Lecture 13: More on Perception of Loudness

We have now seen that perception of loudness is not linear in how loud a sound is, but scales roughly as a factor of 2 in perception for a factor of 10 in intensity. However, we still have not seen how to compare,

- sounds of different frequency as well as intensity
- sounds of different duration
- sounds composed of several sine waves, such as real musical tones or simultaneous tones from several instruments.

Let's talk about these next.

Frequency and Loudness

Tones of the same intensity (power per area), but of different frequency, are perceived as being of different loudness. To simplify the discussion, consider just sustained tones where the pressure is a sine wave. If two sounds have the same intensity and their frequencies lie between about 600 and 2000 Hertz, they will be perceived to be about the same loudness. Outside of this range, that is not the case. For sounds near 3000 to 4000 Hertz, the ear is extra-sensitive; these sounds are perceived as being louder than a 1000 Hertz sound of the same intensity. At frequencies lower than 300 Hertz, the ear becomes less sensitive; sounds with this frequency are perceived as being less loud than a sound of the same intensity and 1000 Hertz frequency. The loss of sensitivity gets bigger as one goes to lower frequencies. Also, at very high frequencies sensitivity is again reduced.

Let us very briefly explain why each of these features is present.

- Hearing below about 300 Hertz becomes inefficient partly because the cochlea does not respond as well here, but also largely because the transmission of the vibrations through the ear bones becomes less efficient at low frequencies. [This is a common problem with impedance matching: below some characteristic frequency of the impedance matcher, it stops working efficiently. We might return to this after talking about resonance and impedance, when we discuss brass instruments.]
- The meatus is a tube, roughly cylindrical and of a certain length. As we will see later, such a tube has a resonant frequency, and sound waves at or near that frequency bounce back and forth several times in the tube before leaving, giving the ear a larger sensitivity to capture that sound. As we will see, the resonant wave length is $\lambda = 4L$ with L the length of the tube. This gives a number around 3500 Hertz, and explains why the ear has a region of especially high efficiency there. This explanation may not

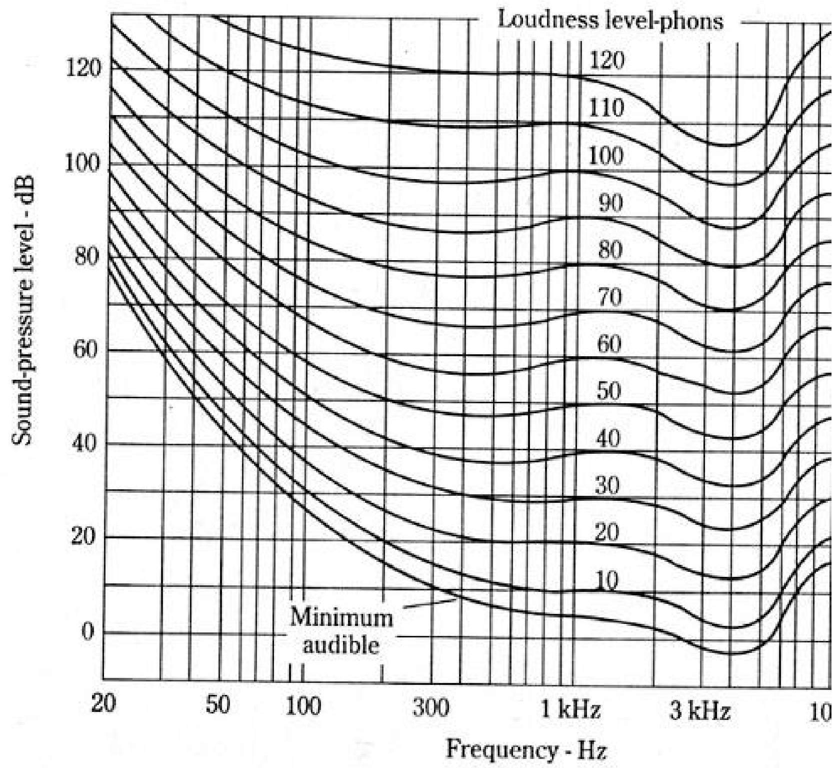

Figure 1: Fletcher-Munson curve, showing what sounds will be perceived as equally loud.

make sense to you now, but we will return to it when we talk about resonances in tubes.

