Chapter 11

Complex Auditory Phenomena

Chapter Outline (Tentative):

I. Pitch Perception
  a. Basic Pitch Dimensions
  b. Pitch and Frequency
    i. Duration and Tone Quality
    ii. Missing Fundamental
      1. Masking and Critical Bands
      2. Masking and the Missing Fundamental

II. Music

III. Auditory Space Perception

IV. Speech

V. Auditory Illusions

Our auditory system provides a rich and complex experience of the world around us that is very different and also, in some cases, very complementary to what is provided by the visual system. For example, music is an auditory experience for which very different from most visual experiences. One of the features of perception that music perception relies on is the perception of pitch. We also use audition to determine the location of sound sources in our environment. This ability is crucial partly because as predators we do not see in all directions so audition helps respond to sounds in areas we cannot see. More importantly, the perception of the direction of a sound source is necessary because we respond slightly faster to auditory stimuli than visual stimuli (REF).

These two examples help illustrate the richness of the auditory experience. It is the purpose of this chapter to describe some of the richness that is provided by audition to our overall perceptual experience. These topics will build upon the material from the last chapter.

Pitch Perception

The first complex auditory experience that will be discussed will be music perception. However, before music perception can be discussed, the more basic experience of pitch perception needs to be covered. It is pitch perception that is the topic of this section of the chapter.

Basic Pitch Dimensions

Frequency is a continuous dimension going progressively from lower values to the upper regions of frequencies that we can hear. Pitch, on the other hand, is a much more complicated psychological experience. For example, pitch values, while going from the lowest pitch to the highest pitch, the pitch names of A through G are repeated, each with their associated sharp and flat values. Each repetition, going from C to C on the piano, represents one octave. Each octave is musically identical to each other octave except that some are perceived to be higher or lower than each other.

The octave is the most fundamental unit of pitch. Each octave represents a doubling of frequency. If one pitch of A is 440 Hz, then the A an octave higher is 880 Hz and the A the next octave is 1760 Hz. [THINK ABOUT THIS]

Pitch and Frequency

The relationship between pitch and frequency is one of the closest relationships between a psychological and physical dimension that has been found. It is a sufficiently close relationship that it is possible to buy tuners for instruments, such as a guitar, that helps a person tune their instrument by simply measuring the frequency of the fundamental tone produced by the instrument. In this section, some of the complexity of the relationship between pitch and frequency will be explored.

Duration and Tone Quality. Open **Interactive Illustration 11.x, Duration and Tone Quality [link to media]**. This illustration will explore one way that frequency is not a completely adequate description for pitch. When the screen comes up there will be a soundwave graph at the top of the screen. Below this graph there are several controls that will be explored, but first just press the **Play** button. A **1000 Hz** tone will be played for **1 msec**. The question is: How does what is being played sound? Like a tone or like a click? Most people report hearing a click. Drag the **Duration** slider out to **1000 msec** and press the **Play** button again. With this duration, the sound now sounds like a tone with a pitch. The only difference in the two tones is their duration. You can play a sequence of tones from short to long by clicking on the **Animate** button centered near the bottom of the screen. This function of the illustration will first play a tone of 1 msec. Then the next tone will be twice as long or 2 msec. Each next tone will be twice as long until a tone of just over 1 second will be played (actually 1024 msec because of all of the doubling). Listen to the sequence and see if you can determine the length of the sound necessary to be able to hear the sound as a tone and not a click. You can follow how long the each tone that is being played it going to be by wanting the values at the end of the **Duration** slider (which you cannot adjust during the animation). You can also adjust the frequency and intensity of the sound being played by using the **Freq** and **Intensity** sliders, respectively.

In previous research on this topic, it seems that the ability to hear a sound as a tone and not a click depends to some extent on the number of cycles present in the sound (Doughty & Garner, 1948, Rossing & Houtsma, 1986). So, you might hear the tone at shorter durations of time for higher frequencies than for lower frequencies. There is also some sense that the reason short duration sounds are hear as clicks even when they are complete cycles of sounds is related to temporal summation. Recall temporal summation from the discussion in vision. Temporal summation refers to the time during with information is averaged together. In this case, it may take a certain amount of time before the sound can be adequately processed to be heard to have a pitch as opposed to just being a click. In any case, this example still shows that it is not sufficient to just have a sound with a certain frequency to have the sound heard as a pitch. In the next example, the relationship between pitch an frequency will become even more complex.

Missing Fundamental. The phenomenon known as the missing fundamental is one of the more curious phenomenon in audition, a field of study replete with curious phenomena (Houtsma & Goldstein, 1972). Open **Interactive Illustration 11.x, Missing Fundamental [link to media]**. When you open the figure you will see in the **Soundwave Graph** at the top of the screen a graph of a complex soundwave made up of the components with the frequency and amplitude that is shown in the **Frequency Spectrum Graph** below it. The components in the soundwave are also shown in the list along the left side of the screen in the **Component** list of check boxes labeled **0** through **11**. The **0** component is the fundamental and all other components are whole multiples of this frequency. So the **1** component has a frequency of twice the fundamental, the **2** component has a frequency of 3 times the fundamental, etc.

Press the **Play** button to start the sound. As with most sounds, the pitch that you hear is determined by the fundamental. Hum with the tone and then click on the **0** check box to remove the fundamental from the sound, but keep humming the pitch. The question is does the pitch change? Click on the **1** check box and see if the pitch changes yet. If you need to you can alternate between your current collection of components of the complex sound and the fundamental by clicking on the **Fundamental Only** check box. Clicking on the **Fundamental Only** check box so that it is selected as shown by the check in the check box will cause the fundamental to be played alone. Clicking on the checkbox again and clearing it will restore the complex tone to the condition it was before. You can restore all of the components to the complex tone by pressing the **Reset** button at the bottom of the check boxes for the component frequencies. You can also listen to any of the components alone without any of the other components by pressing the **Alone** button that is next to check box for the component you want to listen to.

In these cases, the pitch of the complex, though not the timbre, stays the same even though the fundamental which has the same pitch is not present in the complex tone. To push this comparison a little harder, select the **Fundamental Only** option and listen to the pitch, hum again if it helps. Next, select the **1** component alone by clicking on that components **Alone** button. Notice how the pitch jumps to a much higher level. It is still the same pitch because this first harmonic has twice the frequency of the fundamental which makes it one octave higher. Remind yourself how different this is from what happens in the complex tone situation. Hit the **Reset** button and then remove the fundamental and then alternate with the fundamental alone. There is no change in the octave here.

Select the **Fundamental Only** option again and now alternate it with the **2** component by clicking on its **Alone** button. In this case, even the pitches do not match, and yet when the complex had this component as the lowest frequency, the pitch of the complex did not match this component but the fundamental. So, pitch is being determined by something more subtle and complicated than frequency alone (Schouten, 1940).

One possibility might be that even though the fundamental has been removed from the signal, there is still some feature of the sound wave that varies with the same frequency of the fundamental. An illustration will help clarify this point. With the **Interactive Illustration 11.x, Missing Fundamental** still open, click on the **Reset** button to restore all of the frequencies to the stimulus. Recall from Fourier analysis, that the basic shape of the wave is determined by the fundamental. In this illustration, the fundamental is seen in the frequency of the large peaks and troughs. Click the **Fundamental Only** checkbox and see that the rate of peaks and troughs of the fundamental is the same as for the main peaks and troughs of the complex wave with all of the frequencies present. Restore all of the frequencies. Next, click on the **0** checkbox to remove the fundamental from the complex waveform. The waveform becomes more complex, but the pattern repeats itself with exactly the same frequency as the fundamental. There is some repetition that is identical to the fundamental. Click the **1** checkbox. The pattern still has a repetition that is at the same frequency as the fundamental. As you remove the components, you will still see that there is a pattern that repeats as the frequency of the fundamental. Eventually when you have removed so many components, the pattern disappears, if you press the **Play** button for one of these complex waveforms you will no longer hear that is has the same pitch as the fundamental. So, it is possible that the ear is still responding to this variation that has the same frequency as the fundamental and that it is this variation that is responsible for the missing fundamental. It is even possible that the ear is picking this variation up.

Masking and Critical Bands. The idea that the ear recreates the fundamental can be tested by directly suppressing our perception of certain frequencies and not others. To test this idea some new concepts have to be introduced and this section will discuss them before the missing fundamental is returned to. In the visual system we discussed the concept of masking where one stimulus reduces the perception of another stimulus. There is also masking in the auditory system. A tone of one frequency can mask some tones of other frequencies. This is called tone masking.

Open **Interactive Illustration 11.x, Tone Masking** to illustrate this phenomenon. This illustration will allow you to play two tones that will overlap each other. On the top of the screen is represented the **Target** tone or the tone that you are trying to hear. On the bottom of the screen is the **Tone Masker** which will be played to try to cover the **Target** tone. You can play each tone individually by pressing the **Play** buttons on the left side of the screen. The top **Play** button will play the **Target** tone and the bottom **Play** button will play the **Tone Masker**. For the first demonstration, press the **0.5** button on the left had side of the row of buttons labeled **Relative Target Frequency**. The **Tone Masker** will play first and play for ½ of a second. Shortly after the **Tone Masker** plays, the **Target** will play and last for only 1/10th of a second. However, you should have no trouble hearing the playing of the **Target**. Press the **Play** buttons for the **Tone Masker** and the **Target** so that you get a change to hear what each tone sounds like and to be sure that you know what you are listening to. You might need to adjust the intensity of the **Target** tone depending upon your sound system. The intensity chosen works for many systems, but make sure you can hear the target, but that it is not too loud. You do that by pressing the **+** and **–** buttons in the middle on the left hand side of the screen.

Now press the **2.0** button in the row of **Relative Target Frequency** buttons. Most of you will have a hard time hearing the **Target** and many of you will not hear it at all. Press the **Play** button fro the **Target** just to reassure yourself that you can hear the tone. In both cases, the **Tone Masker** frequency was 2000 Hz and the ratio of the two frequencies were the same, 2 to 1. However, in the **0.5** case, the **Target** frequency was 1000 Hz and in the **2.0** case the target frequency was 4000 Hz. This illustration demonstrates one of the curious features of tone masking. A tone will be better able to mask frequencies that are higher than itself better than frequencies that are lower than itself. In the current example, the 2000 Hz **Tone Masker** will better mask tones with frequencies greater than 2000 Hz (or **Relative Target Frequency** values greater than 1) than it will for tones with frequencies less than 4000 Hz (or **Relative Target Frequency** values less than 1). You can play around with this phenomenon to experience this difference for your self. You can adjust each value arbitrarily using the controls to the left of each tone. Some example **Tone Masker** frequencies are in **Set Mask Frequency** menu at the bottom of the screen. **Relative Target Frequency** buttons in the middle of the screen will select **Target** frequencies relative to the current **Tone Masker** frequency however set.

