CHAPTER

5 Hearing

The human auditory system is complex in structure and remarkable in function. Not only does it respond to a wide range of stimuli, but it precisely identifies the pitch and the timbre (quality) of a sound and even the direction of the source. Much of the hearing function is performed by the organ we call the ear, but recent research has emphasized how much hearing depends on the data processing that occurs in the central nervous system as well.

In this chapter you should learn:

- About the human auditory system (ear) and how it functions;
- How large and small numbers are expressed as powers of ten and on a logarithmic scale;
- About critical bands in hearing;
- About binaural hearing and localization;
- How *subjective* attributes of sound and music relate to *physical* parameters.

5.1 ■ RANGE OF HEARING

The range of sound *intensity* (pressure) and the range of *frequency* to which the ear responds, as shown in Fig. 5.1, is remarkable indeed. The intensity ratio between the sounds that bring pain to our ears and the weakest sounds we can hear is more than 10^{12} (1,000,000,000,000). The frequency ratio between the highest and lowest frequencies we can hear is nearly 10^3 (1000) times, or more than nine octaves (each octave represents a doubling of frequency).

Human vision is remarkable, too, but the frequency range does not begin to compare to that of human hearing. The frequency range of vision is about one octave (4×10^{14} to 7×10^{14} Hz, corresponding to wavelengths of 400 to 750 nanometers). Within this one octave range of frequency we can identify more than 7 million different colors (Rossing and Chiaverina 1999). Given that the frequency range of the ear is nine times greater than that of the eye, you can begin to imagine how many sound "colors" might be possible.

In Chapter 1, we learned that *power* is the rate at which work is done; it is equal to work or energy divided by time. We also learned that *pressure* is force per unit area (or force divided by area). In dealing with sound waves (or light waves), it is useful to talk about *intensity*, which is the power per unit area carried by the wave. It is expressed in

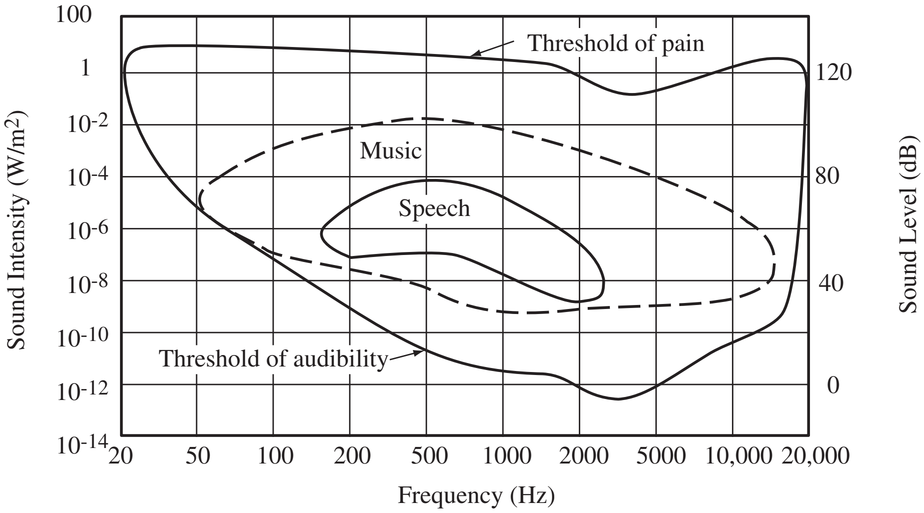


FIGURE 5.1 Range of frequencies and intensities to which the auditory system (ear) responds.

watts/square meter (W/m^2). The intensity of a sound wave multiplied by the area of our eardrum, for example, expresses the amount of power that the sound wave transmits to each ear. It is a small fraction of a watt even for the loudest of sounds.

Because it is rather difficult to measure sound intensity, we generally measure sound pressure instead, as we will discuss in Chapter 6. A microphone gives us an electrical signal proportional to sound pressure. The intensity is proportional to sound pressure squared. We will discuss this in more detail in Chapter 6.

The ear is extremely sensitive to small pressure changes. The pressure change in a very loud sound is still only 10^{-5} normal atmospheric pressure. At some sound frequencies, the vibrations of the eardrum may be as small as 10^{-8} mm, about one-tenth the diameter of the hydrogen atom. It is estimated that the vibrations of the very fine membrane in the inner ear that transmit this stimulus to the auditory nerve are nearly 100 times smaller yet in amplitude (Békésy 1960).

The frequency range of hearing varies greatly among individuals; a person who can hear over the entire audible range of 20–20,000 Hz is unusual. The ear is relatively insensitive to sounds of low frequency; for example, its sensitivity at 100 Hz is roughly 1000 times less than its sensitivity at 1000 Hz. Sensitivity to sounds of high frequency is greatest in early childhood and decreases gradually throughout life, so that an adult may have difficulty hearing sounds beyond 10,000 or 12,000 Hz. (This deterioration of perception of high frequencies, termed *presbycusis*, is compared in Chapter 31 to noise-induced hearing loss.)

Another remarkable quality of the auditory system is its selectivity. From the blended sounds of a symphony orchestra, a listener can pick out the sound of a solo instrument. In a noisy room crowded with people, it is possible to pick out a single speaker. Even during sleep the conditioned ear of a mother can respond to the cry of an infant. We can train ourselves to sleep through the noise of city traffic but to awaken at the sound of an alarm clock or unusual noise.

5.2 ■ STRUCTURE OF THE EAR

For convenience of description it is usual to divide the ear into three sections: the outer ear, the middle ear, and the inner ear (see Fig. 5.2). The *outer ear* consists of the external *pinna* and the *auditory canal* (meatus), which is terminated by the *eardrum* (tympanum). The pinna helps, to some extent, in collecting sound and contributes to our ability to determine the direction of origin of sounds of high frequency. The auditory canal acts as a pipe resonator that boosts hearing sensitivity in the range of 2000 to 5000 Hz.

The *middle ear* begins with the eardrum, to which are attached three small bones (shaped like a hammer, an anvil, and a stirrup) called *ossicles*. The eardrum, which is composed of circular and radial fibers, is kept taut by the tensor tympani muscle. The eardrum changes the pressure variations of incoming sound waves into mechanical vibrations to be transmitted via the ossicles to the inner ear.

The ossicles perform a very important function in the hearing process. Together they act as a lever, which changes the very small pressure exerted by a sound wave on the eardrum into a much greater pressure (up to 30 times) on the oval window of the inner ear. This function, which an engineer might call a mechanical transformer, is illustrated in Fig. 5.3. The lever action of the ossicles provides a factor of about 1.5 in force multiplication, whereas the remaining factor of about 20 in pressure comes from the difference in the areas of the eardrum and round window (the same force distributed over a smaller area results in a greater pressure, as explained in Section 1.6).

Another function of the small bones is to protect the inner ear from very loud noises and sudden pressure changes. Loud noise triggers two sets of muscles; one tightens the eardrum and the other pulls the stirrup away from the oval window of the inner ear. This response to loud sounds, called the *acoustic reflex*, will be discussed in Chapter 6.

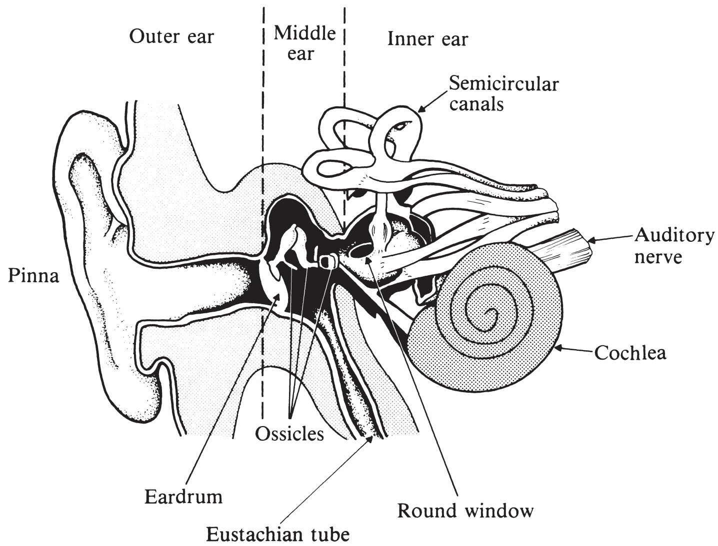


FIGURE 5.2
A schematic diagram of the ear, showing outer, middle, and inner regions. This drawing is not to scale; for purposes of illustration, the middle ear and inner ear have been enlarged.

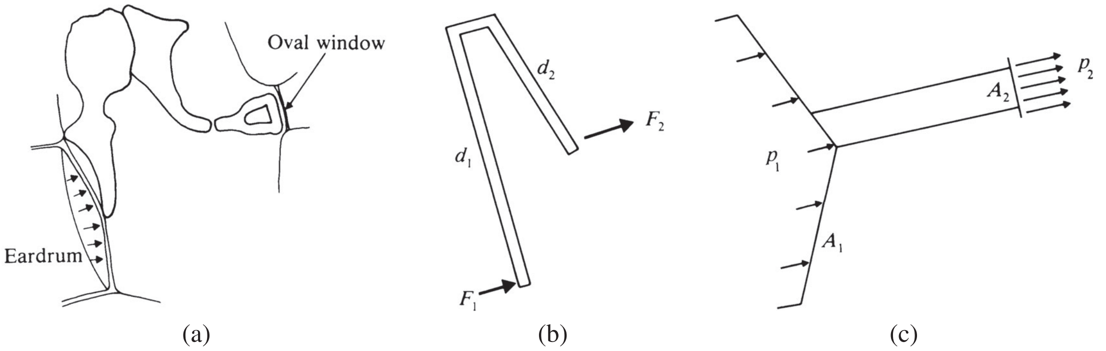


FIGURE 5.3 Pressure amplification by the ossicles. (a) Three bones link the eardrum to the inner ear. (b) Lever action: A smaller force acts through a larger distance, resulting in a larger force acting through a smaller distance. (c) Pressure multiplication by piston action: A small pressure on a large area produces the same force as a large pressure on a small area.

Because the eardrum makes an airtight seal between the middle and outer parts of the ear, it is necessary to provide some means of pressure equalization. The *Eustachian tube*, which connects the middle ear to the oral cavity, is such a safety device. If the Eustachian tube is slow to open, a "popping" may be heard in the ears when the outside air pressure changes, for example, during a rapid change in altitude. It is remarkable that all these middle ear functions take place in a space approximately the size of an ordinary sugar cube!

The marvelously complex *inner ear* contains the *semicircular canals* and the *cochlea*. The semicircular canals contribute little or nothing to hearing; they are the body's horizontal-vertical detectors necessary for balance. The spiral cochlea, a masterpiece of miniaturization, contains all the mechanism for transforming pressure variations into properly coded neural impulses.

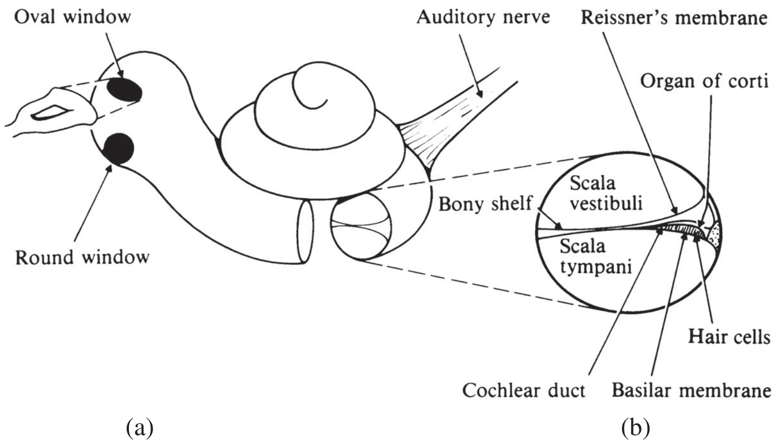


FIGURE 5.4 A schematic diagram of (a) the cochlea; (b) a section cut from the cochlea.

The cross section of the cochlea in Fig. 5.4 shows three distinct chambers that run the entire length: the *scala vestibuli*, the *scala tympani*, and the *cochlear duct*.

The cochlea is filled with liquid and surrounded by rigid bony walls. Actually there are two different liquids, called perilymph (in the canals or scala) and endolymph (in the cochlear duct); the total capacity of the cochlea is only a fraction of a drop. Perilymph is similar to spinal fluid, whereas endolymph is similar to the fluid within cells. The two liquids are kept separate by two membranes: *Reissner's membrane* and the *basilar membrane*. Reissner's membrane is exceedingly thin, approximately two cells thick.

Resting on the basilar membrane is the delicate and complex *organ of Corti*, a gelatinous mass about $1\frac{1}{2}$ in. long. This "seat of hearing" contains several rows of tiny *hair cells* to which are attached nerve fibers. A single row of inner hair cells contains about 4000 cells, whereas about 12,000 outer hair cells occur in several rows. Each hair cell has many hairs, or *stereocilia*, that are bent when the basilar membrane responds to a sound. The bending of the stereocilia stimulates the hair cells, which in turn excite neurons in the auditory nerve.

Modern auditory research has shown that the inner and outer hair cells function quite differently. The inner hair cells are mainly responsible for transmitting signals to the auditory nerve fibers. The more numerous outer hair cells apparently act as biological amplifiers. When their stereocilia are bent in response to a sound wave, the cell changes in length. This pushes against the tectoral membrane, selectively amplifying the vibration of the basilar membrane. It is estimated that the outer hair cells add about 40 dB of amplification, so that hearing sensitivity decreases by a considerable amount when these delicate cells are destroyed by overexposure to noise.

In order to understand how the basilar membrane vibrates, we can see the cochlea uncoiled and simplified in Fig. 5.5. The cochlea then appears as a tapered cylinder divided into two sections by the basilar membrane. (Because the cochlear duct is quite thin, we can ignore it—as a first approximation—and consider the two sections separated by a single membrane.) At the larger end of the cylinder are the oval and round windows, each closed by a thin membrane, and near the far end of the basilar membrane is a small hole called the *helicotrema* connecting the two sections. The basilar membrane terminates just short of the smaller end of the cylinder, so that fluid can transmit pressure waves around the end of the membrane.

When the stapes (stirrup) vibrates against the oval window, hydraulic pressure waves are transmitted rapidly down the scala vestibuli, inducing ripples in the basilar membrane. High tones create their greatest amplitude in the region near the oval window where the basilar membrane is narrow and stiff. On the other hand, low tones create ripples of greatest amplitude where the membrane is slack at the far end (see Fig. 5.6). Thus the initial frequency analysis takes place in the cochlea, although we will see in Chapter 7 that much

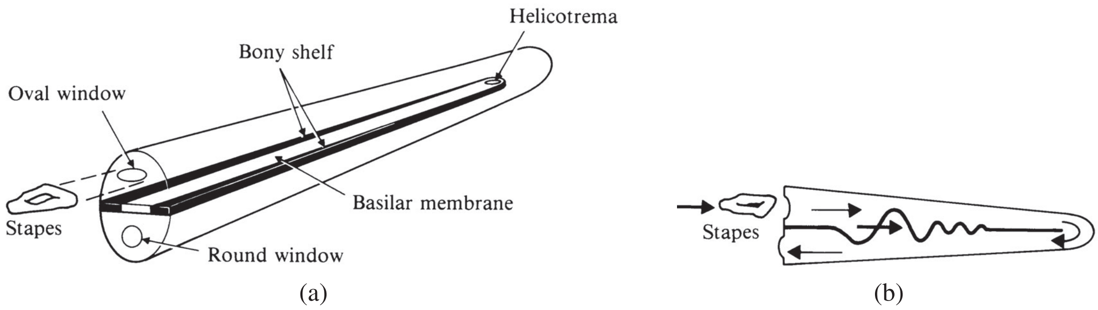


FIGURE 5.5 (a) A schematic diagram of uncoiled cochlea showing the basilar membrane and oval and round windows. (b) When the stapes (stirrup) presses against the oval window, a pressure pulse propagates through the cochlear fluid toward the round window, causing ripples to occur in the basilar membrane.

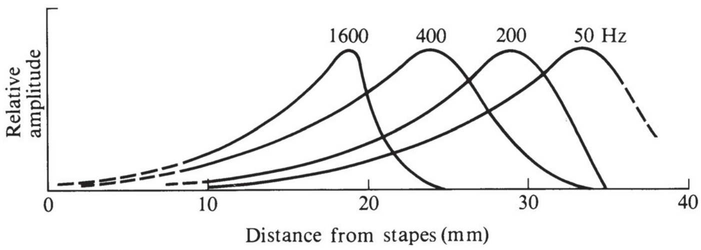


FIGURE 5.6 Basilar membrane displacement amplitude as a function of distance for several different frequencies. (After von Békésy 1960.)

of the sense of pitch is determined in the central nervous system, where the data from the auditory nerve are processed.

The conversion of the mechanical vibrations of the basilar membrane into electrical impulses in the auditory nerve is accomplished in the organ of Corti. When the basilar membrane vibrates, the "hairs" of the hair cells are bent, thus generating nerve impulses that travel to the brain. The impulse rate on the auditory nerve depends on both the intensity and the frequency of the sound.

The overall hearing mechanism is illustrated in Fig. 5.7. Sound waves propagate through the ear canal, excite the eardrum, and cause mechanical vibrations in the middle ear. The

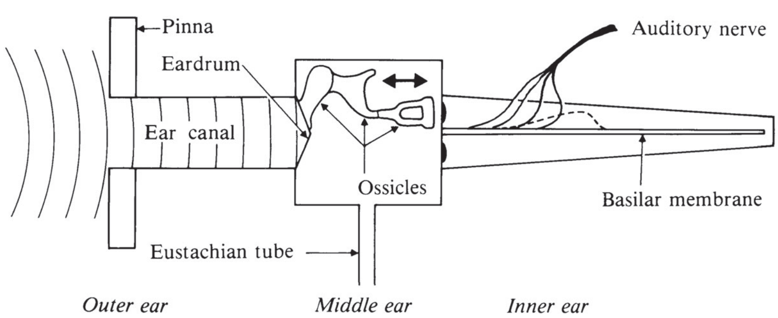


FIGURE 5.7 A schematic representation of the ear, illustrating the overall hearing mechanism. Sound waves in the outer ear cause mechanical vibrations in the middle ear, and eventually nerve impulses that travel to the brain to be interpreted as sound.

stapes vibrating against the oval window causes pressure variations in the cochlea, which in turn excite mechanical vibrations in the basilar membrane. These vibrations of the basilar membrane cause the hair cells to transmit electrical impulses to the brain via the auditory nerve.

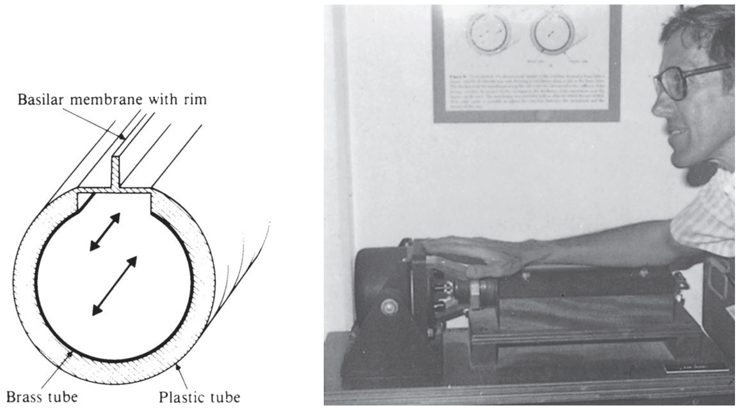
Some sounds are heard through vibrations of the skull that reach the inner ear. Hearing by bone conduction plays an important role in speaking. The sounds of humming or clicking one's teeth are heard almost entirely by bone conduction. (If you stop your ears with your fingers, thus interfering with the air path, the humming may actually sound louder.) During speaking or singing, two different sounds are heard, one by bone conduction and one by air conduction. The recorded sound of your own voice sounds very unnatural to you because only the airborne sound is received by the microphone, whereas you are used to hearing both components in your own voice.

Many researchers have contributed to our understanding of the hearing process, but two scientists deserve special mention: Hermann von Helmholtz and Georg von Békésy.

Hermann Ludwig Ferdinand von Helmholtz (1821–1894) was a physician and a man of many sciences. He did pioneering work in the field of physiology, mathematics, thermodynamics, optics, and acoustics. He invented the ophthalmoscope used to study the interior of the eye and formulated an important theory of color perception. In 1862 he published his monumental book *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, which has been reprinted many times and is useful even today to researchers in psychoacoustics. Helmholtz envisioned the fibers of the basilar membrane as selective resonators tuned, like the strings of a piano, to different frequencies. Thus a complex sound would be analyzed into its various components by selectively exciting fibers tuned to the frequency of one of the components. It turned out that Helmholtz was nearly, but not quite, correct in this assumption, as we shall learn in Chapter 7.

Georg von Békésy, a communications engineer in Budapest, Hungary, became interested in the mechanism of hearing while studying ways to improve telephones. In order to carry out his studies, Békésy carefully removed cochleas from the ears of animal and human cadavers. For his careful and extensive research, he was awarded a Nobel prize in 1961.

In order to illustrate vibrations of the basilar membrane, Békésy built several mechanical models of the cochlea, one of which is shown in Fig. 5.8. A brass tube with a slit at the top is covered by a plastic of varying thickness with a raised ridge. The tube is closed with a piston at one end and a fixed plate at the other, and filled with water. The elasticity of the plastic varies along its length in much the same manner as the basilar membrane. Thus when the piston is driven at various frequencies, the point of maximum excitation moves up and down the tube, which can be felt by placing one's forearm in gentle contact with the ridge of the plastic (Békésy 1960, 1970).



(a) (b)

FIGURE 5.8 (a) Cochlear model constructed by Békésy. (b) The observer notes that the point of maximum sensation moves up and down the forearm as the frequency of the sound changes.

The cochlear model, as well as other instruments used by Békésy, can be seen and operated at a small museum at the University of Hawaii.

Much of Békésy's success was due to the careful techniques he developed for removing the cochleas of fresh cadavers. Working under a microscope with microtools of his own design, he was able to lay open a part of the basilar membrane. The cochlear fluid was drained and replaced by a salt solution with a suspension of powdered aluminum and coal. By observing light scattered from the suspended powder, he discovered an undulation in the basilar membrane when the cochlea was excited by sound.

Békésy studied the ears of many different mammals. An amusing story is told about his excitement in learning that an elephant had died in the Budapest zoo. He traced the carcass to a local glue factory where he was able to recover the elephant's cochleas. To Békésy's delight, traveling waves were observable also in the basilar membrane of the elephant (Stevens and Warshofsky 1965).

5.3 ■ SIGNAL PROCESSING IN THE AUDITORY SYSTEM

Signal processing in the auditory system can be divided into two parts: that done in the peripheral auditory system (ears themselves), and that done in the auditory nervous system (brain). The ears process an acoustic pressure signal by first transforming it into a mechanical vibration pattern on the basilar membrane, as shown in Figs. 5.5 and 5.6, and then representing this pattern by a series of pulses to be transmitted by the auditory nerve. Perceptual information is extracted at various stages of the auditory nervous system.

It is possible, by inserting a tiny electrode into the auditory nerve, to pick up the electrical signals traveling in a single fiber of the auditory nerve from the cochlea to the brain (Tasaki 1954). The signal consists of a series of voltage spikes, each spike corresponding to the stimulation of a hair cell attached to the basilar membrane. The spikes are found to be closely correlated to the mechanical vibration pattern on the basilar membrane up to frequencies of about 4000 or 5000 Hz.

Each auditory nerve fiber responds over a certain range of frequency and sound pressure. Each nerve fiber has a characteristic frequency (CF) at which it has maximum sensitivity. Fibers with a high CF show a rapid rolloff in sensitivity above their CF but a long "tail" below it. A 90-dB stimulus at 500 Hz, for example, causes spikes to appear on all six fibers. By sophisticated techniques such as probing with laser light (Khanna and Leonard 1982) and using the Mössbauer effect (Johnstone and Boyle 1967), it has been found that basilar membrane displacements in live animals show a much sharper frequency response than those of Fig. 5.6 in the cochlea of a dead animal. Rhode and Robles (1974) found that within several hours after death, the basilar membrane response decreases 10–15 dB, the frequency of maximum response shifts downward, and the response curve broadens. In fact, the mechanical frequency response of the basilar membrane in live cochleas is quite comparable to the tuning curves observed in nerve fibers. There is some evidence for sharpening of neural tuning curves further along the neurological pathway, however.

If we were to observe the spikes on a nerve fiber when the stimulus is a tone of a single frequency, we would note that the time between spikes almost always corresponds to one or two or more periods of the tone. Although the nerve fiber does not fire at the peak of every vibration cycle in the basilar membrane, it rarely fires at any other time. The situation is a little more complicated when the stimulus is a complex tone, but still we find that the pattern of spikes on the auditory nerve carries accurate information about the frequency spectrum of the stimulus tone.

