Lecture 18: Digital circuits, sizing, output impedance, rise and fall time

Prof. J. S. Smith

Context

In the lecture, we started discussing how digital gates are built using NMOS and PMOS transistors.

In this lecture, we will continue to discuss the analog properties of digital gates

- Input capacitance
- Output impedance
- Rise and fall times
- Power consumption

Reading

- The midterm exam will be limited to the linear circuit material, and the material from chapters 1-4 in the text.
- Next, we will start with PN diodes again but in forward bias, and develop small signal models: Chapter 6
- We will then take a week on bipolar junction transistor (BJT): Chapter 7
- Then go on to design of transistor amplifiers: chapter 8

Announcements

- The midterm will be held on March 10, in Sibley Auditorium, from 6-8 pm
- The midterm is closed book, one page (one side) of notes.
- There will be a review session by Prof. Smith, Friday March 5, 6-8 pm, in 277 Cory.
- There will be a review session by the TA’s, Sunday March 7, Sunday 6-9pm 277 Cory.
- There will be NO LABS next week
In the last lecture, we discussed how digital gates are improved if both the pull up and pull down are accomplished with switched devices:

- No quiescent current
- High is pulled up all the way to Vdd
- Low is pulled down all the way to ground

If a NMOS and PMOS transistor are hooked up like this, if Vin is near 0 volts the NMOS transistor will be cut off, and the PMOS transistor will be well above threshold and conducting.
OR

- We can also make AND gates and OR gates with switches for both pull up and pull down

```
+1 volt
```

```
Input A
```

```
Input B
```

```
0 volts (ground)
```

NOT Input A

NOT Input B

Output

---

Static CMOS logic gates

- In general, a static CMOS logic gate is made using two networks of FETs: a pull up network, and a pull down network.

---

Pull up/pull down networks

- All of the pull up gates must be made from PMOS because a conductive path through an NMOS transistor could not pull up all the way to Vdd, it would have to remain below Vdd by at least the threshold voltage.

- All of the gates in the pull down network must be NMOS transistors for the same reason, a PMOS transistor would not be able to provide a conductive path for a node all the way to the ground voltage, the gate must be below the source by the threshold amount for the channel to conduct.
The pull up and pull down networks are not limited to simple gates, they can be Boolean combinations, but in general they must be inverting, for example NAND or NOR, if you want to do an AND gate, for instance, you will need to use an inverter, you can't do it in one stage.

The pull up network must be pulling up IF AND ONLY IF the down network is not pulling down!

PMOS is weaker

In this four input NAND gate, you can see that the PMOS transistors are in parallel, but the NMOS transistors are in series. You want to avoid having PMOS transistors in series, because their on resistance is higher due to the lower mobility and saturation velocity of holes. This leads you to use NAND gates rather than NOR gates.
In steady state, the drain current from the NMOS device must go through the PMOS device, and the sum of the voltage drops must equal (Vdd-Vss) so each forms a load line for the other.

For example, with an input voltage of 2 volts, the NMOS transistor is highly inverted, and has very little resistance to ground, but the PMOS transistor is only just above cut off, and is a high resistance path.

At an input of 1.5 volts, the NMOS transistor is conducting well, and the PMOS transistor is still pretty resistive.

At an input of 1 volt, the PMOS transistor is conducting well, and the NMOS transistor is nearing cut off.
CMOS Inverter Load Characteristics

- The plot of the output voltage vs the input voltage:

Balance: pull up vs pull down

- Having an imbalance in the pull up strength and the pull down strength has two consequences:
  - The transition point may not be at half the voltage, which will reduce the margin for one of the states.
  - The rise time may be slower than the fall time, or the fall time may be slower than the rise.

If we were to take our $V_{gs}=1.5$ volt curves, and double the width of the PMOS device, the output would be nearly $V_{dd}/2$ rather than near zero.
Transition voltage

- We can calculate the dividing point between a logic high and a logic low by solving for where the load lines intersect for an output voltage equal to the input voltage. At that point, a series of inverters would put out the same voltage. For higher or lower input voltage, each stage would produce an output closer to the Vdd or Vss.

Logic levels

- Ignoring the effects of channel length modulation, and solving for the voltage $V_M$, which divides the high state from the low state:

$$V_M = \frac{V_{in} + \frac{k_p}{k_n} (V_{DD} + V_T)}{1 + \frac{k_p}{k_n}}$$

This says that the voltage will be between the threshold voltages for the N and P devices, but biased toward the device with larger transconductance $k$.

CMOS logic levels (short channel)

- The drain current for the NMOS device:

$$I_{Dn} = v_{sat,n} W_n C_{ox} (V_{in} - V_{sat,n}) (1 + \lambda n V_{in})$$

- And for the PMOS device:

$$-I_{Dp} = v_{sat,p} W_p C_{ox} [V_{DD} - V_{sat,p} + V_T (1 + \lambda p (V_{DD} - V_{sat,p}))]$$

- And if we define the transconductances:

$$k_n = v_{sat,n} W_n C_{ox} \quad k_p = v_{sat,p} W_p C_{ox}$$

Device width scaling

- To maintain the highest noise margin, we can scale the devices to make their values of $k$ equal.

- For an inverter, that means making the P device a bit wider.

- For other devices, the pull up/pull down will be changed by the number of devices in series or parallel for a particular set of inputs.
Input dependence of rise/fall time

- For example, in the 4 input NAND, the number of parallel P channels depends on the number of inputs which were low (1-4).

So if only one input transitioned Low, it would have considerably less driving capacity than if they all were driven low.

The only time you have to pull the output low is if all inputs are high. Then the pull down is the same as a transistor 4 times as long, at this width.

Channel Lengths

- Notice that for logic devices, the channels are usually made as short as the technology allows. Longer channels would just increase the capacitance and decrease the driving capacity of the gates.

Transient Response

Whenever there is a change at the inputs, the change in the output voltage will require driving current through the switches that are still on, charging the load capacitance.

The time that it takes to charge the capacitance is just the RC time constant of the capacitance together with the ON resistance of the switches. For fast logic, we want low capacitance and low resistance. Since the current through the gate scales with W, we could make that large, but that will also make C large (from the next stage).

Dynamic dissipation

- In a CMOS logic gate, the current from the power supply goes almost entirely to charging the capacitance of the next stage.
- The power it takes is: \( P = V I \)
- The current is supplied at a voltage \( V_{DD} \).
- To find the energy per transition, we integrate over the switching time.

\[
\text{Energy} = \int P(t) dt = V_{DD} \int I(t) dt = V_{DD} Q = CV_{DD}^2
\]

So it helps both to reduce operating voltages, and capacitances.
Level shifting

- In many new VLSI devices, internal voltages are kept low for lower dissipation, but they still need to work with other chips at higher voltages.
- Therefore, level shifters need to be provided at the inputs and outputs.
- High level outputs are essentially amplifiers, which we will be studying soon. Multiple stages may be needed for high currents.
- High level inputs: voltages above the body voltage of the PMOS transistors must be avoided, otherwise the source to body junction will be forward biased, with high currents.

Slow transitions

If the input to a gate is noisy at the point of transition, it may result in a burst of noise pulses due to the high gain of the logic gates at that point.