This asymmetry in tone masking implies a lot about how masking in the auditory system can arise. Examine Figure 11.x, both parts. This figure illustrates what is going on the basilar membrane during tone masking. In Figure 11.x (a) the frequency of the masker is lower than the target tone. The curves in the figure represent the amplitude or maximal motion of the basilar membrane as result of each of the two frequencies used during tone masking. Review **Interactive Illustration 10.x, Traveling Waves and Basilar Membrane**. Figure 11.x (a) shows what happens when the target tone has a lower frequency than the masking tone (e.g, like the **0.5** example above). As in the examples you experienced, the intensity of the masker is much greater than that for the target tone. However, since the target has a lower frequency than the masker, it travels farther down the basilar membrane than the masker so that is has a region of the basilar membrane that it stimulates free from the masker. Compare this circumstance to Figure 11.x (b) where the masker now has the lower frequency. Now the motion of the traveling wave associated with the

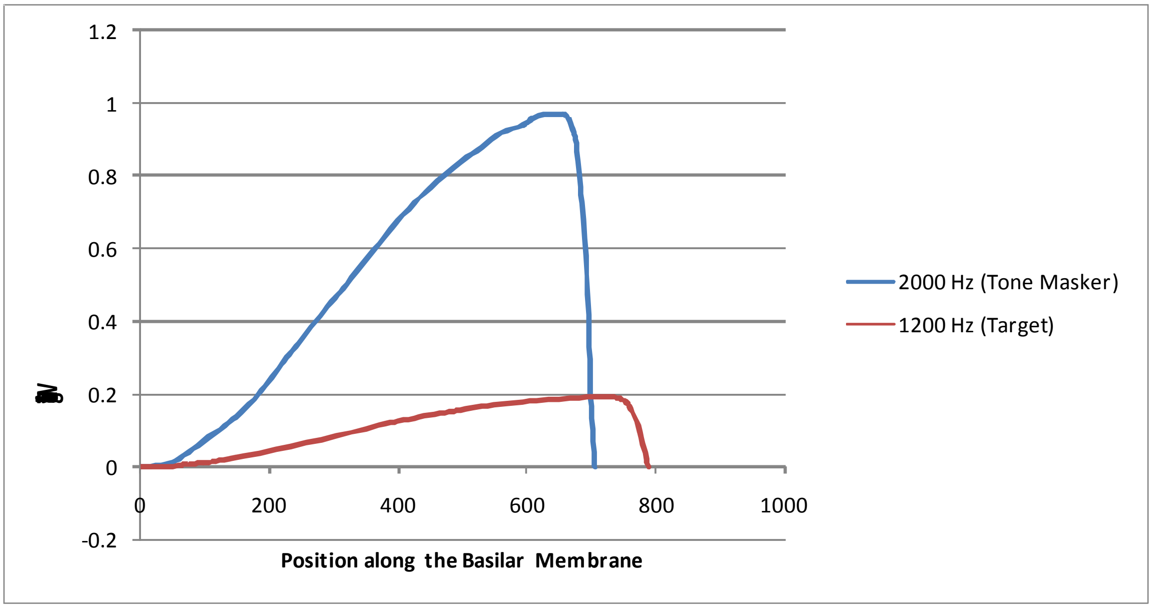

Figure 11.x (a). Illustration of the movement of the traveling waves in tone masking when the Tone Masker is a higher frequency than the Target.

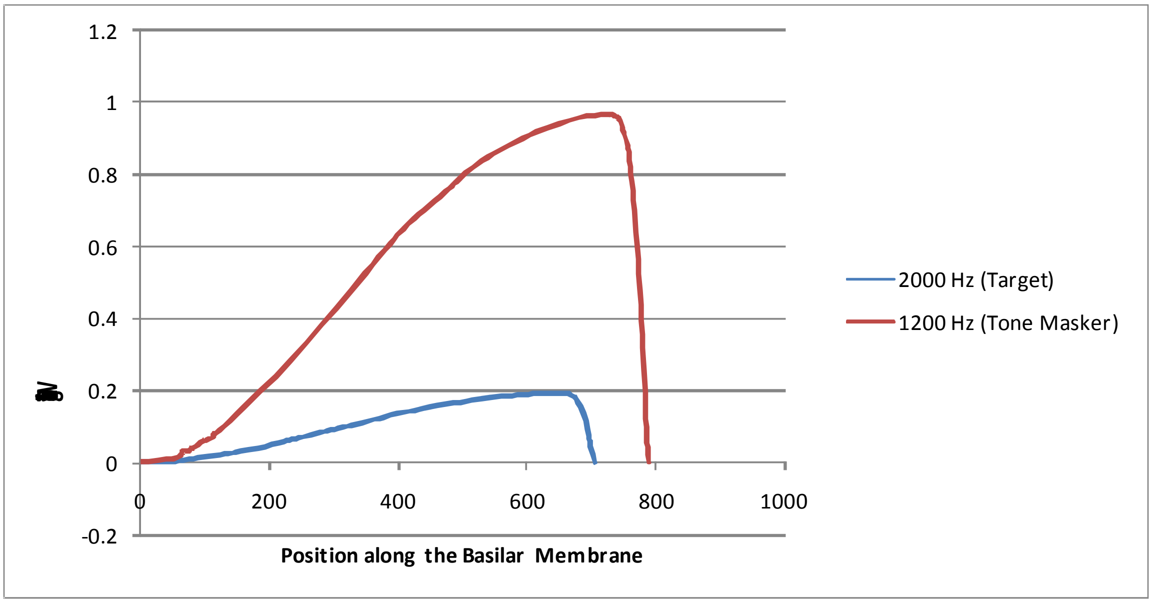

Figure 11.x (b). Illustration of the movement of the traveling waves in tone masking when the Tone Masker is a lower frequency than the Target.

target is completely swamped by the motion associated with the masking tone.

To see this phenomenon in action, open **Interactive Illustration 11.x, The Ear and Tone Masking**. This illustration resembles **Interactive Illustration 10.x, The Ear**, except that you will not be able to see the middle ear by itself. In addition you will be able to stimulate the ear, either when shown completely or when just the inner ear is shown, with two frequencies simultaneously. When the illustration is open, click on the **Freq. 2 High** button and then press the **Start Sound** button. If you have not changed any of the settings, Frequency 1 will be **2500** Hz and Frequency 2 will be **5000** Hz and the amplitude of Frequency 1 will be **0.15** and the amplitude of Frequency 2 will be **1.0**. These settings resemble what is going on in your ear when the masking tone is a higher frequency than the target tone (e.g., Figure 11.x (a)). As you look at the motion of the basilar membrane you can see the large motions caused by the masking tone but you can also see a much smaller motion caused by the target tone farther beyond the end of the motion caused by the masker. To test this observation you can adjust the amplitude of Frequency 1 to **0** and see this motion disappear. Simple press the **Freq. 2 High** button to restore the motion of target tone. Next, press the **Freq. 1 High** button. The amplitude of the **2500** Hz tone will now be **1.0** and the amplitude of the **5000** Hz tone will be **0.15**. This situation resembles what is going on in your ear when the masker has a lower frequency than the target tone (e.g., Figure 11.x (b)). If you look at the motion of the basilar membrane it is hard to tell that there is the motion of the **5000** Hz target tone going on. You can adjust the amplitude of Frequency 2 up or Frequency 1 down to see that this motion is present.

These observations lead to several important conclusions a couple of which will be discussed here. First and foremost, it seems that a great of what happens in masking occurs because of the way the basilar membrane works. One way that masking has been described is by the busy line hypothesis (Patterson & Green, 1978; Zwislocki, 1978). In this hypothesis, masking occurs when the area of the basilar membrane that responds to a tone is already to busy responding to the masker. If you look back at Figure 11.x (b), you will see that the motion of the basilar membrane caused by the masker is so much greater than what is caused by the target tone, even at the location of the peak activity caused by the target tone. In essence, the hair cells in this location of the basilar membrane are already to busy responding to the masking tone to be able to respond to the target tone. The line is busy.

Second, there should be a limit to masking. Look back as Figure 11.x (b) Notice the shape of the traveling wave of the tone masker. The amount of motion of the basilar membrane grows until it reaches the are of peak motion at there area where the resonance of the basilar membrane matches that of the tone. So, as the frequency of the target tone gets higher, the line gets less busy, to use the terminology developed above. Let us try a couple of tests of this idea. First, reopen **Interactive Illustration 11.x, The Ear and Tone Masking**. Using the buttons on the left side of the screen set Frequency 1 to **2000** Hz and Frequency 2 to **4000** and then press the **Freq. 1 High** button to make Frequency 1 the masker. Now press the **Start Sound** button. Only the motion of the masker should be visible in the motion of the basilar membrane. Use the **Frequency 2** slider at the bottom right of the screen and slowly adjust the frequency of Frequency 2, our target, upwards. Eventually you will see a small wrinkle emerge in the motion of the basilar membrane that is caused by the motion of Frequency 2 (the target). So by this frquency, the target should be heard even though we have not increase its intensity, and it is even possible that we are less sensitive to this frequency than to lower frequencies where the tone is masked. Ok, we have a theoretical prediction, not lest us test it. Reopen **Interactive Illustration 11.x, Tone Masking**. Press the **2.0** button to set the frequency of the **Target** so that it is twice that of the **Tone Masker**. In fact, you should have the frequency of the **Tone Masker** of 2000 Hz and the **Target** of 4000 Hz. You might need to lower the intensity of the **Target** tone, with the **–** button, to make sure the **Target** is masked. To test our theory, use the **Target**'s **Frequency** slider in the upper left corner and adjust its frequency to 8000 Hz. Then press the **Play Both** button at the bottom of the screen to play both the **Tone Masker** and the **Target** tone. See if you can hear the **Target** in this case. I am betting you can. It will be a very high pitched little squeak, but it should be audible. So, important elements of masking arise at the basilar membrane and the relative location of the masking stimulus to the tone stimulus seem to be important. These lessons will be important to use as we discuss the next type of masking.

