CHAPTER

3

The Ear and the Perception of Sound

The study of the structure of the ear is a study in physiology. The study of human perception of sound comes under the general heading of *psychology*. *Psychoacoustics* is an inclusive term embracing the physical structure of the ear, the sound pathways, the perception of sound, and their interrelationships. Psychoacoustics, quite a recent term, is especially pertinent to this study because it emphasizes both structure and function of the human ear.

The stimulus sound wave striking the ear sets in motion mechanical movements that result in neuron discharges that find their way to the brain and create a sensation. Then comes the question, "How are these sounds recognized and interpreted?" In spite of vigorous research activities on all aspects of human hearing, our knowledge is still woefully incomplete.

Sensitivity of the Ear

The delicate and sensitive nature of our hearing can be underscored dramatically by a little experiment. A bulky door of an anechoic chamber is slowly opened, revealing extremely thick walls, and three-foot wedges of glass fiber, points inward, lining all walls, ceiling, and what could be called the floor, except that you walk on an open steel grillwork.

A chair is brought in, and you sit down. This experiment takes time, and as a result of prior briefing, you lean back, patiently counting the glass fiber wedges to pass the time. It is very eerie in here. The sea of sound and noises of life and activity in which we are normally immersed and of which we are ordinarily scarcely conscious is now conspicuous by its absence.

The silence presses down on you in the tomblike silence, 10 minutes, then a half hour pass. New sounds are discovered, sounds that come from within your own body. First, the loud pounding of your heart, still recovering from the novelty of the situation. An hour goes by. The blood coursing through the vessels becomes audible. At last, if your ears are keen, your patience is rewarded by a strange hissing sound between the "ker-bumps" of the heart and the slushing of blood. What is it? It is the sound of air particles pounding against your eardrums. The eardrum motion resulting from this hissing sound is unbelievably small—only $\frac{1}{100}$ of a millionth of a centimeter—or $\frac{1}{10}$ the diameter of a hydrogen molecule!

The human ear cannot detect sounds softer than the rain of air particles on the eardrum. This is the threshold of hearing. There would be no reason to have ears more sensitive, because any lower-level sound would be drowned by the air-particle noise. This means that the ultimate sensitivity of our hearing just matches the softest sounds possible in an air medium. Accident? Adaptation? Design?

At the other extreme, our ears can respond to the roar of a cannon, the noise of a rocket blastoff, or a jet aircraft under full power. Special protective features of the ear protect the sensitive mechanism from damage from all but the most intense noises.

A Primer of Ear Anatomy

The three principal parts of the human auditory system, shown in Fig. 3-1, are the outer ear, the middle ear, and the inner ear. The outer ear is composed of the pinna and the auditory canal or auditory meatus. The auditory canal is terminated by the tympanic membrane or the eardrum. The middle ear is an air-filled cavity spanned by the three tiny bones, the ossicles, called the malleus, the incus, and the stapes. The malleus is attached to the eardrum and the stapes is attached to the oval window of the inner ear. Together these three bones form a mechanical, lever-action connection between the air-actuated eardrum and the fluid-filled cochlea of the inner ear. The inner ear is terminated in the auditory nerve, which sends impulses to the brain.

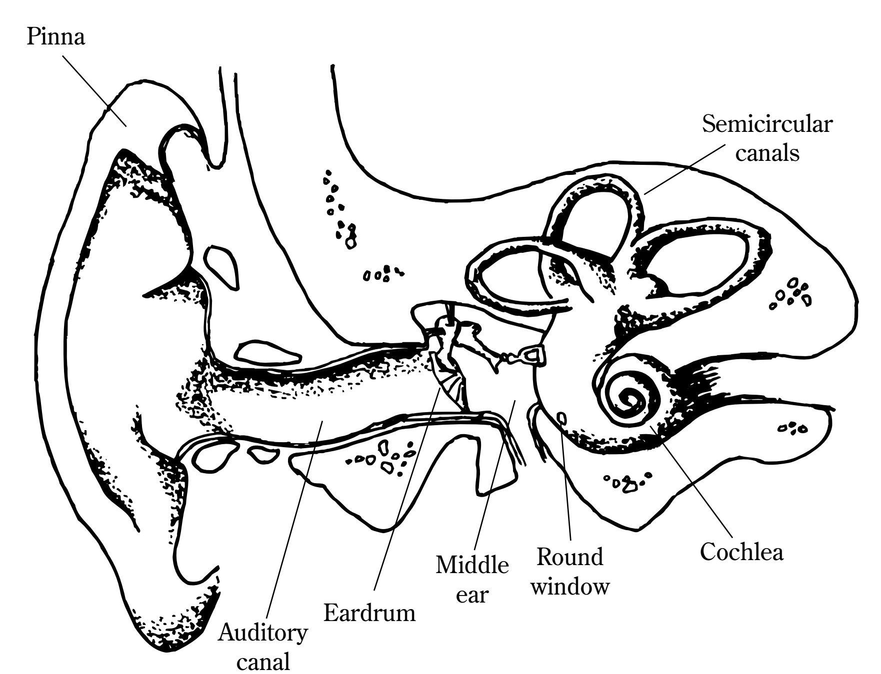

FIGURE 3-1

The four principal parts of the human ear: the pinna, the auditory canal, the middle ear, and the inner ear.

The Pinna: Directional Encoder of Sound

In ancient times, the pinna was regarded as either a vestigial organ or a simple sound-gathering device. True, it is a sound-gathering device. The pinna offers a certain differentiation of sounds from the front as compared to sound from the rear. Cupping your hand behind the ear increases the effective size of the pinna and thus the apparent loudness by an amount varying with frequency. For the important speech frequencies (2,000 to 3,000 Hz), sound pressure at the eardrum is increased about 5 dB. This front-back differentiation is the more modest contribution of the pinna.

Recent research has revealed that the pinna performs a very crucial function in imprinting directional information on all sounds picked up by the ear. This means that information concerning the direction to the source is superimposed on the sound content itself so that the resultant sound pressure on the eardrum enables the brain to interpret both the content of the sound and the direction from which it comes.

Directional Cues: An Experiment

If the equipment is available, a simple psychoacoustical experiment can illustrate how subjective directional impressions result from simple changes in sounds falling on the ear. Listen with a headphone on one ear to an octave bandwidth of random noise centered on 8 kHz arranged with an adjustable notch filter. Adjusting the filter to 7.2 kHz will cause the noise to seem to come from a source on the level of the observer. With the notch adjusted to 8 kHz the sound seems to come from above. With the notch at 6.3 kHz the sound seems to come from below. This experiment[1] demonstrates that the human hearing system extracts directional information from the shape of the sound spectra at the eardrum.

The Ear Canal

The ear canal also increases the loudness of the sounds traversing it. In Fig. 3-2 the ear canal, with an average diameter of about 0.7 cm and length of about 3 cm, is idealized by straightening and giving it a uniform diameter throughout its length. Acoustically, this is a reasonable approximation. It is a pipe-like duct, closed at the inner end by the eardrum.

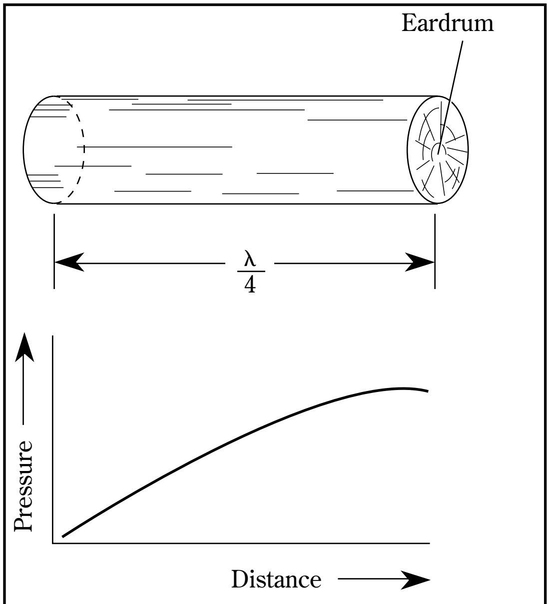

FIGURE 3-2

The auditory canal, closed at one end by the eardrum, acts as a quarter-wavelength "organ pipe." Resonance provides acoustic amplification for the important voice frequencies.

Organ pipes were studied intensely by early investigators when the science of acoustics was in its infancy. The acoustical similarity of this ear canal to an organ pipe was not lost on early workers in the field. The resonance effect of the ear canal increases sound pressure at the eardrum at certain frequencies. The maximum is near the frequency at which the 3-cm pipe is one-quarter wavelength—about 3,000 Hz.

Figure 3-3 shows the increase in sound pressure at the eardrum over that at the opening of the ear canal. A primary peak is noted around 3,000 Hz caused by the quarter-wave pipe resonance effect. The primary pipe resonance amplifies the sound pressure at the eardrum approximately 12 dB at the major resonance at about

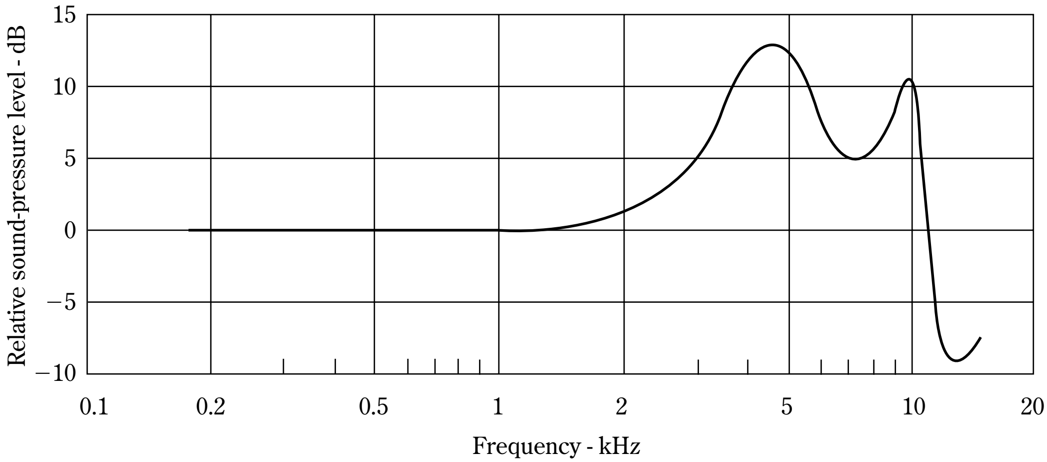

FIGURE 3-3

The transfer function (frequency response) of the ear canal. This is a fixed component that is combined with every directionally-encoded sound reaching the eardrum. See also Figs. 3-15 and 3-16. (After Mehrgardt and Mellart.[2])

4,000 Hz. There is a secondary resonance nearer 9,000 Hz of lower peak pressure.[2]

The Middle Ear

Transmitting sound energy from a tenuous medium such as air into a dense medium like water is a serious problem. Without some very special equipment, sound originating in air bounces off water like light off a mirror. It boils down to a matter of matching impedances, and in this case the impedance ratio is something like 4,000:1. Consider how satisfactory it would be to drive the 1-ohm voice coil of a loudspeaker with an amplifier having an output impedance of 4,000 ohms! Clearly not much power would be transferred.

