

CHAPTER 1

Basic Acoustics and Acoustic Filters

1.1 The sensation of sound

Several types of events in the world produce the sensation of sound. Examples include doors slamming, violins, wind, and human voices. All these examples, and any others we could think of, involve movement of some sort. And these movements cause **pressure fluctuations** in the surrounding air (or some other acoustic medium). When pressure fluctuations reach the eardrum, they cause it to move, and the auditory system translates these movements into neural impulses which we experience as sound. Thus, sound is produced when pressure fluctuations impinge upon the eardrum. An acoustic waveform is a record of

Acoustic medium

Normally the pressure fluctuations that are heard as sound are produced in air, but it is also possible for sound to travel through other acoustic media. So, for instance, when you are swimming under water, it is possible to hear muffled shouts of the people above the water, and to hear noise as you blow bubbles in the water. Similarly, gases other than air can transmit pressure fluctuations that cause sound. For example, when you speak after inhaling helium from a balloon, the sound of your voice travels through the helium, making it sound different from normal. These examples illustrate that sound properties depend to a certain extent on the acoustic medium, on how quickly pressure fluctuations travel through the medium, and how resistant the medium is to such fluctuations.

sound-producing pressure fluctuations over time. (Ladefoged, 1996, and Fry, 1979, provide more detailed discussions of the topics covered in this chapter.)

1.2 The propagation of sound

Pressure fluctuations impinging on the eardrum produce the sensation of sound, but sound can travel across relatively long distances. This is because a sound produced at one place sets up a **sound wave** that travels through the acoustic medium. A sound wave is a traveling pressure fluctuation that propagates through any medium that is elastic enough to allow molecules to crowd together and move apart. The wave in a lake after you throw in a stone is an example. The impact of the stone is transmitted over a relatively large distance. The water particles don't travel; the pressure fluctuation does.

A line of people waiting to get into a movie is a useful analogy for a sound wave. When the person at the front of the line moves, a "vacuum" is created between the first person and the next person in the line (the gap between them is increased), so the second person steps forward. Now there is a vacuum between person two and person three, so person three steps forward. Eventually, the last person in the line gets to move; the last person is affected by a movement that occurred at the front of the line, because the pressure fluctuation (the gap in the line) traveled, even though each person in the line moved very little. The analogy is flawed, because in most lines you get to move to the front eventually. To be a proper analogy for sound propagation, we would have to imagine that the first person is shoved back into the second person and that this crowding or increase of pressure (like the vacuum) is transmitted down the line.

Figure 1.2 shows a pressure waveform at the location indicated by the asterisk in figure 1.1. The horizontal axis shows the passage of time, the vertical axis the degree of crowdedness (which in a sound wave corresponds to air pressure). At time 3 there is a sudden drop in crowdedness because person two stepped up and left a gap in the line. At time 4 normal crowdedness is restored when person 3 steps up to fill the gap left by person 2. At time 10 there is a sudden increase in crowdedness as person 2 steps back and bumps into person 3. The graph in figure 1.2 is a way of representing the traveling rarefaction and compression waves shown in figure 1.1. Given a uniform acoustic medium, we could reconstruct figure 1.1 from figure 1.2 (though note the discussion in the next paragraph on sound energy dissipation). Graphs like the one shown in figure 1.2 are more typical in acoustic phonetics, because this is the type of view of a sound wave that is produced by a microphone – it shows amplitude fluctuations as they travel past a particular point in space.

Sound waves lose energy as they travel through air (or any other acoustic medium), because it takes energy to move the molecules. Perhaps you have noticed a similar phenomenon when you stand in a long line. If the first person steps forward, then quickly back, only a few people at the front of the line may be affected, because people further down the line have inertia; they will tolerate

Time	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
		1	1	1	1	1	1	1							
*	1		2	2	2	2	2	2	X	1	1	1	1	1	1
	2	2		3	3	3	3	3	3	X	2	2	2	2	2
	3	3	3		4	4	4	4	4	4	X	3	3	3	3
	4	4	4	4		5	5	5	5	5	5	X	4	4	4
	5	5	5	5	5		6	6	6	6	6	6	X	5	5
	6	6	6	6	6	6		7	7	7	7	7	7	X	6
	7	7	7	7	7	7	7								7

Figure 1.1 Wave motion in a line of seven people waiting to get into a show. Time is shown across the top of the graph running from earlier (time 1) to later (time 15) in arbitrary units.

An analogy for sound propagation

Figure 1.1 shows seven people (represented by numbers) standing in line to see a show. At time 2 the first person steps forward and leaves a gap in the line. So person two steps forward at time 3, leaving a gap between the second and third persons in the line. The gap travels back through the line until time 8, when everyone in the line has moved forward one step. At time 9 the first person in the line is shoved back into place in the line, bumping into person two (this is symbolized by an X). Naturally enough, person two moves out of person one's way at time 10, and bumps into person three. Just as the gap traveled back through the line, now the collision travels back through the line, until at time 15 everyone is back at their starting points.

