EE204FZ
Lecture 2
p-n Junction Diodes

Zhu DIAO

Email: zhu.diao@mu.ie

Linear Devices

- Linear electronic devices are those whose output signal is a linear function of the input signal (or what we often say, a “straight line”).
- Important properties of linear devices:
  - Output signal is proportional to input signal;
  - Linear devices obey the superposition principle;
  - In AC operation, the input and output signals appear at the same frequency (no frequency conversion).
- Typical linear electronic devices include resistors, capacitors, and inductors when driven at low currents.

Nonlinear Devices

- Nonlinear devices are those whose output signal is not linearly proportional to the input signal.
- Important properties of nonlinear devices:
  - Output signal is not a linear function of input signal;
  - Nonlinear devices in general do not obey the superposition principle;
  - In AC operation, the input and output signals do not appear in the same frequency (frequency conversion).
- Typical nonlinear electronic devices include diodes, transistors, mixers, etc. All devices discussed in this course module are, strictly speaking, nonlinear devices (nonlinearity leads to functionality).

Questions

- Is a p-n junction a diode?
- Is a diode a p-n junction?

What is a Diode?

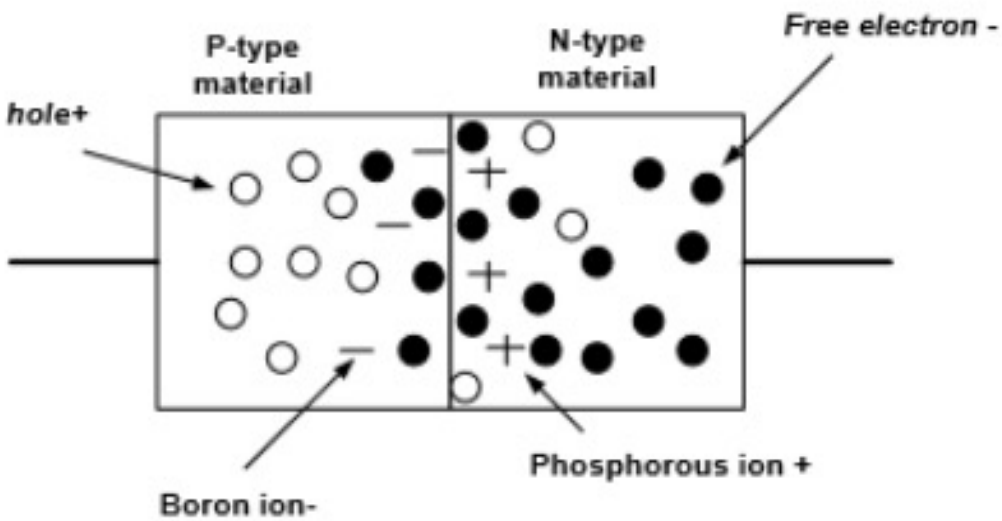
- A diode is a **two-terminal** electronic component that conducts current primarily in **one direction**.
- Diode has low (ideally zero) resistance in one direction and high (ideally infinite) resistance in the other direction.
- Types of diodes:
  - Thermionic diodes (vacuum tubes);
  - Semiconductor diodes:
    - Point-contact diodes;
    - Schottky diodes;
    - **p-n junction diodes.**

p-n Junction

- If a piece of intrinsic silicon is doped so that part is n type and the other part is p type, a p-n junction is formed at the boundary between the two regions (a p-n junction diode is created).
- The p region has many holes (majority carriers) and only a few free electrons (minority carriers) while the n region has many free electrons (majority carriers) and only a few holes (minority carriers).

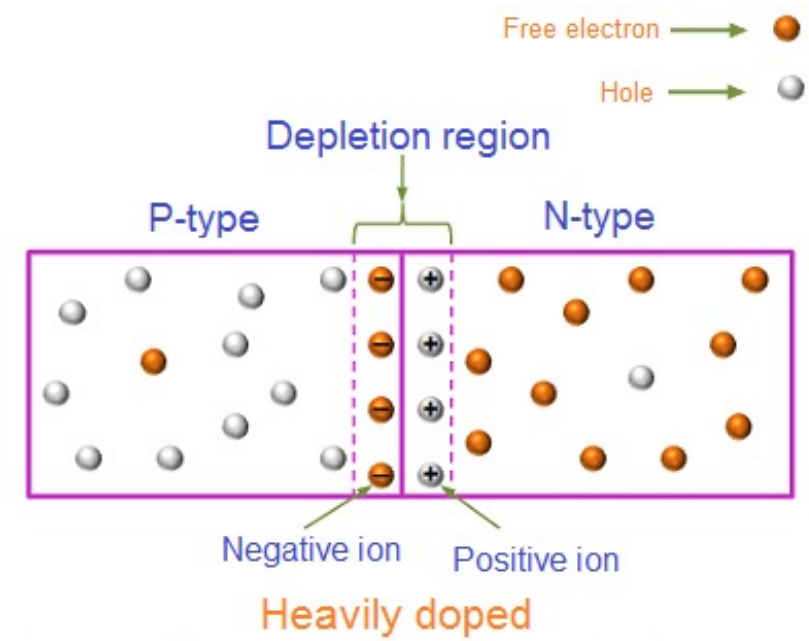
p-n Junction

- The free electrons in the n region randomly drift in all directions. At the instance of p-n junction formation, the electrons near the junction begin to diffuse across and combine with holes in the p region near the junction.
- Before the p-n junction is formed, both sides are neutral in terms of net charge.
- After the p-n junction is formed, the region in the n region which loses free electrons form a layer of positive charges. The region in the p region where holes are combining with electrons forms a layer of negative charges.
- These two layers of charge form the depletion region.



Depletion Region

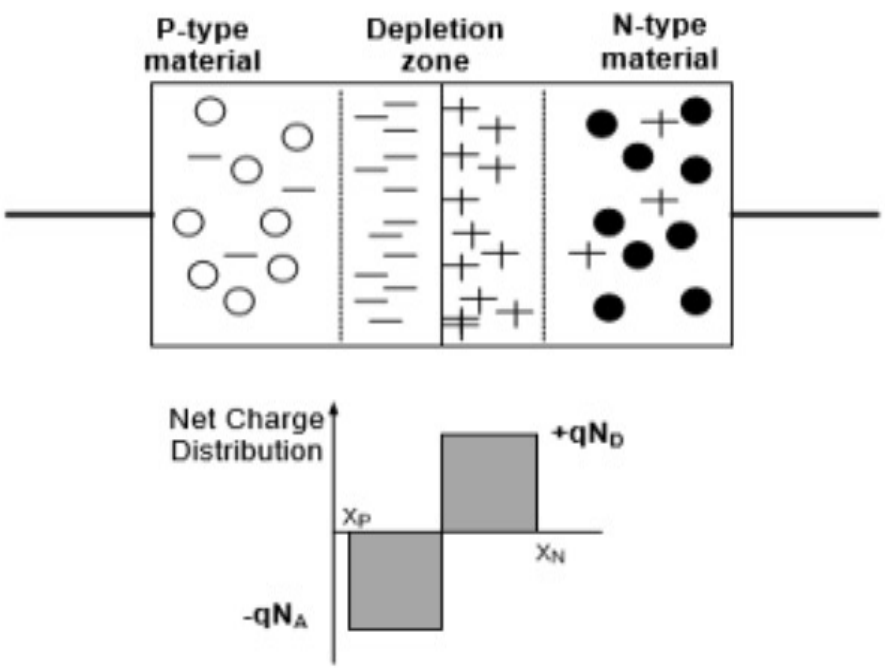
- The depletion region is very thin (~1 μm) compared to the n and p segments. It acts as a barrier to the further movement of majority carriers.
- The barrier potential is the potential difference that leads to the electric field across the depletion region. It is measured in volts (V). A certain amount of electric potential difference equal to the barrier potential (and the correct polarity) must be applied to a p-n junction before electrons can flow across the junction.



Copyright © Physics and Radio-Electronics, All rights reserved

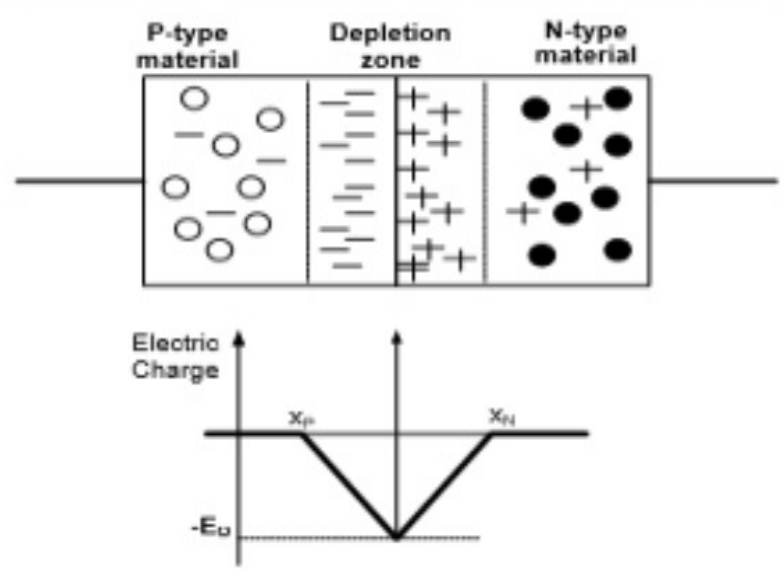
Depletion Region

- As an electric field exists across the junction and the positive and negative charge densities can be separated. The figure here shows the charge distribution in the p-n junction assuming that the p-type and the n-type materials are doped to similar levels. Note if the doping is equivalent on both sides then $x_p = x_N$.

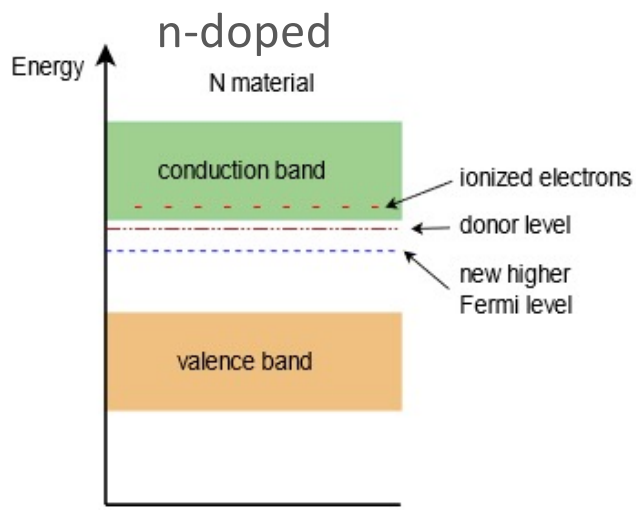


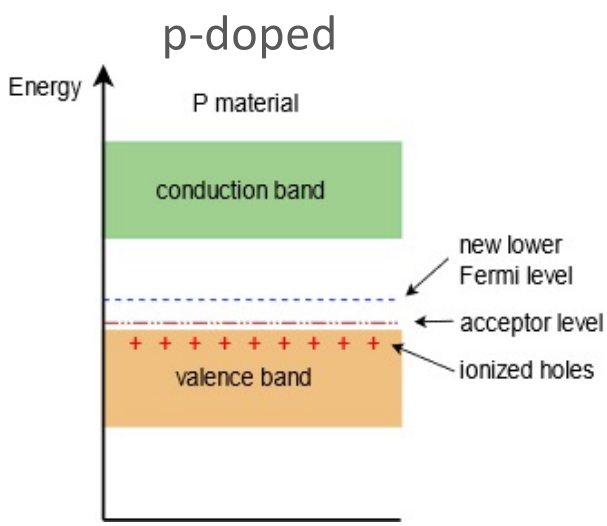
Width of the Depletion Region

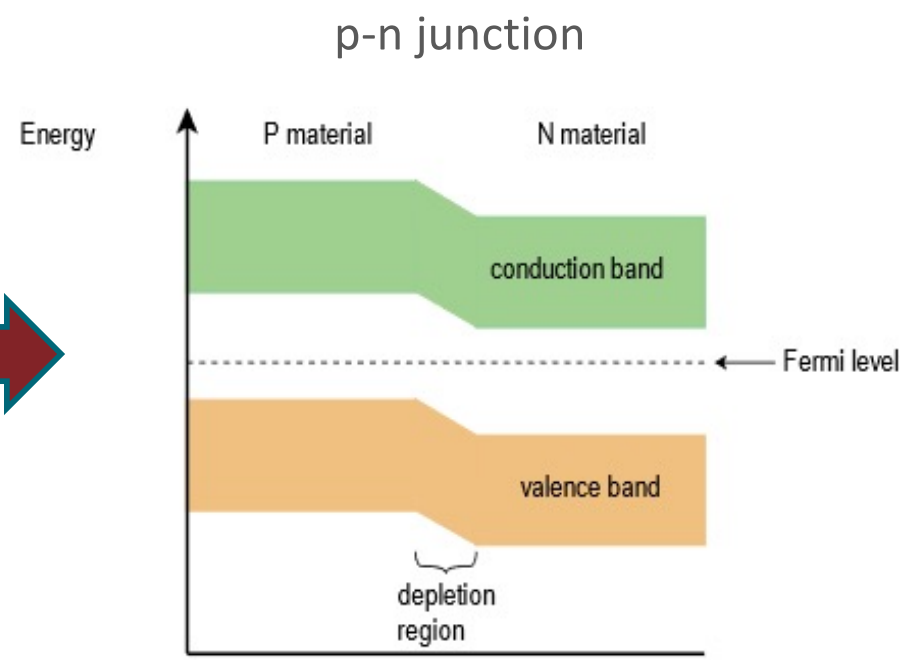
- The depletion region extends into the p-type area to a point designated as x_P and extends into the n-type area to a point designated as x_N. The total width of the depletion region is given by $W = x_P + x_N$.
- The overall charge in the p-type area of the depletion region is equal to the number of ionised acceptor atoms (N_A) per unit area multiplied by the p-type area involved while the charge in the n-type area of the depletion region is equal to the number of ionised donor atoms (N_D) per unit area multiplied by the n-type area involved.



Band Diagram of a p-n Junction



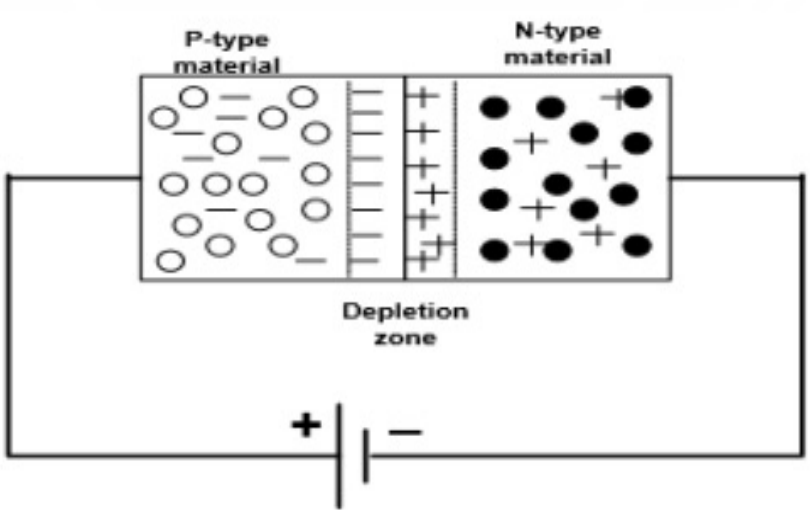




https://eng.libretexts.org/Bookshelves/Electrical_Engineering/

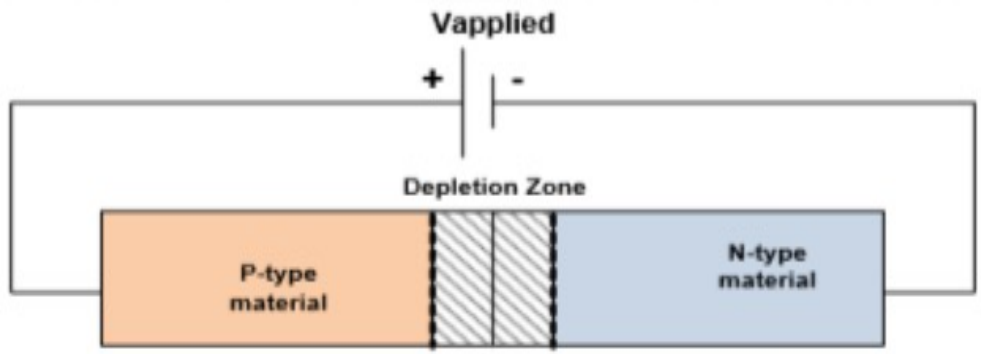
Forward Biasing a p-n Junction

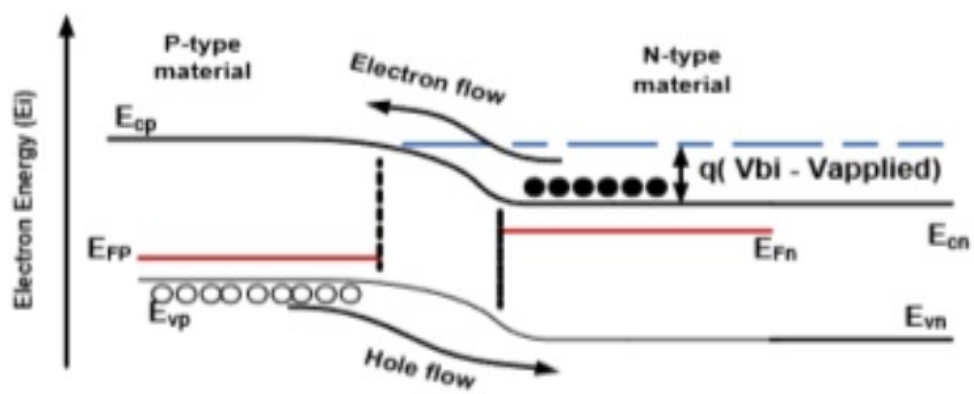
- When a forward bias voltage is applied to a p-n junction, the potential barrier of the p-n junction is lowered allowing electrons and holes to flow across the depletion region. When holes flow from the p-type material across the p-n junction into the n-type material they will then become minority carriers in the n-region. They will also be replenished by the positive terminal of the battery as they diffuse into the n-region. Similarly, when electrons flow from the n-type material across the p-n junction into the p-type material they will then become minority carriers in the p-region. They will also be replenished by the negative terminal of the battery connected to the n-type material and a major diffusion of electrons to the p-region takes place.



Forward Biasing a p-n Junction

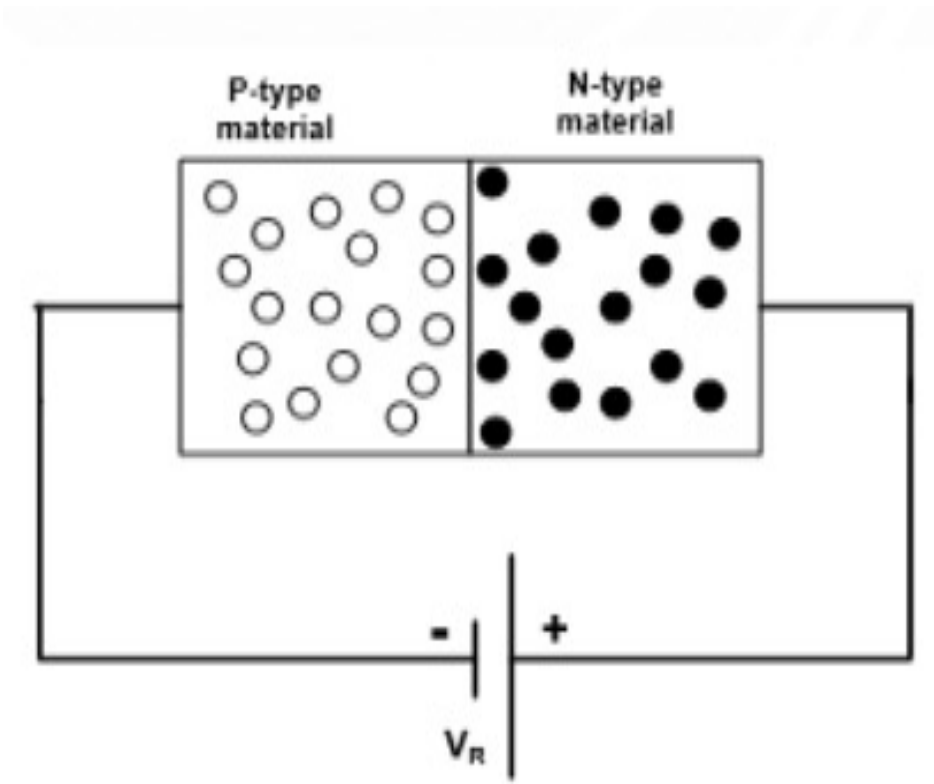
- Note that the potential barrier is much reduced and all the applied potential difference is across the junction.





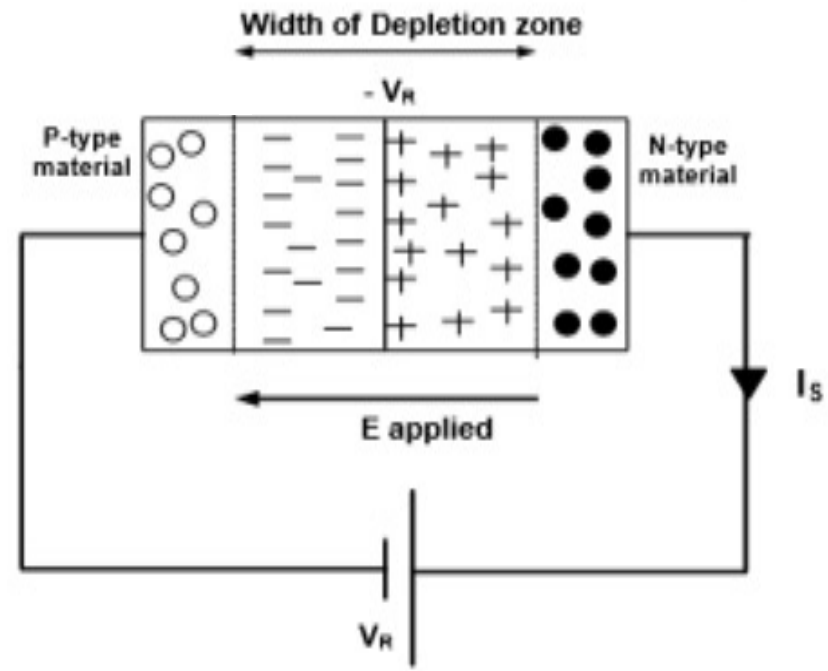
Reverse Biasing a p-n Junction

- Consider the case where a positive electric potential is applied to the n-region with respect to the p-region. In this case the diode is considered to have a reverse bias applied.



Reverse Biasing a p-n Junction

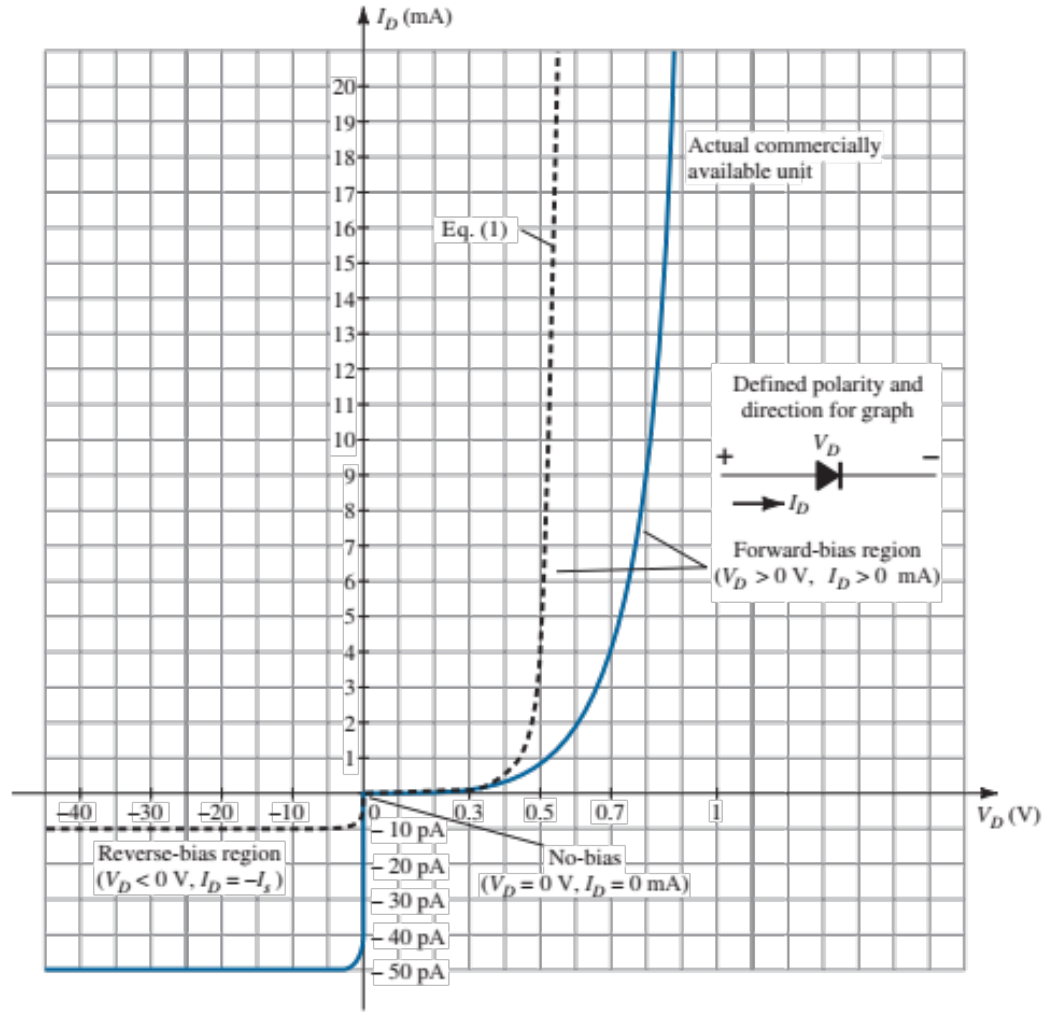
- The negative potential will cause holes in the p-type material to move away from the junction and towards the contact which will result in more exposed negative acceptor ions and thus creating a wider depletion zone. Similarly, the positive potential will cause electrons in the n-type material to move away from the junction and towards the contact which will result in more exposed positive donor ions, and thus creating a wider depletion zone on the n side as well.



(Shockley) Diode Equation

$$I_D = I_s(e^{\frac{V_D}{nV_T}}-1)$$

- I_D is the diode current;
- I_S is the reverse saturation current;
- V_D is the electric potential difference across the diode;
- n is called the ideality factor. It usually takes a value between 1 and 2 (at RT, $n \approx 1$ for Ge and $n \approx 2$ for Si);
- $V_T = \frac{k_B T}{q}$ (~26 mV at RT) is the thermal voltage and $q = 1.602 \times 10^{-19}$ C is the electron charge.



Summary: p-n Junction

- The p-n junction serves an important role in electronic systems and in the understanding of the behaviour of semiconductor devices. It is used in rectification, switching, photonics, as well as being a building block of the Bipolar Junction Transistor (BJT), the Junction Field Effect Transistor (JFET) and the Metal Oxide Semiconductor Field Effect Transistor (MOSFET) which we will cover in this course!

EE204FZ
Lecture 3
Junction Field Effect Transistors (JFETs)

Zhu DIAO

Email: zhu.diao@mu.ie

What is a Transistor?

- There are two basic types of transistors and lots of different ways to build them.
  - A current controlled variable “resistor” -> BJT
  - A voltage controlled variable “resistor” -> FET
- Using “resistor” here is a bit dodgy but basically by applying a given voltage across the device, one can change something in the device that varies the current.

What is a Transistor?

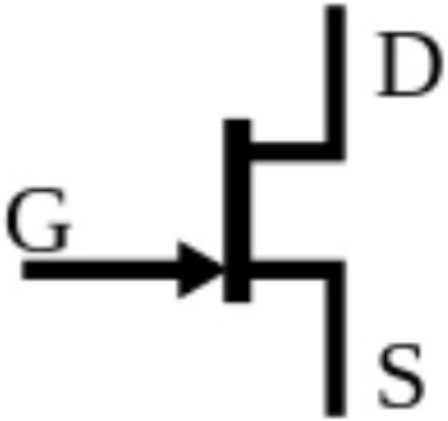
- FETs come in several forms, the most common are
  - Bipolar Junction Transistors (BJTs)
  - Field Effect Transistors (JFETs and MOSFETs)
- Now we also have
  - FinFETs
  - Multi-Gate FETs (MuGFETs)
  - Metal-Semiconductor FETs (MESFETs)
  - And lots of different other types.

Which Type of Transistors is the Best?

BJTs	FETs
First type of transistors, used widely in old designs.	Used widely in modern designs.
Best for low-noise performance.	A bit noisy.
Run the fastest (> 50 GHz) & materials dependent.	Usually slower, but modern FETs can operate at ~ 10 GHz.
Not easy to integrate.	Scales very well with modern fab technology, a very large number of FETs can be packed on the same chip.
Not power efficient.	FETs circuits can be extremely power efficient.
Current controlled.	Voltage controlled.

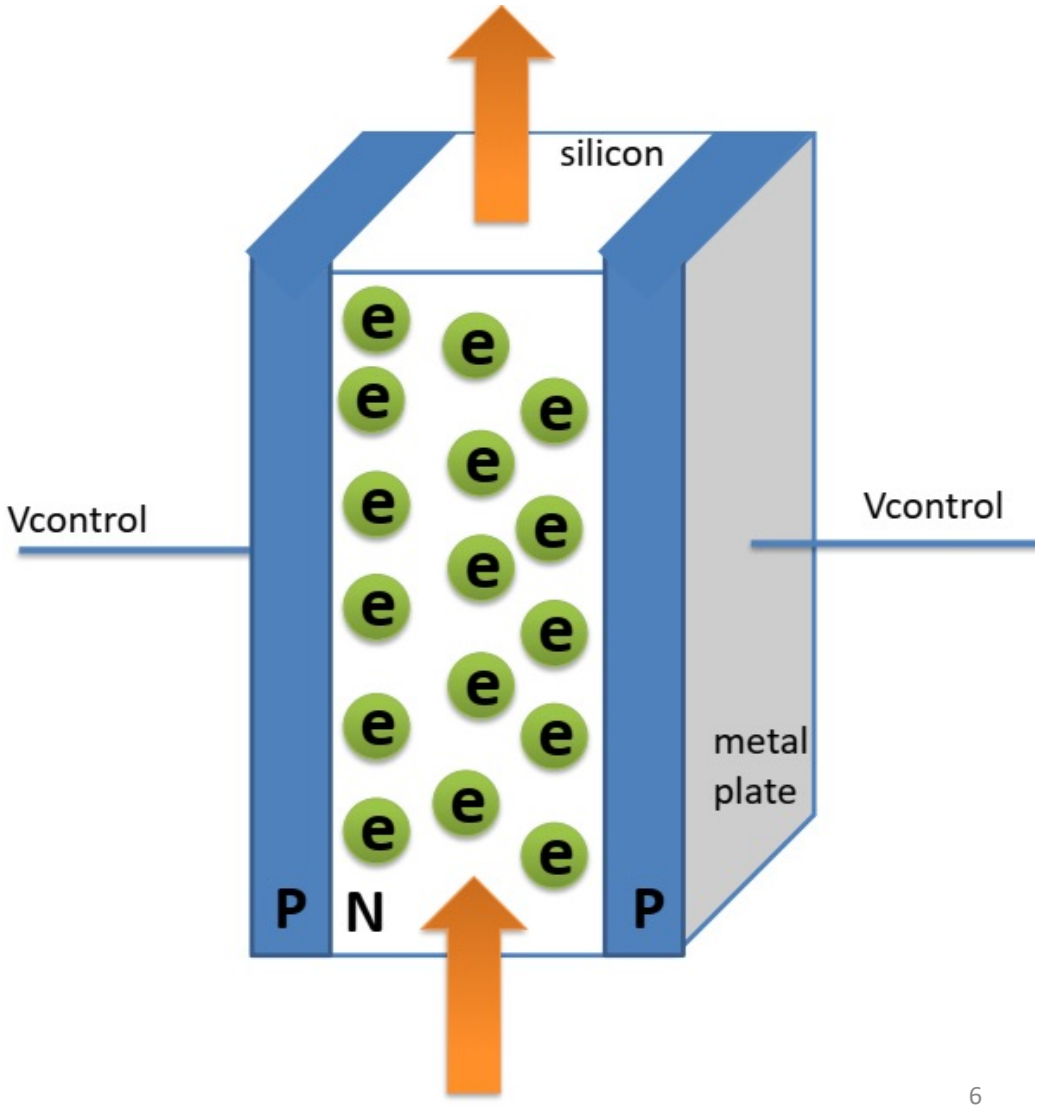
- We use mostly FETs in this course (including the lab sessions).
- When you are doing analogue electronics designs, select the transistor that best suits your applications.

What is a Field-Effect Transistor?



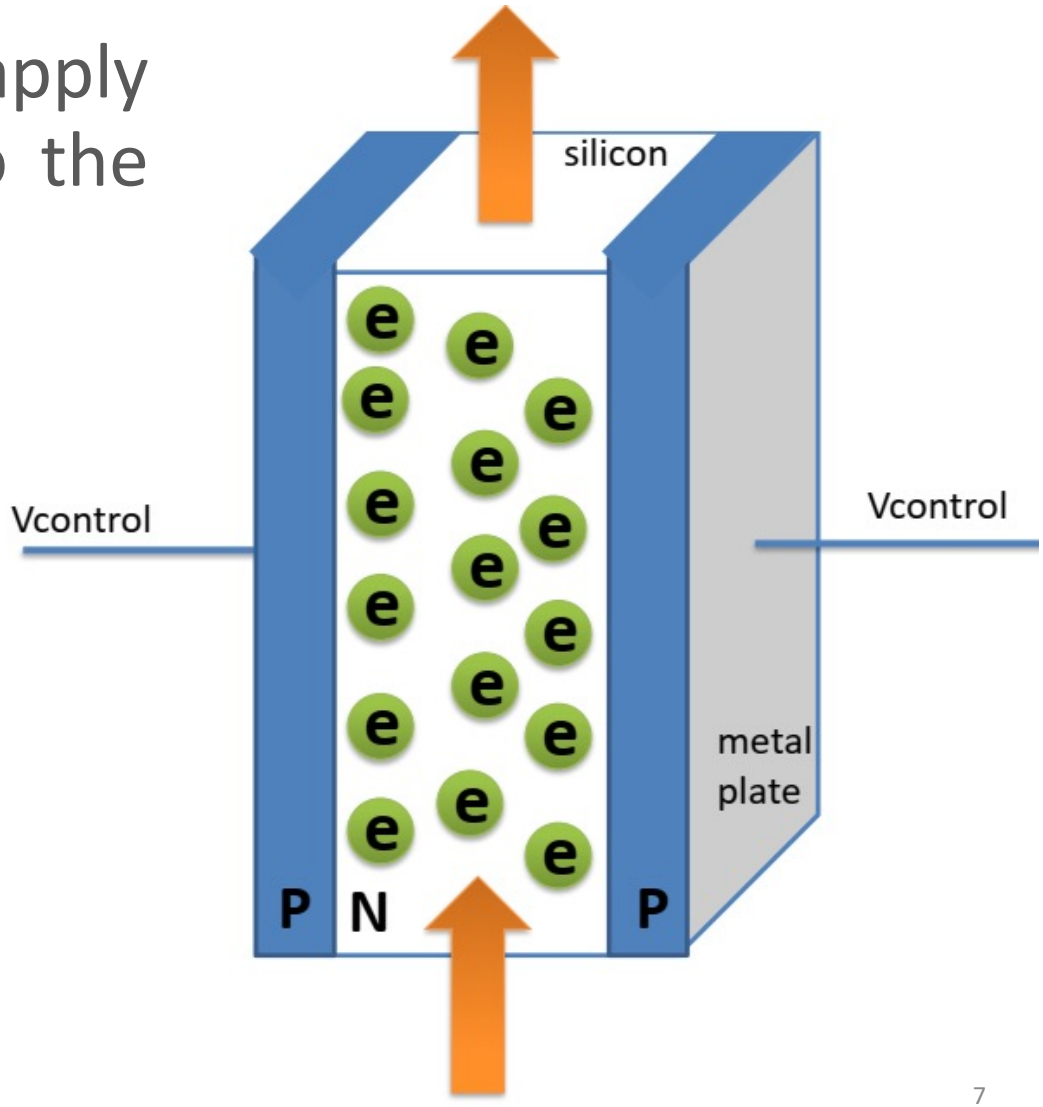
- More seriously, FET is a transistor whose behaviour can be controlled by electric fields. Electric fields are generated by electric potential differences (voltages), so FETs are voltage-controlled devices.

Junction Field Effect Transistors (JFETs)



Junction Field Effect Transistors (JFETs)

- What happens if we apply a negative voltage to the blue plates?

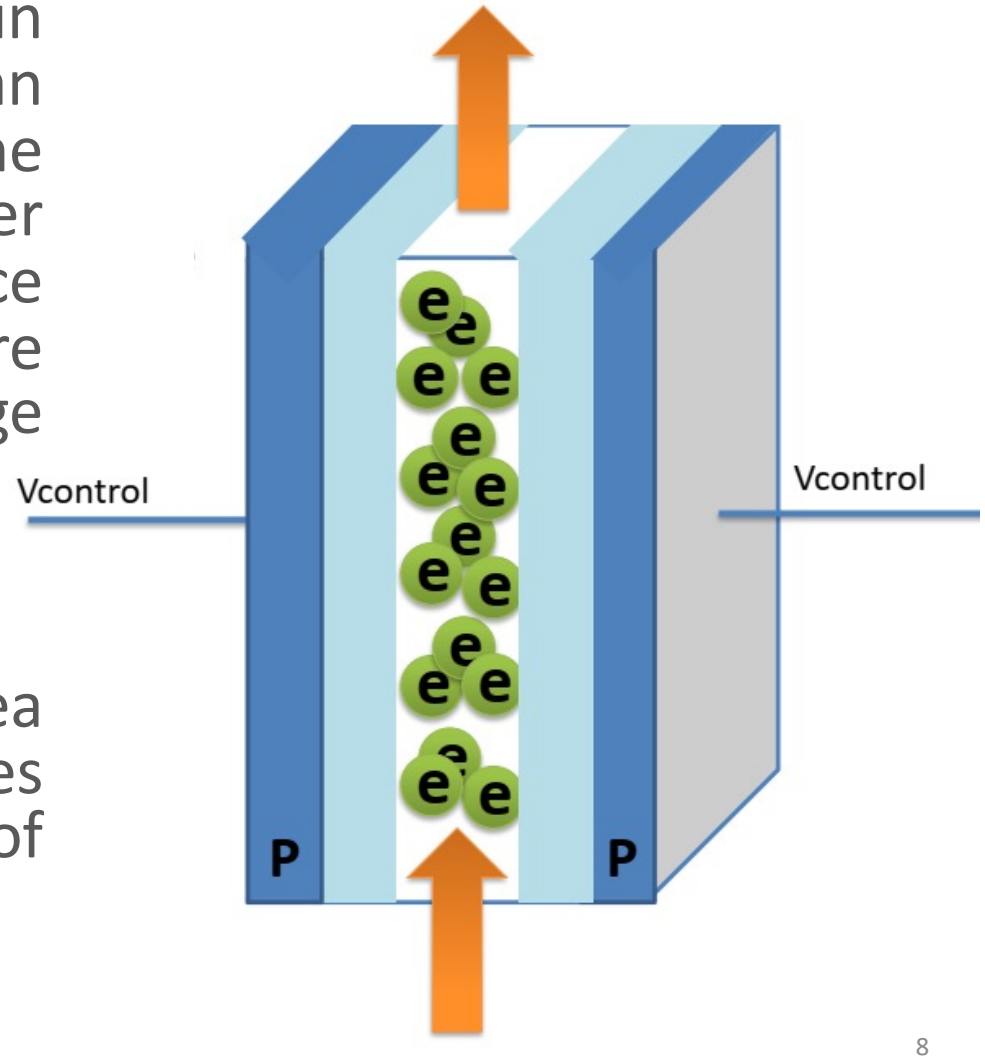


Junction Field Effect Transistors (JFETs)

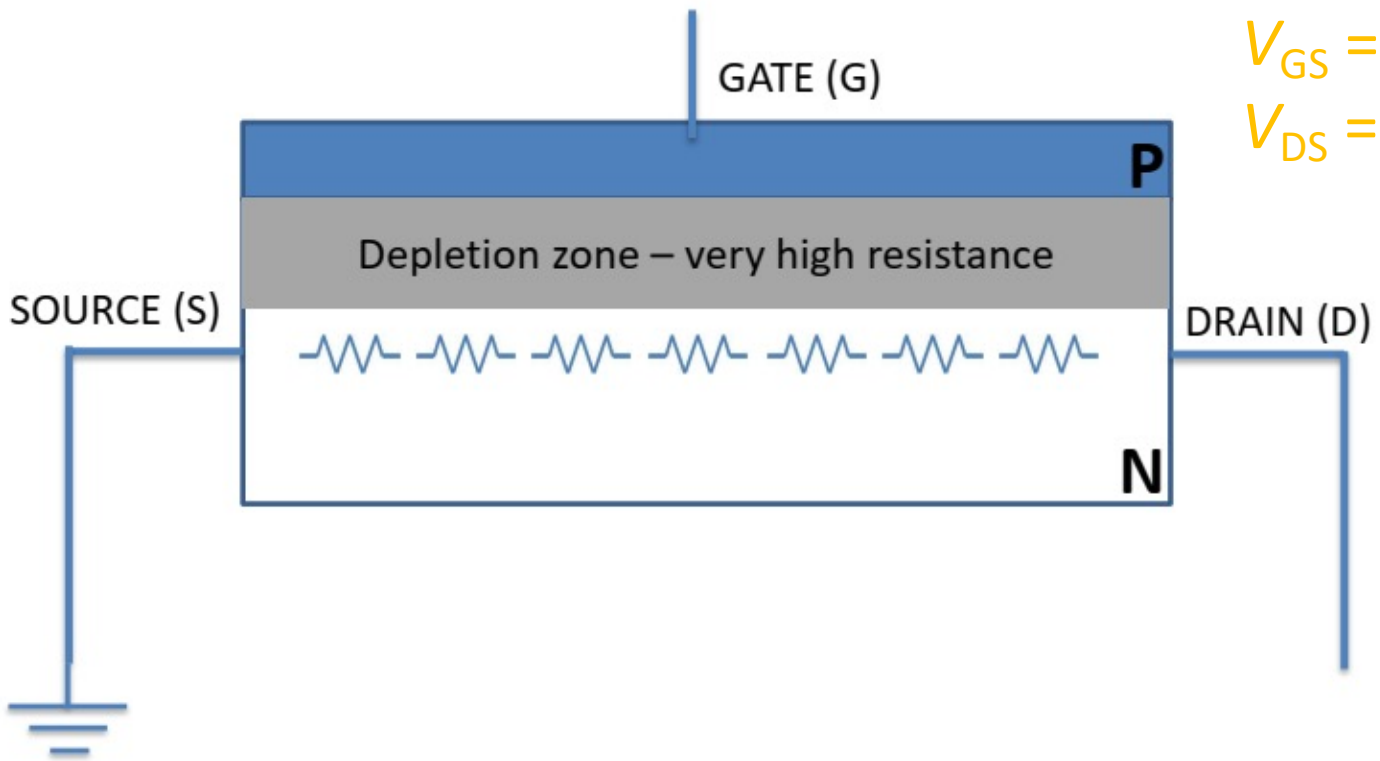
- We've squeezed the area in which the electrons can move by enhancing the depletion region. Higher voltage differences produce bigger fields and expel more and more negative charge carriers.

$$R = \rho \frac{L}{A}$$

- As the cross-sectional area reduces, the resistance goes up, limiting the amount of current that can flow.



$V_{DS} = 0$

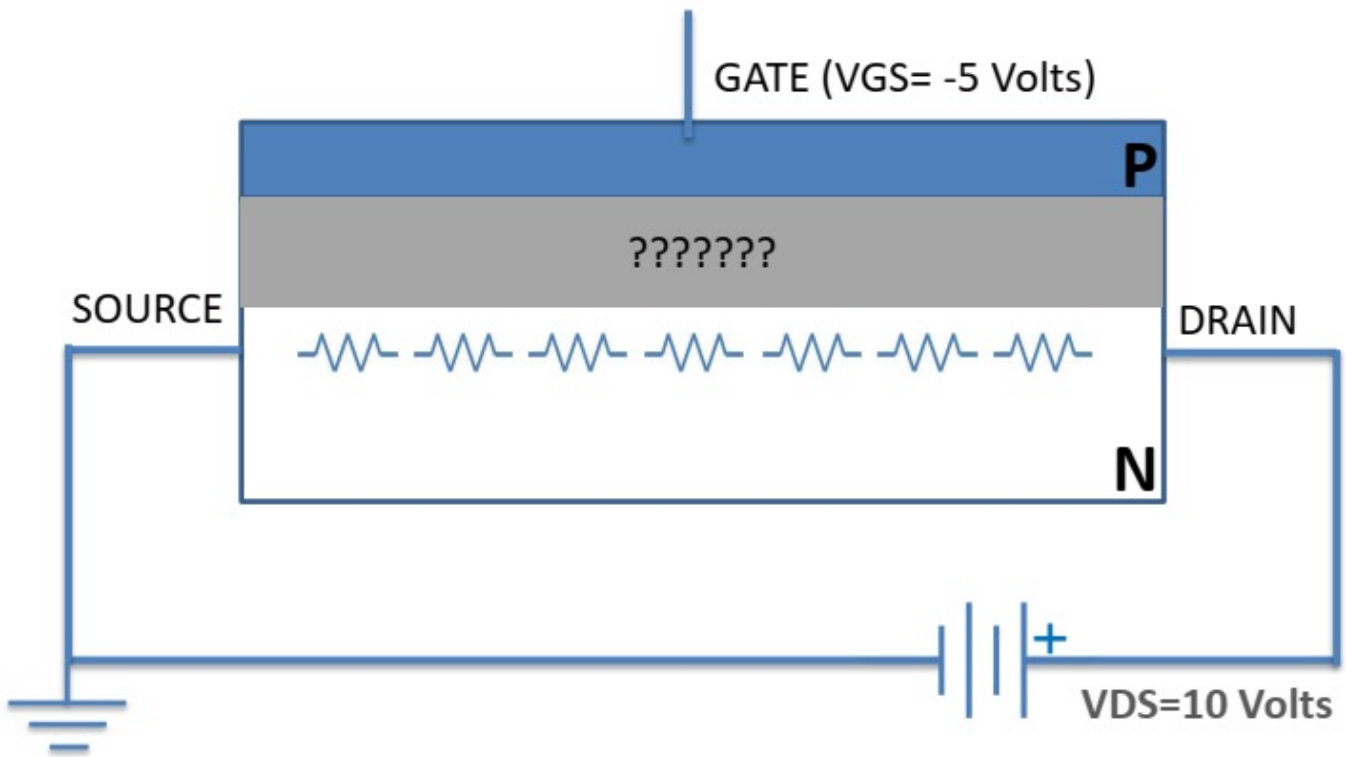


$V_{GS} = V_G - V_S$

$V_{DS} = V_D - V_S$

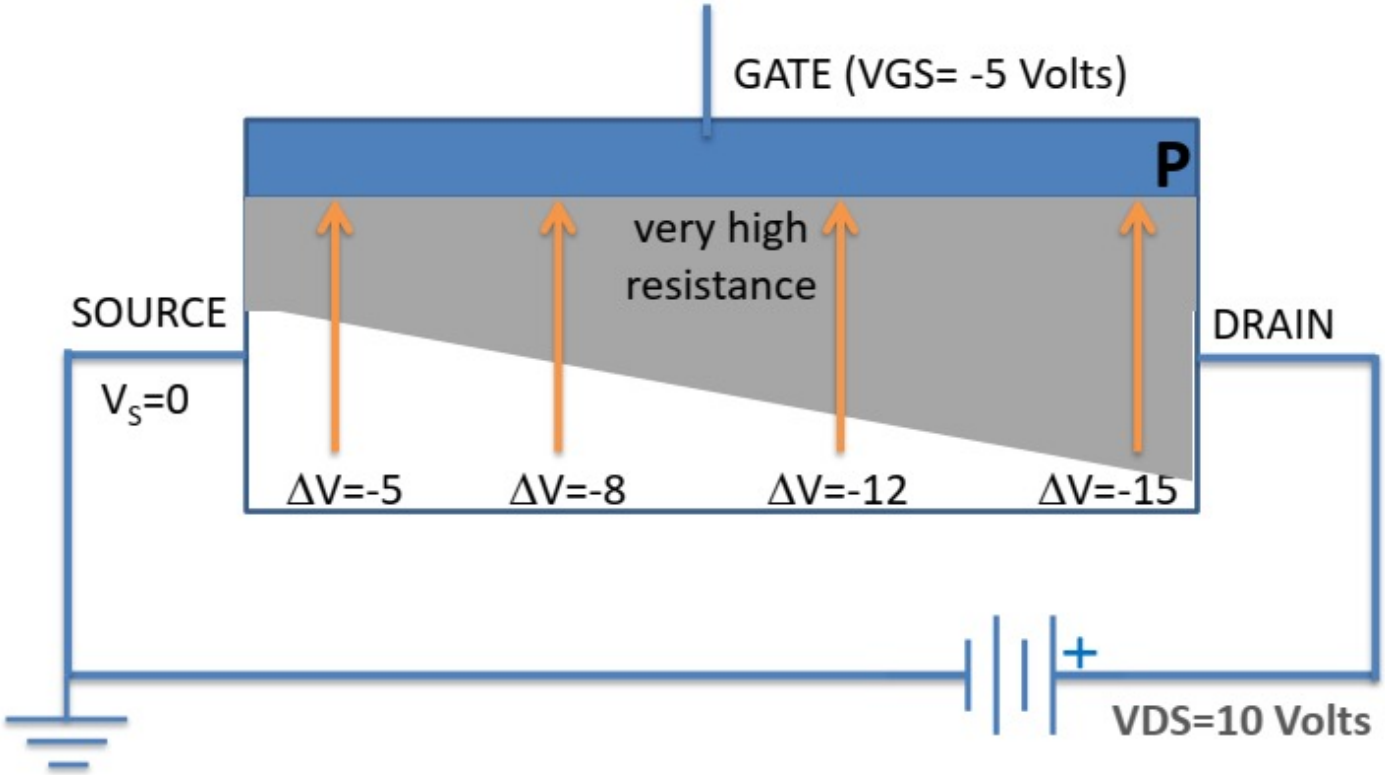
- The p-n junction must be reversed biased (say V_{GS} = -5 V).
- If no current flows, what is the electric potential difference across the p-n junction at different points INSIDE the device?

$V_{DS} > 0$



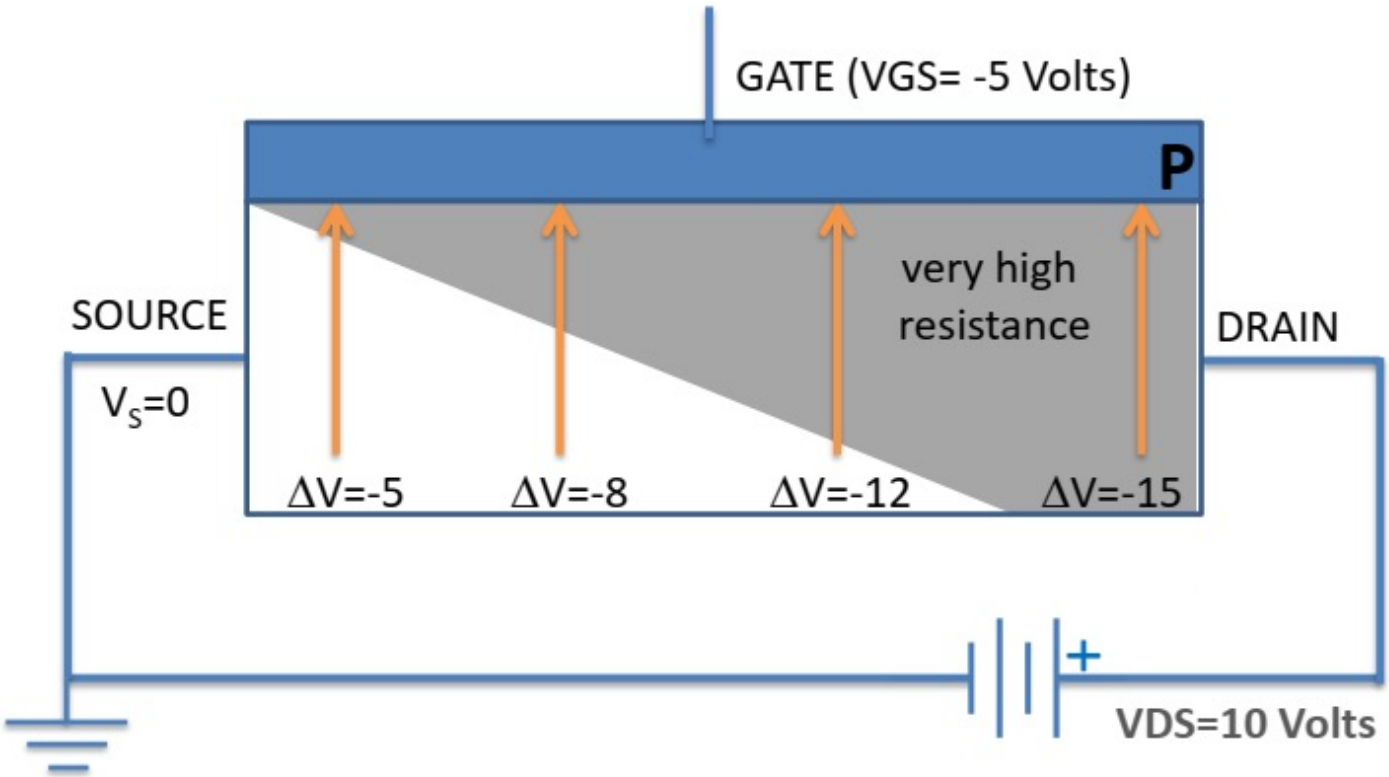
- Now consider if we apply a voltage between drain (D) and source (S), V_{DS} = +10 V, what will be the voltage difference across the p-n junction at different points in the device? What happens to the depletion region?

$V_{DS} > 0$



- Current flows in the channel, so the source terminal is at the ground potential (0 V), and the drain terminal is at $+V_{DS}$ = 10 V. There is a linear fall off in between. If we have too high a V_{DS}, what do you think will happen?

$V_{DS} >> 0$

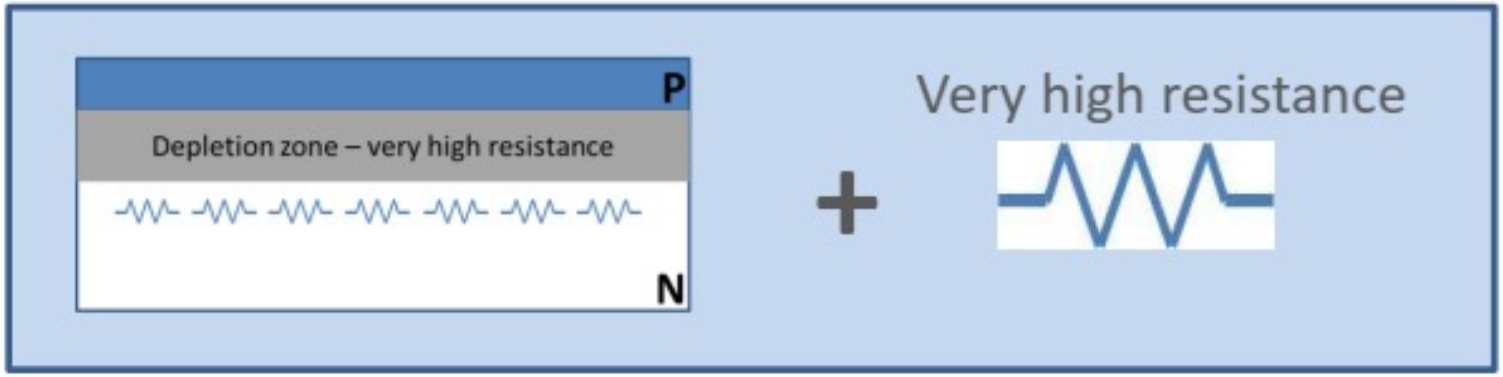


- This is called pinch-off. In theory no current can flow or at least a very high resistance is present. We can say that we've closed off the channel (in practice, the channel is never fully closed). So what do you think happens when no current can flow? ($V = IR$)

Pinch-Off Voltage

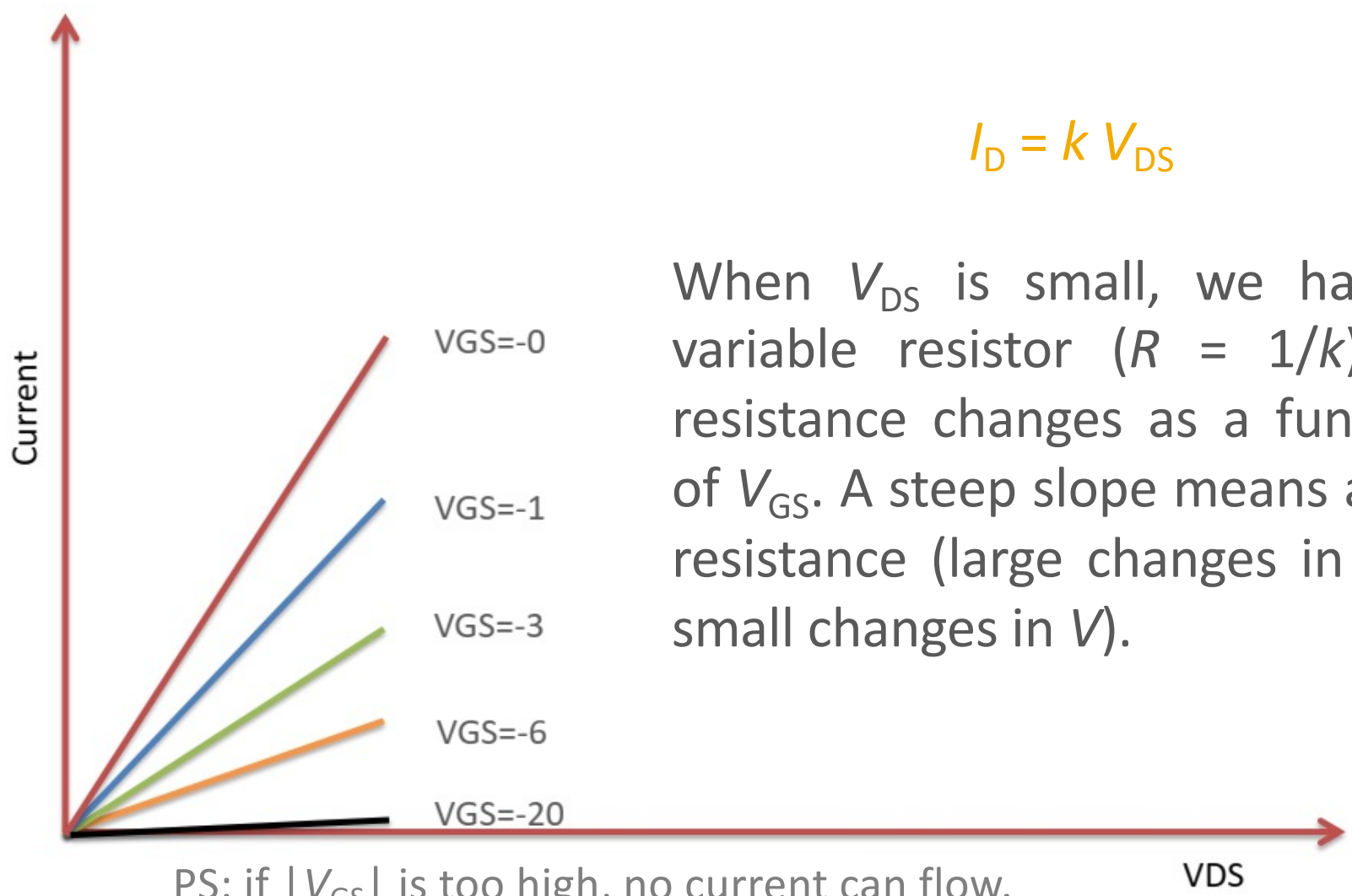
- In a JFET, when V_{GS} = 0, increasing V_{DS} leads to a narrowing of the conduction channel;
- Pinch-off refers to the condition when V_{DS} is increased to a level that the depletion region completely ”blocks” the conduction channel. The level of V_{DS} that establishes this condition is called the pinch-off voltage;
- The pinch-off voltage is also the level of V_{GS} required to completely shut down a JFET (I_D = 0 mA);
- V_P is defined as a negative voltage for n-channel JFETs and a positive voltage for p-channel JFETs.

What Happens at Pinch-Off?



- In reality, pinch-off does not fully shut down the conduction channel;
- So if we took the high resistance bit outside, we'd have a FET that COULD conduct and a high external resistance. In that sense some current will always flow, no matter what the applied voltage is;
- As the applied voltage gets bigger, the "external" resistance gets bigger, but not as fast as the voltage climbs, so we get some increase in current in saturation.

Drain Characteristics (I_D *vs.* V_{DS})



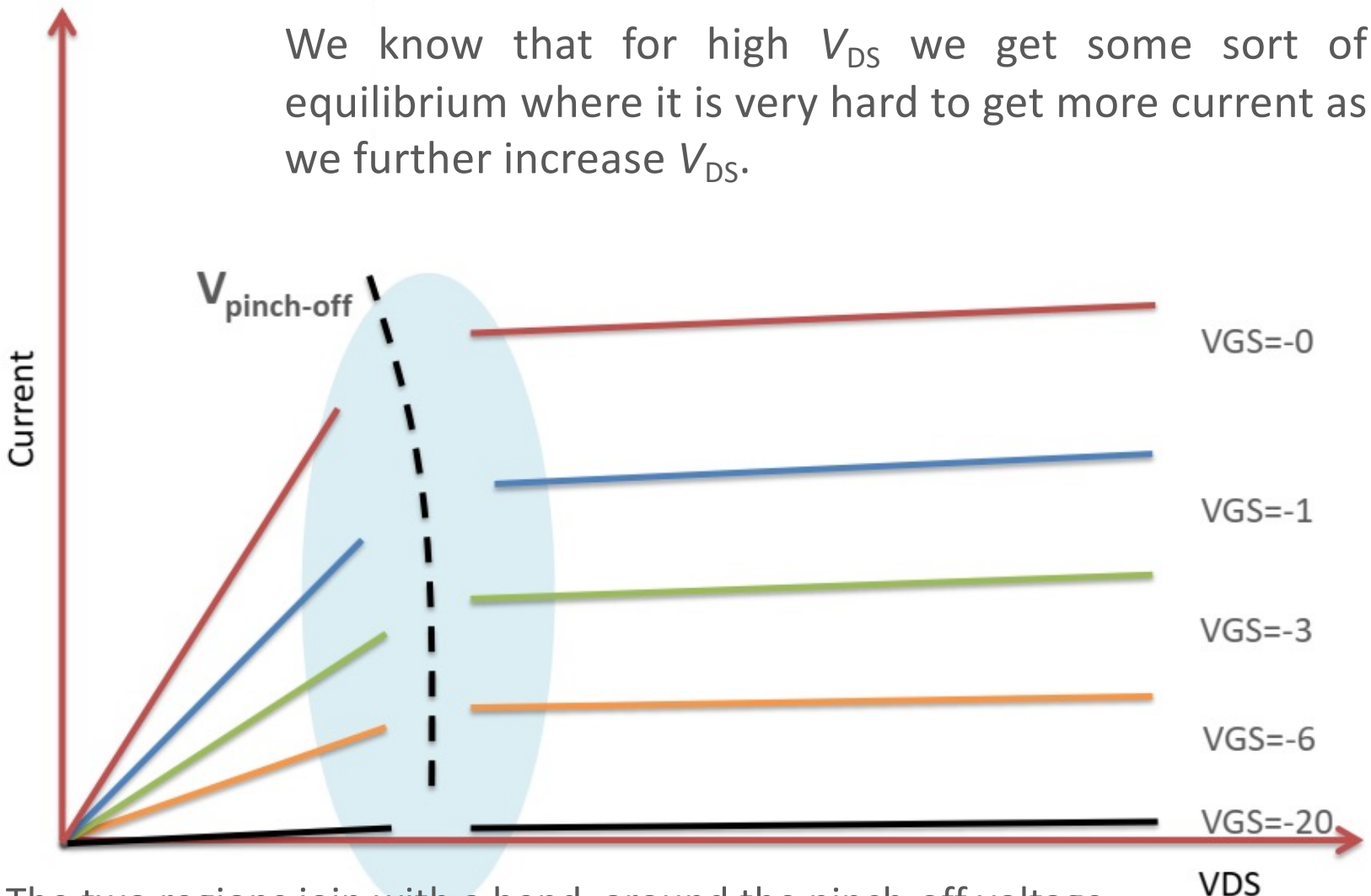
$I_D = k\,V_{DS}$

When V_{DS} is small, we have a variable resistor ($R = 1/k$). Its resistance changes as a function of V_{GS}. A steep slope means a low resistance (large changes in I for small changes in V).

PS: if $|V_{GS}|$ is too high, no current can flow.

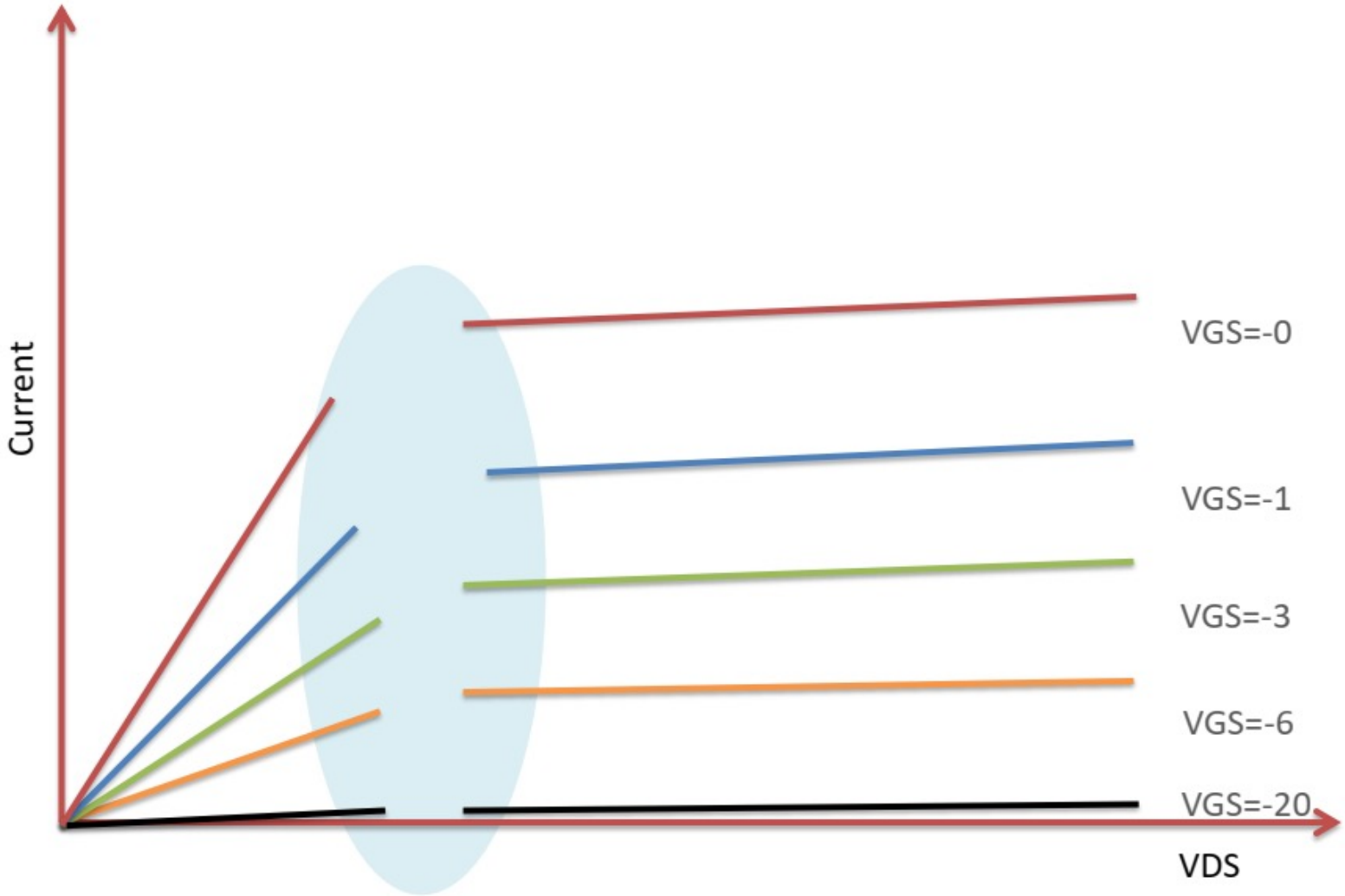
Drain Characteristics (I_D *vs.* V_{DS})

We know that for high V_{DS} we get some sort of equilibrium where it is very hard to get more current as we further increase V_{DS}.



The two regions join with a bend, around the pinch-off voltage.

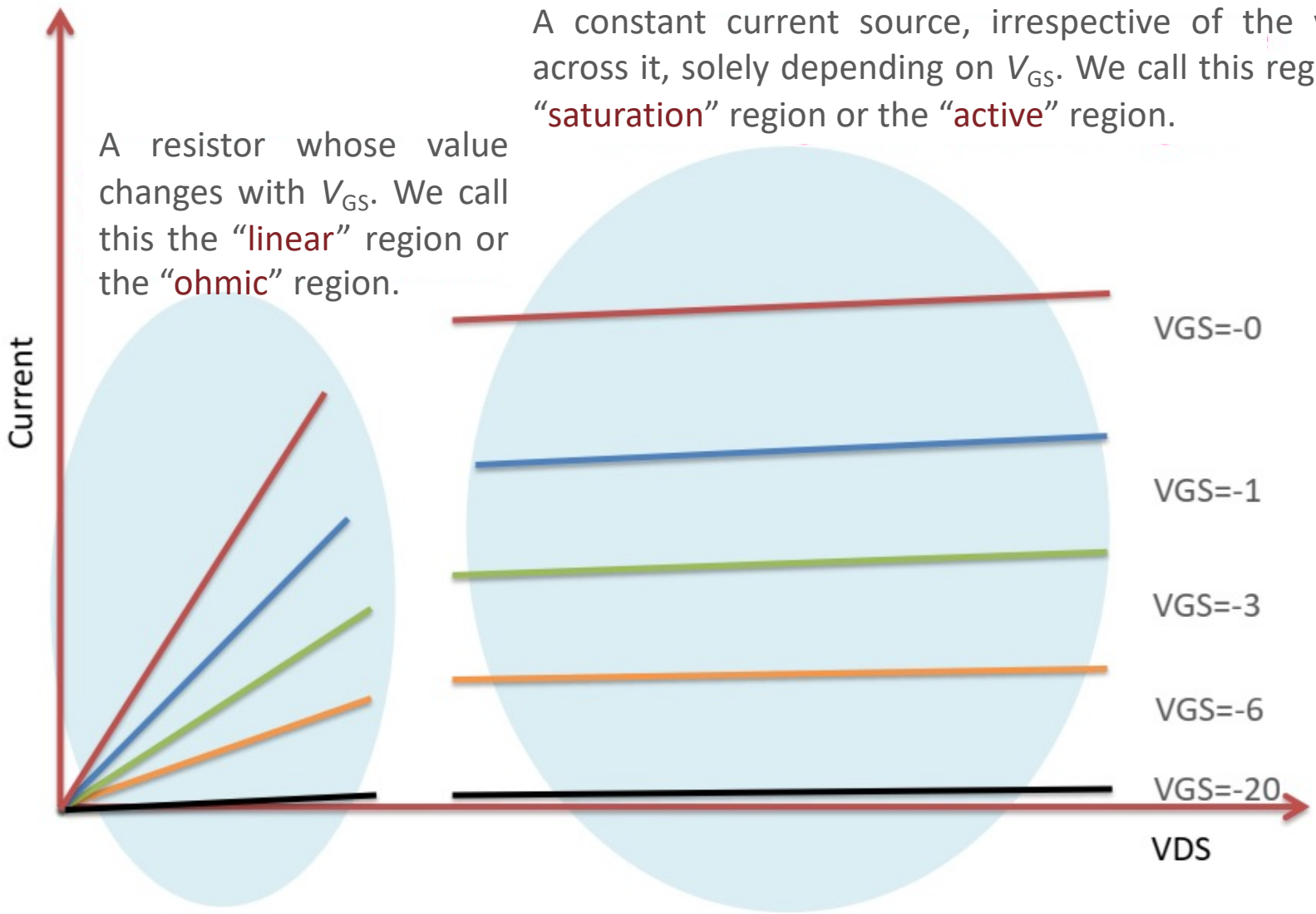
What Do You See in the Drain Characteristics?



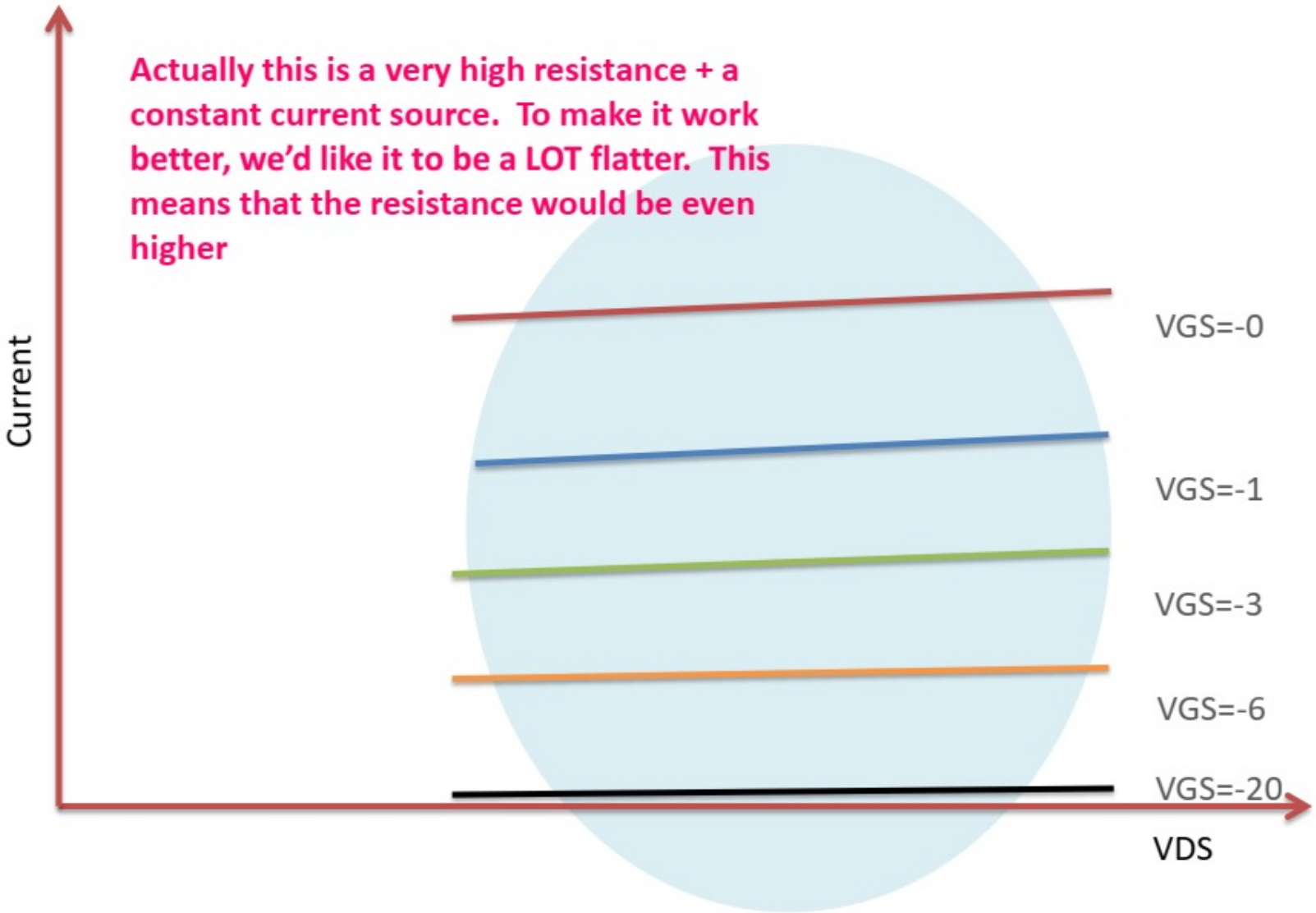
Two Regions in the Drain Characteristics

A constant current source, irrespective of the voltage across it, solely depending on V_{GS}. We call this region the "saturation" region or the "active" region.

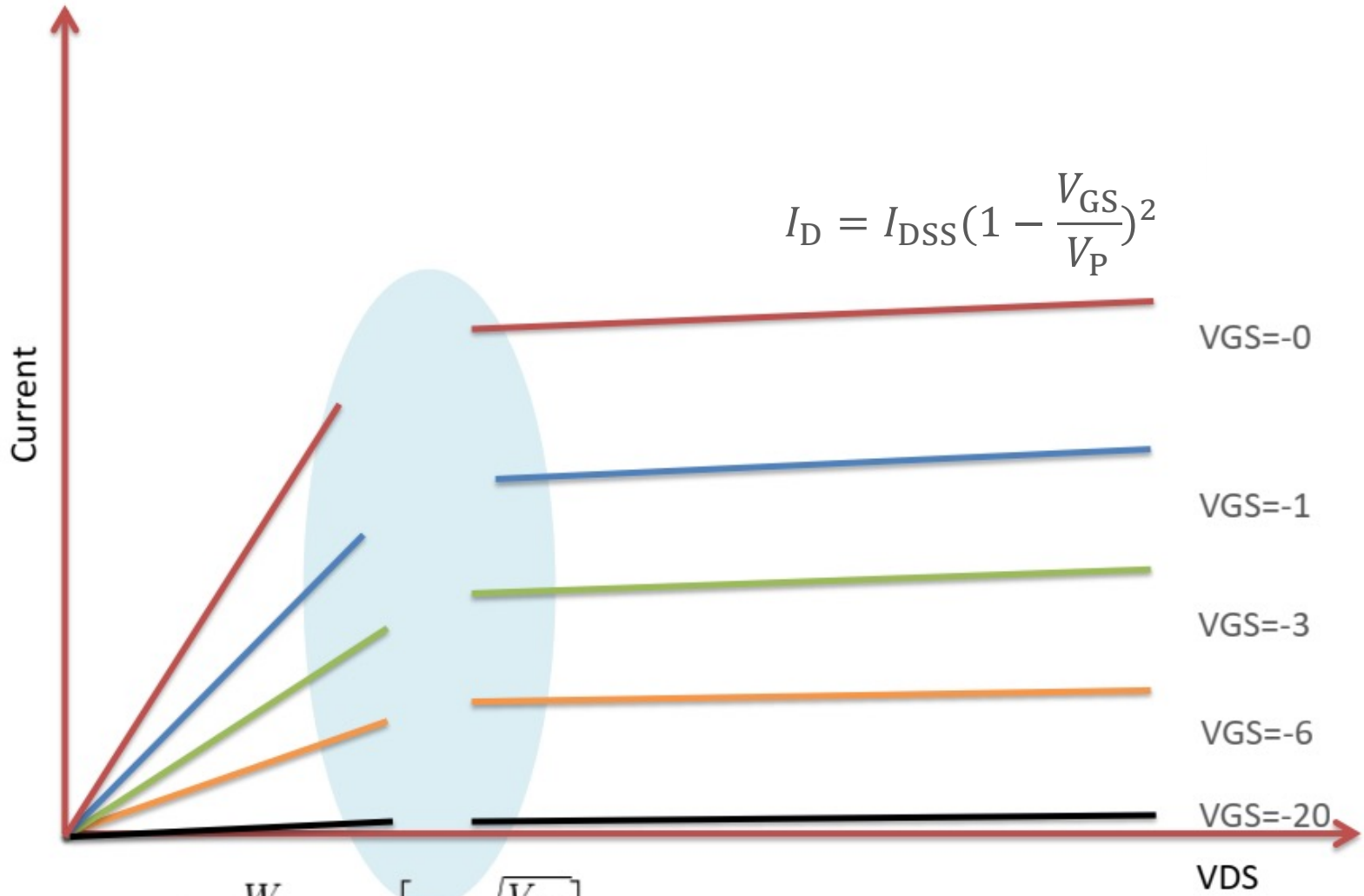
A resistor whose value changes with V_{GS}. We call this the "linear" region or the "ohmic" region.



Saturation Region



Drain Characteristics Equations



$$I_D = I_{DSS}(1 - \frac{V_{GS}}{V_P})^2$$

$$I_D = (2a)\frac{W}{L}qN_d\mu_n\left[1 - \sqrt{\frac{V_{GS}}{V_P}}\right]V_{DS}$$

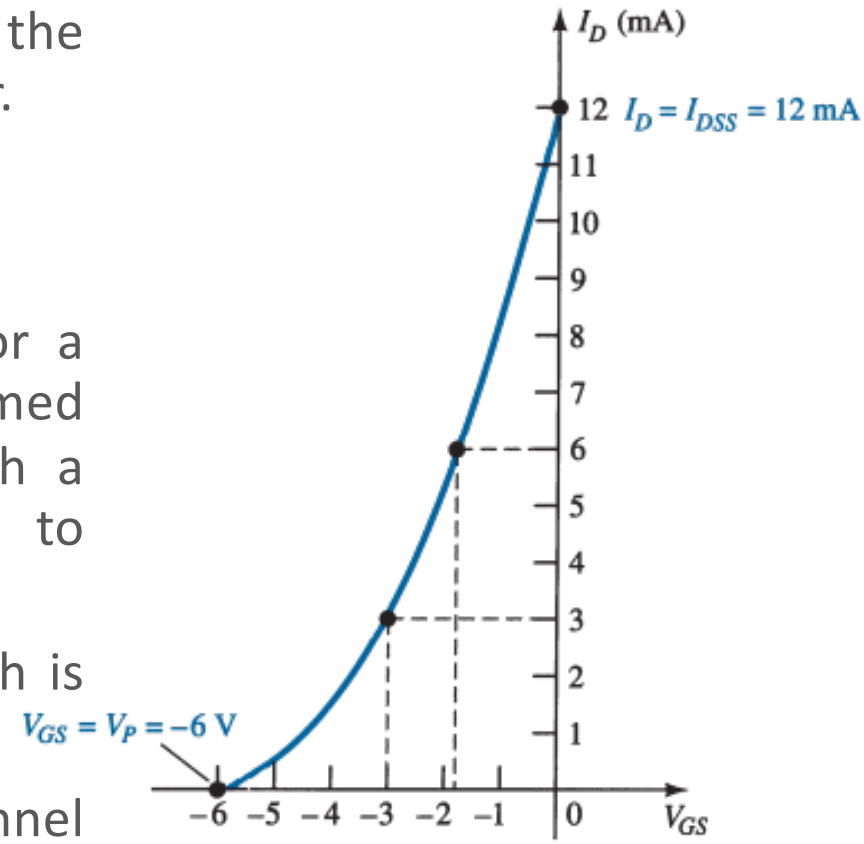
Note: If V_{GS} = const., $I_D = G \times V_{DS}$.

Shockley's (Transistor) Equation

- In the saturation region, the relation between the output current I_D and the input control parameter, V_{GS} is nonlinear.

$$I_D = I_{DSS}(1 - \frac{V_{GS}}{V_P})^2$$

- I_{DSS} is the maximum drain current for a JFET that occurs at V_{GS} = 0. I_{DSS} is named after the "drain-to-source current with a short-circuit connection from gate to source";
- V_{GS} is the gate-to-source voltage, which is the control parameter;
- V_P is the pinch-off voltage. For an n-channel JFET, both V_{GS} and V_P < 0.

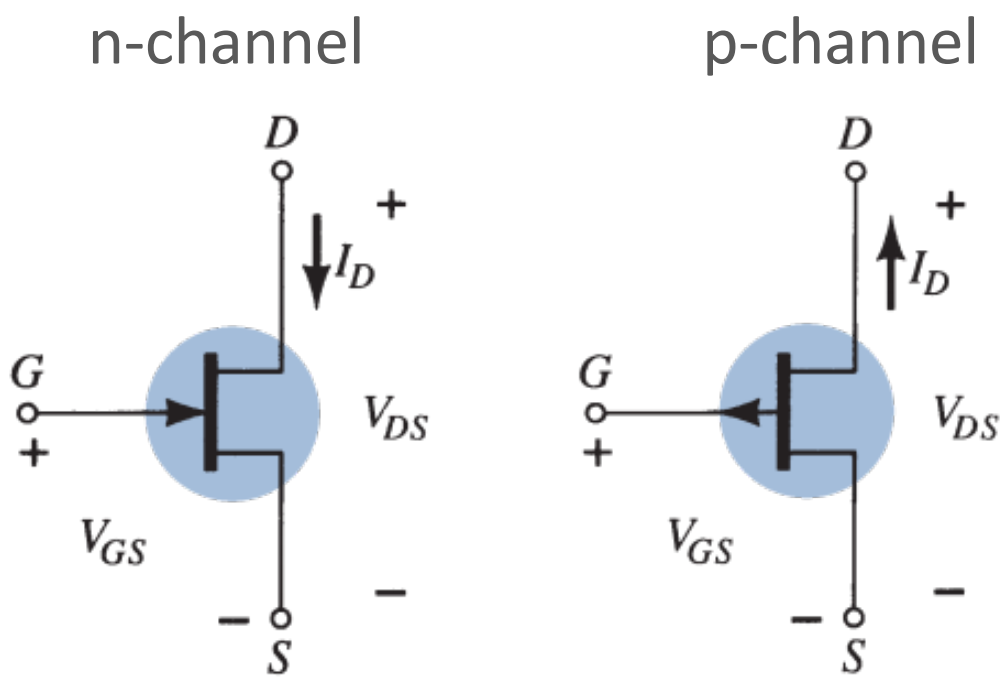


Additional Remarks: JFETs

- JFETs come in n type as well as p type.
- In an n-type JFET, the source is connected to the ground or the low voltage side. In a p-type JFET, the source is connected to the high voltage side.
- The source is named because that is where the electrons come from in an n-type material or where the holes come from in a p-type material. The source is where the majority charge carriers come out and the drain is where they go after leaving the device.
- A good visual explanation of a JFET can be found here (Flash required).

  http://www-g.eng.cam.ac.uk/mmg/teaching/linearcircuits/jfet.html

Circuit Symbols for JFETs



- Pay attention to the directions of the arrows in the symbols. The direction of the arrow represents the forward bias direction of a p-n junction in the FET.

Do We Use JFETs a Lot?

- Well no, it depends a little bit on whom you ask. But in general, we don't, at least not as much as some would think.

- Ideally we'd like the flat part in the saturation region to be very flat. JFETs can be good but they don't get it flat enough… increasing V_{DS} pushes more current through even when a JFET is in saturation. They also never turn off (block all current flows) properly.

- JFETs are mainly used in some niche applications today where high gain has to be combined with low noise, e.g., preamplifiers for interfacing low-level signals.

- However, the most important reason is that MOSFETs are a LOT better, they are especially flat in the saturation region, and they can stop current flowing very well (fully off).

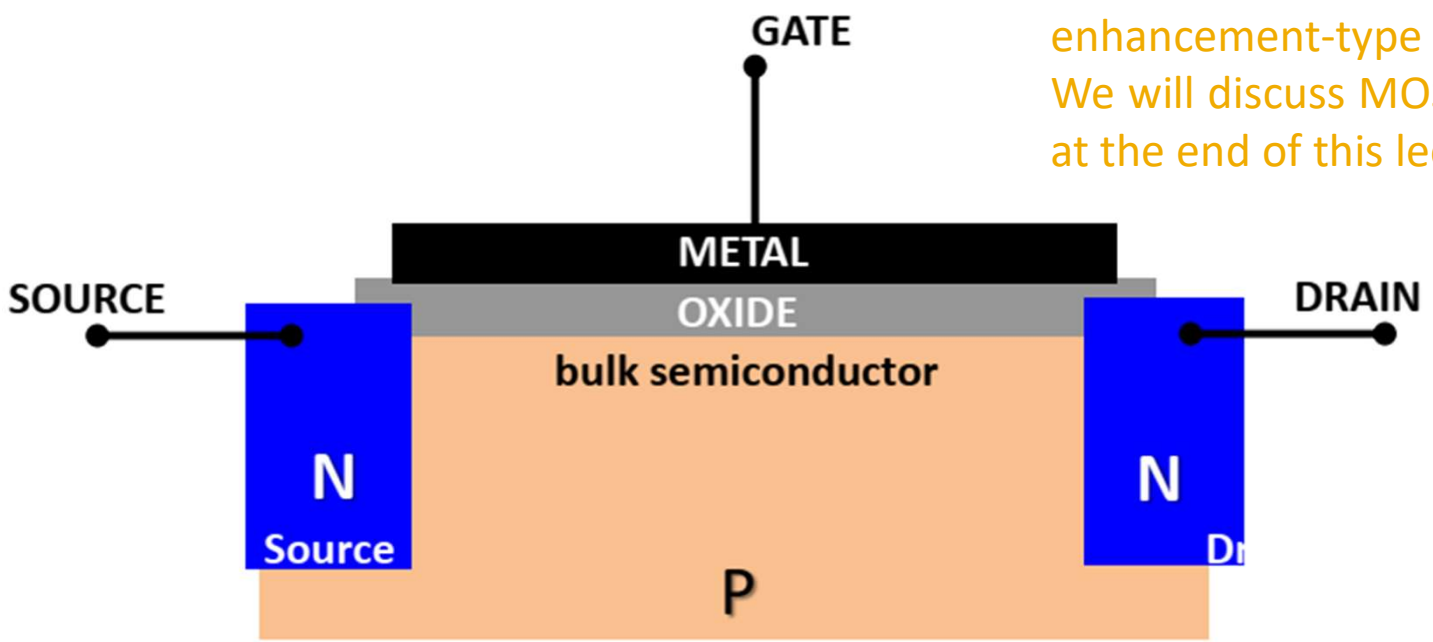
EE204FZ
Lecture 4
Metal-Oxide-Semiconductor Field Effect Transistors (MOSFETs)

Zhu DIAO

Email: zhu.diao@mu.ie

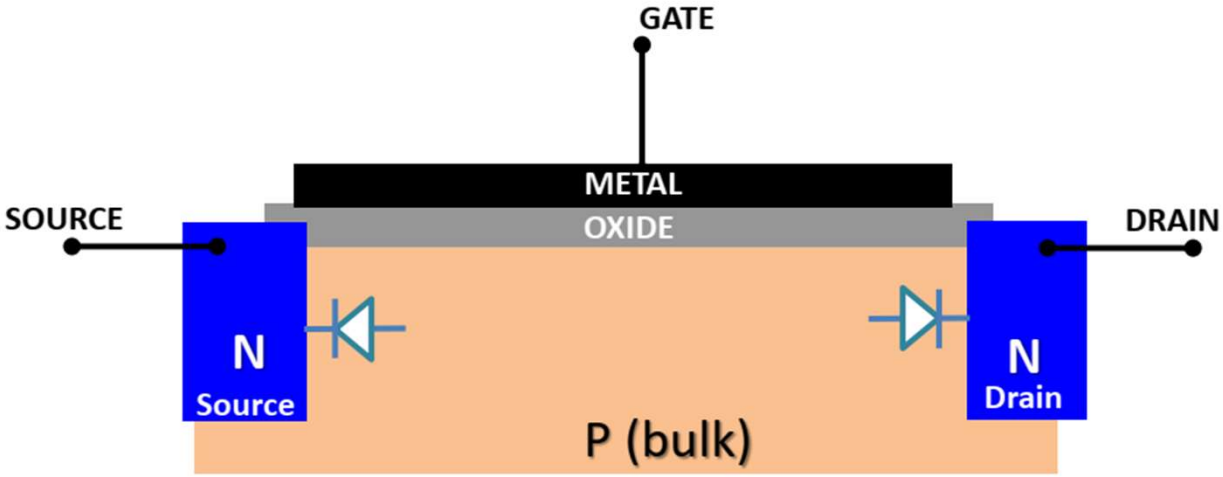
Metal-Oxide-Semiconductor FETs

This is in fact an n-channel enhancement-type MOSFET. We will discuss MOSFET types at the end of this lecture.