- At very high frequencies one goes beyond the frequency where the cochlea is designed to work efficiently. Also, the hair cells responsible for the highest frequencies die with age and exposure to loud sounds, so the high frequency cutoff of the ear tends to move to lower frequencies with age.

We saw that it is hard to answer the question, "how many times louder is sound A than sound B?" However, it is much easier to answer, "is sound A louder, softer, or the same as sound B?" Furthermore, different people will generally give the same answer for the same pair of sounds. That is, if I find a 1000 Hertz tone and a 200 Hertz tone which one person finds to be of equal loudness, another person will also find them to be of equal loudness. (The exception is frequencies above 10 000 Hertz, where some peoples' ears lose sensitivity at a lower frequency than others.)

Adopting 1000 Hertz sine waves as a standard, we can then ask, what intensity must a sound at 100, 200, 300, ... Hertz be, to sound as loud as a 60 dB tone at 1000 Hertz? The answer will be a curve in a plot with frequency on the x and intensity on the y axes. We can also make curves for a 50 dB tone at 1000 Hertz, a 40 dB tone at 1000 Hertz, and so forth. The first people to do this were named Fletcher and Munson, so such a curve is referred to as a Fletcher-Munson curve, and is shown in figure 1. Having just advertized that different people will give the same answer, I nevertheless found at least 2 Fletcher-Munson curves on the web which give slightly different answers. This may be an issue of how modern the equipment used was. There is a beautiful website, linked off the course page entry for this lecture, which lets you measure your own Fletcher-Munson curve.

Let us take some time to explain what the Fletcher-Munson curve means. We will do so by answering some example questions, using the plot.

Q. What must be the sound intensity of a 100 Hertz sine wave, if it is to sound as loud as a 1000 Hertz sine wave of 60 dB intensity?

A. Look on the plot at the curve labeled "60". All points on this line sound as loud as a 1000 Hertz sine wave of 60 dB intensity. (1000 Hertz because it happens to be the standard, 60 dB because the curve is labeled "60".) Find the place this curve intersects the vertical line going up from 100 Hz on the x axis. Now look up what intensity that is, by going horizontally to the y axis: it is most of the way from the 60dB to the 70dB line, so it is about **67 dB**. The relevant points on the Fletcher-Munson curve are circled in green in figure 2.

Q. how loud must a 4kHz sound be, to sound as loud as an 80 dB sine wave at 400 Hertz?

A. First we have to find the point on the graph for 400 Hertz and 80 dB. Find the vertical bar corresponding to 400 Hertz on the x axis, and find where it meets the horizontal bar which is 80 dB on the y axis. We see that the point where they meet is a little above the curve marked "80" but well below the curve marked "90". If we filled in the curves between 80 and 90, it would be about on the "83" curve. A sound of equal perceived loudness must be the same amount up from the 80 curve. So go over to the 4000 Hertz vertical line, and see how far up it you must go to be 3/10 of the way from the "80" curve to the "90" curve. The curve marked "80" meets the vertical bar at 4000 Hertz at about 70 dB on the vertical axis. To go 3/10 of the way to the curve marked "90", we go up to about **73 dB** actual intensity. The relevant points on the Fletcher-Munson curve are put in red squares in figure 2.