The object is to get the feeble energy represented by the vibratory motion of a rather flimsy diaphragm, transferred with maximum efficiency to the fluid of the inner ear. The two-fold solution is suggested in Fig. 3-4. The three ossicles (hammer, anvil, and stirrup) form a mechanical linkage between the eardrum and the oval window, which is in intimate contact with the fluid of the inner ear. The first of the three bones, the malleus, is fastened to the eardrum. The third, the stapes, is actually a part of the oval window. There is a lever action in this linkage with a

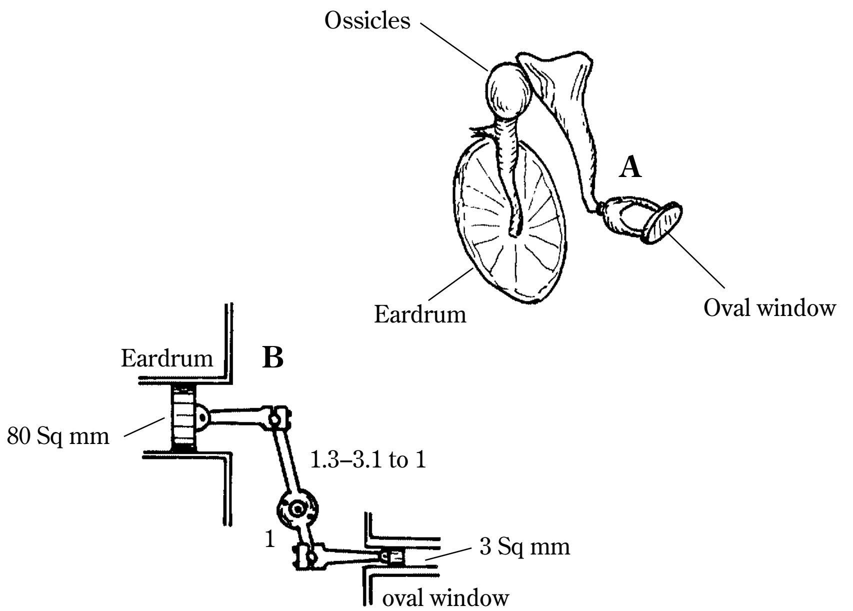

FIGURE 3-4

(A) The ossicles (hammer, anvil, and stirrup) of the middle ear, which transmit mechanical vibrations of the eardrum to the oval window of the cochlea. (B) A mechanical analog of the impedance-matching function of the middle ear. The difference in area between the eardrum and the oval window, coupled with the step-down mechanical linkage, match the motion of the air-actuated eardrum to the fluid-loaded oval window.

ratio leverage ranging from 1.3:1 to 3.1:1. That is, the eardrum motion is reduced by this amount at the oval window of the inner ear.

This is only part of this fascinating mechanical-impedance-matching device. The area of the eardrum is about 80 sq mm, and the area of the oval window is only 3 sq mm. Hence, a given force on the eardrum is reduced in the ratio of 80/3, or about 27-fold.

In Fig. 3-4B, the action of the middle ear is likened to two pistons with area ratios of 27:1 connected by an articulated connecting rod having a lever arm ranging from 1.3:1 to 3.1:1, making a total mechanical force increase of between 35 and 80 times. The acoustical impedance ratio between air and water being on the order of 4,000:1, the pressure ratio required to match two media would be $\sqrt{4{,}000}$, or about 63.2, and we note that this falls within the 35 to 80 range obtained from the mechanics of the middle ear illustrated in Fig. 3-4B.

The problem of matching sound in air to sound in the fluid of the inner ear is beautifully solved by the mechanics of the middle ear. The evidence that the impedance matching plus the resonance amplification of Fig. 3-3 really work is that a diaphragm motion comparable to molecular dimensions gives a threshold perception.

A schematic of the ear is given in Fig. 3-5. The conical eardrum at the inner end of the auditory canal forms one side of the air-filled middle ear. The middle ear is vented to the upper throat behind the nasal cavity by the Eustachian tube. The eardrum operates as an "acoustic suspension" system, acting against the compliance of the trapped air in the middle ear. The Eustachian tube is suitably small and constricted so as not to destroy this compliance. The round window separates the air-filled middle ear from the practically incompressible fluid of the inner ear.

The Eustachian tube fulfills a second function by equalizing the static air pressure of the middle ear with the outside atmospheric pressure so that the eardrum and the delicate membranes of the inner ear can function properly. Whenever we swallow, the Eustachian tubes

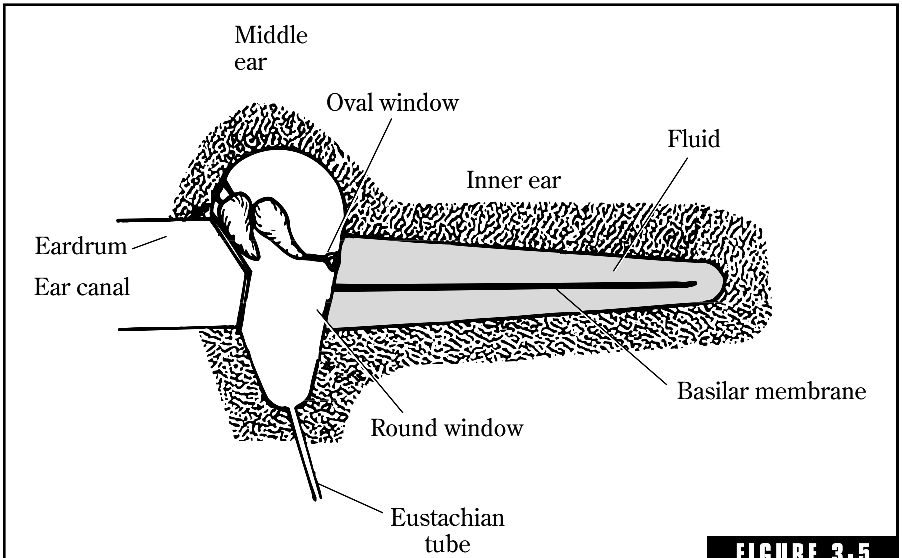

FIGURE 3-5

Highly idealized sketch of the human ear showing the unrolled fluid-filled cochlea. Sound entering the ear canal causes the eardrum to vibrate. This vibration is transmitted to the cochlea through the mechanical linkage of the middle ear. The sound is analyzed through standing waves set up on the basilar membrane.

open, equalizing the middle ear pressure. When an aircraft (at least those without pressurized cabins) undergoes rapid changes in altitude, the occupants might experience momentary deafness or pain until the middle ear pressure is equalized by swallowing. Actually, the Eustachian tube has a third emergency function of drainage if the middle ear becomes infected.

The Inner Ear

Only the acoustical amplifiers and the mechanical impedance matching features of the middle ear have been discussed so far. These are relatively well understood. The intricate operation of the cochlea is still clouded in mystery, but extensive research is steadily adding to our knowledge.

Figure 3-1 shows the close proximity of the three mutually-perpendicular, semicircular canals of the vestibular mechanism, the balancing organ, and the cochlea, the sound-analyzing organ. The same fluid permeates all, but their functions are independent.

The cochlea, about the size of a pea, is encased in solid bone. It is coiled up like a cockleshell from which it gets its name. For the purposes of illustration, this 2¾-turn coil has been stretched out its full length, about one inch, as shown in Fig. 3-5. The fluid-filled inner ear is divided lengthwise by two membranes, *Reissner's membrane* and the *basilar membrane*. Of immediate interest is the basilar membrane and its response to sound vibrations in the fluid.

Vibration of the eardrum activates the ossicles. The motion of the stapes, attached to the oval window, causes the fluid of the inner ear to vibrate. An inward movement of the oval window results in a flow of fluid around the distant end of the basilar membrane, causing an outward movement of the membrane of the round window. Sound actuating the oval window results in standing waves being set up on the basilar membrane. The position of the amplitude peak of the standing wave on the basilar membrane changes as the frequency of the exciting sound is changed.

Low-frequency sound results in maximum amplitude near the distant end of the basilar membrane; high-frequency sound produces peaks near the oval window. For a complex signal such as music or speech, many momentary peaks are produced, constantly shifting in amplitude and position along the basilar membrane. These resonant peaks on the basilar membrane were originally thought to be so broad as to be unable to explain the sharpness of frequency discrimination displayed by the human ear. Recent research is showing that at low sound intensities, the basilar membrane tuning curves are very sharp, broadening only for intense sound. It now appears that the sharpness of the basilar membrane's mechanical tuning curves is comparable to the sharpness of single auditory nerve fibers, which innervate it.

Stereocilia

Waves set up on the basilar membrane in the fluid-filled duct of the inner ear stimulate hairlike nerve terminals that convey signals to the brain in the form of neuron discharges, about 15,000 outer hair cells with about 140 tiny hairs called *stereocilia* jutting from each one. In addition, there are about 3,500 inner hair cells, each having about 40 *stereocilia* attached. These stereocilia are the true transducers of sound energy to electrical discharges. There are two types of hair cells, inner and outer, so-called by their placement and arrangement. As sound causes the cochlear fluid and the basilar membrane to move, the stereocilia on the hair cells are bent, initiating neural discharges to the auditory cortex.

When sound excites the fluid of the inner ear, membrane and hair cells are stimulated, sending an electrical wave through the surrounding tissue. These so-called *microphonic potentials* (analog) can be picked up and amplified, reproducing the sound falling on the ear, which acts as a biological microphone. These potentials are proportional to the sound pressure and linear in their response over an 80-dB range. While interesting, this microphonic potential must not be confused with the *action potentials* of the auditory nerve, which convey information to the brain.

Bending the stereocilia triggers the nerve impulses that are carried by the auditory nerve to the brain. While the microphonic signals are analog, the impulses sent to the acoustic cortex are impulses generated by neuron discharges. A single nerve fiber is either firing or not firing (binary!). When it fires, it causes an adjoining one to fire, and so on. Physiologists liken the process to a burning gunpowder fuse. The rate of travel bears no relationship to how the fuse was lighted. Presumably the loudness of the sound is related to the number of nerve fibers excited and the repetition rates of such excitation. When all the nerve fibers (some 15,000 of them) are excited, this is the maximum loudness that can be perceived. The threshold sensitivity would be represented by a single fiber firing. An overall, well-accepted theory of how the inner ear and the brain really function has not yet been formulated.[3–6]

This has been a highly simplified presentation of a very complex mechanism to which much current research is being devoted. Some of the numbers used and theories discussed are not universally accepted. Popularization of a subject such as the ear is an occupation that might be hazardous to my health, but audio workers must surely be amazed at the delicate and effective workings of the human ear. It is hoped that a new awareness of, and respect for, this delicate organism will be engendered and that damaging high levels of sound be avoided.

Loudness vs. Frequency

The seminal work on loudness was done at Bell Laboratories by Fletcher and Munson and reported in 1933,[7] and refinements have been added by others since that time. The family of equal-loudness contours of Fig. 3-6, the work of Robinson and Dadson,[8] has been adopted as an international standard (I.S.O. 226).

Each equal-loudness contour is identified by its value at 1,000 Hz, and the term *loudness level* in phons is thus defined. For example, the equal-loudness contour passing through 40-dB sound-pressure level at 1,000 Hz is called the 40-phon contour. Loudness is a subjective term; sound-pressure level is strictly a physical term. Loudness level is also a physical term that is useful in estimating the loudness of a sound (in units of sones) from sound-level measurements. The shapes of the equal-loudness contours contain subjective information because they were obtained by a subjective comparison of the loudness of a tone to its loudness at 1,000 Hz.

