SECTION 20

AUDIO SYSTEMS

Although much of this section contains fundamental information on basic audio technology, extensive information on new evolving digital audio formats and recording and reproduction systems has been added since the last edition.

As noted herein, DVD-Audio, for example, is based on the same DVD technology as DVD-Video discs and DVD-ROM computer discs. It has a theoretical sampling rate of 192 kHz with 24-bit processing and can store 4.7 gigabytes on a disc with a choice of two- or six-channel audio tracks or a mix of both. Super Audio CD (SACD) has the same storage capacity. It uses direct stream digital (DSD) with 2.8 MHz sampling in three possible disc types. The first two contain only DSD data (4.7 gigabytes of data on a single-layer disc and slightly less than 9 gigabytes on the dual layer disc). The third version, the SACD hybrid, combines a single 4.7 gigabyte layer with a conventional CD that can be played back on conventional CD players.

MPEG audio coding variations continue to evolve. For example:

- MPEG-1 is a low-bit-rate audio format.
- MPEG-2 extends MPEG-1 toward the audio needs of digital video broadcasting.
- MPEG-2 Advanced Audio Coding (AAC) is an enhanced multichannel coding system.
- MP3 is the popular name for MPEG-1 Layer III.
- MPEG-4 adds object-based representation, content-based interactivity, and scalability.
- MPEG-7 defines a universal standardized mechanism for exchanging descriptive data.
- MPEG-21 defines a multimedia framework to enable transparent and augmented use of multimedia services across a wide range of networks and devices used by different communities. R.J.

In This Section:

On the CD-ROM:

The following are reproduced from the 4th edition of this handbook:

"Ambient Noise and Its Control," by Daniel W. Martin;

"Acoustical Environment Control," by Daniel W. Martin;

"Mechanical Disc Reproduction Systems," by Daniel W. Martin;

"Magnetic-Tape Analog Recording and Reproduction," by Daniel W. Martin.

CHAPTER 20.1
SOUND UNITS AND FORMATS

Daniel W. Martin, Ronald M. Aarts

STANDARD UNITS FOR SOUND SPECIFICATION[1,2]

Sound Pressure

Airborne sound waves are a physical disturbance pattern in the air, an elastic medium, traveling through the air at a speed that depends somewhat on air temperature (but not on static air pressure). The instantaneous magnitude of the wave at a specific point in space and time can be expressed in various ways, e.g., displacement, particle velocity, and pressure. However, the most widely used and measured property of sound waves is *sound pressure*, the fluctuation above and below atmospheric pressure, which results from the wave.

An *atmosphere* (atm) of pressure is typically about 10^5 pascals (Pa) in the International System of units. Sound pressure is usually a very small part of atmospheric pressure. For example, the minimum audible sound pressure (threshold of hearing) at 2000 Hz is 20 μPa, or $2(10)^{-10}$ atm.

Sound-Pressure Level

Sound pressures important to electronics engineering range from the weakest noises that can interfere with sound recording to the strongest sounds a loudspeaker diaphragm should be expected to radiate. This range is approximately 10^6. Consequently, for convenience, sound pressures are commonly plotted on a logarithmic scale called *sound-pressure level* expressed in *decibels* (dB).

The decibel, a unit widely used for other purposes in electronics engineering, originated in audio engineering (in telephony), and is named for Alexander Graham Bell. Because it is logarithmic, it requires a reference value for comparison just as it does in other branches of electronics engineering. The reference pressure for sounds in air, corresponding to 0 dB, has been defined as a sound pressure of 20 μPa (previously 0.0002 dyn/cm^2). This is the reference sound pressure p_0 used throughout this section of the handbook. Thus the sound-pressure level L_p in decibels corresponding to a sound pressure p is defined by

$$L_p = 20 \log (p/p_0) \text{ dB} \qquad (1)$$

The reference pressure p_0 approximates the weakest audible sound pressure at 2000 Hz. Consequently most decibel values for sound levels are positive in sign. Figure 20.1.1 relates sound-pressure level in decibels to sound pressure in micropascals.

Sound power and sound intensity (power flow per unit area of wavefront) are generally proportional to the square of the sound pressure. Doubling the sound pressure quadruples the intensity in the sound field, requiring four times the power from the sound source.