Consider a stimulus consisting of the pure tones C_4 (523 Hz) and C_5 (1046 Hz), spaced one octave apart. Their neural tuning curves (or frequency response curves) shown in Fig. 5.9(a) show very little overlap, so very few hair cells respond to both frequencies. Processing of the one component in the brain is only slightly affected by the presence of the other one.

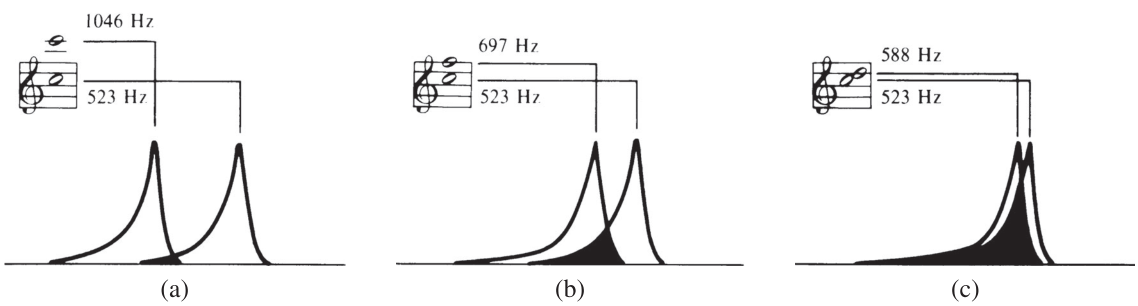


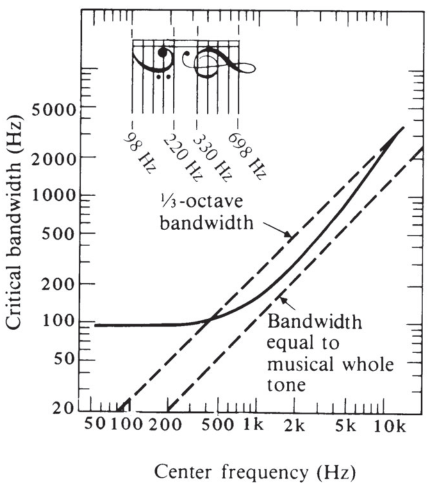
FIGURE 5.9 Frequency response curves for pairs of pure tones. As the interval between them decreases, their response curves show increasing overlap.

As the interval between the two components decreases, the situation changes. Their amplitude envelopes show more and more overlap, as in Fig. 5.9(b) and (c), so an increasing number of hair cells are stimulated by both components. This leads to many interesting auditory phenomena, some of which will be discussed in Chapter 6, 7, and 8.

5.4 ■ CRITICAL BANDS

When two pure tones are so close in frequency that there is considerable overlap in their amplitude envelopes on the basilar membrane, they are said to lie within the same *critical band*. Critical bands are of great importance in understanding many auditory phenomena, such as loudness, pitch, and timbre. They have been defined and measured in a variety of ways (Fletcher 1940; Plomp 1976; Zwicker, Flottorp, and Stevens 1957).

Each critical band may be regarded as a data collection unit on the basilar membrane. About 24 critical bands span the audible frequency range, and the regions on the basilar membrane to which each of these corresponds is about 1.3 mm long and embraces about 1300 neurons (Scharf 1970). The *critical bandwidth* varies with center frequency, as shown in Fig. 5.10, having nearly a constant value at low frequency and being roughly proportional to frequency at high frequency. Bandwidths are found to vary substantially, depending upon the type of experiment.



Center frequency (Hz)	Critical bandwidth (Hz)
100	90
200	90
500	110
1,000	150
2,000	280
5,000	700
10,000	1,200

FIGURE 5.10 Critical bandwidth as a function of the critical band center frequency. Bandwidths are typical of those reported in various experiments.

Critical Bands as Musical Intervals

Over much of the audible range, critical bands are slightly less than $\frac{1}{3}$ octave in width, as indicated in Fig. 5.10. An octave is the musical interval between two tones whose frequencies are in the ratio 2 : 1. The ratio of frequencies of two tones that are

$\frac{1}{3}$ octave apart is $\sqrt[3]{2} = 1.26$. In musical language, $\frac{1}{3}$ octave equals four semitones or a major third. Sound analyzers that measure sound pressure in each of about 30 $\frac{1}{3}$-octave bands are quite common (30 such bands are required to span the audible range as compared to only 24 critical bands because critical bands are substantially greater than $\frac{1}{3}$ octave at low frequency; see Fig. 5.10).

5.5 ■ BINAURAL HEARING AND LOCALIZATION

"Nature," said the ancient Greek philosopher Zeno, "has given us one tongue, but two ears, that we might hear twice as much as we speak." Excellent advice.

The most important benefit we derive from binaural hearing is the sense of localization of the sound source. Although some degree of localization is possible in monaural listening, binaural listening greatly enhances our ability to sense the direction of the sound source.

Lord Rayleigh, who contributed so much to our understanding of acoustics, was one of the first to explain binaural localization of sound. In 1876 Rayleigh performed experiments (which, unknown to him, had been performed nearly a century earlier by Giovanni Venturi, an Italian scientist remembered for his work on fluid dynamics) to determine his ability to localize sounds of different frequencies. He found that sounds of low frequency were more difficult to locate than those of high frequency. According to Rayleigh's explanation, a sound coming from one side of the head produces a more intense sound in one ear than in the opposite ear, because the head casts a "sound shadow" for sounds of high frequency, as shown in Fig. 5.11(a). At low frequency, however, the shadow effect is small because sound waves of long wavelength diffract around the head. At 1000 Hz, the sound level is about 8 decibels greater at the ear nearest the source, but at 10,000 Hz the difference could be as great as 30 decibels.

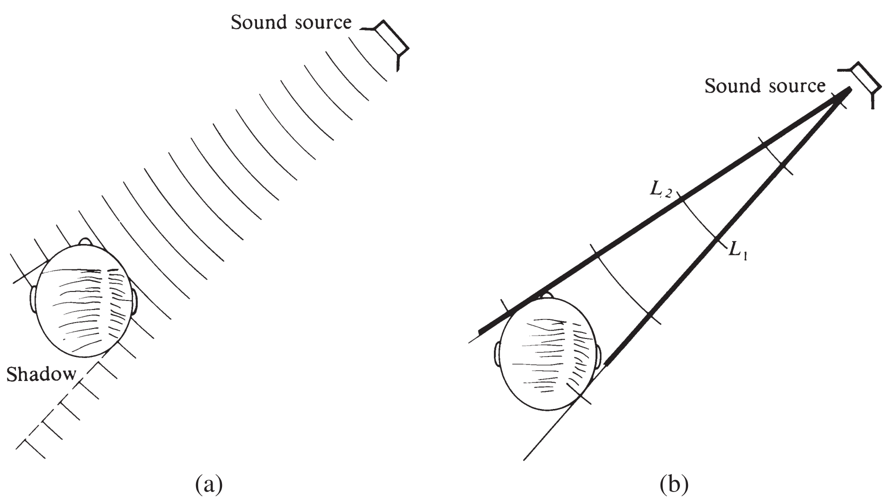


FIGURE 5.11 Localization of a sound source. (a) At frequencies above 4000 Hz, localization is due to intensity difference at two ears. (b) At frequencies below 1000 Hz, localization is due to an interaural time difference between sound traveling paths L_1 and L_2.

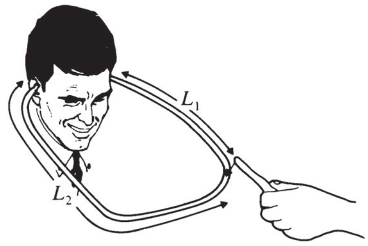


FIGURE 5.12 An experiment illustrating the sensitivity of the ear to interaural time difference. Tapping the tube so that $L_1 = L_2$ causes the sound to appear centered. When $L_2 > L_1$ the sound appears to come from the left.

Sounds of low frequency can be localized, although with slightly less accuracy than those of higher frequency. In 1907, Rayleigh offered a second theory of localization to explain low-frequency effects. A sound coming from the side strikes one ear before the other, and thus the sounds in the two ears will be slightly out of phase, as shown in Fig. 5.11(b). He confirmed this theory by experiments with two tuning forks tuned to slightly different frequencies, so that their relative phases constantly changed. The sound of the beating tone moved from right to left and back again.

Several experiments have confirmed the fact that for frequencies up to about 1000 Hz, localization occurs mainly through detection of the phase difference at the two ears (for steady sounds) or the difference in arrive time (for clicks), as illustrated in Fig. 5.12. Above 4000 Hz, localization by intensity difference takes over. Between 1000 and 4000 Hz, the accuracy of localization declines, with a high error rate around 3000 Hz demonstrating that the two mechanisms do not overlap appreciably.

At high frequencies (about 5000 Hz and upward), the pinna aids in localization of a sound, particularly in distinguishing between sound coming from the front or the back, because it receives sound with slightly greater efficiency from the front. Some animals have the ability to aim their pinnas toward sounds of interest, but human beings must turn the entire head to change pinna orientation.

An important corollary to sound localization is the *precedence effect* (sometimes referred to as the *Haas effect*), which applies to efforts to localize a sound source in a room. If similar sounds arrive within about 35 ms (0.035 s), the apparent direction of the sound source is the direction from which the first arriving sound comes. The ear automatically assumes this to be the direct sound and successive sounds to have been reflected one or more times. This effect will be discussed further in Chapter 23.

5.6 ■ MEASURING SENSATIONS: PSYCHOPHYSICS

Information about the world around us comes from our senses: vision, taste, smell, touch, and hearing. Each of our sensory organs responds to a particular type of stimulus over a limited range of energies. Our eyes, for example, respond to electromagnetic waves over an extremely narrow range of frequency compared to the wide range of electromagnetic radiations all around us.

Perception involves not only the reception of information by the appropriate sensory organ, but the coding, transmission, and processing of this information by the central nervous system. Our understanding of how this is accomplished has advanced remarkably in recent years, but still remains only fragmentary. (This may be due partly to the fact that research in this area involves several disciplines: physics, psychology, physiology, speech and hearing, engineering, mathematics, etc.) An excellent source of information about perception is a collection of articles from *Scientific American* (1972). It appears that many perceptual abilities are intrinsic; others are acquired or developed through experience and training.

The study of the relationships between stimuli and the subjective sensations they produce is the basis of *psychophysics*, so named by a pioneer in the field, G. T. Fechner. Inspired by earlier work on the subject by Ernst Weber, Fechner spent many years trying to determine quantitative relationships between stimulus and perceived sensation, and in 1860 published many of his findings in a monumental book entitled *Elements of Psychophysics*. He summed up much of this work in a simple mathematical law relating sensation to stimulus, which is often referred to today as *Fechner's law*. It expresses the relationship between stimulus and sensation rather simply: As stimuli are increased by *multiplication*, sensations increase by *addition*. For example, as the intensity of a sound is doubled, its loudness increases by one step on a scale. Mathematicians call such a relationship logarithmic; Fechner's law states that sensation grows as the logarithm of the stimulus.

Fechner argued that the same relationship applies to any stimulus and its corresponding sensation: to light and vision, etc. Recent investigations have pointed out its inadequacies; nevertheless, Fechner's law served as a basis for psychophysical theory for nearly a century thereafter. Fechner answered his early critics by saying "The Tower of Babel was never finished because the workers could not reach an understanding on how they should build it; my psychophysical edifice will stand because the workers will never agree on how to tear it down" (Stevens and Warshofsky 1965, 82). We will return to the subject in the following chapters.

5.7 ■ LOGARITHMS IN SOUND AND MUSIC

Although this book employs a minimum of mathematics, there are some times when the use of a little mathematics actually makes things easier to understand. One mathematical tool that should be familiar to everyone who wishes to understand the science of sound (and that should include serious musicians, students of speech and hearing, and anyone interested in sound recording, reproduction, and amplification) is the *logarithm*.

The logarithm to the base 10 of a number x is the power to which 10 must be raised in order to equal x. For example, $100 = 10^2$, so the logarithm of 100 (to base 10) is 2 ($\log 100 = 2$). There are other numbers besides 10 that can serve as a base, of course, but in this book $\log x$ will always mean the logarithm of x to the base 10. Fortunately pressing the log key on most calculators gives the logarithm to base 10.

Why are logarithms so useful in the study of sound? Three applications come to mind:

1. Decibel scales used to express such things as sound level and amplifier gain are based on logarithms.

2. Frequency response of the ear or audio devices are generally expressed on a compressed or logarithmic scale (Fig. 5.1 or 5.11, for example).
3. The keyboard on a piano or other musical instruments is logarithmic.
4. The musical scale is logarithmic (that is, each step is a certain ratio of frequencies).

At this time, therefore, it is appropriate to review or introduce (depending on the background of the reader) some properties of logarithms and logarithmic scales.

Logarithms and Powers of Ten

It is inconvenient to write out numbers such as 1,530,000,000 and 0.000087. These same numbers can be written better as 1.53×10^9 and 8.7×10^{-5}, respectively, because $10^9 = 1{,}000{,}000{,}000$ and $10^{-5} = 0.00001$. Other powers of ten are

$$\begin{aligned} 10^3 &= 1000, \\ 10^2 &= 100, \\ 10^1 &= 10, \\ 10^0 &= 1, \\ 10^{-1} &= 0.1, \\ 10^{-2} &= 0.01, \text{ etc.} \end{aligned}$$

On some electronic calculators, the scientific notation used to display very large and very small numbers expresses 10^3 as E3 and 10^{-3} as E−3, so 1,530,000,000 would be expressed as 1.53 E9 and 0.000087 as 8.7 E−5 (E is an abbreviation for *exponent*, which means the power to which 10 is raised). Other calculators omit the letter E but leave a space between the first part of the number and its exponent (e.g., 1531 becomes 1.531 03 in scientific notation).

To multiply two numbers in scientific or exponential notation, we *add* the exponents; to divide, we *subtract* exponents. Thus

$$\begin{aligned} (10^3)(10^4) &= 10^7; \\ (5 \times 10^2)(3 \times 10^5) &= 15 \times 10^7 = 1.5 \times 10^8; \\ (3 \times 10^{-3})(2 \times 10^5) &= 6 \times 10^2; \\ \frac{10^4}{10^2} &= 10^2; \\ \frac{6 \times 10^5}{3 \times 10^3} &= 2 \times 10^2. \end{aligned}$$

In general, then, $(10^A)(10^B) = 10^{A+B}$, and $10^A/10^B = 10^{A-B}$.

Closely related to exponents and scientific notation are logarithms. As we stated, logarithms are defined as follows: The logarithm to the base 10 of a number x is equal to the power to which 10 must be raised in order to equal x. That is, if $x = 10^y$, then $y = \log x$.

For example, $100 = 10^2$, so $2 = \log 100$ (here $x = 100$, $y = 2$); or $1000 = 10^3$, so $3 = \log 1000$ ($x = 1000$, $y = 3$).

Although logarithms to other bases are used in mathematics, we nearly always use base 10 in acoustics. Most calculators use *log x* to denote the logarithm of x to base 10, and *ln x* to denote the logarithm to the base 2.7183. On some calculators, the inverse logarithm is computed by pressing INV and then log; on others a 10^x key is used.

The following identities are useful for performing calculations with logarithms:

$$\log AB = \log A + \log B;$$

$$\log A/B = \log A - \log B;$$

$$\log A^n = n \log A.$$

The logarithms of some numbers are as follows:

x	$\log x$	x	$\log x$
1	0	6	0.778
2	0.301	7	0.845
3	0.477	8	0.903
4	0.602	9	0.954
5	0.699	10	1.000

Using this table and the identities listed above, we can compute the logarithms of many numbers. For example:

$$\log 400 = \log 4 + \log 100 = 0.602 + 2 = 2.602 \text{ (first identity);}$$

$$\log 2.5 = \log 5 - \log 2 = 0.699 - 0.301 = 0.398 \text{ (second identity);}$$

$$\log 25 = 2 \log 5 = (2)(0.699) = 1.398 \text{ (third identity).}$$

If one remembers that $\log 2 = 0.3$, many logarithms can be estimated closely. For example:

$$\log 4 = \log 2 \times 2 = 0.3 + 0.3 = 0.6;$$

$$\log 5 = \log \frac{10}{2} = 1 - 0.3 = 0.7;$$

$$\log 8 = \log 2 \times 2 \times 2 = 0.3 + 0.3 + 0.3 = 0.9.$$

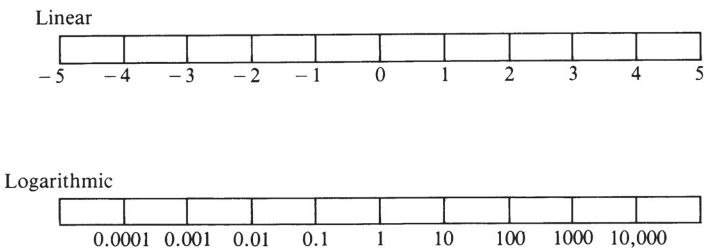


FIGURE 5.13 Linear and logarithmic scales. On the linear scale, moving one unit to the right adds an increment of one; on the logarithmic scale, moving one unit to the right multiplies by a factor of ten.

A *logarithmic scale* is one on which equal distances represent the same factor anywhere along the scale (in contrast to a *linear scale*, on which equal distances represent equal increments). Logarithmic and linear scales are shown in Fig. 5.13.

Sound frequencies are usually represented on a logarithmic scale for reasons that will become clear later on. In Fig. 5.14, the distance from 20 to 200 Hz is the same as from 200 to 2000 Hz or from 2000 to 20,000 Hz.

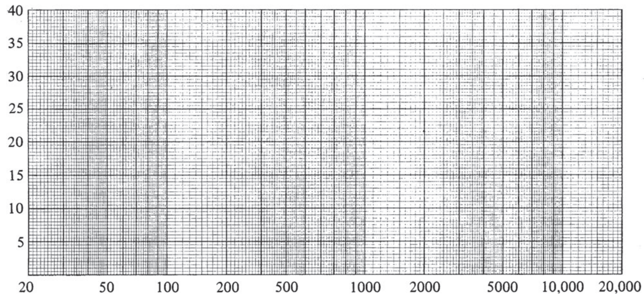


FIGURE 5.14 Graph paper with a logarithmic scale of frequencies. Such graph paper is called *semilog*, because only one axis is logarithmic. On *log-log* graph paper, both axes are logarithmic.

Logarithmic Scales in Music

Musical scales are discussed in Chapter 9. At this time, however, we will mention the *scale of equal temperament*, which is a logarithmic scale of frequency. The most common scale of equal temperament divides an octave (a frequency ratio of 2 : 1) into 12 equal steps. To do this, the octave is divided into frequency ratios of $2^{1/12} = 1.05946$. Going up an octave means traversing 12 such steps; the initial frequency is multiplied by $2^{1/12}$ 12 times, which is equivalent to multiplying it by 2. Most other scales that are used in music use steps of slightly different sizes as one goes up the scale.

Piano tuners, composers of electronic music, researchers of musical instruments, etc., often find it convenient to divide the octave into 1200 equal steps called *cents*. Each cent is 1/100 of a semitone, just as each cent of money is 1/100 of a dollar. Raising the pitch by 1 cent, then, means multiplying the frequency by $2^{1/1200} = 1.000578$. The formula for converting from a frequency ratio to cents (or visa versa), given in Section 9.5, uses logarithms, naturally enough.

5.8 ■ SUBJECTIVE ATTRIBUTES OF SOUND

Four attributes are frequently used to describe sound, especially musical sound. They are loudness, pitch, timbre, and duration. Each of these subjective qualities depends on one or more physical parameters that can be measured. Loudness, for example, depends mainly on sound pressure but also on the spectrum of the partials, the physical duration, etc. Pitch depends mainly on frequency but also shows less dependence on sound pressure, envelope, etc. Timbre is a sort of catchall, including all those attributes that serve to distinguish sounds with the same pitch and loudness. Table 5.1 relates subjective qualities to physical parameters, and is presented here as an introduction to the next three chapters, which discuss loudness, pitch, and timbre in more detail.

TABLE 5.1 Dependence of subjective qualitites of sound on physical parameters

	Subjective Quality			
Physical Parameter	Loudness	Pitch	Timbre	Duration
Pressure	+ + +	+	+	+
Frequency	+	+ + +	++	+
Spectrum	+	+	+ + +	+
Duration	+	+	+	+ + +
Envelope	+	+	++	+

+ = weakly dependent; ++ = moderately dependent; + + + = strongly dependent.
Note: Spectrum refers to the frequencies and amplitudes of all the partials (components) in the sound. The physical duration of a sound and its perceived (subjective) duration, though closely related, are not the same. Envelope includes the attack, the release, and variations in amplitude. These parameters will be discussed in Chapters 6, 7, and 8.

5.9 ■ SUMMARY

The human auditory system responds to pressure stimuli over a range of a million times. The frequency range of hearing extends from 20 to 20,000 Hz for some individuals, substantially less for others. Like other sensations, hearing tends to follow *logarithmic* relationships. The *outer ear* boosts hearing sensitivity in the middle frequency range and aids in determining the direction of a sound. The middle ear contains three small bones, called *ossicles*, which transmit sound pressure from the eardrum to the inner ear. The main part of the inner ear is the *cochlea*, which transforms pressure variations into neural impulses. Much of our ability to determine the direction of a sound source depends on *binaural* hearing, with a different mechanism of localization being dominant at high and low frequencies.

REFERENCES AND SUGGESTED READINGS

Békésy, G. von (1960). *Experiments in Hearing*. New York: McGraw-Hill.

Békésy, G. von (1970). "Enlarged Mechanical Model of the Cochlea with Nerve Supply," in *Foundations of Modern Auditory Theory*, Vol. I. Ed., J. V. Tobias. New York: Academic Press.

Denes, P. B., and E. N. Pinson (1973). *The Speech Chain*. Garden City, N.Y.: Anchor.

Flannagan, J. L. (1972). *Speech Analysis, Synthesis and Perception*. New York: Springer-Verlag.

Fletcher, H. (1940). "Auditory Patterns," *Revs. Modern Physics* **12**: 47.

Helmholtz, H. von (1877). *On the Sensation of Tone as a Psychological Basis for the Study of Music*, 4th ed. Trans., A. J. Ellis. New York: Dover, 1954.

Johnstone, B. M., and A. J. F. Boyle (1967). "Basilar Membrane Vibration Examined with Mössbauer Technique," *Science* **158**: 389.

Khanna, S. M., and D. G. B. Leonard (1982). "Basilar Membrane Tuning in the Cat Cochlea," *Science* **215**: 305.

Plomp, R. (1976). *Aspects of Tone Sensation*. London: Academic Press.

Readings from *Scientific American* (1972). *Perception: Mechanisms and Models*. San Francisco: W. H. Freeman. This collection of readings contains reprints of 38 articles published during 1950–1971 with introductions to each section by R. Held and W. Richards. "The Ear" by Békésy (1957) is included.

Rhode, W. S., and L. Robles (1974). "Evidence from Mössbauer Experiments for Nonlinear Vibration in the Cochlea," *J. Acoust. Soc. Am.* **55**: 588.

Rossing, T. D., and C. J. Chiaverina (1999). *Light Science: Physics and the Visual Arts*. New York: Springer-Verlag.

Scharf, B. (1970). "Critical Bands" in *Foundations of Modern Auditory Theory*, ed. J. Tobias. New York: Academic Press.

Scharf, B., and S. Buus (1986). "Audition I. Stimulus, Physiology, Thresholds" in *Handbook of Perception and Human Performance*. Ed. K. R. Boff, L. Kaufman, and J. P. Thomas. New York: J. Wiley.

Stevens, S. S., and F. Warshofsky (1965). *Sound and Hearing*. New York: Time Life.

Tasaki, I. (1954). "Nerve Impulses in Individual Auditory Nerve Fibres of Guinea Pigs," *J. Neurophysiology* **17**: 97.

Teas, D. C. (1970). "Cochlear Processes," in *Foundations of Modern Auditory Theory*. Ed., J. V. Tobias. New York: Academic Press.

Zwicker, E., G. Flottorp, and S. S. Stevens (1957). "Critical Bandwidth in Loudness Summation," *J. Acoust. Soc. Am.* **29**: 548.

GLOSSARY

auditory canal A tube in the outer ear that transmits sound from the external pinna to the eardrum.

basilar membrane A membrane in the cochlea that separates the cochlear duct from the scala tympani and to which the organ of Corti is attached.

cochlea The spiral organ of the inner ear containing the sound-sensing mechanism.

critical band Frequency band within which two or more tones excite many of the same hair cells on the basilar membrane and thus are difficult to distinguish as separate tones.

eardrum (tympanum) The fibrous membrane that terminates the auditory canal and is caused to vibrate by incoming sound waves.

envelope Time history of the amplitude.

Eustachian tube A tube connecting the middle ear to the oral cavity that allows the average pressure in the middle ear to equal atmospheric pressure.

exponent The number expressing the power to which 10 or some other number is raised.

