Chapter 12
Loudness Perception

Loudness is a familiar property of sound. If you are listening to the radio and you want to increase the loudness you turn up the volume control, i.e., you increase the intensity. Clearly loudness is closely related to sound intensity. But loudness is not the same as intensity. Sound intensity is a physical quantity—it can be measured with a sound level meter. Loudness, by contrast, is a psychological quantity—it can only be measured by a human listener. The purpose of this chapter is to introduce the important elements of sound that contribute to the perception of loudness.

12.1 Loudness Level

On physical grounds alone you can expect that loudness should be different from intensity because the ear does not transmit all frequencies equally, i.e., it does not have a flat frequency response. The outer ear, with its complicated geometry, has several overlapping resonances including the important ear canal resonance near 3 or 4 kHz. Therefore, the magnitude of the input to the nervous system does not depend on intensity alone, the frequency also matters.

The first step on the road to a scale for loudness is a measure of intensity that compensates for the effect of frequency on the loudness of sine tones. Such a measure is the loudness level, and it is expressed in units of phons (rhymes with "Johns"). By definition, any two sine tones that have equal phons are equally loud. The curves of equal phons are called equal-loudness contours. A set of equal-loudness contours is shown in Fig. 12.1.

The contour concept is not unusual. You are probably familiar with contour maps in geography where the vertical and horizontal axes represent north–south and east–west in the usual way, and lines on the map represent equal-elevation contours. Thus one line might be the 100-m elevation line while the next line might be the 110-m line. Equal-loudness contours are like that, except that instead of equal elevation, the lines are equal loudness.

W.M. Hartmann, *Principles of Musical Acoustics*, Undergraduate Lecture Notes in Physics, DOI 10.1007/978-1-4614-6786-1_12,

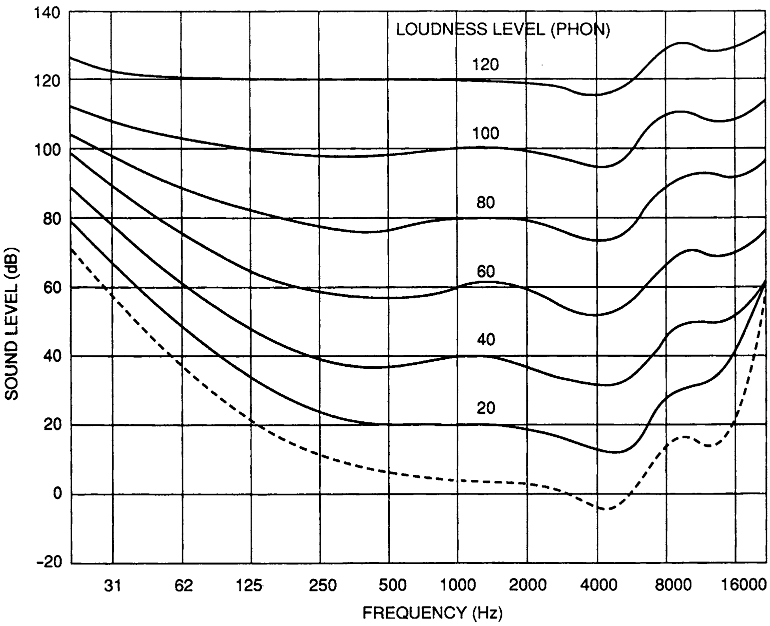

Fig. 12.1 Equal loudness contours come from loudness comparisons made by average human listeners with normal hearing for sine tones of different frequency and sound level. The *horizontal* and *vertical axes* are both physical properties of the tone. The human part appears in the *curves* themselves. Every point on a single curve describes a sine tone with the same loudness. The *dashed line* corresponds to absolute threshold, the weakest sound the average human can hear

To construct an equal-loudness contour requires an experimenter and a listener. The experiment begins when the experimenter presents a sine tone having a frequency of 1,000 Hz and some chosen level, say 20 dB. The listener listens to that tone. Then the experimenter changes the frequency, for instance to 125 Hz, and asks the listener to adjust the level of that 125-Hz tone so that it is just as loud as loud as the 1,000-Hz tone.

