AUDIO DESIGN HANDBOOK

H. A. Hartley

H. A. HARTLEY'S

Audio Design HANDBOOK

GERNSBACK LIBRARY, INC.
NEW YORK 11, N. Y.

Library of Congress Catalog Card No. 57—9007

contents

introduction

THE aim of this book is to bring together in compact form all the material necessary for designing an audio installation of the type known as "high fidelity." Within the limitations of one volume it is impossible for it to be complete; nor is it necessary that it should be so.

There is a vast amount of matter available for reference and study. Apart from the numerous books treating high fidelity as a separate branch of audio engineering, there is the whole range of standard textbooks and innumerable articles in the technical magazines. Some of these are of very great value and the student, engineer and purely amateur enthusiast should not fail to refer to the various indexes published annually listing what has been written. Yet it has to be said that there is an even vaster accumulation of writing on the subject that, while having technical integrity, does not seem to have much practical value.

What value has this book, then, when so much has already been written? The answer to that stems from a conversation I had with the publisher some three years ago. We were discussing various books and scientific papers and he said that he found it very difficult to steer a steady course through so much material—and so many shoals. I said that I, too, had sometimes had to hunt through masses of stuff to get hold of some fact I wanted and very often didn't find it. There and then we resolved to compile a book which would be a handy data book for those interested in audio.

The work was started but before very long we came, independently, to the conclusion that what was wanted was not so much a data book as a guide book, something that would describe how the job was done, with emphasis on the need for the designer to have a clear idea in his mind what he wanted to do before he started in to do it.

This meant that a good deal of the material already got together had to be scrapped and other material fashioned into the new approach. This accounts for the long delay between the original announcement of the book and its actual appearance. While apologizing to those who have been so patient, I think I can fairly say that the book will be more useful to them in its present form.

I describe how my more than 30 years' experience in audio engineering suggests a problem in design should be tackled. I also add from time to time some comments on blind alleys I have explored, and I conclude with a wholly personal chapter on where I think we are all going and why we shouldn't go there! These past 8 years have been exciting years for all audio workers, but because, in the continual search for novelty, certain ideas are being exploited which were tried and rejected thirty years ago (but which are unknown to younger readers), a good-natured word of warning here and there seems desirable.

I have not attempted a thorough treatise on speakers and enclosures. The subject is far too vast even for a whole volume. Meanwhile I offer this book as, I hope, a dependable guide through the intricacies of audio, and where it fails I shall be grateful for comment and criticism and suggestions for its improvement.

H. A. HARTLEY

chapter 1

perception of sound

THE audio engineer is concerned with designing equipment to produce imitations of speech and music. If the imitation bears a close resemblance to the original sound it is said to be "high fidelity." This is a purely subjective definition, for what is high fidelity to one person is an irritating noise to another. In any event, the medium of perception is the human ear, and ears, in sensitivity and frequency response, vary within very wide limits. The "average ear" is as much a myth as an "average person." No person is an average person—he is himself and he has his own particular pair of ears. So curves which give a smooth mean result based on tests of typical cross-sections of listeners represent the *average* of a number of ears, but any one pair of ears may give results considerably different from these curves.

Characteristics of musical instruments

Before the designer can consider ways of dealing with musical sounds he must understand the nature of those sounds. The numerous books on sound reproduction almost invariably give a diagram showing the frequency range of musical instruments; most give the range of fundamental frequencies emitted by various instruments; some include what is generally described as the harmonic range. Most of these are only partially correct and virtually useless.

It is well known that the differing *timbre*[1] of musical sounds and speech is due to the presence of harmonics or *partials*[2] as well as a fundamental frequency. The harmonics have frequencies which are integer multiples of the fundamental frequency and they can vary in number and amplitude. Basic research in this subject was carried out by Lord Rayleigh *(Theory of Sound)*[3] and H. von Helmholtz *(Sensations of Tone)*[4].Reference to these works will give the mathematical and scientific analysis of musical sounds.

Harmonics

A rough guide to the number and amplitude of harmonics in some sound sources is given in Table 1. The relative intensities of harmonics are expressed in figures, but in many cases the figures are guesswork, based on a method of simulating the tone by using a series of tuning forks. With strings, reasonably accurate measurements can be made but it is almost impossible to measure the relative intensities of the harmonics in wind instruments and voices. The method is to set up a series of tuning forks until, when simultaneously activated, they produce a sound which seems to match the original. Admittedly the modern electronic organ had not been dreamed of in Rayleigh's and Helmholtz' day, but their data seem fairly sure.

Modern techniques could produce complete records. Instruments could have their performance recorded in photographed oscillograms and the oscillograms analyzed by harmonic analyzers. However, the writer is not aware of any comprehensive publication of work done on these lines. He has been at some pains to compile all the available information on the nature of the various sources of musical sounds and this data is presented in Fig. 101, which gives a more complete picture of the frequency range of voices and instruments than has hitherto appeared.

For guidance in designing equipment, a dashed line has been added to the illustration(Fig.101)to show what part of the compass of the instruments is not reproduced correctly by apparatus having an upper cutoff at 12,000 cycles. This line could not be a vertical one drawn through the 12,000 ordinate on the frequency

---

[1] i.e. musical quality.

[2] Sometimes called "overtones."

[3] Lord Rayleigh (John William Strutt, 1842-1919). *Theory of Sound,* 1st edition, 1877. 1st American edition, 1945, by special arrangement with the Macmillan Co.

[4] Hermann L. F. von Helmholtz (1821-1894), *On the Sensations of Tone as a Physiological Basis for the Theory of Music,* 1st German edition, 1862. American edition, 1954, Dover Publications.

Table 1—Harmonic Range of Various Sound Sources

Instrument		Fundamental	Harmonics							
			2nd	3rd	4th	5th	6th	7th	8th	Others
	Harp	100	80	56	31	13	3			see note below
	Piano	100	100	9	2	1	.01			" " "
	Violin	100	25	11	6	4	3	2	1.5	
	Flute	100	10							
Flute Pipes of Organ	Stopped: wide	100		25						
	" narrow	100		50		10				
	Open: wide	100	50	10						
	" narrow	100	75	50	40	30	10			
	Oboe	100	80	70	60	50	25	10	5	and others
	Clarinet	100		90	25	50	25	50	5	
	French horn	100	80	75	50	30	10			
	Trumpet	100	80	70	60	50	30	20	5	
	Human voice	100	80	60	50	30	20	10	5	and up to 16th

Note: particularly in the case of the piano the harmonic development depends on the fundamental frequency. For example, C an octave below middle C has 15 harmonics; C an octave above middle C has only 6, and "top C" only five. Other instruments show similar variations, and the harmonic range of stringed instruments depends on whether they are bowed, muted or plucked. Wind instruments have a different harmonic development according to whether they are played "straight" or muted.

The numbers given in Table 1 above represent the relative intensities of the harmonics. Thus, the 6th harmonic of the oboe is 25/100 or 25% of the amplitude of the fundamental.

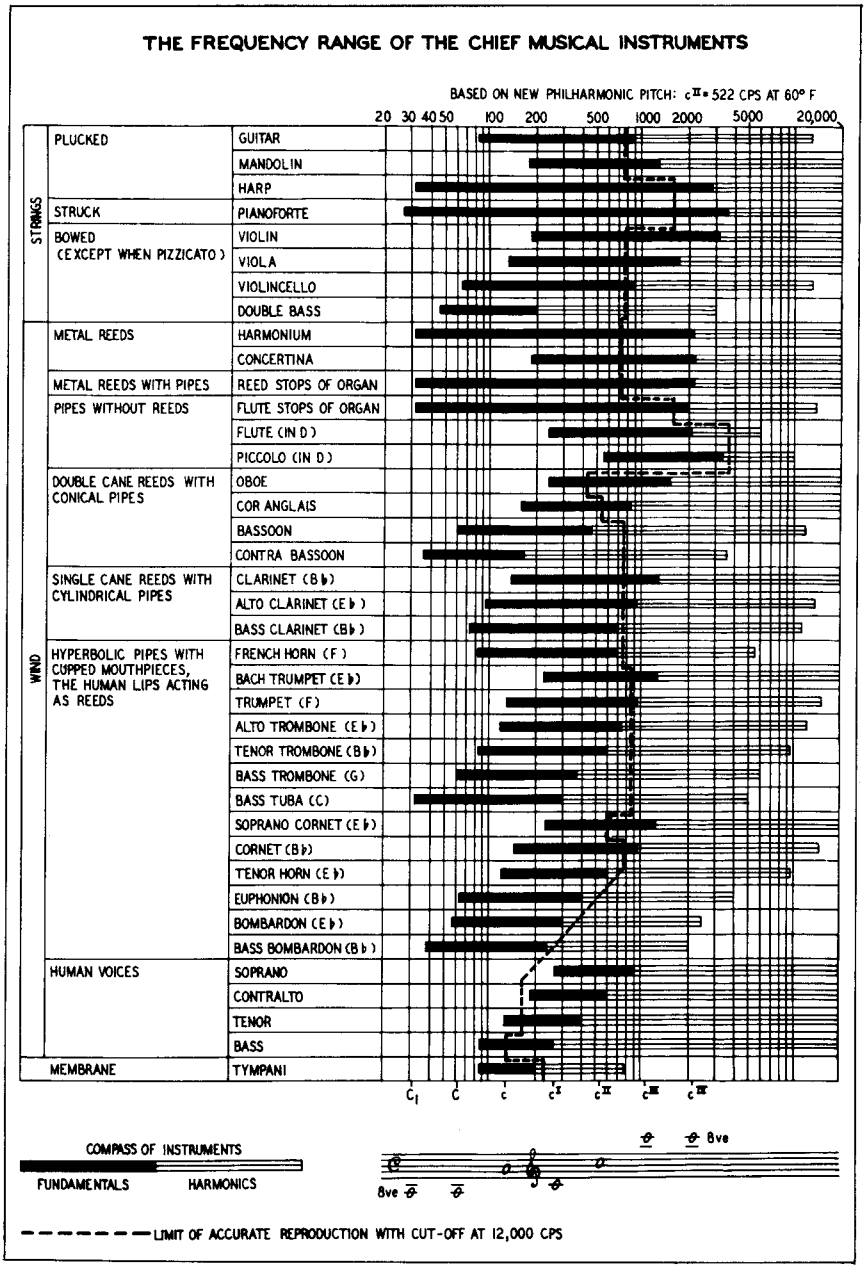


Fig. 101. *In the frequency scale of the above chart the solid line opposite each instrument shows the fundamental range; the harmonic range is depicted by a triple line. See page 8, and the pages following, for the meaning of the dotted line representing a cutoff at 12,000 cps.*

axis since the horizontal line for each instrument represents all its fundamental frequencies and harmonics. Each complex note needs to have all of its harmonics reproduced correctly. The solid parts of the lines represent the fundamental frequencies, but for the lower fundamental frequencies it will be obvious that the "harmonic range" extends downward over the fundamental range shown. The cutoff line (refer to Fig. 101 illustrated on the facing page) must therefore pass through that part of the line which represents the highest frequency at which *all* the harmonics are reproduced.

A little reflection will indicate that not even an overall flat response to 20,000 cps will give near-perfect results in the case of

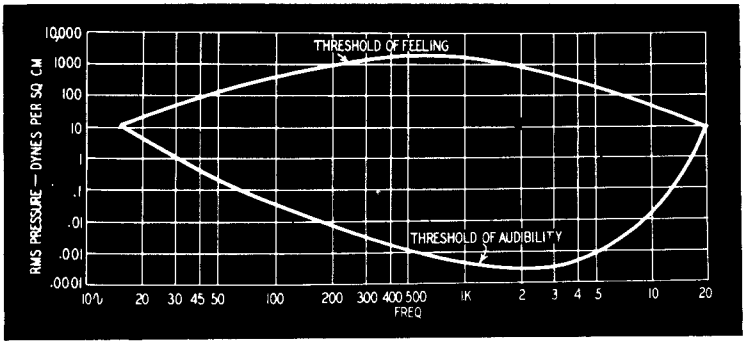


Fig. 102. *Auditory sensation area of the human ear.*

some instruments and all human voices. The situation is further complicated by the fact that a human ear with an upper-limit response of 12,000 cps listening to a live performance in which all the harmonics are present will not hear the same thing as when listening to a reproduced performance having nothing higher than 12,000, due to the fact that the ear is conditioned by what to it are supersonic frequencies. This will be considered further in the next subsection.

Characteristics of the human ear

Sounds too quiet to be heard are neither heard nor felt. As the intensity is increased, the point at which they are just heard is called the *threshold of audibility*. This will depend on the characteristics of any particular ear, but it will also depend on the frequency. The pressure of the sound on the ear (in dynes per square centimeter) can be plotted against the frequency (in cps), and this is shown in Fig. 102. As the pressure on the ear increases

there comes a point when the sensation ceases to be just hearing and becomes both hearing and feeling—in the case of very low frequencies a feeling of throbbing and, when the intensity is high enough, actual pain. This again depends on the ear and the frequency. The point at which feeling rather than hearing becomes dominant is called the *threshold of feeling* and is also plotted in Fig. 102. The area between the two curves is called the *auditory sensation area* of the ear. This is, of course, that mythical "average ear" but all ears follow the same general characteristics, except that individual variation at certain frequencies may be as great as 20 db.

Response of the ear

It will be noticed that the ear is most sensitive to frequencies between 2,000 and 3,000 cps. It so happens that many loudspeakers have a resonance at about 3,000 cps because of defects in construction, the cheapest speakers usually being the worst offenders. While such a combination of characteristics provides an easy way of affording "efficiency" for the nondiscriminating listener, it is precisely the reverse of what is wanted for *fidelity* of performance.

To pure tones the ear has been said to be sensitive to increases of power of from 1 to 3 db. This, however, depends on the frequency of the pure tone. Weber's Law states that in order to produce a perceptible increase in the intensity of a sensation an equal fraction must be added to the previous intensity of the stimulus, whatever its value. The magnitude of the sensation (loudness) produced is proportional to the logarithm of the stimulus (intensity). If I = the original intensity and $\triangle I$ = the just perceptible increase in intensity, then $\triangle I/I$ = the differential sensitivity. The sensation level of a sound is expressed in decibels above the threshold of audibility.

Loudness and intensity

To some readers the above statements may seem obscure, and this may be attributed to the fact that there is a rather widespread confusion as to the exact meanings of loudness and intensity; equally common is the misuse of the term 'decibel'. We frequently read that such and such has a level or magnitude of so many db. This statement is absolutely meaningless, for the unit is one of *comparison* only, and is not absolute. A value expressed in decibels must therefore be given a reference level. The reference level may not actually be specified, but the context must make the implied reference level free from confusion.

The *intensity* of a sound is a real physical quantity that can be defined and measured precisely. It has an objective existence whether anyone is listening to it or not. The *loudness* of the same sound is the effect it has on a listener and is therefore purely subjective, since it depends on the ear and the hearer's motor mechanism. Admittedly the effect of loudnesss, the magnitude of the sensation, is dependent on the primary intensity of the sound, but the loudness itself is not something that can be measured absolutely. It must be referred to some previous level existing in the

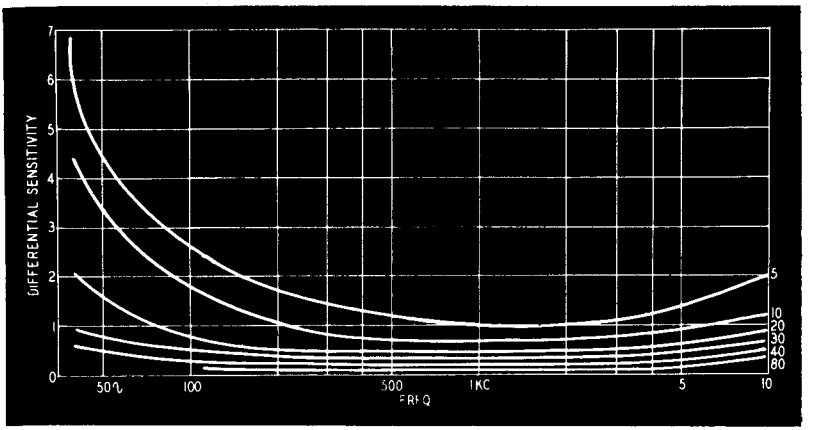


Fig. 103. *Variation of frequency response of the ear at various volume levels. The figures shown at the right-hand ends of the curves are in db.*

hearer's mind. Weber postulates an original intensity; whatever the original intensity may be, an equal proportion of it must be added to produce a perceptible increase. The increase is *perceived* by the hearer as an increase in loudness, and the loudness is proportional to the *logarithm* of the intensity.

Ear response vs. volume level

Let us call the original intensity P_1 (P representing power, for the intensity of a sound is power) and the increased intensity for perceptible increase P_2. Then the loudness is proportional to log $[P_2/P_1]$. The *bel* is the unit of relationship between the two powers and is expressed as the common logarithm of the relationship. The bel is an inconveniently large unit and is therefore usually divided into 10 *decibels*. Hence difference in level in db between two powers P_1 and P_2 equals 10 log $_{10}[P_2/P_1]$. With these definitions in mind a series of curves can be prepared showing the behavior of the ear towards sounds at different volume levels. These are shown in Fig. 103.

These curves show that the frequency response of the ear changes considerably for different volume levels. As could be seen from Fig. 102 the ear is most sensitive to frequencies of 2,000 to 3,000 cps and least sensitive to frequencies at the end of the scale. Fig. 103 shows, however, that this nonlinearity can be overcome by increasing the level of the applied sounds. This argument only applies to pure tones, that is, sounds of sine wave form.

Phonic level

The relation between sensation level and intensity depends on the waveform of the sound. Generally,

Sensation level [in db] $= 10 \log_{10} \times$ intensity [in microwatts per sq. cm.]

If the intensity is such that 1 microwatt moves through 1 square centimeter, which is taken as the comparison intensity, then the sensation level (called the *phonic level)* is zero. This corresponds to the intensity on the ear drum if the pressure of the sound wave is 20 dynes per square centimeter. Reference to Fig. 102 shows that the ear has its widest frequency response at this pressure.

Fig. 103 shows that a phonic level of at least +30 db is necessary for the ear to act as a reasonably linear transducer of sounds. It also shows how much the lower and higher frequencies have to be increased in intensity to give the same loudness sensation at different levels. This is shown for single pure tones. If they are musical sounds (complex waves) the level should be higher. There is also a range of level to consider.

In a concert hall the dynamic range of an orchestra is about 60 db; that of phonograph records is probably about 45 db and of a broadcasting station perhaps 30 db. A level of 30 db would, in a completely quiet listening room, permit adequate reproduction of a broadcast program but phonograph reproduction would require a 45 db level, which at the same time puts the ear into a better condition for high-grade listening. If there is background noise, the level must be set higher. Even so, it will be realized that during the quietest passages the ear itself distorts, but as this also happens in listening to an original performance one need not attach undue importance to this consideration.

Interference

The passage of one sound wave through the air does not affect the passage of other sound waves if the amplitude is small. If a particle of air is acted on by two sound waves, the resultant displacement of the particle is obtained by adding the separate dis-

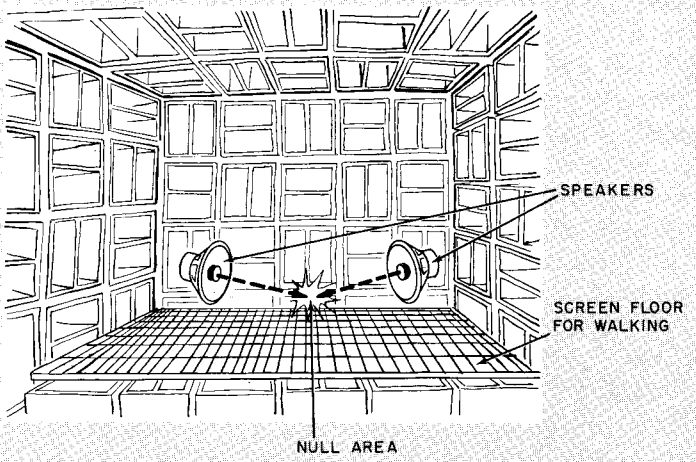
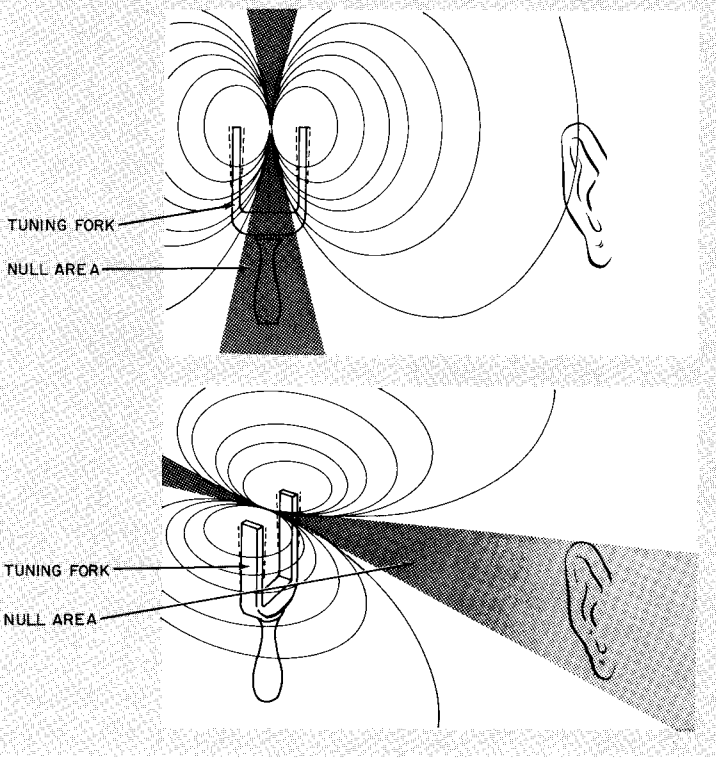


Fig. 104. *Mutual interference of sound waves: a—(above) two speakers working at the same level and frequency in an anechoic chamber; b—(below) null areas produced by the tines of a tuning fork.*



placements vectorially. If the two waves have the same frequency, have amplitudes a_1 and a_2, and are $\phi°$ out of phase, the resultant amplitude $= \sqrt{a_1^2 + a_2^2 + 2a_1a_2 \cos \phi}$. If the two waves are in opposite phase, that is, if a_1 is at greatest positive value when a_2 is at greatest negative value, then $\phi = 180°$, $\cos \phi = -1$, and the above expression reduces to $a_1 - a_2$. If the amplitudes of the two waves are equal, then the resultant amplitude becomes zero. In other words, two sound waves of equal frequency and amplitude but 180° out of phase produce no sound whatever. This phenomenon is called *interference,* and it can be demonstrated very easily with an ordinary tuning fork (Fig. 104-b).

When the fork is struck both tines vibrate at the frequency for which the fork was designed, and while vibrating separate simultaneously and come together simultaneously. Assume the fork to be held vertically with the tines in a north-south axis. When the tines are separating pressure is exerted on the air in northerly and southerly directions creating sound waves travelling outwards; at these instants rarefaction of the air between the tines results in sound waves travelling inwards from east and west. These outgoing and incoming waves extend over spherical zones and on the axes dividing the angles between north-south and east-west they neutralize each other by mutual interference. If you strike the fork and twirl it near one ear you will hear, in each revolution of the fork, four points of maximum sound, four points of no sound, and intermediate waxing and waning of the sound output of the fork. Note particularly that the frequency of waxing and waning is entirely dependent on the speed of the rotation of the fork and has nothing to do with the frequency of the emitted sound.

It will be clear that as the interference is created because the fork has two sound sources, interference can be created by any other two sound sources suitably placed. Two speakers can be placed in a non-reflecting testing chamber, with their axes at an angle. Walking about in the chamber will indicate that at certain points the sound emitted by the two speakers (working at the same frequency and with the same output) is no longer audible. If the input to one of the speakers is reduced the null points will be at some other situations. Now drive one speaker in an ordinary reflective auditorium, such as an ordinary living room, and points where nothing can be heard are fairly numerous. In this case the second sound source is, of course, the reflections from the walls and ceiling (Fig. 104-a).

The formula given earlier for out-of-phase sound waves refers to the lag or lead of one wave with respect to the other (as in ordinary ac calculations) and the lag or lead is determined by the efficiency of the reflecting surface and its distance and angle with respect to the original sound source. The mathematical analysis is comparatively easy but measurements within the test chamber are rather difficult. The matter has been mentioned to explain *standing waves,* which are not waves at all but points of null displacement of the air.

Standing waves

Although the matter of standing waves does not come under the heading of perception of sound, being rather a problem in architectural acoustics, it will be obvious from what has been said that the interference points can be substantially modified by reducing or removing reflections from the walls of the listening room or auditorium. In an ordinary living room not very much can be done except by using carpets, curtains and drapes. Acoustic tiles, such as are used in properly designed nonechoing studios are somewhat costly, difficult to fix and liable to interfere with the domestic decor. Table 2 gives the relative absorption characteristics of various furnishing materials, the figures representing the amount of sound absorbed by the material. An open window obviously reflects nothing, so the figure of merit in this case is unity. All other materials reflect some of the sound, and a study of the table shows that they reflect a fairly high proportion. This accounts for the fantastically elaborate treatment required for anechoic chambers, where the reflection from the walls, floor and ceiling must be as near as possible nonexistent.

Lace and nylon curtains are almost completely transparent to sound and so their presence or absence makes little difference to the properties of the windows or walls behind them. Drapes to be effective should be thick—velvet, cretonne, burlap (art jute) or similar substantial material. Wallpaper affects the properties of the wall material according to the thickness and texture: a thin glazed paper on an insulating panel would impair the absorption especially at higher frequencies; a thick rough-surfaced paper on plaster would increase absorption on a plaster wall. Rooms lined in wood can be treated only with drapes if the wood is used for decorative effect.

Although Table 2 is correct in giving relative absorption for some mean frequency, the absorption depends on the frequency, and absorption of very low frequencies is extremely difficult. Fig.

105 shows the absorption properties of various types of acoustic treatment related to frequency; the fact that the frequency scale is not extended below 125 cps is significant!

Masking effect

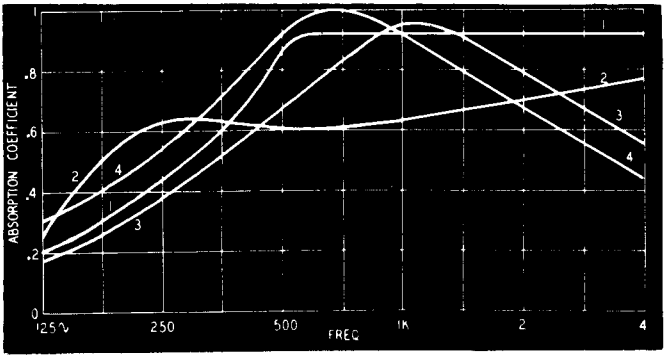
If the two sound waves considered here are not quite the same frequency, *beats* will be created. Assume at any given instant that the positive half-cycles are in synchronism (they will also be in phase); the maximum amplitude of air displacement is $a_1 + a_2$. At the next cycle of a_1 the amplitude a_2 will not be at its maximum; it will lag or lead a_1 and the difference will increase until a whole cycle has been lost or gained on a_1. During this period the sound produced by the combined waves will change from a maximum to a minimum, and the periodicity of the beat will be the difference in frequency between the two waves.

If a certain amount of steady sound or noise is impressed on the ear, it is not so sensitive to other sounds—in the presence of noise the phonic level must be raised. If one sound is being heard and another is increased until the first can no longer be heard, the second is said to *mask* the first. Comprehensive tests were carried out by Bell Telephone Laboratories many years ago, since the matter is important in applied acoustics, and it was found that, whereas a sound of low frequency had to be raised to a high intensity to mask a high-frequency sound, it is easy to mask a sound by one of slightly higher frequency.

Table 2—Absorption properties of the more common furnishing materials within the approximate band of 500 to 1,000 CPS

Open window	1.00
Audiences	0.96 to 0.44
Felt 1-inch thick, ranging from needleloom to piano	0.8 to 0.5
Cork 1½-inches thick	0.32
Insulating boards, such as ½-inch sugar-cane fiber	0.7 to 0.25
Rugs and carpets of various thicknesses	0.3 to 0.2
Velvet curtains	0.25 to 0.2
Cretonne curtains	0.15
Linoleum	0.12
Pine boards	0.06
Plaster on laths	0.034
Glass	0.027
Bricks and hardwall plaster	0.025
Cheesecloth and similar materials	0.025 and less

Fig. 106 gives the masking effects of a 400-cps sound on those of other frequencies. Keeping the intensity of the masking sound constant, the threshold shift denotes the amount the masked sound



1. Acousti-Celotex (cane fiber) 1¼" thick.
2. "Unitone" Acoustic Tile, ⅝" thick.
3. Celotex Absorbex (cement-wood fiber) 1" thick.
4. Asbestos Acoustic Tile with 1" rockwool backing.

Fig. 105. *Absorption properties as related to frequency.*

has to be raised above its normal threshold value to be just audible. For example, a 400-cps sound has to be increased 60 db before it begins to mask a 4,000-cps one. But, as the frequency of the latter falls toward that of the masking sound, the masking effect increases substantially. The fact that the curves cross one another shows that, in case of a complex wave, the ear will hear different harmonics as the intensity is increased.

Combination or subjective tones

If sounds are loud and sustained, *combination* or *subjective* tones are formed. One, at a frequency equal to the difference between the component frequencies, is called the *differential tone;* another has a frequency which is the sum of the component frequencies and is called the *summation tone.* Subjective tones can arise from only one sound wave. At 1,000 cps the second harmonic can be heard at a sensation level of about 50 db; the third, fourth and fifth will be added if the level is raised to 80 db.

If, now, two loud and sustained notes of frequencies f_1 and f_2 are sounded together, the fundamentals will form subjective tones with each other and with the harmonics:

$$f_1,\ f_2,\ f_1 + f_2,\ f_1 - f_2,\ 2f_1,\ 2f_2,\ 2f_1 + f_2,\ 2f_1 - f_2,\ 2f_2 + f_1,\ 2f_2 - f_1,\ 2f_1 + 2f_2,\ 2f_1 - 2f_2,\ 3f_1,\ 3f_2 \text{ and so on.}$$

If the two sounds are complex, the number of subjective tones will be very great. Some of these sounds will be masked by others, so reducing the total audible number which, in any event, is a function of the sound level. See Table 3.

Reaction of the ear to complex sounds

If a filter is used to remove the fundamental of a complex sound, the differential tone that is created by the harmonics will become audible at the fundamental frequency. A commonly met filter of this type will be encountered in the average loudspeaker which cannot reproduce fundamental frequencies at any significant volume below its bass resonant frequency. This does

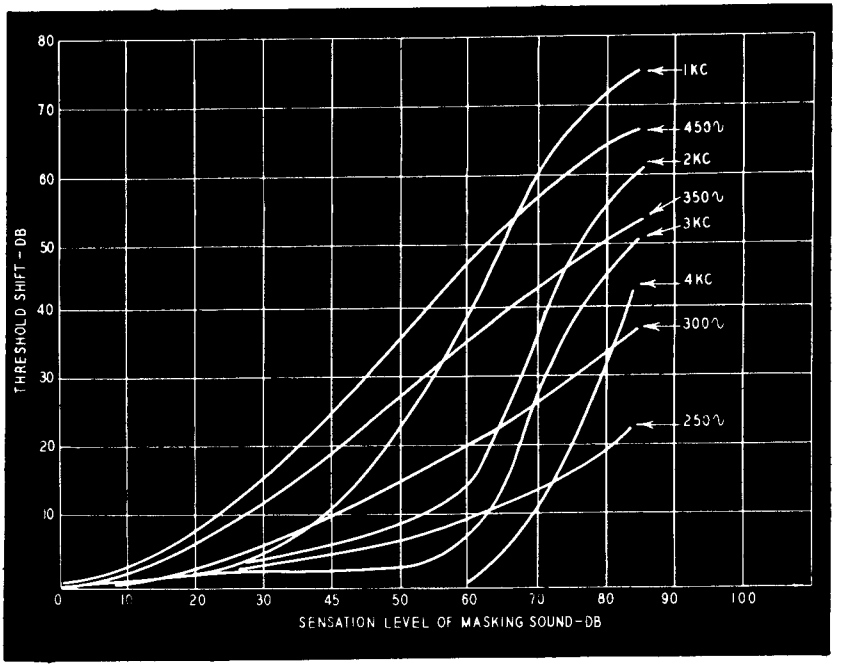


Fig. 106. *Masking effect of a 400-cps sound on those of other frequencies.*

not mean to imply that a loudspeaker without bass resonance is unnecessary for obtaining the best results. There is an audible difference between the aurally reconstructed fundamental bass frequencies and the true bass frequencies that are actually emitted by a loudspeaker.

Suppose a musical instrument sounds a note having a funda-

mental frequency of 50 cps, with harmonics at 100, 150, 200 and so on. Considering only the fundamental and second harmonic, a differential tone will be formed at 50 cps and a summation tone at 150 cps, so the fundamental and third harmonic will be boosted. If the speaker cannot reproduce the 50 cps fundamental, the reproduced series will be 100, 150, etc. which will produce a differential tone of 50 and a summation tone of 250. This gives a series of 50, 100, 150, 200, 250, but the fifth harmonic has been boosted and the original fundamental is only a differential tone from two harmonics. If amplitudes are large, the divergence from reality becomes very great.

At the top end of the frequency spectrum a similar argument applies. Even if the ear could not hear frequencies above 12,000 cycles, it can hear differential tones below 12,000 created by frequencies higher than this. If these unheard frequencies exist in the original performance, the ear will hear the differential tones; if they are cut out by the reproducing equipment, what the ear hears below 12,000 cps is not what it would hear at the original performance.

Basic requirements for fidelity of reproduction

It will now be clear that mere width of frequency response is not enough for high-grade sound reproduction. The frequency response of the reproducing system must also be level. If it is not, peaks and valleys in the response curve will result in nonuniform amplification of both fundamentals and harmonics. The performance of the ear is erratic, even if it conforms to certain general laws, and it is clearly undesirable to complicate the problem further by introducing or failing to eliminate defects in the equipment.

Perfect reproduction of an original performance in an auditorium cannot be achieved in the home. The esthetically equipped expert can only strive to provide a standard which is musically satisfying. This phrase is introduced deliberately because, when science has done its best, the ear itself is the final arbiter and, when the engineer has designed as well as his technique will allow, he must then apply the principles of musical criticism to what he has accomplished.

This being a technical data book, no information is given on musical appreciation beyond the obvious advice that the engineer should make himself thoroughly acquainted with the sounds of musical instruments at firsthand. As a technical worker he should heed the following basic requirements of high-quality reproduction:

1) The overall frequency response of an amplifier should be as wide as the economics of design will permit.
2) Within the predetermined limits of response there should be no audible distortion of any type.
3) The mean level of sound output must be adjusted to enable the ear to function as nearly as possible as a linear transducer.

So far as the first two are concerned, the requirements can be met when designing the amplifier but are not easy to secure from the loudspeaker. Most engineers are compelled to use speakers designed by some other person and all speakers have more or less

Table 3—Effect of a bass filter cutting off fundamental frequency

Subjective tone	A	B
f_1	50	100
f_2	100	150
$f_1 + f_2$	150	250
$f_1 - f_2$	50	50
$2f_1$	100	200
$2f_2$	200	300
$2f_1 + f_2$	200	350
$2f_1 - f_2$	0	50
$2f_2 + f_1$	250	400
$2f_2 - f_1$	150	200
$2f_1 + 2f_2$	300	500
$2f_1 - 2f_2$	100	100
$3f_1$	150	300
$3f_2$	300	450
$3f_1 + f_2$	250	450
$3f_1 - f_2$	50	150
and so on.		

A: subjective tones formed by 50 cps fundamental and second harmonic.
B: subjective tones formed by second and third harmonics.

serious defects. The speaker and its housing should not have a bass resonance and should be free from peaks and valleys in the response curve. Peaks in the region 2,500 to 15,000 cps are particularly objectionable in that they give rise to harsh and "brittle" reproduction. If a suitable speaker cannot be found, it is better

to restrict the overall width of response than to provide opportunity for the speaker to demonstrate its defects. The result will not be fidelity but it will be more acceptable to a musically-minded listener.

So far as the third requirement is concerned, tone-compensated volume controls will neutralize, to some extent, the nonlinearity of the ear at very low and very high phonic levels, but even this is affected by the size and nature of the listening room.

The listening room

Generally speaking, nothing very much can be done in the way of altering the shape or size of a room. The room is there and has to be used, but a great deal can be done to improve its existing acoustical properties. This has to be considered from two points of view—elimination of external noise which would interfere with enjoyment of the music, and modification of the internal reflections and reverberation.

Reflection has been dealt with in Table 2; reverberation should be mentioned as something like an echo which causes the sound to continue after the sound source has ceased to emit sound. Reflection in rooms being constant, the larger the room the longer the reverberation; the more reflective the room the more intense the reverberation. There was the well known lecture room at Harvard University which had a reverberation period of 5.5 seconds for the human voice. Even a slow speaker utters several syllables in this time and the result was an unintelligible jumble of sounds if the room was not full of people. Somewhat primitive treatment helped considerably—putting cushions on the seats to help absorption when the audience was small. Today proper acoustic treatment would get over the difficulty. On the other hand there is a celebrated case in London of a subway tunnel leading to the Science Museum. The reverberation period in this glazed-brick-lined tunnel about 200 yards long is something like 15 to 20 seconds, and has always been a source of amusement for bellowing schoolboys doing a museum jaunt. Nothing can be done, for the proportions of the "room" are hopeless. Fortunately we do not have to listen to our music in subway tunnels, and the ordinary domestic living room can be treated with a considerable measure of success.

Unfortunately, if the music room is the listening room, the piano will have to go somewhere else. It is a strongly resonant device, with tuned resonances over the whole compass. Even

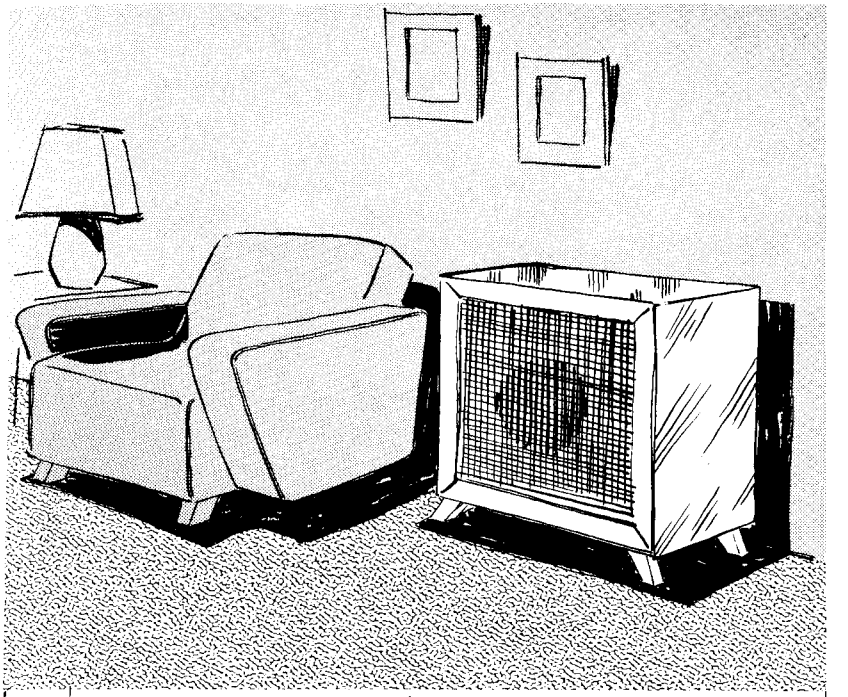
vases and other room ornaments can give rise to resonances. A completely dead room would be lifeless for high-quality reproduction, since "attack" is needed as well as overall frequency response, but undue reflection causes standing waves and reverberation. Unglazed bookcases fairly full of books are good sound absorbers; if glazed, their absorption becomes that of glass. Unupholstered furniture can be assumed to have the same acoustic absorption as pine boards in Table 2. The thick heavy easy chair is better acoustically than modern functional designs. A fitted carpet is more effective than a polished floor with rugs. Uncurtained picture windows are very nearly hopeless. Uncased radiators and heating panels can be very troublesome.

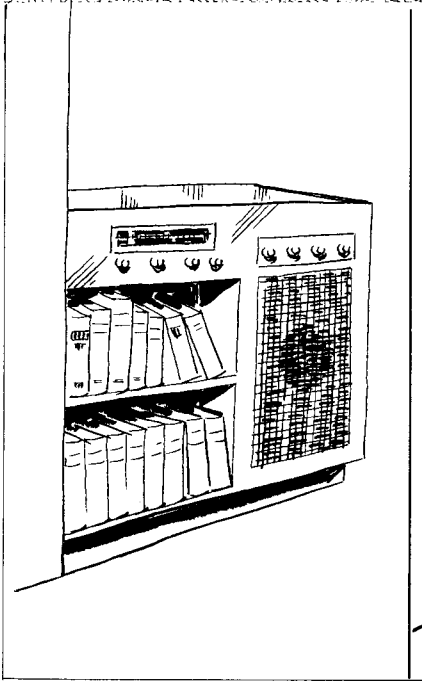
If it is desired to fit acoustic tiles to an otherwise highly reflective wall, the tiles should not be fastened directly to the wall. Wood battens an inch thick should first be applied and the space between the battens filled with Fiberglas or rock wool; then the tiles are applied to cover the whole area. But whether acoustic tiles are used or not, it will pay to give some attention to corners, especially the angles between walls and ceiling. It is not unusual to have the ceiling join the walls with a "cove," a concave plaster surface. This will focus sound as a concave reflector in a car headlamp focuses the light from the bulb. If treatment of the room angles is undertaken, the use of convex corner fillers will help dispersion of the sound to a great extent. In broadcasting studios movable flat or convex reflectors are used to modify the characteristics to suit the program, but these refinements are not necessary for home listening. But angles should be given a convex contour for the best results.

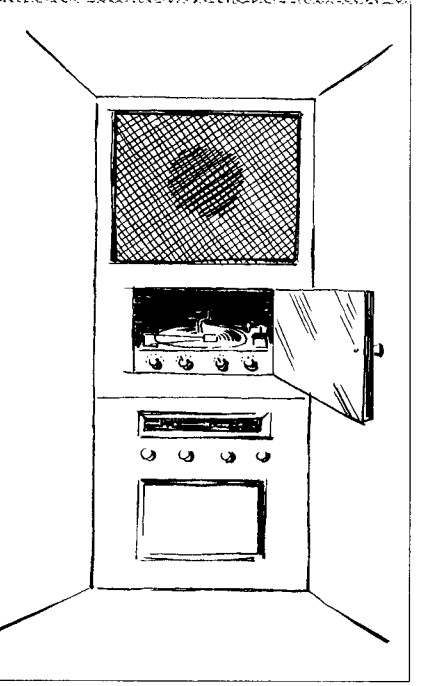
Speaker position

The position of the speaker is very important. Unless mounted in a corner horn, it does not follow that a corner is the best position. Nor, on the other hand, is it likely that a speaker mounted in a closed box forming part of an audio montage across a whole wall will be successful. The present writer may have earned a reputation for being different and difficult, but he feels that an audio system is something which should be tucked away out of sight. A record player and a control box must be conveniently at hand, but they need be no larger than necessary. The speaker and its enclosure have to be heard and obviously should be placed

Fig. 107. *The speaker is sometimes housed in its own cabinet or is made to form an integral unit with the rest of the system.* →







where they sound best, and this does not bear any relationship to the most convenient place for operating the music source. If the speaker has a very good bass response, the audio montage (if part of the speaker cabinet) will be unusable for no turntable spring mounting appears able to insulate the record and pickup from the vibration set up by a speaker in the same housing. A high-fidelity audio system is something to be heard, not seen, and the more that can be put away out of sight the better. This then demands that the speaker should be a separate unit, and being a separate unit it can be placed where it sounds best.

Figs. 107 and 108 illustrate a few of the many arrangements being used and not endorsed by the author.

Multichannel speaker systems

In a multichannel speaker system it is logical to assume that speaker diaphragms should be of similar material and in any event not metallic, which everybody knows has a ringing quality; our diaphragms should be as inert as possible. What is just as important is the spatial relationship of the two or more diaphragms. If you set up two identical speakers some distance apart and drive them with the same signal you will get quite an impressive imitation of stereophonic reproduction. It is not true stereophony since only one channel is used, but the effect is noticeable as long as you are not equidistant from the two speakers. This effect is most noticeable when, if the two speakers are in the two corners at the ends of one wall, you sit near an adjacent wall.

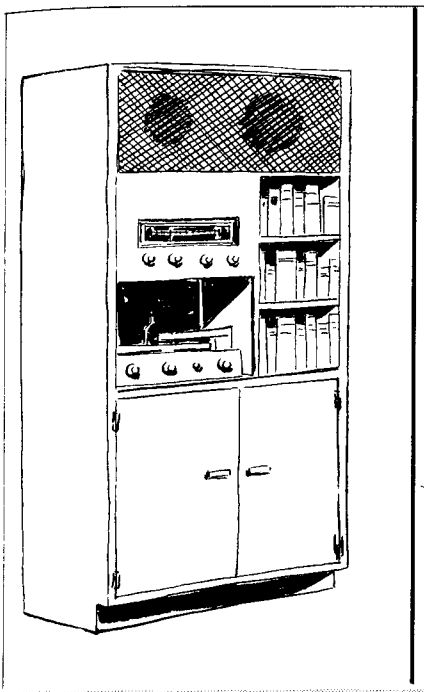
The effect is due to the sound from one speaker being out-of-phase, to some extent, with respect to the other, since the sound takes a little longer to travel from the more distant speaker. The two outputs are combined in the human hearing mechanism to create an illusion of depth; but the effect will only be obtained if the two speakers reproduce the whole frequency range. If one of the speakers is a tweeter and the other a woofer, all you hear are the two separate outputs, the treble coming from one corner, the bass from the other, no matter where you are sitting in the room.

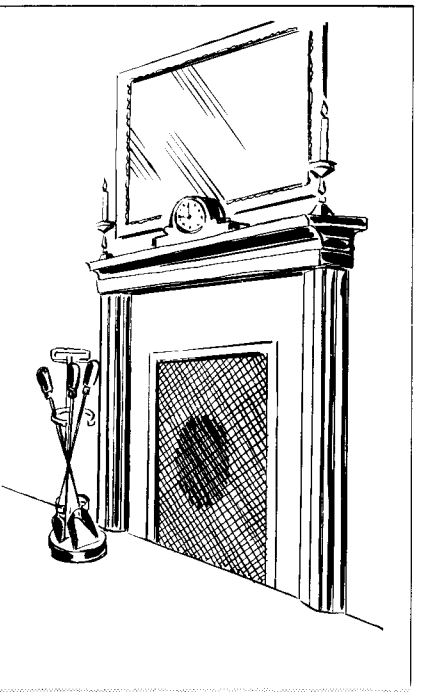
It follows, therefore, that a tweeter-woofer combination must be so disposed that the two sound sources are as close together as possible. Ideally, they should coincide, which accounts for the development of coaxial speakers.

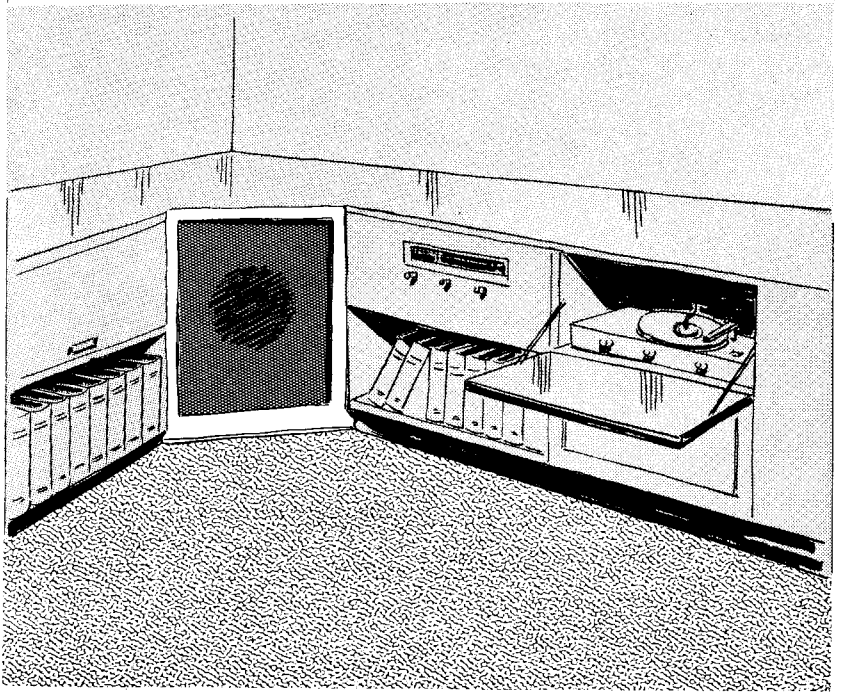
I believe that the best way of laying out a dual range speaker is

---

Fig. 108. *The illustration on the facing page shows a number of possible ways of positioning speakers. (These arrangements are not endorsed by the author.)*







to have each unit horn loaded, so as to avoid interaction between the two units (horns being much more directional at the sound source), and if the tweeter horn can be curved into the mouth of the woofer horn, coaxiality is achieved.

chapter 2

audio amplifiers:
the output stage

A designer of a complete audio installation will often have to acquire as complete entities speakers, pickups, microphones and tape recorders. These will be housed in various ways, the methods of doing which are not within the domain of electronics, and are not critical as to performance, although the housing of a speaker is a problem in acoustics. But he will nave to design the amplifier and an amplifier *can* be designed precisely and absolutely. Its performance can be measured by exact technical procedure. The design of an amplifier is, therefore, an engineering problem.

Design considerations

Many hundreds of different designs have appeared over the years and it is not necessary for the audio hobbyist, technician, or engineer to start from scratch and work the whole thing out from first principles. Nevertheless, unless the general principles are understood and applied, the problem will only be solved, if ever, by trying out circuit after circuit until the desired performance has been secured. This process is not engineering, but its popularity is testified by the constant stream of articles on the subject in technical magazines.

The design must be conditioned by the economic factors preceding the actual technical work. The overriding consideration may be cheapness or it may be the highest possible standard of excellence. The designer of a line amplifier for telephone service will be required to provide an exceptional degree of reliability

over a long period, yet may not have to provide an extremely wide frequency response. An audio engineer may have to produce a circuit which will prove the claims of the sales department of a concern offering high fidelity at a low price, when reliability clearly cannot be a first consideration.

These varying requirements cannot be met by the "perfect" amplifier, although in actual fact a nearly perfect amplifier can be designed. Such an amplifier would be much too expensive for most uses. The designer must, therefore, be sure that the prerequisites have been clearly stated before he proceeds with his work. Given these he should then produce the best possible design for the particular object in view. These remarks may seem commonplace, but it is rather surprising how often the requirements are only specified precisely after a design has been prepared. The man who wishes to create an amplifier for some particular purpose often finds himself under a grave disadvantage. He may consult the usual works of reference to find out just how to put the circuit together and discover so many alternative ways that he does not know which to choose. Then he reads an article in a magazine, written by someone who obviously has had the same sort of problem confronting him, and supposes that the writer has done his work for him. He tries a hookup and finds that the thing doesn't work. But if it doesn't work, it doesn't mean that the writer's design was imperfect. For an article to appear in a responsible magazine the amplifier had to be made, tested, photographed and described with preciseness. When it is copied by the reader, it may be that the copyist's components are all on the edge of the permissible tolerances specified in the original design and the cumulative errors may be enormous. Before such errors can be rectified it is necessary to test the amplifier stage by stage, so investigation of every part of the circuit is a necessary consequence to designing.

An amplifier can be designed exactly by application of the known rules but, before the amplifier can be guaranteed to conform to specifications, it must be ascertained that the constituent parts comply. In view of the high cost of close-tolerance components, it is obviously desirable to evolve a design which is not unduly critical as to values.

