HST.723J, Spring 2005
Lab Project 2

Madhu Shashanka
shashanka@cns.bu.edu

Introduction

The goal of this lab was to introduce the fundamental principles of perceptual audio encoding and the basics of psychoacoustic models used in perceptual audio encoding. In this report, I describe the experiments I conducted to measure my hearing thresholds in quiet and in the presence of different maskers and how they relate to audio encoding.

Both experiments described below involved listening to sound waveforms. The waveforms were created on a PC using MATLAB and sound was generated using a 24-bit digital-to-analog converter in the PC. The electrical signal was then fed via a headphone buffer (TDT HB6) to a sound-proof booth where stimuli were presented via Sennheiser HD580 headphones.

Experiment 1 - Masking Pattern

The aim of this experiment was to measure the masking pattern for a narrowband noise centered at 1 kHz.

Methods

Absolute hearing thresholds were measured in quiet and in the presence of a simultaneous masker using the Method of Adjustment, also known as Bekesy Tracking Method. The target tone was repeatedly played, sweeping in frequency from 100 Hz to 8 kHz in one run and from 8 kHz down to 100 Hz in the second run. The subject's task was to control the loudness of the target tone by continuously pressing/releasing a button so that the tone was just barely detectable.

This was done under two conditions: in quiet and in the presence of a masker. The masker was a narrowband noise with a bandwidth from 950 Hz to 1050 Hz and a spectrum level of 70 dB SPL. The starting value for target tone intensity was 70 dB SPL. Stimuli were always presented to the left ear.

Results

Figure 2 shows thresholds for ascending and descending tone sweeps under the quiet condition. Figure 3 shows results for the condition when the masker is present. Figure 4 shows the average quiet and masked thresholds and the masking pattern is evident in the figure.

Conclusions and Discussion

The masked threshold curve in figure 4 corresponds to a masking pattern. In the presence of the masker, any sound whose level is below the the value of the red curve at its frequency is not detected by the listener.

Digitizing audio leads to quantization noise which is the difference between the analog signal and its digital representation. Increasing the number of bits used to encode the sound results in less quantization noise. However, one can use psychoacoustic knowledge to reduce the number of bits to faithfully represent the signal. Referring to figure 4, one just has to make sure that the quantization noise accompanying a sound like the masker has a level whose value at a given frequency falls below the corresponding value on the red curve. For example, a 30 dB SPL sound at 800 Hz will be inaudible in the presence of the masker. The basic principle in perceptual audio encoding is to use the masking pattern of the stimulus to determine the least number of bits necessary for each frequency sub-band so as to prevent the quantization noise from becoming audible. The number of bits used to encode each frequency sub-band is equal to the least number of bits with a quantization noise that is below the minimum masking threshold for that sub-band.

Experiment 2 - Masking Threshold

It has been observed that tones and noise maskers have different masking effects. In general, noise maskers are more effective. Let the masking signal be represented in the form $A(t)e^{j\omega(t)+\phi(t)}$. For a narrowband gaussian signal, $e^{j\omega(t)}$ is approximately the same as a tone centered at the same frequency. The functions $A(t)$ and $\phi(t)$ are constants in case of tones and they fluctuate with t in the case of noise. The aim of this experiment was to figure out whether the asymmetry effect was due to the amplitude term $A(t)$ or the phase term $\phi(t)$, or a combination of both.

Methods

Masking thresholds were measured for four types of maskers: 1 kHz tone, gaussian noise, multiplied noise and low-noise noise. The target signal was gaussian noise. A two-interval, two-alternative forced choice procedure was followed. Each trial consisted of two 200 ms noise bursts and the task was to decide which interval contained the target signal. All stimuli were gated with 10 ms ramps to avoid spectral splatter. Maskers were presented always at 70 dB SPL. Signal level began at easily detectable levels and was varied adaptively according to a 2-down 1-up rule, tracking the 70.7% correct point on the psychometric function. The signal level was varied in steps of 8 dB until the first two reversals, 4 dB until the next two reversals and 2 dB after that. The run was terminated after ten reversals, and the threshold defined as the mean level in the last six reversals. Two repetitions of each condition was run and presentation order of the conditions was randomized.

Stimuli

Gaussian Noise was generated by digitally filtering a broadband Gaussian noise with a filter centered at 1 kHz with a bandwidth of 20 Hz.

Masker	$A(t)$	$\phi(t)$	Thresholds Run 1 dB SPL	Thresholds Run 2 dB SPL
Tone	Tone	Tone	56.67	55.67
Gaussian Noise	Noise	Noise	62.33	62.00
Multiplied Noise	Noise	Tone	63.00	68.33
Low-noise Noise	Tone	Noise	59.33	57.00

Figure 1: Results of Experiment 2

Multiplied noise was generated by multiplying a sinusoid at 1 kHz with a modulator. The modulator consisted of a low-pass Gaussian noise with a cutoff frequency of 10 Hz at an RMS value of -10 dB (relative to amplitude 1) to which a dc component of value 1 was added.