The type of masking that is of more interest in the present investigation is where a noise stimulus is used to mask a tone. Open **Experiment 11.x, Masking and Critical Bands [link to media]**. This experiment is actually three types of activities built around the same phenomenon, **noise masking [to glossary]**. The first screen is the stimulus set up and you can select the type of activity using the **Type of Activity** menu at the top of the screen. To introduce the phenomenon of noise masking, chose the **Demonstration** option from the **Type of Activity** menu. The screen has many different items on it but for now there are only a few that need to be examined. First, on the lower right side of the screen is a plot of all of the frequencies for a noise sound. The frequency of the waves in the sound is on the x-axis and the amplitude of that frequency is on the y-axis. Noise, or more properly, white noise, is a random collection of all of the possible frequencies of sound, changing randomly over time. If you press the noise's **Play** button which is the bottom of the two play buttons on the lower left side of the screen, you will hear what sounds like static. Static is a type of white noise. While the noise is playing you can press the **Retrigger** button below the noise's spectrum graph or check the **Autotrigger** checkbox to have the graph update itself periodically.. You will see that while all frequencies are present, they change in their amplitude. This change is random and averages out over time. Press the noise's **Stop** button so you don't get too irritated by the sound. It is possible to limit the range of frequencies that are present in the noise stimulus. Click on the **Turn On Filter** checkbox in the upper left corner of the screen. Now the range of frequencies present in the stimulus is limited to the value indicated at the bottom of the **Noise Band** slider on the far left side of the screen. When you first start this activity, this range is set to a bandwidth of **4000** Hz. That means that the lowest frequency present in the noise stimulus is 4000 Hz less than the highest frequency present. The cutoff is not absolutely sharp, but this pretty well defines the range. You can adjust the range of the frequencies using the **Noise Band** slider. As you move the slider down, the width of frequency on the **Noise Masker** spectrum graph gets narrower and the character of the sound changes. It also gets less loud because there are fewer frequencies in the stimulus. This band of noise is centered on one frequency that is the frequency that will have the greatest amplitude. That center frequency for the noise stimulus can be controlled by the **Noise Center Freq** slider next to the **Noise Band** slider.

While the noise sound is playing, you can also play a second sound, the **Stimulus** which is indicated in the **Tone Stimulus** graph at the top of the screen. Start the **Stimulus** sound and while you are in this demonstration mode, adjust the intensity of the stimulus using the **Stimulus Intensity** slider until you can barely hear the tone. When the stimulus has been set at this level, try adjusting the stimulus frequency (the **Stimulus Freq** slider) and/or the size of the **Noise Band** to see what effect these variables have on your ability to hear this stimulus. When you think you have an idea about the effect of noise bandwidth and stimulus frequency on the ability of the noise to mask the stimulus, close this window and the original setup window will still be open.

Now select **Demo Experiment** from the **Type of Activity** menu and press the **Done** button at the bottom of the screen. This demonstration experiment will allow you to explore the parameters of noise masker. It runs very similar to the Frequency Response of the Ear experiment from Chapter 10. There will be two sounds, the noise masker which will come on first and then there will be 10 stimuli each of the same frequency but with each step increasingly less intense. The tone stimulus will be shown in the top graph on the screen and the noise masker will be shown on the middle graph. There are two settings you need to make for each condition, the **Noise Width** and the **Center Frequency** of the noise. These values can be selected using the menu at the top left corner of the screen. (The center frequency value does not change the noise when it is Full Noise, but it does for all other conditions.) For the first run, select **Full Noise** for the **Noise Width** and **1200** Hz for the **Center Frequency** and then press the **Start Series** button. Count how many of the tones you can hear.

When the series is done, a menu will appear to allow you to enter the number of **Steps Heard**. Press the **Data Entered** button to indicate you have made the correct selection the menu and the bottom graph which shows your results will be updated. To complete this experiment, you need to do at least one series for each of the **Noise Width**s for each of the **Center Frequencies**. Run this experiment and then return to the text.

At the end of the demonstration experiment, you might have data that looks somewhat like Figure 11.x. When no noise is used as a masker and then the full noise is used, the center frequency of the noise does not alter the nature of the noise, so the results should be nearly identical for the two "**Center Frequency**" conditions. When there is not any noise, there can be no center frequency, and when all wavelengths in the noise are possibly used, there is not center frequency, either. So the results the number of steps of the tone staircase that is heard should be the same in the two cases. However, for all of the other noise bandwidths, the center frequency has a huge impact on the number of tone loudness steps heard. The tone is 1200 Hz. When the center of the noise band is the same as the tone frequency (the **Noise = Stim** condition) you will hear far fewer noise steps than when the center of the noise band is at 3600 (the **Noise != Stim** condition).

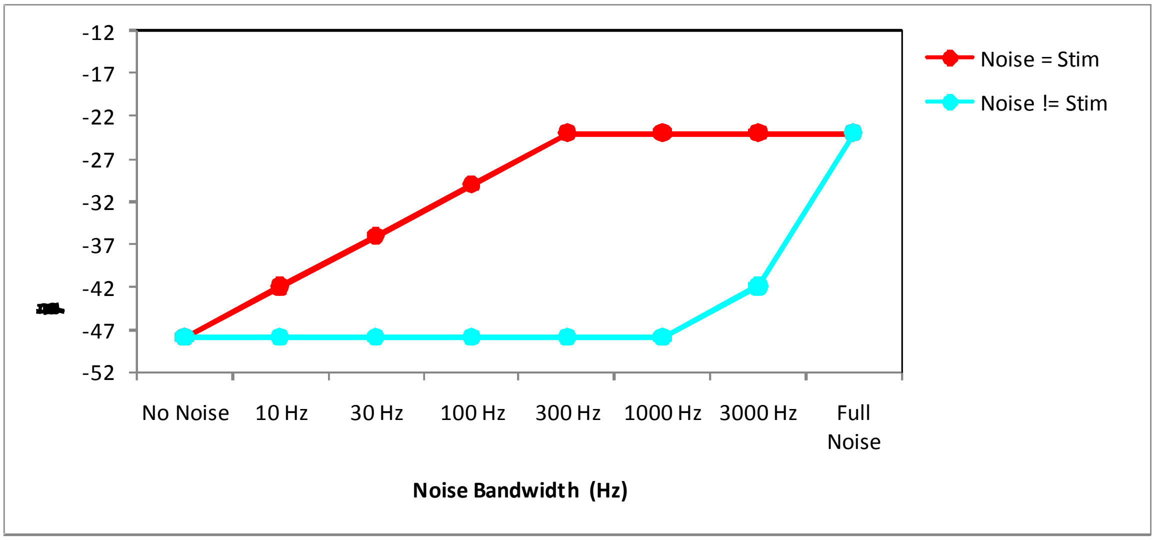

Figure 11.x. Example output from the Masking and Critical Bands demonstration experiment.

Delving into this graph several conclusions can be made. First, and most simply, noise can effectively mask a tone. Not to surprising but the next conclusions are both more surprising and more interesting. First, there seems to be a need for the noise to have the same or similar frequencies as the tone to be an effectively masker. The **Noise = Stim** condition led to much more masking than the **Noise != Stim** condition. Looking more carefully at the figure generated from your results or Figure 11.x we can get some estimate as to the size of the area around the tone that the noise needs to be in to have effective masking. If you look only at the data from the **Noise = Stim** condition, examining the trend starting with the **No Noise** masking condition plotted on the left side of the graph near the origin. Here you are most sensitive to the tone because there is no possible masking other than background noises in your room. Then as you add more noise by adding a wider range of frequencies to the noise, all centered on the tone frequency, you need a more intense tone to be able to hear it. You have more and more masking. But only up to a point. Eventually, the data from the **Noise = Stim** condition should level out. In more careful experiments than we can do here, the date will level out precisely. The width where the graph seems to level out about the width of 300 Hz in this experiment, based on running on students in my classes. Adding more frequencies beyond this bandwidth, so that you are adding frequencies farther and farther from the tone frequency, does not lead to any better masking.

To understand the impact of this number try a very quick experiment. If you have closed **Experiment 11.x, Masking and Critical Bands** or still have the data window open, go back to the main experiment window and run this experiment as a **Demonstration**. Set the **Noise Center Freq** to 1200 and the **Noise Band** to 300. Then press the **Play** button for the **Noise**. Then turn on and off the noise filter using the **Turn On Filter** checkbox at the upper left corner of the screen. When the **Turn On Filter** checkbox is checked you will have the noise limited to mostly +/- 150 Hz about the 1200 frequency of the tone. When the **Turn On Filter** checkbox is not checked, then the full range of frequencies will be played in the noise. Click the **Autotrigger** checkbox under the sound spectrum so you can see the changes in the noise as you check and uncheck the box. What you need to listen for is the intensity of the sound. When the noise has the full range of frequencies, it is much more intense than when it is just the narrow band of the 300 Hz. Yes, still the full range of frequencies is not really any better of a masker than the 300 Hz tone. Adding frequencies of sound to far from the tone makes the noise louder but it does not add to the ability of the noise to mask. This window of frequencies around the tone where noise frequencies can mask a tone is call the **critical band [glossary]**. In one sense, this region of frequencies defines a region of spatial summation along the basilar membrane, recalling that where different frequencies maximally stimulate the basilar membrane is stretched out in order along it (go back to the last chapter if you need a review). So this noise band that is masking the tone will be stimulating a very small region of the basilar membrane centered on the where the tone frequency is stimulating the basilar membrane (Scharf, 1970; Zwicker, Flottorp, & Stevens, 1957).