We can translate the terms of the analogy to sound propagation. The people standing in line correspond to air molecules, the group of them corresponding to an acoustic medium. The gap between successive people is negative air pressure, or rarefaction, and collisions correspond to positive air pressure, or compression. Zero air pressure (which in sound propagation is the atmospheric pressure) is the normal, or preferred, distance between the people standing in line. The initial movement of person one corresponds to the movement of air particles adjacent to one of the tines of a tuning fork (for example) as the tine moves away from the particle. The movement of the first person at time 9 corresponds to the opposite movement of the tuning fork's tine.

some change in pressure (distance between people) before they actually move in response to the change. Thus the disturbance at the front of the line may not have any effect on the people at the end of a long line. Also, people tend to fidget, so the difference between movement propagated down the line and inherent fidgeting

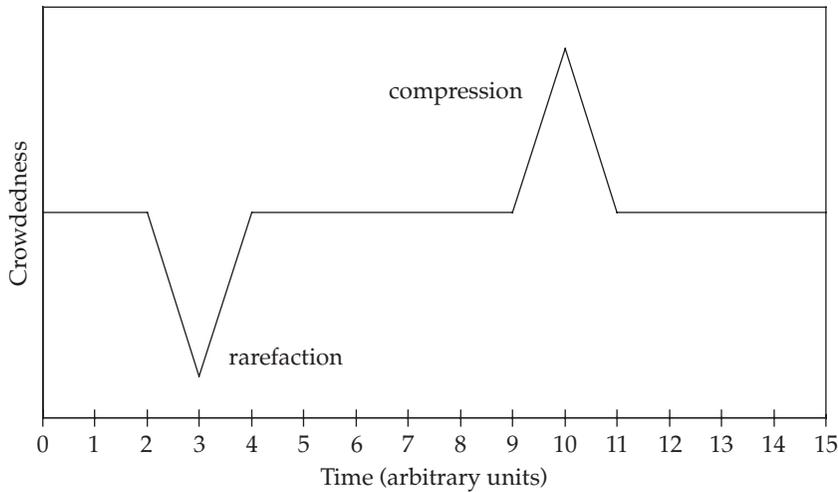


Figure 1.2 A pressure waveform of the wave motion shown in figure 1.1. Time is again shown on the horizontal axis. The vertical axis shows the distance between people.

(the signal-to-noise ratio) may be difficult to detect if the movement is small. The rate of sound dissipation in air is different from the dissipation of a movement in a line, because sound radiates in three dimensions from the sound source (in a sphere). This means that the number of air molecules being moved by the sound wave greatly increases as the wave radiates from the sound source. Thus the amount of energy available to move the molecules (energy per unit surface area on the sphere) decreases as the wave expands out from the sound source, consequently the amount of particle movement decreases as a function of the distance from the sound source (by a power of 3). That is why singers in heavy metal bands put the microphone right up to their lips. They would be drowned out by the general din otherwise. It is also why you should position the microphone close to the speaker's mouth when you record a sample of speech (although it is important to keep the microphone to the side of the speaker's lips, to avoid the blowing noises in [p]'s, etc.).

1.3 Types of sounds

There are two types of sounds: periodic and aperiodic. Periodic sounds have a pattern that repeats at regular intervals. They come in two types: simple and complex.

1.3.1 Simple periodic waves

Simple periodic waves are also called sine waves: they result from simple harmonic motion, such as the swing of a pendulum. The only time we humans get close to producing simple periodic waves in speech is when we're very young. Children's vocal cord vibration comes close to being sinusoidal, and usually women's vocal cord

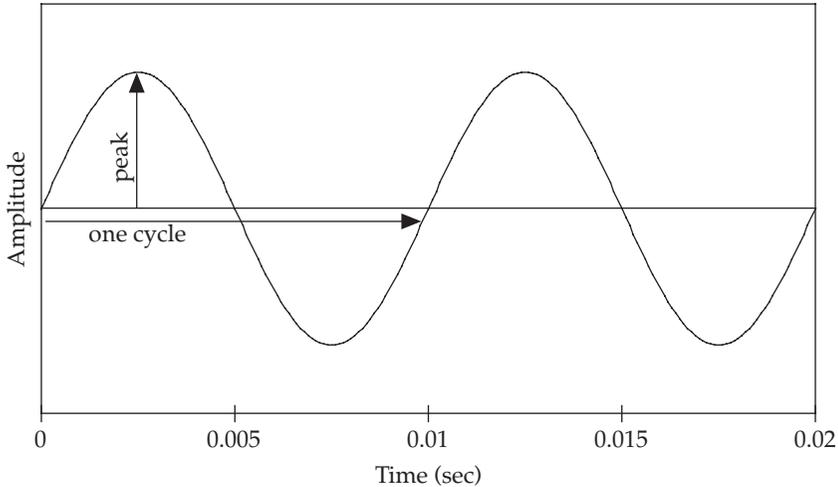


Figure 1.3 A 100 Hz sine wave with the duration of one cycle (the period) and the peak amplitude labeled.

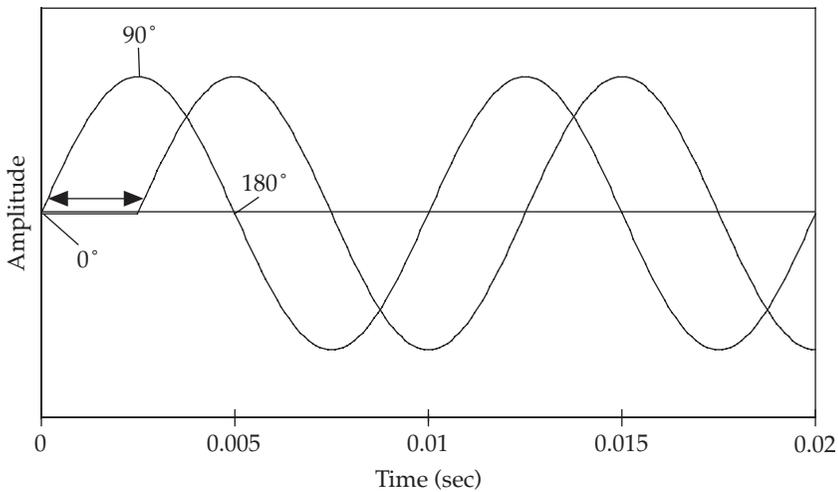


Figure 1.4 Two sine waves with identical frequency and amplitude, but 90° out of phase.

vibration is more sinusoidal than men's. Despite the fact that simple periodic waves rarely occur in speech, they are important, because more complex sounds can be described as combinations of sine waves. In order to define a sine wave, one needs to know just three properties. These are illustrated in figures 1.3–1.4.

