IRE Transactions on AUDIO

Volume AU-5 JULY-AUGUST, 1957 Number 4

Published Bi-Monthly

TABLE OF CONTENTS

PGA NEWS

Audio Sessions at WESCON and NEC 77
PGA Chapter Activities 78
With Other Acoustical and Audio Societies .... *B. B. Bauer* 80

CONTRIBUTIONS

Hi-Fi Philosophy from a European Point of View .... *J. Rodrigues de Miranda* 82
Push-Pull Audio Amplifier Theory .... *M. A. Melehy* 86
Phonetic Typewriter .... *H. F. Olson and H. Belar* 90
The "Stereosonic" Recording and Reproducing System .. *H. A. M. Clark, G. F. Dutton, and P. B. Vanderlyn* 96

CORRESPONDENCE

Letter to the Editor .... *Ted Powell* 111

Contributors 112

PUBLISHED BY THE

Professional Group on Audio

IRE PROFESSIONAL GROUP ON AUDIO

The Professional Group on Audio is an organization, within the framework of the IRE, of members with principal professional interest in Audio Technology. All members of the IRE are eligible for membership in the Group and will receive all Group publications upon payment of an annual assessment of $2.00.

IRE TRANSACTIONS® on AUDIO

Published by the Institute of Radio Engineers, Inc., for the Professional Group on Audio at 1 East 79th Street, New York 21, New York. Responsibility for the contents rests upon the authors, and not upon the IRE, the Group, or its members. Individual copies available for sale to IRE-PGA members at $0.80; to IRE members at $1.20, and to nonmembers at $2.40.

PGA News

AUDIO SESSIONS AT WESCON AND NEC

By the time this issue reaches its readers, WESCON and NEC for 1957, will be a thing of the past but it might be well to point out that Audio Sessions, sponsored by the IRE-PGA, were held at both conferences.

Frank Lennert served as Chairman at WESCON with the following program (Session 38):

1) "General Considerations on Phasing Two-Way Loudspeakers," J. K. Hilliard, Altec Lansing Corp.
2) "A Wide Angle Loudspeaker of a New Type," Leonard Pockman, Ampex Corp.
3) "Simplified Audio Impedance Measurements," Vincent Salmon and Myles R. Berg, Stanford Research Institute.
4) "Multichannel Audio Recorders," W. M. Fujii, Ampex Corp.
5) "Methods of Recording Commercial Stereophonic Masters," R. J. Tinkham, Ampex Corp.

Sponsored by Professional Group on Audio

Phil Williams served as Chairman at NEC with the following program, prepared in cooperation with the IRE-PGA.

1) "Permanent Magnets in Audio Devices," R. J. Parker, General Electric Co., Edmore, Mich.
2) "A Transistorized Decade Amplifier for Low Level Audio Frequency Applications," A. B. Bereskin, University of Cincinnati, Cincinnati, Ohio.
3) "Comparative Tests on Light Weight High Power Sound System," F. C. Fischer and A. A. Gerlach, Cook Research Labs., Morton Grove, Ill.
4) "Modern Practice in Noise Control," R. W. Benson, Armour Research Foundation, Chicago, Ill.

The following account of sound conditioning at the Cow Palace in San Francisco, is reprinted from the June, 1957 issue of the *Grid* published by the San Francisco Section.

Contented Audio?

Ultimate success or failure of the WESCON activities will be determined in August. If audiences at the technical sessions of WESCON can hear the speakers, success. Otherwise, vice versa.

An Audio Problem

When a decision was made to attempt to establish six separate lecture halls in the seating area of the Cow Palace, it was obvious that the problem would hinge on the ability to enclose six groups of approximately 400 seats each in such a way that they would be acoustically isolated both from the exhibit areas and from one another.

This problem was initially investigated by Convention Vice-Chairman, B. M. Oliver, and Walter T. Selsted, research director, Ampex Corporation, alternate Chairman of the Arrangements Committee.

Some of the characteristics of the problem are covered in the following notes by Dr. Vincent Salmon who presented a paper including this topic at a recent Audio Engineering Society meeting.

Problems in architectural acoustics may conveniently be classified into those of interior and exterior acoustics. The interior problems deal with the treatment of rooms so that the contained sound may have the proper level and distribution. This is ordinarily accomplished by the proper shape of the room and correct amount, type, mounting, and placement of absorbing and reflecting surfaces. Exterior acoustics, on the other hand, deals with the isolation of sound between two spaces. For example, in a building we may be interested in having external sound isolated so that, for example, a jet plane from Moffett Field will not disturb a sound recording. Conversely, a building should be isolated also, so that a pipe organ in a bistro will not unduly excite the neighbors at hours when normal people would be asleep.

One of the most common errors made by persons unaccustomed to the problems of acoustics, is assuming that because a material is a good sound absorber it will likewise be a good sound isolator. Nothing could be farther from the truth. A few simple calculations based on admittedly idealized assumptions indicate that if material has an absorption coefficient of 90 per cent, then sound passing through it suffers an attenuation of about 5 db. To achieve an attenuation of 10 db will require 99 per cent absorption; and every 5 db of attenuation will require another 9 on the figure for sound absorption. When it is recalled that most common sound absorbents have average coefficients in the region of 70 to 80 per cent and that it takes at least 3 db to make an easily noticeable change in loudness, it is seen that a completely different viewpoint is needed for problems of interior acoustics.

The basic principle in sound isolation is that of using a heavy, impervious, nonresonant sheet of material in which the mismatch between it and adjacent transmitting media is the basis for lack of transmission. Simple theory indicates that it is this mismatch which is primarily responsible for the transmission loss and the mass per unit area is a primary determinant of this transmission loss. Of course, frequency enters since the wall impedance is usually conceived as being masslike and hence both its impedance and loss will increase with frequency. For a transmission loss greater than 10 db, the loss will increase at a rate of about 6 db per octave for a random incidence of sound.

For a given material of a fixed mass per unit area, much better attenuation is achieved if the material is split up into thinner layers separated by an air space. Here each air space acts as an additional impedance mismatch thus increasing the transmission loss. However, the air space must be sufficiently large so that the air does not become so stiff that it transmits well. For example, a piece of ½-inch plate glass has a transmission loss of about 33 db; if two ¼-inch pieces of the same glass (thus, having the same total thickness) are placed 6 inches apart, the transmission loss is almost 55 db.

In this, however, it is essential that the two halves not be interconnected at portions where they vibrate at greatest amplitudes. This admonition is very necessary in construction of isolating walls in which staggered stud construction is employed. In this means of isolation, there are two sets of studs each supporting a sheet of material used as the outer facing of each wall; the studs must not be interconnected. In fact, many acoustical consultants recommend placing a blanket of material in the space between the walls, not because it adds any particular amount of additional transmission loss, but rather, that it serves to make it more difficult for carpenters to interconnect the two sets of studs.

Similar problems of exterior acoustics exist when it is necessary to isolate a space by means of portable walls or drapes. Here the weight law becomes a primary limitation. If hanging materials are used, they must be heavy and impervious and preferably hung in layers well separated from each other. They do not have to be perfectly flat and, indeed, there is some advantage in having irregularities in them but they must, in no circumstances, be tightly bounded together in any one place. Touching is permitted of course, since ordinarily the contact is of small enough area and with low enough pressure so that the transmission is not unduly increased. A single layer of material having a weight of about 1 ounce per square foot will have at 1000 cycles, a transmission loss of almost 5 db. If two layers of this material are separated by at least 2 inches of air, an optimistic estimate is that at least 10 db will be added to the attenuation. The amount actually realized will depend, of course, upon frequency so no single figure can be quoted unless the nature of the disturbing noises and the intended sound (speech transmission, for example) are known.

When walls of extremely high isolation are employed, it is vitally necessary that all cracks be well sealed. For example, if a wall has a transmission loss of 40 db, then an open area equal to one-hundredth of 1 per cent of the total wall area will transmit as much sound as the wall.

Special Research

According to Selsted's report, it was decided that drape material of some sort should be considered, providing the acoustical isolation proved to be satisfactory. In order to determine this, a method of testing for transmission loss was devised and samples of many kinds of materials were compared by Selsted and assistants from Ampex, Phil Smaller, and James Havstad.

Large sections of these materials were supported over a doorway in a concrete building and tightly sealed around the edges. A sound source capable of producing white noise was used and a calibrated microphone determined the sound level transmitted through the various materials.

Transmission loss was measured over half-octave bands throughout the voice spectrum, thereby permitting an evaluation of the relative isolation provided.

One particular grade of impregnated canvas gave a surprisingly high amount of loss and, therefore, was selected for the installation. It will take 95,000 square feet of this material to satisfy the requirements.

PGA CHAPTER ACTIVITIES

Albuquerque, N. M.

The PGA held a meeting at the home of J. E. Palmer on April 29. At this meeting there was a demonstration and comparison of Klipsch and Carlson type horns.

The annual dinner meeting was held May 16, at the Coronado Club. After the dinner, C. W. Remaley gave a talk and demonstration on "High Intensity Sound."

Cleveland, Ohio

The following report on the activities of the Cleveland Chapter during the 1956–1957 season, has been prepared by Carmen P. Germano, 1956–1957 Chairman of the Cleveland Chapter.

"I would like to call to your attention the successful conclusion of a very fine PGA program in Cleveland during the 1956–1957 season. The article by Herb Heller at the end of this report is a summary of our last meeting which was both informative and entertaining.

"Although not mentioned by Mr. Heller, the three channel stereo demonstration was possible only because of his untiring efforts and a result of a year of experimentations in recording. All who attended were enthusiastic in their reception of all aspects of the meeting.

"I should also like to take this opportunity to recap the entire season. We opened in October, 1956, with an informative paper on tape recording problems by R. F. Dubbe, of Minnesota Mining and Manufacturing. The December meeting featured J. F. Wood of Electro-Voice, Inc., who discussed 'Ceramic Phonograph Pickups.' This was a fine engineering meeting. In February, we were fortunate enough to have Dr. J. L. Hunter of John Carroll University, and H. Mull of the National Advisory Committee for Aeronautics, discuss the acoustic problem (and its treatment) of the Cleveland Public Auditorium (long a controversial subject in this area). Finally, of course, Messers. Roys and Moyer of RCA addressed the group in the meeting previously referred to and adequately covered in the enclosed article.

"I would be indeed derelict, in my duty if I didn't at least mention the fine assistance and cooperation I received during the season. My fellow officers, Mr. Heller, the local radio stations, Case Institute of Technology, and many others were instrumental in making the past year a fine and rewarding experience."

"The Last Word in Sound," by Herbert (H^3) Heller.

"For its gala final event of the 1956–1957 season, the Cleveland PGA Chapter went back to the Tomlinson Hall Ball Room at Case Tech, site of the first Stereo Sound Symposium in 1954. Four radio stations (WDOK, WGAR, WHK, and KYW), the Cleveland Press, college publications, and an internal RCA circular helped us invite a selected segment of the classical music public. By eight o'clock, every one of the 250 seats was taken and over 50 stragglers began to line the walls. Chairman Carmen Germano welcomed the many guests and proceeded with the election of officers by the PGA members present. Elected for the 1957–1958 season by unanimous vote, were Jack Goldfarb of Repco, Chairman; Kenneth Hamann, Cleveland Recording Company, chief engineer, Vice-Chairman; and Ralph DeLany, WHK chief engineer, Secretary.

"Harry Roys, Manager of Engineering at the Indianapolis RCA-Victor Record Division and recent recipient of the PGA Achievement Award, briefly introduced the color movie, 'The Sound and the Story,' showing the intricate steps of the birth of a record from the recording session with the Boston Symphony Orchestra under Charles Munch, to the editing in RCA's New York Studios and to the preparation of stampers for the final pressing. Also covered were quality control and packaging to the final destination on the hi-fi set in the home. The meeting was then opened for a stimulating question session, addressed to Mr. Roys and the second guest speaker, Robert Moyer, Manager of Recording Development, also at Indianapolis. Questions ranged from pinch effect in vertical-lateral two-channel stereo disk recordings all the way back in history, to the availability of cylinder records. There seemed to be no end to the constructive questions the alert audience asked. It led us to wonder how much good will a recording company could reap with a pamphlet answering this obvious thirst for knowledge.

"The last portion of the program consisted of a three-channel stereo demonstration, using a Berlant recorder modified with special three-channel ¼-inch heads, and Enger-Hartley amplifier-speaker systems. Excerpts, all locally recorded with E-V 655C mikes, started with sound effects leading to the grandeur of the 100-piece Cleveland Orchestra at Severance Hall and concluded with scenes from 'Maj. Fair Lady' (an obvious take-off produced here by the Heights Temple). The two and one-half hour program was far from enough for the tireless audience. So, after the official closing of the meeting, we brought the Cleveland Orchestra back to the Ball Room for a few more encores.

"Since this event offered something to everybody in the audience, from the comparative laymen to the advanced audio authorities, it was a gratifyingly successful undertaking, culminating an active audio season in Cleveland. We would like to thank all concerned for making this possible, especially Electro-Voice, Inc., for loan of the mikes, and the Musicians Union and the Cleveland Orchestra for their permission."

Dayton, Ohio

Herman R. Weed, Associate Professor of Electrical Engineering at Ohio State University, presented a paper "Feedback on Push-Pull Amplifiers," at the meeting held on April 24.

The following is an abstract of his paper:

"Most practical power amplifiers used in audio work utilize one or more forms of feedback. Although this feedback is designed to improve the performance of the amplifier, certain undesirable factors must be considered, affecting gain, harmonic distortion, hum, bandwidth, damping.

"Feedback may be divided first into positive or negative, and then into the more specific types of voltage, current, or combination voltage and current feedback. The practical problems: how and from where is the feedback obtained, and what is the effect of each type of feedback on the amplifier performance characteristics?

"The more important methods of obtaining feedback are 1) cathode follower, 2) unby-passed cathode resistor, 3) voltage feedback from secondary of output transformer, 4) plate feedback, and 5) tapped output transformer.

"Three popular circuits are the Williamson type, using output transformer feedback; Williamson Ultra-Linear, employing a tapped output to the screens of the amplifier output tubes; McIntosh amplifier, requiring special input and output transformers to provide both plate and cathode loading.

"One of the outstanding advantages of using secondary output transformer feedback, is the reduction in distortion caused by transformer leakage, with the feedback originating at a point past the output transformer.

"Early amplifier circuits made use of triodes in the output stage because of their greater linearity and less critical matching of the output load. It is apparent that if the screen of a tetrode is tied to the plate, the tube will exhibit triode characteristics, and if tied to the plus *B* supply, will exhibit tetrode characteristics. It is not necessarily true, however, that minimum distortion will result from pure triode operation, or maximum distortion from tetrode operation. The ultralinear circuit compromises the two by providing a tap on the output transformer between the two limits and thereby producing less distortion than with either extreme.

The combination of plate and cathode loading used in the McIntosh amplifier combines the ideal distortion characteristics of the cathode follower amplifier with the more practical plate loaded circuit by providing half the load at each point and obtaining the remaining necessary feedback by resistance coupling from the cathodes of the output tubes to the unbypassed cathodes of the inverter stage."

Long Island, N. Y.

John D. Meehan, chairman of the newly organized Long Island Chapter, reports that they held their final meeting of the season on May 7. At this meeting, Saul J. White, chief engineer of RACON Electric Co. Inc., presented a paper on "The Dynamic Loudspeaker—Can We Improve It?" Included was a demonstration of the RACON 15-inch speaker mounted on a collapsible open (towards the rear) baffle. To augment the base, a cluster of four units was used. Mr. Meehan reports that the response was entirely adequate even when a record of percussion instruments was played.

Philadelphia, Pa.

L. H. Good, chairman of the Philadelphia Chapter reports the following meetings:

On March 13, Richard M. Carrell, of RCA, presented a paper on "Negative Impedance Amplifier." The paper included a demonstration of the effects of using negative impedance amplifiers with various speaker-enclosure systems.

On April 10, Frank H. Slaymaker, chief engineer of the Stromberg-Carlson Company, presented a paper with demonstration on, "Carillons and Bells."

On May 1, at a joint meeting with the Philadelphia Section and Students' and Ladies' night, Dr. D. W. Martin, of the Baldwin Piano Co., presented a paper on "Electronic Music." The paper included a demonstration of the Baldwin organ.

Syracuse, N. Y.

The May 2 meeting of the Syracuse Section was sponsored by the Audio Chapter. At this meeting J. A. Bowman, Special Projects General Manager, Western Electric Company, gave a paper on the "Trans-Oceanic Telephone Cable."

San Francisco, Calif.

On May 21, Dr. Leonard Pockman, Research Physicist, Ampex Corporation, presented the paper "Wide-Angle High-Frequency Loudspeaker." A short resume of this paper appears below:

"*Helium-Filled Horns*—Details of a new high-frequency loudspeaker of unique design were given at the May meeting of the PGA. The unit, of Ampex design, was presented in the demonstration room of the Ampex plant by Dr. Leonard Pockman. Six coaxial horns comprise the device which has a distribution pattern approaching 180° from 1000 to 10,000 cps. Appropriate mixtures of air and helium are used to fill the individual horns and control the velocity of sound transmission to accomplish the result."

WITH OTHER ACOUSTICAL AND AUDIO SOCIETIES

The March, 1956 issue of *The Journal of the Acoustical Society of America*, contains 14 papers and 6 Letters to the Editor, many of which will be of interest to the members of the IRE Professional Group on Audio.

Richard K. Cook, of Bell Telephone Laboratories, describes "Absorption of Sound by Patches of Absorbent Materials." Circular patches, or long strips, of sound absorbent material are on the surface of an infinite perfectly reflecting plane. Exact solutions are found for the absorption of perpendicularly-incident plane waves, and of randomly-incident waves, on a circular piston-like absorber. An exact solution is found for the absorption of perpendicularly-incident plane waves on an infinitely long strip-piston absorber.

Richard E. Werner, of RCA, Camden, N. J., writes on "Effect of a Negative Impedance Source on Loudspeaker Performance." A direct radiator moving coil loudspeaker driven by an amplifier whose output impedance approaches the negative of the blocked voice-coil impedance, can be made to exhibit extended low-frequency response with reduced distortion. The effect of the system is, in some ways, analogous to a manifold increase in loudspeaker efficiency. In a typical case, neutralization of 70 per cent of the blocked voice-coil impedance completely damps the cone resonance, as well as substantially reducing the nonlinear distortion below resonance. When the amplifier is compensated for the falling radiation resistance at low frequencies, uniform output can be obtained to any arbitrary low frequency, limited only by the ultimate power-handling capability of the amplifier and speaker. In this system no additional amplifier power is required at frequencies down to the speaker resonance; additional power is required below that point.

W. J. Toulis, of the United States Navy Electronics Laboratory, San Diego, Calif., contributed an interesting paper entitled, "Radiation Load on Arrays of Small Pistons." The radiation load on a piston in an array is examined, preliminarily, through the correspondence of a piston in an infinitely long tube and in a conical horn to arrays on an infinite plane surface and on a rigid sphere, respectively. The evaluation of the velocity potential for the two arrays suggests that the average radiation load on an array consists of two terms. The principal term arises from a uniform distribution of the total source strength over both active and inactive areas of the array. The secondary term corresponds to the free space load for one of the pistons but reduced in magnitude by a factor proportional to the average separation between adjacent pistons.

"Auditory Adaptation," is an interesting paper contributed by James F. Jerger, of the Audiological Laboratory of Northwestern University. The apparent decline in the ear's response under sustained stimulation, a phenomenon which has been variously labeled "auditory fatigue," "perstimulatory fatigue," and "auditory adaptation," was measured for pure tones over a wide range of frequencies and intensities by the median-plane-localization method. For a given intensity, increasing the frequency from 125 to 1000 cps increased both the initial rate and maximum amount of adaptation. Above 1000 cps further increases in frequency did not appreciably change the adaptation curves.

"Confirmation of the Normal Threshold for Speech on CID Auditory Test W-2," was written by John F. Corso, of the Department of Psychology of the Pennsylvania State University. This study was undertaken to determine the threshold for speech intelligibility for normal ears on the Central Institute for the Deaf (CID) Auditory Test W-2. Data were collected on both right and left ears of 139 subjects, ages 18 to 24 years inclusive, showing negative otological findings, a maximal pure tone hearing loss of 5 db at 500, 1000, or 2000 cps, and a history of minimal noise exposure. The obtained mean threshold for speech, for all ears combined ($N=278$), was 18.55 db re 0.0002 dyne/cm^2, with a standard deviation of 4.46 db. These results confirm the normal value for the threshold of intelligibility for the spondaic words of CID Auditory Test W-2.

The issue contains its usual excellent, "References to Contemporary Papers on Acoustics," by Robert N. Thurston and "Review of Acoustical Patents," by Robert W. Young.

Benjamin B. Bauer

Hi-Fi Philosophy from a European Point of View*

J. RODRIGUES DE MIRANDA†

Summary—**The paper deals with popularizing hi-fi listening facilities in Europe which, due to a difference in broadcasting practice, is easier over there than in the United States. These differences are analyzed and the conclusion of the writer is that a large contingent of hi-fi listeners in Europe could be created by improving the reproduction quality of fm radio sets without increasing the cost too much.**

As a distortion limit for this purpose a figure of 1 per cent is chosen, while the output power should be near the 10-watt figure.