Before returning to the question of where the missing fundamental arises, let us consider some of the important implications, and applications, of the finding of critical bands. One of the important uses of the ideas of critical bands comes in the designing of alerting sounds. Alarms are sounds that need to be heard. If there are other sounds around, the might be able to mask the alerting sound and that make prevent people from responding the an important emergency. I will take one example to make this situation concrete. I will admit that I have a personal experience with this example, not because I had any role in designing them, but that I am old enough to have experience both these sounds designed poorly and designed well. The example I want to talk about is seatbelt alarm in cars. These alarms come on after the car has been started meaning that there can be a fair amount of engine noise that could mask the alarm. Now, it is possible to take two different approaches to making sure the seatbelt alarm is heard. One way is to make the alarm real loud so that it cannot be masked by the background sounds. The first seatbelt alarms in the late 1970's and early 1980's were of this design. Unpleasant to say the least. Drivers were not very happy with this type of car alarm and lots of complaints ensued and drivers tried to find ways to disconnect the alarms. These alarms aren ot much good if they are disconnected. This car alarm design did not take into account the way our auditory system works. The designers ignored or, more probably, did not know about critical bands. So, the second way to design an alarm is to take into account critical mands and make the alarm have a frequency outside the frequencies of the background noise so that no matter how loud the background noise, it will always be heard. Car engine noises and most background noises have low frequencies. So using a tone with a high frequency will not be masked by low frequency noise no matter how loud. That is why car alarms, not just seatbelt alarms, are these high pitched bell-like tones. They are not very loud, but they will not be masked by the noises the car, road, sound system, or occupants make even though they are not very intense. These tones are also a lot more pleasant than the original seat belt alarms for cars. Similar thinking also explains why sirens tend to be high pitched as well.

Critical Bands and the Missing Fundamental. Now it is time to return to the question of if the missing fundamental can be recreated on the basilar membrane. The concept of critical bands will be very helpful here. The logic of our investigation will be this: if the missing fundamental is recreated on the basilar membrane then it can be masked by noise if the noise is in the same critical band as the fundamental. To do this investigation, we need to develop a masking noise that will mask the fundamental but not the overtones that we are interested. In addition, we need to construct a set of overtones with the same pitch as the fundamental we wish to test.

Open **Experiment 11.x, Missing Fundamental Experiment**. This experiment is an informal experiment where you can manipulate the nature of the tones being played and the noise being used to mask and see what happens to what you hear. When you open the experiment, you will see notes written on a musical scale across the top of the screen. When you press the **Play Tune** button in the middle of the screen you will hear, after a few seconds delay, the famous *Ode to Joy* tune by Ludwig van Beethoven. As each note is played, the note will be highlighted on the scale and drawn in orange. The use of the tune will help you hear or not the presence of the missing fundamental by helping you know what note you should be hearing. More about the rule of melody in organizing perception will be discussed below. You can play this tune several ways by selecting one of the checkboxes below the **Play Tune** button. When the **Complete Tone** checkbox is selected, the tune is played using tones made up of a fundamental frequency of the note indicated on the scale and several overtones in some way approximating natural tone, though admittedly the approximation is not great. Selecting the **Fundamental Only** checkbox will cause the tune to be played with only the fundamental frequency being sounded. Using the **Missing Fundamental** checkbox, will lead the tune to be played with a complex of overtones, a fair distance from the fundamental, that has the same pitch as the fundamental. Play the tune a couple of times selecting the **Fundamental Only** and **Missing Fundamental** to see that the tune, though it sounds different, has the same pitches in each case. Hum along with the tune to see that it works. The final checkbox, **Alternate Fund/Miss Fund** will play the tune alternating notes between the fundamental and missing fundamental. In fact, it takes each note and divides in half, playing the first half with the fundamental of the pitch and the second half with the overtones of the same pitch. Play that version and hum along and you can see how the tune is not disrupted by this alternation between the fundamental and overtones for each note. You will also see on the screen that the melody is drawn differently where each note is now divided in two so you will be able to follow what is being played when. It is this version of the tune that is most important for our experiment.

The bottom half of the screen will allow you to play noise to act as our masker. This control of noise looks very similar to what was seen in the demonstration setting of **Experiment 11.x: Masking and Critical Bands**. You can manually control the noise intensity with the **Noise Gain** slider on the left side of the screen, the width of the noise with the **Noise Band** slide next to the right, and the center frequency of the noise with the **Noise Center Freq** slider next to that slider. Remember to use the filter you need to check on the **Turn On Filter** checkbox above the sliders. To see the result of your actions on the noise click on the **Autotrigger** checkbox in the lower right hand corner of the screen.

While I certainly encourage you to play around with the noise yourself so that you can get the best feel of the phenomenon, there are two preset settings for the noise and tune intensity that are available. At the bottom of the screen, you can press the **Mask Fundamental** button to set up the experiment to have a noise to mask mostly the fundamental frequencies of the tone, and you can press the **Mask Overtones** to set up the experiment to mask mainly the overtones of the tune.

Start the experiment selecting the **Fundamental Only** checkbox and pressing the **Mask Fundamental** button. Then press the **Play** button for the noise and then press the **Play Tune** button. Follow the notes on the scale at the top of the screen and you can turn on and off the noise as you wish to ensure that the tune is playing. The fundamental is being effectively masked. You might hear the odd note, but that is probably mostly memory and seeing it played on the screen as will be demonstrated in a bit. No select the **Missing Fundamental** checkbox and play the tune again turning on and off the noise as before. The masker has little if any impact on your earing of the tune. In fact, the tune sounds just the same without the masker as with the masker. Before we tie all this together, select the **Alternate Fund/Miss Fund** checkbox, turn on the noise and play the tune. Now, you will only hear the half of the note played by the overtones. The pitches will be fine but the rhythm will be a bit off in a couple of places where Beethoven used a bit of syncopation to make the rhythm a bit more interesting.

So, masking the fundamental does not alter in any way the perception of the pitch caused by a set of overtones. That is the take home message from this experiment. Even when the region of the basilar membrane where the fundamental is stimulating it is being disturbed in a way that it is not very sensitive to this frequency, your perception of the missing fundamental is not disturbed. As a result, it seems very reasonable to conclude that the missing fundamental must not be the result of the fundamental being recreated by processing in the ear, either outer, middle or inner, or the masking would work. To push this observation a little further, it suggests that pitch perception, even though it is very tied to the frequency of the sound, is some how built up in the brain, in the auditory cortex. Pitch, which seems very sensory, simple, derived straight from the stimulus, now seems to have elements of complex processing derived from our cortical structures (Licklider, 1955; Ritsma & Cardozo, 1963/64).

Application of the Missing Fundamental. It might seem odd that there would be an application for an illusion. But the missing fundamental is very important to your pocketbook. Consider buying speakers for your car or home sound system. The speakers are usually rather expensive, particularly if you want good ones. The reason for the cost is to increase sound fidelity. Better speakers produce a wider range of frequencies. Also, because of resonance, no speaker can produce the entire range of frequencies so you have some speakers for the high frequencies, like tweeters, and some for the low frequencies, like woofers. Now, think of your telephone, whether it be a land line or cell phone. It needs a speaker to reproduce sounds. But the speaker is small and from our discussions of resonance in Chapter 10, it is clear that such speakers will have a relatively high frequency resonance. Also, the quality of the speaker is not terribly expensive. You are not paying a lot of money for that speaker and the sound you get over your phone is not the greatest. In fact, they are so limited that they do not reproduce the fundamentals of the speaking voices you hear on the other end of the line. Yet, I doubt you have any trouble distinguishing male from female speakers you have no trouble hearing the typical fact that most males have lower voices, by about an octave, of most female voices. You hear the proper pitch of the speaker through the missing fundamental. The same phenomenon that has been discussed so much plays a role in every telephone speaker you have ever used. So, one way the missing fundamental helps your pocketbook is that is allows you to not have to by a first class speaker for your phones, keeping the price of the phones down. However, that is the small part of the savings. Since telephones do not reproduce the fundamental frequencies of voices, there is no need to transmit those frequencies over telephone lines or via cell phone towers and satellites. Thus, a lot less information needs to be transmitted over our phone systems making it possible for these phone systems to carry a lot more phone calls on a single phone line to reduce the bandwidth needed for the cell phone communications. That makes the phone system a lot less expensive.

Music

There are many features about music that could be discussed. Music is a complex psychological response to auditory stimuli with specific characteristics. However, the definition of the characteristics of sound that are categorized as music is a fluid definition. Generally, the stimulation must have rhythm, a pitch pattern called melody, and often harmony. However, experiences such as *Stomp* indicate that even rhythm by itself can be considered musical. Given the complexity of music there are several issues that could be discussed. To keep this discussion with the bound appropriate to this text, we will focus on some issues that help us to hear melod.

Background

Before beginning this discussion, it seems appropriate to cover some basics of music and musical scales, at least as they are found in the western traditions of music, to help make the discussions below and even later in some illusions clear. If you have a basic knowledge of music and musical scales, you can skip this section without any loss of understanding later sections. Open **Interactive Illustration 11.x, The Musical Scale**. When you open the screen, you will see a drawing that looks something like a piano, and it is meant to resemble this instrument. As probably most of you know, the musical scale based on a repeating set of notes. The repetition happens around the interval of the octave. Every octave is a doubling of frequency. So, on the piano keys in the interactive illustration, if you press the key on the extreme left, you will play the pitch known as middle C. This frequency is right about 261.63 Hz. If you press the key on the extreme right, you will sound the C that is one octave higher than middle C. This frequency is about 523.25 Hz. While this C sounds higher in pitch than middle C, it also sounds the same note. The next C higher will again double in frequency being about 1046.50 Hz. The octave is the most fundamental musical interval as this interval determines when notes repeat themselves.

The octave is divided into 12 what are called semi-tones or half steps from one note to the octave above it. If you count the keys on the screen, you will see 13 keys, both the white and the black, but the two outer keys are both C so there are only 12 steps on from one C to the next. This makes up what is called the **chromatic scale [to glossary]**. The normal musical scales use only 8 of those keys. There are two main types of scales, major and minor. C Major is the easiest. It makes of the white keys on the piano starting with C and going to the next C. If you press the white keys in order from left to right you will hear the musical scale that you have heard before. The names of the these notes are C, D, E, F, G A, B, and C again. The minor scale differs only in one note. The C minor scale is identical to the C major scale except in the third note. To play the C minor scale, start with C, play the next white note (D), and then the black key just above (E flat). Skip the next white key and pick up with the fourth white key (F) and finish by playing the rest of the white keys.

These basic concepts should help the discuss below.

Melody

Melody is one the basic dimensions of music. But melody is really only a specific example of our general ability to follow lines of music. Imagine listening to an orchestra. Many different instruments playing different patterns of notes, but generally, we have little problem following along with the melody or other lines that we wish to follow. Now, much modern music has a principle line called the melody and most of us follow that line principally. However, in older music there might be several intertwined lines that may have more or less equal roles in the music, this pattern of music is called polyphony and a fugue is a particular example. Even in this case, we can usually follow the lines of music. This ability to follow a given lines of music seems well assisted by similar principles of organization, the Gestalt Laws, that were discussed in Chapter 5 (Deutsch, 1982; Sloboda, 1985).

