from: Audio in Media, Stanley R. Alten, 1981

1

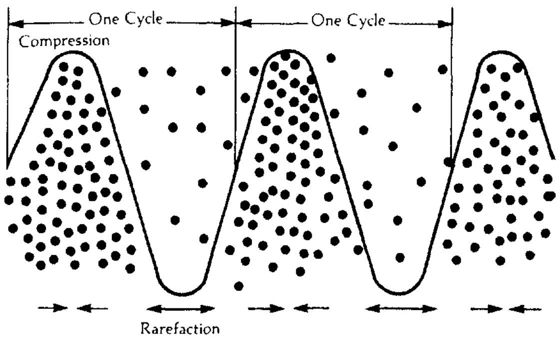
Sound and Hearing: An Introduction

What is sound? The answer depends on who you ask. A physicist will tell you that sound is both a disturbance of molecules caused by vibrations transmitted through an elastic medium (such as air) and the interaction of these vibrations with an environment. That definition does not mean much to the psychologist, who thinks of sound as a human response.

To put the question another way: If a tree falls in a forest and there is no one to hear it, does the falling tree make a sound? The physicist would say yes because a falling tree causes vibrations, and sound is vibration. The psychologist would probably say no because without a perceived sensation there can be no human response; hence, there is no sound. In practical terms, both the physicist and the psychologist are right. Sound is a cause-and-effect phenomenon, and the psychological cannot really be untangled from the physical. Thus, in audio production you need to understand both the objective and the subjective characteristics of sound. Not to do so is somewhat like arguing about the falling tree in the forest—it is an interesting, but unproductive exercise.

The Sound Wave

Imagine yourself listening to a guitarist performing at a live concert. The guitarist plucks a string and it vibrates, setting into motion the air molecules closest to the string. These molecules begin moving to and fro, like a pendulum, bumping into nearby molecules and starting a chain reaction. The chain reaction carries the original vibrations from the guitar string through the air to you. What makes the chain reaction possible is air, or more precisely, a medium with the property of **elasticity.**



1-1 *Sound pressure waves are formed when a vibration causes molecules to bump into nearby molecules, thus compressing them, increasing the sound pressure, and triggering a chain reaction. The original molecules then return to their original position and thereby create a reduction in sound pressure, or rarefaction.*

Elasticity is the characteristic that tends to pull a displaced molecule back to its original position after its initial momentum has caused it to displace the nearby molecules. **Compression** occurs when the disturbed molecules have enough momentum to bump against other molecules. As the molecules return to their original position, a partial vacuum, called a **rarefaction,** is created. When a vibration has passed through one to-and-fro motion, from compression to rarefaction, it has completed one cycle (see **1-1**).

The number of cycles that a vibration completes in 1 second determines its **frequency.** If a vibration completes 50 **cycles per second** (cps), its frequency is 50 **hertz** (Hz); if it completes 10,000 cps, its frequency is 10,000 Hz.*

Frequency and Pitch

Every vibration has a frequency, and generally humans are capable of hearing frequencies from 20 Hz to 16,000 Hz. Frequencies just above or below this range are felt more than heard, if perceived at all. Psychologically, we perceive frequency as **pitch**—the relative highness or lowness of a sound. The more times per second a sound wave vibrates, the higher its pitch. A whistle vibrates more times per second than a fog horn; therefore, its pitch is higher. The G string of a guitar vibrates 196 times per second, so its frequency is 196 Hz. Relative to the A string, which has a frequency of 110 Hz, the pitch of the G string is higher.

*The term "cycles per second" was used to designate frequency until a few years ago, when the term "hertz" was adopted in honor of the nineteenth-century German physicist Heinrich Hertz. We shall use "hertz" throughout this book.

