IRE Transactions on AUDIO

Volume AU-9 NOVEMBER-DECEMBER, 1961 Number 6

Published Bi-Monthly

TABLE OF CONTENTS

CONTRIBUTIONS

CORRESPONDENCE

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 79 Street, New York 21, N.Y. Responsibility for the contents rests upon the authors, and not upon the IRE, the Group, or its members. Individual copies of this issue and all available back issues may be purchased at the following prices: IRE members (one copy) $2.25, libraries and colleges $3.25, all others $4.50. Annual subscription price: non-members $17.00; colleges and public libraries $12.75.

The Editor's Corner

CONFESSIONS OF AN AUDIO ENGINEER

WHEN I was an assistant apprentice I worked like a horse, but didn't get much hay on payday. My boss used to insult my intelligence and waste my time by explaining things that I already knew. Then he wondered why the work didn't get finished. He would give me impractical designs dreamed up by far-out thinkers in the lab; I had to do them over completely to make them operate. After I came up with some pretty neat solutions, he would run them down and say it wasn't what they wanted at all. Sometimes I'd work a whole year and come up with a masterpiece. Then they would say "Very nice, boy, but the management has decided to go into something else." So, it was back to the drawing board.

By the time I became a supervisor, things had changed completely. The assistant apprentices were short on know-how, but long on self-confidence. Let's say there was a simple job that I could have done by myself, easily in a couple of hours. I'd spend half a day selling my assistants on why it should be done at all, why they should do it rather than someone else, why it should be done a certain way rather than in several other ways, and why we needed it by a definite date. When the job was due, I would spend half a day hearing why it wasn't finished, why it doesn't (and can't possibly) operate the way it should, and why we must start over and do it right. The "right" way, of course, was so incredibly complicated that it would take a year to set it up, and another year to make it work. At this stage, somehow, by a combination of diplomacy, flattery, salesmanship, and pleading, we got them to patch it up enough so that it barely passed.

That was only one part of it. Management was even worse. They wanted me to write so many reports, memos and proposals that these alone would take 100 per cent of my time. I had to answer routine mail and telephone calls. There was a constant stream of visitors. I had to recruit new people and keep the old ones satisfied. I was expected to enhance the company image by attending professional meetings, presenting papers, serving on committees, and organizing seminars.

Eventually I came into ownership of a small business, and now I am aware that supervisors and managers have changed for the worse. They don't try to get anything done, but instead, they spend their time organizing everything to the point where it breaks down and has to be reorganized. They travel, give talks, hold seminars, attend conventions, serve on technical committees—in fact, they do everything but work for the company. Together with their assistants, they represent a lot of brain-power, but their ingenuity is directed towards inventing diversions instead of new products. If there were a Nobel Prize for excuses, our research team would have won it long ago.

All things considered, it would be nice to be an assistant apprentice again, where the living is easy. Trouble is that I'm married now, and can't afford it.

(The characters in the above sketch are purely fictional, and the viewpoints expressed are not necessarily the official opinion of the management, the IRE-PGA, or the author.)

—MARVIN CAMRAS

PGA News

CHAPTER NEWS

Boston, Mass.

The Professional Group on Audio of the IRE for the Boston Section dates back to 1948 and credits its formation to an ardent group of "high fidelity" fans who wished to band together in a group to share their common interests. The group contained a small nucleus of those engaged professionally in acoustics and audio. John A. Kessler, the first Chairman, was associated with the M.I.T. Acoustics Laboratory. Others who shared in the formative stages were Ben Drisko and Paul St. George. Since that time, the group has steadily grown until it is now one of the largest groups in the Boston Section.

The chairmanship of the group has been held by John A. Kessler, William G. Burt, Jr., E. B. Dyett, Jr., Benjamin B. Drisko, Weiant Wathen-Dunn, Roger Prager, Richard S. Burwen, Daniel Von Recklinghausen, Donald J. Fritsch and presently by Henry S. Littleboy.

Papers that have been presented before the group have included discussions on loudspeakers by E. Villchur and A. A. Janszen, amplifiers by R. S. Burwen, architectural acoustics by R. H. Hunt and Leo Beranek, and hearing by J. C. R. Licklider.

During the fall of 1960, a lecture series was sponsored by the group together with the Boston Section at John Hancock Hall. The speakers were Dr. Jordan J. Baruch, supervisory engineer at Bolt, Beranek and Newman, Inc.; Edgar M. Villchur, President of Acoustic Research, Inc.; Hermon H. Scott, President and Director of Engineering of H. H. Scott, Inc.; Richard Kaye, Station Manager, WCRB; and Melvin Clark, Jr., Associate Professor of Nuclear Engineering at M.I.T. The sixth lecture was a panel discussion, led by Daniel Von Recklinghausen, Chief Research Engineer of H. H. Scott, Inc.

The 1961–1962 session of PGA will have the following program:

Months	Speaker and Company	Title and Abstract	Place and Time
September 28, 1961	Richard L. Kaye WCRB Station Manager	FM-Stereo Transmission (Transmission using newly authorized multiplex techniques)	Allston Burr B Harvard University 8:00 P.M.
October 26, 1961	Daniel Von Recklinghausen Chief Research Engineer H. H. Scott, Inc.	FM-Stereo Reception (Reception of signals transmitted by use of FM multiplex techniques)	Allston Burr B Harvard University 8:00 P.M.
December 12, 1961	Dr. Allan Grinnell Biological Laboratories Harvard University	Auditory Nervous System of Bats (Results of studies of how bats hear as conducted at Harvard)	M.I.T.
March 22, 1962	Kenneth Stevens Associate Professor of electrical engineering, M.I.T.	Machine Generation and Recognition of Speech (Programming of high-speed digital computers for speech studies)	M.I.T.
May 24, 1962	Amar G. Bose Associate Professor of electrical engineering, M.I.T.	Criteria for the Design of Loudspeakers (Design of loudspeakers using new techniques as developed at M.I.T.	M.I.T.

Chicago, Ill.

The first meeting of the 1961–1962 season was held jointly with the Chicago Chapter of IRE-PGI at Menard's Black Steer Restaurant, Wednesday, September 20, 1961. Earl E. Bockenfeld, Test Equipment Engineer, Hammond Organ Company, spoke on "Techniques for High-Speed Testing of Audio Amplifiers in Production." A summary of his paper is as follows.

The design and capability of a unique test set for measuring the performance of audio power amplifiers are described. The test set is intended for production line and inspection testing and has a test procedure sufficiently simplified to permit operation by nontechnical personnel. One test makes use of a decade switching tube and dry-reed relays in a commutating system to automatically measure frequency response. Another test makes use of a Go-No-Go colored lamp indicator to supplement conventional metering.

On Friday, October 13, 1961, at the Western Society of Engineers Building, a paper was presented on "Stereo Geometry," by Paul W. Klipsch, President, Klipsch & Associates Inc. Stereo geometry, or the localization effect, was measured by producing sounds in a systemic speaker array and determining the accuracy with which observers located the sounds under different conditions. A series of slides depicted different methods of measurement and indicated the type of speaker arrays used and the response of these speakers.

The talk was followed by a stereo demonstration using three large loudspeakers, including a pair of folded corner horns.

CHAPTER OFFICERS AND MEMBERSHIP STATISTICS

The IRE Professional Group on Audio includes seventeen active chapters as listed below. There are approximately 4900 members as of September 30, 1961. Officers are listed from the latest information received by the IRE headquarters to October, 1961.

Chapter	Chairman	Vice-Chairman	Secretary and/or Treasurer
Albuquerque-Los Alamos	John F. Dankworth 1133 S. Monroe St. Albuquerque, N. N.		Dwayne Fry Gen. Chenault N.E. Albuquerque, N. M.
Baltimore	James H. Jackson Bendix Radio Towson 4, Md.	William J. Parrish Westinghouse Air Arm Baltimore 3, Md.	George Kroen WMAR-TV Baltimore, Md.
Boston	Henry S. Littleboy Baird-Atomic 33 University Rd. Cambridge, Mass.		Charles I. Malme 27 Everett St., Apt. 8 Cambridge 38, Mass.
Chicago	James F. Novak Jensen Mfg. Co. 6601 S. Laramie Ave. Chicago 38, Ill.	William C. Wray Motorola Inc. 4501 W. Augusta Blvd. Chicago 51, Ill.	Lane W. Gorton General Radio Co. 6605 W. North Ave. Oak Park, Ill.
Cincinnati	John Brean Baldwin Piano Co. 1801 Gilbert Ave. Cincinnati 12, Ohio		
Cleveland	Henry B. Hoff Ohio Bell Telephone Co. 820 Superior Ave. Cleveland 13, Ohio		
Dayton	Stanley E. Weber WADC, Wright-Patterson AFB, Ohio	Albert Parker (same as Chairman)	Theodore C. Rynda (same as Chairman)
Hawaii	D. H. DaShiell 224 Awakea Rd. Lanikai, Oahu, Hawaii	Iwao Miyake University of Hawaii Honolulu, Hawaii	Daniel L. Pang 1809 Naio St. Honolulu, Hawaii
Houston	James H. Carter The Wrye Co. 2410 West Alabama Houston, Tex.		
Milwaukee	Anton E. Horvath Wisconsin Electric Power Co. Milwaukee 3, Wisc.		
Philadelphia	Karl V. Amatneek 8601 Hull Dr. Philadelphia 18, Pa.	Carl H. Brummer, Jr. 5530 Wayne Ave. Pennsauken 8, N. J.	Hans Dietze RCA Camden, N. J.
San Antonio	Bill Case 122 W. White Ave. San Antonio, Tex.		
San Diego	Roswell W. Austin Scripps Institution of Oceanography San Diego 52, Calif.	Donald M. Lowe USN Electronics Lab. San Diego 52, Calif.	
San Francisco	Ralph L. Brown Lenkurt Electric Co. 1105 County Rd. San Carlos, Calif.	Stanley K. Oleson Stanford Res. Inst. Menlo Park, Calif.	
Syracuse	E. O. Crow General Electric Co. Syracuse, N. Y.	Phillip B. Clark R.D. 3 Skaneateles, N. Y.	Dwight V. Jones General Electric Co. Syracuse, N. Y.
Twin Cities	Robert L. Sell 4740 Girard Ave. S. Minneapolis, Minn.	Richard F. Dubbe Minnesota Mining & Mfg. Co. 900 Bush Avenue St. Paul 6, Minn.	Joseph F. Dundovic Nortronics Inc. 1015 S. 6th St. Minneapolis 4, Minn.
Washington	J. Carlisle Hoadley Diamond Ordnance Fuze Lab. Washington, D. C.	Robert J. Carpenter Natl. Bureau of Standards Washington, D. C.	L. Glenn Whipple Federal Communications Commission Washington, D. C.

Power Requirements for Speech Communication Systems*

R. L. CRAIGLOW†, MEMBER, IRE, N. R. GETZIN†, MEMBER, IRE, AND R. A. SWANSON†, MEMBER, IRE

Summary—**This paper presents a simple method for predicting the power required for audio, AM, DSBSC, and SSB speech communication systems which are peak-power limited and are operating in the presence of "white" Gaussian noise. The effects of simple speech processing such as pre-emphasis, filtering, volume compression, symmetrical peak clipping, and frequency translation are considered both singly and in combination. The method is based on empirical data and is directly applicable to many practical speech communication problems encountered by the radio engineer.**

Introduction

EXTENSIVE studies have been made concerning the intelligibility of speech under various conditions of speech processing and noise. In the great majority of these investigations, the intelligibility has been measured either as a function of the average speech power to the average noise power ratio or in the absence of noise altogether.[1,2] In contrast, most present-day speech communications transmitters are peak-power limited. Moreover, the noise or interference experienced is generally proportional to the bandwidth of the receiver. That is, the noise-power spectrum is relatively flat over the frequency band of interest. The present study has been made under conditions that are closely related to those commonly encountered in practice.

The communications engineer is often interested in determining the transmitter power required to provide satisfactory communications for the task under consideration. The level of intelligibility required depends, of course, on the task. The criterion used in this study is the threshold of perceptibility for rather difficult connected discourse which corresponded to approximately 75 per cent monosyllabic phonetically balanced word intelligibility.[3,4] If a different level of intelligibility is required, an estimate of the required change in signal power is given.

A general block diagram of the type of speech communications systems investigated is shown in Fig. 1. The peak power delivered by the transmitter was held constant, and the noise power spectral density was changed until the threshold of perceptibility was reached. It would be very difficult to present the large amount of experimental data directly in a coherent and understandable form. Fortunately, it has been found possible, within the accuracy of measurement and within the range of systems investigated, to express the results in terms of a sum of correction factors.[5] The required ratio R of the peak instantaneous signal power to the noise power spectral density is therefore given by:

$$R_{\text{in db}} = \frac{S_p}{dN/df} \text{ in db} = 50 + \sum_n C_n \qquad (1)$$

where S_p is the peak instantaneous signal power in watts referred to the input of the receiver, dN/df is the noise power spectral density in watts per cps referred to the input of the receiver, and C_n is the correction factor. The factor of 50 db is added to account for the R required for a flat audio system employing no speech processing.

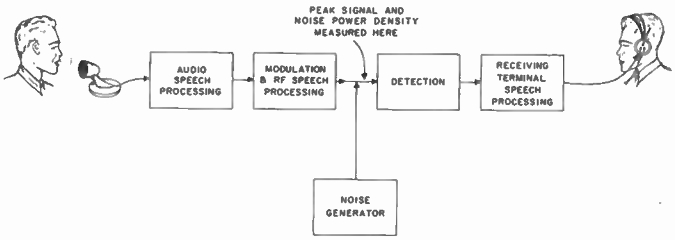


Fig. 1—General block diagram of communication system.

Audio, AM and DSBSC Systems

For audio, amplitude-modulated, and double-sideband suppressed-carrier systems, the required peak instantaneous signal-power-to-noise-power spectral density ratio is given by:

$$\frac{S_p}{dN/df} \text{ in db} = R \text{ in db} = 50 + P + V_A + C_A + W + B + F + A + M \qquad (2)$$

* Received by the PGA, May 31, 1961.

† Research Div., Collins Radio Company, Cedar Rapids, Iowa.

[1] H. Fletcher, "Speech and Hearing in Communications," D. Van Nostrand Co., Inc., Princeton, N. J.; 1953.

[2] S. S. Stevens, Ed., "Handbook of Experimental Psychology," John Wiley and Sons, Inc., New York, N. Y.; 1951.

[3] J. P. Egan, "Articulation Testing Methods II," Psycho-Acoustic Lab., Harvard University, Cambridge, Mass., OSRD Rept. No. 3802, pp. 58–60; November, 1944.

[4] The threshold of perceptibility is that level of intelligibility at which the listener can, with difficulty, determine the gist of connected discourse.

[5] Since such a method of presentation is justified entirely on empirical grounds, great care should be taken in extrapolating results. The interaction of correction factors was less than or equal to twice the standard deviation of the articulation tests.

where the correction factors are associated with the system parameters as follows: P with the amount of pre-emphasis of the speech spectrum employed at the transmitter prior to any nonlinear speech processing,[6] V_A with the amount of audio volume compression at the transmitter prior to any peak clipping,[6] C_A with the amount of hard symmetrical audio peak clipping, W with the equivalent audio band pass of the transmitter *after* the peak clipper,[6] B with the *over-all* transmitter audio band pass, F with the slope of the frequencey response of the over-all *system* including both transmitter and receiver, A with the required word intelligibility, and M with the type of modulation used. The over-all bandwidth of the receiver is always assumed to be equal to the over-all bandwidth of the transmitter. Table I (next page) lists the correction factors and the appropriate graphs for obtaining them.

The amount of pre-emphasis is given in decibels per octave and refers to the slope of the frequency response within the pass band. It will be noted that the slope of the frequency response of the over-all system has no effect on system performance. However, the subjective quality was always judged to be highest when the over-all frequency response was flat (*i.e.*, when any pre-emphasis used at the transmitter is corrected by de-emphasis at the receiver).

The audio volume compressor had an attack time of 0.005 sec and a release time of 0.5 sec. Faster attack and release times will reduce the required peak power and in the limit will approach the performance of a peak clipper. The amount of volume compression and peak clipping are given in terms of the number of decibels that the instantaneous peak signal into the device exceeds the threshold of compression or clipping. In this investigation, the speech level into the system was controlled at a reasonably constant level by the talker. In practice, however, this level will vary considerably from person to person, and a volume compressor will help stabilize this level.

For a DSBSC transmitter, the clipping and volume compression can be performed at RF frequencies. It can be shown, however, that the results obtained are almost the same as if this processing had been performed at audio frequencies.

The power of an AM transmitter is usually stated in terms of the carrier power which is 9 db less than the peak instantaneous power S_p, for 100 per cent modulation, while for a DSBSC system it is usually stated in terms of the peak envelope power which is 3 db less than the peak instantaneous power S_p.

SSB Systems

The same general method of predicting the required peak instantaneous signal power can be used here as in the previous sections. However, the previous results cannot be extended to cover the SSB system, as the SSB envelope is not related to the modulating function in any simple manner. Also, in a SSB system it is desirable to consider the effects of RF volume compression, clipping, and frequency translation. The effects of the various system parameters were again obtained by intelligibility tests and the required ratio of peak instantaneous signal-power-to-noise-power spectral density at the input of the receiver was found to be given by:

$$\frac{S_p}{dN/df} \text{ in db} = R_{SSB} \text{ in db} = 50 + P_S + V_{AS} + C_{AS} + C_{RS} + W_S + B + F + T + A + M \quad (3)$$

where the correction factors are associated with the system parameters as follows: P_S with the amount of pre-emphasis of the speech spectrum employed at the transmitter prior to any nonlinear speech processing,[7] V_{AS} with the amount of audio volume compression at the transmitter prior to any peak clipping, C_{AS} with the amount of hard symmetrical audio peak clipping,[8] V_{RS} with the amount of RF volume compression at the transmitter prior to any RF peak clipping,[9] C_{RS} with amount of hard symmetrical RF peak clipping,[10] W_S with the transmitter pass band *after* all nonlinear processing, B with the *over-all* transmitter band pass, F with the slope of the over-all system frequency response including both transmitter and receiver, T with the amount of frequency translation at the audio output of the receiver caused by receiver detuning, A with the required word intelligibility, and M with the fact that SSB modulation is used. The over-all bandwidth of the receiver is always assumed equal to the over-all bandwidth of the transmitter. Table III (next page) lists the correction factors and the appropriate graphs for obtaining them.

The amount of pre-emphasis, volume compression, and peak clipping are defined as in the previous section. The RF volume compressor used had an attack time of 0.001 sec and a release time of 0.2 sec. The subjective quality of a system employing RF peak clipping was higher than that obtained with a system employing AF peak clipping. Here again, it should be noted that the power of an SSB transmitter is usually specified in terms of the peak envelope power, which is 3 db less than the peak instantaneous power S_p.

[6] The correction factors P, V_A, and W also depend on the amount of audio peak clipping employed as shown in Figs. 2, 3 and 5. Intermediate values may be interpolated.

[7] The correction factor P_S also depends on the amount of AF and RF peak clipping employed as shown in Fig. 9. If both AF and RF clipping are used, the curves can not be interpolated since no data were obtained under such conditions except for 0 db of pre-emphasis.

[8] The correction factor C_{AS} should not be included if RF clipping is also included, as the required correction is included in C_{RS}.

[9] If both AF and RF volume compression are employed, only the larger correction factor V_{AS} or V_{RS} should be included, since they are not additive.

[10] The correction factor C_{RS} also depends on the amount of AF clipping employed as shown in Fig. 13. Interpolation may be used.

TABLE I

CORRECTION FACTORS FOR AUDIO, AM AND DSBSC SYSTEMS

Correction Factor for	Correction Factor Symbol	Correction Factor Reference
1) Pre-emphasis at transmitter prior to any nonlinear processing	P	Fig. 2
2) Audio volume compression	V_A	Fig. 3
3) Audio peak clipping	C_A	Fig. 4
4) Transmitter band pass *after* all nonlinear processing	W	Fig. 5
5) *Over-all* transmitter band pass	B	Fig. 6
6) Over-all *system* frequency response	F	Fig. 7
7) Required word intelligibility	A	Fig. 8
8) Type of modulation used	M	Table II

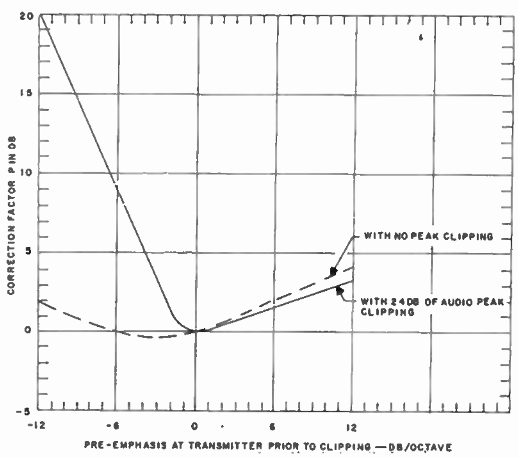


Fig. 2—Pre-emphasis at transmitter prior to any nonlinear processing: audio, AM, and DSBSC systems.

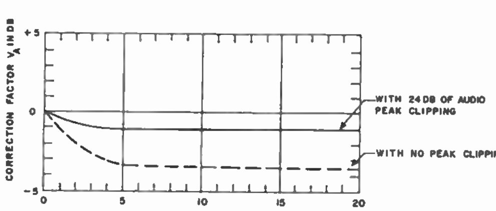


Fig. 3—Audio volume compression, audio, AM, and DSBSC systems.

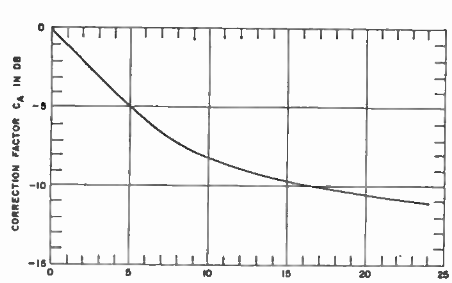


Fig. 4—Audio peak clipping: audio, AM, and DSBSC systems.