Certainly organizing perception is not limited vision. There is a need to take auditory information and make sense of it so it should not be surprising to see that some of these principles apply in audition as well. Let us review a couple and see how they might apply to our ability to follow melodies and other lines in music.

To review, open **Interactive Illustration 11.x, Gestalt Laws of Proximity and Similarity**. This is a repetition of the figure from Chapter 5. In this first example, we will be most concerned about the role of Law of Similarity. Take the **Shape** slider and drag it to the bottom. There will now be alternating columns of squares and circles. Most of you will see this pattern as alternating columns. The squares in a given column are grouped together and the circles in a given column are grouped together. Objects that look similar are grouped together. Similarity can also help up follow musical lines. In many cases, a given musical line, like the melody, are played by a single instrument (or instrument group) or voice(s). For example in a choir, the sopranos usually carry the melody and the fact that these same voices carry the melody throughout help the listener follow the melody.

Other Gestalt laws also play a role in our ability to follow melodies and musical lines. Open **Interactive Illustration 11.x, Gestalt Law of Good Continuation**. This is a repeat of the illustration from Chapter 5 about the Law of Good Continuation. You probably see a curved X figure centered on the screen. Now grab the **No Continuation** slider on the right hand side of the screen and drag it up. The lines separate but instead of the smooth crossing of the lines they bend abruptly in the middle. Rejoin the figure now. It is still possible for the figure to be created in the way you just saw it, but when joined your visual system has strong preference for the perception of two smooth lines crossing each other. A similar grouping of sounds phenomenon happens in our listening to music.

Open **Interactive Illustration 11.x, The Scale Illusion**. At the bottom of the screen press the **L Channel** check box and then press **Start Tune**. The note being played will be highlighted on the scale as it is played. The tones should bounce up and down getting closer together till a note is repeated, then the bouncing gets farther apart. Next press the **R Channel** check box followed by the **Start Tune** button and hear the notes played to the right ear. The pattern is very similar to what is being presented to the right ear. Now press the **Both Channels** check box and press the **Start Tune** button to listen to both channels simultaneously. Headphones work best but this illustration also works with stereo speakers.

What most people hear is rather different from the simple expected combination of the two patterns of notes that have been presented separately to each ear. The pattern of notes played in each ear is show in the interactive figure but they are repeated below with the two patterns played on the same score below in Figure 11.x. The notes played in the left ear are blue and the notes played in the right ear are red. Notice the jumping back and forth of the notes played in the two ears. But also notice, that if you ignore the color, there is both a descending and an ascending scale being played. The descending scale is starts in the right ear and the ascending scale starts in the left ear, but each scale bounces back and forth across to each ear.

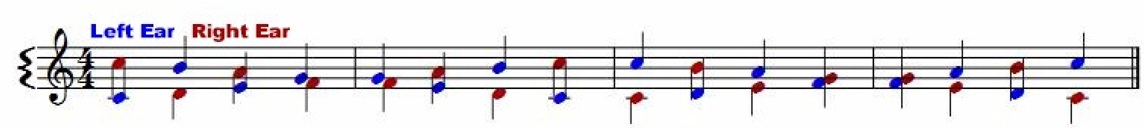

Figure 11.x. The notes played in the scale illusion. The blue notes are played in the left ear and the red notes are played in the right ear.

If the Gestalt Law of Good Continuation plays a role here, one might expect that the brain might group these two scales together so that one might be heard in one ear and the other scale heard in the other ear. In a sense, this pattern is heard but with a slight change. Below in Figure 11.x, one common outcome of the scale illusion is heard. One ear will hear an ascending scale up to the fourth note of the scale. This note will repeat and then the scale will descend back to the starting note. The other ear will hear a descending scale going down 4 notes with that note repeating and then the scale will ascend back to the starting note of the scale. Both of these patterns are then repeated so it sounds like there are two repetitions of the pattern. Good Continuation groups the sounds bouncing back and forth between each ear into organized scales. If you look back at the interactive figure, you might see why the scales seem to turn around where they do even though the actual scale is not completed (Deutsch, 1982). At the fourth and fifth note in each ear, the tone is repeated. The Gestalt Law of Proximity probably plays a role here as two identical notes played in the same ear seem to make sense to group together.

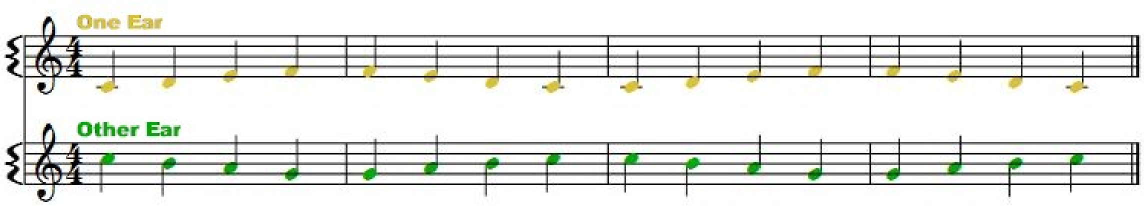

Figure 11.x. One commonly heard outcome of the scale illusion.

To test this idea, let us try a little experiment. Click on the **– Ascend Octave** button. In the original Scale Illusion the ascending and descending scales were in the same octave. Now the ascending scale which starts in the left ear is one octave lower and the two scales will only reach the same note when the descending scale has reached its bottom and the ascending scale has reached it top (both Middle C). Now press the **Start Tune** button. Now most of you should hear an ascending scale in one ear and a descending scale in the other ear. There is no intermediate pitch where the law of proximity can reverse the scale before it is complete. There is no apparent repetition in what you hear either.

To show that brain can, when it does not need to group, quite contentedly hear the scale bouncing between hears, press the **Scale Alt Ear** check box. Now only the ascending scale of the scale illusion will be played, as before, starting in the left ear. Press the **Start Tune** button and listen. You should be able to hear the scale, but it should bounce back and forth between ears. The grouping of the scale illusion seems to be used when necessary to make sense of the auditory input.

From these few examples, we can see that Gestalt-like grouping principles operate in our auditory system as well as in our visual system. From this observation, we can conclude that sensory system in general operate to make sense of the input. To the extent that Gestalt theory is true, the organization of sensory inputs works to see the simplest perception of the incoming stimulus pattern.

Auditory Space Perception

Where does a sound come from? Someone drops their books and you immediately swing your head around and see the poor person looking down at what had recently been a neat stack of books. You are at a party talking to a friend. Someone calls your name and you can look around to directly find the person, all of all the chatter around you who called your name (assuming they meant you instead of someone else who shares your name). We make mistakes in these tasks but auditory information can be very successful in helping us determine the direction of a sound.

As in the perception of depth in vision, we shall approach auditory space perception by discussing the cues, or sources of information, that assist us in determining the perception of the direction and to some extent, the distance of a sound. Also, as in the discussion of depth perception, we can divide these depth cues into different types, in this case, **monaural [to glossary]** and **binaural [to glossary]**. The prefixes mon- and bi- mean one and two respectively just as in monocular and binocular. Aural in this case refers to the ear. So monaural cues require the stimulus to only reach on ear and binaural cues require two functioning ears to work. Each type of cue will be discussed in turn.

Monaural Cues

Two monaural cues will be discussed: relative loudness and the Doppler Shift.

Binaural Cues

The binaural cues require both ears. This statement implies that the stimulus is somehow different to the two ears. At first blush, that statement does not seem to excite much curiosity. In vision, an object that is not on the horopter (see Chapter 8) falls on different parts of the retina. That is the nature of the focusing of the lens and cornea of the image on the retina. Consider the ear. If you need refreshing, reopen **Interactive Illustration 10.x, The Ear**. All sounds that enter the ear follow the same path, in through the external auditory canal, through the ossicles and onto the basilar membrane. Notices than in this illustration there is no control for where the stimulus arises out in the world. The only way to influence there the stimulus falls on the basilar membrane is by adjusting the frequency of the stimulus using the slider at the bottom of the screen. At the single ear, all information about the direction of the sound in the world is lost.

Now, open **Interactive Illustration 11.x, Head in Sound Space** to explore some of the ways that the stimulus could be different between the two ears. The left two thirds of the screen is taken up with a drawn head facing the top of the screen (imagine you are above the person looking down). The sound source is a red dot. Now, in a normal sound source, if not blocked by some object, the sound will go equally in all direction from the sound source. However, for our purposes, we are only concerned with the sound reaching our listener's two ears. These are shown as red lines going from the sound source to each ear. A long each line is a representation of the sound wave traveling to each ear. The sound wave is cyan going to the right ear and green going to the left ear. The same color coding is used in the graphs on the right side of the screen which will be discussed in a minute.

You can move the stimulus by clicking and dragging your mouse over the screen area around the head. The sound source will follow your mouse. So first, click and drag the sound source so that it is straight in front of the listener's head. At the bottom of the screen is some information about the sound travel from the sound source to the head. First, the distance from the sound source to each ear is given. If you have the sound source straight in front of the head, these two numbers will be the same for the two ears. See if you can get those two numbers the same. After the distances, the phase of the sound at the ear is given. Recall that the phase is the position in the sound wave and goes from 0 deg to 360 deg and repeats with each cycle of the wave sound (see **Interactive Illustration 10.x, Sound Basics**). Now in a real sound stimulus, the phase would be constantly changing so this illustration is grabbing the stimulus at one instant. However, if the distance from the sound source to each ear is the same, so will be the phase at each ear. The sounds are in sync. The graphs on the right hand side of the screen replot this information as differences. If the distance from the sound source to the right ear is greater, then a cyan bar will appear over the **Right Ear** labels in the graphs indicating the size of the difference, distance in the top and phase in the lower. If the distance from the sound source to the left ear is greater then green bars appear over the **Left Ear** labels in the graphs.

Now click on the sound source that is straight in front of the listener and slowly drag it in an arc around so that the sound is now directly opposite to the left ear. Notice how the distance from the sound source to the left ear is not smaller than to the right ear, and if you watched the illustration carefully you will notice how the difference in the distance between the two ears increased systematically as the sound source moved from in front of the listener to the left of the listener. There is a systematic relationship between the angle the sound source makes with the head and how much further it has to travel to the farther ear. Fortunately, sound does not travel nearly as fast light and the differences in time it takes the sound to arrive at the two ears is actually perceptible by the brain allowing us to use this difference to help determine the direction of sounds. This binaural cue is called **interaural time of arrival [to glossary]** or **interaural phase difference [to glossary]**. These two binaural cues are nearly the same and will be discussed together.

