Sample Final

- **Google stalking**

  How does Google stalk you? Turns out that you can understand the basics of ad generation and product recommendation on the internet using the ideas we have discussed in class. The key to Google's revenue stream is showing users *relevant* ads, i.e. ads that are likely to engage and interest a given user. So given a product, Google would really like a way to be able to predict whether a user will buy a certain product or not. For this, it is constantly stalking users using cookies, having users sign into their Google+ accounts, recording their browsing history, and so on.

  To be able to recommend relevant products to users[1], Google is constantly collecting information. Here we will understand some toy models related to how this is done.

  Google has a list of 10^3 products that it wants to sell users. It also has collected information about a person's age (a real number between 0 and 100) and their address. A user's address is a very rich source of information about him or her, and it can be converted into three more pieces of information: (1) the average-income level of the neighborhood, (2) the ethnic-mix of the neighborhood and (3) the political leanings of the neighborhood. You can assume these are all also normalized to be real numbers between 0 and 100. Finally, they have a last bit of data: the search query. Let us also represent this as a real number between 0 and 100.

  Let us consider one product, an XBox, that Google wants to study. They want to develop a simple model to understand whether a user is likely to buy an XBox or not (because this will help them show relevant ads). To do this, they run a bunch of experiments of the following nature. For 10^3 users that visit the Google homepage, they record all the above data and also display an ad for an XBox. Then, they record whether the user clicks on the ad or not as $+1$ if the user clicks on the ad, and -1 if the user does not click on the ad.

  1. First, explain how you might set up a linear model to understand the relationship between whether someone clicks on an ad and the demographic information Google has about them.
  2. Now explain how you might use this model to predict the behavior of a new user.

[1] The fancy word for the system that "Google uses to sell you stuff" is a *recommender system*.

- **Limited happiness**

  Thanks to you working on your 16A homework late into the night your roommate is now obsessed with Pharell's "Happy" just as much as you have been in the past two weeks. But unfortunately he/she does not have the file to play the song and you want to send it over. But you have just finished finals and were binge watching YouTube, so you ran out of all the data allocation on your plan. As a result cannot send the entire file — it is too much data. The copy of the song you have is 50,001 samples long, i.e. you have the amplitude of the song at 50,001 time instances. That is 50,001 real numbers, which you cannot afford to transmit. Because of limited reception, your ability to transmit is limited to sending 10,000 numbers. So you want to figure out a way of compressing the song to 10,000 values, while still maintaining the essential characteristics of the song.

  In your toolbox (i.e. in your backpack as you head out on BART[2]) you have:

  – a function that can take any time domain signal and return to you its frequency domain representation (which might include complex numbers). Let n be an odd integer. If you give the function a signal of length n, it will return back to you n real numbers that are the frequency components at the frequencies $\frac{2\pi}{n}k$, where k is an integer that takes values in between $-\frac{n}{2}$ to $\frac{n}{2}$.
  – a function that can take in a vector of length n and pick out the indices and the values of the k largest values in the vector and return these to you a set of tuples. So if you give the function the vector $[9,7,4,1,2,3,8]$ and choose $k = 3$, the function will return to you $(1,9),(2,7),(7,8)$.
  – a function that can take a vector of complex numbers and return a vector of their magnitudes

  Using these two functions describe the strategy you will use to choosing what numbers you will transmit to your roommate. Explain why your strategy is a good one.

[2]bear with me here

- **Spectrum allocation**

  There was a point in time when most television signals used to be transmitted wirelessly using Analog signals. In fact, this is true for some TV sets even today. Different television channels are allocated different parts of the frequency spectrum to transmit in. However, large chunks of this spectrum are unused. This problem explores how we might be able to use them to transmit your own signals.

  You know that Channel A occupies the spectrum from 500MHz to 506MHz, Channel B occupies the spectrum from 510MHz to 516MHz and Channel C occupies the spectrum from 524MHz to 530MHz. The other parts of the spectrum between 500MHz and 530MHz are unoccupied.

