3 HUMAN PERCEPTION AND REACTION TO SOUND

3.1 HUMAN HEARING MECHANISMS

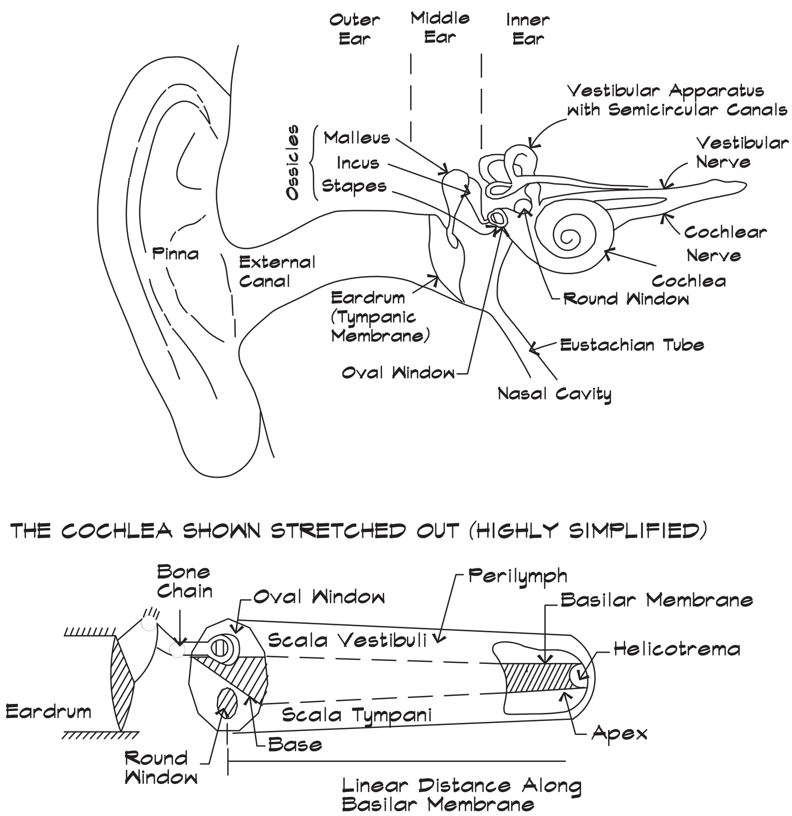
Physiology of the Ear

The human ear is an organ of marvelous sensitivity, complexity, and robustness. For a person with acute hearing, the range of audible sound spans ten octaves, from 20 Hz to 20,000 Hz. The wavelengths corresponding to these frequencies vary from 1.7 centimeters (5/8 inch) to 17 meters (57 feet), a ratio of one thousand. The quietest sound audible to the average human ear, about zero dB at 1000 Hz, corresponds to an acoustic pressure of 20×10^{-6} N/m^2 or Pa. Since atmospheric pressure is about 101,000 Pa (14.7 lb/sq in), it is clear that the ear is responding to extraordinarily small changes in pressure. Even at the threshold of pain, 120 dB, the acoustic pressures are still only about 20 Pa.

The excursion of the ear drum at the threshold of hearing is around 10^{-9} m (4×10^{-7} in) (Kinsler et al., 1982). Most atoms have dimensions of 1 to 2 angstroms (10^{-10} meters) so the ear drum travels a distance of less than 10 atomic diameters at the threshold of hearing. Were our ears only slightly more sensitive, we would hear the constant background noise due to Brownian movement, molecules set into motion by thermal excitation. Indeed, it is thermal motion of the hair cells in the *cochlea* that limits hearing acuity. In very quiet environments the flow of blood in the vessels near the eardrum is plainly audible as a disquieting shushing sound.

The anatomy of the ear, shown in Fig. 3.1, is organized into three parts, termed outer, middle, and inner. The outer and middle ear are air filled, whereas the inner ear is fluid filled. The outer part includes the *pinna*, the fleshy flap of skin that we normally think of as the ear, and a tube known as the *meatus* or auditory canal that conducts sound waves to the *tympanic membrane* or ear drum, separating the outer and middle ear sections. The pinna gathers the sound signals and assists in the localization of the height of a sound source. The 2.7 centimeter (one-inch) long auditory canal acts like a broadband quarter-wavelength tube resonator, whose lowest natural frequency is about 2700 Hz. This helps determine the range of frequencies where the ear is most sensitive—a more or less 3 kHz wide peak centered at about 3400 Hz. The auditory canal resonance increases the sound level at the ear drum around this frequency by about 10 dB above the level at the canal entrance. With the

FIGURE 3.1 A Schematic Representation of the Ear (Flanagan, 1972)



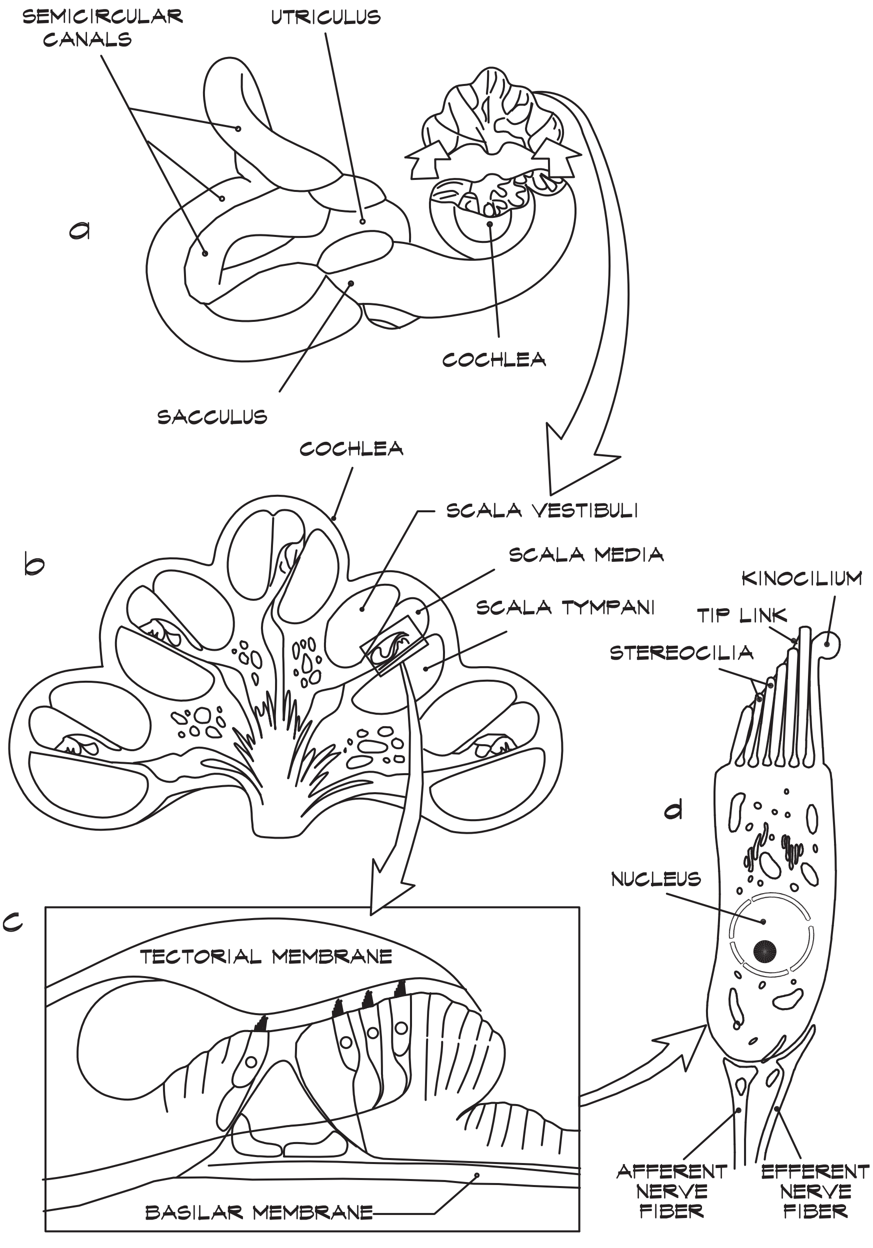
diffraction provided by the pinna and the head, there can be as much as a 15 to 20 dB gain at certain frequencies at the ear drum relative to the free-field level.

The middle ear is an air-filled cavity about 2 cu cm in volume (about the same as a sugar cube), which contains the mechanisms for the transfer of the motion of the eardrum to the cochlea in the inner ear. The ear drum is a thin conical membrane stretched across the end of the auditory canal. It is not a flat drum head, as might be inferred from its name, but rather a tent-like sheath with its peak pointing inward. Near its center, the eardrum is attached to the malleus bone, which is connected in turn to two other small bones. These three, the *malleus* (hammer), *incus* (anvil), and *stapes* (stirrup) act as a mechanical linkage, which couples the eardrum to the fluid-filled cochlea. The stapes resembles a stirrup with its base pressed up against the *oval window*, a membrane that covers the entrance to the cochlea. Because of the area ratio of the eardrum to that of the oval window (about 20 to 1) and the lever action of the ossicles, which produces another factor of 1.5:1, the middle ear acts as an impedance matching transformer, converting the low-pressure, high-displacement motion of the ear drum into a high-pressure, low-displacement motion of the fluid of the cochlea. Atmospheric pressure in the middle ear is equalized behind the eardrum by venting this area to the throat through the *eustation tube*, which opens when we yawn or swallow.

The motion transfer in the middle ear is not linear but depends on amplitude. An *aural reflex* protects the inner ear from loud noises by tightening the muscles holding the stapes to reduce its excursion at high amplitudes, just as the eye protects itself from bright light by contracting the pupil. The contraction is involuntary in both cases and seldom is noticed by the individual. Pain is produced at high noise levels when the muscles strain to protect nerve cells. Unfortunately the aural reflex is not completely effective. There is a reaction time of about 0.5 msec so it cannot block sounds having a rapid onset, such as gunshots and impact-generated noise. A second reason is that the muscles cannot contract indefinitely. Under a sustained bombardment of loud noise they grow tired and allow more energy to pass.

The inner ear, shown in Fig. 3.2, contains mechanisms that sense balance and acceleration as well as hearing. Housed in the hard bone of the skull, the inner ear contains five

FIGURE 3.2 Structure of the Inner Ear (Hudspeth and Markin, 1994)



separate receptor organs, each sensitive to a specific type of acceleration, as well as the cochlea, which detects the loudness and frequency content of airborne sound waves. The sacculus and utriculus include about 15,000 and 30,000 hair cells in planar sheets that react to vertical and horizontal linear accelerations respectively. These organs have the capability of encoding a unique signal for an acceleration in any given direction within a plane. Three *semicircular canals* are arranged to sense the orthogonal directions of angular acceleration. Each consists of a fluid-filled tube interrupted by a diaphragm containing about 7000 hair cells. They provide information on the orientation and acceleration of the human head. The bilateral symmetry of the ears gives us not only backup capability but extra information for the decomposition of motions in any direction.

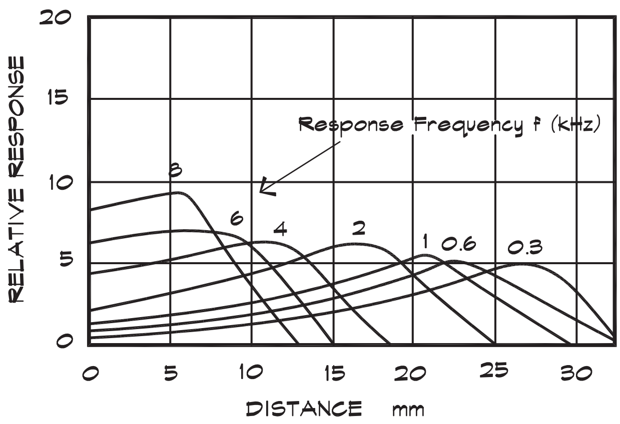
The cochlea is a fluid-filled tube containing the hair cell transducers that sense sound. It is rolled up two and one-half turns like a snail and if we unroll the tube and straighten it out, we would find a narrow cavity 3.5 cm long, about the size and shape of a golf tee scaled down by two-thirds. At its beginning, called the basal end, it is about 0.9 cm in diameter and at the apical end it is about 0.3 cm in diameter. It has two thin membranes running down it near its middle. The thicker membrane is called the *basilar membrane* and divides the cochlea more or less in half, separating the upper gallery (*scala vestibuli*) from the lower gallery (*scala tympani*). Along the membrane lies the *auditory nerve*, which conducts the electrochemical impulses and snakes through a thin sliver of bone called the bony ridge to the brain.

The entrance to the cochlea, in the upper gallery, is the *oval window* at the foot of the stapes. At the upper end of the cochlea near its apex there is a small passageway connecting the upper and lower galleries called the *helicotrema*. At the distal end of the lower gallery near the oval window is another membrane, the *round window*. It acts like the back door to the cochlea, a pressure release surface for fluid impulses traveling along its length and back into the middle ear. The two membranes, the oval window and the round window, seal in the fluid of the cochlea. Otherwise the rest of the cochlea is completely surrounded and protected by bone.

Figure 3.2b shows a cross section of one of the spirals of the cochlea. The upper gallery is separated from a pie-shaped section called the middle gallery (*scala media*) by *Reissner's membrane*. Within this segment and attached to the basilar membrane is the *organ of Corti*, which includes some 16,000 small groups of hair cells (*stereocilia*), arranged in four rows, acting as motion transducers to convert fluid and basilar membrane movement into electrical impulses (Hudspeth and Markin, 1994). The stereocilia are cylindrical rods arranged in a row in order of increasing height and move back and forth as a group in response to pressure waves in the endolymphatic fluid. The hair cells are relatively stiff and only move about a diameter. Through this movement they encode the magnitude and the time passage of the wave as an electrochemical potential, which is sent along to the brain.

Each stereocilia forms a bond between its end and an area on the adjacent higher neighbor much like a spring pulling on a swinging gate (see Fig. 3.2d). When a gate is opened a nerve impulse is triggered and sent to the brain. If the bundle of stereocilia is displaced in the positive direction, toward the high side of the bundle, a greater relative displacement occurs between each stalk and more gates are opened. A negative displacement, towards the short side of the bundle reduces the tension on the biomechanical spring and closes gates. Orthogonal motion results in no change in tension and no change in the signal. The amplitude of the response to sound waves is detected by the number of gate openings and closings and thus the number of impulses sent up the auditory nerve.

FIGURE 3.3 Longitudinal Distance along the Cochlea Showing the Positions of Response Maxima (Hassall and Zeveri, 1979)



As the stereocilia move back and forth they are sometimes stimulated to a degree that pushes them farther than their normal excursion. In these cases a phenomena known as *adaptation* occurs wherein the hair cells acquire a new resting point that is displaced from their original point. The cells find a new operating position and do a recalibration or reattachment of the spring to a gate at a slightly different point on the neighboring cell. Adaptation also suggests a mechanism for hearing loss when hair cells are displaced beyond the point where they can recover due to exposure to loud sounds over a long time period.

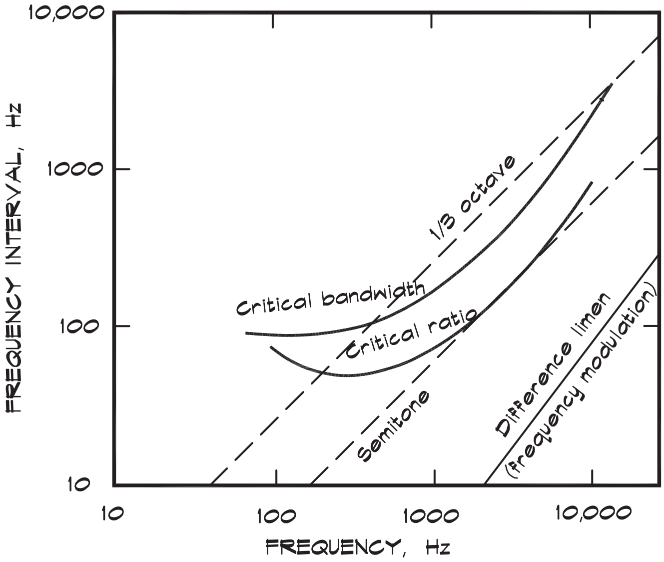
The frequency of the sound is detected by the position of greatest response along the basilar membrane. As a pressure wave moves through the cochlea it induces a ripple motion in the basilar membrane and for each frequency there is a maximum displacement in a certain region. The high frequencies stimulate the area closest to the oval window, while the low frequencies excite the area near the helicotrema. Figure 3.3 illustrates this phenomenon, known as the place theory of pitch detection. It was originated by von Bekesy (1960), and he received a Nobel Prize for his work. The brain can interpret nerve impulses coming from a specific area of the cochlea as a certain sound frequency. There are about 5000 separately detectable pitches over the 10 octaves of audibility.

3.2 PITCH

Critical Bands

Pitch is sensed by the position of maximum excursion along the basilar membrane. This is the ear's spectrum analyzer. There are some 24 discrete areas, each about 1.3 mm long and each contains about 1300 neurons (Scharf, 1970). These act as a set of parallel band-pass filters to separate the incoming sounds into their spectral components. Like electronic filters, the cochlear filters have bandwidths and filter skirts that overlap other bands. When two tones are close enough together that there is significant overlap in their skirts they are said to be within the same *critical band*. The region of influence, which constitutes a critical band, is illustrated in Fig. 3.4. For many phenomena it is about one-third-octave wide. The lower frequencies are sensed by the cochlea at a greater distance from the stapes. The shape of the resonance is not symmetric, having a tail that extends back along the basilar membrane (upward in frequency) from the center frequency of the band. Thus a lower pitched sound

FIGURE 3.4 Critical Bandwidths of the Ear (Kinsler et al., 1982)



can have a region of influence on a higher pitched sound, but not vice versa unless the sounds are quite close in frequency.

The phenomenon of critical bands is of great significance for many aspects of human hearing. They play a role in music by defining regions of consonance and dissonance. They influence the calculation of loudness by determining the method of combination used for multiple tones. They are critical to the phenomenon of masking, explaining many of the varied masking experiments.

Consonance and Dissonance

When two tones are played together, there is a frequency range over which they sound rough or dissonant (Fig. 3.5). Hermann von Helmholtz in his famous book, *On the Sensations of Tone*, hypothesized that the phenomenon of consonance was closely related to the frequency separation of tones and their harmonics. He thought that when two tones or their partials had a difference frequency of 30 to 40 Hz, this caused unpleasant beats. Subsequent experiments by Plomp and Levelt (1965) added some additional factors to his hypothesis.

Plomp's experiments revealed that the maximum dissonance occurs at about 25% of the critical bandwidth. Figure 3.6 gives a graph of consonance and dissonance as a function

FIGURE 3.5 Auditory Perception within a Critical Band (Pierce,1983)

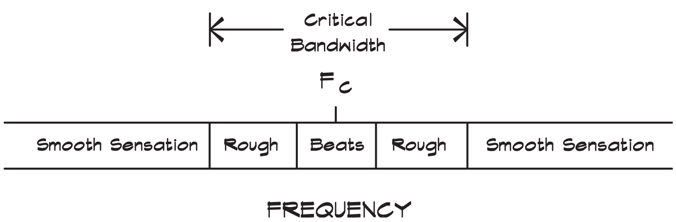
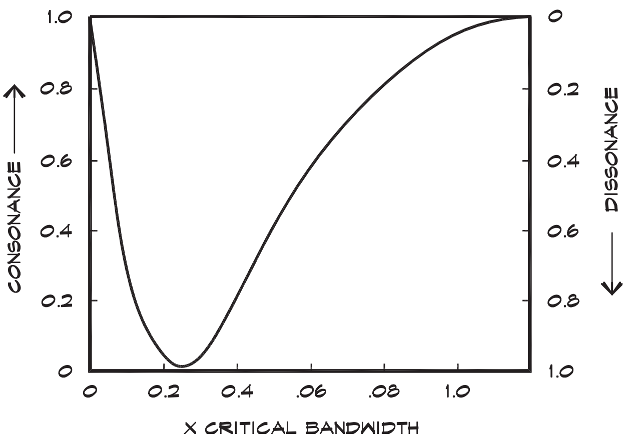


FIGURE 3.6 Consonance and Dissonance as a Function of the Critical Bandwidth (Pierce, 1983)



of the difference frequency, shown in terms of the critical bandwidth. When two tones are very close together the difference frequency is too small to be detected as dissonance but is rather a *tremolo*, a rising and falling of level. As they move apart the two tones interfere in such a way as to produce a roughness. When the frequency difference increases still further, the tones begin to be sensed separately and smoothly. For all frequency differences greater than a critical band, separate tones sound consonant.

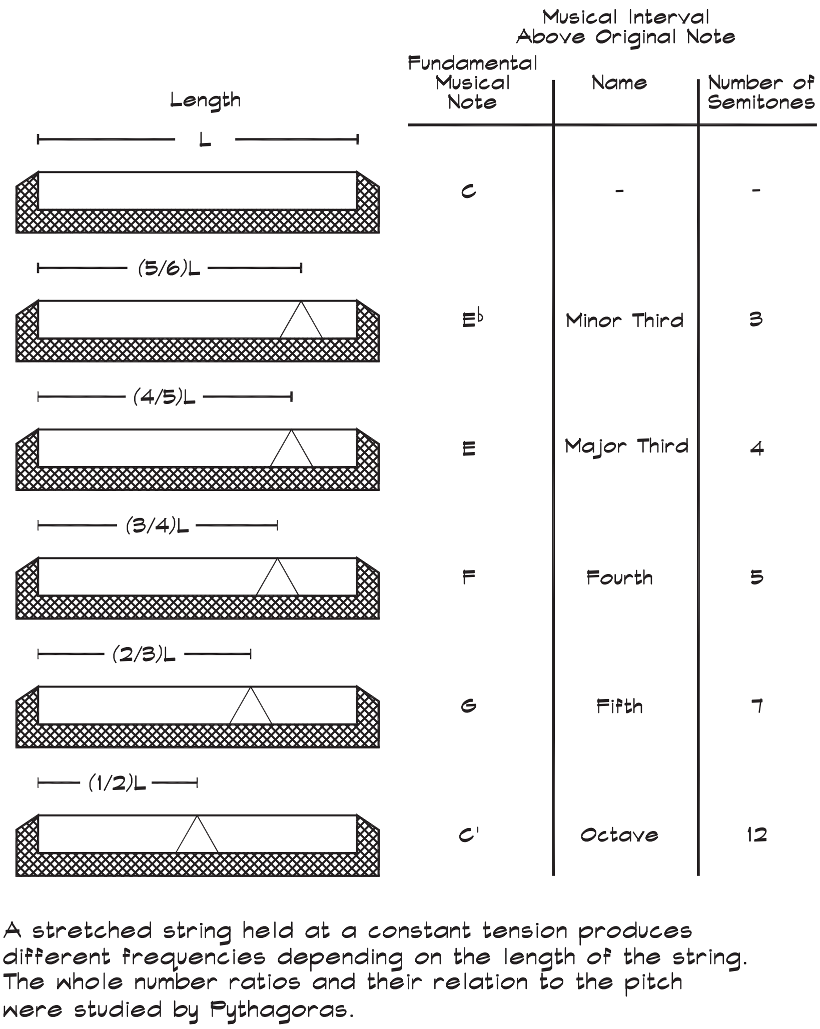
Tone Scales

One of the traditional problems of music is the establishment of a scale of notes, based on frequency intervals, that sound pleasant when played together. The most fundamental scale division is the octave or doubling of frequency, which is the basis for virtually all music. The octave has been variously divided up into 5 notes (pentatonic scale), 7 notes (diatonic scale), or 12 notes (chromatic scale) by different cultures. Western music uses 12 intervals called *semitones*, but a particular piece usually employs a group of seven selected notes, which are designated as a particular key, bearing the name of the starting note. These scales are called major or minor depending on the order of the single or double steps between the notes selected.

Since many of the early instruments were stringed, not only would the player hear the note's fundamental pitch but also its harmonics, which are integer multiples of the fundamental. Note that the second harmonic frequency is twice the fundamental, the third three times and so forth. The first harmonic is the fundamental. Overtones are sometimes taken to mean the same thing as harmonics. In this work overtones are any significant tonal component of the spectrum of a note whether or not this tone has a harmonic relationship to the note's fundamental. It is not uncommon, particularly in percussion instruments such as chimes, to find nonharmonic overtones, some of which change in frequency as they decay. This group of sounds constitutes a musical instrument's spectrum or timbre, the particular character it has that distinguishes its sound from other similar sounds.

Pythagoras of Samos (sixth century BC) is credited with the discovery that when a string is bridged, with one segment having a certain fractional ratio to the overall length, namely 1/2, 2/3, 3/4, 4/5, and 5/6, the two notes had a pleasing sound. These ratios are called the

FIGURE 3.7 Pythagorean Pitch Intervals (Pierce,1983)



perfect intervals in music and traditionally are given special names based on the number of diatonic intervals between them. Figure 3.7 shows the ratios and how they may be obtained from a stretched string and how they relate to the notes we use today. The 2/3 ratio is called a fifth since it has 5 diatonic intervals, the 3/4 ratio is a fourth, 4/5 a major third, and 5/6 a minor third.