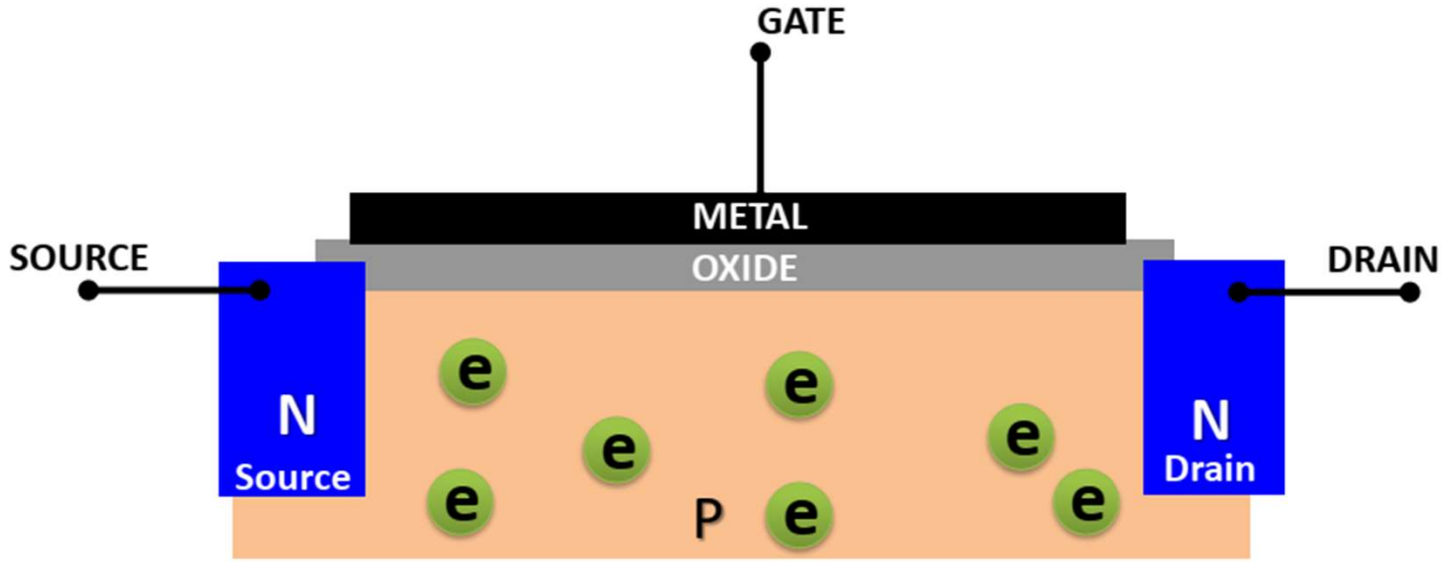
http://www-g.eng.cam.ac.uk/mmg/teaching/linearcircuits/mosfet.html

Metal-Oxide-Semiconductor FETs



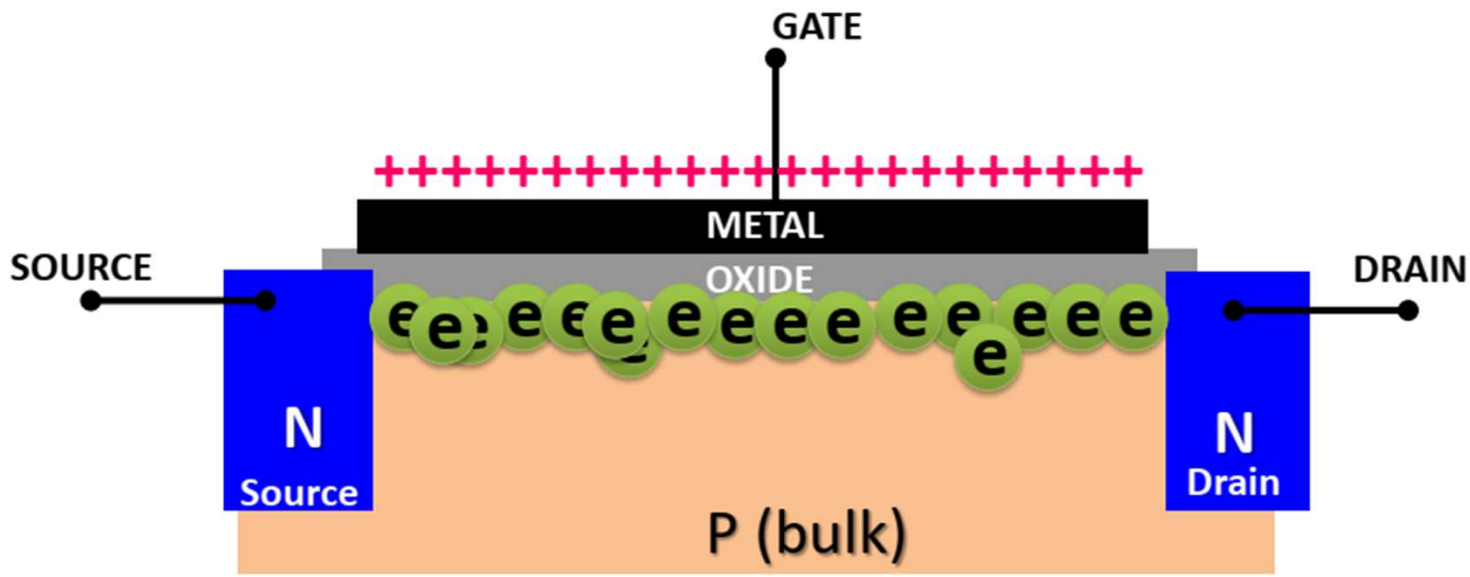
- Now in vague memory of solid-state electronics, a p-n junction is made when p-type and n-type semiconductors come together, and current can't flow through a reverse biased p-n junction (It is a diode). In this scenario, current cannot NORMALLY flow from source to drain.
- Also note, the device is symmetric, so source and drain are swappable until you apply a voltage.

Metal-Oxide-Semiconductor FETs



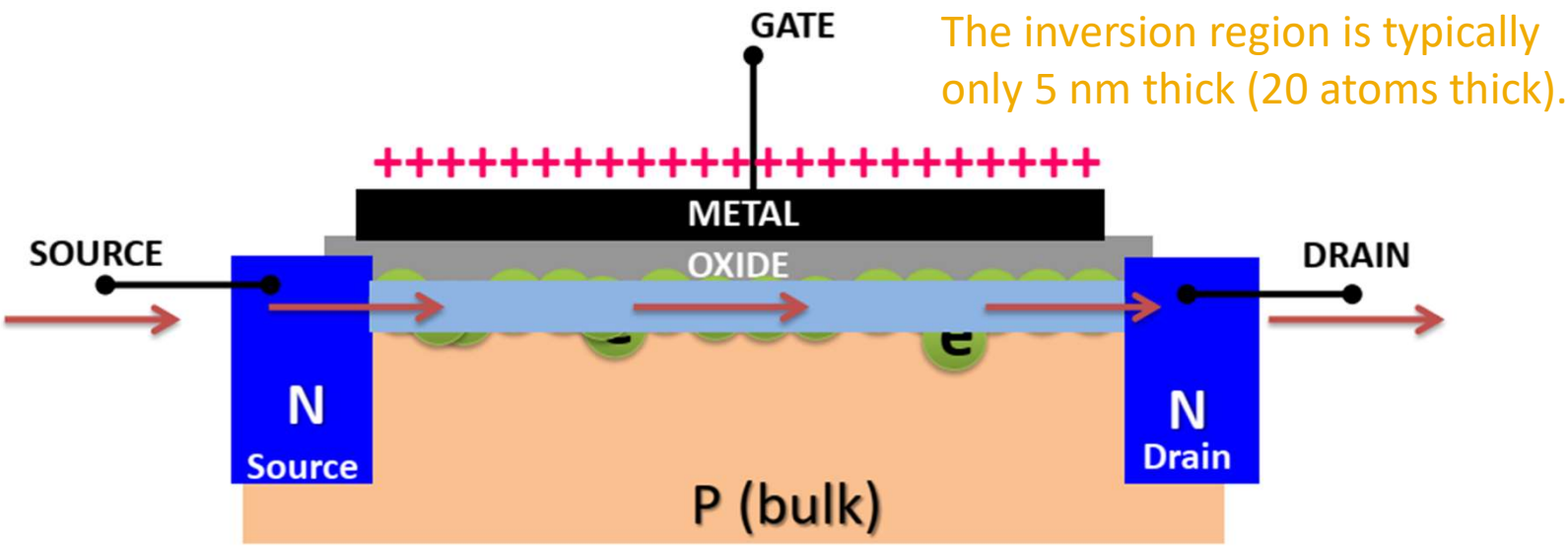
- An n-type material is made when dopants make an excess of electrons. A p-type material is made when dopants make a shortage of electrons. But there are, however, always some electrons in the material.
- Here's a question, how can we make a p-type material n-type?

Metal-Oxide-Semiconductor FETs



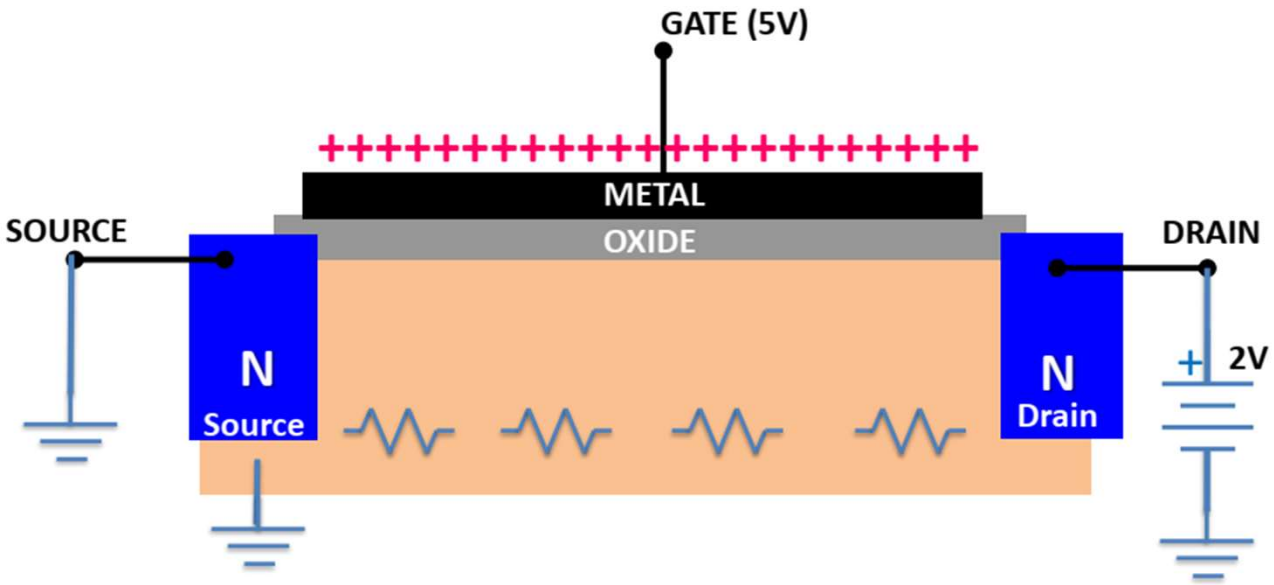
- So all the electrons in the p-type material get attracted to the POSITIVE gate, but the oxide is a good insulator and stops them from getting over to the gate.
- This makes the rest of the material more p-type.
- If enough electrons get attracted, what happens to the p-n junction???

Metal-Oxide-Semiconductor FETs



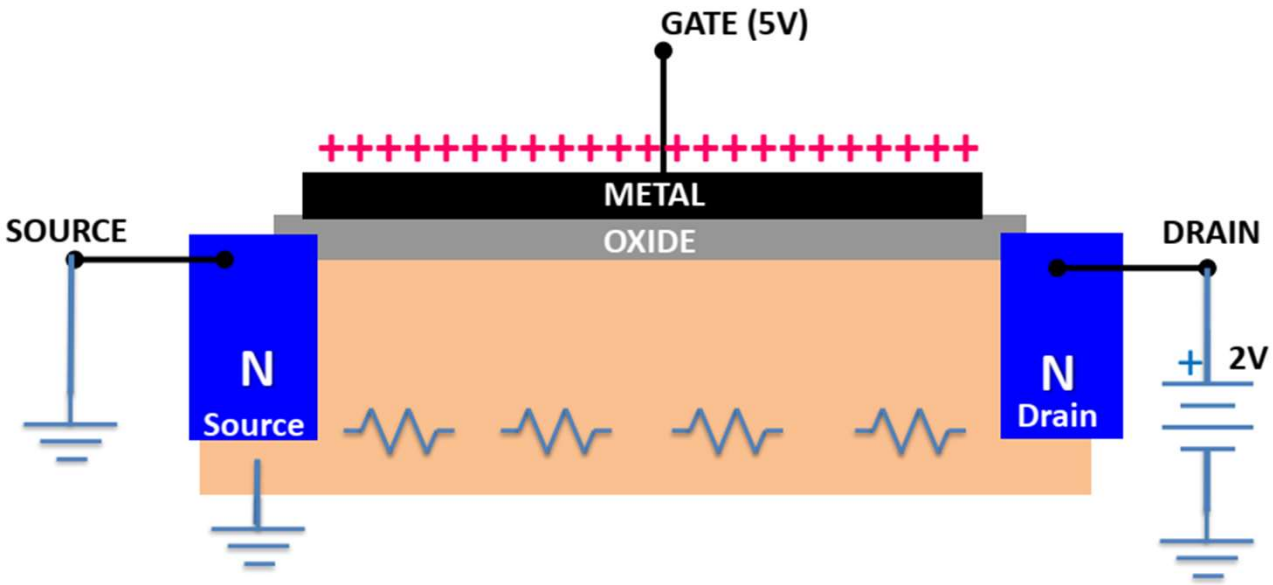
- We call the newly "generated" n-type region the inversion region as we swapped region types. It only lasts as long as the gate is charged.
- There is a minimum gate-bulk voltage difference that needs to be achieved for an inversion region to form. We call this V_T (the threshold voltage).
- So now we have n → n → n, there is no p-n junction, current can now flow.

Metal-Oxide-Semiconductor FETs



- So we've connected the bulk to GROUND.
- We've connected the source to GROUND.
- We've connected the gate to +5 V.
- We've connected the drain to +2 V.
- What happens?

Metal-Oxide-Semiconductor FETs



- So we've connected the bulk to GROUND.
- We've connected the source to GROUND.
- We've connected the gate to +5 V.

An inversion region will form as the electrons are attracted to the gate.

- We've connected the drain to +2 V.

Current will flow.

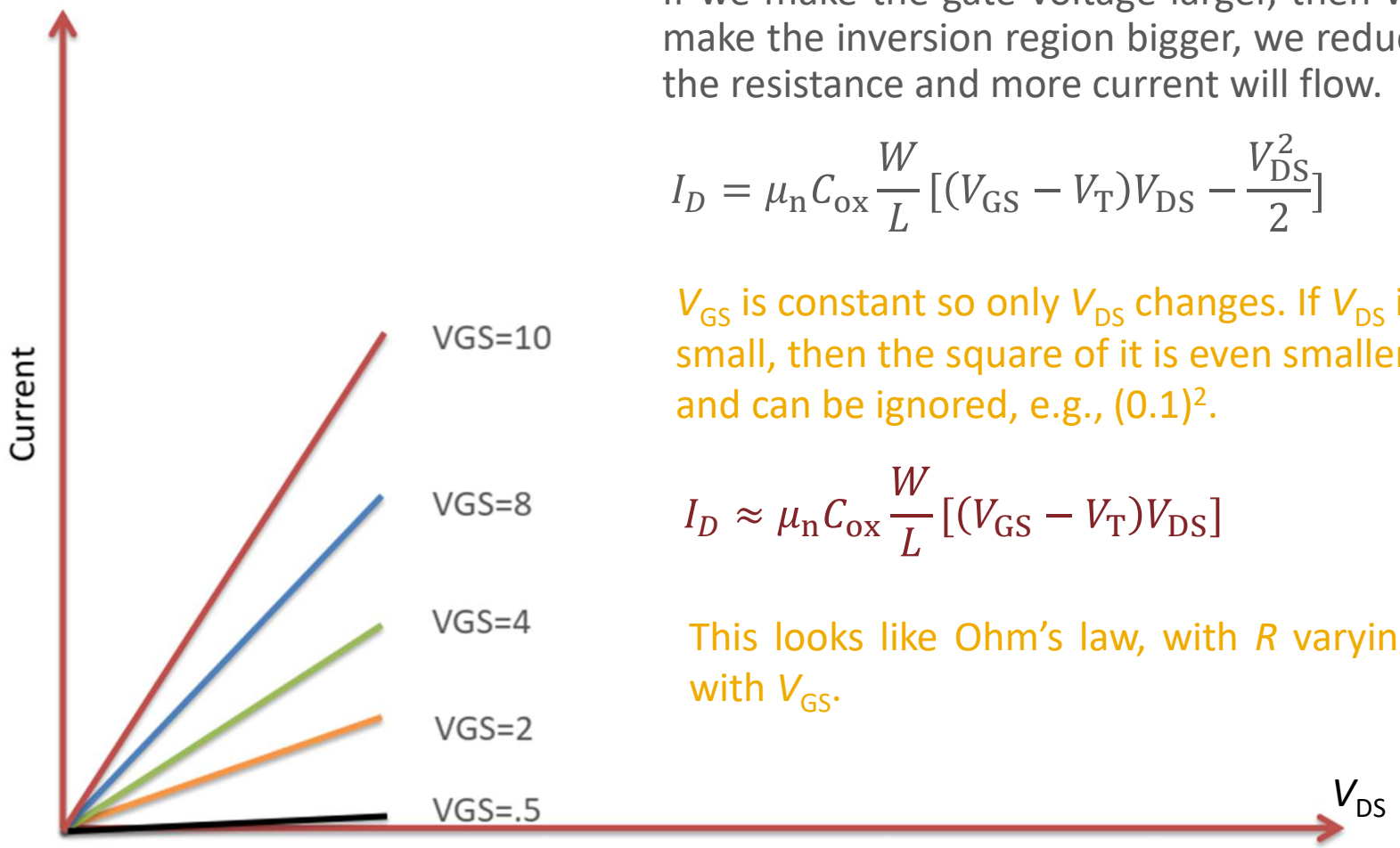
The MOSFET Equation

$$I_D = \mu_{\mathrm{n}} C_{\mathrm{ox}} \frac{W}{L} [(V_{\mathrm{GS}} - V_{\mathrm{T}}) V_{\mathrm{DS}} - \frac{V_{\mathrm{DS}}^2}{2}]$$

- μ_{n}: Mobility (of electrons) at the channel surface;
- C_{ox}: Gate oxide capacitance in F/m^2;
- L and W: Channel length and width, respectively;
- V_{T}: Threshold voltage;
- V_{GS} and V_{DS}: $V_{\mathrm{GS}} = V_{\mathrm{G}} - V_{\mathrm{S}}$ and $V_{\mathrm{DS}} = V_{\mathrm{D}} - V_{\mathrm{S}}$.

When you see a complex equation, try to spot the constants that depend on the construction and physical make up of the device. Even these 'complex' equations are simplifications. Full models of transistors can be 100-page long!

Plotting the MOSFET Equation



If we make the gate voltage larger, then we make the inversion region bigger, we reduce the resistance and more current will flow.

$$I_D = \mu_n C_{ox} \frac{W}{L} [(V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2}]$$

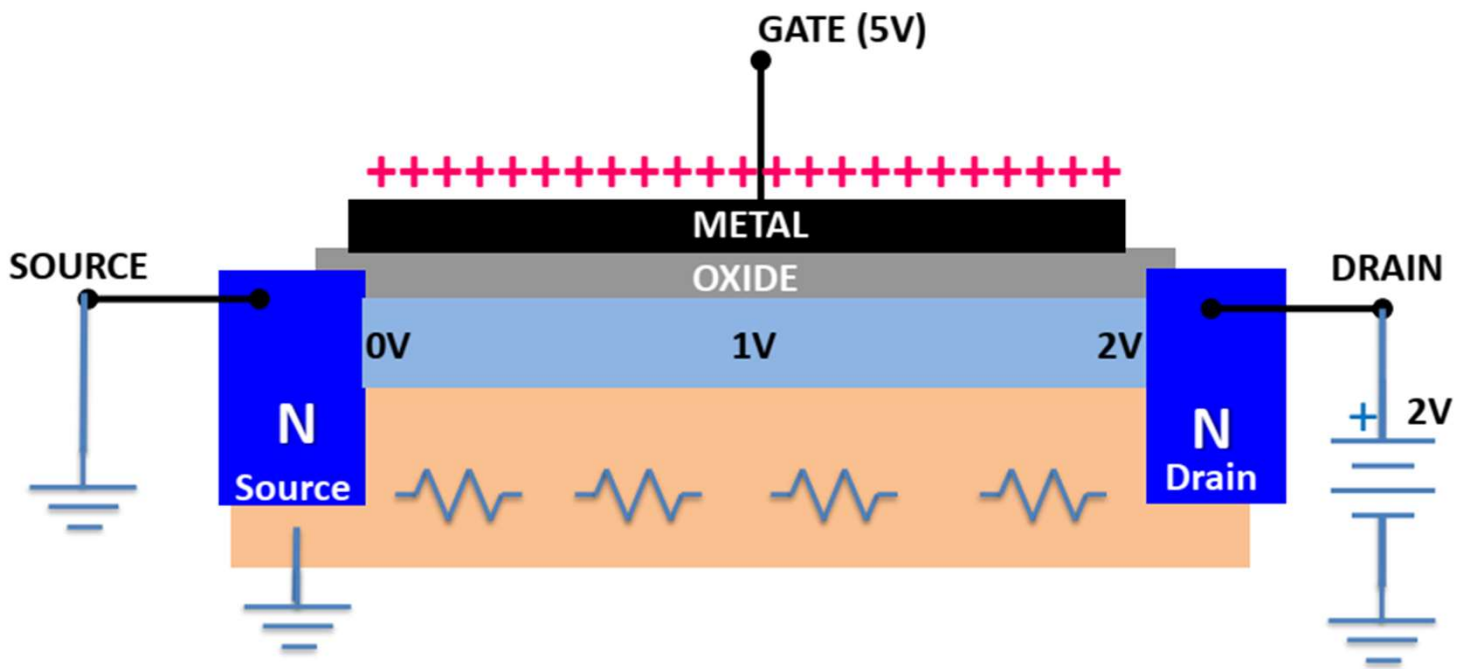
V_{GS} is constant so only V_{DS} changes. If V_{DS} is small, then the square of it is even smaller and can be ignored, e.g., $(0.1)^2$.

$$I_D \approx \mu_n C_{ox} \frac{W}{L} [(V_{GS} - V_T)V_{DS}]$$

This looks like Ohm's law, with R varying with V_{GS}.

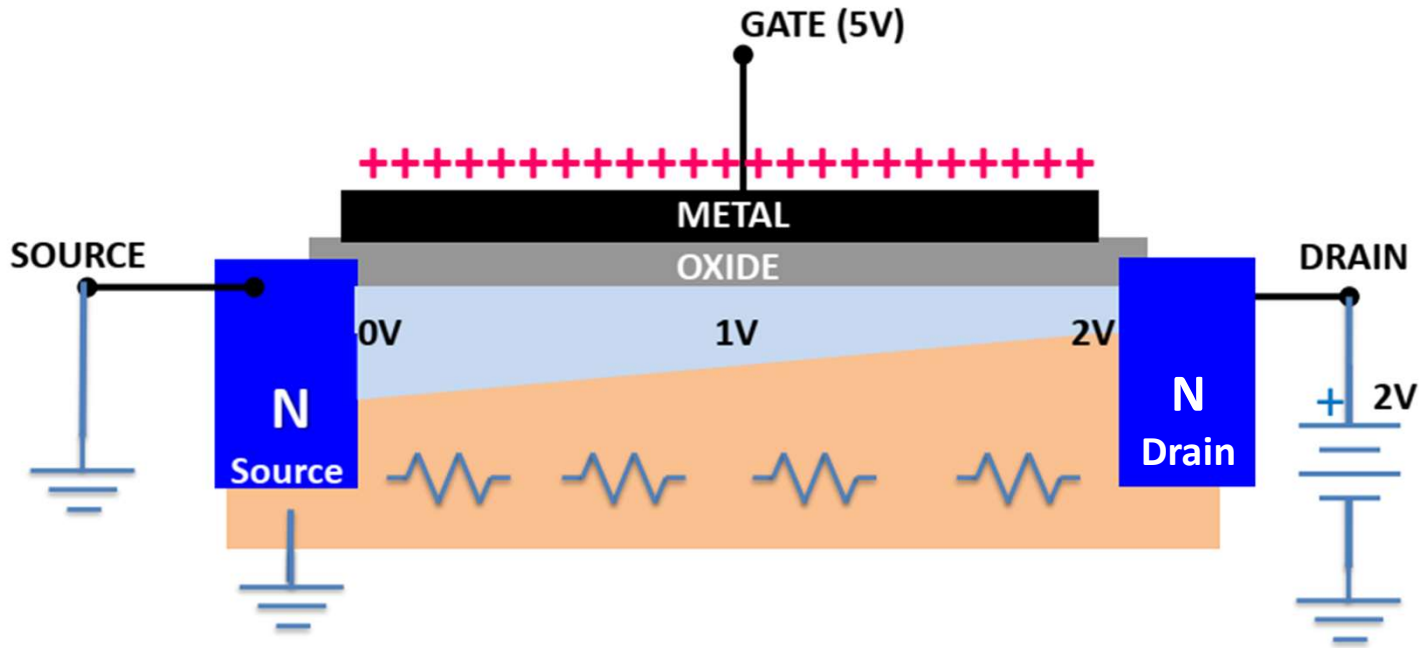
PS: V_{GS} needs to be bigger than V_T to work (typically 0.3 Volts).

The Effect of V_{DS}



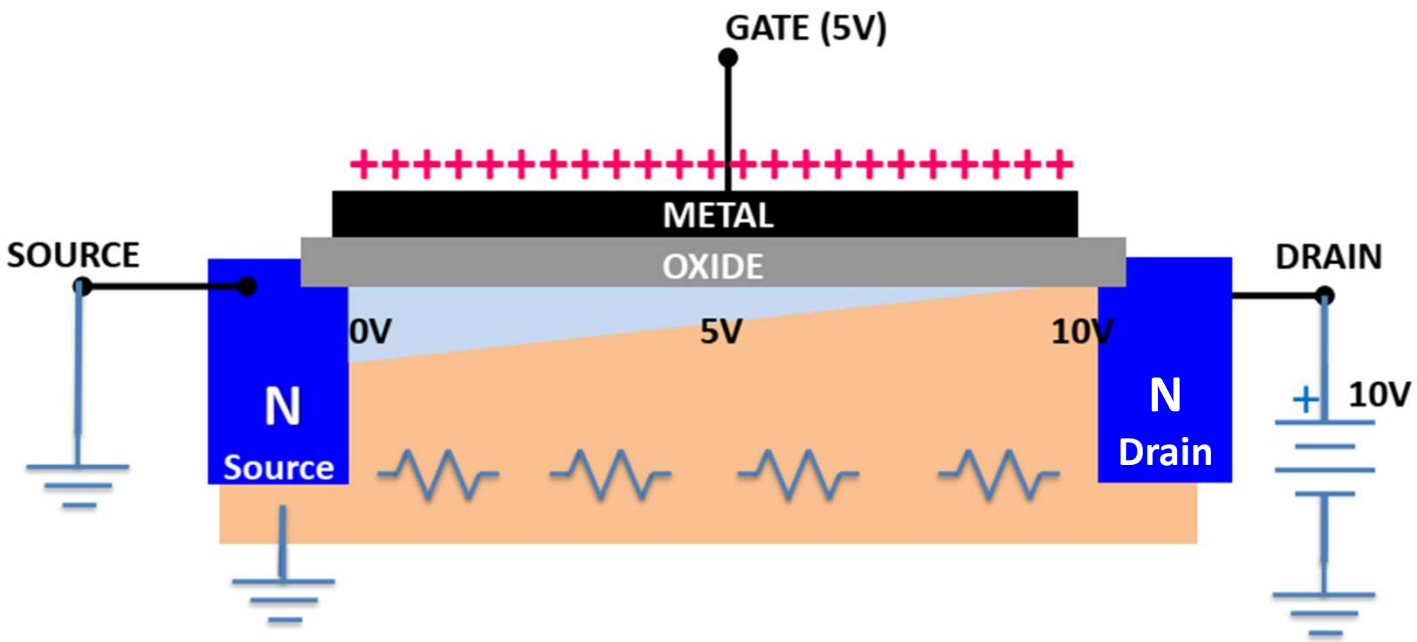
- Internally it's just a resistor, so the voltage at the drain starts at 2 V and drops to zero at the source side. If the resistance is the same everywhere, it will be a linear drop off.
- What does this do to the inversion region?

The Effect of V_{DS}



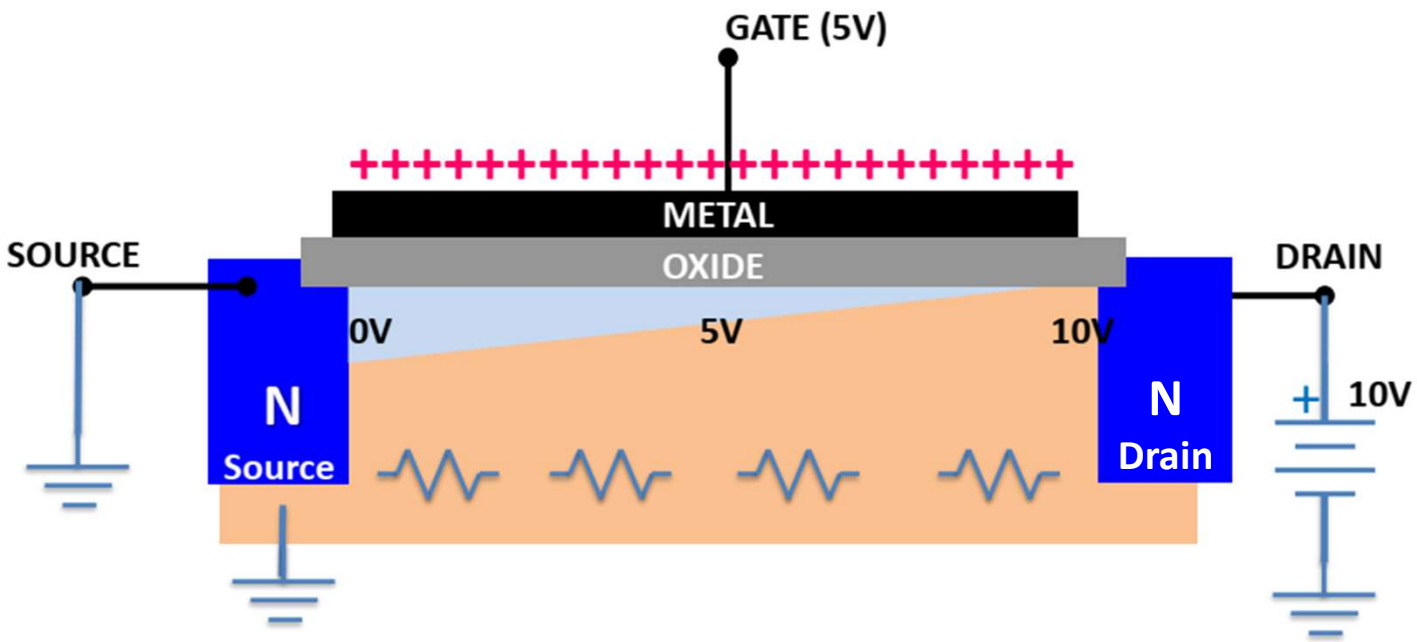
- Inversion regions are bigger if the voltage difference is bigger, so we get a triangular-shaped region.
- The bigger the region, the less resistive it is.

Now Increase V_{DS}



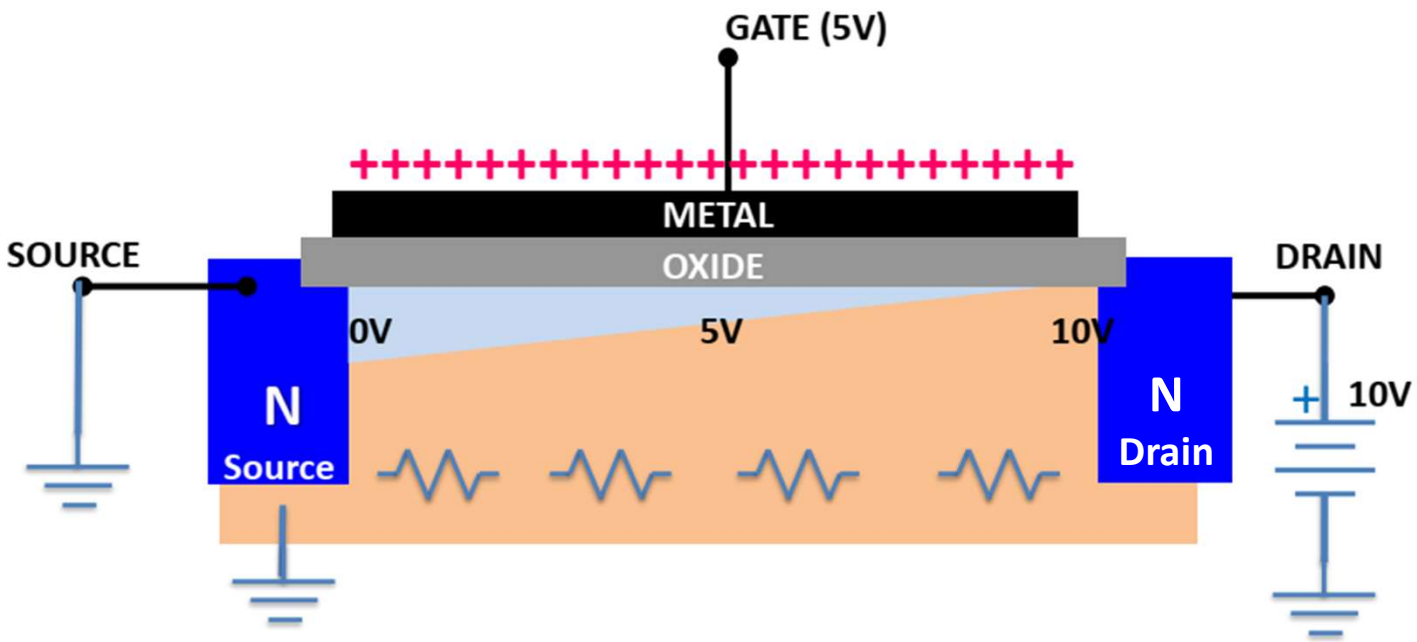
- We have a gap. If the voltage of the bulk material is higher than the gate, then electrons won't want to come to the top. No inversion region.

Channel Pinch-Off



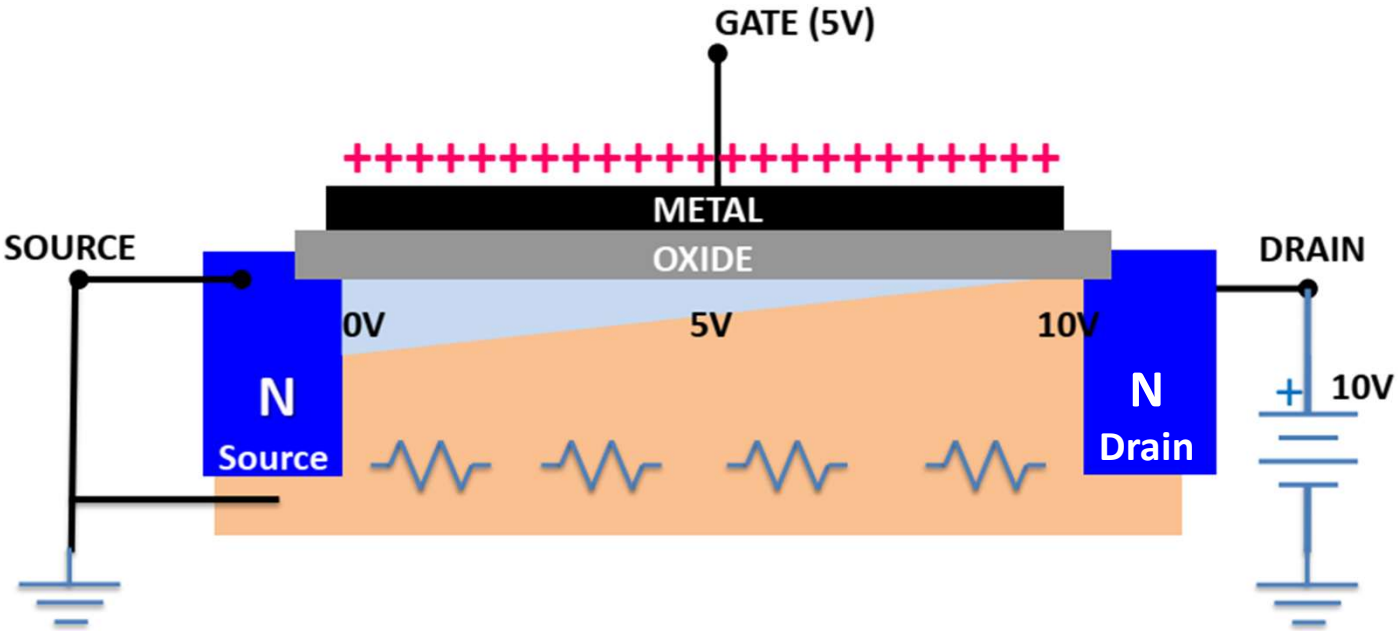
- If the inversion region disappears, the p-n junction returns. The bulk goes back to zero volts (the drain is detached). The inversion region then re-appears and current flows again. When we get these sort of on/off behaviours, in practice we hit an equilibrium between the on/off stages and we get a stable current value. This is the saturation region. The current is ideally stuck at the maximum value that allows an inversion region everywhere.

Channel Pinch-Off



- The equilibrium state is reached when current × (resistance of the inversion region) > (gate voltage - bulk voltage - V_T).
- Current × (resistance of the inversion region) is the source-drain voltage, V_{DS}.

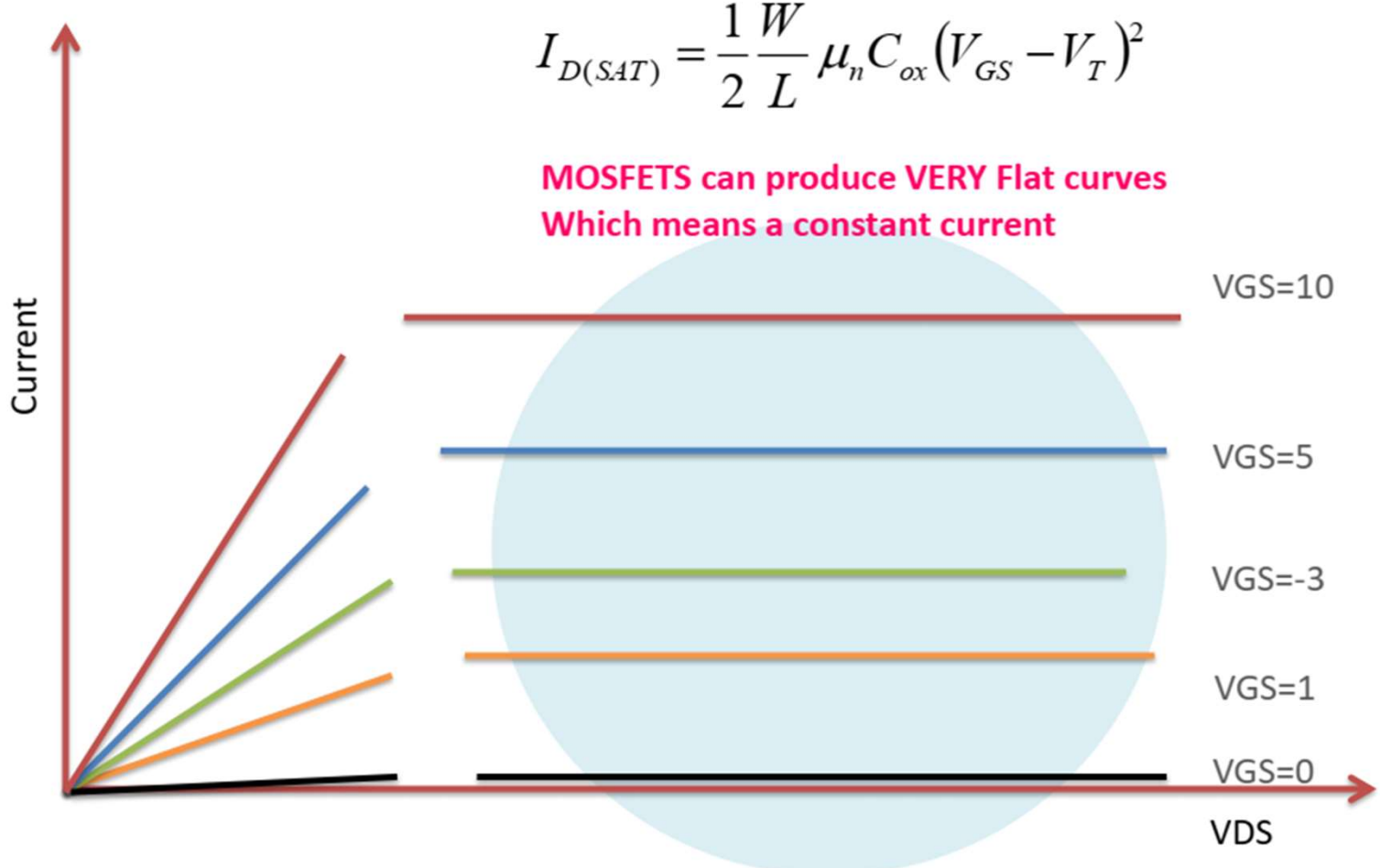
Operation Voltages of MOSFETs



- For simplicity, the bulk is nearly always connected to the source, so now we can say that when
  - $V_{GS} > V_T$: an inversion region forms;
  - $0 < V_{DS} < V_{GS} - V_T$: a current flows in the inversion region;
  - $V_{DS} = V_{GS} - V_T$: we hit the maximum current that can flow;
  - $V_{DS} > V_{GS} – V_T$: no more current should ideally flow.

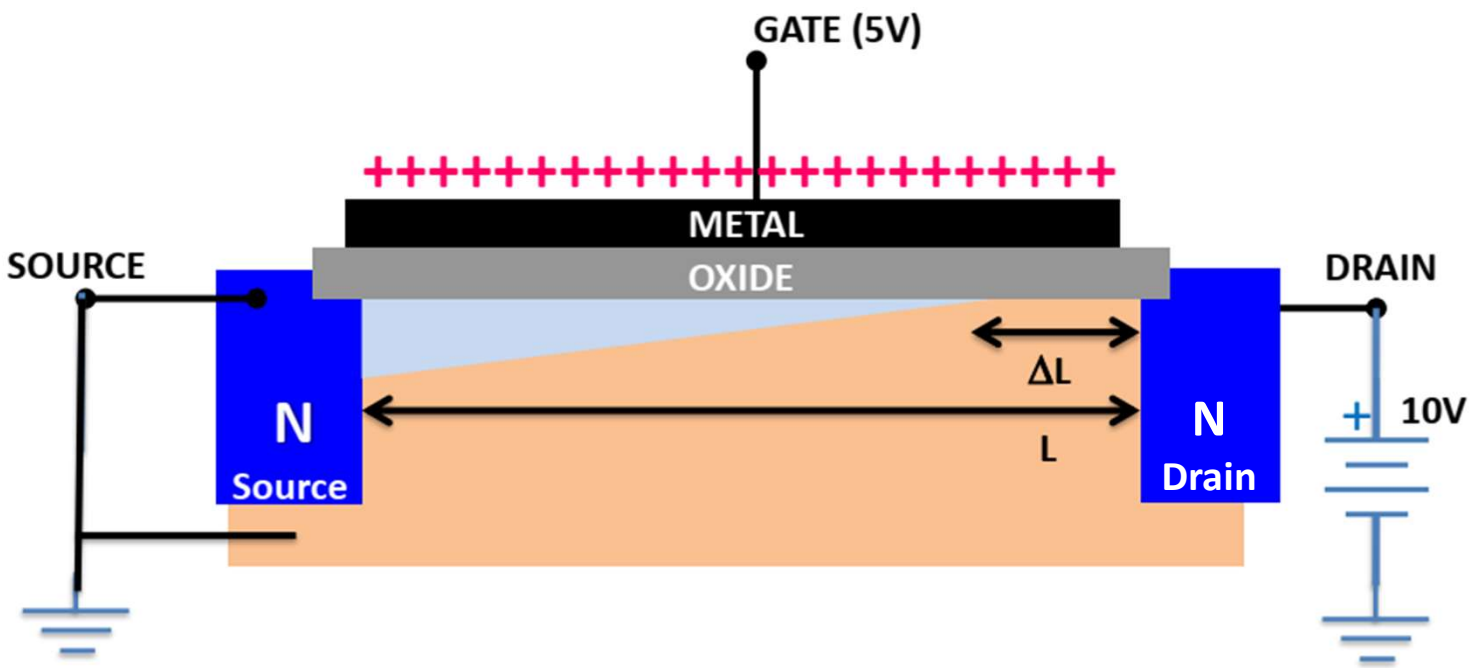
Saturation Region

$$I_{D(SAT)} = \frac{1}{2}\frac{W}{L}\mu_n C_{ox}\left(V_{GS} - V_T\right)^2$$



You can get the $I_{D(SAT)}$ equation by letting $V_{DS} = (V_{GS} - V_T)$.

Channel Length Modulation

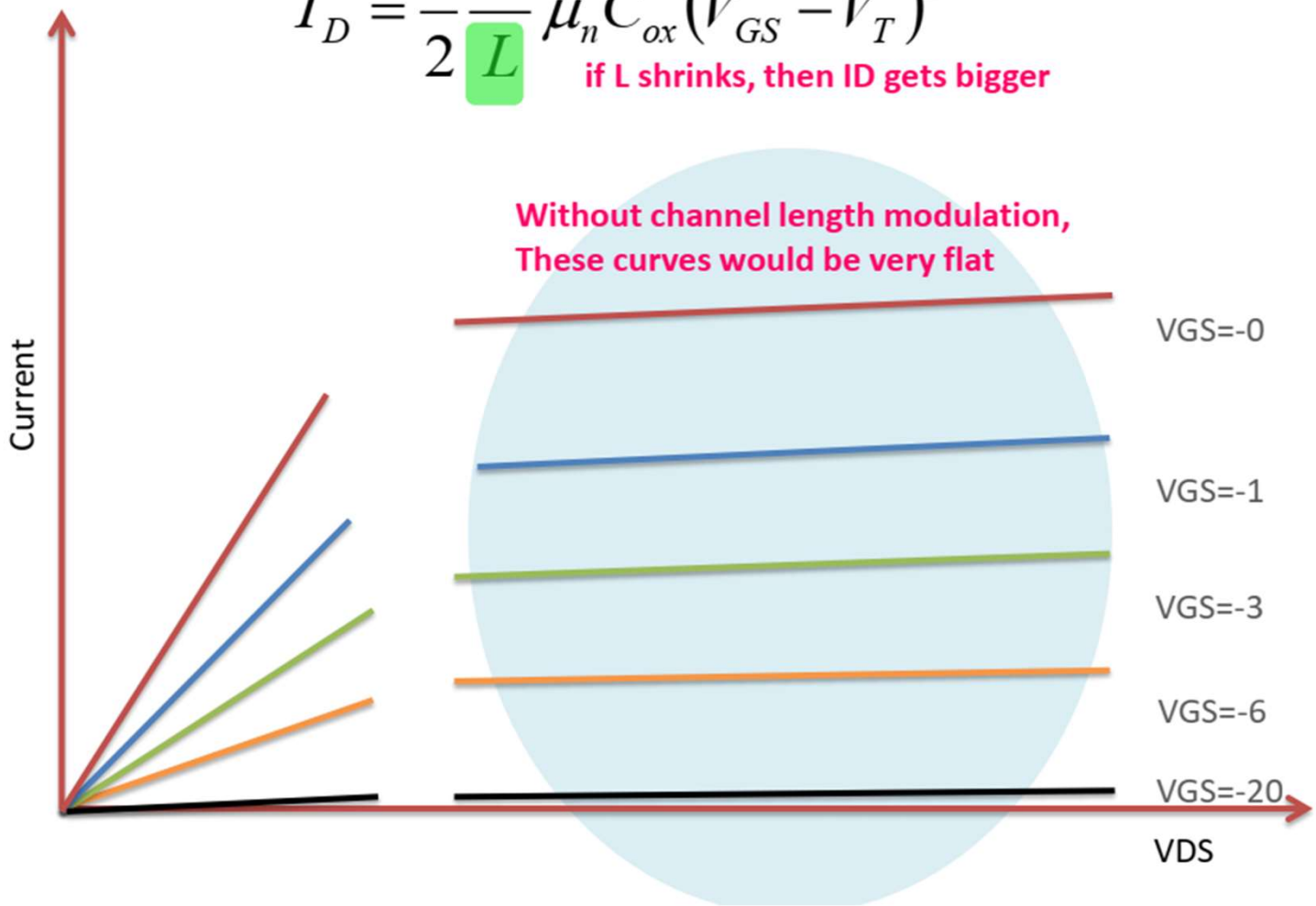


- Effectively as we increase V_{DS} in the saturation region, the channel gets smaller and the gap in between is filled with a depletion region.
- The effective length of the channel becomes (original length - the gap).
- A shorter channel is less resistive ($R = \rho L/A$) so that means more current can flow.

Channel Length Modulation

$$I_D = \frac{1}{2}\frac{W}{L}\mu_n C_{ox}\left(V_{GS} - V_T\right)^2$$

if L shrinks, then ID gets bigger



How do We Avoid it?

- Bigger devices, that make a change in L less of an issue. For example,

- For ΔL = 1 nm,
- L_1 = 50 nm (device 1)
- L_2 = 500 nm (device 2)

- $(L_1 - \Delta L)/L_1$ = 49/50 = 98 % of original value.
- $(L_2 - \Delta L)/L_2$ = 499/500 = 99.8 % of original value.

- There are also other ways being developed in the industry to make this effect less of an issue.

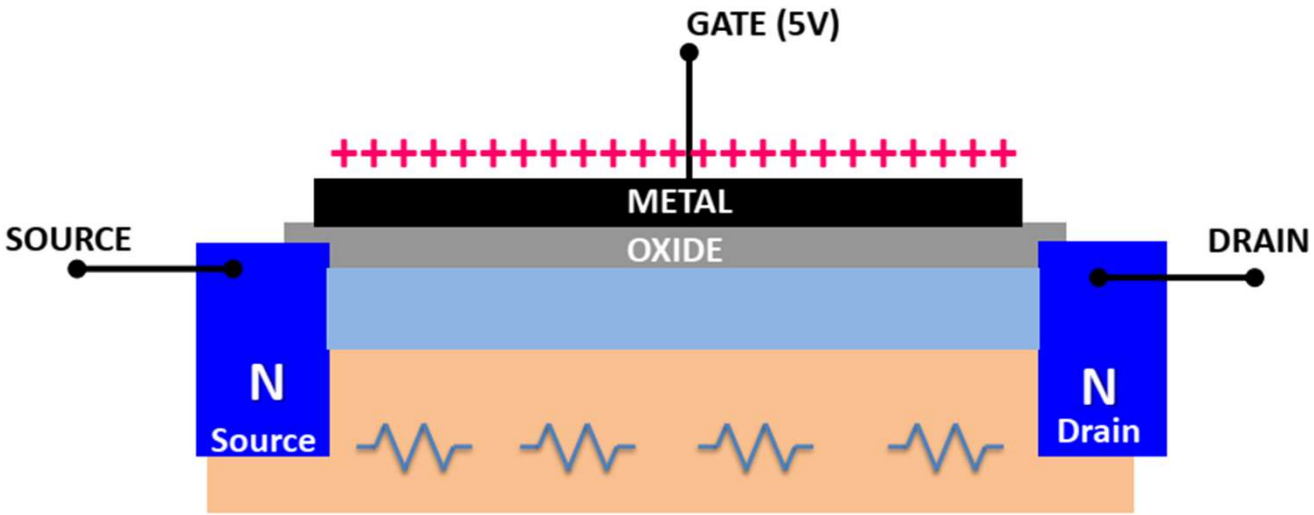
Key Equations for MOSFETs

- $I_D = \mu_{\mathrm{n}} C_{\mathrm{ox}} \frac{W}{L} [(V_{\mathrm{GS}} - V_{\mathrm{T}}) V_{\mathrm{DS}} - \frac{V_{\mathrm{DS}}^2}{2}]$
- $I_{D(\mathrm{sat})} = \frac{1}{2} \mu_{\mathrm{n}} C_{\mathrm{ox}} \frac{W}{L} (V_{\mathrm{GS}} - V_{\mathrm{T}})^2$
- $I_D = I_S$
- $I_G = 0$

- So this comes from the device physics, and different transistors will have different values, but the "shape" of the equation will remain the same for MOSFETs.

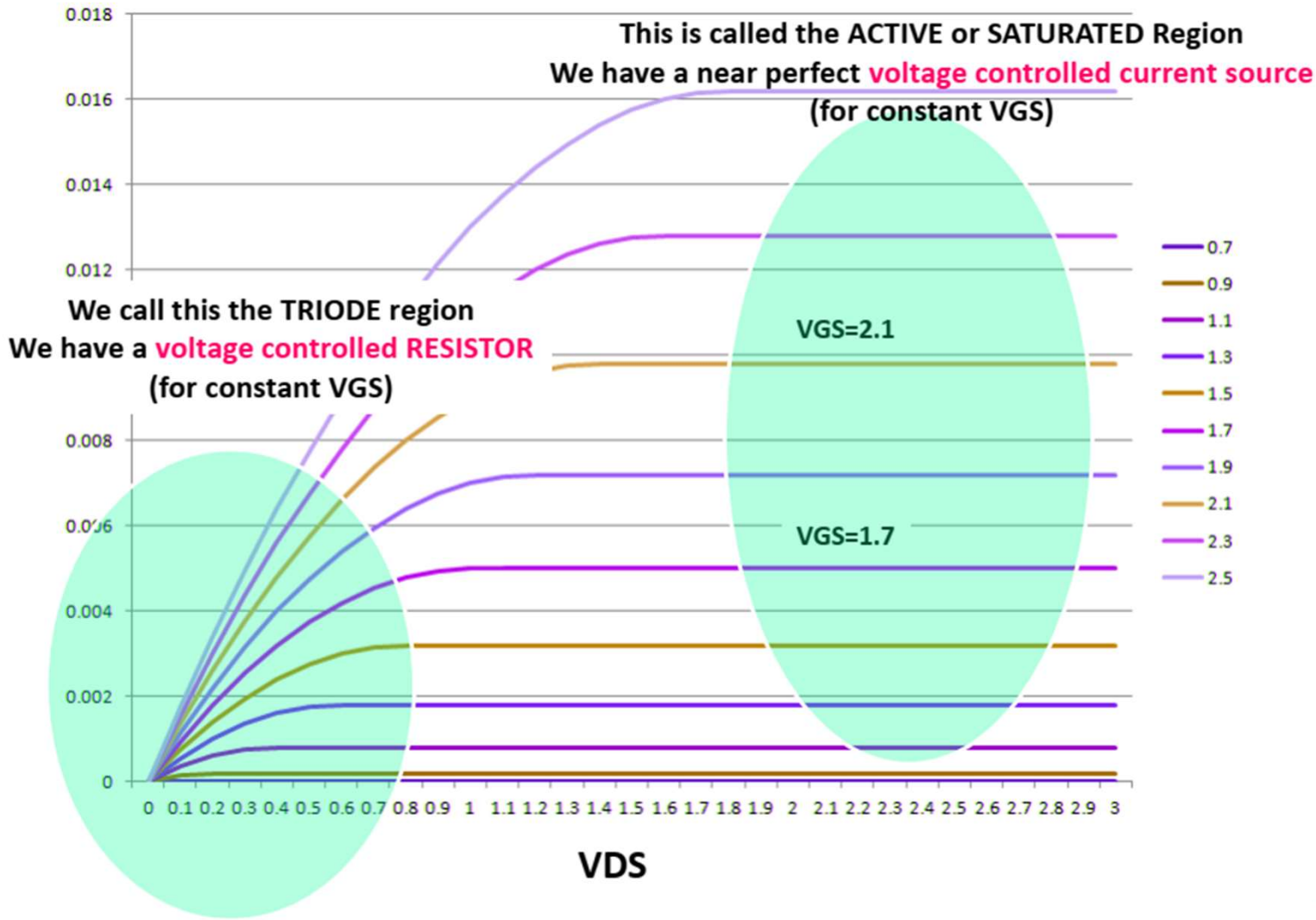
Key Equations for MOSFETs

$$\boxed{I_G = 0} \quad \boxed{I_D = I_S}$$

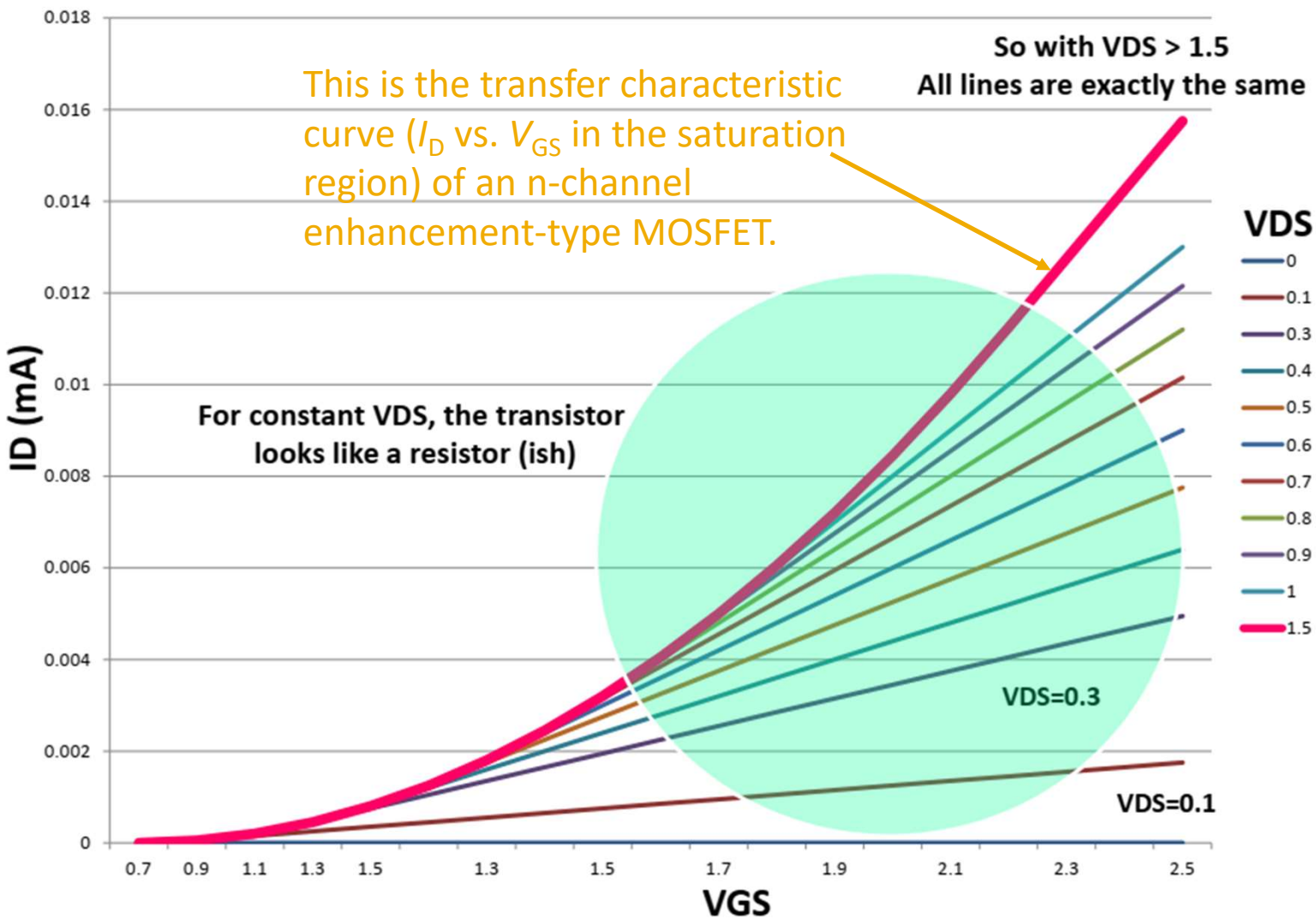


- The gate oxide is a good insulator. It blocks any current flowing from the gate to either SOURCE or DRAIN.
- It also means that any current entering from the SOURCE must go to the DRAIN.

Plotting Key MOSFET Equations

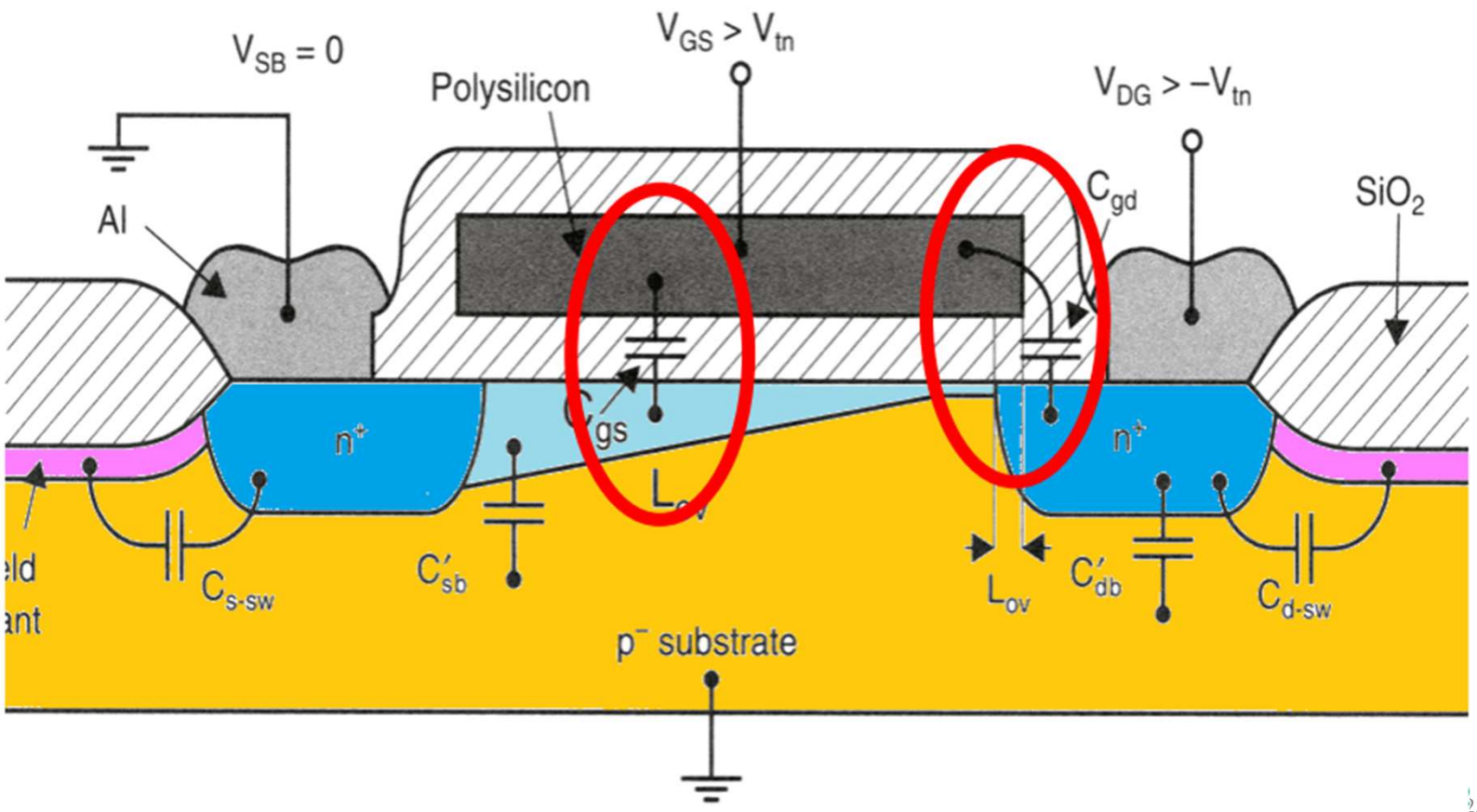


Plotting Key MOSFET Equations (I_D vs. V_{GS})

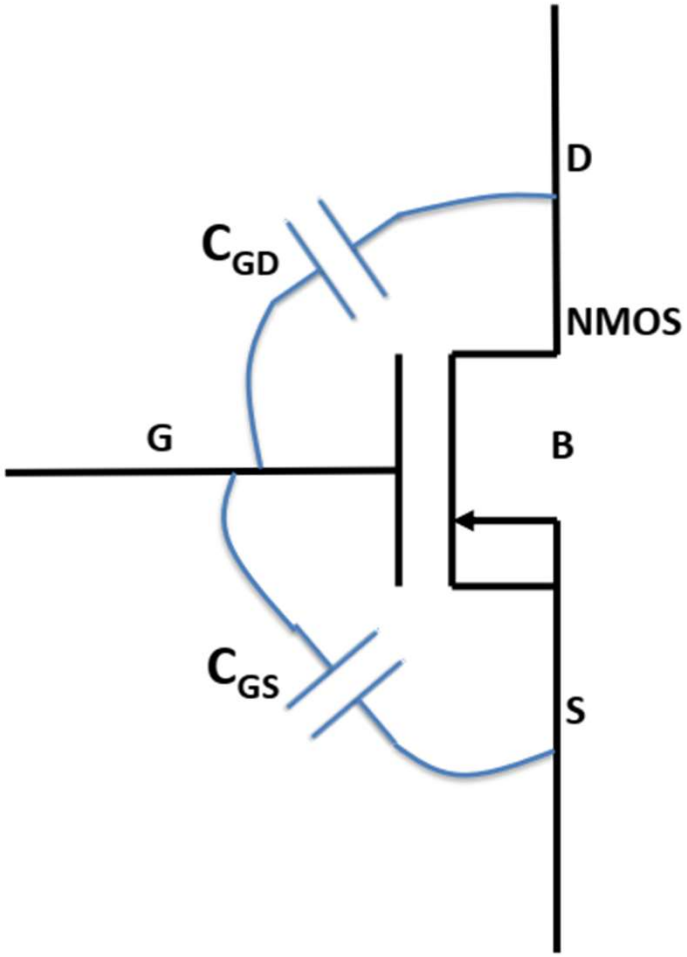


Parasitic Components in MOSFETs

Capacitors form when charged areas are separated by an insulator. Any conducting region overlapping with another is potentially a capacitor.



Parasitic Components in MOSFETs



- There are other parasitic capacitors, but these are the two most important ones. We won't initially worry about them, but we will have to come back to them.
- They do not cause a problem when we are changing things slowly as they charge/discharge rapidly (compared to how fast things change).
- But when we are doing things 100's and 1000 of million times / sec, it is these parasitic capacitances that limit the performance.
- C_{GS} is the biggest in size but C_{GD} has the biggest impact.

Key Points to Remember

- In practice, we can assume NO CURRENT flows into the gate terminal.
- In the saturation region, the curves are VERY FLAT.
- When the device is turned off, they let VERY LITTLE current through.
- The key equation for everything is (when $V_{GS} > V_T$):

$$I_D = \mu_n C_{ox} \frac{W}{L} [(V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2}]$$

- And in the saturation region, i.e., when $V_{DS} > (V_{GS} - V_T)$:

$$I_{D(sat)} = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_T)^2$$

In Practice

- If you are using transistors in a lab, the equations usually are given in the form of curves and you just read the values from the datasheet.
- The equations are scary but if you use them in the most common ways, then they are really easy to understand. All the equations do will help you explain why things are changing.
- If you design computer chips for Intel or Analog Devices, you can choose the key parameters *W*/*L* to change their behaviour to what you need.
- NOTE: The transistor equation here is a very simple equation and the real equations can fill many A4 pages with thousands of parameters. That's simply too complex for humans to get any meaningful info from. Excellent for computers though. 😁

NMOS & PMOS

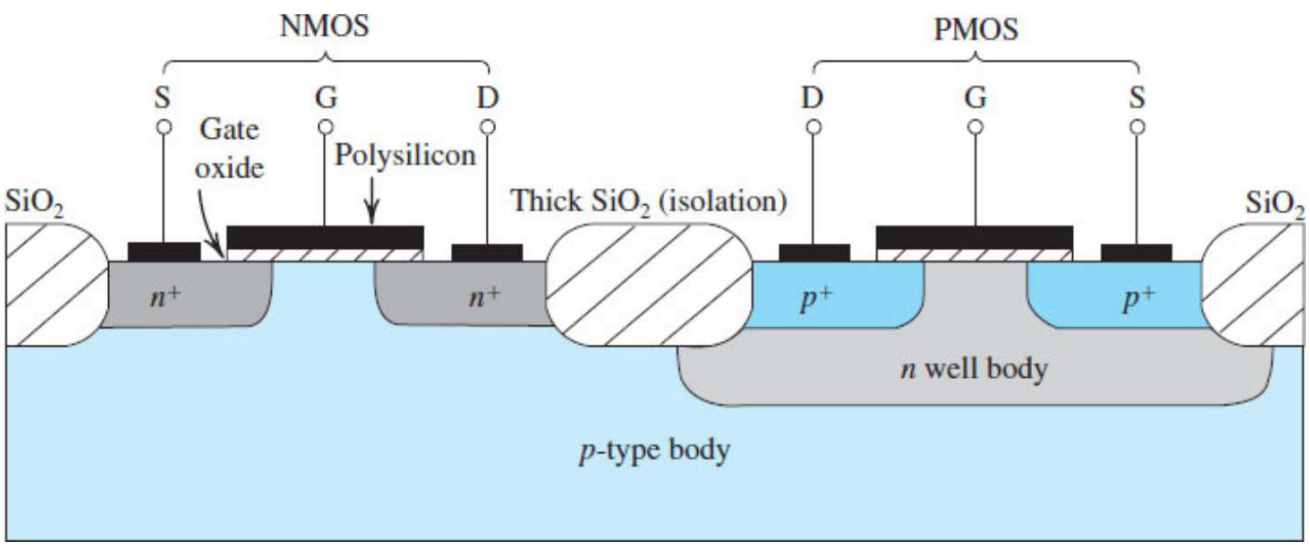
- Based on the majority charge carriers in the channel region, MOSFETs come in two versions:
  - NMOS (n-channel);
  - PMOS (p-channel).

- NMOS and PMOS are complementary devices.

- For enhancement-type MOSFETs:
  - NMOS need the gate voltage to be HIGHER than source to work.
  - PMOS need the gate voltage to be LOWER than source to work.

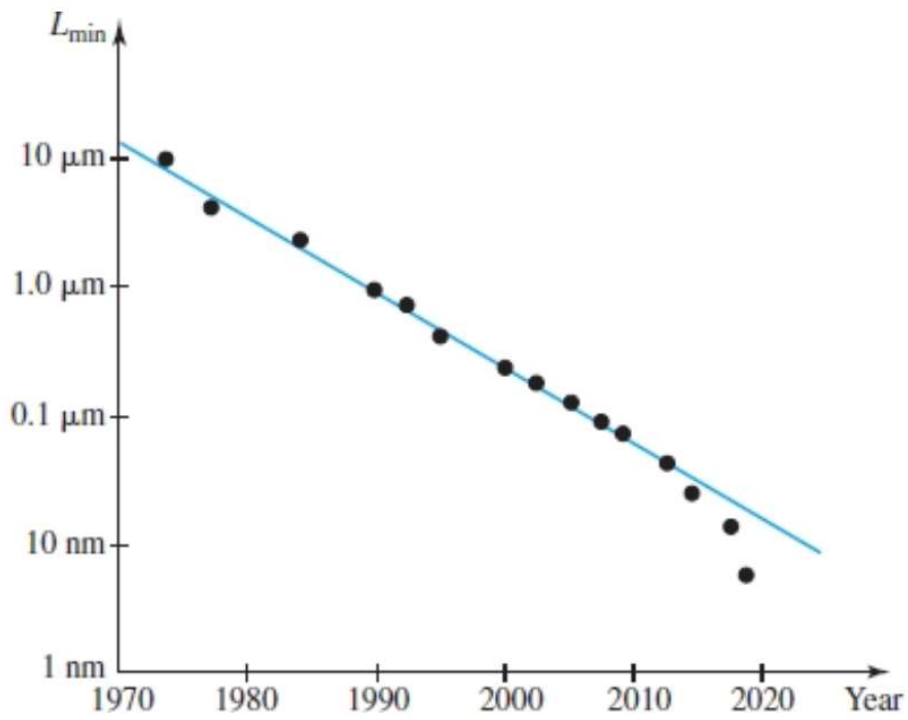
CMOS

- Effective logic circuits can be fabricated by putting n-channel and p-channel MOSFETs onto the same substrate;
- This configuration is called complementary MOSFET arrangement (CMOS) which has significant applications in modern digital electronics.



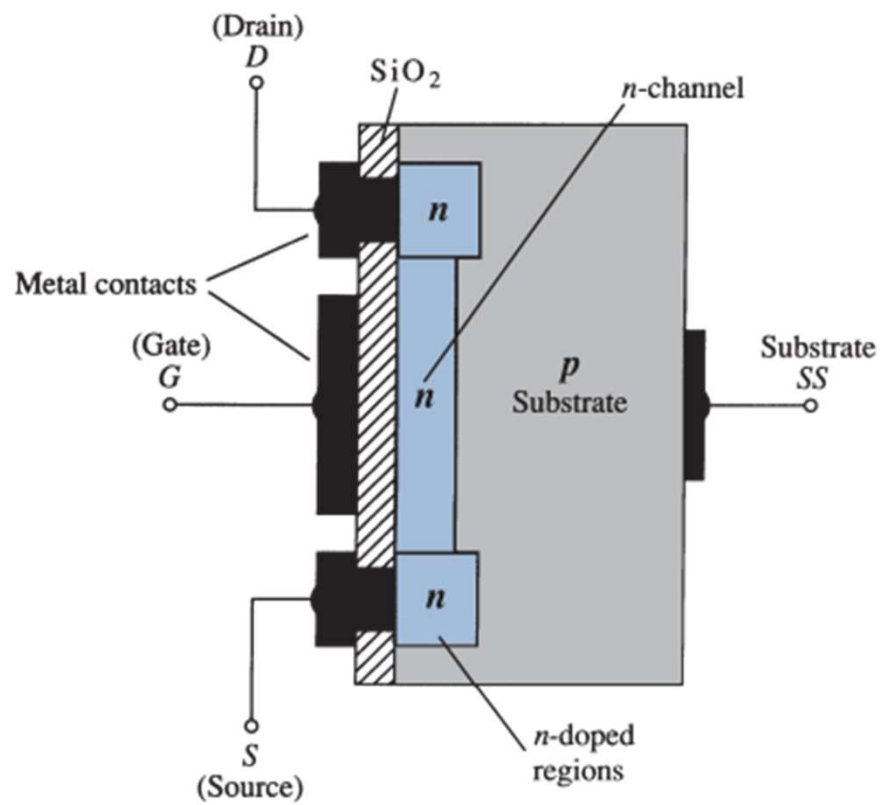
Moore's Law

- Advances in CMOS fabrication technology have led to the reduction in minimum channel length in MOSFETs by 30% every 2 – 3 years;
- This means that one can assemble twice as many transistors on a silicon chip of the same area. This is the famous Moore's law.



Enhancement- & Depletion-Type MOSFETs

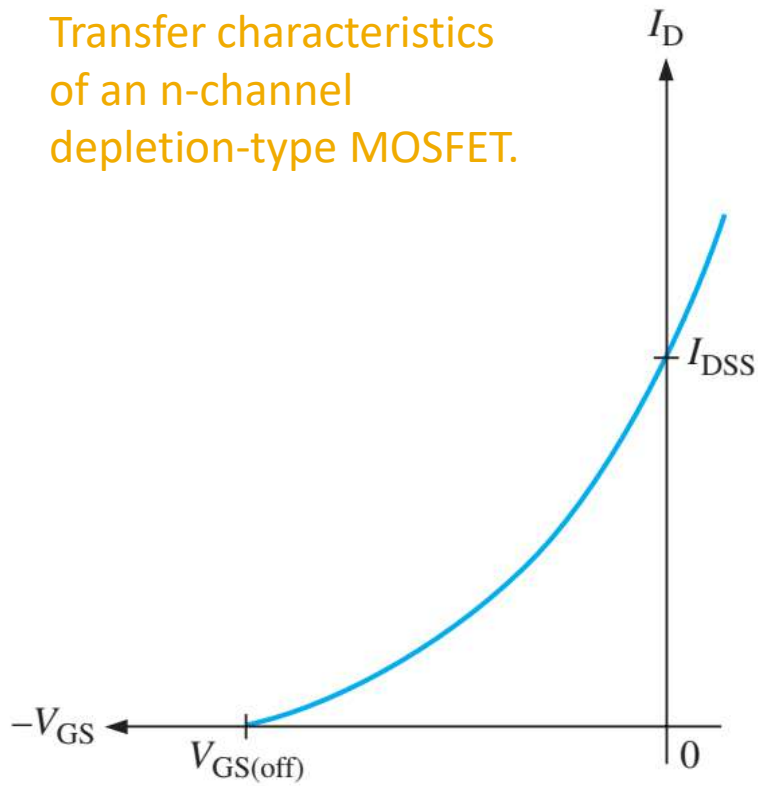
- In all the MOSFETs we discussed so far, the conduction channel is induced by the gate voltage (no channel when V_{GS} = 0). We call these devices enhancement-type MOSFETs.
- There is another type of MOSFET devices where a channel exists under V_{GS} = 0 (it is realised through doping the substrate). These devices are called depletion-type MOSFETs.



Depletion-Type MOSFETs

- Depletion-type MOSFETs are 'normally on' devices (they are conducting when $V_{GS} = 0$);
- Depletion-type MOSFETs can operate in either the depletion-mode or the enhancement-mode (gate voltage may be either positive or negative).

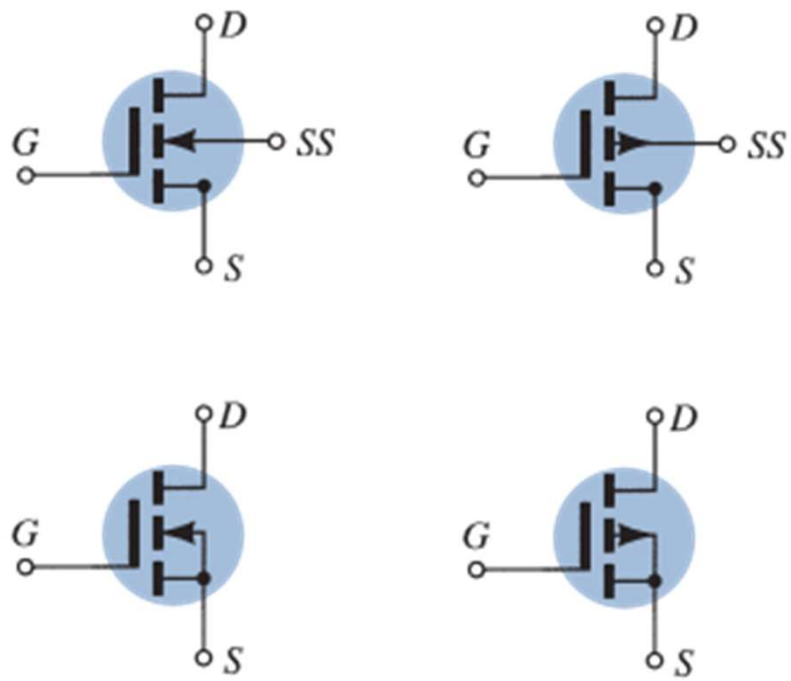
Transfer characteristics of an n-channel depletion-type MOSFET.



Watch Out! It may have already occurred to you. MOSFET equations discussed previously do NOT work for depletion-type MOSFETs.

Circuit Symbols for Enhancement-Type MOSFETs

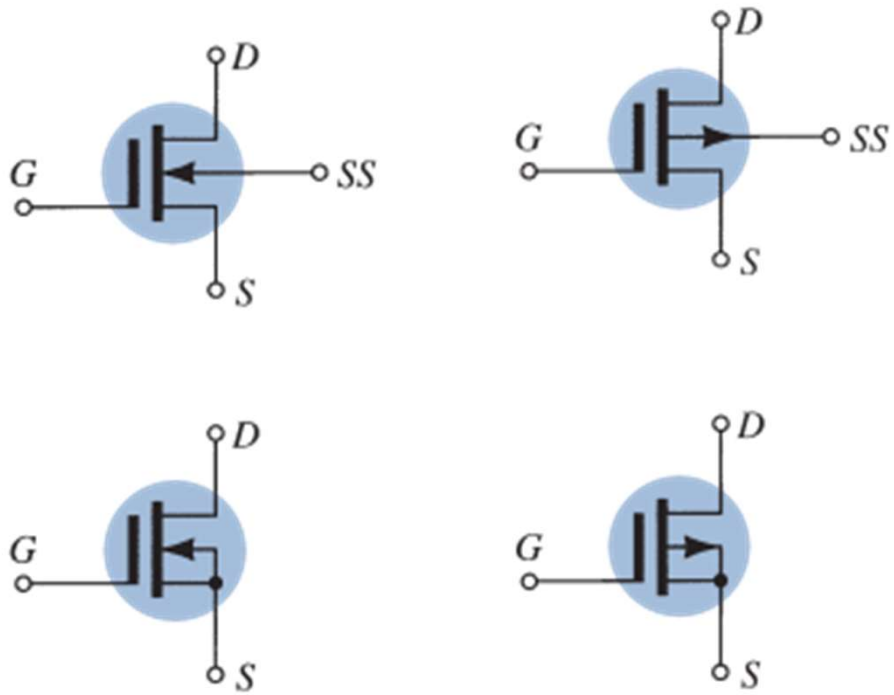
n-channel p-channel



- Pay attention to the directions of the arrows in the symbols. The direction of the arrow represents the forward bias direction of a p-n junction in the FET.

Circuit Symbols for Depletion-Type MOSFETs

n-channel p-channel

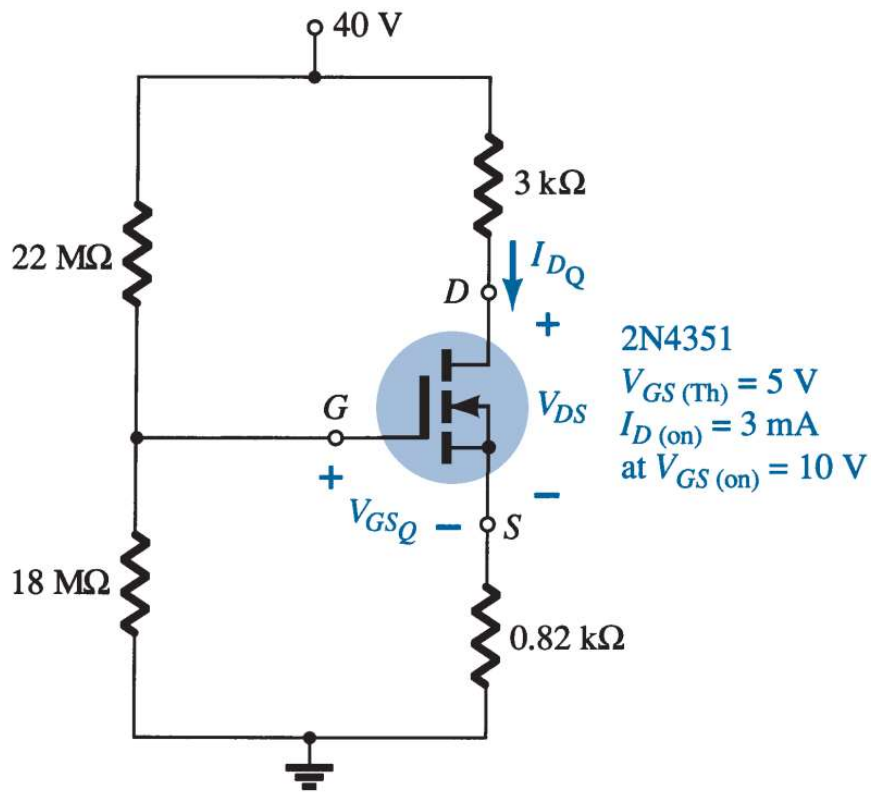


- Pay attention to the difference between enhancement-type and depletion-type MOSFETs. The channels are represented by dashed lines in enhancement-type devices but by solid lines in depletion-type devices.

Example: MOSFET Equations

- Find I_{DQ}, V_{DS} and V_{GSQ}.

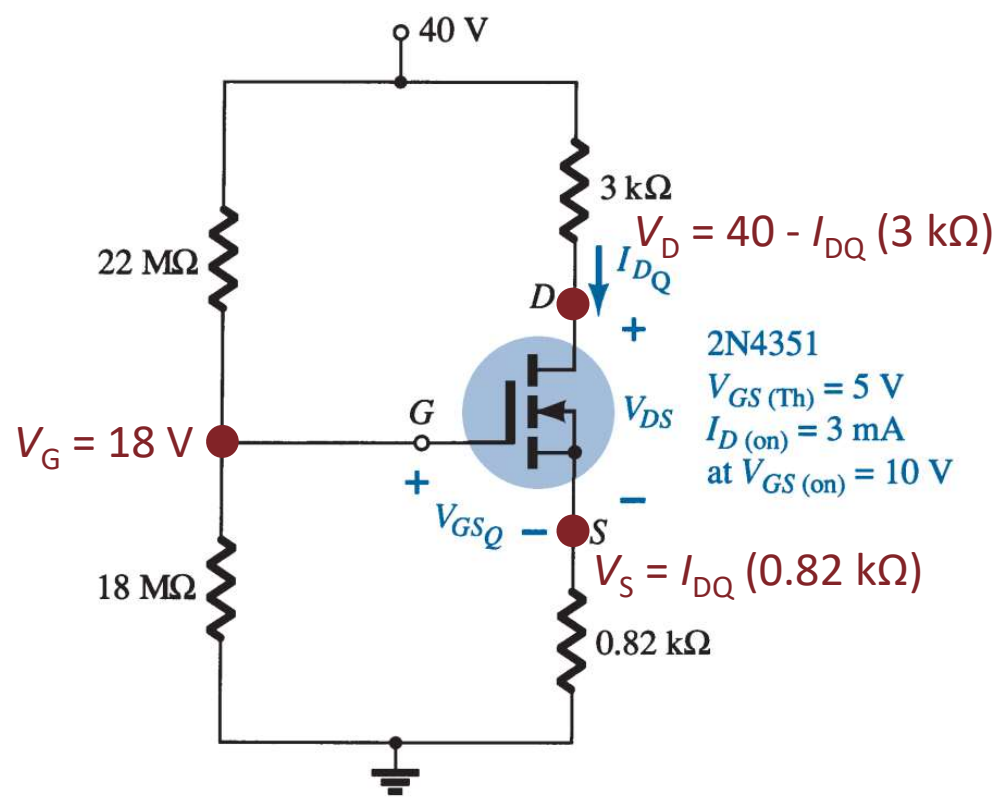
- The symbol "Q" refers to the quiescent (operating) point or simply the Q point. This is the DC operating point of the transistor (to be discussed more in the coming weeks).



Example: MOSFET Equations

- $I_{\mathrm{D}} = K(V_{\mathrm{GS}} - V_{\mathrm{T}})^2$ in which $K = \frac{1}{2}\left(\frac{W}{L}\right)\mu_{\mathrm{n}} C_{\mathrm{ox}}$;
- Insert the information from the data sheet,

$$K = \frac{I_{\mathrm{D}}}{(V_{\mathrm{GS}} - V_{\mathrm{T}})^2}$$
$$= 1.2 \times 10^{-4}\ \mathrm{A/V^2}$$



Example: MOSFET Equations

- We can build an equation on

$$I_{\mathrm{DQ}} = K\left(18 - 820 I_{\mathrm{DQ}} - 5\right)^2$$

- Solving the equation gives:

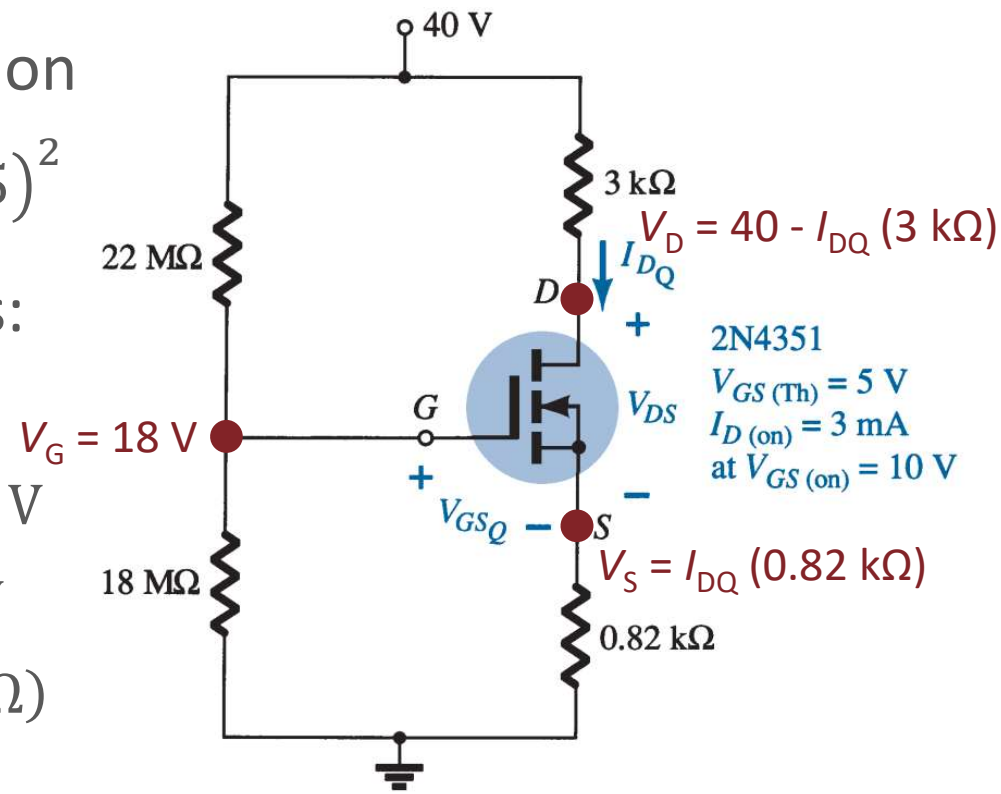
$$I_{\mathrm{DQ}} = 6.73\ \mathrm{mV}$$

$$V_{\mathrm{S}} = I_{\mathrm{DQ}}(0.82\ \mathrm{k\Omega}) = 5.5\ \mathrm{V}$$

$$V_{\mathrm{GS}} = V_{\mathrm{G}} - V_{\mathrm{S}} = 12.5\ \mathrm{V}$$

$$V_{\mathrm{DS}} = 40 - V_{\mathrm{S}} - I_{\mathrm{DQ}}(3\ \mathrm{k\Omega})$$

$$= 14.3\ \mathrm{V}$$



One should double check and make sure that the MOSFET is in the saturation region. How can one do it?

EE204FZ
Lecture 5
Applications of MOSFETs

Zhu DIAO

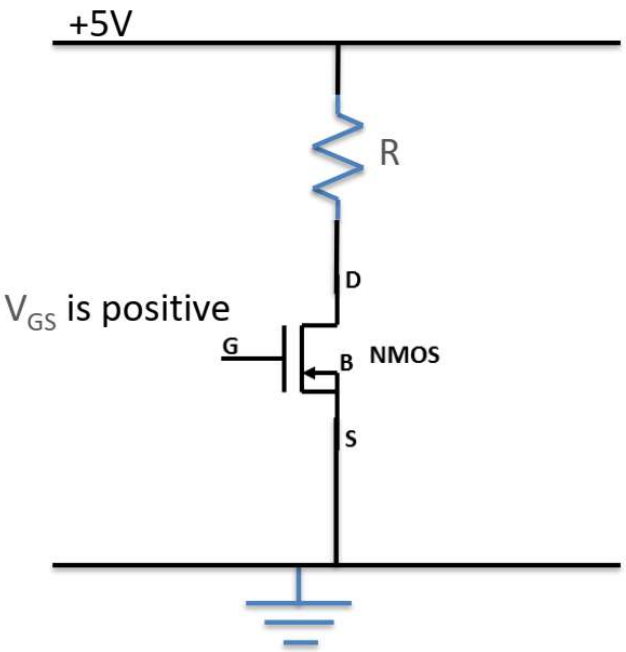
Email: zhu.diao@mu.ie

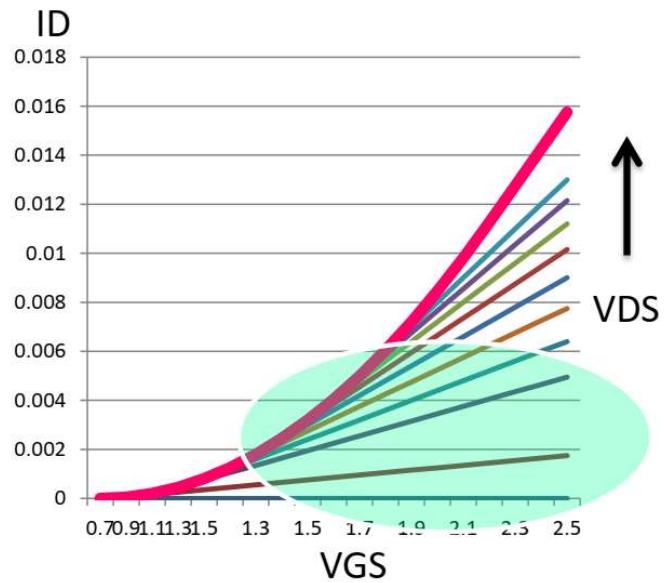
How Can We Use MOSFETs?

- We can use transistors in many ways, but the most common ones are
  - as a resistor;
  - as a switch;
  - as an amplifier (will be discussed in the next section of this course module).
- We will go through each of these briefly and talk about them and show some examples.

As a Resistor

- Consider a device where we want a small DC resistance. Assuming reasonable currents, the voltage dropped across a small resistance will be a small voltage. So V_{DS} is small.
- Normally you would see this in a structure like the following circuit.





Look at the Key Equation

- The first thing is to check what region of the transistor we are in. We said V_{DS} was to be small, so what does this mean?
- If $V_{DS} > V_{GS} - V_T$, then we are in saturation. For a small V_{DS}, we are not likely to be in saturation, but just to be sure, let us pick a voltage for V_{GS} where this is not true:

$$I_D = \mu_n C_{ox} \frac{W}{L} [(V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2}]$$

but $R_{DS} = V_{DS}/I_D$,

$$R_{DS} = \frac{V_{DS}}{\mu_n C_{ox} \frac{W}{L} [(V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2}]}$$

Look at the Key Equation

- Let's make an assumption. The square of a small number is a very small number. Let's remove the $(V_{DS})^2$. It adds a small error, but we can live with that.

$$R_{\mathrm{DS}} = \frac{1}{\mu_{\mathrm{n}} C_{\mathrm{ox}} \frac{W}{L} (V_{\mathrm{GS}} - V_{\mathrm{T}})}$$

in which

  - $\mu_n C_{ox} \approx 30 \times 10^{-6}$ A/V^2 (depending on doping, etc);
  - W/L is a manufacturing feature, but ratios of 5:1 to 100:1 are common. If you are doing IC design, you can control this;
  - V_T is a constant and can range from 0.3 to 2 V.

What Resistances Are Possible?