Pitch also relates to our perception of a sound's tonal characteristic—that is, whether we hear it as tinny, bright, raspy, warm, hissy, and so on. The range of audible frequencies, or **frequency spectrum,** can be divided into sections, each of which has a unique and vital quality. The usual divisions in Western music are called **octaves.** An octave is the interval between any two frequencies that have a ratio of 2 to 1. The range of human hearing covers ten octaves. Starting with 20 Hz, the first octave is 20 to 40 Hz, the second 40 to 80 Hz, the third 80 to 160 Hz, and so on (see the diagram inside the back cover).

The first two octaves are the very low bass. These are the frequencies that give sound power, boom, and fullness. The lowest notes of the piano, organ, and bass are in this range, as are the lower frequencies of traffic, thunder, and explosions. The third and fourth octaves are the upper bass frequencies—the foundation of musical structure. Most tones generated by rhythm instruments such as drums and the bass are in this range. In fact, pitches in the first four octaves are very satisfying to the ear because we perceive them as giving sound an anchor or "bottom." One reason for the bass control on most home receivers and stereo systems is so that we can enhance these foundation frequencies.

If the **low end** of the frequency spectrum gives sound power and fullness, the **midrange**—the fifth, sixth, and seventh octaves—gives it energy. The midrange contains most of the **fundamental** frequencies in sound and music. A fundamental is the lowest or basic pitch of a sound. Due to the sound energy generated, the ear is quite sensitive to the midrange, and long listening sessions during which midrange frequencies predominate can be annoying and fatiguing.

We are also very sensitive to frequencies in the eighth octave, a rather curious range. Pitches in the lower half are extremely unpleasant to our ears; they sound harsh, lispy, and abrasive. The upper half of the eighth octave, however, contains rich and satisfying pitches that give sound definition, clarity, and realism. We perceive the frequencies in the upper eighth octave as being close to the listener, and for this reason it is also known as the **presence range.**

Although the ninth octave and the audible part of the tenth generate only 2 percent of the total power output of the frequency spectrum, they give sound the vital, lifelike qualities of brilliance and sparkle. If there is too much emphasis on this range, however, sound will be hissy. The treble control on most home receivers and stereo systems is used to control ninth- and tenth-octave frequencies.

Amplitude and Loudness

We have noted that vibrations stimulate molecules to move in pressure waves at certain rates, and that rate determines frequency. Vibrations not only affect the molecules' rate of movement, however; they also determine the number of molecules that are set in motion. This number depends on the intensity of a vibration; the more intense it is, the more molecules are displaced. The greater the number of molecules displaced, the bigger is the sound wave. The number of molecules in motion, and therefore the size of a sound wave, is called **amplitude** (see **1-2**). Our subjective impression of amplitude is **loudness**.

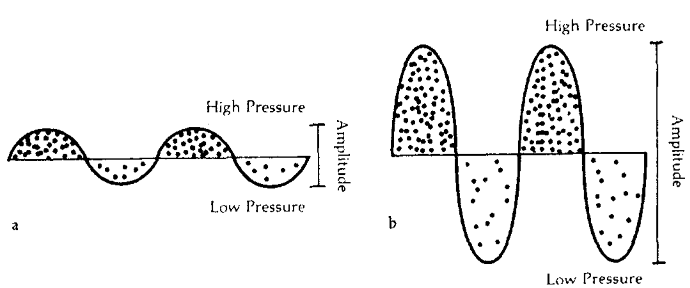
The ear's ability to hear wide variations in loudness is extraordinary. We measure loudness using the **decibel** (dB)—a unit that measures the relative intensity of acoustic pressure or voltage. Acoustic pressure is measured in **sound pressure level** (dB-SPL). Humans have the potential to hear a range from 0 dB-SPL, the threshold of hearing, to 120 dB-SPL, the threshold of pain, and beyond (see **1-3**). This range of quietness to loudness is called **dynamic range**.

Since this range is logarithmic, it means that humans have the capability to hear loudness at a ratio of 1 to 10,000,000 and greater. If the loudness of a sound of 100 dB-SPL were doubled, it would be 103 dB-SPL. Nevertheless, it takes an increase of at least 6 dB for most people to perceive a sound level as doubled.