To realize these properties a special circuit has been designed, *viz.*, a single-ended push-pull output stage with combined positive and negative feedback. A loudspeaker with a voice coil impedance of 800 ohms is being used, making an output transformer superfluous. This direct coupling is one of the necessities to get a cheap but very good setup.

Further, the so-called 3D system and means for getting rid of the keyhole effect are discussed.

IT WAS NOT without hesitation that this title for a paper has been chosen because we, in Europe, do realize that hi-fi is an American child of an electronic father and an acoustic mother. But, although this child has not yet reached the age of full maturity, its offspring can be met already on the other side of the ocean. And this is why there is an excuse to discuss this subject here. The first question arising is, what *is* hi-fi? How can it be recognized? One may state that this term is used for parts of a *reproduction* system of a transmission chain, *e.g.*, for a pickup, an amplifier, and a loudspeaker system. This includes, that as a rule the *source* is given either as a gramophone record, as a tape, or as a detected radio signal. What we want from a complete hi-fi reproducing chain is to give a reproduction as faithful as possible from the original source. One thing is sure: there is not much sense in making the properties of the reproducing equipment substantially better than those of the source. Neither is it sensible to use a good amplifier with a poor pickup or with an inadequate loudspeaker system.

One could say, therefore, that no part of a hi-fi installation should be the weakest link in the transmission chain, in order to be worthy to bear the epithet "hi-fi." For the present let us keep to this philosophy.

One of the best sources for music and speech is fm radio, at least in part of Western Europe. Many of the European broadcast programs are of good quality, particularly the live ones, and several broadcasting companies claim a flat response curve from 20–15,000 cps and a distortion figure of less than 1 per cent. They have the additional advantage of being accessible to a very large public.

* Manuscript received by the PGA, May 29, 1957.
† Philips Industries, Eindhoven, Holland.

Prerecorded tape can be of high quality too, but as yet it is not available to a large extent. In addition, the playback equipment is rather expensive due to the complicated electromechanical parts. Longplaying records are often excellent and are obtainable everywhere, but the maximum obtainable quality, especially freedom of distortion, is certainly less than that of good tape. Because of the high quality of fm programs (fm amplification and detection can easily be accomplished with very little distortion) and also considering the fact that practically all receivers are provided with pickup and tape-recorder connections, we may say that in a large part of Western Europe, prospective buyers of an fm radio set will benefit greatly if this set is provided with a good af section. All this makes us feel convinced that for this part of the world, there really is sense in producing sets in which a medium (or low) power hi-fi amplifier is incorporated—that is to say, if this is allowable from a commercial point of view.

But there are still other reasons, mainly due to differences in the general aspect of broadcasting systems between the U.S. and Europe.

We believe that in Europe the possibility of receiving uninterrupted musical programs, *e.g.*, concerts, operas, etc., is greater than in the U.S.

The desire to listen to such musical programs exists in the U.S., however, due to lack of opportunity of being able to receive uninterrupted musical programs, we believe more use is made of tape and records in the American home, and so the demand has developed of having hi-fi amplifying equipment.

Again, in Europe, with its different broadcasting organization, with the wide and ever-increasing fm coverage, the listening public has learned to take advantage of these possibilities and demand better reproducing quality of radio and audio equipment.

As a result of this demand, the European radio industry has paid a lot of attention to audio circuits and acoustics with the result that today in Europe the cost percentage of the audio part of a radio receiver is much higher than previously and probably much higher than that of an American radio set.

The European acoustical approach has been to try to reduce the "keyhole effect" by using more than one loudspeaker with the additional speakers or some of them located in the sides of the radio cabinet, resulting in such advertising slogans as "3D" etc. It is evident that these speakers have to be of high quality, and in this respect it is interesting perhaps, to mention that the double cone constant-impedance type of loudspeaker is becoming more and more popular in the U.S.A.

As to the enclosure, as long as the loudspeakers have to be incorporated in a radio cabinet, the bass response is limited. In the higher regions also, there is a possibility for improvement. If we try to analyze what is wrong with a one-loudspeaker set from the sound distribution point of view, we find that there are three main points to consider:

1) The *surface brilliancy* of a loudspeaker is often annoying.
2) Beside this effect, the solid angle with which we perceive orchestral music from a loudspeaker is much too small for giving the impression that a real orchestra is playing.
3) In a concert hall we listen to approximately 90–95 per cent indirect and 5–10 per cent direct sound. In a listeners' room we listen to approximately 50 per cent indirect and 50 per cent direct sound.

The use of extra speakers, located in the sides of the radio cabinet, is only a first approach to these problems, but it does result in increasing the ratio indirect-direct sound if the radio is properly located in the room. Of course, if the side speakers radiate towards highly absorbent surfaces such as curtains, the acoustical effect is greatly reduced and the sound balance disturbed.

The better approach to this problem is to use loudspeakers in separate enclosures having the bass speaker mounted in an acoustical cabinet designed for low frequency reproduction and the high-frequency speakers located in the most favorable position in the room. By doing this, full advantage is taken of the possibilities of the audio part of the radio set.

A remarkable improvement can be obtained in this way, especially if the high-tone speakers are placed at some distance from each other and preferably if they radiate towards a reflecting surface (a wall or ceiling) in such a way that the images are sufficiently wide apart. However, the best result is obtained when a small direct radiating source for these frequencies is used as well. To give this thought a follow-up one should like a radio set with a hi-fi amplifier to have means for connecting separate high-tone speakers in such a way that the energy ratio of the speakers can be adjusted according to the type of program.

Having a separate speaker installation accentuates the desire of obtaining a higher quality from the audio part of the radio set.

There are some limiting factors, such as cost, etc. in having in a radio set, a high-power, hi-fi audio part; however, a good compromise solution will be described later.

Where cost is not a necessary limiting factor, the best solution is the use of a tuner and a high-power hi-fi equipment. If and when true stereophonic sources become more widely available, the necessity will arise to duplicate the amplifier and loudspeaker equipment. With stereophony difficulties 1) and 2), fore-going, are automatically solved leaving only, to increase the ratio indirect/direct sound to the proper value, which has to be done experimentally by using reflecting surfaces.

Now let us turn to cost-price considerations. Of course, every company has its own approach to this and in order to give a more detailed example a survey of the trend in development in the Philips factories in Eindhoven, is given below.

We have to investigate first the audio amplifier part of a reproducing system and try to find means to make it both good and inexpensive. The conspicuous component, which limits quality and at the same time has a considerable influence on the price of the amplifier, is the output transformer. In this direction we have launched an attack with interesting results. The first line of attack has been the application of separate channels for the high- and low-frequency ranges, resulting in an audio system, comprising a preamplifier, followed by a frequency dividing network leading into two output stages. The advantages of such a system are:

1) High-low intermodulation is virtually absent if the preamplifier is correctly designed.
2) The low-frequency output transformer can be made to have a high impedance, the stray inductance and capacitance being of no importance. The treble transformer need not have a high inductance, so that stray inductance and capacitance can be kept low.
3) It gives a convenient solution for using separate bass and treble speakers. The cross-over filter can be cheap because it is incorporated in the preamplifier. In sets of this type sold in Europe and to some extent in the U.S., it consists of a double RC filter.
4) High output (for a radio set). Compared to a push-pull output stage with the same output tubes the above mentioned system is less expensive due to the absence of a phase inverter and because of a cheaper output transformer system.

The second approach has been a direct assault on the existence of the transformer itself by using a single-ended push-pull amplifier. This circuit, which in itself is not new, comprises two power tubes connected in series as far as dc is concerned, but in parallel for the delivered audio output power. Hence, the required load impedance is approximately $\frac{1}{4}$ of that of a normal push-pull output stage. If appropriate tubes are chosen, a load impedance for optimum conditions of something like 800 ohms can be obtained. No dc has to pass through this load. These considerations led to examination of the practicability of a loudspeaker with a voice coil of 800 ohms instead of the usual much lower values. This appeared to be quite feasible for a factory with the facilities and the know-how for handling and winding fine gauge wire.

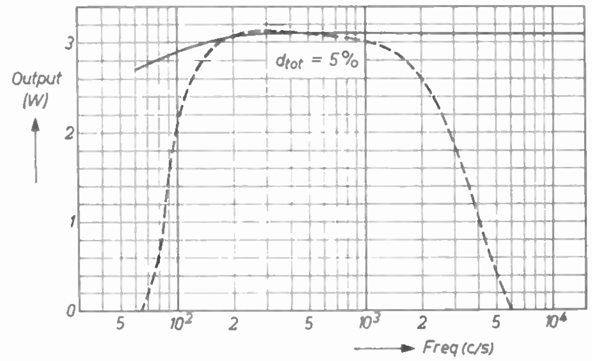

Fig. 1—Power-frequency curve for a simple output stage with transformer-coupled low-impedance loudspeaker (broken line) compared to that of a simple single-ended push-pull output stage without loudspeaker transformer (drawn in full). These curves are given for 5 per cent distortion because this is the usual value for radio sets of this price class.

A third conception has been the combination of the two systems mentioned, *i.e.*, a double amplifier with two single-ended push-pull stages and two high-impedance loudspeakers, one for the high- and one for the low-frequency range. This system results in a very good amplifier but the cost is rather high, at least for radio sets.

The fourth line of attack has been basically an improvement on the single-ended circuit with a preamplifier. In the three systems mentioned above, not much use has been made of feedback because of the loss in sensitivity. Nevertheless the single-ended output stage simply aks for feedback because of the absence of a phase shifting output transformer.

All systems used for getting low distortion figure and a flat frequency response rely on feedback system in one way or another, *e.g.*, such as a large amount of plain negative feedback, feedback for distortion only, and combinations of positive and negative feedback. But in the classic amplifiers the factor that limits the applicable amount of feedback and thus the quality, is always the phase shift and the low-frequency losses of the output transformer. In some cases phase-compensating networks are required.

These limitations are, of course, fully absent if no output transformer is used. That is the reason why several output transformerless circuits have been described in which, however, little or no attention has been paid to cost.

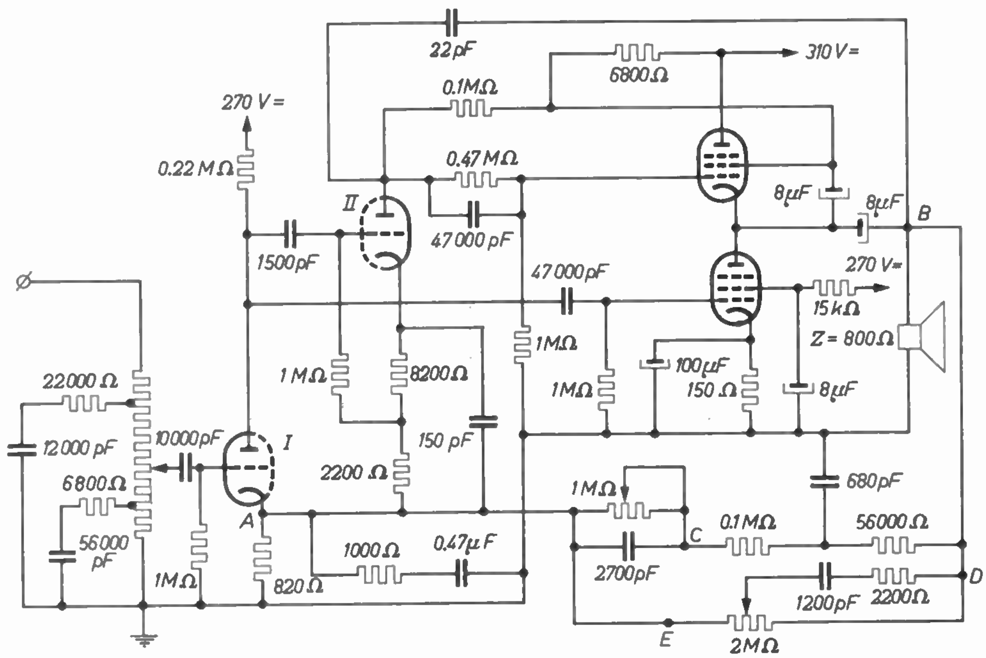

Fig. 2—Circuit diagram with combined positive and negative feedback of a single-ended push-pull output stage capable of delivering 8 watts at 1 per cent in the load of 800 ohms at all frequencies between 80 and 10,000 cps. The positive feedback is obtained through the network between cathode of tube I and earth; the negative feedback by means of the network between points B and A. Tone controls are incorporated in this network.

The combination of a single-ended push-pull amplifier and a high impedance speaker eliminates the use of an output transformer, which allows us to obtain full power even in the lowest and highest frequencies in the frequency spectrum required. In addition, the advantages of a push-pull output stage are retained. A cost-price comparison between a "simplified" circuit of this kind with a "classic" amplifier, showed that the new circuit is *not* more expensive (the cost of an additional power tube and some components being cancelled by the elimination of the output transformer). But much better properties are obtained as shown by the frequency-power curves of Fig. 1.

The high-impedance loudspeaker opens a new way to tackle the quality-cost complex which ultimately led to a really good amplifier suitable for application in radio sets, with an output of 8 watts at 1 per cent distortion. The sensitivity is 20 mv for 50 mw output (which is more than adequate for the audio part of a radio receiver) although only three tubes are used:

> One 12AX7 (preamplifier and phase inverter) and two power tubes more or less equivalent to the 6CW5.

These results have been made possible by over-all negative feedback combined with a positive feedback between phase-inverter and preamplifier. This system, together with even harmonic cancellation in the push-pull output stage and in the preamplifier-phase inverter combination, gives excellent results. The circuit diagram is given in Fig. 2. Negative feedback is applied from the output (point B) to the cathode of the preamplifier (point A) by means of a network in which treble and bass controls are incorporated.

Positive feedback is obtained by letting the cathode current of the preamplifier and the phase-inverter pass through a common resistor. The distortion figures of an amplifier as described above with a setting to deliver 8 watts are shown in Fig. 3. These distortion figures were measured at 90, 1000, and 8000 cps with the tone controls in the position max bass and max treble. Fig. 4 shows the im curve with 1:4 input voltage ratio for frequencies 60 and 10,000 cps, measured with a flat bass response and maximum highs. Finally, Fig. 5 shows the internal resistance as a function of the frequency. For the larger part of the spectrum this resistance is in the neighborhood of 25 ohms (load resistance 800 ohms) while in the lower region, where the speaker resonance is expected to be a pure resistance, it becomes negative. This helps in damping the speaker.

The acoustical results obtained with a radio set equipped with an audio frequency part built according to the principles mentioned above and provided with good loudspeakers are really remarkable. Although this audio part does not show the extremely low distortion figures usually associated with hi-fi amplifiers, it adds so little to distortions already present in the various sources that from a practical angle it might be called a hi-fi radio set.

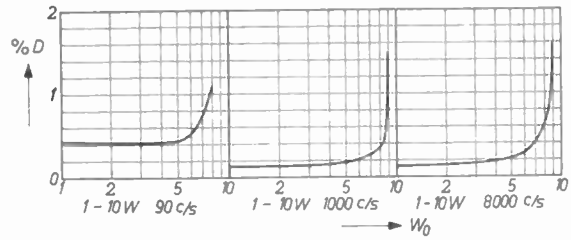

Fig. 3—Distortion as a function of output power for three frequencies (for an amplifier of Fig. 2) (90 cps, 1000 cps, and 8000 cps).

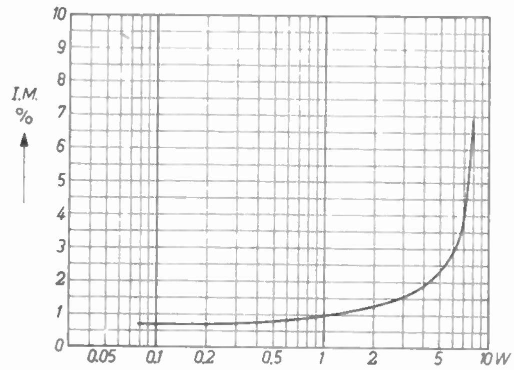

Fig. 4—Intermodulation curve (for an amplifier of Fig. 2). Test frequencies: 60 and 10,000 cps voltage ratio 4:1 at the input. Tone controls in position "flat."

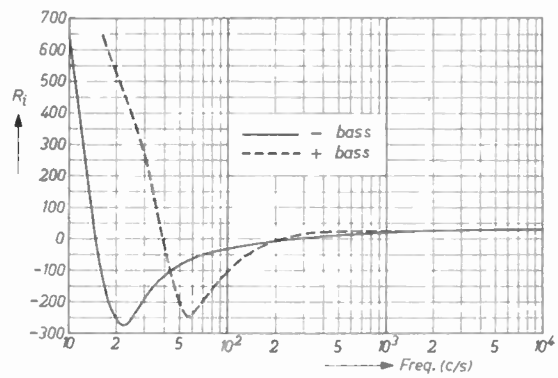

Fig. 5.—Internal impedance of an amplifier of Fig. 2, as a function of frequencies for two positions of the bass control.

Very important for applications in this field is that the price is low as may be deduced from the fact that the number of components is small and that their prices are low, while the difference in price between a high-impedance loudspeaker and a normal one is negligible.

Push-Pull Audio Amplifier Theory*

M. A. MELEHY†

Summary—**Assuming nonlinear tube characteristics, this paper presents: 1) a single mathematical analysis of instantaneous relations applicable to all classes of operation of push-pull audio vacuum tube amplifiers; 2) a mathematical derivation, applicable to all classes of operation, of the composite load line equation and a study of the amplifier frequency response both at medium and low frequencies, with third and higher odd-order effects neglected at low frequencies; 3) an analytical method for finding the maximum magnitude of class AB_1 and AB_2 grid bias beyond which appreciable nonlinear distortion occurs; and 4) a single general condition for matching the load to the tubes irrespective of the class of operation, and the interpretation of this condition for each case.**

Introduction

PUSH-PULL audio amplifiers have been studied by many authors[1–13] who were mainly concerned with graphical techniques and linear operation. The purpose of this paper is to present a single mathematical analysis applicable to all classes of operation of the push-pull amplifier assuming nonlinear tube characteristics. For the derivation of the load line equation and frequency response below the midfrequency range, certain restrictions are placed on the nonlinearity. The analysis takes into consideration the fact that the three transformer circuits involved are all mutually coupled.

Instantaneous Equations

Throughout this paper, the following assumptions are made: 1) the output transformer is lossless and has a unity coupling coefficient; 2) the transformer center tap is exact; 3) the tubes are perfectly matched; and 4) the two input grid-to-cathode instantaneous voltages are of opposite polarity and of equal magnitude.

* Manuscript received by the PGA, May 15, 1957.

† Dept. of Elect. Eng., Michigan State Univ., East Lansing, Mich.

[1] N. W. McLachlan, "The theory of push-pull," *Wireless World*, vol. 22, pp. 629–634; June 13, 1928.

[2] *Ibid.*, pp. 114–118; January 30, 1929.

[3] L. E. Barton, "High audio power from relatively small tubes," Proc. IRE, vol. 19, pp. 1131–1149; July, 1931.

[4] C. E. Kilgour, "Push-pull amplifier graphics," *Electronics*, vol. 6, p. 73; March, 1933.

[5] B. J. Thompson, "Graphical determination of push-pull audio amplifiers," Proc. IRE, vol. 21, pp. 591–600; April, 1933.

[6] J. R. Nelson, "Class B amplifiers considered from the conventional class A standpoint," Proc. IRE, vol. 21, pp. 858–874; November, 1933.

[7] T. McLean, "Analysis of distortion in class B audio amplifiers," Proc. IRE, vol. 24, pp. 487–509; March, 1936.

[8] P. T. Sah, "Quasi-transients in class B audio-frequency push-pull amplifiers," Proc. IRE, vol. 24, pp. 1522–1541; November, 1936.

[9] W. B. Nottingham, "Power conditions for maximum power in class A amplifiers," Proc. IRE, vol. 29, pp. 620–623; December, 1941.

[10] H. L. Krauss, "Class-A push-pull amplifier theory," Proc. IRE, vol. 36, pp. 50–52; January, 1948.

[11] K. R. Sturley, "Push-pull audio-frequency amplifiers," *Wireless Eng.*, vol. 26, pp. 338–343; October, 1949.

[12] A. Peterson and D. B. Sinclair, "A single-ended push-pull amplifier," Proc. IRE, vol. 40, pp. 7–11; January, 1952.

[13] C. Yeh, "Analysis of single-ended push-pull audio amplifier," IRE Trans., vol. AU-1, pp. 9–19; April, 1953.

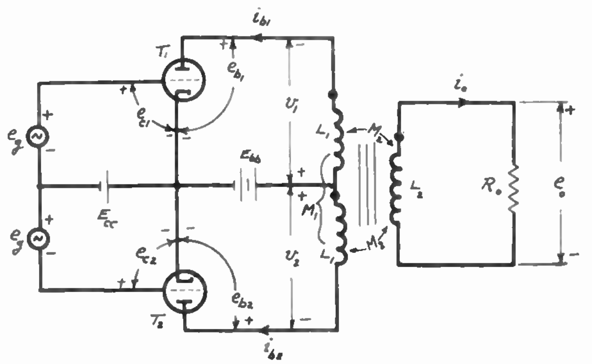

Fig. 1—Circuit diagram for a push-pull amplifier stage.

