dfrac{\ell\omega\Delta t}{2}}{\sin \dfrac{\omega\Delta t}{2}} \right)^2,$$

while $S_t^{(h)}$ multiplies spectra by

$$\left(\frac{\sin \dfrac{\ell h \omega \Delta t}{2}}{\sin \dfrac{h \omega \Delta t}{2}}\right)^2.$$

On the other hand, $F_j$ changes the folding frequency, $f_N$, to a new value $(1/j)^{\text{th}}$ as large as before, namely

$$f_N' = f_N/j,$$

and aliases together the old principal aliases in sets of $j$. The $j$ old principal aliases which have $f$ as their new principal alias are

$$f, 2f_N' - f, 2f_N' + f, 4f_N' - f, 4f_N' + f, \cdots (j \text{ terms}).$$

Our combined operations will, in general, change both amplitudes and principal aliases. If

$$P_{+A}(f) = T(f)P_A(f) + T(2f_N' - f)P_A(2f_N' - f)$$
$$+ T(2f_N' + f)P_A(2f_N' + f) + \cdots$$

where $P_{+A}(f)$ is the new aliased spectrum, $P_A(f)$ is the old aliased spectrum, $f_N'$ is the new folding (Nyquist) frequency, and the summation continues as long as the arguments lie below the old folding frequency, then we call $T(f)$, for any $f$ up to the old folding frequency, the *transmission*. (If there is no new aliasing, the transmission is merely the power transfer function.) The ratio of transmission for a desired frequency to the transmission for an undesired alias of that frequency, such as $T(f)/T(2f_N' - f)$ for example, will be called the *protection ratio*.

Fig. 20 illustrates some simple cases. Curve (b) is the transmission of $S_3$, $F_2 S_3$, or, indeed any $F_j S_3$. All of these are given by

$$\left(\frac{\sin \dfrac{3 \omega \Delta t}{2}}{\sin \dfrac{\omega \Delta t}{2}}\right)^2.$$

Scale (c) illustrates the reduced range of the principal aliases for $F_2$, $F_2 S_3$, or, indeed any $S_j F_2 S_k$. Curve (d) shows the power transfer function of $S_3$, regarded as following $F_2$. In terms of the respective folding frequencies, this curve duplicates curve (b). On an absolute frequency scale it is different. Curve (e) presents the transmission corresponding to $S_3 F_2 = F_2 S_3^{(2)}$ which aliases into curve (d).
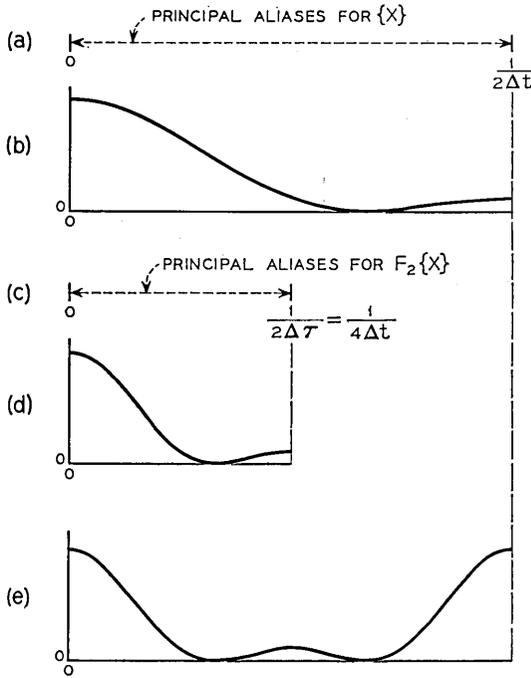
Fig. 20 — Transmission curves for a simple example of smoothing and decimation.

The transmission (= power transfer function) of $S_8$ is shown in Fig. 21.

The smoothing operation described in Section 17 was, in our present notation, $F_j S_k$ where $j = k/4$, which equals $S_4 F_j S_j$. In the case $j = 2$, we have $F_2 S_8 = S_4 F_2 S_2$, and the highlights of transmission and folding pattern are as follows, where the frequency scale has been chosen to make the original folding frequency = 8.

| Original freq. | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| Aliased freq. | 0 | 1 | 2 | 3 | 4 | 3 | 2 | 1 | 0 |
| Transmission | 64 | 26 | zero | 3.2 | zero | 1.4 | zero | 1.0 | zero |

The protection provided for frequency 1 against aliasing from frequency 7 is only in the ratio of 26 to 1. If the spectrum falls toward higher frequencies, this may be enough, but adequacy is far from certain.

Double use of $S_8$ before selecting every second observation would square this protection ratio, giving a ratio near 700. If we are to sum by groups twice, however, we can do better than to use the same length
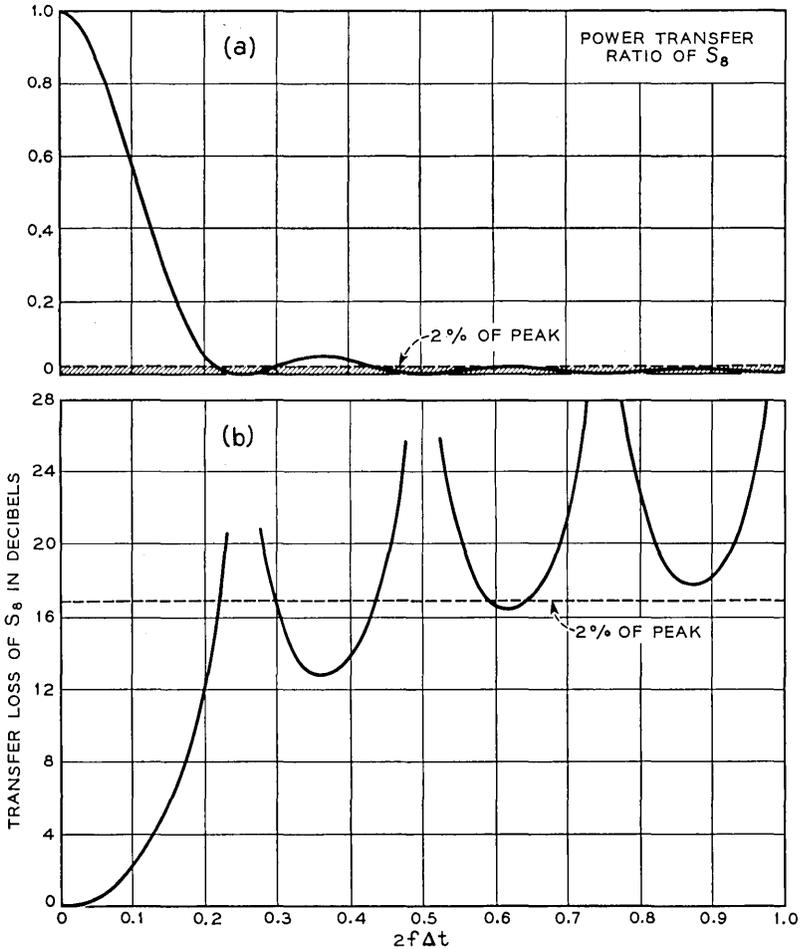
Fig. 21 — Transmission curves for $S_8$.

of group twice. By using slightly different lengths, we can spread out the zeros of the transmission curve, and tend to hold the right-hand end of the transmission curve nearer the origin. Table V shows the transmission and folding behavior of $F_3S_7S_8$, with the original folding frequency taken as 18. Over the lower half of the folded spectrum, this choice yields protection (against aliasing) by a ratio of about 3,000, which should suffice under even moderately extreme circumstances.

If we do not wish to fold quite so far (or in multiples of 2) then $F_2S_5S_6$ gives a protection ratio of 1,500 or better over the lower half of the folded spectrum.

TABLE V — TRANSMISSION AND FOLDING BEHAVIOR OF $F_3S_7S_8$
OR $F_3S_8S_7$ (ORIGINAL FOLDING FREQUENCY = 18)

| Original Frequency | Folded Frequency | Relative Transmission |
|:---:|:---:|:---:|
| 0 | 0 | 3136 |
| 1 | 1 | 2356 |
| 2 | 2 | 946 |
| 3 | 3 | 156 |
| 4 | 4 | 3.5 |
| 5 | 5 | 0.03 |
| 6 | 6 | 3.0 |
| 7 | 5 | 7.4 |
| 8 | 4 | 2.4 |
| 9 | 3 | zero |
| 10 | 2 | 0.04 |
| 11 | 1 | 0.39 |
| 12 | 0 | 1.0 |
| 13 | 1 | 0.17 |
| 14 | 2 | 0.09 |
| 15 | 3 | 0.06 |
| 16 | 4 | 0.12 |
| 17 | 5 | 0.28 |
| 18 | 6 | zero |

So long as we are satisfied with such, only moderately large, protection ratios, and with a substantial fall-off of transmission over the useful part of the spectrum, such repeated *unweighted* summing-or-averagings by groups are likely to be most desirable from the point of view of machine computation. If requirements on the smoothing-and-decimation operation are more stringent, smoothing with a suitably chosen set of graded weights is likely to be required.

B.18 *Modified Pilot Estimation, Cascade Estimation.*

The addition-and-subtraction calculation discussed in Section 18 (i) yields only one estimate per octave, (ii) is unduly sensitive to trends, (iii) involves windows which are broad even for a one-per-octave spacing, and (iv) is deliberately wasteful of data in the interest of computational simplicity. There may well be a desire to correct any or all of these.

The $\delta$ operator, in a notation extending that of the last section, takes the form $\delta = F_2D$, where $F_2$ is the operation of dropping every alternate value and $D$ is the operation of differencing adjoining values. We may replace $\delta$ by $D$ in the computing routine, keeping $\sigma$ the same, with the following effects: (i) loss of symmetry in the procedure, (ii) introduction of numerical factors like 8/15 (to be applied to the sums of squared differences), (iii) near doubling of number of differences available for

squaring and summing. If not enough data is available to give adequate stability to the results of the standard pilot method, then this modification may be worth while.

If we denote the operation of omitting the *other* half of the values by $F_2'$ and introduce $\sigma' = F_2'S_2$ as well as $\sigma = F_2S_2$, then, by using $D$, $\sigma$ and $\sigma'$ successively, we may obtain even more differences for squaring and summing. Table VI shows a convenient pattern of computation, in which sums and differences are *not* located in lines near the lines from which they are derived. Table IX, below, (in Section B.28) provides a numerical example.

TABLE VI — COMPACT CALCULATION ARRANGEMENT FOR COMPLETE VERSION OF PILOT ESTIMATION PROCEDURE

| $X$ | $DX$ | $\left\{{\sigma \atop \sigma'}\right\}X$ | $D\left\{{\sigma \atop \sigma'}\right\}X$ | $\left\{{\sigma \atop \sigma'}\right\}\left\{{\sigma \atop \sigma'}\right\}X$ | $D\left\{{\sigma \atop \sigma'}\right\}\left\{{\sigma \atop \sigma'}\right\}X$ | $\left\{{\sigma \atop \sigma'}\right\}\left\{{\sigma \atop \sigma'}\right\}\left\{{\sigma \atop \sigma'}\right\}X$ |
|---|---|---|---|---|---|---|
| $X_1$ |  | $SX_1$ |  | $SSX_1$ |  | $SSSX_1$ |
|  | $DX_1$ |  | $DSX_1$ |  | $DSSX_1$ |  |
| $X_2$ |  | $SX_3$ |  | $SSX_5$ |  |  |
|  | $DX_2$ |  | $DSX_3$ |  |  |  |
| $X_3$ |  | $SX_5$ |  |  |  |  |
|  | $DX_3$ |  | $DSX_5$ |  |  |  |
| $X_4$ |  | $SX_7$ |  | $SSX_3$ |  |  |
|  | $DX_4$ |  |  |  |  |  |
| $X_5$ |  |  |  |  |  |  |
|  | $DX_5$ |  |  |  |  |  |
| $X_6$ |  | $SX_2$ |  | $SSX_2$ |  |  |
|  | $DX_6$ |  | $DSX_2$ |  |  |  |
| $X_7$ |  | $SX_4$ |  |  |  |  |
|  | $DX_7$ |  | $DSX_4$ |  |  |  |
| $X_8$ |  | $SX_6$ |  | $SSX_4$ |  |  |

$$DX_q = X_{q+1} - X_q, \qquad SSX_q = (SX_{q+2}) + (SX_q)$$
$$SX_q = X_{q+1} + X_q, \qquad DSSX_q = (SSX_{q+4}) - (SSX_q)$$
$$DSX_q = (SX_{q+2}) - (SX_q) \quad SSSX_q = (SSX_{q+4}) + (SSX_q)$$

Two methods are available to cope with trend difficulties. The original series may be differenced, and the differenced series subjected to pilot estimation. (The main disadvantage being some loss in accuracy, etc., at very low frequencies.) Or each column of differences may be "corrected for its mean" adjusting the corresponding divisor from $k$ to $k - 1$. (This is only recommended when the first modification is in use and the column contains all $D$'s, not merely the corresponding $\delta$'s. Instead of

$$\frac{k + 1}{2k} \sum_1^k (D_q)^2$$

then, the corrected sum, corresponding to $(k + 1)/2$ squares, becomes

$$\frac{k + 1}{2(k - 1)} \left[ \sum_1^k (D_q)^2 - \frac{1}{k} \left(\sum_1^k D_q\right)^2 \right].$$

In view of the fact that $\sum D_q$ should equal the last value in the preceding column minus the first value there, a convenient check on the $D_q$ exists.