The surprising thing about the curves of Fig. 3-6 is that they reveal that perceived loudness varies greatly with frequency and sound-pressure level. For example, a sound-pressure level of 30 dB yields a loudness level of 30 phons at 1,000 Hz, but it requires a sound-pressure level of 58 dB more to sound equally loud at 20 Hz as shown in Fig. 3-7. The curves tend to flatten at the higher sound levels. The 90-phon curve rises only 32 dB between 1,000 Hz and 20 Hz. Note that inverting the curves of Fig. 3-7 gives the frequency response of the ear in

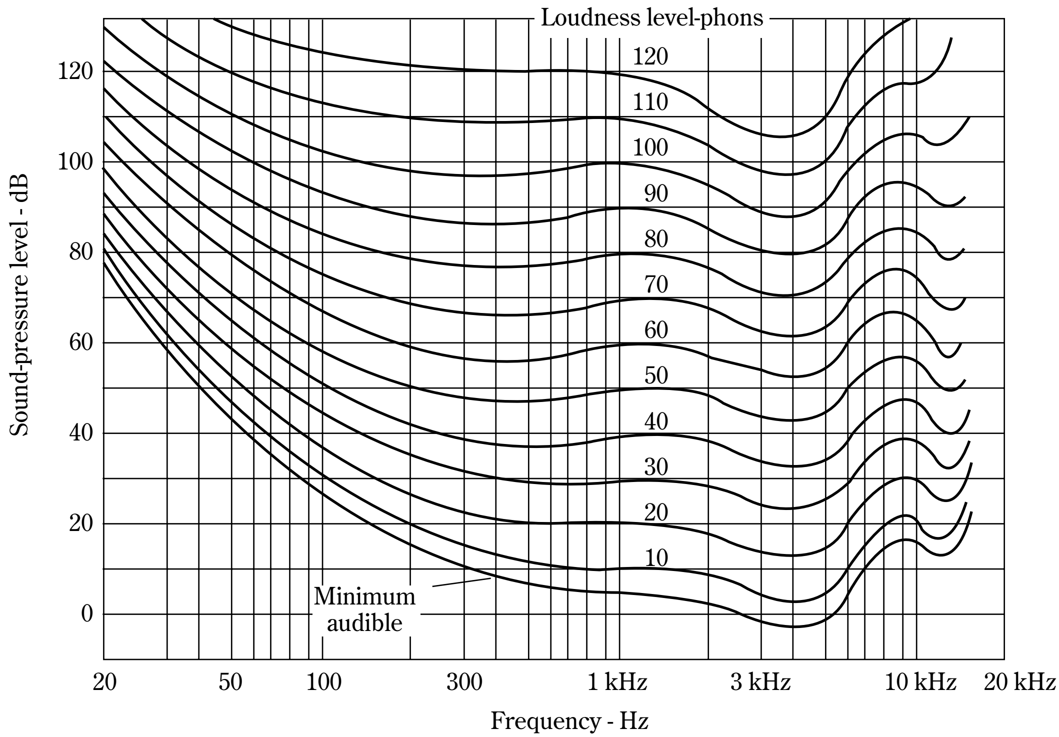

FIGURE 3-6

Equal-loudness contours of the human ear. These contours reveal the relative lack of sensitivity of the ear to bass tones, especially at lower sound levels. Inverting these curves give the frequency response of the ear in terms of loudness level. (After Robinson and Dadson.[8])

terms of loudness level. The ear is less sensitive to bass notes than midband notes at low levels. There are wiggles in the ear's high-frequency response that are relatively less noticeable. This bass problem of the ear means that the quality of reproduced music depends on the volume-control setting. Listening to background music at low levels requires a different frequency response than listening at higher levels.

Loudness Control

Let us assume that the high fidelity enthusiast adjusts the volume control on his or her amplifier so that the level of recorded symphony music is pleasing as a background to conversation (assumed to be

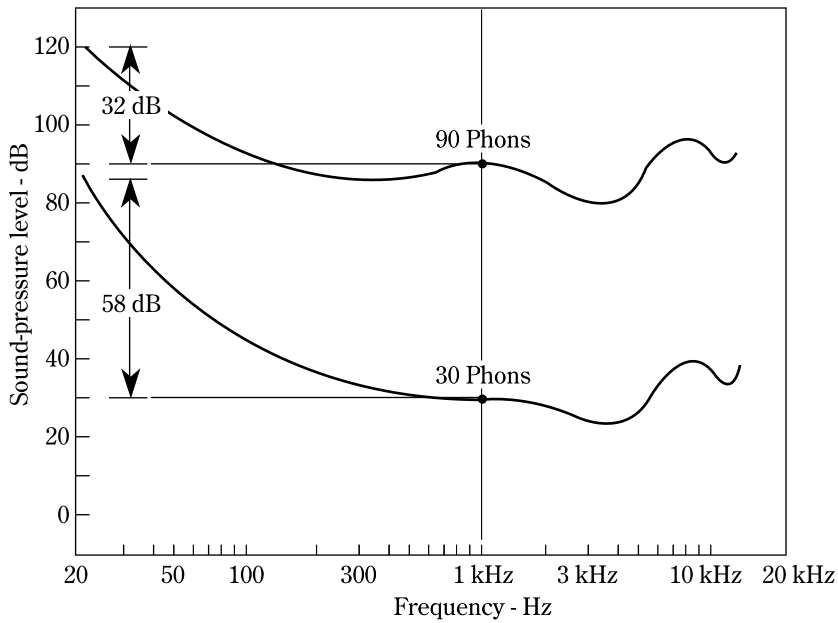

FIGURE 3-7

A comparison of the ear's response at 20 Hz compared to that at 1,000 Hz. At a loudness level of 30 phons, the sound-pressure level of a 20-Hz tone must be 58 dB higher than that at 1,000 Hz to have the same loudness. At 90 phons loudness level, an increase of only 32 dB is required. The ear's response is somewhat flatter at high loudness levels. Loudness level is only an intermediate step to true subjective loudness as explained in the text.

about 60 phons). As the passage was played at something like an 80-phon loudness level in the concert hall, something needs to be done to give the bass and treble of the music the proper balance at the lower-than-concert-hall level. Our enthusiast would find it necessary to increase both bass and treble for good balance.

The loudness control found on many amplifiers adjusts electrical networks to compensate for the change in frequency response of the ear for different loudness levels. But the curve corresponding to a given setting of the loudness control applies only to a specific loudness level of reproduced sound. The loudness control is far from a complete solution to the problem. Think of all the things that affect the volume-control setting in a particular situation. The loudspeakers vary in acoustic output for a given input power. The gain of preamplifiers, power amplifiers, tuners, and phono pickups differs from brand to brand and circuit to circuit. Listening-room conditions vary from dead to highly reverberant. With all of these variables, how can a manufacturer design a loudness control truly geared to the sound-pressure level at the ear of listener *x* with the particular variables of *x*'s equipment and *x*'s listening environment? For a loudness control to function properly, *x*'s system must be calibrated and the loudness control fitted to it.[9]

Area of Audibility

Curves A and B of Fig. 3-8 were obtained from groups of trained listeners. In this case, the listeners face the sound source and judge whether a tone of a given frequency is barely audible (curve A) or

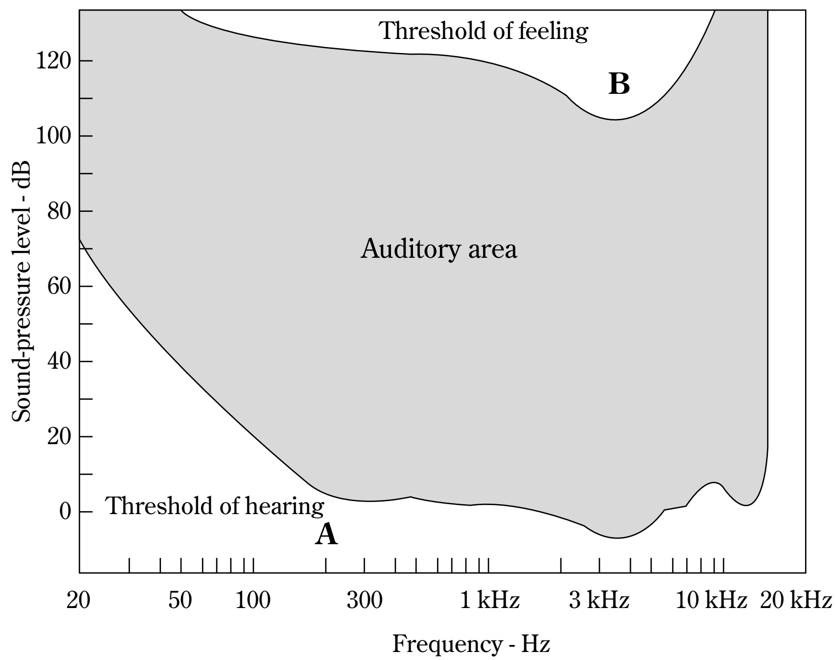

FIGURE 3-8

The auditory area of the human ear is bounded by two threshold curves, (A) the threshold of hearing delineating the lowest level sounds the ear can detect, and (B) the threshold of feeling at the upper extreme. All of our auditory experiences occur within this area.

beginning to be painful (curve B). These two curves represent the extremes of our perception of loudness.

Curve A of Fig. 3-8, the threshold of hearing, tells us that human ears are most sensitive around 3 kHz. Another way to state this is that around 3 kHz a lower-level sound elicits a greater threshold response than higher or lower frequencies. At this most sensitive region, a sound-pressure level of 0 dB can just barely be heard by a person of average hearing acuity. Is it fortuitous that this threshold is at a nice, round, 0-dB level? No, the reference level of pressure of 20 mPa (20 micropascals) was selected for this reason. It is both instructive and comforting to know that a sound-pressure level of 60 dB turns out to be approximately 60 dB above our threshold of hearing.

Curve B of Fig. 3-8 represents the level at each frequency at which a tickling sensation is felt in the ears. This occurs at a sound-pressure level of about 120 or 130 dB. Further increase in level results in an increase in feeling until a sensation of pain is produced. The threshold tickling is a warning that the sound is becoming dangerously loud and that ear damage is either imminent or has already taken place.

In between the threshold of hearing (curve A of Fig. 3-8) and the threshold of feeling (curve B) is the area of audibility. This is an area with two dimensions: the vertical dimension of sound-pressure level and the horizontal range of frequencies that the ear can perceive. All the sounds that humans experience must be of such a frequency and level as to fall within this auditory area. Chapter 5 details more specifically how much of this area is used for common music and speech sounds.

The area of audibility for humans is quite different from that of many animals. The bat specializes in sonar cries that are far above the upper frequency limit of our ears. The hearing of dogs extends higher than ours, hence the usefulness of ultrasonic dog whistles. Sound in the infrasonic and ultrasonic regions, as related to the hearing of humans, is no less true sound in the physical sense, but it does not result in human perception.

Loudness vs. Sound-Pressure Level

The *phon* is the unit of loudness level that is tied to sound-pressure level at 1,000 Hz as we have seen in Figs. 3-6, 3-7, and 3-8. This is useful up to a point, but it tells us little about human reaction to loudness of sound. We need some sort of subjective unit of loudness. Many experiments conducted with hundreds of subjects and many types of sound have yielded a consensus that for a 10-dB increase in sound-pressure level, the average person reports that loudness is doubled. For a 10-dB decrease in sound level, subjective loudness is cut in half. One researcher says this should be 6 dB, others say 10 dB, so work on the problem continues. However, a unit of subjective loudness has been adopted called the *sone*. One sone is defined as the loudness experienced by a person listening to a tone of 40-phon loudness level. A sound of 2 sones is twice as loud, and 0.5 sone half as loud.

