IRE Transactions on AUDIO

Volume AU-8 JANUARY-FEBRUARY, 1960 Number 1

Published Bi-Monthly

TABLE OF CONTENTS

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 fee 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 $0.70; to IRE members at $1.05, and to nonmembers at $2.10.

The Editor's Corner

How to Get Results

A YOUNG man who will go far is Mr. David D., a senior at Manatee High School in Bradenton, Fla. Recently Mr. D. became interested in acoustics and electroacoustics, so he wrote to his congressman in an attempt to acquire information concerning this subject. His congressman forwarded the request to the Federal Communications Commission which in turn suggested The Institute of Radio Engineers. At this point we can envision how the Department of Commerce, the State Department, the Atomic Energy Commission, the Scientific Criminology Laboratories of the Federal Bureau of Investigation, and even Cape Canaveral, might, in turn, have become involved in passing the buck. But someone along the line lacked imagination, and Mr. D's letter fell into the hands of an editor, who discovered that (contrary to popular notion) information on acoustics and electroacoustics was not secret, but could be found in books at libraries and schools. Mr. D. was duly advised of this fact. So far no report of his progress has been received; and it is even possible that Mr. D. is no longer interested in electroacoustics.

—MARVIN CAMRAS, *Editor*

Scrambled Speech

At the old World's Fair in Chicago, in 1933, Bell Telephone had an exhibit of speech scrambling set up with the microphone on one side of the room and the loudspeaker on the other side, so that the person at the mike could not hear what came out of the speaker. The procedure was to ask someone in the audience where he was from, and to pronounce the name of the home town into the microphone so as to amaze the spectators with the outcome. However, it was also the practice to refuse to use certain words. Being nosy, I got into a huddle with the Bell crew after the demonstration, and it seems that a few days before there had been a lady from Oshkosh. When Oshkosh was put in, some of the audience snickered, and some of the women blushed (remember this was over 26 years ago). The crew tried it afterwards and Oshkosh always came out ***kiss. So they made up a list of approved cities, etc. (*Editor's Note:* And also a private list of nonapproved ones.)

EDWARD W. LOGAN, JR.
Memphis, Tenn.

PGA News

ADMINISTRATIVE COMMITTEE MEETING MINUTES

Chicago, Ill., October 12, 1959

The meeting, held in the Time Room of the Hotel Sherman, Chicago, Ill., was called to order at 8:45 P.M. by Chairman Alexander Bereskin. Members and guests present were:

A. B. Bereskin
Marvin Camras
John K. Hilliard
William Ihde (NEC Session Organizer)
J. Ross Macdonald
Albert Meyer (Cincinnati Chapter, PGA)
T. N. Truske (Baltimore Chapter, PGA)
Philip B. Williams

Minutes of the previous meeting of March, 1959, were read by Mr. Camras (Acting Secretary) and were approved unanimously.

Chairman Bereskin said that Dr. Corrington had not yet revised the bylaws to be consistent with the constitution. The Administrative Committee agreed to wait for Dr. Corrington's revisions before taking further action.

Dr. Bereskin next reported on subjects considered at recent professional group meetings. Of interest to PGA was the feeling that certain specialties do not lie wholly within an existing group and, therefore, ought to have a new group of their own. Specifically, Mr. S. A. Cohen of Ampex asked that a Professional Group on Magnetic Recording be formed, since this subject encompasses computers, video recording, instrumentation, and other fields not directly audio. After some discussion, it was the feeling of the committee that few IRE sections could concentrate enough people in such narrow specialty to form a chapter. It was logical to continue the present status, where magnetic recording technology is divided among several groups, each interested in a different aspect. Dr. Hilliard moved that we write to Mr. Cumming at IRE Headquarters and communicate the above feelings. This was seconded by Mr. Williams, and passed by vote.

A discussion was held on a suggestion of Mr. Bauer that IRE Sections which did not have a PGA Chapter should appoint a local PGA representative who would have the status of a chapter chairman. The PGA representative would keep in touch with the national PGA Chapters Chairman and help his section arrange audio programs, obtain tapescripts, and spark the formation of a chapter if sufficient interest developed. Dr. J. R. Macdonald, the Chapters Chairman, said that the suggestion was excellent in principle, but its success depended entirely on finding someone in each section who would cooperate. His experience even in larger sections has made him pessimistic. The Administrative Committee indicated that formal action on its part was not necessary at present.

Dr. Bereskin brought up the matter that necessary trips by the Chairman had not been budgeted, and that it might be reasonable to allow about $250 for the year. Dr. Hilliard made a motion that $250 be allowed. This was seconded by Mr. Williams and passed by the committee.

The Awards Committee report was communicated by S. J. Begun. Recommendations have been made as follows:

For the B. J. Thompson Award: J. S. Aagaard
For the W. R. G. Baker Award: R. E. Werner, J. P. Quitter, and Dome (Delta Sound System)

During the coming IRE National Convention, PGA will be responsible for $2\frac{1}{2}$ sessions as follows: Tuesday A.M., full session; Tuesday P.M., half session together with Professional Group on Broadcasting and Television Receivers; Wednesday P.M., full session planned for a tutorial program on stereo.

A report was read of Dan Martin's committee which suggested topics and authors for the audio section of the 50th anniversary issue of PROCEEDINGS OF THE IRE. These have already been sent to Dr. Goldsmith.

A discussion was held on ways for expanding these TRANSACTIONS. One suggestion was the reprinting of outstanding articles from foreign journals. Another was to have patent reviews on audio.

J. R. Macdonald moved that the next meeting of the PGA Administrative Committee be held in March, 1960 in connection with the IRE Convention. Seconded by Dr. Hilliard. Passed.

The meeting was adjourned at 10:05 P.M.

MARVIN CAMRAS
Secretary pro tem

CHAPTER NEWS

Baltimore

Louis R. Mills of Recording, Inc., spoke on "Comments and Review of the AES Show" and gave a demonstration of stereo microphones on October 20, 1959.

Paul Weathers of Weathers Industries, Inc., described "A New Method of Reducing Turntable Wow and Rumble" and demonstrated his stereo system incorporating a new cartridge and lightweight turntable on December 14, 1959.

Chicago

Roy R. Whymark of Armour Research Foundation presented a talk titled "Sound Vibrations in a Magnetostrictive Tube" on January 8, 1960. The January issue of *Scanfax* gave the following resume of his talk:

"Acoustical elements employing magnetostriction are used in wave filters, delay lines, accelerometers, and to a limited extent in phonograph pick-ups. Essentially, magnetostrictive elements convert electrical to acoustical energy. Whymark's paper indicates on general lines the many factors that influence the sound energy. A magnetostrictive ferrite tube is selected to illustrate the various effects.

"The tube is excited in different mechanical modes, and the vibrations interact in a complicated manner. Mode interactions are studied and their influence is related to such devices as stereo-pick ups, where channel suppression is limited by mode coupling. Deleterious effects also result in delay lines, wave filters, and many other devices using magnetostriction. Electrostatic and magnetostrictive excitations are provided in the experiments. Vibrations are detected using a frequency modulated system. Very small mechanical oscillations are detectable, comparable in amplitude to atomic dimension."

Whymark received his B.Sc. from London University, England, (in 1943). Until recently he was head of the acoustics division, Vickers Group Research Establishment in England. He has specialized in high-power transducer design, underwater acoustics, and physical acoustics.

At Armour, his group activities concern work in UHF sound generation, high-power electroacoustical converters, acoustical defect detection in solid films, ultrasonic remote control, and molecular acoustics. He is a member of the Acoustical Society of America, and the Physical Society of London.

ANNOUNCEMENTS

PGA Audio Sessions at the 1960 IRE National Convention

According to PGA Chairman Alexander Bereskin, PGA has contracted to handle 2½ sessions, which will be held in the Sert Room of the Waldorf-Astoria Hotel, New York, on Tuesday morning and afternoon, March 22, and Wednesday afternoon, March 23. The Wednesday afternoon session is being organized by Ben Bauer as a tutorial session on stereo. The Tuesday afternoon session will be shared with the Professional Group on Broadcast and Television Receivers; it will be organized by Alexander Bereskin and John Rankin. The Tuesday morning session, to be organized by Harry Pearson, will be a general one which should comprise the remainder of the audio papers.

Session 12—Audio

Tuesday Morning, March 22

Chairman: HARRY A. PEARSON, *Sonotone Corp., Elmsford, N. Y.*

12.1. **A Plotter of Intermodulation Distortion,** EDWARD F. FELDMAN, *Panoromic Radio Products, Inc., Mount Vernon, N. Y.*

12.2. **Listener Ratings of Stereophonic Systems,** HARWOOD B. MOORE, *Advanced Product Development Engineering, Radio Receiver Dept., General Electric Co., Utica, N. Y.*

12.3. **Calculations of the Gain-Frequency Characteristic of a Multimesh Transistor Amplifier Stage Using a Programmed Computer,** D. E. BRINKERHOFF, *Delco Radio Div., General Motors Corp., Kokomo, Ind.*

12.4. **Automatic Compensation of an Audio System Spectrum Operating with a Random Noise Input,** CHARLES E. MAKI, *MB Electronics, New Haven, Conn.*

12.5. **An Analysis of Factors Affecting Recording Reliability and Digital Tape Recorders,** KEN TAYLOR, *Ampex Corp., Redwood City, Calif.*

Session 20—Audio and Broadcast and Television Receivers

Tuesday Afternoon, March 22

Chairman: DANIEL W. MARTIN, *The Baldwin Piano Co., Cincinnati, Ohio.*

20.1. **The Present Status of Stereo Broadcasting,** C. G. LLOYD, *GE Special Electronic Components Dept., Auburn, N. Y.*

20.2. **Receiver Design Considerations for Stereophonic FM Multiplex Broadcasting,** C. G. EILERS, *Zenith Radio Corp., Chicago 39, Ill.*

20.3. **The Percival Stereophonic Sound System,** ARTHUR LEVIN, *Cossor, Ltd., Halifax, N. S., Can., and* L. F. BROADWAY, *EMI Electronics, Ltd., Hayes, Middlesex, Eng.*

20.4. **A Continuously Variable Wireless Remote Control for Stereophonic Phonographs,** A. A. GOLDBERG *and* ARTHUR KAISER, *CBS Laboratories, Stamford, Conn.*

20.5. **Automatic Stereophonic Phaser,** B. B. BAUER, A. A. GOLDBERG, *and* G. POLLACK, *CBS Laboratories, Stamford, Conn.*

Session 36—Stereophonic Sound Reproduction

Wednesday Afternoon, March 23

Chairman: BENJAMIN B. BAUER, *CBS Laboratories, Stamford, Conn.*

36.1. **Stereophonic Sound Reproduction,** H. F. OLSON, *RCA Laboratories, Princeton, N. J.*

36.2. **Psychoacoustics of Stereophonic Reproduction,** R. L. HANSON, *Bell Telephone Laboratories, Inc., Murray Hill, N. J.*

36.3. **Some Considerations in Design and Application of a Compatible Magnetic Tape Cartridge,** MARVIN CAMRAS, *Armour Research Foundation of Illinois Institute of Technology, Chicago, Ill.*

36.4. **A $1\frac{7}{8}$ IPS Magnetic Recording System for Stereophonic Music,** P. C. GOLDMARK, C. D. MEE, AND W. P. GUCKENBURG, *CBS Laboratories, Stamford, Conn.*

36.5. **Automated Magnetic Tape Cartridge Mechanisms,** J. D. GOODELL, *CBS Laboratories, Stamford, Conn.*

WESCON PAPERS DEADLINE SET FOR MAY 1, 1960

Authors wishing to present papers at the 1960 Western Electronic Show and Convention technical sessions to be held August 23–26 should register their interest by May 1. Required are 100–200-word abstracts, together with complete texts or detailed summaries. They should be sent to the Chairman of the Technical Program, Richard G. Leitner, WESCON Business Office, 1435 South La Cienega Blvd., Los Angeles 35, Calif.

Selection of papers for the program will be made before June 1 and authors will be advised of acceptance or rejection by that date.

Once again, there will be an IRE-WESCON CONVENTION RECORD published in advance of WESCON by the National Headquarters of the IRE.

Calibration and Rating of Microphones*

WILLIAM B. SNOW†

Summary—Calibration of a microphone consists of measuring its response to some known characteristic of a sound field under specified conditions. Usually the open circuit voltage for a one-microbar sound pressure is determined. Calibrations in an anechoic chamber give plane wave response, while those in a reverberant room give the response to sound arriving from all directions—at random incidence. Techniques have been developed for measuring response with considerable efficiency. Calibrations can be made with pure tones or with wide-band signals such as noise or warbled tones. Complete calibration includes measurements of directivity and impedance as well as linearity of response. From the calibrations it is possible to calculate ratings which give quick and relatively fair comparisons between microphones. The RETMA Rating is particularly effective for the commercial types with impedance below 200,000 ohms. For small crystal and condenser microphones a statement of the noise threshold is more indicative of true performance capability. Although complete specification of microphone performance requires considerable information, the ratings dispel the main ambiguities in response figures arising from differences in impedance, circuits and test sound pressures.

THE subject of microphone calibration and rating has received a great deal of study. Serious efforts have been made to provide standards which will rate microphones in simple and unambiguous terms under all conditions. While the desired simplicity has not been fully attained, effective methods have been developed. Part of the difficulty in obtaining a universal system is that different requirements can best be satisfied by differing methods of specifying the characteristics. In this paper, I shall point out several techniques that have been proposed and used, trying to demonstrate the applicability and advantages of the various systems.

Frequency Response

Definition

A sound wave creates changes in various properties of the supporting medium, such as the pressure, particle velocity, density, or temperature. In a free progressive wave all of these changes in properties are interdependent and directly calculable from each other, so that a measurement of any one would define the sound field. When reflecting surfaces, obstacles, or inhomogeneous conditions in the medium modify the sound field, these simple relationships are altered. Before we can rate microphones we must measure their salient characteristics in responding to these sound field characteristics; that is, we must calibrate them. Most commonly used microphones measure the sound pressure variations, or the pressure gradient developed between two points in space. The human ear also responds to the sound pressure. The usual practice, therefore, is to calibrate microphones in terms of their response to this property of the sound field.

It is generally agreed that the most useful microphone calibration is that which gives the open circuit voltage generated by the instrument in response to a specified sound pressure, over the frequency range of interest. This is defined as frequency response.[1] Usually the pressure is taken as one dyne per square centimeter (one microbar) and the voltage is expressed as decibels relative to one volt. There are several good reasons for this.

1) Although we cannot always isolate the open circuit voltage because of impedance difficulties, we can always measure it readily. This seemingly paradoxical statement will be explained in detail below.

2) If the open circuit voltage and the microphone impedance are known, the microphone performance can be calculated for any condition of loading.

3) The open circuit voltage can be combined with a simple measurement to completely define the performance of the microphone with any load circuit.

4) It corresponds to an effective condition of use. A microphone should face a high impedance to yield maximum signal-to-noise ratio.

5) In microphones connected to vacuum tubes, voltage is all that is needed. It is not necessary to extract power from the sound wave, for the amplifier supplies the energy.

6) When the microphone faces an impedance high compared with its own, variations in microphone impedance do not cause variations in response.

Ambient Conditions for Calibration

The *frequency response* of a microphone is almost always determined by measuring its output in a sound field which has been adjusted to a known value established with the aid of a standard microphone. There are various ways of calibrating standard microphones, but the one in most common use at the present time is the reciprocity method, which makes minimum demands for equipment, computation, and assumption.[1,2]

* Manuscript received by the PGA, October 5, 1959.
† Santa Monica, Calif.

[1] L. L. Beranek, "Acoustic Measurements," John Wiley and Sons, Inc., New York, N. Y.; 1949. A general reference for many of the subjects discussed in this paper.

[2] H. F. Olson, "Acoustical Engineering," D. Van Nostrand Co., Inc., New York, N. Y.; 1957. A general reference for many of the subjects discussed in this paper.

For simplicity, we shall assume that we have such a standard microphone available, accurately calibrated.[3,4] Two types of response measurement, free field calibration, and random field calibration, are found to be of most value. For the former we require a sound field which consists of plane progressive waves, or of spherical waves far enough from a small sound source so that the wave may be considered essentially plane. Reflections of sound from directions other than that directly toward the source must be negligible. The word negligible is used arbitrarily, but the sum of all reflections must be held more than 20 db below the direct wave if they are to cause less than ± 1 decibel error in the response. Such sound fields can be obtained out of doors in a large open space, or in an anechoic chamber. For outdoor measurements, the sound source may be mounted in the plane of the ground, with the microphone suspended above it, or both microphone and source may be mounted on tall poles so that the distance to reflecting objects is large compared to the distance between instruments. While outdoor measurements may be cheaper, conditions are much more satisfactory in an anechoic chamber where the tests are more convenient and are not subject to the vagaries of weather and ambient noise.[5–8]

Anechoic chambers for frequencies above 250 or 300 cps need be only about 5 feet on a side. For the low frequencies a pressure microphone can be calibrated in a closed box driven by a loudspeaker. Directional microphones can be accurately calibrated in a long, hard (concrete or metal walls) tube about 20 inches square and 20 to 25 feet long containing a loudspeaker at one end and a 10-foot absorbing wedge structure at the other. At low frequencies such a tube propagates essentially a plane wave.[9]

When a microphone is used in a reverberant room, sound arrives from all directions at random. This is also frequently true for outdoor conditions where sound sources are distributed in many directions. Under these circumstances what is known as a random incidence calibration will be desired. For simple microphone shapes this calibration can be calculated with reasonable accuracy, or it can be computed from measured directivity patterns. A more direct method is to measure the microphone output vs that of a standard whose random incidence response is known, when both are placed in a random sound field. This type of field can be obtained in properly-designed reverberation chambers.[10,11]

Technique of Calibration

The fundamental calibration arrangement is illustrated by Fig. 1. A sound field is produced by an oscillator or noise generator connected to a loudspeaker; or a mechanical source or airjet might be used with proper stabilizing control. The standard microphone is placed in the sound field at the test point, and the field is calibrated by measuring the sound source input and the standard microphone output with the meters $M1$ and $M2$. Next the test microphone is substituted at the test position and its open circuit voltage is measured for the calibrated sound field. This procedure is repeated for each frequency of interest.

Precautions in the surroundings have been described above. In addition, it is necessary to use a distance D appropriate to the source dimensions to insure adequate approach to plane waves. The size of the source must be small enough compared to the sound wavelength to simulate a point source with D, the distance used. This commonly involves the use of two or more loudspeakers to cover the audio frequency band, for accurate work. These remarks are intentionally only qualitative for microphone calibration requires experience. Representative values of D are 2 to 5 feet. It is usually feasible to hold the loudspeaker diameter to less than a wavelength except above 5000 cps.

Of course in practice there are many tricks to simplify and speed up the measurements and the computations. One of the simplest is to mount both the standard and test instruments on a frame so that they can be alternated. For each frequency, the sound field is set at a known value, using the standard. Then the frame is moved to put the test microphone in position, and the *frequency response* is read directly.

Another effective method (dotted in Fig. 1) is to insert in the sound field a standard microphone which is connected to equipment that controls the output of the sound source.[12] This combination can be made to maintain a constant sound pressure as a function of frequency. Thus, the test microphone is always in a constant-pressure sound field, and it is merely necessary to measure its open circuit voltage. This can be done with automatic equipment which plots a continuous curve giving the *frequency response* directly. A simpler

[3] "Pressure Calibration of Laboratory Standard Pressure Microphones," Amer. Standards Assoc., Z24.4; 1949. (Obtain latest issue.)

[4] "Free-Field Secondary Calibration of Microphones," Amer. Standards Assoc., Z24.11; 1954. (Obtain latest issue.)

