

## 1 Introduction to Transistors

### 1.1 Motivation

We motivate this section by seeking to understanding the speed of digital computers. We can understand computers as a set of blocks that perform digital logic operations, one after the another. You might be familiar with logic in the form of 'AND' and 'OR' operations in 'if' statements while programming. Logic gates are circuits that behave in a manner compatible with those logical operations. For this, high voltages are traditionally used to represent a logical 1 or TRUE and low voltages are traditionally used to represent a logical 0 or FALSE. These logical gate circuits are constructed out of physical devices called transistors. You will learn a lot more about digital circuits and how to construct logical gates out of transistors in 61C.

Here in 16B, we are going to focus on one of the most basic logical blocks: an inverter (a 'NOT' gate). An inverter takes a boolean input and outputs the logical inverse: a 0 (low) maps to a 1 (high) and a 1 (high) maps to a 0 (low). This note will cover the building blocks of inverters, PMOS and NMOS transistors, as well as how to analyze the behavior of these building blocks. The speed of computers is related to how quickly logical blocks can change their state and thus perform logic (for example, how quickly an inverter can output a 0 after its input changes to a 1). The underlying issues that limit the speed of an inverter are the same issues that impact all logical operations, so to understand what is going on, we will focus on inverters for simplicity.

In relation to inverters, we will also demonstrate the oscillator circuit, which oscillates between zero and one. In fact such oscillators are quite common and are used as clocks in all the devices you regularly use! We can analyze the speed and behavior of basic oscillator models to understand the more complex behavior of computers and their speed.

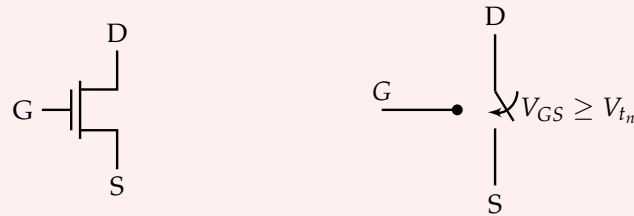
### 1.2 Transistor Switch Models

#### Definition 1 (Terminals of a Transistor)

There are three terminals on a transistor: a source terminal  $S$ , a drain terminal  $D$ , and a gate terminal  $G$ . The voltage on the Gate determines whether the switch connecting the source and the drain is on (closed) or off (open). For transistors, when we talk about the gate voltage, we are talking about the voltage at the gate relative to the voltage at the source. This is usually denoted  $V_{GS} = V_G - V_S$ .

#### Definition 2 (Switch Model of an NMOS Transistor)

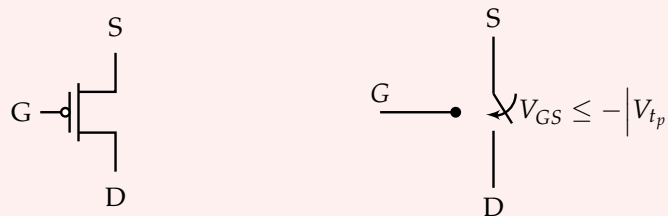
The NMOS switch model is depicted in fig. 1. The switch is on when the voltage between the gate and source,  $V_{GS} = V_G - V_S$ , is above a predetermined threshold  $V_{t_n}$ . Notice the convention: in NMOS Transistors, the source terminal is at the bottom.



**Figure 1:** NMOS Transistor Resistor-switch model.

**Definition 3** (Switch Model of a PMOS Transistor)

The PMOS switch model in fig. 2 is on when the voltage between the gate and source,  $V_{GS} = V_G - V_S$ , is below a predetermined threshold  $V_{t_p}$  (where  $V_{t_p} \leq 0$  by definition). Notice the convention: in PMOS Transistors, the source terminal is at the top.



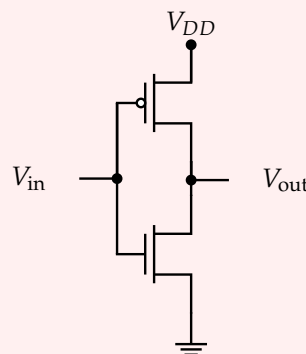
**Figure 2:** PMOS Transistor Resistor-switch model.

In definition 3, we define the condition for a PMOS transistor being “on” in terms of  $V_{GS}$  and  $|V_{t_p}|$  to avoid any sign errors, since  $V_{t_p} \leq 0$ .

### 1.3 Inverters and Oscillators

**Definition 4** (CMOS Inverter)

A CMOS inverter can be created using PMOS and NMOS transistors as shown in fig. 3.



