Final Exam Study Guide:
15-322 Introduction to Computer Music

This document is intended to help you identify and master the main concepts of 15-322, which is also what we intend to test with the final. We make no claims of completeness. Additional lists of main concepts and what you should know can be found in the slides from the Unit 15.

Under what assumptions does sampling capture all of the information in a signal?

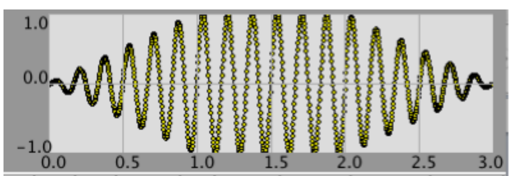
What determines the range of frequencies you can capture?

What determines the signal-to-noise ratio of the captured signal?

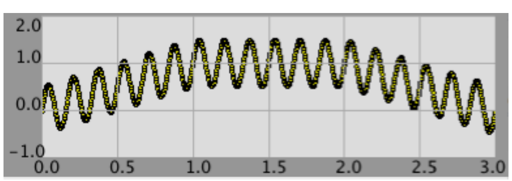
You should know how to represent sinusoids of varying amplitudes and frequencies in a spectrum (frequency domain) plot.

You should know how to interpret a filter frequency response plot, e.g. draw frequency responses for low-pass, band-pass, and high-pass filters. Show the effect of applying a given filter to a given signal using frequency domain representations.

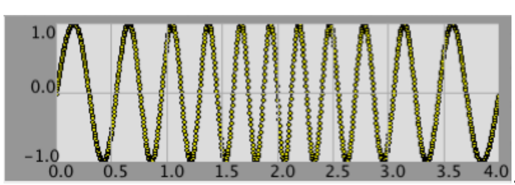
You should know how to write Nyquist expressions to achieve a given control, e.g. write expressions for this:



or this:



or this



.

If you wanted to create reverberation to match that of Carnegie Hall using convolution, what would you do? Describe what measurement you would make in the hall, what sound source and where, where to put microphones, what is the recording called, how do you use it to simulate reverberation? (We're not looking for detailed audio engineering, but conceptually, how does this work?)

What are the main components of Schroeder-style reverb?

What is RT60? (relating to reverberation)

Of the two main types of reverb, Schroeder-style and Convolution, which gives the most accurate model of a room? Which gives the most parametric control? Which is computationally the most efficient?

You should know about amplitude, power, and intensity. What change in power corresponds to 10dB? What change in amplitude corresponds to 10dB? What change in amplitude corresponds to 6dB? (We did not talk about intensity much in class – intensity is power per unit of area, so at a fixed distance from a speaker, intensity is proportional to power.)

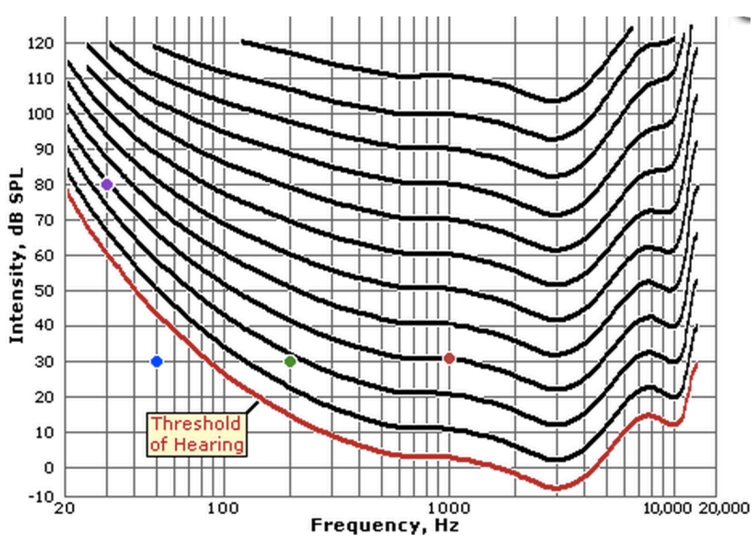
How does intensity decrease with distance (assuming no reflections or absorption)? How does amplitude decrease with distance?

What is the speed of sound?

What is the smallest difference in amplitude we can hear?

What is the smallest difference in frequency we can hear?

Explain the Fletcher Munson graph. How much louder is the threshold of hearing at 100 Hz than at 1000 Hz? If you change the frequency of a sinusoid at 100Hz and 70 dB SPL to 1000Hz, how many dB quieter should you make the sound to give the same perceptual loudness?



What is constant power panning? Why would you use constant power panning? With constant power panning, if a sound has an amplitude of 1 when panned full left, what amplitude do you send to both speakers when panning to the center?