The first is frequency: the number of times the sinusoidal pattern repeats per unit time (on the horizontal axis). Each repetition of the pattern is called a cycle, and the duration of a cycle is its period. Frequency can be expressed as cycles per second, which, by convention, is called hertz (and abbreviated Hz). So to get the frequency of a sine wave in Hz (cycles per second), you divide one second by the period (the duration of one cycle). That is, frequency in Hz equals $1/T$, where T

is the period in seconds. For example, the sine wave in figure 1.3 completes one cycle in 0.01 seconds. The number of cycles this wave could complete in one second is 100 (that is, one second divided by the amount of time each cycle takes in seconds, or $1/0.01 = 100$). So, this waveform has a frequency of 100 cycles per second (100 Hz).

The second property of a simple periodic wave is its amplitude: the peak deviation of a pressure fluctuation from normal, atmospheric pressure. In a sound pressure waveform the amplitude of the wave is represented on the vertical axis.

The third property of sine waves is their phase: the timing of the waveform relative to some reference point. You can draw a sine wave by taking amplitude values from a set of right triangles that fit inside a circle (see exercise 4 at the end of this chapter). One time around the circle equals one sine wave on the paper. Thus we can identify locations in a sine wave by degrees of rotation around a circle. This is illustrated in figure 1.4. Both sine waves shown in this figure start at 0° in the sinusoidal cycle. In both, the peak amplitude occurs at 90° , the downward-going (negative-going) zero-crossing at 180° , the negative peak at 270° , and the cycle ends at 360° . But these two sine waves with exactly the same amplitude and frequency may still differ in terms of their relative timing, or phase. In this case they are 90° out of phase.

1.3.2 Complex periodic waves

Complex periodic waves are like simple periodic waves in that they involve a repeating waveform pattern and thus have cycles. However, complex periodic waves are composed of at least two sine waves. Consider the wave shown in figure 1.5, for example. Like the simple sine waves shown in figures 1.3 and 1.4, this waveform completes one cycle in 0.01 seconds (i.e. 10 milliseconds). However,

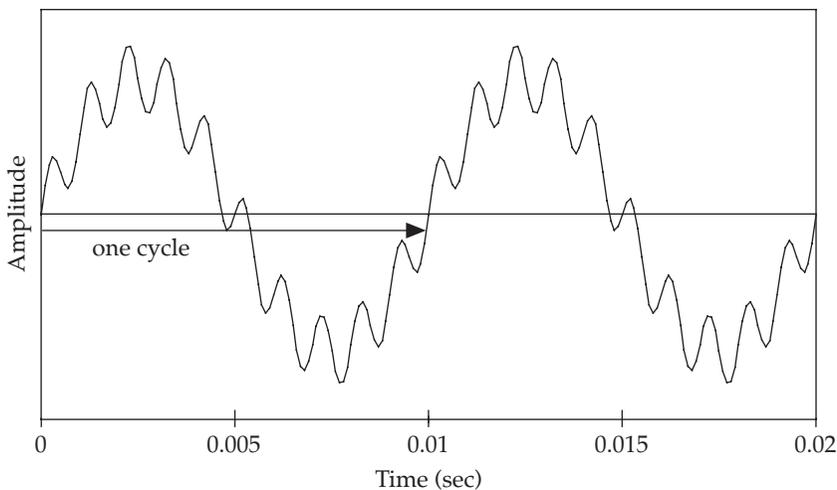


Figure 1.5 A complex periodic wave composed of a 100 Hz sine wave and a 1,000 Hz sine wave. One cycle of the fundamental frequency (F_0) is labeled.

it has an additional component that completes ten cycles in this same amount of time. Notice the “ripples” in the waveform. You can count ten small positive peaks in one cycle of the waveform, one for each cycle of the additional frequency component in the complex wave. I produced this example by adding a 100 Hz sine wave and a (lower-amplitude) 1,000 Hz sine wave. So the 1,000 Hz wave combined with the 100 Hz wave produces a complex periodic wave. The rate at which the complex pattern repeats is called the fundamental frequency (abbreviated F_0).