**Figure 3:** CMOS Inverter

The inverter behaves like a 'NOT' gate. As such, we will also represent inverters using the 'NOT' gate symbol in fig. 4.

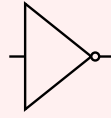


Figure 4: NOT Gate

We can use this circuit element to build oscillators. One way to create oscillators is by connecting together an odd prime number of inverters in a loop. Connecting an inverter in a loop will misbehave<sup>1</sup>. This type of oscillator is called a ring oscillator. By examining the signal in this oscillator after any inverter, we can see that the signal must indeed oscillate between 0 and 1. We can create these inverters physically as shown in Definition 5.

**Definition 5 (Ring Oscillator)**

A *ring oscillator* is created by connecting an odd prime number of inverters in a loop. An example of a ring oscillator using NOT gates is shown in fig. 5, and an equivalent representation using the CMOS inverter diagram is shown in fig. 6.

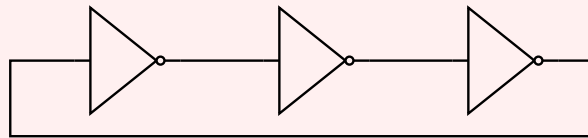


Figure 5: Ring oscillator with 3 inverters

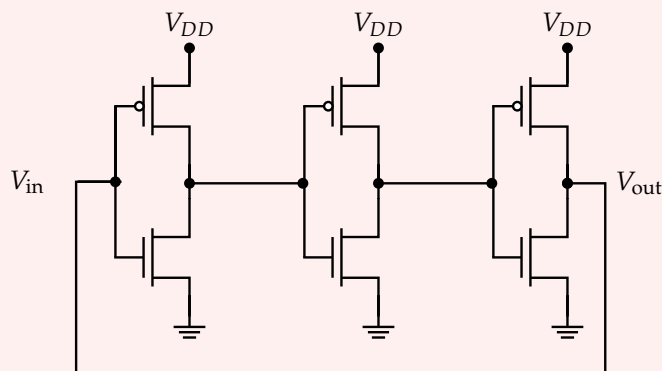


Figure 6: Ring oscillators with 3 CMOS inverters

<sup>1</sup>If we simply have one inverter connected in a loop, we will not have the switching behavior of the oscillator that we desire (depending on if the capacitor is appropriately sized). Since the circuit is fighting between high and low at the output it can stabilize at an intermediate value. In order to allow for oscillations, we need to chain more inverters in a loop. In fact, to prevent undesired behavior, we usually chain together a prime number of inverters. The reason why is related to properties of modulo arithmetic which you will learn in CS70 together with properties of signals studied in EE120.

**Key Idea 6** (Switch Model Contradiction)

Using the switch model of transistors, each inverter switches instantaneously. If each inverter switched instantaneously, then connecting them in a loop with an odd number of inverters would lead to inconsistent behavior. Using the example in fig. 5, if the input to the first inverter is 0, the output of the third inverter would instantaneously change to 1. However, the input of the first inverter and the output of the third inverter share a node in the circuit, so this gives us a contradiction.

Since we can implement this circuit in the real world, there must be some aspect of reality that is missing in the switch model. When such inconsistencies arise, this can be a symptom of failing to properly understand real world behavior. In such cases, we should approach the problem with a more detailed model. The oscillating behavior that we see is actually possible because there is a slight delay between the input and output of the inverters. The slight delay between the transition of the transistor being completely “off” and completely “on” enables the oscillatory behavior that we see.

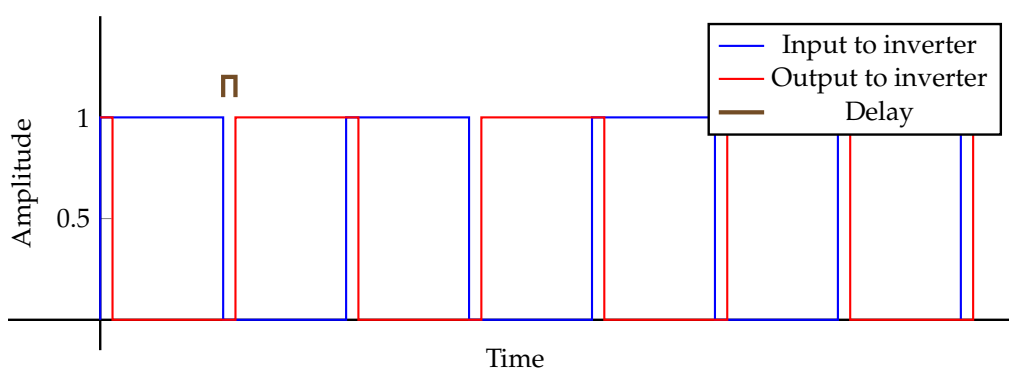


Figure 7: Delay in inverter output for simplified model.

