

2. Physical sound

2.1 What is sound?

Sound is the human ear's perceived effect of pressure changes in the ambient air. Sound can be modeled as a function of time.

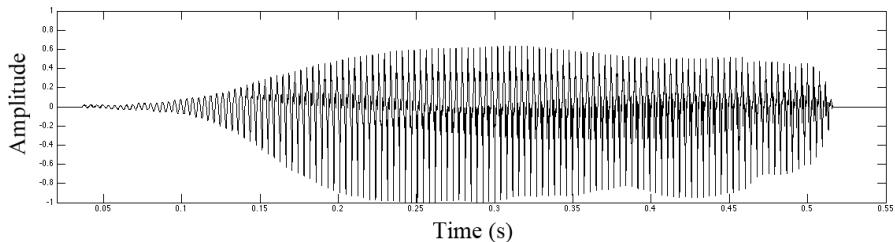


Figure 2.1: A 0.56-second audio clip of an accordion playing C4 (middle C, 261.6 Hz).

When we hear music, we can evaluate its features almost immediately. We can recognize the instrumentation, modality, artist, genre, and perhaps the time and place it was recorded. Graphically, it is difficult to connect this image to what we actually hear: The above graph looks complicated, while the experience of this sound (the *audio signal*) is a single, sustained pitch on an accordion. But when we take a Fourier transform of this clip, we can actually view the frequencies present in a song.

Because pitch and timbre are made up exclusively of the change over time of frequencies and amplitudes, and they tell us so much information about musical features, the Fourier transform is an incredibly useful tool that translates *time domain* signals like music onto an axis of frequencies, i.e., a *frequency domain*. The graph in Figure 2.2 lets our eyes verify what our ears already know: The graph describes the

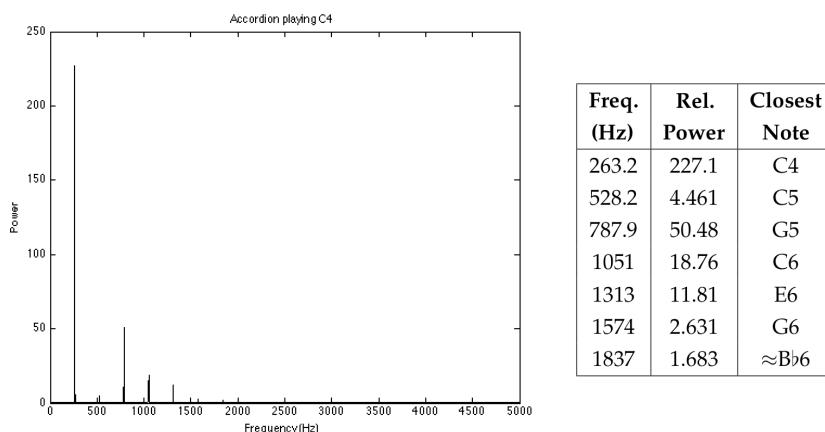


Figure 2.2: The spectrum and listed frequencies attained by the discrete Fourier transform of the clip shown in Figure 2.1. Note the locations of the peaks with respect to frequency.

relative strength of the frequencies present in a signal. In this example, we see the frequency characteristics of an accordion playing C.

The frequencies themselves are not as important as the general shape of the spikes and the distance between them; most cannot distinguish between A and C in isolation, but we do have a relatively easy time identifying the difference between a piano and a violin. This is because of the texture of the instrument's sound, called the *timbre* or *tone color*. When the frequencies are more or less equally spaced from one another, we say that the timbre is *harmonic*, or that we have a *harmonic overtone series*. Explicitly, the Fourier transform of the signal in Figure 2.1 and its graphical representation in Figure 2.2 tell us the signal contains the frequencies 263.2 Hz (C), 528.2 Hz (C), 787.9 Hz (G), 1051 Hz (C), and 1313 Hz (E). The frequencies' respective peak in the graph indicates their loudness; hence, they are decreasing in power.

Middle C is 261.6 Hz, so apparently this accordion is slightly out of tune—but furthermore, its timbre is not perfectly harmonic: the difference between its overtones should be 263.2, but $528.2 - 263.2 = 265$, and $787.9 - 528.2 = 259.7$. There are several possible reasons for why these spikes are not exactly equally spaced. Most likely, it is due to the imperfect physical proportion and construction of the instrument's metal reeds, but it could also be error encountered in the recording process or experimental error.

To interpret how exactly this translates to what our ears hear, we must take into account how certain frequencies are perceived by the brain. Young, healthy human ears can detect frequencies within a range of 20-20,000 Hz, where 20 Hz and 20,000 Hz are threshold and limit values, but our ears are not uniformly sensitive to these frequencies [1]. Within the range of 1000 to 5000 Hz, our ears are especially sensitive, meaning that sounds with frequencies within this range do not have to be as loud for our ears to detect them.

Mathematically, the Fourier transform constructs an *orthonormal basis* that takes a complicated sound wave and reduces it to its component waves, which are all simple sine and cosine waves, or *sinusoids*.¹ It shows us every frequency and its amplitude that is present in a complex sound over an interval of time. The connection between the graph of the transform and its mathematical properties is a giant step towards realizing the Fourier transform and its digital applications. Because sight and sound retrieve giant spheres of information, we have to make decisions about what is important and what we can take for granted. Our brains are so excellent at processing information that we can give certain sensations finer resolution (like an important

¹"Orthonormal" means orthogonal and normal. For a function to be orthogonal to another function, the two functions must be *linearly independent*. The condition of normality is satisfied when each function involved has, in some appropriate sense, energy 1. Finally, a *basis* is a set of functions such that an arbitrary function (within reason) may be written in terms of the basis. See Chapter 7.

Fletcher-Munson equal-loudness curve

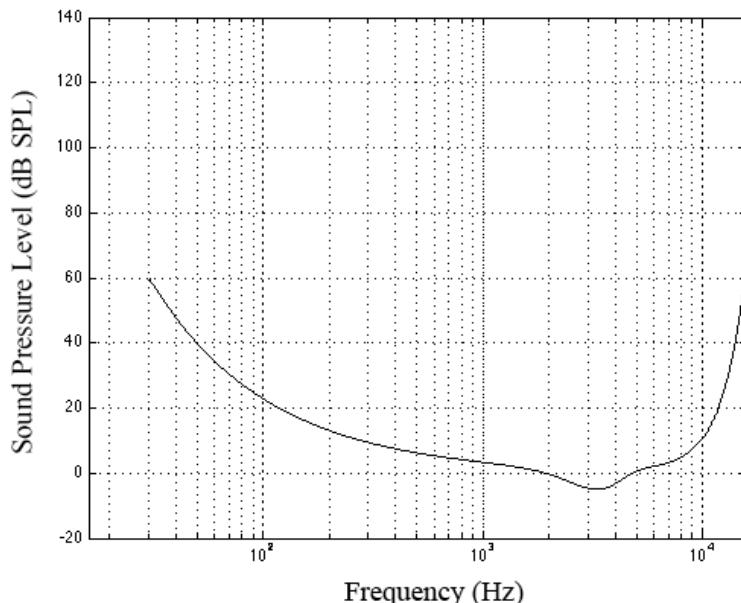


Figure 2.3: We detect frequencies between about 20 and 20,000 Hz as pitched sound. Furthermore, each of these frequencies has a minimum threshold of loudness. This graph, the Fletcher–Munson curve, shows the minimal sound pressure level in decibels (dB) required for the frequency to be heard.

message from our friend), and others none at all (like the hum of the refrigerator).

Consider seeing a relatively involved movie for the second or third time, and noticing things you didn't notice before that now make sense. We seem to prefer movies like these. We may achieve a decent understanding of the plot on the first viewing because we extract salient parts of the dialogue and action and put them in order, but a complex plot can hide clues of outcomes and their rationale all over the film that are more obvious when our brains can support them with familiar elements.

A complex piece of music can be a lot like a complex movie. We perceive both sound and light as signals. When a signal demands

our attention, it is said to have a high amount of *information*. A signal with meaningless content that we don't need or want to listen to is called *noise*. Sound produced by white noise machines, e.g., is random, unpitched, and trivial. It does not contain a message because it is formally disorganized, and it can even help some people sleep because of its uniform randomness. Noise is composed of so many periodic waves that we consider it aperiodic. We cannot extract individual frequencies of noise, as we can in melody or a major chord in music. A signal can be half meaningful and half noise, and our brains are powerful enough to recognize the difference and attempt to separate the two.

Although sine waves are not fun to think about, they substantiate much of the mathematics and physics behind music. The mathematical and physical equations which produced the previous graphs form a basis for the sensation of sound. Many musical concepts are results of mathematical relationships. First, let us examine the basic mathematical structure of sound. Musical form will be addressed in Chapters 3 and 4.

2.2 Simple harmonic motion

Like light, sound is traveling energy, and we can model such energy mathematically with waves. The simplest wave is a *sinusoid*, a trigonometric function such as $\sin(\omega t)$ or $\cos(\omega t)$ where t denotes time and ω specifies how often it repeats itself—its *angular frequency*.² A sinusoidal wave represents the *simple harmonic motion* of an object because its frequency and extreme magnitudes do not change over time.