Fundamental frequency and the GCD

The wave shown in figure 1.5 has a fundamental frequency of 100 Hz and also a 100 Hz component sine wave. It turns out that the fundamental frequency of a complex wave is the greatest common denominator (GCD) of the frequencies of the component sine waves. For example, the fundamental frequency (F_0) of a complex wave with 400 Hz and 500 Hz components is 100 Hz. You can see this for yourself if you draw the complex periodic wave that results from adding a 400 Hz sine wave and a 500 Hz sine wave. We will use the sine wave in figure 1.3 as the starting point for this graph. The procedure is as follows:

- 1 Take some graph paper.
- 2 Calculate the period of a 400 Hz sine wave. Because frequency is equal to one divided by the period (in math that's $f = 1/T$), we know that the period is equal to one divided by the frequency ($T = 1/f$). So the period of a 400 Hz sine wave is 0.0025 seconds. In milliseconds (1/1,000ths of a second) that's 2.5 ms (0.0025 times 1,000).
- 3 Calculate the period of a 500 Hz sine wave.
- 4 Now we are going to derive two tables of numbers that constitute instructions for drawing 400 Hz and 500 Hz sine waves. To do this, add some new labels to the time axis on figure 1.3, once for the 400 Hz sine wave and once for the 500 Hz sine wave. The 400 Hz time axis will have 2.5 ms in place of 0.01 sec, because the 400 Hz sine wave completes one cycle in 2.5 ms. In place of 0.005 sec the 400 Hz time axis will have 1.25 ms. The peak of the 400 Hz sine wave occurs at 0.625 ms, and the valley at 1.875 ms. This gives us a table of times and amplitude values for the 400 Hz wave (where we assume that the amplitude of the peak is 1 and the amplitude of the valley is -1 , and the amplitude value given for time 3.125 is the peak in the second cycle):

ms	0	0.625	1.25	1.875	2.5	3.125
amp	0	1	0	-1	0	1

The interval between successive points in the waveform (with 90° between each point) is 0.625 ms. In the 500 Hz sine wave the interval between comparable points is 0.5 ms.

- 5 Now on your graph paper mark out 20 ms with 1 ms intervals. Also mark an amplitude scale from 1 to -1 , allowing about an inch.
- 6 Draw the 400 Hz and 500 Hz sine waves by marking dots on the graph paper for the intersections indicated in the tables. For instance, the first dot in the 400 Hz sine wave will be at time 0 ms and amplitude 0, the second at time 0.625 ms and amplitude 1, and so on. Note that you may want to extend the table above to 20 ms (I stopped at 3.125 to keep the times right for the 400 Hz wave). When you have marked all the dots for the 400 Hz wave, connect the dots with a freehand sine wave. Then draw the 500 Hz sine wave in the same way, using the same time and amplitude axes. You should have a figure with overlapping sine waves something like figure 1.6.
- 7 Now add the two waves together. At each 0.5 ms point, take the sum of the amplitudes in the two sine waves to get the amplitude value of the new complex periodic wave, and then draw the smooth waveform by eye.

Take a look at the complex periodic wave that results from adding a 400 Hz sine wave and a 500 Hz sine wave. Does it have a fundamental frequency of 100 Hz? If it does, you should see two complete cycles in your 20 ms long complex wave; the waveform pattern from 10 ms to 20 ms should be an exact copy of the pattern that you see in the 0 ms to 10 ms interval.

Figure 1.6 shows another complex wave (and four of the sine waves that were added together to produce it). This wave shape approximates a sawtooth pattern. Unlike the previous example, it is not possible to identify the component sine waves by looking at the complex wave pattern. Notice how all four of the component sine waves have positive peaks early in the complex wave's cycle and negative peaks toward the end of the cycle. These peaks add together to produce a sharp peak early in the cycle and a sharp valley at the end of the cycle, and tend to cancel each other over the rest of the cycle. We can't see individual peaks corresponding to the cycles of the component waves. Nonetheless, the complex wave *was* produced by adding together simple components.

Now let's look at how to represent the frequency components that make up a complex periodic wave. What we're looking for is a way to show the component sine waves of the complex wave when they are not easily visible in the waveform itself. One way to do this is to list the frequencies and amplitudes of the component sine waves like this:

frequency (Hz)	100	200	300	400	500
amplitude	1	0.5	0.33	0.25	0.2

Figure 1.7 shows a graph of these values with frequency on the horizontal axis and amplitude on the vertical axis. The graphical display of component frequencies is the method of choice for showing the simple periodic components of a complex periodic wave, because complex waves are often composed of so many frequency

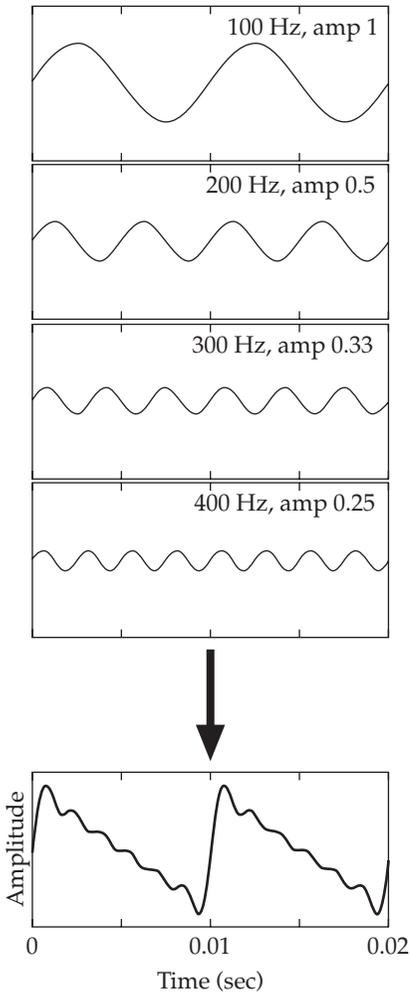


Figure 1.6 A complex periodic wave that approximates the “sawtooth” wave shape, and the four lowest sine waves of the set that were combined to produce the complex wave.

components that a table is impractical. An amplitude versus frequency plot of the simple sine wave components of a complex wave is called a **power spectrum**.

