IRE Transactions on AUDIO

Volume AU-7 **NOVEMBER-DECEMBER, 1959** **Number 6**

Published Bi-Monthly

TABLE OF CONTENTS

PUBLISHED BY THE

Professional Group on Audio

IRE PROFESSIONAL GROUP ON AUDIO

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

IRE TRANSACTIONS® ON AUDIO

Published by the Institute of Radio Engineers, Inc., for the Professional Group on Audio at 1 East 79th Street, New York 21, New York. Responsibility for the contents rests upon the authors, and not upon the IRE, the Group, or its members. Individual copies available for sale to IRE-PGA members at $0.80; to IRE members at $1.20, and to nonmembers at $2.10.

The Editor's Corner

Stereoland

ALICE was greeted by a courteous sales-technician as she entered the high-fidelity auditioning room. The courteous sales-technician nodded his head wisely and began his story. "In ancient times," he explained, "people used one loudspeaker and all the sound seemed to come from a hole in the wall. Now we call it monophony, which suggests such things as monogamy, monotony, monopoly, *et cetera.*"

He continued, "Now I'd like to have you hear stereophonic sound with two loudspeakers." He pressed a button and a loud version of "Gaite Parisienne" issued from two speakers, one on each side of the room.

"Why must it be so loud?" exclaimed Alice.

"I've often wondered myself," said the courteous sales-technician. "Maybe that's what hi-fi means—that the volume must be high. It depends a lot on whether you are a demonstrator or a demonstratee. The demonstrators like it loud, and since they are the ones who own the set, it's usually played loud."

"You'll notice that the orchestra is now spread over the entire wall of the room," he went on.

To Alice, everything seemed to come mainly from one loudspeaker—the closer one. She wondered whether she would hurt the attendant's feelings by telling him. After all, he was so courteous. Finally she told him.

"I'm glad you mentioned it," he said, "You are supposed to stand on an imaginary line equally distant from both speakers." Alice moved to a more strategic location, and zeroed in on the music by leaning first to one side, then to the other, until both speakers were about equally loud. "Now isn't that better?" asked the attendant.

"It is better," she agreed. She listened carefully for a while and tried to imagine the orchestra spread over the front of the room but she could not help thinking that the sounds were concentrated at the two loudspeakers. She remarked, "A while ago you said that 'monophonic' was as if the sound came from one hole in the wall. 'Stereophonic' must then mean two holes in the wall."

"No, no, you don't get it," answered the attendant. "What you are probably noticing is the 'hole in the middle'."

"Quite the opposite," said Alice. "I meant that there was no hole in the middle for music to come out of."

"In hi-fi," said the attendant gravely, "when we talk about a 'hole in the middle' we mean a hole where no music comes out."

Alice said nothing, and was relieved when the attendant continued.

"Anyway," he went on, "you were listening to an elementary system, the Mark I. We hardly recommend it anymore, except to beginners, and to people on a budget. Now our Mark II is a three-channel system with an extra loudspeaker in the middle. It's really two channels, but you get the third channel for nothing by mixing the other two. (Well, almost for nothing. You do have to buy another loudspeaker and mixing network.) Some people call it a 'phantom.' It fills the hole in the middle, you see."

Alice thought for a moment. "I suppose," she said, "that after you've filled the hole in the middle you'll notice holes halfway between the middle and each end. You can fill them with a speaker taking part of its input from each adjacent speaker. This leaves some more holes in between, which can be filled, ***ad infinitum.***"

"It sounds like a good idea," reflected the attendant, "and it would surely sell a lot of speakers. Now let's turn on the Mark II system."

They listened for a while. "You know," said Alice, "everything now comes from the center speaker. I'm not getting much stereo effect at all." The attendant turned a control that reduced the loudness of the center speaker. When the volume from the center speaker was low enough, Alice could hear sounds from the sides again. But now she found that her listening position was even more critical than before. Also, at times the sound image jumped suddenly from one place to another, which was quite disconcerting.

"Maybe we are listening to the wrong kind of record," suggested the attendant. He tried another one, which he explained was made with sounds that came definitely either from one channel or the other, with precautions that no blending or mixing took place to give an "in between" effect. Alice admitted that this was the best she'd heard yet. Strangely enough, it sounded better from two loudspeakers about eight feet apart than it did from the elaborate systems. "It may ***sound*** good," explained the attendant, "but it's only for beginners. Advanced listeners want something more sophisticated."

"Dear me," thought Alice, "I must be only a beginner, because I'd rather listen to this one, regardless of the price."

"We come to the Mark III," said the courteous technician, "and here we have mixed lows and satellite highs."

"Is that why the music seems to jump from one speaker to another sometimes?" asked Alice, "like a satellite in orbit, you know."

"No, it's not that at all," answered the sales-technician. "There are two small side speakers, and they had to call them something. They are stationary satellites, and they handle medium and high notes. The low notes all come out from a single woofer cabinet near the center."

"You mix the lows regardless of which side they come from? Why can you do that?" asked Alice.

"Well, there are theories," said the technician. "There is a theory that people cannot tell the direction of sounds when they are below 250 cycles. Some say 100 cycles. Then there is the Haas effect. It's all very convenient, because in many systems the lows get mixed and we wouldn't be able to separate them anyway. As far as the highs are concerned, you get a stereo effect in almost any part of the room, so the audience is not confined to sitting on a line equidistant from the speakers. Now let's listen." And he turned on the Mark III system.

Alice noted a rather pleasant blended effect to the sound, and indeed it persisted as she moved about. Blending made the location of the various sound sources indefinite, and when she closed her eyes she could place them almost anywhere. She was rather uncomfortable, though, when a vocalist sang, for his voice seemed to come from an area two yards wide. Also, she fancied that the low notes still came from the woofer, in spite of theories and the Haas effect.

"You might also be interested in the Mark IV," said the technician. "This is all integrated into a single cabinet. The loudspeakers are on each side of the cabinet and depend on reflections from the walls of the room to give stereo separation."

Alice listened dutifully, for by this time she was rather tired. The Mark IV sounded much like the satellite system.

"I guess that's about all there is," said the technician. "Other systems are mostly variations of these. Why don't you think about it?" And he excused himself, for there were some cash customers waiting.

Alice reflected for a while. She had come to the demonstrations expecting almost perfect realism. In this she was disappointed. On the other hand, she liked the musical effects which were possible, especially when the performance had been adapted to give unusual and pleasing results, by recording groups who used intuition and experiment.

She had to admit that some of the recordings, though artificial, were as enjoyable for listening purposes as a live performance.

—Marvin Camras, *Editor*

PGA News

CHAPTER NEWS

Chicago

In contrast to extremists who have made loudspeaker enclosures of poured concrete, masonry, or sand, Peter W. Tappan demonstrated effective enclosures made of paper at the PGA meeting in the WSE Auditorium on September 11, 1959. Mr. Tappan is senior research engineer for the Warwick Manufacturing Corporation, Chicago.

Scanfax reviews his talk on "Loudspeaker Enclosure Walls":

In high quality sound reproduction systems, it has been traditional to house the speaker in an enclosure with thick wooden walls. It is the purpose of this paper to show that thick or heavy walls are not always necessary.

The acoustical requirements for enclosure walls and the behavior of typical walls, and the effects on system performance of wall material, structure, thickness, area, shape, curvature, bracing, and damping were discussed.

Results of several experiments were described, and an A-B demonstration of paper versus wooden enclosures were given.

Mr. Tappan is responsible for the acoustical research at Warwick and has worked on the design and development of speaker systems, phonograph pick-ups, and stereophonic and pseudostereophonic equipment.

Dayton

The PGA Dayton Chapter concluded the 1958–1959 year as host to the Dayton Section for a tour and program at the Baldwin Piano Company, Cincinnati, Ohio. Chartered buses provided the transportation with dinner at the Kemper Lane Hotel. Dr. Bereskin, newly elected National Chairman of PGA, addressed the group on the activities and responsibilities of the National and Regional Officers. After dinner, the group proceeded to the Baldwin Piano Company where John Quitter spoke on the history of the company and conducted a tour of the plant in operation. After some color films, the "fabulous" Eddie Osborne, Baldwin organist, demonstrated the Baldwin Organ's features and entertained for the pleasure of the group. Refreshments were served at the conclusion of the entertainment.

Syracuse

"Stereo Components" was the subject of an interesting program presented by A. F. Petrie and P. E. Pritchard of General Electric Company on May 19, 1959.

ANNOUNCEMENTS

During the IRE National Convention, March 21–24, 1960, in New York, PGA expects to participate in three sessions:

1) Tuesday morning, March 22, a full session on Audio,
2) Tuesday afternoon, March 22, a half session combined with Broadcast,
3) Wednesday afternoon, March 23, a full session on Stereophonic Sound.

A brief annual meeting of PGA is planned for Tuesday, March 22. Award winners will be announced, and newly elected officers introduced. We hope to see you there.

The Pyramid Stylus*

C. D. O'NEAL†

Summary—**A new shape for a phonograph needle stylus is described which greatly improves the reproduction possibilities from phonograph records. The new stylus was specifically produced for use with 45°-45° stereo recordings, but will perform equally well on all microgrooved monaural recordings. The scheme described evolves a shape of reproducing stylus having a basic geometry relating to the cutting stylus used in forming all microgrooves. Computations, experimental data and charts are also used to support the improvements to be expected in reproduction performance.**

STEREOPHONIC phonograph records produce more distortion when traced with a given stylus radius than do the more conventional type of laterally modulated monophonic recordings. The higher values of distortion are not a product of the recording technique, but rather the result produced by the reproduction technique.

This paper introduces a new concept in phonograph needle stylus design. Many of the limiting performance parameters normally experienced with the reproduction of sound from phonograph records are minimized by the use of this new stylus. All types of records, whether single track stereo or conventional monaural, are also better reproduced.

This new type of stylus is being introduced at a time when the recommendation of the RIAA Engineering Committee for the future is a 0.5-mil needle tip. This recommendation is for a very good reason, as will be shown. The disadvantages to be found with the 0.5-mil needle tip radius for playing microgrooved records are the major increase of relative pressure per square inch of the needle tip's area on the vinyl record material and the added fragility of the stylus. Pressure per unit area and the fragility factor are each increased by a factor of four, over the 1-mil radius.

Experimental use of a 0.5-mil radius has produced permanent noise in many shallow grooves found on LP records, even with needle pressures of 5 grams. Investigation under a microscope has revealed that scoring occurred on grooves which have their depth radii larger than the 0.5-mil value. These findings indicate another reason why the 0.5-mil radius is not a satisfactory universal stylus for use on both monophonic and stereophonic phonograph recordings.

The needle problems associated with the single track stereo record led our company to explore the possibility of designing a producible needle stylus for desired superior performance parameters. Some of the objectives set up for this project were to include better strength and needle wearability and also better translation performance from the grooves of the recording. Of major importance was the need to extend the useful life of the recording's surface, even to the extent of stopping short of a maximum degree of performance perfection.

In theory, devising a proper performing stylus was found to be less difficult than the production of such a stylus, especially since it was desirable that it be formed from diamond material. Also proper evaluation was found to present a few problems.

For a comprehensive story of this development, certain well-established theoretical precepts underlying the fundamental phonograph needle stylus criteria will need to be reviewed.

All technical writings on the subjects underlying the principles of phonodynamics clearly recognize that the most serious limitation to the ultimate fidelity attainment of phonograph record reproduction exists with the shape of the tip of the reproducer's stylus.

The Pierce-Hunt report, published in 1938, has this to say: "The curve traced by the center of the tip (which we shall assume to be spherical) of a reproducer stylus sliding over a sinusoidal groove surface is not, unfortunately, sinusoidal, and it is embarrassing that the only way to reduce the distortion due to this effect is to reduce the size of the needle tip."

Subsequent writings on this subject have accepted the hemispherical point as an inevitable limitation to the reproduction of phonograph records. In fact, there has been a tendency among writers on this subject to vindicate this misfit of the stylus to the grooves in the records.

What are some of these mechanical factors which cause tracing distortion and what can be done to overcome them?

Although LP lateral recordings and 45-45 stereo recordings are two very different types of groove modulation, they have two things in common: the material of which these records are made and the forming or cutting tool which is used to make their modulated grooves.

Fig. 1 shows the profile relationship between the cutting tool which forms the groove and the "effective" circular contact area of a stylus tip which is used to trace this groove. From the illustration used, it can be seen that the cutter which generates the groove presents a plane cutting surface which is vertical and in the same plane with a radius of the record. This also results in an operating condition where this cutting plane is perpendicular to the axis of an unmodulated groove. These are typical conditions found with laterally modulated LP monaural recordings. The illustration is also a state where the modulation velocity has attained a value which is maximum; this will better illustrate the per-

* Manuscript received by the PGA, August 28, 1959; revised manuscript received, October 20, 1959.
† Fidelitone, Inc., Chicago, Ill.

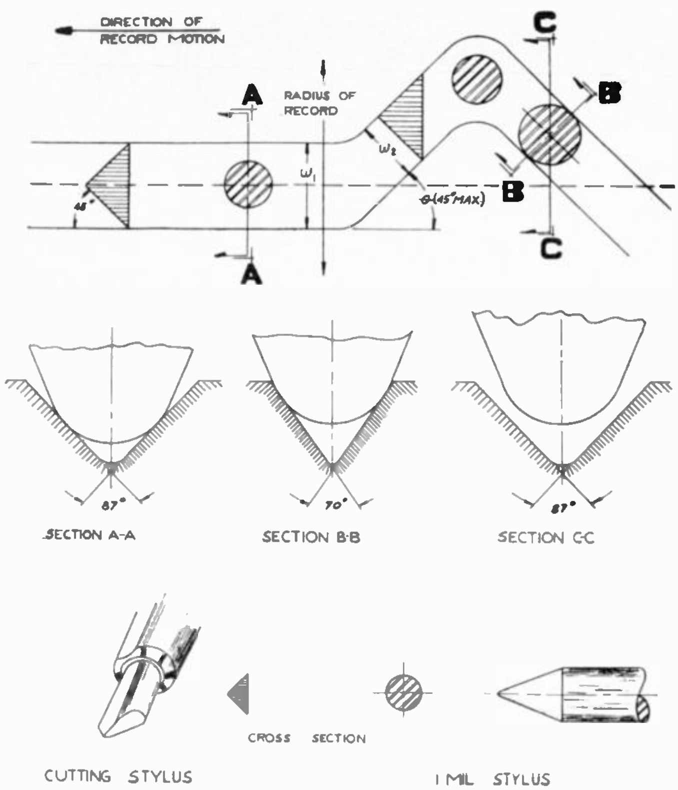

Fig. 1—Plan and sectional views of the assumed reproducing and cutting styli bearing relations, in a typical laterally modulated groove.

formance criteria during recording and reproduction.

Since the stereo 45-45 modulation does not occur in the same plane, as shown in Fig. 1, this illustration is omitted for reasons of simplification. It will, however, be shown later how the accepted technique of presentation will provide a common ground for extrapolating distortion values for either system.

It should be observed that as the lateral reciprocating force is applied to the cutting stylus, in the process of adding modulation, a constriction in width of the groove occurs at points of odd quarter wavelengths. Such a constriction always occurs when the plane cutting surface of the recording stylus moves at any angle θ to the direction of the axis of an unmodulated groove. Such a formed groove, when traced by a hemispherical surface like that of the contact radius of the needle stylus, must lift and fall twice during the tracing of each wavelength. This is commonly known as "pinch effect" and is illustrated in sectional views A-A and B-B in Fig. 1. Observe also that during the "pinch effect" the contact radius of the stylus is continually changing in value, as is also the included groove angle to the hemispherical tip.

In Fig. 2, a typical cross-sectional view of a microgroove is again shown, but in this case the pitch of the groove is running about 260 lines per inch. The centerline separation of each unmodulated groove will, therefore, be 3.85 mils. The nominal width of each groove at the surface of the record, W, will be 2.5 mils, which leaves room, from the separating land area, for peak amplitudes of modulation under 0.7 mil. Since the included angle of the groove, as shown , is 87°, the depth exclusive of the bottom radius will be essentially one-half of the width dimension.

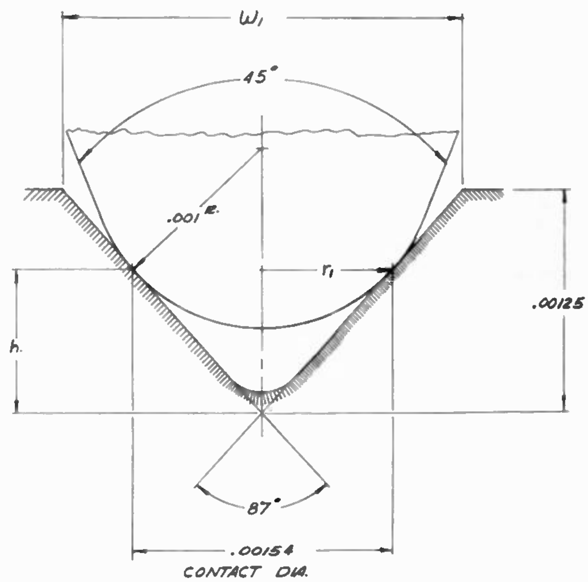

Fig. 2—Sectional view of stylus resting in an unmodulated record groove.

A stylus with a 1.0-mil radius, if placed in such a groove, will engage the two sidewalls which form the groove at two points. These two points lie in a given plane which creates a so-called "contact radius" to the groove. This contact radius is somewhat smaller than the hemispherical radius of the stylus, and its value can be resolved by layout or by calculation in the following manner:

$$r = R \cos 43.5° = 0.00072 \text{ inch} \tag{1}$$

where r is the contact radius and R the stylus radius.

The height at which the contact radius engages the groove sidewalls (h_1) shown in Fig. 2, can be resolved mathematically by

$$h_1 = r \cot 43.5° = 0.000756 \text{ inch.} \tag{2}$$

Now if lateral modulation is added to the unmodulated groove in an amount which can be related to the forward moving velocity of the record (turntable rotation), a condition can be created as illustrated in Fig. 3. The physical shape of the cutting stylus (see Fig. 1) limits the maximum modulation angle θ to 45°. The constricted width W_2 in Fig. 3, under such a maximum condition, will be

$$W_2 = W_1 \cos 45° = 0.00176 \text{ inch.} \tag{3}$$

The included angle of such a groove, relative to the contact points of the stylus shown in Fig. 3, becomes 2θ:

$$2\theta = \text{arc tan} \frac{0.5W_2}{0.5W_1} = \text{arc tan } 0.707 = 35.2°$$

$$2\theta = 70.4°.$$

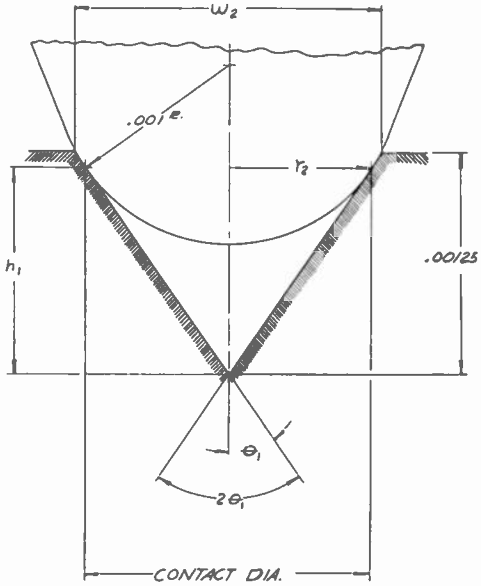

Fig. 3—Sectional view of the same stylus illustrated in Fig. 2 resting in its vertically displaced position which is caused by the condition of maximum pinch effect found on records.

It now becomes evident that as the included angle in the groove becomes smaller, a vertical component of motion is imparted to the stylus tip. The amount of this vertical pinch effect will change the value of the contact radius of the stylus by this relationship:

$$r_2 = R \cos 35.2° = 0.000818 \text{ inch.} \tag{5}$$

This shows a 12 per cent change in contact radius under this condition; Lewis and Hunt [9] point out that second-order distortion might be introduced into lateral motion by the vertical motion if the pinch effect is impeded.