Figure 3-9 shows a graph for translating sound-pressure levels to loudness in sones. One point on the graph is the very definition of the

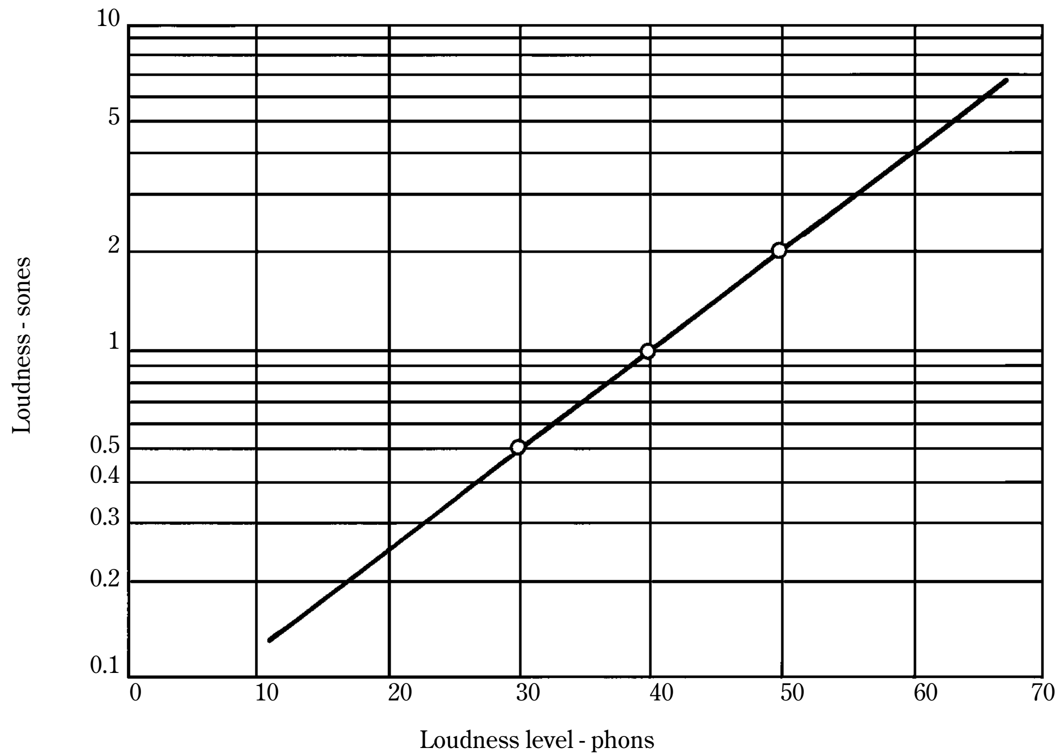

FIGURE 3-9

The graphical relationship between the physical loudness level in phons and subjective loudness in sones.

sone, the loudness experienced by a person hearing a 1,000-Hz tone at 40-dB sound-pressure level, or 40 phons. A loudness of 2 sones is then 10 dB higher; a loudness of 0.5 sones is 10 dB lower. A straight line can be drawn through these three points, which can then be extrapolated for sounds of higher and lower loudness.

As crude as this graph may be, it is a way of getting at the subjective factor of loudness. The value of this line of reasoning is that if a consultant is required by a court to give his or her opinion on the loudness of an industrial noise that bothers neighbors, he or she can make a one-third octave analysis of the noise, translate the sound-pressure levels of each band to sones by the help of a series of graphs such as Fig. 3-9, and by adding together the sones of each band, arrive at an estimate of the loudness of the noise. This idea of being able to add component sones is very nice; adding decibels of sound-pressure levels is a path that leads only to confusion.

Table 3-1 shows the relationship between loudness level in phons to the subjective loudness in sones. Although most audio workers will have little occasion to become involved in phons or sones, it is good to realize that a true subjective unit of loudness (sone) is related to loudness level (phon), which is in turn related by definition to what we can measure with a sound-level meter. There are highly developed empirical methods of calculating the loudness of sound as they would be perceived by humans from purely physical measurements of sound spectra, such as those measured with a sound-level meter and an octave or one-third octave filter.[10]

Table 3-1. Loudness level in phons vs. loudness in sones.

Loudness level (phons)	Subjective loudness (sones)	Typical examples
100	64	Heavy truck passing
80	16	Talking loudly
60	4	Talking softly
40	1	Quiet room
20	0.25	Very quiet studio

Loudness and Bandwidth

In the discussion of loudness we have talked tones up to this point, but single-frequency tones do not give all the information we need to relate subjective loudness to meter readings. The noise of a jet aircraft taking off sounds much louder than a tone of the same sound-pressure level. The bandwidth of the noise affects the loudness of the sound, at least within certain limits.

Figure 3-10A represents three sounds having the same sound-pressure level of 60 dB. Their bandwidths are 100, 160, and 200 Hz, but heights (representing sound intensity per Hz) vary so that areas are equal. In other words, the three sounds have equal intensities. (Sound intensity has a specific meaning in acoustics and is not to be equated to sound pressure. Sound intensity is proportional to the square of sound pressure for a plane progressive wave). The catch is that all three sounds of Fig.3-10A do not have the same loudness. The graph in Fig. 3-10B shows how a bandwidth of noise having a constant 60-dB sound-pressure level and centered on 1,000 Hz is related to loudness as experimentally determined. The 100-Hz noise

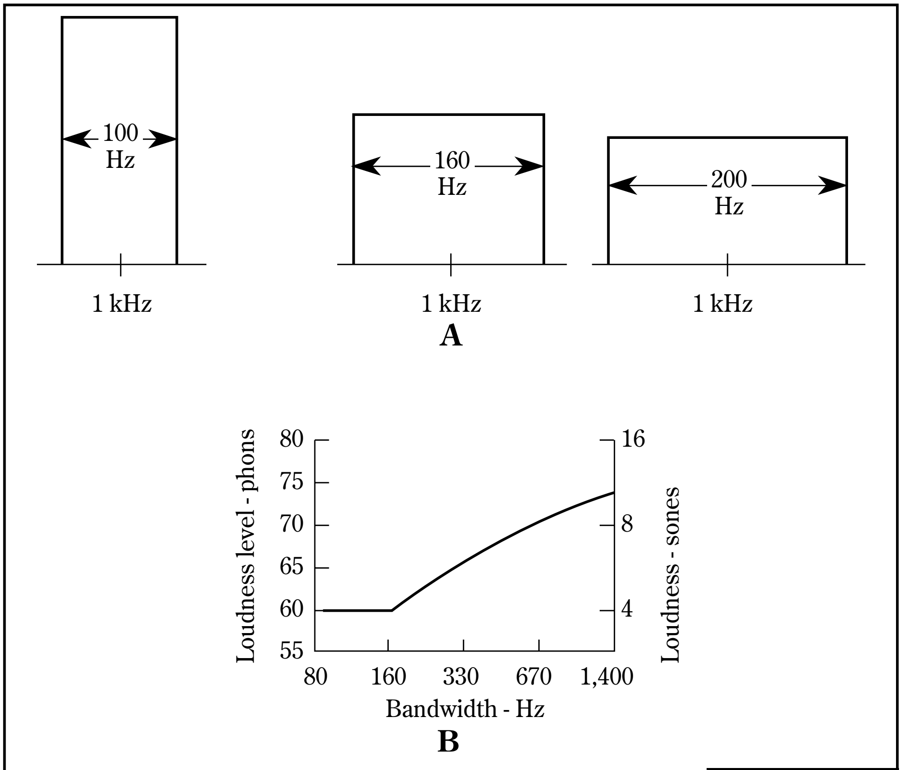

FIGURE 3-10

(A) Three noises of different bandwidths, but all having the same sound-pressure level of 60 dB. (B) The subjective loudness of the 100- and 160-Hz noise is the same, but the 200-Hz band sounds louder because it exceeds the 160-Hz critical band width of the ear at 1,000 Hz.

has a loudness level of 60 phons and a loudness of 4 sones. The 160-Hz bandwidth has the same loudness, but something mysterious happens as the bandwidth is increased beyond 160 Hz. The loudness of the noise of 200-Hz bandwidth is louder, and from 160 Hz or up, increasing bandwidth increases loudness. Why the sharp change at 160 Hz?

It turns out that 160 Hz is the width of the ear's critical band at 1,000 Hz. If a 1,000-Hz tone is presented to a listener along with random noise, only the noise in a band 160 Hz wide is effective in masking the tone. In other words, the ear acts like an analyzer composed of a set of bandpass filters stretching throughout the audible spectrum. This filter set is nothing like that found in the electronics laboratory. The common ⅓-octave filter set may have 28 adjacent filters overlapping at the –3 dB points. The set of critical band filters is continuous; that is, no matter where you might choose to set the signal generator dial, there is a critical band centered on that frequency.

Many years of research on this problem has yielded a modicum of agreement on how the width of the critical-band filters varies with frequency. This classical bandwidth function is shown in the graph of Fig. 3-11. There has been some question as to the accuracy of this graph below about 500 Hz that has led to other methods of measuring the bandwidth. Out of this has come the concept of the *equivalent rectangular bandwidth* (ERB) that applies to young listeners at moderate sound levels.[11] This approach is based on mathematical methods and offers the convenience of being able to calculate the ERB from the equation given in Fig. 3-11.

One-third-octave filter sets have been justified in certain measurements because the filter bandwidths approach those of the critical bands of the ear. For comparison, a plot of one-third-octave bandwidths is included in Fig. 3-11. One-third-octave bands are 23.2 percent of the center frequency. The classical critical-band function is about 17 percent of the center frequency. It is interesting to note that the ERB function (12 percent) is very close to that of one-sixth-octave bands (11.6 percent). This suggests the possibility of one-sixth-octave filter sets playing a larger role in sound measurements of the future.

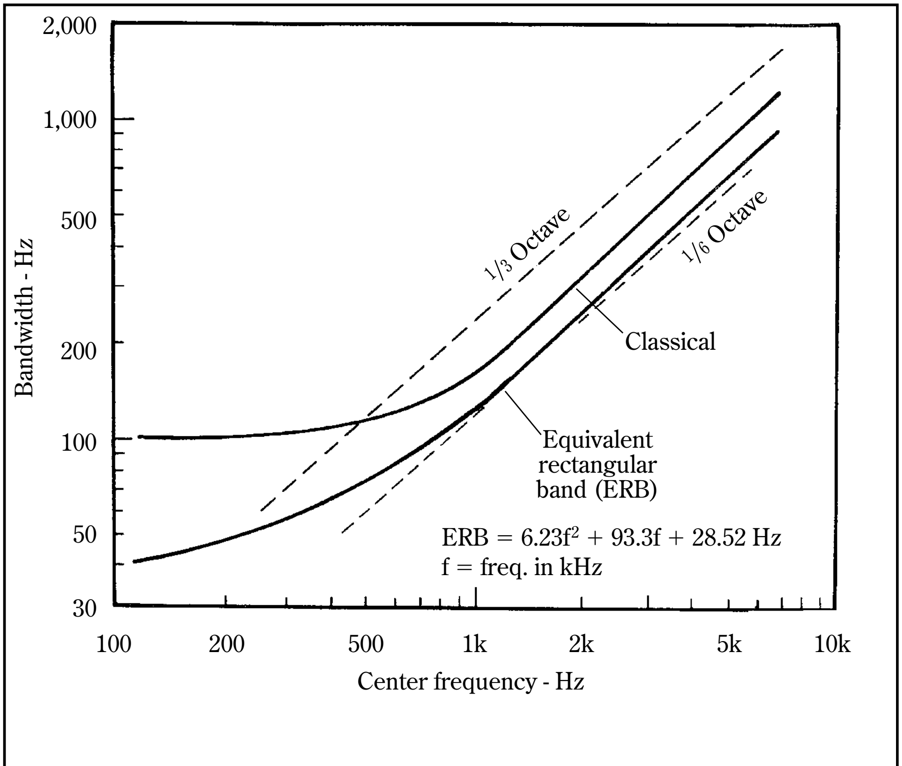

FIGURE 3-11

A comparison of bandwidths of ⅓- and ⅙-octave bands, critical bands of the ear, and equivalent rectangular critical bands (ERB) calculated from the above equation.[11]

Loudness of Impulses

The examples discussed up to this point have been concerned with steady-state tones and noise. How does the ear respond to transients of short duration? This is important because music and speech are essentially made up of transients. To focus attention on this aspect of speech and music, play some tapes backward. The initial transients now appear at the ends of syllables and musical notes and stand out prominently. These transients justify a few words on the ear's response to short-lived sounds.

A 1,000-Hz tone sounds like 1,000 Hz in a 1-second tone burst, but an extremely short burst sounds like a click. The duration of such a burst also influences the perceived loudness. Short bursts do not sound as loud as longer ones. Figure 3-12 shows how much the level of the shorter pulses has to be increased to have the same loudness as a long pulse or steady tone. A pulse 3 milliseconds long must have a level about 15 dB higher to sound as loud as a 0.5-second (500 millisecond) pulse. Tones and random noise follow roughly the same relationship in loudness vs. pulse length.