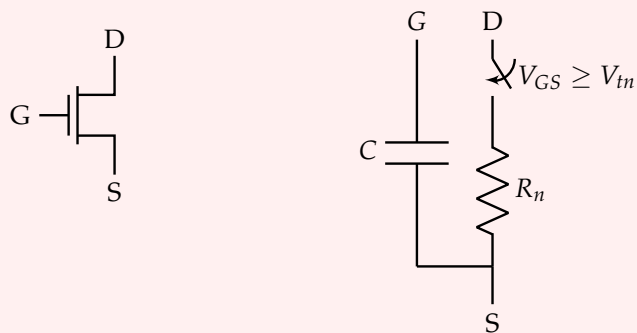
## 1.4 Transistor Resistor-Switch Models

Since our switch model is not enough to understand this delayed behavior, we adopt a more detailed resistor-capacitor model for transistors. We model transistors as having some resistance and some capacitance from their gates.<sup>2</sup> These models are illustrated in definition 7 and definition 8.

**Definition 7** (Resistor-Capacitor Model of an NMOS Transistor)

The diagram in fig. 8 describes the resistor-switch model of an NMOS transistor. The switch is on when  $V_{GS}$  is greater than the threshold voltage. Notice the convention: in NMOS Transistors, the source terminal is at the bottom.

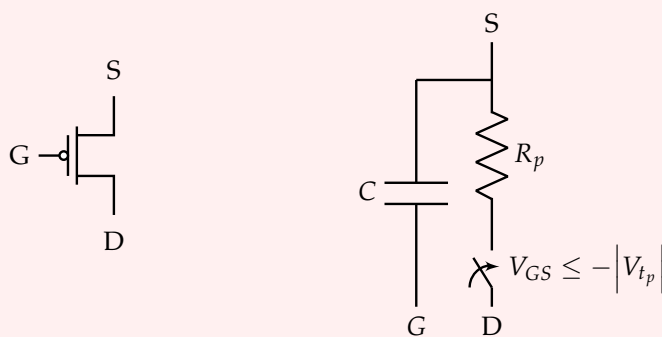
<sup>2</sup>When dealing with these circuits in real-world integrated circuits, we also must deal with the capacitance of the wires.



**Figure 8:** NMOS Transistor Resistor-Capacitor model

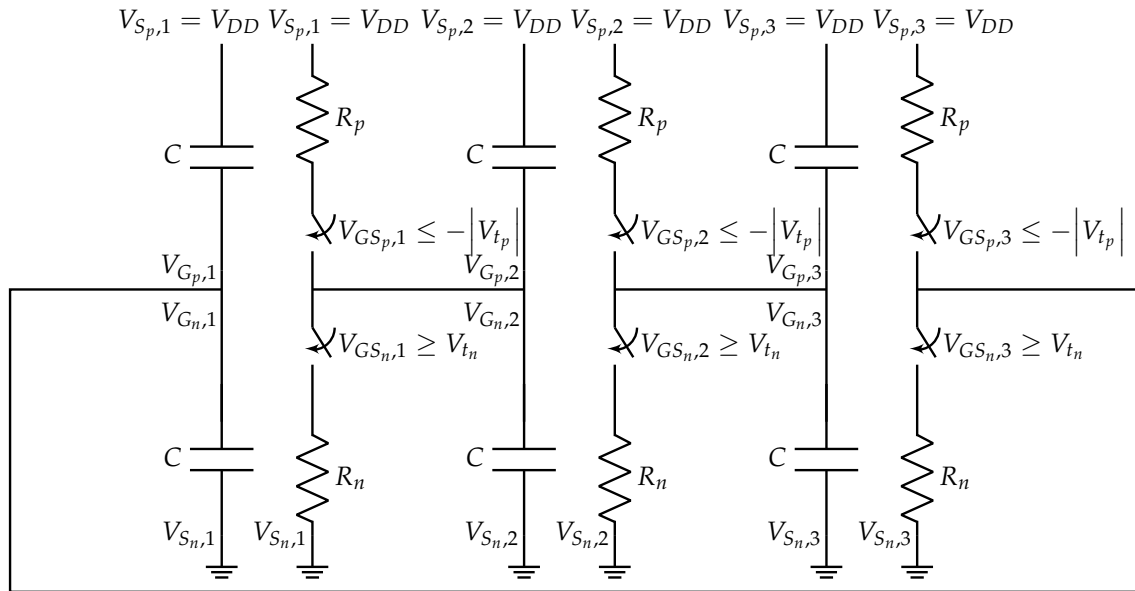
**Definition 8** (Resistor-Capacitor Model of a PMOS Transistor)