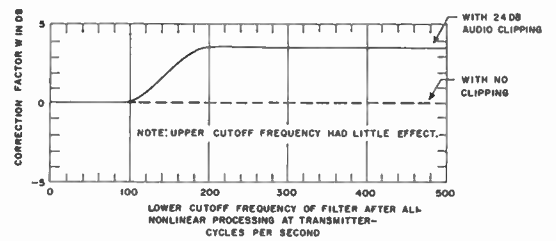


Fig. 5—Transmitter band pass after all nonlinear processing: audio, AM, and DSBSC systems.

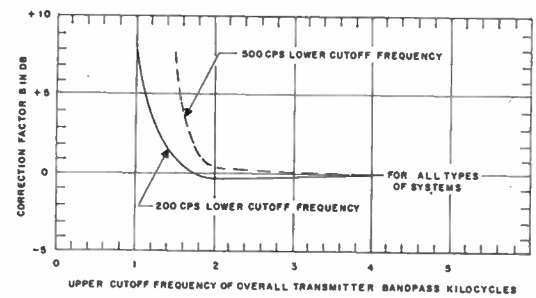


Fig. 6—Over-all transmitter band pass.

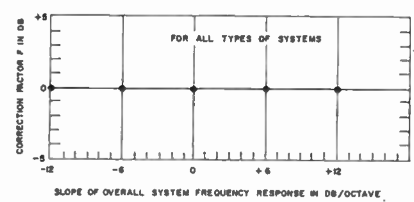


Fig. 7—Over-all system frequency response.

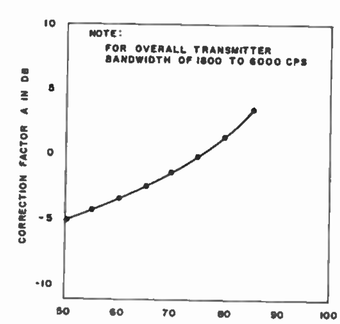


Fig. 8—Required word intelligibility (for W-22 PB word lists of the Central Institute for the Deaf).

TABLE II

CORRECTION FACTOR FOR TYPE OF MODULATION USED

Type of Modulation	Correction Factor in db
1) None (audio system)	0.0
2) Amplitude-modulation	9.0
3) Double-sideband suppressed-carrier	3.0

TABLE III
CORRECTION FACTORS FOR SSB SYSTEMS

Correction Factor for	Correction Factor Symbol	Correction Factor Reference
1) Pre-emphasis at transmitter prior to any nonlinear processing	P_S	Fig. 9
2) Audio volume compression	V_{AS}	Fig. 10
3) Audio peak clipping	C_{AS}	Fig. 11
4) RF envelope compression	V_{RS}	Fig. 12
5) RF envelope clipping	C_{RS}	Fig. 13
6) Transmitter pass band after all nonlinear processing (Effect not investigated. All data given for band pass 300–3100 cps.)	W_S	0 db
7) Over-all transmitter band pass	B	Fig. 6
8) Over-all system frequency response	F	Fig. 7
9) Frequency translation	T	Fig. 14
10) Required word intelligibility	A	Fig. 8
11) Type of modulation (SSB only)	M	2 db

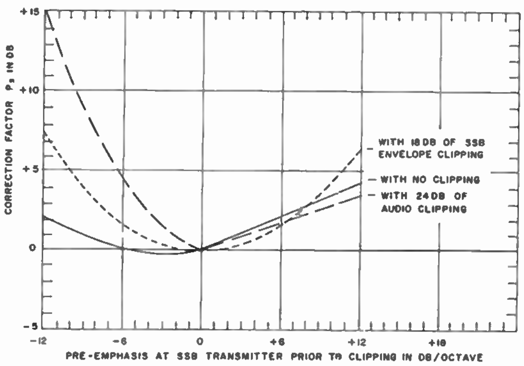


Fig. 9—Pre-emphasis at transmitter prior to any nonlinear processing: SSB systems.

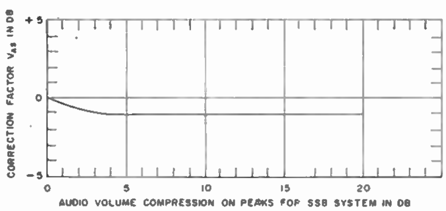


Fig. 10—Audio volume compression: SSB systems.

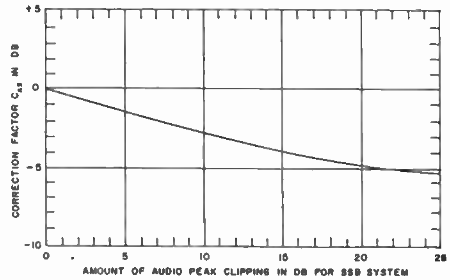


Fig. 11—Audio peak clipping: SSB systems.

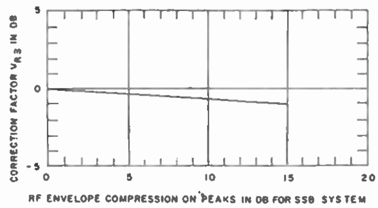


Fig. 12—RF envelope compression: SSB systems.

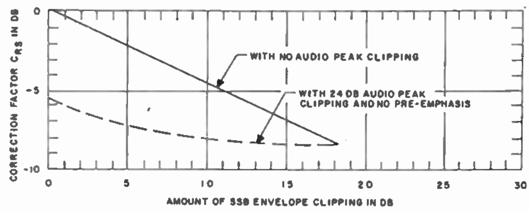


Fig. 13—RF envelope clipping: SSB systems.

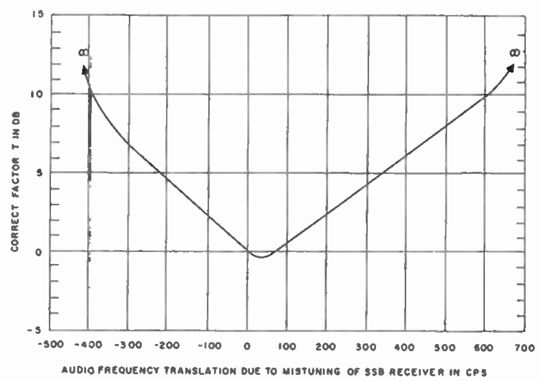


Fig. 14—Frequency translation: SSB systems.

An Example

To illustrate the use of this method of calculating the required transmitter power, consider a SSB system with the following characteristics:

1) A noise-canceling microphone with a frequency response having a slope of +6 db/octave.
2) An audio volume compressor with an attack time of 5 msec, a release time of ½ sec, and a frequency response limited to 500–5000 cps and operating at 10 db of volume compression.
3) An RF volume compressor operating at 10 db of volume compression (1-msec attack time and 0.2-sec release time).
4) An RF clipper operating at 18 db of peak clipping.
5) A filter after the clipper with an equivalent audio band pass of 300–3100 cps.
6) A receiver with an equivalent audio bandwidth of 500–3100 cps and a flat AF response over the bandwidth, but detuned such that the audio spectrum is moved up in frequency by 200 cps.
7) A required word articulation of 60 per cent.

Using (3), the required ratio of peak instantaneous SSB signal power to the noise power spectral density can be calculated as shown in Table IV. The ratio of peak envelope power to noise power spectral density is 3 db less, or 40.9 db. Since the receiver bandwidth is 2600 cps, the total noise in the pass band will be 34.2 db greater than that in a 1-cps bandwidth. The ratio of the peak envelope power to the noise power is therefore 40.9 db, −34.2 db, or +6.7 db.

TABLE IV
Calculation of Required Ratio of the Peak Signal Power to Noise Power Spectral Density

Correction Factor Symbol	Correction Factor in db
Constant from (3)	+50.0
P_S (+6 db/octave with 18-db RF clipping)	+ 1.7
V_{AS} (10 db of audio volume compression)	− 1.0
C_{AS} (no audio clipping)	0.0
V_{RS} (not included[9])	—
C_{RS} (18 db of RF clipping)	− 8.5
W_S (300–3100 cps)	0.0
B (500–3100 cps)	+ 0.2
F (+6 db/octave)	+ 0.0
T (+200-cps frequency translation)	+ 2.5
A (60 per cent word intelligibility)	− 3.0
M (SSB modulation)	+ 2.0
Total $= \dfrac{S_p}{dN/df}$ in db $= R_{SSB}$ in db	+43.9

This method of calculation of the required power has proved very useful in the design of speech communications systems and generally gives an accuracy of ±1 db exclusive of the anomalies associated with the measurement of intelligibility.

Acknowledgment

The authors wish to thank Dr. P. H. Rogers for his advice and encouragement.

Automation of Speech, Speech Synthesis and Synthetic Speech, A Bibliographical Survey from 1950-1960*

PAUL L. SIMMONS†

Summary—This paper presents an enumerative bibliography of the literature on the automation of speech from 1950–1960. The key developments in this area are presented in the introduction. The eventual goal is the fully automated speaking machine. Selections for the bibliography include significant corporate documents, periodical literature and abstracts of research.

RESEARCH in the field of speech automation during the decade of the 1950's has progressed along several lines; first the VOice DEmonstratoR (Voder) which was fathered by H. W. Dudley in 1939 at the New York World's Fair and was a modern automaton with vocal cords of vacuum tubes that produced vowels, consonants and monosyllables. The operator used a keyboard and foot pedal which produced whole sentences of human speech from the mechanism.

Second, the VOice CODER (Vocoder) which had an electrical speech synthesizer similar to the Voder but which made use of control currents from electrically analyzed speech for automatically operating the synthesizer instead of using manual control. The International Dictionary of Physics and Electronics, page 1225, defines it as an instrument which produces synthetic speech, employs recorded voice signals to activate the system instead of the mechanical keys used by the Voder.

Third, the independent research in Great Britain's Post Office Engineering Department, under the direction of Halsey and Swaffield (which is here summarized by A. C. Hales), concludes in its research, in gist, that speech is basically of two types—voiced sounds and hissing; for example, the former might be *or*, *oo* and the latter *ss*. Much of this is redundant in human speech and results in a waste of bandwidth to transmit.

Thus arose the need for a speech compression technique to transmit either natural or synthetic speech. The Air Force at Cambridge Research Center in 1959 held a seminar on Speech Compression and Processing. S. J. Campanella has also surveyed the various compression techniques by type.

H. W. Dudley, in his article on the speaking machine of von Kempelin, traces speech production back to 1791, with the manual fingering of a pneumatic voice producer. The Enclyclopedia Britannica traces the artificial production of vowels, consonants, etc., back to Dr. Eccles, Wheatstone, R. Willis and others, under the topic "Voice Sounds."

In the nineteenth century, various workers built automata which resembled animals and humans, and which played cards and spoke. Finally, it remains for the twentieth century to produce concerted effort towards artificial intelligence, logical machines, reading machines, speaking machines and machines for the automatic recognition of speech, the end product of which will most likely be an automaton which has human-like functions and which can speak, reason, and learn, but which will probably not have a human resemblance.

In this bibliographical survey, an attempt has been made to point out the better known articles on speech synthesis and synthetic speech. It is significant that in this search only one Soviet article was unearthed, by A. A. Pirogov, which is cited here.

All material which seemed pertinent has been given, including abstracts of research, later published as a paper, and corporate documents.

The sources which were searched are listed below:

Applied Sciences and Technology Index, 1958 to June 1960.

Armed Services Technical Information Agency. *Technical Abstract Bulletin.* Arlington, Va. 1957 to 1960.

Bibliographie Linguistique des Annees, 1950 to 1960. Utrecht: Comite International Permanent de Linguistes.

Engineering Index, 1950 to 1959.

Industrial Arts Index, 1950 to 1957.

Proceedings of the Institute of Radio Engineers, 1950 to June 1960. (Section on Acoustics and Audio Frequencies in Abstracts and References).

References to contemporary papers on Acoustics. Published regularly in: *Journal of the Acoustical Society of America*, 1950 to 1960.

Science Abstracts, 1950 to June 1960.

* Received by the PGA, May 1, 1961.
† North American Aviation, Inc., Space and Information Systems Division, Downey, Calif.

Bibliography

Anonymous, Communications, *Aviation Week*, vol. 59, p. 516; August 17, 1953.

Anonymous, "Machine 'talks' to blind and deaf," *Popular Science*, vol. 157; p. 173; September, 1950.

Anonymous, "Symposium on speech analyses and their applications," *J. Acoust. Soc. Jap.*, vol. 14(2), Entire issue, February, 1958. (In Japanese.)

A. S. Abramson, "Vocoder output and whispered speech in a tone language: Thai," *J. Acoust. Soc. Am.*, vol. 31, p. 1568; November, 1959.

A. R. Adolph, "Speech-Pitch Information Extractor," Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, Quarterly Progress Rept. No. 46, pp. 121–130: July 15, 1957.

Air Force, *Proc. Seminar on Speech Compression and Processing*, AF Cambridge Res. Ctr., TR-59-198, L. G. Hanscom Field, Bedford, Mass., September 28–30, 1959. 2 volumes.

G. Arnold, P. Denes, A. Gimson, J. O'Connor and J. Trim, "The synthesis of English vowels," *Language and Speech*, vol. 1, pp. 114–125: February, 1958.

R. E. Bach, Jr., C. R. Howard and S. H. Chang, "Measurement of digital channel capacity for a speech compression system," *J. Acoust. Soc. Am.*, vol. 31, p. 853; June, 1959.

G. M. Badoyannis, "Theoretical and Experimental Research in Communication Theory and Application," Rutgers University, New Brunswick, N. J., Progress Rept. No. 3; 1956. Various pagings.

G. M. Badoyannis, D. Molony, S. Fich, and J. J. Slade, Jr., "Electronically tunable selection amplifier for formant synthesis," *J. Acoust. Soc. Am.*, vol. 28, p. 591; July, 1956.

H. L. Barney, and H. K. Dunn, "Speech synthesis," *In* L. Kaiser, "Manual of Phonetics, Part C: Phonetic Research," North Holland Publishing Company, Amsterdam, 1957.

H. L. Barney, F. E. Haworth and H. K. Dunn, "An experimental transistorized artificial larynx," *Bell Sys. Tech. J.*, vol. 38, pp. 1337–1356; November, 1959.

H. L. Barney, F. E. Haworth, and H. K. Dunn, "New artificial larynx," *J. Acoust. Soc. Am.*, vol. 31, p. 844: June, 1959.

J. Bastian, P. Delattre, and A. M. Liberman, "Silent interval as a cue for the distinction between stops and semivowels in medial position," *J. Acoust. Soc. Am.*, vol. 31, p. 1568; November, 1959.

T. E. Bayston, and S. J. Campanella, "Continuous analysis speech bandwidth compression system," *J. Acoust. Soc. Am.*, vol. 29, p. 1255; November, 1957.

T. E. Bayston, and S. J. Campanella, "Development of a Continuous Analysis Speech Compression System," Melpar Inc., Falls Church, Va., Final Engineering Rept. for July 31, 1957. 1 volume, unpaged.

J. Beck, and E. Schoenhaerl, "A rare kind of substitute voice formation," *Medizinische* (Stuttgart), vol. 38, pp. 1742–1744; September 19, 1959.

C. G. Bell, F. Poza, and K. N. Stevens, "Automatic resolution of speech spectra into elemental spectra," *Proc. Seminar on Speech Compression and Processing*, A. F. Cambridge Res. Ctr., L. G. Hanscom Field, Bedford, Mass., September 28–30, 1959; TR-58-198, paper A-6. 1 volume, 5 pages.

C. G. Bell, F. Poza and K. N. Stevens, "Formant tracking by computer," *J. Acoust. Soc. Am.*, vol. 31, p. 1563; November, 1959.

C. G. Bell, J. M. Heinz, G. Rosen and K. N. Stevens, "Automatic Resolution of Speech Spectra into Elemental Spectra," Res. Lab. of Electronics, M. I. T., Cambridge, Quarterly Progress Rept. No. 54, pp. 161–167; July 15, 1959.

C. G. Bell, H. Fukisaki, J. M. Heinz, A. S. House and K. N. Stevens, "Speech analysis," Res. Lab. of Electronics, M.I.T., Cambridge, Quarterly Progress Rept. No. 57, p. 121; April 15, 1960.

A. R. Billings, "Communication efficiency of Vocoders," *Electronic and Radio Engr.*, vol. 36, pp. 449–453; May, 1959.

A. R. Billings, "Simple multiplex Vocoder," *Electronic and Radio Engr.*, vol. 36, pp. 184–188; May, 1959.

B. P. Bogert, "The Vobanc-voice band compressor," *J. Acoust. Soc. Am.*, vol. 26, p. 137; January, 1954.

J. M. Borst, "Use of spectrograms for speech analysis and synthesis," *J. Audio Engrg. Soc.*, vol. 4, pp. 14–23; January, 1956.

J. M. Borst, and Franklin S. Cooper, "Speech research devices based on a channel Vocoder," *J. Acoust. Soc. Am.*, vol. 29, p. 777; June, 1957.

S. J. Campanella, "A survey of speech band-width compression techniques," IRE Trans. on Audio, vol. AU-6, pp. 104–115: September–October, 1958.

S. J. Campanella, "A digitalized continuous analysis speech compression system," *Proc. Seminar on Speech Compression and Processing*, A.F. Cambridge Res. Ctr., L. G. Hanscom Field, Bedford, Mass., vol. 1, September 28–30, 1959; TR-59-198, paper B-7. 7 pages.

C. Campbell, and J. Murtagh, "The simulated larynx," *Annals of Otology*, vol. 68, pp. 373–392; June, 1959.

S.-H. Chang, "Two schemes of speech compression system," *J. Acoust. Soc. Am.*, vol. 28, pp. 565–572; July, 1956.

S.-H. Chang, H. L. Stubbs, *et al.*, "Visual Message Presentation," Res. Lab. of Electronics, Mass. Inst. Tech., Boston, Scientific Rept. No. 4 for period September 1, 1955 to February 29, 1956; February, 1956. 49 pages.

S.-H. Chang, L. O. Dolansky and C. R. Howard, "A formant tracker for speech sounds," *J. Acoust. Soc. Am.*, vol. 28, p. 161; January, 1956.

S.-H. Chang, R. E. Bach, Jr., C. R. Howard, and R. Sukys, "The Formoder as a tool for speech studies," *Proc. Seminar on Speech Compression and Processing*, A.F. Cambridge Res. Ctr., L. G. Hanscom Field, Bedford, Mass., vol. 1, September 28–30, 1959: TR-59-198, paper B-6. 8 pages.

S.-H. Chang, R. E. Bach, Jr., W. F. King and R. Sukys, "Present status of Formoder," *J. Acoust. Soc. Am.*, vol. 29, p. 1256; November, 1957.

P. Chavasse, "An artificial mouth for acoustic tests," *Compt. Rend. Acad. Sci. (Paris)*, vol. 230, pp. 436–438; January 30, 1950. In French.

E. C. Cherry, "On Human Communication," Technology Press of Mass. Inst. of Tech., Cambridge, 1957.

F. S. Cooper, "Basic factors in speech perception and applications to speech processing," *Proc. Seminar on Speech Compression and Processing*, A.F. Cambridge Res. Ctr., L. G. Hanscom Field, Bedford, Mass., vol. 1, September 28–30, 1959; TR-59-198, paper A-1. 13 pages.

F. S. Cooper, "Spectrum analysis." *J. Acoust. Soc. Am.*, vol. 22, pp. 761–762: November, 1950.

F. S. Cooper, "Some Instrumental Aids to Research on Speech," in "Report of the Fourth Annual Round Table Meeting on Linguistics and Language Teaching," Institute of Languages and Linguistics, Georgetown University, Washington, D. C., pp. 46–53; 1953.

F. S. Cooper, A. M. Liberman and J. M. Borst, "Preliminary studies of speech produced by a pattern playback," *J. Acoust. Soc. Am.*, vol. 22, p. 678; October, 1950.

F. S. Cooper, A. M. Liberman and J. M. Borst, "Synthetic speech: a study of the auditory perception of complex sounds," *Science*, vol. 112, p. 426; October 13, 1950.

F. S. Cooper, E. Peterson, and G. S. Fahringer, "Some sources of characteristic Vocoder equality," *J. Acoust. Soc. Am.*, vol. 29, p. 183; January, 1957.

F. S. Cooper, A. M. Liberman and J. M. Borst, "Interconversion of audible and visible patterns as a basis for research in the perception of speech," *Proc. Natl. Acad. Sci. Wash.*, vol. 37, p. 318; May 25, 1951.

F. S. Cooper, P. C. Delattre, A. M. Liberman, J. M. Borst, and L. J. Gerstman, "Some experiments on the perception of synthetic speech sounds," *J. Acoust. Soc. Am.*, vol. 24; pp. 597–606; November, 1952.

H. J. F. Crabbe, "Electronics and the phonetician: speech analysis and synthesis," *Wireless World*, vol. 65, pp. 289–294; June, 1959.

E. E. David, Jr., "Naturalness and distortion in speech-processing devices," *J. Acoust. Soc. Am.*, vol. 28, pp. 586–589; July, 1956.

E. E. David, Jr., "Signal theory in speech transmission," IRE Trans. on Circuit Theory, vol. CT-3, pp. 232–244; December, 1957.

E. E. David, Jr. and H. S. McDonald, "Note on pitch-synchronous processing of speech," *J. Acoust. Soc. Am.*, vol. 28, pp. 1261–1266; November, 1956.

E. E. David, Jr., M. V. Mathews and H. S. McDonald, "Description and results of experiments with speech using digital computer simulation" *Proc. Natl. Electronics Conf.*, Chicago, Ill., October 13–15, 1958, vol. 14, pp. 766–775, Natl. Electronics Conf., Inc.; March 27, 1959.

E. E. David, Jr., M. V. Mathews and H. S. McDonald, "Digital computer simulation for speech experiments," *J. Acoust. Soc. Am.*, vol. 31, p. 113; January, 1959.

E. E. David, Jr., M. V. Mathews and H. S. McDonald, "A high-speed data translator for computer simulation of speech and television devices," *Proc. Western Joint Computer Conf.*, San Francisco, Calif., pp. 169–172; March 3–5, 1959.

P. Delattre, "The use of the pattern playback in studies of vowel color by synthesis," *J. Acoust. Soc. Am.*, vol. 22, p. 678; October, 1950.

P. Delattre, "Voyelles synthétiques à deux formants et voyelles cardinales," *Maître Phonetique*, vol. 96, pp. 30–36; July–December, 1951. In French.

