Transistor review: Note that this is for a bunch of different classes so it may seem simplistic in some places and cover things you haven't seen yet in other places. I'll note which class information is coming from so you can pick and choose a little but you might want to try to read it all to give yourself a preview.

MODELS:

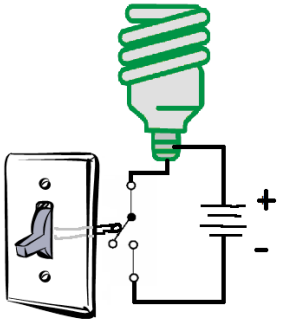
Modeling the behavior of a thing allows you to draw a simple picture on paper but have it represent some (possibly complicated) behavior. You can think of the drawings as black boxes in that they tell you what is happening but don't tell you the details of how it happens. In electronics, when you have something that impedes the flow of current, you call it resistance and when you draw it in a schematic (a picture of a circuit on paper or in software) you draw it like:  . That tells the person looking at it that if there are more electrons on one end of the resistor than the other (that's what voltage is) then some electrons will try to flow from the side that has more electrons to the side that has less but there is something that keeps them from flowing freely – there's something that causes some "impedance" or "resistance". Impedance is a little different than resistance in that it includes inductive and capacitive "resistance" to current changes and is often drawn as a box: 

instead of using the resistance symbol  .

Transistors have many models. When drawing a schematic, you always want to choose the simplest model you can so as to make your calculations the simplest they can be. Since we'll talk about transistor models below, I won't go into details here but, as an example, when you build a transistor in silicon, there is some natural capacitance that just exists. There's no way to get rid of it and no-one has figured out how to build a transistor without some naturally occurring capacitance. The thing is that you don't need to worry about capacitance if your circuit is being run at a very low frequency or is at DC (Remember: Capacitors are open circuits at low frequencies). You DO need to worry about capacitance when you run a circuit at high frequencies. As mentioned, there are quite a few models for a transistor and two of them are: 1) a transistor model without capacitance and 2) a transistor model with capacitance. So, if you are doing circuit analysis on a transistor at DC, you could use the model with the capacitance drawn in but then, when you get to the end, your equations will be messy and you will have to replace all the $(1/j\omega C)$ terms with infinity anyways. If you do that, the capacitance terms will disappear and you will get the result that you would have gotten using the model that doesn't include capacitance. Using a model that is more complicated than needed will put you through more pain during calculations than needed.

In the next few sections I'll go through three different models for transistors and describe where they are used.

TRANSISTORS AS SWITCHES:



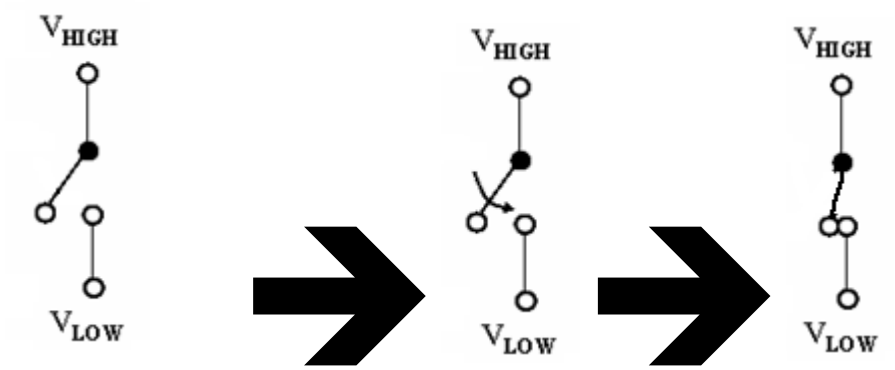
First let's look at a switch. Think of a light switch. A light switch connects or disconnects two wires:

The switch forces a connection between two wires, current flows and the light turns on.

It turns out that **the simplest model** of a transistor is exactly the same as the switch described above and can be simply thought of as a switch. A transistor has three (in this simple model) terminals. Two of them are connected to wires (just like the

light switch example above) and the third terminal opens and closes the switch depending on a voltage applied to it. Opening the switch stops current (open circuit) and a closing the switch lets current flow (short circuit). (Note that I'm ignoring a bunch of details such as the fact that for one well-known family of transistors that many of you may know, you can start and stop current with current into that third terminal, not voltage and some models have four terminals. But we'll keep it simple in this handout and just talk about the three-terminal, voltage in, current out models).

Let's redraw the symbols for the switch model of a transistor:



Switch open – no current Switch closed, current flows

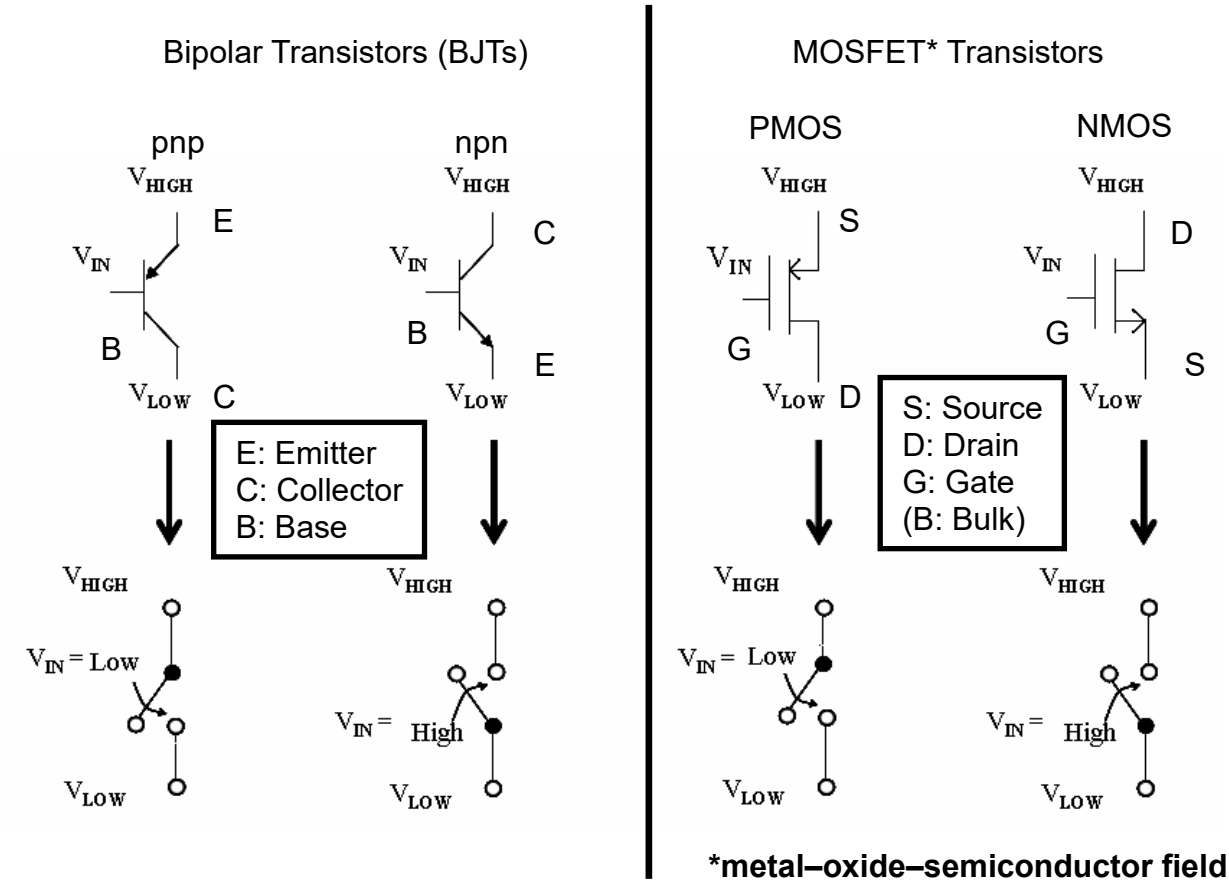
As circuit theory tells you, current will flow from the V_{HIGH} to the V_{LOW}.

We'll be dealing with two types of transistors (switches): 1) Ones that conducts when a high voltage (like VDD) is applied to the third terminal and 2) ones that conducts when a low voltage (like GND) is applied to the third terminal. (The third terminal is the switch input. Not the terminals that are connected to wires that current flow through). The idea of having two different types of transistors, one that conducts on a low input and one that conducts on a high input makes some very interesting things possible. (SEE: **LOGIC CIRCUITS:** on about page 19 or so). You can think of the two models as one where the light turns on when the switch is up and one model where the light turns on when the switch is down.

SYMBOLS FOR THE MORE COMPLICATED TRANSISTOR MODEL:

The switch model can be used for analysis of very low frequency digital circuits or transistors used as switches on a PCB (printed circuit board) that just turn something on or turn something off. It doesn't take into account what happens between the time that it's conducting fully (ON) and when it isn't conducting (OFF). As with all real world events, nothing goes from one state to another instantaneously and that is true with transistors too: They don't go from no conduction to full conduction instantaneously but, when modelling switches, very often you don't care what happens in-between so the switch model is just fine to use.

For example, when calculating how fast a digital circuit can run, unlike switches, you need to know what happens between the OFF and ON states of the transistor. To make it easier to discuss, let's introduce the two types of transistors and their symbols:

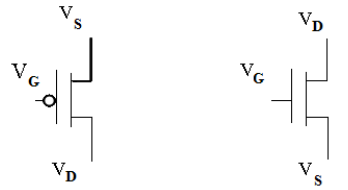


***metal–oxide–semiconductor field-effect transistor**

The two types of transistors we'll see in class: 1) Bipolars (BJTs) and 2) MOSFETs. The difference between these two is how they are constructed in silicon and how the current through them changes when the third terminal (the "switch") voltage is changed (exponential or quadratic) but, in general, they are very much alike. In each type there are two sub-flavors: 1) Ones that conducts when a high voltage (like VDD) is applied to the third terminal and 2) ones that conducts when a low voltage (like GND) is applied to the third terminal. The ones that conduct when a high voltage is applied to the third terminal have names that start with an "N". "npn" for bipolars and "NMOS" for MOSFETs. The ones that conduct when a low voltage is applied to the third terminal have names that start with a "P". "pnp" for bipolars and "PMOS" for MOSFETs.

We'll stick with three terminal model drawings for now though MOSFETS are sometimes drawn with a fourth terminal to the substrate (what's listed as BULK in the above diagram – But we won't have time to talk about that terminal in this handout). Let's give the terminals names so we can talk about voltages and currents on the devices a little easier. In the diagram above, the names of the terminals are spelled out in the black framed box. For BJTs, there is an emitter, a collector and base. The base is the switch and the emitter and collector are connected to the wires that conduct. For MOSFETs, there is a

source, a drain and gate. The gate is the switch and the source and drain are connected to the wires that conduct. The symbols that I have used in the diagram above are the analog symbols for BJTs and MOSFETs. The digital symbols for the MOSFET look like:



PMOS and NMOS

In the analog symbols, for both BJTs and MOSFETs, you always have arrows pointing in the direction of current at the source (MOSFETs) or the emitter (BJTs). In digital symbols there are no arrows but there is a "bubble" on the gate of the PMOS. The symbols for BJTs in digital circuits are the same as the analog symbols. A bubble represents an inverter which is a digital component. Since a bubble (inverter) is a digital part, it isn't used in analog symbols.