  1. You are designing a communication system that uses amplitude modulation. What carrier frequencies between 500MHz and 530MHz might you consider while designing your system, assuming you have access to perfect bandpass filters?
  2. The signal that you want to transmit is in the form of a cosine wave $\cos(\omega t)$. You get design freedom to choose the frequency ω. Assuming you have access to perfect bandpass filters, and you are using amplitude modulation as your communication scheme, what is the maximum ω that your signal can have? What would your carrier frequency be (between 500MHz and 530MHz) to be able to use this signal?

- **Missing women**

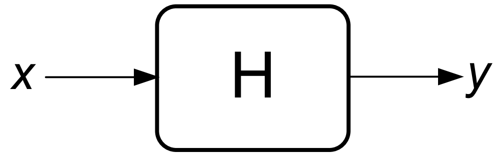
  In 1990, Nobel Laureate economist Amartya Sen noted that many countries in Asia had “missing women”. Sen estimated that more than a hundred million women were "missing" (in the sense that their potential existence had been eliminated either through sex selective abortion, infanticide or inadequate nutrition during infancy). So for instance, the United States has more women than men in the population, but India and China have more men than women[3]. This problem will explore a small toy model related to this phenomena and steady state-populations.

  1. Let n be a time index that denotes the year. Let $g[n]$ denote the total number of women (ages 0 and above) in the population, and similarly let $b[n]$ denote the number of men in the population. Let us say that each year the mortality rate for women is $m_g\%$ and the mortality rate for men in $m_b\%$. Further, let us say that of all the women, $\frac{1}{4}$ are of childbearing age (i.e., in between ages 20-40, where we assume the age range of women to be 0-100). Of those that are of childbearing age, let us say that 10% will actually have children in a given year. In certain Asian populations, because of the factors above, there are 100 girls born for every 111 boys (source UN population division). In our simple model, the men are irrelevant for producing children, and so the fraction of newborns born in a year is only a function of the number of women that year. Set up a simple system of linear equations to model how the two populations evolve in time (i.e. an equation for $g[n+1]$ and $b[n+1]$ in terms of $g[n]$ and $b[n]$.
  2. What is the total number of girls being born each year as a fraction of the current population of girls? If $m_g = 1.3\%$, is the total number of women in the population growing or dying out?
  3. Write this system of equations in matrix format using the variables m_g and m_b
  4. The matrix that relates the populations at time n to the populations at time $n+1$ plays a key role in understanding steady state. Let $m_g = m_b = \frac{100}{40 \cdot 211} = 1.1848\%$. What is the steady-state fraction of men to women in this case?
  5. (Bonus) If $m_g = 1.1848\%$, for what value of m_b does the population of men and women stay roughly equal?

[3] This explanation for why the population of women is smaller than the population of men has been debated by other economists, but it is still the most widely accepted explanation.

- **Discrete-Time LTI System, Impulse Response, Frequency Response**

  Consider a discrete-time LTI system H as shown below.



  You're told that the unit-step response of the system—that is, the response if the input signal is the unit step, $x(n) = u(n)$ for all integers n—is

$$y(n) = \begin{cases} 0 & \text{if } n \leq -2 \\ 1/4 & \text{if } n = -1 \\ 3/4 & \text{if } n = 0 \\ 1 & \text{if } n \geq 1. \end{cases}$$

  1. Determine, and provide a well-labeled plot of, $h(n)$, the impulse response of the filter.
  2. Determine a reasonably simple expression for, and provide a well-labeled plot of, $H(\omega)$, the frequency response of the filter.
  3. Determine the output of the filter if the input is

$$\forall n \in \mathbb{Z}, \qquad x(n) = A + B(-1)^n + C\cos\left(\frac{\pi}{2}n\right),$$

     where A, B, and C are (possibly complex-valued) scalar constants.

