

1 Introduction

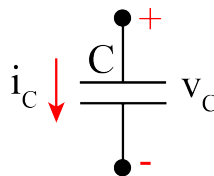
So far we learned that the key enabler of all the digital electronics around us such as smart phones, laptops, etc. is the Transistors! Now that we know how digital circuits and CMOS gates operate, we can ask many questions like:

- How fast a computer operates? (How fast a CMOS gate switch?)
- Why smart phones ran out of the battery quickly?! (How much energy do CMOS gates consume when switching (or even when not)?)
- What limits the complexity of processing can be done with this circuits? How much space do they take up?

We will tackle the first two question regarding switching speed and energy in this lecture, but first we need a quick review on capacitors.

2 Capacitors

Any time we have two conductive materials separated by a dielectric (i.e. an insulator), we have the potential to store electrical charge across the two conductors. This is called a **capacitance** (the device is a **capacitor**) and this way we can store energy. Examples are touch-screen pixels from EE16A labs, two metal wires close to each other, etc.

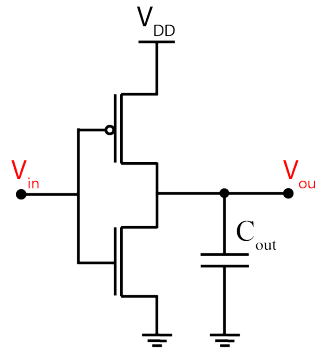


A capacitor's symbol and its voltage/current is shown above. The following is always true about the capacitors:

1. The stored electrical charge can be derived as : $q_C(t) = Cv_C(t) \Rightarrow i_C(t) = C \frac{dv_C(t)}{dt}$ [Amps] (This hold since by definition $i = \frac{dq}{dt}$)
2. $v_C(t)$ can never change instantly (i.e. it can never be discontinuous). To see why, examine 1; if $v_C(t)$ is discontinuous, then $\frac{dv_C(t)}{dt} \rightarrow \infty$ at then discontinuity and $i_c \rightarrow \infty$ which is not feasible (why?).
3. The energy stored in a capacitor at any instant in time is: $U_C = \frac{1}{2}Cv_C^2(t)$ [Joules]

3 Capacitors & Transistors

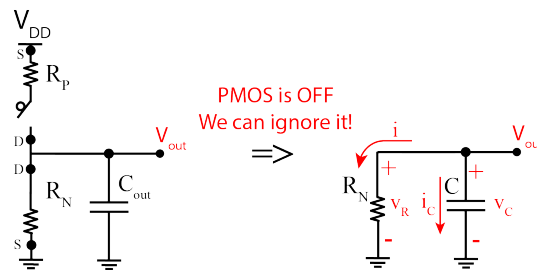
It turns out that whenever we make a transistor, there are always capacitances associated with the nodes. This is unavoidable and unwanted and arises from solid state physics (explained in the last lecture transistor's physics section). In this class, we will model these capacitances as a single capacitor between V_{out} and GND . This capacitance is mainly contributed by the capacitances of NMOS and PMOS transistors of the logic gate (inverter in this example) itself plus the transistor capacitances of the next logical gate loading this stage (we assumed another inverter in this lecture). The actual wires connecting gates together will also add to this capacitance. This is a simple model, but it works surprisingly well enough to answer questions raised at the beginning of this lecture.



4 RC Circuits

Let's look at the inverter illustrated above more carefully and assume $V_{in} = 0V$ and thus $V_{out} = V_{DD}$ and the transistor has been in this state for a long time (since $t = -\infty$).

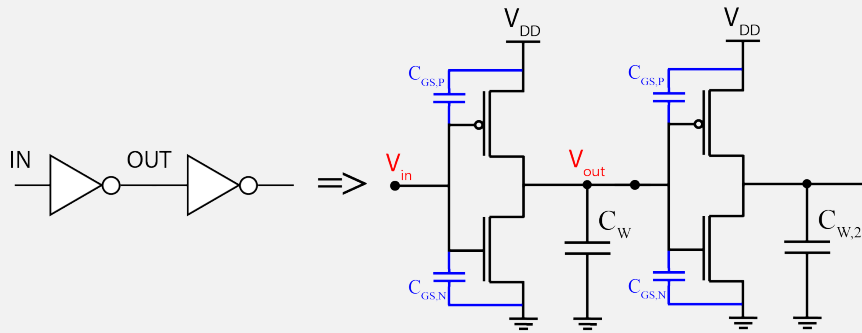
Now, at $t = 0$, we instantly switch V_{in} to V_{DD} . We know, eventually $V_{out} = 0V$. *if there were no capacitance, the output would change instantly because the equations have no concept of time!* but considering the capacitance (C_{out}), things slow down, let's see why? For $t > 0$:



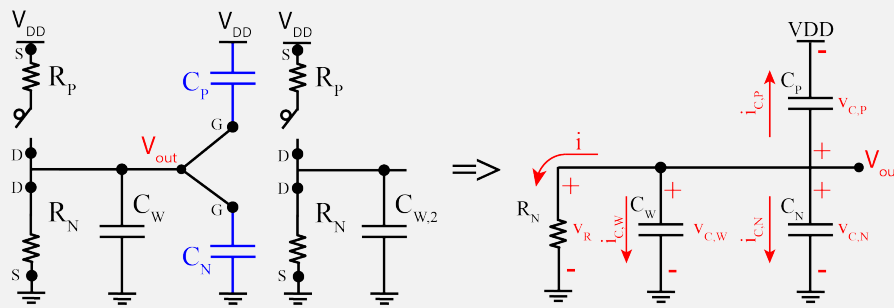
Notice $v_{out}(t) = v_C(t)$ and $i = -i_c = -C \frac{dv_C}{dt}$. Using KVL we know $v_C = v_R$. By substituting capacitor's current equation and ohm's law ($v_R = iR$), we have:

$$v_C = iR = -RC \frac{dv_C}{dt} \Rightarrow v_C + RC \frac{dv_C}{dt} = 0 \Rightarrow v_{out} + RC \frac{dv_{out}}{dt} = 0 \tag{1}$$

Here we will see actually how can transistor capacitors lumped together and modeled as a single capacitor at the output (C_{out}). Assume the logic gate, an inverter, is followed by another inverter:



Using the transistor's model with resistor and capacitor, the circuit model is the following:



Now we can write the KCL and use ohm's law:

$$i + i_{C,W} + i_{C,N} + i_{C,P} = 0 \Rightarrow \frac{v_R}{R} + C_W \frac{dv_{C,W}}{dt} + C_N \frac{dv_{C,N}}{dt} + C_P \frac{dv_{C,P}}{dt} = 0$$

Again, since PMOS is OFF, it does not impact the final circuit model in this case. Now we can replace voltages in terms of v_{out} and V_{DD} :

$$\frac{v_{out}}{R} + C_W \frac{dv_{out}}{dt} + C_N \frac{dv_{out}}{dt} + C_P \frac{d(v_{out} - V_{DD})}{dt} = 0$$

Notice $\frac{d(v_{out} - V_{DD})}{dt} = \frac{dv_{out}}{dt} - \frac{d(V_{DD})}{dt}$ and since V_{DD} is at a fixed value which does not change in time $\frac{d(V_{DD})}{dt} = 0$, thus the RC equation for this circuit can be reduced to:

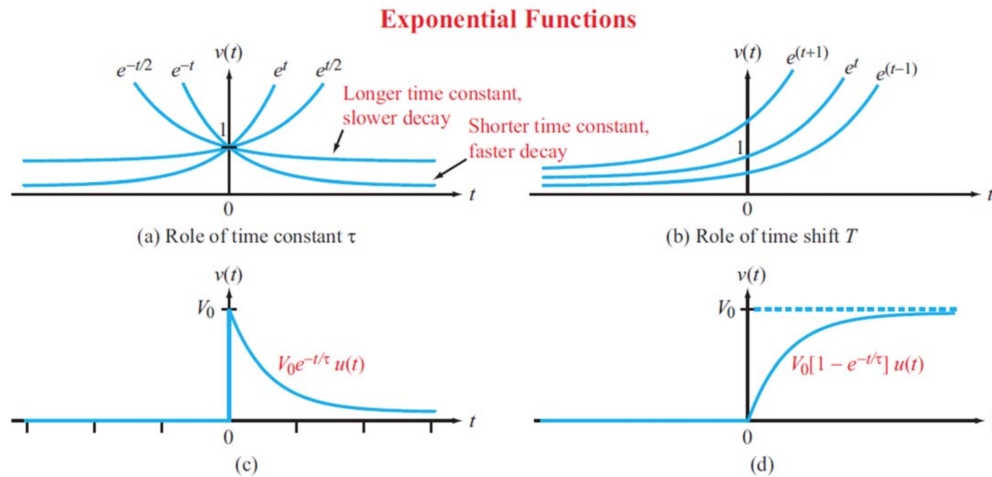
$$\frac{v_{out}}{R} + C_W \frac{dv_{out}}{dt} + C_N \frac{dv_{out}}{dt} + C_P \frac{dv_{out}}{dt} = 0 \Rightarrow v_{out} + R(C_W + C_N + C_P) \frac{dv_{out}}{dt} = 0$$

So we can simply use a single capacitance model with $C_{out} = C_W + C_N + C_P$ and the differential equations is still a homogeneous 1st order D.E.