Both a spring and a tuning fork exhibit simple harmonic motion. Below, we see two states of a vibrating tuning fork called *modes of vibration*. Both of these modes produce sound that is near in tone to a sine wave (or *pure tone*), but as you might have experienced, the tone

²The *frequency* f in Hz. (cycles per second) is related by $\omega = 2\pi f$.



Figure 2.4: A tuning fork and a weighted spring oscillate in simple harmonic motion.

is more metallic and glassy than the electronic sound of a sine wave. When we strike the tuning fork, we experience the *attack* of the sound and then the sound *sustains* (decaying due to frictional forces with time) and eventually *releases*, leaving no sound. In the ideal physical world, i.e., one without the external forces of gravity, friction, and other resistive forces, a spring set into motion could oscillate forever at a uniform amplitude, as could a tuning fork. But that is not what happens in reality. The closest we can get to simple harmonic motion is represented by the curve in Figure 2.5.

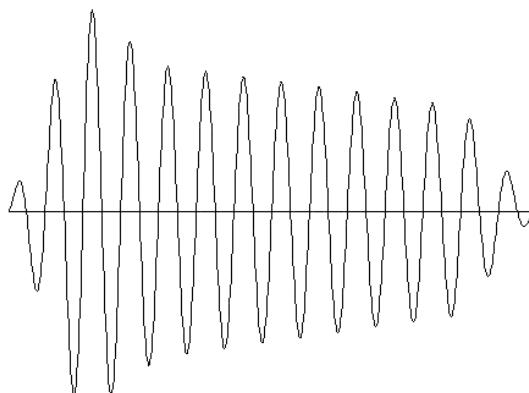


Figure 2.5: A musical wave in reality begins at zero energy, climbs to a maximal energy, and fades to zero energy.

You can see that the amplitude of this wave varies, but the points at which it crosses the horizontal axis, i.e., when its amplitude equals 0, are evenly spaced over time. This means the frequency does not vary, but the intensity of its motion does. True simple harmonic motion can be generated by an oscillator, a computer, or a tuning fork with a driving motor attached to it [2].

Amplitude represents pressure as well as voltage: An audio function models the pressure in the air corresponding to the sound wave as a function of time, and when the signal is electrified, the amplitude represents the (relative) voltage. In acoustics as well as electrical engineering, we call this function a signal, and the amplitude tells us most of the information we need to determine how loud our ears will perceive it to be. Because air is elastic, when a sound wave travels in air, it excites the air molecules and varies the pressure. The amplitude of the graph of a sine wave describes this behavior (see Figure 2.6).

There are three fundamental aspects of a sinusoid of the form $A \sin(\omega t + \phi)$: Its magnitude³ A , frequency ω , and phase ϕ . We have already considered amplitude: It is the pressure. A wave at amplitude 0 means that the system is at normal atmospheric pressure—the pressure of the environment to which our ears have adjusted and no extra pressure is affecting the eardrum at that instant. Frequency can be determined by the number of times per second that the signal has zero pressure, or the rate at which the signal crosses the graph's horizontal axis. Finally, we can determine the phase ϕ at any point t from the time of the next zero crossing, i.e., where the amplitude crosses the horizontal axis.

³Unfortunately, there are quite a few terms that will be used somewhat interchangeably to mean magnitude: Amplitude, height, displacement, energy, power, voltage, pressure, strength, and loudness. Loudness is a perceptual word and since we do not perceive all frequencies as equal (or at all, as in Figure 2.3), this word will be used with caution. Voltage, power, and energy are typically encountered in electrical engineering texts to mean amplitude, though they absolutely do not have equivalent meaning (see Appendix A). Strength and displacement are words used here to denote the magnitude of a wave, i.e., the vertical distance from 0.

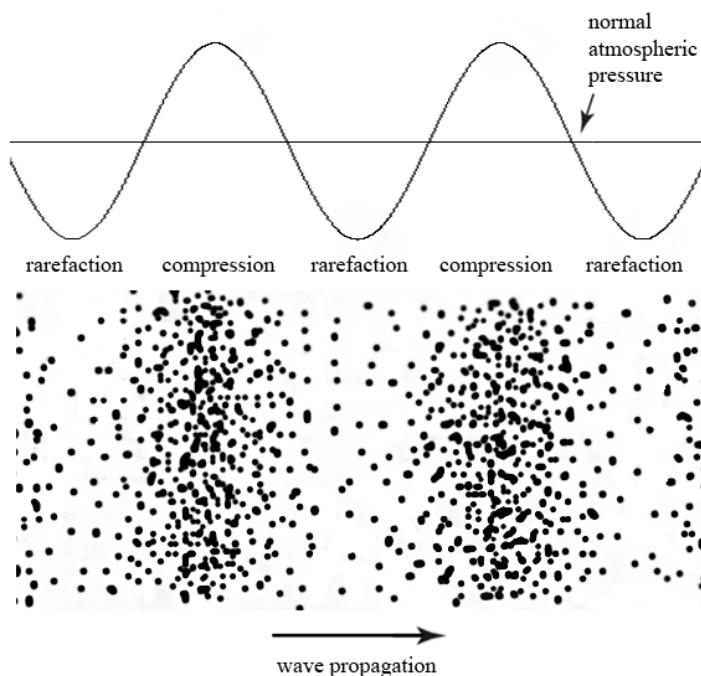


Figure 2.6: The compressions and rarefactions in air resulting from sound waves, shown two ways. The maximal points of the sine wave graph correspond to the most compressed areas of the particle graph, represented by the most densely spaced dots. The minimal points correspond to rarefactions, represented by the least dense spacings of dots. Where the amplitude of the sine wave is 0 represents *normal atmospheric pressure*, where the density of the dots is *average*.

For ease of computation, we only allow amplitude to vary between -1 and 1, so the average value of the amplitude of a simple sine wave is always zero.⁴ It may seem strange that pressure can take on negative values, but it simply means that the sound's pressure is dipping below normal atmospheric pressure. For purposes of standardization, this is defined as the pressure of air at sea level, 101,325 pascals (Pa); but in reality, this is the average atmospheric pressure of our present environment to which our ears have adjusted. Hence, amplitudes higher than 0 imply that the pressure induced by a sound wave is

⁴In electronic reality, sound signals can, however, have nonzero average value due to things like DC offset, uncalibrated equipment, or postproduction changes.

greater than normal pressure (*compression*), and amplitudes below 0 imply that the pressure of a sound wave is less than normal pressure (*rarefaction*). Our ears detect sound by *change* over time in pressure, so a single, isolated amplitude tells us nothing about what we actually hear.

Angular frequency is given by ω in *radians per second* (rad/s), and it is equal to $2\pi f$, where f is *ordinary frequency*, given in *hertz* (Hz). Frequency f is inversely proportional to the time that the sine wave takes to complete one *period* T , as given by the following formula.

$$f = \frac{1}{T}$$

Therefore, $\omega = \frac{2\pi}{T}$.

Phase tells us where the wave is along the course of a single period, taking on angles between 0° and just less than 360° . We are especially interested in phase when we have two waves of identical frequency.

Now let us examine the nature of a simple sinusoid where $\omega = 2\pi$ radians/second (so $f = 1$ Hz), $x(t) = \sin(2\pi t)$.

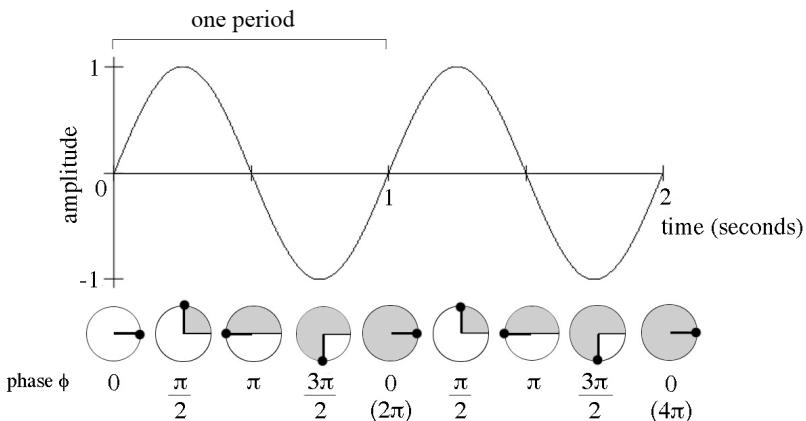


Figure 2.7: A simple sinusoid, $x(t) = \sin(2\pi t)$, with the phase ϕ marked.

Take note of the circular diagrams beneath the graph in Figure 2.7. These circles show different positions along the unit circle. The starting

position where $\phi = 0$ is the rightmost point on this circle, situated at its intersection with the horizontal axis. When we move counterclockwise along the circumference of this circle, we increase the angle relative to this position. When we return to this position, we have moved 360° . In radians, 360° is equal to 2π . We can translate any angle in degrees to radians by multiplying the number of degrees by $\frac{\pi}{180}$, e.g., for a right angle $\phi = 90^\circ$, the equivalent angle in radians is calculated to be $90^\circ \cdot \frac{\pi}{180} = \frac{\pi}{2}$ radians.