The 100-msec region is significant in Fig. 3-12. Only when the tones or noise bursts are shorter than this amount must the sound-pressure level be increased to produce a loudness equal to that of long pulses or steady tones or noise. This 100 msec appears to be the integrating time or the time constant of the human ear.

In reality, Fig. 3-12 tells us that our ears are less sensitive to short transients. This has a direct bearing on understanding speech. The consonants of speech determine the meaning of many words. For instance, the only difference between bat, bad, back, bass, ban, and bath are the consonants at the end. The words led, red, shed, bed, fed, and wed have the all-important consonants at the beginning. No matter where they

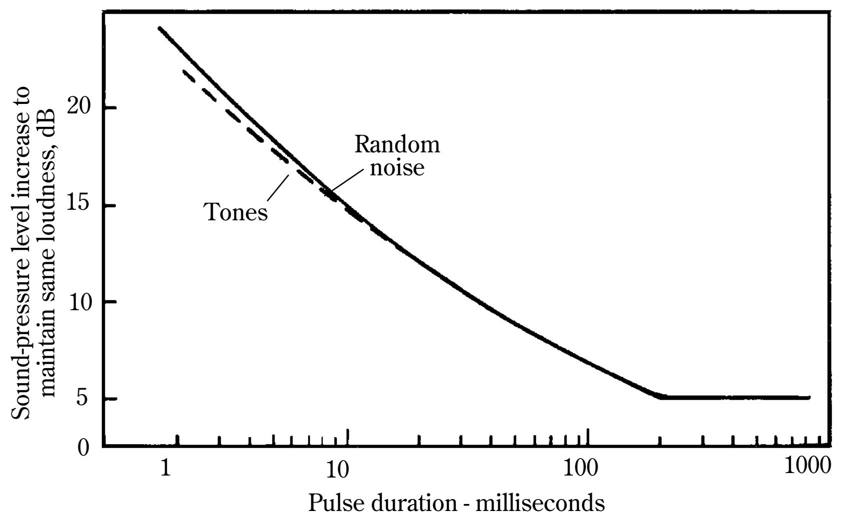

FIGURE 3-12

Short pulses of tones or noise are less audible than longer pulses as these graphs indicate. The discontinuity of the 100- to 200-msec region is related to the integrating time of the ear.

occur, these consonants are genuine transients having durations on the order of 5 to 15 msec. A glance at Fig. 3-12 tells you that transients this short must be louder to be comparable to longer sounds. In the above words, each consonant is not only much shorter than the rest of the word, it is also at a lower level. Thus you need good listening conditions to distinguish between such sets of words. Too much background noise or too much reverberation can seriously impair the understandability of speech because of the consonant problem.[12]

Audibility of Loudness Changes

Modern faders are of the composition type giving gradations in level so small as to be inaudible. Wire-wound faders of early mixing consoles produced discrete steps in level that could be audible. Steps of 5 dB were definitely audible, steps of 0.5 dB were inaudible, but these steps cost too much to produce and 0.5 dB steps were not necessary. Steps of 2 dB, an economic compromise, produced changes in signal level that were barely detectable by an expert ear. Detecting differences in intensity varies somewhat with frequency and also with sound level.

At 1 kHz, for very low levels, a 3-dB change is the least detectable by the ear, but at high levels the ear can detect a 0.25-dB change. A very low level 35-Hz tone requires a 9-dB level change to be detectable. For the important midfrequency range and for commonly used levels, the minimum detectable change in level that the ear can detect is about 2 or 3 dB. Making level changes in increments less than these is usually unnecessary.

Pitch vs. Frequency

Pitch, a subjective term, is chiefly a function of frequency, but it is not linearly related to it. Because pitch is somewhat different from frequency, it requires another subjective unit—the *mel*. Frequency is a physical term measured in cycles per second, now called Hertz. Although a weak 1,000-Hz signal is still 1,000 Hz if you increase its level, the pitch of a sound may depend on sound-pressure level. A reference pitch of 1,000 mels has been defined as the pitch of a 1,000-Hz tone with a sound-pressure level of 60 dB. The relationship between pitch and frequency, determined by experiments with juries of listeners, is shown in Fig. 3-13. Notice that on the experimental curve 1,000 mels coincides with 1,000 Hz, which tells us that the sound-pressure level for this curve is 60 dB. It is interesting to note that the shape of the curve of Fig. 3-13 is quite similar to a plot of position along the basilar membrane of the inner ear as a function of frequency. This suggests that pitch is related to action on this membrane, but much work remains to be done to be certain of this.

Researchers tell us that there are about 280 discernible steps in intensity and some 1,400 discernible steps in pitch that can be

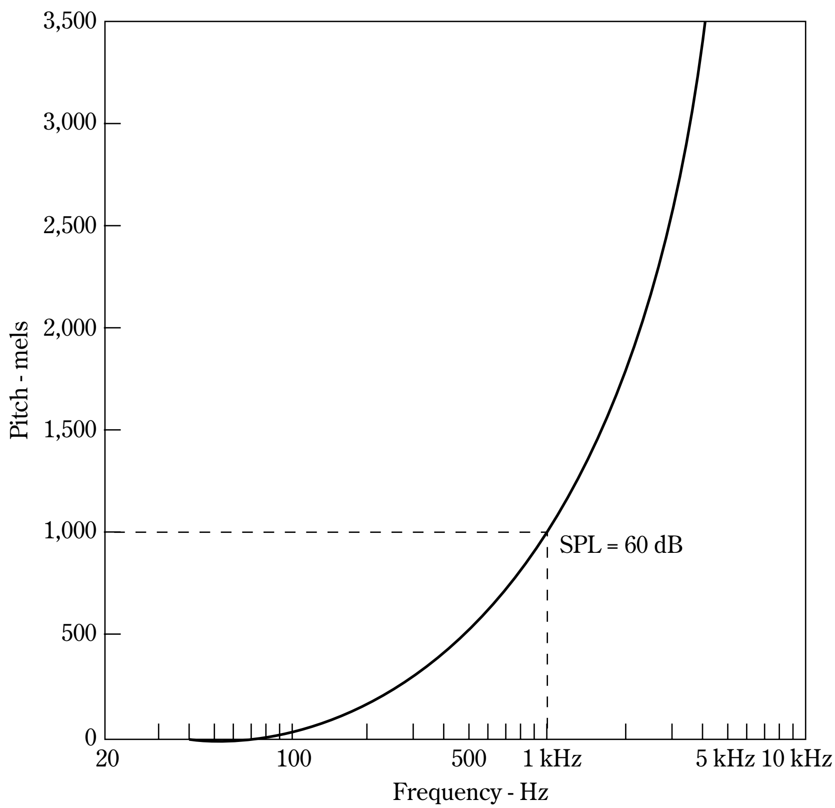

FIGURE 3-13

Pitch (in mels, a subjective unit) is related to frequency (in Hz, a physical unit) according to this curve obtained by juries of listeners. (After Stevens and Volkman.[13])

detected by the human ear. As changes in intensity and pitch are the very stuff of communication, it would be interesting to know how many combinations are possible. Offhand, it might seem that there would be 280 × 1,400 = 392,000 combinations detectable by the ear. This is overly optimistic because the tests were conducted by comparing two simple, single-frequency sounds in rapid succession and bears little resemblance to the complexities of commonly heard sounds. More realistic experiments show that the ear can detect only about 7 degrees of loudness and 7 degrees of pitch or only 49 pitch-loudness combinations. This is not too far from the number of *phonemes* (the smallest unit in a language that distinguishes one utterance from another) which can be detected in a language.

An Experiment

The level of sound affects the perception of pitch. For low frequencies, the pitch goes down as the level of sound is increased. At high frequencies, the reverse takes place—the pitch increases with sound level.

The following is an experiment within the reach of many readers that was suggested by Harvey Fletcher. Two audio oscillators are required, as well as a frequency counter. One oscillator is fed to the input of one channel of a high-fidelity system, the other oscillator to the other channel. After the oscillators have warmed up and stabilized, adjust the frequency of the left channel oscillator to 168 Hz and that of the right channel to 318 Hz. At low level these two tones are quite discordant. Increase the level until the pitches of the 168-Hz and 318-Hz tones decrease to the 150-Hz–300-Hz octave relationship, which gives a pleasant sound. This illustrates the decrease of pitch at the lower frequencies. An interesting follow-up would be to devise a similar test to show that the pitch of higher frequency tones increases with sound level.

Timbre vs. Spectrum

Timbre has to do with our perception of complex sounds. The word is applied chiefly to the sound of various musical instruments. A flute and oboe sound different even though they are both playing A. The tone of each instrument has its own timbre. Timbre is determined by the number and relative strengths of the instrument's partials. Tonal quality comes close to being a synonym for timbre.

Timbre is another subjective term. The analogous physical term is spectrum. A musical instrument produces a fundamental and a set of partials (or harmonics) that can be analyzed with a wave analyzer and plotted as in Fig. 1-15. Suppose the fundamental is 200 Hz, the second harmonic 400 Hz, the third harmonic 600 Hz, etc. The subjective pitch that the ear associates with our measured 200 Hz, for example, varies slightly with the level of the sound. The ear also has its own subjective interpretation of the harmonics. Thus, the ear's perception of the overall timbre of the instrument's note might be considerably different from the measured spectrum in a very complex way.

In listening to an orchestra in a music hall, the timbre you hear is different for different locations in the seating area.[14] The music is composed of a wide range of frequencies, and the amplitude and phase of the various components are affected by reflections from the various surfaces of the room. The only way to get one's analytical hands on studying such differences is to study the sound spectra at different locations. However, these are physical measurements, and the subjective timbre still tends to slip away from us. The important point of this section is to realize that a difference exists between timbre and spectrum.

Localization of Sound Sources

The perception of a direction to a source of a sound is, at least partially, the result of the amazing encoding function of the external ear, the pinna. Sound reflected from the various ridges, convolutions, and surfaces of the pinna combines with the unreflected (direct) sound at the entrance to the auditory canal. This combination, now encoded with directional information, passes down the auditory canal to the eardrum and thence to the middle and inner ear and on to the brain for interpretation.

This directional encoding process of the sound signal is indicated in Fig. 3-14. The sound wavefront can be considered as a multiplicity of sound rays coming from a specific source at a specific horizontal and vertical angle. As these rays strike the pinna they are reflected from the various surfaces, some of the reflections going toward the entrance to the auditory canal. At that point these reflected components combine with the unreflected (direct) component.

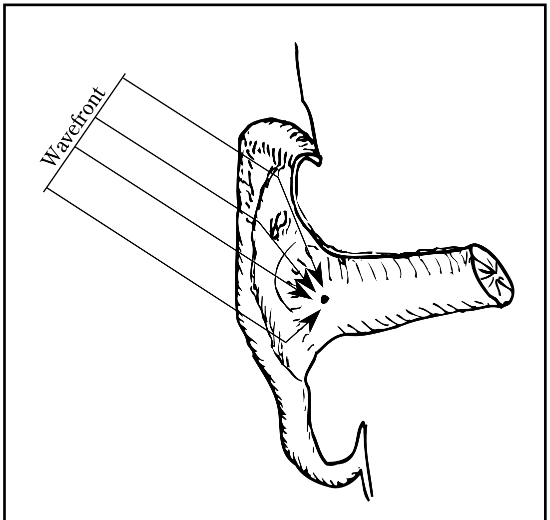

FIGURE 3-14

A wavefront of a sound can be considered as numerous rays perpendicular to that wavefront. Such rays, striking a pinna, are reflected from the various ridges and convolutions. Those reflections directed to the opening of the ear canal combine vectorially (according to relative amplitudes and phases). In this way the pinna encodes all sound falling on the ear with directional information, which the brain decodes as a directional perception.