The present writer has examined over 150 amplifiers which have appeared on the market during the past six years. They show wide variations in design and reflect the requirements placed before the technical staffs of various manufacturers. A general pat-

tern emerges from this study, although designers will adopt widely different methods of achieving a precisely similar objective. They cannot all be right, for the best method is the simplest one which will do what is wanted. A multiplicity of tubes is not necessarily a criterion of excellence; rather it may be the reverse by introducing unreliability. A multiplicity of controls may suggest that the ultimate user has a greater opportunity of modifying the performance to suit his particular requirements. But even this can be dangerous, for an unskillful operator may so adjust them that the

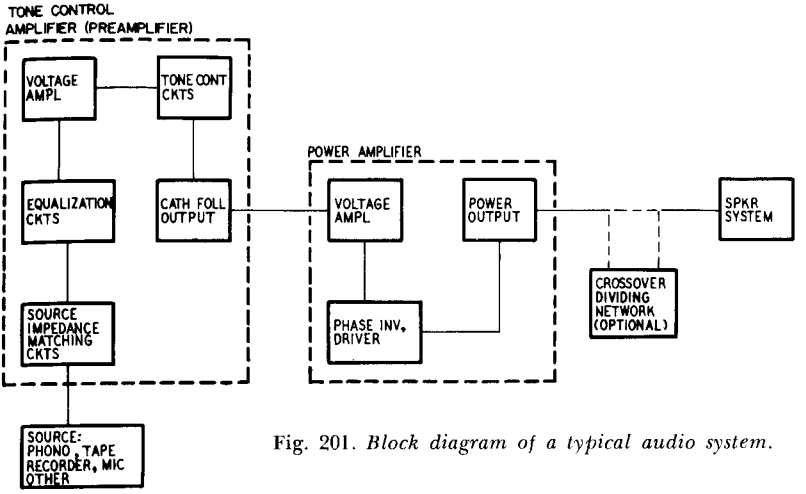


Fig. 201. *Block diagram of a typical audio system.*

resulting performance is something that can only bring discredit on the guiltless designer. Controls are certainly needed to adjust volume and tonal balance, particularly in view of what has already been stated as to the behavior of the human ear, but excessive control facilities, although perhaps good from the "sales angle," are not necessarily good from the point of view of ultimate performance. See Fig. 201 for one possible arrangement of a typical audio system.

Such matters apart, the pattern of the modern amplifier is soon described. The output (power) stage is ordinarily push-pull, sometimes triodes, more often tetrodes are used; the driver stage is almost equally divided between single and double triodes; phase inversion is rarely by transformer. The preceding voltage amplifier is a triode in the great majority of cases—indeed the pentode seems quite out of favor as an audio amplifier in nearly every position. Tone controls are mostly continuously variable. Except for some bias and plate supplies for separate preamplifier units,

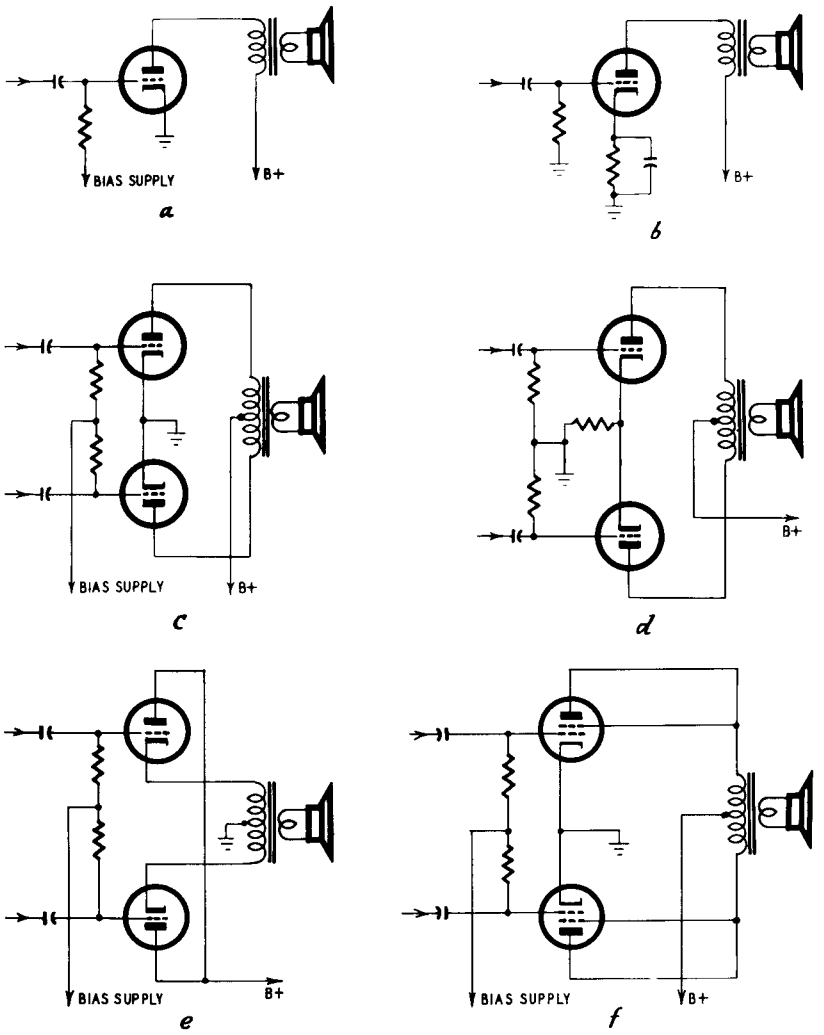


Fig. 202. *Representative triode output stages: a) single-ended, fixed bias; b) single-ended, cathode bias; c) push-pull, fixed bias; d) push-pull, cathode bias; e) cathode follower, push-pull; f) push-pull tetrodes, triode connected*

tube power supply is invariably drawn from thermionic rectifiers.

Consideration of the relative desirability of these techniques is best divided into sections. Accordingly the design of amplifiers will be treated under various headings: power-output stages; driver stages, including phase inversion; voltage amplifiers, including so-called preamplifiers; tone control circuits; output and other audio transformers; rectifiers and power supply.

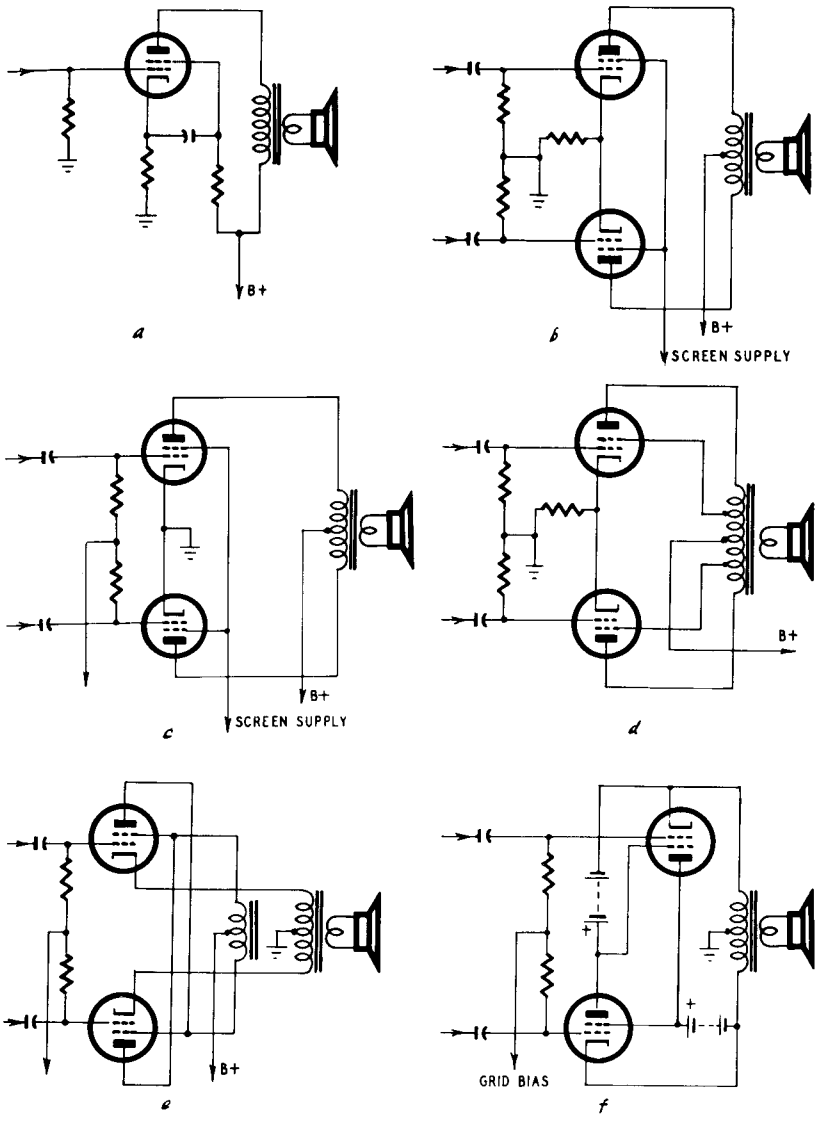


Fig. 203. *Representative tetrode output stages: a) single-ended; b) push-pull, cathode bias; c) push-pull, fixed bias; d) Ultra-Linear; e) unity coupled; f) fundamental Circlotron circuit.*

Power-output stages

A power-output stage can consist of a single triode, tetrode or pentode; push-pull triodes, tetrodes or pentodes; with or without negative feedback. Except in the cheapest equipment, the single

output tube is rare, push-pull operation being almost invariably used. Twenty-five years ago it was not unusual to see triodes used in parallel, obviously to secure greater output power, but as the same two tubes used in push-pull will give even greater power and with less distortion, such applications will not be discussed. See Figs. 202 and 203.

Tubes are paralleled in push-pull stages so as to get greater output; this doubles, triples or quadruples the power available from a single pair of tubes and is a justifiable expedient when large tubes are not available. In this respect British manufacturers have always had a wider range of power triodes than American. Tetrodes can, of course, be used as triodes, but what the American designer lacks is a range of large power triodes. Some data on British power triodes are included in Table 6 as a matter of interest. Table 4 includes characteristics of tetrodes when used as triodes.

In the absence of negative feedback, a triode gives less distortion than a tetrode or pentode and has a better damping effect on the reactive load of a speaker. But the distortion from a single triode is substantial at full output and so the single tube cannot be used in better-grade equipment. From the point of view of efficiency, the tetrode is far better than the triode, so if a little more distortion is tolerated in the final design, the single tetrode output stage is almost universal in inexpensive apparatus. Data for single tetrodes is given in Table 5.

Selecting the tube

Of the large number of amplifiers examined, the great majority uses two 6V6 tubes (as tetrodes) in class-AB_1 push-pull. Less frequently we find the 6L6 or 5881, also in class AB_1; occasionally the 6B4-G (a 6.3-volt-heater octal-base version of the 2A3). As a result of the popularity enjoyed for a considerable time by a certain English circuit the English KT66 is sometimes employed, connected as a triode. One English amplifier available in the United States uses the KT61, connected as a triode. The larger tubes, such as EL34 and 6550 are appearing in more recent designs. We can now examine the behavior of these various output tubes. Data are given in Table 6 for push-pull triodes and in Table 7 for push-pull tetrodes.

The data is that furnished by tube manufacturers; they are average values, since, obviously, tubes can vary within certain limits. The figures given for power output are always based on the assumption that the load is purely resistive, but when a speaker

is used as the load it is both resistive and reactive. This complication will be discussed later.

Tables 4 to 7 are designed to enable the reader to compare performances of various tubes with the minimum of trouble. The facts given are sufficient for designing purposes—a first selection can be made without having to hunt through various tube manuals.

Triodes vs. tetrodes

Used singly, triodes give less distortion than tetrodes but more than is acceptable in high-grade equipment. Triode distortion is mainly second harmonic, which will be cancelled by push-pull operation. Tetrode distortion contains second, third and fourth harmonics. Second and fourth will be cancelled by push-pull working, but not third.

Tube transconductance

The transconductance of the tube is a measure of its grid-voltage—plate-current conversion efficiency but is not of importance except when the minimum number of amplifying stages is demanded on the grounds of cost. In general terms, an output tube with high transconductance requires less input grid swing or plate-supply volts to generate a given amount of audio power than one with lower conductance. In a simple receiver employing the minimum of tubes, the plate-supply voltage is limited and it may be desirable to drive the single output tube direct from the signal rectifier. A tube such as the 50L6 would meet the need, but there is no point in using such tubes in push-pull in more elaborate equipment where voltage amplification can be carried out quite simply.

Tables 6 and 7 show that push-pull triodes do not necessarily give less distortion than push-pull tetrodes, and in some cases are a good deal worse than tetrodes, assuming that the load is purely resistive. For example, push-pull 6L6's with cathode bias deliver 13.3 watts with 4.4% distortion with a plate voltage of 400, when used as triodes. The same tubes as tetrodes in class A_1 produce 18.5 watts with only 2% distortion with a plate voltage of 270, or in class AB_1, 24.5 watts with 4% distortion with a plate voltage of only 360. Admittedly, the choice is not quite so easy as this, since tetrodes are critical as to the value of load resistance and because of their high internal resistance do not offer the same damping to the speaker as triodes. These disabilities can, however, be partially overcome by negative feedback.

Class-AB_2 amplification demands fixed bias, but fixed bias gives less distortion, as a rule, than cathode bias. The simplest to provide, cathode bias is in some measure self-compensating for variations in plate supply voltage; but designers of high-grade and efficient amplifiers should avoid the habit of using cathode bias without thinking of the other method.

Power supply stabilization

Class-AB_1 (and to an even greater extent with class-AB_2) amplification results in wide changes of plate and screen current whether triodes or tetrodes are used. These current changes must not be accompanied by power-supply voltage changes, otherwise the performance deteriorates. Accordingly the supply voltage must remain constant over wide changes of current, and in the case of tetrodes it is highly important that the screen voltage be stabilized. It is easy and customary to connect the screens of the tubes to the same lines as the plates but this is detrimental to the best results.

The impedance of any speaker varies with frequency and will generally be several times the nominal impedance at high frequencies. This is the part of the frequency spectrum which reproduces the harmonics, so harmonic distortion is basically liable to be severe. If the distortion is further increased by tubes operating away off their normal conditions because of heavy signal inputs, the amplifier ceases to be what the designer intended. Many amplifiers, taking the easy way out, could be improved out of all recognition by having the plate and screen supply voltages stabilized. Unfortunately, this adds to the cost of the power supply.

A small point can be brought up in assessing the relative merits of various tubes. The KT66 is, in some quarters, regarded as a mysteriously wonderful tube, but this results from a failure to assess all the circumstances. At least as a tetrode, the 6L6 is not safe to use at higher plate voltages than about 385, owing to the design of the glass pinch inside the tube (Table 7 shows that at 250 to 270 volts the 6L6 is actually better than the KT66), but the difficulty can be overcome completely by using the 807, which has the same characteristics as the 6L6, but which can safely be used at voltages up to 600.

The KT66 characteristics differ from those of the 6L6, 5881 and 807, but whereas those characteristics may be more suitable in certain cases, it is neither a better nor worse tube than currently available American types. Like the 6L6 it is a pre-World War II design.

Single-triode output stages

The data given in Table 4 represent the normal state of the tube in the absence of signal. Taking the 2A3 as an example, a grid bias of —45 volts suggests the tube can accept a signal input of 45 volts peak ac, which means that the control grid will vary in voltage from 0 to —90. The behavior of the tube under these varying conditions is displayed in a family of plate-current—plate-voltage curves, illustrated by Fig. 204. These curves are typical of those appearing in tube manufacturers' manuals.

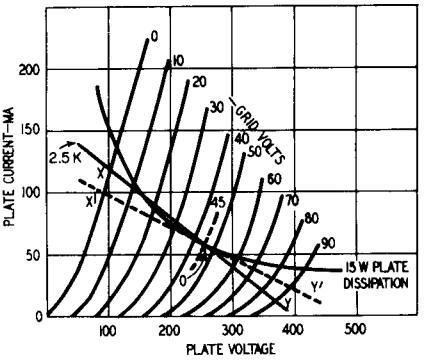


Fig. 204. *Plate-voltage—plate-current curves of the 2A3 showing load line and operating conditions.*

Each curve shows the plate current plotted against plate voltage for grid bias at 10-volt steps between 0 and —90 volts (twice the recommended bias). A short curve is drawn representing the specified normal bias for this tube, —45 volts. Also shown is the maximum permissible plate dissipation, 15 watts, a figure which can be checked by referring to the current and voltage values at any point on the curve.

Load line

No tube can be expected to give good life if operated beyond its maximum plate dissipation rating. As a datum point, therefore, we select that point where the —45-volt curve is intersected by the power dissipation curve. Call this point O. The load line is now drawn; its gradient is determined by the resistance it represents. If a line is drawn joining the points at 250 volts on the voltage axis and 100 ma on the current axis, it would represent a resistance of 2,500 ohms, since:

$$\frac{250 \text{ volts}}{0.1 \text{ amperes}} = 2{,}500 \text{ ohms}$$

Any line drawn parallel to this line would also represent a resistance of 2,500 ohms, so such a parallel line is drawn through the point O, marked XOY. This is the load line for a 2A3 with a purely resistive load of 2,500 ohms. This is illustrated in the drawing of Fig. 204.

A peak signal of 45 volts applied to the grid will cause the operating point to move from O to X, back to O, from O to Y and back to O, but it will be seen that OX does not equal OY. Because of this, distortion will occur for an examination of the figure shows that, whereas a change of voltage on the grid from —45 to O produces a voltage change of 150 on the plate, a change on the grid from —45 to —90 volts produces a voltage change on the plate of only 115. Similarly the current change in the first instance is 60 ma, and in the second 46 ma. In the case of triodes this distortion is mainly second harmonic and can be calculated from the formula that is shown below. The power output can also be calculated.

The point X is the point of minimum voltage and maximum current; Y is the point of maximum voltage and minimum current.

$$\text{Power output} = \frac{[I_{max} - I_{min}] \quad [E_{max} - E_{min}]}{8}$$

$$\%\ \text{2nd harmonic distortion} = \frac{\frac{1}{2}\,[I_{max} + I_{min}] - I_0}{I_{max} - I_{min}} \times 100$$

where I_o = current at the normal operating point O.

In the case illustrated:

$$\text{Power output} = \frac{[117.5 - 12.5] \quad [370 - 105]}{8} = 3.478 \text{ watts}$$

$$\text{2nd harmonic distortion} = \frac{\frac{1}{2}[117.5 + 12.5] - 60}{105} \times 100 = 4.76\%$$

These figures will be seen to be similar to the data given in Table 4.

Note also that the load line passes above the power dissipation curve in the region between —20 and —40 volts. As this happens only momentarily with any musical signal it is of no consequence, but a tube must never be operated in a steady state beyond its maximum rating.

As a matter of interest, a load line may be drawn which is tangential to the dissipation curve; such is shown by X'OY'. The gradient of the line is changed and in fact represents a load of just under 4,000 ohms. The power generated is also changed, as is the distortion, the difference between X'O and OY' not being so great as between XO and OY. For a load of 4,000 ohms the power output is 2.9 watts and the distortion 1.6%

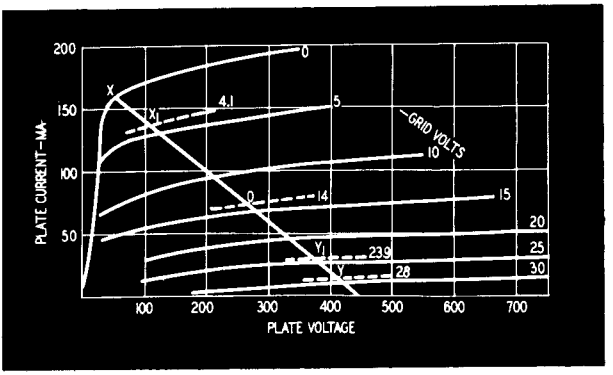


Fig. 205. *Plate-voltage—current curves of the 6L6 showing load line and operating conditions.*

The average plate current changes with application of the signal. It is important to see that the change of current is not accompanied by a change of dc voltage supplied to the plate circuit. For the comparatively small changes of current occurring in class-A power stages, a sufficiently large terminating capacitor in the power supply filter is sufficient to meet the case. In general, anything less than 8 μf is not recommended for good-quality reproduction and can with advantage, be greater.

Slope of the load line

The slope of the load line represents the effective impedance reflected across the output transformer primary. This winding has resistance; call it R' ohms. The slope of the load line is not altered by this, but the actual working plate voltage is less than the supply voltage by the amount $R'I_o$. The transformer primary also has reactance, as has the speaker itself. A reactive load converts the load line into an ellipse. This disadvantage can be overcome to a great extent by increasing the transformer primary inductance. Further information on this point will be given in the section dealing with output transformers.

A simple design trick, which the writer first met 28 years ago,

is available for drawing a suitable load line to give 5% distortion. For this amount of distortion, in Fig. 204 the ratio of XO to OY is as 11 to 9. If now, a straight piece of cardboard in the form of a ruler about 8 inches long is marked with a center zero and is also marked off in divisions to left and right of this center zero in the ratio of 11 to 9, the rule can be placed so that the center rests on the point of operating bias (in the case of the 2A3 at −45 volts) and pivoted about this point until the curves at minimum and maximum bias register on equally marked lines on the rule. A line drawn along the edge of the rule will then give a load line representing a value for 5% distortion. For convenience, the divisions to the left of zero can be made 11 mm and to the right 9 mm. The divisions should be marked 1 to 10 to left and right and should also be subdivided by half-way shorter marks. Visual interpolation will take the place of smaller subdivisions.

Single tetrode output stages

Normal operating data for tetrodes in class A_1 are given in Table 5. Graphical methods are similar to those for triodes but somewhat more complicated. Fig. 205 gives the plate-current–plate-voltage curves for the 6L6 tube. The load line is shown by XOY as before but, as the power output and distortion vary widely with change of load, the tetrode being critical as to correct loading, a different technique must be used.

The load line in Fig. 205 represents a load of 2,500 ohms, as quoted in Table 5. The difference between XO and OY is quite appreciable. This is reflected in the table, which shows that the second-harmonic distortion is 10%. For optimum operation of a tetrode a method of trial and error must be adopted.

Determine the operating point O, whose ordinates are the operating plate voltage and the maximum signal current. From a point just above the knee in the zero-bias curve draw a line through O and continue so that XO is equal to OY. By the standard data provided, O should be on the curve of the specified bias. If on inspection it is found that the change of bias from X to O is *not* the same as that from O to Y, the *load* line is not suitable and another attempt must be made. For example, in Fig. 205, OX and OY are not equal. When the best approximation has been found, the following formula can be evaluated. Point X represents E_{min} and I_{max}; Y represents E_{max} and I_{min}.

$$\text{Load resistance } [R_L] = \frac{E_{max} - E_{min}}{I_{max} - I_{min}}$$

Power output and harmonic distortion

Calculation of the power output and harmonic distortion is a little more complicated. The simplest technique is that usually called the "five selected ordinates method." We have already three ordinates—X, O and Y, the other two are at X_1 and Y_1. These points are the intersections of the load line with the bias curves representing bias of values:

$$V \pm \frac{\sqrt{2}V}{2}$$

where V is the normal bias at point O. Then:

$$\text{Power output} = \frac{[(I_{max} - I_{min}) + \sqrt{2}(I_x - I_y)]^2 R_L}{32}$$

$$\%\ \text{2nd harmonic distortion} = \frac{I_{max} + I_{min} - 2I_o}{I_{max} - I_{min} + \sqrt{2}[I_x - I_y]} \times 100$$

$$\%\ \text{3rd harmonic distortion} = \frac{I_{max} - I_{min} - \sqrt{2}[I_x - I_y]}{I_{max} - I_{min} + \sqrt{2}[I_x - I_y]} \times 100$$

$$\%\ \text{total 2nd and 3rd harmonic distortion} = \sqrt{[\%\ 2nd]^2 + [\%\ 3rd]^2}$$

where I_o, I_x and I_y are the plate currents at points O, X_1 and Y_1 in Fig. 204.

A complete picture of the behavior of the tube can therefore be plotted by evaluating these equations for different load lines. Such curves are given in the RCA tube handbook and need not be repeated here. The third harmonic is liable to be more objectionable than second and low loading of a beam-power tetrode gives less third harmonic distortion than high loading; it also gives less power output. Accordingly, a nice balance must be made between the power output desired and the amount of harmonic distortion permissible. As these tubes give much less third-harmonic distortion than second with low loading, it is clearly advisable to use them in push-pull, which cancels even-harmonic distortion.

Push-pull triode output stages

Fig. 206 shows the plate-current—plate-voltage curves for two 2A3's in push-pull. The upper half of the diagram is identical with Fig. 204 except that intermediate values of bias have been omitted for the sake of clarity. The lower half of the diagram is precisely similar except that it has been turned through 180° and the plate-voltage axes are made to coincide with the operating points, in this case 300 volts, superimposed. The grid-voltage curves are shown in Fig. 206 as dashed lines. From these, derived characteristics (shown as solid lines) are obtained by aver-

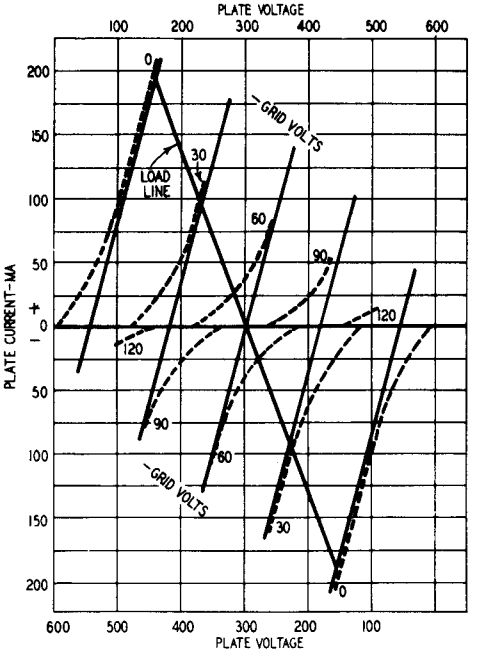


Fig. 206. *Combined characteristics of two 2A3 tubes in push-pull.*

aging the plate-current—grid-potential curves corresponding to the same applied signal. The "curves" shown in the figure are straight lines, and for class-A amplification they will be nearly straight. As one passes to class AB_1, and eventually to AB_2, they will depart more and more from the straight line, but this divergence does not invalidate the basic argument.

For the operating bias, —60 volts, the derived curve represents no signal on the grids, so the two —60-volt curves are averaged. If

the signal on the first grid is 30 volts, the net bias is −30 volts; but the signal on the second grid is −30 volts, so the bias on the second grid is −90 volts. Accordingly, the curves for −30 and −90 volts are averaged. And so on. The derived characteristics represent the relationship between the plate-to-cathode voltage (half the plate-to-plate voltage) and the difference between the two plate currents (twice the alternating current in a plate-to-plate load resistance). Accordingly, the load line shows output voltages and currents when the output transformer has a stepdown ratio of 2.

Matching tubes

Push-pull tubes, for optimum results, should be matched. Maximum power output is obtained when the plate-to-plate load resistance is four times the plate resistance of the composite tube, the composite tube being that whose characteristics are the composite ones shown in the figure, or roughly twice the plate resistance of one of the tubes at the quiescent operating point. Any load resistance less than this will cause a loss of output power, increase distortion, create high peak currents and excessive plate dissipation. The actual value of plate-to-plate resistance in ohms for any load through the point where it intersects the plate-volts axis at the operating voltage is found by projecting the load line to intersect the plate-current axis. Call this point of intersection I amperes and the operating point on the plate-volts axis E volts. Then the plate-to-plate load resistance is 4E/I.

$$\text{The power output} = [I_{max}]^2 \times R_L/8$$

where R_L is the plate-to-plate load resistance. I is, of course, the current at the point where the load line cuts the zero grid-bias curve. Since a perfectly balanced push-pull output stage has no even-harmonic distortion, no calculations of these arise; but the third harmonic can be evaluated by the formula already given, with the notation that the quantity I_0 is now zero, and that "minimum" currents are now negative.

Push-pull tetrode output stages

The same method applies here and composite characteristics can be drawn in a manner similar to that for triodes. Actually, for calculation of the power output of tubes of the 6L6 type, where third-harmonic distortion is small, composite characteristics need not be drawn. In Fig. 204, if the load line is drawn from the

zero grid-volts curve through the operating voltage at zero current, the power output = $\frac{1}{2}$ I_{max} [$E_{max} - E_{min}$].

As with single tubes it is essential that the screen voltage be held constant within narrow limits for maximum power with minimum distortion; with class AB_1, the plate voltage should also be well regulated. In all cases the correct load resistance is much more important than with triodes.

Class B and class AB_2 output stages

These are used for maximum economy of design when large outputs are required. Obviously the plate dissipation of a tube (ac+dc) cannot be exceeded without shortening its life. If the quiescent current is reduced (by increasing grid bias) the dc dissipation is reduced and dissipation for ac power increased. In class-B amplifiers the bias is increased to the cutoff point; one tube handles one-half cycle of the signal, the other deals with the other half cycle. A reasonably close approximation for power output and load resistance can be obtained by assuming that the characteristic curves are parallel straight lines when:

$$\text{Load resistance, plate-to-plate} = 4\,\frac{E_p - E_{min}}{I_{max}}$$

$$\text{Power output from two tubes} = \frac{I_{max}\,[E_p - E_{min}]}{2}$$

where E_p is the plate supply voltage; E_{min} the minimum plate voltage during one cycle, and I_{max} the peak plate current.

The bias must remain constant over wide variations of plate current and so must be provided from a battery or separate power supply.

Class-AB_2 amplifiers are intermediate between class B and class AB_1. The plate current variation is less than with class B (about half as much) and the amplifier is a compromise between the high plate efficiency of class B and the low distortion of class AB_1. Again, fixed bias is essential.

Table 4—Output triodes in class A_1

Type	2A3	KT61	6L6, 6L6-G 5881	EL34
Fil. or heater volts	2.5	6.3	6.3	6.3
Fil. or heater amps	2.5	0.95	0.9	1.5
Max. plate volts	300	350	275	800
Plate dissipation, watts	15	10	12.5	25
Plate resistance, ohms	800	2750	1700	—
Transconductance, μmhos	5250	9860	4700	—
Amplification factor	4.2	—	8	—
	Fixed Bias	Cathode Bias		
Operating plate volts	250	250	250	375
Grid bias volts	—45	—6	—21	—26
Bias resistor, ohms	—	200	490	370
Plate current, ma	60	30	43	70
Load resistance, ohms	2500	5000	6000	3000
Distortion %	5	5	6	8
Power output, watts	3.5	0.7	1.3	6

Type	KT66		6L6, 6L6-G 5881	6B4-G
Fil. or heater volts	6.3		6.3	6.3
Fil. or heater amps	1.27		0.9	1.0
Max. plate volts	400		275	300
Plate dissipation, watts	25		12.5	15
Plate resistance, ohms	1450		1700	800
Transconductance, μmhos	5500		4700	5250
Amplification factor	—		8	4.2
	Cathode Bias		Fixed Bias	
Operating plate volts	250	400	250	250
Grid bias volts	—19	—38	—20	—45
Bias resistor, ohms	315	600	—	—
Plate current, ma	60	63	46	60
Load resistance, ohms	2750	4500	5000	2500
Distortion %	6	7	5	5
Power output, watts	2.2	5.8	1.4	3.5

The tubes in this table are triodes or triode connected tetrodes. Where the listed tube is a tetrode the information given is for maximum, and typical operating conditions when the plate and screen grid of the tube are tied together.

Table 5—Output Tetrodes in Class A_1

Type	6AS5	6AQ5	12AQ5	6W6-GT
Heater volts	6.3	6.3	12.6	6.3
Heater amperes	0.8	0.45	0.225	1.2
Plate volts	150	180	250	200
Screen volts	110	180	250	125
Control grid volts	—8.5	—8.5	—12.5	—8.5
Zero signal plate ma	35	29	45	46
Max. signal plate ma	36	30	47	47
Zero signal screen ma	2	3	4.5	2.2
Max. signal screen ma	6.5	4	7	8.5
Transconductance, μmhos	5600	3700	4100	8000
Load resistance, ohms	4500	5500	5000	4000
Distortion %	10	8	8	10
Power output, watts	2.2	2.0	4.5	3.8

Type	35C5	35L6	50C5	6Y6
Heater volts	35	35	50	6.3
Heater amperes	0.15	0.15	0.15	1.25
Plate volts	110	200	110	135
Screen volts	110	110	110	135
Control grid volts	—7.5	—8	—7.5	—13.5
Zero signal plate ma	40	41	49	58
Max. signal plate ma	41	44	50	60
Zero signal screen ma	3	2	4	3.5
Max. signal screen ma	7	7	8.5	11.5
Transconductance, μmhos	5800	5900	7500	7000
Load resistance, ohms	2500	4500	2500	2000
Distortion %	10	10	9	10
Power output, watts	1.5	3.3	1.9	3.6

Type	50C6	50L6	6V6 and 7C5	
Heater volts	50	50	6.3	
Heater amperes	0.15	0.15	0.45	
Plate volts	200	200	180	315
Screen volts	135	110	180	225
Control grid volts	—14	—8	—8.5	—13
Zero signal plate ma	61	50	29	34
Max. signal plate ma	66	55	30	35
Zero signal screen ma	2.2	2	3	2.2

Table 5 (continued)

Type	50C6	50L6	6V6 and 7C5	
Max. signal screen ma	9	7	4	6
Transconductance, μmhos	7100	9500	3700	3750
Load resistance, ohms	2600	3000	5500	8500
Distortion %	10	10	8	12
Power output, watts	6	4.3	2	5.5

Type	KT61	KT66	6L6, 6L6-G, 5881	
Heater volts	6.3	6.3	6.3	
Heater amperes	0.95	1.27	0.9	
Plate volts	250	250	250	300
Screen volts	250	250	250	200
Control grid volts	—4.4	—15	—14	—18
Zero signal plate ma	40	85	75	51
Max. signal plate ma	?	?	78	54
Zero signal screen ma	7.5	6.3	5.4	3
Max. signal screen ma	?	?	7.2	4.6
Transconductance, μmhos	10500	6300	6000	5200
Load resistance, ohms	6000	2200	2500	4500
Distortion %	8	9	10	11
Power output, watts	4.3	7.25	6.5	6.5

Type	EL37	EL34	6550	
Heater volts	6.3	6.3	6.3	
Heater amperes	1.4	1.5	1.6	
Plate volts	250	250	250	400
Screen volts	250	250	250	225
Control grid volts	—13.5	—13.5	—14	—16.5
Zero signal plate ma	100	100	140	87
Max. signal plate ma	?	?	150	105
Zero signal screen ma	13.5	14.9	12	4
Max. signal screen ma	?	?	28	18
Transconductance, μmhos	11000	11000	11000	9000
Load resistance, ohms	2500	2000	1500	3000
Distortion %	13.5	10	7	13.5
Power output, watts	11.5	11	12.5	20

This table includes tubes commonly used in low-cost low-power audio amplifiers such as may be found in the output stages of radio and television receivers and inexpensive phonographs.

Table 6—Output triodes. Push-pull operating conditions

Type	2A3		KT66	
Working in class	AB_1		AB_1	
Fil. or heater volts	2.5		6.3	
Fil. or heater amps	5.0		2.54	
Plate volts	300		250	450
	Fixed Bias	Cath. Bias		
Grid bias volts	—65		—	—
Bias resistor, ohms	—	780	180	300
Peak volts, grid-to-grid	124	156	40	80
Zero signal plate ma	80	80	110	125
Max. signal plate ma	147	100	?	?
Load resistance, ohms	3000	5000	2500	4000
Distortion %	2.5	5	2	3.5
Power output, watts	15	10	4.5	14.5

Type	EL34		EL84	
Working in class	AB_1		AB_1	
Fil. or heater volts	6.3		6.3	
Fil. or heater amps	3		1.52	
Plate volts	400	430	250	300
Grid bias volts	—29	—32	—	—
Bias resistor, ohms	220	250	270	270
Peak volts, grid-to-grid	62	68	23.1	28.8
Zero signal plate ma	130	128	40	48
Max. signal plate ma	142	134	43	52
Load resistance, ohms	5000	10000	10000	10000
Distortion %	3	1	2.5	2.5
Power output, watts	16	14	3.4	5.2

Type	6L6, 6L6-G 5881	KT88	KT61	6550
	Cathode Bias			Fixed Bias
Working in class	AB_1	A	AB_1	AB_1
Fil. or heater volts	6.3	6.3	6.3	6.3
Fil. or heater amps	1.8	3.6	1.9	3.2
Plate volts	400	485	350	450
Grid bias volts	—	—	—	—46
Bias resistor, ohms	250	560	150	—
Peak volts, grid-to-grid	90	99	23	92
Zero signal plate ma	65	170	63	150

Table 6 (continued)

Type	6L6, 6L6-G 5881	KT88	KT61	6550
Max. signal plate ma	130	180	73	220
Load resistance, ohms	4000	4000	6000	4000
Distortion %	4.4	1–3	2	2.5
Power output, watts	13.3	27	6	28

Tetrode tubes in this table are triode connected. Values are for two tubes. For the KT88 separate bias resistors of 560 ohms are essential.

Table 7—Output tetrodes. Push-pull operating conditions

Type	6L6, 6L6-G, 5881 Fixed Bias			
Working in class	A_1		AB_1	
Heater volts	6.3		6.3	
Heater amps	1.8		1.8	
Plate volts	250	270	360	360
Screen volts	250	270	270	270
Control grid volts	—16	—17.5	—22.5	—22.5
Bias resistor, ohms	—	—	—	—
Peak volts, grid-to-grid	32	35	45	45
Zero signal plate ma	120	134	88	88
Max. signal plate ma	140	155	132	140
Zero signal screen ma	10	11	5	5
Max. signal screen ma	16	17	15	11
Load resistance, ohms	5000	5000	6600	3800
Distortion %	2	2	4	2
Power output, watts	14.5	18.5	24.5	18

Type	6L6, 6L6-G, 5881 Fixed Bias		Cathode Bias	
Working in class	AB_2		A_1	AB_1
Heater volts	6.3		6.3	6.3
Heater amps	1.8		1.8	1.8
Plate volts	360	360	270	360
Screen volts	225	270	270	270
Control grid volts	—18	—22.5	—	—
Bias resistor, ohms	—	—	125	250
Peak volts, grid-to-grid	52	72	40	57
Zero signal plate ma	78	88	134	88
Max. signal plate ma	142	205	145	100

Table 7 (continued)

Type	6L6, 6L6-G, 5881			
	Fixed Bias		Cathode Bias	
Zero signal screen ma	3.5	5	11	5
Max. signal screen ma	11	16	17	17
Load resistance, ohms	6000	3800	5000	9000
Distortion %	2	2	2	4
Power output, watts	31	47	18.5	24

Type	807			6AQ5
Working in class		AB_2		AB_1
Heater volts		6.3		6.3
Heater amps		1.8		.9
Plate volts	400	500	600	250
Screen volts	300	300	300	250
Control grid volts	—25	—29	—30	—15
Bias resistor, ohms	—	—	—	270
Peak volts, grid-to-grid	78	86	78	30
Zero signal plate ma	90	72	60	70
Max. signal plate ma	240	240	200	79
Zero signal screen ma	2	.9	.7	5
Max. signal screen ma	17	16	12	13
Load resistance, ohms	3200	4240	6400	10000
Distortion %	2	2	2	5
Power output, watts	55	75	80	10

Type	KT66			KT61
Working in class		AB_1		AB_1
Heater volts		6.3		6.3
Heater amps		2.54		1.9
Plate volts	500	250	390	275
Screen volts	400	250	275	275
Control grid volts	—45	—17.5	—22.5	—6.7
Bias resistor, ohms	—	100	250	80
Peak volts, grid-to-grid	90	36	70	16
Zero signal plate ma	80	162	104	72
Max. signal plate ma	175	165	125	?
Zero signal screen ma	3	12	5	12
Max. signal screen ma	21	20	18	?
Load resistance, ohms	6000	4000	8000	10000
Distortion %	5	4	6	6.5
Power output, watts	50	17	30	11.5

Table 7 (continued)

Type	EL34			
	Fixed Bias		Cathode Bias	
Working in class	AB_1		Distributed Load	
Heater volts	6.3		6.3	
Heater amps	3.0		3.0	
Plate volts	425	375	430	430
Screen volts	a	b	c	c
Control grid volts	—38	—32	—	—
Bias resistor, ohms	—	—	470 each tube	
Peak volts, grid-to-grid	76	63	45	73
Zero signal plate ma	60	70	125	125
Max. signal plate ma	240	240	130	140
Zero signal screen ma	8.8	9.2	10	10
Max. signal screen ma	50	50	10.2	15
Load resistance, ohms	3400	2800	6600	6600
Distortion %	5	5	0.8	1.3
Power output, watts	55	44	20	37

Type	EL34	6550		
	Cathode Bias	Fixed Bias		Cathode Bias
Working in class	AB_1	AB_1		AB_1
Heater volts	6.3	6.3		6.3
Heater amps	3.0	3.2		3.2
Plate volts	375	400	600	400
Screen volts	b	275	300	300
Control grid volts	—	—23	—33	—
Bias resistor, ohms	130	—	—	140
Peak volts, grid-to-grid	59	46	66	53
Zero signal plate ma	15	180	100	166
Max. signal plate ma	190	270	280	190
Zero signal screen ma	23	9	3	7.5
Max. signal screen ma	45	44	33	39
Load resistance, ohms	3400	3500	5000	4500
Distortion %	5	3	3.5	4
Power output, watts	35	55	100	41

a: common screen resistor of 1,000 ohms.

b: common screen resistor of 470 ohms.

c: individual series screen resistors of 1,000 ohms from 43% tapping of the output transformer.

Table 7 (continued)

Type	EL84		6V6, 7C5	
Working in class	AB_1		AB_1	
Heater volts	6.3		6.3	
Heater amps	1.52		0.9	
Plate volts	250	300	250	285
Screen volts	250	300	250	285
Control grid volts	—	—	—15	—19
Bias resistor, ohms	130	130	200	260
Peak volts, grid-to-grid	16	20	30	38
Zero signal plate ma	62	72	70	70
Max. signal plate ma	75	92	79	92
Zero signal screen ma	7	8	5	4
Max. signal screen ma	15	22	13	13.5
Load resistance, ohms	8000	8000	10000	8000
Distortion %	3	4	5	3.5
Power output, watts	11	17	10	14

Type	KT88			
	Fixed Bias		Cathode Bias	
Working in class	AB_1		AB_1	
Heater volts	6.3		6.3	
Heater amps	3.6		3.6	
Plate volts	460	625	400	475
Screen volts	345	330	255	320
Control grid volts	—48	—45	—	—
Bias resistor, ohms	—	—	440 each tube	
Peak volts, grid-to-grid	98	70	70	98
Zero signal plate ma	100	100	120	160
Max. signal plate ma	240	250	135	180
Zero signal screen ma	7.5	6	7.5	12
Max. signal screen ma	35	32	25	38
Load resistance, ohms	4000	5000	6000	6000
Distortion %	5–7	3.6	3	3
Power output, watts	65	100	34	48

Distortion in the KT88 depends on accuracy of matching.

chapter 3

audio amplifiers: inverters and drivers

THE two tubes in a push-pull power output stage require input signal voltages of equal magnitude but of opposite phase. The signal voltages driving the push-pull stage must be symmetrical with respect to ground: this operation is carried out by a phase inverter. Normally the signal output of the last stage of the voltage amplifier (driving the push-pull tubes) is sufficient to develop maximum power in the output stage, in which case the phase inverter immediately precedes the output stage itself. Where this is not the case (particularly where the output stage consists of two triodes in class AB_2), the penultimate stage is itself push-pull and the phase inverter precedes this stage. Such a penultimate stage is called a driver.

Inductive phase splitters

The simplest method of phase inversion is to use an inductive device, such as a transformer or center-tapped choke. Fig. 301 shows various methods. In Fig. 301-a a transformer with center-tapped secondary is connected in the conventional manner. This method is simple and inexpensive if high-quality reproduction is not essential, but a transformer for high-grade results is rather difficult to design and expensive to make since a wide frequency response with linearity is necessary.

For good bass reproduction the primary inductance must be adequate and since it carries dc the design must be generous. For good treble response the self-capacitance of the windings must be

low and with the generous design postulated for good bass the requirements are in conflict.

These difficulties can be avoided by using a transformer with a high-permeability core, when the windings can be small and a good frequency response secured. Such a transformer must be parallel-fed, as shown in Fig. 301-b. The secondary winding need not be center tapped, for two equal resistors can be connected across

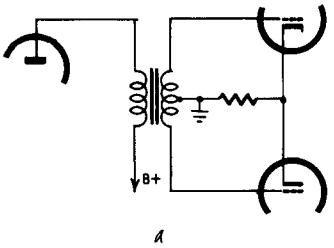


Fig. 301-a. *Inductive phase inverter using center-tapped transformer.*

the winding to give the center tap. Suitable values would be between 100,000 and 250,000 ohms. A resistive center tap could be applied to the transformer in Fig. 301-a.

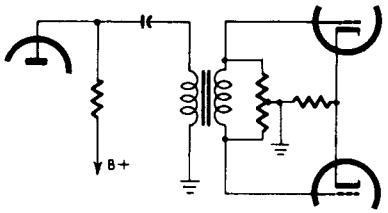
Finally, a center-tapped choke can be used as shown in Fig. 301-c. This method is not very satisfactory owing to the loss of gain as compared with a transformer and the difficulty of getting perfect balance between the upper grid which is coupled directly (through the capacitor) and the lower one, which is coupled through the choke.

Stage gain

The stage gain of Fig. 301-a is the same as though phase inversion were not used and a usual ratio for the transformer would be a stepup of 1 to 2. In Fig. 301-b, where T is the transformation ratio and R the sum of the two resistors across the secondary winding, the load reflected into the previous tube is R/T^2. As the optimum load on the preceding tube is the determining factor for output and distortion, the resistors must be selected to achieve the desired load. In all three methods the output tubes can have common bias or separate bias, as shown in Figs. 301-b and 301-c. A common-bias resistor bypass capacitor is not normally required or desirable. With separate bias, adequate capacitive bypass is necessary to avoid loss of bass.

Vacuum-tube phase splitters

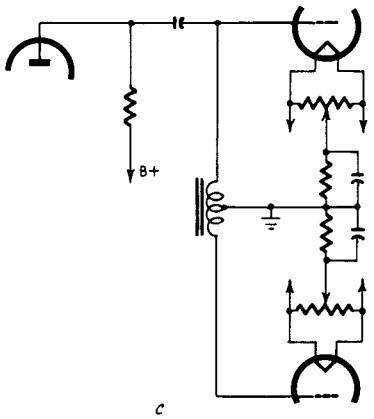
Inductive phase splitters can be avoided by using a vacuum tube for the purpose. In Fig. 302-a the signal is applied to the grid of V1, amplified in the normal manner and passed on to output tube V3. A tap on the grid resistor of V3 is selected so that the signal



b

Fig. 301-b. *Inductive phase inverter using a resistive center-tapped transformer.*

voltage fed to the phase-inverter tube V2 is identical with that fed to V1. If V1 and V2 are matched tubes and their associated R-C networks are identical, V3 and V4 will receive signals equal



c

Fig. 301-c. *Inductive phase inverter using a center-tapped choke.*

in magnitude but opposite in phase. It might be said that V2 takes the place of the transformer of Fig. 301-a. Tubes V1 and V2 can, of course, be twin triodes in one envelope, as can similar pairs in other circuits given.

The phase inverter tube V2 can be eliminated by using the arrangement of Fig. 302-b where the upper output tube is also used

as a phase inverter; its half of the output transformer is shunted by a tapped resistor from the tap of which out-of-phase voltages are fed to the grid of the other output tube. This circuit has the

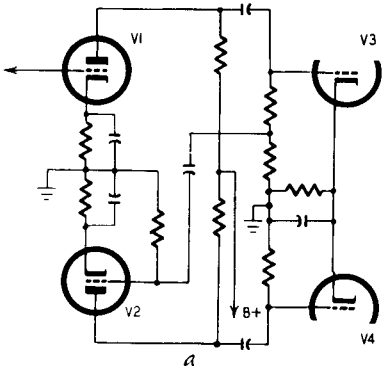


Fig. 302-a. *Vacuum-tube phase inverter with excitation from the grid of the output tube.*

merit of cheapness but it is not recommended when negative feedback is used as complexities of phase shift can occur to cause instability.

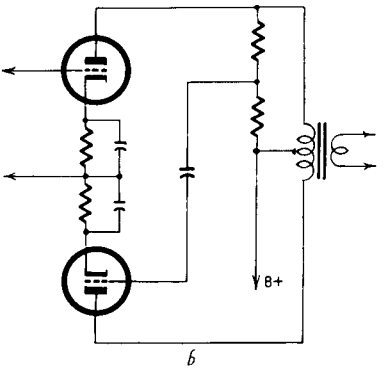


Fig. 302-b. *Vacuum-tube phase inverter with excitation from plate of the output tube.*

These two circuits are not self-balancing and require adjustment for balance when first set up with every tube replacement. Also, because the circuits associated with V2 have two coupling capacitors as compared with one in the plate circuit of V1, they

will be out of balance with the circuit of V1. They are unsuitable for class-AB_2 output stages as the grid–cathode resistance presented to the push-pull tubes is too high.

A driver stage may be needed for these types and the intertube coupling should be a transformer with center-tapped primary and secondary windings or a direct-coupled pair of cathode followers.

Self-balancing inverter

A self-balancing phase inverter is shown in Fig. 303. Resistors R1 and R2 are the ac load on tube V1; R1 and R3 the load on

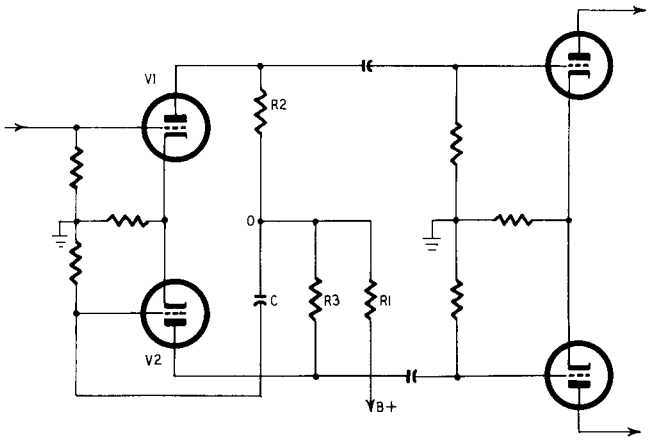


Fig. 303. *Self-balancing phase inverter (floating paraphase).*

V2. R1 is obviously common to both. Unbalance in the plate currents of the two tubes generates a voltage across R1 which is applied to the control grid of V2 through capacitor C. The potential of point O is floating, changing as the plate currents of the two tubes vary with respect to each other; this gives the circuit its name of "floating paraphase."

Schmitt inverter

Fig. 304 shows a different type of self-balancing phase inverter (credited to Schmitt). It is cathode-coupled and exact balance can be obtained by suitable adjustment of the two load resistors. The component values given in the figure have been found to be most suitable and it should be noted that the coupling capacitors in the two grid circuits should have very good insulation resistance. Tube type 6C8-G should be used.

Split-load phase inverter

One of the simplest and most widely used phase splitters appears in Fig. 305. Usually the load on a tube is in the plate circuit; with a cathode follower the load is in the cathode circuit. The circuit of Fig. 305 shows the load divided between plate and cathode circuits, but also included is an unbypassed cathode-bias resistor R2. Now, degeneration (negative feedback) is set up if the bypass capacitor across the bias resistor is omitted, but this only

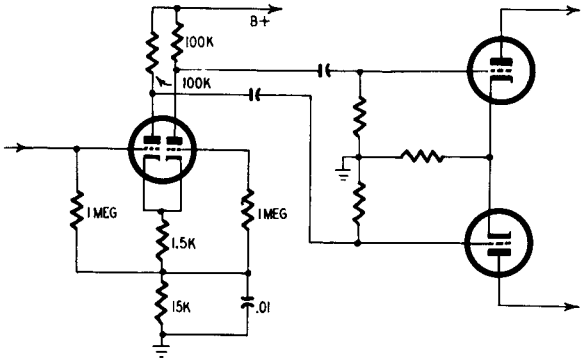


Fig. 304. *Self-balancing phase inverter (Schmitt).*

operates at low frequencies when the electrode–ground capacitances are small in terms of the signal frequencies. At high frequencies these capacitances act as partial bypasses so degeneration is not constant throughout the range of frequencies, with the result that the stage gain is not constant.

Phase splitting occurs when the output from the plate circuit equals that from the cathode circuit: R1 must equal R2 + R3. This does not mean that the resistors must be of very close tolerance for what is required is not accuracy of individual resistors but equality of plate and cathode loads. Similar care in balancing the circuit constants must also be exercised in the case of coupling capacitors C1 and C2 and the grid-circuit resistors R4 and R5 of the following stage. This precaution is necessary not only to secure perfect phase splitting; unbalance can cause instability of the type known as motorboating.