In this example, where the sound level of the 1,000-Hz sine is 20 dB, it is found that the sound level of the 125-Hz sine must be adjusted to 33 dB for equal loudness. These two points (1,000 Hz, 20 dB) and (125 Hz, 33 dB) are on the same equal-loudness contour, and, by definition, this contour is called the 20-phon loudness level. *The frequency of 1,000 Hz is always the reference frequency for the phon scale.* In what follows we will refer to a loudness level in phons by the Greek symbol Φ (Phi, for physical). The level called 0-dB in Fig. 12.1 corresponds to the nominal threshold, 10^{-12} W/m^2.

12.2 Loudness

The loudness level, expressed in phons, Φ, is a frequency-compensated decibel scale, but it is not yet a measure of loudness. A true measure of loudness would scale with the personal sensation of magnitude, but there is no reason to suppose that the numbers 10 phons, 20 phons, etc. will do that. After all, the scale of phons is tied to the decibel scale at 1,000 Hz, and the decibel scale was invented without much consideration of human perception of loudness. A logarithmic scale, like the decibel scale, is convenient mathematically, but careful experiments on human listeners have concluded that loudness is not a logarithmic function of intensity. Loudness does not follow the decibel scale well. Instead, many experiments done over the past 75 years have shown that the rule for loudness is a different kind of function. It is a exponential law, as follows:

$$\Psi \propto 10^{0.03\Phi}, \tag{12.1}$$

where the symbol Greek Ψ (Psi for psychological) stands for loudness. An alternative way to write Eq. (12.1) is as

$$\Psi \propto 1.0715193^{\Phi}. \tag{12.2}$$

These proportionalities don't enable you to calculate any absolute numbers, but they do enable you to calculate ratios. Suppose there are two tones, numbered "1" and "2", that have loudness levels of Φ_1 phons and Φ_2 phons. Then the ratio of the loudnesses is given by

$$\Psi_2/\Psi_1 = 10^{0.03(\Phi_2-\Phi_1)} \tag{12.3}$$

The formula in Eq. (12.3) follows immediately from the proportionality in Eq. (12.1). Ratios can be useful concepts, as shown in the following example.

Example:

Question: There are two sine tones, "1" and "2." These tones have different frequencies and tone 2 is 10 phons greater than tone 1. How much louder is tone 2, compared with tone 1?

Answer: Although the frequencies are different, the fact that loudness levels are given in phons makes it possible to use Eq. (12.3). Plugging into Eq. (12.3) we find:

$$\Psi_2/\Psi_1 = 10^{0.03\cdot(10)} \tag{12.4}$$

or

$$\Psi_2/\Psi_1 = 10^{0.3} \tag{12.5}$$

From the inverse log function on a calculator we find that $\Psi_2/\Psi_1 = 1.995$. That's close to 2. Tone 2 is about two times louder than tone 1.

The example above has illustrated the convenient rule:

Rule 1: *For every increase of 10 phons, the psychological sensation of loudness doubles.*

Loudness vs Sound Level Because we regularly make physical measurements of the sound intensity in terms of the sound level in decibels, it would be really convenient to be able to relate loudness to the sound level, L. There are two requirements that need to be fulfilled in order for us to be able to do that. First, we need to consider a single frequency at a time. Thus we can compare the loudnesses of two tones that have different levels but the same frequency. In other words, we can ask about the effect of turning up the volume control. The second requirement is that the tones involved need to have particular levels and frequencies such that the equal-loudness contours are approximately parallel. A glance at the contours of Fig. 12.1 shows that this second requirement is met by many everyday sounds, having frequencies between 250 and 8,000 Hz and levels greater than 40 dB.