Considering Fig. 1, the instantaneous circuit voltages and currents can, therefore, be expressed as:

$$e_{b1} = E_{bb} - v_1 \tag{1}$$

$$e_{b2} = E_{bb} - v_2 \tag{2}$$

$$v_1 = L_1 \frac{di_{b1}}{dt} - M_1 \frac{di_{b2}}{dt} + M_2 \frac{di_0}{dt} \tag{3}$$

$$v_2 = L_1 \frac{di_{b2}}{dt} - M_1 \frac{di_{b1}}{dt} - M_2 \frac{di_0}{dt} \tag{4}$$

$$0 = M_2 \frac{di_{b1}}{dt} - M_2 \frac{di_{b2}}{dt} + L_2 \frac{di_0}{dt} + R_0 i_0 \tag{5}$$

$$e_{c1} = E_{cc} + e_g \tag{6}$$

$$e_{c2} = E_{cc} - e_g \tag{7}$$

$$M_2 = \sqrt{L_1 L_2}, \quad \text{and} \quad M_1 = L_1 \tag{8}$$

$$i_{b1} = f(e_{c1}, e_{b1}) \tag{9}$$

$$i_{b2} = f(e_{c2}, e_{b2}). \tag{10}$$

From (3)–(5) and (8), it follows that

$$v_1 = -v_2 \tag{11}$$

$$i_0 = -\sqrt{\frac{L_2}{L_1}} \frac{v_1}{R_0} \tag{12}$$

$$i_b = Gv_1 + \frac{1}{L_1} \int_0^t v_1 dt + A \tag{13}$$

where

$$G = \frac{1}{R} = \frac{L_2}{L_1} \frac{1}{R_0} \tag{14}$$

and i_b, defined as the composite plate current, is given by

$$i_b = i_{b1} - i_{b2}. \tag{15}$$

In the midfrequency range, the integral term in (13) becomes negligible. Therefore, the conditions that the tubes are matched requires that the constant, A, be zero; and (13) reduces to

$$i_b = Gv_1 + \frac{1}{L_1}\int_0^t v_1 dt \tag{16}$$

In the midfrequency range, the simplified version of (16) together with (1), (2), (6), (7), (11), and (15), constitute the foundation for Thompson's[5] composite diagrams shown in Fig. 2.

It is to be noticed that all the above equations are valid irrespective of the nonlinearity of the characteristics of the tubes and their classes of operation.

Composite Load Line Equation and Amplifier Frequency Response

The solution of (16) is here accomplished by means of another relation between i_b and v_1. Consider the composite characteristics of Fig. 2. Let these composite curves be expressed by

$$i_b = F(e_{c1}, e_{b1}) = F[(E_{cc} + e_g), (E_{bb} - v_1)]. \tag{17}$$

Assume that around the Q point the i_b function can be represented by a Taylor's series which is convergent in the region of operation. Then

$$i_b = F(E_{cc}, E_{bb}) + \left[\left(e_g\frac{\partial}{\partial e_{c1}} - v_1\frac{\partial}{\partial e_{b1}}\right) + \frac{1}{2!}\left(e_g\frac{\partial}{\partial e_{c1}} - v_1\frac{\partial}{\partial e_{b1}}\right)^2 + \cdots + \frac{1}{n!}\left(e_g\frac{\partial}{\partial e_{c1}} - v_1\frac{\partial}{\partial e_{b1}}\right)^n + \cdots\right] i_b \tag{18}$$

where all the partial derivatives are evaluated at the Q point.

From the assumption that the tubes are matched, it can be shown that at the Q point

$$F(E_{cc}, E_{bb}) = 0 \tag{19}$$

and

$$\frac{\partial^n i_b}{\partial e_{c1}{}^m \partial e_{b1}{}^{(n-m)}} = \frac{2\partial^n i_{b1}}{\partial e_{c1}{}^m \partial e_{b1}{}^{(n-m)}} \quad \text{for } n \text{ odd}$$
$$= 0 \quad \text{for } n \text{ even} \tag{20}$$

where $0 \leq m \leq n$.

Eq. (19) follows from (9), (10), and (15). To establish (20), consider (1,) (2), (6), (7), and (11) from which it follows that

$$e_{b2} = 2E_{bb} - e_{b1} \tag{21}$$

$$e_{c2} = 2E_{cc} - e_{c1}. \tag{22}$$

Hence, (10) can be written as

$$i_{b2} = f[(2E_{cc} - e_{c1}), (2E_{bb} - e_{b1})] \tag{23}$$

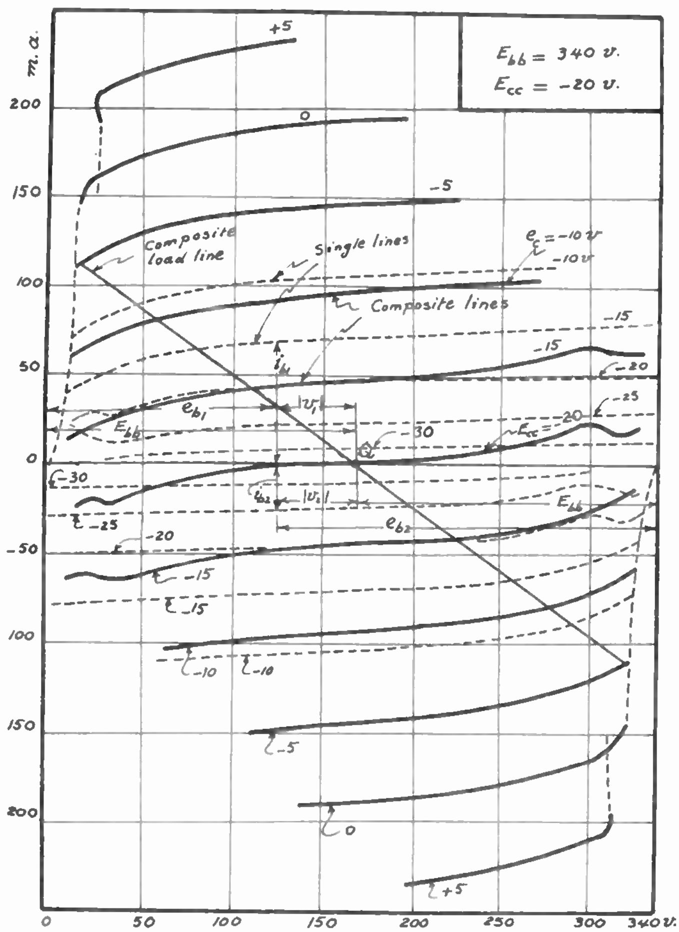

Fig. 2—Composite static plate characteristics for type 6L6 tube.

which at the Q point yields

$$\frac{\partial^n i_{b2}}{\partial e_{c1}{}^m \partial e_{b1}{}^{(n-m)}} = (-1)^n \frac{\partial^n i_{b1}}{\partial e_{c1}{}^m \partial e_{b1}{}^{(n-m)}} \quad (0 \leq m \leq n). \tag{24}$$

Eq. (24) together with (15) establishes (20). Neglecting in (18) the derivatives of the third and higher odd order, and from (16), (19), and (20) and by definition of the single-tube transconductance, g_{m1}, and plate conductance, g_{p1}, it follows that

$$(G + 2g_{p1})v_1 + \frac{1}{L_1}\int_0^t v_1 dt = 2e_g g_{m1} \tag{25}$$

where g_{m1}, and g_{p1} are evaluated at the Q point. If

$$e_g = E_{gm}\cos(\omega t), \tag{26}$$

then the steady-state solution of (25) is

$$v_1 = V_m \cos(\omega t + \phi)$$

where

$$V_m = \frac{2g_{m1}E_{gm}}{\sqrt{\left(\frac{1}{\omega L_1}\right)^2 + \left(\frac{L_2}{L_1}\frac{1}{R_0} + 2g_{p1}\right)^2}} \tag{27}$$

and

$$\phi = \tan^{-1}\left[\frac{1}{\omega L_1\left(\frac{L_2}{L_1}\frac{1}{R_0} + 2g_{p1}\right)}\right].$$

Eqs. (16) and (27) yield

$$i_b = \left[\frac{L_2}{L_1}\frac{1}{R_0}\right]v_1 \pm \frac{V_m}{\omega L_1}\sqrt{1-\left(\frac{v_1}{V_m}\right)^2} \tag{28}$$

which is the composite load line equation. Eq. (28) represents an ellipse. Since v_1 is proportional to i_o [see (12)], (27) is useful for the study of the amplifier response at low frequencies. In the midfrequency range, (28) approaches a straight-line equation, since the term ωL_1 becomes large.

Optimum Bias for Class AB Operation

It has been customary to consider the bias for class AB_1, and AB_2 operation as the e_c intercept of the extension of the linear portion of the tube transfer characteristic. Such a procedure, however, is based on no mathematical foundation. Moreover, for most power tubes, the characteristics in the useful range have no linear portion. It is the purpose here to find the condition from which the grid bias generally can be obtained.

Graphical techniques based on the linearity of the composite load line reveal that the waveform of i_b, and hence the output voltage and current become distorted as the magnitude of the negative bias is increased beyond a certain value referred to as the optimum bias. To find this optimum bias, assume the signal frequency to be in the midfrequency range, and that

$$e_{c1} = E_{cc} + E_{gm}\cos\omega t \tag{29}$$

and at $t=0$, $i_b = I_{bm}$. If i_b had no distortion, then one could write

$$i_b = I_{bm}\cos\omega t \tag{30}$$

so at

$$\omega t = \frac{\pi}{2}, \qquad \frac{di_b}{dt} = -\omega I_{bm}. \tag{31}$$

It is possible to find an exact time derivative for i_b at $t=\pi/2\omega$ in terms of the tube parameters at the Q point. In order to do so, consider (17) which yields

$$di_b = \left(\frac{\partial i_b}{\partial e_{c1}}\right)_{e_{b1}} de_{c1} + \left(\frac{\partial i_b}{\partial e_{b1}}\right)_{e_{c1}} de_{b1}. \tag{32}$$

In the midfrequency range, (16) reduces to $i_b = v_1/R$ and by (1), $i_b R = E_{bb} - e_{b1}$; hence

$$de_{b1} = -R\,di_b. \tag{33}$$

From (29), (32), and (33) it follows that

$$\left(\frac{di_b}{dt}\right)_{t=\pi/2\omega} = \left[\frac{\left(\frac{\partial i_b}{\partial e_{c1}}\right)_{e_{b1}=E_{bb}}}{1 + R\left(\frac{\partial i_b}{\partial e_{b1}}\right)_{e_{c1}=E_{cc}}}\right]\left(\frac{de_{c1}}{dt}\right)_{t=\pi/2\omega}. \tag{34}$$

From (20) and (34) there results

$$\left(\frac{di_b}{dt}\right)_{t=\pi/2\omega} = \frac{-2\omega g_{m1}E_{gm}}{1 + 2Rg_{p1}}. \tag{35}$$

If serious distortion in the output is to be avoided, then (31) and (35) may be equated, yielding

$$(g_{m1})_{\text{at } Q} = \frac{I_{bm}}{E g_{m1}}\left[\frac{1}{2} + R(g_{p1})_{\text{at } Q}\right]. \tag{36}$$

In order to find the optimum bias, consider Figs. 3 and 4, which represent the tube static transfer and plate characteristics, respectively. If the supply voltage E_{bb}, the load resistance per tube R, and the highest allowable total grid voltage e_{cm} are known, then I_{bm} can be found as in Fig. 4, and point S of Fig. 3 is defined. A trial-and-error procedure can be used as in Fig. 3 until a point Q can be found on the transfer characteristic curve such that the slope at Q is equal to tan ϕ_1 with g_{p1} evaluated at Q.

Load Matching

Suppose that E_{cc} and E_{gm} are prescribed; the composite load line is linear but its slope may be varied, and the output distortion is negligible. From Fig. 4, it follows that the output power can be expressed as

$$P_0 = \tfrac{1}{2}(E_{bb} - e_{b1})i_b.$$

For maximum power output under the above conditions, $dP_o/de_{b1} = 0$, which yields

$$\frac{di_b}{de_{b1}} = \frac{i_b}{E_{bb} - e_{b1}} = \frac{1}{R}. \tag{37}$$

Eq. (37) indicates that the slope at P is equal to the magnitude of the slope of the composite load line. This result does not require assuming linearity for the single-tube or composite characteristics and is equally true for class A, AB_1, and AB_2.

In terms of the tube parameters, (37) implies that: 1) for Class A, since i_{b1} and i_{b2} have slopes with nearly equal magnitudes

$$\frac{1}{R} = g_p = 2g_{p1},$$

or

$$R = \tfrac{1}{2}(r_{p1})_{\text{at } Q}; \tag{38}$$

2) for class AB_1 and AB_2, since at P, $i_{b2} = 0$ when e_{c1} attains the highest value

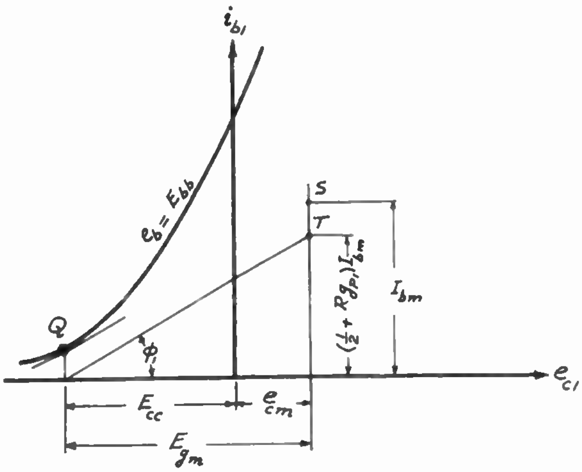

Fig. 3—Static transfer characteristic for $e_b = E_{b1}$.

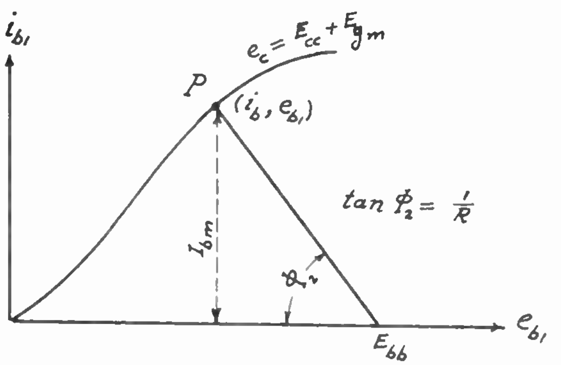

Fig. 4—Static plate characteristic for maximum allowable grid voltage. At point P, $i_b = i_{b1}$.

$$\frac{1}{R} = g_p = (g_{p1})_{\text{at } P}$$

or

$$R = (r_{p1})_{\text{at } P}. \tag{39}$$

Eqs. (38) and (39) are the conditions for matching the load to the tubes for classes A and AB, respectively

Conclusion

If the output transformer is lossless and exactly center tapped and has unity coupling, in the midfrequency range the composite load line approaches a straight line regardless of the nonlinearity of the matched tubes. This conclusion follows from the fact that the integral term of (16) becomes negligibly small.

It can be shown also by (1)–(8) that the composite load line would still be linear at medium frequencies, if the tubes used are both nonlinear and mismatched, provided the output transformer possesses the same properties mentioned above. The mismatching together with the nonlinearity of the tubes, however, gives rise to nonlinear distortion, since the even-ordered partial derivatives of (20) would not vanish identically.

For the derivation of the composite load line and frequency response equations at low frequencies, however, the third and higher odd-order partial derivatives of i_b with respect to e_c and e_b are assumed negligibly small.

Many of the preceding equations, including (16), apply to class B and even class C. However, these two classes are of little use for audio-frequency amplification on account of their generation of excessive nonlinear distortion. (The term class B has occasionally been used in the literature to designate classes AB_1 and AB_2. In this paper, however, class B is meant to designate the operation for which the conduction angle of each tube is 180°.)

For simplicity, the stray and interelectrode capacitances have been neglected. Therefore, the equation describing the frequency response is valid only at low and medium frequencies. This result, however, should be of special interest to high-fidelity designers, since it permits calculation of the transformer primary inductance corresponding to a prescribed low-frequency response.

The content and generality of approach of this paper should pave the way for a parallel quantitative understanding of the transistor push-pull amplifier.

Acknowledgment

The author wishes to express his grateful appreciation to Dr. J. D. Ryder for his very helpful discussions on this paper. The author is also indebted to Dr. M. B. Reed for his most valuable and inspiring suggestions. The very useful comments of Dr. T. Ozker and of L. L. Howard are greatly appreciated.

Phonetic Typewriter*

H. F. OLSON† AND H. BELAR†

Summary—**The important factors involved in the development of a phonetic typewriter are as follows: the particular form in which the words are typed; the means for analyzing the sounds of speech; the identification of the analyzed sounds; the encoding, coding, and decoding of the sounds for the operation of the actuating mechanism; the design of the mechanism for actuating the typewriter. A study has been made of these problems. As a result of this study a simplified model of a phonetic typewriter has been developed incorporating all of these aspects. This model serves to illustrate the principles involved and provides means for further study towards a complete system.**

Introduction

The printing press as developed about 400 years ago made it possible to disseminate information in the form of the printed page. The advent of sound reproduction in various forms made it possible to reproduce sound at the same point or some other point, either at the same time or some subsequent time. Another logical step in the field of communication is the conversion of speech sounds into the corresponding printed words on a printed page. This process involves the use of speech sounds to actuate machinery in accordance with the information carried by the speech. There are many applications for machines that will make it possible for speech to perform certain operations. One of the important and useful machines in this category is a speech typewriter which prints or types on paper the words which are spoken into the machine.

A solution of the problems enumerated above constitutes the development of a phonetic typewriter. Thus, it will be seen that the development of a phonetic typewriter involves many interrelated problems in diverse fields. In a development program of this type it has been found that work on a complete system is often more useful than specialized work on the separate categories. Accordingly, an elementary model of a phonetic typewriter was developed for the purpose of evaluating the problems in the different fields. It is the purpose of this paper to describe this development.

Form of the Typed Presentation

The particular form in which the words are typed is an important aspect of a phonetic typewriter because it involves two interrelated considerations, namely, the complexity of the machine and the fidelity of the typed material. The form in which the words are typed may be divided into many categories, as for example, phonetic, syllables, or words. The phonetic system requires a smaller memory in the machine, but results in an output which requires a knowledge of phonetic symbols similar to the form used in stenography in order to be interpreted. In the other systems, the form may range from recognizable syllables to complete words typed by employing conventional letter symbols. It appears that from practical considerations some form of the latter systems must be used in the phonetic typewriter. The problem then becomes that of deciding upon the form of the presentation from this standpoint.

It is beyond the scope of this paper to present all the considerations with regard to the form of the presentation. Considering both intelligibility, spelling accuracy, and the memory system, the syllable type of presentation shows the greatest promise. In view of this, the phonetic typewriter described in this paper was directed toward the objective of employing and evolving a syllable system.

Description of an Elementary Phonetic Typewriter

The model of the phonetic typewriter which has been developed consists of the following elements: a microphone, an electronic compressor, analyzers, a visual display, encoders, memories, decoders, a typewriter actuating mechanism, and a conventional typewriter.

A schematic diagram of an elementary system which converts the sounds of speech into the corresponding words typed on a page of paper is shown in Fig. 1. The elements of this system will be described in the subsections which follow.

Microphone

The microphone plays an important role in the sound analyzing system because it converts the sounds of speech into the corresponding electrical variations. An important requirement is that this transduction be made with a high order of fidelity. For example, if the microphone introduces spurious components in the conversion from acoustical to electrical variations, the problem of analysis becomes exceedingly difficult if not impossible. The microphone used in the phonetic typewriter is the RCA Type BK-1A dynamic pressure microphone.

Another consideration is the acoustics of the room in which the system is operated. A relatively large value of reflected sound will introduce spurious components in the analyzing system. Therefore, if the sounds of speech are picked up in an ordinary room, the micro-

* Manuscript received by the PGA, April 29, 1957. Sections of this paper appeared in the *Journal of the Acoustical Society of America*. Phonetic typewriter is a term used in this paper to designate a voice-operated machine which types or prints on paper the words spoken into microphone input to the machine.

† RCA Laboratories, Princeton, N. J.

