I547: Audio Music Processing
Homework 5

1. For this problem you may wish to pair up with someone who knows musical notation, though only a little will be necessary. Consider the spectrogram at

   http://music.informatics.indiana.edu/~craphael/acm/mozart_4tet_image.bmp

   in which the solo part is highlighted in yellow. Give your best approximation of appropriate music notation for the solo part. The first note is a quarter note A♭ above the treble staff.

2. This problem uses the 48 KHz "glunker_stew.wav" audio file which is now on the class web page.

   (a) "Robotize" this audio file using a window size of $N = 1024$ and a hop size of $H = N/8$ and listen to the file. Compute the pitch, p_0 in Hz at which you expect the monotone pitch to be sounding and verify that your calculation is correct by listening to a sine tone at that pitch. Take the fft of the robotized audio data and plot out the modulus of the fft over the range $p_0 \pm 100$ Hz. Plot the modulus of the fft of the original data over the same range on the same plot.

   (b) "Whisperize" the data using the same parameters. Produce a plot containing the modulus of the ffts for both the original and modified data.

3. Compute the STFT of the bass oboe data using a window length of $N = 1024$ and a hop size of $H = N/4$. Randomly permute the "columns" (Fourier transforms) of the STFT and regenerate sound using the phase vocoder.

4. Create a version of the "glunker_stew.wav" audio file in which the pitch is twice as high (an octave up), but the rate of speech is the same. Do this by first creating an audio file that is half as long but twice as high and then stretch the file in time.

5. Create a compressed version of the "glunker_stew.wav" version by sythesizing only using the STFT coefficients with the highest moduli. Trying using the top 10%, 5%, 2.5%, and 1%. Submit your code.