The height to which the stylus will rise in the specific condition illustrated in Fig. 3 is

$$h_1 = r_2 \cot 35.2° = 0.001147 \text{ inch.} \tag{6}$$

The change in height over the condition of Fig. 2 is, of course, $h_1 - h$ or 0.391 mil, which represents a peak-to-peak vertical displacement of double frequency which is created by the pinch effect.

If a maximum lateral amplitude of 0.7 mil is accepted for a peak modulation value, then the vertical pinch effect components of modulation can be a very significant value and, of course, the frequency of the pinch effect is always twice the frequency value of the lateral component. Such deductions lead one to suspect large velocity values created by the pinch effect. Such a condition as illustrated in Fig. 3 can occur only at the inside grooves of a recording; even then it cannot occur at very low audio frequencies. The reason for this is that when the modulation angle or so-called clearance angle θ reaches 45°, the modulation velocity exactly equals the moving groove velocity V_g. The lowest frequency on an inside groove at which such a condition of equality can occur is [7]:

$$f = \frac{\omega}{2\pi} = \frac{12314}{6.8} = 1955 \text{ cps} \tag{7}$$

$$= \frac{V_g}{\gamma} = \frac{8.64}{0.0007} = 12314 \text{ radians per second} \tag{8}$$

where V_g is the groove velocity of the record at a given radius from its center hole (inside groove), and γ is the contact radius which has been taken as a nominal value for a 1-mil stylus radius.

The modulation velocity is a function of amplitude and frequency and varies through zero to its maximum value twice for every wavelength, where the zero occurs at 0° and 180° and the maximum at 90° and 270°. Since the pinch effect frequency resulting from a fundamental frequency of 1955 cps is 3910 cps; since it was previously determined that the peak-to-peak amplitude under this condition was 0.391 mil; and since the velocity will be found through wavelength points of 0°, 180° and 360°, then the total pinch effect velocity V_p at such points in time may be resolved by [7]:

$$\begin{aligned} V_p &= a\omega \cos \omega t \\ &= \frac{0.000391 \times 6.28 \times 3910 \cos 0°}{2} \\ &= 4.8 \text{ inches per second.} \end{aligned} \tag{9}$$

One can appreciate why high vertical compliance is equally important with high lateral compliance for a properly functioning phonograph needle system. Additionally, these figures should make it clear why the phonograph pickup must be made as insensitive to the pinch effect modulation as possible.

In lateral types of modulation, as found on conventional LP recordings, the even harmonics are considered to cancel out in a similar way to push-pull audio techniques. Since 45-45 stereo recordings use the vertical as well as the lateral motion to resolve the recorded information, the pinch effect no longer is inconsequential. The 45-45 stereo system is sensitive to all harmonics, odd and even.

We need now to examine some of the parameters of stylus performance which have been accepted as being most important for extrapolating a figure of merit type of distortion value.

The approach taken to the problem has been described by Pierce and Hunt [1] and deals with harmonic values of waveform distortion generated by a circle tracing a groove which has been generated by a plane. The distortion values which will be resolved are close approximations, but cannot be expected to include the yield of the record material to the force of the stylus, which tends to reduce the measured electrical values of distortion below the indicated mathematical values [2]. It is reasonable to recognize that such deformation of record material must take place when a needle stylus is used to trace the groove, since a simple extrapolation will reveal pressures per unit area, which are many tons per square inch, even with 5-gram needle force.

A further condition to be found on the inside radius of a phonograph record is the translation loss[1] which can be in excess of six decibels at ten thousand cycles, even with a new stylus. The amount of this deterioration, in both frequency response and fidelity, is plainly audible to a discriminating listener. Direct measurements of exact amounts of translation loss are difficult, since no standard frequency test records are available with their high frequencies recorded on the inside radius of the record where wavelength will be shortest.

The values of tracing distortion will be given a graphic type of presentation relating to the curve traced by the center of a circle, sliding over a sinusoidal groove surface. The circle exemplifies the contact radius of the phonograph needle stylus, and the sinusoidal groove exemplifies a single frequency recorded on a record. This technique permits a presentation on a single chart of a complete range of tracing distortion resulting from any conditions of groove modulation.

Fig. 4 shows such a graph, adapted for use with LP, 33⅓ rpm records which have lateral modulation. Values for coordinates[2] *ka* and *kr* are found relative to the RIAA recording characteristic at inside and outside diameters on the record.

Fig. 5 shows the RIAA recording characteristic for all microgroove records, both monaural and stereophonic, with the one exception of recording level which is lower on stereo. The continuous line 14 decibels above the velocity characteristic line represents the maximum instantaneous rms program peak velocity.

To make the distortion analysis complete, a separate graph is shown in Fig. 6 for use with the 45-45 stereo type of recording. Such information could have been added to Fig. 4, but would have introduced a confusion of lines.

Referring again to Fig. 4, *ka*, composing the ordinate scale, displays directly the ratio of cyclic stylus velocity to the tangential groove velocity; a uniform scale as appended to the right-hand margin of the chart shows the velocity amplitude in decibels, in reference to the tangential groove velocity taken as zero level. The cutting angle, being maximum at the point where the center line of the laterally driven stylus crosses the center line of the unmodulated groove, is also represented in a scale located on the top right-hand portion of the chart. The tangent of this angle is given directly by *ka*. The heavy solid line extending from the lower left-hand of the chart to the top right-hand, labelled the RIAA recording characteristic, is also used for referencing the frequency which may be read directly in kilocycles from the abscissa scale. This solid line reference scale is

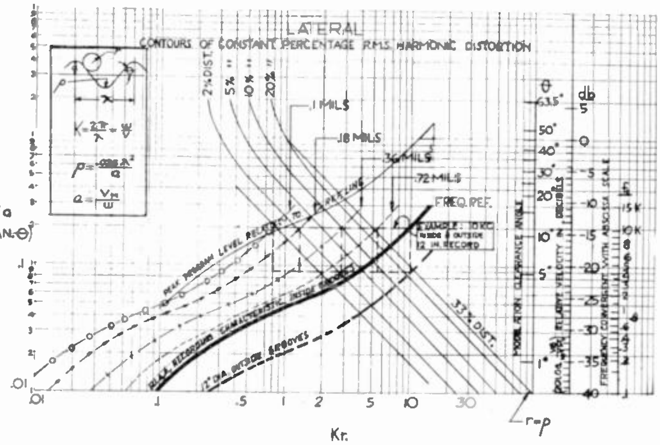

Fig. 4—RIAA recording characteristics loci are superimposed upon values of *ka-kr*. The heavy reference line, which corresponds to the nominal values of RIAA recording characteristic, also references the frequency to be read directly, either on the ordinate frequency scale provided or on an abscissa scale in *kc*.

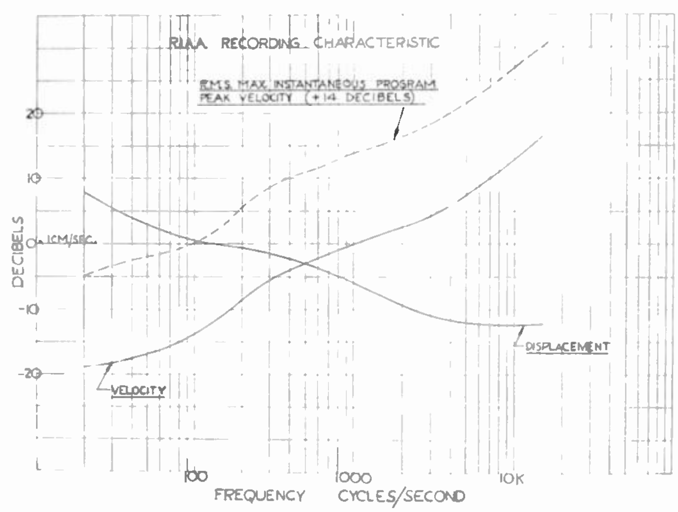

Fig. 5—Standardized RIAA recording characteristics for both stereo and monophonic microgrooved phonograph records.

plotted relative to the groove velocities which exist at the inside radius on a microgroove record. Fourteen decibels above the reference line is a continuous line representing the 14-db rms peak program voltages, which are found on records. The heavy dotted line, which is located approximately 7½ decibels below the reference line, is the relative condition as found on the outside grooves of a 12-inch record. The four separate broken lines, appearing coincident to, and to the left of, the solid reference line, represent four separate styli contact radius conditions. The 45° sloping solid lines, running diagonally from the top left to the lower right of the chart, represent the distortion values found in indicated regions which are coincident to relative frequencies, and various styli contact radii values experienced in playing a single sine wave on a laterally modulated 33⅓ rpm record. The location and value of these distortion lines on the chart are derived from the relationship [1]

[1] In disk recording, translation loss is the loss in output of high frequencies occurring between the outer and inner diameters of the record for a frequency of constant amplitude.

[2] Definitions for *k*, *a* and *r* can be derived from inserted formulas in Fig. 4.

$$H_n = \frac{(kakr)^{n-1}}{n}, \tag{10}$$

where H_n represents the velocity amplitude of any harmonic component, relative to the fundamental component.

As an example of how the chart works, let us determine the harmonic distortion generated by a 1-mil stylus reproducing a single frequency of 10,000 cycles on an LP, $33\frac{1}{3}$ rpm recording at a point near the inside grooves on the record, where the recording velocity is 13 db above 1 cm/sec. [1 kc is the zero db reference of 1 cm/sec (see Fig. 5).]

Solution: First locate the 10-kc point on the solid reference line, which is at $ka = 0.2$; $kr = 10$; move to the left along the $ka = 0.2$ line to the point of intersection with the dashed line representing the 0.72-mil contact radius line (the contact radius for a 1-mil stylus); and discover that this point is in the region of 35 per cent distortion.

Reducing the contact radius to 0.36 mil, we follow $ka = 0.2$ further to the left until it intersects the 0.36-mil dotted line at $kr = 2.8$, and discover that the distortion is reduced to a value of 10 per cent. Following the chart further to the left, to a point where the contact radius line intersects the 0.18-mil dotted line at $kr = 1.4$, we discover the distortion is reduced to under 3 per cent.

If it is desired to find the distortion of the 1-mil stylus on the outside grooves of the 12-inch record instead of on the inside grooves as before, it is only necessary to find the frequency as before on the solid reference line; traverse downward on line $kr = 10$ to $ka = 0.085$ where the dotted reference line is; and then follow to the left on the $ka = 0.085$ line to $kr = 5.4$, which coincides with the 0.72-mil contact radius as found on the outside of the record.

This procedure can be better understood if it is realized that the kr value does not change on the inside or the outside of the record, since the stylus radius is constant until it starts to obtain flattened areas caused by the abrading process. Actually, the distortion lines represented on the chart are only the third harmonic values. The fifth harmonic values add very little to these distortion values, as can easily be seen by the empirical relationship of ka to kr in (10). From relative relationships, it can be shown that when the radius of curvature of modulation ρ is equal to the contact radius r of the stylus, approximately 33 per cent distortion results in laterally modulated records and approximately 50 per cent distortion exists in vertically cut records. (See Fig. 6.)

The value of ρ may be found by

$$\rho = 0.025\lambda^2/A \tag{11}$$

where A is the amplitude and λ the wavelength. If it is desired to find a specific relationship between ka and kr for any given value of distortion in per cent, it may be found by

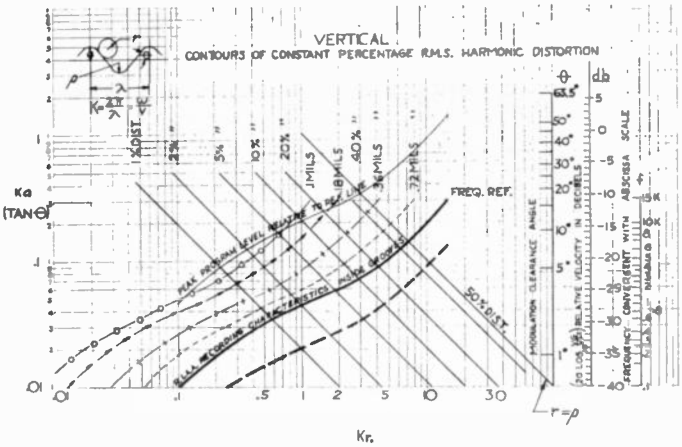

Fig. 6—Same display as presented in Fig. 4, but showing harmonic distortion lines as related vertical recordings such as those exemplified by the new 45°-45° stereo types.

$$ka = \frac{(nH_n)^{1/n-1}}{kr}. \tag{12}$$

In the empirical relationship which numerically exists between the styli contact radii of various sizes and the record groove parameters encountered, high values of third harmonic distortion are produced. It will also be noticed that a drastic reduction in distortion takes place merely by reducing the size of the stylus radius.

Now consider the conditions which result when playing a recording which has a vertical modulation. Fig. 6 shows a chart similar to Fig. 4. The main difference here is where the distortion lines appear relative to the other common variables. Again refer to (10), which is used to determine the harmonic values, and substitute 2 for 3 for the value of n in the formula (as in the lateral recording case); these larger distortion values will result.

The stereo single groove pickup system is sensitive not only to the lateral components of modulation, but also to the vertical components. This being the case, all of the distortion components are added vectorally. Again referring to the original illustration for the LP recording, and only changing the condition to a single track of stereo groove modulation reproducing 10,000 cycles on the inside grooves of a stereo record, it is discovered that the distortion which was 35 per cent for the same relative conditions for lateral recording is now more than 50 per cent. The 0.3-mil contact radius is discovered to reproduce about the same distortion value for stereo reproduction applications as the 0.72-mil contact radius did in lateral recording. The 0.18-mil radius will introduce approximately 5 per cent distortion on the inside grooves of the stereophonic record.

These illustrations show some of the needle problems involved in reproducing stereo records compared with the problems which have been experienced with the LP microgroove records. The values of distortion found for the illustration presented are only normal values commonly found on the inside grooves of a record. Recording rms velocities, 14 db above the values cited, are permitted by specification. The translation losses which would surely result in some maximum conditions of recorded velocities are unpredictable. Such misfits of stylus to the radius of modulation produce sharp cusps or spikes in the response waveform, and may result in mistracking.

While we are dealing only with the generated harmonic values of distortion, it should be borne in mind that with complex tones the intermodulation products of distortion may be as high as 10 times the total values generated as harmonic components. Since intermodulation distortion is inharmonic in nature (it is the sum and difference components of those frequencies mechanically generated), they can cause severe disturbance.

The published theoretical work on disk reproduction generally assumes that the needle point is new. With a worn needle, distortion will be considerably higher, as shown by Bauer [8], who gives relationships between distortion to be expected and the size of the flat on the stylus formed by record abrasion.

Consideration can now be given to a basic modification to the phonograph needle stylus, which will include certain parameters for overcoming some of the faults experienced with the conventional type of hemispherical point.

In view of what has been presented, it would seem logical to use a shape and type of reproducing stylus exactly like the cutting stylus which would overcome the faults found with the conventional hemispherical point. Such a thought is logical, except for the fact that the cutter would probably cut a new groove with its sharp edges. It is evident, however, that the design of a properly formed reproducing stylus should provide for the fact that it must trace the groove as a cutting stylus would, but without its sharp edges.

Fig. 7 shows a groove which is schematic, as was previously presented in Fig. 1; this time, however, a different type of tracing stylus is shown. The reproducing stylus in this illustration is a type of pyramid point, which resembles a cutting stylus without its cutting plane. This shape can trace a record's groove with essentially the same translation accuracy as the cutter; that is, of course, it can do so if the orientation of the stylus is kept within proper limits. The four converging triangular planes, which form the four ridges on the pyramid shape, can be given as small radii as are compatible with the record's resistance to deformation with a given needle force. The included angle of the groove-engaging ridges of the new pyramid stylus is much larger than the conventional 45° type used for the hemispherical point. In fact, the included angle is very nearly the same value as the included angle of the cutting stylus which is used to form the groove. Such a shape provides several valuable features, not the least of which is a much greater strength. Such a shape presents a larger contact area to the sidewalls of the groove, even though the effective engaging radius to the modulation is much smaller than could be practical with a hemispherical point. Such a scheme not only extends the wear life of the stylus and record, but also permits the use of practical values of needle force in the groove. It permits a more accurate tracing of the modulation. The effective contact radius of such a stylus now becomes the radius of the two laterally opposite ridges of the pyramid which are used to engage the sidewalls of the groove.

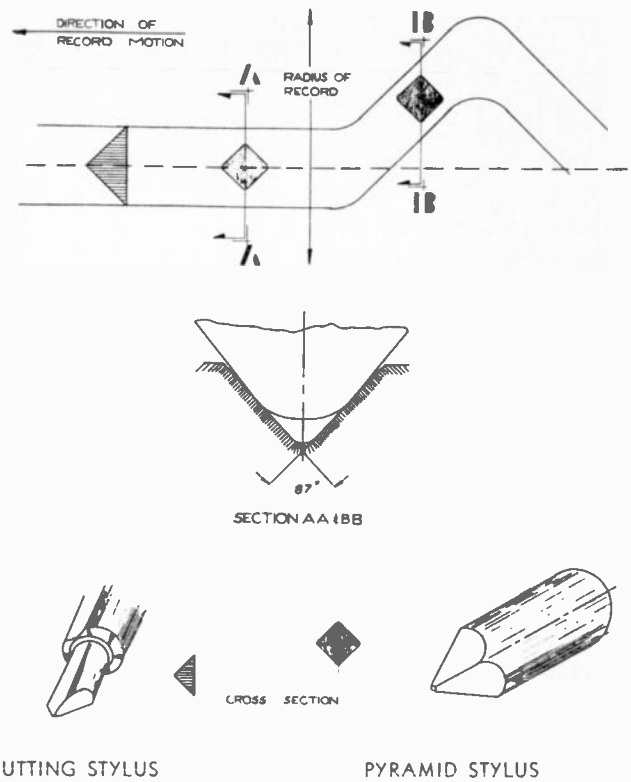

Fig. 7—Plan and sectional views of the assumed reproducing pyramid stylus and cutting stylus bearing relations in a typical laterally modulated groove.

While the pyramid stylus is a desirable improvement over the conventional hemispherical type of stylus, it does, unfortunately, impose additional manufacturing difficulties. This fact naturally will make the pyramid stylus more costly. It is possible, however, that the new pyramid stylus will prove to be cheaper to use than the conventional type, because of the longer service life to be expected for both needle and records. The noticeable improvement in reproduction cannot be valued in dollars. Much of the increased cost of producing a pyramid point in diamond is due to the care required to orient each stone before the process of grinding. In grinding a pyramid the softer planes of the diamond are ground away, leaving the ridges which contact the groove sidewalls coincidental with the crystalographic 111 molecular plane or the hardest diamond plane. Such a requirement for processing is a condition for optimum wear. After the stone has been properly processed and highly polished, it still requires careful orientation in its needle mounting. The rotational horizontal orientation of the stylus in its needle mounting is selected for best performance conditions to be met on the inside grooves of the record, where the reproduction is most critical to best stylus performance parameters.

In conclusion, the following features can be claimed for the pyramid stylus.

1) The pyramid stylus is shaped to reproduce phonograph records with a form which is compatible with the form of the cutting tool which made the original groove. Such a condition minimizes pinch effect, mistracking, second harmonic and needle talk. Reducing pinch effect further reduces the gouging out of grooves during the lifting and falling motion of low compliant needles with massive tone arms playing heavy modulation.

2) Tracing distortion is reduced to the amount theoretically equivalent to a 0.2-mil contact radius. This also insures that the reproduction of phonograph records with a pyramid will not progressively deteriorate in tonal quality as the stylus traces toward the inside grooves on the disk recording. The translation losses are imperceptible.

3) A stylus which is positively driven by smooth, gliding contact with the sidewalls of the record groove always fulfills conditions insuring low background noise and good transient response.

4) Reduced scratch level is impressive even to those who are not essentially discriminating listeners.

