

**This homework is due on Monday, September 9, 2019, at 11:59PM.**

**Self-grades and HW Resubmission are due on Wednesday, September 11, 2019, at 11:59PM.**

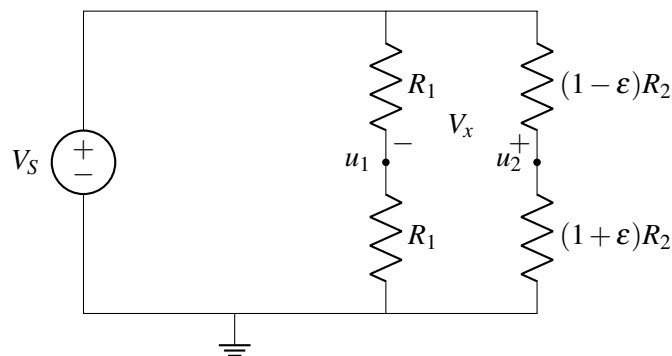
## 1. MEMS Accelerometer

### Optional problem: covering 16a-level circuit analysis and modeling

A micro-electromechanical system (MEMS) accelerometer is a device that can measure acceleration, for example by using a set of strain-sensitive resistors. There are three in every cell phone, detecting the phone's orientation and motion. MEMS accelerometers are made using silicon micromachining. In the accelerometer, a silicon block with a known mass is suspended between springs made of silicon. The compression of the springs can be measured because the resistance of a silicon spring changes when it is compressed. (This occurs because silicon is a piezoresistive material, which we will not talk about in this course.)

Accelerating the device causes the silicon block to move, changing the compression of the attached silicon springs, and therefore changing the resistance across the springs. One of the springs will be compressed while another will be extended, so the resistance of one spring increases while the other decreases. If we measure the changes in the resistance of the springs, then we can understand how the silicon block is moving.

However, the change in resistance is extremely small. For instance, for a change of  $9.8\text{m/s}^2$  of acceleration (equivalent to the Earth's gravitational acceleration,  $g$ ), the resistance only changes by about 4% in our example! To measure such a small resistance change, the resistors are placed in the following configuration known as a Wheatstone bridge:



The resistances of the two resistors on the left, both with the same value  $R_1$ , will remain constant. The two resistances on the right represent the silicon springs. The  $\epsilon$  term represents the fractional change in resistance brought about by movement of the silicon block. For example, if  $\epsilon = 0.01$ , then the resistance of spring being compressed will increase by 1%, while the resistance of the spring being extended will decrease by 1%.

A voltmeter measures the voltage difference  $V_x = u_2 - u_1$  on the device. We use  $V_x$  to determine the change in resistance and hence the acceleration.

- (a) To determine the acceleration, we first need to understand the relationship between our measured voltage  $V_x$  and the resistances of the springs. **What is  $V_x$  in terms of  $R_1$ ,  $R_2$ ,  $\epsilon$ , and  $V_s$ ?**

**Solution:**

Applying the Voltage Divider Rule, we get:

$$V_{u_1} = \frac{R_1}{R_1 + R_1} V_s \quad (1)$$

$$\Rightarrow V_{u_1} = \frac{1}{2} V_s \quad (2)$$

Similarly,

$$V_{u_2} = \frac{(1 + \epsilon)R_2}{(1 + \epsilon)R_2 + (1 - \epsilon)R_2} V_s \quad (3)$$

$$\Rightarrow V_{u_2} = \frac{(1 + \epsilon)}{2} V_s \quad (4)$$

Therefore,

$$V_x = \frac{(1 + \epsilon)}{2} V_s - \frac{1}{2} V_s \quad (5)$$

$$= \frac{\epsilon}{2} V_s \quad (6)$$

- (b) Suppose the minimum voltage the voltmeter can detect is  $V_x = 1\mu\text{V}$ . If this is the minimum  $V_x$ , **what is the minimum measurable resistance change  $\epsilon$  that we can measure?** We are going to make the simplifying assumption that the  $\epsilon$  varies linearly with the acceleration. If each acceleration change of  $9.8\text{m/s}^2$  ( $1g$ ) corresponds to a change in resistance  $\epsilon = 0.04$ , then **what is the minimum acceleration that can be measured by this system?** The answer may be expressed in terms of  $V_s$ .

**Solution:** We first rearrange to have  $\epsilon$  in terms of the other variables:

$$\epsilon = \frac{2V_x}{V_s} \quad (7)$$

Since  $V_x$  is, at a minimum,  $1\mu\text{V}$  or  $1 \times 10^{-6}\text{V}$ , the minimum  $\epsilon$  can be is  $\epsilon = \frac{2 \times 10^{-6}}{V_s}$ .

To see what our minimum acceleration is, we write out an equation for acceleration in terms of  $\epsilon$ . We are given that each acceleration change (let's call it  $\Delta a$ ) of  $1g$  corresponds to a change in resistance  $\epsilon = 0.04$ . We therefore know that 1)  $\Delta a$  and  $\epsilon$  are related linearly (that is,  $\Delta a = m * \epsilon + b$ ) and 2) we have two data points on our line:  $(\Delta a_1, \epsilon_1) = (0, 0)$  and  $(\Delta a_2, \epsilon_2) = (g, 0.04)$ .

Since the line goes through the point  $(0, 0)$ , we know  $b = 0$ . Solving for the slope of the line, we find:

$$m = \frac{1g - 0}{0.04 - 0} = \frac{g}{0.04} \quad (8)$$

This gives us the relationship:

$$\Delta a = \frac{g * \epsilon}{0.04} \quad (9)$$

Plugging in our minimum value of  $\epsilon$ , we find that the minimum acceleration is:

$$\Delta a = \frac{2 \times 10^{-6}}{V_s} \left( \frac{g}{0.04} \right) = \frac{5 \times 10^{-5} * g}{V_s} \quad (10)$$

If  $V_s \approx 1\text{V}$ , this means we can measure a very small change in acceleration with this device.