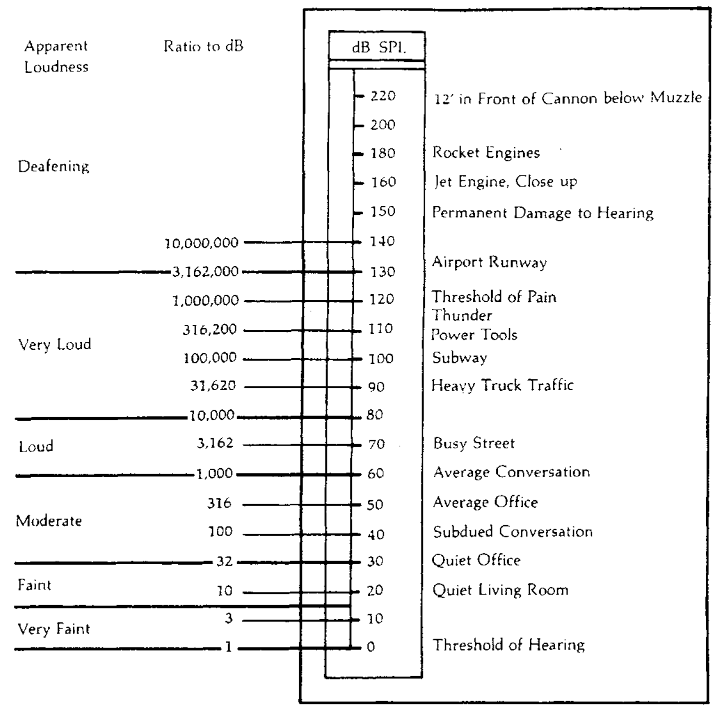
The Ear and Hearing Loss

Although the human ear can hear sounds over a wide dynamic range, hearing loss results if it is exposed to excessively loud levels for too long. Unfortunately, hearing loss from exposure to loud sounds is not regenerative, or self-corrective. The ear is a durable, but delicately integrated mechanism. It is divided into three parts: (1) the **outer ear,** (2) the **middle ear,** and (3) the **inner ear** (see **1-4**).

Sound waves first reach the outer ear, where they are collected and directed to the **auditory canal.** The auditory canal channels the sound waves to the eardrum, which then starts to vibrate. These vibrations are transmitted by three small bones in the middle ear to the inner ear. The inner ear is a spiral



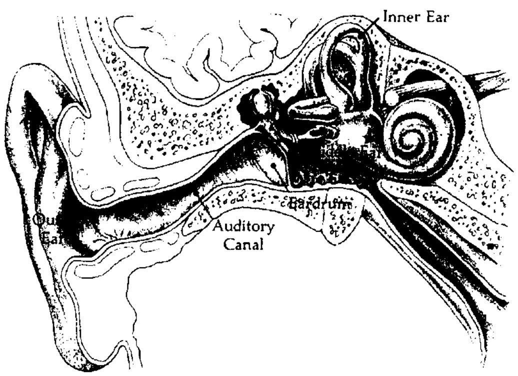
-2 The number of olecules displaced by a bration creates the mplitude or loudness of a und. Since the number of olecules in the sound wave part b is greater than the umber in the sound wave in art a, the amplitude of the und wave in b is greater.



1-3 *Sound pressure levels of common sound sources.*

filled with fluid. Within the spiral is a membrane that contains the **auditory nerve** endings. Vibrations in the fluid excite these nerve endings, which transmit impulses along the auditory nerve to the brain.

Deterioration of the auditory nerve endings occurs through the natural aging process and usually results in a gradual loss of hearing first in the higher frequencies and then in the lower frequencies. Prolonged listening to loud sounds adversely affects the auditory nerve endings and hastens their deterioration. It is not uncommon for young people who are constantly exposed to loud sound levels to have the hearing acuity of a 70-year-old person. To avoid premature deterioration of your auditory nerves, do not expose them to excessively loud sound levels for extended times (see **1-5**).