P. Delattre, F. S. Cooper, A. M. Liberman, and L. Gerstman, "Acoustic loci and transitional cues for consonants," *J. Acoust. Soc. Am.*, vol. 26, p. 137; January, 1954.

P. Denes, "Preliminary investigation of certain aspects of intonation," *J. Acoust. Soc. Am.*, vol. 31, p. 852; June, 1959.

W. C. Dersch, "A voice analog," *J. Acoust. Soc. Am.*, vol. 31, p. 1563; November, 1959.

L. O. Dolansky, "Studio for listening tests," *J. Acoust. Soc. Am.*, vol. 30; pp. 175–181; March, 1958.

H. W. Dudley, "Fundamentals of speech synthesis," Bell Telephone Sys. Tech. Publications, Monograph 2648, Murray Hill, N. J.; 1955. Unpaged.

H. W. Dudley, "Phonetic pattern recognition Vocoder for narrowband speech transmission," *J. Acoust. Soc. Am.*, vol. 30, pp. 733–739; August, 1958.

H. W. Dudley, "Speech analysis and synthesis system," *J. Acoust. Soc. Am.*, vol. 22, p. 410; June, 1950.

H. W. Dudley, "Fundamentals of speech synthesis," *J. Audio Engrg. Soc.*, vol. 3, pp. 170–185; October, 1955.

H. Dudley, and K. H. Davis, "A narrow band speech transmission system employing the recognition of phonetic elements," *J. Acoust. Soc. Am.*, vol. 25, pp. 831–832; July, 1953.

H. Dudley, and T. H. Tarnoczy, "The speaking machine of Wolfgang von Kempelin," *J. Acoust. Soc. Am.*, vol. 22, pp. 151–166; March, 1950.

H. K. Dunn, "The calculation of vowel resonances, and an electrical vocal tract," *J. Acoust. Soc. Am.*, vol. 22, pp. 740–753; July, 1950.

H. K. Dunn, and H. L. Barney, "Artificial speech in phonetics and communications," *J. Speech Res.*, vol. 1, pp. 23–39; March, 1958.

B. H. Edgardh, "The sound-film spectrograph—an instrument for the acoustic analysis of speech and other rapidly changing sounds," *IVA (Stockholm)*, vol. 22, pp. 134–153; May, 1951. (In German.)

M. W. Essigmann, "Visual Message Presentation," Electronics Res. Lab., Northeastern University, Boston, Mass., Scientific Rept. No. 5, for period March 1, through August 31, 1956; September, 1956. 35 pages.

G. Fairbanks, "Systematic research in experimental phonetics: 1. A theory of the speech mechanisms as a servosystem," *J. Speech Disorders*, vol. 19, pp. 133–139; June, 1954.

G. Fairbanks, W. L. Everitt, and R. P. Jaeger, "Method for time frequency compression-expansion of speech," IRE TRANS. ON AUDIO, vol. AU-2, pp. 7–11; January–February, 1954.

R. M. Fano, "The information theory point of view in speech communication," *J. Acoust. Soc. Am.*, vol. 22, pp. 691–696; November, 1950.

C. G. M. Fant, "Status report of speech research at the Royal Institute of Technology, Stockholm, Sweden," *Proc. Seminar on Speech Compression and Processing*, A.F. Cambridge Res. Ctr., L. G. Hanscom Field, Bedford, Mass., vol. 2, September 28–30, 1959; TR-59-198, paper C-6. 12 pages.

G. Fant, "Modern instruments and methods for acoustic study of speech," *Acta Polyt. Scandinavica*, No. 246, vol. 1; 1958. 81 pages.

G. Fant, "On the predictability of formant levels and spectrum envelopes from formant frequencies," in "For Roman Jakobson," Moulton and Company, New York, pp. 109–120; 1956.

G. Fant, "Speech communication research," *IVA (Stockholm)*, vol. 24, pp. 331–337, 1953. (In English.)

J. L. Flanagan, "Automatic extraction of formant frequencies from continuous speech," *J. Acoust. Soc. Am.*, vol. 28, pp. 110–118; January, 1956.

J. L. Flanagan, "Band width and channel capacity necessary to transmit the formant information of speech," *J. Acoust. Soc. Am.*, vol. 28, pp. 592–596; July, 1956.

J. L. Flanagan, "Estimates of the maximum precision necessary in quantizing certain 'dimensions' of vowel sounds," *J. Acoust. Soc. Am.*, vol. 29, pp. 533–534; April, 1957.

J. L. Flanagan, "Extant problems in analysis-synthesis telephony," *J. Acoust. Soc. Am.*, vol. 31, p. 112 January, 1959.

J. L. Flanagan, "Notes on the design of 'terminal-analog' speech synthesizers," *J. Acoust. Soc. Am.*, vol. 29, pp. 306–310; February, 1957.

J. L. Flanagan, "Pitch discrimination for synthetic vowels," *J. Acoust. Soc. Am.*, vol. 29, p. 1255; November, 1957.

J. L. Flanagan, "A resonance-Vocoder and baseband complement: A hybrid system for speech transmission," IRE WESCON CONVENTION RECORD, pt. 7, pp. 5–16; July, 1959.

J. L. FLANAGAN, "A Speech Analyzer for Formant-Coding Compression System," Acoustics Lab., Mass. Inst. Tech., Cambridge, Scientific Rept. No. 4; May, 1955. Various pagings.

J. L. Flanagan, "A resonance-Vocoder and baseband complement: A hybrid system for speech transmission," IRE TRANS. ON AUDIO, vol. AU-8, pp. 95–102; May–June, 1960.

J. L. Flanagan, and A. S. House, "Development and testing of a formant-coding speech compression system," *J. Acoust. Soc. Am.*, vol. 28, pp. 1099–1106; November, 1956.

J. L. Flanagan, and M. G. Saslow, "Pitch discrimination for synthetic vowels," *J. Acoust. Soc. Am.*, vol. 30, pp. 435–442; May, 1958.

J. W. Forgie, and C. D. Forgie, "Results obtained from a vowel recognition computer program," *J. Acoust. Soc. Am.*, vol. 31, p. 844; June, 1959.

A. Fourcin, "Potential dividing function generator for the control of speech synthesis," *J. Acoust. Soc. Am.*, vol. 32, p. 1501; November, 1960.

D. B. Fry, "The experimental study of speech," *Nature* (London), vol. 173; pp. 844–846; May 8, 1954.

D. B. Fry, "Experiments in the perception of stress," *Language and Speech*, vol. 1, pp. 126–152; February, 1958.

O. Fujimura, "Some characteristics of stop consonants," *J. Acoust. Soc. Am.*, vol. 31, p. 1568; November, 1959.

O. Fujimura, "Sound synthesizer with optical control," *J. Acoust. Soc. Am.*, vol. 30, pp. 56–57; January, 1958.

O. Fujimura, T. Ogawa, and S. Hiki, "On the speech sound synthesizer by use of ADP piezo-optics resonators," *J. Acoust. Soc. Jap.*, vol. 14, pp. 129–137; June, 1958.

T. Fukumura, "Some basic properties of quality response to the synthesized complex sound having formant-like spectra," *J. Acoust. Soc. Jap.*, vol. 14, pp. 151–158; June, 1958.

J. A. Fusca, "Speech compression may free channels," *Aviation Week*, vol. 71, pp. 105–109; October 26, 1959.

W. H. George, "Sound recording and 'publication' of acoustical research," *J. Acoust. Soc. Am.*, vol. 31, p. 1030; July, 1959.

L. J. Gerstman, A. M. Liberman, P. C. Delattre, and F. S. Cooper, "Rate and duration of change in formant frequency as cues for the identification of speech sounds," *J. Acoust. Soc. Am.*, vol. 26, p. 952; September, 1954.

L. J. Gerstman, L. Lisker, J. D. O'Connor, A. M. Liberman, P. C. Delattre, and F. S. Cooper, "Initial and intervocalic cues for the perception of liquids and semivowels," *J. Acoust. Soc. Am.*, vol. 29, p. 778; June, 1957.

Great Britain, Post Office Engrg. Dept., "A Summary of Vocoder Research Over the Period 1947 to 1955," by A. C. Hales, *et al.*, Post Office Res. Station, Dollis Hill, London, England, Res. Rept. No. 13961; October 13, 1956. 9 pages, plus unnumbered charts.

M. Halle, and G. W. Hughes, "Vowel Identifier," Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, Quarterly Progress Rept., pp. 109–117; October 15, 1956.

M. Halle and K. N. Stevens, "Analysis by synthesis," *Proc. Seminar on Speech Compression and Processing*, A.F. Cambridge Res. Ctr., L. G. Hanscom Field, Bedford, Mass., vol. 2, September 28–30, 1959; TR-59-198, paper D-7. 4 pages.

M. Halle, and K. N. Stevens, "Speech Communications," Mass. Inst. Tech., Cambridge, Quarterly Progress Rept. No. 56; January 15, 1960. 4 pages.

C. N. Clair, "Factorial analysis of speech perception," *J. Speech Disorders*, vol. 21, pp. 76–87; March, 1956.

C. M. Harris, "A speech synthesizer," *J. Acoust. Soc. Am.*, vol. 25, pp. 970–975; September, 1953.

C. M. Harris, "A study of the building blocks in speech," *J. Acoust. Soc. Am.*, vol. 25, pp. 962–959; September, 1953.

K. S. Harris, "Cues for the identification of the fricatives of American English," *J. Acoust. Soc. Am.*, vol. 26, p. 952; September, 1954.

K. S. Harris, "Some acoustic cues for the fricative consonants," *J. Acoust. Soc. Am.*, vol. 28, pp. 160–161; January, 1956.

K. S. Harris, A. M. Liberman, J. A. Kinney, and H. Lane, "Effect of learning on perception: Discrimination of relative onset time of the formants in certain speech and non-speech patterns," *J. Acoust. Soc. Am.*, vol. 31, p. 1569; November, 1959.

K. S. Harris, A. M. Liberman, H. Hoffman, and B. C. Griffith, "Differential sensitivity to synthetic speech sounds within and between phoneme categories," *J. Acoust. Soc. Am.*, vol. 28, p. 760; July, 1956.

K. S. Harris, H. S. Hoffman, P. C. Delattre and A. M. Liberman, "Effect of the third-formant transitions on the perception of the voice stop consonants," *J. Acoust. Soc. Am.*, vol. 29, p. 777; June, 1957.

K. S. Harris, H. S. Hoffman, A. M. Liberman, P. C. Delattre, and F. S. Cooper, "Effect of third-formant transitions on the perception of voiced stop consonants," *J. Acoust. Soc. Am.*, vol. 20, pp. 122–126; February, 1958.

Z. S. Harris, "Methods in Structural Linguistics," University of Chicago Press, Chicago, Ill.; 1951. 306 pages.

J. M. Heinz, "Model Studies of the Production of Fricative Consonants," Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, Quarterly Progress Rept., pp. 146–149; July 15, 1958.

D. T. Hess, "Terminal Analog Speech Synthesizer," Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, Quarterly Progress Rept. No. 51, pp. 112–114; October 15, 1959.

I. J. Hirsh, and E. G. Reynolds, "The recognition of synthetic and Natural vowels," *J. Acoust. Soc. Am.*, vol. 25, p. 832; July, 1953.
H. S. Hoffman, "Study of some cues in the perception of the voiced stop consonants," *J. Acoust. Soc. Am.*, vol. 30, pp. 1035–1041; November, 1958.
D. T. Hoger, and F. L. Putzrath, "Design and Study of Correlation Instrumentation for Speech Analysis and Synthesis," Radio Corp. of America, Camden, N. J., Scientific Rept. No. 2 (AFCRC-TN-59-566); July 31, 1959. 53 pages.
D. T. Hoger, and F. L. Putzrath, "Study of correlation instrumentation for speech analysis and synthesis," *Proc. Seminar on Speech Compression and Processing*, A.F. Cambridge Res. Ctr., L. G. Hanscom Field, Bedford, Mass., vol. 2, September 28–30, 1959; TR-59-198, paper C-4. 5 pages.
R. A. Houde, "A review of Stromberg-Carlson research program in formant tracking and word recognition," *Proc. Seminar of Speech Compression and Processing*, A.F. Cambridge Res. Ctr., L. G. Hanscom Field, Bedford, Mass., vol.. 1, September 28–30, 1959; TR-59-198, paper B-8. 4 pages.
R. A. Houde, "Speech-Compression-System Study," Stromberg-Carlson Co., Rochester, N. Y., Quarterly Rept. No. 1, July 1, 1959–January 31, 1960. 28 pages.
A. S. House, "A Note on Optimal Vocal Frequency," Acoustics Lab., Mass. Inst. Tech., Cambridge, Scientific Rept. No. 20 (AFCRC-TN-59-171); 1958. 6 pages.
A. S. House, and K. N. Stevens, "Estimation of Formant Band Widths from Measurements of Transient Response of Vocal Tract," *J. Speech Res.*, vol. 1, pp. 309–315; December, 1958.
A. S. House, and K. N. Stevens, "Estimation of Formant Band Widths from Measurements of Transient Response of the Vocal Tract," Acoustics Lab., Mass. Inst. Tech., Cambridge, Scientific Rept. No. 19 (AFCRC-TN-59-107); December, 1958, 7 pages.
C. R. Howard, "Speech Analysis—Synthesis Scheme Using Continuous Parameters," Northeastern University, Boston, Mass., AFCRC-TN-172; November, 1956. 8 pages.
C. R. Howard, S.-H. Chang and M. J. Carrabes, "Analysis and synthesis of formants and moments of speech spectra," *J. Acoust. Soc. Am.*, vol. 28, p. 768: July, 1956.
Hughes Aircraft Co., Los Angeles, Calif., "Twelve-Channel Transistorized Vocoder," Rept. R-114; 1959. 30 pages.
G. W. Hughes, "Identification of speech sounds by means of a digital computer," *J. Acoust. Soc. Am.*, vol. 31, p. 113: January, 1959.
F. Ingemann, "Eight-parameter speech synthesis," *J. Acoust. Soc. Am.*, vol. 32, p. 1501; November, 1960.
F. Ingemann, "Formants as a voicing cue in the perception of [z]," *J. Acoust. Soc. Am.*, vol. 32, p. 1501; November, 1960.
F. Ingemann, "Speech synthesis by rule," *J. Acoust. Soc. Am.*, vol. 29, p. 1255; November, 1957.
G. F. Jacoby, G. J. Duscheck, and F. L. Putzrath, "Design and Study of Correlation Instrumentation for Speech Analysis and Synthesis," Radio Corp. of Amer., Camden, N. J., Scientific Rept. No. 1; March 31, 1959. 84 pages.
K. W. Jarvis, "Vocon—a synthesized control for speech level," *Automat. Elec. Tech. J.*, vol. 3, pp. 55–62; April, 1952.
A. E. Kerby and L. F. Schoene, Jr., "A digital voice data processor," *Proc. Seminar on Speech Compression and Processing*, A.F. Cambridge Res. Ctr., L. G. Hanscom Field, Bedford, Mass., vol. 2, September 28–30, 1959; TR-59-198, paper C-3. 5 pages.
L. G. Kersta, P. D. Bricker and E. E. David, Jr., "Human or machine—a study of voice naturalness," *J. Acoust. Soc. Am.*, vol. 32, p. 1502; November, 1960.
P. J. Klass, "Vocoder increases channels, security," *Aviation Week*, vol. 73, pp. 67–69; July 25, 1960.
H. Koschell, "Die Nachbildung natürlicher Sprachlaute," *Fernmeldetech. Z.* (Berlin), vol. 3, pp. 48–53; February, 1950. (In German.)
L. V. Kriger, "Pitch Extraction for Speech Synthesis with Special Techniques for Use in Digitalized Band-Width Compression Systems," A.F. Cambridge Res. Ctr., Bedford, Mass., TR-59-116; March, 1959. 47 pages.
E. Krocker, "Aufbau und Untersuchung eines Uebertratungssystems für synthetische Sprache," *Technische Hochschule-Wissenschaftliche Zeitschrift* (Dresden), vol. 6, pp. 757–776; 1956–1957. (In German.)
E. Krocker, "Construction and investigation of a transmission system for synthetic speech," *Nachrichtentechnische Zeitschrift*, vol. 7, pp. 553–564; December, 1957. (In German.)
G. Krohm, "Erfahrungen bei der Entwicklung eines Vocoders und Messungen der mit erhaltenen Verständlichkeit," *Z. Angew. Phys.*, vol. 10, pp. 56–65; February, 1958. (In German.)
K. Kuepf-Mueller, and O. Warns, "Sprachsynthese aus Lauten," *Nachrichtentechnische Fachberichte*, vol. 3, pp. 28–31; January, 1956. (In German.)
P. Ladefoged, "The classification of vowels," *Lingua*, vol. 5, pp. 113–127; March, 1956.
W. Lawrence, "The synthesis of speech from signals which have a low information rate," in "Communication Theory," W. Jackson, Ed., Academic Press, New York, N. Y., pp. 460–471; 1953.
W. Lawrence, "Parametric artificial talking device," *J. Acoust. Soc. Am.*, vol. 32, pp. 1500–1501; November, 1950.
W. Lawrence, and L. G. Stead, "The Methods and Purposes of Speech Synthesis," Report of the Colloqium held at Signals Res. and Dev. Establishment, Great Britain, September 8–9, 1955, SRDE Rept. No. 1100; March, 1956. 61 pages.
W. Lawrence, and R. A. Eades, "The Display of Spectral Maxima with a Speech Spectrograph," Signals Res. and Dev. Establishment, Great Britain, SRDE Rept. No. 1087; May, 1955. 6 pages.
A. M. Liberman, "Some results of research on speech perception," *J. Acoust. Soc. Am.*, vol. 29, pp. 117–123; January, 1957.
A. M. Liberman, P. C. Delattre, and F. S. Cooper, "The role of the selected stimulus variables in the perception of unvoiced stop consonants," *Am. J. Psychol.*, vol. 65, pp. 497–516; 1952.
A. M. Liberman, P. C. Delattre, and F. S. Cooper, "Some cues for the distinction between voiced and voiceless stops in initial position," *J. Acoust. Soc. Am.*, vol. 29, p. 1254; November, 1957.
A. M. Liberman, P. C. Delattre, and F. S. Cooper, "Study of one factor in the perception of the unvoiced stop consonants," *J. Acoust. Soc. Am.*, vol. 24, pp. 114–115; January, 1952.
A. M. Liberman, F. Ingemann, L. Lisker, and F. S. Cooper, "Minimal rules for synthesizing speech," *J. Acoust. Soc. Am.*, vol. 31, p. 109; January, 1959.
A. M. Liberman, P. C. Delattre, F. S. Cooper, and L. J. Gerstman, "The role of consonant-vowel transitions in the perception of the stop and nasal consonants," *Psychol. Monogr.*, vol. 68, pp. 1–13; January, 1954.
A. M. Liberman, F. Ingemann, L. Lisker, P. Delattre, and F. S. Cooper, "Minimal rules for synthesizing speech," *J. Acoust. Soc. Am.*, vol. 31, pp. 1490–1499; November, 1959.
J. C. R. Licklider, "Vocal virilization," *J. Acoust. Soc. Am.*, vol. 30, p. 681; July, 1958.
P. Lieberman, "On vowel intonation contours," *J. Acoust. Soc. Am.*, vol. 30, pp. 681–682; July, 1958.
P. Lieberman, "Some aspects of intonation and stress," *Proc. Seminar on Speech Compression and Processing*, A.F. Cambridge Res. Ctr., L. G. Hanscom Field, Bedford, Mass., vol. 1, September 28–30, 1959: TR-59-198, paper A-3. 4 pages.
P. Lieberman, "Vowel Intonation Contours," Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, Quarterly Progress Rept., pp. 156–161; July 15, 1958.
L. Lisker, "Linguistic segments, acoustic segments, and synthetic speech," Linguistic Society of America, *Language*, vol. 33, pp. 370–374; July–September, 1957.
K. Mabuchi, M. Jagi, K. Endo, and J. Oizumi, "On the study of the regions of voiceless consonants by synthesizing," *J. Acoust. Soc. Jap.*, vol. 14, pp. 123–128; June, 1958. (In Japanese.)
P. Marcou, and J. Daguet, "New methods of speech transmission," in "Information Theory," C. Colin, Ed., Academic Press, New York, N. Y., pp. 231–244; 1956.
D. W. Martin, "Thoughts on the imitation of natural sounds," *J. Acoust. Soc. Am.*, vol. 25, pp. 158–159; January, 1953.
Mass. Inst. Tech., Acoustics Lab., Quarterly Rept. for April–June. 1957, J. Doty, Cambridge; July, 1957. 37 pages.
Mass. Inst. Tech., Acoustics Lab., Quarterly Rept. for October–December, 1956, Cambridge; January, 1957. 39 pages.
Mass. Inst. Tech., Acoustics Lab., Quarterly Rept., January–March, 1957, J. M. Heinz and E. R. Rubin, Cambridge, 1957. 25 pages.
Mass. Inst. Tech., Acoustics Lab., "Research on Speech Synthesis," Scientific Rept. No. 17 (AFCRC TN-58-140), Cambridge, March, 1958. 49 pages.
Mass. Inst. Tech., Acoustics Lab., "Speech Compression Research," Final Rept., 1953–1956, Cambridge; February, 1957. 51 pages.
Mass. Inst. Tech., Acoustics Lab., "Speech Compression Research," Status Rept. (AFCRC TN-55-817), Cambridge; February, 1955. 27 pages.
Mass. Inst. Tech., Res. Lab. of Electronics, "Research on speech synthesis," Final Rept., for March 1, 1958–October 31, 1959 (AFCRC TR-60-101), Cambridge, 1960. 15 pages.
M. V. Mathews, "Extremal coding for speech transmission," *J. Acoust. Soc. Am.*, vol. 31, p. 113; January, 1959.
M. V. Mathews, "The effective use of digital simulation for speech processing," *Proc. Seminar on Speech Compression and Processing*, A.F. Cambridge Res. Ctr., L. G. Hanscom Field, Bedford, Mass., vol. 1, September 28–30, 1959; TR-59-198, paper A-4. 6 pages.
W. F. Meeker, C. P. Smith, and W. R. Smith, "An approach to the evaluation of Vocoded speech," *J. Acoust. Soc. Am.*, vol. 32, p. 1502; November, 1960.
W. W. Meeks, J. M. Borst, and F. S. Cooper, "Syllable synthesizer for research on speech," *J. Acoust. Soc. Am.*, vol. 26, p. 137; January, 1954.