## 2. Why guessing and checking is alright in solving differential equations

In lecture (and possibly in other courses), you have seen differential equations solved by looking at the equation, moving parts around, reasoning about it using an analogy with eigenvalue/eigenspaces, and then seeing that the solution that we proposed actually works — i.e. satisfies all the conditions of the differential equation problem. This process should have felt a bit different than how you have seen how systems of linear equations are solved (by doing Gaussian Elimination) where it was clear that every step was valid. Indeed, it *is* different. Although the eigenvalue/eigenspace analogy to differential equations can be made precise and rigorous, doing that carefully is beyond the scope of this course. In effect, all of that reasoning in between seeing the problem and checking the solution can be considered a kind of inspired guessing.

This should lead you to a natural question — how can we be sure that we have found all of the solutions? We’ve checked to see that the solution we found solves the equations, but maybe there are more solutions that are different. How can we be sure? After all, we are using the solution of the differential equation for its *predictive* power — for example, we are using the fact of RC time constants to argue that this limits the speed of digital computation. Making such inferences is only proper if we have indeed found the only solution to the differential equation.

In the mathematical literature, this is sometimes referred to as the problem of establishing the “uniqueness” of solutions. The concept is also very important for us in engineering contexts. You have already seen in EE16A’s touchscreen module that node voltages need not be unique, and that is why you need to specify a ground in your circuit. You also saw this concept in EE16A’s localization module where you learned how to approach inconsistent linear equations by the method of least squares: you started with no solutions, allowed some error and then got infinitely many potential solutions with error. To make the solution unique, you had to specify that you wanted to minimize the size of the hypothesized error.

This problem walks you through an elementary proof of the uniqueness of solutions to a simple scalar differential equation of the form

$$\frac{d}{dt}x(t) = \alpha x(t) \quad (11)$$

with initial condition

$$x(0) = x_0. \quad (12)$$

Being able to do simple proofs is an important skill, not only in its own right, but also for the systematic logical thinking that it exercises. This problem has multiple parts, but the goal is simply to help you see how you could have come up with this proof entirely on your own.

- (a) **Please verify that the guessed solution  $x_d(t) = x_0 e^{\alpha t}$  satisfies (11) and (12).**

**Solution:** Taking the derivative of  $x_d(t)$  with respect to  $t$  gives  $\alpha x_0 e^{\alpha t}$  by the chain rule, and this is equal to  $\alpha x_d(t)$  by inspection. So (11) is satisfied.

Evaluating  $x_d(0) = x_0 e^{\alpha \cdot 0} = x_0$  and so (12) is also satisfied.

- (b) To show that this solution is in fact unique, we need to consider a hypothetical  $y(t)$  that also satisfies (11) and (12).

Our goal is to show that  $y(t) = x(t)$  for all  $t \geq 0$ . (The domain  $t \geq 0$  is where we have defined the conditions (11) and (12). Outside of that domain, we don’t have any constraints. )

How can we show that two things are equal? In the past, you have probably shown that two quantities or functions are equal by starting with one of them, and then manipulating the expression for it using

valid substitutions and simplifications until you get the expression for the other one. However, here, we don't have an expression for  $y(t)$  so that style of approach won't work.

In such cases, we basically have a couple of basic ways of showing that two things are the same.

- Take the difference of them, and somehow argue that it is 0.
- Take the ratio of them, and somehow argue that it is 1.

We will follow the ratio approach in this problem. First assume that  $x_0 \neq 0$ . In this case, we are free to define  $z(t) = \frac{y(t)}{x_d(t)}$  since we are dividing by something other than zero.

**What is  $z(0)$ ?**

**Solution:**

We know  $z(0) = \frac{y(0)}{x_d(0)} = \frac{x_0}{x_0} = 1$  since  $y(0) = x_0$  by (12) and plugging in 0 for  $t$  into the exact expression for  $x_d(t)$ .

- (c) **Take the derivative  $\frac{d}{dt}z(t)$  and simplify using (11) and what you know about the derivative of  $x_d(t)$ .**

(HINT: The quotient rule for differentiation might be helpful since a ratio is involved.)

**Solution:**

The quotient rule tells us how to take the derivative of  $\frac{y}{x_d}$  (we can also view this using the product rule, which is also just another manifestation of the chain rule for differentiation in the multivariate case). The rule applies because the functions involved are differentiable by definition and the denominator is nonzero.

$$\frac{d}{dt}z(t) = \frac{d}{dt} \frac{y(t)}{x_d(t)} = \frac{\frac{d}{dt}y(t)x_d(t) - y(t)\frac{d}{dt}x_d(t)}{(x_d(t))^2} \quad (13)$$

$$= \frac{\alpha y(t)x_d(t) - y(t)\alpha x_d(t)}{(x_d(t))^2} \quad (14)$$

$$= \frac{0}{(x_d(t))^2} = 0 \quad (15)$$

Notice that here, what is important is that both  $y$  and  $x_d$  satisfy (11) and so the numerator in the quotient rule cancels out to zero. The details of  $x_d(t)$  didn't end up mattering.

You should see that this derivative is always 0 and hence  $z(t)$  does not change. **What does that imply for  $y$  and  $x_d$ ?**

**Solution:** Since  $z(t)$  has zero derivative, it cannot change, and hence it stays at its initial value, which is 1. So it is always 1 and hence  $\frac{y(t)}{x_d(t)} = 1$  so  $y(t) = x_d(t)$ .

- (d) At this point, we have shown uniqueness in most cases. Just one special case is left:  $x_0 = 0$ .

Here, the division approach doesn't seem to work because we are not permitted to divide by zero and  $x_d(t) = 0$ .

However, we want to show that  $y(t) = 0$  here as well.

Fundamentally, the argument we want to make is of the "it can't possibly be otherwise" variety. Consequently, a proof by contradiction can be easier to start.

In such proofs, we start by assuming the thing that we want to show is not possible. So assume that  $y(t)$  is not identically 0 everywhere for  $t > 0$ . What does this mean? This means that there is some  $t_0 > 0$  for which  $y(t_0) = k \neq 0$ . (Otherwise, it would be zero everywhere.)

We want to create a contradiction. It is clear that we will have no easy contradiction if we just move forward for  $t > t_0$  because we have no information given about such solutions  $y(t)$  that we can contradict.

What do we know about? We have (12) which says something about  $y(0)$ . This means, that we need to somehow move backward in time from  $t_0$ . That way, we can hope to contradict the initial condition of 0.

What do we have to work with? Well, we just did some work in the previous parts establishing uniqueness of solutions assuming nonzero initial conditions. How can we view what happens at  $t_0$  as a kind of nonzero initial condition?

