University of California
College of Engineering
Department of Electrical Engineering
and Computer Sciences

E. Alon, B. Ayazifar, C. Tomlin **Thurs., Sep. 24, 2015**
G. Ranade **11:00-12:30pm**

EECS 16B: FALL 2015—MIDTERM 1

Important notes: Please read every question carefully and completely – the setup may or may not be the same as what you have seen before. Also, be sure to show your work since that is the only way we can potentially give you partial credit.

NAME	Last First

SID	

Login	

Problem 1: _____/ 13

Problem 2: _____/ 19

Problem 3: _____/ 23

Problem 4: _____/ 10

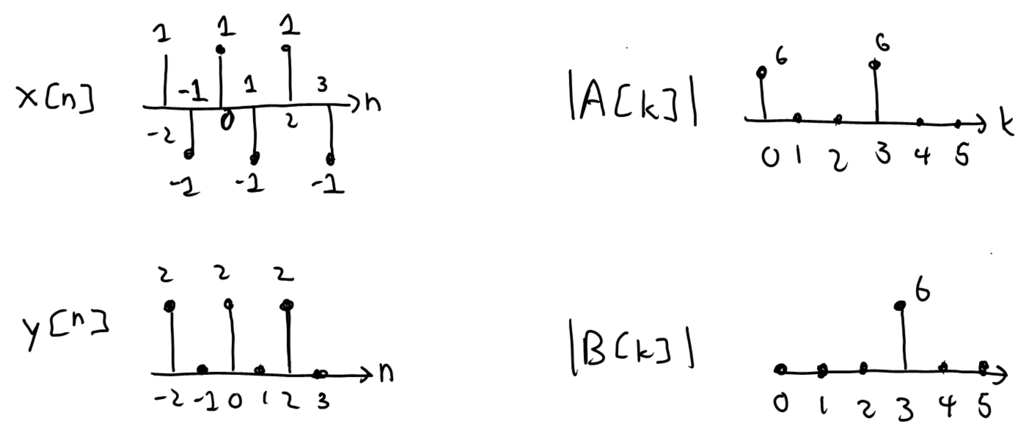
Total: _____/ 65

PROBLEM 1. Signal Analysis with DFT (13 points)

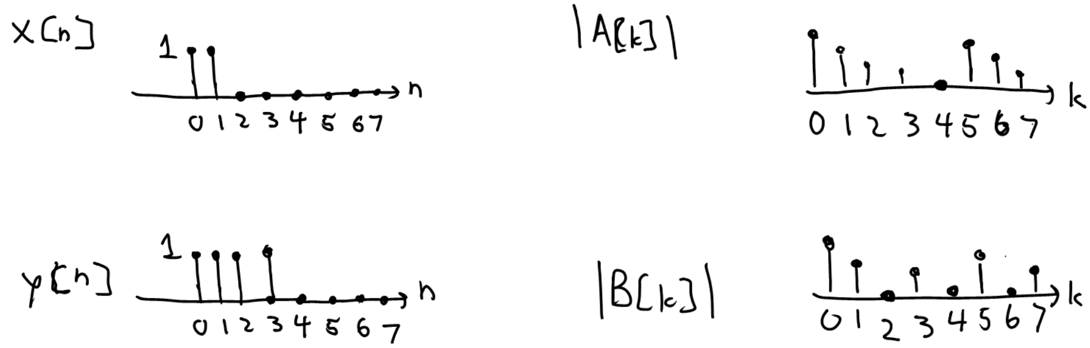
In this problem we will look at a few examples of signal analysis with DFT and how changes in the characteristics of the signals and/or the DFT window impact the resulting frequency-domain representation.

Throughout this problem, you must explain any answers you give in order to receive any credit – i.e., simply guessing an answer will result in zero points.

a) **(3 pts)** Shown below are two time domain signals (x[n], y[n]) and two magnitude plots of DFT coefficients (A[k], B[k]) – which DFT coefficients correspond with which time domain signal? (I.e., does A[k] correspond to x[n], or to y[n]?) Be sure to explain your answer.



b) **(5 pts)** For the x[n]/y[n] and A[k]/B[k] shown below, now which DFT coefficients correspond to which time domain signal?



c) **(5 pts)** If x[n] = $e^{i(4\pi/\mathrm{M})n} + e^{i(6\pi/\mathrm{M})n}$ where M is some positive integer, what is the minimum number of points you must use in your DFT in order for X[k] to contain exactly two non-zero coefficients?

PROBLEM 2. Spectral Leakage and Windowing (19 pts)

As we saw in lecture and homework, if the signal we are taking the DFT of is periodic, but an exact integer period of the signal does not fit within the window we are using, we will end up with what is known as *spectral leakage*. In practice this situation is actually extremely common; e.g., consider a BMI system where the physical mechanisms that set the frequency of an oscillation in say an EEG signal have nothing to do with the sampling rate (and hence time window) of our electronics. In this problem we will examine a technique known as "windowing" that tries to mitigate spectral leakage by multiplying the original signal x[n] with a "window function" w[n]. Conceptually, the window function attempts to taper the signal near the boundaries of the DFT interval in order to make the signal "fit better".

