Psychoacoustics Lab Activity

RealSimple Project*
Ryan J. Cassidy and Julius O. Smith III
Center for Computer Research in Music and Acoustics (CCRMA), and the
Department of Electrical Engineering
Stanford University
Stanford, CA

Abstract

This laboratory exercise, aimed at an early college-level audience, serves as a primer on basic psychoacoustic principles. In addition, an included extended activity involves having the student compute a state-of-the-art loudness profile based on work by leading researchers in the field.

Contents

*Work supported by the Wallenberg Global Learning Network

1 Introduction

In this lab, you will learn some basic concepts associated with psychoacoustics, and perform a lab activity to help demonstrate these ideas. Psychoacoustics is the science of human hearing, with particular attention to the physical principles associated with the organs and tissue used for human hearing, or *audition*. An average human listener can perceive pure tones with frequencies as low as 20 Hz, or as high as 20 kHz (a factor of 1000 different!). The range of loudness levels (more on this to come) that the average human can perceive is even larger: the ratio in pressure between the loudest sound comfortably perceptible, and the quietest perceptible sound is approximately 10^{10}!

In particular, your activity for this lab will be to learn more and experiment with a computer model of loudness. This model will allow you to accurately predict from a recorded sound what its average loudness will be. You will also learn how loudness changes over time as a sound changes over time. Ideally, this lab should follow completion of the monochord laboratory assignment[1], so that you can compute loudness estimates on your own recorded string sounds.

2 Objectives

- Understand the basic principles and facts of psychoacoustics.
- Learn the processes associated with the human perception of loudness.
- Experiment with a computer program designed to perform the various stages of loudness estimation for a variety of recorded sounds.

3 Background and Theory

From prior activities[2], you will recall that the frequency of a sound generally corresponds to its pitch. It is interesting to note that pitch is, in fact, a *perceptual quantity*, which means that different individuals may perceive pitches of sounds in a slightly different manner. Frequency, by contrast, is a *physical quantity*, whose value does not depend on the observer. In psychoacoustics, we often find that things we can measure have both a perceptual and a physical quantity associated with them.

In this lab, the extended activity will focus on a well-known perceptual quantity, *loudness*. This is a concept with which you should be familiar, especially if you have been scolded by parents to keep your music volume down. Loudness is a perceptual quantity, whose corresponding physical quantity is sound intensity.

3.1 Auditory Anatomy

Here we provide a functional description of the human auditory system. In this lab, we focus primarily on the *peripheral* auditory system, or the portion of the auditory system other than the auditory tissue of the spine and brain.

The human auditory periphery may be divided into three regions:

[1] http://ccrma.stanford.edu/realsimple/lab_inst/
[2] http://ccrma.stanford.edu/ jos/lab_inst

1. the outer ear, involving the externally visible portion of the auditory system along the ear canal,
2. the middle ear, which transmits sound vibration from the end of the ear canal to the cochlea, and
3. the inner ear, whose primary organ is the cochlea, responsible for transmitting sound information to the nervous system.

Sound enters the auditory system via the outer ear, a diagram of which is shown in Figure 1. The pinna and the concha comprise the externally visible portion of the outer ear, and serve to focus incoming sound waves on the entrance of the ear canal. Upon traveling through the ear canal, these waves then cause the ear drum to vibrate. Next, the middle ear is responsible for transforming ear drum vibration into a similar vibration on the oval window, which is connected to the cochlea. This transformation is performed by an intricate arrangement of three tiny bones, or ossicles: the malleus (or hammer), the incus (or anvil), and the stapes (or stirrup, see Figure 2). Finally, vibration of the oval window causes a wave to travel along the fluid of the cochlea, the function of which is discussed in the following sections.

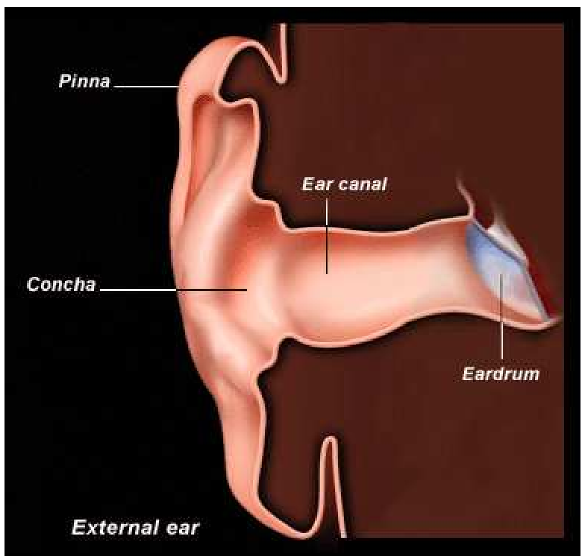

Figure 1: Diagram of the outer ear (re-printed with permission from [2]).