**Apply the change of variables  $t = t_0 - \tau$  to (11) to get a new differential equation for  $\tilde{x}(\tau) = x(t_0 - \tau)$  that specifies how  $\frac{d}{d\tau}\tilde{x}(\tau)$  must relate to  $\tilde{x}(\tau)$ . This should hold for  $-\infty < \tau \leq t_0$ .**

**Solution:**

$$\frac{d}{d\tau}\tilde{x}(\tau) = \frac{d}{d\tau}x(t_0 - \tau) \quad (16)$$

$$= -\alpha x(t_0 - \tau) \quad (17)$$

$$= -\alpha\tilde{x}(\tau) \quad (18)$$

where the second line used the chain rule for differentiation and (11).

This holds for all  $t \geq 0$  which means  $t_0 - \tau \geq 0$  which is the same as  $\tau \leq t_0$ .

- (e) Because the previous part resulted in a differential equation of a form for which we have already proved uniqueness for the case of nonzero initial condition, and since  $\tilde{y}(0) = y(t_0) = k \neq 0$ , we know what  $\tilde{y}(\tau)$  must be. **Write the expression for  $\tilde{y}(\tau)$  for  $\tau \in [0, t_0]$  and what that implies for  $y(t)$  for  $t \in [0, t_0]$ . i.e. What does the expression you have derived for  $\tilde{y}(\tau)$  tell you that  $y(t)$  must be for the interval  $t \in [0, t_0]$ ?**

Note: this is all in the context of a proof by contradiction that we are building up. So the expressions you get for  $\tilde{y}(\tau)$  and  $y(t)$  will have dependencies on  $k$  as well as  $\alpha$ .

*(HINT: This problem does not have any drawn figures associated with it. And yet, everything here is very physical. Functions of time, and so on. To better understand what is going on, it is generally a good strategy to draw caricature diagrams that illustrate every single variable in the problem. For functions of time, draw curves. For any variable, draw the relevant axis and label its origin as well as the direction in which it is increasing. For any value (like  $k$  or  $t_0$ ), draw where it is and mark which axis it is on. And so on. Drawing this yourself will help you understand what is going on in the previous part, this part, and the next part. Drawings will help you keep track of the “forest” and not get lost in the “trees.”)*

**Solution:** We know that  $\tilde{y}$  satisfies  $\frac{d}{d\tau}\tilde{y}(\tau) = -\alpha\tilde{y}(\tau)$  and that  $\tilde{y}(0) = k \neq 0$ . Consequently, by the uniqueness theorem already proved, we know that it must be the case that  $\tilde{y}(\tau) = ke^{-\alpha\tau}$  for the range  $\tau > 0$  as long as the differential equation is valid.

This means that  $y(t) = ke^{-\alpha(t_0-t)}$  as long as  $0 \leq t \leq t_0$ .

- (f) **Evaluate  $y(0)$  and argue that this is a contradiction for the specified initial condition (12).**

**Solution:**

Evaluating this expression at  $t = 0$  gives  $y(0) = ke^{-\alpha t_0}$ . Because  $k \neq 0$ , this means  $y(0) \neq 0$ . This is a contradiction with (12) since that asserts a zero initial condition  $x_0 = 0$ .

Consequently, such a  $y(t)$  cannot exist and only the all zero solution is permitted — establishing uniqueness in this case of  $x_0 = 0$  as well.

Although we gave you lots of guidance in this problem, we hope that you can internalize this way of thinking. This elementary approach to proving the uniqueness of solutions to differential equations works for the kinds of linear differential equations that we will tend to encounter in EE16B. For more complicated nonlinear differential equations, further conditions are required for uniqueness (appropriate continuity and differentiability) and proofs can be found in upper-division mathematics courses on differential equations when you study the Picard-Lindelöf theorem. (It involves looking at the magnitude of the difference of the two hypothetical solutions and showing this has to be arbitrarily small and hence zero. However, the basic elementary case we have established here can be viewed as a building block — the quotient rule gets invoked in the appropriate place, etc. The additional ingredients that are out-of-scope for lower-division courses are fixed-point theorems — which you can think of as more general siblings of the intermediate-value theorem you saw in basic calculus.)

### 3. CMOS Scaling

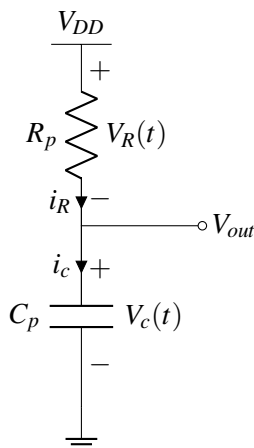
Jerry wants to create a new machine learning accelerator chip using CMOS technology. When designing his chip, he considers the most important parameters of his design to be the amount of energy dissipated when the gate transitions, and the delay time it takes for the output of a gate to hit  $\frac{V_{DD}}{2}$  from either ground or  $V_{DD}$  (i.e. the delay of the gate). These two parameters are very important for CMOS technology, as they determine how quickly the processor can run, and how much power it will consume.

Jerry has access to two different fabrication processes: process A and process B.

Process A uses a supply voltage of  $V_{DD} = 1\text{V}$ . The transistors have a parasitic resistance of  $R_p = 10\text{k}\Omega$ , and the output driven by a representative inverter has a parasitic capacitance of  $C_p = 5\text{fF}$ .

Process B uses a supply voltage of  $V_{DD} = 3\text{V}$ . The transistors have a parasitic resistance of  $R_p = 30\text{k}\Omega$ , and the output driven by a representative inverter has a parasitic capacitance of  $C_p = 1\text{fF}$ .

In order to determine which process is better for the design, Jerry decides to analyze the circuit where the input of an inverter transitions from  $V_{DD}$  to 0. This can be modeled as the following circuit:



Since the input of the inverter is transitioning from  $V_{DD}$  to 0, the initial condition for  $V_c(t)$  is:

$$V_c(0) = 0$$

- (a) In terms of the variables  $V_{DD}$ ,  $R_p$ , and  $C_p$ , **solve for**  $V_{out}(t)$ .

