

## 4 Sound Waves and Sine Waves

John Pierce

### 4.1 Sound and Sine Waves

We are immersed in an ocean of air. Physical disturbances—snapping the fingers, speaking, singing, plucking a string, or blowing a horn—set up a vibration in the air around the source of sound. A sound wave travels outward from the source as a spherical wavefront. It is a *longitudinal* wave in which the pulsating motion of the air is in the direction the wave travels. In contrast, waves in a stretched string are *transverse* waves, for the motion of the string is at right angles to the direction in which the wave travels.

How fast does a sound wave travel? If the air temperature is 20 degrees Celsius, a sound wave travels at a velocity of 344 meters (1128 feet) a second—a little faster at higher temperatures and a little slower at lower temperatures. Sound travels in helium almost three times as fast as in air, and longitudinal sound waves can travel through metals and other solids far faster.

The sound waves that travel through the air cause components of our ears to vibrate in a manner similar to those of the sound source. What we hear grows weaker with distance from the source, because the area of the spherical wavefront increases as the square of the distance from the source, and the power of the source wave is spread over that increasing surface. What actually reaches our ears is complicated by reflections from the ground and other objects. In a room, much of the sound we hear comes to our ears after being reflected from floor, walls, and ceiling.

The vibrations of musical sound are complicated, and the charm of musical sounds lies in their complexity. But most old-time discussions of musical sounds and most old-time experiments with sound waves and with hearing were carried out with a rare and simple sort of sound wave, a sinusoidal wave. How can such discussions and experiments have any relevance to the complicated sounds of music? Chiefly, because the phenomenon of sound propagation in air at intensities encountered in musical performance is a *linear* phenomenon. The undisturbed vibrations of strings or of columns of air are at least approximately linear. Even the vibrations

along the cochlea of the ear are close enough to linear for linear systems ideas to be an appropriate guide to thought. What are the characteristics of linear systems? How can sine waves be useful in connection with the complicated sounds of music?

#### 4.2 Linear Systems

Sine waves are important both mathematically and practically in describing the behavior of *linear systems*.

What is a linear system? The amplifier depicted in figure 4.1 illustrates a linear system. Suppose that an input signal or waveform  $In_1$  produces an output waveform  $Out_1$ , and that an input waveform  $In_2$  produces an output waveform  $Out_2$ . If the amplifier is linear, the combined input waveform  $In_1+In_2$  will produce an output waveform  $Out_1+Out_2$ . The output of a linear amplifier (or of any linear system or phenomenon) for a sum of inputs is the sum of the outputs produced by the inputs separately.

It may be easier to understand if we say that an amplifier is linear if it doesn't produce any distortion. In some real amplifiers there is distortion. We hear things in the output that were not present in the input. Mathematically, a linear system is a system whose behavior is described by a linear differential equation or by a linear partial differential equation. In such an equation the sum of constants times partial derivatives with respect to time and space is equal to 0, or to an input driving function. Some linear, or approximately linear, systems are the following:

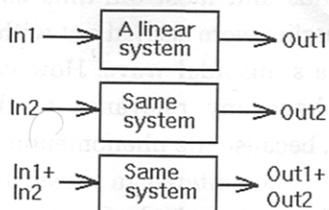
*A sound wave in air* (linear for musical intensities)

*A vibrating string* (linear for small amplitudes)

*A vibrating chime or bell* (ordinarily linear)

*The bones of the middle ear* (linear for small changes in level)

*Vibrations along the basilar membrane of the cochlea* (with some assumptions).



**Figure 4.1** A system is linear if the output due to two overlapping inputs is the sum of the outputs to each input separately.

Tam-tams, cymbals, and some other percussion instruments exhibit a clearly nonlinear phenomenon: an upwelling of high frequencies after striking. Smaller nonlinearities in nearly all musical instruments are responsible for subtle but characteristic musical qualities of the sounds produced. But the most obvious features of the sounds of conventional instruments are consistent with an assumption of linearity.