Having established the first two consonant intervals, several systems were used to fill in the remaining notes to create a musical scale. One was the just scale, which set all note intervals to small integer ratios. Another was the Pythagorean scale, which sought to produce the most equal whole number ratio intervals. Finally was the *equal-tempered scale*, which set the steps between notes to the same ratio. Each system has some problem. The first two do not transpose well to another key, that is, they do not sound the same since the notes have different relationships. The equal-tempered scale introduced about 300 years ago abandoned adherence to perfect integer ratios but chose intervals that did not differ significantly from those ratios. In this scheme, advocated by J. S. Bach, each note is separated from the following one by a factor of $\sqrt[12]{2} = 1.059463$, called a semitone. Every note interval, is in turn, divided into 100 cents so that there are 1200 cents in an octave. The frequency ratio between each cent is the 1200th root of 2, or 1.00057779. In this system a scale may begin on any white or black key on the piano and sound alike. Bach wrote his series, *The Well Tempered Clavier*, which contains pieces in all major and minor keys, in part to illustrate this method of tuning.

Once the system of note intervals had been established, there was the problem of choosing where to begin. For many years there were no pitch standards and, according to Helmholtz (1877), pipe organs were built with A's ranging from 374 Hz to 567 Hz. Handel's tuning fork was measured at 422.5 Hz and that became the standard for the classical composers Haydn, Bach, Mozart, and Beethoven. In 1859 the standard A rose to 435 Hz, set by a French government commission, which included Berlioz, Meyerbeer, and Rossini. The so-called "scientific" pitch was introduced in the early twentieth century with the C note frequencies being integer powers of 2, much like the designation of today's computer memory chips. This resulted in an A of 431 Hz. Later, in 1939, an international conference in London adopted the current standard, A equal to 440 Hz at 20° C.

Tunings still vary with individual instruments and musical taste. The natural frequency in woodwinds rises about 3 cents per degree C due to the increase in the velocity of sound. In stringed instruments the fundamental frequency falls slightly with temperature due to the thermal expansion of the strings. Some musicians raise the pitch of their tunings to get additional edge or brightness. This is an unfortunate trend since it stresses older instruments and adds a shrillness to the music, particularly in the strings.

Pitch

Pitch is the human perception of how high or low a tone sounds, based on its relative position on a scale. Musical pitch is defined in terms of notes however, there are psychoacoustical experiments to measure human perception of relative pitch as well. Absolute pitch discrimination is rather rare occurring in only 0.01 percent of the population (Rossing, 1990). Relative pitch discrimination can be measured by asking subjects to respond when one tone sounds twice as high as another. Like loudness experiments, the results are complex, for while they depend primarily on frequency, they also can vary with intensity and waveform. For example if a 100 Hz tone is sounded at 60 dB and then at 80 dB, the louder sound will be perceived as having the lower pitch. This phenomenon is primarily one that occurs at frequencies below 300 Hz. At the mid frequencies (500 Hz to 3000 Hz), pitch is relatively independent of intensity, whereas at frequencies above 4000 Hz, pitch increases with level.

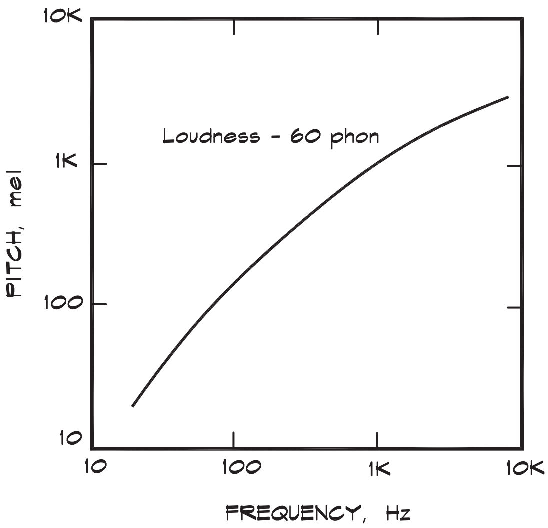
Pitch, as measured in these types of experiments, is different from harmonic relationships found in music. The former is expressed in units of mels, which are similar to sones in that a tone having 2000 mels is judged to be twice as high as one with 1000 mels. The reference frequency for a tone at a given loudness is 1000 Hz, which is defined as 1000 mels. For a given loudness it is possible to define a curve of constant pitch versus frequency as in Fig. 3.8.

3.3 LOUDNESS

Comparative Loudness

Loudness is the human perception of the magnitude of a sound. Early efforts to quantify loudness were undertaken in the field of music. The terms "very loud," "loud," "moderately loud," "soft," and "very soft" were given symbols in musical notation—ff, f, mf, p, and pp, after these words in Italian. But the terms are not sufficiently precise for scientific use,

FIGURE 3.8 Relative Pitch Discrimination vs Frequency (Kinsler et al., 1982)



and depend on the hearing acuity and custom of the person using them. It would seem straightforward to use the measured intensity of a sound to determine its loudness, but unfortunately no such simple relationship exists. Loudness is ultimately dependent on the number of nerve impulses reaching the brain in a given time, but since those impulses come from different regions of the cochlea there is also a variation with the frequency content of the sound.

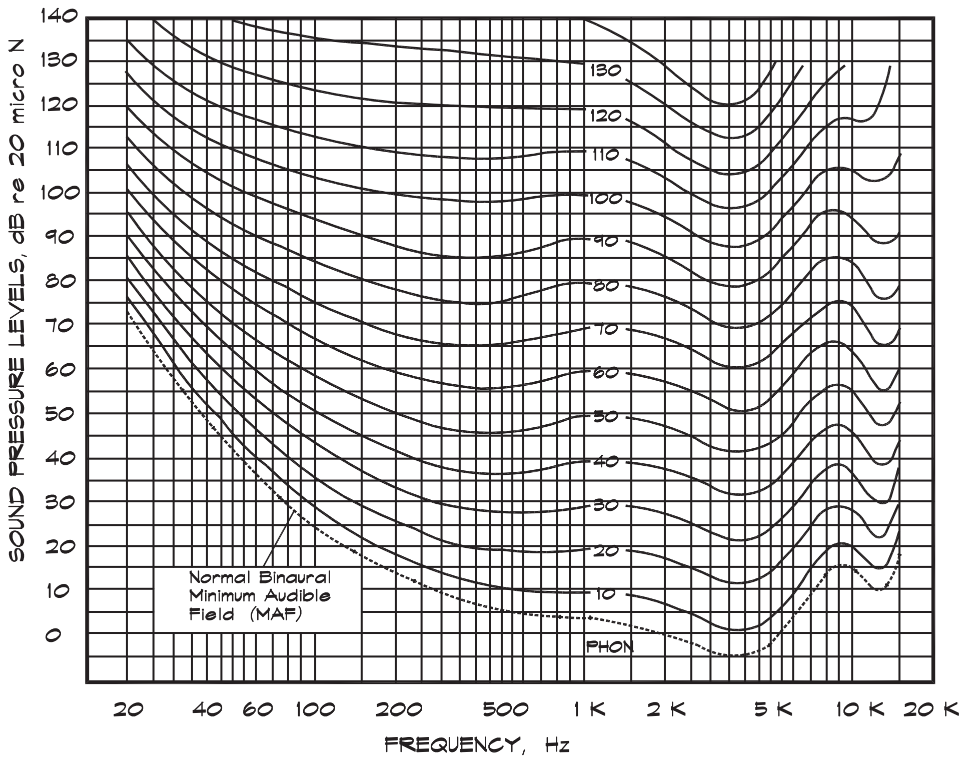
Even when the same sound signal is heard at differing intensities, there will be some variability from listener to listener and indeed even variation with the same listener, depending on his psychological and physiological state. Of general interest is the expected reaction of a listener in a typical environment, which can be determined by testing a number of subjects under controlled conditions. The average of an ensemble of listeners is taken as the result expected from a typical listener, a premise known as the *ergodic* hypothesis in statistics.

Loudness Levels

Comparative loudness measurements were made in the 1920's and 30's by scientists at Bell Laboratories. These tests were done on a group of subjects with acute hearing by presenting them with a controlled set of sounds. Various signals were used, however, in the classic study by Fletcher and Munson, published in 1933, pure tones (sine waves) of short duration were employed. The procedure was to compare the loudness of a tone, presented to the listener at a particular frequency and amplitude, to a fixed reference tone at 1000 Hz having an amplitude that was set in 10 dB intervals to between 0 and 120 dB. Tones were presented to the listeners, by means of headphones, for a one-second duration with a 1.5 second pause in between. Subjects were asked to choose whether a given tone was above, below, or equal to the reference tone, which resulted in a group of loudness-level contours known as the *Fletcher-Munson curves*.

In 1956 Robinson and Dadson repeated the Fletcher-Munson measurements, this time using loudspeakers in an anechoic chamber. The resulting *Robinson-Dadson* curves are shown in Fig. 3.9. The lowest of these curves is the threshold of hearing, which passes through 0 dB at about 2000 Hz and drops below this level at 4000 Hz, where the ear is most sensitive.

FIGURE 3.9 Normal Loudness Contours for Pure Tones (Robinson and Dadson, 1956)



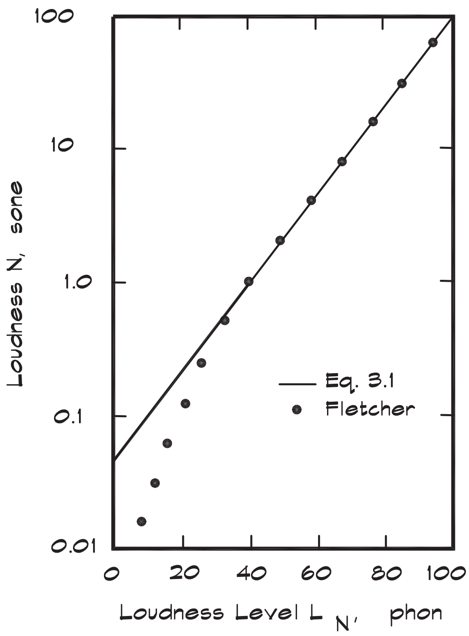
The graph shows that human hearing is significantly less sensitive to low-frequency sounds. At lower frequencies the level rises rapidly as the frequency decreases until, at 30 Hz, the intensity level must be about 65 dB before it is audible. As the intensity of a sound increases, the ear's frequency response becomes flatter. Near 100 dB the ear's response is almost flat except for a small increase in sensitivity around 4000 Hz. These experiments have been repeated many times since the original work was done with various types of signals. They have also been performed on the general population to determine the behavior of hearing acuity with age and other factors.

The curves in Fig. 3.9 are called equal-loudness contours. For any two points along one of these curves the perceived loudness of tones is the same. A loudness level (having units of phons) is assigned to each curve, which is numerically equal to the intensity level of the tone at 1000 Hz. Thus if we follow the 40 phon line we see that a tone having an intensity level of 50 dB at 100 Hz falls on the line and thus has a loudness level of 40 phons. At 1000 Hz the loudness level is equal to the intensity level. At 10,000 Hz the intensity level on the 40 phon line is about 46 dB.

Relative Loudness

The loudness level contours are based on human judgments of absolute loudness. The question asked each subject is whether the tone is louder or quieter than the reference. Another question can be asked, based on a relative comparison, namely, "When is it twice as loud?"

FIGURE 3.10 Loudness vs Loudness Level (Fletcher, 1933)



This gives rise to a measure of relative loudness having units of sones. In this scheme the loudness metric is linear; a sound having a relative loudness of 2 sones is twice as loud as a sound of 1 sone and so forth. The baseline is the 40 phon curve that is given the value of 1 sone. Figure 3.10 shows the relationship between (relative) loudness in sones and loudness level in phons. The result of these measurements is that loudness doubles every 9 phons or about 9 dB at the mid frequencies. A general equation can be written for the relationship between loudness and loudness level of pure tones in the linear region of the curve as follows (Kinsler and Frey, 1962)

$$L_N \cong 30 \log N + 40 \tag{3.1}$$

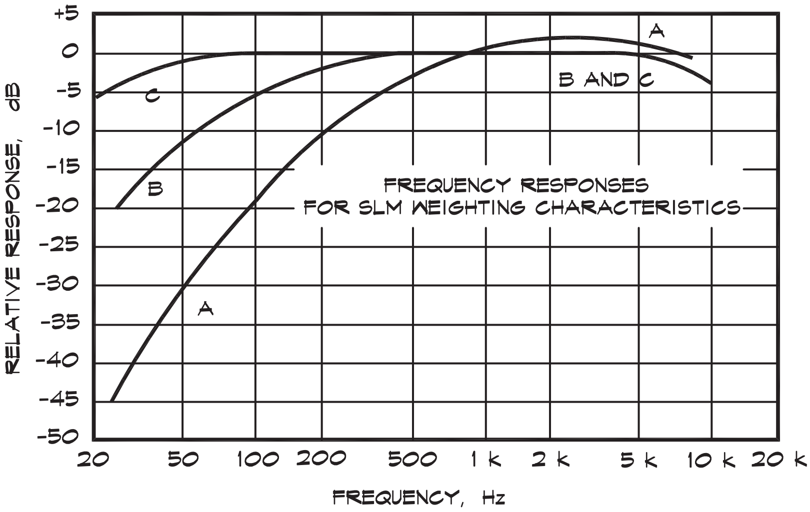
where N = loudness (sones)
L_N = loudness level (phons)

A determination can also be made of the loudness of sounds, having a spectral content more complicated than pure tones. Starting with measured levels in one-third-octave or full-octave bandwidths, various schemes have been proposed by Kryter, Stevens, and Zwicker for the calculation of both loudness and loudness levels. While these are generally more successful than simple electronic filtering in correlating with the human perception of loudness, their complexity limits their usefulness.

Electrical Weighting Networks

Although the Fletcher Munson curves provided an accurate measure of the relative loudness of tones, their shape was too complicated for use with an analog sound level meter. To overcome this problem, electrical weighting networks or filters were developed, which approximate the Fletcher Munson curves. These frequency weighting schemes were designated by letters of the alphabet, A, B, C, and so forth. The A, B, and C-weighting networks were designed

FIGURE 3.11 Frequency Response Characteristics in the American National Standards Specification for Sound Level Meters (ANSI-S1.4 – 1971)



to roughly mirror the 40, 60, and 80 phon lines (turned upside down) shown in Fig. 3.11. The relative weightings in each third-octave band for the A and C filters are set forth in Table 3.1. Since the time of their original development other weighting curves have been suggested. Several D-weighting curves are detailed by Kryter (1970) and an E curve has been suggested, but these systems, along with B-weighting, have not found widespread acceptance.

Only the A and C curves are still in general use. The A-weighted level (dBA) is the most common single number measure of loudness. The C-weighting network is used mainly as a measure of the broadband sound pressure level. Occasionally an $(L_C - L_A)$ level is used to describe the relative contribution of low-frequency noise to a spectrum. The weightings can be applied to sound power or sound pressure levels.

Once a weighting is applied it should not be reapplied. For example, if a recording is made of environmental noise using an A-weighting filter (not a recommended practice) it should not be analyzed by playing it back through a meter using the A-weighting network. It is, however, not uncommon to use a C-weighting filter when recording environmental noise, since it limits the low-frequency sounds that can overload the tape recorder. Replaying a tape thus recorded back through an A-weighting filter would be a reasonable practice so long as the main frequencies of interest were those unaffected by the C-weighting.

It is not uncommon to encounter A-weighted octave or third-octave band levels. These levels, if appropriately designated, are understood to have had the weighting already applied. A-weighted octave-band levels can be calculated from third-octave levels using Eq. 2.62 to combine the three levels within a particular octave band. Likewise overall A-weighted levels can be calculated from A-weighted octave-band levels by using the same formula.

Noise Criteria Curves (NC and RC)

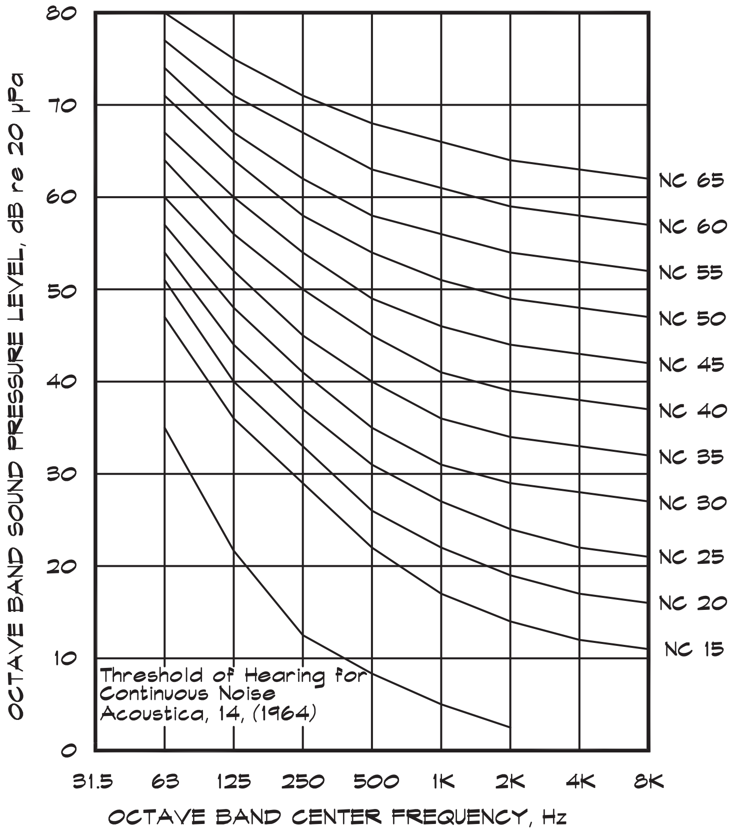
Loudness curves based on octave-band sound pressure level measurements are commonly used in buildings to establish standards for various types of activities. The *noise criterion* (NC) curves shown in Fig. 3.12 were developed by Beranek in 1957 to establish satisfactory

TABLE 3.1 Electrical Weighting Networks

Frequency Hz	A-Weighting Relative Response, dB	C-Weighting Relative Response, dB
12.5	−63.4	−11.2
16	**−56.7**	**−8.5**
20	−50.5	−6.2
25	−44.7	−4.4
31.5	**−39.4**	**−3.0**
40	−34.6	−2.0
50	−30.2	−1.3
63	**−26.2**	**−0.8**
80	−22.5	−0.5
100	−19.1	−0.3
125	**−16.1**	**−0.2**
160	−13.4	−0.1
200	−10.9	0
250	**−8.6**	**0**
315	−6.6	0
400	−4.8	0
500	**−3.2**	**0**
630	−1.9	0
800	−0.8	0
1,000	**0**	**0**
1,250	+0.6	0
1,600	+1.0	−0.1
2,000	**+1.2**	**−0.2**
2,500	+1.3	−0.3
3,150	+1.2	−0.5
4,000	**+1.0**	**−0.8**
5,000	+0.5	−1.3
6,300	−0.1	−2.0
8,000	**−1.1**	**−3.0**
10,000	−2.5	−4.4
12,500	−4.3	−6.2
16,000	**−6.6**	**−8.5**
20,000	−9.3	−11.2

conditions for speech intelligibility and general living environments. They are expressed as a series of curves, which are designated NC-30, NC-35, and so on, according to where the curve crossed the 1750 Hz frequency line in an old (now obsolete) octave-band designation. The NC level is determined from the lowest NC curve, which may be drawn such that no point on a measured octave-band spectrum lies above it. Since the NC curves are defined in 5 dB intervals, in between these values the NC level is interpolated.

FIGURE 3.12 Noise Criterion (NC) Curves (Beranek, 1957)



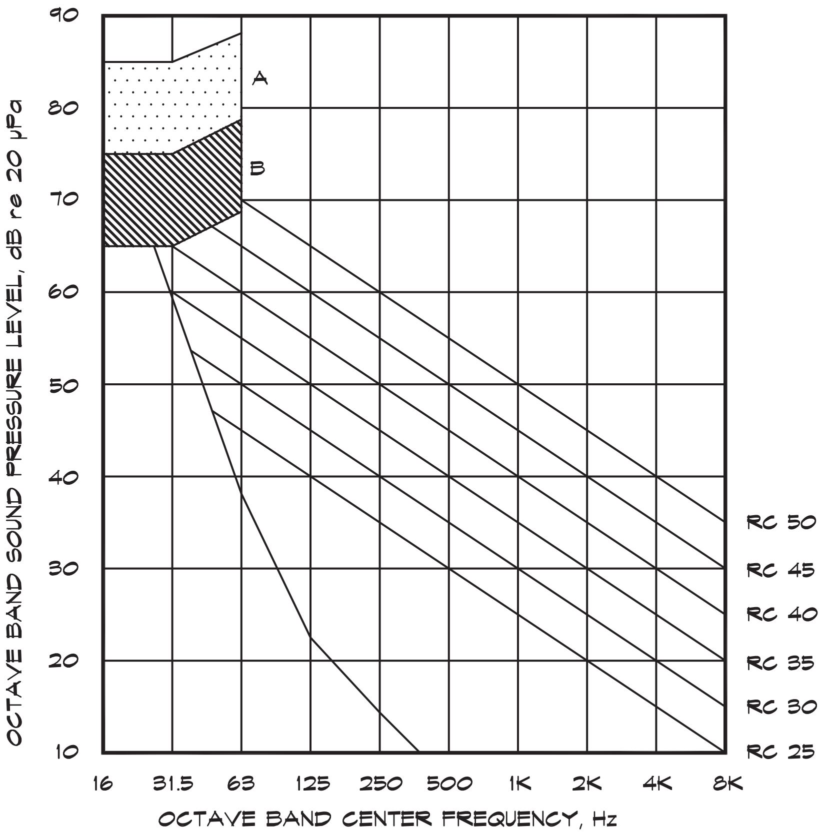
The NC level depends on the actual measured (or calculated) spectrum of the sound but they can be generally related to an overall A-weighted level (Kinsler et al., 1982)

$$NC \cong 1.25\,(L_A - 13) \tag{3.2}$$

where NC = NC level, dB
L_A = sound pressure level, dBA

In 1981 Blazier developed the set of curves, shown in Fig. 3.13, called *room criterion* (RC) curves, based on an American Society of Heating, Refrigeration, and Air Conditioning Engineers (ASHRAE) study of heating, ventilating, and air conditioning (HVAC) noise in office environments. Since these curves are straight lines, spaced 5 dB apart, there is no confusion about the level, which can occur sometimes with NC levels. RC curves are more stringent at low frequencies but include 5 dB of leeway in the computation methodology. The RC level is the arithmetic average of the 500, 1000, and 2000 Hz octave-band values taken from the measured spectrum. At frequencies above and below these center bands a second parallel line is drawn. Below 500 Hz the line is 5 dB above the corresponding RC line and above 2000 Hz it is 3 dB above the line. If the measured spectrum exceeds the low-frequency line the RC level is given the designation R for rumble. If it exceeds the high-frequency line the designation is H for hissy. Otherwise the designation is N for normal.

FIGURE 3.13 Room Criterion (RC) Curves (Blazier, 1981)



Blazier added two other regions to these curves at the low frequencies where mechanical vibrations can be a nuisance in lightweight structures. In region A there is a high probability of noticeable vibrations accompanying the noise while in region B there is a low probability of that occurring, particularly in the lower portion of the curve.

From time to time ASHRAE publishes a guide to assist in the calculation and treatment of HVAC generated noise levels. As part of this guide they include suggested levels of noise for various classifications of interior space. In general these guidelines are applied to noise generated by equipment associated with a building, such as HVAC systems within a dwelling or office, or noise generated in an adjacent room by HVAC, pumps, fans, or plumbing. The standards are not generally applied to noise generated by appliances, which plug into wall sockets within the same dwelling unit. A portion of the 1987 ASHRAE guidelines are shown in Table 3.2.

Just Noticeable Difference

One of the classic psychoacoustic experiments is the measurement of a *just noticeable difference* (jnd), which is also called a difference limen. In these tests a subject is asked to compare two sounds and to indicate which is higher in level, or in frequency. What is found is that the jnd in level depends on both the intensity and frequency. Table 3.3 shows jnd level values at various sound pressure levels and frequencies. For sound levels exceeding 40 dB and at frequencies above 100 Hz, the jnd is less than 1 dB. At the most sensitive levels

TABLE 3.2 Interior Noise Design Goals (ASHRAE, 1987)

	Type of Area	Recommended NC or RC Criteria Range
1	Private Residences	25 to 30
2	Apartments	25 to 30
3	Hotels/motels	
	a Individual rooms or suites	30 to 35
	b Meeting/banquet rooms	25 to 30
	c Halls, corridors, lobbies	35 to 40
	d Service/support areas	40 to 45
4	Offices	
	a Executive	25 to 30
	b Conference room	25 to 30
	c Private	30 to 35
	d Open plan areas	35 to 40
	e Computer equipment rooms	40 to 45
	f Public circulation	40 to 45
5	Hospitals and clinics	
	a Private rooms	25 to 30
	b Wards	30 to 35
	c Operating rooms	35 to 40
	d Corridors	35 to 40
	e Public areas	35 to 40
6	Churches	25 to 30
7	Schools	
	a Lecture and classrooms	25 to 30
	b Open plan classrooms	30 to 35
8	Libraries	35 to 40
9	Concert halls	5 to 15*
10	Legitimate theaters	20 to 30
11	Recording studios	10 to 20*
12	Movie theaters	30 to 35

*Note: Where ASHRAE has recommended that an acoustical engineer be consulted, the author has supplied the NC or RC levels.