Low-noise noise was generated starting with a Gaussian noise signal with a rectangular power spectrum. The following steps were iterated ten times: the envelope of the noise was calculated, representing the absolute value of the analytic signal, and the time waveform was divided by this envelope on a sample-by-sample basis and then restricted to its original bandwidth of 20 Hz by zeroing the corresponding components in the power spectrum.

Results

The table in figure 1 shows the threshold values and average thresholds are plotted in figure 5. Thresholds for Gaussian and multiplied noise maskers are significantly greater than the thresholds for tone and low-noise noise maskers.

Conclusions

We noticed that the threshold for multiplied noise masker was significantly greater than the threshold for low-noise noise masker, and these threshold values are in the same range as the thresholds for Gaussian noise and tone maskers respectively. Multiplied noise has an amplitude function $A(t)$ which is noise-like while low-noise noise has a phase function $\phi(t)$ which is noise-like. Since multiplied noise produces a higher threshold, we can conclude that the asymmetry effect in simultaneous masking is due to the amplitude term $A(t)$.

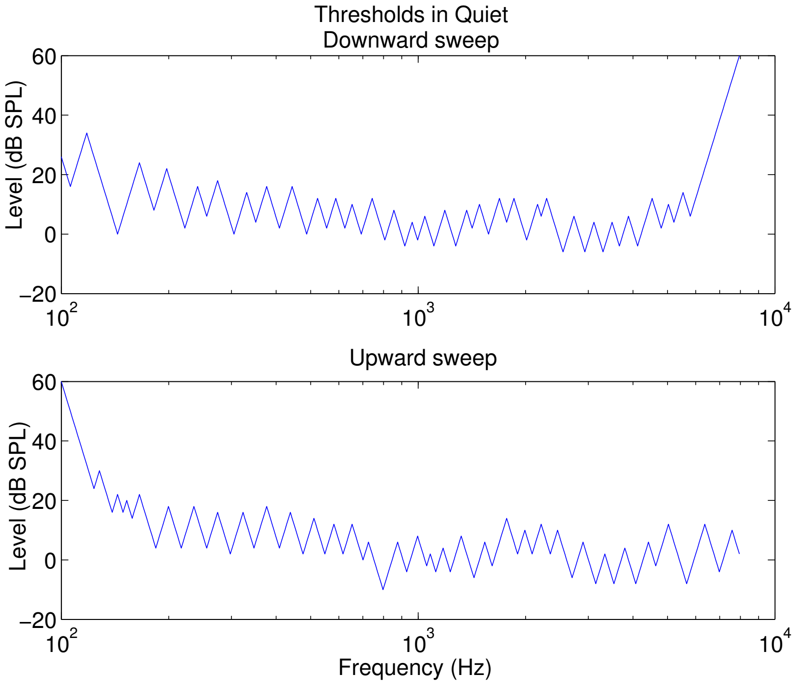

Figure 2: Absolute hearing threshold in quiet - Bekesy Method

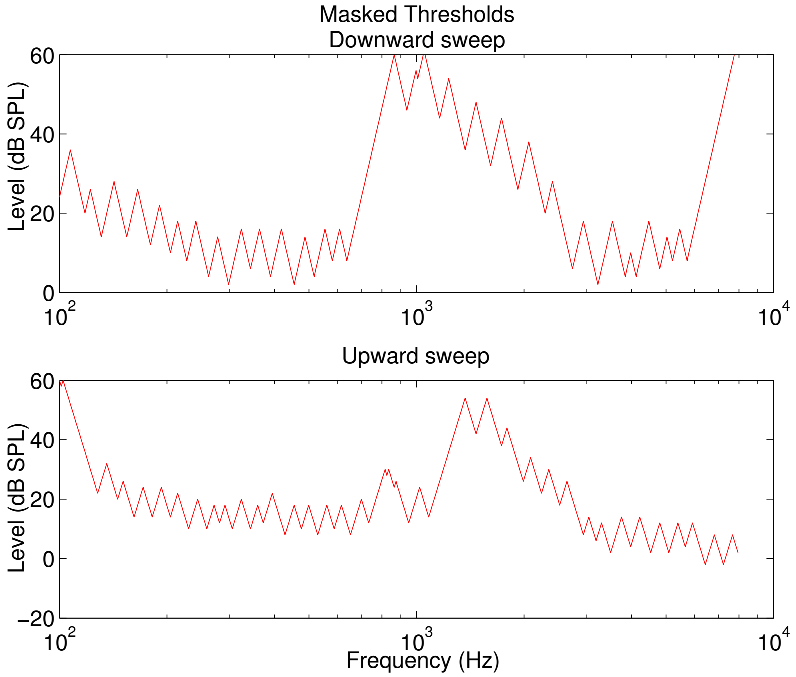

Figure 3: Masked threshold - Bekesy Method. Masker is a narrowband noise with a bandwidth from 950 Hz to 1050 Hz and spectrum level of 70 dB SPL