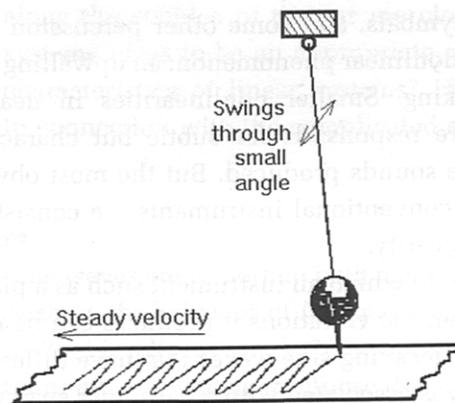
To the degree to which an instrument such as a piano, guitar, bell, or gong is linear, the vibrations it produces can be represented as a sum of slowly decaying sine waves that have different frequencies. Each frequency is associated with a particular spatial distribution of vibrations and has a particular rate of decay. The sound of the wave generated by such a sum of vibrations at different frequencies constitutes a musical tone.

The frequencies of free vibrations of a violin string or the air in a horn predispose the forced (by bowing or blowing) vibrations to have frequencies quite close to those of a free vibration. Skillful bowing of a violin string can give *harmonics*, which are integer multiples of some fundamental frequency. A bugle can be blown so as to produce a rather small number of musical tones, each near a frequency of the free vibration of the air in the tube, again a series of harmonics.

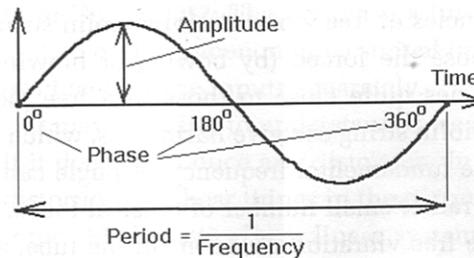
### 4.3 Sine Waves

Because sine waves, and measurements based on sine waves, are pervasive in musical lore, it is important at this point to become well acquainted with a sine wave. Figure 4.2 shows a swinging pendulum that traces out a portion of a sine wave on a moving strip of paper. A true sine wave lasts forever, with its past, present, and future an endless repetition of identical periods or cycles of oscillation. A sine wave can be characterized or described completely by three numbers: the maximum amplitude (in centimeters, volts, sound pressure, or some other unit of measurement), the frequency in Hertz (Hz, cycles per second), and the phase, which specifies the position when the sine wave reaches its peak amplitude. This is illustrated in figure 4.3.

With respect to phase, we should note that the mathematical cosine function is at its peak when the phase is 0 degrees, 360 degrees, 720 degrees, and so on. The mathematical sine function reaches its peak at 90 degrees, 450 degrees, and so on.



**Figure 4.2** A swinging pendulum traces out a sine wave.



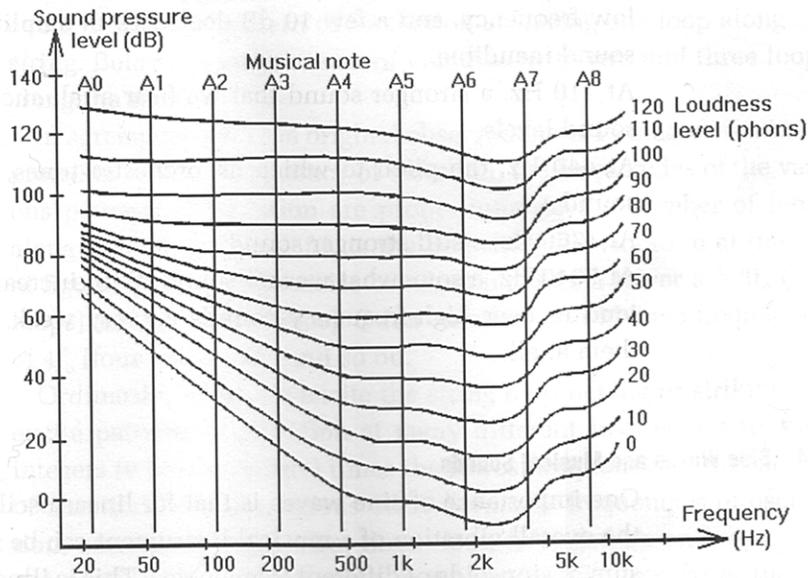
**Figure 4.3** Sine waves are described completely by their frequency (or period), amplitude, and phase.