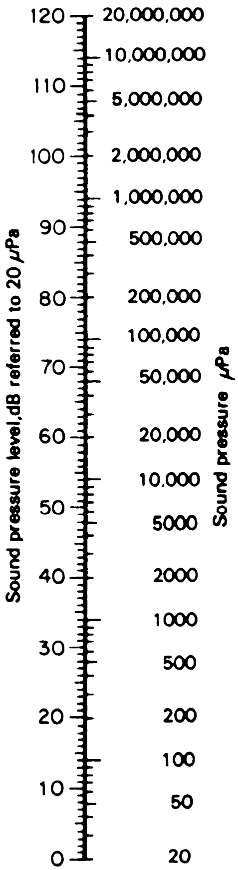

FIGURE 20.1.1 Relation between sound pressure and sound-pressure level.[1]

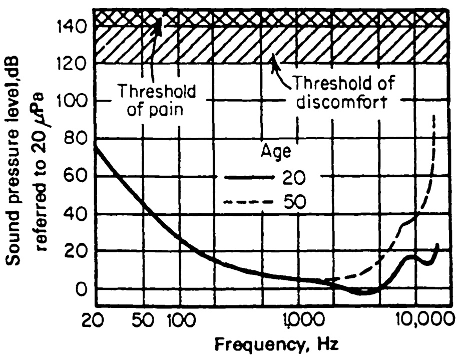

FIGURE 20.1.2 The auditory area.

Audible Frequency Range

The international abbreviation Hz (hertz) is now used (instead of the former cps) for audible frequencies as well as the rest of the frequency domain. The limits of audible frequency are only approximate because tactile sensations below 20 Hz overlap aural sensations above this lower limit. Moreover, only young listeners can hear pure sounds near or above 20 kHz, the nominal upper limit.

Frequencies beyond both limits, however, have significance to audio-electronics engineers. For example, near-infrasonic (below 20 Hz) sounds are needed for classical organ music but can be noise in turntable rumble. Near-ultrasonic (above 20 kHz) intermodulation in audio circuits can produce undesirable difference-frequency components, which are audible.

The *audible sound-pressure level range* can be combined with the audible frequency range to describe an *auditory area*, shown in Fig. 20.1.2. The lowest curve shows the weakest audible sound-pressure level for listening with both ears to a pure tone while facing the sound source in a free field. The minimum level depends greatly on the frequency of the sound. It also varies somewhat among listeners. The levels that quickly produce discomfort or pain for listeners are only approximate, as indicated by the shaded and cross-hatched areas of Fig. 20.1.2. Extended exposure can produce temporary (or permanent) loss of auditory area at sound-pressure levels as low as 90 dB.

Wavelength effects are of great importance in the design of sound systems and rooms because wavelength varies over a 3-decade range, much wider than is typical elsewhere in electronics engineering. Audible sound waves vary in length from 1 cm to 15 m. The dimensions of the sound sources and receivers used in electroacoustics also vary greatly, e.g., from 1 cm to 3 m.

Sound waves follow the principles of geometrical optics and acoustics when the wavelength is very small relative to object size and pass completely around obstacles much smaller than a wavelength. This wide range of physical effects complicates the typical practical problem of sound production or reproduction.

Loudness Level

The simple, direct method for determining experimentally the loudness *level* of a sound is to match its observed loudness with that of a 1000-Hz sinewave reference tone of calibrated, variable sound-pressure level. (Usually this is a group judgment, or an average of individual judgments, in order to overcome individual observer differences.)

When the two loudnesses are matched, the *loudness level* of the sound, expressed in *phons*, is defined as numerically equal to the sound-pressure level of the reference tone in decibels. For example, a series of observers, each listening alternately to a machine noise and to a 1000-Hz reference tone, judge them (on the average) to be equally loud when the reference tone is adjusted to 86 dB at the observer location. This makes the loudness level of the machine noise 86 phons.

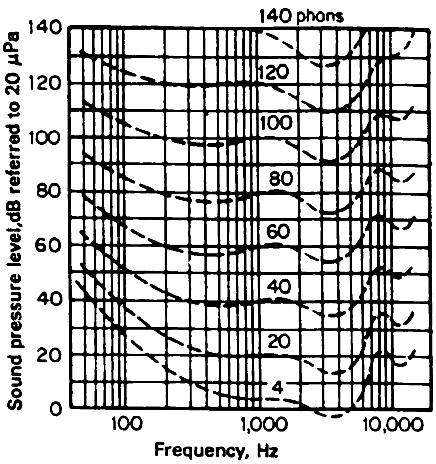

FIGURE 20.1.3 Equal-loudness-level contours.