As a practical matter, therefore, we can rewrite Eq. (12.1) as

$$\Psi \propto 10^{0.03L} \tag{12.6}$$

We expect that this relationship will fail at frequencies below 125 Hz, where the equal loudness contours are not parallel but are squished together. We expect that this relationship will hold good in many practical circumstances. Exercises at the end of the chapter will assume that it holds good.

Given the relationship in Eq. (12.6) it follows that if there are two tones of the same frequency with levels L_1 and L_2, then the loudnesses are related by the ratio

$$\Psi_2/\Psi_1 = 10^{0.03(L_2-L_1)} \tag{12.7}$$

A special case of Eq. (12.7) leads to the convenient rule:

Rule 2: *For every increase of 10 dB, the psychological sensation of the loudness of a tone doubles.*

Rule 2 is more convenient than Rule 1 because it uses decibels instead of phons. But Rule 2 is not as general as Rule 1 because it requires that special conditions hold—constant frequency and parallel equal-loudness contours.

Loudness vs Intensity From Chap. 10, we know that the sound level (L) in decibels can be related to the sound intensity (I) expressed in watts per square meter. The relationship is that $I \propto 10^{L/10}$. For tones in the practical region of frequency and level, it follows from Eq. (12.6) that

$$\Psi \propto I^{0.3}. \tag{12.8}$$

This relationship is a power law. It says that loudness grows as the 0.3 power of the intensity. This statement is close to saying that the loudness grows as the cube root of the intensity. The cube root would be the 1/3 (or 0.333...) power, and that is close to 0.3.

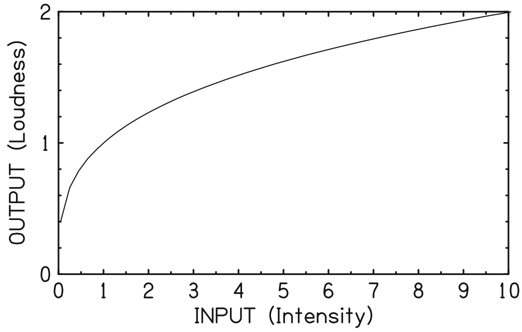

Fig. 12.2 The 0.3 power law is a compressive function. As the input goes from 1 to 10, the output goes from 1 to 2

This is quite an amazing rule. Suppose you have a tone with a sound level of 50 dB. Therefore, the intensity is 10^{-7} W/m^2. If you increase the level to 60 dB the intensity becomes 10^{-6} W/m^2, and the loudness becomes twice as great. If you increase the level by another 10 dB the intensity becomes 10^{-5} W/m^2, and again the loudness doubles. This behavior indicates a highly compressive function. Overall, starting from 50 dB, the intensity has increased 100 times but the end result is only four times louder according to human ears. This behavior is called, "compressive" because the output grows slowly as the input increases. Here, one imagines the input to be the physical intensity, put into the human auditory system, and the output is the listener's sensation of loudness. A graph of a compressive function has a shape like Fig. 12.2.

Example:

Question: There are two sine tones, "1" and "2." These tones have the same frequency, but tone 2 is 300 times more intense than tone 1. How much louder is tone 2, compared with tone 1?

Answer: From the proportionality in Eq. (12.8), you know that

$$\Psi_2/\Psi_1 = (I_2/I_1)^{0.3} \tag{12.9}$$

We are told that $I_2/I_1 = 300$, and so $\Psi_2/\Psi_1 = (300)^{0.3} = 5.5$. Thus tone 2 is about five times louder than tone 1.

12.3 Psychophysics

The delicate matter of scaling a psychological magnitude such as loudness is the subject of the science of psychophysics. The psychophysicist tries to determine the relationship between physical magnitudes, measured with ordinary electronic instrumentation, and psychological magnitudes as reported by human (or animal) observers.

There are several different techniques used to determine psychological magnitudes. In a *direct* technique, the experimenter just asks subjects to rate their sensations of magnitude. For instance, a listener might be presented with 1,000-Hz sine tones having a dozen different intensities. The listener is asked to rate the loudness of the tones. This is an example of a *magnitude estimation* task.