A slight aside here to make the rest of the explanations easier: It turns out that if you understand how BJTs work at the symbol level, you understand 95% of how MOSFETs work and if you understand how MOSFETs work at the symbol level, you understand 95% of how BJTs work. Both BJTs and MOSFETS have two terminals that conduct and one terminal that turns the switch on or off (allows current to flow through the other two terminals or not). Since they behave very much alike and something I say for one will be true for the other, it'd be nice if I could refer to parts of them without having to use separate names for bjts and MOSFETs. The hard part in referring to a particular terminal of a transistor is that the terminals are named differently for BJTs and MOSFETs so from here on out I'm going to use S/E for the source (MOSFET) or emitter (BJT), D/C for the drain (MOSFET) or collector (BJT) and G/B for the gate (MOSFET) or base (BJT). When I use S/E, G/B or D/C, it means that whatever I say is true for both MOSFETs and BJTs.

	Switch terminal	Conducting terminal 1	Conducting terminal 2	P device	N device
MOSFET	Gate	Drain	Source	PMOS	NMOS
BJT	Base	Collector	Emitter	pnp	npn
Combined	G/B	D/C	S/E	P- Transistor	N-Transistor

I'll also use P-transistors to talk about transistors that turn on on a low G/B voltage and N-transistors to talk about transistors that turn on on a high G/B voltage whether they are MOSFETs or BJTs.

If you look at the drawings of the transistors (both MOSFET and BJT) you'll notice that the P-transistors have their S/E at the top (higher voltage) and the N-transistors have their S/E at the bottom (lower voltage). Note that the D/C terminal is also switched. Pretty much everything is opposite with the two flavors of transistors. One (N-transistors) turns on when the G/B voltage goes up and one (P-transistors) turns on when the gate voltage goes down. One has the D/C terminal on the bottom/low voltage (P-transistors) and one has the D/C terminal on the top/high voltage (N-transistors). etc…. Here are the rules:

	P-transistor	N-transistor
Top terminal (higher voltage)	S/E	D/C
Bottom terminal (Lower voltage)	D/C	S/E
Current into	S/E	D/C
Current out of	D/C	S/E
Turns on when	G/B voltage goes down	G/B voltage goes up
BJT only*: Base current	Out of base	Into base

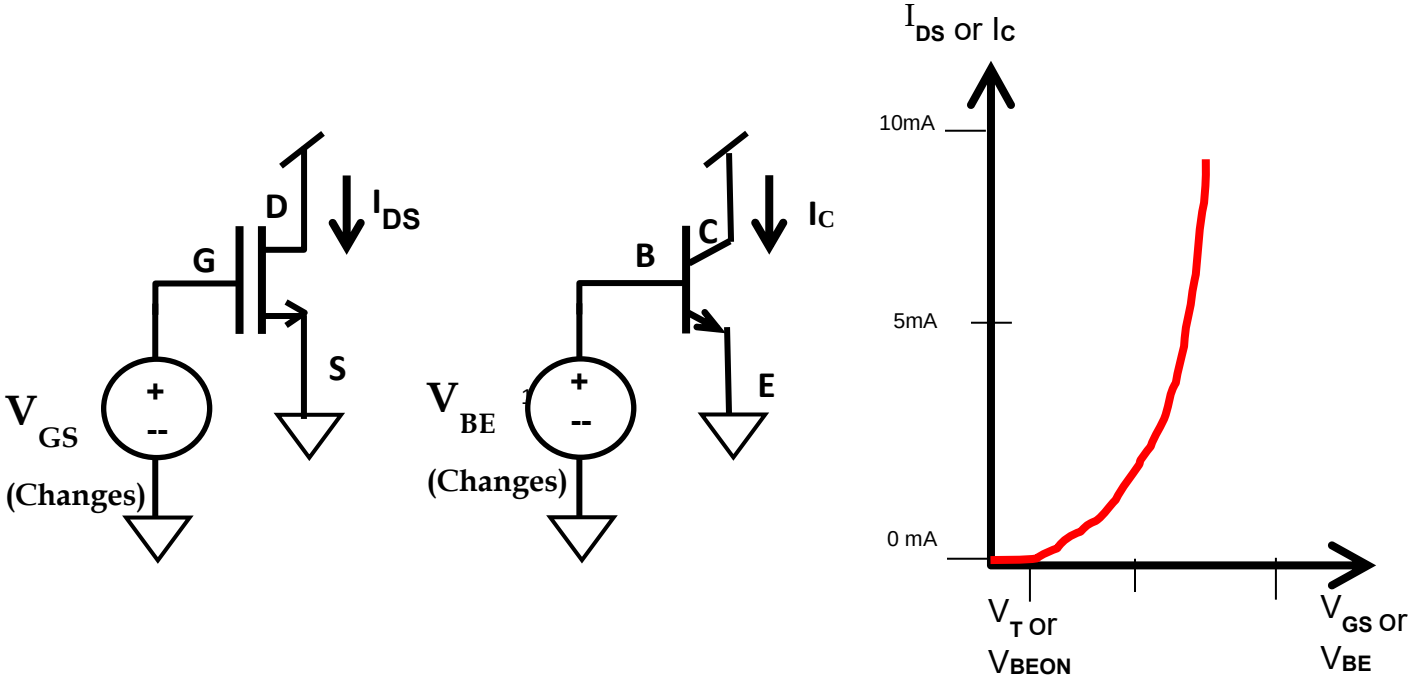
*No current goes into the gate of a MOSFET at low frequencies. At high frequencies, current runs on and off of the capacitance on the gate so it looks like current is flowing through the gate but it actually isn't.

It might be worth the time right now, especially for those that haven't had EE307 or EE308, to summarize the MOSFET and BJT devices once on your own. That should help you solidify the differences between MOSFETs and BJTs, P-transistors and N-transistors and get a review of where and what the names are of the terminals for all these devices.

MORE COMPLICATED TRANSISTOR MODEL – Turning the switch on and off

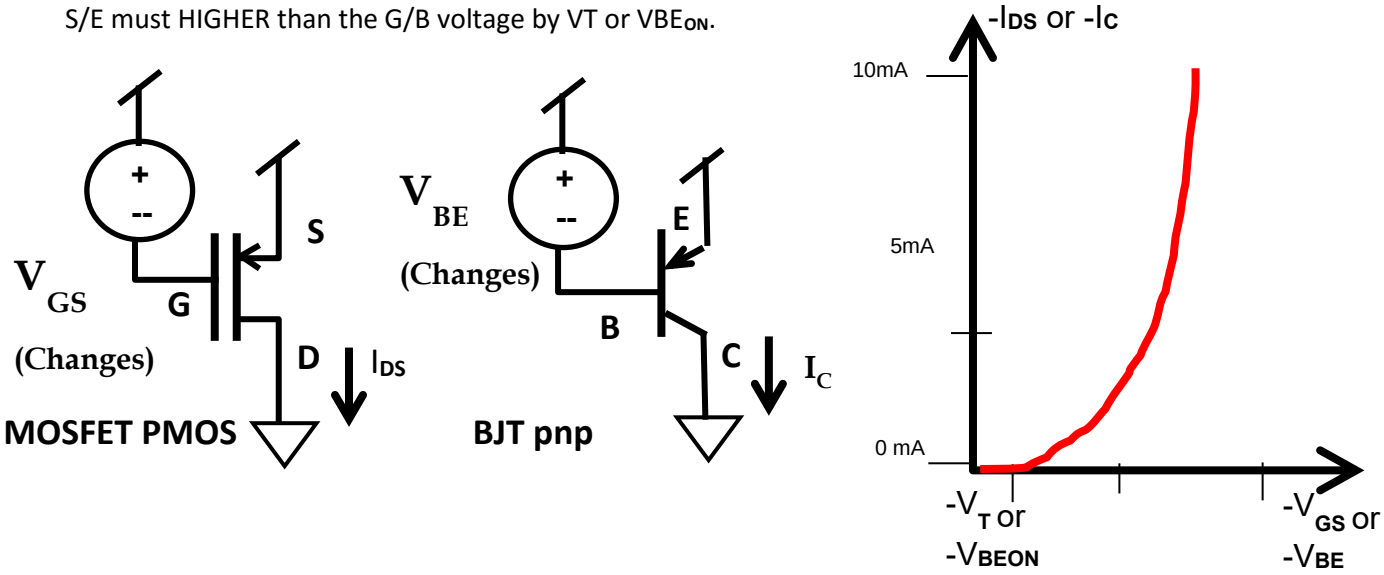
The last section just talked about the naming and symbols of MOSFETs and BJTs. This section will talk about basic turning on and off of transistors and the behavior between being off and fully on. For BJT and MOSFET transistors, behavior means what happens to current through the device. I had a professor once that said that there is really only one thing in electronics and that is electrons. Everything else was just added on to help describe what the electrons were doing. Current is a measure of how many electrons flow through a point per unit time (current is defined as going in the opposite direction of the electron flow). Voltage is just how much an electron wants to get to one place from someplace else – How much the electron wants to go from a higher concentration of electrons to a lower concentration of electrons. Electrons don't like to be in a crowd since they have the same polarity (-) and repel each other so they want to go to a place where there is a lower concentration of electrons. That's all voltage is – electrons wanting to get away from each other to create equilibrium. Resistance is just how hard it is for an electron to get to someplace from someplace. Inductance is an effect where electrons keep flowing as before after a voltage (electron concentration) is changed. Capacitance is where there's an area that has to be filled with electrons before the electrons can freely flow past. So, again, current (electron flow) is really the most important thing to be able to describe in circuits. And, again, the equations we will use to describe transistors will describe the current through the devices.

Let's look at the graphs first. There are two main graphs that you should understand and know for the EE classes at Cal Poly. The first we'll look at tells you about how the switch works. The graph below describes what happens to the current when the G/B voltage is changed relative to the S/E voltage. Remember that the G/B voltage is the input to the third terminal that either makes the transistor conduct or not conduct. It is basically the switch. So this graph tells you what happens when you slowly move the switch from OFF to ON:



There are two **<u>N-transistors</u>** on the left of the graph. One is a MOSFET device and one is a BJT device. N-transistors turn on when their G/B voltage goes up relative to the S/E voltage. MOSFETs turn on when the voltage across the gate and source (the voltage difference between the gate and source voltage) gets above a voltage called the threshold voltage (VT). A BJT turns on when the voltage difference between the base and the emitter is greater than the BJT threshold voltage, VBE_{ON}. Both VT and VBE_{ON} are 100% dependent on who fabricated the transistor. Almost every textbook you've ever seen probably lists both of these voltages as 0.7V. **THAT'S NOT WHAT THEY WILL BE IN THE REAL WORLD!** You'll see values for these voltages from 0.2V to many volts. The textbooks make you think that these will always be 0.7V and that's just not true!!!!! I think they always tell you 0.7V to make grading easy. If you've taken EE348 you'll know that 0.7V would just not work for your experiments. So put in your mind right now that even though your textbooks may use 0.7V, you'll need to check the datasheet (a description of the device put out by the manufacturer) to see what the company really thinks the voltages are and then you'll still have to verify it in lab to see how far off the datasheet is for the batch of transistors that you are using.