**Solution:**

To find an expression for  $V_{out}(t)$ :

$$V_c(t) = V_{out}(t)$$

KCL at  $V_{out}$  (note the direction of  $i_R$ ) yields:

$$i_R(t) = i_c(t)$$

$$\frac{V_{DD} - V_{out}(t)}{R_p} = C_p \frac{dV_{out}(t)}{dt}$$

$$\frac{dV_{out}(t)}{dt} + \frac{1}{R_p C_p} V_{out}(t) = \frac{V_{DD}}{R_p C_p}$$

Using substitution of variables:

$$x(t) = V_{out}(t) - V_{DD}$$

$$V_{out}(t) = x(t) + V_{DD}$$

$$\frac{dV_{out}(t)}{dt} = \frac{dx(t)}{dt}$$

$$\frac{dx(t)}{dt} + \frac{1}{R_p C_p} x(t) = 0$$

$$x(t) = A e^{-\frac{t}{R_p C_p}}$$

Now substituting back to find  $V_{out}$ :

$$V_{out}(t) = V_{DD} + A e^{-\frac{t}{R_p C_p}}$$

Using our initial condition:

$$V_{out}(0) = 0 = V_{DD} + A$$

$$A = -V_{DD}$$

$$V_{out}(t) = V_{DD} \left( 1 - e^{-\frac{t}{R_p C_p}} \right)$$

- (b) Using the expression for  $V_{out}(t)$  that was just calculated, **solve for**  $i_R(t)$ . Keep this expression in terms of the variables  $V_{DD}$ ,  $R_p$ , and  $C_p$ .

**Solution:**

$$i_R(t) = \frac{V_{DD} - V_{out}(t)}{R_p} = \frac{V_{DD} - V_{DD} \left( 1 - e^{-\frac{t}{R_p C_p}} \right)}{R_p} = \frac{V_{DD} e^{-\frac{t}{R_p C_p}}}{R_p}$$

- (c) In the previous part, you should have noticed that  $i_R(t)$  started at some value, and decayed towards 0 as  $t \rightarrow \infty$ .

**Why does this trend make sense? If the voltage were switching to a different level, would the same trend in current hold?** This question is meant to help build intuition and understanding about switching circuits.

**Solution:** In steady state (i.e. when  $t \rightarrow \infty$ ), we know that the voltage will settle to a constant value. Because  $i_C = C \frac{dV_C}{dt}$ , it makes sense that when the voltage is constant, the current through the capacitor will be 0. In general, this trend will hold true: when there is a step voltage change (in this case, the transistor suddenly switching on/off) in a circuit with a resistor and capacitor, the current through capacitor will jump to some level ( $i_0 = i(t=0)$ ), and will then exponentially decay to reach a steady state current of 0.

- (d) Using the values of  $V_{DD}$ ,  $R_p$ , and  $C_p$  from process A, **calculate the time it takes for  $V_{out}$  to reach  $\frac{V_{DD}}{2}$ .**

**Solution:**

We can find the delay time by setting  $V_{out}(t) = \frac{V_{DD}}{2}$ :

$$\frac{V_{DD}}{2} = V_{DD} \left( 1 - e^{-\frac{t}{R_p C_p}} \right)$$

$$\frac{1}{2} = e^{-\frac{t}{R_p C_p}}$$

$$\ln\left(\frac{1}{2}\right) = -\frac{t}{R_p C_p}$$

$$t = -\ln\left(\frac{1}{2}\right) R_p C_p$$

From this, we can say that the delay time to reach  $\frac{V_{DD}}{2}$  for any  $R_p$  and  $C_p$  is:

$$t_d = 0.69 R_p C_p$$

$$t_d = 0.69(10 \times 10^3 \times 5 \times 10^{-15}) = 3.45 \times 10^{-11} \text{s}$$

- (e) Using the values of  $V_{DD}$ ,  $R_p$ , and  $C_p$  from process A, **calculate the total energy delivered by the voltage source,  $V_{DD}$ , while the capacitor is being charged to  $V_{DD}$ .**

For this problem, recall that the instantaneous power delivered by a voltage source is  $P(t) = I(t) \cdot V(t)$ . Note that the current and voltage are functions of time.

Energy can be found by integrating power:

$$E = \int_{t=0}^{t=\infty} P(t) dt$$

Remember that the units of energy are Joules [J], while the units of power are Watts [W], which is energy per time:  $1\text{W} = \frac{\text{J}}{\text{s}}$

**Solution:**



The total energy delivered by the source can be found by integrating the instantaneous power:

$$\begin{aligned}
 U_s &= \int_0^\infty P(t)dt = \int_0^\infty V_{DD}i_R(t)dt \\
 i_R(t) &= C_p \frac{dV_{out}(t)}{dt} \\
 i_R(t) &= C_p V_{DD} \frac{1}{R_p C_p} e^{-\frac{t}{R_p C_p}} = \frac{V_{DD}}{R_p} e^{-\frac{t}{R_p C_p}} \\
 U_s &= \int_0^\infty (V_{DD}) \left( \frac{V_{DD}}{R_p} e^{-\frac{t}{R_p C_p}} \right) dt \\
 U_R &= \int_0^\infty \frac{V_{DD}^2}{R_p} e^{-\frac{t}{R_p C_p}} dt \\
 U_s &= \left( \frac{V_{DD}^2}{R_p} \right) (-R_p C_p) e^{-\frac{t}{R_p C_p}} \Bigg|_0^\infty
 \end{aligned}$$

The total energy supplied by the supply when charging up the capacitor is:

$$U_s = CV_{DD}^2$$

Plugging in component values of process A for the energy dissipation and time delay:

$$U_s = (5 \times 10^{-15})1^2 = 5 \times 10^{-15} \text{J}$$

**Note:** Notice that the energy supplied by the supply when charging the capacitor is  $U_s = CV_{DD}^2$ . However, we know that the energy stored by a capacitor is  $U_{cap} = \frac{1}{2}CV_{DD}^2$ .

*This implies that the energy drawn from the supply in charging the capacitor is twice the energy that the capacitor finally stores when it is charged.* Where did the other half of the energy go? It is dissipated as heat through the resistor when charging the capacitor. If you do a similar calculation to what you just did, where you integrate the energy dissipated through the resistor as the capacitor is charging (i.e. apply  $P = I \cdot V$  to the resistor and integrate), you will find the missing half of the energy.