It is not clear that this kind of experiment should work. As you can imagine, if you ask ten listeners to estimate the magnitude of a tone, you are likely to get ten different answers. However, if the magnitude estimation experiment is continued for many hours over the course of a week or two, certain clear patterns begin to emerge from the data. Although different listeners may use different scales, there are common features (ratios mostly) in their responses.

After the experimenter has found stable numerical responses to tones of different intensities in the estimation task, the experiment is reversed to form a *magnitude production* task. This time, the experimenter gives the listener the numbers, and the listener adjusts a volume control to vary the intensity of the sound. If the listener previously gave numbers on a scale of 1–100 in the estimation task, then the experimenter will give the listener the same range of numbers for the production task. After another week of experimenting, the data from the two tasks are averaged, and the result is a function showing the relationship between physical magnitude and psychological magnitude. For the loudness experiment, the relationship is that given in Eq. (12.1).

The Sone Scale The exponent of 0.3, as in Eq. (12.8), is the basis for the current national and international loudness scale known as the *sone* scale. The reference for this absolute scale is that a 1,000-Hz sine tone with a level of 40 dB SPL shall have a loudness of one sone. Therefore,

$$\Psi(\text{sones}) = \frac{1}{15.849}(\frac{I}{I_0})^{0.3}, \tag{12.10}$$

or

$$\log \Psi(\text{sones}) = -\log(15.849) + 0.03 \times 10 \log \frac{I}{I_0}. \tag{12.11}$$

It follows by definition that any tone with a loudness level of 40 phons has a loudness of one sone, and sones can be calculated from phons by the equation

$$\log \Psi(\text{sones}) = -1.2 + 0.03 L_{\varphi}, \tag{12.12}$$

where L_{φ} is the loudness level in phons.

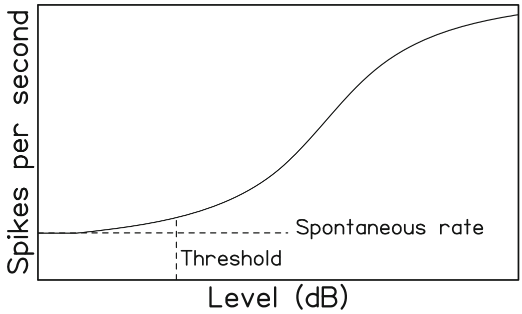

Fig. 12.3 Neural firing rate vs sound level in dB. The threshold is defined as the level where the firing rate becomes 10% greater than the spontaneous rate

12.4 Neural Firing Rate

A scientist finds it natural to try to understand the perception of loudness from what is known about the auditory system and the properties of neurons that process sounds in our brains. It is very compelling to imagine that the loudness sensation is related in some way to the total number of neural spikes transmitted to the central auditory system by the auditory nerve.

As the intensity of a sound increases, the firing rate of neurons in the auditory nerve also increases, but the whole story is not as simple as that. First, even when there is no sound at all, the neurons fire occasionally. This spontaneous firing rate may be about 50 spikes per second. Now suppose that a very weak tone is turned on. For very low intensity levels a neuron's firing rate remains at the spontaneous rate, as shown in Fig. 12.3. As the intensity increases it eventually passes a threshold where the neural firing rate starts to become greater than the spontaneous rate. As the intensity continues to increase, through a range of 30 or 40 dB above threshold, the neuron fires more and more often. But eventually, as the intensity increases still further, the neuron starts to saturate. Its firing rate no longer goes up rapidly with increasing intensity. In other words, neural firing rate looks like a compressive function of intensity. It is a good bet that the loudness of a sound is closely related to the number of neural spikes received by the central auditory system in a 1-s interval, i.e. to the neural firing rate. Therefore, it seems likely that the compressive behavior of loudness observed in the psychology of human hearing is closely related to the compressive behavior of neural firing rate observed physiologically.