Note that $\phi = 0$ is positioned on the circle exactly where $\phi = 2\pi$ and $\phi = 4\pi$. This is true of any even-integer multiple of π ($4\pi, 6\pi, 8\pi, \dots$). Similarly, the trigonometric function of any variable (such as a frequency ω) is the same as that variable phase shifted by an integer-multiple of 2π , i.e.,

$$\begin{aligned}\cos(\omega) &= \cos(\omega + 2k\pi), && \text{and} \\ \sin(\omega) &= \sin(\omega + 2\pi k), && k = 0, 1, 2, \dots\end{aligned}$$

Again, the height of points along the unit circle at angle ϕ is given by $\sin(\phi)$, and the width is modeled by $\cos(\phi)$. The phase at the initial time, $t = 0$, is the angle that the sinusoid is shifted relative to a sine wave with no phase. We represent phase with the Greek letter ϕ , so that a simple sinusoid is formally written

$$x(t) = A \sin(2\pi ft + \phi) = A \sin(\omega t + \phi).$$

We multiply frequency by the quantity 2π because this strengthens the connection between the unit circle and frequency. The angular frequency is commonly used in the Fourier transform and physical science. However, in music, we connect pitch to frequency in hertz (like A440), so we will use $2\pi f$ instead of ω when considering sound musically.

It is impossible to identify the phase of a single sine wave without a reference point. However, it is important for a signal to begin and end at amplitude 0 in order to understand its behavior, because sounds

beginning or ending at a nonzero amplitude surprise our ears to the point that the frequency undergoes distortion. Consider dropping a needle on the record and hearing a fuzzy click. We hear a burst of sound when there is any discontinuity in pressure. This is reflected not only on our basilar membrane, but in the Fourier transform, and it is one type of *clipping*.

2.3 Complex harmonic motion

Now let us examine more complex waves. In reality, virtually every sound is a complex wave. We really only encounter simple waves during hearing tests or in electronic music. In fact, listening to a sine wave for an extended period of time can cause headaches, extreme emotional responses, and hearing damage [3].

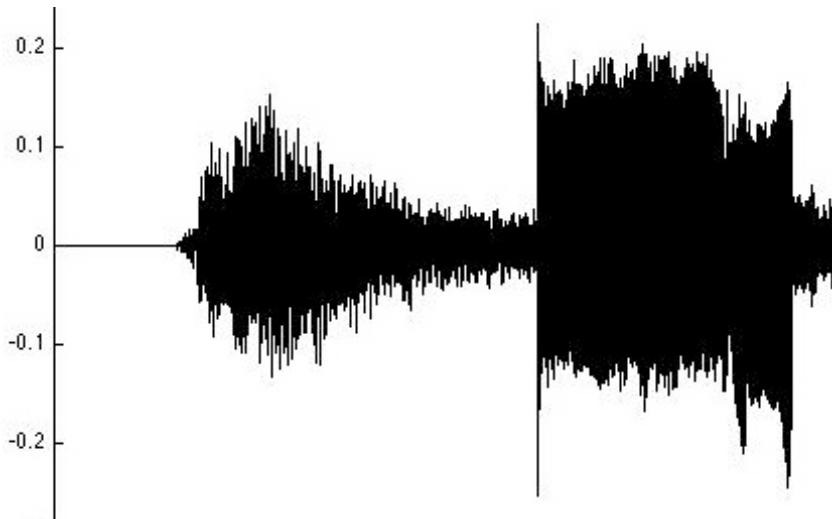


Figure 2.8: A clip of an audio signal.

The horizontal axis of Figure 2.8 is once again time and the vertical axis is amplitude or pressure. Clearly, this is complicated: It is close to impossible for our brains to detect any sort of pattern in this waveform

because there is no clear repetition. Furthermore, there are no distinct frequencies we can pick out because there is no obvious repetition in the sound wave. However, this wave can be decomposed completely into sine waves solely by a Fourier transform. It may require hundreds, even an *infinite* amount of them, but it can be done. Let us look at a simpler—but still complex—wave to illustrate the combination of simple sinusoids, as in Figure 2.9. Our ears can identify a pattern

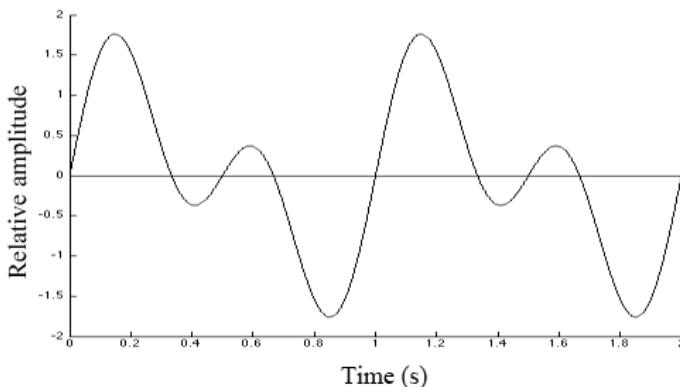


Figure 2.9: The combination of two simple sine waves, $x_1(t) = \sin(2\pi t)$ and $x_2(t) = \sin(4\pi t)$.

because this wave is periodic. This wave is made up of two different sine waves, and they are harmonic relatives of each other: One is twice the frequency of the other. This wave repeats identically after every second. The first sinusoid x_1 has a frequency of 1 Hz (2π rad/s), and the second sinusoid x_2 has a frequency of 2 Hz (4π rad/s). These are called *frequency components* ω_k , where $\omega_1 = 2\pi$ and $\omega_2 = 4\pi$.

We couldn't actually hear these frequencies as pitch in reality because they oscillate too slowly: Our ears only translate frequencies above about 20 Hz as pitched sound to our brains. While using such a low frequency is preferable for ease of visualization and computation, any pair of sine waves whose frequencies have a 2:1 ratio is defined as having the interval of an *octave*. Say that the scale of the horizontal

axis in Figure 2.9 was in milliseconds instead of seconds. Then these sine waves would have the audible frequencies of 1000 and 2000 Hz. Visually, we can see that these waves intersect half of the time that they cross the horizontal axis, and the ratio between these waves' frequency (2:1) emphasizes the inversely proportional mathematical relationship between frequency and time ($f = 1/T$).

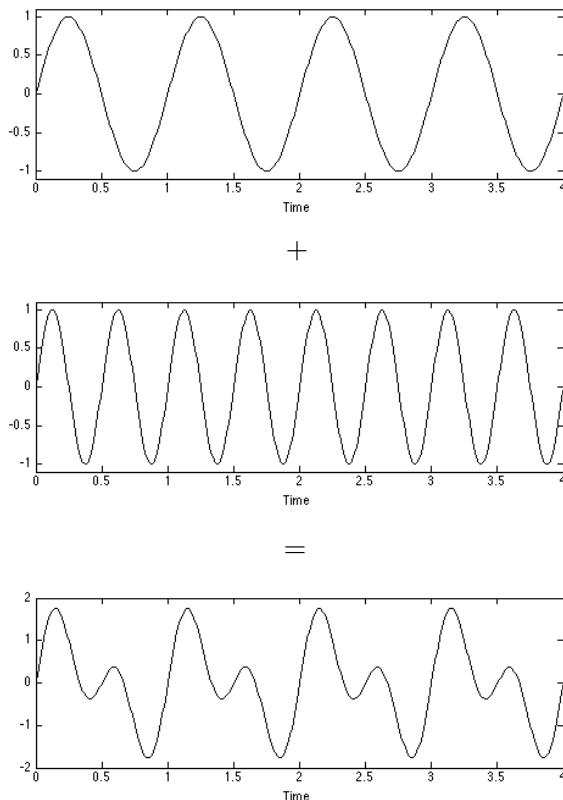


Figure 2.10: The combination of one sine wave of frequency f_1 with another sine wave with frequency $f_2 = 2f_1$ produces the interval of an octave. Note the periodic nature of the resultant wave: It repeats itself identically four times, just like the first wave.

The graph of the signal given in Figure 2.10 can be determined from the graphs of these two sine waves, but picking out even a small handful of simple sinusoids from a complex wave is not a task we want to leave to our senses. We need more advanced computational tools to analyze frequencies contained in a signal that looks like a random string of numbers between -1 and 1 . *That* is what is so very exciting about the power of the Fourier transform.

The physical law at work here is called the *principle of superposition*.

The principle of superposition: Every wave can be represented as a sum of simple sinusoids.

Note that this says nothing about whether this sum has finitely or infinitely many terms. Square and triangle waves, for example, have jagged corners that cannot be represented by a finite amount of sinusoids. The principle of superposition is critical to understanding the concepts in the remainder of this book.

2.4 Harmony, periodicity, and perfect intervals

When two waves have frequencies that are related to each other by a small-number integer ratio like 2:1, we say that they are *harmonic*, that they have a *harmonic relationship*, or that they are *harmonics of one another*. The above example of $f_1 = 1$ Hz and $f_2 = 2$ Hz, i.e., $f_2 = 2f_1$, forms the interval of an octave. Likewise, the octave above any frequency is double that frequency, so we can calculate the frequency f_k that is k -many octaves (the k th octave) above a given frequency f_0 by the equation

$$f_k = 2^k f_0.$$

Letting $k = 0$ returns f_0 , so the frequency 0 octaves above a given frequency is the original or *fundamental* frequency. This is called *unison*, or *perfect unison*. A perfect interval is characterized by a small-number

ratio between the two frequencies, restricted in Western music to 1:1 (P1, perfect unison), 2:1 (P8, perfect octave), 3:2 (P5, perfect fifth), and 4:3 (P4, perfect fourth) [4].