- (f) **Repeat parts (d) and (e), but with the values from process B.**

**Solution:**

Using the equations derived above:

$$\begin{aligned}
 U_s &= (1 \times 10^{-15})3^2 = 9 \times 10^{-15} \text{J} \\
 t_d &= 0.69(30 \times 10^3 \times 1 \times 10^{-15}) = 2.07 \times 10^{-11} \text{s}
 \end{aligned}$$

- (g) **Compare the energy and delay of process A and B.**

**Solution:**

Compared to process B, process A dissipates less energy per transition, but has a longer delay time.

- (h) Jerry's friend Pat tells Jerry that with process B, one can reduce  $V_{DD}$  to 2V. However, the reduction in supply voltage increases the parasitic resistance  $R_p$  to 50k $\Omega$ . **Calculate the new delay and energy.**

**Solution:**

$$U_s = (1 \times 10^{-15})2^2 = 4 \times 10^{-15} \text{J}$$

$$t_d = 0.69(50 \times 10^3 \times 1 \times 10^{-15}) = 3.45 \times 10^{-11} \text{s}$$

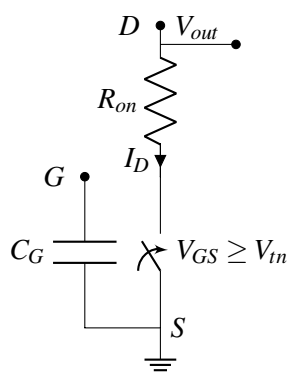
- (i) Based on your previous answers, **which process should Jerry choose to use? Why?**

**Solution:**

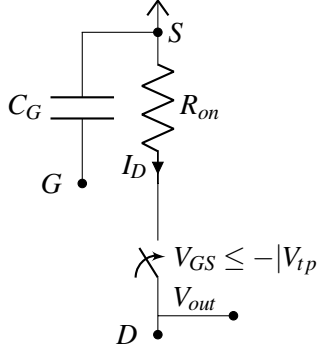
With the new  $V_{DD}$  and  $R_p$  of process B, it ends up that process B and process A have the same delay time. However, process B dissipates less energy per transition with the new  $V_{DD}$  and  $R_p$ , which means Jerry should choose process B which uses the reduced  $V_{DD}$ .

#### 4. Transistor Switch Model

We can improve our resistor-switch model of the transistor by adding in a gate capacitance. In this model, the gate capacitance  $C_G$  represents the lumped physical capacitance present on the gate node of all transistor devices. This capacitance is important as it determines the delay of a transistor logic chain.



a NMOS Transistor Resistor-switch-capacitor model



b PMOS Transistor Resistor-switch-capacitor model. Note we have drawn this so that it aligns with the inverter.

You have two CMOS inverters made from NMOS and PMOS devices. Both NMOS and PMOS devices have an “on resistance” of  $R_{on} = 1 \text{ k}\Omega$ , and each has a gate capacitance (input capacitance) of  $C_G = 1 \text{ fF}$  (femto-Farads =  $10^{-15}$ ). We assume the “off resistance” (the resistance when the transistor is off) is infinite (*i.e.*, the transistor acts as an open circuit when off). The supply voltage  $V_{DD}$  is 1V. The two inverters are connected in series, with the output of the first inverter driving the input of the second inverter (fig. 2).

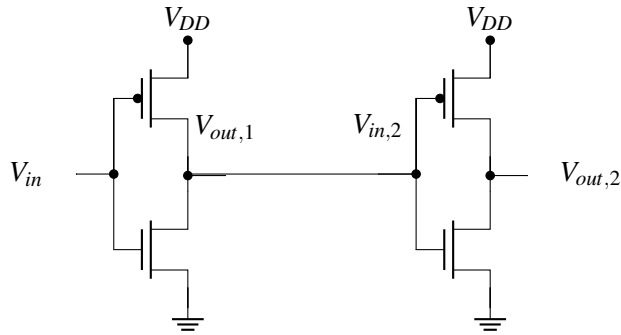


Figure 2: CMOS Inverter chain

- (a) Assume the input to the first inverter has been low ( $V_{in} = 0$  V) for a long time, and then switches at time  $t = 0$  to high ( $V_{in} = V_{DD}$ ). **Draw a simple RC circuit and write a differential equation describing the output voltage of the first inverter ( $V_{out,1}$ ) for time  $t \geq 0$ .** Don't forget that the second inverter is "loading" the output of the first inverter — you need to think about both of them.

**Solution:**

To analyze this circuit as an RC circuit we can recall the transistor switch model. Using this we can see that the first inverter's output appears as a resistor connected to  $V_{DD}$  when the input is low (nmos off, pmos on), or a resistor connected to ground when the input turns high (nmos on, pmos off).

Before  $t = 0$ , the input to the first inverter was low for a long time. This means that for  $t < 0$ , the output of the inverter ( $V_{out,1}$ ) had been held at  $V_{DD}$  for a long time.

At  $t = 0$ , the input goes high, which means that the input inverter's nmos device turns on, connecting  $V_{out,1}$  to ground through a resistance of  $R_{on}$ .

The second inverter "loads" the output of the first inverter. From the notes in the problem, we can model the gates of the transistors as capacitors. These gates together form our capacitive load. The gate of the pmos acts as a capacitor to  $V_{DD}$  and the gate of the nmos acts as a capacitor to ground.

Using this we can draw the following RC circuit:

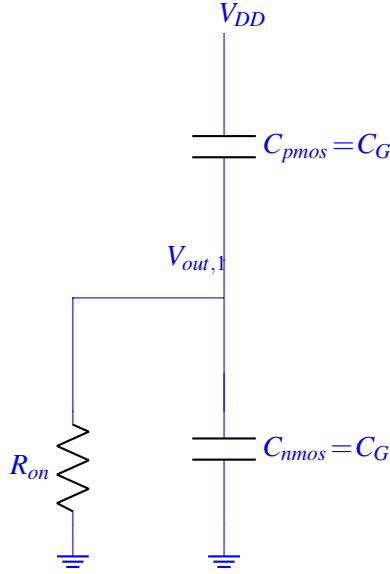


Figure 3: First inverter output at 0

To get the differential equation describing the output of the first inverter at time  $t \geq 0$  let us first think about the behavior of the circuit at and after  $t = 0$ .

Before  $t = 0$  we know that the output  $V_{out,1} = V_{DD}$ . This means that  $C_{nmos}$  is charged, while  $C_{pmos}$  is not as there is no voltage difference across it.

