The Bell System Technical Journal

October, 1932

World-Wide Telephony—Its Problems and Future *

By **BANCROFT GHERARDI** and **FRANK B. JEWETT**

The rapid development of large telephone networks giving a high grade of service between large numbers of telephones in continental areas has laid the foundation for the development of world-wide telephone service. Beginning in 1927 with the establishment of the first commercial telephone circuit between Europe and North America, intercontinental telephone service has, during the past five years, extended rapidly, and with further extensions already definitely planned, will embrace all of the continents and make possible the connection together of practically any two telephones in the world.

Up to the present time radio has been used to overcome the peculiar technical obstacles in the provision of intercontinental circuits. Two portions of the radio spectrum are suitable for this purpose, the long wave providing only a few circuits and the short wave providing for possibly several hundred circuits in the world as a whole. Plans have already been made for the important route between Europe and North America to supplement these with a telephone cable and the use of wire lines for intercontinental routes may become more important in the future.

The full development of intercontinental telephony is affected by a number of general difficulties. Of these the differences in time between different parts of the earth's surface are inherent. Differences in language both affect the ease with which customers can converse over the telephone, and complicate the operating problem. Furthermore, some of the differences in operating and commercial practices in the telephone networks of different continents which have in the past developed largely independently of each other, require consideration in the building up of intercontinental services.

The full development of intercontinental telephony is dependent upon the continued progress in working out these problems and in an extension of the brilliant scientific and engineering achievements which have made possible the present services. It is to be expected that with the further growth of intercontinental service it will be found desirable in the future to adopt a general world-wide plan for the routing of intercontinental messages somewhat comparable to the plans for continental telephone service already under consideration or in use. While political considerations may temporarily affect the form of the world-wide network, ultimately the requirements of economy and good service will no doubt be determining factors in such a plan.

It is to be hoped that the continued closer knitting together of the nations and races of the world by intercontinental telephone circuits will be a great contribution to international friendship and good will.

The authors wish to acknowledge their indebtedness to a number of telephone administrations who have provided them with information regarding present and proposed intercontinental services, supplementing the data previously published. They have also drawn freely on the material presented in the bibliography of Appendix 2, and this material has been of assistance. The authors also express appreciation of the assistance given them by a number of their associates in the American Telephone and Telegraph Company, particularly Messrs. O. B. Blackwell, A. B. Clark, L. Espenschied, O. T. Laube, H. S. Osborne and H. E. Shreeve.

---

* Presented at The International Electrical Congress, Paris, France, July 5–12, 1932.

INTRODUCTION

THE rapidity with which telephone service has been extended to world-wide proportions during the past few years is perhaps one of the most remarkable of man's conquests over time and distance. Already it is a commonplace to hear the human voice from thousands of miles away, over land and sea. Distance and the great natural barriers of the world no longer prevent us from talking with each other. Across oceans and over high sierras the voice now carries its full message. Furthermore, these results have been accomplished within a few years. Telephony has demonstrated its international and intercontinental services. The development and extension of those services lies before us. It seems an appropriate time, therefore, to review some of the problems and possibilities in this extension of the application of the electrical arts to the service of mankind.

World-wide telephony has as its foundation the wire line networks on the various continents with their millions of users. The first step was taken toward overcoming the great barrier presented by the oceans in 1891 when the first submarine telephone cable was laid between Dover and Calais. Further submarine cable developments followed including the laying of a continuously loaded cable between Denmark and Sweden in 1902.

In the meantime the problems of spanning great distances over land were being rapidly solved, both in North America and in Europe, through the application of the loading coil and other transmission improvements. In 1911 service was opened between New York and Denver, a distance of 2000 miles (3300 kilometers). By 1913 the development of underground cable had progressed so that a cable was placed in service between Boston, New York and Washington, a distance of over 420 miles (700 kilometers).

But it required still further improvements in the whole art of long distance telephone transmission, including amplifiers and their application to wire lines before long distance telephony could develop beyond the semi-continental stage to truly transcontinental and transoceanic distances. With improved amplifier elements and the perfection of means for applying repeaters came the opening of transcontinental service in America in 1915 initially between New York and San Francisco, 3200 miles (5300 kilometers). This was followed after the close of the war by the rapid development of telephony of continental scope throughout Europe, stimulated by the close cooperation of the European Administrations through the International Advisory Committee on Long Distance Telephony. In South America the trans-Andean telephone line between Buenos Aires and Santiago was opened in 1928. Two years later came the transcontinental telephone line in Australia.

Overseas radio telephone experiments, in 1915, successfully transmitted the human voice from Washington to Paris and from Washington to the Hawaiian Islands. Commercial development of intercontinental telephony, however, followed somewhat slowly both because of the war and because of the tremendous inherent technical difficulties. However, as an interesting fore-runner of what would come later, a public service radiotelephone system was opened in 1920 linking Catalina Island, off the coast of California, with the wire telephone network of North America. In 1927, after a period of intensive experimentation and development work, commercial service between Europe and North America was opened to the public. This was the first step in the expansion of telephone service from a continental scope to a world-wide scope.

Present Situation and Plans for Future Development

Intercontinental telephony has naturally been dependent upon and been preceded by the development of large networks of telephones to which the intercontinental circuits could be connected. The most highly industrialized parts of the earth's surface today are provided with extensive networks of telephone lines covering great areas and interconnecting large numbers of telephones. This is indicated in Fig. 1 which shows the principal national and continental wire telephone networks, and all individual cities not connected to an extensive interurban network which according to latest reports have more than 10,000 telephones. As a result of the improvements which have been made in these wire networks in the last 15 years, both as to transmission and speed of service, they are today generally available and satisfactory for use in connection with intercontinental service.

For the purpose of this paper circuits of less than 600 miles (1000 kilometers) even if they cross continental boundaries will not be considered intercontinental circuits. In the consideration of world-wide telephony we are interested in the long circuits between distant parts of different continents or between continental areas and distant islands, for which the technique developed for long continental telephone lines is not directly applicable.

The extent to which the continents of the world are already interconnected by telephone circuits and the additional connections contemplated by plans now under way are indicated in Fig. 2. These are such as to make possible conversations between any two continents of the earth either by direct circuit or by switching in Europe or in

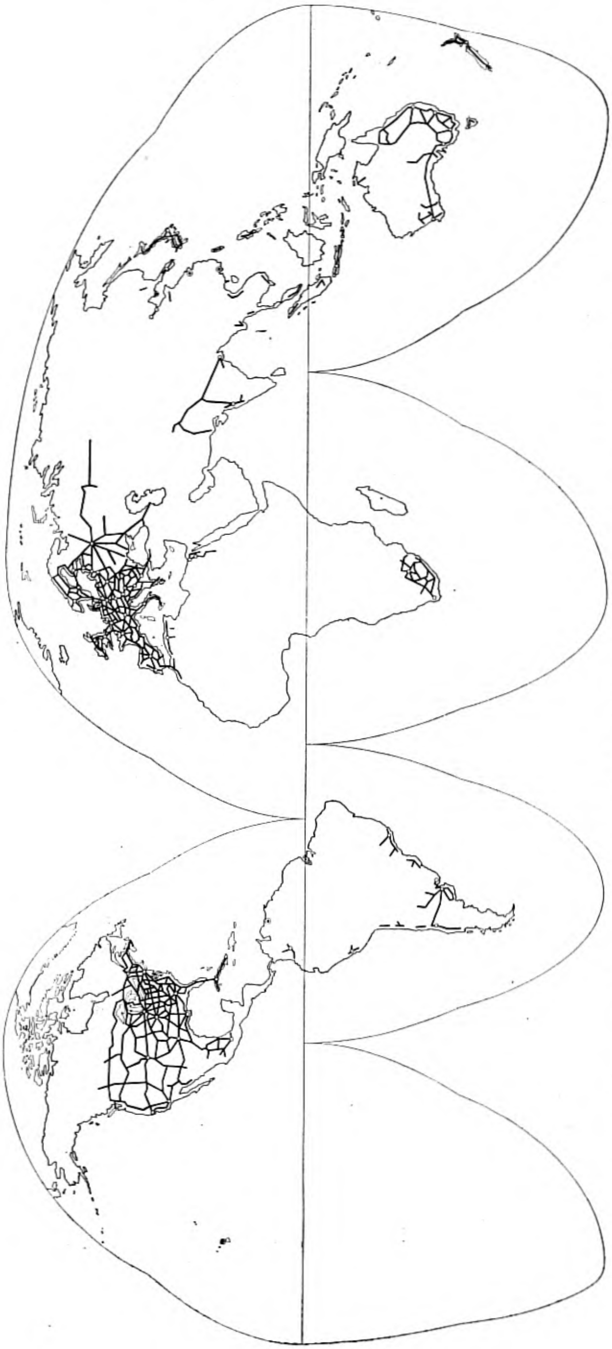
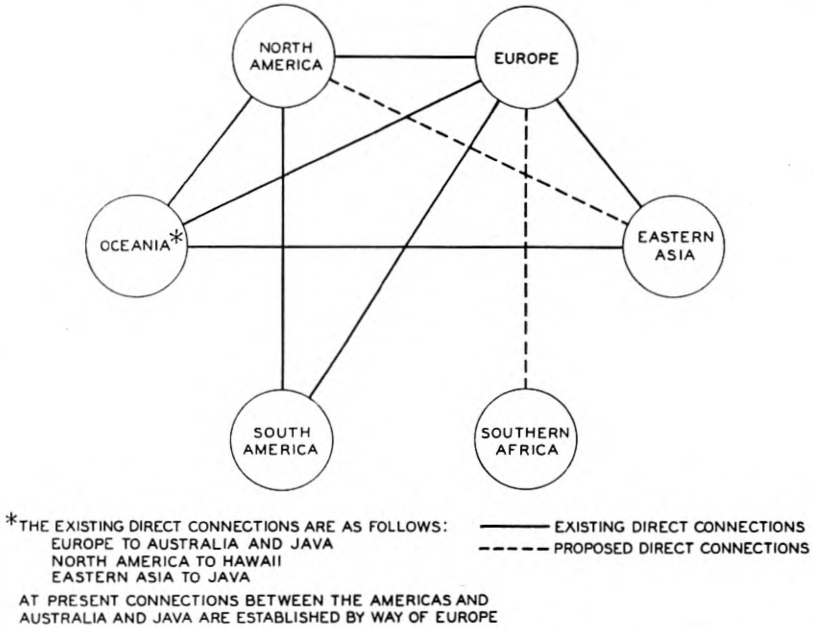


Fig. 1—Wire line telephone networks of the world. Includes indication of areas covered by international, national or interurban networks and individual cities having over 10,000 telephones. Based on information available January 1, 1932.

North America. The existing intercontinental telephone circuits of the world are indicated in Fig. 3. In this figure distinctions are made between full time circuits and part time circuits, that is, those on which terminal apparatus is shared by different points. A distinction is also made between circuits which interconnect wire networks connecting with 20,000 telephones or more, and those terminating in single telephones or networks of less than 20,000 telephones.



*THE EXISTING DIRECT CONNECTIONS ARE AS FOLLOWS:
EUROPE TO AUSTRALIA AND JAVA
NORTH AMERICA TO HAWAII
EASTERN ASIA TO JAVA

AT PRESENT CONNECTIONS BETWEEN THE AMERICAS AND AUSTRALIA AND JAVA ARE ESTABLISHED BY WAY OF EUROPE

Fig. 2—International telephone relations as of January 1, 1932.

There are at present 37 of these intercontinental circuits totalling about 168,000 miles (280,000 kilometers) in length. All are radio circuits. One, in the New York-London group, is a long-wave circuit operating at approximately 60 kilocycles; the others are short-wave circuits in the range between 6000 and 23,000 kilocycles.

Europe and North America, which are the two largest highly industrialized areas of the world, contain about 90 per cent. of the world's telephones. It is natural that intercontinental telephone business in volume should first develop between those two areas Here service is maintained on a 24-hour basis and a group of four circuits is in use. Elsewhere, however, intercontinental connections consist at the present time of a single circuit, or more frequently,

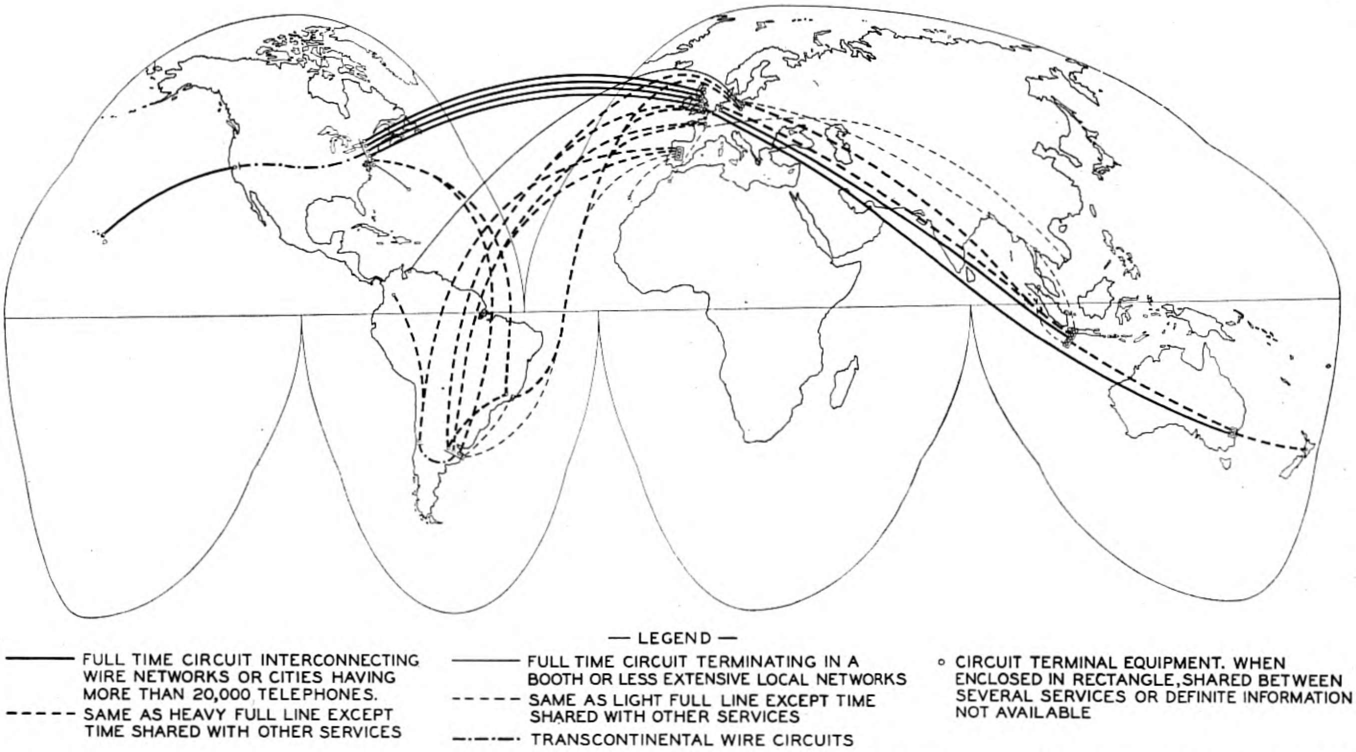


Fig. 3—Long intercontinental telephone circuits of the world. Circuits over 600 miles (1000 kilometers) in length as of January 1, 1932.

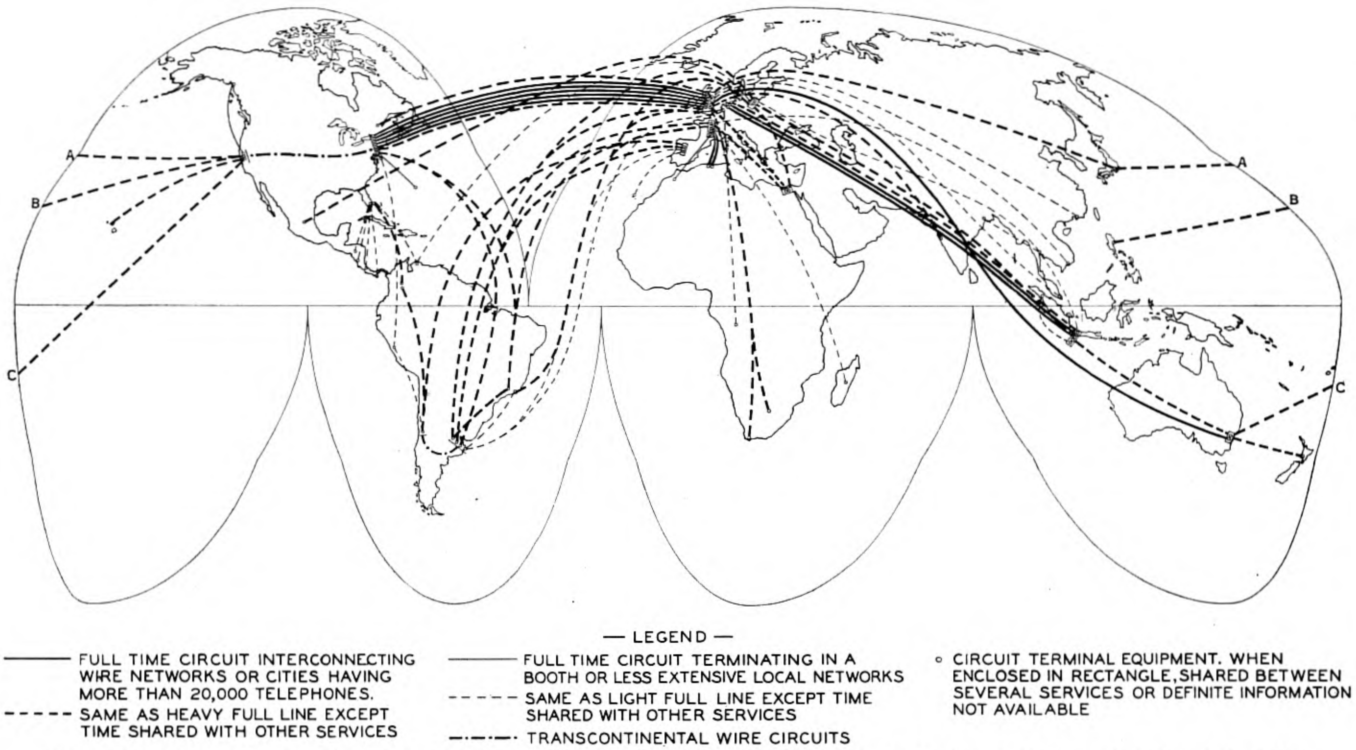


Fig. 4—Existing and proposed long intercontinental telephone circuits of the world. This map is similar to Fig. 3 and includes existing circuits and additional circuits now definitely proposed. Based on information available January 1, 1932.

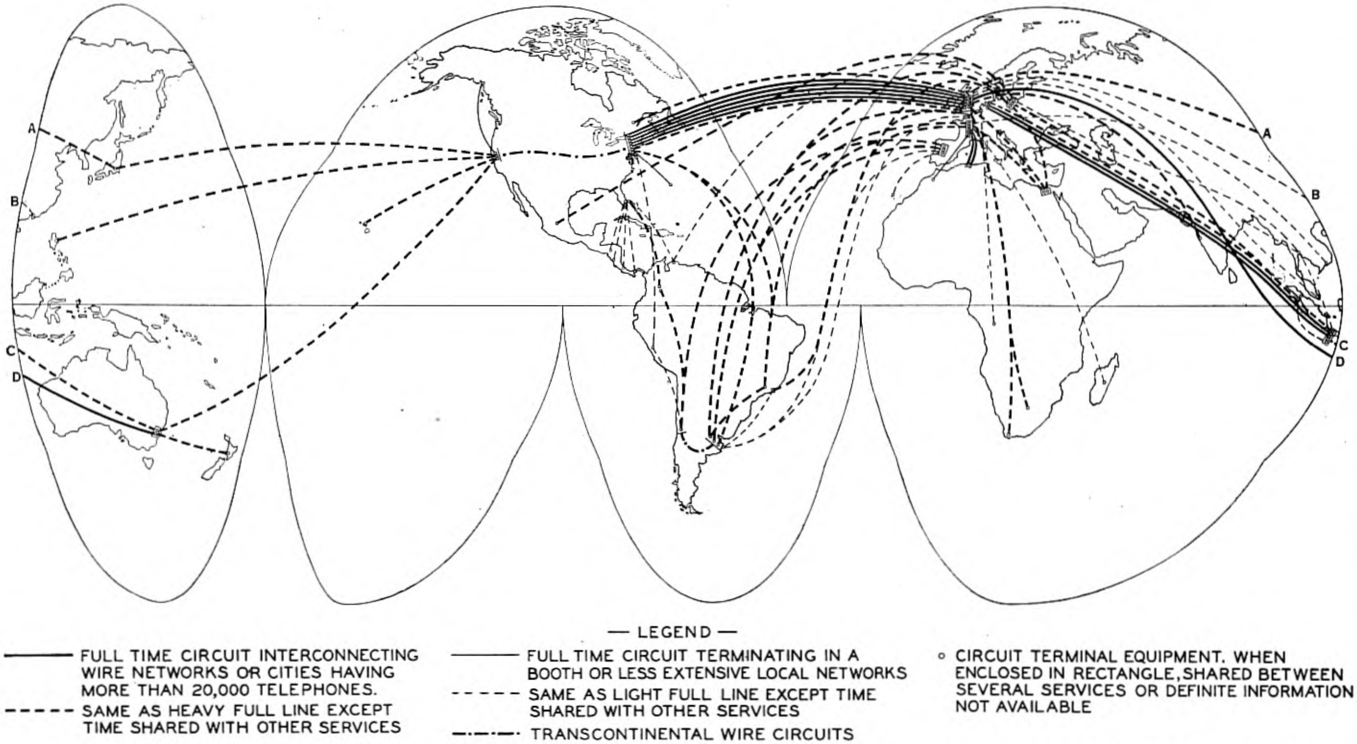


Fig. 5—Existing and proposed long intercontinental telephone circuits of the world. Arranged to show transpacific connections. Based on information available January 1, 1932. Circuits and legend same as shown in Fig. 4.

the part time use of a circuit, the apparatus being arranged to share its time between two or more distant terminals.

One cannot help being impressed by the developments represented by Fig. 3, which have taken place in the short space of five years. We are, nevertheless, in an early stage of development and present service suffers from the limitations of pioneer conditions. Dependence in so many cases upon a single intercontinental circuit which often is shared on a "party line" basis between two or more terminal points, is one of these limitations. Another is the variation in transmission characteristics and the susceptibility to interruptions which in the present stage of development of radio are characteristic especially of short-wave circuits on some of the more important routes. The rapidity with which the first steps have been taken is perhaps the best assurance that advance will continue to be rapid as the commercial demand for increasing amounts of service develops. This is illustrated by Figs. 4 and 5 which show the intercontinental circuits of the world which will exist when the present plans for additional circuits have been completed in so far as these plans are known to the authors.

Fig. 6 shows the extent to which the countries of the world will be tied into one great telephone system when the plans outlined in Figs. 4 and 5 are completed. Fig. 6 also shows the distribution of world trade and population of various groups of countries which are so interconnected. It will be noticed that the world-wide telephone network will include countries having 99 per cent of all the telephones of the world and having 92 per cent of the world's foreign trade. Details of all the present and proposed circuits which have come to the knowledge of the authors are given in Appendix I.

Closely related to the establishment of intercontinental circuits is the establishment of telephone service from European and North American points to a number of passenger vessels normally operating on the North Atlantic and other passenger routes. In addition, work is advancing in the equipment of fishing fleets, tugs, etc., but this service pertains more directly to the continental telephone service.

Technical Problems and Limitations

The unusual technical difficulties of providing satisfactory intercontinental telephone circuits come about not merely because of the distances involved but because these distances include long stretches of sea or undeveloped land. For years prior to the establishment of the first intercontinental circuit, commercial service was given on continental telephone networks over comparable distances, for example, in North America up to about 6000 miles (10,000 kilometers).

WORLD DISTRIBUTION

POPULATION, TELEPHONES, FOREIGN TRADE

For Countries Having Existing or Proposed International Telephone Connections as of January 1, 1932

Country	Total Foreign Trade 1930 (Millions of Dollars)	Total Telephones as of Jan. 1, 1931 (Thousands)	Total Population (Thousands)
EUROPE			
All Principal Countries	30,260	10,620	542,350
NORTH AMERICA			
United States, Canada, Cuba, Mexico, Bermuda	9,500	21,768	153,780
SOUTH AMERICA			
Argentina, Brazil, Chile, Colombia, Uruguay, Venezuela	2,590	595	71,090
OCEANIA			
Hawaii, Australia, New Zealand, Java, Sumatra	2,310	760	69,440
ASIA			
Siam, Indo-China	300	12	31,740
AFRICA			
Morocco, Canary Islands	120	14	5,850
Total in countries having existing connections to international networks	45,080	33,769	874,250
NORTH AMERICA			
Central America and Bahamas	120	18	2,720
SOUTH AMERICA			
Peru	130	14	6,260
OCEANIA			
Philippines	260	26	12,700
ASIA			
Japan, Hong Kong, British India, Malay States	3,810	1,028	410,320
AFRICA			
Algeria, Belgian Congo, Egypt, Union of So. Africa, Madagascar	1,370	197	47,700
Total in countries having prospective international connections	5,690	1,283	479,700
Total, present and prospective international network	50,770	35,052	1,353,950
All Other	4,140	308	622,550
Total World	54,910	35,360	1,976,500

Fig. 6.

The technique of transmission over long continental distances, however, includes as one fundamental element the use of intermediate amplification at frequent intervals of from 45 to 240 miles (75 to 400 kilometers), depending upon the electrical characteristics of the circuit.

Radio Links

All of the existing intercontinental circuits make use of radio for the long transoceanic jumps, and it seems desirable first to consider the technical problems of radio telephony.

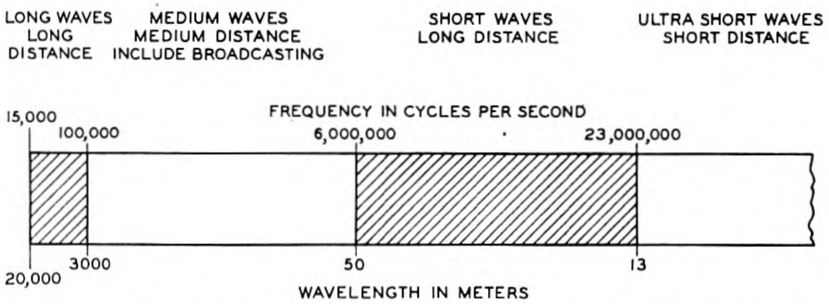


Fig. 7—Radio frequency chart. The bands applicable to intercontinental service are indicated by shaded areas.

Fig. 7 shows the radio spectrum with an indication of the uses made of the various portions of the spectrum. Only the two shaded parts of this spectrum appear to be applicable to intercontinental telephony. One, in the long wave range, extends from about 15 kilocycles to perhaps 100 kilocycles. The second includes the short-wave range from about 6000 to 23,000 kilocycles.

The transmission of these two wave ranges exhibit interesting differences in characteristics. For the short wave-lengths the transmission is frequently referred to as being in the form of "sky waves." This is for the reason that at intermediate distances the waves may practically disappear near the surface of the ground but reappear at greater distances. They appear to have been carried around the curvature of the earth's surface by refraction or reflection from ionized atmospheric layers. While the action in the long-wave range at great distances appears to be also conditioned partly by the ionized layers, the field at the surface of the earth falls off continuously as the distance from the transmitter is increased.

Interesting results are now being obtained in the use for communication purposes of very short radio waves having frequencies above 30,000 kilocycles. The work done to date indicates that these frequencies are not sufficiently deflected from their paths by the atmosphere to follow the earth's curvature. This characteristic appears to prevent the use of these very short radio waves for direct transmission over long distances and limits direct transmission to distances so short that the earth's curvature is not a large factor. For this reason the very high frequency range is not indicated in Fig. 7 as suitable for intercontinental circuits. Where the route is over-land it is possible that such rays may find practical use in forming links in intercontinental circuits particularly where the topography of the country affords advantageous elevated locations for intermediate repeater points. Too little is now known of these very short waves to make their discussion other than speculative.

An important requirement for radio for overseas telephone circuits is the avoidance of overhearing. A number of the overseas circuits now in operation, including the transatlantic group, are equipped with privacy arrangements which so modify the frequency disposition of the voice waves as to prevent overhearing of the conversations by other radio stations not equipped with similar arrangements. Experiments have been made with more elaborate arrangements for obtaining an even higher degree of privacy than that now provided.

Long Wave Circuits

Experience has shown that good results can be obtained in the long-wave range on such circuits as those between North America and Europe which are 3000 miles (5000 kilometers) long and have their transmission paths at a high latitude. A fundamental limitation to the extensive use of long waves, however, is that within the range of suitable frequencies there are theoretically obtainable only about twenty telephone channels between any two points. Actually however, there are not this many available because of the practical limitations imposed by the necessity of sharing this range of frequencies with other types of radio service. The relatively low attenuation of these waves as they are transmitted over the earth's surface makes it appear impractical to use duplicate intercontinental circuits of the same frequency at different points on the earth's surface. The total number of circuits of the long-wave type which can form a part of the ultimate world-wide telephone network seems, therefore, to be very small.

Other limitations in the use of long-wave circuits are that with the amounts of power which it now seems practicable to use (a tube capacity of about 300 kilowatts is used in the transmitter of the present transatlantic long wave circuit) their successful application does not exceed several thousand kilometers and is confined to routes well outside of the equatorial regions. These limitations are necessary to avoid excessive interference from atmospheric disturbances, both because of the relatively large components of these disturbances having frequencies in the range of wave-lengths used by these circuits, and also because of the difficulty of obtaining a high degree of directivity. Advantage has been obtained in the transatlantic long-wave channel by locating directive receiving antennas in the most northerly portions of the United States and Great Britain.

Short-Wave Radio

The short-wave range, as indicated in Fig. 7, covers a very much broader band of frequencies and offers hope for a much larger number of circuits. Only a portion of the entire range can be used at any one time for circuits between any two points because the distance at which the refracted waves reach the earth's surface varies with the frequency. It also varies with the time of day and the season of year, making it necessary, in general, to have about three different wave-lengths allocated for each circuit. In view of this limitation, it appears that this range can provide theoretically perhaps 50 telephone circuits between two given points. This theoretical possibility is reduced by practical considerations and by the necessity of sharing the range with other radio services.

Due, however, to the fact that short waves are more restricted than long waves in distance and in the time of day for which they are effective, it may be possible to use the same frequency for different circuits in several parts of the world without interference. It seems now that the development of the present art may make available throughout the entire world short-wave intercontinental circuits numbered in the hundreds but not in the thousands. This is not an inexhaustible supply of channels for the world to share as is illustrated by the very much larger number of telephone circuits already required for long distance service in the great continental telephone networks. In the United States alone there are about 6500 circuits in long distance service of which 620 are over 600 miles (1000 kilometers) in length.

The fact that by proper selection of frequency it is possible to communicate over the longest distances, gives short waves a flexibility which is unique. They give particularly good results in routes crossing the equator, such as circuits between Europe or North America and South America, for which routes long-wave transmission would be impracticable because of heavy atmospheric disturbances. On the other hand, short-wave transmission is subject to variations and interruptions from fading, distortion, and effects connected with terrestrial magnetic disturbances. These effects appear to be particularly severe on circuits involving routes of transit near the polar regions, such as the route between North America and Europe. Some of these effects are illustrated in Figs. 8 and 9 which give data taken on the United States-Europe route and which show, for comparison, the relatively stable conditions for the long-wave channel. In many respects short-wave and long-wave radio channels admirably supplement each other, being seldom interrupted at the same time.

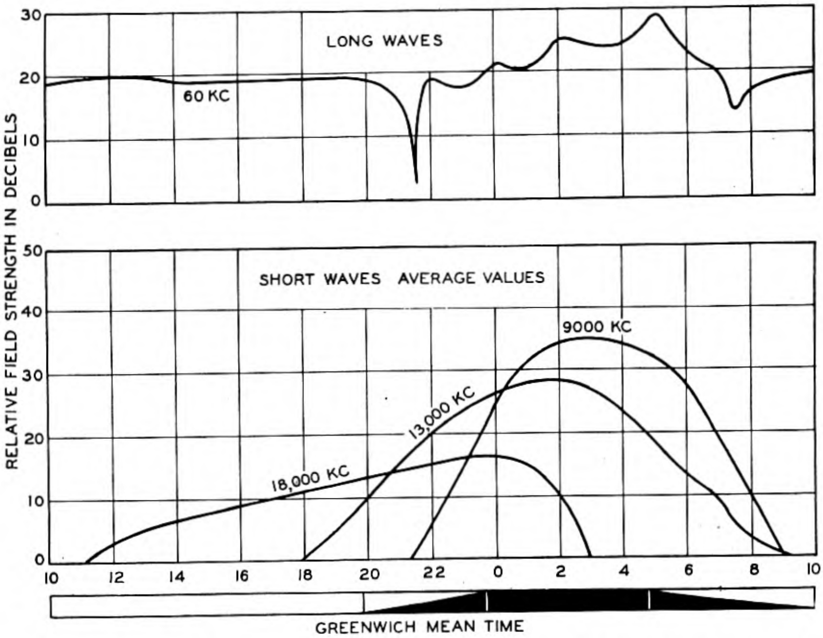


Fig. 8—Diurnal field strength characteristics of long and short-wave transmission over the North Atlantic. Shows effect of sunrise and sunset on the long wave and the necessity with short waves for using different wave-lengths at different times of day.

Short-wave transmission is in an early stage of development and it may be hoped that, with further experience, means will be found to avoid or reduce at least some of the present major limitations.

Wire Circuits

The intercontinental routes offer great difficulties to the placing of wire circuits. However, the limitations of radio, particularly on the intercontinental routes between North America and Europe interconnecting the two largest continental groups of telephones, have naturally led to considerable engineering thought on the possibilities of wire circuits. The most interesting development in this connection is a submarine cable which can be used for long lengths without intermediate repeaters.

Long Submarine Cables Without Intermediate Repeaters

To design long submarine cables without repeaters requires providing the mechanical characteristics necessary for deep sea cables and meeting rather definite limits of overall attenuation. There is a limit to the power which can be imposed on the sending end of the cable to avoid excessive voltage stress on the cable insulation. Even though the cable be adequately shielded from extraneous noises, there is a minimum amount of received power with which satisfactory transmission can be obtained because of the thermal agitation in the conductor itself. These two limits result in a maximum attenuation of about 160 db for satisfactory operation over such a cable.

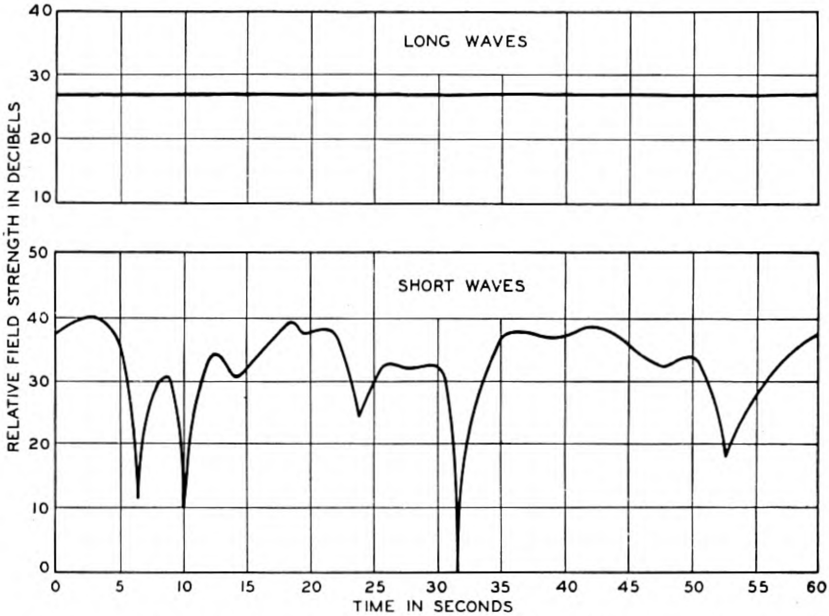


Fig. 9—The upper curve indicates the short period stability of the long waves. The lower curve indicates the rapid variations which are frequently experienced with short-wave transmission.

The actual design of intercontinental telephone cables has been largely directed toward the North Atlantic route between the United States and Europe, including a direct section from Newfoundland to Ireland, a distance of about 2100 miles (3400 kilometers). By the use of a new insulating material, paragutta, with dielectric losses about one-thirtieth those of guttapercha, but with similar mechanical characteristics, and by the use of perminvar, a new magnetic material for loading the cable, it has been possible to design a cable for this section meeting both the mechanical and electrical requirements. With the further development of intercontinental communications, it is proposed to lay such a cable within a few years. With the materials now available, such a cable will provide only one channel of communication, and two-way communication will be carried out by using it as a one-way channel automatically switched in the direction of transmission by the voices of the users. Transatlantic cables providing several telephone channels either by the use of separate conductors or by carrier current methods are beyond the range of the present art.

The direct submarine cable while relatively expensive, is expected to provide a high grade circuit having a stability and a freedom from interruption greater than that provided by the present radio circuits. Proper combination of three different types of circuit, long and short-wave radio and cable, in one circuit group, however, should provide a large measure of assurance of continuous high grade service not dependent on the troubles which may affect any one type of circuit, and at an average expense not greatly above that required for radio circuits alone.

Other Wire Routes

The technical possibilities of wire circuits on intercontinental routes are evidently greatly increased where means can be found to avoid long lengths of submarine cables by the use of intermediate repeater stations. On the direct route between Europe and America, nature has not been kind enough to supply a series of islands at convenient distances and reasonably low latitudes to serve as repeater points. By going north, long lengths of submarine could be avoided. For example, a route through Greenland, Iceland and the Faroe Islands could be laid out with a maximum length of submarine cable of 300 miles (500 kilometers). The obvious difficulties of the route are the great extents of inaccessible and sparsely settled country, and the placing and maintenance of submarine cable under very difficult fog, storm and ice conditions. In the early days of submarine telegraphy such a northern route was seriously considered before the cable art had reached the point of permitting direct cables, but it was never used.

An equally bold solution which has been proposed is to float the desired repeater stations in the open sea. A good deal of ingenuity has been exercised in considering the possibilities of both attended and unattended floating repeater stations and of stations submerged below the action of the waves. So far, however, it is by no means clear that the mechanical difficulties and the problems of maintenance can be dealt with satisfactorily.

As time goes on land lines may become possible for many intercontinental routes where now, because of lack of highways or railways and the wild and unsettled nature of the country, they are out of the question.

OPERATING AND COMMERCIAL PROBLEMS

In addition to the technical problems and limitations associated with the development of a world-wide telephone service which are outlined in the foregoing section, there are numerous operating and commercial problems and still other difficulties of a general nature.

Differences in Time

Limitations of a fundamental character are imposed by the differences in time on different parts of the earth's surface. These differences for a number of the principal metropolitan centers of the world are indicated in Fig. 10. If the business day be assumed to be 8 hours long, it is evident that as a result of the differences in time there is for each city one-third of the earth's surface on which the time is so different from that of the city in question that there is no overlap of the business day. For western Europe this third of the world is for the most part in the Pacific Ocean, so that, as will be noted from Fig. 10, there are few important centers in the world in which the time is more than 8 hours different from that of western Europe. The western part of the United States, however, has time differences of more than 8 hours with a large part of Europe, Asia and Africa. During some portion of the waking day there is an overlap of time between any two points on the earth.

Language

Another limitation, though of a less inherent nature, is the difference in language between different nations. There is as yet no evidence that the electrical requirements for satisfactory transmission of speech are substantially different with different languages, but language differences produce a problem in many intercontinental telephone conversations as often one subscriber is using a language of which he is not wholly master, and this may sometimes apply to both. Under these conditions the transmission requirements of the circuit for a given ease of carrying on conversations is unquestionably greater than when the subscribers are conversing in their native tongue. Hence, there is need in the ultimate development of intercontinental telephone systems for a high standard of transmission.