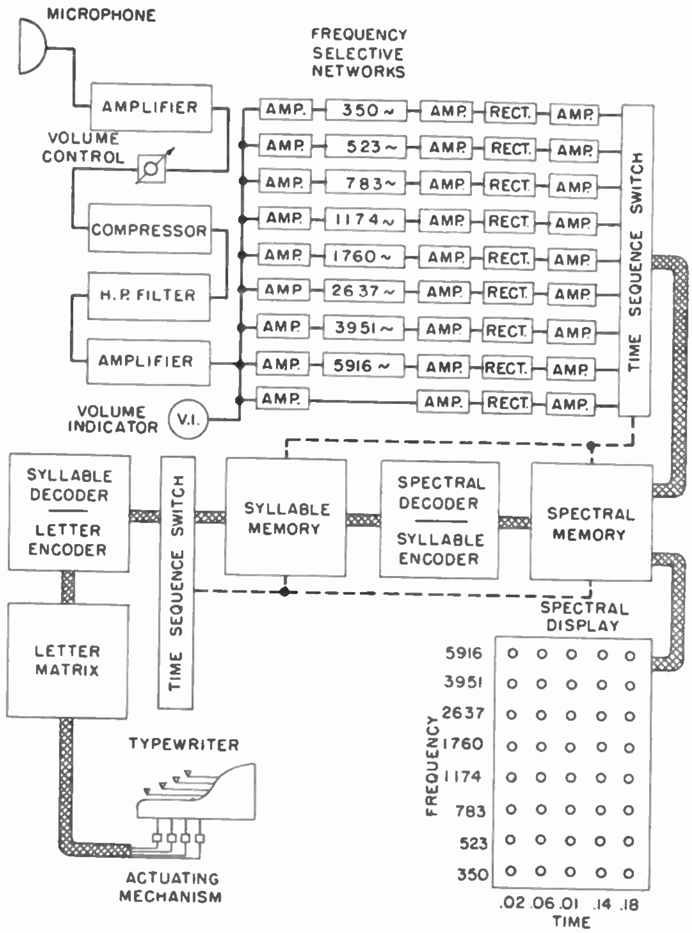

Fig. 1—Schematic diagram of a phonetic typewriter.

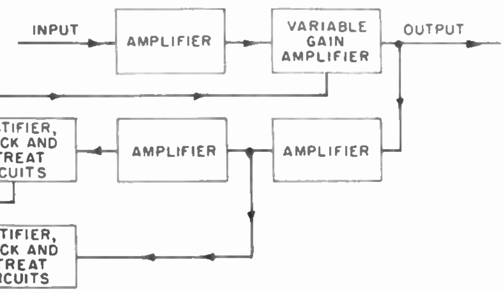

Fig. 2—Schematic diagram of the preamplifier and compressor amplifier.

phone must be operated at a very small distance from the mouth in order to obtain the required ratio of direct to reflected sound.

When a conventional microphone is used at a close range of speech pickup, breath sounds introduce spurious components. Therefore, a special type of wind screen was developed to reduce the effects of breath sounds.

Amplifier and Electronic Compressor

In the course of normal speech, as used in dictation, the amplitude level may vary over wide limits. Since the analysis is a function of the amplitude level, some means must be provided to maintain a relatively constant output level over a wide range of input levels. This can be accomplished by the use of an electronic volume compressor. An electronic volume compressor is a system which reduces the amplification of a vacuum tube amplifier when the signal being amplified is relatively large and increases the amplification when the signal is relatively small. Specifically, the electronic volume compressor employs a vacuum tube amplifier in which the amplification is an inverse function of the general level of the signal. A schematic diagram of the preamplifier and compressor used for the phonetic typewriter is shown in Fig. 2. The electrical signal is fed to the input of the preamplifier. The output of the preamplifier is fed to a two-stage amplifier and rectifier and to a single-stage amplifier and rectifier. The outputs of the two rectifiers are added and applied to the first grids of the output stage of the variable gain amplifier. The tubes in the variable gain amplifier are of the variable transconductance type which means that the amplification of the tube will vary as the voltage applied to the first grid is varied. The output of the rectifiers is a function of the input to the amplifier. The amplification of the amplifier is an inverse function of the bias applied to the first grids of the output stage. Thus, it will be seen, that as the input to the amplifier increases, the negative bias applied to the first grids of the variable gain amplifier will be increased which in turn will decrease the amplification factor of the variable gain amplifier. The net result is a gradual reduction in gain with increase of the input. Typical input-output characteristics of the system shown in Fig. 2 are shown in Fig. 3 It will be seen that a relatively constant output level can be obtained with a wide variation in the input level.

Spectral Analyzer

A system for analyzing and separating the sounds of speech into discrete categories constitutes one of the basic elements of any machine for the control of typing by the sounds of speech. A sound wave may be described in terms of the amplitude and frequency of the components and time. Therefore, the analyzing system may be based upon frequency, amplitude, and time.

The analyzing system is shown in Figs. 1 and 4. As the description will show, it is an amplitude, frequency, and time system. The analyzer follows the compressor amplifier. The input to the analyzer passes through a high-pass filter to an amplifier. The output of the amplifier is fed to the inputs of the amplifiers in channels 1 to 9. The channel amplifiers 1 to 8 are coupled to frequency selective networks. The output of each frequency selective network is followed by another stage of amplification, a rectifier, and a direct current amplifier. The outputs of the direct current amplifiers are used to actuate the relays in the spectral memory system. Channels 1 to 8 inclusive are similar, except for different constants in the frequency selective networks. Each of the frequency selective networks covers a narrow frequency band. The response frequency character-

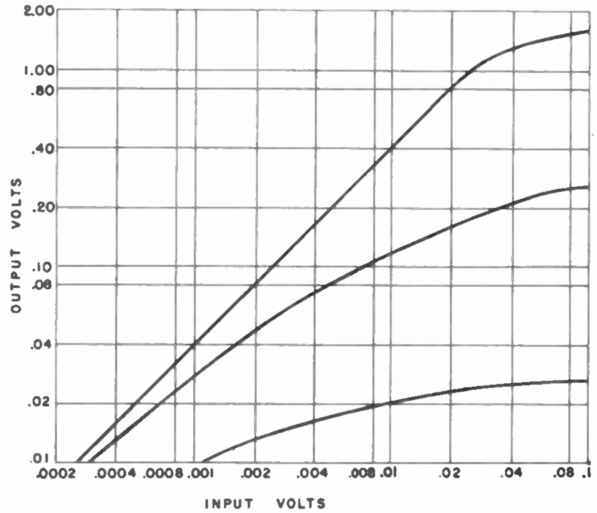

Fig. 3—Input-output characteristics obtained with the amplifier compressor of Fig. 2.

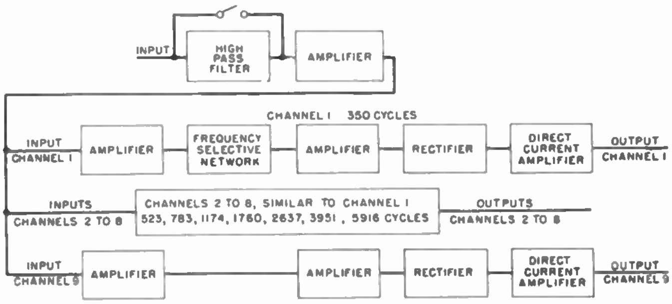

Fig. 4—Schematic diagram of the high-pass filter, amplifiers, frequency selective networks, and direct current output amplifiers.

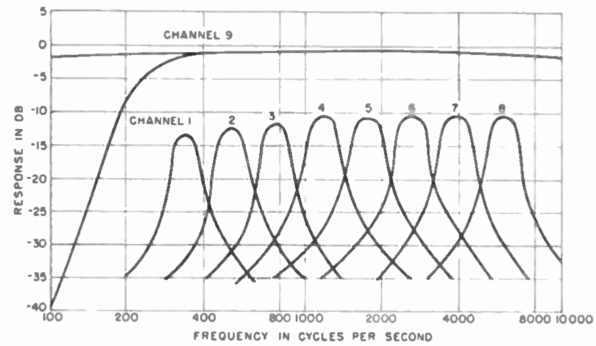

Fig. 5—Response frequency characteristics of the frequency selective amplifier channels 1 to 8 inclusive, the amplifier channel 9, and the high-pass filter.

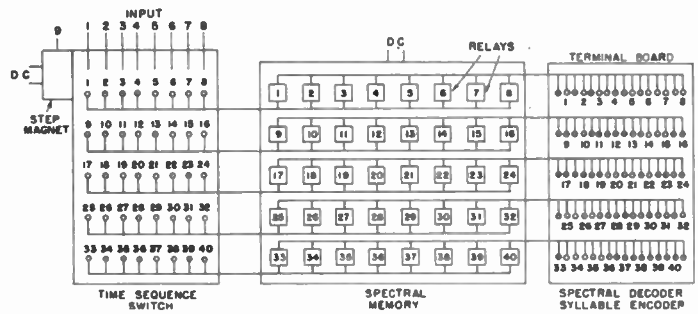

Fig. 6—Schematic diagram of the time sequence switch, the relay memory, the spectral decoder, and syllable encoder.

istics of channels 1 to 8 inclusive are shown in Fig. 5. There are no frequency selective circuits in channel 9 and the response frequency characteristic, as shown in Fig. 5, is practically flat. Channel 9 is used to activate the system. Since there is very little information in speech below 200 cycles, this portion of the range can be eliminated without any appreciable loss. In view of this fact, a high-pass filter is used to reduce the deleterious effects of ambient noises which occur in the low-frequency range. The response frequency characteristic of the high-pass filter is shown in Fig. 5.

Spectral Memory

The outputs of channels 1 to 9 inclusive are fed to the input of the rotary sequence switch shown schematically in Fig. 6. Time is divided into five intervals. The output of channel 9 actuates the stepper in the rotary sequence switch through a relay which sets in motion the ten wiper arms of the switch. The tenth circuit of the switch is used for sequence of operation functions. The wipers make contact with the terminals 1 to 40 inclusive in the time interval 0 to 0.22 second. The wipers stay in contact with terminals 1 to 8 for the time interval 0 to 0.04 second, terminals 9 to 16 for the time interval 0.04 to 0.08 second, terminals 17 to 24 for the time interval 0.08 to 0.12 second, terminals 25 to 32 for the time interval 0.12 to 0.16 second, and terminals 33 to 40 for the time interval 0.16 to 0.20 second. When the wiper arm 1 contacts terminal 1, the coil of the magnet of relay 1 is connected to the output of channel 1; when the wiper arm 2 contacts blade 2, the coil of the magnet of relay 2 is connected to the output of channel 2, etc. If there is sufficient power level available, the armature of the relay will be actuated and the relay will be held in the closed position by the external dc supply until it is released.

Spectral Display

The light display shown in Fig. 1 is connected to the relays of the spectral memory and thereby indicates the relays which have been actuated by the sound. The visible spectral display is useful for analyzing and establishing the code for the system. The use of the light disply in establishing the spectral code will be explained in a later section.

Spectral Decoder and Syllable Encoder

The output of the relay memory is connected to the spectral decoder in the form of a terminal board shown in Fig. 6. By means of the terminal board, the spectral pattern for a particular syllable can be decoded. Furthermore, the syllable can be encoded by means of the terminal board. To summarize: the particular terminals corresponding to the code, as depicted on the light display, can be interconnected on the terminal board and then connected to the input of the syllable memory.

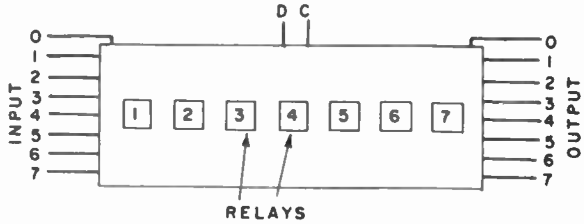

Fig. 7—Schematic diagram of the syllable memory.

Syllable Memory

The syllable memory of Fig. 7 consists of seven relays. Each relay corresponds to a syllable, which means that the memory can store seven syllables. In another experiment, a memory of ten syllables was used. The output of the syllable encoder is fed to the syllable memory. When the syllable corresponding to this code is spoken into the microphone, the corresponding relay in the syllable memory will close and stay closed until it is released following the typing of the syllable. The output of the syllable memory is fed to the syllable encoder.

Syllable Decoder, Letter Encoder, and Letter Matrix

The syllable decoder, letter encoder, and letter matrix consist of a rotary selector switch and a 26×40 matrix (Fig. 8). The output of the syllable memory is connected to the input of the sequence switch. This element of the sequence switch constitutes the syllable encoder. The sequence switch can be connected to the matrix so that up to five letters will be typed for one sweep of the wipers. This element of the sequence switch and matrix constitutes the letter encoder. The ouptut of the matrix is connected to the 26 solenoids of the electrical actuating mechanism of the typewriter.

Typewriter and Actuating Mechanism

The electromagnetic mechanism for actuating each key of the conventional typewriter is of the solenoid and plunger type as shown in Fig. 9. When a current flows in the coil, a force acts upon the plunger which acts to pull it into the coil. A tapered plunger is used in order to obtain a more uniform force over the range of travel of the plunger. The plunger is connected to the key lever of a standard typewriter. In order to speed up the operation of the actuating mechanism, the magnetic return path is through the air. The power input to the solenoid is 50 watts.

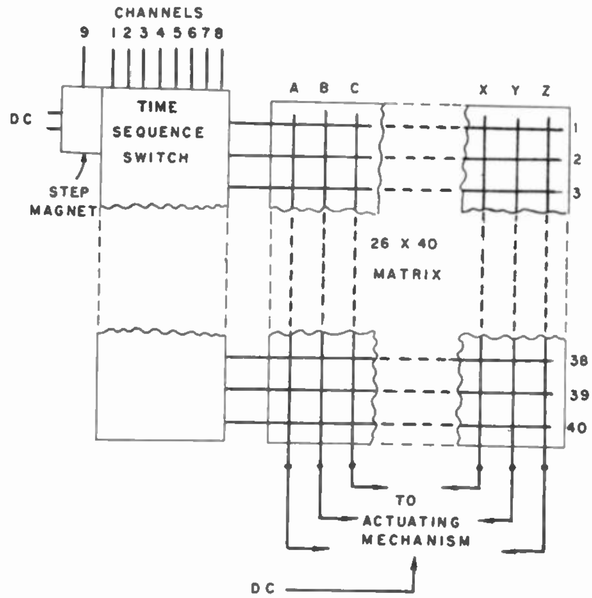

Fig. 8—Syllable decoder, letter encoder, and letter matrix.

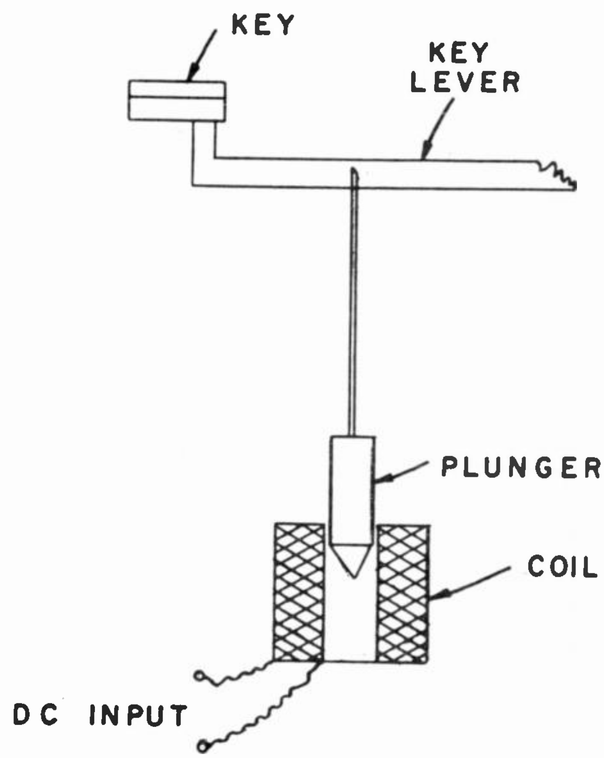

Fig. 9—Key actuating mechanism.

Description of the Operation of the Phonetic Typewriter

The elements of the phonetic typewriter have been described above. The operation of the phonetic typewriter will now be described. To illustrate the operation, it will be assumed that the vowel sound "I" is spoken into the microphone. The output of the microphone is amplified by the preamplifier and the compressor amplifier of Fig. 2 and fed to the frequency selective amplifiers of Fig. 4 and designated as channels 1 to 8 inclusive and the wide-band amplifier designated as channel 9 of Fig. 4. These outputs are fed to the rotary sequence switch of Fig. 6. The output of channel 9 starts the rotary switch when the level of the sound attains a certain specified level. The wipers of the sequence switch start and pass the terminals 1 to 40 in the sequence as outlined in the description relating to Fig. 6. The spectrum display for the vowel "I" will be as shown in Fig. 10. That is, the relays corresponding to these lights will be actuated. The terminal board of Fig. 6

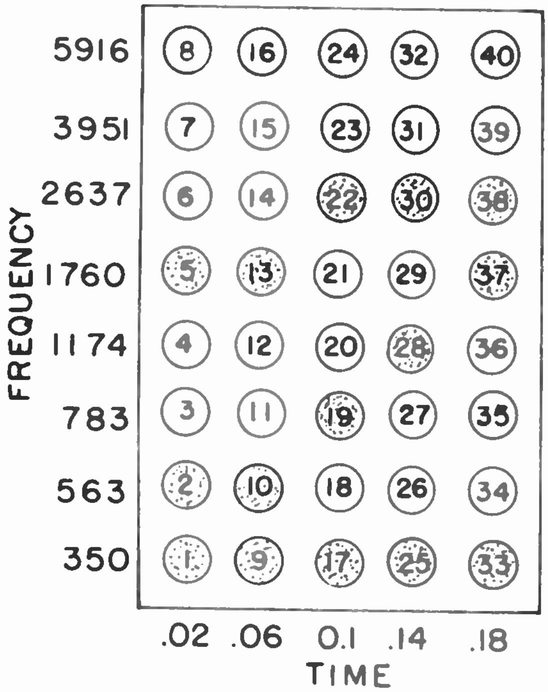

Fig. 10—Frequency-time display of the vowel sound "I" obtained by means of the spectral display section of the apparatus of Fig. 1.

is wired to input 1 of the syllable memory of Fig. 7, so that the armature of relay 1 of the syllable memory will be actuated. Specifically, this is done by connecting the switching circuits of the closed relays 1, 2, 5, 9, 10, 13, 17, 19, 22, 25, 28, 30, 33, 37, and 38, in series with the coil of relay 1 of the syllable memory. The switching circuit of relay 1 of the syllable memory is connected to channel 1 of the syllable decoder of Fig. 8. When the switching circuit of relay 1 of the syllable memory closes, it activates the rotary sequence switch of Fig. 8.

The matrix of Fig. 8 is connected so that the circuit corresponding to the letter "I" will be closed through the actuating mechanism of Fig. 9 which is connected to the key of the letter "I" of the typewriter. A dc current is sent through the solenoid of the actuating mechanism which causes the letter "I" to be typed. Following this, the rotary switch of Fig. 8 returns to the start position, the relays of the spectral memory of Fig. 6 are all released, the relay of the syllable memory of Fig. 7 is released, and the rotary selector switch of Fig. 6 returns to the start position, and the system is ready for the next syllable.

Identification of the Analyzed Sounds

All of the factors involved in the development of a phonetic typewriter have been discussed in considerable detail except the identification of the analyzed sounds. It is beyond the scope of this paper to describe all the aspects of this complex subject. However, the general procedure employed in the identification of the analyzed sounds may be outlined.

The spectral display for the vowel sound "I" is shown in Fig. 10. This sound was used in the description of the

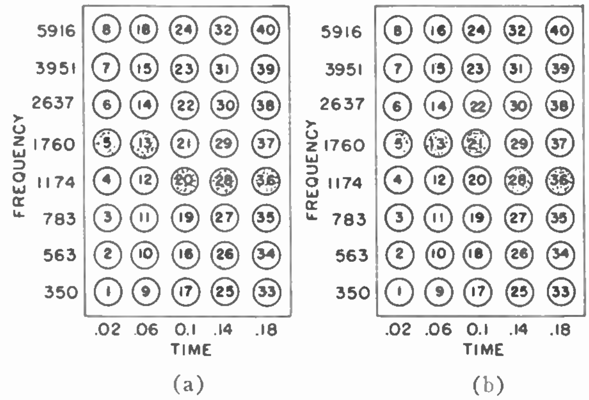

Fig. 11—Two different spectral displays for one part of the spectrum of the same sound.

phonetic typewriter. If it has been established by speaking this vowel sound over, say, 100 times that this is the spectral display, then the connections between the spectral memory and the syllable memory can be made as outlined in the example described in the preceding section.

When the pattern is definite, without any ambiguities, then the identification and corresponding connections are straightforward. The problems in identification and coding arise when the pattern varies. This is illustrated for one part of the spectrum of a sound in Fig. 11. If the time taken for the sound varies, the resultant and corresponding spectral display will be as shown in Fig. 11(a) or 11(b). In Fig. 11(a), relays 5, 13, 20, 28, and 36 are closed. In Fig. 11(b) relays 5, 12, 21, 28, and 36 are closed.

One simple solution to the problem of the connection scheme for the spectral display problem of Fig. 11 is to assume that the sound will be registered on the typewriter for relays 5, 13, 28, 36 being closed. This may well be sufficient because, in general, the sound will have other components which are not shown in Fig. 11 for the sake of simplicity.