For a sound coming directly from the front of the observer (azimuth and vertical angle = 0°), the "frequency response" of the combination at the opening of the ear canal will be that shown in Fig. 3-15. Instead of *frequency response*, a curve of this type is called a *transfer function* because it represents a vector combination involving phase angles.

For the sound at the entrance of the ear canal (Fig. 3-15) to reach the eardrum, the auditory canal must be traversed. As the transfer function at the entrance to the ear canal (Fig. 3-15) and that of the ear canal (Fig. 3-3) are combined, the shape of the resulting transfer function impinging on the eardrum is radically changed. Figure 3-3 showed a typical transfer function of the ear canal alone. It is a static, fixed function that does not change with direction of arrival of the sound. The ear canal acts like a quarter-wave pipe closed at one end by the eardrum exhibiting two prominent resonances.

The transfer function representing the specific direction to the source of Fig. 3-15 combining with the fixed transfer function of the ear canal (Fig. 3-3) gives the combined transfer function at the eardrum of Fig. 3-16. The brain translates this to a perception of sound coming from directly in front of the observer.[2]

The transfer function at the entrance to the ear canal (such as Fig. 3-15) is shaped differently for each horizontal and vertical direction. This is how the pinna encodes all arriving sound enabling the brain to yield different perceptions of direction. The sound arriving at the

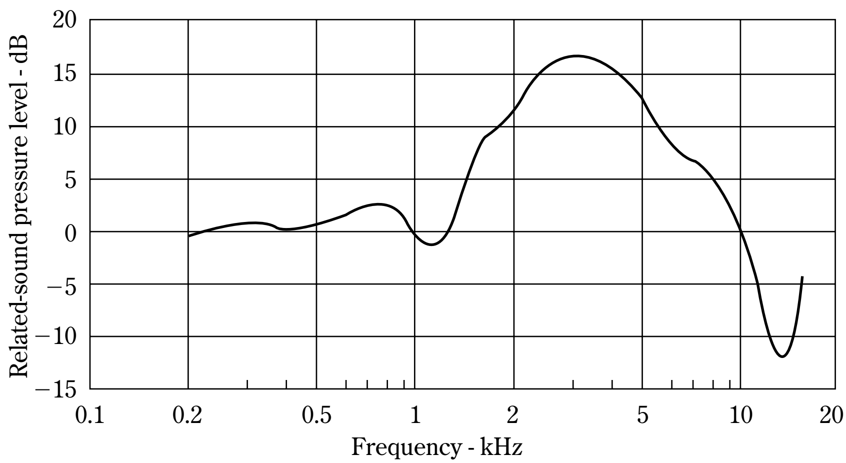

FIGURE 3-15

A measured example of the sound pressure (transfer function) at the opening of the ear canal corresponding to sound arriving from a point immediately in front of the subject. The shapes of such transfer functions vary with the horizontal and vertical angles at which the sound arrives at the pinna. (After Mehrgardt and Mellert.[2])

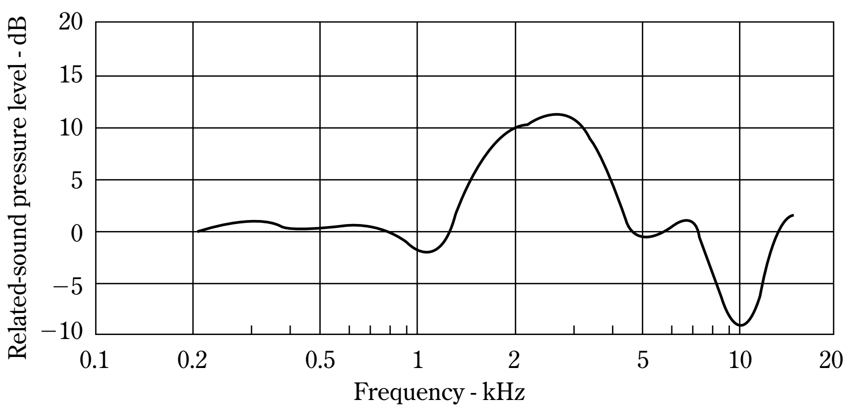

FIGURE 3-16

The transfer function of Fig. 3-15 at the opening of the ear canal is altered to this shape at the eardrum after being combined with the transfer function of the ear canal. In other words, a sound arriving at the opening of the ear canal from a source directly in front of the observer (Fig. 3-15) looks like Fig. 3-16 at the eardrum because it has been combined with the characteristics of the ear canal itself (Fig. 3-3). The brain has no trouble subtracting the fixed influence of the ear canal from every changing arriving sound.

eardrum is the raw material for all directional perceptions. The brain neglects (sees through?) the fixed component of the ear canal and translates the differently shaped transfer functions to directional perceptions.

Another more obvious directional function of the pinna is that of forward-backward discrimination, which does not depend on encoding and decoding. At the higher frequencies (shorter wavelengths), the pinna is an effective barrier. The brain uses this front-back differentiation to convey a general perception of direction.

A crucial question at this juncture is, "How about sounds arriving in the median plane?" The median plane is a vertical plane passing symmetrically through the center of the head and nose. Sources of sound in this plane present identical transfer functions to the two ears. The auditory mechanism uses another system for such localization, that of giving a certain place identity to different frequencies. For example, signal components near 500 and 8,000 Hz are perceived as coming from directly overhead, components near 1,000 and 10,000 Hz as coming from the rear.[15] This is an active area of research that is being continually refined.

The pinna, originally suspected of being only a useless vestigial organ, turns out to be a surprisingly sophisticated sound directional encoding mechanism.

Sound arriving from directly in front of an observer results in a peak in the transfer function at the eardrum in the 2- to 3-kHz region. This is the basis of the successful technique of old-time sound mixers adding "presence" to a recorded voice by adding an equalization boost in this frequency region. A voice can also be made to stand out from a musical background by adding such a peak to the voice response.

Binaural Localization

Stereophonic records and sound systems are a relatively new development. Stereo hearing has been around at least as long as man. Both are concerned with the localization of the source of sound. In early times some people thought that having two ears was like having two lungs or two kidneys, if something went wrong with one the other could still function. Lord Rayleigh laid that idea to rest by a simple experiment on the lawn of Cambridge University. A circle of assistants spoke or struck tuning forks and Lord Rayleigh in the center with his eyes

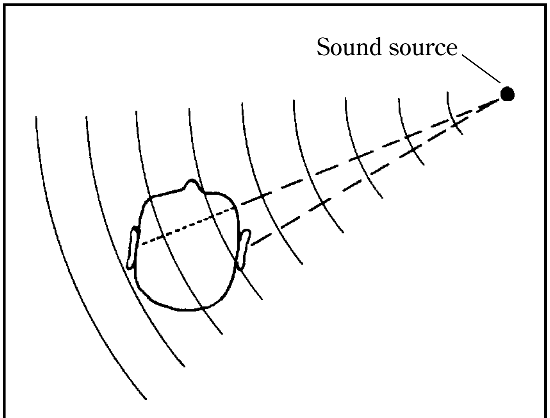

FIGURE 3-17

Our binaural directional sense is dependent in part on the difference in intensity and phase of the sound falling on two ears.

closed pointed to the source of sound with great accuracy, confirming the fact that two ears function together in binaural localization.

Two factors are involved, the difference in intensity and the difference in time of arrival (phase) of the sound falling on the two ears. In Fig. 3-17 the ear nearest the source receives a greater intensity than the far ear because the hard skull casts a "sound shadow." Because of the difference of distance to the source, the far ear receives sound somewhat later than the near ear. Below 1 kHz the phase (time) effect dominates while above 1 kHz the intensity effect dominates. There is one localization blind spot. A listener cannot tell whether sounds are coming from directly in front or from directly behind because the intensity of sound arriving at each ear is the same and in the same phase.

Another method of perception of direction comes into play in a relatively small room. The sound reaches the person over a direct path followed by many reflections from many different directions. The sound that arrives first creates in the hearer the main perception of direction. This has been called *the law of the first wavefront.*

Aural Harmonics: Experiment #1

This experiment, suggested by Craig Stark,[16] can be performed easily with your home high-fidelity system and two audio oscillators. Plug one oscillator into the left channel and the other into the right channel, and adjust both channels for an equal and comfortable level at some midband frequency. Set one oscillator to 24 kHz and the other to 23 kHz without changing the level settings. With either oscillator alone, nothing is heard because the signal is outside the range of the ear. (He notes here, however, that the dog might leave the room in disgust!) When both oscillators are feeding their respective channels, one at 24 kHz and the other at 23 kHz, a distinct 1,000-Hz tone is heard if the tweeters are good enough and you are standing in the right place.

The 1,000-Hz tone is the difference between 24,000 and 23,000 Hz. The sum, or 47,000 Hz, which even the dog may not hear even if it were radiated, is another sideband. Such sum and difference sidebands are generated whenever two pure tones are mixed in a nonlinear element. The nonlinear element in the above experiment is the middle and inner ear. In addition to the intermodulation products discussed earlier, the nonlinearity of the ear generates new harmonics that are not present in the sound falling on the eardrum.

Aural Harmonics: Experiment #2

The distortion introduced by the auditory system cannot be measured by ordinary instruments. It is a subjective effect requiring a different approach. Another demonstration of distortion in the ear can be accomplished by the following method with the same equipment used above, with the addition of a pair of headphones.

First, a 150-Hz tone is applied to the left earphone channel. If the hearing mechanism were perfectly linear, no aural harmonics would be heard as the exploratory tone is swept near the frequencies of the second, third, and other harmonics. If it is nonlinear, the presence of aural harmonics is indicated by the generation of beats. When 150 Hz is applied to the left ear, and the exploratory tone of the right ear is slowly varied about 300 Hz, the second harmonic is indicated by the presence of beats between the two. If you change the exploratory oscillator to a frequency around 450 Hz, the presence of a third harmonic will be revealed by beats.

Experts have even estimated the magnitude of the harmonics by the strength of such beats. The amount of distortion produced in the ear is modest at lower levels but becomes appreciable at high levels. Running the above experiment with tones of a higher level will make the presence of aural harmonics even more obvious.

The Missing Fundamental

If tones such as 1,000, 1,200, and 1,400 Hz are reproduced together, a pitch of 200 Hz is heard. This can be interpreted as the fundamental with 1,000 Hz as the 5th harmonic, 1,200 Hz as the 6th harmonic, etc.

At one time this 200 Hz was called "periodicity pitch" but so-called "pattern" theories dominate today. The auditory system is supposed to recognize that the upper tones are harmonics of the 200 Hz and supplies the missing fundamental that would have generated them. This is a very interesting effect but explanations of it are highly controversial.

The Ear as an Analyzer

Listening to a good symphony orchestra in your favorite concert hall, concentrate first on the violins. Now focus your attention on the clarinets, then the percussion section. Next listen to a male quartet and single out the first tenor, the baritone, the bass. This is a very remarkable power of the human ear/brain combination. In the ear canal, all these sounds are mixed together; how does the ear succeed in separating them? The sea surface might be disturbed by many wave systems, one due to local wind, one from a distant storm, and several wakes from passing vessels. The eye cannot separate these, but this is essentially what the ear is constantly doing with complex sound waves. By rigorous training, a keen observer can listen to the sound of a violin and pick out the various overtones apart from the fundamental!

The Ear as a Measuring Instrument

The emphasis on the distinction between physical measurements and subjective sensation would seem to rule out the possibility of using the ear for physical measurements. True, we cannot obtain digital readouts by looking in someone's eyes (or ears), but the ears are very keen at making comparisons. People are able to detect sound-level differences of about 1 dB throughout most of the audible band if the level is reasonable. Under ideal conditions, a change of a third this amount is perceptible. At ordinary levels, and for frequencies less than 1,000 Hz, the ear can tell the difference between tones separated by as little as 0.3%. This would be 0.3 Hz at 100 Hz and 3 Hz at 1,000 Hz.