Fechner's (Weber's) law An empirical law expressing the way in which sensation varies with stimulus.

hair cells The tiny sensors of sound in the cochlea.

intensity Power per unit area. The intensity of a sound wave is proportional to the square of the sound pressure.

linear scale A scale in which moving a given distance right or left adds or subtracts a given increment.

localization The ability to determine the location or direction of a sound source.

logarithm (of a number) The power to which 10 (or some other base) must be raised to give the desired number.

logarithmic scale A scale on which moving a given distance right or left multiplies or divides by a given factor.

organ of Corti The part of the cochlea containing the hair-cells; the "seat of hearing."

ossicles Three small bones of the middle ear that transmit vibrations from the eardrum to the cochlea.

pinna The external part of the ear.

precedence effect If similar sounds arrive within about 35 ms, the apparent direction is the direction from which the first arriving sound comes.

psychoacoustics The study of the relationships between sound and the sensations it produces. The psychophysics of sound.

psychophysics The study of the relationship between stimuli and the sensations they produce.

Reissner's membrane A membrane in the cochlea that separates the cochlear duct from the scala vestibuli.

scala vestibuli A canal in the ear that transmits pressure variations from the oval window to the cochlear duct.

stereocilia The tiny fibers attached to hair cells that bend and cause electrical signals to be transmitted on the auditory nerve fibers.

REVIEW QUESTIONS

1. What is the intensity ratio between the threshold of pain and the threshold of audibility?
2. What is the frequency ratio between the highest and lowest frequencies we can hear?

3. What is the frequency ratio between the highest and lowest light frequencies we can see?
4. What membrane terminates the outer ear?
5. What are the bones in the middle ear called?
6. What tube connects the middle ear to the oral cavity?
7. What is the main function of the semicircular canal in our inner ear?
8. What is the spiral organ in the inner ear called?
9. What happens to the cilia when the basilar membrane responds to a sound?
10. What part of the basilar membrane responds the most to low-frequency vibrations?
11. In addition to transmission through the outer ear, how else can sound reach the inner ear?
12. At low frequency, the critical bandwidth remains nearly constant. (T or F)
13. How are sounds of low frequency localized?
14. How are sounds of high frequency localized?
15. What is the precedence effect?
16. According to Fechner's law, how do sensations increase as stimuli are increased by multiplication?
17. Pitch depends mainly on what physical parameter?
18. What other physical parameters does pitch also depend on?
19. What is the envelope of a sound?
20. What is a logarithm?

QUESTIONS FOR THOUGHT AND DISCUSSION

1. If everyone's hearing sensitivity were reduced by 10 dB, in what ways would our lives probably be different?
2. What advantage is there in having our various senses respond on a (nearly) logarithmic rather than a linear scale?
3. Before the development of radar, a device used to determine the direction of aircraft consisted of two sound-receiving horns, each of which transmitted sound to one ear. Comment on the effectiveness of such a device.
4. Listen to a tape recording of your own voice and compare its sound to what you hear when you speak and sing. Try to describe the difference in terms of relative balance between high and low frequency components, etc.
5. Compare following distances on the horizontal scale in Fig 5.14.
   (a) 20 to 100
   (b) 100 to 500
   (c) 2000 to 10,000

EXERCISES

1. Assume that the outer ear canal is a cylindrical pipe 3 cm long, closed at one end by the eardrum. Calculate the resonance frequency of this pipe (see Fig. 4.8). Our hearing should be especially sensitive for frequencies near this resonance.
2. At what frequency does the wavelength of sound equal the distance between your ears? What is the significance of this with respect to your ability to localize sound?
3. The effective area of the eardrum is estimated to be approximately 0.55 cm^2. During normal conversation, the sound pressure variations of about 10^{-2} N/m^2 reach the eardrum. What force is exerted on the eardrum (force = pressure × area)?
4. Pressure is force per unit area. Calculate the pressure when a force of 500 N (approximate weight of a 110-lb person) is supported by:
   (a) Spike heels having an area of 10^{-5} m^2 each;
   (b) Standard heels having an area of 10^{-2} m^2 each.
   Comment on the likelihood of denting the floor in each case.
5. Measure the distance between your ears. Divide this distance by the speed of sound (Table 3.1) to find the maximum difference in arrival time $\Delta t = (L_2 - L_1)/v$ that occurs when a sound comes directly from the side.
6. Calculate the difference in arrival time at the two ears for a sound that comes from a 45° direction (from the northwest, for example, when the listener faces north).
7. Perform the following arithmetic operations.
   (a) $(1.6 \times 10^{-8})(5.0 \times 10^3)$
   (b) $\frac{4.5\times10^{-2}}{1.5\times10^{-3}}$
   (c) $1.3 \times 10^3 + 4.3 \times 10^2$
   (d) $4.2 \times 10^2 - 5.4 \times 10^{-2}$

8. Find the following logarithms using the logarithms of the numbers 1–10 and the three identities given.

(a) log 50

(b) log 0.5

(c) $\log 2 \times 10^{10}$

(d) log 16

9. Given $\log x$, find the number x in each case.

(a) $\log x = 0.3$

(b) $\log x = 3.0$

(c) $\log x = 1.3$

(d) $\log x = -0.3$

EXPERIMENTS FOR HOME, LABORATORY, AND CLASSROOM DEMONSTRATION

Home and Classroom Demonstration

1. *Frequency range of audible sound* Use a wide-range loudspeaker, an audio generator, and an amplifier to determine the frequency range of audible sound. (Be careful to avoid harmonic distortion at low frequency, because the harmonics may be heard when the fundamental is inaudible.)

2. *Loudness and sound level* Listen to how much a sound changes in loudness when the sound pressure is doubled and tripled (6-dB and 9.5-dB increase on a sound-level meter). Is it different for a pure (single frequency) tone as compared to broadband noise or music?

3. *Sensitivity to interaural time difference* Place each end of a rubber hose (about 2 to 3 m long) in your ears (see Fig. 5.9). Close your eyes and have someone tap the hose near its center and at varying distances to the left and right of center as you point in the apparent direction of the sound source. How far from the center must the tap occur in order to identify the sound source as *left* or *right*? What, then, is the minimum interaural time difference you can detect?

4. *Time resolution of the ear* Repeat Joseph Henry's experiment (T. D. Rossing, "Joseph Henry and Acoustics," *Physics Teacher* **16**, 600 (1978)). Clap your hands as you move back from a large wall, and note the minimum distance at which a distinct echo can be heard. Determine the time resolution by dividing the distance the sound wave traveled (twice your distance from the wall) by the speed of sound.

5. *Critical bands by masking* Listen to Demonstration 2 on the *Auditory Demonstrations* CD.

6. *Critical bands by loudness comparison* Listen to Demonstration 3 on the *Auditory Demonstrations* CD.

Laboratory Experiments

Critical bands by masking or critical bands by loudness comparison (*Auditory Demonstrations* CD).

Critical bands by loudness comparison (*Auditory Demonstrations* CD).

Binaural localization (*Auditory Demonstrations* CD).

CHAPTER 6 Sound Pressure, Power, and Loudness

In this chapter, we will discuss the quality of loudness and the physical parameters that determine it. The principal such parameter, we learned in Table 5.1, is sound pressure. Related to the sound pressure are the sound *power* emitted by the source and the sound *intensity* (the power carried across a unit area by the sound wave). The output signal of a microphone is generally proportional to the sound pressure, so sound pressure can be measured with a microphone and a voltmeter.

In this chapter, you should learn:

- About sound pressure level, sound intensity level, and sound power level;
- About the decibel scale for comparing sound levels;
- How to combine sound levels from several sources;
- What determines the perceived loudness of a sound;
- How one sound can mask another.

6.1 ■ DECIBELS

Decibel scales are widely used to compare two quantities. We may express the power gain of an amplifier in decibels (abbreviated dB), or we may express the relative power of two sound sources. We could even compare our bank balance at the beginning with the balance at the end of the month. ("My bank account decreased 27 decibels last month.") The decibel difference between two power levels, ΔL, is defined in terms of their power ratio W_2/W_1:

$$\Delta L = L_2 - L_1 = 10 \log W_2/W_1. \tag{6.1}$$

Although decibel scales always compare two quantities, one of these can be a fixed reference, in which case we can express another quantity in terms of this reference. For example, we often express the *sound power level* of a source by using $W_0 = 10^{-12}$ W as a reference. Then the sound power level (in decibels) will be

$$L_W = 10 \log W/W_0 \tag{6.2}$$

EXAMPLE 6.1 What is the sound power level of a loudspeaker that radiates 0.1 W?

Solution $L_W = 10 \log W/W_0 = 10 \log(0.1/10^{-12}) = 10(11) = 110$ dB.

Although L_W is the preferred abbreviation for sound power level, one often sees it abbreviated as *PWL*.

EXAMPLE 6.2 What is the decibel gain of an amplifier if an input of 0.01 W gives an output of 10 W?

Solution $L_2 - L_1 = 10 \log W_2/W_1 = 10 \log 10/0.01 = 10(3) = 30$ dB.

One number fact worth remembering is that the logarithm of 2 is 0.3 (actually 0.3010, but 0.3 will do). Why is this worthwhile? Because $10 \log 2 = 3$, doubling the power results in an increase of 3 dB in the power level. In rating the frequency response of audio amplifiers and other devices, one often specifies the frequencies of the 3-dB points, the upper and lower frequencies at which the power drops to one-half its maximum level.

Actually, remembering that $\log 2 = 0.3$ can be used to estimate other decibel levels. We know that $\log 10 = 1$, so a power gain of 10 represents a power level gain of 10 dB. For a power gain of 5, note that $5 = 10/2$, so if we gain 10 dB (multiplying by 10) and lose 3 dB (dividing by 2), multiplying power by 5 results in a gain of 7 dB. Multiplying by 4 should give 6 dB of gain, because $4 = 2 \times 2$. If multiplication by 2 is equivalent to 3 dB and multiplication by 4 is equivalent to 6 dB, we can probably guess that multiplying by 3 would give about 5 dB (actually 4.8, but 5 is often close enough).

Also, $100 = 10 \times 10$, so a power gain of 100 should represent 20 dB (two 10-dB increases).

Here, then, is a summary of what we have just figured out.

Power ratio:	2	3	4	5	10	100
Decibel gain:	3	5	6	7	10	20

EXAMPLE 6.3 What is the decibel gain when the power gain is 400?

Solution $400 = 2 \times 2 \times 10 \times 10$, so the decibel gain is $3 + 3 + 10 + 10 = 26$ dB.

6.2 SOUND INTENSITY LEVEL

We have just seen how the strength of a sound source can be expressed in decibels by comparing its power to a reference power (nearly always $W_0 = 10^{-12}$ W). Similarly, the sound intensity level at a point some distance from the source can be expressed in decibels by comparing it to a reference intensity, for which we generally use $I_0 = 10^{-12}$ W/m^2. Thus the *sound intensity level* at some location is defined as

$$L_I = 10 \log I/I_0. \tag{6.3}$$

EXAMPLE 6.4 What is the sound intensity level at a point where the sound intensity is 10^{-4} W/m^2?

Solution $L_\mathrm{I} = 10 \log I/I_0 = 10 \log 10^{-4}/10^{-12} = 10(8) = 80$ dB.

Even though they are both expressed in decibels, do not confuse sound power level, which describes the sound source, with sound intensity level, which describes the sound at some point. The relationship between the sound intensity level at a given distance from a sound source and the sound power level of the source depends upon the nature of the sound field. In the following boxes, we consider two cases.

Free Field

When a point source (or any source that radiates equally in all directions) radiates into free space, the intensity of the sound varies as $1/r^2$ (and the sound pressure varies as $1/r$), where r is the distance from the source S. This may be understood as a given amount of sound power being distributed over the surface of an expanding sphere with area $4\pi r^2$ (see Fig. 6.1). Thus the intensity is given by

$$I = W/4\pi r^2, \tag{6.4}$$

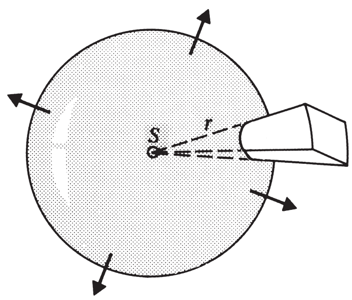


FIGURE 6.1 Spherical sound waves in a free field. The power from source S is distributed over a spherical surface $4\pi r^2$ in area.

where W is the power of the source. An environment in which there are no reflections is called a *free field*. In a free field, the sound intensity level decreases by 6 dB each time the distance from the source is doubled. The sound intensity level (or sound pressure level) at a distance of 1 m from a source in free field is 11 dB less than the sound power level of the source. This is easily shown as follows:

$$I = \frac{W}{4\pi r^2} = \frac{W}{4\pi(1)};$$

$$L_I = 10\log\frac{I}{10^{-12}} = 10\log\frac{W}{10^{-12}} - 10\log 4\pi = L_W - 11 \simeq L_p.$$

Similarly, it can be shown that at a distance of two meters, L_I is 17 dB less than L_W.

Hemispherical Field

More common than a free field is a sound source resting on a hard, sound-reflecting surface and radiating hemispherical waves into free space (see Fig. 6.2). Under these conditions, the sound intensity level L_I and the sound pressure level L_p at a distance of one meter at 8 dB less than the sound power level, once again diminishing by 6 dB each time the distance is doubled. In actual practice, few sound sources radiate sound equally in all directions, and there are often reflecting surfaces nearby that destroy the symmetry of the spherical or hemispherical waves.

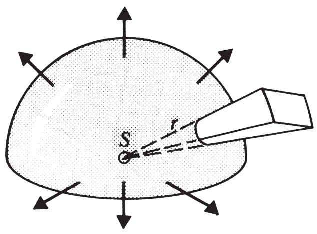


FIGURE 6.2 Hemispherical sound waves from a source S on a hard reflecting surface. The power is distributed over a surface $2\pi r^2$ in area.

EXAMPLE 6.5 If a trombone bell has an area of 0.1 m^2 and the power radiated from the bell during a very loud note is 1.5 W, what is the average intensity and sound intensity level at the bell?

Solution

$$I = \frac{W}{A} = \frac{1.5}{0.1} = 15 \text{ W/m}^2;$$

$$L_I = 10\log\frac{I}{I_0} = 10\log\frac{15}{10^{-12}}$$

$$= 10\log 15 \times 10^{12} = 132 \text{ dB}.$$

EXAMPLE 6.6 The sound pressure level 1 m from a noisy motor resting on a concrete floor is measured to be 95 dB. Find the sound power and the sound power level of the source.

Solution 1

$$L_I = 10 \log \frac{I}{I_0} \simeq L_p = 95 \text{ dB};$$

$$I = I_o \,\text{INV} \log \frac{95}{10} = 3.16 \times 10^{-3} \text{ W/m}^2;$$

$$W = 2\pi r^2 I = 2\pi (1)^2 (3.16 \times 10^{-3}) = 1.98 \times 10^{-2} \text{ W}.$$

Solution 2 For a hemispherical field,

$$L_W = L_p(1 \text{ m}) + 8 = 95 + 8 = 103 \text{ dB};$$

$$W = W_0 \,\text{INV} \log \frac{103}{10} = 1.98 \times 10^{-2} \text{ W}.$$

Demonstration 4 on the *Auditory Demonstrations* CD (Houtsma, Rossing, and Wagenaars 1987) provides test tones to "calibrate" your hearing. Broadband noise is reduced in steps of 6 dB, 3 dB, and 1 dB. You probably noticed that the 1-dB steps are about the smallest steps in which you can notice a difference. The demonstration also records free-field speech at distances of 0.25, 0.5, 1, and 2 m from a microphone. Doubling the distance from the source also reduces the sound level in 6-dB steps, as we have just discussed.

6.3 ■ SOUND PRESSURE LEVEL

In a sound wave there are extremely small periodic variations in air pressure to which our ears respond. The minimum pressure fluctuation to which the ear can respond is less than 1 billionth (10^{-9}) of atmospheric pressure. This threshold of audibility, which varies from person to person, corresponds to a sound pressure amplitude of about 2×10^{-5} N/m^2 at a frequency of 1000 Hz. The threshold of pain corresponds to a pressure amplitude approximately 1 million (10^6) times greater, but still less than 1/1000 of atmospheric pressure.

The intensity of a sound wave is proportional to the pressure squared. In other words, doubling the sound pressure quadruples the intensity. The actual formula relating sound intensity I and sound pressure p is

$$I = p^2/\rho c \tag{6.5}$$

where ρ is the density of air and c is the speed of sound. The density ρ and the speed of sound c both depend on the temperature (see Section 3.5). At normal temperatures the product ρc is around 410 to 420, but for ease of calculation, we often set it equal to 400.

It is useful to substitute for I from Eq. 6.5 setting $\rho c = 400$) in Eq. 6.3: $L_\text{I} = 10 \log p^2/400 I_o = 10 \log p^2/4 \times 10^{-10} = 20 \log p/2 \times 10^{-5}$. The latter expression is defined as the *sound pressure level* L_p (sometimes abbreviated *SPL*, although L_p is preferred).

$$L_p = 20 \log p/p_0, \tag{6.6}$$

TABLE 6.1 Typical sound levels one might encounter

Jet takeoff (60 m)	120 dB	
Construction site	110 dB	*Intolerable*
Shout (1.5 m)	100 dB	
Heavy truck (15 m)	90 dB	*Very noisy*
Urban street	80 dB	
Automobile interior	70 dB	*Noisy*
Normal conversation (1 m)	60 dB	
Office, classroom	50 dB	*Moderate*
Living room	40 dB	
Bedroom at night	30 dB	*Quiet*
Broadcast studio	20 dB	
Rustling leaves	10 dB	*Barely audible*
	0 dB	

where the reference level $p_0 = 2 \times 10^{-5}$ N/m^2 $= 20\mu$Pa (a pascal (Pa) is an alternative name for N/m^2). Note that Eq. 6.6 is the definition of sound pressure level, which is equal to sound intensity level only when $\rho c = 400$ (which would happen at 30° C and 748 mm Hg, for example). However, at ordinary temperatures, the two are so close to each other that they are often considered to be equal and called merely *sound level*. For precise measurement, a distinction should be made, however.

Sound pressure levels are measured by a sound-level meter, consisting of a microphone, an amplifier, and a meter that reads in decibels. Sound pressure levels of a number of sounds are given in Table 6.1. If you have access to a sound-level meter, it is recommended that you carry it with you to many locations to obtain a feeling for different sound pressure levels.

EXAMPLE 6.7 What sound pressure level corresponds to a sound pressure of 10^{-3} N/m^2?

Solution $L_p = 20 \log \dfrac{10^{-3}}{2 \times 10^{-5}} = 34.0$ dB.

EXAMPLE 6.8 How much force does a sound wave at the pain threshold ($L_p \simeq 120$ dB) exert on an eardrum having a diameter of 7 mm?

Solution

$$L_p = 120 = 20 \log \frac{p}{p_0};$$

$$p = p_0 \,\text{INV} \log \frac{120}{20} = 20 \text{ N/m}^2;$$

$$F = pA = 20\pi (3.5 \times 10^{-3})^2 = 1.54 \times 10^{-3} \text{ N}.$$

6.4 ■ MULTIPLE SOURCES

Very frequently we are concerned with more than one source of sound. The way in which sound levels add may seem a little surprising at first. For example, two uncorrelated sources, each of which would produce a sound level of 80 dB at a certain point, will together give 83 dB at that point. Figure 6.3 gives the increase in sound level due to additional equal sources. It is not difficult to see why this is the case, because doubling the sound power raises the sound power level by 3 dB and thus raises the sound pressure level 3 dB at our point of interest. Under some conditions, however, there may be interference between waves from the two sources, and this doubling relationship will not hold true.

When two waves of the same frequency reach the same point, they may interfere constructively or destructively. If their amplitudes are both equal to A, the resultant amplitude may thus be anything from zero up to 2 A. The resultant intensity, which is proportional to the amplitude squared, may thus vary from 0 to 4 A^2. If the waves have different frequencies, however, these well-defined points of constructive and destructive interference do not occur. In the case of sound waves from two noise sources (as in the case of light from two light bulbs), the waves include a broad distribution of frequencies (wavelength), and we do not expect interference to occur. In this case, we can add the energy carried by each wave across a surface or, in other words, the intensities.

In the case of independent (uncorrelated) sound sources, what we really want to add are the mean-square pressures (average values of p^2) at a point. Because intensity is proportional to p^2, however, we can add the intensities. For example, two sources that by themselves cause $L_I = 40$ dB at a certain location will cause $L_I = 43$ dB at the same location when sounded together. (This result is also obtained from the graph in Fig. 6.3.)

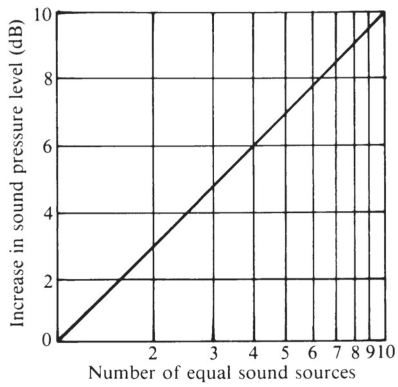


FIGURE 6.3 Addition of equal (uncorrelated) sound sources.

EXAMPLE 6.9 With one violin playing, the sound pressure level at a certain place is measured to be 50 dB. If three violins play equally loudly, what will the sound pressure level most likely be at the same location?

Solution

$$L_p = 10 \log \frac{p_1^2 + p_2^2 + p_3^2}{p_0^2} = 10 \log \frac{I_1 + I_2 + I_3}{I_0}$$
$$= 10 \log \frac{I_1}{I_0} + 10 \log 3$$
$$= 50 + 4.8 = 54.8 \text{ dB}$$

(This result could also be determined from Fig. 6.3.)

EXAMPLE 6.10 If two sound sources independently cause sound levels of 50 and 53 dB at a certain point, what is L_I at that point when both sources contribute at the same time?

Solution

$$50 = 10 \log \frac{I_1}{I_0}$$
$$\text{so} \quad I_1 = I_0 \,\text{INV} \log \frac{50}{10} = (10^{-12})(10^5) = 10^{-7} \text{ W/m}^2;$$

likewise $I_2 = 2 \times 10^{-7}$ W/m^2;

$$L_I = 10 \log \frac{I_1 + I_2}{I_0} = 10 \log \frac{10^{-7} + 2 \times 10^{-7}}{10^{-12}}$$
$$= 10 \log 3 \times 10^5 = 54.8 \text{ dB}.$$

(Note that the answer is *not* $50 + 53 = 103$ dB.)

6.5 ■ LOUDNESS LEVEL

Although sounds with a greater L_I or L_p usually sound louder, this is not always the case. The sensitivity of the ear varies with the frequency and the quality of the sound. Many years ago Fletcher and Munson (1933) determined curves of equal *loudness level* (L_L) for pure tones (that is, tones of a single frequency). The curves shown in Fig. 6.4, recommended by the International Standards Organization, are quite similar to those of Fletcher and Munson. These curves demonstrate the relative insensitivity of the ear to sounds of low frequency at moderate to low intensity levels. Hearing sensitivity reaches a maximum between 3500 and 4000 Hz, which is near the first resonance frequency of the outer ear canal, and again peaks around 13 kHz, the frequency of the second resonance.

The contours of equal loudness level are labeled in units called *phons*, the level in phons being numerically equal to the sound pressure level in decibels at $f = 1000$ Hz. The phon is a rather arbitrary unit, however, and is not widely used in measuring sound. It is

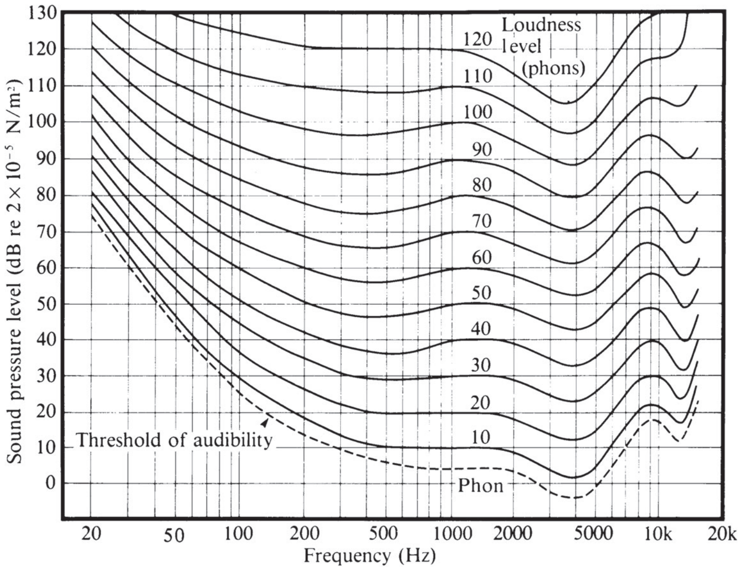
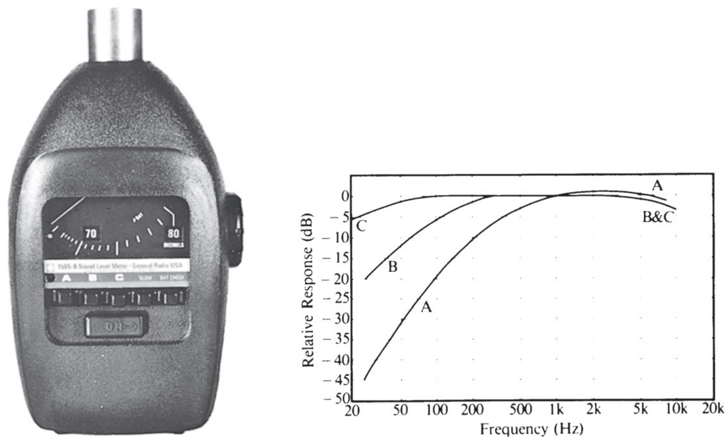


FIGURE 6.4 Equal-loudness curves for pure tones (frontal incidence). The loudness levels are expressed in phons.

important, however, to note the relative insensitivity of the ear to sounds of low frequency, which is one reason why weighting networks are used in sound-measuring equipment.