Q. Which sounds louder: a 60 Hertz, 60 dB tone, or a 2000 Hertz, 45 dB tone?

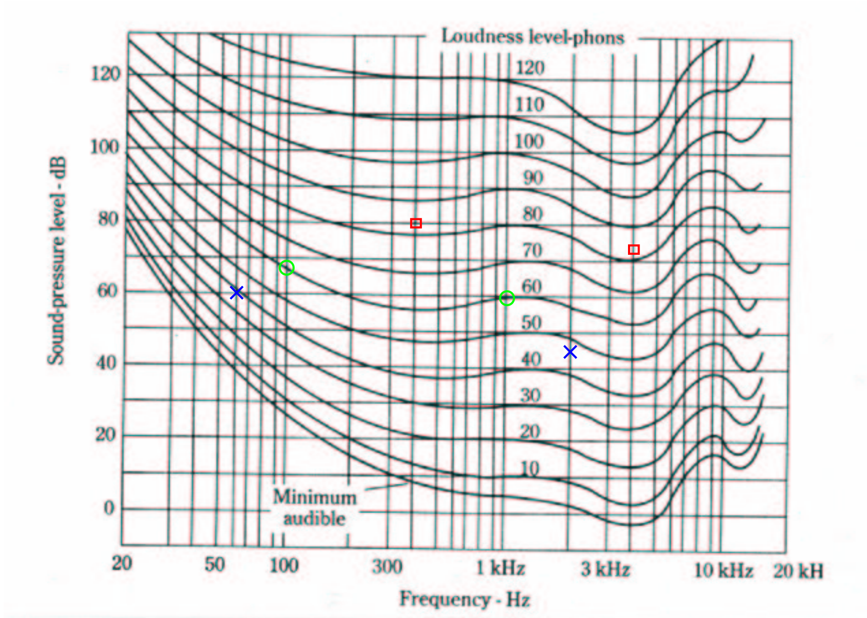

Figure 2: Fletcher-Munson curve marked with the points relevant for the questions in the text: green circles, first question; red squares, second question; blue crosses, third question.

> **A.** The two frequency-loudness points are displayed as blue crosses in figure 2. Note that the cross at 60 Hertz and 60 dB lies just below the curve marked "40" meaning that it is quieter than a 40 dB, 1000 Hertz sound. The cross at 2000 Hertz and 45 dB lies above the line marked "40" so it is louder than such a tone. Therefore, the 2000 Hertz, 45 dB tone sounds louder.

It is usual to define a unit of perceived loudness called a "phon", defined as,

Phon: A tone is x phons if it sounds as loud as a 1000 Hertz sine wave of x decibels (dB).

That is, all the points on the Fletcher-Munson curve labeled 70 are 70 phons loud, and so forth. The phon is defined to take into account the differences in ear efficiency at different frequency. While dB tell the intensity, a physical measure of loudness, phons give a scale which honestly compares how loud sounds will "sound."

While two sounds of the same number of phons are perceived as being equally loud, phons have the same problem as decibels in terms of interpreting just how much louder a sound with more phons will sound. That is, 60 phons does *not* sound twice as loud as 30 phons; in terms of perceptions, they are not a linear scale.

One can over-literally interpret the "rule of thumb" we met previously, that 10× as large an intensity sounds like 2× as loud a sound, and define a unit called a "sone" as follows:

Sone: A tone of x phons is $2^{(x-40)/10}$ sones.

That is,
40 phons is 1 sone (choice of starting point)
50 phons is 2 sones (since it should sound 2× as loud as 40 phons)
60 phons is 4 sones (since it should sound 2× louder still)
70 phons is 8 sones
80 phons is 16 sones
so on. In theory, if our "rule of thumb" that 10× the intensity really sounds 2× as loud, the **sone should correspond with how loud a sound actually "sounds."** That is, twice as many sones should really mean, twice as loud sounding. However I emphasize that this is subject to the caveats about how hard it is to really put a number on how many times louder one sound is than another sound.

What about sounds which are not sine waves?

Any sound, whether periodic or not, can be broken into sine waves of different frequencies. The linearity of sound in air tells us that:

The intensity of a sound is the sum of the intensities of the sine waves which make up the sound. That is the answer to the "physics question" of how sine waves add up into complex sounds. Let us see how to add up intensities to get the total intensity of a sound.

You can describe quantitatively the timbre of a periodic sound by telling the intensity of each harmonic. For instance, suppose a cello plays a tone at 200 Hertz. You can describe (almost[1]) all the information about its timbre by telling how many dB or W/m^2 loud it is in each harmonic, at 200, 400, 600, 800, ... Hertz. Suppose the answer is, 60 dB at 200, 56 dB at 400, 54 dB at 600, 57 dB at 800, and very small at higher frequencies. To find the total intensity in dB, one must convert them into intensities, add the intensitites, and go back to dB:

To add intensities, you must add in W/m^2, not dB. 40 dB plus 40 dB is *not* 80 dB. For the example above, we convert each dB measure into W/m^2: the 60, 56, 54, and 57 dB from the example are (10, 4, 2.5, and 5)$\times 10^{-7}$ W/m^2. Adding gives 21.5×10^{-7} W/m^2, which is 63.3 dB.

