EECS 452 Midterm Exam
Fall 2014

Name: ___________________________________ unique name: _______________

Sign the honor code:

I have neither given nor received aid on this exam nor observed anyone else doing so.

____

Scores:

#	Points
Section I	/40
Section II	/30
Section III	/30
Total	**/100**

NOTES:

- Open book, open notes.
- There are **8** pages including this one.
- Calculators are allowed, but no PDAs, Portables, Cell phones, etc.
- You have 120 minutes for the exam.
- **Be sure to show work and explain what you've done when asked to do so.** You will not receive partial credit without showing work.
- Unless otherwise specified all signed numbers are two's complement numbers.

Section I -- Short answer 40 points

1) X and Y are, respectively, a two's complement 8 bit Q(5) number and a 4 bit Q(3) number: Let Z be the result of multiplying X and Y together. [5]

a) What is the range of values of X (in decimal)?: Min=__-4_, Max=_4-2^(-5)_.

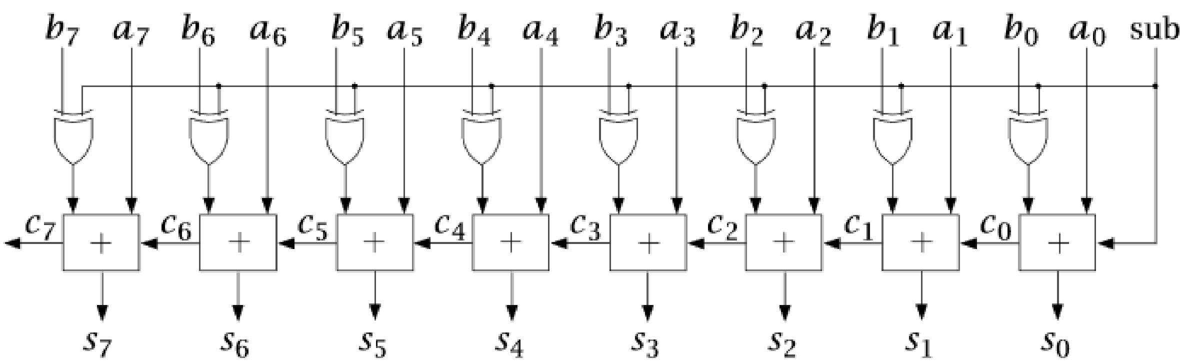
b) What is the range of values of Y (in decimal)?:: Min=__-1_, Max=_1-2^(-3)_.

c) Z is an n bit Q(n-k) number. What are n and k?: n=__12_. k=__4__.

d) What is the range of values of Z (in decimal)?: Min=_-8_, Max=_8-2^(-8)_.

e) How many right shifts are needed to convert Z to a 4 bit Q(3)? _____5______.

2) Consider the following circuit and answer the questions below [5]



a) In one sentence describe the function of this circuit. [1]

This is an 8 bit two's complement ripple adder

b) Let $a_7a_6a_5a_4a_3a_2a_1a_0$=10000001=0x81 and $b_7b_6b_5b_4b_3b_2b_1b_0$=01110111=0x77 and sub=0. What are $c_7c_6c_5c_4c_3c_2c_1c_0$ and $s_7s_6s_5s_4s_3s_2s_1s_0$ in both binary and hexadecimal? [2]

$c_7c_6c_5c_4c_3c_2c_1c_0$=00000111 (0x07) and $s_7s_6s_5s_4s_3s_2s_1s_0 = 11111000$ (0xF8) =-8

c) Repeat (b) when sub=1. [2]

$c_7c_6c_5c_4c_3c_2c_1c_0$=10000001 (0x81) and $s_7s_6s_5s_4s_3s_2s_1s_0 = 00001010$ (0x0A)=10.
Note if we count the carry bit c7 then the result of the subtraction is correct: 246-10=246.