Normally the cathode output is taken directly from the cathode as indicated in Fig. 305 but, particularly with tubes of low amplification factor, it may be better to take the output from the junction of R2 and R3 as shown by the broken line. Exact mathemati-

cal analysis of the alternative connections could be given and prove nothing for in practice the choice can be made only by observing the output of the amplifier on an oscilloscope for a given

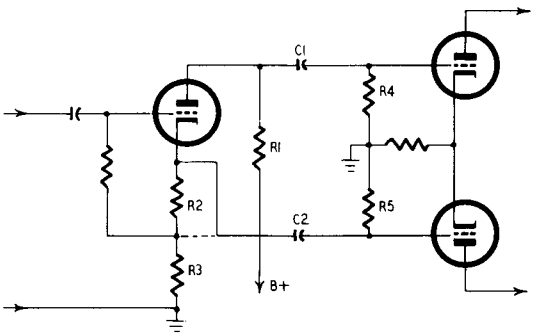


Fig. 305. *Split-load phase inverter.*

signal input. Changing the connection to C2 from the cathode of the phase splitter to the junction of R2 and R3 will then show, for maximum undistorted output, which is better. The writer has

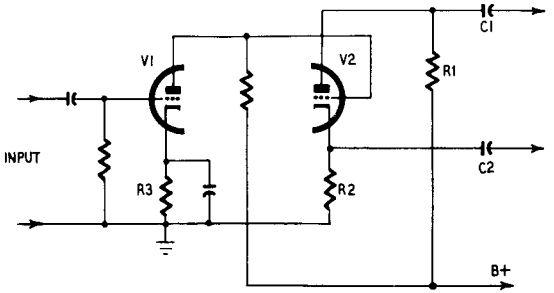


Fig. 306. *Split-load phase inverter with preceding direct-coupled amplifier to give full stage gain.*

found that for tubes such as the 6J5 the best results are obtained when the output is taken from the junction. In either case the stage gain of the phase inverter is somewhat less than 1, usually about 0.8 to 0.9.

Split load inverter with direct-coupled amplifier

This disadvantage of less than unity stage gain can be overcome by a trick which if it involves the use of another tube does not at least involve fitting an extra tube holder in the amplifier chassis. Consider the circuit of Fig. 306 with respect to that of

Fig. 305. Tubes V1 and V2 are the two sections of a double triode, the section V2 representing the inverter tube of Fig. 305. Section V1 is a straight voltage amplifier with a cathode-bias resistor equal to that of R3 in Fig. 305 and a load resistor correct for the tube as an amplifier. This tube is direct-coupled to V2, the phase-inverter tube, and the load of V2 is, as before, made up of R1 and R2.

This device improves the performance of the inverter quite appreciably for the advantages of direct coupling are secured in a greater gain without the use of extra grid-plate coupling capacitors, a matter of importance in reducing phase shift in a feedback amplifier; capacitors C1 and C2 can also be increased to as much as 0.5 μf, a somewhat dangerously high value for the circuit of Fig. 305. It will be appreciated that the high value of cathode resistor makes the split-load inverter particularly suitable for direct coupling. The voltage drop across the cathode resistor is substantial compared with an ordinary amplifier and the cathode is at a comparatively high potential with respect to ground. Tube V2 is therefore in a "natural" condition for having its grid directly coupled to the plate of the preceding tube.

A word of warning is, however, necessary. The correct operation of inverter V2 is dependent on the correct dc potential on its grid. This can only be controlled by the potential on the plate of V1, and in practice it will be found that plate-load and cathode-bias resistors should be somewhat larger than would normally be used for that tube. As with all phase-inverter circuits it is almost essential that symmetrical working be checked by oscilloscope tests.

Cathode-coupled phase inverter

In the paraphase circuit of Fig. 303 resistor R1 is common to both tubes of the inverter. Unbalance in the two plate circuit currents generates a voltage across it which is used to balance the whole circuit. This common load can be used in the cathode circuits of the tubes, creating what is known as the cathode-coupled or long-tail inverter. This is illustrated in Fig. 307. It will be clear that there is a similarity to the split-load inverter because the potential of the cathodes is high with respect to ground. At the same time differences in the plate currents of V2 and V3 set up a voltage across R4, which is applied to the grid of V2 through capacitor C1. A suitable grid voltage for V2 is of the order of 85 volts, bearing in mind the comparatively

high positive potential of the cathodes, and thus suggests that, as in the case of the direct-coupled split-load inverter, the input is direct-coupled to the preceding stage V1, and is so shown in Fig. 307. This inverter has good frequency characteristics but only half the gain of the circuit of Fig. 306. It is a popular circuit in Europe—the well-known Mullard 20-watt amplifier using it —and more recently has come into favor in the United States,

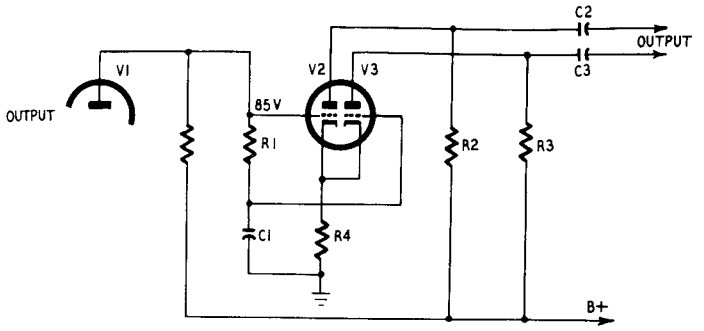


Fig. 307. *Cathode-coupled or "long-tailed" phase inverter showing direct coupling to preceding stage.*

where examples are to be found in the Eico 50- and 60-watt amplifiers, among others.

There is some difference of opinion as to the optimum values of associated components, which in any event depend on the type of double triode used for V2-V3. The following data compares Mullard with Eico practice.

	Tube	R1	R2	R3	R4	C1	C2	C3
Mullard:	ECC83	1 Meg	180K	180K	82K	0.25	0.5	0.5
Eico:	6SN7-GTB	1 Meg	28,750	33K	18K	0.25	0.25	0.25

Drivers

The stage before the output stage is generally called the driver, although the term is more often restricted to the rather special driving conditions in class-AB_2 and class-B amplifiers. In these, a considerably greater signal voltage is required than with class-A or AB_1 output stages. As the tubes of a class-AB_2 output stage are driven into grid current, a low resistance for the grid circuits is essential and an interstage transformer is generally used.

Transformer coupling

In transformer-coupled amplifiers it is customary to use general-purpose triodes having a plate resistance of about 7,500 ohms, such as the familiar 6J5. The output voltage of such a tube can be obtained from its characteristic curves in a manner similar to that for output tubes. The higher the plate voltage the higher the output ac voltage before distortion commences. If one tube is insufficient to drive the output stage to full output, then two should be used in push-pull (the use of a large single power tube would call for a costly design of the interstage transformer). The transformer would then have both primary and secondary windings center tapped. This is actually an advantage since the opposing plate currents in the two halves of the primary winding call for a smaller core than with one tube, and in terms of frequency response, a better performance is obtained from the transformer.

Overall gain

The overall gain from a transformer-coupled triode is about equal to the product of the amplification factor of the tube and the stepup ratio of the transformer. The 6J5 has an amplification factor of 20, and with a plate voltage of 250 and a grid voltage of −8, the plate current is 9 ma. One tube would be quite unable to drive two triodes in class AB_2 and even 9 ma plate current is rather high polarizing for an interstage transformer. In class AB_2, moreover, the driver has to supply power and transformer losses can be appreciable. Two tubes are necessary and, being in push-pull, they make the task of the transformer a simpler one to fulfill. Detailed design of class-B amplifiers is rather complicated and space here does not permit full treatment. The interested reader is referred to the standard textbooks on the subject. In brief:

Class-AB_2 amplifiers operate similarly to class AB_1 and the driving voltage required is less than with class B.

If a transformerless phase splitter is used, remember that inadequate care in design may introduce distortion and instability, mainly through out-of-balance components.

Parasitic oscillation can occur in the grid circuits of the output tubes which can largely be avoided by using a driver transformer having low leakage inductance.

Resistance-coupled driver stages are fairly widely used, particularly when the output stage is two push-pull triodes, or tetrodes or pentodes connected as triodes. This is necessary since

the grid-to-grid swing required for maximum output from two triodes is considerably greater than that for tetrodes or pentodes.

Push-pull drivers

Sometimes a push-pull driver stage is used to swing a pair of tetrodes, apparently on the supposition that there will be less distortion than when using a single tube. This notion is carried to its logical conclusion in designs which show an amplifier en-

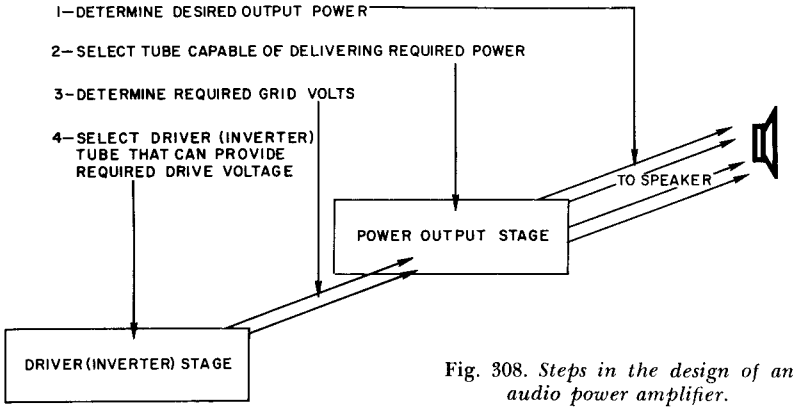


Fig. 308. *Steps in the design of an audio power amplifier.*

tirely push-pull. In theory, perhaps, it could be shown that a wholly push-pull amplifier has less overall distortion than one in which the voltage amplifier is not push-pull, but in practice, if the voltage-amplifying stages are properly designed, the elaboration is unnecessary. It is also undeniable that the simpler an amplifier the less the possibility of distortion creeping in through tube or component deterioration.

Design sequence

Yet the penultimate or driver stage must have sufficient output ac voltage to swing the push-pull output stage and, if the phase inverter is unable to deliver that voltage without distortion, then the driver stage must be push-pull. A particular type of tube can deliver only so many output volts without distortion, whether it is used as a voltage amplifier or phase inverter. This fact is sometimes not appreciated, since technical writers are often asked for an opinion as to the best type of inverter without any mention being made of the output stage of the proposed amplifier. Hence the sequential treatment in this book. The output stage must first be selected. Given this, the required input grid volts are

known. A tube is then selected which will give that voltage output. If, choosing from among the tubes that are normally used as phase inverters, it is found that the desired voltage cannot be obtained, the driver stage must be push-pull. Then, and only then, can the type of phase inverter be selected. (See Fig. 308.)

This is a problem in voltage-amplifier design, which will be discussed in the next chapter. The matter is mentioned at this stage simply because the driver is a voltage amplifier and properly belongs to the next chapter, but may come under the scope of the present chapter because it may be simply a phase inverter if the output stage does not require a large input grid voltage. Suitable tubes for use as phase inverters are listed in Table 10 at the end of Chapter 4.

chapter 4

audio amplifiers: voltage amplification

POWER is required to drive a speaker or a recording head. The power is obtained from the power-output stage, and to get this power, a voltage of the correct magnitude has to be impressed on the grid circuit of the output stage. As the actuating signal device (phono pickup, signal voltage across the detector diode load, etc.) has a voltage output considerably less than required, amplification of the signal voltage is necessary. The voltage amplifier performs this task. It cannot be a *pure* voltage amplifier since the load into which it works does not have infinite impedance, so there will be some power in the output. The output load may be as high as 1 megohm, but is frequently less, and in the case of transformer coupling may be unknown. The number of stages required in a voltage amplifier depends on the degree of amplification required, postulated by the input and desired output voltages and the type of tubes selected. The fewer the tubes, the less likelihood of instability through a common plate supply impedance, but a selection of tubes giving a high stage gain may result in undesirable performance. The design of the amplifier is wholly dependent on the degree of absence of distortion required.

Transformer-coupled stages have already been mentioned. Almost invariably resistance—capacitance-coupled tubes are used in voltage amplifiers, so further discussion on transformer-coupled stages will not be given. Either triodes or pentodes may be used and an examination of their characteristics can be carried out in a manner similar to those of output tubes.

Resistance—capacitance-coupled triodes

It is customary to use either general-purpose (medium-mu with an amplification factor under 50) or high-mu triodes (amplification factor 50 to 100). The tube manufacturers publish resistance—capacitance design charts for all types of tubes, but a word of warning is needed in connection with these. The figures given seem to indicate that a low source impedance is used whereas in practice the source impedance and grid—cathode impedance are usually fairly high (Fig. 401); therefore, the values of associated

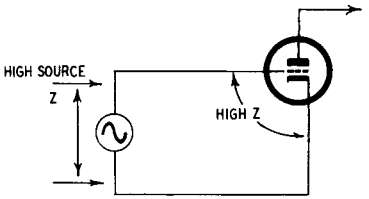


Fig. 401. *Source and grid-to-cathode impedances are generally high.*

components are not necessarily correct. These charts also refer to average tubes and so can be taken only as a rather rough guide in specific cases.

The figures given also assume a harmonic distortion content of 5% and for high-grade equipment this figure is excessive. In view of the fact that distortion occurs when the tube is called upon to deliver its peak voltage, the obvious way to avoid it is to work the tubes well within their limits. Accordingly, for high-grade amplifiers the designer should not expect to get the voltage output from any particular tube that the chart would indicate as being possible.

This does not mean that the overall gain of the amplifier must be determined so that no tube has to produce more than about three-quarters of the possible output voltage. As this limitation usually arises only in the last stage of the voltage amplifier, earlier stages, including so-called preamplifier stages, may be selected on the basis of the charts. If the last tube is unable to deliver the necessary voltage to drive the output stage or driver stage, then the last stage of the voltage amplifier would have to be a push-pull stage and the phase-splitter would be a part of the voltage amplifier.

Design data

For convenience in design the mass of data presented by the tube makers has been edited and put together in different form.

After eliminating some unnecessary information, what is left produces a recognizable pattern. Table 8, giving the correct values for average triodes of the types listed, is subject to the limitations mentioned. As most designers are interested in equipment deriving power from a line-operated supply, battery operating types of tubes have not been included but the relevant information can be extracted from tube-makers' manuals.

The input resistance should not exceed 1 megohm and subsequent grid resistors should be not more than four times the pre-

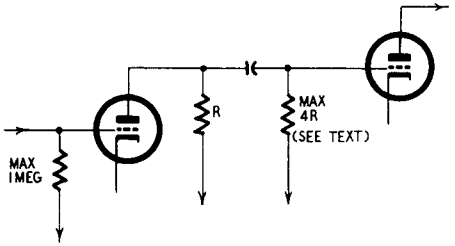


Fig. 402. *Relative values of plate and grid resistors.*

ceding plate resistor (Fig. 402). With high-mu triodes the grid resistor should be less than this. High values of grid resistor tend to produce hum. It should be remembered that output tubes require a lower grid resistor than an ordinary voltage amplifier (the recommended values being given in the tube handbooks) but a simple rule is to keep this as low as 250,000 ohms for high-grade equipment. With this in mind an examination of Table 8 shows that quite a number of alternative usages are ruled out on the grounds that the value of R3 will be too high. If R3 is 0.5 megohm, several opportunities occur of getting a higher output voltage from the preceding tube, but it should be remembered that the 0.22-megohm value is recommended to reduce the effects of reverse grid current in the output stage.

Coupling capacitors

The value of the coupling capacitor C2 is important. If too small, there will be a loss of amplification at low frequencies. The values shown allow for the amplification at 50 cps to be reduced to 0.8 times the amplification at higher frequencies—a loss of 2 db. If this capacitor is reduced by one-half, the 2 db loss will be at 100 cps; if doubled, the 2 db figure is reached at 25 cps, and so on.

If the capacitor is too large and the whole amplifier has a tendency to motorboat, the motorboating frequency will be passed on without attenuation. So a simple way to stop motorboating is to reduce the value of the coupling capacitors. But this must not be done to such an extent that appreciable loss of amplification occurs at low frequencies; the *cause* of the instability must be sought out.

For inexpensive equipment, C2 can be half the values shown; for high-fidelity amplifiers it should be double. Any further increase will not contribute audibly to quality of reproduction but can permit the transfer of subsonic frequencies.

Cl is most conveniently (and least expensively) an electrolytic capacitor. One rated at 10 μf with a working voltage of 25 or 50 will be suitable for all cases.

The supply voltage should be as high as possible, with a maximum of about 300. Examination of the figures for output voltage show clearly that about double the voltage swing can be had by raising the supply voltage from 180 to 300, the gain not being appreciably affected. The greatest output voltage is obtained when the plate load R1 is highest but, as this requires a high following grid resistor, it may not be possible to use this condition with high-fidelity output stages.

A push-pull output stage requires double the input signal voltage needed to drive a single tube so, in selecting a suitable tube from Table 8, the required grid-to-grid voltage must be remembered. If a single tube will not supply this voltage (and for minimum distortion the designer should expect to get only three-quarters of the output voltage E_0 shown in the table) except by using a high plate load, then push-pull working must be used. In general an output stage of two triodes (or two tetrodes connected as triodes) in push-pull will require a push-pull driver stage; two tetrodes can be driven from a single tube.

Resistance—capacitance-coupled pentodes

Table 9 shows suitable operating conditions for resistance—capacitance-coupled pentodes. In a general way considerations are similar to those for triodes, but an additional factor must be taken into consideration—high-frequency attenuation. As will be seen from the data given at the head of the table, the value of the plate load has a bearing on the width of frequency response of the amplifier and R1 must be selected with the performance required in mind.

For a rather low plate load the stage gain is not appreciably greater than that obtainable from a high-mu triode and the ques-

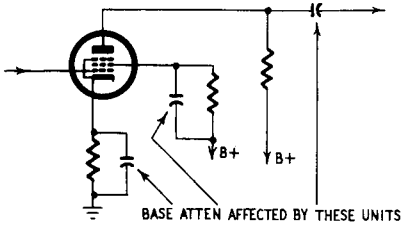
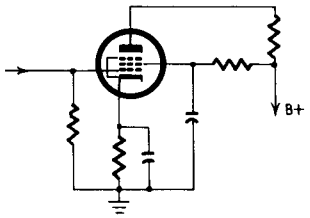


Fig. 403. *Amplification of bass is affected by the size of the coupling, cathode and screen bypass capacitors.*

tion may be asked, why use pentodes at all? This involves consideration of the dynamic characteristics of both types and, con-

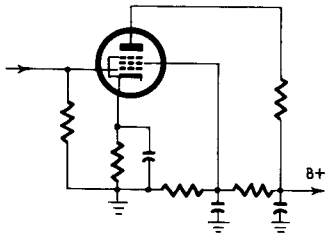


a

Fig. 404-a. *Usual method of applying screen voltage.*

trary to generally held opinion, a pentode as a voltage amplifier can give a good deal less distortion than a triode.

Examination of a tube such as the 6SJ7 used as a triode or as a



b

Fig. 404-b. *Voltage divider for the screen is a better technique.*

pentode will show that, for outputs up to about 10 volts rms, the intermodulation distortion of the pentode is about one-eighth that of the triode, assuming that the operating conditions of both are optimum. At 30-volts output the distortion is about the same; at 60 volts output the pentode has more distortion than the triode. So a pentode is most useful in the earlier stages of an amplifier where the signal voltage to be handled is low. In these positions the pentode has a higher amplification factor and less distortion than the triode and is strongly recommended for high-fidelity amplifiers.

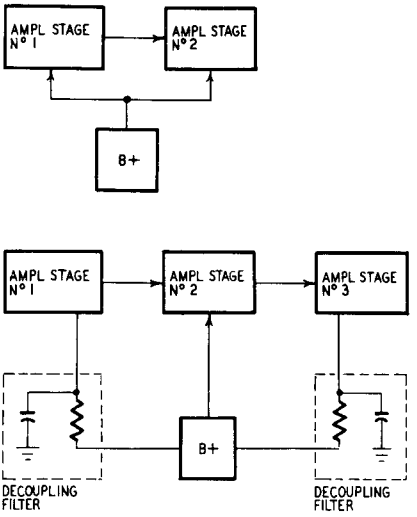


Fig. 405. *Block diagram of decoupling circuits. When two or more amplifier stages are used power supply decoupling is necessary.*

Bass attenuation is affected by the screen bypass as well as the cathode bypass and coupling capacitors. If, instead of a series screen resistance as shown in Fig. 404-a (the usual method of feeding the screen) a potential divider is used as in Fig. 404-b, the loss of bass through inadequate bypassing is very much less. This has the further advantage of stabilizing the screen potential to some extent, a desirable aim with all pentodes.

The cathode bypass will be seen to be greater than that required for triodes, and it is most convenient to use an electrolytic capacitor of 50 μf with a working voltage of 12, or over.

Resistance—capacitance-coupled phase inverters

Table 10 summarizes in convenient form the conditions that have to be met when using twin triodes as amplifying phase inverters. No-gain splitters have already been dealt with in Chapter 3, but the table gives the values required for a circuit of the type shown in Fig. 302-a. For the twin triodes included in Table 10, the values shown in that table should be used for R1, R2, R3 and C but as the output from the two triode sections of the double triode must be equal, the tap on $R3_a$ and $R3_b$ must be adjusted to suit.

The twin triode is an amplifier hence the tap on the grid resistor of one output tube must be selected so that the voltage applied to the second triode of the phase inverter is such that the voltage across the grid resistor of the second output tube is identical. If the voltage gain of the stage is, say 20, then the input sig-

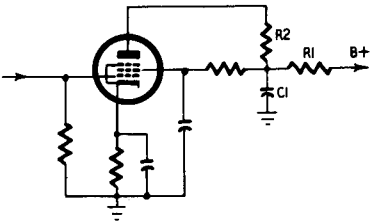


Fig. 406. *RI and CI form a decoupling network.*

nal voltage E_i would be one-twentieth of the output voltage, so the voltage tapped off R3 would be one-twentieth of the total voltage across the two.

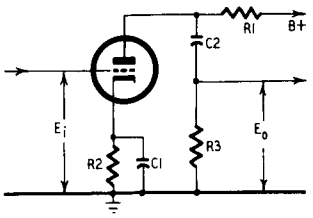
Decoupling

In the discussion on triodes it was stated that the supply voltage should be as high as possible, with a maximum of 300, but this assumes that the supply voltage is applied directly to the load resistor. If an amplifier has only two stages, decoupling is probably unnecessary; but more than two stages are needed in all but the cheapest equipment. Decoupling then becomes essential to overcome common impedance coupling in the power supply. See Fig. 405.

Decoupling involves the use of a series plate resistance R1 with a low impedance bypass to ground, C1, between the supply source and the load resistance R2 of the tube as shown in Fig. 406. As the tube takes plate current, a very appreciable voltage drop takes

place across the decoupling resistance. The higher the value of the decoupling resistance, the better the decoupling (assuming an adequate bypass to ground). For a supply voltage of 300 applied to the plate resistors, it can be taken as a safe working rule that the actual supply voltage should be 400, or a little more, and the bypass capacitor has a capacitance of 8 μf with a dc working voltage of 500.

Table 8—Resistance—capacitance-coupled triodes



C1 and C2 adjusted to give 0.8 E_o at 50 cps.

Data for double triodes (types 6SN7-GTB, 12SN7-GT, 6SL7, and 12AX7) apply to each triode section.

R1 and R3 in megohms; R2 in ohms.

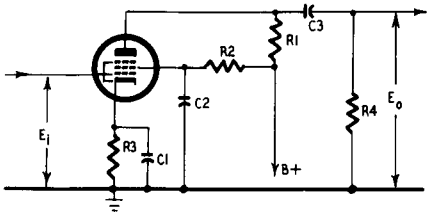
V.G. (voltage gain) = E_o/E_i.

					B+ = 180 volts			B+ = 300 volts		
Type	R1	R3	C1	C2	R2	E_o	V.G.	R2	E_o	V.G.
6BF6		.047	6	.15	2000	32	10	1800	58	10
6SR7	.047	.1	5	.07	2500	42	10	2400	74	11
6ST7		.22	4	.04	3000	47	11	2900	85	11
12SR7		.1	4	.07	3800	36	11	3600	65	12
26C6	.1	.22	3	.03	5100	47	11	5000	85	12
		.47	2	.02	6200	55	12	6200	96	12
		.22	2	.03	8000	41	12	7800	74	12
	.22	.47	1	.02	11000	54	12	11000	95	12
		1.0	1	.01	13000	69	12	13000	106	12
6C4		.047	8	.15	920	20	11	870	38	12
12AU7	.047	.1	6	.1	1200	26	12	1200	52	12
		.22	5	.03	1400	29	12	1500	68	12
		.1	4	.07	2000	24	12	1900	44	12
	.1	.22	3	.03	2800	33	12	3000	68	12
		.47	2	.015	3600	40	12	4000	80	12
		.22	2	.03	5300	31	12	5300	57	12
	.22	.47	1	.015	8300	44	12	8800	82	12
		1.0	1	.007	10000	54	12	11000	92	12
6J5		.047	8	.15	1190	24	13	1020	41	13
6SN7-GTB	.047	.1	6	.07	1490	30	13	1270	51	14
		.22	5	.03	1740	36	13	1500	60	14
12J5		.1	5	.08	2330	26	14	1900	43	14
12SN7-GT	.1	.22	4	.03	2830	34	14	2440	56	14
		.47	3	.02	3230	38	14	2700	64	14
		.22	3	.03	5560	28	14	4590	46	14
	.22	.47	2	.02	7000	36	14	5770	57	14
		1.0	1	.01	8110	40	14	6950	64	14

Table 8 (continued)

Type	R1	R3	C1	C2	B+ = 180 volts R2	E_o	V.G.	B+ = 300 volts R2	E_o	V.G.
6SQ7-GT	.047	.1	8	.06	2600	16	29	1900	31	31
		22	7	.03	2900	22	36	2200	41	39
		.47	6	.02	3000	23	37	2300	45	42
12SQ7-GT	.1	.22	6	.03	4300	21	43	3300	42	48
		.47	5	.02	4800	28	50	3900	51	53
		1.0	4	.01	5300	33	53	4200	60	56
	.22	.47	4	.02	7000	25	52	5300	47	58
		1.0	3	.01	8000	33	57	6100	62	60
		2.2	3	.005	8800	38	58	7000	67	63
6AT6	.1	.1	9	.06	1900	19	30	1500	40	34
6SL7-GT		.22	7	.03	2200	25	35	1800	54	38
		.47	6	.02	2500	32	37	2100	63	41
6SZ7	.22	.22	5	.03	3400	24	38	2600	51	42
6T8		.47	4	.02	4100	34	42	3200	65	46
12AT6		1.0	4	.01	4600	38	44	3700	77	48
12SL7-GT	.47	.47	3	.02	6600	29	44	5200	61	48
		1.0	2	.01	8100	38	46	6300	74	50
		2.2	2	.005	9100	43	47	7200	85	51
6AV6	.1	.1	10	.06	1800	18	40	1300	43	45
12AV6		.22	9	.03	2000	25	47	1500	57	52
12AX7		.47	8	.02	2200	32	52	1700	66	57
	.22	.22	6	.03	3000	24	53	2200	54	59
		.47	5	.02	3500	34	59	2800	69	65
		1.0	4	.01	3900	39	63	3100	79	68
	.47	.47	4	.02	5800	30	62	4300	62	69
		1.0	3	.01	6700	39	66	5200	77	73
		2.2	2	.005	7400	45	68	5900	92	75

Table 9—Resistance—capacitance-coupled pentodes



C1, C2, C3 adjusted to give 0.7 E_o at 50 cps.

Upper limits of level frequency response for values of R1:

R1 = 0.1 megohm—25,000 cps.
R1 = 0.25 megohm—10,000 cps.
R1 = .5 megohm— 5,000 cps.

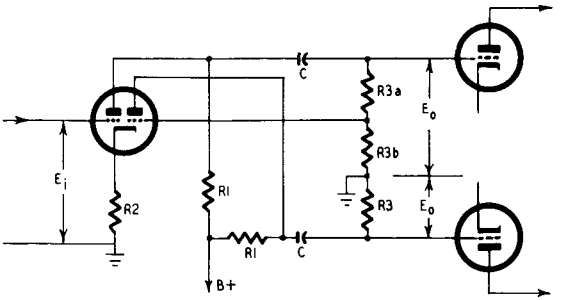
R1, R2, R4 in megohms. R3 in ohms. V.G. (voltage or stage gain) = E_o/E_i.

Table 9 (continued)

Type	R1	R2	R4	C1	C2	C3	B+ = 180 volts R3	B+ = 180 volts E_o	B+ = 180 volts V.G.	B+ = 300 volts R3	B+ = 300 volts E_o	B+ = 300 volts V.G.
6AU6		.12	.1	36	.3	.05	800	57	74	500	76	109
6SH7	.1	.15	.22	33	.25	.03	900	72	116	600	103	145
12AU6		.19	.47	30	.25	.02	1000	81	141	700	129	168
12SH7		.38	.22	25	.2	.02	1500	59	130	1000	92	164
	.22	.43	.47	24	.2	.02	1700	67	171	1000	108	230
		.6	1.0	22	.2	.01	1900	71	200	1100	122	262
		.9	.47	16	.15	.01	3100	54	172	1800	94	248
	.47	1.0	1.0	16	.15	.01	3400	65	232	1900	105	318
		1.1	2.2	15	.15	.005	3600	74	272	2100	122	371
6SF7		.33	.1	16	.4	.05	1000	32	33	750	62	39
12SF7	.1	.5	.22	16	.4	.03	1200	37	45	850	80	46
		.6	.47	16	.4	.02	1300	43	52	900	93	57
		.76	.22	12	.3	.02	1700	37	47	1150	63	62
	.22	.9	.47	12	.25	.015	1700	44	68	1300	78	88
		1.0	1.0	10	.25	.015	1800	47	82	1500	99	97
		1.8	.47	7	.2	.01	3300	38	70	2300	71	82
	.47	2.0	1.0	7	.2	.01	3800	50	85	2500	85	109
		2.2	2.2	7	.2	.005	4000	57	98	2800	105	125
6SJ7		.29	.1	24	.2	.04	760	49	55	500	72	67
12SJ7	.1	.31	.22	22	.2	.04	800	60	82	530	96	98
		.37	.47	20	.2	.02	860	62	91	590	101	104
		.83	.22	17	.15	.02	1050	38	109	850	79	139
	.22	.94	.47	15	.15	.01	1060	47	131	860	88	167
		.94	1.0	15	.15	.01	1100	54	161	910	98	185
		1.85	.47	12	.15	.01	2000	37	151	1300	64	200
	.47	2.2	1.0	12	.1	.005	2180	44	192	1410	79	238
		2.4	2.2	12	.1	.005	2410	54	208	1530	89	263

The 6SH7 and 12SH7 are particularly susceptible to hum when used as high-gain audio amplifiers. When the tubes are used as low level drivers this undesirable characteristic need not be considered. However the tubes should be isolated from external hum sources whenever possible. Two separate cathode connections permit the input and output circuits to be isolated from each other.

Table 10—Resistance—capacitance-coupled phase inverters



This table includes double triodes normally used as voltage amplifiers. Other tubes can also be used; e.g. the double triodes listed in Table 8, where the operating conditions listed in that table should be used.

C is selected to give 0.9 E_o at 50 cps.

$R3_a+R3_b=R3$. $R3_b$ is selected so that the voltage output of each triode of the double triode is equal; e.g. if V.G. = 20 then $R3_b/R3 = 1/20$.

R1, R3 in megohms. R2 in ohms. V.G. (voltage or stage gain) E_o/E_i.

				B+ = 180 volts			B+ = 300 volts		
Type	**R1**	**R3**	**C**	**R2**	**E_o**	**V.G.**	**R2**	**E_o**	**V.G.**
6SN7-GTB	.047	.047	.15	2000	32	10	1800	58	10
		.1	.07	2500	42	10	2400	74	11
		.22	.04	3000	47	10	2900	85	11
	.1	.1	.07	3800	36	11	3600	65	12
		.22	.03	2830	34	14	2440	56	14
		.47	.02	3230	38	14	2700	64	14
	.22	.22	.03	5560	28	14	4590	46	14
		.47	.02	7000	36	14	5770	57	14
		1.0	.01	8110	40	14	6950	64	14
ECC82 12AU7	.047	.047	.15	920	20	11	870	38	12
		.1	.1	1200	26	12	1200	52	12
		.22	.03	1400	29	12	1500	68	12
	.1	.1	.07	2000	24	12	1900	44	12
		.22	.03	2800	33	12	3000	68	12
		.47	.015	3600	40	12	4000	80	12
	.22	.22	.03	5300	31	12	5300	57	12
		.47	.015	8300	44	12	8800	82	12
		1.0	.007	10000	54	12	11000	92	12

Table 10 (continued)

Type	R1	R3	C	B+ = 180 volts R2	E_o	V.G.	B+ = 300 volts R2	E_o	V.G.
ECC83		.1	.06	1800	18	40	1300	43	45
12AX7	.1	.22	.03	2000	25	47	1500	57	52
		.47	.02	2200	32	52	1700	66	57
		.22	.03	3000	24	53	2200	54	59
	.22	.47	.02	3500	34	59	2800	69	65
		1.0	.01	3900	39	63	3100	79	68
		.47	.02	5800	30	62	4300	62	69
	.47	1.0	.01	6700	39	66	5200	77	73
		2.2	.005	7400	45	68	5900	92	75
6SC7		.1	.07	960	17	25	750	35	29
12SC7	.1	.22	.03	1070	24	29	930	50	34
		.47	.02	1220	27	33	1040	54	36
		.22	.03	1850	21	35	1400	45	39
	.22	.47	.02	2150	28	39	1680	55	42
		1.0	.01	2400	32	41	1840	64	45
		.47	.02	3050	24	40	2330	50	45
	.47	1.0	.01	3420	32	43	2980	62	48
		2.2	.005	3890	36	45	3280	72	49

chapter 5

amplifier design

WITH the aid of Tables 4 to 10 it is now possible to work out the approximate design of a whole amplifier. As a working example, let it be supposed that a circuit is required for an amplifier to give 20 watts of audio power for a signal input of 1 volt, using a minimum number of tubes, yet avoiding excessive distortion. An examination of Table 6 shows that most triodes will not meet the power-output requirement, but Table 7 shows that two 6L6's as tetrodes in class-AB_1 will.

From Table 7 it is seen that a peak grid-to-grid input of 57 volts is required to give a power of 24 watts. A phase splitter is required, and if the cathode resistor type of Fig. 304 is used, with a gain of less than unity, the input voltage to the phase splitter must be about 70 volts. Now examine Tables 8 and 9. The grid circuit resistance for 6L6's should be less than 0.5 megohm, so any line of data showing a higher resistance than this for R3 in Table 8 or R4 in Table 9 must be rejected.

Selecting a phase inverter

For 5% distortion in the intermediate amplifying stage the figures for E_o can be accepted, but this distortion is necessarily added to that of the output stage, so it is desirable to assume that E_o can be only 75% of the figures shown. The condition would seem to be met by a type 6BF6 operating at a plate supply voltage of 300 with a load resistor of 0.22 megohm. But the voltage gain is only 12 so a further amplifying stage would be required

to meet the specification for input voltage. A high-mu triode could not be used here, because inspection will show that the required grid impedance for 6L6's cannot be met.

In Table 9 the answer would appear to be found in using a pentode, say a 6SF7, which, with a plate load of 0.22 megohm and a following grid resistor of 0.47 megohm will deliver 71 volts with a stage gain of 82. It would be a mistake to use a 6SJ7 which, with 0.1-megohm plate resistance will give an E_0 of 96 volts with a stage gain of 98 because it is thought that the 71 volts from the 6SF7 should really be only 54 volts on the score of freedom from distortion. It has already been pointed out that a tube such as the 6SJ7 (or the 6SF7 for that matter) will give more distortion than a triode for 60 volts output. Moreover the stage gain of 98 is too high, for if an input of 1 volt is applied to the tube the output will be 98 volts, which is too high for the output stage. It might be argued that the solution is to use a volume control in the grid circuit of the 6SJ7, but the amplifier should be designed to give maximum power with the volume control at maximum, otherwise part of the control is unusable. And a last point is that using a first stage with a no-gain phase splitter involves the use of two tube sockets and two separate tubes. Can this be reduced?

Examination of Table 10 shows that it can. The grid-to-grid input of the output stage is 57 volts; that is, the input for each tube with respect to ground is 28.5 volts. Table 10 shows that the twin triode 6SC7 with a plate resistance of 0.1 megohm will give an E_0 of 50 per triode with a stage gain of 34, the following grid resistance being 0.22 megohm. Since 75% of 50 is 37 volts, each triode is well able to deliver 28 volts to each output tube. As the stage gain is 34 and the required output voltage is 28.5, the input required for maximum power is thus 0.8 volt, a figure sufficiently close to the specification to be acceptable. Referring to the circuit diagram at the head of Table 10, the tap on R3 should be at 220,000/34 (= 6,470) ohms from ground.

Table 7 shows that the output stage needs a plate voltage of 360 for the required performance. But the cathode bias resistor applies a bias voltage to the grids of −22.5, so the total supply voltage required is 382.5. Allowing a margin for the potential drop in the output transformer primary winding the smoothed supply voltage should be 400. This is 100 volts greater than the voltage required for the twin triode and the excess voltage can be profitably disposed of by using decoupling resistors. These should be of a value to create a potential drop of 100 volts with the measured plate cur-

rents of the two triodes. The complete circuit diagram can now be drawn, and is shown in Fig. 501.

While this design is an approximation it is still a necessary first

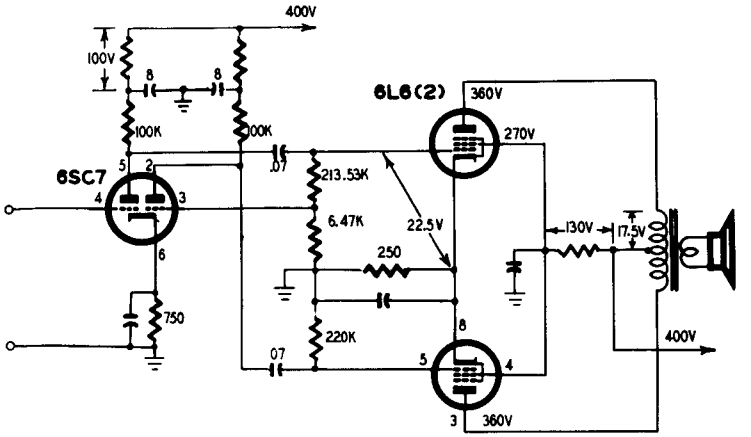


Fig. 501. *Approximate amplifier design.*

attempt at determining the correct tube sequence. Many other factors have to be taken into consideration: no individual tube will perform exactly as stated in the handbooks nor will the values of resistors and capacitors be what they are stated to be unless close-tolerance components are selected. This does not mean that tube and component manufacturers are in any way dishonest. Close-tolerance tubes are not a "commercial proposition" but close-tolerance resistors and capacitors are selected from bulk manufacture by the method of preferred values; as they are the outcome of several selective processes they are necessarily expensive. This introduces a new aspect of the design problem—cost of production for a given set of requirements. Furthermore, the design may lead to unusual values for the resistors and capacitors. For commercial production, a compromise may be made with the use of stock components as near in value to the design as possible.

Checking performance

The approximate design technique just described is based on the assumption that the tubes and components are what they are said to be. The data in Tables 8 and 10 has been derived mathematically, based on the use of preferred values of components of close tolerance, and "average" tubes, that is, tubes built to a pre-

determined specification. When the circuit has been built up into a prototype amplifier the most satisfactory way of proving it is to measure its performance. First check the applied dc voltages with a very high resistance meter. The B+ voltage is to be checked

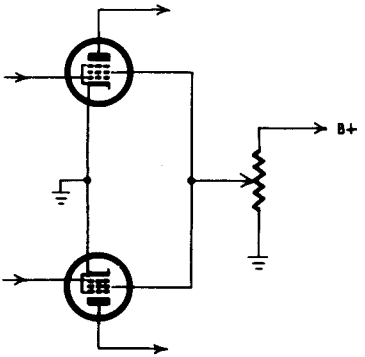


Fig. 502. *Method for determining the value of resistance required to obtain the desired screen grid voltage.*

for 400 volts approximately, but the actual voltage on the output plates with respect to ground should be 360.

Table 7 shows that the screen voltage should be only 270. A series-feed resistance could be used or a potentiometer connected between B+ and ground, the tapping being taken to the screens of the tubes. Knowing the screen current, the tapping point can be calculated but the exact point can more easily be determined by connecting the voltmeter to the screens and altering the potentiometer until the desired voltage is obtained. If a fairly large bleed current is allowed to pass through the potentiometer (by making the total resistance not too high) better regulation of the screen voltage will be obtained, since the screen current variations will be a rather small fraction of the total current passing through the potentiometer. See Fig. 502. This screen feed is not shown in Fig. 501, but is easily understood.

Next, the voltage on the supply end of the plate resistors of the double triode should be set to 300 by selection of suitable decoupling resistors. Using a vacuum-tube voltmeter adjusted to read ac volts, check that with the input of the first tube short-circuited there is no signal voltage reading at the plate of the first tube, at the grids of the output tubes, or at the plates of the output tubes. If the meter shows a reading, then hum or some

other disturbance is actuating the amplifier; this must be traced and eliminated. In the absence of a reading, the input circuit can be opened and an audio oscillator used to inject a signal into the

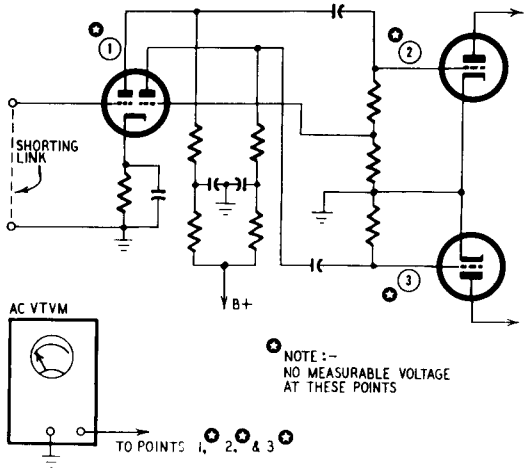


Fig. 503. *Checking the preliminary design.*

amplifier. It was calculated that an input of 0.8 volt was needed for maximum output power, so set the oscillator for an output of 0.8 volt at a frequency of 1,000 cps. The output must be sinusoidal.

Checking stage gain

Applying the vtvm to the plate of the triode directly connected to the input will show if the tube is giving the correct stage gain for the output voltage divided by the input voltage should equal 34. But the reading obtained cannot be accepted without checking the waveform. Harmonic distortion may be present which will give an incorrect reading; (the instrument is calibrated on a sine wave). Freedom from distortion will be shown by shunting the vtvm with an oscilloscope. If the trace is a pure sine wave, no distortion is present and the voltage reading is reliable. The second triode must now be checked and it is of little consequence if the amplification factors of the two triodes are slightly unequal; what is necessary is that the voltages applied to the two output tubes be equal. The vtvm is therefore transferred to the second plate (and also the scope) and, if the output voltage is not identical with that from the first triode section, then the 6,470-ohm resistor must be altered until it is. This voltage will also be 180° out of phase with respect to the other because the capacitor values have negligible

reactance at 1,000 cycles, compared to the resistance values. See Fig. 503.

In the absence of capacitive reactances it would be, for each amplifying tube causes a 180° phase change; the signal from the first triode plate is what might be called the direct signal voltage, but as a part of this voltage is applied to the grid of the second triode section, the output from the plate will be 180° out of phase with respect to the grid input. But the voltage applied to the second grid has passed through C1 which itself causes a phase change, so the output of the second triode is not exactly 180° out of phase with that of the first. If C2 were not present, the phase change caused by C3 would put matters right, but C2 *is* necessary otherwise the grid-cathode resistance for the second triode would be reduced to 6,470 ohms. R_g for both triodes must be equal to ensure equal amplification, and should have a value of about 0.5 megohm. Whatever value is selected for C2, phase balance must be obtained by adjusting the value of C3. This is most easily checked by connecting the oscilloscope across the output grid circuits.

The output stage can now be checked for waveform and power generated. This is most simply done by using a purely resistive load. Table 7 specifies a plate-to-plate load of 9,000 ohms, which is formed by using two 4,500-ohm noninductive resistors in series, with the junction point taken to B+. The rms current through the load is then measured. The power output in watts = I^2R, where, in this case R is 9,000 ohms, I being in amperes. See Fig. 504.

The results of the work described here are not in themselves of very much practical value for the amplifier as it stands is not of very much use. The output load is purely resistive whereas in practice the load is resistive, inductive and capacitive and varies with frequency. An input sensitivity of 0.8 volt is useless with a microphone, tape-recording head or high-fidelity pickup. No provision whatever has been made for a tone or even a volume control. Why then should space be apparently wasted in describing something of no practical worth?

The answer is that too frequently designers and circuit hounds produce the whole outfit complete with gimmicks as a first step and then try to find out why it doesn't work. This can turn out to be an almost impossible task since so many things can go wrong. By making the basic circuit work correctly first, then checking each supplementary circuit as it is added, the trouble can usually be quickly found because the complete hookup has been designed, made and checked stage by stage.

Component tolerance

It is at this point, too, that another factor must be reckoned with. It was stipulated that the components should be close tolerance, otherwise the predicted results cannot be expected. But close-tolerance components are expensive. It is useless to pick components at random out of a wide-tolerance stock for the sum total of errors may be so great as to make the circuit unworkable.

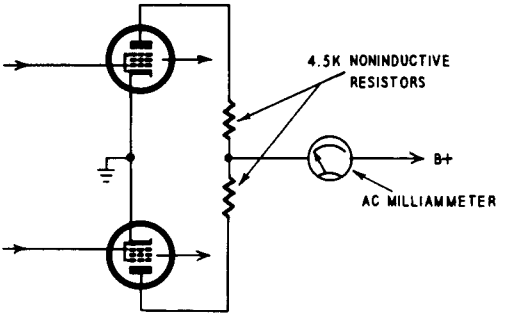


Fig. 504. *A method for determining the power output of the preliminary design.*

Accordingly, by the degree of tolerance permitted in component selection, a margin of performance must be permitted in the circuit design.

The figures given for voltage output, power output and stage gain depend on the values of components being correct. If they are not, the performance of the amplifier will suffer at maximum but probably not at small outputs. In other words, maximum efficiency can be secured only by insistence on exact-value components and, as this is uneconomical, the design must be changed for one of lower overall efficiency. Tubes must be given less work to do so more tubes may be required.

Use of additional stages

There is certainly nothing deplorable in the use of an extra tube—it just happens that it is cheaper to use an extra tube than close-tolerance components. The designer, therefore, uses his discretion to produce the desired results at the lowest *total* cost. This being so, the tests already described should be repeated with components at the limit of the tolerances selected. If the factory is going to use 20% resistors and capacitors, then the circuit must be tested with components plus or minus this percentage, some

being plus and some minus. It may then be found that for tubes selected at random the specified performance cannot be met without adding a stage. If so the extra stage must be added. It may be that the desired output power cannot be obtained from the two output tubes under any circumstances, in which case it might be as well to refrain from calling the amplifier a "25-watt high-fidelity" job and say what it is—possibly a very good amplifier with an "undistorted" output of 15 watts.

Having redesigned the basic circuit on these lines so that an honest performance is obtained with a sine-wave input of 1,000 cps, some other information can now be obtained. The original test was conducted at 1,000 cycles with a resistive output load because that is a set of conditions frequently used to make claims in advertisements of amplifiers. It may produce a good-looking set of figures but it is not the whole story. A high-grade amplifier has to reproduce all frequencies between about 20 and 20,000 cycles with constant output and constant small distortion. So the measurements must be repeated (with components on their outside limits) at spot frequencies of, say, 20, 50, 100, 1,000, 5,000, 10,000 and 20,000 cycles. If the amplifier is not intended to have a wider range than 100 to 10,000 cycles, then measurements outside these limits are unnecessary, but the designer should know what the amplifier is doing outside the specified limits to have some idea of what margin of safety his design possesses.

A circuit such as the cathode–resistance phase splitter of Fig. 305 cannot possibly work with 20%-tolerance components, although such components can easily be used in production. The success of the circuit depends on R1 being equal to R2 + R3 and C1 being equal to C2. Their exact values do not matter in the least—it is necessary only that the plate load be equally divided. Accordingly, in production all that is needed is to select these components in matched pairs.

Using negative feedback

Finally, there is the question of negative feedback, which cuts right across all matters so far considered. Again, for a very good reason, this does not render the work done so far fruitless. Too frequently negative feedback is used as a cure for trouble rather than a technique for improving something. While negative feedback will sometimes counteract many bad features in an amplifier, in cases even a great deal of feedback will produce little improvement, or may even make matters worse. An equivalent amplifier

can be made which will have a better performance without feedback than the defective amplifier with feedback. A competent and conscientious designer should be able to make a very good amplifier without availing himself of the benefits of negative feedback but, when it is applied to an amplifier that is already good, the final results are far in advance of the best designs without feedback. But the design must be right before feedback is applied.

Full development of the design for a high-grade 20-watt amplifier

A few pages back it was said that the work so far done was not of much practical value and that the amplifier so designed would not be of very much use. The reasons for doing what was described lay in stressing the importance of getting back to fundamentals before embarking on a complex circuit. The complete design of a working high-fidelity amplifier may now be undertaken, although the reader is advised to become acquainted with Chapters 6 to 9 before proceeding. Perhaps this section might have been postponed until after Chapter 9 but as the discussion is centered on amplifier design and, as amplifier design is virtually adopting the correct sequence of tubes and the optimum method of coupling them, it is natural to follow on from the consideration of power and voltage-amplifying tubes.

Choosing the output stage

First, then, is the question—what is the best output stage for a high-fidelity amplifier for home use? There is no categorical answer to that question since personal tastes differ and rooms range from the small to the very large. Moreover some speakers are rather insensitive and others produce far more output for less input. But a general idea can be formed from much experience in audio work. It will probably be agreed that among discriminating listeners who aren't deaf and don't want to be deafened and who don't want to interfere with their neighbors' comfort, 20 watts provides a reserve of power for entirely satisfying results with a well-designed speaker in a fairly large living room. This postulates that the *peak* undistorted power is 20 watts; the mean output will probably be no more than 1 watt. And it is this very great difference between mean and peak powers in musical reproduction that causes so much heartache when the genuine music lover hears sounds from his equipment that he knows quite well are not music but distortion. But where does the trouble lie?

Tube characteristics

When dealing with tube characteristics, it was pointed out that the data is average data, typical only of the particular tube being considered. Moreover, these tube characteristics are taken with a sine-wave input; necessarily so, since a static state must be established for the measurements to be recorded. The figures quoted by the tube makers are quite accurate for an average tube and if the tubes are checked by the designer, with sinusoidal input, no discrepancy will be discovered. But whereas the output power is given for a resistive load, and speakers are resistive, inductive and capacitive, so also the output power is given for a sinusoidal input of constant peak value, and in speech and in musical reproduction the peak value is continually changing.

Push-pull output stages can be operated with fixed or cathode bias. Fixed bias is almost always essential with class AB_2 and class B since the tubes are driven into grid current condition. For the more ordinary class-AB amplifiers for domestic use cathode bias is customary since it is considered (quite correctly as it happens) that cathode bias is self-regulating. As the plate current increases the voltage drop across the cathode-bias resistor increases, which results in increased bias voltage and reduction of plate current. See Figs. 505-a,-b,-c.

But things are not always what they seem. If the input is sinusoidal and then slowly increased from a low figure, it will be found that the self-balancing action does take place. If the same series of measurements is taken with fixed bias, it will be found that what was right for small inputs is not right for large inputs and distortion occurs. A class-B amplifier is usually intended for economical production of high audio powers, such as are needed for public-address work, where the highest possible quality is not essential, reasonable intelligibility at the lowest cost being what is required. High-fidelity reproduction at more normal output powers is associated with cathode-biased output stages.