For N-transistors, the G/B voltage is higher than the S/E voltage. Once again P-transistors are opposite. They still turn on by increasing the difference between the G/B and S/E voltages but now the S/E must HIGHER than the G/B voltage by VT or VBE_{ON}.



I'm fudging a bit here. In reality, when you look up the VT or VBEON for a P-transistor, it will already be a negative number. I put the negative signs on the graph just to stress that all of these are negative voltages. In reality, this graph would be flipped over horizontally as well as vertically and in the third quadrant. Note that the current is going from the source to the drain for the PMOS (P-transistor MOSFET) so IDS is negative and, since IC is defined as INTO the collector of the pnp (P-transistor BJT), current coming out of the collector is also defined as negative.

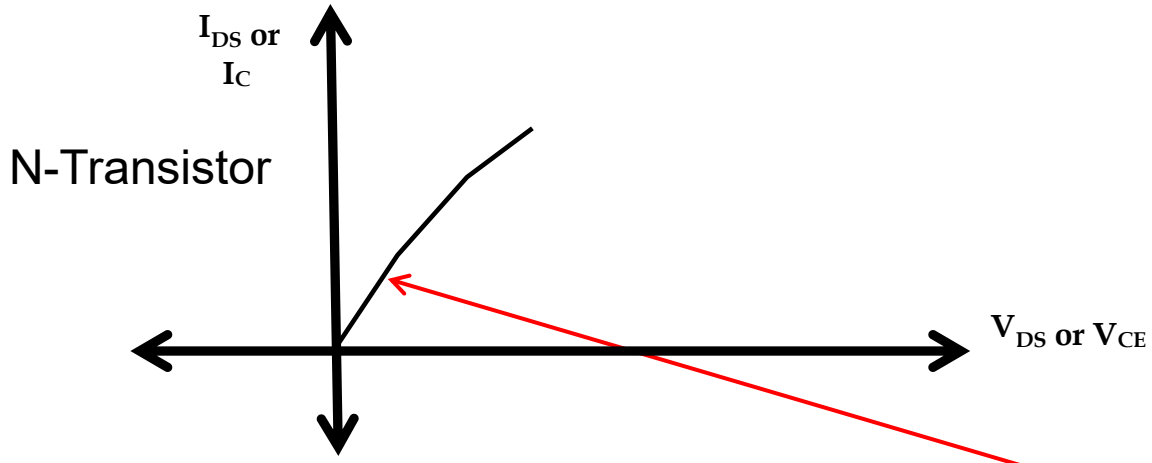
So, in summary:

- The gate (the terminal that acts like a switch) has to be more than VT above the source for an N-transistor MOSFET (NMOS),
- The gate (the terminal that acts like a switch) has to be more than VT below the source for a P-transistor MOSFET (PMOS),
- The base (the terminal that acts like a switch) has to be more than VBEON above the emitter for an N-transistor BJT to be on and
- The base (the terminal that acts like a switch) has to be more than VBEON below the emitter for a P-transistor BJT to be on.

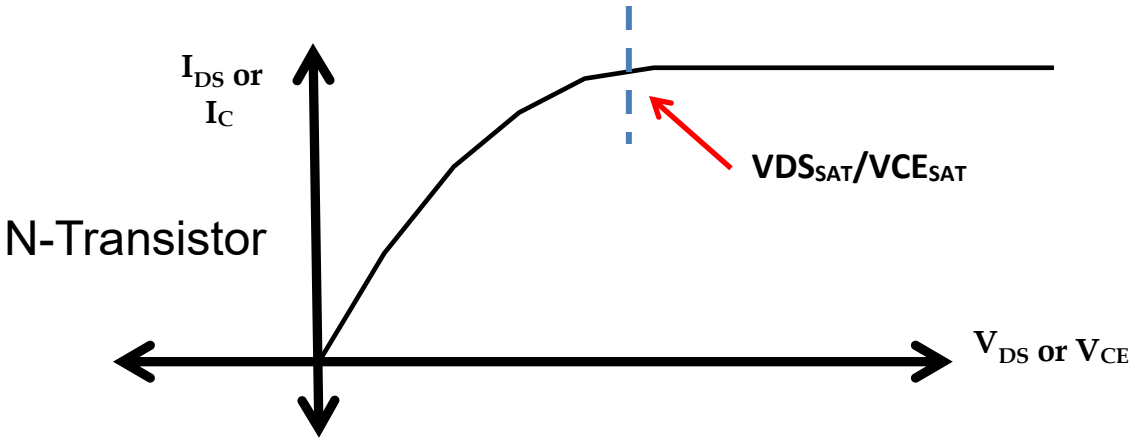
And, to a limit, the bigger that difference is, the more current flows. Also note that there will be different VT values for NMOS and PMOS and that the VT value for the NMOS will be positive and the VT value for PMOS will be negative.

MORE COMPLICATED TRANSISTOR MODEL – Effects of the other two terminals on current

The last section showed you the graph that shows what happens when you change the voltage on the terminal that acts like the light switch. You saw that that voltage must be above a particular level for the transistor to conduct. The other voltage that you have to worry about is across the two terminals that conduct current. Just like a resistor, if there isn't enough voltage across these other two terminals, then there's no reason for the electrons to travel across the transistor. Let's look at the current graph for small voltages across the conducting terminals. Everything I say is true here for MOSFETs and BJTs. Be careful when you look at P-transistors though because, again, everything is flipped and all the signs are switched.



When the voltage across the terminals (VDS for MOSFETs and VCE for BJTs) is small, hardly any current flows (just like a resistor). As the voltage across the terminals increases, the current increases pretty linearly. That is why this region on the graph is often called the linear or resistive region. A resistor would give you a graph that looks about the same. You might think that if you keep increasing the voltage across the conducting terminals, the current would continue to increase. Nope. At some point the voltage across the channel (MOSFET) or base region (BJT) undergoes some changes that stop the current from increasing forever. At that point, where VDS equals a voltage called VDSAT (for MOSFETs) or VCE equals a voltage called VCESAT (for BJTs), the current flattens out and stops increasing as it did in the linear region. This region doesn't look like it's doing much but it's actually the region where you want your transistors to be when you build almost any circuit. I won't talk more about why this region is preferred except to say that in order to draw the curve below, VGS/VBE was held constant and the curve was drawn just changing VDS/VCE. This is very hand-wavey and really only a small part of the story but notice that the highest current is in the region to the right of VDSsat/VCEsat. More current generally means quicker changes and happier digital circuits. The reason that analog circuits like this region is more complex – It turns out that for a small change in VGS/VBE, you get a bigger change in current than in the region to the left of VDSsat/VCEsat. So digital likes that region because it's the highest current for a fixed VGS/VBE and analog likes that region because there's the biggest change in current for a change in VGS/VBE. It is the region that is best for fast digital circuits and analog amplifiers.



TRANSISTOR REGIONS OF OPERATION

A region of operation for a transistor is a region that behaves pretty much the same at any point in that region. That probably won't make sense with just that explanation so look back at the graph at the bottom of the last page. To the left of VDSat / VCEsat, the transistor is behaving like a resistor. To the right of VDSat / VCEsat the transistor is behaving like a current source (close to constant current even when VDS or VCE changes). It would be very difficult to find one equation that neatly covers both of those regions so the smart people that came up with the idea of transistors said, "let's divide things into regions so we can more easily find equations to describe the behavior in each separate section of the graph".

Let's give the regions names before we go on. For EE308 and EE431, we only care about three regions of operations of the MOSFETs and BJTs so I will only name those. In EE306 and EE307 you may see some other regions of operation for BJTs. It's a little out of order but let's name the regions before we define what each region is. I numbered the three regions and noted the names that are used with BJTs and MOSFETs. Since I want to be able to discuss both MOSFETs and BJTs at the same time, I will use the names in the "**What I will call it**" column.

Region	BJT	MOSFET	**What I will call it**	Importance
1	Cutoff	Cutoff, subthreshold	**Cutoff**	No current
2	Saturation	Linear, ohmic, resistive	**Linear**	Not really used in EE308
3	Forward active	Saturation, pinchoff	**Active**	Region we like

Now that I have names for things let's see what those names mean.

Region	**What I will call it**	Requirements (MOSFETs)	Requirements (BJTs)
1	Cutoff	|VGS| < |VT|	|VBE| < |VBEON|
2	Linear	|VGS| > |VT|, |VDS| < |VGS-VT|	|VBE| > |VBEON|, |VCE| < |VCEsat|
3	Active	|VGS| > |VT|, |VDS| > |VGS-VT|	|VBE| > |VBEON|, |VCE| > |VCEsat|

VT, VCEsat and VBEON depend on how the devices are fabricated so you need to read the datasheet to find out their values.

A couple of notes on the last table.

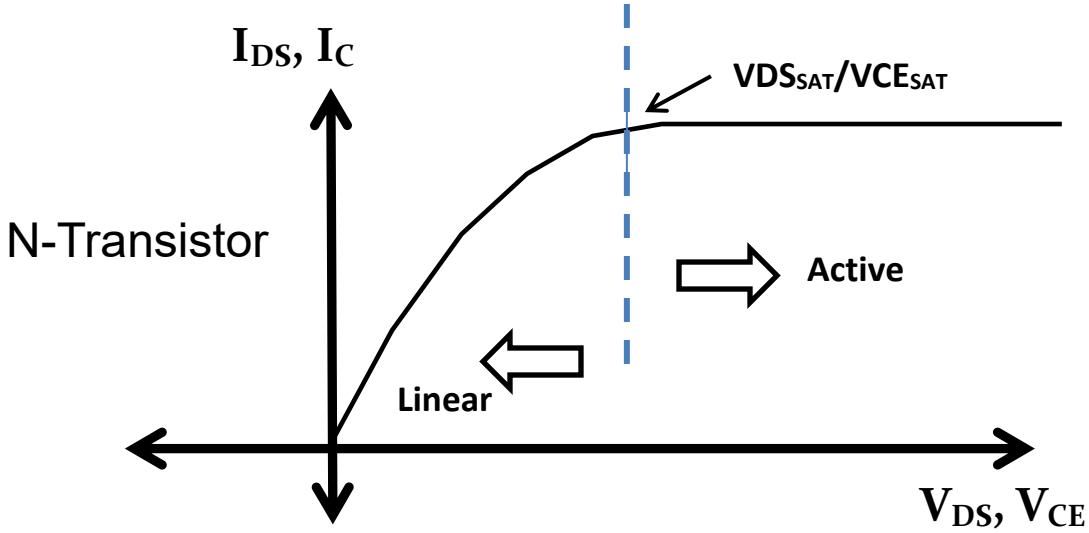
>> First, a comment: I used absolute values around the VGS, VT, VDS, VGS-VT, VBE, VBEON, VCE and VCEsat. For N-transistors, all of these voltages are positive. For P-transistors, all of these voltages are negative. For P-transistors, you can remove the absolute value bars but then you have to flip all the less-than and more-than inequality signs. Please pick a way that makes sense to you and go with it.

>> Second, a summary: There are three regions of operation of interest to us.