Here’s why it is so important that complex periodic waves can be constructed by adding together sine waves. It is possible to produce an infinite variety of complex wave shapes by combining sine waves that have different frequencies, amplitudes, and phases. A related property of sound waves is that any complex acoustic wave can be analyzed in terms of the sine wave components that could have been used to produce that wave. That is, any complex waveform can be decomposed into a set of sine waves having particular frequencies, amplitudes, and phase relations. This property of sound waves is called Fourier’s theorem, after the seventeenth-century mathematician who discovered it.

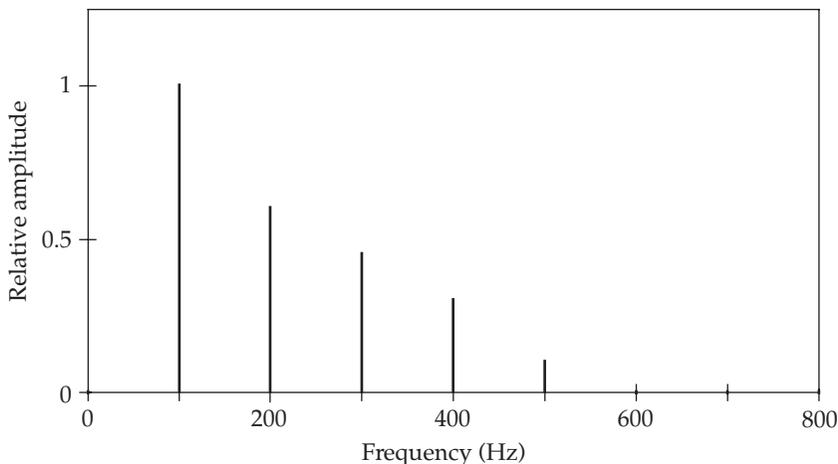


Figure 1.7 The frequencies and amplitudes of the simple periodic components of the complex wave shown in figure 1.6 presented in graphic format.

In Fourier analysis we take a complex periodic wave having an arbitrary number of components and derive the frequencies, amplitudes, and phases of those components. The result of Fourier analysis is a power spectrum similar to the one shown in figure 1.7. (We ignore the phases of the component waves, because these have only a minor impact on the perception of sound.)

1.3.3 Aperiodic waves

Aperiodic sounds, unlike simple or complex periodic sounds, do not have a regularly repeating pattern; they have either a random waveform or a pattern that doesn't repeat. Sound characterized by random pressure fluctuation is called "white noise." It sounds something like radio static or wind blowing through trees. Even though white noise is not periodic, it is possible to perform a Fourier analysis on it; however, unlike Fourier analyses of periodic signals composed of only a few sine waves, the spectrum of white noise is not characterized by sharp peaks, but, rather, has equal amplitude for all possible frequency components (the spectrum is flat). Like sine waves, white noise is an abstraction, although many naturally occurring sounds are similar to white noise. For instance, the sound of the wind or fricative speech sounds like [s] or [f].

Figures 1.8 and 1.9 show the acoustic waveform and the power spectrum, respectively, of a sample of white noise. Note that the waveform shown in figure 1.8 is irregular, with no discernible repeating pattern. Note too that the spectrum shown in figure 1.9 is flat across the top. As we noted earlier, a Fourier analysis of a short chunk (called an "analysis window") of a waveform leads to inaccuracies in the resultant spectrum. That's why this spectrum has some peaks and valleys even though, according to theory, white noise should have a flat spectrum.

The other main type of aperiodic sounds are **transients**. These are various types of clanks and bursts which produce a sudden pressure fluctuation that is not

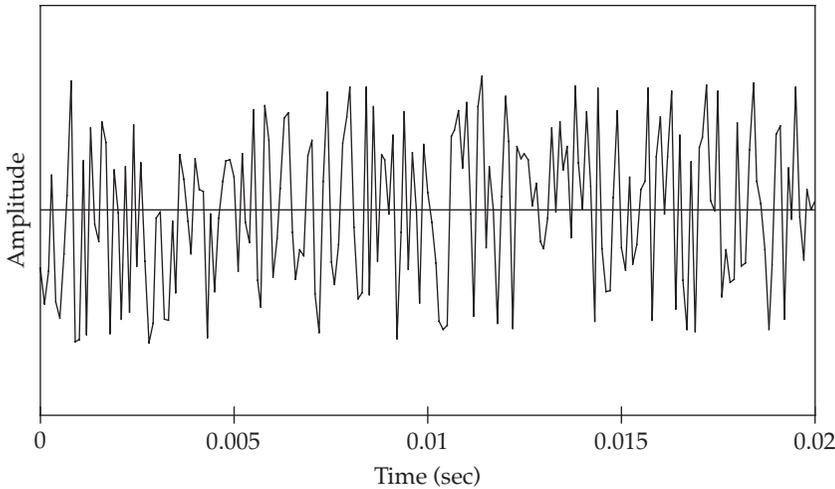


Figure 1.8 A 20 ms section of an acoustic waveform of white noise. The amplitude at any given point in time is random.