[1]The remaining information is the relative phases of the harmonics. The ear gets almost no information from these phases, except perhaps for high harmonics where the critical bands strongly overlap.

Unless several components have almost the same intensity, the intensity in dB is usually almost the same as the intensity in dB of the loudest component.

What about perception?

Widely different frequencies cause excitation in different areas on the cochlea. Therefore, to the extent that perceived loudness is decided by how many nerves are firing, the perceived loudnesses of the different frequency components *do* add up. But if two frequencies are close enough together that their critical bands overlap, then they are exciting some of the same hair cells. All the mechanisms we saw last lecture, involving the way hair cells are ennervated, mean that the number of nerve signals will be less than the sum of the signals if each sound were separate. Therefore,

Sine waves of very different frequency have their perceived loudnesses (in sones, say) add up. Sine waves of nearby frequencies do not and act more like increasing intensity of a single sound.

Unfortunately there is no simple rule for how to add perceived loudnesses, only this rule of thumb. These rules of thumb have some interesting consequences for music.

- An instrument with a rich frequency spectrum (significant intensity in several harmonics) will sound louder than an instrument which makes something close to a sine wave. This is especially so if the spectrum of frequencies reaches over 1000 Hertz, where the ear is more sensitive–especially if it reaches the 3000 Hertz range.
- Adding different instruments with different timbres, and/or playing different notes, increases perceived loudness faster than adding more of the same instrument playing the same note. Ten violins playing the same note only sound about twice as loud as one violin, but ten different instruments combined in harmony (playing different frequencies) will sound more than twice as loud as any one of the instruments (though probably not quite ten times as loud).

Masking

Next, we should ask questions about whether one sound can cover up another so you will not notice it, which is called **masking**.

To warm up, let us ask how much you can change the loudness of a single sound, before you notice that the loudness has changed. That is, how large a change in intensity are you sensitive to? The sound file provided with the lecture contains a tone played 4 times. Each time it changes in loudness (either louder or softer) halfway through the tone. The changes