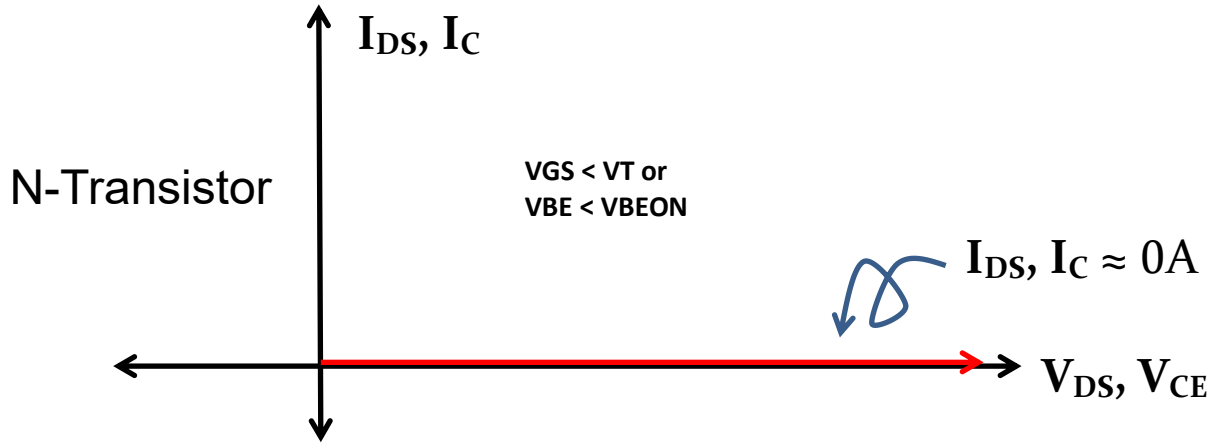
- CUTOFF means that very little current flows (textbooks call it zero but there actually still is current flowing). The device is in cutoff when |VGS|<|VT| (for MOSFETs) or |VBE|<|VBEON| (for BJTs). This is the only region that has the VGS / VBE junctions OFF. The two other regions require this junction to be on which means that |VGS|>|VT| (for MOSFETs) or |VBE|>|VBEON| (for BJTs).

- LINEAR means that the VGS or VBE junction is on (|VGS|>|VT| (for MOSFETs) or |VBE|>|VBEON| (for BJTs)) but there isn't much voltage across the other two terminals. "Much voltage" is defined as less than VDSsat for MOSFETs and less than VCEsat for BJTs. If the VGS / VBE junction is on but |VDS|<|VDSsat| (for MOSFETs) or |VCE|< |VCEsat| for BJTs then the device is in linear. (VDSsat = VGS-VT by definition)
- ACTIVE means that the VGS or VBE junction is on (|VGS|>|VT| (for MOSFETs) or |VBE|>|VBEON| (for BJTs)) and there is much voltage across the other two terminals. "Much voltage" is defined as more than VDSsat for MOSFETs and more than VCEsat for BJTs. If the VGS / VBE junction is on and |VDS|>|VDSsat| (for MOSFETs) or |VCE|>| VCEsat| for BJTs then the device is in active.

Let's look at where those are on the graphs. To the left of the VDsat/VCEsat line is the linear region. To the right of the VDsat/VCEsat line is the active region.



In the third region of operation, cutoff, there is no current so the graph looks like this:



TRANSISTOR REGIONS OF OPERATION AND THEIR CURRENT EQUATIONS

I've described where the regions are on graphs but, as always, there's more. To do calculations on paper, you need equations to describe how a device behaves (what the current is doing) under various operating conditions. By operating conditions I mean the voltages on all the terminals. Note that different voltages on the terminals make the current do different things. In the linear region, the current changes quite a bit with a change in VDS or VCE. In the active region it only changes slightly. When you are designing a circuit, you get to pick the voltages on the terminals (and therefore the region of operation of that device). Setting the terminal voltages is a ¤¤¤**VERY**¤¤¤ important part of design and is called "biasing the circuit" or "biasing the device". In your textbook it may have been called setting the Q-point or setting the operating point. I just tend to call it the DC biasing or just biasing point. I call the act of setting the terminal voltages "biasing". Note that I have just spent an entire paragraph on "biasing". That means it is pretty darn important.

But back to the point of this section: The equations for the different regions of operation. Let's start with the easiest one:

1. **CUTOFF:** IC ≈ 0A (for BJTs), IDS ≈ 0A (for MOSFETs),

Done. So, if |VGS|<|VT| (for MOSFETs) or |VBE|<|VBEON| (for BJTs), this is the current equation → The current equals zero. This is actually a bit of a lie but it's easier for textbooks to ignore that there still is current flowing. The current is just so small compared to when the device is in linear or active, it's usually ignored. I know of circuits designed to run in cutoff so that they use very little power. So, in most cases, it's OK to approximate the current of a device in cutoff as zero, but just know that that's not quite right.

2. **LINEAR:** Not going to bother with the equations.

I know I told you that this is one of the regions of operation I consider important but mostly it's important because you need to be able to identify when a device is in it and then make that device get **OUT** of it (at least in EE308). Since it's really not used very often in circuits, when you see a transistor in linear, you'll need to figure out how to get the device into (probably) active so your circuit will work. The other thing about the equations for transistors that are operating in linear is that they will give you wrong answers. If you go back and look at the graphs that have VDS/VCE on the X-axis, note that a small change in VDS or VCE will make the current of a device operating in the linear region change quite a bit. That means that you need a very exact VDS or VCE value to get anywhere near the right current. Let's say you **DO** have a very exact VDS or VCE to plug into the equations, that is good if you are working on textbook problems or in a simulator but, in the real world, each transistor, no matter how hard you try to match them up, will be slightly different. In a simulator or on paper, all devices are EXACTLY PERFECTLY matched so the math may work out. But in the real world, the current equations for devices in linear will probably give you answers that are pretty far off.

3. **ACTIVE:** $IDS = \frac{k'}{2}\frac{W}{L}(VGS - VTP)^2(1 + \lambda \cdot VDS)$ (for MOSFETs), and

$$I_C = I_S e^{\left(\frac{v_{BE}}{VT}\right)}\left(1 + \frac{VCE}{VA}\right)$$ (for BJTs)

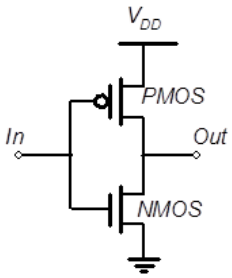
For conducting transistors, this is the region of operation you probably want your transistor to be in.

Once again, if you haven't taken EE307, this might be a good place to again summarize the symbols, graphs and current equations. Try to put it all together.

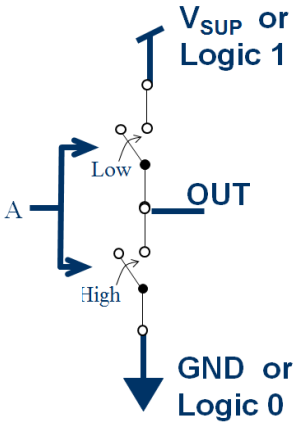
DIGITAL INVERTER AND ITS TRANSFER FUNCTION

We've seen what individual transistor do when you change terminal voltages. Now let's see what you can do with them in a circuit.

Here's the circuit we'll start with:



This is an inverter which is a digital circuit. If you just want to know the simplest operation of this circuit, using a switch model is enough:



Input A is connected to both the top and bottom switches. The top switch closes (conducts) when the voltage to A is low (GND for example) and, for the same A=low input, the bottom switch would be open (not conducting). The other important case is when A is high (Vsup for example) and the bottom switch is conducting (and therefore connecting OUT to GND) and the top switch is OFF. If we replaced the switches with transistors (A PMOS or pnp on top and an NMOS or npn on the bottom) this circuit would behave just like the switch model.

Let's draw the transfer function for the circuit. A transfer function has the input on the X-axis and the output on the Y-axis.

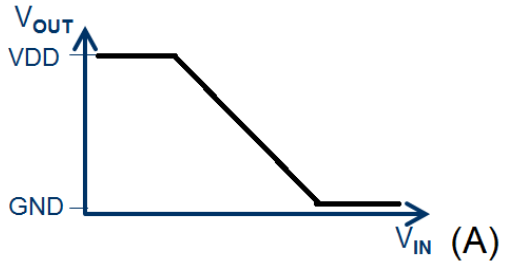
Here's the graph before graphing anything:

When figuring out what a circuit does, it's always good to start with the extremes. Setting the input, A, to the lowest voltage in the system, GND, let's see what "OUT" becomes. If A is zero then the switch that closes on low will close and conduct and the switch that closes on high will open and not conduct. Let's see what that does. Since the top switch conducts on A=0V, it connects the output OUT to the supply voltage Vsup. You'll see many textbooks say: "It creates a low impedance path to the supply voltage". That's just saying that the output is basically connected to Vsup through a wire. If you connect something to Vsup through a wire (with no other influences) that something becomes Vsup. In this case you are connecting Vout to Vsup through a wire so OUT goes to Vsup. But wait – You must also check the other switch because if the other switch is connecting OUT to GND, you have two voltages connected to OUT and that's a problem. But, since the lower switch is open when the input A is low, there is an open circuit (=a cut wire) from OUT to GND. So OUT is only connected to Vsup and will become the voltage Vsup. Let's add that to the transfer function graph:

Now let's look at the other extreme. When the input A is set to the maximum voltage in the system, Vsup, the switch that closes on high will close and conduct and the switch that closes on low will open and not conduct. That means that there is a low impedance path from GND to OUT and the wire from Vsup to OUT is cut (open) and OUT goes to GND. Let's add that to the transfer function graph.

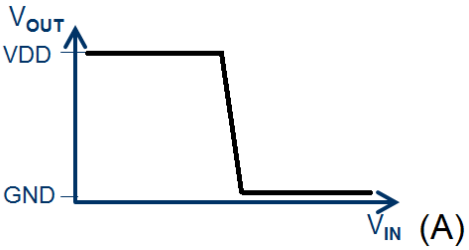
Unfortunately, if you use switches the middle region is a mystery. Switches only tell you on and off but the region in the center can have a time that both are on, both are off, one's conducting a little

and one's conducting a lot…. You can't figure out exactly what's going on there unless you use the more complicated model. We'll fudge now since everyone will, I hope see the details of that center region in either EE307, EE308 or EE431. For now we'll do like many textbooks do and just connect the lines and call it a day:

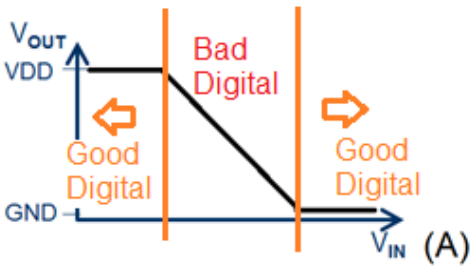


You're probably realizing that just connecting the lines with a sloping line is pretty random. Sometimes the math is just too complicated to draw that part of the graph accurately when doing hand calculations. Anyone know what you do if you need an accurate graph? YOU LET THE SIMULATOR DO IT. The math you learn in these circuit classes are true approximations and are generally used to approximate circuit values when doing preliminary circuit design work. Once you have approximate values, then you go to a simulator and you let it do the hard math. The math isn't totally impossible: If you just want to know the approximate slope of the sloped line at a particular single point, you get the math/tools to calculate it in EE308. Here we won't worry about those calculations – here let's just talk about the graph a little more.