The diagram in fig. 9 describes the resistor-switch model of a PMOS transistor. The switch is on when  $V_{GS}$  is less than the negative threshold voltage. To avoid sign errors, we will define the inequality in terms of  $|V_{tp}|$ . Notice the convention: in PMOS Transistors, the source terminal is at the top.



**Figure 9:** PMOS Transistor Resistor-Capacitor model

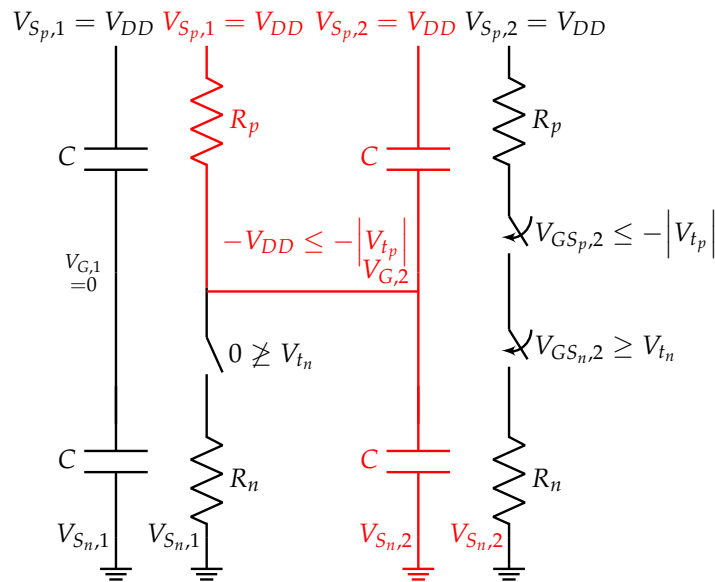
This model for inverters can be used to redraw and analyze our oscillator made out of inverters from fig. 6



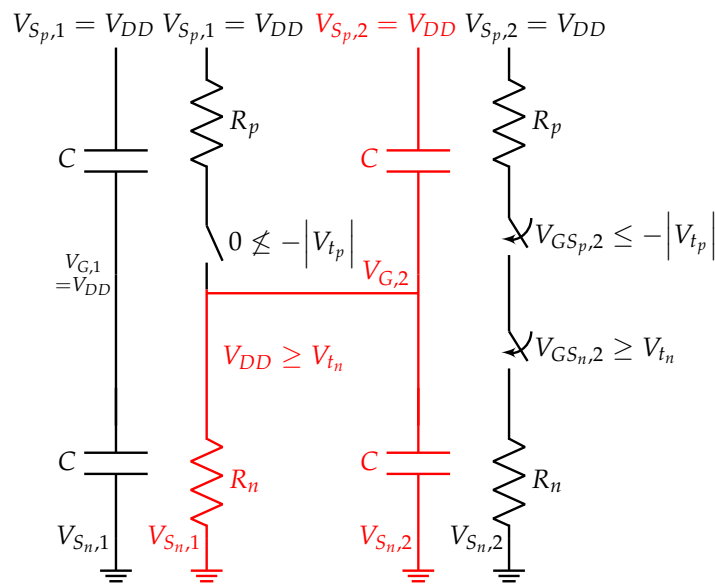
**Figure 10:** Ring oscillator with detailed transistor model:  $V_{G_{S_{n,i}}} = V_{G_{n,i}} - V_{S_{n,i}}$  and  $V_{G_{S_{p,i}}} = V_{G_{p,i}} - V_{S_{p,i}}$ .

With this model, we can see that each inverter drives some capacitance. This means that each inverter is pushing or draining charge from capacitors to cause the output to flip to a 1 (high) or a 0 (low). To get an idea of how fast it takes for the inverter to change signals, let us examine the case of an inverter in the oscillator where the output started at 1, and is switching to 0. To be concrete, we'll center our analysis on the output of inverter 1 (the input of inverter two), which is  $V_{G_{p,2}} = V_{G_{n,2}}$  in fig. 10.