Language differences also complicate the operating problem, as it is important that the operators understand each other easily. This is

Hours Time Difference Between Certain Cities in Various Parts of the World

A → B ↓	Los Angeles San Francisco Vancouver	Chicago New Orleans Mexico City * Winnipeg	New York Washington Montreal Santiago	Buenos Aires Halifax	Rio de Janeiro	Paris London Madrid
Los Angeles San Francisco Vancouver	–	+ 2	+ 3	+ 4	+ 5	+ 8
Chicago New Orleans Mexico City * Winnipeg	– 2	–	+ 1	+ 2	+ 3	+ 6
New York Washington Montreal Santiago	– 3	– 1	–	+ 1	+ 2	+ 5
Buenos Aires Halifax	– 4	– 2	– 1	–	+ 1	+ 4
Rio de Janeiro	– 5	– 3	– 2	– 1	–	+ 3
Paris London Madrid	– 8	– 6	– 5	– 4	– 3	–
Berlin Rome Stockholm Vienna	– 9	– 7	– 6	– 5	– 4	– 1
Cairo Capetown Istambul Moscow	–10	– 8	– 7	– 6	– 5	– 2
Bombay Delhi	+10.5	–11.5	–10.5	– 9.5	– 8.5	– 5.5
Bangkok Singapore	+ 9	+11	+12	–11	–10	– 7
Peiping Shanghai Manila Perth	+ 8	+10	+11	+12	–11	– 8
Tokyo Vladivostock	+ 7	+ 9	+10	+11	+12	– 9
Sydney Melbourne	+ 6	+ 8	+ 9	+10	+11	–10

Notes

Computed from "The Time Zone Chart of the World," published by the Navy Department of the United States.

+ Indicates that cities in the row designated "A" have later clock time than those in the column designated "B" and lie within 180° toward the east.

Fig. 10

A → B ↓	Berlin Rome Stockholm Vienna	Cairo Capetown Istambul Moscow	Bombay Delhi	Bangkok Singapore	Peiping Shanghai Manila Perth	Tokyo Vladivostock	Sydney Melbourne
Los Angeles San Francisco Vancouver	+ 9	+10	−10.5	− 9	− 8	− 7	− 6
Chicago New Orleans Mexico City Winnipeg	+ 7	+ 8	+11.5	−11	−10	− 9	− 8
New York Washington Montreal Santiago	+ 6	+ 7	+10.5	−12	−11	−10	− 9
Buenos Aires Halifax	+ 5	+ 6	+ 9.5	+11	−12	−11	−10
Rio de Janeiro	+ 4	+ 5	+ 8.5	+10	+11	−12	−11
Paris London Madrid	+ 1	+ 2	+ 5.5	+ 7	+ 8	+ 9	+10
Berlin Rome Stockholm Vienna	−	+ 1	+ 4.5	+ 6	+ 7	+ 8	+ 9
Cairo Capetown Istambul Moscow	− 1	−	+ 3.5	+ 5	+ 6	+ 7	+ 8
Bombay Delhi	− 4.5	− 3.5	−	+ 1.5	+ 2.5	+ 3.5	+ 4.5
Bangkok Singapore	− 6	− 5	− 1.5	−	+ 1	+ 2	+ 3
Peiping Shanghai Manila Perth	− 7	− 6	− 2.5	− 1	−	+ 1	+ 2
Tokyo Vladivostock	− 8	− 7	− 3.5	− 2	− 1	−	+ 1
Sydney Melbourne	− 9	− 8	− 4.5	− 3	− 2	− 1	−

− Indicates that cities in the row designated "A" have earlier clock time than those in the column designated "B" and lie within 180° toward the west.

This does not take account of the change in date in crossing the International date line.

* Time after April 1, 1932.

Fig. 10 (continued).

accomplished to a large extent at the present time by providing bi-lingual operators at the terminals of international circuits. Even so, on built-up connections with switches in several countries, operation is cumbersome since the terminal operators may not be able to talk to each other directly. This is one of the problems to which the International Advisory Committee on Long Distance Telephony has been giving attention. It is out of the question in a world-wide telephone system to expect the operator handling the call to be able in all cases to talk directly with the distant subscriber. However, it seems important from the standpoint of giving the highest grade of service that ultimately she should be able in all cases to talk directly and easily with the operator at the distant national terminal, no matter where this may be.

Operating and Commercial Practices

The continental telephone networks have in the past developed to a large extent independently of each other, and it is therefore natural that there should now exist differences in operating and commercial practices, some of which must be considered in giving intercontinental service. In Europe and in the United States, to take, for example, the two largest networks, the point of view in the development of the telephone systems has been somewhat different. In Europe emphasis has been laid on the continuous use of the long toll circuits, developing for them as large a message capacity as possible, while in the United States emphasis has been given to the rapid completion of all calls. This difference in point of view naturally led to differences in practice. For example, the classification of service based upon urgency and the limitation in length of conversations generally in use in Europe are not found in American practice. Also, an important factor in American practice for connections over long distances is the so-called person-to-person service in which a specified person is called rather than a specified telephone number, and the order is considered satisfied only when the person called is brought to the telephone. This service has not been generally used in European practice. These serve as illustrations of the type of difference in practice which must be adjusted between the administrations involved in establishing new intercontinental services, and it is evident that as these services become more used and more nearly universal, these adjustments will become increasingly important.

A consideration which might be important in the development of intercontinental services in some cases is a difference in standards of transmission or of speed of completion of calls. In recent years there has been in all large telephone networks a trend towards higher standards of service and this has been a favorable factor in making possible the beginnings of an intercontinental service.

General Discussion and Conclusions

The extension of overseas telephony during the past five years has already linked together into one system all of the largest continental telephone networks of the world and with the completion of further extensions now under way, this world-wide system will include all but two of the wire networks which give access to more than 20,000 telephones in all six continents of the earth. True, many of these overseas connections are as yet but slender threads of conversation, important perhaps not so much because of the communication which they now handle but because they represent the first realization of great possibilities in the achievement of a world-wide telephone system closely linking together the continents of the earth. In the words of Mr. Walter S. Gifford, President of the American Telephone and Telegraph Company, in the Annual Report of that company for the year 1926, the ultimate ideal of a world-wide system is that it shall enable "anyone anywhere to pick up a telephone and talk to anyone else anywhere else, clearly, quickly and at a reasonable cost."

What are the obstacles to the realization of this ideal for world-wide telephony? Intercontinental service is subject to extraordinary technical and operating difficulties and as yet only the first steps in overcoming these difficulties have been taken. The quality of service at the present time both as regards transmission, continuity and speed of service is not comparable with that given today on the large continental telephone networks, but is more comparable to the standards of service on long continental connections in the early days of their development 15 years ago. The costs of intercontinental service are materially higher than the costs for the longest connections on continental networks even where distances are similar. These facts indicate that to a large extent the future development of intercontinental telephony is dependent upon a continuance of those brilliant advances in the communication art which have made our present intercontinental circuits possible. Further technical developments making it possible to improve the quality and reduce the cost of intercontinental services will, as they become available, have a tremendous effect.

The development of overseas services of large magnitude will require a closer coordination between the telephone plants and practices of the various continental wire networks than exists at the present time. The circuits which will form the connecting link between the subscriber and the terminals of the overseas circuits must have such transmission characteristics as to provide for satisfactory operation of the complete connection, including the overseas circuit and the continental extensions at both ends. In one respect this will impose requirements more severe than necessary where continental service alone is in question. We refer to the velocity of transmission which determines the elapsed time between the speaking of a word at one end of a circuit and its reception at the other end. Losses of power and electrical distortions in circuits may, within limits, be compensated for but time once lost in the propagation of the conversation over the circuit cannot be regained.

Equally important is the closer coordination of operating methods and commercial practices in so far as they affect intercontinental communications. It is natural that continental telephone networks developing more or less independently should represent somewhat different solutions of the operating and commercial problems involved in giving telephone service. Intercontinental service brings new problems and requires the development of new operating methods and new commercial practices designed to simplify and expedite the handling of these connections.

In the closer cooperation between telephone administrations and the consideration of their joint problem which comes with the development of intercontinental telephone service, it is increasingly important that they have commonly accepted methods of measuring and of expressing all of the quantities affecting the types and grade of service to be given. The International Advisory Committee on Long Distance Telephony (The C. C. I.) which has been active in facilitating the cooperation of European administrations in the improvement of international telephony in Europe, has included in its program the development of internationally accepted terms and units which will be of help to the nations of the world in their rendering of intercontinental service.

The commercial success of an intercontinental telephone service depends upon the existence at its terminals of wire networks by means of which large numbers of telephone subscribers can be given connection to the intercontinental circuits. The ultimate development of a really universal service depends, therefore, in part on the creation of large national or continental wire telephone networks in the areas where these do not now exist. It is only in so far as this takes place that it will become practical to realize world-wide telephony.

As world-wide telephony overcomes these obstacles and develops in magnitude and completeness, what form will the system take? It is, of course, too early to give a categorical answer to this question, but present development gives us some indications from which to judge the future.

Considering first the field of use of the various types of circuit, it seems that radio circuits will for a long time fill an important field in the provision of intercontinental circuits. This is particularly true of the short-wave systems which seem to be best adapted for pioneer work such as is going on at the present time on light traffic routes. Here their imperfect reliability is more than offset by their flexibility and relative economy. Long-wave radio will, no doubt, continue to be valuable for certain routes where the direction of transmission is east and west and at a high latitude, but the limit, which is apparently inherent, on the total number of circuits of this type that can be used simultaneously in the world would seem to prevent them from supplying any large part of the world's future needs for intercontinental circuits. Radio telephony, both short-wave and long-wave, must compete for wave-lengths with other forms of radio service. It is evident from the important part that radio must continue to play in world telephony that the increasing needs for wave-lengths for this rapidly growing service will require special consideration in future international radio conferences. Wire circuits which in the present pioneer stage are just beginning to enter the scene, will undoubtedly become more important for the principal circuit groups as intercontinental communication develops.

As the amount of intercontinental traffic builds up, and as the technical form and best routing of the telephonic relations between continents become established, it is to be expected that experience will show the advisability of adopting a fundamental routing and switching plan similar to those plans which have already been considered and put into use for some of the large continental networks. It is now too early to suggest in any detail the form of such a switching plan, but it seems that while political considerations may temporarily affect the form of the network, ultimately the requirements of economy and good service which have determined the form of continental plans now in use will be weighty factors in the planning of a world-wide fundamental switching plan. This gives a clue to some of its characteristics.

The splitting of circuit time between different terminal wire networks, while a valuable expedient for offering service under pioneer conditions, will naturally disappear generally as sufficient traffic develops to justify a full time circuit on a given route. This may be expected to be followed by the development of small circuit groups between the more important continental telephone networks, each group being operated as a unit between fixed terminal points. The advantages, while circuit groups are small, of concentrating intercontinental traffic as far as possible between the same terminals rather than diverting it to individual circuits between different terminal points, are great. As an illustration, between the United States and Europe where the volume of traffic has already led to the development of a circuit group, the operation of the six circuits now contemplated as a unit is estimated to give for the same grade of service a capacity one third greater than would be afforded by six separate circuits of the same character between different terminal points. The inclusion in one group of circuits of three different types, short-wave radio, long-wave radio and cable, will afford a continuity of service and an insurance against interruption far beyond what could be achieved with single circuits. A single circuit group between the two continental networks is also advantageous from a service standpoint because of the simpler operating arrangements.

While the first stage in the development of intercontinental business appears to indicate the concentration of intercontinental traffic in so far as the extent of continental networks makes this practicable, a second stage in the development will naturally be the establishment of additional circuit groups between other points of the networks. This becomes economical and desirable from a service standpoint when the original circuit group becomes large enough to permit of subdivision without great loss in efficiency and when the amount of traffic which can be conveniently handled through additional points in the continental wire networks is sufficient to fill the time of a group of several circuits. Hence, it is to be expected that the ultimate switching plan for world-wide telephony will include between large continental networks, such for example as exist in Europe and in the United States, a number of groups of overseas circuits between different terminal points selected so as to handle the traffic most conveniently and economically. Such a plan would also necessarily provide arrangements for the use of alternate routes. In any case the ultimate best plan from the standpoint of service and economy will depend upon the volume of traffic.

The technical achievements which have made possible the linking together of the continents of the earth with telephone circuits are in a high degree romantic. What may be accomplished for the benefit of mankind by the continued development of this world-wide telephone network depends upon what is said over the telephone circuits. It is the hope of scientists and engineers who have been engaged in this work that this closer knitting together of the nations and races of the world will be a great contribution to international friendship and good will. Mr. Coolidge, President of the United States in 1927, on the occasion of the opening of the telephone connection between the United States and Spain, gave apt expression to this thought when he said, "I believe it to be true that when two men can talk together the danger of any serious disagreement is measurably lessened, and that what is true of individuals is true of nations. The international telephone which carries the warmth and the friendliness of the human voice will always correct what might be misinterpreted in the written word."

APPENDIX I

INTERCONTINENTAL TELEPHONE CIRCUITS OF THE WORLD
JANUARY 1, 1932

Existing Circuits[1]

Circuit Group Index No.	Circuit Designation	Ownership	Distance Kilometers*	Service Date	Rates[2]	Time Sharing Arrangements
	NORTH AMERICA—EUROPE					
101	(1) London—New York (Long Wave)	British P. O.—Am. Tel. and Tel. Co.	5550	1/ 7/27	$30.00	Not shared
	(2) London—New York	British P. O.—Am. Tel. and Tel. Co.	5550	6/ 6/28	$30.00	Not shared
	(3) London—New York	British P. O.—Am. Tel. and Tel. Co.	5550	6/ 1/29	$30.00	Not shared
	(4) London—New York	British P. O.—Am. Tel. and Tel. Co.	5550	12/ 1/29	$30.00	Not shared
	NORTH AMERICA—SOUTH AMERICA					
201	Buenos Aires—New York	Compania Internacional de Radio (I. T. and T. Co.)—Am. Tel. and Tel. Co.	8580	4/ 3/30	$30.00	At New York with 202
202	New York—Rio de Janeiro	Am. Tel. and Tel. Co.—Companhia Radio Internacional do Brasil (I. T. and T. Co.)	7810	12/18/31	$30.00	At New York with 201 At Rio de Janeiro with 802
	EUROPE—SOUTH AMERICA					
301	Buenos Aires—London	Compania Internacional de Radio (I. T. and T. Co.)—British P. O.	11140	12/12/30	£ 6/0	At Buenos Aires with 307, 802 At London with 302, 303
302	Buenos Aires—London	Transradio Internacional, Compania Radiotelegrafica Argentina S. A.—British P. O.	11140	12/12/30	£ 6/0	At Buenos Aires with 305, 308, 312, 801 At London with 301, 303

---

* 1 kilometer ≈ .6 mile.

[1] See notes at the end of this appendix.

APPENDIX I (*Continued*)

Circuit Group Index No.	Circuit Designation	Ownership	Distance Kilometers	Service Date	Rates[2]	Time Sharing Arrangements
	EUROPE— SOUTH AMERICA (*cont.*)					
303	London—Rio de Janeiro	British P. O.—Companhia Radiotelegraphica Brasileira (Transradio)	9290	5/21/31	£ 6/0	At London with 301, 302 At Rio de Janeiro with 306, 309, 314, 801
304	Buenos Aires—Paris....	Compania Internacional de Radio (I. T. and T. Co.)—Cie. Gle. de T. S. F.	11060	6/11/30	150 G. F.	At Buenos Aires with 311 At Paris with 305, 306
305	Buenos Aires—Paris....	Transradio Internacional Compania Radiotelegrafica Argentina S. A.—Cie. Gle. de T. S. F.	11060	2/ 1/29	150 G. F.	At Buenos Aires with 302, 308, 312, 801 At Paris with 304, 306
306	Paris—Rio de Janeiro...	Cie. Gle. de T. S. F.—Companhia Radiotelegraphica Brasileira (Transradio)	9170	3/31/30	750 F.	At Paris with 304, 305 At Rio de Janeiro with 303, 309, 314, 801
307	Berlin—Buenos Aires...	German P. O.—Compania Internacional de Radio (I. T. and T. Co.)	11920	9/10/30	120.00 M.	At Berlin with 308, 309 At Buenos Aires with 301, 802
308	Berlin—Buenos Aires...	German P. O.—Transradio Internacional Compania Radiotelegrafica Argentina S. A.	11920	12/10/28	120.00 M.	At Berlin with 307, 309 At Buenos Aires with 302, 305, 312, 801
309	Berlin—Rio de Janeiro..	German P. O.—Companhia Radiotelegraphica Brasileira (Transradio)	10030	3/21/30	120.00 M.	At Berlin with 307, 308 At Rio de Janeiro with 303, 306, 314, 801
310	Berlin—Maracay, Ven. .	German P. O.—Venezuelan Govt.	8300	9/13/31	120.00 M.	Not shared
311	Buenos Aires—Madrid. .	Compania Internacional de Radio (I. T. and T. Co.)—Compagna Telephonica Nationale d'Espagna (I. T. and T. Co.)	10060	10/12/29	150 G. F.	At Buenos Aires with 304 At Madrid with 313
312	Buenos Aires—Madrid. .	Transradio Internacional Compania Radiotelegrafica Argentina S. A.—Compania Transradio Espanola (Transradio)	10060	1929	150 G. F.	At Buenos Aires with 302, 305, 308, 801 At Madrid, with 314

APPENDIX I (*Continued*)

Circuit Group Index No.	Circuit Designation	Ownership	Distance Kilometers	Service Date	Rates [2]	Time Sharing Arrangements
	EUROPE—SOUTH AMERICA (*cont.*)					
313	Madrid—Santiago......	Compagna Telephonica Nationale d'Espagna (I. T. and T. Co.)—Compania Internacional de Radio, S. A. (I. T. and T. Co.)	10700	4/17/31	165 G. F.	At Madrid with 311 At Santiago with 803, 804
314	Madrid—Rio de Janeiro	Compania Transradio Espanola—Companhia Radiotelegraphica Brasileira (Transradio)	8140	1930		At Madrid, with 312 At Rio de Janeiro with 303, 306, 309, 801
	EUROPE—AFRICA					
401	Casablanca—Paris.....	Moroccan Tel. & Tel. Admin.—Cie. Gle. de T. S. F.	1820	11/ 3/30	111 F.	At Paris with 503
402	Madrid—Teneriffe.....	Compagna Telephonica Nationale d'Espagna (I. T. and T. Co.)—Compagna Telephonica Nationale d'Espagna (I. T. and T. Co.)	1820	1/22/31	30 Pesetas	At Madrid with Madrid-Mallorca
	EUROPE—ASIA—OCEANIA					
501	London—Sydney.......	British P. O.—Amalgamated Wireless Australasia	17000	4/30/30	£ 6/0	Note 3
502	(1) Amsterdam—Bandoeng...........	Neth. Govt.—Neth. Indies Tel. Admin.	11740	1/ 8/29	33 Florins	At Amsterdam—Not shared At Bandoeng—Note 5
	(2) Amsterdam—Bandoeng...........	Neth. Govt.—Neth. Indies Tel. Admin.	11740	12/ 1/29	33 Florins	At Amsterdam—Not shared At Bandoeng—Note 5
503	Paris—Saigon.........	Cie. Gle. de T. S. F.—Cie. Gle. de T. S. F.	10120	4/11/30	450 F.	At Paris with 401
504	Bandoeng—Berlin......	Neth. Indies Tel. Admin.—German P. O.	10830	12/29/29	96.00 M.	At Bandoeng—Note 5 At Berlin with 505
505	Bangkok—Berlin.......	Siamese Govt.—German P. O.	8610	4/15/31	96.00 M.	At Bangkok—Note 3 At Berlin with 504

APPENDIX I (*Continued*)

Circuit Group Index No.	Circuit Designation	Ownership	Distance Kilometers	Service Date	Rates [2]	Time Sharing Arrangements
	NORTH AMERICA—ASIA OCEANIA					
601	Honolulu—San Francisco	Mutual Tel. Co.—Am. Tel. and Tel. Co.	3850	12/23/31	$21.00	Not shared
	NORTH AMERICA					
701	Hamilton (Bermuda)—New York	Imp. & Int. Comm. Ltd.—Am. Tel. and Tel. Co.	1280	12/21/31	$15.00	Not shared
	SOUTH AMERICA					
801	Buenos Aires—Rio de Janeiro	Transradio Internacional Compania Radiotelegrafica Argentina S. A.—Companhia Radio telegraphica Brasileira (Transradio)	1970	1931		At Buenos Aires—with 302, 305, 308, 312 At Rio de Janeiro—with 303, 306, 309, 314
802	Buenos Aires—Rio de Janeiro	Compania Internacional de Radio (I. T. and T. Co.)—Companhia Radio Internacional do Brazil (I. T. and T. Co.)	1970	12/12/31	75 G. F.	At Buenos Aires with 307, 301 At Rio de Janeiro with 202
803	Bogota—Santiago	All America Cables Inc. (I. T. and T. Co.)—Compania Internacional de Radio S. A. (I. T. and T. Co.)	4260	8/ 1/31	105 G. F.	At Santiago with 313
	ASIA—OCEANIA					
901	Sydney—Wellington	Australian P. O.—New Zealand Govt.	2210	11/25/30		Note 3
902	Bandoeng—Bangkok	Neth. Indies Tel. Admin.—Siamese Govt.	2360	4/15/31	37.50 Florins	At Bandoeng—Note 5 At Bangkok—Note 3
903	Bandoeng—Sydney	Neth. Indies Tel. Admin.—Australian P. O.	5460	12/23/30	73.00 Florins	At Bandoeng—Note 5 At Sydney—Note 3
904	Bandoeng—Medan (Sumatra)	Neth. Indies Tel. Admin.—Neth. Ind. Tel. Admin.	1410	9/16/31	12.00 Florins	At Bandoeng—Note 5

APPENDIX I (*Continued*)

Intercontinental Telephone Circuits of the World
January 1, 1932

Proposed Circuits[1]

Circuit Group Index No.	Circuit Designation	Ownership	Distance Kilometers	Service Date[4]	Rates[2]	Time Sharing Arrangements
	North America—Europe					
101	(5) London—New York (Long Wave)	British P. O.—Am. Tel. and Tel. Co.	5550	1934	$30.00	Not shared
	(6) London—New York (Cable)	Note 3—Am. Tel. and Tel. Co.				Not shared
102	London—Montreal	British P. O.—Canadian Marconi Ltd.	5220	1932	£ 6/0	At London—with 404 At Montreal—Not shared
103	Berlin—Mexico City	German P. O.—Mexican Govt.	9720			Note 3
	North America—South America					
203	Lima—New York	Compania Peruana de Telefonos, Limitada (I. T. and T. Co.)—Am. Tel. and Tel. Co.	5920	1932		Note 3
204	Bogota—Miami	Compania Telefonica Central (Assoc. T. & T. Co.)—Am. Tel. and Tel. Co.	2440	1932		Note 3
205	Maracay—Miami	Compania Annonima Nacional Telefonos de Venezuela—Am. Tel. and Tel. Co.	2190	1932		Note 3

APPENDIX I (*Continued*)

Circuit Group Index No.	Circuit Designation	Ownership	Distance Kilometers	Service Date[4]	Rates[2]	Time Sharing Arrangements
	EUROPE—SOUTH AMERICA					
315	Madrid—Rio de Janeiro	Compagna Telephonica Nationale d'Espagna (I. T. and T. Co.)—Companhia Radio Internacional do Brasil (I. T. and T. Co.)	8140	1932	150 G. F.	At Madrid with 311, 313 At Rio de Janeiro with 202, 802
316	Brussels—Buenos Aires	Belgian Govt.—Transradio Internacional Compania Radiotelegrafica Argentina S. A.	11320	1932	150 G. F.	At Brussels—with 406 At Buenos Aires with 302, 305, 308, 312, 801
	EUROPE—AFRICA					
401	Casablanca—Paris	Moroccan Tel. & Tel. Admin.—French P. T. T.	1820			Not shared
403	Capetown—London	Overseas Com. Co. of So. Africa—British P. O.	9680	1932		At London—with 301, 302, 303
403A	Johannesburg—London	South African P. O.—British P. O.	9080			Note 3
404	Cairo—London	Marconi Radio Tel. of Egypt—British P. O.	3510	1932		At London—with 102 At Cairo—Note 3
405	(1) Algiers—Paris	French P. T. T.—French P. T. T.	1350	1932		Not shared
	(2) Algiers—Paris	French P. T. T.—French P. T. T.	1350	1932		Not shared
406	Brussels—Leopoldville	Belgian Govt.—Note 3	6000	1932		At Brussels—with 316
407	Berlin—Cairo	German P. O.—Marconi Radio Tel. of Egypt	2910			Note 3
408	Paris—Tananarive	French P. T. T.—French P. T. T.	8760			Note 3
409	Cairo—Paris	Marconi Radio Tel. of Egypt—French P. T. T.	3210			Note 3

APPENDIX I (*Continued*)

Circuit Group Index No.	Circuit Designation	Ownership	Distance Kilometers	Service Date[4]	Rates[2]	Time Sharing Arrangements
	EUROPE—ASIA—OCEANIA					
502	(3) Amsterdam—Bandoeng..........	Neth. Govt.—Neth. Indies Tel. Admin.	11740		33 Florins	Note 3
506	Bombay—London......	Indian Radio Teleg. Co.—British P. O.	7300			At London—with 101
507	London—Singapore.....	British P. O.—Imp. & Int. Comm. Ltd.	12390			Note 3
508	London—Tokyo.......	British P. O.—Japanese Govt.	9560			Note 3
509	Hong Kong—London...	Hong Kong Tel. Co.—British P. O.	9630			Note 3
	NORTH AMERICA—ASIA—OCEANIA					
602	Manila—San Francisco.	Philippine L. D. Tel. Co. (Assoc. T. & T. Co.)—Am. Tel. and Tel. Co.	11220			Note 3
603	San Francisco—Tokyo..	Am. Tel. and Tel. Co.—Japanese Govt.	8260			Note 3
604	San Francisco—Sydney.	Am. Tel. and Tel. Co.—Australian P. O.	11950			Note 3
	NORTH AMERICA					
702	Miami—Tegucigalpa (Honduras).........	Am. Tel. and Tel. Co.—Tropical Radio Teleg. Co.	1480	1932		At Miami with 703, 704, 705
703	Miami—Managua (Nicaragua).........	Am. Tel. and Tel. Co.—Tropical Radio Teleg. Co.	1610	1932		At Miami with 702, 704, 705

APPENDIX I (*Continued*)

Circuit Group Index No.	Circuit Designation	Ownership	Distance Kilometers	Service Date[4]	Rates[2]	Time Sharing Arrangements
	NORTH AMERICA (*cont.*)					
704	Miami—San Jose (Costa Rica)........	Am. Tel. and Tel. Co.—Tropical Radio Teleg. Co.	1800	1932		At Miami with 702, 703, 705
705	Miami—Panama.......	Am. Tel. and Tel. Co.—Tropical Radio Teleg. Co.	1870	1932		At Miami with 702, 703, 704
706	Juneau—San Francisco .	Note 3—Am. Tel. and Tel. Co.	2530			
	SOUTH AMERICA					
804	Lima—Santiago.......	Compania Peruana de Telefonos, Limitada (I. T. & T. Co.)—Compania Internacional de Radio, S. A. (I. T. & T. Co.)	2470			At Lima—Note 3 At Santiago with 803, 313
	ASIA—OCEANIA					
905	Bangkok—Calcutta.....	Siamese Govt.—Indian Radio Co.	1620			Note 3

Notes

1. The table includes only circuits over 1000 km. in length. All circuits are short wave radio circuits except as noted in the New York—London group.
2. Rates given are for an initial 3 minute period. Abbreviations used are as follows: G. F. = Gold Francs, F. = French Francs, M. = Marks.
3. Definite information is not available.
4. Probable service date.
5. There are five transmitters at Bandoeng which serve the six circuits shown in the table on a shared basis.

APPENDIX II

SELECTED BIBLIOGRAPHY

The following bibliography does not purport to be complete but is selected from among the articles bearing on the subject of "World-Wide Telephony" which have come to the attention of the authors.

ANGWIN, A. S. Chairman's Address, Wireless Section, *Journal of Institution of Electrical Engineers*, December 1931, p. 17–35.

BAILEY, AUSTIN, DEAN, S. W. AND WINTRINGHAM, W. T. Receiving System for Long-wave Transatlantic Radio Telephony. Institute of Radio Engineers, *Proceedings*, Vol. 16, No. 12, December 1928, p. 1645–1705.

BEER, C. A. AND EVANS, G. T. Transatlantic Telephony; the Post Office Differential Voice-Operated Anti-Singing Equipment. *Post Office Electrical Engineers Journal*, Vol. 20, April 1927, p. 65–72.

BLACKWELL, O. B. Transatlantic Telephony—the Technical Problem. American Institute of Electrical Engineers, *Journal*, Vol. 47, No. 5, May 1928, p. 369–373. *Bell System Technical Journal*, Vol. 7, April 1928, p. 161–167.

BLACKWELL, O. B. The Time Factor in Telephone Transmission. *Bell System Technical Journal*, Vol. 11, No. 1, January 1932, p. 53–66. American Institute of Electrical Engineers, *Transactions*, Vol. 51, No. 1, March 1932.

BUTTNER, H. H. Role of Radio in the Growth of International Communication. *Electrical Communication*, Vol. 9, No. 4, April 1931, p. 249–254. Institute of Radio Engineers, *Proceedings*, Vol. 19, No. 1, January 1931, p. 51–58.

CRAEMER, P. Weltfernsprechnetz in seinen geographischen Bedingtheiten. World Engineering Congress. *Proceedings*, Tokio, 1929, Vol. 19, p. 497–505.

CRAEMER, P. Weltfernsprechverkehr. *Elektrotechnische Zeitschrift*, Vol. 50, No. 27, July 4, 1929, p. 959–963.

CRAEMER, P., AND FRANKE, A. Länderkarten des Europäischen Fernsprechnetzes. Sonderheft des "Europäischer Fernsprechdienst." May 1931.

ERIKSON, P. E. Service Telephonique International. *Journal Telegraphique*, Vol. 55, No. 12, December 1931, p. 381–385.

ESPENSCHIED, LLOYD, ANDERSON, C. N., AND BAILEY, AUSTIN. Transatlantic Radio Telephone Transmission. Institute of Radio Engineers, *Proceedings*, Vol. 14, No. 1, February 1926, p. 7–56. *Bell System Technical Journal*, Vol. 4, No. 3, July 1925, p. 459–507.

ESPENSCHIED, LLOYD AND WILSON, W. Overseas Radio Extensions to Wire Telephone Networks. Institute of Radio Engineers, *Proceedings*, Vol. 19, No. 2, February 1931, p. 282–302. *Bell System Technical Journal*, Vol. 10, No. 2, April 1931, p. 243–264.

GHERARDI, BANCROFT AND JEWETT, F. B. Telephone Communication System of the United States. World Engineering Congress Proceedings, Tokio, 1929, Vol. 19, p. 413–496. *Bell System Technical Journal*, Vol. 9, No. 1, Jan. 1930, p. 1–100.

HEISING, R. A. Production of Single Sideband for Transatlantic Radio Telephony. Institute of Radio Engineers, *Proceedings*, Vol. 13, No. 3, June 1925, p. 291–312.

LEE, A. G. Radio Communication Services of the British Post Office. Institute of Radio Engineers, *Proceedings*, Vol. 18, No. 10, October 1930, p. 1690–1731.

LEE, A. G. Transatlantic Telephony. *Telegraph and Telephone Journal*, Vol. 12, April 1926, p. 150–151; vol. 13, February 1927, p. 92–93. *Post Office Electrical Engineers Journal*, Vol. 19, April 1926, p. 74–75.

LITZINGER, L. Möglichkeiten eines transkontinentalen Fernsprechweitverkehrs zwischen Europa und Asien. Europäischer Fernsprechdienst, No. 5, July 1927, p. 3–14.

MÜLLER, E. Auf dem Wege zum Oceanfernsprechkabel. Europäischer Fernsprechdienst, No. 20, November 1930, p. 379–385.

OSWALD, A. A. AND SCHELLENG, J. C. Power Amplifiers in Transatlantic Radio Telephony. Institute of Radio Engineers, *Proceedings*, Vol. 13, No. 3, June 1925, p. 313–361.

PILLIOD, J. J. Intercontinental Radiotelephone Service from the United States. American Institute of Electrical Engineers, *Transactions*, Vol. 51, No. 1, March 1932.

SHAUGHNESSY, E. H. Rugby Radio Station. *Post Office Electrical Engineers Journal*, Vol. 19, January 1927, p. 373–382. *Electrical Review*, Vol. 98, April 30, May 7, 1926, p. 716–719, 753–756. *Electrician*, Vol. 96, April 23, 1926, p. 468–469.

TOWNSHEND, H. International Telephony. *Telephone and Telegraph Journal*, Vol. 15, February 1929, p. 87–89; March 1929, p. 106–108; April 1929, p. 128–129; May 1929, p. 148–150.

Transoceanic Telephone Service; a symposium. Contents: Miller, T. G. General Aspects. Bown, Ralph. Short-Wave Transmission.* Oswald, A. A. Short-Wave Equipment.* Cowan, F. A. Short-Wave Stations. American Institute of Electrical Engineers, *Transactions*, Vol. 49, No. 2, April 1930, p. 621–643. **Bell System Technical Journal*, Vol. 9, April 1930.

WAGNER, K. W. Plan einer Fernsprechkabelverbindung zwischen Europa und Amerika. Reprint from Sitzungsberichten der Preussischen Akademie der Wissenschaften-Physikalisch-Mathematischen Klasse, Vol. 7, 1929.

WAGNER, K. W. Transozeanische Fernsprechkabel, *Elektrische Nachrichten Technik*, Vol. 6, No. 4, April 1929, p. 125–138.

WATERSON, K. W. Transatlantic Telephone Service—Service and Operating Features. American Institute of Electrical Engineers, *Journal*, Vol. 47, April 1928, p. 270–273. *Bell System Technical Journal*, Vol. 7, No. 2, April 1928, p. 187–194.

WILSON, W. AND ESPENSCHIED, LLOYD. Radio Telephone Service to Ships at Sea. *Bell System Technical Journal*, Vol. 9, No. 3, July 1930, p. 407–428.

WRIGHT, S. B. AND SILENT, H. C. The New York-London Telephone Circuit. *Bell System Technical Journal*, Vol. 6, No. 4, October 1927, p. 736–749.

Long Distance Telephone Circuits in Cable *

By A. B. CLARK and H. S. OSBORNE

This paper first very briefly reviews the history of long distance telephone cables in the United States. A statement is then given of the basis of the electrical design of present day cables, followed by a discussion of the standards applied to cable circuits and the application of cables to the telephone needs of the country. While the present system is satisfactory for the circuits now used in cable up to distances of 1800 miles (3000 kilometers) or more, it would not be satisfactory for the much greater distances expected for the future, both for continental and intercontinental service. The paper closes with a brief account of the progress which has been made in the development of a cable carrier telephone system which is expected to be satisfactory for any distances which may in the future be spanned by telephone circuits in cable.

The authors wish to acknowledge their appreciation of the assistance given them by a number of their associates in the American Telephone and Telegraph Company, particularly Messrs. L. G. Abraham and R. M. Goetchius.

DURING the last two decades there has been, in all countries which have a large telephone development, a remarkable increase in the use of long distance telephone circuits in cable resulting in the building up of such large networks of long distance telephone cables as those which today cover the continent of Europe and a part of the continent of North America. This paper discusses the technical problems encountered in this development and the solutions applied in the development of the long distance cable network of the United States of America, using this as an illustration because it is the telephone plant with which the authors are most familiar.

IMPORTANT CABLE DEVELOPMENT MILESTONES IN THE UNITED STATES

The early long distance telephone circuits in the United States were practically all open wire. In fact, great care was exercised in laying out open-wire circuits to eliminate the necessity for using even short stretches of cable. Cable began to be considered seriously for long distance service when loading became available. In 1902 the first commercially loaded cable circuits in the United States were installed between New York City and Newark, N. J., a distance of about 11 miles (17 kilometers). Other cables rapidly followed, until in 1906 loaded cables were installed between New York and New

---

* Presented at The International Electrical Congress, Paris, France, July 5–12, 1932.

Haven, a distance of 79 miles (127 kilometers), and between New York and Philadelphia, a distance of 87 miles (140 kilometers). These 1906 cables contained No. 14 A. W. G. conductors (1.6 millimeters in diameter) and were loaded with 250-millihenry inductance coils spaced about 6000 feet (1830 meters) apart. By means of these cables it was possible to obtain transmission equivalents low enough so that these circuits furnished what was then considered very good long distance service. These early cables consisted entirely of physical circuits.

In order to obtain more circuits from the same number of wires, much development effort was then spent on methods for phantoming cable circuits and loading the phantoms. A number of problems were encountered and solved, the most serious being the problem of avoiding undue crosstalk between the circuits due to fortuitous unbalances introduced at the loading points and between the cable conductors. The problem was finally solved and in 1910 a short cable was installed between Boston and Neponset, Mass., providing loaded phantom circuits as well as side circuits, thus increasing the number of circuits 50 per cent. This was followed by a rapid extension of the use of this type of cable.

The ultimate in large gauge loaded cables was achieved in 1914 when the installation was completed of underground cable from Boston to Washington, a distance of 450 miles (724 kilometers), New York City being about at the midpoint. Some of the conductors in the cables were No. 10 A. W. G. (2.6 millimeters in diameter) while others were No. 13 A. W. G. (1.8 millimeters in diameter). Most of the loading consisted of 200-millihenry coils on the sides and 135-millihenry coils on the phantoms, spaced 7400 feet (2255 meters) apart. The 1000-cycle losses per mile (per kilometer) on these circuits were .050 (.031) db for the sides and .040 (.025) db for the phantoms for No. 10 A. W. G. and .085 (.053) db and .070 (.043) db for the side and phantom, respectively, of No. 13 A. W. G.

The vacuum tube telephone repeater was demonstrated as a great success when the New York-San Francisco telephone line was officially opened January 25, 1915. When this device was applied to the then available loaded cable circuits various imperfections, unimportant on non-repeatered circuits, produced serious effects, some of which had already been encountered in the work leading up to the loading of the open wire transcontinental line. Among these were the impedance characteristics of the cable circuits which were irregular due in part to insufficient stability and uniformity in the capacitances of the individual loading sections. These impedance irregularities prevented good repeater balances being obtained and consequently restricted the amplification which could be utilized on two-wire circuits.

It was soon realized that even if the loading were made very uniform two-wire repeatered circuits would be restricted in their transmission ranges, partly because of the unbalances encountered at repeaters and partly because of the tendency of the circuits to crosstalk into each other. Experiments were therefore begun utilizing the four-wire circuit method. When four-wire circuits were first set up using the large gauge loaded circuits between Boston and Washington great difficulty was experienced in obtaining even reasonably uniform attenuation at different frequencies. It then became apparent that smoother (more uniform) circuits would be necessary. It also became evident that higher cutoff frequency and higher velocity loading was necessary, in order to widen the effective transmission band and reduce delay distortion and echo effects. As a first step in this direction a system of loading in which inductance coils of 175 millihenries on the side circuits, spaced 6000 feet (1830 meters) apart, was introduced, primarily for two-wire circuits. This was commonly referred to as medium-heavy loading, having a cutoff frequency of approximately 2800 cycles. Experiments on long four-wire circuits with this loading which were specially set up for test confirmed the previous ideas as to the seriousness of echo effects and delay distorton effects and made it apparent that a much lighter weight and higher cutoff loading would be necessary for great distances. Accordingly, a system of loading known as H-44-25 (6000 feet—1830 meters—spacing with 44-millihenry coils on the sides and 25-millihenry coils on the phantoms) was introduced. This was the beginning of modern long distance cable circuits in America.