What are the three types of data compression? Define these terms and give an example of each.

What is frequency masking? Explain the roll of frequency masking in MP3 music compression.

You should be familiar with the steps of MP3 encoding. E.g. what is the purpose of Huffman Coding in MP3?[1]

Why does MP3 encode music using a frequency-based representation rather than a time-based one?

Suppose you are concerned about quantization noise using an 8-bit DAC. What is the improvement in S/N ratio if you upgrade to a 12-bit DAC (assume perfect, noise-free analog circuits)?

Why do many telephone systems use (8-bit) µLaw rather than 8-bit PCM? Explain an advantage of µLaw over PCM and an advantage of PCM over µLaw.

Roughly how many data bits are on an audio CD (assume 80 minutes)? Hint: 16-bit stereo at 44100 Hz.

Use Nyquist patterns to generate these (infinite) sequences:

1 2 3 1 2 3 1 2 3 …
1 1 2 1 2 3 1 2 3 4 1 2 3 4 5   1 1 2 1 2 3 1 2 3 4 1 2 3 4 5   …
1 2 3 3 2 1 1 2 3 3 2 1 …

Write a Nyquist function to produce an FM tone with approximately N harmonics starting at F Hz and with amplitude A, where N, F, and A, are parameters.

What is computer accompaniment?

You should be familiar with the topics in the music understanding lecture. E.g. define the following terms: style recognition, onset detection, score alignment, music fingerprinting, music recommendation, music distribution, human computer music performance (of popular music).

Multiplication in the time domain is equivalent to what in the frequency domain?

Multiplication in the frequency domain is equivalent to what in the amplitude domain?

Multiplying by a constant in the time domain is equivalent to what in the frequency domain? (Careful: you might think the answer has to do with convolution, but this is a special case with an especially simple answer.)

---

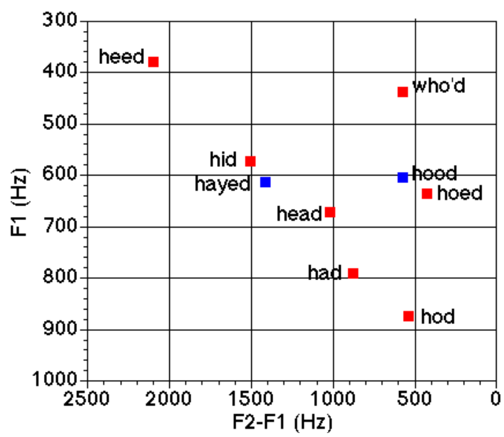
[1] Answer: to efficiently encode samples or groups of samples that have very non-uniform distributions (because smaller values are much more likely).

Adding two functions in the time domain is equivalent to what with the functions in the frequency domain?

How would you accomplish time stretching with granular synthesis?

Given a time-stretching method, how would you accomplish pitch shifting (without stretching or shrinking)?

How does sampling synthesis achieve control of duration and pitch? How does sampling avoid the problem that pitch-shifting a recording can sound unnatural if you pitch shift by a large amount?



What are the formant frequencies for the “oo (who)” and “ee” vowels?

You should understand the components of physical models such as the clarinet and bowed string models discussed in class. E.g. identify the waveguide, lumped filter, and non-linear elements in the clarinet and string models. What do these correspond to in corresponding acoustic instruments?

How does Karplus Strong synthesis work? Describe the elements and what they do? E.g. how is pitch determined, and what causes the string to decay?

You should understand the concepts of spectral interpolation synthesis, e.g. how are spectra selected? Why are there two tables in the synthesizer? What goes into the tables?

What are HRTFs?

Why are HRTFs usually used with headphones?

John Chowning made the point that things do not necessarily seem more distant just because you lower the amplitude. What cues are available for distance?

---

1. In terms of what is sound intensity measured? Write the expression relating this measure with pressure.

2. In three sentences, summarize the functionalities of the outer ear, middle ear and the inner ear.

3. What is the difference between *reverberation* and *echo*?

4. If two sine tones are very close together in frequency, the total loudness perceived is less than the sum of the two loudnesses that would be heard from the tones played separately. Why does this happen?