In this section we've been working with an inverter – A digital circuit. If you are designing a digital circuit, you want a circuit that always gives you an output of either a high or a low voltage. Voltages in-between are confusing as to whether they are supposed to be a logic 1 or a logic 0. Looking at the graph at the top of this page, you see that the output is clearly high to the left and clearly low to the right but you also see a pretty big region that is neither. This graph is of a circuit that is not a very good digital circuit. Half of the time it's not clear whether the output is a logic 1 or a logic 0. This is a much better digital circuit:



Now the majority of the time it's pretty clear whether the output is a logic 1 or a logic 0. There's still that center region that is undefined but it's much smaller. For this circuit, a digital circuit, you want to operate as much of time as possible in the left and right regions where it's clear what the output is in terms of logic values.



ANALOG SECTION OF THE INVERTER TRANSFER FUNCTION AND SMALL SIGNAL VALUES:

In this section we'll use the same circuit as in the last section… If you're saying: "But that was a digital circuit and you're talking about analog now" you have a good point BUT you'll see in a minute that the inverter is just the digital name for that circuit and that exact same circuit can be called an inverting amplifier when looked at as an analog circuit. But a few other things first…

We've identified the important points on the transfer curve for digital. Now let's look at analog. Analog (EE308) is interested in how much a voltage changes when another voltage changes. The most common pair of voltages are input and output of a circuit. In digital you just cared what the output voltage was. You wanted to know if it was bigger than the threshold or smaller. That's all you needed to know. In analog, you want to know how much the output moves when you change the input. That may sound familiar:

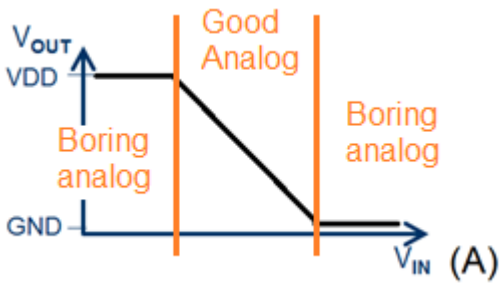
Value of interest: Vout/Vin

I'm pretty sure all of you have seen this in your textbooks and recognize this as "gain". This is WRONG!!!!!!! And one of my major pet peeves with textbooks that talk about gain. This equation is not gain. Here is the equation for gain:

$$Gain = A_V = \frac{\Delta Vout}{\Delta Vin}$$

Gain is how much the output changes for a change in the input.

Going back to the transfer function graph, can you see what the gain is in the digital friendly regions? How much does Vout change when Vin changes in the digital friendly regions? If you said zero, you are right. The output doesn't change at all even when the input changes in the good digital regions. Those regions are not very interesting to analog people. Now let's look at the region on the transfer function graph that is bad for digital. All the sudden it's interesting because the gain isn't zero. In this case it's some negative number because Vout goes lower when Vin gets higher. This is the region that analog designers like because the output changes with a change in the input. So the bad region for digital is the good region for analog and the bad region for analog is the good region for digital. And now you can see that the inverter can also be used as an inverting amplifier – you just operate in a different region of the transfer function. (Inverting means that the gain is negative – The output decreases when the input increases or vice versa).



As mentioned, in EE308 you get the tools to calculate the slope of the line at **a single specific point** on the line that connects the two flat nice logic lines (of the transfer function of the inverter) but calculating the entire line should be left to the simulators. The simulator calculates individual points and then connects them for you. So, again, let's not worry about calculating the shape of the line when working by hand.

One aside before going on: A_V is the symbol for voltage gain. It's used when looking at voltage in, voltage out systems. There are other systems that are commonly used but not talked about much in EE308 due to lack of time:

Gain name	Circuit input	Circuit output	Notation
Voltage gain	Voltage	Voltage	A_V
Current gain	Current	Current	A_i
Transconductance	Voltage	Current	Gm
Transresistance or transimpedance	Current	Voltage	Rm or Zm

In EE211 you saw ideal op-amps and they were considered voltage in voltage out circuits and therefore you talked about A_V.

Back to gain discussion: As mentioned, gain is the change in the output voltage over the change in the input voltage. Not the output voltage over the input voltage!!!!! Gain is a small signal value – Hopefully, in EE306, you heard the words "small signal". It was probably pretty fuzzy as to what that meant. EE308 will try to clear that definition up but one part of the definition of "small signal" is that the values (the small signal values) are all a change in something over the change in something else. All "small signal" values that you'll see are changes in something over the changes in something else. That's part of the small signal definition. When you started small signal analysis in EE306 it probably didn't make much sense. Let's take a moment to show one example of what the math related to small signal signals looks like. Small signal analysis is an analog technique that describes what is happening at a single point on a curve. We already said that gain is a change over another change:

$$Gain = A_V = \frac{\Delta Vout}{\Delta Vin}$$

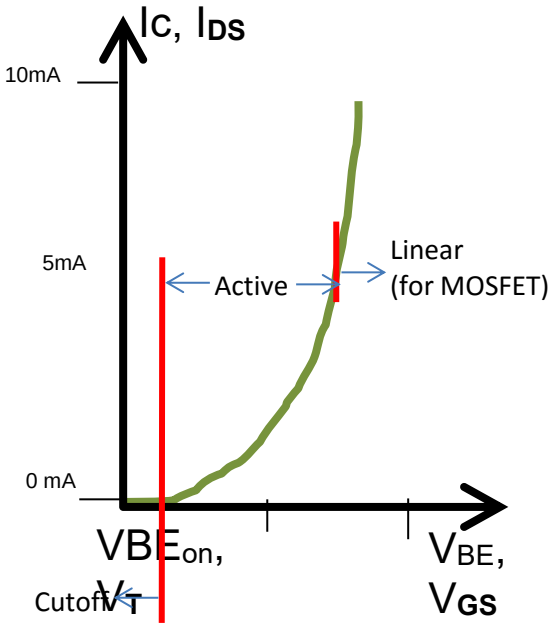
Generalizing this we can write:

$$Gain = \frac{\Delta output}{\Delta input}$$

Let's do the math for a BJT (the math for a MOSFET is the same except you use the MOSFET current equation instead of the BJT current equation). Here's the BJT current equation for a BJT in the active region:

$$I_C = I_S e^{\left(\frac{v_{BE}}{VT}\right)}\left(1+\frac{VCE}{VA}\right)$$

This describes what the current is for a particular VBE value. It's important that you see that IC is the output of a single transistor and that VBE is the input to the transistor.



When you change VBE, IC changes. If you make that into and equation you get:

$$Gain = \frac{\Delta output}{\Delta input} = \frac{\Delta IC}{\Delta VBE} = slope$$

Looking back at the table you can see that a transistor, on its own, because it has voltage in, current out is a transconductance amplifier and the gain, ΔIC/ΔVBE, is the "transconductance" of the transistor. … Take a second look at how the slope (ΔIC/ΔVBE) changes for different VBE values. So, really, there is no single answer for the gain (slope) of this graph across the entire active region. The slope changing throughout the active region is why it is so hard to draw the graph accurately by hand. What is done in analog is that you bias your circuit to operate around a single point on the curve and then use the values at that central point to describe your amplifier. If you've taken EE306 or EE308, you know that the first step of building a circuit is to "bias" the circuit to conduct a certain current by setting node voltages to values that cause the circuit to conduct that current. For this example, you must assume that the node voltages have been set to put the transistor at a point in the active region. We need to use a current equation and, if we don't know what region of operation the transistor is in, we can't pick the correct current equation.

One more piece before we do the math. Since we have set our transistor to be operating at a single point (which means that you've set VBE to conduct a particular IC through your transistor), we need to do something with the gain equation. The gain equation uses deltas which need two points. We only have one operating point (or Q-point). Think back to calculus and remember what happens when you take the delta in the denominator and make it smaller and smaller until it is infinitesimal. As the delta gets smaller and smaller, it gets close to a single point and that is the definition of a derivative. Remembering calculus, the gain equation becomes:

$$Gain = \frac{\Delta output}{\Delta input} = \frac{\Delta IC}{\Delta VBE} = slope = \frac{\partial IC}{\partial VBE}$$

Here's the math:

$$Gain = \frac{\partial IC}{\partial VBE} = \frac{\partial\left[I_S e^{\left(\frac{v_{BE}}{VT}\right)}\left(1+\frac{VCE}{VA}\right)\right]}{\partial VBE} = \frac{I_S e^{\left(\frac{v_{BE}}{VT}\right)}\left(1+\frac{VCE}{VA}\right)}{VT} = \frac{IC}{VT}$$

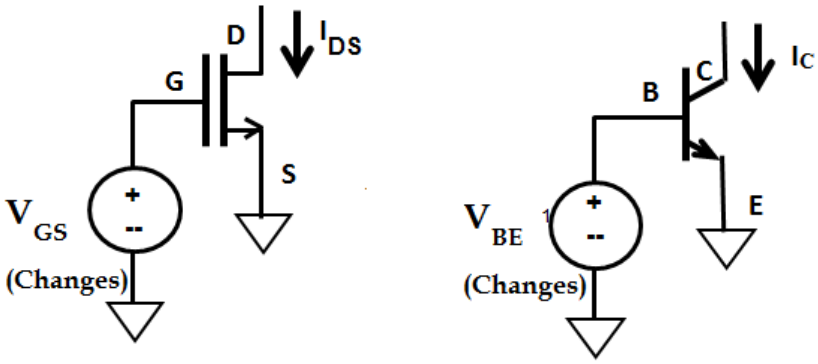
This says that the slope at the point you biased your transistor at is whatever the current you set through your transistor over the constant VT≈25mV or 26mV. All that work and now you've found the transconductance (gain) of a single transistor when you've set a specific current of IC through it.

This is how you use derivatives to find the change in various voltages and currents relative to each other in EE308. I just wanted to give you a taste (or reminder) here so the term small signal might make a little more sense to you. Small signal is really a derivative around your operating point.

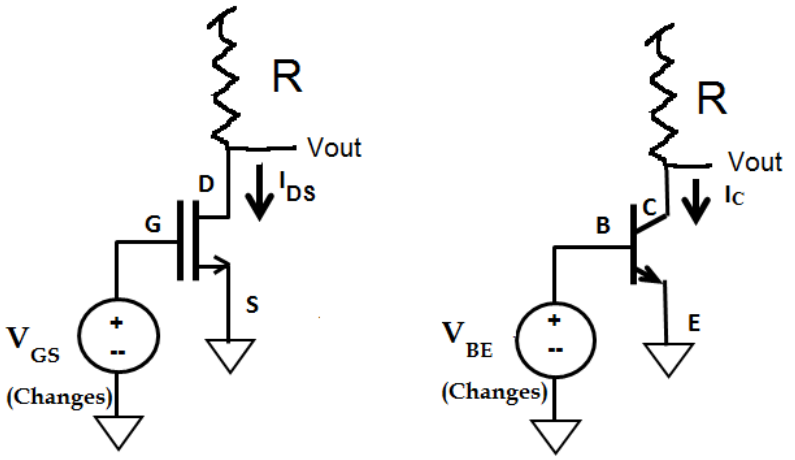
MAKING A VOLTAGE AMPLIFIER:

The last section talked about transconductance of a transistor. Remember that transconductance tells you how much the output current changes when the input voltage is changed. And remember that MOSFETs and BJTs are described by a current equation (which means that they are current output devices) and the equation says that if you change the input voltages, the (output) current changes. It's great to have a transconductance device but we want a voltage amplifier.