ELECTRICAL DESIGN OF TOLL CABLE CIRCUITS

It is impracticable in a short paper such as this to deal fully with all of the considerations which determine the electrical design of cable circuits. There are set down here, however, the important characteristics of the types of circuit which have been adopted for use in the United States and the reasons why certain arrangements were selected. Many specific designs which will meet the transmission objectives are, of course, possible. The designs selected have been based upon the aim of obtaining the desired results in the most economical manner, including the advantage which comes from concentrating on a small number of types of circuits, rather than on a larger number designed to meet accurately the requirements of different types of situations. In other countries where the ratios of costs for different parts of cable systems are different it is to be expected that the most economical designs will differ.

Cable Constants

The toll cable which is standard in the Bell System plant has capacitance of .062 microfarad per mile (.038 mf. per km.) for the side circuits and about .100 microfarad per mile (.062 mf. per km.) for the phantoms. There appears to be little to gain for voice-frequency circuits by varying materially from these capacitance figures.

With respect to size of wire, No. 19 A. W. G. conductors (.9-millimeter diameter) are well suited for both two-wire and four-wire circuit operation. No. 16 A. W. G. conductors (1.3-millimeter diameter) have been employed to a considerable extent in the past. In new cables conductors of this gauge are, in general, used only for relatively short non-repeatered circuits, or for program circuits. At the present time the possible economic advantage of using smaller sizes of wire than No. 19 A. W. G. is so small that it is considered to be outweighed, in general, by the greater complexity and variability of the circuits which would result.

Side Circuits and Phantoms

With the exception of program transmission circuits, multiple twin quads, utilizing both the side and phantom circuits, are used exclusively.

The ratio between the capacitances of phantoms and side circuits is about 1.6, while the ratio of resistances is 0.5. Because of this the phantoms which are loaded for the same cutoff frequency as the side circuits have lower attenuation and lower impedance than the side circuits. For repeatered circuits the phantoms and sides are so operated that they give substantially equal transmission results, this being desirable for flexibility reasons.

Two-Wire and Four-Wire Circuits

As is well known, two-wire repeatered circuits are more economical for the shorter distances while four-wire circuits are necessary to economically provide satisfactory transmission for longer distances. It is possible for terminating business to design two-wire cable circuits which will deliver good telephone service for distances of at least 900 miles (about 1500 kilometers). However, to meet the transmission standards current in the United States four-wire circuits are generally more practicable for distances more than two or three hundred miles (a few hundred kilometers).

Inductance and Spacing of Loading Coils

Theoretically the inductance and spacing of loading coils might be varied for each circuit length in order to obtain the most economical design. Studies which have been made, however, indicate that, taking into account the advantage of flexibility, it is desirable to use only two types of spacings of loading coils and only two general types of loading units. In the Bell System the toll loading coil spacings are 3000 feet (915 meters) and 6000 feet (1830 meters), loading with these spacings being designated B and H, respectively. For two-wire circuits, 88-millihenry loading coils for the sides and 50-millihenry coils for the phantoms are used with both spacings, giving loadings designated as B-88-50 and H-88-50. The choice between these is dictated by the repeater spacing. If less than about 45 miles (72 kilometers), H-88-50 loading is used; if greater, B-88-50 is used. Two-wire H-172-63 loading was used in the past for two-wire circuits but this has now been given up for new work in favor of the wider frequency band B and H-88-50 systems.

For four-wire circuits, as stated above, the standard loading is H-44-25 meaning, of course, 6000-foot (1830-meter) spacing of coils with 44-millihenry coils on the sides and 25-millihenry coils on the phantoms.

Important Transmission Characteristics of Loaded Cable Conductors

The characteristics of loaded cable circuits depend principally upon the electrical constants of the cable conductors, the inductance of the loading coils and their spacing. Some of the more important transmission characteristics of the loaded cable systems employing cables and loading coils of the type just described are given in Table I.

Spacing of Repeaters and Automatic Transmission Regulators

Repeaters are spaced as close to 50 miles (80 kilometers) apart as practicable. In the past variations upward from this to about 60 miles (100 kilometers) were allowed but it is now believed best to avoid such long spacings. Where the location of cities or other geographical situations make it desirable, spacings less than 50 miles (80 kilometers) are used.

Automatic transmission regulators are provided for circuits in aerial cables in excess of 50 to 100 miles (80 to 160 kilometers) in length and are preferably placed at every second repeater station. The devices are such, however, that satisfactory results may be obtained if regulators are as far apart as about 150 miles (250 kilometers), while under certain conditions circuit flexibility considerations call for regulators at adjacent repeater stations. With cable entirely underground transmission regulators will function satisfactorily from the electrical standpoint spaced as far apart as 300 miles (500 kilometers) although circuit flexibility considerations usually call for breaking circuits up into shorter regulator sections. In general, regulators are provided for all circuits in underground cable in excess of approximately 180 miles (300 kilometers) in length.

TABLE I

Characteristics of 19-Gauge Repeatered Toll Cable Circuits

	Two Wire				Four Wire H-44-25	
	B-88-50		H-88-50			
	Side	Phantom	Side	Phantom	Side	Phantom
Characteristic Impedance—Ohms at 1000 Cycles	1,560	930	1,120	670	800	450
Attenuation at 1000 Cycles						
db Per Kilometer at 55° F.	.17	.15	.22	.19	.30	.25
db Per Mile at 55° F.	.28	.23	.35	.30	.47	.39
Nominal Velocity—						
Kilometers Per Second	16,000	17,000	23,000	24,000	31,500	33,000
Miles Per Second	10,000	10,500	14,300	15,000	19,000	20,000
Cutoff Frequency of Loading—Cycles Per Second	5,600	5,900	4,000	4,200	5,600	5,900
Attenuation Change at 1000 Cycles—						
db Per Kilometer						
Aerial (55°–109° F.)	.019	.016	.025	.022	.034	.029
Underground (55°–73° F.)	.007	.006	.009	.007	.011	.010
db Per Mile						
Aerial (55°–109° F.)	.031	.026	.041	.035	.055	.046
Underground (55°–73° F.)	.011	.009	.014	.012	.018	.015

Gains of Repeaters

In four-wire circuits the spacing of the repeaters and their gains depends largely on "one-way circuit" considerations. The maximum levels are fixed by the repeater and loading coil capacities to handle speech waves without distortion, the lower levels being fixed primarily by noise. The lower levels also depend somewhat on crosstalk considerations, particularly crosstalk between circuits transmitting in opposite directions but this crosstalk is usually not controlling. The upper level limit used is + 10 db while the lower is − 24 db, these being referred to the level at the transmitting end of the toll circuit as zero.

In two-wire circuits, the above "one-way circuit" considerations are unimportant in determining the levels, these being largely fixed by "two-way" considerations. The important "two-way" considerations are crosstalk on the one hand and proper control of circulating currents which might cause singing or serious echoes on the other hand. These considerations keep the repeater gains so low that the limit of repeater and loading coil capacity is not important nor is the lower limit at which noise would become serious.

The most important crosstalk consideration on two-wire circuits is "near-end" crosstalk. The arrangement of repeater gains which gives lowest near-end crosstalk is one in which the output levels of the repeaters are the same throughout the circuit until the receiving end is reached, at which point the repeater gain is reduced or loss inserted as necessary to give the desired overall net loss for the circuit.

This arrangement is not best from the standpoint of the circulating currents, however. From this standpoint the best setup is one in which the repeater output levels "taper" considerably from the sending end to the receiving end of the circuit.

In the Bell System plant a compromise is made between these two considerations which calls for layout of gains about as follows: At the transmitting end the transmitting repeater gain is made such that the outgoing level is + 3 db. As the transmission passes through other repeaters the upper level is allowed to drop about ½ db per repeater for average temperature conditions. Finally at the receiving end of the circuit gain or loss is introduced to give the required net loss for the circuit. Of course, the application of the above rules for laying out repeater gains results in giving individual repeaters different gains in the two directions.

Smoothness of Impedance

For two-wire circuits it is important that the cables have a "smooth" impedance-frequency characteristic. To attain this result, loading coil inductances are held within close manufacturing tolerances while the cable capacitance variations are also carefully controlled. In the field care is, of course, taken with the spacing of the loading coils. Following are some representative figures for side circuits of fractional deviation from the average values per loading section or per loading coil:

	H Spacing	B Spacing
Representative deviation of cable capacitance *	.013	.018
Representative deviation of loading coil spacing *	.005	.005
Representative deviation of loading coil inductance *	.007	.007
Total deviation	.016	.020

* Representative deviation is the square root of the mean square of the individual deviations.

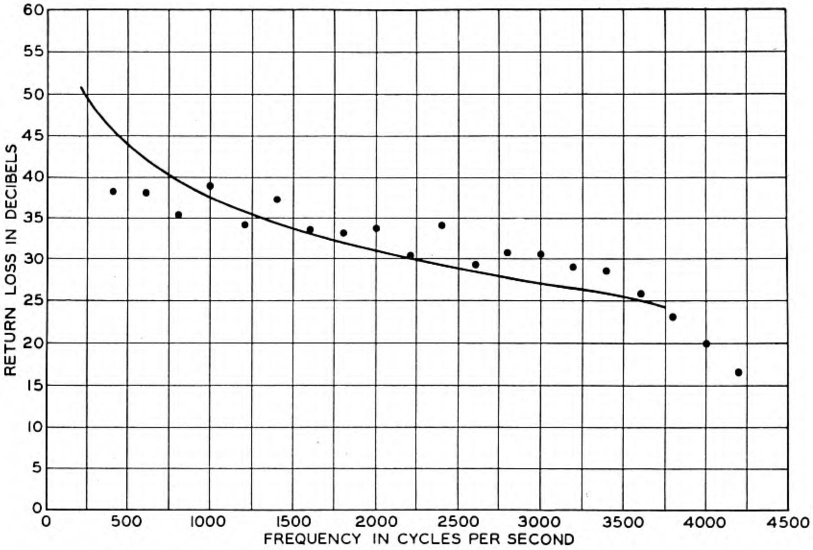


Fig. 1—Representative return losses on 19-ga. B-88-50 side circuits. 37 per cent of the measurements show lower return losses. ● 26 circuits Newark to Princeton.

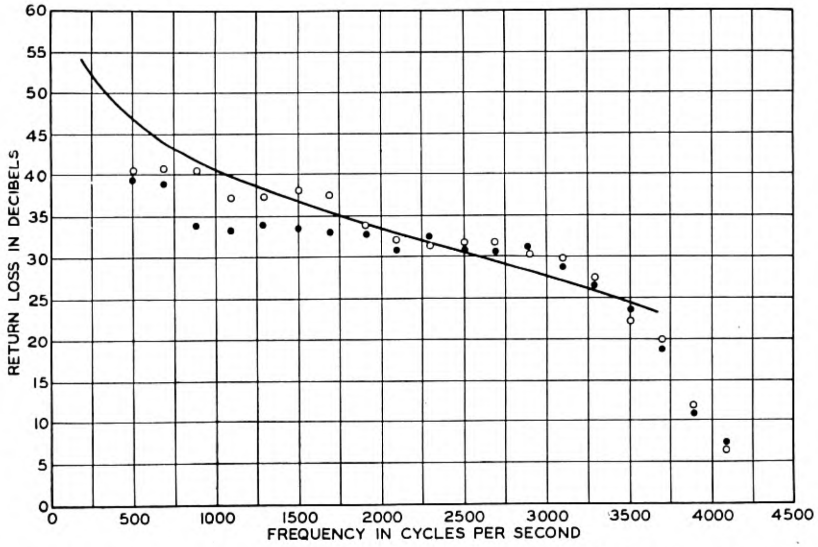


Fig. 2—Representative return losses on 19-ga. H-88-50 side circuits. 37 per cent of the measurements show lower return losses. o 26 circuits Princeton to Newark. ● 26 circuits Princeton to Philadelphia.

Representative return loss versus frequency curves are shown in Figs. 1 and 2 for B-88-50 and H-88-50 circuits, respectively, the points plotted being corresponding measurements on actual field facilities.

In the case of four-wire circuits, impedance irregularities are not so serious. However, for practical reasons, the same tolerances are usually followed for the several parts.

Control of Crosstalk

For two-wire circuits, the important crosstalk is near-end while for four-wire circuits it is far-end. For both of these, crosstalk between circuits within a single quad is greatest but crosstalk between circuits in different quads is also important. For two-wire circuits, in order to avoid long crosstalk exposures between any two circuits, it is the practice to carry three circuits together in a single quad only in a single repeater section, the circuits being systematically mixed at each repeater station. In the case of four-wire circuits this mixing is done only at the ends of regulator sections. In both outside cables and in the office cable, care is exercised to segregate the oppositely bound pairs of four-wire circuits because of the relatively large level differences.

In the outside cables control of crosstalk involves adjustments of the fortuitous unbalances in the loading units and unbalances between circuits in the loading sections. The following table shows the standards for phantom-to-side crosstalk expressed in decibels ordinarily worked to for the component parts of cable circuits:

	Two-Wire				Four-Wire	
	B-88		H-88		H-44-25	
	Avg.	Max.	Avg.	Max.	Avg.	Max.
Repeater section outside plant*	78		82		76	
Loading coil	96		96		96	
Cable proper per load section	93		94		94	
Toll office*	79		82		80	
Office wiring	80		86		82	
Office equipment	86		84		84	
Repeaters		91		91		74

* For two-wire circuits on the average these values must be decreased about 8 db to compare with overall value; for four-wire circuits about 9 db should be added to these values to compare with the overall.

Note: All values in table are for a frequency of 1250 cycles.

Control of Delay Distortion

In the case of two-wire circuits, since only relatively short distances are involved, delay distortion does not enter as a design factor. In the case of long four-wire circuits delay distortion is very important and to avoid this, very light weight and high cutoff loadings are used. Delay distortion may be reduced by employing correcting networks which introduce distortion counter to that introduced by the line. Such networks, in addition to their cost, have the disadvantage that they increase the total delay of the circuit. Their use is not necessary for the lengths of circuits and types of construction now used for message telephone circuits in the United States.

PERFORMANCE CHARACTERISTICS

Minimum Working Net Loss

In determining the numbers of circuits of different types of construction to be provided in a proposed cable for long distance work, it is necessary to take into account the limiting lengths for which the different types of circuit will meet the transmission requirements established for different conditions. This limitation for repeatered circuits is not set by attenuation but rather by singing margin, crosstalk or echoes. The lowest net loss at which a circuit equipped with repeaters may be operated without passing the limiting requirements for any of these characteristics is called the minimum net loss of the circuit. Since it is desirable to keep the net loss of a circuit at any time at least as great as the minimum net loss, an allowance for the probable circuit variations is added to the minimum net loss to obtain the minimum working net loss. This minimum working net loss is in general not directly proportional to the length of the circuit although in some cases within the important range of distances it can be considered to be proportional to a sufficient degree of accuracy.

It is evident that the minimum working net loss is a characteristic of fundamental importance in the design of new toll cables as well as in determining the operating limitations of existing cable circuits. Circuits having the same type of loading and the same spacing of repeaters may have widely different minimum net losses. This can come about because of differences in the accuracy with which crosstalk coupling is reduced in the manufacture and installation of the cable and its associated loading coils and apparatus, differences in the degree of uniformity of these characteristics and differences in the perfection of balance and matching of impedance between the cable conductors and the associated equipment. In the United States the study of the crosstalk, impedance and echo results obtained in various toll cables has led to the development of standard practices as to uniformity of construction and detailed specifications of the equipment with which the minimum net loss to be obtained with new toll cables can be closely predicted.

Performance Characteristics of Two-Wire Circuits

Fig. 3 shows, for B and H-88-50 two-wire circuits used exclusively for terminating business, the minimum working net losses for various distances as limited separately by crosstalk, echoes and singing. Of course, the lower the net loss the greater is the tendency toward excessive echoes, crosstalk or possibility of singing. For any given case the most exacting limitation controls the minimum working net loss.

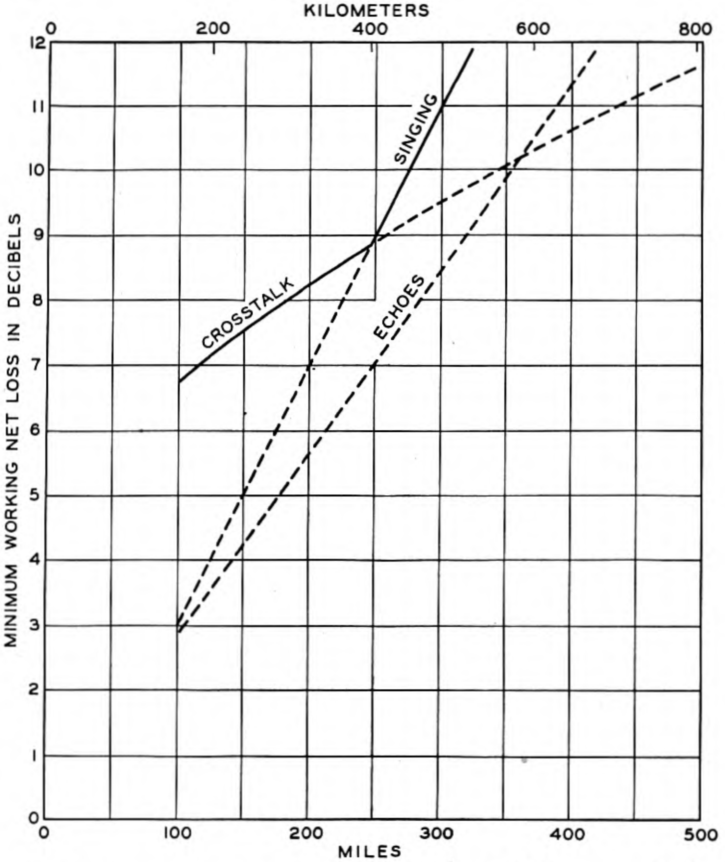


Fig. 3—Minimum working net loss of 19-ga. B and H-88-50 two-wire facilities versus circuit length for terminating business. These apply to either 50-mile (80 km.) sections of B-88-50 or 45-mile (72 km.) sections of H-88-50, whichever has the more severe limitations.

These curves include allowance for the unavoidable variations in the net loss which occur from time to time due to repeater battery variations, residual variations left over after pilot wire regulators have removed major transmission variations due to temperature changes in the cables, humidity effects in office wiring, etc.

Performance Characteristics of Four-Wire Circuits

Fig. 4 shows, for four-wire circuits used exclusively for terminating business, the minimum working net losses for various distances as limited separately by crosstalk and echoes. The possibility of singing does not enter as a limitation on these circuits. In these curves suitable allowance has also been made for the effect of the unavoidable transmission variations.

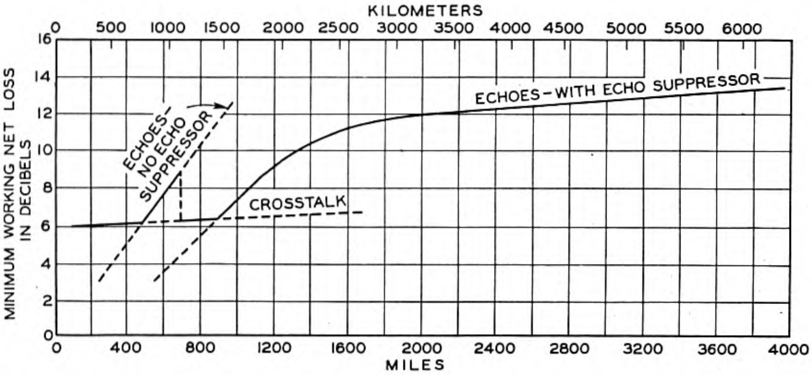


Fig. 4—Minimum working net loss of 19-ga. H-44-25 four-wire facilities versus circuit length for terminating business.

It will be observed that crosstalk is controlling only for the shorter distances up to a little over 400 miles (640 kilometers). Up to this length meeting proper crosstalk limits requires that the net loss be kept above about 6 db.

Echo constitutes the important limitation to the net loss of four-wire circuits. As is well known, echo suppressors go a long way toward eliminating echo effects but do not remove these effects, which remain the most important limiting factor on the longer circuits.

Transmission Requirements

The transmission requirements established for toll circuits are based on the provision of adequate transmission for the complete connection between any two points in the United States and the southern part of Canada, and also between any point of the country and the terminals of the intercontinental circuits. In order that this may be accomplished in an economical and orderly way a general toll switching plan has been adopted for the entire continental area. This plan establishes a fundamental basis for the routing of connections involving more than one toll circuit through the establishment of about 150 important switching centers to which all of the 2500 toll centers of the country will be directly connected. These 150 switching centers are interconnected by groups of high grade toll circuits either directly or for distant parts of the country through the intermediary of "regional centers" of which there are eight in the continental area.

By means of this plan it is possible to allocate each group of toll circuits to one of several broad classifications depending upon its position in the general toll switching plan, and to apply standard transmission requirements to each such broad classification. These requirements include the requirement that the effective net loss of all direct circuits shall not exceed 9 db, and that circuits designed for use in switched connections shall have minimum working net losses not exceeding 3 db for end links, 4 db for circuits between regional centers and 3.5 to 4 db for the remaining intermediate links interconnecting the important switching centers. When several circuits are connected together to form a long switched connection, the overall crosstalk effects are not appreciably increased over the effects of an individual circuit. Singing effects will usually not be limiting since long switched connections are seldom established without at least one intervening four-wire circuit. On the other hand echo effects increase fairly rapidly even with circuits equipped with echo suppressors and therefore, in selecting facilities which will meet the requirements of minimum working net loss previously specified, the echo effects are generally controlling on long connections. Since these circuits are also used for direct circuit connections, higher net losses, which will be satisfactory from the crosstalk and singing standpoints in the terminating condition, are obtained by adding pads at one or both terminals of the circuit for this condition.

A more complete statement of the transmission requirements applied to toll cable circuits in the United States is given in Table II. For convenience there are also given in this table the current transmission requirements for international circuits adopted by the C. C. I. In order to make the comparison as nearly as possible on a comparable basis the international circuit requirements of the C. C. I. are compared with the requirements for American toll circuits interconnecting two regional centers, and the requirements for the national terminal are compared with the American requirements for connections from the

TABLE II

TRANSMISSION STANDARDS

Bell System	*C. C. I.*
Number of Regional Centers—8	Number of European National Outlets—29

NET LOSS

Circuits Between Regional Centers (On effective transmission basis)	*International Circuits* (Based on 800-cycle transmission equivalent)

	Terminating Business	Switched Business (Via Net Loss)*
Future Plant	9.0 db	3.0 db
Existing Plant	9.0–11.0	3.0–5.0 db

Terminating or Switched Business	*Future Plant*	*Existing Plant*
Two-Wire	8.7 db	11.3 db
Four-Wire	6.9 db	9.6 db

MAXIMUM OVERALL NET LOSS BETWEEN SUBSCRIBERS†

Regional Center Transmitting Loss	16 db
Regional Center Receiving Loss	12 db
Circuit Between Regional Centers (Via Net Loss)	3–5 db
Total	31–33 db

National Transmitting Loss	17 db
National Receiving Loss	11 db
International Circuit	7–11 db
Total	35–39 db

FREQUENCY BAND WIDTH

250 to 2750 cycles Loss at extreme frequencies 10 db greater than at 1000 cycles. For narrower bands the effective transmission equivalent reflects the effect on transmission of band width.	300 to 2400 cycles Loss at extreme frequencies 8.7 db greater than at 800 cycles.

DELAY DISTORTION

Direct circuit connection (terminal circuit)—20 milliseconds. Via circuits (in switched connections) 10 milliseconds. (Differences between 1000 cycle delay and the highest frequency effectively transmitted.)	*Overall Connection* 30 milliseconds as the difference between the time of propagation for the highest frequency effectively transmitted and the time of propagation for 800 cycles.

NON-LINEAR DISTORTION

No fixed limits at present on message circuits but consideration is being given to methods of measuring and evaluating the impairment due to this effect.	No fixed limits at present on message circuits but consideration is being given to methods of measuring and evaluating the impairment due to this effect.

* Via net loss is the loss which the circuit contributes to an overall connection when switched to toll lines at both ends.

† The regional center and national outlet transmitting and receiving losses are the maximum losses between the most distant subscriber in the particular area served by the regional center or national outlet and the particular regional center or national outlet.

TABLE II—(*Continued*)

Bell System	*C. C. I.*
TRANSMISSION VARIATIONS (AVERAGE OF THE TWO DIRECTIONS)	
250-mile (400-kilometer) circuits ± 1.5 db, 1000-mile (1600-kilometer) circuits ± 3.0 db (these limits are exceeded approximately 5 per cent of the time).	No limits.
ECHOES	
Echoes are limiting in accordance with curves representing the results of experience as to permissible loudness of echo without interference for different times of propagation. See page 129 of the "Red Book" of the C. C. I. On four-wire circuits echo suppressors are employed for all circuits used for switched business in excess of 270 miles (430 kilometers) of H-44-25 facilities. Four-wire circuits used for terminal business only are equipped with echo suppressors when they exceed approximately 650 miles (1000 kilometers) of H-44-25 facilities. Echoes on circuits so equipped are limiting as indicated in Fig. 4 of this paper.	Echoes are limiting in accordance with curves representing the results of experience as to permissible loudness of echo without interference for different times of propagation. See page 129 of the "Red Book" of the C. C. I.
SINGING MARGIN (TERMINAL CONDITION)	
10 db for two and four-wire circuits.	Two-wire 6.8 db. Four-wire 9.0 db.
CROSSTALK	
For quiet circuits: Average—70 db. Maximum—60 db. (Based on 1 per cent chance of exceeding 60 db.)	54 db at least.
MINIMUM WORKING NET LOSS*	
4 db.	No limits established.
TIME OF PROPAGATION	
For continental communication—250 milliseconds. Continental circuits in an intercontinental connection—100 milliseconds. For continental communication: Delay between echo suppressors—100 milliseconds.	For continental communication—250 milliseconds. Continental circuits in an intercontinental connection—100 milliseconds. For continental communication: Delay between echo suppressors—100 milliseconds.
NOISE (INCLUDING BABBLE) TERMINAL CONDITION MEASURED AT THE TOLL SWITCHBOARD	
+ 26 db (200 noise units) referred to reference noise (approximately 10 noise units). For noise values greater than above limit N. T. I.'s are applied.	*2 Millivolts (Approximately 160 Noise Units)* (Equivalent intensity of 800-cycle tone of the limiting voltage across a receiver whose impedance is adjusted to 600 ohms.)

* For more detailed discussion refer to text.

subscriber to a regional center. As there are eight regional centers in the United States and 29 national networks which constitute the European continental network, the areas involved in the comparison are not wholly comparable.

The application of these transmission requirements leads to the following approximate limits for the use of the different standard types of toll cable construction in the toll circuits of the United States:

19-gauge B and H-88-50 two-wire circuits	250 miles (400 kilometers) when used exclusively for terminating business
19-gauge B and H-88-50 two-wire circuits	135 miles (220 kilometers) for circuits used for switched business
19-gauge H-44-25 four-wire circuits without echo suppressors	650 miles (1050 kilometers) when used exclusively for terminating business
19-gauge H-44-25 four-wire circuits without echo suppressors	270 miles (430 kilometers) for circuits used for switched business

Use of Toll Cables in the United States

In the application of toll cables to meet the service requirements of the United States use is generally made of cables 2⅝ inches (67 millimeters) in outside diameter, although to some extent use has been made of cables 3⅛ inches (79 millimeters) in outside diameter. A large proportion of the new toll cables contain one of the combinations of gauges indicated in the following table:

TABLE III

	16-Gauge Quads	16-Gauge Pairs	19-Gauge Quads	22-Gauge Quads
Cable 1......	0	6	148	1
Cable 2......	19	6	114	1

The present toll cable routes in the United States are indicated on the map shown in Fig. 5, together with probable future extensions. The existing network includes 13,000 miles (21,000 kilometers) of route, 21,000 miles (34,000 kilometers) of cable, and when fully equipped, 5,000,000 miles (8,000,000 kilometers) of circuit.

The division of the cable between various types of construction is indicated in the following table:

Underground in ducts................................	46%
Aerial..	50%
Buried underground..................................	4%

As an illustration of the practical application of toll cable, some information is given below regarding the makeup of some of the circuits out of New York City. There are associated with the New York toll board 235 groups of circuits. In order to avoid too cumber-

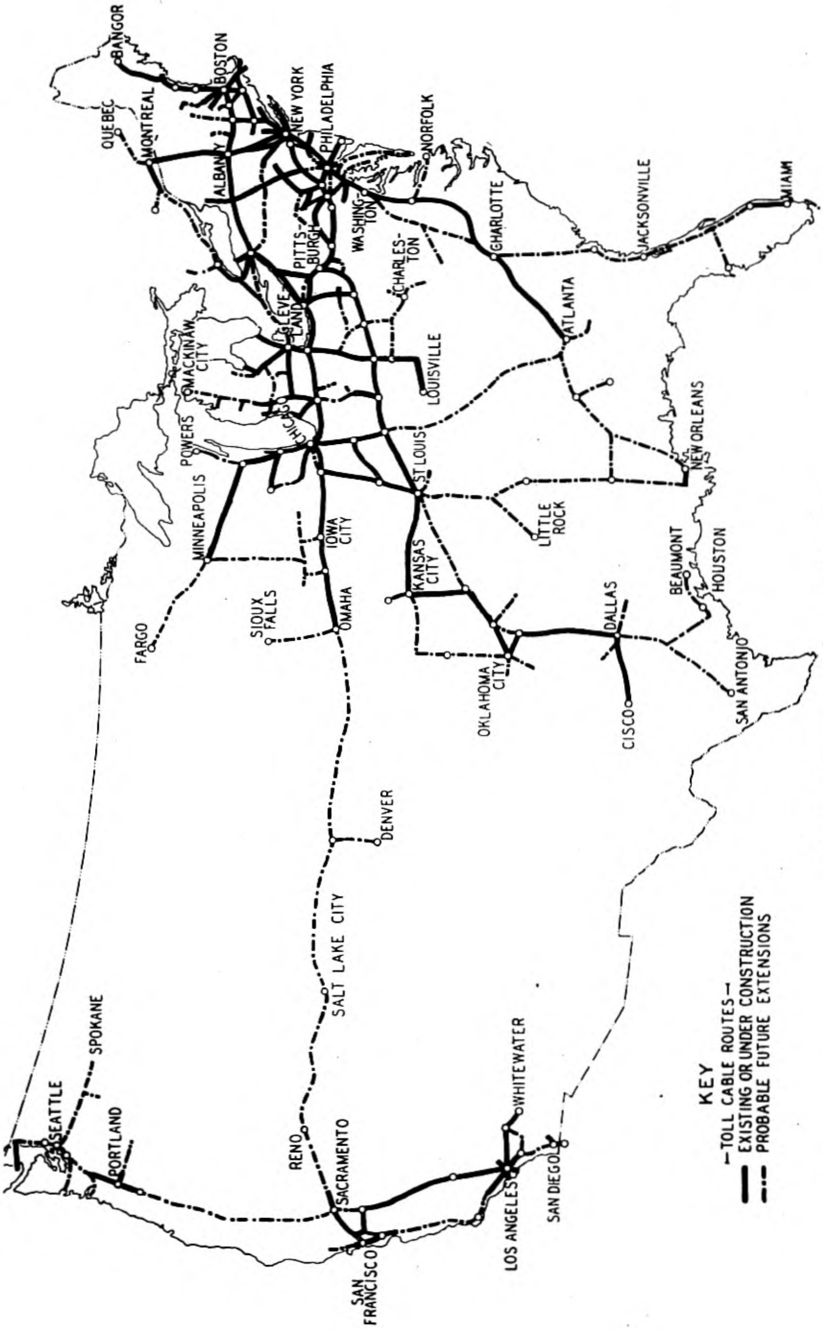


Fig. 5—Toll cable routes—Bell System.

some an illustration, we shall consider only circuits from New York to other regional centers and primary outlets. Chicago has a larger number of connections to these points but New York is chosen because of its interest in connection with international service. These circuits are indicated on the map in Fig. 6.

Of the 64 circuit groups to regional centers and primary outlets, 43 are less than 870 miles (1400 kilometers) in length, 95 per cent of the circuit mileage being cable. The makeup of a few of these groups chosen for the purpose of illustration is given in Table IV.

The remaining circuit groups indicated in Fig. 6, 870 miles (1400 kilometers) and more in length, and 21 in number, have the makeups and electrical characteristics indicated in Table V.

It will be noted that the average circuit in this classification is made up of about one half four-wire toll cable circuit and the other half of carrier telephone superimposed upon open wire. Since carrier circuits have, in general, electrical characteristics comparable to the four-wire cable circuits and have a relatively high velocity of propagation, the combination of four-wire with carrier results in very satisfactory electrical characteristics, even for the longest circuits.

A point of interest in connection with the table is the time of propagation. This is, in all cases, well within the provisional limit of 100 milliseconds adopted by the C. C. I. for the time of propagation of the continental terminating circuits of an intercontinental connection.

FUTURE REQUIREMENTS OF VERY LONG CABLE CIRCUITS

At the present time cable up to lengths of 1800 miles (3000 kilometers) is used in the regular routine in the United States and gives a satisfactory performance. Tests have been made on longer four-wire cable circuits up to lengths exceeding 3600 miles (6000 kilometers). These tests show that for such lengths, and particularly for the much greater lengths which may result from the development of intercontinental telephone service, the present design of toll cable circuits would not be entirely satisfactory.

More Effective Echo Suppressors

Circuits 3600 miles (6000 kilometers) long when equipped with ordinary echo suppressors fail to give the net loss of 9 db which has been set up as a design objective. A more effective type of echo suppressor is necessary to work such a circuit at as low a net loss as 9 db. Experiments have been made with an echo suppressor of a type which changes its sensitivity automatically, depending upon the

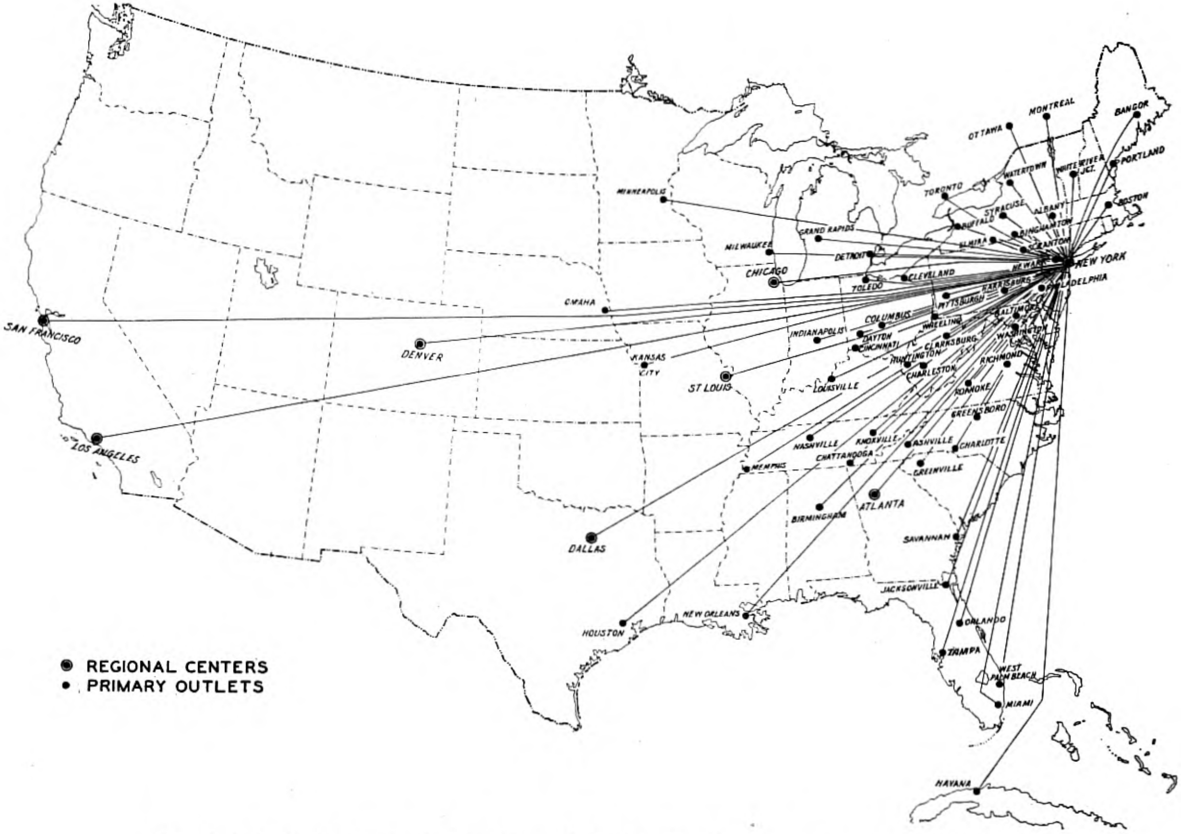


Fig. 6—Direct circuit groups from New York to regional centers and primary outlets.

TABLE IV

TYPICAL CIRCUITS FROM NEW YORK TO REGIONAL CENTERS AND PRIMARY OUTLETS LESS THAN 870 MILES (1400 KILOMETERS)—JANUARY 1932

City	Class*	Number of Circuits	Length		Make-Up	Effective Net Loss—db	
			Kilometers	Miles		Terminal Condition	Via Condition
Philadelphia	Via	24	140	87	Four Wire H-44-25	6	0
	Term	299			Two Wire Older Types of Loading	12†	—
Boston	Via	17	367	228	Four Wire H-44-25	9	3
	Term	85			Two Wire Older Types of Loading	12†	—
Washington	Via	7	359	223	Four Wire H-44-25	8	2
	Term	44			Four Wire H-44-25	8	—
Pittsburgh	Via	15	587	365	Four Wire H-44-25	8	2
	Term	24			Two Wire H-44-25	8	—
Baltimore	Via	6	299	186	Four Wire H-44-25	8	2
	Term	32			Four Wire H-44-25	8	—
Albany	Via	12	249	155	Two Wire H-44-25	8	2
	Term	23			Two Wire H-172-63	12†	—
Cleveland	Via	8	817	507	Four Wire H-44-25	9	3
	Term	18			Four Wire Older Type Loading	13†	—
Detroit	Via	9	1,090	677	Four Wire H-44-25	9	3
	Term	17			Four Wire Older Type Loading	12†	—
Toronto	Via	7	803	499	Four Wire H-44-25	9	3
	Term	9			Two Wire H-172-63	11†	—

* Via—Circuits used for switched business.
Term—Circuits used for terminal business only.

† With further growth of plant these circuits will be brought within present limits by reloading or by replacing with new circuits, limiting the older types of loading to shorter distances.