3.2 The Cochlea

The principle auditory organ of the inner ear is the cochlea, shown in the inner ear diagram of Figure 3. Though spiral in shape, it is easiest to think about the cochlea by imagining it has been "unrolled." When thought of in this way, the cochlea contains three parallel ducts running along its length: the scala vestibuli, the scala media (or cochlear duct), and the scala timpani (see Figure 4). The second of these, the scala media, contains a fluid (called *endolymph*) that is chemically quite different from that of the other two (called *perilymph*). The three ducts are separated by two membranes: Reissner's membrane, and the acoustically more significant Basilar Membrane (BM). A diagram of these features is shown in the cross section of Figure 4.

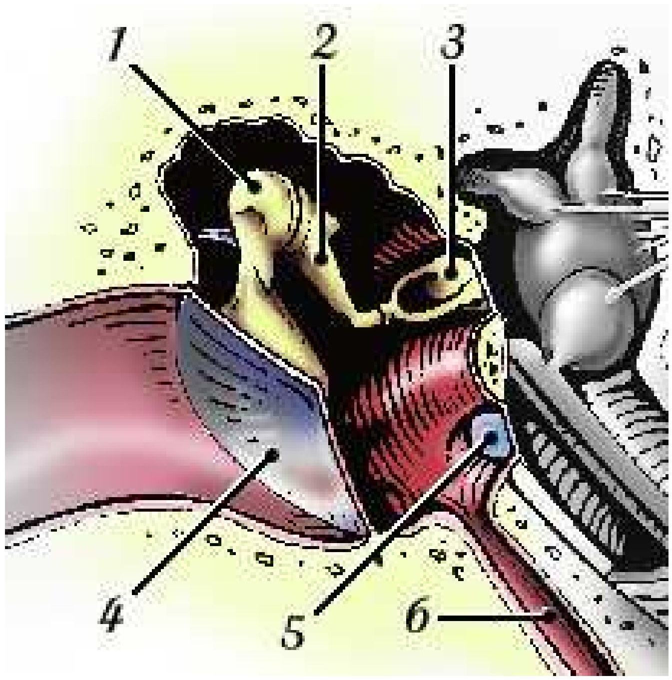

Figure 2: Diagram of the middle ear (re-printed with permission from [2]).

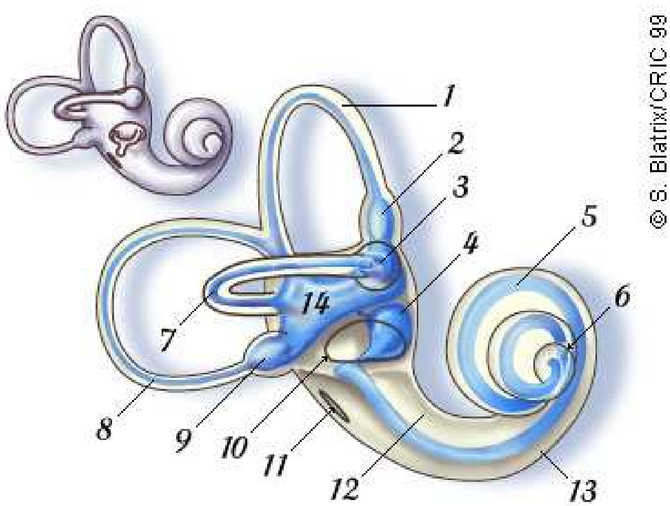

Figure 3: Diagram of the inner ear. The cochlea, with its spiral shape, is the primary auditory organ of the inner ear (re-printed with permission from [2]).

