Handbook of Sound Reproduction

EDGAR M. VILLCHUR*

Chapter 10—Part 2. Loudspeakers.

Continuing the discussion of the basic principles of one of the most important elements of a sound system, with a few hints on optimum conditions for speaker operation.

Harmonic Distortion

Non-linearity in loudspeakers is normally much greater than that of any other component in the reproducing chain. While amplifier advertisements vie with one another in comparing fractional distortion percentages at rated output, the best commercial speakers available cannot even approach such excellence. The manufacturer's published distortion *vs.* frequency and output curves of a speaker in the 150 dollar class is shown in *Fig.* 10—7. This speaker is probably one of the top quality units available, but if its distortion curve were presented as a performance index of the most humble of amplifiers, that amplifier would undoubtedly be scorned by the audio market.

A major source of harmonic distortion is the non-linearity of the cone and voice-coil suspensions, particularly at large excursions. As these suspensions are stretched their restraining forces increase instead of remaining constant. —beyond a certain point the suspensions will not give at all without tearing—and cone displacement ceases to be proportional to the magnetomotive force.

Harmonic distortion is also produced if the excursion of the voice coil takes it into a region where the total magnetic flux through which it moves is reduced. The instantaneous voice coil displacement no longer follows the signal because the magnetomotive force at the extreme positions is reduced.

These mechanical sources of distortion generally affect both halves of the cycle equally and the generated harmonics are therefore of odd orders, predominantly the third. Any method for reducing voice-coil excursion without decreasing output will reduce harmonic distortion. The more efficient the coupling to the air the less will be the displacement required of the voice coil for a given radiated acoustical power; the type of speaker mounting used is thus extremely important relative to distortion.

Both of the above types of non-linearity are most prominent in the bass frequencies due to the increased excursion. Voice-coil displacement for the same signal amplitude increases as the frequency is lowered, and below the frequency of ultimate air load resistance displacement increases inversely as the *square* of the frequency. At a given frequency distortion increases with signal amplitude, also because of the increased excursion.

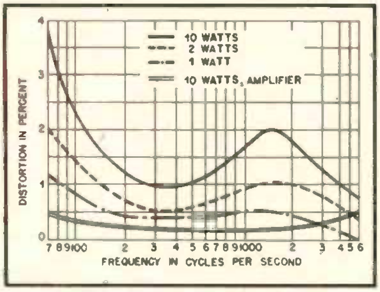

Fig. 10—7. Harmonic distortion of a high quality coaxial speaker, compared to that of an equivalent-quality amplifier.

Subharmonic Distortion

Figure 10—8 illustrates how flexing of the cone in response to voice-coil motion may produce subharmonics. When the voice coil moves forward, a stiff rim suspension may cause the cone to bend in either direction. When the voice coil moves forward a second time, however, the cone is already in transverse motion, returning from its first position of flexure, and the resulting momentum added to the flexing force causes it to bend the opposite way. The cone thus completes one cycle of motion during the time required for the voice coil to complete two. The frequency of flexure of the cone is one half that of the frequency of vibration of the voice coil.

Subharmonic formation is relatively minor in loudspeaker performance, and is discouraged by a highly compliant rim suspension and by cone design which discourages flexure.

Intermodulation

The fact that amplitude distortion is produced by a loudspeaker means that intermodulation will also exist, providing that the different frequencies involved are passed through the same distorting system. In the case of the speaker this intermodulation may be illustrated physically. When the voice coil is driven by a low-frequency note into a position where the suspensions exercise more than normal restraint on motion, or where the magnetic field is