Let's use the following values:

- $\mu_n C_{ox} \approx 30 \times 10^{-6}$ A/V^2 (depending on doping, etc);
- $W/L = 50$;
- $V_T = 0.5$ V.

$$R_{DS} = \frac{1}{(30 \times 10^{-6})(50)\,(V_{GS} - 0.5)}$$

So therefore for

$V_{GS} = 1$ V	$R_{DS} = 1333\ \Omega$
$V_{GS} = 2$ V	$R_{DS} = 444\ \Omega$
$V_{GS} = 5$ V	$R_{DS} = 148\ \Omega$

If we double W/L, we halve the resistance. Hence, we can make pretty accurate low-value resistors.

What About High-Value Resistances?

- Here we are talking about the saturation region (since if we are in the triode region, $V_{GS} \uparrow$, $R_{DS} \downarrow$);
- We have a large value of V_{DS}, so we need to use the saturation region equation. Ideally, the maximum current is given by

$$I_{\mathrm{D(sat)}} = \frac{1}{2}\mu_{\mathrm{n}} C_{\mathrm{ox}} \frac{W}{L}(V_{\mathrm{GS}} - V_{\mathrm{T}})^2$$

and the resistance is

$$R_{\mathrm{DS}} = \frac{V_{\mathrm{DS}}}{\frac{1}{2}\mu_{\mathrm{n}} C_{\mathrm{ox}} \frac{W}{L}(V_{\mathrm{GS}} - V_{\mathrm{T}})^2}$$

- This is not working out very nicely for us, but let us plug in some numbers as from before.

What About High-Value Resistances?

- Let's use the following values:
  - $\mu_n C_{ox} \approx 30 \times 10^{-6}$ A/V^2 (depending on doping, etc);
  - $W/L = 50$;
  - $V_T = 0.5$ V.

Remember, $V_{DS} > V_{GS} - V_T$ for saturation region.

	V_{DS} = 1 V	V_{DS} = 2 V	V_{DS} = 3 V	V_{DS} = 5 V
V_{GS} = 1 V	5333	10666	15999	26666
V_{GS} = 2 V	592	1185	1777	2692
V_{GS} = 3 V	213	426	619	1065
V_{GS} = 5 V	65	130	195	325

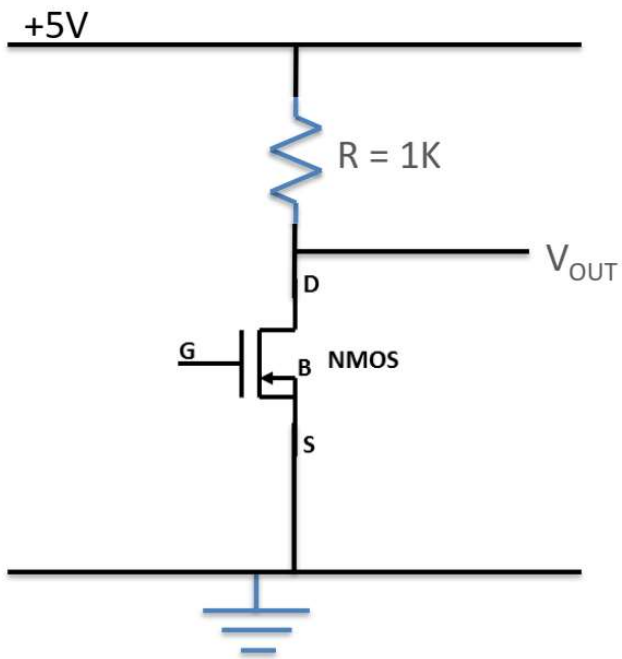
- We can't get good high valued DC resistance from a MOSFET, except under one condition.....

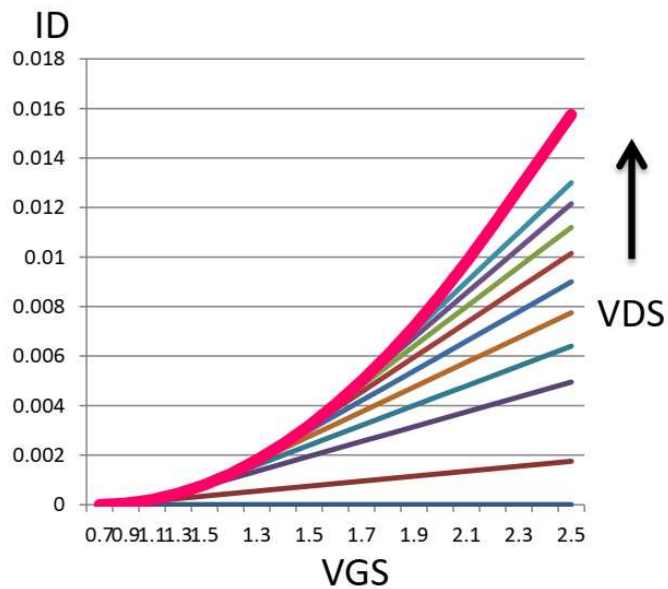
Summary: The Transistor as a Resistor

- If V_{GS} is less than the threshold voltage:
  No current flows, a very high resistance (MΩ);
- If $V_{GS} > V_T$, but V_{DS} is small and we are not in saturation:
  A small, controllable resistance (50 - 500 Ω). This can be made smaller by making V_{GS} bigger or by making the device wider (W/L bigger);
- If $V_{GS} > V_T$, but V_{DS} is large and we are in saturation:
  A very changeable resistor – depending on V_{GS} and V_{DS} and not very large (50 - 50000 Ω).
- So why bother,
  - It is voltage controlled, so we can change it;
  - It can be precise, we can get a resistance exactly TWICE of another, etc;
  - It is smaller than the area we would need to make a resistor in silicon.

The Transistor as a Switch

- In this scenario, we turn the transistor ON as far as possible and then turn it OFF as far as possible. Consider the following circuit:

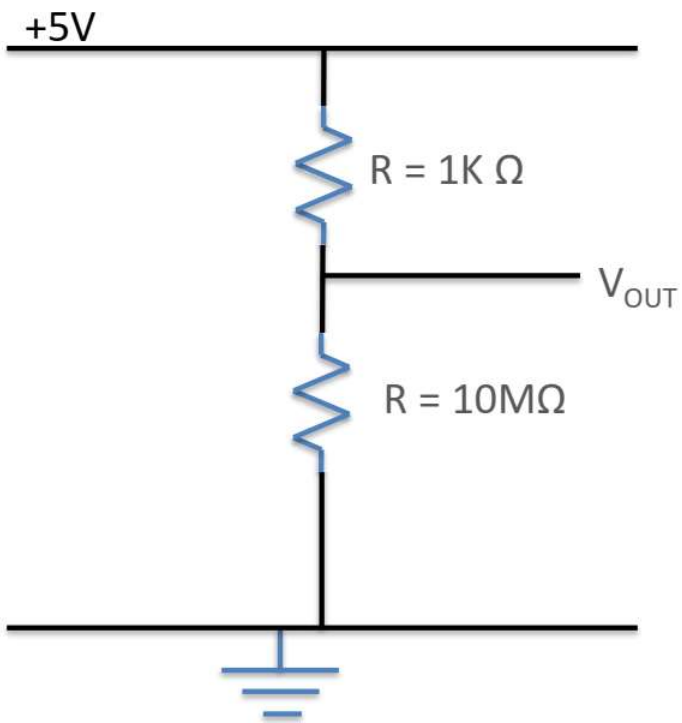




- What happens if I connect G to ground?
- What happens if I connect G to +5 V?

The Transistor as a Switch (OFF)

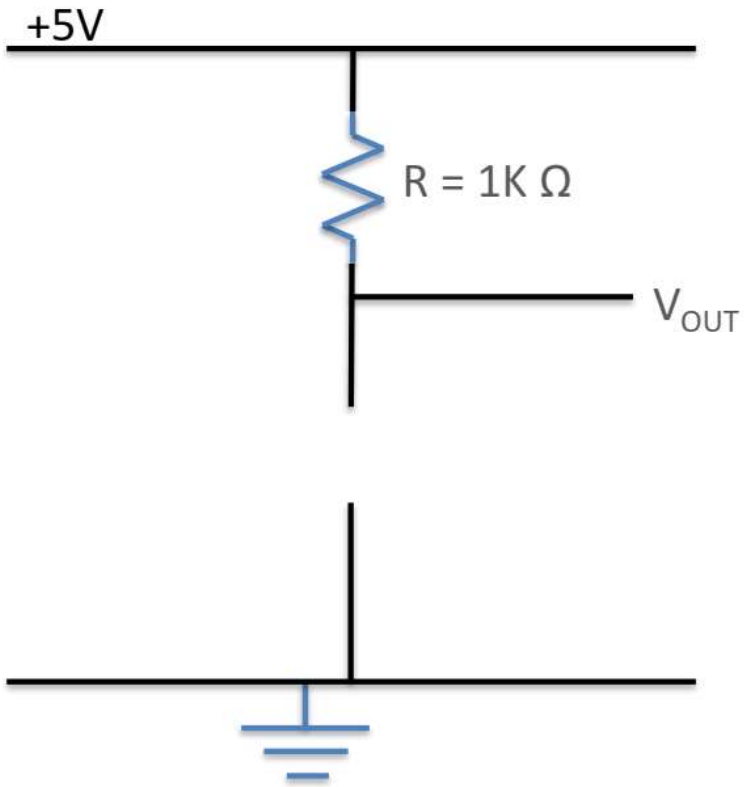
- When the gate is connected to ground, there is no inversion region. There are two back-to-back diodes. No current flows. A very high resistance.



What voltage will I see on V_{OUT}?

The Transistor as a Switch (OFF)

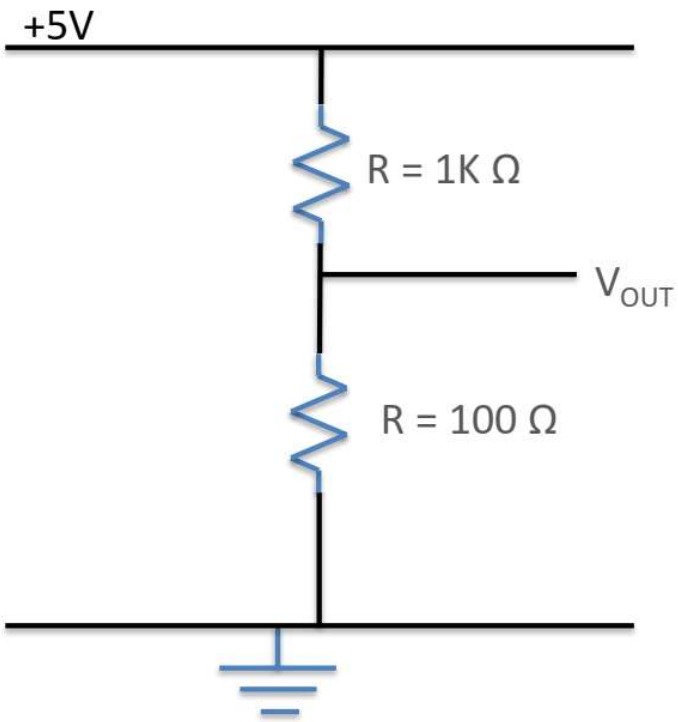
- An easier way to view this is to assume that a turned-off MOSFET is an open circuit.



What voltage will I see on V_{OUT}?

The Transistor as a Switch (ON)

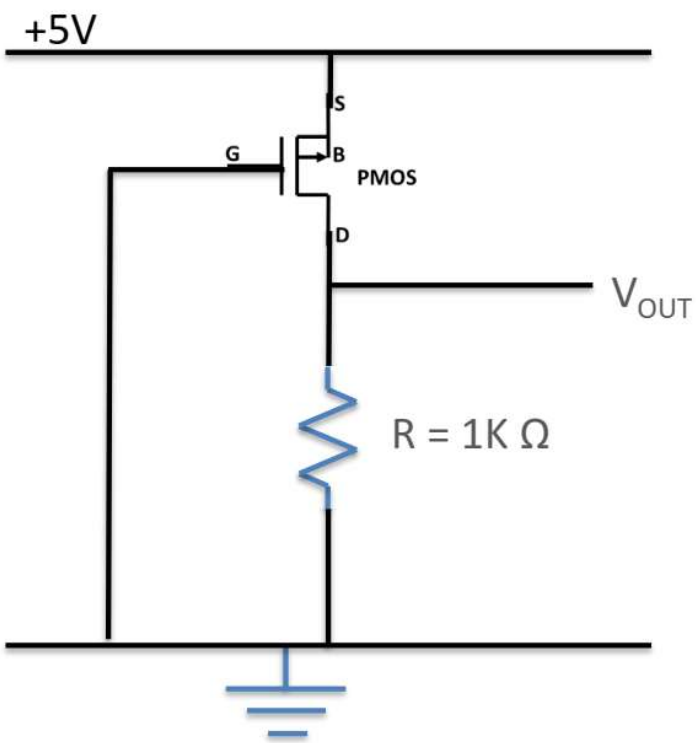
- Now when the GATE voltage is made positive with respect to the source/bulk ($V_{GS} > V_T$), we get an inversion region and we get a resistor defined by the characteristic equation (say 100 ohms). It could be lower depending on the transistor.

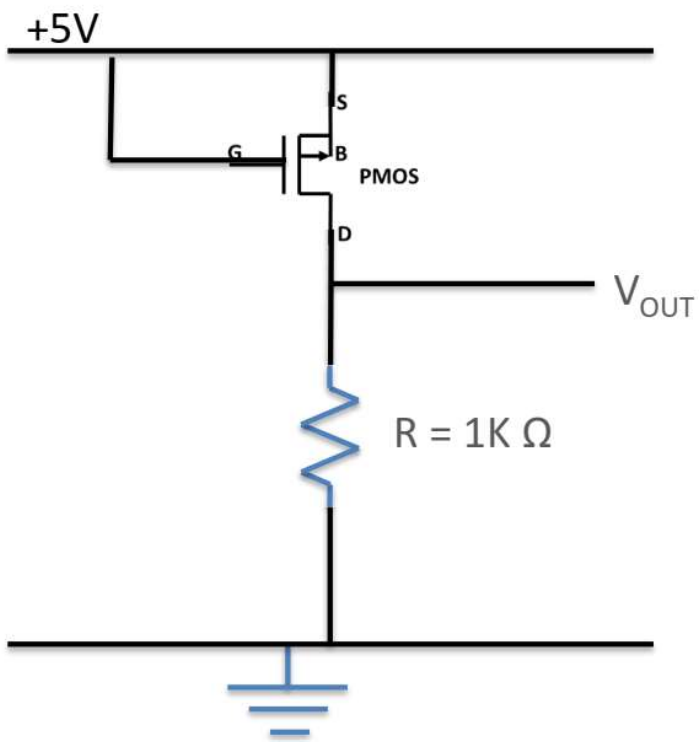


What voltage will I see on V_{OUT}?

Looking at the PMOS

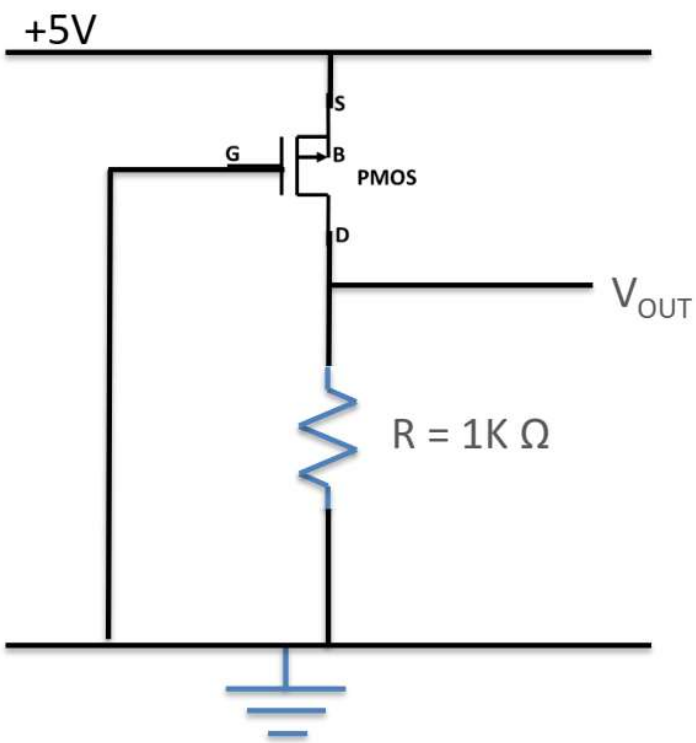
- What resistance will the transistor display for the following circuits?



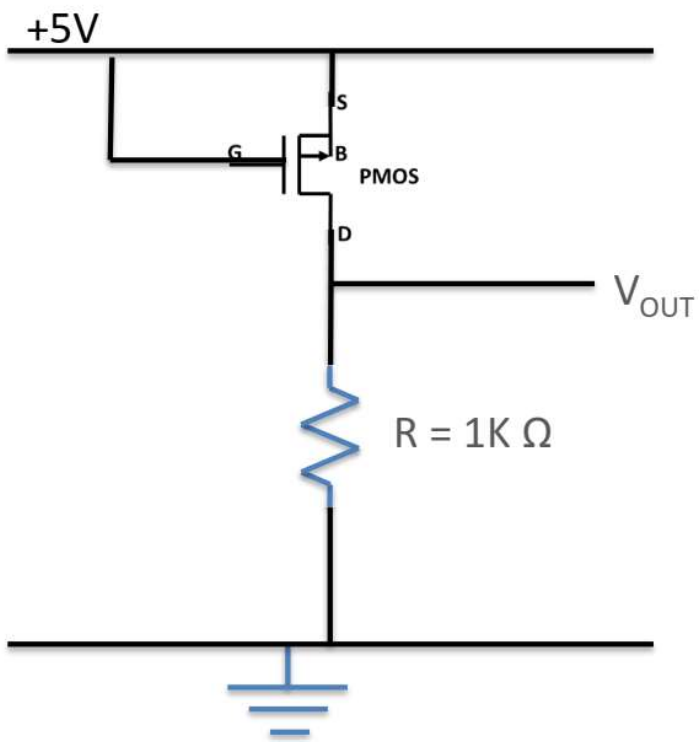


Looking at the PMOS

- What resistance will the transistor display for the following circuits?

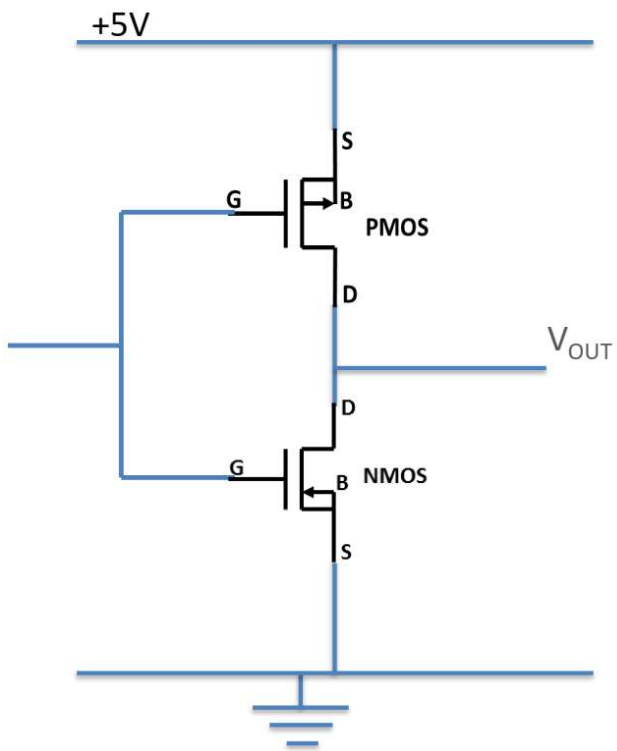


ON



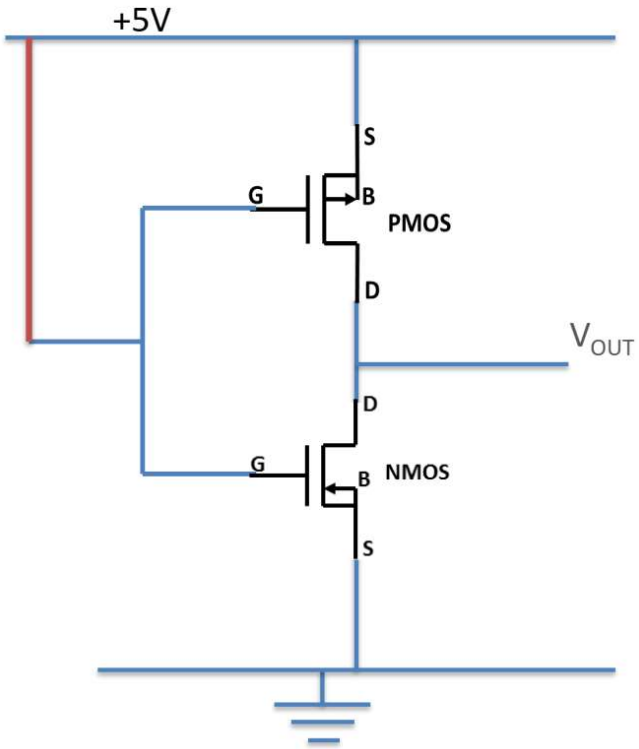
OFF

Consider the Following Circuit

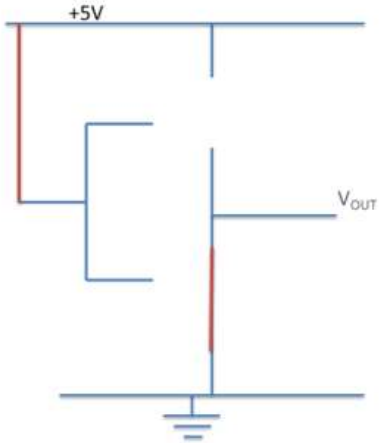


- What happens if I connect G to ground?
- What happens if I connect G to +5 V?

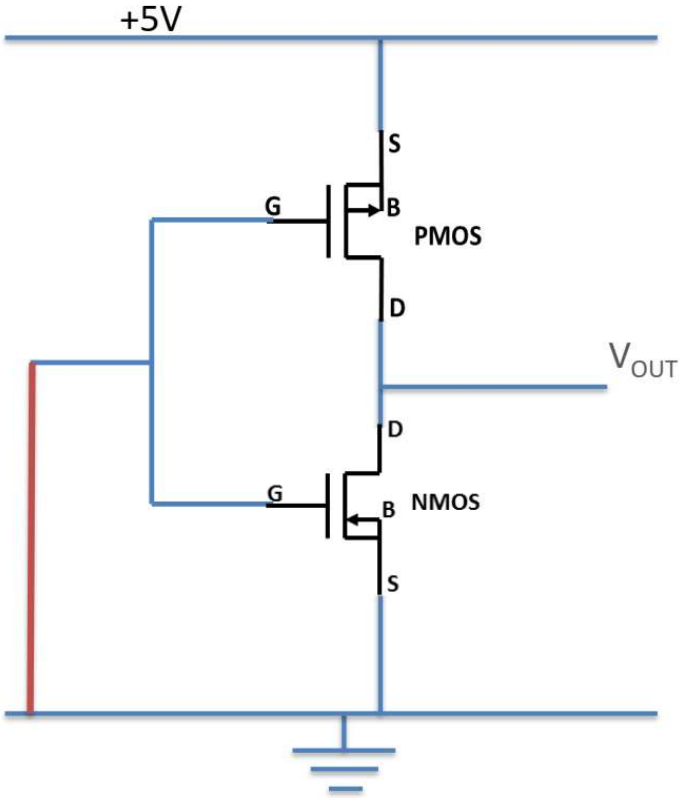
Gate is Connected High



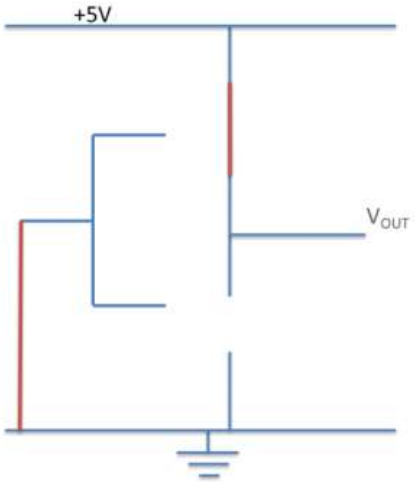
- When the gate is connected to 5 V, the NMOS has V_{GS} = 5 V, greater than V_T, turns on fully (saturated);
- The PMOS has V_{GS} = 0 V, less than $|V_T|$, turns off fully (no inversion region);
- The output V_{OUT} goes to ZERO.



Gate is Connected Low



- When the gate is connected to 0 V, the NMOS has V_{GS} = 0 V, less than V_T, turns off fully (no inversion region);
- The PMOS has V_{GS} = -5 V, $|V_{GS}|$ greater than $|V_T|$, turns on fully (saturated);
- The output V_{OUT} goes to 5 V.



A Table of the Results

	V_{OUT}
$V_G = 0$ V	5 V
$V_G = 5$ V	0 V

- If we considered a binary system to have voltages of 0 and 5 V to represent binary 0 and 1. What would this table correspond to in terms of a logic gate?

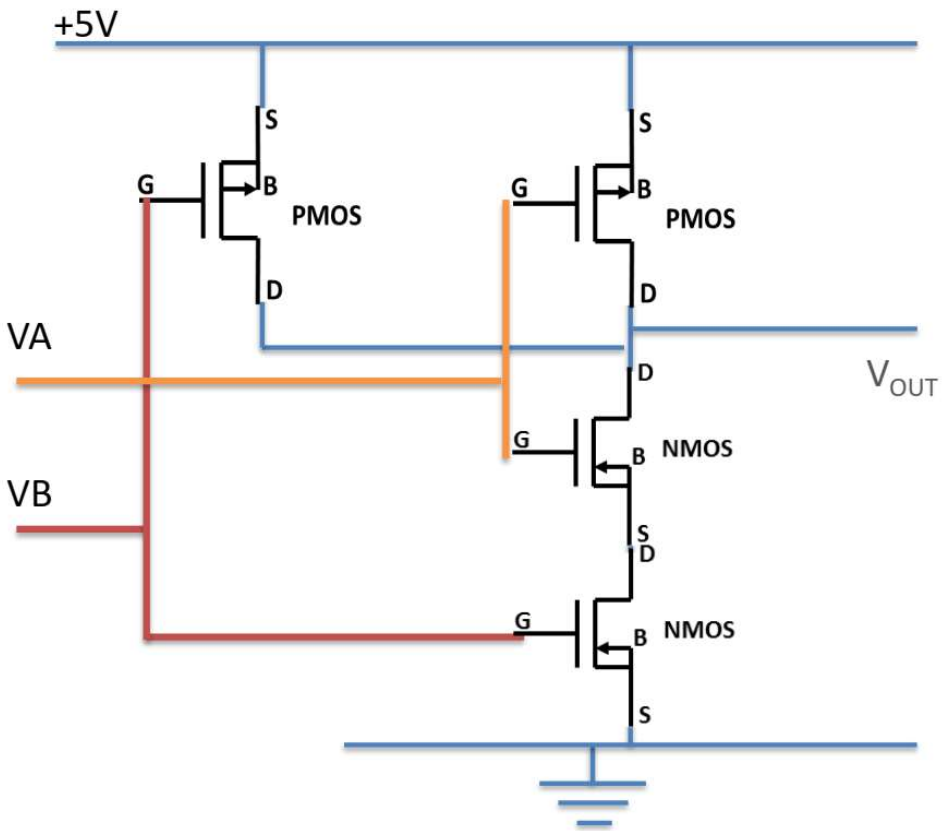
A Table of the Results

	V_{OUT}
$V_G = 0$ V	5 V
$V_G = 5$ V	0 V

- If we considered a binary system to have voltages of 0 and 5 V to represent binary 0 and 1. What would this table correspond to in terms of a logic gate?

This is an inverter (NOT gate)!

Consider the Following Circuit



- What happens to V_{OUT} as V_A and V_B take either 0 or 5 V and in different combinations?

A Table of the Results

V_A (V)	V_B (V)	V_{OUT} (V)
0	0	
0	5	
5	0	
5	5	

- If we considered a binary system to have voltages of 0 and 5 V to represent binary 0 and 1. What would this table correspond to in terms of a logic gate?

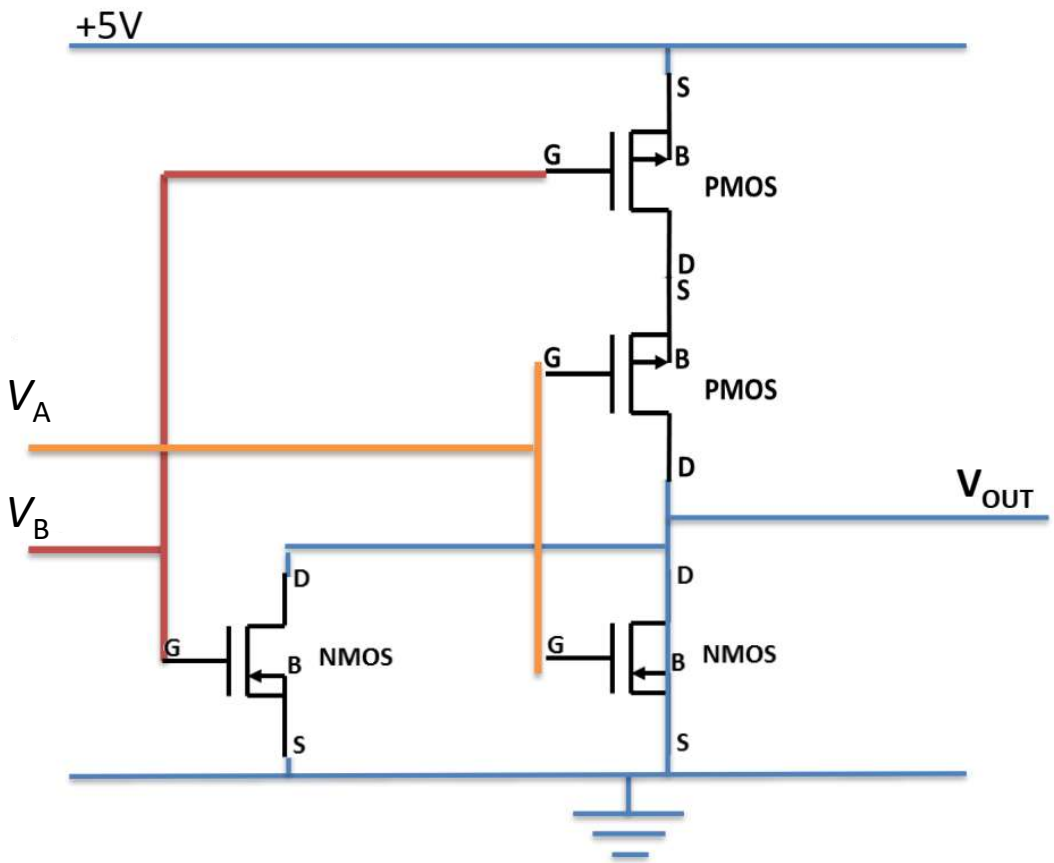
A Table of the Results

V_A (V)	V_B (V)	V_{OUT} (V)
0	0	
0	5	
5	0	
5	5	

- If we considered a binary system to have voltages of 0 and 5 V to represent binary 0 and 1. What would this table correspond to in terms of a logic gate?

This is a NAND gate.

Consider the Following Circuit



- What happens to V_{OUT} as V_A and V_B take either 0 or 5 V and in different combinations?

A Table of the Results

V_A (V)	V_B (V)	V_{OUT} (V)
0	0	
0	5	
5	0	
5	5	

- If we considered a binary system to have voltages of 0 and 5 V to represent binary 0 and 1. What would this table correspond to in terms of a logic gate?

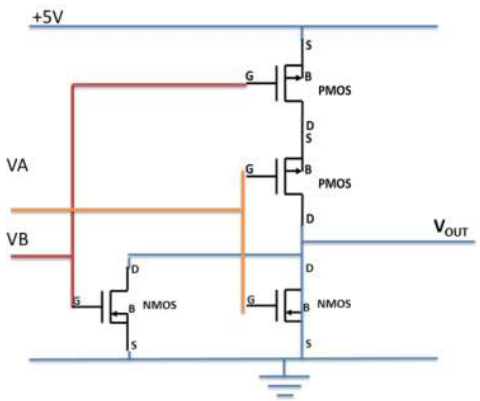
A Table of the Results

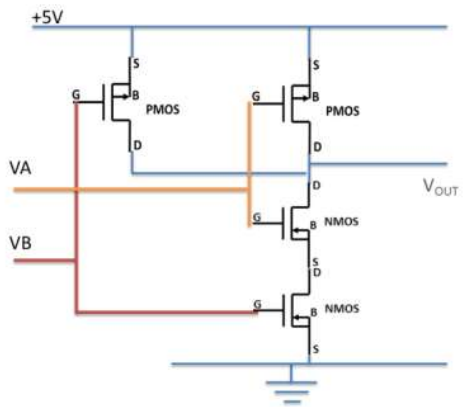
V_A (V)	V_B (V)	V_{OUT} (V)
0	0	
0	5	
5	0	
5	5	

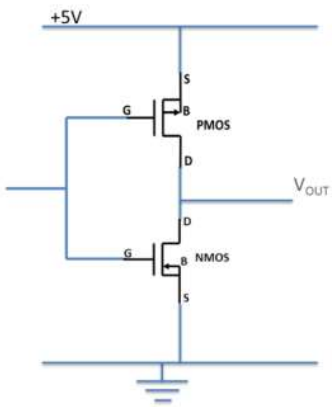
- If we considered a binary system to have voltages of 0 and 5 V to represent binary 0 and 1. What would this table correspond to in terms of a logic gate?

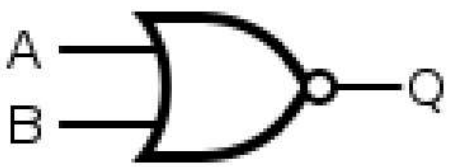
This is a NOR gate.

CMOS Logic

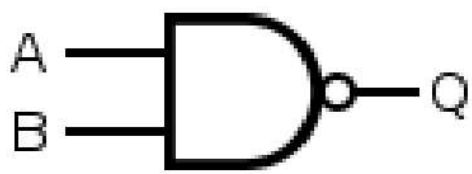




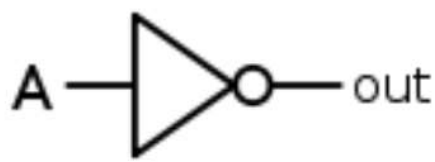




NOR



NAND



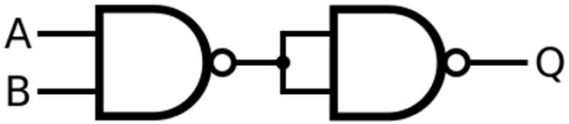
NOT (Inverter)

- These are the three fundamental logic gates built using CMOS transistors. Why the “C” in “CMOS”, because we must use a PMOS AND an NMOS to make them work.

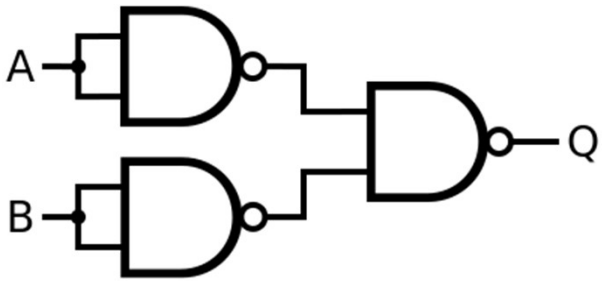
NAND & NOR Logics

- Both NAND and NOR gates have the property of functional completeness. They are both "universal gates";
- Any Boolean expression can be re-expressed by an equivalent expression utilising only NAND or NOR gate;
- Some examples using of NAND logic:

NOT



AND



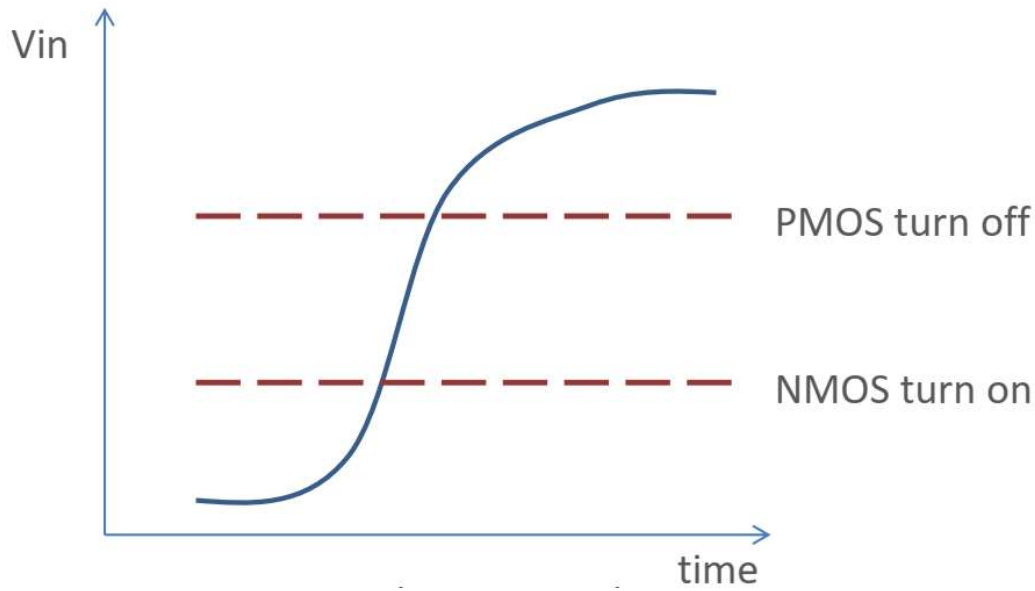
OR

Most solid-state drives (SSDs) use non-volatile NAND flash memory.

Problems & Benefits

- When looking at any logic gate implementation, we need to figure out a couple of things first:
  - How fast can it run?
  - How much power does it use?
- The answers for CMOS are quite good in both cases but not perfect and we need to touch on these.

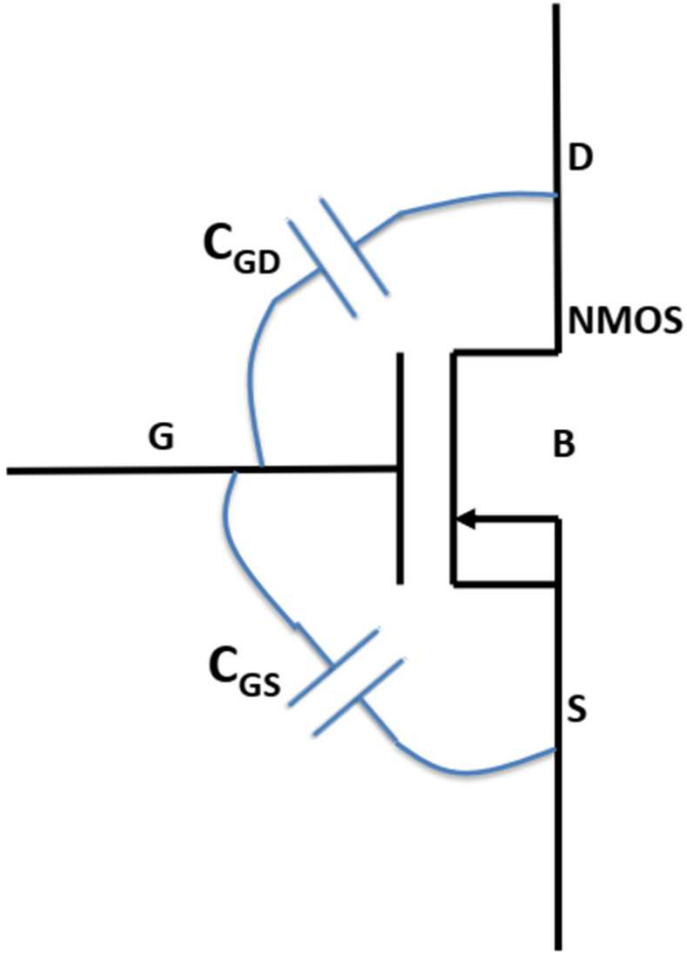
Starting and Falling Time Delays



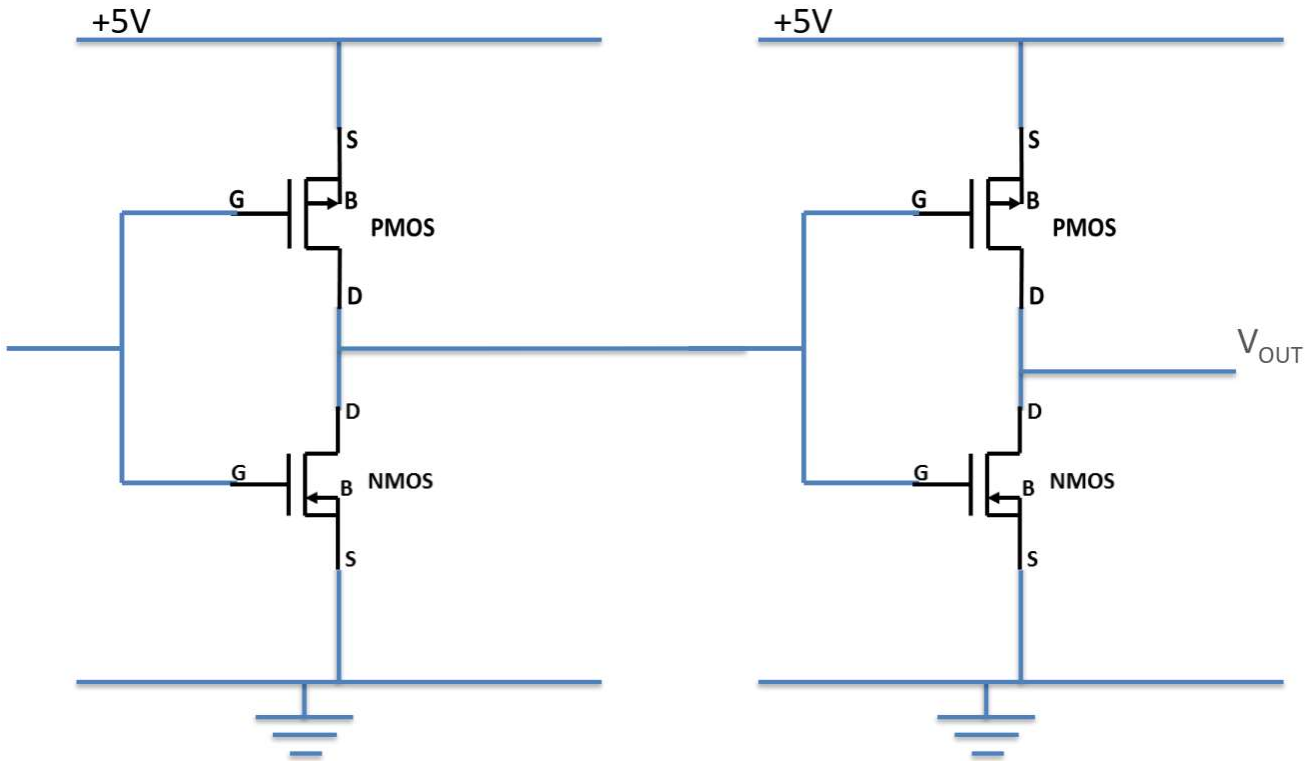
- Imagine we send a square wave (zoomed in) into an inverter, and we see the voltage ramping up from ZERO to 5 V. Clearly the transistors change state when the inversion region appears and disappears;
- It is going to take some time for the input signal to go from ZERO to $V_{T(NMOS)}$ and then some more time for it to get to (5 - $V_{T(PMOS)}$);
- When the NMOS turns on, the output voltage starts dropping, but it only gets properly low when the PMOS turns off.

Those Darn Parasitics

- The problem is mostly because of the parasitic capacitances. They take time to charge and discharge;
- To charge/discharge a capacitor you need to add or remove charge (through current). The bigger the capacitance, the more charge you need to add or remove;
- If you want it to change faster, you need to be able to push a lot of current in or out of the GATE during changeover. It is only for a short period of time.

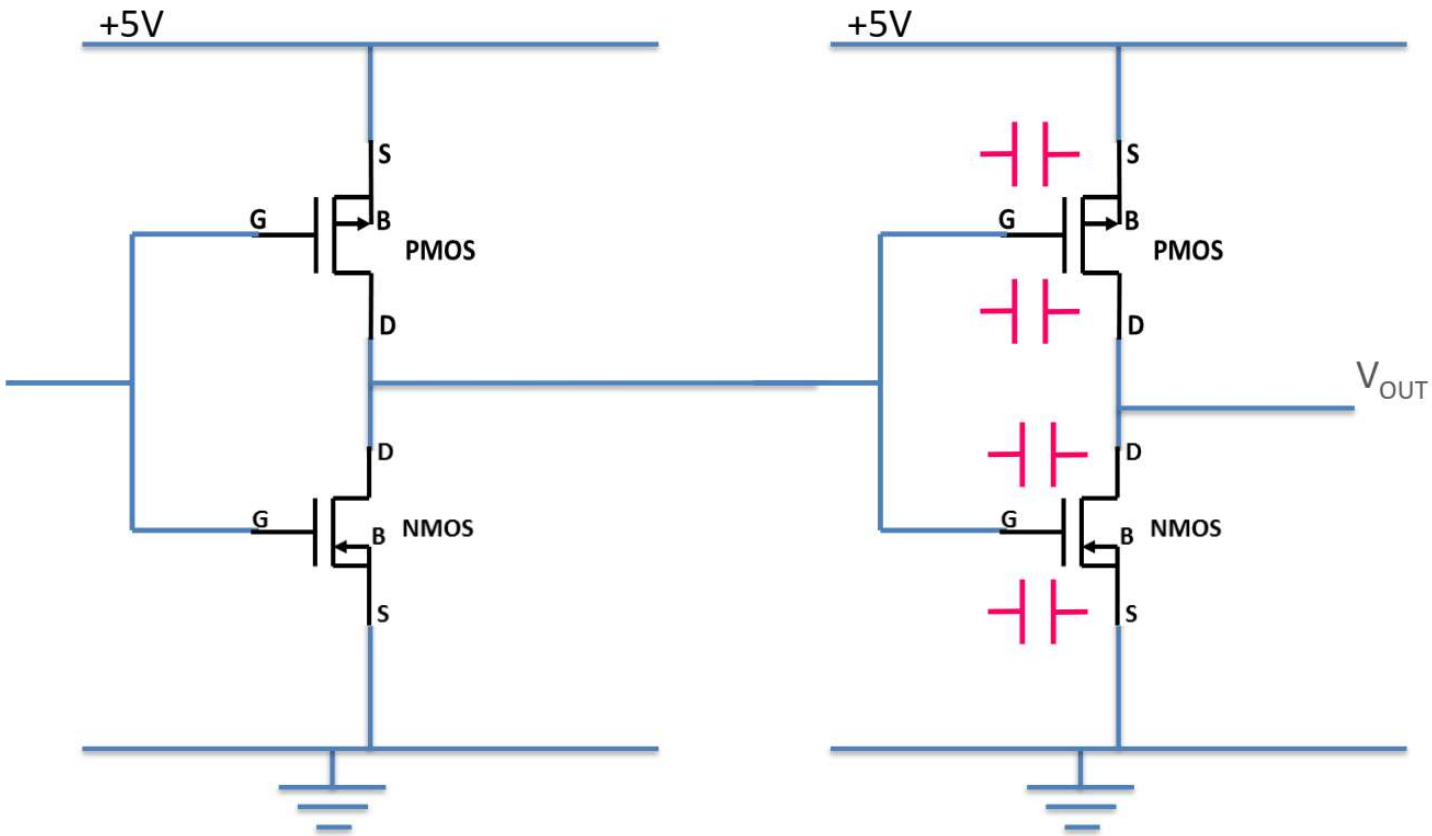


Those Darn Parasitics



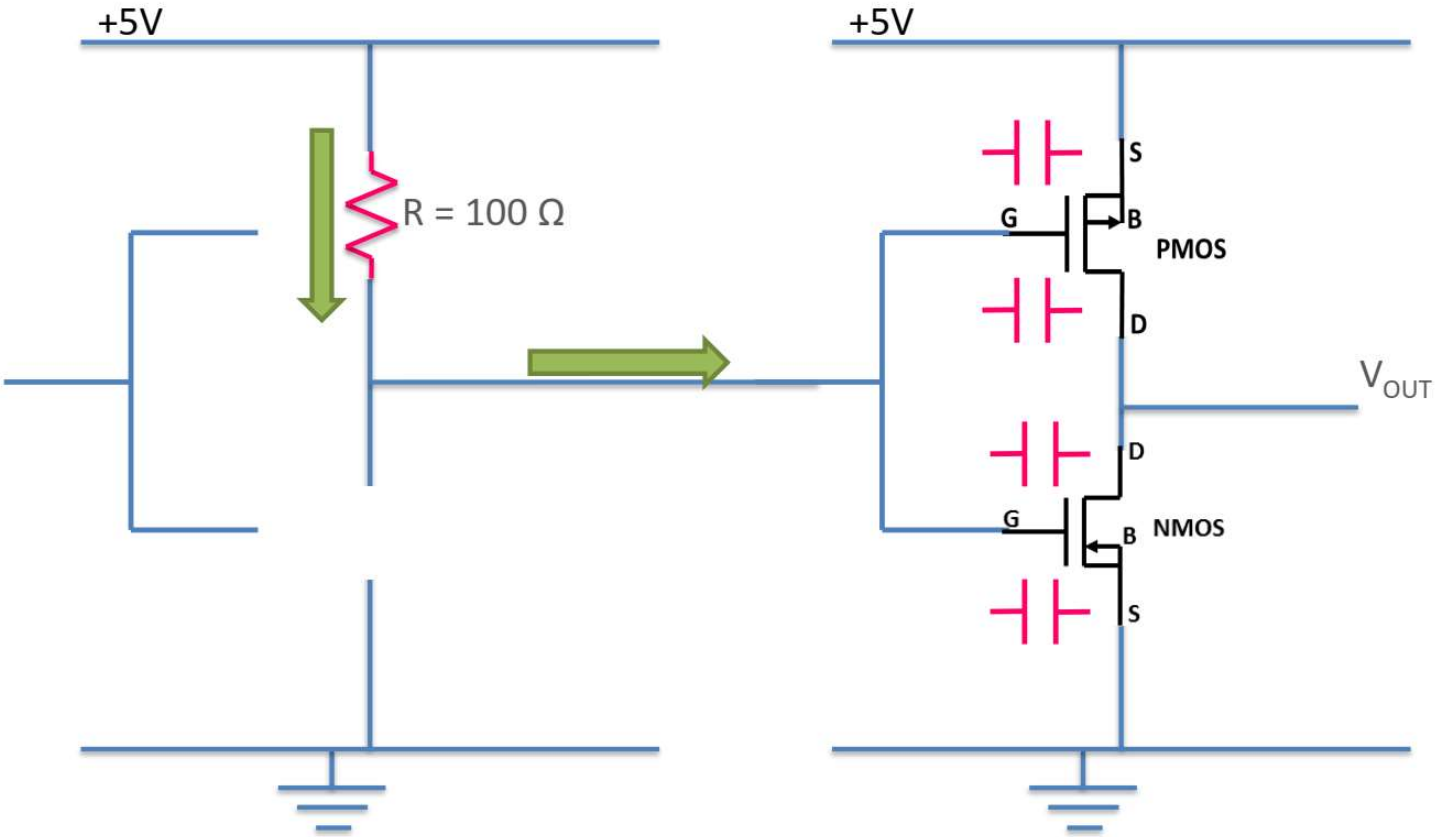
- We have two inverters connected together, but let's just look at the middle.

Those Darn Parasitics



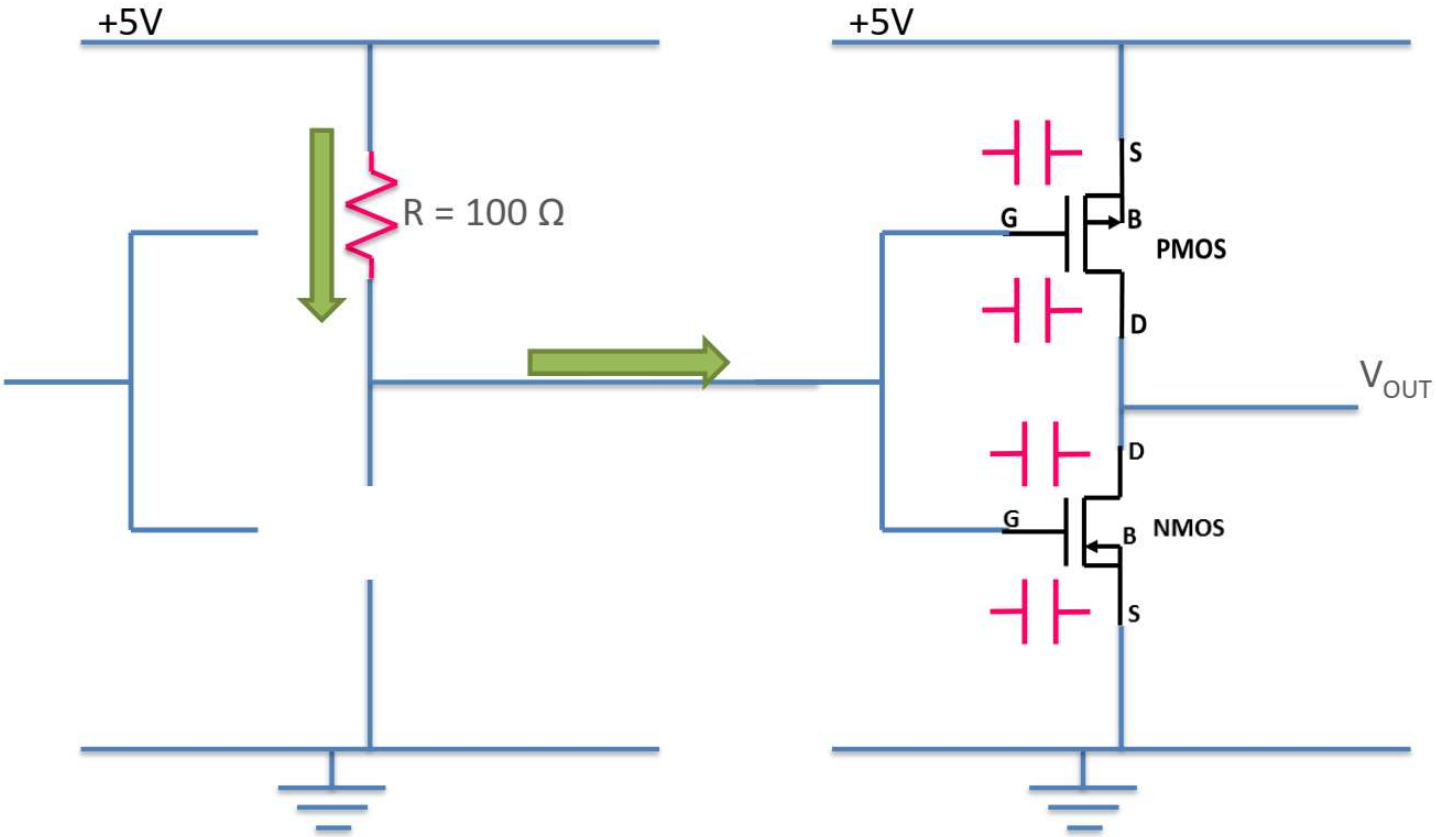
- We have an output from the middle of the other inverter, connecting to a gate which has 4 parasitic capacitors.

Those Darn Parasitics



- Now the first inverter has switched, and we have the NMOS off and the PMOS on. The current through the PMOS low resistance flows into the gate of the next inverter, charging/discharging the capacitors. The speed that we can change the voltage is proportional to *RC* time constants.

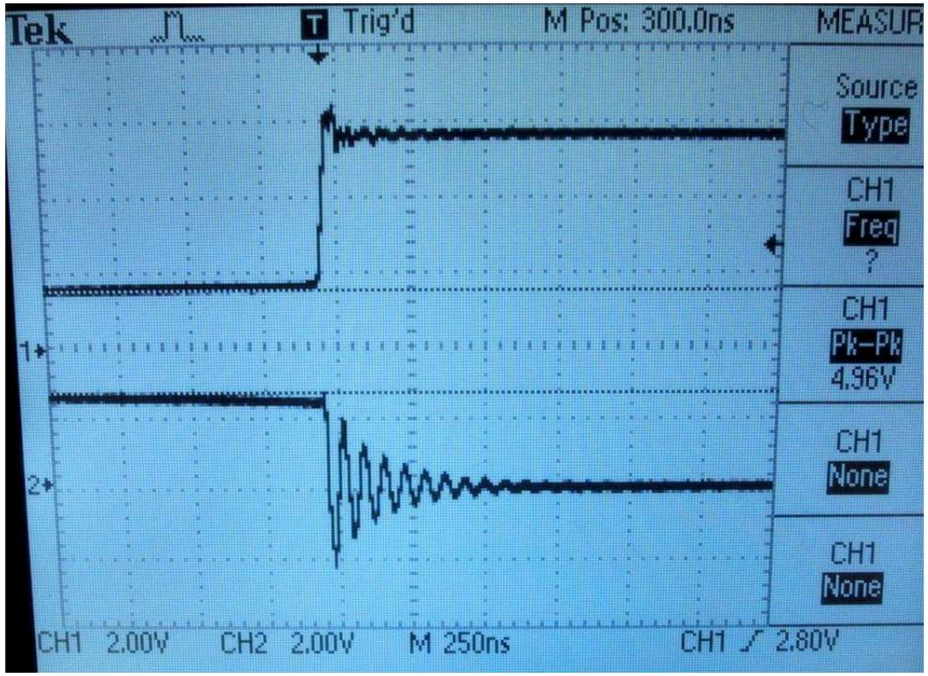
Those Darn Parasitics



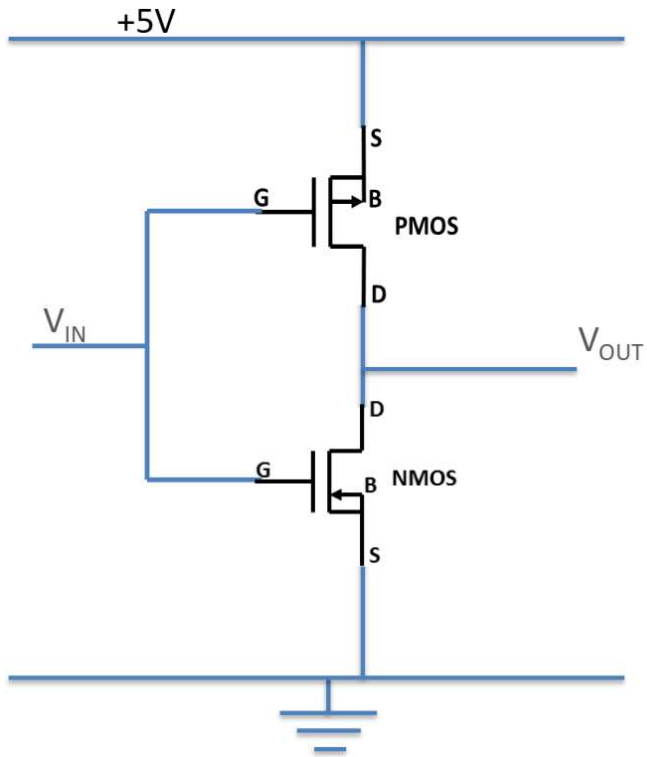
- Consider a transistor with the following values: C_{GS} = 27 pF, C_{GD} = 3 pF (which is very big), and R = 100 ohm (small R, big current);
- $\tau \approx$ 6 ns (or ~200 MHz).

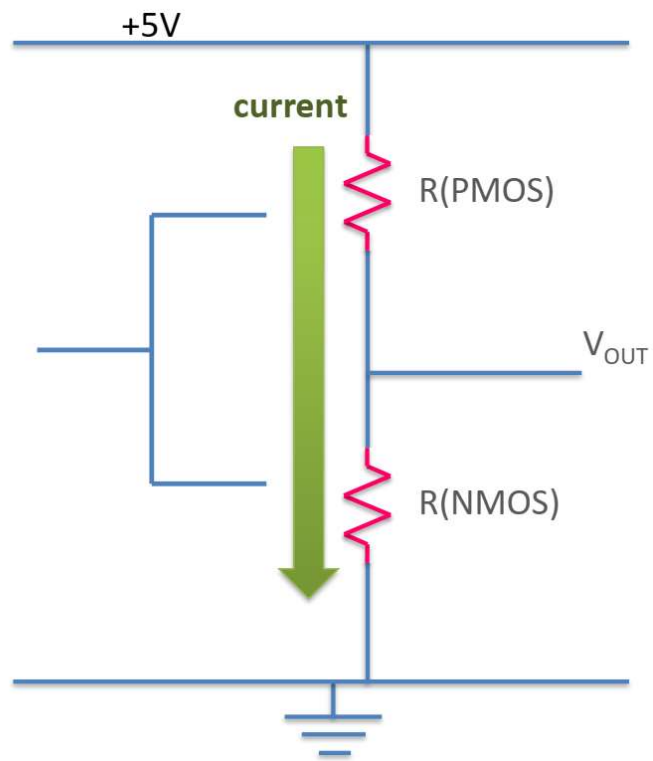
Even More Darn Parasitics

- The other thing that happens is that we meet up with some parasitic inductors (from the wires). *RLC* combinations produce oscillations which we can see on the oscilloscope as **RINGING**. It fades away but can be big. So we can't look at the output until a little time after the switch.



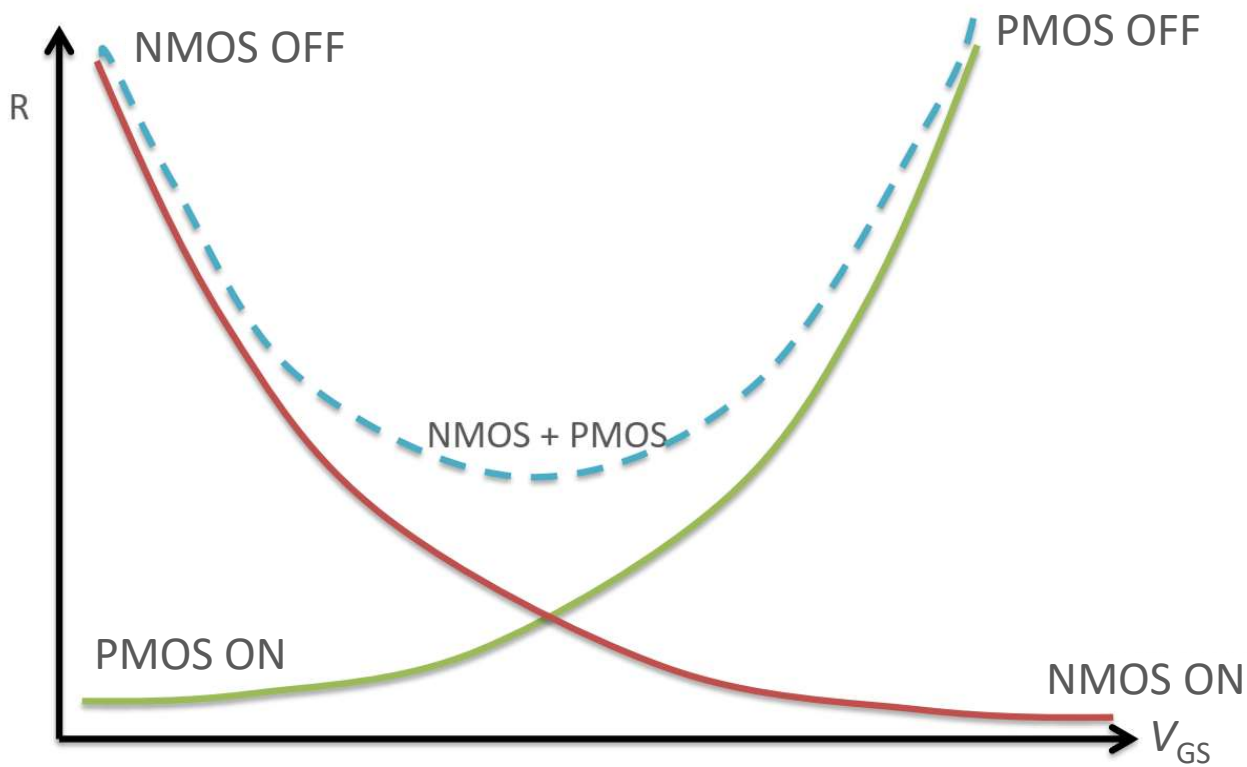
Current Flows





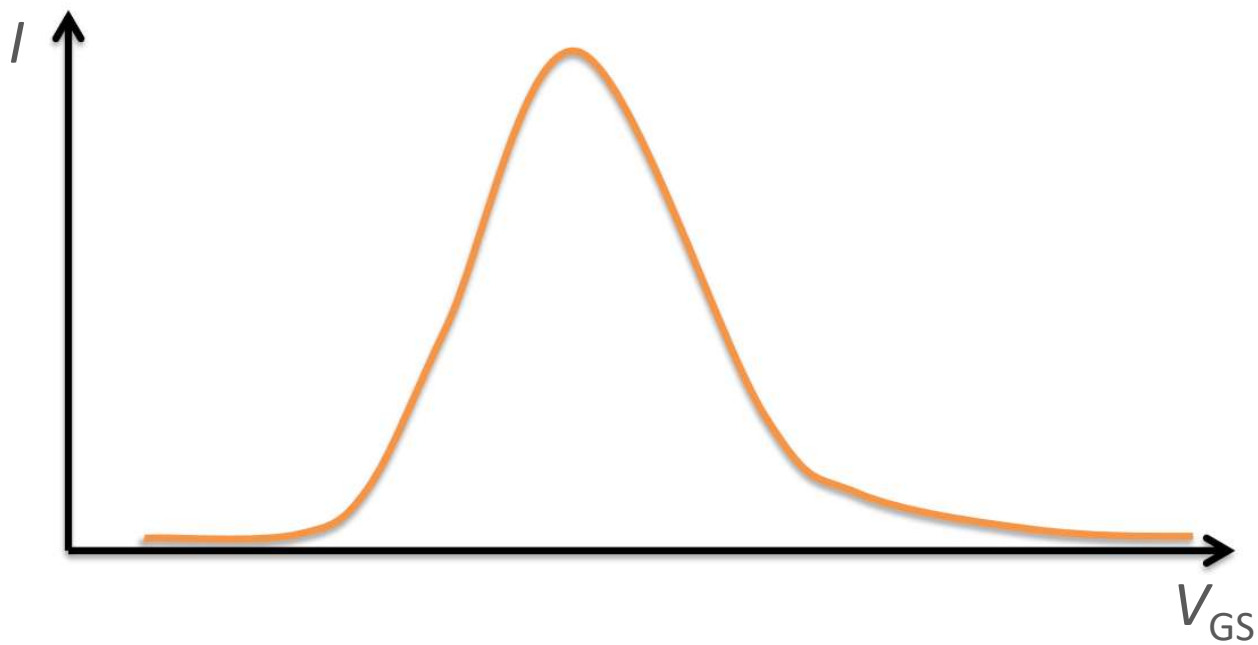
- The basic equation is still ***P = VI***.
  The voltage is fixed but the resistance will change as we turn transistors on and off. The current depends on both resistors.

The Resistance of the Transistors



- When we are at either extremes of V_{GS}, the resistance is very high, however, in the middle the resistance could be "low", say 1 – 10 kΩ.

The Current Therefore



- The current is very small at either extremes, but it is much bigger in the middle. Not big, but still noticeable.
- There is no current at either extreme values of V_{GS}, so with no current, there is no power consumed.

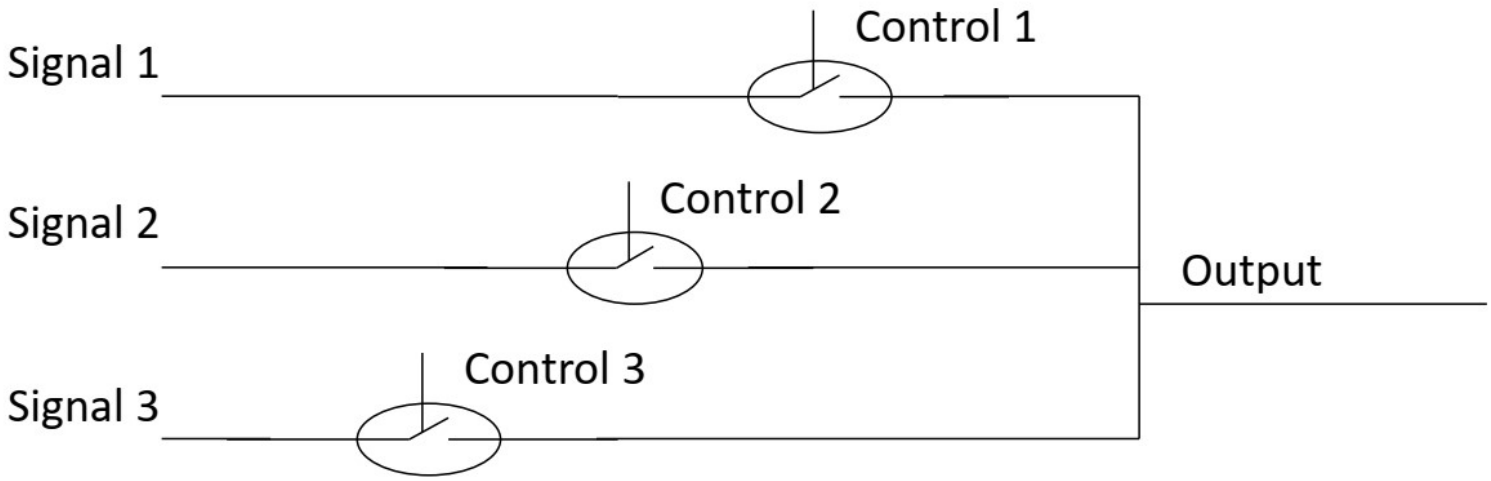
What is Important

- Do NOT set the gate voltage for an inverter to be in the middle, it will consume a lot of power and will eventually burn the transistors;
- Power is only consumed when we have mid-range values for V_{GS}. This means we only spend power in a CMOS logic gate when we are changing values. There is no power consumed when we are at either extremes;
- Move fast from one extreme to the other, the less time we spend in the middle, the less power we will consume.

- Our choices are not straightforward, it will depend, but we can
  - Have less capacitance on our output (reducing *RC* speeds up switching);
  - Decrease *R* so we can switch faster (bigger spikes of current, more power);
  - Increase *R* so we have lower current spikes when we are changing (slower).

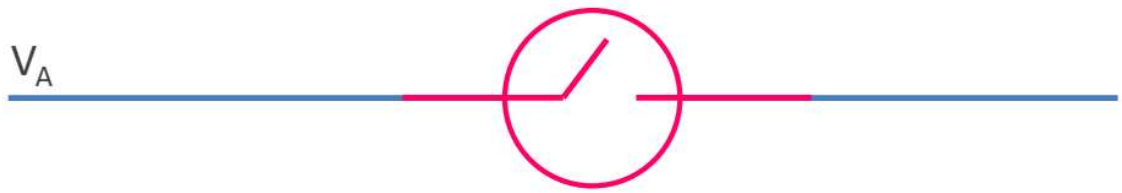
Signal Gates or Switches

- In analog circuitry, it is always very useful to be able to disconnect signals and reconnect them.
- For example, the multiplexor (mux).

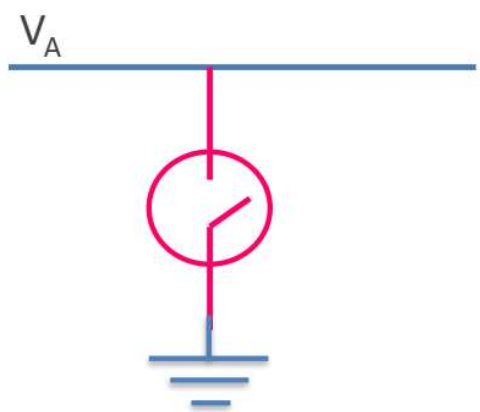


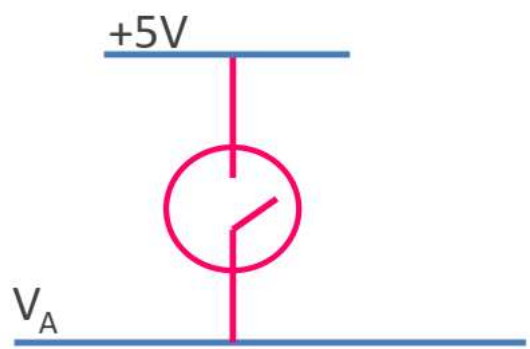
Types of Switches

- The make/break switch:



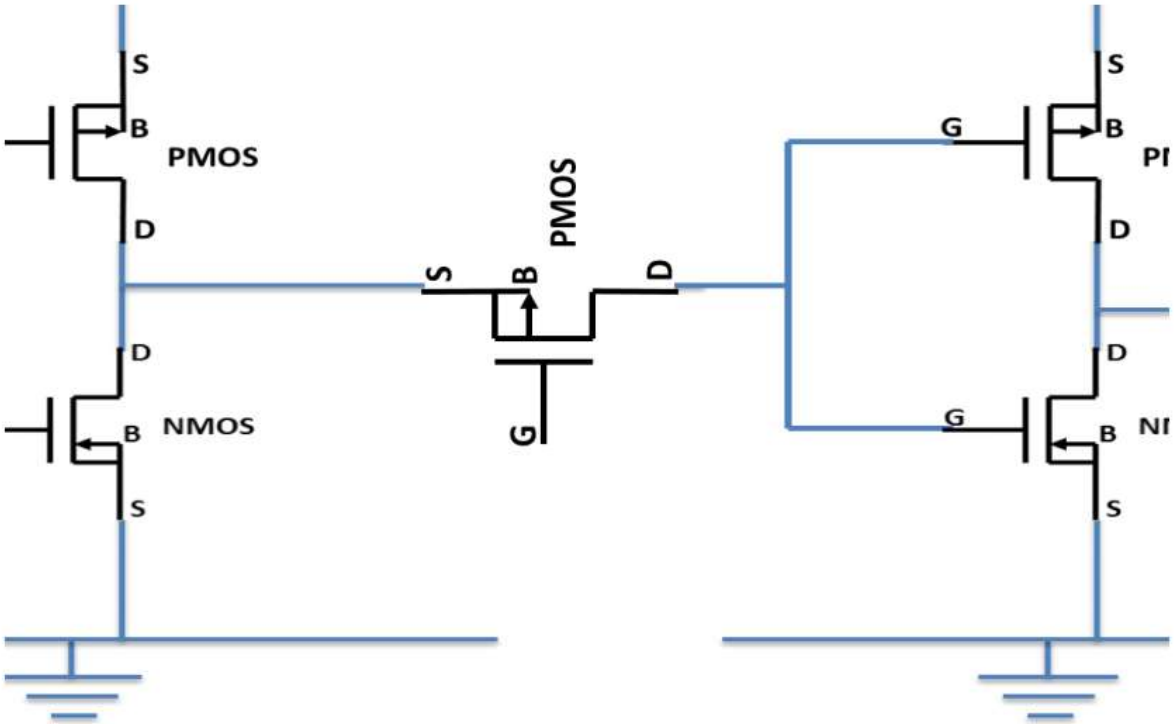
- The pull-down/pull-up switch:





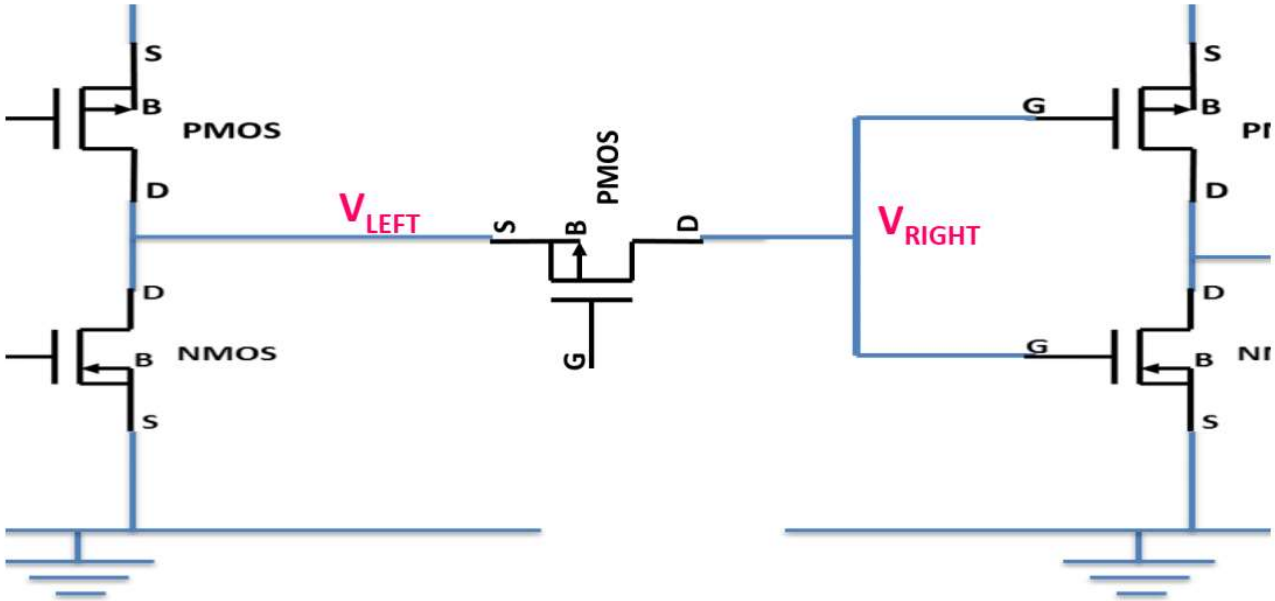
Make/Break Switches

- What we need is a HIGH IMPEDANCE when the switch is open and a LOW IMPEDANCE when the switch is closed.
- We could use a PMOS or an NMOS to achieve this.



Make/Break Switches

- What voltage on the GATE do I need to turn this switch ON (closed) if the output of the LEFT inverter is 5 V?



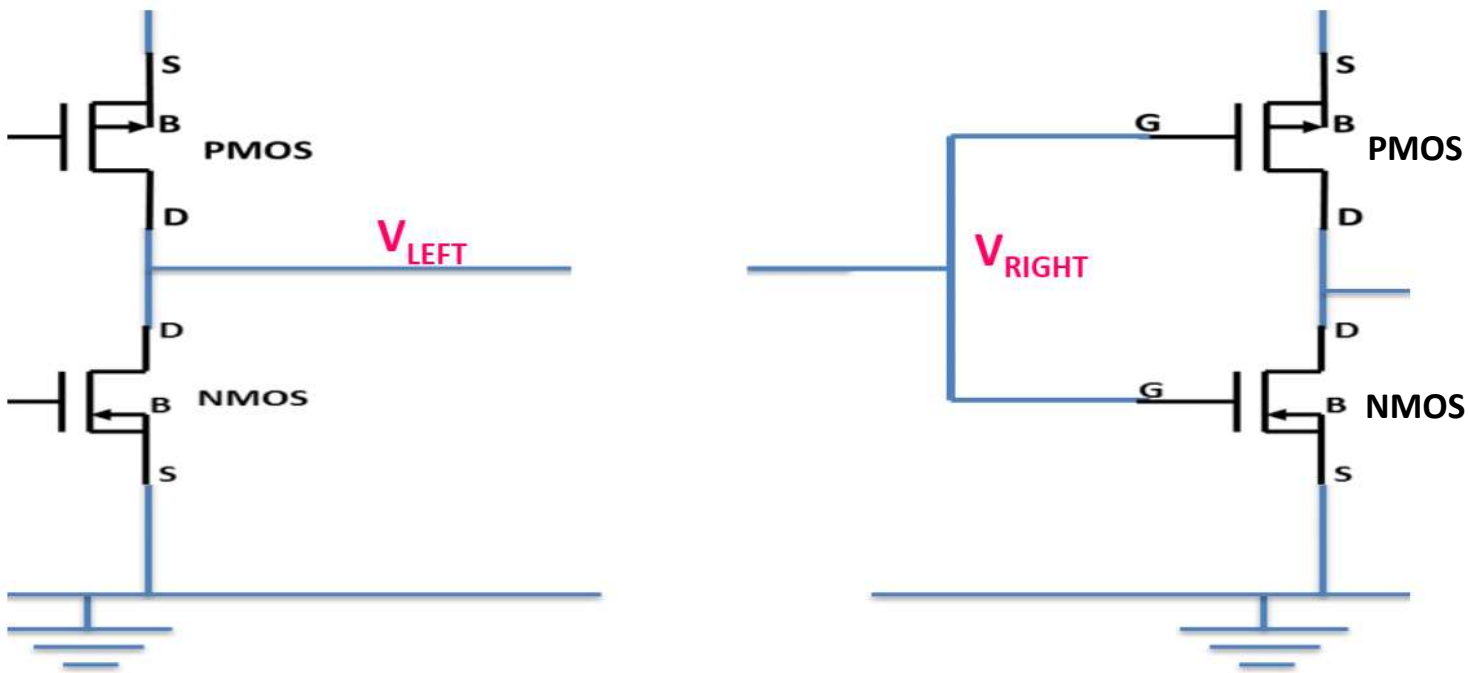
- If the output of the inverter (LEFT) is now zero, what voltage do I need to ensure that the switch is CLOSED?

First Problem

- Basically, if the input goes too low, the PMOS will begin to turn off even when we have the gate voltage at its smallest value;
- The reverse also happens. If we have an NMOS, if the input goes too high, we can never set the GATE high enough to get the inversion region to form;
- If we get even close to the threshold voltage, we're going to find the resistance beginning to climb really quickly.

Make/Break Switches

- We have made a very good disconnect. What voltage do we now have on the gate for the second inverter?



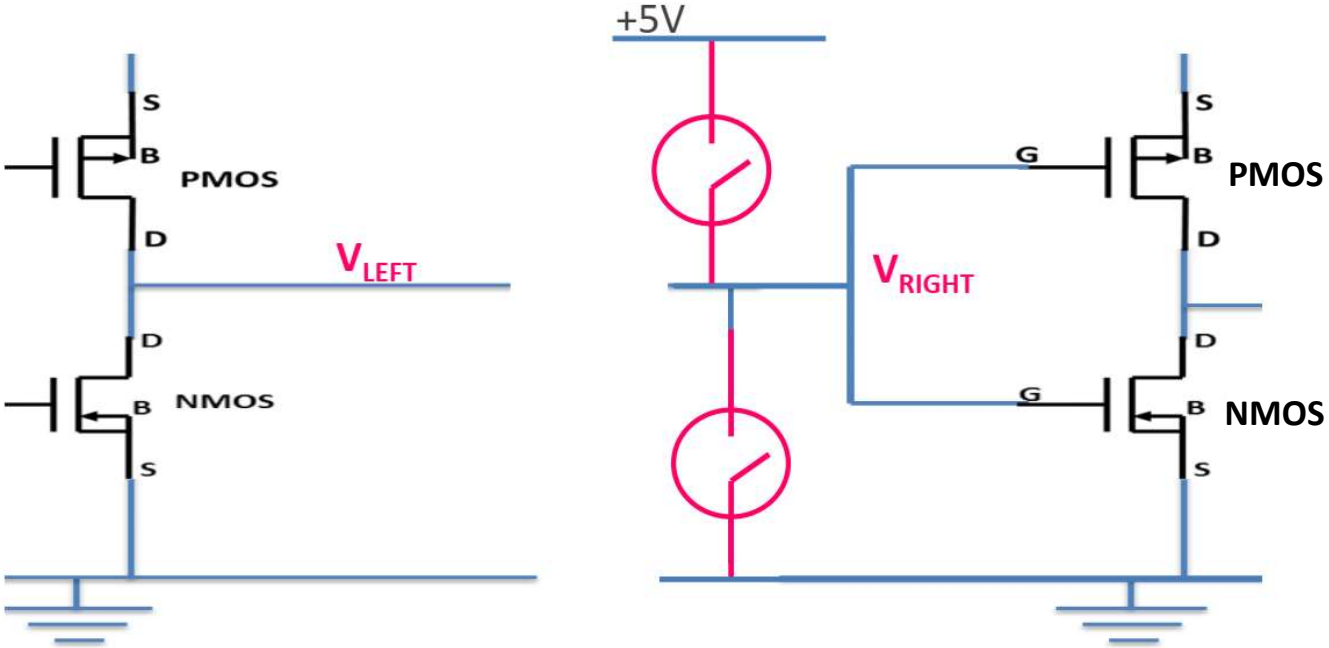
- Are you really sure that you know the answer to this one?

Second Problem

- A good switch means that the right-hand side is now FLOATING;
- Floating means that it could be any voltage and has a likely tendency to drift;
- In practice it will pick up charge from somewhere, radio emissions, parasitic capacitances, and will end up taking a voltage. As it has no DC path for current to flow, the charge has nowhere to go, so even a tiny charge can hold a voltage;
- Enough to trigger a MOS transistor;
- The one thing you do know for sure, IT IS NOT LIKELY TO BE ZERO. Unless you set it to something, it could be anything.

Pull-Ups & Pull-Downs (2nd Problem)

- When we have floating wires, we like to set them to a known voltage. Normally we connect them to ground (pull-down) or we connect them to the max voltage (pull-up).

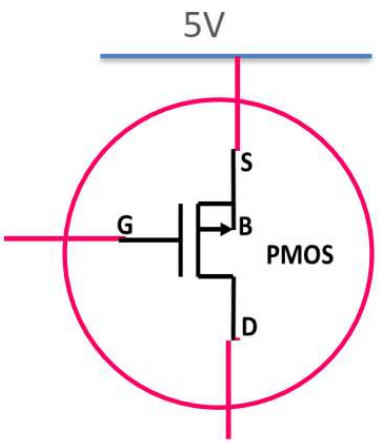


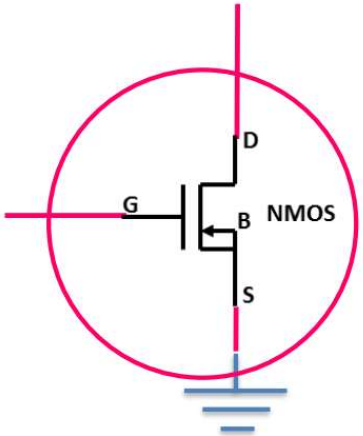
- Depending on which switch is closed, V_{RIGHT} is either shorted to GROUND or to 5V.

Pull-Ups & Pull-Downs (2nd Problem)

To make a pull-up, we tend to use a PMOS. Why? Because we need to know V_{GS}, and if the source is floating, we're confused. In a PMOS, the source is connected to V_{DD}.

To make a pull-down, we tend to use an NMOS. Why? Because we need to know V_{GS}, and in an NMOS, the source is connected to GROUND. No confusion.

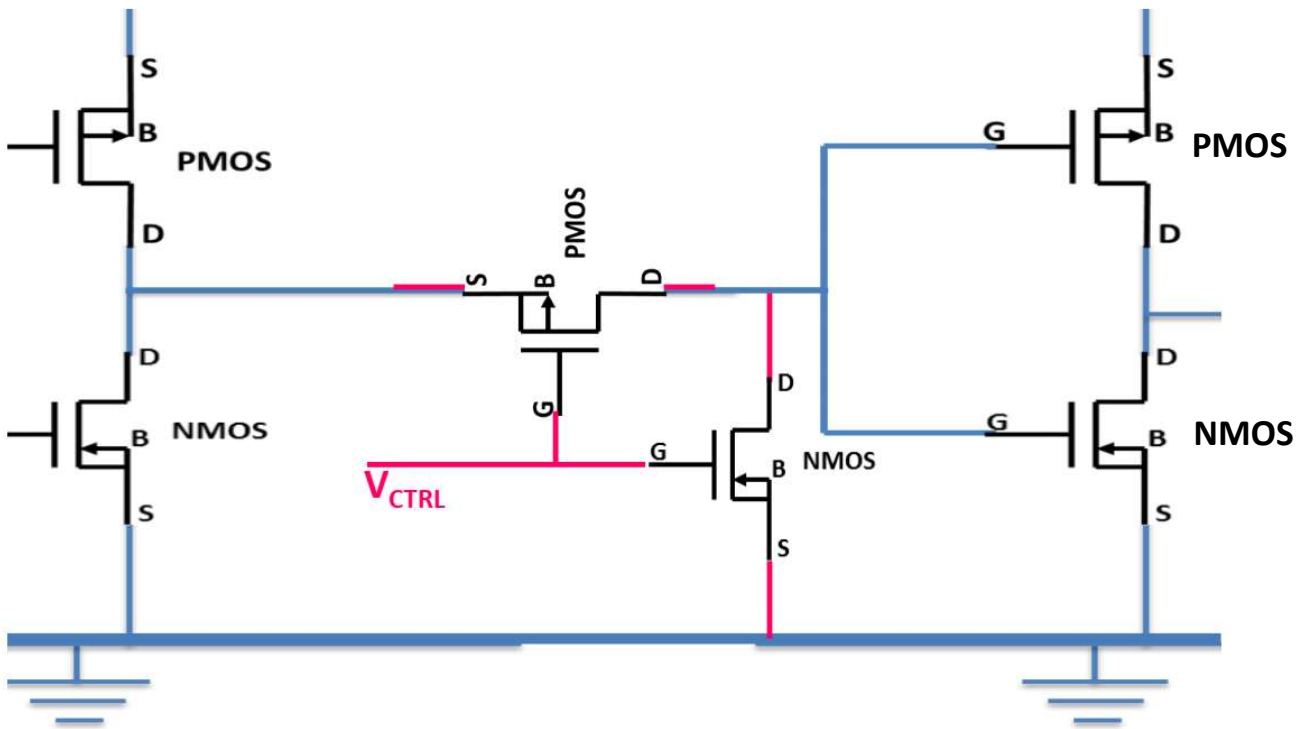




- In this scenario, does the ON resistance of the switch matter?

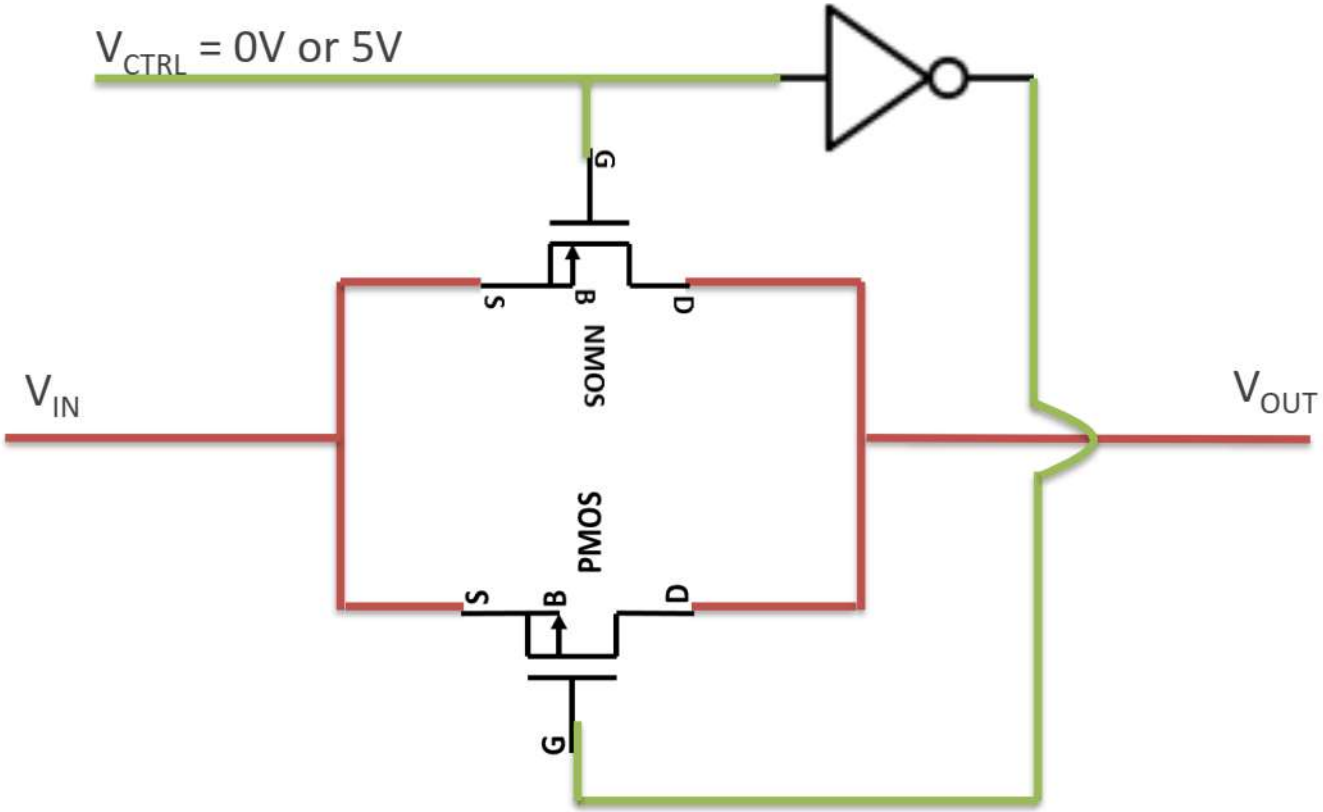
The Final Make/Break Solution

- When we have floating wires, we like to set them to a known voltage. Normally we connect them to ground (pull-down) or we connect them to the max voltage (pull-up).