Melpar Inc., "Development of a Continuous Analysis Speech Compression System," Final Engineering Report. Falls Church, Va.; July 31, 1957. Unpaged.

W. Meyer-Eppler, "Synthetische Sprache," *Phys. Z.* (Berlin), vol. 29, p. 609; April, 1950. (In German.)

R. L. Miller, "Auditory test with synthetic vowels," *J. Acoust. Soc. Am.*, vol. 23, pp. 632–633; September, 1951.

R. L. Miller, "Auditory tests with synthetic vowels," *J. Acoust. Soc. Am.*, vol. 25, pp. 114–121; January, 1953.

R. L. Miller, "Improvements in the Vocoder," *J. Acoust. Soc. Am.*, vol. 25, p. 832; July, 1953.

I. H. Mol, "Some notes on the Vocoder problem," *Het P. T. T. Bedrijf* (The Hague), vol. 9, pp. 114–118; May, 1957.

W. A. Munson, and H. C. Montgomery, "A speech analyzer and synthesizer," *J. Acoust. Soc. Am.*, vol. 22, p. 678; September, 1950.

K. Nakata, "Synthesis of nasal consonants by a terminal analog synthesizer," *J. Radio Res. Labs.* (Tokyo), vol. 6, pp. 243–254; April, 1959. (In Japanese.)

K. Nakata, "Synthesis of Nasal Consonants by Terminal Analog Synthesizer," Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, Quarterly Progress Rept. 15, pp. 104–107; April, 1958.

K. Nakata, "Synthesis and Perceptions of Nasal Consonants," Acoustics Lab., Mass. Inst. Tech., Cambridge, Scientific Rept. No. 21 (AFCRC TN-58-763); February 19, 1959. 6 pages.

K. Nakata, "Synthesis and perception of nasal consonants," *J. Acoust. Soc. Am.*, vol. 31, pp. 661–666; June, 1959.

K. Nakata, and K. N. Stevens, "Synthesis of Nasal Consonants by Terminal Analog Synthesizer," Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, Quarterly Progress Rept. No. 50, pp. 144–146; July 15, 1958.

Northeastern University, Electronics Res. Lab., "Speech Analysis," AFCRC TR-58-107, Boston, Mass.; February 28, 1958. 35 pages.

Northeastern University, Electronics Res. Lab., "Speech Analysis," Final Rept. for period January 15, 1958–August 31, 1959. (AFCRC TN-59-187), Boston, Mass.; August 31, 1959. 69 pages.

Northeastern University, Electronics Res. Lab., "Speech Analysis," Final Scientific Rept. for period March 1, 1957–January 31, 1958, Boston, Mass.; February 28, 1958. 31 pages.

Northeastern University, Electronics Res. Lab., "Speech Analysis," Scientific Rept. 1, for period January 15, 1958–June 30, 1958, (AFCRC TN-58-181), Boston, Mass. 34 pages.

D. H. Obrecht, "Some acoustic correlates of velarization in Lebanese Arabic," *J. Acoust. Soc. Am.*, vol. 31, p. 1568; November, 1959.

J. Oizumi, and E. Kubo, "Synthesizing speech," *J. Acoust. Soc. Jap.*, vol. 10, pp. 155–158; September, 1954. (In Japanese.)

E. Peterson, and F. S. Cooper, "Peakpicker: A bandwidth compression device," *J. Acoust. Soc. Am.*, vol. 29, p. 777; June, 1957.

G. E. Peterson, "Basic physical systems for communication between two individuals," *J. Speech Disorders*, vol. 18, pp. 116–120; June, 1953.

G. E. Peterson, and D. L. Subrahmanyam, "Evaluation of time frequency scanning for narrow band speech transmission" *J. Acoust. Soc. Am.*, vol. 31, p. 113; January, 1959.

G. E. Peterson, E. Sivertsen, and D. L. Subrahmanyam, "Intelligibility of diphasic speech," *J. Acoust. Soc. Am.*, vol. 28, pp. 404–411; May, 1956.

G. E. Peterson, W. S.-Y. Wang, and E. Sivertsen, "Segmentation techniques in speech synthesis," *J. Acoust. Soc. Am.*, vol. 30, pp. 682–683; July, 1958.

G. E. Peterson, W. S.-Y. Wang, and E. Sivertsen, "Segmentation techniques in speech synthesis," *J. Acoust. Soc. Am.*, vol. 30, pp. 739–742; August, 1958.

G. E. Peterson, W. S.-Y. Wang, and E. Sivertsen, Rept. No. 1, Parts 1 and 2. University of Michigan, Ann Arbor; May, 1958. 28 pages.

A. A. Pirogov, "A harmonic system for compressing speech spectra," *Electronics Express*, vol. 1, pp. 20–23; May, 1959.

A. A. Pirogov, "Harmonic systems for speech band-width compression," *Elektrosvyaz*, vol. 13, pp. 8–17; March, 1959. (In Russian.)

C. W. Poppe, and P. J. Suhr, "An automatic voice read-out system," *Proc. Eastern Joint Computer Conf.*, Washington, D. C., December 9–13, 1957; pp. 218–222.

R. K. Potter, and J. C. Steinberg, "Toward the specifications of speech," *J. Acoust. Soc. Am.*, vol. 22, pp. 807–820; November, 1950.

F. Poza, "Formant Tracking," Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, Quarterly Progress Rept., pp. 150–151; October 15, 1959.

V. A. Purto, "Investigation of the principles and development of devices for oral output of information," in U. S. Joint Publication Res. Service, Soviet Developments in Information Processing and Machine Translation, Abstracts of the Conference on Mathematical Linguistics, Leningrad, U.S.S.R., April 15–21, 1959, Series 893-D; August 31, 1959. Various pagings, in English.

M. Rajkovic, "Speaking machines, phonetographs," *Tesla*, vol. 6, pp. 10–15; January–February, 1959. [In Yugoslavian.]

H. Rindfleisch, "Ein einfaches Verfahren zur Erzeugung von synthetischer Sprache," *Technische Hausmitteilungen des Nordwestdeutschen Rundfunks*, vol. 3, pp. 176–177; 1951. (In German.)

G. Rosen, "Dynamic analog speech synthesizer," *J. Acoust Soc. Am.*, vol. 30, pp. 201–209; March, 1958.

G. Rosen, "Dynamic Analog Speech Synthesizer," Acoustics Lab., Mass. Inst. Tech., Cambridge; October 1, 1957. 9 pages.

G. Rosen, "Dynamic Analog Speech Synthesizer," Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, Quarterly Progress Rept., pp. 102–104; April 15, 1958.

G. Rosen, "Dynamic Analog Speech Synthesizer," Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, Quarterly Progress Rept. No. 52, pp. 142–144; January 15, 1959.

G. Rosen, "Dynamic Analog Speech Synthesizer," Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, Quarterly Progress Rept. No. 53, pp. 171–174; April 15, 1959.

G. Rosen, "Experiments with a dynamic analog of the vocal tract," *Proc. Seminar on Speech Compression and Processing*, A.F. Cambridge Res. Ctr., L. G. Hanscom Field, Bedford, Mass., vol. 1, September 28–30, 1959; TR-59-198, paper A-5. 10 pages.

G. Rosen, "A Prototype Section for a Dynamic Speech Synthesizer," Mass. Inst. Tech., Cambridge, Scientific Rept. No. 1 (AFCRC TN-55-196); February, 1955. 37 pages.

G. Rosen, "Fricative production by a dynamic analog synthesizer," *J. Acoust. Soc. Amer.*, vol. 31, p. 1564; November, 1959.

G. Rosen, and J. M. Heinz, "Analog Speech Synthesizer," Acoustics Lab., Mass. Inst. Tech., Cambridge, Quarterly Rept., for April–June, 1956, p. 24; July, 1956.

G. Rosen, and K. N. Stevens, "Apparatus for the generation of psycho-acoustic stimuli varying in complexity from noise bursts to syllables," *J. Acoust. Soc. Am.*, vol. 31, p. 114; January, 1959.

G. Rosen, K. N. Stevens, and J. M. Heinz, "Dynamic analog of the vocal tract," *J. Acoust. Soc. Am.*, vol. 28, p. 767; July, 1956.

H. Rubin, "Simulation of a Vocoder," *Proc. Seminar on Speech Compression and Processing*, A.F. Cambridge Res. Ctr., L. G. Hanscom Field, Bedford, Mass., vol. 1, September 28–30, 1959; TR-59-198, paper B-4. 9 pages.

M. R. Schroeder, "New approach to time domain analysis and synthesis of speech," *J. Acoust. Soc. Am.*, vol. 31, pp. 852–853; June, 1959.

M. R. Schroeder, "Vocoders for military use," *Proc. Seminar on Speech Compression and Processing*, A.F. Cambridge Res. Ctr., L. G. Hanscom Field, Bedford, Mass., vol. 1, September 28–30, 1959; TR-59-198, paper B-2. 7 pages.

M. R. Schroeder, and E. E. David, Jr., "Vocoder and its application to the transmission of high-quality speech over narrow-band channels," *J. Acoust. Soc. Am.*, vol. 31, p. 113; January, 1959.

M. R. Schroeder, and E. E. David, Jr., "A Vocoder for transmitting 10 kc/s channel," *Acustica*, vol. 10, pp. 35–43; 1960.

M. R. Schroeder, and T. H. Crystal, "Auto-correlation Vocoder," *J. Acoust. Soc. Am.*, vol. 32, p. 913; July, 1960.

G. N. Sholes, "Synthesis of final [z] without voicing," *J. Acoust. Soc. Am.*, vol. 31, p. 1568; November, 1959.

Signals Research and Development Establishment. Technical Summary No. 24, SRDE speech band-width compression project, Great Britain; January 20, 1958. Various pagings.

E. Sivertsen, "Segment inventories for speech synthesis," *J. Acoust. Soc. Am.*, vol. 32, p. 919; July, 1960.

F. H. Slaymaker, "Bandwidth compression by means of Vocoders," IRE TRANS. ON AUDIO, vol. AU-8, pp. 20–25; January–February, 1960.

F. H. Slaymaker, "Speech band-width compression with Vocoders," 1959 IRE NATIONAL CONVENTION RECORD, pt. 7, pp. 41; July, 1959.

C. P. Smith, "The analysis and automatic recognition of speech sounds," *Electronic Engrg.*, vol. 24, pp. 368–372; August, 1953.

C. P. Smith, "An approach to speech bandwidth compression," *Proc. Seminar on Speech Compression and Processing*, A.F. Cambridge Res. Ctr., L. G. Hanscom Field, Bedford, Mass., vol. 2, September 28–30, 1959; TR-59-198, paper C-2. 7 pages.

C. P. Smith, "A Method for Speech Data Processing by Means of a Digital Computer," A.F. Cambridge Res. Ctr., Cambridge, Mass., Rept. ERD-CRRB-TM-5-8-103; 1958. Various pagings.

C. P. Smith, "Modifications Proposed for Improvement of Vocoder Performance," A.F. Cambridge Res. Ctr., Cambridge, Mass., Report ERD-CRRC-TM-58-110; June, 1958. Various pagings.

C. P. Smith, "A phoneme detector," *J. Acoust. Soc. Am.*, vol. 23, pp. 446–451; July, 1951.

C. P. Smith, "Speech Data Reduction: Voice Communication by Means of Binary Signals at Rates under 1000 Bits per Second," A.F. Cambridge Res. Ctr., Electronic Res. Dir., Communications Lab., Bedford, Mass., TR-57-111; May, 1957. 116 pages.

C. P. Smith, "Studies of a method for digital voice communication," *J. Acoust. Soc. Am.*, vol. 31, p. 1563; November, 1959.

C. P. Smith, "The Use of Digital Computers in Speech Analysis and Synthesis," A.F. Cambridge Res. Ctr., Bedford Mass., TR-59-958; November, 1959. 13 pages.

C. P. Smith, "Report on visit to speech research groups in Britain," A.F. Cambridge Res. Ctr., Bedford, Mass., Rept. ERD-CRRS-TM-58-106; April, 1958. Various pagings.

L. G. Stead and E. T. Jones, "The SRDE speech bandwidth compression project," *Proc. Seminar on Speech Compression and Processing*, A.F. Cambridge Res. Ctr., L. G. Hanscom Field, Bedford, Mass., vol. 1, September 28–30, 1959; TR-59-198, paper B-5. 12 pages.

J. C. Steinberg, and W. A. Munson, "Speech and music," in "Electrical Engineers' Handbook," John Wiley and Sons, Inc., New York, N. Y.; pp. 12-19 to 12-29, 1950.

K. N. Stevens, "Frequency discrimination for damped waves," *J. Acoust. Soc. Am.*, vol. 24, pp. 76–79; January, 1952.

K. N. Stevens, "Effect of duration upon vowel identification," *J. Acoust. Soc. Am.*, vol. 31, p. 109; January, 1959.

K. N. Stevens, "The Role of Duration in Vowel Identification," Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, Quarterly Progress Rept., pp. 136–139; January 15, 1959.

K. N. Stevens, "Synthesis of speech by electrical analog devices," *J. Audio Engrg. Soc.*, vol. 4, pp. 2–8; 1956.

K. N. Stevens, "Synthesis of Speech Sounds by an Electromechanical Vocal Cord," in Monograph Series on Language and Linguistics, Georgetown University, Washington, D. C., vol. 4, pp. 53–58; 1953.

K. N. Stevens, "Toward a model for speech recognition," *J. Acoust. Soc. Am.*, vol. 32, pp. 47–55; January, 1960.

K. N. Stevens, and K. Nakata, "Synthesis and Perception of Fricative Consonants," Res. Lab. of Electronics, Mass. Inst. Tech., Cambridge, Quarterly Progress Rept. No. 50, pp. 150–152; July 15, 1958.

K. N. Stevens, and R. P. Bastide, "Electrical synthesizer of continuous speech," *J. Acoust. Soc. Am.*, vol. 27, p. 207; January, 1955.

K. N. Stevens, C. G. Bell, P. T. Brady, J. M. Heinz, and G. Rosen, "Automatic resolution of speech spectra into elemental spectra," *J. Acoust. Soc. Am.*, vol. 3, p. 844; June, 1959.

K. C. H. Stockhoff, "A twelve-channel transistorized Vocoder," *Proc. Seminar on Speech Compression and Processing*, A.F. Cambridge Res. Ctr., L. G. Hanscom Field, Bedford, Mass., vol. 1, September 28–30, 1959; TR-59-198, paper B-3. 4 pages.

A. J. Strassman, and K. C. Stockhoff, "Military applications for speech compression techniques," *Proc. Fourth National Convention on Military Electronics*, Washington, D. C., June 27–29, 1960; pp. 75–81.

P. Strevens, "The performance of Lawrence's six-parameter speech synthesizer," *Revista do Laboratoria Fonetica Experimentia Coimbra*; 1958. Unpaged; in Spanish.

R. Sukys, R. E. Bach, Jr., and S. H. Chang, "Supplementing the Formoder with fixed Vocoder-like bands," *J. Acoust. Soc. Am.*, vol. 31, pp. 113–114; January, 1959.

H. Sund, "A sound spectrometer for speech analysis," Ingeniors Vetenskaps Akademien (Sweden), Electrical Engineering Series, *Acta Polyt.* 228, vol. 8, No. 4; 1957. 66 pages.

J. Swaffield, "Some progress with Vocoder-type systems," *Proc. Seminar on Speech Compression and Processing*, A.F. Cambridge Res. Ctr., L. G. Hanscom Field, Bedford, Mass., vol. 1, September 28–30, 1959; TR-59-198, paper A-7. 6 pages.

R. V. Tait, "A dental appliance for the production of artificial voice," *Proc. Roy. Soc. Med.* (London), vol. 52, pp. 747–749; September, 1959.

F. Vilbig, "Improvement and simplification of the Scanvocoder and its connection to a correlation pulse code system," *J. Acoust. Soc. Am.*, vol. 28, p. 768; July, 1956.

F. Vilbig, "Method for measuring fine details of differences in the structure of various speech sound spectrums," A.F. Cambridge Res. Ctr., Communications Lab., Bedford, Mass., TR-56-122; November, 1956. 10 pages.

F. Vilbig, "Method for measuring fine details of differences in the structure of various speech sound spectrums," *J. Acoust. Soc. Am.*, vol. 29, p. 183; January, 1957.

F. Vilbig, "Speech compression," *J. Acoust. Soc. Am.*, vol. 28, p. 160; January, 1956.

F. Vilbig, and K. H. Haase, "Some systems for speech-band compression," *J. Acoust. Soc. Am.*, vol. 28, pp. 573–577; July, 1956.

J. Von Redlich, and R. W. Wendahl, "On the relative loudness of synthesized vowels," *J. Acoust. Soc. Am.*, vol. 32, p. 919; July, 1960.

W. S.-Y. Wang, and G. E. Peterson, "Segment inventory for speech synthesis," *J. Acoust. Soc. Am.*, vol. 30, p. 683; July, 1958.

W. S.-Y. Wang, and G. E. Peterson, "Segment inventory for speech synthesis," *J. Acoust. Soc. Am.*, vol. 30, pp. 743–746; August, 1958.

O. Warns, "The intelligibility of synthetic speech," *Frequenez*, vol. 11; pp. 169–175; June, 1957. (In German.)

O. Warns, "Über die synthese von sprache aus lautin und lautkembinationen," Dissertation, Technische Hochschule, Darmstadt; 1957. (In German.)

W. Wathen-Dunn, "The speech research program of the Air Force Cambridge Research Center," *Proc. Seminar on Speech Compression and Processing*, A.F. Cambridge Res. Ctr., L. G. Hanscom Field, Bedford, Mass., vol. 2, September 28–30, 1959; TR-59-198, paper C-1. 6 pages.

E. S. Weibel, "Vowel synthesis by means of resonant circuits," *J. Acoust. Soc. Am.*, vol. 27, pp. 858–865; September, 1955.

E. S. Weibel, "Vowel synthesis using resonant circuits to represent formants," *J. Acoust. Soc. Am.*, vol. 26, p. 952; September, 1954.

M. Weinstock, "The Army and speech compression," *Proc. Seminar on Speech Compression and Processing*, A.F. Cambridge Res. Ctr., L. G. Hanscom Field, Bedford, Mass., vol. 2, September 28–30, 1959; TR-59-198, paper C-8. 7 pages.

R. W. Wendahl, A. L. Case, and R. W. Stream, "Effect of duration upon the intelligibility of synthetic vowels," *J. Acoust. Soc. Am.*, vol. 32, pp. 919–920; July, 1960.

F. Winckel, "Analyse und Synthese der Sprachlaute," *Funk U. Ton* (Berlin), vol. 8, pp. 50–54, 99–102; 1954. (In German.)

F. Winckel, "Die menschlische stimme in elektrischen modellvorstellungen," *Elektro-Tech. Z.*, vol. 73A, pp. 708–711; November 11, 1952. (In German.)

F. Winckel, "Elektrische Abildung von Sprachlauten," *Elektron*, vol. 5, p. 185; 1951. (In German.)

J. Wiren, "Identification of turbulent speech sounds," *J. Acoust. Soc. Am.*, vol. 29, p. 1255; November, 1957.

S. L. Witryal, and W. A. Kaess, "Individidual Differences in the Perception of Synthetic Speech," Connecticut University, Conn.; December 31, 1956. 12 pages.

S. L. Witryal, and W. A. Kaess, "Individual differences in perception of synthetic speech," *Percept. Mot. Skills*, vol. 8, pp. 215–218; 1958.

A Dual Channel Transistor Power Amplifier*

ALEXANDER B. BERESKIN†, FELLOW, IRE

Summary—**This paper describes the theory and performance of a dual channel transistor power amplifier capable of developing 30 watts in each channel with less than 1% intermodulation distortion. Full power is developed at frequencies below 20 cycles per second. Detailed transformer construction and amplifier assembly information is presented along with a list of parts and purchasing information.**

THE BASIC circuit of the amplifier[1,2] described in this paper was presented at the 1957 IRE National Convention. It has been redesigned to take advantage of transistor improvements and price reductions that have been made during the intervening years. In this paper it is presented with a power capability of 30 watts in each channel. The amplifier can be built as a single channel amplifier or as a double channel amplifier one channel at a time. In either case, the two channel power transformer, which is the only element common to the two channels, should be used to facilitate future conversion to two channel operation.

The basic power amplifier circuit shown in Fig. 1 has three common emitter stages and is capable of accepting in excess of 25 db of over-all voltage feedback. In this circuit a low level *n-p-n* transistor TR1 is direct coupled to the driver *p-n-p* transistor TR2. The driver transistor is in turn coupled through the driver transformer *DT* to the output stage *p-n-p* transistors TR3 and TR4. An output autotransformer *OT* is used to match the load R_L to the output transistors. With this arrangement the collector of TR4 is at approximately the same dc potential as the emitter of TR1, and a direct coupled over-all feedback network, consisting of a resistor R_{fb} and a capacitor C_{fb} in parallel, is connected between them. The resistor controls the low and middle frequency overall feedback while the capacitor controls the high-frequency over-all feed-back. Obviously the same general arrangement could be used if TR1 were a *p-n-p* transistor and TR2, TR3, and TR4 were *n-p-n* transistors.

The 10 K resistor R_1 completes an internal negative feedback loop that involves TR1, TR2, and the driver transformer. The input must be increased by roughly 10 db to compensate for the effect of this resistor. The two resistors R_2 also provide negative feedback in this loop, but the capacitor C_2 limits their effect to the very low-frequency range. One of these resistors R_2 also appears

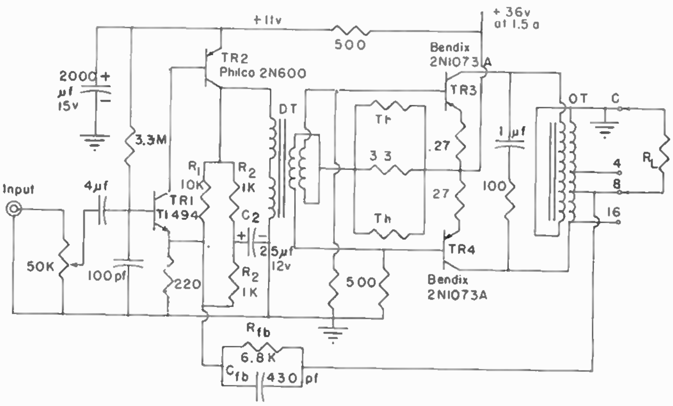


Fig. 1—Power amplifier circuit.

as a damping resistor across the primary of the driver transformer at higher frequencies.