Sound level meters have one or more weighting networks, which provide the desired frequency responses. Generally three weighting networks are used; they are designated A, B, and C. The C-weighting network has an almost flat frequency response, whereas the A-weighting network introduces a low-frequency rolloff in gain that bears rather close resemblance to the frequency response of the ear at low sound pressure level. A sound level meter is shown in Fig. 6.5, along with the frequency responses of A-, B-, and C-weighting networks.



(a) (b)

FIGURE 6.5 Sound-level meter with the frequency response of its A-, B-, and C-weighting networks. (Photography courtesy of GenRad, Inc.)

Measurements of sound level are usually made using the A-weighting network; such measurements are properly designated as L_p(A) or SPL(A) in dB, although the unit dBA or dB(A) is often used to denote A-weighted sound level. Inside a building, the C-weighted sound level may be substantially higher than the A-weighted sound level, because of low-frequency machinery noise, to which the ear is quite insensitive. Many sound-level meters have both fast and slow response, the slow response measuring an "average" level.

Although it is difficult to describe a sound environment by a single parameter, for many purposes the A-weighted sound level will suffice. At low to medium sound levels, it is reasonably close to the true loudness level so that dBA (easily measured with a sound-level meter) may be substituted for phons without too much error.

There are many examples of interesting sound environments to measure. In the classroom, one can ask the entire class to shout loudly, then half the class to do so, one-fourth of the class, etc. The sound level should drop about 3 dB in each step. One can also measure traffic noise, noise near a construction site, sound level at a concert, noise in an automobile, and so on. In each case the A-weighted sound level should be measured, although it may be interesting to measure the C-weighted level (which places more emphasis on sounds of low frequency) as well.

6.6 ■ LOUDNESS OF PURE TONES: SONES

In Chapter 5, we mentioned Fechner's law, relating sensation to stimulus. The logarithmic relationship in that law was found to provide only a rough approximation to listeners' estimates of their own sensations of loudness. In an effort to obtain a quantity proportional to the loudness sensation, a loudness scale was developed in which the unit of loudness is called the *sone*. The sone is defined as the loudness of a 1000-Hz tone at a sound level of 40 decibels (a loudness level of 40 phons).

For loudness levels of 40 phons or greater, the relationship between loudness S in sones and loudness level L_L in phons recommended by the International Standards Organization (ISO) is

$$S = 2^{(L_L - 40)/10}. \tag{6.7}$$

A graph of Eq. 6.7 is shown in Fig. 6.6. An equivalent expression for loudness S that avoids the use of L_L is

$$S = Cp^{0.6}, \tag{6.8}$$

where p is the sound pressure and C depends on the frequency.

Equations 6.7 and 6.8 are based on the work of S. S. Stevens, which indicated a doubling of loudness for a 10-dB increase in sound pressure level. Some investigators, however, have found a doubling of loudness for a 6-dB increase in sound pressure level (Warren 1970). This suggests the use of a formula in which loudness is proportional to sound pressure (Howes 1974):

$$S = K(p - p_0), \tag{6.9}$$

where p is sound pressure and p_0 is the pressure at some threshold level.

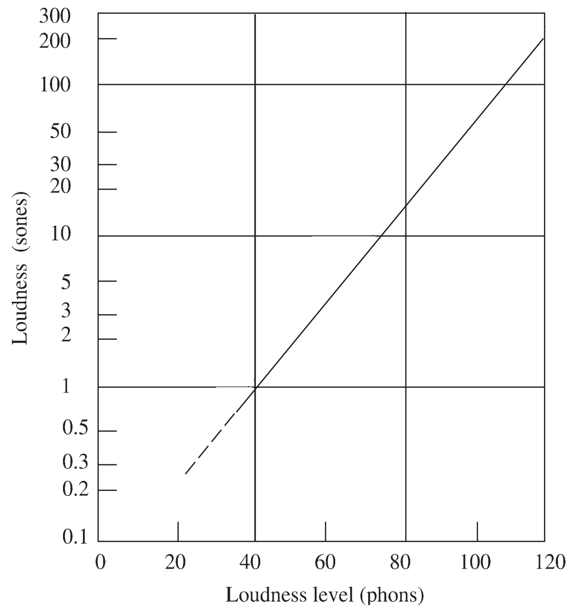


FIGURE 6.6 The relationship between the loudness (in sones) and the loudness level (in phons) from Eq. 6.7.

EXAMPLE 6.11 Find the loudness level and the loudness of a 500-Hz tone with $L_p =$ 70 dB.

Solution From Fig. 6.4, the loudness level is $L_L = 74$ phons.
The loudness is: $S = 2^{(74-40)/10} = 10.6$ sones.

6.7 ■ LOUDNESS OF COMPLEX TONES: CRITICAL BANDS

As pointed out in Table 5.1, loudness depends mainly on sound pressure, but it also varies with frequency, spectrum, and duration. We have already seen how loudness depends on frequency; now we will consider its dependence on the spectrum of the sound.

If we were to listen to two pure tones having the same sound pressure level but with increasing frequency separation, we would note that when the frequency separation exceeds the *critical bandwidth*, the total loudness begins to increase. Broadband sounds, such as those of jet aircraft, seem louder than pure tones or narrowband noise having the same sound pressure level. Figure 6.7 illustrates the dependence of loudness on bandwidth with fixed sound pressure level and center frequency. Note that loudness is not affected until the bandwidth exceeds the critical bandwidth, which is about 160 Hz for the 1-kHz center frequency shown.

One way to estimate the critical bandwidth is to increase the bandwidth of a noise signal while decreasing the amplitude in order to keep the power constant. When the bandwidth is greater than a critical band, the subjective loudness increases above that of a reference noise

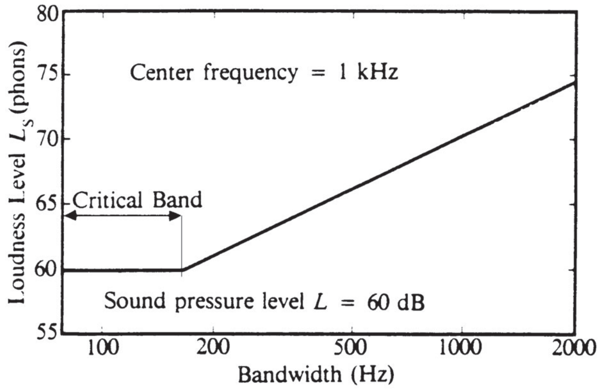


FIGURE 6.7 The effect of bandwidth on loudness.

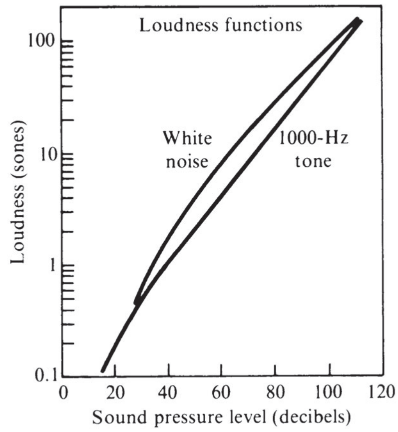


FIGURE 6.8 Loudness of white noise compared to that of a 1000-Hz tone at the same sound pressure level. Fifteen subjects judged the noise presented binaurally through headphones (Scharf and Fishken 1970).

signal because the stimulus now extends over more than one critical band (Demonstration 3 in Houtsma, Rossing, and Wagenaars 1988).

The perceived loudness of broadband (*white*) noise is compared to that of a 1000-Hz tones have the same SPL in Fig. 6.8. At a sound pressure level of 55 dB, the white noise is judged to be about twice as loud as the 1000-Hz tone, but at higher and lower levels the difference is substantially less (Scharf and Houtsma 1986).

The dependence of loudness on stimulus variables, such as sound pressure, frequency, spectrum, duration, etc., appears to be about the same whether the sound is presented to one ear (monaurally) or to both ears (binaurally). However, a sound presented to both ears is judged nearly twice as loud as the same sound presented to one ear only (Scharf and Houtsma 1986).

6.8 ■ LOUDNESS OF COMBINED SOUNDS

The loudness of combined sounds is a subject of considerable interest. How many violins must play together, for example, in order to double the loudness? Or, how does the loudness of traffic noise depend on the number of vehicles? We stated in Section 6.4 that the intensities (or mean-square pressures) from two or more uncorrelated sound sources add together to give a total intensity. The loudness is not necessarily additive, however. Accepted methods for combining loudness are given in the following box.

When two or more tones are mixed, the way in which their individual loudnesses combine depends on how close they are to each other in frequency. We can have three different situations:

1. If the frequencies of the tones are the same or fall within the critical bandwidth, the loudness is calculated from the total intensity $I = I_1 + I_2 + I_3 + \cdots$. If the intensities I_1, I_2, I_3, etc., are equal, the increase in sound level is as shown in Fig. 6.3. The loudness may then be determined from the combined sound level.
2. If the bandwidth exceeds the critical bandwidth, the resulting loudness is greater than that obtained from simple summation of intensities. As the bandwidth increases, the loudness approaches (but remains less than) a value that is the sum of the individual loudnesses:

$$S = S_1 + S_2 + S_3 + \cdots. \tag{6.10}$$

3. If the frequency difference is very large, the summation becomes complicated. Listeners tend to focus primarily on one component (e.g., the loudest or the one of highest pitch) and assign a total loudness nearly equal to the loudness of that component (Roederer 1975).

To determine the loudness of sones of a complex sound with many components, it is advisable to measure the sound level in each of the ten standard octave bands (or in thirty $\frac{1}{3}$-octave bands). Octave bands are frequency bands one octave wide (that is, the maximum frequency is twice the minimum frequency). Octave-band analyzers, available in many acoustic laboratories, usually have a filter that allows convenient measurement of the sound level in standard octave bands with center frequencies at 31, 63, 125, 250, 500, 1000, 2000, 4000, 8000, and 16,000 Hz. Once these levels have been measured, a suitable chart (see ISO Recommendations No. 532) can be used to find the loudness in sones.

Is this seemingly complicated procedure necessary? For precise determination of loudness, yes. For estimating loudness, no. A pretty fair estimate of loudness can be made by using an ordinary sound level meter to measure the A-weighted sound level. To estimate the number of sones, let 30 dBA correspond to 1.5 sones and double the number of sones for each 10-dBA increase, as shown in Table 6.2. This procedure works quite well at low to moderate levels, because the A-weighting is a reasonable approximation to the frequency response of the ear.

Because the previous paragraphs have dealt with numbers, formulas, and graphs, it is appropriate to make a few comments on how they apply to music, environmental noise,

TABLE 6.2 Chart for estimating loudness in sones of complex sounds from A-weighted sound levels

L_p(A)	30	40	50	55	60	65	70	75	80	85	90	dB
S	1.5	3	6	8	12	16	24	32	48	64	96	sones

and audiometric measurements. It should be emphasized that loudness is subjective, and its assessment varies from individual to individual. On the average, a sound of four sones sounds twice as loud as a sound of two sones, but some listeners may regard it as three times louder or one and a half times louder.

Interesting examples illustrating the importance of loudness phenomena in music appear throughout the literature. Roederer (1975) discusses the selection of combinations of organ stops. Benade (1976) describes how a saxophone was made to sound louder at the same sound pressure level by a change in timbre.

6.9 ■ MUSICAL DYNAMICS AND LOUDNESS

Variations in loudness add excitement to music. The range of sound level in musical performance, known as the *dynamic range*, may vary from a few decibels to 40 dB or more, depending on the music (loud peaks and pauses may cause the instantaneous level to exceed this range). The approximate range of sound level and frequency heard by the music listener is shown in Fig. 6.9.

Composers use dynamic symbols to indicate the appropriate loudness to the performer. The six standard levels are shown in Table 6.3.

Measurements of sound intensity of a number of instrumentalists have shown that seldom do musical performers actually play at as many as six distinguishable dynamic levels, however. In one study, the dynamic ranges of 11 professional bassoonists were found to vary from 6 to 17 decibels with an average of 10 dB (Lehman 1962). A 10-dB increase

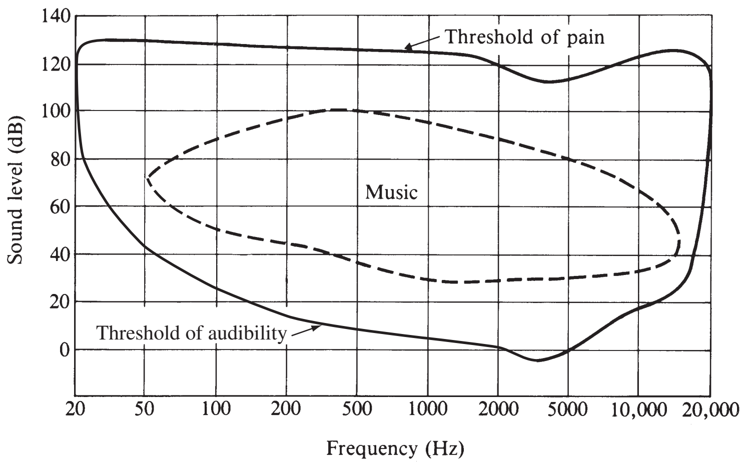


FIGURE 6.9 Approximate range of frequency and sound level of music compared to the total range of hearing.

TABLE 6.3 Standard levels of musical dynamics

Name	Symbol	Meaning
Fortissimo	*ff*	Very loud
Forte	*f*	Loud
Mezzo forte	*mf*	Moderately loud
Mezzo piano	*mp*	Moderately soft
Piano	*p*	Soft
Pianissimo	*pp*	Very soft

TABLE 6.4 Dynamic ranges of musical instruments

Instrument	Average dynamic range (dB) (Clark and Luce 1965)	Maximum dynamic range (dB) (Patterson 1974)
Violin	14	40
Viola	16	
Cello	14	
String bass	14	30
Recorder		10
Flute	7	30
Oboe	7	
English horn	5	
Clarinet	8	45
Bassoon	10	40
Trumpet	9	
Trombone	17	38
French horn	18	
Tuba	13	

in sound level, you will recall from Section 6.4, is usually said to double the loudness (expressed in sones). Most listeners would have considerable difficulty identifying six different levels within a dynamic range of 10 dB. Dynamic ranges of several instruments are given in Table 6.4.

The dynamic ranges in Table 6.4 are for single notes played loudly and softly. Several instruments have much more sound power near the top of their playing range than near the bottom. (Fortissimo on a French horn, for example, is found to be nearly 30 dB greater at C_5 than at C_2, although the difference between *ff* and *pp* on any note of the scale may be 20 dB or less.)

Measurement of the dynamic ranges of various musical instruments and players is an instructive and relatively easy experiment for the reader to perform. The dynamic ranges of most players we have measured fall close to those reported by Clark and Luce (1965).

6.10 ■ MASKING

When the ear is exposed to two or more different tones, it is a common experience that one may mask the others. *Masking* is probably best explained as an upward shift in the hearing threshold of the weaker tone by the louder tone and depends on the frequencies of the two tones. Pure tones, complex sounds, narrow and broad bands of noise all show differences in their ability to mask other sounds. Masking of one sound can even be caused by another sound that occurs a split second after the masked sound.

Some interesting conclusions can be made from the many masking experiments that have been performed:

1. Pure tones close together in frequency mask each other more than tones widely separated in frequency.
2. A pure tone masks tones of higher frequency more effectively than tones of lower frequency.
3. The greater the intensity of the masking tone, the broader the range of frequencies it can mask.
4. If the two tones are widely separated in frequency, little or no masking occurs.
5. Masking by a narrow band of noise shows many of the same features as masking by a pure tone; again, tones of higher frequency are masked more effectively than tones of lower frequency than the masking noise.
6. Masking of tones by broadband ("white") noise shows an approximately linear relationship between masking and noise level (that is, increasing the noise level 10 dB raises the hearing threshold by the same amount). Broadband noise masks tones of all frequencies.
7. *Forward masking* refers to the masking of a tone by a sound that ends a short time (up to about 20 or 30 ms) before the tone begins. Forward masking suggests that recently stimulated cells are not as sensitive as fully rested cells.
8. *Backward masking* refers to the masking of a tone by a sound that begins a few milliseconds later. A tone can be masked by noise that begins up to ten milliseconds later, although the amount of masking decreases as the time interval increases (Elliott 1962). Backward masking apparently occurs at higher centers of processing where the later-occurring stimulus of greater intensity overtakes and interferes with the weaker stimulus.

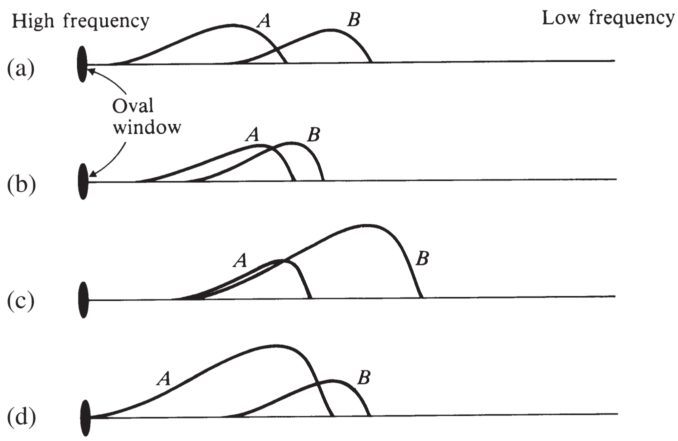


FIGURE 6.10 Simplified response of the basilar membrane for two pure tones *A* and *B*. (a) The excitations barely overlap; little masking occurs. (b) There is an appreciable overlap; tone *B* masks tone *A* and somewhat more than the reverse. (c) The more intense tone *B* almost completely masks the higher-frequency tone *A*. (d) The more intense tone *A* does not completely mask the lower-frequency tone *B*.

9. Masking of a tone in one ear can be caused by noise in the other ear, under certain conditions; this is called *central masking*.

Some of the conclusions about masking just stated can be understood by considering the way in which pure tones excite the basilar membrane (see Fig. 5.6). High-frequency tones excite the basilar membrane near the oval window, whereas low-frequency tones create their greatest amplitude at the far end. The excitation due to a pure tone is asymmetrical, however, having a tail that extends toward the high-frequency end as shown in Fig. 6.10. Thus it is easier to mask a tone of higher frequency than one of lower frequency. As the intensity of the masking tone increases, a greater part of its tail has amplitude sufficient to mask tones of higher frequency.

6.11 ■ LOUDNESS REDUCTION BY MASKING

Sounds are seldom heard in isolation. The presence of other sounds not only raises the threshold for hearing a given sound but generally reduces its loudness as well (this is sometimes called *partial masking*).

Figure 6.11 shows how white noise reduces the apparent loudness of a 1000-Hz tone. Compared to the tone in quiet (see Fig. 6.8), the loudness functions in white noise are steeper. Rising form an elevated threshold, the partially masked tone eventually comes to its full unmasked loudness when the noise level is less than 80 dB. In more intense noise, the loudness does not reach its full unmasked value, but the function approaches the same slope as the function without masking noise (Scharf and Houtsma 1986).

6.12 ■ LOUDNESS AND DURATION: IMPULSIVE SOUNDS AND ADAPTATION

How does the loudness of an impulsive sound compare to the loudness of a steady sound at the same sound level? Numerous experiments have pretty well established that the ear averages sound energy over about 0.2 s (200 ms), so loudness grows with duration up to this

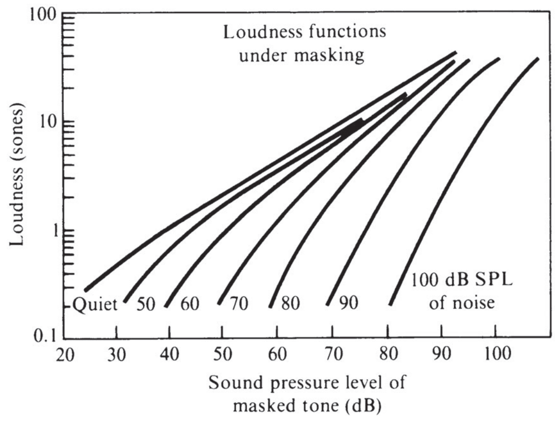


FIGURE 6.11 Loudness functions for a 1000-Hz tone partially masked by white noise at various sound pressure levels. Subjects adjusted the level of the tone in quiet so that it sounded as loud as the tone with noise. (After Scharf 1978.)

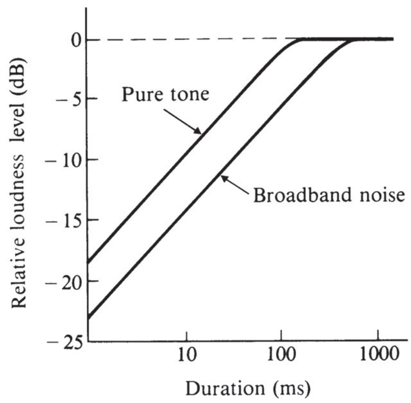


FIGURE 6.12 Variation of loudness level with duration. (After Zwislocki 1969.)

value. Stated another way, loudness level increases by 10 dB when the duration is increased by a factor of 10. The loudness level of broadband noise seems to depend somewhat more strongly on stimulus duration than the loudness level of pure tones, however. Figure 6.12 shows the approximate way in which loudness level changes with duration.

The ear, being a very sensitive receiver of sounds, needs some protection to avoid injury by very loud sounds. Up to 20 dB of effective protection is provided by muscles attached to the eardrum and the ossicles of the middle ear. When the ear is exposed to sounds in excess of 85 dB or so, these muscles tighten the ossicular chain and pull the stapes (stirrup-shaped bone) away from the oval window of the cochlea. This action is termed the *acoustic reflex*.

Unfortunately the reflex does not begin until 30 or 40 ms after the sound overload occurs, and full protection does not occur for another 150 ms or so. In the case of a loud impulsive sound (such as an explosion or gunshot), this is too late to prevent injury to the ear. In fact a tone of 100 dB or so preceding the loud impulse has been proposed as a way of triggering the acoustic reflex to protect the ear (Ward 1962). It is interesting to speculate what type of protective mechanism, analogous to eyelids, might have developed in the auditory system had the loud sounds of the modern world existed for millions of years (earlids, perhaps?).

Like most other sensations, loudness might be expected to decrease with prolonged stimulation. Such a decrease is called *adaptation*. Under most listening conditions, however, loudness adaptation appears to be very small. A steady 1000-Hz tone at 50 dB causes little adaptation, although the loudness of a tone that alternates between 40 and 60 dB appears to decrease in loudness over the first two or three minutes, as do tones within about 30 dB of threshold (Scharf and Houtsma 1986).

Exposure to a loud sound affects our ability to hear another sound at a later time. This is called *fatigue* and may result in both a temporary loudness shift (TLS) and a temporary threshold shift (TTS). TLS and TTS appear to be greatest at a frequency a half octave higher than that of the fatiguing sound. Noise-induced TTS is discussed in Chapter 31.

6.13 ■ SUMMARY

Each of the quantities sound pressure, sound power, and sound intensity has an appropriate decibel level that expresses the ratio of these quantities to appropriate reference levels.

Sound pressure can be measured directly by a sound level meter, which may offer one to three different frequency weightings. Loudness level (in phons) expresses the sound pressure level of an equally loud 1000-Hz tone, whereas the loudness (in sones) expresses a subjective rating of loudness. Expressing the loudness of complex tones is fairly subtle, involving critical bandwidth, masking one tone by another, etc. The loudness of impulsive sounds increases with their duration up to about 0.2 s. The dynamic range of music covers about 40 dB, although individual instruments have dynamic ranges considerably less than this. Composers use six standard levels to indicate loudness. The ear is partially protected from loud sounds by the acoustic reflex.

REFERENCES AND SUGGESTED READINGS

Benade, A. H. (1976). *Fundamentals of Musical Acoustics*. New York: Oxford.