12.5 Excitation Patterns

We must remember that the neurons described above in the discussion of neural firing rate are originally excited by a vibration pattern (excitation pattern) on the basilar membrane. The neurons that are active are those at places on the basilar membrane that are excited by frequencies in the tone. If the tone is a sine tone, there

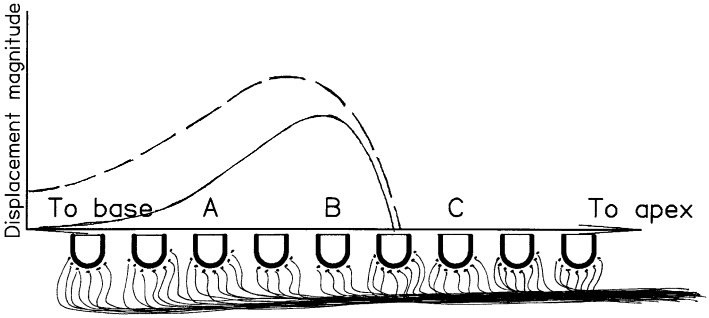

Fig. 12.4 Cartoon showing basilar membrane motion for a sine tone (*solid*) and a more intense sine tone (*dashed*). Three representative hair cells are given labels A, B, and C

is only one frequency and the excitation pattern has a single peak, as shown by the solid line in Fig. 12.4. Therefore, only a few neurons are excited. As the intensity grows, the excitation pattern becomes somewhat wider, as shown by the dashed line, and additional neurons are excited. This is a second mechanism for transmitting the sensation of loudness to the brain. Not only do active neurons become more active (e.g., the neuron at location B) but additional neurons (e.g., the neuron at location A) start to fire beyond their spontaneous rate. The additional spikes contributed by the neuron at A will help somewhat to generate the loudness sensation when the neurons near location B saturate. However, because of neural compression, a law of diminishing returns applies when the only neurons that contribute are those that respond to a single frequency. To increase the total firing rate a lot, to get a really loud sound, you need more than one frequency. You need a complex tone. For instance, adding a low-frequency component should excite the neuron at location C. A broadband tone with components at frequencies that span the audible range excites all places on the basilar membrane and should be particularly loud. The compressive character of neural firing means that the best way to generate loudness is to spread the available power over a broad frequency band rather than over a narrow band where it excites only a few neurons. An example of this effect appears in the next section.

12.6 Complex Sounds

To review where we have been in this chapter, recall that the chapter began with sine tones with particular levels (L in dB) and frequencies (f) and put these tones on equal-loudness contours denoted by loudness level (Φ in phons). Next, we related the loudness level (Φ) to the sensation of loudness (Ψ) through an exponential law. Finally we identified a practical region of levels and frequencies where the equal-loudness contours were approximately parallel, and we related loudness (Ψ) to level (L) and then to intensity (I). We found that loudness was proportional to the 0.3 power of the intensity. We related this power law to facts about the auditory nervous system.

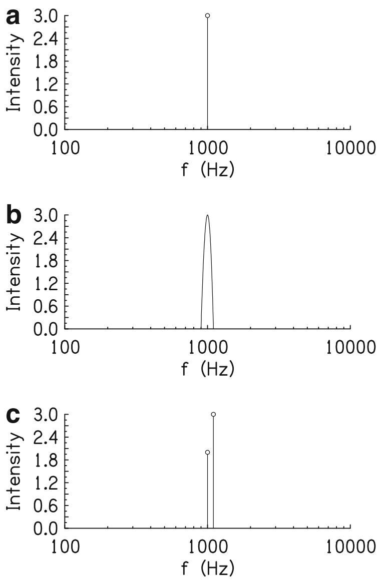

Fig. 12.5 Spectra for loudness calculations: (**a**) sine, (**b**) whistle, (**c**) two closely-spaced sines, (**d**) two separated sines, and (**e**) complex periodic tone. The *horizontal axis* is a logarithmic frequency scale, an approximation to the scale of the real basilar membrane

After all this work, it is a little disappointing to realize that the only sounds that we have considered so far are sine tones (Fig. 12.5a). The real world consists of complex sounds, made up of many sine tone components. Fortunately, it is not too hard to use what we already know to make good guesses about the loudness of complex sounds. The key to our approach is to assume that the sensation of loudness ought to depend on the total number of neural spikes received by the brain. We consider three kinds of sounds in turn.