Perfect unison is trivially the smallest integer ratio. Two frequencies separated by an octave are in the ratio 2:1, the second smallest integer ratio. The *perfect fifth* has a ratio very close to 3:2 in equal temperament, and it is exactly 3:2 in just intonation and Pythagorean tuning. A *perfect fourth* has a 4:3 ratio: It is the *inversion* of the perfect fifth, sounded by moving up an octave and down a perfect fifth. In fact, as in the discussion later of musical temperament and tuning systems, *every* note of the 12-tone Pythagorean scale can be attained by moving in perfect fifths, but these intervals are related to f_0 by increasingly larger integer ratios.

Moreover, the smaller the integer ratio between two frequencies (and two periods, thereby), the more pleasant or consonant we find their interval. In the introduction to this chapter, I mentioned the harmonic overtone series and its relationship to timbre in music: Musical instruments are constructed to have tones containing integer-related frequencies. To back up a little bit: When we hear A at 440 Hz on a piano, we do not just hear the frequency 440 Hz. If this were the case, it would sound no different from an electronic beep caused by an oscillator or ideal tuning fork. When we hear A440 from a piano, we actually hear a whole *spectrum* of other frequencies resulting from the resonance of the piano and the nature of the *fixed string*. Musical, pitched instruments like the piano, with the exception of percussion instruments, generate overtone series that are very-nearly harmonic, regardless of the pitch played. The modes of vibration on circular membranes have *Bessel function ratios*.

We define *nodes* as zero *crossings* (i.e., where $x(t)$ is zero) of the horizontal axis by a wave and *antinodes* as areas of maximal compression and rarefaction. A node exists on an instrument at a point or region that stays stationary while the rest of the instrument vibrates. Nodes and antinodes are used when we consider *standing waves* which

occur within all musical instruments and rooms. A standing wave is produced when a sound wave's forward velocity is the same as its backwards velocity, hence it stands still with respect to position. For example, in a violin, waves move back and forth at the same velocity along a string fixed at both ends and therefore are only displaced up and down. Standing waves occur when the wavelengths of a given frequency are in integer proportion to the dimensions of a string, room, or column of air. They cause feedback in a recording studio because they do not die as quickly as waves of other frequencies.

Helmholtz resonance can be witnessed when air is blown across a small opening of an otherwise closed cavity, like a bottle of water or the cracked window of a moving car. The frequency produced is inversely proportional to the volume of this cavity and proportional to the cross sectional area of the opening, so frequencies of larger cavities like the interior car are low. The formula to calculate the Helmholtz resonance ω_H is given by

$$\omega_H = \sqrt{\gamma \frac{A^2 P_0}{m V_0}}$$

where γ is the adiabatic index of specific heats (1.4 for dry air), A is the area of the opening, P_0 is the initial pressure of the air inside of the cavity, m is the mass of air in the neck of the opening, and V_0 is the initial volume of air inside of the cavity. So, widening the crack of the car window will increase the angular frequency of the Helmholtz resonance, and reducing the amount of water in the bottle (hence increasing V_0) will reduce the characteristic Helmholtz resonant frequency ω_H .

Lightly placing one's finger at any of the nodes on a fixed string does noticeable things to its harmonics. Doing so at the halfway point on a guitar string, for example, causes the odd harmonics to drop out and the octaves above the fundamental to be very clear. Figure 2.11 depicts the first four modes of vibration of a fixed string and the dots highlight their nodes.

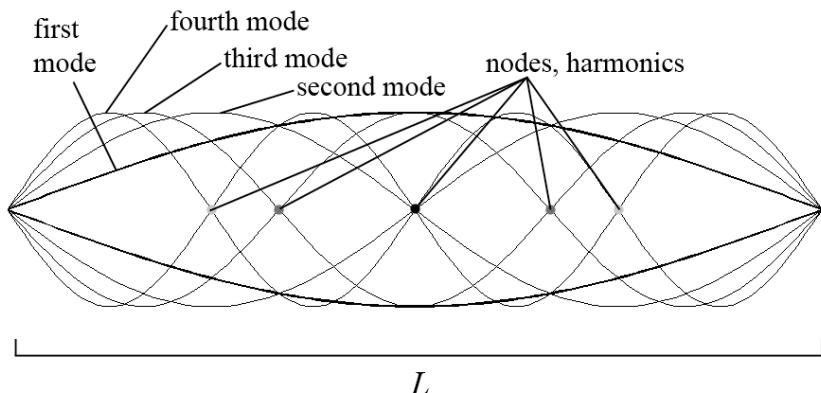
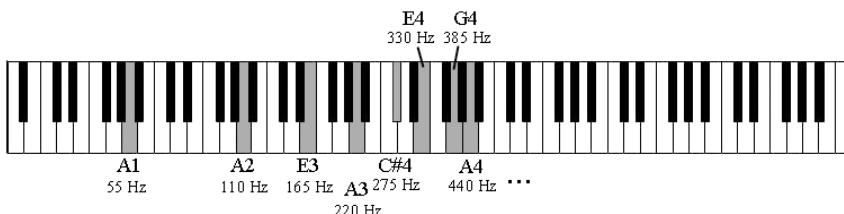


Figure 2.11: The first four modes of a fixed string. Because a string is secured at both ends, its overtone series is defined according to its length, and the wavelengths of the frequencies it contains are restricted to integer divisions of that length, i.e., $f:2f:3f$, and so on. A non-integer ratio would result in an impossible scenario: A string loose at one end.

By paying extra attention to the tone of musical instruments, the individual overtones may be realized. If you have access to a piano, try this experiment. Find a way to depress every key on the piano except for the second-to-bottom A, using books or a friend's arms. Do this slowly so that the keys do not trigger sound. When the piano is silent, strike the A with considerable force and listen closely. You should be able to hear at least two octaves and a perfect fifth (an E) higher than this note. These are the first three overtones of the fundamental frequency, 55 Hz, shown in the following table.

Frequency	Note name	Ratio to 55 Hz	Interval
55 Hz	A	1:1	Perfect Unison
110 Hz	A	2:1	Perfect Octave
165 Hz	E	3:1	Perfect Fifth
220	A	4:1	Perfect Octave
275	C♯	5:1	Major Third
330	E	6:1	Perfect Fifth
385	G	7:1	Minor Seventh
440	A	8:1	Perfect Octave



The intervals above are in their simplest forms. As you can see, the interval between E3 at 165 Hz and A1 at 55 Hz spans a perfect octave *and* a perfect fifth. Frequencies separates by octaves sound so similar that we actually call all octaves by the same note name, so in most cases this relaxed terminology of reducing intervals that span more than an octave is acceptable.⁵

This series hypothetically continues forever to include the ratios 9:1, 10:1, and so on. But the first few partials of any instrument have more energy than higher ones, so those are the ones we predominantly perceive—even though removing the higher ones would affect the perceived timbre.

⁵Particularly in the genre of jazz, the intervals of the ninth, eleventh, and thirteenth are used with some frequency. However, their sonority is similar to the interval minus an octave, i.e., a ninth has similar quality to a second, a eleventh to a fourth, and a thirteenth to a sixth.

Harmonicity—and furthermore, the Western conceptualization of consonance—in music is manifested by simple mathematical relationships. We will say more about consonance and dissonance in the fourth chapter on auditory perception.

2.5 Properties of waves

When waves interact with other waves or with media like walls, water, and hot air, they exhibit to some degree the properties of reflection, refraction, interference, and damping. Some of these we observe on a daily basis, like echoes, but some are quite rare, like cancelation. Understanding the properties of waves helps avoid unwanted sounds (noise) and improve the desired message (signals), and all of the properties are direct consequences of the behavior of amplitude, phase, and frequency in response to the physical world.

Before we begin to explain these properties, three more features of waves useful to understand are *wavelength*, *amplitude envelopes*, and *crests* versus *troughs*. Wavelength λ is the distance in meters that a wave of frequency f travels away from its source in one period T . We calculate wavelength λ with the equation

$$\lambda = vT = \frac{v}{f},$$

where v is the velocity of sound. In dry, room-temperature (68° Fahrenheit) air, the speed of sound is about 343 meters per second (m/s), and a 1000 Hz wave would therefore have a wavelength of

$$\frac{343 \text{ m/s}}{1000 \text{ s}^{-1}} = 0.343 \text{ m.}$$

Note that frequency ($f = 1/T$) can be notated either as hertz or s^{-1} .