Clark, M., and D. Luce (1965). "Intensities of Orchestral Instrument Scales Played at Prescribed Dynamic Markings," *J. Audio Eng. Soc*. **13**: 151.

Egan, J. P., and H. W. Hake (1950). "On the Masking Pattern of a Simple Auditory Stimulus," *J. Acoust. Soc. Am*. **22**: 622.

Elliott, L. L. (1962). "Backward and Forward Masking of Probe Tones of Different Frequencies," *J. Acoust. Soc. Am*. **34**: 1116.

Fletcher, H., and W. A. Munson (1933). "Loudness, Definition, Measurement and Calculation," *J. Acoust. Soc. Am*. **6**: 59.

Hawkins, J. E., Jr., and S. S. Stevens (1950). "Masking of Pure Tones and Speech by Noise," *J. Acoust. Soc. Am*. **22**: 6.

Houtsma, A. J. M., T. D. Rossing, and W. M. Wagenaars (1988). *Auditory Demonstrations* (Philips Compact Disc #1126-061 and text). Woodbury, N.Y.: Acoustical Society of America.

Howes, W. L. (1974). "Loudness Function Derived from Data on Electrical Discharge Rates in Auditory Nerve Fibers," *Acustica* **30**: 247.

International Standards Organization, Standards R131, 26, R357 and R454 (International Organization for Standardization, 1 Rue de Narembé, Geneva, Switzerland).

Lehman, P. R. (1962). "The Harmonic Structure of the Tone of the Bassoon," Ph.D. thesis, University of Michigan, Ann Arbor.

Littler, T. S. (1965). *The Physics of the Ear*. New York: Macmillan.

Patterson, B. (1974). "Musical Dynamics," *Sci. Am.* **231**(5): 78.

Plomp, R. (1976). *Aspects of Tone Sensation*. London: Academic.

Roederer, J. G. (1975). *Introduction to the Physics and Psychophysics of Music*, 2nd ed. New York: Springer-Verlag.

Scharf, B. (1978). "Loudness," in *Handbook of Perception*, Vol. 4. Ed., E. C. Carterette and M. P. Friedman. New York: Academic Press.

Scharf, B., and D. Fishken (1970). "Binaural Summation of Loudness," *J. Experimental Psych*. **86**: 374–79.

Scharf, B. and A. J. M. Houtsma (1986). "Audition II. Loudness, Pitch, Localization, Aural Distortion, Pathology," in *Handbook of Perception and Human Performance*. Ed., K. R. Boff, L. Kaufman, and J. P. Thomas. New York: J. Wiley.

Stevens, S. S., and G. Stevens (1975). *Psychophysics Introduction to Its Perceptual, Neural and Social Prospects*. New York: Wiley.

Ward, W. D. (1962). "Studies on the Aural Reflex III. Reflex Latency as Inferred from Reduction of Temporary Threshold Shift from Impulses," *J. Acoust. Soc. Am*. **34**: 1132.

Warren, R. M. (1970). "Elimination of Biases in Loudness Judgements for Tones," *J. Acoust. Soc. Am*. **48**: 1397.

Zwislocki, J. J. (1969). "Temporal Summation of Loudness: An Analysis," *J. Acoust. Soc. Am*. **46**: 431.

GLOSSARY

acoustic reflex Muscular action that reduces the sensitivity of the ear when a loud sound occurs.

auditory fatigue Change in loudness of a sound that follows a loud sound.

critical bandwidth The frequency bandwidth beyond which subjective loudness increases with bandwidth (see also definition in chapter 5).

decibel A dimensionless unit used to compare the ratio of two quantities (such as sound pressure, power, or intensity), or to express the ratio of one such quantity to an appropriate reference.

intensity Power per unit area; rate of energy flow.

intensity level $L_I = 10 \log I/I_0$, where I is intensity and $I_0 = 10^{-12}$ W/m^2 (abbreviated *SIL* or L_I).

loudness Subjective assessment of the "strength" of a sound, which depends on its pressure, frequency, and timbre; loudness may be expressed in sones.

loudness level Sound pressure of a 1000-Hz tone that sounds equally loud when compared to the tone in question; loudness level is expressed in phons.

masking The obscuring of one sound by another.

phon A dimensionless unit used to measure loudness level; for a tone of 1000 Hz, the loudness level in phons equals the sound pressure in decibels.

sone A unit used to express subjective loudness; doubling the number of sones should describe a sound twice as loud.

sound power level $L_W = 10 \log W/W_0$, where W is sound power and $W_0 = 10^{-12}$ W (abbreviated *PWL* or L_W).

sound pressure level $L_p = 20 \log p/p_0$, where p is sound pressure and $p_0 = 2 \times 10^{-5}$ N/m^2 (or 20 micropascals) (abbreviated *SPL* or L_p).

white noise Noise whose amplitude is constant throughout the audible frequency range.

REVIEW QUESTIONS

1. In what units is sound intensity measured?
2. What reference level is used to measure sound intensity level?
3. What is meant by a free field?
4. How large is the "just noticeable difference" in sound level?
5. How much does the sound level decrease in a free field when the distance from the source is doubled?
6. In air, how does ρ change as the temperature increases?
7. In air, how does c change as the temperature increases?
8. In air, how does ρc change with temperature?
9. What is the approximate sound level in normal conversation?
10. If each of two sound sources alone produces a sound level of 55 dB at a certain point, what will the level most likely be at that point if both sources are active?
11. In what units is loudness level expressed?
12. In what units is loudness expressed?
13. What generally happens to loudness as the bandwidth of a noise source is increased while the sound level stays constant?
14. By approximately how many decibels must the A-weighted sound level increase in order to double the loudness of a complex tone?
15. What is the average dynamic range of a single note played on a musical instrument?
16. What is backward masking?
17. Is it easier for a tone of lower frequency to mask a tone of a higher frequency, or vice versa?
18. Does a given tone generally sound louder or less loud against a background noise as compared to the same tone in a quiet setting?
19. How does loudness depend upon duration at constant sound level?

QUESTIONS FOR THOUGHT AND DISCUSSION

1. Which will sound louder, a pure tone of $L_p = 40$ dB, $f = 2000$ Hz, or a pure tone of $L_p = 65$ dB, $f = 50$ Hz?
2. If two identical loudspeakers are driven at the same power level by an amplifier, how will the sound levels

due to each combine? Does it make a difference whether the program source is stereophonic or monophonic?

3. How is it possible for one sound to mask a sound that has already occurred (backward masking)? Speculate what might happen in the human nervous system when such a phenomenon occurs.

4. How low long must a burst of broadband (white) noise be in order to be half as loud as a continuous noise of the same type?

5. Why do community noise laws generally specify maximum L_p(A) rather than L_p(C)?

EXERCISES

1. What sound pressure level is required to produce minimum audible field at 50, 100, 500, 1000, 5000, and 10,000 Hz?
2. What sound pressure level of 100-Hz tone is necessary to match the loudness of a 3000-Hz tone with $L_p = 30$ dB? What is the loudness level (in phons) of each of these tones?
3. With one violin playing, the sound level at a certain place is measured as 50 dB. If four violins play equally loudly, what will the sound level most likely be at this same place?
4. If two sounds differ in level by 46 dB, what is the ratio of their sound pressures? their intensities?
5. A loudspeaker is supplied with 5 W of electrical power, and it has an efficiency of 10% in converting this to sound power. What is its sound power level? If we assume that the sound radiates equally in all directions, what is the sound pressure level at a distance of 1 m? at a distance of 4 m?
6. A 60-Hz tone has a sound pressure level of 60 dB measured with C-weighting on a sound level meter. What level would be measured with A-weighting?
7. Find the sound pressure and the intensity of a sound with $L_p = 50$ dB.
8. What is the decibel gain when the power gain is 30? when it is 50?
9. According to Fig. 6.6, what is the loudness level that produces a loudness of 10 sones? 100 sones?

EXPERIMENTS FOR HOME, LABORATORY, AND CLASSROOM DEMONSTRATION

Home and Classroom Demonstration

1. *The decibel scale* Demonstration 4 on the *Auditory Demonstrations* CD (Houtsma, Rossing, and Wagenaars 1988). Broadband is reduced in steps of 6 dB, 3 dB, and 1 dB. This is followed by free-field speech, recorded at distances of 0.25, 0.5, 1, and 2 m from a microphone.

2. *Frequency response of the ear* Demonstration 6 on the *Auditory Demonstrations* CD. Tones having frequencies of 125, 250, 500, 1000, 2000, 4000, and 8000 Hz are decreased in 10 steps of −5 dB each in order to determine thresholds of audibility at each frequency.

3. *Loudness scaling* Demonstration 7 on the *Auditory Demonstrations* CD. Listeners are asked to rate the loudness of 20 test tones in comparison to a reference tone. These ratings are plotted against the sound level of each test tone to establish an average loudness scale. If done as a class demonstration, better statistics are obtained by combining all the responses on a single graph.

4. *Critical bands by loudness comparison* Demonstration 3 on the *Auditory Demonstrations* CD. The bandwidth of a noise burst is increased while its amplitude is decreased to keep the power constant. When the bandwidth is greater than a critical band, the subjective loudness increases above that of a reference noise burst, because the stimulus now extends over more than one critical band.

5. *Critical bands by masking* Demonstration 2 on the *Auditory Demonstrations* CD. A 2000-Hz tone is masked by spectrally flat (white) noise of different bandwidths. You expect to hear more steps in the 2000-Hz tone staircase when the noise bandwidth is reduced below the critical bandwidth.

6. *Temporal integration* Demonstration 8 on the *Auditory Demonstrations* CD. Bursts of broadband noise having durations of 1000, 300, 100, 30, 10, 3, and 1 ms are presented at eight decreasing levels. A graph of a number of steps heard as a function of duration should give an indication of integration time (see Fig. 6.12).

7. *Asymmetry of masking* Demonstration 9 on the *Auditory Demonstrations* CD. This demonstration compares the mask-

ing of a 2000-Hz tone by a 1200-Hz tone with the masking of a 1200-Hz tone by a 2000-Hz tone.

8. *Backward and forward masking* Demonstration 10 on the *Auditory Demonstrations* CD. This demonstration of masking by nonsimultaneous tones compares forward masking (masking tone before the test tone) with backward masking (masking tone after the test tone). Forward masking is more robust, of course, but the amazing thing is that backward masking occurs at all!

9. *Familiarity with sound levels* Use an inexpensive sound-level meter to determine the sound level of as many different sounds as possible (music, broadband noise, traffic noise, conversation, etc.).

10. *Asymmetry of masking with two oscillators* Two audio generators can be used to show that a tone of lower frequency masks a tone of higher frequency more effectively than the other way around.

Laboratory Experiments

Sound level (Experiment 10 in *Acoustics Laboratory Experiments*)

Loudness level and audiometry (Experiment 12 in *Acoustics Laboratory Experiments*)

CHAPTER 7

Pitch and Timbre

Pitch has been defined as that characteristic of a sound that makes it sound high or low or that determines its position on a scale. For a pure tone, the pitch is determined mainly by the frequency, although the pitch of a pure tone may also change with sound level. The pitch of complex sounds also depends on the spectrum (timbre) of the sound and its duration. In fact, the pitch of complex sounds has been one of the most interesting objects of study in psychoacoustics for several years.

In this chapter you should learn:

- About pitch scales and pitch discrimination;
- How pitch depends on frequency, sound level, duration, timbre, and competing sounds;
- About theories of pitch;
- About pitch of complex tones and virtual pitch;
- About absolute pitch;
- About timbre;
- About spectral analysis of complex tones;
- About analytical and synthetic listening.

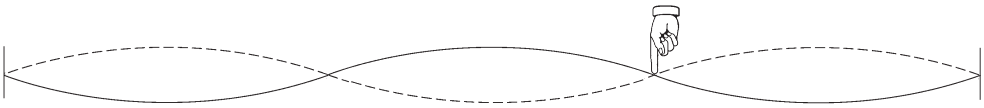
7.1 ■ PITCH SCALES

The American National Standards Institute (1960) defines *pitch* as "that attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from low to high." This definition probably leads most of us to think of a musical scale. Are there other pitch scales besides musical scales? Is there a subjective scale of pitch similar to the sone scale of loudness discussed in the previous chapter?

Pitch is a subjective sensation. Two persons hearing the same sound may assign it different may assign it different positions on a pitch scale. In fact, some listeners may assign a different pitch to a sound depending upon whether it is presented to the right or left ear (this is called *binaural diplacusis*).

The basic unit in most musical scales is the *octave*. Notes judged an octave apart have frequencies nearly (but not always exactly, as we will see) in the ratio 2:1. As early as the sixth century B.C., according to legend, Pythagoras of Athens noted that if one segment of a string is half as long as the other, the pitches produced by plucking the two segments have a special similarity. Errors of one octave are frequently made in judging the pitch of a musical note. (If you don't believe this, ask a musician to whistle a note and then to name the octave in which the note lies.)

From Chapter 7 of *The Science of Sound*, Third Edition. Thomas D. Rossing, Richard F. Moore, Paul A. Wheeler.



Pythagoras discovers the octave (ca 600 B.C.).

In music, the octave is subdivided in different ways, as we shall see in Chapter 9. Western music normally divides the octave into 12 intervals called *semitones;* these are given note names (A through G with sharps and flats) and designated on musical staves.

Psychophysical Pitch Scales

Various attempts have been made to establish a psychophysical pitch scale. If an average listener were allowed to listen to a tone of 4000 Hz followed by a tone of low frequency and then asked to tune an oscillator to a pitch halfway between, a likely choice would be something around 1000 Hz. On a scale of pitch, then, 1000 Hz is judged as halfway between 0 and 4000 Hz. The unit used for subjective pitch is the *mel;* the scale is arranged so that doubling the number of mels doubles the pitch. From 0 to 2400 mels spans the frequency range 0 to 16 kHz; the correspondence between mels and hertz is shown in Fig. 7.1.

Another psychophysical scale is based on critical bands of hearing. A critical bandwidth is designated one *bark*. Interestingly enough, it turns out that one bark is very nearly equal to 100 mels, so the two scales are actually quite similar.

A numerical scale of pitch (in mels) is not nearly so useful as a numerical scale of loudness (in sones), however. Pitch is usually related to a musical scale where the octave, rather than the critical bandwidth, is the "natural" pitch interval.

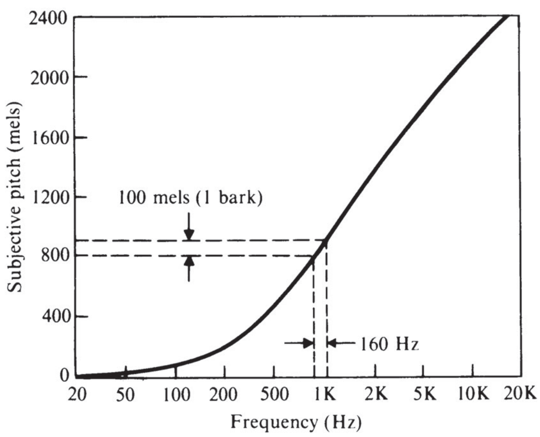


FIGURE 7.1 Pitch scale versus frequency scale. On the pitch scale, 100 mels is close to the width of the critical band, which is 160 Hz at a center frequency of 1000 Hz (dashed lines). (From Wightman and Green 1974. Reprinted by permission of the Am. Inst. of Physics.)

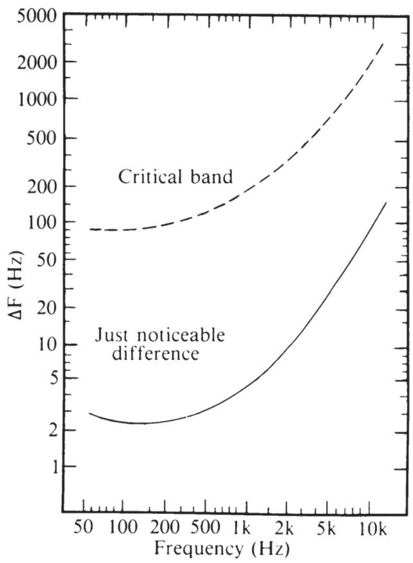


FIGURE 7.2 Just-noticeable difference (jnd) in frequency determined by modulating the frequency of a tone at 4 Hz. Note that the jnd at each frequency is nearly a constant percentage of the critical bandwidth. (From Zwicker, Flottorp, and Stevens, 1957.)

7.2 ■ PITCH DISCRIMINATION

The ability to distinguish between two nearly equal stimuli is often characterized, in psychophysical studies, by a *difference limen* or *just-noticeable difference* (jnd). Two stimuli will be judged the same if they differ by less than the jnd.

The jnd for pitch has been found to depend on the frequency, the sound level, the duration of the tone, and the suddenness of the frequency change. It also depends on the musical training of the listener and to some extent on the method of measurement. Figure 7.2 shows the average (of four subjects) for pure tones at a sound level of 80 dB. From 1000 to 4000 Hz, the jnd is approximately 0.5 percent of the pure tone frequency, which is about one-twelfth of a semitone. Sometimes the term *frequency resolution* is used to denote the jnd divided by the frequency ($\Delta f/f$).

By comparing the upper and lower curves in Fig. 7.2, we can see that critical bandwidth is roughly equal to 30 difference limens or jnd's at all center frequencies. This remarkable result suggests that the same mechanism in the ear is responsible for critical bands and for pitch discrimination. It is quite likely related to regions of excitation along the basilar membrane (see Section 5.4).

It is interesting to compare pitch discrimination to color discrimination. Whereas the visible spectrum extends over one octave (violet light has roughly twice the frequency of red) and includes 128 just noticeable differences (distinguishable hues or colors), the auditory spectrum covers about 10 octaves with 5000 jnds.

7.3 ■ PITCH OF PURE TONES

We have already noted that pitch depends mainly on frequency; pitch scaling with frequency was discussed in Section 7.1. We now consider the pitch dependence of pure tones on other physical quantities such as sound pressure, duration, envelope, and the presence of other sounds.

Pitch and Sound Level

Early experiments on pitch versus sound level reported substantially larger pitch dependence on sound level than more recent studies do. Early work by Stevens (1935) indicated shifts in pitch as large as two semitones (apparent frequency changes of 12%) as the sound level of pure tones increased from 40 to 90 dB. Tones of low frequency were found to fall in pitch with increasing intensity; tones of high frequency rise in pitch with increasing intensity, and tones of middle frequency (1–2 kHz) show little change. (This has sometimes been referred to as *Stevens' rule*.) Stevens found the maximum downward shift with sound level at 150 Hz and the largest upward shift with sound level around 8000 Hz.

It now appears that the effect is small, even for pure tones, and varies from observer to observer; in one experiment, for example, five musically trained subjects hear pitch lowerings that varied from 0 to 75 cents (75 cents = $\frac{3}{4}$ semitone) when a 250-Hz tone increased from 40 to 90 dB (Ward 1970). Whereas pitch changes for individuals tend to follow Stevens' rule, then, averaging over a group of observers makes the changes less significant. Figure 7.3 shows the pitch shifts of pure tones with frequencies from 200 Hz to 6000 Hz averaged over 15 subjects.

The small pitch changes shown in Fig. 7.3, as well as the larger changes described by early investigators, are for pure tones. Less is known about the effect for complex tones. Studies with musical instruments have generally shown very small pitch change with intensity (around 17 cents for an increase from 65 to 95 dB, for example). Whether the pitch of a complex tone rises or falls with increasing intensity appears to depend on which partials (above or below 1000 Hz) are predominant (Terhardt 1979).

In contrast with the results in Fig. 7.3, however, increasing the amplitude of short tone bursts causes a downward shift in pitch over a wide range of frequency. Similar results are found in experiments using 12-ms bursts (Doughty and Garner 1948) and 40-ms bursts (Rossing and Houtsma 1986).

A phenomenon of pitch change that has been observed during reverberant decay may be due in part to pitch change with sound level, although other effects appear to contribute as well. This phenomenon is quite apparent when one is listening to pipe organ music in

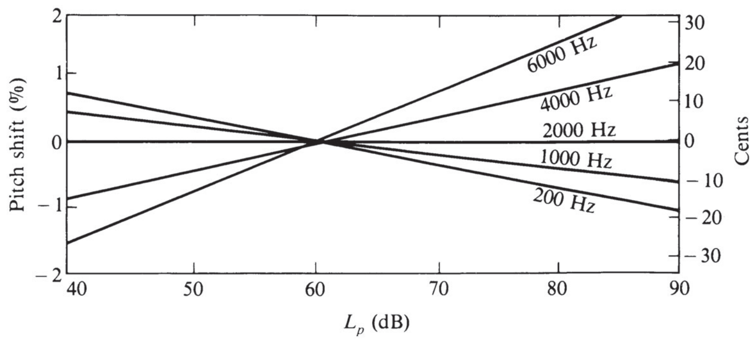


FIGURE 7.3
Pitch shift of pure tones as a function of sound pressure level. Shifts are shown in both percent and cents (100 cents = 1 semitone). The curves are based on data from 15 subjects. (After Terhardt 1979.)

churches with substantial reverberation; the pitch often appears to rise as the sound level diminishes after a loud chord ends (Parkin 1974).

It is fortunate for performing musicians and listeners alike that the change in pitch with sound level for complex tones is much less than was reported from early experiments with pure tones. Musical performance would be very difficult if substantial changes of pitch occurred during changes in dynamic level.

Pitch and Duration

How long must a tone be heard in order to have an identifiable pitch? Although early experiments by Savart (1840) indicated that a sense of pitch develops after only two cycles, most subsequent investigations indicated that a longer duration is required (see Fig. 7.4). Very brief tones are described as *clicks*, but as the tones lengthen, the clicks take on a weak sensation of pitch, increasing in strength upon further lengthening.

The transition from click to tone depends on sound level; if the tone does not begin abruptly, but rather with a soft onset, tone-recognition times as short as 3 ms are possible, which is shorter than the attack time of most musical instruments (Winckel 1967).

It has been suggested that the dependence of pitch on duration follows a sort of "acoustical uncertainty principle" $\Delta f \, \Delta t = K$, where Δf is the uncertainty in frequency and Δt is the duration of a tone burst. Under optimum conditions, K can be less than 0.1 (Majerník and Kalužný 1979). When the tone duration falls below 25 ms, the pitch may appear to change, although slightly different results are reported by various investigators (Rossing and Houtsma 1986).

The ear has an especially high sensitivity for detecting frequency changes of pure tones. The jnd for frequency change in pure tones Δt is less than for noise, provided that the amplitude of the pure tone remains constant. Even with a band of noise as narrow as 10 Hz at a center frequency of 1500 Hz (which sounds like a pure tone of varying amplitude), Δt will be six times greater than for a pure tone of 1500 Hz (Zwicker 1962).

Pitch and Envelope

The perceived pitch of a short exponentially decaying sinusoidal tone is found to be consistently higher than a simply gated sine tone with the same frequency and energy (Hartmann 1978). Rossing and Houtsma (1986) found the same effect for tones with rising exponen-

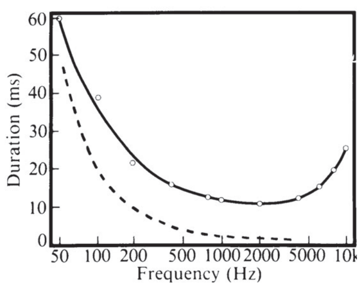


FIGURE 7.4 The duration required for a given tone to produce a definite pitch. The solid line is from the data of Bürck, Kotowski, and Lichte (1935); the dashed line is the duration of two cycles (Savart 1840).

tial envelopes, and found that the pitch shift depends on the sound pressure level as well as the rate of rise or fall of the tone envelope.

The reason for the envelope dependence of pitch is not clear, but it appears to be related to the pitch shift with intensity discussed earlier in this section. It certainly is an effect that musicians should take into account when dealing with the pitch of percussion instruments.

Effect of Interfering Sounds

Sounds are seldom heard in isolation. Another factor that influences the pitch of pure tone is the presence of other interfering sounds. Experiments both with a second interfering tone and with interfering noise can be summarized as follows:

1. If the interfering tone has a frequency below that of the test tone, an *upward* shift always occurs.
2. If the interfering tone frequency is above that of the test tone, a *downward* shift is observed at low frequencies.
3. Interfering noise always causes an upward pitch shift if it has a lower frequency than the test tone (but if it has a higher frequency, the shift can occur in either direction).
4. The pitch shift increases with the amount by which the interfering tone or noise amplitude exceeds that of the test tone (Terhardt and Fastl 1971).

7.4 ■ PITCH OF COMPLEX TONES: VIRTUAL PITCH

When the ear is presented with a tone composed of exact harmonics, it is easy to predict what pitch will be heard. It is simply the lowest common factor in these frequencies, which is the fundamental. The ear identifies the pitch of the fundamental, even if the fundamental is very weak or missing altogether. For example, if the ear hears a tone having partials with frequencies of 600, 800, 1000, and 1200 Hz, the pitch will nearly always be identified as that of a 200-Hz tone, the *missing fundamental*. This is an example of what is called *virtual pitch*, since the pitch does not correspond to any partial in the complex tone. The ability of the ear to determine a virtual pitch makes it possible for the undersized loudspeaker of a portable radio to produce bass tones and also forms the basis for certain mixture stops on a pipe organ.