[5] L. L. Beranek and H. P. Sleeper, Jr., "Design and construction of anechoic sound chambers," *J. Acoust. Soc. Amer.*, vol. 18, p. 140; 1946.

[6] M. S. Corrington, R. L. Libbey, and S. V. Perry, "The AF anechoic chambers at Cherry Hill," IRE Trans. on Audio, vol. AU-4, pp. 161–166; November–December, 1956.

[7] R. L. Berger and E. Ackerman, "The Penn State anechoic chamber," *Noise Control*, vol. 2, pp. 16–21, 63; September, 1956.

[8] B. G. Watters, "Design of wedges for anechoic chambers," *Noise Control*, vol. 4, pp. 32–37; November, 1958.

[9] R. S. Dadson and E. G. Butcher, "Recent work at the National Physical Laboratory on the free field calibration of microphones," in "Proceedings of the First ICA-Congress: Electro-Acoustics," C. W. Kosten and M. L. Kasteleyn, Eds., p. 103; 1953.

[10] J. H. Botsford, R. N. Lane, and R. B. Watson, "A reverberation chamber with polycylindrical walls," *J. Acoust. Soc. Amer.*, vol. 24, pp. 742–744; November, 1952.

[11] D. R. McAuliffe, "Design and performance of a new reverberation room at Armour Research Foundation, Chicago, Illinois," *J. Acoust. Soc. Amer.*, vol. 29, p. 1270; December, 1957.

[12] A. L. Seligson, "Free-field technique for secondary standard calibration of microphones," *J. Audio Eng. Soc.*, vol. 4, p. 110; July, 1956.

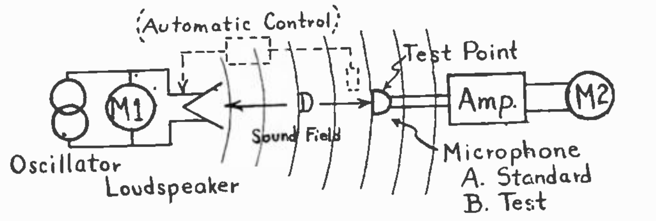


Fig. 1—Basic calibration arrangement. A sound field is generated by an oscillator and loudspeaker (or noise generator). The sound field is measured by a standard microphone. Then the output of the test microphone is read using the same sound field. A standard microphone can be used to control the sound field automatically, reducing the testing time or allowing more data to be taken. A calibrated amplifier and meter $M2$ are required.

method employs the automatic sweeping of the test tone and recording of the microphone output, but omits the automatic control. A template is prepared of the sound field as measured by the standard microphone. This is laid over the recorded curves for comparison.

Insert Voltage Test Method

Measurement of open circuit voltage is done by the insert voltage method based upon the Thévenin Theorem. As far as a load circuit is concerned, any generator can be replaced by a circuit consisting of a voltage equal to the open circuit generator voltage, acting in series with the impedance of the generator looking back into its output terminals.[13]

Schematically, the application to microphone calibration is shown by Fig. 2. The circuit must be set up so that $R1$ is much smaller than the microphone impedance (usually one to ten ohms) so that the voltage across it depends only upon the resistance of $R1$. Thus, this voltage is "inserted" in the microphone circuit in series with the microphone impedance, and acts as if it were generated by the microphone. First, the switch S is thrown to connect the oscillator output to the amplifier. Readings on $M2$ and $M3$ are recorded, $M3$ giving the output of the standard microphone for the loudspeaker output resulting from the input read on $M2$. Now the switch is thrown to connect the oscillator to the attenuator, cutting off the loudspeaker tone and applying voltage across $R1$ in series with the microphone. The attenuator is adjusted until $M3$ gives the same reading as before, signifying that the microphone circuit is supplying the same voltage to its amplifier as when it was excited acoustically. The value of this voltage is obtained from the reading of $M1$, the setting of the attenuator, and the attenuation of the voltage-divider network $R2$–$R1$. Since it is introduced into the circuit in series with the actual microphone impedance, and yields the same amplifier output as the acoustical signal, it is equal to the open circuit voltage of the microphone. From the *frequency response* calibration of the standard microphone the value of sound pressure is obtained.

[13] Beranek, *op. cit.*, p. 598.

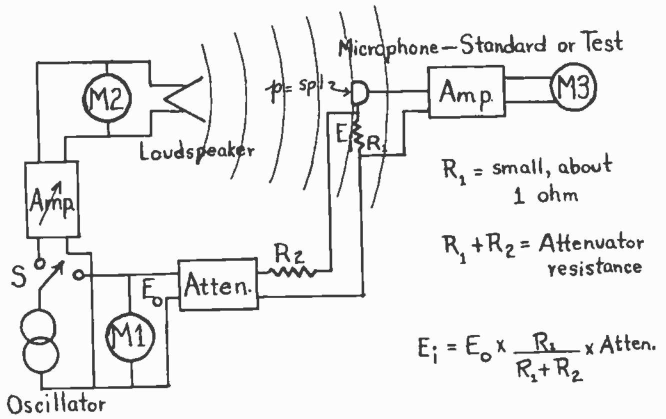


Fig. 2—Thévenin theorem, or insert voltage, schematic arrangement. The open circuit voltage of the microphone is measured by determining an electrically inserted voltage, E_i, which gives the same output as the acoustical signal. The microphone amplifier and meter $M3$ need not be calibrated, or uniform with frequency.

Now the test microphone is substituted and the process is repeated with the same value of loudspeaker input, $M2$. This will yield a new value of $M3$, and of insert voltage across $R1$ to produce it. The new insert voltage is the response of the test microphone, *i.e.*, the open circuit voltage for a known sound pressure. Usually all values are recorded in db, giving voltage level and sound pressure level. A series of these measurements over the frequency range gives the *frequency response* calibration of the test microphone.

Notice that this measurement makes no demands whatever on the characteristics of the loudspeaker, microphone amplifier or indicating meters $M2$ and $M3$. Its accuracy depends only upon knowledge of the attenuation system and its input voltage read on $M1$. This substantiates the statement made above, that while we cannot always isolate the open circuit voltage, we can measure it, in the sense that we have determined another voltage which is exactly equal to it. And it vastly simplifies the equipment problem.

A noise generator can be used in place of the single frequency oscillator shown in Fig. 1, if this is desired. When noise is used, however, it is important to remember that the voltage inserted in series with the microphone must have exactly the same frequency characteristic as the loudspeaker output. When noise is used, it is ordinarily confined to narrow frequency regions with filters, which makes this a reasonably valid assumption. If a mechanical source is employed, the filtering must be done in the microphone circuit, and corrections must be determined from insert voltage calibrations of the amplifier with standard and test microphones.

Although in audio frequency studies the phase of the response is not usually needed, it can be measured by this method. A phase indicator would be added to $M3$ and phase shifting networks would be required in the attenuator circuit to operate on the insert voltage. These measurements would usually be both arduous and tedious. Probably *easier* phase-testing methods could be devised to supplement the amplitude measurements carried out by the insert method.

The various semiautomatic and automatic methods differ in detail from the schematic of Fig. 2, but their fundamental operation is the same. Initial system calibrations are set up by the insert voltage method, and hold only for those particular components. When an unknown microphone is calibrated, it is necessary to make an acoustical run, followed by an insert voltage run to determine the effect of the measuring amplifier upon the microphone.

Certain precautions must be observed. A microphone circuit has low output, and ground loops introducing hum must be avoided. Shielding must be adequate to avoid hum pickup or actual signal cross-feed from the high-level oscillator circuits. In particular, though, it is necessary to insure that the voltage is inserted "electrically close" to the microphone, so that it actually simulates an internally generated voltage without being affected by stray impedances. This is usually simple for low impedance microphones; it may not be for capacitive types with long cables, as shown in Fig. 3.

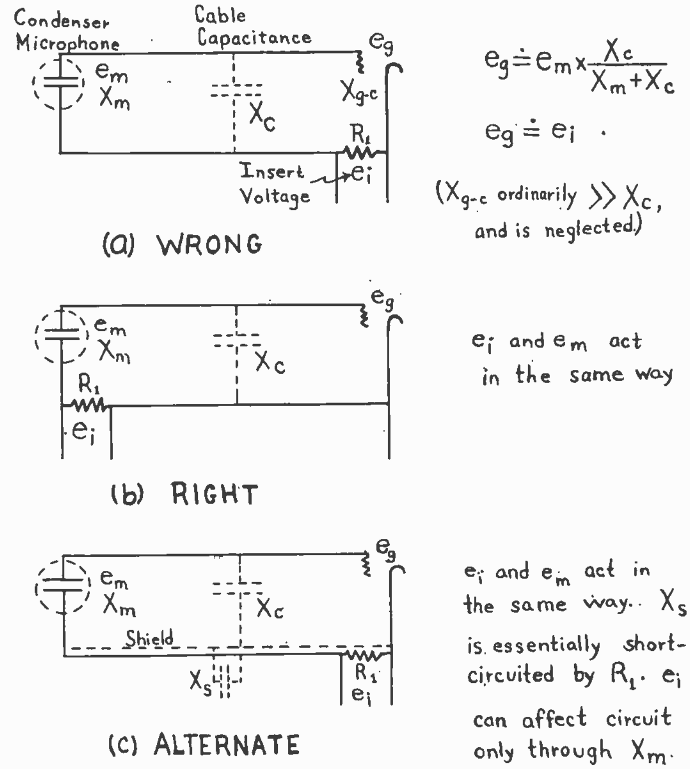


Fig. 3—Insert voltage precautions. (a) The insert voltage is in series with the capacitance of the microphone and the cable capacitance in parallel, whereas the generated microphone voltage is in series with these two reactances in series. The grid voltage will be higher for the insert voltage than for the generated voltage. (b) Insert voltage and generated voltage are applied to the circuit at the same point and are equivalent. (c) The high-potential side of the insert voltage is shielded from the central cable conductor, and in effect is applied as in case b. Reactance X_S is shunted by a very low resistance and can be neglected.

Directivity

Choice of Orientation

Nearly all microphones have directivity in at least part of the operating frequency range, unavoidable because of practical design considerations. In addition, many microphones are given intentional directive characteristics. Therefore, it is necessary to decide which angular orientation is to be used for the calibration, and this depends upon probable use. For example, a microphone ordinarily used for close pickup, with its diaphragm perpendicular to the direction from the source, should be calibrated with the same orientation (called normal sound incidence, normal indicating perpendicular, not "most used"). Because of sound diffraction, microphones with uniform pressure response tend to have a broad peak at high frequencies for free-field normal incidence, but this orientation gives maximum signal-to-noise ratio. If it is desired to have a more uniform frequency response, the microphone should be mounted, both for calibration and use, with the diaphragm parallel to the direction to the source (parallel or 90° sound incidence).

Directivity Measurements

When complete directivity information is desired, the procedure is similar to the free-field method prescribed above. The instrument is rotated successively to various angles around its principal axes, and frequency response runs are made at each position. Alternatively, the frequency is set at various values, and response is measured while the microphone is rotated. It can be seen that the complete measurement of directivity characteristics is a time-consuming activity. Fortunately, the directivity characteristic of microphones of a given geometry does not vary much from unit to unit. Therefore, the elaborate directivity measurements are necessary only on occasional samples, and it is ordinarily necessary to measure the directivity pattern only in one or two important planes for individual instruments.

Directivity measurements make severe demands upon the quality of the sound field. For example, if a reflection sound pressure of -20 db exists, and the microphone has a directivity discrimination of 20 db, but happens to be pointed toward the reflection, the output due to the direct and the reflected pressures will be the same. Since the phase may be anything, the net microphone output may vary from a very low value to double the proper value. For a reflection to be -20 db at a surface, its absorption must be 99 per cent. It is evident that particular care must be exercised during directivity measurements.

Impedance

Microphone impedance can be measured on appropriate impedance bridges, or by the various simplified methods described in the literature. The impedance includes the ac electrical resistance, plus the reflected mechanical and radiation resistance and reactance. Consequently, the measurements should be made with a completed microphone operating under normal conditions for high accuracy. Knowledge of the impedance is useful for several reasons. When we know the impedance as a function of frequency, and the open circuit voltage (*frequency response*), we can compute the microphone performance for any condition of loading. Second, the impedance tells us what type of amplifier must be used with the microphone. Third, the fundamental limit to microphone sensitivity is self-noise.

Self Noise—Resistance Noise

This limiting noise arises from the resistive portion of the impedance, and usually is easier to calculate than to measure. The resistance will ordinarily be a function of frequency, and can be expressed as either an equivalent series or a parallel element, depending on which is more convenient.[14] The self-noise is usually calculated as an equivalent noise voltage that acts in series with the effective resistance component of the microphone in determining the noise voltage applied to the external circuit. The reactive component of microphone impedance must be properly included in the calculation. Noise voltage is dependent upon absolute temperature, but the variation over the ordinary room temperature range is less than 1 db, and can usually be neglected.

It is convenient to remember that the equivalent noise voltage in one ohm for one cycle bandwidth is −198 db re one volt. For any generator, add 10 log (resistance) and 10 log (frequency bandwidth in cps). These values can be read on a slide rule or obtained from a table of power ratio vs db. As an example, consider a 30-ohm microphone in the octave band between one and two kc. We have:

$$\begin{aligned} \text{noise voltage} &= -198 + 10\log 30 + 10\log 1000, \\ &= -198 + 15 + 30, \\ &= -153 \text{ db re 1 volt.} \end{aligned}$$

If the microphone can be mounted in a very quiet place, the self-noise output can be measured. This is frequently feasible at the higher frequencies where acoustical shielding is relatively effective, and background noise tends to be low. Another method which approximates the desired condition is to employ a dummy microphone which has the same impedance (static) but does not generate a voltage. In general microphones are inefficient and the dynamic impedance contributions are small.

Amplifier Coupling

In spite of the importance of a knowledge of microphone impedance as detailed above, it should be remembered that this knowledge is not required for experimental use of the instrument. By means of the insert voltage measurement previously described, the correct *frequency response* of the microphone with any coupling circuit is readily obtained. The ratio of the output to the inserted voltage gives the amplification of the amplifier working from the microphone as a generator. This may be quite different from the characteristic of that amplifier working from a resistive generator. It will be seen that this method obviates any impedance measurements or calculations but gives complete information on the output of the amplifier in terms of sound pressure as well as in terms of the generated microphone voltage. In many cases this is sufficient. If the *frequency response* of a microphone is known it can be connected to any amplifier. When the amplification *of this combination* is measured by the insert voltage method the microphone response in db can be added directly to the measured amplification in db to give the amplifier output in terms of sound pressure level.

Linearity

Over the usual dynamic range of sound pickup, microphone linearity of response is excellent. Nevertheless, for a complete characteristic measurement, nonlinearity should be determined. This is not an easy test. It is necessary to have an intense sound field in which the test sound pressure is sufficiently pure; and since both the air and the sound-generating loudspeaker are very apt to distort the wave, obtaining sufficient purity of tone requires extreme precautions. It is usual, for this type of measurement, to employ some form of tuned acoustic cavity to suppress unwanted components. The microphone output may be given as a ratio of fundamental to harmonic voltages, using an harmonic analyzer; or an intermodulation measurement may be made. The latter is the more difficult test because it requires a driving field with two pure frequencies. An approximate method consists in plotting the input vs the output and selecting the point where the output falls significantly "below the line"; for example, 0.5 or 1 db.

In practice, unless a microphone is in bad repair, it will have negligible distortion compared to other parts of the system up to a sound pressure level of 120 db for frequencies in the audible range. The standard microphones used on sound level meters are satisfactory to 140 or 150 db. Special microphones are linear to above 200-db sound pressure level.

Microphone Ratings

The fundamental performance data enable us to calculate or measure what the characteristics of the microphone will be when it is considered as an element of a system. A microphone rating system which would sum up all needed knowledge about a microphone in a single figure is obviously very desirable. In the general case this appears visionary. But in the application of microphones to sound recording, broadcasting, sound rein-

[14] W. B. Snow, "Audio frequency input circuits," *J. Audio Eng. Soc.*, vol. 1; January, 1953.

forcement and noise measurement, many features of operation have become relatively standard, and rating systems have become useful.

Romanow-Hawley Rating

Strenuous attempts in this direction have been made and the most complete method is due to Romanow and Hawley, who developed a rating system for all elements of a sound-reproducing assembly.[15,16] They based their system upon the idea that if, at each junction point, the energy transferred is used as the criterion, data which is independent of impedance level will result. The system prescribes methods of testing each part which result in ratings which can be added together to get an over-all system response directly.

For the microphone the required data are the open circuit voltage response and the impedance. The rating is defined in terms of the power available to a resistive load equal to the nominal microphone impedance, which is taken as the impedance at 1000 ± 100 cycles. Power available is the power that a voltage equal to the microphone open circuit voltage would deliver to a resistance equal to the nominal microphone impedance when acting in series with a generator of the same resistance. The voltage is taken as that corresponding to reference sound pressure (0.0002 microbar) and the power with reference to one milliwatt. In addition, there is a coupling factor which takes into account the differences in impedance between microphone and amplifier, plus cable impedance. This method is quite general and can be used at any frequency. When the microphone and loudspeaker impedances are nearly constant with frequency, their ratings may be added to the amplifier ratings to gain a reasonably accurate idea of the over-all system response. If reactive microphones or loudspeakers are used, the coupling factors are important and must be calculated.

RETMA Rating

A simplified method has been standardized by the Radio Electronic Television Manufacturers Association (now Electronic Industries Association) for use with microphones for sound systems.[17,18]

This uses the above rating system as a base but assigns rating impedances in a few steps from 38 to 100,000 ohms for microphones, and calculates only at 1000 cycles. The problem of coupling factors is handled for high impedance microphones by assigning them 100,000 ohms resistance for rating impedance, regardless of the actual impedance.

A Microphone System Rating, G_M, is computed as follows:

$$G_M = \left(10 \log_{10} \frac{E^2/4R_{MR}}{p^2}\right) - 44 \text{ db},$$

where

E = open circuit voltage generated by the microphone,
p = sound pressure in microbars,
R_{MR} = microphone rating impedance.

This G_M is the ratio in db relative to 0.001 watt and 0.0002 microbar of the electric power available from the microphone to the square of the undisturbed sound field pressure in a plane progressive wave at the microphone position. In terms of the *frequency response* described earlier, it is.

$$G_M = (20 \log E/p) - 10 \log R_{MR} - 50 \text{ db}.$$

The first term is the usual microphone sensitivity in volts per microbar, expressed in db. In many instances catalog information gives sensitivity in volts per *10* microbars. In these cases, 20 db must be subtracted before applying the formula.

Representative Examples—Broadcast Microphones

The best way to demonstrate the usefulness of the RETMA Rating is to calculate a few examples. Table I gives data for four types representative of microphones used for broadcast and sound recording pickup today. The column headings are in accordance with the formulas given above except for the last one. This represents the net power available using the real impedance rather than the rating impedance, and making allowance for the capacitive generator in the case of the condenser microphones.

The dynamic and velocity microphones represent the class of inherently low-impedance instruments operating on magnetic principles and having essentially resistive impedance. The 1-kc response covers a range of 30 db; but this is because of impedance differences. The real difference, as shown by "Net power available," is only 2 db. G_M indicates 1 db because of the arbitrary assignment of rating resistance. It is obvious that the rating system is useful for these microphones.