Another solution to the problem of the connection scheme for the spectral display of Fig. 11 is to assume that the sound will be registered on the typewriter for relays 5, 13, 28, and 36 all being closed and for either 20 or 21 being closed.

In general, speech sounds consist of more than one frequency band. The vowel sounds contain three or more distinct frequency bands. In these cases, there are many lights in the spectral display that are lighted every time that the vowel sound is spoken. The identification under these conditions can usually be resolved by using only the relays in the spectral display that close every time the vowel sound is spoken.

There is another important identification means for the sounds of speech, namely, the spectral and amplitude characteristics of the starting of a speech sound. The unique property of this part of a speech sound is that it remains practically invariant under almost all conditions.

Fig. 12—Phonetic typewriter.

The identification of the sounds of speech is a problem which must be solved by working with the system. Most of the problems of identification can be resolved after a few hours of analysis of the speech sounds required.

Phonetic Typewriter Assembly

The elementary development model of the phonetic typewriter is shown in Fig. 12. The cabinet houses all the equipment except the microphone and typewriter. The spectrum display, volume control, and volume indicator are visible in the cabinet.

Performance of the Elementary Phonetic Typewriter

Two series of words were selected for use in the performance tests of the elementary version of the phonetic typewriter. For the seven-syllable memory the following words were used: I—can—see—you—type—this—now. For the ten-syllable memory the following words were used: are—see—a—I—can—you—read—it—so—sir. In either of these series it is possible to arrange the words in various permutations to make up a large number of intelligible sentences. When the code is set up for a particular person, quite reliable performance can be obtained if care is used to insure clear enunciation of the words spoken into the microphone. For example, using several hundred series of words, the typing accuracy was such that 98 per cent of the words were typed correctly.

Conclusions and Observations

The simplified version of the phonetic typewriter which has been described will type up to ten syllables. This machine demonstrates and illustrates the principles involved in the electronic and mechanical translation of the sounds of the voice into the corresponding typed syllables. Considerable additional work will be required in the development of an analyzing system which will provide more reliable performance, particularly when the words are not enunciated with care. A phonetic typewriter will probably be a personal type instrument. This simplifies the problem of obtaining reliable performance. Nevertheless, the problems of analysis are still very complex, particularly with regard to the analysis of words which exhibit very small differences in the sounds. Other areas which will require considerable development work involve the memories, codes, and switching systems.

The "Stereosonic" Recording and Reproducing System*

H. A. M. CLARK†, G. F. DUTTON†, AND P. B. VANDERLYN†

Summary—The paper reviews briefly the history of stereophonic reproduction. The principal basic systems with their underlying ideas are described and compared. Some account is given of the supposed mechanism of natural binaural listening from the viewpoint of direction localization.

The principles and practice are discussed of a particular system for domestic use, derived from the early work of Blumlein, and characterized by the use of spaced loudspeakers driven in phase, to which the name "stereosonic" has been given. The aims of this system are defined, and the mathematical theory involved in its use is developed. Limitations and sources of error in the results achieved are described.

Equipment used in the making of master recordings and some of the problems of studio technique involved are described. Consideration is given to the form which a domestic stereophonic record should take, and the standards to which such a record should conform, together with the requirements which these impose on the reproducing equipment.

List of Principal Symbols

θ = Angle between radius vector of the sound source and the median plane.
θ_t = True angle.
θ_a = Apparent angle.
h = Effective human-ear spacing.
v = Velocity of sound in air.
ϕ = Phase difference of sound at the listener's two ears.
ω = Angular frequency.
ψ = Semiangle subtended at the listener by a pair of loudspeakers.
μ = Difference of the time of arrival of the sound at the listener's two ears.
L, R = Peak amplitudes at listener's ears due to left and right loudspeakers.
x = Half-distance between loudspeakers.
r = Projection of the listener's displacement from the median plane onto the loudspeaker base-line, expressed as a fraction of x.
d = Distance between loudspeaker base-line and the listener.
y_a = Fractional distance of apparent source from central position.
y_t = Fractional distance of true source from central position.
β, γ = Angles subtended by the left- and right-hand loudspeaker with respect to the median plane of the listener.
τ = Time-of-arrival difference of the sound from left- and right-hand loudspeakers at the listener caused by path differences.
μ_L, μ_R = Semitime differences at the listener's ears due to sounds from left and right loudspeakers.
ϕ_L, ϕ_R = Resultant phase of the sound at the left and right ears relative to the center position of the listener.

I. The Importance of Stereophony in Sound Reproduction

NORMAL listening is always binaural, and the two ears are used in conjunction by the brain to interpret the sounds heard. If all the individual sound sources, such as voices and instruments, and the reverberant sounds, are recorded or transmitted by a single channel before being reproduced by one or more loudspeakers, it is not possible to use the binaural sense of location to differentiate between the various sounds as is done in direct listening.

In order to add some spatial realism using only one channel, a number of schemes have been tried, such as the use of spaced loudspeakers supplied through frequency-dividing networks, or several loudspeakers fed through delay networks, but none of these is universally successful on all types of program.

An honest assessment of the sound quality obtainable through a single channel, in which the frequency, harmonic, intermodulation, and noise distortions have been reduced to the limits detectable by the human ear, reveals that the reproduction is still far from lifelike. The addition of real binaural listening conditions seems to be the most likely factor remaining.

Many tests have now been made which all show that, if a system is used in which the relative sounds at the two ears are each reproduced as similarly as possible to those heard in direct listening, *i.e.*, with the correct relative amplitude and phase differences for each sound source, the reproduced sound takes on a new quality which is nearer to realism than can be obtained in any other way.

For a given amount of "ambient" or reverberant sound, it is possible to obtain a better definition, *i.e.*, ability to distinguish individual instruments or voices in a chorus. A soloist can be heard clearly through an accompaniment, without having to resort to making the soloist appear unnaturally close to the listener.

The problem of hearing clear bass parts free from the unduly long low-frequency reverberation which is so

* Manuscript received by the PGA, June 24, 1957. Reprinted with corrections from *Proc. IEE*, vol. 104, pt. B; 1957.
† Electric and Musical Industries, Ltd., Hayes, Middlesex, Eng.

commonly encountered, is considerably eased, and full bass response without "boom" can be obtained. It is noticeable that the reproduced sound level can be raised higher than with single-channel reproduction before the average listener becomes irritated. The ability to do this permits a closer reproduction of the original dynamic range of the music in cases where this is great. It also seems possible to obtain a given subjective loudness for less acoustic power than with single-channel operation.

All this is quite aside from any benefits which may accrue from the ability to reproduce illusion of movements and the spatial relationships so essential to drama and opera.

The requirements then seem to be to reproduce, with reasonable economy, in a system suitable for home use, the same sounds at each ear separately which would have been heard by direct listening in an optimum position in front of the performers. This requirement can be limited to sounds arriving from directions covering an angle of, say, 90° in front of the observer for most practical purposes. No method as yet has been devised to meet this requirement from all horizontal directions, but neither this nor the need for the indication of vertical angular direction seems to be of major significance.

II. History

An explanation of the principles by which the normal human being uses both his ears for the purpose of spatial location has been sought since the classic studies in acoustics of the mid-nineteenth century. Attempts to reproduce this natural ability at a distance were made early in the development of telephony. One of the earliest accounts was given in *L'Électricien* of 1881 by E. Hospitalier, who described the sound system installed in the Paris Opera House.[1] This basically consisted of a pair of microphones spaced on opposite sides of the stage, which supplied at a distance a number of pairs of telephones held to the ears of the listeners, who received an impression of localization.

It was probably not for many years, until sound reproduction for entertainment purposes was receiving more attention, that any serious development work along these lines was undertaken. It was well understood, however, that the use of two similar microphones spaced by a distance equivalent to that between a human being's two ears, operating through two separate channels into a pair of headphones, did, in fact, give a remarkably accurate impression of spatial location. The use of headphones nevertheless robs the observer of the ability to locate sources outside the horizontal plane and prevents discrimination between front and rear, since, in these cases, it is necessary to move the head relative to the sound field.

In the 1920's it became customary to use loudspeakers for sound-reproducing systems instead of head telephones, and the problem of reproducing sound with directional sense immediately became more complex, for reasons which are discussed later in the paper.

One of the earliest workers in this field was the late Blumlein. He was convinced at the time that the main contribution towards the sense of spatial location was provided by the difference in time of arrival at the two ears which, at low frequencies, can be interpreted in terms of phase differences.

He invented a system[2] by which the phase difference between the outputs of two similar microphones spaced by a distance, small, compared with the wavelength, can be converted into two in-phase outputs of different amplitudes driving two loudspeakers spaced by an appreciable distance. If the relative amplitudes corresponding to a given phase difference at the microphones are correctly proportioned, the phase differences at the two ears of the person listening to the two loudspeakers, will be a close simulation of the phase differences which the listener would have heard had he been at the microphone position.

Blumlein showed how the correct relative loudspeaker inputs could be derived from two closely spaced pressure microphones by the use of suitable modifying networks which he called "shuffling" networks. He also showed how velocity microphones could similarly be used, and solved the mathematical relationship between the various parameters in order to obtain the correct apparent position of the source with any given loudspeaker arrangement. In addition, he described a means of recording two channels on a single disk, and, in fact, disk records of this type were pressed and reproduced in 1933. Experimental motion-picture films were made incorporating twin-track optical recordings, and these were reproduced along with large-screen projection in 1935. The industry was not prepared to introduce the wide screen at this date, and therefore further work on the application was abandoned.

Blumlein's method has the advantage of being a "free field" system, the observer being able to move his head without producing a corresponding movement of the apparent source. It relies on the interaction of sounds from the two loudspeakers at the listener's ears, and cannot utilize headphones.

One of the first successful public demonstrations of stereophonic sound reproduction was given by Bell Laboratories in America in 1934.[3] The system used three microphones at the center and extremities, respectively, of the sound stage, and they drove, through three identical channels, three loudspeakers similarly disposed on the reproducing stage. Using such a three-channel system, fairly accurate location was obtained in the azimuth sense, and comparatively effective location was possible in depth. In 1939, RCA used a similar ar-

[1] E. Hospitalier, "The telephone at the Paris opera," *Sci. Amer.*, vol. 45, p. 422; December 31, 1881.

[2] A. D. Blumlein, British Patent No. 394325.

[3] H. Fletcher, "Auditory perspective," *Bell Sys. Tech. J.*, vol. 13, pp. 239–244; April, 1934.

rangement with three channels for the stereophonic recording of a sound film entitled *Fantasia*.

Since the war, demonstrations have been given by Philips at Eindhoven, of a system using two microphones placed in an artificial head, the outputs of which supply two widely-spaced loudspeakers. In such a system, the phase differences of the microphone outputs become of no particular significance when reproduced at the two spaced loudspeakers, but the relative amplitudes of the microphone outputs due to masking by the artificial head are reproduced at the loudspeakers, and provide a measure of spatial location to an observer in front of them. Such an effect, of course, can take place only at the upper frequencies, for example, those above about 700 cps, and any directional effect at the low frequencies is neglected.

More recently, several American companies have made recordings using two widely-spaced microphones driving spaced loudspeakers via the medium of twin-track magnetic tapes, and one has issued some disks carrying the two channels, one on the outer half and the other on the inner. This has the serious disadvantage of halving the playing time.

The advent of magnetic tape, providing two synchronous but otherwise independent channels, made possible the commercial exploitation of Blumlein's work, resulting in the "stereosonic" system described in this paper.

This was demonstrated to members of the profession and representatives of the Press in April, 1955. In April, 1956, a full-scale public demonstration at the Royal Festival Hall was given to an audience of 1800.

III. Comparison of Basic Systems

In the preceding brief historical review, the systems employed fall into three basic types. These are described below.

A. The "Wavefront" System

If an infinite number of microphones, placed in a vertical plane between the source of sound and the listener, were to be connected to an infinite number of loudspeakers in corresponding positions, clearly the radiated wave could be reproduced unaltered and true binaural audition would be preserved.

A line, rather than an area, of microphones and loudspeakers would give perfect location in a horizontal direction, which is adequate for most purposes since location in a vertical plane seems to be impossible except by inclining the head.

An approximation to this condition, with a corresponding limitation in the accuracy of the results, can be obtained using a finite number of microphones and loudspeakers. Contemporary film sound systems use several channels in this manner. The system described by Fletcher also used this principle. Fig. 1 shows the practical arrangement.

Detailed tests were made with a loudspeaker located

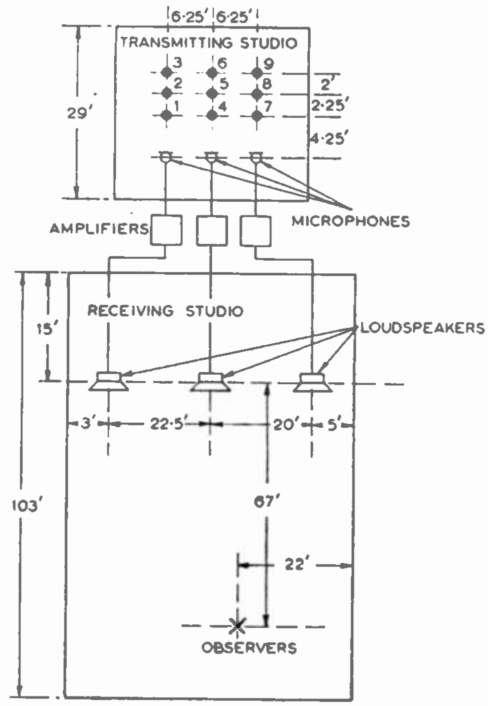

Fig. 1—Arrangement used by Bell Laboratories in 1934.

at one of nine positions in the transmitting studio, and the apparent position as estimated by observers in the listening studio corresponded fairly accurately in breadth, and to a reasonable degree in depth, when the observers were not too far from the center line. Very high-fidelity channels were used, and the quality of results obtained by this system was said to be better than anything ever heard prior to that date.

The use of three independent amplifier and loudspeaker channels, although suitable for public use, is quite uneconomical for domestic use. The ultimate simplification of the wavefront system to two channels can, under certain conditions, give pleasing results. For this purpose, two microphones are placed about ten feet apart in front of the sound source, driving two identical amplifying and recording or transmitting systems with two loudspeakers at a similar separation. There is a decided tendency, however, for the sounds to appear to be coming from the two separate speakers, with a gap in between in which the sound is weak.

B. The Reproduction of Correct Sound Pressures at the Ears

A somewhat different approach to that of the wavefront conception is to consider local conditions at the observer's two ears. A system has been devised which relies on a consideration of these differences. The approach is based on the assumption that low-frequency sounds play little or no part in directional localization and that the principal clue to direction is the intensity difference at high frequencies produced by the shadowing effect of the head. It is shown, if two pressure microphones are used as artificial ears in a dummy head, intensity differences between the outputs occur in accordance with the classic measurements described by Fletcher. It is demonstrated then, that if these outputs are applied to a pair of widely-spaced loudspeakers placed symmetrically with respect to an observer, intensity differences do occur at his ears, although they are not as great as the original differences at the microphones. In connection with this system, attempts have been made to show that time differences are less effective than intensity differences in producing impressions of directional localization, and that incorrect time differences can be compensated by modifying the intensity differences. Then it is argued, that since the major factor in localization, *i.e.*, intensity differences at high frequencies, is reproduced, natural binaural listening is simulated.

The consideration of local conditions at the ears forms the basis of the "stereosonic" system, described in detail in Section IV. This springs from the original Blumlein invention. It recognizes that, since each loudspeaker communicates with both ears, differences in magnitude of the sound pressures at the loudspeakers at low frequencies produce phase and not magnitude differences at the ears, since the contributions from the two loudspeakers arrive at slightly differing times. A pair of directional microphones is employed effectively at a single position to produce two outputs in phase, differing in amplitude according to the direction of the sound source. These are applied to a pair of spaced loudspeakers so as to produce at the ears a time difference independent of frequency at low frequencies, and an intensity difference using the shadowing effect of the head at high frequencies. It is claimed that this represents the nearest approach yet made to natural listening conditions.

C. *Pseudo-Stereophonic Systems*

No detailed description will be given of systems in which a single microphone output supplies two or more recording channels at relative levels which are controlled manually in the dubbing stage. Such means are used in the dialog sequences of wide-screen motion pictures and give an illusion of movement to single voices, etc., but clearly cannot give a simultaneous representation of the direction of a large number of sources as is required for most musical programs.

IV. Theory of the "Stereosonic" System

Before discussing the theory in detail, it will be advisable to review briefly what is known of the mechanism of spatial auditory location in the human being.

A. *Mechanism of Angular Localization*

Rayleigh[4] in 1896, and Stewart[5] in 1920, both carried out experiments which demonstrated that intensity differences at the ears were insufficient to account for location at the lower frequencies, and that the phase differences had to be taken into account, although above about 1000 cps the intensity differences were necessary to avoid ambiguity. In spite of this and the work of others, there has been some reluctance in many quarters to accept the importance of phase differences.

Banister[6] in 1931, and the Medical Research Council[7] in 1932, gave prominence to the idea that the time difference rather than phase difference may be the element detected, in which case there is no need for limitation of the effect to the lower frequencies.

In spite of this, and much other and more recent work, no exact explanation can be given yet for the mechanism of sound location. The authors believe that the two principal quantities used by an observer to estimate the angle of arrival of a sound wave are first, the difference in time of arrival of a wavefront, and second, the difference in intensity at the two ears. Of these, by far the more important is the time difference. For sinusoidal waves, a constant time difference at the two ears is equivalent to a relative phase difference proportional to the frequency of the original sound. At low frequencies, where this phase difference is less than, say, π radians, the direction of arrival could be deduced from it. For such a deduction to be possible, the times of passage of each wavefront past both ears must be identifiable to the observer. As the frequency is increased, *i.e.*, the wavefronts follow each other more closely, the point is eventually reached when the wavefront arrives at one ear before the preceding one has reached the other. Since there is nothing to distinguish between the wavefronts, ambiguities arise and it is impossible to interpret the observed time differences uniquely in terms of direction of arrival. These ambiguities start to occur when the ear spacing becomes equal to a half-wavelength. At higher frequencies the head becomes an appreciable obstacle and produces an intensity difference, the magnitude of which can be used to assess the direction of arrival. Also, it may assist in resolving the ambiguities mentioned, and allowing the observed time differences to be interpreted at higher frequencies. Recent work in America[8] shows that the direction of arrival of pure tones can be judged unambiguously with reasonably constant accuracy up to 1200 cps. At this frequency the ear spacing approximates to a whole wavelength, suggesting that the first ambiguity due to confusion of wavefronts is overcome, possibly with the assistance of intensity differences. In this case, the brain has to choose between two possible directions, widely separated and on opposite sides of the head.

Where the sound waveform is complex, as in the majority of natural cases, the shape of the modulation en-

[4] Lord Rayleigh, "Theory of Sound, Vol. II," Macmillan Co., New York, N. Y., p. 440.
[5] G. W. Stewart, "Binaural location of pure tones—I and II," *Phys. Rev.*, vol. 15, pp. 425–445; May, 1920.
[6] H. Banister, "The basis of sound localization," Physical Society Discussion on Audition, London, June, 1931.
[7] J. H. Shaxby and F. H. Gage, "The localization of sounds in the median plane," Medical Res. Council Special Rep Series No. 166, H.M. Stationery Office, London, Eng.; 1932.
[8] T. T. Sandel, D. C. Teas, W. F. Feddersen, and L. A. Jeffress, "Localization of sound from single and paired sources," *J. Acoust. Soc. Amer.*, vol. 27, pp. 842–852; September, 1955.

velope probably assists in identifying the wavefronts and making possible the interpretation of observed time differences.

With pure tones at high frequencies, when it is necessary to rely on intensity differences to judge direction, accuracy deteriorates, since the amplitude differences at the ears, characteristic of the source position, may be modified by stationary waves and reflections from walls and other obstacles.

1) Cochlear Response: When a sound wave arrives at the ear, the immediate result is the production of an electrical waveform corresponding to that of the instantaneous sound pressure. This is known as the "cochlear response," and it can be detected by amplifying the potential differences between suitably placed electrodes. The signal, in this form, is unsuitable for analysis by the brain, and its primary purpose seems to be to initiate an electrochemical response in adjacent nerve fibers containing the original information in pulse coded form.

2) Action Potential: This secondary signal is called the "action potential" and it differs strikingly from the cochlear response. It consists of short pulses, of constant duration and amplitude, apparently occurring at random intervals. The average frequency of these pulses is related to the intensity of the original stimulus rather than its frequency. This latter is determined in some other way, possibly by observing which part of the basilar membrane is responding, although it must be admitted that pitch discrimination is more accurate than can be accounted for by simple resonance. Observations on single nerve fibers[9] show that successive pulses are always separated in time by an integral multiple of the period of the stimulus and that they occur at a particular point on the cochlear response waveform.

When a pulse has been initiated the particular fiber in which it occurs remains inactive for a short interval, known as the "refractory period," which varies according to the strength of the stimulus. This mechanism limits the pulse rate in a single fiber to a few hundred per second. There are, however, many such nerve fibers associated with each ear, and it is known that at least one fiber will respond at each cycle of the incoming stimulus, up to a frequency of 1500 cps, perhaps higher.