TABLE 3.3 Minimum Detectable Changes (JND) in Level for Sine Waves, dB (Pierce, 1983)

Freq.	Signal Level, dB											
Hz	5	10	20	30	40	50	60	70	80	90	100	110
35	9.3	7.8	4.3	1.8	1.8							
70	5.7	4.2	2.4	1.5	1.0	.75	.61	.57	1.0	1.0		
200	4.7	3.4	1.2	1.2	.86	.68	.53	.45	.41	.41		
1000	3.0	2.3	1.5	1.0	.72	.53	.41	.33	.29	.29	.25	.25
4000	2.5	1.7	.97	.68	.49	.41	.29	.25	.25	.21	.21	
8000	4.0	2.8	1.5	.90	.68	.61	.53	.49	.45	.41		
10,000	4.7	3.3	1.7	1.1	.86	.75	.68	.61	.57			

TABLE 3.4 Minimum Detectable Changes (JND) in Frequency for Sine Waves, Cents (Pierce, 1983)

Frequency	Signal Level, dB										
Hz	5	10	15	20	30	40	50	60	70	80	90
31	220	150	120	97	76	70					
62	120	120	94	85	80	74	61	60			
125	100	73	57	52	46	43	48	47			
250	61	37	27	22	19	18	17	17	17	17	
500	28	19	14	12	10	9	7	6	7		
1000	16	11	8	7	6	6	6	6	5	5	4
2000	14	6	5	4	3	3	3	3	3	3	
4000	10	8	7	5	5	4	4	4	4		
8000	11	9	8	7	6	5	4	4			
11,700	12	10	7	6	6	6	5				

(greater than 60 dB) and frequencies (1000–4000 Hz), the jnd is about a quarter of a dB. When the jnd is 0.25 dB it means that we can notice a sound, with the same spectrum, which is 13 dB below the level of the background. This has important implications for both privacy and intelligibility in the design of speech reinforcement systems.

The jnd values in frequency for sine waves are shown in Table 3.4. Like the jnd in level, it is also dependent on both intensity and frequency. At 2000 Hz, where we are most sensitive, the jnd is about 3 cents (0.3% of an octave) or about 0.5% of the pure tone frequency for levels above 30 dB. This is about 10 Hz. Some trained musicians can tell the difference between a perfect fifth (702 cents) and an equal tempered fifth (700 cents), so that greater sensitivity is not uncommon. Note that these comparisons are done by sounding successive tones or by varying the tone, not by comparing simultaneous tones where greater precision is obtainable by listening for beats. Piano tuners who tune by ear use this latter method to achieve precise tuning.

Environmental Impact

Environmental Impact Reports (EIR) in California or Environmental Impact Statements (EIS) for Federal projects are prepared when a proposed project has the potential of creating a significant adverse impact on the environment (California Environmental Quality Act, 1972). Noise is often one of the environmental effects generated by a development, through increases in traffic or fixed noise sources. Impact may be judged either on an absolute scale through comparison with a standard such as a property line ordinance or a Noise Element of a General Plan, or on a relative scale through changes in level. In the exterior environment the sensitivity to changes in noise level is not as great as in the laboratory under controlled conditions. A 1 dB change is the threshold for most people. Since the change in level due to multiple sound sources is equal to $10 \log N$, where N is the ratio of the new to the old number, it takes a 1.26 ratio or a 26% increase in traffic passing by on a street to produce a 1 dB change. Table 3.5 shows a general characterization of human reaction to changes in level.

TABLE 3.5 Human Reaction to Changes in Level

Change in Level (dB)	Reaction
1	Noticeable
3	Very Noticeable
6	Substantial
10	Doubling (or Halving)

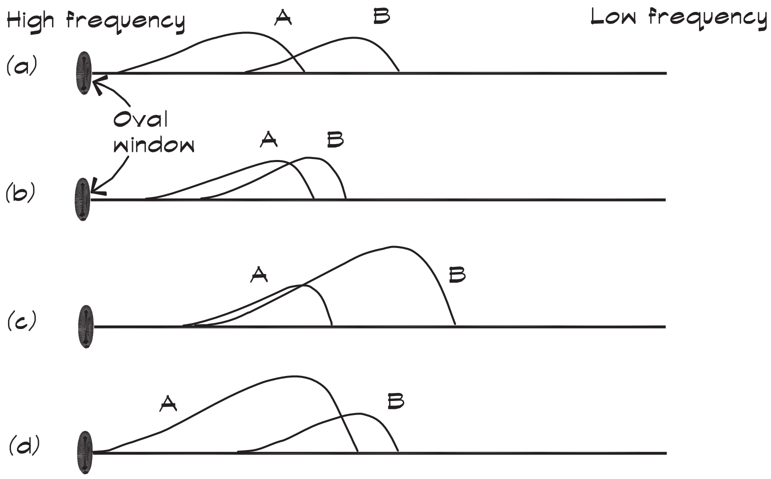
The characterizations listed in Table 3.5 are useful in gauging the reaction to changes in environmental noise. If a project increases the overall noise level at a given location by 1 dB or more, it is likely to be noticeable and could potentially constitute an adverse environmental impact. A 3 dB increase is very noticeable and, in the case of traffic flow, represents a doubling in traffic volume.

3.4 INTELLIGIBILITY

Masking

When we listen to two or more tones simultaneously, if their levels are sufficiently different, it becomes difficult to perceive the quieter tone. We say that the quieter sound is *masked* by the louder. Masking can be understood in terms of a threshold shift produced by the louder tone due to its overlap within the critical band on the cochlea. Figure 3.14 illustrates this principle. The figure is helpful in understanding the findings of experiments associated with

FIGURE 3.14 Overlap of Regions of the Cochlea (Rossing, 1990)

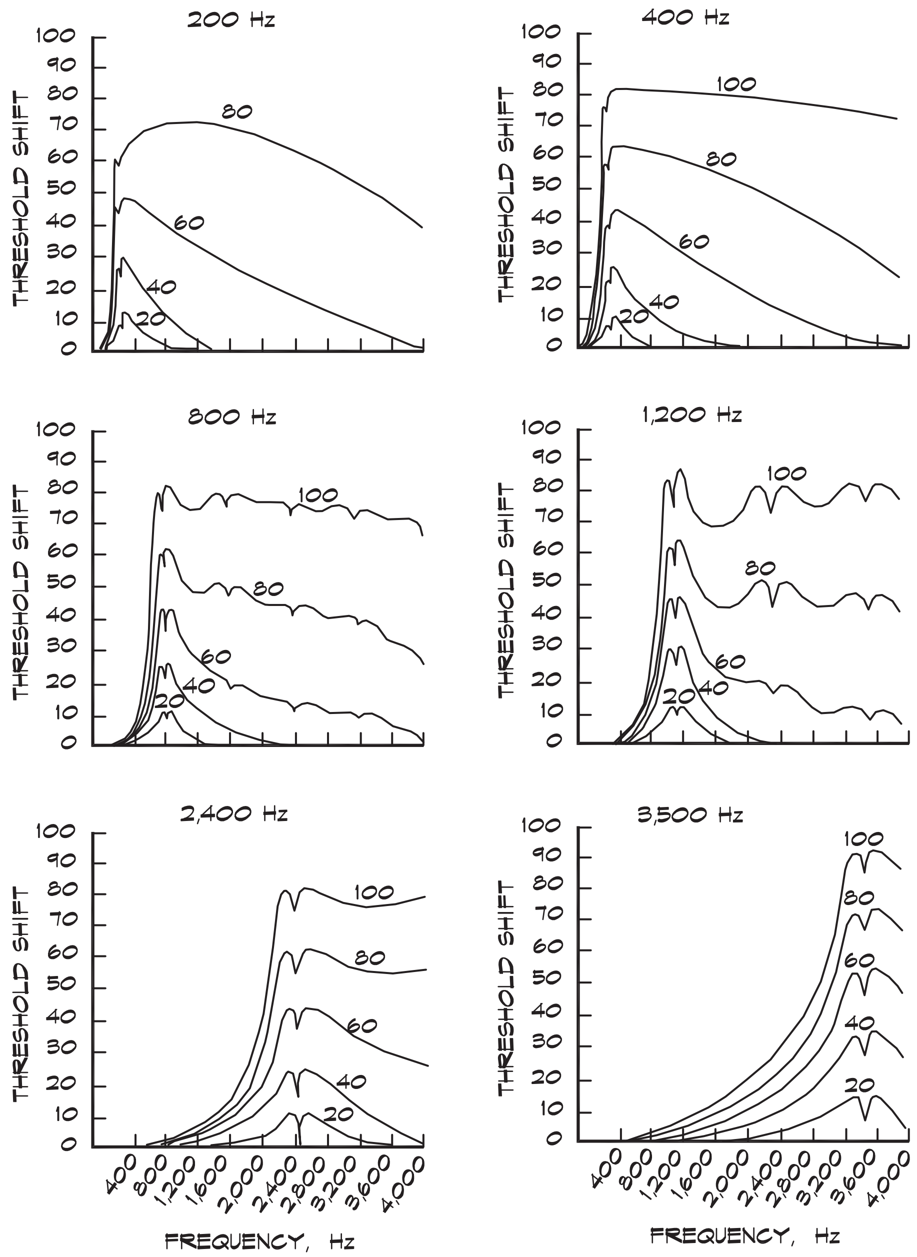


Simplified response of the basilar membrane for two pure tones A and B. (a) The excitations barely overlap; little masking occurs. (b) There is an appreciable overlap; tone B masks tone A and somewhat more than the reverse. (c) The more intense tone B almost completely masks the higher-frequency tone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

masking. Tones that are close in frequency mask each other more than those that are widely separated.

Tones mask upward in frequency rather than downward. The louder the masking tone the wider the range of frequencies it can mask. Masking by narrow bands of noise mimics that of pure tones and broad bands of noise mask at all frequencies. Early experiments on masking the audibility of tones in the presence of noise were performed by Wegel and Lane (1924) and subsequently published by Fletcher (1953). Two tones are presented to each subject. The first is a constant masking tone at a given level and frequency. A second tone

FIGURE 3.15 Pure Tone Masking Curves (Fletcher, 1953)



is introduced at a selectable frequency and its level is reduced until it is no longer audible. Based on these types of tests a series of masking curves can be drawn, which are shown in Fig. 3.15. Each curve shows the threshold shift of the masked tone or the difference between its normal threshold of audibility and the new threshold in the presence of the masking tone. For example in the second curve the masking tone is at 400 Hz. At 80 dB it induces a 60 dB threshold shift in an 800 Hz tone. From Fig. 3.9 we can see that the threshold of tonal hearing at 800 Hz is 0 dB, so the level above which the 800 Hz tone is audible is 60 dB. The fine structure on the masking curves is interesting. Around the frequency of the masking tone there are little dips in threshold shift—which is to say the ear becomes more sensitive. These dips repeat at the harmonics of the masking frequency. The reason is that when the two frequencies coincide, beats are generated that alert us to the presence of the masked tone.

As the level of the masking tone is increased, the breadth of its influence increases. A 400 Hz masking tone at 100 dB is effective in swamping the ear's response to 4000 Hz tones, while a 40 or 60 dB masking tone does little at these frequencies. At low levels the bandwidth of masking effectiveness is close to the critical bandwidth. High-frequency masking tones have little or no effect on lower frequency tones. A 2400 Hz masking tone will not mask a 400 Hz tone no matter how loud it is. This graphically illustrates the effect of the shape of the cochlear filter skirts.

Masking experiments also can be used to define critical bands. In the standard masking test a tone is not audible in the presence of masking noise until its level exceeds a certain value. The masking noise can be configured to have an adjustable bandwidth and the tests can be repeated for various noise bandwidths. It has been found (Fletcher and Munson, 1937) that masking is independent of noise bandwidth until the bandwidth is reduced below a certain value. This led to a separate way of measuring critical bandwidths, which gave results similar to those achieved using consonance and dissonance.

Masking is an important consideration in architectural acoustics. It is of particular interest to an acoustician whether speech will be intelligible in the presence of noise. In large indoor facilities, such as air terminals or sports arenas, low-frequency reverberant noise can mask the intelligibility of speech. This can be partially treated by limiting the bandwidth of the sound system or by adding low-frequency absorption to the room. The former is less expensive but limits the range of uses. Multipurpose arenas, which are hockey rinks one day and rock venues the next, should have an acoustical environment that does not limit the uses of the space.

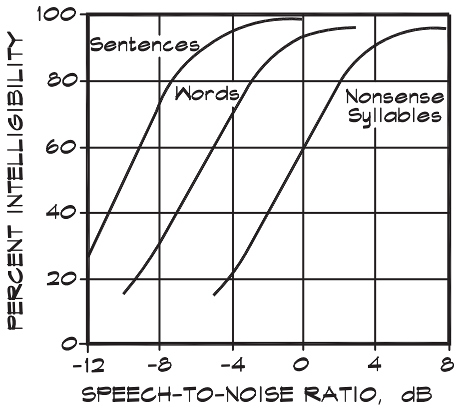
Speech Intelligibility

Speech intelligibility is a direct measure of the fraction of words or sentences understood by a listener. The most direct method of measuring intelligibility is to use sentences containing individual words or nonsense syllables, which are read to listeners who are asked to identify them. These can be presented at various levels in the presence of background noise or reverberation. Both live and recorded voices are used, however recorded voices are more consistent and controllable.

Three types of material are typically used: sentences, one syllable words, and nonsense syllables, with each type being increasingly more difficult to understand in the presence of noise. In sentence tests a passage is read from a text. In a word test individual words are read from a predetermined list, called a closed response word set, and subjects are asked to pick the correct one. A modified rhyme test uses 50 six-word groups of monosyllabic rhyming or similar-sounding English words. Subjects are asked to correctly identify the spoken word

FIGURE 3.16 Results of a Typical Intelligibility Test (Miller et al., 1951)

Intelligibility of different types of test materials in the presence of noise. Speech intelligibility is shown as a function of speech-to-noise ratio for sentences, monosyllabic words, and nonsense syllables.



from the list of six possible choices. A group of words might be: sag, sat, sass, sack, sad, and sap. Tests of this type lead to a measure of the fraction, ranging from 0 to 1, of words that are correctly identified. Figure 3.16 shows some typical results.

The degree to which noise inhibits intelligibility is dependent on the *signal-to-noise ratio*, which is simply the signal level minus the noise level in dB. When the noise is higher than the speech level, the signal-to-noise ratio is negative. A signal-to-noise ratio is a commonly used concept in acoustics, audio, and electrical engineering. It is called a ratio since it represents the energy of the signal divided by the energy of the noise, expressed in decibels. For a typical test the noise is broad band steady noise such as that produced by a waterfall.

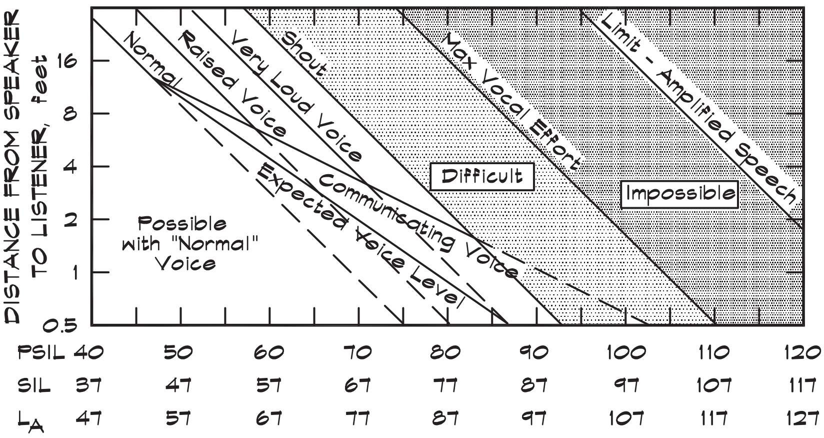
What is apparent from Fig. 3.16 is that even when the signal-to-noise is negative, speech is still intelligible. This is not surprising since the brain is an impressive computer, which can select useful information and fill in the gaps between the words we understand. For most applications if we can grasp more than 85–90% of the words being spoken we achieve very good comprehension—virtually 100% of the sentences. With an understanding of more than 60% of the words we can still get 90% of the sentences and that is quite good. If we understand fewer than 60% of the words the intelligibility drops off rapidly.

Speech Interference Level

Signal-to-noise ratio is the key to speech intelligibility, and we obtain more precise estimates of the potential interference by studying the background noise in the speech frequency bands. The *speech interference level* (SIL) is a measure of a background noise's potential to mask speech. It is calculated by arithmetically averaging separate background noise levels in the four speech octave bands, namely 500, 1000, 2000, and 4000 Hz. The SIL can then be compared to the expected speech sound pressure level to obtain a relevant speech to noise ratio.

Figure 3.17 shows the expected distance over which just-reliable communications can be maintained for various speech interference level values. The graph accounts for the

FIGURE 3.17 Rating Chart for Determining Speech Communication (Webster, 1969)



expected rise in voice level, which occurs in the presence of high background noise. Note that these types of analyses assume a flat spectrum of background noise, constant speech levels, and non-reverberant spaces.

The *preferred speech interference level* (PSIL) is another similar metric, calculated from the arithmetic average of background noise levels in the 500, 1000, and 2000 Hz octave bands. The PSIL can also be used to obtain estimates of the intelligibility of speech in the presence of noise.

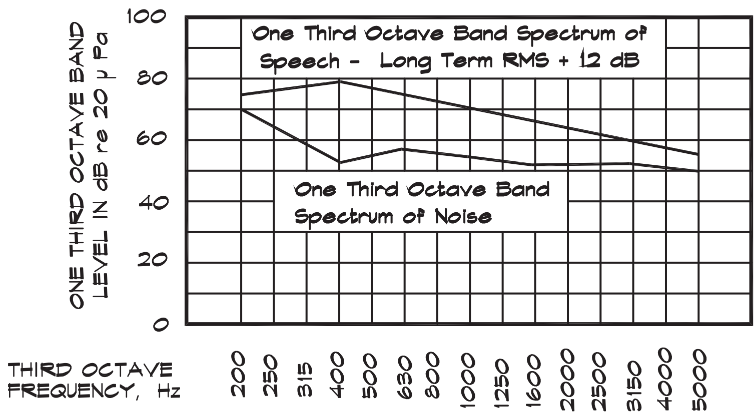
Articulation Index

The *articulation index* (AI) is a detailed method of measuring and calculating speech intelligibility (French and Steinberg, 1947). To measure the AI, a group of listeners is presented a series of phonemes for identification. Each of the test sounds consists of a logatom, or structured nonsense syllable in the form of *a consonant-vowel-consonant* (CVC) group embedded in a neutral carrier sentence, which cannot be recognized from its context in the sentence. An example might be, "Now try pom." The fraction of syllables understood is the AI.

Developed by researchers at Bell Laboratories in the late 1920s and early 1930s, including Fletcher, French, Steinberg, and others, it also included a method of calculating the expected speech intelligibility by using signal-to-noise ratios in third-octave bands, which are then weighted according to their importance. In this method speech intelligibility is proportional to the long-term rms speech signal plus 12 dB minus the noise in each band. The proportionality holds provided the sum of the terms falls between 0 and 30 dB. Figure 3.18 shows the calculation method along with the weighting factors used in each band. In each of 15 third-octave bands the signal-to-noise ratio is multiplied by a factor and the results are added together.

AI calculations can be made even when the spectrum of the background noise is not flat and is different from that of speech. It also accounts in part for the masking of speech by low-frequency noise. AI uses the peak levels generated by speech as the signal level and the energy average background levels as the noise. Consequently the signal-to-noise ratios are somewhat higher than other methods, such as SIL, which are based on average speech levels for the same conditions.

FIGURE 3.18 An Articulation Index Calculation (Kryter, 1970)



BAND	SPEECH PEAKS MINUS NOISE (dB)	WEIGHT	COLUMN 2 x 3
200	5	0.0004	0.0020
250	12	0.0010	0.0120
315	18	0.0010	0.0180
400	26	0.0014	0.0364
500	23	0.0014	0.0322
630	18	0.0020	0.0360
800	17	0.0020	0.0340
1000	16	0.0024	0.0384
1250	15	0.0030	0.0450
1600	14	0.0037	0.0518
2000	12	0.0037	0.0444
2500	10	0.0034	0.0340
3150	8	0.0034	0.0272
4000	6	0.0024	0.0144
5000	5	0.0020	0.0100
			AI = 0.4358

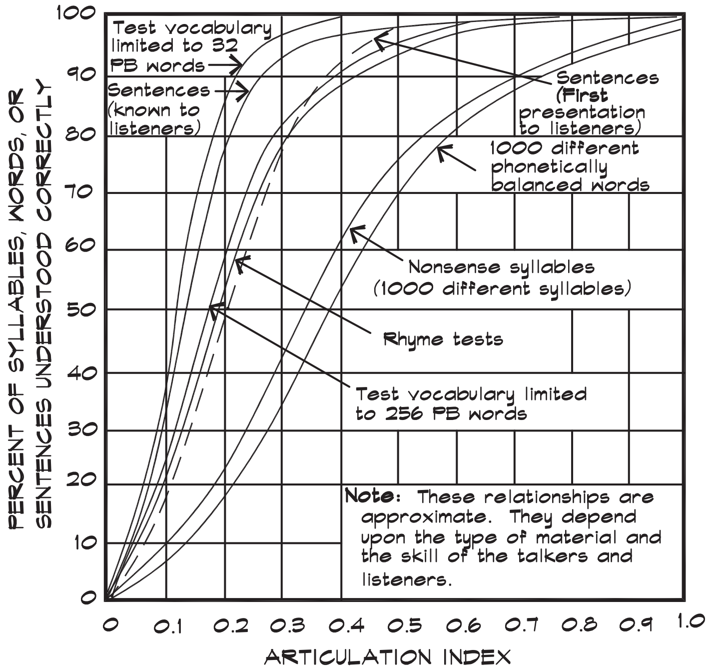
A typical AI calculation using the data shown in the graph. The one-third octave band spectrum of speech is used plus 12 dB and the noise spectrum is subtracted in each band. The result is multiplied by a weighting factor according to the band and the result is summed to obtain an overall Articulation Index figure.

The result of an AI calculation is a numerical factor, which ranges from 0 to 1, with 1 being 100% word or sentence comprehension. Beranek (1947) suggested that a listening environment with an AI of less than 0.3 will be found to be unsatisfactory or marginally satisfactory, while AI values between 0.3 to 0.5 will generally be acceptable. For AI values of 0.5 to 0.7 intelligibility will be good, and above 0.7 intelligibility will be very good to excellent. Figure 3.19 shows the relation between the AI and other measures of speech intelligibility.

AL_{CONS}

The *articulation loss of consonants* (AL_{CONS}), expressed as a percentage, is another way of characterizing the intelligibility of speech. Similar to the articulation index, it measures the proportion of consonants wrongly understood. V. Peutz also found that the correlation

FIGURE 3.19 Relation between AI and Other Speech Intelligibility Tests (ANSI S3.5, 1969)



between the loss of consonants (in Dutch) is much more reliable than a similar test with vowels. He published (1971) a relationship to predict intelligibility for unamplified speech in rooms, which had been studied much earlier at Bell Labs.

$$AL_{CONS} = \frac{200\ r^2\ T_{60}^2}{V} \tag{3.3}$$

Beyond the limiting distance $r_\ell = 0.21\sqrt{V/T_{60}}$

$$AL_{CONS} = 9\ T_{60} \tag{3.4}$$

where T_{60} = reverberation time, (s)
V = room volume, (m^3)
r = talker to listener distance, (m)

Privacy

The inverse of intelligibility is *privacy*, and articulation index is equally useful in the calculation of privacy as it was for intelligibility. Both are ultimately dependent on signal-to-noise ratio. Chanaud (1983) defined five levels of privacy, which are shown in Table 3.6, and has related them to AI in Fig. 3.20.

TABLE 3.6 Degrees of Acoustical Privacy (Chanaud, 1983)

Degree of Privacy	Acoustical Condition	Possible Subjective Response
Confidential Privacy	Cannot converse with others. Cannot understand speech of others. May not be aware of presence of others. May not hear activity sounds of others. Confidential conversations possible. No distractions.	Complete privacy. Sense of isolation. No privacy complaints expected.
Normal Privacy	Difficult to converse with others. Occasionally hear the activity sounds of others. Aware of the presence of others. Speech and machines audible but not distracting. Confidential conversations possible only under special conditions.	Sense of privacy. Some isolation. No privacy complaints expected.
Marginal Privacy	Possible to converse with others by raising voice. Often hear activity sounds and speech of others. Aware of each others presence. Conversations of others occasionally understood.	Sense of community. Sense of privacy weakened. Some privacy complaints expected.
Poor Privacy	Possible to converse with others at normal voice levels. Activity sounds, speech, and machines will be continually heard. Continually aware of each others presence. Frequent distractions.	Sense of community. Loss of privacy. Some loss of territory. Privacy complaints expected.
No Privacy	Easy to converse with others. Machine and activity sounds clearly audible. Total distraction from other tasks.	Sense of community. Sense of intrusion on territory. No sense of privacy. Many privacy complaints expected.

3.5 ANNOYANCE

Noisiness

Comparative systems, similar to those used in the judgment of loudness, have been developed to measure *noisiness*. Subjects were asked to compare third-octave bands of noise at differing levels based on a judgment of relative or absolute noisiness. Somewhat surprisingly, the results shown in Fig. 3.21 (Kryter, 1970) differ from a loudness comparison. Relative noisiness is described by a unit called noys, a scale that is linear in noisiness much like the sone is for loudness. It is converted into a decibel-like scale called *perceived noise level* (PNL) with units of PNdB, by requiring the noisiness to double every 10 dB. The perceived noise level scale is used extensively in the evaluation of aircraft noise. The work of other investigators: Ollerhead (1968), Wells (1967), and Stevens (1961) is also shown in the figure.