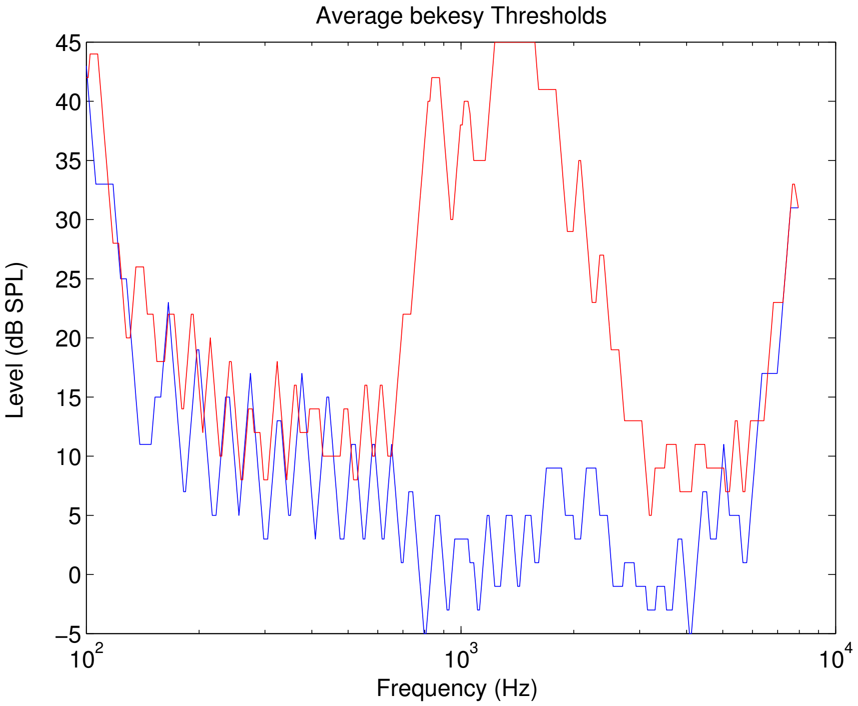

Figure 4: Average Bekesy Thresholds - Masked (red) and Quiet (blue)

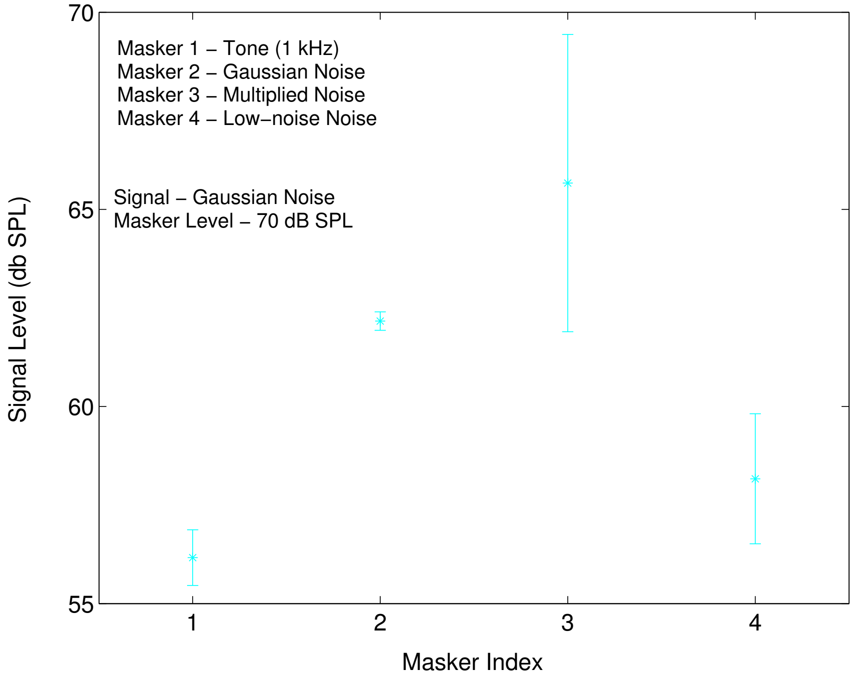

Figure 5: Masked thresholds with different maskers