3) Consider adding three 8 bit Q0 two's complement numbers from left to right: 0xA3+0x7E+0x04. Answer the following [5]:

a) What is the result of the addition without saturation in hexadecimal and in decimal? [2]

0xA3+0x7E=10100011+01111110=00100001 (the carry bit c7 of adder will be 1 but there is no overflow since adding neg+pos number)

0xA3+0x7E+0x04=00100001+00000100=00100101 which is 0x25 or 37 in decimal

b) What is the result of the addition with saturation at each step? [2]

As long as the logic detects that a neg+pos number were being added so that it does not incorrectly declare overflow followed by saturation there will be no saturation applied to either step.

c) Is only one of (a) and (b) right, are both right or are both wrong in terms of giving the correct value of the sum? [1]

Both are correct.

4) An ADC with 8 bit quantizer resolution has a signal-to-quantization-noise-power-ratio (SQNR) equal to 2 (or 3dB). How many more bits of resolution would have to be added to the ADC to raise the SQNR to at least 20 (or 14dB)? [5]

Recall from lecture 4 that we gain 6dB for every bit we add to the ADC. 20.0 corresponds to 10*log10(20)=13dB so we would need 2 more bits, which would bring us to 15dB or a SQNR equal to 10^(15/10)=31.6 (adding only 1 more bit would only bring us to 9dB or SQNR equal to 7.9).

5) An audio signal is bandlimited to 20KHz and is to be digitized for processing using a ADC. Answer the following [5].

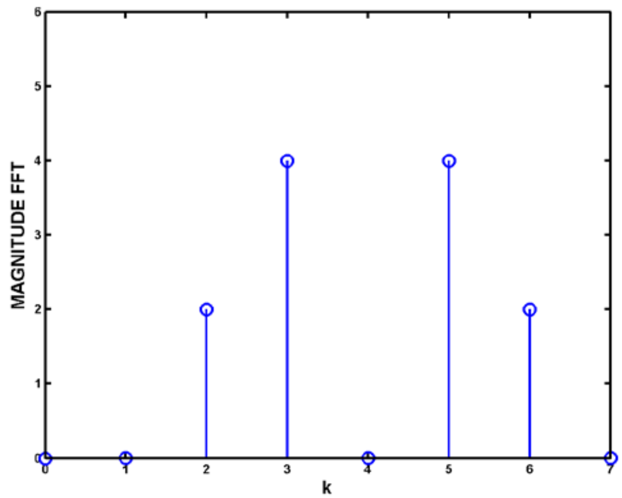
a) What is the lowest sampling frequency that suffices to recover information about the audio signal from its samples? [2]

An audio signal occupies most of the band, i.e. it has a low pass spectrum. Hence, the Nyquist rate of 40kHz is the lowest applicable sampling frequency.

b) If spectrum analysis of this signal is to be performed using a N-point radix 2 DFT what is the minimum value of N required in order that the DFT have a frequency resolution of 10Hz or better? [3]

The frequencies in the DFT are spaced F_s/N Hz apart. Hence for this to be equal to or less than 10Hz we require $N \geq 40{,}000/10$=4000. Hence a 4096 point radix 2 FFT would work.

6) A real-valued time domain waveform that is bandlimited to 1kHz is sampled at the Nyquist sampling rate. Its 8 point FFT has a magnitude spectrum shown below. The following questions are about the frequency components contained in the time domain waveform. [5]



a) How many sinusoidal components are present in the waveform? [2]

Two sinusoids are present and their frequencies correspond to the two DFT indices 2 and 3. The DFT indices 5 and 6 contain energy in the conjugate symmetric frequencies of the two sinusoids and do not correspond to additional sinusoids.

b) What are the frequencies (in Hz) of these sinusoids? [2]

The Nyquist sampling frequency is Fs=2*1000Hz=2kHz. The frequency spacing in the FFT corresponds to Fs/N=2000/8=250Hz. Hence the two sinusoidals are at 500Hz and at 750Hz.

c) What are the amplitudes of these sinusoids? [1]

The magnitude of each line in the FFT magnitude plot is N/2=4 times the amplitude of the sinusoid at that frequency. Hence the lower frequency sinusoid has amplitude 0.5 and the higher frequency sinusoid has amplitude 1.0.