Noisiness is affected by a number of factors that do not influence loudness (Kryter, 1970). Two that do affect loudness are the spectrum and the level. Others that do not

FIGURE 3.20 Level of Privacy vs Articulation Index (Chanaud, 1983)

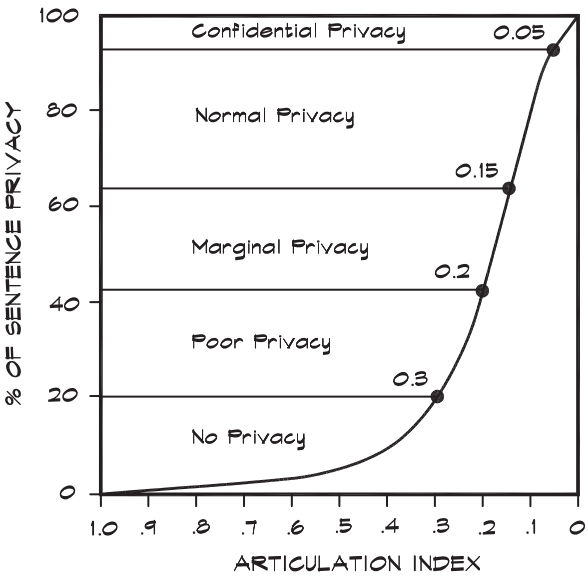
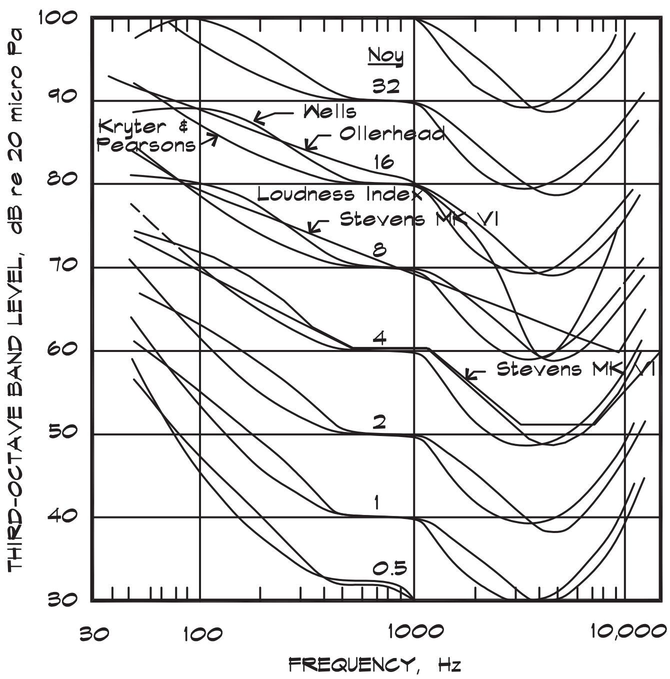


FIGURE 3.21 Equal Noisiness Contours of Various Authors (Kryter, 1970)



include: 1) spectrum complexity, namely the concentration of energy in pure tone or narrow frequency bands; 2) total duration; 3) in nonimpulsive sounds, the duration of the increase in level prior to the maximum level, called onset time; and 4) in impulsive sounds, the increase in level in a 0.5 second interval. Although these factors normally are not encountered in architectural acoustics, they contribute to various metrics in use in the United States and in Europe, particularly in the area of aircraft noise evaluation. For further details refer to Kryter (1970).

Time Averaging - L_{eq}

Since the duration of a sound can influence its perceived noisiness, schemes have been developed to account for the tradeoff between level and time. Some of these systems are stated implicitly as part of a particular metric, while others appear in noise standards such as those promulgated by the Occupational Safety and Health Administration (OSHA) or in various noise ordinances.

To a casual observer the simplest averaging scheme would appear to be the arithmetic average of measured levels over a given time period. The advantage of this type of metric is that it is simple to measure and is readily understandable to the layman. Two disadvantages of the arithmetic average are: 1) when there are large variations in level it does not accurately account for human reaction to noise, and 2) for doing prediction calculations on moving sources it is enormously cumbersome.

When a sound level varies in time it is convenient to have a single number descriptor, which accurately represents the effect of the temporal variation. In 1953, Rosenblith and K.N. Stephens suggested that a metric be developed, which included frequency weighting and a summation of noise energy over a 24-hour period. A number of metrics have since evolved that include some form of energy summation or energy averaging. The system most commonly encountered is the equivalent level or L_{eq} (pronounced ell-ee-q), defined as the steady A-weighted level that contains the same amount of energy as the actual time-varying A-weighted level during a given period. The L_{eq} can be thought of as an average sound pressure level, where the averaging is based on energy. To calculate the L_{eq} level from individual sound pressure level readings, an average is taken of the normalized intensity values, and that average is converted back into a level. Mathematically it is written as an integral over a time interval

$$L_{eq} = 10 \ \log \frac{1}{T} \int\limits_{t=0}^{T} 10^{0.1 \ L(t)} \tag{3.5}$$

or as a sum over equal-length periods

$$L_{eq} = 10 \ \log \frac{1}{N \, \Delta t} \sum_{i=1}^{N} 10^{0.1 \ L_i} \Delta t = 10 \ \log \frac{1}{N} \sum_{i=1}^{N} 10^{0.1 \ L_i} \tag{3.6}$$

where L_{eq} = equivalent sound level during the time period of interest (dB)
$L(t)$ = the continuous sound level as a function of time
L_i = an individual sample of the sound level, which is representative of the i th time period Δt

It has been found that human reaction to time-varying noise is quite accurately represented by the equivalent level. L_{eq} emphasizes the highest levels that occur during a given time period, even if they are very brief. For example it is clear that for a steady noise level that does not vary over a time period, the L_{eq} is the same as the average level L_{ave}. However if there is a loud noise, say 90 dBA for one second, and 30 dBA for 59 seconds, then the L_{eq} for the minute time period would be 72.2 dBA. The L_{ave} for the same scenario is 31 dBA. The L_{eq} is much more descriptive of the noise experienced during the period than the L_{ave}. When equivalent levels are used in environmental calculations, they are often based on a one-hour time period. When they begin and end on the hour they are called hourly noise levels, abbreviated HNL.

Twenty - Four Hour Metrics – L_{dn} *and CNEL*

One metric enjoying widespread acceptance is the L_{dn} or *day-night level*, which was recommended by the U.S. Environmental Protection Agency (EPA) for use in the characterization of environmental noise (von Gierke, 1973). The L_{dn}, or as it is sometimes abbreviated, the DNL, is a 24-hour L_{eq} with the noise occurring between the hours of 10 PM and 7 AM the next day increased by 10 dB before averaging. The L_{dn} is always A-weighted but may be measured using either fast or slow response.

$$L_{dn} = 10 \log \left\{ \frac{1}{24} \left[\sum_{i=8}^{22} 10^{0.1\, HNL_i} + (10) \sum_{i=23}^{7} 10^{0.1\, HNL_i} \right] \right\} \tag{3.7}$$

Each of the individual sample levels in Eq. 3.7 is for an hour-long period ending at the military time indicated by the subscript. The multiplier of 10, which is the same as adding 10 dB, accounts for our increased sensitivity to nighttime sounds.

Another system, the Community Noise Equivalent Level (CNEL), is in use in California and predates the day-night level. It is similar to the L_{dn} in that it is an energy average over 24 hours with a 10 dB nighttime penalty; however, it also includes an additional evening period from 7 P.M. to 10 P.M., with a multiplier of 3 (equal to adding 4.8 dB).

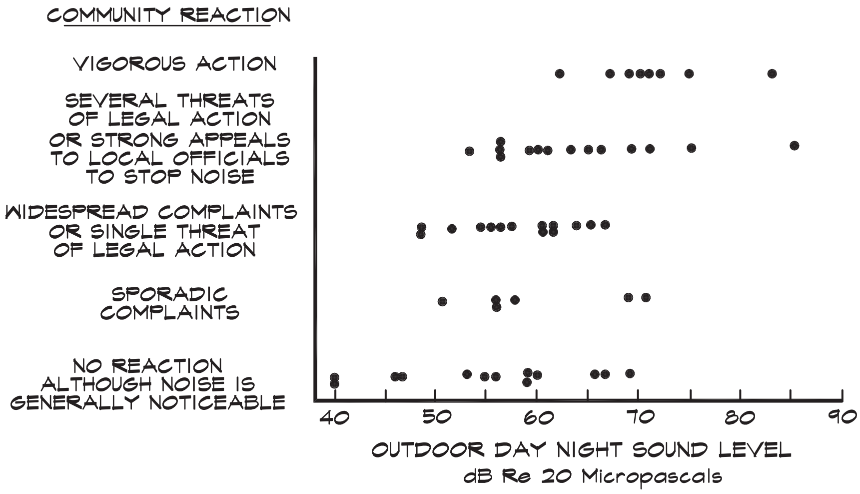
$$CNEL = 10 \log \left\{ \frac{1}{24} \left[\sum_{i=8}^{19} 10^{0.1\, HNL_i} + 3 \sum_{i=20}^{22} 10^{0.1\, HNL_i} + 10 \sum_{i=23}^{7} 10^{0.1\, HNL_i} \right] \right\} \tag{3.8}$$

Since the CNEL includes the extra evening factor it is always slightly higher than the L_{dn} level over the same time period. For most cases the two are essentially equal. Like the L_{dn}, the CNEL is A-weighted but is defined using the slow response.

Annoyance

The annoyance due to a sound can be highly personal. Any sound that is audible is potentially annoying to a given individual. Studies of annoyance have generally been based on the aggregate response of people exposed to various levels of noise. Much of the work in this field was done in the area of aircraft noise, and much of it is based on exterior noise levels. The U.S. EPA, following the mandate of the Noise Control Act of 1972, undertook a study of both the most appropriate metric to use for environmental noise and also the most appropriate levels. Since aircraft noise varies both from aircraft to aircraft and from day to day, the EPA study (von Gierke, 1973) recommended a 24-hour metric, namely the day-night level. They then developed recommendations on levels appropriate for public health and welfare (EPA Levels Document, 1974). Figure 3.22 shows part of the results of that study, specifically, the community reaction to exterior community noise of various types.

FIGURE 3.22 Community Reaction to Intrusive Noise (EPA Levels Document, 1974)



The data were very scattered. As a result of the wide variations in response an attempt was made to include factors other than level, duration, and time of day to normalize the results. A number of corrections were introduced, which were added to the raw day-night levels. These are listed in Table 3.7. Once the corrections had been applied, the data were replotted. The result is shown in Fig. 3.23, and the scatter is considerably reduced.

The final recommendations in the EPA Levels Document, for the levels of noise requisite to protect public health and welfare, are summarized in Table 3.8. It is interesting to note that the recommended exterior noise level of L_{dn} 55 does not guarantee satisfaction, and indeed according to the Levels Document it still leaves 17% of the population highly annoyed. Another interesting finding from this document, shown in Fig. 3.24, is that in one aircraft noise study, which related annoyance and complaints, the number of complaints lag well behind the number of people who are highly annoyed.

Satisfactory levels of interior noise are less well defined. Statutory limits in multifamily dwellings in California (CAC Title 24) are set at a CNEL (L_{dn}) 45 for noise emanating from outside the dwelling unit. It should be emphasized that statutory limits do not imply happiness. Rather they are the limits at which civil penalties are imposed. Many people are not happy with interior noise levels of L_{dn} 45 when the source of that noise is outside of their homes.

The 1987 ASHRAE guide suggests an NC 25 to 30 (30 to 35 dBA) as appropriate for residential and apartment dwellings. The EPA aircraft noise study (von Gierke, 1973) indicates that a nighttime level of 30 dBA in a bedroom would produce no arousal effects. Their recommendation of a maximum exterior L_{dn} of 60 dBA was based, in part, on a maximum interior level of 35 dBA at night with closed windows. The same reasoning, when applied to the Levels Document recommendations of L_{dn} 55 dBA, would yield a maximum nighttime level of 30 dBA with windows closed. Note that most residential structures provide about 20–25 dB of exterior to interior noise reduction with windows closed and about 10–15 dB with windows open. For purposes of this brief analysis maximum levels are taken to be 10 dB greater than the L_{eq} level. Van Houten (Harris, 1994) states that levels of plumbing related noise between 30 and 35 dBA in an adjacent unit can be "a source of

TABLE 3.7 Corrections to Be Added to the Measured Day-Night Sound Level (DNL) to Obtain the Normalized DNL (EPA Levels Document, 1974)

Type of Correction	Description	Correction (dBA)
Seasonal Correction	Summer (or year-round operation).	0
	Winter only (or windows always closed).	−5
Correction for Outdoor Noise Level Measured in Absence of Intruding Noise	Quiet suburban or rural community (remote from large cities and industrial activities).	+10
	Normal suburban community (not located near industrial activities).	+5
	Urban residential community (not located adjacent to heavily traveled roads or industrial activities).	0
	Noisy urban residential community (near relatively busy roads or industrial areas).	−5
	Very noisy urban residential community.	+10
Correction for Previous Exposure and Community Attitudes	No prior experience with intruding noise.	+5
	Community has had some previous exposure to intruding noise but little effort is being made to control the noise. This correction may also be applied in a situation where the community has not been exposed to the noise previously, but the people are aware that bona fide efforts are being made to control it.	0
	Community has had considerable previous exposure to intruding noise and the noise maker's relations with the community are good.	−5
	Community is aware that operation causing noise is very necessary and it will not continue indefinitely. This correction can be applied for an operation of limited duration and under emergency circumstances.	−10
Pure Tone or Impulse	No pure tone or impulsive character.	0
	Pure tone or impulsive character present.	+5

FIGURE 3.23 **Community Reaction to Intrusive Noise vs Normalized DNL Levels (EPA Levels Document, 1974)**

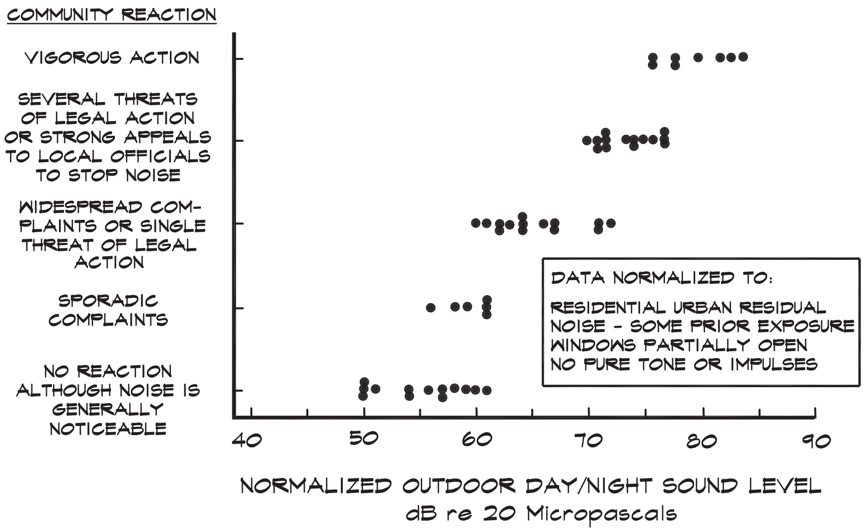
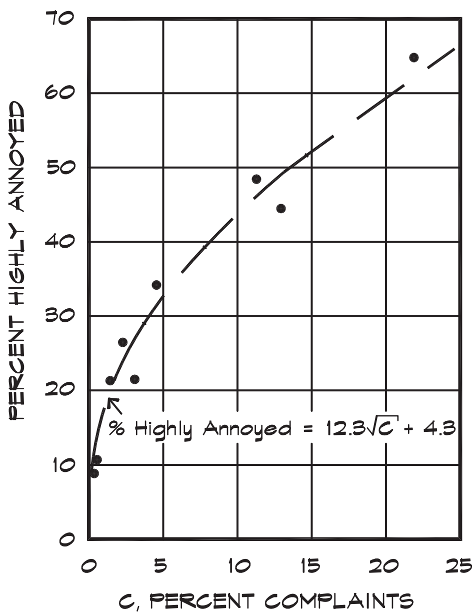


TABLE 3.8 **Summary of Noise Levels Identified as Requisite to Protect Public Health and Welfare with an Adequate Margin of Safety (EPA Levels Document 1974)**

EFFECT	LEVEL	AREA
Hearing loss	**$L_{eq(24)}$ < 70 dBA**	**All areas**
Outdoor activity interference and annoyance	**L_{dn} < 55 dBA**	**Outdoors in residential areas and farms and other outdoor areas where people spend widely varying amounts of time and other places in which quiet is a basis for use.**
	$L_{eq(24)}$ < 55 dBA	**Outdoor areas where people spend limited amounts of time, such as school yards, playgrounds, etc.**
Indoor activity interference and annoyance	**L_{eq} < 45 dBA**	**Indoor residential areas.**
	$L_{eq(24)}$ < 45 dBA	**Other indoor areas with human activities such as schools, etc.**

concern and embarrassment." In the author's practice in multifamily residential developments, intrusive levels generated by activities in another unit are rarely a problem below 25 dBA. At 30 dBA they are clearly noticeable and can be a source of annoyance, and above 35 dBA they frequently generate complaints.

FIGURE 3.24 Percentage Highly Annoyed as a Function of Percentage of Complaints (EPA Levels Document, 1974)



3.6 HEALTH AND SAFETY

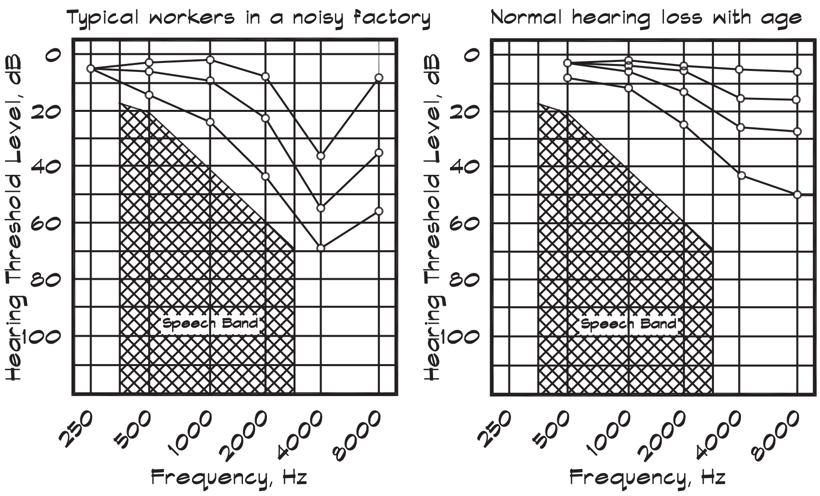
Hearing Loss

Noise levels above 120 dB produce physical pain in the human ear. The pain is caused by the ear's unsuccessful attempt to protect itself against sound levels about 80 dB above the auditory threshold by reducing its own sensitivity through use of the aural reflex. Exposure to loud noise damages the cochlear hair cells and a loss of hearing acuity results. If the exposure to noise is brief and is followed by a longer period of quiet, the hearing loss can be temporary. This phenomenon, called *temporary threshold shift* (TTS), is a common experience. The normal sounds we hear seem quieter after exposure to loud noises. If the sound persists for a long time at a high level or if there is repeated exposure over time, the ear does not return to its original threshold level and a condition called *permanent threshold shift* (PTS) occurs. The damage is done to the hair cells in the cochlea and is irreversible. The process is usually a gradual one that occurs at many frequencies, but predominantly at the upper end of the speech band. The loss progressively inhibits the ability to communicate as we age.

Human hearing varies considerably with age, particularly in its frequency response. In young people it is not uncommon to find an upper limit of 20 to 25 kHz, while in a 40 to 50 year old an upper limit of between 10 kHz and 15 kHz is more normal. Most hearing losses with age occur at frequencies above 1000 Hz, with the most typical form being a deepening notch centered around 3500 Hz. Noise induced hearing loss contributes to presbycusis, which is hearing loss with age. Figure 3.25 shows some typical hearing loss curves, one set for workplace noise induced loss, and the other for age-related loss. Scientists do not agree to what extent presbycusis should be considered "natural" and to what extent it is brought on by environmental noise.

A task group was appointed by the EPA to review the research on the levels of noise that cause hearing loss (von Gierke, 1973). It published, as part of its final report, the relationship

FIGURE 3.25 Progressive Hearing Loss Curves (Schneider et al., 1970)



between daily noise exposure and noise induced hearing loss for the most sensitive 10% of the population. Based on this and other data the EPA task force recommended no more than a day-night level of 80 dBA to protect the population from adverse health effects on hearing. The EPA Levels Document (EPA, 1974) went further and recommended an 8-hour L_{eq} level of no more than 75 dBA to protect public health for purposes of hearing conservation alone. These studies found no physiological effects for levels below 70 dBA.

The U.S. Occupational Safety and Health Administration (OSHA) has set legal standards concerning noise exposure of workers in the workplace. The legal limit is 90 dBA (as measured using the slow meter response) for an 8-hour workday with a 5 dBA per time halving tradeoff. This means that a worker may be exposed to no more than 85 dBA for 16 hours, 90 dBA for 8 hours, 95 dBA for 4 hours, 100 dBA for 2 hours, 105 dBA for 2 hours, 110 dBA for 1 hour, or 115 dBA for any time. The OSHA standard exposure is based on a finding by the American Academy of Otolaryngology—Head and Neck Surgery (AAO-HNS) that a hearing loss is significant only when the average hearing threshold at 500, 1000, 2000, and 4000 Hz has increased 25 dB. Clearly OSHA standards are much less restrictive than EPA recommendations, which were established without consideration of cost.

Workplace noise limits are characterized in terms of a noise dose, which is a relative intensity multiplied by a time, expressed as a percentage of an allowable limit. Over a given time period the dose can be calculated from levels, L_i measured in a series of intervals T_i

$$D = (100/T_n) \sum_{i=1}^{N} (T_i)\, 10^{\left[(L_i - L_c)/(10)\, q\right]} \tag{3.9}$$

where D = noise dose expressed as a percentage of the allowable daily dose
T_n = normalization time period, usually 8 hours
T_i = duration of the i^{th} time period (hours)

L_i = A-weighted, slow meter response, sound pressure level for the i^{th} time interval lying within the range of 80 to 115 dBA
L_c = criterion noise level, usually 90 dBA
q = nondimensional parameter, which determines the exchange rate over time; e.g., $q = 5/(\log 2)$ for a 5 dB per time halving exchange rate, $q = 3/(\log 2)$ for a 3 dB exchange rate
N = total number of intervals

The OSHA standards are expressed in terms of a total allowable level, which is based on 90 dBA over an 8-hour day and the 5 dB exchange rate. In these terms

$$L_{TWA} = 90 + (q)\log(D/100) \tag{3.10}$$

where L_{TWA} = time weighted average level (dBA) normalized to 8 hours
D = noise dose as defined above
q = exchange rate factor = 16.61 for the 5 dB rate

A 3 dB exchange rate assumes that hearing damage is proportional to the total energy of sound impacting the ear over a working lifetime. The 5 dB exchange rate allows more energy to impact the ear based on the understanding that the noise is intermittent and gives the ear some time to recover.

OSHA standards limit the time-weighted-average level to 90 dBA and requires that hearing protection be worn and that administrative and other controls on equipment be initiated. At 85 dBA OSHA requires that periodic hearing tests be performed on workers and that records on these tests be maintained.

3.7 OTHER EFFECTS

Precedence Effect and the Perception of Echoes

When a sound is reflected off a wall or other solid surface, the returning sound wave can be perceived as an echo under certain conditions. If the delay time between the initial sound and a second sound is decreased, the echo eventually disappears. A simple experiment can be carried out by clapping hands 15 meters (50 feet) or more away from a large flat wall and listening for the echo. At about 6 meters from the wall, the echo goes away and we hear a single sound. This is known as the precedence, or Haas effect (Haas, 1951), although its recognition predates Haas. The American scientist, Joseph Henry, demonstrated a similar effect at the Smithsonian in the 1840s (Davis and Davis, 1987). What happened in the hand clapping experiment is that the reflected sound finally fell below the level/delay threshold of perceptibility.

The perceptibility of a single echo is shown in Fig. 3.26 as measured in an anechoic space using speech at 70 dB and a loudspeaker placed in front. Figure 3.27 includes the results of the experiments of Haas, which were performed using a pair of loudspeakers, separated by 90°, placed in front of a listener. The source material was speech at different speaking rates presented at equal levels from each loudspeaker. One loudspeaker was delayed with respect to the other and the subjects were asked whether they felt disturbed by the reflection or echo. When the delay exceeded about 65 msec an annoying echo was perceived. At delay times less than 50 msec, echoes were perceived, which were not annoying even in cases where the reflection was 5 to 10 dB stronger than the primary sound (Blauert, 1983).

