Midterm Solutions
November 6, 2017

- The total number of points available on this midterm is 100.
- In solving the problems on this midterm, clearly show your work and clearly indicate your answer. For all problems on this test, answers should be expressed in the simplest form possible.
- There are a total of 4 problems, each with several subparts. Please verify your exam to make sure you have the entire exam (total of 4 problems, total length 4 pages).
- $\log_{10}(2) \approx 0.3$
- An octave corresponds to a factor of 2 change in frequency. For example, given a tone at 200 Hertz, a tone an octave above this tone would have a frequency of 400 Hz. Similarly, a tone an octave below 200 Hz is at 100 Hz.

.

1. Discrete Time Fourier Transform (25 points)

Note: In all parts of this problem, sequences that are a function of n are defined only at integer values of n.

Let $x(n)$ be a signal with DTFT $X(\omega)$. In addition, let $v(n)$ be as follows:

$$v(n) = \frac{e^{-j\pi n}x(n) + x(n)}{2}$$

a) (8 points) Let $V(\omega)$ denote the DTFT of $v(n)$. Express $V(\omega)$ in terms of $X(\omega)$.

b) (9 points) Now consider the sequence $y(n) = v(2n)$. For example, if $v(n)$ is a sequence of the form

$\underline{a}$ b c d e f g. . .

then $y(n) = v(2n)$ would be

$\underline{a}$ c e g . . .

Find $Y(\omega)$, the DTFT of $y(n)$, expressed in relation to $V(\omega)$, using the most simplified form of this expression as possible. As noted in part a), $V(\omega)$ denotes the DTFT of $v(n)$.

c) (8 points) Let $g(n) = x(2n)$ and let $G(\omega)$ denote the DTFT of $g(n)$. Express $G(\omega)$ in terms of $X(\omega)$.

Answer

a) The DTFT of x(n) is denoted by $X(\omega)$.

According to the DTFT frequency shift theorem, the DTFT of $e^{-j\pi n}x(n) = X(\omega + \pi)$. (The DTFT shift theorem is very fast to derive.)

Using result of the DTFT shift theorem combined with the linearity property gives

$$V(\omega) = \frac{X(\omega + \pi) + X(\omega)}{2}$$

b) As provided in the problem statement, y(n) = v(2n). First consider the DTFT of y(n):

$$Y(\omega) = \sum_{n=-\infty}^{\infty} y(n)e^{-j\omega n} = \sum_{n=-\infty}^{\infty} v(2n)e^{-j\omega n}$$

Now consider $V(\omega)$, the DTFT of the original $v(n)$:

$$V(\omega) = \sum_{n=-\infty}^{\infty} v(n)e^{-j\omega n}$$

Since we know from the definition of $v(n)$ that $v(n)$ is nonzero only for n even, the above summation is equivalent to

$$V(\omega) = \sum_{n=-\infty}^{\infty} v(2n)e^{-j\omega 2n}$$

Thus, comparing the above equation to:

$$Y(\omega) = \sum_{n=-\infty}^{\infty} v(2n)e^{-j\omega n}$$

means that

$$Y(\omega) = V\left(\frac{\omega}{2}\right)$$

c) x(2n) is a sequence which contains only the even indexed samples of x(n), shifted to locations (relative to their location in x(n)) $n/2$, where n (with reference to the original sequence x(n)) is even. v(n) is a sequence in which the even indexed samples are equal to the corresponding samples of x(n). The odd indexed samples of v(n) are equal to zero. v(2n) is a sequence which contains only the values that were at the even indexed samples of v(n), though shifted in location (from n to n/2, where n is even). This means that x(2n) = v(2n). Therefore, since g(n) =x(2n) from the defintion in the problem statement, it must also be true that g(n)= v(2n). This means that G(ω) = V(ω /2). Applying this equation in the expression found in part (a) gives:

$$G(\omega) = \frac{X\left(\frac{\omega}{2} + \pi\right) + X(\frac{\omega}{2})}{2}$$

Note that the answer to this question is *not* simply that G(ω) = X(ω /2), as this would ignore the fact that the g(n) contains only the values that were previously located at the *even* indexed location of x(n). The values of x(n) at *odd* index locations have been removed. Those odd indexed values, however, did contribute to X(ω), the DTFT of x(n).