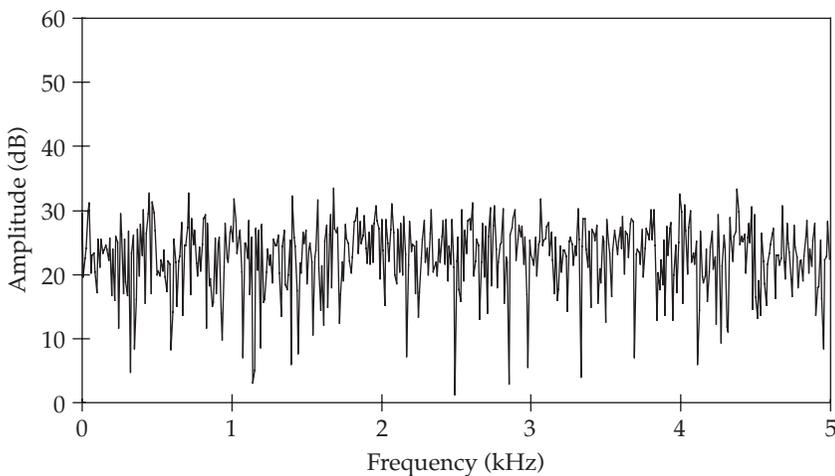


Figure 1.9 The power spectrum of the white noise shown in figure 1.8.

sustained or repeated over time. Door slams, balloon pops, and electrical clicks are all transient sounds. Like aperiodic noise, transient sounds can be analyzed into their spectral components using Fourier analysis. Figure 1.10 shows an idealized transient signal. At only one point in time is there any energy in the signal; at all other times pressure is equal to zero. This type of idealized sound is called an “impulse.” Naturally occurring transients approximate the shape of an impulse, but usually with a bit more complicated fluctuation. Figure 1.11 shows the power spectrum of the impulse shown in figure 1.10. As with white noise, the spectrum is flat. This is more obvious in figure 1.11 than in figure 1.9 because the “impulseness” of the impulse waveform depends on only one point in time, while the “white noisiness”

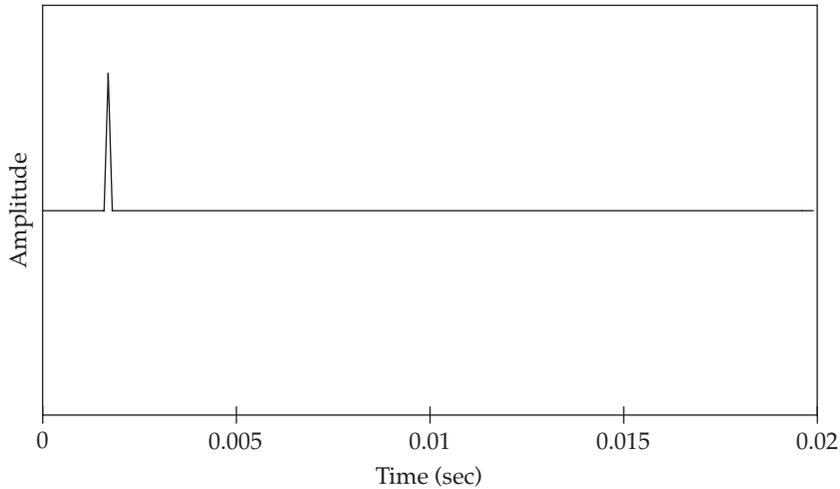


Figure 1.10 Acoustic waveform of a transient sound (an impulse).

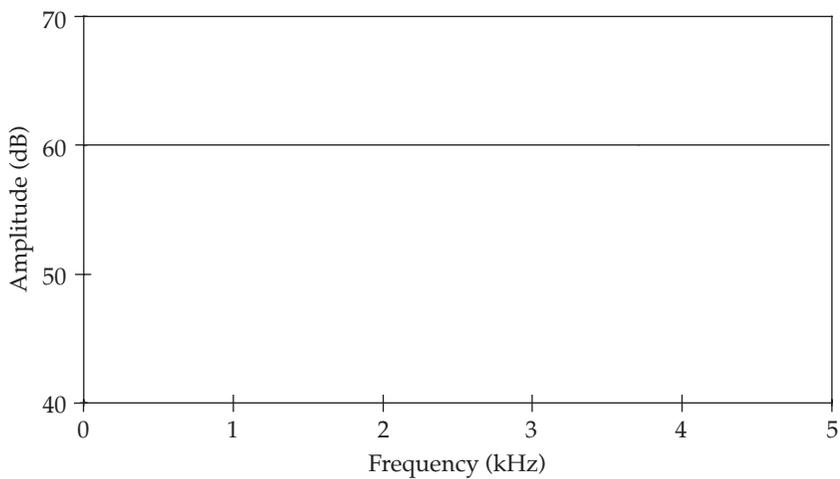


Figure 1.11 Power spectrum of the transient signal shown in figure 1.10.

of the white noise waveform depends on every point in time. Thus, because the Fourier analysis is only approximately valid for a short sample of a waveform, the white noise spectrum is not as completely specified as is the impulse spectrum.