1-4 *Cross section of the human ear.*

1-5 *Hours of exposure to high sound levels permitted by the U.S. Government and the British Occupational Hygiene Society.*

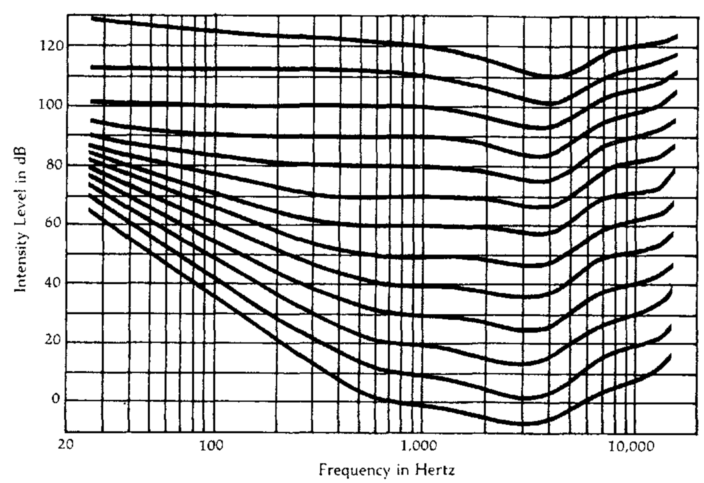
U.S. Government—Occupational Safety and Health Act		*British Occupational Hygiene Society*	
Sound Level (dB)	*Daily Permissible Hours of Exposure*	*Sound Level (dB)*	*Daily Permissible Hours of Exposure*
90	8	90	8
92	6	91	6
95	4	93	4
97	3	94	3
100	2	96	2
105	1	99	1
110	½	102	½
115	¼	105	¼

Frequency and Loudness

It is natural to assume that the greater the amplitude, the louder is the sound. This is not always so, however. If a guitarist plucks all six strings equally hard, you do not hear each string at the same loudness level. The high E string (328 Hz) sounds louder than the low E string (82 Hz). To make the low string sound as loud, the guitarist would have to pluck it harder.

This suggests that the high E string may sound louder because of its higher frequency. If you sound three tones—50 Hz, 1,000 Hz, and 15,000 Hz—at a fixed loudness level, however, the 1,000-Hz tone sounds louder than either the 50-Hz or the 15,000-Hz tone.

These observations show that the human ear is not equally sensitive to all audible frequencies; we do not hear low and high frequencies as well as we hear middle frequencies. This, oddly enough, is the principle of *equal*, rather than unequal, loudness (see **1-6**), and it has important implications for recording and listening. For example, if you have the loudness of a sound at a high level during recording and at a low level during playback, both low and high frequencies disappear. The converse is also true: If loudness is low when recording and high when playing back, the low and high frequencies predominate. We have noted that the bass and treble controls on a home receiver and stereo allow you to adjust the volume of low and high frequencies; the equal loudness principle is one reason they were put there.



1-6 ***Equal loudness*** *curves illustrate our relative lack of sensitivity to low and high frequencies compared with middle frequencies. A 50-Hz sound would have to be at 50 dB to seem as loud as a 1,000-Hz sound at 0 dB. If music is recorded at 100 dB and played back at 50 dB, the bass frequencies and the frequencies between 3,000 and 4,000 Hz sound weaker. (Based on Fletcher-Munson.)*

Velocity

Although frequency and amplitude are the most important physical components of a sound wave, a third component—**velocity, or the speed of a sound wave**—should be mentioned. Velocity usually has little impact on pitch or loudness and is relatively constant in a controlled environment; sound travels 1,130 feet per second at sea level and 70 degrees Fahrenheit. Velocity changes significantly in very high or low temperatures, however, increasing as air warms and decreasing as air cools. For every change of 1 degree Fahrenheit, the speed of sound changes 1.1 feet per second. If velocity increases or decreases greatly, the pitch of a sound rises or falls proportionately.