5) Heavy modulation will increase the tangential drag on a needle stylus. This condition often results in a type of waveform distortion called tangential distortion which is noticeable with pickups permitting a fore and aft motion. Any tangential weakness in the design of any pickup system will show improved performance with the new pyramid stylus.

6) Phonograph record disks last longer and retain lower background noise when played exclusively with pyramid styli.

Appendix

There have been expressions of interest received concerning experimental findings with the pyramid stylus. This section is added to present additional substantiating evidence and to reply to certain questions which have been asked.

Fig. 8 shows calculated relative harmonic distortion in per cent, which exists between generated harmonic components with changing values of styli contact radii, which reproduce recorded single sine waves. A 1-mil stylus was taken as unity and various smaller radii were used for comparison as described on the graph. All distortion values used in calculating the percentage relationships were direct readings taken from the chart in Fig. 6.

Fig. 9 shows a plotting of the average values found in Fig. 8. Consideration concludes that a contact radius of zero would be expected to produce a zero value of harmonics (disregarding record groove deformation). It is interesting to note that the average calculated values found in Fig. 8 fall very close to a straight line; this points directly to a value of zero distortion at zero radius in Fig. 9. Attention should also be drawn to the percentages of relative response improvement, which not only exist with the highest frequencies, but also exist uniformly over most of the audio spectrum, irrespective of the modulation velocity or groove velocity involved.

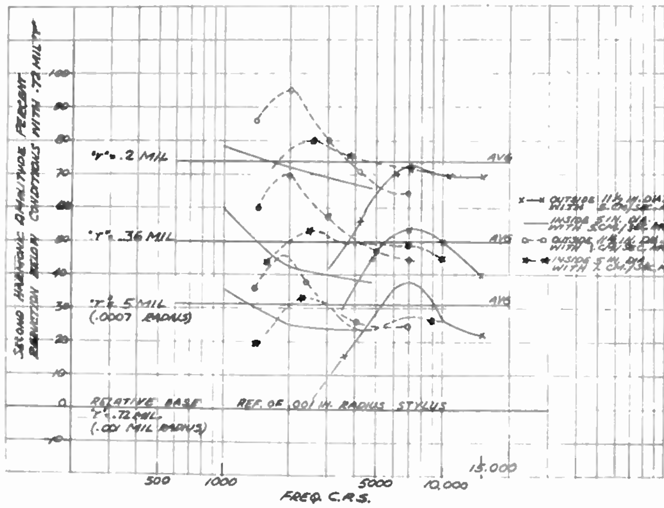

Fig. 8—A comparison of the contact radius of a 1-mil stylus with smaller contact radii playing recordings which have maximum and nominal values of recording velocity, at the locations of maximum and minimum groove velocity on the record. Comparisons are made with a series of overlaying curves which represent the percentage reduction of second harmonic distortion as found with stereo reproduction.

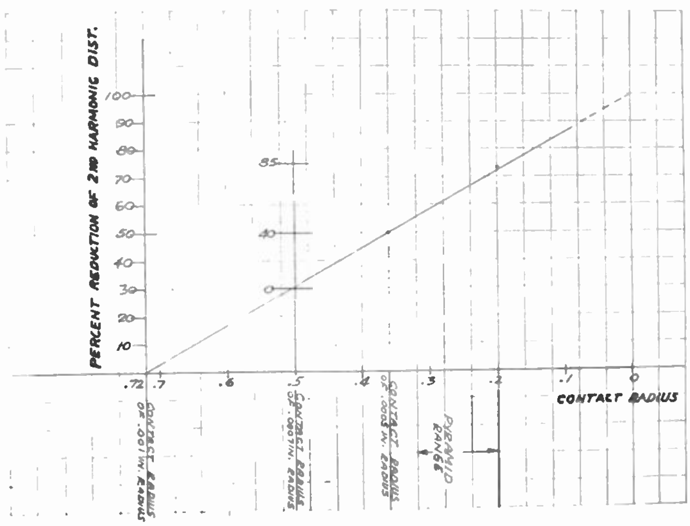

Fig. 9—Average relative percentage values of reduced distortion, with values of contact radii reduced to values generated by a 1-mil radius point (0.72-mil contact radius).

Some of the experimental findings are as follows.

1) A properly dimensioned pyramid stylus, having well polished ridge contact radii of 0.2 mil, is less destructive to a microgrooved disk recording than a conventional 0.7-mil radius. A 0.3-mil contact radius produces less damage than a 1.0-mil tip radius. These

findings are irrespective of needle force, length of play or general handling.

2) Frequency response curves, which are run with available test recordings, do not expose the true value of increased high frequency response obtained with the pyramid stylus playing commercial phonograph recordings of music. This is because frequency test recordings have their highest frequencies recorded on the outside grooves, where higher groove velocities produce longer wavelengths. A 12-inch acetate recording disk was used to record an 8-kc signal on the maximum and minimum recording diameters. Efforts were made to record this frequency with a modulation velocity of 5 cm/sec. A 0.7-mil stylus introduced a translation loss of 6 db on this recording, and the pyramid stylus showed a 1-db loss. This experiment requires repeating, however, because the test was made before proper stylus mounting procedures were developed.

3) Distortion readings were made with a model 300A Hewlett-Packard Harmonic Wave Analyzer using existing banded frequency test recordings of both monaural and stereophonic varieties. Here again such conventional test recordings were inadequate for the purpose, since they did not permit the true testing of conditions which exist near the center grooves. These tests did, however, result in establishing a technique for determining the best value of included angle and other related factors of mounting. These adverse conditions of measurement did, however, show a definite reduction of second harmonic on monaural recording with stereo pickup and somewhat less reduction of third harmonic components. Since a stereo pickup was being used for these measurements, a reduced second harmonic indicated less pinch effect. A reduced third harmonic indicated better riding of the centerline of the stylus on the centerline of the groove.

4) Stereo recordings played 300 times exclusively with pyramid styli at 5-gram needle force have shown no noticeable increased background noise due to record wear.

5) Stylus wear tests were conducted using metal pyramid styli, since no wear was detected using diamond styli. This may be due to the very careful orientation required in grinding which results in the extreme closeness of ridge radii to the hardest molecular crystalographic plane. Using metal styli permitted a saving in testing time and a good means of exploring the degree of performance deterioration with amount of point material removed. It was found that after playing each stylus the same number of times, the 0.7-mil stylus was beginning to shoulder, but the pyramid had not shouldered. Harmonic distortion measurements made at 5 kc with the 0.7-mil worn stylus produced a measured 12 per cent distortion while the pyramid produced a 4 per cent value. This data should only be accepted for relative evaluation, since better mounting techniques were developed after these tests were conducted.

Acknowledgment

The author wishes to express appreciation to R. Simon, of Battelle Memorial Institute, for his careful reading and useful comments on the paper. Acknowledgment is also made to the useful assistance, in some of the laborious calculations, of L. J. Holc. D. Clarke was very helpful through his assistance in obtaining carefully controlled experimental data.

Bibliography

[1] J. A. Pierce and F. V. Hunt, "Distortion in sound reproduction from phonograph records," *J. Soc. Motion Picture Engrs.*, vol. 31, pp. 157–186; August, 1938.
[2] H. E. Roys, "Distortion in phonograph reproduction," *RCA Rev.*, vol. 14, pp. 397–412; September, 1953.
[3] M. S. Corrington and T. Murakami, "Tracing Distortion in Stereophonic Disc Recording," 1958 IRE National Convention Record, pt. 7, pp. 73–81.
[4] P. C. Goldmark, B. B. Bauer and W. S. Bachman, "The Columbia compatible stereophonic record," IRE Trans. on Audio, vol. AU-6, pp. 25–28; March–April, 1958.
[5] H. F. Olson, "Acoustical Engineering," D. Van Nostrand Co., Inc., New York, N. Y.
[6] F. Langford-Smith, "Radiotron Designers Handbook," Published by the Wireless Press for Amalgamated Wireless Valve Co., Pty. Ltd., Sydney, Australia; 1952. Reprinted and distributed by R.C.A.
[7] W. S. Bachman, "Phonograph dynamics," *Electronic Industries*, pp. 86–89, 124–190.
[8] B. B. Bauer, "Distortion in phonograph reproduction caused by needle wear," *J. Acoust. Soc. Amer.*, vol. 16, pp. 246–253; April, 1945.
[9] F. V. Lewis and J. A. Hunt, "A theory of tracing distortion in sound reproduction from phonograph records," *J. Acoust. Soc. Amer.*, vol. 12, pp. 348–365; January, 1941.

Time-Frequency Scanning in Narrow-Band Speech Transmission*

D. L. SUBRAHMANYAM† AND G. E. PETERSON‡

Summary—**Two basic sampling methods for the transmission of intelligible speech with reduced channel capacity have previously been studied. The channel vocoder developed by Dudley is based on frequency band separation or quantization, and Schisser, and Fairbanks, Everitt, and Jaeger have developed a technique of time sampling. Sampling in both time and frequency offers a third major possibility which should be investigated.**

For this study a narrow-band speech transmission system was constructed which scans the time-frequency plane. The system scans in either a sinusoidal or in a sawtooth manner over a frequency range of 200 to 7000 cps. In order to achieve the scanning, a variable carrier frequency oscillator, having a frequency shift of ±30 per cent of the carrier frequency and an amplitude modulation of ±2 db, has been developed. The process of scanning leaves empty spaces in the time-frequency plane, and a four-channel time-delay system which employs the method of dielectric recording has been used to fill the gaps in the time-frequency plane by repeating the signal samples.

When signal reiteration is employed with this system, a score of 75 per cent of monosyllabic phonetically balanced word lists was obtained with a scanning filter of 1000-cps bandwidth and a sinusoidal scanning rate of 30 times per second. This intelligibility is appreciably higher than that which can be achieved with a fixed 1000-cycle filter located in the frequency region of maximum intelligibility.

I. Channel Capacity in Speech Transmission

Introduction

EXPERIMENTAL studies on the transmission of speech indicate that essentially complete intelligibility is obtained with a system having a frequency range of 100 to 7000 cps, and a signal-to-noise ratio of 40 or 50 db.[1] The transmission of uncoded speech with full fidelity in the frequency range of 100 to 7000 cps and a signal-to-noise ratio of 40 db requires a channel capacity of the order of 90,000 bits per second.

The speech signal is generated by a complicated mechanism whose acoustical output is a time function of the basic parameters of the system. According to estimates from speech analyses and measurements on actual narrow-band transmission systems, the information rate required to transmit the essential "information bearing elements" of speech is considerably less than the total channel capacity required for the full-band speech signal.[2]

Systems for the transmission of speech with low channel capacity requirements may be classified into two general groups: 1) speech parameter systems, in which speech is coded in terms of speech parameters, which may be either physiological or acoustical and, 2) physical parameter systems, in which the speech signal is sampled periodically in some combination of the time, frequency, and amplitude dimensions.

Time-Frequency Plane

Signals may be described in either the time or the frequency domain. The transformation from the description in terms of time to the description in terms of component frequencies is based on the Fourier reciprocal integral relations

$$S(t) = \int_{-\infty}^{+\infty} S(f)e^{j2ft}df$$

and

$$S(f) = \int_{-\infty}^{+\infty} S(t)e^{-j2\pi ft}dt. \quad (1)$$

The analysis in time is performed by an ideal oscilloscope, and that in frequency by a system of infinitely graded filters with no damping.

In a two-dimensional orthogonal coordinate system with frequency and time as coordinate axes, a sinusoidal frequency $f=f_0$, which extends in time from $-\infty$ to $+\infty$, is a line parallel to the time axis; an impulse function at time $t=t_0$, which extends in frequency from $-\infty$ to $+\infty$, is a line parallel to the frequency axis. The time-frequency plane is represented in Fig. 1, which may be referred to as the "diagram of information."

Gabor[3] has shown that the signal which occupies the minimum area $(\Delta f)(\Delta t)=\frac{1}{2}$ in the time-frequency plane is a harmonic oscillation modulated by a probability function. In complex form, the signal may be expressed as

$$\psi(t) = e^{-\alpha^2(t-t_0)^2}e^{j(2\pi f_0 t+\phi)} \quad (2)$$

where

α = sharpness factor of the pulse,
t_0 = time of occurrence of the peak of the pulse,
f_0 = frequency of the modulating oscillation, and
ϕ = phase constant of the modulating oscillation.

* Manuscript received by the PGA, August 26, 1959; revised manuscript received October 7, 1959. Presented at the Convention of the Acoustical Society of America, Chicago, Ill., November 20, 1958.

† C.E.E.R.I., Pilani (Rajasthan), India; formerly with Speech Research Laboratory, University of Michigan, Ann Arbor, Mich.

‡ Speech Research Laboratory, University of Michigan, Ann Arbor, Mich.

[1] N. R. French and J. C. Steinberg, "Factors governing intelligibility of speech sounds," *J. Acoust. Soc. Am.*, vol. 19, pp. 910–919; January, 1947.

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

[3] D. Gabor, "Theory of communication," *J. IEE*, vol. 93, pt. III, pp. 429–457; November, 1946.

The spectrum of the signal has the same analytical form and is given by

$$\Phi(f) = e^{-(\pi/\alpha)^2(f-f_0)^2} e^{-j(2\pi t_0 f-\phi)}. \quad (3)$$

The constant α is related to Δf and Δt by

$$\Delta t = \sqrt{(\pi/2)}1/\alpha \quad \text{and} \quad \Delta f = \alpha/\sqrt{2\pi}. \quad (4)$$

The signal represented by (2) is called the "elementary signal." The elementary signal may be represented as a rectangle with sides Δt and Δf centered around (t_0, f_0) as shown in Fig. 1. Any other signal within the interval (t_1, t_2) and confined to the frequency range (f_1, f_2) may be expanded in terms of elementary signals such that corresponding rectangles occupy the appropriate time-frequency region.

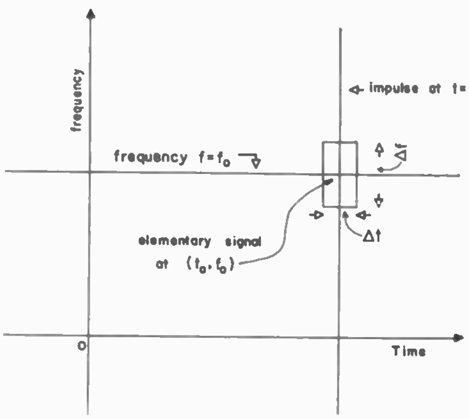

Fig. 1—Time-frequency plane showing an impulse function at $t=t_0$, a harmonic oscillation of frequency $f=f_0$, and an elementary signal at (t_0, f_0).

Speech Parameter Systems

From the point of view of acoustics, speech production consists of: 1) the generation of complex air vibrations, either random or periodic at the glottis or higher along the vocal passageways; 2) selective transmission of the vibrations by the vocal tract due to the natural resonances of the cavities; and 3) radiation of the modified sound from the openings of the mouth and nose. In the process of speech production the parameters of the source, and of the transmission and radiating systems, are normally changing in time.

An acoustical analysis of the speech signal in the time-frequency plane may be made with a sound spectrograph.[4] Fig. 2 shows the spectrogram of the phrase "narrow-band speech transmission." Due to the selection of harmonics by resonances of the vocal tract, vowel and vowel-like sounds appear as bars, which are known as formants, and are especially prominent in the first syllable of *narrow*, *band*, and *transmission*. The positions and movements of these formants provide much information about the time function of the articulatory patterns of speech. The random striations in the spectrogram, in general, represent fricative sounds in speech. The [*s*]'s in Fig. 2 illustrate this type of pattern particularly well. Sometimes a voicing accompanies friction and in such cases the striations may be superimposed upon formant patterns. For plosives, gaps or blank spaces appear in the spectrogram, as for the [*p*] of *speech*. When the plosive is voiced, a low-frequency resonance results as shown in the [*b*] of *band*.

Narrow-band systems of the speech parameter type[5–7] may employ physiological data such as the positions of the tongue, lips, soft-palate, etc., and acoustical data such as the fundamental voice frequency, the type of exciting source, position and movement of formants, etc. The data thus obtained form the control signals which are transmitted over the channel. The control currents at the receiving end operate on a synthesizer to recreate the signal.

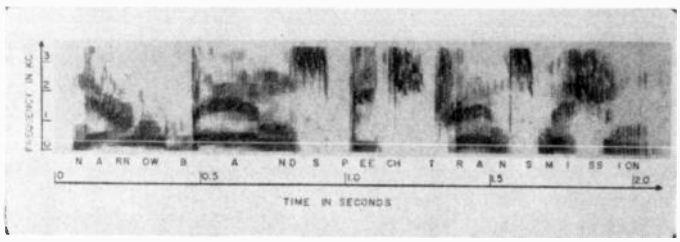

Fig. 2—Spectrogram of the phrase "narrow-band speech transmission."

Physical Parameter Systems

Spectrographic displays of speech show that only a limited extent of the time-frequency plane is normally involved in speech. For example, the plosive bursts cover the spectrum but the plosive gaps leave much unused time; the vowel and vowel-like sounds are relatively continuous but involve only a limited use of the frequency domain.

Narrow-band systems of the physical parameter type carry out a systematic sampling of the speech signal in the time-frequency plane. Low-pass and band-pass filters are techniques of sampling parallel to the time axis,[8] and interrupted speech[9] and time reiteration[10] are methods of sampling parallel to the frequency axis.

From the examination of sound spectrograms it ap-

[4] W. Koenig, H. K. Dunn, and L. Y. Lacy, "The sound spectrograph," *J. Acoust. Soc. Am.*, vol. 17, pp. 19–49; July, 1946.

[5] W. Lawrence, "The synthesis of speech from signals which have a low information rate," in "Communication Theory," Academic Press, Inc., New York, N. Y., pp. 460–471; 1953.

[6] 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.

[7] C. R. Howard, S. H. Chang, and M. J. Carrabes, "Analysis and synthesis of formants and moments of spectra" (abstract), *J. Acoust. Soc. Am.*, vol. 28, p. 768; July, 1956.

[8] H. Dudley, "Remaking speech," *J. Acoust. Soc. Am.*, vol. 11, pp. 169–177; October, 1939.

[9] G. A. Miller and J. C. R. Licklider, "The intelligibility of interrupted speech," *J. Acoust. Soc. Am.*, vol. 22, pp. 167–173; March, 1950.

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

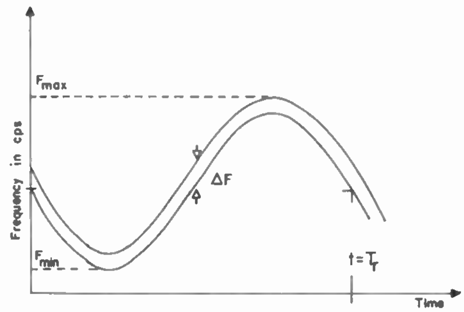

Fig. 3—Sinusoidal scanning in the time-frequency plane.

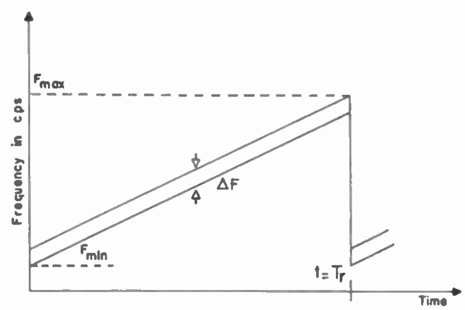

Fig. 4—Sawtooth scanning in the time-frequency plane.