Conduction transfer occurs in both the input and output of the output stage, and in order to avoid conduction transfer notches,[3] the secondaries of the driver transformer *DT* and the two windings of the output autotransformer *OT* are bifilarly wound.

The driver transformer is one of the major reactive components in the feedback loop and it is desirable that it have low primary-secondary leakage reactance. The low primary-secondary leakage reactance is attained by winding the primary in two equal sections with a bifilar secondary sandwiched between them. The primary of the driver transformer carries dc current and this must be taken into account in its design.

The output transformer winding consists of a relatively small number of turns of parallel wires suitably insulated from each other. In general just two parallel copper wires with Heavy Formvar insulation are adequate. The ground return is used as the common output terminal and taps are brought out on one section of the bifilar winding to match 4, 8, and 16 ohm loads. The laminations in this transformer should be fully interleaved.

The operating points of TR1 and TR2 are controlled by the resistor R_B in the base circuit of TR1. This resistor should be chosen to produce 11 volts between the emitter of TR2 and ground when the supply voltage is 36 volts. This automatically sets the collector current of TR2 at about 50 ma and this is adequate to drive the output transistors to saturation. Normal operation produces very little temperature rise in TR1 and TR2

* Received by the PGA, August 21, 1961; revised manuscript received September 12, 1961. This material also appeared in the IRE STUDENT QUARTERLY, September, 1961.

† University of Cincinnati, Cincinnati, Ohio.

[1] A. B. Bereskin, "A high power high quality transistor audio power amplifier," 1957 IRE NATIONAL CONVENTION RECORD, pt. 7, pp. 149–161.

[2] A. B. Bereskin, U. S. Patent No. 2,932,800; April 12, 1960; assigned to the Baldwin Piano Company, Cincinnati, Ohio.

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

so that stabilization beyond the effect of the 500 ohm dropping resistor is not necessary.

The output transistors TR3 and TR4 operate over a relatively wide temperature range and require bias stabilization. The bias is provided by the voltage developed across the 3.3 ohm base return resistor and the dc resistance of the driver transformer secondary windings by the current flowing through the two 500 ohm resistors connected between the two bases and ground. Temperature compensation is provided by the two thermistors *Th* connected in parallel with the 3.3 ohm base return resistor. These thermistors have a resistance of 5 ohms at 25°C and a −3.9 per cent/°C temperature coefficient of resistance. They are thermally coupled to the collectors of TR3 and TR4. Additional bias stabilization is obtained from the two 0.26 ohm resistors in series with the emitters of TR3 and TR4 at the expense of approximately 6 db decrease in the loop gain.

The 100 pf capacitor connected between the base of TR1 and ground provides stabilization for frequencies in excess of 100 kc. The series combination of 0.1 μf and 100 ohms connected between the collectors of the two output transistors makes the amplifier open circuit stable. With some loudspeaker loads this combination may be removed, while with others it is essential in order to keep the system from oscillating at frequencies in excess of 100 kc.

The power supply used with this amplifier is shown in Fig. 2. It consists of one power transformer *PT* with two single phase full wave bridge rectifiers. Simple choke input filters are used. Fuses are provided in the ac line and also in the individual dc channels. Pilot lights have been incorporated in the ac and dc sections of the power supply.

The low level frequency response characteristics in in Fig. 3 were obtained in each case with the input potentiometer set for one half the output, at 1 kc, that resulted when it was set at its maximum position. This is not necessarily the same position in each case since the shunting effect of the amplifier input varies with the amount and type of feedback used. The value of 0.774 volts was chosen as the 0 db reference because it corresponds to 1 mw at 600 ohms and is the reference used on the scale of the VTVM.

Data for curve *A* was obtained with all local and over-all feedback removed except for the 0.27 ohm emitter resistors in the output stage. In this case response is within 3 db of the 1 kc value between 100 cycles and 15 kc. At the low end the response is dropping off at 9 db/octave while at the high end it is dropping off at about 13 db/octave.

Adding the 10 K resistor R_1 changed the response to that of curve *B*. The input signal required was 11 db higher than in the previous case. The low frequency end of this curve has the same shape as the previous curve but has been displaced toward the left by about $1\frac{1}{2}$ octaves. The high frequency end of the curve remains

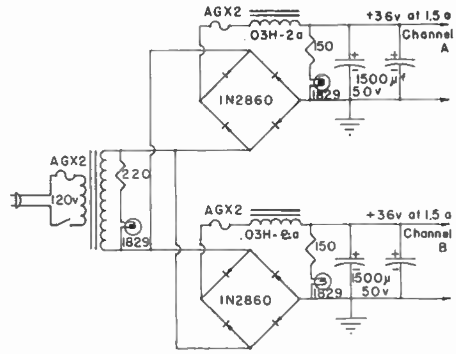
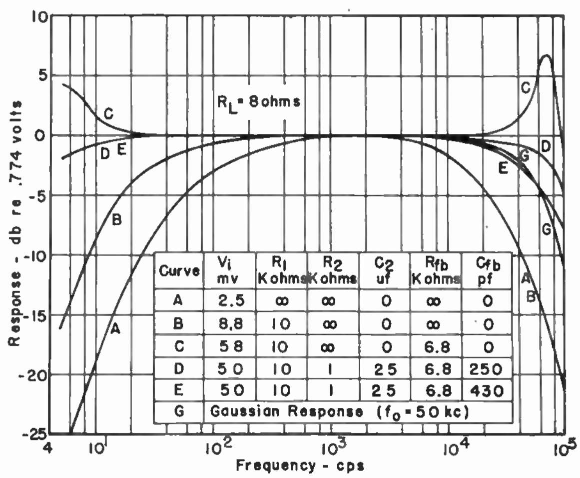


Fig. 2—Power supply.



Curve	V_i mv	R_1 Kohms	R_2 Kohms	C_2 uf	R_{fb} Kohms	C_{fb} pf
A	2.5	∞	∞	0	∞	0
B	8.8	10	∞	0	∞	0
C	58	10	∞	0	6.8	0
D	50	10	1	25	6.8	250
E	50	10	1	25	6.8	430
G	Gaussian Response (f_0 = 50 kc)					

Fig. 3—Low level frequency response.

unchanged showing that the drop off in high frequency response, up to 100 kc, is due almost entirely to the output stage which is not affected by R_1. The response is now within 3 db of the 1 kc value between 25 cycles and 15 kc.

The further addition of R_{fb} = 6.8 K between the 8 ohm output tap and the emitter of TR1 results in the response shown by curve *C*. The input had to be increased by an additional 16.5 db to obtain this curve. The response is now essentially flat between 20 cycles and 20 kc, but there are undesirable humps at both the low- and high-frequency ends. These could result in both low and high-frequency instability.

The low-frequency response was corrected by inserting resistors R_2 and capacitor C_2 to reduce the low-frequency loop gain. The high-frequency response was corrected by using a 250 pf capacitor for C_{fb}. This combination gave rise to curve *D*. The shunting effect of R_2 across the primary of *DT* reduced the middle frequency loop gain so that the input signal required was now only 15 db above that of curve *B*.

Uniform frequency response at high frequencies is not as important as good transient response. To insure good transient response, C_{fb} was changed to 430 pf because it was determined experimentally to be the value necessary just to eliminate ringing with a 5 kc square wave. The resulting response characteristic is shown as curve E. Good transient response is insured if the frequency response is Gaussian

$$\left(A = A_0 e^{-.346\left[\frac{f}{f_0}\right]^2}\right)$$

and to see how closely the experimental curve approached this response a curve for Gaussian response with $f_0 = 50$ kc has been drawn in as curve G. Within the limitations of its shape, curve E appears to be as good an approximation to curve G as could be expected. Response characteristic data was also obtained for 20 per cent and 80 per cent settings of the input potentiometer. For these two cases the low frequency response was identical to curve E and the high frequency response was slightly higher than curve E, being everywhere within $\frac{1}{2}$ db of curve G.

The 5 kc square wave response of the amplifier is shown in Figs. 4(a)–(c). Fig. 4(a) is for 20 volts, 2 volts, and 0.2 volts peak to peak signal across an 8 ohm resistive load. Slight dissimilarities are evident between the leading and trailing edges of the square wave. The dissimilarities are due to the fact that, since class B operation is employed, only one output transistor is working at a time. There are bound to be differences in the frequency response of individual transistors so that the best value of C_{fb} for one transistor would not necessarily be best for the other unit. The value of C_{fb} used is therefore a compromise between the best leading and trailing edges of the 5 kc square wave. Fig. 4(b) is for an RCA 515S2 loudspeaker, housed in a 10 cubic foot base reflex cabinet, connected to the 16 ohm tap of the amplifier. Fig. 4(c) is for a GE S1201D-7 loudspeaker connected to the 8 ohm tap of the amplifier. This speaker was housed in a simple wall baffle. The RCA 515S2 is a compound speaker having low and high frequency units with a very simple crossover network. With this speaker, the amplifier is stable without the 0.1 μf-100 ohm combination connected between the two collectors. The GE S1201D-7 speaker is a single cone unit and requires the 0.1 μf-100 ohm combination for high-frequency stability. If the 0.1 μf-100 ohm combination can be removed, additional high frequency power is available for the load. All tests were made with this combination connected.

The harmonic distortion of the amplifier is shown in Fig. 5 for various frequencies and various values of power. The apparent increase in distortion at very low power levels is part fact and part fiction. The fictional part is due to the fact that the harmonic distortion meter cannot distinguish between residual hum and noise and the harmonic distortion due to the application of a signal. The residual hum and noise is 74 db below 32 watts. For the 0.5 watt output condition this corresponds to 0.16 per cent which must be combined with

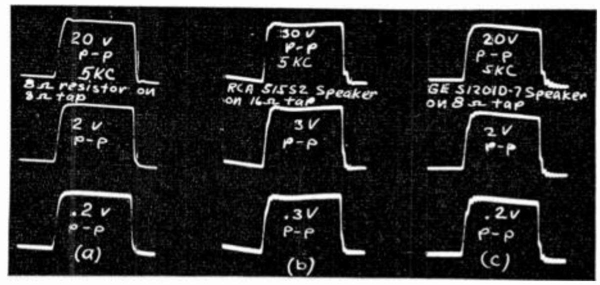


Fig. 4—5-kc square wave response.

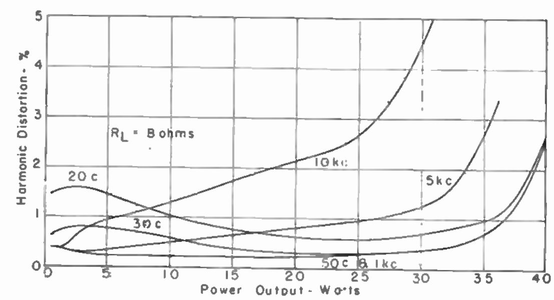


Fig. 5—Harmonic distortion characteristics.

the 0.12 per cent distortion in the signal generator. The measured distortion for this condition is 0.40 per cent for frequencies of 50 cycles and above. The factual part of the distortion increase can be attributed to the fact that for low signal levels, with class B operation, the curved lower portion of the transfer characteristic is used. This results in both higher inherent distortion and lower loop gain. The lower loop gain in turn reduces the effectiveness of the over-all feedback. In any case the distortion is low enough to classify this as a good 30 watt amplifier.

The curves of intermodulation distortion, shown in Fig. 6 for 80 per cent 70 cycle signal and 20 per cent 5 kc signal, confirm the fact this is an excellent 30 watt amplifier. It should be noted that the intermodulation distortion is plotted as a per cent of the smaller of the two signals.

The output impedance of the amplifier is shown as a function of frequency in Fig. 7. Over most of the frequency range the output impedance is roughly one fourth of the nominal tap impedance. As explained in a former paper,[1] the nominal tap impedance is not determined by maximum power considerations but by power transistor dissipation, the power capability of the driver transistor, and the rectifier supply current and voltage rating. Load impedances other than the nominal values may be used on any of the taps within the other limitations specified above.

The feedback circuit elements were worked out empirically for one channel, and the same nominal values were then applied to the other channel. The only exception was the value of C_{fb} which was carefully adjusted for best 5 kc square wave response. The value of C_{fb} was

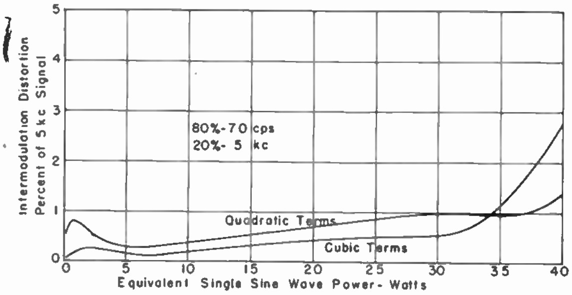


Fig. 6—Intermodulation distortion.

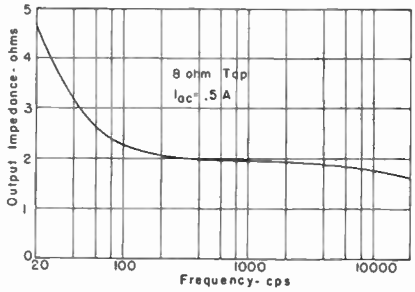


Fig. 7—Output impedance characteristic.

20 pf higher for the second channel than for the first one. A spot check was run on all performance characteristics with substantial agreement between the two channels.

Thermal Design

In order to achieve good electrical performance with a transistor amplifier, strong consideration must be given to its thermal design. For ideal class *B* operation, maximum device dissipation occurs when the peak signal is 63.6 per cent of the supply voltage. At this condition the dissipation per device is 0.1015 V^2_{CC}/R_L. For a supply or voltage of 36 volts and an equivalent load resistance of 16 ohms, this corresponds to 8.25 watts per device.

Ideal class *B* operation is too severe for small signal operation so that a small amount of positive bias is generally necessary. If the zero signal current is 100 ma, maximum collector dissipation could easily be 10 to 12 watts per transistor. This power must be dissipated without permitting the collector temperature to exceed 100°C and preferably without adding to the temperature of the lower level transistors and the power rectifiers.

A 17×10×4-inch 14 Ga aluminum chassis was used to investigate the temperature rise with various transistor mounting schemes. Silicone grease was used at all thermal contacts and the power dissipation per unit was 10 watts.

1) The four transistors were uniformly spaced on a 4×17-inch side of the chassis and electrically insulated but thermally coupled to the chassis with anodized aluminum wafers. The temperature rise at the transistor base was 29°C while the coolest spot on the chassis was 10°C above ambient.

2) Each of the four transistors was attached directly to a 3×4×3/16-inch aluminum plate and the aluminum plates were mounted on one 17×4-inch side of the chassis. A sheet of 0.004-inch bond paper was placed between the aluminum plates and the chassis to supply thermal coupling but electrical insulation. The temperature rise at the transistor base was 25°C while the coolest portion of the chassis was 8.5°C above ambient.

3) Each of the transistors was mounted directly on a Delco 7270725 heat sink. Two of the heat sinks were attached with nylon insulators to each of the 10×4-inch sides of the chassis with 1-inch clearance between the chassis bottom and the heat sink bottom to permit free circulation of air. The temperature rise at the transistor base was 22°C and the coolest portion of the chassis was 1.5°C above ambient.

4) The heat sinks of part 3, with transistors attached directly to them were attached back to back, two in a row, and mounted vertically 1¼ inches above the 10×17-inch side of the chassis. The heat sinks were insulated from each other with nylon insulators. The temperature rise of the transistor base was 22°C as before but there was no noticeable temperature rise on any of the 4-inch sides of the chassis.

This last method was considered to be the best thermal design since it kept the transistor base temperature as low as any of the other methods and did not produce any chassis temperature rise to interfere with the heat sinking of the driver transistor and the current rating of the power rectifiers. A 14×10×3-inch 16 Ga aluminum chassis was now adequate since it had sufficient surface area for components and was relieved of the need of acting as a power transistor heat sink.

Heat sinking of the 2N600 driver transistors is particularly important since they are being operated at about 0.5 watts collector dissipation. These transistors were stud mounted to ⅛-inch brass plates that were cut and drilled to fit the anodized aluminum power transistor insulating wafers. The back of the brass plates were lapped with fine sand paper to eliminate high spots and burrs. Two of these assemblies were thermally coupled but electrically insulated from one of the 10×3-inch sides of the chassis with the anodized aluminum wafers.

The TI 494 transistor is silicon and dissipates only 6 mw so that it does not require heat sinking. The 1N2860 rectifiers operate within a chassis maintained essentially at room temperature and can therefore operate at full rating. These rectifiers along with the pilot lights and the 500 ohm bias and dropping resistors are mounted at the chassis end most distant from the 2N600 transistors to minimize their effect on the operation of these transistors.

The thermistors used for bias stabilization of the power transistors normally come as disks with wires soldered to the two sides of the disk. To obtain good thermal coupling between the thermistors and the power transistor heat sinks they were modified by removing one of the wires and soldering them flat to an $\frac{1}{8}$-inch brass plate similar to the one used for the 2N600 transistors. This operation was performed on a hot plate that just barely brought the brass plates to solder melting temperature and was then turned off to avoid over heating the thermistors. Each of these brass plates was then thermally coupled but electrically insulated from the corresponding heat sink by an anodized aluminum wafer. Excellent temperature control was obtained in this manner.

Transistor Complement

It was mentioned previously that the transistor complement could contain either *p-n-p* or *n-p-n* transistors in the output stage. From a practical point of view, however, only the *p-n-p* transistors need to be considered since *n-p-n* transistors of this dissipation rating are available only in silicon and are quite expensive. This dictates that the driver transistor be of the *p-n-p* type and the low level transistor be of the *n-p-n* type.

Since the low level and driver transistors are physically much smaller than the output transistors, it is to be expected that types will be available which will have considerably better frequency response than that of the output transistor. This permits the high-frequency response to be limited mainly by the output transistors and leads to a more gradual reduction in response and variation in phase than if several stages were contributing to the drop off simultaneously.

The original amplifier[1] used Delco 2N174 transistors with a beta cutoff frequency of 7 kc. This is still a very fine transistor for the purpose. All recorded tests on the present amplifier were made with Bendix 2N1073A transistors. The second channel had one 2N1073A and one 2N1073, selected for high breakdown voltage, which had reasonably well matched transfer characteristics. Matching of the transfer characteristics serves to keep down the distortion at low signal levels. On the other hand the use of the 0.27 ohm emitter resistors serves to reduce the need for transistor matching. A judicious balance between these two conditions must be achieved.

The manufacturer specifies an alpha cutoff frequency of 1.5 mc and a low current beta of 50 for the 2N1073A transistor. This corresponds to a beta cutoff frequency of about 30 kc. Additional transistors that might be considered for the output stage are the RCA 2N1906 and the TI 2N1046 series.

A Sylvania 2N68 transistor with a beta cut off frequency of 12 kc was used in the original amplifier.[1] The Philco 2N600 transistor, with a beta cutoff frequency of about 100 kc is a much better choice at the present time. Very few other transistors will satisfy the driver requirements as well as the 2N600.

Any high frequency *n-p-n* transistor will satisfy the requirements of the low level stage. The relatively inexpensive TI 494 transistor, with an alpha cutoff frequency of 20 Mc, satisfies these requirements admirably and removes the need to design additional temperature stability into this stage.

Iron Core Components

The Stancor C2685 filter chokes used in the power supply are the only iron core components available commercially at the present time. Some transformers relatively close to the power transformer may be found commercially, but the driver and output transformers will have to be fabricated. The windings involved are relatively simple and should not cause an undue amount of trouble to anyone really interested in completing the construction of this amplifier.

1) The power transformer uses a $1\frac{3}{4}$-inch stack of EI-$1\frac{3}{8}$-inch laminations. The type of laminations used is not critical. The primary contains 310 turns of No. 20 HF wire while the secondary contains 114 turns of No. 17 FV wire. The laminations on this transformer may be interleaved three at a time if desired.

2) The output transformer uses a $1\frac{1}{2}$-inch stack of EI-$1\frac{1}{8}$-inch grain oriented laminations with Mil-T holes. The winding consists of 212 bifilar turns of No. 20 HF wire with taps at 106 and 150 turns on one file only. This coil must be tightly wound to fit in the window space available. This transformer uses 103 lamination sets which must be completely interleaved.

3) The driver transformer uses a $\frac{7}{8}$-inch stack of EI-$\frac{7}{8}$-inch laminations. The laminations of this transformer must not be interleaved. The E sections should be butted against the I sections without any additional spacers. The type of laminations used here is not critical either. All of the wire used on this transformer is No. 24 FV. The primary consists of two 250 turn sections connected in series. The secondary consists of 90 bifilar turns sandwiched between the two sections of primary. The end of one of the bifilar secondaries is connected to the start of the other one to form the center tap.

For transformer winding hints and a comparison of iron characteristics the reader is referred to an earlier paper.[4] The *H* holes described in that paper correspond to what is now called a Mil-T hole.

Chassis Layout

The chassis top layout is shown in Fig. 8. The power transformer was the only iron core component mounted on top of the chassis. This was necessary because this transformer was too large to fit within the 3-inch chassis depth. The chassis space around the heat sinks is relatively clear in order to provide free flow of cooling air.

[4] A. B. Bereskin, "Build it yourself," IRE Student Quarterly, pp. 15–37; September 1956.

The heat sinks themselves are $1\frac{1}{4}$ inches above the chassis to permit air to get into the "smokestack" region between the two back to back transistors. The heat sink tops, which are the tallest components on the chassis, are $4\frac{1}{4}$ inches above the surface of the chassis. The input potentiometers were intended only for channel balancing purposes and were therefore placed relatively inaccessibly below the right hand heat sinks close to the input phono jacks.