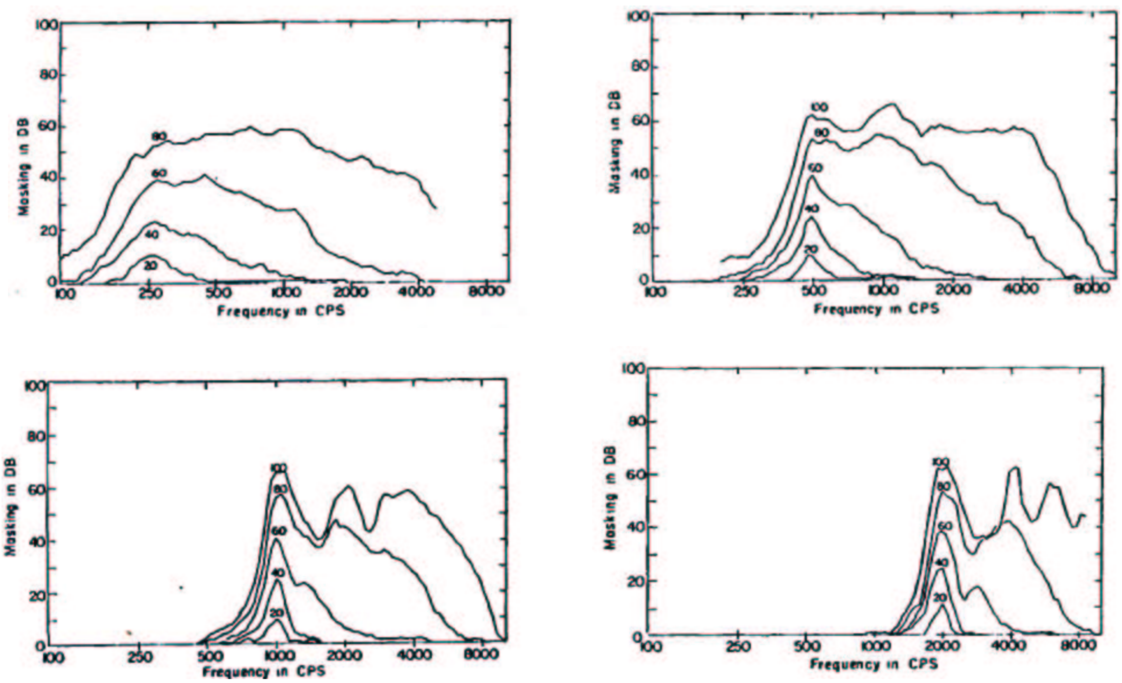

Figure 3: Masking curves, where the loud tone is 250 Hz (top left), 500 Hz (top right), 1000 Hz (bottom left), and 2000 Hz (bottom right).

are by 10%, 5%, 20%, and 40%. Which tones get louder, and which get softer? The 10% tone is hard to tell; the 40% tone is easy. The 5% tone is probably impossible. Therefore, your sensitivity to a change in intensity is about 10 or 20 percent in intensity, or around 0.5 dB.

If I play one tone at frequency f, and I turn on another tone at a very slightly different frequency f', I can tell the second tone is there when the intensity is about 1/100 of the first intensity. We will see why it is 1/100 instead of 1/10 (as we just found) when we talk about beats. Notice that if the first tone is, say, 60 dB, that means that a 40 dB tone at close to the same frequency will be completely covered up.

On the other hand, if I play a high frequency at a loud intensity and a much lower frequency at a very small intensity, since the lower frequency sound excites a completely different spot on the cochlea, I should be able to hear the soft tone despite the loud one. Not only should I, but I actually can. Therefore, whether one tone covers up another or not, really depends on both the loudnesses and the frequencies of the two tones.

When a loud tone makes it impossible for your ear to notice a soft tone, we say that the loud tone is **masking** the soft tone, which is **being masked**. Again, this is an issue of perception, so one has to study it on human subjects. Again, whether a soft tone shows up underneath a loud tone can vary from subject to subject, particularly on their level of musical training. Nevertheless, one can make plots of what "typical" listeners can or cannot hear. Such plots are shown in figure 3, which requires even more clarification than the Fletcher-Munson curve.

Suppose I play a tone of 500 Hertz and 60 dB. What sounds can I hear at the same time, and what sounds will be drowned out? Find the curve on the 500 Hertz plot (upper right in the figure) which is labeled "60." Everything below this curve is drowned out by the tone; everything above it can be heard. For instance, for a tone of 1000 Hertz, the curve is at the 20 dB level. Therefore, we find that 1000 Hz sounds of more than about 20 dB are audible, sounds of less than 20 dB are not.

The 20 and 40 dB curves can be thought of as mapping out the critical band of the masking tone. Wherever the loud tone is vibrating the cochlear membrane, another tiny sound which tries to vibrate it at the same place will not be noticable. The curves for very loud sounds, 80 and 100 dB, look very different; in particular there are high frequency bumps and features, often at multiples of the main pitch. We will learn more about why when we talk about aural harmonics.

Things to know about masking are,

- Generally, playing one tone masks tones of very nearby frequencies if the nearby pitch is more than 20 dB softer.
- Masking is stronger on the high frequency side of the loud tone than on the low frequency side. That is, a deep, loud tone covers up high pitched, soft tones, but a high pitched, loud tone does not cover up deep, soft tones very much.
- Very loud tones have funny effects.
- Noise with many frequencies masks more effectively than pure tones.

There are also some funny and unexpected kinds of masking which tell you that it is only partly in the ear, and partly in the brain:

- A loud sound can mask soft sounds even if they come up to 0.4 seconds after the loud sound turns off.
- A loud sound can mask a soft sound if it turns on within .040 seconds of the start of the soft sound.
- A loud tone played only in one ear can mask a soft tone played only in the other ear.

The last thing to discuss about loudness is very short sounds. The ear seems to "add up" the loudness of a sound over about 1/10 of a second, so a very short sound of a given intensity does not sound as loud as a longer sound of the same intensity. There is a sound file to illustrate this, but it is difficult to make it work well because of speaker transients.