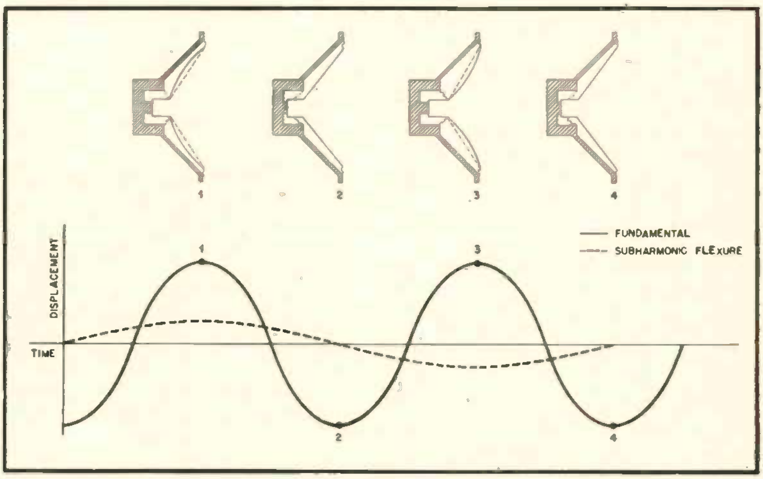

Fig. 10—8. Formation of loudspeaker subharmonics.

* *Contributing Editor,* AUDIO ENGINEERING.

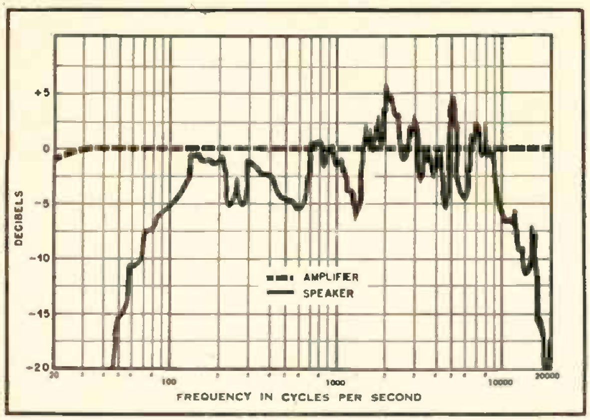

Fig. 10—9. Manufacturer's published on-axis frequency response curve for a high-quality speaker, compared to curve for equivalent-quality amplifier.

weaker, the peaks of the low-frequency signal will be flattened. But the restraint of the suspensions will also limit the excursion of high-frequency vibrations at this moment, and any high-frequency signals will be reproduced with less amplitude during the peaks of the low-frequency signal. Since the peaks are usually flattened during both halves of the cycle, the higher-frequency signal will be amplitude modulated at twice the low-frequency rate. The sidebands, or sum and difference frequencies associated with such modulation, may be expected to be at frequencies equal to that of the higher frequency plus and minus twice the lower frequency.

A second type of intermodulation that occurs in speakers is caused by the Doppler effect. If the cone is stimulated by a low-frequency signal it will at times be approaching and at other times receding from the listener. A high-frequency note superimposed on this slowly oscillating cone will have its pitch, relative to the listener, alternately raised and lowered. This constitutes frequency modulation of the high note by the lower frequency, creating sidebands. The Doppler type of intermodulation will be greatest when a cone with large excursion in the bass is simultaneously used to reproduce the upper treble. The effect, however, is much less serious than that of amplitude modulation, and like the latter it is alleviated by reduced voice-coil excursion.

Speaker systems which assign different portions of the sound-frequency spectrum to separate voice coil and cones discriminate against intermodulation between the distorted signals of each portion.

Frequency Response

The frequency-response curve of a speaker is commonly plotted as sound pressure created by the speaker (in dynes/cm^2 converted to db) *vs.* the frequency of a constant-amplitude signal input. It is a very difficult measurement to make, since it involves the acoustical conditions of the space into which the speaker radiates. Speaker frequency-response curves are often more indicative of general trends than exact performance, and can only expect to be duplicated under the same acoustical and electrical conditions that existed when the curve was made.