7) A filter has transfer function equal to $H(z) = 1 + z^{-2} - 2\,z^{-4} + z^{-6}$. Answer the following [5]

a) What is the impulse response h[n] associated with this filter? [2]

n=	0	1	2	3	4	5	6	7…
h[n]=	1	0	1	0	-2	0	1	0…

b) What is the order of this filter. Is it FIR or IIR? [2]

The order is 6 and it is FIR.

c) If a discrete time signal is input to this filter, and the signal sampling rate is 8kHz, what is the group delay of this filter at frequency F=0 Hz? [1]

The phase of this FIR is not linear since the impulse response is neither symmetric nor antisymmetric about any time index. Hence, the filter does not have constant group delay. You could find the group delay at f=0 by computing the phase phi(f)= tan^(-1)[Im(H(e^(j2 pi f))/Re(H(e^(j2 pi f))] of H(e^{j2 pi f)) and differentiating it, dividing by 2 pi Fs, and evaluating it at f=0. This would give you the answer that the group delay is equal to 0 at f=0. The group delay would be equal to 4Ts if H(z) were z^(-2)-2z^(-4)+z^(-6) – the additional “1” in the stated H(z) was an unfortunate typo in this problem. Due to this typo no-one got points taken off this question.

8) Shown below is a filter constructed of two second order IIR (biquad) filters like the ones studied in Lec 11. Assume that the poles of the two filters are not at the same location inside the unit circle. For a), b) and c) state whether the filter can be equivalently represented as shown. [5]

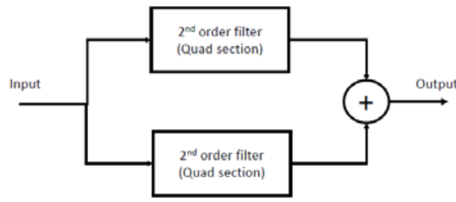
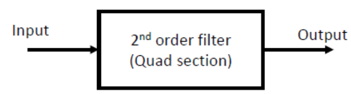


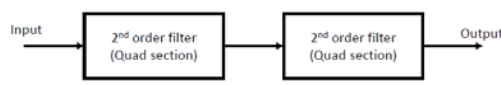
Figure 1 A filter composed of 2nd order sections.

a) Is the filter shown below equivalent to the filter in Fig 1, perhaps with different zero and pole locations than the 2nd order filters in Fig 1? [2]



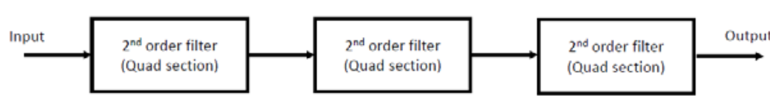
Your answer: Y N. Explain.
The answer is No. Fig 1 shows the output to be the sum of two second order filters. Let the top filter in Fig 1 have z-transform H1(z)=B1(z)/A1(z) and the bottom filter have z-transorm H2(z)=B2(z)/A2(z) . The transfer function is H1(z)+H2(z) = [B1(z)A2(z)+B2(z)A1(z)]/[A1(z)A2(z)] which is 4th order if the poles in H1 and H2 are distinct.

b) Is the filter shown below equivalent to the filter in Fig 1, perhaps with different zero and pole locations than the 2nd order filters in Fig 1? [2]



Your answer: Y N. Explain.
The answer is Yes. As explained for part a) the filter in Fig 1 is equivalent to a 4th order filter and, using the methods we discussed in class, this can be represented as a cascade of two 2nd order filters.

c) Is the filter shown below equivalent to the filter in Fig 1, perhaps with different zero and pole locations than the 2nd order filters in Fig 1? [2]



Your answer: Y N. Explain.
The answer could be yes or no. Yes if you say that one of the three 2nd order filters is all-pass (e.g. H3(z)=1). Otherwise the answer is no since the cascade of 3 2nd order filters is a 6th order filter.