The eminent Harvey Fletcher[17] has pointed out how the remarkable keenness of the human ear saved the day in many of his researches in synthesizing musical sounds. For example, in his study of piano sounds, he initially postulated that all that is necessary is to measure the frequency and magnitude of fundamental and harmonics and then combine them with the measured values of attack and decay. When this was done, the listening jury unanimously voted that the synthetic sounds did not sound like piano sounds but more like organ tones. Further study revealed the long-known fact that piano strings are very stiff and have properties of both solid rods and stretched strings. The effect of this is that piano partials are *nonharmonic!* By correcting the frequencies of what were assumed to be harmonics in integral multiples, the jury could not distinguish between the synthetic piano sounds and the real thing. The critical faculty of the ears of the jury in comparing sound qualities provided the key.

An Auditory Analyzer: An Experiment

Knowledge of the ear's filterlike critical bands leads to the tantalizing idea of analyzing continuous noises such as traffic noises, underwater background noises, etc., by using the ear instead of heavy and expensive sound-analyzing gear. This must have occurred to Harvey Fletcher, who first proposed the idea of critical bands, and to many investigators in this field who have dealt with critical bands through the years.

The general approach is illustrated in Fig. 3-18.[18] A tape recording of the noise to be analyzed is played back and mixed with a tone from a variable-frequency oscillator. The combination is amplified and listened to with a pair of headphones having a flat frequency response. The oscillator is set, say, at 1,000 Hz and its output

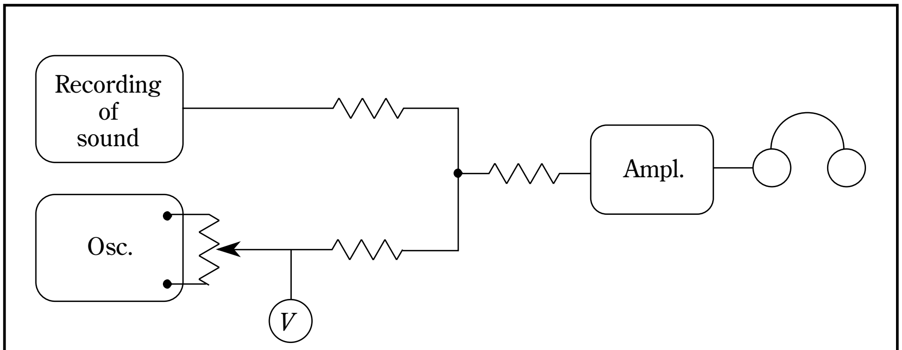

FIGURE 3-18

Equipment arrangement for using the critical bands of the human ear for sound analysis.

adjusted until the tone is just hidden or masked by the noise. Only the noise in the critical band centered on 1,000 Hz is effective in masking the tone. If the noise is expressed in sound-pressure level of a band 1 Hz wide, the voltage of the tone then corresponds to the 1-Hz sound-pressure level of the noise at the masked point. Adjusting the voltage until the tone is just masked should yield one of the points on our noise spectrum graph. For convenience, let us assume that this voltmeter is calibrated in dB referred to some arbitrary base such as 1 volt (dBv). Referring to Fig. 3-11, note that the critical band centered on 1,000 Hz is 160 Hz wide. This can also be expressed in decibels by taking $\log_{10}160 = 22$ dB; this 22 dB, representing the width of the critical band as it does, must be subtracted from the voltmeter reading in dB. This gives one point on the noise spectrum graph. By repeating the process for other frequencies, a series of points is obtained that reveal the shape of the noise spectrum. If the recording and the entire measuring system (including the observer's ears) were calibrated, the absolute levels for the noise spectrum could be obtained.

The important point here is that there is such a set of filters in our head that could be put to such a task, not that this method will ever replace a good sound level meter equipped with octave or one-third octave filters. Surely human variables would far exceed sound-level meter fluctuations from day to day, and what the observer eats for breakfast has no effect on the sound-level meter, although it might affect the dependability of the readings made with physiological equipment.

Meters vs. the Ear

There still remains a great chasm between subjective judgments of sound quality, room acoustics, etc., and objective measurements. Considerable attention is being focused on the problem. Consider the following descriptive words, which are often applied to concert-hall acoustics[19,20]:

warmth	clarity
bassiness	brilliance
definition	resonance
reverberance	balance

fullness of tone	blend
liveness	intimacy
sonority	shimmering

What kind of an instrument measures warmth or brilliance? How would you devise a test for definition? Progress, however, is being made. Take definition for instance. German researchers have adopted the term *deutlichkeit*, which literally means clearness or distinctness, quite close to definition. It can be measured by taking the energy in an echogram during the first 50 to 80 milliseconds and comparing it to the energy of the entire echogram. This compares the direct sound and early reflections, which are integrated by the ear, to the entire reverberant sound. This relatively straightforward measurement of an impulsive sound from a pistol or pricked balloon holds considerable promise for relating the descriptive term definition to an objective measurement. It will be a long time before all of these and a host of other subjective terms can be reduced to objective measurements, but this is a basic problem in acoustics and psychoacoustics.

There comes a time at which meter readings must give way to observations by human subjects. Experiments then take on a new, subjective quality. For example, in a loudness investigation, panels of listeners are presented with various sounds, and each observer is asked to compare the loudness of sound A with the loudness of B or to make judgments in other ways. The data submitted by the jury of listeners are then subjected to statistical analysis, and the dependence of a human sensory factor, such as loudness, upon physical measurements of sound level is assessed. If the test is conducted properly and sufficient observers are involved, the results are trustworthy. It is in this way that we discover that there is no linear relationship between sound level and loudness, pitch and frequency, or between timbre and sound quality.

The Precedence Effect

Our hearing mechanism integrates sound intensities over short intervals and acts somewhat like a ballastic measuring instrument. In simpler terms, in an auditorium situation, the ear and brain have the remarkable ability to gather all reflections arriving within about 50 msec after the direct sound and combine (integrate) them to give the impression that all this sound is from the direction of the original source, even though reflections from other directions are involved. The sound energy integrated over this period also gives an impression of added loudness.

It should not be too surprising that the human ear fuses all sounds arriving during a certain time window. After all, our eyes fuse a series of still pictures at the cinema, giving us the impression of continuous movement. The rate of presentation of the still pictures is important; there must be at least 16 pictures per second (62-millisecond interval) to avoid seeing a series of still pictures or a flicker. Auditory fusion works best during the first 20 or 30 milliseconds; beyond 50 to 80 milliseconds discrete echoes dominate.

Haas[21] set his subjects 3 meters from two loudspeakers arranged so that they subtended an angle of 45 degrees, the observer's line of symmetry splitting this angle. The conditions were approximately anechoic. The observers were called upon to adjust an attenuator until the sound from the "direct" loudspeaker equaled that of the "delayed" loudspeaker. He then proceeded to study the effects of varying the delay.

A number of researchers had previously found that very short delays (less then 1 msec) were involved in our discerning the direction of a source by slightly different times of arrival at our two ears. Delays greater than this do not affect our directional sense.

As shown in Fig. 3-19, Haas found that in the 5 to 35 msec delay range the sound from the delayed loudspeaker has to be increased more than 10 dB over the direct before it sounded like an echo. This is the precedence effect, or Haas effect. In a room, reflected energy arriving at the ear within 35 msec is integrated with the direct sound and is perceived as part of the direct sound as opposed to reverberant sound. These early reflections increase the loudness of the sound, and as Haas said, result in "...a pleasant modification of the sound impression in the sense of broadening of the primary sound source while the echo source is not perceived acoustically."

The transition zone between the integrating effect for delays less than 35 msec and the perception of delayed sound as discrete echo is gradual, and therefore, somewhat indefinite. Some peg the dividing line at a convenient 1⁄16 second (62 msec), some at 80 msec, and some at 100 msec beyond which there is no question about the discreteness of the echo. In this book we will consider the first 30 msec as in Fig. 3-19, the region of definite integration.

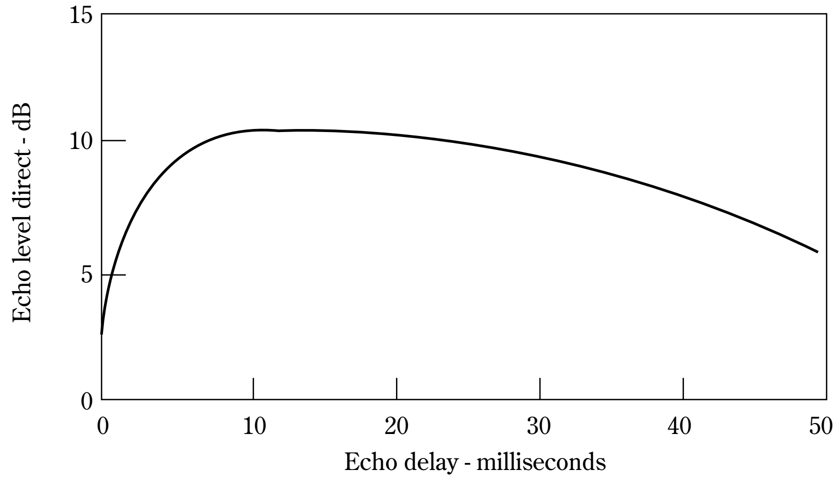

FIGURE 3-19

The precedence effect (Haas effect) in the human auditory system. In the 5 to 35 msec region, the echo level has to be about 10 dB higher than the direct sound to be discernible as an echo. In this region, reflected components arriving from many directions are gathered by the ear. The resulting sound seems louder, because of the reflections and appears to come from the direct source. For delays 50 to 100 msec and longer reflections are perceived as discrete echoes. (After Haas.[21])

Perception of Reflected Sound

In the preceding section, "reflected" sound was considered in a rather limited way. A more general approach is taken in this section. It is interesting that the loudspeaker arrangement Haas used was also used by dozens of other researchers and that this is basically the familiar stereo setup; two separated loudspeakers with the observer (listener) located symmetrically between the two loudspeakers. The sound from one loudspeaker is designated as the *direct* sound, that from the other, the delayed sound (the *reflection*). The delay injected between the two signals and their relative levels is adjustable. Speech is used as the signal.[22]

With the sound of the direct loudspeaker set at a comfortable level, and with a delay of, say 10 ms, the level of the reflected, or delayed, loudspeaker sound is slowly increased from a very low value. The sound level of the reflection at which the observer first detects a difference in the sound is the threshold of reflection detection. For levels less than this, the reflection is inaudible; for levels greater than this, the reflection is clearly audible.

As the reflection level is gradually increased above the threshold value, a sense of spaciousness is imparted to the combined sound. This sense of spaciousness prevails, even though the experiment is conducted in an anechoic space. As the level of the reflection is increased about 10 dB above the threshold value, another change is noticed in the sound; a broadening of the sound image and possibly a shifting of the image toward the direct loudspeaker is now added to the increasing spaciousness. As the reflection level is increased another 10 dB or so above the image broadening threshold, another change is noted; *discrete echoes* are heard.

This is all very interesting, but what practical value does it have? Consider a specific example: a listening room in which recorded music will be played. Figure 3-20 contains answers to the effect of sound reflected from floor, ceiling, and walls being added to the direct sound from the loudspeakers. Reflections below the threshold of perception are unusable; reflections perceived as discrete echoes are also unusable. The usable area is the unshaded area between those two threshold curves, A and C. Simple calculations can give estimates of the level and delay of any specific reflection, knowing the speed of sound, the distance traveled and applying the inverse square law. Figure 3-20 gives the subjective reactions the listener will probably have to the combination of any reflection and the direct sound.

To assist in the "simple" calculations mentioned previously, the following equations can be applied:

$$\text{Reflection delay} = \frac{(\text{reflected path, ft}) - (\text{direct path, ft})}{1{,}130 \text{ ft/sec}}$$

This assumes 100% reflection at the reflecting surface.

$$\text{Reflection level at listening position} = 20 \log \frac{\text{direct distance, ft}}{\text{reflection distance, ft}}$$

This assumes the inverse square propagation.