Here's a question for you: If you have some current, how do you turn that into a voltage? Think about an equation that relates voltage and current. Hopefully you came up with V=IR. The "Common source amplifier" (MOSFET) or "Common emitter amplifier" (BJT) does just that: It takes the current out of a transistor and puts it through a resistor. Here's a MOSFET and BJT transistor with a current output (transconductance devices):



Now, to change the current output into a voltage, let's force the current out of the transistor across a resistor:



Let's see what happens here. If I raise the gate or base voltage, VGS / VBE increases. The current equations say that if VGS or VBE increases, the current increases:

$$IDS = \frac{k'_P}{2}\frac{W}{L}(VGS - VTP)^2\left(1 + \lambda \cdot |VDS|\right)$$ (for MOSFETs)

$$I_C = I_S e^{\left(\frac{v_{BE}}{VT}\right)}\left(1 + \frac{VCE}{VA}\right)$$ (For BJTs)

If the current increase through the resistor, the voltage drop across the resistor increases and Vout goes down. In summary, if VG or VB goes up, Vout goes down. That means that these are inverting amplifiers.

You might want to know how much Vout changes for a change in VG or VB but that is too complex for this handout. For a complete description, take EE308 but here, I'll just show you a hand wavey version of what that gain value is. Note that now we have voltage input voltage output circuits so this is not a transconductance circuit anymore: It's a voltage amplifier.

This isn't 100% true but let's assume that all the current that goes through the transistor also goes through the resistor. We already figured out how much the output current changes through the transistor for a change in the input voltage of a BJT so let's use the BJT version of our voltage amplifier to show the gain calculation. We know that the current changes IC/VT for a change in VBE of a transistor when the transistor is in the active region.

$$Gain = \frac{\Delta output}{\Delta input} = \frac{\Delta VOUT}{\Delta VBE} = \frac{\Delta(IC \times R)}{\Delta VBE} = \frac{\Delta IC}{\Delta VBE} \times R = \frac{IC}{VT} R$$

And that's the (hand wavey version of) gain equation of the common emitter amplifier.

In general, if you want a voltage output, you take a transistor, which is a current output device, and put that current across a device that acts resistively. As a reminder, if you aren't operating the device in the active region then you get very little gain (take EE308 for details). Note that the bigger the "R" value is in the voltage gain equation for a BJT common emitter amplifier is bigger the bigger "R" is. There's a limit though and the bigger you make "R", the lower Vout is and the closer your transistor goes to the linear region. To keep good gain you must make the value of "R" small enough that the Vout keeps the transistor in the active region.

MAKING A DIGITAL (CMOS) CIRCUIT:

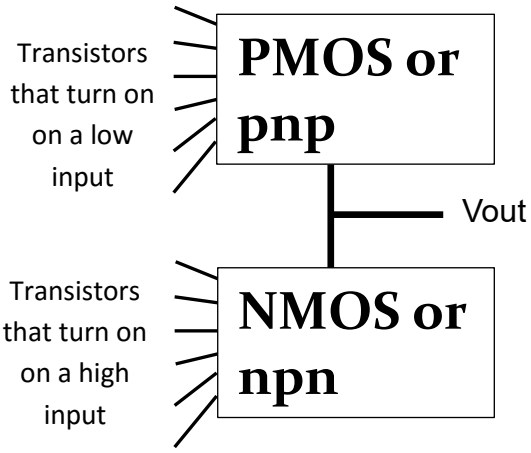
The last section talked about gain which is an analog value. This section we'll move back to digital and see how transistors are used as switches to implement digital circuits.

USING TRANSISTORS AS SWITCHES TO MAKE DIGITAL CIRCUITS:

So what can you do with this simple "switch" model of a transistor? The answer is: build digital circuits (EE307). Transistors in digital circuits (generally) either conduct the most they can or don't conduct at all – just like a switch. Conducting some but not the maximum that a transistor can conduct

really doesn't have much use in most digital circuits. (Again, I am generalizing – My thesis was on a way to implement digital circuits with transistors that run in that middle region between off and MAX current. That design technique is very powerful but wasn't mentioned in EE307. If interested, please ask me… But here let's again generalize and say, for digital circuits you generally want either max current or no current – transistor ON or transistor OFF).

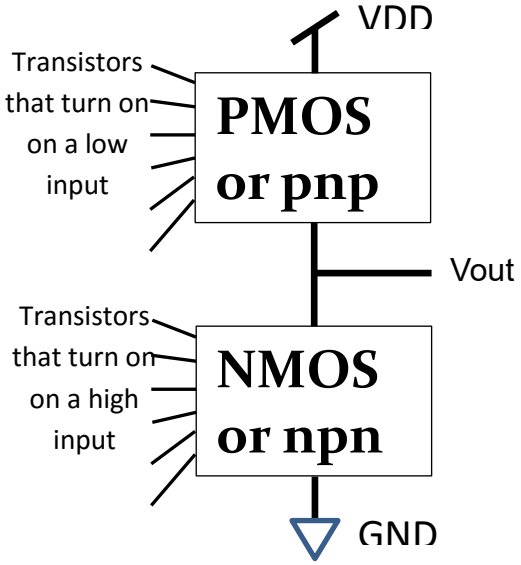
There are 10 or 15 well-known ways to create digital gates (like AND, OR, NAND, etc.) with transistors and more ways that aren't as well known. Examples of ways to implement gates with transistors you may have seen in EE307 include pseudo-NMOS, pass transistor, TTL, etc.. We'll talk about a way to implement digital gates called "CMOS". Where I went to school, CMOS just meant you used both types of MOSFET transistors in a circuit (PMOS and NMOS), but many if not most textbooks use CMOS to mean a specific design technique where you use a bunch of turn-on-on-high transistors (NMOS) below the output and a bunch of turn-on-on-low (PMOS) transistors above the output:



One thing to remember here is that the circuit as drawn to the right will not work. Just like a light that won't turn on unless there's power, almost ALL circuits (and all circuits that we'll talk about in EE308) **NEED POWER!!!!**

The correct way to draw that last circuit would be with VDD and GND:

These are the symbols for VDD and GND that I'll probably use through most of the class though other professors and textbooks may use variations.



LOGIC CIRCUITS:

For those of you that haven't had CPE133 or CPE129, here's a VERY quick picture of gates and how circuits are made from gates. Let's make a sentence:

"Blue is green"

That's false. In digital electronics, a system has a voltage that represents FALSE and a voltage that represent TRUE. In most circuit in EE307, VDD represented TRUE and GND represented FALSE but that is actually a choice of the circuit designer and is NOT always true. For example, the RS232 ports that have disappeared from many computers uses -3V to -15V to represent TRUE and 3V to 15V represents FALSE. Since there's a variety of voltages that can represent TRUE and a variety of voltages that can represent FALSE, it's easier to come up with a way to talk about TRUE and FALSE that doesn't require knowing the voltage that the designer picked to represent each. That's where the '1's and '0's are used in digital systems. '1' represents TRUE (no matter what voltage in the system is used) and '0' represents FALSE (no matter what voltage in the system is used). That allows you to talk about the function of the circuit or system without having to get into the electrical details. '1' is often called "Logical 1" and '0' is often called "Logical 0" to make it clear that it's not the voltages that the speaker is talking about.

So back to "Blue is green"… In a program you might write:

If X==Y then…

If X equals Blue and Y equals Green, then you get a statement that is either true or false. In this case it's FALSE. The program would then do:

If '0' then… OR If "Logical 0" then… OR If FALSE then…

Which says "If something-that's-not-true, do something". This won't do that something. You can get even more complicated sentences like:

If (((A==B) and (C==D)) or E) then…

The equal or not equal determination is more than I want to do in this handout so let me get rid of those and give you a different equation to use for our circuit example:

If (NOT(A and B) or NOT(C or NOT D)) then…

You've seen that IF just cares if the logic equation is true (evaluates to a logic 1) or false (evaluates to a logic 0) so here let's just make a circuit that evaluates the logic equation and ignore "IF" for now. Here's the logic equation presented with AND symbols (the dot) and OR symbols (the +) instead of the words AND and OR:

$$Y = \overline{(A \bullet B)} + \overline{(C + \overline{D})}$$

There's an order of operations in logic just as there is in math. Here's a list from the strongest (done first) to the weakest (done last):

1. Stuff inside parenthesis gets done first before outer layers.
2. NOT (negative sign in math) is most tightly bound and gets applied first.
3. AND (multiply in math) gets applied next.
4. OR (addition in math) gets applied last.

In getting our circuit we'll need one more piece of information and that's DeMorgan's laws:

$$\overline{A \bullet B} = \overline{A} + \overline{B}$$

$$\overline{A + B} = \overline{A} \bullet \overline{B}$$

DeMorgan's say that if you have a bar over ANDed things, you can separate that out to NOTed ORed things. You can also go backwards: If you have a string of ORed things, you can NOT them all, AND them all and then put a bar over them and you won't have changed the equation's value. This is true for the second case where you have an ORed equation with a bar over it. I'll assume that you've seen DeMorgan's and will stop here with the explanation.

To convert this equation into a CMOS circuit you must massage the logic equation into a form that:

1. has a single bar over the entire equation and
2. under that bar there can't be a bar that goes over more than one input.

So let's massage this equation a bit so that it fits that form:

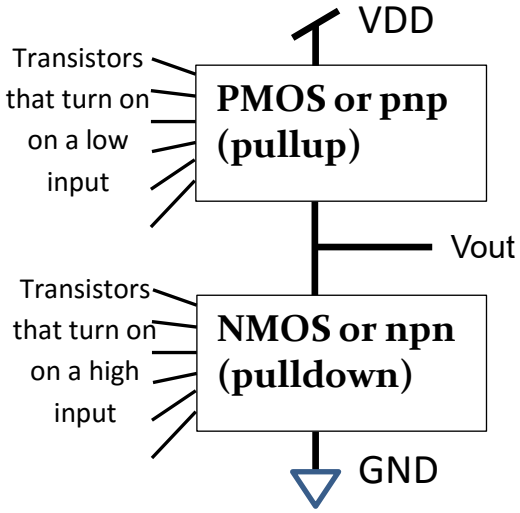
$$Y = \overline{(A \bullet B)} + \overline{(C + \overline{D})} = \overline{\overline{(A \bullet B)} + \overline{(C + \overline{D})}}$$

I negated the equation twice so it's the same. If I say not not something it means "something". I now have a bar over the whole equation but under that top bar, I have a bar going over more than one input. For example, the second bar goes over all four inputs. So now let's use DeMorgan's to get rid of the long bars under the top bar (which we want to keep) and the rule that says not not is is just… is:

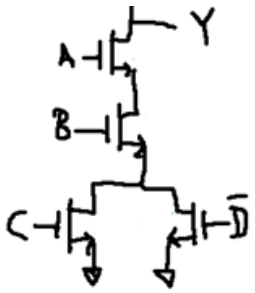
$$Y = \overline{\overline{(A \bullet B)} + \overline{(C + \overline{D})}} = \overline{\overline{(A \bullet B)} \bullet \overline{(C + \overline{D})}} = \overline{(A \bullet B) \bullet (C + \overline{D})}$$

This is the form you need to create the pullup (or PMOS) network and the pulldown (or NMOS) network. Let's do the NMOS pulldown network first. More rules:

1. ANDed terms go in series (That way, there will only be current through them when they are both on = both have high input).
2. ORed terms go in parallel (That way, there will be current through when either of them are on = one or both have high inputs).