At  $t = 0$ , when the input to the first inverter changes (input switches to high), the nmos will turn on, discharging the  $V_{out,1}$  node. Thus  $V_{out,1}$  will eventually discharge to zero in steady state.

We know the voltage across  $C_{pmos}$  is  $V_{out,1}(t) - V_{DD}$  and the voltage across  $C_{nmos}$  is  $V_{out,1}(t)$ . Using this information we can set up a differential equation to solve for  $V_{out}(t)$ :

$$I_{c_{pmos}} = C_{pmos} \frac{d}{dt} (V_{out,1}(t) - V_{DD}) \quad (19)$$

$$I_{c_{nmos}} = C_{nmos} \frac{d}{dt} V_{out,1}(t) \quad (20)$$

$$I_{R_{on}} = \frac{V_{out,1}(t)}{R_{on}} \quad (21)$$

$$I_{c_{pmos}} + I_{c_{nmos}} = -I_{R_{on}} \quad (22)$$

$$C_{pmos} \frac{d}{dt} (V_{out,1}(t) - V_{DD}) + C_{nmos} \frac{d}{dt} V_{out,1}(t) = -\frac{V_{out,1}(t)}{R_{on}} \quad (23)$$

$$C_{pmos} \frac{d}{dt} V_{out,1}(t) + C_{nmos} \frac{d}{dt} V_{out,1}(t) = -\frac{V_{out,1}(t)}{R_{on}} \quad (24)$$

$$(C_{pmos} + C_{nmos}) \frac{d}{dt} V_{out,1}(t) = -\frac{V_{out,1}(t)}{R_{on}} \quad (25)$$

$$\frac{d}{dt} V_{out,1}(t) = -\frac{V_{out,1}(t)}{R_{on}(C_{pmos} + C_{nmos})} \quad (26)$$

$$\frac{d}{dt} V_{out,1}(t) = -\frac{V_{out,1}(t)}{2R_{on}C_G} \quad (27)$$

- (b) Given the initial conditions in part (a), **solve for**  $V_{out,1}(t)$ .

**Solution:** We know that the solution to a differential equation of the form

$$\frac{d}{dt}V_{out,1}(t) = -\frac{V_{out,1}}{R_{on}(2C_G)}$$

is

$$V_{out,1}(t) = ke^{-\frac{t}{R_{on}(2C_G)}}$$

Plugging in the initial condition  $V_{out,1}(0) = V_{DD}$  we find that  $V_{out,1}(t) = V_{DD}e^{-\frac{t}{R_{on}(2C_G)}}$ .

- (c) **Sketch the output voltage of the first inverter, showing clearly (1) the initial value, (2) the initial slope, (3) the asymptotic value, and (4) the time that it takes for the voltage to decay to roughly 1/3 of its initial value.**

**Solution:**

(1) We know that the output of our inverter started with the initial value  $V_{DD}$ .

(2) Since the differential equation tells us the change in value of  $V_{out,1}(t)$  at time  $t$  we can simply plug in  $t = 0$  into our differential equation to get the initial slope:

$$\frac{d}{dt}V_{out,1}(t) = -\frac{V_{out,1}(0)}{R_{on}(C_{nmos} + C_{pmos})} \quad (28)$$

$$\frac{d}{dt}V_{out,1}(t) = -\frac{V_{DD}}{R_{on}(C_{nmos} + C_{pmos})} \quad (29)$$

Thus the initial slope is  $-\frac{V_{DD}}{R_{on}(C_{nmos} + C_{pmos})} = -\frac{V_{DD}}{R_{on}(2C_G)}$

(3) Since the input to the inverter changed from high to low we know the output of the first inverter ( $V_{out,1}$ ) is going to go to 0 in steady state, as this node will be discharged by the first inverter's nmos transistor.

Alternatively, we can find the asymptotic value by plugging in  $t = \infty$  to the solution we found for  $V_{out,1}(t)$  to find  $V_{out,1} = V_{DD}e^{-\frac{\infty}{R_{on}(2C_G)}} = 0$ .

(4) To approximate when the output will decay to  $\frac{1}{3}$  its original value, we use the fact that  $e^{-1} = \frac{1}{e} \approx \frac{1}{3}$ . We thus want to find when  $V_{out,1} = V_{DD}e^{-1}$ . This will occur when the  $e$  term is raised to  $-1$ , which occurs when  $t = R_{on}(2C_G) = 2 * 10^{-12}$ .

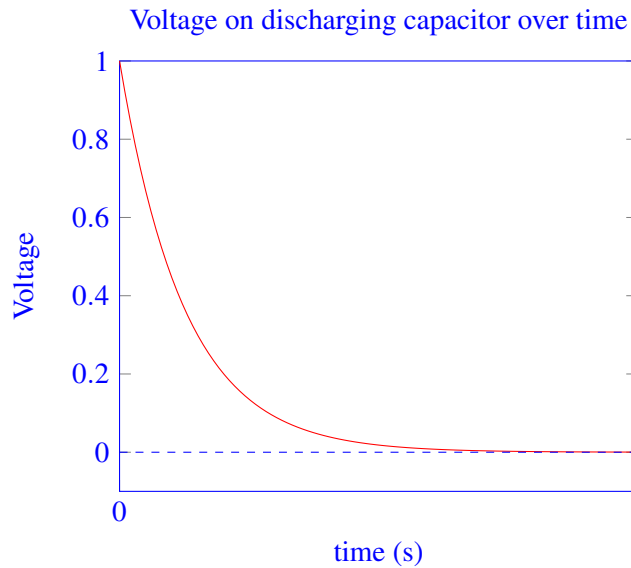


Figure 4

- (d) A long time later, the input to the first inverter switches low again.

**Solve for  $V_{out,1}(t)$ .**

**Sketch the output voltage of the first inverter ( $V_{out,1}$ ), showing clearly (1) the initial value, (2) the initial slope, and (3) the asymptotic value.**

**Solution:**

We know that after a long time, the output of the first inverter has stabilized to 0. When the input switches low again, the input inverter's nmos device turns off, while the input inverter's pmos device turns on. This connects the  $V_{out,1}$  node to  $V_{DD}$ , as shown in fig. 5.