FIGURE 3.26 Absolute Perceptibility of a Delayed Signal (Kuttruff, 1973)

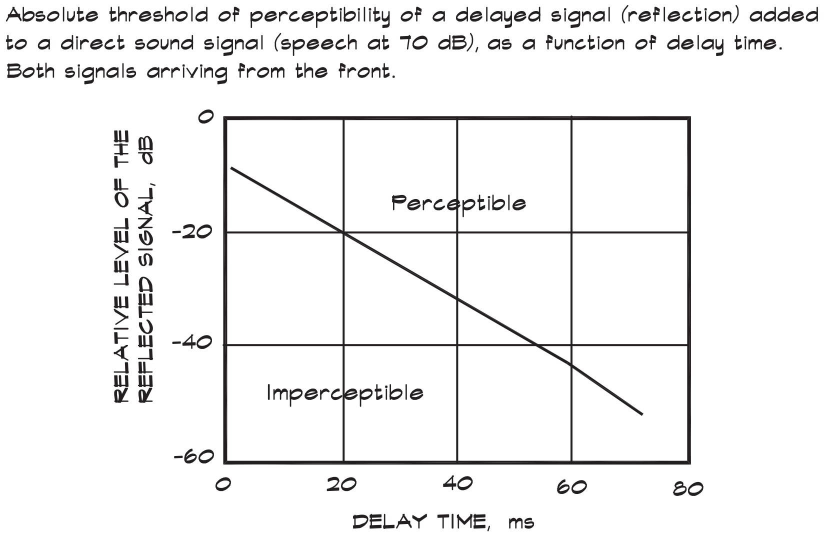
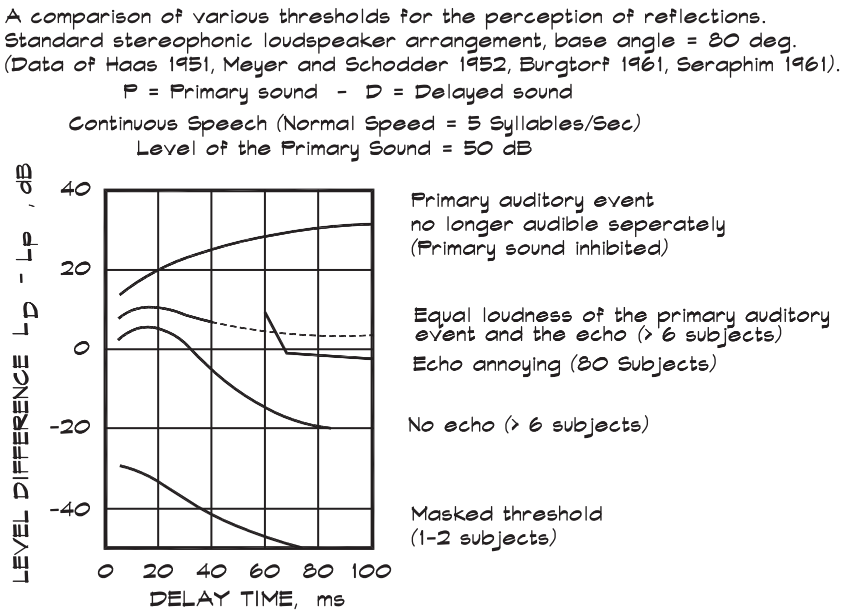


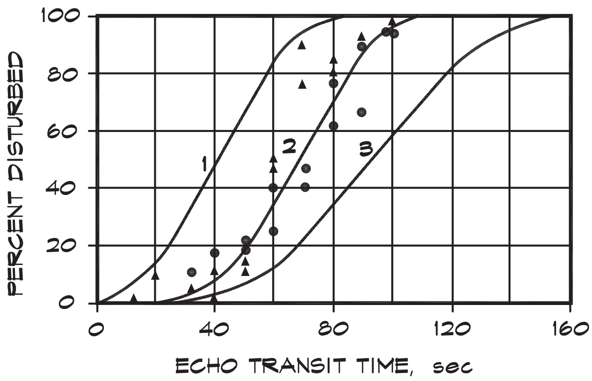
FIGURE 3.27 Thresholds for Perception of Reflections (Blauert, 1983)



Where the delayed sound was less than 30 msec after the initial event, the two sounds merged and no echo was perceived, even for sounds 5 or more dB above the initial sound.

The precedence effect is of considerable importance in architectural acoustics both for the natural reinforcement of live sounds coming from reflecting surfaces as well as for electronic reinforcement of speech or music. If a sound, originating from a performer on stage, is amplified and projected to an audience from a front loudspeaker, the image will

FIGURE 3.28 Disturbance Due to an Echo (Kuttruff, 1973)



Percentage of listeners disturbed by a delayed signal at the same level as the undelayed sound signal (speech). The abscissa is the delay time.

1. 7.4 syllables per second
2. 5.3 syllables per second
3. 3.5 syllables per second

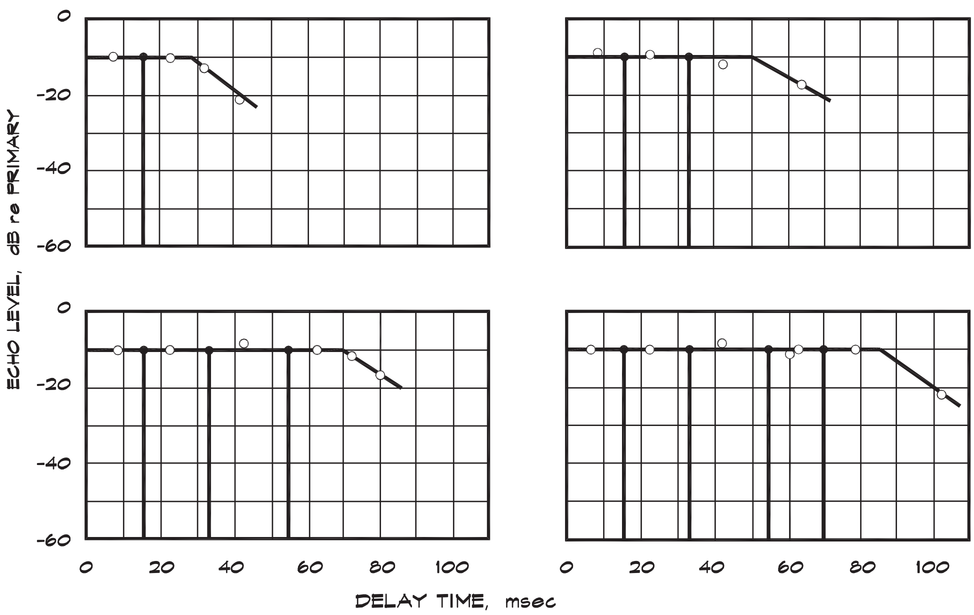
appear to come from the performer so long as the delay time is sufficiently short. The tradeoff between delay time and level, necessary to preserve the illusion of a single source, depends on the type of sound, i.e. speech or music, and the direction of the loudspeaker or reflector. If a reinforcing sound is within 25 msec of the initial sound, then speech is clearly understood. For music, a delay of 35 msec is normally not a problem even in rapidly played passages. For romantic music, delays as high as 50 msec can be tolerated. The three curves in Fig. 3.28 give the results of different rates of speaking on listener disturbance due to delayed echoes. The experiment is done with two loudspeakers, having the same level, placed in a relatively dead listening room (0.8 s reverberation time). The figure shows the importance of eliminating long-delayed reflections, and indeed even a single reflection, for the intelligibility of speech.

Data taken by Seraphim (1961), which were reproduced by Kuttruff (1973), led to the belief that when a series of delayed reflections arrives at a listener, the so-called Haas zone, where only a single sound is perceived, could be extended in time. Figure 3.29 gives the results of Seraphim's experiments using sounds coming only from the front in an anechoic environment. Here the threshold of perceptibility of the delayed sound is constant with delay time even when delays extend to 70 msec. Olive and Toole (1989) pointed out that this experiment, which was carried out in rather unrealistic conditions with all echoes having the same level, has been misapplied to reflections coming from many directions.

This is not to say that temporal forward masking does not occur for sounds having the same spectral content. Data published by Olive and Toole (1989) in Fig. 3.30 relate the absolute thresholds of perception for single lateral reflections to different types of source signal. In general the perception thresholds for speech and percussive reflections decrease with delay time, while those for music remain flat. The normal perception of a separate reflection, which would occur at a given delay time at much lower levels of reflected sound (Fig. 3.30), is masked. Experiments such as these indicate the importance of the smoothness

FIGURE 3.29 Absolute Perceptibility of a Delayed Signal (Kuttruff, 1973)

Absolute threshold of perceptibility of a delayed (reflection) being added to a sound field consisting of the direct sound plus one, two, three, and four reflections at fixed delay times and relative levels, which are denoted by the vertical lines. The original sound signal is speech. All sound components arriving from the front.



of the series of early reflections, following the arrival of the initial sound, in a critical listening space. These effects will be discussed in more detail in Chapt. 21.

Perception of Direction

The perception of direction is controlled by two factors: 1) the interaural delay time between the ears, and 2) the level difference created by the interaction between the head and the ears. When a sound originates to the left of the head in Fig. 3.31, it arrives at the left ear about a millisecond before the right ear. The high-frequency components are louder for the left ear than the right ear due to the shielding provided by the head itself. The brain uses the combination of level and time differences to decode the sound source direction.

When two sounds arrive at the listener the perceived direction is determined by first sound to arrive, even when the second sound is as much as 10 dB stronger. For equal level sources, delay gaps as low as one millisecond can bias the perceived direction to one side. This is called the precedence effect. Figure 3.32 shows the results of experiments (Madsen, 1970) on the tradeoff between time delay and intensity difference using two loudspeakers. At delay times below 1 msec, the higher level loudspeaker tends to dominate. Between 1 and 30 msec the precedence effect controls the direction, and above this point the two loudspeakers are increasingly perceived as separate sources.

When two sounds arrive at a listener simultaneously, the louder sound determines the direction. The apparent direction of a sound coming from two equidistant loudspeakers can be controlled by adjusting the level (panning) between them. Figure 3.33 contains the results of listening experiments performed by de Boer (1940) and Wendt (1963) by varying the level

FIGURE 3.30 Absolute Thresholds for a Single Lateral Anechoic Reflection (Olive and Toole, 1989)

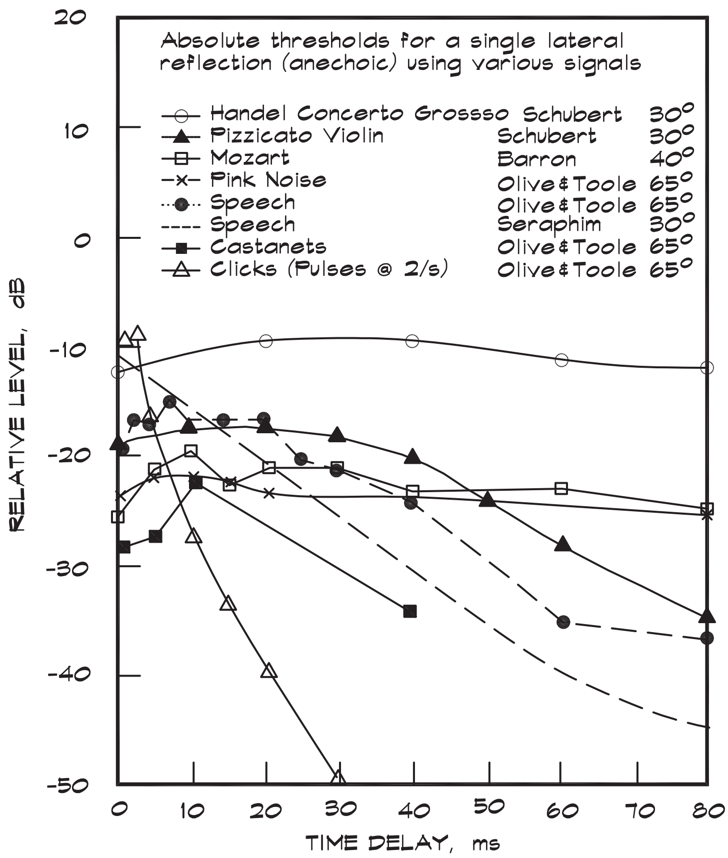
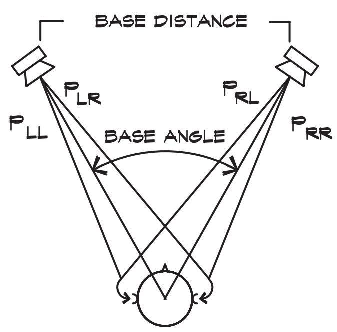


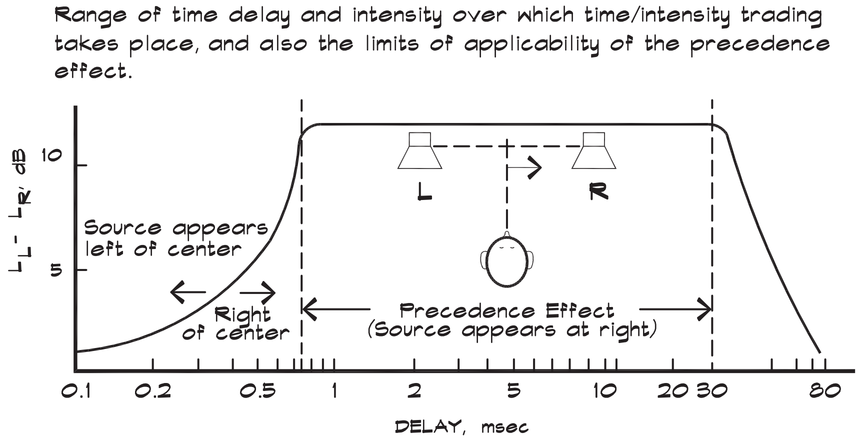
FIGURE 3.31 Head Geometry Relative to a Stereo Source (Blauert, 1983)



between two loudspeakers positioned 60° apart in the horizontal plane. When loudspeakers were placed 90° apart the error in the perceived direction increased significantly (Long, 1993). In most recording studios and mixdown rooms, loudspeakers spacing has standardized to a 60° spacing. Here the stereo image can be maintained and comfortably manipulated with panning.

In the vertical plane the ability of the brain to interpret time delays is much weaker, since our ears are on the sides of our heads. Results of localization tests in the vertical plane

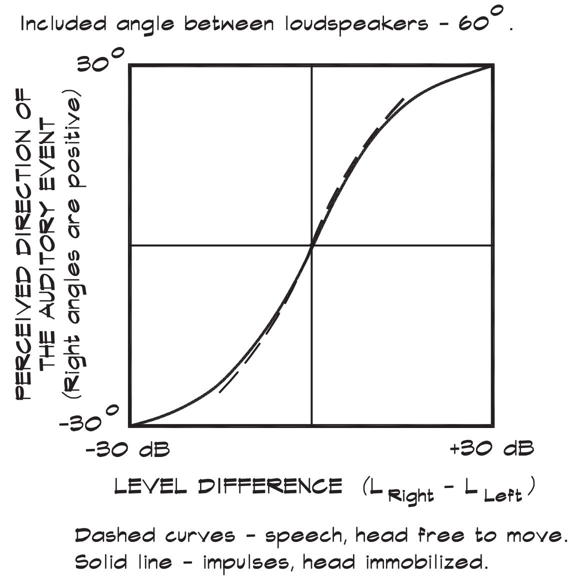
FIGURE 3.32 Perception of Source Direction with Delay (Madsen, 1970)



show a greater error and greater tolerance of wide loudspeaker placement. Our inability to precisely locate a vertical source makes realistic sound reinforcement systems possible. A properly designed loudspeaker cluster located above a stage can be used to augment the natural sound of the performers while maintaining the illusion that all the sound is coming from the stage.

The level-delay tradeoff has been carefully studied (Meyer and Schodder, 1952) by asking subjects to indicate the level difference at which the sound seemed to come from midway between a pair of stereo loudspeakers for various delays (Fig. 3.34). Study of this experiment is most helpful in the design of sound systems for it shows how far one can go in raising the loudspeaker level to augment the natural sound. It also shows how the apparent

FIGURE 3.33 Perception of Source Direction with Level (de Boer, 1940 and Wendt, 1963)



direction of sound can be moved about by using two loudspeakers and adjusting the time delay and level between them.

Clearly a stereo image, where a sound is perceived as originating between two loudspeakers, is difficult to maintain. The center image shifts to one side when one sound arrives only a few milliseconds earlier. Thus true stereo imaging is limited to a relatively small listener region close to the centerline between two carefully balanced loudspeakers. In a large room, such as a church or theater, a true stereo image can seldom be achieved.

Directional cues are best introduced by placing loudspeakers near the location of origin of the sound. For example, in motion picture sound systems, three loudspeaker clusters are arranged behind the screen in a left-center-right configuration, and the sound is panned to the proper level during the mix. In theme park attractions localization loudspeakers are placed in or near animatronic figures to provide a directional cue, even when most of the sound energy may be coming from a separate loudspeaker cluster.

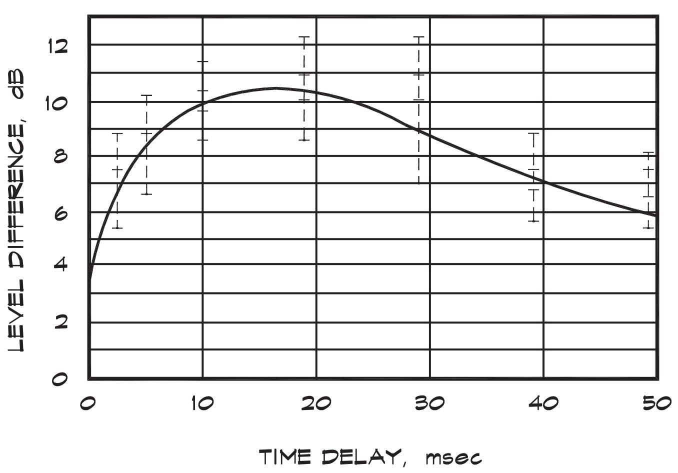
Binaural Sound

It is possible to reproduce many of the three-dimensional spatial attributes we hear in real life by recording sound using a dummy head with microphones in the ears and listening to the sound through stereo headphones, one for each microphone. This recording technique is referred to as dummy head stereophony or binaural reproduction, and is used in the study of concert hall design as well as in highly specialized entertainment venues. The results are startlingly realistic, particularly when the sound sources are located behind and close to the head.

When sounds are recorded binaurally, events that occur on the side or to the rear of our head are clearly localized. Sound sources located in front sound like they originate inside our head, overhead, or even behind. Several explanations for this phenomenon have been offered: 1) the effects of the pinnae are not duplicated when the playback system is a pair of headphones, 2) headphones affect the impedance of the aural canal by closing off the tube, and 3) the cues available from head motion are not present.

FIGURE 3.34 Equal Loudness Curve for Delayed Signals (Kuttruff, 1973)

Critical level difference between a delayed and undelayed signal which results in a perceived equal loudness of both signals (speech). (Meyer and Schodder, 1952)



This page intentionally left blank

4 ACOUSTIC MEASUREMENTS AND NOISE METRICS

4.1 MICROPHONES

Both microphones and loudspeakers are transducers—electromechanical devices for converting sound waves into electrical signals and vice versa. Microphones sense small changes in sound pressure through motion of a thin diaphragm. Cone loudspeakers create changes in pressure through the motion of a diaphragm driven by a coil of wire, immersed in a magnetic field. Since both microphones and loudspeakers operate in a similar manner, microphones can be used as loudspeakers and loudspeakers as microphones. Even the human eardrum can act as a loudspeaker.

The most common types of microphones in use are: 1) dynamic, 2) condenser, 3) electret, 4) ceramic, and 5) ribbon. All microphones consist of a diaphragm, which moves back and forth in response to changes in pressure or velocity brought about by a sound wave, and electronic components that convert the movement into an electric signal. Microphones are characterized by a sensitivity, which is the open circuit output voltage produced by a given pressure, expressed in decibels re 1 V/Pa. A one-inch diameter instrumentation microphone might produce 54 mV for an rms pressure of 1 Pa, yielding a sensitivity of $20 \log [(54\,\mathrm{mV}) / (1\,\mathrm{Pa})]\,[(1\,\mathrm{Pa}) / (1\,\mathrm{V})] = -25\,\mathrm{dB}$. Note that 1 Pa is the sound pressure that corresponds to the 94 dB sound pressure level generated by standard pistonphone calibrators.

A dynamic microphone, illustrated in Fig. 4.1, operates on the same principal as a loudspeaker. A diaphragm moves in response to the changes in sound pressure and is mechanically connected to a coil of wire that is positioned in a magnetic field. The induced current, produced by the motion of the coil, is the microphone's output signal. Both the diaphragm and the coil must be very light to produce adequate high-frequency response. Most dynamic microphones produce a very low output voltage; however, since the electrical output impedance is low, the microphone can be located relatively far away from the preamplifier. Dynamic microphones are rugged and are primarily used in sound reinforcement applications, where low fidelity is good enough. One manufacturer of dynamic microphones used to demonstrate its product's toughness by using the side of it to pound a nail into a block of wood.

A condenser microphone, in Fig. 4.2, consists of a thin stretched stainless-steel diaphragm that is separated from a back plate by a narrow air gap. The two parallel plates

FIGURE 4.1 Dynamic (Moving Coil) Microphone (Rossing, 1990)

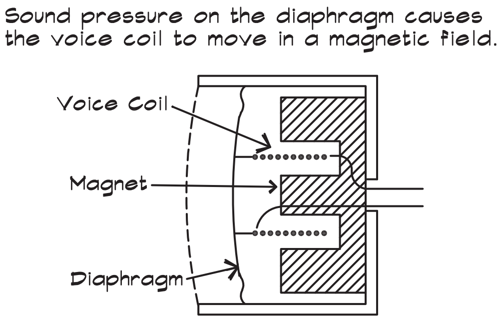
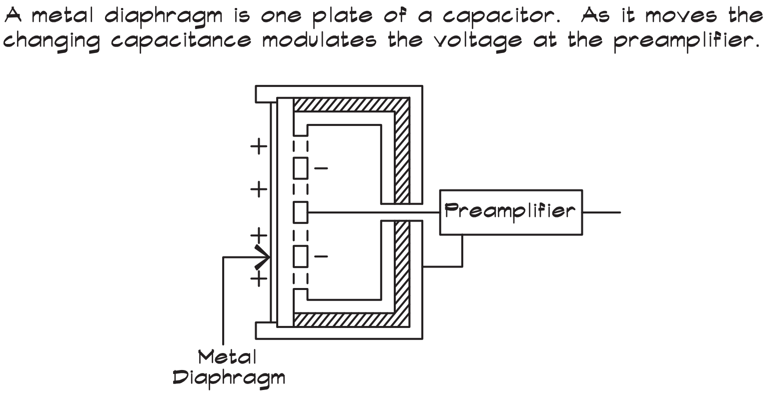


FIGURE 4.2 Condenser Microphone (Rossing, 1990)



become a capacitor when a DC polarizing voltage, typically 150 to 200 V, is applied. Motion of the diaphragm generates an electrical signal by varying the capacitance and thus the voltage between the plates. These microphones are very sensitive and accurate and have excellent frequency response characteristics. They are less rugged than dynamics and require a source of the polarizing voltage.

An electret microphone, in Fig. 4.3, is another form of condenser, which is sometimes called an electret condenser. It includes a thin polymeric diaphragm, where the polarizing voltage is not externally applied but is built into the polymer so that it is permanent. Otherwise the microphone operates in much the same way as the condenser does.

The ceramic microphone, in Fig. 4.4, has a diaphragm that is mechanically coupled to a piezoelectric material. A piezoelectric generates a voltage when strained. Many such materials exist such as lead zirconate titanate, called PZT, barium titanate, and rochelle salt. These microphones are more rugged than the capacitive types, are less sensitive, and do not require an external polarization voltage.

A ribbon microphone, sometimes referred to as a velocity microphone, works by suspending a thin metallic foil in a magnetic field. Figure 4.5 shows an example. The conducting ribbon is light enough that it responds to the particle velocity rather than the pressure. Since the ribbon is open to the back and shielded on the sides by the magnet, these microphones have a bidirectional polarity pattern. Ribbon microphones are very sensitive to moving air

FIGURE 4.3 Electret Condenser Microphone (Rossing, 1990)

A thin metallized plastic diaphragm is tightly stretched across a perforated backing plate. The holes in the back plate couple to an air cavity.

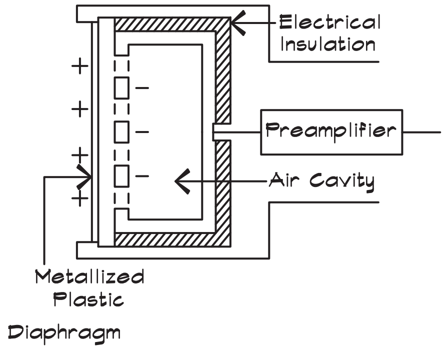


FIGURE 4.4 Ceramic Microphone (Rossing, 1990)

Sound pressure on the diaphragm causes deformation of the crystal, generating an electrical signal.

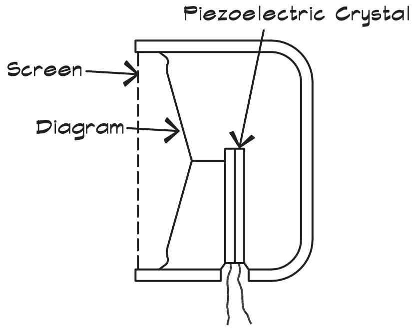


FIGURE 4.5 Ribbon Microphone (Rossing, 1990)

A lightweight ribbon diaphragm moves in a magnetic field, thus generating an electrical signal.

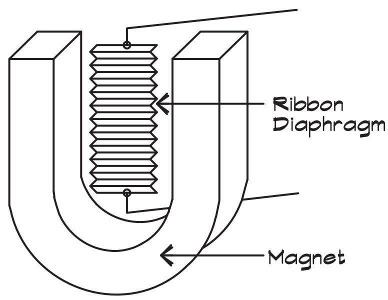
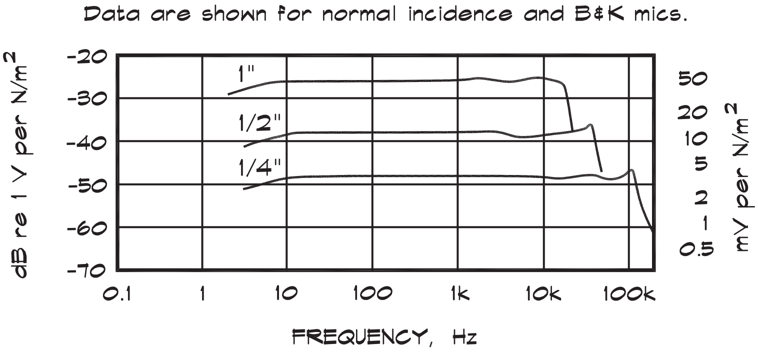


FIGURE 4.6 Sensitivity of Condenser Microphones (Hassall and Zaveri, 1979)



currents as well as high sound pressure levels. An unsuspecting acoustician, seeking to determine the characteristics of a reverberation chamber, once fired a blank pistol in a room full of ribbon microphones, quickly converting them into expensive paperweights. Due to the fragility of this type of microphone, its use is limited to the studio.