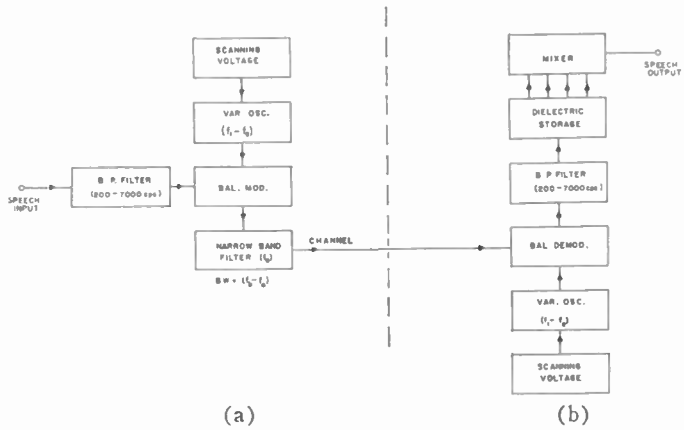

Fig. 5—Schematic block diagram of the narrow-band speech transmission system. (a) Analyzer (transmitter); (b) synthesizer (receiver).

pears that scanning in the time-frequency plane might provide more adequate and less discontinuous speech samples for reiteration than sampling in time or frequency alone. Accordingly, a system was developed which scans the time-frequency plane in a periodic manner. Investigations of the scanning method for a narrow-band speech transmission system are described in the remaining sections of this paper. Section II explains the narrow-band scanning method in the time-frequency plane and the experimental technique employed. Section III briefly describes the scanning filter of the narrow-band speech transmission system, and Section IV describes the dielectric storage system employed for reiteration. Section V outlines the electrical performance of the over-all system, and an evaluation of the system in transmitting speech is presented in Section VI.

II. Narrow-Band Scanning Method

In the system which was constructed, the time-frequency plane is scanned in either a sinusoidal or sawtooth manner. Fig. 3 shows the sinusoidal scanning, and Fig. 4, the sawtooth scanning. In Figs. 3 and 4, the range of frequency scanned is $(F_{max}-F_{min})$ cps, and the bandwidth of the scanning filter is ΔF cps. The scanning frequency is the reciprocal of T_r, where T_r is the repetition period.

Experimental Technique for Scanning the Time Frequency-Plane

Fig. 5 shows the block diagram of the experimental system. The input band-pass filter selects frequencies in the range of 200 to 7000 cps from the speech signal. The output of this filter modulates a variable carrier oscillator. The control voltage, which is either sinusoidal or sawtooth, controls the frequency of the variable carrier oscillator. The narrow-band filter selects the lower sideband from the output of the balanced modulator.

At the synthesizer (receiver) the received signal undergoes demodulation with a variable carrier oscillator whose frequency is made to vary in synchronization with the variable carrier oscillator at the transmitter. A band-pass filter of 200 to 7000 cps separates the speech signal from the output of the demodulator.

The process of sampling in the time-frequency plane leaves blank spaces, and a dielectric storage system is used to fill the empty spaces by repeating the signal that has occurred at preceding instants. The storage system has four playback heads and hence the original signal is repeated three times. The output of the storage system is added in a four-channel mixer to complete the reconstruction of the speech signal.

Parameter Relations in the Narrow-Band Scanning Sysstem

Let the frequency of the carrier oscillator be varied from f_1 cps to f_2 cps. The scanning rate f_r is the reciprocal of T_r. These are shown in Fig. 6 for sawtooth scanning.

Let the width of the narrow-band filter be (f_b-f_a) cps where f_b and f_a, expressed in cps, are 3-db points of the narrow-band filter, and f_0, expressed in cps, is its center frequency.

From the geometrical representation of Fig. 6(a) it follows that

$$(f_1 - F_{min}) = f_b$$

and

$$(f_2 - F_{max}) = f_a. \tag{5}$$

From the geometry of Fig. 6(b) it is seen that the number of repetitions required is given by

$$N = T_r/x. \tag{6}$$

Again from Fig. 6(b)

$$\frac{f_b - f_a}{x} = \tan\theta = \frac{F_{max} - F_{min}}{T_r}. \tag{7}$$

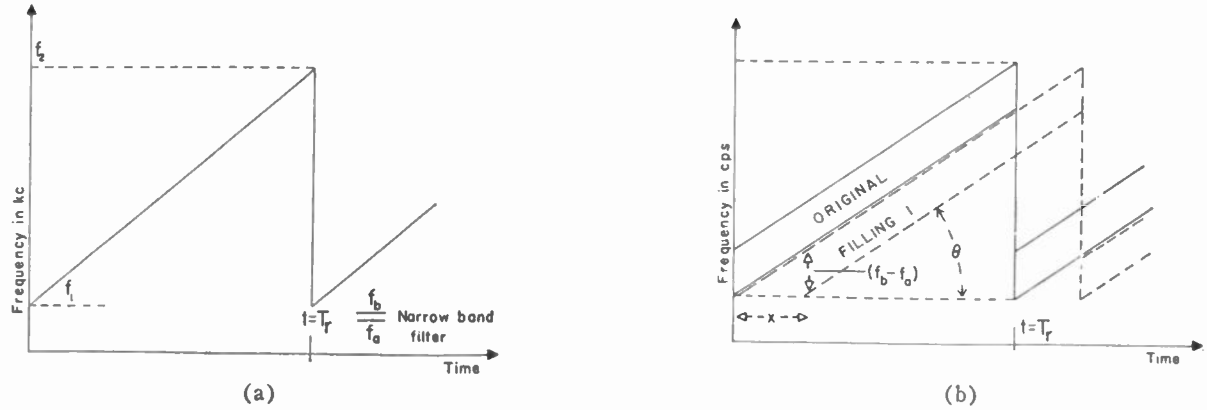

Fig. 6—Sawtooth variation of the carrier frequency and sawtooth scanning in the time-frequency plane. (a) Variation of carrier frequency; (b) Time-frequency plane of speech signal.

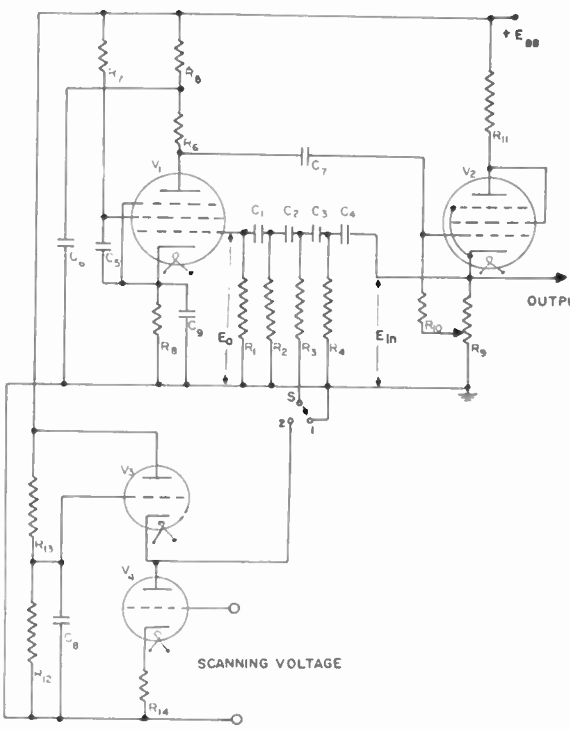

Fig. 7—Electronically controlled variable phase-shift oscillator.

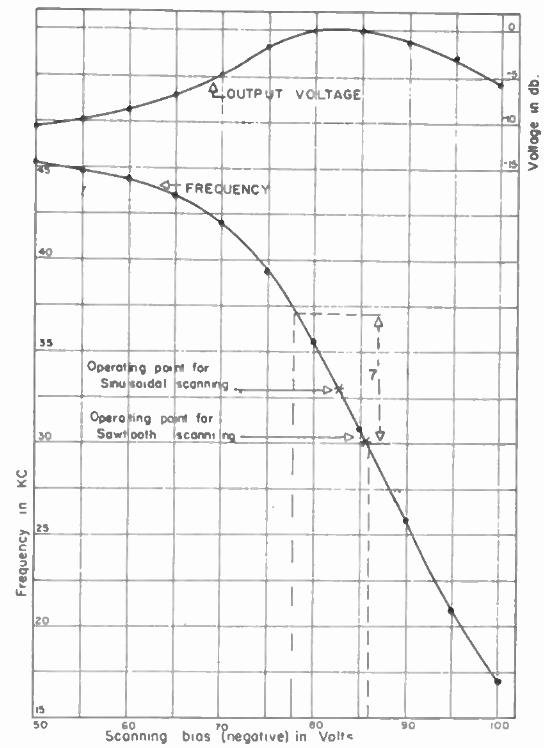

Fig. 8—Plot of oscillator frequency and output voltage as a function of scanning bias.

Therefore,

$$N = \frac{F_{\max} - F_{\min}}{f_b - f_a}. \tag{8}$$

Thus, the total number of repetitions required at the receiver to fill the time-frequency plane is equal to the channel-width reduction factor.

III. The Heterodyne Filter

The Variable Frequency Carrier Oscillator

The relations given in (5) determine the range of carrier frequency scan necessary for varying the audio frequency signal from 200 to 7000 cps. The carrier frequency scan is given by

$$\begin{aligned}(f_2 - f_1) &= (F_{\max} - F_{\min}) - (f_b - f_a)\\ &= (7000 - 200) - \Delta F \end{aligned} \tag{9}$$

where ΔF is the bandwidth of the narrow-band filter.

The variable frequency carrier oscillator, of the resistance-capacitance type, employs a four mesh phase-shift network in the feedback circuit.[11,12] Fig. 7 shows the basic circuit of such an oscillator.

Fig. 8 shows plots of the oscillator frequency and its output voltage as a function of the grid voltages (scanning voltage) of V_4 (of Fig. 7). Actually the curves of Fig. 8 are for the variation of both R_3 and R_4.

A carrier frequency range from 30 kc to 37 kc was selected in order to keep the lower sideband out of the audio range (above 20 kc). The operating points for sinusoidal and sawtooth scanning are indicated in Fig. 8.

[11] E. L. Ginzton and L. M. Hollingsworth, "Phase-shift oscillators," Proc. IRE, vol. 29, pp. 43–49; February, 1941.

[12] M. Artzt, "Frequency modulation of resistance capacitance oscillators," Proc. IRE, vol. 32, pp. 409–414; July, 1942.

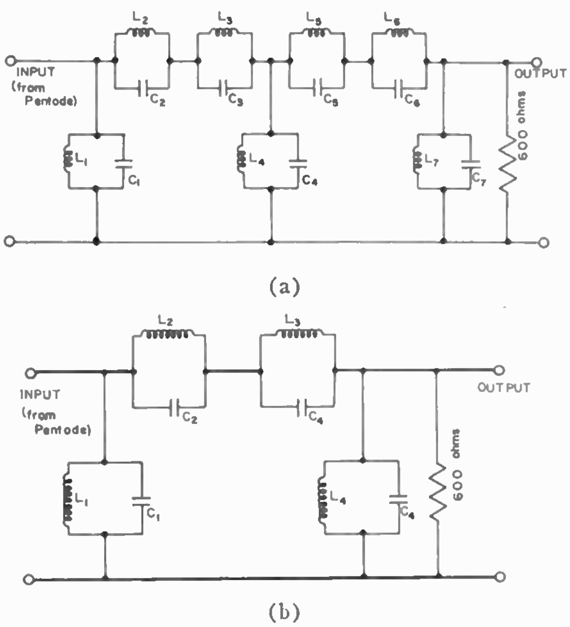

Fig. 9—(a) Section I of the narrow-band filter. (b) Section II of the narrow-band filter.

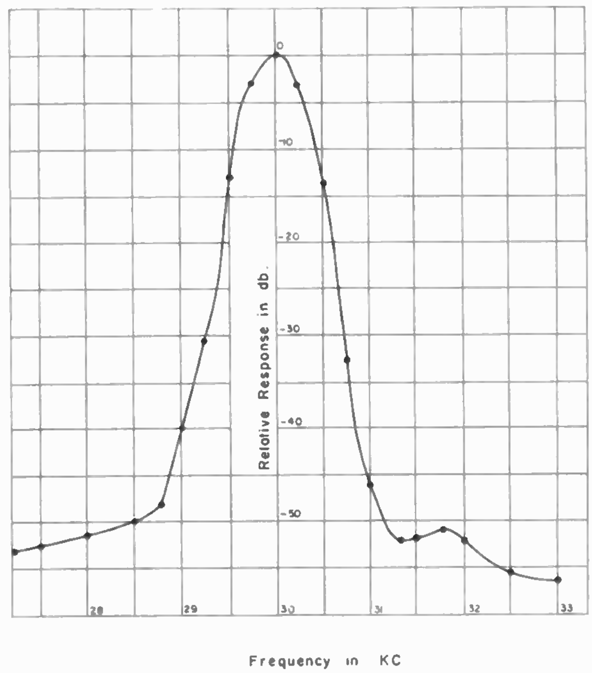

Fig. 10—Response curve of the narrow-band filter.

These are fixed by applying the required negative voltages or scanning bias. The scanning voltage is superimposed on the scanning bias. The amplitude of the scanning voltage determines the frequency range scanned. The amplitude modulation of the oscillator over the frequency range of interest is less than $\frac{1}{4}$ db. The complete circuit of the phase-shift oscillator is included in Fig. 11.

The Balanced Modulator

The variable carrier oscillator is modulated by the speech signal. A balanced ring modulator circuit[13,14] was used to aid in suppressing the carrier.

The Narrow-Band Filter

The center frequency of the filter was chosen to be 30 kc, and the bandwidth at the 3-db points to be 500 cps. The lower and upper half-power frequencies are $f_a = 29.75$ kc, and $f_b = 30.25$ kc. In order to reduce further the carrier frequency, the upper wing of the filter should be steep. A second major consideration in the design of the narrow-band filter was its bandwidth in relation to the scanning rate. The response of a linear resonant system to a sinusoidal driving function having a linear variation of frequency with time has been studied previously.[15] For the case of a Gaussian filter, the relative amplitude A_0 of the filter is 92 per cent when

$$S/b^2 = 0.3 \tag{10}$$

where S is the sweep rate in radians per second per second, b is the bandwidth of the filter in radians per second, and A_0 is defined so that it has a value of unity for an infinite pulse length and zero sweep rate. Eq. (10) sets the upper limit for the rate of scanning when the bandwidth of the filter is fixed. The lower limit for scanning is set by the structure of the speech signal and the nature of speech perception.

The narrow-band filter uses the design of the Tchebycheff approximation in both pass-and-stop-bands; it consists of two sections and each section is designed to be driven by a constant current source. Fig. 9 shows the components of the two individual sections, and Fig. 10 shows the response of the cascaded filter.

The Demodulator

The demodulation of the single-sideband system is accomplished by superimposing upon the single sideband a local oscillation whose frequency is varying in the same manner and in synchronism with that at the analyzer.[16,17] In the balanced demodulator the oscillator input is applied in push-pull, whereas the signal input is applied in parallel. If the signal voltage is small in comparison with the oscillator voltage, the output of the balanced frequency converter circuit consists of the required audio signal as well as other harmonic frequencies. The latter are eliminated by the 200- to 7000-cps band-pass filter.

Circuit of the Narrow-Band Scanning System

Fig. 11 shows the schematic circuit diagram of the narrow-band scanning system. The tubes V_1 and V_4

[13] A. Hund, "Frequency Modulation," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 152–154; 1942.

[14] F. E. Terman, "Radio Engineers Handbook," McGraw-Hill Book Co., Inc., New York, N. Y., pp. 552–553; 1943.

[15] H. W. Batten, R. A. Jorgensen, A. B. Macnee, and W. W. Peterson, "The response of a panoramic receiver to CW and pulse signals," PROC. IRE, vol. 42, pp. 948–956; June, 1954.

[16] Terman, *op. cit.*, pp. 567, 578.

[17] E. W. Herold, "The operation of frequency converters and mixers for superheterodyne reception," PROC. IRE, vol. 30, pp. 84–103; February, 1942.

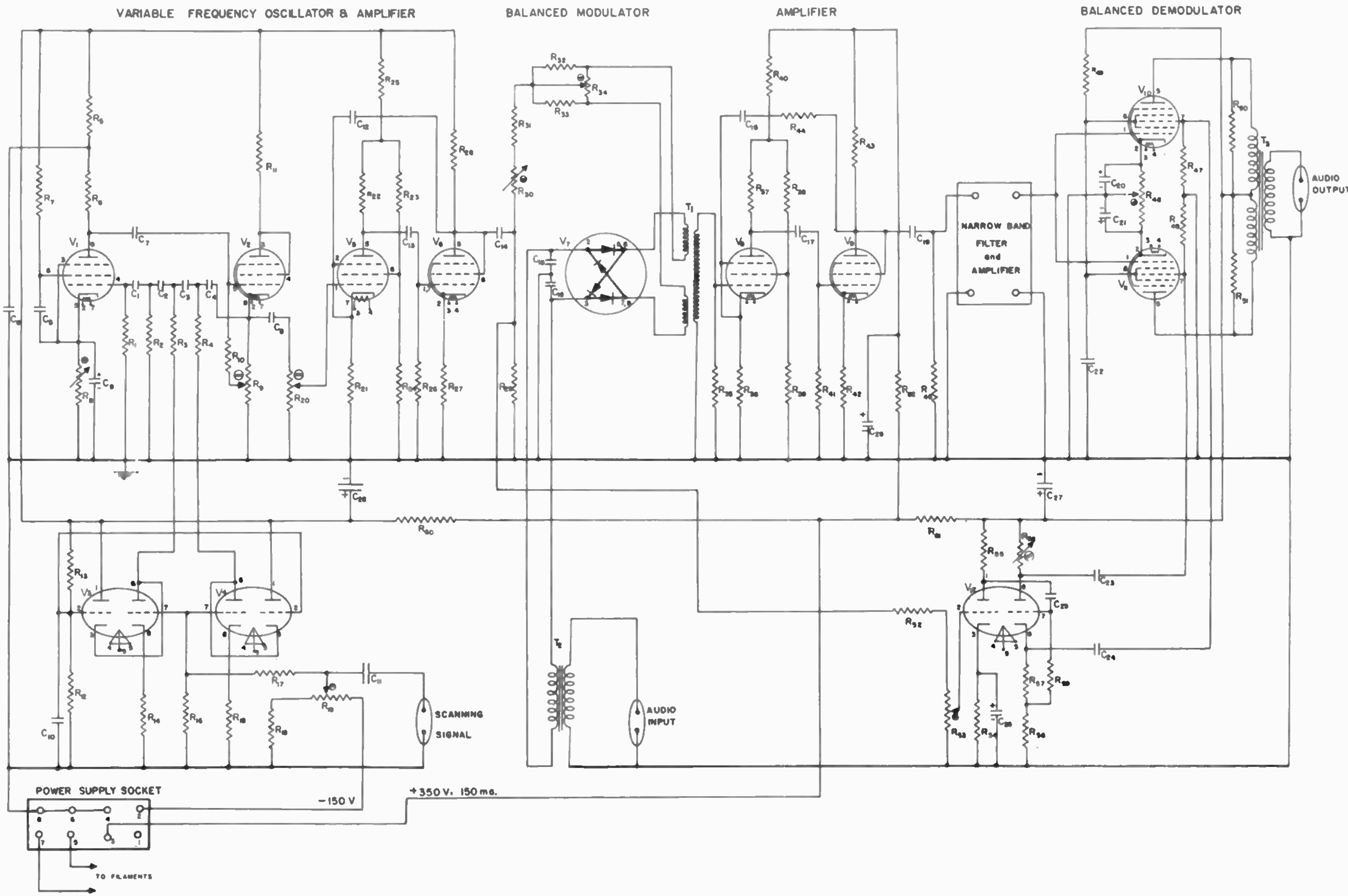

Fig. 11—Schematic circuit diagram of the narrow-band scanning system.

constitute the variable frequency carrier oscillator; the tubes V_3 and V_4 operate as variable resistances in series with R_3 and R_4 of the four mesh phase-shift network of the oscillator. The scanning bias is adjusted by varying the potentiometer R_{19}. The scanning voltage is superimposed upon the scanning bias; the amplitude of the scanning voltage determines the frequency range scanned.