TABLE V

DIRECT CIRCUIT GROUPS FROM NEW YORK TO REGIONAL CENTERS AND PRIMARY OUTLETS EXCLUDING THOSE LESS THAN 870 MILES (1400 KILOMETERS)—JANUARY 1932

City	Number of Circuits	Circuit Make-Up								Effective Net Loss db		1000-Cycle Delay Milliseconds
		Kilometers				Miles						
		Four Wire H-44-25	Type C Carrier	Other	Total	Four Wire H-44-25	Type C Carrier	Other	Total	Terminal Condition	Via Condition	
Atlanta	9	1,010	606	—	1,616	623	380	—	1,003	9	3	36
Birmingham	3	1,010	843	—	1,853	623	527	—	1,150	9	3	37
Chicago	59 *	1,400	—	—	1,400	—	—	—	—	9	3	46
Dallas	4	1,605	1,084	—	2,689	998	673	—	1,671	10	4	58
Denver	3	1,685	1,465	—	3,150	1,048	910	—	1,958	10	4	61
Havana	3	359	2,300	—	2,659	221	1,430	—	1,651	10	4	22
Houston	3	1,605	1,210	—	2,815	998	751	—	1,749	10	4	58
Jacksonville	8	1,010	—	632	1,642	623	—	398	1,021	9	3	37
Kansas City	5	2,010	—	—	2,010	1,250	—	—	1,250	10	4	67
Los Angeles	7	1,485	4,110	—	5,595	922	2,554	—	3,476	11	5	66
Memphis	3	1,100	863	—	1,960	684	534	—	1,218	9	3	40
Miami	19	1,010	1,215	—	2,225	622	759	—	1,382	9	3	38
Minneapolis	5	1,685	647	—	2,332	1,048	401	—	1,449	9	3	39
New Orleans	7	1,010	1,425	—	2,435	623	890	—	1,513	9	3	39
Omaha	2	2,175	—	—	2,175	1,350	—	—	1,350	10	4	72
Orlando	3	1,010	—	900	1,910	623	—	559	1,182	10	4	39
St. Louis	13	1,605	—	—	1,605	998	—	—	998	9	3	53
San Francisco	6	1,830	3,430	—	5,260	1,137	2,133	—	3,270	11	5	75
Tampa	5	1,010	931	—	1,941	623	585	—	1,208	9	5	37
Tulsa	2	1,605	703	—	2,308	998	438	—	1,436	10	4	56
West Palm Beach	9	1,010	1,105	—	2,115	623	690	—	1,313	9	3	38

* 45 Circuits for Terminal Business Only.
Note: All of these circuits are equipped with echo suppressors.

amount of noise on the circuit at any given time, the less noise the more sensitive the device. It has been found possible to adjust this device to sufficient sensitivity to permit working a 3600-mile (6000-kilometer) circuit with 9 db net loss. When noise is added to the circuit the sensitivity of the device, of course, diminishes but the added echo is largely masked by the increased noise.

Delay Distortion

For a 3600-mile (6000-kilometer) H-44-25 circuit the difference in the delay at 1000 cycles and at 3000 cycles amounts to about 0.025 second. Experiments which have been made on circuits introducing very little non-linear distortion indicate that this amount of delay distortion by itself is not particularly serious. However, on the long four-wire circuits where non-linear distortion is also present, the effect of delay distortion becomes more pronounced so that it becomes quite objectionable. Delay distortion correctors would therefore be required for H-44-25 circuits of this length, although for circuits of 1800 miles (3000 kilometers) they do not appear necessary.

Time of Propagation

When a long connection is built up using cable circuits, the delay proper, quite apart from delay distortion, becomes important. For a 3600-mile (6000-kilometer) length of H-44-25 circuit equipped with delay equalizers the time of propagation is about one-quarter of a second in each direction. This time of propagation is generally considered about all that should be tolerated and is the C. C. I. tentative limit for a complete connection.

Adverse Interaction of Two Echo Suppressors

When voice-operated devices are introduced on very long cable circuits another complication results. Assume, for example, that the 3600-mile (6000-kilometer) connection is made up of two links, each equipped with an echo suppressor of the usual type, either mechanical relay or vacuum tube operated. Assume that the echo suppressors are 1800 miles (3000 kilometers) apart, the delay between these devices being one-eighth of a second for each direction of transmission. When conversations are carried on over this circuit it is found that occasionally when the speakers at the two ends utter words at nearly the same instant both echo suppressors respond, each echo suppressor blocking one direction of the circuit. Consequently, certain words or parts of words are lost. In telephone conversations it is found that with such a circuit arrangement if the time of propagation in each direction between echo suppressors does not exceed about 0.1 second, the amount lost apparently is not a serious handicap. The C. C. I. has tentatively recommended that 0.1 second be taken as a limit towards which it is desirable to work if practicable.

Better Cable Circuits Desirable for Extreme Distances

These difficulties with the present type of cable construction all become more pronounced if, instead of a 3600-mile (6000-kilometer) circuit, consideration is given, for example, to possible future intercontinental circuits in all cable construction. Such a circuit between San Francisco and Istanbul, for example, would be about 10,000 miles (16,000 kilometers). The time of propagation for such a connection would be about .6 of a second in each direction so that two-way conversations would be seriously impeded. Serious difficulties would also be experienced with the voice-operated devices and because of the accumulated distortions, including non-linear effects. While, therefore, it is possible that a circuit of this sort could be used for two-way telephony between Istanbul and San Francisco, the imperfections of such a circuit would be so outstanding as to warrant a serious effort to develop something better if this could be done at a reasonable cost.

Telephone Carrier in Cable

In order to obtain better transmission results over very long cable circuits in an economical manner the development of a carrier system for cables has been actively undertaken. The development work of the Bell System has now been carried to the point where it seems assured that it will be successful and that telephone carrier will have an important field of use in long distance cables on heavy routes.

This carrier system uses non-loaded cable conductors whose velocity is very high as compared to voice-frequency loaded circuits, the effective circuit velocity including delays introduced by apparatus being about 100,000 miles (160,000 kilometers) per second. An experimental trial system has been set up by looping circuits back and forth in a cable so as to produce the equivalent of the system shown in Fig. 7. It will, of course, be understood that this figure represents merely the experimental setup and should therefore not be considered the ultimate in such matters as carrier channels per pair, repeater spacing, etc.

Talking tests which were made using this experimental setup showed very satisfactory quality of transmission and no appreciable interference between circuits. In addition to testing each of the nine telephone circuits shown in the sketch, these nine circuits were con-

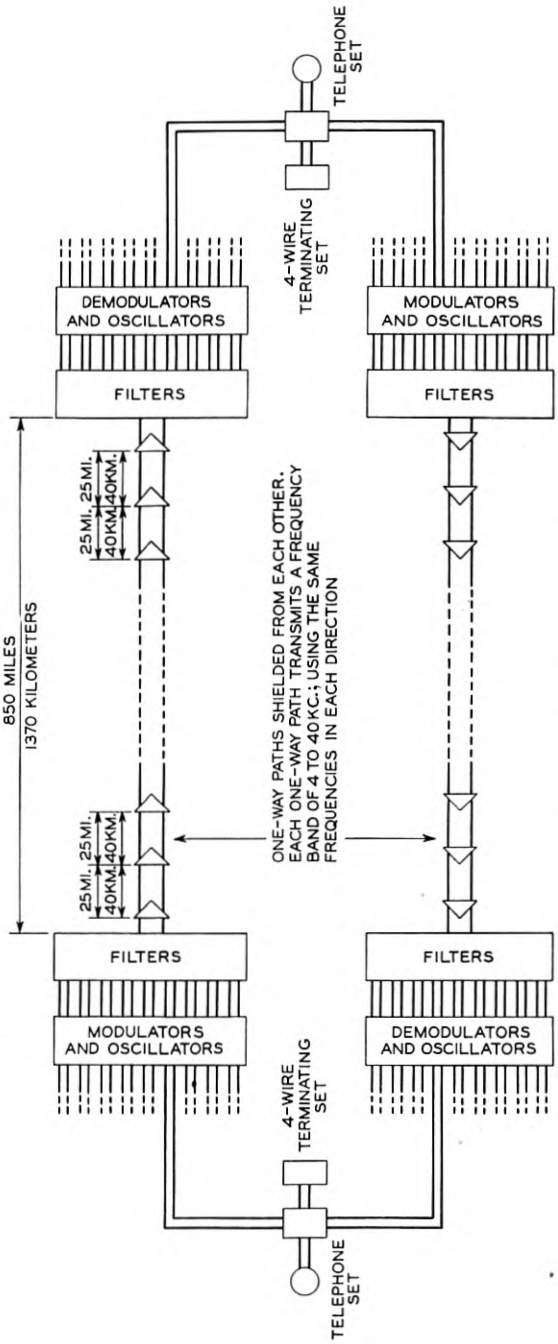


Fig. 7—Experimental cable carrier telephone circuit.

nected in tandem giving an overall length of 7500 miles (12,000 kilometers) of circuit. Conversations over this 7500-mile (12,000-kilometer) circuit were very satisfactory. In fact, the transmission quality was not greatly impaired even when a 15,000-mile (24,000-kilometer) length of one-way circuit was established by connecting all of the links in tandem.

While the development of this carrier system is far from completion and it is not clear at the present time how far it can be applied to other than heavy traffic routes, it is certain that wherever this form of construction is justified, distance no longer remains as a limiting factor.

In the fifty years since the first International Electrical Congress at Paris, the new art of telephone communication has passed through many stages of development and during the past thirty years a new art, making possible communication in cable over long distance, has been born and brought to maturity. This has been made possible by a number of important and fundamental developments such as loading, quadded cable and telephone repeaters. While the development has been rapid, particularly during the past twenty years, it is not too much to expect that the next twenty or thirty years will witness an even greater and more rapid technical development and expansion in the use of long distance toll cables in all parts of the world, associated with a continued increase in the service rendered to mankind by long distance telephone communication.

BIBLIOGRAPHY

BLACKWELL, O. B. The Time Factor in Telephone Transmission. *Bell System Technical Journal*, Vol. XI, January, 1932, pp. 53–66.

CARSON, J. R. Building Up of Sinusoidal Currents on Long Periodically Loaded Lines. *Bell System Technical Journal*, Vol. III, October, 1924, pp. 558–566.

CLARK, A. B. Telephone Transmission Over Long Cable Circuits. *Transactions of American Institute of Electrical Engineers*, Vol. 42, February, 1923, pp. 86–97. *Electrical Communication*, February, 1923, pp. 26–40. *Bell System Technical Journal*, Vol. II, January, 1923, pp. 67–94.

CLARK, A. B. AND C. W. GREEN. Long Distance Cable Circuit for Program Transmission. *Bell System Technical Journal*, Vol. IX, July, 1930, pp. 567–594.

CLARK, A. B. AND R. C. MATHES. Echo Suppressors for Long Telephone Circuits. *Transactions of American Institute of Electrical Engineers*, Vol. 44, April, 1925, pp. 481–490.

CRISSON, G. Irregularities in Loaded Telephone Circuits. *Bell System Technical Journal*, Vol. IV, October, 1925, pp. 561–585. *Electrical Communication*, Vol. 4, October, 1925, pp. 98–110.

DEMAREST, C. S. Telephone Equipment for Long Cable Circuits. *Transactions of American Institute of Electrical Engineers*, Vol. 42, June, 1923, pp. 742–752. *Bell System Technical Journal*, Vol. II, July, 1923, pp. 112–140.

LANE, C. E. Phase Distortion in Telephone Apparatus. *Bell System Technical Journal*, Vol. IX, July, 1930, pp. 493–521.

NYQUIST, H. AND S. BRAND. Measurement of Phase Distortion. *Bell System Technical Journal*, Vol. IX, July, 1930, pp. 522–549.

OSBORNE, H. S. A General Switching Plan for Telephone Toll Service. *Bell System Technical Journal*, Vol. IX, July, 1930, pp. 429–447.

SHACKLETON, W. J. AND I. G. BARBER. Compressed Powdered Permalloy—Manufacture and Magnetic Properties. *Transactions of American Institute of Electrical Engineers*, Vol. 47, No. 2, April, 1928, pp. 429–436.

SHAW, THOMAS AND WILLIAM FONDILLER. Development and Application of Loading for Telephone Circuits. *Bell System Technical Journal*, Vol. V, 1926 pp. 221–281. *Transactions of American Institute of Electrical Engineers*, Vol. 45, February, 1926, pp. 268–292.

STEINBERG, JOHN C. Effects of Phase Distortion on Telephone Quality. *Bell System Technical Journal*, Vol. IX, July, 1930, pp. 550–566.

ZOBEL, O. J. Distortion Correction in Electrical Circuits with Constant Resistance Recurrent Networks. *Bell System Technical Journal*, Vol. VII, July, 1928, pp. 438–534.

The Conception and Demonstration of Electron Waves *

By C. J. DAVISSON

An attempt is made in this article to trace the growth of our ideas regarding the electron from their inception less than a hundred years ago to the present day. The discussion begins with a consideration of the vague and tentative deductions concerning an ultimate electrical charge which became possible when Faraday revealed the laws of electrolytic conduction; it touches upon the clarification of the conception of the electron as a charged particle capable of independent existence and subject to the laws of classical electrodynamics which was effected at the close of the last century and the beginning of the present one by the researches of J. J. Thomson and others; it indicates the difficulties in which this conception became involved, and the attempts made by Planck, Bohr and others to extricate it from them. The latter part of the paper is devoted to the amplified conception of the electron which has been developed during the last decade—a conception in which electrons are recognized as having, in different circumstances, the properties of both waves and particles.

INTRODUCTION

IT is my purpose in this report to describe a few experiments, typical of several hundred now recorded, in which streams of electrons exhibit the properties of beams of waves. It seems desirable, however, to begin by briefly reviewing various steps in the development of our conception of the electron before about the year 1925. It is against this background only that the phenomena revealed by the experiments to be described appear in true relief.

The idea that electric charge is granular was not new at the time of the first International Electrical Congress in 1881. Faraday had determined and announced the laws of electrolytic conduction fifty years earlier, and it was recognized, by some at least, that these laws suggested the existence of an elementary charge or atom of electricity. An estimate of the magnitude of this natural and presumably ultimate unit of charge had indeed been made a few years prior to the Congress by the Irish physicist Stoney from such data as were then available. The word "electron" to designate the hypothetical atom of electricity was not, however, introduced until the year 1891. The concept of the electron gained rapidly in sharpness in the decade next following—the last of the century—not so much indeed from the introduction of new ideas concerning it as from experimental evidence in support of ideas already held.

It was no new idea, for example, that neutral atoms contain positive and negative charge in equal amounts, and that the ions of an electrolyte are merely atoms or groups of atoms in which these charges are

---

* Presented by title at The International Electrical Congress, Paris, France, July 5–12, 1932.

unbalanced in one direction or the other by one or more electronic units. Yet this remained largely a speculation until the study of the conductivity imparted to gases by X-rays made all other views untenable. Ions of both signs are formed at a uniform rate within the body of a gas subjected to this then newly discovered radiation; their charges are ionic; they move through the gas by diffusion and under the influence of an impressed electric field; they disappear through recombination.

Other ideas now familiar were not entirely novel even in 1890; for instance, the idea that positive and negative electrons possess mass as well as charge—that those of one sign are more massive than the other—that within the atom the lighter revolve about the heavier ones. Weber, following on Ampere, had pictured a mechanism of this kind to explain the magnetic properties of materials. All these notions became much more plausible, however, when Lorentz showed (as he did in 1897) that the splitting and polarization of spectral lines by a magnetic field might be explained as the effect of the field upon the period of revolving particles such as Weber had assumed, and that from the magnitude of this so-called "Zeeman effect" one might actually calculate the ratio of the charge of the particle to its mass. The value so found was greater by a factor 2000, or thereabouts, than the similar ratio for hydrogen ions in electrolysis. If the charge of the particle were indeed the electronic charge, then the mass of the particle was about 1/2000 only of the mass of the hydrogen atom—a highly acceptable conclusion. The concept of the electron had gained much in definiteness, and so also had that of the atom.

But more illuminating still was the discovery made in the same year that the trajectories followed by cathode rays in traversing electric and magnetic fields are exactly those to be expected if these rays are streams of swiftly moving negatively charged particles with a charge to mass ratio amounting again, as in the foregoing instance, to about 1/2000 that of the hydrogen ion. There could be little doubt that this was the very particle inferred by Lorentz from the "Zeeman effect." Cathode rays were certainly streams of free negative electrons—unattached to atoms. This supremely important discovery was made by J. J. Thomson in England and by Wiechert in Germany.

The conception of the negative electron as a subatomic particle possessing mass as well as charge, capable of independent existence, and subject to the laws of classical electrodynamics now seemed clearly established. If various of the simple relationships were at this time sensed rather than demonstrated—such, for example, as the exact identity of the charge to mass ratios of the Zeeman effect particle and the cathode particle—there was, nevertheless, full confidence that these relationships would be confirmed by more exact measurements. And this indeed proved to be true. The anticipated details of the picture as then blocked in have since been supplied by a series of precision experiments in which Millikan's measurement of the absolute magnitude of the electronic charge is preeminent.

The turn of the century was a time of high hope. The key had been found, it appeared, to an understanding of vast ranges of phenomena; given the electron, electrodynamics and sufficient mathematics, *all* electrical and magnetic phenomena must become explicable. It seemed not too daring even to have thoughts concerning the structure of the atom. But this, as it turned out, was mostly illusion; every success of the electron theory of this period was matched by an equally conspicuous failure. Metallic conductors were pictured as containing atmospheres of free electrons with the properties of a monatomic gas. The drift of this electronic gas under the influence of an impressed field constituted the electric current. The form of Ohm's law was neatly explained, but not so the direct proportionality between the resistivity of a pure metal and its absolute temperature. The thermionic emission of electrons could be explained, apparently, in all its details, but the distribution of energy in the black body spectrum could not. The explanation of the simple Zeeman effect was most gratifying and reassuring, but the simple numerical relationships among the frequencies of line spectra remained as baffling as ever—and this, in spite of the considerable success which Drude and others had achieved in explaining the optical properties of materials in terms of electrons elastically bound within atoms.

The impasse was finally breached by Planck who showed, in 1905, that the black body spectrum could be explained if one were willing to assume that materials contain electric oscillators which emit and absorb energy only in amounts proportional to their frequencies. The conception of the electron was unaltered—not even involved, perhaps—but an oscillator, which might be a vibrating electron, was conceived to behave in a manner contrary to electrodynamical principles. A success had been achieved at the cost of violence to classical ideas regarding the production of electromagnetic radiation.

The next assault—a brilliant tour de force by Bohr—achieved its first objectives at a stride, but at a sacrifice of electrodynamical principles greater even than Planck's. Bohr showed in 1911 that by combining the idea of a concentrated atom nucleus required by Rutherford's experiments on the scattering of alpha rays with the heterodox idea of Planck, and with new devices of his own invention, one could conceive an atom model capable of yielding precisely the Rydberg constant and the complete spectrum of atomic hydrogen. The casualties included two properties previously allotted to the electron as a matter of course: the property of radiating energy during orbital motion, and the property of revolving about the nucleus in an orbit consonant with classical dynamics and determined by initial conditions which might be regarded as arbitrary. Bohr excluded from the infinity of such orbits all but a special series. The motion of the electron remained planetary, but all else was new and bizarre.

A great initial success had, however, been attained and hope of a thorough conquest of spectra ran high,—too high as it now appears, for beyond a few other quantitative successes, further achievements were qualitative to a greater or less extent and consequently less impressive. It turned out also that advancement in the elucidation of spectra could be made only at the cost of an ever increasing array of special rules and prohibitions—additional equipment of the same arbitrary nature as that of Bohr's original postulates. Out of this necessity appeared the one new idea regarding the electron which had emerged in twenty years—the idea advanced by Goudsmit and Uhlenbeck that the electron spins and possesses in consequence a magnetic moment.

It was recognized a decade ago by Bohr and others that the attack upon the atom, despite its propitious beginning, had in a considerable measure failed; and this because it had lacked, so to speak, a proper base of operation. It was felt that the many arbitrary rules and restrictions required to correlate the data of spectroscopy must flow in a natural and unforced way from fundamental mechanical principles as yet undiscovered. A system of mechanics was envisaged which would degenerate to classical mechanics for large scale phenomena, but which would present an entirely different aspect on the atomic scale, and be capable, of course, of explaining atom dynamics as revealed by spectra.

Attempts to discover these underlying principles led to the formulation by Heisenberg in 1925 of what is known as matrix-mechanics, and led L. de Broglie in the same year to put forward his ideas concerning a so-called wave-mechanics. These proposals are said to be statements in different forms of one and the same principle—so far, at any rate, as applications to atom dynamics are concerned. It is the wave-mechanics, however, which has appealed most strongly to the physicist, and it is with this only that I will here concern myself.

The basic idea of the wave-mechanics was supplied by a paradoxical situation which had existed for some years in the theory of optics. It was well established experimentally that a beam of monochromatic light can impart to individual electrons in matter amounts of energy proportional to its frequency. The factor of proportionality between these quantities—the energy imparted to the electrons and the frequency of the light—is the same as that obtained from the black body spectrum for the factor of proportionality between the energy quantum of the Planck oscillator and its frequency. The relation between the energy ϵ imparted to the electron and the frequency ν of the light is expressed, that is, by the formula $\epsilon = h\nu$, where h is the so-called Planck constant. When one tries to visualize the mechanism back of this phenomenon, he is led inevitably to a corpuscular theory of light. No other view appears adequate to explain this central fact of photoelectricity and others related to it.

On the other hand, the phenomena of interference and diffraction disposed long ago, as is well known, of an earlier corpuscular theory of light in favor of the wave theory. The demands of these phenomena are as insistent today as every they were, so that the situation comes to this, that one class of optical phenomena indicates clearly that light is a corpuscular radiation, and another indicates no less clearly that it is a propagation of waves. It is hopeless to try explaining photoelectric phenomena in terms of nothing but waves, and it is equally hopeless trying to devise a purely corpuscular interpretation of interference and diffraction.

It was de Broglie's brilliant idea that a situation similar to this might exist in regard to electrons, that electron streams like beams of light might possess in different circumstances the properties both of wave trains and of particle streams. If this were true, and if the wave aspect alone were adequate to explain the behavior of electrons in atoms, then the unhappy state of affairs which existed in regard to the interpretation of spectroscopic data might be remedied.

The formula $\epsilon = h\nu$ expresses, as we have seen, a certain correlation between the corpuscular and the wave properties of light; if the light regarded as a beam of waves is of frequency ν, then when it is regarded as a stream of corpuscles, the corpuscles or photons are of energy $\epsilon = h\nu$. A second correlation follows at once from this one and from the relation which is known to exist between the transfer of energy and of momentum by a beam of light. This second correlation relates the momentum p of the photon to the wave-length λ in vacuo of the associated waves, and is expressed by the formula $p = h/\lambda$. Or if we write σ to represent wave number—the number of waves per cm.—then the two correlations are expressed by the symmetrical formulæ

$$\epsilon = h\nu,$$
$$p = h\sigma.$$

In developing his idea of a possible wave aspect of the electron, de Broglie was led to the conclusion (partly, it appears by intuition and partly by considerations based on relativistic mechanics) that if electrons possess wave properties, the correlation between their wave and corpuscular aspects will be expressed by these same two formulæ.

It would lead us too far afield to follow even cursorily the further development of de Broglie's idea toward its original objective of explaining the behavior of atoms as disclosed by their spectra. These interesting matters must be left with the mere statement that the mathematical researches of Schrödinger and others have led to a conception of the atom in which standing wave patterns replace the permitted electron orbits of the Bohr model, and from which it is possible to derive certain laws of spectra which previously could be given only as empirical rules.

The immediate object which de Broglie had in view in postulating a wave aspect of the electron has thus been attained, and its attainment argues strongly, of course, for the soundness of the underlying conception. It is not this spectroscopic evidence, however, which reveals most clearly the wave as a real and actual property of electrons, but the more direct and unequivocal evidence supplied by experiments described in following paragraphs in which streams of electrons are diffracted by crystals.

It was implicit in de Broglie's earliest writings regarding electron waves that a stream of electrons moving with uniform speed along parallel lines will exhibit in appropriate circumstances the properties of a beam of monochromatic waves. De Broglie's first step was, indeed, to associate a train of plane parallel waves with an electron moving with uniform speed along a straight line. It remained, however, for the young German physicist Elsasser to point out the logical conclusion to which these speculations lead, and to indicate the crucial experiment by which they might be tested: to wit, that a beam of electrons scattered by an appropriate grating should exhibit the phenomenon of diffraction, and that the appropriate grating for this purpose is a crystal, since the wave-lengths calculated from de Broglie's formula for electrons of moderate speeds are like those of X-rays, of the order of one Angstrom unit.

It is the demonstration of this phenomenon—the diffraction of electrons by crystals—which constitutes now, as has been intimated, the chief experimental evidence in support of de Broglie's conception, and it is with demonstrations of this kind that I am here primarily concerned. The first of these was made by Davisson and Germer, who showed in 1927 that beams of electrons are diffracted by a crystal of nickel, and that the wave-lenghts λ deduced from the diffraction-patterns for beams of electrons of various speeds v agree with those calculated from de Broglie's formula $\lambda = h/p = h/mv$. A second and independent demonstration was made by G. P. Thomson, who showed later in the same year that beams of high speed electrons are diffracted on transmission through thin films of polycrystalline metal, and that electron wave-lengths computed from patterns so obtained verify the de Broglie relationship.

It will be well, before considering these earliest experiments in more detail, to present certain others, made more recently, of which the interpretations are more simple. The simplest experimental result which suggests that electrons should be regarded as waves rather than as particles, is perhaps the regular reflection of a beam of electrons from the face of a crystal. The experimental arrangement used in demonstrating this phenomenon is indicated on the left in Fig. 1. Electrons emitted by a hot filament are accelerated and formed into a beam, and this beam is directed against the face of a crystal target at a known angle of incidence. The target in this case is of nickel and its face is parallel to one of the principal sets of atom planes of the crystal (111 planes). The surface is etched and presents to the incident beam a multitude of crystal facets parallel to the plane of the target face. Some of the electrons on striking these facets are scattered without loss of energy. The distribution of these full speed electrons in and near the plane of incidence is then determined by explorations with a

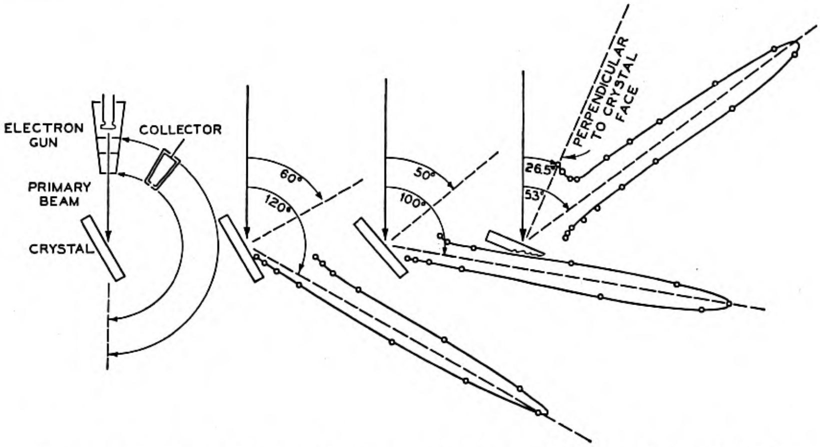


Fig. 1—Schematic diagram of apparatus for determining angular distribution of electrons scattered without loss of energy by a metallic crystal, and curves revealing specular reflection of 20-volt electrons.

Faraday collector as indicated. The results of investigations of this kind, exhibited by polar diagrams in the same figure, show clearly that the incident beam is regularly reflected as if from the crystal facets.

The difficulty in explaining this result by the simple concept of electrons as particles is, that the surface of the crystal is much too coarse grained to serve as a reflector for particles as small as electrons; the diameter of the electron is of the order 10^{-13} cm., whereas, the diameters of atoms are of the order 10^{-8} cm.—greater by a factor 10^5—and this also is the order of the distance of least separation of atoms in the crystal face. It is hard to imagine how such a surface can appear smooth to the incident electrons. On the older views regarding interactions between electrons and atoms, the fate of an incident electron should be much the same as the fate of a comet plunging into a region densely packed with solar systems; the electron *might* emerge from the crystal without loss of energy after a fortunate encounter with a single atom, but its direction of departure should be a matter of private treaty between the individual electron and the individual atom—in particular, it should not be influenced by the arrangement of atoms in the crystal. The fact that nearly all of the full speed scattered electrons move away in the direction of regular reflection from the crystal face means that three atoms at least are involved in the action, since this number is required to fix the plane of the reflecting surface. The simple observation described above is thus inexplicable in terms of atoms and electrons and their interactions as previously conceived.

It is interesting to try to imagine how this phenomenon would have been interpreted, had it been observed ten years ago. With de Broglie's speculations before us, we recognize it at once as one of the circumstances in which the electron streams exhibits the properties of a wave train. The only restriction to this interpretation is that the wave-length must be assumed small compared to the linear dimensions of the reflecting surface. If wave-lengths are given correctly by de Broglie's formula they are, as has been mentioned, of the same order as those of X-rays. This explains why reflections such as exhibited in Fig. 1 are obtained from the face of a crystal, but not from a polycrystalline surface, however highly polished; the reflected beam is synthesized, so to speak from a multitude of scattered waves spreading out from atoms regularly arranged in layers parallel to the surface—lacking this regularity the synthesis does not occur.

The specular reflection of X-rays from a crystal face is, as we know, selective in the following respect: the intensity of the reflected beam passes through sharp maxima as the glancing angle θ passes through values which satisfy the Bragg formula $\sin \theta = n\lambda/2d$, where λ represents wave-length, d the distance between atom planes parallel to the surface, and n an integer. We expect this to occur with reflection of electrons, if the de Broglie waves are scattered by underlying atom layers as well as by the outermost. I will show later that the reflection portrayed in Fig. 1 is, indeed, selective. But with low-speed electrons, there are confusing complications which we can avoid by dealing with electrons of considerably higher speed; I will, therefore, begin by displaying the selective reflection of electrons accelerated through thousands rather than tens of volts only.

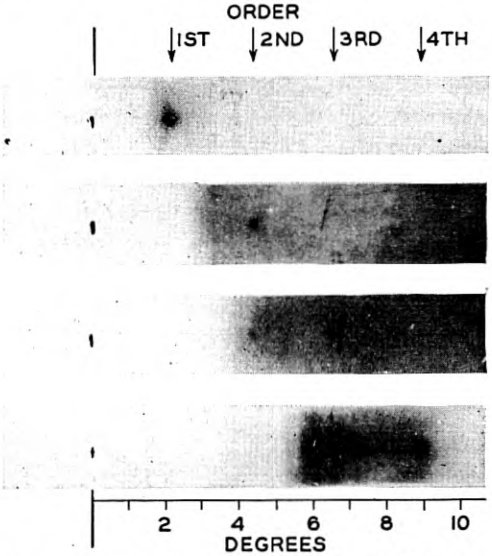


Fig. 2—Photographic record of selective reflection of 53-kv. electrons by a crystal of iron—(100) face.

In Fig. 2 we have a photographic record of the reflection of a beam of 53 kilovolt electrons from a crystal of iron. The experimental arrangement is essentially the same as that indicated in Fig. 1, though devices rather more elaborate are required for producing electron beams of such speeds. A photographic plate set at right angles to the direction of the primary beam and 30 or 40 cms. beyond the crystal replaces the exploring electrode. Each strip in Fig. 2 is the record of reflection at a particular angle of glancing. After each exposure the crystal is turned through the position of grazing (zero glancing angle), the primary beam falls directly upon the plate, and its direction is thus recorded. These fiducial marks appear in a column on the left. The glancing angles are proportional (since all of them are quite small) to the distances from the spots formed by the primary beam to the edges of the respective fogged regions; the fogging is produced by general or random scattering and its sharp cutoff on the left marks the intersection of the plane of the crystal with the photographic plate. If reflections were specular but non-selective, a spot due to the reflected beam would appear on each strip as far to the right of the fogging edge as the fiducial spot is to the left. This is not what is observed; there is strong specular reflection at a series of equally spaced angles, and at adjacent angles weaker reflection which apparently is not regular. In other ranges of angle, there is no reflection at all. This is exactly the phenomenon observed with X-rays, the apparently non-regular reflection is ascribed in the latter case to regular and selective reflection from parts of the crystal which are displaced somewhat from the mean orientation of the crystal as a whole—ascribed, that is, to imperfections in the crystal. The same explanation applies here.

Here then is an experiment in which a stream of electrons exhibits the properties of a beam of waves. The mere occurence of specular reflection is, as we have seen, incompatible with the idea that electrons are simple particles, with such properties as are commonly ascribed to particles. The further observation that the reflection is selective in accordance with the Bragg law amounts to a demonstration that we are dealing with trains of waves—or, at least, to a demonstration of the convenience of this conception—for the Bragg law is simply and accurately explained as a consequence of interference among scattered waves expanding from regularly disposed centers, such as the atoms of a crystal.

The data of the experiment enable us to calculate the length of the waves. From the Bragg law $\lambda = 2d \sin\theta/n$; the reflections are from the (100) atom planes of iron for which $d = 1.43 \times 10^{-8}$ cm. or 1.43 Ångstrom units; the value of $\sin\theta/n$ deduced from Fig. 2 and related data is 0.0189, so that $\lambda = 2 \times 1.43 \times 0.0189 = 0.054$ Ångstrom units. *This is the experimentally determined wave-length of 53 kv. electrons.*

We compare this with the *theoretical* wave-length computed from the de Broglie formula, $\lambda = h/p$. The momentum p of a particle of rest-mass m and charge e which has been accelerated from rest through a potential difference V in absolute units is given in relativistic mechanics by

$$p = (2Vem)^{1/2}\left[1 + \frac{Ve}{2mc^2}\right]^{1/2},$$

where c represents the velocity of light. Writing this into the de Broglie formula and evaluating constants for the electron one obtains as a close approximation.

$$\lambda = \left(\frac{150}{V}\right)^{1/2}[1 - 4.9 \times 10^{-7}V]$$

for V in volts. Thus the theoretical wave-length of 150-volt electrons is one Ängstrom unit or 10^{-8} cm., and the wave-length of the 53 kv. electrons employed in this experiment is 0.0546 Ängstroms which is in good agreement with the value found experimentally.

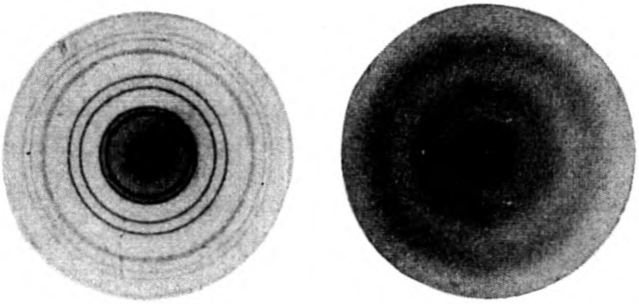
These results which are from previously unpublished data by Davisson and Germer constitute perhaps the simplest demonstration of the wave aspect of electrons and verification of the de Broglie relation—the simplest, at any rate, in which use is made of crystal diffraction.

E. Rupp has demonstrated the diffraction of electrons by an ordinary ruled optical grating and has obtained thereby values of electron wave-lengths which agree with the theoretical values within the rather wide limits of error of this border line experiment. This is, in fact, an experiment more immediately intelligible than any involving the use of crystals. But Rupp's photographic plates exhibiting these results are not very impressive, and, therefore, I have not arranged for their reproduction in this report.

The Bragg reflection, though the easiest to interpret among the types of crystal diffraction, is not the most easily demonstrated, nor the most striking, nor yet the type which yields most information concerning the diffracting crystal and the atoms composing it. It is the Hull-Debye-Scherrer type of diffraction which is in these respects preeminent, and the type which has been most thoroughly investigated. A mass of finely divided crystals of random orientation is placed in the path of a beam of monochromatic radiation; a photographic plate set at right angles to the incident beam receives and records the radiation from the diffracting material. It is evident perhaps that the pattern produced by an aggregate of a very great number of crystals oriented at random will be the same as that generated by a single crystal turned into equally many random positions. Thus, if the material is iron, the single crystal in a particular orientation gives rise to the first order diffraction spot of Fig. 2; rotate the crystal about the primary beam and this spot generates a circle or ring—so also for the second and higher spots, each of them generates a ring.

But the atoms comprising the crystal may be regarded as arranged in many different sets of planes—there are, in fact, an infinite number of such arrangements. Of these, one is unique in having a greater spacing between planes than any other—a greater value of the constant *d*. The first order ring due to these planes is the smallest in the pattern; the first order rings due to other arrangements follow in the order of decreasing values of *d*. These first order rings plus their companions of higher orders constitute the complete pattern of the aggregate. From the sequence of ring diameters, one infers the arrangement of the atoms in the crystal; from the scale of the pattern, the ratio of crystal spacing or "constant" to radiation wave-length; from the relative intensities of the rings, the way in which the scattering power of the atom varies with angle. The ring pattern is thus a storehouse of information concerning both crystal and wave.

Patterns of this type for electrons were obtained first by G. P. Thomson; it was thus, in fact, that Thomson made his demonstration of electron waves. A great many such patterns have since been obtained and studied. Two beautiful examples by Wierl are reproduced in Fig. 3. The one on the left records the diffraction of a beam of 36 kv. electrons by a thin film of polycrystalline silver, the one on the right is for 45 kv. electrons diffracted by carbon tetrachloride vapor. The form of the pattern for silver—the particular sequence of ring diameters which it displays—is characteristic of the so-called face centered cubic arrangement of atoms, and agrees with the pattern similarly obtained with X-rays. The scale of the pattern with other data of the experiment, including the scale factor of the silver crystal, yields an "observed" wave-length of 36 kv. electrons in agreement with the value computed from the de Broglie formula.



Ag CCl_4

Fig. 3—On left—diffraction pattern produced by transmission of 45-kv. electrons through thin silver foil—by R. Wierl. On right—pattern obtained by transmission of 36-kv. electrons through CCl_4 vapor—by R. Wierl.

The pattern for carbon tetrachloride is bracketed with that for silver because the chlorine atoms in this molecule have the same arrangement as the atoms in the crystal of silver. Four atoms only are required to fix a simple crystal arrangement; the chlorine atoms in CCl_4 occupy the corners of a regular tetrahedron, and this figure determines the face centered cubic arrangement. The chlorine atoms, to which nearly the whole of the scattering is due, may thus be properly thought of as forming exceedingly small crystals of this structure; the vapor is thus a crystal aggregate similar to that of polycrystalline silver. The marked differences between the two patterns are due to the disparity in the number of atoms forming the individual crystals in the two cases. The number in the chlorine crystals—four each—is the smallest possible, and the "resolving power" of the crystal, regarded as an optical instrument, is in consequence the least possible. The form of the pattern for CCl_4 accords with the tetrahedral arrangement of the chlorine atoms; the scale of the pattern with the calculated wave-length of the electrons fixes the length of the tetrahedron edge.

The purpose of this report is accomplished with the description of these few representative experiments which reveal and demonstrate—so far as demonstration is possible—a wave aspect of electrons in conformity with de Broglie's conception. The experiments do not tell us in what medium, if any, the waves occur, with what speed they are propagated, whether they are longitudinal or transverse and capable of polarization—they tell us merely that when electrons reach an element of space from a given source simultaneously over different paths the resultant intensity (the number of electrons traversing the element per unit time, as we think of it) is not the arithmetic sum of various scalar contributions as we naturally expect it to be, but is given instead by the square of the sum of contributions which are vector quantities—in precisely the manner with which we are familiar in optics. We must make the addition as if we were dealing with superposed trains of waves—with due regard for phase as well as amplitude. This kind of addition is characteristic of trains of waves. When we find quantities which add together in this particular way we conclude that the quantities are, in fact, trains of waves; this is our reason for regarding light and X-rays as wave phenomena and it is our reason also for so regarding electrons.

Having demonstrated the convenience, if not indeed the necessity, of regarding electrons in certain circumstances as waves rather than as particles, we enquire naturally if these waves are refracted on passing from one medium to another like the waves of light and X-rays, and whether or not they are polarizable. We rather expect to find that they do exhibit refraction, for if the wave-length of a beam of electrons in vacuo is given by $\lambda = (150/V)^{1/2}$ we expect that on passing into a metal of which the thermionic work function is φ the wave-length of the beam will be altered to $\lambda' = [150/(V + \varphi)]^{1/2}$. We expect, in other words, that the metal will have for electrons accelerated through V volts a refractive index given by $\mu = \lambda/\lambda' = (1 + \varphi/V)^{1/2}$. Something of this kind is indeed found experimentally, though the phenomenon is less simple than we have here imagined.