If a strong fundamental is not essential for perceiving the pitch of a musical tone, the question arises as to which harmonics are most important. Experiments have shown that for a complex tone with a fundamental frequency up to about 200 Hz, the pitch is mainly determined by the fourth and fifth harmonics. As the fundamental frequency increases, the number of the dominant harmonics decreases, reaching the fundamental itself for $f_0 =$ 2500 Hz and above (Plomp 1967). Consider, for example, a tone A_3 with a frequency $f_0 = 200$ Hz; if the fourth and fifth harmonics were raised in frequency, the pitch of the tone would most likely appear to rise even though the fundamental remained at 220 Hz.

When the partials of the complex tone are not harmonic, however, the determination of virtual pitch is more subtle. According to current theories of pitch, the ear picks out a series of nearly harmonic partials somewhere near the center of the audible range, and determines the pitch to be the largest near-common factor in the series (Goldstein 1973).

Several demonstrations of virtual pitch are presented by Houtsma, Rossing, and Wagenaars (1987).

Musical examples of the ability of the auditory system to formulate a virtual pitch from near harmonics in a complex tone are the sounds of bells and chimes. In each case the pitch of the *strike note* is determined mainly by three partials that have frequencies almost in the ratio 2 : 3 : 4, as we shall see in Chapter 13. In the case of the bell, there is usually another partial with a frequency near that of the strike note, which reinforces it. In the case of the chime, however, there is none: The pitch is purely subjective.

The following two sections discuss theories of pitch perception and their historical development, including some important experiments that led to our present understanding of pitch of complex tones (which are so important in music and speech). You may read them carefully, skim them, or skip directly to Section 7.7.

7.5 ■ SEEBECK'S SIREN AND OHM'S LAW: A HISTORICAL NOTE

About the middle of the eighteenth century, A. Seebeck performed a series of experiments on pitch perception that produced some significant, if surprising, results. As a source of sound, Seebeck used a siren consisting of a rotating disc with periodically spaced holes that created puffs of compressed air at regular intervals, as shown in Fig. 7.5(a). Seebeck noted that this siren produced sound with a very strong pitch corresponding to the time between puffs of air. Doubling the number of holes, as shown in Fig. 7.5(b), raised the pitch exactly an octave, as expected.

However, using a disk with unequal spacing of holes, as shown in Fig. 7.5(c), produced an unexpected result: the pitch now heard matched that heard with the siren in (a). This may be understood by studying the corresponding waveforms (amplitude versus time) and spectra (amplitude versus frequency) shown in Fig. 7.5. In (a) the spectrum has components at the fundamental frequency $1/T$ (where T is the period) and its harmonics ($2/T$, $3/T$, etc.). In (b) the fundamental frequency is twice as great ($2/T$), and the harmonics occur at

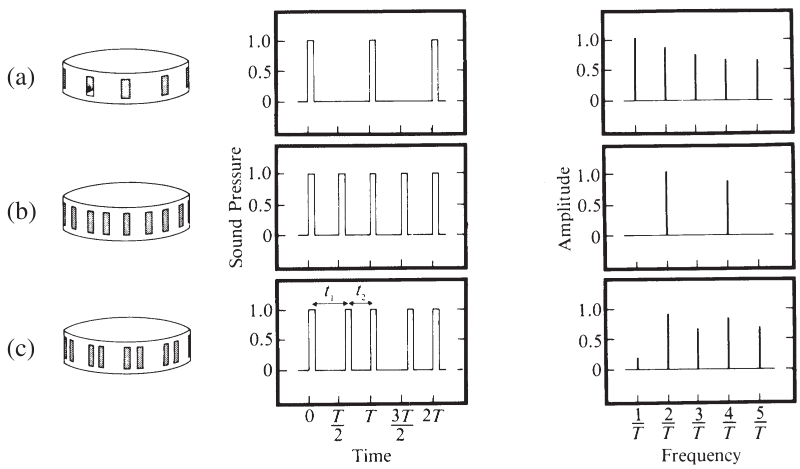


FIGURE 7.5 Three different sirens used by Seebeck along with the waveforms and spectra of sound they generate. (After Wightman and Green 1974.)

$4/T$, $6/T$, etc.; thus, the pitch is an octave higher. In (c) where the spacing between puffs is alternately t_1 and t_2, the period of repetition is $T = t_1 + t_2$; thus, harmonics occur at the same frequencies as in (a), although the fundamental is very much weaker. The pitch therefore matches that of case (a), although the quality or timbre of the sound is quite different.

It is quite easy and instructive to repeat Seebeck's experiment using an electronic pulse generator to generate the wave forms shown in Fig. 7.5. What one hears in the case of waveform (c) is two tones an octave apart, the lower tone becoming softer as $t_2 \rightarrow t_1$, disappearing rather abruptly when $t_2 = t_1$, whereas the upper tone remains relatively constant in loudness.

About the time Seebeck was performing his experiment, G. S. Ohm adapted Fourier's theorem on spectrum analysis (see Section 7.10) to acoustics and formulated what is often known as *Ohm's acoustical law* (or *Ohm's second law*, his first law having dealt with electric circuits). Ohm believed that a pitch corresponding to a certain frequency could be heard only if the acoustic wave contained power at that frequency. Thus, he criticized Seebeck's interpretation of his siren experiment that periodicity, rather than fundamental frequency, determines pitch. In the case of the waveform shown in Fig. 7.5(c), however, the sensation of pitch is far too strong to be explained on the basis of the weak component or partial at the fundamental frequency, and thus Ohm's law is contradicted. Ohm finally suggested that the phenomenon was due to an acoustical illusion (Wightman and Green 1974).

In his monumental work *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, H. von Helmholtz (1877) supported Ohm's position, adding the important idea of distortion products generated in the ear. For pure tones, these distortion products would be primarily harmonics of the pure tone (harmonic distortion). For the waveforms shown in Fig. 7.5(a) and (c), however, distortion would produce sum and difference tones, resulting in the generation of a strong fundamental, since difference tones between all the adjacent partials would be at this frequency.

Experiments with filtered sound by H. Fletcher (1934) and others appeared to support Helmholtz. When the lower harmonics of a complex tone are filtered out, the pitch remains the same. This phenomenon can be demonstrated by recording the sound of a musical instrument and playing it back through a high-pass filter to remove the fundamental (and even the lower harmonics). The missing fundamental is supplied by the ear of the listener.

7.6 ■ THEORIES OF PITCH: PLACE PITCH VERSUS PERIODICITY PITCH

Two major theories of pitch perception have gradually developed on the basis of numerous experiments in many different laboratories. They are usually referred to as the place (or frequency) theory and the periodicity (or time theory). Before discussing these theories, let us briefly review the relationship between frequency and period.

A periodic waveform is one that repeats itself after a certain interval of time, called the period T. The reciprocal of the period is the fundamental frequency f_1. If the waveform is complex, it can be resolved into a spectrum of partials with frequencies $2f_1$, $3f_1$, etc., called the harmonics (see Section 2.7). A periodic waveform need not have energy at its fundamental frequency f_1, as will become apparent later in this chapter. In a pulse waveform, the fundamental *frequency* is not necessarily the same as the pulse *rate*. The waveform in Fig. 7.5(c), for example, has $2/T$ pulses per second, although its fundamental frequency $f_1 = 1/T$ is only half as great. In determining pitch, the ear apparently performs *both a time analysis and a frequency analysis* of the sound wave and reaches its final decision after a considerable amount of computation!

The idea that vibrations of different frequencies excite resonant areas on the basilar membrane is often referred to as the *place theory* of hearing. According to this theory, the cochlea converts a vibration in time into a vibration pattern in space (along the basilar membrane), and this in turn excites a spatial pattern of neural activity. The place theory explains many aspects of auditory perception but fails to explain others.

Helmholtz regarded the basilar membrane as a frequency analyzer, with transfer fibers "tuned" to resonate at frequencies determined by their length, mass, and tension. A complex wave of sound pressure would excite regions of the basilar membrane corresponding to the frequencies of its components or partials, the higher frequencies acting on regions near the oval window, and the lower frequencies acting closer to the far end where the membrane is thick and loose. (Helmholtz was nearly correct; later investigations showed that the individual fibers are not free to resonate, but the membrane as a whole can create the effect of resonances.)

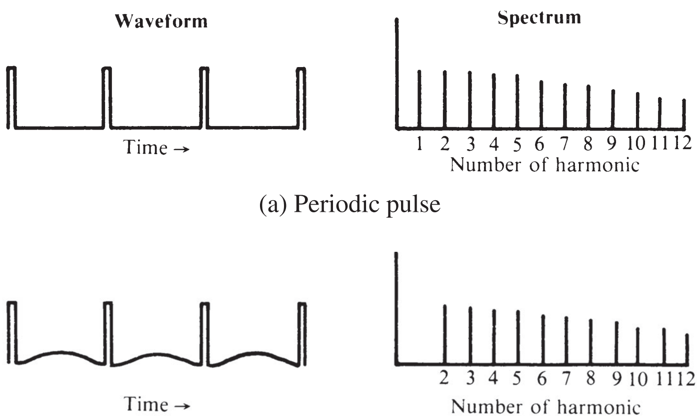
In his experiments with cochleas removed from human cadavers, Békésy provided support for the place theory of pitch perception. By ingenious and careful experiments, he directly observed wavelike motions of the basilar membrane caused by sound stimulation. Just as Helmholtz had suggested, the place of maximum vibration moved up and down the basilar membrane as the frequency of the sound wave changed (see Figs. 5.6 and 5.8).

More recent experiments have pointed to limitations in the place theory of pitch perception, however. One difficulty is in explaining fine frequency discrimination. In order to respond to rapid changes in frequency, a resonator must have considerable damping. But damping decreases selectivity, that is, the ability to discriminate between small differences in frequency. Another difficulty arises in attempting to explain why we hear a complex tone as one entity with a single pitch.

According to the *periodicity theory* of pitch, the ear performs a *time* analysis of the sound wave. Presumably the time distribution of the electrical impulses carried by the auditory nerve has encoded into it information about the time distribution of the sound wave. This information is decoded by a process called *autocorrelation* (to be discussed in Section 8.13) in the central nervous system.

In the late 1930s, J. F. Schouten and his colleagues in the Netherlands performed experiments that supported the periodicity theory of pitch. Schouten studied stimuli, such as those shown in Fig. 7.6, in which the pitch corresponds to the repetition rate of the pulses, 200 Hz. In the waveform shown in Fig 7.6(b), the fundamental component has been canceled out by addition of an out-of-phase signal of 200 Hz; the pitch remains unchanged at 200 Hz, the frequency of the missing fundamental. Schouten then added a pure tone of 206 Hz. If a distortion product of 200 Hz were actually present in the ear, as suggested by the hypothesis of Helmholtz, beats should be heard at a frequency of six per second. No beats were heard.

Schouten continued his experiments with a type now called pitch-shift experiments. Using amplitude modulation, he produced complex waveforms in which the frequencies of individual components could be shifted by the same amount, thus leaving the spacing between components undisturbed. For instance, a carrier frequency of 1200 Hz modulated by a 200-Hz signal produces components at 1000 Hz and 1400 Hz (called *sidebands*) along with the 1200-Hz component. Such a waveform, shown in Fig. 7.7, has a clear pitch of 200 Hz. If the carrier frequency is changed to 1240 Hz, however, the components are shifted to 1040, 1240, and 1440 Hz. The pitch is now found to shift to about 207 Hz, even though the difference frequency remains at 200 Hz. This experiment can be repeated in the laboratory using a generator with provision for amplitude modulation or an electronic music synthesizer. (See Demonstration 21, Houtsma, Rossing, and Wagenaars 1987).

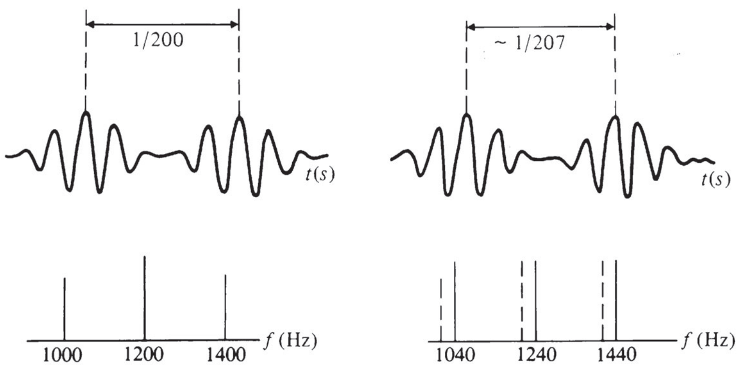


(a) Periodic pulse

(b) Periodic pulse without fundamental

FIGURE 7.6 Cancellation of the fundamental frequency of a complex signal. Part (a) shows a periodic pulse train and its spectrum. By appropriate adjustment of phase and amplitude, the fundamental may be canceled as shown in (b). In both cases, however, the pitch of the signal corresponds to the fundamental. (After Schouten 1940.)

The virtual pitch can be estimated by dividing the component frequencies by successive integers 5, 6, and 7 to obtain a "nearly common factor." In this case $1040/5 = 208$, $1240/6 = 206.7$, and $1440/7 = 205.7$. Averaging these three factors together gives 207 Hz, which the auditory system accepts as the frequency of the missing fundamen-



(a) (b)

FIGURE 7.7 Waveforms for pitch-shift experiments of the Schouten type: (a) carrier of 1200 Hz modulated at 200 Hz; (b) carrier at 1240 Hz modulated at 200 Hz.

tal. Using 4, 5, and 6 or 6, 7, and 8 leads to less consistent trial factors, so the auditory system prefers the 207-Hz factor.

Schouten explained the pitch-shift phenomena as due to synchronous firing in the auditory nerve due to an unresolved "residue" of high-frequency components. These components, too close in frequency to be resolved on the basilar membrane, retain the periodicity of the original tone envelope. Schouten's residue theory of pitch provided a reasonable alternative to the distortion hypothesis of Helmholtz, but subsequent experiments (e.g., Plomp (1967), Ritsma (1967)) showed that the pitch of complex tones is determined by the low-frequency (resolved) components rather than by the high-frequency (unresolved) residue. An excellent historical review of the subject is given by Plomp (1967).

The importance of some sort of *central* pitch processor in the nervous system was illustrated by experiments in which a single harmonic of a missing fundamental was presented to one ear and a different harmonic to the other ear (Houtsma and Goldstein 1972). The resulting virtual pitch heard this way (*dichotic* presentation) appeared to be as strong as when both harmonics were presented to the same ear (*monotic* presentation). In both monotic and dichotic presentations, the virtual pitch tends to deteriorate with increasing harmonic number.

One might correctly conclude from the foregoing discussion that both the place and periodicity theories of pitch have validity. Clues from both frequency and time analyses of the sound are used to determine pitch, although one or the other may predominate under certain conditions. For low-frequency tones, the time (periodicity) analysis appears to be more important, whereas at high frequencies, the frequency analysis in the basilar membrane (*place clues*) plays a more important role. The relative importance of each type of clue and the frequency range over which the clues predominate are still under study.

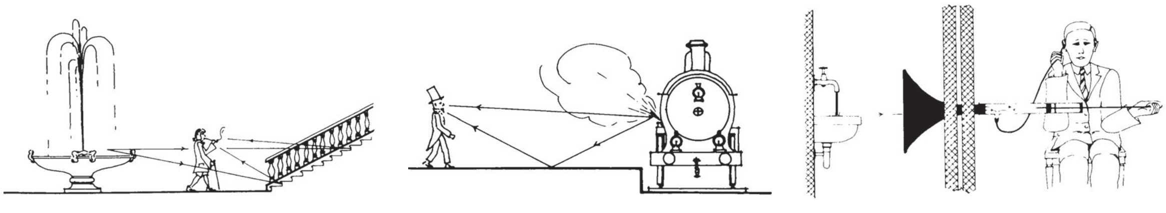
Modern Theories of Pitch

Modern theories of pitch, given such names as *optimum processor theory* (Goldstein 1973), *virtual pitch theory* (Terhardt 1974), and *pattern transformation theory* (Wightman 1973), describe how the ear-brain processor determines the pitch of complex tones. Each of them has attractive features. A detailed discussion of them is beyond the scope of this book.

Quite a few experiments have been conducted to evaluate the predictions of these theories (references are given in Scharf and Houtsma (1986) and in Houtsma and Rossing (1987)). Some of these experiments compare the observed pitch shifts in the complex tone with those observed in the partials due to masking noise, amplitude envelope change, intensity change, etc. Others compare complex tones made up of high and low partials, partials of unequal amplitude, etc. The general conclusion appears to be that none of the current pitch theories is completely successful in explaining all the experiments.

Repetition Pitch: A Demonstration of Pitch

In 1693, astronomer Christiaan Huygens, standing at the foot of a staircase at the castle at Chantilly de la Cour in France, noticed that sound from a nearby fountain produced a certain pitch. He correctly concluded that the pitch was caused by periodic reflections of the sound against the steps of the staircase. *Repetition pitch*



(a) (b) (c)

FIGURE 7.8 Examples of repetition pitch (from Bilsen and Ritsma 1969/1970): (a) Huygens (1693) observed periodic reflections of the noise of a fountain against the steps of a staircase; (b) Minnaert (1941) observed the interference of the hissing sound from a locomotive with its reflection from a platform; (c) Hermann (1912) observed interference between the noise of running water and its reflection in a tube of adjustable length.

due to interference between noise and its delayed repetition is discussed by Bilsen and Ritsma (1969/1970), who describe several historical examples, including those shown in Fig. 7.8.

Repetition pitch can be demonstrated in a number of ways, including the two shown in Fig. 7.9. In both cases, broadband noise is combined with identical noise

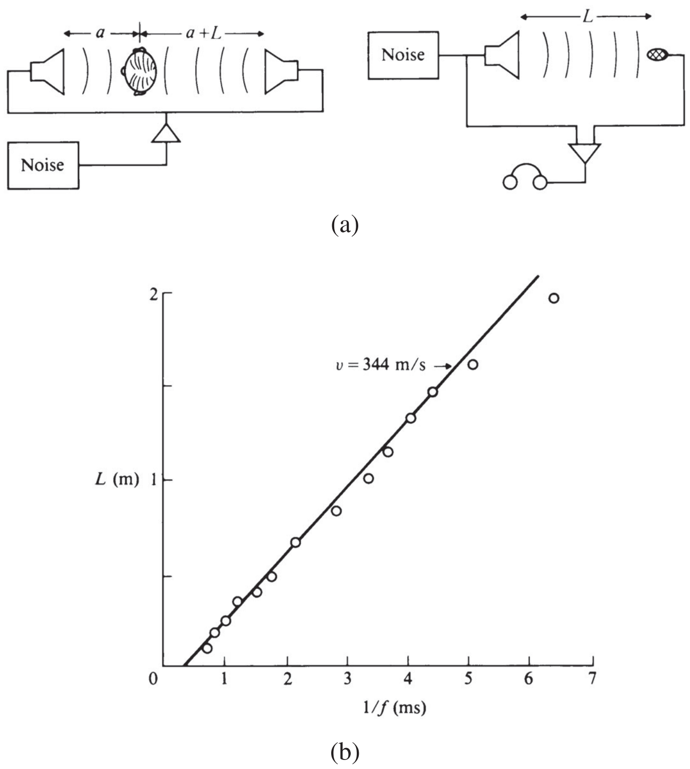


FIGURE 7.9 (a) Two ways to demonstrate repetition pitch by combining noise with similar noise delayed by time $T = L/v$; (b) Pitch change with time delay using the second arrangement in (a). (After Rossing and Hartmann 1975.)

that has traveled a distance L farther and thus is delayed by a time $T = L/v$, where v is the speed of sound. The perceived pitch corresponds to a frequency $f = 1/T = v/L$ and is easy to identify for time delays of 1 to 7 ms. Some blind persons make use of the phenomenon to locate obstructions by observing the interference between the direct and reflected sound.

7.7 ■ ABSOLUTE PITCH

A subject that has held considerable fascination, but also causes no small amount of controversy, is *absolute pitch*. The term refers to the ability to recognize and define (e.g., by naming or singing) the pitch of a tone without the use of a reference tone. This ability is often compared to absolute recognition of color (e.g., green) without any comparison to a standard spectrum. Whereas absolute color recognition is possessed by about 98% of the population (only 2% being partially or totally colorblind), absolute pitch recognition is rare (less than 0.01% of the population appears to have it).

Absolute pitch contrasts with relative pitch, which most persons have to some degree. Nearly all persons can tell whether one tone is higher than another; persons with some musical experience or training can recognize intervals between tones with varying degrees of precision. Someone with a well-trained ear can tell when the frequency of a second tone deviates a little as one percent from the expected interval, although these judgments are not as accurate as they are consistent. For example, the frequency that a person judges, with great consistency, to be an octave above a 1000-Hz tone may actually be 2060 Hz. Relative pitch, in fact, is a remarkable sensory ability that has no counterpart in our other senses. We cannot judge a color that has twice the frequency of a reference color; the only comparable judgment in the visual domain might be selection of a complementary color, and few people develop the ability to do that with great accuracy.

Psychologists have studied absolute pitch for at least 75 years, and during that time there has been considerable discussion and some controversy concerning its origin. In particular, there is less than unanimous agreement as to whether absolute pitch is inherited, acquired, or possibly both. At least four different theories about absolute pitch have developed (Ward 1963):

1. *Heredity theory.* The faculty for developing absolute pitch is inherited, just as the ability for color identification is (unless one inherits colorblindness). The child, so gifted, learns pitch names in early life just as color names are learned.
2. *Learning theory.* The opposite point of view, that absolute pitch can be acquired by almost anyone by diligent and constant practice, is not too widely held.
3. *Unlearning theory.* The ability to develop absolute pitch is nearly universal, but is simply trained out of most people at an early age (by emphasis on relative pitch, for example).
4. *Imprinting theory.* Imprinting is a term used to describe rapid irreversible learning that takes place at a specific developmental stage (used to explain, for example, why duck-

lings will follow for the rest of their life the first moving object they see after hatching). Proponents of this theory feel that nearly all children could be taught absolute pitch at the appropriate age of development.

Bachem (1955) distinguishes between *chroma* and *tone height* as two separate components of pitch. All A's up and down the scale have the same chroma, or quality, but differ in tone height (possessors of absolute pitch frequently make octave errors in identifying tones). Above about 5000 Hz, chroma tends to become fixed, whereas tone height continues to increase, so that absolute pitch identification is not possible.

At least one person with absolute pitch has reported a change in his internal pitch standard with time (Vernon 1977). At age 52 he noted a tendency to identify keys one semitone higher than they should be. He was troubled because he heard the overture to Wagner's "Die Meistersinger" in the "effeminate" key of C$^\sharp$ rather than the "strong and masculine" key of C. By age 71, however, it had moved still further into the sturdier key of D! The shift of internal pitch standard may have been due to a change in elasticity of the basilar membrane with age; in other words a tone of a given frequency was invoking maximum activity at a different place on the basilar membrane than in earlier years.

Speakers of *tone languages*, in which a speech sound can take on several different meanings depending on its tone (see Section 15.8), appear to have a knack for absolute pitch. Vietnamese and Mandarin speakers repeat words on different days with pitches within a semitone, demonstrating a remarkably precise and stable absolute pitch template in producing words (Deutsch, Henthorn, and Dolson 1999).

However it develops, absolute pitch is a remarkable ability. Absolute pitch (inherited or acquired) may continue to be a controversial subject for some time to come, because of the obvious difficulty of experimenting with human subjects in isolation. If one really wants a child to acquire absolute pitch (it has disadvantages as well as advantages!), one should probably begin as early as possible to play find-the-note games on the piano.

7.8 ■ PITCH STANDARDS

The advantages of a universal pitch standard are so obvious that it is quite remarkable that for so many years there was none. Pipe organs were built with A's tuned all the way from 374 to 567 Hz (Helmholtz 1877). In 1619, Praetorius suggested a pitch of 424 Hz; Handel's tuning fork reportedly vibrated at 422.5 Hz. This pitch standard prevailed, more or less, for two centuries, and it is the pitch standard for which Hayden, Mozart, Bach, and Beethoven composed.

Early in the nineteenth century pitch began to rise, probably due to an increased use of brass instruments, which were found to sound more brilliant at the higher pitch. In 1859 a commission appointed by the French government (which included Berlioz, Meyerbeer, and Rossini) selected 435 Hz as a standard. Early in the twentieth century a *scientific pitch*, with all the C's being powers of 2 (128, 256, 512, and so on), appeared; this leads to about 431 Hz for A. Unfortunately, tuning forks made to this standard are still being distributed by scientific and medical supply houses.