The data on condenser microphones is different. Such microphones have a relatively high response, but also a high impedance. Both instruments have R_{MR} of 10^5 ohms, and a G_M of -150, in spite of the fact that one has 8.5 times the impedance of the other. The actual power that could be drawn is represented by the last column, where the figures are -162 and -171 respectively. It would be even lower if the microphone were connected to a 100,000-ohm load. Thus, the rating gives a false impression of the microphone in comparison to

[15] F. F. Romanow and M. S. Hawley, "Proposed method of rating microphones and loudspeakers for systems use," Proc. IRE, vol. 35, pp. 953–960; September, 1947.
[16] Beranek, *op. cit.*, ch. 13.
[17] "Microphones for Sound Equipment," Radio Electronic Television Manufacturers Association Standard, No. SE105; August, 1949.
[18] "Radiotron Designer's Handbook," F. Langford-Smith, Ed., 4th ed., distributed in U. S. by Tube Division, RCA, Harrison, N. J., chs. 18 and 19; 1953.

TABLE I

G_M RATINGS OF REPRESENTATIVE BROADCAST MICROPHONES

Type	1-kc response	Impedance	R_{MR}	G_M	Net power available
Dynamic	−88	30 ohms resistive	38	−154	−153
Dynamic with transformer	−58	30,000 ohms resistive	40,000	−154	−153
Velocity (ribbon) with transformer	−81	250 ohms resistive	150	−153	−155
1-inch condenser	−50	50 mmf capacitance $-j\,3.2\times10^6$	100,000	−150	−162
$\frac{5}{8}$-inch condenser	−50	6 mmf $-j\,26\times10^6$	100,000	−150	−171
Condenser with cathode follower	−50	10,000 ohms resistive*	9600	−140	−135

* Permissible load resistor.

TABLE II

G_M RATINGS OF OTHER REPRESENTATIVE MICROPHONES

Type	1-kc response	Impedance	R_{MR}	G_M	Net power available
Dynamic (repeated)	−88	30	38	−154	−153
2-inch Rochelle salt crystal	−58	2200 mmf* capacitance $-j\,72{,}000$	40,000	−154	−153
2-inch Rochelle salt crystal, 25-foot cable	−60*	2950 mmf* capacitance $-j\,54{,}000$	40,000	−156	−154
1-inch hypothetical titanate, 6-foot cable	−65	580 mmf $-j\,280{,}000$	100,000	−165	−166
$\frac{1}{2}''\times\frac{1}{2}''\times\frac{1}{4}''$ variable reluctance	−76†	1850 ohms inductive	2400	−160	−158

* Subject to wide variations with temperature.
† 2000-ohm resistive load.

the others. Of course, condenser microphones are not used as power converters in practice. They are directly connected to vacuum tube circuits, usually cathode followers, to obtain extra-high impedance. The last line in Table I shows data for the case where a cathode follower is included as an integral part of the microphone, on the assumption of a generator resistance of 1000 ohms and a permissible load resistance of 10,000 ohms, reasonable figures for this class of device. G_M is now −140 and available power −135 db (re 1 milliwatt for 0.0002 microbar sound pressure). The microphone is now a better power converter than the others. However, this is fictitious, because the vacuum tube is supplying the power. Is this combination superior to the other microphones in the table? The rating system does not give the full answer for this type of instrument, although it tends in the right direction.

Representative Examples—Other Types

Table II illustrates similar calculations on other types. The data for the 30-ohm dynamic microphone is repeated for comparison. A 2-inch Rochelle Salt unit, used extensively for sound measurement work, and representative of good crystal microphones, is shown next. Although this is a capacitive microphone, its impedance is below 100,000 ohms and G_M rates it well. The same microphone is also shown with a 25-foot cable. With the large capacitance of Rochelle Salt, reasonable cable lengths are feasible. It must be pointed out that the data are for 77°F temperature. The capacitance of Rochelle Salt varies greatly with temperature, and so these figures are optimum. The output *with cable* can be as much as 6 db lower at 30° and 110°F.

A hypothetical lead zirconium titanate microphone is shown. This would be a broadcast-fidelity unit, highly resistant to temperature and humidity effects. Probably no general purpose microphone would be built with lower capacitance. The G_M is a reasonable rating for this unit.

Finally, data is shown for a "Tiny Tim" microphone only $\frac{1}{2}\times\frac{1}{2}\times\frac{1}{4}$ inch in dimensions, operating on the variable reluctance principle over the voice frequency range. It has a medium impedance, and therefore a higher *sensitivity* number (−76) than the dynamic or velocity instruments. The G_M, however, is 6 db lower, reflecting the very small size of the unit.

The G_M Rating

This rating is seen to be very effective in describing the inherent capability of microphones of all types when their impedance values lie below about 200,000 ohms. This includes all instruments that can be used alone as practical circuit elements. With the rating, statements must be made about frequency range covered and directionality.

Other Rating Methods

Alternate rating concepts have been proposed. One due to Massa[19] gives the maximum electrical power the microphone will deliver to a resistive load for a given

[19] F. Massa, "Microphone efficiency: a discussion and proposed definition," *J. Acoust. Soc. Amer.*, vol. 11, p. 222; 1939.

acoustical power intercepted by the diaphragm. This is akin to the system described above, and results using this method show that most microphones are only of the order of 1 per cent efficient. Baerwald[20] has suggested a different measure, where the absolute efficiency represents the amount by which the microphone output exceeds its self-noise. While the following does not follow Baerwald's exact method, it is based upon the same concept.

Octave Band Noise Threshold

High output is important in a microphone but more important, for most applications, is high signal-to-noise ratio. This can be specified in various ways, but they all depend upon a definition of the signal level. The self-noise of the microphone and amplifier system that actually sets the lower limit, can be described in terms of an equivalent sound pressure level that would produce the same output. Any signal lower than this would be usable only with difficulty. While it cannot be called the limit of detection, it is a practical working threshold.

The background noise in locations where microphones are used is usually specified in octave bands. Therefore, it has been found convenient to give the equivalent noise threshold in the same terms. Since the rating frequency is 1 kc, the octave band 700 to 1400 cps will center around it. Levels in this band will be essentially the same as in the standard 600–1200 cps band, and the latter was used in the calculations to agree with measured results.

Table III shows octave band-noise threshold levels for the various microphones when they are operated into a high impedance. If the microphone is immersed in a sound field of this sound pressure level (spl) in the 600–1200 cps octave band, half the output power will be caused by the sound and half by the self-noise of the system. For the low-impedance types, sufficient step-up in the input transformer can be obtained to make tube noise negligible. Therefore, the self-noise arises in the microphone resistance and the threshold is calculated as follows, from data on *response* and impedance, using the 30-ohm dynamic instrument as an example:

$$\text{noise level} = -198+10\log 30+10\log(600\text{ cps}),$$
$$= -198+15+28 = -155\text{ db re 1 volt;}$$

$$\text{sensitivity} = -88\text{ db re 1 volt for 1 microbar, which is 74 db spl;}$$
$$= -88-74 = -162\text{ for 0.0002 microbar, which is 0 db spl;}$$

signal must be $-155-(-162)=7$ db higher to equal noise;

octave band threshold therefore = 7 db spl.

[20] H. G. Baerwald, "The absolute noise level of microphones," *J. Acoust. Soc. Amer.*, vol. 12, p. 131; 1940.

TABLE III

OCTAVE BAND-NOISE THRESHOLDS

600–1200 cps band
db above 0.0002 microbar

Type	Threshold db
Dynamic, 30 ohm*	7 calculated
Dynamic, 30,000 ohm*	7 calculated
Velocity, 250 ohm*	9 calculated
1" condenser, 50 mmf†	7 measured
1" condenser with grid resistor, 50 mmf†	15 measured
⅝" condenser, 6 mmf†	14 measured
2" and 1" rochelle salt crystal on sound level meter‡	9 measured
1" titanate, 6-foot cable‡	16 estimated
½"×½"×¼" variable reluctance*	11 calculated

* Impedance transformed to a level making tube noise negligible.
† With associated cathode follower coupling tube.
‡ 10 megohm input to amplifier.

For the low-impedance types, the thresholds rise about 3 db if they are matched, instead of being operated open-circuit.

In the case of the capacitance microphones the thresholds are calculated in the same way, but the self-noise must be measured because it is amplifier noise and depends upon the coupling and circuitry. The values used are representative of modern instrumentation.

If the three tables are compared it will be found that the four low-impedance instruments are rated relatively in about the same way by the G_M and threshold methods. It turns out that this is also true of the crystal microphones, since they have been designed in an impedance range to compete with the magnetic types.

The performance of the condenser microphones is determined by the amplifier performance that has been attained. At present, they are seen to be in the same range as the magnetic units. Further reduction in noise threshold would result if leakage-current paths and other noise-producing effects in the tubes could be reduced.

It should be mentioned that the crystal and titanate types could probably be made to have the same threshold as the 1-inch condenser microphone if they were designed for higher output and lower capacitance, to be used with a cathode follower. Such materials have a volume resistivity that would set an ultimate limit, however. Reduction of threshold for the magnetic types requires improvements in the instruments, such as greater magnetic flux density.

CONCLUSION

The G_M system of rating true microphone sensitivity is a very useful technique, especially for the magnetically operated instruments whose impedance is largely resistive. It is also effective for capacitive microphones with capacitance above 500 mmf, which includes the general purpose types. The microphones which have attained wide use have similar G_M ratings. The G_M system is particularly suitable for broadcasting and sound recording situations.

With instruments of small capacitance, such as condenser microphones, the G_M value does not give a true picture of performance. It rates the microphones as relatively more sensitive than they are.

Since microphones are frequently used with a high impedance termination, and the condenser microphones must be used in this way, no power is supplied by the unit. A method of rating which gives the octave-band sound-pressure level that produces a signal just equal to the self-noise of the microphone and amplifier in the same band is a useful device that will handle all cases. It corresponds to G_M ratings for the resistive microphones where the amplifier noise contribution can be made negligible by using input transformers which raise the resistance noise voltage above tube noise voltage. For the capacitive types it gives a rating that depends upon the amplifier noise and microphone *response*, with some effect from the microphone capacitance. It provides a true picture of the signal-to-noise capability of the microphone as actually used, but is more complicated to measure than G_M. The threshold system has particular advantages for sound-level measurement since it gives directly the lower limit of feasible tests.

Complete specification of microphone performance requires a great deal more information than a single rating number, as outlined above. Nevertheless, the rating numbers are extremely useful because they dispel the main ambiguities in response figures which arise from different impedance levels, operating circuits, and definitions of terms used in presenting the data.

Stereophonic Projection Console*

B. B. BAUER† AND G. W. SIOLES†

Summary—**A system is described for home stereophonic reproduction from a single cabinet by reflection of sound from the room boundaries. The effect is emphasized by using loudspeakers which maintain a uniform cardioid directional pattern over the useful frequency range. The directional properties are obtained with acoustic phase-shift networks.**

Introduction

FOR quite some time, it has been evident that a stereophonic sound system contained in a single cabinet would be of considerable interest to many users. This is because the correct installation of a multiunit system may be in conflict with the seating arrangement, the artistic decor, or the space limitations of a typical living room. There are two ways of providing stereophonic performance with a single unit instrument:

1) by making it sufficiently large to establish a suitably long base-line for the sound sources, or
2) by causing the sound to be reflected from the room boundaries.

As the former approach involves dimensional problems, the second one was chosen. The instrument to be described is provided with sufficiently directional radiators so that those sounds which are important for directional perception will be radiated preferentially toward the room boundary and thence reflected to the listener. A lesser portion of sound reaches the listener directly and the combined direct and reflected sounds give an over-all result which is comparable to that obtained with multiunit systems.

* Manuscript received by the PGA, September 21, 1959.
† CBS Labs., Stamford, Conn.

Projection Console

Several attempts have been made in the past to produce single unit stereophonic reproducers. Camras[1] used two loudspeakers at the opposite end of the cabinet, divergently inclined towards the rear, so that the wave would reach the listener after reflections from the back and the side walls. De Boer[2] arranged two loudspeakers facing outwardly at the end of the cabinet and used two doors opening forward so that the direction of the high frequencies could be controlled by the angular position of the doors. Levy, Sioles, Carlisle, and Brociner[3] described a similar arrangement with the doors interposed between the loudspeakers and the listeners. Such systems are effective when the wavelength of sound is smaller than the dimensions of the loudspeakers and the doors. Their effectiveness diminishes at longer wavelengths which still are important for stereophonic reproduction. The "projection console" to be described depends upon dipole radiators and acoustic phase-shift

[1] M. Camras, U. S. Patent 2,710,662; 1955.
[2] K. de Boer, U. S. Patent 2,610,694; 1952.
[3] S. Levy, G. W. Sioles, V. Brociner, and R. Carlisle, "Listener reaction to stereophonic reproduction by reflected sound," *J. Acoust. Soc. Amer.*, vol. 31; 1959.

networks to obtain suitable directional properties at wavelengths comparable with or longer than the dimensions of the baffle. Therefore, its operation, to a great degree, is independent of the physical dimensions of the instrument.

The principle of the projection console is shown in Fig. 1. In this arrangement, the horizontal radiation patterns of the loudspeakers are caused to be of a heart-shape (cardioid), although other members of the limaçon family may be chosen. For centrally located listeners and with each channel operating independently, the direct sounds are considerably attenuated because of the angular relationship with the radiation pattern. The sounds from the reflected image are oriented at a more favorable angle and therefore will be relatively more intense. The combined direct and reflected sounds will seem to arrive from a point intermediate between the source and the image. With the console placed in moderately live rectangular rooms, the sounds from the left and right channels will appear to arrive from near the corners. On the other hand, identical signals fed into both channels will result in a symmetrical pattern arriving to the listener simultaneously from both loudspeakers and their images, and the combined effect will be that of a central sound arriving from the front. With varying mixtures of central and side sounds, a "wall of sound" is created in front of the listeners.

Consider now the reception for an observer to one side of the room. As the observer moves toward the right, for example, the intensity from the right loudspeaker will be increased, and that from its image will be diminished. Consequently, the over-all intensity will not be greatly changed. Close to the boundary, the radiation from both the direct source and its image will become nearly equal, and the combined sound will appear to be arriving from between the two, *i.e.*, again from the right corner of the room. At the same time the radiation from the left-hand image will be increased in a compensatory manner. To a considerable extent, therefore, the projection console will tend to provide equally intense sound from both channels over a large area, this type of action having been identified with the performance of the "Isophonic"** stereophonic system described previously.[4]

Directional Operation

A method for obtaining directional operation of loudspeakers at wavelengths longer than its dimensions, is shown in Fig. 2. The loudspeakers are installed at one end of the cabinet, but instead of being completely enclosed, access is provided to the rear of the piston by means of a "tunnel" terminated by pads of "ozite" or similar material. Consider only that part of the system

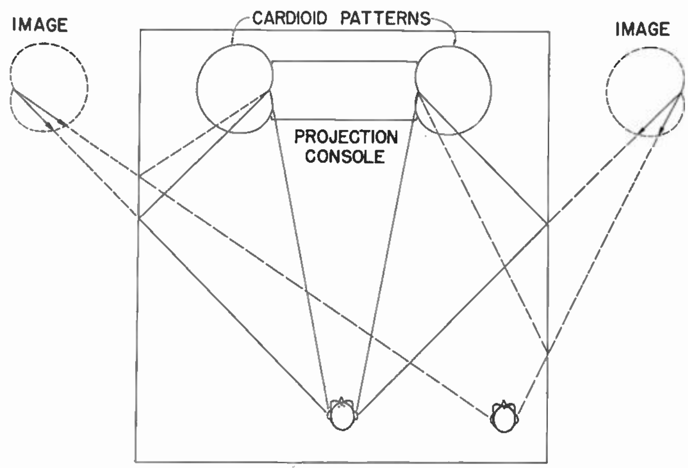


Fig. 1—The principle of operation of the projection console.

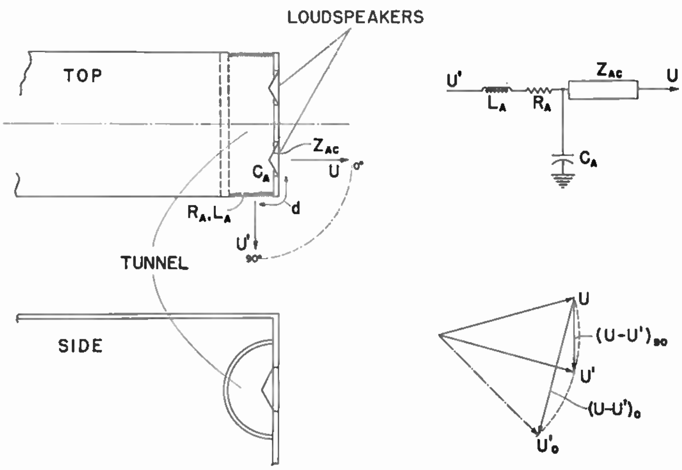


Fig. 2—Obtaining directional operation of loudspeakers at low frequency.

on one side of the line of symmetry; the loudspeaker presents an acoustical impedance Z_{AC} and it produces a volume current U at the front of the piston. The back current is modified by the tunnel and its termination. From the equivalent network point of view, the tunnel is an acoustical compliance C_A and the ozite pads form an acoustical resistance R_A and inertance L_A, these elements being arranged in a relationship shown at the upper right hand side of Fig. 2. With the proper choice of constants, as shown below, such a network can be designed to shift the phase of sound through an angle proportional to frequency without substantial change in its magnitude.[5] The combined effect at a remote point in space of the direct and the phase-shifted volume currents, U and U', can be analyzed in terms of two angles on the phasor diagram of Fig. 2, each of which is proportional to frequency: 1) the angle spanned by the arrows U and U' produced by the phase-shift network,

** Registered in the U. S. Patent Office.

[4] B. B. Bauer, "Broadening the area of stereophonic perception," *J. Audio Engrg. Soc.*, vol. 8; April, 1960.

[5] B. B. Bauer, "Uniphase unidirectional microphones," *J. Acoust. Soc. Amer.*, vol. 13, p. 41; 1941.

which is independent of the orientation of the listener; and 2) the angle spanned by the arrows U' and U_0' caused by the unequal distance between the sources of U and U' and the observer. This distance is designated as d, and in the case of the observer stationed towards the 0° incidence angle the full delay due to this distance comes into play, corresponding to the time d/c_v seconds where c_v is the velocity of sound. As the observer moves along the dotted line towards the direction designated as 90°, making an angle θ with the arrow 0°, the effective distance diminishes approximately as the cosine of this angle, and hence, in general, the delay will be (d/c_v) cos θ seconds, corresponding to a phase angle $\phi=(\omega d/c_v)$ cos θ radians. The directional pattern arising from this action may be ascertained as follows: let U represent the relative phase of the volume current at the front of the piston and U_0' represent the phase of U' for the observer at 0°. The combined volume current is the difference between U and U_0' portrayed by the phasor $(U-U_0')$. As the observer moves at an angle θ towards the 90° arrow and beyond, the phasor U_0' moves on its own arc toward U' and thence toward U following the cosine law. The length of the phasor $(U-U_0')$ will, therefore, follow the approximate generalized form $(k+\cos\theta)$ where k is that portion of the phase shift attributable to the action of the phase-shift network expressed as a fraction of that part attributable to the maximum delay caused by the distance d. This is the equation of a limaçon, but if $k=1$, then the equation portrays a cardioid, and if $k=0$, it defines a figure-eight pattern.