If, as may not be too unlikely, none of these modifications suffices, as will surely be the case if more than one estimate per octave is required, something better than the modified pilot estimation method, without going all the way to the detailed method's equi-spacing in frequency, may be desired. Such an intermediate method should give roughly constant spacing on a logarithmic scale, and provide reasonably clean windows, with about a 100-to-1 ratio between major lobes and minor lobes. It would be useful for high quality pilot estimation, and might sometimes suffice for the complete analysis.

Easy calculation and a roughly logarithmic scale both favor a cascade process, in each of whose cycles some computation is carried out on given values, and then half as many values are computed ready for the next cycle. We may expect, then, that it will be possible to think of each cycle as having three phases:

(a) computation of estimates

(b) smoothing of given values

(c) deletion of alternate smoothed values.

Our main attention needs to be given to the last two phases, since the first is a branch which can be changed rather freely.

We are concerned, therefore, with smoothing procedures to precede halving of the folding frequency. We have, then, to choose two frequencies averaging to the new folding frequency such that we plan to make estimates based on the smoothed (and halved) values up to the lower of these, while effectively eliminating frequencies above the higher. If the folding frequency at the start of the present cycle is $f_0$, these two frequencies can be written as $\alpha f_0$ and $(1 - \alpha)f_0$. In the next cycle, then, we anticipate estimation up to $\alpha f_0$. In the present cycle, we anticipate estimation up to $2\alpha f_0$ and must consequently cover the octave from $\alpha f_0$ to $2\alpha f_0$.

In the choice of $\alpha$ we must balance two considerations of computing effort. For a given number of lags (a given number of multiplications in forming mean lagged products per cycle) our estimates are spaced a fixed fraction of $f_0$. The larger $\alpha$ and $2\alpha f_0$, the more of these we may use. Contrariwise, the smaller $\alpha$, the easier it is, in terms of computational effort, to provide smoothing which suppresses the whole interval from $2\alpha f_0$ to $f_0$ by a factor of about 100 in comparison with what it does to any frequency between 0 and $\alpha f_0$.

Trial suggests that a value of $\alpha$ near $\frac{1}{3}$ is reasonable. We wish, therefore, a smoothing which suppresses frequencies between $(\frac{2}{3})f_0$ and $f_0$.

Two elementary smoothings with zeros in this range are running means-or-sums of 2 (with a zero at $f_0$) and running means-or-sums of 3 (with a zero at $(\frac{2}{3})f_0$). If, for convenience, we use sums, these operations multiply the spectrum by, respectively,

$$2 + 2 \cos \pi f/f_0,$$

and

$$3 + 4 \cos \pi f/f_0 + 2 \cos 2\pi f/f_0.$$

If both are applied, the factor takes on the following values:

| $f/f_0 =$ | 0.0 | 0.3 | $\frac{1}{3}$ | 0.5 | 0.6 | 0.62 | $\frac{2}{3}$ | 0.8 | 0.82 | $\frac{5}{6}$ | 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| factor = | 36 | 15 | 12 | 2 | 0.23 | 0.08 | 0 | 0.146 | 0.147 | 0.143 | 0 |

The result is a protective factor (against the effects of aliasing on halving) of just over 100 for $\alpha = 0.3$ and of about 80 for $\alpha = \frac{1}{3}$. This should be entirely satisfactory for most pilot purposes. (If further protection is needed, $Z_q = 0.6X_{q-1} + X_q + 0.6X_{q+1}$, which has zero transmission at $f/f_0 = 0.815$, could also be applied.)

If we wish to reduce the number of additions, we might smooth by threes, and then combine smoothing by twos and halving, e.g.,

$$X_q^* = X_{q-1} + X_q + X_{q+1},$$

$$Z_q = X_{2q}^* + X_{2q+1}^*,$$

which requires 2.5 additions per $X$-value. If we rearrange, however, to

$$\widetilde{X}_{2q} = X_{2q} + X_{2q+1},$$

$$2\widetilde{X}_{2q} = \widetilde{X}_{2q} + X_{2q},$$

$$Z_q = X_{2q-1} + 2\widetilde{X}_{2q} + X_{2q+2},$$

we find only 2 additions per $X$-value. Thus this type of smoothing and halving is computationally quite simple.

If we use this smoothing-and-halving process cycle after cycle we must, after obtaining our spectral estimates for the series actually processed in a given cycle, adjust them for the effects of all the smoothings in preceeding cycles. Near zero frequency this factor is $(36)^{-d}$ for $d$ previous smoothings. At a few selected frequencies the factors are as follows:

| $f/f_0 =$ | 0 | 0.075 | 0.15 | 0.225 | 0.30 |
|---|---|---|---|---|---|
| $j = 1$ | $(36)^{-1}$ | $1.05(36)^{-1}$ | $1.23(36)^{-1}$ | $1.60(36)^{-1}$ | $2.4(36)^{-1}$ |
| $j = 2$ | $(36)^{-2}$ | $1.06(36)^{-2}$ | $1.29(36)^{-2}$ | $1.81(36)^{-2}$ | $2.95(36)^{-2}$ |
| $j = 3$ | $(36)^{-3}$ | $1.07(36)^{-3}$ | $1.31(36)^{-3}$ | $1.86(36)^{-3}$ | $3.10(36)^{-3}$ |
| $j = 4$ | $(36)^{-4}$ | $1.07(36)^{-4}$ | $1.31(36)^{-4}$ | $1.87(36)^{-4}$ | $3.16(36)^{-4}$ |
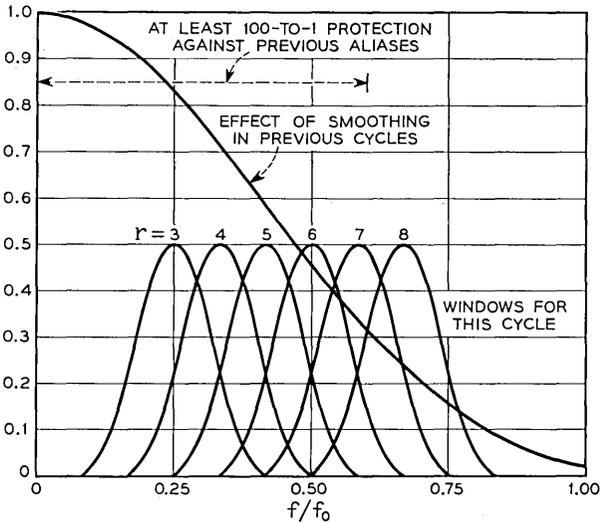
Fig. 22 — Windows for $m = 12$, and smoothing from previous cycles.

When an ordinary calculation with $m = 12$, for example, and hanning is used at each cycle, the spectral windows of the more relevant esti-mates and the effect of smoothing in previous cycles appear as in Fig. 22. Clearly the estimates for $r = 3, 4, 5$ and probably 6 will be quite usable. The estimate for $r = 7$ may well be usable, but its window extends up to a point where protection against aliasing is beginning to be much reduced. Computation for all $r$ from 0 to 12 is probably worthwhile, lower values of $r$ providing rough checks for later cycles and higher values indicating the extent of the danger from aliasing (to estimates of the next cycle) during the smooth-and-halve phase of the present cycle.

The resulting spectral windows are shown for parts of 4 cycles in Fig. 23. The windows 6, 5, 4, 3 $\sim$ 6', 5', 4', 3' $\sim$ 6'', 5'', 4'', 3'' $\sim$ 6''', 5''', 4''', 3''', and so forth, will give a fairly effective set of coarsely spaced spec-tral estimates.

B.19 *Rejection Near Zero Frequency*

We come now to the details of compensation, on the average, for non-zero but constant averages or for averages changing linearly with time. (As a convenient shorthand we will refer to these as "constant" and "linear" trends.) The $X_h$, on a sample of which our calculations are to be based, can each usefully be regarded as a sum of two terms: (i) a fixed (but unknown) trend and (ii) a sample from a stationary ensemble with

average zero. These terms are added together and are statistically inde-
pendent (since one is fixed).
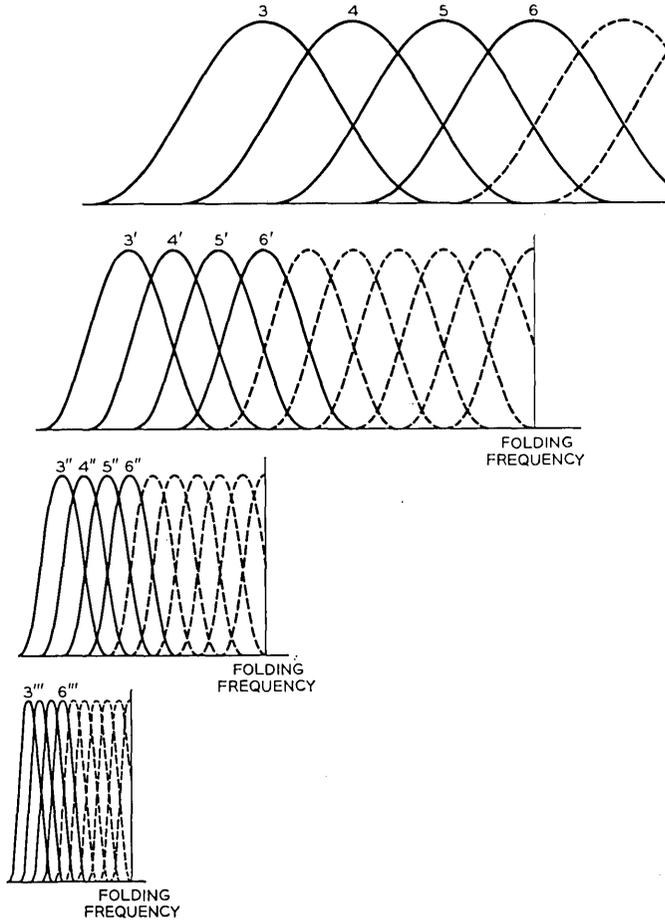


Fig. 23 — Spectral windows in successive cycles (smoothing corrected for).

Consider any quadratic expression in the $X_h$. If we introduce the two
terms for each $X_h$, we may write the quadratic expression as a sum of
three parts: (i) a quadratic in the trend (fixed, of course), (ii) an ex-
pression linear both in the trend and in the stationary fluctuations (and
hence of average value zero), and (iii) a quadratic in the stationary
fluctuations. The first and last of these three parts are, of course, just
what would have arisen had the trend alone or, respectively, the fluctua-
tions alone been present. Hence, since the average value of the middle

part vanishes, the average value of the whole quadratic is the sum of the average value for the trend alone and the average value for the fluctuations alone. To study the efficacy with which a quadratic expression rejects a trend in the presence of fluctuation, *we have only to study its behavior in the presence of trend alone.* After this, we shall want to study the behavior for fluctuations alone of those expressions which satisfactorily reject trend.

(The variance of the quadratic expression will be determined from the middle and last parts. If third moments of the fluctuations vanish, as will be the case for Gaussian ensembles, the contributions will come from these parts separately, the covariance between parts vanishing.)

A few formulas will be useful in discussing possible rejection techniques. These are conveniently derived on the basis of $n = 2k + 1$ available equi-spaced values extending from $-k$ through 0 to $+k$. (The indexing of the values is only a convenience, and cannot affect any essential result. The limitation to an odd number of points is essentially a convenience also — we shall replace $2k + 1$ by an unrestricted $n$ rather freely.)