Section II – longer answers 30 points

1) Consider two variables `(Int16)x` and `(Int16)y`. `x` is Q(8) and `y` is Q(8), both are in two's complement format. Suppose `x = 0x0400` and `y = 0x0300`. [15]

   a) What are the decimal values of `x` and `y`? What is their decimal product? [3]

   i) x = 4.0, y = 3.0, x*y = 12.0

   ii) It seems too good to be true! In Q8, the bottom 8 bits (which are all zeros) for both x and y are the fractional part of the value. So $x=2^2=4$, $y=2^1+2^0=3$.

   b) What is the product of `x` and `y` in hexadecimal two's complement Q(6)? [3]

   iii) Result size was not specified so 32 or 16 bit answers were accepted.

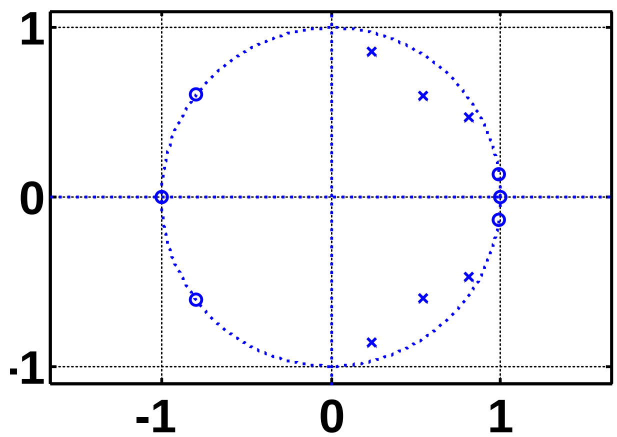
   iv) The product is `0x000C0000 >> (8+8-6) = 0x0300`. (or `0x00000300`) (or `0x300`)

   v) Note that the product is in Q6 while the multiplier y is in Q8. They are both given the 16-bit representation of `0x0300`, however they are different "values" because of how we interpret them.

   c) Write no more than three lines of C-code that multiply `x` by `y` and convert the result back to a 16-bit two's complement Q6 number (with rounding.) Place the resultant product into `(Int16)result`. [9]

```
Int32 temp = (Int32)x * (Int32)y;
Int16 result = (Int16)( ( (temp + (1<<9) ) >> 10);
// Note: (1<<9) == 0x0200
// Int16 cast may be omitted
// Int32 casts may not (for a 16-bit system architecture.)
```

2) A filter has the pole-zero constellation shown below. Answer the following. [15]

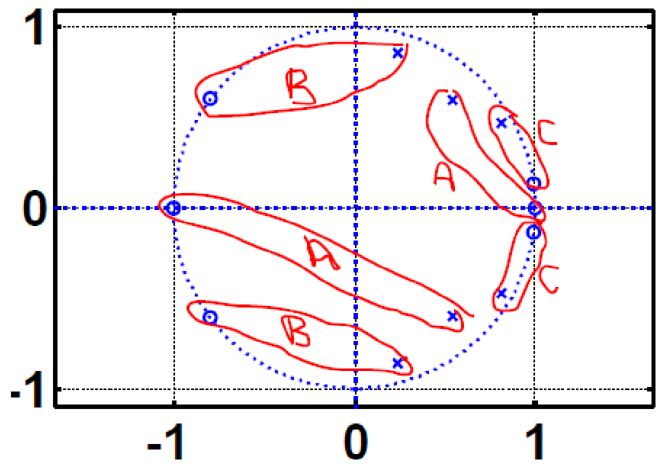


a) What is the filter order and how many 2nd order biquad sections would you need to implement this as a cascade of 2nd order filters as in Lab 5? [5]