The relative amplitudes of sine waves are often expressed in terms of *decibels (dB)*. If wave 1 has a peak amplitude of vibration of  $A_1$  and a reference wave or vibration has a peak amplitude of vibration of  $A_2$ , the relationship in decibels of vibration  $A_1$  to vibration  $A_2$  is given by

$$20 \log_{10} (A_1/A_2). \quad (4.1)$$

A sound level in decibels should always be given as decibels above some reference level. Reference level is often taken as a sound power of a millionth of a millionth of a watt per square meter. A person with acute hearing can hear a 3000 Hz sine wave at reference level. Reference level is sometimes also taken as a sound pressure of 0.00005 newtons, which is almost exactly the same reference level as that based on watts per square meter.

In many experiments with sound we listen to sounds of different frequencies. It seems sensible to listen to sinusoidal sound waves in an orderly fashion. We will use the diagram shown in figure 4.4 to



**Figure 4.4** The equal loudness curves link combinations of sound pressure level and frequency that are heard as equally loud.

guide our listening. Here frequency in Hertz is plotted horizontally. Nine vertical lines are shown, spaced one octave apart at frequencies from 27.5 Hz (A0), the pitch frequency of the lowest key on the piano keyboard, to 7040 Hz (A8), above the topmost piano key.

The curves shown are equal loudness curves. Along a particular loudness curve the various combinations of frequency and level give sounds that are judged to have the same loudness. The constant loudness curves crowd together at low frequencies. At low frequencies, a small change in amplitude results in a large change in loudness. There is some crowding together at about 4000 Hz.

We can listen to tones at a chosen frequency given by one of the vertical lines at six different amplitudes, each successively 10 dB below the preceding amplitude. This tells us how a sinusoidal sound of a given frequency sounds at six sound levels 10 dB apart. Of course, the sine wave sounds fainter with each 10 dB decrease in amplitude. What we hear depends on the initial sound level, and that depends on the audio equipment and its loudness setting. But, roughly, this is what we hear:

At 27.5 Hz, a weak sound that disappears after a few 10 dB falls in level. The constant loudness curves are crowded together at this

low frequency, and a few 10 dB decreases in amplitude render the sound inaudible

At 110 Hz, a stronger sound that we hear at all successively lower sound levels

At 440 Hz, the pitch to which an orchestra tunes, a still stronger sound

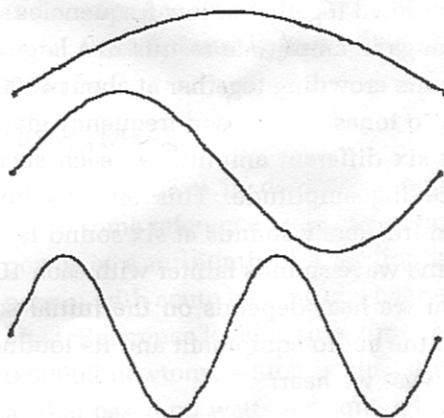
At 1760 Hz, a still stronger sound

At 7040 Hz, a somewhat weaker sound. With increasing age people tend to hear high-frequency sounds as very weak, or not to hear them at all.

#### 4.4 Sine Waves and Musical Sounds

One importance of sine waves is that for linear oscillating systems, the overall vibration of a musical instrument can be regarded as the sum of sinusoids of different frequencies. This is illustrated in figure 4.5, which shows several patterns of oscillation of a vibrating string.

In the vibration at the top, the deviation of the string from straightness varies sinusoidally with distance along the string. The center of the string vibrates up and down, with a sinusoidal displacement as a function of time, and the oscillation falls smoothly to 0 at the ends. At any instant the variation of displacement with distance along the string is sinusoidal. We can think of the oscillation of the string as corresponding to a traveling sine wave of twice the length of the string, reflected at the fixed ends of the string. We



**Figure 4.5** Some modes of vibrations of a stretched string. Different modes have different numbers of loops: from top to bottom, here, one, two, three. The frequencies of vibration are proportional to the number of loops.

can describe this pattern of oscillation as having one loop along the string. Below we see patterns of vibration with two and three loops along the string.