2. Sampling (22 points)

Consider a continuous time signal that is then sampled at a rate of 20.48 kHz. A segment of length 1024 samples is selected and the 1024-point DFT is computed.

a) (5 points) What is the time duration of the segment?

b) (5 points) What is the frequency resolution of the DFT, where for this problem frequency resolution is defined as the spacing, in Hertz, between two adjacent DFT bins?

c) (12 points) Now consider a continunous time signal signal that consists solely of a pure sinusoid at 15 KHz. Suppose that the sinusoidal signal is then sampled at 20.48 KHz. The sampling process does not include any filtering. Sampling in this case consists simply of measuring the exact value of the input signal at intervals of T seconds, where T is the inverse of the sampling frequency. Then, the sampled signal is reconverted to analog form by passing it through an ideal D/A converter. What frequency, in Hertz, will be output from the D/A converter?

Answer

a) The time duration spanned by the segment is 1024/20.48kHz = 50 msec

b) The frequency resolution of the DFT is Fs/N = 20 Hz.

c) A sinuusoid at 15 KHz corresponds to a transform that has a pair of delta functions at positive and negative 15 KHz. Sampling in the time domain corresponds to multiplication by a train of delta functions spaced at T =1/20480 seconds. In the frequency domain, this corresponds to convolution with a train of delta functions with spacing 20.48 KHz. This will replicate the transform at intervals equal to the sampling frequency. Thus, the frequency component originally present at 15 KHz will be replicated at (and therefore aliased to) −5.48 KHz, and the frequency component originally present at −15 KHz will be aliased to 5.48 KHz. There will be other copies as well (e.g., at positive and negative 35.48 KHz, etc.), but those other copies will be eliminated by the low pass filter used in reconstruction, which will have a cutoff at 10.24 KHz. Thus, the output of the D/A converter will be a sinusoid at 5.48 KHz.

3. Linear Predictive Coding and Autocorrelation (28 points)

Consider a vowel sound sampled at 8 KHz. A segment, denoted $s(n)$, is extracted using a Hamming window. The magnitude of the DTFT of this segment is shown in Figure 1 below (assume that in the figure all the harmonics shown are equally spaced):

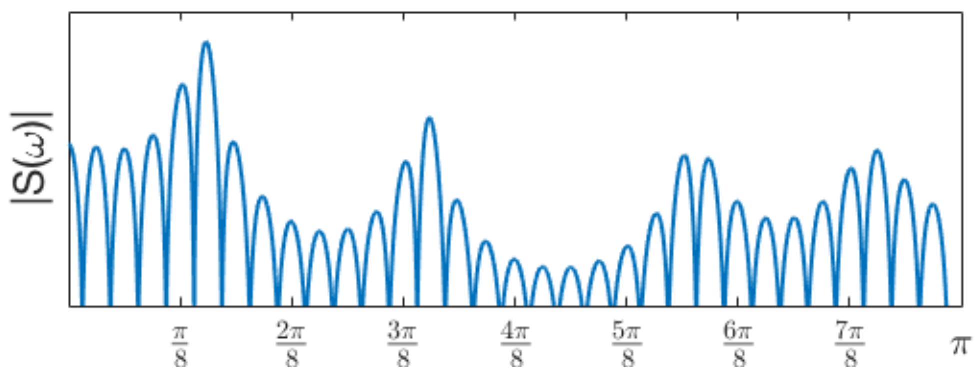

Figure 1. Spectrum of vowel segment s(n)

a) (8 points) Given the lobes shown in Figure 1 above, and the fact that this spectrum was obtained by taking the DTFT of a signal that had been subject to a Hamming window, what is the minimum window length that was needed to ensure that the lobes associated with the harmonics shown in the figure (which for each harmonic correspond to the result of convolving the DTFT of the window function with the delta function associated with that harmonic) remain non-overlapping? You should specify the window length in samples, not seconds.

For parts b) through d) of this problem, consider the following:

The time domain signal representing the vowel segment is passed through the filter denoted “System 1” as shown in Figure 2 below, and is then down-sampled to a sampling frequency of 1600 Hz. System 1 is a filter. The downsampler downsamples by a factor of 5. This means that it simply keeps every fifth sample and discards the intervening samples. For example, if the input to the downsampler is a sequence

1 2 3 4 5 6 7 8 9 10 11 12 . . .

the output is a sequence

1 6 11 . . .