The frequency response of the best speakers is, as might be expected, far more erratic than the response of electronic amplifiers. *Figure* 10—9 is the manufacturer's published on-axis frequency-response curve for a typical high-quality speaker, compared to the curve for an equivalent-quality commercial amplifier. It is common practice for speaker frequency response to be described numerically, as the upper and lower frequency limits which the speaker reproduces. Variation within these limits is sometimes as much as plus or minus 10 db or more.

From the point of view of listening, more variation in the tonal color of reproduction may be expected in changing from one make or type of speaker system to another than in changing any of the other audio components. In speakers of the same general price range this is usually not so much a function of the range of frequencies reproduced as it is dependent upon the emphasis and deficiencies at different points on the response curve. The greatest evils of this erratic response are associated with bass resonance, tending to produce boominess, and cone break-up resonances in the low highs, tending to produce shrillness. At its worst, accentuated resonant response creates a situation where signals anywhere near resonance set the speaker to sounding its lone, prolonged note in chorus. Low-frequency signals may become loud thumps or roars whose pitch is difficult to distinguish.

For the speaker to radiate low frequencies efficiently, large masses of air must be moved. The excursion of the cone has a limited range, and so systems designed for good low-frequency performance generally take full advantage of what motion there is by having a large cone area in contact with the air. (Several twelve- or fifteen-inch speakers in parallel may even be used for extreme low-frequency reproduction.) A large cone and extended travel calls for a large voice coil. The voice coil must be long enough to utilize efficiently the entire magnetic field over the extended path, and of sufficient radius for rigid coupling to the cone. Thus the mass of the moving system, when designed from the viewpoint of low-frequency reproduction, will be comparatively high. This is not particularly disadvantageous in the low range.

For high-frequency reproduction it is especially important that the mass of the moving system be low. Small, rigid cones or diaphragms and small, light voice coils are therefore suitable. The excursion of the voice coil will be greatly reduced, because of the inverse relationship between frequency and displacement, and also because of the fact that typical program material has somewhat less energy content in the high-frequency range. The magnetic field can therefore be concentrated over a smaller area, and the required size of the magnet structure is less.

The contradictory requirements of low- and high-frequency reproduction may be compromised in a single speaker unit, or two or three speakers can be used, each designed for optimum performance within its frequency range. Horn loading is often used for the high range because of the reduction in size and weight of the vibrating diaphragm made possible by the more efficient coupling to the air, and because the size of the required horn is conveniently small.

Transient Response

The accuracy with which a speaker can follow sudden starts, stops, and changes of the electrical signal is directly proportional to the damping and to the evenness of the frequency response.

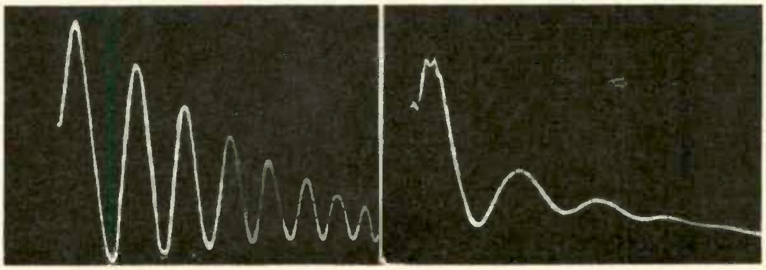

Fig. 10—10 (A) left. Impedance and damping characteristics of undamped speaker (8-inch). The oscillogram represents back e.m.f. produced by the voice coil for instantaneous d.c. stimulation, removed after the first quarter-cycle. (B) right. Same as (A), using edge-damped loudspeaker (12-inch W. E. 728B).

High-frequency transients will, of course, require good high-frequency response, though the transient may be associated with lower frequencies.

A type of transient distortion commonly associated with loudspeakers is hangover, the tendency of the speaker cone to continue to vibrate after the signal has stopped. Hangover is especially likely to occur at a signal frequency to which some part of the mechanical or acoustical system of the speaker and enclosure exhibit sympathetic resonance, and where, as a result, there is a peak in the frequency-response curve. The effect is that of smearing one note onto the next, destroying the clarity and distinctness of the various voices of the program material.