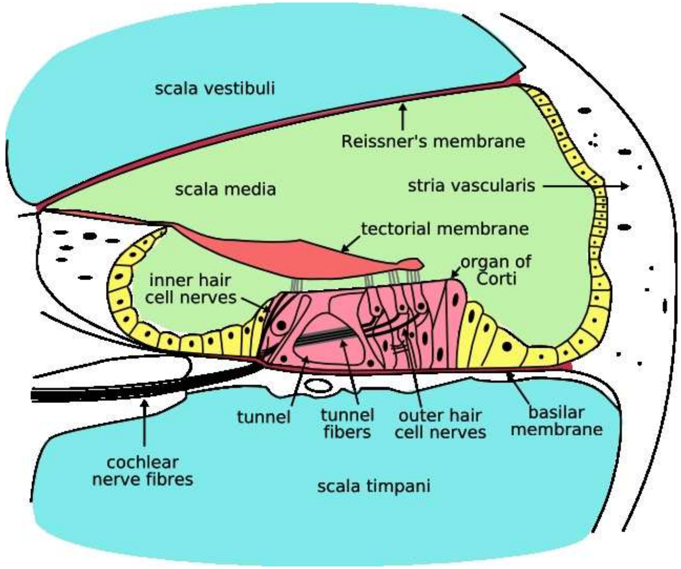

Figure 4: Cross-section of the cochlea (re-printed with permission from [1]).

3.3 Place Theory of Sound Perception

Upon being launched into the scala vestibuli by the oval window, sound waves travel to the end of the cochlea, and return through the scala timpani, where they cause vibrations of the round window (see Figure 3). In the process, however, they also cause the basilar membrane to vibrate in a specific way: sound energy is said to be organized *tonotopically* along the basilar membrane. This means that energy occupying a given spectral range is said to cause displacement of a corresponding region of the BM. In other words, a sinusoidal sound of a given frequency will cause a corresponding peak vibration at a unique place on the BM. The effect of different BM vibrations for sounds of various spectral contents is demonstrated by an online animation.

3.4 Cochlear Excitation Patterns and the Excitogram

Motion of fluid in the three cochlear ducts, along with the consequent displacement of the BM, results in the displacement of stereocilia connected to hair cells attached to the BM. Most importantly, it is these hair cells which transmit auditory signals to the brain via the auditory nerve. The average neural activity in response to a steady sound as a function of frequency is referred to as an *excitation pattern.* For a non-steady, or time-varying sound, we can find the excitation pattern as it changes as a function of time. We refer to a plot of this changing excitation pattern as a *cochlear excitogram.*

3.5 Equivalent Rectangular Bandwidth (ERB)

As discussed in Section 3.4, excitation patterns are functions of frequency—that is, they give the average neural activity for each audible frequency. When plotting excitation patterns and excitograms, it is preferable to view frequency in terms of *Equivalent Rectangular Bandwidth (ERB).* A corresponding frequency expressed in terms of *ERB-rate* is scaled in a manner that gives a more psychoacoustically appropriate weighting to a given frequency range than would be obtained if Hz were used to measure frequency. The formula to compute a frequency in terms of ERB-rate from a given frequency f in Hz is:

$$ERB_{\text{rate}} = 21.4 \log_{10} (0.00437f + 1) . \quad (1)$$

4 Procedure

4.1 Creating a Cochlear Excitogram

1. If you have not already done so, install the program Octave on your computer. Start the program by typing `octave` on the command line. Also, download the archive of source code required for this lab[3], and uncompress the archive into the directory in which you will be running Octave.

2. We need to load a sound into the Octave environment. For this purpose, we have created an Octave function to load a sample sound for you. On the Octave command line, type the following:

   ```
   > x = loadLoudnessTestSound();
   ```

   This creates a test sound, and stores it in the Octave variable `x`.

3. Next, it is necessary to apply a filter to the sound to simulate the effects of transmission through the outer and middle ear. To do this, enter the following command:

   ```
   > xFilt = transmissionOuterMidEar(x);
   ```

4. We next wish to determine the short-time spectra of the signal as they vary over time, which may be found using a form of Short-Time Fourier Transform (STFT). Compute this by typing the following:

   ```
   > xSTFT = cochlearInput2fftInt(xFilt);
   ```

5. Finally, create and plot a cochlear excitogram using the following command:

   ```
   > [xExcitogram,bandCentersHz] = fftInt2ExcitPat(xSTFT);
   ```

   In this command, the excitogram is stored in the Octave variable `xExcitogram`, and the variable `bandCentersHz` gives the frequencies at which the excitogram is computed. This command may require a minute or so to finish on your computer.