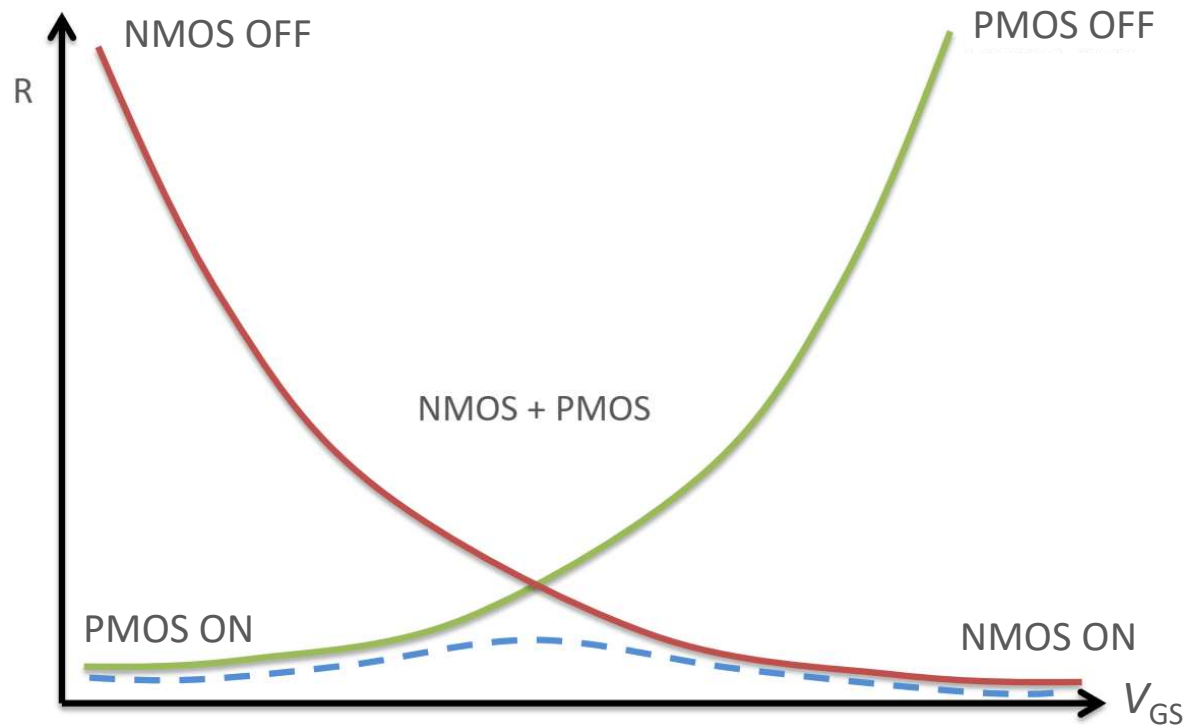
Two Parallel Switches

- See if you can figure how this circuit works:



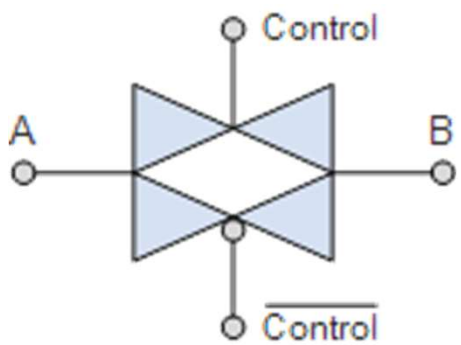
Back to the 1st Problem: A Good Switch

- Here we have two transistors in parallel. For V_{CTRL} = 5 V, when one MOSFET is turning off, the other is turning on. Parallel resistances are always smaller than the smallest of the individual resistances, so we get something like below FOR ALL VALUES OF INPUT VOLTAGE.



A Transmission Gate

- This combination of transistor switches and an inverter is called a transmission gate and it has its own symbol:



Control	A	B
1	0	0
1	1	1
0	0	High-Z
0	1	High-Z

- When drawn like this, it is understood that when the control voltage goes HIGH, the switch is CLOSED (signal at terminal A is transmitted to terminal B).

Design Issues

- Threshold Voltage

Transmission gates work best when the threshold voltage is not in the centre of the range (e.g., 2.5 V for a 5 V range) because you want one transistor definitely ON when the other is turning off. This is a minor issue.

- ON Resistance

The on resistance of the transmission gate can be made very low. This is useful as we do not wish to add more resistance to a circuit, as all nodes have some capacitance and the charging up time is based on *RC*.

How low does it need to go? That depends on the application, the speed needed to charge/discharge capacitive loads, etc. Sometimes it's ok for it to be quite high. As the Gate is only ever ZERO or MAX VOLTAGE, the only way to control the resistance is *W*/*L*. Or in other words, big devices = low resistance. Big fast devices cost more to make. It's a trade-off.

Design Issues

- Clock Feedthrough

There is an issue with something called clock feedthrough which is based upon capacitances and transistor speed. It is only a significant issue at high frequencies. Basically, what happens is that the junction capacitances charge and discharge at different rates and hence the voltages may change at different rates.

A good solution is to keep the sizes of the devices equal so that capacitances remain approximately the same. For highly sensitive systems, further factors can also be considered.

This is not a major issue for this course.

Example 1: Switch Design

- Design a switch capable of being driven by a 0/2.5 V logic signal that can switch an input signal that ranges from 0.3 to 2.2 V. The effective resistance of the switch when open should exceed 100 MΩ and when closed should present no more than 25 Ω.

PMOS	NMOS
V_T = -0.8 V	V_T = 0.8 V
$\mu_p C_{ox}$ = 30 μA/V^2	$\mu_n C_{ox}$ = 90 μA/V^2

- Minimum device length = 0.18 μm.

Example 1: Switch Design

- First: The voltage range of the input is such that it goes below (0 + V_T) for the NMOS and (2.5 - V_T) for the PMOS. A single transistor switch would be unsuitable for the task. The transistor would be turned off by the input signal. Therefore, we must use a transmission gate.
- Second: the off-resistance of a transmission gate is the parallel resistance of two devices with no inversion region, hence two reverse bias diodes. It's going to be big. So let's say the off condition has been satisfied.

Example 1: Switch Design

- Third: The ON resistance is the parallel combination of two resistances. Considering the curves of the transistor resistances, we have two situations
  - The extremes, one very large and one very small resistance.
  - The centre, two small but significant resistances in parallel.
- In calculating these resistances, assume the device is in the triode region (nearly all MOSFET switches operate in the triode region). For example, a very low resistance ideally (< 20 ohm) and a small current of 10 mA (which is a large current for MOSFETs) lead to a V_{DS} of <200 mV.

$$R_{\mathrm{DS(Triode)}} = \frac{1}{\mu_{\mathrm{n}} C_{\mathrm{ox}} \frac{W}{L} (V_{\mathrm{GS}} - V_{\mathrm{T}})}$$

Example 1: Switch Design

- Pick the extreme, from the equation, the PMOS will have a bigger resistance due to the smaller mobility ($\mu_p C_{ox}$ = 30 μA/V^2). If we satisfy the PMOS-on extreme, then the NMOS-on side will be satisfied, too.

$$R_{\mathrm{DS(Triode)}} = \frac{1}{\mu_{\mathrm{p}} C_{\mathrm{ox}} \frac{W}{L} (V_{\mathrm{SG}} - |V_{\mathrm{T}}|)} < 20\ \Omega$$

- Now at the extreme, V_{signal} = 2.2 V, V_G = 0 V, V_T = -0.8 V so

$$V_{\mathrm{SG}} - |V_{\mathrm{T}}| = (2.2\ \mathrm{V} - 0\ \mathrm{V}) - |-0.8\ \mathrm{V}| = 1.4\ \mathrm{V}$$

- Therefore,

$$\frac{L}{W(30\times10^{-6})(1.4)} < 20 \text{ or } \frac{W}{L} > 1190$$

Example 1: Switch Design

- Pick the centre, here we need to calculate both the NMOS and the PMOS transistor resistances.
- In the centre, V_{signal} = 1.25 V, so
  - NMOS: $V_{\text{GS}} - V_{\text{T}}$ = (2.5 V – 1.25 V) – (0.8 V) = 0.45 V;
  - PMOS: $V_{\text{SG}} - |V_{\text{T}}|$ = (1.25 V – 0 V) - |-0.8 V| = 0.45 V.
- For this example, it just happens to be the same for P and N.
- Therefore,

$$R_{\text{PMOS}} = \frac{L}{W(30 \times 10^{-6})(0.45)} = 74074(\frac{L}{W})$$

$$R_{\text{NMOS}} = \frac{L}{W(90 \times 10^{-6})(0.45)} = 24691(\frac{L}{W})$$

Example 1: Switch Design

- When V_{signal} is at the centre, the total resistance of the switch is the parallel resistance of the two devices, so

$$R_{\text{parallel}} = \frac{1}{\frac{1}{R_{\text{PMOS}}} + \frac{1}{R_{\text{NMOS}}}} < 20$$

$$\therefore \frac{1}{\frac{W}{74075L} + \frac{W}{24691L}} < 20$$

$$\therefore \frac{L}{W}(18518) < 20$$

$$\frac{W}{L} > 926$$

Example 1: Switch Design

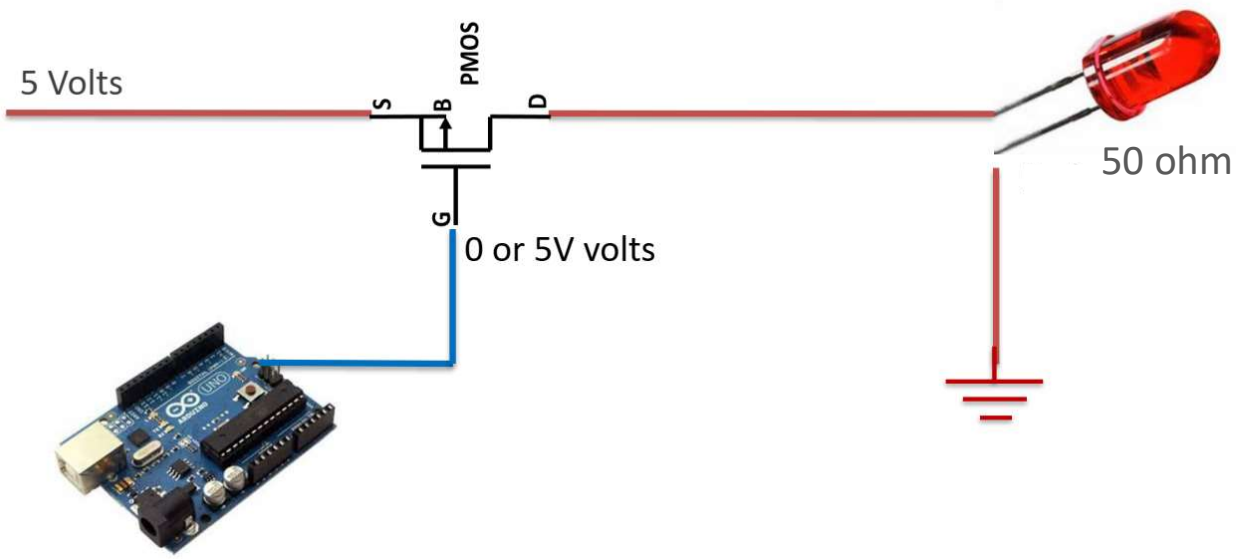
- Now we have two constraints for us to satisfy our maximum resistance requirements:
  - Centre: $W/L > 926$;
  - PMOS extreme: $W/L > 1190$.
- The PMOS extreme is the tougher one, so set W/L to be 1250. It should be safe enough.
- Use minimum length for L, giving

$$L = 0.18\ \mu\text{m and } W = 225\ \mu\text{m}$$

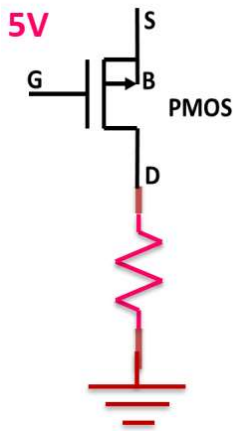
  This is a BIG device.
- Finally, we did not factor in temperature or variations in process (mobility), which would require us to make the device even bigger and test more extremes and corners.

Example 2: Switching On/Off an LED

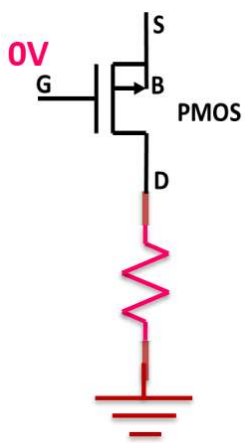
- I have an Arduino that has a pin that can go from 0 to 5 V as I change it from digital ZERO to digital ONE. I want to turn on an LED powered by a 5 V battery. In the following circuit, calculate the power lost in the transistor switch if the LED has an effective resistance of 50 ohm. If I leave the switch on (digital ONE), how much power is used in the LED? If I leave the switch off 50% of the time, how much power is used in the LED?



Example 2: Switching On/Off an LED



- Let's look at this, $V_{GS} = V_G - V_S$. For the transistor to work, the magnitude of $|V_{GS}| > |V_T|$ (negative for PMOS), so that means that V_{GS} will be equal to zero when V_G is equal to 5. There is no way it can be less than zero if the gate is at 5 Volts.



- Let's look at this, $V_{GS} = V_G - V_S$. If $V_G = 0$ and $V_S = 5$ V, then we have $V_{GS} = -5$ V, definitely ON. The PMOS switch should always be closed with the smallest resistance. The PMOS is likely in the triode region (LED resistance is significant).

Example 2: Switching On/Off an LED

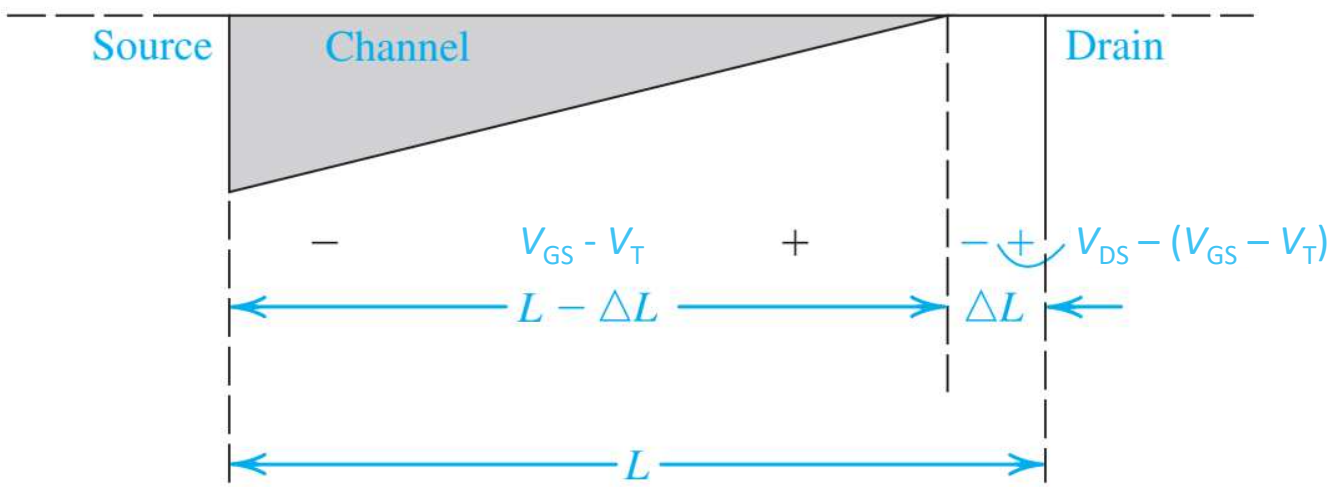
$$R_{\mathrm{DS(Triode)}} = \frac{1}{\mu_{\mathrm{p}} C_{\mathrm{ox}} \frac{W}{L} (V_{\mathrm{SG}} - |V_{\mathrm{T}}|)}$$

- Let's assume that for the PMOS:
  - $\mu_p C_{ox}$ = 30 μA/V^2;
  - W/L = 1000;
  - V_T = -0.8 V.
- The MOSFET resistance is

$$R_{\mathrm{DS(Triode)}} = \frac{1}{(30 \times 10^{-6})(1000)(5 - 0.8)} = 7.94\ \Omega$$

- If the LED has a resistance of 50 ohm, about 14% of the voltage (power) would be lost in the transistor WHEN ON, and 86% in the LED.
- When OFF, no current flows, no power lost.

A Word on Channel-Length Modulation



- In the saturation region, when V_{DS} further increases, the channel pinch-off point moves slightly away from the drain and towards the source;
- The effective channel length becomes $(L - \Delta L)$;
- The electric potential difference across the channel remains constant at $(V_{GS} - V_T)$. This is sometimes called the "overdrive voltage".

A Word on Channel-Length Modulation

- For small ΔL, $I_D \propto \frac{1}{L-\Delta L} = \frac{1}{L}(1+\frac{\Delta L}{L})$;
- We assume that $\frac{\Delta L}{L} = \lambda V_{DS}$ in which λ is the channel-length modulation parameter;
- The MOSFET equation in the saturation region is

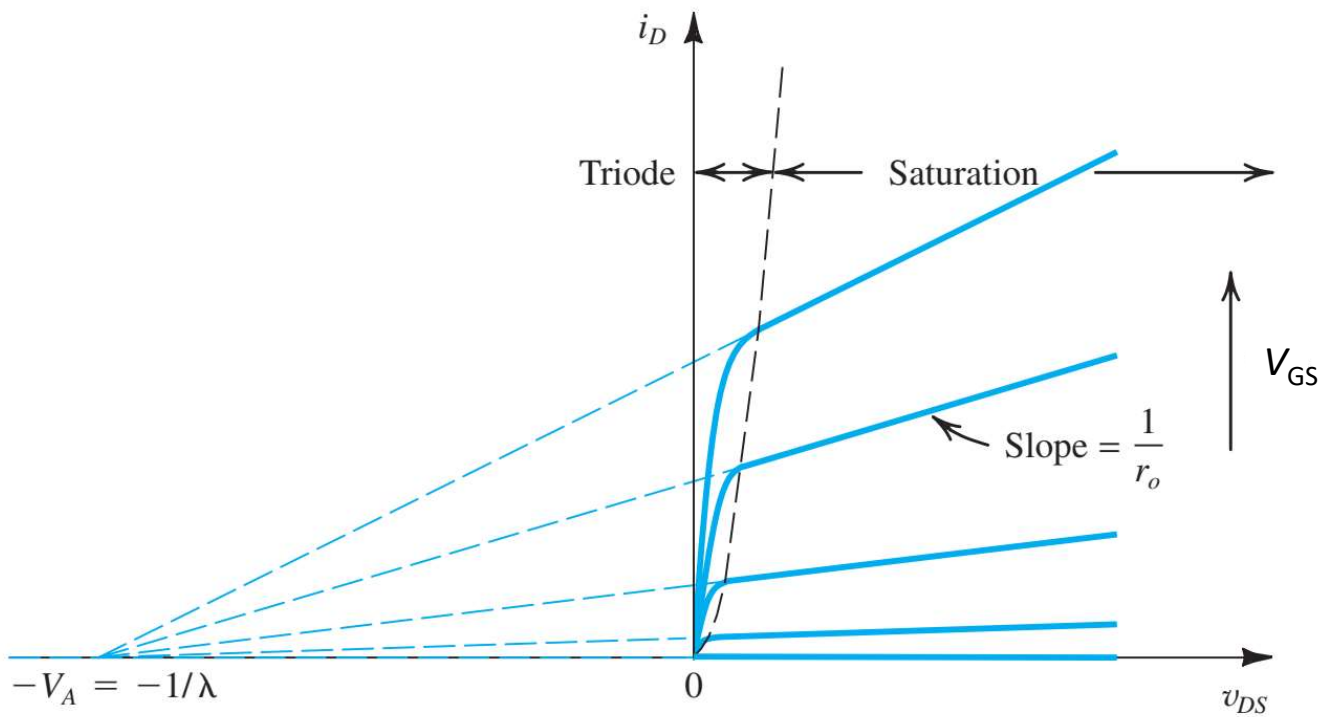
$$I_{\mathrm{D(sat)}} = \frac{1}{2}\mu_{\mathrm{n}} C_{\mathrm{ox}} \frac{W}{L}(V_{\mathrm{GS}} - V_{\mathrm{T}})^2(1+\lambda V_{\mathrm{DS}})$$

- The drain current is not a constant anymore but increases with V_{DS}. The output resistance of the MOSFET is not infinity anymore but

$$r_o = (\frac{\partial I_D}{\partial V_{DS}})^{-1} = \frac{1}{\lambda I_D}$$

and I_{D} is the drain current without channel-length modulation being taken into account.

A Word on Channel-Length Modulation



- If the I_D vs. V_{DS} curve in the saturation region is extrapolated, it intercepts the V_{DS} axis at a point $V_{DS} = -V_A$. V_A (a positive value) is referred to as the Early voltage (after Jim Early).

EE204: Analog Electronics

Models and Biasing

What is a model?

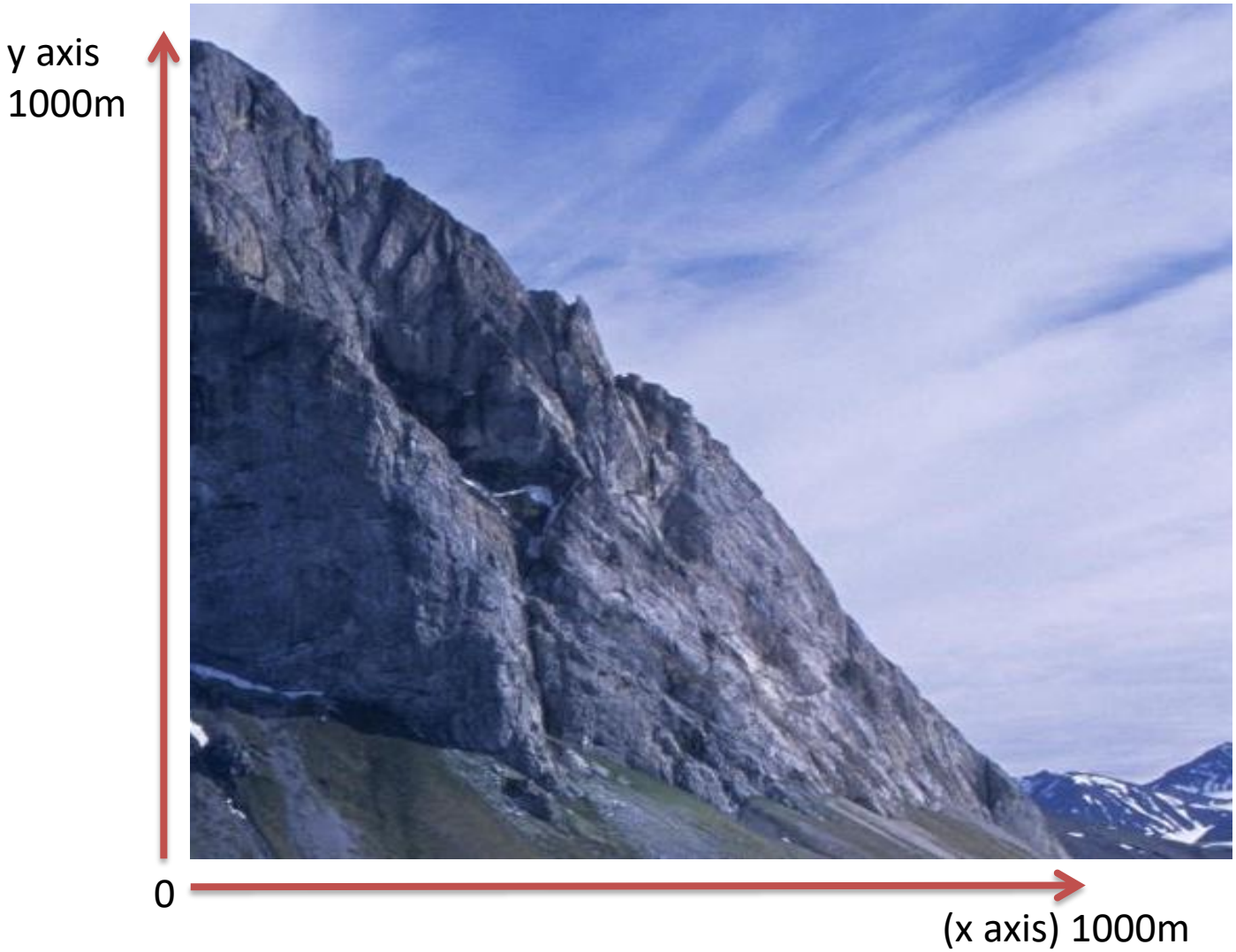
A model tells us the **important information** in an **easy-to-use** format

Models are **simplifications**. The complexity of the model depends on how **precisely** we want to represented the real behaviour

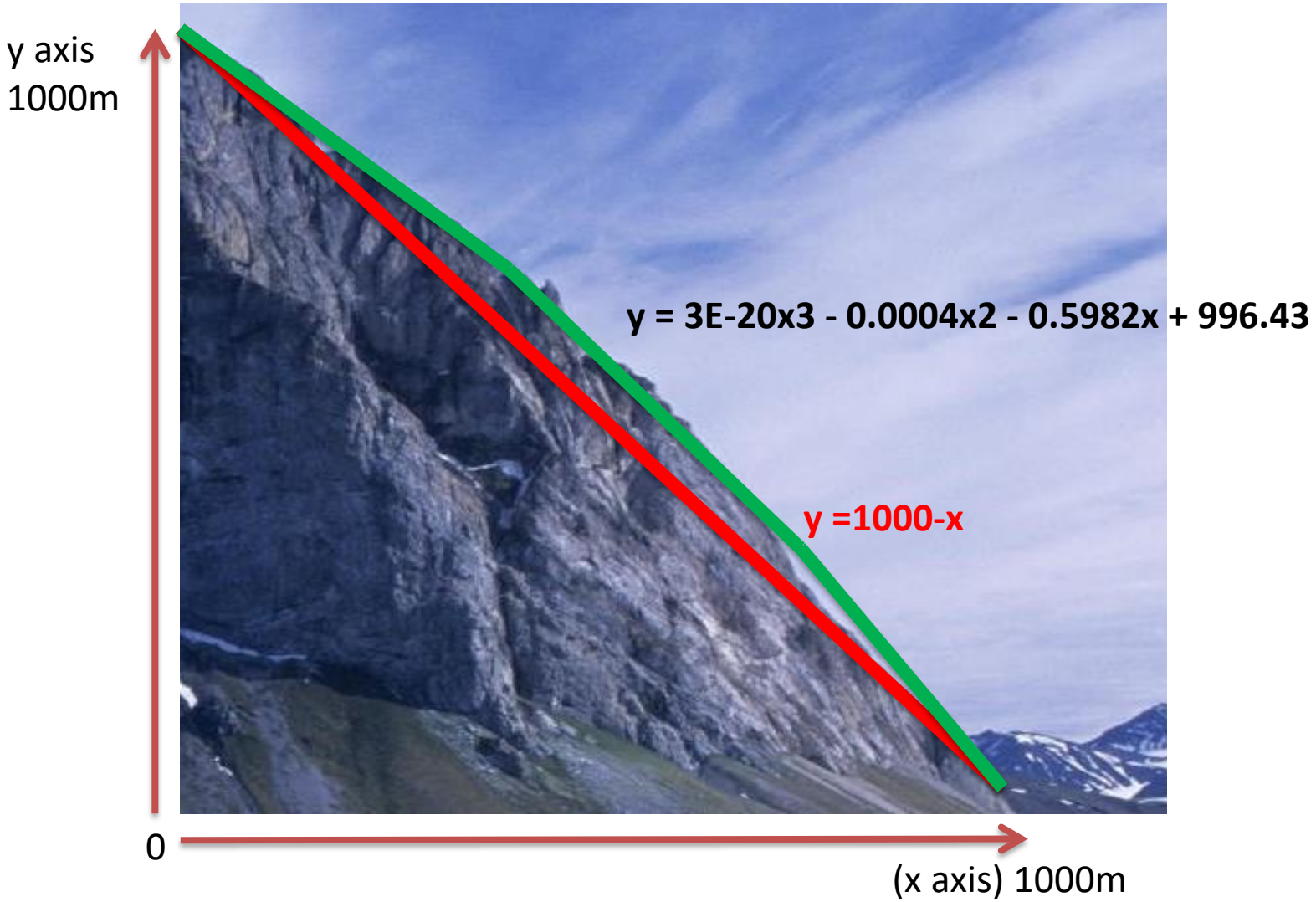
Models can be derived from the **underlying physics** (theoretical), but more commonly are found my **taking measurements** (empirical) and ideally combining the two approaches.

Models are **NOT** perfect, but they can be useful

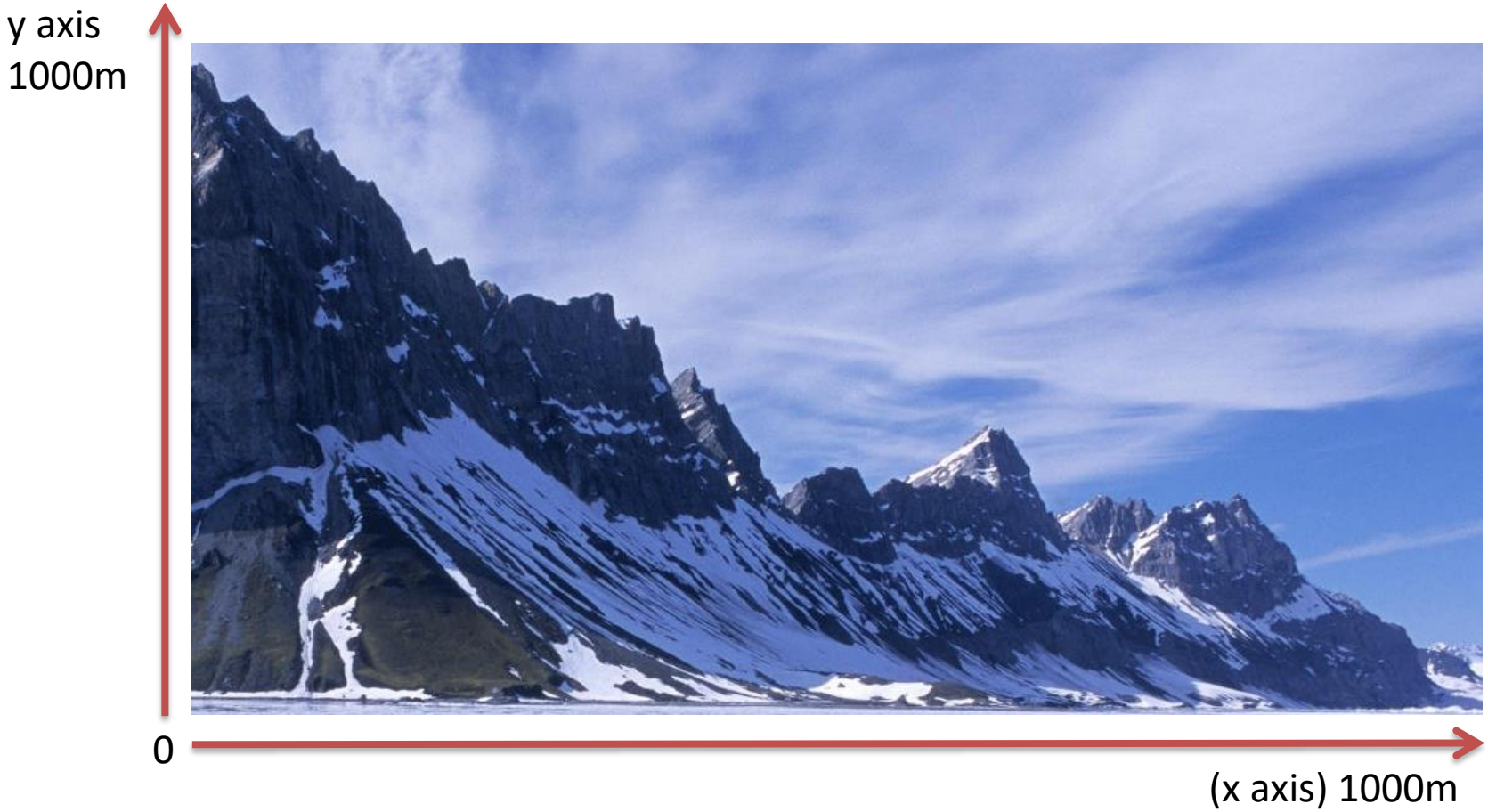
What sort of curve (we call it a model) could we use for this mountain to describe it's slope?



We like using **MATHS** as the language to explain things in a repeatable fashion, and it helps solve problems as well

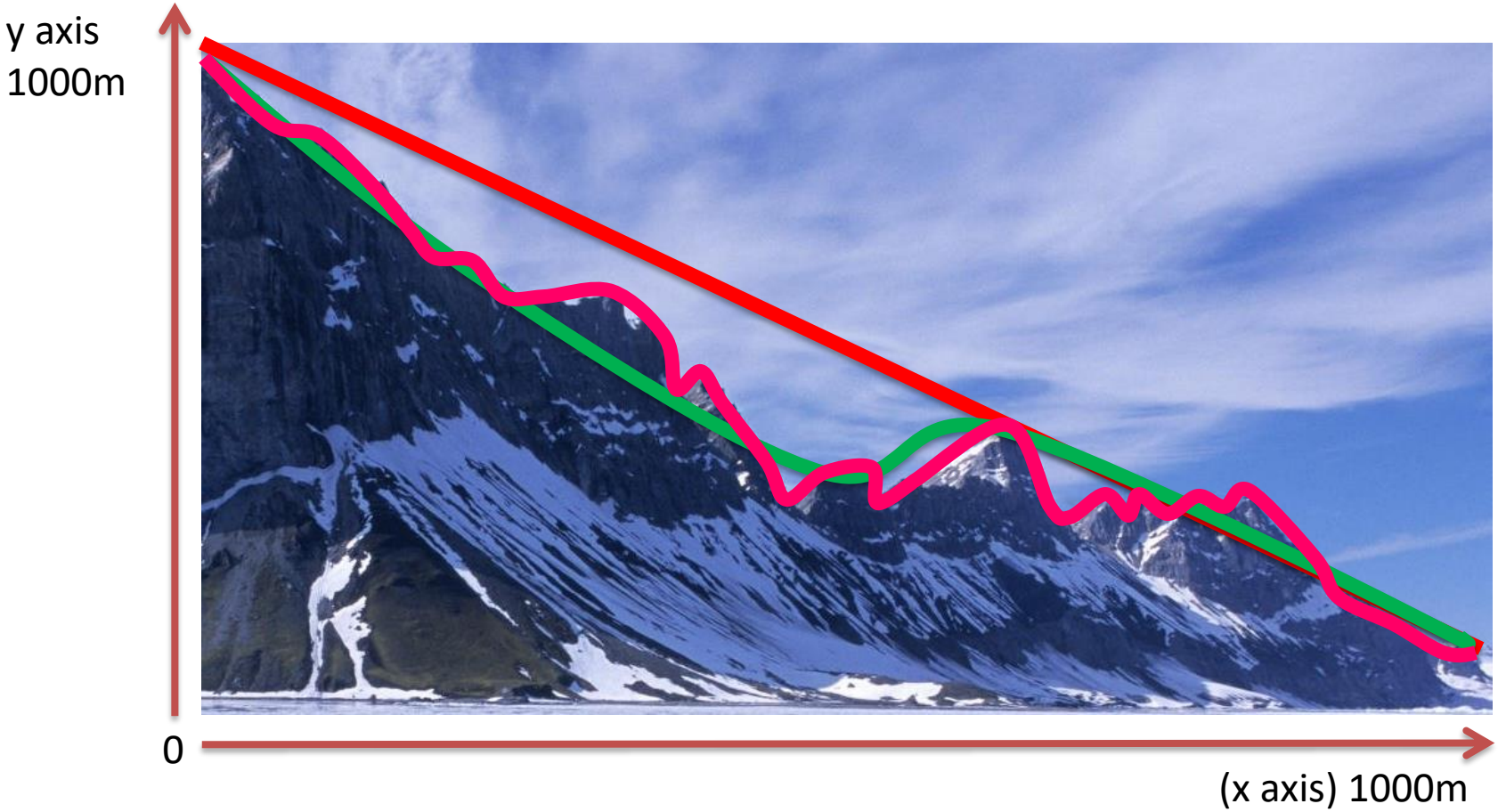


What sort of Models could we use for this mountain?



A choice: **simplicity versus accuracy**

Reality is messy, our tools only work well with simple models



Question: for your application... Do you really need to know about all the bumps and dips???

What is a model?

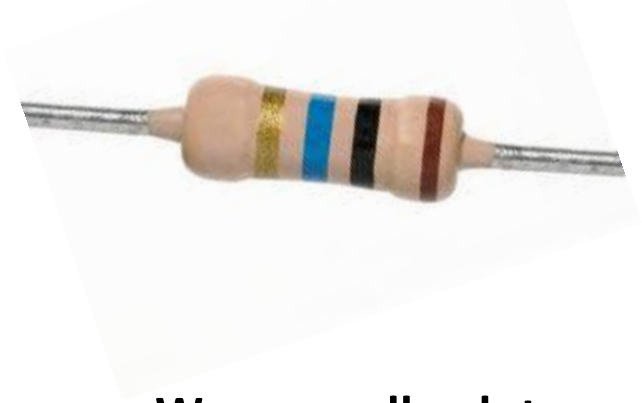
It's a tool to make our **lives easy**... Simple brains

It's a **mathematical representation** of reality that allows us to use maths to extract some useful information.

It's imperfect, it has errors. It will never match reality perfectly... But it's a good first guess..

A good model is "**close enough**" but "**simple enough**" to use

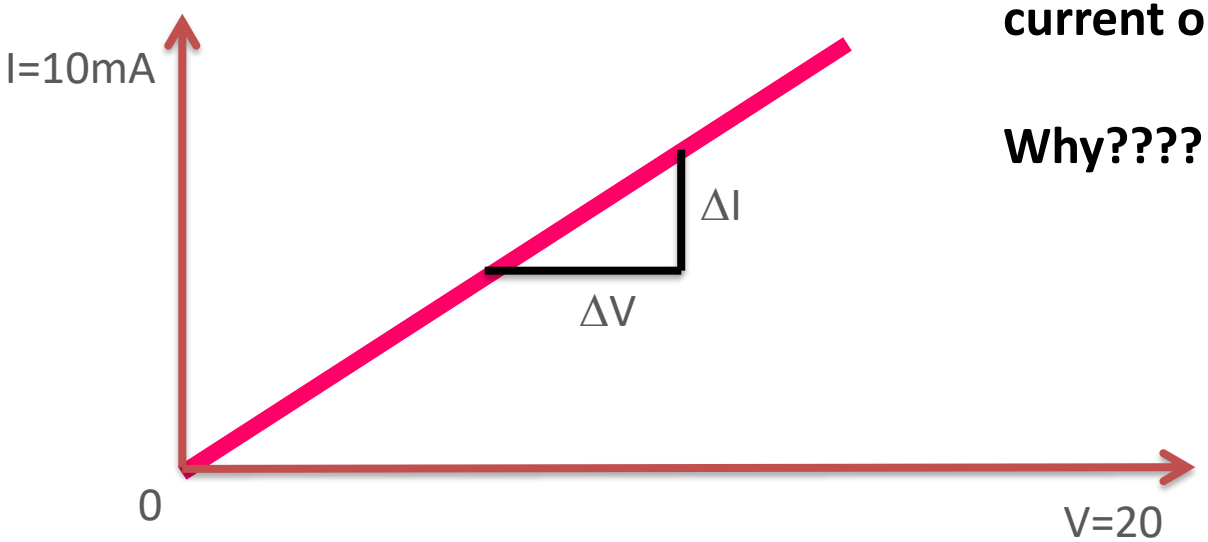
Looking at a Resistor



Ohms Law **V = I * R**

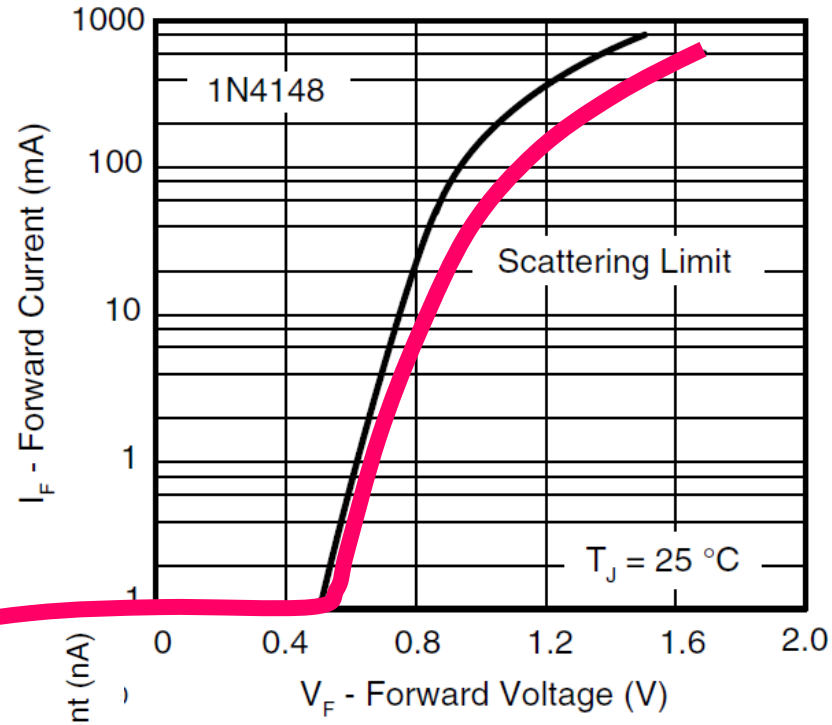
We normally plot voltage on X axis and current on Y.

Why????

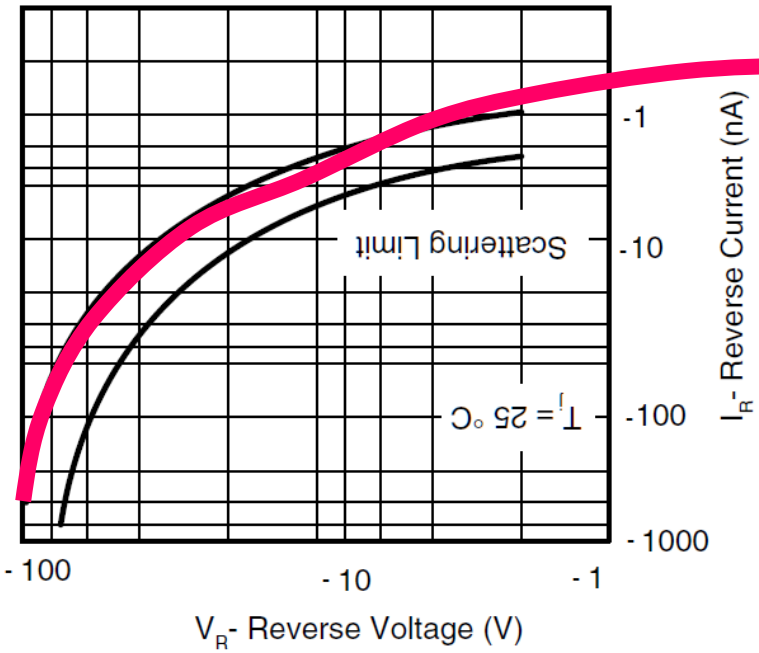


What does the slope of the line correspond to??

Looking at a Diode



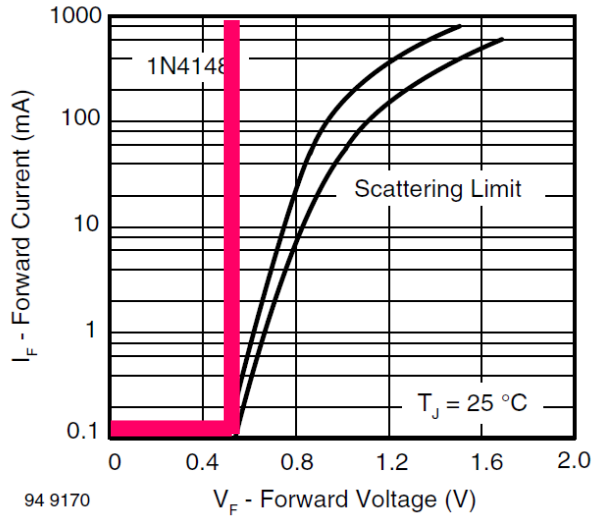
Reverse current is 1000 times smaller than forward current

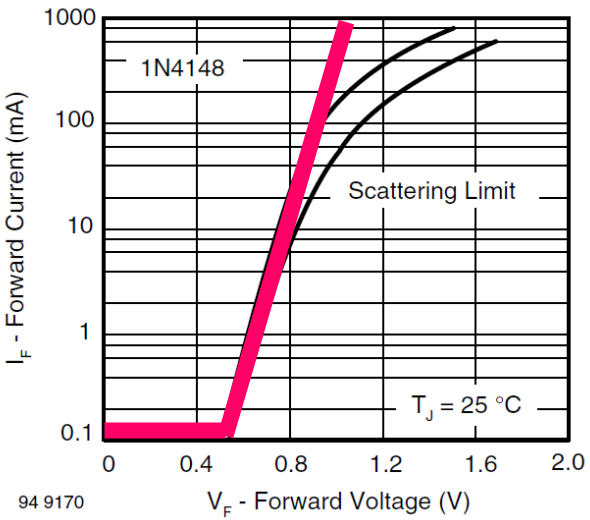


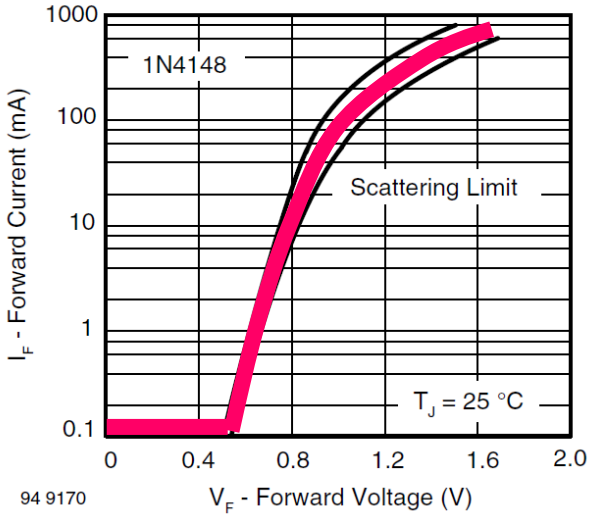
$$I_D = I_S\left(\exp\left(\frac{V_D}{nV_T}\right)-1\right)$$

$$n = \text{ideality factor}$$

One device, multiple models







Which model is most accurate?

Which model is the best to use?

A circuit-based model

We have tools to solve circuits, and sometimes we can see ways to simplify circuits to help us solve them for a specific purpose.

So can we take a mathematical expression and convert it to a circuit representation.

Lets add some new circuit elements to our collection

Independent Sources

$I_{out} = 0.3$ (or any constant)

$V_{out} = 0.3$ (or any constant)

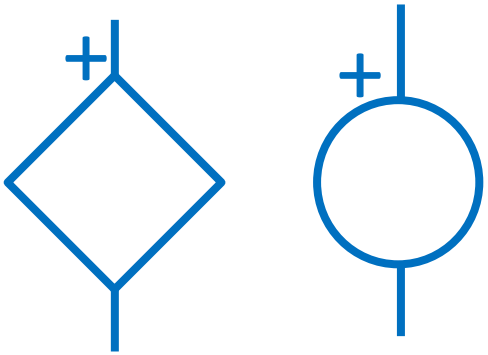
These are constant valued sources. Nothing effects what they do. An independent voltage source is a battery.

Independent current sources are rarer.

Dependent Sources

Voltage Controlled Current Source

$I_{out} = k\ V_{control}$

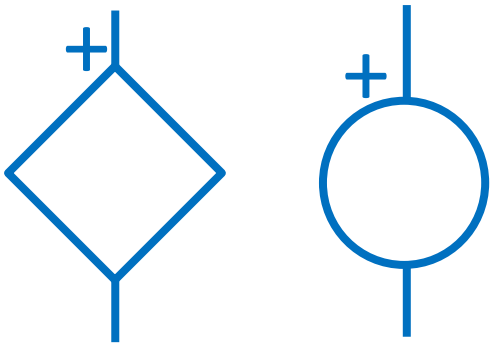


Voltage Controlled Voltage Source

$V_{out} = k\ V_{control}$

Current Controlled Current Source

$I_{out} = k\ I_{control}$



Current Controlled Voltage Source

$V_{out} = k\ I_{control}$

Don’t worry about how these are made, they are normally useful quirks of physics. With these, you need to write out what their controlling signal is as symbols don’t tell you enough

Making a circuit model (1)

I am told that for this “device”, the following happens

the current through the device is unknown, but the voltage across the device is a constant value **k**

$$\Delta V = k$$
$$I = \text{unknown}$$

This is the expression for a constant voltage source, or a battery.

We have no information about what the current through the device is.

Making a circuit model (2)

I am told that for this “device”, the following happens

the voltage across a device is unknown, but the current through this device is **k** amps

$$\Delta V = \text{unknown}$$
$$I = k$$

This is the expression for a constant current source

We have no information about what the voltage dropped across the device is.

Making a circuit model (3)

I am told that for this “device”, the following happens

the voltage across this device is proportional (k) to the current flowing through it.

$$\Delta V = k*I$$

This is the expression for a resistor, where resistance R=K

Making a circuit model (4)

I am told that for this “device”, the following happens

the current through the device is proportional (**k**) to the voltage applied across it.

$$I = k * \Delta V$$

This is again Ohm’s law but for a conductance. The conductance of a device is inversely proportional to it’s resistance (R=1/k)

Making a circuit model (5)

I am told that for this “device”, the following happens

the voltage across a device is unknown, but the current through the device is **k** times the current I_{in} that is being measured somehow.

$$\Delta V = \text{unknown}$$
$$I = k*I_{in}$$

This is the expression for a current controlled current source.

NOTE: I_{in} is not the same as the current through the device. It’s from somewhere else in the circuit.

Making a circuit model (6)

I am told that for this “device”, the following happens

the voltage across a device is unknown, but the current through the device is **k** times the voltage V_{in} that is being measured somehow.

$$\Delta V = \text{unknown}$$
$$I = k*V_{in}$$

This is the expression for a voltage controlled current source.

NOTE: V_{in} is not the same as the voltage across the device. It’s from somewhere else in the circuit.

Making a circuit model (7)

I am told that for this “device”, the following happens

the current through the device is unknown but the voltage is **k** times the value of Vin, which is measured somehow.

$$\Delta V = k*V_{in}$$
$$I = \text{unknown}$$

This is the expression for a voltage controlled voltage source

NOTE: V_{in} is not the same as the voltage across the device. It’s from somewhere else in the circuit.

Making a circuit model (8)

I am told that for this “device”, the following happens

the current through the device is unknown but the voltage is **k** times the value of Vin, which is measured somehow.

$$\Delta V = k*I_{in}$$
$$I = \text{unknown}$$

This is the expression for a current controlled voltage source

NOTE: I_{in} is not the same as the current through the device. It’s from somewhere else in the circuit.

Making a circuit model (the fun part)

The important thing to understand here is that we have two types of devices – those that produce current and those that produce voltages.

You cannot successfully mix current and voltage output devices. The question "3 amps + 4 volts =" is meaningless.

So imagine that you have two forms of equations that you can use

$$I_{XY} = 3 + V_X*(0.2) - 13*V_1 + 5*V_2 + 0.1*I_3$$

$$V_{XY} = 0.2 + 5\ I_{XY} + 13*V_1 - 2*V_2 - 1.1*I_3$$

Making a circuit model (the fun part)

$$I_{XY} = 3 + V_{XY}*(0.2) - 13*V_1 + 5*V_2 + 0.1*I_3$$

Something that adds a current that is proportional to the current I_3

Something that adds a current that is proportional to the voltage V1 and V2

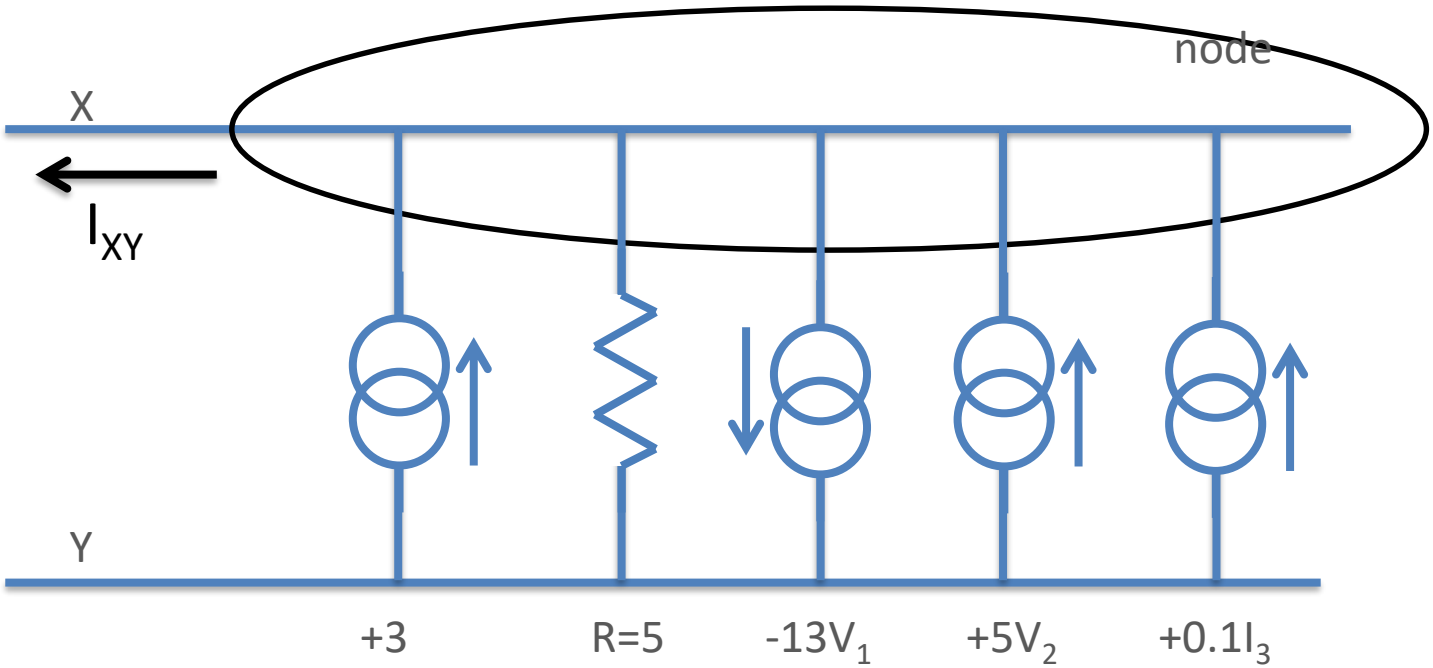
Something that adds a current that is proportional to the voltage across the points XY

Something that adds a constant current… easy

Making a circuit model (the fun part)

$$I_{XY} = 3 + V_{XY}*(0.2) - 13*V_1 + 5*V_2 + 0.1*I_3$$

Remember, KCL says that currents add at a node. Be careful with the arrows.



Making a circuit model (the fun part)

$$V_{XY} = 0.2 + 5\,I_{XY} + 13*V_1 - 2*V_2 - 1.1*I_3$$

Something that adds a voltage that is proportional to the current I_3

Something that adds a voltage that is proportional to the voltage V1 and V2

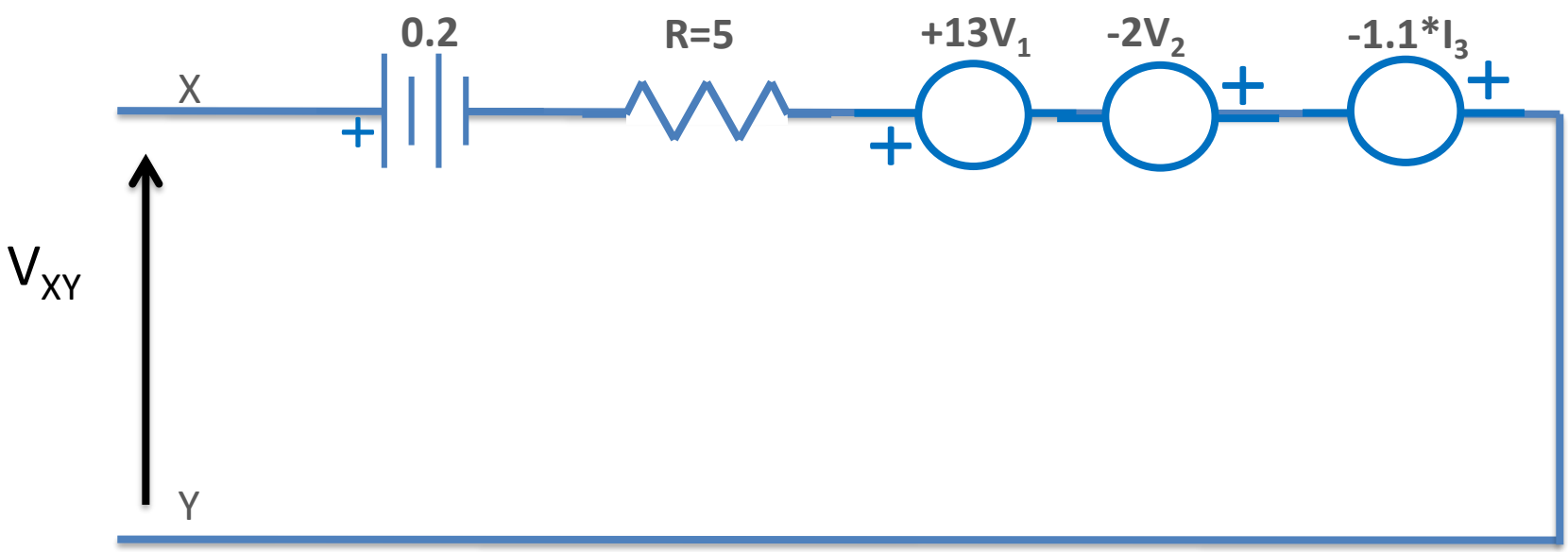
Something that adds a voltage that is proportional to the current flowing through it

Something that adds a constant voltage… easy

Making a circuit model (the fun part)

$$V_{XY} = 0.2 + 5\,I_{XY} + 13*V_1 - 2*V_2 - 1.1*I_3$$

Remember, KVL says that voltages add in a loop.
Be careful of the signs.



Making a circuit model

Circuit models have a directly equivalent mathematical model. We use which ever one is easiest.

In circuit models, we will attempt to never use $()^2$ or anything complicated. Our aim is simplicity

Remember all models are simplifications. Computers are great at crunching equations.

They'll give you the answer to 6 decimal places but they can't tell you if it's the right answer to the problem you wanted to solve.

Biasing and Small Signal Behaviour

What is a small signal?

It’s a change in your signal that is small.

What is small?

Small enough that you don’t get any weird behaviours

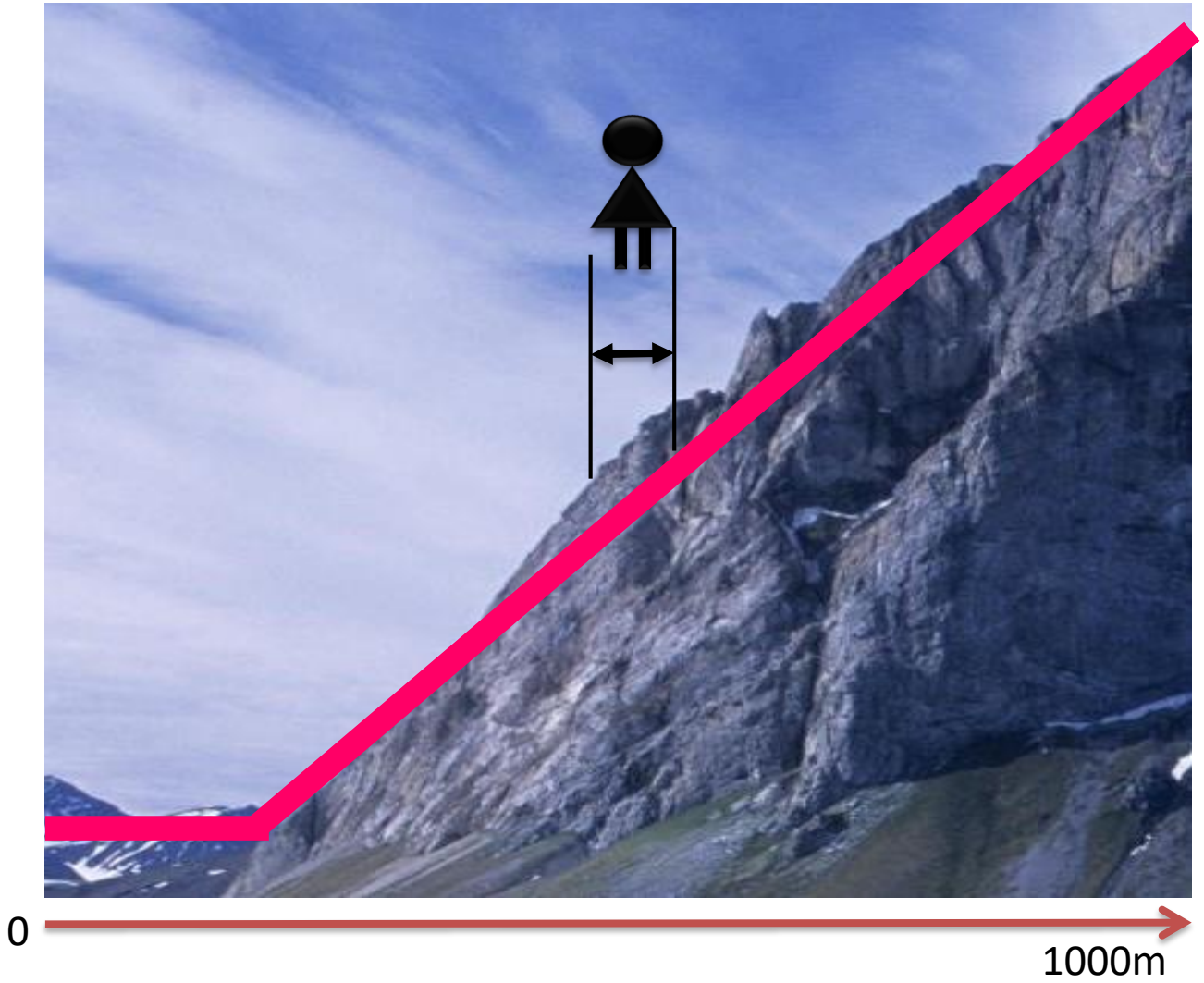
Seriously... What is small??

I am being serious... There is no good definition

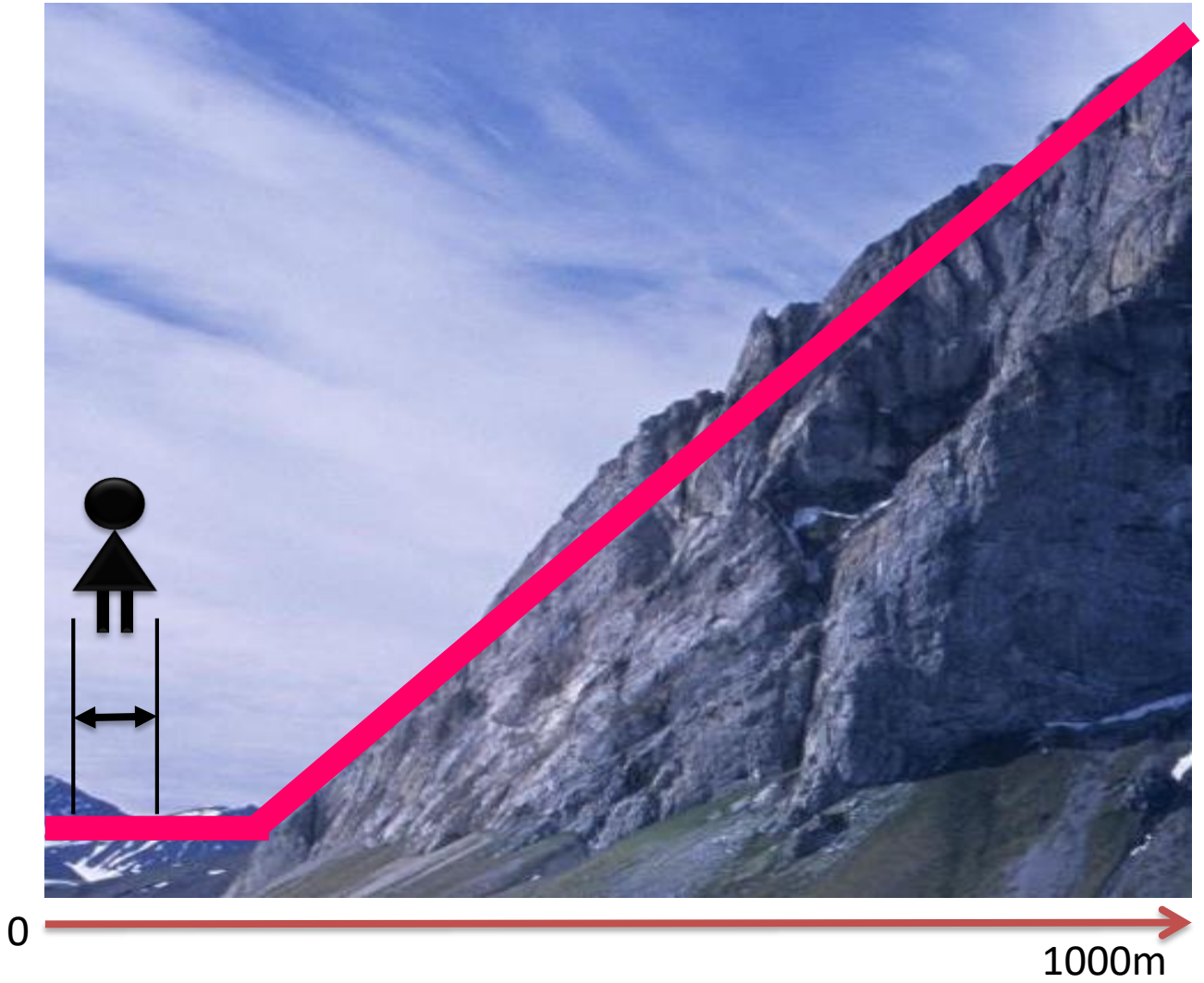
Got an example?

Usually these are AC signals, for example microphone audio output signal (mV)

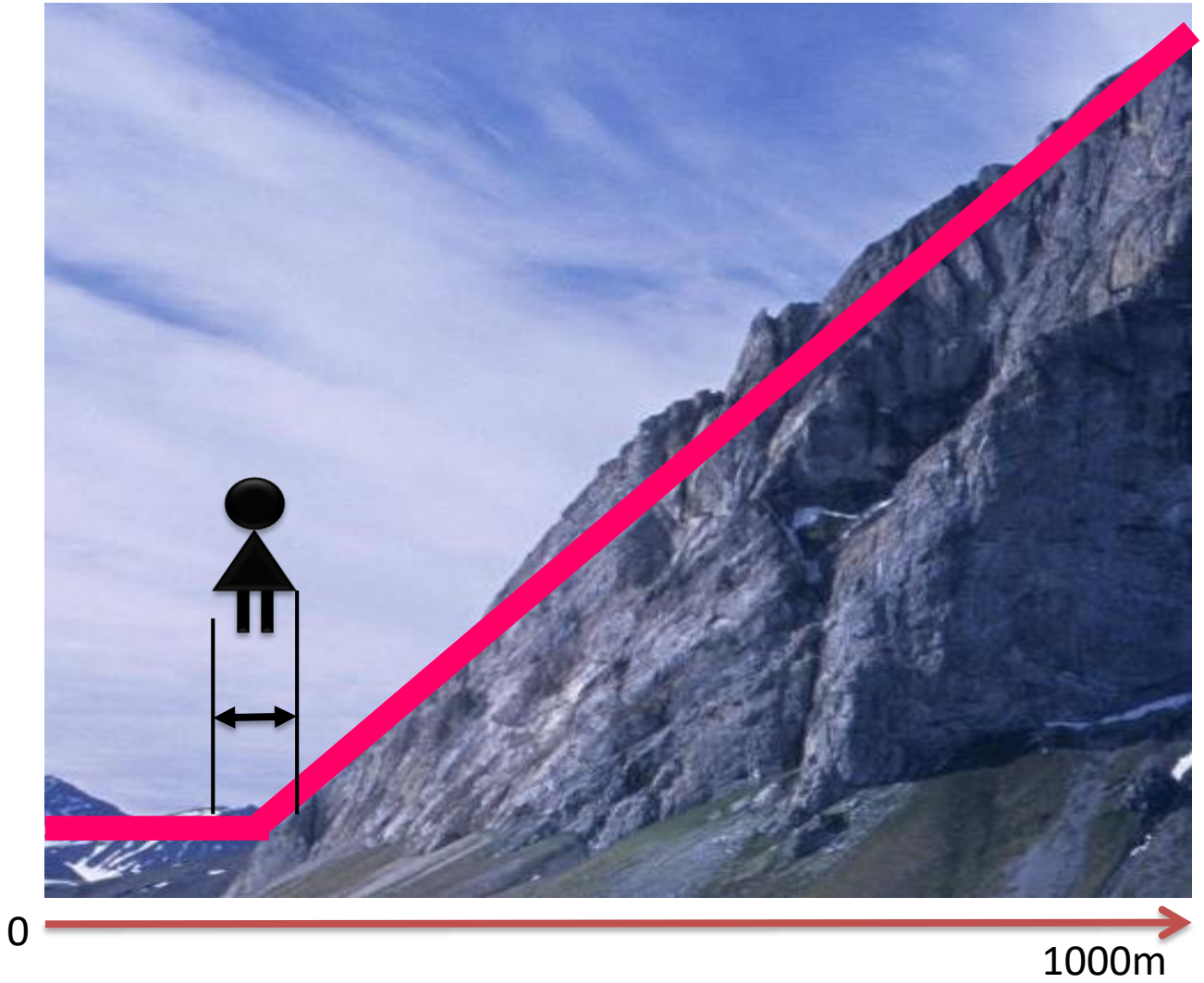
What slope does the person experience?
How far can they go before they experience a different slope?



What slope does the person experience?
How far can they go before it changes significantly?



What slope does the person experience?
How far can they go before the slope changes significantly?

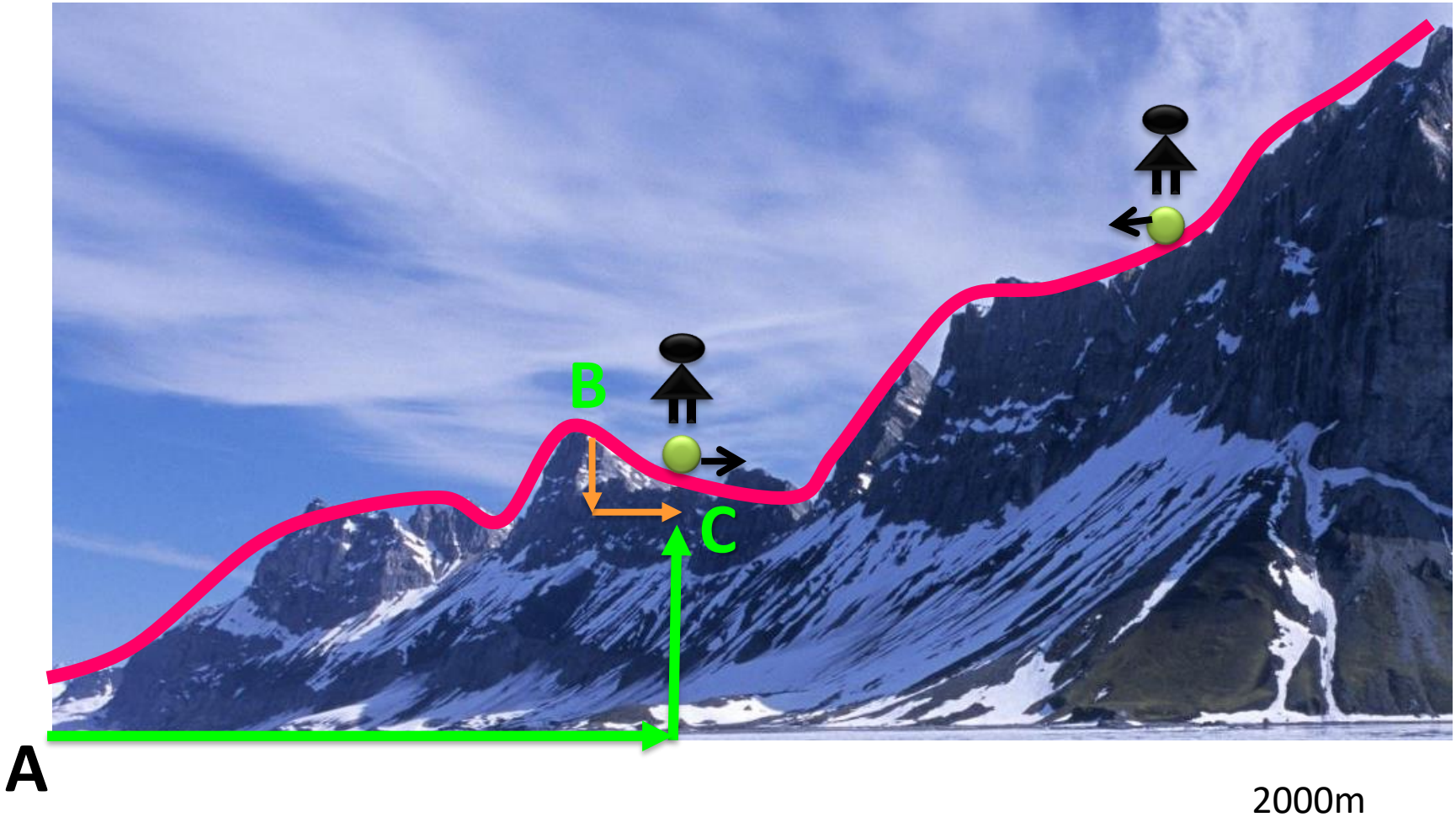


As I walk left to right… there will be points where I may end up walking downhill.

What does this mean??

Over a very small distance, everything looks like a straight line

The slope of the line at each point can of course change.

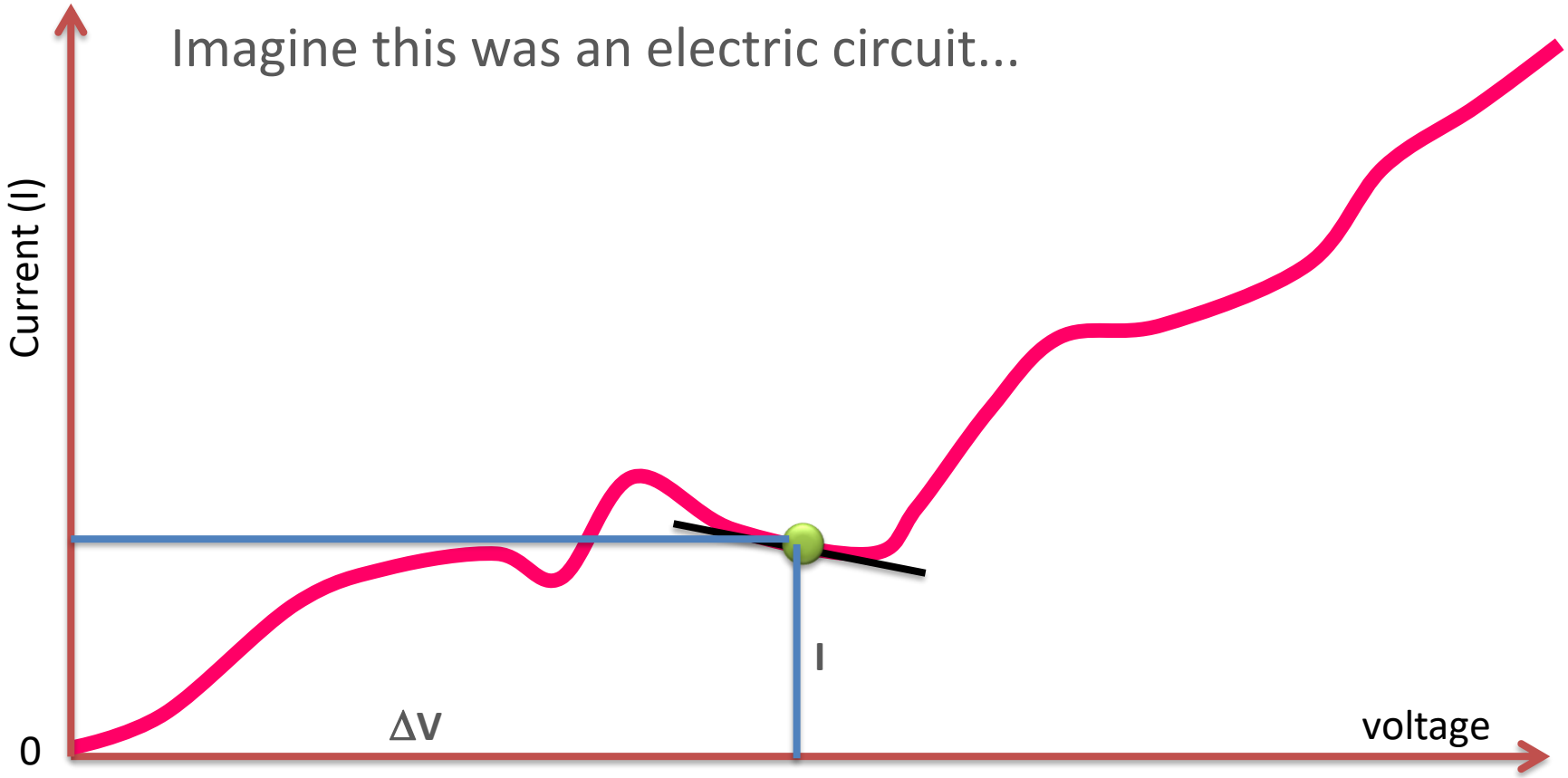


So am I walking uphill or not… it depends on your perspective..

If I am starting at A….

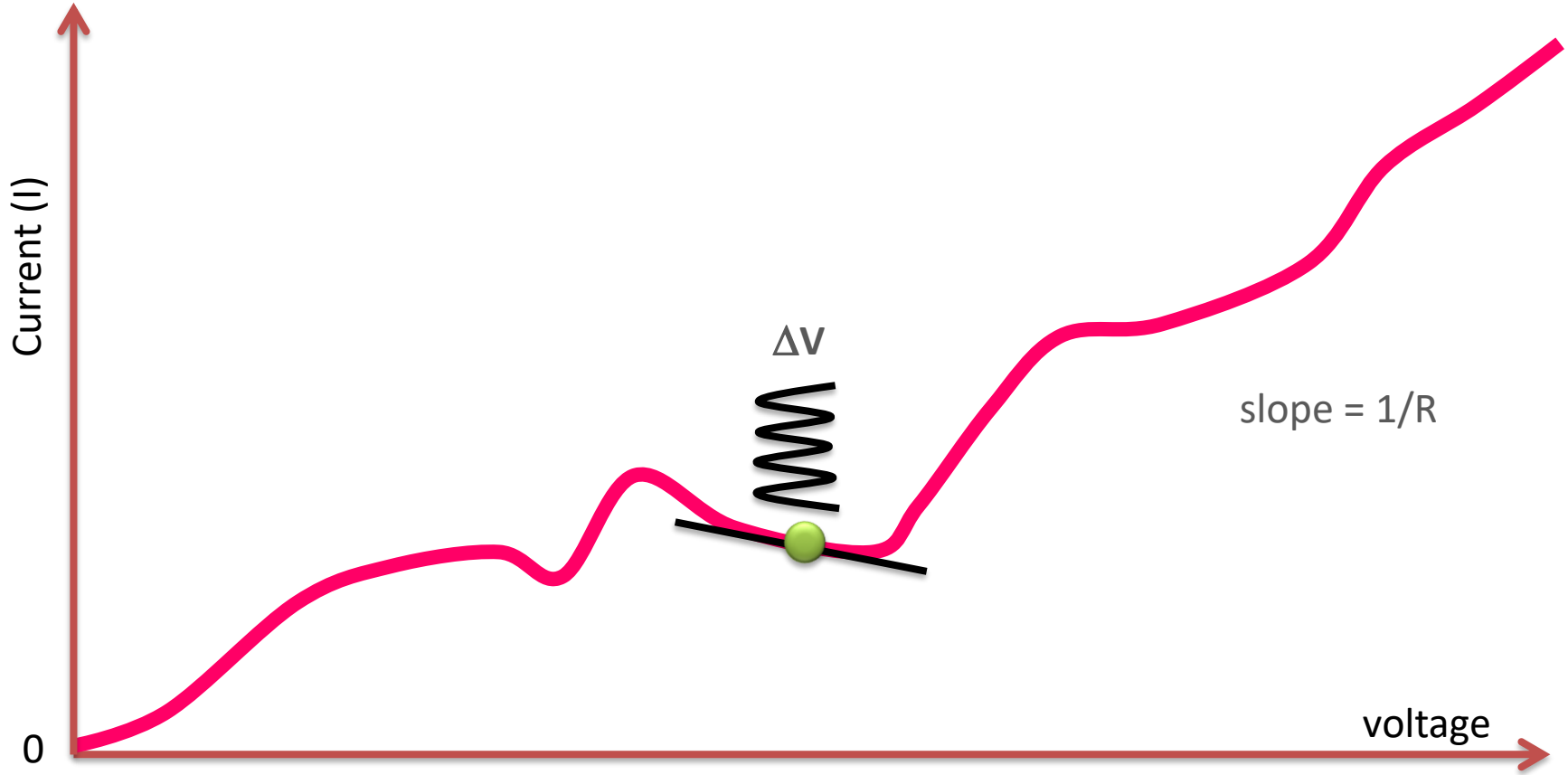
If I am starting at B…

The same applies to Voltages and Currents



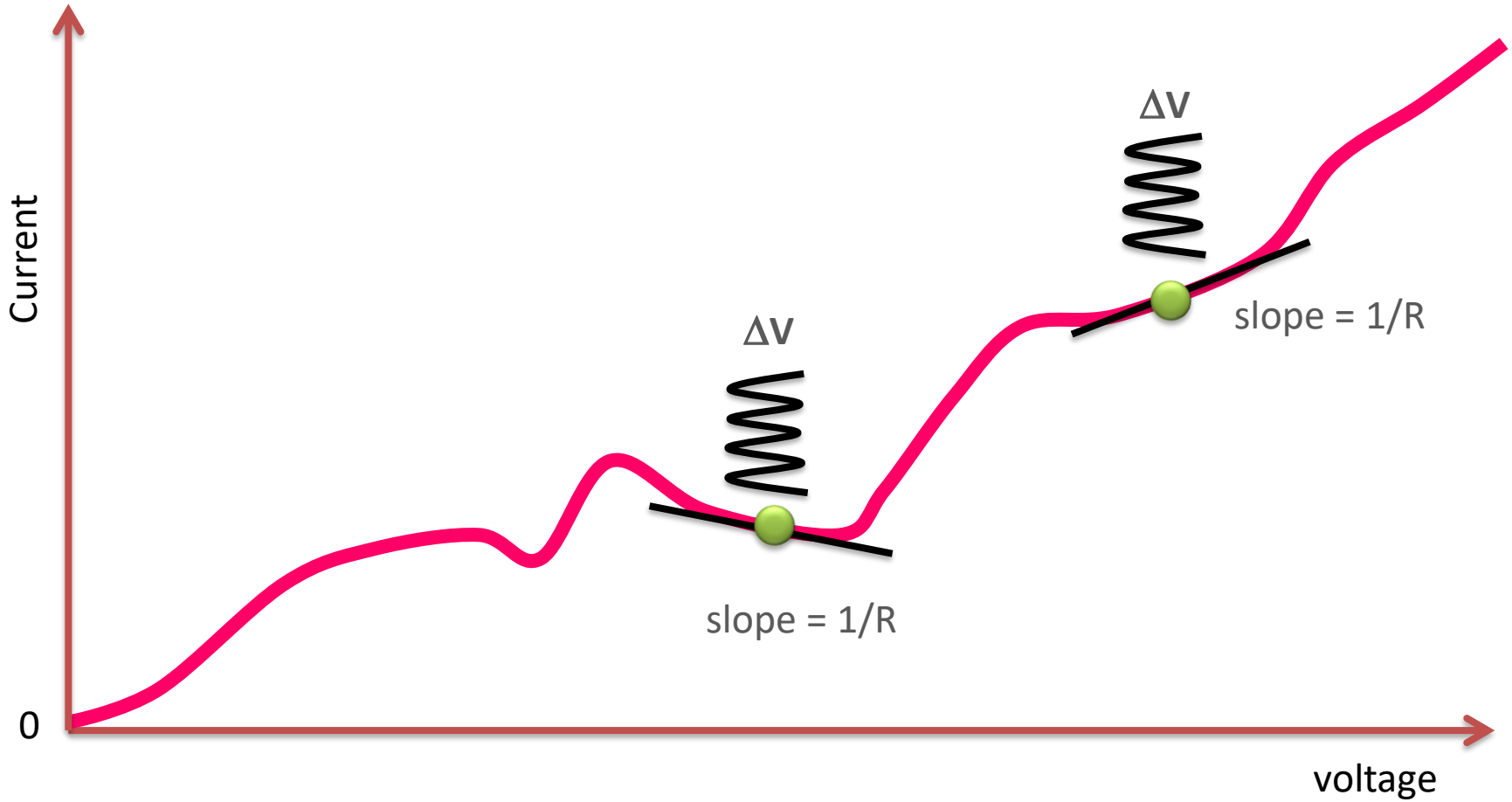
If I take the total voltage dropped, and the total current...

I get a value of resistance given by $\Delta V = I*R$



Now imagine the numbers where R=100, V=10, I= 0.1

Looking at this curve, what would happen to the current if I increased the voltage a little… would it go up or down?



A small change in voltage causes a small change in current

$$\Delta V = R_{small} (\Delta I)$$

This Resistance to small changes is not always positive and it's clearly not always the same for different parts of the curve and it's not the same as the resistance for large signals

Small Signals and Large Signals

Small Signals are small changes that are so small that they stay is a region on the curve where everything looks like a **straight line** (we call straight-line regions **LINEAR**) **(this is the correct exam-paper answer)**

Small Signals are always considered to be time-varying signals with zero average values – often picked up from the sensors and need to be amplified.

Small Signal Resistance is given by

$$\Delta V = R_{small} (\Delta I)$$

Small signals use lower case letters (v, i)
Large signals use upper case letters (V,I)

Large signals are any and all signals that are not small
Large signal Resistance is given by

$$\Delta V = R (I)$$

What is Biasing

The **Bias Point** is where you **get the behaviour** you want for your small signals. You choose it and then you design your biasing. A good bias point is far away from sharp changes in your curve.

Biasing is the process of **getting to the point** in the curve that you want. The curve represents the behaviour of the device

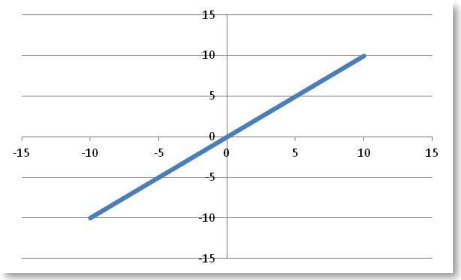
Biasing is done through the use **of DC signals**.

Small Signals are small changes that are so small that they stay in a linear region on the curve, anything between 10 and 100 times smaller than the big signals is a good definition of small.

Your real signal is your **small signal + biasing**. Or else it's just a very small signal and won't let you do much.

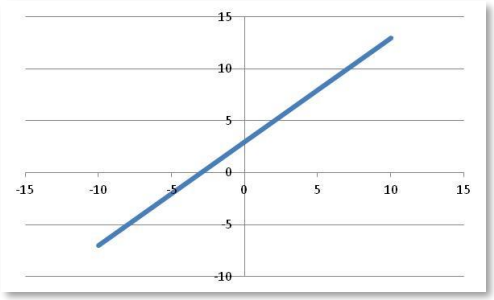
Some more curves or models….

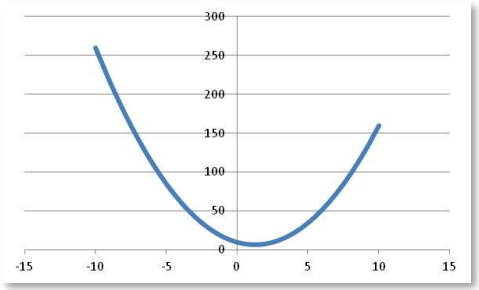
Linear, Quadratic, Cubic, etc...



Linear, special term **AFFINE** as it goes through the origin. It does not need to be 1:1, just a straight line

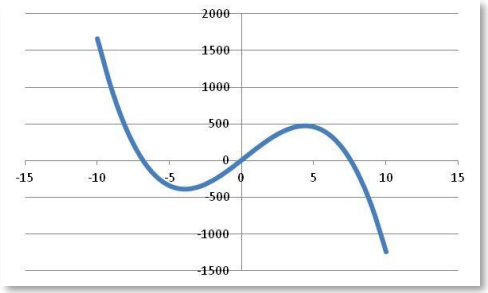
Linear

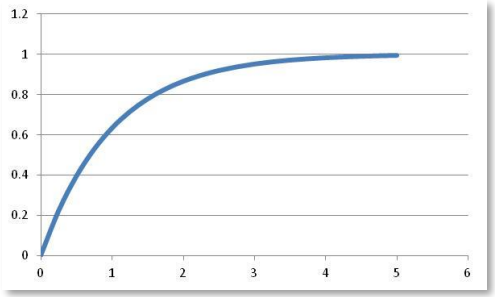




Quadratic

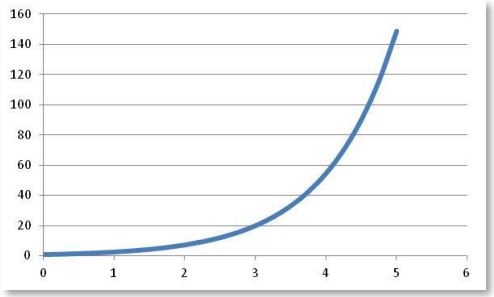
Cubic



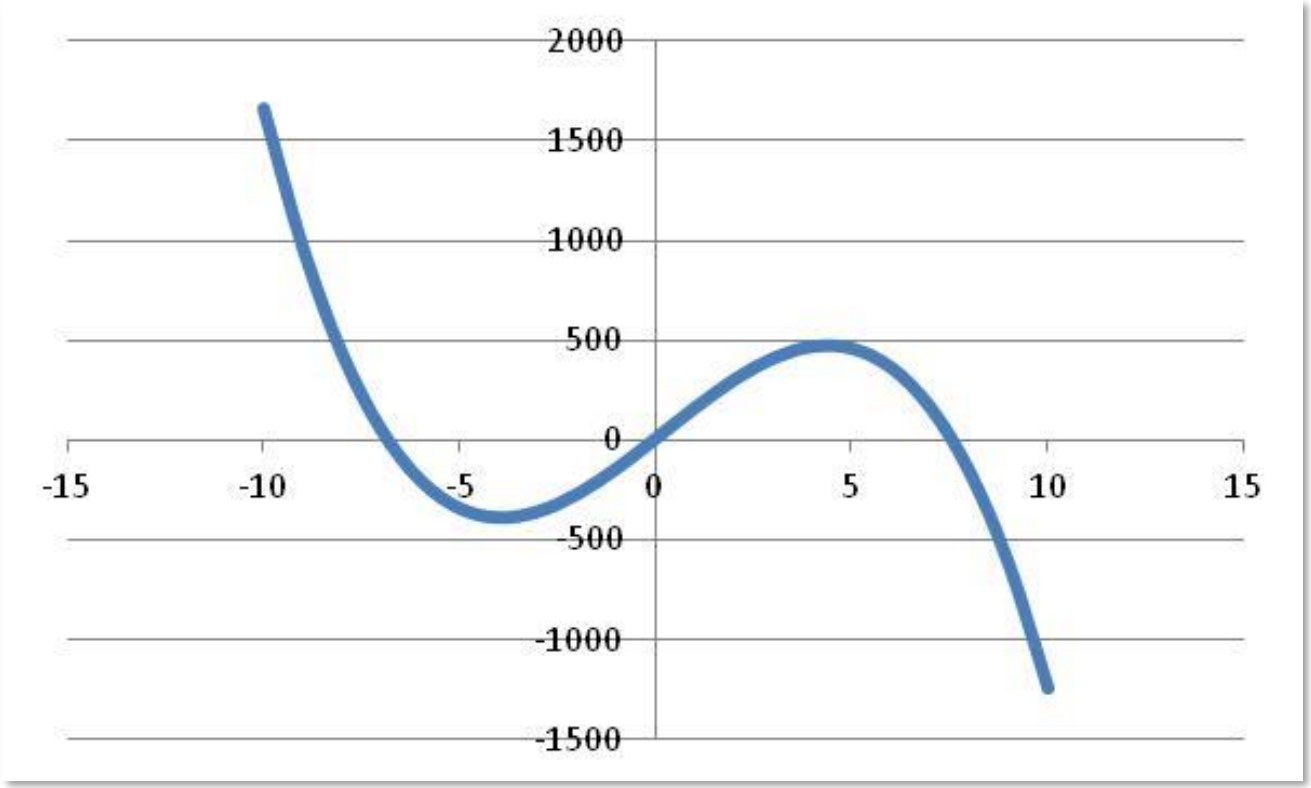


Asymptotic exponential but there are other forms that do this too

Exponential



Piecewise Linear

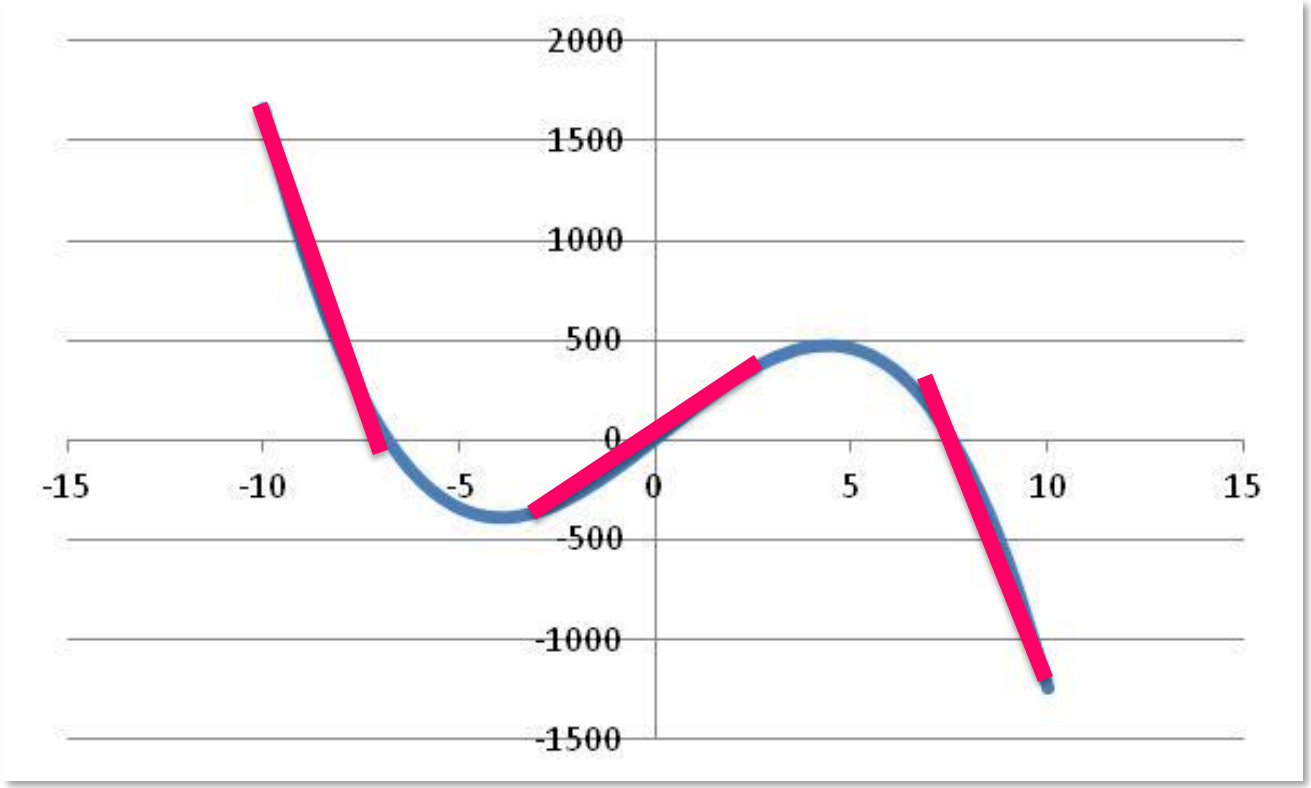


We say something is piecewise linear if we can remake the curve using a bunch of linear sections. It cannot ever be a perfect fit unless the pieces are infinitely small. The smaller the pieces, the better the bit.