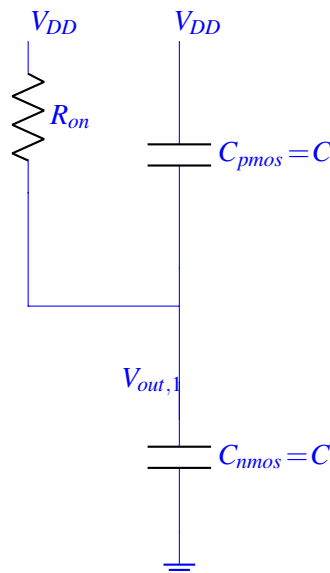


Figure 5: Inverter output at 1

To set up the differential equation, we apply KVL and KCL again:

$$I_{c_{pmos}} = C_{pmos} \frac{d}{dt} (V_{out,1}(t) - V_{DD}) \quad (30)$$

$$I_{c_{nmos}} = C_{nmos} \frac{d}{dt} V_{out,1}(t) \quad (31)$$

$$I_{R_{on}} = \frac{V_{out,1}(t) - V_{DD}}{R_{on}} \quad (32)$$

$$I_{c_{pmos}} + I_{c_{nmos}} = -I_{R_{on}} \quad (33)$$

$$C_{pmos} \frac{d}{dt} (V_{out,1}(t) - V_{DD}) + C_{nmos} \frac{d}{dt} V_{out,1}(t) = -\frac{V_{out,1}(t) - V_{DD}}{R_{on}} \quad (34)$$

$$C_{pmos} \frac{d}{dt} V_{out,1}(t) + C_{nmos} \frac{d}{dt} V_{out,1}(t) = -\frac{V_{out,1}(t) - V_{DD}}{R_{on}} \quad (35)$$

$$(C_{pmos} + C_{nmos}) \frac{d}{dt} V_{out,1}(t) = -\frac{V_{out,1}(t) - V_{DD}}{R_{on}} \quad (36)$$

$$\frac{d}{dt} V_{out,1}(t) = -\frac{V_{out,1}(t) - V_{DD}}{R_{on}(C_{pmos} + C_{nmos})} \quad (37)$$

$$\frac{d}{dt} V_{out,1}(t) = -\frac{V_{out,1}(t) - V_{DD}}{2R_{on}C_G} \quad (38)$$

We will use substitution of variables:

$$x(t) = V_{out,1}(t) - V_{DD} \quad (39)$$

$$V_{out,1}(t) = x(t) + V_{DD} \quad (40)$$

$$\frac{d}{dt} x(t) = \frac{d}{dt} V_{out,1}(t) \quad (41)$$

Substituting in:

$$\frac{d}{dt} x(t) = -\frac{x}{2R_{on}C_G} \quad (42)$$

$$x(t) = Ae^{-\frac{t}{2R_{on}C_G}} \quad (43)$$

Substituting again for  $x(t)$ :

$$V_{out,1}(t) = V_{DD} + Ae^{-\frac{t}{2R_{on}C_G}}$$

Using the initial condition  $V_{out,1} = 0$  (as the input to the first inverter was high for a long time before switching low) implies  $A = -V_{DD}$ . Thus:

$$V_{out,1}(t) = V_{DD} \left( 1 - e^{-\frac{t}{2R_{on}C_G}} \right)$$

(1) Because the input to the first inverter was high for a long time, we know the initial value of  $V_{out,1}(t) = 0$ . This was the initial condition applied to the solution of the differential equation, above.

(2) To find the initial value of the slope we can plug in  $t = 0$  to the above differential equation:

$$\frac{d}{dt} V_{out,1}(t) = \frac{(V_{DD} - V_{out,1}(0))}{R_{on}(2C_G)}$$

where  $V_{out,1}(0) = 0$ . Thus our initial slope is  $\frac{(V_{DD})}{R_{on}(2C_G)}$ . Notice this slope is positive while the previous part had a negative slope.

(3) Since the input to the inverter changed from low to high and the input inverter's pmos is now on, we know the output of the first inverter is going to go to  $V_{DD}$  in steady state.

Alternatively, we can find the asymptotic value by plugging in  $t = \infty$  to the solution we found for  $V_{out,1}(t)$  to find  $V_{out,1} = V_{DD} \left(1 - e^{-\frac{\infty}{R_{on}(2C_G)}}\right) = V_{DD}(1 - 0) = V_{DD}$ .

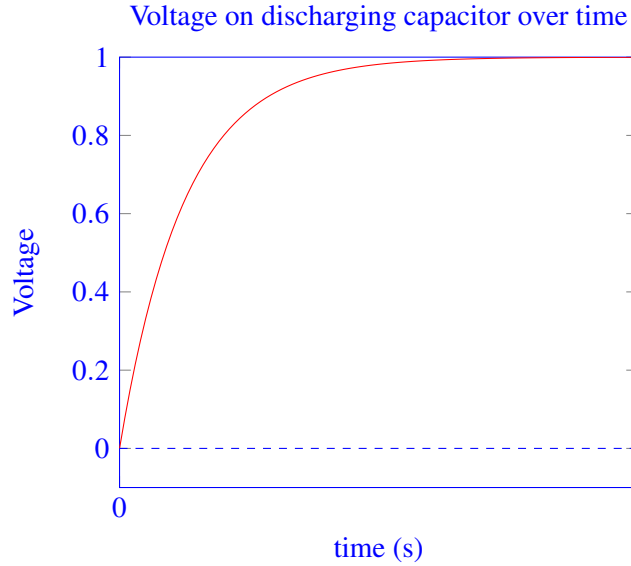


Figure 6

- (e) For each complete input cycle described above ( $V_{in} = 0V \rightarrow 1V \rightarrow 0V$ ), **how much charge is pulled out of the power supply?** Give both a symbolic and numerical answer. **Consider only the charge needed to charge up the  $V_{out,1}$  node.**

**Solution:**

To find the charge required from the supply, we can integrate the current required from the supply during each phase of the cycle ( $Q = \int_0^\infty I_{V_{DD}}(t)dt$ ).