Therefore, provided that the brain can recognize pairs of pulses produced at the two ears by the same sound wave, the original time-difference information is available to it.

3) Time-Comparison System: It is clear that, if the above reasoning is correct, each cycle of the incoming wave gives rise to a pulse of action potential up to frequencies of about 1500 cps, and that such pulses occur at a definite point in the cycle. This happens at both ears, and if there were some means of measuring the time differences between their production, the angle of arrival

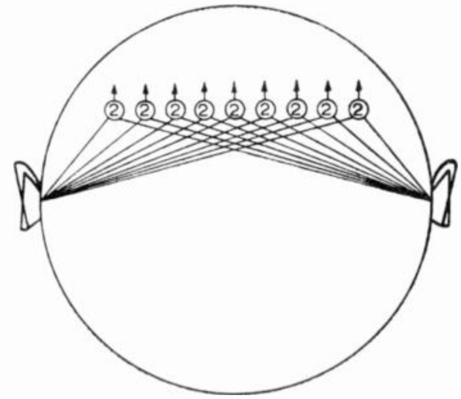

Fig. 2—Schematic of possible time-comparison system.

of the sound could be deduced. A theory has been put forward by Jeffress[10] which suggests how this is done.

It must be noted that transmission of a pulse along a nerve fiber is not simple electric transmission like that along a telephone cable. It is electrochemical in nature; pulses travel without attenuation as they are self regenerating, and at a relatively slow speed. The rates of transmission as measured do not exceed about 10^4 cm, *i.e.*, several times slower than the velocity of sound in air.

The theory then postulates the existence of a nerve combination such that one nerve requires to be stimulated by two others simultaneously before it will respond. Engineers will recognize this combination as being analogous to the "logical-and gate" ubiquitous in computer circuits. It is suggested that a number of these nerve combinations are spread across the brain and stimulated by pairs of nerve fibers of appropriate length, connected to the two ears. Fig. 2 shows this schematically, in engineering terms, remembering that transmission times are proportional to the physical lengths of the different connections. A response from one of these "gates" corresponds to a pair of pulses separated by a definite time interval arriving from the two ears. Since the lengths of the nerve fibers are significant, it is clear that the position of the nerve combination in the brain associates it with a definite time interval between stimuli, and thus with a definite direction of arrival of the original sound. In the figure, a response from the left-hand gate corresponds to a sound wave arriving from the right side of the observer and vice versa.

Some evidence in support of this theory has recently been published. Experiments on cats are described[11] in which clicks, separated by a known time interval, were supplied to the two ears independently and the response on the left and right lobes of the brain observed. Maximum response was obtained from the right lobe when the click supplied to the left ear was advanced by a

[9] R. Galambos and H. Davis, "The response of single auditory nerve fibres to acoustic stimulation," *J. Neurophys.*, vol. 6, pp. 39–57; 1939.

[10] L. A. Jeffress, "A place theory of sound localization," *J. Comp. Phys. Psych.*, vol. 41, pp. 35–39; 1948.

[11] M. R. Rosenzweig, "Cortical correlates of auditory localization and of related perceptual phenomena," *J. Comp. Phys. Psych.*, vol. 47, pp. 269–276; August, 1954.

time corresponding to a sound coming from the left of the subject, and vice versa.

4) Application of the Theory: The physiology of the brain is insufficiently known to confirm or deny the existence of such a mechanism as that postulated, nevertheless its extreme simplicity makes the idea attractive. The theory forms a useful working basis, and many of the observed facts can be fitted into it. It demonstrates clearly the importance of the first wavefront to arrive at the ears; this has the best chance of triggering a pulse, since the inhibitory mechanism is inoperative. The difficulty of localizing pure tones at high frequency is also shown. Pairs of pulses originating from different cycles of the incoming wave will produce spurious responses from the gates, which the brain will have to discard. It is thought that, under these conditions, intensity differences play an important part. With complex high-frequency waveforms, however, pulses will tend to be grouped round prominent features in the waveform, and this will again enable the time comparison mechanism to operate.

Altogether, the best chance of a clear direction seems to be at those lower frequencies where the action-potential pulses can only be paired in one way to give a time difference corresponding to a possible angle of arrival. This does not mean that the most accurate indication will be obtained at the lowest frequencies. The slower rate at which the waveform crosses the zero-pressure axis and the higher threshold, as shown in the Fletcher-Munson curves, combine to produce a loss of resolution that appears as an error in estimated direction. Nevertheless the sense of direction is very strong at low frequencies, and the authors do not subscribe to the view that low frequencies play no part in angular localization.

B. The Aim of the "Stereosonic" System

The aim of the "stereosonic" system is to reproduce at the ears of the listener, in as large an area as possible in front of a pair of loudspeakers, the same vector sound pressures as he would have experienced by direct listening in a corresponding position in front of the sound stage. In other words, although the spacing between the loudspeakers must, in general, be less than the width of the original sound source (except for some solo instruments and unaccompanied solo voices), if the apparent angular width of the reproduced sound is the same as that subtended at the optimum position for direct listening, the listener will imagine that he has been conveyed to the optimum position for direct listening in the recording studio. So far as is known, the ratio of reverberant to direct sound provides the only clue to distance of the source, except for any possible change in quality due to absorption of high frequencies at a distance. In general, the reverberation of the average domestic room is small compared with that of concert halls used for large musical combinations, and so the sense of distance thus conveyed is not substantially altered.

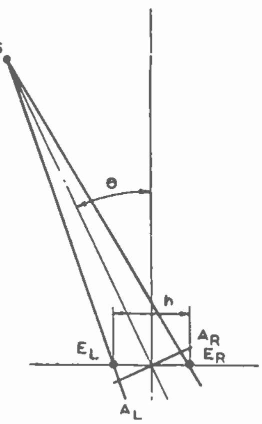

Fig. 3—Time differential of sound at two ears from oblique source.

It is impossible to achieve this aim with perfection, but an attempt is made in the system to reproduce the actual relative phases; thus, the interaural time differences at the lower frequencies and the relative intensities at the upper frequencies.

It should also be clearly understood that the system is intended to operate only with spaced loudspeakers and thus transmit the spatial effect when listening in free space. The system will not function correctly if the two channels are connected to the left and right headphones.

C. Mathematical Theory

1) Vector Pressures at Ears in Direct Listening: Fig. 3 shows an actual source of sound S before two ears E_L and E_R, spaced by a distance h in such a position that its direction is at an angle θ to the face-on position. The sound to E_R, however, will travel a distance E_RA_R further than, and that to E_L a distance E_LA_L less than, the average. If v is the velocity of sound in air, the sound will take a time $(h \sin \theta)/2v$ to travel from A_L to E_L. Thus, the time interval between the arrivals of sound at the two ears will be $(h \sin \theta)/v$. Hence, if h is small compared with the distance from the source, the magnitude at each ear will be the same, but there will be a phase difference

$$\phi = \frac{\omega h \sin \theta}{v} \tag{1}$$

where ω is 2π times the frequency of the sound wave.

At higher frequencies, when the wavelength is short compared with h, the phase angle will be large and ambiguous but the time delay will be the same. Owing to the masking effect of the head, the magnitude is not subject to accurate calculation but has been determined experimentally.[12,13]

[12] F. A. Firestone, "The phase difference and amplitude ratio at the ears due to a source of pure tone," *J. Acoust. Soc. Amer.*, vol. 2, pp. 260–270; October, 1930.

[13] L. J. Sivian and S. D. White, "Minimum audible sound fields," *J. Acoust. Soc. Amer.*, vol. 4, pp. 288–320; April, 1933.

If these phase differences and magnitudes can be simulated in reproduction, the received sound will appear to come from the same angle θ.

2) Reproducing System: For domestic purposes it is generally permissible to use two loudspeakers only, and they must be operated in such a way as to reproduce as nearly as possible the required vector pressures at the ears. Since the bulk of acoustic energy in music and speech occurs below about 700 cps, it is important that the system should operate effectively at low frequencies.

In Fig. 4, W_L and R_R represent two similar loudspeakers supplied with inputs which are in phase but of different magnitudes. The loudspeakers subtend an angle 2ψ at a centrally placed listener whose ears E_L and E_R are separated by a distance h, which is assumed to be small compared with the loudspeaker distance. From (1) the phase difference between E_L and E_R for sound coming from either speaker will be $2\omega\mu = (\omega h \sin \psi)/v$, where 2μ is the difference in time of arrival at E_L and E_R.

Let the average instantaneous sound pressure at both ears be $L \sin \omega t$ from W_L and $R \sin \omega t$ from W_R. The sound pressures at each ear can be calculated as follows:

	Average	At E_L	At E_R
From W_L	$L \sin \omega t$	$L \sin \omega(t+\mu)$	$L \sin \omega(t-\mu)$
From W_R	$R \sin \omega t$	$R \sin \omega(t-\mu)$	$R \sin \omega(t+\mu)$

Total pressure at

$$E_L = L \sin \omega(t + \mu) + R \sin \omega(t - \mu)$$
$$= \sqrt{(L^2 + R^2 + 2LR \cos 2\omega\mu)}$$
$$\cdot \sin \left[\omega t + \arctan \left(\frac{L - R}{L + R} \tan \omega\mu \right) \right].$$

Similarly, the total pressure at E_R is

$$\sqrt{(L^2 + R^2 + 2LR \cos 2\omega\mu)}$$
$$\cdot \sin \left[\omega t - \arctan \left(\frac{L - R}{L + R} \tan \omega\mu \right) \right].$$

Hence the phase difference between E_L and E_R is ϕ_2, where

$$\phi_2 = 2 \arctan \left(\frac{L - R}{L + R} \tan \omega\mu \right)$$

$$= 2 \arctan \left(\frac{L - R}{L + R} \tan \frac{\omega h \sin \psi}{2v} \right).$$

When $\omega\mu$ and $\phi_2 \ll \pi/2$

$$\phi_2 = \frac{L - R}{L + R} \frac{\omega h \sin \psi}{v}. \tag{2}$$

Thus, if the loudspeakers are supplied in phase with correct relative amplitudes, the phase difference ϕ_2 at the ears can be made to be the same as that from a sound source at an angle within $\pm\psi$. In addition, if L or R is permitted to become negative, *i.e.*, to have a phase reversal, the apparent direction lies outside the limit of $\pm\psi$ at frequencies where this analysis is valid. This phenomenon is not readily observable in practice, possibly because it is accompanied by a certain degree of amplitude cancellation. It has, however, been observed under laboratory conditions.[8]

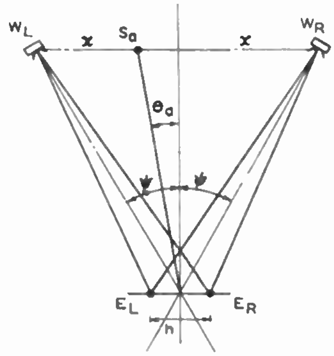

Fig. 4—Apparent source deduced from the time differential of sound produced at two ears by two loudspeakers operating at different sound levels.

3) Microphone Systems. When Blumlein's first experiments were carried out in 1929–30, the only readily available microphones were pressure-operated and had substantially circular polar diagrams. In order to obtain the required inputs to the loudspeakers, the first arrangement to be used consisted of two such pressure microphones separated by about 8 inches (*i.e.*, typical distance between ears). The outputs were, therefore, the same in magnitude, for any source angle, but there was a phase difference given by (1). To convert this into the required amplitude difference demanded by (2), an ingenious circuit was used in which the two microphone outputs were first summed and differenced to produce two new voltages in quadrature. The difference voltage was integrated, *i.e.*, multiplied by a factor proportional to $1/\omega$ and rotated by 90°. The sum voltage was uniformly attenuated by a suitable amount. These two modified voltages were again summed and differenced, giving two final voltages, in phase but of different amplitudes, which it can be shown are of such values that, when applied to two loudspeakers, will, according to (2), produce at the ears the same phase angles as those at the microphones.

Fig. 5 represents two velocity microphones, *e.g.*, ribbon type, with their axes of maximum response at 90°. Such microphones have a response which is proportional to the sine of the angle between the source and the plane of the ribbon. A source S_t is assumed to be at an angle θ_t from the median axis of the microphones. The outputs from the microphones will be in phase at all values of θ_t if the microphones are placed on the same vertical axis but their outputs L and R are given by

$$L \propto \sin (45° + \theta_t)$$
$$R \propto \sin (45° - \theta_t)$$

whence it can be shown that

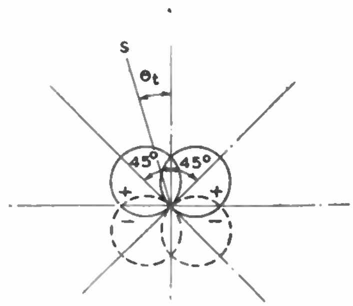

Fig. 5—Polar characteristics of a pair of velocity microphones at 90°.

$$\frac{L - R}{L + R} = \frac{\sin \theta_t}{\cos \theta_t} = \tan \theta_t. \qquad (3)$$

4) Performance of Complete System at Low Frequencies: If the outputs of the microphones are connected (after suitable amplification) to the loudspeakers, the resultant phase difference between the sounds at the ears will be

$$\phi_2 = \frac{L - R}{L + R} \frac{\omega h \sin \psi}{v} \qquad (2)$$

$$= \tan \theta_t \frac{\omega h \sin \psi}{v}$$

from (3).

The ears will interpret this as an apparent sound source at an angle θ_a, where $\sin \theta_a = \phi_2 v / \omega h$, from (1); *i.e.*

$$\sin \theta_a = \tan \theta_t \sin \psi. \qquad (4)$$

Fig. 6 shows the apparent angle plotted against the true angle for various values of ψ. It will be seen that, when the listener is at such a distance that the loudspeakers subtend an angle of 120°, the apparent angle is very close to the true angle up to ±35°. When the listener is at other distances from the loudspeakers, but still on the center line, the apparent source remains at the same fraction of the total loudspeaker spacing, thus showing that a correctly proportioned sound picture is presented although the angular scale has been altered. The angular distortion is such that the apparent source appears to be somewhat nearer the center than it should be over most of the range.

5) Performance at High Frequencies: If the velocity microphones have a constant polar characteristic at all frequencies, the ratio of voltages applied to the loudspeakers is independent of frequency. At frequencies above about 700 cps, however, the foregoing analysis will not hold, since the phase angle between the pressures at E_L and E_R will be ambiguous. At these frequencies the masking effect of the head, however, causes the left ear to be more affected by the left loudspeaker than by that on the right and vice versa, and it easily can be demonstrated experimentally that directional information can be conveyed at the higher frequencies.

Subjective tests were made with a number of observers using two loudspeakers supplied with known

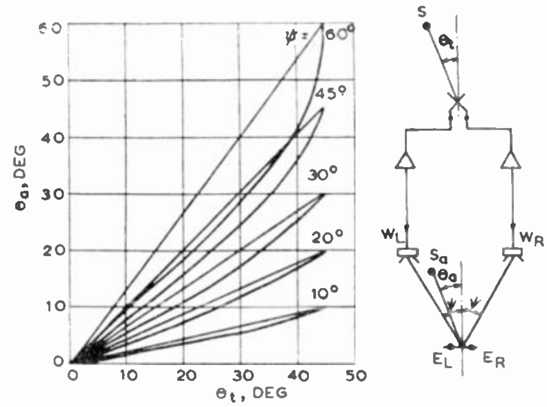

Fig. 6—Apparent angle plotted against true angle for velocity microphones at 90°.

relative voltages from a source of recorded music. First, a filter was inserted passing all frequencies up to 700 cps from a variety of sources of sound including male and female speech, solo, orchestral, and brass-band music. The experiment was repeated using all frequencies above 600 cps. Quite definite location within about ±2° was obtained in each case, but whereas at low frequencies the angle was in agreement with that predicted from (1) and (2) for a given loudspeaker ratio, that obtained at high frequencies was greater. The relationship obtained is in agreement with that published by other workers who rely primarily on intensity differences.[14] By introducing a factor of approximately 0.7 into the ratio $(L-R)/(L+R)$ above 700 cps, the results for high and low frequencies can be brought into line, except for extreme positions of the source.

D. Basis of Recording System

To meet the requirements set out in Section IV-B, the microphone system can consist of a pair of velocity microphones with mutually perpendicular axes of maximum response. The outputs from these microphones can be summed and differenced and the difference channel subjected to a loss of 3 db at all frequencies[15] above 700 cps. Then these voltages are summed and differenced again and used to supply two identical loudspeaker channels. If the loudspeakers are equidistant from the observer and subtend an angle of approximately 90° at him, he will receive substantially correct directional information at all frequencies within the spectrum. At greater distances from the loudspeakers the listener will observe the same apparent position of the source, relative to the loudspeakers.

An important feature of a system using two coincident velocity microphones at 90° is that the rms sum of the outputs of the loudspeakers is constant for a source at a constant distance from the microphones, regardless of its direction. It is this feature which enables the system to reproduce a uniform sound field between the

[14] K. De Boer, "Stereophonic sound reproduction," *Philips Tech. Rev.*, vol. 5, pp. 107–114; April, 1940.
[15] P. B. Vanderlyn, British Patent Application No. 23989, 1954.

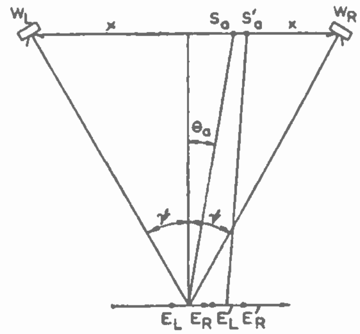

Fig. 7—Change of position of apparent source with displaced listening position.

loudspeakers without the tendency to leave a "hole" in the center. This latter effect can be very pronounced with the spaced microphone system.

E. Sources of Error

Although the two amplifying channels and the microphones and loudspeakers have been assumed to be distortionless in the above explanation, it is clear that they cannot be matched perfectly in amplitude and phase, particularly the transducers. The effect of the small inevitable differences in frequency response can be calculated from the preceding theory, but such transducers will, in general, introduce phase shifts before amplitude differences become very pronounced.

Calculations show that a phase shift in one channel produces no change in the central position, *e.g.*, when the two channels are at the same level. A phase shift in one channel of as much as 90° can introduce error factors of 2 at small values of θ_t, 1.5 at $\theta_t = 26°$ and 1.2 at $\theta_t = 39°$.

1) Effect of Asymmetric Listening Position: The preceding calculations have assumed an observer symmetrically placed with regard to the loudspeakers and facing the center line. If the observer turns his head, the apparent source will appear to remain in the same position, as in natural listening, but if he moves to either side of the center line, the apparent position will move in the same direction, *e.g.*, from S_a to S_a' in Fig. 7.

The reasons for this are, first, that the loudness from W_R has increased and that from W_L decreased, and their relative phases are altered owing to the difference in the path lengths. Second, the angles subtended by the loudspeakers at the observer also have become unequal. The effect of these inequalities makes the calculation of apparent angle exceedingly laborious, but it is possible, by making simplifying assumptions, to appreciate their general effect. Fig. 8 shows the position of the apparent source of Fig. 7 for various listening positions plotted against the positions of the real source with respect to the microphones. In Fig. 8, r is the projection of the observer's displacement on the loudspeaker base-line, expressed as a fraction of x. The calculations ignore the effect of the path difference on the

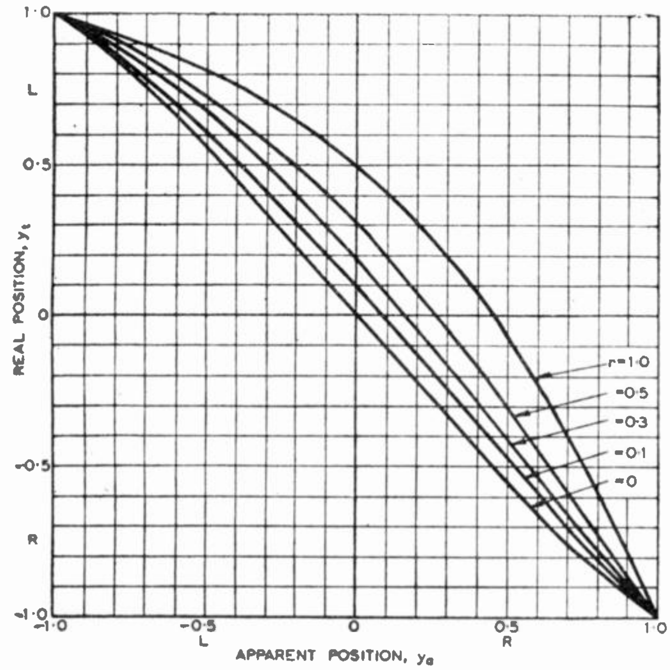

Fig. 8—Apparent position plotted against true position for asymmetric listening position.

relative phases at the listener of the sound from W_L and W_R; *i.e.*, it applies where such a path difference is small compared with a wavelength. The curves bear out the statement that the apparent source moves in the direction of the observer.