The first formulas relate to constants and linear trends, and are as follows:

$$\text{ave } \{C_r \mid X_t \equiv 1\} = \frac{1}{2k + 1 - r} \sum_{-k}^{k-r} (1)(1) = 1,$$

$$\text{ave } \{C_r \mid X_t \equiv t\} = \frac{1}{2k + 1 - r} \sum_{-k}^{k-r} h(h + r)$$

$$= \frac{1}{6} (2k(k + 1) - r(2k + 1) - r^2),$$

$$= \frac{1}{12} n^2 K_r,$$

where

$$K_r = 1 - \frac{1}{n^2} - 2 \frac{r}{n} - 2 \frac{r^2}{n^2} \approx 1,$$

whence,

$$\text{ave } \{C_r \mid X_t \equiv \alpha + \beta t\} = \frac{1}{2k + 1 - r} \sum_{-k}^{k-r} (\alpha + \beta h)(\alpha + \beta h + \beta r)$$

$$= \alpha^2 + \frac{\beta^2}{12} n^2 K_r,$$

while

$$\frac{1}{c + d + 1} \sum_{-c}^{d} (\alpha + \beta h) = \alpha + \beta \frac{d - c}{2},$$

$$\frac{1}{a + b + 1} \sum_{-a}^{b} \frac{1}{c + d + 1} \sum_{-c}^{d} (\alpha + \beta h)(\alpha + \beta g)$$

$$= \alpha^2 + \alpha\beta \left( \frac{d - c + b - a}{2} \right) + \beta^2 \frac{(d - c)(b - a)}{4}.$$

From the first two formulas, as combined in the fourth, we learn what dependence on $\alpha$ and $\beta$ is required if we are to reject both constants and linear trends for any fixed value of $r$. The fifth formula, as developed into the last, shows that this cannot be done by subtracting the product of any two simple, equally weighted means of the $X$'s.

Suppose now that $k = 3j - 2$, so that the means of the lower one-third of all, of the upper one third of all, and of all $X$'s are, respectively

$$\overline{X^-} = \frac{1}{2j - 1} \sum_{-k}^{-j} X_h,$$

$$\overline{X^+} = \frac{1}{2j - 1} \sum_{j}^{k} X_h,$$

$$\bar{X} = \frac{1}{6j - 3} \sum_{-k}^{k} X_h,$$

and so that, when $X_t = \alpha + \beta t$, we have

$$\overline{X^-} = \alpha - \beta \left( \frac{k + j}{2} \right) = \alpha - (2j - 1)\beta = \alpha - \frac{n}{3} \beta,$$

$$\overline{X^+} = \alpha + \beta \left( \frac{k + j}{2} \right) = \alpha + (2j - 1)\beta = \alpha + \frac{n}{3} \beta,$$

$$\bar{X} = \alpha.$$

Thus if we wish to reject a constant we may choose

$$E_{0r} = (\bar{X})^2 \text{ (independent of } r\text{)}$$

and have

$$\text{ave } \{C_r \mid X_t \equiv \alpha\} \equiv \text{ave } \{E_{0r} \mid X_t \equiv \alpha\}.$$

This choice, however, produces no rejection of $X_t \equiv \beta t$ at all.

If we wish to reject either a constant or a linear trend or both in a

simple way, we may choose

$$E_{1r} = \bar{X}^2 + \frac{3}{16} K_r \cdot (\overline{X^+} - \overline{X^-})^2$$

for which

$$\text{ave } \{C_r \mid X_t \equiv \alpha + \beta t\} \equiv \text{ave } \{E_{1r} \mid X_t \equiv \alpha + \beta t\}.$$

An alternate calculation, nearly equivalent to the use of $E_{0r}$ is illustrated by

$$\frac{1}{2k + 1 - r} \sum_{-k}^{k-r} (X_h - \bar{X})(X_{h+r} - \bar{X})$$

$$= \frac{1}{2k + 1 - r} \sum_{-k}^{k-r} X_h X_{h+r} - \bar{X}^2$$

$$- \bar{X} \left( \frac{r\bar{X} - \sum_{k-r+1}^{k} X_h}{2k + 1 - r} + \frac{r\bar{X} - \sum_{-k}^{-k+r-1} X_h}{2k + 1 - r} \right)$$

$$= \frac{1}{2k + 1 - r} \sum_{-k}^{k-r} X_h X_{h+r} - \bar{X}^2 - \bar{X} Q_r,$$

where

$$Q_r = \frac{1}{2k + 1 - r} \left[ 2r\bar{X} - \sum_{k-r+1}^{k} X_h - \sum_{-k}^{-k+r-1} X_h \right].$$

Thus subtracting the mean of all the observations from each observation before forming the $C_r$'s is equivalent to using

$$E_{2r} = \bar{X}^2 + \bar{X} Q_r.$$

Similarly, fitting and subtracting the elementary-least-squares straight line corresponds to

$$\hat{\beta} = 3 \frac{\sum_{-k}^{k} h X_h}{k(k + 1)(2k + 1)} = \frac{12}{n(n^2 - 1)} \sum_{-k}^{k} h X_h,$$

and

$$\frac{1}{2k + 1 - r} \sum_{-k}^{k-r} (X_h - \bar{X} - \hat{\beta}h)(X_{h+r} - \bar{X} - \hat{\beta}(h + r))$$

$$= \frac{1}{2k + 1 - r} \sum_{-k}^{k-r} X_h X_{h+r} - \bar{X}^2 - \bar{X} Q_r$$

$$+ \frac{\hat{\beta}}{n - r} \sum_{k-r+1}^{k} (h + r) \cdot (X_h - X_{-h}) - \frac{n^2 \hat{\beta}^2}{12} \left[ \frac{2(n^2 - 1)}{n(n - r)} - K_r \right].$$

Thus linear fitting corresponds to the use of

$$E_{3r} = \bar{X}^2 + \bar{X}Q_r - \frac{\hat{\beta}}{n - r} \sum_{k-r+1}^{k} (h + r) \cdot (X_h - X_{-h})$$
$$+ \frac{n^2\hat{\beta}^2}{12} \left[ \frac{2(n^2 - 1)}{n(n - r)} - K_r \right].$$

and if we wish to simplify computation by neglecting "end corrections", consider

$$E^*_{3r} = \bar{X}^2 + \bar{X}Q_r + \frac{n^2\hat{\beta}^2}{12} K_r.$$

Another version can be obtained by recalling the usual formulation of a sample estimate of a covariance between two unlagged variates. If we write

$$\overline{X^-}_{(r)} = \frac{1}{2k + 1 - r} \sum_{-k}^{k-r} X_h$$

$$\overline{X^+}_{(\cdot)} = \frac{1}{2k + 1 - r} \sum_{-k}^{k-r} X_{h+r}$$

then we are led to use the expression

$$\frac{1}{2k + 1 - r} \sum_{-k}^{k-r} (X_h - \overline{X^-}_{(r)})(X_{h+r} - \overline{X^+}_{(r)})$$
$$= \frac{1}{2k + 1 - r} \sum_{-k}^{k-r} X_h X_{h+r} - \overline{X^-}_{(r)} \overline{X^+}_{(r)}$$

which corresponds to the use of

$$E_{4r} = \overline{X^-}_{(r)} \overline{X^+}_{(r)}$$

or, in the case of a linear trend, perhaps, to

$$E_{5r} = \overline{X^-}_{(r)} \overline{X^+}_{(r)} + \frac{n^2\hat{\beta}^2}{12} \left[ K_r + \frac{3r^2}{n^2} \right].$$

(The $E$'s corresponding to fitting a separate straight line to $X_{-k}$, $\cdots$, $X_{k-r}$ and to $X_{-k+r}$, $\cdots$, $X_k$ do not seem worth writing down.)

Having now obtained a collection of quadratic expressions which, when subtracted from the corresponding $C_r$, remove the average effect of trend, *we have now to learn what effect these subtractions will have on the contribution of the fluctuations to the average values of the corresponding quantities.* As noted at the beginning of this section, it suffices to consider pure fluctuations, so that we need only study the

spectral windows corresponding to the $E_{kr}$, and to their finite cosine transforms.

If we consider white noise (of unit variance) for a moment, the $X_h$ become independent with average zero and unit variance and we can calculate the average values of the various $E_{kr}$ in an elementary way. This is of interest, because these average values for white noise are exactly proportional to the integrals of the corresponding spectral windows. We find the following results:

| Quadratic expression | Average value for unit white noise | Approximation to this |
|:---:|:---:|:---:|
| $E_{0r}$ | $\dfrac{1}{n}$ | $\dfrac{1}{n}$ |
| $E_{1r}$ | $\dfrac{1}{n}\left(1 + \dfrac{9K_r}{8}\right)$ | $\sim \dfrac{2}{n}$ |
| $E_{2r}$ | $\dfrac{1}{n}$ | $\dfrac{1}{n}$ |
| $E_{3r}^{*}$ | $\dfrac{1}{n}\left(1 + \dfrac{n^2}{n^2 - 1}K_r\right)$ | $\sim \dfrac{2}{n}$ |
| $E_{4r}$ | $\dfrac{n - 2r}{(n - r)^2}$ | $\sim \dfrac{1}{n}$ |
| $E_{5r}$ | $\dfrac{n - 2r}{(n - r)^2} + \dfrac{n}{n^2 - 1}\left(K_r + \dfrac{3r^2}{n^2}\right)$ | $\sim \dfrac{2}{n}$ |

The quadratic expressions which eliminate the average effect of a constant trend (one constant) have spectral windows integrating approximately proportional to $1/n$, while those which compensate for general linear trends (two constants) integrate approximately proportional to $2/n$. This is simple, and seems straightforward.

It is possible, however, to eliminate the average effect of a centered linear trend (one with $\alpha = 0$) by subtracting a windowless quadratic expression, one whose average vanishes for all *stationary* ensembles. We need only consider

$$X_{-2}X_0 - 2X_{-1}X_1 + X_0X_2$$

whose average value clearly vanishes under stationarity, but whose value when $X_t \equiv \alpha + \beta t$ is $2\beta^2$, to see that this is so. (Cursory inquiry into variability suggests that such use of windowless quadratics may

increase the variability of the finally resulting spectral density estimates.)

Leaving such possibilities aside, let us compare the $E_{kr}$ for $k = 0, 1, 2, 3, 4, 5$. The corresponding spectral windows can be written down with the aid of the formulas of Section A.6, but as they are rather complicated, we shall avoid doing so. The three choices eliminating constant trends, $E_{0r}$, $E_{2r}$ and $E_{4r}$, all have integrated spectral windows approximating $1/n$. The dependence on $r$ is in any case weak, and where present increases with $r$. For $E_{0r}$ there is no dependence on $r$.

We are, of course, more concerned with the spectral windows associated with the modified $V_r$'s rather than with those associated with the modifying $E_{kr}$. If we use $E_{0r}$, then the change in spectral window is the same for each lag. Consequently, only the $R_{00}(f)$ window is affected, the $R_{cr}(f)$ windows corresponding to the other $V_r$ vanish. The situation is more complicated for the other cases, and no clear advantage over the use of $E_{0r}$ appears.

The spectral window corresponding to $\bar{X}^2$ is, of course, $(\text{dif } nf / \text{dif } f)^2$ and is both always positive and well concentrated near zero. (If we replaced $\bar{X}$ by the average over a graded data window we could decrease the corresponding spectral window for $f$ beyond the first side lobe of dif $nf$, but the concomitant broadening of the main lobe would make the result much less useful.) Clearly the use of $E_{0r}$ will be quite satisfactory.

Comparing $E_{1r}$, $E_{3r}$ and $E_{5r}$ is not so simple. $E_{3r}^*$ has the smaller integrated spectral window for $r/n$ small, but the simplicity of calculation of $E_{1r}$ will often outweigh this fact. (If economizing on the spectral window is important, we could use a windowless quadratic to eliminate the average effect of $\beta$.) Accordingly, the use of $E_{1r}$ is recommended, unless it is simpler to subtract a fitted linear function from all $X_t$.

(There is no Section B.20.)

## B.21 *Sample Computing Formulas*

Only the formulas for correction for prewhitening and for correction for the mean require discussion. The factor for $1 \leqq r \leqq m - 1$ is exactly the reciprocal of the prewhitening transfer function, calculated at the nominal frequency of the estimate. The factor for $r = m$, and the main portion of that for $r = 0$, differ only in the selection of the frequency at which the prewhitening transfer function is evaluated. Since, in each case, the nominal frequency is at one edge of the band of frequencies covered, the frequency of evaluation was displaced from the nominal frequency toward the center of the corresponding band. The choice of a point one-third of the way across the band was somewhat arbitrary.

Finally, there is the question of compensation for the correction for the mean. The raw estimate for $r = 0$ would naturally be thought of as corresponding to the interval from $-1/(4m)$ to $+1/(4m)$ cycles per observation, while the hanned estimate would cover from $-1/(2m)$ to $+1/(2m)$ cycles per observation. Since $n$ degrees of freedom are associated with the entire interval from $-\frac{1}{2}$ to $+\frac{1}{2}$ cycle per observation, the hanned estimate for $r = 0$ is associated with $n/m$ degrees of freedom. One of these has been eliminated by correction for the mean, as would also have been the case had we used $E_{2r}$ or $E_{4r}$, so that we need to compensate for a reduction in the ratio

$$\frac{\dfrac{n}{m} - 1}{\dfrac{n}{m}} = \frac{n - m}{n}$$

whose reciprocal is the first compensating factor for $r = 0$. (Had we used $E_{1r}$, $E_{3r}$, or $E_{5r}$, we would have had to compensate for the loss of two degrees of freedom by a factor $n/(n - 2m)$.)