Evidence of refraction is contained in the experimental results displayed in Fig. 4. The ordinates of this curve are proportional to the

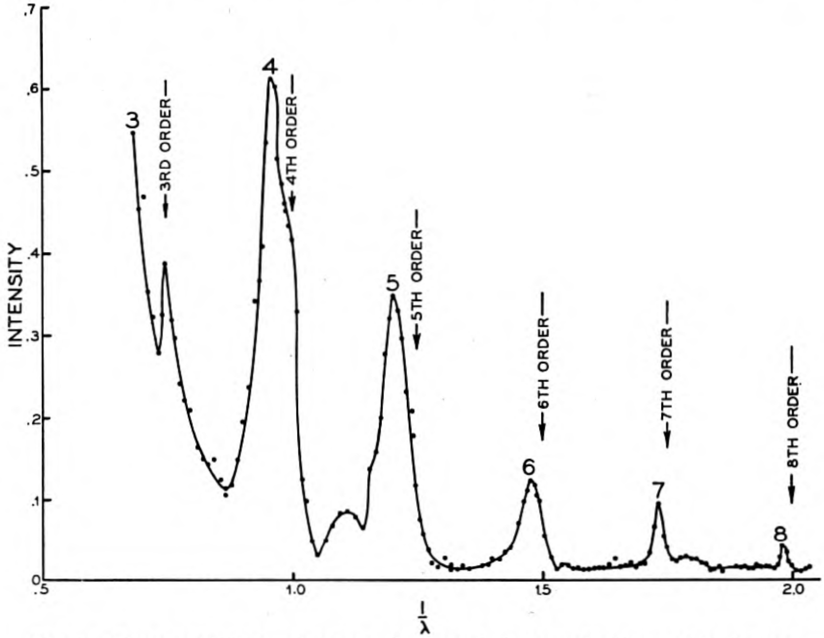


Fig. 4—Curve exhibiting selective reflection of electrons incident at 10 degrees on nickel crystal—(111) face. Departures from simple Bragg law reveal refraction.

intensity of the beam regularly reflected at 10 degrees incidence from a nickel crystal; the abscissa are proportional to the reciprocal of the wave-length of the incident beam. The maxima in the curve represent selective Bragg reflections of a sequence of orders, but are displaced somewhat to the left from the positions calculated from the simple Bragg formula, and indicated in the figure by arrows.

The simple Bragg formula is derived with the assumption, however, that the refractive index of the crystal is unity. The more general formula for reflection from a crystal face is $n\lambda = 2d(\mu^2 - \cos^2\theta)^{1/2}$ which reduces to the familiar form when the index μ is equal to one. Thus if λ_1 represents the wave-length at which Bragg reflection of a given order is expected ($\mu = 1$) and λ the wave-length at which the reflection maximum is actually observed, then the refractive index of the crystal satisfies the formula

$$(\mu^2 - 1) = \sin^2 \theta \left(\frac{\lambda^2}{\lambda_1^2} - 1 \right)$$

Or if in agreement with our assumed dispersion formula $\mu^2 - 1 = \varphi / V$ the constant φ will be given by $\varphi = (V_1 - V) \sin^2 \theta$ where V_1 and V are the voltages corresponding to wave-lengths λ_1 and λ.

On our simple view of the matter, we expect this formula to yield the same value of φ for all orders of reflection and for all angles of incidence. This expectation is not, however, realized. From Fig. 4 we obtain for the value of φ for nickel 14 or 15 volts. But under other conditions values are obtained as low as 10 volts and as high as 20 or 25 volts. The results shown in Fig. 2, if rightly interpreted, also are incompatible with the assumed law of dispersion. Thus, if θ represents the calculated glancing angle at which the nth order Bragg reflection occurs for a given wave-length when $\mu = 1$ ($\varphi = 0$), then for $\varphi \neq 0$ this reflection is to be expected at angle θ_1 such that

$$\frac{\sin \theta}{\sin \theta_1} = \left(1 - \frac{4d^2\varphi}{150n^2} \right)^{1/2}.$$

It will be noted that the right hand side of the equation does not involve the speed or wave-length of the incident electrons from which we conclude that relative defections from the simple Bragg law should be as conspicuous for high speed electrons as for low, and this we find not to be the case; the reflections recorded in Fig. 2 conform to the simple law, as if φ were equal to zero. The situation then is this, that while we have clear evidence that electron waves are refracted, the laws of their refraction are evidently not simple, and are yet to be discovered.

I turn now to the polarization of electron beams. Rupp, in a recent series of remarkable experiments, has shown that a beam of electrons may, in appropriate circumstances, exhibit an asymmetry with respect to its direction of motion. In these experiments, beams of electrons are diffracted by thin films of gold and annular patterns are obtained like the one for silver shown in Fig. 3. They differ, however, from the patterns ordinarily obtained in that the individual rings are not uniformly dark all around, or as we say, in all "azimuths." In a certain azimuth the density of each ring is at a maximum, in the opposite azimuth (180° away) it is at a minimum—the rings are stronger or denser on one side of the pattern than on the other. This signifies a non-uniformity or asymmetry in the number of electrons scattered into the various segments of a ring, and is due to an asymmetry in the source of the electrons constituting the beam incident upon the gold foil. These, instead of coming directly from a filament or from a high voltage discharge, as is ordinarily the case, are electrons which have already been scattered through 90 degrees by a metal target. The primary beam of electrons moved in a line parallel to the diffracting film and at some distance away; it fell upon a metal target; some of its electrons were scattered through 90° by this target, forming a secondary beam which was the one diffracted by the film of gold. Azimuths about a secondary beam formed in this way are not indistinguishable: one of them is unique in containing the primary beam which fell upon the target. It is in this azimuth that the electrons are most copiously scattered by the diffracting film, and the rings are most dense.

The sense of the effect may be stated in a different way. If the density of the rings were independent of azimuth, the mean total deflection from the primary beam (the deflection at the target plus the deflection at the film) of the electrons forming the pattern would be 90°: but actually, the mean deflection is greater than 90°. The electrons in the secondary beam have a bias toward a further deflection in the same sense as that which they received at the target.

Rupp has further shown that the effect of passing the secondary beam (the beam from the target to the film) through a *longitudinal* magnetic field is, so to speak, to rotate the azimuth of polarization; the azimuth of maximum density is displaced from its original position (without magnetic field) in a sense which depends upon the direction of the field and by an amount proportional to the length and intensity of the field.

The effect of a *transverse* magnetic field is different for different azimuths; if the direction of the field is normal, to the plane of the primary beam incident upon the target and the secondary beam the effect is nil. If it is parallel to this plane the polarization of the secondary beam decreases as the field strength is increased; at a critical intensity depolarization is complete, and at intensities which are higher still, the polarization is reversed—the azimuths of maximum and minimum density are interchanged. In these tests with transverse magnetic fields the force on the electron due to its motion through the magnetic field is balanced by the force due to a suitably adjusted electrostatic field.

The results obtained by Rupp in these experiments may be explained by means of the concepts of electron spin and corresponding magnetic moment. The spin-axes of the electrons of the primary beam we imagine to be oriented at random. But among the electrons scattered by the target this is no longer true; those scattered in any small solid angle have a non-uniform distribution in orientation; the resultant of all their spins is a vector normal to the plane containing the primary beam and the direction of scattering. Thus, the resultant spin vector is everywhere tangent to circles about the primary beam, in the same general way as the magnetic vector in the field about a wire carrying an electric current. If the polarization process is one of selection—if, that is, the scattering material selects from the incident electrons those of a particular orientation to scatter in a particular direction, then it is to be expected that double scattering will exhibit just the type of asymmetry which Rupp actually observes.

The effect of a magnetic field upon a spinning electron is to cause its spin vector to precess about an axis parallel to the direction of the field. The same statement may be made in regard to the resultant spin vector of an assemblage of electrons such as that which constitutes the beam incident upon the gold foil in Rupp's experiments. This action is competent to explain the various changes which Rupp observes in the state of polarization of the primary beam when it has traversed the differently directed fields. The magnitude of the effect yields a value of the ratio of the magnetic moment of the electron to its angular momentum and this agrees well with the theoretical value of this ratio, e/mc.

It will have been remarked perhaps that in picturing these polarization effects, we have reverted to the conception that electrons are particles. It is difficult to see how they are to be explained in terms of the characteristics of waves. There is, however, no particular reason for making the effort. Here, as in other circumstances, the choice of conception is entirely a matter of convenience.

An Expansion for Laplacian Integrals in Terms of Incomplete Gamma Functions, and Some Applications*

By EDWARD C. MOLINA

INTRODUCTION

LAPLACE has given us, in the Théorie Analytique des Probabilités, Book I, Part II, Chapter I, a method of approximating by means of series to the value of a definite integral of the type

$$I = \int_{x}^{x_2} y_1^{\theta_1} y_2^{\theta_2} \cdots y_n^{\theta_n} \phi dx,$$

where y_1, y_2, $\cdots$ y_n are functions of x whose exponents θ_1, θ_2, $\cdots$ θ_n are of the same order of magnitude as a large number θ. The last function in the integrand, ϕ, embraces all factors whose exponents are of low order of magnitude compared with θ. The integral here considered must not be confused with the well known "Laplacian Transform" integral which is also embodied in the Théorie Analytique.

When the function $\phi(x)$ is other than a mere constant, the Laplacian method as presented in the Théorie Analytique does not, in certain cases, give us a series which reduces to its first term as θ approaches infinity. The object of this paper is to present a modification of the method which gives the desired result and to present two applications of considerable practical importance. The modification consists in divorcing the function ϕ from the factors of the integrand raised to high powers and associating ϕ with the factor (dx/dt) which makes its appearance with the Laplacian change of variable from x to t.

DEDUCTION OF MODIFIED EXPANSION

Setting $\theta_s/\theta = r_s$ we have

$$y_1^{\theta_1} y_2^{\theta_2} \cdots y_n^{\theta_n} = (y_1^{r_1} y_2^{r_2} \cdots y_n^{r_n})^{\theta} = y^{\theta},$$

(1)
$$I = \int_{x}^{x_2} [y(x)]^{\theta} \phi(x) dx,$$

which we shall write in the form

$$I = \int_{x}^{x} [y(x)^{N} \phi_{w}(x) dx,]$$

---

* Presented by title at International Congress of Mathematicians, Zurich, Switzerland, September, 1932.

where $N = \theta + w$, $\phi_w(x) = [y(x)]^{-w}\phi(x)$. Introducing the parameter w does not constitute an essential modification of the procedure given by Laplace, but we shall find that w plays an important part when we come to applications of the expansion presented in this paper.

In what follows it is assumed that $y(x)$ is a positive monotonically increasing or decreasing function in the range x_1, x_2; the extension to a range of integration divisible into subranges within which $y(x)$ is monotonic will be obvious.

Let X be that one of the two limits x_1, x_2 for which $y(x)$ has its greatest value; set $y(X) = Y$. Assume with Laplace that

$$[\log Y - \log y(x)]^{\frac{1}{u+1}} = (x - X)g(x, X), \tag{2}$$

u being a positive number or zero and g a function of which $(x - X)$ is not a factor. Set

$$y = Ye^{-t^{u+1}}, \tag{3}$$

$$\phi_w(x)(dx/dt) = A_0 + A_1\frac{t}{1!} + A_2\frac{t^2}{2!} + A_3\frac{t^3}{3!} + \cdots. \tag{4}$$

These two equations give (certain well-known conditions being fulfilled)

$$I = Y^N \sum_{s=0} A_s \int_{t_1}^{t_2} \frac{t^s e^{-Nt^{u+1}} dt}{s!}. \tag{5}$$

Finally set $Nt^{u+1} = T$ and we obtain the *modified Laplacian expansion*

$$I = Y^N \sum_{s=0} \left(\frac{1}{N}\right)^{H_s} B_s[P(H_s, T_2) - P(H_s, T_1)] \tag{6}$$

wherein

$$H_s = \frac{s+1}{u+1},$$

$$B_s = \left(\frac{A_s}{u+1}\right)\frac{\Gamma(H_s)}{s!}, \qquad = A_s \quad \text{for} \quad u = 0;$$

$$T_1 = N[\log Y - \log y(x_1)], \qquad = 0 \quad \text{if} \quad X = x_1;$$

$$T_2 = N[\log Y - \log y(x_2)], \qquad = 0 \quad \text{if} \quad X = x_2;$$

$$P(H_s, T') = \frac{\int_0^{T'} T^{H_s - 1} e^{-T} dT}{\Gamma(H_s)},$$

$$T' = T_1 \text{ or } T_2.$$

For the computation of $P(H_s, T')$ recourse may be had to the extensive "Incomplete Gamma Function" Tables edited by Karl Pearson which give the ratio of the incomplete to the complete function. When H_s is a positive integer we have

$$P(H_s, T') = \sum_{k=H}^{\infty} \frac{T'^k e^{-T'}}{k!},$$

the well known Poisson Exponential Binomial Limit for which short tables will be found in Pearson's "Tables For Statisticians And Biometricians" and in T. C. Fry's "Probability and Its Engineering Uses."

Applications

I

Consider the incomplete Beta Function

$$I(p) = \int_0^p (1-x)^{n-c} x^{c-1} dx$$

for positive integral values of n and c such that n is large compared with c. Its *modified Laplacian expansion* is

$$I(p) = (c-1)!(n+w)^{-c} \sum_{m=0} (n+w)^{-m} A(m, c-1) P(c+m, T_2),$$

where $T_2 = -(n+w)\log(1-p)$ and, setting $w = (d-c+1)/2$,[1]

$$A(0, c-1) = 1,$$

$$A(1, c-1) = \binom{c}{1}\frac{d}{2},$$

$$A(2, c-1) = \binom{c+1}{2}\left[\frac{(c-1)+3d^2}{12}\right],$$

$$A(3, c-1) = \binom{c+2}{3}\left[\frac{(c-1)d+d^3}{8}\right],$$

$$A(4, c-1) = \binom{c+3}{4}\left[\frac{5(c-1)^2-(2-30d^2)(c-1)+15d^4}{240}\right].$$

As many more coefficients as one desires can be obtained by means of the equations

$$A(m, c-1) = \sum_{k=0}^{m} Q(m, k)\binom{c-1+m}{m+k},$$

$$Q(m+1, k) = (m+k)Q(m, k-1) + (w+k)Q(m, k),$$

$$Q(o, o) = 1.$$

---

[1] See Appendix II for the details of the analyses.

The expansion given above for $I(p)$ is of great practical value in connection with the evaluation of the incomplete binomial summation. We know [2] that

$$P(c;\,n,\,p) = \sum_{x=c}^{n} \binom{n}{x} p^x(1-p)^{n-x}$$
$$= \frac{n!}{(c-1)!(n-c)!}\int_0^p x^{c-1}(1-x)^{n-c}dx;$$

therefore, we have

$$P(c;\,n,\,p) = \frac{n!}{(n+w)^c(n-c)!}\sum_{m=0}\left(\frac{1}{n+w}\right)^m A(m,\,c-1)P(c+m,\,T_2).$$

This approximation formula is submitted for ranges of values of c and n such that c is of the order of magnitude of $\sqrt{n}$ or less. But for values of $P(c;\,n,\,p)$ which are of most interest, say,

$$.0001 < P < .9999,$$

c is of the same order of magnitude as

$$a = np$$

when n is a large number. Therefore, the formula is particularly valuable for values of p which are of the same order of magnitude as $1/\sqrt{n}$ or less.

The first table at the end of this paper indicates the degree of accuracy obtainable by taking $d = 0$ and using only one or two terms of the approximation for $P(c;\,n,\,p)$. The table also indicates the result of taking d such that $T_2 = a = np$, and using only one, two, three or four terms of the approximating series.

I am indebted to Miss Elizabeth McCusker and Miss Catherine Lennon of the Department of Development and Research of the American Telephone and Telegraph Company for the computation work involved in the construction of the tables at the end of this paper.

II

Consider the Bessel Function

$$I_n(b) = \frac{(\frac{1}{2}b)^n}{\sqrt{\pi}\Gamma(n+\frac{1}{2})}\int_0^\pi (\text{sinx})^{2n}\, e^{b(\text{cosx})}\, dx,$$

---

[2] See Laplace, "*Théorie Analytique des Probabilités,*" 1st edition, p. 151.

for b large compared with n, n being a positive integer. Writing

$$e^{-b}I_n(b) = \frac{(2b)^n}{\sqrt{\pi}\Gamma(n+\frac{1}{2})}\int_0^\pi (\tfrac{1}{2}\,\mathrm{sinx})^{2n}[e^{-(\frac{1}{2}-\frac{1}{2}\cos x)}]^{2b}dx,$$

leads us to the *modified Laplacian expansion*

$$e^{-b}I_n(b) = \frac{\left(\dfrac{2b}{2b+w}\right)^n}{\sqrt{\pi(2b+w)}}P(n+\tfrac{1}{2},\,2b+w)S_n,$$

where

$$S_n = \sum_{m=0}\left(\frac{1}{8b+4w}\right)^m \frac{A_{2n+2m}}{2(2n)!}\,\frac{P(n+m+\frac{1}{2},\,2b+w)}{P(n+\frac{1}{2},\,2b+w)}\,\frac{n!}{(n+m)!}.$$ [3]

To determine the coefficients A_{2n+2m} we proceed as follows: For the integral now under consideration

$$\begin{aligned}
&N = 2b+w, \qquad x_1 = 0, \qquad x_2 = \pi,\\
&y(x) = e^{-(\frac{1}{2}-\frac{1}{2}\cos x)}, \qquad X = 0, \qquad Y = 1,\\
&(\log Y - \log y)^{\frac{1}{2}} = (\tfrac{1}{2}-\tfrac{1}{2}\cos x)^{\frac{1}{2}} = xg(x),\\
&g(x) = \frac{1}{x}\sin\frac{x}{2} = \sum_{k=0}\frac{(-1)^k x^{2k}}{2^{2k+1}(2k+1)!},
\end{aligned}$$

so that

$$\begin{aligned}
&u = 1, \qquad y(x) = e^{-t^2}, \qquad T_1 = 0, \qquad T_2 = 2b+w,\\
&\phi_w(x) = (\tfrac{1}{2}\,\mathrm{sinx})^{2n}e^{w(xg)^2},\\
&B_{2n+2m} = [n!\Gamma(n+\tfrac{1}{2})/(n+m)!(2n)!2^{2m+1}]A_{2n+2m}.
\end{aligned}$$

The relations above give

$$\phi_w(x)(dx/dt) = 2t^{2n}(1-t^2)^{n-\frac{1}{2}}e^{wt^2}.$$

Let

$$\sum_{s=0}A'_{2n+2s}t^{2n+2s}/(2n+2s)! = 2t^{2n}(1-t^2)^{n-\frac{1}{2}},$$

so that

$$\begin{aligned}
A'_{2n} &= 2(2n)!,\\
A'_{2n+2s} &= \frac{(-1)^s(2n-1)(2n-3)\cdots(2n-2s+1)(2n+2s)!}{2^{s-1}s!};
\end{aligned}$$

then

$$\begin{aligned}
\phi_w(x)(dx/dt) &= e^{wt^2}\sum_{s=0}A'_{2n+2s}\frac{t^{2n+2s}}{(2n+2s)!}\\
&= \sum_{k=0}\sum_{s=0}t^{2n+2s+2k}A'_{2n+2s}w^k/k!(2n+2s)!
\end{aligned}$$

---

[3] An expansion of this type for Bessel functions was derived many years ago by Hadamard; see Watson, "A Treatise on the Theory of Bessel Functions," pp. 204 and 205.

and

$$A_{2n+2m}/(2n + 2m)! = \sum_{k=0}^{m}(w^k/k!)A'_{2n+2m-2k}/(2n + 2m - 2k)!$$

We must now assign a value to w. One which suggests itself is to take w such that the second term in our expansion for $I_n(b)$ vanishes; in other words, such that A_{2n+2} shall be zero. This merely requires that we set $w = n - 1/2$, giving

$A_{2n} = 2(2n)!, \qquad A_{2n+2} = 0$
$A_{2n+4} = -(2n - 1)(2n + 4)!/2, \qquad A_{2n+6} = -(2n - 1)(2n + 6)!/3$
$A_{2n+8} = (2n - 1)(2n - 5)(2n + 8)!/16$, etc.
For $n = 1$ and $w = 1/2$ we have

$$A_2 = 4, \qquad A_4 = 0, \qquad A_6 = -360, \qquad A_8 = -13440, \cdots$$

$$e^{-b}I_1(b) = \left[\left(\frac{2b}{2b + .5}\right) P(3/2,\, 2b + .5)/\sqrt{\pi(2b + .5)}\right] S_1$$

wherein

$$S_1 = 1 - 15(8b + 2)^{-2}\frac{P(7/2,\, 2b + .5)}{P(3/2,\, 2b + .5)} - 140(8b + 2)^{-3}\frac{P(9/2,\, 2b + .5)}{P(3/2,\, 2b + .5)} - \cdots$$

The degree of accuracy obtainable when only one, two or three terms of S_1 are made use of is indicated by the second table at the end of this paper. For some values of b, ranging from 6 to 16, comparison is made between the successive approximations for $e^{-b}I_1(b)$ and its true value as given by Watson in his Theory of Bessel Functions. The last three columns of Table II give the figures for $w = 0$ instead of $w = 1/2$.

Appendix I

When the expansion of $\phi_w(x)(dx/dt)$ is not obvious we may proceed as does Laplace in expanding (dx/dt).

Equations (2) and (3) give

$$x = X + tg^{-1}. \tag{7}$$

This last equation, together with the Lagrange-Laplace expansion theorem, gives for a function of x the expansion

$$f(x) = f(X) + \sum_{s=1}\frac{t^s}{s!}\left[\frac{d^{s-1}}{dx^{s-1}}\left(\frac{1}{g}\right)^s f'(x)\right]_{x=X}.$$

But $df(x)/dt = f'(x)(dx/dt)$. Therefore, taking $f(x)$ such that $f'(x) = \phi_w(x)$,

$$\phi_w(x)\frac{dx}{dt} = \sum_{s=0}\frac{t^s}{s!}A_s,$$

$$A_s = [D_x{}^s g^{-(s+1)}\phi_w(x)]_{x=X}, \qquad D_x{}^s = \frac{d^s}{dx^s},$$

$$= Y^{-w}\sum_{v=0}\frac{w^v}{v!}[D_x{}^s(x-X)^{(u+1)v}g^{-(s+1)+(u+1)v}\phi(x)]_{x=X},$$

or

$$\frac{A_s}{s!} = Y^{-w}\sum_{v=0}\frac{w^v}{v!}\frac{[D_x{}^{s-(u+1)v}\phi(x)g^{-(s+1)+(u+1)v}]_{x=X}}{[s-(u+1)v]!}. \tag{8}$$

Beta Function

For the incomplete Beta Function we have,

$$X = 0, \qquad u = 0, \qquad Y = 1, \qquad g(0) = 1,$$

$$\phi(x) = x^{c-1}(1-x)^{-c} = \sum_{r=0}\binom{c-1+r}{r}x^{c-1+r}.$$

Therefore,

$$\frac{A_s}{s!} = \sum_{v=0}^{s-(c-1)}\frac{w^v}{v!}\sum_{r=0}\binom{c-1+r}{r}\frac{[D_x{}^{s-v}x^{c-1+r}g^{-(s+1-v)}]_{x=0}}{(s-v)!},$$

or

$$\frac{A_s}{s!} = \sum_{v=0}^{s-(c-1)}\frac{w^v}{v!}\sum_{r=0}^{s-(c-1)-v}\binom{c-1+r}{r}\frac{[D_x{}^{s-(c-1)-v-r}g^{-(s+1-v)}]_{x=0}}{[s-(c-1)-v-r]!}.$$

Since $A_s = 0$ for $s < c-1$, set $s = c-1+m$; then interchanging v and $m-v$ we obtain

$$\frac{A_{c-1+m}}{(c-1+m)!} = \sum_{v=0}^{m}\frac{w^{m-v}}{(m-v)!}\sum_{r=0}^{v}\binom{c-1+r}{c-1}\frac{[D_x{}^{v-r}g^{-(c+v)}]_{x=0}}{(v-r)!}. \tag{9}$$

To evaluate (9) we may have recourse to the formula [4]

$$D_x{}^M g^{-S} = S\binom{S+M}{M}\sum_{K=1}^{M}\left[\frac{(-1)^K\binom{M}{K}D_x{}^M g^K}{(S+K)g^{S+K}}\right], \tag{10}$$

taking

$$M = v - r,$$

$$S = c + v.$$

---

[4] See formula 25, page 15, of Schlömilchs Höheren Analysis, Band II, Dritte Auflage, 1879. I am indebted to Mr. J. Riordan of the Department of Development and Research, American Telephone and Telegraph Company, for calling my attention to this formula.

Bessel Function

For the Bessel integral with b large compared with n we have

$$X = 0, \qquad u = 1, \qquad g(0) = 1/2, \qquad Y = 1,$$

$$\begin{aligned}\phi(x) &= \left\{\left[\sin\left(\frac{x}{2}\right)\right]^2\left[1 - \sin^2\left(\frac{x}{2}\right)\right]\right\}^n = [(xg)^2 - (xg)^4]^n \\ &= \sum_{r=0}^{n}\binom{n}{r}(-1)^r(xg)^{2n+2r}.\end{aligned}$$

Substituting in (8) we have

$$\frac{A_s}{s!} = \sum_{v=0}\frac{w^v}{v!}\sum_{r=0}^{R}(-1)^r\binom{n}{r}\frac{[D_x^{s-2v}x^{2n+2r}g^{-(s-2n-2v-2r+1)}]_{x=0}}{(s-2v)!},$$

where $2R \not> s - 2n - 2v$; note that if $R > n$, then $\binom{n}{r} = 0$.

Since $A_s = 0$ for $s < 2n$ and g is an even function of x, set $s = 2n + 2m$. Then Leibnitz's theorem and interchanging v with $m - v$ give

$$(11) \qquad \frac{A_{2n+2m}}{(2n+2m)!} = \sum_{v=0}^{m}\frac{w^{m-v}}{(m-v)!}\sum_{r=0}^{v}(-1)^r\binom{n}{r}\frac{[D_x^{2v-2r}g^{-(2v-2r+1)}]_{x=0}}{(2v-2r)!},$$

where again we may have recourse to (10) for the differentiations of negative powers of $g(x)$.

APPENDIX II

Writing the incomplete Beta Function in the form

$$I(p) = \int_0^p (1-x)^{n+w}x^{c-1}(1-x)^{-(c+w)}dx,$$

we now have

$$\begin{aligned}&N = n + w, \qquad x_1 = 0, \qquad x_2 = p, \\ &y(x) = (1 - x), \qquad Y = 1, \qquad X = 0, \\ &\phi_w(x) = x^{c-1}(1-x)^{-(c+w)}, \\ &\log Y - \log y = -\log(1 - x), \\ &\qquad\qquad = xg(x), \qquad g(x) = \left(1 + \frac{x}{2} + \frac{x^2}{3}\cdots\right), \\ &u = 0, \\ &y(x) = e^{-t}, \\ &T_1 = 0, \\ &T_2 = -(n + w)\log(1 - p).\end{aligned}$$

From these equations we derive

$$\begin{aligned}\phi_w(x)(dx/dt) &= e^{wt}(e^t - 1)^{c-1} \\ &= \sum_{k=0}^{c-1} (-1)^k \binom{c-1}{k} e^{(w+c-1-k)t} \\ &= \sum_{m=0} \left(\frac{t^s}{s!}\right)\left[\sum_{k=0}^{c-1} (-1)^k \binom{c-1}{k} (w+c-1-k)^s\right], \\ & \qquad\qquad s = c-1+m.\end{aligned}$$

Therefore (see any work on finite differences),

$$A_{c-1+m} = \Delta^{c-1}(w)^{c-1+m} = (-1)^m \Delta^{c-1}(-w-c+1)^{c-1+m}.$$

Now

$$\begin{aligned}\Delta^{c-1}(w)^{c-1+m+1} &= w\Delta^{c-1}(w)^{c-1+m} + (c-1)\Delta^{c-2}(w+1)^{c-1+m} \\ &= (w+c-1)\Delta^{c-1}(w)^{c-1+m} + (c-1)\Delta^{c-2}(w)^{c-2+m+1}.\end{aligned}$$

Increasing c by 1, decreasing m by 1 and setting

$$A_{c-1+m} = (c-1)!A(m,\ c-1),$$

we have

$$A(m,\ c) - A(m,\ c-1) = (w+c)A(m-1,\ c)$$

or

$$\Delta_c A(m,\ c-1) = (w+c)A(m-1,\ c)$$

where the subscript c implies that Δ_c operates on c. Assume that

$$A(m,\ c-1) = \sum_{k=0}^{m} Q(m,\ k)\binom{c-1+m}{m+k},$$

where the coefficients $Q(m,\ k)$ are functions of w; then

$$\begin{aligned}\Delta_c A(m+1, c-1) &= \sum_{k=0}^{m} Q(m,\ k)(w+c)\binom{c+m}{m+k} \\ &= \sum_{k=0}^{m+1} Q(m,\ k)\left[(m+k+1)\binom{c+m}{m+k+1} + (w+k)\binom{c+m}{m+k}\right] \\ &= \sum_{k=0}^{m+1} [(m+k)Q(m,\ k-1) + (w+k)Q(m,\ k)]\binom{c+m}{m+k},\end{aligned}$$

from which, by finite integration with reference to c, we obtain

$$A(m+1, c-1) = \sum_{k=0}^{m+1} [(m+k)Q(m,\ k-1) + (w+k)Q(m,\ k)]\binom{c-1+m+1}{m+1+k}$$

so that

$$\begin{aligned}Q(m+1,\ 0) &= wQ(m,\ 0), \\ Q(m+1,\ k) &= (m+k)Q(m,\ k-1) + (w+k)Q(m,\ k), \\ & \qquad\qquad 0 < k < m+1, \\ Q(m+1,\ m+1) &= (2m+1)Q(m,\ m).\end{aligned}$$

Since

$$A(0,\ c-1) = 1 = \binom{c-1}{0},$$

we know that $Q(0, 0) = 1$. Hence

$$A(1, c-1) = w\binom{c}{1} + \binom{c}{2},$$

$$A(2, c-1) = w^2\binom{c+1}{2} + (3w+1)\binom{c+1}{3} + 3\binom{c+1}{4},$$

$$A(3, c-1) = w^3\binom{c+2}{3} + (6w^2+4w+1)\binom{c+2}{4} + (15w+10)\binom{c+2}{5} + 15\binom{c+2}{6}.$$

$$A(4, c-1) = w^4\binom{c+3}{4} + (10w^3+10w^2+5w+1)\binom{c+3}{5} + (45w^2+60w+25)\binom{c+3}{6} + (105w+105)\binom{c+3}{7} + 105\binom{c+3}{8},$$

etc.

As yet the value of w has not been specified. Since when m is an odd number and $w = -(c-1)/2$, the equation

$$\Delta^{c-1}(w)^{c-1+m} = (-1)^m\Delta^{c-1}(-w-c+1)^{c-1+m}$$

gives us

$$\Delta^{c-1}\left(-\frac{c-1}{2}\right)^{c-1+m} = 0,$$

let us write $w = (d/2) - (c-1)/2$, leaving d arbitrary; we obtain

$$A(0, c-1) = 1,$$

$$A(1, c-1) = \binom{c}{1}\frac{d}{2},$$

$$A(2, c-1) = \binom{c+1}{2}\left[\frac{(c-1) + 3d^2}{12}\right],$$

$$A(3, c-1) = \binom{c+2}{3}\left[\frac{(c-1)d + d^3}{8}\right],$$

$$A(4, c-1) = \binom{c+3}{4}\left[\frac{5(c-1)^2 - (2-30d^2)(c-1) + 15d^4}{240}\right],$$

etc. Finally, for $d = 0$, or $w = -(c-1)/2$,

$$A(0, c-1) = 1, \qquad A(1, c-1) = 0,$$

$$A(2, c-1) = \binom{c+1}{3}\frac{1}{4}, \qquad A(3, c-1) = 0,$$

$$A(4, c-1) = \binom{c+3}{5}\left(\frac{5c-7}{48}\right), \qquad A(5, c-1) = 0$$

etc.

TABLE I

INCOMPLETE BINOMIAL SUMMATION

$$P(c;\, n,\, p) = \frac{n!}{(c-1)!(n-c)!}\int_0^p x^{c-1}(1-x)^{n-c}dx$$

np	n	p	c	$P(c;\, n,\, p)10^6$	$w = -(c-1)/2$		w Such that $T_2 = np$			
					$(P - P_1)10^6$	$(P - P_2)10^6$	$(P - P_1)10^6$	$(P - P_2)10^6$	$(P - P_3)10^6$	$(P - P_4)10^6$
5	50	.1	2	966,214	78	0	− 77,687	4,340	− 198	8
			4	749,706	428	0	− 51,250	2,399	− 90	3
			6	383,877	518	0	− 802	520	− 2	0
			8	122,145	273	0	9,911	676	34	2
			10	24,538	77	0	4,058	426	34	2
			12	3,220	13	0	789	117	13	1
	100	.05	2	962,919	19	0	− 36,837	986	− 22	0
			4	742,161	101	0	− 23,907	535	− 10	0
			6	384,001	123	0	− 190	123	0	0
			8	127,960	68	0	5,209	175	4	0
			10	28,188	21	0	2,378	124	5	0
			12	4,274	4	0	547	40	2	0
10	100	.1	2	999,678	25	0	− 98,756	7,193	− 460	27
			4	992,164	244	0	− 155,547	14,675	− 1,085	68
			6	942,423	739	0	− 152,886	14,185	− 997	58
			8	793,949	1,268	1	− 94,502	7,232	− 430	22
			10	548,710	1,399	2	− 25,277	2,065	− 83	4
			12	296,967	1,054	2	11,472	1,244	43	3
			14	123,877	563	1	15,423	1,511	112	7
			16	39,890	221	1	8,242	1,102	112	10
			18	10,007	65	0	2,832	495	65	7
			20	1,979	15	0	699	152	24	3

TABLE I—*Continued*

np	n	p	c	$P(c;\, n,\, p)10^6$	$w = -(c-1)/2$		w Such that $T_2 = np$			
					$(P - P_1)10^6$	$(P - P_2)10^6$	$(P - P_1)10^6$	$(P - P_2)10^6$	$(P - P_3)10^6$	$(P - P_4)10^6$
	200	.05	2	999,596	6	0	− 47,020	1,644	− 51	1
			4	990,951	59	0	− 72,566	3,261	− 116	3
			6	937,657	177	0	− 70,458	3,100	− 104	3
			8	786,695	300	0	− 43,386	1,575	− 45	1
			10	545,290	330	0	− 11,460	462	− 9	0
			12	300,244	252	0	6,002	309	5	0
			14	129,892	140	0	8,241	402	15	0
			16	44,355	58	0	4,750	318	16	0
			18	12,089	18	0	1,811	160	10	0
			20	2,664	4	0	509	56	4	0

$P - P_m$ = error incurred by using only the first m terms of the series for $P(c;\, n,\, p)$ presented in this paper.

TABLE II

BESSEL FUNCTION
$e^{-b}I_1(b)$

b	P	$w = n - 1/2 = 1/2$ for $n = 1$			$w = 0$		
		$\left(\frac{P-P_1}{P}\right)$	$\left(\frac{P-P_2}{P}\right)$	$\left(\frac{P-P_3}{P}\right)$	$\left(\frac{P-P_1}{P}\right)$	$\left(\frac{P-P_2}{P}\right)$	$\left(\frac{P-P_3}{P}\right)$
6	.152052	−.007497	−.001457	−.000331	−.071134	−.004188	−.000701
8	.134143	−.004043	−.000587	−.000098	−.051474	−.002187	−.000261
10	.121263	−.002529	−.000293	−.000038	−.040358	−.001344	−.000125
12	.111464	−.001731	−.000167	−.000018	−.033198	−.000911	−.000070
14	.103698	−.001259	−.000104	−.000010	−.028198	−.000658	−.000042
16	.097350	−.000958	−.000070	−.000006	−.024510	−.000497	−.000029

P = value of $e^{-b}I_1(b)$ from Table II of Watson's Theory of Bessel Functions.

$\left(\frac{P-P_m}{P}\right)$ = proportional error incurred by using the first m terms of the series for $e^{-b}I_1(b)$ presented in this paper.

Contemporary Advances in Physics, XXIV

High-Frequency Phenomena in Gases, First Part

By KARL K. DARROW

This is an account of the behavior of conducting gases subjected to high-frequency electrostatic fields—behavior which can be interpreted, in many cases with striking success, by supposing that the free electrons wandering in the gas are set into motion by the field, and oscillate and drift according to laws which can be derived from our knowledge of the response of free electrons to steady fields. When a constant magnetic field coexists with the high-frequency forces, the phenomena become more complicated, but remain predictable. There are also peculiar phenomena indicating that the electrons in a conductive gas have certain natural frequencies of oscillation. Applications are made to the absorption of radio-frequency waves in ionized gases.

IN this article I will describe some of the phenomena which are observed when the voltage across a region filled with gas is varying quite rapidly. Considered as a function of time, the voltage may be periodic, a sine-wave with uniform amplitude and of a frequency somewhere between a few thousands and a few hundreds of millions of cycles per second. It may be a succession of highly-damped wavetrains, each commencing with a rapid rise of voltage and continuing in oscillations of high frequency but swiftly declining amplitude, which die away into nothing and after an interval (it may be of a few hundredths or a few thousandths of a second) are followed by another train. It may be a brief irregular spasm of electromotive force, of which the highest value of the voltage is measured or merely guessed. In the gas itself there may be the phenomenon of sudden and violent breakdown; or the establishment of a self-sustaining luminous discharge, like in aspect to a glow or an arc; or merely a vibratory motion of electrons, freed by other agencies and set in motion by the oscillating field.

One might regard this as a subject which physicists had better leave alone, until they have full understanding of the seemingly much easier problems of discharges across a gas exposed to a constant voltage. So great are the apparent advantages of steady or "direct-current" discharges for the student, so great the apparent inconveniencies of high frequencies, that one might reasonably think it futile to assail the latter with weapons which have not yet overcome the former. Thus, in a self-sustaining glow maintained by a constant voltage, there is a peculiar distribution of space-charge, which distorts the field between the electrodes in a characteristic way, and manifests itself by a striking subdivision of the gas into zones of light and zones of darkness, differing much from one another in their electrical state as well as in their aspect. This distribution is not established instantly; would one not do better to examine it at leisure and understand it in full for each separate value of voltage, before beginning to study discharges in which it is continually changing over from the form appropriate to one voltage into the form appropriate to another? or discharges in which the voltage is varying so rapidly, that the gas is always in a state of transition, never even approaching the equilibrium appropriate to any steady value of voltage whatever?

Well, this is not necessarily to be supposed! it may be that when the voltage across the gas varies with extreme rapidity, the gas itself enters into a sort of equilibrium-state, perhaps even a more intelligible state than that which it attains when it is allowed an ample time to adjust itself to a constant value of voltage. The simplicity of certain empirical laws of the high-frequency glow suggests this, as also does the aspect of the glow. There is also the following argument, rather paradoxical in sound, perhaps, but forcible. A self-sustaining direct-current discharge, a glow or an arc, involves a steady outflow of electrons from its cathode. This outflow must be maintained by agents coming out of the gas itself—photons and positive ions and excited atoms, which are generated in the gas by the discharge and fall upon the cathode. Of the relative prominence of these agents, little is known—it forms one of the major problems of the steady discharge in gases; but at least it is certain, that they owe their existence and their effectiveness largely to the distribution of space-charge in the gas. The distribution of the space-charge therefore is controlled by the requirement (or seems controlled—one never knows what is cause and what is effect) that the electron-outflow from the cathode must be kept at a level suitably high. Now in the high-frequency discharge, the demand for electrons from the cathode is attenuated or abolished; witness the fact that such discharges may be maintained when the electrodes are separated by insulators from the gas! The peculiar disposition of the space-charge is therefore not demanded, at least not to so great an extent; conditions are intrinsically simpler.