In agreement with the original observations of Pythagoras as interpreted in terms of frequency of vibration, the frequencies of the various patterns of vibration are proportional to the number of loops along the string. Thus, if  $f_0$  is the frequency for vibration at the top of figure 4.5, the frequencies of vibration shown lower are  $2f_0$  (two loops) and  $3f_0$  (three loops). Other "modes" would have frequencies of  $4f_0$  (four loops),  $5f_0$ , and so on.

Ordinarily, when we excite the string by plucking or striking, we excite patterns of vibration at many different frequencies that are integers (whole numbers) times the lowest frequency. In one period of duration,  $1/f_0$ , the various other *harmonic* frequencies of oscillation, corresponding to two, three, four, five, and so on loops, will complete two, three, four, five, and so on oscillations. After the period  $1/f_0$ , the overall oscillation will repeat again, "endlessly" in an ideal case of no decay in amplitude.

We have considered various aspects of sine waves that we hear. Wavelength is an aspect of sinusoidal sound that is associated with a sound wave traveling through air. The wavelength of a sinusoidal sound is sound velocity divided by frequency. As noted in section 4.1, the velocity of sound in air is 344 meters/second (1128 feet/second). In table 4.1, wavelength is tabulated for various frequencies (and musical pitches). We see that in going from the lowest key on the piano, A0 (frequency 27.5 Hz) to A7 (the highest A on the keyboard (frequency 3520 Hz), the wavelength goes from 41 feet (12.5 meters) to 0.32 foot (0.1 meter). Actual musical tones include harmonics whose wavelengths are much shorter than that of the fundamental or pitch frequency.

For some musical instruments (including some organ pipes and the clarinet), the sounds produced contain chiefly odd harmonics of a fundamental frequency. This happens whenever one end of a tube is closed and the other end is open. If  $f_0$  is the fundamental frequency of a closed organ pipe, the chief frequencies present are  $f_0$ ,  $3f_0$ ,  $5f_0$ ,  $7f_0$ , and so on.

We can represent the sustained sound of a musical instrument by a sum of sine waves with many harmonic frequencies. But we hear the sound as a single musical tone with a pitch that is given by the *pitch frequency*, the frequency of which the frequencies of all the partials are integer multiples. The pitch of a musical sound de-

**Table 4.1** Musical notes, frequencies, and wavelengths

NOTE NAME	FREQUENCY (HZ)	WAVELENGTH (FT.)
A0	27.5	41
A1	55	20.5
A2	110	10.25
A3	220	5.1
A4	440	2.56
A5	880	1.28
A6	1760	0.64
A7	3520	0.32

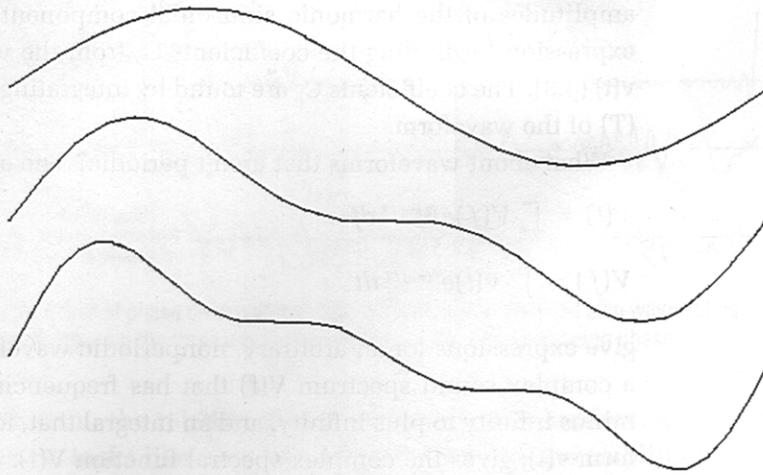
depends on the simple harmonic relation among the many frequencies present. The musical quality of the overall sound depends in part on the relative intensities of the various harmonics, and in part on how they are excited initially (the *attack* quality of the sound). (We will discuss the topics of pitch and quality further in later chapters.)