(There is no Section B.22.)

## DETAILS FOR PLANNING

B.23 *Duration Requirement Formulas*

We are now in a position to assemble and modify formulas from a number of sections as a basis for formulas expressing explicit requirements. In the process we shall have to give explicit definitions for certain concepts. The first of these is *resolution*. If we hann or hamm, we obtain estimates every $1/(2T_m)$ cps. Adjacent estimates have very considerably overlapping windows, and consequently the estimates have substantially related sampling fluctuations and refer to overlapping frequency regions. It would be a clear mistake to consider these estimates as completely resolved. When we come to next-adjacent estimates, however, the situation is quite different. The overlap is small, the covariance being about 5 per cent of either variance for a moderately flat spectrum. We shall consequently treat such pairs of estimates as completely resolved, and place

$$(\text{resolution in cps}) = 2\frac{1}{2T_m} = \frac{1}{T_m \text{ in seconds}}.$$

We can express the stability, so far most often expressed as the number of elementary frequency bands or equivalent degrees of freedom

associated with each estimate, in terms of the spread in db of an interval containing, with prescribed probability, the ratio of true smoothed power to estimated smoothed power. Reference to Table II in Section 9 shows the inter-relations to be, approximately

$$k = 1 + \frac{250}{(80 \text{ per cent range in db})^2},$$

$$k = 1 + \frac{400}{(90 \text{ per cent range in db})^2},$$

$$k = 1 + \frac{625}{(96 \text{ per cent range in db})^2},$$

$$k = 1 + \frac{840}{(98 \text{ per cent range in db})^2}.$$

We shall write our combined formulas in terms of the 90 per cent range. Formulas for other per cent ranges are easily obtained by replacing 400 by the appropriate constant. It must be emphasized that when we use a 90 per cent range we only have 9 chances in 10 of finding *each* individual estimate correspondingly close to its average value and that if we have, say, 30 estimates, we are quite sure that at least one will be more discrepant than this.

We recall that we adopted (see end of Section 6 of Part I)

$$T'_n = (\text{total length of record}) - \frac{p}{3} T_m,$$

where $p$ was the number of pieces, and, for design purposes.

$$k \approx \frac{2T'_n}{T_m}.$$

Hence,

$$(\text{duration in seconds}) = T_n = \left(\frac{T'_n}{T_m} + \frac{p}{3}\right) T_m$$

$$= \left(\frac{k}{2} + \frac{p}{3}\right) \frac{1}{(\text{resolution in cps})}$$

$$= \left(\frac{1}{2} + \frac{200}{(90 \text{ per cent range in db})^2} + \frac{(\text{pieces})}{3}\right) \Big/ (\text{resolution in cps}).$$

Writing (length of each piece)·(number of pieces) for (duration) and

solving for the number of pieces yields

$$\text{(number of pieces)} = \frac{\dfrac{1}{2} + \dfrac{200}{(90 \text{ per cent range in db})^2}}{(\text{length of each piece})(\text{resolution in cps}) - \frac{1}{3}}.$$

These relations are general, and apply equally to analog processed continuous records or digitally processed equi-spaced records, provided the spectrum is, or can be prewhitened to be, reasonably flat.

B.24 *Digital Requirement Formulas*

If we are to use equi-spaced digital analysis, if we can provide frequency cutoff easily, and if we need to cover frequencies up to some $f_{\max}$, then we can probably take our folding frequency at about $\frac{3}{2} f_{\max}$. The necessary number of lags $m$ then follow from

$$\Delta t = \frac{1}{3 f_{\max}} = \frac{1}{2 f_N},$$

$$m = \frac{T_m}{\Delta t} = 3 T_m f_{\max}.$$

The necessary number of data points follows from

$$n = \frac{T_n}{\Delta t} = \left( T'_n + \frac{p}{3} T_m \right) 3 f_{\max}$$

$$= 3 T'_n f_{\max} + p T_m f_{\max}$$

$$= \left( 3 \frac{T'_n}{T_m} + p \right) (T_m f_{\max}),$$

and the rough number of multiplications is

$$mn = 9 T_m T'_n (f_{\max})^2 + 3p(T_m f_{\max})^2.$$

The quantity $T_m f_{\max}$ can be written as

$$\frac{f_{\max}}{1/T_m} = \frac{\text{maximum frequency}}{\text{resolution}} = \text{(number of resolved bands)},$$

so that the number of data points becomes

$$n = \left( 3 \frac{T'_n}{T_m} + p \right) \text{(number of resolved bands)}$$

$$= (1.5k + p) \text{ (number of resolved bands)},$$

TABLE VII

Second differences of Brouwer's data and the add-and-subtract pilot estimation process as applied to them. (Block 2 only.)

| Date | 10F* | diff. | 2nd diff. = $X_q$ | $\delta$ | $\sigma$ | $\delta\sigma$ | $\sigma\sigma$ | $\delta\sigma\sigma$ | $\sigma\sigma\sigma$ | $DX_q$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1853.5 | 56 | | | | | | | | | |
| | | 0 | | | | | | | | |
| 1854.5 | 56 | | −4 | | | | | | | |
| | | −4 | | 4 | −4 | | | | | 4 |
| 1855.5 | 52 | | 0 | | | | | | | |
| | | −4 | | | | 7 | −1 | | | −2 |
| | 48 | | −2 | | | | | | | |
| | | −6 | | 7 | 3 | | | | | 7 |
| | 42 | | +5 | | | | | | | |
| | | −1 | | | | | | −2 | −4 | −9 |
| | 41 | | −4 | | | | | | | |
| | | −5 | | 11 | 3 | | | | | 11 |
| | 36 | | +7 | | | | | | | |
| | | +2 | | | | −9 | −3 | | | −30 |
| 1860.5 | 38 | | −23 | | | | | | | |
| | | −21 | | 40 | −6 | | | | | 40 |
| | 17 | | +17 | | | | | | | |
| | | −4 | | | | | | | | −19 |
| | 13 | | −2 | | | | | | | |
| | | −6 | | −4 | −8 | | | | | −4 |
| | 7 | | −6 | | | | | | | |
| | | −12 | | | | 18 | 2 | | | +7 |
| | −5 | | +1 | | | | | | | |
| | | −11 | | +8 | 10 | | | | | +8 |
| 1865.5 | −16 | | +9 | | | | | | | |
| | | −2 | | | | | | −23 | −19 | −20 |
| | −18 | | −11 | | | | | | | |
| | | −13 | | +13 | −9 | | | | | +13 |
| | −31 | | +2 | | | | | | | |
| | | −11 | | | | −3 | −21 | | | −1 |
| | −42 | | +1 | | | | | | | |
| | | −10 | | −14 | −12 | | | | | −14 |
| | −52 | | −13 | | | | | | | |
| | | −23 | | | | | | | | +15 |
| 1870.5 | −75 | | +2 | | | | | | | |
| | | −21 | | +7 | 11 | | | | | +7 |
| | −96 | | +9 | | | | | | | |
| | | −12 | | | | −17 | 5 | | | −20 |
| | −108 | | −11 | | | | | | | |
| | | −23 | | +16 | −6 | | | | | +16 |
| | −131 | | +5 | | | | | | | |
| | | −18 | | | | | | 4 | 14 | +4 |
| | −149 | | +9 | | | | | | | |
| | | −9 | | −9 | 9 | | | | | −9 |
| 1875.5 | −158 | | 0 | | | | | | | |
| | | −9 | | | | −9 | 9 | | | +4 |
| | −167 | | +4 | | | | | | | |
| | | −5 | | −8 | 0 | | | | | −8 |
| | −172 | | −4 | | | | | | | |
| | | −9 | | | | | | | | +8 |
| | −181 | | +4 | | | | | | | |
| | | −5 | | −10 | −2 | | | | | −10 |
| | −186 | | −6 | | | | | | | |
| | | −11 | | | | 6 | 2 | | | +15 |

TABLE VII, CONTINUED

| Date | 10F* | diff. | 2nd diff. = $X_q$ | $\delta$ | $\sigma$ | $\delta\sigma$ | $\sigma\sigma$ | $\delta\sigma\sigma$ | $\sigma\sigma\sigma$ | $LX_q$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1880.5 | −197 | | +9 | | | | | | | |
| | | −2 | | −14 | 4 | | | | | −14 |
| | −199 | | −5 | | | | | | | |
| | | −7 | | | | | | 1 | 5 | −6 |
| | −206 | | −11 | | | | | | | |
| | | −18 | | +33 | 11 | | | | | +33 |
| | −224 | | +22 | | | | | | | |
| | | +4 | | | | −19 | 3 | | | −30 |
| | −220 | | −8 | | | | | | | |
| | | −4 | | +8 | −8 | | | | | +8 |
| 1885.5 | −224 | | 0 | | | | | | | |
| | | −4 | | | | | | | | |
| 1886.5 | −228 | | | | | | | | | |

*(Brouwer's notation) is the fluctuation in the earth's rotation. Here 10F is the fluctuation expressed in tenths of seconds of time.

and the rough number of multiplications becomes

$$nm = (4.5k + 3p)(\text{number of resolved bands})^2$$

$$= \left[\frac{9}{2} + \frac{1800}{(90 \text{ per cent range in db})^2} + 3(\text{pieces})\right]$$

$$\cdot (\text{number of resolved bands})^2.$$

These last formulas assume satisfactory shaping of frequency cutoff, and constant resolution up to the maximum frequency of interest. Particular situations may deviate from this in either direction.

(There are no Sections B.25, B.26, B.27.)

B.28. *Analysis of Example C*

It is most desirable that an account of this sort include enough details of a numerical example to allow those readers who wish to do so, to follow through and check. This is impractical when even a few tenths (or even a few hundredths) of a million multiplications are involved. Example C, however, offers us an opportunity to present such details for one example, even if it is quite atypical.

We remarked in Section 28 that the add-and-subtract pilot estimation procedure of Section 18 *might* be applied to the second differences of Brouwer's values (themselves the differences ephemeris time *minus* mean solar time). During the 131 years from 1820 to 1950 astronomical techniques have improved, and observational errors seem, according to Brouwer's own analysis, to have somewhat decreased. In

TABLE VIII — SUMS OF SQUARED DIFFERENCES

|  | $\delta$ column | $\delta\sigma$ column | $\delta\sigma\sigma$ column | (*) | (†) |
|---|---|---|---|---|---|
| Block 1 | 8615 | 2083 | 321 | −502 | −971 |
| Block 2 | 4130 | 1230 | 550 | −9 | −70 |
| Block 3 | 3447 | 2591 | 273 | 1557 | −244 |
| Block 4 | 1553 | 304 | 516 | −162 | 283 |
| Noise | 5/2 | 3/4 | 3/8 | 0 | 0 |
| Flat | 1 | 1 | 1 | 0.7 | 0.85 |

(*) Sum for $\delta\sigma$ column less 0.3 times sum for $\delta$ column.
(†) Sum for $\delta\sigma\sigma$ column less 0.15 times sum for $\delta$ column.

order to reflect this fact, and to provide some external estimate of error we shall study the second differences in 4 separate blocks, 4 separate time intervals of 32 consecutive years each. In view of the fact that Brouwer estimates the mean observational error to have decreased from 0.38 to 0.17 seconds of time over the period, it will clearly suffice to work in units of 0.1 second of time. Table VII shows, for the second time block, calendar dates, Brouwer's values, their first and second differences (the latter being the series we consider as the $X_q$) and the results of the add-and-subtract pilot estimation model. In this table, sums and differences are shown on lines half-way between the lines containing the entries from which they are formed. The last column shows the result of completing the $\delta X_q$ column to a $DX_q$ column. The resulting sums of squared differences are shown in Table VIII for each of the four blocks, together with comparative values.

These comparative values show the anticipated relative sizes of such sums of squared differences in case the spectral density were

(1) proportional to $(1 - \cos \pi f/f_N)^2 = (1 - u)^2$, the shape assumed for the "noise" component,

(2) flat, the shape assumed for the other component according to the second model,

where we have introduced $u$ as an abbreviation for $\cos \pi f/f_N$. The successive columns have average sums of squared differences which are easily seen to be obtained by multiplication of the spectrum by (when we start with 32 $X$'s):

$$16 \,(2 - 2 \cos \pi f/f_N) = 32 \,(1 - u),$$

$$8 \,(2 + 2 \cos \pi f/f_N) \,(2 - 2 \cos 2\pi f/f_N) = 64(1 + u - u^2 - u^3)$$
$$= 64 \,(1 + u) \,(1 - u^2),$$

$$4 \,(2 + 2u) \,(4u^2) \,(16 \, u^2 - 16 \, u^4) = 512(u^4 + u^5 - u^6 - u^7),$$

and integration.