This feature of high-frequency discharges—their competence to do without electrons from the cathode—requires attention. It is derived from a fundamental principle, which one is all too likely to forget if one has long been occupied with steady currents: the principle of Maxwell, that an electric field when varying in time is equivalent to a current. Let us apply this principle to a current in a circuit (Fig. 1) composed of a source E of electromotive force; wires leading (one of them through a galvanometer G) from the opposite poles of the source to electrodes A and C; and a gas between the electrodes. For convenience let us imagine the gas enclosed in a cylindrical vertical tube, the electrodes near its ends.

Take up first the direct-current discharge; say the cathode is above; consider any surface, S, cutting across the gas which the tube encloses. After the steady state is established, there will be positive ions crossing the surface on their way toward the cathode (upward) and negative ions crossing it on their way toward the anode (downward). We may speak of these, both classes, as ions going in the "right" direction. The total charge which they carry, as many of them as cross in unit time, will be the current across the surface —if there are no ions going in the "wrong" directions. But in a luminous discharge there are generally positive ions crossing toward the anode and negative ions crossing toward the cathode. The sum of the charges which they carry must be subtracted from the former sum; or, to express the idea better, they must be added with a

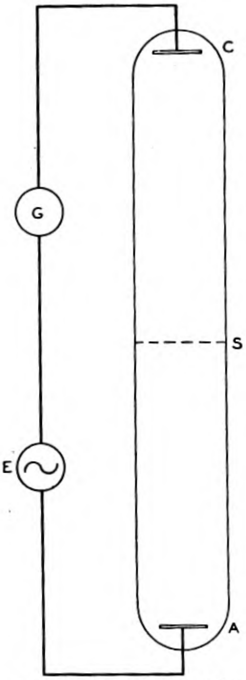


Fig. 1—Illustrating an internal-electrode tube for the transmission of high-frequency current through gas.

negative sign. We have to form a fourfold sum: charges borne by positive ions going up and charges borne by negative ions going down, with *plus* sign; charges borne by negative ions going up and charges borne by positive ions going down, with *minus* sign. The terms of this fourfold sum may vary from one to another of the surfaces intersecting the tube (in general, they do) but the sum remains the same; it is the current through the gas. It is also the current flowing through the plane of the cathode-surface (it may then consist of two terms only, electrons emerging from the cathode metal and positive ions impinging on it; or it may involve the two additional terms, because of electrons falling upon the cathode and positive ions rebounding from it). It is also the current flowing through the wires, and the current measured by the galvanometer at *G*.

Now for a high-frequency discharge, this must be modified. The fourfold sum aforesaid is not the whole of the current. Or rather, we might call it the whole of the current, but then we should be compelled to say that the current across the various cross-sections of the circuit is not the same, and is not measured by the galvanometer. We do better to follow Maxwell, or rather the usage which developed out of Maxwell's theory: call it the "net convection current" and introduce the name "displacement current" for the quantity which must be introduced, in order to get a sum which is the same for all the cross-sections of the circuit and equal to the reading of the galvanometer. Nor is this "sum" to be obtained by simple addition. We are obliged to take into account another complication, from which direct-current discharges are free: the necessity of distinguishing between a current-component which is proportional to the voltage and a current-component which is proportional to the rate of change of the voltage; or, to take the simple case of a sinusoidal voltage, the current-components which are respectively "in phase" and "in quadrature" therewith. The displacement-current belongs entirely to the latter component, while the convection-current may belong partly to the one and partly to the other, as I will presently illustrate.

It is, then, *not* required that the net convection-current should be the same all through the gas; much less, that it should be the same as the current reported by the galvanometer. At sufficiently high frequencies, electrons and positive ions may oscillate in the interspace between anode and cathode without reaching either, if they start their independent careers far enough off from both. Can one attain a condition in which there is a convection-current of oscillating ions in the middle of the gas, but none whatever in the vicinity of the electrodes, and no electrons escape from the metal of the cathode into the gas? Without an impervious screen between the metal and the gas this would, I suppose, be impracticable; but if there be such a screen—if for instance the electrodes are outside of the glass-walled tube containing the gas, instead of being inside—it can be achieved. The ions in the imprisoned gas can be set into oscillatory motion by an alternating voltage applied to electrodes outside; indeed, a self-sustaining luminous glow-discharge can be maintained within the tube; the charged particles which vibrate in the shining gas never enter the metal parts of the circuit nor transfer their charges to these. Only by invoking Maxwell's concept of the displacement current are we able to contend that there is continuity of current-flow around the circuit and across the gap between the anode and the cathode.

The so-called "current through the gas"—i.e. the reading of the galvanometer in series with the circuit—is therefore a datum decidedly remote from the phenomena within the gas. Even in the simplest of all cases, that of a direct-current discharge, it merely gives the sum of the four terms aforesaid, not the value of any individually. In the much more intricate case of the high-frequency discharge, it is a combination of the four terms aforesaid and the displacement-current. This is probably why it figures so little in accounts of the high-frequency discharge.

If the current through the gas is hard to interpret when measured, the voltage across it may present an easier problem, or may on the other hand be unmeasurable altogether. The simplest case of all is that of a gas in contact with electrodes (as in the tube of Fig. 1) between the two of which an undamped sinusoidal voltage is applied. An alternating-current electrostatic voltmeter, of one type or another, shunted across the electrodes, will give the "effective" or "root-mean-square" value of the potential-difference between them. It is often the maximum value of the voltage which the experimenter wants, or thinks that he wants; to get it he must multiply the reading of the voltmeter by $\sqrt{2}$. This is not the proper factor unless the voltage is truly sinusoidal; the observer must find out about this (by using an oscillograph to make the waveform visible, or by hunting with a wavemeter for harmonics of the fundamental frequency) and use a different factor if the waveform is distinctly not a pure sine-function. If the electrodes are outside of the tube containing the gas, some part of the potential-gradient between them is located in the insulating walls, and does not act upon the gas. If the voltage is applied as a sequence of highly-damped short wavetrains with intervals between, the problem of determining its maximum value is a serious one. Perhaps the best available methods have never been applied in experiments in this field; at any rate, many a physicist has contented himself with measuring something which he thought to be proportional to the maximum voltage, without knowing what value to assign to the factor of proportionality.

The toughest problem of all, in respect of measuring current and voltage—or let me say, in respect of finding something significant to measure—is forced upon us by a form of discharge which Hittorf invented (this is probably a more suitable word than "discovered") in 1884. He wound a wire spirally around a tube containing air at low pressure, connected the spiral across a Leyden jar and the terminals of an induction-coil, and thus sent through the wire a sequence of current-pulses which were highly damped wavetrains; they incited a brilliant glow in the tube. Within the spiral, and therefore pervading the gas in the tube, there was of course a magnetic field parallel to the axis and alternating its direction with the alternations of the pulsing current. There was also a circular electric field due to the varying magnetic field, pointing alternately clockwise and counter-clockwise around the axis. Also there was a varying electric field due to the alternating potential-differences between the windings of the spiral. All of these three must have influenced the mobile ions of the gas! Often, as Thomson was later to stress, the discharge takes the form of a brilliant ring, thus suggesting that it is the second of the fields which dominates. But the problem of their relative responsibilities is a subtle and very difficult one; and the "ring discharge" is as troublesome to elucidate by theory, as it is easy to realize in practice.

I have written thus far as though we were concerned entirely with two sorts of stable conditions: the self-sustaining luminous high-frequency discharge, and the oscillations of ions in a gas where the supply of ions is maintained not by the alternating voltage but by some other agency acting independently. We are concerned with these, and with a third matter as well: the process of "breakdown," the sudden onset of the self-sustaining discharge which occurs when the amplitude of the high-frequency voltage across a gas hitherto tranquil is elevated past a critical value. This critical value or "breakdown-potential" is the most frequently measured of all the measurable qualities of the discharge; though strictly speaking it is scarcely a quality of the discharge, but rather of the tranquil gas of which it marks the oncoming transformation. Presumably it is preceded by an intermediate state, in which the oscillating voltage both displaces the ions in the gas, and gives them energy enough to form others; but of this we as yet know little.

I will arrange the optics much as I have previously arranged them in the description of direct-current discharges:[1] first, the observations on gases independently ionized and exposed to an oscillating voltage which displaces the independently-formed ions; then the observations on breakdown or transition, in which the oscillating voltage is raised to such a value that it initiates and maintains a self-sustaining discharge; and finally, the observations on the self-sustaining discharge itself. Beforehand, though, it will be convenient to derive some equations describing the presumable behavior of ions in gases exposed to alternating fields.

First, consider a charged particle free to move in a vacuum, and exposed to an alternating electric field. One might expect it to oscillate to and fro across a fixed point, like a simple harmonic vibrator. This however occurs only in a special case. In general, the particle will oscillate about a point which glides at uniform speed along the direction of the field. For, let us write down the equation for the acceleration of the particle, denoting by e its charge, by m its mass, by E_0 the amplitude and by ν the frequency of the field E, which last we suppose directed along the axis of x:

$$m(d^2x/dt^2) = eE = eE_0 \sin 2\pi\nu t. \tag{1}$$

Integrating once, we obtain for the speed of the ion,

$$dx/dt = \frac{1}{2\pi\nu}\frac{eE_0}{m}(1 - \cos 2\pi\nu t) + v_0, \tag{2}$$

and integrating a second time, we obtain for its displacement from the point which it occupied at $t = 0$:

$$x = \left(\frac{1}{2\pi\nu}\frac{eE_0}{m} + v_0\right)t - \frac{1}{4\pi^2\nu^2}\frac{eE_0}{m}\sin 2\pi\nu t. \tag{3}$$

The first term on the right of equation (3) represents the steady drift of the centre of oscillation, the second term the oscillation to and fro across this centre.

The speed of the steady drift is the coefficient of t in equation (3). Its value depends on that of the constant of integration v_0, which, as one sees from equation (2), is the speed (or rather, the x-component of the velocity) of the ion at the instant $t = 0$ when the electric field passes through zero. If this speed just happened to be equal to $eE_0/2\pi\nu m$, and directed in the sense opposite to that in which the increasing field was about to draw the ion, the particle would oscillate about a fixed point; the amplitude of its vibrations would amount to $eE_0/4\pi^2\nu^2 m$, its maximum kinetic energy to $e^2E_0^2/8\pi^2\nu^2 m$. But this is

---

[1] In a recent book, "Electrical Phenomena in Gases," abstracted in the July 1932 *Bell Sys. Tech. Jour.*, and to which I refer at various points in this article.

a very improbable event; it is difficult to conceive of any mechanism plausible or unplausible which would endow all the ions in a gas with such an initial speed. We can scarcely make use of any but the most general formulæ; among which I choose those for the drift-speed u and the maximum kinetic energy K_m of the ions:

$$\begin{aligned} u &= \frac{1}{2\pi\nu}\frac{eE_0}{m} + v_0, \\ K_m &= \frac{1}{2}m\left[v_0 + \frac{1}{\pi\nu}\frac{eE_0}{m}\right]^2. \end{aligned} \tag{4}$$

If we knew the distribution-in-speed of the ions at the instant $t = 0$, and they were in a vacuum (i.e., never collided with atoms), we could predict their future motions. But they are not in a vacuum, and we do not know their distribution-in-speed at $t = 0$. What use, then, can be made of the equations?

Little or nothing, I am afraid, of such definiteness as to permit of exact and verifiable predictions about high-frequency discharges! After all, it is of the essence of a discharge that the phenomena occur in a gas where molecules and ions make collisions with each other; inferences drawn from the assumption that ions never collide with molecules are not likely to be close to truth. Nevertheless we may make some deductions of general value.

Thus: the expressions for the amplitude of the vibration of the ion, the constant speed of the point about which it vibrates, and the maximum kinetic energy acquired by the ion, all of them decrease when m is increased. All of them either are inversely proportional to m, or else involve a term inversely proportional to m. Therefore all of them are much smaller for positive ions than for free electrons. If a gas contains ions of both these kinds, the free electrons have by far the most energy, oscillate the farthest and drift the fastest. On this account we may often pretend that all the ions in the gas are stationary, excepting only the free electrons; and I shall often make this pretence.

Again: the fact that the center of oscillation of each ion drifts (except in a special case, presumably very rare) suggests that even in a vacuum it is impossible to confine the ions to any restricted part of the region between the electrodes.

Finally: the value given in (4) for the maximum kinetic energy of the ions, with the particular choice of zero for the value of v_0, may be taken as an indication of the order of magnitude of the energy which an ion might acquire in a gas not so dense but that occasionally it might run without a collision for a time as long as the duration of one oscillation. But for more than the order-of-magnitude we should certainly not rely on it!

Since allowance should be made for the collisions of the ions with molecules of the gas, let us do what is customary in physics: rush to the extreme, exaggerate the effect instead of neglecting it, and suppose that collisions are so frequent that their net influence upon a wandering ion amounts to a force of the type called "viscous" or "frictional"—a force proportional and oppositely directed to the velocity of the particle. Into the equation of motion we introduce a new term, so that it takes the form:

$$m(d^2x/dt^2) + g(dx/dt) = eE = eE_0 \sin nt \tag{5}$$

g standing for the new "coefficient of friction," and n being written for $2\pi\nu$ so as to avoid constant repetition of the factor 2π. One integration is sufficient to bring out the important point; we obtain:

$$\frac{dx}{dt} = \frac{eg}{m^2n^2 + g^2} E - \frac{me}{m^2n^2 + g^2} \frac{dE}{dt}. \tag{6}$$

If now we multiply each member of the equation by Ne — N standing for the number of ions per unit volume—each becomes equal to the current-density borne by ions in the gas.

Notice now an important thing about this expression for current-density. It consists of two terms, one of which is proportional to the fieldstrength E, while the other is proportional to the time-derivative of the fieldstrength—two terms, therefore, of which one is in phase with E and the other in quadrature with E. The first we may set down in the following equation:

$$\textit{Current-density in phase with fieldstrength} = \frac{Ne^2g}{m^2n^2 + g^2} E = \sigma E \tag{7}$$

and we may apply the name "conductivity" to the coefficient σ. The second must be treated differently. By itself it is not the whole of the current-density in quadrature with the fieldstrength; there is the displacement-current also to be taken into account, equal to $(1/4\pi)(dE/dt)$. Then we must write:

Current-density in quadrature with fieldstrength

$$= \frac{1}{4\pi}\left(1 - \frac{4\pi Ne^2m}{m^2n^2 + g^2}\right)\frac{dE}{dt} = \frac{\epsilon}{4\pi}\frac{dE}{dt} \tag{8}$$

and we may apply the name "dielectric constant" to the coefficient ϵ.

Say now that we can ionize a gas abundantly, as for instance by maintaining an intense direct-current discharge across it between auxiliary electrodes; and that we can apply a small oscillating voltage transversely, between a pair of electrodes facing one another like the plates of a condenser; and that we can measure, in the circuit containing this oscillating voltage and these condenser-plates, the components of current in phase and in quadrature with the field—then by equations (7) and (8), we can determine N and g. Or if we can measure either component by itself, and N in some independent way, we can determine g and test the equations by comparing its value with one derived from our knowledge of the flow of ions through gases in a steady field. I will describe how this is done, after giving equations for a more special and a more general case.

If the coefficient g is very large compared to mn—we shall later see that this is the same as saying: if an ion makes very many collisions with molecules during a single cycle of the applied field—the component of convection-current in quadrature with fieldstrength almost vanishes. The ions describe simple-harmonic vibrations about fixed centres (there is no possibility of a steady drift of the centre of vibration), the velocity being in phase with the field; the amplitude of vibration is easily shown to be given by this expression:

$$\text{Amplitude} = eE_0/2\pi\nu g = \mu E_0/2\pi\nu. \tag{9}$$

I take this opportunity of mentioning that the quantity e/g, the quotient by E of the drift-speed attained by the ion under a *steady* field of strength E, is called the "mobility" of the ion at the pressure in question, and is denoted by μ.

If the equation of motion of the ion includes in addition to the two terms in the left-hand member of equation (5) a third term fx (I have written it out with this additional term farther along in the article, as equation 26), the expressions for the current-components in phase and in quadrature with field strength involve the coefficient f. The formulae for σ and ϵ become the following:

$$\sigma = \frac{Ne^2g}{[(f/n) - mn]^2 + g^2} \qquad \epsilon = 1 - \frac{4\pi Ne^2[m - (f/n^2)]}{[(f/n) - mn]^2 + g^2}. \tag{10}$$

The term in question and its coefficient f would naturally be introduced if we were dealing with bound electrons, but it seems odd to postulate such a term in speaking of freely-moving electrons or ions; nevertheless there is reason to do so, as I will later mention.[2]

When a gas containing free electrons is traversed by electromagnetic waves, there is at every point an oscillating electric field having the frequency of the waves, and the electrons vibrate according to the

---

[2] There is also an allowance for polarization which should be made if extreme accuracy is desired. See Appleton & Chapman, *l.c. infra.*

foregoing laws. The waves are then absorbed—that is to say, their amplitude falls off exponentially as they progress—and their phase-velocity is altered. In the particular case in which g and therefore σ are negligibly small, there is no absorption, and the phase-velocity is altered in the ratio $1 : \sqrt{\epsilon}$—a fact which is described by saying that the index of refraction of the electron-populated gas is equal to $\sqrt{\epsilon}$. Since ϵ is less than unity according to equation (8), the waves go faster than they would if there were no free electrons roaming the gas. Formulae for index of refraction and absorption-coefficient, in the general case in which g and σ do not vanish, are given further along in this article (equations 15, 16). The transmission of radio waves through the atmosphere of the earth, in which free electrons are present in a concentration varying greatly with altitude, is much affected by this refraction.

EXPERIMENTS ON OSCILLATION OF ELECTRONS IN IONIZED GASES

There is a method which, in principle, is adequate for measuring the dielectric constant and the conductivity of an ionized gas subjected to high-frequency vibrations; adequate therefore, in principle, for testing the expressions (7) and (8), for evaluating the quantities g and N which figure in those expressions. I will describe the method, first in the form in which it was applied to ionized air by Szekely, with results which are of some interest in themselves but mainly serve to illustrate how great a gulf intervenes between "adequate in principle" and "adequate in practice."

The gas—air at some pressure of the order of a few hundredths of a millimeter of mercury—was contained in a long tube, having electrodes near its ends whereby a direct current could be passed through the tube, maintaining the air in a state of intense ionization. Inside the tube and near its middle were a pair of parallel plane electrodes, rectangular in shape (18 by 42 mm.) and 4.5 mm. apart; to these the high-frequency voltage was applied; I will speak of them hereafter as "the ionization-condenser." When the direct current was flowing, a self-sustaining glow-discharge existed in the tube, and the ionization-condenser was submerged in the negative glow thereof.

The ionization-condenser was shunted by an adjustable condenser outside of the tube, and the two of them by an inductance and an adjustable resistance; all this constituted a circuit—I will call it hereafter the "high-frequency circuit"—coupled at one place to a source of high-frequency E.M.F., at another place to a detector. With such an arrangement, when one varies the capacity of the adjustable condenser and leaves everything else unchanged, the current in the circuit (measured by the detector) passes through a maximum at a certain setting of the capacity. At the attainment of this maximum, the system is said to be in resonance.

In Szekely's experiments, the adjustable condenser was set to produce resonance in three different conditions: in case (*a*) the ionization-condenser was entirely disconnected; in case (*b*) it was connected, but there was no direct-current discharge flowing in the tube, and the air was not ionized; in case (*c*) it was connected, and the air was ionized by the direct-current discharge. The values C_a, C_b, C_c of the adjustable capacity were measured at resonance under the three conditions; their differences give the values of the capacity of the ionization-condenser (plus that of its leads, which apparently is a large part of the total) when the gas is ionized and when it is not. Out of these data one may calculate the dielectric constant of the ionized gas. As for the conductivity: the values i_{mb} and i_{mc} of the high-frequency current at resonance are different in cases (*b*) and (*c*), smaller in the latter. After measuring i_{mc}, one may return to condition (*b*)—shutting off the direct-current discharge—restore the resonance, and reduce i_{mb} to the value just found for i_{mc} by adding resistance to the circuit. From the amount ρ of added resistance which is necessary to achieve this, one may calculate the resistance between the condenser-plates when the air is ionized; and thence—taking account of the size and shape of the plates and the distance between them—the conductivity of the ion-populated air.

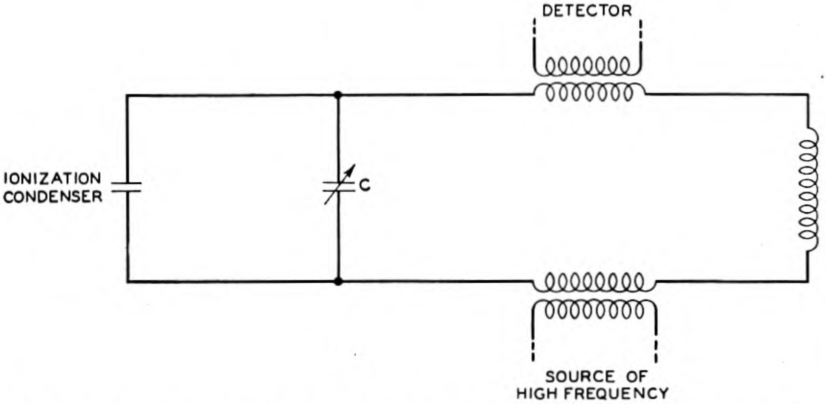


Fig. 2—Illustrating a method for measuring dielectric constant and conductivity of an ionized gas (between the plates of the "ionization-condenser").

Everything thus is apparently provided for testing the expressions and determining the constants derived from the theory of electrons oscillating in the ionized gas. But, the results of the experiments were not agreeable. True, in one important respect there was concordance with theory. Szekely measured the resistance of the ionization-condenser (or, to express it better, the resistance of the ionized air between the plates) for various values of the direct current sustaining the ionization, and various frequencies ranging from one to five millions. Plotting against the square of the frequency those values of resistance which belonged to one and the same strength of direct current, she obtained ascending straight lines (Fig. 3). Now, this agrees with the equation (7) supplied by the theory, according to which the reciprocal of the conductivity of an ionized gas is a linear ascending function of frequency squared, if N be constant;

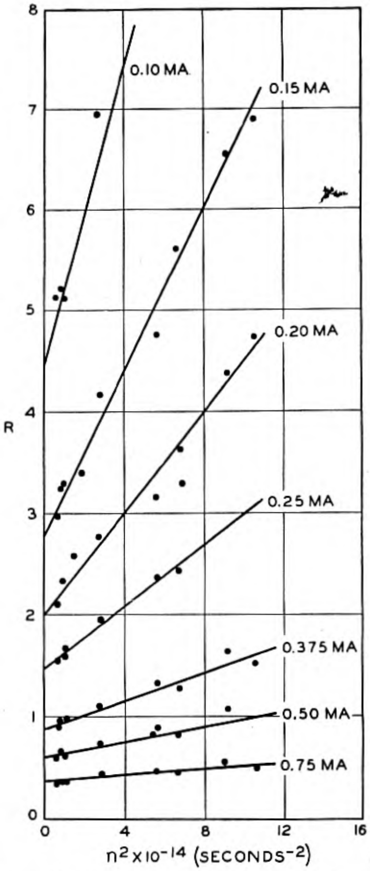


Fig. 3—High-frequency resistance of ionized rarefied air, plotted against square of frequency. (A. Szekely; *Annalen der Physik.*)

but the values computed from these curves for N and g are not very plausible, as I will later stress. What is much more serious: the capacity of the ionization-condenser is increased by ionizing the air; the dielectric constant of the ionized air is greater than unity, instead of less! The method therefore, adequate as it seems in principle, suffers from some defect or defects, which it is important to discover.

One of these defects was recognized towards 1925 by Appleton, who in previous work by the same sort of method had obtained the same bothersome result: when the air was ionized, the capacity of the ionization-condenser usually went up instead of down (although he, and van der Pol before him, did have the satisfaction of observing a decline of the capacity, when the strength of the direct current and hence the degree of ionization were relatively low). This however he explained by taking into account the space-charge sheaths of positive ions which form upon plates immersed in a strongly-ionized gas, or for that matter upon the walls of the tube containing the gas, if they are allowed to assume the potential which they naturally seek.[3]

If each of the two plates of the ionization-condenser is overspread by a positive-ion sheath of thickness x (x being less than half the distance d between the two plates, otherwise the sheaths would overlap) the system behaves not like a single condenser but like two in series. The capacity of each of the two, according to Appleton and Childs, varies inversely as x; their formula for each is $AK/3\pi x$, K standing for what they define as "the effective dielectric constant of the sheath" and A for the area of the plate; it differs from the customary formula for the capacity of a plane condenser—$AK/4\pi x$—because of the distribution of charge in the volume between the plate and the outer edge of the sheath. The capacity of the two in series is $AK/6\pi x$. If x is sufficiently small, this will be considerably larger than the capacity $A/4\pi d$ of the ionization-condenser as it was when the gas was not yet ionized; and the change in capacity occurring when ionization is started will consist chiefly of the substitution of the value $AK/6\pi x$, for the value $A/4\pi d$—even the influence of the layer of ionized gas between the outer surfaces of the sheaths will be minor.

If these ideas are correct, then when x is small the change of capacity of the ionization-condenser should vary inversely as x, provided we can vary x without altering K. Now when the potential V of the plates relative to the bulk of the ionized gas is varied, and the intensity of the ionization in the gas remains the same, the thickness of the sheath varies as $V^{3/4}$ while the current-strength across it remains

---

[3] "Electrical Phenomena in Gases," pp. 355–371.

unchanged.[4] If K remains the same, then the increment of capacity due to the ionization should vary as $V^{3/4}$; its logarithm should be a linear function of the logarithm of V, the line having slope $-3/4$. This inference was tested by Appleton and Childs, with the favorable result displayed in Fig. 4.

The space-charge sheaths thus furnish an explanation for the unexpected sign of the change in capacity occurring when the gas is suddenly filled with ions—an explanation releasing us from having to assume that the dielectric constant of the gas rises above unity, in contradiction to the simple theory.

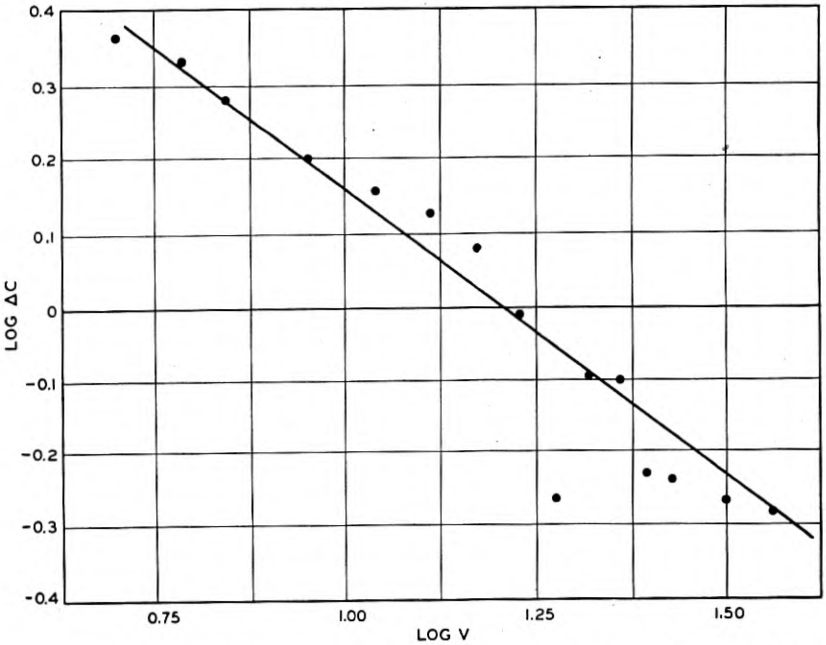


Fig. 4—Alteration in capacity of a stratum of ionized gas, ascribed to the formation of positive-ion space-charge-sheaths along the surfaces of the electrodes bounding the stratum. The slope of the line is $-\frac{3}{4}$. (E. V. Appleton & E. C. Childs; *Philosophical Magazine.*)

Could the sheaths be eliminated? If x is allowed to increase indefinitely, several things happen; in particular, the assumption that the thickness varies as $V^{3/4}$ and the assumption that $K/6\pi x$ is much larger than $1/4\pi d$ depart further and further from the truth, and

---

[4] "Electrical Phenomena in Gases," equation (182), page 360. The equation is the familiar "three-halves-power equation," which derives its usual name from the relation between voltage and current-density with which we are not here concerned. It is valid only if the ions mostly cross the sheath without collisions, an important restriction.

when x passes $d/2$ the sheaths overlap and the conditions change entirely. Other things being equal, x increases as the number of ions per unit volume of the gas outside the sheath decreases; and it is actually found that as the direct current maintaining the ionization is reduced, the increase of capacity due to the ionization is also reduced, passes through zero, and becomes a diminution as the simple theory indicates.

One might on the other hand get rid of the sheaths altogether, by adjusting the (mean) potential of the plates to so low a value that x vanishes, instead of leaving it to seek its own level. This has recently been done by Childs, though for some reason he chose to study not the dielectric constant, but the conductivity of the ionized air. A galvanometer was coupled into the high-frequency circuit, which was tuned to resonance when the mean potential of the plates had been so adjusted that the visible sheaths had just vanished. The reading of the galvanometer was taken; the current maintaining the ionization (it was 300-cycle A.C., instead of D.C.) was discontinued, and various high resistances were connected in parallel with the condenser plates, until the galvanometer resumed its former reading; the value of resistance then existing was taken as that of the ionized air between the condenser-plates. The edge-correction of the condenser was determined by filling the tube with alcohol of known conductivity; it was found that the conductance of the condenser had to be multiplied by 0.57 to be converted into conductivity of the medium between the plates. The values of conductivity (I will presently state the conditions more precisely) were of the order of 10^{-14} E.M.U.

A value of conductivity by itself, obtained at a single frequency, is theoretically the value of a combination of N and g; to determine either, without the aid of a simultaneous measurement of dielectric constant, one must have an independent value of the other. Childs evaluated N by the Langmuir probe-method.[5] Working with air at 1 mm. pressure and a frequency about one million, for three values of the 300-cycle A.C. maintaining the ionization (5 ma., 10 ma., 15 ma.) he obtained three values of the electron-concentration N (they were 9.6, 24 and 36 times 10^7 per cc.), and substituted them, not into the general equation (7) for σ, but into an approximate form thereof:

$$\sigma = Ne^2/g, \tag{11}$$

believing g to be so large that mn is small by comparison. The values of g thus obtained were 2.5, 3.1 and 2.5 times 10^{-18}. Despite

---

[5] "Electrical Phenomena in Gases," pp. 351–352.

the large differences between the values of N, those of g show no trend, which is as it should be.

Another way of determining g without bothering about N has just been carried into effect by Appleton and Chapman. Returning to equation (7): one readily sees that, if it be possible to vary g without altering N, the conductivity should pass through a maximum when $g = mn$. Now according to the simple interpretation the coefficient g, being the measure of a sort of friction which the electrons suffer when travelling through a gas, should increase with the pressure. If therefore one can vary the pressure of the gas sufficiently while applying the same high frequency, and maintaining the same degree of ionization, one should find the conductivity passing through a maximum. Moreover, anything proportional to the conductivity of the gas should pass through a maximum, anything inversely proportional to the conductivity should pass through a minimum. This conveniently makes it unnecessary to measure the conductivity (or anything else) absolutely; it is sufficient to choose something which varies (say) inversely with conductivity, plot it as function of pressure, and locate the minimum of the curve. For the pressure corresponding to this minimum, the value of g is equal to the product of the mass of the electron by the applied frequency.

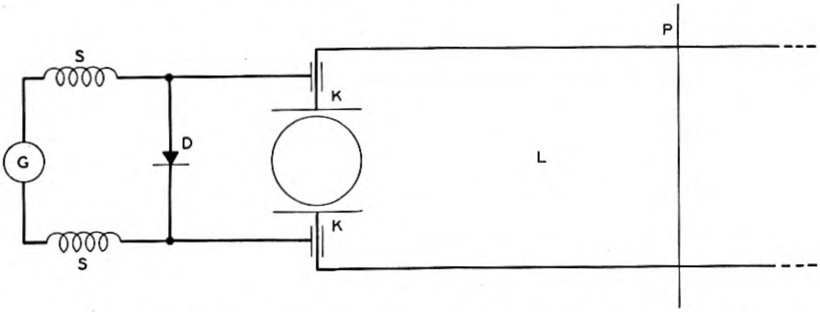


Fig. 5—Illustrating another scheme for estimating the dielectric constant and conductivity of an ionized gas (in the tube between the plates K), with Lecher wires and movable bridge P replacing the adjustable condenser of Fig. 1. (E. V. Appleton & F. W. Chapman; *Proc. Phys. Soc. London.*)

The apparatus and the circuit of Appleton and Chapman are shown in Fig. 5; as one notices, the plates of the ionization-condenser are now outside the tube and the ionized gas, and the high-frequency circuit now comprises long parallel wires, known (after the man who first used them as portions of high-frequency circuits) as Lecher wires. The current through the galvanometer to the left (shunted by a galena detector) is the measured quantity; the change which occurs in the reading of the galvanometer, when the discharge and the ionization commence in the tube (resonance being restored in the circuit, by shifting the "bridge" across the wires), is the quantity which supposedly varies inversely as conductivity. In the curves of Fig. 6, one sees the minima. Such curves were plotted for four frequencies—340, 500, 550 and 625 millions—and the maxima occurred at the pressures 0.08, 0.11, 0.12 and 0.15 mm. Hg. The values of mn corresponding were 1.9, 2.8, 3.1 and 3.5 times 10^{-18}; these by the theory are the values of g for the corresponding pressures. They are roughly proportional to the pressure, as they should be. (Incidentally, the resonance-frequency of the system was shifted, when the ionization commenced, in the proper sense—the sense corresponding to a dropping of the dielectric constant below unity.)

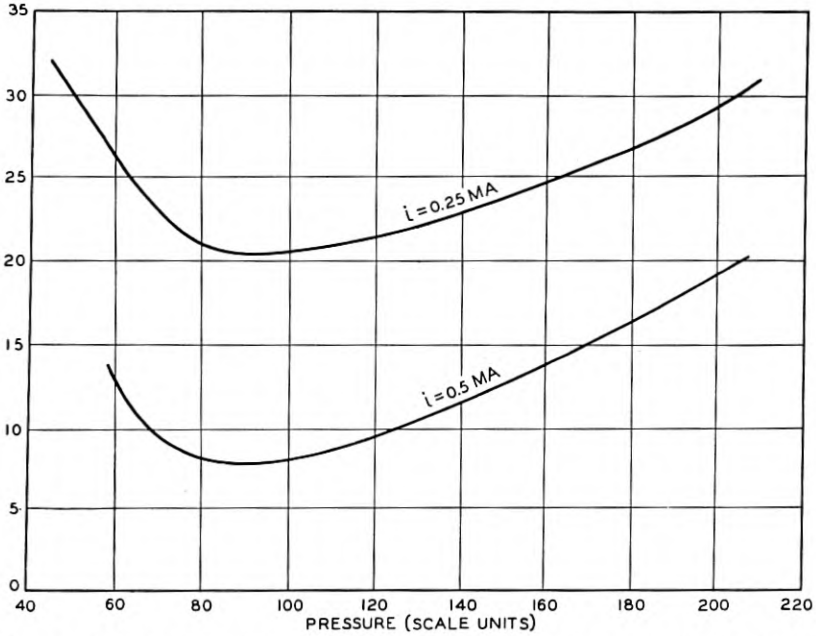


Fig. 6—Evidence of a maximum of conductivity occurring when the frequency is related in a particular way to the frictional coefficient g. (Appleton & Chapman.)

Before considering further data, let us compare these values of g with those deduced from observation (or theory) of electrons drifting through the same or a similar gas under a constant field. In such a condition, they attain a "terminal drift-speed" u given by the equation,

$$gu = eF, \tag{12}$$

F standing for the fieldstrength. Despite the aspect of the equation, u is not proportional to F; the coefficient g must be regarded as a function of F (over wide ranges of fieldstrength it is proportional to the square root thereof). Though this is a fact of experience, it will be helpful to develop the theory to some extent.

Picture the drifting of electrons through a gas in the customary (though far too primitive) way. Imagine the gas as a congregation of elastic spheres, with which the electrons make elastic impacts. When these latter enter the region where the constant and uniform field is applied, they are speeded up in the direction of the field; but owing to the deflections and the losses of energy which they suffer at their impacts, their velocities become and remain almost isotropic, and their average energy approaches but does not surpass a certain limiting value determined by the fieldstrength. After they have progressed sufficiently far, they form an "electron-gas" mingled with the atoms of the material gas; this electron-gas drifts slowly along towards the positive electrode, but its individual corpuscles have (as a rule) random velocities tremendously in excess of that comparatively modest drift-speed, just as the molecules of the air have random velocities many times greater than the speed of the wind. Let me denote the drift-speed by u, the mean speed of the random motions of the electrons by ω. Now it may easily be shown[6] that if the simple picture of the molecules as big elastic spheres and the electrons as little ones is acceptable, u and ω and the mean-free-path of the electrons l are related by the equation:

$$u = \theta eEl/m\omega, \tag{13}$$

θ here standing for a numerical factor not very different from unity, the exact value of which depends on the underlying assumptions made in the statistical analysis of the motions of molecules and electrons. The value 0.8 for θ is probably as good as any, but it would be pointless to spend time deciding between different values, since the molecules cannot be represented exactly as elastic spheres, and the degree of their deviation from this simple model would affect the quantity θ.

Now comparing the last two equations, we find:

$$\begin{aligned} g &= m\omega/\theta l \\ &= m/\theta\tau = mZ/\theta, \end{aligned} \tag{14}$$

τ and Z here representing respectively the mean duration of the free flight of an electron between consecutive impacts, and the number of impacts made by an electron in unit time. (I introduce these

---

[6] "Electrical Phenomena in Gases," pp. 174–175.

symbols because they are used in some of the original articles, not because they add anything to the fundamental ideas.)

The value of ω may be determined by the Langmuir probe-method; values for l, the mean-free-path of the electrons, are supplied by various methods of varying reliability. Childs, in his experiment which I have quoted, obtained with the probe the three values 1.3, 1.5, 1.7 times 10^8 (cm./sec.) for ω, these corresponding to the three cited values of the current maintaining the ionization; combining these with estimates [7] of l, he obtained for $m\omega/l$ the values 3.2, 3.8, $4.3 \cdot 10^{-18}$. Comparing these with the values found for g, and remembering the manifold chances of faults in the assumptions, one is favorably impressed with the agreement.

Could we not evaluate g directly, by measuring the drift-speed u of the electrons exposed to a certain constant fieldstrength F, and forming the ratio of eF to u to which, according to equation (12), g is equal? Here we must be careful. According to equation (13), g depends on the vivacity of the random motions of the molecules, whereof ω is the mean speed; now, ω depends on the fieldstrength: we must select such a value of F, that the random agitation of the electrons shall be the same as it is in the actual gas on which the high-frequency field is imposed. Now this actual gas is subjected to a constant field, that which maintains the ionization. It is natural and simple to assume, that this field controls the value of ω, the effect of the high-frequency field on the mean speed of random agitation of the electrons being presumably slight. The measurement of the drift-speed should therefore be made in the very gas under the very same constant field, in which the high-frequency phenomena are observed.

Let me denote by i the steady current along the length of the tube, due to the constant field; by N, the number of electrons per unit volume; by u, their drift-speed; by A, the cross-sectional area of the tube. It can then be readily seen that i is equal to $NeuA$.[8] The method is consequently simple in principle; in practice, the chief difficulty apparently is that N is not the same near the walls of the tube as along its axis, so that probe-measurements should be made at a number of distances from the axis and the results averaged. Appleton and Chapman determined N at a single point of the tube; the value of the ratio eF/u came out equal to $2.3 \cdot 10^{-18}$—again, a remark-

---

[7] Apparently he multiplied by $4\sqrt{2}$ the value of the mean-free-path of molecules of nitrogen, this being the simple gas presumably most like air; at any rate he used the value .036 cm.