#### 4.5 Fourier Analysis

Most musical instruments produce sounds that are nearly periodic. That is, one overall cycle of the waveform repeats, or nearly repeats, over and over again. Looking at this in another way, traditional musical tones, of the voice or of an instrument, are periodic, or nearly periodic. Hence, it is pertinent to consider the general qualities of periodic sounds. Any periodic waveform can be approximated by a number of sinusoidal components that are *harmonics* of a *fundamental frequency*. That fundamental frequency may or may not be present in the sound. It is the reciprocal of the period of the waveform measured in seconds.

This is illustrated in figure 4.6 by three approximations of a sawtooth waveform. In approximating a sawtooth waveform we add harmonic-related sine waves whose frequencies are  $f_0$ ,  $2f_0$ ,  $3f_0$  and so on, and whose amplitudes are inversely proportional to the frequencies. Three sine waves give a very poor approximation to a sawtooth. A better approximation is given by 6 sinusoidal components, and a still better approximation by 12.





**Figure 4.6** Representation of a sawtooth wave as the sum of one, two, and three sinusoids.

A true sawtooth waveform is a succession of vertical and slanting straight-line segments. A Fourier series approximation to a true sawtooth waveform that uses a finite number of harmonically related sine waves differs from the sawtooth waveform in two ways. In a gross way, the approximation gets better and better as we include more and more terms. But there is a persistent wiggle whose amplitude decreases but whose frequency increases as we add more and more terms. We will see later that the ear can sometimes hear such a wiggle, as well as a pitch associated with a true sawtooth waveform. Remember, from the equal loudness contours, that we can hear only up to a given frequency, so if we add enough harmonic sinusoids to our approximation of any wave, we can get perceptually as close as we like.

A fitting of the sum of harmonically related sine waves to a periodic waveform is called *Fourier analysis*.

Mathematically, the Fourier series transform is defined by the equations

$$v(t) = \sum_{n=-\infty}^{\infty} C_n e^{j2\pi nt/T} \quad (4.2)$$

$$C_n = \frac{1}{T} \int_{-T/2}^{T/2} v(t) e^{j2\pi nt/T} dt. \quad (4.3)$$

These describe the representation of a periodic time waveform  $v(t)$  in terms of complex coefficients  $C_n$  that represent the phases and

amplitudes of the harmonic sinusoidal components (4.2), and the expression for finding the coefficients  $C_n$  from the waveform signal  $v(t)$  (4.3). The coefficients  $C_n$  are found by integrating over the period ( $T$ ) of the waveform.

What about waveforms that aren't periodic? The equations

$$v(t) = \int_{-\infty}^{\infty} V(f)e^{j2\pi nt/T} df \quad (4.4)$$

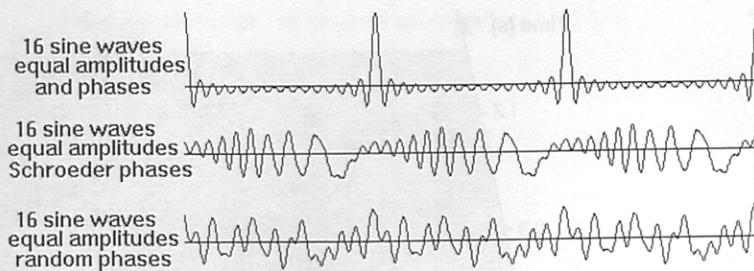
$$V(f) = \int_{-\infty}^{\infty} v(t)e^{j2\pi nt/T} dt \quad (4.5)$$

give expressions for an arbitrary, nonperiodic waveform in terms of a complex sound spectrum  $V(f)$  that has frequencies ranging from minus infinity to plus infinity, and an integral that, for a given waveform  $v(t)$ , gives the complex spectral function  $V(f)$ . Such an overall resolution of a complete waveform into a spectrum is of limited use in connection with music. For example, we could in principle find the spectrum of a complete piece of music. This would tell us very little that we would care to know. Today, most Fourier analyses of waveforms are performed by computer programs, using a discrete definition of the Fourier transform.