Piecewise Linear – small pieces

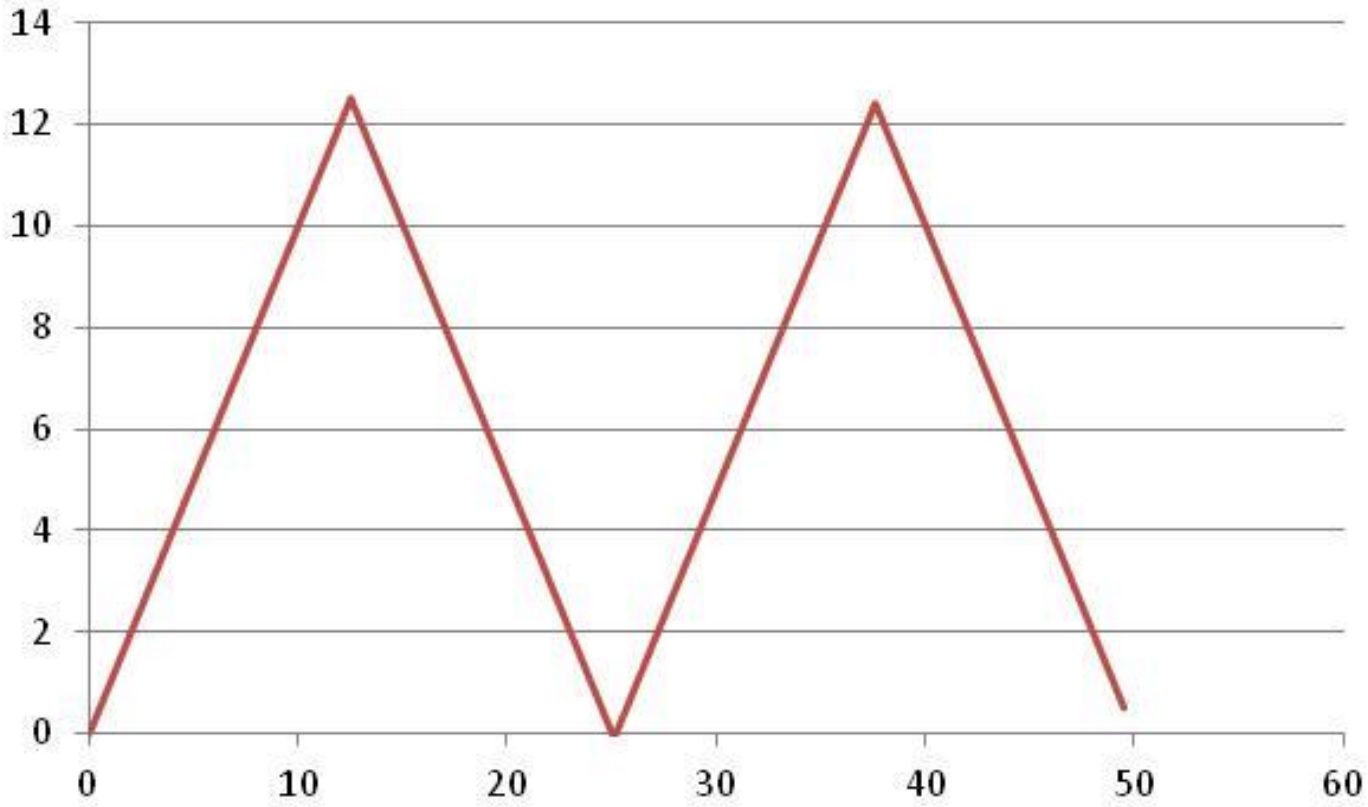
This should remind you of numerical techniques for doing integration (simpson's rule)

Piecewise Linear – small pieces



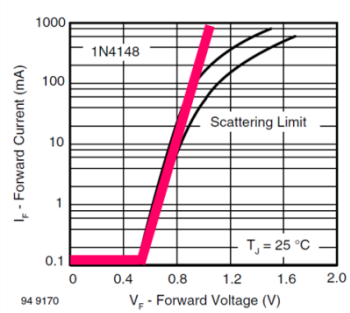
Usually when we are doing piecewise linear approximations... We ignore the messy bits and only look at whether there are large regions that can be well approximated by a straight line... Otherwise it's not that useful to us.

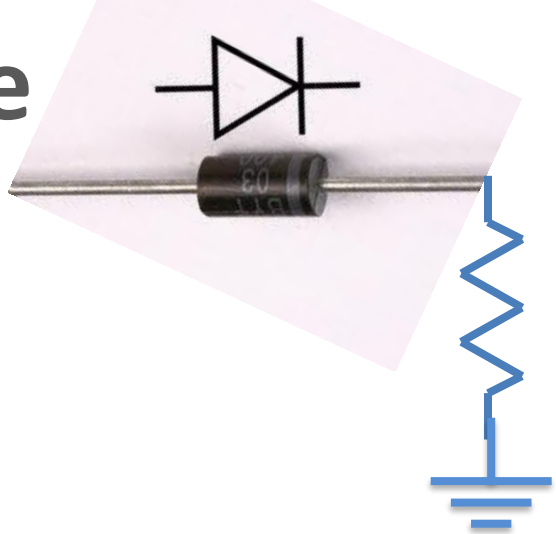
Piecewise Linear – small pieces



Some shapes are very easy to represent as piece-wise linear, provided you avoid wondering what happens at the junctions ☺

An example



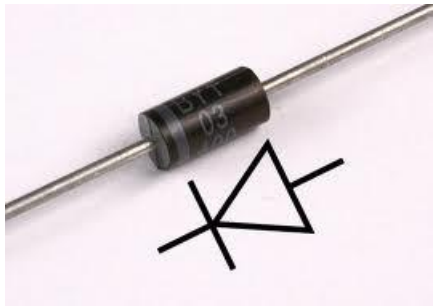


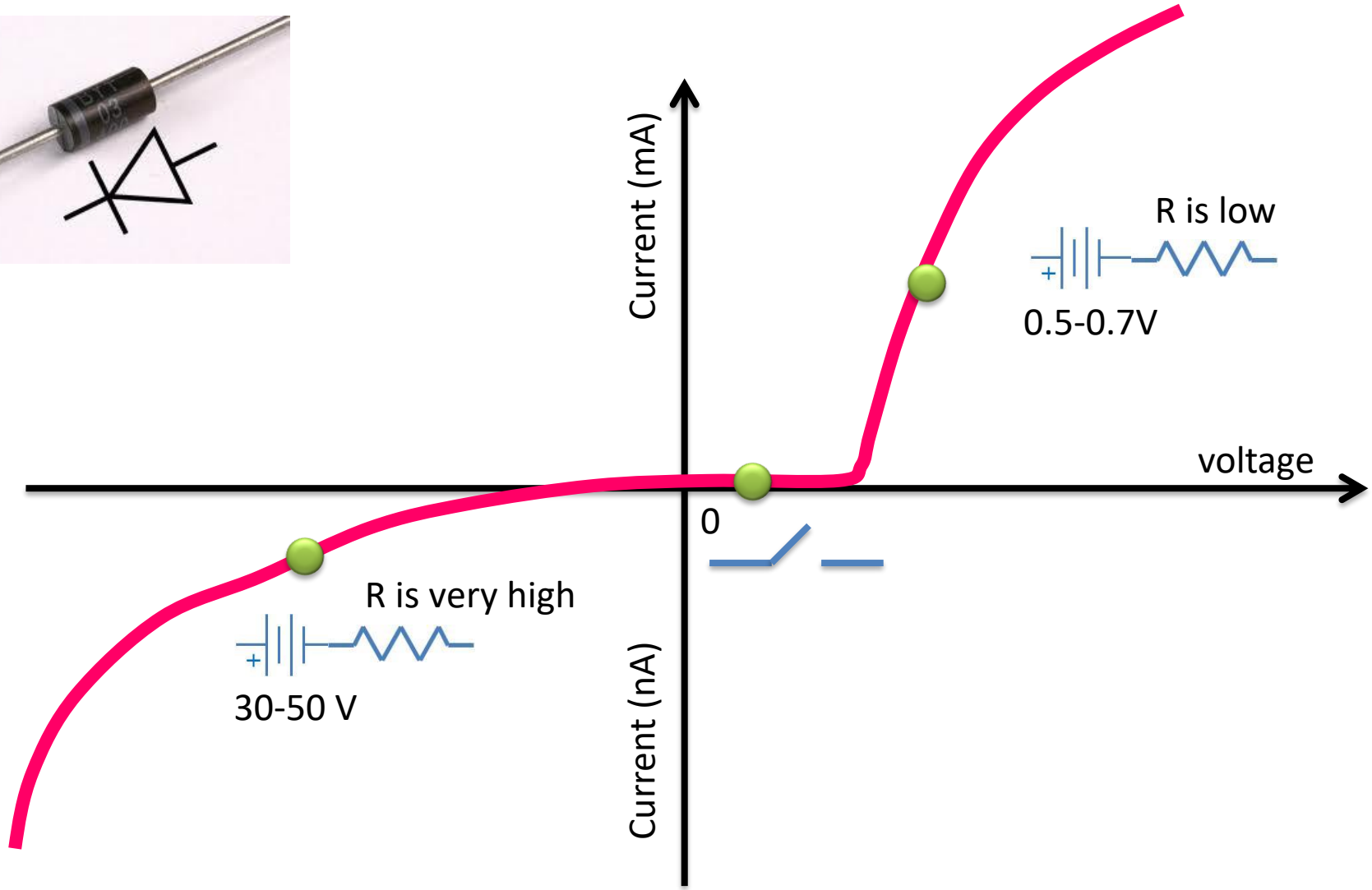
Threshold voltage = 0.7 Volts
On resistance of 1 ohm
Load Resistance of 100 ohms

I input a sinusoidal signal of amplitude +/- 100mV,
what voltage do I see on the resistor? Sketch!

I input a sinusoidal signal of amplitude +/- 100mV with a +1Volt constant offset
what voltage do I see on the resistor? Sketch!

Looking at a Diode





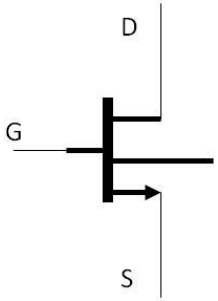
Equations with more than 1 input

For diodes, capacitors, resistors, inductors, the current is dependent only one factor, the voltage applied across it.

There are some devices where the current that flows through it is dependent on **MORE THAN ONE VALUE**. The most important one of these is a transistor

More on this later, but one type of transistor has the following equation for current flow

$$I_D = \mu_n C_{ox} \frac{W}{L}\left[(V_{GS} - V_{tn})V_{DS} - \frac{V_{DS}^2}{2}\right]$$



Scary Solid-state equations

$$I_D = \mu_n C_{ox} \frac{W}{L}\left[(V_{GS} - V_{tn})V_{DS} - \frac{V_{DS}^2}{2}\right]$$

$$I_D = I_S\left(\exp\left(\frac{V_D}{nV_T}\right) - 1\right)$$

You will not be asked to derive them or memorise them

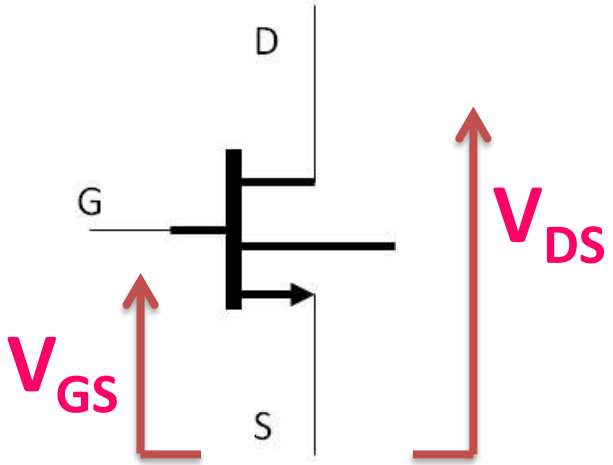
You do need to be able to interpret them, which bits are constant and which bits cause a change

They explain behaviours but they're not perfect either

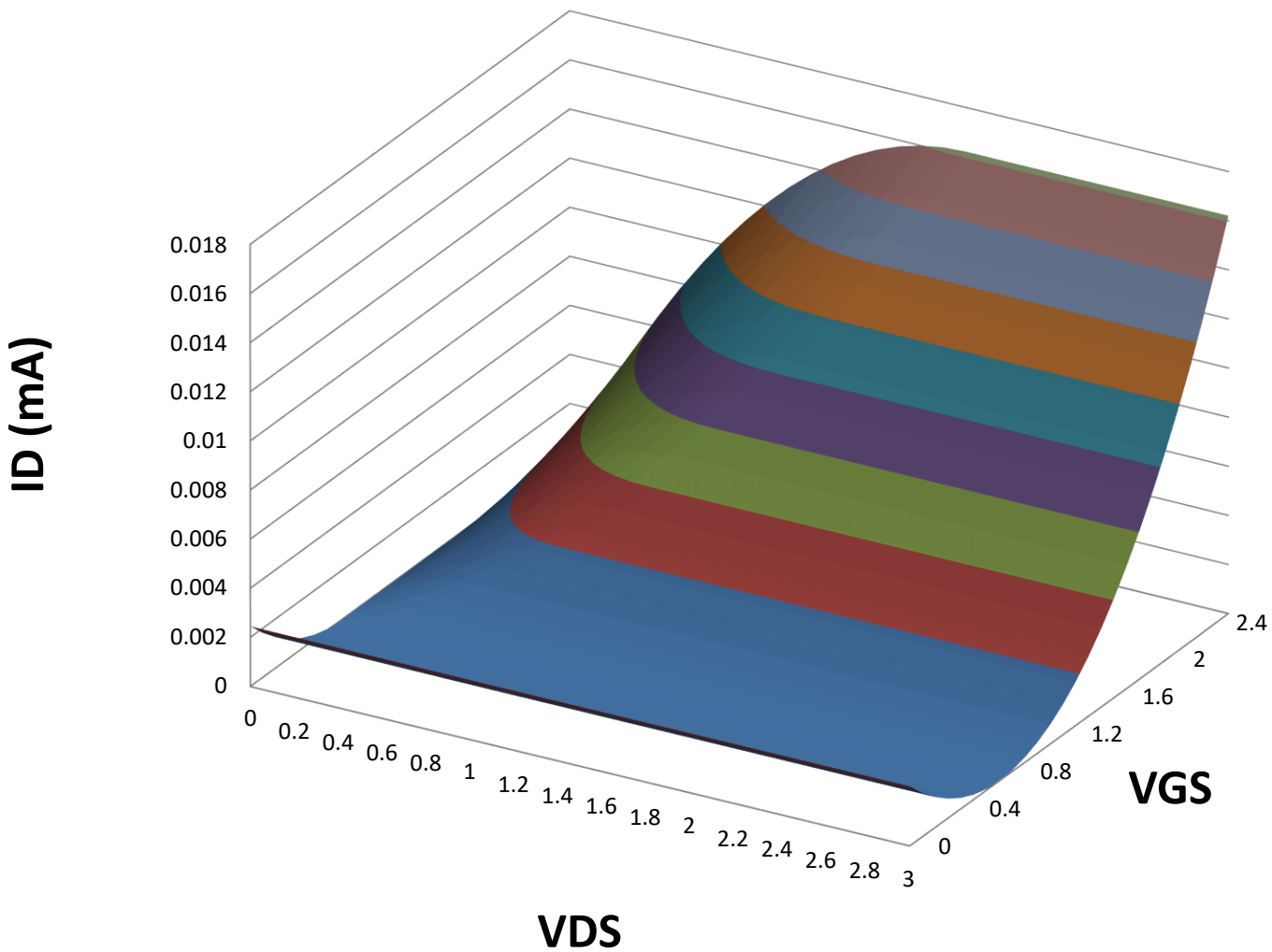
Explaining this equation – just a little

These are constants that depend on the construction and physical make up of the transistor. They do not change

$$I_D = \mu_n C_{ox} \frac{W}{L}\left[(V_{GS} - V_{tn})V_{DS} - \frac{V_{DS}^2}{2}\right]$$

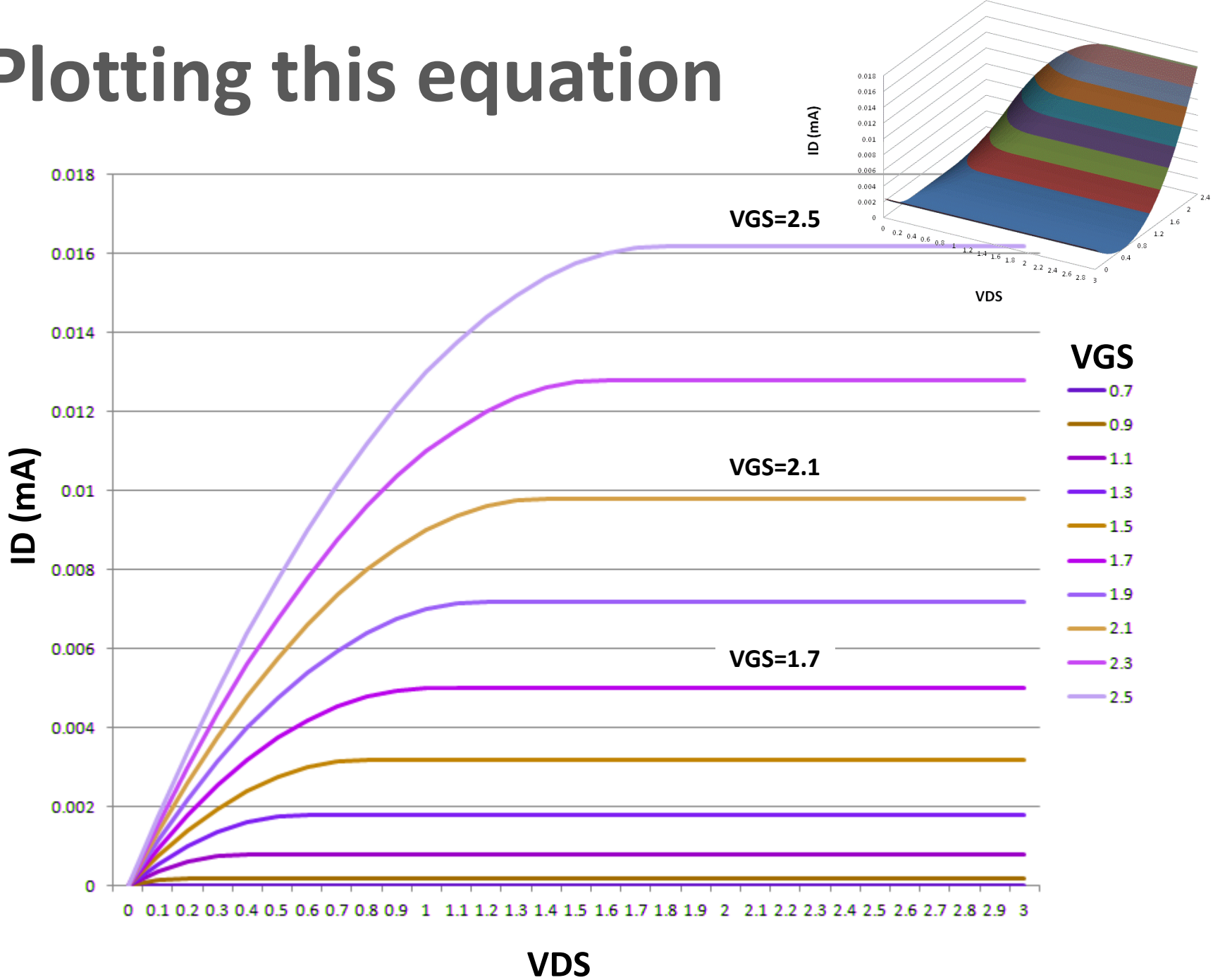


Plotting this equation



This was drawn from the equations.. Can anyone spot the problem???

Plotting this equation



3D Models

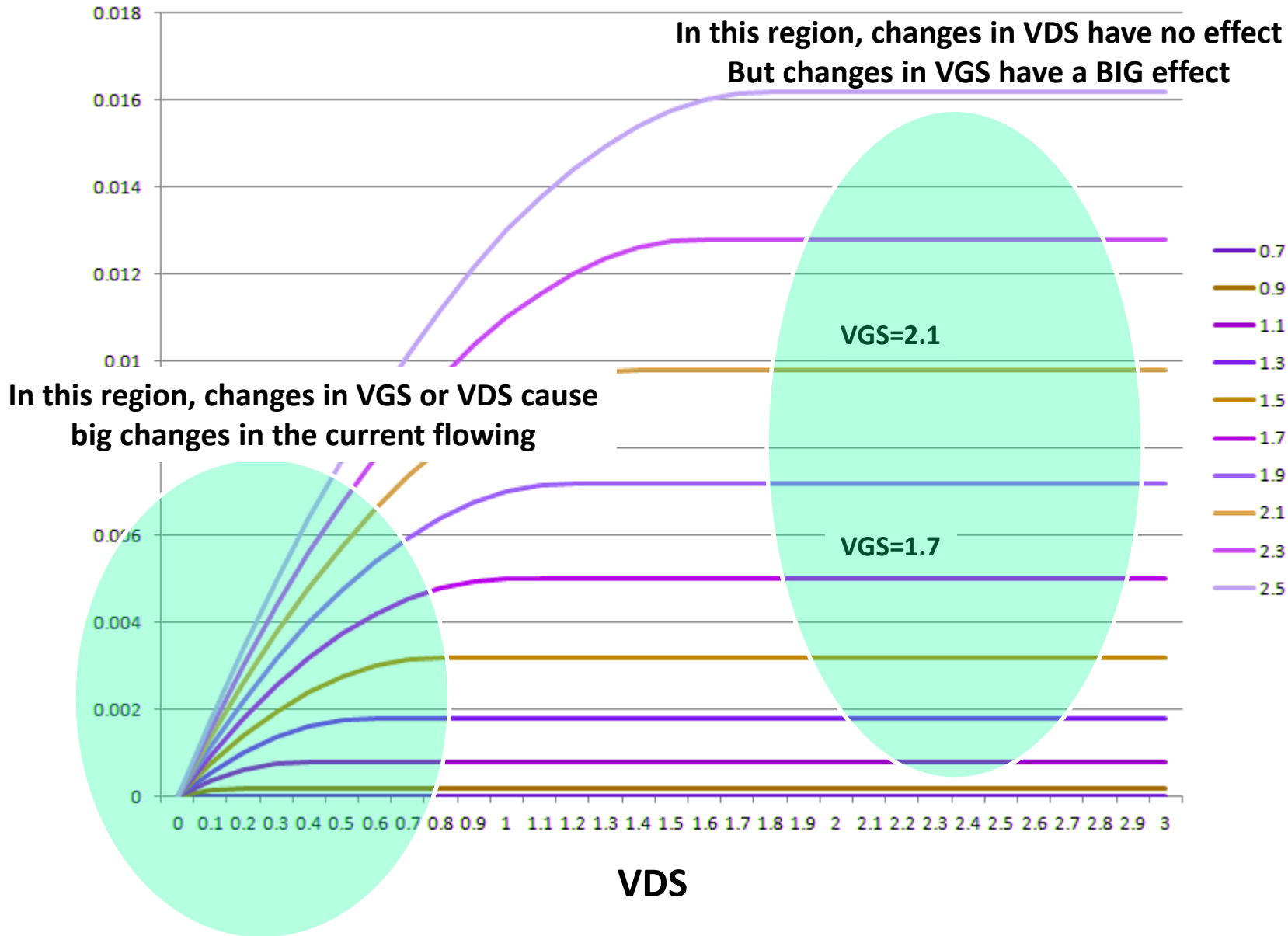
Much harder to visually interpret

These lines are smooth but reality definitely does not provide us with smooth lines

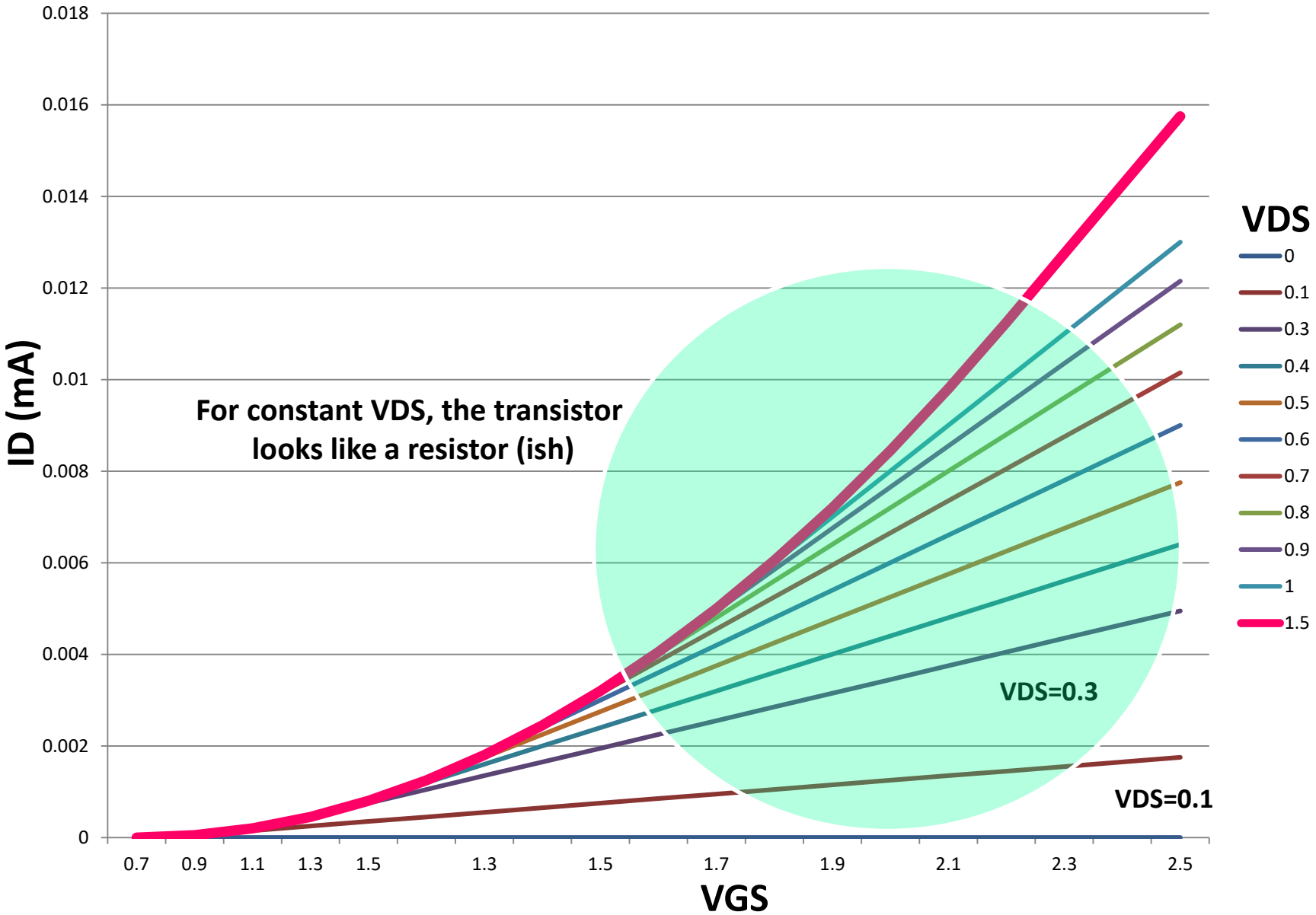
What we are looking for is something that we can use in these models... Can we get them to do something interesting or useful?

So lets look at the plots and see if anything jumps out (straight lines are always useful).

Plotting this equation



Plotting this equation (different X Axis)



Now we are going to use the different behaviours for small and large signals to make amplifiers

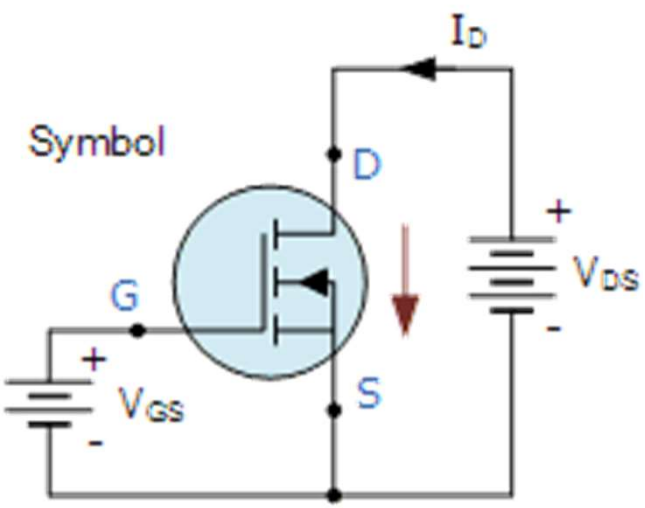
EE204FZ
Lecture 6
Models & Biasing
Supplementary Information

Zhu DIAO

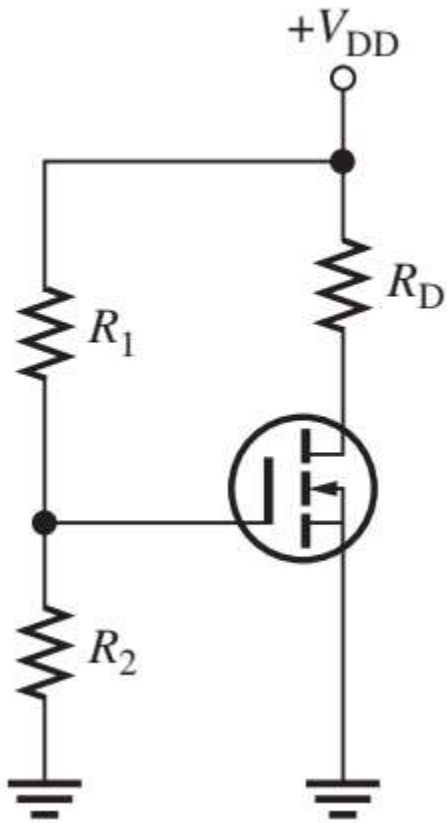
Email: zhu.diao@mu.ie

Typical Biasing Schemes for MOSFETs

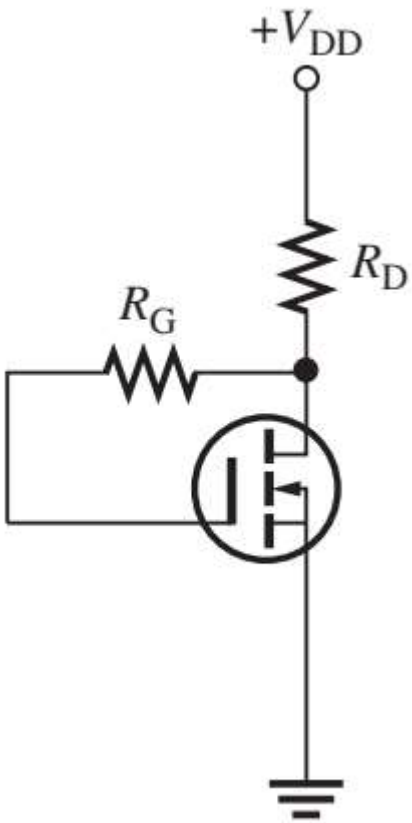
Fixed biasing



Voltage-divider biasing

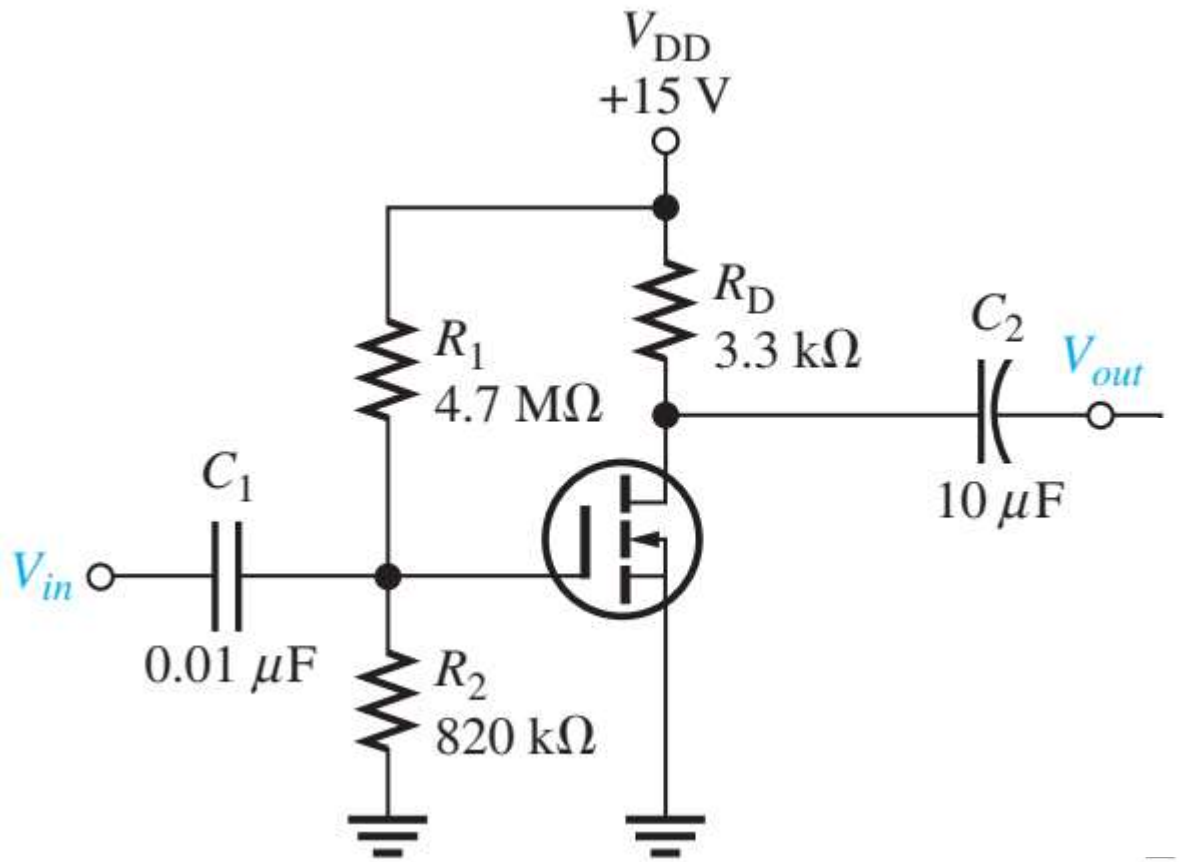


Drain-feedback biasing



Small Signal?

- Capacitors can be used to decouple the DC bias.

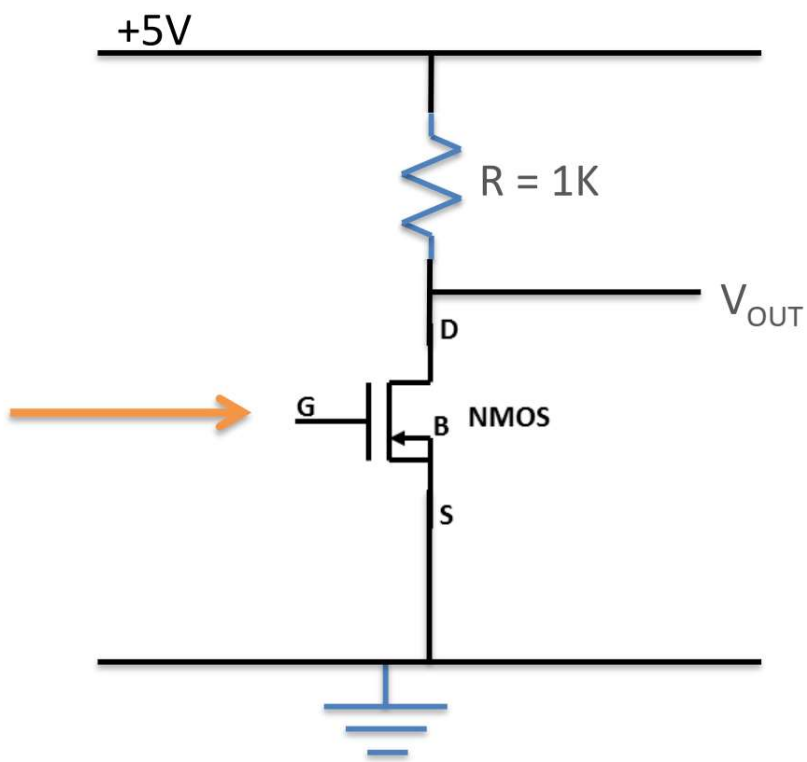


EE204FZ
Lecture 7
FET Amplifiers

Zhu DIAO

Email: zhu.diao@mu.ie

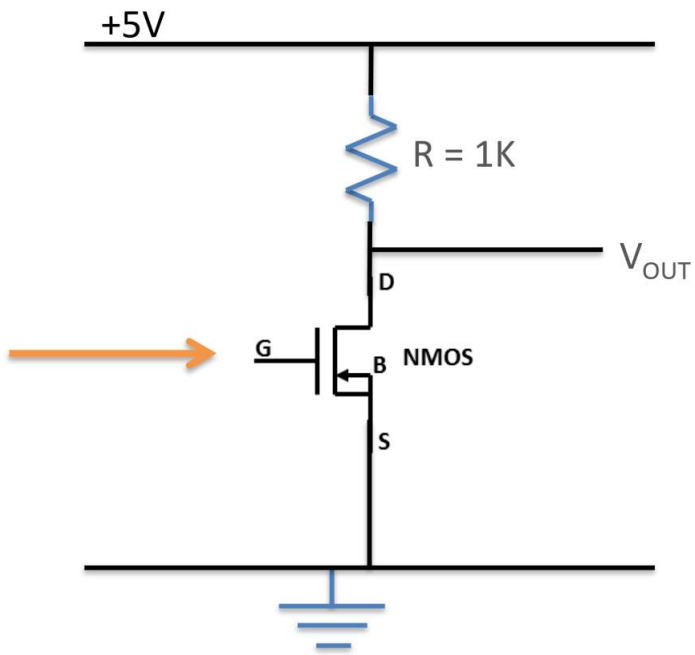
Amplification



$$I_{\mathrm{D(sat)}} = \frac{1}{2}\mu_{\mathrm{n}} C_{\mathrm{ox}} \frac{W}{L}(V_{\mathrm{GS}} - V_{\mathrm{T}})^2$$

- Let's imagine that the ideal MOSFET is in saturation, V_{DS} has very little effect on current flows, but V_{GS} sets the current.
  - What happens to I_D if one increases V_{GS}?
  - What happens then to V_{OUT}?

Amplification



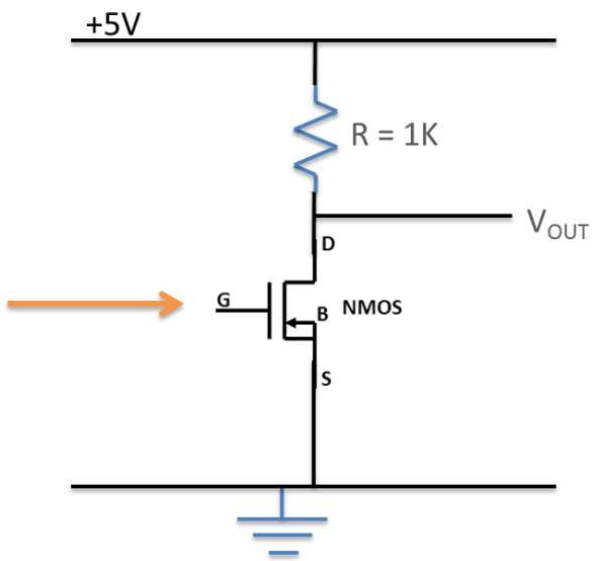
Assume that

- $\mu_n C_{ox} = 50 \times 10^{-6}\ \mathrm{A/V^2}$;
- W/L = 100;
- V_T = 1 V.

$$I_{D(sat)} = (0.0025)\,(V_{GS} - 1)^2$$

- Let's imagine now that one has V_{GS} = 2 V, so $I_{D(sat)}$ = 2.5 mA.
  - What voltage is V_{OUT} at?
- If V_{GS} goes up 1% to 2.02 V,
  - What is I_D?
  - What value is V_{OUT}?

Amplification



$$I_{\mathrm{D(sat)}} = (0.0025)\,(V_{\mathrm{GS}} - 1)^2$$

$$\Delta I_{\mathrm{D(sat)}} \approx (0.0025)(2)(V_{\mathrm{GS}} - 1)\Delta V_{\mathrm{GS}}$$

$$\Delta V_{\mathrm{out}} \approx -(2.5)(2)(V_{\mathrm{GS}} - 1)\Delta V_{\mathrm{GS}}$$

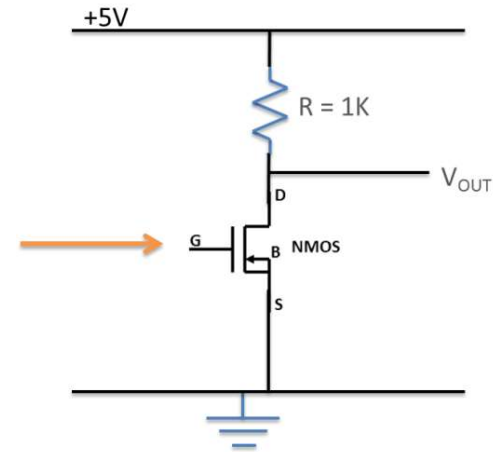
$$\frac{\Delta V_{\mathrm{out}}}{\Delta V_{\mathrm{GS}}} \approx -(2.5)(2) = -5$$

$$A \approx -5$$

- This is a bit of a hack, with assumptions everywhere, but basically we get an amplification of the change in the gate signal.
- More correctly, this needs to be worked out a bit more carefully.
- NOTE: Gain is usually given the symbol of *A* for Amplification.

Equations Are for Fun!

$$V_{\text{out}} = 5 - RI_{\text{D(Sat)}}$$
$$= 5 - R\frac{1}{2}\mu_{\text{n}}C_{\text{ox}}\frac{W}{L}(V_{\text{GS}} - V_{\text{T}})^2$$



Consider a small change in V_{GS}, now calculate the new V_{out}:

$$V_{\text{out(new)}} = 5 - R\frac{1}{2}\mu_{\text{n}}C_{\text{ox}}\frac{W}{L}(V_{\text{GS}} + \Delta V_{\text{GS}} - V_{\text{T}})^2$$

Let's find the change in V_{out} before doing all the multiplications, etc.

$$\Delta V_{\text{out}} = V_{\text{out(new)}} - V_{\text{out}}$$
$$= \left[5 - R\frac{1}{2}\mu_{\text{n}}C_{\text{ox}}\frac{W}{L}(V_{\text{GS}} + \Delta V_{\text{GS}} - V_{\text{T}})^2\right] - \left[5 - R\frac{1}{2}\mu_{\text{n}}C_{\text{ox}}\frac{W}{L}(V_{\text{GS}} - V_{\text{T}})^2\right]$$
$$= \left[-R\frac{1}{2}\mu_{\text{n}}C_{\text{ox}}\frac{W}{L}(V_{\text{GS}} + \Delta V_{\text{GS}} - V_{\text{T}})^2\right] - \left[-R\frac{1}{2}\mu_{\text{n}}C_{\text{ox}}\frac{W}{L}(V_{\text{GS}} - V_{\text{T}})^2\right]$$

Equations Don't Need to be Remembered!

$$\Delta V_{\text{out}} = \left[-R\frac{1}{2}\mu_{\text{n}}C_{\text{ox}}\frac{W}{L}(V_{\text{GS}} + \Delta V_{\text{GS}} - V_{\text{T}})^2\right] - \left[-R\frac{1}{2}\mu_{\text{n}}C_{\text{ox}}\frac{W}{L}(V_{\text{GS}} - V_{\text{T}})^2\right]$$

$$= -R\frac{1}{2}\mu_{\text{n}}C_{\text{ox}}\frac{W}{L}[(V_{\text{GS}} + \Delta V_{\text{GS}} - V_{\text{T}})^2 - (V_{\text{GS}} - V_{\text{T}})^2]$$

Now the messy expansion. But let's just look at that bit:

$$(V_{\text{GS}} + \Delta V_{\text{GS}} - V_{\text{T}})^2$$

$$= (V_{\text{GS}})^2 + (\Delta V_{\text{GS}})^2 + (V_{\text{T}})^2 + 2\Delta V_{\text{GS}}V_{\text{GS}} - 2V_{\text{GS}}V_{\text{T}} - 2\Delta V_{\text{GS}}V_{\text{T}}$$

$$(V_{\text{GS}} + \Delta V_{\text{GS}} - V_{\text{T}})^2 - (V_{\text{GS}} - V_{\text{T}})^2$$

$$= [(V_{\text{GS}})^2 + (\Delta V_{\text{GS}})^2 + (V_{\text{T}})^2 + 2\Delta V_{\text{GS}}V_{\text{GS}} - 2V_{\text{GS}}V_{\text{T}} - 2\Delta V_{\text{GS}}V_{\text{T}}]$$

$$-[(V_{\text{GS}})^2 - 2V_{\text{GS}}V_{\text{T}} + (V_{\text{T}})^2]$$

Amplification

Let's start doing cancellation:

$$(V_{\mathrm{GS}} + \Delta V_{\mathrm{GS}} - V_{\mathrm{T}})^2 - (V_{\mathrm{GS}} - V_{\mathrm{T}})^2$$
$$= [\cancel{(V_{\mathrm{GS}})^2} + (\Delta V_{\mathrm{GS}})^2 + \cancel{(V_{\mathrm{T}})^2} + 2\Delta V_{\mathrm{GS}} V_{\mathrm{GS}} - \cancel{2V_{\mathrm{GS}} V_{\mathrm{T}}} - 2\Delta V_{\mathrm{GS}} V_{\mathrm{T}}]$$
$$-[\cancel{(V_{\mathrm{GS}})^2} - \cancel{2V_{\mathrm{GS}} V_{\mathrm{T}}} + \cancel{(V_{\mathrm{T}})^2}]$$
$$= 2\Delta V_{\mathrm{GS}} V_{\mathrm{GS}} - 2\Delta V_{\mathrm{GS}} V_{\mathrm{T}} + (\Delta V_{\mathrm{GS}})^2$$

Going all the way back, we get:

$$\Delta V_{\mathrm{out}} = (-R)\frac{1}{2}\mu_{\mathrm{n}} C_{\mathrm{ox}} \frac{W}{L} [2\Delta V_{\mathrm{GS}} V_{\mathrm{GS}} - 2\Delta V_{\mathrm{GS}} V_{\mathrm{T}} + (\Delta V_{\mathrm{GS}})^2]$$

The change in V_{out} is solely dependent on V_{GS}. Interesting.

Amplification

- Now for amplification:

$$\frac{\Delta V_{\text{out}}}{\Delta V_{\text{GS}}} = (-R)\frac{1}{2}\mu_{\text{n}}C_{\text{ox}}\frac{W}{L}\left(2V_{\text{GS}} - 2V_{\text{T}} + \Delta V_{\text{GS}}\right)$$

$$\approx (-R)\frac{1}{2}\mu_{\text{n}}C_{\text{ox}}\frac{W}{L}2(V_{\text{GS}} - V_{\text{T}})$$

Assume ΔV_{GS} is small.

- The amplification can be expressed as:

$$A \approx (-R)\,\mu_{\text{n}}C_{\text{ox}}\frac{W}{L}(V_{\text{pinch-off}})$$

$V_{\text{pinch-off}} = V_{\text{GS}} - V_{\text{T}}$

- After all that, we can conclude:
  - The amplification is negative, which means if the input signal goes up, the output goes down;
  - If in saturation, the gain is determined by the V_{GS} voltage only, and the value of resistance R.

Amplification

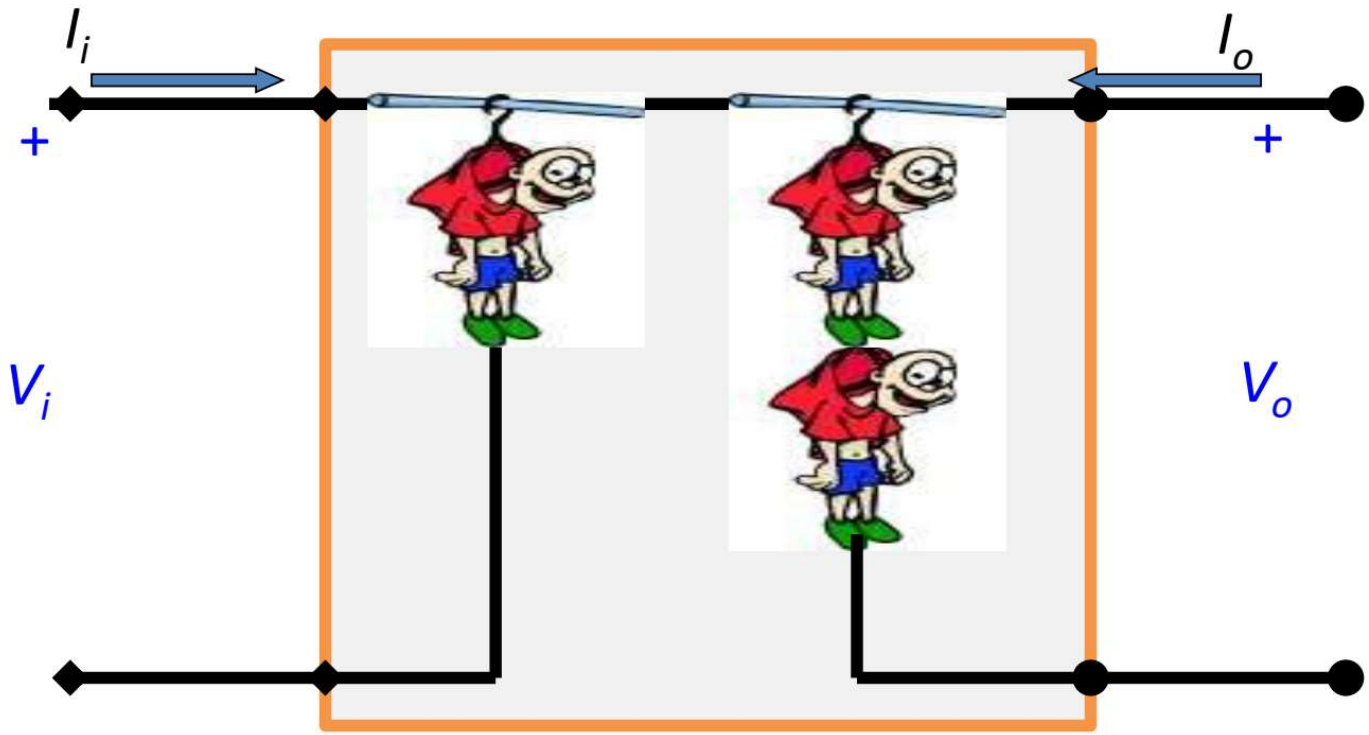
- And the biggest lesson of all is:

5 pages of maths later, I work out the equations for just one transistor and one resistor. How the *cough* *cough* do I manage something with more than that? Never mind thousands and millions of transistors.

The Trick

- What we need to do is to simplify our circuit as much as possible. To do that we split any circuit into two parts and deal with them one at a time.
  - The large signal (big voltages);
  - The small signals.
- Sometimes we only need to look at the big voltage swings, sometimes we need to focus on the small signals.
- For the small signals, we have a trick called "two-port networks", or black-box design.

Examples of a Black-Box



- It doesn't matter what is really inside the box, just that we can define ratios between input and output voltages, and input and output currents.

Two-Port Systems

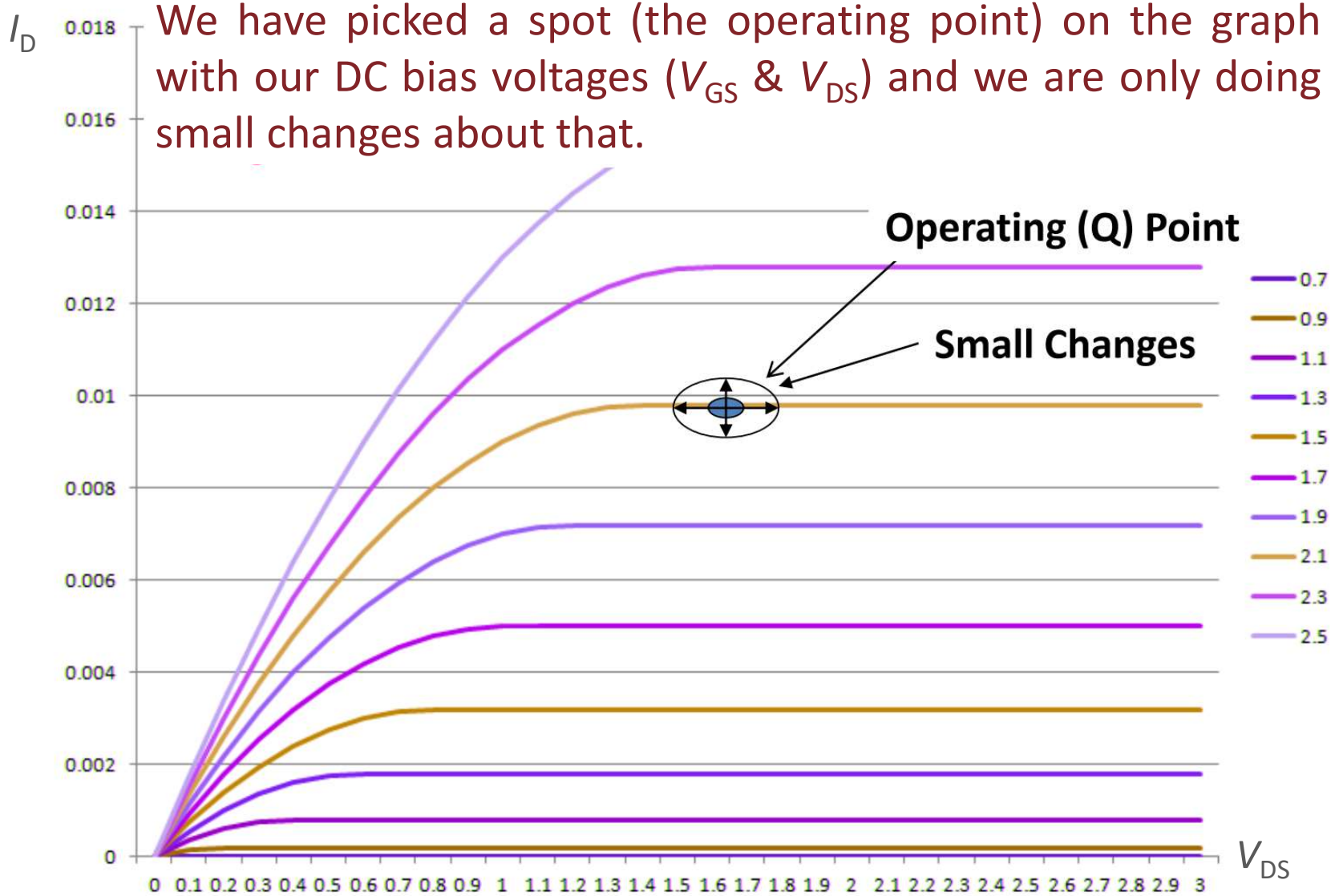
- There are many two-port "electrical" models that can be used to describe the transistor.
- Your choice of model often depends on the application that you need, large signal, small signal, whether you are looking for input or output impedance.
- For small signal the most common models are
  - H-parameter *
  - Hybrid-π *
  - T Equivalent Circuit
  - r_e Model
- For large signal
  - Ebers-Moll *
  - Charge Control Approach

Two-Port Systems

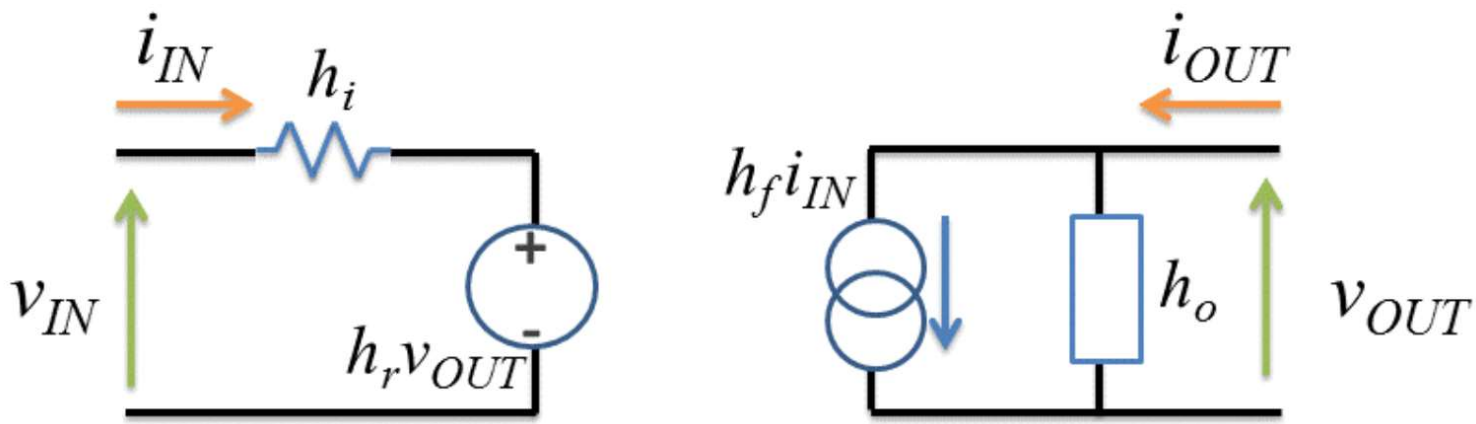
- In this course we are only concerned about SMALL SIGNALS.
- The key thing to remember is that amplification is normally small signals (often very very small) trying to get bigger.
- We are only going to look at 2 models which are related to each other:
  - H-parameter
  - Hybrid-π

Small Signal Analysis

We have picked a spot (the operating point) on the graph with our DC bias voltages (V_{GS} & V_{DS}) and we are only doing small changes about that.



H-Parameter Model



where

$$v_{\mathrm{IN}} = h_i(i_{\mathrm{IN}}) + h_r(v_{\mathrm{OUT}})$$
$$i_{\mathrm{OUT}} = h_f(i_{\mathrm{IN}}) + h_o(v_{\mathrm{OUT}})$$

Note:

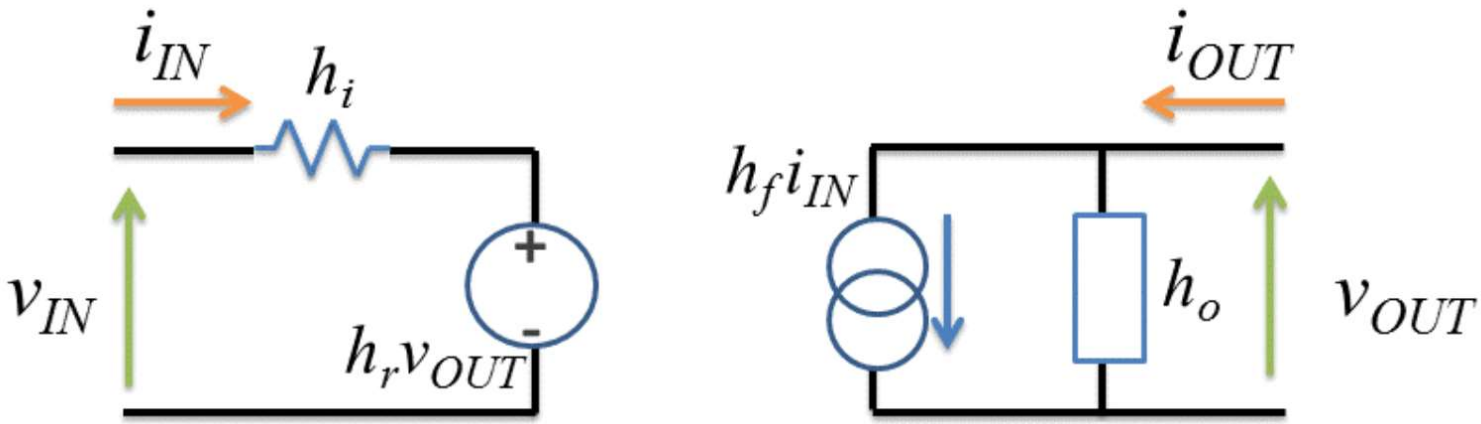
- Capital letters refer to large signals and DC;
- Small letters refer to small signals.

Reading the Expression

$$v_{\mathrm{IN}} = h_i(i_{\mathrm{IN}}) + h_r(v_{\mathrm{OUT}})$$
$$i_{\mathrm{OUT}} = h_f(i_{\mathrm{IN}}) + h_o(v_{\mathrm{OUT}})$$

- $h_i = \frac{v_{\mathrm{IN}}}{i_{\mathrm{IN}}}\Big|_{v_{\mathrm{OUT}}=0}$: h_i is the value of v_{IN} divided by i_{IN} when $v_{\mathrm{OUT}} = 0$;
- $h_o = \frac{i_{\mathrm{OUT}}}{v_{\mathrm{OUT}}}\Big|_{i_{\mathrm{IN}}=0}$: h_o is the value of i_{OUT} divided by v_{OUT} when $i_{\mathrm{IN}} = 0$.

Defining the H-Parameters



$$h_i = \frac{v_{\mathrm{IN}}}{i_{\mathrm{IN}}}\Big|_{v_{\mathrm{OUT}}=0}$$

Input Resistance (ohm)

$$h_r = \frac{v_{\mathrm{IN}}}{v_{\mathrm{OUT}}}\Big|_{i_{\mathrm{IN}}=0}$$

Reverse Transfer Voltage Ratio

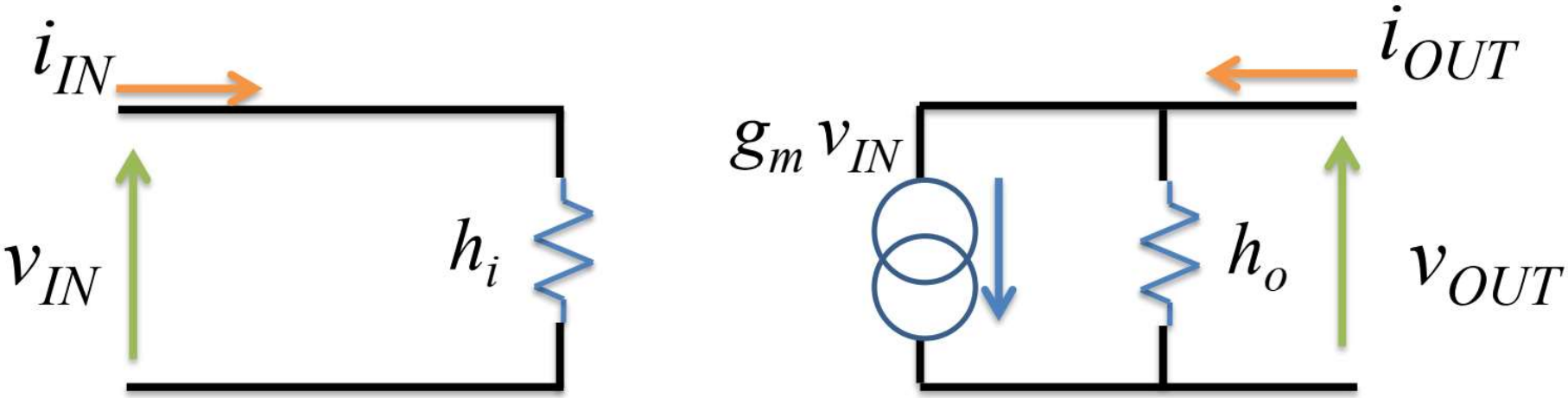
$$h_f = \frac{i_{\mathrm{OUT}}}{i_{\mathrm{IN}}}\Big|_{v_{\mathrm{OUT}}=0}$$

Forward Transfer Current Ratio

$$h_o = \frac{i_{\mathrm{OUT}}}{v_{\mathrm{OUT}}}\Big|_{i_{\mathrm{IN}}=0}$$

Output Conductance or Admittance (1/ohm)

Hybrid- π Parameter Model

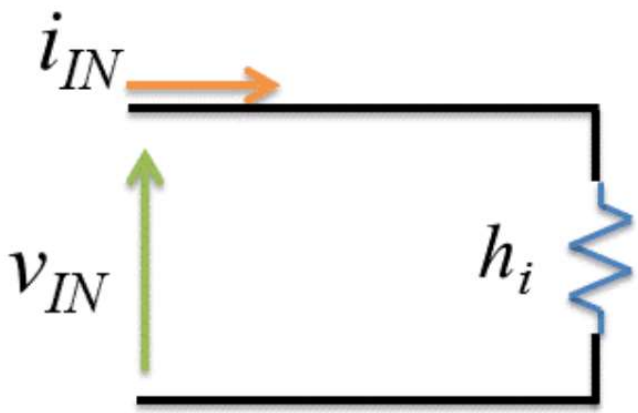


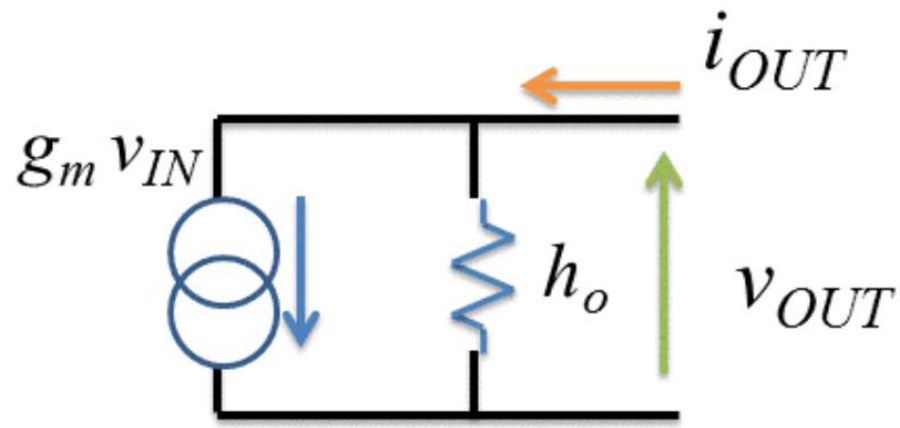
Where

$$i_{\mathrm{OUT}} = g_m(v_{\mathrm{IN}}) + (v_{\mathrm{OUT}})/h_o$$

This is very similar to the h-parameter model except h_o is drawn as a resistor, and the current source is now controlled by v_{IN}.

Defining the Hybrid- π Parameters





$$g_m = \frac{i_{\mathrm{OUT}}}{v_{\mathrm{IN}}}\Big|_{v_{\mathrm{OUT}}=0}$$

Forward Transconductance

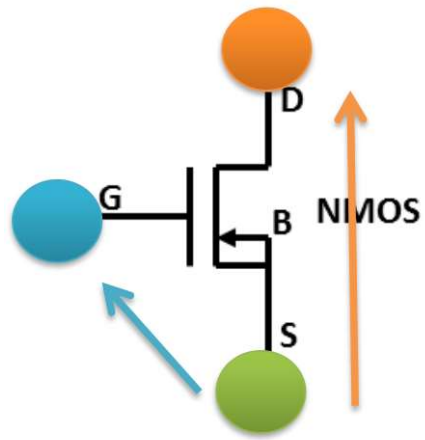
$$h_o = \frac{v_{\mathrm{OUT}}}{i_{\mathrm{OUT}}}\Big|_{v_{\mathrm{IN}}=0}$$

Output Resistance (ohm)

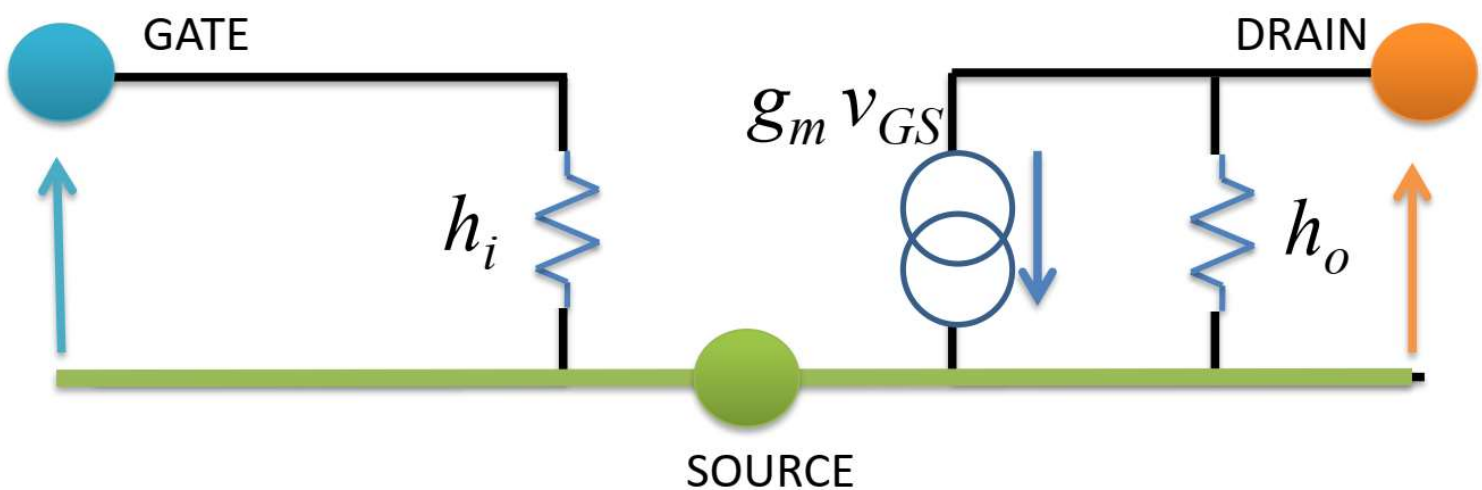
$$h_i = \frac{v_{\mathrm{IN}}}{i_{\mathrm{IN}}}\Big|_{v_{\mathrm{OUT}}=0}$$

Input Resistance (ohm)

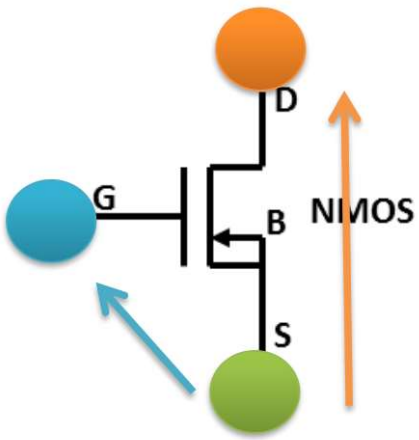
Applying the Hybrid-π Model to a MOSFET



We swap the transistor for the small signal model then carry out our analyses.

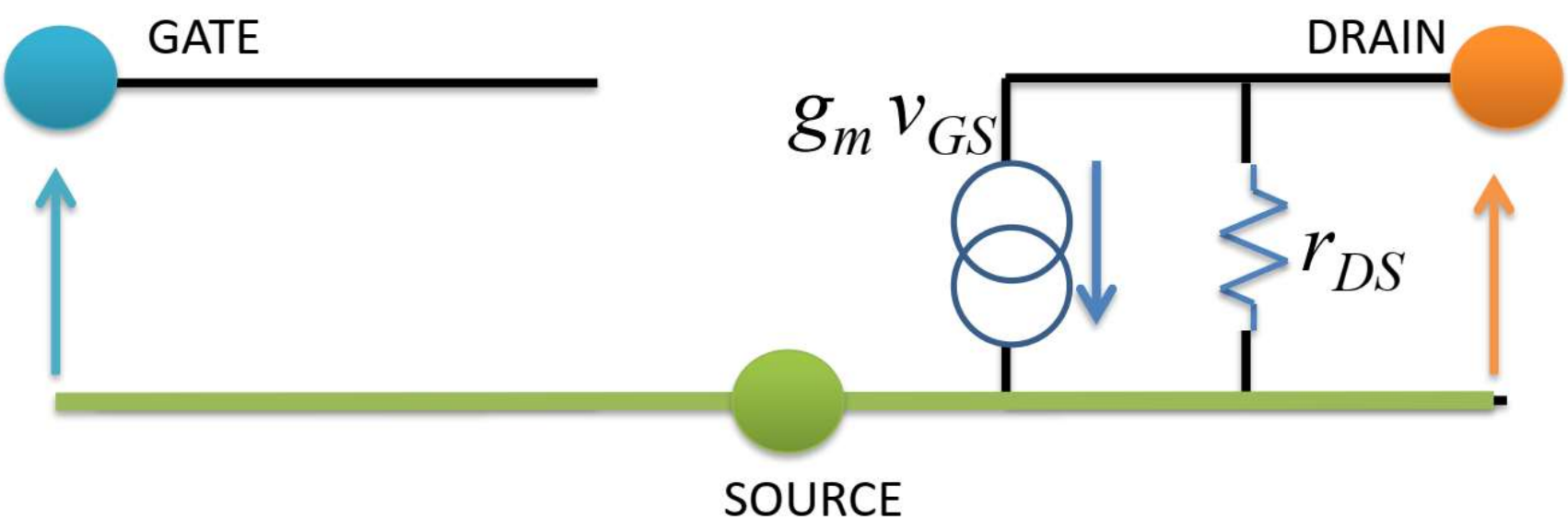


Applying the Hybrid-π Model to a MOSFET

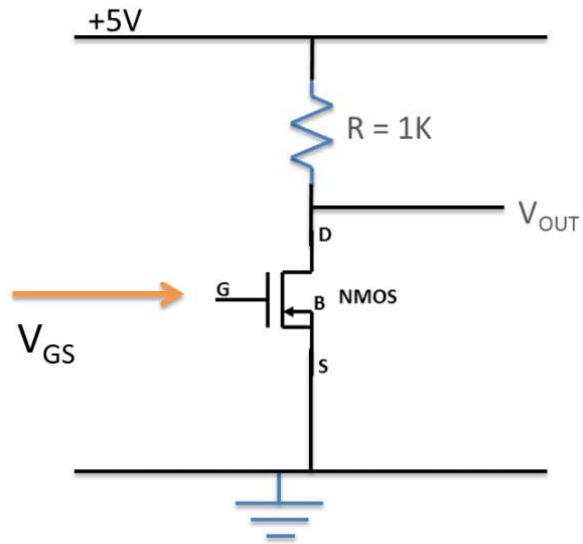


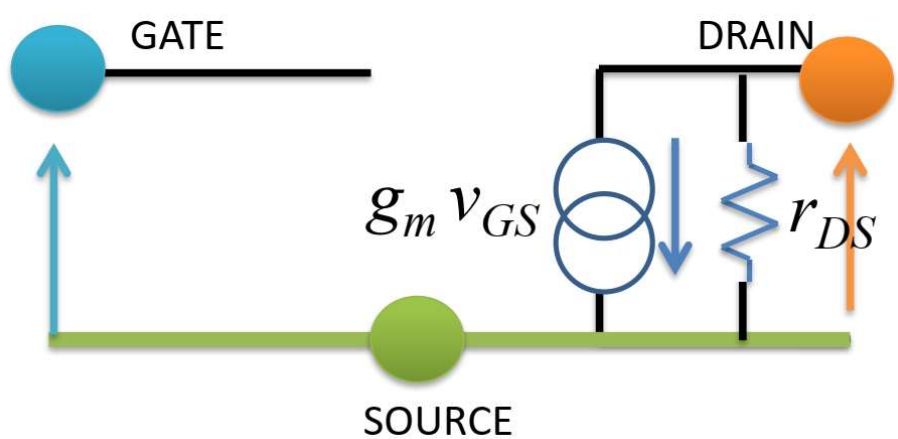
The input resistance of a MOSFET gate is infinite.

The small signal variation of a MOSFET in saturation is given by the slope of the flat-part of the curve, it is called r_{DS} and is either infinite (ideal MOSFETs), or very high (real MOSFETs).

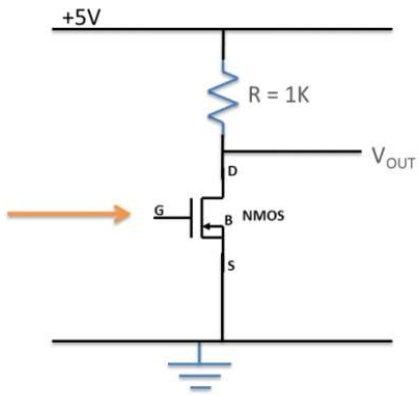


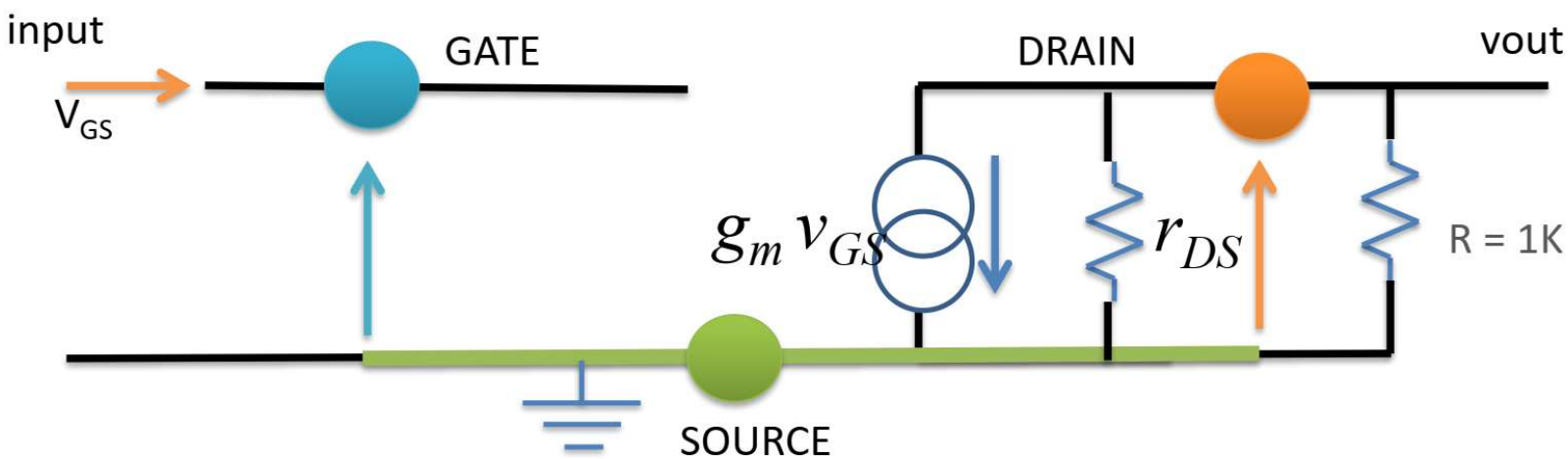
Example: Small Signal Equivalent Circuits





Example: Small Signal Equivalent Circuits





Rules When Constructing Equivalent Small Signal Circuits

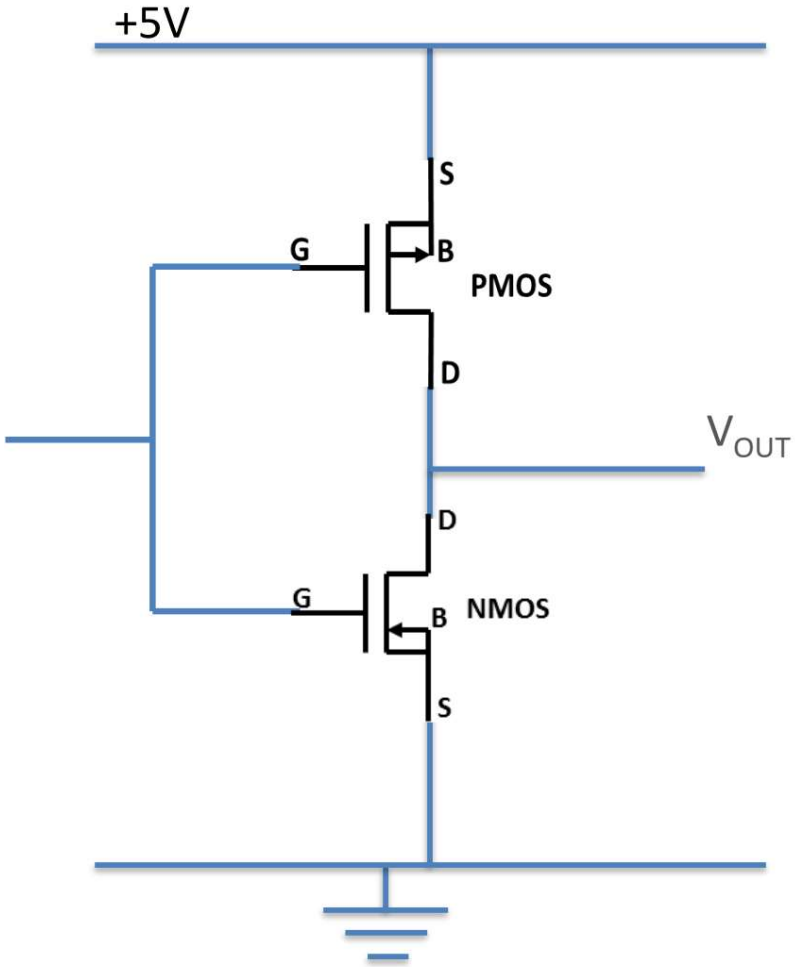
- Start with the transistor model;
- Identify the gate, drain and source;
- All fixed voltages are treated as small-signal grounds. That means the +5 V is also a ground. Think why this is true!!!
- Connect things up, remembering the "new" grounds are all the same place;
- PMOS and NMOS have the same small signal model, perhaps with different values of parameters;
- r_{DS} can often be ignored but it is best to leave it in.

Important Note

- Solving small signal equivalent circuits is much easier than doing a normal transistor analysis;
- However, the results are ONLY ACCURATE for small signals (AC);
- To understand what is happening with your DC signals, you need to carry out a DC analysis of your circuit as well;
- If your AC signals become large, then they'll work sort of like what you expected from the small signal analysis, but they will start experiencing different amplification (usually less) as they move away from your operating point.

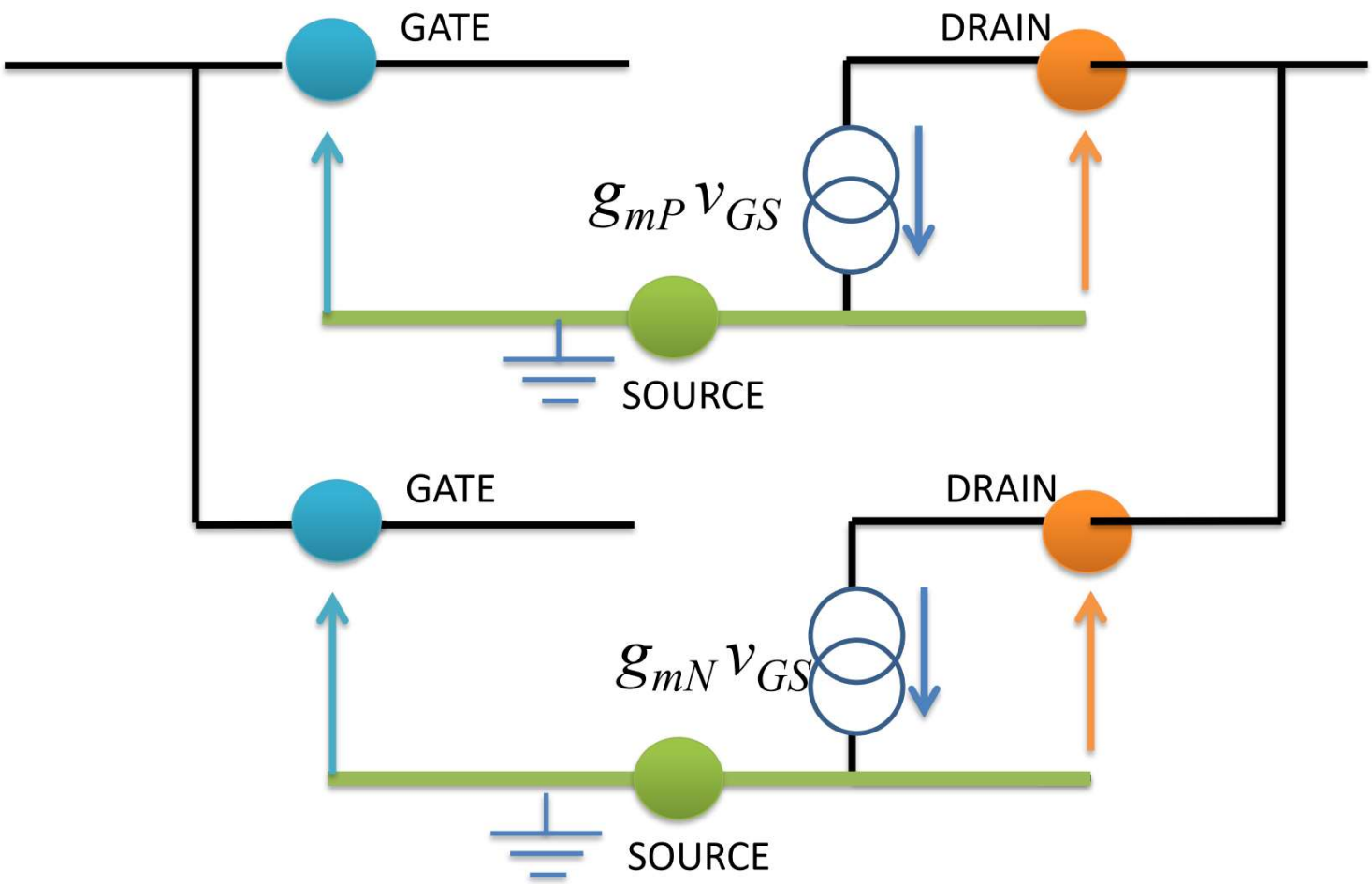
Example: The Invertor

- For fun, draw the small signal equivalent circuit of an inverter.



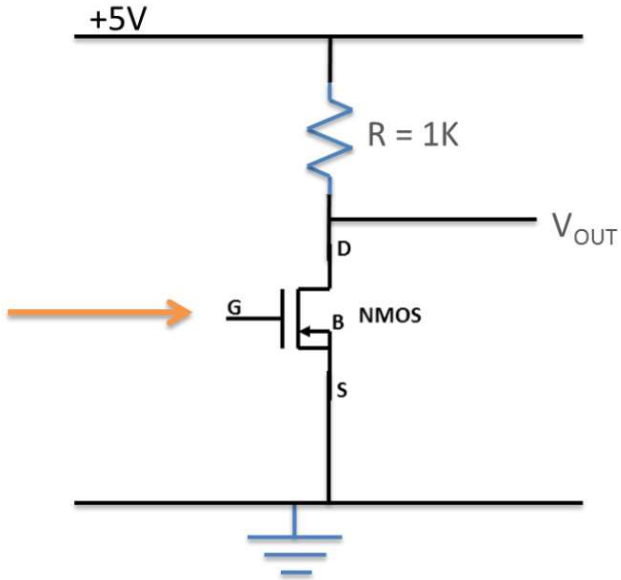
Example: The Invertor

- Only problem is that they are in parallel.



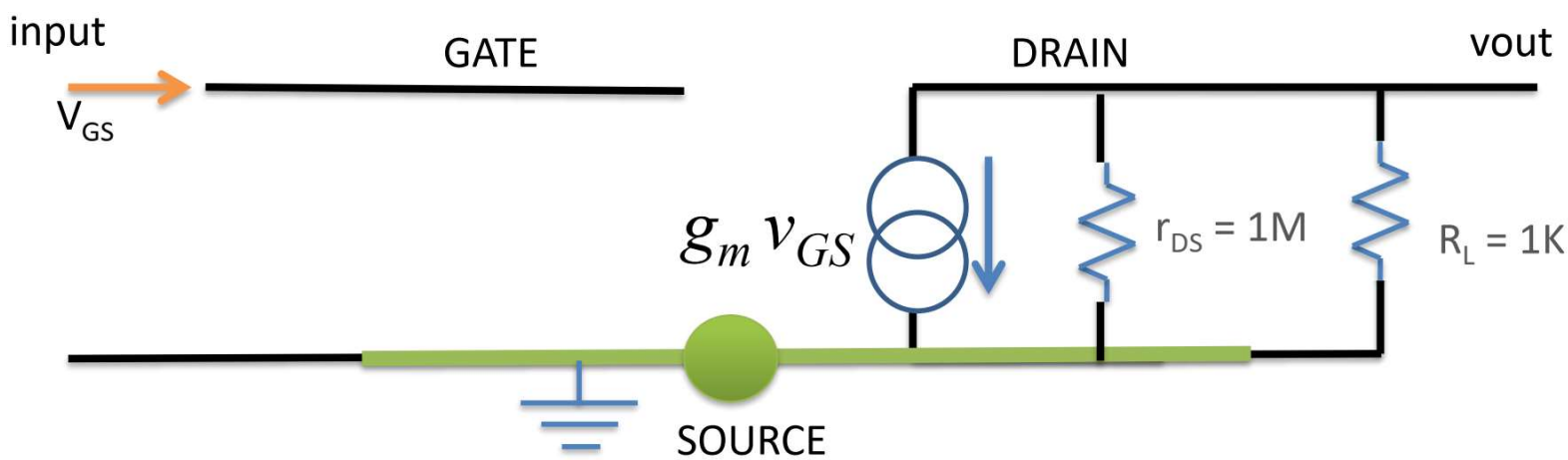
Example: Let's Calculate Gain

- A ZVN3306 n-MOSFET was used with a 1 kohm load resistor. Assuming that we are in saturation and dealing with small signals, calculate the gain a small signal would experience.

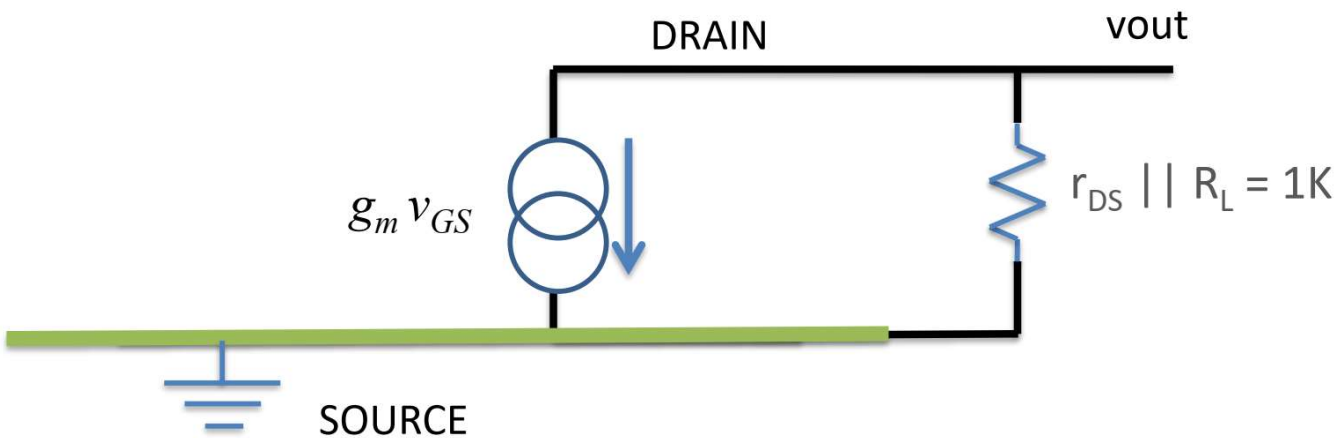


- The datasheet gives the following values:
  - g_m = 0.15 Siemens;
  - r_{DS} > 1 Mohm.

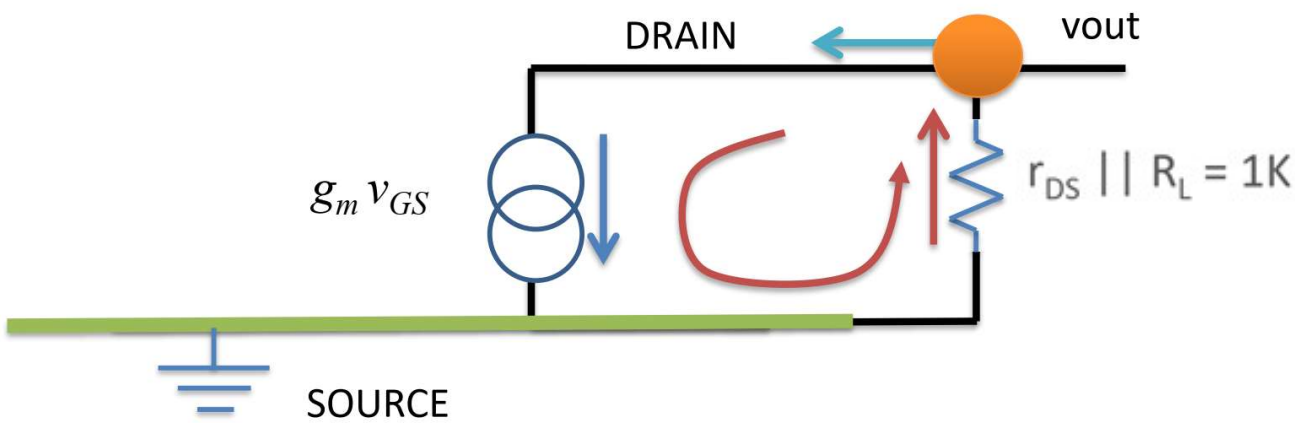
Example: Let's Calculate Gain



- r_{DS} is much much bigger than R_L, so it can be ignored.
- Secondly, we only need to look at the right-hand side as nothing is happening on the left-hand side.