The chassis bottom layout is shown in Fig. 9 in spread-out fashion. A considerable amount of equipment has been assembled in a relatively small space but all components are accessible for ease of servicing. Transistors TR1 are not shown on this diagram since they are soldered directly to the terminal strip on the far right.

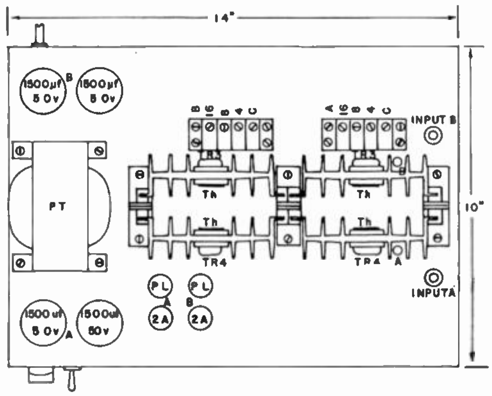


Fig. 8—Chassis top layout.

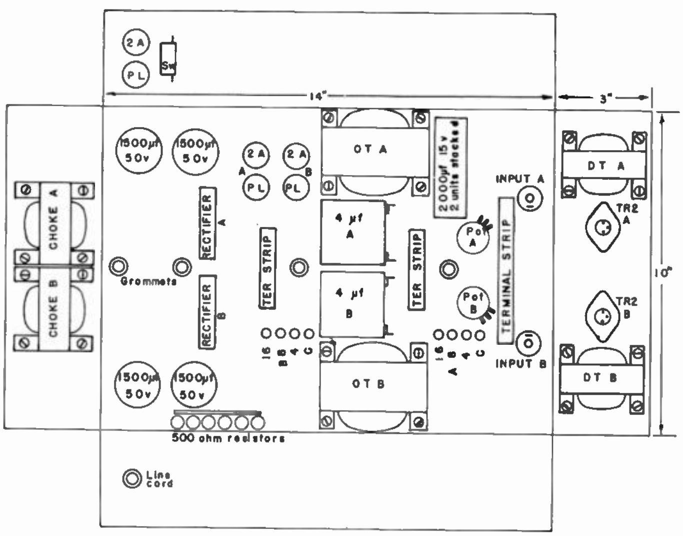


Fig. 9—Spread out view of chassis bottom layout.

It will be noticed that the driver transformer cores have been placed at right angles to all other cores and relatively distant from each other to avoid hum pickup and crosstalk. The residual hum of the amplifier is 74 db below 32 watts while the crosstalk is 70 db below 32 watts.

Initial Adjustment

When this amplifier is first turned on, the ac input should be controlled with a Variac and brought up to operating level gradually while the current in the output stage is observed. In order to avoid excessive voltage on TR2, the resistor R_B should be initially set at 1 megohm and should then be increased until the emitter of TR2 is at 11 volts when the dc supply is 36 volts. The output transistor bias and balance can be checked by measuring the voltage across the 0.27 ohm resistors. This voltage should correspond approximately to 100-ma current flow for both transistors. If substatial unbalance exists, the 500 ohm bias resistors should be modified slightly or different transistors should be used. The use of transistors other than those specified may lead to substantially different values of bias and feedback components.

Reducing the input coupling capacitor below 4 μf is not recommended since it leads to low frequency instability when the input potentiometer is set at zero. If the input capacitor is of the metal can variety, its capacitance to the can will be about 100 pf, and then the 100 pf input circuit capacitor may be omitted. If paper tubulars are used, the 100 pf capacitor must be used.

Summary

The amplifier described in this paper has been employed in a home stereo system in conjunction with a transistor phono preamplifier[5] developed previously. The two units are compatible and provide excellent reproduction of the recorded signal. The performance characteristics presented in this paper amply testify to the quality of the amplifier. A photograph of the complete amplifier, which weighs 28 pounds, is shown in Fig. 10.

Where to Buy

The net prices shown in the "List of Material" were obtained from a late catalog but may be subject to some variation depending on where the material is purchased. Most of these items are available from your local radio parts or electrical supply distributor. If you live in an area that does not have these distributors the parts may be purchased from one of the several mail order houses that operate on a national scale.

The iron core material may be purchased from some local transformer manufacturer or from Thomas and Skinner, Inc., 1120 East 23 St., Indianapolis 7, Ind. The Thomas and Skinner price list specifies a minimum order charge of $10.00 and a minimum item charge for two or more items of $5.00 per item, so that it would be desirable for two or more people to pool their lamination orders to meet these minimums.

[5] A. B. Bereskin, "A transistorized stereo preamplifier and tone control for magnetic cartridges," IRE Trans. on Audio, vol. AU-8, pp. 17–20; January–February, 1960.

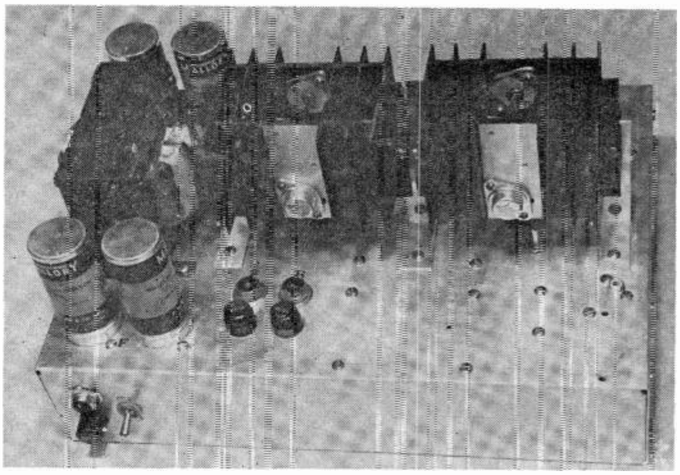


Fig. 10—Complete amplifier.

The Fenwell, ZB05J1 thermistor has been discontinued since this paper was written. The Fenwell KDO5L1 thermistor is available and will work equally well. Stocks of the ZB05J1 may still be available at some of the parts distributors.

LIST OF MATERIAL

Quantity	Item	Total Net Cost
1	14×10×3-inch 16 Ga aluminum chassis	$ 2.78
4	2N1073A Bendix transistors	29.40
2	2N600 Philco transistors	6.90
2	TI 494 Texas Instrument Transistors	5.50
8	1N2860 RCA Silicon Rectifiers	5.92
4	2B05J1 or KD05L1 Fenwall Thermistors	5.00
4	Transistor mounting kits with anodized aluminum washers	1.80
4	No. 7270725 Delco Heat Sinks	3.00
8	No. 7269634 Delco insulating spacers for heat sinks	.40
4	1500 μf-50 volt electrolytic capacitors	12.08
2	2000 μf-15 volt electrolytic capacitors	4.30
2	25 μf-12 volt electrolytic capacitors	1.88
2	4 μf-150 volt metallized paper capacitors	7.74
2	.1 μf-200 volt paper capacitors	.48
3	Fuse mounts	1.23
3	Pilot light assemblies	2.13
3	1829 pilot lights	.95
1	SPST power switch	.52
1	Line cord	.21
2	Phono jacks	.32
2	50 K linear potentiometers	1.82
2	4 screw-terminal binding posts	.68
6	500 ohm-10 watt wire wound resistors	3.36
2	100 ohm-10 watt wire wound resistors	1.12
1	Tube Transistor Silicone Compound	1.47
	Assorted resistors, capacitors, grommets, terminal strips, and hardware	
2	C2685 Stancor Filter Chokes	6.96
$5\frac{1}{2}$ lb.	EI-$1\frac{3}{8}$-inch HS Radio 6AAS M-15 Laminations (26 Ga)	1.76
7 lb.	EI-$1\frac{1}{8}$-inch Mil-T hole Orthosil 3XAAS M-6 Laminations (29 Ga)	3.14
$2\frac{1}{2}$ lb.	EI-$\frac{1}{6}$-inch HS Radio 4AAS M-19 Laminations (26 Ga)	.78
2 lb.	No. 20 HF wire	2.40
1 lb.	No. 17 FV or HF wire	1.14
1 lb.	No. 24 FV wire	1.34

Eight Cardinal Points in Loudspeakers for Sound Reproduction*

PAUL W. KLIPSCH†, SENIOR MEMBER, IRE

Summary—Good audio sound quality in stereophonic reproduction may be summarized in the "eight cardinal points of loudspeakers," which are: 1) minimal distortion, 2) optimum size, 3) avoidance of rattles, 4) avoidance of shadows, 5) avoidance of cavities, 6) wide spacing, 7) proper number, and 8) toe-in.

I. Distortion

The DISTORTION of a speaker is so closely related to power output and size that these functions will have to be considered together. The entire first section deals with distortion as a function of power and size. Other qualities pertinent to size only are then considered separately.

Defining distortion as the generation of frequencies not originally present eliminates the waveform changes due to simple anomalies of frequency response. Then, distortion and frequency response may be considered separately. So defined, distortion means the production of harmonics of the original sounds, and the generation of modulation products between different original sounds. Harmonic distortion is not particularly objectionable, *per se*, unless of severe order and magnitude, but the nonlinearity which causes harmonic distortion produces the other forms. Modulation distortion is the interaction between 2 tones, like 261.63 (middle C) and 329.63 (E above middle C), giving rise to sum and difference tones of 68.00 and 591.26 which are discordant with the original tones and with 69.30 (C#) and 587.33 (D). This illustrates simple "intermodulation." Another form, frequency modulation, takes place when a sheet of air vibrating at large amplitude at a low frequency carries with it a higher frequency which is then caused to flutter. For example, a sheet of air moving 0.21 in at 50 cps would frequency modulate a higher frequency by an rms value of 0.35 per cent (peak value 0.50 per cent), which is enough to be quite harsh to the listener.[1]

Distortion may be minimized by using large radiating surfaces. There are two ways to accomplish this. One may employ a direct radiator of large area, or employ a horn of adequate mouth area. The radiation of one acoustic watt at 30 cps was accomplished by Kellogg,[2] who employed 56 speaker units of 8-in diameter, approximately 14 square ft of cone area, limiting diaphragm motion to 0.08 in.

With a horn, the same effective mouth area is required, with the horn function acting as a transformer, so that a relatively small diaphragm performs small excursions at relatively high pressures, the transformation to high volumetric velocity at low pressure at the mouth being performed by the horn. Kellogg proposed four drive units of 8-in size, limiting excursion to only 0.04 in, in a horn. To conserve size, the horn was admitted into the corner of a room so that the actual mouth was augmented by the mirror images formed by reflections from the three mutually perpendicular wall surfaces. (Kellogg illustrates the arrangement using two walls and the ceiling for one arrangement and two walls and the floor for another.)

Since "intermodulation" was not part of the audio vocabulary at the time of Kellogg's work, it seems significant that Kellogg chose 0.04–0.08 inch as the limiting diaphragm motion. This seems to be a safe excursion in terms of generation of modulation distortion of both kinds.

Choice of the direct radiator or horn is dependent on both economic and functional factors. The optimum horn is a complicated structure and is expensive, but its power output capacity is not limited by the size of the horn and the transient response is superior. The direct radiator is simpler and less costly; in optimum size it will be about twice as big as the optimum horn, but its size may be reduced below optimum with sacrifice of power output capacity, whereas the horn must be optimum or sacrifice many functional attributes including tonal range, distortion and power limits.

Power Required

Sound pressure measurements at various recording sessions indicate peak sound intensity levels of the order of from 100 to nearly 120 db, or 20 to nearly 200 dynes per square centimeter at the listener's location. Massa[3] notes "peak sound pressure of the order of 100 dynes per square centimeter are necessary to secure realistic musical reproduction" (Massa Chart 72).

To produce 100-db sound intensity outdoors in proximity to a ground plane (semi-infinite space) at a distance of 10 ft from the loudspeaker requires a sound power output of 14 acoustic watts. Indoors, a "typical" living room of nearly optimum size, say 16×24 ft with 9-ft ceiling, would have a volume of approximately 3500 cubic ft. Optimum reverberation time for this size room would be between 0.8 and 1.0 sec (Massa Chart 70).

* Received by the PGA, August 21, 1961.
† Klipsch and Associates, Hope, Ark.

[1] P. W. Klipsch, "Subjective effects of frequency modulation distortion," *J. Audio Engrg. Soc.*, vol. 6, p. 143; April, 1958.

[2] E. W. Kellogg, "Means for radiating large amounts of low frequency sound," *J. Acoust. Am. Soc.*, vol. 3, pp. 94–110; July, 1931.

[3] F. Massa, "Acoustic Design Charts," Blakiston Co., Philadelphia, Pa.; 1942.

To produce 100 dynes per square centimeter in such a room requires 0.8 acoustic watt output (Massa Chart 72).

Power Available

Direct Radiator: To produce 1 acoustic watt output from one side of a 10-in piston at 30 cps, 1.4-in peak amplitude or 2.8 in excursion is required. For 0.28 in power output is 0.01 w. For 0.21-inch excursion, power output is 0.0056 w (Massa Chart 64). This excursion produced severe frequency modulation distortion.

Limiting diaphragm excursion to about 1/16 in (using 0.060 in), the power output from a 10-in-diameter piston is approximately 0.0005 w at 30 cps. To produce 1 acoustic watt at 30 c, limiting a direct radiator cone excursion to 0.060 in would require a 66-in-diameter piston, or the equivalent of 44 cones of 10-in diameter (Massa Chart 64).

If the back radiation can be utilized by a "reflexed"-type enclosure, half as much cone area is required. For a realistic size of speaker, it is preferable to permit the bass response to droop rather than to permit distortion to rise.

Harmonic distortion of a single frequency in a "long throw" direct radiator can be held to very low values; measurements of a typical bass unit at $\frac{5}{8}$-in excursion showed negligible harmonic distortion. But when two frequencies are mixed, modulation distortion becomes dominant at high output levels (0.01 acoustic watt). This explains the muddy and even gravelly quality of a small direct radiator driven to excessively high output levels.

Horn Speaker: In a horn-type bass speaker, the power available may be calculated from the throat area, piston size and permissible amplitude, or one may set the requirements and find the piston size. For example, for 32.7 cps, 0.030-in peak diaphragm motion (0.060 excursion), 1 acoustic watt output and 84 sq in throat area, Massa's Chart 78 gives a piston diameter of 14 in.

The harmonic distortion at 500 cps due to compression in the horn can be calculated.

Measurements of power in the 30–500-cps range, the 500–5000-c range and 5000 up range, have been made. They indicate that over 9/10 of the power is concentrated in the lower band, and over half the power below 200 cps. Using 0.2 w as the maximum power required at 500 cps when the maximum bass power is 1 w, the power per square inch of horn throat is 0.0024 w. The nominal cutoff due to taper of one high-quality bass horn is 40 cps, so f/fc for 500 c is 12.5. The harmonic distortion is of the order of 0.5 per cent (Massa Chart 79). Negligible modulation distortion results from this low order of nonlinearity.

Amplifier Power Required

Efficiencies of direct radiators vary from 0.1 per cent to about 5.0 per cent. Horns range from 20 to 50 per cent usable. At one extreme is the "long throw" direct radiator which is weighted to produce a low resonant frequency. Comparison of one of these with a low-efficiency horn of the combined horn-direct-radiator type (5 to 10 per cent) indicated an efficiency ratio of 100, leading to the conclusion that the "loaded" direct radiator was from 0.05 to 0.1 per cent efficient. Recommended amplifiers ranged from 60 to 100 w, though the observed distortion changed from fine to gross at about 25-w input (0.025-w output).

In well-designed horns, the typical product of efficiency times the ability to absorb power from the amplifier (a function of impedance mis-match) should yield an "efficacy" of 20 per cent; that is, an acoustic power output from the speaker of 2 w from a 10-w amplifier. Thus, amplifier power output required for one acoustic watt is of the order of 5 electrical watts. Actually 115-db peak acoustic output in a "typical" living room required 0.25 electrical watt peak. This has been demonstrated many times to skeptical audiences.

General Considerations

At frequencies below which a resistive load applies to the diaphragm, means should be provided to limit cone motion. The simplest means is a closed back air chamber, the stiffness of which takes command when other loads approach zero. Lack of such control was a fault of one corner-horn back-loading design, in that below horn cutoff the cone was limited only by its own mass-compliance values.

A more complex below-cutoff control can be achieved with a ported back air chamber; applied to one design of enclosure-type speaker, this porting was of an elongated narrow tube which permitted back pressures to augment front radiation between 30 and 50 c, but which unloaded the diaphragm only below about 10 c.

Realization of the importance of the back air chamber dates back into the 1930's with Bostwick's work at the Bell Telephone Laboratories. A specific application to horn speakers was made in 1940 and reported in 1941,[4] wherein the specific size of a tight air chamber is related to horn throat size and taper rate.

The importance of the back air chamber in both horn type and direct radiators has been widely recognized as a means to limit cone motion and limit distortion.

II. Size

Optimum speaker size is tied to the power output requirements and how much distortion is to be tolerated.

Direct Radiators

The bass size requirement for 1 acoustic watt output at 30 c and 0.060 maximum motion was established as a 66-in piston in an infinite baffle. If modulation distortion is permitted to rise to the 1.5 per cent level, a

[4] P. W. Klipsch, "A low frequency horn of small dimensions," *J. Acoust. Soc. Am.*, vol. 13, pp. 137–144; October, 1941.

½-in excursion would reduce the required piston area by a factor of 8 and a piston of 24-in diameter would suffice. A 14-in piston (15-in drive unit) would have to perform an excursion of about 1.2 in to radiate 1 w at 30 c; this excursion is not possible for any modern 15-in driver unit regardless of harmonic distortion.

The power output of a 14-in cone at 30 c varies as the excursion as follows:

Excursion Inches	Output, Watts
1.2	1.0
0.6	0.25
0.12	0.01
0.06	0.025

Any reduction from optimum size represents some sort of sacrifice. Usually, the most satisfactory compromise is to sacrifice bass response while retaining the lowest possible distortion. The bookshelf speaker typically uses a 12-in driver unit (10-in effective piston diameter) in a small box of either total enclosure or vented type. As the box is usually 1.5 to 2 cubic ft, it does not confine radiation to 2 π solid angle as would an infinite baffle, so a further sacrifice is entailed. If one limits FM distortion to acceptable levels and corresponding cone motion to 0.06-in, the power output at the bottom edge of the bass range is about 0.0005 w, enough for maximum sound pressure levels of between 80 and 90 dbi in a "typical" room.

Excessive speaker size results in separation of bass and treble events, and the 66-in piston, or equivalent area in a group of small speakers, will produce undesirable spatial distribution. The shortest wavelength such a bass unit should be required to radiate would be 94 in corresponding to 140 cps. Such a bass unit might be combined with an 8-in midrange unit operating up to 2000 or 3000 cps and a group of 1-in tweeters arrayed on a portion of a spherical surface.

To afford a desirable size, one recent design uses a single 15-in drive unit, in a 6.7-cubic-ft box, so ported that back pressure is available for radiation below 50 cps, but with the port in the form of an elongated tube with dissipation so that unloading takes place only below about 10 cps. This bass unit can radiate nearly 0.025 w at 0.60-in diaphragm excursion at 30 cps. Its polar or spatial pattern is satisfactory up to 500 cps and tolerable to 1000. Horn-type mid range and tweeter units are used.

Horns

The optimum size horn can be calculated based on the data of Wente and Thuras,[5] wherein a horn mouth of $\frac{1}{6}$ wavelength was found to give satisfactory throat impedance.

A frequency of 32.7 c is the lowest that will be demanded to be reproduced at high power levels; the longer organ pipes radiate mainly harmonics, and the ear begins to respond to longer wavelengths as pulses rather than tones. The wavelength of a tone of 32.7 cps is 34.5 ft. For a horn to radiate into hemispherical space, the mouth diameter should be 34.5/6 or 5.7 ft, with area of approximately 25 square ft. By corner placement, the solid angle of radiation and required area are reduced by a factor of 4.

Given the mouth size and taper rate, the horn bulk may be computed. Without going into the details contained in earlier papers, the corner bass horn optimum bulk is 16 cubic ft.

The horn unloads at some frequency; for a large bass horn with large drive unit, this effective cutoff is sometimes as much as 10 per cent lower than the cutoff computed for taper.[6] Wherever the cutoff does occur, means should be provided to limit diaphragm motion. The natural means is an air chamber, the volume of which is computed from the taper rate and throat size.

Attempts to deviate appreciably from the optimum bulk have resulted in compromised performance. On the large side, one may cite the experiment with a bass horn 18 ft long with 8×10-ft mouth; the spatial and temporal separation of the bass output from any feasible treble system precluded reproduction of speech and other sounds containing complicated transients.

On the small side, the corner-horn back-loading system was developed for a total bulk of 10 cubic ft or less.[7] The cutoff was 45 c instead of the intended 30, and it unloaded below 50 so that diaphragm motions were not confined to the desired 1/16-in excursion especially below the cutoff. The 10 cubic ft of occupied space proved attractive and a licensee sold large numbers of these boxes. A small peak at 100 c afforded a sensation of heavy bass output, but actually the response was down 10 db at 50 cps, and below the 45-c cutoff the diaphragm was limited by its own mass-compliance values rather than by any load imposed by the box.

Size of horn relative to the wavelengths to be reproduced is not subject to variation beyond small limits. It is better to adhere to the optimum size, or forego the advantages of the horn. A direct radiator is less critical.

Transient response in well-designed horns provides peak powers far above the capabilities of direct radiators; the peak output of the optimum size of horn is limited by available input power. Naturalness of music reproduction depends on the dynamic range, and horn performance has been the best means of attaining such dynamic range.

Horns cost more than direct radiators utilizing the same amount of drive mechanisms, but the durability and low obsolescence may more than offset the cost over a period of years.

Frequency response was almost omitted from this

[5] E. C. Wente and A. L. Thuras, "Loudspeakers and microphones," in "Symposium on Auditory Perspective," *Elec. Engrg.*, vol. 53, pp. 17–24; January, 1934.

[6] P. W. Klipsch, "A note on acoustic horns," Proc. IRE, vol. 33, pp. 447–449; July, 1945.

[7] P. W. Klipsch, U. S. Patent No. 2,731,101; January 17, 1956.

consideration because it rates about last in importance. Yet, more effort is spent to gain "flat" frequency response than in optimizing other values.