During the input step from  $V_{in} = 0$  to  $V_{in} = V_{DD}$ , we know that the voltage is  $V_{out,1}(t) = V_{DD}e^{-\frac{t}{2C_G R_{on}}}$ . We then get:

$$\begin{aligned} I_{V_{DD}} &= I_{C_{pmos}} = C_G \frac{d}{dt} (V_{out,1}(t) - V_{DD}) \\ &= C_G \frac{-1}{2C_G R_{on}} V_{DD} e^{-\frac{t}{2C_G R_{on}}} \end{aligned}$$

Thus:

$$\begin{aligned} Q_{0 \rightarrow 1} &= \int_0^\infty I_{C_{pmos}}(t)dt = \int_0^\infty C_G \frac{-1}{2C_G R_{on}} V_{DD} e^{-\frac{t}{2C_G R_{on}}} dt \\ &= C_G \cdot V_{DD} e^{-\frac{t}{2C_G R_{on}}} \Big|_0^\infty \\ &= C_G \cdot V_{DD} (0 - 1) = -C_G \cdot V_{DD} \end{aligned}$$



During the input step from  $V_{in} = V_{DD}$  to  $V_{in} = 0$ , we know that the voltage is  $V_{out,1}(t) = V_{DD} \left(1 - e^{-\frac{t}{2R_{on}C_G}}\right)$ . The current from the supply will be equal to the sum of the resistor current and PMOS gate capacitor current:  $I_{V_{DD}} = I_R + I_{C_{pmos}}$ . We get:

$$\begin{aligned} I_{V_{DD}} &= I_{C_{pmos}} + I_R = C_G \frac{d}{dt}(V_{out,1}(t) - V_{DD}) + \frac{V_{out,1}(t) - V_{DD}}{R_{on}} \\ &= -C_G \frac{-1}{2C_G R_{on}} V_{DD} e^{-\frac{t}{2C_G R_{on}}} + \frac{V_{DD} \left(1 - e^{-\frac{t}{2R_{on}C_G}}\right) - V_{DD}}{R_{on}} \\ &= -C_G \frac{-1}{2C_G R_{on}} V_{DD} e^{-\frac{t}{2C_G R_{on}}} + \frac{-V_{DD} e^{-\frac{t}{2R_{on}C_G}}}{R_{on}} \end{aligned}$$

Thus:

$$\begin{aligned} Q_{1 \rightarrow 0} &= \int_0^\infty I_{C_{pmos}}(t) + I_R(t) dt = \int_0^\infty -C_G \frac{-1}{2C_G R_{on}} V_{DD} e^{-\frac{t}{2C_G R_{on}}} + \frac{-V_{DD} e^{-\frac{t}{2R_{on}C_G}}}{R_{on}} dt \\ &= -C_G V_{DD} e^{-\frac{t}{2C_G R_{on}}} \Big|_0^\infty + \frac{-V_{DD}}{R_{on}} \cdot -1 \cdot 2R_{on} C_G e^{-\frac{t}{2R_{on}C_G}} \Big|_0^\infty \\ &= -C_G V_{DD} (0 - 1) + 2C_G V_{DD} (0 - 1) \\ &= -C_G V_{DD} \end{aligned}$$

The total charge is thus

$$Q_{total} = Q_{0 \rightarrow 1} + Q_{1 \rightarrow 0} = -C_G V_{DD} - C_G V_{DD} = -2C_G V_{DD}$$

Note that the current direction for  $I_{V_{DD}}$  was pointing into the  $V_{DD}$  source, so the charge represents the charge moved **into** the power supply. As the question asks for the charge pulled out of the power supply, we know:

$$Q_{\text{pulled out of power supply}} = -Q_{total} = 2C_G V_{DD} = 2(1\text{fF} \cdot 1\text{V}) = 2\text{fC}$$

#### Alternative solution:

During the input step when  $V_{in} = 0$ , note that  $V_{out,1}$  is connected to  $V_{DD}$  through the input inverter's pmos. Thus, during this phase, the power supply is supplying charge to change the node voltage at node  $V_{out,1}$ . We use the equation

$$Q = CV$$

noting that  $V_{initial} = 0$  and  $V_{final} = V_{DD}$ .

Thus:

$$\begin{aligned} Q &= CV \\ &= (C_{nmos} + C_{pmos}(V_{DD} - 0)) \\ &= 2C_G V_{DD} \\ &= 2(1\text{fF} \cdot 1\text{V}) = 2\text{fC} \end{aligned}$$

During the input step when  $V_{in} = 1$ , note that  $V_{out,1}$  is connected to ground through the input inverter's nmos. In this case, the supply is not providing any charge to the  $V_{out,1}$  node. Rather, the charge on this node is being moved to ground through the nmos. Thus there is no charge during this input step.

For the entire input cycle, we thus find that  $Q = 2C_G V_{DD} = 2fC$ .

## 5. Write Your Own Question And Provide a Thorough Solution.

Writing your own problems is a very important way to really learn material. The famous “Bloom’s Taxonomy” that lists the levels of learning is: Remember, Understand, Apply, Analyze, Evaluate, and Create. Using what you know to create is the top level. We rarely ask you any homework questions about the lowest level of straight-up remembering, expecting you to be able to do that yourself (e.g. making flashcards). But we don’t want the same to be true about the highest level. As a practical matter, having some practice at trying to create problems helps you study for exams much better than simply counting on solving existing practice problems. This is because thinking about how to create an interesting problem forces you to really look at the material from the perspective of those who are going to create the exams. Besides, this is fun. If you want to make a boring problem, go ahead. That is your prerogative. But it is more fun to really engage with the material, discover something interesting, and then come up with a problem that walks others down a journey that lets them share your discovery. You don’t have to achieve this every week. But unless you try every week, it probably won’t ever happen.

## 6. Homework Process and Study Group

Citing sources and collaborators are an important part of life, including being a student! We also want to understand what resources you find helpful and how much time homework is taking, so we can change things in the future if possible.

- (a) **What sources (if any) did you use as you worked through the homework?**
- (b) **Who did you work on this homework with?** List names and student ID’s. (In case of homework party, you can also just describe the group.)
- (c) **How did you work on this homework?** (For example, *I first worked by myself for 2 hours, but got stuck on problem 3, so I went to office hours. Then I went to homework party for a few hours, where I finished the homework.*)
- (d) **Roughly how many total hours did you work on this homework?**

### Contributors:

- Alex Devonport.
- Kris Pister.
- Regina Eckert.
- Anant Sahai.
- Nikhil Shinde.
- Kyle Tanghe.
- Sidney Buchbinder.