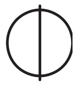
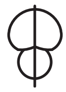
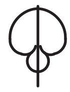
Frequency Response

Instrumentation microphones, so called because they can be calibrated using a pistonphone calibrator, are cylindrical and come in nominal sizes: one-inch (actually 0.936 in or 23.8 mm), half-inch (12.7 mm), and quarter-inch (6.5 mm) diameters. The size of a microphone affects its performance. Small microphones can measure sounds at higher frequencies and generally are less directional and less sensitive since they have a lower surface area. A one-inch instrumentation microphone, for example, might be able to measure levels as low as 0 dBA, while having an upper frequency limit of 10 kHz. A half-inch microphone might be good to 10 dBA and 30 kHz, and a quarter-inch microphone typically can measure down to 20 dBA and as high as 70 kHz. Examples of their response curves are given in Fig. 4.6.

Directional Microphones

Microphones, like sound sources, can have a response that varies with angle, which is represented by a polar diagram with angles measured relative to the normal to the diaphragm. Ideally, instrumentation microphones are nondirectional; however, at high frequencies there is some self shielding and loss of sensitivity, which is often greatest at a 120° to 150° angle of incidence. The polar diagrams for several types of microphones are shown in Table 4.1. Directional microphones are not used for precision measurements, but are quite useful for recording and sound-reinforcement systems. When the microphone capsule is smaller than a quarter wavelength, it is not directional; however, directivity can be built in by manipulating the construction of the housing. Figure 4.7 illustrates the design of a cardioid housing. By leaving an opening at the rear, sound coming from the rear arrives at the front and back of the diaphragm at the same time, thus canceling. Sound arriving from the front takes some additional time to reach the rear of the microphone diaphragm. By carefully attenuating selected frequencies traveling along certain paths the sound entering the rear cavities can be delayed so that it arrives close to 180° out of phase and does not cancel out the frontal sound.

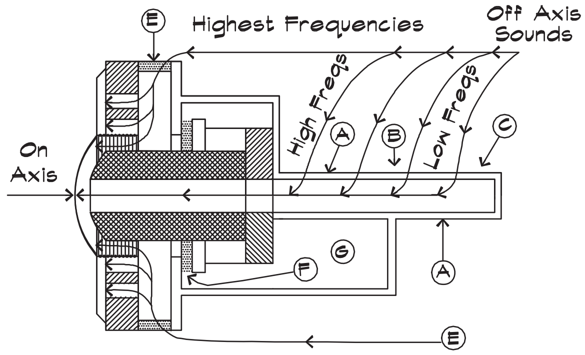
TABLE 4.1 Directional Characteristics of Microphones (Shure Inc., 2002)

	Omni-Directional	Bidirectional	Cardioid	Hypercardioid	Super-Cardioid
Polar Response Pattern					
Polar Equation	1	$\cos\theta$	$1/2(1 + \cos\theta)$	$1/4(1 + \cos\theta)$	$0.37 + 0.63\cos\theta$
Pickup Arc 3 dB Down	360°	90°	131°	105°	115°
Pickup Arc 6 dB Down	360°	120°	180°	141°	156°
Relative Output At 90° (dB)	0	$-\infty$	−6	−12	−8.6
Relative Output At 180° (dB)	0	0	$-\infty$	−6	−11.7
Angle at Which Output = 0	--	90°	180°	110°	126°
Random Energy Efficiency	0 dB	0.333 −4.8 dB	0.333 −4.8 dB	0.250* −6.0 dB	0.268** −5.7 dB
Distance Factor	1	1.7	1.7	2	1.9

* Minimum random energy efficiency for a first-order cardioid.
** Maximum front to total random energy efficiency for a first-order cardioid.

FIGURE 4.7 Cross-Section of an Electrovoice Variable-D Cardioid (Burroughs, 1974)

The letters indicate paths available for rear entry.



Highly directional microphones can be made using a series of openings in a tube, or a group of different length tubes, leading to the diaphragm. These so-called shotgun microphones work because sounds arriving on axis and entering through the holes combine in the tube in the proper phase relationship. Sounds arriving from the side and traveling down the tube combine with a random phase relationship that attenuates the signal at the diaphragm.

Directional microphones are very important in sound reinforcement systems. They selectively amplify sound coming from one direction, ideally from the user, and attenuate sound from other directions. This reduces feedback and allows a greater system gain. Properly designed directional microphones should have a consistent directivity pattern over a range of frequencies, otherwise they would color the off-axis sound. The more directional a microphone, the greater the coloration and the greater the directional lobing. Sometimes highly directional microphones can generate more system feedback than cardioid microphones, due to the influence of off-axis lobing patterns. In general, the less directional the microphone the more natural sounding it is.

Sound Field Considerations

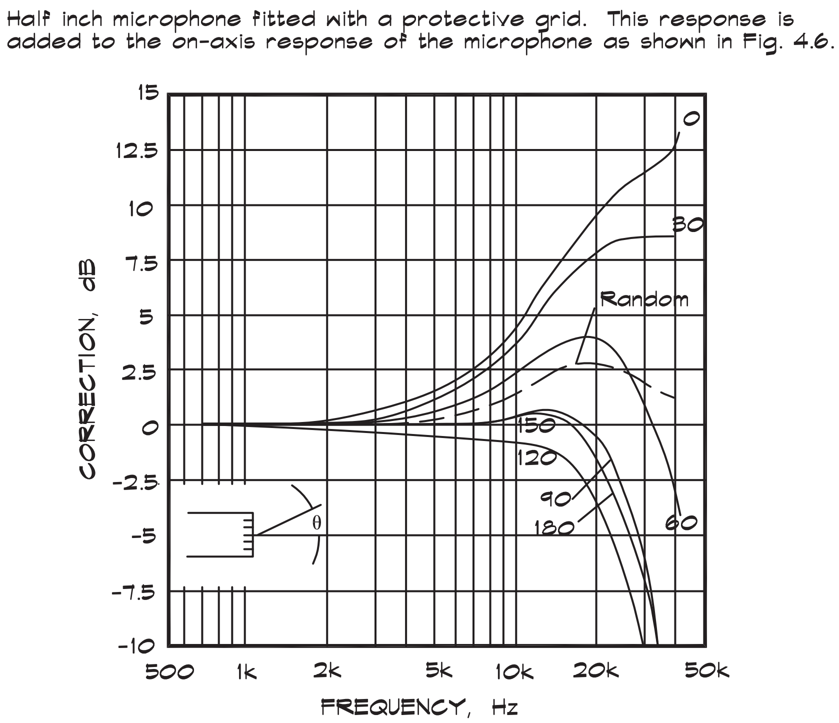
Microphone directivity sometimes influences the method of making measurements, even with instrumentation microphones. Typical microphones have their greatest sensitivity for sound incident on the diaphragm at 0°, called normal incidence. When the sound is traveling in a direction that is parallel to the plane of the diaphragm, at 90° to the normal, it is called grazing incidence. Most microphones have an angle for which their response is the flattest, usually 0° or 90°, but sometimes it can be another angle.

Microphones are described by their preferred type of sound field; for example, free field, random incidence, or pressure field. All microphones respond to pressure, but their sensitivity can be adjusted to produce the flattest response for a given angle of incidence or type of sound field. A free field is characterized by direct, unimpeded propagation of the wave from the source to the receiver. A diffuse or random field is one where the sound arrives from every direction with equal probability, and in a pressure field the sound pressure has the same magnitude throughout the space.

For a half-inch instrumentation microphone, below 5000 Hz all orientations produce a frequency response that is flat to within 2 dB. If a measurement is being made in a free field above 5000 Hz, the microphone should be oriented so that its flattest response direction is used, but this may vary with frequency, as can be seen in Fig. 4.8. Different standards organizations make different recommendations for proper free-field measurements (Fig. 4.9). IEC standards specify that the meter be switched to frontal mode and be oriented for normal incidence. ANSI standards require the selection of the random mode and an orientation of 70° to 80° to the source. For moving sources the microphone should be oriented for grazing incidence so that the directivity does not change with the motion of the source. This is achieved by angling the microphone upward.

When measurements are being done indoors, the random correction should be selected. Measuring with a free-field microphone in a diffuse field or with a random-incidence microphone in a free field yields only small inaccuracies, usually at high frequencies. The most accurate results will be obtained by using the setting appropriate to the type of sound field, but the differences are generally small.

FIGURE 4.8 Free Field Correction Curves for a Microphone (Bruel and Kjaer, 1986)



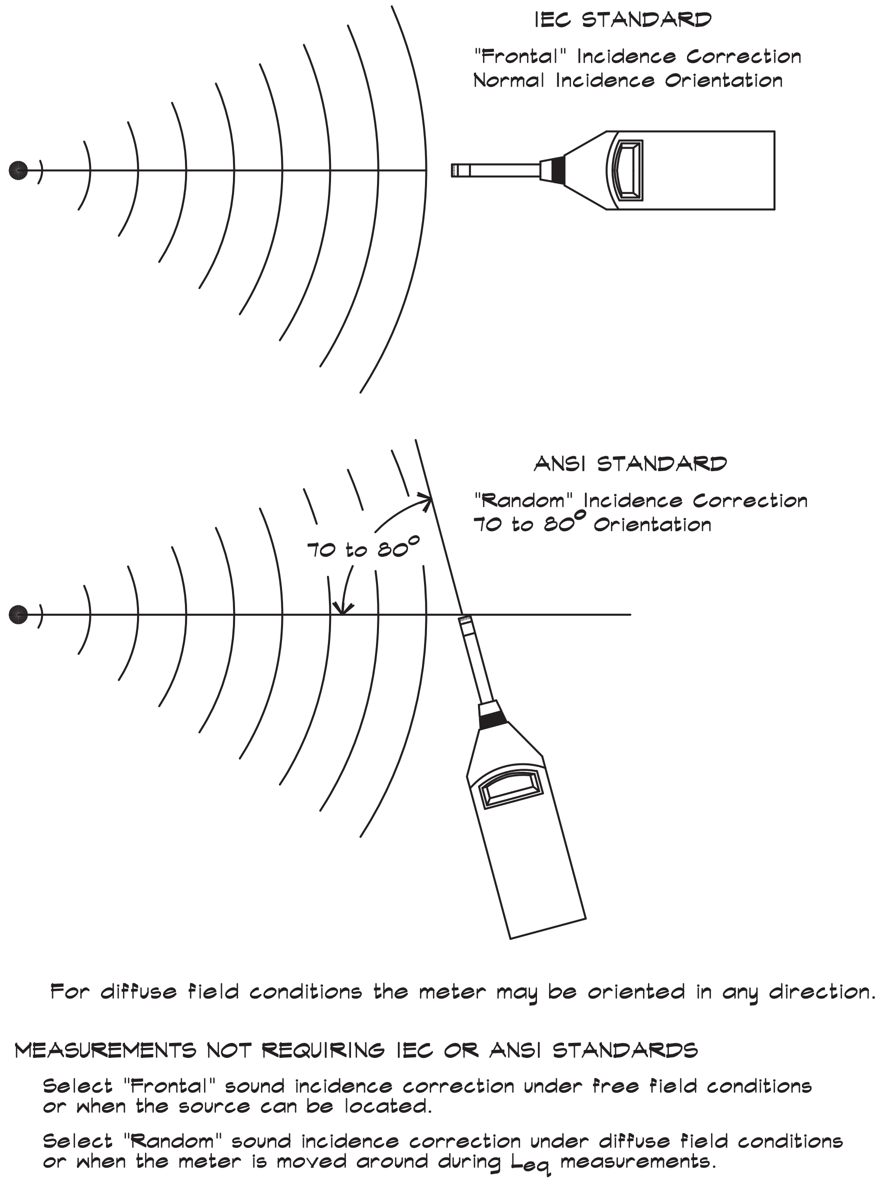
4.2 SOUND LEVEL METERS

The sound level meter, such as that shown in Fig. 4.10, is the fundamental acoustical instrument. Meters are battery powered and have become increasingly sophisticated, frequently containing internal processing, which automates many of the measurement functions. The individual controls vary from meter to meter; however, in general, there is a commonality of features. The basic controls allow for a selection of time weightings—fast, slow, and impulse—each of which represents a different ballistic time constant. Several frequency weightings are available: linear (unweighted), A-weighted, C-weighted, and a band limited linear scale. Frequency bandwidths may be selected from all pass, octave, and third-octave bands. There is a range selection that determines the highest and lowest levels measurable by the meter. Depending on the meter, there may be various types of automatic processing.

The internal parts of a meter include a microphone, preamplifier, various filters, a range control, time averager, and level indicator. The filters sometimes are contained in a separate module that may be attached to the meter, or are an integral part of the meter itself. On most hand-held sound level meters the filter selection is made manually. Where a group of filters operate simultaneously and display a number of levels on a bar graph in real time, the meter is called a spectrum analyzer or real-time analyzer.

Sound level meters are classified into three different groups by accuracy. Each class has a slightly different tolerance allowed in its precision. These standards are defined by the

FIGURE 4.9 Free Field Sound Measurements (Bruel and Kjaer, 1986)



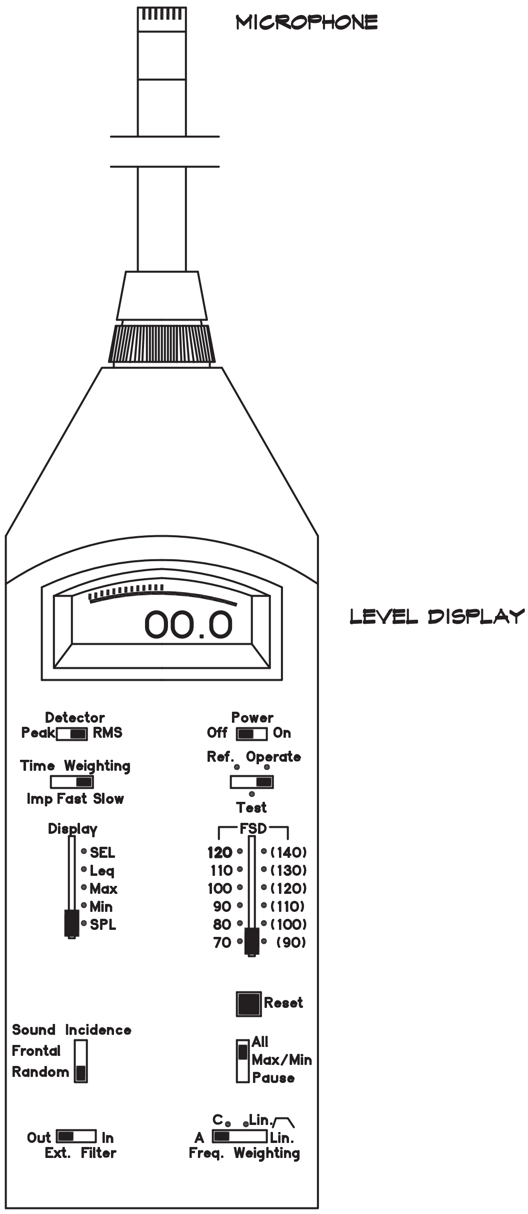
American National Standard Specification for Sound Level Meters, ANSI S1.4-1983.

Class 0	Laboratory	±0.2 dB 22.4 – 11200 Hz
Class 1	Precision	±0.5 dB 22.4 – 11200 Hz
Class 2	General Purpose	±0.5 dB 63.0 – 2000 Hz
		±1.0 dB 22.4 – 11200 Hz

Meter Calibration

Sound level meters should be calibrated before use, using a pistonphone calibrator placed over the microphone. These calibrators generate a steady tone, usually at 1000 Hz, by means

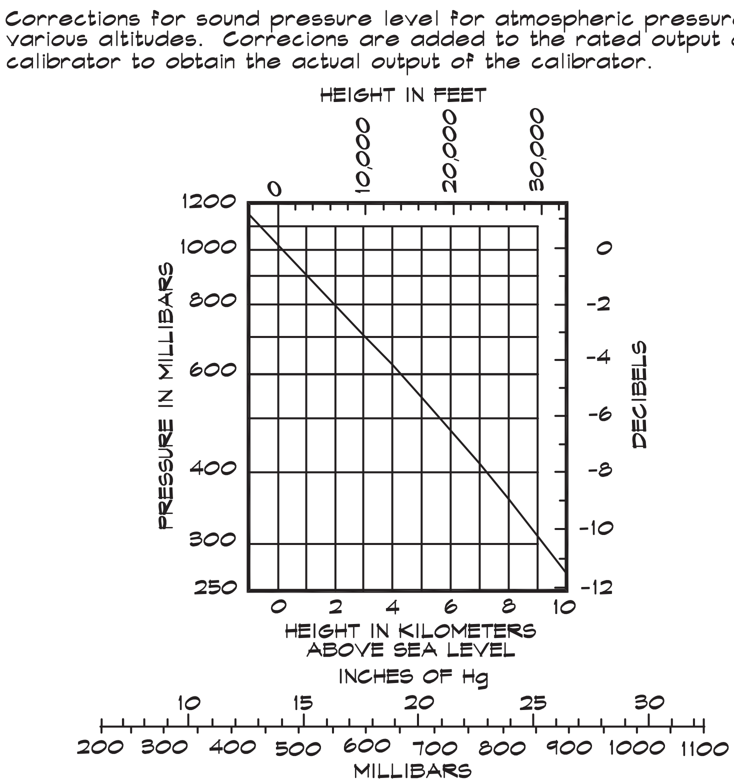
FIGURE 4.10 Sound Level Meter



of an oscillating piston in one end of a small cavity. The calibrator produces a nominal 94 dB, or with some calibrators a 114 dB, pure tone signal. The meter is adjusted to the proper level using a screw adjustment.

Pistonphone calibrators produce changes in volume in the cavity, which can be translated into changes in pressure using an equation of state. Most calibrators are set to produce the reference level at normal atmospheric pressure of 1013 millibars (1.01×10^5 Pa). Since atmospheric pressure varies, there is a correction given in Fig. 4.11 that must be applied according to altitude. This is the same correction as the term $10\ \log(\rho_0 c_0/400)$ in Eq. 2.67, including a density that changes with altitude.

FIGURE 4.11 Sound Level Meter Calibration Corrections (Peterson and Gross, 1974)



Calibrators themselves should be calibrated periodically against a microphone of known sensitivity. Since microphones are used to calibrate calibrators and vice versa, we encounter a classic chicken and egg conundrum; that is, how do we calibrated the original reference? The original microphone must be calibrated using another microphone in what is called a reciprocity calibration. The microphones used are identical and both transducers are used as loudspeakers and microphones in this technique. Refer to Kinsler et al. (1982) for further details.

Meter Ballistics

Early sound level meters were equipped with a d'Arsinval galvanometer, which responds to a voltage and indicates the sound level with a needle pointer. These early meters were very sensitive and tended to chatter or move back and forth rapidly. Electrical damping was added, which slowed the needle's response and made it more readable. The choice of the damping resistor in the indicator circuit, along with the capacitance of the microphone, set the exponential time constant of the circuit. Three response speeds are now used—*slow, fast*, and *impulse*. The slow setting has a time constant of 1000 ms (1 second), while for fast response it is 125 ms. A time constant has a precise mathematical meaning in engineering. In one time constant the value rises to $(1 - 1/e)$ or falls to $1/e = 1/2.718$ of its steady value. If a sound is instantaneously raised to a certain level the meter will rise to within 2 dB of the actual level in one time constant. Standard practice is to use 200 ms tone bursts at 1000 Hz to test a meter's response, since real sine waves have a finite rise time. The fast meter response must read within 2 dB of the steady level, and the slow meter response must be between 3 to 5 dB of the steady level (ANSI S4.1).

The rise time for fast and slow response is about the same as the fall time, so for integrating sound level meters, which measure L_{eq} levels, either fast or slow response gives about the same result. Some metrics, such as the CNEL level in California, require a particular response time, in this case, the slow response. For general use, the fast response is preferred. Impulse response is only employed to measure impact noise and other rapidly rising waveforms. The impulse time constant for a rising signal is 35 ms and for a falling signal is 1500 ms. Thus the meter holds the reading near its maximum level.

Meter Range

Sound level meters have an adjustable scale that allows the range of measurable levels to be set. If the range is set too low, then when a high level event occurs the meter will overload and not yield an accurate reading. If the range is too high, the indicated level will not fall below a certain value, and quiet events will not be measured accurately. Most meters have an overload indicator that signals the user to change the range. The range should be set as low as possible without tripping the overload indicator.

Detectors

There are two types of detector circuits found on most meters, peak, and rms (root mean square). Peak circuits sense the maximum amplitude present in the waveform. Mean-square detectors measure the time average of the square of the wave. Since the energy in the wave is proportional to the mean-square value, the rms detector is the most commonly used setting. Peak amplitudes are often of interest in vibration measurements. Peak-hold circuits, which capture the highest level during the measurement period, are utilized in the measurement of special sources such as sonic booms, where the wave shapes are not sinusoidal.

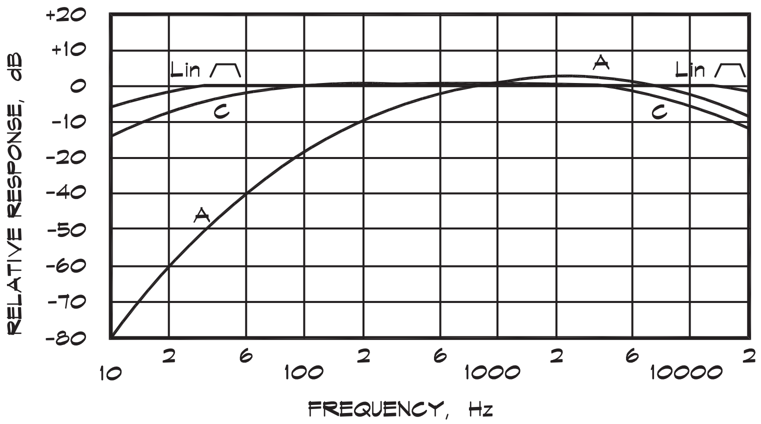
Filters

Sound meters come equipped with various selectable filters. The simplest is the linear filter, which passes sounds within the overall band limits of the instrument, for example 5 Hz to 100 kHz. This is not of particular interest in architectural acoustics, since it includes sounds that are well beyond our hearing capability. A second selection, the band-limited linear setting, includes a bandpass filter between 20 Hz and 20 kHz, and is quite useful for recording, since it blocks out low-frequency sounds that would otherwise overload a tape recorder. The characteristics of this filter along with the A and C weighting networks are shown in Fig. 4.12. Octave and third-octave bandwidth filters are also available. The standard frequency ranges have been given in Table 2.1. Filters may be cascaded, for example both octave band and A-weighting may be applied, yielding an A-weighted octave-band level. It is preferable to use the linear or band-limited linear settings when narrow-band filtering is done. This yields a consistent measurement methodology that does not require undue bookkeeping.

4.3 FIELD MEASUREMENTS

Field measurements are a critical part of architectural and environmental acoustics. Even with the simplest sources, care must be taken to follow proper procedure. A meter appropriate to the task must be selected. For environmental survey work a meter, tripod, calibrator, windscreen

FIGURE 4.12 A, C, and Lin Weighting Characteristics (Bruel and Kjaer, 1979)



(to reduce wind generated noise), logbook, distance measuring device (tape or rolling ruler), and watch are the standard kit. A small screwdriver is used to set the calibration. Spare batteries are a good idea. If they are left in the original packaging they can be distinguished from used ones. A camera is handy to record any unusual features of the site.

Headphones sometimes are included for listening to the sound being measured through the meter. They are essential for tape recording. Sometimes extraneous noise occurs that is not audible except through headphones. An example is arcing of the microphone, which can be caused by high humidity. Arcing produces a spurious popping sound that affects the data. Thus headphones are recommended when the relative humidity exceeds 90%.

For all measurements a record should be kept, noting the following information where it is relevant:

1) Location
2) Source description
3) Pertinent source details (e.g., manufacturer, model, operating point conditions)
4) Date and time
5) Engineer
6) Source dimensions and the radiating surfaces
7) Distance and direction to the source or a description of the measurement location
8) Meter settings
9) Background noise levels
10) Any unusual conditions
11) Time history
12) Measured data

Sources, which are outdoors and well away from reflecting surfaces, are the most straightforward. If the source is a piece of mechanical equipment the measurement position is selected based on the number of locations necessary to characterize the directivity of the source. For estimation of far-field levels from near-field measurements, data should be taken no closer than the largest dimension of the source, unless the area of the source is taken into account, by using Eq. 2.91.

The measurement distance for source characterization in a free field should be greater than a wavelength. For frequencies of 100 Hz the minimum distance is about 11 ft (3.4 m), while for 50 Hz the distance is about 22 ft (1.7 m). The danger of taking measurements too close is the possibility of including energy from only a portion of the source. If the source includes several separate pieces of equipment, the overall level will not be accurately represented if measurements are made too close to one individual component. Sometimes sound waves close to a source are not planar or are nonpropagating. Low-frequency emissions from large transformers are a good example of this type. Often low-frequency measurements require multiple samples and the microphone locations should be at least $\frac{\lambda}{4}$ apart.