If each output tube is biased separately, the bias resistor must be shunted by a suitable capacitance to avoid loss of bass. If the two tubes have a common bias resistor this is not necessary; indeed it is *essential* to omit the capacitor. The resistor and capacitor form a network which has a time constant; for good bass reproduction the impedance at low frequencies must be low to avoid attenuation (at high frequencies the impedance is negligible) and this implies a long time constant. When the chapter

on negative feedback is studied, it will be found that a long time constant is needed in the interstage couplings to reduce phase change at low frequencies. This helps to create instability in the form of motorboating. Since the large capacitors have a low impedance for low frequencies, the long time constant of the bias network has an equally serious drawback when the input to

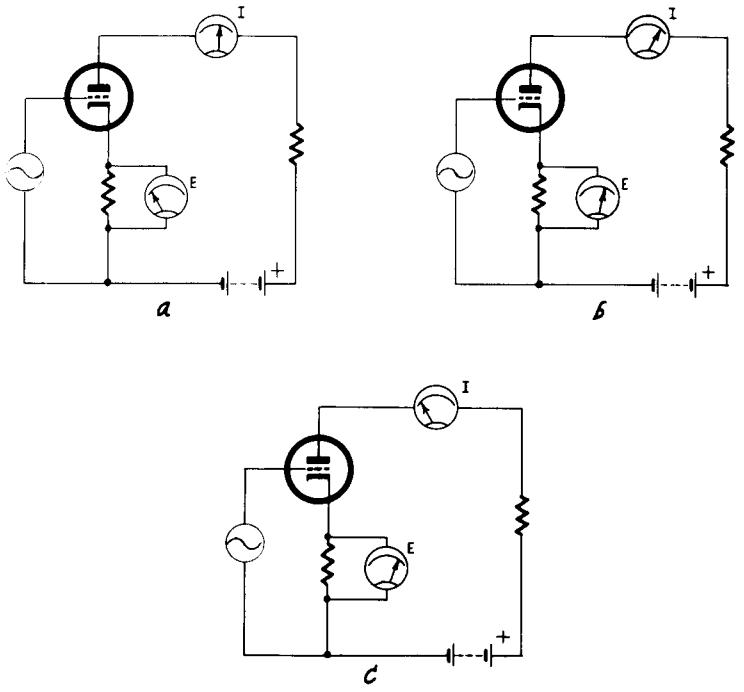


Fig. 505. *Self regulating action produced by cathode bias: a) with constant sinusoidal input, bias voltage and plate current remain at a relatively steady level; b) as amplitude of input signal rises, plate current increases and the bias becomes more negative; c) the change in bias tends to reduce the plate current and the tube tries to return to its original operating point.*

the output stage is far from constant and also includes transient peaks of maximum amplitude. Music is neither loud nor soft all the time; generally it varies very considerably from something just above audibility to deafening crashes, with a mean output from a 20-watt amplifier of something of the order of 1 watt or a little less. So long as the mean input is such as to produce 1-watt output, the output will be undistorted and it will be obvious that the plate (and screen) current will be constant and the bias constant, too—in fact, fixed bias. But if the output

tubes are used in class AB then an increase of input will produce an increase in plate and screen currents. This will be compensated by the increase of bias resulting from the increase of current through the bias resistor *so long as the change is gradual.* If a sudden increase of input is applied, shorter than the time constant of the bias network, there will not be time for the bias to change and the condition of fixed bias will continue in the peak, with consequent distortion. Hence the desirability of using a common bias resistor and no bypass capacitor because the time constant will be much shorter than when separate bias resistors are used, for these must be bypassed to avoid bass attenuation. See Figs. 506-a,-b.

Triode vs. tetrode output

A realization of this quite serious snag led to purists insisting that triodes in class A were the only tubes to use for the best possible results. In class A the plate current does not change with variation in input so the problem does not arise. Yet the poor efficiency of triodes and the limited power available from class-A amplifiers are drawbacks just as serious. Hence the wide use of tetrodes in class AB, but a keen pair of ears can always hear the cleanness in the peaks of a triode class-A amplifier and the distortion in those from a tetrode class-AB amplifier—*unless certain remedial measures are taken.*

The first has already been given—a common bias resistor without bypass. But it must be pointed out that a smoothing capacitor is shunted across the power supply and is consequently shunted across the bias resistor too. It is literally true that a substantial reduction in the capacitance of the power supply filter will help by reducing the time constant of the bias network, but such a reduction will impair the voltage-regulating properties when a large capacitor is used and make matters worse at all levels. Tetrodes and pentodes with constantly changing plate and screen currents demand good voltage regulation from the power pack. It will be found in practice that distortion will be reduced on peaks if the power supply filter capacitor is increased to something of the order of 50 μf. But the best way of all is to regulate the voltage by ancillary means such as the method given in Fig. 908 in Chapter 9. This permits a smaller value of capacitor in the power supply.

Next comes the use of negative feedback and, as Chapter 7 makes clear, this must be used with discretion. Negative feedback

can be used to such an extent that the distortion will be reduced to nothing; unfortunately the gain of the amplifier will also be reduced to nothing. Striking a happy mean between adequate gain and adequate absence of distortion, a point is reached at which, to avoid excessive phase change—which would change nega-

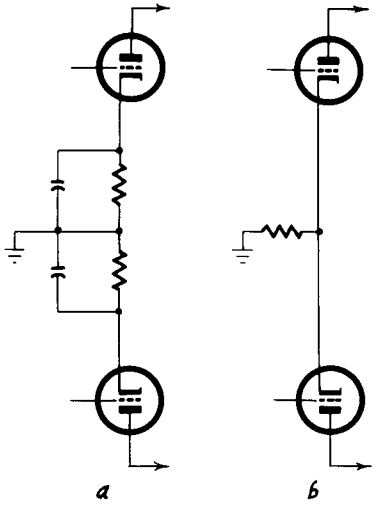


Fig. 506. *Bypassing and cathode bias: a) when separate bias resistors are used bypassing can be used to an advantage; b) when a common bias resistor is used the bypass capacitor should be omitted.*

tive into positive feedback at low frequencies—the number of amplifying stages must be reduced to the minimum possible number. Or, if several stages are desired, the feedback must not be taken over them all. A lengthy practical experience in amplifier design will convince the technical worker who has no preconceived bees in his bonnet that certain simple rules should be observed;

1. To secure the advantages of negative feedback, the minimum number of amplifying stages should be used. This postulates—

2. The output stage should be tetrodes or pentodes for the sake of their greater efficiency (resulting in a smaller required input voltage), provided steps are taken to reduce their effective plate resistance to give a good damping factor to the speaker.

3. The use of pentode voltage amplifiers both for their greater freedom from distortion and their greater voltage amplification per stage.

4. That high values of negative feedback should not be used to counteract amplifier shortcomings since this would reduce the overall gain too much and perhaps require an extra stage with the inevitable complications and instability.

Ultra-Linear operation

For something like 15 years the writer marketed a 20-watt amplifier consisting of a pentode voltage amplifier, a no-gain phase inverter and a pair of push-pull tetrodes. It gave this output power with only 0.7-volt input (an easy output for any tone control amplifier) and is still considered to have a very clean sound on the transient peaks. This is primarily due to a stabilized power supply and careful design. But it could be improved still more by comparatively recent developments variously called Ultra-Linear or distributed-load operation. With proper design, two tetrode-pentode tubes of 25-watt plate dissipation under Ultra-Linear operation will give 20 watts output power with appreciably less than 1% total distortion under actual working conditions, and 30 watts with only 1% total distortion. This quality standard is good enough to satisfy the most critical listener and will be adopted for the present discussion. It has been found that the Mullard EL34/6CA7 tube is particularly suitable for the type of operation now suggested, but other new tube types are becoming available which, no doubt, will give equally good results.

Positioning the transformer screen-grid taps

The most noticeable departure from standard circuitry when using Ultra-Linear operation is that the screen grids of the output tubes are taken to taps on the output transformer primary—the load is "distributed." The reason for this must be understood so that the design of the circuit can be made with intelligence. There is nothing mysterious in the operation if it is realized that an Ultra-Linear amplifier is a special type of feedback amplifier. Negative feedback is used to reduce distortion and plate resistance, making the pentodes behave like triodes. But the power efficiency of the multi-electrode tube is retained, particularly from the point of view of required input grid volts. It is generally recognized by expert designers that greater stability and overall freedom from distortion are more readily obtained if the negative feedback is not just slung back in one large chunk from the output transformer secondary to amplifier input. It is usually best to counteract output stage distortion by a separate feedback loop.

This can be done by the simple circuit of Fig. 704-b in Chapter 7, and this was the usual method of applying feedback when the idea first broke into then-current circuitry thinking. Feedback over several stages was a much later development.

Ultra-Linear operation involves the application of *nonlinear* negative feedback to the output stage through the medium of the screen grids. This results in a tube condition somewhere between that of a pentode and a triode, and is controlled by the position of the taps on the output transformer primary. In Fig.

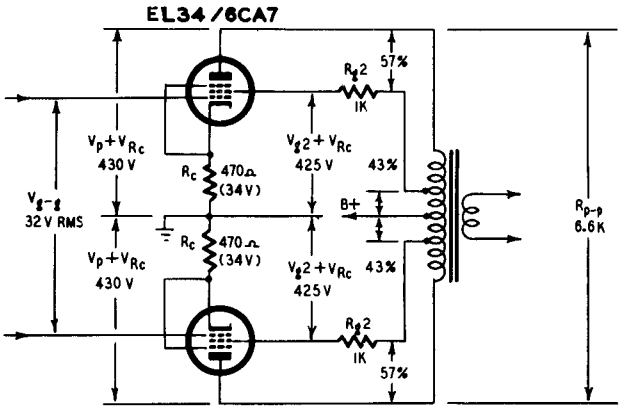


Fig. 507. *Ultra-Linear connection of output tubes.*

507 it will be seen that, if the screens are connected to the outer ends of the winding, they are virtually strapped to the plates and the tubes behave as triodes. If they are connected to the center tap, the tubes behave as pentodes. An intermediate position indicates distributed-load operation.

If the common winding, that is the portion between the center and screen taps, is about 20% of the whole, the distortion is about the same as the triode condition (although the plate impedance is not that of a triode). A definite improvement is achieved by increasing the common winding to some figure between 40 and 45%. Experiment shows that the optimum figure is 43%. Since this type of output stage is basically of the negative feedback type, it can be considered that the position of the tap controls the amount of feedback. The 43% tap being determined and the operating conditions laid down, a series of curves can be drawn to show the performance of the pair of tubes (Figs. 508 and 509). The recommended plate voltage is 400 and, making allowance

for the voltage drop across the output transformer primary and bias resistor, a B-plus supply of 440 volts is suggested. The function of the resistors R_g2 in Fig. 507 is simply to provide the correct screen voltage of 395. Fig. 508 shows the power output available with various plate-to-plate loads and, as the top curve indicates, the optimum load is 6,600 ohms. The bottom curve shows that at this load resistance the distortion is also least.

Driving voltage required

Fig. 509 shows what input grid-to-grid voltage is required for various output powers, and as the originally desired output was 20 watts the input volts will be 32 rms. However all alternating current has a peak as well as a mean value and the peak value is $\sqrt{2}$ times the rms value. The peak input voltage will therefore be $32 \times \sqrt{2} = 45.25$. From Fig. 509 it is seen that an input voltage of 45 gives a power output of 32.5 watts, but the curves showing harmonic distortion give a warning that cannot be ignored. Earlier in this discussion it was stated that the distortion at 20 watts output would be only 0.8% and the total distortion curve of Fig. 508 confirms this. If this were the only distortion curve shown, the designer might reasonably say that he would be willing to increase the distortion a little for the sake of having an amplifier with 30 watts output. (The extra 10 watts makes it just that much more attractive as a selling proposition, and it doesn't cost anything to get that extra power.) But the aim is to produce a real high-fidelity amplifier, and what happens if those extra 10 watts are taken?

Assume that it is decided to have a total distortion content of only 1%, which on the face of it seems a pretty good specification. This corresponds with an output power of 31.5 watts, so there appears to be 1.5 watts in hand. This lines up with an input voltage of 43.5 rms, and that has a peak value of 61.5 volts. Now look at the harmonic analysis of the total distortion. It is a very great help to designers that tube manufacturers have got into the habit of providing harmonic-distortion curves of output tubes since hi-fi became a force in the land. They don't print these data sheets just to make the tube handbooks bulkier; the curves are put there to be used, and used they must be if an amplifier is going to be good to listen to at high levels.

With the nominal output of 20 watts, the peak input of 45 volts develops 0.6% fifth-harmonic distortion at 32.5 watts output and the peak output for a nominal 30-watt amplifier will be right off the diagram. Fifth-harmonic distortion is the very devil. The

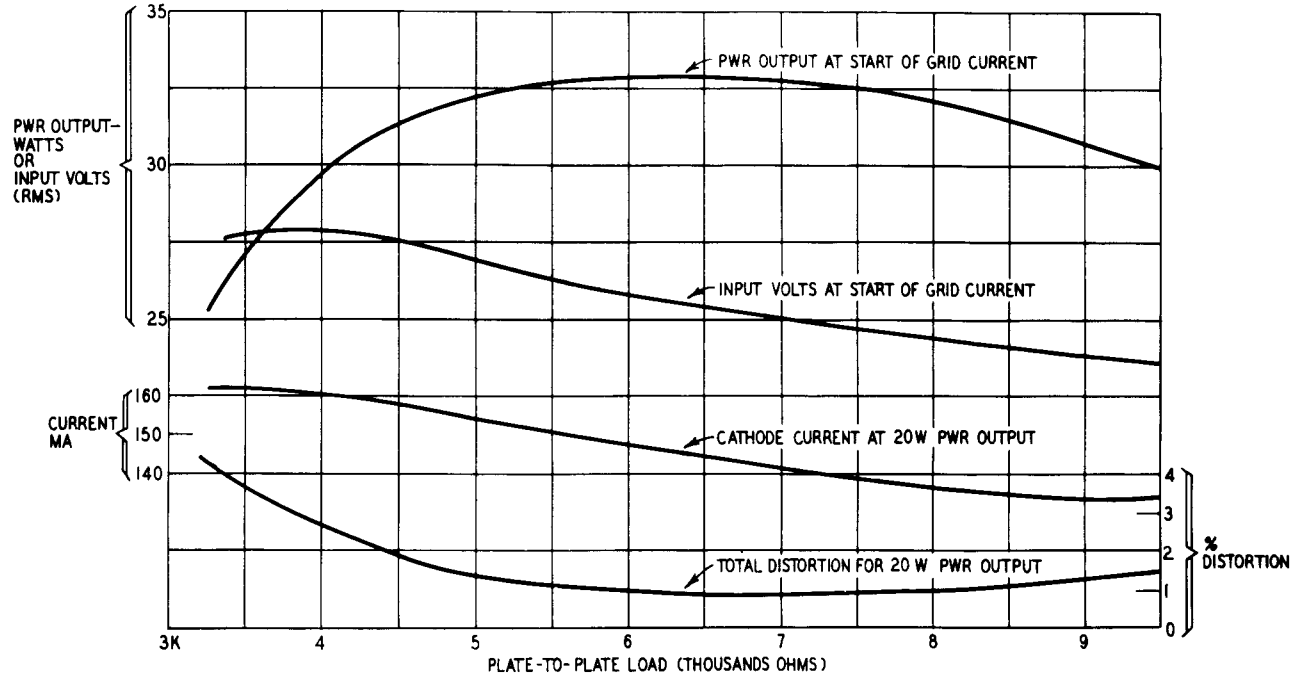


Fig. 508. *Power output as related to plate-to-plate loads in Ultra-Linear operation.*

ear is fairly accommodating to second and third, but even 0.5% fifth harmonic is intolerable to a musical ear. An input of 54 volts, gives nearly 1.3% fifth-harmonic distortion and a peak input of 61 volts would result in unbearable reproduction from a natural-sounding point of hearing. The curves of Fig. 509 have been extended to the right only sufficiently to show that at powers in excess of about 30 watts the distortion gets out of hand.

This being so, it may be asked why even 0.6% fifth harmonic is considered acceptable for the present design. The answer is, of course, that further feedback will be applied to the amplifier to reduce this. The fifth-harmonic distortion, as the principle component of the total distortion at high outputs, will be reduced most but, if the amplifier is supposed to have a 30-watt (or even greater) output, a considerable amount of feedback will have to be used. We are back to what has been pointed out several times—that the function of feedback is to improve an existing good amplifier, not to convert a bad one into something that can be listened to. Designs, therefore, which claim high output powers "with less than 1% total distortion" must be considered on their innate merits. And if the total distortion is mostly fifth harmonic, it will sound pretty bad to a keen pair of ears, even with considerable feedback. The present design contains 0.6% fifth (peak) and feedback will reduce it to about 0.2%, a just acceptable figure. If the distortion without feedback is something like 2%, feedback cannot reduce it to an acceptable figure unless most of the amplifier gain is reduced to an absurdly small amount. With which thought the discussion of the amplifier design can be resumed.

As overall negative feedback will be used, it is necessary to provide more gain in the voltage amplifier than would be required without feedback, if the input to the whole amplifier is to be kept reasonably small. The aim is to keep the number of stages at a minimum and the amplification per stage at a minimum. A no-gain phase inverter, therefore, seems undesirable, and so the cathode-coupled phase inverter is adopted. In constructing the prototype amplifier to prove the design under discussion, the writer happened to have on hand the Mullard high-mu double triode ECC83 and this was found to be eminently satisfactory. The American equivalent (12AX7) would be equally suitable. The first stage should be a pentode both on the score of high gain and freedom from distortion. Unfortunately, quite a number of pentodes are noisy and liable to introduce hum. Within the writer's experience the lowest-noise pentode he has met is the

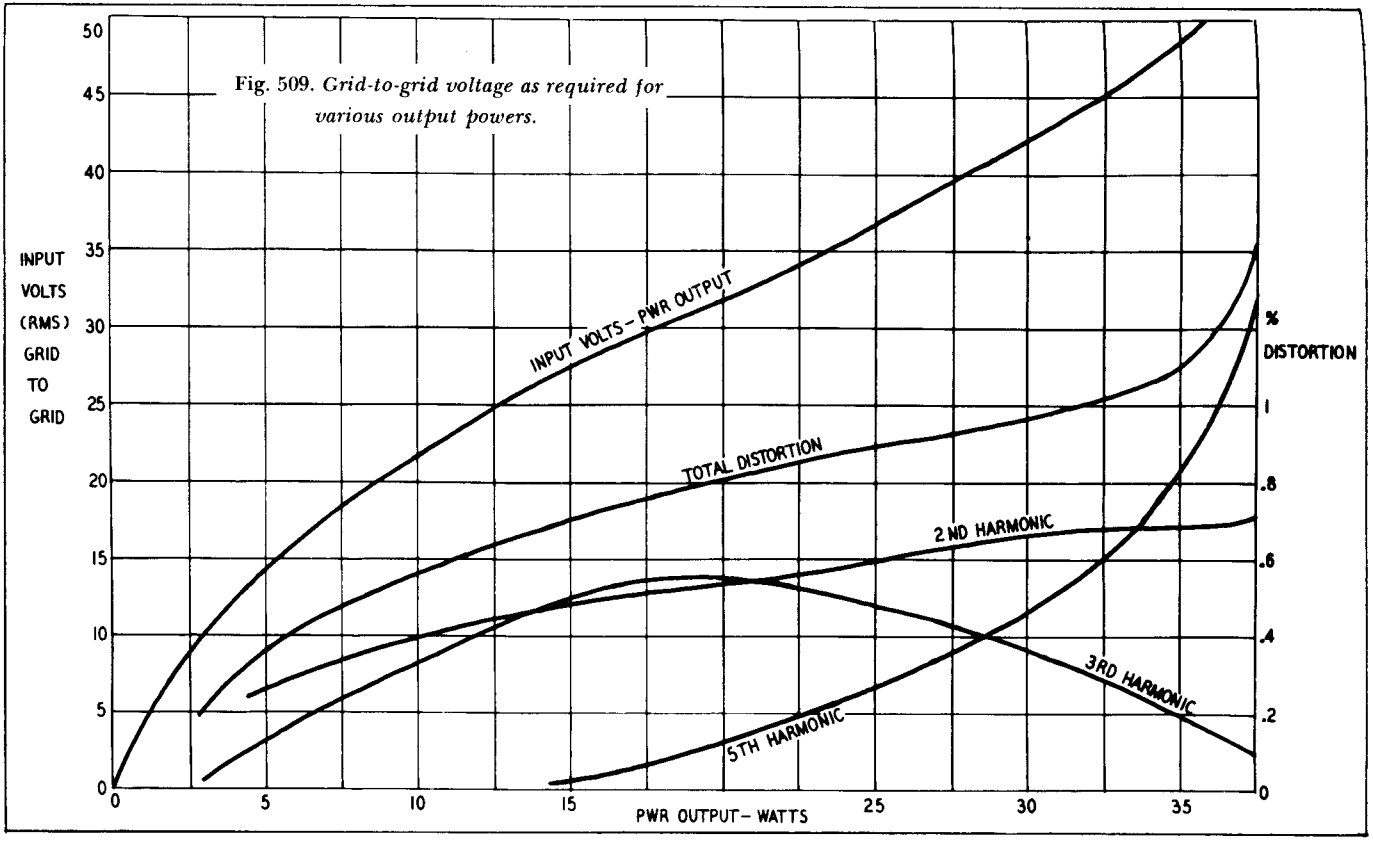


Fig. 509. *Grid-to-grid voltage as required for various output powers.*

Mullard EF86, which was developed several years ago for critical amplifier positions. The amplifier has a high gain and any noise or hum generated in the first-stage grid circuit will become unbearable in the output. Just any old pentode will not do; it must be genuinely designed for low noise.

As will be seen when the chapter on feedback is read, the basic cause of instability is due to phase change in the amplifier transforming negative into positive feedback. This is best avoided by reducing the number of stages to the minimum and reducing the R-C networks, too (in number but not in time constant). An excellent way of doing this is to use direct coupling between tubes and, as the cathode-coupled phase inverter conveniently lends itself to this, the first amplifying stage is direct-coupled to the phase inverter.

The output transformer must have been designed for Ultra-Linear operation and the taps for the screens must be at 43% of each half of the primary winding. Suitable components are available from Acro, Dyna Co., Partridge and others. The power pack is not shown in the circuit diagram (Fig. 507) but any good design will prove satisfactory, particularly if voltage stabilization has been included. The smoothed voltage output must be 440, with a current of rather more than 140 ma, to which, of course, must be added the current required for operation of the tone control preamplifier, recording amplifier, radio unit, etc. For the usual 500-volt 250-ma rectifier tube, capacitor-loaded, a power transformer secondary of 410–0–410 volts with a capacitor of 8 μf will prove just right. This allows for a 25-volt drop across the smoothing choke, which should have a maximum resistance of 200 ohms and an inductance of at least 10 henries at 180 ma.

Fig. 510 may now be considered in some detail. Starting at the input, capacitor C_X is not required if the preamplifier includes a coupling capacitor after the last tube's plate. If the amplifier were fed with a transducer having continuity to dc, the grid circuit of the first tube would be short-circuited if C_X were not there. Resistor R_X serves to reduce the input to the amplifier. The two resistors R_X and R1 form a potentiometer and could be replaced by a 1-megohm volume control. The amplifier has a sensitivity such that an input of 220 mv gives an output of 20 watts, and 300 mv-input just produces overload (in terms of the distortion discussed earlier). If the normal output of the preamplifier, including its tone control circuits, is greater than 220 mv, then R_X should be of such value as just to fully load the amplifier. It is assumed that the volume control will be located in the preamplifier.

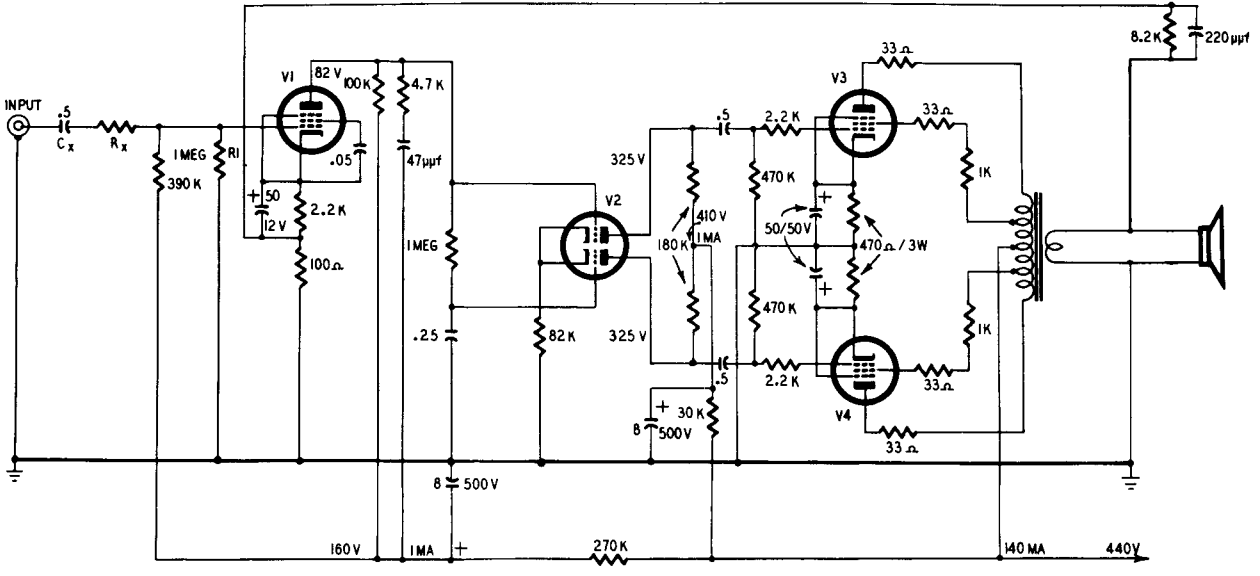


Fig. 510. *The complete amplifier less power supply.*

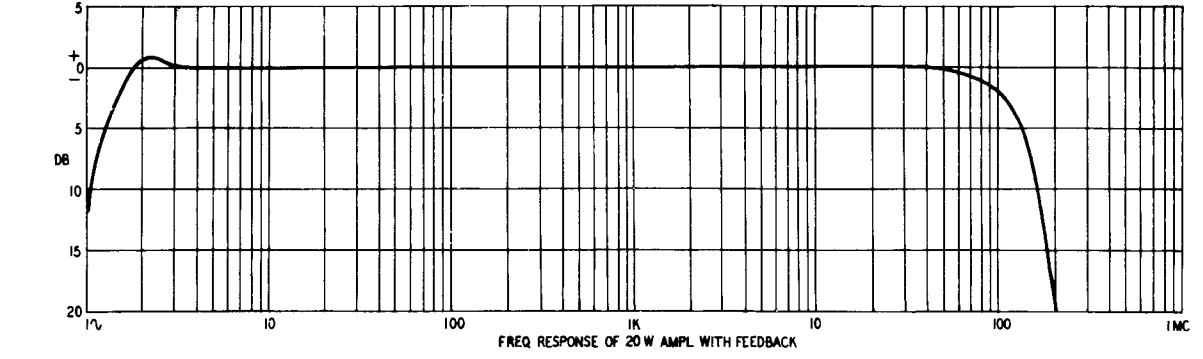
As the first tube (EF86) has been chosen for low noise, it is clearly absurd to introduce noise from other components. All resistors associated with the first stage should be deposited-carbon high-stability type, as also should the feedback resistor of 8,200 ohms. The noise level in the amplifier is 90 db below 20 watts, and the use of ordinary carbon resistors in the first stage could reduce this figure appreciably. It seems a pity to "spoil the ship for a ha'p'orth of tar."

The 4,700-ohm 47-$\mu\mu$f network shunting the plate load resistor of the first tube requires some explanation. In Chapter 7 it is pointed out that an amplifier to which adequate feedback is to be applied must have a very much wider frequency response than would seem to be necessary. Specially designed bass and treble cutoff circuits are required beyond the postulated range. In the bass this is required to prevent the passage of positive feedback caused by phase change of negative, but the present design has very little bass phase change because it has direct coupling and few stages. In the extreme top, however, a cut is needed since there is a tendency for the amplifier to be too good for the applicaton of feedback. Accordingly the supersonic treble is cut by the treble-cut network shown. In the performance curves of Fig. 511, it will be seen there is an upward trend in the phase change between 10,000 and 100,000 cycles although far below instability even at 100,000. Without the treble-cut network this would be greater, but it will be seen that the frequency response of the amplifier is not impaired, for it is virtually up to the 100,000-cycle value.

The phase-inverter plate-load resistors, the coupling capacitors and the grid resistors of the output stage should be closely matched. Their exact value is not critical but they should be matched. In the event of resistors not being dead-matched, the slightly higher values should be included in the circuits associated with that half of the phase inverter which is direct-coupled to the plate of the first stage.

The bypass capacitors of the output stage bias resistors are necessary because a common bias resistor is not practicable with the inherent feedback characteristics of distributed-load working. It is inevitable, therefore, that in this design there will have to be a time constant in the cathode circuits of the output tubes. Whether this matters or not can be disposed of by examining the output trace on an oscilloscope with square-wave input. In the present design it does not matter.

The 33-ohm resistors shown connected to the plates and screens



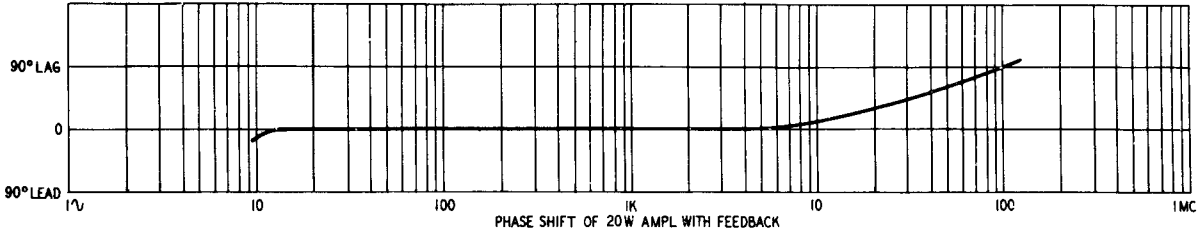


Fig. 511. *Performance curves for the amplifier of Fig. 510.*

of the output tubes are a foible of the present writer. Acting as high-frequency oscillation stoppers, they should be as small as possible and soldered directly to the appropriate pins of the tube sockets. They may not be necessary, but high-efficiency output tubes are prone to oscillate at supersonic frequencies and the stoppers can do a useful job without impairing the performance of the amplifier.

chapter 6

audio transformers

IRON-CORED transformers are used between stages and almost invariably as a means of coupling the output stage to the speaker or other power-actuated device. A good transformer is a somewhat expensive component and because it costs more to design and make, the low-cost ones are generally poor. Because of this, the interstage transformer has gradually dropped out of use in these days of better amplifiers, although it is desirable with class-AB_2 amplifiers, and essential for class-B amplifiers.

In the early days of the development of better reproduction, dynamic speakers were sometimes wound with high-impedance voice coils, a difficult and troublesome business. Before that, of course, electromagnetic speakers had high-impedance windings. These high impedance types did not require an output transformer between them and the output stage but their technical limitations drove them out of use. Thus the output transformer has become an essential component in most audio installations.

The design of a suitable audio transformer is not at all a difficult matter, but the actual manufacture of a high-grade component at a reasonable price calls for a somewhat specialized production technique and it is rare for a general electronics manufacturer to "roll his own." The specialist transformer manufacturer is, therefore, the usual source of supply. He will provide exactly what the designer wants if the designer knows what he wants.

Transformer characteristics

Transformers for audio frequencies are far from being ideal, simply because it is impossible to make a perfect transformer. The

ideal transformer would have infinite reactance at all frequencies but zero winding resistance. The windings would have no self-capacitance; there would be no leakage inductance (another way of saying that the coupling between the windings would be perfect) and there would be absolutely no losses in the core. In audio

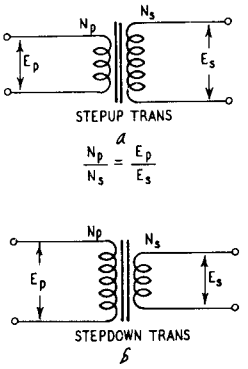


Fig. 601. *Relationship between primary and secondary windings of a transformer: a) stepup transformer; b) stepdown transformer.*

engineering the widest frequency band usually required is between 20 and 20,000 cycles, so a less-than-perfect transformer can be used quite successfully. The designer must know just what imperfections can be allowed. This involves some consideration of what the transformer has to do and how it does it.

A transformer usually has two windings, primary and secondary, but these are not always separate. A single-winding transformer with a tap in the winding is called an autotransformer. An ac voltage applied to the primary winding will be stepped up or down according to the ratio between the turns in the primary and secondary windings. If the primary winding has 1,000 turns and the secondary 2,000, then a voltage x applied to the primary will be stepped up to a voltage 2x across the secondary. If an autotransformer has a winding of 1,000 turns with a tap at 500, and if the primary voltage x is applied across the 500-turn section, the voltage across the whole winding will be 2x. But only exactly if it is an ideal transformer. See Fig. 601.

Transformer efficiency

The imperfections in a practical transformer may be lumped together and expressed as a percentage efficiency. The ideal transformer would be 100% efficient and the ratio of secondary volts to primary volts will be equal to the ratio of secondary turns to primary turns. But if the efficiency is less than 100% (and in small power transformers it may be as low as 85%), it is obvious that

extra turns have to be added to the secondary winding to get the desired output voltage. How many extra turns are needed depends on the size of the core, that is, on the cross-section of that part of the core which is surrounded by the windings.

Design considerations

To make the matter simple, the case of power transformers for the 50- or 60-cycle power line, such as are used in amplifier power supplies, can be considered first. There is no basic difference between these and transformers for a wide range of audio frequencies but, since they have to cater only to a single low frequency, they are easier to consider and design. Current passing through the primary winding sets up a magnetic field which results in magnetic flux in the core. If too much flux is created in the core, the core will become saturated and cease to perform efficiently; iron losses will be set up. The core must therefore be adequate to carry the total power transformed without saturation of the iron; the core size determines the number of turns per volt required for the primary winding.

A useful empirical formula, applicable to most normal core shapes, gives the area of cross-section of the core (in square inches) and reads:

$$\frac{\sqrt{VA}}{5.58}$$

where VA is the output in volt-amperes (assuming a working frequency of 60 cps). The number of primary turns required is obtained from the fundamental transformer equation:

$$N = \frac{10^8 E}{4.44\, fBA}$$

where N equals primary turns; f equals frequency in cycles; B equals maximum flux density and A is the area of the core cross-section in square inches. E, of course, is the applied ac voltage.

From this equation it is evident that the smaller the core, the more turns are required; that is, the turns per volt rise as the core size decreases. Now there is an optimum shape of core lamination and a small core results in a small lamination with a small window to accommodate the winding. But as more turns have to be placed into the small window, the gauge of the wire for the winding must be smaller and the dc resistance will rise. So power is wasted by heating the winding and the efficiency goes down. This loss of efficiency is due to copper loss (so called to distinguish it from

losses due to the iron core) and must be compensated for by increasing the number of turns on the secondary winding beyond the number that would be required by a perfect transformer.

Summarized, the considerations for a suitable design of transformer for a 50- or 60-cycle line are: the core should be of adequate size for the power transformed with due regard to the permeability and B-H of the core material; the primary turns must be properly selected for the core size required; and the secondary turns must be increased, beyond the theoretical number of turns required, to compensate for losses in the transformer.

It will be obvious that the larger the core the greater the winding space available, the heavier the gauge of wire that can be used and the smaller the copper losses. Accordingly, the larger the transformer the greater is its efficiency. Large industrial transformers are produced with an efficiency above 99%; very small ones may have an efficiency as low as 80%.

As another consideration a transformer for a 25-cycle power line must have a larger core than one for 50, because more turns per volt are required to avoid saturation, and one for 100 cycles can have a smaller core. We can now consider transformers for the whole audio-frequency range.

Audio transformers are used for power transference either as interstage or output couplings. There is also the special case, at the input of an amplifier, of small transformers for impedance matching between a microphone, pickup or recording head and the first tube, but such are called upon to handle very little power. However they introduce another problem: being used at a point which is followed by the whole gain of the amplifier, they must clearly avoid injecting any signal except that which originates in the voltage-creating device connected to them. They are susceptible to any stray interference in their field. A common type of interference is hum set up by the fields of other inductive devices such as power transformers and smoothing chokes. The input transformer must, therefore, have as small an external field as possible and must be completely shielded. A material of high permeability, such as Mumetal, is used for the core laminations as well as the screening box into which the transformer must be fitted.

The great advantage of high-permeability core materials is that a much smaller winding is needed to secure a given inductance. If the winding is smaller, its self-capacitance is appreciably less. What does this mean in actual practice? An audio transformer is an impedance-matching device, its stepup and stepdown character-

istics enabling different impedances to be accurately matched. In a power transformer the concern is impedance matching at one frequency and this can be translated simply into voltage alteration to suit the circuit requirements.

The audio transformer must match specified impedances at all frequencies if distortion is to be avoided. At low frequencies an insufficient primary inductance will cause a loss of voltage developed across the windings; at high frequencies the self-capacitance of the winding itself will shunt down the voltage while leakage inductance will prevent voltage being transferred from one circuit to the other. High-permeability cores will help at both ends of the frequency scale but, as high permeability cores are easily saturated, they can be used only for transformers handling small power.

If used for interstage coupling, the plate current of the preceding tube must be kept out of the primary winding by parallel fed connections.

Interstage transformers

Interstage transformers are also, generally, low-level transformers. It is obviously desirable that the transformation ratio should be constant over the whole frequency range of the amplifier; without particular care being taken in the design of the transformer this will not occur.

The stepup ratio at low frequencies can be expressed by the following equation:

$$\text{Stepup ratio at low frequencies} = \frac{\text{Stepup ratio at mid-frequency}}{\sqrt{1+\left[\frac{R}{\omega L}\right]^2}}$$

where R is the plate resistance of the previous tube, plus the primary resistance; L is the effective primary inductance and ω is equal to $2\pi f$. When $\omega L = 2R$, the attenuation will be 1 db; when $\omega L = R$, the attenuation will be 3 db.

At high frequencies, the reactance of the leakage inductance, the self-capacitance of the windings, and the value of R will form a low-Q resonant circuit so that the response of the transformer will peak at a frequency determined by the values of the components; beyond this resonant frequency the response will fall rapidly. The frequency of the peak can be raised by lowering the self-capacitance and leakage inductance of the windings, which

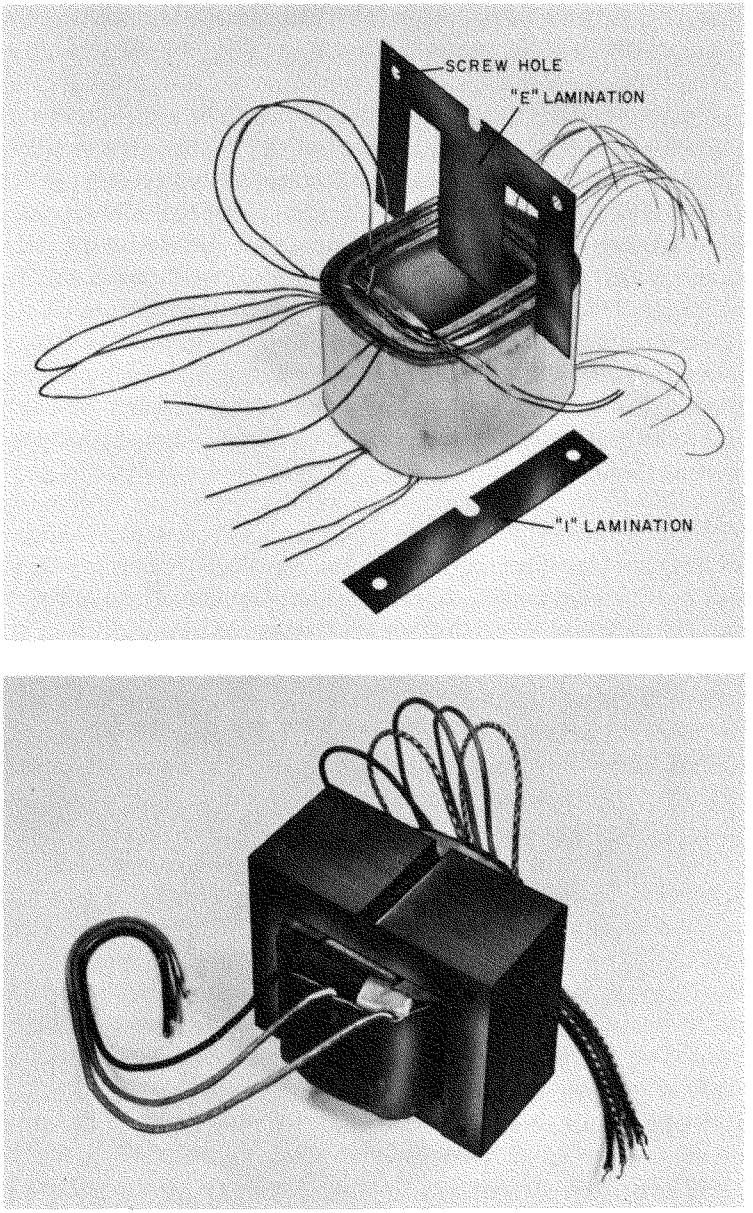


Fig. 602. *Internal view of a high-quality audio output transformer: a) single-core lamination and windings; b) assembly of complete core and windings. (Photos courtesy* Dyna Co.)

can be achieved by winding in sections with primary and secondary sections interposed to give very good coupling.

As the leakage inductance is proportional to the square of the turns, the response can be extended by decreasing turns but, as the primary inductance must be maintained for good bass response, the solution is to use high-permeability core material. As compared with silicon-steel stampings, Mumetal will increase the primary inductance of the same windings over twenty times. To remove the peak from the response of the transformer, this can be done by winding with resistance wire, which lowers the Q; but if this is overdone, the response will fall off too severely.

Output transformers

These general considerations also apply to output transformers but, as these are called upon to handle quite appreciable audio power plus the steady dc of an output tube or the out-of-balance dc of push-pull tubes, easily saturated high-permeability cores are unsuitable and recourse has to be made to silicon steel (or one of the new types of core recently introduced). As the leakage inductance and self-capacitance of the windings normally will be considerable—added to which is the difficulty that the secondary winding has comparatively few turns—attention must be paid to the method of winding and excellence of coupling between the windings. If, moreover, the output transformer is of the multi-ratio type, the coupling may not be constant over the range of tappings and the leakage inductance will therefore vary according to the tapping used for any particular secondary load. (The internal construction of a high-quality output transformer is shown in Figs. 602-a,-b.)

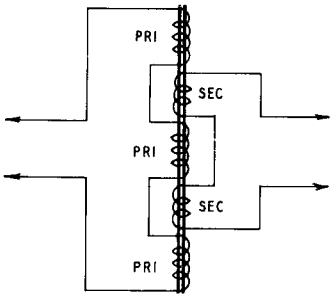


Fig. 603. *Method of alternating primary and secondary windings to reduce leakage inductance.*

Given a single ratio output transformer and choice of a heavy-gauge wire for the secondary winding, it is usual and convenient

to split the primary into three sections with a two-section secondary sandwiched between the primary sections. This will give reasonably low leakage inductance but the self-capacitance of the primary may be fairly high, and with considerable potential difference between adjacent turns. See Fig. 603.

The most perfect method that has been evolved is to wind the primary in one-turn-thick discs, which are then assembled with one-turn-thick secondary discs alternately along the core, but this is a very special method of construction. In general the primary should be wound into thin "pies", and assembled on the core, interleaved by insulating discs and secondary pies. This is a very good construction but the cost of manufacture often puts it out of consideration. The average output transformer is layer wound in sections with convenient interleaving of the secondary winding. See Fig. 604.

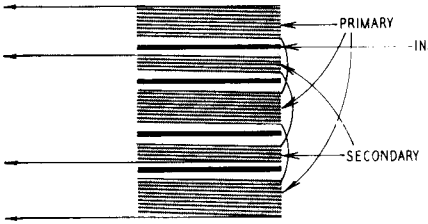


Fig. 604. *Special method of constructing a transformer in which the primary and secondary windings are interleaved.*

The preceding paragraphs outline the things to look for in transformer construction and, as these components are generally bought from outside sources, what must be done is to check their performance on receipt. This does not imply that the transformer manufacturers supply unreliable data as to the performance of their products, but too frequently the conditions of test and measurement are not specified in the technical data accompanying the component, or the conditions are not what the designer has in mind. It should be standard practice, therefore, to measure the performance of every transformer acquired for use.

Testing and measuring audio transformers

The first test is for ratio, and the nominal ratio is that at mid-frequency. This, for high accuracy, should be done on a special type of ac bridge but a simple test is quite sufficient for all ordinary purposes. Referring to Fig. 605-a, a 1,000-cps voltage is applied to the transformer primary through an input potentiometer. If the voltage source is controllable, the potentiometer is not required. A vtvm is now connected across the primary winding and

the voltage read; the meter is transferred to the secondary winding and the voltage read again. The transformation ratio is therefore the secondary reading divided by the primary reading; it is as simple as that. Remember, however, that this ratio represents *voltage* stepup or stepdown. Impedances vary as the *square* of the turns ratio, resistance and inductance being stepped up if it is a stepup transformer, capacitance being stepped down; and the reverse with a stepdown transformer.

Testing transformer frequency response

The frequency response of an output transformer can be checked by the method of Fig. 605-b. R1 represents the internal resistance of the output tube or tubes (not the plate-to-plate load which is the load reflected on to the tubes by the secondary load multiplied by the square of the transformation ratio). R2 is the equivalent resistance of the speaker, cutter, or recording head. The input voltage is held constant in magnitude but varied from the lowest

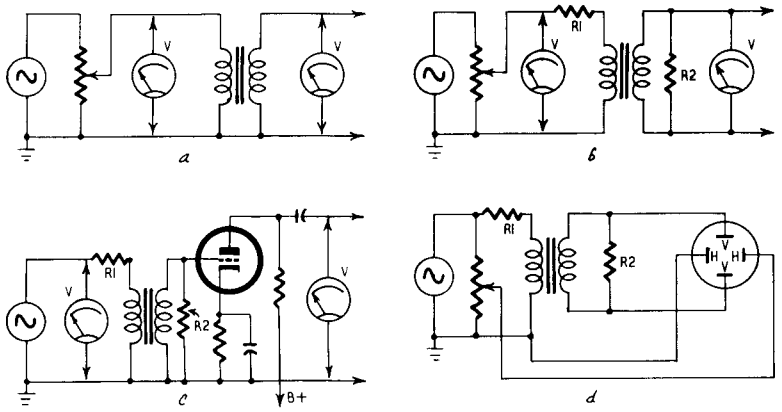


Fig. 605. *Testing and measuring audio transformers: a) for turns ratio; b) for frequency response; c) with vacuum-tube load; d) for distortion.*

to the highest frequencies to be handled by the transformer. For each frequency the secondary voltage is read and the ratio of secondary to primary volts plotted on logarithmic paper against frequency. In this way the frequency response is determined under the working conditions that would exist were the secondary load purely resistive. If the output transformer has a multitapped secondary, the frequency response should be taken on each tapping.

Checking interstage transformers

An interstage transformer should be checked with the tube that will follow it, as shown in Fig. 605-c. Here R1 represents the plate resistance of the preceding tube and R2 is the grid load resistor, if used. If not used, R2 is to be omitted. Readings are taken as outlined earlier and the frequency response plotted. Naturally, in the case of both output and interstage transformers, the applied voltage should be equivalent to the peak voltage that would be applied to the transformer when actually working in the complete amplifier. As the two tests just described do not reveal distortion, if present, the next test must be for distortion under maximum working conditions.

The simple test for this is to use an oscilloscope. If a sine-wave voltage is applied to the primary, the secondary should show a sine wave on the oscilloscope. Experienced operators will detect any distortion in the trace but it is not always easy to determine whether a given trace is a sine wave. If the circuit of Fig. 605-d is used, observation of the trace is much simpler. Instead of using the internal time base of the scope, the input voltage is applied to the horizontal plates of the c-r tube, the width of trace being adjusted by the high-resistance potentiometer across the input. As before, R1 is the equivalent resistance of the preceding stage and R2 is the load resistance. If an interstage transformer is being tested and no secondary load is intended, R2 will be omitted. The secondary winding voltage is applied to the vertical plates of the c-r tube.

The trace will be a straight line at 45° if no waveform distortion is present. A good transformer will not cause distortion at high frequencies apart from treble attenuation due to leakage inductance and self-capacitance of the windings but waveform distortion can occur at low frequencies. When the core approaches saturation, the magnetizing current waveform will be far removed from a sine wave and, although the input voltage will be sinusoidal, the voltage across the primary will not. As a consequence the voltage across the secondary will have the same type of distortion but magnified by the transformation ratio. This will show up on the trace as a splitting of the center part of the straight line into a bump on either side and displaced with respect to each other. If the primary inductance is too low, the straight line will be converted into a closed loop.

Parasitic oscillation

Transformers can be a contributing factor to parasitic oscilla-

tion. Interelectrode capacitance in tubes setting up Miller effect can cause feedback at supersonic frequencies. A transformer in the grid circuit of an output tube has inductance and capacitance, constituting a dynamic impedance. The plate circuit also includes an inductive reactance in the shape of the output transformer and feedback will occur depending on the gain of the tube and the

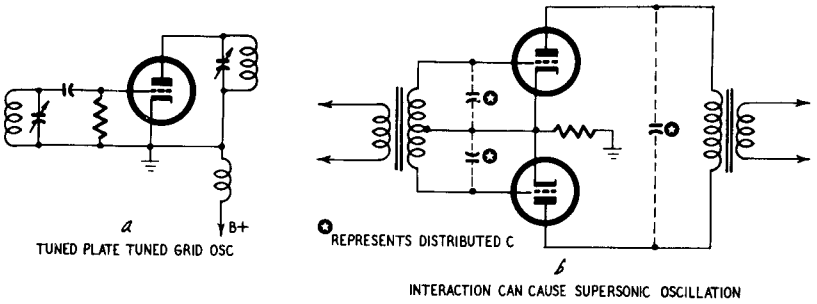


Fig. 606. *The use of interstage transformers can cause parasitic oscillation. Note the similarity between the tuned-grid tuned-plate oscillator (a) and the push-pull output stage using an input transformer (b).*

values in the grid and plate circuits. This can occur even when the interstage coupling is R-C, but it is more likely to happen with transformer coupling and still more likely with circuits designed for class-B amplification (Fig. 606). Presence of oscillation can be

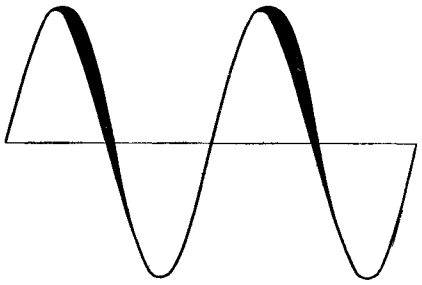


Fig. 607. *Spreading of the sine wave indicates oscillation.*

detected by a spread of the sinusoidal trace seen when the amplifier output is examined with an oscilloscope, using its internal time base in the normal manner (Fig. 607).

Because Miller effect is unavoidable, class-B interstage transformers must be very carefully designed for the precise tube and

circuit conditions under which they will operate and such transformers must be tested under their normal operating conditions. This tendency to oscillate does not imply that there is a defect in the transformer; it is merely a liability incurred by the use of transformers and resulting from tube properties. However, interstage transformers have to be used between tubes, so the final test is that which displays the presence of supersonic oscillation. This is naturally dependent on the inductive and capacitive characteristics of the transformer.

Specification of a transformer in terms of a power-frequency response is sometimes given by the manufacturer. This is a more useful indication than a plain statement of frequency response at some predetermined small input voltage.

Detection of shorted turns, determination of leakage inductance and self-capacitance of windings, core losses and other relevant data can be determined by the regular methods described in any standard textbook on electronic procedure. It seems hardly necessary to expand this chapter by detailing these methods. The tests given are the most important ones from the amplifier designer's point of view and it can be assumed that responsible manufacturers avoid sending out transformers with shorted turns. Again it is not of great importance to have precise values of leakage inductance or self-capacitance; what the designer wants to know is the frequency response of the transformer under working conditions and the tests given here will tell him that.

chapter 7

negative feedback

FEEDBACK can occur inadvertently or be applied deliberately; it can be positive or negative, voltage or current feedback. A typical case of positive feedback is the ordinary electron-tube oscillator, where the plate circuit is tightly coupled to the grid circuit so that the gain of the circuit is infinitely increased.

Feedback through the tube capacitances can be such as to set up self-oscillation; this, again, is positive feedback. Positive feedback is associated with increased gain, negative feedback with

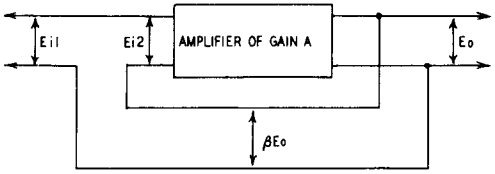


Fig. 701. *Block diagram showing the application of negative feedback.*

decreased gain. Yet the application of excessive negative feedback can cause an amplifier to oscillate at a low frequency (motorboating) or at a high frequency, but it is not the negative feedback itself that has caused the instability but the phase change in the amplifier resulting in the feedback becoming positive at low or high frequencies.