Experience indicates that both direct radiators and horns are limited to a range of about 4 octaves if reasonably flat response and efficiency appropriate to the type are demanded. One multihorn designed affords 4 octaves woofer range, 3.3 octaves midrange and about 1.5 octaves tweeter range, with crossovers at 500 and 5000 cps. These values result in low modulation distortion if individual units are "clean."

If a midrange flat to within a 10-db peak-to-trough ratio can be achieved, it should be considered good. One midrange achieved 4.5 db over the 500–4000 range; by contrast, a reflexed type displayed 28 db ratio with a resultant "nasal" quality on voice and "twangy" quality on piano reproduction.

The flatest tweeter ever tested by this writer was one using ionized air. At the time its reliability was poor, and a piston-horn type tweeter was adopted for production, the choice being that of a unit most nearly resembling the ionized air reference unit.

Size of bass, midrange and tweeter are related to the wavelengths to be handled. It is usually best to use single units for a given function; two or more speakers radiating the same wavelength produce "dipole effects" or peculiar spatial radiation patterns. Hence, optimum size is as important in midrange and tweeter units as in bass sections; a single unit of optimum size is superior to two units of equivalent size.

III. Rattles

Avoidance of rattles should go without saying. Yet, one has only to look at and feel the many "audio furniture" cabinets with sliding, or worse yet, slatted doors which rattle to the touch. Acoustic feedback from speakers to electronic components in the same cabinet falls in this same fault, which is mainly the result of letting decor take precedence over function.

This is not a problem.

Other requirements such as wide spacing and toe-in require that speakers be separate from electronics, and doors on the separate speakers become unnecessary. A speaker design should include, besides the basic function of transforming electrical to acoustic power, the solidity of structure which permits accuracy of reproduction without introducing spurious sounds due to rattles.

IV. Shadows

Since high frequencies (short wavelengths) do not turn corners, it is important that the sound source is not shadowed by backs of chairs, other people, or the listener's own knees. Optimum tweeter height appears to be about 44 in from the floor for living room use, with listeners seated. A height of 30 in has proved a reasonably satisfactory compromise. For the center stereo speaker, ceiling height has solved the problem of location in some environments, and the stereo illusion has been retained. As Snow[8] states " . . . localization in the vertical plane is poor. The loudspeakers can therefore be placed above or below the stage level without loss of illusion provided high fidelity of reproduction is maintained." Care must be exercised to avoid separation of bass and treble in any one loudspeaker system: usually 2 ft or 2 msec are tolerable separations. The various speaker systems in a stereo array must be separated by much more than that, and the heights of the various speaker systems may differ by several feet. It is generally better to elevate the speakers than to suffer the effects of shadows, but this may demand a compromise with the treatment of cavities (Section V).

There have been some rather bizarre speaker designs introduced in the last dozen or so years. One involved locating a coaxial speaker behind a slitted board. The reflections within the so-formed chamber and the limited egress port resulted in gross measured peak-to-grough response anomalies. The ostensible purpose of the arrangement was to improve bass response, but the deterioration of both treble and bass indicates the "design" was based on wishful thinking rather than on basic physics. This same design violated principles discussed in Section V.

Application of common sense seems to be the best criterion in speaker design in these phases relative to rattles and shadows. The basic principles are solidity of structure and ability to "see the tweeter" from any listener location. Grille cloth of good quality does not obscure the radiators acoustically, but posts or large supports for the cloth, or boxed-in tweeters may be expected to produce poor area coverage and anomolous frequency response. The best tweeter is substantially in the open.

V. Cavities

The effect of a cavity formed by mounting a speaker box on 14-in legs was evaluated and reported in a paper.[9] The response loss at 50 c was approximately 24 db. The next 2 octaves—50 to 200 c—exhibit peak-to-trough response ratios in excess of normal. This is about the simplest form of cavity and response deterioration which can be caused by it.

A more complicated cavity structure has been employed to enhance bass radiation through peaking due to resonance. The same thing can be accomplished by applying a speaker driving unit to an undamped pipe; resonances will heavily re-enforce some sounds, and the speaker is definitely louder. In the absence of some reference, like original sound, some listeners will make a first impression choice of the louder of two speakers. One particular structure consisted of several coupled

[8] W. B. Snow, "Basic principles of stereophonic sound," *J. Soc. Motion Picture and Television Engrs.*, vol. 63, pp. 567–589; November, 1953. (See especially p. 579.)

[9] P. W. Klipsch, "Corner speaker placement," *J. Audio Engrg. Soc.*, vol. 7, pp. 106–109, 114; July, 1959.

air chambers, some in back of the drive unit, exhausting to space through the front air chambers. The response between 100 and 1000 c consisted of eight major peaks with intervening troughs deeper than 20 db and some exceeding 30 db, yet response at 30 c was down 25 db.

The cavitity of these types should not be confused with the basic back air chamber which is a fundamentally correct principle of physics. A good air chamber design is capable of calculable performance. For example, to achieve a minimum total acoustic reactance it was possible to design an air chamber with a volume of 2.9 times the product of the horn throat area and the taper rate.

A definite function was sought and achieved. On the other hand, the experimental variation of size and number of air chambers without a basic understanding of the principles will necessarily result in a series of resonances—something the acoustic physicist tries to avoid in a speaker.

In the building of experimental speakers, the variables within an element and the number of elements are so great that the odds are about one in a million that one hits it right on any one trial. Experimentation with the benefit of so-called "theory" changes the odds by a factor of several thousand. This intimates why basic knowledge of cavities is essential to success; it also intimates why so many built-in speakers are unsuccessful.

Stereo

The next three sections deal specifically with stereophonic sound reproduction.

From the literature on this subject, it should be evident that speakers which yield superior results for monophonic[10] reproduction should be excellent for stereo. Basic principles of good sound reproduction are as applicable to stereo as to monophonic.

VI. Spacing

In 1934 the Bell Telephone Laboratories reported the first full-scale demonstrations of stereophonic sound reproduction.[5] A speaker array 42 ft wide was used. In 1953, Snow further reported on the developments arising from cinema experience.[8]

Recent experience has proved the feasibility of speaker arrays presenting 90° angle subtended at the listener. By contrast, angles less than about 10° provide no stereo effect at all.

Stereo, even in the extent of a mere right-left sense and a degree of ambiency, requires more than 10° subtended angle.

Stereo in the sense of reproducing the original spatial relationships requires larger angles. Large angles are best served with flanking speakers spaced to the corners, even to the extent of using the long wall of an oblong room.

Examples

In an auditorium 50 ft wide and 75 ft long, three speakers were arrayed across the available frontal 50 ft with flanking units in corners. Accurate localization of sounds was possible even in the front and back rows of seats, and listeners on the flanks near the walls localized sound sources with only slightly less accuracy than those located on the axis of symmetry.

In a 14- ×15-ft living room with the three-speaker array on the 15-ft wall, localization is good in over half the room area including the extreme flanks.

A 16- ×25-ft room with 9-ft ceiling was acoustically treated to achieve a reverberation time of about 0.8 sec at 500 cps. The speaker array is corner-to-corner on the long wall. Recording and playback in the same room is possible, and comparisons of playback with original sound is good. Stereo geometry studies conducted in this studio give rise to criteria of accuracy of localization which prove the advantages of wide speaker spacing, a center unit for focus and toe-in.[11]

VII. Number of Speakers

The Bell Telephone Laboratories used three speakers, the center either in an electrically independent channel or bridged across two stereo channels. No subsequent development has succeeded in obviating the need for three speakers, except where only the simple right-left effect is sought. For the hearing mechanism to be able to fuse the sounds into a curtain resembling the original, a source in the central area is needed.

It should be obvious that this center speaker, being the focus of attention, should be the very best that can be obtained. Its midrange and tweeter should exhibit the same high quality demanded of the corner flanking speaker systems.

As Snow[8] points out, the center system may be smaller than the flanking units, but since the localization of a soloist must tend toward the center, the quality of the entire treble range must be at least equal to that of the other speakers in the stereo array.

Snow, both in the cited reference[8] and in his patent,[12] points out that the center speaker reduces "lateral shifting of the virtual sound source with changes in the observer's position." This is true whether the center speaker is part of an electrically independent channel or is bridged across two stereo channels.

Filling of the "hole in the middle" becomes incidental; when the virtual sound sources correspond to the original locations, the resulting "accuracy of geometry" includes eliminating the hole in the middle. Without the

[10] Snow, *op. cit.*, proposed "Monophonic" as the opposite of "Stereophonic" in the same way that "Monaural" and "Binaural" are opposite pairs.

[11] P. W. Klipsch, "Stereo geometry," unpublished.

[12] W. B. Snow, U. S. Patent No. 2,137,032; November 15, 1938.

center speaker, the absence of a hole in the middle becomes a rare "special case" for perhaps just one observing location and perhaps not even one location.

This is true regardless of the stereo recording technique or number of microphones. No bridging of microphones can fill the hole in the center if there is no speaker there to reproduce it except sometimes for one unique observing location. The center speaker is capable of affording the "solid sound curtain."

VIII. Toe-In

Slightly later than the famed Symposium, Snow showed the desireability to angle or "toe-in" the outside speakers.[8,12] The effect is to reduce the shift of the virtual sound sources for different observing locations, thus supplementing the function of the center speaker.

Even prior to the revival of the bridged center speaker, this writer advocated corner speaker spacing and toe-in for stereo.[13] The effect was to enable listeners at the flanks to hear stereo, but listeners at all points experienced the hole in the middle. The virtual sound sources remained at one side or the other. But without the toe-in, the listener on the flank heard only the nearest speaker. Thus, toe-in accomplishes its function even with only two speakers.

[13] P. W. Klipsch, "Experiences in stereophony," *Audio*, vol. 39, pp. 16–17, 41–42: July, 1955.

The principle has to do with projecting more sound energy toward the opposite side of the listening area. A natural "floodlight" spatial radiation pattern is desirable—the "spotlight" pattern would be fatal—and the toe-in for the bast majority of suitable listening rooms becomes 45°.

The center speaker is still necessary. Combined with toe-in of flanking units, the error figures[11] derived for a central listening position were of the order of 0.12 and increased to 0.16 for observers at the flanks of the listening area. With the center speaker off and the toe-in eliminated, the error was 0.62. For the spacings used, a "live" error figure of 0.04 obtained. Error figures for three independent channels and for the two channels with bridged center speaker differed by insignificant amounts.

IX. Conclusion

The requirements of good audio and good stero may be summed up in the eight cardinal points.

Large books could be written on these points. The foregoing is an attempt to touch the high spots, while affording some basic bibliography for one who would delve further.

"Colorless" Artificial Reverberation*

M. R. SCHROEDER† and B. F. LOGAN†, MEMBER, IRE

Summary—**Electronic devices are widely used to introduce in sound signals an artificial reverberation subjectively similar to that caused by multiple reflections in a room. Attention is focused on those devices employing delay loops. Usually, these devices have a comb-like frequency response which adds an undesired "color" to the sound quality. Also, for a given reverberation time, the density of echoes is far below that encountered in a room, giving rise to a noticeable flutter effect in transient sounds. A class of all-pass filters is described which may be employed in cascade to obtain "colorless" reverberation with high echo density.**

Introduction

ELECTRONIC devices are widely used today to add reverberation to sound. Ideally, such artificial reverberators should act on sound signals exactly like real, three-dimensional rooms. This is not simple to achieve, unless one uses a reverberation chamber or the electrical equivalent of a three-dimensional space. Reverberation chambers (and plates[1]) are preferred by broadcast stations and record manufacturers because of their high quality and lack of undesirable side effect, but they are not truly *artificial* reverberators.

In this paper, we shall focus our attention on electronic reverberators consisting of delay-lines, disk or tape-delay, and amplifiers. Electronic reverberators are both cheaper than real rooms and have wider applicability, notably in the home (unless one wants to convert the basement into a reverberation chamber). They can also be employed to increase the reverberation time of auditoriums, thereby adapting them to concert hall use, without changing the architecture.

* Received by the PGA, December 8, 1960; revised manuscript received, February 27, 1961. Reprinted with permission from the *J. Audio Engrg. Soc.*, vol. 9, pp. 192–197; July, 1961.
† Bell Telephone Labs., Inc., Murray Hill, N. J.

[1] W. Kuhl, "The acoustical and technological properties of the reverberation plate," *European Broadcasting Union Rev.*, Pt. A—Technical, No. 49; May, 1958.

Before attempting the difficult task of reproducing room characteristics by delay-lines, it is wise to recall some of the important properties of large rooms.

The Frequency Response of Large Rooms

A room can be characterized by its normal modes of vibration. It has been shown[2] that the density of modes is nearly independent of room shape and is proportional to the square of the frequency:

$$\text{number of modes per cps} = \left(\frac{4\pi V}{c^3}\right) f^2.$$

Here V = the volume, c = the velocity of sound, and f the frequency.

Above a certain critical frequency,[3] given by

$$f_c = 2000\sqrt{T/V} \text{ (reverberation time } T \text{ in seconds, } V \text{ in m}^3\text{)},$$

the density of modes becomes so high that many modes overlap. In this frequency range, which is of prime interest for large rooms, the concept of individual normal modes loses its practical (though not its theoretical) significance. The behavior of the room is governed by the collective action of many simultaneously excited and interfering modes resulting in a very irregular amplitude-frequency response.[4,5] However, the fluctuations are so rapid (on the frequency scale) that the ear, in listening to a non-steady sound, does not perceive these irregularities.[6] (The response fluctuations can be heard by exciting the rooms with a sinewave of slowly varying frequency and listening with one ear.) When the room response is measured using, instead of a sinewave, a psychoacoustically more appropriate test signal, such as narrow bands of noise, the response would indeed be much smoother.

It is this apparent smoothness of a room's frequency response which people have found particularly difficult to imitate with artificial reverberators. In this paper we shall describe electronic reverberators which have perfectly flat amplitude-frequency responses. Thus, they not only overcome this long-standing difficulty but are actually superior to rooms *in this one respect.*

However, a flat frequency response is not the only requirement for a high-quality reverberator. Before we can hope to successfully design one, we must also know something about the *transient* behavior of rooms.

[2] P. M. Morse and R. H. Bolt, "Sound waves in rooms," *Rev. Mod. Phys.*, vol. 16, pp. 69–150; April, 1944.
[3] M. R. Schroeder, "Die statistischen Parameter der Frequenzjurven von grossen Räumen," *Acustica*, vol. 4, Beiheft 2, pp. 594–600; 1954.
[4] E. C. Wente, "Characteristics of sound transmission in rooms," *J. Acoust. Soc. Am.*, vol. 7, pp. 123–126; October, 1935.
[5] H. Kuttruff und R. Thiele, "Über die Frequenzabhängigkeit des Schalldrucks im Räumen," *Acustica*, vol. 4, Beiheft 2, pp. 614–617; 1954.
[6] A. F. Nickson and R. W. Muncey, "Frequency irregularity in rooms," *Acustica*, vol. 5, pp. 44–48; 1955.

The Transient Behavior of Rooms

How does a room respond to excitation with a short impulse? If we record the sound pressure at some location in the room as a function of time, we first observe an impulse corresponding to the direct sound which has traveled from the sound source to the pick-up point without reflection at the walls. After that we see a number of discrete low-order echos which correspond to one or a few reflections at the walls and the ceiling. Gradually, the echo density increases to a statistical "clutter." In fact, it can be shown[7] that the echo density is proportional to the square of the elapsed time:

$$\text{number of echos per second} = \frac{4\pi c^3}{V} t^2.$$

The time after which the echo response becomes a statistical clutter depends on the width of the exciting impulse. For a pulse of width Δt, the critical time after which individual echos start overlapping is about

$$t_c = 5 \cdot 10^{-5}\sqrt{V/\Delta t} \ (V \text{ in m}^3).$$

Thus, for transients of 1-msec duration and a volume of 10,000 m³ (350,000 ft³), the response is statistical for times greater than 150 msec. In this region, the concept of the individual echo loses its practical significance. The echo response is determined by the collective behavior and interference of many overlapping echos.[8]

Another important characteristic of large "diffuse" rooms is that all modes have the same or nearly the same reverberation time and thus decay at equal rates as evidenced by a straight-line decay when plotting the sound level in decibels vs elapsed time.

Still another property of acoustically good rooms is the absence of "flutter" echos, *i.e.*, periodic echos resulting from sound waves bouncing back and forth between parallel hard walls. Such periodicities in the echo response are closely associated with one-dimensional modes of sound propagation which can be avoided by splaying the walls and placing "diffusors" in the sound path.

The Conditions to Be Met by Artificial Reverberators

After this brief review, we are in a position to formulate conditions to be met by artificial reverberators.

1) The frequency response must be flat when measured with narrow bands of noise, the bandwidth corresponding to that of the transients in the sound to be reverberated. This condition is, of course, fulfilled by re-

[7] L. Cremer, "Die wissenschaftlichen Grundlagen der Raumakustik," Band 1 ("Geometrische Raumakustik"), S. Hirzel Verlag, Stuttgart, Germany, vol. 1, p. 27; 1948.
[8] R. C. Jones, "Theory of fluctuations in the decay of sound," *J. Acoust. Soc. Am.*, vol. 11, pp. 324–332; January, 1940.

verberators which have a flat response even for sinusoidal excitation.

2) The normal modes of the reverberator must overlap and cover the entire audio frequency range.

3) The reverberation times of the individual modes must be equal or nearly equal so that different frequency components of the sound decay with equal rates.

4) The echo density a short interval after shock excitation must be high enough so that individual echos are not resolved by the ear.

5) The echo response must be free from periodicities (flutter echos).

In addition to these five conditions, a sixth one must be met which is not apparent from the above review of room behavior but easily violated by electronic reverberators:

6) The amplitude-frequency response must not exhibit any apparent periodicities. Periodic or comb-like frequency responses produce an unpleasant hollow, reedy, or metallic sound quality and give the impression that the sound is transmitted through a hollow tube or barrel.

This condition is a particularly important one because long reverberation times are achieved by circulating the sound by means of delay in feedback loops. The responses of such loops, which are the equivalent of one-dimensional sound transmission, are inherently periodic and special precautions are required to make these periodicities inaudible.

In the following, a basic reverberator is described which fulfills conditions 1), 3) and 6) ideally. By connecting several of these reverberating elements in series, conditions 2), 4) and 5) can also be satisfied without violating the others.

Two Simple Reverberators

The simplest reverberator consists of a delay-line, disk, or tape-delay which gives a single echo after a delay time τ. Its impulse response is

$$h(t) = \delta(t - \tau), \tag{1}$$

where $\delta(t)$ is the Dirac delta-function (an ideal impulse). The spectrum of the delayed impulse is

$$H(\omega) = e^{-i\omega\tau}, \tag{2}$$

where ω is the radian frequency. The absolute value of $H(\omega)$ is one. This means that all frequencies are passed equally well and without gain or loss.

In order to produce multiple echos without using more (expensive) delay, one inserts the delay line into a feedback loop, as shown in Fig. 1, with gain g of magnitude less than one (so that the loop will be stable). The impulse response, illustrated in Fig. 1(b), is now an exponentially decaying repeated echo:

$$h(t) = \delta(t - \tau) + g\delta(t - 2\tau) + g^2\delta(t - 3\tau) + \cdots \tag{3}$$

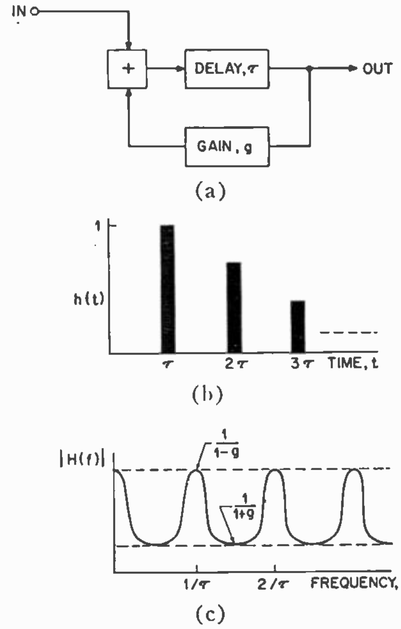


Fig. 1—(a) Delay in feedback loop. (b) Impulse response. (c) Frequency response. Simple reverberators with exponentially decaying echo response. Frequency response resembles comb.

The corresponding complex frequency response is

$$H(\omega) = e^{-i\omega\tau} + ge^{-2i\omega\tau} + g^2e^{-3i\omega\tau} + \cdots, \tag{4}$$

or, using the formula for summing geometric series,

$$H(\omega) = \frac{e^{-i\omega\tau}}{1 - ge^{-i\omega\tau}}. \tag{5}$$

By taking the absolute square of $H(\omega)$, one obtains the squared amplitude-frequency response:

$$|H(\omega)|^2 = \frac{1}{1 + g^2 - 2g\cos\omega\tau}. \tag{6}$$

As can be seen, $|H(\omega)|$ is no longer independent of frequency. In fact, for $\omega = 2n\pi/\tau$ ($n = 0, 1, 2, 3, \cdots$), the response has maxima (for positive g) given by

$$H_{\max} = \frac{1}{1 - g}, \tag{7}$$

and, for $\omega = (2n+1)\pi/\tau$, minima given by

$$H_{\min} = \frac{1}{1 + g}. \tag{8}$$

The ratio of the response maxima to minima is

$$H_{\max}/H_{\min} = \frac{1 + g}{1 - g}. \tag{9}$$

For a loop gain of $g = 0.7$ (-3 db), this ratio is $1.7/0.3 = 5.7$ or 15 db!

The amplitude-frequency response of a delay in a feedback loop has the appearance of a comb with periodic maxima and minima, as shown in Fig. 1(c). Each response maximum corresponds to one normal mode. The natural frequencies are thus spaced $1/\tau$ cps apart.

The 3-db-bandwidth of each peak is approximately

$$\Delta f = \frac{-\ln g}{\pi\tau}, \tag{10}$$

where "ln" denotes the logarithm to the base $e = 2.718 \cdots$. Converting to logarithms to the base 10 (log), one obtains

$$\Delta f = \frac{1}{20\pi \log e} \frac{-\gamma}{\tau} = 0.0367 \frac{-\gamma}{\tau}, \tag{11}$$

where γ is the loop gain in decibels: $\gamma = 20 \log g$. For $\gamma = -3$ db, the bandwidth is about $0.11/\tau$ or only one-ninth of the spacing of the natural frequencies. The subjective effect of this resonant response is the hollow or reedy sound quality mentioned above.