When we recall that

$$\int_0^{f_N} u^k \, df$$

vanishes for odd $k$, and is equal to $f_N$, $f_N/2$, $3f_N/8$, $5f_N/16$ and $35f_N/128$ respectively, for $k = 0, 2, 4, 6$ and 8, we easily obtain the values given for comparison in Table VIII (after removing $32f_N$ as a common factor).

The last two columns of this table represent attempts to combine sums of squared differences so as to estimate the first component free of the "noise" component whose spectrum is proportional to $(1 - u)^2$. The attempts appear quite useless.

One reason for the lack of success is easy to find. There are only 4 values of $\delta\sigma\sigma$ to square and sum for each 32-year block. This means no more than 4 degrees of freedom, and, consequently very poor stability. We can partially correct this difficulty by going over to the modified add-and-subtract method in which all possible differences of a given sort are calculated.

Table IX presents the compact calculation for block 1, arranged as in Table VI. Here $(\sigma)$ stands for first $\sigma$ and then $\sigma'$. The mean squares for each column of differences lead to the results summarized in Table X for all four blocks. The difference estimates are now seen to be more stable and to increase somewhat with decreasing frequency, although not as much as would be expected even for a flat component.

The analysis is doing better, but is not yet satisfactory. One likely reason for this appears when we inquire what sort of spectral windows go with (\*), (†) and (‡). The windows corresponding to $D(\sigma)$ and to (\*) are shown in Fig. 24. The amount of negative area near $f/f_N = 1$ required to compensate for the rather large pickup of the "noise" component by the $D(\sigma)$ column is quite substantial, suggesting probable increased variability.

We can reduce our difficulties from such causes by using only slightly more complex processes. We can probably use the $D$ column satisfactorily as an indication of the noise component.

We need to obtain two other composite measures of the spectrum, both of which avoid large values of $f/f_N$, one of which is concentrated near $f/f_N = 0$ and the other of which avoids $f/f_N = 0$. If we can obtain a smoothed and decimated sequence which avoids the upper part of the spectrum, then sums and differences of such values will have mean squares with the appropriate properties. The results of Section B.18 suggest trying $F_2S_2S_3$ as the operation generating the modified sequence, to which $D$ (differencing) and $S_2$ are then to be applied.

## TABLE IX

Compact calculation applied to second differences (in tenths of seconds of time) for block 1 (values in Brouwer's Table VIII(c) used where appropriate). (Entries arranged as in Table VI.)

| $X_q$ | $D$ | $(c)$ | $D(\sigma)$ | $(\sigma)^2$ | $D(\sigma)^2$ | $(\sigma)^3$ | $D(\sigma)^3$ | $(\sigma)^4$ | $D(\sigma^4)$ |
|---|---|---|---|---|---|---|---|---|---|
| −27 | | −6 | | −10 | | −10 | | −9 | |
| | +48 | | 2 | | 10 | | 11 | | 13 |
| +21 | | −4 | | 0 | | +1 | | 4 | |
| | −27 | | 4 | | −6 | | −5 | | |
| −6 | | 0 | | −6 | | −4 | | | |
| | +8 | | 0 | | 13 | | 12 | | |
| +2 | | 0 | | 7 | | 8 | | −3 | |
| | −8 | | −1 | | −12 | | | | |
| −6 | | −1 | | −5 | | | | | |
| | +12 | | −4 | | 6 | | | | |
| +6 | | −5 | | 1 | | −6 | | −4 | |
| | +2 | | 6 | | 1 | | 8 | | |
| +8 | | 1 | | 2 | | 2 | | | |
| | −16 | | 5 | | 4 | | 1 | | |
| −8 | | 6 | | 6 | | 3 | | 5 | |
| | +1 | | −10 | | | | | | |
| −7 | | −4 | | | | | | | |
| | +13 | | 3 | | | | | | |
| +6 | | −1 | | −4 | | −5 | | −7 | |
| | 0 | | 3 | | 3 | | 3 | | |
| +6 | | 2 | | −1 | | −2 | | | |
| | −17 | | −3 | | −3 | | 24 | | |
| −11 | | −1 | | −4 | | 22 | | 20 | |
| | +16 | | 23 | | 6 | | | | |
| +5 | | 22 | | 2 | | | | | |
| | −9 | | −42 | | −1 | | | | |
| −4 | | −20 | | 1 | | −5 | | −2 | |
| | +8 | | 15 | | 20 | | 8 | | |
| +4 | | −5 | | 21 | | 3 | | | |
| | −2 | | 16 | | −46 | | −7 | | |
| +2 | | 11 | | −25 | | −4 | | −1 | |
| | +6 | | | | | | | | |
| +8 | | | | | | | | | |
| | −20 | | | | | | | | |
| −12 | | 15 | | 11 | | 10 | | 26 | |
| | +23 | | −19 | | −12 | | 6 | | |
| +11 | | −4 | | −1 | | 16 | | | |
| | −23 | | 18 | | 7 | | −23 | | |
| −12 | | 14 | | 6 | | −7 | | 9 | |
| | +31 | | −29 | | 4 | | | | |
| +19 | | −15 | | 10 | | | | | |
| | −36 | | 27 | | −4 | | | | |
| −17 | | 12 | | 6 | | 5 | | 21 | |
| | +27 | | −18 | | −19 | | 11 | | |
| +10 | | −6 | | −13 | | 16 | | | |
| | −21 | | 6 | | 24 | | −18 | | |
| −11 | | 0 | | 11 | | −2 | | 14 | |
| | +16 | | 10 | | | | | | |
| +5 | | 10 | | | | | | | |
| | +12 | | −11 | | | | | | |

TABLE IX, CONTINUED

| $X_q$ | $D$ | $(\sigma)$ | $D(\sigma)$ | $(\sigma)^2$ | $D(\sigma)^2$ | $(\sigma)^3$ | $D(\sigma)^3$ | $(\sigma)^4$ | $D(\sigma^4)$ |
|---|---|---|---|---|---|---|---|---|---|
| +17 | | −1 | | 10 | | 7 | | 10 | |
| | −45 | | 8 | | −13 | | −4 | | |
| −28 | | 7 | | −3 | | 3 | | | |
| | +36 | | −14 | | −3 | | −20 | | |
| +8 | | −7 | | −6 | | −17 | | −14 | |
| | +6 | | 1 | | 15 | | | | |
| +14 | | −6 | | 9 | | | | | |
| | −33 | | −5 | | −9 | | | | |
| −19 | | −11 | | 0 | | −9 | | 0 | |
| | +28 | | 33 | | −17 | | 18 | | |
| +9 | | 22 | | −17 | | 9 | | | |
| | −7 | | −32 | | 29 | | −14 | | |
| +2 | | −10 | | 12 | | −5 | | 4 | |

TABLE X — MEAN SQUARES FOR DIFFERENCE COLUMNS AND SUITABLE LINEAR COMBINATIONS THEREOF FOR THE ANALYSIS ILLUSTRATED BY TABLE IX

| Block | $D$ | $D(\sigma)$ | $D(\sigma)^2$ | $D(\sigma)^3$ | (*) | (†) | (‡) |
|---|---|---|---|---|---|---|---|
| 1 | 485 | 283 | 234 | 176 | −8 | −57 | −115 |
| 2 | 252 | 207 | 206 | 317 | 56 | 55 | 166 |
| 3 | 207 | 258 | 145 | 148 | 134 | 21 | 24 |
| 4 | 76 | 41 | 80 | 106 | −5 | 34 | 60 |
| Ave. of 2, 3 and 4.......................... | | | | | 51 | 38 | 78 |
| Noise Flat | 5/2 1 | 3/2 2 | 3/2 4 | 3/2 8 | 0 1.4 | 0 3.4 | 0 7.4 |

(*) Mean square for $D(\sigma)$ less 0.6 times that for $D$.
(†) Mean square for $D(\sigma)^2$ less 0.6 times that for $D$.
(‡) Mean square for $D(\sigma)^3$ less 0.6 times that for $D$.

The corresponding windows are, for the square of a single value of $S_2(F_2)S_2S_3X_q$,

$$4u^2(2 + 2u)(1 + 4u + 4u^2) = 8(u^2 + 5u^3 + 8u^4 + 4u^5),$$

and, for the square of a single value of $D(F_2)S_2S_3X_q$,

$$(4 - 4u^2)(2 + 2u)(1 + 4u + 4u^2) = 8(1 + 5u + 7u^2 - u^3 - 8u^4 - 4u^5),$$

to which should be compared that for the square of a single value of $DX_q$
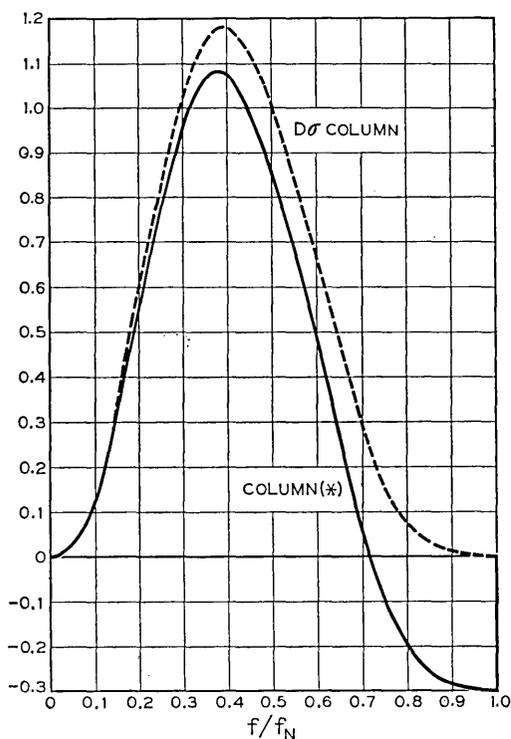
$$(2 - 2u) = 2(1 - u).$$

Fig. 24 — Windows corresponding to $D(\sigma)$ and to (*).

If now we form all values of $S_2(F_2)S_2S_3X_q$, $D(F_2)S_2S_3X_q$ and $DX_q$ (Table XI illustrates the parallel calculation of $F_2$ and $F_2'$ quantities), and form the corresponding mean squares, we obtain the results shown in Table XII.

If we compare the $S$-$0.2\delta$ column with the last column, we find a definite tendency for the last to be smaller. As indicated by the last two lines of the table, neither of these columns reflects, on the average, a "noise" component, while, on the average, both reflect a "flat" component equally. The natural conclusion is that the observed spectrum is more peaked toward $f/f_N = 0$ than would be expected from a mixture of "flat" and "noise" components. It might seem natural to conclude that the first model is to be preferred.

However, more careful examination shows that not only do the $\delta$ mean squares appear to decrease with improvements in astronomical technique, but (leaving aside block 1) the values in the S-$0.2\delta$ column are decreasing and those in the D-$0.6\delta$ column may well be doing the

## TABLE XI

Calculation of $S_2(F_2)S_2S_3$ and $D(F_2)S_2S_3$ for the second differences of Brouwer's fluctuations in the earth's rotation. (Portion of block 2.)

| Date | $X_q$ | $S_2X_q$ | $F_2S_2S_3X_q$ | $DF_2S_2S_3X_q$ | $S_3F_2S_2S_3X_q$ | $F_2S_2S_3X_q$ | $DF_2'S_2S_3X_q$ | $S_3F_2'S_2S_3X_q$ |
|---|---|---|---|---|---|---|---|---|
| 1854.5 | −4 | | | | | | | |
| 1855.5 | 0 | −6 | | | | | | |
| | | | −3 | | | | | |
| | −2 | +3 | | | | | | |
| | | | | +10 | +4 | 2 | | |
| | +5 | −1 | | | | | | |
| | | | +7 | | | | −14 | −10 |
| | −4 | +8 | | | | | | |
| | | | | −26 | −12 | −12 | | |
| | +7 | −20 | | | | | | |
| | | | −19 | | | | +5 | −19 |
| 1860.5 | −23 | +1 | | | | | | |
| | | | | +20 | −18 | −7 | | |
| | +17 | −8 | | | | | | |
| | | | +1 | | | | +9 | −5 |
| | −2 | +9 | | | | | | |
| | | | | −4 | −2 | 2 | | |
| | −6 | −7 | | | | | | |
| | | | −3 | | | | +1 | 5 |
| | +1 | +4 | | | | | | |
| | | | | +2 | −4 | 3 | | |
| 1865.5 | +9 | −1 | | | | | | |
| | | | −1 | | | | −11 | −5 |
| | −11 | 0 | | | | | | |
| | | | | −17 | −19 | −8 | | |
| | +2 | −8 | | | | | | |
| | | | −18 | | | | −12 | −28 |
| | +1 | −10 | | | | | | |
| | | | | +6 | −30 | −20 | | |
| | −13 | −10 | | | | | | |
| | | | −12 | | | | 18 | −22 |
| 1870.5 | +2 | −2 | | | | | | |
| | | | | 15 | −9 | −2 | | |
| | +9 | 0 | | | | | | |
| | | | 3 | | | | 8 | 4 |
| | −11 | +3 | | | | | | |
| | | | | 14 | 20 | 6 | | |
| | +5 | +3 | | | | | | |
| | | | 17 | | | | 21 | 33 |
| | +9 | +14 | | | | | | |
| | | | | −4 | 30 | 27 | | |
| 1875.5 | 0 | +13 | | | | | | |
| | | | 13 | | | | −23 | 31 |
| | +4 | 0 | | | | | | |
| | | | | −15 | 11 | 4 | | |
| | −4 | +4 | | | | | | |
| | | | −2 | | | | −3 | 5 |
| | +4 | −6 | | | | | | |
| | | | | | | 1 | | |
| | −6 | +7 | | | | | | |
| 1880.5 | 9 | | | | | | | |

