I547: Audio Music Processing
Homework 4

1. Consider the first (y) and second (z) "differences" of a sequence x:

$$\begin{aligned} y_j &= x_j - x_{j-1} \\ z_j &= y_j - y_{j-1} \end{aligned}$$

   for $j = 1, 2, \ldots$. These are the discrete analogs of the first and second derivatives.

   (a) Write $y = a * x$ and $z = b * x$ where a and b are short filters.

   (b) Compute the frequency responses of these filters by padding them with 0's and taking the FFT. Which frequencies of the signal x are suppressed and amplified by using these filters? Justify your response using plots of the frequency response of the filters.

   Sol: $y_j = 1x_j + -1 * x_{j-1}$ *so* $y = a*x$ *where* $a = (a_0, a_1) = (1, -1)$. *By similar reasoning* $z_j = 1x_j - 2x_{j-1} + 1x_{j-2}$ *so* $z = a * x$ *where* $a = (a_0, a_1, a_2) = (1, -2, 1)$. *When looking at the frequency response these filters amplify higher frequencies.*

2. We know that the FFT transforms convolution into multiplication. The reverse relationship:

$$x \cdot a \overset{\text{FFT}}{\leftrightarrow} X * A$$

   also holds where $x \overset{\text{FFT}}{\leftrightarrow} X$ and $a \overset{\text{FFT}}{\leftrightarrow} A$ and $\cdot$ denotes pointwise multiplication.

   (a) What is the FFT of an N-point Hann (raised cosine) window for *all* values of the frequency variable. Explain why this is consistent with your understanding of the FFT.

   (b) Suppose we have a sine wave that oscillates an integral number of times per N points. What is the effect in the frequency domain of multiplying the sine by the Hann window function before taking the FFT (i.e. the usual windowing procedure)?

   Sol: Since the window function is explicitly given as a sum of cosines that oscillate integral numbers of times per FFT length, we can simply "read off" the FFT. Writing the Hann window as

$$w(j) = 1/2 - 1/2cos(2\pi j/N)$$

   In particular, the Hann window is a sum of a constant (the cosine that oscillates 0 times, and the cosine that oscillates onces (with phase $-\pi$*). Remembering the factor of 2 difference in scaling, and using the complex conjugate symmetry of real FFTs gives between the 0,N/2 coefficients and the other coefficients gives*

$$W(n) = N/2 \times \begin{cases} 1 & n = 0 \\ -1/2 & n = 1 \\ -1/2 & n = N-1 \end{cases}$$

3. Initialize y_1 and y_2 and define y_j for $j = 3, 4, \ldots$ by

$$y_j = 1.5y_{j-1} - .8y_{j-2}$$

   The resulting y can be written as $y_j = ar^j \sin(j\theta + \phi)$. Solve for r and θ.

   Recalling that $y_j = 2r\cos(\theta)y_{j-1} - r^2 y_{j-2}$ *generates a sine wave with angular frequency* θ, $y_j = ar^j \sin(\theta j + \phi)$, *we just need to equate coefficients. This gives* $r = \sqrt{.8}$ *and* $\theta = \cos^{-1}(1.5/(2\sqrt{.8})$

4. Write an R program to filter the bass oboe data using a filter that "passes" only frequencies between 1000 and 3000 Hz. Listen to the result.

   *Since the bass oboe data is sampled at 8kHz, 1000 Hz corresponds to bin 1000*1024/8000 = 128, while 3000 corresponds to 384. Using this just proceed as the low pass filter done in class.*

5. Consider the random walk model

$$y[i] = y[i-1]/2 + \epsilon[i]$$

where $\epsilon = (\epsilon[1]\epsilon[2]\ldots)$ is a vector of independent 0-mean random numbers (eg rnorm(n)).

(a) Write this equation as

$$a * y = \epsilon$$

for some filter a where $*$ denotes convolution.

$a = (a_0, a_1) = (1, -1/2)$

(b) Express the FFT of y in terms of the transforms of a and ϵ

$Y = E/A$ *where* Y, E, A *are the FFTs of* y, ϵ, a.

(c) Describe the spectral qualities of y.

$A(n) = 1 - e^{-2\pi i n/N} bydirectcalculation.$ *If we look at* $1/Mod(A)$ *we see that the frequency response decreases from 0 all the way to the Nyquist frequency.*

(d) Create y by setting the first value to 0 and running the eqaution. Listen to the result, and compute the FFT.

6. Imagine a simple room that produces an echo. The echo comes $D = .125$ secs after the original sound, but decreases in amplitude by a factor of $F = .8$. This echo is now *part* of the sound in the room, so D seconds later we will hear another version of the original echo, decreased in amplitude by another factor of F. This process will continue indefinitely, though the echos eventually become inaudible as their amplitudes decrease. Explicitly model this process by an AR filter with input signal x and output signal y, for audio data sampled at 8kHz. Run this on the bass oboe data to see the effect.

$$y_n = y_{n-1000} + x_n$$