- **Exponentially-Weighted Moving-Average Filter (FIR Case)**

  One of the tools used by stock traders to determine trending behavior in the prices of stocks that interest them is the *Exponentially-Weighted Moving Average (EWMA)* filter. The idea is to determine a moving average value such that the more recent closing stock prices have progressively higher weights in the computation of that weighted average value. Let $x(n)$ denote the closing price of Stock X on day n. Let $y(n)$ denote the corresponding EWMA value; that is, $y(n)$ is the output of the EWMA filter at the close of trading on day n. We can model the input-output behavior of an M-day EWMA filter using the equation

  $$\forall n \in \mathbb{Z}, \quad y(n) = \beta \sum_{m=0}^{M-1} \alpha^m x(n-m),$$

  where β is a positive normalization constant, and $0 < \alpha < 1$ denotes a weight factor. Clearly, α^m decreases as m increases, which means that the farther back in time we go, the less the influence of that day's closing price is on the current day's EWMA value.

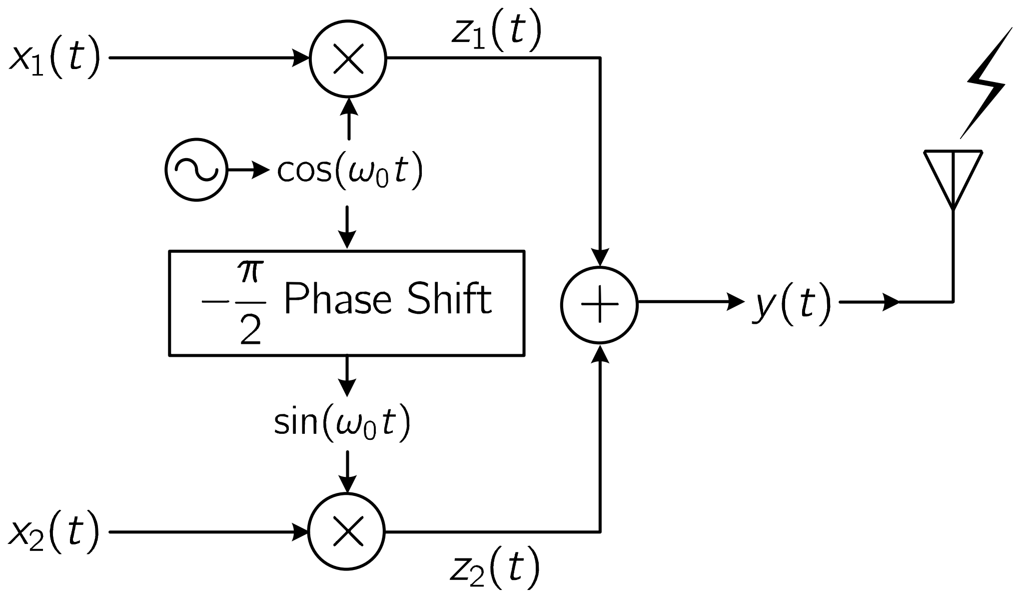
  1. Determine the normalization constant β such that if the input signal is given by $x(n) = 1$ for all n, then the output signal y is identical to the input signal. The β that you determine must be in the form of a reasonably simple expression involving α.
  2. For $\alpha = 1/2$ and β computed according to the expression you arrived at in part (1), determine, and provide a well-labeled plot of, the impulse response values $h(n)$ for all integers n.
  3. In terms of β and α, determine a reasonably simple expression for $H(\omega)$, the frequency response of the EWMA filter H.
  4. We call $H(0)$ the *DC gain* of the filter H—that is, the gain that the filter applies to a zero-frequency complex-exponential input signal given by $x(n) = e^{i0n} = 1$ for all integers n. For the value of β that you determined in part (1), what is the corresponding DC gain of the filter?
  5. Suppose the price of a hypothetical stock oscillates between \$1 and 50 cents on alternate days. In particular,

     $$x(n) = \begin{cases} \$1 & \text{if } n \text{ is even} \\ \$0.5 & \text{if } n \text{ is odd.} \end{cases}$$

     For $\alpha = 1/2$ and β determined according to part (1), provide a reasonably simple expression for, and a well-labeled plot of, the output values $y(n)$ for a *ten-day* EWMA filter.