   You can create a plot of the excitogram using the following:

   ```
   > excitPatPlot(xExcitogram, bandCentersHz);
   ```

   The plot should resemble that shown in Fig. 5.

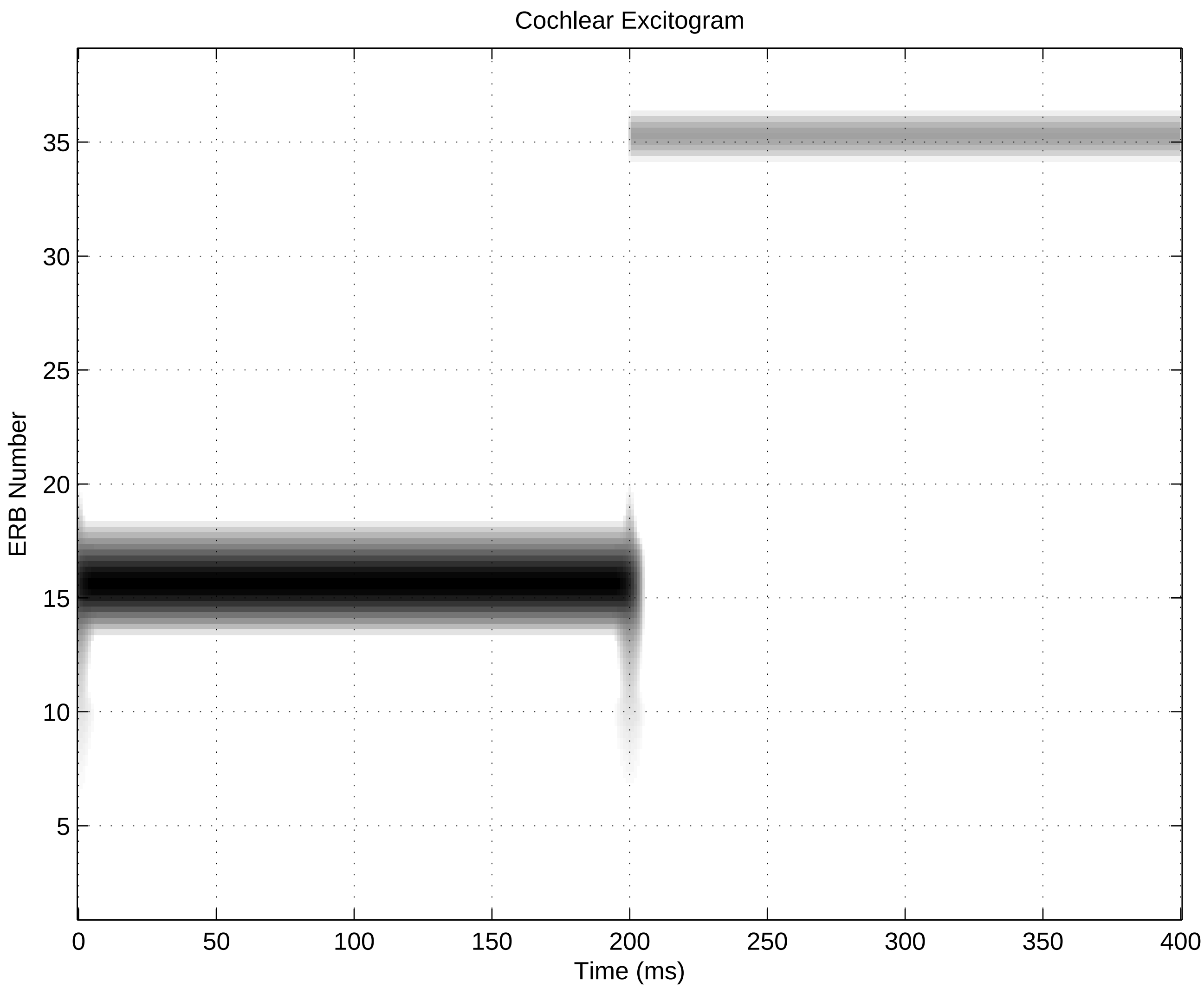

Figure 5: A sample cochlear excitogram.