In 1939 an International Conference in London unanimously adopted 440 Hz as the standard frequency for A_4, and this is almost universally used by musicians. A few orchestras have once again begun a “pitch-raising” game by tuning to 442 or even 444 Hz for greater brightness. This is unfortunate, however, because instruments designed to play well at one pitch may not retain their tone or intonation at another (this is especially true of woodwinds). Singers of today sing the arias of Mozart and Beethoven about a semitone above the pitch for which they were written; most violins of the old masters have already had to be strengthened by adding stouter bass bars and necks to accommodate the increased string tension of today’s pitch standard.

Tuning forks have served as convenient pitch standards since the time of Handel. More recently, quartz crystals have provided us with a more precise standard for measuring frequency as well as time. Electronic frequency counters and stroboscopic tuners have made it possible for every physics laboratory as well as every band or orchestra to have precise and dependable frequency standards. The United States Bureau of Standards broadcasts an exceedingly precise 440-Hz tone on its short wave radio station WWV for checking local standards.

The frequency of most musical instruments changes with temperature, and those using wood and gut also change with humidity. The velocity of sound increases about 0.6 m/s for each degree Celsius, so the pitch of a wind instrument rises about 3 cents ($\frac{3}{100}$ of a semitone) per degree of temperature rise (the slight lowering of pitch due to expansion in length is negligible). String instruments generally fall in pitch due to relaxing tension as temperature rises.

7.9 ■ TIMBRE OR TONE QUALITY

The word *timbre*, borrowed from French, is used to denote the tone quality or tone color of a sound. The American National Standards Institute (1960) defines it: “Timbre is that attribute of auditory sensation in terms of which a listener can judge two sounds similarly presented and having the same loudness and pitch as dissimilar.” An explanatory note is added: “Timbre depends primarily on the spectrum of the stimulus, but it also depends upon the waveform, the sound pressure, the frequency location of the spectrum, and the temporal characteristics of the stimulus.” This definition suggests that judgment of timbre must take place under conditions of equal loudness and pitch (and probably equal duration as well), and so Pratt and Doak (1976) have suggested an alternative definition: “Timbre is that attribute of auditory sensation whereby a listener can judge that two sounds are dissimilar using any criteria other than pitch, loudness or duration.”

Timbre may be described as a multidimensional attribute of sound (Plomp 1970); it is impossible to construct a single subjective scale of timbre of the type used for loudness (sones) and pitch (mels), for example. Two recent attempts to construct subjective scales, by asking listeners to rate various verbal attributes of steady sounds, are illustrated in Fig. 7.10. Each investigator found the dull–sharp (brilliant) scale the most significant.

In discussing timbre, and especially in reading about the many experiments on timbre described in the literature, it is important to distinguish between the timbre of *steady* complex tones and those that include *transients* or other variations with time. Plomp (1970) suggested the possibility of using *tone color* to refer to the perceptual differences between steady complex tones; this suggestion has not been widely accepted, however.

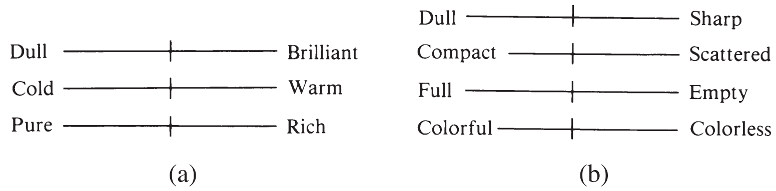


FIGURE 7.10 Subjective rating scales for timbre: (a) Pratt and Doak (1976); (b) von Bismarck (1974).

A thorough investigation of the timbre of steady tones was carried out by Helmholtz (1877). Helmholtz demonstrated that the sounds of most musical instruments (including the vocal folds or cords) consist of series of harmonics that determine the timbre. Furthermore, he carefully described a way in which the ear could comprehend timbre. On the basis of his experiments, he formulated the following general rules:

1. Simple tones, such as those of tuning forks and widely stopped organ pipes, have very soft, pleasant sound, free from roughness but dull at low frequencies.
2. Musical tones with a moderately loud series of harmonics up to the sixth (such as those produced by the piano, the French horn, and the human voice) sound richer and more musical than simple tones, yet remain sweet and soft if the higher harmonics are absent.
3. Tones consisting of only odd harmonics (narrow stopped organ pipes, clarinet) sound hollow and, if many harmonics are present, nasal. When the fundamental predominates, the quality of tone is rich; when the fundamental is not sufficiently strong, the quality of tone is poor.
4. Complex tones with strong harmonics above the sixth or seventh are very distinct, but the quality of tone is rough and cutting.

Helmholtz continued with careful experiments to determine the dependence of timbre on the relative phases of the harmonics. Using electrically driven tuning forks and tuned resonators (of the type we now call Helmholtz resonators), he concluded that timbre does not depend on phase differences between the harmonics. Unfortunately, Helmholtz could detect only very slow changes in phase in his experiments (a limitation that he apparently recognized), and thus some interesting dynamic phase effects were overlooked. So thorough were the studies of Helmholtz, that until 1950 very little new information of significance appeared in the literature.

Before continuing the discussion of timbre, we will investigate the Fourier analysis of a tone.

7.10 ■ FOURIER ANALYSIS OF COMPLEX TONES

The determination of the harmonic components of a periodic waveform is called *Fourier analysis*, after the mathematician Joseph Fourier (1768–1830), who formulated an important mathematical theorem: *Any periodic vibration, however complicated, can be built up from a series of simple vibrations, whose frequencies are harmonics of a fundamental frequency, by choosing the proper amplitudes and phases of these harmonics*. Constructing a

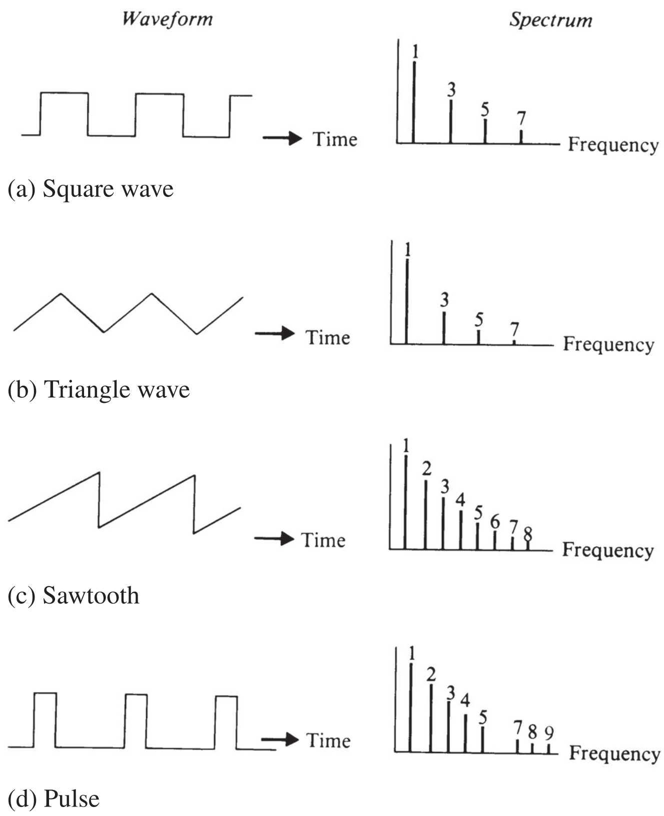


FIGURE 7.11 Spectra of complex waveforms. The square wave and triangle wave are missing all the even-numbered harmonics.

complex tone from its harmonics (the opposite of Fourier analysis) is called *Fourier synthesis*. The terms *spectrum analysis*, *harmonic analysis*, and *sound analysis* are sometimes used to describe Fourier analysis applied to sound. A specification of the strengths of the various harmonics (usually in the form of a graph) is called a *spectrum*.

Spectra of four different complex waveforms are shown in Fig. 7.11. Although they sound rather harsh and unmusical (with the exception of the flutelike triangle wave), these waveforms are frequently used to create sound in electronic music synthesizers (see Chapter 27). The square wave, for example, is composed of only odd-numbered harmonics with amplitudes in the ratio $1/n$. Thus, if the fundamental has frequency f and amplitude A, the other harmonics in the spectrum will have frequencies of $3f$, $5f$, $7f, \ldots$, and amplitude $A/3$, $A/5$, $A/7, \ldots$. The triangle wave has odd harmonics with amplitudes in the ratio $1/n^2$ (that is, A, $A/9$, $A/25, \ldots$). The sawtooth wave, on the other hand, has both odd-numbered and even-numbered harmonics with amplitudes in the ratio $1/n$ (A, $A/2$, $A/3, \ldots$).

Figure 7.12 illustrates how Fourier analysis works. The first six harmonics of a sawtooth wave are shown individually and collectively. Note that when combined *in the proper*

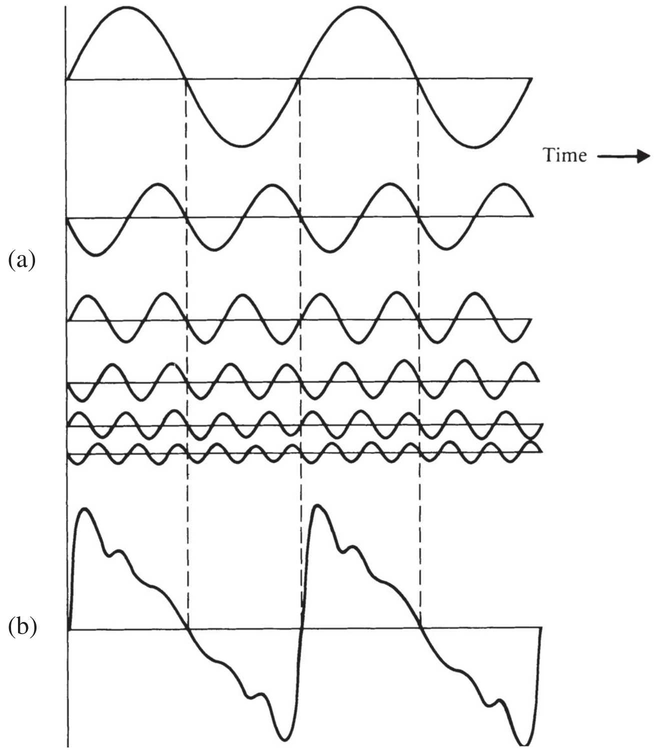


FIGURE 7.12
Fourier synthesis of a sawtooth wave: (a) first six harmonics; (b) sum of the first six harmonics.

phase, the first six harmonics approximate a sawtooth wave, although "wiggles" occur that diminish with the addition of higher harmonics.

Textbooks present many "typical" spectra of musical instruments. It should be emphasized, however, that sound spectra from a given instrument vary widely according to the way in which the instrument is played (soft, loud, high, low, or midrange) and how the sound is recorded (near field, far field, reverberant field, direction of microphone from the instrument, etc.).

One way to determine the spectrum of harmonics is by direct computation from the recorded waveform. One of the earliest instruments developed for recording waveforms was the *phonodeik* designed by D. C. Miller (1916), which used a vibrating mirror to direct a beam of light onto a moving film. Most of the sound spectra in early publications were calculated from phonodeik recordings.

Modern spectrum analyzers are of two types: digital and analogue. Digital-spectrum analyzers begin by sampling one period of the wave at regular intervals and feeding these samples into a digital computer. The computer then calculates the amplitude and phase of each harmonic.

Analogue spectrum analyzers use filters or other electronic circuits to isolate the harmonics one after another. If this is done very rapidly (in a few milliseconds), the analyzer

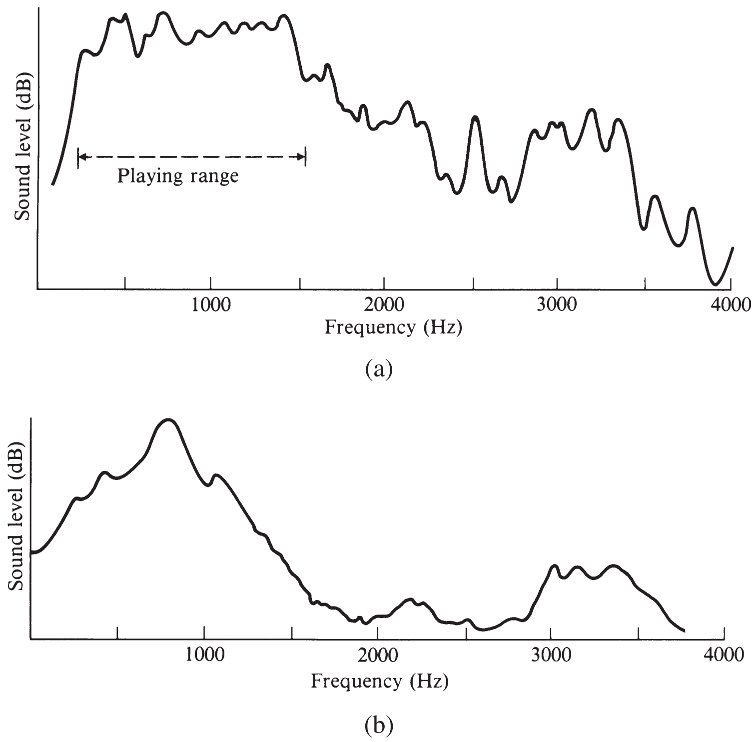


FIGURE 7.13
Time-averages of spectra; (a) clarinet; (b) tenor singing "ah."

is called a *real-time* spectrum analyzer, which is very useful for studying changing sounds or spectra during attack and decay of sounds.

Some interesting information about timbre can be obtained by averaging many spectra (from the same instrument, for example). Figure 7.13 shows averages of 512 spectra of a clarinet and a male voice. In each case, the pitch is varied by playing (singing) up and down the scale during the recording. The significance of the various maxima will be become clear after reading about woodwind instruments (Chapter 12) and voice formants (Chapter 15).

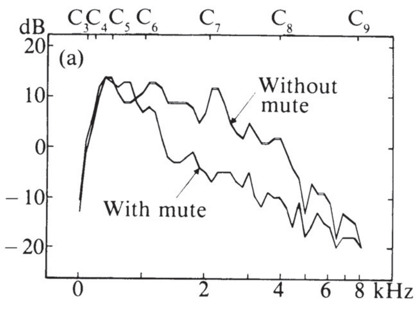


FIGURE 7.14
Long-time average spectra of a violin with and without a mute. (From Jansson and Sundberg 1975.)

Long-time-average spectra have been used extensively at the Royal Institute of Technology in Stockholm to study musical instruments and the singing voice. A long-time-average spectrum contains information on the written music, the performance, the musical instrument, and the room in which it is played. The effect of varying any one of these factors can be studied by holding the others constant. Figure 7.14 shows the long-time-average spectra of a violin played with and without a mute, for example.

It should be mentioned that whereas the effects of phase on timbre are small for steady tones, the ear is in fact quite sensitive to *changes* in phase, especially if they take place at a regular rate. This is illustrated by the phenomenon described as *second-order beats*, to be discussed in Chapter 8.

7.11 ■ TIMBRE AND DYNAMIC EFFECTS: ENVELOPE AND DURATION

In Sections 7.9 and 7.10, the discussion focused on the timbre of steady complex tones. Transients and other dynamic effects, however, play an important role in determining the timbre of musical and speech sounds, as you can prove to yourself by two simple experiments.

Record the sounds of a number of different musical instruments. In the first experiment, play the tape backward (so that the attack transient occurs at the end). You will hear some curious effects. For example, a piano played backward sounds like a reed organ or a harmonium. (This is illustrated in Demonstration 29, Houtsma, Rossing, and Wagenaars 1987). For a second experiment, cut and splice the tape so that the attack transient is removed. Without attack transients, a remarkable similarity is noted between dissimilar pairs of instruments, such as a French horn and a saxophone and even a trumpet and an oboe.

Berger (1963) performed an experiment in which the sounds of various instruments were presented with the first and last half seconds removed; using 30 band students as a jury of listeners, he obtained the *confusion matrix* shown in Table 7.1. Note that with the transients removed, the sound of an alto saxophone was correctly identified by only four

TABLE 7.1 Listener judgments of recorded wind-instrument tones presented with first and last half seconds removed (Berger 1963)

	Response										
Stimulus	Flute	Oboe	Clarinet	Tenor saxophone	Alto saxophone	Trumpet	Cornet	French horn	Baritone	Trombone	No answer
Flute	1	2		1	6	5	4			4	7
Oboe		28									2
Clarinet	1	1	20	4	3						1
Tenor saxophone			25	2	1						2
Alto saxophone				3	4		1	11	5	5	1
Trumpet	8				6	2	3	4	1	3	3
Cornet		1				12	15				2
French horn	1			2	3			5	6	6	7
Baritone			1	1	2	3	2	4	7	3	7
Trombone	2	1		5	3			1	5	9	4

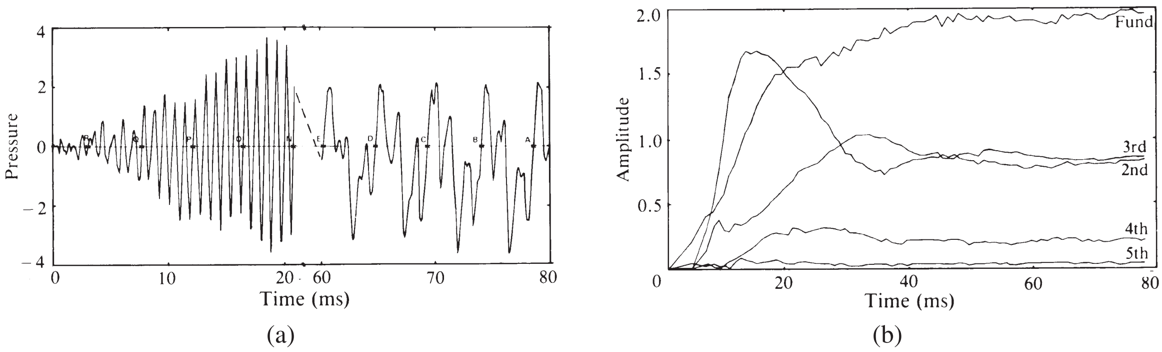


FIGURE 7.15 (a) The waveform of an attack transient. (b) Amplitudes of the first five harmonics of the attack transient of a 110-Hz diapason organ pipe. (From Keeler 1972.)

jurists, whereas eleven jurists thought it was a French horn. Also surprising is the confusion of tenor saxophone with clarinet, because the "woody" tone of a clarinet emphasizes odd-numbered harmonics.

During attack, the various partials of a musical sound may develop at different rates. Figure 7.15 shows the attack waveform of an organ pipe tones, along with the onset of the first five harmonics. During the attack transient, the waveform is not exactly periodic; note the difference from cycle to cycle. Note that the second harmonic of the organ pipe develops slightly faster than the others; in other wind instruments, the fundamental is often found to lead.

Strong and Clark (1967) performed some interesting experiments in which they interchanged spectra and time envelopes of wind instrument tones. They synthesized many tones, each time using the envelope characteristic of one instrument with the spectrum of another, and asked listeners to identify the instrument. They found that in the cases of the oboe, clarinet, bassoon, tuba, and trumpet, the spectrum is much more important than the envelope; in the case of the flute, the envelope is more important than the spectrum; in the cases of the trombone and French horn, spectrum and envelope appear to be of comparable importance. The general principle seems to be that the spectrum takes on greatest importance when it has a maximum in a unique location within its playing range.

7.12 ■ VIBRATO

Vibrato is widely used to enhance musical performance, both instrumental and vocal. In order to avoid misunderstanding, it is important to carefully define vibrato.

The definition recommended by the American National Standards Institute (1960) is "The vibrato is a family of tonal effects in music that depend on periodic variations of one or more characteristics in the sound wave." The important note is added: "When the particular characteristics are known, the term 'vibrato' should be modified accordingly: e.g., frequency vibrato, amplitude vibrato, phase vibrato and so forth." In keeping with this recommendation, we use the term *frequency vibrato* to refer to frequency modulation (FM) and *amplitude vibrato* to refer to amplitude modulation (AM). In practice it is virtually impossible to have frequency vibrato without amplitude vibrato because of the effect of room resonances and resonances in the source instrument. Amplitude vibrato without frequency vibrato is possible (in the case of a vibraphone with resonators that open and close periodically, for example), but would be the exception rather than the rule.

Unfortunately, some texts use the term *vibrato* to refer to frequency vibrato but the term *tremolo* to refer to amplitude vibrato. This is unfortunate, not only because frequency vibrato and amplitude vibrato nearly always coexist in musical performance, but also because the term *tremolo* is generally used in music to refer to something else: rapid back-and-forth strokes of a violin bow or rapid alternation between two notes.

Vibrato was studied extensively some 40 to 50 years ago by C. E. Seashore and colleagues at the University of Iowa, and many of their findings are confirmed by more recent experiments (Ward 1970). Vibrato appears to vary with individual performers, an “average” rate for both singers and instrumentalists being around 7 Hz. Singers seem to use a slightly greater depth of frequency vibrato than instrumentalists do, however.

You can perform interesting experiments on vibrato using an audio generator with provision for frequency modulation (many generators can be frequency modulated by a second oscillator), or with an electronic music synthesizer. Try varying both the *rate* and the *depth* of frequency modulation. You will probably find that with modulation rate in the rage of 1 to 5 Hz, you can recognize the periodicity of pitch change (most clearly around 4 Hz). Beginning at about 6 Hz, however, the tone takes on a single average pitch with intensity fluctuations at the frequency of the vibrato. At a still higher rate (around 12 Hz), the sound becomes a rather unpleasant confusion of more than one tone. It is not difficult to see why performers choose a vibrato rate around 7 Hz.

The parameters of a natural vibrato fluctuate slightly during the duration of a tone. Tones from electronic instruments, which have a fixed rate and depth of vibrato, sound artificially rigid. Analyses of the vibrato used by opera singers Maria Callas and Dietrich Fischer-Dieskau show that both singers use deep vibratos (Winckel 1975). The rates of vibrato and trill used by Callas were the same, and in fact her transition from vibrato to trill was made with no change of phase. When trained singers sing duets, they reportedly adjust their vibratos to have identical rate and phase (but not necessarily depth); the adjustment is most likely subconscious (Winckel 1975).

Vibrato is said to cover up small errors in frequency. Fletcher, Blackham, and Geertsen (1965) found that the vibrato of many violinists apparently centers 15 to 20 cents above the target pitch. Vibrato makes identification of vowel sounds more difficult and tends to conceal formant frequencies of singers that may deviate substantially from the corresponding formant frequencies of normal speech (Sundberg 1975).

7.13 ■ BLEND OF COMPLEX TONES

Our auditory system has the ability to listen to complex sounds in different modes. When we listen *analytically*, we hear the different partials separately; when we listen *synthetically* or holistically, we focus on the whole sound and pay little attention to the partial sounds. Listeners differ in the degree to which they listen analytically or synthetically. If a two-tone complex of 800 and 1000 Hz is followed by one of 750 and 1000 Hz, for example, an analytic listener will hear one partial go down in pitch; a synthetic listener will hear a virtual pitch rising a major third from 200 to 250 Hz (Demonstration 25, Houtsma, Rossing, and Wagenaars 1987).

A tone with several harmonic partials, whose frequencies and relative amplitudes remain steady, is generally heard as a single tone, even if the total intensity changes. However, when one of the harmonics is turned off and on, it stands out clearly (Demonstration 1, Houtsma, Rossing, and Wagenaars 1987). The same is true if one of the harmonics is given a vibrato (i.e., its frequency, its amplitude, or its phase is modulated at a slow rate).

One of the most remarkable feats of our auditory system is its ability to single out complex tones from a complex background, such as the sounds of different instruments in a symphony orchestra or conversation at a cocktail party, for example. In the former case, the ear interprets certain partial tones as belonging to one particular instrument, other partials as belonging to another instrument. In other words, it looks for familiar or likely sets of partial tones and fuses these together into a single complex tone at the same time it hears a blend of many instrument sounds. As we have seen, the mechanism for this analysis is partially understood, but much research remains to be done.

Erickson (1975) addresses this subject from the standpoint of a composer. In an enlightening chapter, "Some Territory Between Timbre and Pitch," he discusses three ways in which a complex sound can be heard: (1) as a chord; (2) as a pitch (with timbre); (3) as a sound (an unpitched sound without definite pitch or pitches such as the sound of a bass drum). These three concepts can be represented as the apexes of a triangle (see Fig. 7.16) with the grey areas between them represented by the sides of the triangle.

Transformation from a chord to the fused condition described as a sound, for example, is illustrated by the music of Edgard Varese. A pitch (with timbre)-to-chord transformation occurs in the unusual chanting of Tibetan lamas recorded and described by Smith, Stevens, and Tomlinson (1967). The chanting is done in such a way that certain harmonics of the voice become audible as separate pitches, giving the effect of one person singing a continuous chord.

It is well known that the partials in a piano tone are stretched further apart than partials in a true harmonic series (see Chapter 14). Stretching the partials even further apart causes the sounds to become bell-like or chime-like. More surprising, perhaps, is the observation that compression of the partials also produces bell-like timbres (Slaymaker 1970). Individual partials, in both cases, can be singled out more easily than the harmonic partials of the usual musical tone; the transformation can be described as going from pitch (with timbre) to an inharmonic chord as the partials are stretched or compressed beyond certain limits.