TABLE XII — MEAN SQUARES OF COLUMNS FOR SECOND
DIFFERENCES OF BROUWER'S DATA

| For | $S$ | $D$ | $\delta$ | $S$-0.2$\delta$ | $D$-0.6$\delta$ | 2.5 ($D$-0.6$\delta$) |
|---|---|---|---|---|---|---|
| Block 1 | 83 | 228 | 484 | −14 | −62 | −155 |
| Block 2 | 290 | 178 | 252 | 240 | 27 | 68 |
| Block 3 | 153 | 168 | 207 | 112 | 44 | 110 |
| Block 4 | 62 | 47 | 75 | 47 | 2 | 5 |
| Average (2, 3, 4) | 168 | 131 | 178 | 133 | 24 | 61 |
| Noise | 1 | 3 | 5 | 0 | 0 | 0 |
| Flat | 28 | 12 | 2 | 27.6 | 10.8 | 27.6 |

$S$ = mean square value of $S_2(F_2)S_2S_3X_q$ .
$D$ = mean square value of $D(F_2)S_2S_3X_q$ .
$\delta$ = mean square value of $DX_q$ .

same. The suggestion is — which, after we have seen another block or two or three (of 32 years each), we may be able to confirm or deny — that the lower frequency component, as well as the noise component, is decreasing as astronomical technique improves. If this be true, then most of the low-frequency power in blocks 2 and 3 may represent observational and reductional sources of variation rather than changes in the rotation of the earth.

While more refined analyses might show something more, 131 values are not a great number, especially since technique has changed during their measurement, and it is likely that we shall have to wait a while for a more definite answer to this question.

This example was included, not because of the peculiar importance of the irregularities in the rotation of the earth, but rather to illustrate certain general points, particularly these:

(1) While careful spectral analysis requires a very considerable amount of arithmetic, there are situations where simple analysis will yield useful results. (All the values in this section were obtained with pen or pencil and an occasional use of a slide rule. The use of simple differencing and summing techniques to produce moderately sharp windows has been studied by the Labroustes.[41, 42, 43] Although they felt themselves interested in line spectra, their methods, properly reinterpreted, are easily and directly applicable to continuous spectra.)

(2) The windows corresponding to the add-and-subtract pilot analysis are broad and the estimation far from exhaustive. (The procedure is intended for use in planning prewhitening, not as a tool for answering questions of even moderate difficulty.)

(3) Subdivision of the data before analysis is quite often helpful.

(If blocks 2, 3, and 4 had been analyzed as one unit, it would have been easy to jump to a conclusion which now seems dangerous.)

(4) Spectral analysis can lead to results unsuspected before the calculations were made. (The decrease in amplitude of the low-frequency component which Brouwer's data now suggests was not at all suspected until the values given in Table XII were pulled together.)

### INDEX OF NOTATIONS

In case a notation is not widely used, the Sections in which it is used are specified in parentheses:

$a$ = an integer (A.6 and B.19),

$a_0$, $a_1$, $a_2$ = real constants (B-8.4),

$a_{i0}$, $a_{ij}$ = assorted constants (13 and B.5),

$A$ = various constants (27, B.9, B.15),

$A\left(f; \dfrac{1}{\Delta t}\right)$ = infinite Dirac comb made up of unit $\delta$-functions spaced $1/\Delta t$
in frequency (see A.2 or Table IV for formula),

$b$ = an integer (A.6 and B.19),

$B$ = auxiliary quantity defined in B-8.5 (B.8),

$B(t)$, $B_i(t)$, $\tilde{B}(t)$ = a data window (graded) (10 and B.10),

$c$ = an integer (A.6 and B.19),

$c_0$, $c_1$, $\cdots$, $c_k$ = real constants (B.15),

$C_{ij}$ = autocovariance corresponding to $t_i$ and $t_j$ (1),

$C(\tau)$ = autocovariance at lag $\tau$,

$C_{00}(\tau)$ = autocovariance estimated from record of finite length,

$C'_{00}(\tau)$, $\tilde{C}_{00}(\tau)$, etc. = hypothetical or actual analogs of $C_{00}(\tau)$,

$C_i(\tau) = D_i(\tau)C_{00}(\tau)$ = modified autocovariance estimate (defined in 4),

$C_r$ = sample autocovariance at lag $r\Delta\tau$ (usually = $r\Delta t$ in practice),

$d$ = an integer (A.6, B.18 and B.19),

$D$ = the operation of differencing adjoining values,

$D_i(\tau)$ = a prescribed even function of time shift, often a lag window
(cp. 4, 5, B.4, B.5),

$D_{ei}(\tau)$ = lag window equivalent of $B_i(\tau)$ (see B-4.11 for representation),

$e$ = 2.71828182845 $\cdots$ ,

$E_{kr}$ = $k^{\text{th}}$ alternative correction at lag $r$ (19, B.19),

$f$ = frequency (in cycles per second where not otherwise specified), in
equi-spaced discrete situations $f = r/(2m \cdot \Delta t) = (r/m)f_N$ ,

$f_0$ = a particular frequency (in Section B.18 the folding frequency of
the cycle in question),

$f_1$ = a frequency, often the nominal frequency of a smoothed estimate,

$f_N$ = Nyquist or folding frequency = $1/(2 \cdot \Delta t)$,

$f_N{}^*$ = effective Nyquist or folding frequency = $1/(2 \cdot \Delta\tau)$, (13),

$F_j$ = the operation of retaining only every $j^{th}$ value ($j$ = 2, 3, $\cdots$) (B.19, B.28),

$F_2'$ = the operation of retaining the alternate values dropped by $F_2$ (B.19, B.28),

$g$ = an index running from $-c$ to $+d$ (A.6 and B.19),

$G(t)$ = a function of the time, Fourier transform of $S(f)$ (correspondence also required with subscripts 0, 1, 2),

$\tilde{G}(t)$ = the Fourier transform of a particular box-car function (B.12),

$h$ = an integer, generally satisfying $\Delta\tau = h \cdot \Delta t$, in A.6 and B.19 an index running through indicated ranges,

$H_i(f; f_1)$ = $Q_i(f + f_1) + Q_1(f - f_1)$ = power transfer function (from power at absolute frequency $f$ to estimate at nominal frequency $f_1$),

$H_i(f - f_1)$ = special form of $H_i(f; f_1)$,

$H_Y(f; f_1)$ = power transfer function defined in B-10.2,

$i = \sqrt{-1}$,

Subscript $i$, values often 0, 1, 2, 3, 4, usually identifies quantities or functions associated with $i^{th}$ window pair,

$j$ = an integer, often such that $3j - 2 = k$ or $\ell$ (B.19),

Subscript $j$, values often 0, 1, 2, 3, 4, alternative to subscript $i$,

$k$ = usually number of equivalent degrees of freedom or number of elementary frequency bands; in Section 17, length of group averaged,

$\ell$ = an integer; in 18 an exponent of 2, in A.6 satisfying $n = 2\ell + 1$,

$L$ = inductance (in example of 27),

$m$ = integer, number of longest lag (longest lag = $m \cdot \Delta\tau$, usually = $m\Delta t$),

$M(t, \tau)$ = bilinear monomial in the $X$'s
    = $X(t + \tau/2) \cdot X(t - \tau/2)$,

$n$ = usually one less than the number of discrete data points, in A.6 and B.19 an integer = $2\ell + 1$,

$n'$ = effective length of record ($n' \cdot \Delta t = T_n'$),

$p$ = usually the number of pieces of record; also = $i\omega$ (A.1 only); a real number (usually integral, A.6),

$p_0$ = a constant power density (B.9),

$p_0$, $p_1$, $p_2$, $\cdots$ = ordinates in general example (9),

$P(f)$ = density of power spectrum (normalized so that variance = $\int_0^\infty 2P(f)\,df$), also with various subscripts,

$P_a(f)$ = aliased density of power spectrum (periodic in frequency, cp. 12),

$P_A(f)$ = principal part of aliased density of power spectrum (confined to $|f| \leqq 1/(2 \cdot \Delta\tau)$, cp. 12),

$P_Y(f_1)$ = estimate of spectral density based on filtered signal (B.10),

$P_{ei}(f)$ = spectral density estimated with the data window $B_i(t)$,

$P_{00}(f)$ = symbolic Fourier transform of $C_{00}(\tau)$,

$P_{0A}(f)$ = aliased estimate of smoothed $P(f)$,

$P_{i1}(f) = H_i(f; f_1)P(f)$, roughly a filtered version of $P(f)$ (subscript $i$ interchangeable with $j$, also 1 with 2),

$P_{iA}(f_1)$ = aliased estimate of smoothing of $P(f)$ by $H_i(f; f_1)$,

$P_{iA1}(f) = H_i(f; f_1) \cdot P_A(f)$ = aliased filtered spectral density,

$P_{iAk}(f)$ = corrected aliased estimate of $P(f)$ smoothed by $Q_i(f)$,

$P_{out}(f; f_1) = |\, Y(f; f_1)\,|^2 \cdot P(f)$ = spectral density of output,

$\tilde{P}(f)$ = a perfectly known power spectrum (B.12),

$\tilde{P}_A(f)$ = either the aliased spectrum corresponding to $\tilde{P}(f)$ (B.12) *or* the aliased spectrum of the $\tilde{X}_q$ series (15, 17),

$\tilde{P}_{i1}(f)$ = an auxiliary quantity defined in B-8.4,

$q$ = index running as indicated,

$Q_r$ = an auxiliary quantity (B.19),

$Q_0(f)$ = Fourier transform of centered box-car function of length $2T_m$ (See also $Q_i(f)$ for $i$),

$Q_i(f)$ = Fourier transform of $R_i(\tau)$ (see B.5 for $i = 0, 1, 2, 3, 4$ and B-8.7 for another choice),

$Q_0(f; \Delta\tau)$ = spectral window corresponding to (Fourier transform of) a discrete box-car function,

$\qquad\qquad = \Delta\tau \cdot \cos \tfrac{1}{2}(w \cdot \Delta\tau) \sin (mw \cdot \Delta\tau)$,

$Q_{ei}(f)$ = spectral window corresponding to (Fourier transform of) $D_{ei}(\tau)$,

$Q_{0A}(f)$ = aliased form of $Q_0(f) = Q_0(f; \Delta\tau)$,

$Q_{iA}(f)$ = aliased form of $Q_i(f)$,

$\tilde{Q}_0(f), \tilde{Q}_i(f)$ = auxiliary quantities (B.8),

$r$ = integer index almost always running from 0 or 1 to $m$ or $m - 1$,

$R$ = resistance (in example of Section 27),

$R_{ik}(f)$ = spectral window associated with $Q_i(f)$ and the sequence $E_{k0}$, $E_{k1}, \cdots, E_{km}$,

$S_j$ = operation of summing by (overlapping) sets of $j$ each (B.17, B.28),

$S(f)$ = function of frequency, Fourier transform of $G(t)$ (also with subscripts 1, 2, $\cdots$),

$\tilde{S}(f)$ = a particular box-car function (B.12),

$t$ = time in the sense of epoch (also with various subscripts),

$T$ = a time, usually positive,

$T_m$ = half-length of box-car function of time, *or* greatest lag used or considered,

$T_n$ = length of record,

$T'_n$ = effective length of record, approximately $T_n - \dfrac{p}{3} T_m$,

$u$ = general real variable,

$U_r$ = corrected estimate of smoothed power density (at nominal frequency $r/(2m \cdot \Delta t)$),

$V_r$ = raw estimate of smoothed power density at nominal frequency $r/(2m \cdot \Delta t)$,