Example: Let's Calculate Gain



- Now there is nothing connected to V_{OUT}, so all the current leaving the ORANGE node must be matched with that entering:

$$I_R = -g_m v_{GS}$$

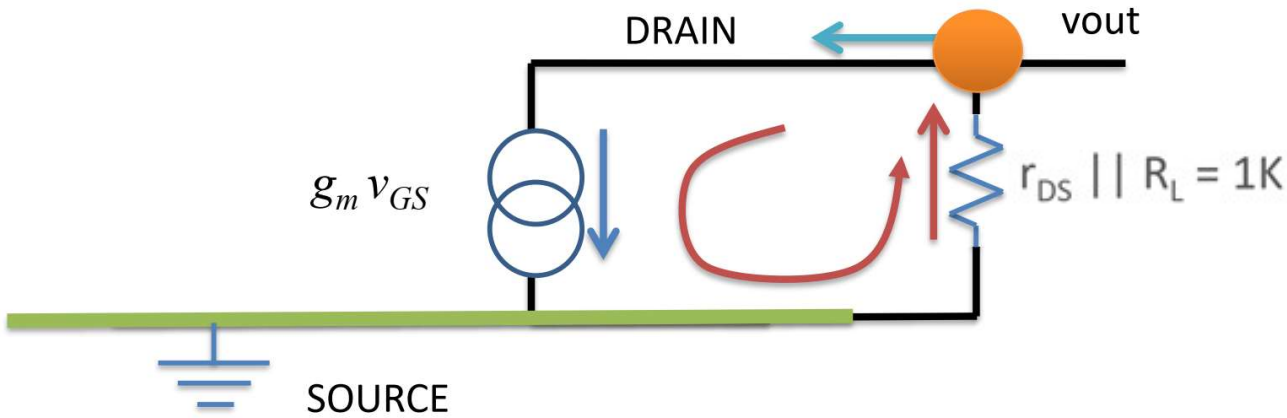
- The output small signal v_{out}:

$$v_{out} = -R_L I_R = -R_L g_m v_{GS}$$

- The gain is therefore:

$$A = \frac{v_{out}}{v_{GS}} = \frac{-R_L g_m v_{GS}}{v_{GS}} = -R_L g_m$$

Example: Let's Calculate Gain



$$A = \frac{v_{\text{out}}}{v_{\text{GS}}} = \frac{-R_{\text{L}} g_{\text{m}} v_{\text{GS}}}{v_{\text{GS}}} = -g_{\text{m}} R_{\text{L}}$$

- Plugging in the numbers we get:

$$A = -(1000)(0.15) = -150$$

- This is a large gain, and if you get too much gain, you might experience clipping.

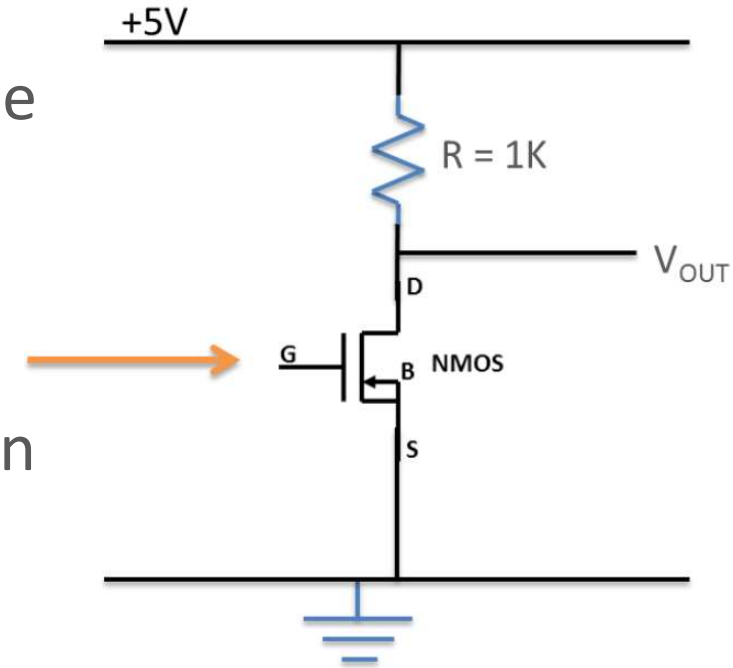
Clipping: A Small Signal Failure

- The DC value of V_{OUT} is given by the DC values for the gate voltage:

$$I_{D(sat)} = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS} - V_T)^2$$

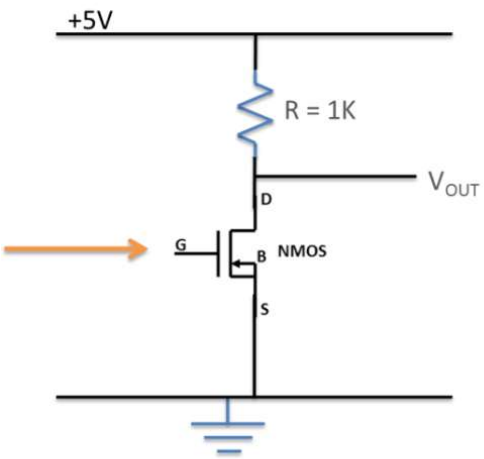
- The value of V_{OUT} is therefore given by:

$$V_{OUT} = 5 - R_L I_{D(sat)}$$

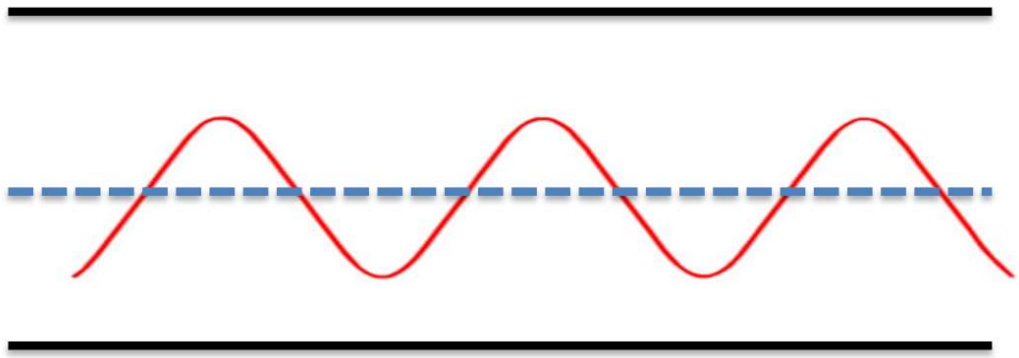


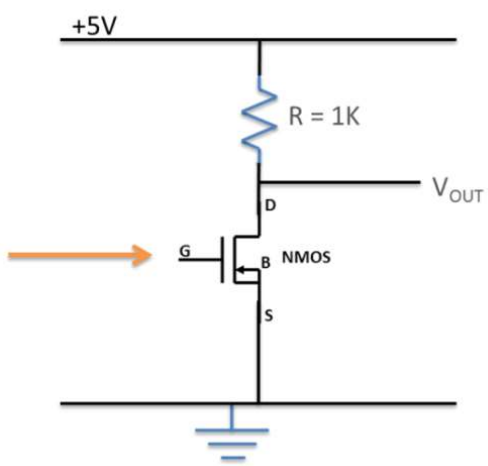
- What is the biggest and smallest values V_{OUT} can take?
- If we get a signal on the gate that is amplified and the output goes too high, what will happen?

Clipping

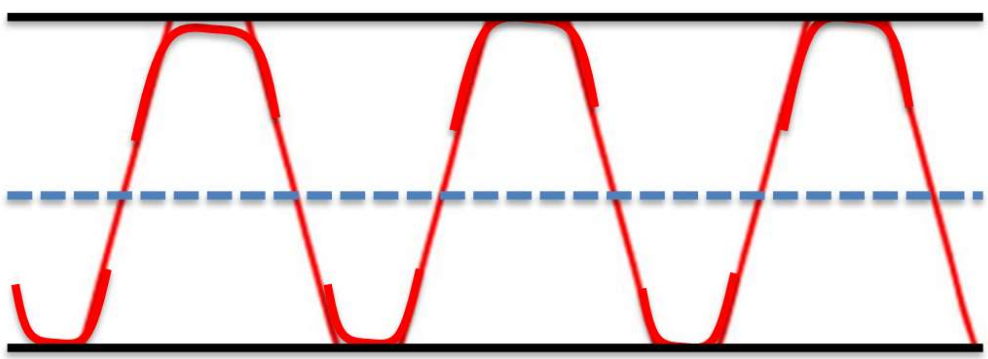


Using a small signal.

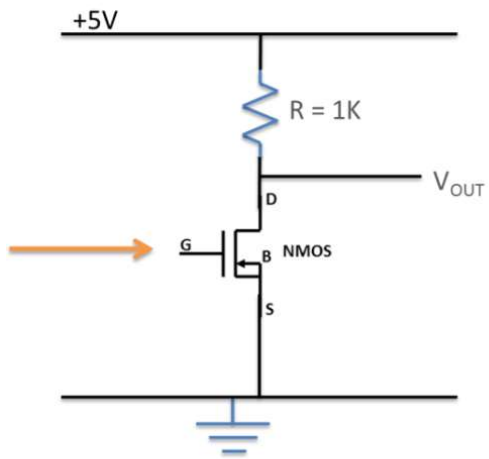




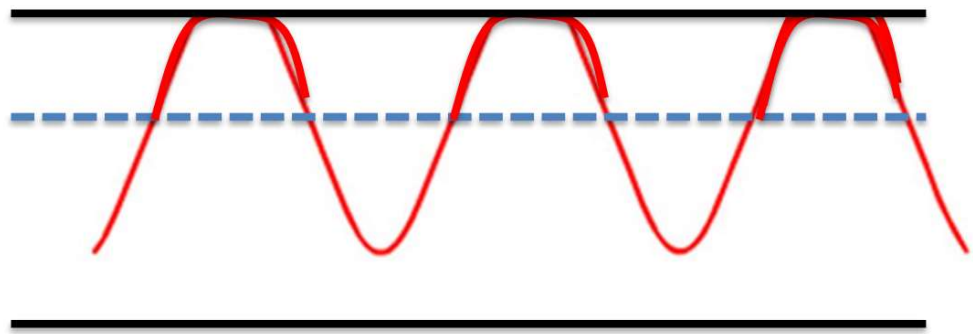
Signal gets clipped at the top and bottom.

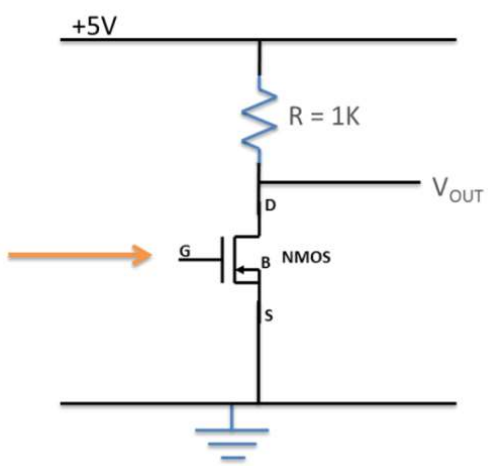


Clipping

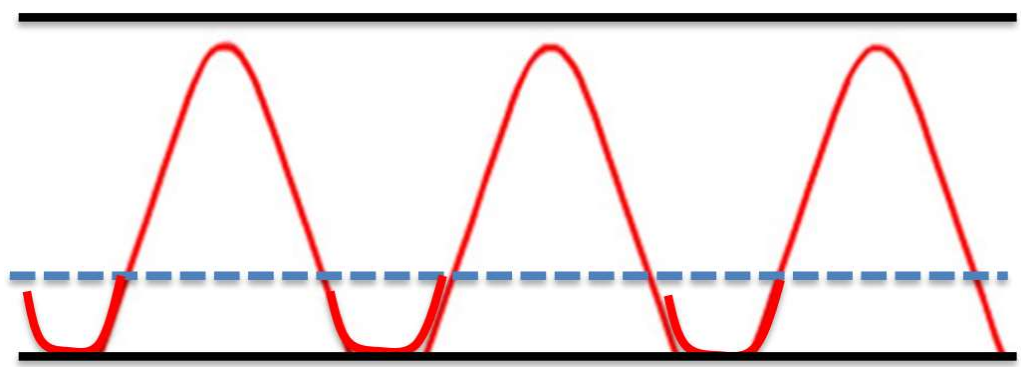


Centre V_{OUT} on the high side, clipping on one side.

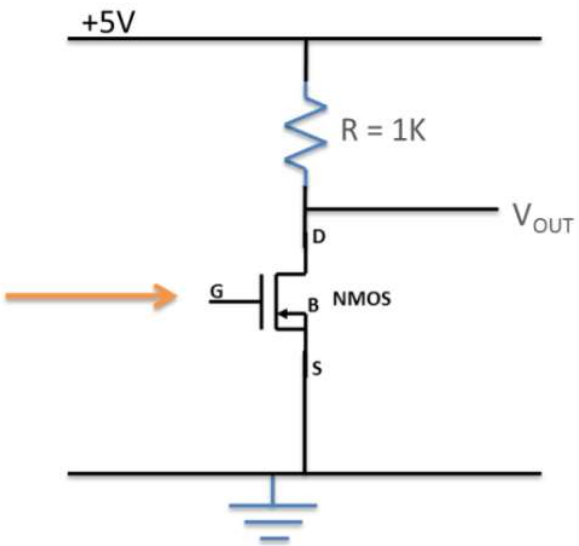


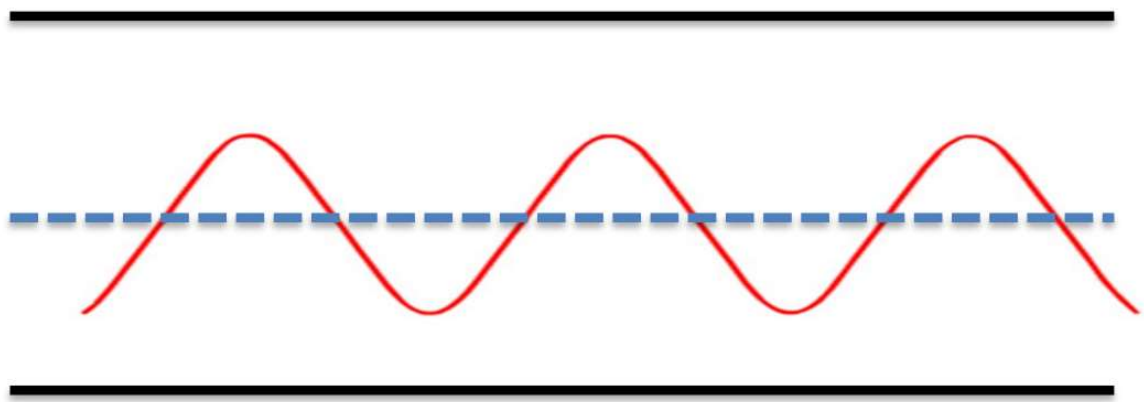


Centre V_{OUT} on the low side, clipping on the other side.



What's a Good "Small Signal"?

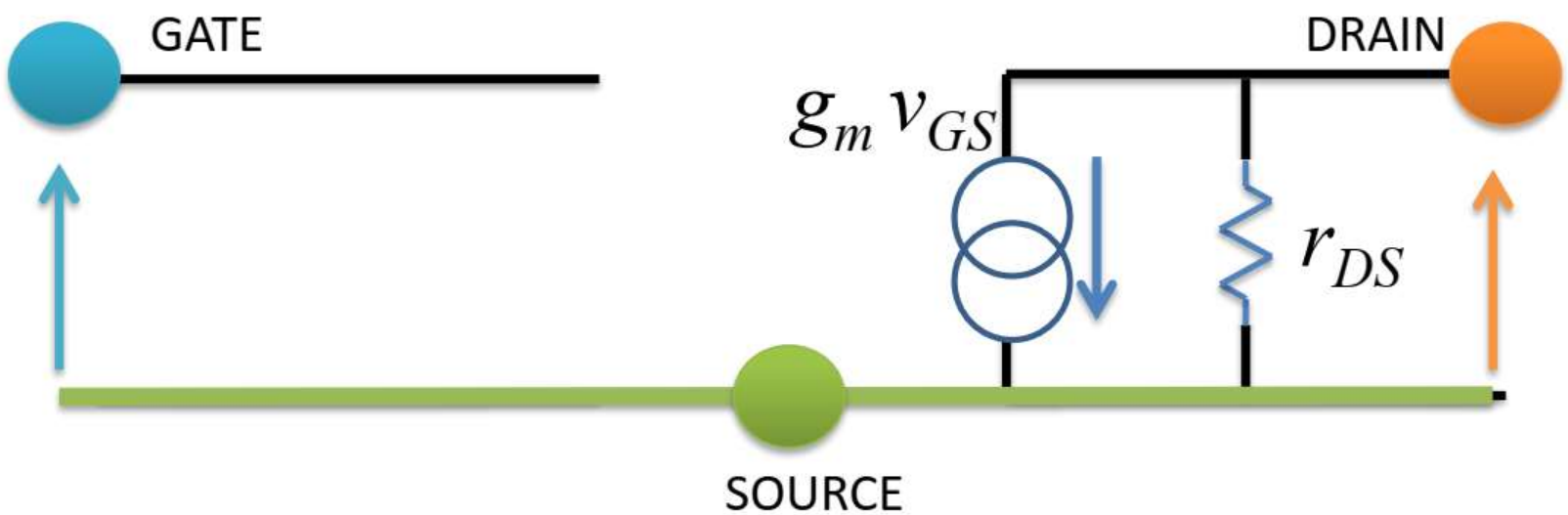




- In an audio microphone application, a small signal is ~10 mV.
- In a mobile phone receiver, a large signal is 10 mV, a small signal is 1 µV.
- In these cases, you can use a gain of 100 and still have small enough for the output signal and no clipping. It all depends on your application.
- You can use special pre-amps or low-noise amps that concentrate on very weak signals and boost them a little, so you can use other circuits that concentrate on producing larger, more powerful outputs… sensitivity *vs.* power.

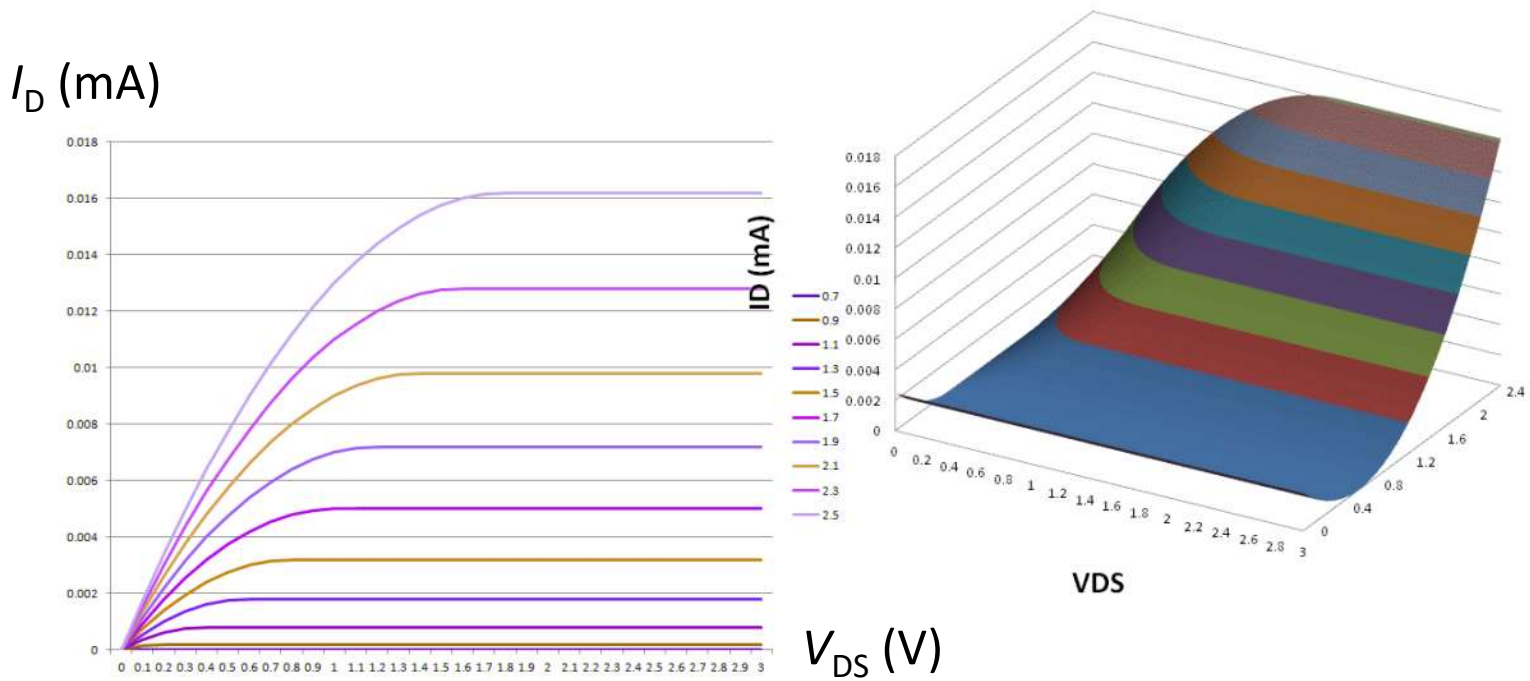
The Importance of Bias

- Biasing is important. It controls
  - When the output will clip;
  - The amplification we can achieve as it controls g_m;
  - It also controls the value of r_{DS}.



What is Really Biasing?

- A lot of it is picking where on the transistor curve we want to be. And that means choosing values of V_{GS}, I_D and V_{DS}.



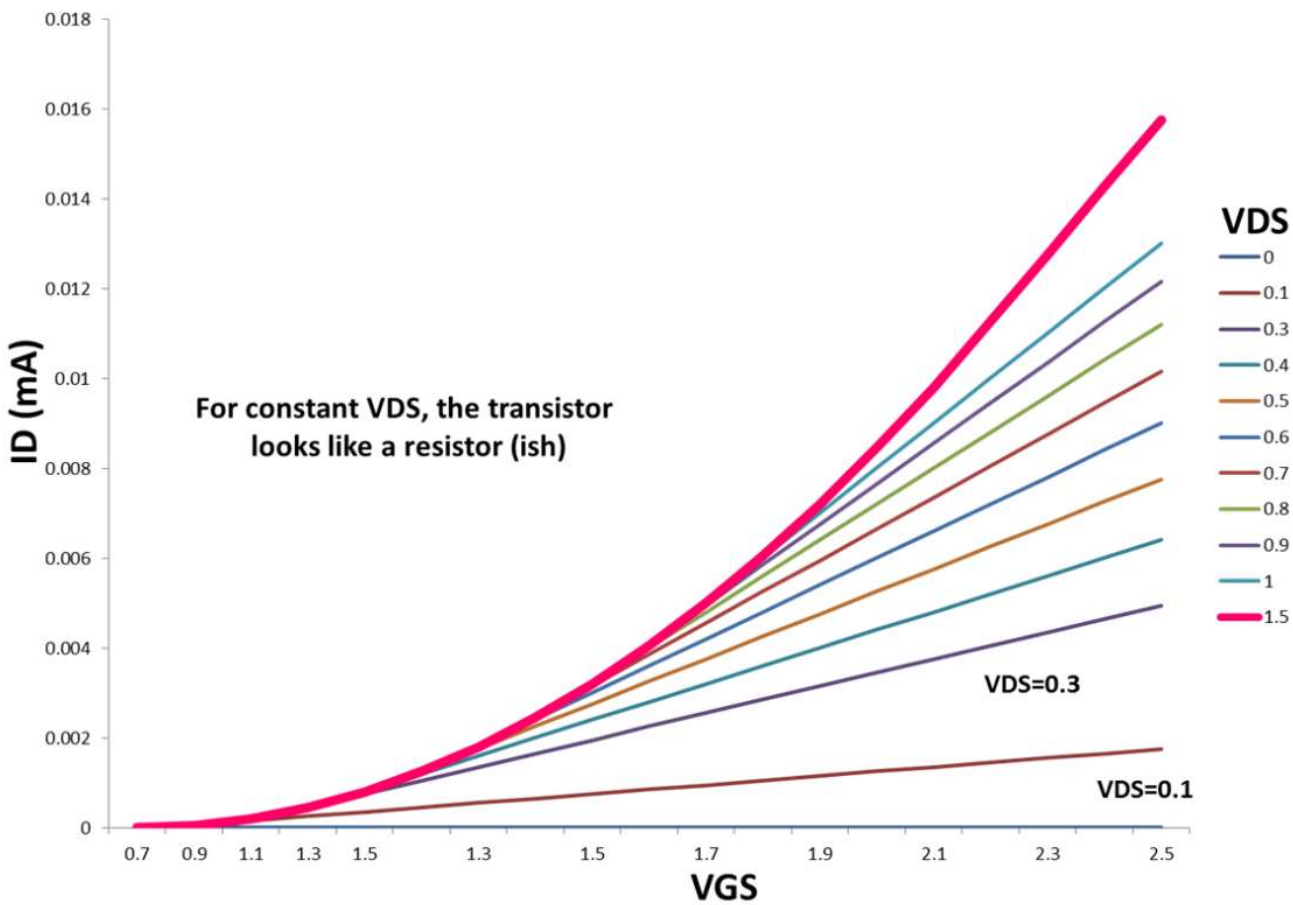
- In the analogy, biasing is picking what part of the mountain I want to walk a few steps back and forth on (saying go 3 km to the east and 500 metres up).

What is Really Biasing?

- Biasing uses DC voltages to control the values of small-signal parameters.
- The next few slides have a bit of maths but the important thing is the relationships between the key small-signal parameters and the DC values:
- $I_{D(\mathrm{sat})} = \frac{1}{2}\mu_{\mathrm{n}} C_{\mathrm{ox}} \frac{W}{L}(V_{\mathrm{GS}} - V_{\mathrm{T}})^2$
- $g_m = \mu_{\mathrm{n}} C_{\mathrm{ox}} \frac{W}{L}(V_{\mathrm{GS}} - V_{\mathrm{T}}) = \sqrt{2\mu_{\mathrm{n}} C_{\mathrm{ox}} \frac{W}{L} I_{D(\mathrm{sat})}} = k\sqrt{I_{D(\mathrm{sat})}}$
- $r_{\mathrm{DS}} = \frac{1}{\lambda I_{D(\mathrm{sat})}}$

Biasing and g_m

- $g_m = \frac{dI_D}{dV_{GS}}$, g_m is the local slope of the curve when I_D is plotted as a function of V_{GS}.



Biasing and g_m

- For this analysis, we ignore the channel length modulation (since it is a small effect):

$$I_{D(\text{sat})} = \frac{1}{2}\mu_\text{n} C_\text{ox} \frac{W}{L}(V_\text{GS} - V_\text{T})^2$$

- Let's look at the small-signal output:

$$g_m = \frac{dI_{\text{D(sat)}}}{dV_\text{GS}} = \frac{d}{dV_\text{GS}}\left[\frac{1}{2}\mu_\text{n} C_\text{ox} \frac{W}{L}(V_\text{GS} - V_\text{T})^2\right]$$

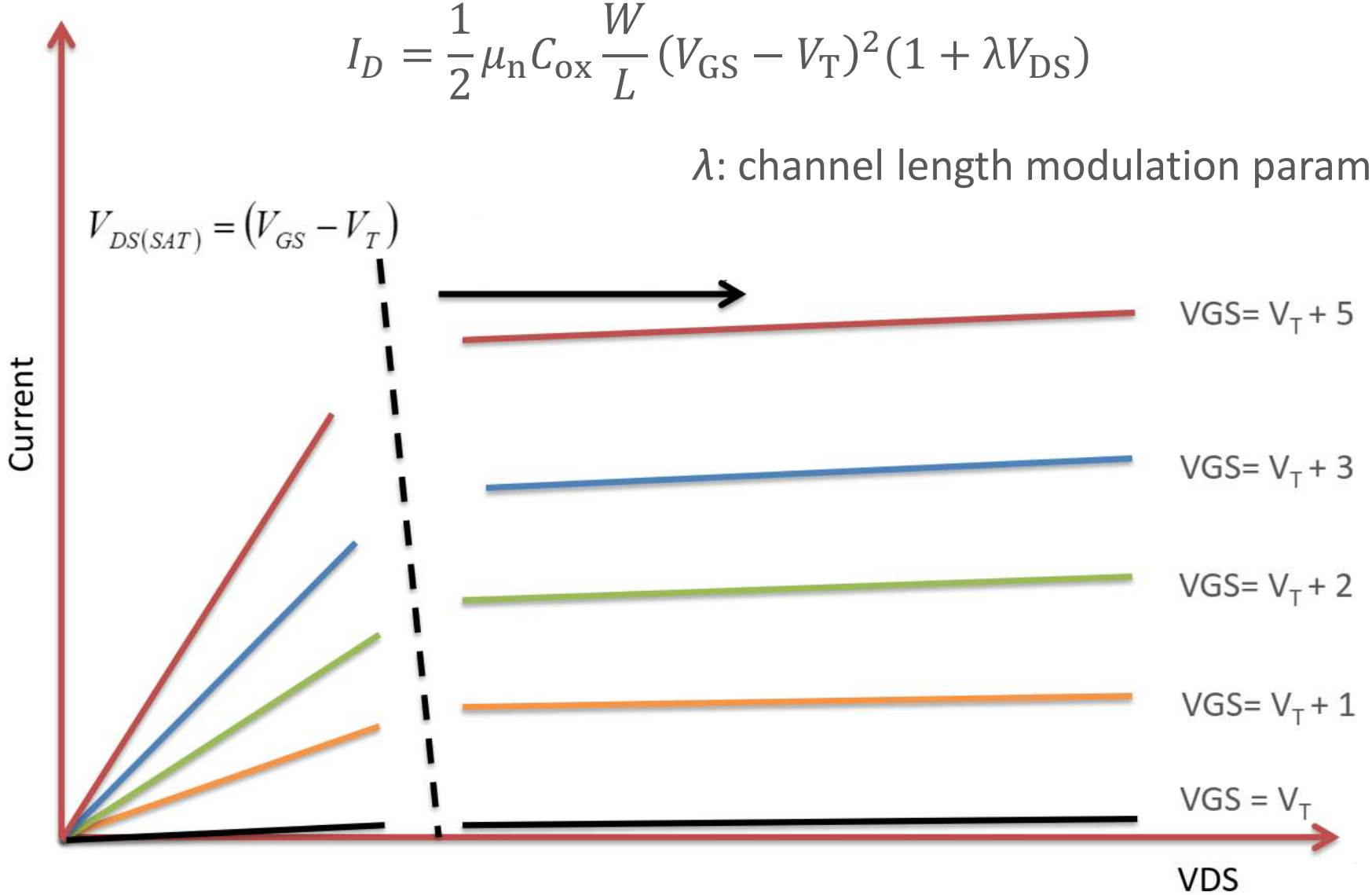
$$= \left(\frac{1}{2}\mu_\text{n} C_\text{ox} \frac{W}{L}\right) 2(V_\text{GS} - V_\text{T}) = \mu_\text{n} C_\text{ox} \frac{W}{L}(V_\text{GS} - V_\text{T})$$

$$= \sqrt{2\mu_\text{n} C_\text{ox} \frac{W}{L} I_{D(\text{sat})}} = k\sqrt{I_{D(\text{sat})}}$$

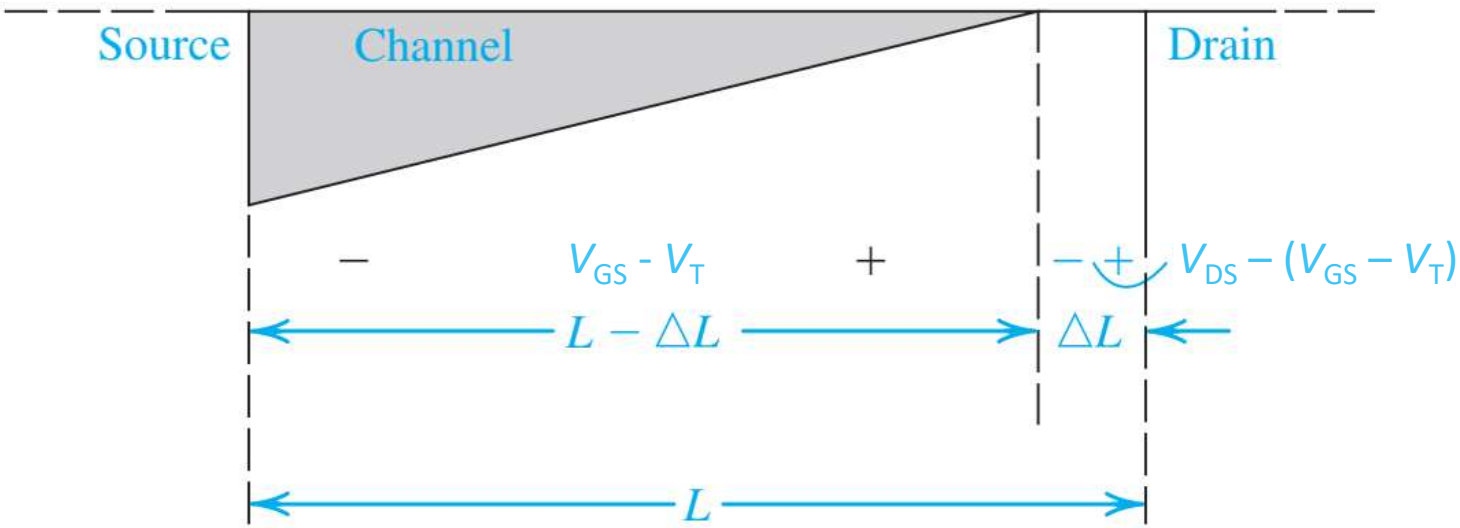
Biasing and r_{DS}

$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS} - V_T)^2(1 + \lambda V_{DS})$$

λ: channel length modulation parameter



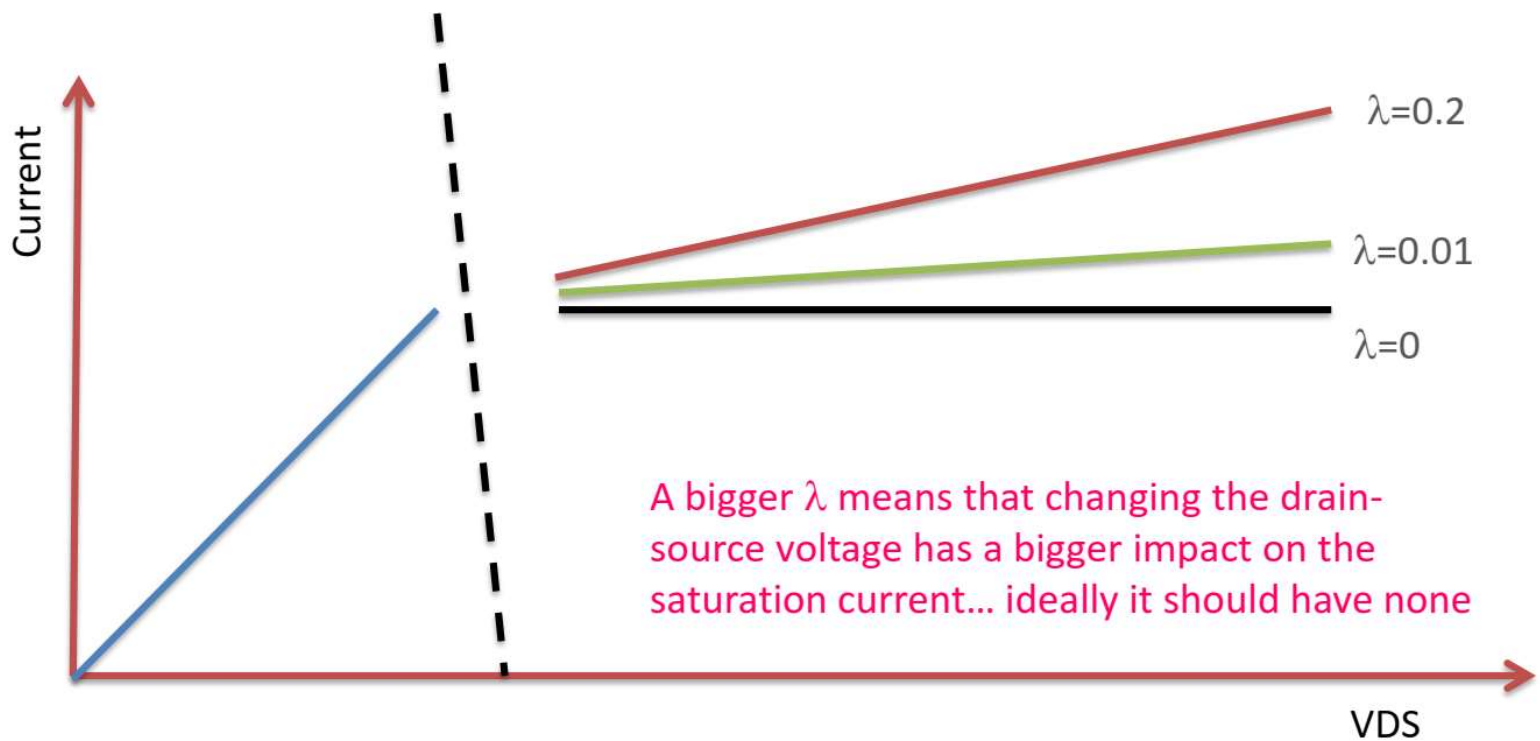
Recall Channel-Length Modulation



- In the saturation region, when V_{DS} further increases, the channel pinch-off point moves slightly away from the drain and towards the source;
- The effective channel length becomes $(L - \Delta L)$;
- The electric potential difference across the channel remains constant at $(V_{GS} - V_T)$. This is sometimes called the "overdrive voltage".

A Few Extra Words on λ

- λ is the "channel length modulation parameter" that indicates how far we are away from an ideal MOSFET device in saturation. It is a constant value that varies for each transistor type manufactured. It gets bigger for small devices, especially in modern processes. For modern devices it could be anywhere from 0.01 to 0.2 depending on device length (longer devices, smaller λ).



Back to r_{DS}

- Let's look at the small-signal output impedance:

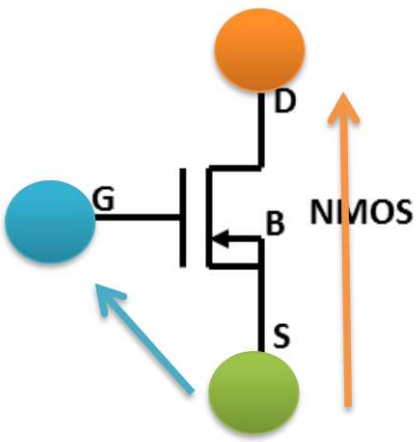
$$\frac{1}{r_{\mathrm{DS}}} = \frac{dI_{\mathrm{D}}}{dV_{\mathrm{DS}}} = \frac{d}{dV_{\mathrm{DS}}}\left[\frac{1}{2}\mu_{\mathrm{n}}C_{\mathrm{ox}}\frac{W}{L}(V_{\mathrm{GS}} - V_{\mathrm{T}})^2(1 + \lambda V_{\mathrm{DS}})\right]$$

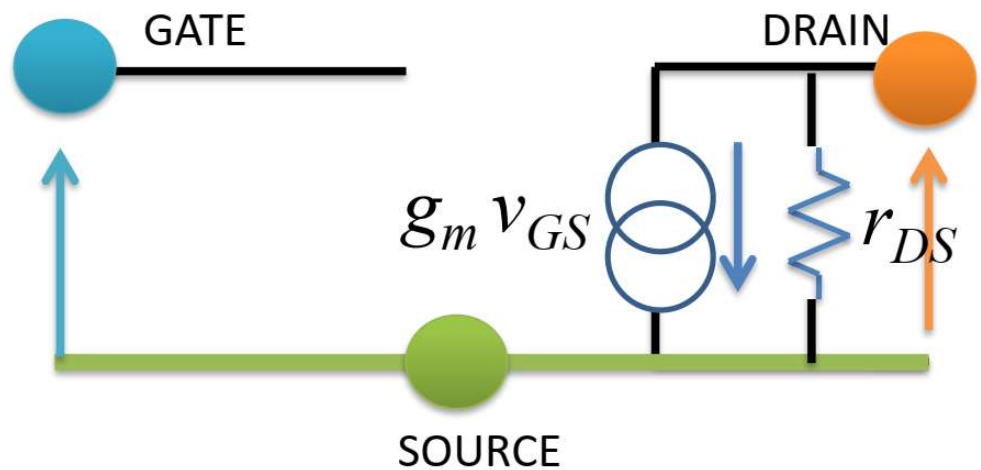
$$\frac{1}{r_{\mathrm{DS}}} = \lambda\frac{1}{2}\mu_{\mathrm{n}}C_{\mathrm{ox}}\frac{W}{L}(V_{\mathrm{GS}} - V_{\mathrm{T}})^2 = \lambda I_{D(\mathrm{sat})}$$

$$r_{\mathrm{DS}} = \frac{1}{\lambda I_{D(\mathrm{sat})}}$$

and $I_{D(\mathrm{sat})} = \frac{1}{2}\mu_{\mathrm{n}}C_{\mathrm{ox}}\frac{W}{L}(V_{\mathrm{GS}} - V_{\mathrm{T}})^2$.

Summary: Small-Signal Parameters





- $I_{D(\text{sat})} = \frac{1}{2}\mu_{\text{n}} C_{\text{ox}} \frac{W}{L}(V_{\text{GS}} - V_{\text{T}})^2$
- $g_m = \mu_{\text{n}} C_{\text{ox}} \frac{W}{L}(V_{\text{GS}} - V_{\text{T}}) = \sqrt{2\mu_{\text{n}} C_{\text{ox}} \frac{W}{L} I_{D(\text{sat})}} = k\sqrt{I_{D(\text{sat})}}$
- $r_{\text{DS}} = \frac{1}{\lambda I_{D(\text{sat})}}$

Back to Biasing

- $I_{D(\text{sat})} = \frac{1}{2}\mu_{\text{n}} C_{\text{ox}} \frac{W}{L}(V_{\text{GS}} - V_{\text{T}})^2$
- $g_m = \mu_{\text{n}} C_{\text{ox}} \frac{W}{L}(V_{\text{GS}} - V_{\text{T}}) = \sqrt{2\mu_{\text{n}} C_{\text{ox}} \frac{W}{L} I_{D(\text{sat})}} = k\sqrt{I_{D(\text{sat})}}$
- $r_{\text{DS}} = \frac{1}{\lambda I_{D(\text{sat})}}$
- The reason all of this is interesting is the equation for gain:

$$\text{gain} = A = -g_m R_{\text{L}}$$

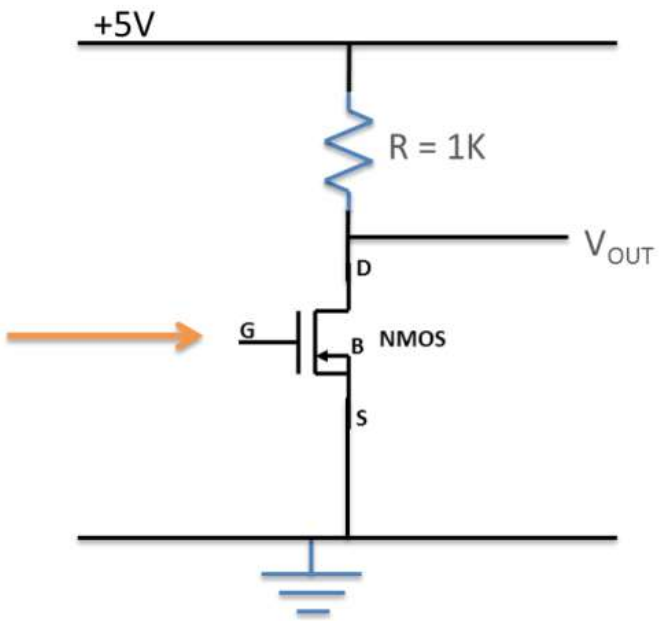
- When in saturation, everything is controlled by the $I_{\text{D(sat)}}$ current which is controlled only by the DC bias voltage on the gate V_{GS}. Changing that, we change the current.
- Triode region has different equations but we almost never use it for amplifiers.

The Effect of Varying V_{GS}

- $I_{D(\text{sat})} = \frac{1}{2}\mu_{\text{n}} C_{\text{ox}} \frac{W}{L}(V_{\text{GS}} - V_{\text{T}})^2$
- $g_m = \mu_{\text{n}} C_{\text{ox}} \frac{W}{L}(V_{\text{GS}} - V_{\text{T}}) = \sqrt{2\mu_{\text{n}} C_{\text{ox}} \frac{W}{L} I_{D(\text{sat})}} = k\sqrt{I_{D(\text{sat})}}$
- $r_{\text{DS}} = \frac{1}{\lambda I_{D(\text{sat})}}$

- Let's increase V_{GS}:
  - V_{GS} ↑
  - $I_{D(\text{sat})}$ ↑
  - $g_m = k\sqrt{I_{D(\text{sat})}}$ ↑
  - |gain|=$|A| = g_m R_{\text{L}}$ ↑

What Happens with V_{OUT}?



$$V_{\mathrm{OUT}} = V_{\mathrm{DD}} - R_{\mathrm{L}} I_{\mathrm{D(sat)}} = 5 - R_{\mathrm{L}} I_{\mathrm{D(sat)}}$$

- If we increase V_{GS}:
  - V_{GS} ↑
  - $I_{D(\mathrm{sat})}$ ↑
  - $R_{\mathrm{L}} I_{\mathrm{D(sat)}}$ ↑
  - V_{OUT} ↓

We have to be careful. Too much current will push V_{OUT} down unless we reduce R_L at the same time. But we need R_L for gain ($A = -g_m R_L$).

Increase Gain While Maintaining V_{OUT}?

- We are playing with the following two equations:

$$V_{\mathrm{OUT}} = 5 - R_{\mathrm{L}} I_{\mathrm{D(sat)}} \text{ and } |\text{gain}| = |A| = g_m R_{\mathrm{L}}$$

- If I_D increases, R_L must decrease, but if we decrease R_L the gain goes down. If we increase R_L, then I_D must go down. What do we do???
- The key is that the relation between g_m and $I_{D(sat)}$ is not linear!

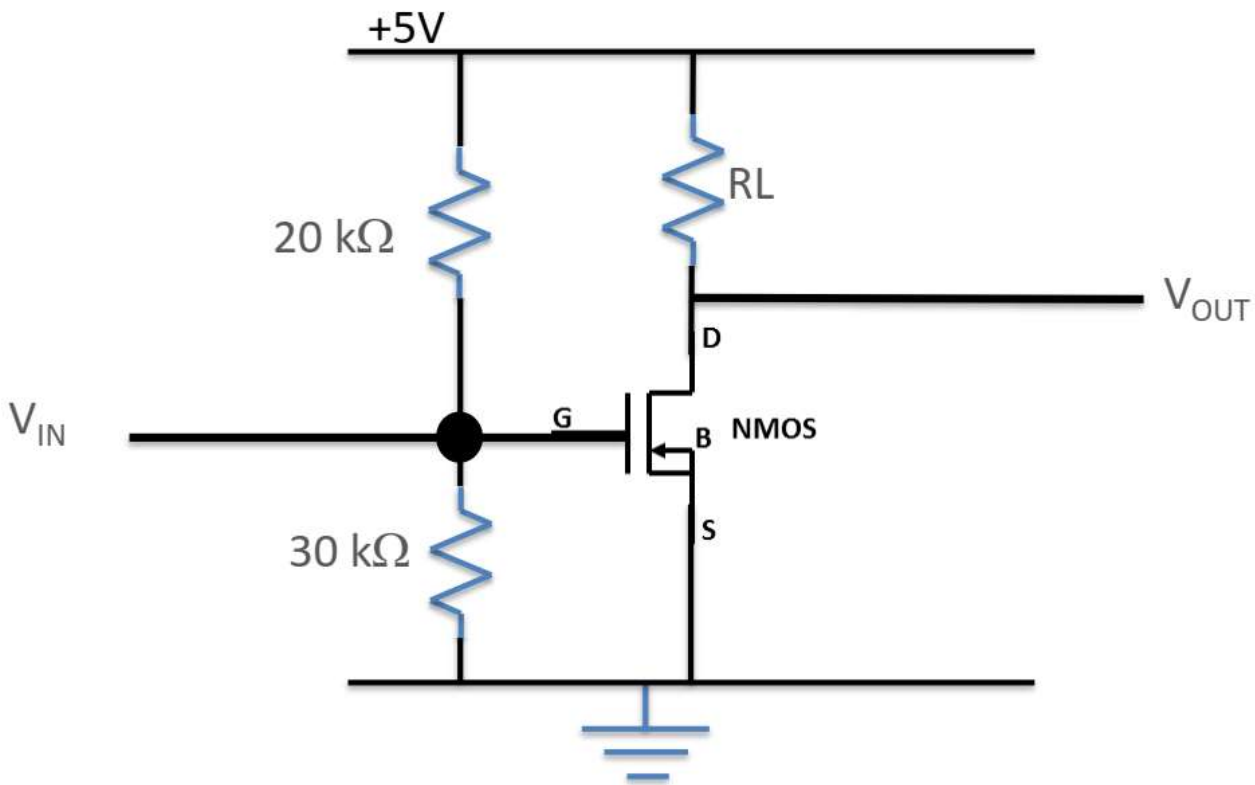
$$V_{\mathrm{OUT}} = 5 - R_{\mathrm{L}} I_{\mathrm{D(sat)}} \text{ and } |\text{gain}| = k\sqrt{I_{D(\mathrm{sat})}}(R_{\mathrm{L}})$$

- What we could do is to decrease I_D (for example by a factor of 2), increase R_L (by a factor of 2), and gain would then actually go up by a factor of $\sqrt{2}$ without affecting V_{OUT} (no change).

Is Decreasing $I_{D(sat)}$ Bad?

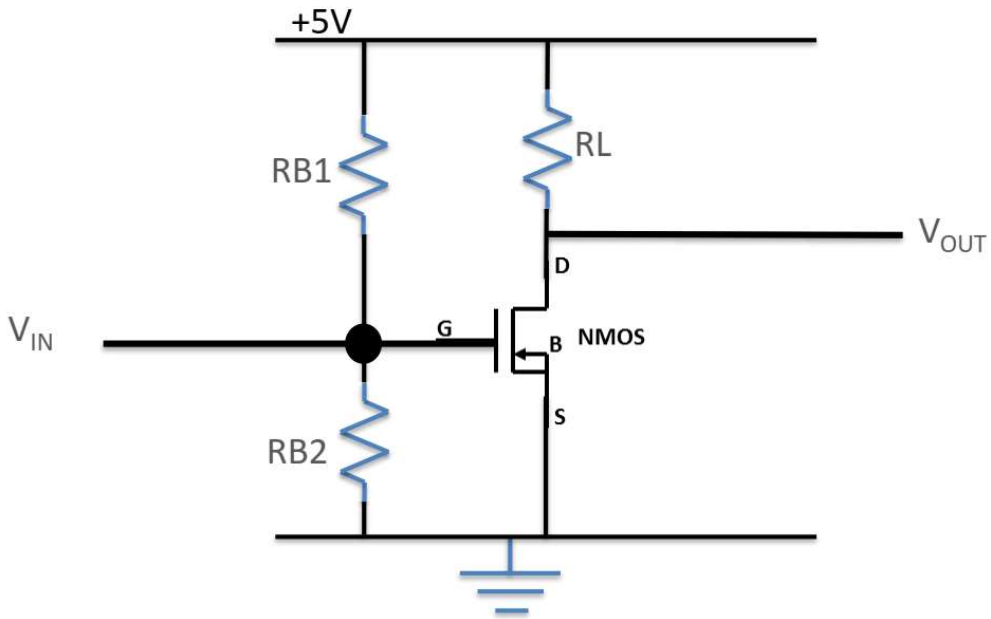
- IT CAN BE!!!!
- There is a limit to how far we can decrease I_D before we get to an unacceptable value of V_{GS} and start turning off the device. V_{GS} must be greater than V_T, we do not want to be close to that boundary as our small signal will be changing V_{GS}. So stay away!!!
- A very low I_D will make it difficult to run the transistor fast (handle high frequency signals) as we have no current to rapidly charge/discharge capacitors connected to the load.
- This is less of an issue for tiny signals but a serious issue for anything you'd want to do in a project.

How Do We Set a Bias Voltage?



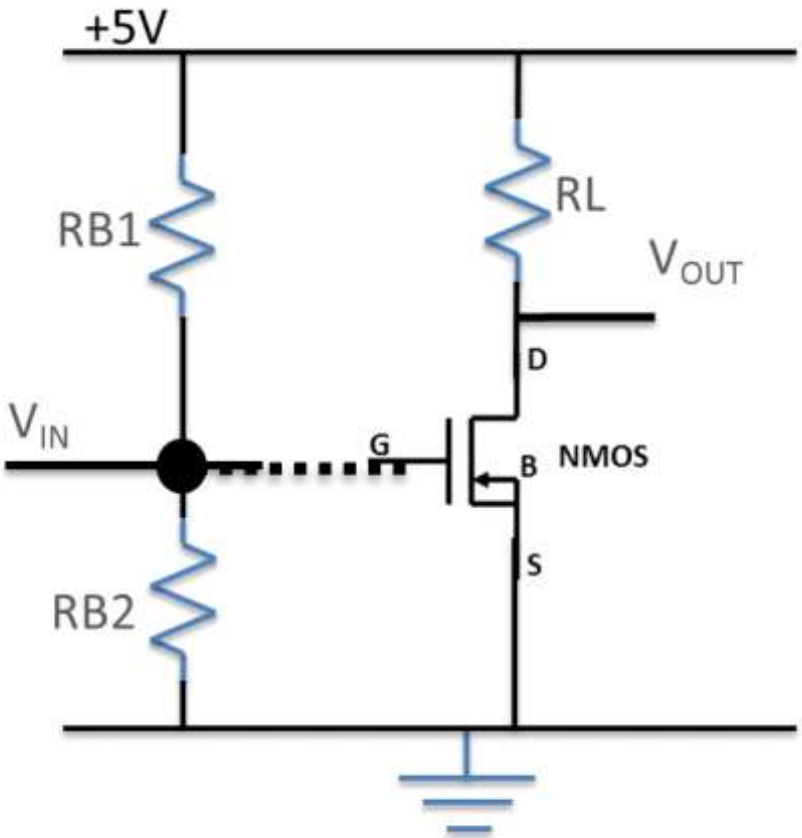
- Remember the voltage-divider biasing scheme? This is a very common biasing circuit for a CMOS device. What voltage is on the gate of the transistor?

Voltage-Divider Biasing



- What current flows into the gate of the CMOS transistor?
- What happens if one connects an input signal that has a
  - low resistance?
  - high resistance?
- What happens if one connects an output that has a
  - low resistance?
  - high resistance?

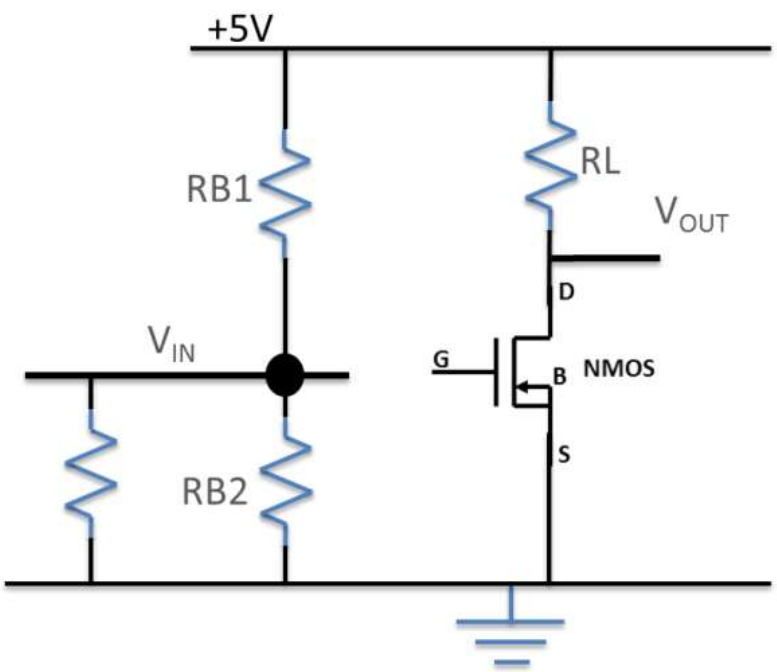
How is V_{GS} Set?

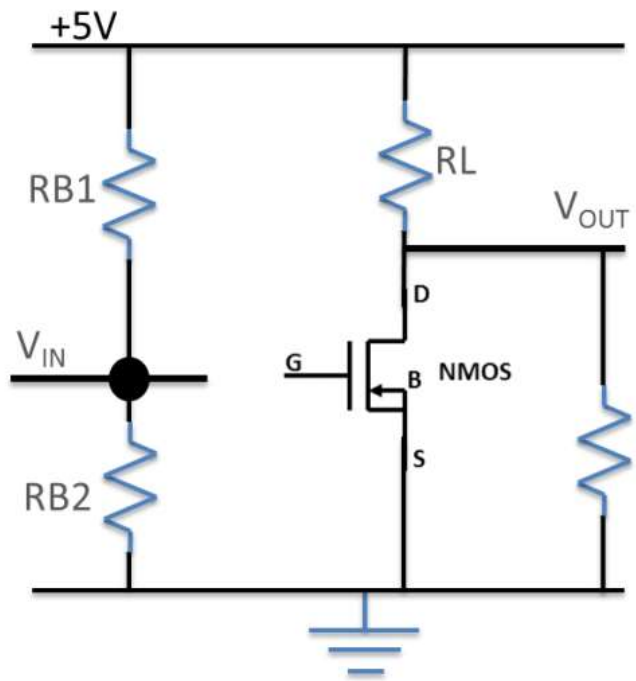


If no current flows, it's effectively a break. The voltage is set by the resistors, not by the transistor.

$$V_{\mathrm{GS}} = \frac{R_{\mathrm{B2}}}{R_{\mathrm{B1}} + R_{\mathrm{B2}}}(5\ \mathrm{V})$$

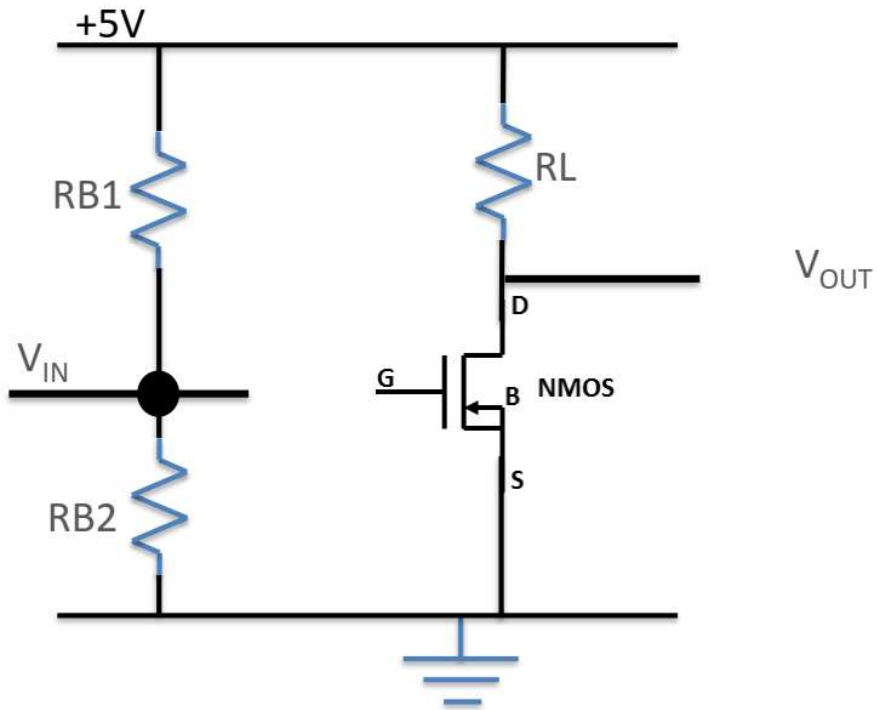
What Happens with Load Resistances?





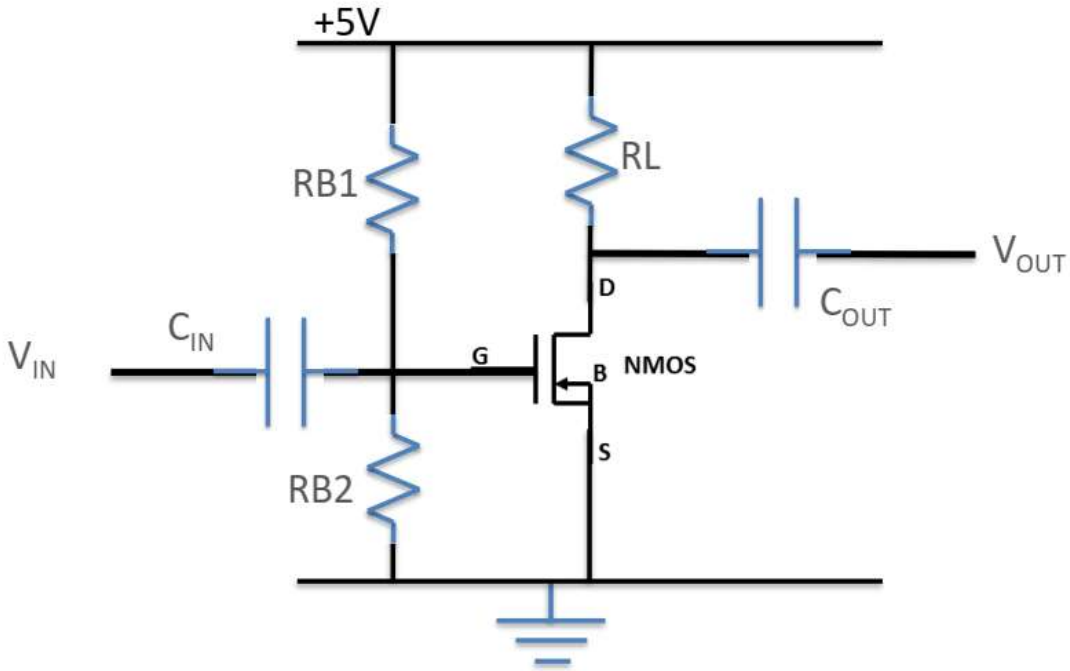
- Connecting resistors to the inputs and the outputs could change our values for V_{OUT} and V_{IN} badly, messing up our transistor. Not desirable.

More on Voltage-Divider Biasing



- Assuming that no current flows back into the V_{in}, how much current do you need to set the voltage on the gate?
- What value should the resistors be?
- Why? (think power efficiency)

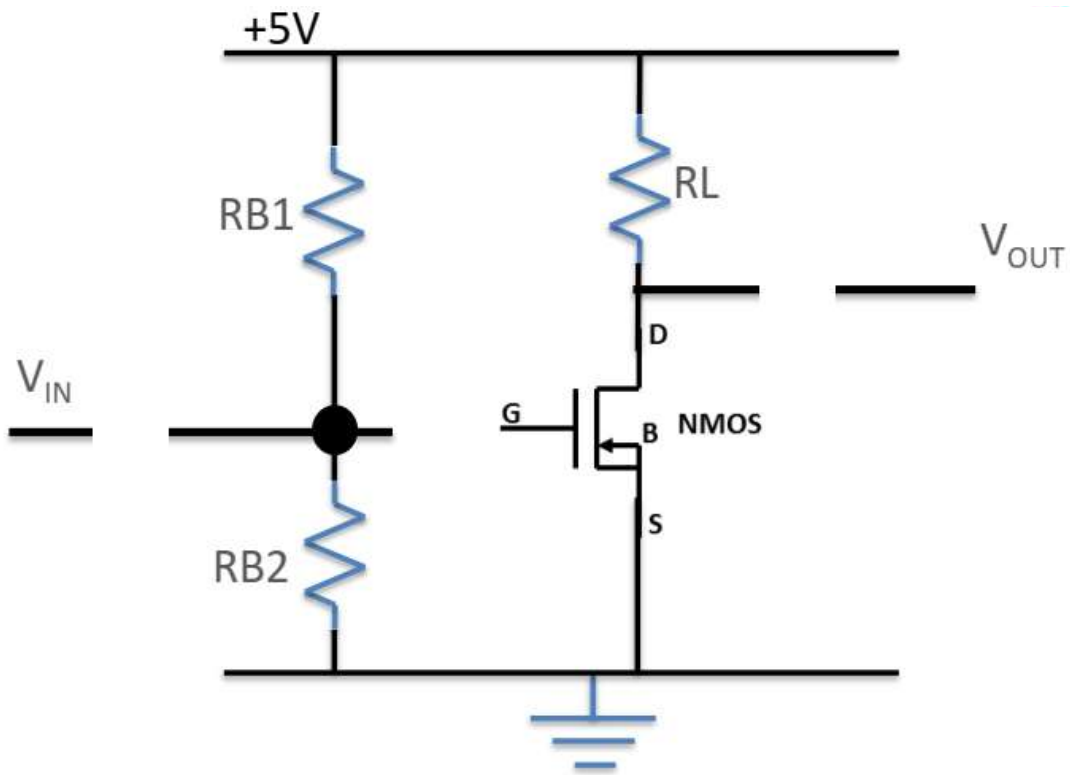
(De-)Coupling Capacitors



If in doubt, throw a big capacitor in, commonly 1 μF or 10 μF used.

- These capacitors can also be used at the input and output.
- Capacitors block DC, so basically you are splitting the DC circuit. Then the internal DC values don't make a difference to what happens before or after. Sometimes this is a good thing, sometimes not.
- How well do they block? It depends on your frequency of operation. Higher frequencies go with smaller capacitors. Impedance of a capacitor is $1/(j\omega C)$.

The DC Part of the Circuit



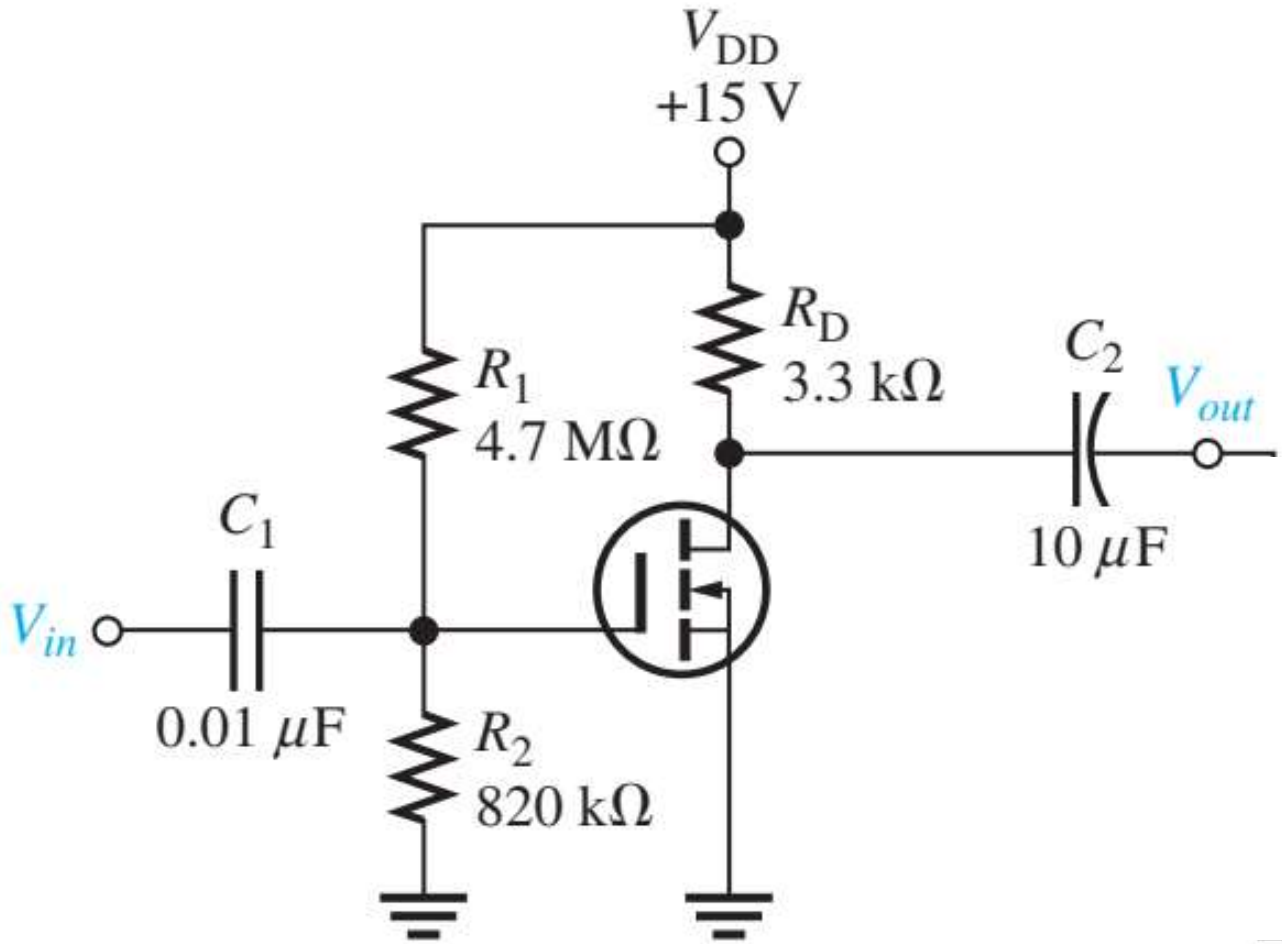
If in doubt, throw a big capacitor in, commonly 1μF or 10 μF used.

- Now it really doesn't matter what you connect to the output or the input. The biasing is isolated and protected!!!

$$V_{\mathrm{GS}} = \frac{R_{\mathrm{B2}}}{R_{\mathrm{B1}} + R_{\mathrm{B2}}} V_{\mathrm{DD}}$$

Recall This Circuit?

- Capacitors can be used to de-couple the DC bias.

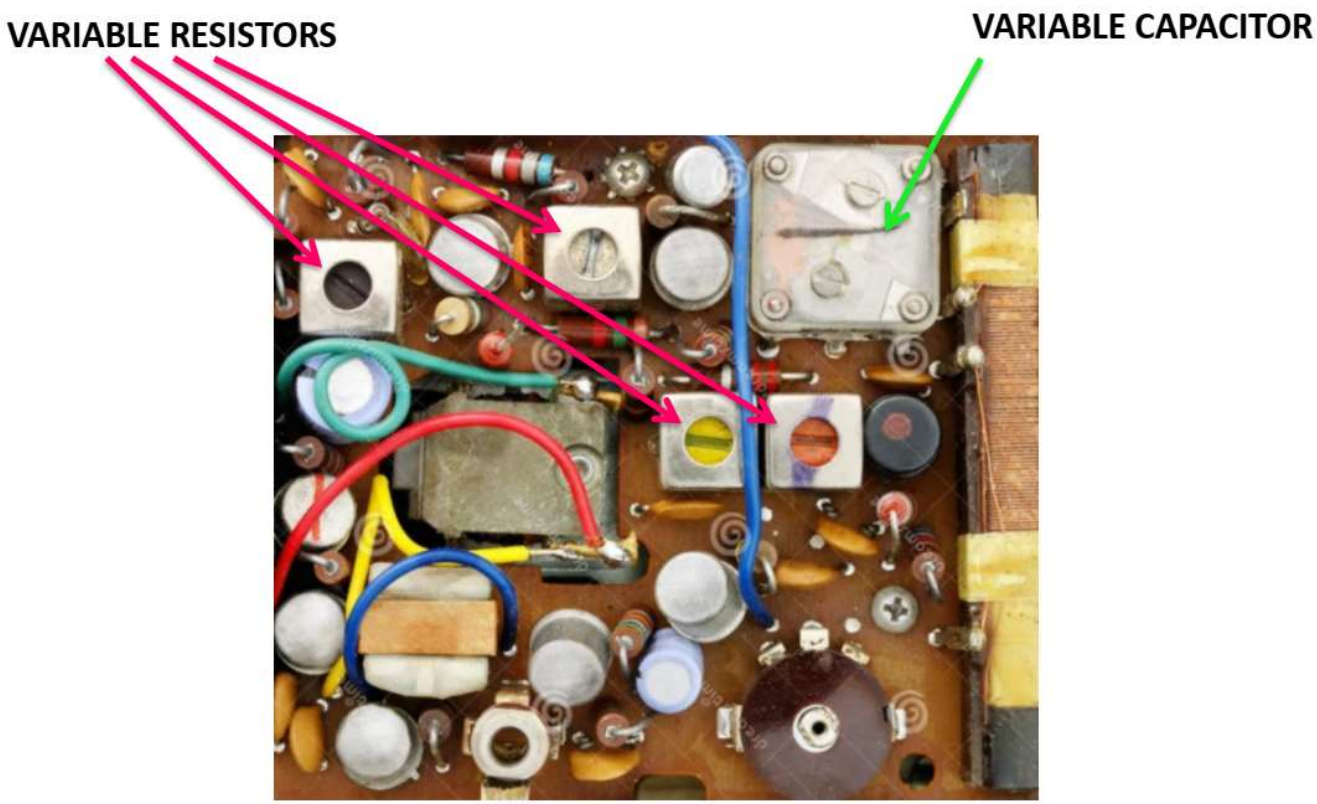


How to Choose a Bias Voltage?

$$V_{\mathrm{OUT}} = 5 - R_{\mathrm{L}} I_{\mathrm{D(sat)}} \text{ and } |\text{gain}| = |A| = g_m R_{\mathrm{L}}$$

- There are two settings you need to adjust, and they are interlinked. There is no perfect answer.
- A common process goes like this:
  - Start with a reasonable value of I_D for your transistor;
  - What value of V_{GS} gives me that current;
  - What value of R_L gives me V_{OUT};
  - What is my value of GAIN;
  - If the GAIN is too low, adjust V_{GS} and R_L to try to get a good compromise.
- Software design tools can assist you in getting the precise answer but it is always a compromise and that is why it takes an engineer to pick the BEST COMPROMISE. There is no single correct answer.
- That's why it's called design.

Use of Trimmers



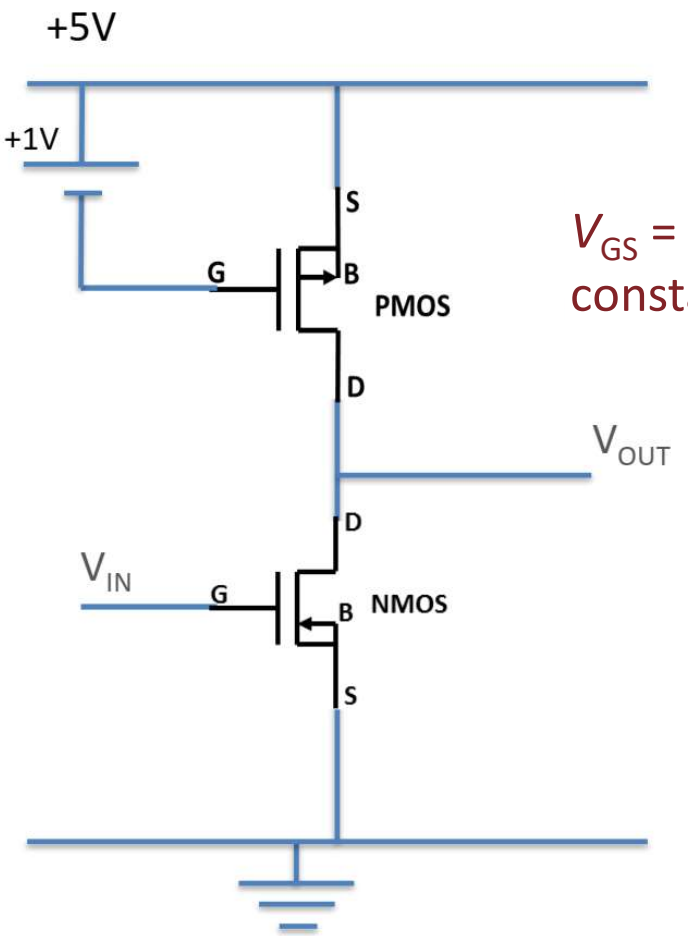
- An old radio circuit board. With discrete components one doesn't always get things right the first time, so tweaking used to be an important part of the manufacturing process.

Some Terminology

- You'll see this mentioned in different places, just so you know…
  - GND Ground;
  - V_{CC} positive voltage for BJT circuits (or older circuits), originally means voltage common collector;
  - V_{DD} positive voltage for CMOS circuits;
  - V_{EE} negative voltage (normally for BJT or op-amp circuits);
  - V_{SS} negative voltage for CMOS circuits (not seen that often). It is usually just the ground in a single-supply circuit.

Trick 1: Constant Gate Voltage

- What is the small-signal equivalent circuit for the top transistor?

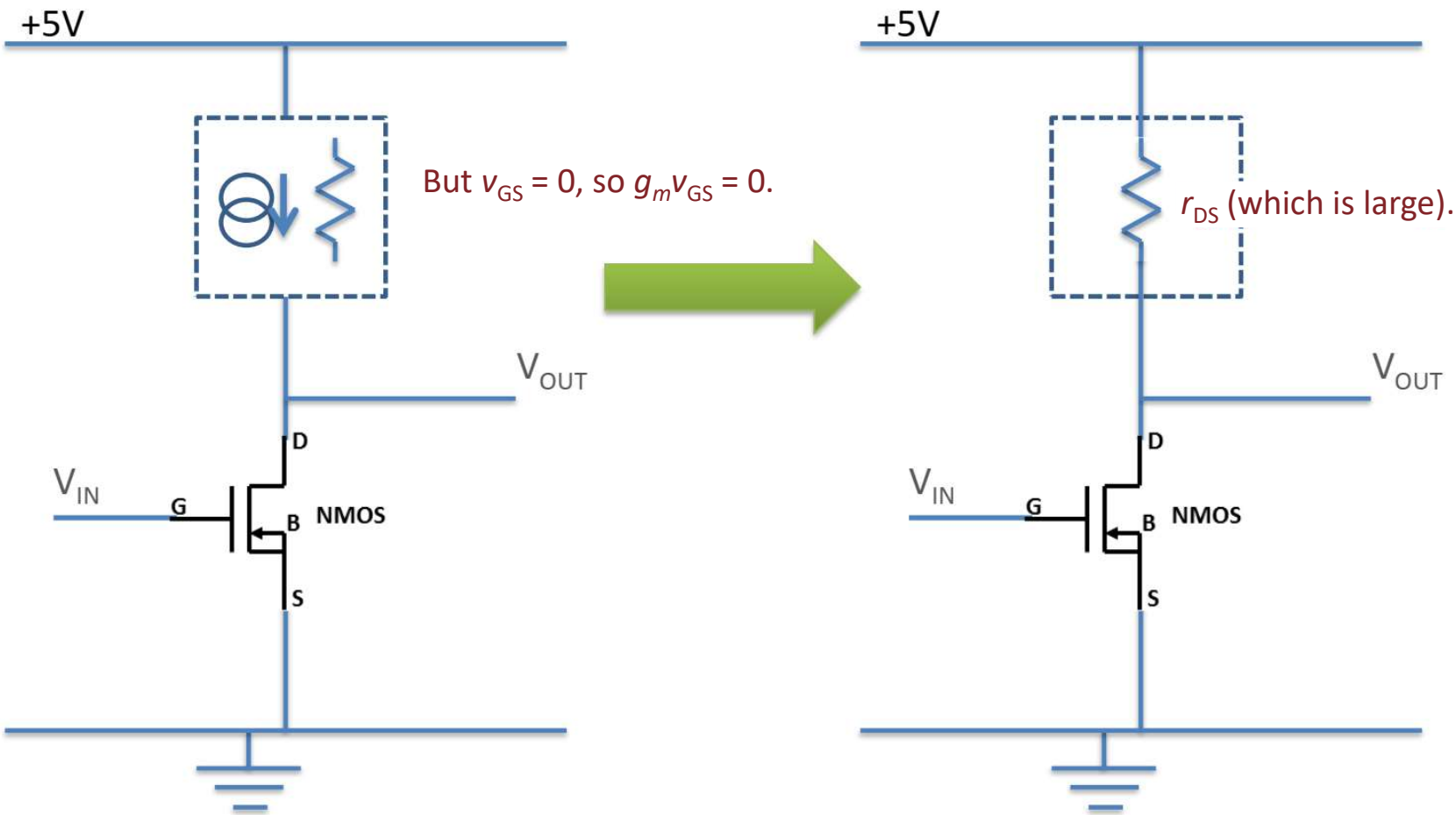


V_{GS} = -1 V (PMOS). But it is constant, so v_{GS} = 0.

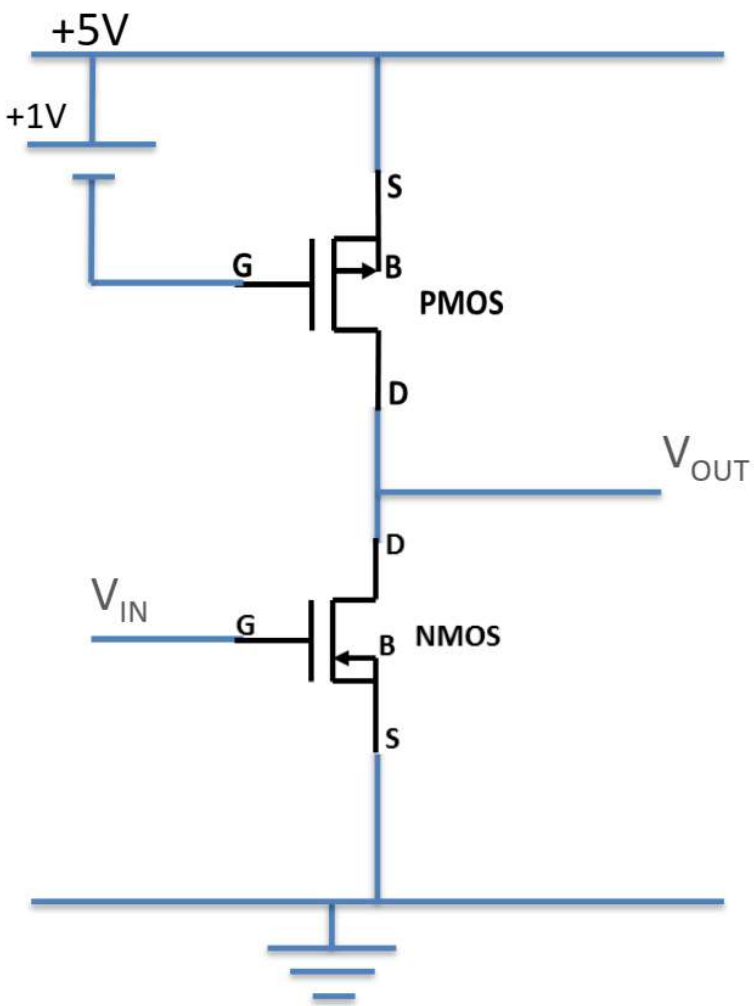
This is used very widely in real circuits, but is not a core part of this module. Just try to follow the concepts and be able to talk at a hand-wavy level.

Trick 1: Constant Gate Voltage

- What is the small-signal equivalent circuit for the top transistor?



Trick 1: DC Resistance

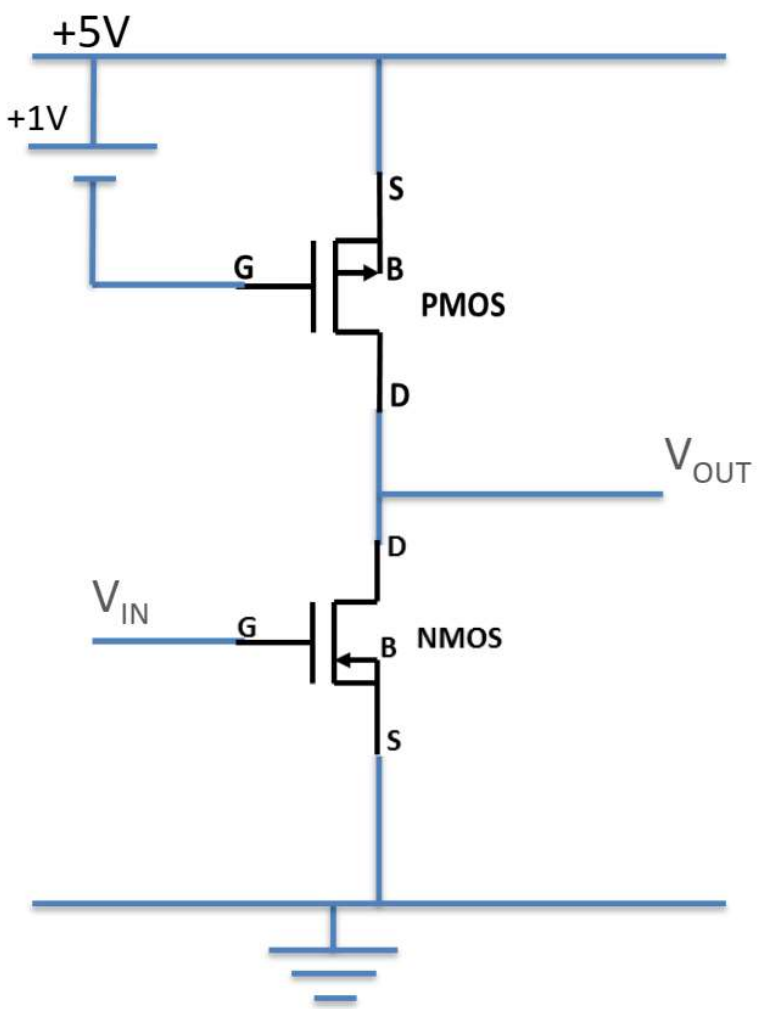


- Let's assume that we can put the PMOS in the saturation region (this can be done by adjusting, e.g., V_{IN} to the NMOS and V_{GS} for the PMOS):

$$R_{DS} = \frac{|V_{DS}|}{\frac{1}{2}\mu_p C_{ox}\frac{W}{L}(|V_{GS}| - |V_T|)^2}$$

- As we saw earlier, R_{DS} tends to be smaller, in the order of 100 - 10000 ohms and can be controlled;
- The voltage V_{GS} we pick sets the value of R_{DS}.

Trick 1: DC Resistance *vs*. AC Resistance

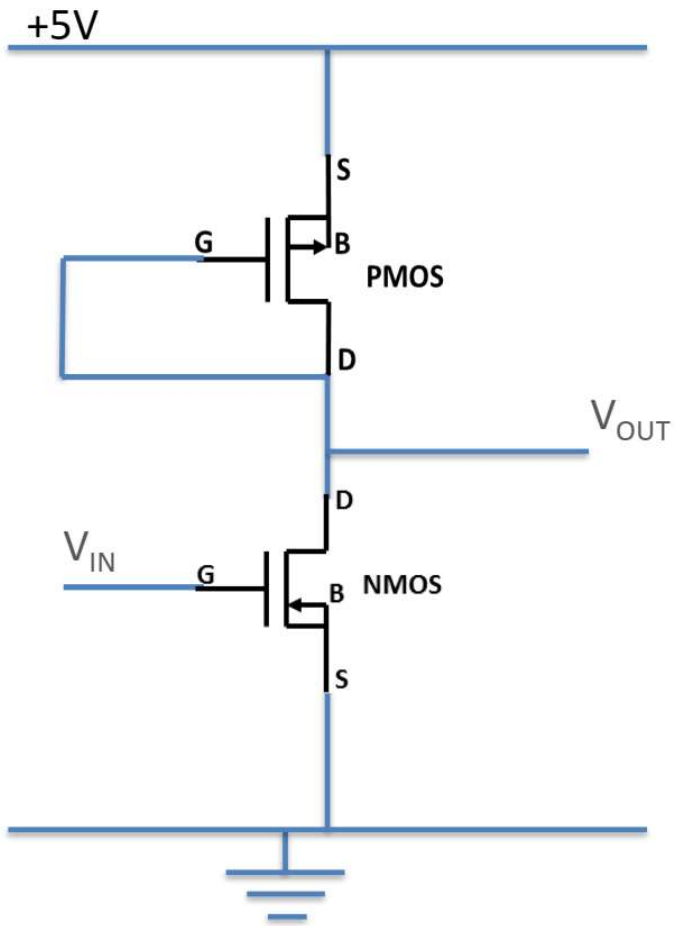


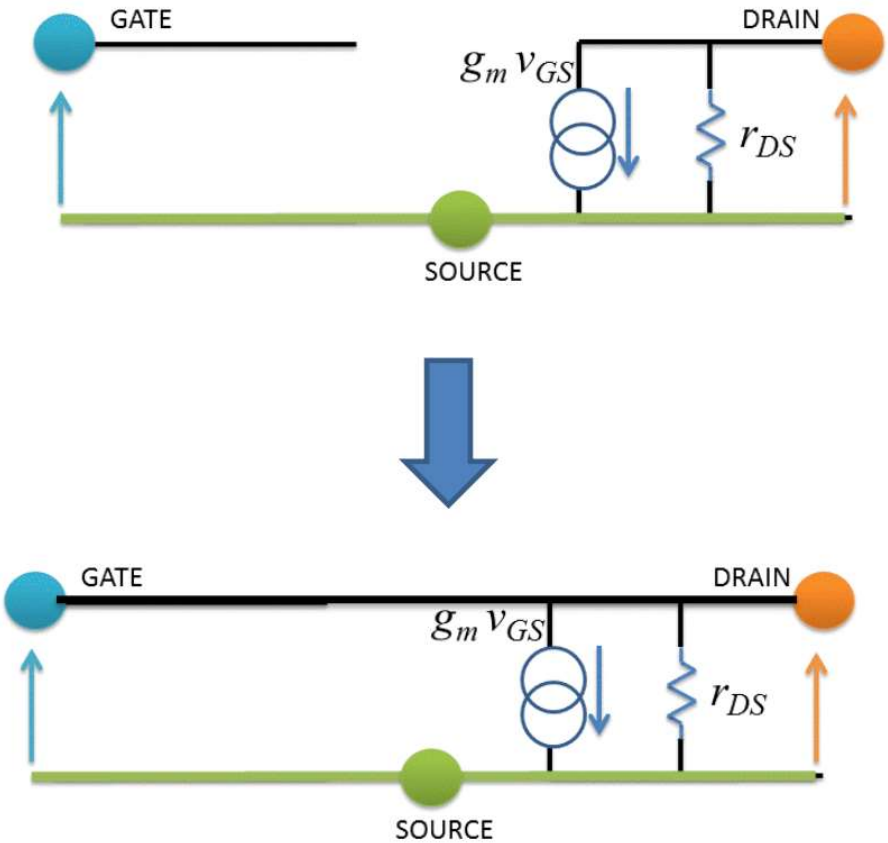
$$R_{\mathrm{DS}} = \frac{|V_{\mathrm{DS}}|}{\frac{1}{2}\mu_{\mathrm{p}} C_{\mathrm{ox}} \frac{W}{L} (|V_{\mathrm{GS}}| - |V_{\mathrm{T}}|)^2}$$

$$r_{\mathrm{DS}} = \frac{1}{\lambda I_{D(\mathrm{sat})}} = \frac{1}{\frac{\lambda}{2}\mu_{\mathrm{p}} C_{\mathrm{ox}} \frac{W}{L} (|V_{\mathrm{GS}}| - |V_{\mathrm{T}}|)^2}$$

- Basically, we can get a device which has a different LARGE VALUE/DC resistance from it's small-signal resistance;
- We can tune this value by the V_{GS} of this transistor;
- This is used a lot in microchip design as transistors are smaller than large resistors.

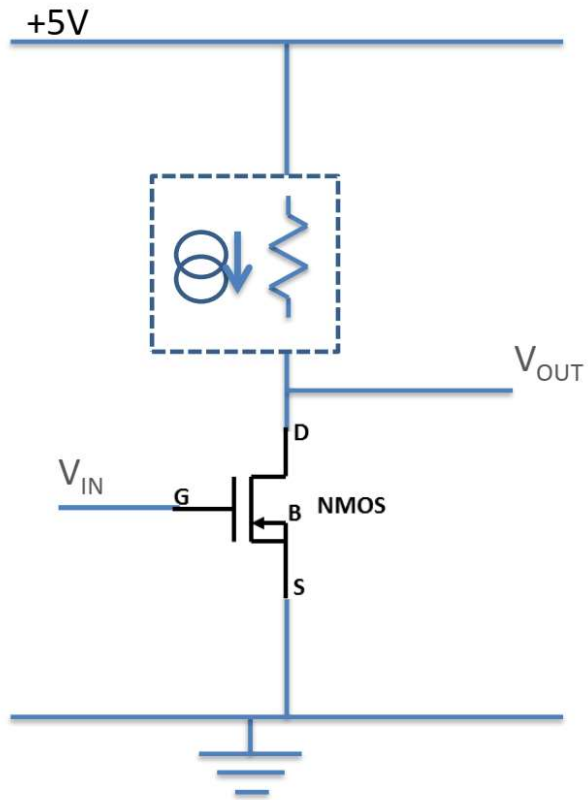
Trick 2: Gate -> Drain

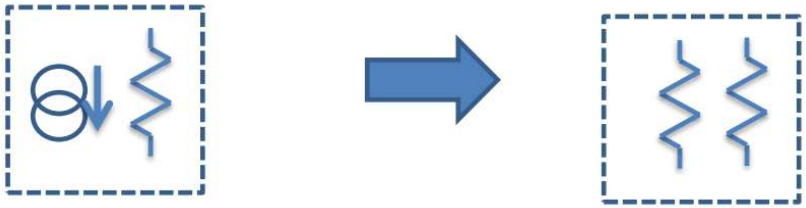




- In this configuration, saturation is guaranteed as $|V_{DS}| = |V_{GS}| > (|V_{GS}| - |V_T|)$.