The analysis required to proportion the acoustical network so as to shift the phase of sound without substantially altering its magnitude, is shown in Fig. 3. Assuming that the radiation impedance at the port outlets is small enough to be lumped together with L_A and R_A, it is simple to show that the rear volume current U' is given by

$$U' = U/(1 - {}^2L_AC_A + j\omega C_AR_A). \tag{1}$$

What needs to be done is to equate the denominator of (1) to a unit phasor operating at an angle $\phi=k\omega d/c_v$. It is known that for moderate angles, $\sin\phi=\varphi\cdots$, and $\cos\phi=1-\phi^2/{}^2!\cdots$. Taking cognizance of the similarity between these expressions and the terms of equal coefficient in (1), we at once arrive to the solutions $\omega C_AR_A\approx k\omega d/c_v$, and, hence,

$$RA \approx kd/C_Ac_v \tag{2}$$

$$1-\omega^2L_AC_A \doteq 1 - k^2\omega^2d^2/c_v{}^22!,$$

and hence,

$$L_A \doteq k^2d^2/C_Ac_v{}^22! \tag{3}$$

For cardioid operation, k is simply taken as unity. It should be pointed out that theoretical determinations serve merely as a guide and some experimentation is needed to obtain the desired polar pattern. The network is operative up to the frequency where $d\approx\lambda/4$, where λ is the wavelength of sound. At higher frequency, the desired directional operation is obtained by proper cone configuration.

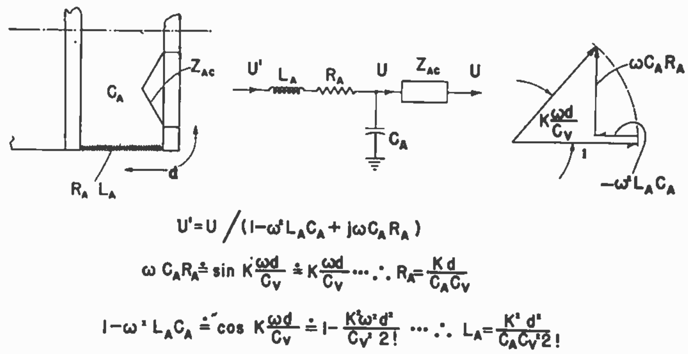


Fig. 3—Phase-shift network analysis.

Fig. 4 gives an example of the type of performance that has been obtained at 400 cps with a particular loudspeaker and tunnel configuration, using pad resistance R_A equal to 0.1 acoustical ohms (cgs system), which produces a cardioid pattern. L_A is principally defined by the radiation mass at the ports. With the pads removed, $(R_A\to 0)$, the polar pattern approximates the cosine function, deviating from it, however, because of dissymmetry between the front and back acoustical paths. However, even this type of pattern has been found to be rather effective in producing satisfactory directional effects in the projection console. On the other hand, if both ports are closed $(R_A=\infty)$, then the polar pattern approximates a circle, and the stereophonic performance becomes considerably less effective.

It has been shown previously[6] that stereophonic performance in average size rooms appears to be determined mainly by sound frequencies of 250 cps and above. Therefore, it becomes convenient to restrict the cardioid operation of the projection units down to approximately 250 cps. Frequencies below 250 cps may be reproduced by separate nondirectional loudspeaker units. The frequency response of a single channel in the completed unit is shown in Fig. 5. The curves at 0°, 90°, and 135° indicate that the directional properties are retained at all the frequencies of interest. Above 500 to 800 cps, the phase-shift network becomes inoperative and directivity is maintained because of diffraction. The low-frequency channel provides nondirectional operation, and while its response appears to be lower than that of the high frequency channel, actually this is not the case because the total radiation of the latter is subject to a random radiation coefficient of the cardioid pattern which is about 4.8 db.

[6] P. C. Goldmark and J. M. Hollywood, "Psychoacoustics applied to stereophonic reproduction systems," *J. Audio Engrg. Soc.*, vol. 7, pp. 72–74; April, 1959.

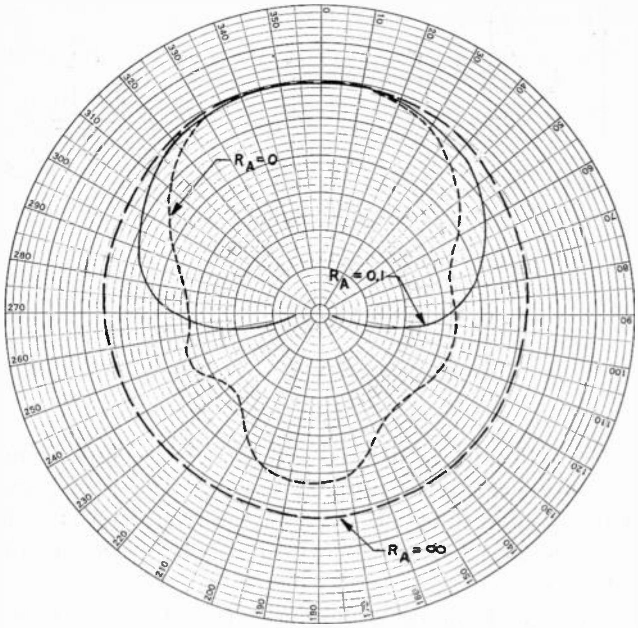


Fig. 4—Polar patterns obtained with $R_A=0.1$ acoustical ohms (cardioid); $R_A=0$ (cosine) and $R_A=\infty$ (circular) at 400 cps.

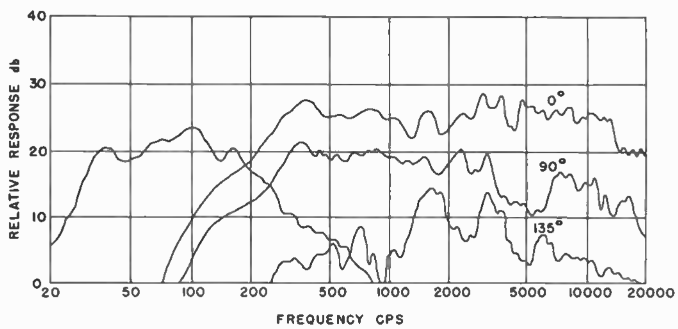


Fig. 5—Over-all frequency response of the system.

Listening Tests

A number of tests were performed to compare the performance of the projection console with a conventional two-unit loudspeaker system placed near the corners of the room. The schematic arrangement is shown in Fig. 6. The stereophonic effect created by the projection console, compared favorably with the two-unit system. For example, in playing a selection in which a trumpet solo instrument is played on the right channel, the direction sensed with the two-unit instrument was from the area 1. With the projection console, there was a slight shift to the right, namely to the location 1′. With a guitar playing at the left channel, appearing from the area 2, with the projection console, the direction appeared to shift somewhat to the right, to the area 2′. A central sound, such as a solo voice, appeared to arrive from the direction 3 and 3′ for both arrangements. Localization for intermediate points was quite good, many observers being unable to tell which of the two systems was in operation.

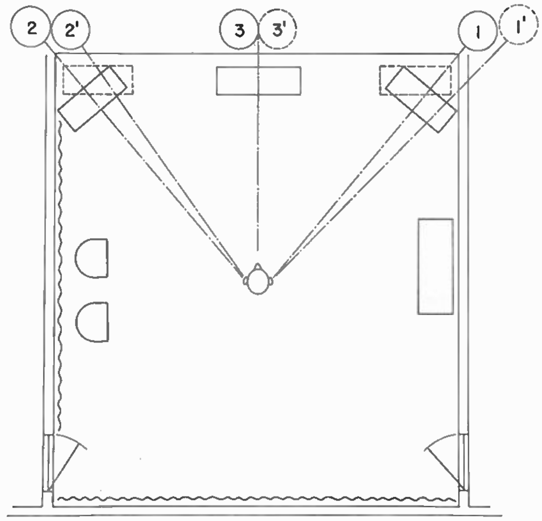


Fig. 6—Arrangement for comparative listening tests.

Acknowledgment

The authors wish to acknowledge with thanks the participation of Dr. Peter C. Goldmark, President and Director of Research of CBS Laboratories, and Mr. William S. Bachman, Director of Engineering of Columbia Records, who have helped to guide this development towards its final goal.

A Transistorized Stereo Preamplifier and Tone Control for Magnetic Cartridges*

ALEXANDER B. BERESKIN†

Summary—**RIAA equalization with ±4 db bass control and ±8 db treble control have been achieved, along with negligible hum, noise and distortion, in a transistorized stereo preamplifier developed for use with magnetic cartridges. Simple circuit modifications adapt this preamplifier for use with most magnetic cartridges.**

Introduction

THE stereo phono preamplifier and tone control described in this paper was developed for use with magnetic cartridges. Output specifications for these cartridges are usually in terms of the voltage developed across a relatively high load resistance. This type of specification is generally compatible with vacuum tube preamplifiers. In the case of transistorized preamplifiers the cartridge acts as a current source so that its electrical parameters are also of considerable importance.

A survey of several magnetic cartridges indicated that the GE GC7 and CL7 cartridges have relatively high resistance, and inductance values of 2200 ohms and 0.50 henries respectively. The Shure M3D and M7D cartridges have 330 ohms and 0.365 henries while the GE VR227 and VR225 cartridges have 1640 ohms and 0.460 henries. The basic design was made for GC7 and CL7 units having the highest resistance and inductance values and provisions were made for adapting the amplifier to cartridges with other values of resistance and inductance.

The standard RIAA playback characteristic is shown in Fig. 1. This characteristic is defined on the basis of circuits with time constants of 3180, 318, and 75 μsec and corresponds to corner frequencies of 50, 500, and 2120 cycles with basic slopes of −6 db per octave between 50 and 500 cycles, and above 2120 cycles. It can be seen from this characteristic that the required variation in response between 20 cycles and 12,000 cycles is 34 db.

Basic Amplifier

The basic amplifier that has been developed is shown in Fig. 2. This amplifier uses two low-noise RCA 2N220 transistors. RCA 2N175 transistors are electrically equivalent to the 2N220 and could have been used equally well. Frequency-sensitive collector-to-base feedback is used in both stages and a frequency-sensitive interstage network is used to provide tone control. The input and interstage coupling capacitors are also carefully proportioned to control the low-frequency response.

The base-to-ground resistors are selected to make the dc collector-to-ground voltage approximately one-half of the collector supply voltage.

The frequency-sensitive feedback in the collector-to-base circuits reduces the gain with increasing frequency but its effect is constant for frequencies in excess of 1 kc. These feedback circuits provide greater low-frequency gain variation than that required by the RIAA characteristic, so bass tone control is obtained by counteracting this trend in gain with the interstage 100-K bass control potentiometer and 0.05-μf capacitor. Maximum bass response results when the bass control potentiometer and capacitor are shorted.

Treble control is obtained in the interstage circuit by bypassing portions of a 10-K potentiometer with a 0.015-μf capacitor. Maximum treble response is obtained when the 10-K treble control potentiometer is completely bypassed.

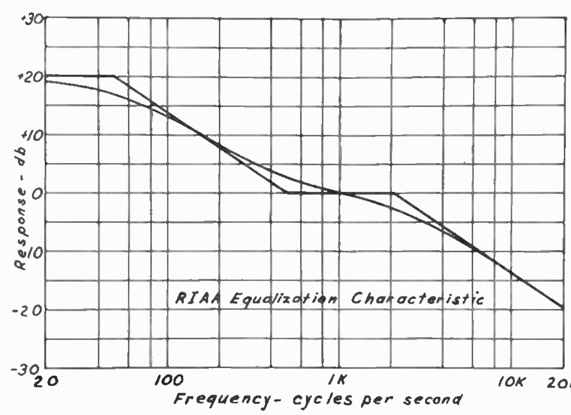


Fig. 1.

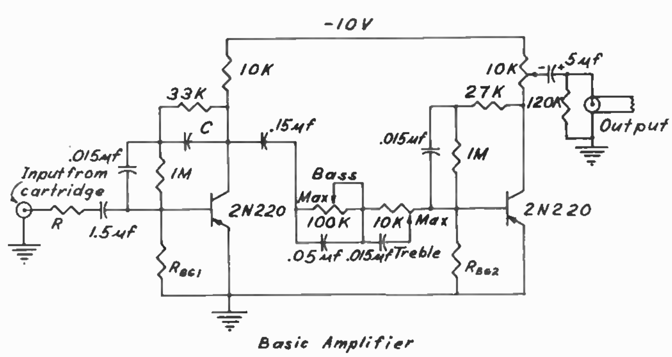


Fig. 2—Basic amplifier.

* Manuscript received by the PGA, October 9, 1959. Presented at the Natl. Electronics Conf., Chicago, Ill., October 12, 1959.
† University of Cincinnati, Cincinnati, Ohio.

An analytical approach to the determination of the compensation elements would require the simultaneous consideration of the cartridge, the feedback circuits, the interstage elements, and the two transistors each of which may be represented rather accurately by the Hybrid π equivalent circuit diagram. The complexity would be reduced to a certain degree by the fact that the low-frequency compensation is negligible at high frequencies and vice-versa. The analytical approach is still a formidable one, and therefore the actual approach used was to estimate the required values and then to use the experimental approach to determine their exact values. It was determined, however, from the differential equations of the cartridge and input circuit, assuming that the transistor input circuit was a 3900-pf capacitor in parallel with a 3380-ohm resistor, that the system was slightly overdamped and should therefore be free of electrical oscillations.

The experimental setup consisted of a 6-ohm source of signal connected between the cartridge common terminal and ground. The free ends of the cartridge windings were connected to the input circuits of the transistor amplifiers. This method was sufficiently accurate for all preliminary measurements but final adjustments were made on the basis of the London Stereophonic Frequency Test Record (PS131) which is guaranteed by the manufacturer to be within $\pm\frac{1}{2}$ db of the RIAA characteristic. The test record covers only the 40-cycle to 12,000-cycle range but the experimental setup is satisfactory for investigating frequency response in both directions beyond this range. The experimental setup also permits the measurement of voltage amplification and distortion.

The component values determined by this method are shown in Fig. 2. Both R and C are zero for the GC7 and CL7 cartridges. Proper RIAA equalization resulted when the unshorted portion of the bass control potentiometer was 50 K and the unbypassed portion of the treble control potentiometer was 3 K. The actual response obtained is shown by the middle curve in Fig. 3. All points on this curve are contained within a 1.6-db range. The relative effects of imperfect equalization and imperfect recording levels were not assessed but it is clear that even perfect equalization could result in a 1-db variation in the response characteristic. The maximum bass-maximum treble and minimum bass-minimum treble curves show that the tone controls permitted roughly $\pm$4-db bass and $\pm$8-db treble control.

The effect of frequencies below 40 cycles and above 12 kc could not be assessed with the test record so the signal generator setup was used for this purpose. The curve in Fig. 4 was obtained for the RIAA setting of the tone controls. This curve confirms the lack of electrically caused oscillation and shows that, neglecting possible mechanical resonances, there is no sudden deviation from the RIAA characteristic below 40 cycles and above 12 kc.

Alternate Cartridges

When the GE cartridge was replaced with the Shure M3D cartridge without changing the tone control settings, and both R and C were zero, the upper curve in Fig. 5. was obtained. Serious response deviations are evident in this curve. For $R=1.5$ K and $C=0$ the next lower curve resulted. The curve for $R=2.2$ K and $C=0$ has good low- and middle-frequency response but, due to the reduced value of cartridge inductance, it has excessive response at the high frequencies. Increasing C to 680 pf resulted in the fourth curve from the top while

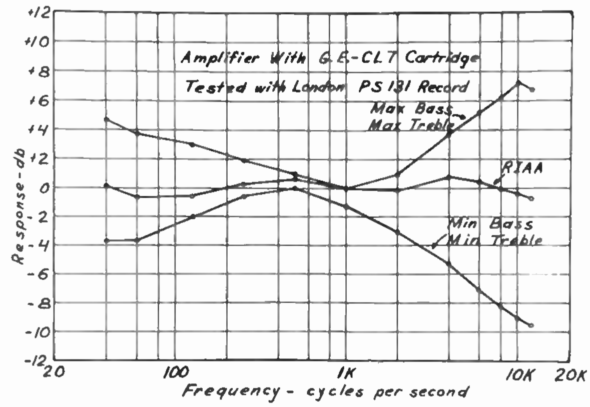


Fig. 3.

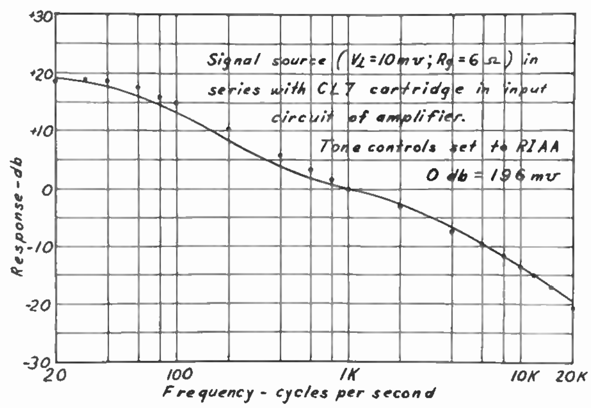


Fig. 4.

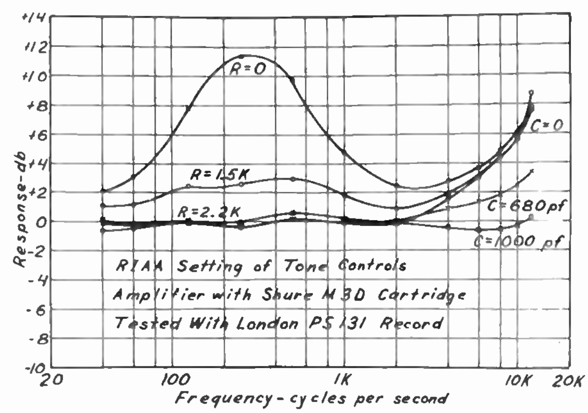


Fig. 5.

increasing C to 1000 pf resulted in the lowest curve. On this last response characteristic all points fall within a 1.2-db range. The net output is now the same as it was for the GE cartridge and is about 35 mv for an rms stylus velocity of 1 cm/second at 1 kc. Normal stereophonic record material results in an output voltage, measured with an RCA WV-74A High Sensitivity AC VTVM, which fluctuates between 50 and 100 mv and which occasionally reaches 300 mv.

Similar results were obtained with a GE VR227 cartridge for $R=560$ ohms and $C=680$ pf.

Noise and Distortion

The residual random noise in the output, measured with a 15-kc (white noise) band-pass filter (−3 db at 25 cycles and 13 kc) and an rms responding VTVM was 170 μv for one channel and 110 μv for the other channel. One transistor yielding noise values of 350 μv was discarded. These values are more than 60 db below the 1-kc signal output with a stylus velocity of 5 cm/second rms. An appreciable portion of this noise was judged to be low-frequency ($1/f$) noise. This was confirmed when the band-pass filter was replaced by one complementing the Fletcher-Munson equal-loudness contour (20 db above threshold). This filter has a white-noise bandwidth of 10 kc referred to the 1-kc response. In this case the measured random noise signals were respectively 54 and 49 μv referred to the 1-kc response of the system. These values are more than 70 db below the 1-kc signal output with a stylus velocity of 5 cm/second rms.

With a well-shielded cartridge there was no measurable hum at power frequency or at harmonics of the power frequency. The amount of hum output during regular use would therefore depend entirely on the cartridge shielding and connections and to the stray fields present in the neighborhood of the cartridge.

The distortion was measured over the 30-cycle to 10-kc range with a HP Model 330B Distortion Analyzer. Since this distortion analyzer is built to operate with a minimum input signal of 1 volt it was necessary to operate the amplifier with considerably higher input and output voltages than would ever be encountered in practice. In order to interpret correctly the extremely low measured values of distortion, the residual distortion in the signal source was also measured and both curves are shown in Fig. 6. It is interesting to note that over most of the range the output distortion was lower than the input distortion. This is not as unreasonable as it appears to be at first consideration. The harmonics contained in the input signal are amplified, except at the very low frequencies, by a minimum of 3 db less than the fundamental. If the amplifier itself does not introduce any new distortion it can be expected that the output distortion will be less than the input distortion. For the 30-cycle signal, the second harmonic would be amplified by $1\frac{1}{2}$ db less than the fundamental so that there should not be too great a difference between the input and output distortions. For the 100-cycle signal, the second harmonic is amplified 4.5 db less than the fundamental and the output distortion is about 4 db less than the input distortion. For a 1-kc signal, the second harmonic is amplified 3 db less than the fundamental and the output distortion is exactly 3 db less than the input distortion. It is only for signal frequencies above 5 kc that the output distortion distinctly rises above the input distortion but a 1-volt output at these frequencies is certainly unreasonable. In any case distortion is definitely not a problem in this amplifier. The residual random noise, indicated by the dashed curve, would be insignificant even at one tenth the output signal.