An amplitude envelope describes the general shape of the amplitude over time for a given wave. Attack, decay, sustain, and release are the four general qualities of an amplitude envelope, and they are most

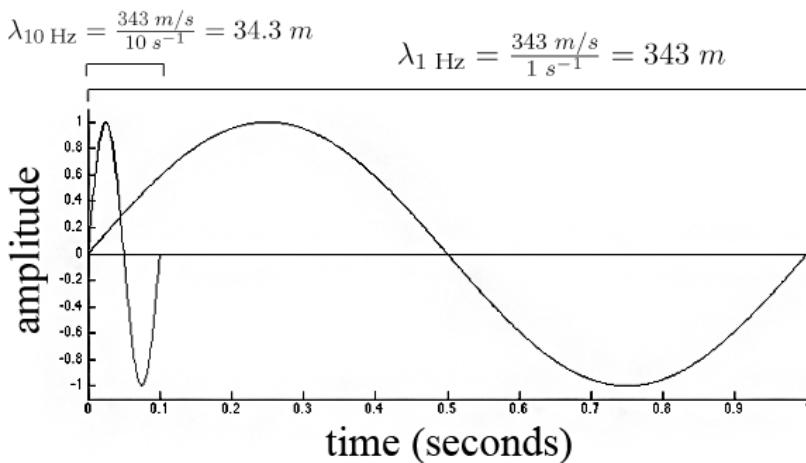


Figure 2.12: The wavelengths of two pure tones, 1 Hz and 10 Hz, calculated by the formula $\lambda = \frac{v}{f}$. Since hertz are measured in inverted seconds (s^{-1}), this is the same as multiplying the speed of sound by the duration of one period (0.1 and 1 seconds, respectively). Notice that $\lambda_{10\text{Hz}}$ is one-tenth of the length of $\lambda_{1\text{Hz}}$.

often ordered respectively for acoustic instrument examples. They are all notated one of three ways: As an instant to refer to the instant at which they begin (the onset time), as an interval to mean the interval over which they occur, or as a rate defining the speed at which they happen. It is easy to understand them graphically, as in Figure 2.13.

However, an amplitude envelope is rarely this simple, and the one defining the shape of the frequency domain is not described the same way. Attack time and attack rate are particularly meaningful to the mathematics of music, especially when attempting to extract features from music. Attack almost solely defines where onsets exist. Onsets help us identify the location of beats and important events like the beginning of choruses or verses in musical signals.

Finally, *crests* and *troughs* occur at the antinodes of a wave or fixed string. These are simply synonyms for maximum and minimum values in pressure.

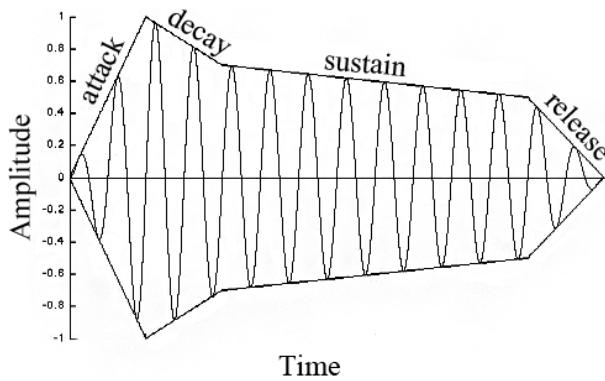


Figure 2.13: A general attack-decay-sustain-release envelope, or *ADSR* envelope: The first onset of a note is the attack; the movement from the peak of the attack to the sustain is the decay; the duration a note is held is shown in the sustain; and the final decrease is the release, where the note is no longer being played. This envelope also describes reverberation.

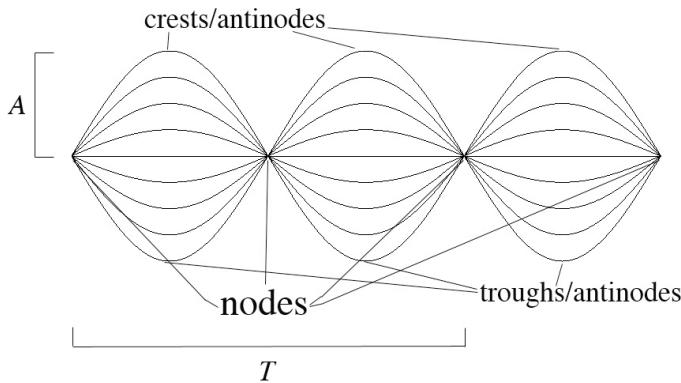


Figure 2.14: The nodes, antinodes, crests, and troughs of a sine wave, shown with eight different amplitudes. The antinodes are located at the crests and troughs, i.e., areas of extreme compression and rarefaction, and the nodes are located at normal atmospheric pressure. Along a string, the nodes are located where the string does not move. These positions are according to its length.

The nodes are located where the amplitude is 0. Both ends of the string are therefore nodes. Antinodes occur in exactly the opposite places: Where the magnitude (absolute value) of the amplitude is locally maximal—i.e., the magnitude is greater than both of the leftmost and rightmost magnitudes. These extreme regions are also called *com-*

pressions (where maximal) and *rarefactions* (where minimal). Skipping a jump rope creates one antinode (we only count an antinode once per extrema), two nodes, one compression, and one rarefaction.

Reflection

The property of reflection can be readily observed when loud sounds initiate in rooms with hard surfaces. We experience reflection when sound in a room *reverberates* or *echoes*. Bats use the reflection of sound to aid their night vision using *echolocation*, calculating their distance to objects from their own position to a high level of precision by emitting a chirp and measuring the time that it takes for its reflection to be heard. [5]. The variables in echolocation are: The speed of sound v equal to 343m/s (also written c_v , though the notation c in physics is typically reserved for the speed of light), the round-trip time that the sound takes to hit the object and reflect back, and the speed at which the observer (the bat) is traveling. Since sound travels relatively quickly to the speed of the bat, this third variable is reasonably negligible and it is unlikely that the bat is taking note of this at all. So, if the sound takes 5 seconds to reflect back to the bats ears, the object is $(343\text{m/s}) \cdot (5\text{s})/2 = 857.5$ meters away. This is divided by 2 because it took 5 seconds for the round-trip time, so it took 2.5 seconds for the sound to travel to the object.

The human brain perceives sonic events that begin less than about one-tenth of a second (0.1 s) apart to be part of the same sound [3]. So, reflections of sound over short distances are perceived as a single signal because they happen within 0.1 seconds of one another. Therefore, the minimum distance that a sound can travel in order for an echo to be perceived is therefore about $(343 \text{ m/s}) \cdot (0.1 \text{ s})/2 = 17.15$ meters, again dividing by 2 because it has to make a round trip. Sounds beginning greater than 0.2 seconds apart are separated by the brain, and between 0.1 and 0.2 seconds is an *interval of confusion* or *roughness*.

The myth that a duck's quack does not echo was only recently debunked by the Acoustics Research Centre at the University of Salford in 2003 [6]. Their best guess as to why this was ever a myth is that quacks may be difficult to detect because they do not have a sharp attack like lightning or handclapping, and furthermore, ducks are usually in water or the air, not in tunnels where echoes are often observed.

Reflection can be a useful property of sound when recording in noisy environments. During sporting events like football and basketball games, you may observe a few people wearing headphones on the sidelines holding a large, clear, circular object with a microphone at the center. This is a circular parabola, designed much like a satellite dish. The microphone is placed at the parabola's focus, a point through which all waves that hit the parabola reflect and travel. At the Exploratorium Museum in San Francisco, there are two large parabolas about eight feet in diameter. They are installed vertically so that museum visitors can sit inside them on seats strategically placed so one's ears are very close to the focal point. The parabolas face each other, but are about 50 feet apart, making it seem irrational that soft sounds could be effectively transmitted over such a distance in the popular, noisy museum. Surprisingly, speech barely louder than a whisper can be clearly heard at the other end. The same idea applies to satellite dishes, but their foci extend far beyond the rim of the dish to compensate for the great distance to their signals' sources in outer space.

Refraction

When sound travels from one region to a region with a different density or stiffness, refraction and dispersion occur[4]. In waveguide synthesis, these regions are called *scattering junctions*. The denser a region, the less room the closely spaced particles have to move around. Sound waves can become more excited in stiffer mediums due to improved elasticity

[7]. Therefore, sound travels more quickly in stiff, light solids than in liquids or gases. Measurements taken with a contact mic on the metal interior of a brass horn, for example, are much richer (more partials are articulated) than measurements taken from the air outside of the horn. The type of wood used in the body of a violin influences how much the violin amplifies its sound, and the propagation of sound in spruce (a common wood in violins) is twice as fast along the grain (3000 m/s) as it is across it (1500 m/s) [8]. Refraction is an especially important property to consider for submariners, architects, and materials scientists.

The speed of sound depends on the *bulk modulus* B of a medium (a number representing its elasticity or stiffness) and the *density* ρ of a medium.

$$c_v = \sqrt{\frac{B}{\rho}}$$

So, it increases with stiffness, and decreases with density. Listed on the next page are different speeds of sound inside of various media [9]. The bulk moduli of woods are given parallel to (along) the grain. All of these numbers are variable, and the numbers are averaged if a range was given.