While you are looking at this figure, notice the faint cyan arches bouncing a way from the left side of the head. Your head is a solid object and solid object will reflect some of the sound that falls on them. So the sound will be a little bit fainter to the far ear. How much fainter will also depend upon how direct the sound is to the two ears. When the sound is straight in front of you, the two sounds will have the same intensity to the two ears. When the sound is to the side there will be the biggest difference in the sound intensities between the two ears. This intensity difference is called the **interaural loudness difference [to glossary]**.

Interaural Time of Arrival/Interaural Phase Difference. Time of arrival refers to when the sound first reaches the ear. Phase refers to the difference in the phase as the sound continues at each ear. So technically time of arrival is mostly when the sound is of a very short duration, say a click. The phase difference will be used for tones that continue a while. Recall how auditory transduction occurs. The cilia are moving back and forth and in one direction, they are pulled apart and the tip links are pulled taut opening ion channels which is when the transduction happens. As a result, recall that the firing of the neuron happens at one particular phase of the tone. What this means is that differences in phase at the two ears means that neurons carrying the information to the brain will also be out of sync with each other firing at different times. It seems that is this difference in firing rate that is picked up by the brain.

To experience how this binaural cue works, open **Interactive Illustration 11.x, Time of Arrival and Phase**. To best use this illustration, you will need some type of headphones. If you do not have headphones try to place your speakers as far apart as possible so that you are as close to between the two speakers as you can be. The center of the screen shows a diagram of the waveform that will be presented.

Initially, a click will be shown. Since this auditory space cue is binaural, separate signals will be generated for each ear. In essence, this example is in stereo. The signal to the left ear is drawn in green and the signal for the right ear is drawn in cyan. Press the **Play** button to see the initial stimulus, a click, that is used in this illustration. These conventions for drawing the separate signals for the left and right ear will be true for all of the demonstrations that use stereo. The signal for the right ear will be drawn slightly lower on the graph so that they will never completely overlap.

Now let us further explore interaural time of arrive and phase as a cue for the direction of a sound. We will start with clicks where the binaural cue is time of arrival. When you first pressed the **Play** button, the clicks to each ear, 0.5 msec in duration, are played simultaneously. If you have headphones, the sounds should seem to come from the middle of your head. If you are using speakers, the sound should appear to come from halfway between the two speakers. To simplify the rest of the discussion, I will assume that you are wearing headphones, but the results should be similar for a pair of speakers.

Now move the **Time of Arrival** slider at the bottom of the screen slightly to the right so that the value displayed at the end of the slider reads **00.50 msec**. When these values are positive, the signal to the right ear will be presented first. The sound signal to the left ear will arrive later by the time period indicated on the slider. Looking at the graph of the two signals, the green line for the left ear shows the click being generated when the signal for the right ear is done. Now press the **Play** button. The click will now sound like it arises somewhere nearer to your right ear. Going back to **Interactive Illustration 11.x, Head in Sound Space**, this outcome can be predicted. Sounds that arise from the right hand side of our head will arrive first to the right ear. **Interactive Illustration 11.x, Time of Arrival and Phase** simply simulates this occurrence leading to the outcome you have experienced (Hafter & Dye, 1983).

Now adjust the **Time of Arrival** slider to **-00.50** where the sound will now arrive at the left ear first. Make a prediction about what you should experience and then press the **Play** button. The position of the sound should reverse in your head and now sound closer to your left ear. You can have the computer manipulate the **Time of Arrival** automatically for you by pressing the **Animate** button. The **Animate** button will adjust the relative delay between the left and right ear signals, first making the left ear signal later and then reversing the pattern so that the right ear is after the left ear signal. It repeats this pattern until you press the **Stop** button. You should experience the click first moving towards your right ear and then reverse direction and move back to the middle of your head and then towards your left ear. At the extreme values of **Time of Arrival** you might hear two clicks. When the time of arrival is two great, the brain does not experience them as a single stimulus, but just like when disparity gets to great, you will experience two separate clicks. So there is a limit to time of arrival. These limits will be important later.

Now let us change the **Type of Stimulus** to a **Tone** using the menu in the lower right corner of the screen. The situation is similar but not identical for clicks. The **Time of Arrival** slider is now called the **Phase** slider. The important difference in a tone that last longer than a click is actually the phase of the sound as it arrives at the two ears. When you first bring up the **Tone** stimulus, the waveforms arriving at the two ears will arrive with identical phases. Press the **Play** button to hear the sound with the **Phase** of **0**. The sounds, like the clicks with a time of arrival of 0, should sound like one sound arising from the middle of your head. Press the **Stop** button when you know what you are experiencing.

Now adjust the **Phase** slider so that it reads **72 deg**. The tone to the right ear will precede the tone in the left ear by a phase of 72 deg. Looking at the graph of the sound waves, you should see the peaks for the right ear (cyan) occurring before the peaks for the left ear (green). When you press the **Play** button you should now ear the tone occurring towards your right ear. The tones are actually started at the same time in both ears, only the phase differs. So, in this way we can be sure that the change in the perceived direction of the tone is due solely to the phase of the sounds in each ear and not to time of arrival. Change the **Phase** to a **-72 deg** and the tone will now sound closer to your left ear. Pressing the **Animate** button, when the tone is stopped, will start a sequence similar to the clicks. One-half second tones will be played in each ear with the phase gradually changing the phase to **90 deg** and then to **-90**

deg and back again until you press the **Stop** button. The sound should be experienced as moving through your head, first to the right, then to the left, and then back to the right (Blauert, 1983).

Just as with the clicks, there is a limit to how phase can be used to determine the direction of a sound. Sound cycles are very repetitive which provides a limit to how useful phase is for perceiving direction. The limits on the **Phase** slider are from **-360 deg** to **360 deg**, recall that phase goes from 0 to 360 deg like in a circle. Press the **Center** button to set the **Phase** to **0 deg** or centered then press the **Play** button. Now gradually drag the **Phase** slider to the right towards **360 deg**. As you drag the slider, the sound should appear to move to the right, sound towards the left and then sound in the middle of your hear when you reach **360 deg**. If you look at the soundwave graph and then press the **Center** button it will not look like the graph has changed. Being 360 deg out of phase is the same as being in phase so the sound should sound like it is coming from the same direction. This limit has important implications that will be discussed later.

Interaural Loudness Difference. Now open **Interactive Illustration 11.x, Relative Loudness** to illustrate how interaural loudness differences can help us perceived the direction of a sound. This illustration will only use tones which will simplify our discussion. When you first play the tones, the tones in both ears, which are the same frequency, are at the same intensity and will sound like one tone arising in the middle of your head. Press the **Play** button at the bottom of the screen and listen to the sound. Then, drag the **Left** tone's **Intensity** slider, which is labeled in green, and lower the intensity. As you lower the intensity of the intensity of the tone in the left ear, the sound you experiencie should appear to move towards your right ear. When there is no intensity in the left ear, the sound will appear to come straight from your right side. If you replace the **Left** tone's **Intensity** to 1 and lower the **Right** tone's **Intensity** (labeled in cyan) to 0, the situation will be reversed. The tone will appear to move towards your left ear. As you can see, sound intensity can play a role in our perception of the direction of a sound (Blauert, 1983).

Applications of these Auditory Space Cues. The use of the term stereo, short for stereophonic, for generating separate signals for the left and right ear should not surprise you. The perception of depth that arises from generating separate images for each eye was called stereopsis (see Chapter 8). As a reminder, stereo comes from the Greek word for solids so stereopsis was defined as solid vision. Stereophonic means solid sound. The choice of the word for solid might be a bit odd for both stereopsis and stereophonic, but in both cases, the use of the prefix stereo is related to the attempt to give a more realistic impression of how the stimulus, whether it be an object or a symphony is spread out in the world.

The development of multitrack recording of which stereo is the simplest form, goes back to 1939 and the Disney movie *Fantasia*. The goal was to draw the audience more into the movie by having sounds arise from all around the audience member making them feel more like they were in the middle of the music. These auditory space cues of interaural phase differences and interaural loudness differences are at the heart of the experience that arises from these multi-track recordings. Movies routinely use many speakers in a theater and it is increasingly common to have multiple speakers in homes (the 5.1 and 7.1 surround sound systems). Thus, you might hear a car or horse come from behind you to into the screen. Perhaps one of the most dramatic uses of auditory space perception was the liquidation of the Jewish ghetto scene in *Schindler's List*. In this scene, Steven Spielberg stages a riot all through the area of the city that they were filming. Then they filmed their scenes in the middle of this organized chaos. Thus, wile the audience watched the horror on the screen, they were buffeted by the sounds of the off screen actions.

Why two binaural auditory cues? As in depth perception, having more auditory cues does allow us to have a richer perception of auditory space. The two auditory cues can work together and complement each other and help disambiguate confusing stimuli. Open **Interactive Illustration 11.x, Binaural Cues for Direction**. This illustration will allow you to combine bother interaural phase difference and loudness differences to test how they can combine to us perceive differences in direction. The layout of this illustration is similar to **Interactive Illustration 11.x, Time of Arrival and Phase**. However, there are some additional controls in the lower right corner of the screen. Most important for our current use are the check boxes for **Interaural Phase Diff** and **Interaural Loudness Diff**. These check boxes play a role in which of these binaural cues will be used as the direction of the sound is simulated to come from the left or the right. Mostly easily, press the **Animate** button to start the sound moving back and forth, starting in the middle and then going right. When it reaches the end, it will then reverse course and goes left. If you have both binaural cues selected, you should have a good perception of the sound moving back and forth. Selecting only one of the cues will reduce this perception. This activity is in many ways reminiscent of **Interactive Illustration 8.x, Pictorial Cues to Depth** where the addition of the depth cues lead to a generally stronger perception of depth.