It is important to note that a waveform, that is, a plot one cycle long of amplitude versus time, is a complete description of a periodic waveform. A spectrum gives a complete description of a waveform, consisting of two numbers for each single frequency. These two numbers can describe the real and imaginary parts of a complex number, or they can describe the amplitude and phase of a particular frequency component. Conversion back and forth from complex number representation to amplitude and phase representation is accomplished simply.

In plots of spectra of sound waves, the phase of the spectral components is seldom displayed. What is plotted against frequency is usually how the amplitude varies with frequency. The amplitude is often given in decibels. Or the square of the amplitude is plotted versus frequency. This is called a power spectrum.

Is the phase of a Fourier component important? Figure 4.7 shows 4 periods of waveforms made up of 16 sinusoidal components with harmonic frequencies ( $f_0, 2f_0, 3f_0, \text{ etc.}$ ) having equal amplitudes but different phases. The waveforms look very different. The top-most waveform is a sequence of narrow spikes with wiggles in between. In the center waveform the phases have been chosen so as to make each repeated cycle of the waveform look like a sinusoid of decreasing frequency, also called a chirp. In the waveform at the



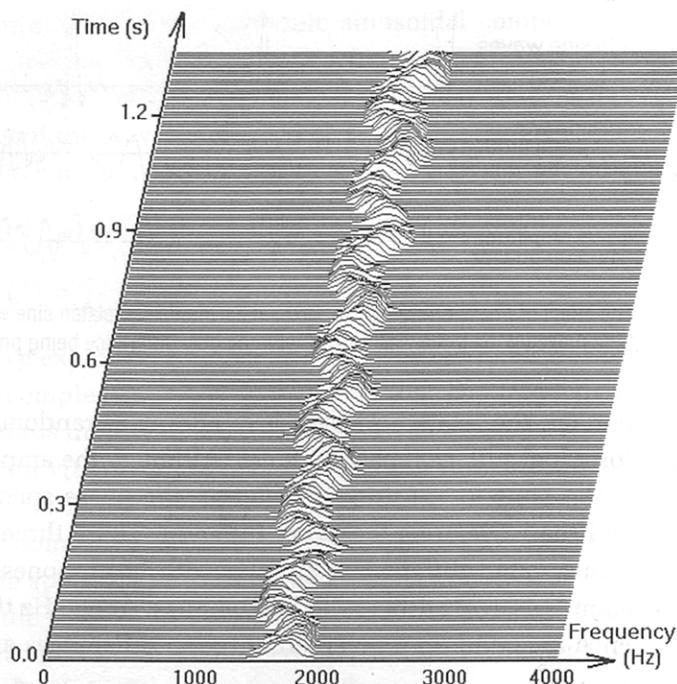
**Figure 4.7** The effect of phase on waveform. Sixteen harmonically related sine waves of equal amplitude make up the three waveforms, with the only difference being phase.

bottom, the relative phases were chosen at random, and the waveform looks like a repeating noise. Although the amplitude spectrum is the same for all three waveforms, the phase spectra are different and the waveforms look very different. These three different waveforms sound different at 27.5 Hz with headphones. At 220 Hz the sounds scarcely differ with headphones. At 880 Hz there is no difference in sound. In a reverberant room, differences are small even at 27.5 Hz.

Partly because we don't listen through headphones, and partly because most pitches are higher than 27.5 Hz, most plots of spectra take no account of phase.

It can be important to know how the frequency content of a musical sound changes with time. Many sustained musical sounds have small, nearly periodic changes of amplitude (tremolo) or of frequency (vibrato). And there are attack and decay portions of musical sounds. As an example of the importance of this, Jean-Claude Risset and Max Mathews found in 1969 that in the sounds of brassy instruments, the higher harmonics rise later than the lower harmonics. This is useful, indeed necessary, in synthesizing sounds with a brassy timbre. How can we present a changing spectrum in a way that is informative to the eye? One way of representing changing spectra is to plot successive spectra a little above and to the right of one another, so as to give a sense of perspective in time. Figure 4.8 shows successive spectra of a sine wave with a little vibrato that shifts the peak a little to the left, then back, repeating this pattern periodically.