Mechanical, acoustical, and electrical damping all reduce hangover. *Figure* 10—10 illustrates the effect of internal speaker damping, achieved by coating the rim suspension with a viscous material. (Mechanical friction changes with velocity, while the viscous friction of the coating tends to be independent of velocity, as pure electrical resistance is independent of current.) Acoustical and electrical damping, upon which most speaker systems primarily rely, involve characteristics of the associated mounting device and amplifier as well as of the speaker itself. Low amplifier output impedance presents a heavy electrical load to the speaker as a generator, quickly bringing unauthorized motion to a halt. High magnetic flux in the voice coil gap, which produces a strong back e.m.f., aids this damping action.

Power Capability

Speakers are rated as to their ability to handle steady power within given distortion percentages, and also as to their power capacity for peaks of short duration. These ratings refer only to input electrical power and not to output acoustical power. If the input power is quite a bit less than the maximum power rating of the speaker, distortion is less, indicating that the use of a speaker system over-rated as to power is advantageous. When several speaker units are connected together in a properly matched multispeaker system the power capacities of each unit are added together for the total power rating.

Speaker Efficiency

The efficiency of a speaker is defined as the output acoustical power divided by the input electrical power. Direct-radiator loudspeakers have efficiencies ranging from about 2 to 7 per cent. For an input of 10 electrical watts a 3-per-cent efficient speaker will radiate 0.3 acoustical watts, while a 6-per-cent efficient speaker will require only 5 watts of electrical input to produce the same acoustical output. It is evident that the power rating of the speaker must be considered in terms of speaker efficiency; an efficient 10-watt speaker may be able to handle more *acoustical* power than an inefficient 20-watt unit.

Important determinants of speaker efficiency are the strength of the magnetic field in the gap relative to the mechanical impedance of the moving system, and the length of wire in the voice coil. Speaker efficiency is also highly dependent upon the acoustical coupler—horn couplers may increase efficiency to from 25 to 50 per cent. A common "intuitive" fallacy is the assumption that large speakers require more electrical driving power than small ones for the same acoustical output. The opposite is usually true.

Fig. 10—11. Typical speaker frame and voice-coil suspension. (Courtesy Stephens Mfg. Corp.)

Although speaker efficiency has no direct bearing on performance (especially when internal viscous damping is used) the same construction features that make for high electro-mechanical efficiency create good electrical damping. An efficient speaker system also allows the amplifier to operate at a lower electrical power level for the same sound power. Good mechanico-acoustical efficiency is always advantageous, as a given amount of acoustical energy can be radiated with less voice-coil excursion and therefore less distortion.

Speaker efficiency may be simulated by an apparent loudness created by resonances and distortion in the mid-frequency range.

Radiation Pattern

The radiation pattern of a vibrating loudspeaker cone is determined by the combination of effects of the circular ring elements composing this cone. (See Chapter 3.) A given point in space will be subject to sound radiation from all of the adjacent vibrating ring sources. The additive or cancelling effects of the sound from these sources will be affected not only by the relative position of each ring, but also by the time delay involved in the travel along the cone from the voice coil to the ring element.

Fig. 10—12. Voice coil and seamless molded cone for use in woofer. (Courtesy Stephens Mfg. Corp.)

The smaller a speaker cone is in relation to the wave length being radiated the less cancellation will result from radiation from different points on the ring elements, and the broader the radiation pattern will be. A large speaker, therefore, has a very directive radiation pattern concentrated around the axis at the high frequencies. The use of small treble speakers, or of separate small cones or diaphragms for treble reproduction, broadens the high-frequency radiation pattern.