In this case, a gate voltage of 0 at the input of inverter 1 means the NMOS transistors are off and the PMOS transistors are on, giving fig. 11. The capacitors depicted are the gate capacitances of inverter 2, and the resistances are the transistor switch resistances of inverter 1. The input of inverter 1 ( $V_{G_{,1}}$ ) sets the output of inverter 1 via the inverter 1 switches, and this output is the input of inverter 2.



**Figure 11:** Inverter output when  $V_{G,1} = 0$  and therefore,  $V_{G,2} = V_{DD}$ .



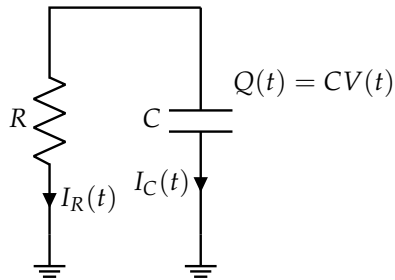
**Figure 12:** Inverter output when  $V_{G,1} = V_{DD}$  and therefore,  $V_{G,2} = 0$ .

Let's use an example to clarify what's happening in the figures above, by considering inverters 1 and 2 (see fig. 11 and fig. 12). Suppose that the input to the inverter chain,  $V_{G,1}$ , starts off at 0 V. Then, if this input is held for enough time, the inverter 1 output will become 1 V (and, since the inverter 1 output voltage is the same as the voltage on the gate capacitance of inverter 2, we can say that  $V_{C_{2,n}}$  for the NMOS capacitor is 1 V, or  $V_{DD}$ .)

We can see the behavior of the circuit and what components we need to consider, as they are highlighted in red in fig. 11 and fig. 12. By plugging in  $V_{G,1} = 0$  and  $V_{G,1} = V_{DD}$  respectively, we can see which components will form a complete circuit. Using this we can analyze the voltage (as a function of time) at

the input of inverter 2, i.e.  $V_{G,2} = V_{G_{p,2}} = V_{G_{n,2}}$ . We can now perform some circuit analysis to determine  $V_{G,2}$  in fig. 12 above.

## 1.5 RC Model Example



**Figure 13:** Capacitor discharging through circuit

Consider the circuit above. We can apply the capacitor current equation, Kirchhoff's Current Law (KCL), and Node Voltage Analysis (NVA) to analyze this circuit. By KCL,

$$I_C(t) = -I_R(t) \quad (1)$$

Using Ohm's Law, we know

$$I_R(t)R = V_R(t) \quad (2)$$

From NVA, we know that  $V_R(t) = V_C(t)$  since the resistor and capacitor share the same node. Combining the above equations, we see

$$I_C(t) = -\frac{1}{R}V_C(t) \quad (3)$$

Now, we can incorporate the capacitor current equation to achieve a differential equation:

$$C \frac{dV(t)}{dt} = -\frac{1}{R}V_C(t) \quad (4)$$

$$\frac{dV(t)}{dt} = -\frac{1}{RC}V_C(t) \quad (5)$$

Now, we can solve for  $V_{G,2}(t)$  in the circuit from fig. 12. The red portion of that circuit is redrawn in fig. 14.



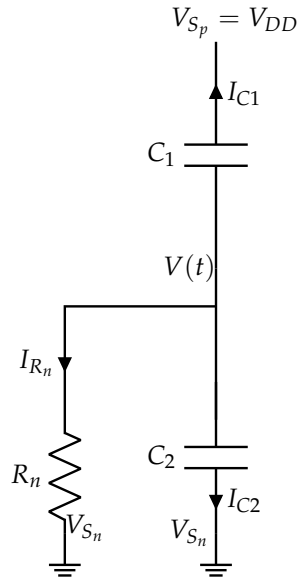


Figure 14: Inverter output at 0

In fig. 14, the inverter has just switched from outputting 1 to outputting 0. This means that the voltage  $V(t)$  started at  $V_{DD}$  and decreases to 0 at steady state. We know the voltage across  $C_1$  is  $V(t) - V_{DD}$  and the voltage across  $C_2$  is  $V(t)$ . Using this information we can set up a differential equation to solve for  $V(t)$  (which is analogous to  $V_{G,2}(t)$  in fig. 12):

$$I_{C1} = C_1 \frac{d}{dt}(V(t) - V_{DD}) \quad (6)$$

$$I_{C2} = C_2 \frac{d}{dt}V(t) \quad (7)$$

$$I_{R_n} = \frac{V(t)}{R_n} \quad (8)$$

$$I_{C1} + I_{C2} + I_{R_n} = 0 \quad (9)$$

$$C_1 \frac{d}{dt}(V(t) - V_{DD}) + C_2 \frac{d}{dt}V(t) + \frac{V(t)}{R_n} = 0 \quad (10)$$

$$C_1 \frac{d}{dt}(V(t) - V_{DD}) + C_2 \frac{d}{dt}V(t) = -\frac{V(t)}{R_n} \quad (11)$$