This is a **first order differential equation (D.E.)**. The general form is: $\frac{dx}{dt} + ax = 0$. Because in this form the right hand side (RHS) is zero, it's called a **homogenous 1st order D.E.**. The notation of derivatives are sometimes shown by dots as well ($\frac{dx}{dt} = \dot{x}$, $\frac{d^2x}{dt^2} = \ddot{x}$, and so on).

5 Solving 1st Order D.E.

As stated above these equations have a general form of: $\frac{df}{dt} + af = 0$. If we rewrite it in the following format: $\frac{d}{dt}(f(t)) = -af(t)$. What will this equation remind you from linear algebra topics studied in 16A? **Eigenvalues!**. Remember the eigenvectors of a matrix A were solutions of $Av = \lambda v$ equation. In the differential equations, $\frac{d}{dt}$ can be seen as a derivative operator and thus here we are looking for eigenfunctions of the derivative operator with the eigenvalues of $-a$. It turns out that the eigenfunctions are in the general form of an **exponential function**: $f(t) = e^{kt}$. Next figure illustrates various forms of these functions for different k values ($u(t)$ denotes a step function in this figure):



The unique property of the exponential functions is that:

$$\frac{df(t)}{dt} = k \cdot e^{kt} = kf(t)$$

Hence a function in the form of $x(t) = c_1 e^{kt} + c_0$ can be an answer to our 1st order D.E.:

$$\frac{df(t)}{dt} = -af(t) \Rightarrow c_1 k e^{kt} = -ac_1 e^{kt} - ac_0$$

Since this equation should hold for every t , $c_0 = 0$ and $k = -a$ (Prove this yourself). **Thus the solution to $\frac{df}{dt} + af = 0$ is in the form of ce^{-at} , where c can be any number.**

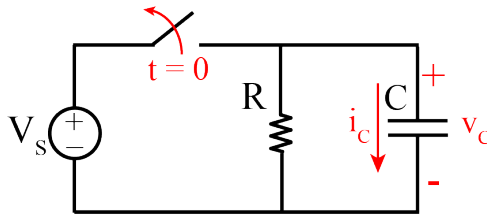
Going back to the RC equation for v_{out} , the solution for $t > 0$ will be:

$$v_{out}(t) = v_{out}(0)e^{-t/RC}$$

where $v_{out}(0)$ is the output voltage at $t = 0$ (remember the voltage across a capacitor should transit continuously). In the case of an inverter $v_{out}(0) = VDD$ since input was low before switching. The value of RC is called **time constance** and shown with $\tau = RC$. This parameter determines how fast the exponential decays.

6 Natural Response of a Charged Capacitor

Let's analyze the RC circuit, where a capacitor is charged to V_S and for $t > 0$ it starts discharging through a resistor of R .



As we already saw, the voltage over the capacitor is:

$$v_C(t) = V_S e^{-t/\tau}$$

where $\tau = RC$. Capacitor's current can be derived using $i_C(t) = C \frac{dv_C(t)}{dt}$ as:

$$i_C(t) = C \frac{dv_C(t)}{dt} = -C \frac{V_S}{\tau} e^{-t/\tau} = -\frac{V_S}{R} e^{-t/\tau}$$

Stored energy in the capacitor can be expressed as:

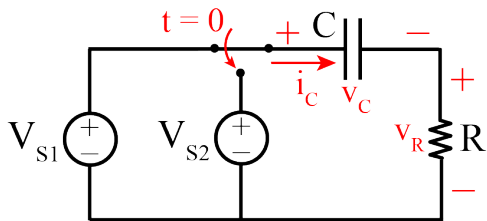
$$U_C(t) = \frac{1}{2} C v_C^2(t) = \frac{1}{2} C V_S^2 e^{-2t/\tau}$$

And the power dissipated in the circuit is:

$$P(t) = i_C(t) v_C(t) = -\frac{V_S^2}{R} e^{-2t/\tau}$$

7 General Solution of RC Circuits with DC Sources

Now consider a new circuit where R and C are connected in series, the capacitor is charged to V_{S1} and at the time $t = 0$, we switch the capacitor's terminal to another voltage source with V_{S2} volt.