Trick 2: Gate -> Drain





- But in this case,

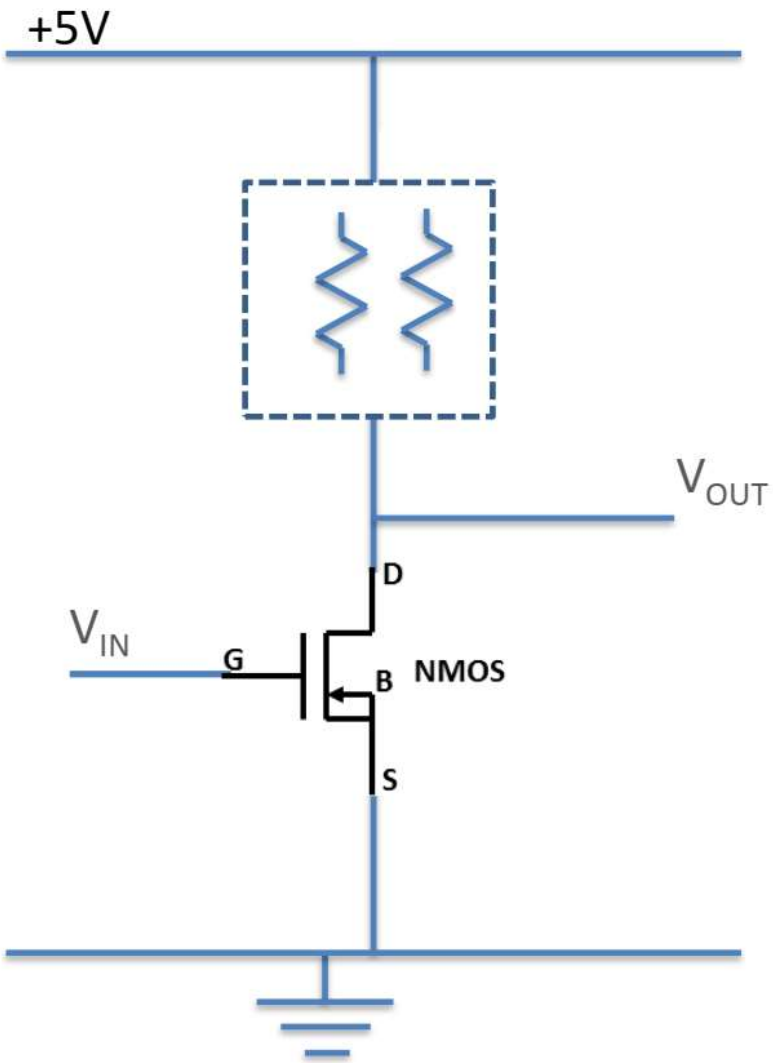
$$i_{\mathrm{D}} = g_m v_{\mathrm{GS}} = g_m v_{\mathrm{DS}}$$

And Ohm's law says:

$$i_{\mathrm{D}} = g_m v_{\mathrm{GS}} = \frac{1}{r_{\mathrm{gm}}} v_{\mathrm{DS}}$$

- Actually, what we have here is that instead of a current source, we have a resistance of $1/g_{\mathrm{m}}$.
- So the transistor, for small signals, looks like two parallel resistors.
- $1/g_{\mathrm{m}}$ will be the smaller one so it will dominate.

Trick 2: Gate -> Drain



$$r_{\mathrm{DS}} = \frac{1}{\lambda I_{\mathrm{D(sat)}}}$$

$$I_{\mathrm{D(sat)}} = \frac{1}{2}\mu_{\mathrm{p}} C_{\mathrm{ox}} \frac{W}{L}(|V_{\mathrm{GS}}| - |V_{\mathrm{T}}|)^2$$

and

$$r_{\mathrm{gm}} = \frac{1}{g_m}$$

So the combined resistance is:

$$\frac{1}{r_{\mathrm{T}}} = \frac{1}{r_{\mathrm{DS}}} + \frac{1}{r_{\mathrm{gm}}} = \frac{1}{r_{\mathrm{DS}}} + g_m$$

$$r_{\mathrm{T}} \approx r_{\mathrm{gm}} = 1/g_m$$

Trick 2: DC Resistance

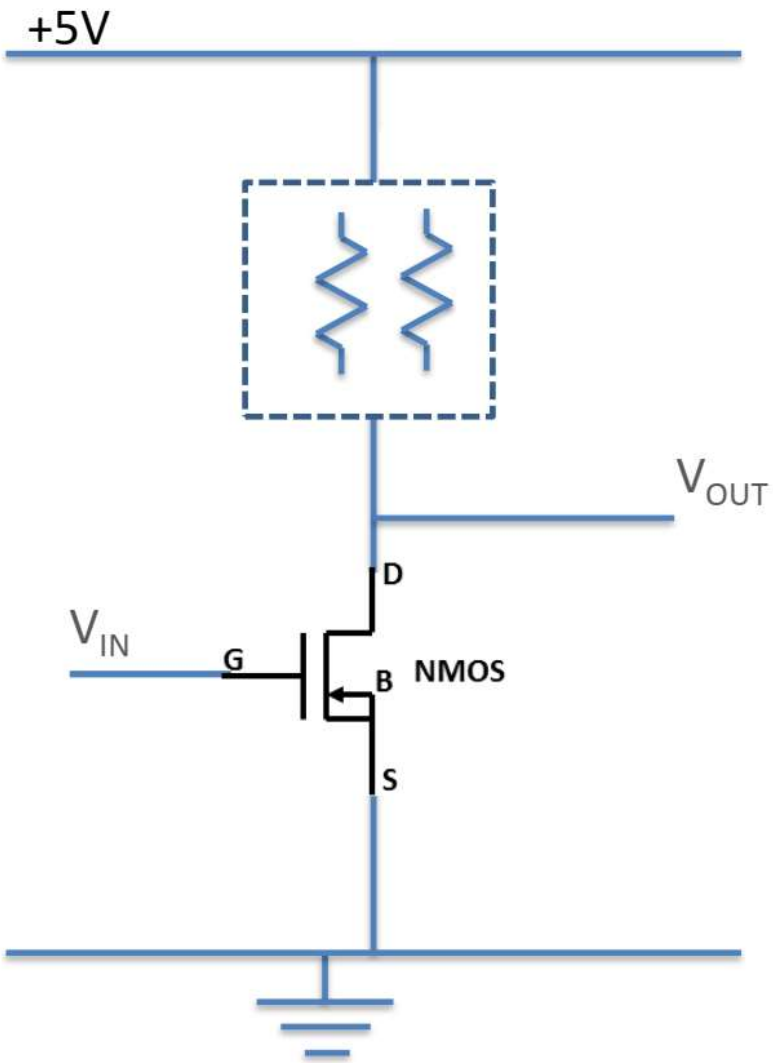
- We have the small signal resistance, but what is the DC resistance for biasing purposes? Keep in mind that $V_{GS} = V_{DS}$.

$$r_{\mathrm{T}} \approx \frac{1}{g_m} = \frac{1}{\mu_{\mathrm{p}} C_{\mathrm{ox}} \frac{W}{L} (|V_{\mathrm{DS}}| - |V_{\mathrm{T}}|)}$$

$$R_{\mathrm{T}} = \frac{V_{\mathrm{DS}}}{I_{\mathrm{DS}}} = \frac{|V_{DS}|}{\frac{1}{2} \mu_{\mathrm{p}} C_{\mathrm{ox}} \frac{W}{L} (|V_{\mathrm{GS}}| - |V_{\mathrm{T}}|)^2} = \frac{2L}{\mu_{\mathrm{p}} C_{\mathrm{ox}} W} \frac{|V_{\mathrm{DS}}|}{(|V_{\mathrm{DS}}| - |V_{\mathrm{T}}|)^2}$$

- Again we have different values for *R* and *r*, but both are controlled by V_{DS}. V_{DS} can change depending on our value of drain voltage. For small signals it's good (V_{DS} may be viewed as a constant). For large signals it is not so good (the change in V_{DS} has to be taken into account). However, this circuit is very easy to build (just use a wire).

Trick 2: Gate -> Drain



In summary,

$$r_{\mathrm{T}} \approx \frac{1}{g_m} = \frac{1}{\mu_{\mathrm{p}} C_{\mathrm{ox}} \frac{W}{L} (|V_{\mathrm{DS}}| - |V_{\mathrm{T}}|)}$$

$$R_{\mathrm{T}} = \frac{V_{\mathrm{DS}}}{I_{\mathrm{DS}}} = \frac{2L}{\mu_{\mathrm{p}} C_{\mathrm{ox}} W} \frac{|V_{\mathrm{DS}}|}{(|V_{\mathrm{DS}}| - |V_{\mathrm{T}}|)^2}$$

The combined small-signal resistance is:

$$\frac{1}{r_{\mathrm{T}}} = \frac{1}{r_{\mathrm{DS}}} + \frac{1}{r_{\mathrm{gm}}} = \frac{1}{r_{\mathrm{DS}}} + g_m$$

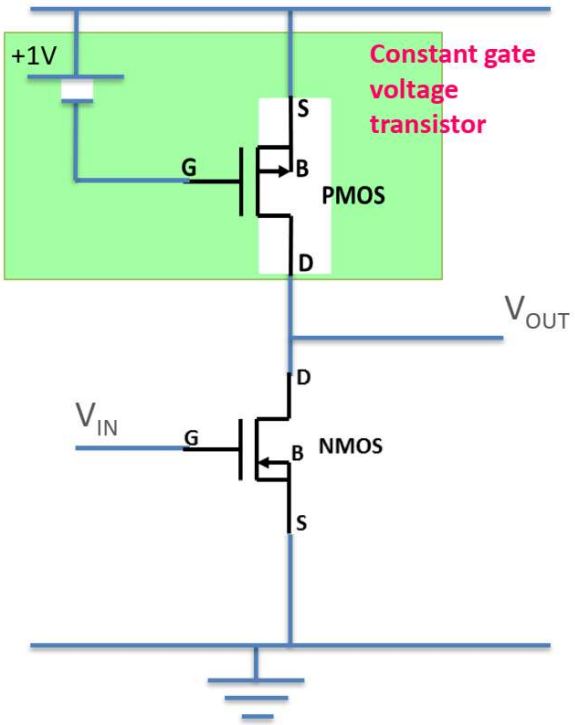
$$r_{\mathrm{T}} \approx r_{\mathrm{gm}} = 1/g_m$$

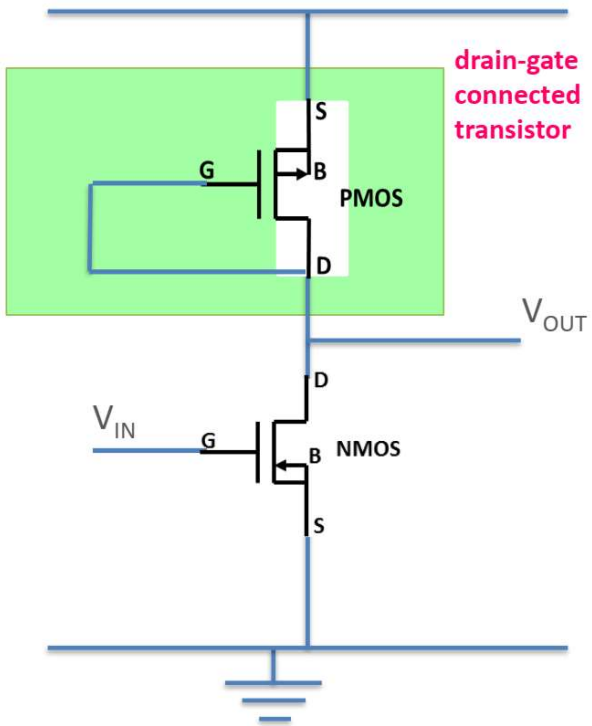
Trick 2: Gate -> Drain

- Let's plug in some numbers for a drain-connected device:
  - $|V_{DS}|$ = 3 V;
  - $|V_T|$ = 0.5 V;
  - W/L = 20;
  - $\mu_p C_{ox}$ = 30 μA/V^2.
- $r_T \approx \frac{1}{g_m} = \frac{1}{\mu_p C_{ox} \frac{W}{L}(|V_{DS}|-|V_T|)} = \frac{1}{(30\times10^{-6})(20)(3-0.5)} \approx 667\ \Omega.$
- $R_T = \frac{2L}{\mu_p C_{ox} W} \frac{|V_{DS}|}{(|V_{DS}|-|V_T|)^2} = \frac{2}{(30\times10^{-6})(20)} \frac{3}{(3-0.5)^2} \approx 1600\ \Omega.$
- We can make the DC and AC resistances different, depending on how we want to manipulate gain versus DC bias conditions.

Comparison: Trick 1 *vs*. Trick 2

- Both will provide us with different *R* and *r*. The **constant gate voltage** solution requires a biasing circuit or a battery. The **drain-gate connected transistor** requires only a wire. It is guaranteed to be in saturation but as the drain voltage changes, V_{GS} changes so you get a small wobble on *r*. Probably not enough to worry about in most cases. This version is used more in microchips as it is smaller.

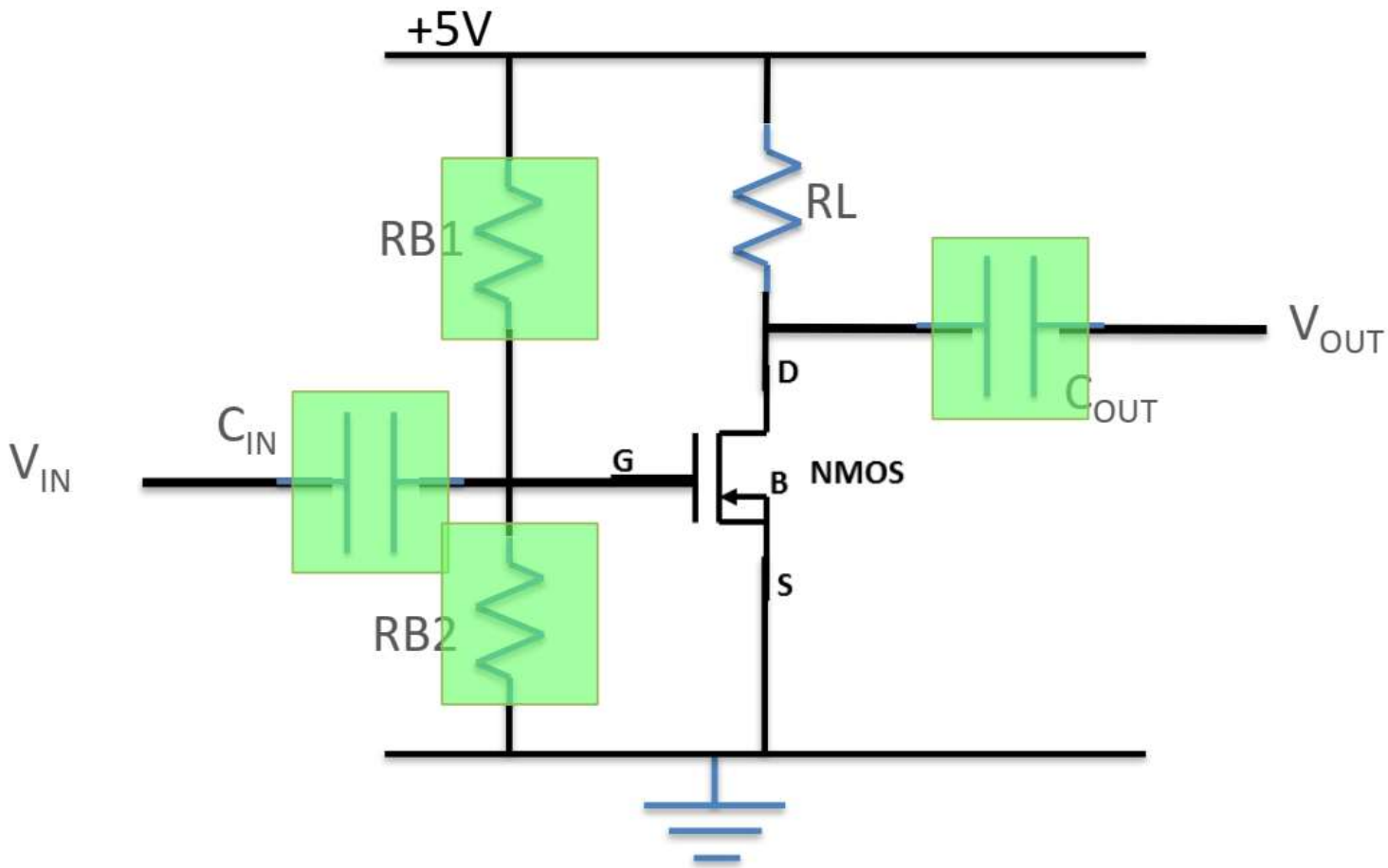




Back to Calculation Again!

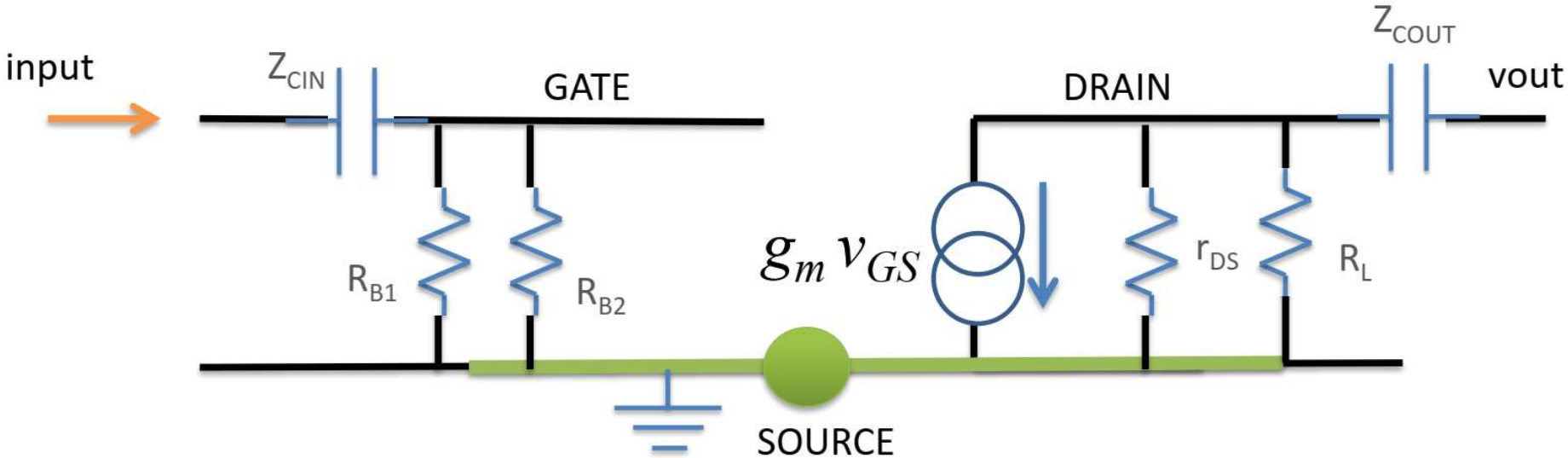
- What do those resistors and capacitors do in an FET amplifier?
- This is an important bit of our course module.

(De)-Coupling Capacitors & Bias Resistors



- Let's draw the small-signal equivalent circuit.

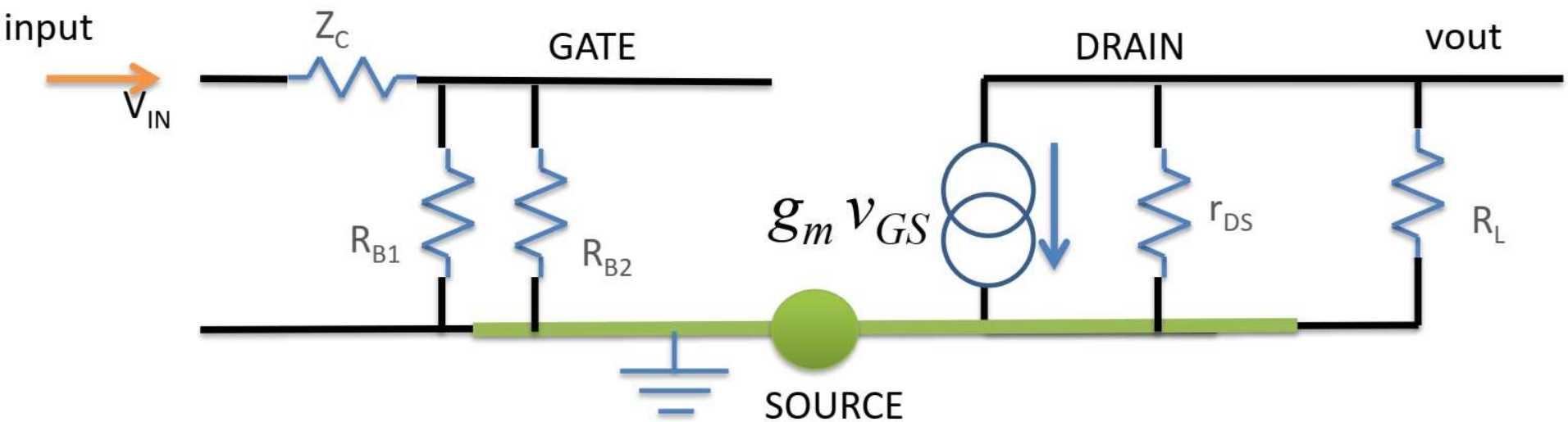
Small-Signal Analysis



- The two bias resistors go from GATE to V_{DD} and EARTH which are both small-signal earths.
- The capacitors present impedances to the AC signal, so let's add them in, too.

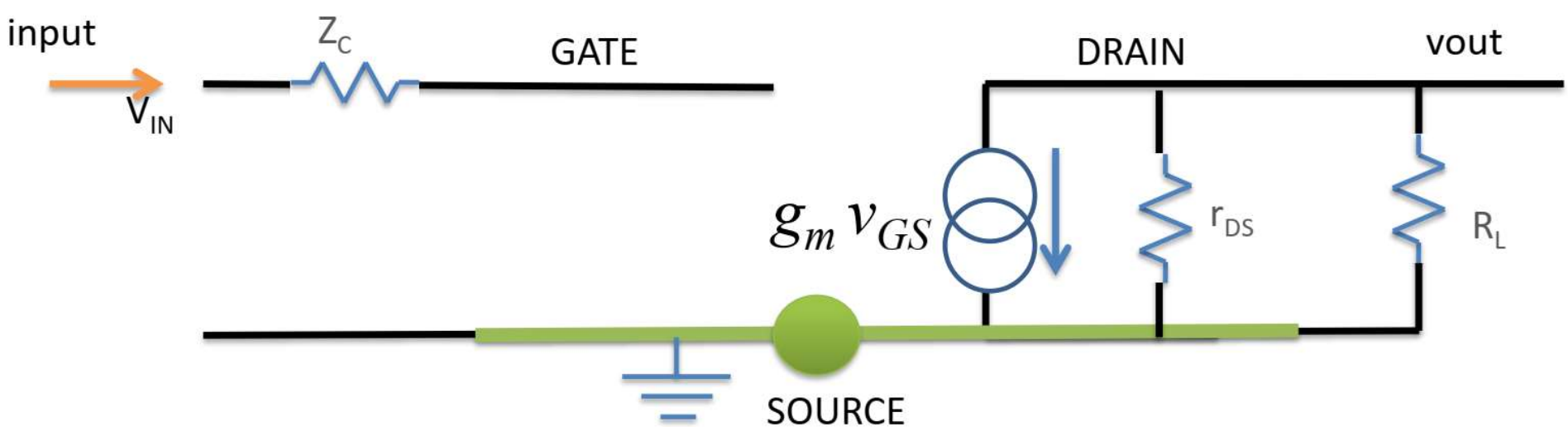
Small-Signal Analysis

- We'll be ignoring the output capacitance for the moment.
- Simplifying the input capacitor down to its impedance, $Z_C = 1/(2\pi f C)$, we get:



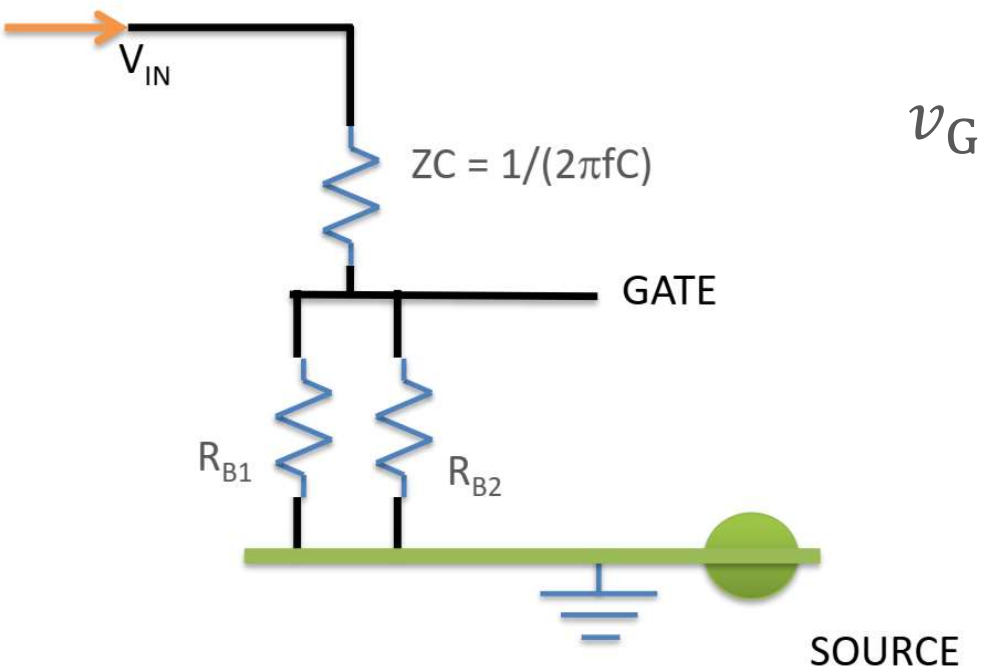
Small-Signal Analysis

- If we pick R_{B1} and R_{B2} to be very large, then we could ignore them. In this case, it wouldn't matter what the capacitor value was, no current flows.
- So we go back to the original circuit:



Small-Signal Analysis

- Let's consider that the impedance of a capacitor can be very large for very small frequencies, so that assumption is not always true. Let's explore this situation.
- Redraw the circuit a little to emphasise the voltage divider. What we really want is the gate voltage.



$$v_{\mathrm{GS}} = \frac{R_{\mathrm{B1}}||R_{\mathrm{B2}}}{\frac{1}{j\omega C} + R_{\mathrm{B1}}||R_{\mathrm{B2}}} v_{\mathrm{IN}}$$

$$= \frac{1}{\frac{1}{j\omega C(R_{\mathrm{B1}}||R_{\mathrm{B2}})} + 1} v_{\mathrm{IN}}$$

Small-Signal Analysis

- We also know that without the capacitor, we have

$$A = \frac{v_{\mathrm{OUT}}}{v_{\mathrm{GS}}} = \frac{-R_{\mathrm{L}} g_{\mathrm{m}} v_{\mathrm{GS}}}{v_{\mathrm{GS}}} = -g_{\mathrm{m}} R_{\mathrm{L}}$$

- But we want the "real gain" from v_{IN} to v_{OUT}, so let's swap this into our equation:

$$v_{\mathrm{GS}} = \frac{1}{\frac{1}{j\omega C(R_{\mathrm{B1}}||R_{\mathrm{B2}})} + 1} v_{\mathrm{IN}}$$

$$v_{\mathrm{IN}} = [\frac{1}{j\omega C(R_{\mathrm{B1}}||R_{\mathrm{B2}})} + 1] v_{\mathrm{GS}}$$

Small-Signal Analysis

- Back to the gain equation:

$$A_{\text{real}} = \frac{v_{\text{OUT}}}{v_{\text{IN}}} = \frac{v_{\text{OUT}}}{[\frac{1}{j\omega C(R_{\text{B1}}||R_{\text{B2}})} + 1]v_{\text{GS}}} = \frac{1}{\frac{1}{j\omega C(R_{\text{B1}}||R_{\text{B2}})} + 1}\frac{v_{\text{OUT}}}{v_{\text{GS}}}$$

- But $A = v_{\text{OUT}}/v_{\text{IN}}$,

$$A_{\text{real}} = (A)\frac{1}{\frac{1}{j\omega C(R_{\text{B1}}||R_{\text{B2}})} + 1}$$

- The original GAIN (A) is being scaled by the reduction being produced by the voltage divider, nothing more complicated than that.

Small-Signal Analysis

- Let's look at some numbers to see what this means. Let R_{B1} and R_{B2} = 2 MΩ, and let C = 1 μF, then

$$A_{\text{real}} = (A)\frac{1}{\frac{1}{j\omega C(R_{\text{B1}}||R_{\text{B2}})}+1} = (A)\frac{1}{\frac{1}{j\omega(1\mu)(1\text{M})}+1}$$

$$= (A)\frac{1}{\frac{1}{j\omega}+1}$$

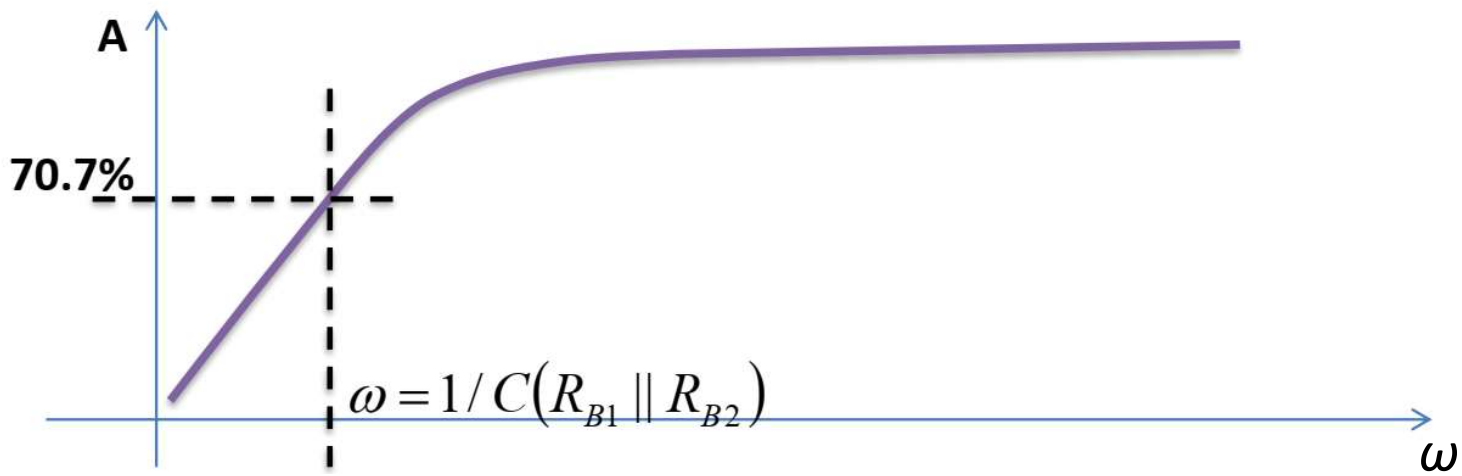
- We can conclude that

$$A_{\text{real}} = (A)\frac{1}{1-j\frac{1}{\omega}} \text{ and } |A_{\text{real}}| = |A|\frac{1}{\sqrt{1+(\frac{1}{\omega})^2}}$$

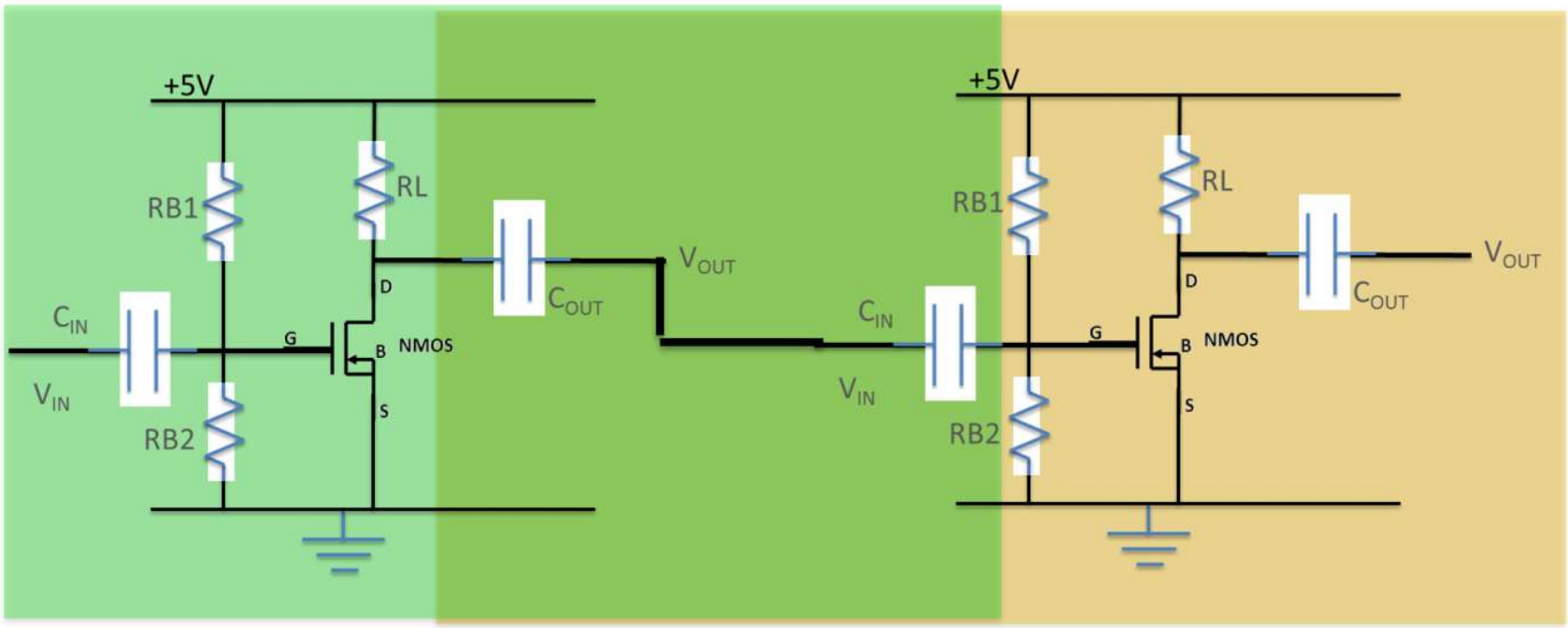
Small-Signal Analysis

$$|A_{\text{real}}| = |A| \frac{1}{\sqrt{1 + (\frac{1}{\omega})^2}}$$

- If the frequency is equal to zero, then $1/\omega$ is infinite, the gain is being divided by $1/\omega$ which will give zero;
- As the frequency becomes > 1 Hz, then the effect becomes very small very quickly. At 1 kHz, $1/[(2\pi)1000] = 1.6 \times 10^{-4}$, $[1/(1.6 \times 10^{-4})]^2 = 2.5 \times 10^{-8}$;
- So we get a curve like this.



The Output Capacitance



- If you are connecting two transistors together, then you don't need two coupling capacitors for the output and the input, you can just use one;
- In that case, imagine that there is no output capacitor and it is just an input one, the effect is similar, low frequency signals cannot get through, the familiar high-pass transfer function;
- This double hurts our low frequency signals, which may be a problem or a benefit.

(De)-Coupling Capacitors: How to Pick?

$$\omega = 1/[C(R_{\mathrm{B1}}||R_{\mathrm{B2}})]$$

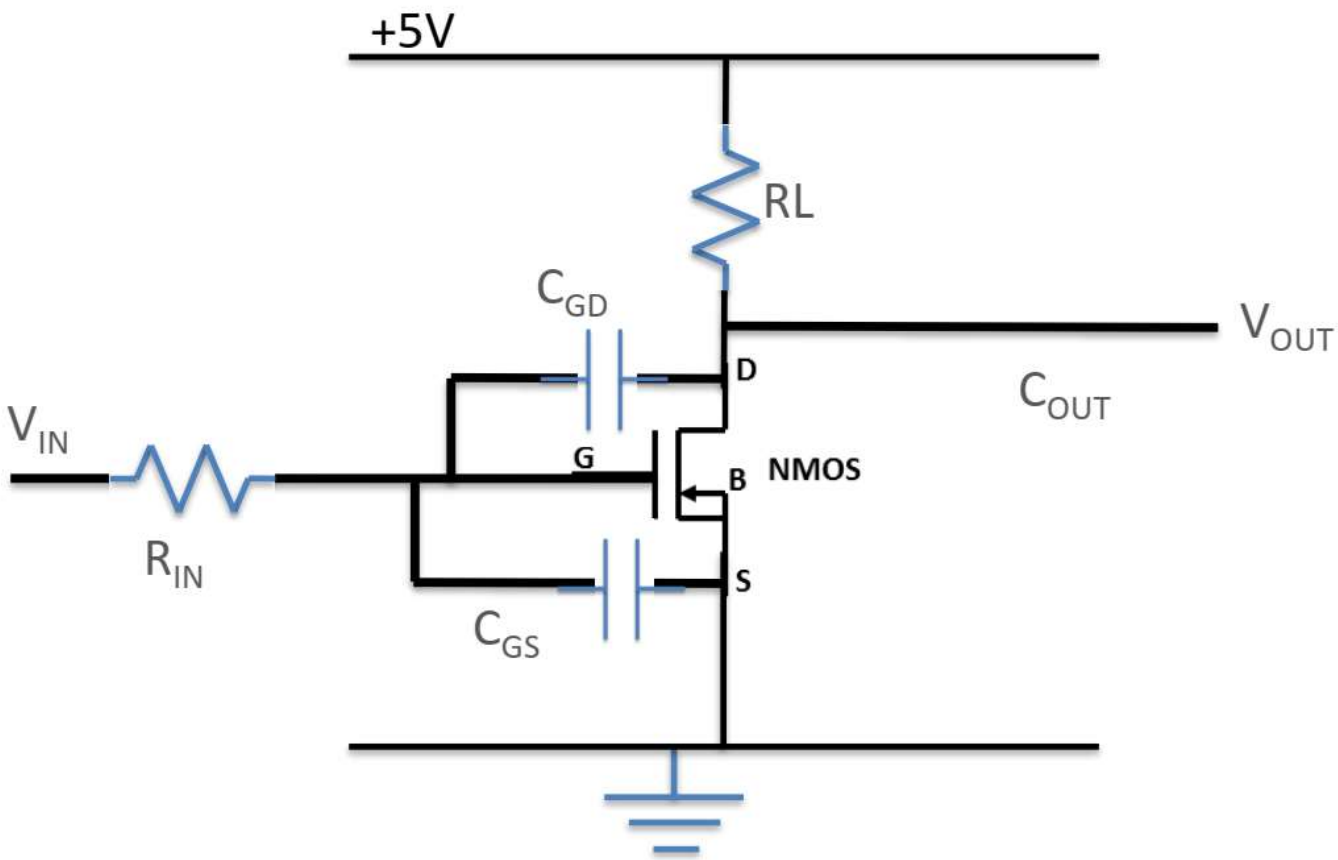
- First, it is good to pick large values of bias resistors so that they take very little current. Any current they take is wasted.
- You should then pick your capacitor value so that the gain is back to normal BEFORE you get to frequencies of interest.
- What is the lowest frequency of interest? It depends on your application, but for example:

Lowest human-detectable audio signal = 20 Hz

- If you are taking sensitive measurements, you may want to avoid the hum from AC electricity which is everywhere, at 50 Hz and its harmonics at 150 and 250 Hz. This characteristic here works in our favour.

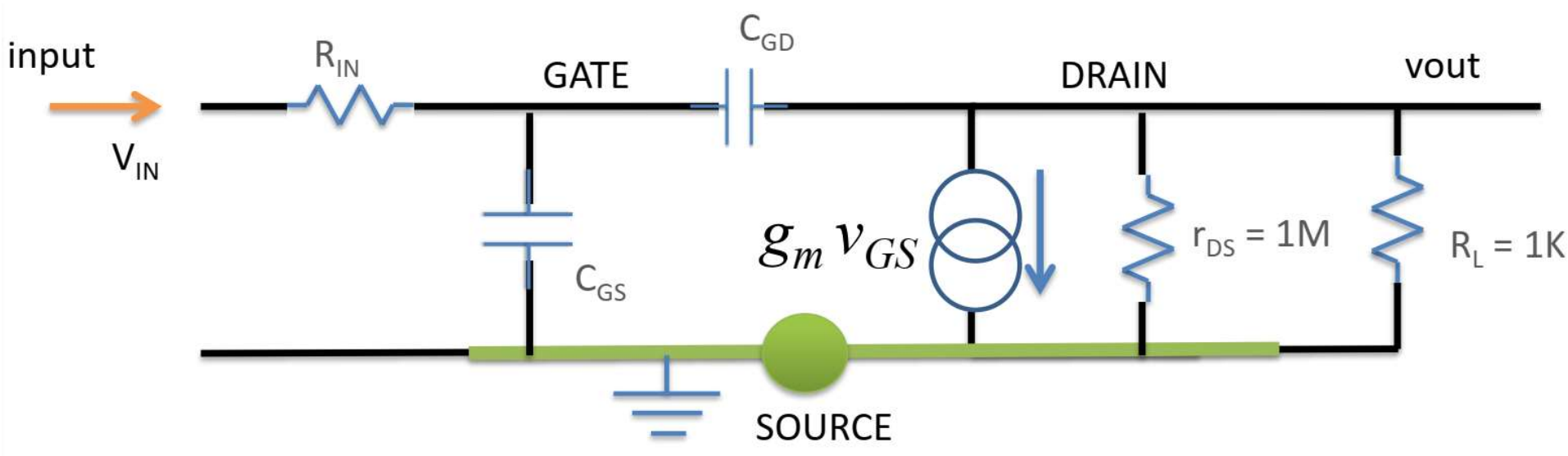
Power-line interference = 50 Hz and harmonics

There Are More Capacitors!



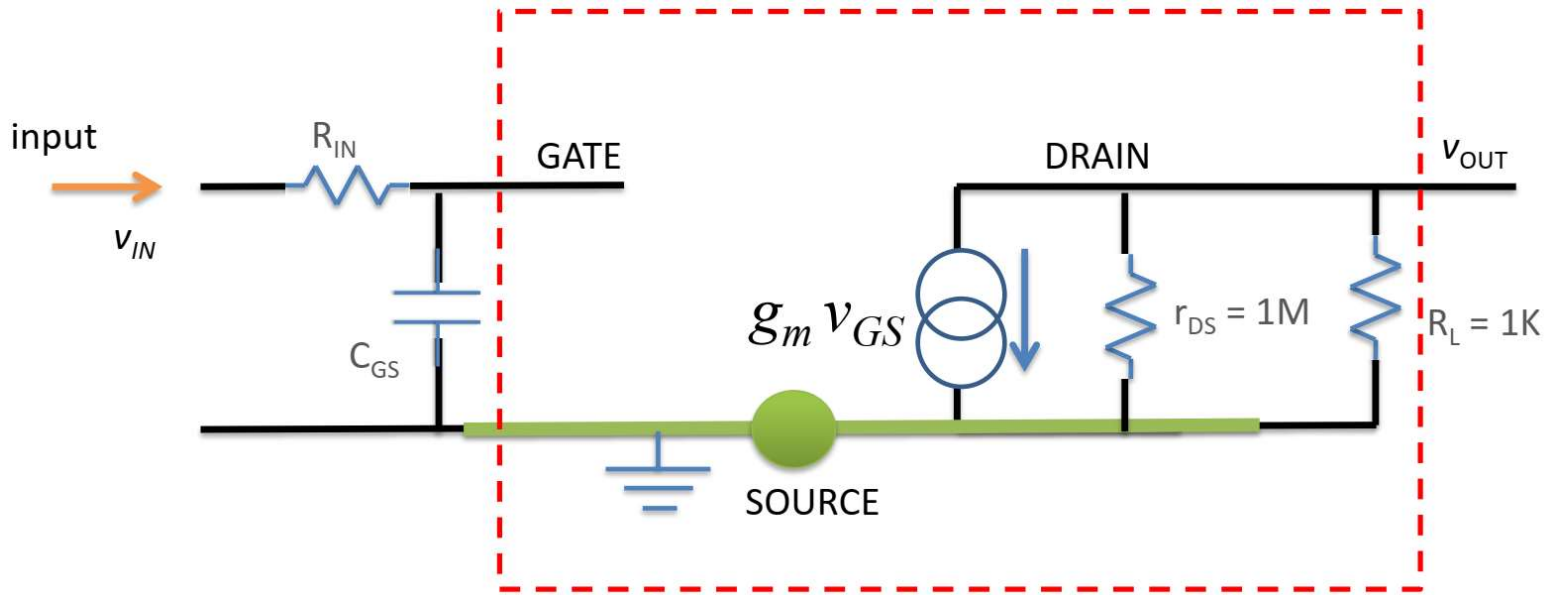
- Here we have two parasitic capacitances (C_{GD} and C_{GS}). We are ignoring any bias resistors or coupling capacitors;
- We have an R_{IN} which corresponds to a source resistance belonging to the input circuit. The resistance of the circuit usually creates the signal.

Small-Signal Analysis



- In this case, we need to look at what these small parasitic capacitors will do to our signal.
- Let's look at them separately.

C_{GS} First



- Let's divide the circuit into two parts and recall what we've already learnt that the gain of the part in the dashed-box is given by $-g_m R_L$.
- But the value of V_{GS} is given by the voltage divider:

$$v_{GS} = v_{IN}\left(\frac{Z_{CGS}}{R_{IN} + Z_{CGS}}\right) = v_{IN}\left[\frac{1}{(R_{IN}/Z_{CGS}) + 1}\right]$$

C_{GS} First

$$A_{\text{real}} = \frac{v_{\text{OUT}}}{v_{\text{IN}}} = \frac{v_{\text{OUT}}}{[\left(\frac{R_{\text{IN}}}{Z_{\text{CGS}}}\right) + 1]v_{\text{GS}}} = \frac{1}{\left(\frac{R_{\text{IN}}}{Z_{\text{CGS}}}\right) + 1}\left(\frac{v_{\text{OUT}}}{v_{\text{GS}}}\right)$$

$$A_{\text{real}} = A\frac{1}{1 + j\omega R_{\text{IN}}C_{\text{GS}}}$$

- Let's just look at the magnitude of the gain:

$$|A_{\text{real}}| = |A|\frac{1}{\sqrt{1 + (\omega R_{\text{IN}}C_{\text{GS}})^2}}$$

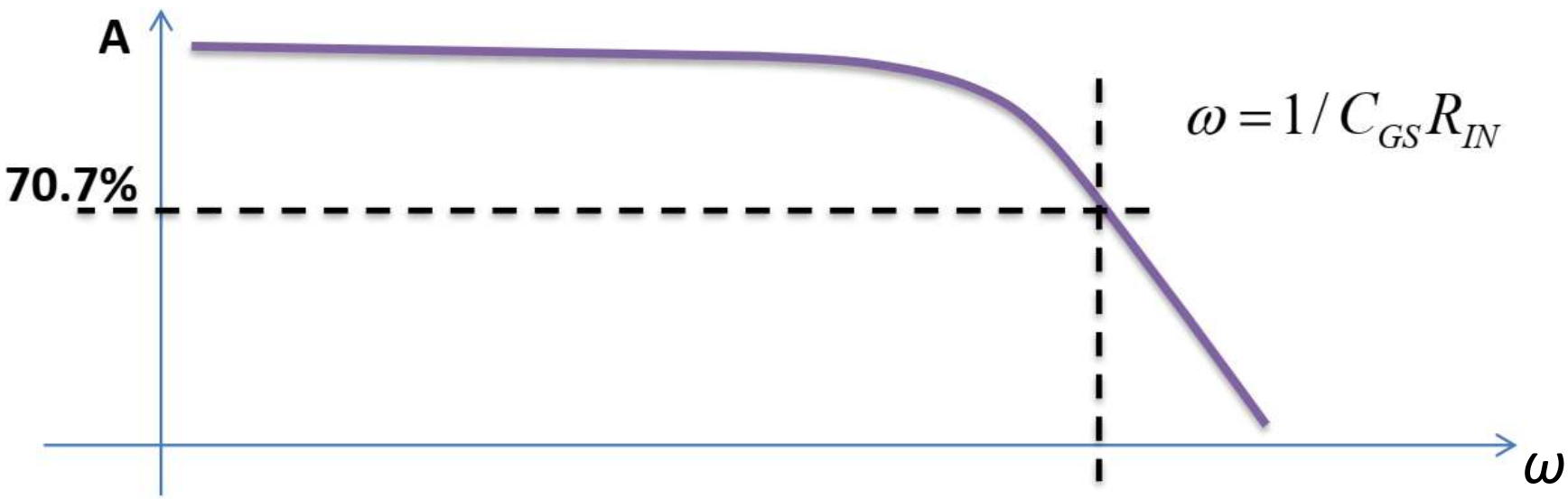
The Impact of C_{GS}

$$|A_{\text{real}}| = |A| \frac{1}{\sqrt{1 + (\omega R_{\text{IN}} C_{\text{GS}})^2}}$$

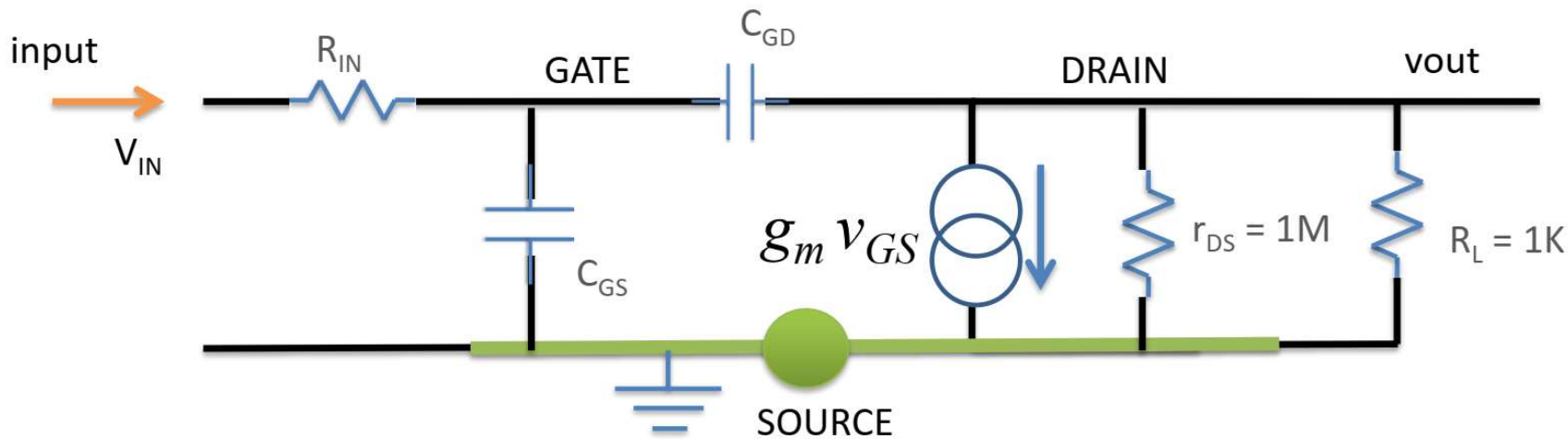
- Without looking at this too carefully, we know:
  - if $\omega << 1/(R_{IN}C_{GS})$, then $A_{real} = A$;
  - if $\omega = 1/(R_{IN}C_{GS})$, then $A_{real} \approx 70.7\%$ of A;
  - if $\omega >> 1/(R_{IN}C_{GS})$, then $A_{real} \approx 0$.
- C_{GS} stops the circuit working at high frequencies. Any capacitor, from wires, or breadboards that connects the gate to source will reduce the high frequency performance of the circuit.
- A large value of R_{IN} will cause the cut-off frequency to be lowered which leads to a narrower bandwidth.

The Impact of C_{GS}

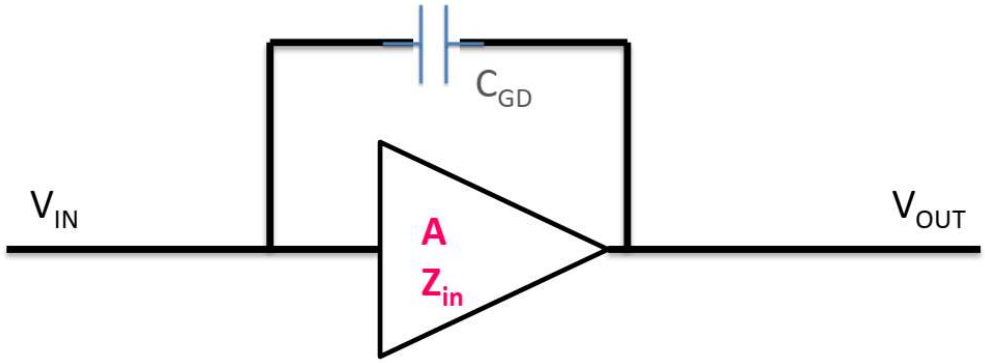
$$|A_{\text{real}}| = |A| \frac{1}{\sqrt{1 + (\omega R_{\text{IN}} C_{\text{GS}})^2}}$$



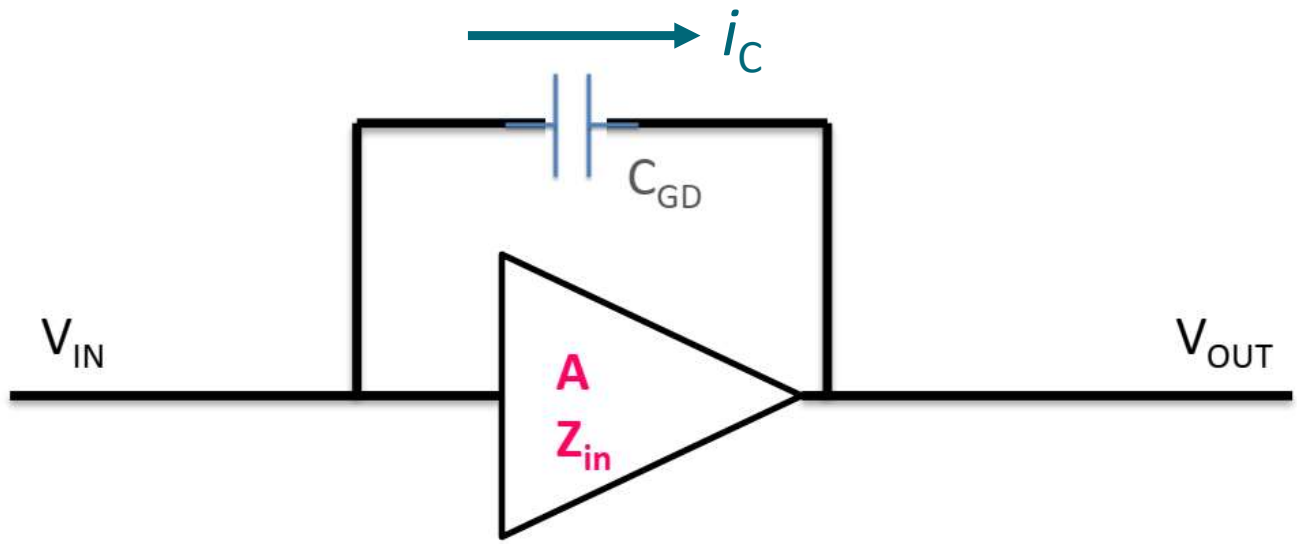
Now for C_{GD}



- Let's ignore C_{GS} as we know what it does.
- Let's imagine we have a simple amplifier, with gain A and impedance Z_{IN}.



Now for C_{GD}



- In a MOSFET, we know that Z_{IN} is very large. The current going through C_{GD} is given by the equation

$$i_C = \frac{v_{IN} - v_{OUT}}{Z_{CGD}} = \frac{v_{IN} - Av_{IN}}{Z_{CGD}} = (1 - A)\frac{v_{IN}}{Z_{CGD}}$$

Now for C_{GD}

$$i_C = (1 - A)\frac{v_{IN}}{Z_{CGD}}$$

- If we can define an effective "input impedance", and i_c, in this case, will become the input current:

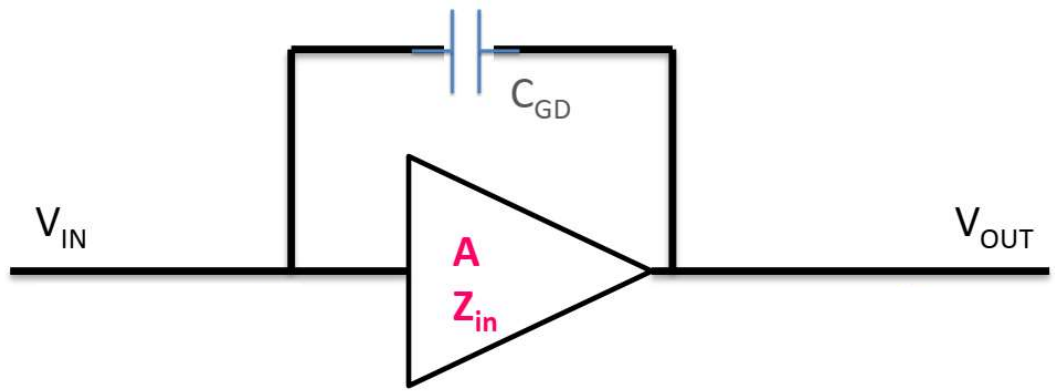
$$Z_{EQ} = \frac{v_{IN}}{i_{IN}} \approx \frac{v_{IN}}{i_C} = \frac{Z_{CGD}}{1 - A}$$

- We can define an equivalent input capacitance C_{EQ}, whose value is given by

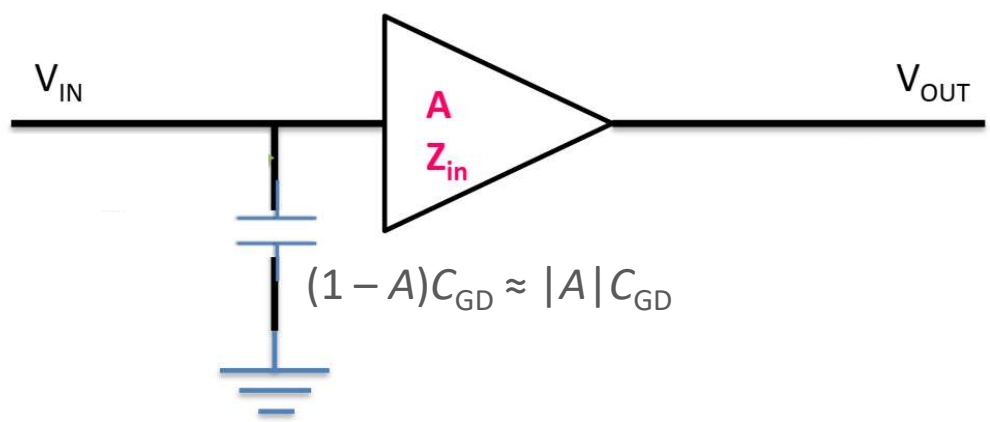
$$Z_{EQ} = \frac{Z_{CGD}}{1-A} \text{ and } C_{EQ} = (1 - A)C_{GD}$$

- The equivalent input capacitor C_{EQ} is actually the same as the C_{GD} capacitor multiplied by (1 - A) of the circuit. This is called the Miller effect.

The Miller Effect

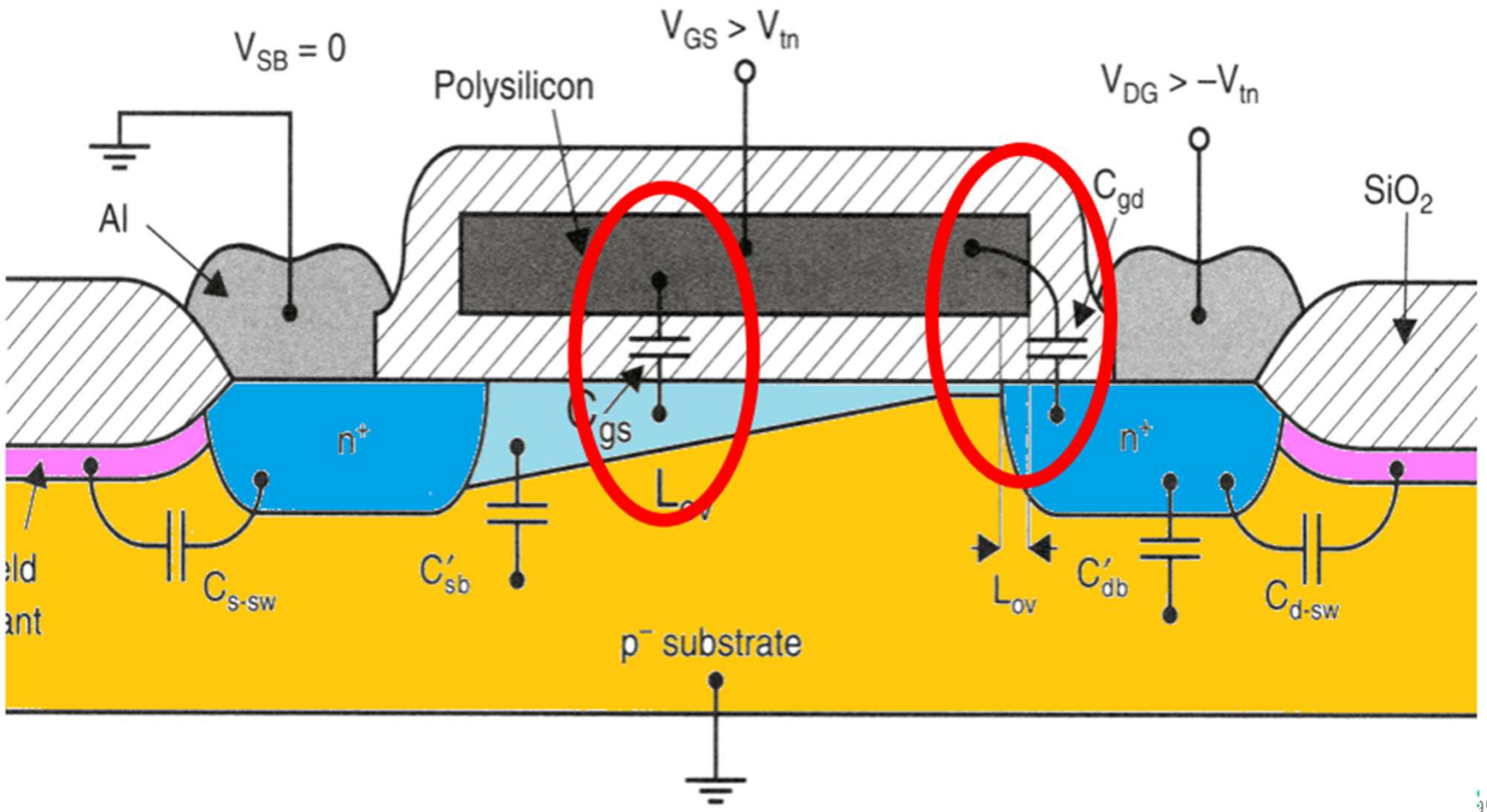


- The Miller effect says that from the perspective of the INPUT, an impedance (capacitor, inductor, resistor) going from input to output is scaled up by the gain of the amplifier. This is a big problem for high-gain systems.
- This is an approximation and does not cover all effects of that impedance.

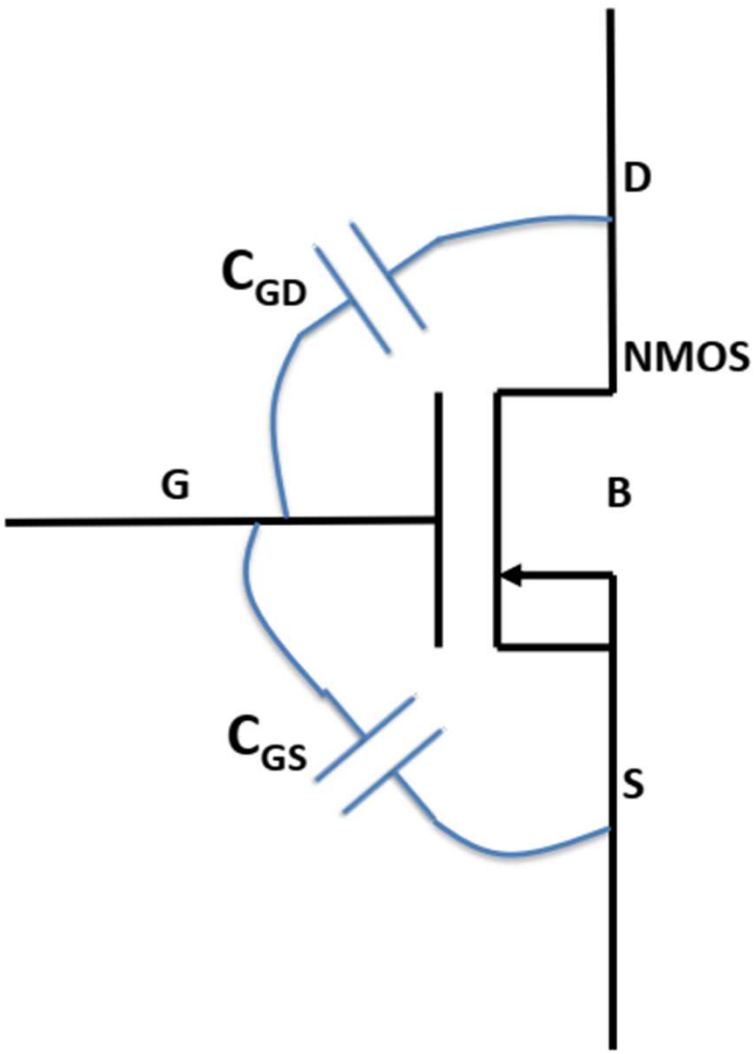


Parasitic Components in MOSFETs

Capacitors form when charged areas are separated by an insulator. Any conducting region overlapping with another is potentially a capacitor.

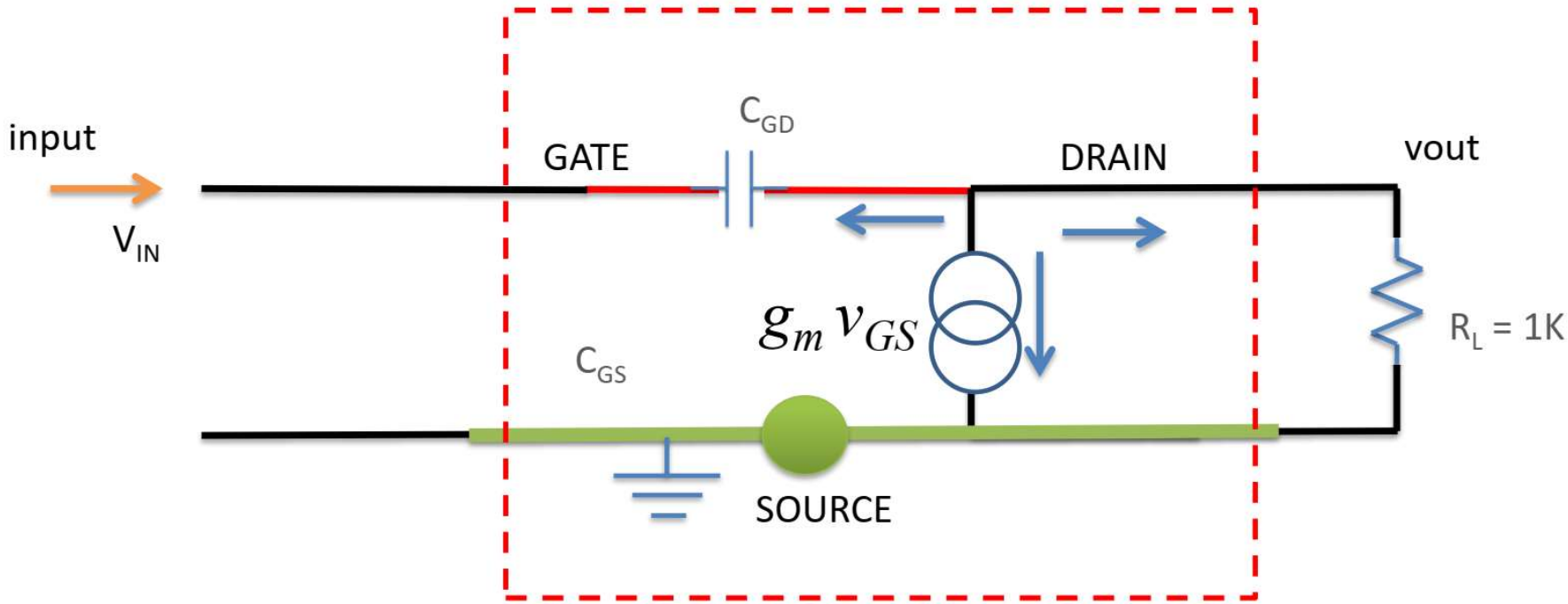


Parasitic Components in MOSFETs



- There are other parasitic capacitors, but these are the two most important ones. We won't initially worry about them, but we will have to come back to them.
- They do not cause a problem when we are changing things slowly as they charge/discharge rapidly (compared to how fast things change).
- But when we are doing things 100's and 1000 of million times / sec, it is these parasitic capacitances that limit the performance.
- C_{GS} is the biggest in size but C_{GD} has the biggest impact.

Now for C_{GD}: A Different Analysis



- Let us conduct a nodal analysis at the drain:

$$i_{\mathrm{DG}} + g_m v_{\mathrm{GS}} + i_{\mathrm{RL}} = 0$$

$$\left(\frac{v_{\mathrm{DS}} - v_{\mathrm{GS}}}{Z_{\mathrm{CGD}}}\right) + (g_m v_{\mathrm{GS}}) + \left(\frac{v_{\mathrm{DS}}}{R_{\mathrm{L}}}\right) = 0$$

Now for C_{GD}: A Different Analysis

- Now remember $v_{IN} = v_{GS}$ and $v_{OUT} = v_{DS}$, and we want to find gain $A = v_{OUT}/v_{IN}$.

$$\left(\frac{v_{\mathrm{DS}} - v_{\mathrm{GS}}}{Z_{\mathrm{CGD}}}\right) + (g_m v_{\mathrm{GS}}) + \left(\frac{v_{\mathrm{DS}}}{R_{\mathrm{L}}}\right) = 0$$

$$v_{\mathrm{DS}}\left(\frac{1}{Z_{\mathrm{CGD}}} + \frac{1}{R_{\mathrm{L}}}\right) = v_{\mathrm{GS}}(-g_m + \frac{1}{Z_{\mathrm{CGD}}})$$

- Therefore, $A = v_{OUT}/v_{IN} = v_{DS}/v_{GS}$ is given by

$$A = \frac{-g_m + \frac{1}{Z_{\mathrm{CGD}}}}{\frac{1}{Z_{\mathrm{CGD}}} + \frac{1}{R_{\mathrm{L}}}} = \frac{-g_m R_{\mathrm{L}} + \frac{R_{\mathrm{L}}}{Z_{\mathrm{CGD}}}}{1 + \frac{R_{\mathrm{L}}}{Z_{\mathrm{CGD}}}}$$

Now for C_{GD}: A Different Analysis

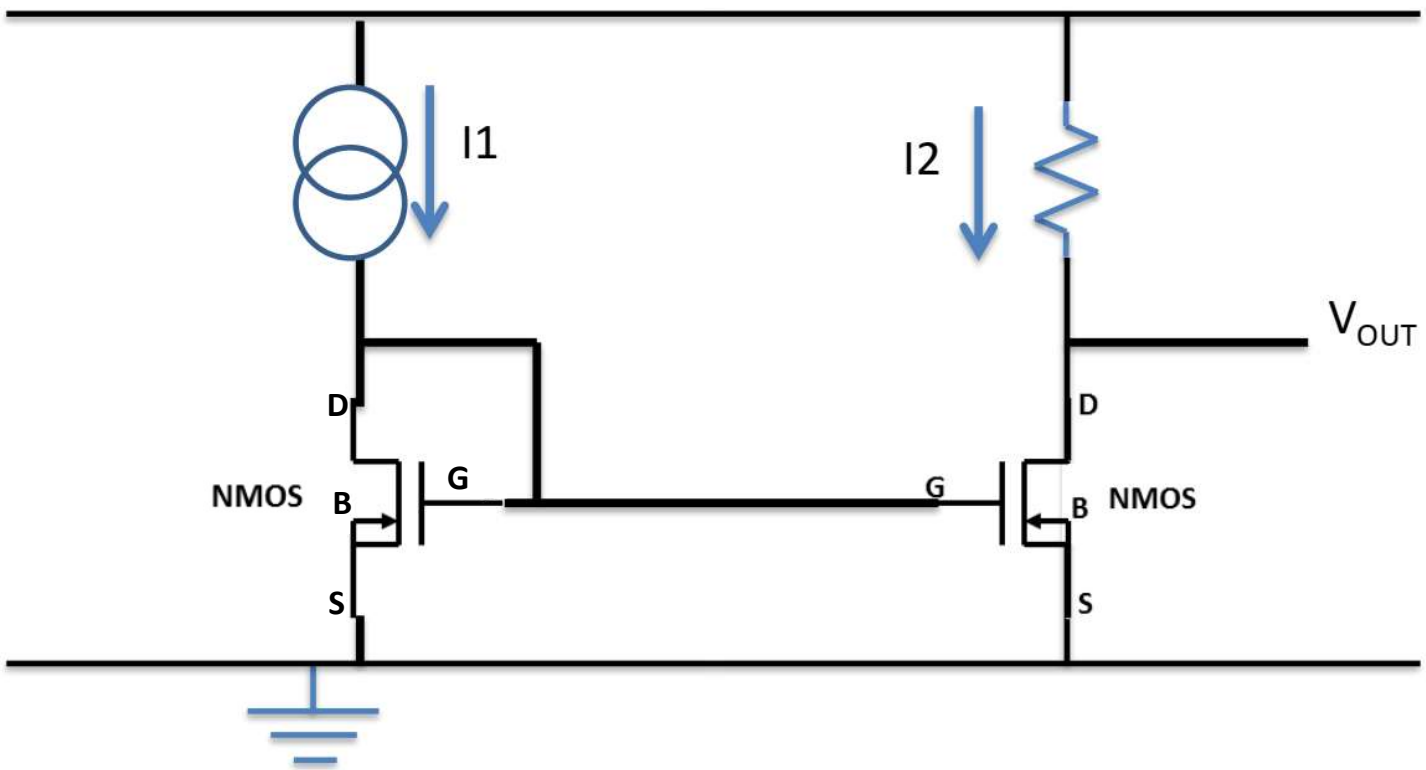
$$A = \frac{-g_m R_L + \frac{R_L}{Z_{CGD}}}{1 + \frac{R_L}{Z_{CGD}}} = \frac{-g_m R_L + j\omega C_{GD} R_L}{1 + j\omega C_{GD} R_L}$$

- This is messy, but let's see what it says. If C_{GD} is small, then the gain goes back to $-g_m R_L$.
- Let's see what happens when it is not small (or $\omega R_L C_{GD} >> 1$):

$$A \approx \frac{-g_m R_L + j\omega C_{GD} R_L}{j\omega C_{GD} R_L} = 1 - \frac{g_m}{j\omega C_{GD}}$$

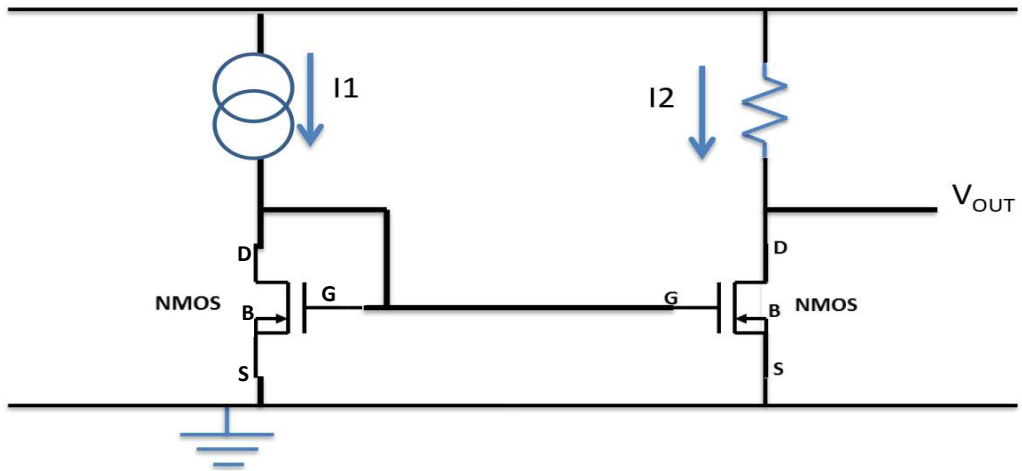
- The capacitor will neutralise the gain and if the frequency is high enough, it will short gate to drain and we will get a copy of the input at the output with no gain. Before that, we get amplification, but it reduces rapidly as the Miller effect tells us.
- In practice, C_{GD} will reduce high frequency gain BEFORE C_{GS} does.

What is Happening Here?



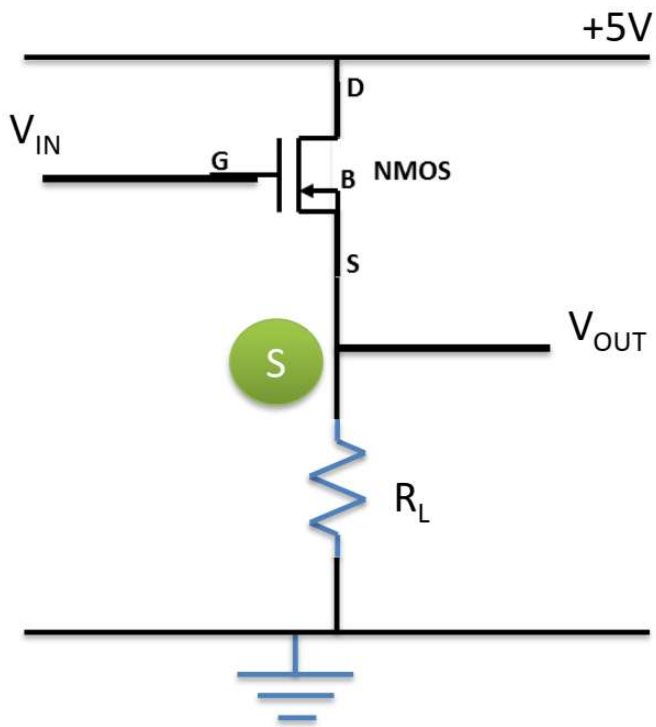
$$I_{D(\mathrm{sat})} = \frac{1}{2}\mu_{\mathrm{n}} C_{\mathrm{ox}} \frac{W}{L}(V_{\mathrm{GS}} - V_{\mathrm{T}})^2$$

Current Mirror



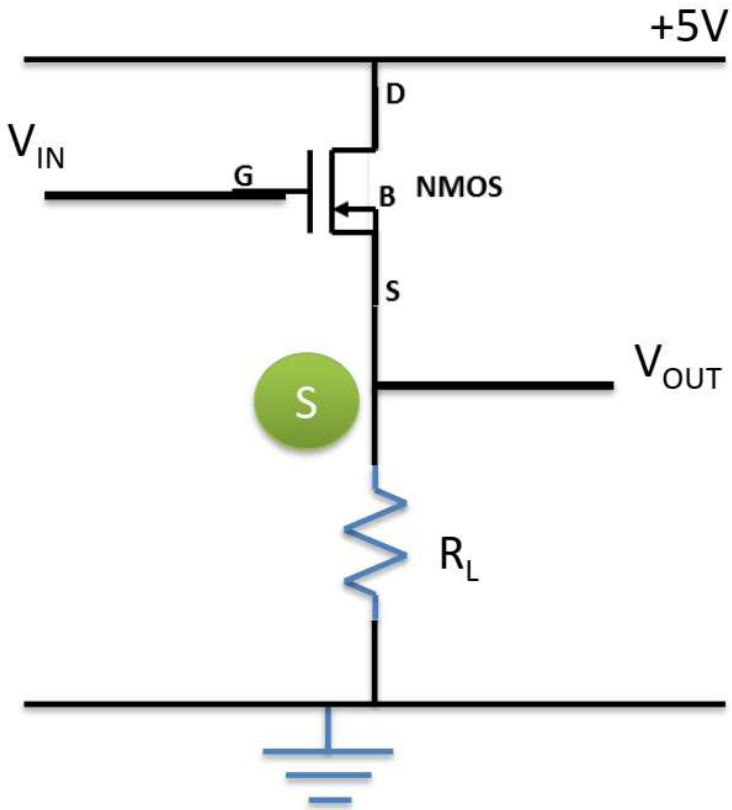
- The transistor equation says that if I_D flows, there must be a V_{GS};
- The two transistors see the exact same V_{GS};
- The transistor equation says that given a value of V_{GS}, a current I_D will flow;
- In an ideal case (provided that $V_{T1} = V_{T2}$), I_D does not depend on V_{DS}, so therefore $I_1 = I_2$;
- Where there is a dependency on V_{DS} (real case), I_2 will differ slightly from I_1;
- This is called a current mirror or current multiplier.
- How would you make it multiply the current? For example, by a factor of 4?

Source Follower Circuit



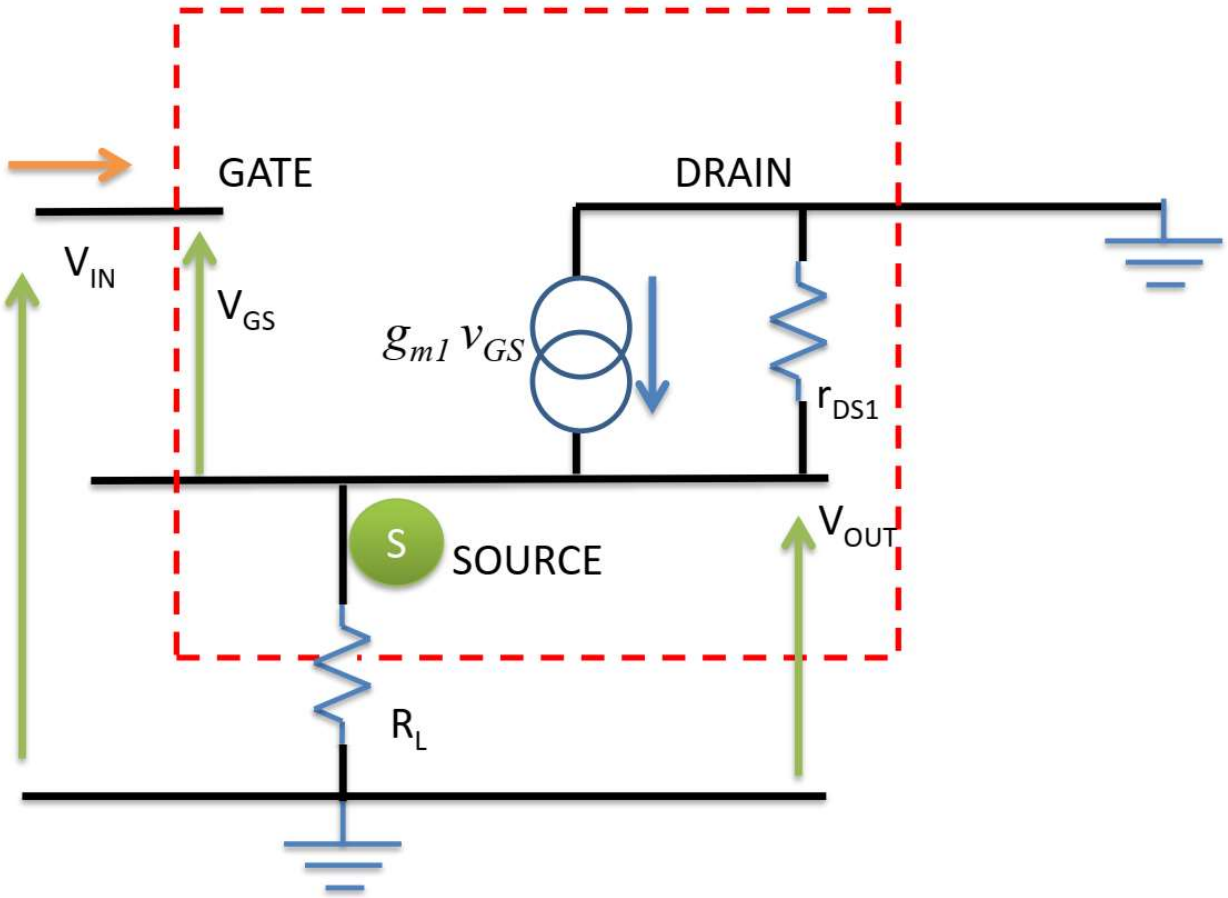
- The previous circuits are called COMMON SOURCE designs as the source is common to both the gate and the drain;
- A transmission gate or switch is typically called COMMON GATE;
- This circuit configuration is called COMMON DRAIN or SOURCE FOLLOWER.

Source Follower Circuit



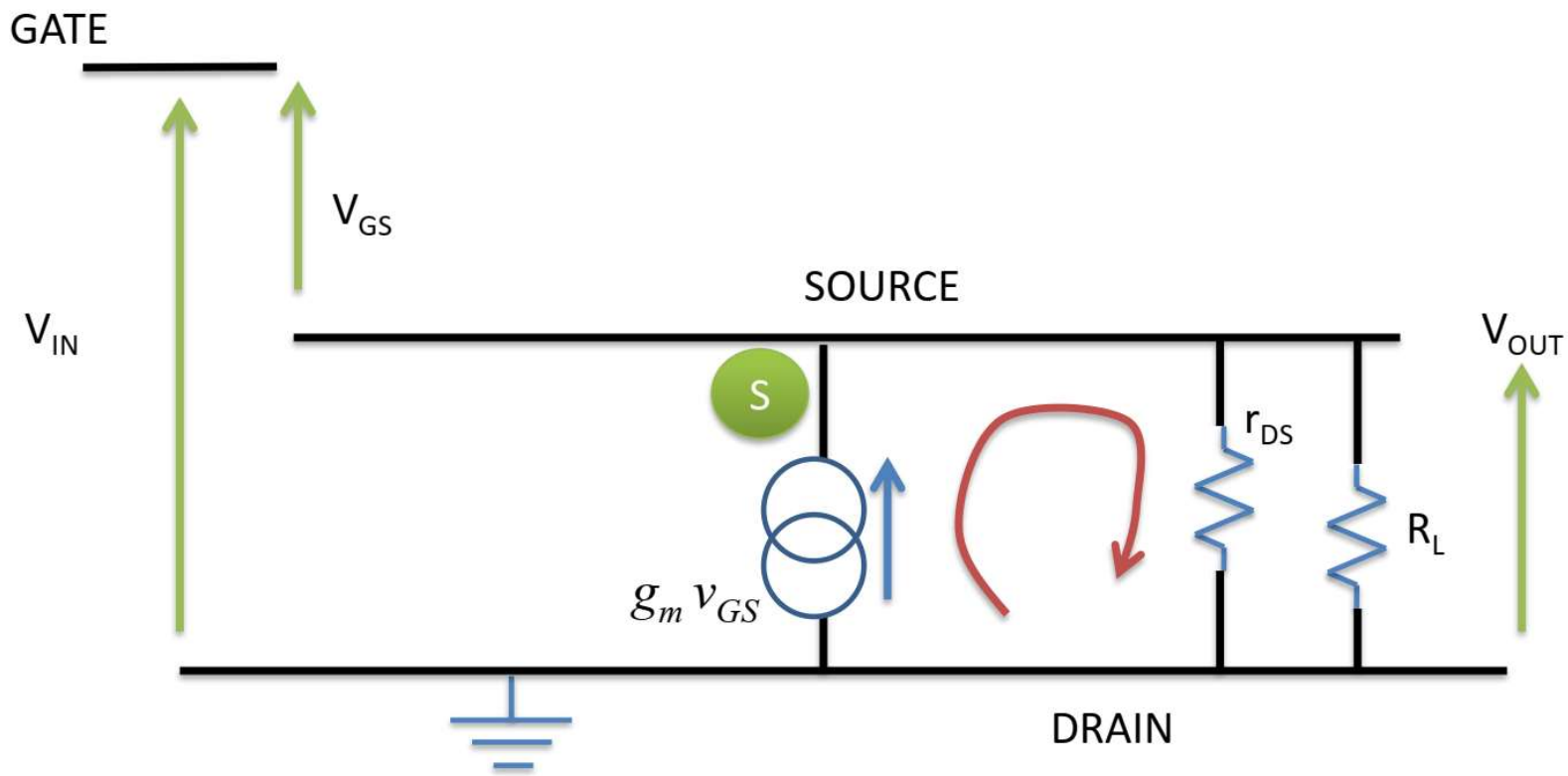
- This circuit uses an NMOS with the drain connected to the positive voltage and a resistor connected to the source. The output is taken at the source.
- Let's work out the gain.

Source Follower Circuit



- The circuit is messy but there is one thing you should note: V_{IN} is **NOT** V_{GS} in this case.

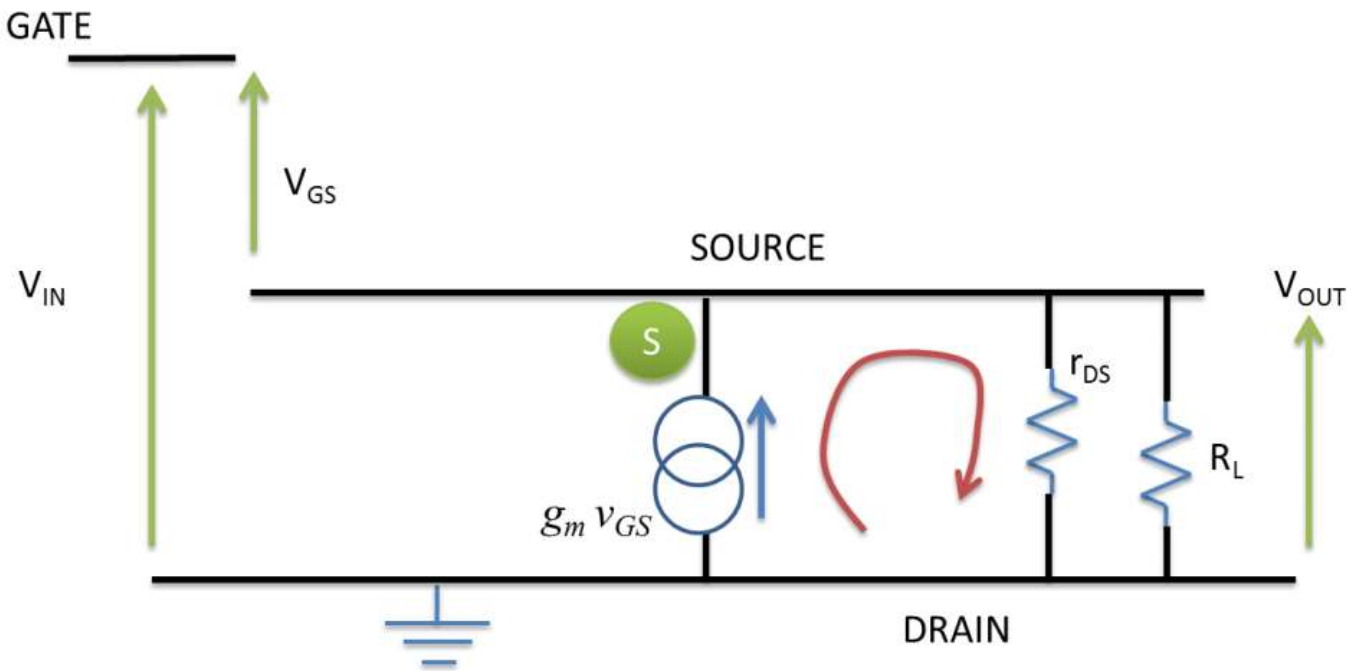
Source Follower Circuit



$$v_{\mathrm{GS}} = v_{\mathrm{IN}} - v_{\mathrm{OUT}}$$

$v_{\mathrm{OUT}} = i(r_{\mathrm{DS}}||R_{\mathrm{L}})$ (note that there is no "-" sign)

Source Follower Circuit

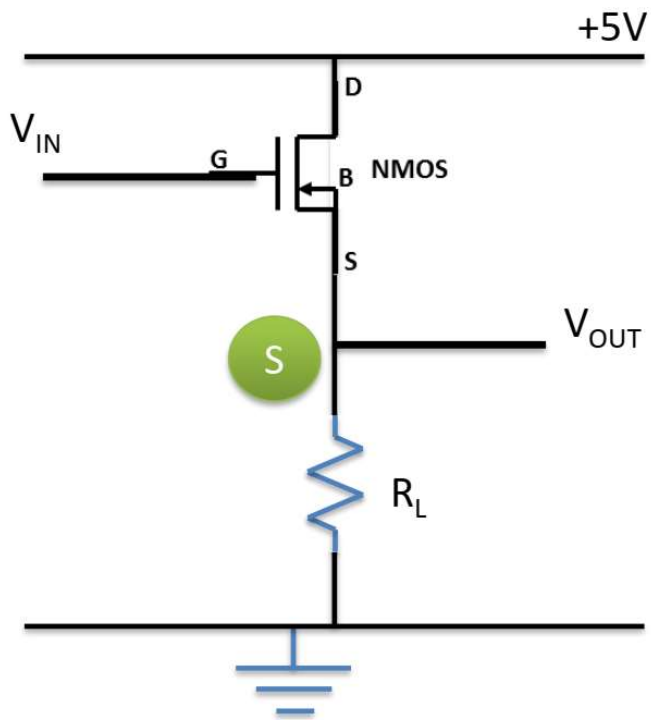


$$v_{\mathrm{OUT}} = i(r_{\mathrm{DS}} \parallel R_{\mathrm{L}}) = g_m v_{\mathrm{GS}}(r_{\mathrm{DS}} \parallel R_{\mathrm{L}}) = g_m(v_{\mathrm{IN}} - v_{\mathrm{OUT}})(r_{\mathrm{DS}} \parallel R_{\mathrm{L}})$$

$$v_{\mathrm{OUT}}[1 + g_m(r_{\mathrm{DS}} \parallel R_{\mathrm{L}})] = v_{\mathrm{IN}} g_m\,(r_{\mathrm{DS}} \parallel R_{\mathrm{L}})$$

$$A = \frac{v_{\mathrm{OUT}}}{v_{\mathrm{IN}}} = \frac{\dfrac{v_{\mathrm{IN}} g_m\,(r_{\mathrm{DS}} \parallel R_{\mathrm{L}})}{1 + g_m(r_{\mathrm{DS}} \parallel R_{\mathrm{L}})}}{v_{\mathrm{IN}}} = \frac{g_m\,(r_{\mathrm{DS}} \parallel R_{\mathrm{L}})}{1 + g_m(r_{\mathrm{DS}} \parallel R_{\mathrm{L}})}$$

Source Follower Circuit



- Are we in saturation?

$$V_{DS} > V_{GS} - V_T$$

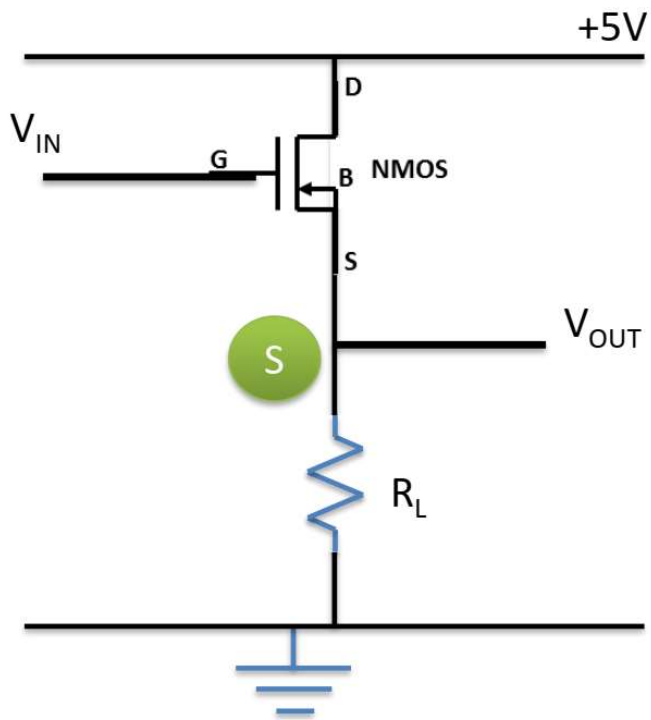
- Yes, as the drain voltage is always higher than the gate voltage (V_D = 5 V), this is true.

$$A = \frac{g_m(r_{DS} \parallel R_L)}{1 + g_m(r_{DS} \parallel R_L)} \approx 1$$

V_{IN} goes up
I_D goes up
$I_D R_L$ goes up
V_{OUT} goes up

- This circuit clones a copy of the input to the output. There is no inversion.
- To understand why it is useful, let's look at the DC situation.