Complete Amplifier

The complete circuit diagram is shown in Fig. 7. The power supply used is an unconventional one developed by the author some time ago for use with low-level transistor amplifiers. A particularly interesting property

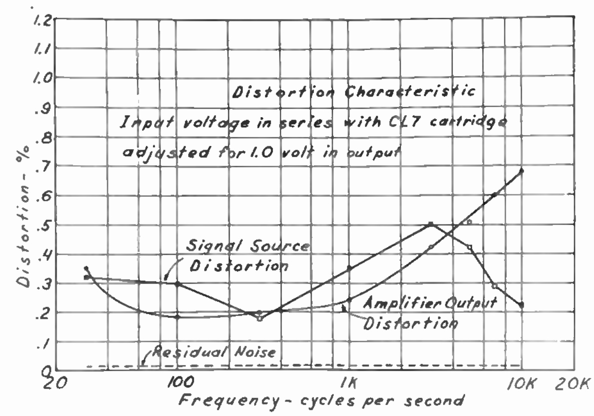


Fig. 6.

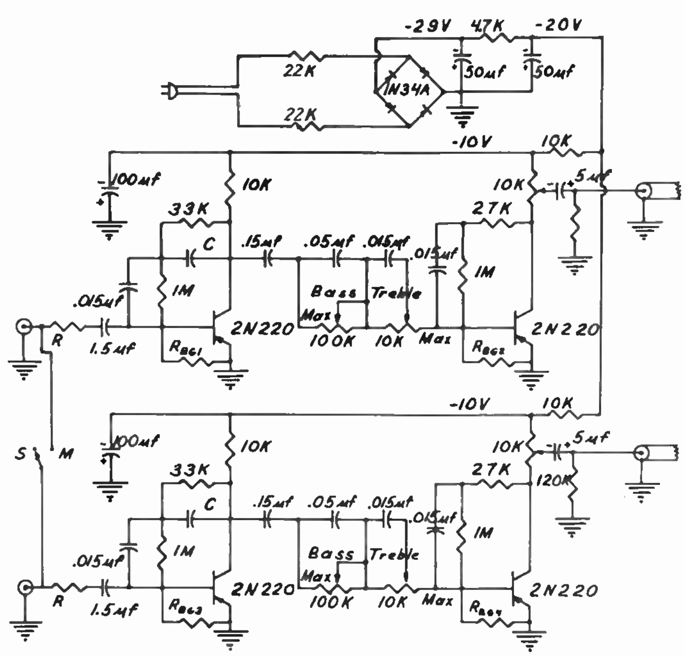


Fig. 7—Complete amplifier.

of this power supply is that either ac line can be shorted to ground without changing the dc output voltage. Other more conventional power supplies could be used equally well.

Two of the base-ground resistors were omitted altogether and the other two were 33 K and 39 K, but the bias condition would not have been seriously affected if they had been omitted also.

A switch has been provided for the paralleling of the two inputs when monophonic records are used. Each channel still has its own tone and volume controls.

Conclusion

Very satisfactory results have been obtained with this amplifier in practice. In addition to providing RIAA equalization and volume control, the amplifier supplies ±4 db bass control and ±8 db treble control. In cases where more output is needed to drive a power amplifier, an additional stage with flat response can be added. The components R and C must be correctly chosen for the specific cartridge used. These components may be soldered into the circuit or built into a selector switch.

Bandwidth Compression by Means of Vocoders*

FRANK H. SLAYMAKER†

Summary—**Speech information may be transmitted over a bandwidth one tenth of that required for the original speech if attention is directed toward reproducing the envelope of the power density spectrum rather than the waveform itself. Pitch information can be transmitted independently of the spectrum information and the two sets of signals combined at the receiving end to resynthesize the original speech sounds. In the vocoder the power spectrum is analyzed and synthesized by means of band-pass filters. The energy for the voiced sounds is obtained from an oscillator called a buzz source, and for the fricative consonants the energy is obtained from a white noise source.**

Introduction

In discussing bandwidth compression by means of vocoders, let us look first at the situation which makes it desirable to compress the bandwidth occupied by speech. The compression of the speech bandwidth is totally unlike the process of cutting down on radio bandwidth by means of single sideband modulation; it is also unlike the compression of the television bandwidth by making use of the similarities between successive frames in the television picture. Bandwidth compression in speech is possible primarily because the acoustic wave as emitted by the mouth is more complex than the actual information required to control the speech mechanism.

Information Rates

Table I shows the channel capacity necessary for the transmission of English words. The assumption made in this table is that the channel necessary to handle the original speech wave had a bandwidth of 3000 cycles and a signal-to-noise ratio of 30 db. The maximum channel capacity in bits per second according to Shannon[1] can be expressed as $W \log_2 (1+P/N)$ where W is the bandwidth in cps and P/N is the ratio of the signal power to the noise power. It follows then that the channel capacity (of a 3000-cps, 30-db channel) is 30,000 bits per second. The same bit rate, but a much larger bandwidth, is required if the original speech wave is quantized to the maximum number of levels consistent with the 30-db signal-to-noise ratio; the wave is sampled at a rate double that of the highest frequency appearing in the speech wave, and the level information is sent by a binary-pulse-code scheme. Such a system or channel, however, can transmit not only speech information but all the other sounds which can be handled in a 3000-cycle, 30-db channel. These sounds might be music, sound effects, or even telemetering data.

TABLE I

Comparative Channel Capacities Necessary for Speech Transmission

Method of Coding	Necessary Channel Capacity
Digitized Speech Wave Form	30,000 bits per second
Phoneme Coded Speech	60 bits per second
Word Coded Speech (at 120 words per minute)	
Vocabulary of 2 words	2 bits per second
Vocabulary of 8000 words	26 bits per second
Vocoder	2000 bits per second
Teletype (at 120 words per minute)	75 bits per second

* Manuscript received by the PGA, September 23, 1959. This paper was presented at the 1959 IRE National Convention.

† Stromberg-Carlson Co., Rochester, N. Y.

[1] C. E. Shannon and W. Weaver, "The Mathematical Theory of Communication," University of Illinois Press, Urbana, p. 113; 1949.

There are, however, only a limited number of sounds that can be differentiated in English speech. These individual sounds, called phonemes by the phoneticians, form the essential building blocks of our spoken speech. In English there are approximately 40 of these sounds; the precise number depends upon who makes the list and how many diphthongs are given individual identities. A six-bit code will serve to identify 64 different phonemes. Then, if we assume that we can talk at a rate of about 10 phonemes per second, it would take only 60 bits per second to transmit the phonemic information in our speech. Even so, such a system will permit the sending of phonemic combinations which do not exist in the English language. If we code the English words individually, rather than the basic speech sounds, we find even lower rates of information transmission. The bit rate required is a function of the size of the vocabulary from which we choose the words. In other words a child saying a given sentence is transmitting less information than an adult, who may have a vocabulary of several thousand words, saying the same sentence.

Assuming for the moment that a good average talking rate is 120 words per minute and that our vocabulary is limited to two words, a limited conversation such as "Yes, No, Yes, No, No, No, Yes," and so forth requires only two bits per second to carry the information. On the other hand, to select two words per second from a vocabulary of 8000 words requires 26 bits per second. Eight thousand words is larger than the average vocabulary of "Johnny Q. Public" but falls short of that of some of our more advanced scholars. The 2000-bits-per-second rate for a vocoder in Table I is a mid-value figure. Vocoders can be built to use a lower bit rate at some sacrifice in the quality and intelligibility of the signal, and they can also be built requiring higher rates.

To send information by teletypewriter we find that, at a rate of 120 words per minute which is about twice the normal rate for a teletypewriter, we will need 75 bits per second.

Notice the extreme discrepancy between the 30,000-bits-per-second and the 26-bits-per-second rate required for selecting words from a vocabulary of 8000 words. It is no wonder that as our communication channels get more and more crowded one looks with considerable skepticism at the use of a 30,000-bits-per-second channel to carry information which can be sent through a 26-bits-per-second channel. The problem now becomes two-fold: 1) how to extract the information from the speech wave, and 2) how to transmit the information over a reduced bandwidth but still retain the ability for a person to talk into a microphone and listen to the information reproduced as normal speech.

Identification of Speech Information

Fig. 1 shows two oscillograms of the sound "ē," as in "see," spoken by two different individuals. It is quite apparent that the two oscillograms are so dissimilar that there seems to be very little hope in using the curve of sound-pressure vs time as the source of the information. Fig. 2 shows the same sound of "ē" spoken by the same two speakers plotted in a different fashion. In Fig. 2 the intensity is plotted with respect to frequency, and it is apparent that there are peaks or resonances occurring at practically the same frequencies for both speakers. The first peak occurs at about 300 cycles, the second

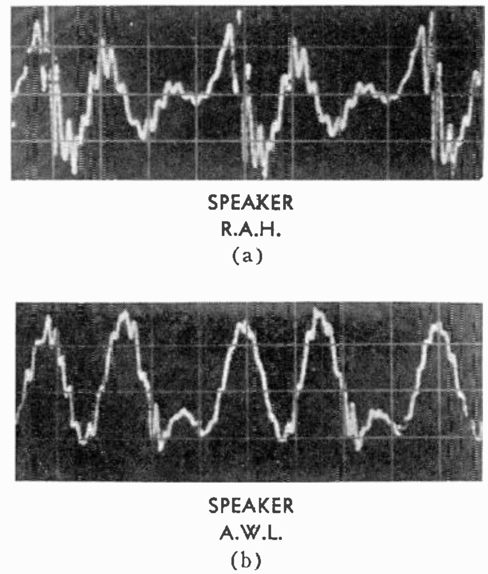


Fig. 1—Oscillogram of the sound "ē" as in "see" spoken by two different speakers.

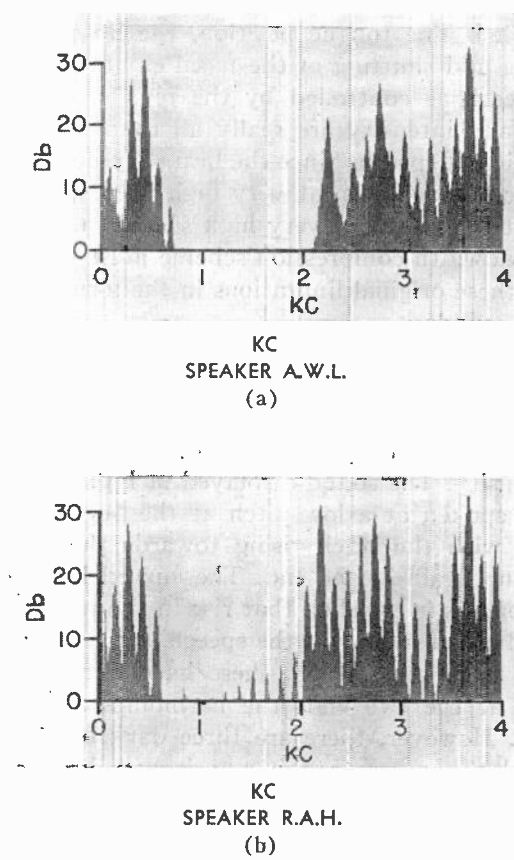


Fig. 2—Sound spectrograph sections of the sound "ē" as in "see" spoken by two different speakers.

one at about 2200, the third at about 2800, and the fourth of about 3600 cycles. These resonances, called formants by the phoneticians, occur in characteristic positions for the different speech sounds. Information in speech is carried in the gross envelope of the power density spectrum rather than in the sound-pressure vs time function. Notice that even though the rough envelope shape of these two spectra are the same, the individual harmonics making up the spectra are different and occur in different places. In fact, almost any sound from noise to music can be substituted for the original series of harmonics and the result will be intelligible as long as the envelope of the power density spectrum has the same shape.

Fig. 3 shows a cross section of the vocal tract. Air coming up the windpipe passes through the vocal folds, or vocal cords, in the larynx and produces a series of pulses which, if analyzed as a Fourier series, result in a fundamental and a series of harmonics. This complex nonsinusoidal sound is modified by passing through the vocal tract, past the tongue and the lip openings, and also (for some sounds), through the nasal cavities which can be shut off by the velum. The vocal tract acts as a complicated filter that emphasizes some of the harmonics in the original output of the vocal folds and supresses others. Other sounds, such as "sh," "s," "f," "t," and "p," do not require the vocal folds to be excited but are generated by turbulences in the vocal tract. Speech, then, is produced by a fairly simple structure in which there are very few parameters that need to be controlled. The tongue position, the lip position, the opening and shutting of the nasal cavity, the pitch of the speech as controlled by the tension of the vocal cords, and intensity are really all that enter into the formation of speech. Since the brain can neither control the vocal mechanism at very high speeds, nor can we understand speech at very high speeds, it follows that our bandwidth compression scheme must take into account these original limitations in the formation of the speech sounds.

Fig. 4 is a sound spectrogram showing that the positions of the resonances or formants are independent of the voice pitch. The spectrogram is a plot of frequency vs time where the darkness of the trace is an indication of intensity. The sound displayed in Fig. 4 is "ĕ" as in "met" spoken at a low pitch at the beginning of the figure, with the pitch rising towards the center and dropping again at the end. The curved lines, running more or less in parallel, that rise in the middle are the individual harmonics of the speech wave. In the center where the pitch is high, these harmonics are further apart; at the two ends the harmonics are closer together. However, there are three dark bands running through the spectrogram in a horizontal direction which indicate the increased intensity of the sound output at the frequencies corresponding to the formants of the vocal tract. These three formants remain fixed in frequency regardless of the pitch of the original voice. Fig. 5 is a spectrogram showing how these formants or resonances move when words are spoken. In this figure, the words "one," "two," and "three" are shown as spoken by two different speakers. Notice the similarity between the two sets of traces; this indicates how the filtering mechanism, which our vocal tract might be called, changes its characteristics to form the words. At the beginning of the "two" there is a burst of high-frequency energy starting very sharply, which is the "t" sound, followed by the formation of the broad bands of resonances at the time that the vocal folds are excited.

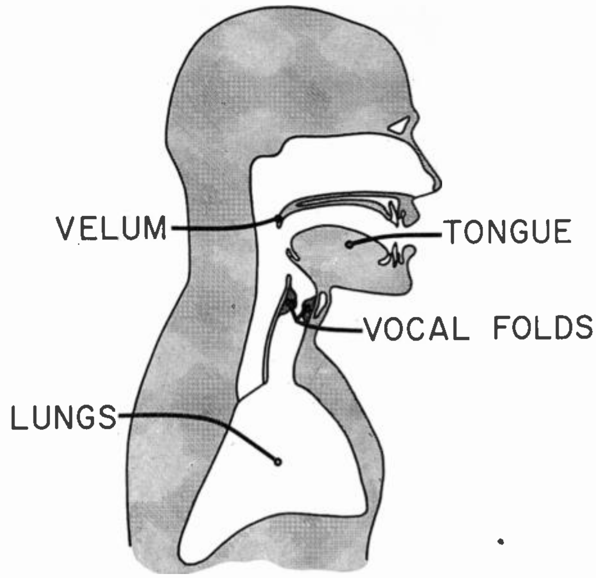


Fig. 3—Cross section of the vocal tract.

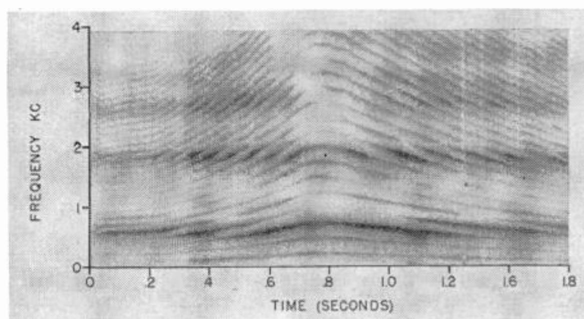


Fig. 4—Sound Spectrogram of the sound "ĕ" as in "met" spoken with changing pitch.

The Vocoder

The vocoder, which was invented by Homer Dudley about 20 years ago, accomplishes with electrical filters, nonsinusoidal oscillators, and noise sources, the same sort of thing that occurs in our vocal tract. Since the speech information is contained in the power spectrum, the first tasks for the vocoder are to analyze the original

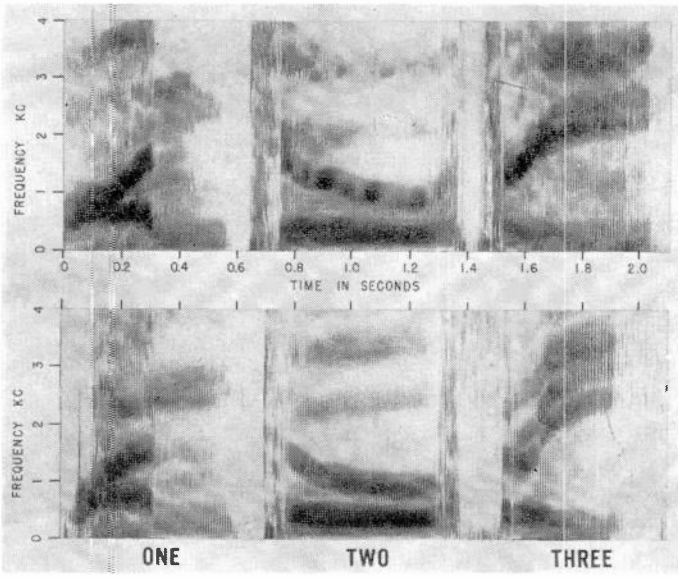


Fig. 5—Sound spectrograms of the words "one," "two," "three," as spoken by two different speakers.

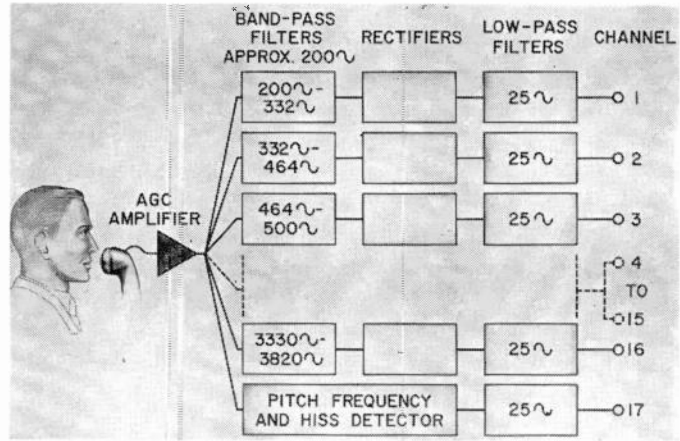


Fig. 6—Block diagram of a vocoder analyzer.