1. Narrow-band sounds: If you whistle, you create a tone that is almost a sine tone, but it fluctuates in time. This fluctuation appears spectrally as a narrow band of noise, centered on a frequency that corresponds to the pitch of the whistled tone (Fig. 12.5b). The neurons that respond to the narrow band of noise are essentially the same neurons that would respond to a sine tone and we feel confident about using the entire sine-tone model to calculate the loudness of a narrow-band noise. All the components that make the tone noisy are compressed in the same way.

A second example of narrow-band sounds is a few (for instance two) sine tones with frequencies that are close together, e.g., 1,000 and 1,100 Hz (Fig. 12.5c). These tones excite mostly the same neurons and the excitations from both tones are compressed together. Thus if one tone has an intensity of 2 units and the other tone has an intensity of 3 units, the loudness will be proportional to $(2+3)^{0.3} = 5^{0.3}$ or 1.62.

2. Separated sine tones: If there are two sine tones with quite different frequencies then two quite different regions of the basilar membrane will be active and quite different sets of neurons will be excited. For instance, the frequencies might be 500 and 2,000 Hz (Fig. 12.5d). Excitations from these two sets of neurons will be compressed separately. Thus if one tone has an intensity of 2 units and the other tone has an intensity of 3 units, the loudness will be proportional to $(2)^{0.3} + (3)^{0.3} = 1.23 + 1.39 = 2.62$. You will notice that the loudness of the separated tones is considerably greater than the loudness of tones that are close together in a narrow band, i.e., 2.62 is greater than 1.62.

3. Harmonic tones: Let's consider a periodic complex tone with a fundamental frequency of 500 Hz and 16 harmonics (Fig. 12.5e). This tone has components at 500, 1,000, 1,500, ..., 8,000 Hz. To find the loudness of this tone we need to add up the contributions from all 16 components. What is interesting about this tone is that the low-frequency components 500 and 1,000 Hz are separated by a factor of 2 and excite different neurons in the auditory system. But the highest-frequency components, 7,500 and 8,000 Hz, are separated by only a factor of 1.07, and they excite mostly the same neurons. Thus, both kinds of addition—separated components and narrow band—must be done for such a sound.

12.7 Critical Band

The loudness calculations above point up the need for a way to decide whether the spectral components of a sound will excite the same neurons in the auditory system or different neurons. Strong compression occurs when only the same neurons are involved. Thoughts like these led to the concept of the *critical band*. If two spectral components are separated in frequency by more than a critical band, they excite different neurons. Two components that are separated by less than a critical band excite the same neurons. It should be evident that the all-or-nothing character of this definition of the critical band is only an approximation to what is really a continuum of interaction along the basilar membrane and in higher auditory centers. Nevertheless, psychoacoustical experiments have found reasonable reproducibility in measurements of the critical band width.

One way to measure the width of a critical band is by loudness, as suggested above. As the frequency separation of two tones increases, the loudness starts to increase when the separation exceeds a critical band width. Another way to measure is by *masking*. A weak tone is masked by an intense band of noise if the frequency of the tone and the frequencies of the noise fall within the same critical band. Masking means that the weak tone cannot be heard because of the intense noise. It is an amazing fact that if the weak tone and the intense noise band are separated in frequency by more than a critical band then there is almost no masking at all!