In order to take account of the effect of path differences it is necessary to fix the scale of the layout shown in Fig. 7, so that a definite phase shift can be associated with any given listening position at a particular frequency. For this purpose, let d equal 10 feet to represent a typical domestic arrangement. A listening position is chosen corresponding to $r = 0.5$, and the apparent position is plotted against frequency in Fig. 9 for various values of the true angle θ_t. It will be seen that, as the frequency increases from zero, deviations from Fig. 8 are initially small, but they increase rapidly above 100 cps, reaching a maximum in the region of 200 cps beyond which the function approaches its initial value once more before reaching a second maximum at about 600 cps and again returning to its original value.

The calculations have not been made for frequencies above which the theory given in Section IV-B, is valid.

The function may reach a finite maximum, which in all cases is seen to lie outside the loudspeaker base-line, or it may become imaginary. In this case, the phase difference at the ears is so large that it cannot be interpreted in terms of any real angle of arrival. These regions are the ones in which the path difference approximates to an odd number of half-wavelengths, the sound from the two loudspeakers arriving substantially out of phase at the two ears.

Fortunately, at these critical frequencies the resultant magnitude becomes small, thus allowing the remaining sound, which has small phase errors, to predominate.

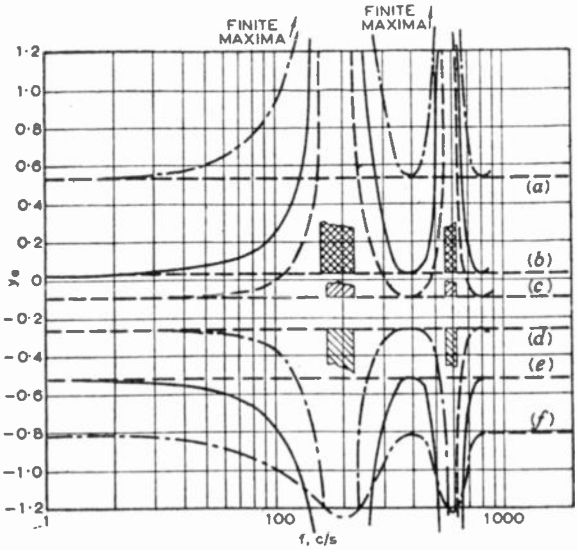

Fig. 9—Apparent position as a function of frequency for asymmetric listening position. (a) $\theta_t = 36.9°$. (b) $\theta_t = 18.4°$. (c) $\theta_t = 11.3°$. (d) $\theta_t = 0$. (e) $\theta_t = -18.4°$. (f) $\theta_t = -36.9°$.

The importance of scale is obvious from the curves; halving all dimensions will allow the system to operate to twice the frequency before the anomalous regions are approached.

The effect of the offset listening position is worst for true angles near the center, when L and R have about the same amplitude. (The extreme positions, $\theta_t = \pm 45°$, are unaffected, since only one loudspeaker is operating.)

V. Recording Equipment

A. Microphone Combinations

1) Pressure Microphones: As stated in 3) of Section IV-C, the earliest experiments were necessarily carried out with pressure microphones. These had substantially spherical polar diagrams, at least, at low frequencies. The problem of producing magnitude differences was solved by giving them a small but definite spacing, summing and differencing their outputs, and integrating the difference vector before recombining to drive the loudspeakers. The significant parameter in such a case is the actual spacing employed. If this exceeds about a quarter-wavelength of the highest working frequency, the resultant outputs to the loudspeakers become of opposite polarity and the system breaks down. Small spacing is desirable on this score. However, if the spacing is too small, the difference vector will diminish until it is comparable with the normal random differences in the microphones. Moreover, the amplification in the difference channel has to be increased to compensate, in addition to having a characteristic inversely proportional to frequency as already mentioned. Low-frequency electrical noise introduced by the amplifiers thus becomes the factor limiting the reduction in microphone spacing. A way out of this difficulty was sought by using more than one pair of microphones to cover the required frequency band, but, as can be imagined, this arrangement was cumbersome in practice and complicated theoretically. Pressure microphones were abandoned finally after a composite microphone system had been tried, using two pairs of crystal elements, one pair being spaced 8 cm apart and operating up to 1000 cps and the other being spaced 1 cm apart and operating from 1000 cps upwards.

2) Velocity Microphones: By this time, ribbon microphones were becoming available. As is well-known, these have a cosine-law polar characteristic in the horizontal plane. A pair of these can be used in two ways to produce the required outputs for driving a pair of loudspeakers. For example, they may be used with one microphone having its axis directed to the center of the sound stage and the other as close as possible but with its axis at right angles. In this condition, the outputs from the microphones will be in phase and invariant with frequency (insofar as their chracteristics are flat), and will correspond to the sum and difference vectors described in 1) of Section V-A, after manipulation but before recombination to drive the loudspeakers.

Alternatively, they may have their axes equally inclined to the center of the sound stage. In this condition, their outputs are suitable for direct amplification and reproduction by a pair of spaced loudspeakers. Such outputs may be considered to be of "left and right" type rather than "sum and difference," as in the previous instance.

In either case, it will be noted that the operation of integration that had to be performed in the case of pressure microphones is unnecessary. Consequently, although ribbon microphones are generally less sensitive than pressure types, the over-all signal/noise ratio of the system is improved.

3) Other Types of Microphones: It will be apparent that any pair of microphones having a polar characteristic other than circular can be used in this way to produce "left and right" type signals of different amplitudes at a pair of loudspeakers. Whether these combinations are of value depends on the uniformity with which their polar characteristics can be maintained with frequency, as well as the degree within which the microphones can be matched in amplitude and phase.

Combinations of microphones having dissimilar polar characteristics may also be employed to produce "sum and difference" type outputs, provided that their amplitude and phase characteristics are well matched. Combinations of polar diagrams which appear to show promise are those of circular with cosine, which would give the same directional response as the pressure microphones used in the early experiments, and cardioid with cosine, which would give substantially a one-sided version of the same thing.

4) Current Practice: The combination of two cosine microphones has so much to recommend it compared with other known arrangements that the principal effort has been concentrated on its improvement. If a pair of ribbon microphones is used, or even a specially-designed double-element ribbon type, the difficulties encountered are still quite serious. At low frequencies, the fundamental resonances of the two ribbons cause relative phase shifts unless these are adjusted to occur at the same frequency and to have the same degree of damping, while at high frequencies the ribbon microphone having the theoretical cosine-law polar diagram still has to be designed. Departure from constancy of polar diagram with frequency will upset the requirement that the amplitude differences for any given angle of arrival shall be independent of frequency. Nevertheless, with all their limitations, some very acceptable recordings have been made using ribbon types.

Within recent years the condenser microphone has undergone considerable development, enabling directional characteristics to be obtained from a double-diaphragm element of small size. Since these elements are substantially free from mechanical resonances within the audible spectrum, their amplitude and phase characteristics show great uniformity. This is fortunate as the polar characteristics show some variation from sample to sample, particularly at low frequencies. Thus selection has to be principally on the basis of polar response. Individual adjustments to match their sensitivities can be made easily. Pairs of these microphones, mounted with their elements as close together as possible in a single cylindrical holder, have been used with some success. Like all arrangements with one element above the other, it is necessary to insure that the common axis is perpendicular to the plane in which the various sounds lie, otherwise the vector relationships are upset.

B. Studio Techniques

The problems encountered in making a "stereosonic" record are not all of an engineering nature. The finished record must have a pleasing tonal balance, and the apparent spatial distribution of the instruments, if not of prime importance, must not sound unnatural. Using a single pair of microphones, it is often extremely difficult to satisfy the required conditions. In single-channel recording, good results have been achieved sometimes, using only one microphone, but the occasions when this is possible are the exception rather than the rule. Generally, in order to achieve a proper balance, multi-microphone techniques have to be adopted. For obvious reasons this cannot be done when the direction of the various sources is significant. It is, therefore, necessary to rely on correct disposition of the performers in the recording studio to achieve the required tonal balance. In this connection, it may be remarked that the double-cosine combination has two working arcs each 90° wide on opposite sides of the common axis. This has been found useful in practice, particularly for large choral works, in which the orchestra has been arranged on the one side, with choir and soloists on the other. The adoption of such unusual layouts, however, often makes the task of the conductor more difficult, since he may not be able to see all the performers without turning around. Considerations of this sort, which may be termed the technicalities of musical performance, often preclude the use of the purely engineering solution and make the task of the recording engineer even more complicated.

Under difficult circumstances it is possible to envisage the need for more than one microphone pair, but it is realized that the use of additional microphones may introduce as many problems as it solves.

C. Microphone Amplifiers and Equalizers

Amplifier design follows conventional lines. Of course, it is essential that pairs of amplifiers used in this type of recording have accurately matched frequency and phase characteristics, and that their gain be constant over long periods. In general, however, modern designs employing negative feedback have no difficulty in achieving the required consistency. The same requirement for accurate matching applies to any microphone equalizers that are used, *e.g.*, for bass correction when working in close proximity to an artist.

D. Mixers

As stated in Section V-B, the use of a single microphone pair for stereophonic recording is almost mandatory. It is possible nevertheless, to foresee circumstances in which it may be necessary to use a second pair or to introduce extraneous effects.

Therefore, some form of mixer is necessary, although this will be, generally, a simple device compared with those used in conventional recording, which may have to cater for relatively large numbers of microphone channels.

A stereophonic mixer in current use does permit the use of two microphone pairs with independent level controls. Also, it includes means for providing the correction at high frequencies referred to in Section IV-D. This requires that the signals, if not of the sum-and-difference type, shall have been subjected to a sum-and-difference operation.

A further facility is provided which gives an additional degree of flexibility in the recording studios. This consists of a differential control permitting the relative amplification of the "sum and difference" channels to be altered over a limited range. This control allows the apparent angle of the total reproduced sound to be modified and can be used to good effect in several ways. The microphones may have to be withdrawn from the artists in order to get more "ambience" into the recording. This will reduce the apparent stage width, but the full width can be restored by suitably attenuating the sum channel relative to the difference. Similarly, if the need for great "presence" calls for a very close microphone position, the reproduction may cause a solo instrument to sound much too large, and this can be corrected by attenuating the difference channel relative to the sum.

The signals are finally subjected to a second sum-and-difference operation to convert them to "left and right" type for application to the input of a twin-track tape recorder. The replay amplifiers of the machine drive two matched loudspeakers suitably adapted for stereophonic monitoring.

A simplified diagram showing the operations performed in a typical recording channel is given in Fig. 10.

An interesting feature of the control system is the peak-level indicator. The use of two such meters, one on each channel, is unsatisfactory owing to the difficulty of observing them simultaneously. In this equipment, a peak-level indicator circuit[16] with the ability to measure the true level of short transients, is connected to each recording channel and the meter on the mixer panel can be switched to either channel at will. Alternatively, for general use, a circuit is provided whereby the meter is switched electronically to whichever channel has the higher level. The operator then uses his main gain control in the ordinary way to avoid overloading either channel on the tape.

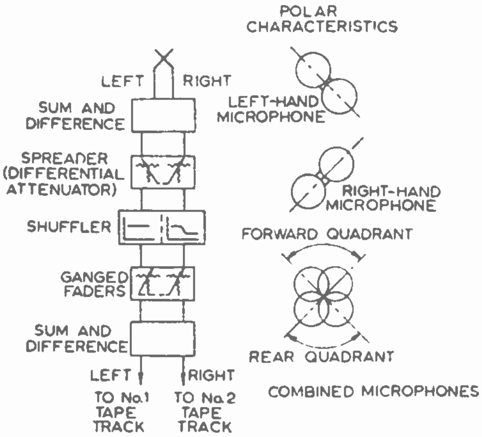

Fig. 10—Functional schematic of recording system.

E. Magnetic-Tape Recording Machines

In normal record production it is the practice, irrespective of the final form of the commercial article, to make master recordings on magnetic tape for convenience of play-back, editing, etc. These advantages apply equally to stereophonic recording. The system described employs professional machines of high quality, modified by the addition of twin-track heads and twin-head amplifier channels.

A standard tape speed of 15 inches per second has been chosen for two reasons:

1) Satisfactory standard of master quality can be obtained at this speed.

2) The majority of single musical works can be recorded within the full 11-inch diameter tape spools and thus simplify the copying process.

The record heads are provided with bias from a common oscillator to avoid heterodyne beats due to residual crosstalk.

VI. Recording and Reproducing Methods

A. Choice of Media

It is essential that the recording of the two channels should be made on a medium which allows the two tracks to be permanently linked together so that the correct phase relation between the signals is maintained throughout any copying process and replay operation.

In 1931, when Blumlein conceived the basic idea of his system, the only recording medium capable of giving reasonable results was the wax-cut disk with its electrolytic copying process. He, therefore, proposed applying the two signals from the microphone system to the complex transducer so arranged as to cut a single groove. If the axes of movement of the transducer armatures carrying left and right signals are inclined at 45° to the surface of the wax and are thus at 90° to each other, the resulting lateral cut can be arranged to represent the sum of the two channels, and the "hill and dale" to represent the difference. Such a disk could be played on a normal gramophone to give the equivalent reproduction of a single-channel recording. The early experimental samples of the complex-cut disk suffered from considerable interaction at the higher frequencies and from excessive background noise. The fine-groove technique may simplify the problem, but to produce a complex-cut disk to the standards demanded by the modern record is very difficult, and the problem has not been solved satisfactorily yet.

In the United States, stereophonic disk records have been marketed in which the lateral-cut tracks are disposed in two concentric bands. This record is replayed with two pickup heads mounted on a common arm. The operation of locating both pickup styli in their correct grooves is a delicate one. The frequency response and distortion characteristic of the two tracks will differ quite appreciably when there is a large difference between the track diameters.

There have been various suggestions for modifying the frequency band of the signals of one of the channels so that both channels can be recorded on one groove by means of a single transducer. One such method, proposed by Livy,[17] suggests that one set of signals is arranged to modulate a carrier, then to select the lower sideband and apply this, together with the normal signal band of the other channel, to a wide-range cutter head. Reproduction is effected by using a wide-range pickup, separating the two frequency bands and applying the upper band to modulate a carrier of the same frequency as that used for recording. The lower sideband produced by this modulator contains the same frequency components as the original signal. Livy also proposed to record the carrier frequency on the disk so

[16] M. T. Terry, British Patent Application No. 12839, 1956.

[17] W. H. Livy, British Patent No. 612163.

that it would be used during reproducing to control the frequency of the local oscillator. This double-bandwidth system demands a very high performance in terms of frequency response, and thus if a frequency range up to 10 kc be required, the full bandwidth of the system will be at least 20 kc. This wide bandwidth is very difficult to attain, particularly at the inner groove diameters.

If recordings are made side by side on a continuous film, the synchronization problems are largely overcome, and while it is recognized that optical methods have been used, the recent development of the magnetic-tape process makes it a more attractive system.

1) Crosstalk: The over-all crosstalk between channels should be better than 30 db at all frequencies. In the production of a commercial tape record, there are at least four stages where crosstalk can take place: master recording, master replay, copy recording, (commercial tape record), and copy replay.

If copy masters are used, two more stages of crosstalk are introduced. In practice, it has been found that the same amount of crosstalk is introduced during the recording process as during the reproducing process.

In well-designed heads a crosstalk of -50 db can be attained, and if the crosstalk at the various copying stages can be added arithmetically the over-all figure for a commercial tape record will be -38 db, or if a copy master is used, -34 db. If it is required to replay either twin-track or single-channel half-track tape records from the same head system, the crosstalk of this replay head must be better than -55 db, and preferably better than -60 db, in the 100–3000 cps region.

Crosstalk can take place owing to the mutual inductance of the heads and owing to leakage fields from one track to the head on the other track. The mutual inductance can be reduced by placing magnetic screens between the heads and separated from them by brass or aluminum spacers. These screens should be large enough to shield the magnetic circuit and the windings. The usual back gap should be eliminated, and the head dimension should be kept as small as possible.

A useful reduction of crosstalk can be achieved by either series or parallel cross-connection of the windings, so that a small signal from one head is injected in antiphase to the leakage signal on the other head. If carefully adjusted the series cross-connection can reduce the residual crosstalk by at least 10 db.

Appreciable crosstalk may occur if the front edge of the outer magnetic screening shield is too close to the magnetic tape. The flux from one track enters the shield and then passes to the magnetic yoke of the other head operating on the other track. A clearance between the tape and the shield of 0.2 inch is sufficient to avoid this effect.

B. Track Standards

The early experimental, stereophonic, magnetic-tape recordings were made with the heads displaced by $2\frac{1}{16}$ inches. This enabled the conventional half-track heads to be used on the existing recording and replay machines, and the separation between the heads was sufficient to make crosstalk negligible. When copy tapes were made for issue as commercial records, it was soon obvious that the use of in-line heads would make for easier acceptance as an international standard for domestic and commercial tape records. Also it was decided to change to in-line heads on the master recordings, so that the process of editing was simplified. Furthermore, the use of in-line heads eliminates phase shift between the signals on the two tracks owing to small variations of elasticity of the tape along its length.

Standardization has been effected on the designation of the tracks, which is as follows:

If the tape moves from left to right and with the active side facing away from the observer the top track shall be designated *No. 1 track* and shall carry the recording for the left-hand channel as viewed from the audience. The bottom track shall be designated *No. 2 track* and shall carry the recording of the right-hand channel.

The replay response characteristic (100 μsec) and the track dimensions are in accordance with Amendment No. 1 to B.S. 1568: 1953, relating to tape speeds of $7\frac{1}{2}$ inches per second. The track dimensions are as shown in Fig. 11.

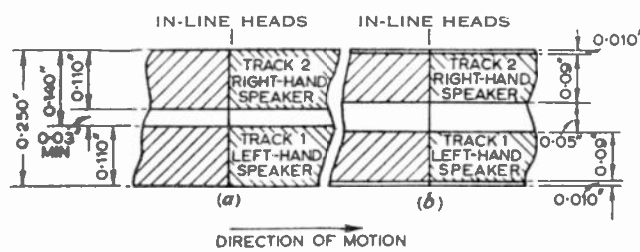

Fig. 11—Track standards. (a) Record-head track. (b) Replay-head track (views looking on coated side of tape).

C. General Requirements for Domestic Operation

The tape transport mechanism may be of the conventional type, but the associated twin-track magnetic replay head must conform to the standards laid down above. Since the signals from the two tracks on the tape are left and right, it is essential that the gains of the two replay channels should be closely matched, and for this purpose preset controls should be provided for initial adjustment. It is a further requirement that the equality of gain should be maintained for all settings of the main coupled gain control. Large relative phase shifts in the two channels should be avoided, particularly at low frequencies.

The underlying principle of the system demands that the loudspeakers should possess, as far as possible, uniform polar response characteristics in the horizontal plane at all frequencies. It has been suggested that departure from uniformity can be an advantage in maintaining the apparent position of the source for a wide range of listening positions. Such nonuniform polar characteristics cannot reasonably be attained except at high frequencies. Even if it were possible to extend the directional characteristics down to the low frequencies, the beneficial results claimed would not be realized on account of the path differences as outlined in 1) of Section IV-E. At the high frequencies nonuniform polar characteristics have the disadvantage that the over-all tonal balance will vary with the listener's position. Furthermore, the directional response will emphasize any background hiss and will tend to identify the two loudspeakers as separate sound sources and thus, interfere with the illusion. In addition, the domestic user might have difficulty in positioning the loudspeakers with sufficient accuracy.

If cone loudspeakers are used the rear radiation should be suppressed, and this can be conveniently done by using a closed box baffle. Such baffles can have quite small volume, and they can be designed to occupy a very small floor area. Any loss of bass response due to the small enclosed volume can be compensated electrically in the power amplifiers.

1) A Commercial Model: One form of domestic reproducer has been described by Smith and Martin,[18] but for the sake of completeness a brief general description may be of interest. The complete machine consists of two consoles, each containing a loudspeaker group and a power amplifier. One of these consoles also contains a tape deck with its associated head amplifiers, tone and gain controls. Fig. 12 is a schematic. Bass and

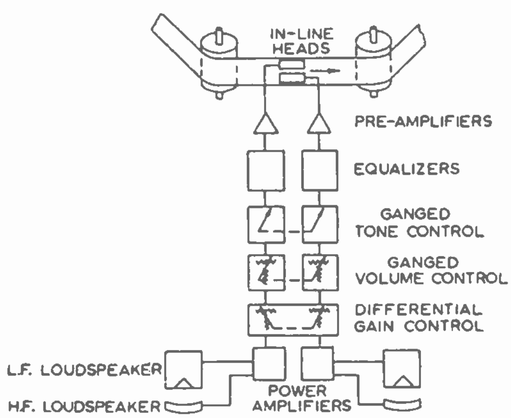

Fig. 12—Functional schematic of reproducer.

treble ganged tone controls are provided, in order to allow some adjustment for the different acoustic conditions which may be met.

In order to maintain uniformity between the two channels the controls are of the stepped-switch type. The main gain control consists of two ganged switches with 12 steps of 3 db each. There is a differential gain control on the panel to take up any slight drift in amplifier gains and to allow for asymmetry in the listening room. A preset continuous control on the amplifier chassis balances the gain during factory adjustment. The power amplifiers are rated at a peak output level of 10 watts.