1.4 Acoustic filters

We are all familiar with how filters work. For example, you use a paper filter to keep the coffee grounds out of your coffee, or a tea ball to keep the tea leaves out of your tea. These everyday examples illustrate some important properties of acoustic filters. For instance, the practical difference between a coffee filter and a

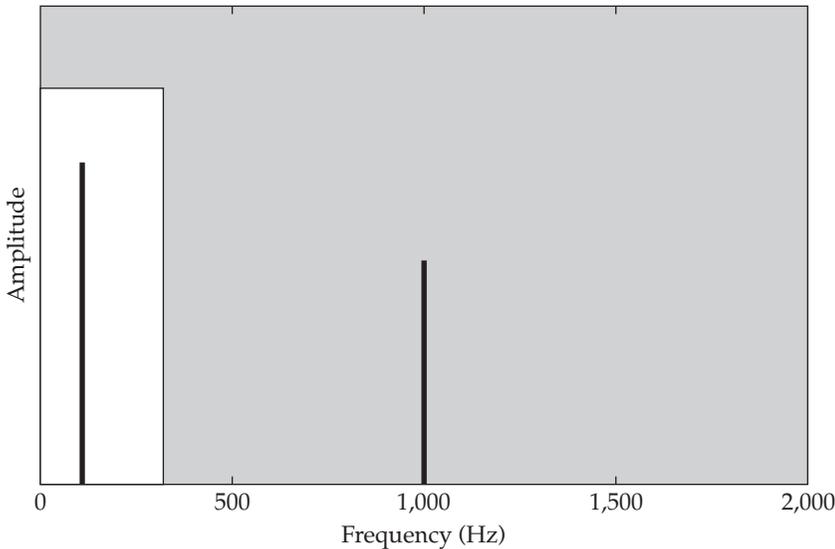


Figure 1.12 Illustration of the spectrum of a low-pass filter.

tea ball is that the tea ball will allow larger bits into the drink, while the coffee filter captures smaller particles than does the tea ball. So the difference between these filters can be described in terms of the size of particles they let pass.

Rather than passing or blocking particles of different sizes like a coffee filter, an acoustic filter passes or blocks components of sound of different frequencies. For example, a **low-pass** acoustic filter blocks the high-frequency components of a wave, and passes the low-frequency components. Earlier I illustrated the difference between simple and complex periodic waves by adding a 1,000 Hz sine wave to a 100 Hz sine wave to produce a complex wave. With a low-pass filter that, for instance, filtered out all frequency components above 300 Hz, we could remove the 1,000 Hz wave from the complex wave. Just as a coffee filter allows small particles to pass through and blocks large particles, so a low-pass acoustic filter allows low-frequency components through, but blocks high-frequency components.

You can visualize the action of a low-pass filter in a spectral display of the filter's response function. For instance, figure 1.12 shows a low-pass filter that has a cutoff frequency of 300 Hz. The part of the spectrum shaded white is called the **pass band**, because sound energy in this frequency range is passed by the filter, while the part of the spectrum shaded gray is called the **reject band**, because sound energy in this region is blocked by the filter. Thus, in a complex wave with components at 100 and 1,000 Hz, the 100 Hz component is passed, and the 1,000 Hz component is blocked. Similarly, a **high-pass** acoustic filter blocks the low-frequency components of a wave, and passes the high-frequency components. A spectral display of the response function of a high-pass filter shows that low-frequency components are blocked by the filter, and high-frequency components are passed.

Filter slopes

The low-pass filter illustrated in figure 1.12 has a very sharp boundary at 300 Hz between the frequencies that are blocked by the filter and those that are passed. The filter has the same effect on every component below (or above) the cutoff frequency; the slope of the vertical line separating the pass band from the reject band is infinitely steep. In real life, acoustic filters do not have such sharp boundaries. For instance, it is more typical for the transition between pass band and reject band to extend over some range of frequencies (as in the band-pass filter illustrated in figure 1.13), rather than to occur instantaneously (as in the low-pass filter illustration). A very steep slope is like having very uniform-sized holes in a tea ball. A shallow filter slope is like having lots of variation in the size of the holes in a tea ball. Some particles will be blocked by the smaller holes, though they would have got through if they had found a bigger hole.

Band-pass filters are important, because we can model some aspects of articulation and hearing in terms of the actions of band-pass filters. Unlike low-pass or high-pass filters, which have a single cutoff frequency, band-pass filters have two cutoff frequencies, one for the low end of the pass band and one for the high end of the pass band (as figure 1.13 shows). A band-pass filter is like a combination of

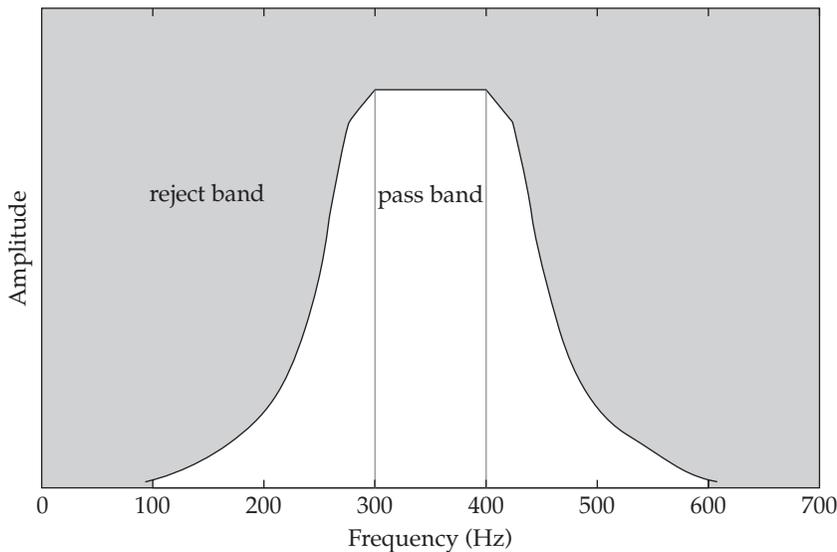


Figure 1.13 Illustration of a band-pass filter. Note that the filter has skirts on either side of the pass band.