All-Pass Reverberators

In our search for better reverberators, we discovered that a certain mixture of the output of the multiply delayed and the undelayed sound resulted in an equal response of the reverberator for all frequencies. The mixing ratio that accomplishes this and results in unity gain for all frequencies is $(-g)$ for the undelayed sound and $(1-g^2)$ for the multiply delayed sound. The corresponding circuit is shown in Fig. 2. Its impulse response is given by

$$h(t) = -g\delta(t) + (1-g^2) \cdot [\delta(t-\tau) + g\delta(t-2\tau) + \cdots]. \tag{12}$$

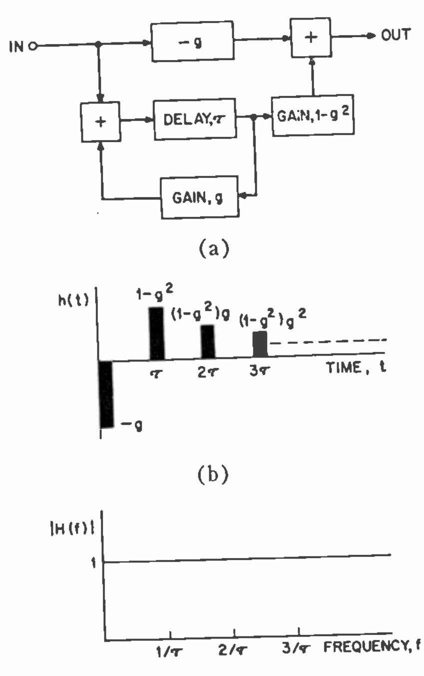


Fig. 2—(a) All-pass filter. (b) Impulse response. (c) Frequency response. Modification of simple reverberator. By adding proper amount of undelayed signal, frequency response of reverberator becomes flat (all-pass reverberator).

The corresponding frequency response is

$$H(\omega) = -g + (1-g^2)\frac{e^{-i\omega\tau}}{1-ge^{-i\omega\tau}}, \tag{13}$$

or

$$H(\omega) = \frac{e^{-i\omega\tau} - g}{1 - ge^{-i\omega\tau}}, \tag{14}$$

or

$$H(\omega) = e^{-i\omega\tau}\frac{1 - ge^{i\omega\tau}}{1 - ge^{-i\omega\tau}}. \tag{15}$$

What is the absolute value of this $H(\omega)$? The first factor on the right has, of course, absolute value one. The second factor is the quotient of two conjugate complex vectors, *i.e.*, its absolute value is also one. Thus,

$$|H(\omega)| = 1. \tag{16}$$

In other words, the addition of a suitably proportioned undelayed path has converted the comb filter (6) into an all-pass filter (16). This is not a mere academic result. The conversion of a comb filter into an all-pass filter is accompanied by a marked improvement of the sound quality from the hollow sound of the former to the perfectly "colorless" quality of the latter.

Now we are in possession of a basic reverberating element which passes all frequencies with equal gain and thus fulfills conditions 1) and 6) above. The spacings and decay rates of the normal modes (though no longer "visible" as resonant peaks of the amplitude-frequency response) are the same as those for the previously discussed comb filter. Thus, condition 3), requiring equal decay rates for the normal modes, is also fulfilled.

Whether the normal modes overlap (condition 2) can no longer be judged on the basis of the amplitude-frequency response because it is constant. However, the phase-frequency response still reflects the distribution of normal modes and thus must conform to condition 2). The phase-lag of $H(\omega)$ as a function of frequency is, with (15),

$$\phi(\omega) = \omega\tau + 2\arctan\frac{g\sin\omega\tau}{1 - g\cos\omega\tau}. \tag{17}$$

A more convenient quantity to consider is the rate of change of phase-lag with respect to radian frequency:

$$\frac{d\phi}{d\omega} = \frac{1-g^2}{1+g^2-2g\cos\omega\tau}\tau, \tag{18}$$

which has exactly the same dependence on ω as the square amplitude-frequency response $|H(\omega)|^2$ of the corresponding comb filter [see (6)]. The physical significance of $d\phi/d\omega$ is that of the envelope or "group" delay of a narrow band of frequencies around ω. According to

(18), for a loop gain of $g=0.7$ this envelope delay fluctuates as much as 32:1 for different frequency bands, with the long delays occurring, of course, for frequencies near the natural frequencies, $2\pi n/\tau$ ($n=0, 1, 2, \cdots$), of the filter. The half-width of the envelope delay peaks is the same as that for squared amplitude [see (10)]. Thus, for a loop gain of -3 db, only one-ninth of all frequency components suffer a large envelope delay, while the remaining frequencies are much less delayed. This constitutes a very unequal treatment of different frequency components and violates condition 2).

The remaining two conditions, 4) and 5), are also violated as we shall see immediately. The relationship between reverberation time T (defined by a 60-db decay) and the two parameters of the reverberator, the delay τ and the loop gain γ in decibels, is as follows. For every trip around the feedback loop the sound is attenuated γ db. Thus, the 60-db decay time is

$$T = \frac{60}{-\gamma}\tau. \tag{19}$$

For $\gamma=-3$ db,[9] we have $T=20\cdot\tau$. Thus, in order to achieve, for example, 2 seconds of artificial reverberation, the loop delay must be 0.1 sec. With this loop delay the basic reverberating element shown in Fig. 2 produces one echo every one-tenth of a second. This constitutes a most undesirable periodic flutter echo. Also, the echo density (ten echos per second) is much too low to give a continuous reverberation. Thus, conditions 4) and 5) are violated.

How can one obtain a less periodic time response and a greater echo density without giving up the all-pass characteristic?[10] If several all-pass feedback loops with incommensurate loop delays are connected in *series*, as illustrated in Fig. 3, the combined frequency response remains flat, while the echo response becomes aperiodic and the echo density increases.

In addition, a better coverage of the frequency axis with normal modes is achieved. In fact, the envelope delay response of the series connection is a sum of terms like (18) with different τ's. Since each of these terms "covers" only one-ninth of the frequency axis, at least five all-pass feedback loops in series are required. On the other hand, one can also show that *too many* all-pass feedback loops in tandem are bad because they lead to a very unnatural, nonexponential reverberation which builds up to its maximum intensity rather slowly before it starts decaying. We shall spare the reader the mathematical details of this peculiar reverberation because he has suffered already too much, we are afraid.

Fig. 4 shows the impulse response of five all-pass filters connected in series with loop delays of 100, 68, 60, 19.7, 5.85 msec. The loop gains are +0.7, −0.7, +0.7, +0.7, +0.7, respectively. This combination of delays and gains was arrived at after considerable experimentation observing the response to a variety of sounds, both on the oscilloscope and by listening, and using a smooth envelope of the decay as a criterion. The appearance of the echo response is quite random and not unlike that of real rooms.[11,12] An increase in pulse density with increasing time can also be noticed.

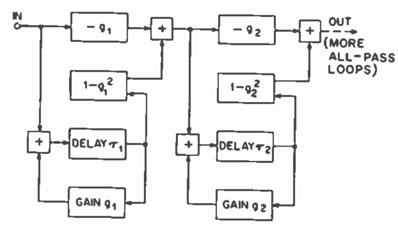


Fig. 3—Series connection of several all-pass reverberators with incommensurate delays to make echo response aperiodic and increase echo density.

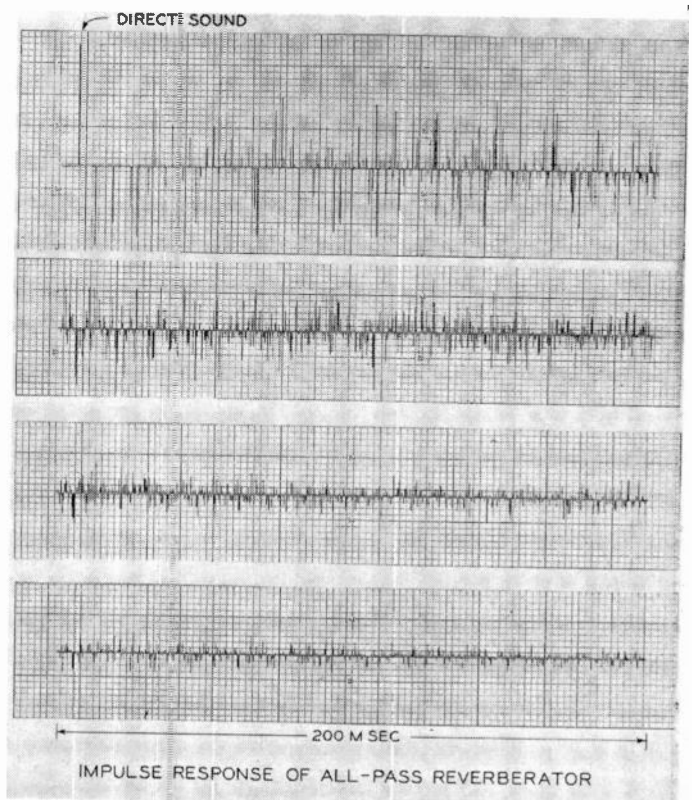


Fig. 4—Echo response of all-pass reverberator consisting of five simple reverberators connected in series, as shown in Fig. 3.

[9] We do not consider open-loop gains greater than 0.7 (−3 db), because in practice it is difficult to maintain the desired closed loop characteristics with gains too close to unity.

[10] N. V. Franssen, "Neuentwicklung in der räumlichen Schallwiedergabe," Bericht über die 4. Tonmeistertagung, Detmold, Germany; October, 1957, p. 10; Frenssen has suggested the use of multiple feedback, proportioning the gains to give a flat open-loop frequency response. This allows larger loop gains without incurring instability. However, a feedback loop around an all-pass filter results in a nonflat frequency response, unless supplemented by a direct path of suitable gain as described above.

[11] E. Meyer and R. Thiele, "Raumakustische Untersuchungen in zahlreichen Konzertsälen und Rundfunkstudios unter Anwendung neuerer Messverfahren," *Acustica*, vol. 6, Beiheft 2, pp. 425–444; 1956.

[12] G. R. Schodder, "Über die Verteilung der energiereicheren Schallrückwürfe in Sälen," *Acustica*, vol. 6, Beiheft 2, pp. 445–465; 1956.

Digital Computer Simulation

The impulse response shown in Fig. 4 was obtained from a large digital computer (IBM 7090) in conjunction with special digital-to-analog conversion and plotting apparatus. In addition to using the computer as a "draftsman," many of the actual reverberation experiments were performed with the help of the digital computer. In this research method (sometimes called "digital simulation"[13]), ordinary ("analog") tape recordings of the sound to be reverberated are prepared and converted into digital tapes by means of special conversion equipment. The computer then reads the digital tape and acts upon it exactly like any desired real equipment would act on the sound signal. To facilitate the programming of the computer, our engineers make use of a special translation program, developed at Bell Telephone Laboratories, which translates their block diagrams into the computer language. Thus, the computer first "compiles" its own program on the basis of the block diagram information and then acts on the sound as the block diagram would do. It then prepares one (or several) digital output tapes which are converted back into analog tape recordings and evaluated by listening. Needless to say, this is a very powerful research tool, especially when complex equipment is to be evaluated! In this manner we have studied the subjective quality of a great variety of reverberators with both flat and nonflat frequency responses.

Application to Quasi-Stereophony

The late Holger Lauridsen[14] of the Danish State Radio has discovered a method of splitting a single audio signal into two "quasi-stereophonic" signals which give the listener all the "ambience" of multichannel stereophony but permits, of course, no correct localization of individual sound sources. In order to achieve this, Lauridsen has used delay networks connected to form a pair of interleaved comb filters.[15] However, these comb filters give rise to unpleasant sound qualities—not quite as pronounced as in artificial reverberators, but nevertheless easily perceptible. We have overcome this disadvantage by using a pair of all-pass filters, like the one shown in Fig. 2, to split the single-channel audio signal. This idea is described in greater detail in a forthcoming publication.[16]

[13] E. E. David, Jr., "Digital simulation in research on human communication," Proc. IRE, vol. 49, pp. 319–329; January, 1961.

[14] H. Lauridsen, "Nogle forsøg med forskellige former rum akustik gengivelse," *Ingeniøren*, no. 47, p. 906; 1954.

[15] M. R. Schroeder, "An artificial stereophonic effect obtained from a single audio signal," *J. Audio Engrg. Soc.*, vol. 6, pp. 74–79; April, 1958.

Summary

Several reverberators of the all-pass type were successfully simulated on a digital computer (IBM 7090). Others were instrumented with delay-lines and tape-delay. No coloration of the reverberation sound was detected in any of these electronic reverberators. Our listening experience with all-pass reverberators indicates that the problem of unequal response to different frequencies has been solved and sound "coloration" completely eliminated. Audible flutter echos have been avoided by the use of several all-pass feedback loops with "incommensurate" delays in series.

The application of all-pass reverberators to the problem of increasing the reverberation time of auditoriums and concert halls by purely electroacoustic means[17] remains to be studied. Here the flat frequency response is particularly important for two reasons: 1) It minimizes acoustic feedback problems. The "ringing" and instability due to the unavoidably irregular frequency response of the room can be reduced by shifting all frequency components of the reverberated sound by a small constant amount.[18] 2) A flat response of the reverberator contributes to the high sound quality required in concert hall applications. Ultimate acceptance of electroacoustic techniques in concert halls and opera houses is assured only if the artificial effects are not recognized as such by the music-loving public.

Acknowledgment

We thank Miss N. Nordahl for assistance in the computer simulation experiments and Messrs. A. J. Prestigiacoma, F. K. Harvey and J. G. Cisek for conducting a variety of auxiliary experiments.

[16] M. R. Schroeder, "Improved quasi-stereophony and colorless artificial reverberation," *J. Acoust. Soc. Am.*, vol. 33, pp. 1061–1064; August, 1961.

[17] R. Vermeulen, "Stereo reverberation," *J. Audio Engrg. Soc.*, vol. 6, pp. 124–130; April, 1958.

[18] M. R. Schroeder, "Stop feedback in public address systems," *Radio Electronics*, vol. 31, pp. 40–42; February, 1960. See also "Improvement of acoustic feedback stability in public address systems," *Proc. 3rd Internatl. Congr. Acoustics*, Elsevier Publishing Co., Amsterdam, The Netherlands, vol. 2, pp. 771–775; 1961.

Correspondence

Comments on "Transient Distortion in Loudspeakers"*

The above paper[1] rings a long unrung bell. My 1920–1922 laboratory notebooks, for my work on electrical recording for the Brunswick Phonograph Company, have many notations headed "Transient Distortions," applying to a number of elements in the chain of more or less resonant devices between the recording microphone and the listener's ears. Whether I coined that term I do not know, but so far as I am aware, my use of it was pretty early in the art of electrical reproduction of sound.

My remedy at that time resorted to antiresonant circuitry for suppression of resonant peaks in the recording system, as well as for those of the reproducing system of Brunswick phonographs. Although these (filters) left some side effects, on the whole they worked pretty well! They might, with modern improvement in filter circuitry, still be useful for modern loudspeakers.

Benjamin F. Miessner
680 N.E. 105th St.
Miami Shores, Fla.

* Received by the PGA, August 21, 1961.

[1] R. J. Larson and, A. J. Adducci, IRE Trans. on Audio, vol. AU-9, pp. 79–85; May–June, 1961.

Contributors

Alexander B. Bereskin (A'41–M'44–SM'46–F'58), for a photograph and biography please see page 94 of the May–June, 1961, issue of these Transactions.

❖

R. L. Craiglow

Robert L. Craiglow (A'50–M'56) was born in Columbus, Ohio, on May 11, 1923. He received the B.S.E.E. degree from The Ohio State University, Columbus, in 1946.

He then joined the Collins Radio Company, Cedar Rapids, Iowa, where he worked on frequency control devices. He was a member of the Group on Frequency Control Devices of the Research and Development Board from 1951 to 1953, and Chairman of the Group on Frequency Control Devices of the Technical Advisory Panel on Electronics, Department of Defense, from 1954 to 1957. Since 1959, as a member of the Research Division of Collins Radio Company, he has been engaged in studies relating to speech processing and bandwidth compression.

❖

N. R. Getzin

Norman R. Getzin (S'57–M'60) was born in Oak Creek, Wis., on March 2, 1927. He received the B.S. degree in education from Milwaukee State Teachers College, Milwaukee, Wis., in 1951, the M.S. degree in education from Northwestern University, Evanston, Ill., in 1956, and the B.S.E.E. degree from the University of Illinois, Urbana, in 1957.

In 1959, he joined the Collins Radio Company, Cedar Rapids, Iowa, where he has been engaged in research on speech compression techniques as a member of the Applied Research Division. Hs is presently enrolled in graduate studies at the Iowa State University, Iowa City.

Mr. Getzin is a member of Eta Kappa Nu.

❖

P. W. Klipsch

Paul W. Klipsch (A'34–M'44–SM'45) was born in Elkhart, Ind., on March 9, 1904. He received the B.S.E.E. degree from New Mexico State College in 1926 and the degree of Engineer from Stanford University, Stanford, Calif., in 1934.

He was employed in the Testing Department of the General Electric Company from 1926 to 1928; the Angelo Chilian Nitrate Corporation, Tocopilla, Chile, from 1928 to 1931; the Independent Exploration Company, Houston, Tex., from 1934 to 1936; the Subterrex Company, Houston, from 1937 to 1941; and the U. S. Army Ordnance Department. He is presently the owner of Klipsch and Associates, Hope, Ark., manufacturers of loudspeakers.

Mr. Klipsch is a Fellow of the Audio Engineering Society and a member of the AIEE, the Acoustical Society of America, Tau Beta Pi, and Sigma Xi. He is a Licensed Professional Engineer in Texas and Arkansas.

❖

Benjamin F. Logan (M'58), for a photograph and biography please see page 59 of the March-April, 1961, issue of these TRANSACTIONS.

❖

Manfred R. Schroeder, for a biography and photograph please see page 60 of the March-April, 1961, issue of these TRANSACTIONS.

❖

Paul L. Simmons was born in Hutchinson, Kan., on February 12, 1929. He received the B.A. degree in 1953 from the University of Denver, Denver Colo. After serving two years in the U. S. Army, 1953-1955, he earned the M.S. degree in library science from the University of Denver in 1956.

He served as Serials Librarian at the Long Beach State College Library, Long Beach, Calif., from 1956-1958. From 1958 to 1961 he was at System Development Corporation, Santa Monica, Calif., as Reference Librarian. He is presently Library Research Analyst for Space and Information Systems Division of NAA, Downey, Calif. Mr. Simmons has made bibliographic searches on Soviet aircraft and missiles, mechanical translation, computers in psychology, Soviet computers, computers in medicine, oceanography, information processing in command and control systems, and satellite surveillance and tracking systems in connection with corporate projects.


P. L. SIMMONS

Richard A. Swanson (S'58-M'59) was born in Wahoo, Neb., on June 30, 1937. He received the B.S.E.E. degree from the University of Nebraska, Lincoln, in 1959.

He then joined the Research Division of the Collins Radio Company, Cedar Rapids, Iowa, where he has done experimental research in speech processing techniques and has conducted propagation studies using the Echo balloon and moon as passive reflectors. At the present time, he is engaged in graduate study at the University of Nebraska.


R. A. SWANSON

Index
to
IRE Transactions
on
Audio
Volume Au-9, 1961

IRE Professional Group on Audio
Combined Index for 1961

Compiled by D. W. Martin

Classification of Subjects

1. IRE-PGA
   - 1.1 General
   - 1.2 Constitution and By-Laws
   - 1.3 National and Regional Meetings
   - 1.4 Chapters
   - 1.5 Membership
   - 1.6 Transactions
   - 1.7 People
   - 1.8 Humor
2. Bibliographies, Reviews, Standards, Tapescripts
   - 2.1 Bibliographies
   - 2.2 Reviews
   - 2.3 Standards
   - 2.4 Tapescripts
3. Sound Systems
   - 3.1 General
   - 3.2 Stereophonic, Binaural, and Spatial Effects
   - 3.3 Military
4. Microphones
   - 4.1 General
   - 4.2 Condenser
   - 4.3 Magnetic
   - 4.4 Crystal
   - 4.5 Moving-Coil
   - 4.6 Ribbon
   - 4.9 Special Microphones
5. Amplifiers
   - 5.1 General
   - 5.2 Preamplifiers and Voltage Amplifiers
   - 5.3 Power Amplifiers
   - 5.4 Frequency-Range Dividing Networks
   - 5.5 Transistorized Amplifiers
   - 5.6 Filters
   - 5.9 Special Amplifiers
6. Loudspeakers
   - 6.1 General
   - 6.2 Direct-Radiator Units
   - 6.3 Horn-Driver Units
   - 6.4 Horns, Enclosures, Baffles
   - 6.9 Special Types
7. Disk Recording and Reproduction
   - 7.1 General
   - 7.2 Disks
   - 7.3 Recording
   - 7.4 Pickups and Tone Arms
   - 7.5 Pre-emphasis and Postequalization
   - 7.9 Special Mechanical Recorders
8. Magnetic Recording and Reproduction
   - 8.1 General
   - 8.2 Tape and Wire
   - 8.3 Recording and Erasing
   - 8.4 Playback
   - 8.5 Pre-emphasis and Postequalization
   - 8.9 Special Magnetic Recorders
9. Acoustics
   - 9.1 General
   - 9.2 Room Acoustics
   - 9.3 Sound Waves and Vibrations
   - 9.4 Speech
   - 9.5 Music
   - 9.6 Hearing
   - 9.7 Psychoacoustics
10. Broadcast Audio
11. Audio Measuring Equipment and Techniques
12. Electronic Musical Instruments

TRANSACTIONS OF THE IRE
CONTENTS

January–February, 1961

AU-9-1

March–April, 1961

AU-9-2

May–June, 1961

AU-9-3

July–August, 1961

AU-9-4

September–October, 1961

AU-9-5

November–December, 1961

AU-9-6

AUTHOR INDEX FOR 1961 TRANSACTIONS ON AUDIO

ANALYTIC SUBJECT INDEX FOR 1961 IRE TRANSACTIONS ON AUDIO

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 Professional Groups Secretary, The Institute of Radio Engineers, Inc., 1 East 79 Street, New York 21, N.Y.