Medium	$B (\times 10^9 \text{ N/m}^2)$	$\rho (\text{kg/m}^3)$	c_v
Dry air (20° C)	0.000142	1.21	343 m/s
Water (25° C)	2.15	965	1493 m/s
Salt water (25° C)	2.34	1022	1533 m/s
Ebony	13.8	1200	3391 m/s
White oak	11	770	3780 m/s
Honduras mahogany	10.4	650	4000 m/s
Indian rosewood	12.0	740	4027 m/s
White ash	12.2	750	4033 m/s
Engelmann spruce	9.0	550	4036 m/s
Red maple	11.3	675	4092 m/s
Black cherry	12.2	630	4401 m/s
Steel	200	7820	5057 m/s
Glass	70	2600	5189 m/s
Brazilian rosewood	16.0	830	5217 m/s
Diamond	442	3500	11238 m/s

Table 2.1: The speed of sound c_v in common acoustic materials is given by $c_v = \sqrt{\frac{B}{\rho}}$ where B is the stiffness and ρ is the density.

The bulk modulus B describes the volumetric elasticity (three-dimensional), while *Young's modulus* describes the tensile or linear elasticity (two-dimensional). For the different types of wood, B is actually Young's modulus. Both are ratios of stress to strain, measuring the resistance of a material to uniform compression. Hence, both the bulk modulus and Young's modulus represent the *inverse of compressibility*.

Reverberation

Reverberation refers to sound reflecting against walls, refracting into absorbent material, and dissipating in air after its origination. We talk about reverberation especially in *room acoustics*, where a recording studio should ideally have no reverberation, but a cathedral may have a lot of reverberation. Its graphical representation is different: Once

again, the horizontal axis is time and the vertical axis is amplitude, but this is not the signal itself. Instead, Figure 2.15 shows us the amplitude of events over time.

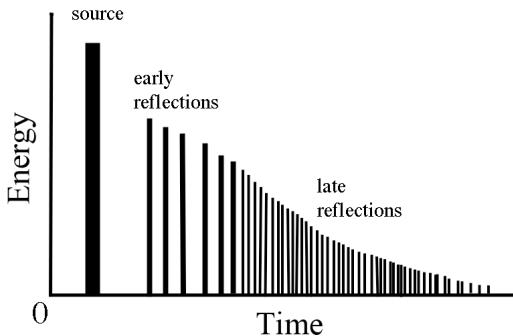


Figure 2.15: Reverberation is typically generalized as three main events: The source signal, its early reflections (with thicker vertical lines), and its late reflections (the thinner vertical segments). The time of the early reflections with respect to the source sound is how a source will seem near or far from the location of the observer.

The first event is the original sound—the source signal. As the source sound propagates in the room, it bounces off each of the walls. The first time it does this and returns to the receiver is depicted in the early reflections event. In the above graph, there are six early reflections representing six walls or surfaces, which is typical in a rectangular room with four walls, a floor, and a ceiling. The later the reflection, the farther the surface is from the receiver. The weaker the reflection, the longer the sound has traveled and the higher absorbency of the surface material. The late reflections event depict the later bounces off of these surfaces with gradually less energy.

The *frequency response* of a reverberant space like a room or musical instrument is calculated by exciting the space with all frequencies in its range at a constant pressure and transforming this recording with a Fourier transform to deduce its resonant frequencies, which appear as peaks in the frequency response. This can be done by exciting the instrument with a sine *sweep* (a pure tone that oscillates from low to

high frequencies) and recording the instrument's vibration. In room acoustics, the frequency response is typically calculated by playing a burst of white noise, because it has equal energy at all frequencies. The Fourier transform of white noise is perfectly flat, reflecting the equal power of the frequencies, and the Fourier transform of the recording of the white noise in a room will have bumps where the room is resonating or attenuating sound on a frequency basis. This burst of white noise is also called an *impulse*, and a space's reaction to it is called an *impulse response*. Impulses can also be taken with other loud, brief, noisy things like fireworks, balloon pops, and handclapping, though their frequency response is naturally more variable than that of white noise.

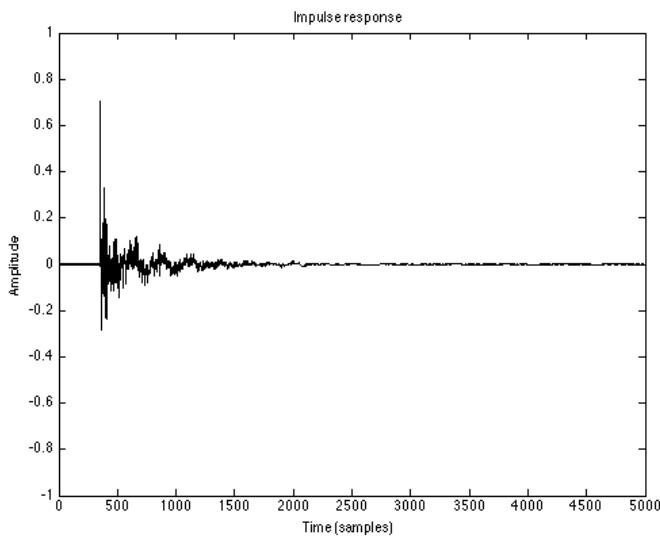


Figure 2.16: The impulse response of a room is a function of time. The general shape of the amplitude envelope is decreasing, but there are peaks where the sound is reflecting off of surfaces in the room. The horizontal axis is in samples, not seconds, so this impulse response lasts less than half of a second.

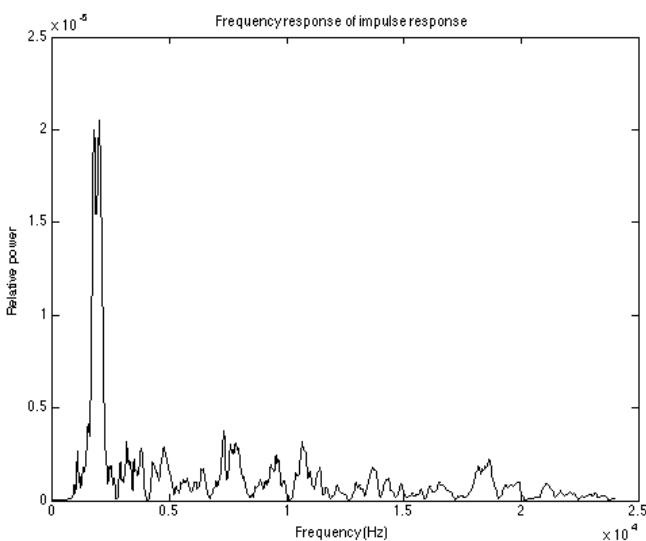


Figure 2.17: The frequency response of the same room as in Figure 2.15. This is a function of frequency. Rooms typically have resonances in the lower frequency range because of their larger dimensions compared to musical instruments.

To calculate the frequency response, we simply take the Fourier transform of the impulse response, depicted in Figure 2.17.

Room acoustics is a constantly expanding field of research in the scientific study of sound. A standard way of measuring the reverberation of an acoustic space is by calculating the RT_{60} , the time that a sound (typically wide-band or narrowband noise) takes to decay by 60 decibels in that space. Architectural structures built for a musical purpose like auditoriums and studios use room acoustics to choose materials and dimensions that will best amplify or attenuate certain frequencies. The table and graph depicted in Figures 2.18 and 2.19 define the absorption that some materials have with respect to sound, and you can see a direct correlation between the hardness of the material and how much sound it absorbs.

Material.	Frequencies in cycles/sec.		
	250.	500.	1,000-2,000.
Ordinary wall and ceiling surfaces:			
Lime plaster	0·02-0·03	0·03-0·04	0·03
Hard plaster	0·01-0·02	0·01-0·02	0·02-0·03
Unpainted brick	0·03	0·03	0·05
Wood-paneling, 3-ply	0·01-0·02	0·01-0·02	0·01-0·02
Curtains:			
Cretonne	—	0·15	—
Medium weight	—	0·2-0·4	—
Heavy, in folds	—	0·5-1·0	—
Floor coverings:			
Wood block in mastic	0·03	0·06	0·10
Cork carpet, $\frac{1}{4}$ in. thick	0·03	0·07	0·20
Porous rubber sheet, $\frac{1}{4}$ in. thick	0·05	0·05	0·20
Axminster carpet, $\frac{1}{4}$ in. thick	0·05	0·10	0·35
" " on $\frac{1}{4}$ -in. felt underlay	0·05	0·40	0·65
" " on $\frac{1}{4}$ -in. rubber "	0·05	0·20	0·45
Turkey carpet, $\frac{1}{2}$ in. thick	0·10	0·25	0·30
" " on $\frac{1}{2}$ -in. felt underlay	0·30	0·50	0·65
Special absorbents:			
Acoustic plasters ($\frac{1}{2}$ to 1 in. thick) on stone	0·15	0·25	0·30
Fibre boards, plain, $\frac{1}{4}$ in. thick, on battens	0·30-0·40	0·30-0·35	0·25-0·35
Medium efficiency tiles, on battens	0·40	0·40	0·50
High-efficiency tiles, with perforated surfaces, on battens	0·50	0·80	0·85
Acoustic felts, 1 in. thick, perforated covers on hard surface	0·30	0·70	0·80
Acoustic felts, $\frac{1}{2}$ in. thick, on battens	0·25	0·45	0·70
Wood wool-cement board, 1 in. thick, on battens	0·30	0·60	0·70
Sprayed asbestos, 1 in. thick	0·50-0·60	0·65-0·75	0·60-0·75
Slag wool or glass silk about 2 in. thick, on battens	0·70	0·85	0·90
Cabot quilt, 3-ply, two layers	0·40	0·70	0·70
Individual objects in open-window units (ft.):			
Audience per person	4·3	4·7	5·0
Chairs, bent ash	0·16	0·17	0·21
Cushions, hair, $2\frac{3}{4}$ sq. ft. under canvas and plush	1·1	1·8	1·5

Figure 2.18: The absorption coefficients of various materials for the frequencies 250 Hz, 500 Hz, and 1000-2000 Hz. This table comes from Alexander Wood's *The Physics of Music* [10].