Inharmonicity in the partials of a complex tone appears to be detected in a different way for low and high harmonics. For low harmonics, the inharmonic partial appears to "stand

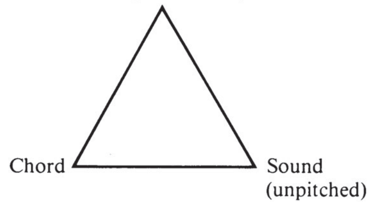


FIGURE 7.16 Three ways in which a complex sound can be heard. (From Erickson 1975.)

out" when it is mistuned by an amount that varies from 1 to 3% in different subjects. For high harmonics, on the other hand, the mistuning is detected as a kind of beat, or roughness, presumably reflecting a sensitivity to changing phase of the mistuned harmonic relative to the other harmonics (Moore, Peters, and Glasberg 1985).

7.14 ■ SUMMARY

Pitch has been defined as the characteristic of a sound that gives it the sensation of high or low. It is determined mainly by the frequency of a tone, but sound level, spectrum, and duration also influence pitch. Early models for pitch perception regarded the basilar membrane as a frequency analyzer of high resolution (place theory), but more recent studies have shown that much of the determination of pitch is contributed by a temporal analysis in the central nervous system (periodicity pitch). The ear is able to assign a pitch to complex sounds composed of inharmonic partials and even to some presentations of wideband noise. Some persons have the ability to identify pitch independent of a reference pitch (absolute pitch).

Timbre or tone quality depends on the frequency of a tone, its time envelope, its duration, and the sound level at which it is heard. Any complex waveform that is periodic can be constructed from simple tones with the right frequency and phase; determination of the spectrum of simple tones is called *spectrum analysis* or *Fourier analysis*. Under most conditions, the timbre of a complex sound is insensitive to the phase of its components. Periodic variation of the frequency and amplitude, called *vibrato*, lends warmth and blend to musical tones. A vibrato rate of about 7 Hz is common in musical performance.

REFERENCES AND SUGGESTED READINGS

American National Standards Institute (1960). "USA Standard Acoustical Terminology." S1.1–1960.

Appel, K., J. Gastineau, C. Bakken, and D. Vernier (1998). *Physics with Computers*. Portland, Ore.: Vernier Software.

Bachem, A. (1955). "Absolute Pitch," *J. Acoust. Soc. Am*. **27**: 1180.

Bassett, I. G. (1982). "Vibration and Sound of the Bass Drum," *Percussive Notes* **19**(3): 50.

Berger, K. W. (1963). "Some Factors in the Recognition of Timbre," *J. Acoust. Soc. Am*. **36**: 1888.

Bilsen, F. A., and R. J. Ritsma (1969/1970). "Repetition Pitch and Its Implication for Hearing Theory," *Acustica* **22**: 63.

Bismarck, G. von (1974). "Timbre of Steady Sounds: A Factorial Investigation of Its Verbal Attributes," *Acustica* **30**: 146.

Bürck, W., P. Kotowski, and H. Lichte (1935). "Die Horbarkeit von Laufzeitdifferenzen," *Elektrotechn. Nachr.-Techn*. **12**: 355.

Deutsch, D., T. Henthorn, and M. Dolson (1999). "Absolute Pitch Demonstrated in Speakers of Tone Languages," *J. Acoust. Soc. Am*. **106**: 2267 (abstract).

Doughty, J. M., and W. M. Garner (1948). "Pitch Characteristics of Short Tones II: Pitch as a Function of Duration," *J. Exp. Psychol*. **38**: 478.

Erickson, R. (1975). *Sound Structure in Music*. Berkeley: University of California. (See especially Chapter 2.)

Evans, E. F., and J. P. Wilson, eds. (1977). *Psychophysics and Psychology of Hearing*. London: Academic Press.

Fletcher, H. (1934). "Loudness, Pitch and Timbre of Musical Tones and Their Relation to the Intensity, the Frequency, and the Overtone Structure," *J. Acoust. Soc. Am*. **6**: 153. See also *Phys. Rev*. **23**: 427 (1924).

Fletcher, H., E. D. Blackham, and O. N. Geertsen (1965). "Quality of Violin, Viola, Cello and Bass-Viol Tones," *J. Acoust. Soc. Am*. **37**: 851.

Goldstein, J. L. (1973). "An Optimum Processor Theory for the Central Information of Complex Tones," *J. Acoust. Soc. Am.* **54**: 1496.

Grey, J. M. (1977). "Multidimensional Perceptual Scaling of Musical Timbres," *J. Acoust. Soc. Am.* **61**: 1270.

Hartmann, W. M. (1978). "The Effect of Amplitude Envelope on the Pitch of Sinewave Tones," *J. Acoust. Soc. Am.* **63**: 1105.

Helmholtz, H. L. F. von (1877). *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, 4th ed. Trans. A. J. Ellis. New York: Dover, 1954.

Houtsma, A. J. M., and J. L. Goldstein (1972). "The Central Origin of the Pitch of Complex Tones: Evidence from Musical Interval Recognition," *J. Acoust. Soc. Am.* **51**: 520.

Houtsma, A. J. M., and T. D. Rossing (1987). "Effects of Signal Envelope on the Pitch of Short Complex Tones," *J. Acoust. Soc. Am.* **81**: 439.

Houtsma, A. J. M., T. D. Rossing, and W. M. Wagenaar (1987). *Auditory Demonstrations* (Philips Compact Disc #1126-061 and text). Melville, N.Y.: Acoustical Society of America.

Jansson, E. V., and J. Sundberg (1975). "Long-Time-Average Spectra Applied to Analysis of Music. Part I: Method and General Applications," *Acustica* **34**: 15.

Keeler, J. S. (1972). "Piecewise-Periodic Analysis of Almost-Periodic Sounds and Musical Transients," *IEEE Trans. Audio and Electroacoustics* **AU-20**: 338.

Majerník, V., and J. Kalužný (1979). "On the Auditory Uncertainty Relations," *Acustica* **43**: 132.

Miller, D. C. (1916). *Science of Musical Sounds*. New York: Macmillan.

Moore, D. C. J., R. W. Peters, and B. R. Glasberg (1985). "Threshold for the Detection of Inharmonicity in Complex Tones," *J. Acoust. Soc. Am.* **77**: 1861.

Parkin, P. H. (1974). "Pitch Change During Reverberant Decay," *J. Sound and Vibration* **32**: 530.

Plomp, R. (1967). "Pitch of Complex Tones," *J. Acoust. Soc. Am.* **41**: 1526.

Plomp, R. (1970). "Timbre as a Multidimensional Attribute of Complex Tones," in *Frequency Analysis and Periodicity Detection in Hearing*. Eds., R. Plomp and G. Smoorenburg. Leiden: Sijthoff.

Pratt, R. L., and P. E. Doak (1976). "A Subjective Rating Scale for Timbre," *J. Sound and Vibration* **45**: 317.

Ritsma, R. (1967). "Frequencies Dominant in the Perception of the Pitch of Complex Sounds," *J. Acoust. Soc. Am.* **42**: 191.

Rossing, T. D., and W. M. Hartmann (1975). "Perception of Pitch from Time-Delayed White Noise," AAPT meeting, Anaheim, California (January 1975).

Rossing, T. D., and A. J. M. Houtsma (1986). "Effects of Signal Envelope on the Pitch of Short Sinusoidal Tones," *J. Acoust. Soc. Am.* **79**: 1926.

Savart, F. (1840). "Uber die Ursachen der Tonhohe," *Ann. Phys. Chem.* **51**: 555.

Scharf, B., and A. J. M. Houtsma (1986). "Audition II: Loudness, Pitch, Localization, Aural Distortion, Pathology," in *Handbook of Perception and Human Performance*, Vol. 1. Eds., K. R. Boff, L. Kaufman, and J. P. Thomas. New York: J. Wiley.

Schouten, J. F. (1940). "The Perception of Pitch," *Philips Tech. Rev*. **5**: 286. (This article summarizes much of the contents of several articles in *Proc. Ned. Acad. Wet*. **41** and **43**.)

Slaymaker, F. H. (1970). "Chords from Tones Having Stretched Partials," *J. Acoust. Soc. Am.* **47**: 1569–1571.

Smith, H., K. N. Stevens, and R. S. Tomlinson (1967). "On an Unusual Mode of Chanting by Tibetan Lamas," *J. Acoust. Soc. Am.* **41**: 1262.

Stevens, S. S. (1935). "The Relation of Pitch to Intensity," *J. Acoust. Soc. Am.* **6**: 150.

Strong, W., and M. Clark, Jr. (1967). "Perturbations of Synthetic Orchestral Wind-Instrument Tones," *J. Acoust. Soc. Am.* **41**: 277.

Sundberg, J. (1975). "Vibrato and Vowel Identification," in Report STL-QPSR 2-3/1975, Speech Technology Laboratory, Royal Institute of Technology, Stockholm.

Terhardt, E. (1974). "Pitch, Consonance, and Harmony," *J. Acoust. Soc. Am.* **55**: 1061.

Terhardt, E. (1979). "Calculating Virtual Pitch," *Hearing Research* **1**: 155.

Terhardt, E., and H. Fastl (1971). "Zum Einfluss von Störtönen und Störgeräuschen auf die Tonhöhe von Sinustönen," *Acustica* **25**: 53.

Vernon, P. E. (1977). "Absolute Pitch: A Case Study," *Br. J. Psychol*. **68**: 485.

Ward, W. D. (1963). "Absolute Pitch," *Sound* **2**(3): 14; **2**(4): 33.

Ward, W. D. (1970). "Musical Perception," in *Foundations of Modern Auditory Theory*, Ed. J. V. Tobias. New York: Academic.

Wightman, F. L. (1973). "The Pattern-Transformation Model of Pitch," *J. Acoust. Soc. Am.* **54**: 407.

Wightman, F. L., and D. M. Green (1974). "The Perception of Pitch," *Am. Scientist* **62**: 208.

Winckel, F. (1967). *Music, Sound and Sensation*. Trans., T. Binkley. New York: Dover.

Winckel, F. (1975). "Measurement of the Acoustic Effectiveness and Quality of Trained Singers' Voices," 90th meeting of Acoustical Society of America, San Francisco.

Zwicker, E. (1962). "Direct Comparisons between the Sensations Produced by Frequency Modulation and Amplitude Modulation," *J. Acoust. Soc. Am.* **34**: 1425.

Zwicker, E., G. Flottorp, and S. S. Stevens (1957). "Critical Bandwidth in Loudness Summation," *J. Acoust. Soc. Am.* **29**: 548.

GLOSSARY

absolute pitch The ability to identify the pitch of any tone without the aid of a reference.

analytic listening Listening to a complex tone in a way that individual components or partial tones are heard as separate entities.

bark An interval of frequency equal to a critical bandwidth.

critical band The frequency bandwidth at which subjective response (to loudness, pitch, etc.) changes rather abruptly (see Chapter 6).

distortion An undesired change in waveform. Two common examples are harmonic distortion and intermodulation distortion. *Harmonic distortion* means that harmonics are generated by altering the waveform in some way ("clipping" the peaks, for example). *Intermodulation distortion* refers to the generation of sum and difference tones.

envelope The amplitude of a tone as a function of time.

Fourier analysis, or **spectrum analysis** The determination of the component tones that make up a complex tone or waveform.

Fourier synthesis The creation of a complex tone or waveform by combining its spectral components.

fundamental The lowest common factor in a series of harmonic partials. The fundamental frequency of a periodic waveform is the reciprocal of its period.

harmonic A partial whose frequency is a multiple of some fundamental frequency.

inharmonic partial A partial that is not a harmonic of the fundamental.

just noticeable difference (jnd) or **difference limen** The minimum change in stimulus that can be detected.

mel The unit of subjective pitch; doubling the number of mels doubles the subjective pitch for most listeners. The critical band is about 100 mels wide.

octave The basic unit in most musical scales. Notes judged an octave apart have frequencies nearly in the ratio 2:1.

partial tone (or **partial**) One of the components in a complex tone (it may or may not be a harmonic of the fundamental).

period The smallest increment of time over which a waveform repeats itself.

periodic quantity One that repeats itself at regular time intervals.

periodicity pitch Pitch determination on the basis of the period of the waveform of a tone.

phase The fractional part of a period through which a waveform has passed, measured from a reference.

pitch An attribute of auditory sensation by which sounds may be ordered from low to high.

place theory of pitch A view of the basilar membrane as a frequency analyzer of high resolution; pitch is determined by sensing the place on the basilar membrane that has maximum excitation.

repetition pitch Pitch sensation created by the interference of a sound with a time-delayed repetition.

residue theory of pitch A view that components of a tone that cannot be resolved by the basilar membrane (the residue) are analyzed in time by the central nervous system.

semitone One step on a chromatic scale. Normally $\frac{1}{12}$ of an octave.

spectral dominance A view that certain partials dominate in the determination of the pitch of a complex tone.

spectrum The "recipe" for a complex tone that gives the amplitude and frequency of the various partials.

strike note Note heard when a bell or chime is struck.

subjective pitch Pitch determined to have a frequency that does not correspond to that of any partial.

synthetic (holistic) listening Listening to a complex tone in a way that focuses on the whole sound rather than the individual partials.

timbre An attribute of auditory sensation by which two sounds with the same loudness and pitch can be judged dissimilar.

transient A sound that does not reoccur, at least on a regular basis.

tristimulus diagram A way of representing timbre graphically in terms of the relative loudness of three different parts of the spectrum.

vibrato Tonal effect in music resulting from periodic variation of amplitude, frequency, and/or phase.

virtual pitch Subjective pitch created by two or more partials in a complex tone (two examples are the "missing fundamental" of a filtered tone and the strike note of a bell).

REVIEW QUESTIONS

1. What is the basic unit in most musical scales and what frequency ratio does it represent?

2. How does the jnd for pitch generally compare to the critical bandwidth at the same frequency?

3. On what physical parameter(s) does pitch depend?

4. How does the pitch of a 200-Hz tone depend on sound level?

5. How does the pitch of a tone change if noise of a lower frequency is added?

6. What pitch will generally be heard when tones of 800, 1000, and 1200 Hz are sounded together?

7. What is meant by *absolute pitch*?

8. What is the frequency of A_4, according to the International pitch standard?

9. Does the timbre of a complex tone depend on the relative phases of its harmonics?

10. What frequencies would appear in the spectrum of a 100-Hz square wave?

11. Does playing a tone backward change its spectrum?

12. Does playing a tone backward change its timbre?

13. Vibrato is generally defined as a periodic change in
(a) frequency
(b) amplitude
(c) phase
(d) timbre
(e) all of these

14. A preferred vibrato rate is
(a) 3 Hz
(b) 7 Hz
(c) 15 Hz
(d) 100 Hz
(e) depends upon the frequency of the tone

15. A single harmonic in a complex tone can be made to stand out by
(a) turning it on and off
(b) modulating its frequency
(c) modulating its amplitude
(d) mistuning it
(e) any of these

QUESTIONS FOR THOUGHT AND DISCUSSION

1. A "tonic" chord in the key of A consists of tones with frequencies of 440, 550, and 660 Hz. When such a chord is played on the piano or by three instruments, why is this not heard as a single tone with a pitch of 110 Hz (the "missing fundamental")?

2. Have you ever experienced the pitch change during reverberation described by Parkin? Would this effect be apparent on a recording with reverberation?

3. Discuss the advantages and disadvantages to a performing musician of possessing absolute pitch.

4. Try to account for the most prevalent "confusions" in Berger's experiment (Table 7.1) in the identification of instrument tones without the transients.

5. Why is it virtually impossible to have frequency vibrato in a musical instrument without amplitude vibrato?

EXERCISES

1. At what point would you divide a 65-cm guitar string (as Pythagoras did) so that the two segments sound pitches one octave apart?

2. From Fig. 7.2, find the jnd at frequencies 200, 1000, and 5000 Hz.

3. By referring to Fig. 7.2, show that the critical band comprises roughly 30 jnds. (Compare them at 200, 1000, 5000, and 10,000 Hz, for example.)

4. According to Fig. 7.3, how many cents does the pitch of a 200-Hz tone fall when the sound pressure level is changed from 50 to 90 dB?

5. In Fig. 7.5(c) let $t_1 = 7$ ms and $t_2 = 3$ ms. Determine $1/T$, $2/T$, and $3/T$. What is the frequency of the pitch that would be heard? What is the pulse rate?

6. From Fig. 7.15(b), determine the approximate rise times of the first and second harmonics of a diapason organ pipe.

7. From Fig. 7.1, determine the number of mels in an octave from

(a) C_3 (131 Hz) to C_4 (262 Hz);

(b) C_4 to C_5 (523 Hz);

(c) C_5 to C_6 (1046 Hz).

8. If a pure tone with a frequency of 800 Hz is modulated at 150 Hz, what sidebands are produced? According to the theory discussed in Section 7.6, what virtual pitch will probably be heard? (Try dividing by various sets of integers such as 4, 5, 6 and 5, 6, 7, etc.)

9. If the steps in Fig. 7.8(a) are 30 cm deep, what pitch would most likely be heard?

10. Compare the tension in a violin string tuned to a standard A (440 Hz) with the tension in the same string tuned to match Handel's tuning fork (422 Hz). (See Section 3.2.)

11. What are the frequencies of the first four partials in a 300-Hz square wave?

12. What is the frequency of the maximum sound level in the spectrum of Fig. 7.13(b)?

EXPERIMENTS FOR HOME, LABORATORY, AND CLASSROOM DEMONSTRATION

Home and Classroom Demonstration

1. *Dependence of pitch on intensity* Demonstration 12 on the *Auditory Demonstrations* CD (Houtsma, Rossing, and Wagenaars 1987). Tone bursts having frequencies of 200, 500, 1000, 3000, and 4000 Hz are presented at two levels 30 dB apart. Does the pitch of the second tone sound lower, higher, or the same as the first pitch?

2. *Dependence of pitch on intensity* An audio generator is connected to an amplifier and loudspeaker so that pure tones (sine waves) of different frequencies (200 to 4000 Hz) can be heard different sound levels. Does the pitch rise, fall, or stay the same when the intensity is increased?

3. *Pitch salience and tone duration* Demonstration 13 on the *Auditory Demonstrations* CD. Tones of 300, 1000, and 3000 Hz are presented in bursts of 1, 2, 4, 8, 16, 32, 64, and 128 periods. How many periods are necessary to establish a sense of pitch?

4. *Influence of masking noise on pitch* Demonstration 14 on the *Auditory Demonstrations* CD. A 1000-Hz tone, 500 ms in duration and partially masked by noise low-pass filtered at 900 Hz, alternates with an identical tone presented without masking noise. The tone partially masked by noise of lower frequency generally appears slightly higher in pitch. Do you agree?

5. *Octave matching* Demonstration 15 on the *Auditory Demonstrations* CD. A 500-Hz tone alternates with another tone that varies from 985 to 1035 Hz in steps of 5 Hz. Which one sounds like a correct octave? Most listeners select a tone somewhere around 1010 Hz, which illustrates our preference for "stretched" octaves.

6. *Stretched and compressed scales* Demonstration 16 on the *Auditory Demonstrations* CD. Another demonstration illustrating preference for stretched intonation. A melody is played in a high register with an accompaniment in a low register.

7. *Difference limen, or jnd* Demonstration 17 on the *Auditory Demonstrations* CD. Ten groups of four tone pairs are presented. In each tone pair the second tone pair may be higher or lower than the first (write down which you hear). The frequency difference decreases with each group.

8. *Difference limen, or jnd* The difference limen for frequency is conveniently demonstrated by a two-tone switching generator (such as the Automated Industrial Electronics 2TSG-1) or a computer by switching back and forth between tones of frequency f and $f + \Delta f$.

9. *Seebeck's siren* The waveforms shown in Fig. 7.1 are generated electronically (with a pulse generator (see T. D. Rossing, "Seebeck's Siren," *Phys. Teach*. **17**: 352 (1959)) or with a computer), displayed on an oscilloscope, and fed to an audio amplifier and loudspeaker. The abrupt disappearance of the $1/T$ tone when $t_2 - t_1 = 0$ is rather dramatic.

10. *Virtual pitch* Demonstration 20 on the *Auditory Demonstrations* CD. A complex tone consisting of 10 harmonics of 200 Hz is presented, followed by the same tone without the fundamental, with the two lowest harmonics, etc. Does the pitch of the complex tone change?

11. *Shift of virtual pitch* Demonstration 21 on the *Auditory Demonstrations* CD. The 800-, 1000-, and 1200-Hz partials in

a complex tone are shifted upward in steps of 20 Hz. The virtual pitch is heard to rise. Shifting to 850, 1050, and 1250 Hz, for example, generally produces a shift in virtual pitch from 200 to 210 Hz, as can be determined by matching to a spectral pitch.

12. *Shift of virtual pitch* Schouten's pitch-shift experiment (Fig. 7.3) can be done with an amplitude-modulated audio signal. Some signal generators may require a band-reject filter to eliminate leakage of the modulation tone into the output.

13. *Masking spectral and virtual pitch* Demonstration 22 on the *Auditory Demonstrations* CD. The Westminster chime melody is played with pairs of tones. The first tone of each pair is a pure tone, the second a complex tone with the same pitch. Low-pass noise masks only the pure-tone notes, whereas high-pass noise masks only the virtual pitch of the complex tone.

14. *Virtual pitch with random harmonics* Demonstration 23 on the *Auditory Demonstrations* CD. The Westminster chime melody is presented with various harmonics of a missing fundamental.

15. *Strike note of a chime* Demonstration 24 on the *Auditory Demonstrations* CD. An orchestral chime is struck eight times, each time preceded by cue tones equal to the first eight partials of the chime. In most orchestral chimes the virtual pitch of the strike note lies between the second and third partial.

16. *Analytic versus synthetic pitch* Demonstration 25 on the *Auditory Demonstrations* CD. A two-tone complex of 800 and 1000 Hz is followed by one of 750 and 1000 Hz. Do you hear the pitch go up or down? If you listen analytically, you will hear one partial go down in pitch; if you listen synthetically you will hear the virtual pitch go up a major third (from 200 to 250 Hz).

17. *Scales with repetition pitch* Demonstration 26 on the *Auditory Demonstrations* CD. Repetition pitch can be demonstrated by playing scales or melodies with pairs of pulses having appropriate time delays between members of a pair.

18. *Repetition pitch* Repetition pitch can be demonstrated using a tape recording made with a movable microphone (see Fig. 7.9). As a home experiment, moving the head between two loudspeakers (first drawing in Fig. 7.9) works well.

19. *Circularity in pitch judgment* Demonstration 27 on the *Auditory Demonstrations* CD. The *Shepherd scale*, which demonstrates circularity in pitch judgment, is an auditory analog to the ever-ascending staircase visual illusion.

20. *Spectrum analysis* Use a tunable bandpass filter to present each of the harmonics of a square wave, both on an oscilloscope and aurally through headphones or through an audio amplifier and loudspeaker. Do the same for sustained tones from musical instruments recorded on a loop of tape in order to play back continuously.

21. *Fourier analysis* An FFT analyzer or a PC with an FFT card can be used to display the spectra of various waveforms and musical instrument sounds.

22. *Fourier synthesis* Fourier synthesis using separate oscillators plus a mixer (the Pasco 9300 combines them in one unit) or a computer is entertaining as well as instructive. Observe the synthesized waveforms both visually (on an oscilloscope) and audibly (on headphones or loudspeaker).

23. *Effect of spectrum on timbre* Demonstration 28 on the *Auditory Demonstrations* CD. A carillon bell and a guitar tone are synthesized in eight steps by adding successive partials. How many partials are necessary to make the sound source recognizable?

24. *Effect of tone envelope on timbre* Demonstration 29 on the *Auditory Demonstrations* CD. Piano tones, heard backward, do not sound like piano tones, even through the spectrum remains unchanged. This demonstrates the significant influence of temporal envelope (including attack and decay) on timbre.

25. *Change in timbre with transposition* Demonstration 30 on the *Auditory Demonstrations* CD. A three-octave scale, synthesized by transposing the highest note of a bassoon, sounds different from a scale played on a bassoon.

26. *Canceled harmonics* Demonstration 1 on the *Auditory Demonstrations* CD. When the amplitudes of all 20 harmonics in a tone remain steady, we tend to hear the tone holistically (as a single, complex tone). When a harmonic is canceled and restored, it calls attention to itself and we tend to listen analytically. This is demonstrated for harmonics 1 through 10.

Laboratory Experiments

Perception of pitch (Experiment 13 in *Acoustics Laboratory Experiments*)

Sound spectra (Experiment 11 in *Acoustics Laboratory Experiments*)

Tones, vowels, and telephones (Experiment 22 in *Physics with Computers* by Appel et al.)

The demonstration experiments on the *Auditory Demonstrations* CD can be used as laboratory experiments.