High-frequency sound radiated from the center ring elements of the cone travels along two paths, that of the air and that along the sides of the cone itself. The velocity of travel in the paper cone is on the order of two times the velocity in air, and so the phase of sound from the outer rings—relative to that from the center—is affected by both the angle of the cone and the velocity of sound in the particular cone material. Factors which tend to delay radiation from the rim relative to that from the center broaden the radiation pattern. Shallow cones therefore have a broader pattern than deep ones of the same diameter. The flared cone, due to increased rigidity, has a sharper pattern than that of the conical cone. The use of corrugations and of relatively soft cone material, both of which slow up sound propagation, broaden the angle of sound radiation.

Special procedures and devices to broaden high-frequency radiation are often employed. More than one high-frequency speaker may be used, and the units arranged in an arc. A second method of diffusion is to load the high frequency radiator acoustically with a multicellular horn whose individual mouths are fanned out. Whatever method is used, the results are stated most rigorously by a polar graph which relates angle of radiation, frequency, and the sound intensity level.

LOUDSPEAKER CONSTRUCTION

The Speaker Frame

The main requirement of a good speaker frame is rigidity and absence of resonant behavior. To these ends it is heavy and sometimes made of cast metal. A typical commercial speaker frame (with voice coil suspension in place) appears in *Fig.* 10—11.

Rim Suspensions

The most common type of rim suspension is that employing a corrugation in the one-piece cone, as in *Fig.* 10—12. Another effective but expensive method makes use of a separate rim of kidskin, cloth, or soft leather.

Voice-Coil Suspensions

The type of centering suspension in current general use is the corrugated disc illustrated in *Fig.* 10—11. Formerly the most popular type was the slotted disc, whose appearance gave the voice coil suspension the name of "spider."

Cones

Although cheaper cones are sometimes made by rolling and glueing a paper development of the cone form, the best type is moulded in one unbroken section. (See *Fig.* 10—12.) A mixture of pulp and water is drawn through a master screen in the shape of the cone, leaving a deposit of pulp which is removed when dry. Hard, springy cone material may create an impression of greater volume due to more accentuated break-up resonances and increased distortion.

Voice Coils

The design of voice coils is concerned with factors of mass, resistance, and efficient and linear use of the narrow air gap.

Low resistance is most important in low-frequency speakers, for reasons of efficiency and electrical damping (the d.c. resistance of the wire appears in series with the source impedance of the amplifier), and copper wire is therefore commonly used. The added mass is not an important consideration in woofers. Speakers which must also reproduce high frequencies, however, require as light a voice coil as possible, and therefore often make use of higher resistance aluminum wire for their voice coils.

The most efficient use of the air gap is made by square wire or by edgewise wound ribbon wire. The dividend of additional use of the available space may be taken in lower resistance or in reduced thickness for the same resistance.

The excursion of the voice coil must not take it into a region where the magnetic flux through which it moves is reduced. If the voice coil is made longer than the gap length the average field strength affecting the voice coil will be the same, because as one end of the coil moves into a weaker part of the field the other end is moving into a stronger field area. Uniformity of gap flux may also be maintained by a system which makes the voice coil shorter than the length of gap. If the excursions of the voice coil are small it will remain within the strong uniform section of the field at all times.

The system which uses the smaller voice coil is more appropriate for high-frequency reproduction, and the one with the larger voice coil is suited for low frequencies. These two systems, contradictory to one another in a speaker with a single voice coil, may both be used in a dual or multiple unit.

Multispeaker Circuits

Speakers may be connected in series, parallel, or series-parallel, and the total nominal impedance is calculated by the same method that is used for simple resistive networks. For example, two 8-ohm speakers in series must be fed from the 16-ohm amplifier tap, and a second such 16-ohm speaker combination, connected in parallel with the first, brings the total impedance down to 8 ohms again.