[8] *Electrical Phenomena in Gases*, p. 207, pp. 232–233; on the two latter pages I describe Killian's application of the method to mercury vapor.

able agreement with the values deduced from the high-frequency phenomena![9]

(It is rather surprising to realize that in such discharges, we must not conceive the electrons as moving with a steady velocity along the axis of the tube, on which a sinusoidal oscillation parallel to the field is superposed. If we could follow the wanderings of an individual electron, the sinusoidal oscillation would be as little obvious as the steady drift. Only the rapid zigzag motions of the corpuscles would be conspicuous; what I have been calling a vibratory motion is, in truth, only a slight bias of these random flights, just as the apparent steady drift is itself a slight bias.)

Another experiment, capable in principle of testing the expressions for dielectric constant and conductivity and of giving the values of N and g, consists in sending electromagnetic waves of the frequency desired across a stratum of ionized gas, and measuring their index of refraction and their absorption in the stratum. The former of these two has not, so far as I know, been measured; but apparatus for determining the second (it is that of Hasselbeck's experiment) is shown in Fig. 7: one sees the paraffin lenses which convert a diverging beam of Hertzian waves into a parallel beam which is sent through the ionized gas, and others which reconvert the parallel beam into one which converges upon a bolometer. Part of the beam is reflected from a semi-transparent mirror onto another bolometer, so that the ratio of the intensities of the waves before and after the passage through the gas can be determined without regard to fluctuations.

The formulae for refractive index n_0 (I add the subscript because of having already used n, the conventional symbol, for another purpose) and absorption-coefficient k are familiar to everyone who has studied the theory of absorption and reflection of light by metals, for in the classical theory of metals such a substance is conceived exactly as we are now conceiving an electron-populated gas. We have:[10]

$$n_0^2 - k^2 = \epsilon, \qquad n_0 k = 2\pi\sigma/n, \tag{15}$$

and putting for ϵ and σ the expressions (7, 8), we get:

$$n_0^2 - k^2 = 1 - \frac{4\pi N e^2 m}{m^2 n^2 + g^2}, \qquad n_0 k = \frac{2\pi}{n}\frac{N e^2 g}{m^2 n^2 + g^2}, \tag{16}$$

solving which equations for k, and putting $m/\theta\tau$ for g (according to equation 14) and n_1^2 for the combination $4\pi N e^2/m$ (we shall meet it

---

[9] Similar observations by Jonescu and Mihul on air and on hydrogen, subjected in some cases to magnetic field, have recently been published, but with scant detail.

[10] See for instance P. Drude, "Treatise on Optics" (page 361 of the English translation).

later in another meaning), and α for the ratio n_1/n, we obtain:

$$k^2 = \frac{1}{2(1 + n^2\theta^2\tau^2)} \{\sqrt{[1 + n^2\theta^2\tau^2(1 - \alpha^2)]^2 + n^2\theta^2\tau^2\alpha^4} + [1 + n^2\theta^2\tau^2(1 - \alpha^2)]\}. \quad (17)$$

By measuring k with waves of a single frequency, and measuring in addition either of the two quantities N and g, it is possible to evaluate the other of the two by means of this equation; provided the equation is correct, a presumption which should be tested by making the measurements at several different frequencies.

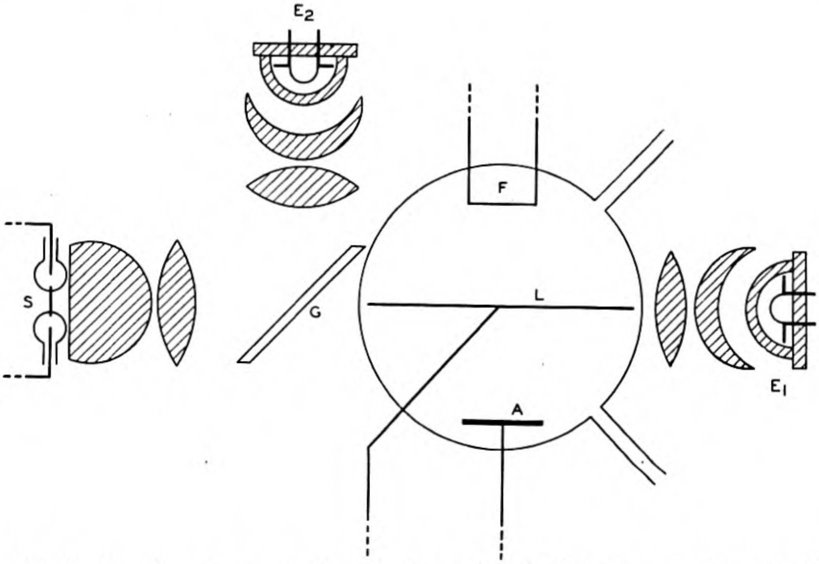


Fig. 7—Apparatus for measuring the absorption of Hertzian waves by ionized gas. *S*, source of waves; *G*, semitransparent mirror; E_1, E_2, receivers; *A*, *F*, anode and cathode of ionizing discharge; *L*, wire probe. (W. Hasselbeck; *Annalen der Physik.*)

Dänzer sent 4-cm. waves (of frequency $7.5 \cdot 10^9$)through a stratum of neon-helium mixture excited by a low-frequency discharge; measured the absorption, and measured also g by the method later employed by Appleton and Chapman. For N he then computed the value $1.3 \cdot 10^{12}$ (electrons per cc.)—a value of which, in the lack of further knowledge concerning electrons in this gas (he does not even state the pressure) one can only say that it is probably of the right order of magnitude. Much more extensive was the work of Hasselbeck, who used various frequencies ranging from 4.8 to 1.44 times 10^9, and introduced a probe into the gas in order to measure N and the mean speed ω of the electrons by Langmuir's method—it was a wire stretching clear across the globular tube, being made of this form so that it might give average values appropriate to the whole of the discharge. The best of his data refer to neon-helium mixture at various pressures of the order of a few tenths of a millimeter; there are data also for argon and nitrogen.

From equation (16), k should be proportional to N so long as n_0 does not depart too much from 1 and other things remain unchanged. Hasselbeck varied N by varying the current maintaining the ionization, measured it with the probe and measured the absorption. So long as the pressure of the gas was below a certain value, of slightly over 0.1 mm. Hg, the curves of k vs N were indeed straight lines rising from the origin (Fig. 8), though at higher pressures they were concave-upward. As for the curves of absorption *vs* frequency, some display maxima in the range over which they are plotted, some do not; but it is possible to suppose that each is characterized by a single maximum, which moves toward higher frequencies when the pressure or the ionization is augmented, and in some of Hasselbeck's experiments happened to be out of the range.

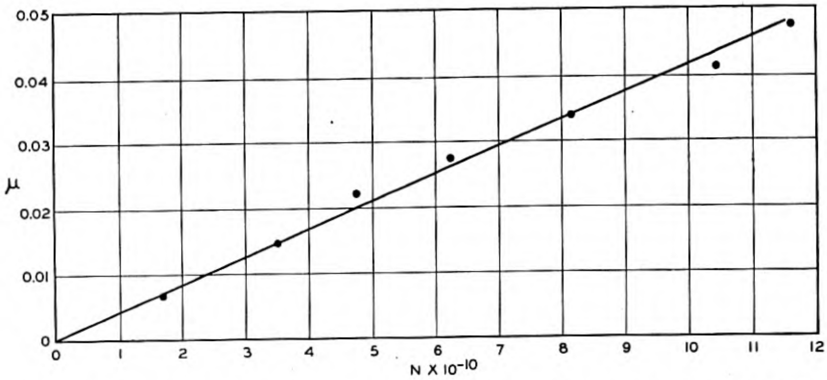


Fig. 8—Showing the proportionality of absorption of Hertzian waves to number of free electrons per cc. of absorbing gas (neon-helium mixture, pressure 0.18 mm. Hg). (W. Hasselbeck; *Annalen der Physik*.)

The comparison of the absorption-*vs.*-frequency curves with so complicated an expression as that of equation (17) is no simple matter, nor is Hasselbeck's description of his procedure entirely clear. It appears, however, that the curves obtained in neon-helium mixture at a pressure of 0.87 mm. Hg depart but little from the form of the expression on the right of (17), and that what departure there is may not be significant. However the curves obtained at the much higher pressure of 21 mm. Hg differ very seriously from the theoretical form.

Reverting to the question of testing equation (8) for the dielectric constant, that is, for the current in quadrature with the voltage: one might try the test with a gas-free space, populated by electrons derived from a hot filament or from some other source. This in fact has been tried, the electrons being shot across the interspace between the plates of the "ionization-condenser," in directions nearly parallel to the plane of these plates; from the (controllable) speed of these electrons, and from the charge which they bear per unit time to a collector located beyond the condenser, one may estimate the number-per-unit-volume heretofore denoted by N.

Since the density of the gas may easily be made so low that few electrons collide with even one atom between the condenser-plates, one expects the coefficient g to vanish. Strictly speaking, it does; and yet there is in effect a component of current in phase with voltage. The electrons, being pulled aside by the transverse high-frequency field as they traverse the condenser, acquire kinetic energy; and this absorption of energy from the high-frequency circuit produces the same reaction, and is measured in the same way, as an ordinary conductance. Benner, who developed these ideas, derived (with certain simplifying assumptions) these expressions for the coefficients ϵ and σ of a cloud of streaming electrons which individually take the time T to cross the condenser:

$$\epsilon = 1 - 4\pi \frac{Ne^2}{mn^2}\left(1 - \frac{\sin nT}{nT}\right), \qquad \sigma = \frac{Ne^2}{mn^2T}(1 - \cos nT), \quad (18)$$

and they have been tested with satisfactory results by Jonescu and Mihul.

We turn now to the possibility for which preparation was made in equations (10); that an electron in an ionized gas may experience, in addition to the force due to the applied field (eE) and the quasi frictional force due to the gas ($g\dot{x}$) yet a third force proportional to its distance from some fixed point. One is accustomed to postulate this last for electrons bound to molecules or atoms. (Such electrons, by the way, contribute to the current-component in quadrature with field strength, so that the dielectric constant of a gas in absence of ionization is not quite unity as I have been writing). But to imagine such a force acting on free electrons must seem strange—as if one were denying them the quality of freedom. Still, for a cloud of electrons mingled with positive ions there is such a force, and therefore a "natural frequency"; and another natural frequency in addition, if there happens to be a magnetic field.

Considering the former first: suppose that initially we have a uniform distribution of electrons, N per unit volume; and that at a certain moment it is suddenly distorted, by shifting every particle a distance ξ parallel to the x-direction, this distance being a function of the original position x of the particle. Fix attention on a column of unit cross-section; initially, between two planes x and $x + dx$, there were Ndx electrons; they suddenly move over and occupy the space between the planes $x + \xi$ and $x + dx + \xi + (d\xi/dx)dx$, so that the density between these two planes, or let me say the number of electrons per unit volume at $x + \xi$, is given by the formula:

$$N'_{x+\xi} = N\left(1 - \frac{d\xi}{dx}\right). \tag{19}$$

We now introduce Poisson's relation between net density ρ of electric charge and space-derivatives of field strength. As by assumption all shifts of electrons are parallel to the x-direction, so also is the field-vector; we put X for its magnitude, and write Poisson's equation thus:

$$dX/dx = 4\pi\rho. \tag{20}$$

If the electrons were the only charged particles, we should have to put $N'e$ for ρ; and there would be a field of strength different from zero and varying from place to place, even when the distribution of the electrons was uniform. If however there is also a uniform distribution of positive ions, N per unit volume, and this is not affected when that of the electrons becomes non-uniform, then for ρ we need set only the second term on the right-hand side of (19), multiplied by e; we get

$$dX/dx = -4\pi Ne(d\xi/dx), \tag{21}$$

and integrating, with the boundary-condition [11] $X = 0$ at $\xi = 0$,

$$X = -4\pi Ne\xi, \tag{22}$$

so that an electron shifted from its original location, by virtue of such a mass-distortion of the formerly uniform distribution, is indeed subjected to a restoring force proportionate to its shift.

Putting $m(d^2\xi/dt^2)/e$ for X in equation (22), we get:

$$d^2\xi/dt^2 = -(4\pi Ne^2/m)\xi = -n_1^2\xi, \tag{23}$$

showing that there is a tendency to oscillations—"plasma-electron oscillations," as Tonks and Langmuir call them—of frequency ν_1 thus given:

$$\nu_1 = n_1/2\pi = \sqrt{Ne^2/\pi m} = 8980\sqrt{N}. \tag{24}$$

---

[11] This seems to be demanded by symmetry if ξ is a sinusoidal function of x: otherwise the case is more obscure.

This expression is strictly valid only when the gas does not interfere at all with the motion of the electrons; otherwise the frequency is reduced, in the same way as the natural frequency of a pendulum or a circuit is lowered by damping. The equation of motion of the electrons in a high-frequency field is as follows:

$$m(d^2x/dt^2) + g(dx/dt) + fx = eE \sin nt; \qquad f = 4\pi Ne^2. \quad (25)$$

The solution has already been indicated (equations 10).

Attempts to discover this natural frequency have been made in two ways: by examining curves of σ or ϵ or other correlated quantities plotted against frequency or against degree of ionization, and by searching for electromagnetic waves due to oscillations arising of themselves in highly-ionized gases.

The most thorough experiments by the former way are due to Tonks. He placed a tube containing a mercury arc between condenser-plates attached to long parallel wires, these being crossed by a movable bridge including a thermocouple, and coupled to an oscillating circuit. After establishing fixed values of the frequency and the current-strength in the latter circuit, he shifted the bridge until the thermocouple reported a maximum of current, and measured this maximum value; it and the shift of the bridge (the zero from which this latter is measured is unimportant) were plotted as functions of the current-strength in the mercury arc, which controls the number of free electrons per cc. A natural frequency is indicated by a minimum in the former of the curves, and in the latter curve a peculiar crinkle, similar to that which appears in a dispersion-curve in the neighborhood of an absorption-frequency, and in the curve with black dots in Fig. 9.

Embarrassingly it turned out that there were two, and indeed sometimes three, minima in the one curve and crinkles in the other. To these the corresponding values of N were correlated, being obtained by the Langmuir probe-method. The comparison with theory may then be made in either of two ways: by putting the value of the applied high frequency into equation (24), computing N, and comparing it with the observed values of N at the minima; or alternatively, by putting the observed values of N into equation (24), computing ν_1, and comparing its values with that of the applied frequency ν. As an example of the result of the first procedure: in one experiment, the applied frequency was $1.59 \cdot 10^8$; this should coincide with the plasma-electron frequency, resonance should occur, when $N = 0.63 \cdot 10^9$ electrons per cc.; the values of N at the two minima which were observed were 0.27 and 0.86 times 10^9. The latter is illustrated by a graph in Tonks' article; it turns out that when the comparison is made at various frequencies between $1.59 \cdot 10^8$ and $3.66 \cdot 10^8$, the curves of ν_1 *vs.* N agree only fairly with those of ν *vs.* N; where ν_1 varies as the square root of N, ν varies as the powers 0.42 and 0.45 for the two minima regularly appearing.

Another instance of a crinkle in a curve of the second type aforesaid —a curve in which the shift of the bridge across the Lecher wires, necessary to restore resonance in the circuit which includes the ionization-condenser, is plotted against N or something varying with N—has been published by Appleton and Chapman. It refers to ionized air; the applied frequency is $3.75 \cdot 10^8$, the value of N is $1.62 \cdot 10^9$, and these two stand to one another in substantially the relation of equation (24). Yet other instances have been published by H. Gutton, and these offer difficulties.

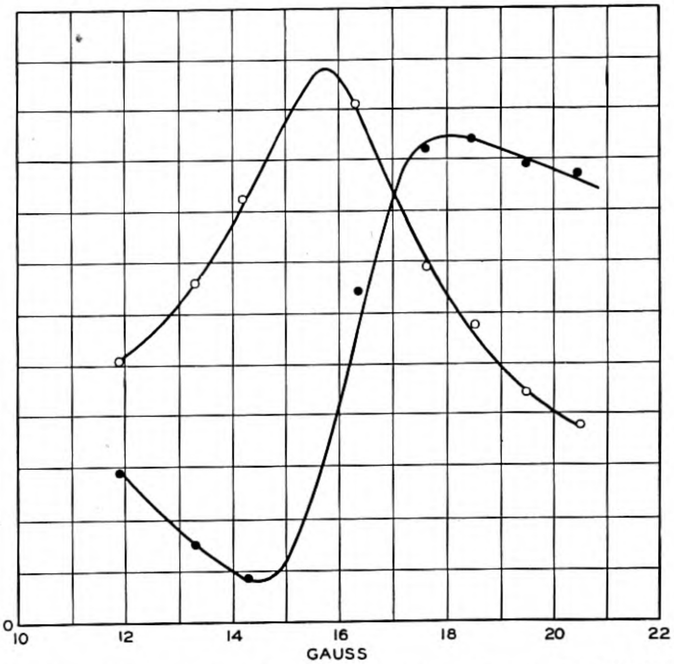


Fig. 9—Illustrating variation of dielectric constant (curve with full circles) and conductance (curve with hollow circles) of ionized gas in the neighborhood of a natural frequency, here due to a constant magnetic field. (S. Benner; *Naturwissenschaften.*)

Gutton in his later work (with ionized hydrogen at the low pressure of 0.0004 mm. Hg) followed the method which I have just described, measuring the shift y of the Lecher bridge required to restore resonance, and the mean-square current I^2 traversing the Lecher bridge at resonance. He did not measure N, but something probably (though not certainly) proportional to N, the factor of proportionality not known: the direct current i flowing across the gas between two probes inserted on opposite sides of the tube and maintained at a constant P.D. Plotting y and I^2 against i, he observed a crinkle in the former curve, a minimum in the latter—evidence of a natural frequency at a particular value of electron-concentration.

In his earlier work (which I mention because the curves are often reproduced) Gutton connected the plates of the ionization-condenser to one another through a thermocouple, and to this circuit coupled an oscillator of which the frequency could be varied. For each of a number of values of the ionizing-current in the discharge-tube he varied the frequency until resonance was declared by a maximum of the mean square of the current in the thermocouple; he measured this maximum I^2 and the wave-length λ of the oscillations. Not having any quantitative measure of the ionization against which to plot I^2 and λ, he plotted one against the other—I^2 as ordinate, λ as abscissa—and obtained curves of the curious appearance shown in Fig. 10, in which the arrow indicates the sense in which the ionization increases along each curve. The start is made from the wave-length (408 cm.) at which the system is in resonance when the gas is not ionized. At low pressure, the curve bends first to the left and downward; this signifies that as the ionization increases the dielectric constant of the gas is falling and the conductivity rising, as by the theory they should. Then at an unknown but seemingly sharply-marked value of ionization, the curve bends sharply to the right; and this signifies the same as the crinkle and the minimum in the other more-fully-comprehended curves which I have been discussing.

It is still dubious whether the natural frequency so revealed is that which the foregoing theory predicts. Neither in his earlier nor in his later experiments did Gutton measure N (the estimate of its value which he once makes is derived in an indirect and fallible way). Measuring the values of i at which the resonance appeared in his later work, and comparing them with the corresponding values of the frequency, he found ν^2 proportional to $i^{3/4}$; the range of frequencies was comparatively small (not quite 4 : 1) but if the result is certain and i is truly proportional to N, equation (24) is contradicted. Appleton and Chapman testify that they found the resonance-effect observed by Gutton, but at values of N (which they measured by the Langmuir method) entirely too great to permit of regarding it as the plasma-electron frequency, from which they believe it to be distinct.

As for the search for spontaneous oscillations probably having the proper plasma-electron frequency, it is at present even less advanced.

Electromagnetic waves of high frequency often emanate from gases intensely ionized by a flow of direct current, and their wave-lengths have several times been measured, but for one reason or another the comparison with equation (24) was not and cannot be made. In a mercury-vapor arc Tonks and Langmuir detected oscillations of frequencies nearly as high as 10^9, accompanied by others ranging downwards as low as 10^6; the former, they believe, were plasma-electron vibrations; but unluckily they were unable to estimate N with any degree of exactness.

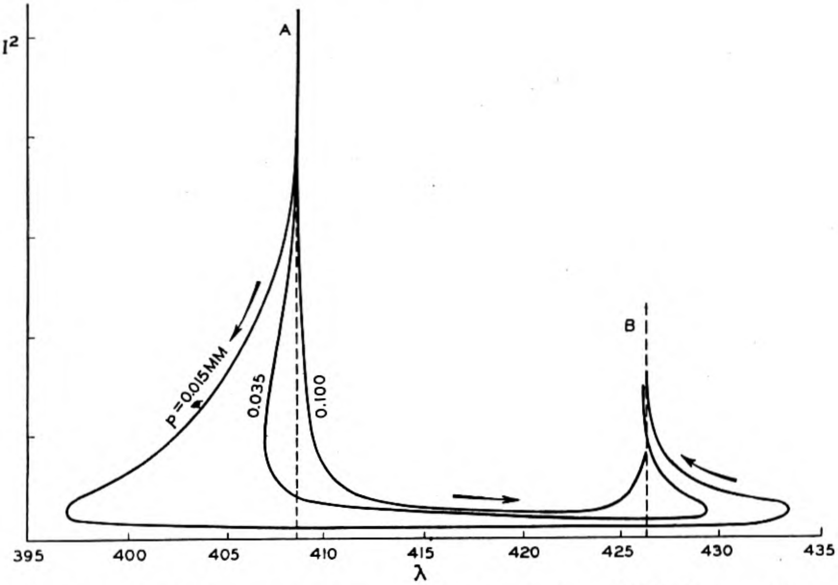


Fig. 10—Correlation of quantities serving as measures of high-frequency conductance and dielectric constant of rarefied ionized hydrogen. (H. Gutton; *Comptes Rendus.*)

If a region populated with free electrons is pervaded by a constant magnetic field, yet another natural frequency exists. Say, to begin with, that there is no electric field; then any moving electron, instead of continuing in a straight line, is constrained to describe a spiral path (the axis of the spiral being parallel to the magnetic field). The velocity of the electron affects the curvature of the helix, *but not the time of traversing a single winding* thereof. The reciprocal of this time, the number of windings described by any moving electron per second, is the "natural frequency" aforesaid, and is given by the formula:

$$\nu_H = n_H/2\pi = eH/2\pi mc, \qquad (26)$$

H standing as usual for the magnetic field strength. Since this frequency remains the same however much the speed of the electron may change, and therefore is the same for all the electrons in the region in question and for all values of electric field strength, we may expect it to be important when an ionized gas (or a volume containing free electrons but no atoms) is exposed to a high-frequency field; in curves of dielectric constant and conductance vs. frequency, we may look for peculiarities when $\nu = \nu_H$. The precise theory, I must add, is not simple when collisions of electrons with atoms must be taken into account (the fundamental equations were given long ago by Lorentz); but under certain restrictions—according to Appleton and Childs, the number of electrons per cc. and the number of collisions per cycle of the high frequency must be kept under certain limits—it leads to the inference that there should be a maximum of conductance at $\nu = \nu_H$.

This maximum is manifested by the sharp and striking minimum seen near the middle of the curve in Fig. 11. For obtaining these data, Appleton and Childs had an arrangement similar in the main to that of Fig. 5, excepting that the detecting galvanometer was coupled (through an amplifier) between the far ends of the wires of the bridge, and there was a magnetic field of adjustable strength parallel to the axis of the tube. Though I have spoken thus far of what should be observed when H is held constant and ν is varied through the value given by equation (26), it is more convenient in practice to hold the frequency constant and vary H through the corresponding value. The curve of Fig. 11 is accordingly a curve of galvanometer-reading

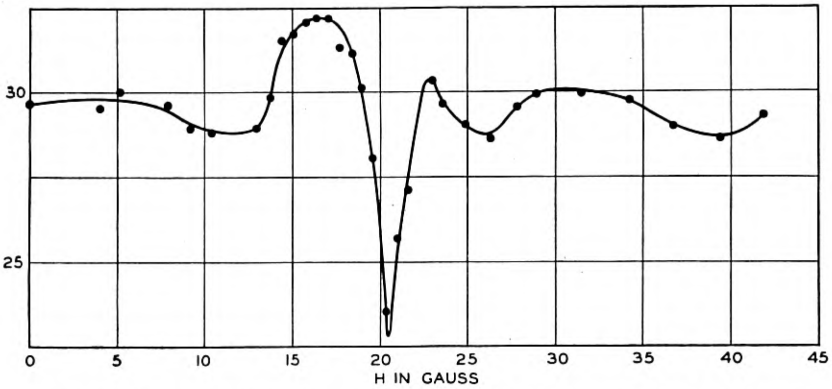


Fig. 11—Evidence of a natural frequency produced in an ionized gas by a constant magnetic field. (Appleton & Chapman.)

vs. H; the remarkable minimum occurs at a value of H departing by less than 2 per cent from $2\pi mc\nu/e$. These data were obtained at frequency $5.46 \cdot 10^8$; others observed at $2.96 \cdot 10^8$ and $3.45 \cdot 10^8$ yielded agreements almost as good, or better.

Maxima of conductivity of ionized air have lately been found by Jonescu and Mihul, at the predicted values of field strength H, for various frequencies of the order 10^8. Earlier Benner, in a note deplorably brief, showed not only a curve of conductivity (or rather, of something proportional to conductance) displaying a maximum, but a curve of dielectric constant displaying a crinkle like that of a dispersion-curve in the neighborhood of a region of anomalous dispersion. These are the curves of Fig. 9, already introduced into this article to illustrate how these quantities vary in the vicinity of a natural frequency. The "ionization-condenser" of Benner's experiment consisted of the grid and plate of a triode, the space between them populated with electrons emitted from the filament; it is to be inferred that the tube contained some gas, but unfortunately nothing definite is said about the kind or amount. The maximum of the one curve and the crinkle in the other occurred at a value of H some twelve per cent higher than the predicted value aforesaid.

On applying a longitudinal magnetic field of about 21 gauss to the tube of ionized hydrogen with which he had observed the peculiar natural frequency mentioned above, H. Gutton found this one replaced by two, well marked and well separated, one being shifted toward higher frequencies from the original value and one toward lower. The same phenomenon has been observed by Tonks, in his studies of the natural frequencies which he identifies with the plasma-electron oscillations predicted by the theory culminating in equation (24). The doubling of the resonance is analogous to the Zeeman effect, and is amenable to theory.

The earth's magnetic field imprints a natural frequency upon the electrons populating the air, in particular the upper strata thereof; this affects the transmission of radio waves in curious ways, which were foretold in this journal seven years ago by Nichols and Schelleng, and in England by Appleton and Barnett.

In the second part of this article, I will treat of the conditions under which a high-frequency field may initiate and maintain a luminous discharge in a gas, and of the laws of these discharges.

REFERENCES

The theory underlying equations (1) to (4) is well known; it is given in considerable detail by G. Mierdel, *Ann. d. Phys.*, **85**, 612–640 (1928), with some illustrative curves, and in several other papers. The theory underlying equations (5) to (10) apparently originated, like so much else in electronics, with H. A. Lorentz; it was developed with especial reference to the transmission of radio waves by W. H. Eccles, *Proc. Roy. Soc.*, **A 87**, 79–99 (1912); later enunciations, with various modifications, may be found in the papers of J. Salpeter, *Phys. ZS.*, **14**, 201–203 (1913); E. V. Appleton & F. W. Chapman, *Proc. Phys. Soc. Lond.*, **44**, 246–254 (1932), and in several papers cited in the next paragraph.

Experiments designed to measure dielectric constant and conductivity of ionized gas for high-frequency fields have been performed by A. Szekely, *Ann. d. Phys.* (5), **3**, 112–132 (1929); E. V. Appleton & E. C. Childs, *Phil. Mag.* (7), **10**, 969–994 (1930); Appleton & Chapman, *l.c. supra;* E. C. Childs, *Phil. Mag.* (7), **13**, 873–887 (1932); H. Gutton, *Annales de physique* (10), **13**, 62–129 (1930); H. Gutton & J. Clement, *C. R.*, **184**, 441–443, 676–678 (1927); V. Jonescu & C. Mihul, *C. R.*, **192**, 343–345 (1931.)

Experiments on the absorption of Hertzian waves by ionized gases: H. Dänzer, *Ann. d. Phys.* (5), **2**, 27–62 (1929); W. Hasselbeck, *ibid.* (5), **12**, 477–502 (1932).

Experiments on the reaction of electron-streams projected through very rarefied gas to a high-frequency field: L. Bergmann & W. Düring, *Ann. d. Phys.* (5), **1**, 1041–1068 (1929); S. Benner, *ibid.* (5), **3**, 993–996 (1929) (theory); V. Jonescu & C. Mihul, *C. R.*, **191**, 1436–1438 (1930).

Equation (24) for the so-called "plasma-electron" natural frequency appears to have been first derived by J. J. Thomson, and shortly afterward by L. Tonks & I. Langmuir, *Phys. Rev.* (2), **33**, 195–210, 990 (1929); see also Thomson, *Phil. Mag.* (7), **11**, 697–735 (1931). Experiments on this, and natural frequencies of possibly different origin: L. Tonks, *Phys. Rev.* (2), **37**, 1458–1483 (1931); H. Gutton, *Annales de physique* (10), **13**, 62–129 (1930); C. Gutton, *ibid.* (10), **14**, 5–14 (1930); Appleton & Chapman, *l.c. supra.*

Natural frequency or frequencies appearing when a constant magnetic field is imposed upon the gas: S. Benner, *Naturwiss.* **17**, 120–121 (1929); Appleton & Chapman, H. Gutton, L. Tonks, *ll.cc. supra.* The influence, upon the transmission of radio waves, of the magnetic field of the earth, was inferred independently in America by H. W. Nichols & J. C. Schelleng, *Bell Syst. Tech. Jl.*, **4**, 215–234 (1925) and in England by E. V. Appleton & M. A. F. Barnett, *Electrician*, **94**, 398 (1925).

Transformer Coupling Circuits for High-Frequency Amplifiers

By A. J. CHRISTOPHER

This article deals with the use of transformer type of coupling circuits in high-frequency amplifiers to transmit efficiently voltages or currents between certain limiting frequencies while attenuating those above and below the limiting frequencies. The similarity of these coupling circuits to band-pass filters is shown and the conditions to be satisfied in order that they may act as such are covered. Means of obtaining uniformly high amplification over relatively wide frequency bands are explained. Typical conditions under which these coupling circuits have been employed and factors affecting their performance are discussed.

I. Introduction

THE designer of high-frequency amplifiers is often confronted with the problem of obtaining, with a given number of amplifying tubes and coupling circuits in cascade, maximum voltage amplification over a predetermined band of frequencies, and high attenuation to all voltages outside of the desired band of frequencies. To secure a large voltage amplification the most convenient, economical and practical arrangement for coupling the various stages of the amplifier is by means of the step-up transformer. By adding condensers in parallel with one or more of the windings of the transformer a frequency discrimination characteristic is obtained which can be controlled to a large extent by the proper choice of the transformer constants and the tuning capacitances. In the case of the usual type of transformer coupling with the secondary winding tuned to resonance at a given frequency the voltage amplification for a single stage depends on the resistance of the secondary winding, the conductance of the grid circuit of the second tube and the size of the tuning condenser. With proper choice of the transformer constants very large amplification can generally be obtained over a relatively narrow band of frequencies. To obtain high amplifications over wider frequency bands other factors must be taken into consideration.

Possibly the most important of these factors is the impedance of the circuit into which the secondary winding operates. In the case of an interstage transformer with the secondary winding connected directly to the grid circuit of a three-element tube this impedance depends on the electrode capacities of this tube, the amplification factor, the plate impedance and the impedance connected to the output terminals of the tube. Where the amplification factor and plate circuit terminating impedance are low, as in the case of tubes operating as demodulators, the effective input impedance of the tube is comparatively high and depends mainly on the electrode capacities. The input impedance of a shielded-grid tube is likewise comparatively high and depends mainly on the capacity between the grid and filament terminals. The input impedance of a tube decreases with increase in frequency and at very high frequencies it becomes so low that difficulty is experienced in obtaining any increase in amplification in the coupling circuit. In the case of the single tuned transformer, previously referred to, the effective input capacity of the tube may act as part of the tuning condenser with the result that at the resonance frequency a substantial amplification can usually be obtained. However, when it is desired to transmit a broad band of frequencies the magnitude of the effective input capacity becomes a very important controlling factor. For a definite band width the possible voltage amplification varies inversely as the value of the effective capacity across the secondary winding.

A study of the various factors entering into the use of a suitable coupling circuit for successive stages of a high frequency amplifier has indicated that best results are obtained if the design of the transformer is based upon the principles of the broad-band filter. Definite relations are obtained between the constants of the transformer windings, the tuning capacities, and the impedances of the circuits between which it operates. The structure is thus essentially a band-pass filter and has all of its elements properly proportioned to provide the desired band selectivity, but it still retains the form and the functions of a transformer.

II. Theory and Method

Fig. 1 shows a simple transformer type of coupling circuit connected between the plate circuit of an amplifying tube and the input circuit of a second tube. The primary and secondary circuits of the trans-

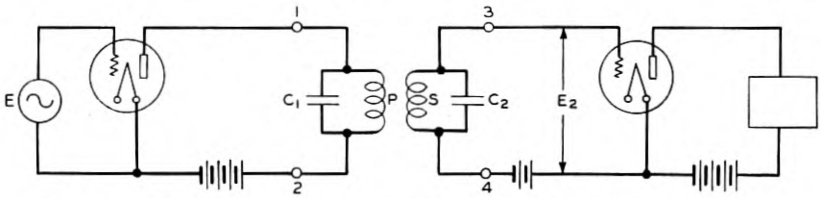


Fig. 1—Circuit schematic of an amplifier using a transformer type of coupling circuit.

former are adjusted to resonate at the same frequency. The determination of the constants for the circuit so that it conforms not only to the requirements relating to the band selectivity and to the voltage transformation ratio, but also to requirements of transmission efficiency, is explained in the following mathematical analysis.

For convenience in analyzing the essential parts of Fig. 1 they are shown in a simplified schematic form in Fig. 2. In this schematic, the sending-end impedance Z_0 corresponds to the effective plate impedance of the first vacuum tube and the electromotive force E_1 corresponds to the internal plate voltage of this tube. The transformer formed by inductances P and S is replaced by the well known "T" network.[1] The capacity C_1 is provided by a separate condenser, while that of C_2 is equal to the effective distributed capacity of winding S plus the effective input capacity of the second vacuum tube.

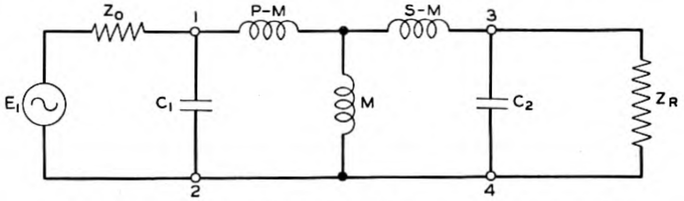


Fig. 2—"T" network of a transformer coupling circuit.

The method used in computing the output voltage of a transmission circuit such as is shown in Fig. 2 consists in determing the "image" impedances corresponding to each pair of terminals. These image impedances are determined from the open and short-circuit impedances, measured at the terminals of the network,[2] by the following relationships:

$$Z_{1,\,2} = \sqrt{Z_0'Z_S'}, \tag{1}$$

and

$$Z_{3,\,4} = \sqrt{Z_0''Z_S''}, \tag{2}$$

where Z_0' and Z_S' equal respectively the open and short-circuit impedances and $Z_{1,\,2}$ the image impedance at terminals 1, 2 and Z_0'' and Z_S'' equal the open and short-circuit impedances and $Z_{3,\,4}$ the image impedance at terminals 3, 4.

Since the resistance elements of a transformer of this type can be made very small in comparison to the reactance elements, it is practical to eliminate them in computing the open and short-circuit impedances.

---

[1] "Telephone Transformers," by W. L. Casper, *A. I. E. E. Jour.*, March, 1924, Vol. XLIII, No. 3.

[2] "Transmission Characteristics of Electrical Wave Filters," by O. J. Zobel, *Bell Sys. Tech. Jour.*, October, 1924, Vol. III, No. 4.

With the assumption that the elements of the coupling circuit are pure reactances and that the two anti-resonant circuits are at resonance at the same frequency, f_0, the following expression for $Z_{1,\,2}$ as a function of frequency f was derived with the use of equations 1 and 2.

$$Z_{1,\,2} = \frac{1}{2\pi f C_1}\sqrt{\frac{1 - K^2}{\left(1 - K - \frac{f_0^2}{f^2}\right)\left(\frac{f_0^2}{f^2} - 1 - K\right)}} \tag{3}$$

and $Z_{3,\,4}$ is the same expression with C_2 substituted for C_1 or

$$Z_{3,\,4} = Z_{1,\,2} \times \frac{C_1}{C_2} = Z_{1,\,2}\frac{S}{P},$$

where K is the coefficient of coupling between inductances P and S, and f_0 is the common resonance frequency of the two anti-resonant circuits in Fig. 1.

The lower and upper cut-off frequencies f_1 and f_2 are related to the resonance frequency f_0 as follows:

$$\begin{aligned} f_0 &= f_1\sqrt{1 + K} \\ &= f_2\sqrt{1 - K}. \end{aligned} \tag{4}$$

The geometrical mean frequency

$$\begin{aligned} f_m &= \sqrt{f_1 f_2} \\ &= \frac{f_0}{\sqrt[4]{1 - K^2}}. \end{aligned}$$

Substituting the above expressions for f_m in place of f in equation (3), we find that at the geometrical mean frequency the image impedance in terms of f_1 and f_2 has the value

$$\frac{1}{2\pi(f_2 - f_1)C_1}, \tag{5}$$

which will be denoted by $Z_{1,\,2}$.

It will be noted upon examining equation (3) that $Z_{1,\,2}$ and $Z_{3,\,4}$ are resistive between f_1 and f_2 and are minimum at the geometrical mean frequency. They increase uniformly on either side of the geometrical mean frequency and are infinite at f_1 and f_2. For frequencies below f_1 and above f_2 they are reactive. It can therefore be seen that this type of coupling circuit inherently possesses the characteristics of a band-pass filter. However, in order that the band characteristic may be properly developed, the magnitudes of the various elements must be proportioned with respect to the terminal impedances as well as to the limiting frequencies. This is done by making the image impedances about equal to the respective terminal impedances at the mean frequency f_m of the band. Preferably the impedances should be matched at both ends, but if the transformer normally operates with one end substantially open-circuited, as in an amplifier, it is sufficient to effect the matching at the other end. The following formulas giving the constants of the various elements in terms of the limiting frequencies were derived from the foregoing equations.

$$K = \frac{f_2^2 - f_1^2}{f_2^2 + f_1^2}, \tag{6}$$

$$f_0 = \frac{\sqrt{2} f_1 f_2}{\sqrt{f_1^2 + f_2^2}}, \tag{7}$$

$$C_1 = \frac{1}{2\pi(f_2 - f_1)\, Z_{1,\,2}}, \tag{8}$$

$$P = \frac{1}{4\pi^2 f_0^2 C_1}, \tag{9}$$

$$S = \frac{1}{4\pi^2 f_0^2 C_2}, \tag{10}$$

or

$$S = P \times \frac{C_1}{C_2}. \tag{11}$$

As previously mentioned the above equations are based on the ideal condition which assumes the elements to be pure reactances. Actually, a small amount of resistance is present in each element of the transformer which tends to reduce the width of the transmission band. Consequently f_1 and f_2 should be assumed slightly lower and higher respectively than the lower and upper frequencies of the desired transmitted band in order to insure uniform voltage amplification. It is advantageous to make the impedance $Z_{1,\,2}$ approximately 0.8 of the terminal impedance Z_0, in which case $Z_{1,\,2}$ and Z_0 will be equal at two frequencies near the band limits and will not be greatly mismatched at the geometrical mean frequency. This tends to improve the uniformity of transmission within the band.