- **Quadrature multiplexing**

  The figure below illustrates a transmission scheme known as *quadrature multiplexing*, by which we can embed two bandlimited continuous-time signals on the *same* carrier frequency. This method—which is also referred to as *quadrature-carrier multiplexing* and *quadrature amplitude modulation (QAM)*—allows the information-bearing signals to have overlapping footprints on the frequency spectrum.



  The trigonometric identities below may be useful to you in one or more parts of this problem:

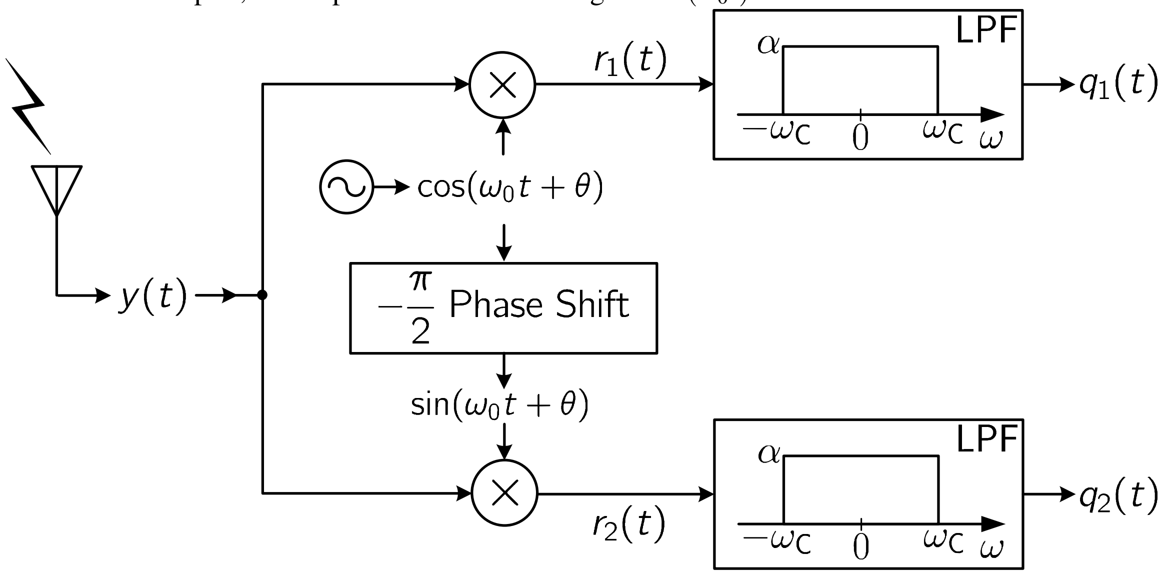
$$\begin{aligned} \cos a \cos b &= \frac{1}{2}\Big[\cos(a+b) + \cos(a-b)\Big] \\ \sin a \cos b &= \frac{1}{2}\Big[\sin(a+b) + \sin(a-b)\Big]. \end{aligned}$$

  (a) Show that the transmitted signal y can be expressed in the amplitude-phase form

$$\forall t \in \mathbb{R}, \quad y(t) = A(t) \cos\big(\omega_0 t + \phi(t)\big),$$

  and determine how each of the signals x_1 and x_2 contributes to the amplitude A and the phase ϕ of the transmitted signal y. In other words, express each of $A(t)$ and $\phi(t)$ in terms of $x_1(t)$ and $x_2(t)$.

The figure below shows a receiver demodulator for the QAM transmission scheme described above. Note that the local oscillator at the receiver has a phase error of θ (radians) with respect to its transmitter counterpart, which produces the carrier signal $\cos(\omega_0 t)$.



The passband gain of each ideal low-pass filter is $\alpha = 2$, and its cutoff frequency is much smaller than the carrier frequency, yet sufficient to pass the baseband content of interest; that is, $A < \omega_c \ll \omega_0$.

(b) Determine a reasonably simple expression for the output signals q_1 and q_2. Your answers should be in terms of the information-bearing signals x_1 and x_2, and the phase error θ.