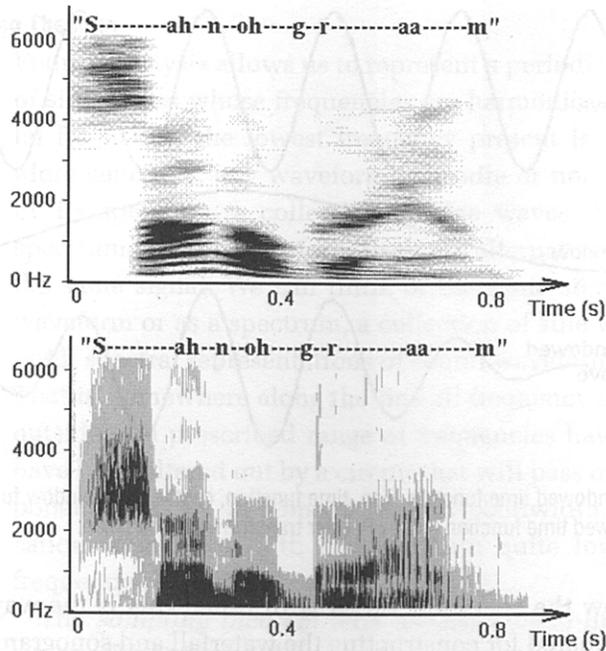
There is another way of representing changing spectra, a representation by sonograms (also called spectrograms). This is particularly valuable in studying very complicated sounds such as speech. A sonogram of speech is shown in figure 4.9. The amplitude at a given



**Figure 4.8** A "waterfall spectrum" representation of a sinusoidal sound with a slow sinusoidal variation of frequency with time.

frequency is represented by darkness (pure white represents zero amplitude). Distance from the bottom represents frequency. Time is plotted left to right. The two sonograms are of the same speech sound. In the upper sonogram, resolution is good in the frequency direction—we can see individual harmonic tracks—but it is blurred in the time direction. In the lower sonogram the resolution is good in the time direction—we can see individual pitch periods representing the vibrations of the vocal folds—but it is fuzzy in the frequency direction. Resolution can't be sharp in both directions. If we want precise pitch, we must observe the waveform for many periods. If we want precise time, we must observe the waveform for only part of a period. In general, the product of resolution in frequency and resolution in time is constant. This is a mathematical limitation that has nothing to do with the nature of the sound source.

Fourier analysis, the representation of a periodic waveform in terms of sine waves, is an essential tool in the study of musical sound. It allows us to determine the frequency components of a sound and to determine how those components change with time. Is the waveform or the spectrum better? If you are looking for a weak

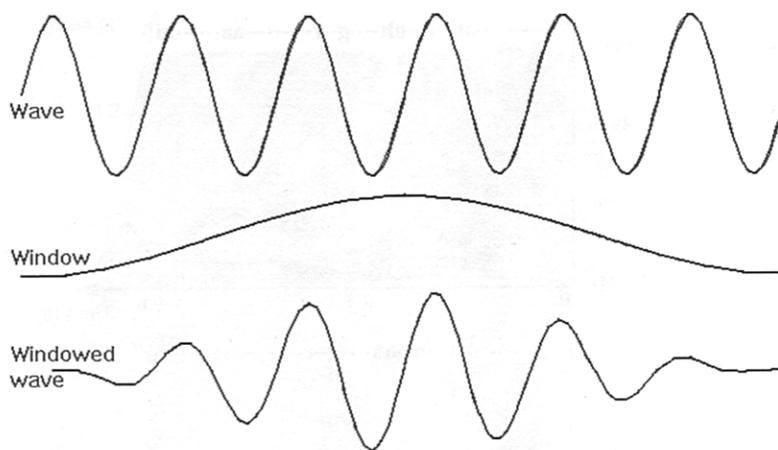


**Figure 4.9** Spectrograms in which amplitude or intensity is represented by degree of darkness.

reflection following a short sound (as in radar), the waveform is better. But suppose you want to find the sound of a tin whistle in the midst of orchestral noise. You may have a chance with a spectral analysis that sharply separates sound energy in frequency. You won't have a chance by just looking at the waveform. So both waveforms and spectra are legitimate and useful ways of depicting sounds.