However, these two binaural cues also complement each other in another way. As mentioned above each of these binaural cues have limitations. Reopen **Interactive Illustration 11.x, Head in Sound Space**. Drag the sound source so that is it s straight out from the left ear so that the difference in the distance and phase between the two ears is at a maximum. Now slowly drag the **Frequency** slider to the right increasing the frequency of the sound. The distance difference graph will be unchanged, but the phase difference graph will show an increase in the phase difference between the two ears and then, suddenly the phase difference will go back to 0 deg difference between the two ears. Leave the slider right ear and look carefully at the sound waves as they travel from the sound source to the two ears. Particularly examine the sound wave as it crosses the head. If you look carefully, you will see that there is one complete cycle of the sound wave as it crosses the head from left to right which is why the sound has the same phase as it reaches the two ears. But, when the sound has the same phase at the two ears, that should be consistent with sounds coming from directly in front or behind the person. So, for a tone that uses phase differences between the two ears, there is an upper limit of frequency that can be used. Once the frequency gets so high that the wavelength gets long enough so that it is the length of the head or shorter, then phase differences between the two ears will not be useful as an auditory cue to sound direction. This problem is not relevant for interaural time of arrival. In clicks, it is when the sounds first arrive that important, and that difference does not depend upon frequency.

So, if interaural phase difference works best low frequencies, what about interaural loudness differences. The situation is in many ways the opposite. Think of it this way: imagine waves. Now imagine a small stone in the path of these waves. If the wavelength of your waves is small, then this small stone will disrupt the waves. Cause a break in the waves. This of ripples that are so easily disrupted. Now if the waves have a very large wavelength, this small stone in no way disrupt the waves. Think of tides, waves with very long wavelengths. Even continents do not really disrupt the tides much. For another analogy, think of sitting in your room with the door closed and someone is playing loud music in another building. Perhaps it is a concert on another part of the campus. What instrument do you hear the most clearly? Most people respond the base drums or base guitar. Instruments that make low frequency noises. These sounds are note as easily blocked by the intervening walls and doors.

Looking back at **Interactive Illustration 11.x, Head in Sound Space** examine the intensity of the sound reflections as you adjust the **Frequency** slider. As you make the frequency higher, the reflections get more intense meaning that there is less of the sound traveling to the opposite ear. As you lower the frequency, these reflections will get less intense and even disappear, indicating that the sounds at the two ears have more and more similar intensity until they are the same. The same phenomenon happens for sounds crossing from one side of our head to the other. So, interaural loudness difference works better the higher the frequency. Thus, the very frequencies that interaural phase difference does not work well for are the same frequencies that interaural loudness difference works well for. So, here is another way these two binaural cues work together.

There are some interesting comparative implications from this discussion. Humans have fairly moderate sized heads as far as animals go. Just within mammals, our heads are much larger than mice heads and much smaller than elephant heads. A mouse does not have much head at all and almost no distance between the two ears. Not much size to reflect back sounds. Interaural loudness difference would not seem to be much use as a cue of sound direction for a mouse. However, their ears are quite close together meaning the sound would have to be a much higher frequency before interaural phase difference would no longer be of any use. Conversely with the elephant, their heads are much larger and should be a much more effected block to a much wider range of frequencies making interaural loudness differences much greater than we experience. But their ears are much farther apart making interaural phase differences much less useful. Just are our heads are intermediate in size and, thus, we can use both types of binaural cues, in mice are much more responsive to phase differences and elephants to loudness differences (REF).

The Complexity of Having Two Ears. Recall from Chapter 8 on depth perception, that having two eyes made our visual world much more complex. There are now two images of the objects, one for each eye, and they do not fall on the same place in the two eyes. This is disparity. The disparity leads to the depth perceived from stereopsis, but also means that the brain has to deal with these two images. If the disparity is small enough, then we can fuse the two images and see only one. If the disparity is too large, the brain tries to suppress one of the images, binocular suppression, but that does not always work. Sometimes we experience double images, diplopia.

Well, we have two ears as well and the sound is different in each ear. It is usually a little later in one ear and usually a little softer. The question is does the brain have to do any of the similar sort of actions as in vision so that we don't walk around thinking we are hearing echoes all the time? One phenomenon in hearing suggests that the brain does try something similar to binocular suppression. It seems that the brain suppresses the perception of sounds arriving slightly later. This is known as the suppression of echoes or the **precedence effect [to glossary]**. You have already experienced this phenomenon. Reopen **Interactive Illustration 11.x, Time of Arrival and Phase**. Leave the **Type of Stimulus** to be a **Click**. The press the **Animate** button. The click to one ear will arrive sooner than to the other ear. Still for most of the delays, you will hear it as one click. The suppression of later arriving sounds by the precedence effect is leading you to hear only one click.

Now open **Interactive Illustration 11.x, The Precedence Effect**. In this illustration you can experiment with parameters of a sound in a way to allow you to test and experience the precedence effect. The screen will look similar to **Interactive Illustration 11.x, Time of Arrival and Phase**. You will play a very short tone, 1 msec, so short that it sounds like a click. Recall from earlier in this chapter the discussion of sound duration and whether a sound is experience at a tone or a click. At either end of the tone, the sound intensity is ramped from the peak intensity to no intensity. The leading ramp is very brief, 1 msec. The duration of the ramp at the end of the sound can be adjusted using the **End Ramp** slider near the bottom of the screen. Adjust this slider so that it reads about **100 msec**. You don't have to be real accurate, just in the general area. The waveform of the sound will be updated as you drag the slider. You will see the ending of the sound gradually decline over that approximately 100 msec until there is no intensity left. We can this of this end ramp as a sort of simulation of echoes. Not completely accurate but it will do the job. Now press the **Play** button and listen to the sound. The sound should be quite brief sounding. Like a click with a pitch to it, but very short. Of course the sound is not very long, it is only 2 msec longer than the **End Ramp** duration. Now click the **Reverse** checkbox which is right below the **Play** button. The sound will not be played backwards with the sound now gradually increasing up to the main part of the sound and with a very short tailing off. Press the **Play** button again. With these settings, the sound should seem to last a good bit longer. To really test the difference that the order that the sound is played on try pressing the **Play Pair** button to the right of the **Reverse** checkbox. This button will play the sound first in the forward order with the long trailing sound and then reverse the sound and play it again so that the sound gradually ramps up to full intensity. This quick paring of playing the sound forwards and backwards should really help you hear how the later arriving sounds, particularly if less intense, are suppressed and less noticeable. As a result of this suppression, you can see how we hear far fewer echoes than are actually present. Just as the visual system suppresses many of the double images arising from two eyes, the auditory system suppresses many of the echoes that occur because of all of the solid surfaces that are in most environments.

Now press the **Reset** button at the bottom center of the screen and we will not explore how the precedence effect plays a roles in suppressing the sound from the far ear helping us hear only one sound instead of two. Drag the **Delay** slider at the bottom of the screen to the right so that it reads a positive **5**. The right sound wave, drawn in cyan, will now be sounded first, starting 5 msec before the left sound wave, drawn in green. Next, drag the **Left Gain** slider, in the upper left region of the screen, so that it reads about **0.25**, meaning the left channel signal will be ¼ as intense as the right channel signal. You do not need headphones for this demonstration, even though different signals will be being generated to the two ears; in fact, it works better without headphones. At this point your setup should look something like Figure 11.x.

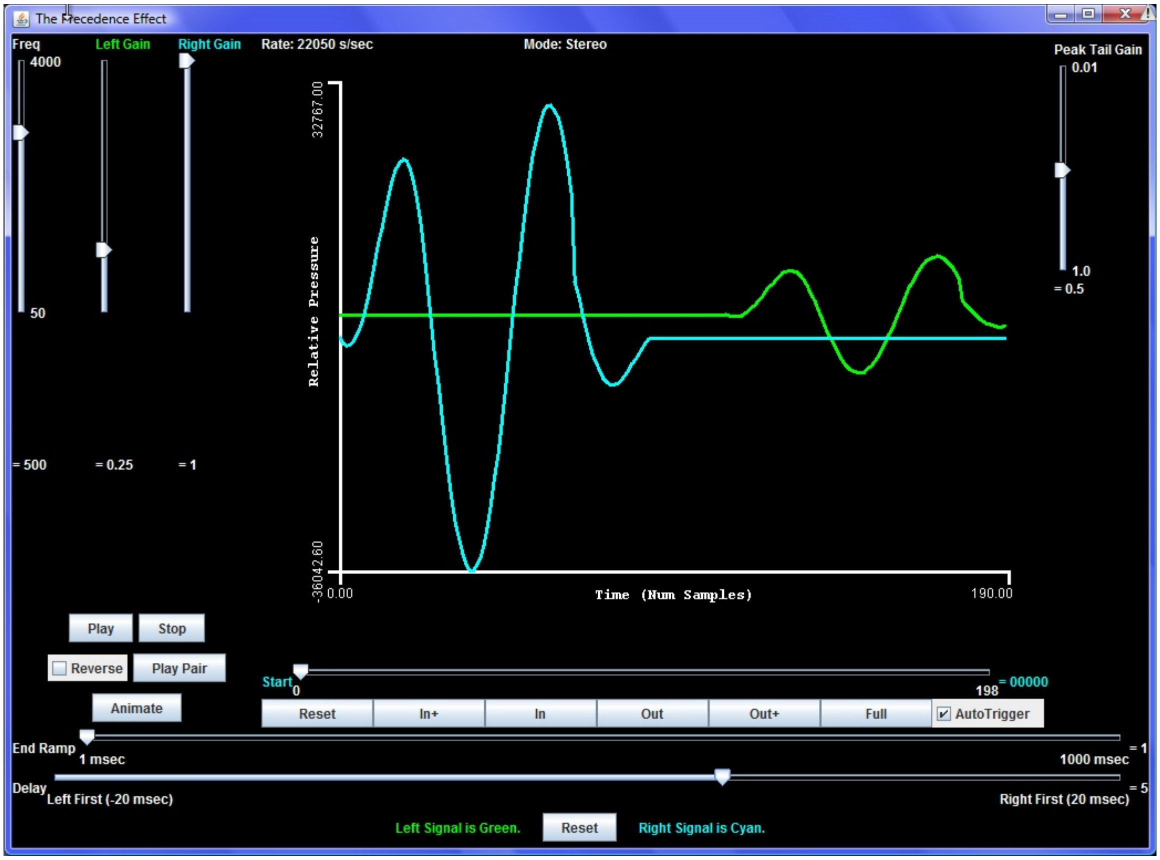

Figure 11.x. Setup of Interactive Illustration 11.x, The Precedence Effect to test binaural supression.

Now press the **Play** button. Press it several time. Most people will report hearing a single click arising from the direction of the right speaker. Next, click on the **Reverse** checkbox. In this situation, the softer sound from the left speaker will be played first. Press the **Play** button again. I suspect you will no hear two clicks, one from each speaker though they may be more difficult to localize, at least the first click. Two sounds, one arriving first, but you only hear the first one. Here the precedence effect is supporting a sort of binaural suppression to keep us from hearing two sounds because we have two ears.