Some sources are simply too large to conveniently get away from them. A good example is a refinery or a power plant. In such cases noise levels should be taken at regular distance intervals around the source and the results logged, according to where they were taken. Measurement locations should be spaced so that there is no more than a few decibels difference from one location to the next.

Measurements that are made to characterize a source rather than a location should be taken well away from reflecting surfaces. A minimum distance of $\frac{\lambda}{4}$ is recommended. If octave-band measurements are being taken and the 63 Hz band is of interest, then a distance of 4 to 5 feet is appropriate. Measurements will include reflections from the ground or other reflecting surfaces. Reflections from the observer can cause high-frequency comb filtering (Fig. 4.13), so the common practice is to hold the meter so that the microphone is extended away from the body or to support the instrument on a tripod.

An accurate measurement for source characterization is also difficult if the source receiver distance is too great. Even if the line-of-sight path is unimpeded, wind, atmospheric turbulence, ground cover, and air attenuation all play an important role in determining the measured noise level given off by a fixed source. At distances greater than 60 m (200 ft), noise level measurements can be dependent on wind velocity and direction. At distances greater than 150 m (500 ft), sound levels can be greatly influenced, even on a calm day, by ground cover, atmospheric turbulence, and air attenuation. At greater distances, thermal inversion layers can also be a major contributor. For all these reasons it is difficult to perform characterization measurements at large distances (say $>$ 60 m or 200 ft) from the source. Such measurements may be representative of a noise environment at a particular location under the measurement conditions, but may not be sufficiently accurate to characterize the source.

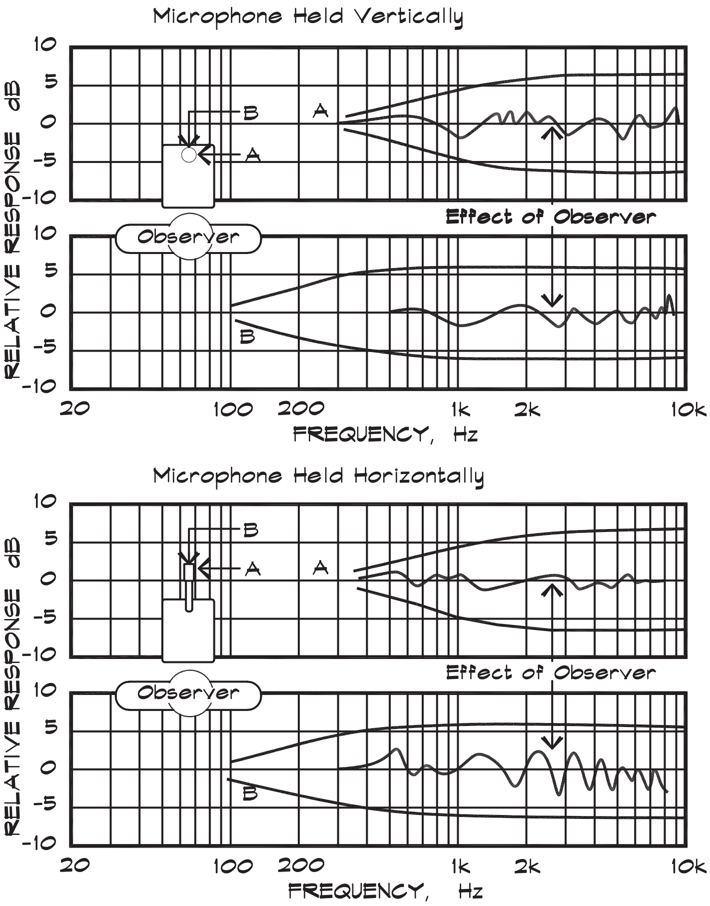
Background Noise

If there is a significant background (ambient) level present it should be measured. For a steady background it is best to turn off the source to be measured and note the ambient separately in all frequency bands of interest. The actual source-generated level then can be calculated from

$$L_{Source} = 10 \log \left[10^{0.1 L_{Tot}} - 10^{0.1 L_{Amb}} \right] \tag{4.1}$$

where L_{Source} = source sound pressure level (dB)
L_{Tot} = total combined source + ambient sound pressure level (dB)
L_{Amb} = ambient sound pressure level (dB)

FIGURE 4.13 Effect on Frequency Response as a Result of the Microphone Position (Petersen and Gross, 1974)



If the sound source cannot be turned off, it may be possible to measure the ambient noise level at a location that is similar to the location of interest but is away from the influence of the source. Locations may be available in shielded areas or, if the ambient noise is due to a roadway, at another site that is the same distance from the roadway.

When the background noise is variable and the source is steady, it is often easiest to measure the minimum combined level at a time when the ambient is quiescent. This gives an accurate source level if the ambient is sufficiently low. When the ambient is quiet, usually 10 dB below the source, its contribution can be ignored. With a varying ambient, if the source can be turned off, the minimum ambient can be recorded and then the minimum combined level measured. This gives a good value for the source level after adjustment using Eq. 4.1 as long as the minimum ambient levels are repeatable.

If the ambient is relatively steady and close to the source level, it can be measured separately using an averaging meter on the L_{eq} setting. The combined level then is measured in the same way and the source level calculated as before. This technique is also useful if the source, or background level, varies periodically, as it might with a pump motor or multiple sources such as fans or pumps, which produce beats. In taking data of this type, it is important to average over several beat cycles so that variations are properly taken into account.

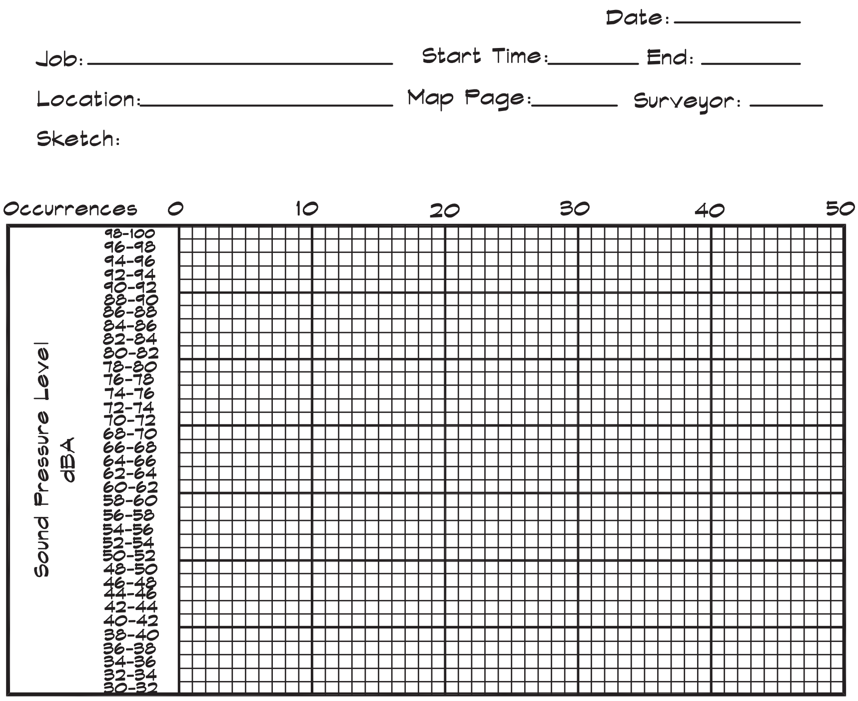
When the source level is less than the ambient, accurate measurements are difficult unless both the source and ambient levels are very steady. Even in these cases long averaging times are required to get good results. If the source is steady and the ambient varies, the minimum level gives the most accurate source level.

Time-Varying Sources

When traffic or other time-varying sources are to be measured, certain additional steps are useful. Although integrating meters are highly accurate, the nature of their output (i.e., one number) is sometimes not ideal, particularly when the data must be presented to a nontechnical audience. In these cases a log sheet such as that shown in Fig. 4.14 is helpful. In taking the data the meter is read at regular intervals, usually 5 or 10 seconds apart, and a notation is made on the log of the level that the meter shows at the interval mark. A representative number of samples are taken as determined either by the metric or the time period. One advantage to this methodology lies in the ability of the user to analyze the sampled data and extract more than one metric from the record. It also allows the engineer to ignore spurious signals such as barking dogs or aircraft flyovers that may not be relevant to the data being collected. Recording data, either on tape or in a recording sound level meter for later analysis, is another way of accomplishing the same goal. Data can be regularly sampled, and average levels calculated over a fixed time period and saved internally on a storage device for later analysis.

When a single moving source is to be measured, data are taken at a standard distance, say 15 m (50 ft), under prescribed conditions of velocity or acceleration. Data may be analyzed internally within the meter, or captured on a digital or analog recording device, or displayed

FIGURE 4.14 Noise Survey Log



graphically on a strip chart. When a recording is made, the calibration should flow through to all devices downstream of the meter. A tone is introduced using a pistonphone calibrator and is recorded along with the data. The meter range may then be adjusted by a known amount to accommodate the actual range of the data. A record should be made in a log or on the strip chart or verbally on tape noting the change in scale.

Both analog and digital recording devices can overload when signal levels exceed their dynamic range. When digital devices run out of headroom the resultant sound is most unpleasant. Analog tape recorders overload by producing a nonlinear or compressed version of the actual signal. If a two-channel device is available, the data may be recorded simultaneously on both channels at different level settings. This technique allows the data having the greater signal-to-noise ratio to be used, while retaining a margin of safety on the attenuated channel in case of overload.

Diurnal (24-Hour) Traffic Measurements

If a diurnal noise metric such as an L_{dn} or CNEL is to be measured, the ideal methodology is to position monitoring equipment at the location of interest for the entire 24-hour period. Often this is not practical due to the security, financial, or technical difficulties involved. In such cases a good estimate of the actual metric can be obtained by short-term monitoring if the hour by hour distribution of traffic is known or can be approximated. Measured distributions (Wyle, 1971) are given for urban traffic in Fig. 4.15 and for highway traffic in Fig. 4.16. The interesting feature about these data is that although they were taken 10 years apart they are almost identical. This implies that average diurnal traffic patterns are relatively stable.

If the reference L_{eq} level is known for the passage of one vehicle then the L_{eq} for N_h identical vehicles over the same time period is

$$L_{eq} = L_{ref} + 10 \log N_h \tag{4.2}$$

FIGURE 4.15 Typical Hourly Distribution of Total Daily Urban Vehicle Traffic (Wyle Laboratories, 1971)

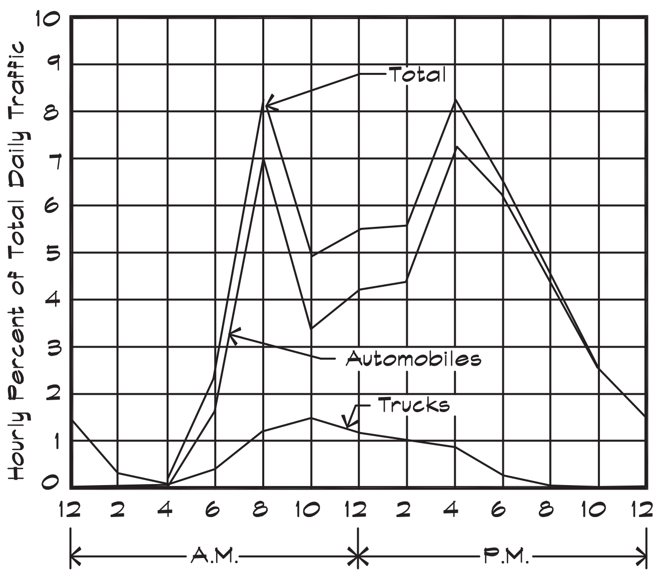
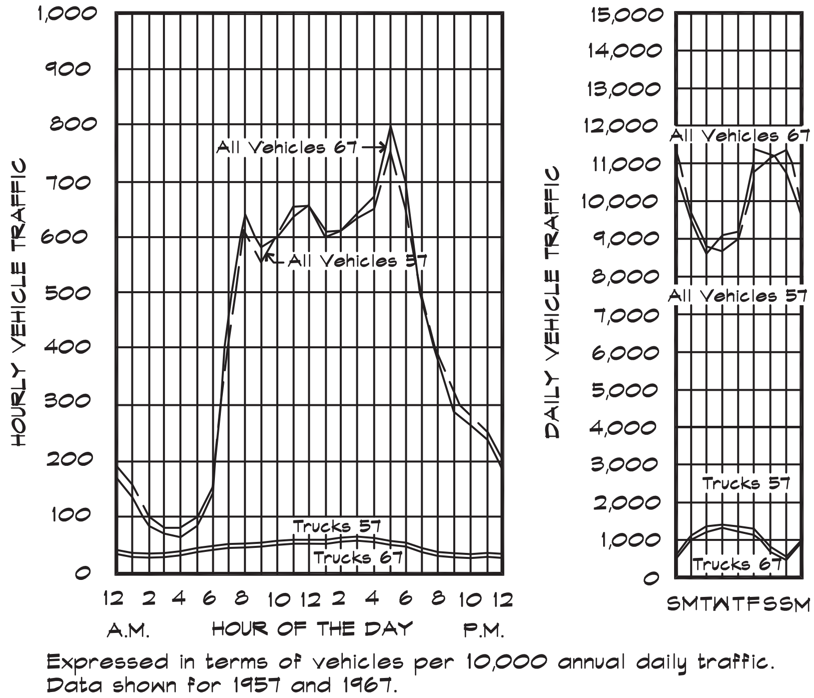


FIGURE 4.16 Hourly and Daily Variations in Intercity Highway Traffic in California (Wyle Laboratories, 1971)



where L_{eq} = equivalent sound level during the time period of interest (dBA)

L_{ref} = equivalent sound level for one vehicle passage during the time of interest (dBA)

N_h = number of like vehicles passing the measurement point during the time period of interest (usually one hour)

Assume that we can obtain the L_{eq} level for a given hour by direct measurement at a site. This can be accomplished by measuring over an hour period or by sampling the noise over a shorter time period and by assuming that the sample is representative of the hour period. Once the data have been obtained for the known hour, they can be adjusted for the time of day in which they were measured using standard distributions such as those in Figs. 4.15 and 4.16 or the actual site-specific traffic distribution, if it is known.

A traffic calculation uses a weighted hourly number of vehicles passing a point that yields the L_{dn} or CNEL level if inserted into Eq. 4.2. Thus

$$L_{dn} = L_{ref} + 10 \log N_{dn\,ave} \tag{4.3}$$

where L_{dn} = day night noise level (dBA)

L_{ref} = equivalent sound level for one vehicle passing by during an hour period (dBA)

$N_{dn\,ave}$ = weighted average number of like vehicles passing the measurement point during an equivalent hour

The day-night average number can be calculated from the distributions for urban and highway conditions using Eq. 3.7 for L_{dn} or Eq. 3.8 for CNEL.

$$N_{dn\,ave} = \left\{ \frac{1}{24} \left[\sum_{i=8}^{22} N_i + (10) \sum_{i=23}^{7} N_i \right] \right\} \tag{4.4}$$

where N_i = number of vehicles passing the measurement point during the i th hour

Finally by subtracting Eq. 4.2 and 4.3 we can obtain the difference in decibels between an L_{eq} level in any particular hour and the day-night level over a 24-hour period for a known traffic distribution.

$$L_{dn} \cong L_{eq}(h) + C(h) \tag{4.5}$$

where L_{dn} = day - night noise level (dBA)

$L_{eq}(h)$ = equivalent sound level for a given hour, h (dBA)

$C(h) = 10 \log \dfrac{N_h}{N_{dn\,ave}}$

= correction (dB) for the hour, h, based on the appropriate traffic distribution

The result is given in Table 4.2 for the Wyle urban and highway distributions for L_{dn}. The CNEL for these distributions is about 0.5 dB higher. If the traffic pattern at a particular site differs from those given here and is known, a similar calculation can be done for the specific distribution.

Included in these approximations is the assumption that the traffic speed and other factors that affect traffic noise, such as truck percentage, remain nearly the same over a 24-hour period. On crowded city streets this may not be the case. If traffic is free-flowing during the measurement period this method gives a conservative (high) estimate of the L_{dn} level. If traffic is slowed due to congestion, the noise levels will not be representative of a free-flowing condition.

If readings are taken during congested periods, the method will underestimate the actual 24-hour levels. If traffic slows significantly during rush hour, measurements made during off-peak periods, when traffic is flowing freely, will yield a result that is somewhat higher than the actual L_{dn} value.

The distribution of truck traffic over the day does not exactly track the automobile distribution. A similar calculation can be undertaken that includes truck percentages, with a knowledge of the difference between the reference level for trucks and cars. Naturally this introduces additional complexity. Based on 24-hour measurements, the method has been found to yield levels within one or two dB of the actual values, even without inclusion of a separate truck percentage distribution.

4.4 BROADBAND NOISE METRICS

At first glance the number and variety of acoustic metrics is overwhelming. In no other science are there as many different fundamental ways of measuring and characterizing the

TABLE 4.2 Approximate Conversion from Leq to Ldn or CNEL (Based on the traffic distributions shown in Figs. 4.15 and 4.16)

Hour	Highway Distribution		Urban Vehicle Distribution	
	CNEL − L_{eq} (dB)	L_{dn} − L_{eq} (dB)	CNEL − L_{eq} (dB)	L_{dn} − L_{eq} (dB)
1	8.2	7.7	10.9	10.4
2	10.4	9.9	15.1	14.6
3	11.2	10.7	16.9	16.4
4	11.6	11.1	19.9	19.4
5	10.6	10.1	9.5	9.0
6	8.2	7.7	6.5	6.0
7	3.6	3.1	2.6	2.1
8	1.6	1.1	0.8	0.3
9	2.0	1.5	1.7	1.2
10	1.9	1.4	2.9	2.4
11	1.6	1.1	2.7	2.3
12	1.5	1.0	2.6	2.2
13	1.9	1.4	2.5	2.1
14	1.8	1.3	2.4	1.9
15	1.6	1.1	1.5	1.0
16	1.4	0.9	0.8	0.3
17	0.7	0.2	1.3	0.8
18	1.3	0.8	1.8	1.3
19	2.7	2.2	2.7	2.2
20	4.0	3.5	3.6	3.1
21	5.0	4.5	4.3	3.7
22	5.3	4.8	5.6	5.1
23	5.9	5.4	6.9	6.4
24	7.1	6.6	7.8	7.4

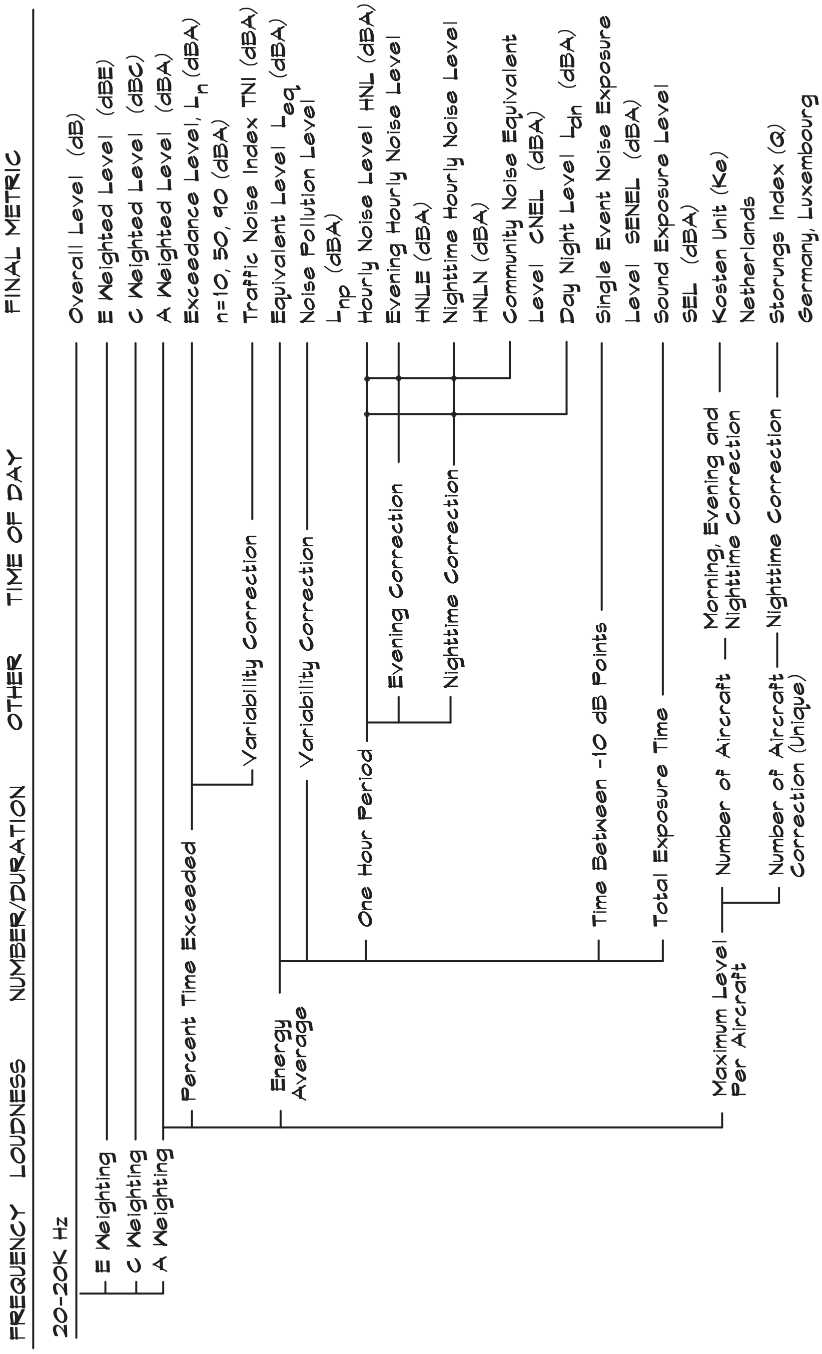
basic parameters. In physics the kilogram, meter, and second do not change. In electronics the volt, ohm, and ampere are stable and well defined. In environmental acoustics, however, different countries, states, cities, and counties often use different measurement schemes, which may not be directly convertible from one to another.

Even though the absolute number of metrics is large, the number of types of corrections applied to the measured data is rather modest. For example, a frequency correction for the loudness of a sound is included in most sound metrics but there are a number of ways to account for it, including A-weighting, NC curves, noys, and so on. The fundamental types of corrections include bandwidth, loudness, source number or duration, time of day, variability, onset, and pure tone content. The way each is included in a particular metric varies, but several are usually included in some fashion.

Bandwidth Corrections

The first correction category is the bandwidth of the measurement. Generally this is either wide band (i.e., 20 to 20 kHz) or band limited to octave or third-octave bandwidths. Narrow-band or chirped (swept) filters are also employed but the other corrections are seldom

FIGURE 4.17 Broadband Noise Metrics



applied to these measurements. Several metrics based on wide-band measurements are shown in Fig. 4.17. The loudness corrections in these measurements are applied by means of electronic filters, such as the A-weighting network, which are included in the meter itself. Subsequent corrections can be applied internally by the meter or can be added by a separate calculation.

Duration Corrections

One of the earliest metrics for describing traffic generated noise was the L_{10} (pronounced *ell-ten*) level. An L_n level is defined as the A-weighted sound level exceeded n % of the time during the measurement period. The L_{10} level is close to the maximum level occurring during a time period and its use reflects the fact that the highest levels are the most annoying. L_{10} levels are measured by using a histogram sampling technique, either manually or internally within the meter. If a histogram of measurements is made and there are 100 total samples, the L_{10} level is determined by counting 10 (10% of the total) measurements down from the highest level.

In a similar fashion the other excedance levels of interest can be determined. The L_{50} level or median is sometimes used. The L_{90} level is frequently used to characterize the residual background. L_n levels are expressed as whole numbers. From the statistical distribution of noise levels that can be characterized as normally distributed, certain relationships can be developed relating excedance levels to L_{eq} levels. For example, the energy average level, expressed in terms of the mean value (Barry and Reagan, 1978), is

$$L_{eq} = L_{50} + 0.115\,\sigma^2 \tag{4.6}$$

where L_{eq} = equivalent sound level (dB)
L_{50} = mean value sound level (dB)
σ = standard deviation of the sound levels (dB)

For a normal distribution, the L_{50} level and the L_{10} level are related

$$L_{10} = L_{50} + 1.28\,\sigma \tag{4.7}$$

The relationship between L_{10} and L_{eq} can be obtained

$$L_{eq} = L_{10} - 1.28\,\sigma + 0.115\,\sigma^2 \tag{4.8}$$

where L_{eq} = equivalent sound level (dB)
L_{10} = sound level exceded 10% of the time (dB)
σ = standard deviation of the sound levels (dB)

The standard deviation of highway traffic noise is usually 2 to 5 dB, so the L_{10} level is higher than the L_{eq} level. For traffic noise, the L_{eq} level is about equal to the L_{20} level. Not all outdoor noise distributions are normal, so these equations should be used carefully as general estimates of the actual values.

Variability Corrections

Metrics have been developed that include a term for the variability of the sound, the theory being that the more variable the sound distribution, the more annoying it is. The noise pollution level is one of these and is used to characterize community noise impacts. It is defined as

$$L_{NP} = L_{eq} + 2.56\,\sigma \tag{4.9}$$

where L_{NP} = noise pollution level (dBA)
L_{eq} = equivalent sound level (dBA)
σ = standard deviation of the sound levels (dBA)

Note that the noise pollution level uses A-weighting.