In applying the above equations, C_1 is determined first from equation (8). P is then determined from equation (9). From a knowledge of the effective distributed capacity of S and the input capacity of the second tube in Fig. 1, S is determined from equation (10). The windings P and S are then arranged with respect to each other to satisfy the coupling coefficient of equation (6). To facilitate adjustment C_2 may include a small adjustable condenser to compensate for variations in winding capacities as well as in the input capacity of the vacuum tube.

The following analysis shows that maximum voltage amplification is obtained with practically uniform transmission for all frequencies within the transmitted band when the secondary circuit is terminated only in its tuning condenser.

Referring to Fig. 2 and assuming that $Z_{1,\,2} = Z_0$ and $Z_R = Z_{3,\,4}$, we find that the voltage drop across Z_R is

$$E_R = \frac{1}{2} E_1 \sqrt{\frac{Z_{3,\,4}}{Z_{1,\,2}}}$$

or

$$E_R = \frac{E\mu}{2} \sqrt{\frac{Z_{3,\,4}}{Z_{1,\,2}}} \,. \tag{12}$$

Now from Thevenin's Theorem,[3] assuming $Z_0 = Z_{1,\,2}$, we can obtain the following expression for the output current for any value of the impedance Z_R:

$$I_r = \frac{E_2}{Z_R + Z_{3,\,4}}, \tag{13}$$

where E_2 is the open-circuit voltage at terminal 3, 4 and I_r is the current flowing in the terminating impedance Z_R.

The actual voltage across the terminating impedance Z_R is

$$E_R = I_R Z_R = \frac{E_2 Z_R}{Z_R + Z_{3,\,4}} \,.$$

Where Z_R is equal to $Z_{3,\,4}$ the output voltage has the value

$$E_R = \frac{E_2}{2} \,.$$

The effect of matching the terminating impedance to the impedance $Z_{3,\,4}$ is thus to cut the output voltage in half.

Then

$$\frac{E_2}{E} = \mu \sqrt{\frac{Z_{3,\,4}}{Z_{1,\,2}}} = \mu \sqrt{\frac{C_1}{C_2}}, \tag{14}$$

where μ = voltage amplification constant of the first tube.

Generally the effect of leaving the output end of a filter open-

---

[3] "Transmission Circuits for Telephone Communication," by K. S. Johnson, p. 79.

circuited would be to introduce irregularities of transmission within the band. However, if the band is relatively narrow the resistances of the filter elements are sufficient to smooth out these irregularities and in certain cases the transmission may be made more uniform by the omission of the impedance Z_R. This has been found to be the case with transformers of the type described here.

Equation (14) shows that the voltage ratio E_2/E which is the ratio of the open-circuit voltage of the tuned circuit at terminals 3, 4 to the applied grid voltage of the first tube is directly proportional to $\sqrt{\frac{C_1}{C_2}}$ over the transmitted band. By proper choice of the constants of the coupling circuit so that $Z_{1,2}$ will equal Z_0 at a frequency lower and higher than the geometrical mean frequency this voltage ratio will remain practically constant for all frequencies within the band.

The voltage ratio E_2/E for any frequency ω outside the transmission band may be obtained from the following equation:

$$E_2/E = u\left[\frac{Z_1}{Z_2} \times \frac{Z_3}{Z_4} \times Z_5/Z_6\right],$$

where

$$Z_1 = \frac{1}{j\omega C_2}; \qquad Z_2 = j\omega(S - M) + Z_1; \qquad Z_3 = \frac{Z_2 j\omega M}{Z_2 + j\omega M};$$

$$Z_4 = Z_3 + j\omega(P - M); \qquad Z_5 = \frac{Z_4 \frac{1}{j\omega C_1}}{Z_4 + \frac{1}{j\omega C_1}}; \qquad Z_6 = Z_5 + Z_0.$$

P, S and M are in henrys and C_1 and C_2 in farads.

As the effective voltage amplification of the network within the transmission band is directly proportional to

$$\sqrt{\frac{C_1}{C_2}}$$

it is evident that since C_1 is fixed by equation (8), the value of C_2 must be kept to a minimum to realize maximum voltage step-up. Consequently, greater amplification will be obtained if the second tube in Fig. 1 has minimum input capacitance. The effective distributed capacitance of the winding is kept at a low value by placing the winding in narrow grooves on a spool.

In actual use the maximum amplification realized at the higher frequencies is somewhat less than that shown in equation (14) since the conductance component of the input impedance of the second tube reduces the effective voltage at this point. It is therefore of importance that the conductance component as well as the capacitance component of the input impedance of the second tube be kept to a minimum if maximum amplification is to be obtained. If the second tube in Fig. 1 is operated as a negative grid bias detector the ratio of E_2/E may be made considerably larger than if it is operated as an amplifier. The application of a large negative grid bias and the use of a by-pass condenser in the plate circuit of the tube to improve its modulation efficiency usually results in a very small input capacitance and conductance. It is therefore sometimes desirable in multistage amplifiers to place the selective circuits between the last amplifier tube and the detector tube. With the advent of the shielded-grid tube having the characteristic of low input capacitance and conductance, transformers of this type are particularly adapted to high frequency amplifiers such as used for the intermediate frequency amplifier of a superheterodyne receiver.

III. Practical Applications

Fig. 3 shows the transmission characteristic of a simple transformer coupling circuit when operating from a balanced resistance of 1000 ohms into the grid circuits of push-pull amplifying tubes. Stability was more important in this particular case than maximum voltage step-up so that a fixed condenser C_2 was added externally across the

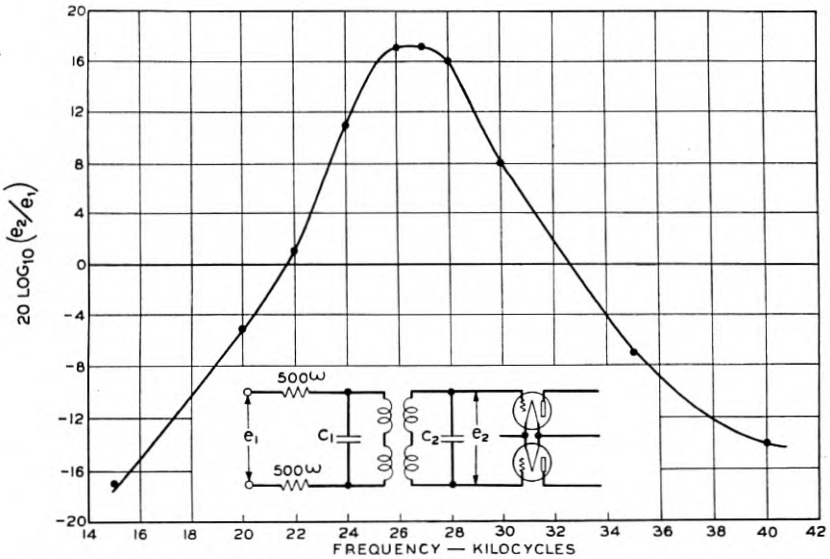


Fig. 3—Typical transmission characteristic of a transformer coupling circuit.

grid circuit to prevent the operation of the transformer from being affected by variations in the grid circuit capacities.

More than one coupling circuit has been used to obtain the necessary selectivity. Several of these have been connected together with either series or shunt condensers to obtain the equivalent of a multi-section band-pass filter. Networks of this type in which the elements have been chosen in accordance with the previously mentioned equations have been employed as band filters because of their simplicity and compactness, and the relatively low cost of the inductance elements. They have been used to connect two equal or unequal impedances as well as to operate from an impedance directly into the grid circuit of a vacuum tube.

One of these networks was used between the plate circuit of a shielded-grid vacuum tube and the grid circuit of a second shielded-grid vacuum tube at 84 kilocycles. The gain characteristic of a stage of this type is shown in Fig. 4. The capacitance C_1 represents the capacitance between the plate and shield of the first tube plus the winding capacitance and capacitance C_2 represents the winding capacitance and effective input capacitance of the second tube. It will be of interest to know that the voltage amplification obtained over the transmitted band was approximately equal to the amplification of the first shielded-grid tube. Consequently, the transformer network

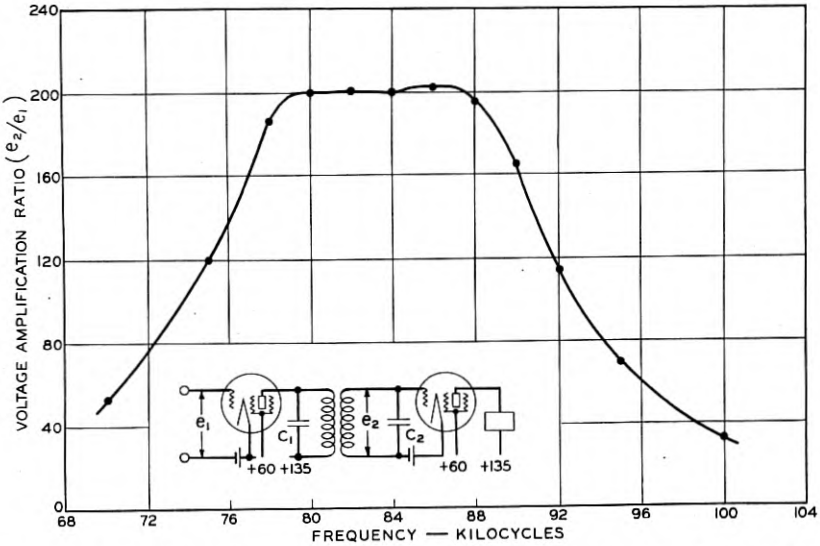


Fig. 4—Transmission characteristic of a transformer coupling circuit operating between shielded-grid tubes.

efficiently transmitted the internal plate voltage of the first tube to the grid circuit of the second tube. The internal plate impedance of the particular tube used was approximately 400,000 ohms. A two-stage amplifier consisting of two shielded-grid amplifying tubes and two transformer coupling circuits connected in cascade gave a voltage amplification of approximately 40,000 times over the frequency band shown.

Another type of transformer coupling circuit which was employed in the intermediate frequency amplifier of a high quality superheterodyne radio receiver is shown in Fig. 5. The circuit schematic of this transformer and its transmission characteristic are shown in Fig. 6. It will be noted that two transformers are mounted separately and electrically connected by a series condenser C_2'. C_1 and L_1 were determined from equations (8) and (9) and $L_1 = L_2 = L_3$. $C_2' = \frac{C_1}{2}$. The capacitance across L_4 was equal to the winding capacity and effective input capacity of the second tube. The second tube of Fig. 6 was the second detector of the intermediate frequency amplifier.

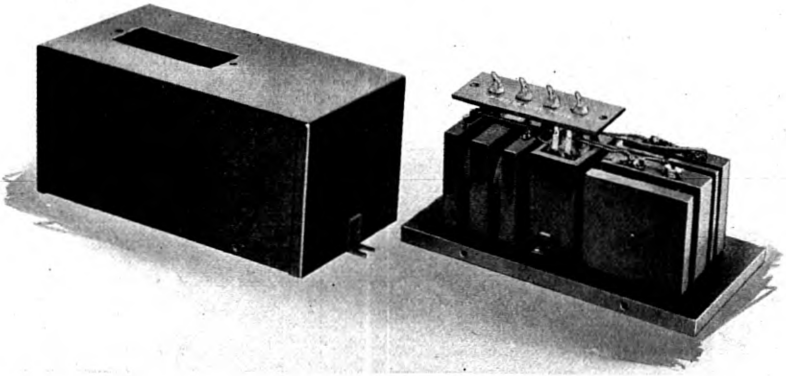


Fig. 5—Coupling circuit consisting of two-tuned transformers connected in cascade.

The elements of a coupling circuit consisting of three transformers and their associated condensers for operation over the carrier frequency range of 50 to 150 kilocycles are shown in Fig. 7. The windings of stranded wire are applied in narrow grooves to reduce the dielectric losses of the insulation between layers. The ratio of the reactance to the effective resistance for these coils varies from approximately 150 at 50 kilocycles to approximately 240 at 150 kilocycles. The halves of the winding connected to the grid circuit of the balanced tubes

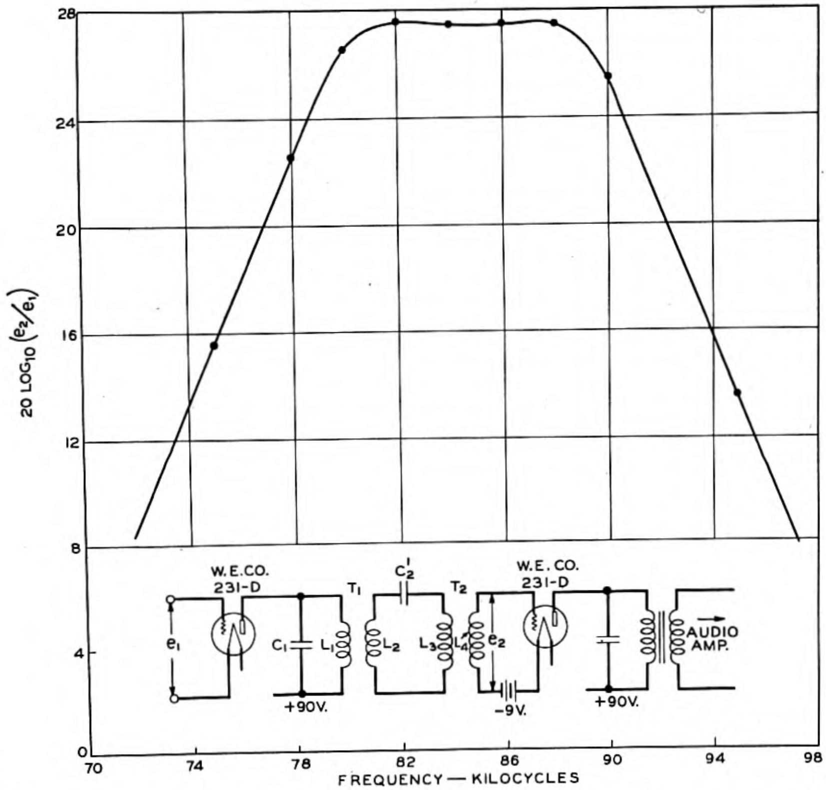


Fig. 6—Transmission characteristic of a double transformer type of tuned circuit.

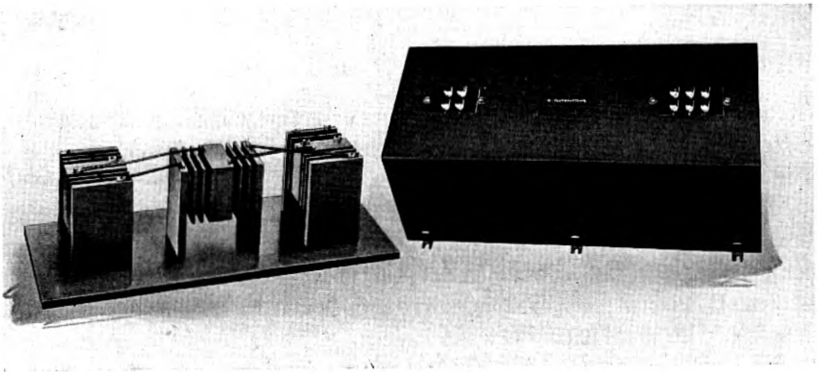


Fig. 7—Coupling circuit consisting of three-tuned transformers in cascade.

were wound in the two grooves of the last spool type assembly in such a manner as to maintain satisfactory balance for this type of operation. Four adjustable condensers were added externally to these transformers as shown in Figs. 8 and 9. The capacity of the condensers was adjustable so that the transmitted band could be located anywhere between 50 and 150 kilocycles.

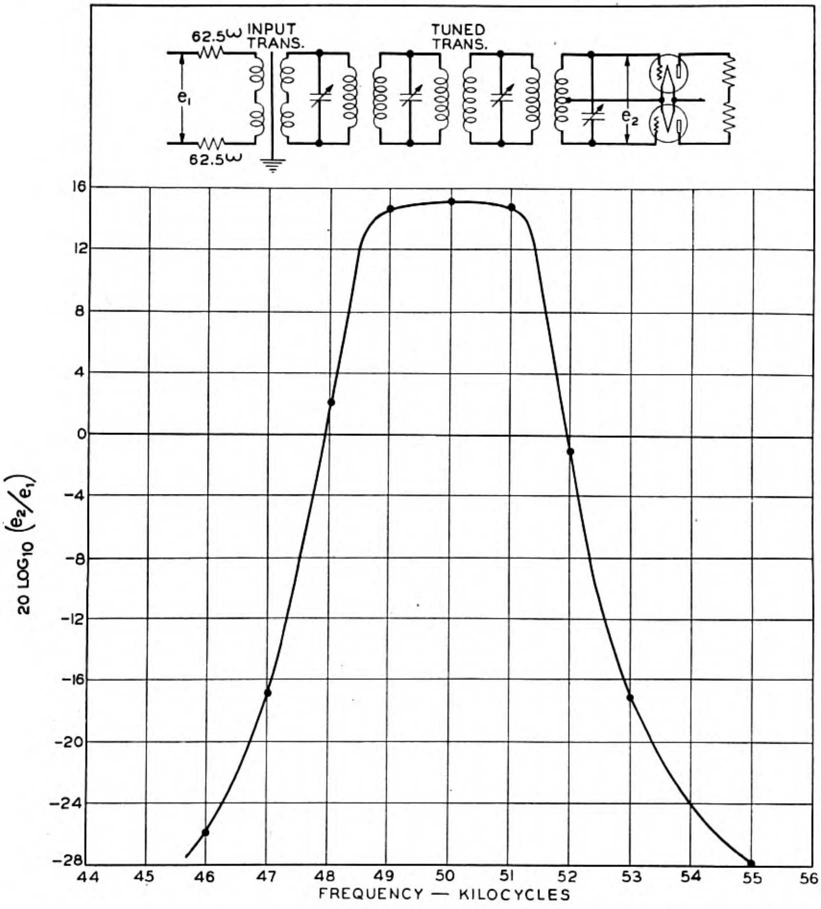


Fig. 8—Transmission characteristic of a triple transformer type of tuned circuit.

Fig. 8 shows the circuit schematic and transmission characteristic of the same transformer network at 50 kilocycles when operating from a non-inductive resistance into the grid circuit of push-pull modulator tubes. A step-up transformer was used between the resistance and the tuned transformer to match impedances and improve the efficiency of transmission. Fig. 9 shows the transmission characteristic and circuit schematic of the same tuned transformer when operating between two non-inductive resistances. Repeating coils were used

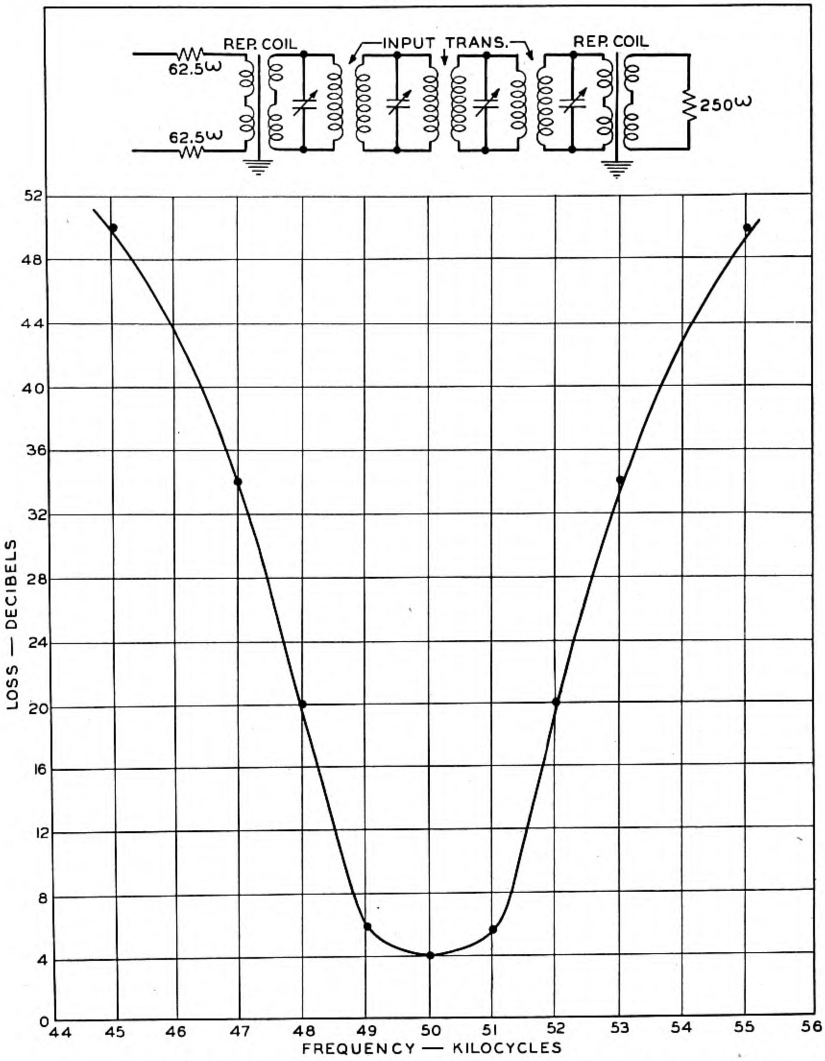


Fig. 9—Transmission loss characteristic of a triple transformer type of tuned circuit.

between the transformer and the terminated resistances to match impedances. Although the width of the transmission band at 50 kilocycles is only 2500 cycles it is considerably wider at 150 kilocycles as one would expect from the equation showing the relation of band width and capacity. This, however, was not an objectionable feature in the circuit in which the transformer was employed. In order to maintain a constant band width irrespective of its location, the capacity must remain constant and the self impedance of the windings and the coupling coefficients changed or the inductance elements maintained at a constant value and the capacities and coupling coefficients changed.

IV. Conclusions

Transformer types of coupling circuits having the inductive and capacitive elements proportioned as explained in this paper are essentially band-pass filters and are therefore particularly adapted to high-frequency amplifiers or circuits where it is necessary to transmit efficiently frequencies within a desired band while strongly attenuating all frequencies outside the band. By proper choice of the transformer constants and the condensers it has been shown that uniformly high voltage amplification was obtained over relatively wide frequency bands. It has also been shown that maximum uniform voltage amplification for a given frequency band was obtained when the output terminals of a transformer coupling circuit were terminated only in a condenser.

A few applications of transformer coupling circuits have been discussed and the individual characteristic shown. It should be understood, however, that these coupling circuits are not limited to the frequency bands illustrated but may be efficiently used at higher frequencies and over wider transmission bands.

The author wishes to acknowledge the helpful suggestions of Mr. H. Whittle of the Bell Telephone Laboratories who was associated in the development of transformer coupling circuits for purposes described in this article.

Abstracts of Technical Articles from Bell System Sources

The Manufacture of Rubber-Covered Wires for Telephone Installations.[1] S. E. Brillhart. Rubber-insulated wires are extensively used by the telephone companies in connecting up apparatus and equipment which is exposed to varying climatic conditions in the same fashion in which rubber-covered wires are employed by other electrical industries. In order to meet all size, strength, and electrical requirements the various wires available for use must differ widely from one another and from commercial rubber-insulated wires, in the character of insulation with which they are covered, as well as the properties of the conductors.

At one extreme this diverse group includes wires for such use as telephone drops which extend from the cable terminals to the buildings in which the stations are located. These drop wires must be capable of carrying heavy snow and ice loads in winter and also be able to withstand exposure to summer heat and strong sunlight in hot climates. They are Nos. 14 and 17 B. & S. gauge hard-drawn copper and bronze conductors with insulations of relatively high quality containing more than 30 per cent of rubber and covered with a weatherproofed cotton braid. At the other end the group contains wires for connecting up instruments within buildings. Being supported frequently and protected from exposure, they are made from No. 22 gauge conductors, insulated with a thinner wall and covered with a colored glazed cotton braid.

A plant is located at the Point Breeze Works, Baltimore, in which certain unique methods and departures from conventional methods of manufacturing rubber-covered wires have been reduced to practice on a large scale. This paper purposes to describe the plant.

Some Recent Developments in Underground Conduit Construction in the Bell System.[2] A. L. Fox. In the past the type of joint made by trowelling cement mortar around the abutting ends of multiple clay conduit has not been entirely satisfactory because in some cases it permitted infiltration of sand and silt which obstructed the ducts.

Various types of joints have been investigated, including a modified

---

[1] *Mech. Engg.*, June, 1932.

[2] *Bell Telephone Quarterly*, July, 1932.

form of the old trowelled joint, cement mortar collars molded in place, and joints made with different plastic materials such as asphalt compounds spread on a fabric backing to facilitate application and retention. None of these provided the tight seal desired.

A satisfactory joint has been developed by encasing the junction in a mortar bandage consisting of a wide band of cement mortar enclosed in cheese-cloth and held tightly against the conduit by tapes passing completely around the joint and secured on top. A separate strip of cheese-cloth is imbedded in the center of the mortar to prevent slumping. With cement mortar of proper consistency and an admixture to insure the desired plasticity sufficient cement paste comes through on the inside when the bandage is applied to provide a tight bond. A strip of paper placed in the bandage under the cheese-cloth in the outer side helps to distribute the pressure of the tying tapes and assists in retaining water in the mortar thereby aiding the hardening and increasing the strength of the joint.

The results of hydrostatic tests of these joints show that they are practically watertight. Their use is expected to effect savings in conduit construction since their high strength permits in many cases omission of the concrete base and further savings may accrue through the increased speed with which conduit can be laid and joined and the fact that the trench can be back-filled immediately without danger of injuring the joints. Other savings in the labor of rodding ducts will be realized because the new type of joint is siltproof.

The Depth of Origin of Photoelectrons.[3] HERBERT E. IVES and H. B. BRIGGS. Previous work has shown that the photoelectrons from a silver plate covered with an equilibrium film of alkali metal follow the wave-length distribution of energy just above the silver surface, i.e., in the alkali metal. This question has been further investigated with particular references to alkali metal films in their early stages of development, where their average depth is less than one atom. Computations made on the absorption of light just within the silver surface show that there should be very definite and striking differences in the wave-length distribution of photoemission if emission occurs due to light absorption in the silver, as contrasted with emission from a film on the silver. Experimental tests made with sodium and cæsium films show that in the earliest measurable state the emission exhibits characteristics peculiar to the light absorption in silver, and that as the films build up the emission becomes characteristic of the energy above the silver. It is concluded that the photoelectrons originate

---

[3] *Phys. Rev.*, June 1, 1932.

partly in the underlying metal and partly in the alkali metal film, the relative proportions varying with the film thickness.

The Lapel Microphone and Its Application to Public Address and Announcing Systems.[4] W. C. JONES and D. T. BELL. Many speakers find it difficult to use the conventional type of microphone, because of the restrictions that it imposes upon their freedom of movement. A microphone, known as the lapel microphone, designed to be attached to the speaker's clothing, has been developed for overcoming these limitations.

The vibratory structure of the lapel microphone is designed to have low mass and stiffness, and to resonate at a comparatively high frequency. The resilient support of the diaphragm adds sufficient mechanical resistance to prevent the occurrence of a prominent peak in the response at the resonance frequency. Means are provided for reducing extraneous noise to a minimum. A part of the sound reaching the microphone, due to body vibration, is rich in low frequencies and must be attenuated, otherwise the quality of transmission will be unnatural. This attenuation is accomplished in the coupling transformer, which, together with the apparatus required for suppressing clicks, for indicating when the circuit is in operation, etc., is mounted in a control cabinet. A flexible cord connects the microphone to this cabinet.

It is expected that the lapel microphone will find application in theaters, churches, convention halls, lecture and banquet rooms, and the like, where public address systems are now employed. It also can be applied in connection with other sound recording and reproducing equipment where the background noise, characteristic of carbon microphones, is not a limiting factor.

Vacuum Tube and Photoelectric Tube Developments for Sound Picture Systems.[5] M. J. KELLY. This paper reviews some recent vacuum tube and photoelectric cell developments which are of interest in sound recording and reproduction systems. An indirectly heated cathode triode is described, in the output circuit of which the current components due to the a-c. power supply of the heater have been reduced approximately 20 decibels below previously obtained levels. This tube makes it possible to use an a-c. supply in amplifiers having flat frequency characteristics with over-all gains of the order of 100 decibels. The microphonic disturbances in vacuum tubes are discussed. A measuring system for evaluating the microphonic noise

---

[4] *Jour. S.M.P.E.*, September, 1932.

[5] *Jour. S.M.P.E.*, June, 1932.

currents is described, and the characteristics of a filamentary cathode tube of low microphonic noise level are given. The characteristics of a double anode, thermionic, gas-filled, rectifier tube for use in a d-c. power supply unit for the sound lamp and vacuum tube filaments of reproducing systems are given. A photoelectric cell of high sensitivity for use in sound reproduction work is described.

Analysis and Reduction of Output Disturbances Resulting from the Alternating-Current Operation of the Heaters of Indirectly Heated Cathode Triodes.[6] J. O. McNally. This paper discusses the disturbance currents in the output circuits of indirectly heated cathode triodes, introduced by the use of alternating current in the heaters. It indicates that the disturbance currents are introduced into the output circuit by (1) the electric field of the heater, (2) the magnetic field of the heater current, and (3) the resistance between heater and grid and between heater and plate, and the capacitance between heater and grid and heater and plate.

The outputs due to the electric field between cathode and plate are produced by the "grid" action of the heater and heater leads. The frequency of the output is chiefly that of the heater supply. The outputs are shown to be effectively reduced by electrostatically shielding the heater.

Disturbance currents of the frequency of the heater supply, and of double this frequency are shown to be produced by the magnetic field. The double-frequency component is shown experimentally to be proportional to the square of the heater current. The following means of reducing the magnetic field are discussed: (1) the adoption of a heater geometry which produces a smaller field in the space between the cathode and the plate, (2) the use of a magnetic shield around the heater system, and (3) the use of a lower current, higher voltage heater.

The ways in which disturbance currents are introduced by leakage resistances and capacitances between heater and grid and heater and plate are indicated, and experimental verification is given for the case of resistance between the grid and heater.

Use has been made of this disturbance current analysis in the development of an extremely low disturbance output tube, which is described.

Fourier Series in Three Dimensions.[7] W. O. Pennell. The classical Fourier Series represents a function in a given interval and then repeats the same values in the next and subsequent intervals. In

---

[6] *Proc. I.R.E.*, August, 1932.

[7] *Am. Math. Monthly*, May, 1932.

other words if $S(x)$ is the classical Fourier Series representing $f(x)$ in the interval $0 < x < a$ then $S(x) = f(x - na)$ where n takes on the values $n = 0, \pm 1, \pm 2, \cdots$ corresponding to the various intervals $na < x < (n + 1)a$.

The author has shown how a generalized Fourier Series $S_1(x)$ may be obtained representing a function in the intervals $na < x < (n + 1)a$ as follows:

$$S_1(x) = b^n f(x - na),$$

where n takes on the values $n = 0, \pm 1, \pm 2, \cdots$, and b is any real constant.

In this paper is described a still more general Fourier Series $S_{11}(x)$ representing a function in the intervals $na < x < (n + 1)a$ as follows:

$$S_{11}(x) = [b^n f(x - na)]_{n\psi},$$

where corresponding to the above intervals n takes on the values $n = 0, \pm 1, \pm 2, \cdots$, b is any real constant, and the subponent notation $n\psi$ denotes the rotation of the plane of the curve about the X axis through an angle $n\psi$ with the XY plane.

Current Propagation in Electric Railway Propulsion Systems. JOHN RIORDAN.[8] This paper presents a systematic method of attack, based on the superposition theorem on the problems of electric railway propulsion systems arising from the presence of tracks and other leaky conductors. The treatment is limited to systems in which the tracks and other leaky conductors may be represented with sufficient accuracy by a single conductor, but includes series and shunt discontinuities in this equivalent conductor. The general equations of current propagation in a single conductor in the presence of a conductor carrying a fixed current are taken in a form similar to the transmission equations ordinarily employed for power transmission and telephone lines, apparently due to H. Pleijel, Report to Swedish Royal Railway Administration, 1919. Though these equations are not rigorous they have been found to agree with experimental observations within engineering accuracy.

The starting point of the treatment is the development of the properties of a basic circuit consisting of a straight conductor of finite length connected at its terminals to a parallel leaky conductor or track which is continuous and infinite. The circuit involves the greatest degree of continuity in the track, subject to the connection of other

---

[8] Presented at the A.I.E.E. Summer Convention, Cleveland, Ohio, June 20–24, 1932. To be published in *A.I.E.E. Transactions.*

conductors, since the only discontinuities are those involved in the connection of the conductor to the track; it is also conveniently adapted to modifications for discontinuities. The general single series discontinuity is a fundamental point of departure in the treatment of such apparatus as track booster transformers. The basic circuit modified by discontinuities as required gives directly the propulsion circuit impedances needed in the railway network impedance diagram; it may also be employed in the construction of a particular kind of cumulative induction curve for neighboring communication lines, which takes into account the distribution of current along the track.

The method of building up complex propulsion systems from basic circuits is illustrated by examples chosen for their practical importance, but the paper does not give detailed procedure for engineering application of the method and its results.

Kennelly-Heaviside Layer Studies Employing a Rapid Method of Virtual-Height Determination.[9] J. P. Schafer and W. M. Goodall. This paper describes a new method of determining the virtual height of the ionized regions by visual observations of the received pulse pattern on a cathode ray oscillograph tube, both for single frequencies and for two frequencies simultaneously. A résumé of the data obtained during observations of some three hundred hours is given. The frequencies used for these tests were 1604 kc, 2398 kc, 3256 kc, 4795 kc, and 6425 kc. A number of the tests included measurements made upon two frequencies in rapid rotation. The more important results may be summarized as follows:

1. On a large number of occasions during the night, a phenomenon has been observed apparently indicating an increase in the density of ionization in the lower layer. This is important because the ionization is usually assumed to decrease during the night hours.
2. Reflections are often observed simultaneously from both ionized layers. An explanation of this phenomenon is given.
3. The virtual heights of the reflecting layers are rarely duplicated from day to day for a given time and frequency.
4. Large numbers of multiple reflections are frequently obtained representing a path distance of over 5000 km. for the last reflection. This fact indicates that the multiple-hop mode of propagation is probable for long-distance transmission.

The Principles of the Light Valve.[10] T. E. Shea, W. Herriott, and W. R. Goehner. The light valve has been used very widely as the

---

[9] *Proc. I.R.E.*, July, 1932.

[10] *Jour. S.M.P.E.*, June, 1932.

modulating device in systems of film sound recording. In this paper the principles of operation of the light valve are discussed, and those engineering factors which prescribe limitations on performance and indicate operating advantages are described in detail. The type of distortion which results when a light valve is overloaded is depicted both for single-plane and two-plane valves. Finally, a new type of light valve having advantages from the standpoints of weight, size, and stability of operation is described.

Economic Control of Quality of Manufactured Product.[11] W. A. SHEWHART. This book of 501 pages is an exposition of the technique developed within the Bell Telephone System for securing economic control of quality of manufactured product at every stage in the process of fabrication all the way from raw materials to finished product. It is divided into seven parts, the first of which is devoted to a general survey of the characteristics of a controlled quality, the scientific basis for attaining control, and the economic advantages to be derived. In the second part, after a definition of what is meant by quality, the methods of presenting data are discussed in detail, both graphically and by means of statistics such as the average, standard deviation, skewness, flatness, and correlation coefficient. Part III presents the necessary and sufficient conditions for the specification of a controlled quality. The allowable sampling fluctuations in statistics are indicated in the next part, illustrated by experiments made under conditions known to be controlled. This part constitutes a survey of the present status of sampling theory. Part V takes up the problem of specifying standard quality and indicates the important changes that should be made in many kinds of specifications in order to secure the greatest assurance of uniform quality at minimum cost. Five practical criteria for determining whether or not the quality under consideration differs from standard by more than an amount that should be left to chance are presented in Part VI. In Part VII, after a summary of the fundamental principles, consideration is given to the problem of sampling, and finally a control program is presented which shows the relation of control to research, design, development, production, and purchasing. Several appendices give the original experimental data on which some of this work was based and a bibliography. Throughout the book the fundamental principles are amply illustrated by practical examples.

[11] Published by D. Van Nostrand Company, New York, 1931.

Contributors to this Issue

A. J. CHRISTOPHER, B.S. in E.E., Carnegie Institute of Technology, 1921; Engineering Department, Western Electric Company, 1921–25; Bell Telephone Laboratories, 1925–. Mr. Christopher has been engaged in voice and carrier frequency transformer design and more recently concerned with the development of repeating, retardation, and induction coils.

A. B. CLARK, B.E.E., University of Michigan, 1911; American Telephone and Telegraph Company, 1911–. Toll Transmission Development Engineer, 1928–. Mr. Clark's work has been largely concerned with toll telephone and telegraph systems.

KARL K. DARROW, B.S., University of Chicago, 1911; University of Paris, 1911–12; University of Berlin, 1912; Ph.D., University of Chicago, 1917; Western Electric Company, 1917–25; Bell Telephone Laboratories, 1925–. Dr. Darrow has been engaged largely in writing on various fields of physics and the allied sciences.

C. J. DAVISSON, B.Sc., University of Chicago, 1908; Ph.D., Princeton University, 1911; Instructor in Physics, Carnegie Institute of Technology, 1911–17; Engineering Department of the Western Electric Company, 1917–25; Bell Telephone Laboratories, 1925–. Dr. Davisson's work since coming with the Bell System has related largely to thermionics and electronic physics.

BANCROFT GHERARDI, B.Sc., Polytechnic Institute, Brooklyn, N. Y., 1891; M.E., Cornell University, 1893; M.M.E., Cornell University, 1894. New York Telephone Company, Engineering Assistant, 1895–99; Traffic Engineer, 1899–1900. New York and New Jersey Telephone Company, Chief Engineer, 1900–06. New York Telephone Company, and New York and New Jersey Telephone Company, Assistant Chief Engineer, 1906–07. American Telephone and Telegraph Company, Equipment Engineer, 1907–09; Engineer of Plant, 1909–18; Acting Chief Engineer, 1918–19; Chief Engineer, 1919–20; Vice President and Chief Engineer, 1920–. Mr. Gherardi is a Past President of the American Institute of Electrical Engineers and is now President of the American Standards Association.

Frank B. Jewett, A.B., California Institute of Technology, 1898; Ph.D., University of Chicago, 1902. American Telephone and Telegraph Company, Transmission and Protection Engineer, 1904–12. Western Electric Company, Assistant Chief Engineer, 1912–16; Chief Engineer, 1916–21; Vice President and Chief Engineer, 1921–22; Vice President, 1922–25. International Western Electric Company, Vice President, 1922–25. American Telephone and Telegraph Company, Vice President, and Bell Telephone Laboratories, President, 1925–. Dr. Jewett is a Past President of the American Institute of Electrical Engineers.

Edward C. Molina, Engineering Department of the American Telephone and Telegraph Company, 1901–19, as engineering assistant; transferred to the Circuits Design Department to work on machine switching systems, 1905; Department of Development and Research, 1919–. As Switching Theory Engineer, Mr. Molina has made contributions to the theory of probability and its applications to telephone problems, such as the efficiency of various trunking arrangements and the significance of data derived from samples.

H. S. Osborne, B.S., Massachusetts Institute of Technology, 1908; Austin Research Fellow in Engineering, 1908–10; Eng.D., 1910. American Telephone and Telegraph Company, Engineering Department, 1910–19; Department of Development and Research, 1919–20; Department of Operation and Engineering, 1920–. Mr. Osborne is Transmission Engineer and as such is responsible for assisting the Associated Companies in connection with telephone and telegraph transmission and protection matters.