Where remote speakers are used it is occasionally desirable to furnish separate volume controls for the individual speakers. This may be accomplished by an L-type level control, illustrated in

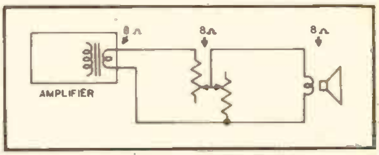

Fig. 10—13. L-type level control for attenuating volume at speaker.

Fig. 10—13, in which the two sections are tapered in such a way that the impedance presented to the speaker source remains relatively constant. Attenuation is accomplished by wasting power in the control, and its power rating should be as great as that of the amplifier. A simple heavy-duty potentiometer, with a resistance several times the nominal impedance of the speaker, may also be used, but at the expense of impedance mismatching.

Bass-treble combinations, poetically referred to as "woofer-tweeter" systems, require some sort of dividing network to separate the different portions of the frequency spectrum. The simplest of these networks is shown at (A) in *Fig.* 10—14. The tweeter is protected from damage by low-frequency signals, but treble signals are fed to both woofer and tweeter. Cone break-up and intermodulation in the woofer is thus discouraged only to the extent that part of the treble signal is by-passed through the tweeter.

The more complicated but more effective inductance-capacitance networks of (B) and (C) in *Fig.* 10—14 produce a much greater separation between bass and treble. Here a roll-off is introduced at the top of the woofer range and at the bottom of the tweeter range, at the rate of 6 db per octave (in terms of power) for the single capacitor-inductor network, and 12 db per octave for the double element network. The *crossover* frequency is that frequency at which the value of the various elements, including the speakers themselves, cause the input power to divide equally between the two speakers. Since the power to each speaker is halved at this point, the attenuation in both treble and woofer circuits is 3 db.

The equations for calculating the values of L and C for given crossover frequencies appear below the diagrams. The equation for the R-C network is a mathematical statement of the fact that the impedance of the capacitor is equal to that of the speaker at crossover. The impedance of the combination is treated as the impedance of the woofer.

The equations for the single element L-C network are derived from expressions which state that the impedance of the capacitor is equal to that of the tweeter at crossover, and that the impedance of the inductor is equal to that of the woofer at crossover:

$$\frac{1}{2\pi fC} = Z_{tweeter} \qquad 2\pi fL = Z_{woofer}$$

The impedance of the combination is treated as the impedance of one speaker, if woofer and tweeter have the same value, or as the average between woofer and tweeter impedances if these differ.

(*Continued on page 63*)

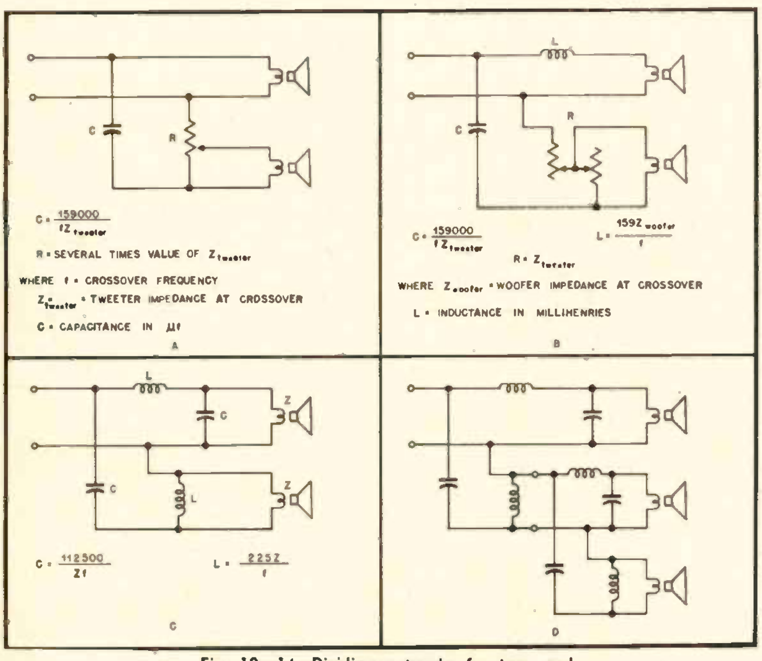

Fig. 10—14. Dividing networks for two- and three-way speaker systems.