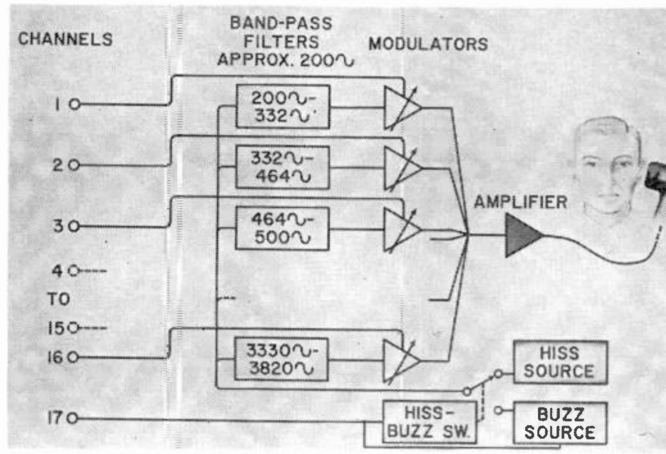


Fig. 7—Block diagram of a vocoder synthesizer.

speech sounds and to obtain a measure of the shape of the power spectrum. Fig. 6 shows a series of band-pass filters which are contiguous and cover the entire speech band. The outputs of these filters are rectified and passed through low-pass filters to an output which gives an analog measure of the energy in each of the pass bands in question. There is also a circuit which measures the pitch of the original voice and indicates whether the sound was voiced, *i.e.*, whether the vocal folds were vibrating, or whether the vocal tract was excited merely by turbulence. The analog signals from each of these channels are now passed to the vocoder synthesizer (Fig. 7), and used to modulate output channels having exactly the same frequency ranges as the original band-pass filters in the analyzer. The synthesizer is in reality a talking machine containing a very nonsinusoidal oscillator, called the buzz source, that serves the same function as the vibrating vocal folds in the throat. The synthesizer also contains a source of thermal noise called the hiss source that produces the sounds generated by air turbulance in the vocal tract. There is a switch that changes back and forth from the buzz source to the hiss source, depending upon whether the analyzer indicates voiced sounds or not. The output of whichever source is chosen is fed into the band-pass filters of the synthesizer; the outputs of the synthesizer filters, as modulated by the controlling signals, are combined to form speech once more. The filters in parallel modify the sounds from the buzz and hiss sources in essentially the same way as did the vocal tract.

In any practical communications scheme, the controlling signals for the vocoder must be transmitted on one channel rather than by 17 different channels, as indicated in Figs. 6 and 7. Some form of multiplexing is required. Fig. 8 shows a time-division multiplex scheme in which some form of a scanning switch, either mechanical or electrical, scans the output of the vocoder channels and sends the information in sequence along the transmission link. At the receiving end the equipment demultiplexes the signals and transmits it to the synthesizer. Fig. 9 shows another form of multiplex that is similar to a carrier telephony system, in which individual oscillators, each representing one of the channels of the vocoder, have the appropriate information applied to the output of the oscillator by some modulation system. The demultiplexing equipment here would consist of a series of filters to pick out the required channel information. The pitch extractor circuitry in the vocoder, or any speech-compression system for that matter, can be represented by the block diagram of Fig. 10. In this diagram the fundamental frequency is emphasized by means of a filter. If the fundamental is not present in the original input material, it must be regenerated by combining the harmonics. The signal which now has a very heavily emphasized fundamental is clipped, differentiated, and rectified. The rectified pulses are then used to trigger a single-shot multivibrator, and the output of the multivibrator is smoothed or integrated. The output from the integrator is an analog signal in which the amplitude is a function of the original fundamental frequency of the speech signal.

Fig. 11 shows two basic ways of emphasizing the fundamental frequency for application to the pitch measuring circuits. The Stromberg-Carlson Company has developed vocoders using both of these methods. The first one is a fairly simple method in which a filter is chosen to select the band which is most likely to contain the fundamental for a given speaker. Two different filters are usually used, one for female voices and one for male voices. If the female-voice filter is used for a male speaker, the equipment is likely to select the second harmonic of the male voice as the fundamental, and the vocoder will speak an octave higher than the original speaker. To avoid this operational difficulty, another approach is to use a variable-frequency band-pass filter which starts from rest at a low-frequency position and moves upward until it reaches the peak response at the lowest frequency being radiated by the voice. Such a system avoids the necessity of having a male-female switch on the analyzing equipment. With either type of fundamental-frequency filter the pitch indication is obtained in the same way.

Since the vocoder involves modulation devices that have a limited dynamic range, it is vital that the speech be fed into the vocoder analyzer at a level which is consistent with the normal operating levels of the equipment. A skillful human operator, such as the operator at the controls of a broadcast transmitter, could do the job, but usually a voice-operated gain-controlling circuit, such as that shown in Fig. 12, is included in the vocoder. This gain-controlling circuit embodies a conventional automatic gain control which reduces the gain when the output becomes large, and also has a gain-increasing circuit which increases the gain when sounds that pulse at the speech rate are introduced into the system. The device is quite empirical in

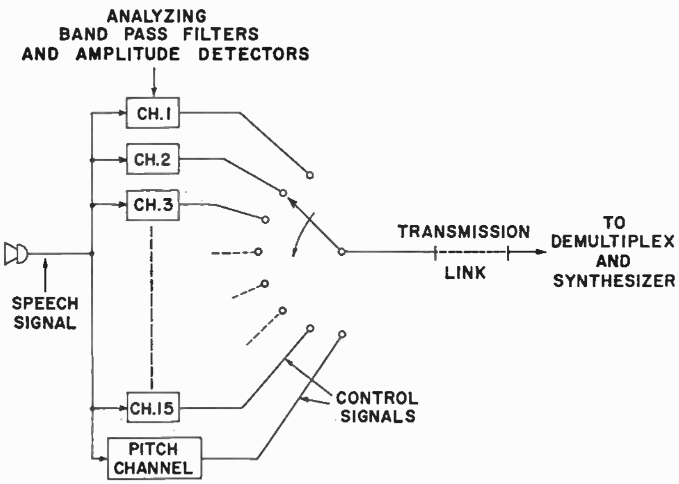


Fig. 8—Time-division multiplex applied to the transmission of the vocoder output signal.

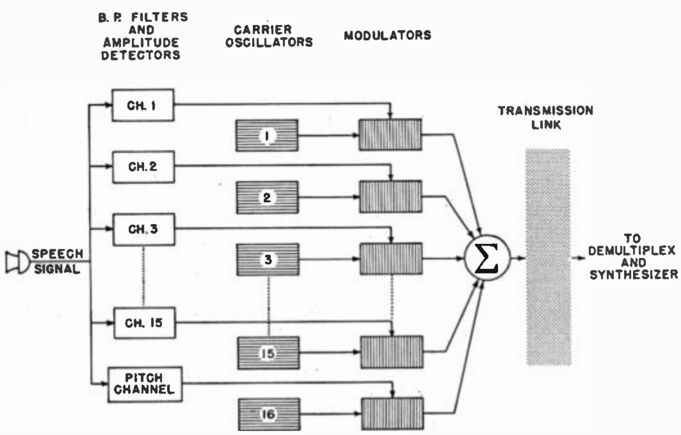


Fig. 9—Frequency multiplex applied to the transmission of the vocoder output signal.

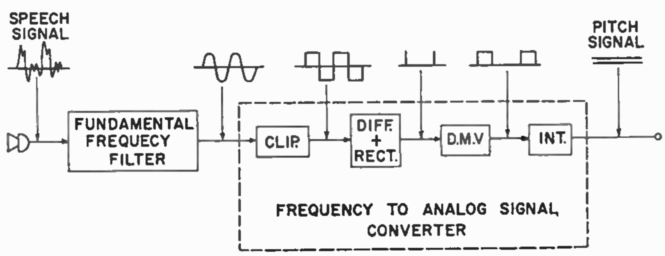


Fig. 10—Pitch extractor.

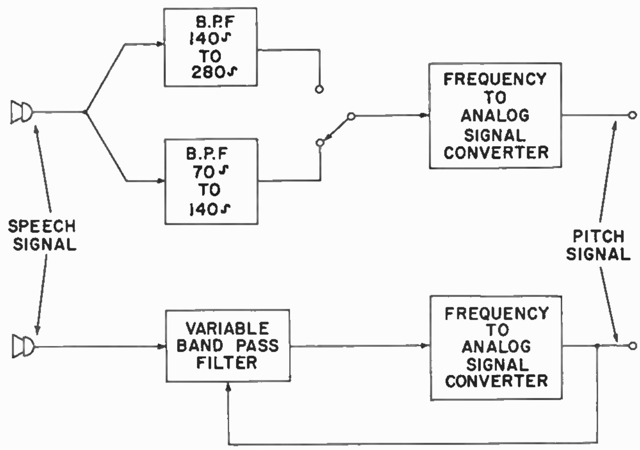


Fig. 11—Two fundamental-frequency filters.

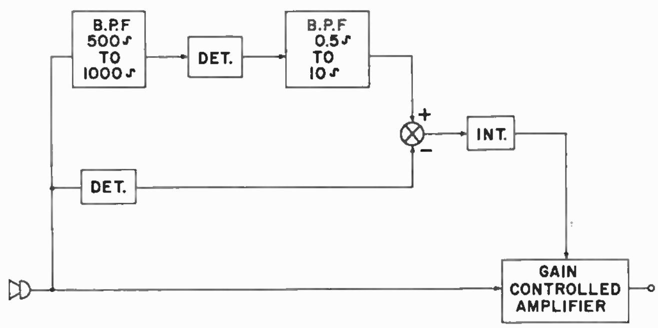
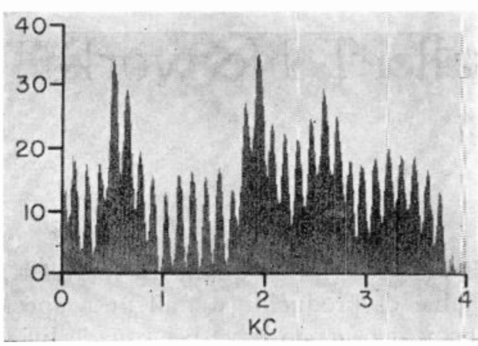


Fig. 12—Voice-operated gain controlling circuit.

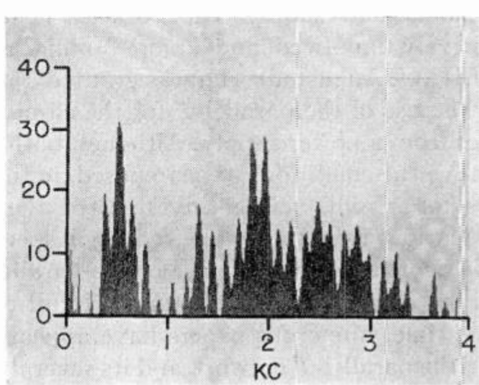
its operation and design, and although it can increase its gain very rapidly when a person starts to talk into it, once the gain level has been established the time constants are quite long and the gain does not change for long periods.

The next question one might ask is: "How well does the vocoder work?" Fig. 13 shows two frequency spectra, one made with the vocoder, and the other with straight unvocoded speech. The sound is "ĕ" as in the word "shed." Notice that the main formants occur at the same place, at about 600, 1800, and 2600 cycles, and that the sound as resynthesized on the vocoder is quite intelligible. The extraneous lower-amplitude peak at about 1200 cycles on the vocoded speech might be caused by a channel that has its gain set too high. The extra peak might also be caused by a transition between the situation that can occur at low frequencies where a single harmonic falls into each channel of the vocoder and the situation at some higher frequency where two harmonics appear in each channel. In the latter situation, the analyzer indicates more energy in the channel transmitting the two harmonics, and the gain of the corresponding synthesizer channel is, of course, increased. This transistion point shifts up and down in frequency as the voice pitch changes; because the ear interprets continuous speech rather than instantaneous spectra, the extra formant does no real harm. If the extra peak is caused by a channel with its gain adjusted too high during alignment procedures, the result would be similar to the sound of speech reproduced by a loudspeaker having a peak in its frequency response characteristic. The ear is very well conditioned to ignoring such variations in frequency response. Fig. 14 shows the comparison between the spectrograms for the words "shed light" as voiced by the original speaker and by the vocoder. Notice that at the beginning there is a burst of sound occurring above about 1700 cycles, which gives the "sh" sound; this is followed by the formants characteristic of the sound "ĕ," occurring at about 500 cycles and about 1800 cycles, which show the same positions in both spectrograms. The "d" sound is characterized by the sudden cessation of the voiced sounds; and unless one says "da" or "duh," the "d" does not have a characteristic sound of its own. The "l" sound of light is rather low in intensity and is followed by high-intensity formants indicated in the "ahee" sound. Notice that in both the vocoded and unvocoded speech the formants move in the same direction and occupy the same frequency positions. There is a silence at the beginning of the "t" sound followed by a burst of high-frequency energy. In other words, as represented pictorially by the spectrogram, the two sounds would be read as "shed light" by anyone who could read a spectrogram. Audibly the two sound like the same words, and the speaker's voice can also be recognized.

In addition to the actual resynthesis of speech by the vocoder, which is useful for communication and bandwidth compression, the vocoder permits several interesting demonstrations which are both entertaining and are also a help in understanding what elements in the speech convey the types of information of interest. For instance, by stopping the sending pitch information,



KC
UNVOCODED
(a)



KC
VOCODED
(b)

Fig. 13—Comparison of sound spectra of the sound "ĕ" as in "shed"; one of normal speech and the other after passing the sound through a vocoder.

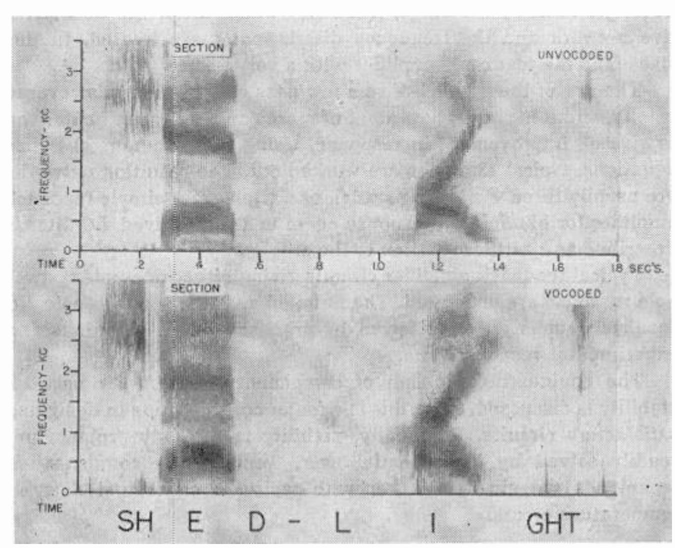


Fig. 14—Comparison of two sound spectrograms of the words "shed light"; one of normal speech and the other after passing the sound through a vocoder.

the vocoder can be made to talk always in a monotone. As far as articulation goes, it is possible to understand the words just as well with or without the pitch information. The type of information considered in Table I would not be affected. In other words the phonemes and the words can still be recognized. However, much of the emotional content of the speech is gone, and the ability to recognize the individual speaker is also considerably diminished, if not lost completely. It is also possible to illustrate very forcefully by means of the vocoder the premise that it is the gross envelope of the power density spectrum that conveys the speech information rather than the sound pressure as a function of time. This situation can be demonstrated by substituting music for the buzz source in the vocoder synthesizer. The result is what would happen if a human had a pipe organ or a symphony orchestra in his throat instead of the vocal folds. The sound produced is that of a pipe organ or orchestra talking, if one can imagine such a thing.

Such demonstrations do show that intelligibility, that is, the transmission of speech information, and naturalness are in many ways independent. Vocoders can be built with the pitch changes and individual voice inflections reproduced so accurately that individual voices can be recognized and identified. By transmitting less pitch information, naturalness suffers, but not the intelligibility. By using fewer channels, and thus transmitting less spectrum information, both naturalness and intelligibility suffer. The bandwidth compression is achieved by transmitting only the information that is necessary for reconstruction of the speech sounds by the receiving equipment. This information is carried in the spectrum shape, fundamental pitch, and intensity.

Bibliography

[1] R. J. Halsey and J. Swaffield, "Analysis-synthesis telephony, with special reference to the vocoder," *J. IEE (London)*, pt. 3, vol. 95, pp. 391–410; September, 1948.
[2] H. Dudley, "Remaking speech," *J. Acous. Soc. Amer.*, vol. 11, pp. 169–177; October, 1939.
[3] H. Dudley, "The carrier nature of speech," *Bell Syst. Tech. J*, vol. 19, pp. 495–515; October, 1940.

Design and Use of *RC* Parallel-*T* Networks*

GIFFORD WHITE†

Summary—The *RC* parallel-*T* network with a transmission null at f_0 is described and the symmetrical lattice approach to its analysis is outlined, following a notation of Guillemin. The selection of design parameters for various principal applications, with relevant references to published work, is given.

The common applications requiring a response curve symmetrical about f_0, such as the single-frequency notch filter, the ac derivative network and the frequency discriminator are treated. In this class falls the feedback amplifier with a very narrow notch.

The use of the parallel-*T* as a low pass or a high pass is covered briefly, and it is shown how a net of more complexity can be derived to give an improvement in response. Using the symmetrical lattice equations, typical examples are worked out. The resulting networks are usually three *T* nets in parallel, or a triple-*T*. A simple feedback amplifier for obtaining a response equal to an *m*-derived *LC* filter is described as a further solution to the problem.

Typical feedback amplifier circuits giving either one-pole or two-pole response are presented. The detailed analysis of a two-pole *RC* feedback net is given, followed by practical design equations and experimental response data.

The engineering problem of component selection for network stability is discussed, since this is a major consideration in designing satisfactory circuits. Frequently, stability is the only problem not readily solved by the potential user. Temperature compensation techniques are given, together with typical experimental data on temperature errors.

* Manuscript received by the PGA, September 29, 1959. This paper was presented at the NEC, Chicago, Ill., October 12, 1959.
† White Instrument Labs., Austin, Tex.

I. Introduction

THE first description of the *RC* parallel-*T* circuit capable of producing a null at a specified frequency appears to have been made by B. D. H. Tellegen[1] in 1937 in connection with the problem of radio interference elimination. He also disclosed several *RLC* circuits of varied configuration, all capable of producing the familiar notch response. The work by Tellegen has been relatively unnoticed, and a more familiar reference is that of H. W. Augustadt who was granted a patent in 1938[2] on the use of the parallel-*T* for the elimination of harmonics from a power supply. Although both of these patents gave the conditions to be imposed on the circuit constants for a null, neither investigated the full response equations over the whole frequency spectrum.

In 1938 Scott described the use of the parallel-*T* in a feedback amplifier[3] to produce a tuned-circuit response. Since that time, numerous papers have appeared, each analyzing the parallel-*T* network and its several applica-

[1] B. D. H. Tellegen, "Star and Delta Connection of Impedances," U. S. Patent No. 2,093,665; September 21, 1937.
[2] H. W. Augustadt, "Electric Filters," U. S. Patent No. 2,106,785; February 1, 1938.
[3] H. H. Scott, "A new type of selective circuit," Proc. IRE, vol. 26, pp. 226–235; February, 1938.

tions. The wide interest shown in the parallel-T network, as evidenced by the number of papers concerning it, arises from its unusual properties. It has a sharp and stable transmission null at a predetermined frequency in an unbalanced form as required by electronic circuits. Being composed of R and C alone, high impedance levels are easily accommodated, and nets with null frequencies as low as a fraction of a cycle per second are practical. It is quite linear with voltage, as contrasted with LC circuits for low-frequency use. In addition to its usefulness in practical applications, the mathematical description of its behavior yields to standard analytical techniques. Complete response equations and sets of design charts for the better known applications are in print and readily available. However, there is still some misunderstanding of the advantages and limitations on the use of the parallel-T in the solution of circuit problems, and some remarks on old applications may be helpful. In addition, some extensions of the technique can be described which increase the utility of this versatile RC net in the solution of practical problems.