Before the switch is flipped at time $t = 0$, $v_C(0) = V_{S1}$ (why?). Now we can write KVL for $t > 0$ to find the differential equation associated with this circuit:

$$-V_{S2} + v_R + v_C = 0 \Rightarrow -V_{S2} + i_C R + v_C = 0 \Rightarrow -V_{S2} + RC \frac{dv_C}{dt} + v_C = 0 \Rightarrow \frac{dv_C}{dt} + \frac{1}{RC} v_C = \frac{V_{S2}}{RC}$$

This D.E. is inhomogeneous in this case since the RHS is non-zero. The general form of an inhomogeneous 1st order D.E. is:

$$v_C' + a v_C = b$$

There are two ways to derive the solution to this equation:

$$1. v_C(t) = v_C(0)e^{-t/RC} + \frac{b}{a}(1 - e^{-t/RC})$$

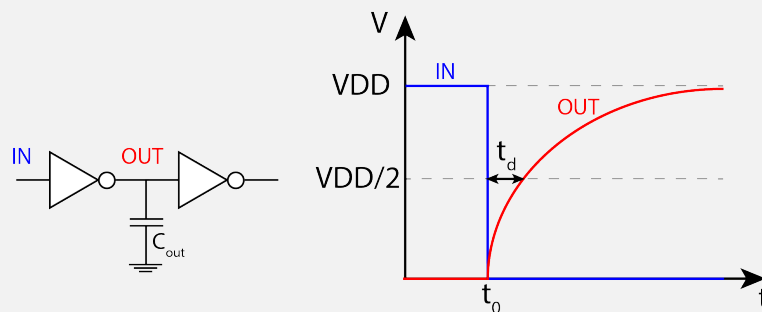
The first term is the decaying exponential of the initial voltage due to V_{S1} source. The second term is the exponential of the capacitor charging up to V_{S2} . Notice that at $t \rightarrow \infty$, the capacitors will look like an open (because there is only a DC source), so by KVL $V_{S2} = v_C(\infty)$ (Note that $b/a = V_{S2} = v_C(\infty)$). In other words, the solution is the linear superposition of the initial charge of the capacitor discharging and the V_{S2} supply charging the capacitor up to V_{S2} .

$$2. v_C(t) = v_C(\infty) + [v_C(0) - v_C(\infty)]e^{-t/RC}$$

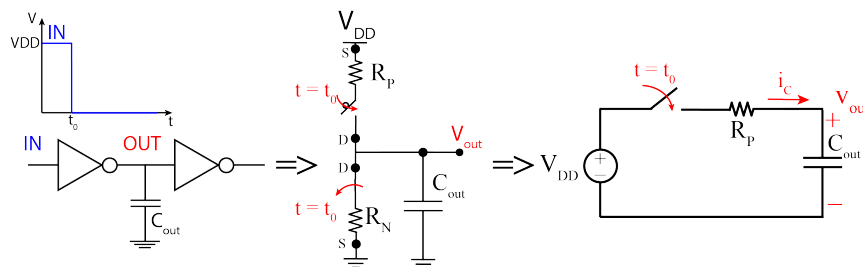
The first term is the steady state value and the other term is the transient response of $v_C(t)$. This is equivalent to the previous solution, but is often faster to write if you know $v_C(0)$, $v_C(\infty)$, and RC . All three of them are sometimes discoverable by inspection.

Finally, knowing how RC circuits work and their transient response, we can answer the questions we asked at the beginning of this lecture. Regarding the computation speed, we learned that the inverter or any other logic gate cannot switch infinitely fast. Due to the logic delays to settle the output voltage the computation cannot run very fast neither!

The delay of each logic gate is always defined as being from 50% of the input signal to the 50% of the output. In CMOS gates, we know that values will be finally either $0V$ or V_{DD} and thus the delay, t_d is normally defined as the time difference of input and output reaching $V_{DD}/2$. This definition is used because we do not just have one gate, but actually many of them in a chain, and we want to be able to say that the delay of the chain is the sum of the delays of the individual stages (which implies that you need to select delay as being 50% to 50%).



Going back to the second question on the energy required for a logic gate to operate, let's see how much energy we need to switch an inverter gate (the simplest CMOS gate)?



The energy required to charge the capacitor to V_{DD} is provided by the voltage source (V_{DD}). There are two approaches to calculate it:

1. We know the total charge needed to charge up the capacitance to V_{DD} is $Q_C = C_{out} \cdot V_{DD}$. By definition the energy required to move the Q charge between two points with V_{DD} voltage across is $Q \cdot V_{DD}$. Thus total dissipated energy by the supply to charge the capacitor is $W_{DD} = C_{out} V_{DD}^2$.
2. We know the power given by any element to the circuit is $i(t)v(t)$. Also, the energy can be calculated by integrating power over time, hence:

$$\begin{aligned} W_{DD} &= \int i_{DD}(t)v_{DD}(t) dt = \int i_C(t)V_{DD} dt = V_{DD} \int \frac{dq_C(t)}{dt} dt = V_{DD}(q_C(\infty) - q_C(0)) \\ &\Rightarrow W_{DD} = V_{DD}(C_{out}V_{DD} - 0) = C_{out}V_{DD}^2 \end{aligned}$$

Question: What's the energy stored in the capacitor after charging it to V_{DD} ? Where does the rest of energy dissipated by the battery go?

Consequently, any computation and process step requires certain amount of energy and this energy is drawn from the supplies (like batteries).