At 1000 Hz the decibel and phon levels are numerically identical, by definition. However, at other frequencies sinewave tones may have numerically quite different sound- and loudness-levels, as seen in Fig. 20.1.3. The dashed contour curves show the decibel level at each frequency corresponding to the loudness level identifying the curve at 1000 Hz. For example, a tone at 80 Hz and 70 dB lies on the contour marked 60 phons. Its sound level must be 70 dB for it to be as loud as a 60-dB tone at 1000 Hz. Such differences at low frequencies, especially at low sound levels, are a characteristic of the sense of hearing. The fluctuations above 1000 Hz are caused by sound-wave diffraction around the head of the listener and resonances in his ear canal. This illustrates how human physiological and psychological characteristics complicate the application of purely physical concepts.

Since loudness level is related to 1000-Hz tones defined physically in magnitude, the loudness-level scale is not really psychologically based. Consequently, although one can say that 70 phons is louder than 60 phons, one cannot say *how much* louder.

Loudness

By using the phon scale to overcome the effects of frequency, psychophysicists have developed a true loudness scale based on numerous experimental procedures involving relative-loudness judgments. *Loudness*, measured in *sones*, has a direct relation to loudness level in phons, which is approximated in Fig. 20.1.4. (Below 30 phons the relation changes slope. Since few practical problems require that range, it is omitted for simplicity.) A loudness of 1 sone has been defined equivalent to a loudness level of 40 phons. It is evident in Fig. 20.1.4 that a 10-phon change doubles the loudness in sones, which means *twice as loud*. Thus a 20-phon change in loudness level quadruples the loudness.

Another advantage of the sone scale is that the loudness of components of a complex sound are additive on the sone scale as long as they are well separated on the frequency scale. For example (using Fig. 20.1.4), two tonal components at 100 and 4000 Hz having loudness levels of 70 and 60 phons, respectively, would have individual loudnesses of 8 and 4 sones, respectively, and a total loudness of 12 sones.

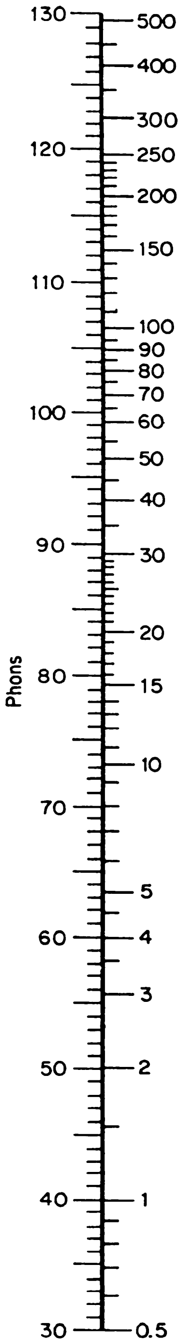

FIGURE 20.1.4 Relation between loudness in sones and loudness level in phons.[3]

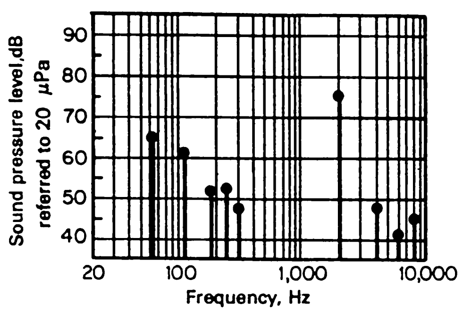

FIGURE 20.1.5 Typical line spectrum.[1]

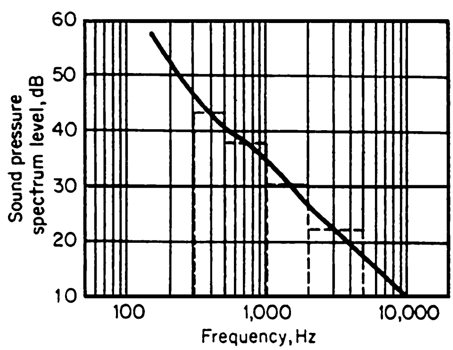

FIGURE 20.1.6 Continuous-level spectrum curve for a motor and blower.[1]

Detailed loudness computation procedures have been developed for highly complex sounds and noises, deriving the loudness in sones directly from a complete knowledge of the decibel levels for individual discrete components or noise bands. The procedures continue to be refined.