Occupational and Recreational Deafness

The hearing of workers in industry is now protected by law. The higher the environmental noise, the less exposure allowed (Table 3-2).

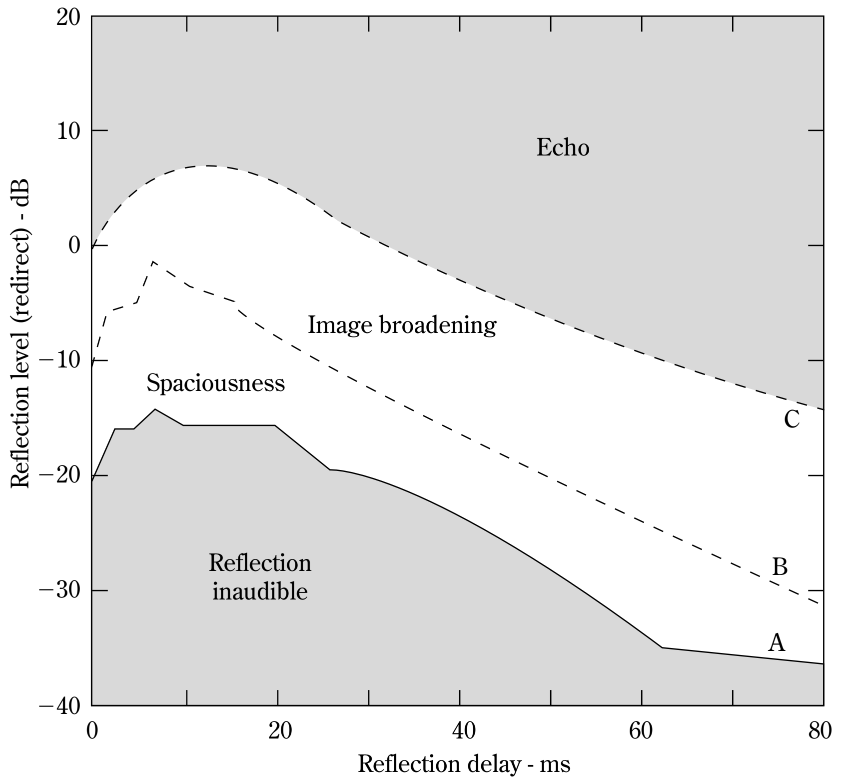

FIGURE 3-20

The effects of lateral reflections on the perception of the direct sound in a simulated stereo arrangement. These measurements were made in anechoic conditions, lateral angles 45–90 degrees, with speech as the signal. (A) Absolute threshold of audibility of the reflection. (B) Image shift/broadening threshold (A & B After Olive and Toole,[22] and Toole.[23]) (C) Lateral reflection perceived as a discrete echo (After Meyer and Schodder,[24] and Lochner and Burger[25]).

Researchers are trying to determine what noise exposure workers are subjected to in different plants. This is not easy as noise levels fluctuate and workers move about, but wearable *dosimeters* are often used to integrate the exposure over the work day. Industries are hard pressed to keep up with changes in regulations, let alone the installation of noise shields around offending equipment and keeping ear plugs in or ear muffs on the workers. Nerve deafness resulting from occupational noise is recognized as a distinct health hazard.

Table 3-2. OSHA permissible noise exposure times.*

Sound pressure level, dB, A-weighting, slow response	Maximum daily exposure hours
85	16
90	8
92	6
95	4
97	3
100	2
102	1.5
105	1
110	0.5
115	0.25 or less

*Reference: OSHA 2206 (1978)

It is especially bad when one works all day in a high-noise environment, then engages in motorcycle or automobile racing, listens to a 400-watt stereo at high level, or spends hours in a discotheque. The professional audio engineer operating with high monitoring levels is risking irreparable injury to the basic tools of the trade—his ears. As high-frequency loss creeps in, the volume control is turned up to compensate, and the rate of deterioration is accelerated.

The key to conservation of hearing is the audiogram. Comparing today's audiogram with earlier ones establishes the trend; if downward, steps can be taken to check it. The audiogram of Fig. 3-21, which looks something like the Big Dipper

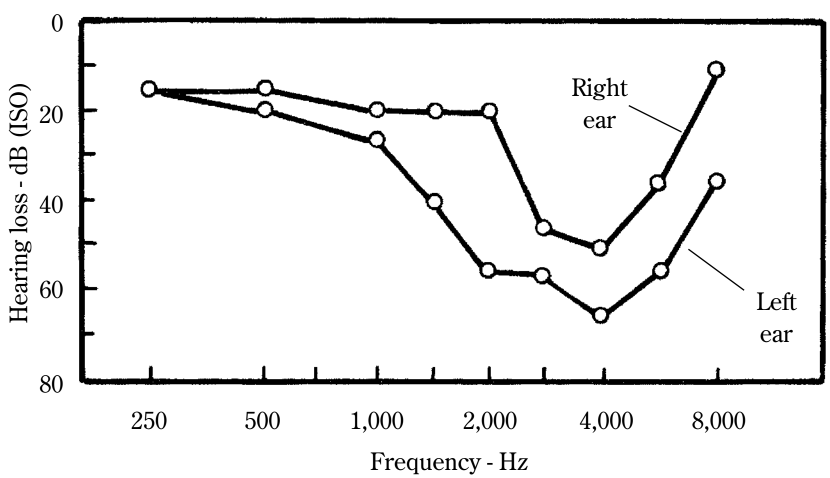

FIGURE 3-21

Audiograms showing serious loss centered on 4 kHz, presumably resulting from years of exposure to high-level sound in the control room of a recording studio.

constellation, is that of a 50-ish sound mixer in a recording studio. The indications are that this loss, centered on 4 kHz, is the accumulation of many years of listening to high-level sounds in the control room.

Summary

- The ear is sensitive enough to hear the tattoo of air particles on the eardrums in the quietude of an anechoic chamber.
- The auditory canal, acting as a quarter-wave pipe closed at one end by the eardrum, contributes an acoustical amplification of about 10 dB, and the head diffraction effect produces another 10 dB near 3 kHz. These are vital speech frequencies.
- The leverage of the ossicle bones of the middle ear and the ratio of areas of the eardrum and oval window successfully match the impedance of air to the fluid of the inner ear.
- The Eustachian tube and round window provide pressure release and equalization with atmospheric pressure.
- Waves set up in the inner ear by vibration of the oval window excite the sensory hair cells, which are connected to the brain. There is a "place effect," the peak of hair cell agitation for higher frequencies being nearer the oval window, and low frequencies at the distal end.
- The area of audibility is bounded by two threshold curves, the threshold of audibility at the lower extreme and the threshold of feeling or pain at the loud extreme. Our entire auditory experience occurs within these two extremes.
- The loudness of tone bursts decreases as the length of the burst is decreased. Bursts greater than 200 msec have full loudness, indicating a time constant of the ear at about 100 msec.
- Our ears are capable of accurately locating the direction of a source in the horizontal plane. In a vertical median plane, however, localization ability is less accurate.

- Pitch is a subjective term. Frequency is the associated physical term, and the two have only a general relationship.
- Subjective timbre or quality of sound and the physical spectrum of the sound are related, but not equal.
- The nonlinearity of the ear generates intermodulation products and spurious harmonics.
- The Haas, or precedence, effect describes the ability of the ear to integrate all sound arriving within the first 50 msec, making it sound louder.
- Although the ear is not effective as a measuring instrument yielding absolute values, it is very keen in comparing frequencies, levels, or sound quality.
- Occupational and recreational noises are taking their toll in permanent hearing loss. Definite precautionary steps to minimize this type of environmentally caused deafness are recommended.

Endnotes

[1]Bloom, P.J., *Creating source illusions by special manipulation*, J. Audio Eng. So., 25 (1977) 560-565.

[2]Mehrgardt, S and V. Mellert, *Transformation characteristics of the external human ear*, J. Acous. Soc. Am., 61, 6 (1977) 1567-1576.

[3]Moore, Brian C.J., *An introduction to the psychology of hearing*, New York, Academic Press (1982).

[4]Pickles, James D., *An introduction to the physiology of hearing*, 2nd. ed., San Diego, Academic Press (1988).

[5]Tobias, Jerry V. Ed., *Foundations of modern auditory theory*, Vol. 1, New York, Academic Press (1970).

[6]Tobias, Jerry V., Ed., *Foundations of modern auditory theory*, Vol. 2, New York, Academic Press (1972).

[7]Fletcher, H. and W.A. Munson, *Loudness, its definition, measurement, and calculations*, J. Acous. Soc. Am., 5 (1933) 82-108.

[8]Robinson, D.W. and R.S. Dadson, *A re-determination of the equal-loudness relations for pure tones*, British J. of Appl. Psychology., 7 (1956) 166-181. (Adopted by the International Standards Organization as ISO-226.)

[9]Toole, Floyd E., *Loudness—Applications and implications for audio*, dB Magazine, Part I: 7, 5 (May 1973) 7-30, Part II: 7, 6 (June 1973) 25-28.

[10]Zwicker, C.G., G. Flottorp, and S.S. Stevens, *Critical bandwidths in loudness summation*, J. Acous. Soc. Am., 29, 5 (1957) 548-557.

[11]Moore, Brian C.J. and Brian Glasberg, *Suggested formulae for calculating auditory filter bandwidths and excitation patterns*, J. Acous. Soc. Am., 74, 3 (1983) 750-753.

[12]Meyer, Erwin, *Physical measurements in rooms and their meaning in terms of hearing conditions*, Proc. 2nd Int. Congr. on Acous. (1956) 59-68.

[13]Stevens, S.S. and J. Volkman, *The relation of pitch to frequency: a revised scale*, Am. J. Psychology, 53 (1940) 329-353.

[14]Plomb, R. and H.J.M. Steeneken, *Place dependence of timbre in reverberant sound fields*, Acustica, 28, 1 (1973) 50-59.

[15]Blauert, Jens, *Spatial Hearing*, Cambridge, MIT Press (1983).

[16]Stark, Craig, *The sense of hearing*, Stereo Review, (September 1969) 66, 71-74.

[17]Fletcher, Harvey, *The ear as a measuring instrument*, J. Audio Eng. Soc., 17, 5 (1969) 532-534.

[18]Everest, F. Alton, *The filters in our ears*, Audio, 70, 9 (1986) 50-59.

[19]Schroeder, M.R., D. Gottlob, and K.F. Siebrasse, *Comparative study of European concert halls: correlation of subjective presence with geometric and acoustic parameters*, J. Acous. Soc. Am., 56, 4 (1974) 1195-1201.

[20]Hawkes, R.J. and H. Douglas, *Subjective experience in concert auditoria*, Acustica, 28, 5 (1971) 235-250.

[21]Haas, Helmut, *The influence of a single echo on the audibility of speech*, J. Audio Eng. Soc., 20, 2 (1972) 146-159. (This is an English translation from the German by Dr. Ingr. K.P.R. Ehrenberg of Haas' original paper in Acustica, 1, 2 (1951).

[22]Olive, S.E. and F.E. Toole, *The detection of reflections in typical rooms*, J. Audio Engr. Soc. 37 (1989) 539-553.

[23]Toole, F.E., *Loudspeakers and rooms for stereophonic sound reproduction*, Proc. AES 8th International Conference, Washington, D.C., 3-6 May, 1990, pp 71-91.

[24]Meyer, E. and G.R. Schodder, *On the influence of reflected sounds on directional localization and loudness of speech*, Nachr. Akad. Wiss., Göttingen, Math., Phys., Klasse IIa, 6 (1952) 31-42.

[25]Lochner, J.P.A. and J.F. Burger, *The subjective masking of short time delayed echoes by their primary sounds and their contribution to the intelligibility of speech*, Acustica, 8 (1958) 1-10.