II. The RC Parallel-T Notch Filter

The simplest use of the parallel-T net is as a filter with a symmetrical notch response, where a single frequency is to be completely removed from the signal spectrum. Here the narrowest possible notch is usually wanted. A general symmetrical parallel-T is sketched in Fig. 1, after Valley and Wallman.[4] The parameter n fixes the internal ratios of the RC components. To produce a null at a frequency f_0, the RC product must meet the equation $(2\pi f_0 rC)^2 = n$. For a zero resistance source and an open circuit as a load, Valley and Wallman derived the condition for the narrowest notch, which occurs at $n=1$. The major objection to this solution is that the source impedance usually is not negligible. Furthermore an improvement in Q can actually be obtained by deliberate loading. Cowles[5] presents the response equations in complete form covering the loaded cases. If we impose the restriction on the load resistances, as shown in Fig. 1, to give equal high- and low-frequency response, the general response equation is

$$\beta = \frac{1}{1 - j\dfrac{\rho}{Q(\rho^2 - 1)}}$$

where

$$\rho = f/f_0.$$

The concept of Q is extended to the notch case by taking Q to be the reciprocal of the bandwidth at the 3-db points of the response, analogous to the usual tuned circuit response. Fig. 1 shows a set of symmetrical responses for various resistance loadings of a parallel-T in which $n=2$. The situation on notch width is displayed more clearly if we plot the Q of the response versus dc (or high-frequency) transmission, as in Fig. 2, for networks with various values of n. Note that any particular network has a maximum value of Q at some specific loading condition, and that improvements in Q come at an expense in insertion loss. The values of the dc and high-frequency transmission (which are assumed equal) are used as voltage response parameters. The source and load resistances consequently must be related to the network resistance parameter in a way to satisfy the symmetry of Fig. 1.

A usual value of Q for a parallel-T is 0.25, with a limiting upper value of 0.5 under the restrictions assumed here. Lower values than 0.25 may be useful for special

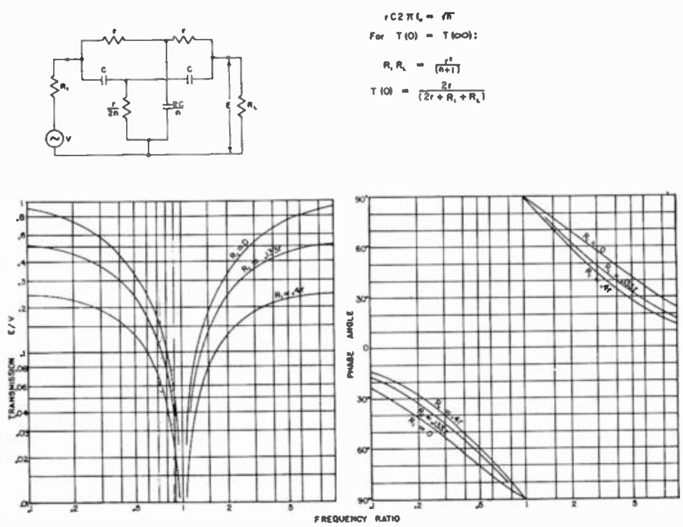


Fig. 1—Parallel-T and typical symmetrical notch response.

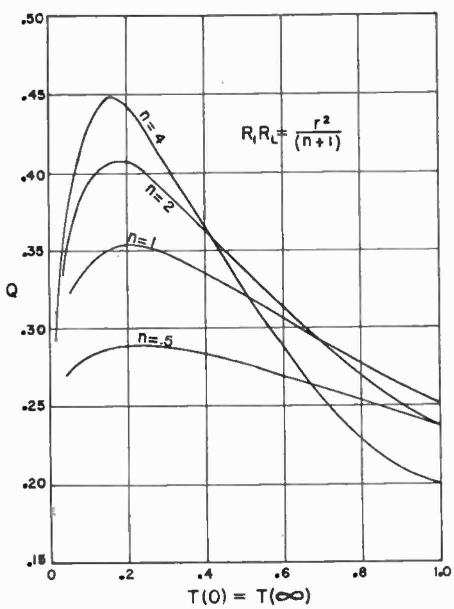


Fig. 2—Q of a parallel-T notch response vs network parameters.

[4] G. Valley and H. Wallman, "Vacuum Tube Amplifiers," M.I.T. Rad. Lab. Series, McGraw-Hill Book Co., Inc., New York, N. Y., vol. 18; 1948.

[5] L. G. Cowles, "Parallel-T RC networks," Proc. IRE, vol. 40, pp. 1712–1717; December, 1952.

situations. For example, in a feedback stage giving a single-tuned response, the Q of the response is approximately GQ_0, where G is the feedback gain and Q_0 is the Q of the feedback net as defined above. However, the response far away from the notch frequency is $1/G$ and it is desirable to have a large G for best skirt selectivity. For a given tuned response Q, it may be best to use a large value of stage gain and a net of intentionally low Q_0. In connection with this application, it is not necessary and often not convenient to design for a low source impedance to feed the parallel-T. The loading may be quite large, as in transistor circuits, so long as the symmetry condition on source and load resistances is satisfied. For heavy loading as is often convenient in transistor coupling nets, a value of n different from the usual $n=1$ or $n=2$ usually will be preferred.

If the parallel-T is to be a frequency elimination filter to remove one component from a complex spectrum, a Q of even 0.5 is inconveniently broad. A simple active circuit using one cathode follower is described by Sallen and Key.[6] A typical circuit is shown as the first half of the active filter of Fig. 5. Under the assumption that the network impedance is very much higher than the dynamic resistance of the cathode follower, the response is identical in form with that of the net itself except that the Q has been increased. The width of the notch is easily decreased by a factor of 10, giving a Q of 3 or more. A penalty must be paid in the depth of the notch attenuation because any residual unbalance at the null frequency is multiplied by the same factor as the Q. Even so, the circuit arrangement has high practical utility.

Other typical uses of the symmetrical notch circuit such as discriminators and derivative networks[7] have been described elsewhere.

III. Low-Pass Filter

The parallel-T is frequently used as either a high-pass or a low-pass filter. In Fig. 3 are shown response curves of typical nets with the circuit constants arranged to give a low-pass response. Source and load resistances must be properly chosen to give some attenuation at the high-frequency end. Note that the curve is like that of a simple m-derived filter except for the long frequency scale. The lower curve is for a net with $n=1$ as is most commonly used. The middle curve is for $n=2$, which is near optimum for this set of loading conditions. The pass band is appreciably wider, showing that for a proposed application, it is worth the time required to find the net that will be optimum for the assigned terminal conditions. The full response equations are given by Cowles.[5]

[6] R. P. Sallen and E. L. Key, "A practical method of designing RC active filters," IRE Trans. on Circuit Theory, vol. CT-2, pp. 74–85; March, 1955.

[7] G. E. White and M. Relson, "Alternating Current Rate Circuits," U. S. Patent No. 2,446,567; August 10, 1948.

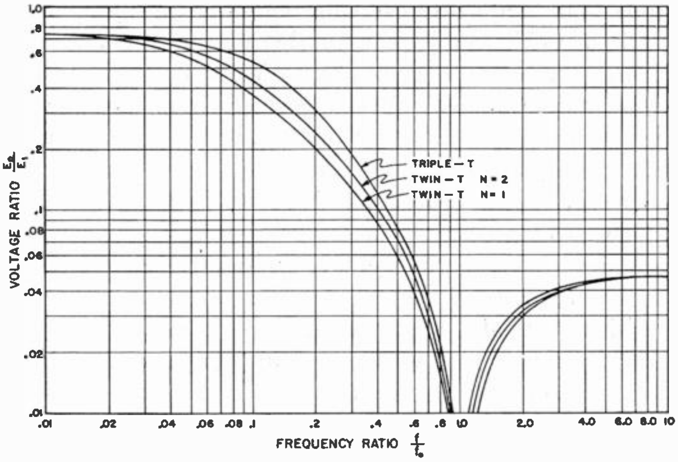


Fig. 3—Low-pass filter responses.

The usual objection to the parallel-T as an m-derived filter substitute is that the region between pass and cut-off is much too wide for most applications. This raises the question whether or not a more complex circuit might have better behavior. With more circuit parameters at the free disposal of the designer, a better shape might be given to the pass region. Of the several methods available for investigating circuits of the parallel-T type, symmetrical lattice theory[8] appears to offer the easiest solutions. Ordinary circuit equations lead to a quick solution of the simple parallel-T, but with the expectation of more elements in the network, mesh analysis leads to cumbersome algebraic forms. The lattice methods permit the use of auxiliary graphical aids so that intuition can be brought into play, and the final step of writing the full algebraic response equations can be deferred until most of the investigation is complete. A major limitation is that networks derived in this way will be symmetrical.

In Fig. 4 is shown a primitive lattice form which contains one more RC loop in the series arm than the simple parallel-T. Following a common method for lattice analysis,[8] the source resistance is taken as 1 ohm and transformed into the lattice leaving a current source as the driving function. The 1 ohm taken from the load leaves a residual r_L which is positive or negative depending upon whether or not the actual load resistance R_L is greater or less than 1 ohm. The transmission ratio is given by

$$T(\omega) = (Z_b - Z_a)\frac{r_L}{2r_L + (Z_b + Z_a)}.$$

This equation is general for the symmetrical lattice with source and load resistance. For the usual case where the source and load resistances are both 1 ohm, the response equation simplifies to

$$T(\omega) = (Z_b - Z_a)/2.$$

[8] E. A. Guillemin, "Synthesis of Passive Networks," John Wiley & Sons, Inc., New York, N. Y.; 1957.

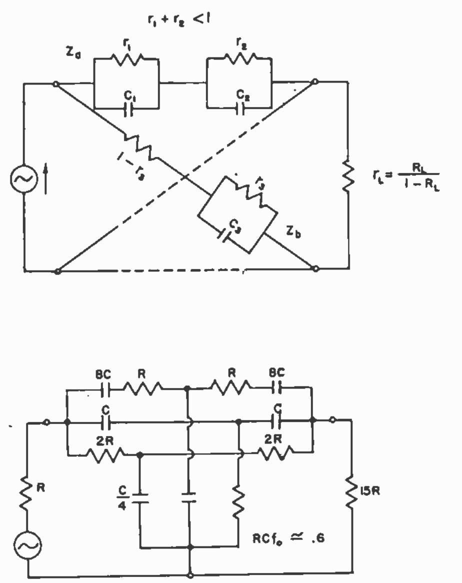


Fig. 4—Typical triple-T low-pass filter derived by lattice analysis.

The attractiveness of this design method lies in part in the ease with which the equations for the individual Z_a and Z_b arms can be written, usually by inspection, or sketched as polar diagrams as an aid in arriving at a desired solution. Further, the effect of each circuit parameter is separated in the equations so that the result of varying each element is readily apparent.

As an example of one problem solved in this fashion, it was desired to find a network with a low-pass response like those of Fig. 3, with the requirement of maximum flatness. Further constraints were that the dc response ratio would be 0.75, and the high-frequency response ratio would be 0.05. These latter two ratios limit the freedom of choice in the source and load resistances, but in most applications the load resistance can be set at any reasonable value. The requirement for maximum flatness will affect the choice of the internal resistance-capacitance products, and although an endless number of solutions can be found with maximum flatness, one will have a maximum bandwidth in the pass band.

In order to find this optimum network, a numerical approach is easier to use than an analytical solution because of the high degree of the explicit response polynomials. In one procedure, a set of values is assigned to the resistances in the series arm of the primitive lattice shown in Fig. 4. Note that the resistance of the cross arms is 1 and the resistance of the series arms is less than 1 because of the assumption of a 1-ohm source. Various ratios are then taken for the capacitors in the series arms, and the response is plotted for each from the equation for $T(\omega)$. The impedance components of the cross arm Z_b are constrained by the requirement that at some real frequency $Z_a = Z_b$ because this is the null condition. This determines Z_b as a function of the notch frequency f_0. For some ratio of the two series arm capacitors the response will be maximally flat. Other values are taken for the series arm resistors and again maximally flat solutions are found. Of the maximally flat solutions, the circuit with the widest pass band is selected as the best solution.

The remaining process is that of finding the unbalanced T equivalents of the lattice by the equivalent circuits shown in Guillemin.[8] Not all lattices have unbalanced T equivalents, and even when they exist, ingenuity may be required to find a practical T circuit. The solution to the problem outlined above is shown in Fig. 4, and the response is plotted in Fig. 3.

In spite of the complexity of the design process, the nets usually are practical to construct, and the improvement in the width of the pass band is appreciable. There is, of course, no difficulty in transforming a given design into different impedance levels or different frequency scales. In a typical use of this filter, it might follow a demodulator operating on a 60- or 400-cps carrier. The notch frequency is set at either the fundamental or the second harmonic, depending upon the type of demodulator.

If active circuit solutions can be tolerated, a response much like that of a conventional m-derived low-pass LC filter can be obtained with the use of cathode follower circuits. Such a filter is shown in Fig. 5. A narrow notch circuit is followed by a low-pass circuit,[6] with the individual responses as shown in the curves. The product curve is shown as the complete response. The output contains a dc offset, so that a more complex circuit than that shown must be used if the dc component is to be recovered. In some situations, as in an electrocardiograph, the dc component is discarded and this filter offers a good solution to the power frequency interference problem that besets most EKG work. Full design equations will be found in Sallen and Key.[6]

IV. Band-Pass Filters

The parallel-T offers a very useful solution to the problem of designing band-pass filters at low frequencies. Inductors of high quality at frequencies below 100 cps usually exhibit nonlinearities, and become useless at frequencies of the order of 10 cps for filter design purposes. LC sections at low-audio and subaudio frequencies consequently are not practical. The use of the parallel-T in a feedback amplifier to simulate a single-tuned LC circuit is widely known and practiced since its first publication by Scott.[3] Valley and Wallman[4] give useful data not only for the single-tuned amplifier, but for the more general case of stagger-tuning. By proper frequency and Q selection, a band-pass filter of any number of poles can be produced.

A feedback amplifier stage is shown in its elementary form in Fig. 6. If the beta network is a parallel-T net, the single-tuned response results. If Q_0 is the effective Q of the network as plotted in Fig. 2, and G is the open loop gain through the net and amplifier at very high- or very low-frequencies, then the response Q is given ap-

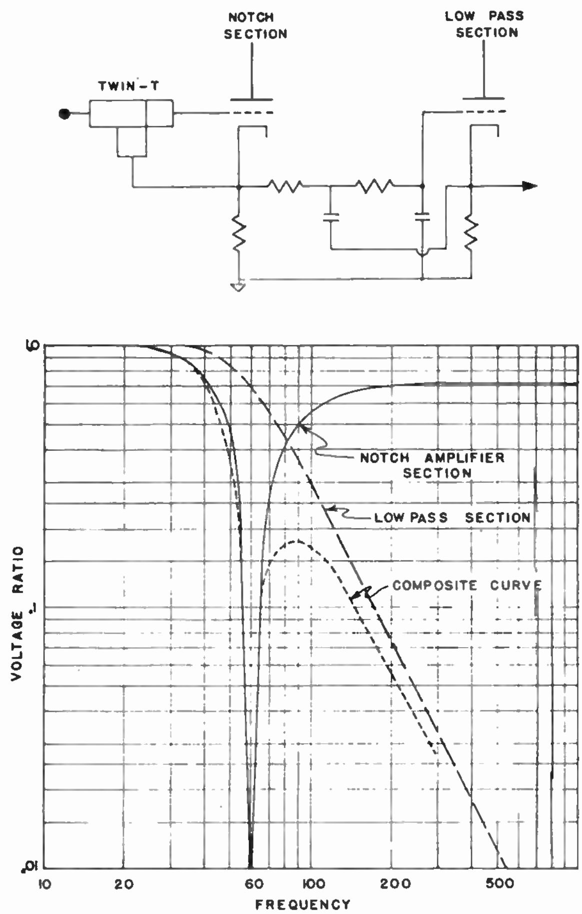


Fig. 5—Active low-pass filter with m-derived response.

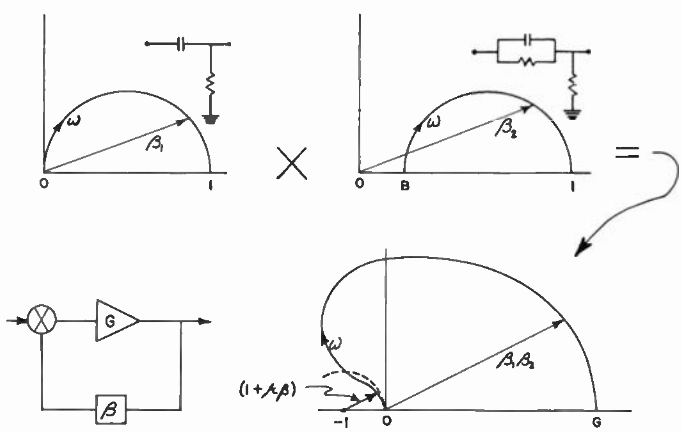


Fig. 6—Beta networks for a 2-pole feedback net.

proximately by GQ_0. Note that the open loop feedback gain usually is not the same as the forward or signal gain. With Q_0 of the order of 0.2 it is quite practical to obtain a stable response Q up to 50 (which implies a bandwidth of 2 per cent). Practical circuits in many forms have been published.[3,4] The single-tuned response usually is used where a single frequency predominates in the signal and no wide-band modulation components are present. The center frequency may be made as low as a fraction of a cycle per second.

The single-tuned circuit response is of limited utility because the double-tuned or two-pole response is usually the minimum that can be considered a true band-pass selectivity. In LC circuit equivalents, this is a half-section of a constant-k filter. It is easily approximated by stagger-tuning two single-tuned stages in cascade[4] and in either case it is easy to arrange for a region of flat response in the pass band. It is not generally known that a net can be designed to give a two-pole response through feedback around a single amplifier stage. For ease of analysis, the low-pass form of the network may be derived first, and then the band-pass net may be obtained by a frequency transformation. As an initial understanding of the problem, consider the shape imposed on the beta network response, as shown in Fig. 6.

The two-pole low-pass amplitude function starts at the origin with zero slope, and rises to a maximum for a Tchebycheff response, or is flat for an appreciable region for a Butterworth response. To obtain these effects by feedback, we note that the plot of the beta network response in polar form must be tangent to the unit circle centered at -1. For a Tchebycheff response, the beta locus must cut into this unit circle, since the amplifier response is the reciprocal of the $1+G\beta$ vector. For the maximally flat or Butterworth response, the beta locus and the unit circle must remain tangent for the greatest possible distance. In the cutoff region, the $1+G\beta$ vector becomes larger than unity and eventually must be very large compared to unity. For physical realizability, the $G\beta$ locus must swing clockwise and terminate on the real axis.

Of the simple networks which might have a transfer function of the necessary character, consider the high pass nets of Fig. 6. Let the response of net 1 be given by

$$\frac{j\tau_1\omega}{1+j\tau_1\omega}$$

and that of net 2 by

$$\frac{B+j\tau_2\omega}{1+j\tau_2\omega}.$$

The parameter B is the dc response of net 2, and we are assuming that both have unity response at high frequency, and no interaction when connected together. In an actual problem the loading of net 1 by net 2 will be appreciable and will modify the expression for the cascaded response, but the form of the expressions will remain the same. Under these simplifications, we have the beta response to be

$$\beta = \frac{j\tau_1\omega}{1+j\tau_1\omega}\left(\frac{B+j\tau_2\omega}{1+j\tau_2\omega}\right).$$

The over-all response is then

$$\frac{1}{V} = 1 + G\beta = 1 + G\frac{j\tau_1\omega(B + j\tau_2\omega)}{(1 + j\tau_1\omega)(1 + j\tau_2\omega)}$$

where V is the response of the completed circuit.