SOUND HANDBOOK

(from page 32)

The equations for the dual element L-C network are derived from expressions which state that the impedance of each capacitor and each inductor is equal to the impedance of each voice coil at crossover, multiplied by the square root of two:

$$\frac{1}{2\pi fC} = \sqrt{2}\,Z \qquad 2\pi fL = \sqrt{2}\,Z$$

The impedance of the combination is treated as the impedance of one speaker.

When the woofer and tweeter voice coils have different values, and the circuit of (C) in *Fig.* 10—14 is used, the values of L and C associated with each speaker must be calculated independently. In such a case the rated input impedance of the whole network, as in the case of the single-element network, is treated as the average between the rated impedances of the two speakers.

Where the upper range is again divided up between speakers the network of (D) in *Fig.* 10—14 is used. This circuit is derived from that of *Fig.* 10—15 (C); the input leads to the latter's tweeter are treated as the source for the new two-way system of middle and high range speakers. Using the same equations as for the original two-way system the values for the sub-network may be calculated independently, on the basis of the new crossover. The total impedance is still the average.

The value of speaker impedance used for all of the above calculations will probably be fairly close to the nominal impedance, depending on the crossover frequency used; the value for Z at particular frequencies may be estimated, requested from the speaker manufacturer, or measured as in *Fig.* 10—15. The variable resistor in this test set-up is adjusted until the voltmeter indications are the same in both positions.

The crossover frequency is selected on the basis of the characteristics of each speaker. The advantages of a low crossover frequency are that intermodulatory products between the bass signals of probable greatest distortion, and the higher-frequency signals are avoided, the woofer can operate over the frequency range in which it moves rigidly, without break-up, and the radiation pattern of higher frequencies is not restricted by the large diameter woofer. The disadvantages are that the crossover network becomes bulkier and more expensive, and that the responsibilities placed upon the tweeter become much greater. Crossover points from several hundred to several thousand cps have been used successfully. When the crossover point occurs at a frequency region in which the output of one or both of the speakers is significantly accentuated (a common situation where the woofer is used for low treble signals) this condition may be relieved by a suitable shift of the roll-off of that speaker away from the crossover point. This is achieved simply by calculating L and C of the affected branch on the basis of the shifted crossover. Similarly, a

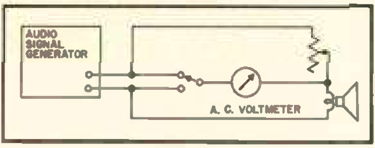

Fig. 10—15. Method of measuring unknown speaker impedance at a given frequency. The speaker impedance is numerically equal to the value of the resistance when the voltage drops across each are equal.

significant "hole" in the response curve may be filled in by a corresponding shift in the opposite direction.

An array of speakers assigned to different portions of the frequency spectrum may produce an unbalanced tonal structure due to varying efficiency from one unit to the other. Horn-loaded tweeters used alongside direct-radiator woofers, for example, normally require provision for attenuating the treble signal, and a horn loaded woofer should be given a correspondingly efficient tweeter. An L-type level control between the speaker (normally the tweeter) and the network feeding it [See (B) in *Fig.* 10—14] accomplishes this result without affecting operation otherwise.

The coils of the dividing network must have high power-handling capacity, and must not introduce non-linearity into the circuit. Non-metallic cores, which avoid the non-linear and power dissipating effects of core saturation and eddy currents, are suited to such requirements, and are made possible by the low values of inductance called for. Winding data appears at (A) and (B) in *Fig.* 10—16.