$w$ = a chance quantity (B.6),

$W$ = a frequency band width (B.9),

$W_e$ = equivalent width, often of $P_{i1}(f)$ defined in Section 8,

$W_{\text{main}}$ = width of main lobe,

$W_{\text{side}}$ = width of side (unsplit) lobe,

$W(t)$ = impulse response of linear transmission system,

$W(t; f_1)$ = in B.10 the impulse response of filter with transfer function $Y(f; f_1)$,

$x$ = usually a real variable, in B.6 a chance quantity (random variable),

$X_q$ = $q$th value of discrete, equi-spaced time series,

$X(t)$ = value of time function,

$\bar{X}$ = in Section 1 an average value (along an infinite function or across an ensemble), in 19 ff. the mean of the observed $X_q$'s,

$\overline{X^+}$, $\overline{X^-}$ = means of end thirds of observed values,

$\tilde{X}_q$, $X_q^*$ = $q$th values of linearly transformed time series,

$y$ = in Section B.6 a chance quantity (random variable),

$Y(f)$ = steady-state transfer function corresponding to linear transformation or linear transmission system,

$z$ = in Section B.6 a chance quantity (random variable),

$\alpha$ = a real constant (15) *or* an indeterminate (B.8), *or* a certain fraction (B.18) *or* an unknown constant (B.19),

$\alpha_i$ = factor indicating extent of end effect losses (8, B.6) defined by B-6.9,

$\beta$ = real constant (15) *or* an indeterminate (B.8) *or* a constant defined by B-8.8, *or* an unknown (slope) constant (B.19),

$\hat{\beta}$ = estimate of $\beta$(B.19),

$\gamma$ = real constant (15),

$\Gamma(f)$ = power-variance spectral density,

$\Gamma_{\Delta t}(f)$ = power-variance spectral density in the equi-spaced discrete case,

$\delta$ = operation of forming alternate differences,

$\delta'$ = operation of forming alternate differences complementary to $\delta$,

$\Delta f$ = a change in frequency, sometimes the width $1/(2T'_n)$ of an elementary frequency band,

$\Delta t$ = time interval, usually that between data values,

$\Delta \tau$ = time interval, usually that between lags used ($= h \cdot \Delta t$),

$\epsilon$ = a small real number (A.2),

$\xi$ = real, time-like variable of integration (A.3),

$\lambda$ = real, time-like variable of integration (A.3) *or* a real constant (15),

$\mu$ = real constant (15),

$\nu$ = real constant (15),

$\omega$ = angular frequency (in radians per second unless otherwise specified) always = $2\pi f$ [with any sub- or superscripts],

$\varphi$ = $2\pi f T_m$ , a normalized frequency (B.8),

$\Phi(f, \lambda)$ = an auxiliary function (B.6) (defined in B-6.3),

$\psi$ = a phase angle (A.6),

$\chi_k{}^2$ = a quantity distributed as chi-square on $k$ degrees of freedom (9),

$\sigma, \sigma'$ = complementary operations of summing by adjacent parts and then omitting every alternate sum,

$\tau$ = time difference or lag,

$*$ = sign of convolution,

superscript $*$ = sign of complex conjugate (A.3),

$\nabla(t; \Delta t)$ = infinite Dirac comb approximating the constant unity [formula in A.2 or Table IV],

$\nabla_m(t; \Delta t)$ = finite Dirac comb approximating a unit-height, centered, box-car function [formula in A.2 or Table IV].

REFERENCES (PARTS I AND II)

1. N. R. Goodman, *On the Joint Estimation of the Spectra, Cospectrum and Quadrature Spectrum of a Two-Dimensional Stationary Gaussian Process*, Scientific Paper No. 10, Engineering Statistics Laboratory, New York University, 1957. (Ph.D. Thesis, Princeton University).
2. G. I. Taylor, Statistical Theory of Turbulence, Proc. Roy. Soc. (London), **A151,** pp. 421–478, 1935.
3. S. O. Rice, Mathematical Analysis of Random Noise, B.S.T.J., **23,** pp. 282–332, July, 1944; **24,** pp. 46–156, Jan. 1956. (This article also appears in *Selected Papers on Noise and Stochastic Processes*, edited by N. Wax, Dover Publ., New York, 1954, pp. 133–294.)
4. J. W. Tukey, The Sampling Theory of Power Spectrum Estimates, in *Symposium on Applications of Autocorrelation Analysis to Physical Problems*, Woods Hole, June 13, 1949. NAVEXOS-P-735, Office of Naval Research.
5. M. S. Bartlett, Periodogram Analysis and Continuous Spectra, Biometrika, **37,** pp. 1–16, 1950.
6. W. Marks and W. J. Pierson, The Power Spectrum Analysis of Ocean-Wave Records, Trans. Am. Geophysical Union, **33,** pp. 834–844, 1952.
7. H. Press and J. C. Houbolt, Some Applications of Generalized Harmonic Analysis to Gust Loads on Airplanes, J. Aero. Sciences, **22,** pp. 17–26, 1955.
8. H. A. Panofsky, Meteorological Applications of Power Spectrum Analysis, Bull. Am. Meteorological Soc., **36,** pp. 163–166, 1955.
9. G. P. Wadsworth, E. A. Robinson, J. G. Bryan, and P. M. Hurley, Detection of Reflections on Seismic Records by Linear Operators, Geophysics, **18,** pp. 539–586, 1953. (Also see later reports, Geophysical Analysis Group, M.I.T.)
10. M. S. Bartlett and J. Mehdi, On the Efficiency of Procedures for Smoothing Periodograms from Time Series with Continuous Spectra, Biometrika, **42,** pp. 143–150, 1955.

11. M. S. Bartlett, *An Introduction to Stochastic Processes*, Cambridge Univ. Press, Cambridge, England, 1955.
12. U. Grenander and M. Rosenblatt, *Statistical Analysis of Stationary Time Series*, Wiley, New York, 1957.
13. J. W. Tukey and C. P. Winsor, *Note on Some Chi-Square Normalizations*, Memorandum Report 29, Statistical Research Group, Princeton, 1949.
14. G. Groves, Numerical Filters for Discrimination Against Tidal Periodicities, Trans. Am. Geophysical Union, **36**, pp. 1073–1084, 1955.
15. H. R. Seiwell, The Principles of Time Series Analysis Applied to Ocean Wave Data, Proc. Nat. Acad. Sciences, **35**, pp. 518–528, Sept., 1949.
16. H. R. Seiwell and G. P. Wadsworth, A New Development in Ocean Wave Research, Science, **109**, pp. 271–274, 1949.
17. D. Brouwer, A Study of Changes in the Rate of Rotation of the Earth, Astronomical J., **57**, pp. 126–146, 1952.
18. Brouwer, pp. 135, 138.
19. G. A. Campbell and R. M. Foster, Fourier Integrals for Practical Applications, B.S.T.J., **7**, pp. 639–707, 1928; also, Van Nostrand, New York, 1942.
20. E. A. Guillemin, *Introductory Circuit Theory*, Wiley, New York, 1953, p. 257.
21. L. Schwartz, *Theorie des Distributions*, Vol. 1 and II, Hermann et Cie, Paris, 1950. (See also Halperin; Friedman; and Clavier, noted in Bibliography.)
22. E. A. Guillemin, *The Mathematics of Circuit Analysis*, Wiley, 1949, pp. 485–501.
23. H. Nyquist, Certain Topics in Telegraph Transmission Theory, Trans. A.I.E.E., pp. 617–644, April, 1928.
24. J. M. Whittaker, *Interpolatory Function Theory*, Cambridge Univ. Press, Cambridge, England, 1935, Chapter IV.
25. W. R. Bennett, Methods of Solving Noise Problems, Proc. I.R.E., **44**, 1956.
26. R. W. Hamming and J. W. Tukey, Measuring Noise Color, unpublished memorandum.
27. C. L. Dolph, A Current Distribution for Broadside Arrays Which Optimizes the Relationship between Beam Width and Side-Lobe Level, Proc. I.R.E., **34**, pp. 335–348, 1946.
28. P. Jaquinot, Quelques recherches sur les raies faible dans les spectres optiques, Proc. Phys. Soc. (London), **63B**, pp. 969–979, 1950.
29. P. Boughon, B. Dossier, and P. Jaquinot, Determination des fonctions pour l'amelioration des figures de diffraction dans le spectroscope, C. R. Acad. Sci. Paris, **223**, pp. 661–663, 1946.
30. L. Isserlis, On a Formula for the Product-Moment Coefficient of Any Order of a Normal Frequency Distribution in Any Number of Variables, Biometrika, **12**, pp. 134–139, 1918.
31. H. Hotelling, Relations Between Two Sets of Variates, Biometrika, **28**, pp. 321–377, 1936.
32. H. E. Kallmann, Transversal Filters, Proc. I.R.E., **28**, pp. 302–310, 1940.
33. H. Wold, *A Study in the Analysis of Stationary Time Series*, 2nd Edition, Almquist and Wiksell, Stockholm, 1954.
34. Daniel B. DeLury, *Values and Integrals of the Orthogonal Polynomials up to n = 26*, Toronto, University Press, 1950.
35. Ronald A. Fisher and Frank Yates, *Statistical Tables for Biological, Agricultural and Medical Research*, Edinburgh, Oliver and Boyd, 4th Ld. 1953, (especially Table XXII).
36. William E. Milne, *Numerical Calculus*, Princeton University Press, 1949.
37. Cornelius Lanczos *Tables of Chebyshëv Polynomials $S_n(x)$ and $C_n(x)$*, Nat'l. Bur. Stds. App'd. Math. Series 9, Washington, Gov't Print. Off., 1952, (especially pp. xiv to xvi).
38. Cornelius Lanczos, Trigonometric Interpolation of Empirical and Analytical Functions, J. Math. Phys., **17**, pp. 123–199, 1938.
39. Robert C. Minnick, Tshebycheff Approximations for Power Series, J. Assoc. Comp. Machinery, **4**, pp. 487–504, 1957.
40. J. C. P. Miller, Two Numerical Applications of Chebyshëv Polynomials, Proc. Roy. Soc. Edinb., **62**, pp. 204–210, 1946.
41. H. Labrouste and Y. Labrouste, Harmonic Analysis by Means of Linear Combinations of Ordinates, J. Terr. Magnetism, **41**, 1936.

42. H. Labrouste and Y. Labrouste, Analyse des graphiques resultant de la super-position de sinusoides, Mem. Acad. Sci. Inst. France, (2) **64,** No. 5, 1941. (See also Ann. de l'Inst. de Phys. du Globe de Univ. de Paris, **7,** pp. 190–207; **9,** pp. 99–101; **11,** pp. 93–101; **14,** pp. 77–91, for much of this material.)
43. H. Labrouste and Y. Labrouste, *Tables Numeriques precedes d'un expose de la methode d'analyse par combinaison d'ordonnees,* Presses Univ., Paris, 1943.

## BIBLIOGRAPHY

In addition to the articles cited in the text, the reader's attention is called to the additional titles below.

P. Boughon, B. Dossier, and P. Jaquinot, Apodisation des raies spectrales au moyen d'ecrans absorbents, J. Recherches C.N.R.S., No. **11,** pp. 49–69, 1950.

P. A. Clavier, Some Applications of the Laurent Schwarz Distribution Theory to Network Problems, *Proceedings of the Symposium on Modern Network Synthesis, II,* Polytechnic Institute of Brooklyn, 1955.

L. Couffignal, La Methode de H. Labrouste pour la recherche des periods, Ciel et Terre, **66,** pp. 78–86, 1950.

B. Friedman, *Principles and Techniques of Applied Mathematics,* Wiley, New York, 1956, Chapter 3.

I. Halperin, *Introduction to the Theory of Distributions,* University of Toronto Press, Toronto, 1952.

B. Mazelsky, Extension of the Power Spectrum Methods of Generalized Harmonic Analysis to Determine Non-Gaussian Probability Functions of Random Input Disturbances and Output Responses of Linear Systems, J. Aeronautical Sciences, **21,** pp. 145–153, March, 1954.

H. A. Panofsky and I. van der Horen, Spectra and Cross-Spectra of Velocity Components in the Mesometeorological Range, Quart. J. Roy. Meteorol. Soc., **81,** pp. 603–605, 1955.

H. A. Panofsky and R. A. McCormack, The Vertical Momentum Flux at Brookhaven at 109 Meters, Geophys. Res. Papers (USAF, Cambridge Research Center), No. **19,** pp. 219–230, 1952.

H. Press and J. W. Tukey, Power Spectral Methods of Analysis and Their Application to Problems in Airplane Dynamics, *Flight Test Manual, NATO, Advisory Group for Aeronautical Research and Development,* IV-C, pp. 1–41, June, 1956. (Reprinted as Bell System Monograph No. 2606.)

# Recent Monographs of Bell System Technical Papers Not Published in This Journal*

ABRAHAMS, E.

Donor Electron Spin Relaxation in Silicon, Monograph 2874.

ANDERSON, O. L., CHRISTENSEN, H. and ANDREATCH, P., JR.