While System 1 is a filter, there is no additional filter included in the box in Figure 2 labeled "Downsampler (Decimator)". The downsampled signal is then passed through an LPC predictor characterized by the following H(z).

$$\mathrm{H(z)} = \frac{G}{1 - \sum_{k=1}^{p} a_k z^{-k}}$$

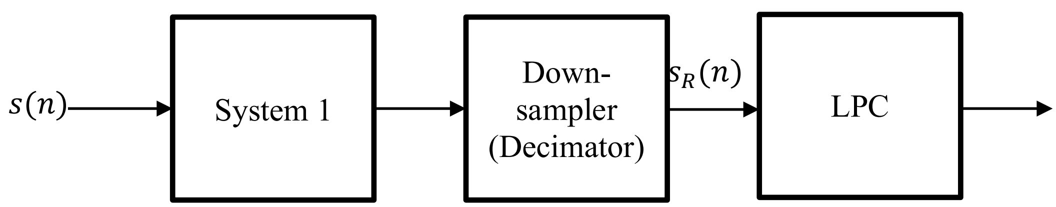

Figure 2.

b) (5 points) What filter should be used in System 1 to ensure that there is no aliasing arising due to the subsequent downsampling step in which only every fifth sample is kept? You should assume that the filter is ideal (i.e., the frequency response can contain vertical edges). Your answer should specify the type of filter as well as the specific attributes.

c) (10 points) Using the second-order linear predictive analysis, find the predication coefficients a_k, assuming that $R(0) = 6, R(1) = -4, R(2) = 1$.

d) (5 points) Using the result from part c), and assuming a sampling frequency of 1600 Hz, find the formant frequenc(y)(ies) in **Hertz** associated with the second order model associated with the R(0), R(1), and R(2) values provided in part c).

Answer

a) The spectrum of the signal shows that the 4th harmonic appears at $\frac{\pi}{8}$, meaning that the fundamental frequency (and therefore the spacing between the lobes shown in the figure) is $F_0 = \frac{1}{4} \times \frac{\pi}{8} = \frac{\pi}{32}$. To resolve the harmonics and ensure the mainlobes arising from the convolution of the transform of the window with the source function are distinct, the width of the main lobe of the DTFT of the window should be smaller than the fundamental frequency. Because the width of the main lobe of the transform of the Hamming window is $8\pi/L$ (where L is the length of the window in samples),

$$\frac{8\,\pi}{L} \leq \frac{\pi}{32} \Rightarrow L \geq 256\ (samples)$$

This problem did not require specifying the length of the window in seconds, but, that computation is as follows: Because the sampling frequency is 8 kHz, this window length corresponds to 32 msec.

b) System1 should be a low-pass filter with a cut-off frequency of 800 Hz. This is because the post-downsampler sampling rate will be 8000/5 = 1600 Hz. Thus, to avoid aliasing, it is necessary to ensure that the input to the downsampler does not have any frequencies higher than 800 Hz.

c) This can be solved by using the normal equations and left multiplying by the inverse of R:

$$a = R^{-1}r = \begin{bmatrix} 6 & -4 \\ -4 & 6 \end{bmatrix}^{-1} \begin{bmatrix} -4 \\ 1 \end{bmatrix} = \frac{1}{6^2 - (-4)^2} \begin{bmatrix} 6 & 4 \\ 4 & 6 \end{bmatrix} \begin{bmatrix} -4 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -0.5 \end{bmatrix}$$

Thus, $a_1 = -1, a_2 = -0.5$

d) The poles of the vocal tract transfer function are located at the z values where $1 - \sum_{k=1}^{2} a_k z^{-k} = 0$. That is, the poles are located at the roots of the following quadratic equation:

$$z^2 + z + 0.5 = 0$$

This results in a pair of complex conjugate poles at $z = -0.5 \pm 0.5j$, and hence there is one formant. The formant frequency is at $\omega = \frac{3\pi}{4}$, which, converted into Hertz, is $\frac{3\pi}{4} \times \frac{1600\ Hz}{2\pi} = 600\ Hz$

4. Windowing (25 points)

Consider a discrete-time triangular window $w_T(n)$ specified as follows, where N is even:
:

$$w_T(n) = \begin{cases} \dfrac{n}{N/2}, & n = 0, \dots, \dfrac{N}{2} \\ \dfrac{N-n}{N/2}, & n = \dfrac{N}{2} + 1, \dots, N-1 \end{cases}$$

Note that since N is even, the number of nonzero values in $w_T(n)$ will be $N-1$, which is odd.

a) (15 points) Find $W_T(\omega)$, the DTFT of the window $w_T(n)$. Note that this problem is not asking you to find the absolute value of $W_T(\omega)$. It is asking you to find $W_T(\omega)$, which may be complex.

b) (10 points) Assume a large N and a small ω. Determine (with a full justification of your answer) the roll-off rate of the sidelobes of the DTFT triangular window (again assuming you are examining the rolloff in where ω is small, and also assuming a large N). Your answer should be expressed in terms of in dB per octave. Your answer should be a number (e.g., X dB per octave, where X is a number), and should not contain any unevaluated log functions.