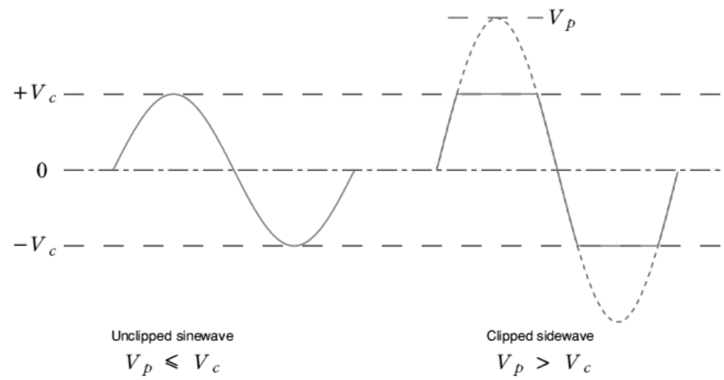
(c) In this part, assume that the phase error is very small ($|\theta| \ll 1$).

  (i) Simplify your expressions for q_1 and q_2 in part (b) based on the assumption that the phase error is very small.

  (ii) Suppose the information-bearing signals x_1 and x_2 are of equal magnitudes at each time instance: $|x_1(t)| = |x_2(t)|$, for all t. What's the largest allowable phase error θ such that the interference of x_2 in q_1 (or, equivalently, the interference of x_1 in q_2) does not exceed 1% (i.e., -40 dB)? Express your answer in radians and degrees.

Audio Distortion Generator

As we discussed in lecture, "distortion" is a common effect used in rock music to create a fuzzy guitar sound. In practice, this distortion is achieved by "clipping" the sound waves as shown below:



For this problem we will develop a couple of different circuits that will enable us to create a variable amount of distortion by having an adjustable clipping level (V_c in the figure above). In particular, you are given an audio input voltage source V_{audio}, a voltage source with value V_{clip} that sets the voltage at which the **output** of your circuit should clip, and a pair of power supply voltage sources V_{dd} and V_{ss}, where $V_{ss} = -V_{dd}$ and you can choose V_{dd} to have whatever value you would like. You should assume that both of these voltage sources have source resistances of 10kΩ, and that the output of the circuit drives a 100Ω load resistance. As indicated in the figure, you can further assume that V_{audio} can take on values anywhere in the interval -5V to 5V, and that V_{clip} can range from 0V – 5V. Finally, when the circuit is not creating distortion (i.e., the output is not clipped), its output should simply reproduce the input with a gain of 1.

(a) Given the description above, write an equation that expresses V_{out} (the output of the distortion circuit) as a function of V_{audio} and V_{clip}.

(b) Implement the distortion generation circuit using only the provided voltage sources along with op-amps and resistors.

(c) After having successfully designed and built this distortion generation circuit, you start considering selling them to other people to make some money on the side. It turns out however that the op-amps you used are relatively expensive, and so you start asking around to see if there any other components available that might let you build the circuit while using fewer op-amps.

In this context, one of your colleagues tells you a component called a “diode”, whose symbol is shown below. These diodes are made out of semiconductors (often silicon, which is the same thing we use to build the vast majority of the integrated circuits embedded in the products you own), they (ideally) have the very interesting characteristic that when the voltage across them is greater than ~0.7V (i.e., Vdiode > 0.7V), the diode behaves like a short circuit (i.e.. 0Ω resistor), while when the voltage is less than ~0.7V (i.e., Vdiode ≤ 0.7V), the diode behaves like an open circuit (i.e., a resistor with infinite resistance).

For the sake of simplicity, assuming that we always wanted the output of our distortion circuit to clip when |Vaudio| > 0.7V, design a new distortion generation circuit using resistors, diodes, and only a single op-amp.

(d) (BONUS) Still using only one op-amp but assuming that you have ideal voltage sources with values of $\pm V_{clip}$, and that the diode become short circuited whenever V_{diode}>0V (instead of V_{diode} > 0.7V in part c), augment your diode-based circuit so that it regains the ability to adjust the voltage at which clipping occurs based on the value of V_{clip}.