TYPICAL FORMATS FOR SOUND DATA

Sound and audio electronic data are frequently plotted as functions of frequency, time, direction, distance, or room volume. Frequency characteristics are the most common, in which the ordinate may be sound pressure, sound power, output-input ratio, percent distortion, or their logarithmic-scale (level) equivalents.

Sound Spectra

The frequency spectrum of a sound is a description of its resolution into components of different frequency and amplitude. Often the abscissa is a logarithmic frequency scale or a scale of octave (or fractional-octave) bands with each point plotted at the geometric mean of its band-limiting frequencies. Usually the ordinate scale is sound-pressure level. Phase differences are often ignored (except as they affect sound level) because they vary so greatly with measurement location, especially in reflective environments.

Line spectra are bar graphs for sounds dominated by discrete frequency components. Figure 20.1.5 is an example.

Continuous spectra are curves showing the distribution of sound-pressure level within a frequency range densely packed with components. Figure 20.1.6 is an example. Unless stated otherwise, the ordinate of a continuous-spectrum curve, called *spectrum level*, is assumed to represent sound-pressure level for a band of 1-Hz width. Usually level measurements L_{band} are made in wider bands, then converted to spectrum level L_{ps} by the bandwidth correction

$$L_{ps} = L_{\text{band}} - 10 \log(f_2 - f_1) \text{ dB} \tag{2}$$

in which f_1 and f_2 are the lower- and upper-frequency limits of the band.

When a continuous-spectrum curve is plotted automatically by a level recorder synchronized with a heterodyning filter or with a sequentially switched set of narrow-bandpass filters, any effect of *changing* bandwidth on curve slope must be considered.

Combination spectra are appropriate for many sounds in which strong line components are superimposed over more diffuse continuous spectral backgrounds. Bowed or blown musical tones and motor-driven fan noises are examples.

Octave spectra, in which the ordinate is the sound-pressure level for bands one octave wide, are very convenient for measurements and for specifications but lack fine spectrum detail.

Third-octave spectra provide more detail and are widely used. One-third of an octave and one-tenth of a decade are so nearly identical that substituting the latter for the former is a practical convenience, providing a 10-band pattern that repeats every decade. Placing third-octave band zero at 1 Hz has conveniently made the band numbers equal 10 times the logarithm (base 10) of the band-center frequency; e.g., band 20 is at 100 Hz and band 30 at 1000 Hz.

Visual proportions of spectra (and other frequency characteristics) depend on the ratio of ordinate and abscissa scales. There is no universal or fully standard practice, but for ease of visual comparison of data and of specifications, it has become rather common practice in the United States for 30 dB of ordinate scale to equal (or slightly exceed) 1 decade of logarithmic frequency on the abscissa. Available audio and acoustical graph papers and automatic level-recorder charts have reinforced this practice. When the entire 120-dB range of auditory area is to be included in the graph, the ordinate is often compressed 2:1.

Response and Distortion Characteristics

Output-input ratios versus frequency are the most common data format in audio-electronics engineering. The audio-frequency scale (20 Hz to 20 kHz) is usually logarithmic. The ordinate may be sound- or electric-output level in decibels as the frequency changes with a constant electric or sound input; or it may be a ratio of the output to input (expressed in decibels) as long as they are linearly related within the range of measurement.

When the response-frequency characteristic is measured with the input frequency filtered from the output, a distortion-frequency characteristic is the result. It can be further filtered to obtain curves for each harmonic if desired.

Directional Characteristics

Sound sources radiate almost equally in all directions when the wavelength is large compared to source dimensions. At higher frequencies, where the wavelength is smaller than the source, the radiation becomes quite directional.

Time Characteristics

Any sound property can vary with time. It can build up, decay, or vary in magnitude periodically or randomly. A reverberant sound field decays rather logarithmically. Consequently the sound level in decibels falls linearly when the time scale is linear. The rate of decay in this example is 33 dB/s.

REFERENCES

1. Harris, C. M. (ed.), "Handbook of Acoustical Measurements and Noise Control," Chaps. 1 and 2, McGraw-Hill, 1991.
2. Acoustical Terminology (Including Mechanical Shock and Vibration), S1.1–1994, Acoustical Society of America, 1994.
3. ANSI Standard S3.4-1980 (R1986).