The filter order is 6 (6 poles and 6 zeros). Hence you would need 3 biquads to implement as a cascade of 2nd order filters.

b) On the figure please circle the pole-zero pairs that would constitute each 2nd order biquad section that you would use to implement this filter in cascade form as in Lab 5. Indicate how you would order these 2nd order sections for the cascaded implementation by drawing a block diagram with input on the left and output on the right? [5]

There are two zeros that are real and simple (not double zeros). Hence, using the algorithm specified in class and in lab, you obtain:

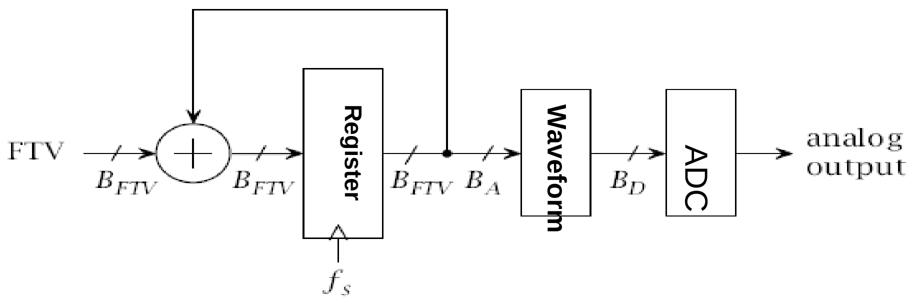


The grouping's labeled A, B, C should be cascaded in that order from left to right with input on the left and output on the right of the cascade.

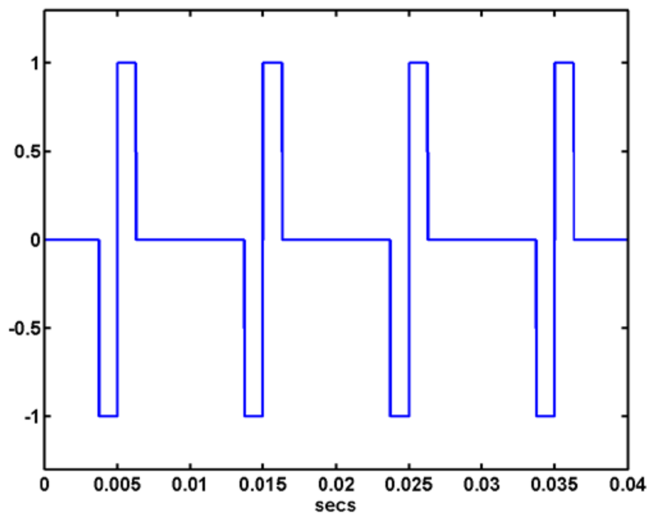
c) Would any of the poles result in 2^{nd} order biquad sections whose denominator coefficients would need to be scaled in order to implement in Q(15) on a 16 bit processor? If so, which poles need attention and how would you do the scaling?

The two pole pairs closest to the unit circle (in groupings A and C) would need attention since they have real parts that are greater than ½.

Section III Longer answers 30 points



1) The DDS system above is to be used to generate a periodic waveform. The CPU clock frequency is fs=56MHz, the register length is B_{FTV}=32 bits, and the waveform table has B_A=8 bit addressable space to store samples. The desired periodic waveform is shown below, where the duty cycle is equal to 1/3 (i.e., in any period only 1/3 of the samples are non-zero). Assume that the ADC is ideal (cardinal series reconstruction). Please answer the following [15].



a) What values need be stored in the waveform table? Draw a sketch being careful to label all axes, transition indices, and amplitudes. [5]

You simply need to note that this is periodic with period 0.01 and sketch the first period over indices n=0,1,2…255. The first 84 (1/3 of 256) entries of the table are zeros. The entries between indices 85 and 127 are -1. The entries between 128 and 170 are +1. The entries between 171 and 255 are zero. Note that there are only 3 amplitude values stored in the waveform table -1,0 and 1.