Answer

a) As provided in the problem statement, the number of nonzero values in the window will be $N-1$. Note that $w_T(0) = 0$. The peak value of 1 will occur at N/2. The other values will be symmetric about this location; i.e.:

$$w_T(1) = w_T(N-1) = 2/N$$
$$w_T(2) = w_T(N-2) = 4/N$$
$$w_T(3) = w_T(N-3) = 6/N$$
$$\text{etc.}$$

The triangular window can be shifted by 1 to place the first nonzero value at the origin:

$$y(n) = w_T(n+1)$$

$y(n)$ is a triangular window that has nonzero values in the range [0, N-2] and a maximum at N/2 -1. This can be expressed as the result of a convolution of two rectangular windows of length N/2 that have their first nonzero values value at the origin. We know that convolving two sequences of length $N/2$ will result in a sequence of length of $\frac{N}{2}+\frac{N}{2}-1=N-1$.That is,

$$y(n) = w_R(n) * w_R(n)$$

where

$$w_R(n) = \begin{cases} \frac{1}{\sqrt{N/2}}, 0 \le n \le \frac{N}{2}-1 \\ 0, \text{otherwise} \end{cases}$$

The value of $\frac{1}{\sqrt{N/2}}$ arises because at the point of maximum overlap (i.e., when all N/2 values in the rect functions are overlapping), the result of the convolution will be

$$\frac{N}{2}\frac{1}{\sqrt{\frac{N}{2}}}\frac{1}{\sqrt{\frac{N}{2}}} = 1$$

The DTFT of $w_R(n)$ can be found as

$$W_R(\omega) = \sum_{n=-\infty}^{\infty} w_R(n)e^{-j\omega n} = \sum_{n=0}^{\frac{N}{2}-1} \frac{1}{\sqrt{N/2}}e^{-j\omega n} = \frac{1}{\sqrt{N/2}}\frac{1-e^{-j\omega\frac{N}{2}}}{1-e^{-j\omega}}$$
$$= \frac{1}{\sqrt{N/2}}\frac{e^{-j\omega N/4}(e^{j\omega N/4}-e^{-j\omega N/4})}{e^{-j\omega/2}(e^{j\omega/2}-e^{-j\omega/2})} = \frac{e^{j\omega\left(\frac{1}{2}-\frac{N}{4}\right)}}{\sqrt{N/2}}\frac{\sin\left(\frac{N\omega}{4}\right)}{\sin\left(\frac{\omega}{2}\right)}$$

Since convolution in the time domain corresponds to multiplication in the frequency domain, $Y(\omega)$, the DTFT of $y(n)$ can be written as:

$$Y(\omega) = W_R(\omega) \times W_R(\omega)$$

Therefore:

$$Y(\omega) = \frac{e^{j\omega\left(1-\frac{N}{2}\right)}}{N/2}\left(\frac{\sin\left(\frac{N\omega}{4}\right)}{\sin\left(\frac{\omega}{2}\right)}\right)^2$$

Then we can apply the DTFT shift theorem to find the DTFT of the window that is not centered at the origin, i.e., that spans the range $0 \le n \le N-1$.

$$W_T(\omega) = e^{-j\omega 1} Y(\omega) = e^{-j\omega\frac{N}{2}} \frac{1}{N/2} \left(\frac{\sin\left(\frac{N\omega}{4}\right)}{\sin\left(\frac{\omega}{2}\right)} \right)^2, -\pi \leq \omega \leq \pi$$

b) For small ω and large N, the effects of aliasing will be negligible. Since $\sin(x) \approx x$ for small x, the envelope of the DTFT will be inversely proportional to ω^2. Thus, when ω is doubled (which is the change in frequency corresponding to one octave), the magnitude of the DTFT will be reduced by a factor of 4. This means that, in linear terms, the rolloff occurs at rate of a factor of 4 per octave. In dB terms, this corresponds to 20 $\log_{10}$ (0.25). Since, as stated at the top of this exam $\log_{10}(2) \approx 0.3$, then:

$\log_{10}(0.5) \approx -0.3$

and therefore

$\log_{10}(0.25) = \log_{10}(0.5/2) \approx -0.6$

and

$20 \log_{10}(0.25) \approx -12$

Thus, the rolloff of the sidelobes occurs at 12 dB per octave.