The capacitors must be non-polarized since they are in an a.c. circuit, and should have a working voltage rating of about 50 volts for safety. Although non-polarized electrolytic capacitors (two electrolytics connected "back-to-back") have been used in dividing networks, this practice is not recommended because of the lower reliability of electrolytics from the point of view of accuracy of rated value, changes of value with time, and failure in service. Good-quality paper or oil-filled capacitors are suitable.

Experimental Trends in Speaker Design

The logical extension of the use of negative feedback in amplifiers is to include the speaker moving system within the feedback loop. If the source of feedback voltage is an independent generating coil in the loudspeaker, inaccuracies in voice-coil motion as well as electrical inaccuracies in the amplifier will be corrected. Such a design was patented as long ago as 1925,[2] but has not, to date, been brought to commercial practicability. A successful feedback loudspeaker would go far in bringing instantaneous voice-coil velocity to a closer relationship with the electric signal. (The feedback principle has already been applied, with great success, to disc recording heads.) Paradoxically, however, the acoustic output of a voice coil with constant-velocity response would exhibit bass losses, due to the progressive drop of air-load resistance at lower frequencies. The shape of the acoustic frequency-response curve at the low end, corresponding to that of the graph of air-load resistance *vs.* frequency, would require that compensatory bass boost be introduced, and that the amplifier have a power capability sufficient to handle this extra boost.

A revolutionary principle of speaker design has recently been introduced in France. The "ionic" loudspeaker is a direct electro-acoustic device that bypasses the usual electro-mechanical stage. Agitation of the air molecules is produced by an electrostatic field rather than by the mechanical pushing and pulling of an intermediary cone or diaphragm, and since no mechanical moving parts are necessary the entire suspension system—the weakest link in loudspeaker design—may be eliminated.

The molecules which are controlled electrostatically are in an acoustical chamber which contains the stimulating electrodes. Since air molecules in their normal, uncharged state are insusceptible to the influence of electric fields, the particles in the chamber are first ionized, or given a positive charge. This is done by heating the chamber, increasing random molecular movements and inducing collisions which knock off orbital electrons from some of the molecules. Once there are a few charged particles, an alternating field of supersonic frequency is able to complete the ionization process by more of these electrically created collisions.

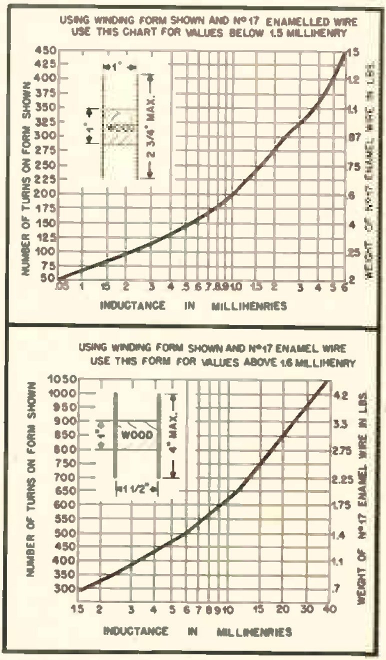

Fig. 10—16. Winding data for dividing network coils. Brass screws are used for the coil form. (From "Crossover Networks for Speaker Systems," Courtesy University Loudspeakers, Inc.)

[2] J. P. Maxfield, *et. al.*, U. S. Patent No. 1,535,538.

The signal to be reproduced is applied as an amplitude-modulating voltage to the supersonic field, and the acoustical vibrations are coupled to the room through a horn. The field is of radio frequency, a fact which introduces an interesting feature. The output of an AM radio receiver requires only sufficient r.f. amplification before it is applied to the speaker, dispensing with the need for a detector and audio amplification stages.

References

John G. Frayne, and H. Wolfe, *Elements of Sound Recording,* Chapter 30.

Jensen Mfg. Co., *Impedance Matching and Power Distribution in Loudspeaker Systems,* Technical Monograph No. 2.

H. F. Olson, *Elements of Acoustical Engineering,* Chapter 6.