Technique for Connecting Electrical Leads to Semiconductors, Monograph 2875.

ANDREATCH, P., JR., see Anderson, O. L.

BALLHAUSEN, C. J., See Liehr, A. D.

BENES, V. E.

A Sufficient Set of Statistics for a Simple Telephone Exchange Model, Monograph 2851.

BENSON, K. E., see Pfann, W. G.

BOGERT, B. P.

Response of an Electrical Model of the Cochlear Partition, Monograph 2876.

BOHM, D., HUANG, K., and PINES, D.

Subsidiary Conditions in the Collective Description of Electron Interactions, Monograph 2877.

BRADY, G. W.

Structure of Tellurium Oxide Glass, Monograph 2878.

HOBSTETTER, J. N., see Greiner, E. S.

HOPKINS, I. L., and HAMMING, R. W.

  **On Creep and Relaxation,** Monograph 2884.

HROSTOWSKI, H. J. and KAISER, R. H.

  **Infrared Absorption of Oxygen in Silicon,** Monograph 2885.

HUANG, K., see Bohm, D.

HUYETT, M. J., see Sobel, M.

JONES, W. D., see Turrell, G. C.

KAISER, R. H., see Hrostowski, H. J.

LAX, M., and SACHS, R.

  **Frequency Dependence of AC Resistance of Thin Semiconducting Films,** Monograph 2886.

LEBOWITZ, J. L. and FRISCH, H. L.

  **Model of Nonequilibrium Ensemble: Knudsen Gas,** Monograph 2898.

LIEHR, A. D. and BALLHAUSEN, C. J.

  **Intensities in Inorganic Complexes,** Monograph 2887.

MAKI, A., see Turrell, G. C.

MARCATELLI, E. A.

  **Heat Loss in Grooved Metallic Surface,** Monograph 2888.

McGLASSON, J., JR., see Peterson, J. W.

MERTZ, P.

  **Information Theory Impact on Modern Communications,** Monograph 2900.

NIELSEN, E. G.

  **Behavior of Noise Figure in Junction Transistors,** Monograph 2889.

PEARSON, G. L., see Wertheim, G, K.

PETERSON, J. W., McGLASSON, J., JR., and HITTINGER, W. C.

Some Aspects of Alloying onto Germanium Surfaces, Monograph 2890.

PFANN, W. G., BENSON, K. E., and WERNICK, J. H.

Aspects of Peltier Heating at Liquid-Solid Interfaces in Germanium, Monograph 2891.

PINES, D., see Bohm, D.

POLLAK, H. O., see Gilbert, E. N.

POMEROY, A. F.

Waveguide Loss Charts, Monograph 2899.

PRESTIGIACOMO, A. J.

Plastic-Tape Sound Spectrograph, Monograph 2892.

PRINCE, M. B., see Veloric, H. S.

RICE, S. O.

Distortion in a Noise-Modulated FM Signal by Nonlinear Attenuation, Monograph 2859.

RIGROD, W. W.

Noise Spectrum of Electron Beam in Longitudinal Magnetic Field, Monograph 2860.

RIORDAN, J.

The Numbers of Labeled Colored and Chromatic Trees, Monograph 2896.

SACHS, R., see Lax, M.

SOBEL, M., HUYETT, M. J., and GUPTA, S. S.

Selection and Ranking Problems with Binomial Population, Monograph 2787.

STORKS, K. H., see Wright, J. P.

SUNDE, E. D.

Self-Timing Regenerative Repeaters, Monograph 2864.

THURMOND, C. D., see Dacey, G. C.

TURRELL, G. C., JONES, W. D., and MAKI, A.

Infrared Spectra and Force Constants of Cyanoacetylene, Monograph 2893.

VELORIC, H. S., and PRINCE, M. B.

High-Voltage Conductivity-Modulated Silicon Rectifier, Monograph 2865.

WAITE, T. R.

Theoretical Treatment of the Kinetics of Diffusion-Limited Reactions, Monograph 2895.

WERNICK, J. H., see Pfann, W. G.

WERTHEIM, G. K. and PEARSON, G. L.

Recombination in Plastically Deformed Germanium, Monograph 2897.

WRIGHT, J. P. and STORKS, K. H.

X-ray Fluorescent Spectra for Analysis of Preservatives in Wood and Paper, Monograph 2894.

# Contributors to This Issue

R. B. BLACKMAN, A.B., California Institute of Technology, 1926; Bell Telephone Laboratories, 1926–. From 1926 to 1936, Mr. Blackman was engaged in physical research in hearing, acoustics and electro-mechanical filters. Since 1936, he has been engaged in applied mathematical research, and in military problems, specializing in linear circuits, and in data-smoothing and prediction. He is a member of Tau Beta Pi and Institute of Radio Engineers.

DAVID J. BRANGACCIO, B. of M.E., 1942, New York University; Bell Telephone Laboratories, 1934–. Mr. Brangaccio engaged in work in television research and later with wide-band amplifiers before transferring to the Radio Research Department in 1945. He worked with traveling wave tubes and later took charge of a group making experimental microwave tubes. After work on the development of a broadband oscilloscope tube in 1954 he is again concerned with traveling wave tubes.

A. FRANCIS DIETRICH, Bell Telephone Laboratories, 1942–. During 1942 and 1943, Mr. Dietrich worked on radar projects. Since 1943 he has been engaged with experimental studies in broadband systems and with terminals and repeaters for microwave radio and waveguide applications.

DAVID H. EVANS, B.S., 1948, Lehigh University; Ph.D., 1953, Brown University; Bell Telephone Laboratories, 1953–. Since joining the technical staff of Bell Laboratories, Mr. Evans has engaged chiefly in mathematical consulting on electronic switching with members of the switching systems development department. Member of Operations Research Society of America, Institute of Management Sciences, Phi Beta Kappa, Tau Beta Pi and Sigma Xi.

A. C. GILMORE, Western Electric Company, 1916–1925; Bell Telephone Laboratories, 1925–1958. Mr. Gilmore died in January, 1958, after 41 years of Bell System service. His early work was the design of central

office and crossbar system equipment. After 1940 he was concerned with military and defense projects, including mobile fire control radar, air warning control centers and the Army telephone central office.

PRESTON R. GRAY, Bell Telephone Laboratories, 1924–. For many years Mr. Gray was engaged in testing work for step-by-step system circuits and analyzation of orders for manual and dial systems. During World War II he worked on the design of switchboards for the armed forces. Until 1951 Mr. Gray was concerned with designing circuits for special service desks and announcement machine systems. After working on the development of circuits for the Army switchboard, Mr. Gray became leader of a group for the analyzation and development of special circuits.

ESTILL I. GREEN, A.B., 1915, Westminster College; B.S.E.E., 1921, Harvard University; Hon. D. Sc., 1956, Westminster College; American Telephone and Telegraph Company, 1921–1934; Bell Telephone Laboratories 1934–. Vice President in charge of Systems Engineering since 1955, Mr. Green has engaged in a wide range of Bell System activities. For many years he took part in planning the development of new transmission systems and services and facilities for special customers. In 1939 he was appointed Test Engineer, responsible for the development of transmission circuit field testing apparatus. During World War II Mr. Green was in charge of the development of radar test equipment and other military developments. As Director of Transmission Apparatus Development in 1948, he headed the development of systems components including electronic parts for transistorized systems. In 1953 Mr. Green was named Director of Military Communications Systems, in charge of planning and development in that area. Mr. Green holds more than 70 patents and is the author of many articles in scientific and personnel fields. Fellow of American Institute of Electrical Engineers and Institute of Radio Engineers and member of Acoustical Society of America, American Physical Society, Operations Research Society of America, and American Association for the Advancement of Science.

W. S. IRVINE, B.S. in E.E., 1936, Union College; New Jersey Bell Telephone Company, 1936–1944; Bell Telephone Laboratories, 1944–. With New Jersey Bell, Mr. Irvine was concerned with central office maintenance and was an outside plant engineer. At Bell Laboratories, he took part in airborne radar projects with the trial installation group. After 1945 Mr. Irvine was concerned with circuits standards and design

of step-by-step and PBX circuits. In 1951 he transferred to military work and took part in the development of the Army switchboard. Member of Eta Kappa Nu.

L. A. MEACHAM, B.S. in E.E., 1929, University of Washington; Certificate of Research, 1930, Cambridge University; Bell Telephone Laboratories, 1930–. Mr. Meacham, station development engineer at Bell Laboratories, engaged in frequency standards research until 1940. He then worked for two years on the application of electronic devices to telephone switching. From 1941 to 1953 Mr. Meacham was concerned with radar development, studies of pulse code modulation, and underwater sound. Since then he has engaged in studies of transistorized telephone sets. Mr. Meacham, who has 52 patents, received the Eta Kappa Nu Recognition of Outstanding Young Electrical Enginners in 1939. Fellow of Institute of Radio Engineers, member of Phi Beta Kappa, Sigma Xi, Eta Kappa Nu and Tau Beta Pi.

J. R. POWER, B.S. in E.E., 1927, Carnegie Institute of Technology; Bell Telephone Laboratories, 1927–. Mr. Power was concerned with the design of special motors and generators, chiefly for use with motion pictures and television, until 1935 when he began studies of noise reduction. During World War II he worked on high-power loud speaker systems for the government. After the war Mr. Power took part in the development of hearing aids, artificial larynges, and later, telephone ringers. His recent work has been with transistorized telephone sets. Member of Acoustical Society of America, American Association for the Advancement of Science, Tau Beta Pi and Eta Kappa Nu.

GORDON RAISBECK, B.A., 1944, Stanford University; Ph.D., 1949, Massachusetts Institute of Technology; Bell Telephone Laboratories, 1949–. Mr. Raisbeck has been an instructor at Stanford Univ., at M. I. T., and at Drew Univ., and served in the U. S. Navy as a Radio Technical Officer during World War II. His early work at Bell Laboratories was in research in acoustics and underwater sound. From 1950 to 1953 he concentrated on research in transistor circuits and then took his present post in charge of a group engaged in transmission line research. Throughout his Laboratories career, Mr. Raisbeck has done work on information theory. He was a Rhodes Scholar at Oxford University in 1947. Senior member of Institute of Radio Engineers, member of American Mathematical Society, Mathematical Association of America, Society for Industrial and Applied Mathematics, American Management Association and Sigma Xi.

W. T. READ, JR., B.S., 1944, Rutgers University; M.S., 1948, Brown University; National Defense Research Committee, 1943–1946; Bell

Telephone Laboratories, 1947–. With the N. D. R. C. at Princeton, Mr. Read was engaged in shock-wave tests and he supervised part of the air blast measurements at the Bikini atom bomb tests. With Bell Laboratories he concentrated on stress analysis before turning to studies in dislocation theory in crystals and electrical and mechanical effects of dislocations in semiconductors. Since 1950 Mr. Read has been concerned with problems in semiconductor device theory. Member of Phi Beta Kappa.

J. W. SULLIVAN, B.E.E., M.Sc., 1952, Ohio State University; Bell Telephone Laboratories, 1952–1956. Mr. Sullivan was engaged in research and development in microwave tubes during his association with Bell Laboratories. Member of Institute of Radio Engineers, Tau Beta Pi, Eta Kappa Nu and Sigmi Xi.

JOHN W. TUKEY, Sc.B. Chem., 1936, Sc.M., 1937, Brown University; M.A., 1938, Ph.D., 1939, Princeton University. Mathematics Department, Princeton U., 1939–. Bell Laboratories, 1945–. Mr. Tukey's work has covered development of new statistical techniques; broad systems analysis and synthesis problems in studies of highly complex weapons systems; and other problems with mathematical or statistical aspects. He is the author of numerous articles, and is a co-author with W. G. Cochran and Frederick Mosteller of *The Statistical Problems of the Kinsey Report.* Mr. Tukey is co-inventor with C. E. Shannon and J. R. Pierce of a cathode ray device. He is a member of the American Statistical Association, American Association for Advancement of Science; American Society for Quality Control, Sigma Xi, Biometric Society, American Mathematical Society, Mathematical Association of America, Institute of Mathematical Statistics, N. Y. Academy of Sciences, Econometric Society, American Society for Human Genetics, Association for Computing Machinery, Royal Statistical Society, American Association of University Professors, Operations Research Society of America, American Association of Mathematics Teachers, Society for Industrial and Applied Mathematics, International Statistical Institute, and the Cosmos Club.

FRED WEST, B.E., 1928, Johns Hopkins University; Bell Telephone Laboratories, 1928–. With the station development department since joining Bell Laboratories, Mr. West has worked on studies of transmission systems and components and on the development of methods for measuring characteristics of station apparatus. He later took part in developing test equipment for manufacturing control and during World War II in development of deception devices. He is now engaged in exploratory development of station apparatus.