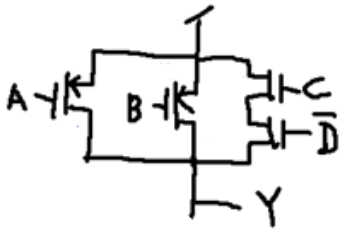
That gives the following pulldown network:



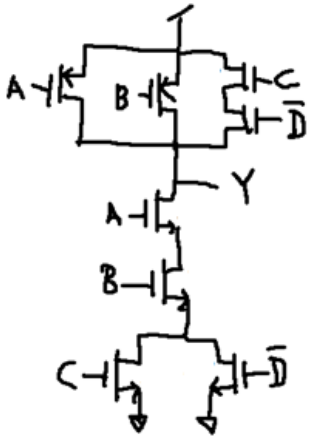
Now for the pullup network. It needs to pullup when the pulldown network isn't pulling down. In other words, it needs to turn on a path of conducting (or low impedance as hopefully your textbook called it) transistors so the output is connected to VDD (but not to GND). More rules:

1. ANDed terms go in parallel (That way, there will be current through when either of them are on= input low. When input1 OR input2 are low).
2. ORed terms go in series (That way, there will only be current through them when they are both on= input high. When input1 AND input2 are low).

Doing that you get the pullup network:



Put them together and you get a CMOS logic circuit that implements $Y = \overline{(A \bullet B)} + \overline{(C + \bar{D})}$:



The way CMOS works is that when there is a path that is conducting current, that means that there is a low impedance path ➔ The path that the current flows through is a low impedance path. A low impedance path is like a wire. When you connect something to a voltage through a wire, the something becomes that voltage. In CMOS' case, the pulldown network can make a path from GND to the output ➔ Connects the output to GND through a wire (low impedance path) to GND and makes the output become GND. OR the pullup network can make a path from VDD to the output ➔ Connects the output through a wire to VDD and makes the output become VDD. The two networks basically tie a wire from the output to either GND or VDD but never (intentionally) both at the same time. Having both the pullup network and pulldown network conduct at the same time would give you a wire from VDD to GND and probably some smoke. Connecting VDD to GND causes a lot of current to flow and that means fried parts. You can check and make sure it only conducts and pulls Y down to GND (I'm taking GND as logical 0) when the original equation gives you Y=0.

One thing I forgot to mention: Basic logic gates such as NAND, AND, 3input NAND, OR, XNOR, etc. can all be designed following these same rules. The example I did here was of a random logic equation and when a random logic equation is put into transistors as we did, the resulting circuit is usually called a "complex logic gate".

I also used analog transistor symbols. These, technically, should be drawn as digital transistor symbols.

There are many ways to build a logic gate and we just did one. The advantages of CMOS is that it is relatively fast, robust against environmental noise, easy to design as compared to some of the other methods of gate design in transistors, relatively easy to automate and lower power than some other families. It is, by far the most commonly used logic family. Ask me if you'd like to know more about that but I'm guessing at something like upwards of 95% of the digital chips made use CMOS.

CAPACITANCE AND DELAY:

Finally the last topic. The speed of a circuit can't be described unless the intrinsic (inherent) capacitance is understood. Capacitance is always mentioned when discussing how fast a circuit can change values but delay is a combination of the capacitance on the node where the value needs to change and the resistance in the path of the current on to or off of the capacitance. So, to understand delay, first you need to know where the capacitance comes from. Here are the cross sections of an NMOS (MOSFET) and an npn (BJT):

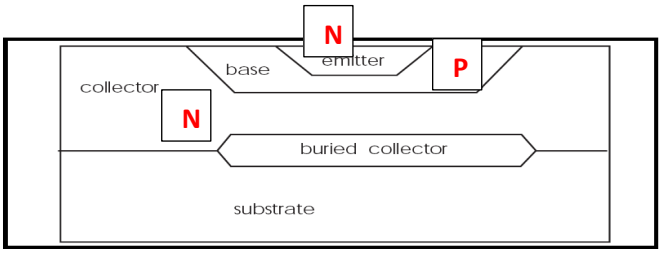
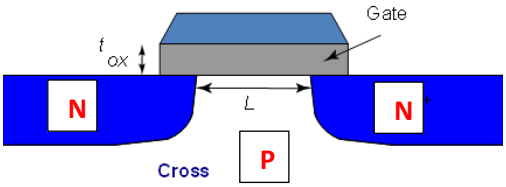


Figure 13-7: Vertical Transistor (SUBS = +1)



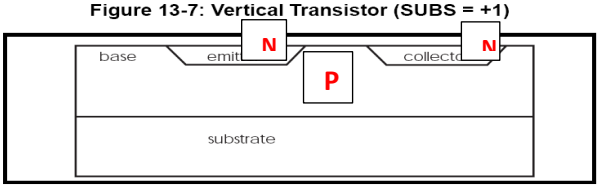


Figure 13-8: Lateral Transistor (SUBS = -1)

MOSFET cross section BJT cross section – Top is usual, bottom is more for BiCMOS

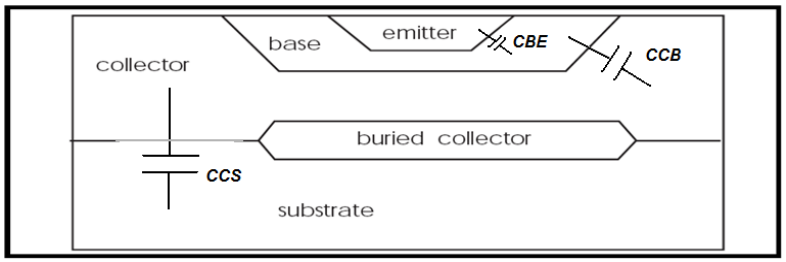
If you have had EE306 the P and N regions should make sense. (If anyone hasn't had EE306, I'm not sure that you should be in any of the classes that I'm asking you to do this reading for). The capacitance that comes directly from the transistor comes from these PN junctions. (In the MOSFET, there's an additional capacitance from the gate but I'll save details on that until class).

There are basically two types of capacitances associated with PN junctions:

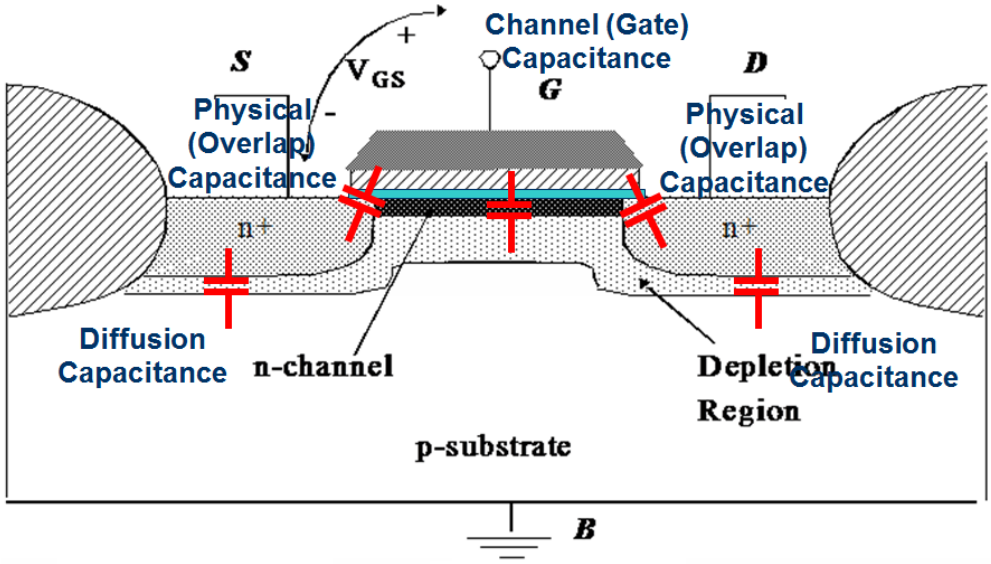
1. Junction capacitance: This is related to the depletion region and has to do with PN junctions that are reverse biased.
2. Diffusion capacitance: Also called charge storage capacitance and has to do with PN junctions that are forward biased.

A capacitor is two conductors with a non-conductor between them. The conductors are close enough to see each other but the charge on one conductor can't flow to another due to the insulator between them. It is an electric field. If you want to simplify Cs, Ls and Rs: Cs are the effects from electric fields, Ls are the effects from magnetic fields and Rs are just how hard it is for an electron to get from one place to another. Impedance is also how hard it is for an electron to get from one place to another but it includes electric field and magnetic field effects. When you have an OFF biased PN junction you have n-material, p-material and a depletion region between that keeps the electrons from freely flowing. This is how the junction capacitance works. Conductor (N or P material), insulator (depletion region), and another conductor (P or N material).