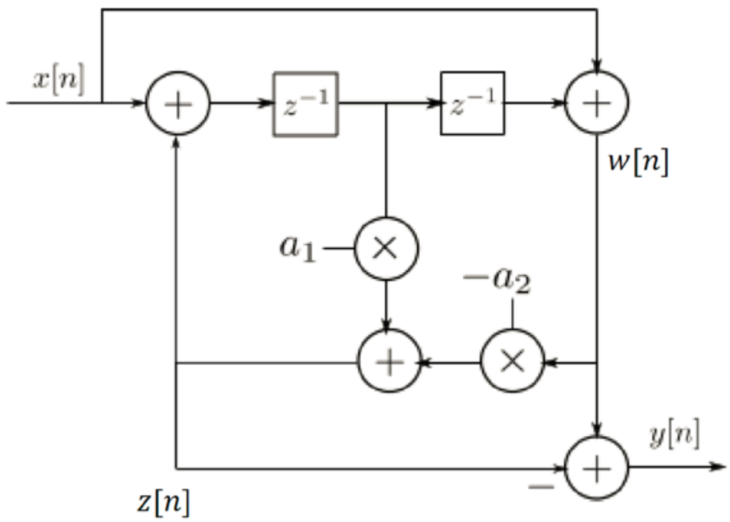
b) What value should FTV be set to? [5].

Fo=FTV(Fs/2^BFTV)=or FTV=2^(32)*100/(56*10^6)=7669.6 so set FTV to 7669 or 7670.

c) What is the minimum number of bits of resolution B_D needed for the waveform table? [5]

Only need 2 bits of resolution since the entries in the waveform table take on only three different values -1,0,1.

2) A filter is shown in the figure below. The input is x[n], the output is y[n] and w[n], z[n] are internal states. Please answer the following. [15]



a) At what points in the filter do we have to be concerned about overflow if we use Q(15) arithmetic? [5]

All the summers will require attention.

b) Express y[n] in terms of x[n], z[n] and w[n] in both the time domain and the Z-transform domain. [5]

In terms of x, x, and w: y[n]=w[n]-z[n] or Y(z)=W(z)-Z(z).

c) Express w[n] in terms of x[n] and z[n] and express z[n] in terms of x[n] and w[n] in both time domain and Z-transform domain. Based on these equations can you tell if the filter is FIR or IIR? [5].

w[n]=x[n]+ (x[n-2]+z[n-2]) or W(z)=X(z)(1+z^(-2))+Z(z)z^(-2)
z[n]=a1(x[n-1]+z[n-1])-a2w[n] or Z(z)=a1 z^(-1)(X(z)+Z(z))-a2 W(z)
We can take the equation for W(z) and substitute it into the equation for W(z) obtaining:
Z(z)=a1 z^(-1)(X(z)+Z(z))-a2 (X(z)(1+z^(-2))+Z(z)z^(-2))
Express this equivalently as an equation for Z(z) in terms of X(z):
Z(z) (1-a1 z^(-1)+a2 z^(-2))=X(z)(-a2+a1 z^(-1) - a2 z^(-2))
Or equivalently, we see that Z(z) is the result of an IIR filter applied to X(z)
Z(z)= (-a2+a1 z^(-1) - a2 z^(-2)) /(1-a1 z^(-1)+a2 z^(-2)) X(z).
Therefore, we can conclude that the filter above is an IIR filter, since, from part (b) and the expression for W(z)
Y(z)=W(z)-Z(z) = X(z)(1+z^(-2))+Z(z)z^(-2)-Z(z)
= 1+z^(-2)-(1-z^(-2))(-a2+a1 z^(-1) - a2 z^(-2)) /(1-a1 z^(-1)+a2 z^(-2)) X(z)
= H(z) X(z)

where H(z) is an IIR filter since it has poles.