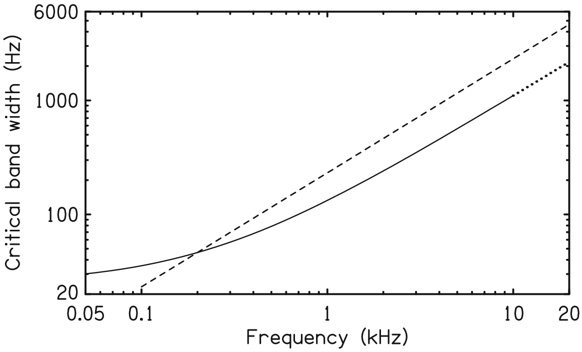

Fig. 12.6 Critical band widths are shown by the *solid line*. They were measured using bands of masking noise having a notch with variable width. One-third octave widths are shown by the *dashed line*. Because the plot is a log–log plot, the one-third octave reference is a *straight line*

Critical band widths determined from masking experiments are given in Fig. 12.6. The dashed line is a reference showing bandwidths that are one-third of an octave wide. Evidently the critical band is wider than 1/3 octave at low frequencies and narrower than 1/3 octave at high frequencies.

Exercises

Exercise 1, Equal loudness contours
What are the axes of the equal loudness contours?

Exercise 2, Phons
What is the loudness level in phons of a 300-Hz tone having a sound level of 40 dB?

Exercise 3, Sub-threshold or supra-threshold
Is it possible to hear a 200-Hz tone with a sound level of 10 dB?

Exercise 4, Threshold
In terms of the equal loudness contours, why is the threshold of hearing said to be 0 dB?

Exercise 5, Following the curves
What is the level of a 4,000-Hz tone that is as loud as a 125-Hz tone having a sound level of 50 dB?

Exercise 6, How to make loudness
Here is a loudness competition. You start with a 1,000-Hz sine tone having an intensity of 10^{-6} W/m^2. To increase the loudness, you are allowed to add another sine tone, also with an intensity of 10^{-6} W/m^2. Your goal is to choose the frequency of that added tone in order to lead to the greatest possible loudness. Which frequency do you choose? (a) 100 Hz, (b) 1,000 Hz, (c) 1,100 Hz, (d) 6,000 Hz.

Exercise 7, The compressive law of loudness

(a) Show that Eq. (12.8) agrees with the idea that increasing the intensity by a factor of 10 doubles the loudness. (b) If the intensity of tone 2 is 50 times greater than the intensity of tone 1, how much louder is tone 2 compared to tone 1?

Exercise 8, The compressive law used in reverse

How much must the intensity of a tone be increased in order to create (a) four times the loudness; (b) 16 times the loudness?

Exercise 9, Why is there a valley?

Why are equal loudness contours lowest in the frequency range from 3 to 4 kHz?

Exercise 10, Loudness and dB

The band conductor asks the trombone player to play louder. The trombone player increases the level by 7 dB. How much louder is the sound?

Exercise 11, Listening to your car radio

You are driving at high-speed on a highway listening to your car radio. There is a lot of road noise, and you need to turn up the level of your radio to hear the music. When you exit the highway and come to a stop at the traffic signal, your radio sounds much too loud. How does this relate to neural firing? It should be evident that the total neural firing rate in your auditory system is less when you are stopped than when you are driving at high speed. How can you explain the fact that the radio is louder when you are stopped?

Exercise 12, Terminology

Bozo objects to the last sentence of the exercise immediately above. He says that the radio isn't really louder when you are stopped—it just *sounds* louder. Set Bozo straight on the scientific definition of loudness.

Exercise 13, Bathroom fans

Consumers Reports (January 2004) rates the noise of bathroom fans in units of sones. Quiet fans are rated from 0.5 to 1.2 sones, and the loudest are rated at 4 sones. How much louder is a 4-sone fan compared to a 0.5-sone fan?

Exercise 14, Contour maps

As you look at a contour map, you see that at some places on the map the contours are far apart and at other places the contours are close together. From the spacing of the contours, you learn whether the terrain is rather flat or whether it is steeply rising. What can you infer from the equal loudness contours in Fig. 12.1. Compare low-level signals at 62 Hz with low-level signals at 1,000 Hz.

♠