The tubes V_5 and V_6, designed as a negative feedback amplifier with output impedance of 600 ohms, isolate the balanced modulator from the variable frequency oscillator. The output of the balanced modulator is amplified by the two-stage negative feedback amplifier, V_8 and V_9. The output of this amplifier is fed into the narrow-band filter. In the present investigations, the output of the narrow-band filter is fed in parallel to the number 1 grids of V_{10} and V_{11} (pentagrid converters). The number 3 grids of these tubes are fed in push-pull by the second half of the twin triode V_{12}, which is used as a phase splitter. The first half of V_{12} is used as a conventional triode amplifier, and has its input from the isolating amplifier (V_5 and V_6) through the voltage divider R_{52} and R_{53}. Also, in these investigations, the variable carrier oscillator at the analyzer serves as the variable carrier oscillator at the synthesizer. The output of the system is obtained from the 600-ohm secondary of the audio transformer T_3.

IV. The Dielectric Storage System

It is seen from (10) that a filter of 500-cps bandwidth can be swept at a maximum rate of about 70 times per second. This corresponds to a scanning period of approximately 14 msec. In order to fill the blanks completely under these conditions it would be necessary to repeat the received signal 13 times; *i.e.*, a storage system with 14 playback heads, each playback head separated from its neighbor by 1 msec, is required, as indicated by (8). Consequently, the minimum time delay in the storage system required for a scanning filter of 500-cps bandwidth is 1 msec.

In order to obtain time delays of the order of 1 msec with existing high-quality magnetic recorders and playback heads, the tape speeds must be enormously high. At these high speeds, out-of-contact recording must be employed which introduces problems of level and equalization. As a result, other signal recording methods were considered. Dielectric recording[18,19] offered the possibility of small components, so that closer head spacings could be employed; it also provides greater flexibility and ease of reproduce head adjustment.

Dielectric recording makes use of the principle that a charge, proportional to an applied voltage, can be stored on a moving dielectric medium. A voltage proportional to the charge can then be induced in a capacitance-coupled pickup probe. The over-all performance of the dielectric recorder is determined by the process employed for transferring the charge onto the surface of the dielectric medium, the characteristics of the medium, and the process employed for recovering the stored charge from the medium.

Recording Process

In the process of recording, a charge proportional to a signal voltage is transferred to the dielectric medium. This is achieved by creating a conducting path of ionized air by means of a stable RF corona discharge with the help of an auxiliary electrode; the signal voltage is then used to control the drift of ions in the gap.[20] The surface of the moving dielectric medium is charged to the average potential of the region, *i.e.*, to the instantaneous value of the signal voltage.

Basic considerations include the ionization of air at atmospheric pressure and the recombination properties of such ionized air. In a qualitative way it can be said that the total conductivity of the ionized region controls storing of the charge on the dielectric medium and erasing of the previously stored charge; the sharpness of the boundary of the ionized region determines the minimum wavelength response, and the fluctuations of the potential of the conducting region near the surface of the medium determines the noise level on the dielectric.

Dielectric Medium

The charge stored on the surface of the dielectric medium decays, due to the volume resistivity and the surface resistivity of the medium. The decay of the charge due to volume resistivity is independent of both the geometry of the medium and the distribution of the charge.[21] Lucite, which is used in this application, has a decay time of about 80 hours and is quite adequate for short-time storage. The decay time due to surface resistivity depends on both the geometry and the distribution of the surface charge. For a given dielectric material, the decay time due to surface resistivity rapidly decreases with decrease of wavelength on the medium. This is of practical importance in the design of a recording system since it determines the surface velocity of the recording medium for a specified high-frequency response.

Surface resistivity is influenced by the nature of the surface, its cleanliness, relative humidity, etc. Practical expedients, such as coating the surface with silicon oil, may be employed to increase surface resistivity.

Reproducing Process

The process of reproducing consists of obtaining an induced voltage on a pickup probe located a short distance from the surface of the moving dielectric medium.

[18] V. Anderson, "A recording-reproducing system using a dielectric storage medium" (abstract), *J. Acoust. Soc. Am.*, vol. 27, p. 1004; September, 1955. See also Scripps Institution of Oceanography, San Diego 52, Calif., Ref. SIO 55-6.

[19] V. C. Anderson, "Dielectric recorder," *Rev. Sci. Instr.*, vol. 28, pp. 504–509; July, 1957.

[20] B. L. Loeb, "Basic Processes or Gaseous Electronics," University of California Press, Los Angeles, Calif., pp. 2–4; 1955.

[21] W. R. Smythe, "Static Dynamic Electricity," McGraw-Hill Book Co., Inc., New York, N. Y., p. 451; 1939.

The wavelength response of the pickup probe depends on its diameter and also on the ratio of probe separation to the thickness of the dielectric medium. The total induced charge, and hence the output voltage, depends on the radius of the probe wire.

Four-Channel Time-Delay System

The storage system employed in the present study makes use of the dielectric recording and reproducing described above, and contains four independent channels.

Recording Head Assembly

The recording head design is based on a Stupakoff seal. A 5-mil platinum wire is mounted axially in the inner tube of the seal. Approximately 4000 volts at 150 kc is applied to the outer shell. The capacity between the outer shell and the inner tube couples the high-frequency high voltage to the 5-mil wire and establishes the stable RF corona discharge between the protruding tip of the wire and the Lucite on the aluminum drum. Surrounding the platinum wire is a stainless steel ring to which the signal to be recorded is applied. The constructional details of a recording head are shown in Fig. 12.

Pickup Head Assembly

The pickup probe consists of a wire with a diameter of about 15 mils; the wire is shielded and the tip of the probe is located at a distance of about 10 mils from the moving dielectric medium. The voltage output of the probe is proportional to the induced charge. Fig. 13 shows a playback head assembly.

Mechanical Assembly

The mechanical unit of the four-channel time delay system is shown in Fig. 14. An aluminum drum 12 inches in diameter and 5 inches in height is fitted with a Lucite cylinder of $\frac{1}{8}$-inch wall thickness on which the charge is stored. The steel shaft is coupled to a $\frac{3}{4}$-hp, 1740-rpm induction motor. A circular aluminum disk which has four concentric circular tracks is mounted concentric with the shaft and its housing. One recording head assembly and one erasure head assembly, each similar in construction and containing four heads, are located in tracks 1 and 2. In each of the four tracks, a playback head assembly is located. Each playback head assembly is movable over almost 360° in its respective track and provides delays ranging from 0.69 to 27 msec. It is possible to feed a signal to all four recording heads in parallel; also, it is possible to feed a signal successively from one playback head to the next to obtain a time delay of approximately four times that available with a single track.

Electronics

Fig. 15 shows the circuit diagram of the high-voltage power supply. It is designed around a tuned plate oscillator employing the beam power tube 6L6 and a high-voltage transformer.[22,23] A second oscillator circuit supplies power to the erase heads. Fig. 16 shows the circuit diagram of the recording amplifier; this amplifier delivers about 280 volts rms to the audio ring.

In order to avoid the shunting effect of stray capacitances on the pickup probe, a high impedance preamplifier is mounted on the horizontal bar. It is a cathode follower, as shown in Fig. 17.

The frequency response of each channel with high frequency equalization is from 70 to 10,000 cps. The signal-to-noise ratio is approximately 47 db, and the system is linear within ± 1 db in that range. The output of the four-channel dielectric storage system is mixed in a conventional four-channel mixer as indicated in the block diagram of Fig. 5.

V. Performance Characteristics

The response curves of the narrow-band speech transmission system were obtained for sinusoidal and random noise inputs. These curves show the behavior of the system when scanning the time-frequency plane in either a sinusoidal or sawtooth manner.

Sinusoidal Input

It has already been noted that the scanning bias determines the quiescent or static frequency of the variable carrier oscillator, and the peak-to-peak amplitude of the scanning voltage determines the range of frequency scanned. In the sinusoidal scanning of the speech signal over the range of 200 to 7000 cps, the scanning bias is adjusted so that the quiescent carrier frequency is 33.5 kc. As the frequency of a sinusoidal input signal is varied, the output follows the response of the narrow-band filter. Fig. 18(a) shows the spectrogram of the output when the signal input is 3.5 kc and when there is no scanning voltage. It shows a horizontal line corresponding to 3.5 kc.

Random Noise Input

If the scanning bias is adjusted so that the carrier frequency is 33.5 kc, and random noise is fed into the system, it is seen from the schematic block diagram of Fig. 5 that the output following the audio filter (of the synthesizer) should contain a band of frequencies of 500-cps width centered around 3.5 kc. The spectrogram obtained under the above conditions is shown in Fig. 18(b).

Spectrograms of the output of the system following the audio filter when the scanning signal is sinusoidal in the range of 200 to 7000 cps are shown in Fig. 18(c) and Fig. 19(a). Fig. 18(c) shows the result when the time-frequency plane is scanned in a sinusoidal manner at the rate of 5 times per second, and Fig. 19(a) shows the result for a rate of 30 times per second.

A spectrogram of the output of the system with sawtooth scanning at the rate of 6 times per second is shown in Fig. 19(b).

[22] O. H. Schade, "Radio frequency operated high voltage supplies for C.R. tubes," Proc. IRE, vol. 31, pp. 158–163; April, 1943.

[23] R. S. Mautner and O. H. Schade, "Television high voltage RF supplies," *RCA Rev.*, pp. 43–81; March, 1947.

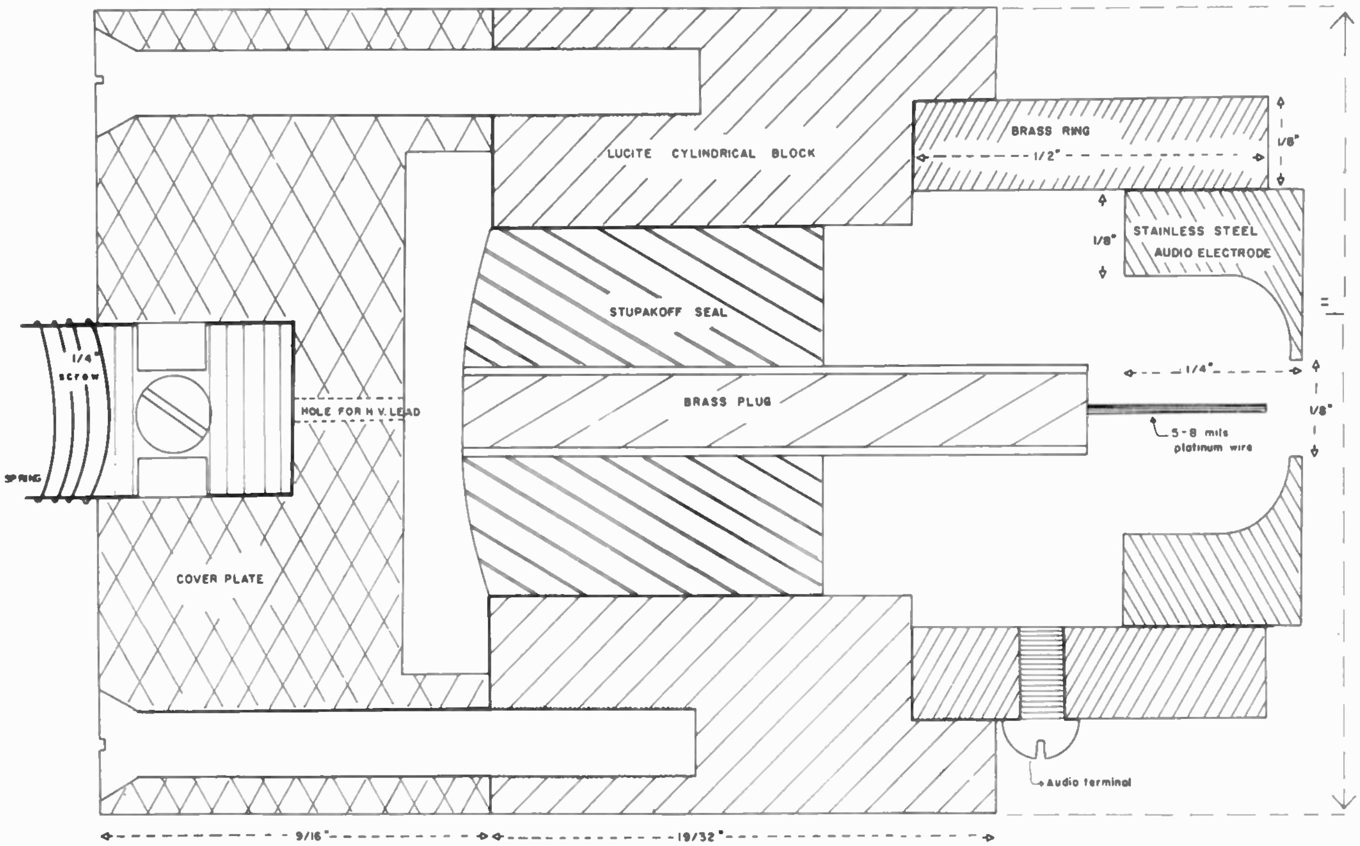

Fig. 12—Constructional details of the recording head.

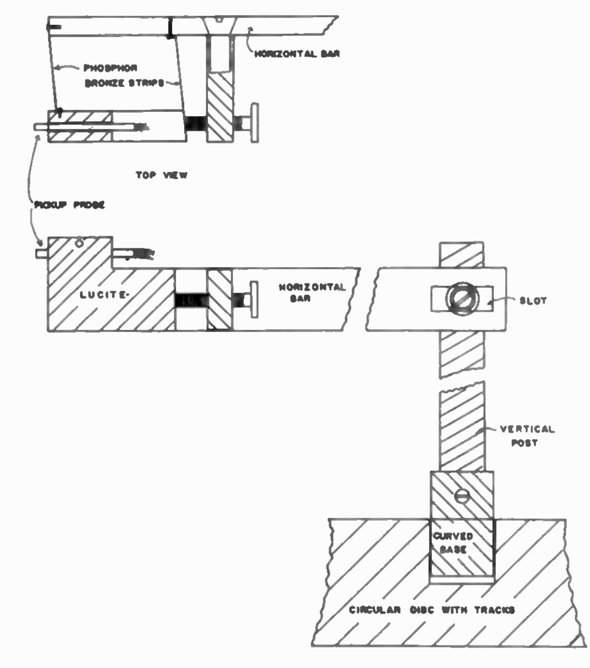

Fig. 13—Constructional details of the playback head assembly.

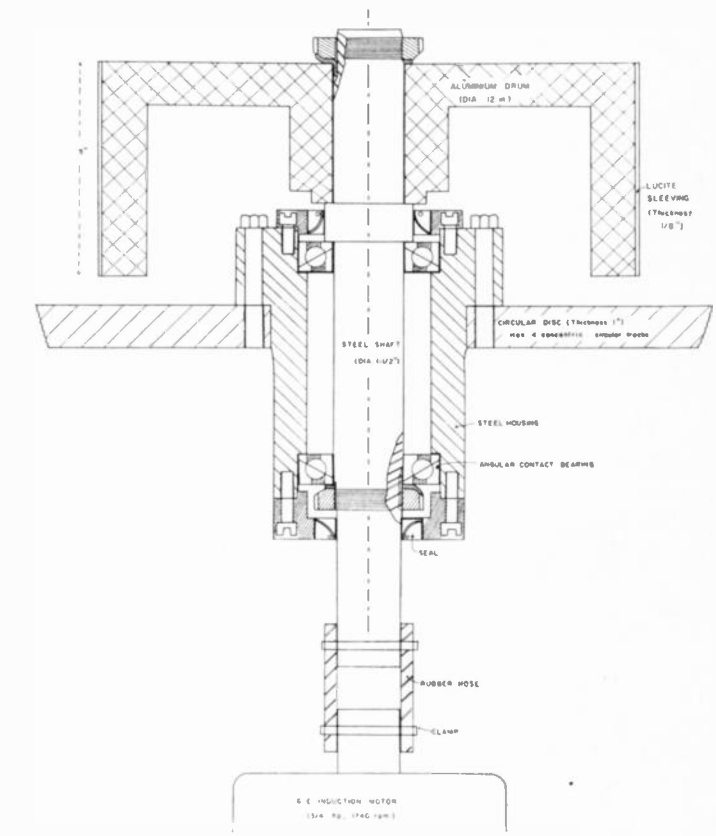

Fig. 14—Vertical section of the four-channel time-delay system.

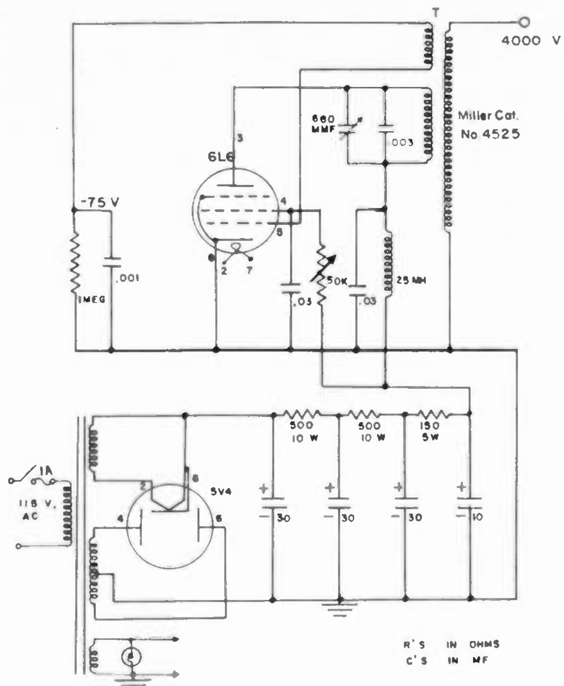

Fig. 15—Schematic circuit diagram of the high-voltage power supply.

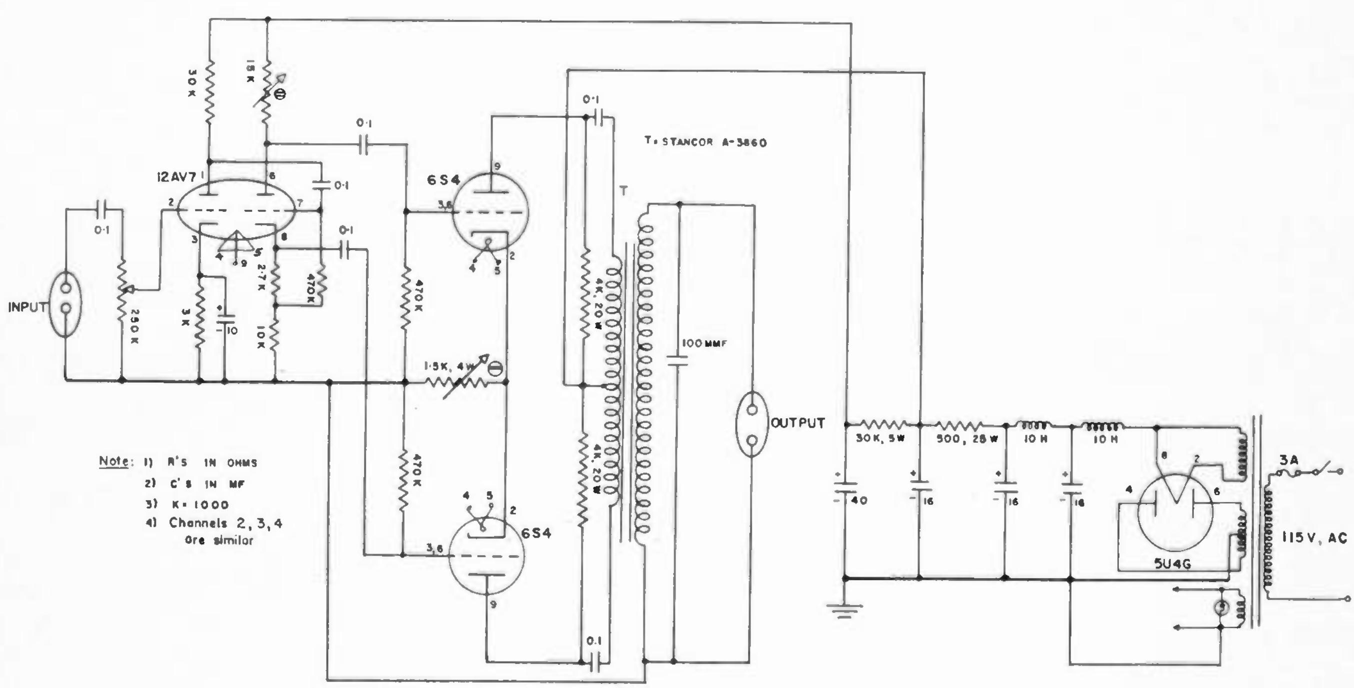

Fig. 16—Schematic circuit diagram of the recording amplifier.