5. Why is white noise called "white"?

6. What happens if a set of partials is modulated in frequency by a common temporal envelope?

7. What is the *missing-fundamental* phenomenon?

---

1. What are the two aspects of Lord Rayleigh's Duplex Theory of sound localization?

   a. Interaural Time Difference and Interaural Level Difference
   b. Interaural Time Difference and Interaural Frequency Difference
   c. Interaural Transfer Difference and Interaural Frequency Difference
   d. Interaural Transfer Difference and Interaural Elevation Difference

Answer: _______

2. Which of the following has the greatest effect in forming the Head Related Transfer Function (HRTF)?

   a. Eyes
   b. Ears
   c. Hair
   d. Mouth

Answer: _______

3. Which of the following is NOT a technique originally developed for voice synthesis?

   a. Linear Prediction Coding (LPC)
   b. VOSIM
   c. Formant Wave-Function Synthesis (FOF)
   d. Digital Pulse Code Modulation (DPCM)

Answer: _______

4. In Linear Panning between 2 speakers, a value α, between 0 and 1, is the portion of the sound coming from the left speaker, and (1- α) is the portion of the sound coming from the right speaker. What is the main problem with using Linear Panning to make a sound "move" from left to right (or vice versa)?

a. Mid-range frequencies are attenuated
b. It's electronically hard to implement
c. The power of the signal changes with α, so the perceived loudness of the sound varies with pan position
d. Instead of moving from left to right (or vice versa), the sound will seem to move from above to below (or vice versa)

Answer: _______

8. What is the hardware arrangement for sound delivery, given a goal of creating the best illusion of 3-D sound using the HRTF?

a. Headphones

b. Small speakers a few feet away from the ears, such as those of a desktop computer.

c. Large wall-mounted stereo speakers.

d. 5 or 6 surround / wraparound speakers, such as a Dolby 5.1 arrangement in a room.

Answer: _______

9. What defines the human range of hearing (typically), and about how large is it?

a. quiet whisper – loud shout, about 60 dB

b. softest perceivable sound – loudest painless (or softest painful) sound, 120 dB

c. eardrum resonance – eardrum rupture, about 200 dB

d. there is no limit to human hearing in either loudness or softness, the only reason some noises escape our perception is because they are below the noise threshold of the air around us.

11. There are five formants in the vowel sounds produced by the human voice, but the most important ones that human listeners use to understand one another are

a. The bottom two (in frequency)

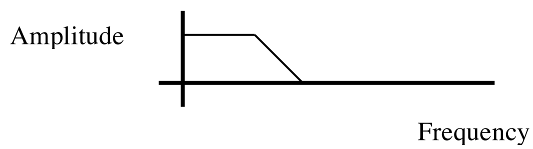
b. The top two (in frequency)

c. The second and third (in frequency, from lowest to highest)

d. The highest and lowest (in frequency).

Answer: ________

4. Here is the frequency response for a given filter (only magnitude is shown):



What kind of filter is this?

a. Low-pass
b. High-pass
c. All-pass
d. Forward-pass

Answer: ________

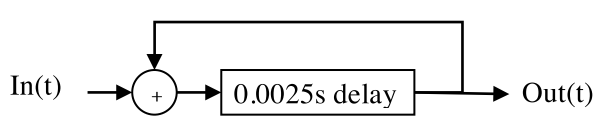
5. Which of the following phenomena is responsible for the behavior of a comb filter?

a. Low-pass filtering
b. High-pass filtering
c. Constructive interference
d. Superposition

Answer: ________

6. A circuit (see figure) is constructed such that its output is defined recursively by:

Out(t + 0.0025 seconds) = In(t) + 0.5 * Out(t).



This circuit behaves as a:

a. Low-pass filter with the -3 dB point at 250 Hz
b. Low-pass filter with the -3 dB point at 400 Hz
c. Comb filter strengthening integer multiples (harmonics) of 250 Hz
d. Comb filter strengthening integer multiples (harmonics) of 400 Hz

Answer: ________

7. From an original waveform a frequency spectrum is generated by means of a Short-Time Fourier Transform. Were a new waveform to be generated from that spectrum by some inverse operation of the Short-Time Fourier Transform, the new waveform would

a. Compare favorably with the original waveform
b. Be recognizable as a highly distorted version of the original waveform
c. Be completely unrecognizable noise
d. The Short-Term Fourier Transform is a mathematical "one-way" function, there is no way to invert it and go from spectra to waveforms.

Answer: ________

8. Which of the following is **not** a cue for judging how far away a sound was generated?

a. Diminished amplitude / diminished energy
b. Shifted phase
c. Loss of high-frequency components
d. Loss of low-amplitude frequencies below noise or hearing threshold.

Answer: ________

9. Which cue is **most** indicative of a sound being generated by a source in motion?

a. Increasing amplitude
b. Decreasing amplitude
c. Doppler-shifted frequencies
d. Frequency-variable phase shifts

Answer: ________

4. Which of the following will the sound generated by the Nyquist code

```
(play (fmosc c5 (mult (pwl 0.05 2000 1 1000 1) (osc c5))))
```

sound most like?

a. A click
b. Hitting a metallic pipe
c. A singer's voice
d. A brass instrument

Answer: _________________________