What we actually do in Fourier analysis of musical sounds is to use a computer program, called a fast Fourier transform (FFT). The analysis produces a spectrum that gives both amplitude and phase information, so that the waveform can be reconstructed from the spectrum obtained. Or the amplitude alone can be used in a spectral plot. Of an actual sound wave, we take the spectrum of a selected or *windowed* portion of a musical sound that may be several periods long.

Figure 4.10 illustrates the process of windowing. At the top are a few periods of a sine wave. In the center is a windowing function. This is multiplied by the overall waveform to give the windowed portion of the waveform, shown at the bottom. In analyzing the waveform, a succession of overlapping windows is used to find out



**Figure 4.10** Windowed time function. Top, time function, center, time window function, bottom, windowed time function whose Fourier transform is to be taken.

how the spectrum varies with time. This is the way the data were prepared for constructing the waterfall and sonogram plots of figures 4.8 and 4.9.

A strict reconstruction of the waveform from the spectrum obtained from any one window would repeat over and over again, but such a reconstruction is never made. In constructing the variation of the spectrum with time, or in reconstructing the waveform from the spectra of successive windowed waveforms, each windowed waveform is limited to the time duration of the window. Such a reconstruction necessarily goes to 0 at the ends of a particular window, where the window and the windowed waveform go to 0. The analysis of a succession of overlapping windowed waveforms makes it possible to construct an overall spectrum that varies with time, and from this overall spectrum the waveform itself can be reconstructed.

Fourier analysis is a mathematical verity. It is useful in connection with musical tones because the ear sorts sounds into ranges of frequency, and tampering with the sound spectrum has clear effects on what we hear and identify. Consider the sound of a human voice. If we remove or filter out low frequencies, the sound becomes high and tinny, but its musical pitch does not change. If we filter out the high frequencies, the voice becomes dull. Nonperiodic fricative (noise) components of a sound are identified through the higher frequencies of their spectra. If we filter out the higher frequencies, we can't tell *f* (as in fee) from *s* (as in see).

#### 4.9 Closing Thoughts

Sine waves and Fourier analysis are powerful resources in all studies of musical sounds. Most musical sounds are produced by linear systems, or by approximately linear systems (freely vibrating strings or columns of air), or by forced vibrations of such systems that reflect their linear nature. Parts of our organs of perception are approximately linear. Higher-level processing, while not linear in itself, reflects the approximately linear processing that is carried out at lower levels.

Analyses based on sine waves are of crucial value in understanding the nature of musical sounds and of their perception. Sine waves, however, are not music. Nor are they even musical sounds by themselves. Rather, they are ingredients out of which musical sounds can be concocted, or through which musical sounds can be analyzed and studied. What we do with sine waves, how we use them, must be guided by the capabilities and limitations of human perception.

#### References

- Bracewell, R. N. (1986). *The Fourier Transform and Its Applications*. Second edition. New York: McGraw-Hill. An advanced book on Fourier analysis.
- Chui, C. K. (1992). *An Introduction to Wavelets*. Boston: Academic Press.
- , ed. (1992). *Wavelets: A Tutorial in Theory and Applications*. Boston: Academic Press. Gives more sense of actual application.
- Risset, J. C., and M. V. Mathews. (1969). "Analysis of Musical Instrument Tones." *Physics Today*, 22(2): 23–40.
- Schafer, R. W., and J. D. Markel, eds. (1979). *Speech Analysis*. New York: IEEE Press. Contains many early papers on speech and on the phase and channel vocoders.
- Steiglitz, K. (1996). *A Digital Signal Processing Primer*. Menlo Park, CA: Addison Wesley. An excellent introductory reference to Fourier analysis, sinusoids, and linear systems.