We form the power response by multiplying this equation by its conjugate, obtaining

$$\frac{1}{VV^*} = \frac{1+(\tau_1+\tau_2BG\tau_1)^2\omega^2 - 2\tau_1\tau_2(1+G)\omega^2+\tau_1^2\tau_2^2(1+G)^2\omega^4}{1+(\tau_1^2+\tau_2^2)\omega^2+\tau_1^2\tau_2^2\omega^4}$$

$$= \frac{1+b_2\omega^2+b_4\omega^4}{1+a_2\omega^2+a_4\omega^4}.$$

We may find the constraints on the a and b coefficients to have a maximally flat response approximating a Butterworth shape. One condition for maximal flatness will be $a_2 = b_2$, as can be seen by requiring that the second derivative with respect to ω^2 be zero at $\omega = 0$. This condition leads to

$$\frac{\tau_1}{\tau_2} = \frac{2(1 - B)}{B(BG + 2)}.$$

For $a_2 = b_2$, the response equation may be simplified to

$$\frac{1}{VV^*} = \frac{1 + (\tau_1^2 + \tau_2^2)\omega^2 + \tau_1^2\tau_2^2(1 + G)^2\omega^4}{1 + (\tau_1^2 + \tau_2^2)\omega^2 + \tau_1^2\tau_2^2\omega^4}.$$

For simplicity we will assume that the cutoff frequency occurs at $\omega = 1$, and here the power response will be $\frac{1}{2}$. We obtain as an additional relation

$$1 + \tau_1^2 + \tau_2^2 = \tau_1^2\tau_2^2(G^2 + 2G - 1) = 0.$$

This may be rearranged as

$$\tau_1^2\tau_2^2 = \frac{\tau_2^2(1 + \tau_2^2)}{\tau_2^2(G^2 + 2G - 1) - 1}.$$

We are approximating the classical Butterworth form

$$\frac{1}{VV^*} = 1 + \omega^4.$$

Hence we wish a_2 and a_4 to be as small as possible, consistent with other requirements. For a minimum a_4, we set

$$\frac{\partial\tau_1^2\tau_2^2}{\partial\tau_2^2} = 0 = \lfloor\tau_2^2(G^2 + 2G - 1) - 1\rfloor - (1 + \tau_2^2)\tau_2^2(G^2 + 2G - 1)$$

and

$$\tau_2^2 = \frac{1 + (G^2 + 2G)^{1/2}}{G^2 + 2G - 1}.$$

This is also the value of τ_1^2, as can be verified by substitution and $\tau_1 = \tau_2$ for minimum a_4.

For large G, $\tau_2 \doteq G^{-1/2}$. Also with G large and $\tau_1 = \tau_2 = \tau$,

$$B \doteq (2/G)^{1/2}.$$

Hence,

$$\frac{1}{VV^*}\ \frac{1 + 2\tau^2\omega^2 + \tau^4(1 + G)^2\omega^4}{1 + 2\tau^2\omega^2 + \tau^4\omega^4}.$$

For large G, the response equation reduces to

$$\frac{1}{VV^*} = \frac{1 + 2\omega^2/G + \omega^4}{1 + 2\omega^2/G + \omega^4/G^2}.$$

This approximates the Butterworth form quite well in the pass band for typical values of G, with increasing departure in the cutoff region. However, the error is only of the order of $1/G$. Obviously the extreme attenuation in an amplifier of this form cannot be more than $1/G$ with a β of the assumed form.

In principle, it is possible to compare the classical Tchebycheff form[7] with the original response equation for the filter and thus derive constraints on the parameters to approximate the Tchebycheff (or double humped) response form. The results are difficult to set down in closed form, and since much of the design process is empirical, the Butterworth solution is usually an adequate beginning for experimental data. It can be inferred from the Nyquist plot of Fig. 6 that the Tchebycheff response can be obtained in two ways. If a Butterworth network is placed in the feedback path, and the feedback gain is reduced, then the G-beta plot shrinks and the unit circle cuts it in two places. Since $1+G\beta$ has a magnitude less than 1 when it lies inside this unit circle, the gain which is its reciprocal must have a response peak, corresponding to a Tchebycheff form. The bandwidth to the 3-db point has been increased in the process.

Alternately, we may hold the feedback gain constant and emphasize the cardioid character of the G-beta plot by using a smaller B. We again get a peaked Tchebycheff form, but the bandwidth has not been increased as much as in the first instance.

The advantage of the Tchebycheff response is that the skirt of the selectivity curve falls more steeply than in the Butterworth response for a given amount of bandwidth. Whether or not it should be used depends upon the demands of the problem. In most cases, the maximally flat Butterworth response is preferred.

To transform the constants to a specific cutoff frequency ω_c, we select new values of τ_1 and τ_2, so that the new values are $\omega_c\tau_1' = \tau_1$ and $\omega_c\tau_1' = \tau_2$. To transform the low-pass networks to the band-pass case, with a response symmetrical about a specified ω_0, we substitute for ω the transformation variable $\omega/\omega_0 - \omega_0/\omega$.

To maintain a band-pass bandwidth of ω_c, we substitute $\omega_c\tau$ for Q, where Q has the usual meaning for each part of the beta network (since each half now has a second order response equal to a properly chosen LC pair).

If the β networks are to be constructed of RC elements only, the β_1 net is a parallel-T with a null at f_0. The β_2 net is a bridged-T or a parallel-T adjusted for a minimum response equal to B. In a practical parallel-T network, the Q usually lies between 0.1 and 0.3, imposing a severe limitation on design flexibility.

The β nets also may be made of LC components. The β_1 net will be a bridged-T with a null at f_0, and the β_2 a bridged-T with a minimum response B at f_0. It should be noted that the values of Q needed for a given bandwidth are much less than needed to construct an equivalent passive filter. However, any nonlinearities of the inductors will be emphasized by the amount of the feedback gain, and hence the results may not be useful.

A typical amplifier and band-pass feedback network are shown in Fig. 7. The design parameters of this circuit are as follows:

$$G = 120 \text{ (effective feedback gain)},$$

$$Q_1 = Q_2 = 0.20 \text{ (parallel-}T\text{ networks)},$$

$$\text{forward gain} = 80.$$

By using the conditions on the circuit parameters and the low-pass to band-pass variable substitutions, we have $Qf_c/f_0 \doteq G^{-1/2} = 0.091$, or $f_c/f_0 = 0.45$. The bandwidth of the filter is 45 per cent, or from about $0.80 f_0$ to $1.25 f_0$. The parameter B, representing the amount of unbalance of the parallel-T which has no null, is $B \doteq (2/G)^{1/2} = 0.13$.

The response curve of a band-pass filter with a center frequency of 1 cps is shown in Fig. 7. The largest capacitor in the feedback net is about 1 mfd when typical resistance values are 0.5 mfd. Filters with center frequencies as low as 0.05 cps are in use. Spectrum analyzers are typical problems which make use of this type of active filter, since many real time analysis problems involve signals at frequencies lower than can be handled by conventional selective filters.

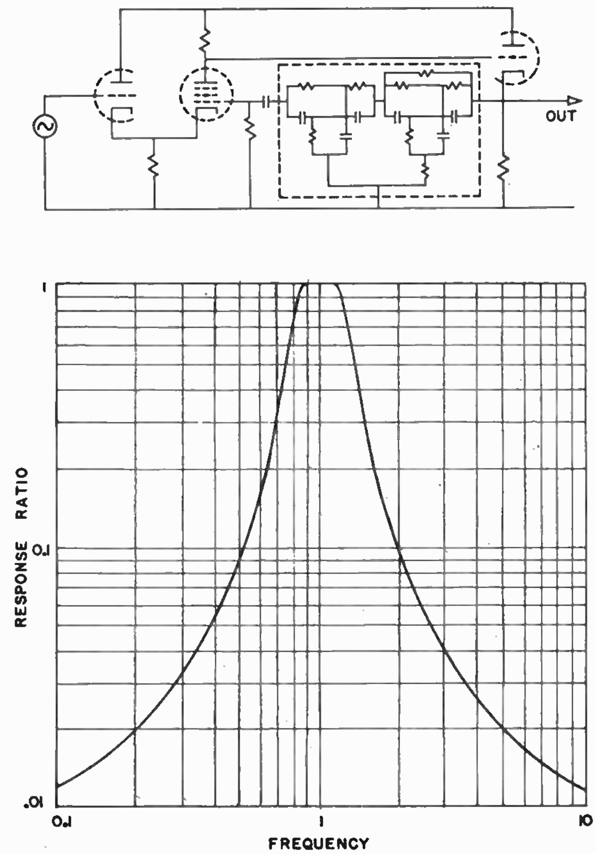


Fig. 7—Band-pass feedback net with a 2-pole response.

V. Component Stability Problems

In the construction of null networks, the RC products inside the net must have a specified value within very close tolerances in order that the responses have a good null. Feedback circuits are particularly demanding in this respect, because any imperfections of the null are multiplied by the feedback gain. Not only must the null be good initially, but it must be maintained against temperature variations and the effects of time. The parallel-T has acquired an unjustly poor reputation with many circuit designers because solutions to the stability problem were not worked out.

If the notch transmission is to be less than 0.001 (60 db) which is a typical requirement, then the RC product must approximate the required value with the same order of accuracy. It is not economical to adjust both R and C to within 0.1 per cent of the theoretical value, but a trimming procedure in either R or C can be devised to bring the network into balance by experimental procedures. These procedures can be set on a production basis and present no problem other than the maintenance of test voltages of accurate frequency and very clean waveform. The initial accuracy of the null adjustment is thus easily assured.

Not so apparent is a solution to the problem of temperature stability. Components must of course have very low drift with time and must show good retrace characteristics with temperature cycling. Ideally, the temperature coefficients of the capacitors should be equal to that of the resistors and opposite in sign. The choice of components involves economics since the most stable components generally are the most expensive.

For problems involving only straight filtering or only low-gain feedback circuits, a combination of plastic film capacitors (such as Mylar) and precision film resistors is usually adequate. For example, the low-pass nets of Section II and the wide-band feedback nets of Section IV usually can be made in this way. Film capacitors are small in size and relatively low in cost. Standard composition resistors and impregnated paper capacitors generally are not acceptable.

For the highest accuracy, a combination of silvered mica capacitors, wire-wound resistors, and a small amount of negative coefficient resistance trimmer gives good performance at null frequencies above 100 cps. The negative coefficient compensating resistors are necessary because silvered mica capacitors usually have a small positive coefficient. If performance above room temperature is important, then the combination should not be used at low null frequencies because of the behavior of mica as a dielectric. Its apparent leakage increases rapidly with temperature at low frequencies causing a loss of null in the network, and the effect is not compensated by any simple means. The performance at 400 cps and higher is generally quite good.

For null frequencies below 100 cps, a combination of polystyrene capacitors, wire-wound resistors, and a small percentage of positive coefficient resistance gives a superior performance. Because of component cost and bulk, this method of construction is used only where extreme stability is required.

In Fig. 8 are plotted typical temperature runs of three combinations of components, each within its best frequency range. The errors have been plotted as the transmission ratio with the test voltage set at the correct frequency. No attempt has been made to show the phase angle of the error voltage. With the proper selection of components, feedback circuits based on the parallel-T network provide stable filters with characteristics not obtainable from passive LC circuits. They deserve wider application in low-frequency instrumentation problems.

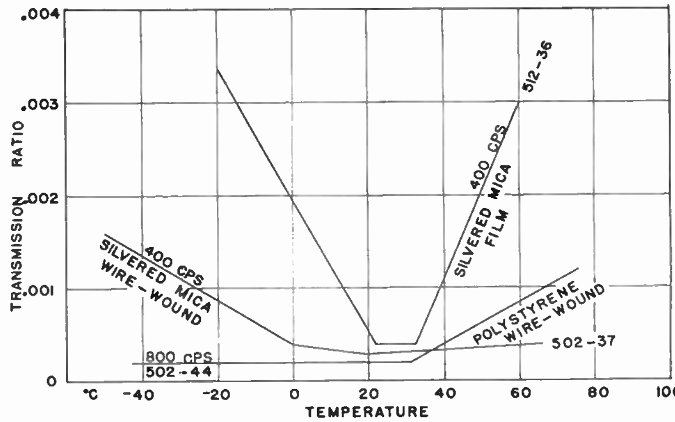


Fig. 8—Typical temperature errors of parallel-T networks.

Correspondence

Comments on "Nonlinear Distortion Reduction by Complementary Distortion"*

It was with great interest that I read J. Ross Macdonald's article in the September–October Transactions on Audio.[1] There are two reasons for this interest. One is that practical techniques for this sort of distortion reduction have been used by myself and others for some years. The second is that having thought about this problem a bit, I have a couple of comments on this type of analysis.

The type of circuit used to obtain complementary distortion cancellation is shown in Fig. 1. A simple voltage amplifier is followed by a cathode follower. It is very convenient to use a connection of this sort because of the common availability of two triodes in a single envelope. There are distinct advantages to the use of two tubes as a single amplifier. In the first place, the gain of the combination is essentially independent of the load and the output impedance of the amplifier is low. In the second place, there is the advantage of distortion reduction by complementary distortion. The voltage amplifier and the cathode follower have transfer curves with opposite curvature and the total distortion produced in the combination is considerably less than that found in either the voltage amplifier or the cathode follower alone. It may at first seem somewhat of a surprise to find that distortion decreases as one proceeds down the line of amplifier stages in such a connection but apparently it is exactly the sort of phenomena which Macdonald describes which is taking place. If some care is taken in selecting the operating points of the two stages, it will be found that a very low value of total harmonic distortion may be achieved.

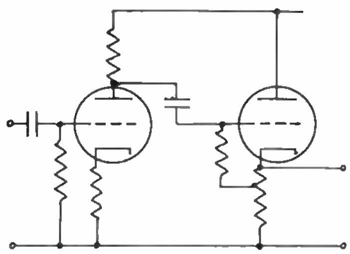


Fig. 1.

I would like at the same time to present the following comment on Macdonald's statements. He states that it is not possible to cancel a square law distortion component with a second square law component of distortion and this is certainly the case. This result would seem to indicate that since it is indeed not a square law distortion of one sign which cancels a square law distortion of another sign, but rather some other form of distortion, we should then not expect to form a complementary transfer characteristic in this manner. The required transfer characteristic will have to be formed out of many higher-order terms. This result is indicated by the analysis. However, it is unfortunate to say that such cancellation is impossible. Certainly we can conceive of the correctly shaped transfer curve such that the over-all transfer characteristic is a straight line. The shape of this required transfer characteristic would probably be described by an infinitude of harmonic terms, of course. But it is fortunate that such a characteristic may be very much easier to synthesize than to analyze.

The fact that Macdonald did not work out an infinitude of terms (and have a very long paper indeed) should not be taken as an absolute denial to the perfectionist that it is possible to resurrect a distorted signal by complementary distortion.

R. A. Greiner
Dept. of Elec. Engr.
University of Wisconsin
Madison, Wis.

* Received by the PGA, December 10, 1959.

[1] J. R. Macdonald, "Nonlinear distortion reduction by complementary distortion," IRE Trans. on Audio, vol. AU-7, pp. 128–133; September–October, 1959.

Contributors

Benjamin B. Bauer (S'37–A'39–SM'44–F'33) was born on June 26, 1913, in Odessa, Russia. He graduated from Pratt Institute, Brooklyn, N. Y., in 1932 and received a degree of E.E. from the University of Cincinnati, Cincinnati, Ohio, in 1937.

B. B. Bauer

He joined Shure Brothers, Inc., Chicago, Ill., manufacturers of electro-acoustical devices, as a co-op student in 1936, and became chief engineer in 1940 and vice president in 1950. In 1957, he joined CBS Laboratories, Stamford, Conn., and was appointed a vice president in 1958. He is in charge of the Audio, Acoustics, and Magnetics Research in the Laboratories. He is responsible for the development of many electroacoustical devices such as microphones, phonograph pick-ups, and tape-recording heads, and holds over 30 patented inventions in the fields of audio, acoustics, and magnetic recording.

Mr. Bauer is a Fellow of the Acoustical Society of America, and the Audio Engineering Society. He is also associate editor of the *Journal of the Acoustical Society of America* and, in 1955, as co-founder and past National Chairman (and present Secretary) of the IRE Professional Group on Audio, he was recipient of the Group's "Achievement Award." He is a member and recipient of the "Recognition Award" of Eta Kappa Nu, and a member of Tau Beta Pi and Sigma Xi.

❖

George W. Sioles (M'52) was born on October 15, 1924 in Peekskill, N. Y. He entered Columbia University, New York, N. Y., in 1942 and interrupted his studies in March, 1943 to join the U. S. Army Air Corps, where he flew combat missions as lead bombardier-navigator of a B-29 squadron. He re-entered Columbia in 1946, and received the B.S. degree in chemical engineering in 1949.

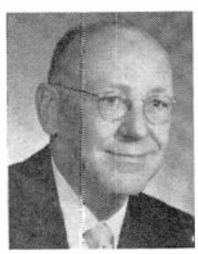
Following this, he set up his own company for the design and installation of high-fidelity equipment. From 1951 to 1953, he was with the Special Devices Center of the Office of Naval Research, and from 1953 to 1957, was with University Loudspeakers, where he became Manager of Engineering. He joined CBS Laboratories, Stamford, Conn., in January, 1959, to take charge of the Group for Transducer Research, and has recently been appointed head of the Section on Audio Research.

G. W. Sioles

Mr. Sioles is a member of the Acoustical Society of America and the Audio Engineering Society.

❖

William B. Snow (A'26–VA'39–SM'50) was born on May 16, 1903, in California. He received the B.S. degree in electrical engineering in 1923, and the E.E. degree in 1925, both from Stanford University, Stanford, California.



W. B. Snow

He worked in the Acoustical Research Department at Bell Telephone Laboratories, New York, from 1925 to 1946, being very active in the study of stereophonic sound reproduction and recording. From 1941 to 1945, he was under a leave of absence to join the U. S. Navy Underwater Sound Laboratory at New London, Connecticut, where he became assistant director and specialized in acoustical measurements. In 1946 he joined the Kellex Corporation (now Vitro Corporation), where he was appointed director of Physical Research and Development in 1950. Since 1953 he has carried on a private consulting practice in acoustics and noise measurement and control in Santa Monica, California.

Mr. Snow is a Fellow of the AES and the Acoustical Society of America, a member of AIEE and AAAS. In 1956 he received the PGA Service Award.

Frank H. Slaymaker (M'45–SM'59) was born in Linwin, Neb., on April 22, 1914. He received the B.S. degree in electrical engineering from the University of Nebraska, Lincoln, in 1941 and the Professional E.E. degree in 1946.

F. H. Slaymaker

In 1941 he became a research engineer at the Stromberg-Carlson Company, a division of General Dynamics Corporation, Rochester, N. Y., where he currently holds the position of manager of the Electro-acoustics Laboratory. He is a registered Professional Engineer in the state of New York, and holds numerous patents in the fields of telephone switching, noise reducing microphones, ultrasonic transducers, ultrasonic echo ranging systems and electronic carillons.

Mr. Slaymaker is a member of Sigma Xi and a Fellow of the Acoustical Society of America.

❖

Gifford E. White (S'41–A'42–SM'47) was born in San Saba, Tex., on February 17, 1912. He received the B.A. and M.A. degrees in physics from the University of Texas, Austin, Tex., in 1939. He later did two years of graduate work in electrical engineering at the Massachusetts Institute of Technology, Cambridge.

G. E. White

He has been employed by Humble Oil and Refining Company, Sperry Gyroscope Company, Statham Instruments, Inc., and is presently the owner of White Instrument Laboratories, Austin, Tex. He holds numerous patents in the physical instrumentation field.

Mr. White is a member of the American Physical Society, the Society of Exploration Geophysicists, Eta Kappa Nu, and Sigma Xi.

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.

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.