How many sound objects do you see on the excitogram? Approximately what ERB numbers do these sound objects correspond to? What frequencies in Hz do these ERB numbers correspond to? Which sound object creates the largest cochlear excitation? Which creates the smallest cochlear excitation?

4.2 Creating a Time-Varying Loudness Profile

This section assumes you have just created a cochlear excitogram in Octave as described in Section 4.1.

1. The first step in creating a time-varying loudness profile from a cochlear excitogram involves creating a specific loudness-gram. You can do this in Octave using the following command:

   ```
   > xSpecLoudgram = excitPat2SpecLoud(xExcitogram, bandCentersHz);
   ```

2. Second, to compute and plot the instantaneous, short-term, and long-term loudness profiles from the specific loudness-grams, issue the following command:

   ```
   > specLoud2LoudProf(xSpecLoudgram, bandCentersHz);
   ```

   A sample loudness profile, plotted in both sones and phons, is shown in Figure 6.

5 Acknowledgements

Thanks to Patty Huang, who provided initial Matlab code for time-varying loudness computation.

References

[1] *Image:Cochlea-crosssection.png*, Available online at http://en.wikipedia.org/wiki/Image:Cochlea-crosssection.png.

[2] R. Pujol, V. Reclar-Enjalbert, and T. Pujol, *Promenade 'round the Cochlea*, Available online at http://www.iurc.montp.inserm.fr/cric/audition/english/start2.htm.

[3] http://ccrma.stanford.edu/ jos/psychoacoustics/loudspec.zip

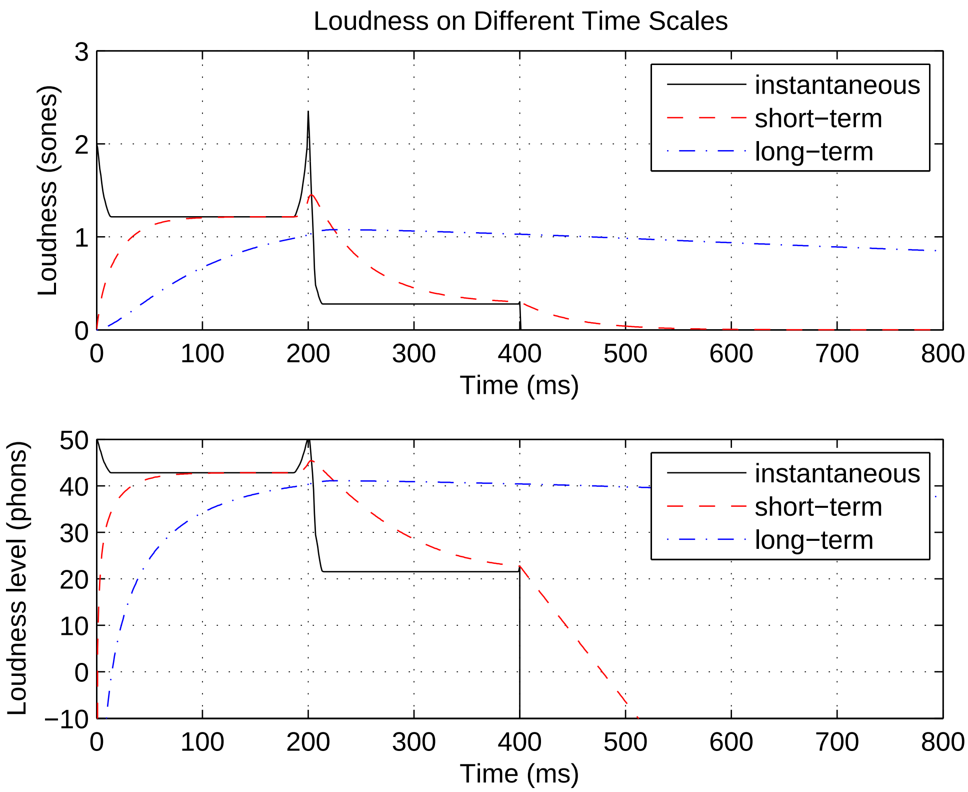

Figure 6: Sample instantaneous, short-term, and long-term loudness profiles.