Fig. 701 shows any amplifier with a gain of A. In the absence of feedback $E_{i1} = E_{i2}$ and $A \times E_i = E_o$. Now feed back a portion

β of the output voltage E_o. Obviously β cannot be greater than unity and if the whole of E_o were fed back the input to the amplifier would be so great that overloading would be inevitable. β is therefore a fraction of E_o. The gain of the amplifier with β E_o feedback is

$$\frac{E_o}{E_{i1}} = \frac{E_o}{E_{i2} + \beta E_o}$$

Another way of expressing this is to say that the amplification after feedback is:

$$\frac{A}{1 + \beta A}$$

where A is the gain of the amplifier without feedback.

Similarly:

$$\text{distortion with feedback} = \frac{\text{distortion without feedback}}{1 + \beta A}$$

so there is the basic principle: negative feedback decreases the gain of an amplifier and decreases the distortion in the same proportion. In case the sign preceding the factor βA or βE_o is thought to be wrong when compared with formulas given in standard textbooks, remember that for negative feedback β is negative. The amplification formula is frequently given as:

$$\frac{A}{1 - \beta A}$$

but, if β itself carries a negative sign, the factor must become $+\beta A$.

Negative feedback stabilizes the gain of an amplifier and consequently improves the frequency response. Without feedback the gain will vary with frequency due to the presence of capacitive and inductive reactances and so will the phase shift. It is easy to design an amplifier with a flat response from, say, 100 to 10,000 cycles without feedback. Below and above this range the gain will fall and the phase shift will increase. If the phase shift becomes greater, negative feedback, when applied to the amplifier, tends toward positive feedback, so the gain lost by negative feedback is reduced and the width of response without loss of gain is greater. Considered this way it will be obvious that to depend entirely on phase-changed feedback to give a wide response is not sound practice, for so much feedback may have to be used that negative is changed to positive feedback at low and high frequencies, causing oscillation and instability. Feedback *can*

make a poor amplifier somewhat better but the true purpose of negative feedback is to make a good amplifier very much better.

This is possible because of other attributes. Negative voltage feedback decreases the effective plate resistance of the output tube or tubes in the same ratio as it reduces gain.

Provided the feedback voltage is in series with the input voltage (as shown in Fig. 701) the input resistance of the amplifier will be multiplied by the same factor, $(1 + \beta A)$.

Stability

Phase shift in the amplifier will result in instability if the amount of feedback or the phase change is too great. A single stage of resistance–capacitance coupling, with adequate screen and bias bypassing, will have a phase change of not more than $\pm 90°$ at the highest and lowest frequencies. If the bypassing is not adequate, the phase change can increase up to a maximum that is always a little less than $\pm 180°$. If there is more than one stage, then the total phase change will be the sum of the phase changes of each stage. An output stage with transformer has a phase change that reaches a maximum of 90° at low frequencies and 180° at high frequencies.

A tube without associated reactances changes the phase exactly 180° for all frequencies and a direct-coupled amplifier can have not only a very wide frequency response, but no phase change at the low frequency end except the reversals effected by the tubes themselves. If, therefore, feedback is used, all that is necessary is to make sure that the feedback has the correct phase to provide negative feedback and low frequency instability or motorboating cannot arise.

Phase shift and time constants

With an R-C amplifier involving phase change in each stage it is necessary to keep the phase change as small as possible, otherwise instability will occur with quite small amounts of feedback. To avoid this calls for much longer time constants than those given in Tables 8 to 10. Attenuation of the response at each end increases phase shift; longer time constants contribute to wide frequency response. If, in the interests of economy, the time constants are kept short and it is hoped to widen the response by applying negative feedback, it will be found that the desired amount of feedback may not be attainable owing to instability resulting from the large phase change introduced by the interstage coupling. In

other words, the great advantages of negative feedback are only achieved if the amplifier to which the feedback is to be applied has been properly and generously designed in the first place.

Since the design of an amplifier can be calculated, the conditions for stability can also be calculated. These can be displayed in a Nyquist diagram. The method is fully described in standard textbooks and technical journals; a useful summary will be found in Langford Smith's *Radiotron Designer's Handbook,* Fourth *Edition,* pp. 356 *et seq.* This book also gives references to the best articles on the subject. As it is the aim of the present work to avoid mathematical computation, some notes on practical steps to avoid or overcome instability are given in the next section.

Application of negative feedback

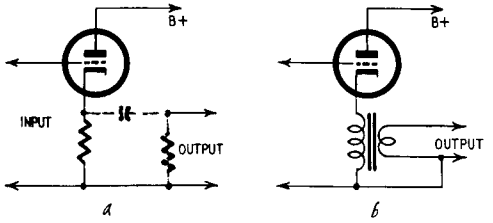
The simplest application of feedback to a single stage is found in the cathode-loaded or "cathode follower" tube, shown in Fig. 702-a. This is simply a case where an ordinary triode plate—cathode circuit has had the plate load transferred from the plate to the cathode end of the circuit. The basic circuit, as shown, postulates a conductive grid circuit, and when the tube is used as an amplifier, suitable negative bias must be applied, say by a bias battery.

To call the tube an amplifier is, however, hardly correct, for the gain of the stage is less than unity since there is inherent 100% negative voltage feedback. The output source impedance is very low, the input impedance high and distortion is practically nonexistent, so the value of the cathode follower lies in its action as an impedance transformer. The output terminal impedance is approximately equal to the reciprocal of the mutual conductance of the tube, in parallel with the cathode load. If this is used to feed a low value load impedance also, the power will be severely limited, and the distortion will rise too.

The cathode follower can be R-C coupled to the next stage by inserting a coupling capacitor and shunt resistance. It will be appreciated that the whole of the dc losses and the power output must be dissipated by the cathode resistor.

One way of solving both the bias and output coupling problems is to replace the cathode coupling resistor with the primary of an output transformer, as shown in Fig. 702-b, the secondary being connected to the load, which would normally be a speaker. This usage, because of the low output impedance characteristics of the cathode follower, gives extremely good speaker damping but as

there is no gain in such an output stage it follows that the necessary amplification must be made up by an additional stage of voltage amplification at quite a high level. The power output of the tube will be the same as that when used normally but, as the gain is less than unity, the input voltage required for normal operation must be multiplied by the amplification factor of the tube. The dc resistance of the transformer primary will con-

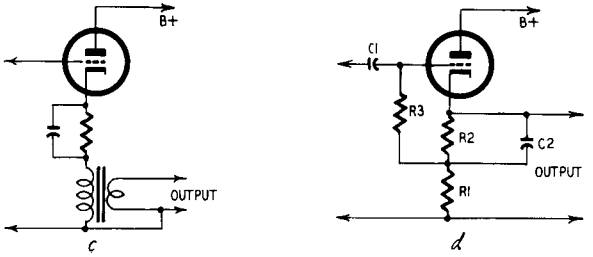


Figs. 702-a, -b. *Simple current feedback: (a) cathode follower; (b) cathode follower arrangement in which the primary of the output transformer replaces the cathode resistor.*

tribute toward the bias resistance required, but an additional bias resistor, with appropriate bypass capacitance may need to be inserted between the cathode and the transformer primary. See Fig. 702-c.

Simple current feedback

A simple way of applying feedback (in this case, current feedback) to any R-C stage is to omit the cathode bias resistor bypass capacitor. In Fig. 702-a, if the grid circuit is returned to ground,



Figs. 702-c, -d. *Cathode circuit arrangements: (c) bias resistor in series with primary of output transformer; (d) split cathode load.*

the cathode resistor will be the bias resistor, but this will generally produce far too much bias on the tube causing it to operate at a very low plate current. To overcome this, the cathode load can be divided, as in Fig. 702-d, where R1 and R2 form the cathode load (replacing the plate load) and R2 is the bias resistor

proper. C2 may or may not be needed as a bypass. The grid return would then be taken to the junction of R1 and R2. The blocking capacitor, C1, may be the plate coupling capacitor from the previous stage. If the load resistor is divided equally between plate and cathode circuits, the phase inverter circuit of Fig. 305

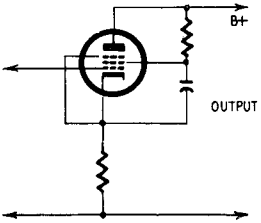


Fig. 703. *Method of connecting a series feed resistor.*

is formed. The split-load phase inverter is therefore partly a cathode-follower stage.

Triode-connected tetrodes and pentodes

The foregoing points refer to triodes. With tetrodes or pentodes, if the plate and screen grids are connected to the same B-plus

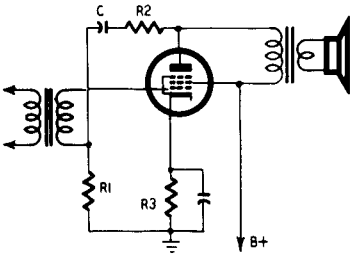


Fig. 704. *Applying feedback across a single tube output stage.*

supply directly, they will behave like triodes because the cathode load is common to both plate and screen circuits. For pentode operation the screen must be given an independent potential, most easily done by a series-feed resistance, suitably bypassed to cathode as shown in Fig. 703.

Single-stage feedback

Negative feedback over the output stage only is easy to arrange

and cannot introduce instability. It may include the output transformer or not, as desired. Fig. 704 shows feedback from the plate of the output tube to the interstage transformer. The ratio

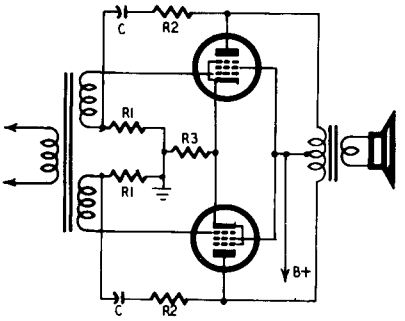


Fig. 705. *Applying feedback across a push-pull output stage.*

of R1 to R2 determines the amount of voltage fed back and blocking capacitor C should be chosen to give the minimum possible reactance phase shift for those frequencies normally on the flat part of the response curve without feedback.

The arrangement is simply duplicated in mirror image for a push-pull output stage, as shown in Fig. 705, but note that the

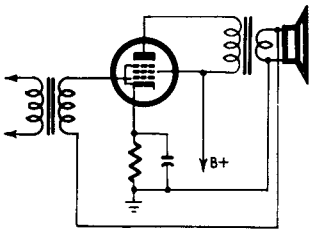


Fig. 706. *Applying feedback across the output stage and output transformer.*

input transformer must have separate secondaries as the two feedback resistors R1 cannot be merged. Bias resistor R3 is, of course, common to both tubes and does not normally require bypassing. Precisely similar circuits can be used for tetrodes or pentodes. If it is desired to include the output transformer in the feedback circuit, the feedback potential divider is hardly neces-

sary since there is appreciable stepdown in the output transformer. The ground terminal of the input transformer can be taken directly to one terminal of the output transformer secondary, as shown in Fig. 706. In this case a push-pull output stage does not require a two-winding input secondary, the center tap of the secondary winding being taken to the output secondary. The phase of feedback must be watched, for connection of the output transformer secondary the wrong way will result in positive feedback.

Two-stage feedback

Examples of feedback over two stages are shown in Figs. 707 and 708. Feedback can be taken from the output plate to the

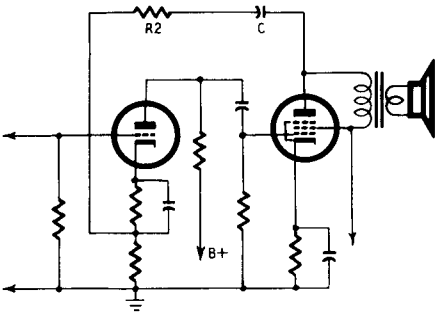


Fig. 707. *Applying feedback across two stages.*

previous stage cathode (Fig. 707), or from the output transformer secondary to the previous stage cathode (last part of Fig. 708 and first part of Fig. 707) or grid (Fig. 708). It is desirable to be able

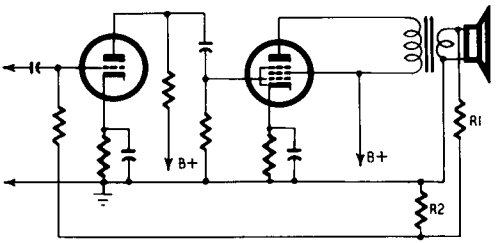


Fig. 708. *The secondary of the output transformer is part of the feedback loop.*

to adjust the feedback voltage, as in the arrangement of Fig. 704, so a potential divider is connected across the secondary winding,

represented by R1 and R2 in Fig. 708. In a developmental or experimental amplifier this can conveniently be a potentiometer or

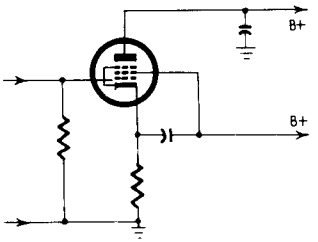


Fig. 709. *Rf bypass to ground in an early stage increases stability in feedback amplifiers.*

one of the resistors can be variable, which will allow a preset of either minimum or maximum feedback.

All these arrangements are equally suitable for push-pull working and, provided the reactances (and phase shift) of the first stage are suitably arranged, instability will not occur with reasonable amounts of feedback.

By methods similar to those of Figs. 704, 705, etc., feedback can be applied over three stages, or more, but instability will become

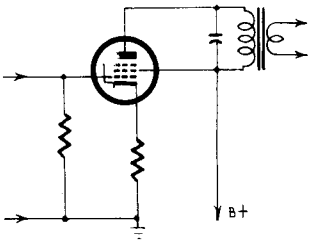


Fig. 710. *A capacitive shunt is sometimes placed across the primary of the output transformer.*

troublesome unless certain precautions are taken. As has been explained, this is because of cumulative phase shifts due to bass and treble cutoff in the successive stages of the amplifier. Supersonic oscillation may not be heard in the speaker yet it may rise to such a value as to burn out the speaker voice coil. The first step in designing any multistage amplifier is to make sure that the output stage has no tendency whatever to oscillate at high

frequencies in the absence of feedback. The oscilloscope is the best indicator of parasitic oscillations. Similarly there must be absolutely no tendency whatever to motorboating through common impedance coupling in the plate supply. The amplifier must be rock-steady without feedback under all possible working condi-

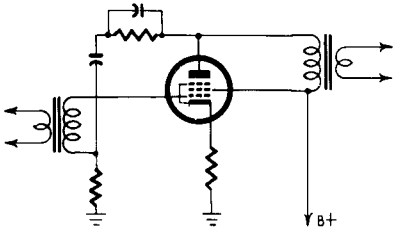


Fig. 711. *The diagram shows a capacitive shunt across the feedback resistor.*

tions so that, if instability appears when feedback is applied, it is known for certain that it is feedback that has caused the oscillations; otherwise the cause and cure can never be found.

Amount of feedback

The desired amount of feedback may not be possible if the design of the amplifier is inadequate. Instability can certainly be cured by reducing the amount of feedback and the easiest way

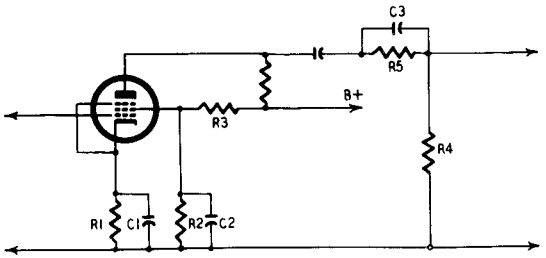


Fig. 712. *Bass cut filter for feedback amplifiers.*

out may be to make sure that the performance of the amplifier with reduced feedback is adequate. Otherwise, simple corrective devices include connecting a very small capacitor between the plate of an early stage and ground (Fig. 709) shunting the primary

of the output transformer by a suitable capacitor (value to be found by experiment) (Fig. 710) or shunting the feedback resistor by a capacitor (Fig. 711).

Staggered response

A more satisfactory method is to stagger the response of the several stages. In a three-stage amplifier, for example, by giving

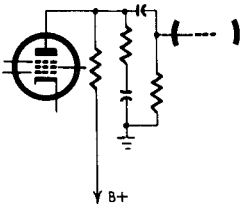


Fig. 713. *Treble cutoff filter is placed between plate and ground.*

the intermediate stage a narrower response than the other two, the phase shifts of the various stages will occur at different frequencies. Although the final response may not be as wide as one would have hoped for, the somewhat narrower response may permit a degree

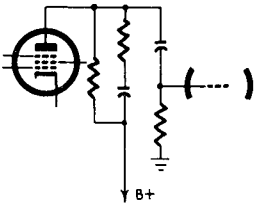


Fig. 714. *Treble step filter is shunted across the plate load.*

of feedback not otherwise obtainable, thus securing important feedback advantages with but slight loss of transmission characteristics (a convenient term for the frequency–amplitude characteristic of the amplifier).[1]

Modification of the normal amplifier response involves shifting

[1] H. W. Bode has evolved a method of some precision (see: *Bell System Technical Journal*, July, 1940, p. 421: "Relations Between Attenuation and Phase in Feedback Amplifier Design" and "Network Analysis and Feedback Amplifier Design"; D. Van Nostrand Co.) whereby the appropriate response characteristics are determined in a mathematical device called the complex frequency plane, which can be applied to provide an amplifier with good feedback and good stability. A simplified treatment assuming that the attenuation characteristics are straight lines is described by V. Learned (Proceedings of the IRE, July, 1944, p. 403: "Corrective Networks for Feedback Circuits"), but this needs some skill in application.

the frequency at which either bass or treble cutoff operates. Simple bass cutoff can be achieved by suitable choice of the cathode and screen bypass capacitors (C1 and C2 in Fig. 712), treble cutoff by the network between plate and ground in Fig. 713. But for the best results it is necessary to provide a "step" in the bass and treble attenuation. A bass step can be introduced by a network of the type shown in Fig. 712, a treble step by an R-C network shunted across the plate load as in Fig. 714.

It may be thought that the information just given is not very practical because specific values are not given. This is true but an empirical statement that such and such values of components are required is also of little practical value. It must be realized at all times that feedback is not just something that provides an easy way out for the shortcomings of careless design. It has been made clear that the response of the amplifier has a direct bearing on the behavior of the feedback loop. The amplifier *is* part of the feedback loop for the signal enters the amplifier, passes through it and is fed back to the input; the loop contains the whole circuit from original input to fed-back input. So, then, every reactance in the amplifier has to be considered in relation to the design of the complete feedback loop and the only way to do that is to calculate the complete circuit. This involves mathematical processes.

Stabilized feedback amplifier

It is obvious that the response of the amplifier must be known both as to voltage and phase change. Treble and bass boost or

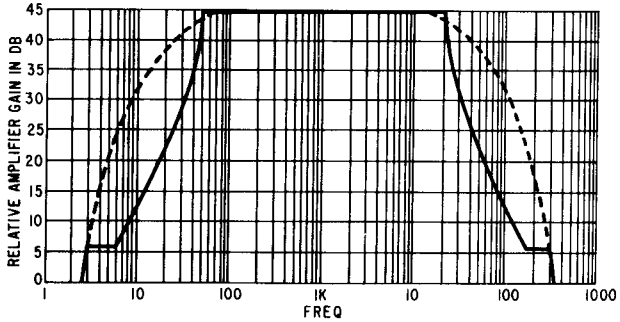
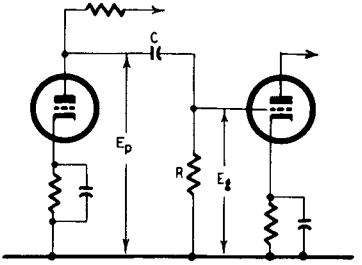


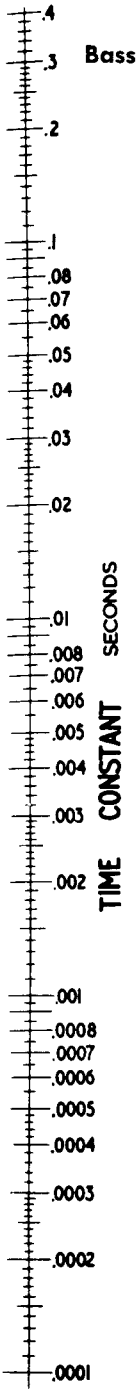
Fig. 715. *Ideal response of a stable feedback amplifier.*

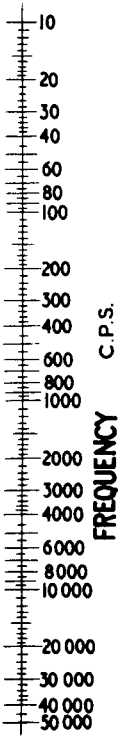
attenuation are also involved. The circuits which achieve this are filters, and filters are associated with what are usually called tone controls. These matters are dealt with in the next chapter.

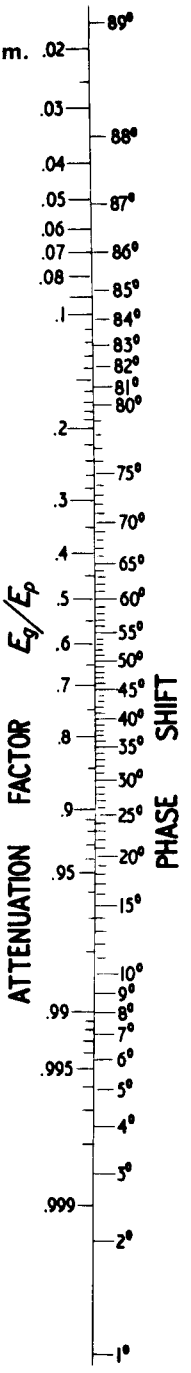
Bass attenuation and phase-shift nomogram.



Time constant in seconds = R × C where R is in ohms and C is in farads or R is in megohms and C is in microfarads.







The ideal response for a stable feedback amplifier is shown in Fig. 715. The dotted response curve represents the response expected from a well-designed three-stage resistance-coupled amplifier. At 6 db per octave attenuation of each stage, the attenuation over the three stages is 18 db per octave. The example given shows a flat response from 50 to 20,000 cps, with appreciable response at 20 and 40,000 cps. If substantial feedback, say, of the order of 30 db is applied to such an amplifier, phase shift of the attenuated response outside the desired limits of 50 to 20,000 cps will cause instability. To overcome this, the attenuation immediately outside the designed limits must be sudden and severe, as shown by the solid line, followed by a step, as indicated by the level parts of the solid response curve. Such attenuation characteristics are usually only obtainable by using L–C or L–C–R filters. In most cases these refinements are not necessary, as the degree of feedback is usually less; but where a three or more stage amplifier is to be improved by substantial feedback (in the present illustration there is also a nearly 10 db margin of safety, to guard against conditional instability) severe attenuation with a step filter is essential. It will be noticed that for a flat response from 50 to 20,000 cps the designer of a successful feedback amplifier is concerned with the response over a range as wide as 3 to 300,000 cps. The danger of instability is lessened by strict attention to the phase shift of each stage over the whole frequency response. Reference can be made to the phase shift nomogram shown on page 125.

chapter
8

filters and tone-controls

A NETWORK is a circuit made up of two or more components —resistors, capacitors, inductors or the mutual inductance between two inductors. These component parts of the network are called elements or parameters.

When the parameters are constant and independent of the current through them, the network is said to be linear. If any of the parameters are nonlinear, (such as amplifying tubes, iron-cored inductors, electrolytic capacitors, barretters, glow lamps, thermistors) the network is nonlinear.

If there is no source of energy within the network, it is said to be passive; if the network contains any sort of generator, it is called an active network.

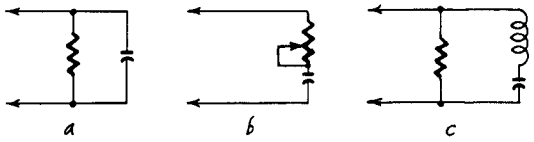
If the network has only two points at which power can be applied, such as the three diagrams in Fig. 801, it is called a two-terminal network; if it has four, i.e. has an input and an output, as in the examples of Fig. 802, it is referred to as a four-terminal network.

Even a simple potentiometer or volume control is a four terminal network, for the circuit of Fig. 802-b can have the upper right-hand resistor omitted and the junction of the other two is clearly the tapping point of the potentiometer. Circuits a and b of Fig. 801 will be recognized as, for example, the cathode resistor and bypass capacitor in an audio amplifier and a variable treble cut, connected between plate and ground of an amplifying tube. The circuit of Fig. 802-a is typically an intermediate-

frequency transformer in a superheterodyne receiver, and the mutual inductance between the two inductors is one parameter of the network.

Transducers

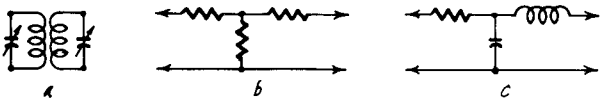
A four-terminal network is equivalent to a transmission system and has an optimum input and output load. Such a system is sometimes called a transducer, for example, when the term is ap-



Figs. 801-a,-b,-c. *Various arrangements for two-terminal networks.*

plied to a phonograph pickup or a loudspeaker. Although these devices have only two apparent terminals their equivalent circuits can be drawn in the form of a four-terminal transmission system.

If the network is purely resistive—that is, made up of non-inductive noncapacitive linear resistors—its behavior is constant



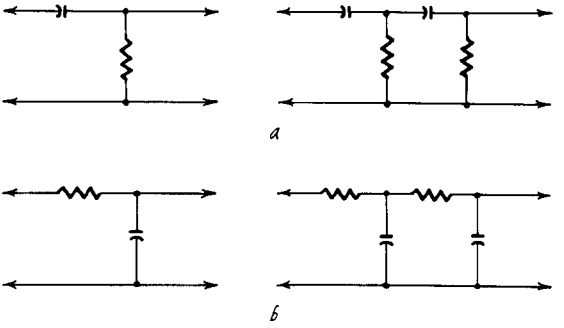
Figs. 802-a,-b,-c. *Four-terminal networks.*

for dc or ac of any frequency and the impedance of the network is the same as its resistance. With the introduction of capacitance or inductance, the reactance of the parameters, whose value is a function of the frequency of the applied voltage or current, means that the impedance of the network must be evaluated by Ohm's law for ac. All reactive networks are filters in a sense because a filter is a circuit which discriminates between different frequencies. Filters are designed to do certain types of discrimination and are conveniently divided into four groups.

Filters

A low-pass filter is one which transmits all low frequencies up to a specified limit and then attenuates frequencies higher than that. A high-pass filter is one which attenuates all frequencies below a specified limit while passing on without change those above it. A bandpass filter passes a specified band of frequencies

without change while attenuating all frequencies below and above the specified band. A trap filter attenuates a specified band of frequencies while passing on those above and below the specified band. By these definitions, a treble-cutting device is a low-pass filter; a bass-cutting control is a high-pass filter; the device which



Figs. 803-a,-b. *Single-section and two-section high-pass and low-pass filters.*

would produce a response of the curve of Fig. 715 is a bandpass filter and a scratch or heterodyne whistle filter is a trap filter.

Since a purely resistive network does not discriminate between frequencies, it is obvious that modification of the response of

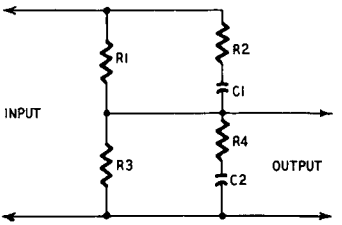


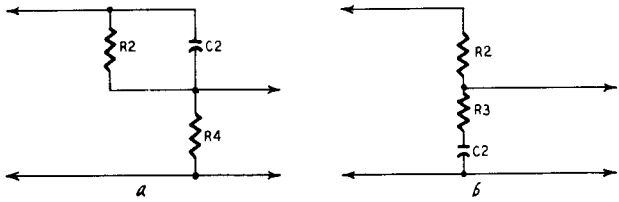
Fig. 804. *Basic tone control circuit.*

an amplifier with respect to frequency cannot be done by purely resistive methods; these are restricted to control of input or output regardless of frequency. Modification of the response with respect to frequency involves the use of capacitance or inductance or both, usually with resistance, to modify the effect of the impedances of the reactive parameters. For most purposes of tone control it is theoretically immaterial whether capacitors or inductors are used. The practical choice is usually capacitors because they are cheaper and much less liable to pick up hum; they are also more compact. But phase change occurs when either

capacitors or inductors are used, the phase change with an inductor being the reverse of that with a capacitor.

Attenuation and boosting

Confusion is sometimes caused by incorrect use of the term "boosting". One hears frequently of treble or bass boost but on examination it is found that that which is called boosting is nothing of the sort. To boost infers that something has been given a push, as it were. A high-pass filter is not a treble booster since it operates by cutting the bass. If, for some specified pur-



Figs. 805-a,-b. *Tone control circuits: a) bass cut; b) treble cut.*

pose, the treble must be emphasized, then it can be done only by employing an additional stage of amplification and wasting the bass that could be amplified by that stage—but no frequencies have been boosted. Genuine boosting can be achieved, of course, in some such way as replacing the resistive load of an amplifying tube by a load consisting of an inductor tuned by a capacitor. According to the values selected, the circuit will have a maximum impedance at its resonant frequency; this impedance may have a value which gives the maximum amplification from the tube but it may not be the value which gives the minimum distortion. The actual value of impedance and the sharpness of resonance naturally depend on the Q of the circuit, and in a radio-frequency amplifier it is often desirable that the Q should be as great as possible. On the other hand a sharp resonance in an audio amplifier is usually very objectionable; the Q can be lowered by shunting the tuned circuit with a resistor.

It is generally the aim of an audio designer to construct an amplifier with as flat a response as possible so tuned L-C circuits are not used except in such specialized applications as neutralizing a dip in the response of a speaker or providing a trap filter for needle scratch, caused by some resonance in the pickup system, (always excepting, of course, L and C components in filters and attenuators). Accordingly, the tone control in an amplifier is

nearly always of the untuned type and so is not boosting in the literal sense. There is, however, a loose convention that an ampli-

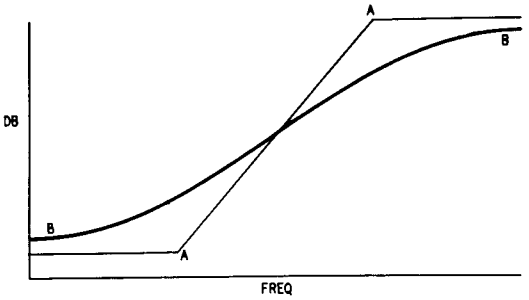


Fig. 806. *Ideal and practical attenuation curves: AA) ideal characteristic; BB) actual characteristic of an R-C filter. The slope of AA is 6 db per octave.*

fying tube in association with a high or low-pass filter is called a booster since the tube is a generator.

The rate of attenuation in a simple high-pass or low-pass filter is 6 db per octave, two frequencies an octave apart having a frequency ratio of 2:1. If a steeper cut than this is required, then the filter must have two or more sections; each section will cut at the rate of 6 db per octave, and for a perpendicular cut in the response

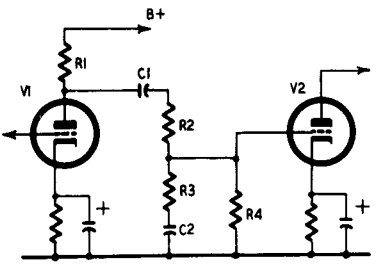


Fig. 807. *Active low-pass filter provided by two tubes and associated networks.*

an infinite number of sections would be required. Fig. 803 shows simple and two-section high-pass and low-pass filters. The first circuit of Fig. 803-a can be understood to represent the coupling capacitor and following grid resistor of an R–C coupled amplifier. Any alteration in value of the coupling capacitor can have no effect on the actual slope of attenuation; it can merely move the attenuation higher up or lower down the frequency scale of the amplifier. An amplifying tube can separate the two sections of the

filter in the second circuit. Similar considerations apply to the circuits in Fig. 803.

The basic circuit for all tone-control circuits employing resistance and capacitance is given in Fig. 804. The values of the components may vary between zero and infinity, these limits being the equivalent of short- or open-circuiting any of the components. Applying this to the elimination of certain components produces

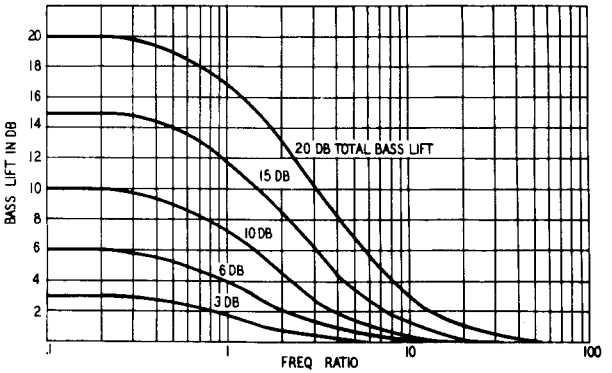


Fig. 808. *Response curves of the filter shown in Fig. 807 for different degrees of amplification.*

the circuits of Fig. 805 which show bass- and treble-cut circuits. A simple rule to determine the purpose of any network is to remember that a series capacitor will not pass low frequencies if it is of small capacitance, nor will a small shunt capacitor short-circuit

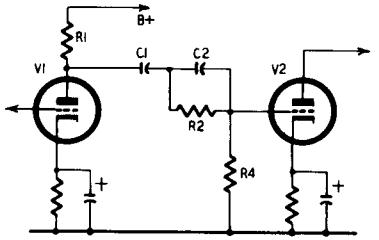


Fig. 809. *Active high-pass filter provided by two tubes and associated networks.*

low frequencies. The capacitor of Fig. 805-a is in series between input and output of the filter; in Fig. 805-b it is shunted across input and output.

In view of what has already been stated as to the 6 db slope produced by a capacitor it might be thought that the circuits of Fig. 805 will give that degree of attenuation; in practice they do not. In Fig. 806, the response AA, having a slope of 6 db per octave, is what would be obtained ideally by the circuit of Fig. 805-a. In

practice what one gets is a response shown by BB, at no point of which is the gradient as steep as 6 db per octave. The circuits of Fig. 805 are of no value without associated amplifying tubes and these provide the terminal impedances of the networks. Obviously

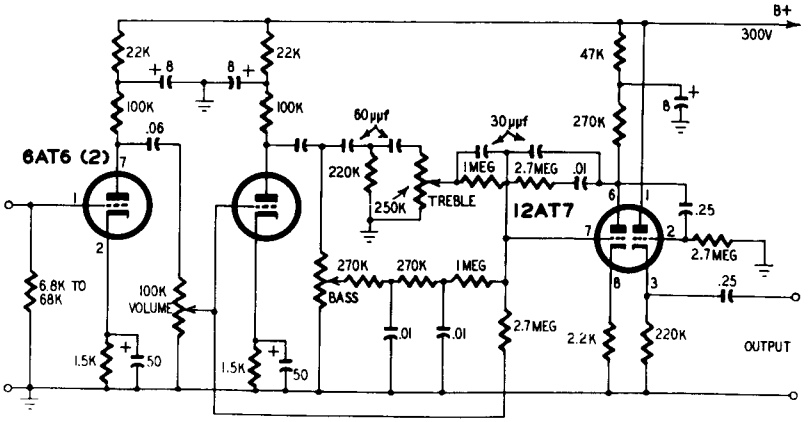


Fig. 810. *Tone-control amplifier giving smooth control of bass and treble.*

the terminal impedances, being part of the filter, have a bearing on its behavior, and as the terminal impedances will necessarily vary

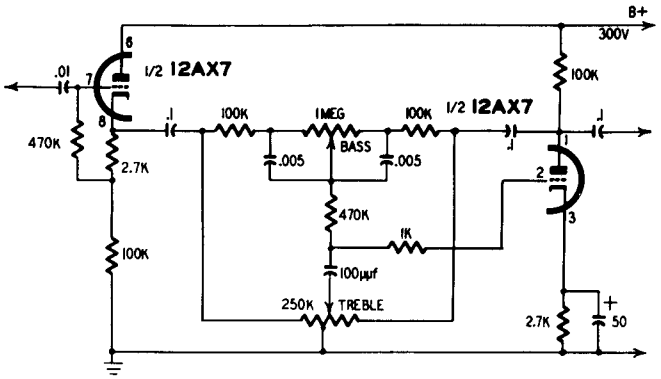


Fig. 811. *Simple tone-control amplifier.*

with frequency as with the types of tubes used and *their* parameters, the design of the filter must take into account the types of tubes used and the conditions under which they are operated.

Fig. 807 shows two tubes and associated networks to provide an active low-pass filter. This, as has already been explained, is us-

ually called bass boosting, but as the boost is obtained by non-amplification of higher frequencies, it is desirable to call the circuit what it is. R1 is the plate load of the first tube and is in parallel with the internal plate impedance of the tube, with the series network R2–R3–C2, and the second tube's grid resistor R4. The value

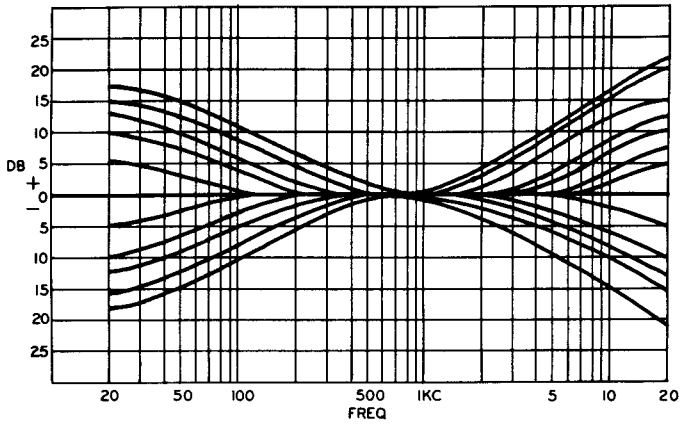


Fig. 812. *Response curves of the tone-control amplifier shown in Fig. 811.*

of R1 is determined by the tube selected for the first amplifying stage; that of R4 must not exceed the grid–cathode resistance specified for the second tube.

The impedance of C2 decreases as frequency rises, so acting as a partial short circuit to ground. R4 should be many times the value of R3 (at least 20 times) otherwise it will act as a short across the circuit R3–C2. C1 is selected to give adequate bass response, as detailed in earlier chapters, but may be chosen to give bass attenuation below those frequencies it is desired to pass through the filter. The reactance of C1 is assumed to be very small as compared with R4. It may happen, however, that the combined impedances of C1, R3 and C2 are so low that, being shunted across the plate load R1, they cause distortion through inadequate load on the tube; this is counteracted by inserting R2. If the first stage is required to deliver only a small output it may be omitted, but the value of R4 must not be less than twice the plate resistance of the first tube; for higher outputs and when R4 is not high enough R2 must be included. This point is mentioned because it is frequently omitted in published tone-control circuits when its presence is essential, which results in a tedious hunt for distortion that ought

not to be there. R2 has the further advantage of improving the ratio between bass and treble.

The amount of boost is a direct function of the equivalent values in the circuits of Fig. 805. It will be noticed that the slope of the curves in db per octave is a function of the total boost. In Fig. 806 the point where curves AA and BB intersect is mid-frequency, and this is where the slope is defined. Curve AA is shown with a slope of 6 db per octave between the limit frequencies, but the practical curve BB only reaches its maximum

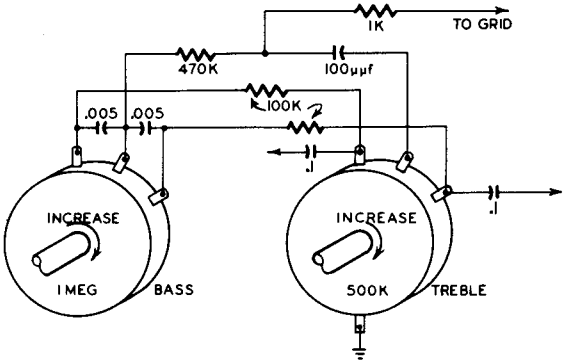


Fig. 813. *Construction details of the center section of the tone-control amplifier shown in Fig. 811.*

slope at the mid-frequency. In Fig. 808 the family of curves has the following maximum slopes:

Total bass lift (db):	3	6	10	15	20
Slope at mid-frequency (db/octave):	1	2	3	4	5

The horizontal axis of Fig. 808 is described as the frequency ratio; the curves can be moved to the left or right by altering the value of C2. The amount of bass compensation can be varied by shunting C2 with a variable resistor.

The companion circuit to Fig. 807, this time for an active high-pass filter, is shown in Fig. 809, and, by suitable choice of components a set of curves can be drawn which are a reversal of those shown in Fig. 808. As before, altering the value of C2 moves the curves to the left or right; the amount of treble gain can be varied either by changing R2 or inserting a variable resistor in series with C2 and the junction of R2 and R4.

Scattered through technical handbooks and magazines are a

host of circuits giving continuously or stepped variable combined bass and treble controls. These differ in complexity and behavior, although the final result is almost always satisfactory. Nothing would be gained by reproducing these various tone-control amplifiers, which is what they really are, but the writer submits Fig. 810 as a design which he has used with great satisfaction for a number of years. This gives a very smooth control of bass and treble and is particularly suitable for the proper compensation of various phonograph recording characteristics as well as improving the response of broadcast transmissions. The incorporation of a cathode-follower output enables the amplifier to be constructed as a separate unit and used fairly remotely from the main amplifier.

Another circuit, originally described by P. J. Baxendall (*Wireless World,* October, 1952) and further dealt with by B. T. Barber (*Audio Engineering,* September, 1953) is shown in Fig. 811. It has the merit of greater simplicity and a somewhat greater amount of available control. This can be assembled into a very small unit through the introduction of a combined component embracing the whole center network (Centralab's Junior Compentrol C3-300). The writer has found the circuit of Fig. 811 very convenient indeed both as to capabilities and compactness.

The response curves of Fig. 812 illustrate the wide range of control which the circuit makes available. Fig. 813 shows the construction detail of the center section of the control amplifier when the Compentrol is not used.

chapter 9

amplifier power supplies

RECTIFIED ac for the plate supply of amplifier tubes can be obtained from high-vacuum rectifier tubes, from gas tubes (mercury vapor), from metal (selenium or copper oxide) rectifiers or from vibrators (for operation from low-voltage storage batteries such as for car radios). In all cases the rectified ac must be filtered and smoothed so that it is steady dc, otherwise hum, usually the second harmonic of the supply frequency, will be injected into the amplifier. Almost invariably high-vacuum rectifiers are used for domestic equipment, presumably on the score of cheapness, but metal rectifiers have much to commend them.

The operating conditions for rectifier tubes are given in some detail in tube manufacturers' handbooks and there is no point in repeating such information here. This chapter will be concerned with explaining the various types of rectifier circuits and considerations of filtering and smoothing. It was made clear when discussing output tetrodes in high-fidelity amplifiers that a stabilized plate and/or screen supply was essential for the best results, so methods of obtaining this are described.

Half-wave rectifier

The simplest circuit is the half-wave rectifier, shown in Fig. 901; circuit *a* shows a thermionic rectifier, *c* a metal rectifier, both with a capacitive load. The waveform of the output of the rectifier tube is shown at *b*; that of the metal rectifier with capacitive load at *d*.

With a resistive or inductive load the waveform is the same as that of the tube rectifier.

Since half the sine wave input of the power line is suppressed, there is necessarily an interruption of the supply which must be

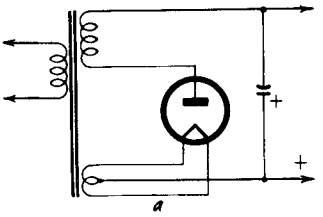


Fig. 901-a. *Half-wave thermionic rectifier.*

smoothed out in the filtering system. Imperfect smoothing will result in a hum at the line frequency.

The rectified output can be made continuous (although varying between peak and mean voltage) by using two rectifiers in oppo-

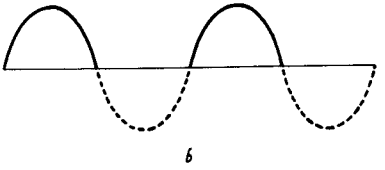


Fig. 901-b. *Half-wave rectifier output waveform.*

site phase. These are ordinarily combined in one envelope to form the usual full-wave rectifier, as shown in Fig. 902, where, again,

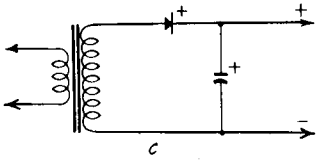


Fig. 901-c. *Half-wave supply using metal rectifier.*

circuits for thermionic and metal rectifiers are illustrated, together with the waveforms of the respective rectifiers. Imperfect smoothing, as mentioned earlier, will obviously result in hum at double

the supply frequency. The higher the frequency of the ripple to be smoothed, the easier it is to smooth it, so full-wave rectification is preferred to half-wave in audio design.

Bridge rectifier

All types of rectifiers have a stated maximum power output. If the current taken is reduced, the voltage output can be increased

Fig. 901-d. *Output waveform of metal rectifier with capacitive load.*

and *vice versa*. If the required power is beyond the resources of a given rectifier, then it must be duplicated or a larger type chosen. Connection can be made in parallel (as in the similar case of paralleled output tubes) or in bridge fashion.

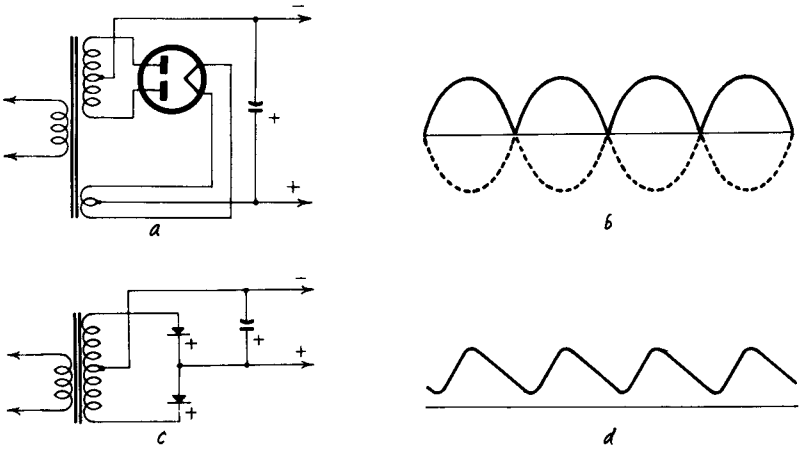


Fig. 902. *Full wave rectifiers (a and c) and their output waveforms (b and d).*

Bridge connections are shown in Fig. 903 for both thermionic and metal rectifiers. From an examination of Fig. 903 observe that four half-wave rectifier tubes must be used, the required connections being impossible with two full-wave units. Furthermore, three separate filament or heater windings must be provided in the power transformer. As this increases the cost of the transformer considerably, it is customary to use two full-wave rectifiers in parallel when one filament winding is all that is needed. With metal rectifiers bridge connection is widely used and very con-

venient, for the circuit arrangements of Fig. 904 can be arranged in a single stack of elements. The waveforms of the output of a bridge rectifier are identical with those of Fig. 902.

Voltage doubler

The voltage-doubling arrangement is very convenient for high values of rectified volts because the transformer secondary need have only half the turns required for a half-wave, full-wave or bridge rectifier; but since the input volt-amperes to the rectifier

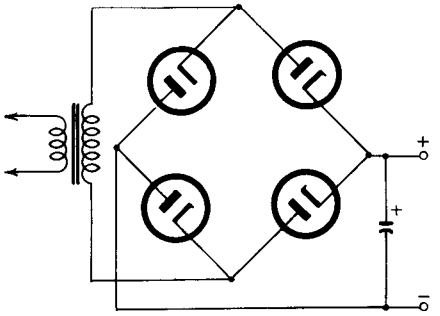


Fig. 903. *Thermionic bridge rectifier.*

must be the same, the current is double. Voltage-tripling, or -quadrupling circuits are also available. Voltage-doubling circuits

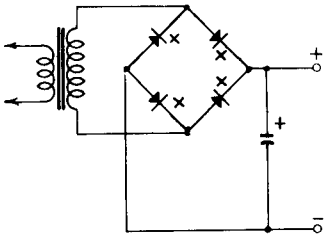


Fig. 904. *Metal rectifier bridge connection.*

are shown in Fig. 905. Since the reservoir capacitors are in series their working voltage need be no greater than that required for ordinary rectifiers.

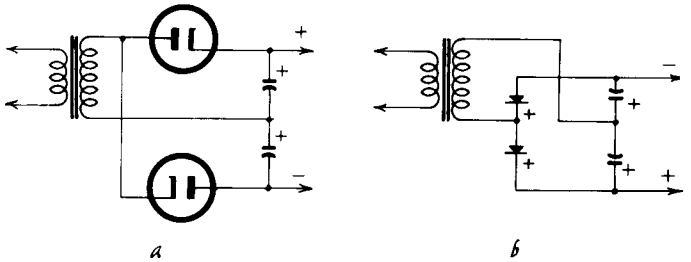
Filters

Rectifiers work into a filter whose function is to smooth out the ripple on the rectified output voltage. The part of the filter

directly connected to the rectifier can be either a capacitor or an inductor, these alternative methods being termed capacitor input or choke input respectively. The circuits of Figs. 901 to 905 show capacitor input arrangements with no further filtering. The choke input equivalent of Fig. 902 is given in Figs. 906-a, -b, where the choke can be connected in either the positive or negative leg.

Capacitor input filter

For capacitor input, the output voltage of the rectifier is a function of the capacitance of the input or "reservoir" capacitor. The actual voltage will, of course, also be governed by the current taken from the rectifier but, for a given current and power transformer secondary volts, the rectified volts will increase as the capacitance of the reservoir capacitor is increased. If this is increased too much the rectifier will be seriously overloaded during parts of the supply cycle.

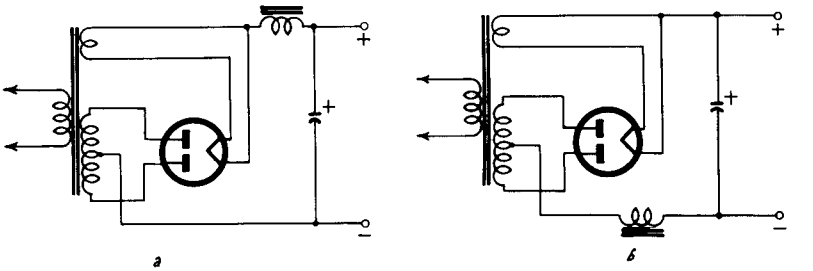


Figs. 905-a,-b. *Voltage-doubler circuits: a) using thermionic rectifiers; b) using metal rectifiers.*

With capacitor input, the ac volts applied to the rectifier, thermionic or metal, need never be greater and are usually slightly less than the dc volts required. With a full-wave rectifier of the type shown in Fig. 902 the required voltage must be applied to each plate so the voltage across the center-tapped transformer secondary will be double that required for a half-wave or bridge rectifier and four times that needed in voltage doubler. But as the total volt-amperes is the same in all cases, the current will be proportionately reduced. The design of the transformer cannot, however, be undertaken without a knowledge of the value of the reservoir capacitor. The necessary information can be obtained from the data books of tube manufacturers, and, if data for metal rectifiers is not obtainable, it can be taken as similar to that for thermionic rectifiers.

Choke input filter

With choke input, the ac voltage to be applied to the rectifier is appreciably higher than the dc voltage required. This seeming disadvantage is outweighed by the good voltage regulation of the choke-input system. This inductor has a minimum critical value according to the current demand (usually given for maximum rating in the tube manufacturers' data books) but beyond this, increase of the choke inductance produces less ripple in the rectified current; a choke of infinite inductance would produce ripple-free current. The inductance of the choke is dependent on the amount of direct current passing through it and variation of inductance through fluctuations of current is restricted by using a choke having a gap in the core. So-called "constant-inductance" chokes must necessarily have appreciable gaps, but a smaller gap



Figs. 906-a,-b. *Choke-input filters: a) choke in the positive leg; b) choke in the negative leg.*

than would normally be necessary can also be used. The critical value of inductance required decreases as the load current rises, so the characteristics of a choke using a smaller-than-normal gap (one having an appreciable change of inductance with change of current) can be utilized. Such a component is usually called a swinging choke. It can be used only immediately after the rectifier; further chokes in the filtering and smoothing system should have nearer to "contant-inductance" characteristics to insure adequate smoothing for the varying currents resulting from the behavior of the output stage of the amplifier.

In class-AB_2 and class-B amplifiers the conditions of minimum current are such that an unduly high value of inductance would be required, in which case it is customary to add a bleeder resistance to the load to keep the current at a figure large enough to enable a reasonable sized inductor to be used. Even with class-AB_1 am-

plifiers using tetrodes, a bleeder potentiometer may be desirable for the screen supply, not only to permit a lighter rating of swinging choke inductance to be used, but to insure reasonable stability of the screen voltage, the variations in screen current then being negligible as compared with the steady current drawn by the bleeder.