Interference

So far we have discussed the propagation and perception of sound under ideal conditions. However, in any form of communication generated by humans—spoken, written, played, painted, projected, or broadcast—some interference is not only possible but also probable. Interference may be external or internal. External interference is caused by something outside the transmitting medium—such as people munching popcorn during a movie, coughing during a concert, or shouting during a speech—or it may be due to a scratchy record, a smeared painting, or a torn page.

Internal interference is directly related to the transmitting medium—such as an out-of-tune instrument, a hoarse announcer, a "snowy" TV screen, or a noisy sound system. In audio, two of the most common types of internal interference are **distortion** and **noise**.

DISTORTION Distortion is any undesirable change in the **timbre**—the quality or color—of a sound. It can make a violin sound screechy, a trumpet raspy, or a voice scratchy. Usually, however, distortion is not so obvious but quite subtle; after you listen for a period of time it becomes apparent through a feeling of annoyance, uneasiness, or fatigue. Distortion exists in several forms: **harmonic distortion, intermodulation distortion, frequency distortion, loudness distortion, transient distortion,** and **spatial distortion**.

Harmonic Distortion Harmonic distortion is actually quite unharmonious. It occurs when the audio system introduces sounds into a recording that were not present originally. This usually happens when the input and output of a sound system are *nonlinear*—that is, do not change in direct proportion to each other.

Intermodulation Distortion Intermodulation distortion occurs when two or more different frequencies pass through an amplifier at the same time and interact to create combination tones and dissonances that are unrelated to the original sounds. Audio systems are most vulnerable to intermodulation distortion when frequencies are far apart, as when a piccolo and a baritone saxophone are playing at the same time. Intermodulation distortion usually occurs in the high frequencies because they are weaker and more delicate than the low frequencies.

Frequency Distortion Frequency distortion is caused when frequencies that are present in the input source are not present, or not equally reproduced, in the output. This usually occurs with inexpensive audio systems that cannot reproduce all audible frequencies at the same loudness, or that are incapable of reproducing the bass and treble frequencies.

Loudness Distortion Loudness distortion arises when a signal is recorded or played back at a level of loudness that is greater than the sound system can handle.

Transient Distortion Transient distortion relates to the inability of an audio component to record or reproduce **transients**—sounds with sudden, explosive attacks and quick decays, such as drumming, crashing, and popping. Sometimes transient distortion produces a ringing sound.

Spatial Distortion Spatial distortion is a factor in stereophonic, quadriphonic, and binaural recording and reproduction, where the position of sounds between two or four loudspeakers coming from front, rear, left, center, and right is important to aural balance (see Chapters 6 and 7). In stereo, for example, the sound of a vocalist usually comes from between the two loudspeakers and the accompaniment comes from the left, from behind, and from the right of the vocalist. If it sounds like the vocalist is coming from the left or right of center or like the accompaniment is in front of the vocalist, the problem is spatial distortion.

NOISE Noise is any unwanted electric or electromagnetic disturbance, other than distortion, that is heard at the output of an audio system. Three types of noise are: (1) **equipment noise** generated by components in the system or by fluorescent lights, power cables, or air conditioners; (2) **tape noise** generated by the recording tape (see Chapter 11); and (3) **system noise**, a combination of equipment and tape noise.

For a sound system to be professionally acceptable, the difference between the loudest sound it can record or reproduce and the inherent system noise generated by the sound must be as great as possible. This difference is known as **signal-to-noise** (S/N) **ratio** and is measured in decibels. Most professional audio systems have S/N ratios of at least 55 to 1. This means that it is possible to produce 55 dB of sound before the system generates 1 dB of noise. S/N ratios actually are expressed in negative numbers; hence, a S/N ratio of −55 to 1 is better than a S/N ratio of −50 to 1.

from Audio in Media, Stanley R. Alten, 1981