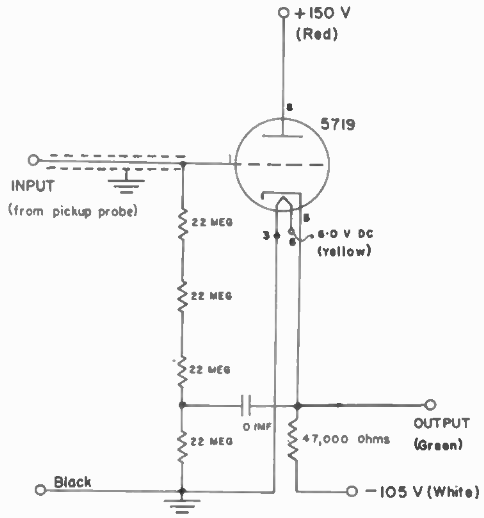

Fig. 17—Schematic circuit diagram of the preamplifier for the pickup probe.

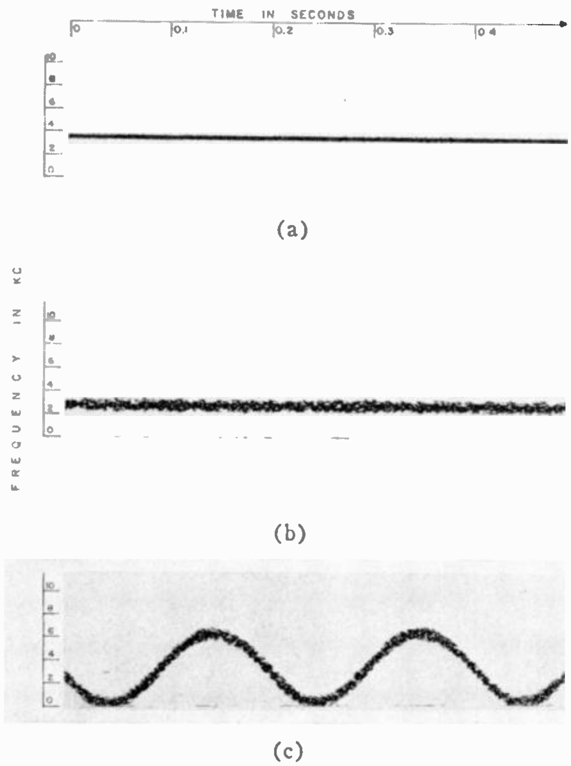

Fig. 18—Sound spectrograms showing the output of the narrow-band scanning system. (a) 3.5-kc sinusoidal input. No scanning. (b) Random noise input. No scanning. (c) Random noise input. Sinusoidal scanning at 5 times per second.

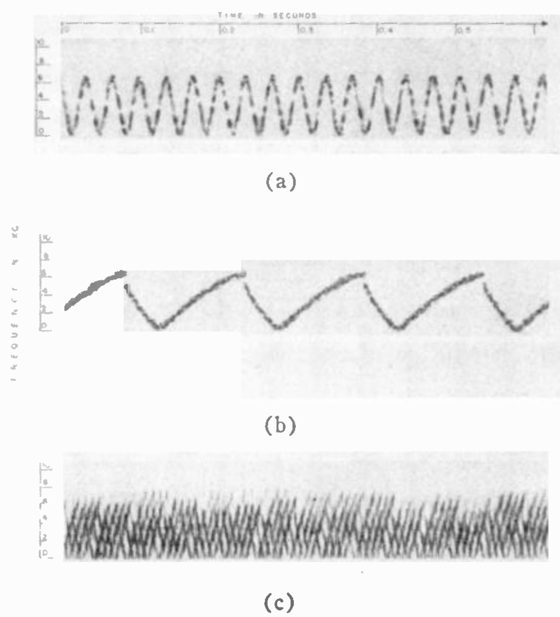

Fig. 19—Sound spectrograms of the output of the narrow-band transmission system with random noise input. (a) Sinusoidal scanning at 30 times per second. (b) Sawtooth scanning at 6 times per second. (c) Sawtooth scanning at 30 times per second. Repeated three times.

Response Characteristics of the System with Storage

In order to fill (or partially fill) the gaps between successive scans in the complete narrow-band speech transmission system, the output of the audio filter is fed into the dielectric storage system. The storage system has four playback heads, and the signal is repeated three times. The spectrogram of Fig. 19(c) shows the output when the input to the system is random noise, and the scanning voltage is sawtooth. The three repetitions of the received signal are well illustrated in the figure.

VI. Evaluation of the Narrow-Band Speech Transmission System

Intelligibility tests are generally employed in rating speech transmission systems. The intelligibility tests employed in the present study make use of the phonetically balanced and randomized Harvard PB lists (50 words each).[24] The intelligibility score for any given

[24] J. P. Egan, "Articulation testing methods," *The Laryngoscope*, vol. 58, pp. 955–991; September, 1948.

TABLE I

The Per Cent of PB Words Correct with a Frequency Range of 200 to 7000 CPS, a 500-CPS Scanning Filter, and a Sawtooth Scanning Voltage

Scanning Rate (Per Second)	Time Delay Between Adjacent Channels in Milliseconds	Intelligibility Score (Per Cent Correct)			
		D.J.S.	G.E.P.	G.G.F.	P.R.
15	7.5	6	8	6	4
		16	24	6	8
30	7.5	5	9	2	5
		14	8	8	6
50	5	10	2	0	6
		6	6	10	4

TABLE II

The Per Cent of PB Words Correct with a Frequency Range of 200 to 7000 CPS, a 500-CPS Scanning Filter, and a Sinusoidal Scanning Voltage

Scanning Rate (Per Second)	Time Delay Between Adjacent Channels in Milliseconds	Intelligibility Score (Per Cent Correct)			
		D.J.S.	G.E.P.	G.G.F.	P.R.
$1\frac{7}{8}$	$133\frac{1}{3}$	—	—	—	—
$7\frac{1}{2}$	$66\frac{2}{3}$	4	2	—	—
10	$7\frac{1}{2}$	22	4	2	12
15	$7\frac{1}{2}$	32	18	12	18
		40	26	20	14
30	5	6	4	2	4
		10	8	8	6
50	$2\frac{1}{2}$	6	2	4	0
		16	10	2	6

TABLE III

The Per Cent of PB Words Correct with a Frequency Range of 200 to 7000 CPS, an Effective Scanning Filter Width of 1000 CPS, and a Sinusoidal Scanning Voltage

Scanning Rate (Per Second)	Time Delay Between Adjacent Channels in Milliseconds	Intelligibility Score (Per Cent Correct)			
		D.J.S.	G.E.P.	G.G.F.	P.R.
$3\frac{3}{4}$	$66\frac{2}{3}$	4	4	0	6
15	$33\frac{1}{3}$	44	20	16	32
20	15	57	32	28	20
30	15	82	87	63	73
		82	76	58	60
60	15	46	46	46	44
		66	66	48	58
100	$2\frac{1}{2}$	44	38	36	44
		44	48	40	30
200	$1\frac{1}{4}$	4	4	0	4

condition of the transmission system is the percentage of words received correctly.

Procedure

The Harvard PB lists were recorded by a trained speaker through the narrow-band speech transmission system under various conditions. For each condition the output of the system was recorded on a Console Ampex, Model 300 Magnetic Tape Recorder operating at 15 inches per second. The recorded signals were reproduced to the subjects individually over a pair of earphones. Four subjects trained in phonetics and having the same dialect of general American speech were used in these listening tests. Each subject was allowed to select a gain setting which he believed would afford maximum intelligibility for him. In recording, all gains were kept constant throughout the system; for each subject all tests were reproduced with constant gain.

Results

The results of the intelligibility tests are shown in Tables I–III. In each case the effective speech range scanned is 200–7000 cps. A sawtooth scanning voltage was employed in the experiment recorded in Table I and sinusoidal scanning was employed for the two remaining sets of data. Both the actual and the effective filter widths were 500 cps for Tables I and II, but the effective filter width was 1000 cps for Table III. In order to conduct the tests whose results are shown in Table III, the scanning bias and the scanning range of the system were adjusted so that the time-frequency plane was scanned over a range of only 3500 cps. The lists were next played through the system at one half of the recorded speed. The recorded output of the system was played at twice the recorded speed for the listening tests, and thus the original time and frequency ranges were restored. By this technique, the effective bandwidth of the filter and the effective scanning rate were doubled.

Discussion

The results given in Tables I and II show that the intelligibility for a 500-cps bandwidth scanning filter is not very high. In both cases, the intelligibility score reaches a maximum at the rate of 15 times per second;

the score for a sinusoidal scanning is higher than for sawtooth scanning. Although the four subjects differ in their intelligibility score, the trend of variation for each subject is the same. The poorer intelligibility score in the case of sawtooth scanning is probably due to the abrupt transition of the sawtooth wave which produces transients in the filter; whereas, in the case of sinusoidal scanning the change of slope is gradual and, therefore, the filter transients are less prominent.

In the case of a 1000-cps bandwidth filter, the results of Table III show an intelligibility score ranging from 61 to 82 per cent at a scanning rate of 30 times per second; these data are plotted in Fig. 20. In this case also, the individual subjects differ in their scores, but the trend of variation for each subject is the same.

It is interesting to note that in the case where a 1000-cps bandwidth scans the frequency range of 200 to 7000 cps, about 1/7 of the speech signal is selected at every instant of time, and the remaining 6/7 or nearly 86 per cent is rejected. Under these conditions an average maximum intelligibility of 75 per cent is obtained when the scanning is sinusoidal at a rate of 30 times per second. This represents a channel capacity reduction of seven; *i.e.*, seven different messages, each having an average intelligibility of about 75 per cent on monosyllabic PB word lists, could be transmitted over a channel of 7000-cps bandwidth.

Extensive studies of speech intelligibility have shown that the speech range of 200 to 7000 cps can be subdivided into 20 frequency bands such that each band contributes equally to the intelligibility score.[25] According to these data a band-pass filter of width 500 cps, when located in the optimum frequency region for intelligibility, makes a contribution to speech intelligibility of about 25 per cent; a 1000-cps bandwidth filter makes a contribution to speech intelligibility of about 35 per cent.

When the narrow-band filter of the present study was located in the frequency region of maximum intelligibility, *i.e.*, centered at approximately 2000 cps, the intelligibility scores for both the 500-cps filter and the effective 1000-cps filter were found to be somewhat higher than that predicted. These results are probably due to the relatively slow slopes of the filter and to the fact that the listeners were highly trained.

While other narrow-band systems which have been demonstrated provide equal intelligibility with even greater channel capacity reduction, it should be noted that this is the first systematic experimental study of this general type of speech reduction system. Simple improvements that are immediately obvious to us are triangular scanning with rounded corners and an adequate number of signal repetitions (as 13 instead of 3 for a 200–7000-cps range and a 500-cps filter).

[25] H. Fletcher and R. Galt, "Perception of speech and its relation to telephony," *J. Acoust. Soc. Am.*, vol. 22, pp. 89–151; March, 1950.

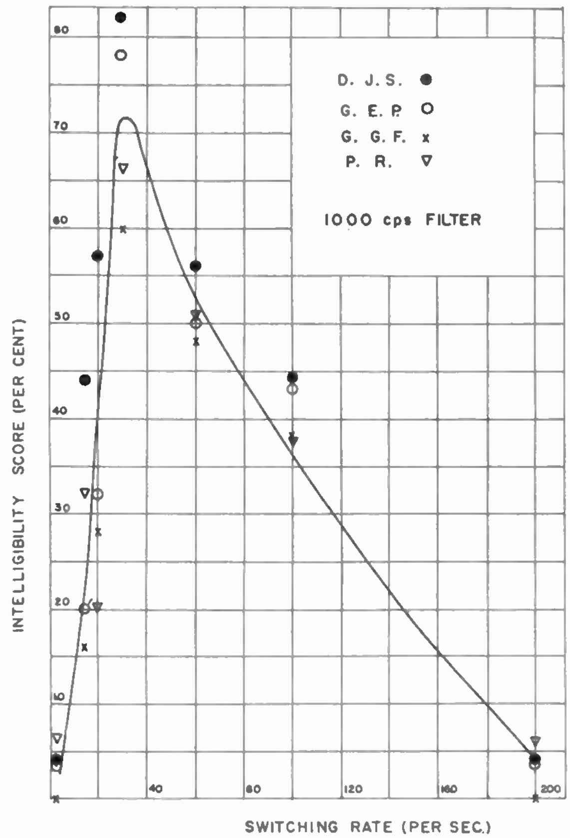

Fig. 20—Intelligibility score vs switching rate (1000-cps filter).

The data presented are for only two different filter widths. The data suggest that the maximum intelligibility achieved is associated with scanning rate in relation to filter transient response. As the filter bandwidth is increased, transient distortions should decrease for a given scanning rate. Bandwidth saving, however, decreases accordingly. If the optimum scanning rates for each of a sequence of filter widths were determined, it should be informative to determine the point of maximum increment in intelligibility as a function of filter bandwidth. One might then consider whether the associated scanning rate is of significance to auditory theory.

VII. Acknowledgment

The authors wish to express their thanks to D. J. Sharf, G. G. Freeman, and P. Rettich for their valuable assistance in evaluating the intelligibility of the narrow-band system. We should also like to express our appreciation to Dr. V. C. Anderson, senior research physicist at the Marine Physical Laboratory of the Scripps Institution of Oceanography, San Diego, Calif. Dr. Anderson advised us in the design of the dielectric recorder and also aided us in achieving proper operation of the instrument. The authors gratefully acknowledge the Horace H. Rackham School of Graduate Studies for a grant from the Michigan Alumni Fund which made it possible to construct the dielectric audio storage system employed in this study.

Circuits for Three-Channel Stereophonic Playback Derived from Two Sound Tracks*

PAUL W. KLIPSCH†

Summary—**Derivation of a third playback channel from two stereo sound tracks may be accomplished by several means. The center channel may be derived by recombination prior to power amplification; acoustically, after amplification, by using two center speakers; and in a variety of phase relationships, including the limiting case of equal signals (monophonic) in which either sum or difference combination may be chosen by polarity selection.**

THE underlying philosophy of a derived third channel from a two-track stereo sound source is that if two microphones are properly placed relative to each other and to a distributed sound source, their combined output would be that of a single microphone midway between them. This single microphone equivalent output may be derived by combination. An actual third microphone output mixed into the two sound tracks is also recoverable by suitable recombination. The peaks and cancellations resulting from the summation of outputs of two microphones seem to be of the same order as those due to standing-wave patterns in an ordinary room.

It will be shown in a later paper[1] that three channels derived from two sound tracks differ from three independent channels in that there is an increase in signal mutuality[2] amounting to a few decibels in some cases. For example, consider a typical source-microphone geometry with three independent channels. The center channel contains only what is heard by the center microphone, but the center microphone includes flanking sounds appropriate to the flanking microphones. In the derived three-channel system, the center channel contains about one decibel more sound derived from the flanks than does the independent three-channel system, due to mixing flanking and center microphones to produce the center channel. Experience indicates that the derived three-channel system can approach the independent three-channel system to a highly satisfactory degree if a considerable number of details are observed. This paper calls attention to some of these details as they relate to recombination on playback.

Randomness between stereophonic signals is shown to apply in some cases, but the center microphone or a sound source equidistant from two microphones is a different case, and may require a more sophisticated recombination circuit for recovery of this signal.

Consider a single simple sound source in space, with two microphones at unequal distances from such source. The phase relation of the output of the two microphones is dependent upon the difference between the distances between the sound source and the two microphones, and if this difference is more than a wavelength or so, the phase[3] relation may be considered random. This would take place, for example, if one microphone were 10 feet from the source, the other 20 feet, and the wavelength 3.75 feet corresponding to a frequency of 300 cycles. Here the phase difference would be 2.66 complete cycles or 960°. The greater the total phase shift, the greater the probability of complete randomness.

This gives rise to the concept that for purposes of analysis, one may consider any given frequency component in a sound source when picked up by two microphones as producing two isoperiodic signals of random phase. To simplify further for analysis, consider the two signals to be of equal amplitude.

It may be stated that the time-average amplitude of the combination of two isoperiodic vibrations of equal amplitude and random phase is $\sqrt{2}$ or about 1.41 times the value of either amplitude.[4,5]

This may be indicated pictorially by a pair of equal amplitudes, A and B, where A is of arbitrary phase (say zero angle for reference) and B is of random phase angle. Fig. 1 shows A and 0°, and B at any of several equally probable angles.

The average value of the sum of A and B is found, for example, by taking a number of equally spaced values of the angle of B (all equally probable because of the random nature of the situation), summing each with A, and dividing by the number of samples.

The mathematics is involved and should be sought in the literature.[4,5]

This sum is found to be representable as a pair of vectors at right angles as in Figs. 2(a) or 2(b).

Following polyphase transformer diagrams, a pair of electrical signals may be considered as derived from

* Manuscript received by the PGA, June 25, 1959; revised manuscript received, October 21, 1959.

† Klipsch and Associates, Hope, Ark.

[1] P. W. Klipsch and R. C. Avedon, "Signal Mutuality and Crosstalk in Two and Three Track Three Channel Stereo Systems," presented at Convention of Audio Engineering Society, New York, N. Y.; October 8, 1959.

[2] Crosstalk is the inadvertent transfer of signal from one channel to another; signal mutuality is the natural consequence of one microphone in a stereo array picking up signals pertinent to other microphones.

[3] For use herein the term "phase" will be used to apply to the angular relation between two directed quantities, where the angle may be any value, and "polarity" applies to the special cases where phase angles are confined to 0° and 180°.

[4] Lord Rayleigh, "On the resultant of a large number of vibrations of the same pitch and of arbitrary phase," *Phil. Mag.*, vol. 10, p. 73; 1880. ("Scientific Papers," vol. 1.) Also, "Theory of Sound," McMillan and Co., Ltd., London, Eng., vol. 1, pp. 35–42, 1877; 2nd ed., 1894.

[5] P. W. Klipsch, "Some aspects of multiple recording in seismic prospecting," *Geophys.*, vol. 1, pp. 365–377; October, 1936.

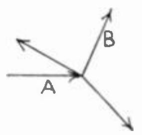

Fig. 1—Addition of two isoperiodic vibrations of random phase. The probable average value is shown to be $\sqrt{2}$ times the value of either vibration, equivalent to the two vibrations taking place at 90° phase shift.

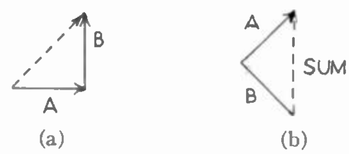

Fig. 2—Two isoperiodic vibrations added at 90°.

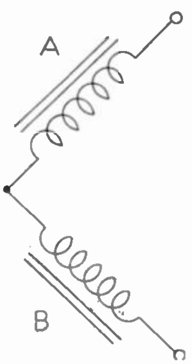

Fig. 3—Two output transformers delivering signals A and B at a 90° phase angle may be illustrated thus.

transformers arranged to show their phase relationship by their geometric arrangement as illustrated in Fig. 3.

In the derivation of three stereophonic sound channels from two sound tracks there are several courses open. Steinberg and Snow[6] merely stated that their center channel was a 6-db mixture of the two sound tracks. The author used a recombination circuit in which the center channel is -3 db with respect to the flanking independent channels as shown in Fig. 4(a) in which the vector values are indicated in Fig. 4(b).

Alternatively,[7] a simpler method using only two amplifiers was shown which was indistinguishable audibly in the tests conducted (which ignored monophonic conditions and conditions approaching monophonic). This is shown in Fig. 5.