Two-stage filter

The ripple-removing filter is generally of the inductance–capacitance type since heavy current through a resistor would cause a serious voltage drop. Except for the cheapest equipment, this

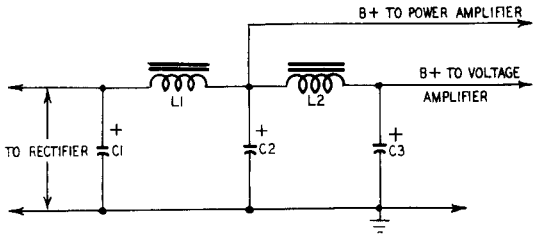


Fig. 907. *Two-stage capacitor-input filter.*

would be two-stage (see Fig. 907). The first capacitor C1 must be rated to withstand the peak voltage of the rectifier and to avoid possible breakdown should be generous, with a test voltage three times the dc working volts. C2 has only small ripple on it and so

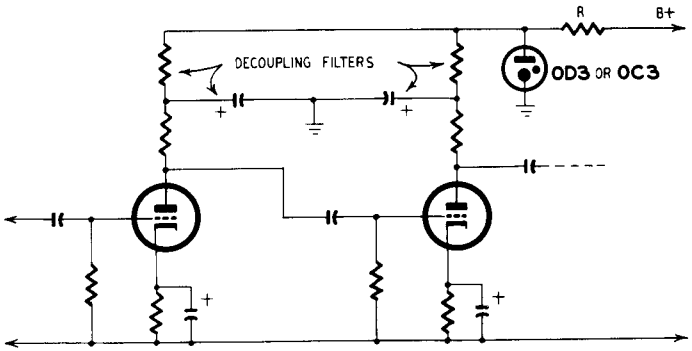


Fig. 908. *Decoupling network using a V-R tube in place of capacitive bypassing.*

the test voltage need be only double the dc working volts (i.e. smoothed plate-supply voltage). If only for economy's sake, to

avoid having two chokes of generous design to carry the whole plate current of the amplifier, the output stage could be fed from the first stage of the filter. It may even be desirable that it should be done this way, for it will be obvious that the impedance of the circuit C1–L1–C2 is common to all plate circuits; if all plate supplies are drawn from the second stage the impedance of the section L2–C3 is added to the common impedance. This may give rise to low-frequency instability in the amplifier, usually prevented by decoupling circuits, so it is desirable to reduce the common impedance as much as possible. As two filter stages are essential for high-gain amplifiers, the best way out is to take the output plate supply from the first section of the filter. Yet hum must be avoided and adequate inductance and capacitance are necessary. Note that there are two resonant circuits in Fig. 907, C1–L1–C2 and C2–L2–C3. These must resonate at a frequency well below the supply frequency to avoid any undesirable resonance effects. Minimum values of inductance are:

For frequency (cps)	25	50	60	100	120
L1 or L2 =	56	14	7	3.5	1.8

$$\times \frac{C1 + C2}{C1 \times C2} \text{ or } \frac{C2 + C3}{C2 \times C3}$$

where C is in μf and L is in henries.

These figures should be exceeded as far as the economics of the design will allow. The inductance obviously is taken at the full current passing through the chokes. As the current through L2 is generally quite low, it is quite feasible to have the inductance high, even as much as 100 henries. It is rarely desirable to have the inductance of L1 less than 15 to 20 henries if the equipment is capable of reproducing very low frequencies. With a single smoothing stage for the output tubes, C2 can very well be as great as 24 or 32 μf. C3 need rarely be more than 8 μf.

However, the audio designer must also consider various other factors such as physical size, price, availability, etc. In preamplifier stages taking low plate current, a third filter stage of the resistance–capacitance type can very well be added as additional smoothing for a sensitive part of the amplifier.

Decoupling filters

From what has already been said, this R–C filter may introduce instability through common impedance in the plate-supply circuit. Also, if the whole amplifier has considerable gain, there may still be insufficient smoothing to eliminate all hum. The amplifier

circuits so far shown have not included decoupling filters, simply on the score of simplicity of illustration. Fig. 908 shows the first two stages of a multistage amplifier with the decoupling filters included in the plates of the tubes. Resistor R in the plate-supply line can be considered as the filter resistor for the third filter section. Normally the tube end of this resistor would have a capacitor connected to ground as the low-impedance ac shunt. If instead of the capacitor a gas tube, such as the 0D3 or 0C3 is used, considerable benefits are secured. The tube acts as a voltage regu-

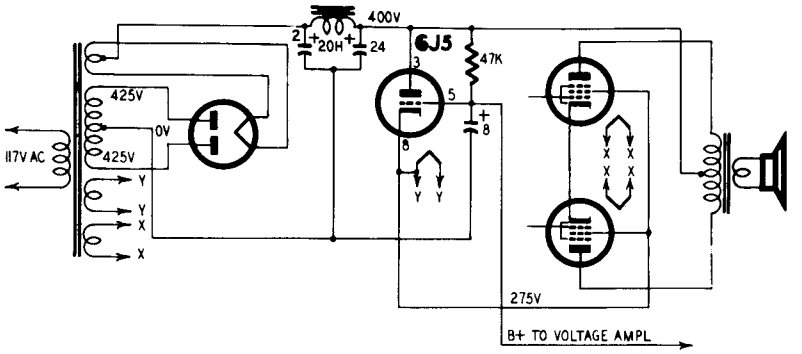


Fig. 909. *Simple form of stabilized supply using a triode.*

lator since gas tubes have the property of maintaining a constant output voltage by varying their internal resistance as the input volts vary; it is therefore an almost perfect shunt to ground for ac, which in itself is a varying voltage. So although the gas tube is primarily a voltage regulator, it has the further merit of acting as a complete hum eliminator and a contributor to stability because of its very low impedance.

As mentioned in the chapter dealing with tetrode output stages, distortion is avoided if the plate and screen voltages can be stabilized. Actually the screen voltage is much more critical than the plate voltage, and a stabilized voltage supply can be obtained by using one (or two in series for voltages beyond the control capabilities of one) gas tube shunted across the supply lines. More refined control can be obtained by using triodes or pentodes in conjunction with gas tubes. It must be said, however, that the writer considers many of the stabilizing circuits somewhat complicated and puts forward a very simple circuit which he has used with great success for many years (see Fig. 909). This involves only a 6J5 tube and a few small components as extras but works ex-

tremely well. And if even this simple arrangement appears too much trouble, it can be taken as axiomatic that no high-fidelity amplifier can give its best unless at least part of the supply from the power unit is properly stabilized. Tubes must work under their optimum conditions, and, if voltages are all over the place as a result of the applied signal causing changes in current through the smoothing filters, precise performance is impossible.

Iron-cored inductor characteristics

There is no technical difficulty in designing a choke, output transformer or power transformer to any given performance, but since the manufacture of these calls for skill and experience it is customary for the audio designer to acquire these components

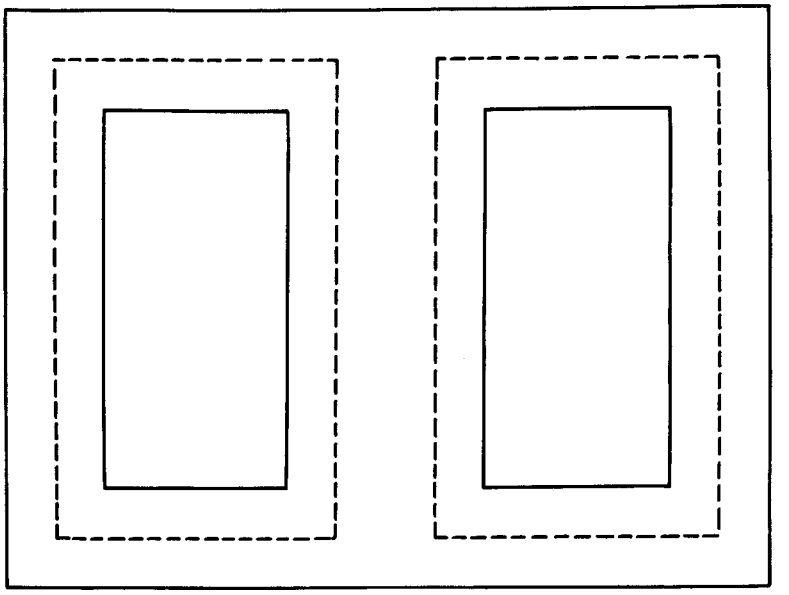


Fig. 910. *Typical transformer or choke construction. The mean magnetic path is shown by the dashed lines.*

from specialist manufacturers. The designer should, however, have some knowledge of the properties and characteristics of these components to ensure a sound choice being made.

The inductance L of an iron-cored coil is given by the formula

$$L = \frac{3.2\,(n)^2 \mu a}{10^8\, l} \text{ henries}$$

where n is the number of turns; μ is the permeability of the core to ac; a is the effective cross-section of the coil in square inches and l is the length of the magnetic circuit in inches (shown as a dashed line in Fig. 910).

The permeability depends on the core material used, which, except for specially prepared molded cores, consists of a stack of iron alloy laminations, insulated on one side to reduce losses. Laminations come in pairs, either as E's and I's or T's and U's. The center section of the core should be twice the width of the

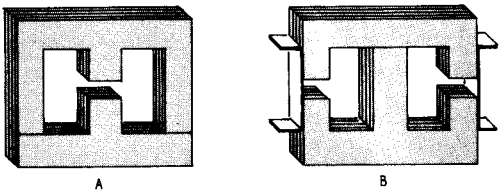


Fig. 911. *Several possible arrangements of laminated-iron cores.*

two side limbs, a seemingly obvious requirement. Some laminations are produced which are badly designed in this respect.

If the choke (or af transformer, which is simply a choke with a secondary winding) carries no dc the laminations are interleaved, making for easier handling, and a push-pull output transformer can have the same arrangement and work satisfactorily if the output tubes are balanced for plate current.

It will be obvious that the size of the core has a bearing on the value of inductance produced by a given number of turns but an even greater effect on the size of wire required to get the same turns into the winding window available. A small core postulates a finer gauge of wire with a substantial increase in dc resistance. When a choke has to carry dc, the voltage drop due to a high resistance winding may not be acceptable. For ac considerations, the core must not be too small if substantial power is handled. This is similar to the case of an output transformer which, apart from any dc present, has to transform the ac power generated by the output stage. See Fig. 911.

For output as well as power transformers, the optimum core cross-section area in square inches A is given by

$$A = \frac{\sqrt{\text{output in volt-amps}}}{5.6}$$

When dc passes through a choke, or af transformer, the ac permeability falls. This could be overcome by increasing the size of the core, but the simplest way of increasing the inductance is to provide a gap in the core. This is achieved by placing a layer of insulating material between the two bunched pairs of the magnetic circuit. The determination of the optimum gap dimensions depends on the amount of dc passing through the winding; for currents up to about 2.5 ma in windings of about 1,000 turns little is gained by having a gapped core, but heavier currents demand it. Typical usage of gapped chokes is in the smoothing chokes of power supply units.

chapter
10

speakers and enclosures

THE audio engineer is not usually concerned with the design of speakers; this is a specialist's job and the designer of sound equipment generally makes his choice from the numerous types and models available. Although it would, therefore, be outside the scope of this book to deal with speaker design, the audio technician requires some guidance in making the best choice.

Published data on speakers are generally entirely misleading, not through any inherent dishonesty on the part of the sponsors of such information, but simply because there is no agreed basis on which measurements should be made. The curves and other information published in advertisements and catalogs is probably quite accurate but can still be meaningless under actual working conditions. For example, it is quite customary to show the response curve of the acoustic output of a speaker with a steady input of 1 watt of audio power. Such a curve is about as useful as the response curve of an amplifier with a resistive load; neither condition is met in actuality—it is a laboratory statement only. Even then, such a curve is of little use unless it can be compared with the curves of other speakers taken under identical conditions of input, loading and acoustic surroundings. Still the comparative curves would have no practical value if nothing is said about the position of the measuring microphone. If it is stated that the microphone is on the axis, then the static response of the speaker can be assumed only for axial reception. The performance of a speaker under static conditions should be shown in a set of polar

radiation curves and these curves can only be compared with another set of polar curves, provided the measurements were all taken in the same measuring room on speakers identically mounted and given precisely the same input.

Limitations of frequency response curves

The accompanying figures illustrate the performance of the author's 215 speaker. They are used, not because of any desire to publicize this particular speaker, but because the data is available and known to be accurate. It is presented as an example of the information required before a speaker's performance can be assessed in the *static* state, and the engineer or music lover who wishes to assess the performance of other speakers is quite reasonably entitled to ask for the same type of data from their manufacturers. Such curves begin to tell the story of what the speaker will do, but it is only the beginning of the story.

Fig. 1001 shows the frequency response of the Hartley 215 speaker on the axis with 1 or 4 volts input. The lower curve shows that the sound output of the speaker is within ± 5 db from 40 to 20,000 cps with a sinusoidal ac input of 1 volt rms. This seems to be a pretty good performance. And it is a fact that this particular speaker, in spite of being a simple straightforward single-unit system, has an unusually wide frequency response, and the curve shown compares very favorably with curves of other speaker systems, even when multi-channel and expensive. But that is not the whole story. Fig. 1001 (upper curve) also shows the axial response with 4 volts input and now it is found that the response between 40 and 20,000 cps is within the limits of ± 12.5 db. The inference is that this particular speaker introduces distortion as the input is increased, and if other speakers did not do so the designer would have something to worry about. Now the possibility of other speakers distorting under these conditions can be demonstrated only by a similar pair of curves, and it is only commercial common sense to emphasize the good features of a product and not say very much about weak points. What is much more difficult to combat is the lack of understanding of nontechnical buyers examining a highly technical proposition, and a speaker *is* a highly technical proposition.

The curves of Fig. 1001 were compiled from measurements made in a properly designed anechoic chamber. The speaker was mounted in the author's nonresonant box baffle, and do therefore truly represent the speaker's performance. Other manufacturers also present frequency response curves of their speakers and, if

they are also taken in anechoic chambers, they can be fairly compared with the curves of Fig. 1001, provided the mounting or enclosure of the speaker has not been designed to remedy shortcomings in the speaker unit itself. It would, of course, be quite fair to have the speaker unit so mounted, but it should be clearly understood that the unit must also be mounted in an exactly similar enclosure in the final listening room if the results predicted by the response curve are to be obtained. Any other form of mounting or enclosure will render the data useless.

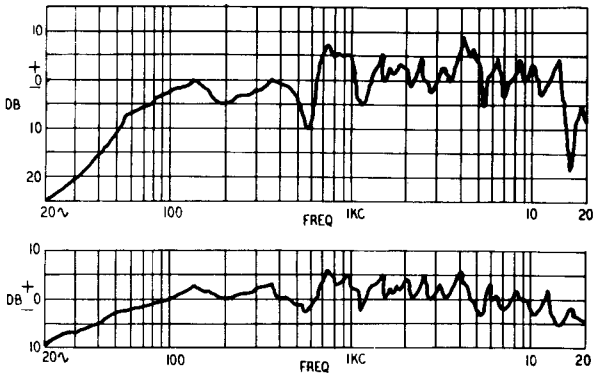


Fig. 1001. *Frequency response curves of a speaker with 4 volts input (upper curve) and with 1 volt input (lower curve).*

These complications, if they may be called such, although they are obviously just applied common sense, are not generally appreciated by a nontechnical user. It is because of this that manufacturers of high-grade speakers are hesitant about publishing response curves of any sort. They can be so easily misunderstood. The writer publishes the lower curve and does not hesitate to do so because it looks good. If he didn't publish any other he would not have offended against a code of business ethics, because the curve is a true statement of fact—so far as it goes. It is the truth, but not the whole truth; yet it is nothing but the truth. But Fig. 1001 is actually a reprint of a figure that has gone into his actual sales literature, and the 4-volt curve was shown, too, in an attempt to tell the whole truth about what his speaker does with varying inputs. Some know-alls at once said that obviously the speaker can't stand high inputs, but what really matters is the question whether it is any worse on high inputs than other speakers. Research and development work over the past year or

so have proved that it is not, but the reason for the disparity has been discovered and proves to be very interesting.

Before dealing with this, however, the matter of polar curves can be disposed of. Fig. 1002 shows polar curves for the 215 speaker, derived from measurements taken in the same anechoic chamber and with the speaker mounted in the same nonresonant box baffle. They depict a plane section of the front hemisphere of projected sound output. But since the curves have been staggered to avoid a continual criss-cross, which would only add confusion to the picture, they should not be taken as absolute with respect to each other. The 8,000-cycle curve is true to scale and conforms to the axial frequency response of Fig. 1001. The circular co-ordinates read from 0° to 180° clockwise, so the 90° ordinate is the horizontal axis of Fig. 1001. Reference to that figure shows where the 90° points of 2,000, 4,000, 13,000 and 20,000 cycles should cut the axial ordinate.

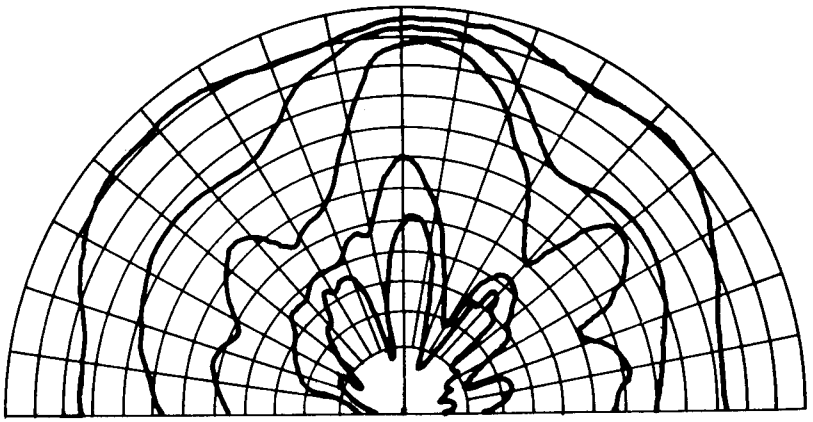


Fig. 1002. *Polar response curves. The normally overlapping curves have been separated. From the outer curve toward the center, the curves represent the response at 2, 4, 8, 13 and 20 kc. Radial ordinates at 10° intervals; circular ordinates at 5 db intervals. The 8 kc curve is in its proper position and the relative displacements of the other curves can be obtained from the lower curve of Fig. 1001.*

The useful information given by Fig. 1002 is that at 2,000 cycles the speaker propagates sound almost equally in all directions. Although the diagram is only a section of the hemisphere, because the speaker is circular, any section on the axis would give the same results. (In practice, room reflections would cause variations, but these curves were taken in an anechoic chamber to give a

picture of what the speaker does under ideal conditions.) At 2,000 cycles the variation in output at any point in front of the speaker is within ±5 db. At 4,000 cycles the variation is within ±12 db and at 8,000 it is within ±16 db. At 13,000 cycles it is within ±17 db and at 20,000 within ±13 db. As might be expected, the response at the highest frequencies falls off rapidly at 90° off the axis. Many speakers are far more directional, which means, in polar-diagram terms, that the curves will have a pear-shaped bunching centered on the axis and with a spread of about 20° on either side. As a matter of simple personal convenience one does not expect to have to locate a group of music listeners in a narrow funnel in the listening room, so it seems essential, for more spatial listening, to demand polar curves and see that they do spread out in some way suggestive of Fig. 1002.

Distortion arising from cone deformation

Figs. 1001 and 1002 give the actual sound output of the speaker as detected by the microphone and measured by the gear attached to the microphone. The microphone detects the sound—the total sound—whether it is distorted or not. Consequently the curves just discussed do not reveal the presence of distortion. It is clear from Fig. 1001 that the speaker illustrated does something when the input is increased which results in a wider variation of response at 4 watts compared with that at 1 watt. Why? The Hartley 215 speaker, whatever its faults, does have an extremely free suspension and can reproduce very low bass without bass resonances cropping up to cause that not unusual one-note thump (and to overcome which the bass-reflex cabinet has become so widely used). Moreover, special steps have been taken to insure that other forms of distortion, such as treble-modulated bass, do not creep in. It can be assumed that the distortion suggested by Fig. 1001 does not arise through faults in the suspension system (and this is proved by the smoothness of both curves below 100 cycles). Therefore, it must come from the cone-coil assembly. Again, in this speaker very great care has been taken to avoid distortion arising through a weak coil assembly or a weak joint between the coil and the cone, as has been done by the designers of all high-grade speakers. A weak coil-cone joint usually gives a well-marked peak between 2,000 and 4,000 cycles (usually at 3,000). This well-known fact is turned to good use in producing cheap speakers for cheap radios which have no real top; the peak at 3,000 gives a synthetic brightness to the reproduction. The sole

remaining component that can introduce distortion is the cone itself.

Fig. 1003 shows three ways in which the cone can be deformed to give distortion and generate spurious sounds. Assume the cone has a free suspension, which means that the periphery of the cone is not restricted by the conventional molded corrugated paper surround. Make up a simple straight-sided cone out of drawing paper. If the mouth of the cone is laid on a flat table and pressure applied to the apex, the cone will be found to be very strong. Squeeze the cone across a diameter of the front opening and it will be very weak. Bend the outer edge between the fingers of two hands and it will again be found to have no strength. Holding the apex in one hand, bend the edge of the cone toward the apex with the other hand, and once more it will be found that the cone is weak. In fact the only condition in which the cone has strength is when the outer edge is resting on the table; that is the condition of perfect loading, when the cone is resisted, as it were, by the counter pressure of the table.

When a speaker is working, the only loading on the cone is that afforded by the air, but as air is an elastic substance the loading is not perfect. The loading, or *damping* as it is usually called, should be as great as possible (hence a good speaker has a good damping factor) not only to reduce deformation as much as possible but also to return to normalcy as quickly as possible any deformation that may occur. Since the damping is not perfect, deformation will occur, and it can occur in three different ways (Fig. 1003). A straight-sided cone will take up shapes as indicated in Figs. 1003-a and b, where the rounded points inside and outside the normal circle are called nodes and those points on the circle antinodes. The number of nodes varies as the frequency of the applied signal.

If the speaker under examination, is driven by an audio oscillator driving an amplifier and viewed by fluorescent lighting, the nodes can be seen very easily. With fluorescent tubes a static picture of the cone can be seen at line frequency and at the second, third, fourth and so on harmonics of the line frequency. With ordinary tungsten lamps the nodes are seen as blur images, and no movement appears at the antinodes other than the back-and-forward movement of the cone itself. The blur images of the nodes are superimposed in the blur image of the whole cone. The stroboscopic effect of the fluorescent tube will be effective only at line frequencies or their harmonics; at other frequencies

the blur image will appear. An exponential cone has a much flatter mouth than a straight-sided cone, and therefore is more rigid across a diameter. But it is weaker axially just because it is flatter. Nodes will therefore appear in an axial direction, as shown in Figs. 1003-c and -d. Finally, wave motion will travel along the cone from apex to mouth as shown in 1003-e.

Loose vs. tight suspension

If the cone as finally located in the complete speaker is not free-edge, then there will be an unavoidable bass resonance, due

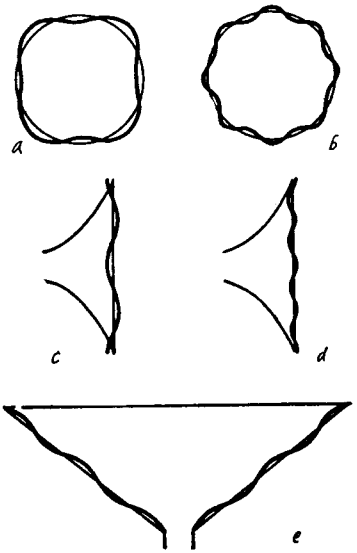


Fig. 1003. *Types of cone deformation causing distortion: a, -b) typical nodes formed in a straight-sided conical diaphragm; c, -d) exponential diaphragms displaying axial nodes; e) wave motion along the surface of the cone.*

to the limited movement possible. But the restraining of the back-and-forward motion is also accompanied by restraint of the noding found with a free-edge cone. It would seem, therefore, that tight suspension may be undesirable from the bass point of view but very desirable from the noding angle. Unfortunately things are not always what they seem, and the mere fact that the cone wants to node and cannot means that the energy thus suppressed at the edge of the cone must be dissipated somewhere

else. It is dissipated by general cone breakup, but as each would-be node involves a sector of the cone the breakup also occurs in sectors. Since the area of the sector is less than the whole, the spurious noise created by cone breakup will have higher frequencies than the noise produced by nodes.

Detailed analysis of speaker behavior cannot be given in the course of a single chapter. The argument can best be summarized by saying that the more irregular response curve of a speaker driven by a high input is due to cone breakup. This is not revealed as distortion by a measuring microphone, but by a more irregular curve resulting from the fundamental frequency sound output being converted into a harmonic series of outputs. Actually, the distortion can be seen on an oscilloscope if the output of a very high-grade microphone is connected to the vertical axis and the audio oscillator to the horizontal. If the oscilloscope is used as a monitor, the input to the speaker can be kept down to the level where no spurious harmonics appear in the sound output. If the response curve of the speaker is taken simultaneously by the usual methods, the overall response may prove to be very disappointing.

Various procedures are often taken to strengthen the body of the cone. A common method is to mold a series of concentric corrugations along the whole cone surface. These unquestionably stiffen it diametrically but can do practically nothing to prevent wave motion of the type illustrated in Fig. 1003-e. There are grounds for believing, too, that suppression of the nodes by this means can result in complementary breakup between adjacent corrugations. Transverse wave motion can be counteracted by stiffeners running from apex to outer edge. It is worth pointing out that as long ago as 1924, long before the days of electrodynamic speakers, a British electromagnetic unit, the Celestion, had a 15-inch paper diaphragm on the front of which was cemented a bamboo spiral running from apex to outer edge, and on the back was a complete set of radial stiffeners, also of bamboo. The problem, therefore, was realized and an attempt made to solve it 34 years ago, because of the generally held belief that the perfect diaphragm must also be a perfect piston; that is, something that does not deform and does not produce spurious harmonics.

Power-handling capacity

There is a direct relationship between the diameter of the cone and its displacement at low frequencies. For a given input the

larger the cone the smaller the displacement (Fig. 1004). Note that the curves marked with cone diameter refer to cones of that diameter and *not* to the overall diameter of the speaker, and also that the flux density in the magnet gap must be constant for all sizes. Obviously the higher the flux density the more sensitive the speaker and the greater the movement for a given input. In comparative measurements of this type only one variable can be considered, and the variable is the cone diameter. Roughly speaking, it will be seen from the chart that a 15-inch cone is displaced only half as much as a 10-inch cone at any frequency, but little is gained by increasing the cone diameter to 18 inches.

Although, therefore, it would seem desirable to make the cone as large as possible, for the sake of power-handling capacity at low frequencies without due displacement of the cone, there are

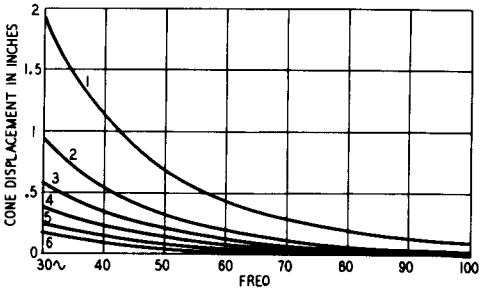


Fig. 1004. *Comparative cone displacement (for fixed input) of paper cones of various diameters: curve 1 for 5-inch, 2 for 8-inch, 3 for 10-inch, 4 for 12-inch, 5 for 15-inch and 6 for 18-inch cone. The curves are valid only for inputs below the power-handling capacity of the speaker.*

serious disadvantages attending the use of a large cone. A large cone must obviously weigh more than a small one; it may weigh considerably more if the thickness of the cone material has to be increased to give the same rigidity as the smaller one. This will result in a loss of high-frequency propagation, but it will also spoil the attack of the speaker, its response to transients, since the inertia of the heavy cone opposes the ideal instantaneous movement necessary to reproduce a steep wave front. Conventionally the cure for lack of top in a large bass speaker is to use a tweeter and dividing network. The latter serves two useful functions—diversion of high-frequency power from the woofer (which would otherwise be wasted) and diversion of low-frequency power from

the tweeter to prevent electromechanical overload of the small-diaphragm speaker. Recourse to a multi-channel speaker system and dividing network does not, however, remove the inertia of the woofer's heavy cone. That can be done only by reducing the weight of the cone. And, as paper is almost invariably used for speaker cones, it means reducing the thickness, reducing the rigidity and increasing breakup with large inputs.

Cone materials

Paper is used because it is cheap, easy to work and lends itself to mass production. It appears to have no other advantages that can be determined by laboratory investigation. Other materials, such as metal and bakelite, have been used from time to time, with varying degrees of success. Metal-cone speakers were introduced very many years ago by two British manufacturers and one still exists used as a doublet, with one unit working into the back of the other. Modern tweeters have small light aluminum cones, as do several horn-loaded speakers in the United States and Britain. But horn loading introduces a different set of working conditions. Electrostatic speakers invariably have aluminum diaphragms, either stretched-flat diaphragms or pleated sound emitters. The distinguishing characteristics of all metal-diaphragm speakers is the coloration, sometimes slight, inseparable from the use of metal (which property is exploited in the brass section of the orchestra). The only bakelite diaphragm speakers that appeared over a long period were the British Hartley-Turner speakers of pre-World War II days. Their manufacture was discontinued in that form because of the intractability of the material.

The one outstanding break from traditional cone material is found in the recently introduced Hartley-Luth 220 speaker, which has a cone of a highly specialized synthetic which has no natural ring, being almost inert acoustically. This material, details of which have not been divulged, has unusual physical properties. It is stated to consist of a core of cellular material which is left in an unimpregnated condition so far as the air cells are concerned, but the bonding material is allowed to penetrate the fibers to form a hard lattice-work frame. On the outer surfaces a complete skin is formed of the very hard bonding material. A section of the cone material somewhat resembles a lattice-work girder—two planes joined by a criss-cross of stiff struts. This form of girder is used in engineering structures to save weight without sacrificing stiffness, and the application of the principle to an

acoustic diaphragm produces some interesting results. Cones made in this way are little heavier than paper of similar thickness but mechanically about four times stronger. To this extent they more nearly approach the ideal of a perfect piston, and the response does not vary with increased input (see Fig. 1001). The curves of Fig. 1004 do not apply since it is found that the rigidity of the cones results in greater resistance from the air. In other words, the air loading is higher and the output rises very considerably at low frequencies, due entirely to the absence of noding and cone breakup.

The Hartley 215 speaker was often criticized because of its seeming lack of bass. In part this was due to the absence of a noticeable bass resonance, but mainly due to the 9-inch diaphragm not moving enough air at the lowest frequencies. This objection could easily be overcome by using two units side by side, connected in simple series, thus increasng the diaphragm area by 100%. It is generally recognized that better bass is usually obtained by using several small speakers rather than one large one of equivalent diaphragm area. In the development stages of the Hartley-Luth 220 it was thought that a simple improvement might result from using two units to give still more bass, but actual experiment proved that the increase was barely noticeable. The unavoidable explanation could only be that the paper cone of the 215 did not act as a true rigid piston at low frequencies with large inputs and so lost some output, whereas the rigid cone of the 220 converted the electrical input into mechanical pressure on the air without loss. In specific terms, direct comparison between the 215 and the 220 revealed that for fixed input the much more rigid cone of the 220 extended the output downward by more than an octave. This can be expressed in another way by saying that the output of the 220 at 25 cycles was somewhat greater than that of the 215 at 50 cycles, and the higher frequencies up to about 200 cycles were equally improved. The curves of Fig. 1004, therefore, which are derived from data originally compiled by Massa, apply only to paper cones. The 9-inch synthetic cone of the 220 has a greater bass output than a 15-inch paper cone, especially when driven hard. The work involved in preparing this new type of speaker seems to prove beyond argument that the first requirement of a speaker cone is that it should not deform when actuated by the voice coil.

Visual appraisal of a speaker

An electrodynamic speaker can be fairly well appraised just by looking at it and feeling it. If the cone has a rather narrow angle,

it will be directional and have a poor polar diagram: the flatter the cone the wider the dispersion of sound. If the cone gives when pushed in one place with a finger, as is almost inevitable with paper cones, then the cone will distort with high inputs, giving uneven response and introducing harmonics either from nodes or suppressed nodes. The molded surround is usually left thinner than the main cone body for the sake of greater flexibility. But, if the projecting corrugation feels harder than the two-inward pointing corrugations, then the hard corrugation will resonate independently of the main body of the cone at a frequency dependent on the size of the cone and the lack of stiffness of the inward corrugations. Examination of the cone and surround under a strobe light, with the speaker driven by an oscillator and amplifier, will reveal at which frequency this occurs.

The outer and inner surfaces of the cone can be grasped between the thumb and finger of both hands and the whole assembly moved backward and forward. Measurement of the total possible displacement, which, of course, is limited by the design of the outer surround and the spider washer, will permit estimating the power handling capacity by reference to Fig. 1004 if the cone is made of paper. It must be understood, however, that the Massa data only applies to the power-handling capacity of the speaker. It reveals that for a given limit of excursion a speaker will handle more power as the cone size is increased, but the lower frequency limit is determined by the design of the suspension. If the bass resonant frequency of the speaker is, say, 50 cycles, then it will have a vastly increased output at that frequency and practically no output at a fundamental frequency below that. The output will consist mainly of third harmonic, a little second and no measurable fundamental. This is usually called frequency doubling, but it is more correct to call it frequency tripling. The Massa figures would therefore apply only to frequencies above 50 cycles *for that particular design of suspension.* If the cone is tapped with the finger tip, a sound will be heard resulting from the contact between the finger and the cone. But, with the ear in the cone, a low note will also be heard, and its frequency is the bass resonant frequency of the suspension system.

There are various ways of extending the seeming response of a single-unit speaker having a paper cone, such as adding a subsidiary cone, introducing compliances of one form or another to various parts of the cone or voice coil. These do what they are said to do, but whether the end product is entirely free from

distortion is a matter of some doubt. Detailed discussion of these various devices goes beyond the scope of this chapter.

Speaker impedance

The impedance of a speaker varies with frequency, as shown in Fig. 1005. The peak in the lower register occurs at the resonant frequency of the suspended system. Provided the cone and voice coil are moving freely, this peak is not unduly prominent and the curve is then said to represent the free impedance of the speaker.

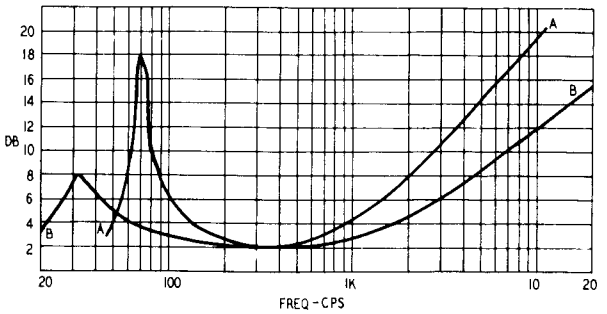


Fig. 1005. *Impedance curves of dynamic speakers: AA) curve of a conventional design with cone surround resonance at 70 cps; BB) curve of speaker with free-edge cone. Voice coil, cone compliance and rear suspension (spider washer) resonance of 32 cps.*

If the vibrating system is prevented from moving, the resultant curve represents the blocked impedance of the speaker. At middle and high frequencies there is little difference between the blocked and free impedances, but at the resonant frequency the difference is considerable. Inadvertent blocking of the movement of the cone can be caused by tightness of the cone surround and the rear suspension washer or spider. The usual molded corrugations in the outer periphery of the cone do not give much freedom of movement and as a consequence the spider is often of a design that would give little freedom if the cone were free-edge. Adoption of a free-edge cone removes one source of blocked movement but this is not enough if the spider washer does not also give free movement of the suspended system.

Moving any speaker diaphragm by hand demonstrates that the farther the cone is moved from the mean position the harder it becomes to move until the point of maximum excursion is reached when it will not move at all. This increasing resistance is depicted by the peak in the impedance curve. The height of the

peak is reduced by free suspension, although introducing free suspension also introduces a whole host of other difficulties, such as modulation of the treble by the bass, and the effect known as electro-mechanical rectification.

The rise in the treble is mainly due to the inductance of the voice coil and will naturally be greater with higher inductance. It is conventional to build speakers with a higher impedance rating than is necessary or desirable and, like many conventions, it is not based on what might be called absolute thinking. For the standard of reproduction demanded of inexpensive mass-produced equipment it is convenient to be able to buy speakers of standardized impedances. But these standards were adopted long before the desire for high-fidelity reproduction became widespread. More or less automatically, low-fidelity standards were carried over into the high-fidelity speaker industry with, at least in the writer's opinion, unfortunate results.

The impedance curve of any speaker can be taken with the usual impedance-bridge technique and for comparison purposes speakers should be mounted on a standard flat and inflexible baffle in the open air. Sound reflections from the walls of a measuring room can cause blocking of the free movement of the cone.

However, there are simpler methods well within the capabilities of the worker not equipped with full laboratory facilities. These are shown in Fig. 1006. In all cases an audio oscillator with a constant sinusoidal output is required. The first method shown requires an ac ammeter and ac voltmeter; these are connected as in Fig. 1006-a. As the input is varied from the lowest to the highest frequencies, note is taken of the current through the speaker and the voltage drop across it. Call these, respectively, I and E. The impedance Z is then equal to E/I. The readings can be taken at 10-cycle intervals from 10 to 100 cycles; then at 100-cycle intervals to 1,000 cycles, and finally at 1000-cycle intervals to 20,000 cycles. In particular, note carefully when E rises steeply at the bass resonant frequency and note the frequency. If there are other peaks, they should also be noted. The final result when plotted as a curve will not be exactly like Fig. 1005, which shows smoothed curves of trends, but will have a number of small peaks which represent resonances in various parts of the speaker, even of the cone basket.

Ac ammeters are not always available, so two voltmeters can be used, as in Fig. 1006-b. The method is simply a comparison of voltage drops across two resistors. In the figure, resistor R should

have the same dc resistance as the speaker voice coil. If the meters are dc units and dc passes through the speaker and resistor R, the voltage drop will be the same across each. When ac is applied, the impedance of the pure resistance R is constant at all frequencies but the speaker has reactance so its impedance will vary with frequency. Apply various frequencies as indicated earlier and note the readings on the two meters. The impedance is then given by the formula

$$Z_s = R\ (E_s/E_r)$$

These two methods have the disadvantage that meters have to be observed and a certain amount of simple calculating has to be done. The continuously visual method shown in Fig. 1006-c in-

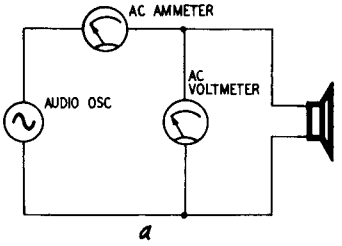


Fig. 1006-a. *Determining speaker impedance by using an ammeter and voltmeter.*

volves the use of an oscilloscope. Resistor R is, as before, a non-inductive resistor having the same dc resistance as the speaker

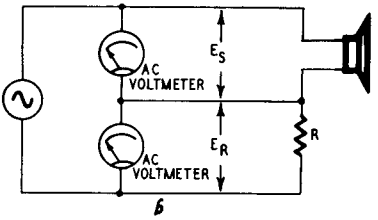


Fig. 1006-b. *Speaker impedance can be measured with the use of two ac voltmeters.*

voice coil. Resistor R_s, shown dotted, is a precisely similar component but it is not in use all the time. To set up the circuit, connect R and R_s as shown but omit the speaker. Isolate the oscil-

loscope circuit from the oscillator by a good-quality audio transformer. Apply a signal and adjust the oscilloscope amplifier's sensitivity controls until the trace is a straight line inclined at 45° to the two axes. (The internal time base of course, is not used.) Now take out R_s and connect the speaker.

If and when the speaker acts as a pure resistance, the trace will be a straight line. Inductive and capacitive components in the speaker will cause the line to broaden into a narrow ellipse; the major axis of this ellipse is the indicator of what follows. As the frequency is varied so the impedance of the speaker varies and this causes a change in tilt of the trace. For any point on the trace

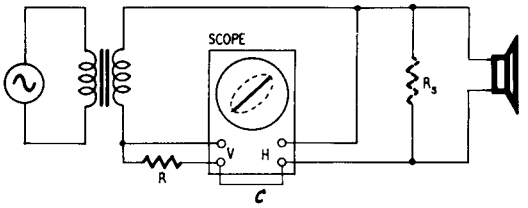


Fig. 1006-c. *Technique for measuring speaker impedance using an oscilloscope.*

or the major axis of the trace, the respective magnitudes of the speaker impedance and the resistance of R will be indicated by the coordinates on the horizontal and vertical axes of the oscilloscope graticule. The method is very sensitive and placing the hand in front of the speaker is enough to show the change of impedance. It is the most convenient way of observing the effect of different types of enclosures or modifications of enclosures as well as studying the effect of altering room furnishings and damping.

The impedance for which output transformer calculations are made is taken as that at 400 cycles. The impedance curve, as well as the frequency response curve, will vary with input because the air or mechanical loading of the speaker will change with change of input. Published equivalent circuits of a loudspeaker do not normally take account of this complication, so a mathematical examination of a speaker based on an equivalent circuit must be accepted with caution.

Since the impedance varies so widely with frequency, it will be obvious that the loading on the output stage is far from constant. And if the output transformer has been designed on the basis of the nominal impedance of the speaker, mis-matching will be considerable at the bass end and also in the treble. It may, therefore,

be desirable to select a transformer ratio which provides a mean value of load at bass, mid and high frequencies. This mismatching is not so important with triodes as with tetrodes or pentodes (which latter can be made more accommodating with negative feedback), but a transformer ratio correct for the impedance at 400 cycles, apart from causing distortion in the bass through overload, is the cause of what is usually known as "pentode quality" in the highs. It is not the pentode which causes distortion; it is simply that the harmonic distortion caused by incorrect loading is particularly noticeable at high frequencies.

The baffle and speaker impedance

The cure for these troubles is, of course, to use a speaker with as nearly level an impedance as possible; but as many popular speakers do not have this desirable property, various devices are

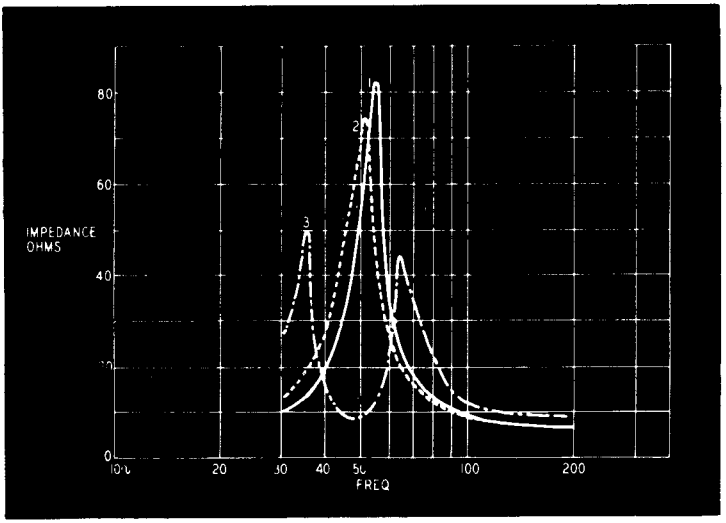


Fig. 1007. *Effect of baffles on speaker impedance: 1) 12-inch speaker in free air; 2) mounted on a flat baffle; 3) in a ported enclosure.*

adopted to overcome the speaker's defects. With a view to neutralizing the bass resonance, an acoustical phase inverter, popularly known as a bass-reflex cabinet, is widely used. This device was first described by A. L. Thuras in U. S. patent 1869178, (1932), and became generally available when the patent expired. It is generally assumed that it will improve the bass response and increase the power-handling capacity of the speaker, increase the acoustical damping at low frequencies and decrease the impedance

at the bass resonant frequency while reducing the amplitude of the cone movement at that frequency. These claims may generally be accepted in an engineering sense, but many listeners of good musical education are convinced that the final effect is that of a strongly resonating system. Certainly the housing itself is a resonator for it is designed to resonate at a frequency determined by the characteristics of the speaker, but it is a matter of esthetics to decide whether the final results are free from distortion. (See Fig. 1007.)

Fig. 1008 gives a cross-section of a typical bass-reflex housing which must be constructed in a rigid manner and lined with sound-absorbing material to cut down internal reflections. The upper opening will be circular with a diameter equal to that of the speaker cone; the port or vent below must have an area equal to the area of the speaker opening. The volume of the housing assumes no absorbent lining, but the volume of the speaker itself must be allowed for. For what it is worth, the expression for the volume of the cabinet in cubic inches is:

$$\pi r^2 \left(\frac{1.84 \times 10^8}{\omega^2} \times \frac{1}{1.7r + l} + l \right)$$

were r is the radius of the speaker cone in inches; l, the length of tunnel in inches; $\omega = 2\pi \times$ frequency of vent resonance.

Fig. 1009 gives the external dimensions of a typical bass-reflex enclosure.

Summary of design features for baffle-loaded speakers

1. A single speaker capable of reproducing the freqency range needed for high-fidelity reproduction is possible but requires great skill to design.

2. A complex wave is reproduced by a single diaphragm because of cone breakup. For best results this breakup must be carefully controlled. The apex of the cone reproduces the highs; the entire cone moves as a piston to reproduce the bass.

3. The harder the cone material, the better the treble and bass response—but this interferes with cone breakup. Concentric ridges in the cone do not appear to affect the breakup, and improve the bass response by stiffening the cone radially and so counteract the development of nodes.

4. Large cones node radially more easily than do small ones; exponential cones node axially. Nodes reduce output at the frequencies where they occur and introduce undesirable harmonics.

5. The cone material should be acoustically inert. Metal diaphragms produce a characteristic *ringing* coloration.

6. Narrow-angle cones cause excessive focusing of the high frequencies. Wide-angle cones give better diffusion but tend to node more easily.

7. Large cones are too massive for good treble response and transient reproduction. Their size can cause phase distortion of the low-frequency components of a heavy transient.

8. Small cones cannot reproduce low frequencies with sufficient output unless the suspension is free enough to enable them to move the same amount of air as a large one.

9. Subsidiary tweeter cones give more treble, but the undamped outer rim can cause feathery reproduction if driven hard. The narrow angle of the tweeter cone also causes excessive focusing of the highs.

10. Large straight-sided cones cause the enclosed air to resonate at its own natural frequency, causing a superimposed hoot. Exponential and small cones do not suffer audibly from this defect.

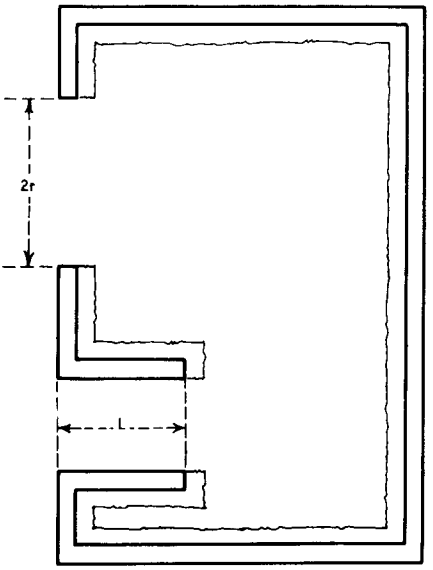


Fig. 1008. *Cross-section of a typical bass-reflex enclosure. The jagged line represents the acoustic lining material.*

11. Treble response can be improved by specially designed voice coils, where the diaphragm is relieved of the weight of the coil at high frequencies.

12. The voice coil should be as light and as rigid as possible, with only two layers of wire. The wire should be double-silk-in-

sulated to provide a good key for the varnish. A 1-inch coil is more rigid than a 2-inch but a 2-inch pole piece allows for a greater flux density in the gap.

13. A tight suspension gives a high bass-resonant frequency, and frequencies below this will be reproduced mainly as third harmonic. Free suspension can introduce intermodulation distortion and nonlinear distortion.

14. To avoid intermodulation distortion the voice coil must cut constant flux at all points in its excursion. The coil must be substantially longer than the gap.

15. To avoid nonlinear suspension distortion the magnetic field must be symmetrical about the gap. If the center pole is shouldered, a flat front magnet plate cannot give a symmetrical field.

Dividing networks

Except in highly specialized designs such as the author's own speakers, it is not possible to achieve a wide frequency response without major irregularities in a single-speaker unit. The design of a speaker to give a good output at low frequencies is a flat contradiction of that required for high frequencies and most single units are a compromise design between these two extremes.

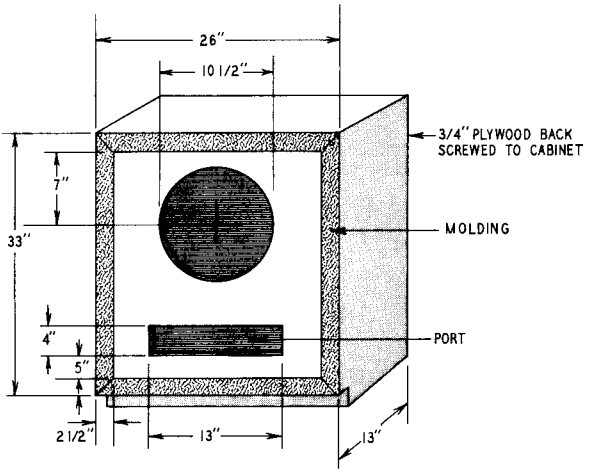


Fig. 1009. *External view of a typical bass-reflex enclosure.*

There is no fundamental difference between the design of a dynamic tweeter and a woofer: one has a small very light diaphragm; the other a large, reasonably inflexible one which cannot

reproduce satisfactorily above about 3,000 cycles. The power-handling capacity of a speaker for low frequencies is a function of the size of the diaphragm (since adequate bass reproduction can be obtained only if enough air is physically moved) so the small cone of a tweeter makes such a speaker have no power-handling capabilities below about 300 cycles. Using woofer-tweeter combinations means, therefore, that the bass must be kept out of the tweeter and the treble kept out of the woofer. This is done by using suitable high- and low-pass filters.

Combinations of such filters are called dividing or crossover networks. The former term is the better one since the purpose of the network is to divide the output of the amplifier into two parts, the treble end and the bass end. The middle frequencies are reproduced by the top part of the response of the low-pass filter, the bottom part of the response of the high-pass filter and a nonuniform portion of the combined responses centered on the crossover

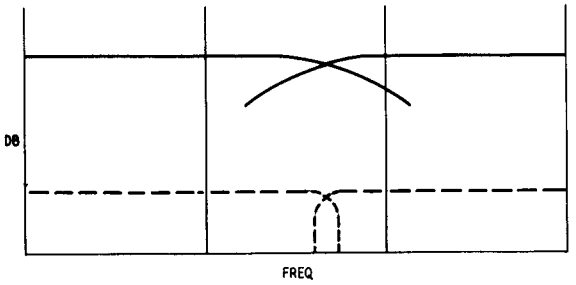


Fig. 1010. *Response curve of a typical dividing network.*

frequency, as shown by the solid curves in Fig. 1010. Experience has shown that the most satisfactory results are obtained with an attenuation of 12 db an octave beyond where attenuation begins, in spite of the fact that there is a definite dip in the response of the network at the crossover frequency. It might be thought that if the two filters had sharp cutoffs, as shown by the broken curves, the greater uniformity would be an advantage, but to get such an abrupt attenuation calls for filters having many stages. A more serious disadvantage is that a sudden switch from woofer to tweeter, as would happen with filters of such characteristics, is audibly distressing.

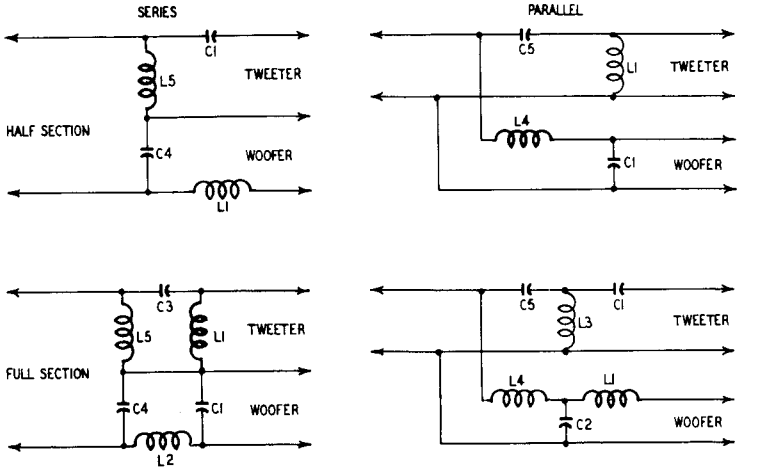
The two speakers cannot have precisely the same characteristics. If, as in the case of wide-range reproduction, two speakers of differing characteristics have to be used, it is better that the merging from one into the other be done gradually; but it must not be done

too gradually, otherwise the filtering of each channel will not be sufficiently effective.