The traffic noise index (TNI) is another metric that includes a term for the variability of the noise environment. In this metric the variability is characterized in terms of the difference between the L_{10} and the L_{90} levels. The traffic noise index is given by

$$\mathrm{TNI} = 4(L_{10} - L_{90}) + L_{90} - 30(\mathrm{dBA}) \tag{4.10}$$

where TNI = traffic noise index (dBA)
L_{10} = level exceeded 10% of the time (dBA)
L_{90} = level exceeded 90% of the time (dBA)

Both the noise pollution level and the traffic noise index were developed for use in characterizing traffic noise and are not as accurate in predicting human reaction to other environmental noise sources.

Sound Exposure Levels

Metrics that utilize the format of energy times time are called exposure levels and are expressed in decibels with a reference period time of one second. There is considerable usefulness in such metrics in that they contain all the energy that occurs during a given event packed into a period one second long. The sound exposure level (SEL) is one such metric and is defined as

$$\mathrm{SEL} = 10 \log \left[\sum_{i=1}^{N} 10^{0.1 L_i} \right] \tag{4.11}$$

where SEL = sound exposure level (dBA)
L_i = sound level for a given one - second time period (dBA)
N = number of seconds during the measurement period

The SEL can be measured directly by many sound level meters. The meter can be set to display the SEL, which is internally computed, following the initiation of the measurement, by pushing the meter reset button. The L_{eq} can be calculated from the SEL for a given time period T

$$L_{eq} = \mathrm{SEL} - 10 \log(T) \tag{4.12}$$

where SEL = sound exposure level (dBA)
L_{eq} = equivalent sound level for a given time period (dBA)
T = time (s)

When there are several events, the L_{eq} level can be calculated from the SEL levels for each event. The SEL levels are combined using Eq. 2.62 and the L_{eq} level is calculated using Eq. 4.12. If both the L_{eq} and the SEL are measured simultaneously, the measurement time period can also be calculated using Eq. 4.12.

Single Event Noise Exposure Level

The single event noise exposure level (SENEL) is similar to the SEL in that it sums the energy times the time associated with an event. Originally, it was developed to measure the noise energy of the flyby of a single aircraft. In such measurements it is sometimes difficult to tell when to begin and when to stop the readings. If the data are recorded on a strip chart or tape recorder it is unclear at what point on either side of the peak to stop adding up the energy. To short cut the process the SENEL was developed. This metric is the exposure level contained in the top 10 dB of a single event sound level record. The duration of the event in a SENEL is the time between the two points at which the level falls 10 dB below the maximum. Figure 4.18 shows the L_{eq} for a triangular sound pattern. The SEL or SENEL may be calculated from these L_{eq} levels by using an equation similar to 4.12, where the time period is equal to the pulse duration τ. Once the SENEL is known, the L_{eq} can be calculated for any period of time containing the event.

$$L_{eq} = \mathrm{SENEL} - 10\log(\mathrm{T}) \tag{4.13}$$

where SENEL = single event noise exposure level (dBA)
L_{eq} = equivalent sound level for a given time period (dBA)
T = time period for which the L_{eq} is to be calculated (s)

Note that it is necessary that the time period T in both Eqs. 4.12 and 4.13 be equal to or greater than the time period over which the SEL or SENEL was measured; otherwise, the event is not accurately represented. As with SEL, if several events occur within a given time period, then the individual SENEL levels may be combined using Eq. 2.62. An equivalent level can be calculated using Eq. 4.13 from the combined SENEL level.

4.5 BAND LIMITED NOISE METRICS

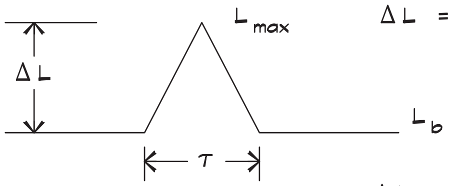
Techniques used for measurements employing octave band or other bandwidth filters vary little from those described for measuring broadband levels. Care must be taken in measuring low-frequency sounds so that the appropriate spacing between the source, reflecting surface, and the measurement location is observed. Sufficient sampling time is also a factor with low-frequency measurements because some low-frequency sources produce beat frequencies, which may be on the order of 1 Hz or less and may vary slowly over time.

Figure 4.19 shows a summary of the types of metrics obtained from octave-band measurements. As with the broadband systems there are a number of different metrics; however, the number of correction categories is relatively small. A loudness can be measured using electronic filters such as the A-weighting network. The A-weighted octave-band spectrum is useful as an aid in the determination of the frequency band making the most significant contribution to the overall A-weighted noise level. If most of the A-weighted energy is contained in one frequency band, then noise control efforts should be concentrated there.

A simple unweighted octave-band level is the basis for a number of metrics that determine the loudness by a direct comparison of the measured data to a standard curve of values. Several standards have been developed over the years, having to do principally with heating, ventilating, and air conditioning (HVAC) noise. The NC and RC curves are described in Chapt. 3.

FIGURE 4.18 L_{eq} Levels for Various Time Patterns (US EPA, 1973)

For triangular shaped patterns - time period = τ



$$\Delta L = L_{max} - L_b$$

$$L_{eq} = L_b + 10 \log \frac{10}{2.3\,\Delta L}\left(10^{\frac{\Delta L}{10}} - 1\right)$$

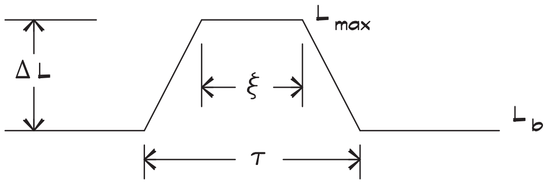
When ΔL is greater than 10 dB the following approximation is accurate.

$$L_{eq} = L_{max} - 10 \log \frac{2.3\,\Delta L}{10}$$

When there are a series of n identical triangular time patterns as shown above occurring within an interval T then

$$L_{eq} = L_b + 10 \log \left[1 + \frac{n\tau}{T} \left\{ \frac{\left(10^{\frac{\Delta L}{10}} - 1\right)}{2.3} - \frac{\Delta L}{10} \right\} \right]$$

For trapezoidial shaped patterns - time period = τ



$$L_{eq} = 10 \log \frac{1}{\frac{(\tau - \xi)\,\Delta L}{10} + \frac{\xi}{2}} \left[10^{\frac{L_b}{10}} \frac{(\tau - \xi)}{2.3} \left(10^{\frac{\Delta L}{10}} - 1\right) + 10^{\frac{L_{max}}{10}} \frac{\xi}{2} \right]$$

When ΔL is greater than 10 dB and ξ is small compared to τ.

$$L_{eq} = L_{max} - 10 \log \frac{2.3\,\Delta L}{10} + 10 \log \xi$$

When there is a series of n identical trapezodial time patterns as shown above occurring within an interval T then

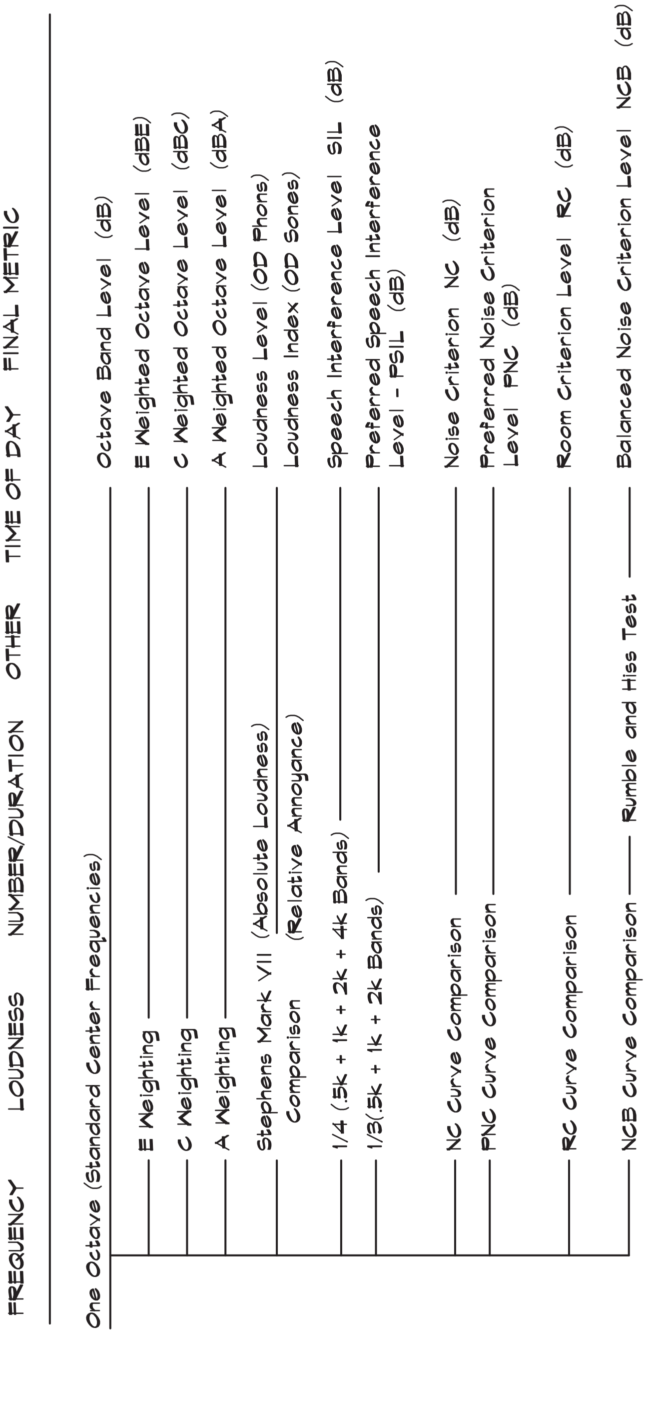
$$L_{eq} = L_{max} + 10 \log \frac{n\tau}{2.3\,T} + 10 \log n\xi$$

For a series of triangular pulses the last term above can be omitted.

Preferred Noise Criterion (PNC) Curves

PNC curves were introduced by Beranek in 1971 and are a revision of his earlier (1957) NC curves. PNC curves altered the high- and low-frequency octave values somewhat. The difference between the two has not been sufficient to result in the wide acceptance of the PNC version. The PNC curves are shown in Fig. 4.20. The use of the PNC curve is similar to that of the NC curve in that the PNC level is determined using the method of tangency.

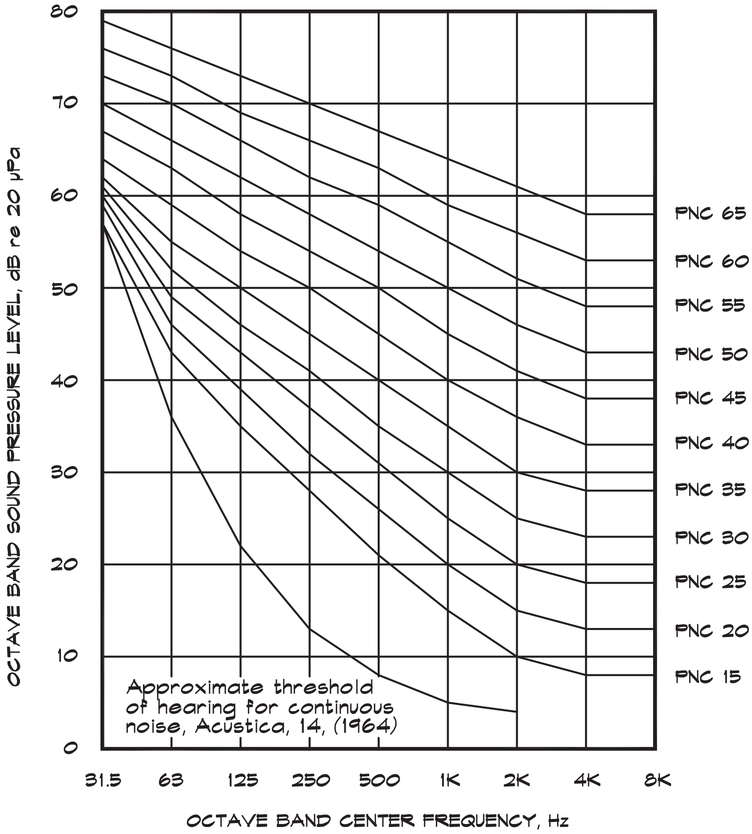
FIGURE 4.19 Octave Band Noise Metrics



Balanced Noise Criterion (NCB) Curves (Beranek, 1989)

In 1989, Beranek introduced another version of his 1957 NC curves, which he suggested for application to unoccupied rooms. These NCB levels, given in Fig. 4.21, are similar to the NC curves; however, the frequency range extends to 16 Hz and the metric calls for the calculation of the speech interference level (SIL) from the noise spectrum. This is rounded

FIGURE 4.20 Preferred Noise Criterion (Beranek, 1971)



to the nearest dB and compared with the NCB curve designation, which is also characterized by its SIL. If the measured SIL is equal to or below the curve designation, then the noise level meets the NCB criterion for speech interference.

Next comes a test for rumble or low-frequency energy. To check for this condition, 3 dB is added to the measured SIL and the NCB curve corresponding to this level is overlaid on the measured data. Where there are excedances of the new curve in the octave bands below 1 k Hz, they must be reduced to the elevated curve levels to comply with the standard.

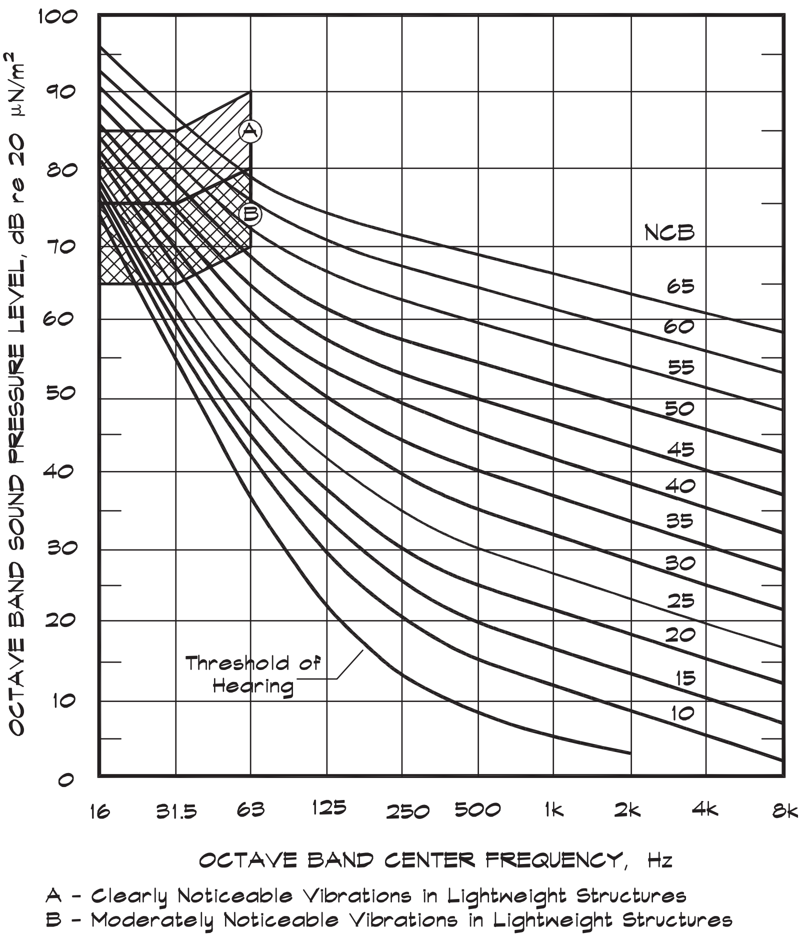
Finally there is the NCB test for hiss or high-frequency annoyance. An NCB curve is selected that provides the best fit to the measured data in the 125–500 Hz bands. Then this curve is plotted against the measured data. If the measured data exceed it in any of the three bands above 1 k Hz, then they must be reduced to meet the hiss criterion.

In occupied spaces Beranek calls for the measurement or estimation of noise levels due to normal work activities, which are to be combined with the unoccupied (HVAC) levels before a comparison to the NCB curve is made.

Other Octave-Band Metrics

Other systems exist for the determination of loudness based on measured octave-band data. They are based on empirical tests of relative or absolute comparisons presented to listeners in much the same way that the Fletcher-Munson experiments were done.

FIGURE 4.21 Balanced Noise Criterion Curves (NCB) (Beranek, 1989)



Robinson and Whittle (1962) constructed relative loudness curves in a very similar way. Stevens (1972) developed a series of systems for the calculation of loudness from octave band and other narrower bandwidth data. These systems rarely are encountered in architectural acoustics.

Octave-Band Calculations

It is frequently necessary to obtain an overall A-weighted level from unweighted octave-band data. The calculation is done by first adding the corrections for A-weighting, given in Table 3.1, to the level in each octave, and then by combining the A-weighted octave-band levels together using Eq. 2.62.

Occasionally it is necessary to generate an octave-band spectrum to match a given A-weighted level. This is straightforward if the spectrum shape of the sound source can be obtained. For example, let us assume that it is known that recorded music has a given octave-band spectrum and that this spectrum generates an overall A-weighted sound pressure level of 70 dBA. If we wish to obtain the octave-band spectrum of music that will yield an overall A-weighted level of some other level, for example 80 dBA, it is only necessary to add the difference between 80 and 70 to each octave-band level. It is assumed that the spectrum shape does not change with level for this source. It is useful to prepare normalized spectra for standard sources, which, when added to the overall A-weighted sound level, will yield an unweighted octave-band level having the same overall value.

If there are two sources present at the same time and we know the octave-band spectrum levels of each source independently, the spectrum for the two sources combined is obtained by applying Eq. 2.62 to the pairs of levels in each octave.

Third-Octave Bandwidth Metrics

Third-octave band metrics are similar to octave-band levels—they are simply a thinner slice of the same pie. They can be combined into groups of three centered around the octave-band center frequencies using Eq. 2.62 to obtain octave-band levels.

A summary of various third-octave and narrow band metrics is shown in Fig. 4.22. As with the octave-band metrics there are different versions of loudness and annoyance comparisons. One of these, the perceived noise level (PNdB) developed by Kryter (1970), has been used as the basis for several of the standard metrics for characterizing aircraft noise.

Aircraft Noise Rating Systems

Aircraft noise ratings vary principally in the methodologies they use for adjusting for the number of aircraft, the addition of pure tone corrections, and the inclusion of nighttime penalties. An excellent review of aircraft metrics was prepared by (Schuller et al., 1995). He summarizes the descriptors using the equation

$$\mathrm{Level} = A \log \left(\sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{n_i}\ \mathrm{w_i}\ 10^{\mathrm{L_i}/B} \right) - C \tag{4.14}$$

where A, B, C = constants
i = aircraft type category index
N = total number of aircraft type categories
$\mathrm{n_i}$ = number of noise events for aircraft category i per 24 - hour day
$\mathrm{w_i}$ = penalty (or weighting) factor for aircraft operation i
$\mathrm{L_i}$ = single event noise level for aircraft category i

The parameters used in Eq. 4.14 for various environmental metrics are given in Table 4.3.

Most of the metrics used for aircraft correlate well with the simpler $\mathrm{L_{dn}}$ level, which is the most commonly used system in the United States. For estimation purposes the following formulas may be used:

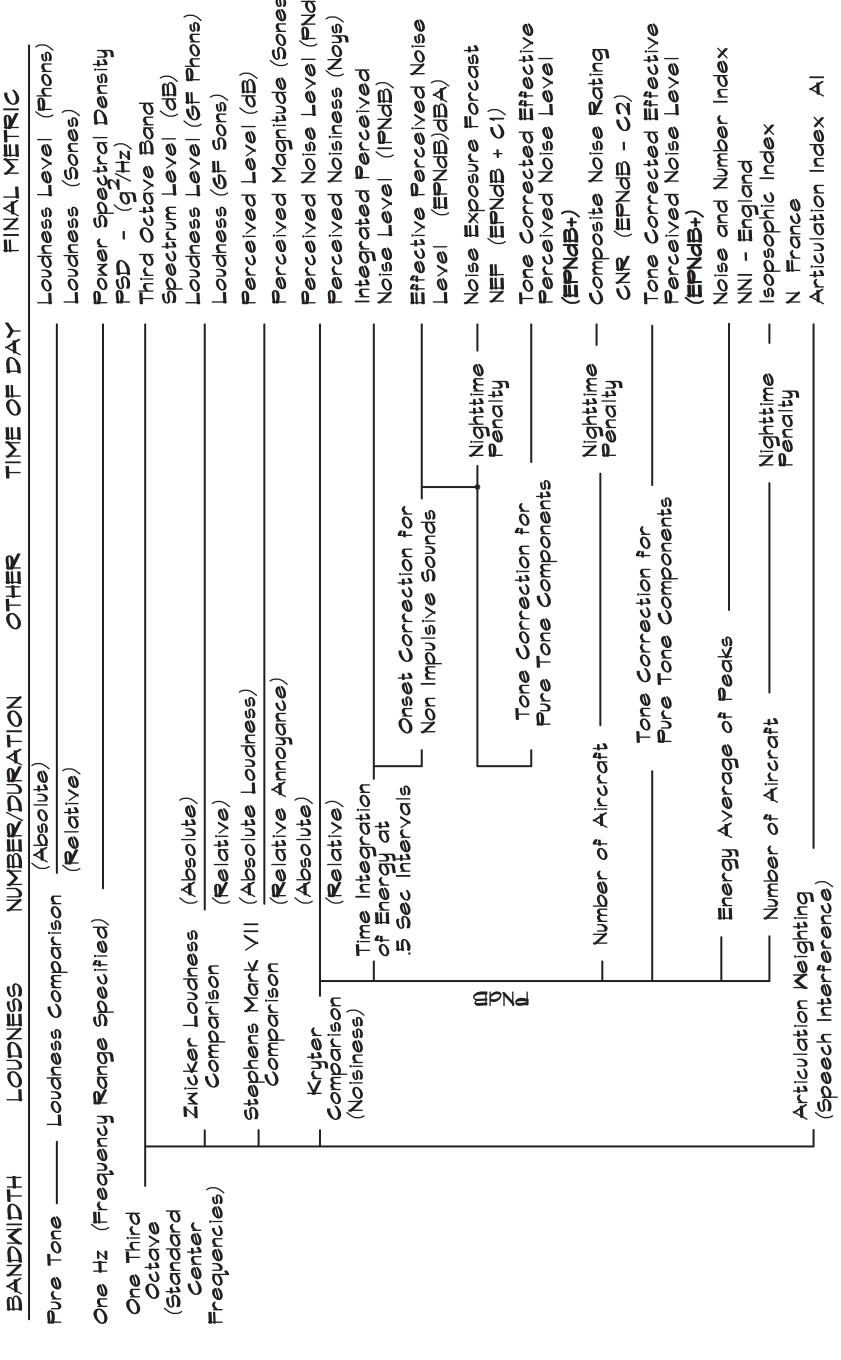
$$\mathrm{L_{dn}} \cong \mathrm{CNEL} \tag{4.15}$$

$$\mathrm{L_{dn}} \cong \mathrm{NEF} + 35(\pm 3) \tag{4.16}$$

$$\mathrm{L_{dn}} \cong \mathrm{CNR} - 35(\pm 3) \tag{4.17}$$

Similar relationships can be derived for the other metrics currently in use.

FIGURE 4.22 Narrow Band Noise Metrics



Narrow-Band Analysis

The analysis of sound in frequency bands of one-third octave and less is often useful for the detailed analysis of room acoustics and vibration. Instruments used for this type of measurement in real time are called spectrum analyzers or real-time analyzers (RTA). Two types of meters are most frequently encountered, those having a constant percentage bandwidth

TABLE 4.3 Parameters Used in Equation 4.14 (Schuller et al., 1995)

Metric	Constants A	Constants B	Constants C	Day Interval (Hours)	W_i	Morning, Evening Interval (Hours)	W_i	Night Interval (Hours)	W_i	L_i (dB)
Ke	20	15	105.8	08–18	1	06–08, 18–23	2–8	23–06	10	L_{ASmx}
L_n	10	10	44	06–23	0			23–06	1	L_{AE}
L_d	10	10	47.6[1]	07–23	1			23–07	0	L_{AE}
Q	13.3	13.3	65.7[2]	06–22	1			22–06	5[3]	L_{ASmx}
IP	10	10	49.4	07–22	1			22–07	10	L_{AE}
L_{dn}	10	10	44	06–23	0			23–06	1	L_{AE}
CNEL	10	10	44	07–19	1	19–22	3	22–07	10	L_{AE}
L_{24h}	10	10	44	00–24	1					L_{AE}
NNI	10	10	$80 - 5 \log \sum_{i=1}^{N} n_i w_i$	06–18	1			18–06	0	L_{pnmx}
NEF	10	10	88	07–22	1			22–07	16.7	L_{epn}

Single Event Noise Level Descriptors
L_{AE} = A-weighted sound exposure level
L_{ASmx} = Maximum S (slow) A-weighted sound level
L_{pnmx} = Maximum perceived noise level
L_{epn} = Effective perceived noise level

1) 7 hour night from 00.00 to 06.00 and 23.00 to 24.00 hours on a given day
2) 16 hour daytime period from 07.00 to 23.00 hours on a given day
3) Seperate calculations are specified for day and night. Values shown here are for calculations with emphasis on the contributions from nightime flight operations, Q_n. The weighting penalty includes a multiplication by the duration, in seconds, beween the first and last times that the instantaneous A-weighted sound level is within 10 dB of the maximum A-weighted sound level.