In this arrangement, accurate geometric reproduction was obtained.[8] The arrangement has some interesting aspects; one oddity is that for perfectly balanced monophonic tracks, the center channel goes dead. This was found to give more accurate geometry in the case of a two-source sound, like a two-loft pipe organ, than the system of Fig. 4, because the center contributed too little to be heard in the normal stereo listening area. This is as it should be, since there is no sound source in the middle.

When a source of sound is precisely equidistant from two microphones (as in the center of the stage expressed by a line precisely midway between microphones), the

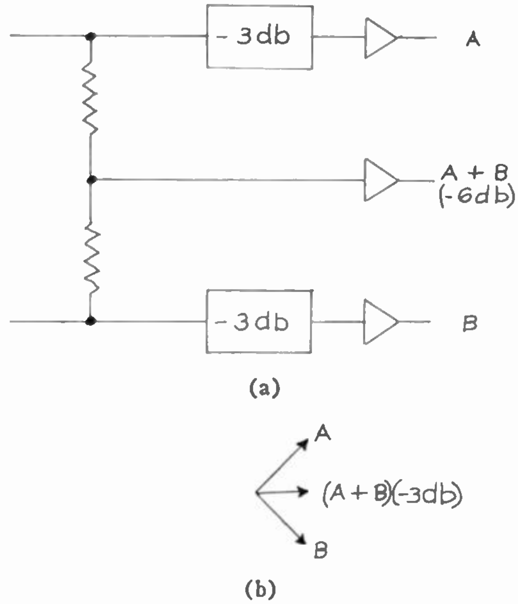

Fig. 4—Derivation of a sum signal, for center channel stereo use, from two sound tracks to be used to feed a third sound channel. The vector relations involved are shown in (b). It should be noted that the recombination circuit drops the center channel 6 db, which is too much, so flanking tracks are dropped 3 db. Experience indicates the desirability of separate volume controls to accommodate for all environments.

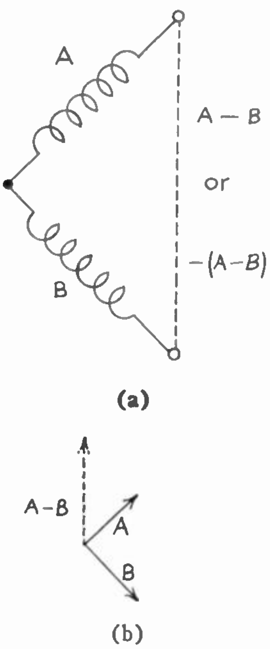

Fig. 5—Derivation of a difference signal from two sound tracks. (1957 Klipsch-Eargle circuit.)

[6] J. C. Steinberg and W. B. Snow, "Physical factors," Symp. on Auditory Perspective, *Trans. AIEE*, vol. 53; 1934. Also, *Elect. Engrg.*, vol. 53, pp. 9–32, 216–219; Janaury, 1934.

[7] P. W. Klipsch, "Stereophonic sound with two tracks, three channels by means of a phantom circuit (2PH3)," *J. Audio Engrg. Soc.*, vol. 6, pp. 118–123; April, 1958.

[8] P. W. Klipsch, "Wide stage stereo," IRE Trans. on Audio, vol. AU-7, pp. 93–96; July-August, 1959.

equal amplitude and phase conditions in recording are such as to cancel on playback; a center microphone introduced equally into the two sound tracks would likewise not be reproduced. And, for monophonic reproduction, balance may be detected by the silencing of the center channel.[9]

There are two approaches to eliminate this null at balance for zero phase-shift signals. One is to make a universal phase shift of, say, 90° to one of the signals; the other is to apply 180° shift or polarity reversal. The 90° method[8] involves a phase-shift network shown in Fig. 6(a) with its vector diagram in Fig. 6(b).

The 180° shift, or polarity reversal, may be accomplished in several different ways. Thompson[7,10] used two center speakers (Fig. 7), feeding the left center from the left track, etc. The summation is acoustical and takes place in front of two identical and properly polarized speakers. For monophonic, this results in dipole effects and an anomalous radiation pattern, but has been demonstrated as highly satisfactory.[11]

A second way involves a transformer (which must be of special design to preserve tonal quality and retain low distortion), shown in Fig. 8.

The circuit of Fig. 8 may be regarded as a variation of that of Fig. 6 in which the frequency at which $\pi/2$ phase shift occurs has been lowered to an infrasonic value; in the ultimate limit the mutual inductance M and capacitance C are increased as much as possible, C being effectively infinite as a short circuit.

A third way, again requiring a high quality additional transformer, is shown in Fig. 9.

In all of the arrangements of Figs. 7–9, balance or focus controls have not been shown, but may be applied.

In the case of the Thompson method (Fig. 7) the two center speakers may be connected to lower voltage taps on the output transformers of the respective amplifiers, or suitably designed asymmetric minimum impedance pads[12] may be inserted.

In the case of the added transformer (Figs. 8 and 9), again the choice of output transformer taps or resistive pad (previous remarks still pertinent) may be employed for a fixed installation and the taps may be preferable; however, when changes must be made, an attenuator proves most convenient and, with proper design, is found audibly indistinguishable from the performance using taps.

The transformer is necessarily a specialized design requiring close coupling between the two windings, perhaps even involving bifilar windings to achieve low leakage inductance, and requiring enough iron and copper for maximum primary inductance. Such a transformer is to be offered as a component.[13]

The addition of a transformer in any of the circuits of Figs. 6, 8, or 9 involves additional exciting current, and thus demands superlative quality. If one wants to avoid an additional transformer, the ideal way to achieve polarity reversal would be within the existing transformers of the power amplifiers.

This may be accomplished by grounding the center tap or 4-ohm tap instead of the normally grounded end of the secondary winding. This leaves available the 4, 8, and 16-ohm taps for the flanking speakers, and leaves available voltages corresponding to plus and minus values obtainable at the 0 and 16-ohm taps.

This requires relatively simple alteration of the amplifiers, but owing to the necessity of changing the feedback loops also, the alteration should be undertaken only by the factory or with factory cooperation and advice, or by someone qualified to readjust the feedback loops for proper gain, minimum distortion, and freedom from instability.

Some amplifiers are already wired this way,[14] with the center tap grounded, so the particular type can be used right off the shelf. Fig. 10 shows the output circuitry, but not the internal feedback loops which would differ for various amplifiers. To carry this to a (perhaps) ultimate status, one might build an amplifier with 32, 8, 0 ohms output, or 8, 0, 8 ohms with grounded center tap to accommodate a 16-ohm speaker in 2PH3 output configuration with minimum impedance mismatch.[15]

Still another way to retain polarity in a 2PH3 system with two amplifiers is to reverse polarity before amplification, as shown in Fig. 11.

This arrangement is particularly simple if one uses a preamplifier which has a "phase" switch. But if one foregoes the complexity of such a preamplifier by feeding directly from a stereo tape machine to the power amplifiers, the polarity reversal requires a transformer or an active network.

There are other methods of achieving three-channel stereo derived from two tracks. One method uses one large full-range speaker in the center with a pair of bass-restricted outrigger speakers of smaller size and limited cost. The circuits of Figs. 6 and 8–10 may be employed, with the addition of suitably chosen series capacitors feeding the small flanking speakers to limit the bass fed

[9] This difference-cancellation phenomenon is useful at times; alignment of tape heads is more precise by null than by maximum indications. It provides an excellent test for balance.

[10] H. Thompson, personal communication.

[11] This experiment and demonstration was conducted at the studios of Klipsch and Associates, Inc., and also at various other locations where valuable audience reactions were derived when it was compared with other derived third-channel systems.

[12] P. W. Klipsch, "The trouble with attenuators," *Audiocraft*, vol. 3, pp. 26–27, 29; November, 1958. This paper criticized particularly the use of pads for attaining balance between woofer and tweeter, for example, and especially constant impedance T pads, or pads exhibiting a high impedance looking back. Applied to the whole speaker system, and using an L pad with minimum impedance looking back, the equivalent generator impedance for a 16-ohm load may have maximum values of the order of 4 ohms, an entirely tolerable value especially for speakers of high inherent damping (horn-type).

[13] Electro-Voice transformer XT-1.

[14] In the McIntosh amplifier, the ground strap on the secondary terminal may be removed and any other secondary tap grounded.

[15] Three 16-ohm speakers would impose approximately 11 ohms load on two amplifiers which may be of 8 ohms nominal impedance.

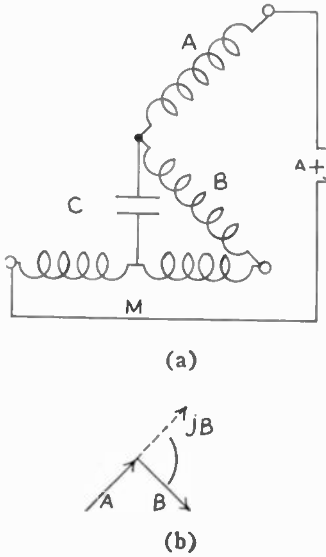

Fig. 6—Derivation of a complex sum signal involving rotating one signal through 90° before addition.

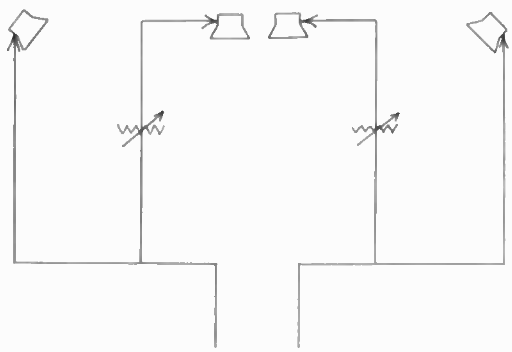

Fig. 7—Derivation of a sum signal for center channel stereo use by acoustic means.

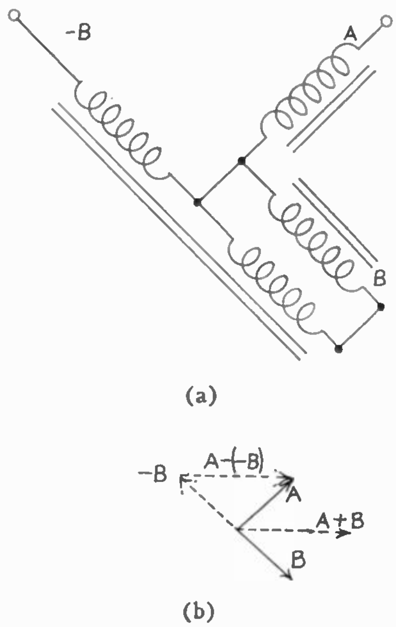

Fig. 8—Extension of system of Fig. 6 to achieve the 90° phase shift (modified Klipsch-Eargle circuit).

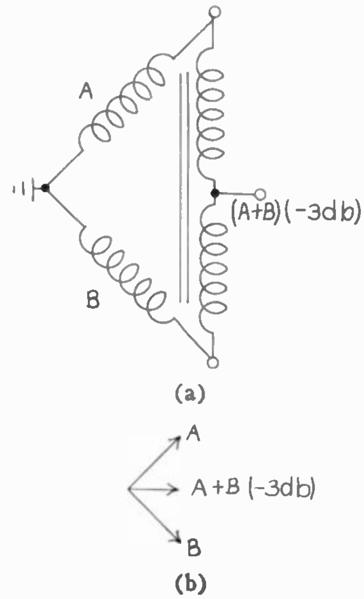

Fig. 9—Another way of deriving a sum signal using an additional transformer.

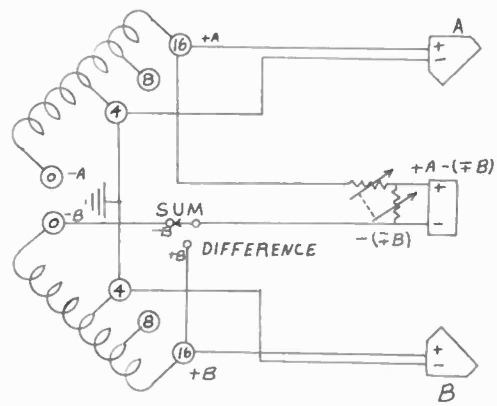

Fig. 10—Simple alteration of power amplifier to derive a sum or difference signal for center channel use. This method obviates the use of an added transformer. Note preservation of correct polarities in all three speakers.

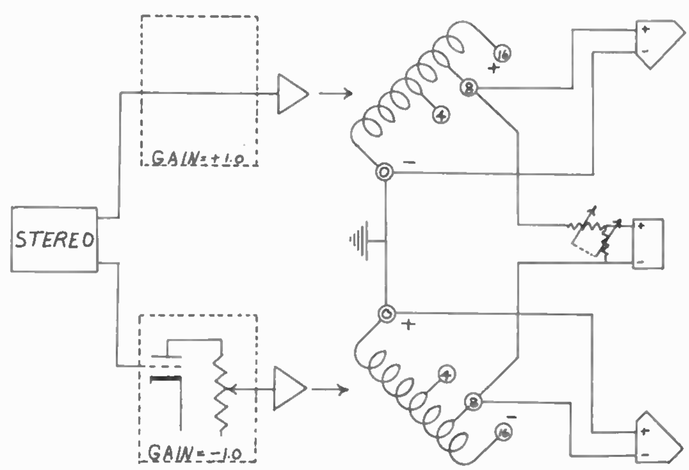

Fig. 11—Polarity reversal in preamplifier for just one track, together with corresponding polarity connections from amplifiers to speakers. This affords additive center channel and consistent flanking polarities. Subtractive polarity switching requires switching in both preamp and output. A "best match" is specifically indicated for the usual-type amplifier to the 10.6-ohm nominal load imposed by three 16-ohm speakers on two amplifiers.

to them to limit distortion. While the sacrifice of non-corner placement of the bass speaker is tonally considerable, stereo geometry is retained, especially if the small flanking speakers are corner placed.

It should be remembered that stereo geometry is concerned with the derivation of the signals or channels and with speaker spacing; tonality depends on speaker quality and placement. Accurate stereo geometry could be attained with three speakers costing $8.00 each; accurate tonality would be attained with speakers costing upwards of 100 times as much. Thus stereo geometry requires a colinear array of three speakers; stereo geometry with tonal fidelity requires that these speakers comprise corner units of superb quality for the flanking elements, and use of a center speaker of substantially equal quality even if the bass is restricted. Experience shows that the less bass restriction in the center speaker, the better. It was once thought that a 300-cycle bass cutoff would suffice, and excellent stereo was obtained with such center speakers. But extension to 100 cycles turned out to be an obvious improvement, and there appears to be a worthwhile further improvement in extending the bass in the center unit to 50 cycles. So far very few people have felt like altering their music room to provide a corner for the center speaker, but with flanking corner speakers a wall-type center speaker offers superb stereo, both in geometry and tonality.

If a center microphone is mixed normally into the two sound tracks, the sum recombination is indicated for its recovery. There are examples of recordings where the center channel was recorded into the two sound tracks as a difference signal, in which case the difference recombination is necessary for its recovery. Therefore, in all the schemes providing for a sum recombination, means should be provided for difference recombination also; that is, a polarity switch should be provided. Such a means is shown in Fig. 10.

The question has been asked: What happens when one microphone signal is some 20 or 30 decibels higher than the others, as for an example, a solo instrument close to it? Experience indicates that the whole situation becomes essentially monophonic to the corresponding channel. As little as 6-db gain increase in one channel can cause the effect of concentrating all attention on this channel. Apparently differences of this magnitude are not practiced in the recordings within the experience of this writer.

At the 1959 High Fidelity Show in New York, the system shown in Fig. 10 was used. Most of the time it could be demonstrated that there was no choice between the sum or difference recombination. One tape, "Porgy and Bess" (a new 4-track tape), required a sum combination to focus the soloist when playing the first half, and a difference combination when turned over to play the second half. Just what peculiar manipulations of the center channel polarities were performed during recording can only be guessed.

How far toward infinity one may multiply the possible forms of two-track three-channel stereo is not foreseeable. It is believed this paper with its references constitutes a fundamentally sound basis from which to project the future.

Characteristics of Degenerative Amplifiers Having a Base-Emitter Shunt Impedance*

WILLIAM D. ROEHR†

Summary—**In amplifiers having emitter degeneration, an impedance is sometimes used between base and emitter. A common case occurs when several emitter followers are used in cascade. The resistors become necessary in order to provide some measure of stability.**

In an audio amplifier of similar design it became necessary to know what effect this shunt resistor would have upon the input impedance of the stage. As a result, the analysis given in this paper was performed. Experimental measurements which support the resulting equation are given.

This analysis and these measurements led to the discovery of a circuit which would exhibit a high ac input impedance, yet the resistors in the dc base circuit could be kept low to provide good stability. Although it can be shown that circuit power gain is the same regardless of whether a resistor is used in series with the emitter or the base to obtain a high input impedance, considerations of dc stability and distortion demand a more thorough investigation of this problem.

Design principles are outlined for this high input impedance stage and an example is worked out in detail. Supporting measurements are given.

* Manuscript received by the PGA, July 6, 1959. This paper appeared in *Electronic Equipment Eng.*, vol. 7, pp. 51–54; July, 1959.
† Motorola, Inc., Phoenix, Ariz.

In amplifiers which are driven from an emitter-follower, it is generally necessary, for stability reasons,[1] to connect a resistor from base to emitter of the driven transistor. Such a configuration is shown in Fig. 1. The input impedance of Q_1 can be calculated from well-known equations providing the emitter load impedance is known. This load is, of course, the input impedance of Q_2, but no formulas have been published which account for the presence of R_B. It is the purpose of this paper to develop a useful equation for the input impedance of Q_2 and to show substantiating measurements.

An important by-product of this development is the discovery of a method of designing an amplifier whose input impedance greatly exceeds the dc resistance in the base circuit. A high input impedance is taken for granted when tubes are used, but is difficult to achieve with germanium transistors, because of the high I_{CO} encountered. Stability considerations generally limit the base bias resistors to a few thousand ohms. A circuit having an input impedance of 100 k ohms is easily designed, using the technique to be described. A high impedance input circuit is very desirable in power amplifiers if sources such as crystal pickups, radio tuners, or tube preamplifiers are to be accommodated without adverse loading effects. Other applications, such as ac electronic voltmeters, could readily use such a circuit.

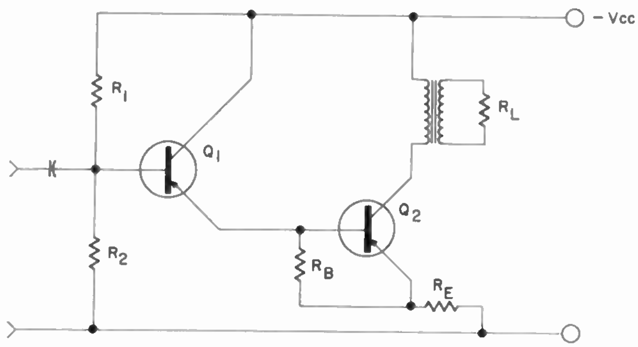

Fig. 1—Direct coupled circuit.

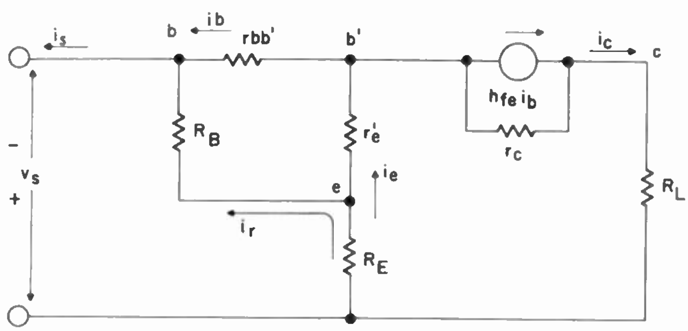

Fig. 2—AC equivalent circuit.

Analysis

The analysis can be easily carried out as follows. Fig. 2 shows the equivalent ac circuit of Q_2. For purposes of this discussion, the effect of r_c will be neglected and R_L will be assumed zero. Cases where this is not true will be discussed later.