Single-stage resistance-capacitance filters, as mentioned in an earlier chapter, give an attenuation of 6 db per octave; but dividing networks for loudspeakers usually require a steeper slope than this, so inductors are used instead of resistors. Such L/C filters give slopes of 12 db and 18 db per octave respectively for half section and full section types. The points at which attenuation begins will obviously be farther apart with 12 db filters than with 18 db if a constant level at crossover frequency is to be observed. In practice, it is found that the dip at crossover frequency will be about 3.5 db. Fig. 1011 gives the circuits of dividing networks of both series and parallel types.

Constant-resistance network

Another type of dividing network, details of which are given in Fig. 1012, is the so-called constant-resistance network. It will be



$L1 = R/\omega$ $L2 = 2R/\omega$ $L3 = R/2\omega$ $L4 = 1.6R/\omega$ $L5 = 0.63R/\omega$
$C1 = 1/\omega R$ $C2 = 2/\omega R$ $C3 = 1/2\omega R$ $C4 = 1.6/\omega R$ $C5 = 0.63/\omega R$

Where L is in henries, C is in farads. R = voice coil resistance.
$\omega = 2\pi \times$ crossover frequency.

Fig. 1011. *Dividing network circuits (series and parallel types).*

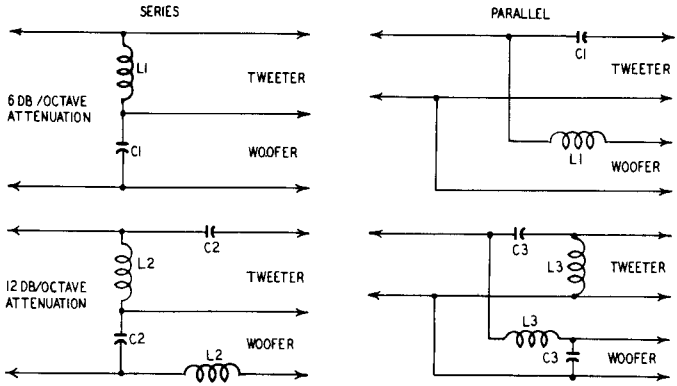
understood that the load on the output stage should be constant for all frequencies and we know that the impedance of a single-unit speaker varies considerably with frequency. Adoption of a

constant-resistance network does *not* mean that circuit gymnastics have produced a two-way speaker system which presents a constant impedance at all frequencies, for the constant-resistance network has a constant input resistance only for purely resistive loads on the output terminals of the network. The problem of speaker impedance varying with frequency remains as before.

The advantage of the constant-resistance network is that the capacitors and inductors can have more convenient values throughout, as can the inductors. This simplifies design and manufacture and accounts for the commercial popularity of this type of network.

Horn-loaded speakers

Horn-loaded speakers are a different technical proposition; the prime object of fitting a horn to a speaker unit is to increase its electroacoustic efficiency. A horn speaker does not perform fundamentally better or worse than a baffle-loaded speaker, it is merely more efficient. But to secure adequate bass reproduction calls for



$$L1 = R/\omega \qquad L2 = R/\sqrt{2}\omega \qquad L3 = \sqrt{2}R/\omega$$
$$C1 = 1/\omega R \qquad C2 = \sqrt{2}/\omega R \qquad C3 = 1/\sqrt{2}\omega R$$

Where L is in henries, C is in farads. R = voice coil resistance.
$\omega = 2\pi \times$ crossover frequency.

Fig. 1012. *Constant-resistance dividing networks (series and parallel types).*

as inconveniently bulky a horn as with baffle speakers. To reproduce a 50-cycle note with negligible loss calls for a baffle 12 feet square or a horn 22 feet long with a flare circumference of 24 feet. Neither of these is conveniently accommodated in an ordinary living room so designers have had recourse to folding horns to make them more compact. A horn, folded or straight, must follow

certain laws, otherwise the advantages of horn loading are lost. Some cabinet type folded horns are virtually acoustical labyrinths and may have undesirable acoustic effects on the reproduction.

Provided the length of the horn can be accommodated in the auditorium (and in an exponential horn the first few feet are related to the high-frequency part of the spectrum), there is no point in having multiple-unit reproducers. But since the problem is always that of squeezing the unit into a reasonable space, the horn assembly can be divided into a straight tweeter and a folded woofer, with each section driven by an appropriate speaker unit. In any event, a horn-loaded speaker is very directional at high frequencies (indeed is very much more directional than a direct radiator or baffle-loaded speaker). As it is desirable to separate the highs, a multicellular tweeter horn or acoustic-lens diffuser can be used to advantage.

The limitation of a horn speaker in the bass is the flare cutoff frequency, and it is important that this should be higher than the natural resonant frequency of the speaker unit; indeed it is important that the amplifier itself should have a bass cutoff equal to the flare cutoff, otherwise there will be an appreciable tendency for the speaker to develop harmonics of the frequencies the horn cannot reproduce. If low frequencies are desired and a large horn is not, the solution appears to be that of P. W. Klipsch, who has developed a design wherein an air chamber behind the speaker unit, resonating at about the flare cutoff frequency, provides a "capacitive" reactance about equal to the "inductive" reactance of the horn at cutoff frequency. In this case frequencies lower than that of the flare cutoff frequency can be reproduced without appreciable distortion because this chamber takes control.

For a straight horn the cross-section can be circular or square; the flare must increase according to the exponential law selected and sudden changes of curvature are detrimental. For a folded horn the same rule holds, so for best results the horn cannot be built up out of flat pieces of material; yet it nearly always is. The cabinet type of folded horn is, therefore, some way short of a perfect horn even if the increase of cross-section area is carefully maintained (which it must be). As any speaker is already a far-from-perfect transducer, the purist may consider it undesirable to introduce further distortion through using an imperfect horn and may prefer to generate more audio watts in the amplifier to make up for the lower efficiency of the direct radiator. A "folded" baffle does not require the same precision of design as a folded horn.

chapter 11

measurements and testing

TECHNICAL workers in electronics frequently use meters without stopping to think what the meters are and what they are supposed to do. It is quite true, of course, that as the purpose of a meter is to give an indication of circuit conditions, mere observation of the meter reading ought to suffice. For example, it may be argued that a car driver need not know how his speedometer works since all he wants to know is the speed of the car, but it doesn't concern him very much whether the speedometer shows 50 or 55 mph, for the instrument may be inaccurate, the drive gearing may be wrong for the size of tires used and the tires themselves may not be inflated to the correct pressure. In electronics measurements, one has to be more careful about the accuracy of readings. Sometimes the reading may be wrong due to an inaccurate meter and sometimes the wrong answer will be given by an accurate meter because it is improperly used. Such unhappy cases are avoided by having some knowledge of meters and other indicators and how they work.

Meters for voltage, current and resistance

Voltmeters, ammeters and ohmmeters are all ammeters because they measure current; their scales may be calibrated in volts or ohms but they are still ammeters. These remarks apply, of course, to the usual moving-coil meters on every service technician's bench. There are other meters which are not ammeters, such as electrostatic voltmeters for measuring high voltages; these are pure voltmeters.

Moving coil meter

Current passing through the pivoted coil of the meter (to which is attached the pointer) creates a magnetic field which is acted on by the magnetic field of the permanent magnet built into the meter. The meter, in fact, is an electric motor but without a commutator because continuous rotation is not required. The armature is restrained by the springs controlling the movement (these springs return the pointer to zero in the absence of an applied current) and the angular displacement of the armature is proportional to the current passing through its coil. Such a meter will give a readable deflection on dc only.

To use the meter as a voltmeter a fixed resistance R is connected in series with the coil, as shown in Fig. 1101-a. Application of a voltage to the two components causes a current to pass through the circuit of a value:

$$I = \frac{V}{R + R_a}$$

where R_a is the internal resistance of the meter. As R is increased so the current through R_a falls—hence it can be considered a multiplying resistance. If various values of R are selected by a switch, the meter circuit becomes a multirange voltmeter.

In Fig. 1101-b, absence of shunt resistor R would cause the whole current to pass through the meter, which is now acting as a straight ammeter. The internal resistance of the meter will limit the current passing through it and the resistor. If the current is greater than the meter can show, then the meter is shunted by R, which diverts a proportion of the current. Strictly speaking, R is still a multiplying resistance since it permits the meter to register a higher current than is possible in its absence, but it is usually called a shunt.

Ac meters

For measuring ac voltage and current there are several types of meters, but it is now customary to use a dc moving-coil meter in conjunction with a copper oxide rectifier, as in Fig. 1101-c. The current-carrying capacity of the meter is determined by its full-scale deflection. But the rectifier also has a maximum current capacity, and if the resistance of the meter is too high (or a fuse in the meter is blown so that the meter is open-circuit) the rectifier will be destroyed, owing to the full circuit voltage being applied to it. Voltage multipliers can be inserted in the rectifier circuit at

R, but increased current ranges can only be provided by using an input current transformer, shown dotted, if a linear scale is to be maintained.

If the dc scale of the meter is used, the current registered will be the mean value ac current passed. But what is wanted is the rms value, which is higher than the mean value in the ratio of

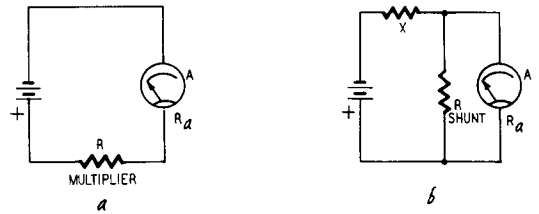


Fig. 1101-a, -b. *To use the ammeter as a voltmeter (left) a series resistor is inserted. The meter range is extended (right) through the use of a shunt resistor.*

1.11; a 100 μa-scale deflection will represent an rms current of 111 μa. The scale will be linear for ac as well as dc except for very small currents, when the rectified alternating (direct) current tends to be proportional to the square of the alternating current. This

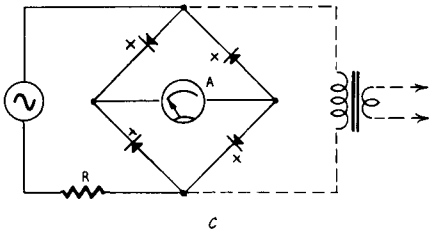


Fig. 1101-c. *Dc meter, in conjunction with rectifiers, can be used to measure ac voltage and current.*

causes a closing up of the scale at the low end. This is why the rectifier cannot be expected to give linear rectification if it is shunted by resistance for the measurement of heavy currents, which explains the use of a current transformer for this purpose. Similarly, linear scales for low ac voltages can be secured only by using an input stepup voltage transformer.

Given a high-grade rectifier, ac meters of this type can be depended on for good accuracy with sinusoidal inputs up to 100,000 cps. Thereafter the response will fall off until at 1 mc the read-

ing will be 20% low. For audio measurements, therefore, the meter can be taken as accurate, provided the current measured is sinusoidal. In-phase second-harmonic distortion will give a reading which is low up to about 2%; in-phase third harmonic will give a meter error up to +5%, but third harmonic 180° out of phase will give a reading up to −9% low. Rectifier instruments should not be used for measurement of distorted ac waveforms.

Measuring resistance

By Ohm's Law we know that R (ohms) = E (volts)/I (amperes). This applies also to ac provided it is remembered that the resistance may consist of noninductive, noncapacitive resistance plus

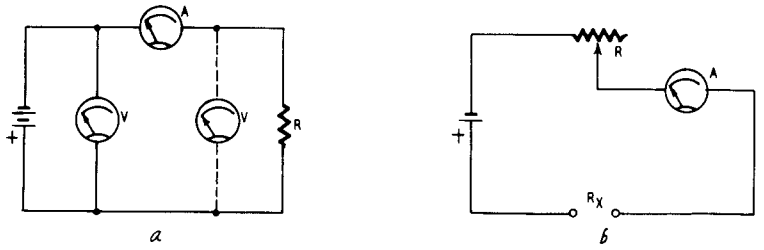


Fig. 1102-a, -b. *Measuring resistance by voltmeter-ammeter method (left) or by using an ohmmeter (right).*

the impedances of any inductances and capacitances in the circuit, the values of which depend on the frequency of the ac. So far as dc is concerned, it is necessary only to measure the voltage applied across a resistor and divide by the current passing through it to get the value of the resistance of such a circuit as is shown in Fig. 1102-a, but a word of warning is necessary. If the current passing through the voltmeter is appreciable as compared with that passing through R, then the reading will not be accurate. Therefore, the voltmeter must be connected on the battery side of the ammeter. If, however, the resistance of the ammeter is appreciable as compared with R, then the voltmeter must be connected as shown dotted, for it is necessary to know the voltage actually applied to R.

On this simple basis an ohmmeter can be constructed. The battery voltage can be fixed (and is invariably connected inside the instrument case), but as the battery will deteriorate with age, some compensation must be provided for this. In the circuit of Fig. 1102-b, if the terminals R_x are short-circuited, the only resistance in the circuit is R plus the internal resistance of the

meter. Accordingly R can be adjusted to give full-scale deflection and this point will indicate zero ohms across terminals R_x. If the short is removed and a resistor is connected across the terminals, the current passing through the meter will fall and the higher the resistance, the smaller will be the deflection.

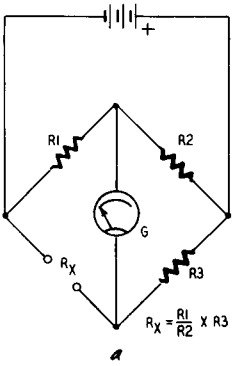


Fig. 1103-a. *Wheatstone bridge for measurement of dc resistances.*

A resistance scale can be drawn on the meter which will read from right to left, the reverse of current and voltage readings, and

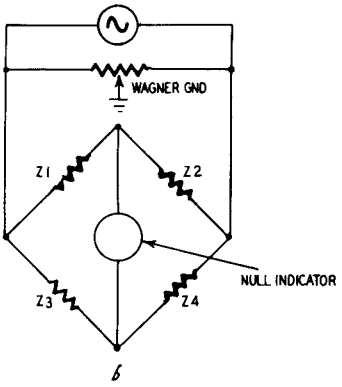


Fig. 1103-b. *Ac bridge.*

the scale will also become increasingly crowded toward the left. The useful range can be extended by having several resistance ranges, and the multiplier this time requires a greater voltage

from the battery, so that a measurable current can be passed through a higher value of resistance.

Naturally it is desirable that the scales should have a simple multiplying factor, say 10 times, so that only one resistance scale need be used. This calls for quite careful design of the complete meter. On any range the preliminary routine must be to short the resistance terminals and adjust zero (full-scale deflection) before making any resistance measurements.

Accurate measurement of current and voltage is merely a matter of selecting a suitably accurate meter, but there is no such thing as a highly accurate ohmmeter because of the bad scale characteristics. Accurate measurement of resistance is best done on a bridge.

Bridges

The circuit of Fig. 1103-a shows the Wheatstone bridge for measurement of dc resistances. R1 and R2 are usually banks of precision resistors with values of 1, 10, 100 and 1,000 ohms; R3 is a decade resistance box with a minimum value of 1 and a maximum value of, usually, 11,111 ohms. Galvanometer G is used as an indicator of balance. If the ratio of the unknown resistance R_x to the variable resistance R3 is the same as the ratio of R1 and R2 the bridge will balance. Therefore, if R1 and R2 have the same value, the value of the unknown resistance will be found by determining what value of R3 is needed for balance. If R_x is outside the limits of value of R3, balance can be found by changing the ratio of R1 to R2. For all cases:

$$R_x = \frac{R1}{R2} \times R3$$

The limitations of the Wheatstone type of bridge are found in measuring low resistances, because of the resistance of the leads and contacts, and in very high resistances, because of the very high R1/R2 ratio required, and the fact that no galvanometer has a sufficiently high resistance to match the bridge resistance. For such cases a dc vacuum-tube voltmeter is a better indicator and, of course, the battery voltage can be raised. Low resistances are best measured on a Kelvin bridge.

An ac bridge has basically the same form, as shown in Fig. 1103-b, except that an oscillator is used as the energizing source and a null indicator (headphones, c-r tube or ac vacuum-tube voltmeter) for balance. Here, as before, the bridge is balanced

when $Z1/Z2 = Z3/Z4$, but attention has to be paid to the nature of the impedances. If Z1 and Z2 are purely resistive, as they often are, then Z3 and Z4 must both be identical kinds of impedance: if Z3 is inductive, Z4 must be inductive; if Z3 is capacitive, Z4 must be capacitive. A complete bridge requires, therefore, standard resistors, capacitors and inductors to be used for Z4 where resistance, capacitance and inductance have to be measured. The other extra feature is the potentiometer across the oscillator labeled "Wagner ground." Both oscillator terminations have capacitance to ground which may make balance impossible; the Wagner ground enables a balance to be secured without interfering with the operation of the bridge.

There are many types of ac bridge, each having been devised for some special application. Standard inductors, for example, are costly and cumbersome and can be avoided by using a Maxwell bridge. This compares unknown inductors with standard capacitors using resistors for Z1 and Z4, with capacitance and resistance in Z2. The Wien bridge measures capacitance in terms of resistance and frequency or frequency in terms of the other quantities. Details of these and other specialized types can be found in any standard textbook on ac bridge methods.

Vacuum-tube voltmeters

To measure the actual dc voltage on the plate of an amplifying tube, we must remember that the plate current passes through the plate-load resistor and also the decoupling resistor, if one is in the circuit. If an ordinary moving-coil voltmeter is used to measure this voltage, the current drawn by the voltmeter will cause a potential drop across the load and other resistors and give a grossly inaccurate reading. The error will be diminished if the meter only draws a small current. Meters of the type described as "20,000 ohms-per-volt" give a reasonably correct measurement of plate voltages for general service work. Some modern meters have sensitivities up to 100,000 ohms-per-volt. But for design purposes the accuracy is not good enough—the meter should draw practically no current from the circuit under investigation. A vacuum-tube voltmeter is such a device.

Vacuum-tube voltmeters are of differing types, some for dc measurements, some for ac; combined instruments contain facilities also for measuring resistance. As the application of the vtvm to the circuit has virtually no effect on its behavior, it is a very simple matter to check all tube voltages and also gain, stage by

stage. An audio oscillator provides the input signal for the amplifier, adjusted to the rms value postulated by the design. Voltage amplification is then measured throughout the amplifier by probing at successive grids and plates. If the oscillator input is held constant and varied throughout the frequency range of the amplifier, it is possible to take a response curve of the amplifier stage by stage, thus locating at what point, if any, the performance is defective. Response measurements of amplifiers should, as pointed out in an earlier chapter, be taken with a resistive output load.

Oscillators, oscilloscopes and other devices

No audio design or measurement work can be done without a signal source. An audio oscillator with a frequency range of about 30 to 30,000 cycles is an essential piece of equipment for the audio engineer. Various types are available and the comparatively recent R-C oscillators are considerably simpler and cheaper than the much older beat-frequency kind. The requirements are that the output should be quite sinusoidal at all frequencies; that the stability of frequency should be reasonably good; that the output should be capable of suitable attenuation. A useful feature is the ability to switch from sine-wave output to square-wave when desired. The square-wave output is generally achieved by using a clipped sine wave; for amplifier testing this can be considered quite good enough for all ordinary purposes.

Waveform investigation is impossible without a cathode-ray oscilloscope, and the larger the tube the greater the ease of examining the waveform. When the output of the amplifier is being examined, it is desirable to make a continual check that the input is true to specifications. This can be done by having an oscilloscope on both input and output. Separated oscilloscope traces are not easy to compare, so the use of a double-beam oscilloscope has great advantages in direct comparison. The input waveform can be given the same amplitude on the tube as the output by suitable adjustment of the internal amplifiers of the oscilloscope. Double-beam oscilloscopes are not so common as to be easily come by, but an ordinary oscilloscope can be converted into a pseudo double-beam instrument by using an electronic switch which switches the oscilloscope from the input to the output so rapidly that the tube screen persistence gives a continuous picture. A separate provision for adjusting the compared voltages is provided for in the electronic switching circuit.

Amplifier frequency response

The frequency response of an amplifier can be displayed on the c-r tube by using a sweep generator to sweep the whole frequency range sufficiently rapidly to give a steady trace on the screen. In this way adjustments to improve the frequency response are instantaneously observable. But this will not at the same time check for waveform distortion; an amplifier must be checked separately for frequency response and freedom from distortion.

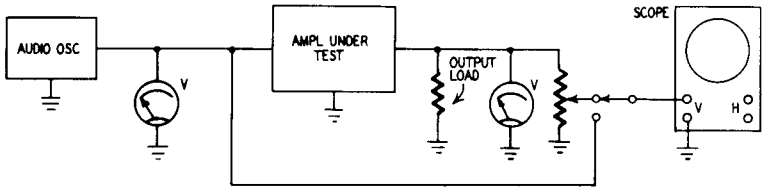


Fig. 1104. *Block diagram of layout for measuring performance of an amplifier.*

Harmonic analyzers are desirable but not essential pieces of equipment in an audio laboratory for the amplifier designer's task is not to produce certain percentages of harmonic distortion but to eliminate it as far as possible. Knowledge of what various waveforms mean as to harmonic content is necessary to interpret the pattern and a selection is given starting on page 182 in this chapter.

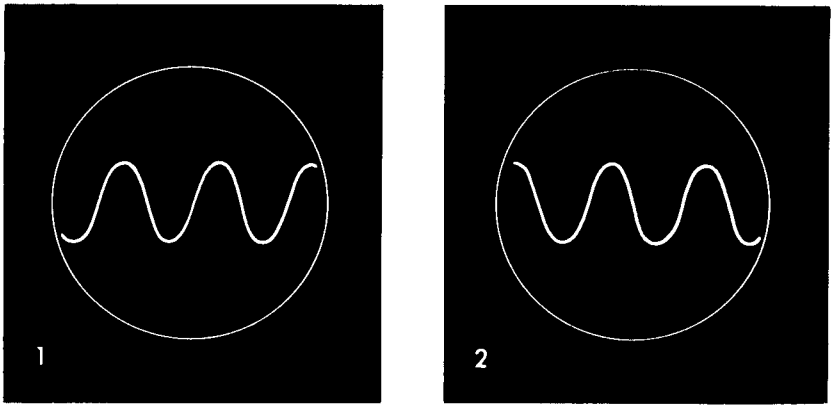
Measurement techniques

Testing an amplifier for overall performance should be done on the basis of the amplifier delivering its maximum undistorted output. Performance under any other conditions is quite misleading. In the block diagram of Fig. 1104 an audio oscillator is shown energizing the amplifier under test. If the oscillator has no output meter, an ac (rectifier type) voltmeter to measure the input volts must be added as shown. Every component and lead attached to the input of the amplifier must be properly shielded, otherwise hum will be introduced which will nullify the measurements.

A resistive load is connected across the output and this resistance must be capable of dissipating the full power output of the amplifier. This would suggest a wirewound resistor, but as this is inductive its impedance will vary with frequency. The load resistor is best made up of paralleled composition resistors of the highest ratings obtainable, unless, of course, noninductive wire-

wound units are available (easily made by winding on to a former a doubled resistance wire so that the two free ends are at the same end of the former).

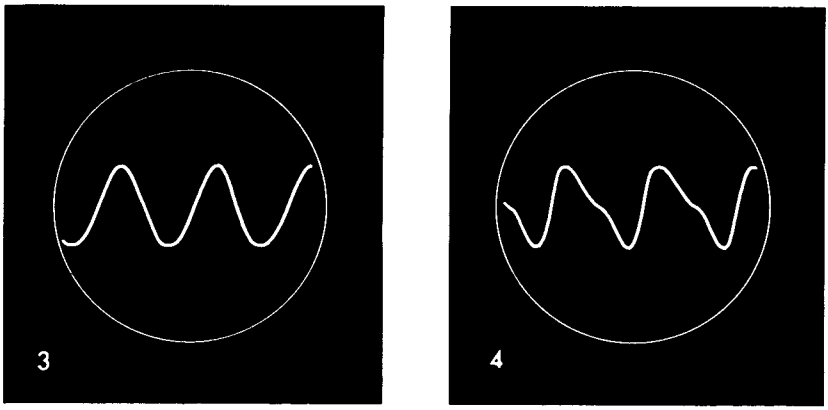
First test for stage gain, at, say, 1,000 cycles. Set the oscillator to this frequency and the oscillator output to give the designed rms input to the amplifier. Using a vtvm, touch successive grids and plates up to and including the output-stage grid circuit to make sure that the amplification per stage is as designed. The value of the load resistor is known and if the voltage across it is taken on an ac voltmeter (again a rectifier type) the current can be calculated by $I = E/R$, and the output watts roughly by $W = I \times E$. If the watts produced are too high, then there is too much gain in the amplifier; if too low, the gain is not enough. Generally speaking a



Trace 1) *Sine wave;* Trace 2) *15% 2nd harmonic +90° out of phase.*

little too much gain is usually of no importance; too little gain can be very tiresome. (These remarks must be modified by the feedback setting.)

An oscilloscope should now be connected as shown in Fig. 1104. For a single-beam type, a switch should be included to permit examination of the input and output waveforms; a double-beam oscilloscope will show input and output simultaneously. Check that the output at full watts is sinusoidal. At the same time note the voltmeter reading and put index marks on the scope reticule to denote the vertical height of the wave pattern. The output voltmeter can now be removed. Without bothering to reset the time base of the oscilloscope for other frequencies, sweep the oscillator down to the lowest frequency and up to the highest the amplifier is designed to reproduce. Nonsetting of the time base gives a mud-



Trace 3) *15% 2nd harmonic −90° out of phase;* Trace 4) *30% 2nd harmonic in phase.*

dled oscillogram but makes it easier to see if the height has varied at all. As the height of the pattern is proportional to the voltage it is a very simple matter to plot output voltage against frequency and so make a quick response curve of the amplifier. If the frequency response is what the designer intended, examination of the waveform at all frequencies can now be undertaken.

Keeping the input constant at the prescribed voltage and using the oscilloscope time base in the normal manner, check that a sine wave output exists at all frequencies. If everything appears to be in order, the audio oscillator can now be switched to square-wave output and the examination repeated. It is necessary to point out that a perfect square wave cannot be expected all over the frequency scale. At the low end, departure from the square is to be expected, but certain features of waves that are not square cannot be passed (see the waveforms given in this chapter). At high frequencies the thing is impossible unless the amplifier has an infinitely wide response. If the designed response of the amplifier is to be flat up to 50,000 cycles, the square wave will begin to lose its squareness at about 5,000. A square wave consists of a fundamental and an infinite series of harmonics; absolute squareness is not required, fortunately, but enough harmonics must be amplified to give reasonable squareness. If the amplifier cuts off at 50,000 cycles and the square-wave input has a frequency of, say, 30,000, it will be obvious that not even the second harmonic will be amplified. In practice the oscilloscope will show very nearly a sine wave at all square-wave inputs above about 15–20,000 cycles. But note particularly that as the square wave degenerates gradu-

ally into a sine wave, it must not, at some point, take on an irregular form for this will show that the amplifier is distorting transients very badly. The transition must be smooth and gradual.

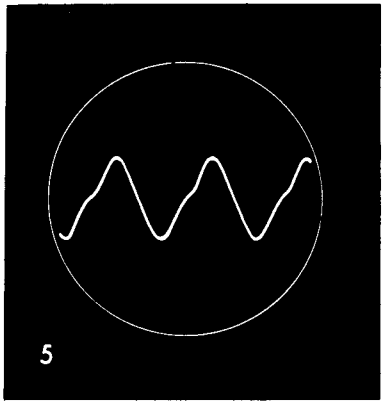
Where the secondary of the output transformer of an amplifier provides alternative taps for speakers of various impedances, the tests given here must be repeated for each tap, using the appropriate load resistance.

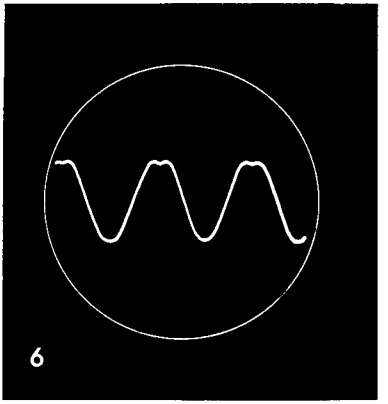
Interpretation of waveforms

The oscilloscope will display any distortion present in the output of an amplifier, but if the screen is small and the trace not finely focused, distortion which would be audible to a trained ear may not be perceptible to the eye. Some skill is needed to interpret what is seen on the screen. The waveforms of typical traces (starting on page 182) should be studied in conjunction with these notes.

Small amounts of harmonic distortion will not be seen except on a 5 or 7 inch tube with very good focusing, and even then it may be desirable to have a pure sine wave inscribed on the removable tube reticule for comparison purposes. If such is used, then the horizontal and vertical shifts and amplifiers must be adjusted so that a sine-wave input coincides with the reference waveform. This refinement can, however, be dispensed with (and would be out of place with a small oscilloscope) if the amplifier under test is driven to overload. The increased harmonic distortion will then be readily seen and its nature determined. Such traces would not be correct for the amplifier under normal operating conditions, but they will indicate wherein the suspected impure sine wave departs from the perfect form.

Trace 1 shows a sine wave; with a small oscilloscope it may be desirable to alter the sweep frequency so that only one complete wave is displayed at any frequency; this will make it easier to spot departures from the perfect form, but involves considerable adjustment to the time base. As a general rule not more than three complete waves should be displayed otherwise it will be impossible to detect irregularities in the waveform except at maximum and minimum points; the accompanying traces generally give two waveforms. The appearance of a pure sine wave should be memorized for the particular oscilloscope in use, for faults in the oscilloscope may distort the trace. The instrument should, of course, be free from innate distortion, but if it is not quite perfect, a note should be taken of the shape of the sine wave trace.



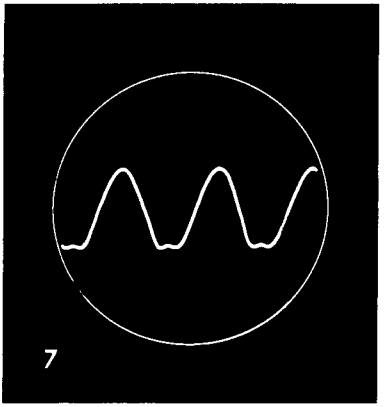


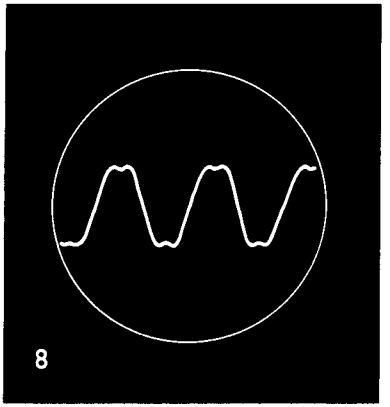
Trace 5) *30% 2nd harmonic 180° out of phase;* Trace 6) *30% 2nd harmonic +90° out of phase.*

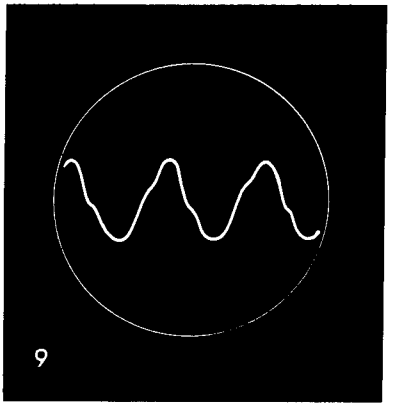
Second harmonic distortion

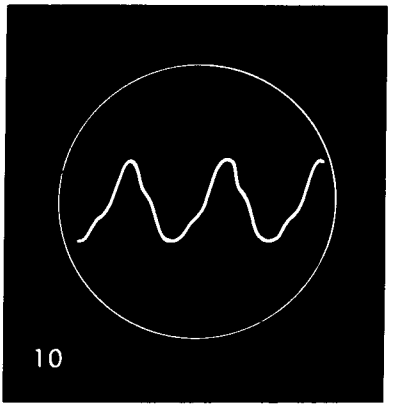
5% second harmonic distortion will hardly be perceptible at any phase displacement of the harmonic. If the second harmonic is out of phase with the fundamental there will be a gradual widening of the top or bottom of the waveform as shown in traces 2 and 3 which represent 15% second harmonic with phase differences of ±90°. Even harmonic distortion is present in single output tube amplifiers and in out-of-balance push-pull amplifiers to a small extent. Traces 2 and 3 show the result of overloading (as, in fact, do all the other traces) that produces the characteristics of second harmonic distortion.

Trace 7) *30% 2nd harmonic −90° out of phase;* Trace 8) *15% 3rd harmonic in phase.*









Trace 9) *15% 3rd harmonic +90° out of phase;* Trace 10) *15% 3rd harmonic −90° out of phase.*

If the second harmonic is in phase with the fundamental or 180° out of phase (being the reciprocal of in phase, as −90° is the reciprocal of +90°) the sides of the wave will begin first to lean to one side and straighten and then take on a curvature opposite to that of a sine wave. If the harmonic is in phase, the lagging side will begin to kink, if 180° out of phase the leading side will kink. Trace 4 shows 30% second harmonic in phase with the fundamental; trace 5 shows 30% second harmonic 180° out of phase. If the second harmonic is 90° out of phase, but with 30% amplitude, the tops and bottoms will flatten as shown in traces 6 and 7, these being emphasized conditions of traces 2 and 3.

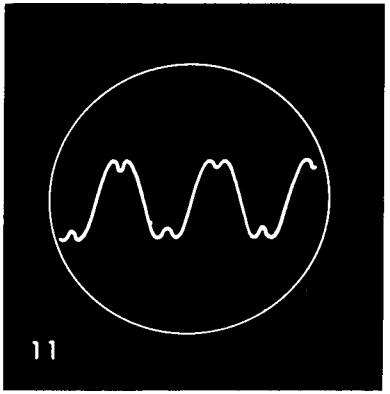
Third harmonic distortion

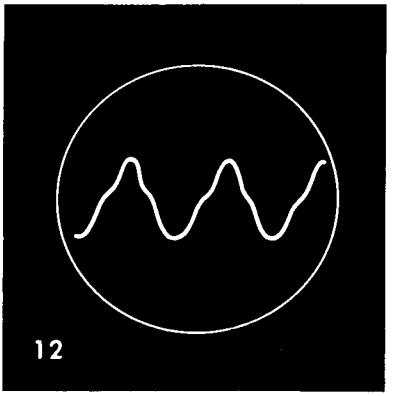
5% third harmonic in phase with the fundamental will need careful observation to see the difference from a sine wave, the tops and bottoms are both slightly rounded, but if 180° out of phase, the sides of the waves will be straighter. Notice that in-phase third harmonic tends to broaden both tops *and* bottoms, in contrast to second harmonic which broadens tops *or* bottoms. As the amplitude is increased to 15%, the in-phase third harmonic gives the shape at trace 8; an increase to 30% produces trace 11. This follows a similar sequence of flattening tops or bottoms of the second harmonic ±90° out of phase, but the flattening starts earlier.

Third harmonic 180° out of phase gives straighter sides, as has been stated, and these develop into the characteristic shape of trace 12, which represents 30% third harmonic. Note particularly, however, that the straight sides perceptible with 5% third har-

monic are quite symmetrical when the harmonic is 180° out of phase, for even with the kinks of trace 12 each half cycle is symmetrical. However, if the third harmonic is ±90° out of phase the straight sides will not be distinguishable from excessive second harmonic in phase for small percentages.

An increase of amplitude reveals the essential differences in traces 9 and 10, which display 15% third harmonic 90° out of

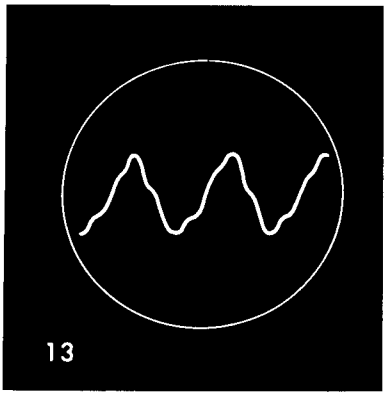


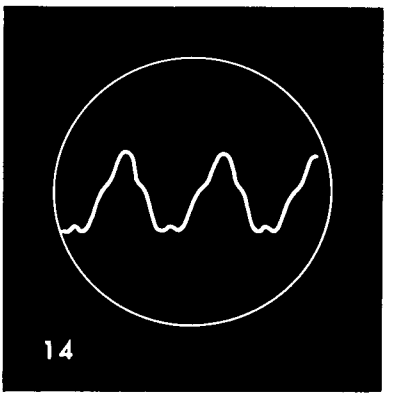


Trace 11) *30% 3rd harmonic in phase;* Trace 12) *30% 3rd harmonic −90° out of phase.*

phase. In the second harmonic, traces 4 and 5, the kink is in the middle of the side of the wave; with third harmonic it is nearer the top and bottom. Also traces 4 and 5 represent 30% second harmonic while 9 and 10 represent 15% third harmonic; distortion

Trace 13) *15% 4th harmonic in phase;* Trace 14) *15% 4th harmonic +90° out of phase.*





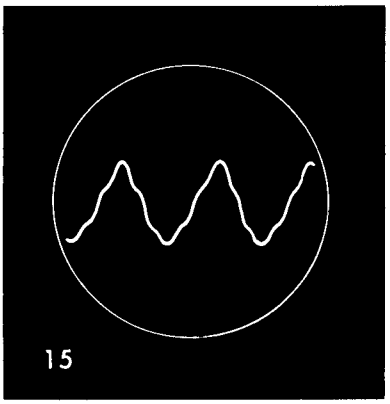
becomes evident much sooner with third harmonic to the ear as well as to the eye.

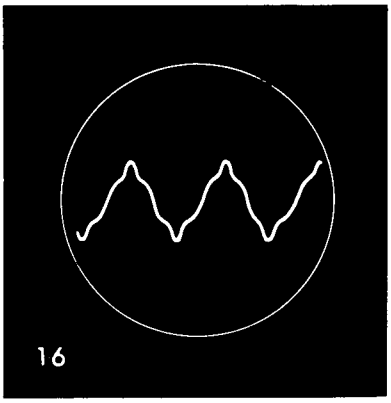
Fourth harmonic distortion

5% of fourth harmonic is distinctly noticeable. Traces 13 and 14 show 15% in phase and +90° out of phase; the traces would be reversed for reciprocal harmonic content. Fourth harmonic would be present in single tube output stages, but generally in much smaller percentage. A combination of 5% second and fourth harmonic would give a trace somewhat similar to those of traces 13 and 14. These traces obviously would not appear when investigating a properly balanced push-pull amplifier.

Fifth harmonic distortion

Fifth harmonic is soon perceived. Trace 15 gives the trace for only 5% fifth harmonic in phase with the fundamental; with increase, the form in trace 16 develops, which shows 15% of fifth harmonic. If the fifth harmonic is 180° out of phase with the fundamental, the patterns of traces 17 and 18 appear, representing 5% and 15% respectively. Note that the form of trace 14 is repeated in trace 18 except that the flattish bottoms of trace 14





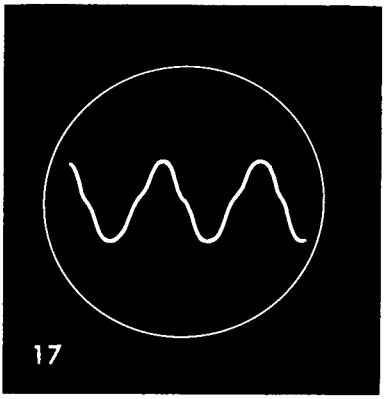
Trace 15) *5% 5th harmonic in phase;* Trace 16) *15% 5th harmonic in phase.*

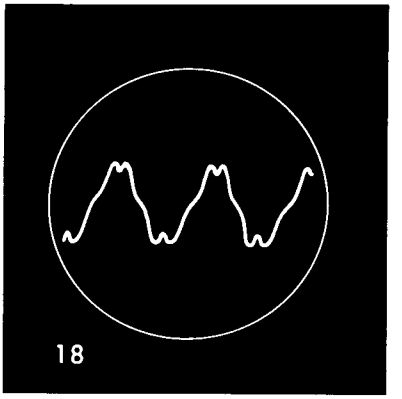
appear also in the tops of trace 18 as well as the bottoms; as a similar relation exists in traces 6 and 8 it will be seen that there is a "family likeness" in odd harmonics as there is in even harmonics. Combinations of even harmonics (as also combinations of odd harmonics) display the family pattern but with less total harmonic

content than the traces from separate harmonics, conforming to the generally recognized principle that wide range reproduction must avoid high harmonic distortion; a wide range speaker very quickly reveals distortion in an amplifier of limited performance.

Second plus third harmonic distortion

This combination is always present with single tetrode output stages. 5% of second and third is shown in trace 19 (both in phase

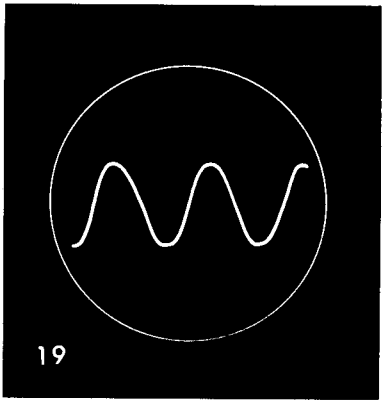


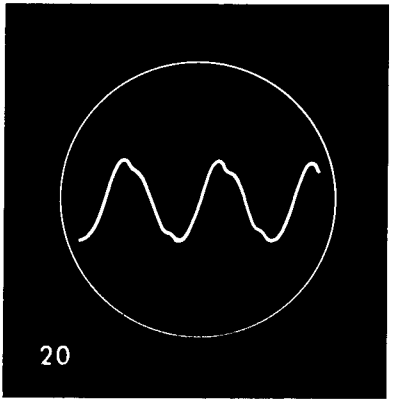


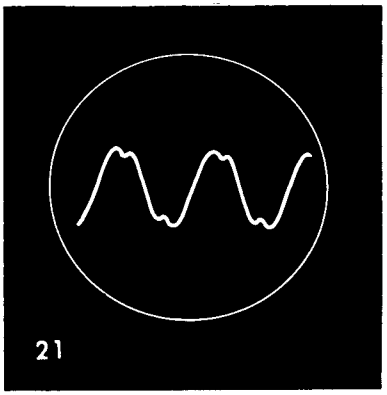
Trace 17) *5% 5th harmonic 180° out of phase;* Trace 18) *15% 5th harmonic 180° out of phase.*

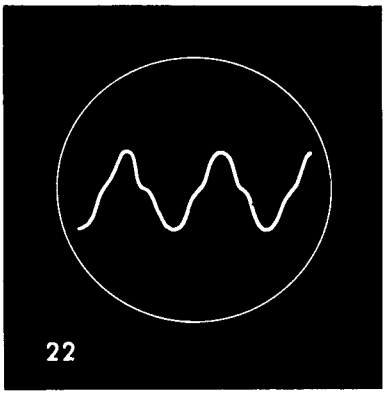
with the fundamental) and if the amplitude of the third harmonic is greater than that of the second, as is usual with tetrodes, the very characteristic shape of trace 20 develops. Increasing the

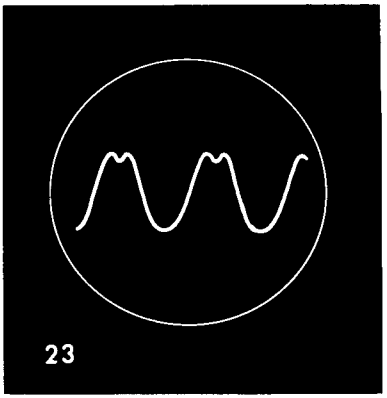
Trace 19) *5% 2nd harmonic in phase and 5% 3rd harmonic in phase;*
Trace 20) *5% 2nd harmonic in phase and 15% 3rd harmonic in phase.*



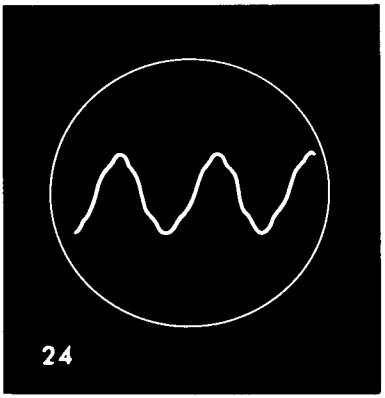


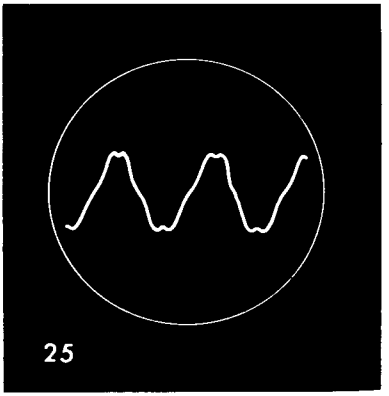


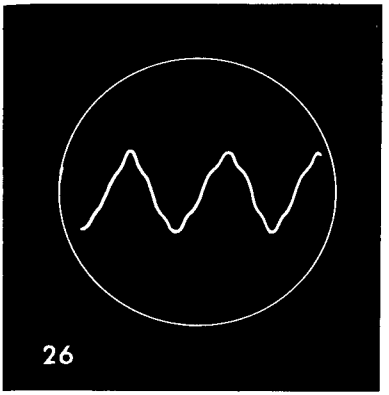


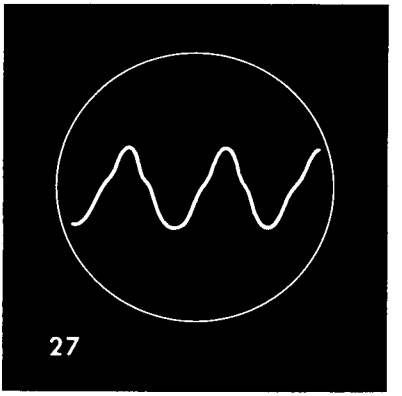


Trace 21) *5% 2nd harmonic in phase and 30% 3rd harmonic in phase;* Trace 22) *5% 2nd harmonic in phase and 30% 3rd harmonic 180° out of phase;* Trace 23) *15% 2nd harmonic −90° out of phase and 15% 3rd harmonic in phase;* Trace 24) *5% 3rd harmonic in phase and 5% 5th harmonic in phase;* Trace 25) *5% 3rd harmonic in phase and 5% 5th harmonic 180° out of phase.*

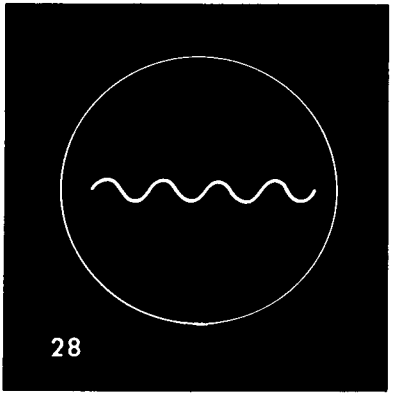


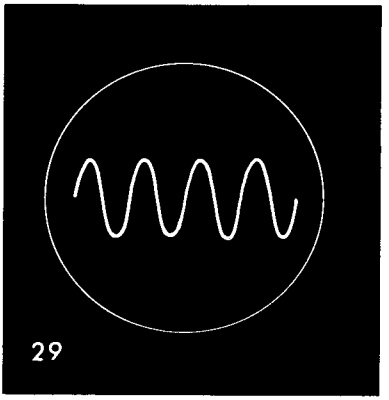


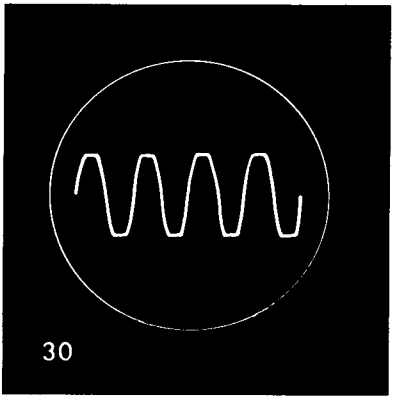




Trace 26) *5% 3rd harmonic 180° out of phase and 5% 5th harmonic in phase;* Trace 27) *5% 3rd harmonic 180° out of phase and 5% 5th harmonic 180° out of phase;* Trace 28) *waveform when amplifier is set to maximum undistorted output;* Trace 29) *slight overloading;* Trace 30) *more overloading.*



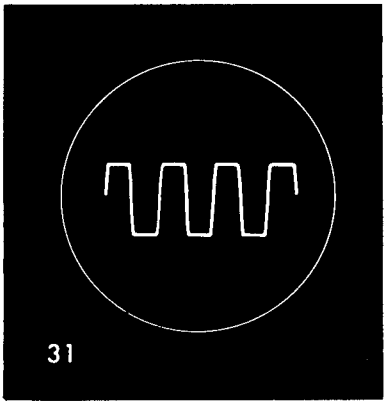


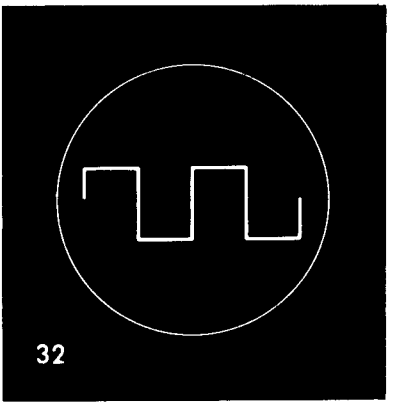


third harmonic content still more produces trace 21, which differs from trace 11 in that the two peaks are not the same height. If the third harmonic only is out of phase with the fundamental, trace 22 is produced, but this is not the same as trace 12, for the halfway kinks are not symmetrical as they are in trace 12. If the second harmonic is 180° out of phase and the third harmonic is in phase, traces 19 to 22 are reversed. Trace 23 represents 15% each of second and third harmonic, and differs from trace 6 in that the bottom troughs are wider.

Third plus fifth harmonic distortion

This combination is present in all push-pull tetrode amplifiers



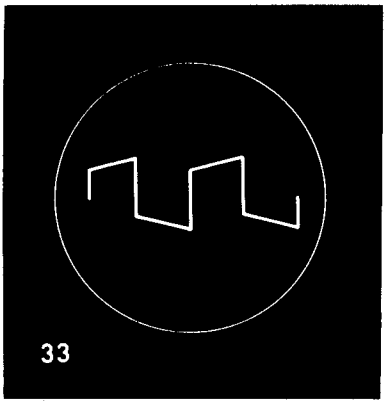


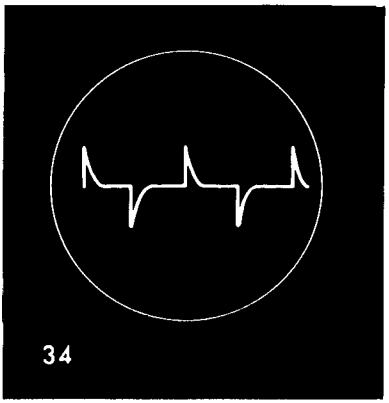
Trace 31) *Severe overloading;* Trace 32) *square wave.*

when overloaded. This type of distortion is soon seen and heard. Trace 24 shows 5% of each in phase with the fundamental; trace 25 when the fifth harmonic is 180° out of phase; trace 26 when the third harmonic is 180° out of phase; trace 27 when both the third and fifth harmonics are 180° out of phase. If either harmonic is only 90° out of phase, the tops and bottoms will be tilted to the right or left. It will be noticed that the divergence from the sine wave form is less when both harmonics are in phase with each other even if out of phase with the fundamental. As phase shift is directly associated with attenuation, in-phase conditions are secured by having an amplifier with wide frequency response.