Let's assume that our signal of interest is $x[n] = e^{i\pi n/2}$, and that we will be taking a DFT of length 6 over the interval n = 0, 1, … , 5. The magnitudes of the DFT coefficients for this signal are |X[0]| = |X[3]| = 1.41, |X[1]| = |X[4]| = 3.86, and |X[2]| = |X[5]| = 1.04.

In this problem we'll examine the simplest type of window known as a "boxcar". Specifically, we will be multiplying x[n] by a window $w[n] = 1$ for $0 \leq n < 4$ and zero otherwise.

a) **(2 pts)** Sketch the original signal x[n] and the new windowed signal $\hat{x}[n]$

$$\hat{x}[n] = x[n] \cdot w[n] = e^{i\pi n/2} \text{ for } 0 \leq n \leq 3 \text{, zero otherwise}$$

over the same interval we used to take the DFT as the original signal.

b) **(7 pts)** Compute the DFT coefficients $X[\mathrm{k}]$ of the new signal for the same length-6 interval as the original signal.

c) **(5 pts)** Assuming that your answer to part (a) was such that $|X[0]| = |X[3]| = 0$, $|X[1]| = |X[4]| = 3.35$, $|X[2]| = |X[5]| = 0.9$, explain how applying the window $w[n]$ captures the nature of the original input signal $x[n]$ better than not windowing.

d) **(5 pts)** Defining the energy of a signal y[n] as $E_y = \sum_{n=0}^{5} |y[n]|^2$, compare the energy of the original signal x[n] against the energy of windowed signal $\hat{x}[n]$. Based on this comparison, comment on how well windowing x[n] preserves the original signal.

PROBLEM 3. Three Letter Acronyms (23 points)

In this problem we will explore how using either DFT or SVD to analyze (and perhaps even compress by keeping track of only non-zero coefficients/components) a signal of interest can give us two different views of the same underlying characteristics of a signal. Throughout this problem, we will be considering the signal $x[n] = \cos(\omega_0 n) + \cos(2\omega_0 n)$, where $\omega_0 = (2\pi/10)$.

a) **(1 pts)** If every sample of the signal requires 4 bytes of memory and we keep 1000 total samples of the signal, how many bytes will it take to store the signal?

b) **(8 pts)** Compute the DFT coefficients $X[k]$ corresponding to the signal $x[n]$ for a DFT interval of n=0, 1, … , 9. How many non-zero coefficients are there? If each coefficient and its corresponding frequency index (i.e., $X[k]$ and k) require a total of 12 bytes of memory to store, how many bytes does it take to completely represent the signal?

c) **(5 pts)** Now let's use a procedure similar to what we did with the neural data in the lab to construct a matrix based off of this signal that we can then analyze with SVD. Specifically, imagine that we construct a matrix A by taking 10-sample long intervals of the signal x[n] and using each one of those to populate the rows of the matrix – i.e.,:

$$A = \begin{bmatrix} x[0] & x[1] & \ldots & x[9] \\ x[10] & x[11] & \ldots & x[19] \\ & & & \end{bmatrix}$$

If we were to take the SVD of the matrix A, how many non-zero singular values will the matrix have in this case?

d) **(3 pts)** Given your answer to part c), if every singular value or entry in a singular vector requires 4 bytes of memory to store, how many total bytes does it take to completely represent the signal?

e) **(6 pts)** If we were to construct a new matrix B using the same approach as we had taken to construct A in part b), but using intervals of length 5 (instead of length 10), how many non-zero singular values will the new matrix B have?

PROBLEM 4. Fortune Telling (10 pts + BONUS 5 pts)

Using your newfound knowledge of PCA and the SVD, you decide to play a little trick on a few (or even a few hundred) of your closest friends and gather some data that will allow you to make an educated prediction about whether or not they will buy successfully buy a house in the ultra-competitive Bay Area real estate market. Knowing that each of these pieces of information in some way are likely to impact this, you collect the following pieces of information from each of your friends:

- Age (in years)
- How many years they have been with their current partner (0 if they are single)
- Age of their partner (0 if they are single)
- How many siblings they have
- Approximate salary (in $/year)
- Approximate salary of their partner (in $/year, 0 if single)
- Their height (in m)
- The height of their partner (in m, 0 if single)
- The number of times per week they go clubbing (0 if not single)
- Their overall satisfaction with life (rated on a scale from 0-10, with 10 being satisfied to the point of annoying everyone around them)

For the rest of this problem, let's assume that you successfully gathered this information from 100 of your friends.

a) **(3 pts)** Describe how you would construct a matrix A out of the data collected above that we might later be able to analyze in order to make a prediction about whether your friend will buy a house. You should arrange A such that the information from each friend is arranged in a row of the matrix. Be sure to indicate what the dimensions of the matrix would be.

b) **(7 pts)** Assuming that the matrix has one very dominant singular value and that you find that the data is actually indicative by running a few test cases where you know whether your friend bought a house or not, for any one particular friend i whose data you have collected, describe how you would use the matrix A and their individual data vector a_i^T to predict whether or not they will have buy a house in the next 5 years?

c) **(BONUS: 5 pts)** Using the same method as part b), suggest some other prediction you might be able to make about your friends, what data you might need to make it accurately, and under what conditions it is likely for the prediction to be accurate. Note that most of the bonus credit will be assigned to addressing the final issue (i.e., under what conditions your prediction is likely to be accurate).