I547: Audio Music Processing
Homework 3

1. In the following examples suppose we are given $x(n)$ for $n = 0, \ldots, N-1$ where $N = 1024$ and suppose X is the finite Fourier transform of x, $x \overset{FFT}{\leftrightarrow} X$. Give the modulus and argument of the complex numbers $X(0) \ldots X(N/2)$ when

   (a) $x(n) = \cos(2\pi 5n/N + 1)$
   $X(5) = e^i$ (Mod(X(5)) = 1, Arg(X(5)) = 1)
   all others 0

   (b) $x(n) = 4\sin(2\pi 5n/N) + 3\sin(2\pi 6n/N) + 2\cos(2\pi 8n/N)$

   $$\begin{aligned} X(5) &= 4 * e^{-i\pi/2} \\ X(6) &= 3 * e^{-i\pi/2} \\ X(8) &= 2 * e^{-i0} \end{aligned}$$

   all others 0

   (c) $x(n) = 4\sin(2\pi 5n/N) + \cos(2\pi 5n/N)$
   Mod(X(5)) = $\sqrt(17)$ Arg(X(5)) = $\tan^{-1}(-4)$
   all others 0

   (d) $x(n) = 3\cos(2\pi 8n/N - 1) + 3\cos(2\pi 8n/N + 1)$
   Mod(X(8)) = $6\cos(1)$ Arg(X(8)) = 0
   all others 0

   No coding is necessary or desirable.

2. Download the Bass Oboe data from the class website and examine using the program spectrum_movie.r.

   (a) Describe the effect of varying the FFT length in this program.

   (b) Window a representative (non-silence) piece of the audio data before computing the FFT using a Hann (raised cosine) window. Compare the spectra with and without this window and submit your plots.

   (c) Identify *through visual inspection alone* the locations (eg. frames 10 - 20) and musical pitches (eg. "middle c") of the first 5 notes. By "frame" I mean an individual analysis segment or Fourier transform.

   F below Middle C, Ab below middle C (repeated), F below Middle C, Eb below middle C, Db below middle C,

3. Using approximately 1 second of sound starting 64000 samples into the bass oboe data, create a Hann-windowed fft. Reconstruct the sound as a sum of sine waves using

   (a) all N/2 sines — okay to ignore the DC (non-oscillating) component

   (b) the N/4 sine waves having the biggest amplitude

   (c) the N/4 sine waves having the smallest amplitude

   (d) the N/10 sine waves having the biggest amplitude

   (e) the N/100 sine waves having the biggest amplitude

   Listen to each sound and describe what you hear.

   You should reconstruct like we did using the sawtooth.r program but only using the the sine waves whose amplitudes are bigger or smaller than some threshold. To find the appropriate threshold take the collection of moduli and sort them, choosing the appropriate value. For instance, for part b), you would only use the values whose amplitudes are greater than the N/4th sorted amplitude (sorted in decreasing order).

4. Using the idea of the timbre_copy.r example, create a 16 sinusoid model of bass oboe note. You might want to "page through" the sound using spectrum_movie.r to find a suitable location.

   (a) Using your timbre model, recreate "Amazing Grace" starting on $C = 4402^{-21/12}$.

   (b) Using your timbre model, recreate the expressive octave leap of the previous assignment.

   Submit your code and be prepared to play your examples in class.

   If $a_1, \ldots, a_{16}$ are the amplitudes for harmonics 1, ..., 16, as taken from the sample spectrum. create the desired frequency function, f, as before. You'll have different frequency functions for the two parts of the problem. The frequency function should be sampled at the same rate as the audio you will create. Then create the phase function $\phi = \text{cumsum}(f)/SR$. Now the desired audio will be

$$g(t) = \sum_{k=1}^{16} a_k \sin(2\pi k\phi)$$

   It should be easy to verify if the solutions are correct by listening.

5. From a single note on the "octaves" data compute two representative timbre models from two different spectra, as you did in the previous problem. Call these a^1, a^2 where $a^1 = (a_1^1, \ldots, a_{16}^1)$, and similarly for a^2. Create a timbral vibrato by oscillating back and forth smoothly between the two timbres defined by a^1 and a^2 for a given frequency.

   Suppose we want to create sound a frequency f with a timbral vibrato at rate 5 Hz. Let the mixing function be

$$m(t) = \frac{1 + \sin(2\pi 5t)}{2}$$

   which oscillates between 0 and 1 at a rate of 5 Hz. Then get the timbral vibrato as

$$g = \sum_{k=1}^{1} (a_k^1 m(t) + a_k^2 (1 - m(t))) \sin(2\pi k f t)$$

   At the extreme points of the sinusoidal oscillation we get the two original timbres defined by a^1 and a^2, interpolating smoothly between these.