The loop equations are

$$v_s = R_E i_e + i_r(R_E + R_B) \tag{1}$$

$$v_{be} = i_e r_e' + i_b r_{bb}' = i_r R_B. \tag{2}$$

Solving (2) for i_r, find

$$i_r = \frac{i_e r_e' + i_b r_{bb}'}{R_B}.$$

Noting that $i_e = (h_{fe}+1)i_b$, write

$$i_r = \frac{i_b}{R_B}\left[(h_{fe}+1)r_e' + r_{bb}'\right].$$

Now the term in brackets is the input impedance of a transistor in the common-emitter configuration with the collector circuit short-circuited. Thus, written simply,

$$i_r = \frac{i_b h_{ie}}{R_B}. \tag{3}$$

This fact may well have been deduced by simple reasoning. Now examine (1). Writing $(h_{fe}+1)i_b$ for i_e and substituting (3) into (1), it becomes

$$v_s = R_E(h_{fe}+1)i_b + i_b\frac{h_{ie}}{R_B}(R_E + R_B). \tag{4}$$

By definition, $R_{in} = v_s/i_s$. Also, $i_s = i_r + i_b$. Using these relations and (3) and (4),

$$R_{in} = \frac{v_s}{i_s} = \frac{v_s}{i_r + i_b} = \frac{R_E(h_{fe}+1)i_b + i_b\dfrac{h_{ie}}{R_B}(R_E+R_B)}{\dfrac{i_b h_{ie}}{R_B} + i_b}.$$

Cancelling the common i_b term plus slight simplification yields the final result:

$$R_{in} = \frac{R_E(h_{fe}+1) + h_{ie}\left(1 + \dfrac{R_E}{R_B}\right)}{1 + \dfrac{h_{ie}}{R_B}}. \tag{5}$$

Inspection reveals that if R_B is open, the expression becomes the familiar one for an amplifier with emitter degeneration. If R_E is zero, the input impedance becomes simply the parallel combination of R_B and h_{ie}. The manner in which the input resistance is reduced by R_B may seem surprising.

[1] W. D. Roehr, "A two watt transistor audio amplifier," IRE Trans. on Audio, vol. AU-7, pp. 125–128; September-October, 1959.

TABLE I
INPUT RESISTANCE* OF CIRCUIT SHOWN IN FIG. 3

Transistor Number	h_{fe}	h_{ie}	$h_{ie}\left(1+\frac{R_E}{R_B}\right)$	$R_E(h_{fe}+1)+h_{ie}\left(1+\frac{R_E}{R_B}\right)$	$\frac{h_{ie}}{R_B}$	Measured R_2	R_1	R_1	Per Cent Error
1	92	2.8	7.8	101	5.00	16.8	16.2	16.2	0.0
2	105	3.1	8.6	115	5.52	17.6	17.0	17.3	−1.7
3	119	3.8	10.6	131	6.78	16.8	16.2	15.9	1.9
4	127	3.7	10.3	138	6.61	18.1	17.4	17.1	1.8
5	94	2.85	7.9	103	5.08	17.0	16.4	16.2	1.2
6	104	3.2	8.9	114	5.72	16.9	16.3	16.2	0.6

* All resistance values are in kilohms.

MEASUREMENTS

To verify (5), the test circuit of Fig. 3 was used. Emitter current was adjusted to 1 ma and oscillator voltage adjusted for 100 mv on V with R_s set to zero. Then R_s, a decade box, was adjusted until V read 50 mv; thus, the circuit input resistance was equal to R_s plus the internal resistance of the generator which measured 500 ohms.

The computed values of R_{in} were found by measuring h_{ie} and h_{fe} under test conditions of $I_E=1$ ma and $V_{CE}=6$ volts. Using (5), R_{in} was computed. This computation yields the value at node 2, which will be called R_2. The circuit R_{in} at node 1, called R_1, as seen by R_s includes the shunting effect of the 470-k ohm bias resistor. The measurements and resulting computations are shown in Table I. Percentage error was taken as $[R_1(\text{calc})-R_1(\text{meas})/R_1(\text{meas})]\times 100$. These results are well within experimental error; indeed, the equipment used does not justify such low error. Type 2N654 and 2N655 transistors were used for these measurements.

THE GENERAL CASE

The effect of r_c may be accounted for in two ways. If the circuit under consideration were an emitter follower, then r_c would be essentially shunted across the input. The result of a load in the collector circuit is to lower the transistor input impedance from h_{ie}, the value obtained with output shorted, to one determined by R_L and r_c. In any event, the transistor will exhibit some input impedance which can be called Z_{be}. An examination of (1) and (3) shows that they are perfectly general; therefore, Z_{be} may be substituted for h_{ie}. There is also no reason why impedance cannot be substituted for resistance in general. Then

$$Z_{in}=\frac{Z_E(h_{fe}+1)+Z_{be}\left(1+\frac{Z_E}{Z_B}\right)}{1+\frac{Z_{be}}{Z_B}} \qquad (6)$$

where all quantities may be a function of frequency.

THE RESISTANCE MULTIPLYING CIRCUIT

A look at the circuit of Fig. 3 and the data in Table I reveals a somewhat startling result. If the transistor were removed from its socket, the input impedance of the circuit would be 1600 ohms, but with the circuit operating, the impedance increases to approximately 16 k ohms. This effect suggests that a circuit might be designed which would exhibit high input impedance, but the dc resistance in the base circuit could be kept low to provide good stability. Such a circuit is shown in Fig. 4. It will be noted that, for ac R_E is the parallel combination of R_1, R_2, and R_5, and R_B is the parallel combination of R_3 and R_4. For dc, R_E is just R_1, while R_B is the parallel combination of R_2+R_3 and R_4+R_5.

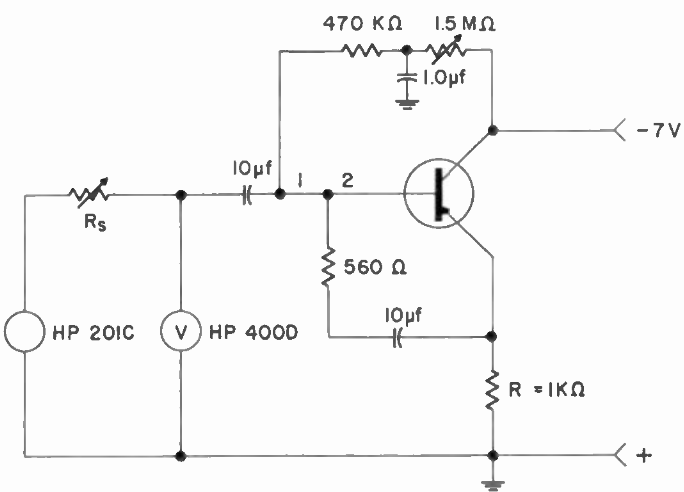

Fig. 3—Input impedance test circuit.

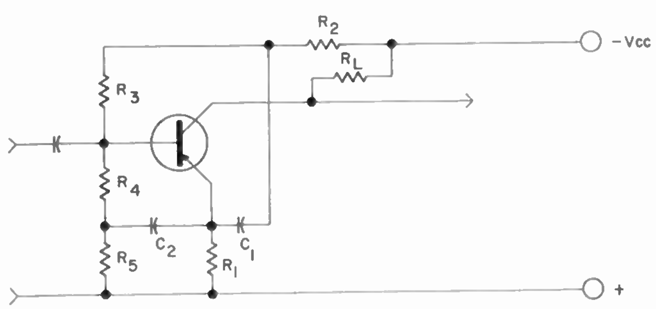

Fig. 4—Resistance multiplying circuit.

Before considering a design method for this circuit, an expression for power gain should be derived which might aid in the choice of component values. The basic equation for power gain (see Fig. 2) is

$$G_E = \frac{i_c^2 R_L}{i_s^2 R_{\text{in}}}. \tag{7}$$

Substituting the previously derived expressions for i_s and R_{in},

$$G_E = \frac{i_c^2 R_L}{\left(\frac{i_b h_{ie}}{R_B} + i_b\right)^2 \left[\frac{R_E(h_{fe}+1) + h_{ie}\left(1 + \frac{R_E}{R_B}\right)}{1 + \frac{h_{ie}}{R_B}}\right]}.$$

Cancelling common terms and replacing i_c/i_b by h_{fe} (this assumes $R_L \ll r_c$),

$$G_E = \frac{h_{fe}^2 R_L}{\left[R_E(h_{fe}+1) + h_{ie}\left(1 + \frac{R_E}{R_B}\right)\right]\left[1 + \frac{h_{ie}}{R_B}\right]}. \tag{8}$$

In cases where high input resistance is desired, the dominant term in the denominator will be $R_E(h_{fe}+1)$. The addition of R_B will then have little effect upon power gain, especially when $R_B > R_E$ and $R_B > h_{ie}$, conditions which should be strived for in practice.

Design Principles

A few principles to consider when designing the resistance multiplying stage will next be outlined. The first problem is to gain an idea of the required values for R_E and R_B to permit the design value of input impedance to be obtained. The available dc supply, required voltage output, and dc stability will fix the maximum values for R_1 and the dc base resistance. The output should preferably be taken from R_L, so that changes in the succeeding stage will not cause large input impedance fluctuations. Once the dc conditions are set, the problem is to determine the position of the taps for the capacitors on the base bias resistors. Eq. (5) may be simplified for this purpose. Assuming the major part of the input impedance is due to the $R_E h_{fe}$ term, write (5) as

$$R_{\text{in}} = \frac{h_{fe} R_E}{1 + \frac{h_{ie}}{R_B}}. \tag{9}$$

Establish a ratio (K) between R_B and h_{ie}. Then, solving (9) for R_E and using K for h_{ie}/R_B,

$$R_E = \frac{R_{\text{in}}(1 + K)}{h_{fe}}. \tag{10}$$

Now as K is made smaller, obviously R_E can be smaller for a given input impedance. As (8) shows, this will result in more power gain. However, a lower limit is placed upon K, dictated by the stability of the circuit and I_{CO} of the chosen transistor. It is necessary to start somewhere; probably a good compromise value for K is 0.5. Then R_3 and R_4 are fixed; from (10) R_E may be computed and, hence, R_1, R_2, and R_5 estimated.

These principles may best be clarified by an example. The design requirements are outlined below, together with the transistor data and available supply voltage.

Amplifier Characteristics:

Output voltage—1 v peak across 2.2 K ohm load
Input impedance—100 k ohms or greater
DC stability—100 μa I_{CO} should cause no performance degradation.

Fixed Conditions:

Supply voltage—24 volts
Transistor Specs at $V_C = 6$ V and $I_E = 1$ ma: $h_{fe} = 100$, $h_{ie} = 3000$ ohms.

Choose $K = 0.5$, *i.e.*, $R_B = 6$ k ohms; from (10) $R_E = 100(1.5)/100 = 1.5$ k ohms. As a trial, make $R_1 = R_2 = R_5 = 4700$ ohms, then $R_E = 1570$ ohms. A 1.0-ma emitter current would easily allow 1 volt output at low distortion across a 2.2-K load. Then $V_E \approx V_B = 4.7$ volts. By inspection,

$$V_B = \frac{V_{cc}(R_4 + R_5)}{R_2 + R_3 + R_4 + R_5},$$

ignoring I_B, which is normally small. If R_4 is chosen as 10 k (so that the parallel combination of R_3 and R_4 will be 6 k or more), from Shea,[2] the current stability factor,

$$S_I \lessapprox \frac{R_4 + R_5}{R_1} = \frac{10 + 4.7}{4.7} = 3.1,$$

which should be satisfactory. Solving for R_3,

$$R_3 = \frac{V_{cc}(R_4 + R_5)}{V_B} - (R_2 + R_4 + R_5)$$

$$R_3 = \frac{24(14.7)}{4.7} - (4.7 + 14.7) = 56 \text{ k}.$$

Since maximum S_I is 3.1, emitter current will increase to 1.31 ma when maximum I_{CO} is flowing. Minimum V_{CE} should be about 1 volt dc plus the required peak output voltage swing—in this case, 2 volts. Emitter voltage is $(4.7)(1.31) = 6.3$ v. Therefore, the drop across R_L should not exceed $24 - (6.3 + 2) = 15.7$ v. $I_C \approx I_E + I_{CO} = 1.31 + 0.10 = 1.41$ ma, then maximum $R_L = 15.7/1.41 = 11.3$

[2] R. F. Shea, *et al.*, "Transistor Circuit Engineering," John Wiley and Sons, Inc., New York, N. Y., ch. 3; 1957.

k ohms. 10 k is a convenient value and allows some safety factor. The circuit is now designed as shown below.

Using exact expressions and values as shown in Fig. 5,

$$R_L(\text{ac}) = \frac{R_L R_L'}{R_L + R_L'} = \frac{(2.2)(10)}{2.2 + 10} = 1.8\text{ k}$$

$$R_E = \frac{1}{1/R_1 + 1/R_2 + 1/R_5} = \frac{4.7}{3} = 1570\text{ ohms}$$

$$R_B = \frac{R_3 R_4}{R_3 + R_4} = \frac{(56)(10)}{56 + 10} = 8.45\text{ k}$$

$$S_I \approx \frac{\dfrac{(R_2 + R_3)(R_4 + R_5)}{R_2 + R_3 + R_4 + R_5}}{R_1} = \frac{\dfrac{(4.7 + 56)(10 + 4.7)}{4.7 + 56 + 10 + 4.7}}{4.7}$$

$$= \frac{11.8}{4.7} = 2.5.$$

Substituting into (5) and (8) yields the results shown in Table II.

Calculation of the required values for C_1 and C_2 is a difficult operation and will not be shown here. A general rule can be stated, however. Their reactance should be small compared to R_2 and R_5 at all frequencies of interest. 15-μf units were used in the example. Data obtained are shown in Table II. All the original specifications were met or exceeded. Using a 120-k ohm source impedance, frequency response was down 1 db at 25 cps and 23 kc. Bandwidth becomes greater as the source impedance is reduced.

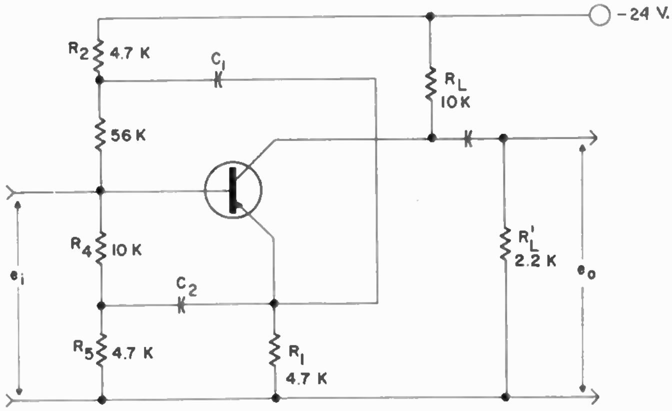

Fig. 5—Completed circuit.

TABLE II

Characteristic	$Z_{in}(K\Omega)$	G_E	S_I	Distortion At 1V Peak Output At 1000 cps	
				No I_{CO}	120 μa Simulated I_{CO}
Measured	118	18 db	2.8	0.4 per cent	0.3 per cent
Calculated	120	19 db	2.5	—	—

Conclusions

Design equations have been developed showing input impedance and power gain for a base input amplifier having a shunt resistor from base to emitter. Experimental measurements support these equations, which will be found useful when designing emitter follower circuits or other circuits having a high input impedance.

Correspondence

Compatible Stereo Sound*

With reference to Lamberty's contribution[1] it will be of interest to note that British Patent No. 612,163 of application date 1946 by W. H. Livy of E.M.I. Studios, London, discloses this method.

E. Fowler
Manager, Recording Studios
E.M.I. Records, Ltd.
London N.W. 8, Eng.

* Received by the PGA, September 21, 1959.

[1] B. J. Lamberty, "A compatible method of recording and reproduction of stereo sound," IRE Trans. on Audio, vol. AU-6, p. 89; July-August, 1958.

Contributors

Paul W. Klipsch (A'34–M'44–SM'45), for a photo and biography please see page 40 of the March–April, 1959, issue of these TRANSACTIONS.

Creon D. O'Neal (A'52–M'55) was born on April 9, 1908, in Cumberland, Md. He received the A.B. degree from Fairmont State College, Fairmont, W. Va., in 1931.

C. D. O'NEAL

From 1931 to 1938, he taught physical and mathematical sciences, and was in private business from 1938 to 1943. In 1943, he joined the Philco Corporation, Philadelphia, Pa., where he served with the research, engineering and government divisions, respectively. He is presently employed as the director of electronic research at Fidelitone, Inc., Chicago, Ill., which he joined in 1958.

He has worked with condensers, quartz crystals, radar scanners, phonograph pickups, record changers, electrostatic speakers, audio amplifiers, acoustic torpedoes, and the Sidewinder Missile.

Mr. O'Neal was chairman of the Philadelphia Chapter of the PGA in 1957, and has served on various standardization committees for both civilian and military engineering societies.

Gordon E. Peterson was born in Danville, Ill., on October 14, 1913. He attended DePauw University, Greencastle, Ind., as a Rector Scholar, and received the A.B. degree in mathematics in June, 1935. His graduate study was in phonetics and electrical engineering at Louisiana State University, Baton Rouge, where he received the M.A. degree in 1937 and the Ph.D. degree in 1939.

G. E. PETERSON

From 1939 to 1943 he was director of the Speech and Hearing Clinics at Ball State Teachers College, Muncie, Ind. Thereafter he spent a year teaching phonetics at Northwestern University, Evanston, Ill. From 1944 to 1946 he was a research fellow in the Psycho-Acoustic Laboratory at Harvard University, Cambridge, Mass., where he conducted research on communication systems.

Since 1946 Dr. Peterson's major efforts have been devoted to research on speech analysis and speech synthesis, with particular reference to automatic speech recognition. He was a member of the technical staff of the Bell Telephone Laboratories until 1953, when he joined the faculty of the University of Michigan. At the present time he is professor of speech and electrical engineering, Director of the Speech Research Laboratory, and Chairman of the interdisciplinary degree program in Communication Sciences at the University of Michigan.

Dr. Peterson is a Fellow of the Acoustical Society of America and the American Speech and Hearing Association, and a member of Phi Kappa Phi and Sigma Xi.

William D. Roehr (S'57–M'58), for a photo and biography please see page 135 of the September–October, 1959 issue of these TRANSACTIONS.

D. L. Subrahmanyam was born in Bangalore, India on July 17, 1921. In 1944 he received the B.S. degree with honors in physics, and in 1945 the M.S. degree in physics from the University of Mysore, Bangalore, India.

D. L. SUBRAHMANYAM

From 1945 to 1948 he was a research assistant in the Electrical Communication Engineering Department of the Indian Institute of Science, Bangalore, India. During this period he worked on the instrumentation and measurement techniques for studying acoustics of rooms. He was awarded the associateship of the Indian Institute of Science for this work. From 1948 to 1951 he worked in the Research Department of All India Radio, Delhi, India.

In 1953, Dr. Subrahmanyam obtained the S.M. degree in electrical engineering from Massachusetts Institute of Technology, Cambridge, Mass. and in 1954 the Ph.D. degree from the University of Michigan. During this period, he investigated various aspects of acoustical properties of sound absorbing materials, absorption of sound over audience, and narrow-band speech transmission systems.

Since 1957, he has served as a senior scientific officer in the Central Electronics Engineering Research Institute, Pilani, Rajasthan, India, and is the head of the Audio and Acoustics Division.

Dr. Subrahmanyam is a member of Sigma Xi.

Index

to

IRE Transactions

on

Audio

Volume Au-7, 1959

IRE Professional Group on Audio
Combined Index for 1959

Compiled by D. W. Martin

Classification of Subjects

TRANSACTIONS OF THE IRE
CONTENTS

January–February, 1959

AU-7-1

March–April, 1959

AU-7-2

AUTHOR INDEX FOR 1959 IRE TRANSACTIONS ON AUDIO

ANALYTIC SUBJECT INDEX FOR 1959 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 Technical Secretary, Institute of Radio Engineers, Inc., 1 East 79th Street, New York 21, N.Y.