In order to give the best horizontal polar distribution, the elliptical cone loudspeakers are mounted in closed rigid box baffles of $3\frac{1}{2}$ feet^3 capacity, the major axis of the cone ellipse being vertical. The frequency range of this speaker is limited at the upper end to about 5000 cps in which region the electrostatic speaker takes over and continues beyond 15 kc. The electrostatic speaker consists of a curved metal back plate 24 inches long by $1\frac{1}{4}$ inches wide, over which is laid a membrane metallized on the side away from the back plate. This speaker is driven from a separate amplifier and is supplied with a bias potential of 300 volts.

2) Domestic Listening Conditions: In general, "stereosonic" tape records are balanced with regard to tonal quality and perspective, for reproduction in the average home lounge in which the reverberation time is of the order of 0.5 second or less. This reverberation period is short compared with that in the studio or concert hall in which the recordings will have been made; it is not likely to interfere with the listener's impression of the original studio conveyed to him by the stereophonic effect; and there will be no confusion due to the two distinct reverberations which might blur the detail. Provided that the walls and the floor are reasonably absorbent, it is found that, in spite of errors due to asymmetry, one can move about over a large floor area in front of the loudspeakers without losing the major benefits of this type of reproduction. As stated earlier in the paper, the optimum position for listening is the apex of an equilateral triangle with the base-line formed by the line joining the two loudspeakers; this is the arrangement used for monitoring during recording. In the average room it is satisfactory to set the loudspeakers 10 or 12 feet apart, but in a restricted space a very fair performance can be obtained down to distances of as little as 4 or 5 feet.

D. Operation of the System in Large Halls

Special difficulties attend the reproduction of "stereosonic" records on a large scale, apart from the general fact that, like all other types of record, they are balanced primarily for domestic conditions.

The most serious errors are caused by increased path differences. These occur, as outlined in Section IV-E., down to the lower frequencies and over a much greater proportion of the listening area than is the case when operating at the scale for which the system was designed.

In a small room reflections from the walls, etc., decay rapidly and follow each other at such short intervals

[18] M. B. Martin and D. L. A. Smith, "The design of magnetic recording and reproducing equipment for domestic use," *J. Brit. IRE.* vol. 16, pp. 65–77; February, 1956.

that the resulting sounds appear to the listener to coalesce. In large halls, the relatively undamped primary reflections from the surroundings arrive at the listeners' ears at intervals sufficiently great to cause an appreciable distortion of the sound picture.

Ideally, the system requires that the loudspeakers shall be uniformly radiating point sources. In the domestic case, a reasonable approximation to this can be achieved because single small units can be constructed to handle the required power. In large-scale reproduction, multiple units or single units of appreciable size are necessary; this departure from the ideal can cause a depreciation of the stereophonic definition.

Appendix

Change of Position of Apparent Source Under Asymmetric Listening Conditions

In the asymmetric case, there are three principal reasons for departure from the simple theory applicable to central listening. These are as follows:

1) Operation of the inverse-square law to alter the magnitudes of the sound pressures from the loudspeakers at the listener.

2) The phase difference due to the unequal path lengths from the loudspeakers to the listener.

3) The unequal angles subtended by the loudspeakers at the listener.

The method of calculation of apparent position is shown below. Fig. 13 shows diagrammatically a source S on a sound stage PQ, towards which a microphone pair M is directed. The microphone outputs are reproduced on a similar stage by a pair of loudspeakers, W_L, W_R, assumed to have uniform polar characteristics over the frequency range considered. The virtual image S_a thus produced is observed by a listener O displaced from the central position C.

Then let

$2x$ = Distance between loudspeakers W_L and W_R.
d = Distance between listener O and base-line.
rx = Displacement of listener parallel to base-line.
y_a = Fractional distance of apparent source S from central position.
y_t = Fractional distance of real source from central position.
L, R = Magnitudes of sound pressures due to W_L and W_R at C.
β, γ = Angles subtended at listener by W_L and W_R.
τ = Difference in time of arrival of sounds from W_L and W_R at O due to path difference.
μ_L, μ_R = Time differences at ears E_L and E_R relative to O of sounds from W_L and W_R.
ϕ_L, ϕ_R = Resultant phase relative to O of sound at left and right ears.
θ_a = Apparent angle of arrival of sound at O.
θ_t = True angle of source with respect to microphones for magnitudes L and R from the loudspeakers.

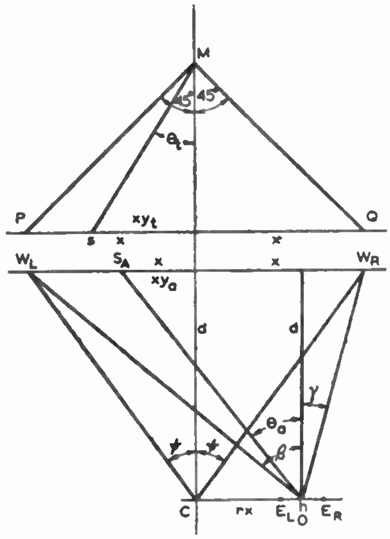

Fig. 13—Position of apparent source in the asymmetric case.

Then:

$$OW_L = d\sqrt{[(1+r)^2 \tan^2\psi + 1]}$$
$$OW_R = d\sqrt{[(1-r)^2 \tan^2\psi + 1]}$$
$$\tau = \frac{OW_L - OW_R}{v}$$
$$= \frac{d}{v}\left\{\sqrt{[(1+r)^2\tan^2\psi+1]} - \sqrt{[(1-r)^2\tan^2\psi+1]}\right\}$$
$$\mu_L = \frac{h\sin\beta}{2v}$$
$$\mu_R = \frac{h\sin\gamma}{2v}.$$

Relative pressures at the listener's ears are as follows:

from

$$\begin{array}{lcc} & \text{at } E_L & \text{at } E_R \\ W_L & \dfrac{L\sin\omega(t-\tau+\mu_L)}{\sqrt{[(1+r)^2\tan^2\psi+1]}} & \dfrac{L\sin\omega(t-\tau-\mu_L)}{\sqrt{[(1+r)^2\tan^2\psi+1]}} \end{array}$$

from

$$\begin{array}{lcc} W_R & \dfrac{R\sin\omega(t-\mu_R)}{\sqrt{[(1-r)^2\tan^2\psi+1]}} & \dfrac{R\sin\omega(t+\mu_R)}{\sqrt{[(1-r)^2\tan^2\psi+1]}}. \end{array}$$

The resultant phases at the ears relative to O, derived from the above, are:

$$\phi_L = \arctan\left\{\frac{\dfrac{L\sin\omega(\mu_L-\tau)}{\sqrt{[(1+r)^2\tan^2\psi+1]}} - \dfrac{R\sin\omega\mu_R}{\sqrt{[(1-r)^2\tan^2\psi+1]}}}{\dfrac{L\cos\omega(\mu_L-\tau)}{\sqrt{[(1+r)^2\tan^2\psi+1]}} + \dfrac{R\cos\omega\mu_R}{\sqrt{[(1-r)^2\tan^2\psi+1]}}}\right\}$$

$$\phi_R = -\arctan\left\{\frac{\dfrac{L\sin\omega(\mu_L+\tau)}{\sqrt{[(1+r)^2\tan^2\psi+1]}} - \dfrac{R\sin\omega\mu_R}{\sqrt{[(1-r)^2\tan^2\psi+1]}}}{\dfrac{L\cos\omega(\mu_L+\tau)}{\sqrt{[(1+r)^2\tan^2\psi+1]}} + \dfrac{R\cos\omega\mu_R}{\sqrt{[(1-r)^2\tan^2\psi+1]}}}\right\}$$

whence it can be shown that the total resultant phase difference at the ears is:

$$(\phi_L - \phi_R) = \arc\tan\left\{\frac{\dfrac{L^2 \sin 2\omega\mu_L}{(1+r)^2\tan^2\psi+1} - \dfrac{R^2 \sin 2\omega\mu_R}{(1-r)^2\tan^2\psi+1} + \dfrac{LR2 \sin\omega(\mu_L-\mu_R)\cos\omega\tau}{\sqrt{[(1+r)^2\tan^2\psi+1]}\sqrt{[(1-r)^2\tan^2\psi+1]}}}{\dfrac{L^2 \cos 2\omega\mu_L}{(1+r)^2\tan^2\psi+1} + \dfrac{R^2 \cos 2\omega\mu_R}{(1-r)^2\tan^2\psi+1} + \dfrac{LR2 \cos\omega(\mu_L-\mu_R)\cos\omega\tau}{\sqrt{[(1+r)^2\tan^2\psi+1]}\sqrt{[(1-r)^2\tan^2\psi+1]}}}\right\}.$$

Making the approximation, on which the system is based, that $\omega\mu_L$ and $\omega\mu_R$ are small, thus implying also that ϕ_L and ϕ_R are small:

$$(\phi_L - \phi_R) \to \tan(\phi_L - \phi_R)$$

$$\to \left\{\frac{\dfrac{L^2 2\omega\mu_L}{(1+r)^2\tan^2\psi+1} - \dfrac{R^2 2\omega\mu_R}{(1-r)^2\tan^2\psi+1} + \dfrac{LR2\omega(\mu_L-\mu_R)\cos\omega\tau}{\sqrt{[(1+r)^2\tan^2\psi+1]}\sqrt{[(1-r)^2\tan^2\psi+1]}}}{\dfrac{L^2}{(1+r)^2\tan^2\psi+1} + \dfrac{R^2}{(1-r)^2\tan^2\psi+1} + \dfrac{LR2\cos\omega\tau}{\sqrt{[(1+r)^2\tan^2\psi+1]}\sqrt{[(1-r)^2\tan^2\psi+1]}}}\right\}.$$

But

$$(\phi_L - \phi_R) = \frac{\omega h \sin\theta_a}{v}.$$

therefore

$$\theta_a = \arc\sin\left[\frac{v}{\omega h}(\phi_L - \phi_R)\right].$$

Substituting for μ_L and μ_R in the expression for $(\phi_L - \phi_R)$:

$$\theta_a = \arc\sin\left\{\frac{\dfrac{(L/R)\sin\beta}{(1+r)^2\tan^2\psi+1} - \dfrac{(R/L)\sin\gamma}{(1-r)^2\tan^2\psi+1} + \dfrac{(\sin\beta-\sin\gamma)\cos\omega\tau}{\sqrt{[(1+r)^2\tan^2\psi+1]}\sqrt{[(1-r)^2\tan^2\psi+1]}}}{\dfrac{L/R}{(1+r)^2\tan^2\psi+1} + \dfrac{R/L}{(1-r)^2\tan^2\psi+1} + \dfrac{2\cos\omega\tau}{\sqrt{[(1+r)^2\tan^2\psi+1]}\sqrt{[(1-r)^2\tan^2\psi+1]}}}\right\}.$$

The fractional distance of the apparent source from the central position is then given by

$$y_a = \frac{d\tan\theta_a}{x} - r = \frac{\tan\theta_a}{\tan\psi} - r.$$

Similarly, the fractional distance of the real source from the central position before the microphones is given by

$$y_t = \frac{\tan\theta_t}{\tan 45°} = \tan\theta_t$$

45° being half the working angle of the microphones.

Acknowledgment

The authors wish to acknowledge the assistance of many members of the technical staff at the Recording Studios and the Research Laboratories of Electric and Musical Industries, Ltd., and Messrs. D. E. Broadbent and B. Shackel for helpful discussion on the physiology of hearing.

They also wish to thank the Directors of Electric and Musical Industries, Ltd., for permission to publish the paper.

Correspondence

*PGA Editor**

The Australian audio series[1] was most interesting and appears to follow along the British school of thinking.

About the ear's resolving powers, it may be that the odd harmonics and especially the higher-order components, are the ones the ear is most sensitive to. Hence, amplifier harmonic and intermodulation distortion ratings might correlate more with subjective evaluation if the various harmonics were specified. Before the war, one hi-fi addict stated to the writer that he could detect the difference between a medium-mu and hi-mu triode; and direct and condenser coupling between two stages by a listening test. It may be significant that a hi-mu triode tends to generate more higher-order odd harmonics. Quite possibly, coupling condensers' dielectric hysteresis effects produce some odd-ball traces of distortion the ear can detect.

With regard to test methods, an interesting variant upon Maxwell's sine-wave-burst transient-test method, is to use an audio sweep generator and gradually increase the sweep speed. An interesting "shock excitation" action can sometimes be seen on the scope as the circuit starts to "ring" as the sweep rate increases.

As pointed out by an RCA paper[2] many years ago, audio transducers can also generate considerable fm distortion, particularly loudspeakers. Quite possibly, audio transformers and coupling condensers might generate traces of this type of distortion. (Wow, flutter, and hash are a form of this type.) Hence audio fm distortion effects might be investigated.

The adoption of some sort of an inductive speaker "dummy" load for rating amplifiers would be a help, particularly since pentodes and feedback are widely used in the more popular equipment. Such a combo looks a lot less inviting when working into a speaker load than when feeding a dummy resistor. Pentodes with feedback seem to always sound a bit "dirty" to many hi-fi hounds and for this reason they prefer the costlier triode setups with less traces of phase-shift ringing.

Ted Powell
Great Neck, L. I., N. Y.

* Received by the PGA, August, 2, 1957.

[1] IRE Trans., vol. AU-5, March-April, 1957.

[2] G. L. Beers and H. Belar, "Frequency-modulation distortion in loudspeakers," Proc. IRE, vol. 31, pp. 132–138; April, 1943.

Contributors

H. A. M. Clark received the B.S. degree in engineering. In 1928, he joined the research department of Columbia Gramophone Co., Ltd. In 1931, he became associated with Electric & Musical Industries, Ltd., and carried out research work on recording, including stereophonic methods, television, studio equipment, and transmitters and radar systems. In 1953, he was appointed Technical Manager of the Record Division, and in 1957, became Technical Director of EMI Internation Ltd.

H. A. M. Clark

Mr. Clark is a member of the IEE.

✣

J. Rodrigues de Miranda was born on December 24, 1905, in Amsterdam, Holland. He attended the Technical High School in Delft, Holland from 1922 to 1927 when he graduated as an electrical engineer. From 1927 to 1938, he was the manager of the technical department of a wholesale firm. In 1938, he joined Philips Industries, Eindhoven, the Netherlands, where his work has been mainly with audio and acoustical research and development. Since 1951, Mr. de Miranda has been the manager of the reproduction advisory group of the Radio Apparatus Laboratory at Philips.

J. R. de Miranda

Mr. de Miranda is a member of the Nederlands Radio Genootschap, the Koninklÿk Instituut van Ingenieurs, and the Audio Engineering Society.

✣

G. F. Dutton received the B.Sc. and Ph.D. degrees. He joined the Designs Division of Gramophone Co., Ltd., in 1931, carrying out development work on studio acoustics, loudspeakers, and components associated with the gramophone technique. After World War II, he was responsible for the development of magnetic recording machines and industrial equipment. In 1955, Dr. Dutton joined the Record Division of the Electric and Musical Industries Ltd., Hayes, Middlesex, England, to take charge of the Advance Development Group.

G. F. Dutton

He is a member of the IEE.

✣

Mahmoud Ahmed Melehy (A'53) was born on June 10, 1926, in Dessamy, Giza, Egypt. In 1947, he received the B.S. degree in electrical engineering from Cairo University, Cairo, Egypt. He then joined the Egyptian State Telephone and Telegraph Service. In 1949, he received the M.S. degree in electrical engineering from The Ohio State University, Columbus, Ohio. In 1950 and 1952, respectively, he obtained the M.S. degree in mathematics and the Ph.D. degree in electrical engineering, both from the University of Illinois, Urbana, Ill.

M. A. Melehy

From 1952 until 1955, Dr. Melehy taught electrical engineering at the University of Arkansas, Fayetteville, Ark.; Heliopolis University, Cairo, Egypt; and the University of Alaska, College, Alaska. Since 1955, he has been an assistant professor of electrical engineering at Michigan State University, East Lansing, Mich. During the summer of 1956, he joined as a senior engineer the Computer Department of the Research Laboratories, Division of Bendix Aviation Corporation, Detroit, Mich.

Dr. Melehy is a member of the AIEE, Eta Kappa Nu, and the American Society for Engineering Education.

✣

Harry F. Olson (A'37–VA'39–SM'48–F'49) was born on December 28, 1902, in Mt. Pleasant, Iowa. He received the B.S. degree in 1924, the M.S. degree in 1925, the Ph.D. degree in 1928, and the E.E. degree in 1932, from the University of Iowa. From 1928 to 1930, he was in the Research Department of Radio Corporation of America; from 1930 to 1932, in the Engineering Department of RCA Photophone; from 1932 to 1941, in the Research Division of RCA Manufacturing Company; and since 1941, with the RCA Laboratories. He is Director of the Acoustical and Electro-mechanical Laboratory of the RCA Laboratories.

H. F. Olson

For his contributions in the field of audio engineering, Dr. Olson has received the following honors: the Modern Pioneer Award of the National Association of Manufacturers in 1940, the John H. Potts Medal of the Audio Engineering Society in 1952, the Samuel L. Warner Medal of the Society of Motion Picture and Television Engineers in 1955, the John Scott Medal of the City of Philadelphia in 1956, and the Achievement Award of the Professional Group on Audio of the Institute of Radio Engineers in 1956.

Dr. Olson is a member of Tau Beta Pi and Sigma Xi. He is a Fellow of the Society of Motion Picture and Television Engineers, the American Physical Society, the Acoustical Society of America, and the Audio Engineering Society. He is a past president of the Acoustical Society of America and is the present chairman of the Administrative Committee of the IRE Professional Group on Audio.

✣

P. B. Vanderlyn joined the research laboratories of Electric & Musical Industries, Ltd. in 1935, where he has been engaged in stereophonic recording and reproducing systems. His work at EMI has included as well, audio and video equipment for television studios, radar projects, etc. At the present time, he is in charge of the acoustics group at the EMI Research Laboratories.

P. B. Vanderlyn

Mr. Vanderlyn is an associate member of the IEE.

INSTITUTIONAL LISTINGS

The IRE Professional Group on Audio is grateful for the assistance given by the firms listed below, and invites application for Institutional Listing from other firms interested in Audio Technology.

ALLIED RADIO CORPORATION, 100 N. Western Ave., Chicago 80, Illinois
Everything in Radio, Television, and Industrial Electronics

ALTEC LANSING CORPORATION, 9356 Santa Monica Blvd., Beverly Hills, California
Microphones, Speakers, Amplifiers, Transformers, Speech Input

AUDIOPHILE RECORDS, Saukville, Wisconsin
High Quality Disc Recordings for Wide Range Equipment

BALLANTINE LABORATORIES, INC., Fanny Rd., Boonton, New Jersey
Electronic Voltmeters, Decade Amplifiers, Voltage Calibrators, Multipliers, Shunts

ELECTRO-VOICE, INC., Buchanan, Michigan
Microphones, Pickups, Speakers, Television Boosters, Acoustic Devices

FAIRCHILD RECORDING EQUIPMENT CO., Whitestone 57, New York
Audio Amplifiers, Moving Coil Cartridges, Transcription Arms, Turntables, Professional Disc Recorders

FREED TRANSFORMER COMPANY, INC., 1718 Weirfield St., Brooklyn 27, New York
Transformers, Reactors, Filters, Magnetic Amplifiers, and Laboratory Test Equipment

(Please see back cover for additional listings)

INSTITUTIONAL LISTINGS (Continued)

JENSEN MANUFACTURING COMPANY, 6601 South Laramie Ave., Chicago 38, Illinois
Loudspeakers, Reproducer Systems, Enclosures

KNOWLES ELECTRONICS, INC., 9400 Belmont Ave., Franklin Park, Illinois
Miniature Microphones and Receivers, Special Recorder and Audio Devices

JAMES B. LANSING SOUND, INC., 3249 Casitas Ave., Los Angeles 39, California
Loudspeakers and Transducers of All Types

SONOTONE CORPORATION, P.O. Box 200, Elmsford, New York
Ceramic Phonograph Cartridges

STROMBERG-CARLSON COMPANY, Special Products Div., 1700 University Ave., Rochester 10, New York
High-Fidelity Equipment; Intercommunication, Public Address, and Sound Systems; Electronic Carillons

TELEX, INC., Telex Park, St. Paul 1, Minnesota
Subminiature Magnetic Receivers, Microphones, Hi-Fi Headsets, Encapsulation, Packaged Circuitry

UNITED TRANSFORMER COMPANY, 150 Varick St., New York, New York
Manufacturers of Transformers, Filters, Chokes, Reactors

UNIVERSITY LOUDSPEAKERS, INC., 80 South Kensico Ave., White Plains, New York
High Fidelity, Commercial-Industrial, and Military Speakers and Accessories

Charge for listing in six consecutive issues of the TRANSACTIONS—$75.00.
Application for listing may be made to the Technical Secretary, Institute of Radio Engineers, Inc., 1 East 79th Street, New York 21, N.Y.