Much of the results depicted in Figures 2.18 and 2.19 come from some of the first discoveries concerning room acoustics by Wallace Clement Sabine (1868-1919) of Harvard University, who found that

$$T = 0.161 \frac{V}{AS}$$

where T is the reverberation time, V is the volume of a reverberant space in cubic meters, A is the average absorption coefficient, and S is the surface area of the material. A is strictly less than 1 because 100

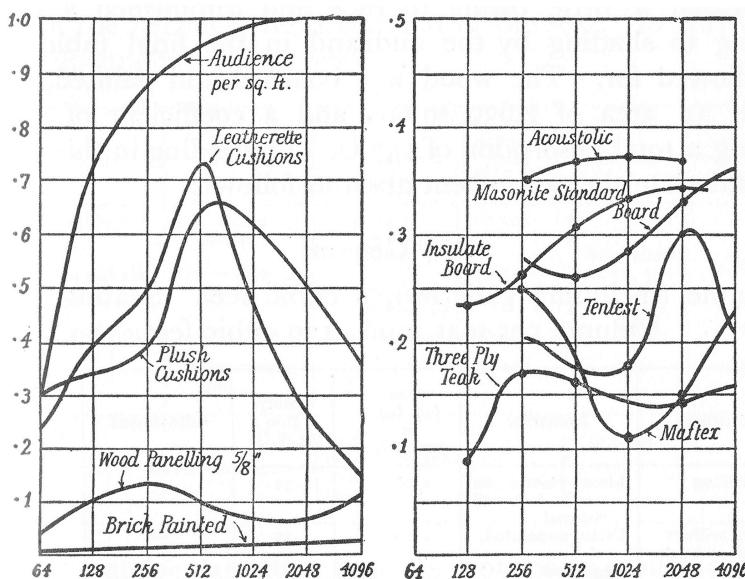


Figure 2.19: The absorption curves of various materials over the frequency range 64-4096 Hz [10].

percent absorptive material does not exist, but as a point of reference, one square meter of 100% absorptive material is called *1 sabin*.

Interference

Let us begin with a theorem from mathematics.

Theorem: For any real numbers $a, b \in \mathbb{R}$,

$$|a + b| \leq |a| + |b|.$$

Proof: Let $a, b \geq 0$. Then the result $|a + b| = |a| + |b|$ is immediate. Let $a, b \leq 0$. Likewise, it is clear that $|a + b| = |a| + |b|$. Finally, let a be the opposite parity (sign) of b . Then $|a + b| < |a| + |b|$. So, for all values of a, b in \mathbb{R} , $|a + b| \leq |a| + |b|$.

Constructive interference is only satisfied by the maximal case, $|a + b| = |a| + |b|$, and otherwise destructive interference is occurring.

There are two types of interference in sound: Constructive and destructive. Constructive interference occurs when two waves, call them $x_1(t)$ and $x_2(t)$, interact such that

$$|x_1(t) + x_2(t)| = |x_1(t)| + |x_2(t)|.$$

In words, the magnitude of their sum is equal to the sum of the magnitude of each wave. It can be shown that the sign of the two waves must be the same.

Destructive interference is the exact opposite of constructive interference. Destructive interference is such that

$$|x_1(t) + x_2(t)| < |x_1(t)| + |x_2(t)|.$$

For this to be true, the signs of the two waves must be opposite, as given in the above proof. Therefore, when these waves interact, they have a detrimental effect on the overall pressure of the air through which they propagate. Constructive or destructive interference can only occur when the waves intersect at the same location, whether at a single point or set of points, and at the same instant or same interval of time.

Cancellation is a result of *completely* destructive interference, wherein $|x_1(t) + x_2(t)| = 0$. Consider two identical sinusoids, $x_1(t) = x_2(t) = A \sin(2\pi ft + \phi)$. Now imagine that you have two speakers facing each other, located exactly at an integer multiple of the wavelength of the sinusoids ($\lambda = v/f$, remember) apart from one another, both connected to your CD player in stereo. Let one channel be $x_1(t)$ and the other be $x_2(t)$. When you press play, the waves travel from the speaker to the opposite speaker at the same time. Because they are placed an integer multiple (4 times) of their wavelengths apart, the crests of one wave will occur exactly where the troughs of the second wave occur. These two waves are called *completely out of phase* from each other:

Their phases are different by π radians, or 180° . This is the most that two waves can be out of phase, even though a circle contains 360° . The waves become reflections of each other because they are exact opposites, flipped about the horizontal axis, and become *in phase* with each other when there is no difference (angle) between their respective phases.

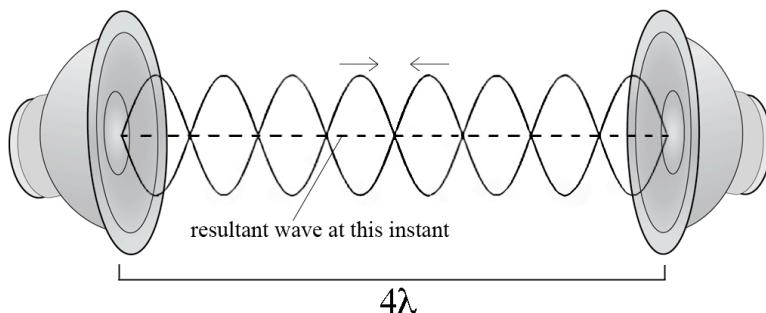


Figure 2.20: Two speakers exhibiting completely destructive interference. The superposition of their respective sounds waves is shown by the dotted line. Since one wave is moving to the left and the other to the right at the same speeds, the wave is 0 only 2 times per period (much like a regular sine wave), but it does not travel—it stands. Hence, the result is a standing wave, which causes acoustic feedback.

So, their superposition is zero everywhere at certain instants—at their compressions and rarefactions, to be more precise. These two waves form what is called a *standing wave*. A standing wave occurs when two sine waves of equal frequency travel at the same velocity in opposite directions, so their velocities, v_1 and v_2 , sum to zero where $v_1 = -v_2$ (i.e., the wave does not propagate—it *stands*). This happens in musical instruments along a fixed string or in a column of air. A standing wave is perfectly stationary, but its amplitude changes periodically at the same frequency of the two waves.

Standing waves cause acoustic feedback because they resonate in an acoustic space and thus have more sustain, causing a microphone and speaker to continuously receive and transmit them when recording. Say that a room is $27' \times 25' \times 12'$. Then the frequencies with

wavelengths equal to 27 feet, 25 feet, or 12 feet (41.3 Hz, 44.6 Hz, and 92.9 Hz) will stand in this room, and furthermore, integer multiples of these frequencies will also cause feedback (though, to a lesser degree) because their reflections will mirror their propagations. This means that larger rooms will have very low-frequency resonances and small corridors (and musical instruments) will have higher frequency resonances, because a wavelength of 30 centimeters translates to about 1143 Hz.

Resonators and noise-canceling headphones are effective in diminishing the power of undesirable frequencies. Resonators tuned to the undesired frequency will capture and reduce the frequency by creating a standing wave: The wave carrying the undesired frequency is attracted to the resonator, and the resonator absorbs and dissipates the wave by cancelation. Noise-canceling headphones detect noise via an exterior microphone near the ear. The noise is directed to an electric circuit that transforms it into an antinoise signal—a signal exactly out of phase with the detected noise. This antinoise signal is played through the headphones to cancel the noise.

The inverse square law

All forms of radiation obey the inverse square law, which simply says that the farther you are away from a source of energy, the less intense the energy will be. In a uniform medium, a source propagates in all directions equally, so we model this motion in three dimensions as a sphere. The intensity I at a radius r from a sound source with original power P will be $I = \frac{P}{4\pi r^2}$, because the surface area of a sphere is given by $4\pi r^2$.