General observations

If an amplifier under test is set to maximum undistorted output power and the output applied to the oscilloscope, the trace will

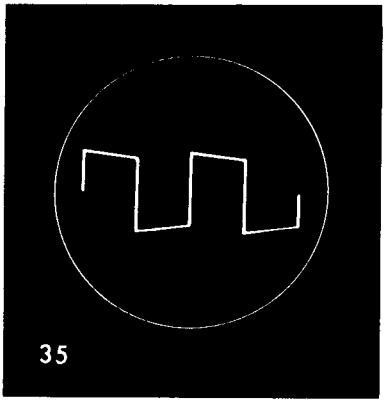


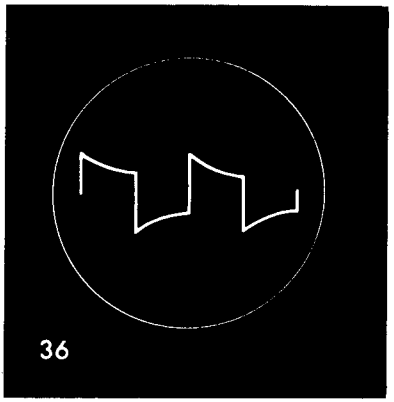


Trace 33) *Effect of long time constant on a square wave;* Trace 34) *Effect of short time constant on a square wave.*

appear as in 28. The oscilloscope vertical amplifier must at no time be overloaded itself and is best set to give a trace not more than about one third of the screen diameter in height. If, now, the input to the amplifier under test is increased, distortion will begin to appear, as in trace 29; with further increase of input the trace will change to 30 and then 31, which is a fairly good imitation of a square wave; this denotes very considerable distortion. An almost perfect square wave can be obtained from a fundamental and twelve odd harmonics, and traces 30 and 31 might be considered roughly as a sort of summation of all the earlier traces. This square-wave output can be used as a means of checking the per-

Trace 35) *Effect of an amplifier (with good response) on a square wave;* Trace 36) *Effect of an amplifier (with bass attenuation) on a square wave.*

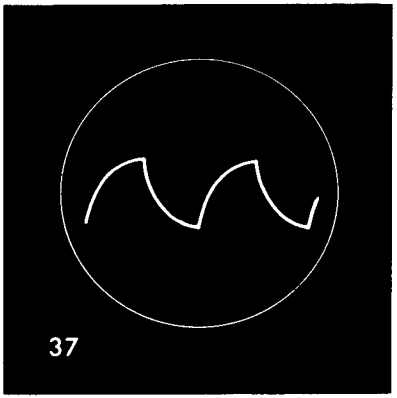




formance of the amplifier under normal conditions by injecting a square wave from a suitable input oscillator.

Square wave testing

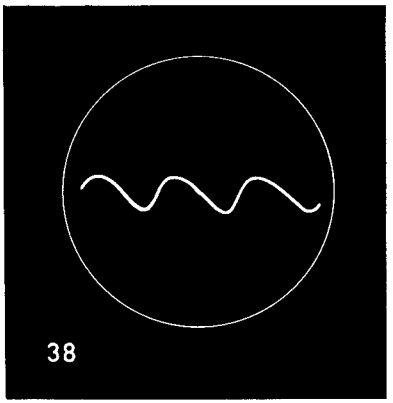
The oscillator should be applied to the oscilloscope for checking; trace 32 gives the appearance of the square wave on the screen. An amplifier without feedback can hardly be expected to give a true square wave at any frequency, for if a simple R-C coupling is inserted between the oscillator and the oscilloscope, a trace like 33

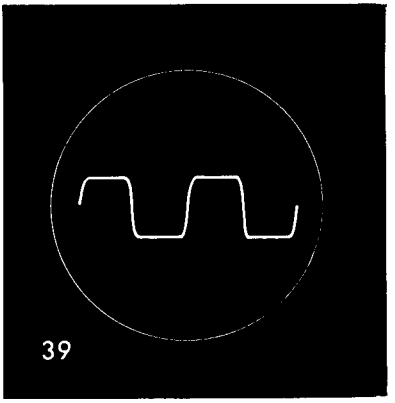


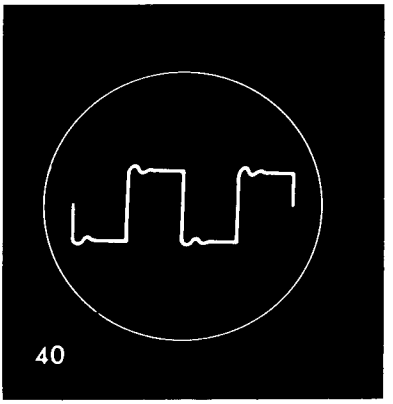
Trace 37) *Effect of an amplifier (with treble attenuation) on a square wave, typical of an amplifier without negative feedback, at 3,000 cps.*

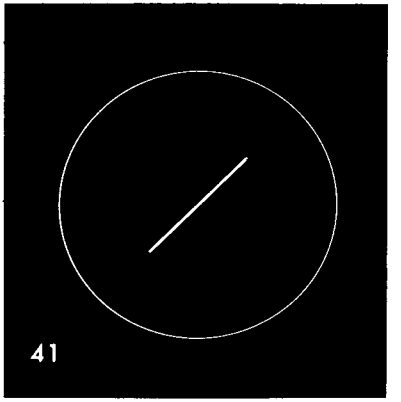
will be obtained with a long time constant and one like 34 for a short time constant. It might be supposed, therefore, that a complete amplifier with perhaps three R-C couplings will look very

Trace 38) *Effect of an amplifier without negative feedback on square wave at 10,000 cps;* Trace 39) *Effect of an amplifier with negative feedback on square wave at 10,000 cps.*

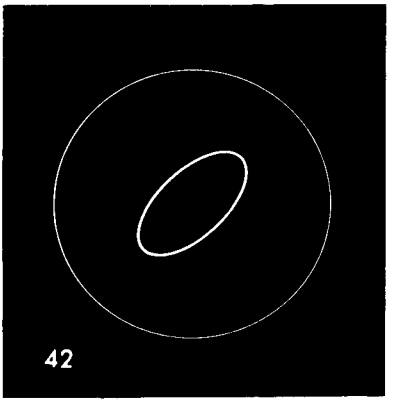


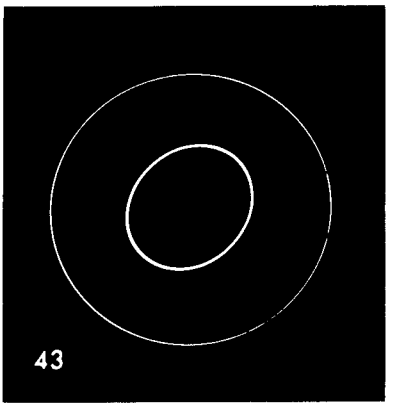


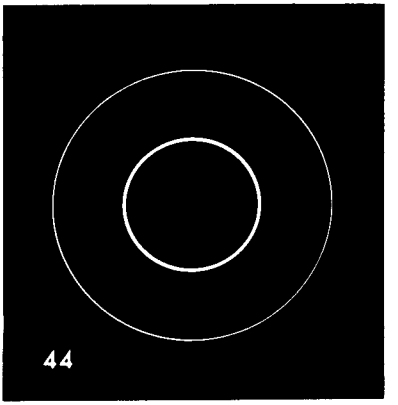




Trace 40) *Parasitic oscillation revealed by square wave;* Trace 41) *Two voltages in phase;* Trace 42) *Two voltages 30° or 330° out of phase;* Trace 43) *Two voltages 60° or 300° out of phase;* Trace 44) *Two voltages 90° or 270° out of phase.*





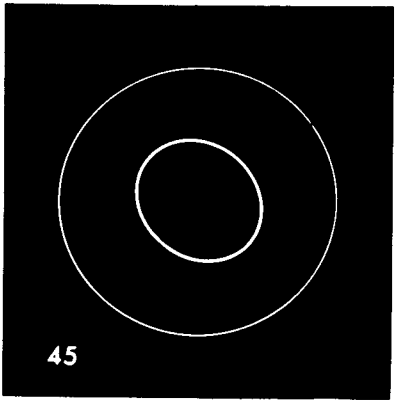


bad indeed. In practice it is not quite like this, for the trace depends on the phase shift at various frequencies and phase shifts will sometimes cancel.

First take an amplifier without feedback. At 50 cps the trace will appear as in 33 if the bass response is exceptionally good; if poor, like 34. In general it will appear like trace 36. At the usual mid-frequency, 1,000 cps, where the amplifier is usually flat for some octaves below and above, there will be the nearest approach to a square wave, but with some rounding of the leading upper and lower corners. At about 4,000 cps, the shape will have changed to trace 37, owing to a loss of high harmonics, and at 10,000 cps the output will be nearly sinusoidal, as in trace 38.

Addition of the right amount of feedback will secure a trace like 35 at 50 cps. As the frequency rises the form will change to a perfect square wave at 1,000 cps which should be maintained until at least 3,000 cps. Thereafter the shape will gradually spread until at 10,000 cps the trace of 39 will be developed. After that, the squareness will depart until a sine wave is reached.

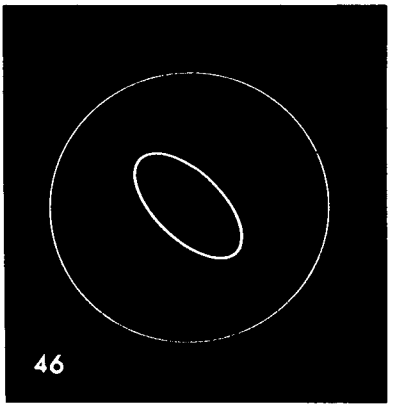
The oscilloscope is unquestionably the surest way of setting the amount of feedback. If the feedback is too great, motorboating will start, indicated by the pattern jumping up and down on the



Trace 45) *Two voltages 120° or 240° out of phase.*

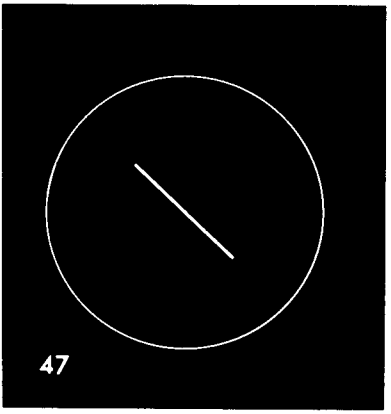
screen. Reducing the feedback will increase the gain and increase harmonic distortion. Adjustment of the feedback is probably most easily settled by driving the amplifier into overload and increasing the feedback until gain and distortion is reduced to the desired level. Such adjustments are best carried out using a sine-wave in-

put; but when the best setting is thought to have been reached it is necessary to switch to square-wave input, for the amplifier may be only "conditionally stable" and may break into oscillation by the shock excitation of the transient front of the square wave. The amplifier should also be switched off and the tubes allowed to cool,



Trace 46) *Two voltages 150° or 210° out of phase.*

then switched on again to see if there is instability on warming up. If there is, the margin of safety is inadequate.



Trace 47) *Two voltages 180° out of phase.*

Distortion of the waveform as in trace 40 is known as ringing. It is due to parasitic oscillation and must be located and cured.

Phase measurement

The oscilloscope is the most convenient instrument for displaying the phase relationships of two voltages. One voltage is applied across the horizontal plates, the other across the vertical plates. The time base is not used. The horizontal and vertical amplifiers should be adjusted so that the spot displacement on the tube is equal on both axes. When the two voltages are in phase, the trace is as shown in 41; when 180° out of phase, trace 47 applies. The circular trace 44 indicates that the two voltages are 90° or 270° out of phase; an elliptical trace indicates intermediate phase differences. Trace 42 represents 30° or 330°; 43 is for 60° or 300°; 45 shows 120° or 240° and 46 is 150° or 210°.

The phase/frequency characteristic of an amplifier can be easily taken by connecting the input of the amplifier (output of the sinusoidal oscillator) to the horizontal plates of the oscilloscope, and the output to the vertical plates. The oscillator is then run over the frequency scale and the phase shifts of the whole amplifier noted for each frequency. This provides the preliminary information before planning feedback circuits to improve the response.

Perfection of phase inversion stages can be checked by connecting the output-tube grids to a horizontal and vertical plate, the other plate terminals being common grounded to the amplifier ground. The trace should be exactly as 47; if not, the parameters of the phase inversion stage must be modified until the phase inversion is exactly 180°.

chapter
12

high fidelity – hail and farewell

THE preceding chapters of this book are what might be called engineering writing; they deal with facts which can be tabulated and illustrated by charts. They can be confirmed by making experiments and taking measurements or by reference to other authorities. They constitute the hard core of audio engineering—and of high-fidelity reproduction. In one respect they resemble the engine of an automobile in that they provide the power to do the work, but unlike automobile engineering there is no general agreement as to what sort of power is wanted. But as the music lover doesn't want just a turntable, an amplifier and a speaker sitting naked and unashamed on the floor, so the car owner wants something put round his power unit.

It is generally supposed that the customer gets what he wants in the end, but it would be more accurate to say that in these days of mass production, with consequent heavy outlays on tooling and assembly equipment, the auto manufacturer can only try and guess what sort of car would be salable. Having made the decision, he molds the public mind by the usual publicity methods into accepting his design. There is a great deal of hardware hung around the modern automobile chassis that performs no useful function but which has to be paid for by the customer, and in hi fi (the contraction is used deliberately) a similar state of affairs prevails. But whereas the demand in automobiles is toward a push-button state of affairs, hi fi seems to involve a multiplicity of controls which are supposedly intended to remedy shortcomings in the original source of signals but which, more often, are used to

create a sort of noise which is deemed to be "stunning," although it is liable to stun in another sense.

In this chapter I am going to write in the first person, because some of the things here are not engineering facts. They can be disputed. You can agree with them or say they are nonsense, but one basic logical fact is that any argument in which the disputants have not clearly understood the fundamental assumptions that lead to the argument can never be resolved. As to my qualifications to write on what must be a very thorny problem, I can only say that I was one of the pioneers of high-quality sound reproduction and since 1925 I have had no other professional interest. In those many years I have met users of audio equipment in different countries, have noted what they think and what they want, have tried to give them what they want and have not been unsuccessful. From all this practical experience a pattern has emerged, and I propose to deal with some of the "intangibles" of sound reproduction as I see them and as many others see them too.

High fidelity

I invented the phrase "high fidelity" in 1927 to denote a type of sound reproduction that might be taken rather seriously by a music lover. In those days the average radio or phonograph equipment sounded pretty horrible but, as I was really interested in music, it occurred to me that something might be done about it. I investigated the nature of sound, the behavior of the human ear, the minimum requirements of good sounding equipment, and finally produced something which did appeal to others also interested in musical reproduction. As the weakest link in the chain of reproducing equipment was the speaker, my work naturally was devoted to improving speakers. Nothing that had been done before was allowed to influence the ultimate design. If it was good, it was retained; if it wasn't, it was rejected. As most of my friendly competitors preferred to improve existing ideas, the result was that my ideas and designs were something unorthodox. In the last count I found that I was quite successful as a rather small manufacturer but there was never a wide enough appeal for the mass market, and the reason was very simple indeed. A high-fidelity sound system should not add to nor take away from the musical signal put into it. If a radio broadcast was of superb quality or a record an outstanding example of good recording, the results were magnificent, by the standards of those days. But in those days (and even occasionally today) most of the radio

transmissions were distorted by the poor technical quality of telephone land lines and most of the records were even worse. The consequence was that my "high-fidelity" equipment made a lot of the music sound awful, and most people said they would rather have something moderately acceptable than a rare item of real enjoyment.

Britain, therefore, was the pioneer country for high-fidelity. But it came too early and it is quite true, so far as my memory serves, that in all that long period between 1927 and 1939 I sold not a single speaker in the United States. The arrival of new techniques in recording provided a much better source of raw material for a high-fidelity installation, and this provided the impetus for a development which has now gotten somewhat out of hand. The object of this chapter, therefore is to try and bring some sort of order out of chaos.

I did not invent the word "audiophile" but I am grateful to the man who did because it is a very convenient term to apply to a certain type of audio equipment user. I do not want it to be thought that I have any derogatory opinions of audiophiles. On the contrary I greatly value their existence if only because they are good for trade. But that they exist emphasizes the logical need in this argument for a clear definition of both parties to the discussion. I shall call a man who has no interest in the mechanics of audio reproduction but is concerned only with the most faithful reproduction of existing music a "music lover"; the man who is more interested in stunts or the overpowering reproduction of drums, triangles and, if it comes to that, the components of "concrete music," an audiophile. In doing just that I imply that his major interest is in the means and not the end.

There is nothing discreditable in this. There is no absolute law which says that a long-haired musician is any more desirable a creature than a gimmick hound. So long as we remain an ostensibly free people we have a right to amuse ourselves as we want to, but it should be a point of honor that our amusement doesn't impinge on the free enjoyment of others. Some audiophiles make their presence heard over a very wide area. Then there are the folk in between. A music lover has every right to interest himself in the machinery that produces his music but, if he takes the advice of an audiophile, he may find himself landed with an outfit that doesn't give him what he wants. Some people are perfectly happy with a 10-watt amplifier feeding a modest speaker system. If it sounds good, then why should they waste money on something

more elaborate? Some audio engineers are very fond of music, and design equipment to satisfy themselves. Others are not very interested in real music but get a kick out of devising complicated systems that emit perfectly overwhelming sounds. Let everyone do as he will; but my motto is "live and let live" and the aim of this chapter is to guide the reader in the way *he* wants to live, not the way the other man says he should.

Some thoughts on frequency range

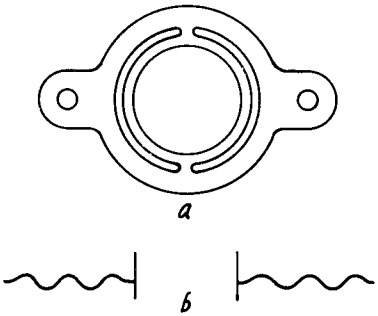
I believe the first serious analysis of the requirements for good musical reproduction was made by me and published in *The Wireless World* in 1932. My conclusion was that the *undistorted* range of a high-fidelity system should be 32 to 9,000 cycles. On the face of it, this seems a preposterously low stipulation in view of the stupendous increase in circuit and speaker refinement since those far-off days, but there are two things to bear in mind—one is the operative word "undistorted" and the other is that in those days wide-range radio reception was out of the question because of adjacent-channel interference. FM hadn't been thought of. But even to achieve that comparatively narrow frequency band without distortion is still a difficult proposition, in fact nearly impossible. We must put up with some distortion, but we can try and eliminate the sort that is objectionable.

More about speakers

Starting at the bottom and working up toward the top, the first thing a music lover will object to is a bass resonance. Every speaker has a bass resonance of some sort, although over 30 years ago it was no uncommon thing to have experimental speakers that were quite free from this distressing fault. The cone-coil assembly was suspended by threads and moved quite freely, too freely in fact for the coil had a habit of getting out of the magnetic gap and staying out. One British manufacturer had the bright idea of using brass wires to insure constancy of centering, but he overlooked the fact that the wires turned out to be very effective resonators. In a practical speaker it is necessary to provide some sort of restoring force so that at the termination of a transient signal of high amplitude the voice coil returns to its normal position.

In most speakers this is very simply achieved by molding a set of corrugations onto the outer edge of the cone and cementing the surround to the cone basket. When this is done, the spider at the

apex of the cone can be a design like Figs. 1201-a-b, where a is usually stamped out of thin bakelite sheet and b is molded from buckram and then varnished. Obviously there is no need for



Figs. 1201-a,-b. *Various types of spiders: a) concentric type suitable for use with a molded paper outer suspension; b) molded corrugated type now widely used, also suitable for molded paper outer suspension.*

greater freedom of movement in the spider than is possible from the outer suspension. This system of suspension will inevitably introduce a well-marked bass resonance, depending on the size of the cone and the flexibility of the suspension system, with a reso-

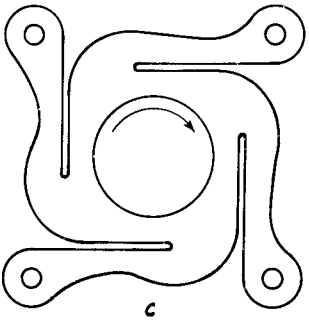


Fig. 1201-c. *Tangential arm spider producing rotary cone motion when used with a free edge suspension.*

nant frequency somewhere in the region of about 50 to 150 cycles.

If an attempt is made to avoid this by having a free outer suspension (and very many magazine articles have been written to show just how this can be done), the limiter then becomes the

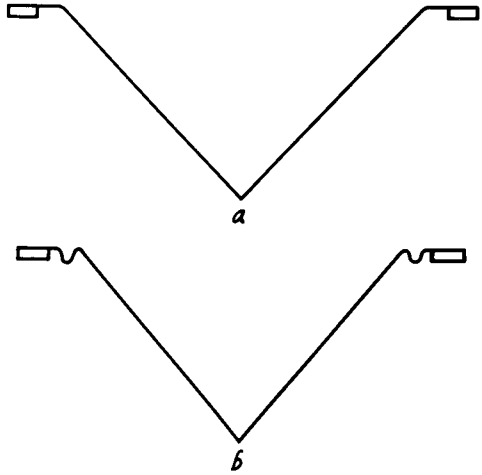
spider. Concentric spiders of the type shown in Figs. 1201-a,-b must obviously have a somewhat limited travel and toward the end of the travel the resistance to motion increases very rapidly since the annular components of the design cannot increase in size. This will still produce a highly emphasized bass resonance. In most speakers which have a "free" outer suspension, the surround is usually a fabric ring as shown in the cross-section of Fig 1202-a. Thin sheet rubber was actually used way back in 1925 and decently buried after tests showed it to be unsuitable; yet some speakers are offered today with such a surround. The drawback of the flat fabric ring is that it gives very little unrestricted movement.

The original Hartley-Turner speaker had a radically different shape of spider (Fig. 1201-c). The tangential arms are clearly of fixed length but, if the center of the spider is moved forward as though you were pulling the diagram out of the page, they will allow the center of the spider to move forward a quite appreciable amount *if the center of the spider is allowed to rotate,* as shown by the arrow in the center of the diagram. If the cone to which the voice coil is fastened is attached to the basket by the usual integrally molded paper surround, it cannot rotate and the design has the same limitations as those of Figs. 1201-a,-b, so provision is made to enable the cone to rotate on its axis. This is achieved by fitting a fabric surround with a U molded into it (as shown in Fig. 1202-b) which provides plenty of play to accommodate the rotary motion of the cone. As the cone rotates, it is almost all around the circle, working against the weave in bias, and all woven fabrics can be stretched on the bias. This provides a cushioning effect and the movement of the cone is not decisively snubbed at one particular point. This accounts for the lower peak in Fig. 1005 in chapter 10, and the feature was so successful that in 30 years it has never been found necessary to modify it.

When I first put this design on the market, I claimed with some justification that I had produced a speaker which had virtually no bass resonance, but I had also to consider what the implications of adopting this design really were. It needed no great technical know-how to appreciate that, as the cone was almost literally free-edge, it was free to move as it wanted to. And one thing that a cone does want to do is to node at bass frequencies. To provide the stiffest possible cone meant the use of bakelized paper, much harder and stronger than ordinary molded paper-pulp cones. Such paper comes in flat sheets, so the cone blank had to be cut from

the sheet and joined up as a cone. Experiments indicated that the strongest cone should have an included angle of 90° and should be as small as possible. If it were too small, the power-handling capacity at low frequencies would be negligible; if too large, noding troubles would be severe. The final compromise was a diameter of 8 inches. And as the bakelite paper was hard, the top was good. As the cone angle was comparatively narrow, the speaker was very directional.

This account of early work is not meant to prove what a clever person the author is; the purpose is to illustrate that all audio



Figs. 1202-a-b. *Free edge suspension: a) usual type consisting of a flat felt, woven cloth, thin or sponge rubber ring. Movement is rather restricted; b) free edge suspension permitting rotary motion arising from tangential arm spider. Incorrect material or its application will result in cone drooping.*

work has to be a compromise. In striving for perfect bass reproduction I had introduced the liabilities of limited power handling at low frequencies and directional treble response. Other manufacturers had decided to have a bass resonance and reduce excessive response on the axis by using exponential cones, which, because they were made of softer paper, had less top. Who was right? There is no answer to that question for it all depends on the user's taste in sound reproduction.

Channels of purity

Of course you will say that, if I had known in those early days

what we all know now, I would have devised the near-perfect speaker for bass and another near-perfect unit for treble. But that is just what I did—in 1930! I had traveled sufficiently far along the high-fidelity road to appreciate that the requirements for good bass and good treble reproduction were inherently conflicting. So I constructed a nicely enlarged speaker having an 18-inch cone and restricted the freedom of the cone edge so as to suppress nodes, which could be done because the cone movement would be much less than with an 8-inch cone. To this I added a tweeter speaker with a wide-angle cone of the same material (to maintain constant coloration of the sound emitted) and a carefully designed dividing network. That speaker was displayed at the National Radio Exhibition in London in 1931 and aroused a lot of interest. It was, I think, the first multi-channel speaker put into production, and I let it run for a year, during which time many people bought it and thought it wonderful. I then withdrew it because I thought it sounded horrible, and the main reason was that the large cone speaker had an air-column resonance within the cone that to my ears was unendurable. I also disliked the sound coming from two separate and distinct sources.

In these sophisticated days it is quite obvious that the widest frequency response is obtained from a three- or four-way speaker system, if mere width of frequency response is what is wanted. The audiophiles want it but I have no real evidence that music lovers want it. Let us consider why. As we grow older, the frequency response of our ears contracts. I am quite certain that my ears had a wider frequency response in 1930 than they have now, and I will go so far as to say that my wartime experiences in England in several blitzes didn't make matters any better. My wide-response ears in 1931 caused me to withdraw my unique "duplex" speaker because they heard things I didn't want to hear, because they weren't in the original music. My impaired hearing of 1957 (although I can still hear 22,000 cycles at about 40 db) still rejects what I hear from a modern multi-channel system for precisely the same reason. What I find unacceptable is the sort of bass I hear from a large-cone speaker, the sort of treble I get at the same time from a tweeter of different diaphragm material, and a lack of continuity of coloration over the whole system.

Speaker coloration

Don't tell me there is no coloration in a modern speaker. I can hear it, and why shouldn't I? What makes the difference between

a Stradivarius or Amati and a mass-produced fiddle is the stuff of which the instruments are made. What makes the difference between the woodwind and brass of the orchestra is the stuff of which the instruments are made. What makes the difference between a Steinway and a Bechstein is the stuff of which they are made. You just don't get away from it, and the sound you get from a paper-coned woofer is quite different from the sound you get from an aluminum-cone tweeter or an electrostatic speaker for that matter. I can hear it, and music lovers can hear it. But I suspect the audiophile doesn't bother. He seems to want to reproduce the sound of a triangle more triangular than the real thing. I don't mind one little bit if that is what he wants; all I beg is that he doesn't call it high fidelity.

I am quite convinced, from over 30 years of listening to real music and reproduced music, that if we could get *perfect* reproduction between the limits of 32 and 9000 cycles we should, as music lovers, be quite satisfied. But a long time has passed since that *desideratum* was laid down, and I would say now that I should like to have a very fine response from 20 to 12,000 cycles. I say that because I have this past year found means of extending what I can reproduce with my speakers, and the extension both ways has a profound effect on the enjoyment to be had from the reproduced music.

Don't get the impression that, because of what I have just said in the way of disliking multi-channel speakers, I deliberately make do with speaker designs that can only reproduce up to 12,000 cycles. The published response curve of my 215 speaker (Fig. 1001 in chapter 10) shows that it goes on quite gaily up to something like 20,000 cycles. But simply because I designed the thing I am more entitled to criticize it than anyone else, for commercial considerations do not enter into the matter. Perhaps I shouldn't foul my own nest but, if I am a purveyor of high-quality reproduction, I see no reason for becoming a purveyor of low-fidelity writing. The 215 speaker does have a useful response at 20,000 cycles and thereby competes very favorably with two-and three-way systems, in spite of its modest appearance. But . . .

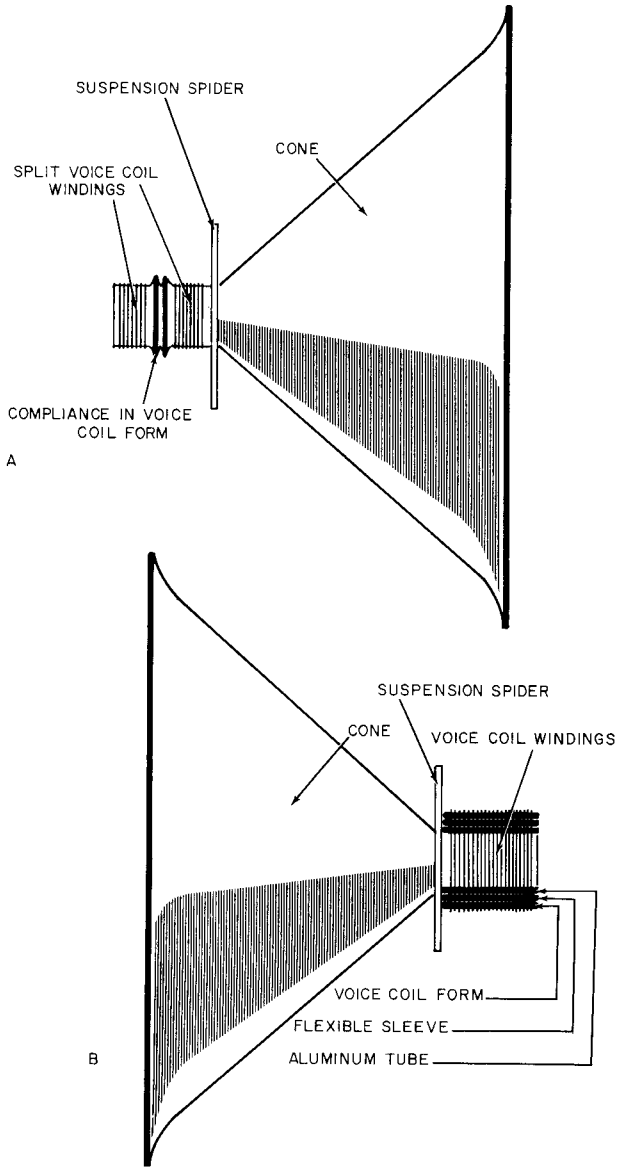
The cone, the coil and distortion

I have already explained in Chapter 10 that cone deformation causes distortion. In a more specific way it can be stated categorically that cone breakup produces upper harmonics of the fundamental frequency applied to the speaker. The cone of my old

bakelite-coned speaker did not tend to break up and because of that it seemed to be somewhat lacking in high treble response. My first effort to improve the top was to use the compliance type of voice coil. Two forms of this are given in Fig. 1203. The first method depends on the fact that at high frequencies the part of the coil nearest the cone is the only one that moves, the elastic compliance between preventing transmission of the coil movement to the remainder of the coil. The second method is the one I adopted. Here the cone is attached to an aluminum tube surrounded by a rubber sleeve which is compressed to some degree by the voice-coil former which carries the winding proper. At low frequencies the whole assembly moves as an integral unit; at high frequencies currents are induced in the single-turn "secondary" winding (the aluminum tube) which alone drives the cone, the elasticity of the rubber compliance isolating the movement from the "primary" voice-coil winding.

This certainly gave me more top, but I was still bedeviled by the directional properties of the narrow-angle straight-sided cone. In an experimental way I succeeded in producing a small number of molded bakelite cones of exponential shape, but they didn't turn out a regular manufacturing proposition. I therefore turned my attention to seeing what could be done in the way of improving the more ordinary paper cone. Any attempt to stiffen it by applying a hard lacquer only succeeded in increasing the mass so much that any potential improvement brought about by the treatment was nullified by the increased weight. Then it occurred to me that, as the apex of the cone is the part that produces the top, it seemed a sensible thing to separate the apex from the rest of the cone by introducing a cone compliance, just as I had done with the voice coil. So I removed a thin ring of paper about halfway down the cone and replaced it with a film of elastic material in a solvent which was allowed to dry out, leaving the two halves of the cone separate but flexibly coupled. I took a response curve and found the top response had shot up. This was first carried out in 1938 and during the next 12 months I had produced quite a lot of speakers which gave great satisfaction. So it seemed that the idea really worked.

After the war I resumed manufacture and still found the idea worked, which tempted me to think about invading the United States market. The 215 speaker was introduced in 1949, the first of many subsequent arrivals from Britain and Europe. The claims made for the speaker were borne out in practice, and it did give



Figs. 1203-a,-b. *Two methods of introducing a compliance into a voice coil; a) split voice coil on a form with center compliance; b) compound voice coil with intermediate compliance.*

a frequency response comparable with more complicated speaker systems without incurring the disabilities of the multi-channel idea.

But in every audio box of ointment lurks a fly. I happened to have one 215 speaker that had escaped the depredations of war, and several made between 1946 and 1950 when I started a new line of investigation in 1956. The cone compliance of the prewar speaker was harder than the paper it joined; the compliances of the 1946–50 period were a good deal stiffer than those of 1956, yet measurement showed that all speakers had the same response in the upper register. The information given in Fig. 1001 (Chapter 10) was available to me in 1952 and set up a train of thought but no opportunity arose for a more thorough investigation. The findings of 1956 proved that new light was needed on an absurd state of affairs. Work on the development of a more rigid cone suggested that an exhaustive examination of the behavior of the paper cone should be undertaken. In due course it was discovered that the addition of the "compliance" to the original paper cone had simply resulted in a stiffness pattern being applied to the paper so that the sections of the cone between the stiffer parts vibrated at harmonic frequencies of the applied signals. The original measurement was, in its way, perfectly authentic. A microphone was placed before the speaker and the measured output plotted against the frequency of the input to the speaker, and the result was the response curve of the speaker. Unfortunately it was not appreciated that the result could be completely misleading, and it would not be out of order to presume that very many other speakers have inadvertently misleading response curves too! A microphone is not a harmonic analyzer.

With the 220 speaker referred to in Chapter 10 ready for production, it was only natural to run comparative tests for response between the 215 and the 220. It was found then that the 220 had appreciably less response above 16,000 cycles than the 215, and this could well be due to the fact that the cone, without any form of compliance, was not breaking up. The next step was to try out the new speaker on all kinds of "guinea pigs"—music lovers, audiophiles and audio dealers who could run the thing on A–B tests against everything in their showrooms. Without exception the reports were that, apart from the quite astounding bass from such a comparatively small speaker (the invariable comment was that the stiff 9-inch cone gave more than a paper 15-incher), the top response was incredible. We did not reveal

that the speaker had less top than the 215, we just said "here is a new speaker—do what you like with it." My colleagues and I have proved to our own satisfaction that what happens after about 15,000 cycles doesn't matter, and the less there is the better. But if the response is outstandingly good below that figure and free from distortion, then there is a quality of definition that doesn't seem to be attainable by any other method.

Once again I do insist that here is no propaganda for a new design of mine. When all is said and done, the customer will buy what he likes best, but what must be clearly appreciated is that what he likes may not necessarily be accurate reproduction. I have satisfied myself that near-perfect reproduction up to 12,000 cycles will enthrall a genuine music lover (my own personal choice of a pickup doesn't do much over that figure but is supremely good up to that point), and any reproduction up to the frequencies that only dogs can hear not only contributes nothing to the quality but adds a good deal of distraction by reproducing scratches, noises and so on that are not in the original music.

But I cannot admit to any liking for a bass cutoff. In spite of what theory may suggest, I find that the ability of the reproducing system to go down without distortion to below the audible limit does improve what is above that limit. I do not profess to explain why; I just know that it is so. But it calls for a super-excellent turntable. My extremely expensive American record player was found to have such a rumble at 18 cycles that it could not be used, for the rumble alone was enough to overload the speaker, giving a cone excursion of ½ inch. Another point to watch is the way the speaker is mounted or housed. A speaker such as I have just described is really moving things at low frequencies—the air, the box containing it, the floor, walls, ceiling, windows, everything within range. If you go to church, you must have felt the building tremble when the open diapaison of the pedal organ gets going. That is what happens when a speaker with a cone that doesn't fold up is driven by a mere 20-watt amplifier. Very serious attention must be given to the business of mounting and housing.

Audio furniture problems

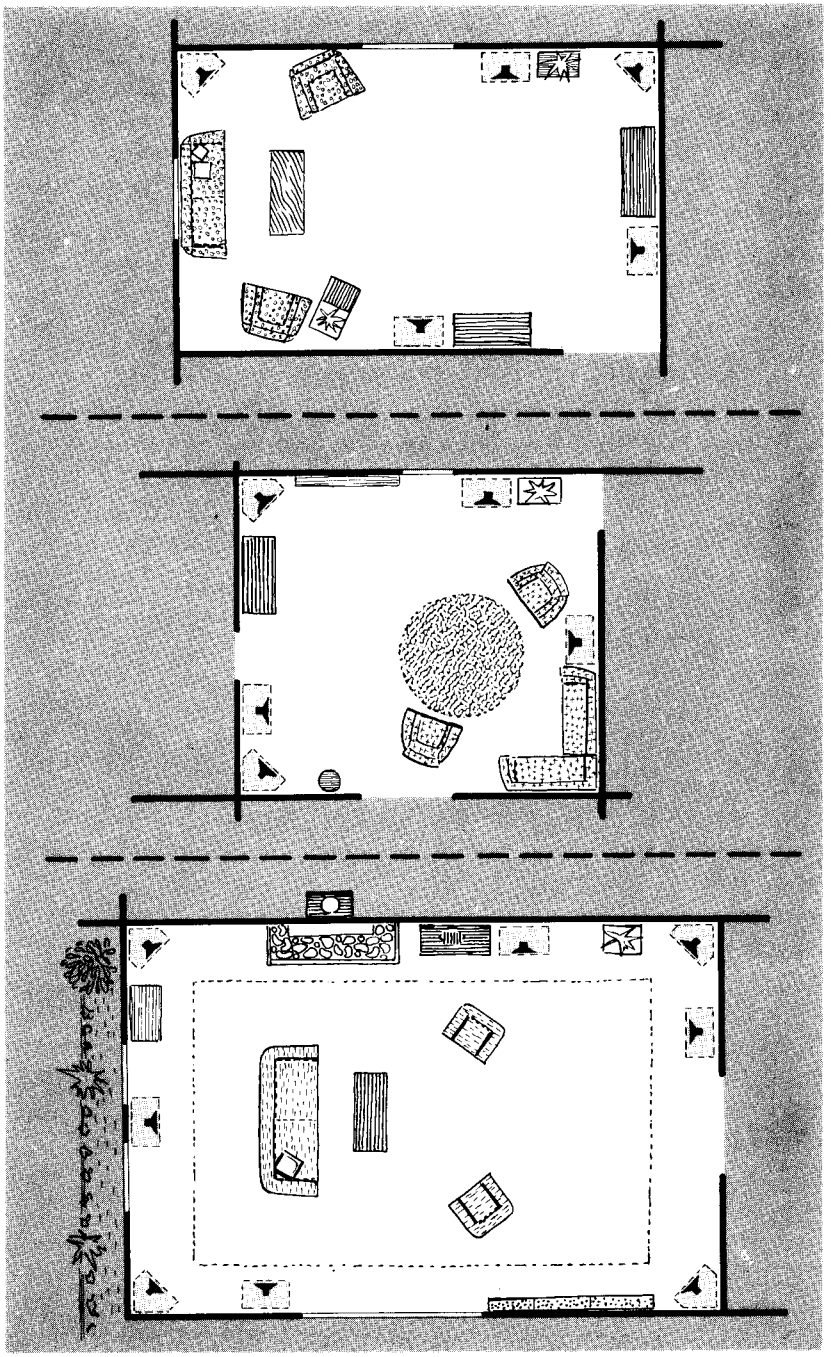
It might be thought that as I am a professional audio engineer my listening room, music room—call it what you like—has a most imposing array and display of gadgetry. Something like those wonderful montages you see in the magazines from time to

time. Well, it hasn't because I haven't got a music room. If you visited my London apartment,or my country place where I have plenty of room to move around and do things, you wouldn't see any audio equipment at all. It is all carefully hidden away because in my own home I am not interested in the means, just the end product. To me audio is just a room service like central heating. But while very ready to concede that a man is perfectly entitled to cover the whole of a wall with an audio omnibus, there are several disadvantages in doing that, and those disadvantages are rarely pointed out, possibly because they are inimical to the "planned-layout" idea. I suggest, therefore, before you do plan your layout you consider these points.

Any combined radio–phonograph is unusable unless there is a bass cutoff somewhere in the system. It may be the pickup or the speaker but bass cutoff there must be, otherwise acoustic–mechanical feedback will make the whole system unstable. I recall one lecture I gave at MIT and laid out the demonstrating equipment on the very solid bench of the lecture theatre. I was behind the bench and the speaker was facing the audience. I put on my first record and wondered what had gone wrong. In a panic I checked everything I could think of while the audience throughly enjoyed themselves and then took a walk around the bench and saw the speaker cone nearly leaping out of its box. So we put the turntable on chunks of sponge rubber and moved the speaker off the bench, and everything was just what we wanted. The concrete floor did not feed back vibration from the speaker into the turntable. Now you will say that had the turntable been properly spring-mounted to begin with this wouldn't have happened. But I assure you that I have never found any spring mounting that was good for frequencies below about 80 to 100 cycles, and if there is motor rumble added the whole thing becomes impossible to use. So the first way of making a wall-covering installation work is to put in a bass cutoff.

Next, it is very tempting, if you have such a layout, to put the speaker into its compartment so that it is nicely tucked away. There is only one place, as a rule, in a room where the speaker sounds best. It can be found by trial and error (Fig. 1204) as it is almost impossible to plot the acoustic characteristics of the listening room with ordinary measuring equipment. So the speaker in a suitable enclosure must have a long connecting lead and the thing moved about until the best results are obtained. You can be pretty sure it won't be flat against the wall at the exact point

Fig. 1204. *Possible speaker locations. Speaker positioning must be determined by trial and error.*



you have marked in the layout. But over and above this is the question of where you are going to put the sound that comes from the back of the speaker diaphragm. Are you going to use an infinite baffle, meaning a closed box, or a reflex housing which allows the back sound to come out of a port? Let us consider these alternatives. But first let us consider the simplest way of all of mounting a speaker—on a flat baffle.

Bass and baffles

Fig. 1205 shows the bass lost by using finite baffles of various sizes. If the baffle is "bent" into a box, then the equivalent baffle

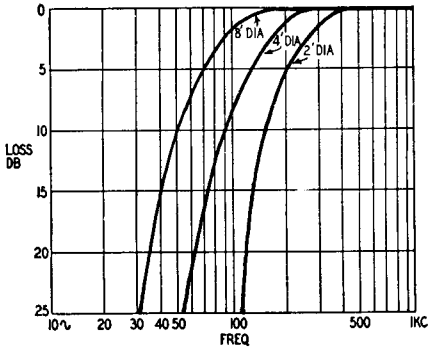
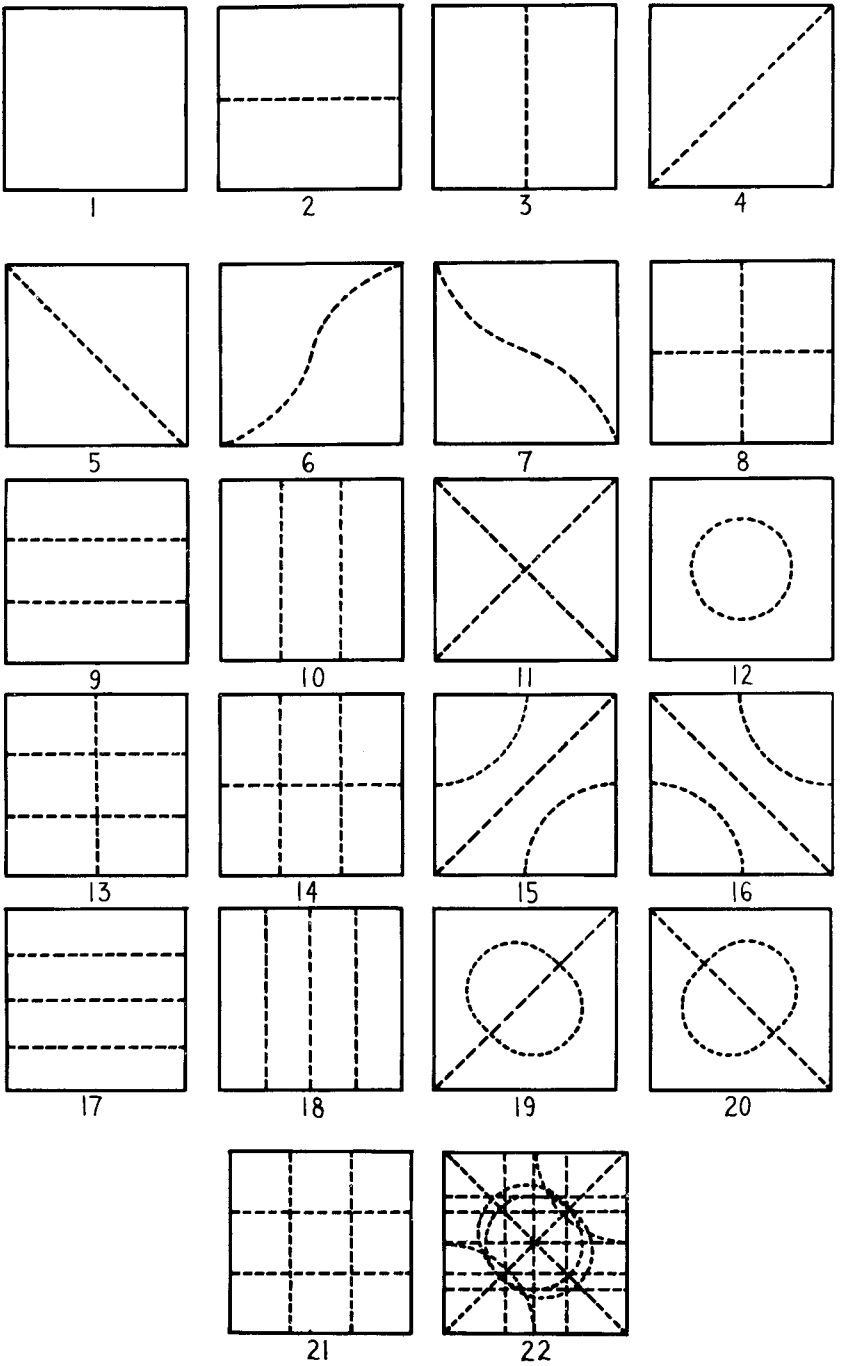


Fig. 1205. *Bass loss with finite baffles.*

diameter is the shortest path from the front edge of the speaker opening across the front and side and the diagonal path from the back edge to the back of the speaker opening. As pointed out in Chapter 10, the minimum size for adequate reproduction of a 50-cycle note is 12 feet, and where can this be stowed away in an ordinary room? Fig. 1205 shows where a flat baffle bends. The fundamental frequency referred to in that figure depends, of course, on the size of the baffle; in other words, the baffle's bass resonant frequency. Bending the baffle into box shape raises the resonant frequency since the edges of the box are stiff, but each side and the front will resonate independently. Hence the suggestion that the best box baffle should be either doubleskinned and the space between filled with sand, or the box should be made of brick or concrete. There are two forms of resonance in a speaker enclosure—resonance of the air within the box and resonance of the box itself. If the box is closed, making it "infinite," the disturbance of the air inside is even greater and will actually interfere with the movement of the diaphragm.

Fig. 1206. *Nodal lines in a flat baffle: 1) fundamental with no node, baffle being rigid for power applied to the speaker; 2–7) 2nd harmonic nodes; 8) 3rd harmonic nodes; 9–12) 4th harmonic nodes; 13–16) 5th harmonic nodes; 17–20) 6th harmonic nodes; 21) 7th harmonic nodes, 22) complex pattern formed by 2nd to 7th harmonic nodes.*

→


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22

The flexing of a speaker baffle is a function of the power applied to it by the power fed into the speaker. Harmonics of the natural fundamental resonant frequency of the baffle develop along nodal lines in the baffle (Fig. 1206). If the fundamental is 60 cycles the second harmonic is 120, the third 180 and so on. The baffle would node something like the last sketch in the figure shown on page 215 with all harmonics present up to the seventh—a nasty sight. If the flat baffle is formed into a cubical box, the front and all the sides will node in a somewhat similar manner for each of the elements of the box is itself a membrane. The front, however, is fixed on four edges, the four sides on three edges. If the box is a closed box, all six membranes are fixed on four edges. The nodal patterns will be identical in all six sides with a closed box, but with an open box the four sides will equal each other but differ from the nodes of the front. How much effect this can have on smoothing the overall response is almost impossible to estimate; a rectangular box will give different nodal patterns for front, top and bottom and sides.

All this virtually implies that a speaker with really first-class bass response down to the lowest frequencies is almost impossible to use in an enclosure, and the only practicable baffle is a brick wall with a hole in it. And that is really the ideal way of mounting a speaker, even if the back sound is wholly lost—except in the next room. But fortunately there actually are ways of housing such a speaker, and the method I evolved (shown in Fig. 1207) is the device I christened the "Boffle." This is a stage-by-stage absorption of the sound from the back of the speaker diaphragm, the essential feature being that slabs of air alternate with slabs of sound-absorbing material. As the sound-absorbing material is carried round the frames that support it and jammed between the frames and the sides of the box, the box is damped too. All that comes out of the back, which is not closed, is a low-pitched rumble severely attenuated. Recent development work has indicated that carefully calculated filters can be arranged so that a delay is applied to the escaping bass sound to bring it into phase with what comes from the front of the speaker, so reinforcing the bass. Unfortunately at the time of writing I have insufficient data to be able to give a practical design. It will come as soon as I have satisfied myself that it does what I want it to do.

It will be clear that the straightforward flat or box baffle can't be used in that wall montage because there is no place for the back sound to go. A reflex enclosure has a port which lets it out

so that type could be used. The design data has been given in Chapter 10; here I need only say that, if the enclosure is properly designed, then the effect the enclosure has on the speaker is to reduce the peak in the impedance curve indicating bass resonance by replacing it with two other peaks of lesser magnitude one above and one below the resonant frequency. The size and port dimensions of the enclosure are dependent on the size of the speaker and its bass resonant frequency, and to get the desired result the design *must* be right. It follows that you can't just go into a store, buy a speaker unit you rather fancy and a reflex housing that appeals to your eye and your pocket and put the two together. They may be completely unmatched. Yet people do that every day and once again they are free to do it if they feel that way, but the result is not high fidelity. And whether the matching is right or wrong, it is perfectly obvious that the speaker in its enclosure is a system of resonators. I have always been of the opinion that a high-fidelity reproducing system must be absolutely and completely acoustically inert. There are so many things that can't be inert, as I have already indicated. Why make life more difficult by employing resonators?

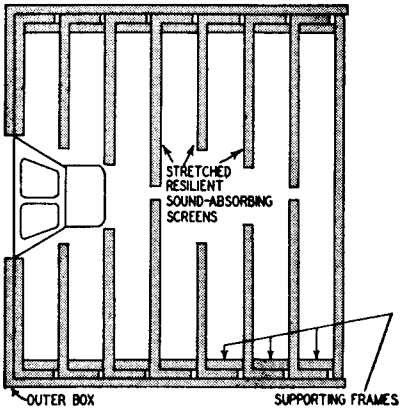


Fig. 1207. *The Hartley Boffle. Each set of four sound absorbing screens forms an acoustic filter, the example shown comprising a 2-stage filter. Note that sound absorbing material is nipped between the screen frames and sides of the box to damp out resonances in the box sides and adding the frame thickness to the box wall thickness.*

I sincerely think that the large entirely complete wall layout system is musically worthless. By all means have one if you want it, as something to play with. But, if your basic aim is to *reproduce* music, then you will not only save a lot of money but get better quality sound if you think in terms of the amplifier being a device that needs no control and so can be hidden away in a cupboard; the radio tuner and the record player (with, of course, the control amplifier) as something that can be put together into quite a small box that can be closed up when not wanted, and the speaker either mounted in the wall or in some enclosure that does not resonate itself.

I have pointed out in Chapter 10 that to get good bass reproduction from a horn-loaded speaker, the horn must be at least 22 feet long and have a flare opening of 24 feet. You can fold the 22 feet of length into something more compact but I don't know of any way of folding the circumference of the flare into something smaller. Another point: I have shown you how a flat or box baffle nodes and resonates, and this can be overcome to some extent by very rigid and strong construction, damped with sound-absorbing material if necessary. Exactly the same applies to a horn, for all the various panels and deflectors will resonate in precisely the same way. But the internal surfaces of the horn must be smooth. As in a folded horn one piece of wood may form both the inside and outside of the horn and as both surfaces can be part of the inside it is impracticable to try to damp out the resonances by applying sound-absorbing material. The wood must be thick and rigid enough not to resonate. I would advise that, when examining any folded-horn enclosure offered to you, thump the assembly in various places with your closed fist. You will not find anything made of wood that is acoustically dead, and the speaker will make it acoustically alive.

This is no condemnation of well-designed speakers properly mounted in folded-horn cabinets. If the cabinet is well made and the speaker has a fairly level response at all frequencies, the combination will sound pretty good except with high power at low frequencies. Then it is extremely difficult to damp out resonances in a complex structure. For this very reason I become more and more convinced that complexity in a high-fidelity sound system is a mistake. The more parts there are, the more the chances of distortion being introduced through something not working as it ought to.

index

D

E

F

G

H

I

L

M

N

O

P

R

S

T

U

V

W

PRINTED IN THE UNITED STATES OF AMERICA