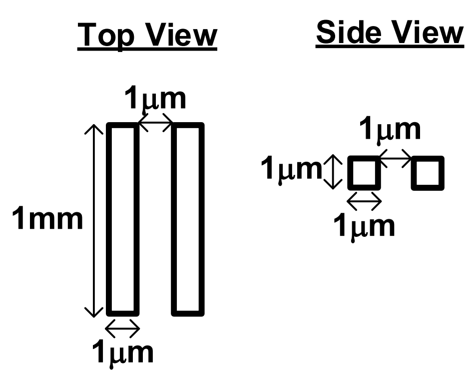
Scaling of Wires in an Integrated Circuit

One of the main factors behind the dramatic increase in available computing capabilities is that the semiconductor (chip) industry – and in particular companies designing processors like Intel, AMD, Qualcomm, Nvidia, etc. – has been able to follow a principle known as "Moore's Law", which states that the number of components (transistors) that can be integrated on to a single chip with a certain total cost should roughly double every 2-3 years. In this manner, the cost of each component would decrease by ~2X in that same time frame, and hence every generation we are able to integrate on to our chips ever more sophisticated designs with ever-improving capabilities while charging roughly the same for that part as we charged for the previous, less capable generation of parts.

Technologically speaking, the ability to integrate more components has been driven by what is known as "scaling", where in moving from one generation to the next, all of the components in the chip can be "scaled" such that they are physically smaller by some scale factor than they were before. For example, if the dimensions of every component (and the spacing between the components) were scaled by a factor of $1/\sqrt{2}$ (~0.7), then a square chip could fit twice as many components in to the same area.

In this problem we will take a brief look at what this process of scaling does to the resistance and capacitance of metal wires that we might have in our chip to make connections between the various components. As we will see further in EE16B, the resistance and capacitance of these wires can have a direct influence on both the performance (clock speed) and power consumption of the chips that we build.

(a) Let's start by looking at two metal wires on a first generation chip shown below. Let's assume that the left hand wire carries some signal we're interested in (e.g., the DRAM column we saw from homework 3), and the right hand wire is ground. Using the dimensions shown below and assuming that the resistivity of the metal (copper) is 17 nΩ*m, what is the electrical resistance between the bottom of the signal wire and the top of the signal wire?



(b) For the same wire configuration as part (a) and assuming that the wires have silicon dioxide in between them ($\varepsilon = 3.9 * 8.85e\text{-}12$ F/m), what is the capacitance between the signal wire and the ground wire?

(c) Now let's look at what happens when we scale these wires to the next technology generation – i.e., the dimensions of all of the wires and the spacing in between them are multiplied by $1/\sqrt{2}$. Now what is the resistance between the bottom of the signal wire and the top of the signal wire? Similarly, what is now the capacitance between the signal wire and the ground wire?

(d) The case we examined in part (c) is for what's known as a "local wire", where the distance between the components it connects is reduced as the dimensions of the components themselves are reduced. However, some wires are what's known as "global wires" whose length does not reduce as the dimensions of the components are scaled; these wires are required to make connections between components that are at fixed positions relative to the overall dimensions of the chip (e.g., imagine one block at one corner of the chip that needs to communicate with another block that exists at another corner of the chip).

Using the same assumptions as part c) but with the length of both the signal and the ground wires fixed at 1mm, now what are the resistance and capacitance values?

(e) One of the biggest issues faced by the people who build these chips is that this scaling makes delivering the power supply voltage on the chip increasingly difficult. (As a side fun fact, Elad's Ph.D. thesis was directly related to this topic.) To get a glimpse of why this is the case, for the fixed length wire from part (d), if we always need 10mA of current to flow through the signal wire, how does the magnitude of the voltage dropped from the top of the wire to the bottom of the wire scale as we move from one technology generation to the next (i.e., scale all of the other dimensions by $1/\sqrt{2}$ for each generation)?