However, because we measure the intensity of sound with decibels (dB), which are a logarithmic unit, the inverse square law returns a different equation for sound waves than the one listed above. Intensity, as we will explore in more detail in Chapter 3, is proportional to the

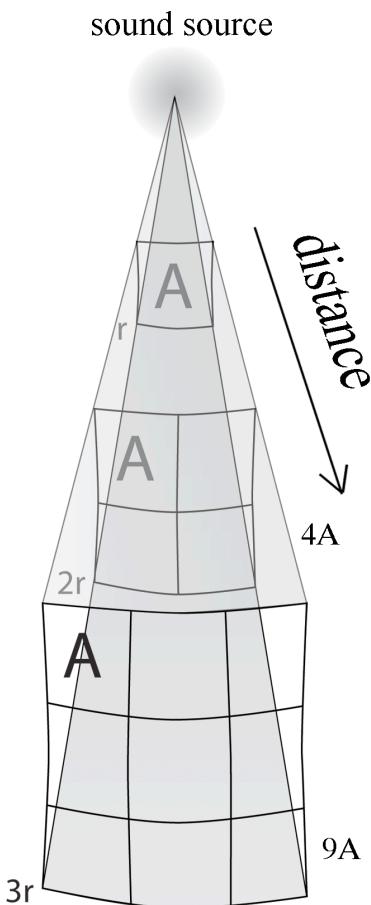


Figure 2.21: Three-dimensional depiction of the inverse-square law

square of sound pressure, P^2 . Therefore, source intensity becomes proportional to I/r^2 at a distance of r from it, and the pressure is then P/r , not P^6 . We say "proportional to" in mathematics when a ratio exists between two quantities, but their relationship is not necessarily the same in every scenario, i.e., the ratio may fluctuate. The number of days that it rains in a year, for example, is proportional to the annual

⁶Note that only intensity and pressure diminish with distance, not frequency or wavelength. Red does not get any "less red" the farther we are away from it.

inches of rain accumulated in a year, but x -many days of rain does not necessarily mean y -many inches of rainfall.

The brain treats the ears like two distinct microphones, not as a total or average [4]. It relies heavily upon the inverse square law to detect the proximity of sources, while the distance between the ears and the physicality of the pinnae (the flaps of skin external to the skull) provide information about the sources' directionality. The primary function of hearing, or of any sensation for that matter, is to alert the hearer of threat to its survival. The sensation of sound can quickly activate our adrenal glands, and can thereby serve to inform the proper fight or flight response. Music, too, has the power to elicit very strong emotional reactions, including fear and anger.

Another aspect of sound that requires a physical explanation is our ability to hear sounds from sources outside of the room. In my office, I can hear footsteps approaching from around the corner and the elevator bell, even though the elevator is located far down the hall. But these sounds all seem to be coming from my doorway. This can be explained by *Huygens' principle*, depicted in Figure 2.22.

Huygens' principle: Every point of a moving wave is also the center of a new source, each propagating a fresh set of waves in all directions.

This is also known as *diffraction*, and explains why the sound from a loudspeaker can be heard at locations behind it, above it, and to its left and right. This is true of light as well, but because the wavelength of light is so short (about 390 to 750 nanometers) due to its very large frequency (400-790 Terahertz, where 1 Terahertz (THz) = 10^{12} Hz), it is not as easily perceived.

The elasticity of air is what allows sound and light to vibrate, so there is neither light nor sound in a vacuum. Other qualities of air and the Earth's atmosphere can have interesting effects on traveling waves.

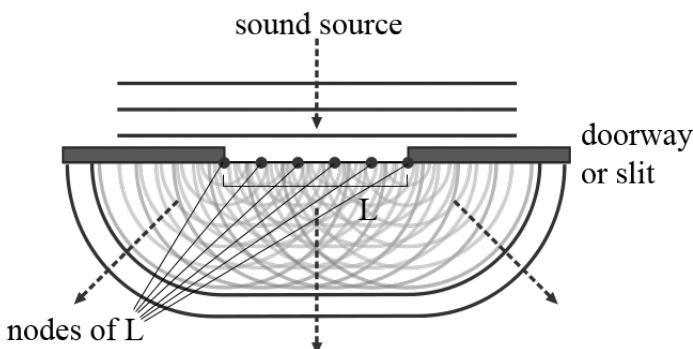


Figure 2.22: Sounds originating on the other side of an open doorway will appear to originate from the doorway itself, states Huygens' Principle.

The effects of temperature, humidity, velocity, and altitude

Waves move differently depending on the media through which they travel, as stated in the discussion of refraction. Temperature, humidity, and altitude are all directly related to atmospheric pressure, which is itself a result of gravity. Sound waves vibrate easiest in high-pressure areas where there are fewer forces working against their energy. Since pressure decreases as elevation increases, sound waves tend toward the ground. Sound travels slowly and loses energy faster in hotter temperatures because heat rises. As the hot air moves upwards, it takes some of the sound waves' energy with it.

Temperature's effect on pitch is most noticeable in wind instruments, due to the expansion of their bores from heat. A flute, for example, rises in pitch about 0.002 Hz for every 1°C (1.8°F) rise in temperature. The tuning of piano strings increases about 0.00001 Hz for each 1°C increase in temperature because hotter strings expand.

The velocity of a sound can be calculated as before, in the section on refraction, but also from the derivative of the pressure of a medium with respect to its density:

$$v = \sqrt{\frac{\partial p}{\partial \rho}} = \sqrt{\frac{B}{\rho}}$$

where p is the pressure of a medium ρ is once again its density. Therefore, the bulk modulus can be determined by

$$B = v^2 \rho = \frac{\partial p}{\partial \rho} \rho.$$

In 0% humidity (dry) air,

$$v = 331.3 \sqrt{1 + \frac{t}{273.15}}$$

for temperature t in degrees Celsius.

As mentioned in the discussion of refraction, a sound's propagation speed is dependent upon the medium through which the sound travels. A final way of calculating these speeds is with the *Mach number* of a given medium, where the Mach number of dry air is 1. A Mach number greater than 1 indicates that sound is traveling at a supersonic speed. We can calculate the Mach number with the equation

$$M = \sqrt{\frac{2}{\gamma - 1} \left[\left(\frac{q_c}{p} + 1 \right)^{\frac{\gamma-1}{\gamma}} - 1 \right]}$$

where M is the Mach number, q_c is impact pressure of the medium, p is the pressure of the medium, and γ is the ratio of the pressure of the medium to heat—the volume constant. This equation comes from Bernoulli's principle in fluid dynamics.

Humidity has a small but detectable effect on sound propagation, due to the presence of lighter and more elastic water molecules in the air. As you may guess, the velocity of sound increases in humid air, up to 0.6%. Since the density of air is lower at higher altitudes than at sea level or below it, the speed of sound decreases as its altitude increases.

The speed by which sound travels affects the sound's volume at a given distance, and unsurprisingly, the faster sound travels, the better it maintains its original intensity. Wind will additively or subtractively affect the speed, working as you may suspect: When wind is blowing in the direction of the sound, it increases its velocity, and thus its loudness.

Finally, when an observer or sound source is moving, the *Doppler effect* causes the wavelength of sounds to change. As a source moves closer to an observer, every period of the sound wave gets increasingly shorter, causing the period to get smaller and frequency to get larger. Conversely, waves moving away from an observer will have increasingly larger periods, causing frequency to decrease. Austrian physicist Christian Doppler witnessed and quantified this in 1842 with the mathematical formula

$$f_o = \left(\frac{c + v_o}{c + v_s} \right) f_s,$$

where f_o is the frequency heard by the observer, c is the speed of sound in the medium (343 m/s in air), v_o is the velocity of the observer, v_s is the velocity of the sound's source, and f_s is the frequency of the source. The observer's velocity v_o will be positive if the observer is moving *towards* the source, and v_s will be positive if the source is moving *away* from the observer.

2.6 Chapter summary

This chapter investigated the four main properties of waves—*amplitude*, *frequency*, *period*, and *phase*—and their behavior in an ideal world and in reality. Also distinguished were the terms *signal* and *noise*. These terms are important for understanding the compression of information, a topic raised at the end of Chapter 5.

Sound waves can be decomposed into a sum of simple sine waves, says *the principle of superposition*. It is easy to realize the amplitude (A), frequency (f), period ($1/f$), and phase (ϕ) of a simple sine wave: It is of the form $A \sin(2\pi ft + \phi)$. When two waves are related in frequency by a small-number integer ratio, we say that the (musical) interval between them is *harmonic*. Amplitude explains the amount of *pressure* induced by a sound wave in a medium. Sound can only be heard when the pressure of a medium is varying, as changing pressure signifies

a disturbance in the eardrum. For example, when we go to places at high altitudes, there is a lower pressure, but no specific sounds are associated with this change.

The phase of a wave in relation to the phase of another wave determines how two or more waves will *interfere* with each other when they interact in a medium. When the resultant wave is less in amplitude than the sum of the amplitudes of the original waves, it is interfering destructively. Otherwise, it is undergoing constructive interference. Waves traveling in exactly opposite directions and at identical frequencies create *standing waves*. For the same reason, standing waves also happen in musical instruments, on a fixed string, and in columns of air.

Sound waves reflect off of objects, refract in different media, and lose energy as they dissipate (the *inverse square law*). All of this is *reverberation*, which describes the behavior of sound after it originates from a source like a speaker or musical instrument. Every point through which a sound wave travels is also the source of a new set of waves, states *Huygens' principle*, but this source is not thought of in the same way as a speaker or musical instrument.

A hotter environment will increase the frequency of the sound passing through it. Increasing the humidity and density of a medium increases the velocity of sound waves. Sounds moving towards an observer will have increasingly shorter wavelengths and thus higher frequencies, and the converse is true (the *Doppler effect*). Finally, sound waves are attracted to high-pressure areas where their energy will be most conserved, and since pressure decreases with altitude, all sound waves tend towards the ground.