Speech

Auditory Illusions

The discussion of vision ended with an examination of illusions, so it seems appropriate to end the discussion of audition with a couple of intriguing illusions. The first illusion to be discussed is the Shepherd Illusion or Shepherd Scale (REF). This is the same psychologist who is behind the Shepherd Tables that were examined in Chapter 9. This illusion is in some ways an auditory analogue to a visual illusion, the never ending staircase shown in Figure 11.x. Follow around this staircase. If you go counter-clockwise, each step goes up but at the end of the staircase you are right back at the beginning. This illusion falls into the class of impossible figures.

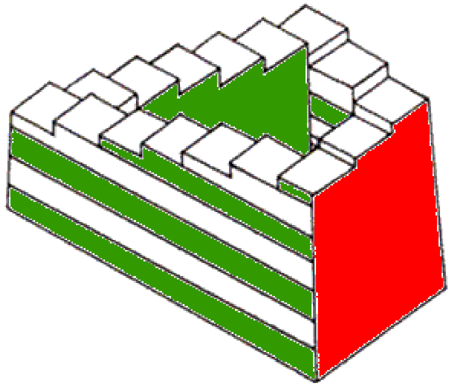

Figure 11.x. The never ending staircase.

The Shepherd scale is very similar. Open **Interactive Illustration 11.x, Shepherd Illusion**. Across the top half of the screen are the notes that will be played. When you press the **Start Tune** button at the bottom of the screen the tune will start with the note on the far left hand side of the screen (the note of C) and then play through all the notes on the scale. The tune will be repeated until the **Stop** button is pressed. Across the bottom half of the screen is a spectral analysis of the note being played. One difference from the spectrums that have been shown before is that the frequencies have been converted to logarithms. As each note in the tune is played the spectrum for that note will be shown on the lower half of the screen.

With this introduction go ahead and press the **Start Tune** button and listen. Like the ever ascending stair case each note sounds like a step up from the previous note. However, you will probably also notice that the notes do repeat themselves. If you wish, you can click on and select the **Follow Notes** check box in the lower left hand corner of the screen and see the note being played highlighted. If you select the **Fundamental Only** check box, you can compare how the illusion sounds with the fundamentals of the notes indicated on the scale. In this case, you can clearly hear the tones drop at the beginning of the scale. Selecting the **Complete Tone** check box will restore the illusion.

While illusions are always fun, they are only useful is they help us understand something about the operation of our sensory systems. To understand this, it is important to understand the construction of these notes, sometimes called Shepherd Tones.

First, these tones are composed of a series of octaves. Recall from above, an octave is always the doubling of the previous frequency. So if the lowest frequency of a Shepherd Tone is 100 Hz, the next will be 200 Hz and the third will be 400 Hz, and so forth. This explains part of the reason that the frequencies on the interactive figure's spectrum are plotted in logarithms. In logarithms to the base 10, a doubling of a value is always an increase of a logarithm of 0.3. So the logarithm of 100 is 2. The logarithm of 200 is 2.3 and the logarithm of 400 is 2.6. Thus, each frequency in the note will be plotted an equal distance apart.

Second, the amplitude of each note is carefully selected. The amplitude of each component are determined according to the normal distribution. If you need to review the normal distribution refer back to Chapter 2 about Signal Detection Theory. A copy of the normal curve is shown in Figure 11.x. On this Figure, the y-axis is the probability of a value and x-axis is the variable being measured. The mean is in the center of the curve and the dashed lines show standard deviations which are measures of how far one is from the mean. In a Shepherd Tone, the x-axis is the log of the frequency value. For the current demonstration, the mean was set to be the logarithm of 440 Hz which is the A above middle C. The y value is the intensity of the frequency. The standard deviation value was set to 0.6 or the logarithm of 2 octaves.

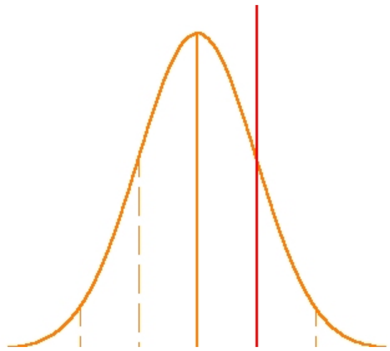

Figure 11.x. The normal distribution with standard deviations indicated.

Finally, to build a note, you start with one frequency, say middle C which has a frequency of about 262 Hz. Then you build the other frequencies all those that fall within 2 standard deviations below and above the mean value on the distribution. That will give you the frequencies and their intensities for the note on the left hand side of the curve. The next note is D which has a frequency of about 294 Hz. This note is built and played and so on for each note. The frequencies nearest 440 Hz will always be the most intense and the frequencies are always between 27.5 Hz and 7040 Hz, 4 octaves below and above 440 Hz, respectively. As the notes goes up the scales, The high frequencies will get too high and be dropped off and new frequencies at the low end of the scale will be added.

You can watch all of this action on the spectrum at the bottom of the screen. If you have stopped the tune, go ahead and restart it and the illusion will be clear. As the tune plays, the intensity of each frequency in a note is highest near the middle of the spectrum and falls off to each side. You can see the normal distribution plotted on the screen. Now, look just at the middle frequencies that have the highest intensity. Don't follow any individual frequency element. As each note is played, these middle frequencies step to the right towards higher frequencies. Here is the ever ascending part of the illusion. It always happens. Now track one frequency element. Pick one that is at the left hand side of the screen so you can see the complete progress. As each successive Shepherd Tone is played, the frequency will step to the right towards a higher frequency. The amplitude will first grow as it approaches the mean of 440 Hz and then decline as it gets further above 440 Hz in frequency. Then this element will be dropped when it gets above 7040 Hz. At this point a new frequency will be added that is far below but just above 27.5 Hz. Here is the repetition of the notes. A frequency element starts at the bottom frequency level travels to the highest frequency level and then is dropped off to be replaced by a frequency at the lowest level.

The sliders on the left hand side of the screen will allow you to adjust the intensity and speed of the tune, the **Intensity** and **Tempo** sliders, respectively. You can also try the illusion on a selection of different major and minor scales using the **Scale to Play** menu in the upper left corner of the screen.

The next illusion is the Octave Illusion, another illusion discussed by Diana Deutsch (REF). It is similar to the scale illusion, simpler to set up, but more complicated it its outcome. Open **Interactive Illustration 11.x, Octave Illusion**. If you have headphones, this illustration works best with them. If you do not have headphones, set you speakers as far apart as possible. The setup of the illusion is both shown on the screen and illustrated in Figure 11.x. In each ear, the note C is played over an over. In the left ear, first middle C is played and then the C one octave higher. These two notes are repeated 6 times. On the right ear, the order is reverse, first the C above middle C is played and then middle C is played. Again, this pattern is repeated 6 times. So, when middle C is played in the left ear, the C above middle C is played in the right ear, and vice versa. Press the **Start Tune** button and see what you hear.

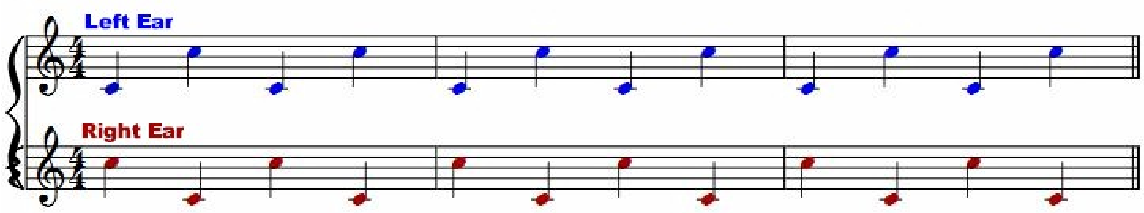

Figure 11.x. The notes played during the Octave Illusion.

People have different experiences in this illusion, but the most common outcome is shown in Figure 11.x. On ear will hear the middle C and the other ear will hear the C an octave higher. So far this is very like the scale illusion and we can think of the Gestalt Law of Proximity playing a role. The middle C's are gathered perceptually into one ear, and the C an octave higher is gathered into the other ear. However, look more closely at Figure 11.x which is showing this common perception of the octave illusion. When you ear the middle C, you don't hear the C an octave higher in the other ear. Most people report hearing the two C's not only in two ears, but alternating in time. So, say you hear the middle C's in the left ear. You will hear just that middle C and the next note will be the C an octave higher in the other ear. If you look at the two Figures illustrating the Octave illusion, you will see that you seem to perceive only half of the notes. So in addition to some sort of Gestalt process, there is some sort of suppression going on.

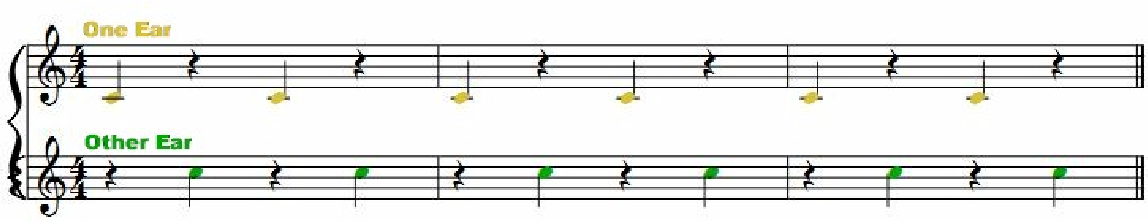

Figure 11.x. One common experience of the octave illusion.

Play around with the illusion. Some questions you might consider, does the illusion depend upon the interval being an octave. You can change the interval two ways. In the lower left hand corner of the screen is the **Change Interval** menu. You can pick different intervals from an octave down to a semitone (a half step or the distance between two adjacent keys on the piano). The other way are to use **Raise Lower**, **Lower Lower**, **Lower Upper**, and **Raise Upper** buttons across the bottom of the screen. The buttons will adjust the upper or lower of the pair of notes by a semitone. If you lower the lower note enough you could make the interval two octaves and see what happens then. In addition, you can make the notes different in the two ears. The **Raise L Half Step**, **Lower L Half Step**, **Lower R Half Step** and **Raise R Half Step** buttons will change all the pitches in the indicated ear by a semitone in the indicated direction. Sliders on the right control the speed (**Tempo**) and intensity.