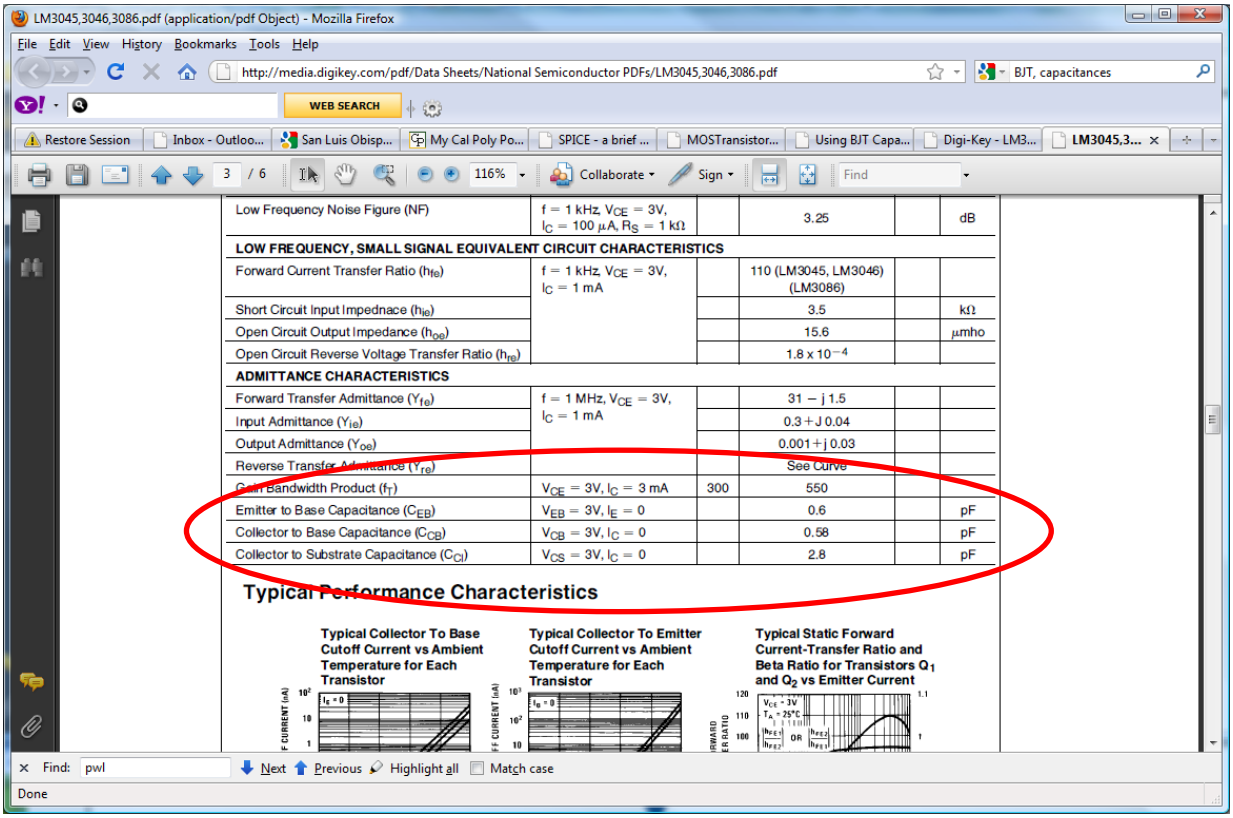
The Diffusion capacitance is a little more complex. Under steady state there will be a concentration of minority carriers in the base and a concentration of minority holes in the emitter. When the base voltage increases, the minority carriers in both the base and emitter must change and both the emitter and base must change their charge by the same amount. Until the charge is changed, the circuit isn't in steady state and this delay while electrons move around is a capacitance-like behavior. So, Even though the base-emitter junction is forward biased, its capacitance-like behavior can't be ignored. The diagram below shows all the BJT capacitances. Each PN junction contributes some capacitive effect so there are three capacitors; CCS (collector substrate), CCB (collector base) and CBE (base emitter).



Here's the capacitances for a MOSFET:



Data sheets will often give you information on each PN junction's capacitance. For example, here's the capacitance for National Semiconductor's LM3045/LM3046/LM3086 Transistor Arrays:



Here's the capacitance graphs for the transistors, NPN – MPSA06, PNP – MPSA56:

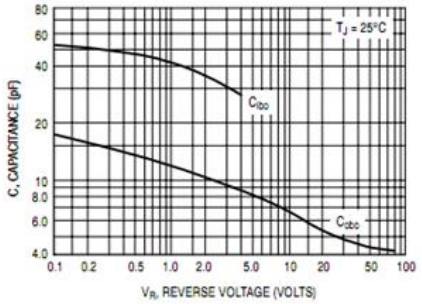


Figure 4. MPSA05/06 Capacitance

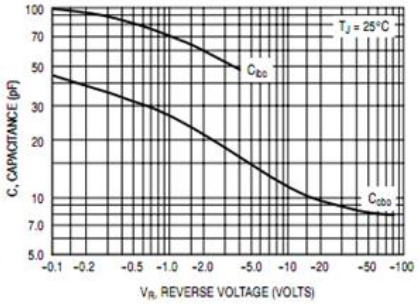


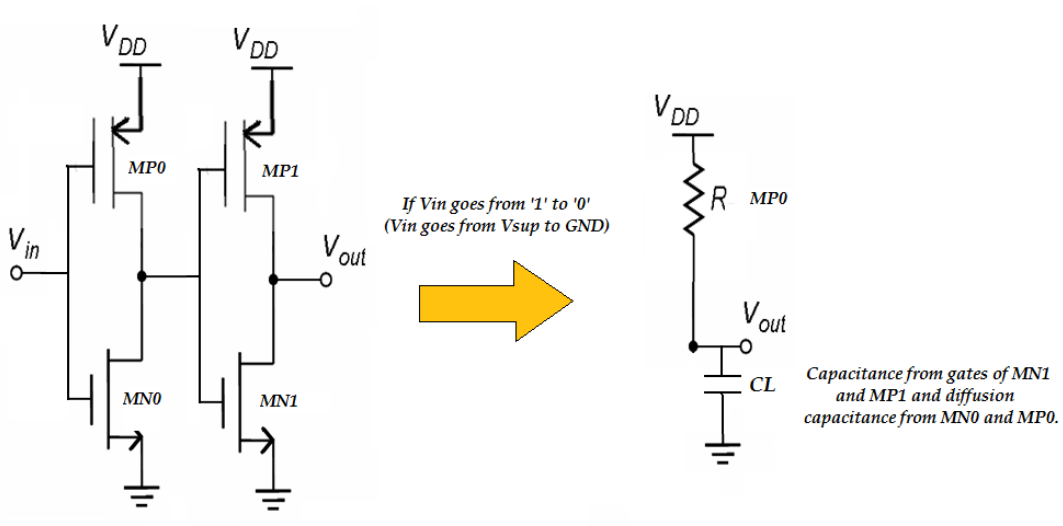
Figure 5. MPSA55/56 Capacitance

These are the graphs for the 56/06 transistors that I sometimes use in EE348. Note that capacitances are not constant. As the base voltages changes the depletion regions change and the capacitance changes.

OK. That's where the capacitance on the transistors come from but there are other contributors to capacitance on a chip. For example, wires couple to other wires, the substrate and to transistors. So the description above is simplified and used to help people are trying to do approximate C calculations on paper. To really get an accurate C value you generally have to use software tools.

Let's look at the delay on an inverter. This is a digital gate. Analog delay and digital delay are calculated differently so **THIS METHOD IS NOT VALID FOR ANALOG CIRCUITS**! You probably talked about delays in EE307. There was a threshold voltage that, if the output of your circuit rose above or

dropped below, you were happy and you said your circuit had switched from a logic '1' to a logic '0' or vice versa. You were worried about a capacitance that was charging up through a resistance:



Hmmm… The second circuit doesn't look much like the first…. Remember back in the first part of this handout where models were discussed? The main rule of using a model to help calculate some value is that you use the simplest model available so as to make your math as easy as possible. The second circuit is, indeed, the first circuit but some simplifications have been done. When you make Vin go to GND, the NMOS MN0 turns off an acts like an open circuit. If it isn't conducting then it doesn't need to be in the picture you use to analyze the behavior. In the picture above, there's some voltage across the PMOS and some current flowing through it so it can be modeled as a resistor (A real resistor – not a small signal resistor). Another thing you can simplify is the second inverter. When trying to find out how fast the node between the inverters change, all you need to know is the resistance in the path of whatever is bringing current on to that node and the capacitance at that node. No current goes into the gates of the second inverter (MN1 and MP1) but the gates (bases if these were BJTs) contribute some C to that node. CL is the capacitance contributed by the gates of the second inverter and MP0 and MN0. The value of CL can be calculated but the point of this section is that the two inverters in series shown above can be modeled as the R that the PMOS is acting like onto a C which is the sum of all the C at the node between the inverters. We won't calculate the C and R value in this handout – We'll learn how to that in EE308 or if you take EE307 from me. And EE431 will go into depth on that too. But let's look how the delay is calculated.

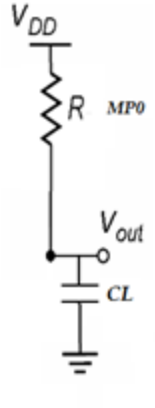
For digital, you find the speed that a circuit can switch by changing its input instantaneously and then watching the node of interest. When Vin is changed instantaneously from VDD to GND, the node between the two inverters has to switch from GND to VDD. What is important to you is not when that node gets to VDD but when it crosses a threshold voltage that designates the change between a logic 1 and a logic 0. In real life that threshold is actually determined by the when next circuit (the second inverter in this case) starts to see its input as having changed from a 1 to a 0 or a 0 to a 1. So, in this case, the first inverter starts to switch and the threshold exists at the point when the second inverter sees its input as having switched logic levels. Though it hardly ever happens in the real world, that threshold voltage was probably Vm=VDD/2 in your textbooks (it's traditional and makes the math easier) so let's use that. In EE307, you probably calculated:

$$Vout = VDD \times \left(1 - e^{\frac{-t}{RC}}\right) \quad \Rightarrow \quad \frac{VDD}{2} = VDD \times \left(1 - e^{\frac{-t}{RC}}\right) \quad \Rightarrow \quad \frac{1}{2} = e^{\frac{-t}{RC}}$$

$$\Rightarrow \quad t = -RC\ln\left(\frac{1}{2}\right) \quad \Rightarrow \quad t = 0.69RC$$

This is the time between the time when the input changes until the time the output gets to VDD/2. The R is whatever resistance the PMOS MP0 looks like and C is C_L, the capacitance from everything at the node between the two inverters. This is called the propagation delay (rise time and fall time are something different) and it is a digital metric that describes how fast a device transitions.

Analog is a different animal so there's a whole different metric you look at. You still use an R and a C in the analog delay and, though the C calculation is the same as the digital case, the R calculation is different. I won't cover how to calculate the R value for an analog circuit here (that's an EE308 topic) but, in order to do the math, I'll assume we have the "analog" R. I'll use the same circuit in my analog delay example but I will call the (digital) inverter circuits, inverting amplifiers to make them sound more analogy. (Inverters can be pretty good inverting amplifiers!). The next question is: At the possibility of depressing you, do you all remember Bode plots? If you look at the first inverting amplifier as being "loaded" by the second (loading means that the second inverting amplifier is putting C on that node and making the transition slower. It can also mean extra R but in this case C is all that matters) you get the same model:



If you look at this circuit and remember EE211, you'll see that then equation for Vout as it goes from GND to VDD looks like:

$$H(j\omega) = \frac{\frac{1}{j\omega C_L}}{\frac{1}{j\omega C_L} + R} = \frac{1}{1 + j\omega RC_L}$$

If you remember your Bode plot stuff you can see that there is a pole at: $\omega = \frac{1}{RC_L}$. After a pole, the magnitude starts to drop at 20dB/decade. In analog terms that means that you circuit works as expected up to $\omega = \frac{1}{RC_L}$ but after that the signals can't get through your circuit fast enough to swing to the maximum and minimum voltages. There's just not enough time for the capacitor to charge and

discharge to get to those voltages. What that says about your inverter is that if you put a sine wave into it, it will start to underperform at $\omega = \frac{1}{RC_L}$. That's how fast your analog circuit can run….

IN CONCLUSION:

This handout has gone through **A LOT** of material. I'm sure some of you are kind of reeling from information overload. That's OK. Ask questions and I'm happy to go through the handout if that's what you'd like to do. We could even schedule a time and have a group review session…

For those of you that are in EE431 but haven't had one or both of the EE classes, I'd like for you to come talk to me and officially go through the handout. For those of you in EE308 that haven't had EE307, the same goes but just the parts that we won't get in EE308.

For EE308, there's a homework that this reading should help with. For EE431, I'll post the EE308 homework so you can get some practice but it won't be due.