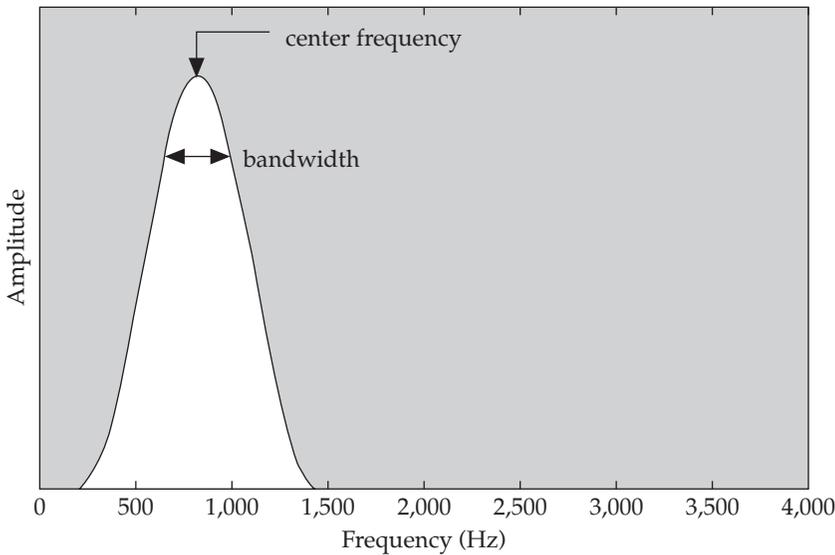


Figure 1.14 A band-pass filter described by the center frequency and bandwidth of the filter.

a low-pass filter and a high-pass filter, where the cutoff frequency of the low-pass filter is higher than the cutoff frequency of the high-pass filter.

When the high and low cutoff frequencies of a band-pass filter equal each other, the resulting filter can be characterized by its **center frequency** and the filter's **bandwidth** (which is determined by the slopes of the filter). Bandwidth is the width (in Hz) of the filter's peak such that one-half of the acoustic energy in the filter is within the stated width. That is, considering the total area under the curve of the filter shape, the bandwidth is the range, around the center frequency, that encloses half the total area. In practice, this half-power bandwidth is found by measuring the amplitude at the center frequency of the filter and then finding the width of the filter at an amplitude that is three decibels (dB) below the peak amplitude (a decibel is defined in chapter 3). This special type of band-pass filter, which is illustrated in figure 1.14, is important in acoustic phonetics, because it has been used to model the filtering action of the vocal tract and auditory system.

Exercises

Sufficient jargon

Define the following terms: sound, acoustic medium, acoustic waveform, sound wave, compression, rarefaction, periodic sounds, simple periodic wave, sine wave, frequency, Hertz, cycle, period, amplitude, phase, complex periodic wave, fundamental frequency, component waves, power spectrum, Fourier's theorem, Fourier analysis, aperiodic sounds, white noise, transient, impulse, low-pass filter, pass band, reject band, high-pass filter, filter slope, band-pass filter, center frequency, bandwidth.

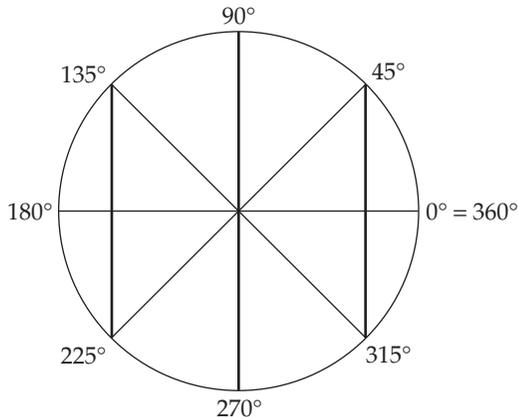


Figure 1.15 Degrees of rotation around a circle.

Short-answer questions

- 1 What's wrong with this statement: You experience sound when air molecules move from a sound source to your eardrum.
- 2 Express these times in seconds: 1,000 ms, 200 ms, 10 ms, 1,210 ms.
- 3 What is the frequency in Hz if the period of a sine wave is 0.01 sec, 10 ms, 0.33 sec, 33 ms?
- 4 Draw a sine wave. First, draw a time axis in equal steps of size 45, so that the first label is zero, the next one (to the right) is 45, the next is 90, and so on up to 720. These labels refer to the degrees of rotation around a circle, as shown in figure 1.15 (720° is twice around the circle). Now plot amplitude values in the sine wave as the height of the vertical bars in the figure relative to the line running through the center of the circle from 0° to 180°. So the amplitude value at 0° is 0; the amplitude at 45° is the vertical distance from the center line to the edge of the circle at 45°, and so on. If the line descends from the center line (as is the case for 225°), mark the amplitude as a negative value. Now connect the dots freehand, trying to make your sine wave look like the ones in this chapter (don't use your ruler for this).
- 5 Draw the waveform of the complex wave produced by adding sine waves of 300 Hz and 500 Hz (both with peak amplitude of 1).
- 6 Draw the spectrum of a complex periodic wave composed of 100 Hz and 700 Hz components (both with peak amplitude of 1).
- 7 Draw the spectrum of an acoustic filter that results from adding two band-pass filters. Filter 1 has a center frequency of 500 Hz and a bandwidth of 50 Hz, while filter 2 has a center frequency of 1,500 Hz and a bandwidth of 150 Hz.