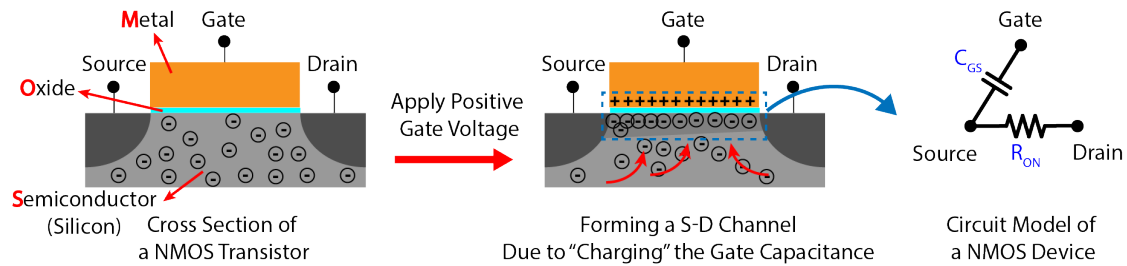
$$C_1 \frac{d}{dt}V(t) + C_2 \frac{d}{dt}V(t) = -\frac{V(t)}{R_n} \quad (12)$$

$$(C_1 + C_2) \frac{d}{dt}V(t) = -\frac{V(t)}{R_n} \quad (13)$$

$$\frac{d}{dt}V(t) = -\frac{V(t)}{R_n(C_1 + C_2)} \quad (14)$$

This is exactly the same form of differential equation that we got for the discharging capacitor circuit, just with a different value for capacitance! Thus, we have shown that we can boil this inverter circuit down to a capacitor discharging through a resistor. (You can take a similar approach to show that an inverter that switches from 0 to 1 is akin to *charging* a capacitor through a resistor).

## 1.6 OPTIONAL: Charge Puddle Model of a Transistor



**Figure 15:** A toy view of a physical transistor to illustrate the "puddle" model of transistor operation. The "puddle" is the group of negative charges that have accumulated on the interface of the semiconductor and the oxide. [https://en.wikipedia.org/wiki/Threshold\\_voltage](https://en.wikipedia.org/wiki/Threshold_voltage) has an interesting animation that literally shows the "puddle" growing as the gate voltage changes.

The underlying physical mechanisms governing the behavior of a transistor are out of scope for this course since it requires physics. However, there is a "charge puddle model" that gives a heuristic sense for why transistors work this way.<sup>3</sup> Look at fig. 15. The idea of this model is that the gate of the transistor is like one terminal of a capacitor with the other part being in the silicon between the source and drain terminals. When the voltage on the capacitor is high enough in the right direction, a "puddle" of charge carriers forms in the silicon to balance out the charge being put on the gate. When the puddle is large enough (hence the finite threshold), it connects the source and the drain, allowing current to flow between them. The "source" is the terminal that can be viewed as where the relevant charge carriers spill from to form the puddle. For NMOS, these carriers are electrons having a negative charge. For PMOS, these carriers are called "holes" and they have a positive charge. Actually understanding this properly requires more physics but it might help some of you remember the difference between PMOS and NMOS.

## 2 Check your Understanding

Read through these simple questions to check your basic understanding of the notes.

- Where transistors are used? Why they are so important?
- What are the names of the transistor's terminal?
- What is the difference between NMOS and PMOS?
- What is a CMOS inverter?
- Describe how ring oscillator works.
- What contradiction do we encounter if we try to describe the ring oscillator with the idealized switch model?
- What is the RC model for a transistor?

<sup>3</sup>When we say "heuristic", we generally mean "intuitive", and the word is used to allude to the fact that there's some preciseness missing that makes it not fully accurate or fully rigorous.

**Contributors:**

- Anish Muthali.
- Neelesh Ramachandran.
- Nikhil Shinde.
- Anant Sahai.
- Aditya Arun.
- Nikhil Jain.
- Chancharik Mitra.
- Matteo Guarrera.