Source Follower Circuit



- V_{GS} is constant for a given drain current. The source voltage will be lower than the gate by whatever voltage is required by the drain current. More current, a greater voltage drop will result.
- If we only have small AC signals changing V_{IN}, then I_D will be constant and V_{OUT} will be some voltage less than V_{IN}. The circuit is used as a level shifter to drop the voltage of the input.
- A PMOS would lift the voltage ($V_{GS} < 0$).

$$V_{\mathrm{OUT}} = V_{\mathrm{IN}} - V_{\mathrm{GS}} \text{ where } I_{\mathrm{D(sat)}} = \frac{1}{2}\mu_{\mathrm{n}} C_{\mathrm{ox}} \frac{W}{L}(V_{\mathrm{GS}} - V_{\mathrm{T}})^2$$

Looking up the value of V_{GS} for I_D on a datasheet will give you approximate voltage drop.

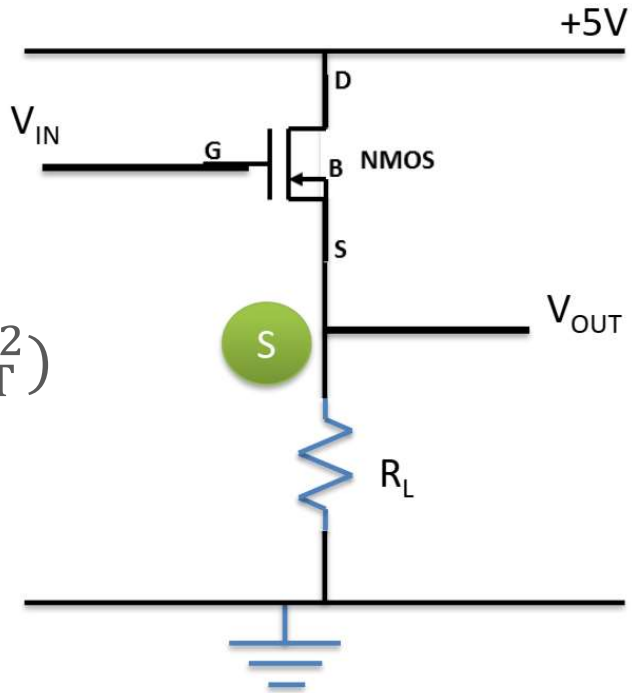
Source Follower Circuit

$$V_{\mathrm{OUT}} = R_{\mathrm{L}} I_{\mathrm{D(sat)}}$$

$$V_{\mathrm{OUT}} = R_{\mathrm{L}} \frac{1}{2} \mu_{\mathrm{n}} C_{\mathrm{ox}} \frac{W}{L} (V_{\mathrm{GS}} - V_{\mathrm{T}})^2$$

$$V_{\mathrm{OUT}} = R_{\mathrm{L}} \frac{1}{2} \mu_{\mathrm{n}} C_{\mathrm{ox}} \frac{W}{L} (V_{\mathrm{GS}}^2 - 2V_{\mathrm{GS}} V_{\mathrm{T}} + V_{\mathrm{T}}^2)$$

And $V_{\mathrm{GS}} = V_{\mathrm{IN}} - V_{\mathrm{OUT}}$

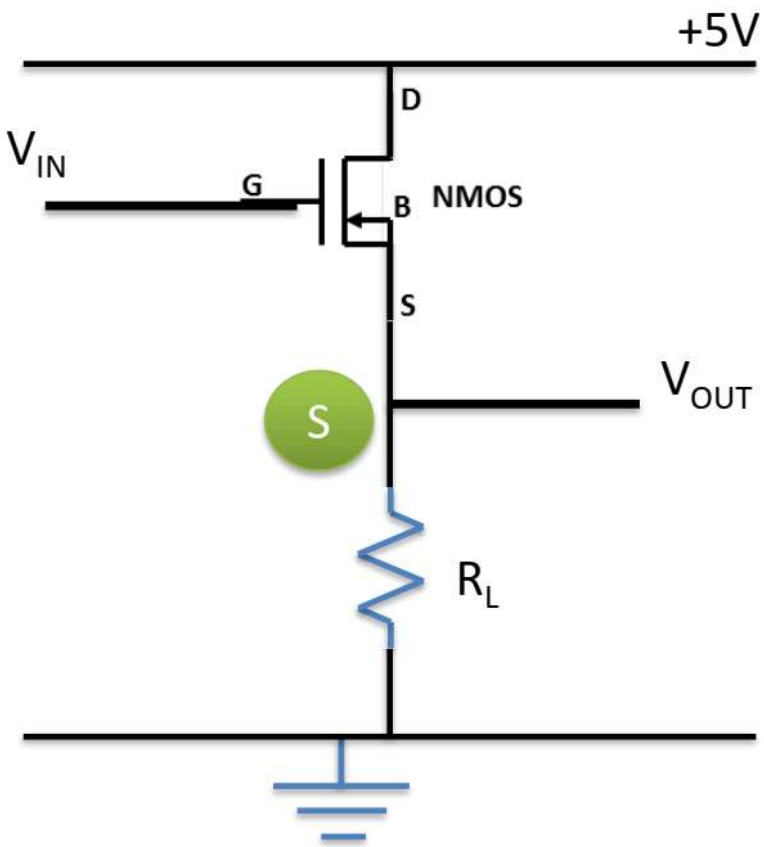


$$V_{\mathrm{OUT}} = R_{\mathrm{L}} \frac{1}{2} \mu_{\mathrm{n}} C_{\mathrm{ox}} \frac{W}{L} [(V_{\mathrm{OUT}} - V_{\mathrm{IN}})^2 + 2(V_{\mathrm{OUT}} - V_{\mathrm{IN}}) V_{\mathrm{T}} + V_{\mathrm{T}}^2]$$

$$V_{\mathrm{OUT}} = \frac{R_{\mathrm{L}}}{2} \mu_{\mathrm{n}} C_{\mathrm{ox}} \frac{W}{L} [(V_{\mathrm{OUT}}^2 - 2V_{\mathrm{OUT}} V_{\mathrm{IN}} + V_{\mathrm{IN}}^2) + 2(V_{\mathrm{OUT}} - V_{\mathrm{IN}}) V_{\mathrm{T}} + V_{\mathrm{T}}^2]$$

No easy analysis is possible. There is a complicated quadratic expression to find out what V_{IN} and V_{OUT} finally settle at. Computers love this stuff. However, R_{L} lets you tweak what that level will be.

Applications of Source Followers



- Level shifters: as already explained. A copy of the AC signal is passed through at a higher/lower voltage depending on if you use a PMOS or NMOS circuit;
- Buffers: in this case imagine you have a weak signal generating the input. The output is being driven by the 5 V supply, it can supply much greater current to R_L (if you need it there) and can charge capacitors connected to V_{OUT} more rapidly. It protects a weak/fragile input from whatever needs to use it.

Summary

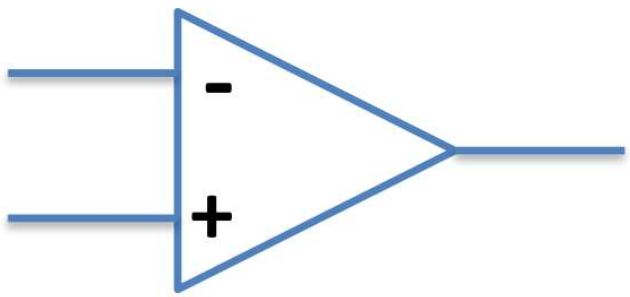
- Common Source:
  - Amplifier, gain = $-g_m R_L$;
  - Can also be used as a current mirror circuit, a current multiplier, or a current source.
- Common Gate:
  - Rarely used as an amplifier (except for some high-frequency applications);
  - Now used mostly as switches.
- Common DRAIN / Source Follower:
  - A level shifter;
  - A buffer to protect weak signals.

EE204FZ
Lecture 8
Operational Amplifiers

Zhu DIAO

Email: zhu.diao@mu.ie

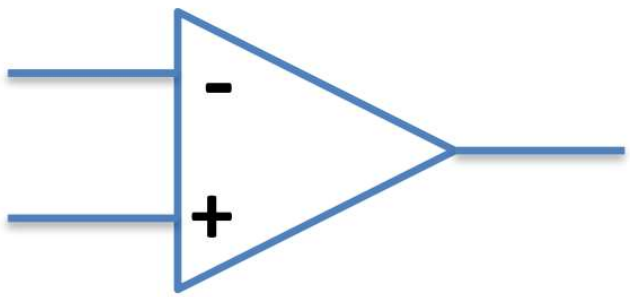
What is an Ideal Op-Amp?



- They amplify the DIFFERENCE of two inputs;
- They have INFINITE input impedance;
- They have ZERO output impedance;
- They have INFINITE gain.

Some other characteristics also include zero offset voltage, infinite bandwidth, etc.

What is a “Practical” Op-Amp?



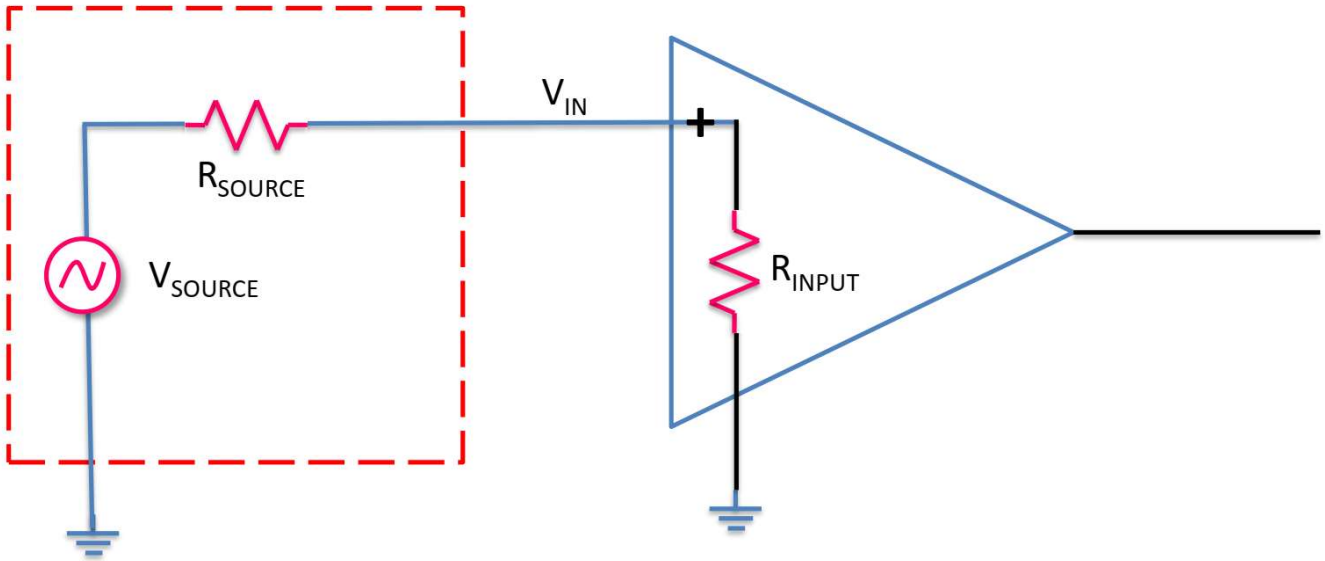
- They amplify the DIFFERENCE of two inputs;
- They have very high input impedance

  (typically allowing 1 – 2 pA of input current, or R > 1 TΩ);
- They have very low output impedance

  (good ones can change output voltage at 500 V/µs);
- They have nearly INFINITE gain

  (up to 200k for low-speed op-amps, > 1k for high-speed devices).

Why is High Input Impedance Important?



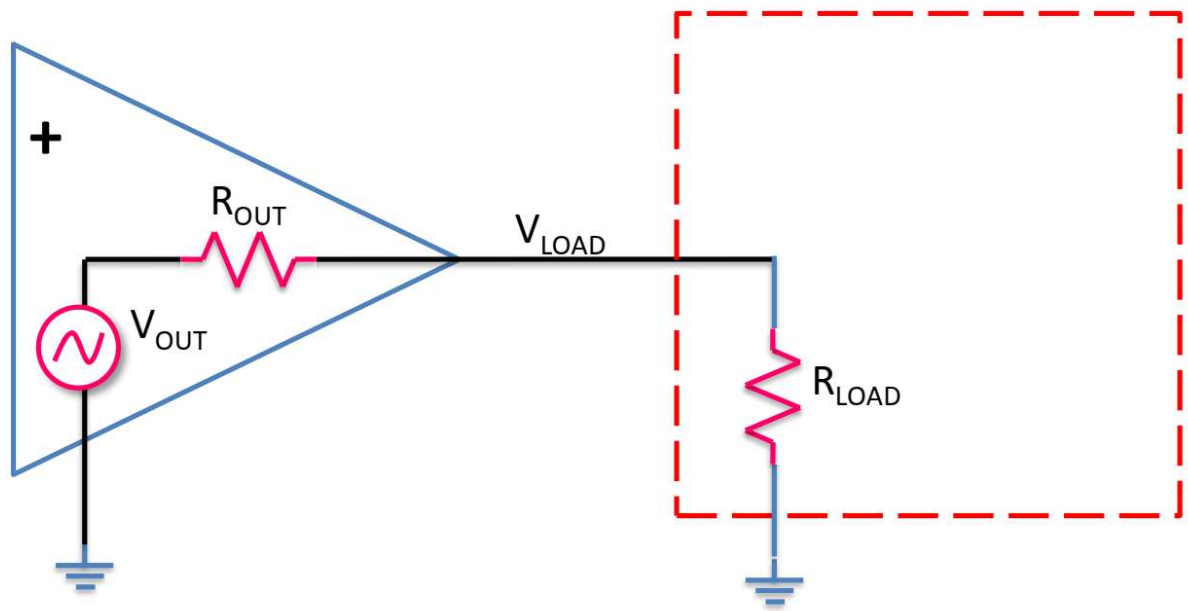
$$V_{\mathrm{IN}} = V_{\mathrm{SOURCE}} - R_{\mathrm{SOURCE}} I_{\mathrm{IN}}$$

$$= V_{\mathrm{SOURCE}} - R_{\mathrm{SOURCE}} \frac{V_{\mathrm{SOURCE}}}{R_{\mathrm{SOURCE}} + R_{\mathrm{IN}}}$$

$$= V_{\mathrm{SOURCE}} \left(1 - \frac{R_{\mathrm{SOURCE}}}{R_{\mathrm{SOURCE}} + R_{\mathrm{IN}}}\right)$$

- To make $V_{\mathrm{IN}} = V_{\mathrm{SOURCE}}$, we need to make I_{IN} zero, or make R_{IN} infinitely large.
- Pulling current from a high output impedance source drops voltage.

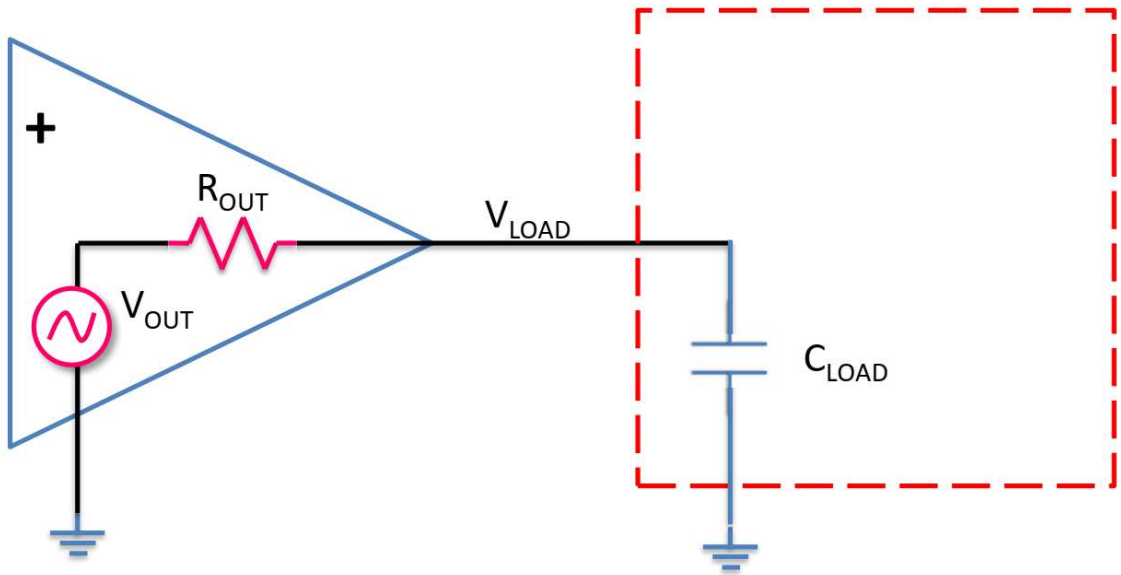
Why is Low Output Impedance Important?



$$V_{\mathrm{LOAD}} = V_{\mathrm{OUT}} - R_{\mathrm{OUT}} I_{\mathrm{OUT}} = R_{\mathrm{LOAD}} I_{\mathrm{OUT}}$$

- Like last time, we can make $V_{LOAD} = V_{OUT}$ if we make I_{OUT} equal to zero. This occurs if R_{LOAD} is very large. This can occur if we are connecting to a high impedance input (e.g., a FET). If R_{OUT} was very large, then by the resistor divider rule, almost no signal would be seen at R_{LOAD}.

Why is Low Output Impedance Important?

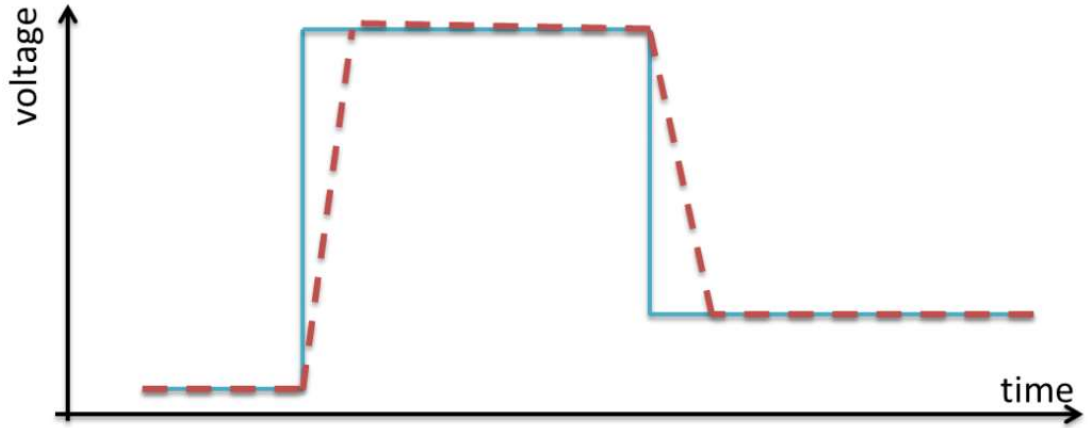


- If we have a capacitor as the output, then the capacitor will need to charge/discharge. The time it takes to get to 97% is approximately 6 time constants. Of course it depends on where you start and end. However, as a time constant is in this case

$$\tau = R_{\mathrm{OUT}} C_{\mathrm{LOAD}}$$

- Reducing R_{OUT} will mean that we get more current and thus more rapid charging/discharging of the capacitors. R_{OUT} = 0 would mean infinitely fast changes in the output voltage.

What is the Slew Rate?

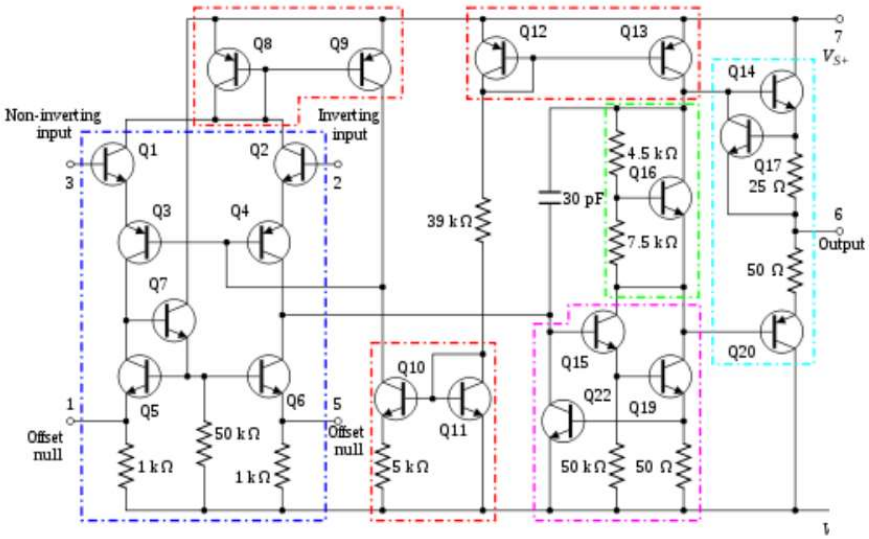


- When the output is about to change, the presence of capacitors and inductors will slow down the change – capacitors are usually the problem.
- The slew rate is the fastest speed at which an output can change. It is of course dependent on the amount of capacitance connected to the output.
- There is no such thing as infinitely fast slew rates as that would mean zero output impedance (infinite current delivery) or zero capacitance.

LM741 slew rate = 5 V/µs at 100 pF

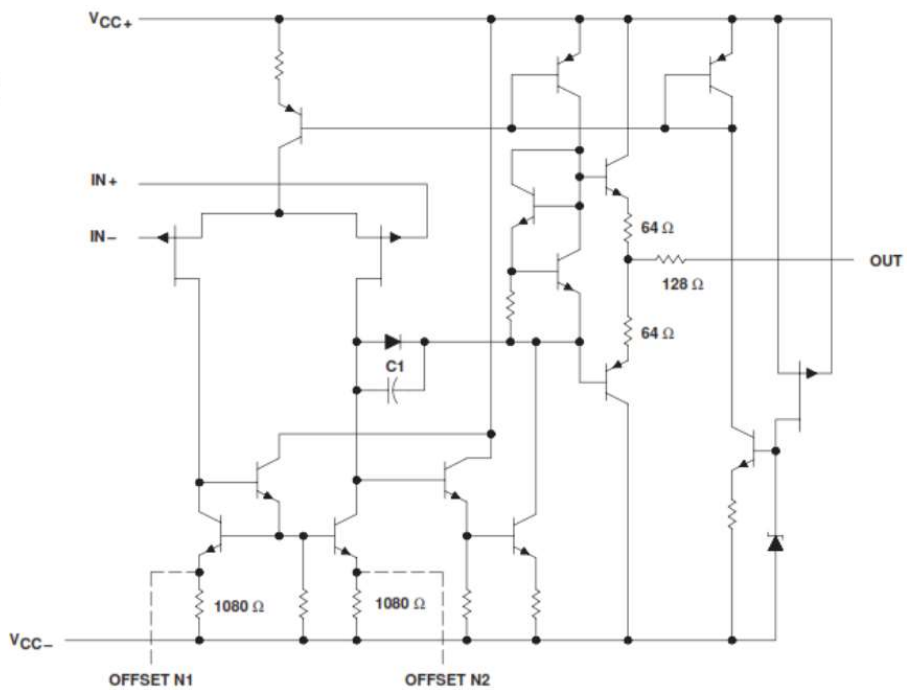
- **Note**: a 1 MHz signal with 5 V_{pk} goes from 0 to 5 V in a quarter period or 0.25 µs, so this op-amp is at least 4 times too slow for a 1 MHz signal of this amplitude.

What is Inside an Op-Amp?

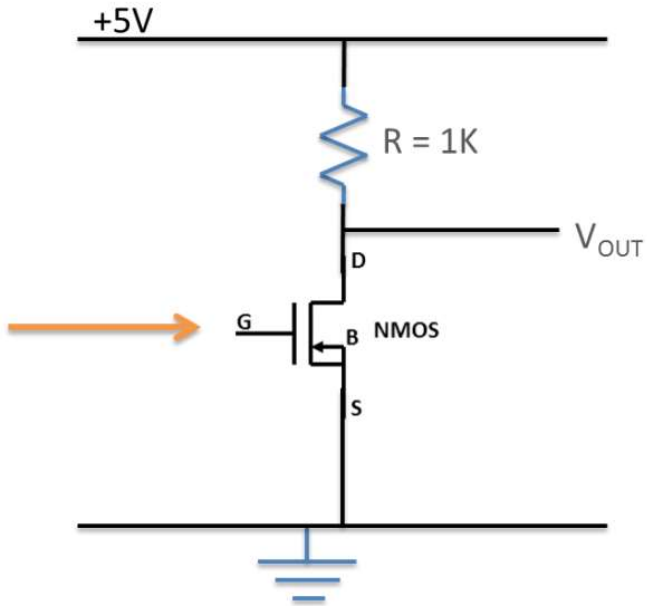


BJT implementation (LM741)

CMOS implementation (TL081)



The Key Design Philosophy for Op-Amps



- The gain of this amplifier is $-g_m R_L$.
- Imagine I want a gain of 10.0, how do I get it?
- Each of these is dependent on temperature and other characteristics.
- Resistors are 5% accurate if you are lucky.
- Manufacturing effects could add another few percent error.
- In silicon technology, ±33% is considered good.

What is a Typical Op-Amp Gain?

- We don't know… nor do we really care.
- We ask for it to be LARGE, and ideally as large as possible. How large depends on the application. The first place to look is the datasheet. It will depend on frequency, but more about that later.

Gain=100K
at low freq.

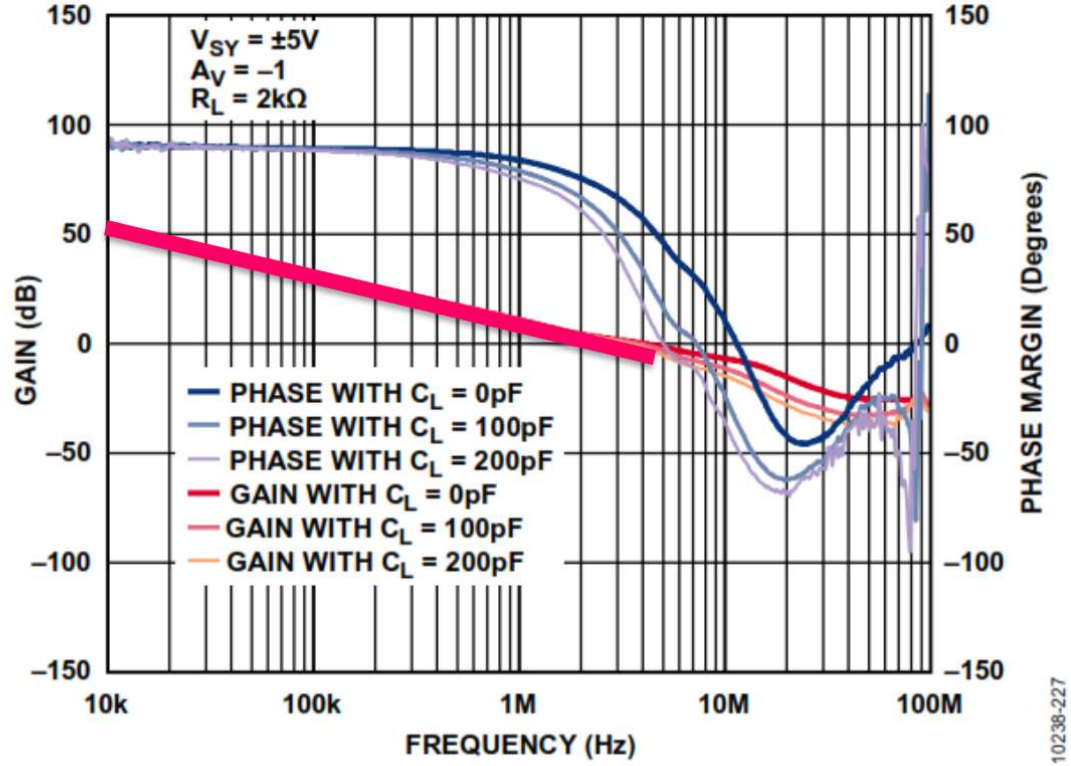
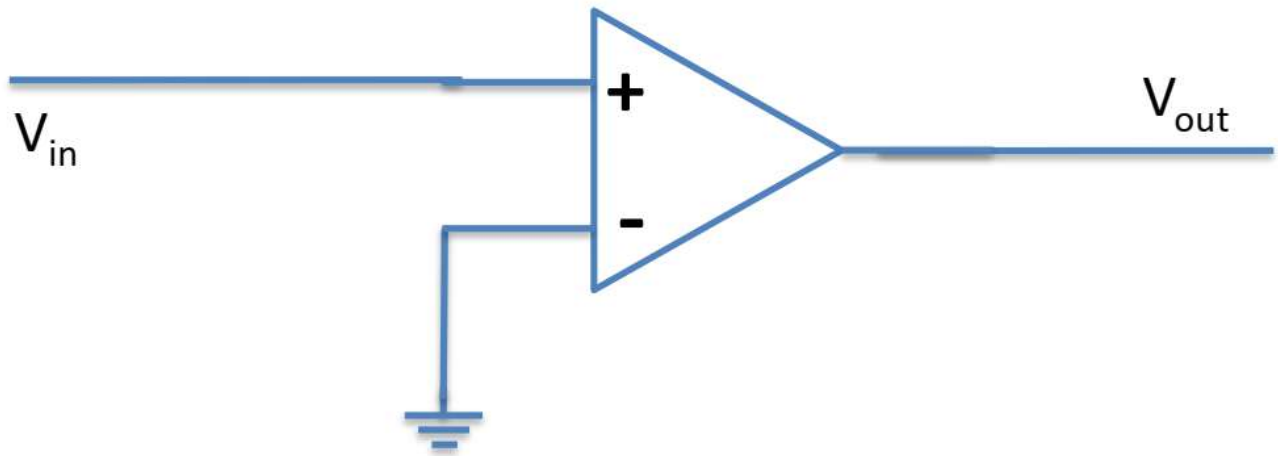


Figure 28. Open-Loop Gain and Phase vs. Frequency, V_{SY} = ±5 V

The Key Design Philosophy for Op-Amps

- We do not care what the values of our parameters are, provided that
  - THE GAIN IS BIG ENOUGH;
  - THE INPUT IMPEDANCE IS HIGH ENOUGH;
  - THE OUTPUT IMPEDANCE IS LOW ENOUGH.
- With all of these, the aim is to design/use an op-amp that goes so far beyond our requirements, that we don't mind if it is 50% out.

Using an Op-Amp

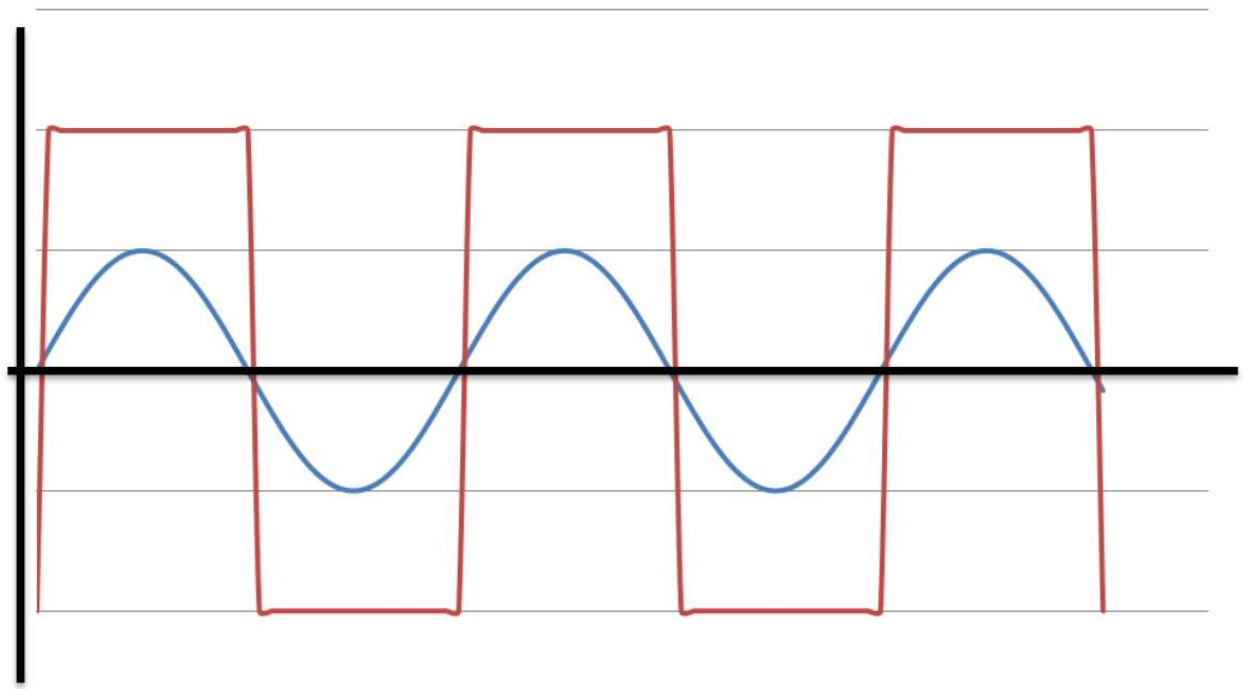


- Consider a zero-mean sinusoidal signal (V_{IN}). The output of the op-amp is given by the expression

$$V_{\text{OUT}} = A(V_{+} - V_{-})$$

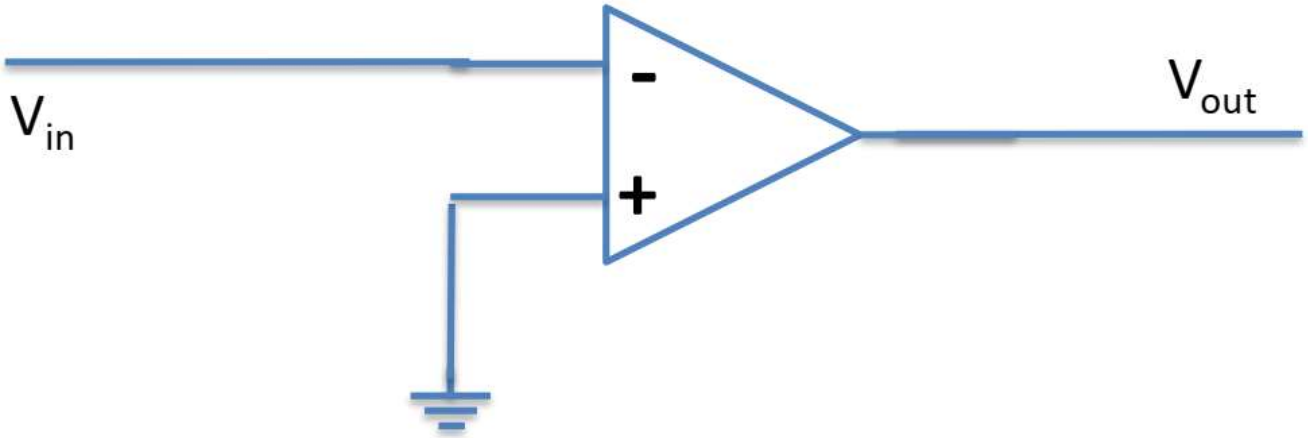
- If A is equal to 100k, and the maximum output voltage is (say) ±10 V, then the output will look like this.

Using an Op-Amp



- If V_{CC} = 10 V and the gain is 100k, then even 0.1 mV of a difference between the inputs and the outputs would be sufficient to make the output hit the +/- voltage rails.
- This use of an op-amp works very well as a comparator. Is A>B?

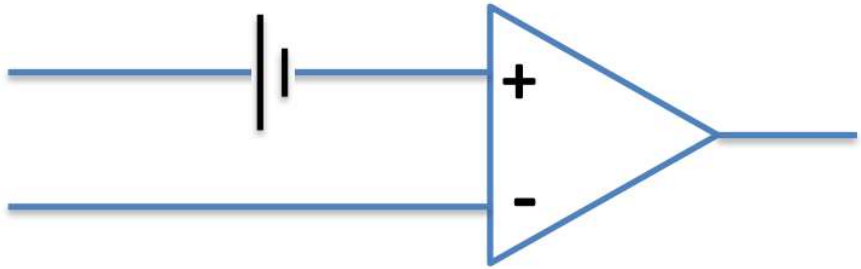
Using an Op-Amp



- If I swap the inputs so that V_{IN} goes into the MINUS pin, what do you think the output would look like now?

Other Characteristics: Input Offset Voltage

- Due to manufacturing tolerances and component variations, op-amps will never have perfectly matched inputs. One input will be higher in electric potential than the other so that even if both are equal externally, internally there would be a difference. This is called the input offset voltage and it looks like a small voltage added to the PLUS input.



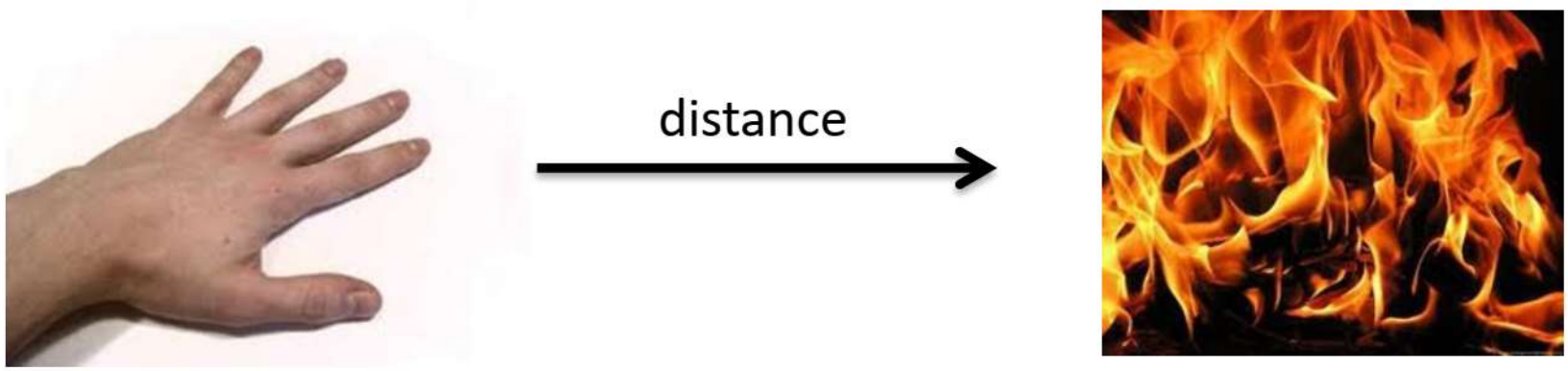
- This limits how small a difference you can detect reliably, or if you wish to take very sensitive measurements. There are circuit tricks to make it smaller, but you need to pick special op-amps.

Chopper stabilized op-amps: < 1 μV
Best bipolar op-amps: 10 – 25 μV
High-speed op-amps: 100 – 2000 μV
Trimmed CMOS op-amps: < 1000 μV

General purpose op-amps: 50 – 500 μV
Best JFET input op-amps: 100 – 1000 μV
Cheap CMOS op-amps: 5000 – 50,000 μV

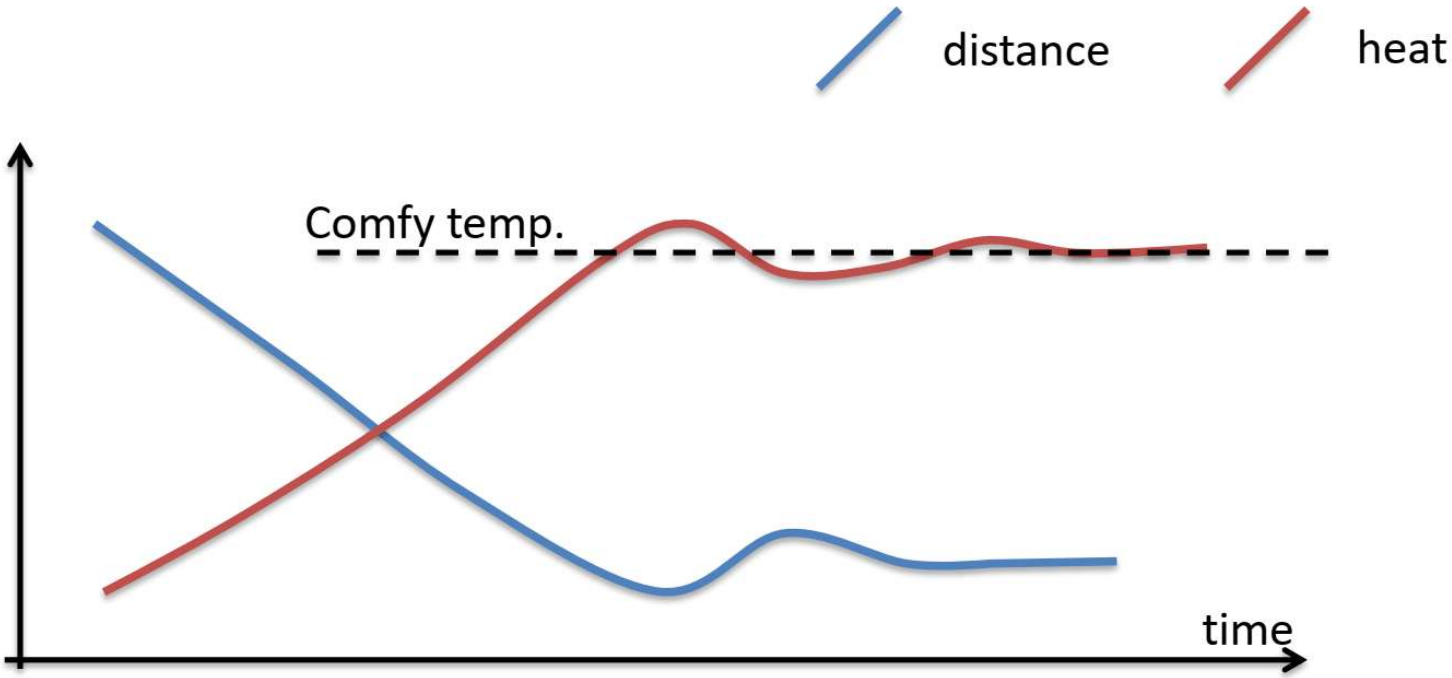
Feedback

- This will be covered in dedicated modules later. To make it simple:
- Feedback is a process in which information about the past and/or current values of the output of a system influences the current or future values of that output.



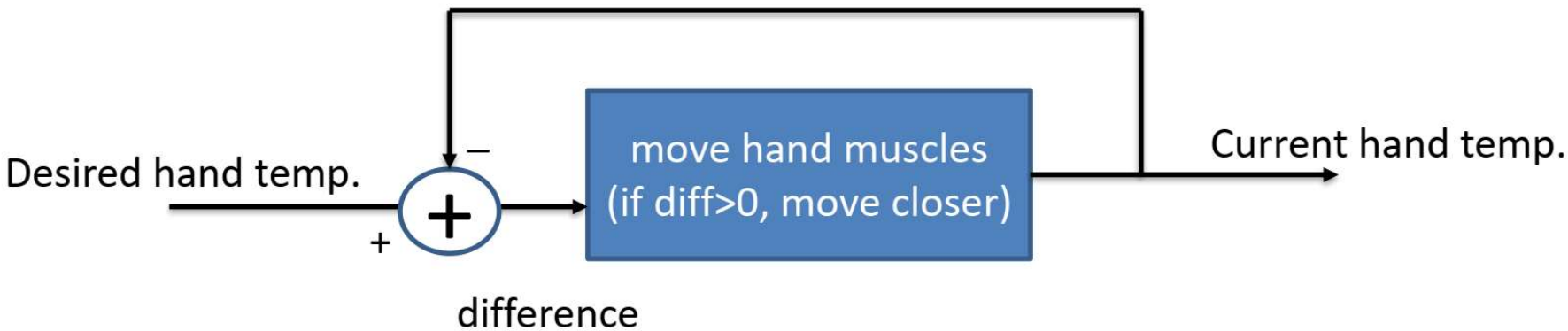
- Consider you have a hand near a fire. Your hand is cold and you want to warm it up to a comfortable temperature. You decide to move your hand closer to the fire (variable is distance). How do you choose when to stop?

Feedback



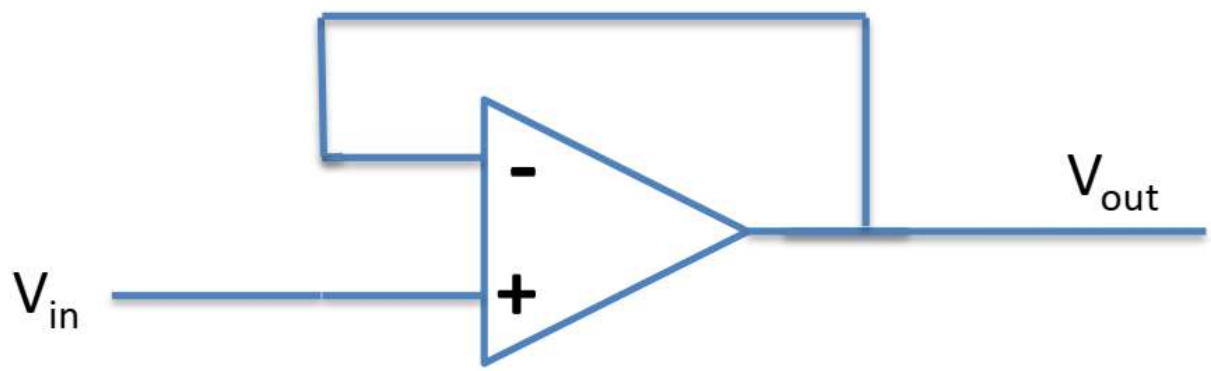
- You move your hand forward and it gets warm;
- You keep moving it. It gets warmer;
- If you move too far and it gets too warm, you pull your hand back;
- Eventually you'll move back and forward a little until it is perfect.

Feedback



- Where the feedback signal works to limit the increase in the output, we call this negative feedback.
- The other type of feedback is called positive feedback and instead of the difference we get the sum. What effect would this have in our example?

In an Op-Amp



Let's explore this scenario:

$$V_{\mathrm{OUT}} = A(V_{+} - V_{-}) = A(V_{\mathrm{IN}} - V_{\mathrm{OUT}})$$

We have

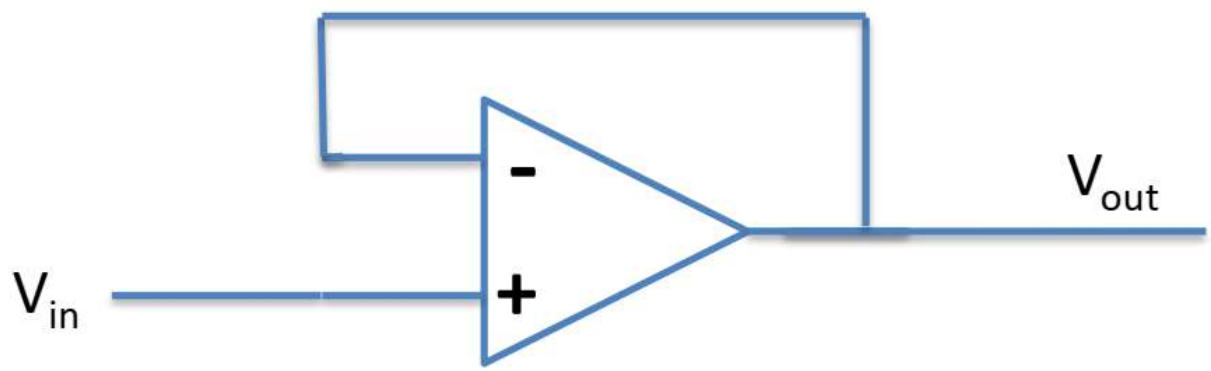
$$V_{\mathrm{OUT}} = \frac{A}{1+A} V_{\mathrm{IN}}$$

If A is very large (100k), then pretty much this says $V_{\mathrm{OUT}} = V_{\mathrm{IN}}$.

- If V_{OUT} increases above V_{IN}, then the difference is negative and V_{OUT} falls.
- As V_{OUT} falls, it drops below V_{IN}, and V_{OUT} increases. The balance point is when $V_{\mathrm{OUT}} = V_{\mathrm{IN}}$.

This circuit is called a voltage follower.

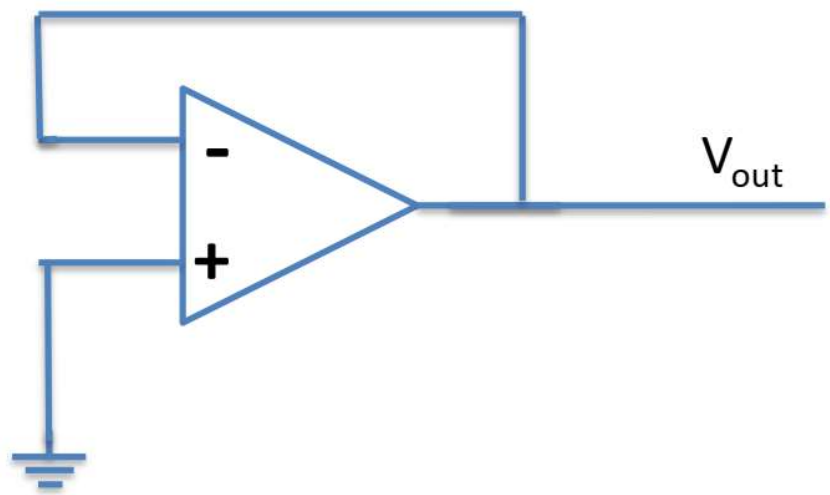
Using Feedback to Simplify Designs



- In any scenario where there is negative feedback and the output voltage is not maxed out, we can assume that the difference in the two inputs (PLUS and MINUS) is very very small.

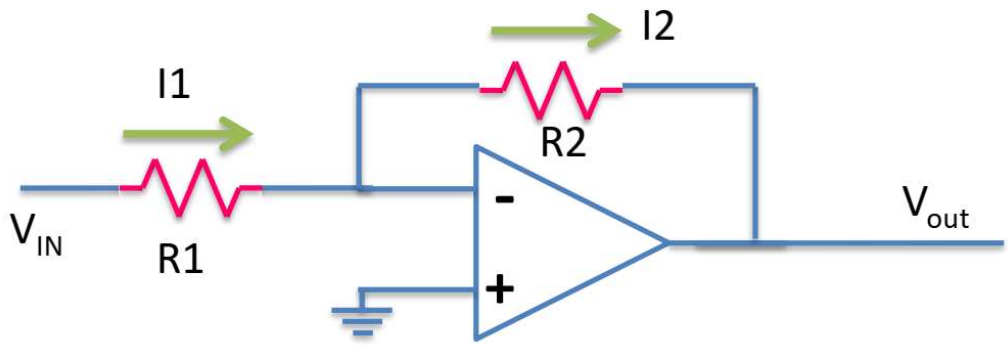
Assumption: The PLUS and MINUS inputs have the same voltage when there is a negative feedback loop.

Virtual Earth (Ground)



- If the assumption is true that PLUS and MINUS have the same voltage, then if I connect one to ground, then the other must also have the same voltage – i.e., zero. We call the MINUS pin in this scenario a virtual earth (ground).
- To get a negative feedback to work we commonly use the MINUS pin of the op-amp, thus we must then connect the earth to the PLUS pin.

An Amplifier



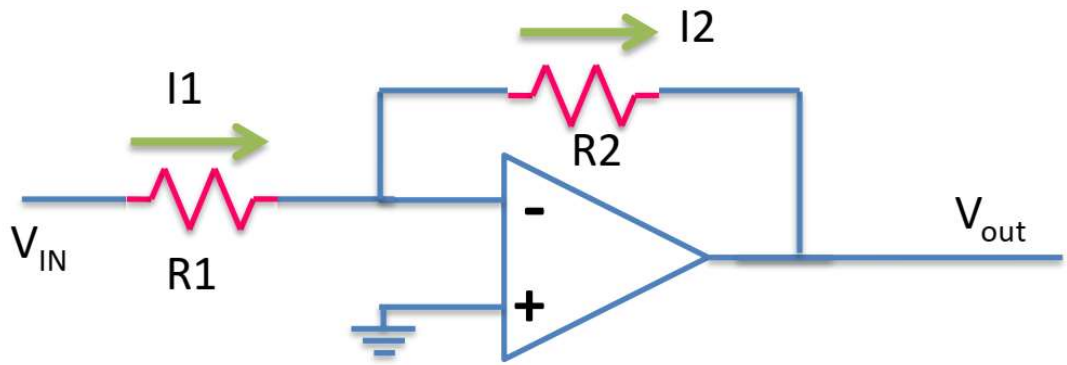
- In this circuit, the PLUS pin is connected to EARTH.
- There is a feedback path, so that means the MINUS pin has ZERO voltage.
- V_{IN} is not zero, so through R_1 flows current I_1, $I_1 = V_{\mathrm{IN}}/R_1$.
- But no current can flow into an ideal op-amp (infinite input impedance), therefore the same current must flow through R_2, $I_2 = I_1$.
- A current through a resistor drops voltage I_2R_2, so

$$V_{\mathrm{OUT}} = V_{-} - I_2R_2 = V_{-} - I_1R_2 = -I_1R_2$$

- Insert the expression for I_1:

$$V_{\mathrm{OUT}} = -\left(\frac{V_{\mathrm{IN}}}{R_1}\right)R_2 = -V_{\mathrm{IN}}\left(\frac{R_2}{R_1}\right)$$

An Amplifier



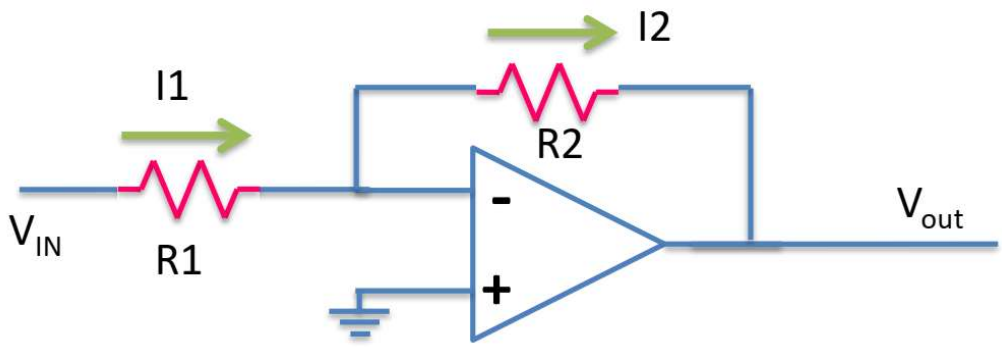
$$V_{\mathrm{OUT}} = -\left(\frac{R_2}{R_1}\right)V_{\mathrm{IN}}$$

Therefore, gain is

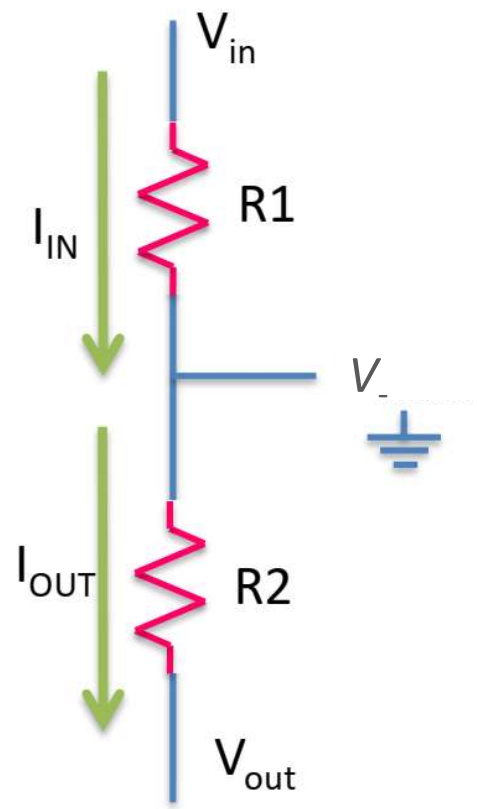
$$A = -\frac{R_2}{R_1}$$

- We call this the closed-loop gain of the system – closed because there is feedback. The gain without feedback is called open-loop gain.
- Also remember we have some assumptions here, e.g., $V_+ = V_-$, virtual earth, ideal op-amp, etc.

An Amplifiers: In Plain English

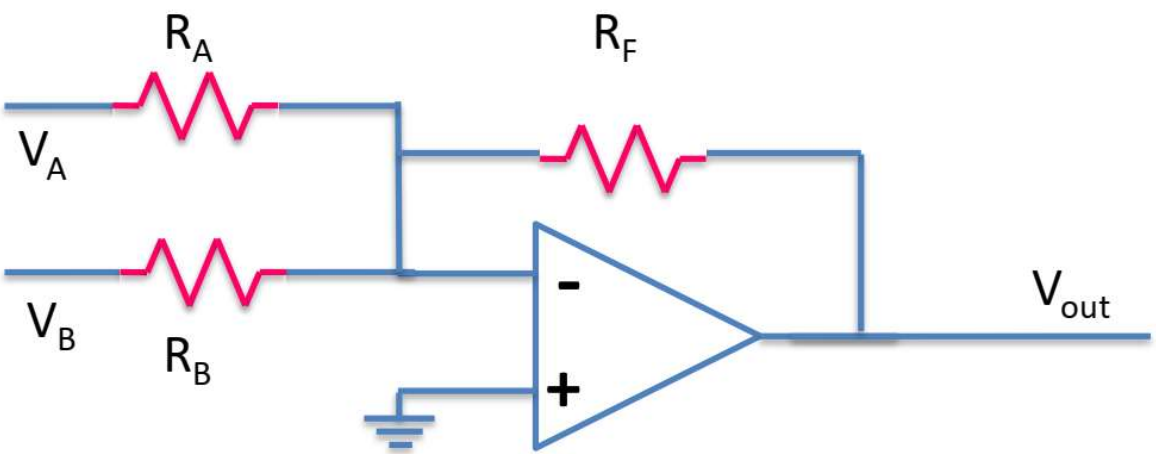


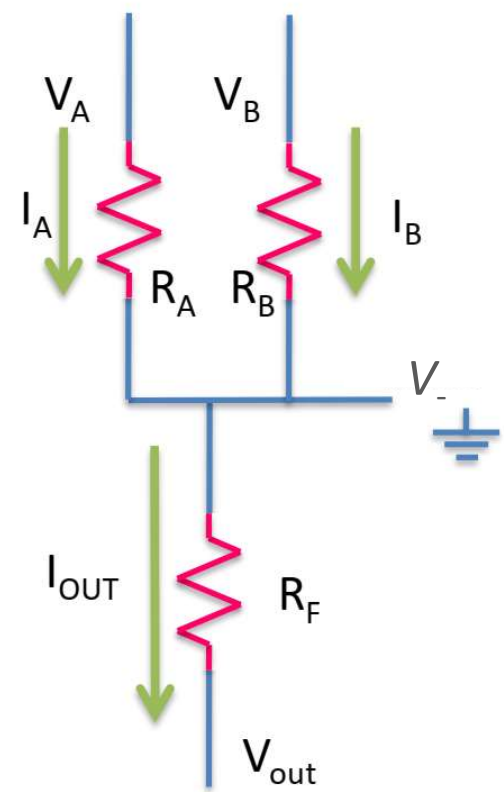
- If V_{OUT} gets too negative, V_- will be negative. Then $A(V_+ - V_-)$ becomes a positive number and V_{OUT} will try to rise, and we get this bigger/smaller oscillation until they balance ($V_+ = V_- = 0$).
- So as the PLUS pin is at zero volts, the feedback will always work to force the MINUS pin to zero volts. It will do this by changing the value of V_{OUT} to whatever value is needed to achieve this.
- Consider them vertically stacked instead. The middle point is zero, but no current can escape. In this scenario, the current is V_{IN}/R_1, but it must then flow through the second resistor. V_{OUT} will keep changing until it is the right value to get the voltage divider right.



$$I_{\mathrm{IN}} = I_{\mathrm{OUT}} \text{ and } \frac{V_{\mathrm{IN}} - 0}{R_1} = \frac{0 - V_{\mathrm{OUT}}}{R_2}$$

Another Amplifier: Two Inputs





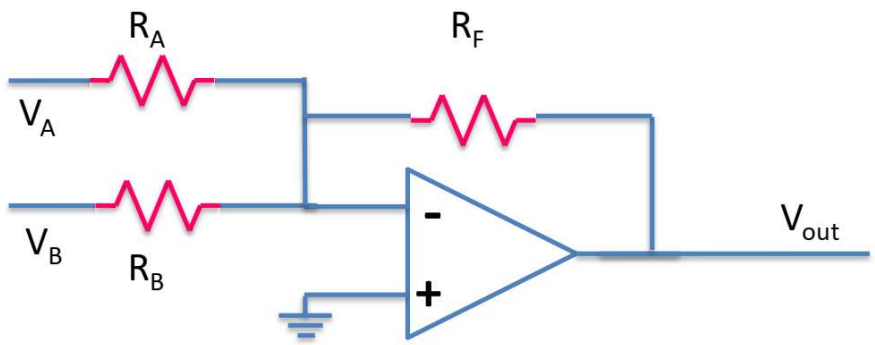
- What we have here are two inputs, but the theory is the same. V_- must stay at zero, and V_{OUT} moves.
- It is harder to do a divider for voltage here, so let's do nodal analysis with currents. What goes in must come out:

$$I_A = V_A/R_A \text{ and } I_B = V_B/R_B$$

$$I_{OUT} = I_A + I_B$$

$$V_{OUT} = -I_{OUT}R_F$$

Another Amplifier: Two Inputs



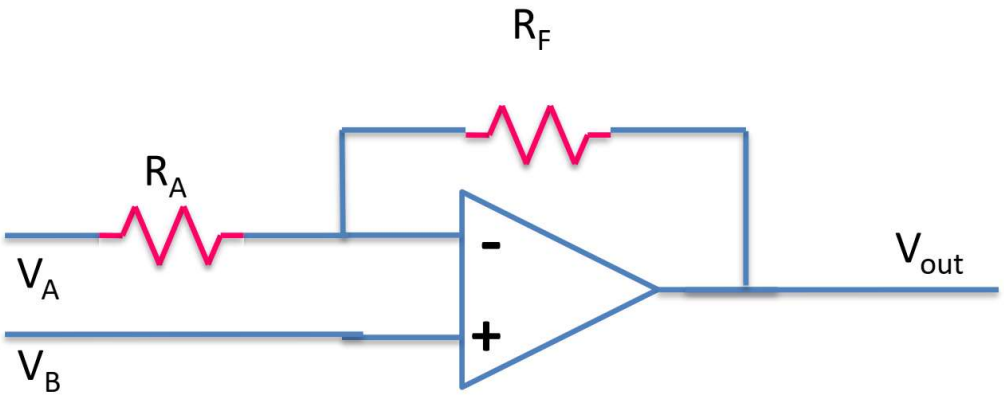
$$V_{\mathrm{OUT}} = -R_{\mathrm{F}}\left(\frac{V_{\mathrm{A}}}{R_{\mathrm{A}}} + \frac{V_{\mathrm{B}}}{R_{\mathrm{B}}}\right) = -[V_{\mathrm{A}}\left(\frac{R_{\mathrm{F}}}{R_{\mathrm{A}}}\right) + V_{\mathrm{B}}\left(\frac{R_{\mathrm{F}}}{R_{\mathrm{B}}}\right)]$$

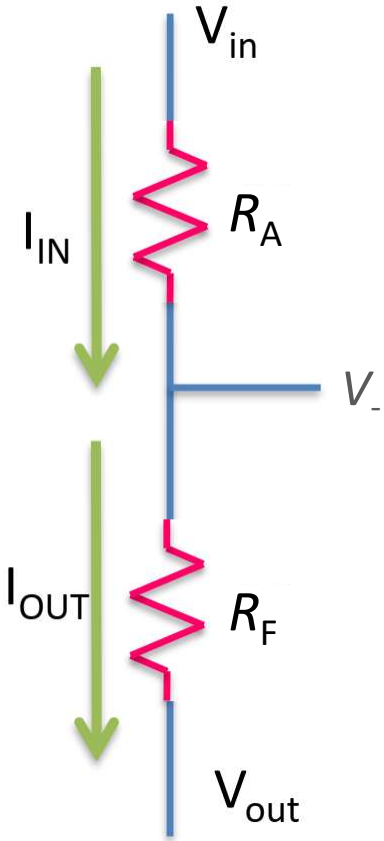
- Let us imagine that all the resistors are equal, then what we get is:

$$V_{\mathrm{OUT}} = -(V_{\mathrm{A}} + V_{\mathrm{B}})$$

- This gives us the function of summation, and the design is called an inverting summer (summing amplifier). You can use multiple branches to add more voltages together.

Another Amplifier: Two Inputs





- In this case, we do not have V_+ set to the ground. What will happen here?
- Same analysis: But this time V_- will try to match V_+ which is equal to V_B.

$$I_{\mathrm{IN}} = \frac{V_{\mathrm{A}} - V_{\mathrm{B}}}{R_{\mathrm{A}}} \text{ and } I_{\mathrm{OUT}} = \frac{V_{\mathrm{B}} - V_{\mathrm{OUT}}}{R_{\mathrm{F}}}$$

$$I_{\mathrm{IN}} = I_{\mathrm{OUT}}$$

Another Amplifier: Two Inputs

$$I_{\mathrm{IN}} = \frac{V_{\mathrm{A}} - V_{\mathrm{B}}}{R_{\mathrm{A}}} \text{ and } I_{\mathrm{OUT}} = \frac{V_{\mathrm{B}} - V_{\mathrm{OUT}}}{R_{\mathrm{F}}}$$

$$I_{\mathrm{IN}} = I_{\mathrm{OUT}}$$

$$\frac{V_{\mathrm{A}} - V_{\mathrm{B}}}{R_{\mathrm{A}}} = \frac{V_{\mathrm{B}} - V_{\mathrm{OUT}}}{R_{\mathrm{F}}}$$

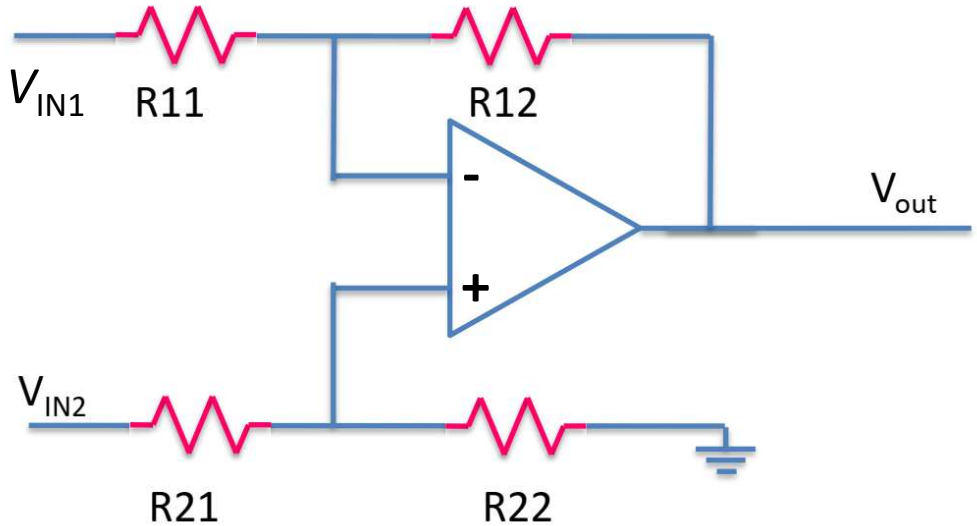
$$V_{\mathrm{OUT}} = -V_{\mathrm{A}}\left(\frac{R_{\mathrm{F}}}{R_{\mathrm{A}}}\right) + V_{\mathrm{B}}\left(1 + \frac{R_{\mathrm{F}}}{R_{\mathrm{A}}}\right)$$

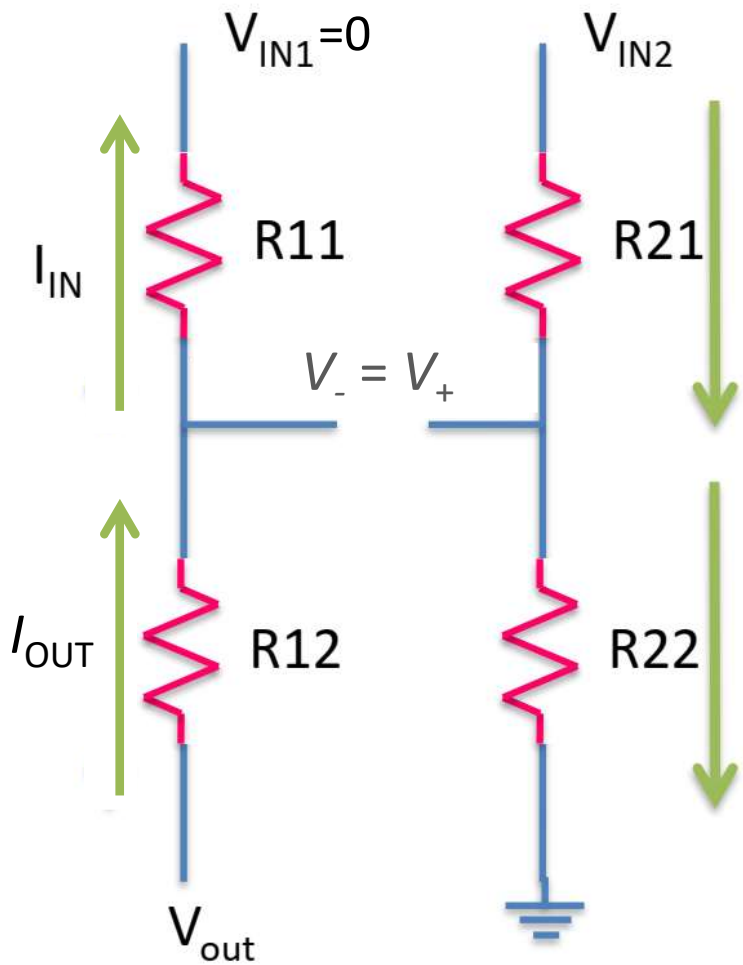
$$= (V_{\mathrm{B}} - V_{\mathrm{A}})\left(\frac{R_{\mathrm{F}}}{R_{\mathrm{A}}}\right) + V_{\mathrm{B}}$$

$$V_{\mathrm{OUT}} \approx (V_{\mathrm{B}} - V_{\mathrm{A}})\left(\frac{R_{\mathrm{F}}}{R_{\mathrm{A}}}\right)$$

- This makes a subtraction function which is accurate if R_F/R_A is very big. We should make R_F big and R_A small.

Another Configuration





- This circuit is a bit tricky. We will use the superposition theorem to solve it.
- If $V_{IN2} = 0$, $V_{OUT} = -\left(\frac{R_{12}}{R_{11}}\right) V_{IN}$.
- If $V_{IN1} = 0$, we can write down two equations:

$$V_+ = V_{IN2}\frac{R_{22}}{R_{21}+R_{22}} \text{ and } V_- = V_{OUT}\left(\frac{R_{11}}{R_{11}+R_{12}}\right)$$

Another Configuration

$$V_+ = V_-$$

$$V_{\mathrm{IN2}} \frac{R_{22}}{R_{21} + R_{22}} = V_{\mathrm{OUT}} \frac{R_{11}}{R_{11} + R_{12}}$$

$$V_{\mathrm{OUT}} = \left(\frac{R_{22}}{R_{11}}\right)\left(\frac{R_{11}+R_{12}}{R_{21}+R_{22}}\right)V_{\mathrm{IN}} \quad \text{(when } V_{\mathrm{IN1}} = 0\text{)}$$

- When V_{IN1} and V_{IN2} are both present, the V_{OUT} would be the sum of the two cases:

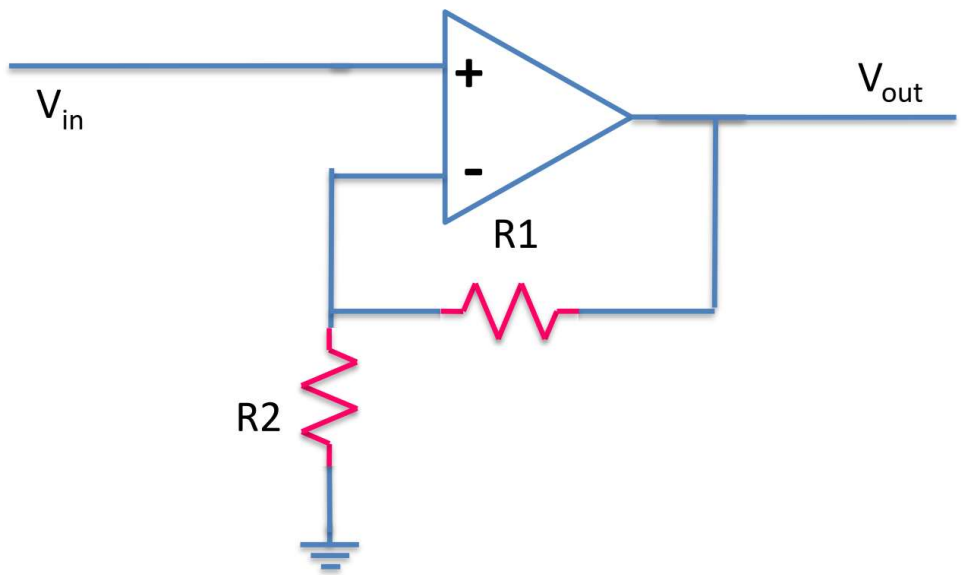
$$V_{\mathrm{OUT}} = -\left(\frac{R_{12}}{R_{11}}\right)V_{\mathrm{IN1}} + \left(\frac{R_{22}}{R_{11}}\right)\left(\frac{R_{11} + R_{12}}{R_{21} + R_{22}}\right)V_{\mathrm{IN2}}$$

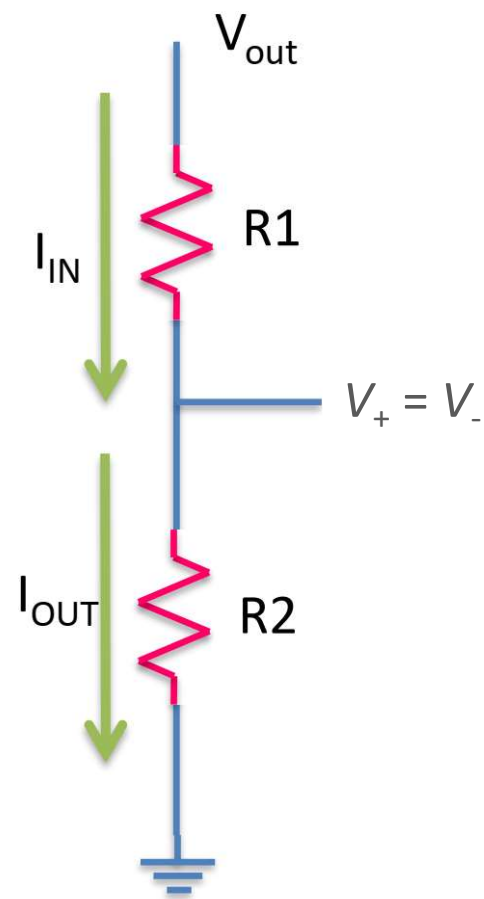
- If we assume that $R_{11} = R_{21}$ and $R_{12} = R_{22}$,

$$V_{\mathrm{OUT}} = \left(\frac{R_{12}}{R_{11}}\right)(V_{\mathrm{IN2}} - V_{\mathrm{IN1}})$$

- We amplify the difference, and the exact ratio and scaling depend on the resistors.

A Different Configuration





- This is a different configuration for the resistors but the principle remains the same.
- V_+ and V_- need to stay very close to each other if there is working negative feedback.

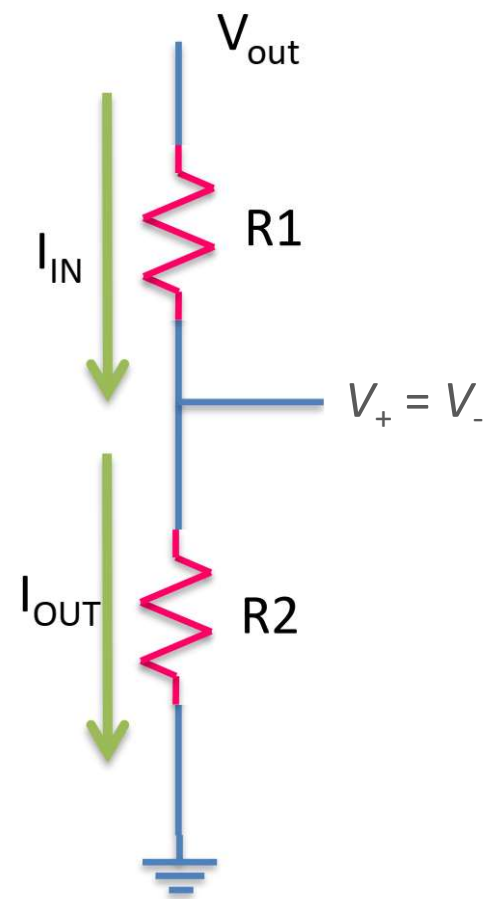
Non-Inverting Amplifier

$$I_{\mathrm{IN}} = \frac{V_{\mathrm{OUT}} - V_{-}}{R_1} \text{ and } I_{\mathrm{OUT}} = \frac{V_{-} - 0}{R_2}$$

$$I_{\mathrm{IN}} = I_{\mathrm{OUT}}$$

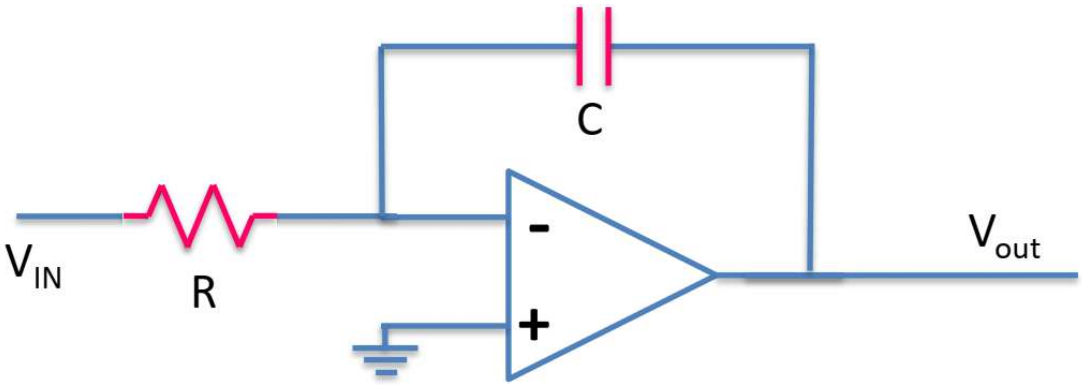
$$\frac{V_{\mathrm{OUT}} - V_{-}}{R_1} = \frac{V_{-}}{R_2} \text{ and } V_{-} = V_{+} = V_{\mathrm{IN}}$$

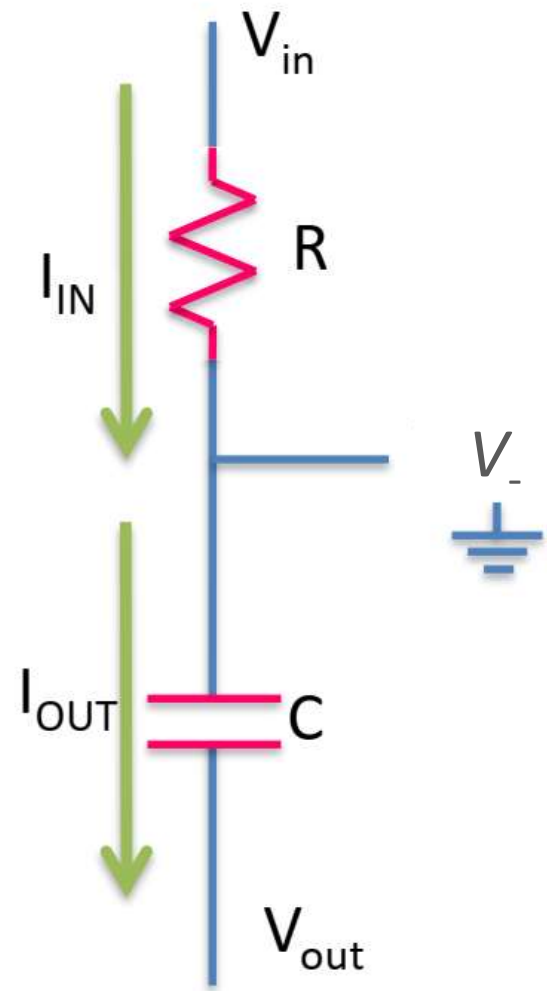
$$V_{\mathrm{OUT}} = \left(1 + \frac{R_1}{R_2}\right) V_{\mathrm{IN}}$$



- The important thing to note here is that there is NO MINUS SIGN.
- This is a non-inverting amplifier and the gain is given as shown. If you want a gain of 1.00, make R_1 = 0 or R_2 = infinity.

An Amplifier: With a Capacitor



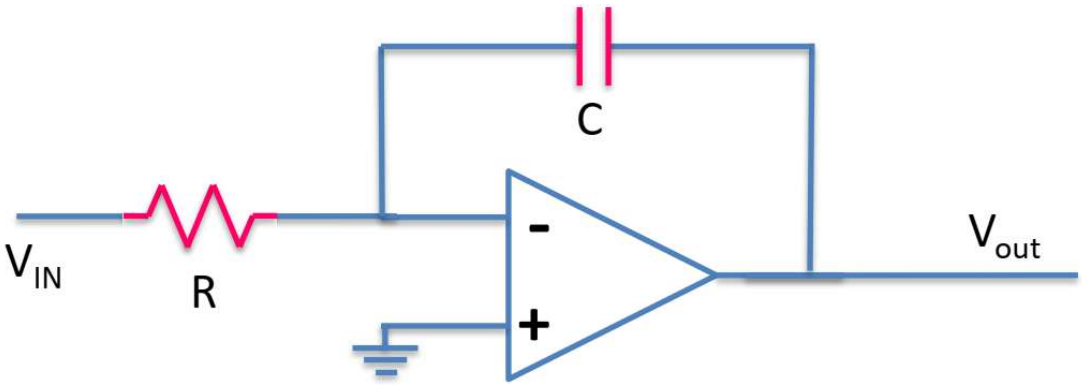


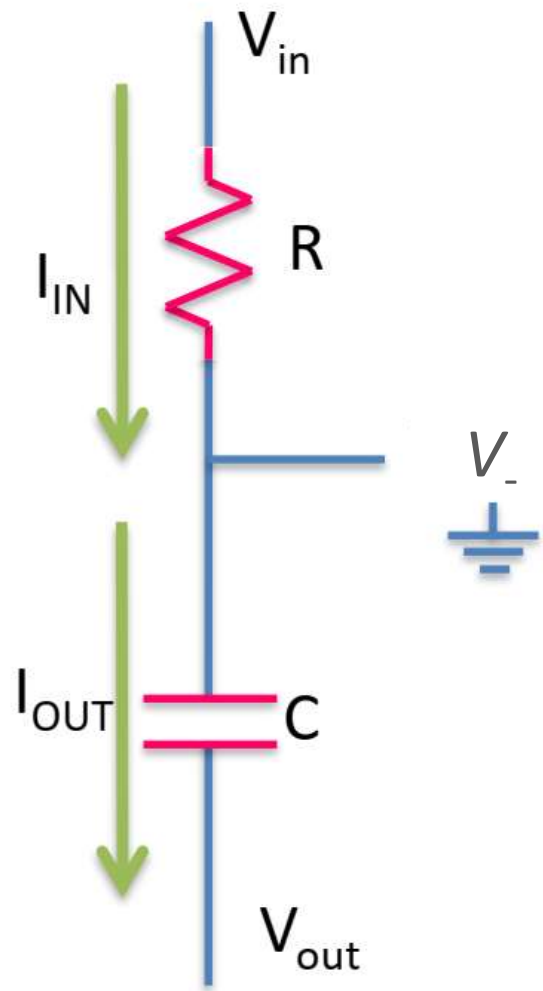
- There are two ways to look at this circuit. The first is to treat it as a normal amplifier.

$$A = -\frac{Z_C}{R} = -\frac{1}{R}\frac{1}{j\omega C} = -\frac{1}{j\omega RC}$$

- We have a frequency dependent gain. When frequency is zero, the gain is infinite. When frequency is high, the gain is low.

An Amplifier: With a Capacitor





- A second way is to look at the current:

$$I_{\mathrm{IN}} = \frac{V_{\mathrm{IN}} - V_{-}}{R} \text{ and } I_{\mathrm{OUT}} = C\frac{d(V_{-} - V_{\mathrm{OUT}})}{dt}$$

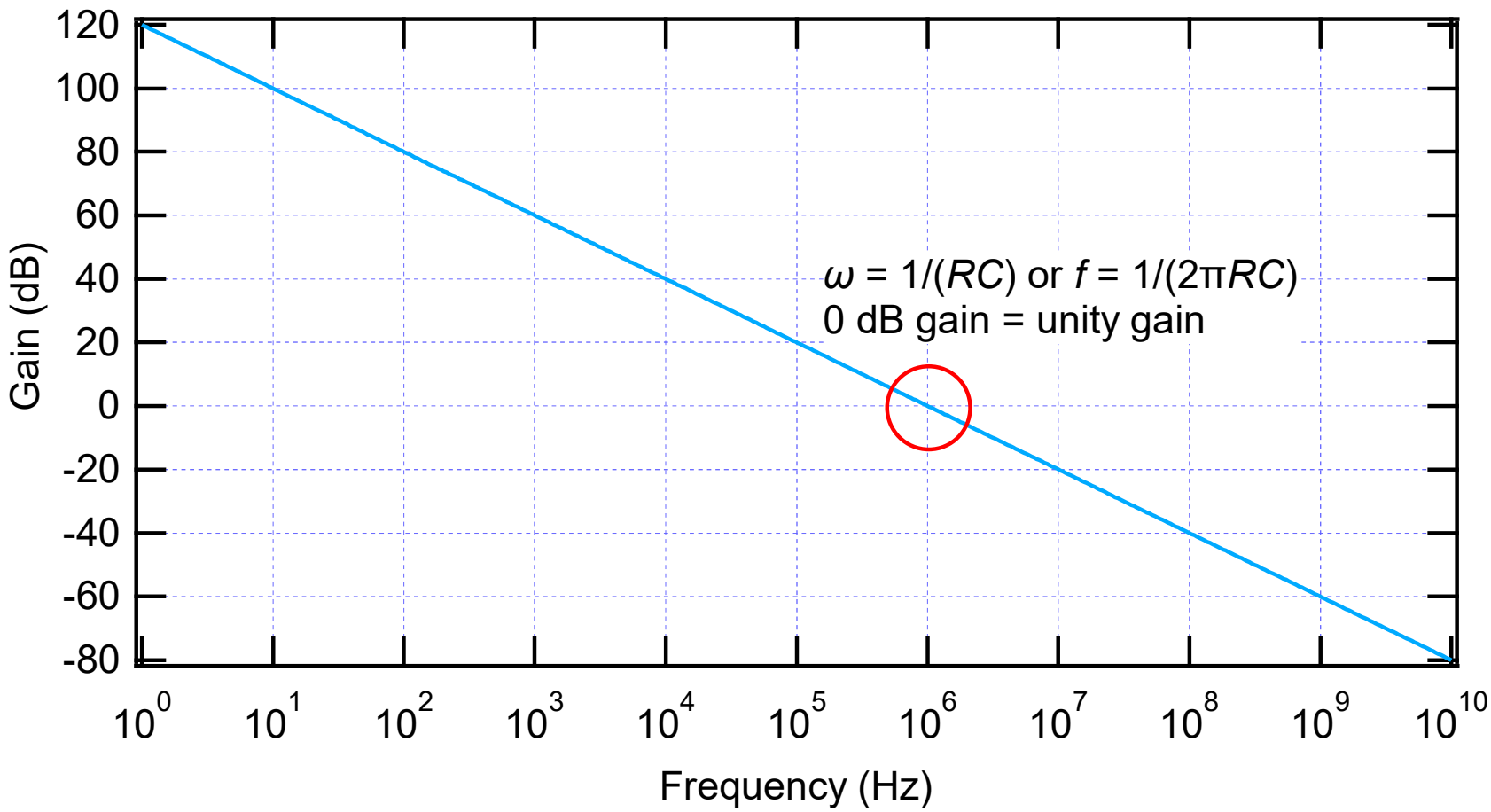
$$\frac{V_{\mathrm{IN}}}{R} = -C\frac{dV_{\mathrm{OUT}}}{dt}$$

$$V_{\mathrm{OUT}} = -\left(\frac{1}{RC}\right)\int V_{\mathrm{IN}}\,dt$$

- V_{OUT} is the INTEGRAL of V_{IN}, this gives us an INTEGRATOR function.

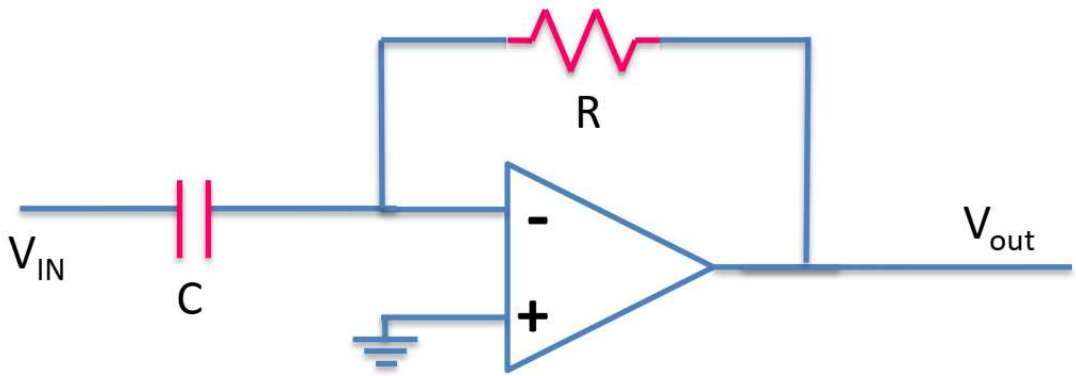
An Integrator

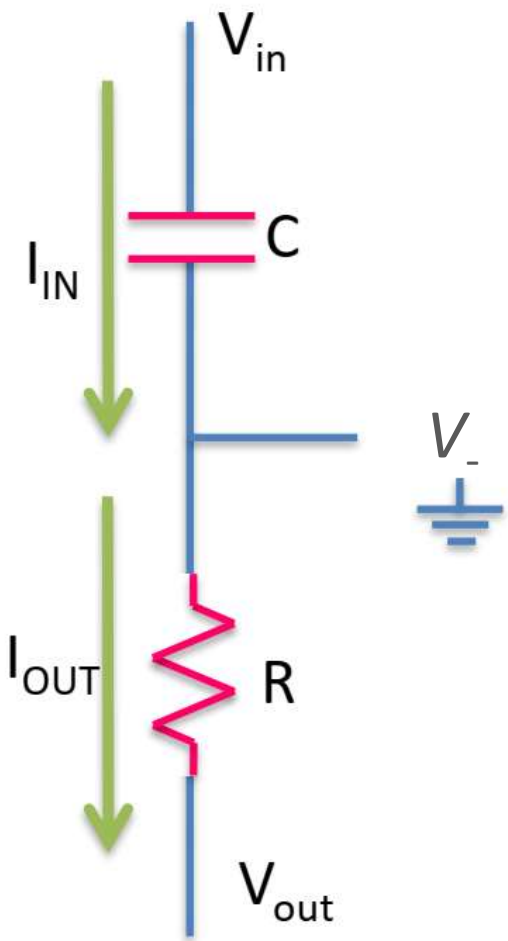
$$|A| = 1/(\omega RC)$$



Attention! The plot is in the log-log scale.

Another Amplifier: With a Capacitor





- Swap the capacitor and the resistor. Let's redo the calculation:

$$I_{\mathrm{IN}} = C\frac{d(V_{\mathrm{IN}} - V_{-})}{dt} \text{ and } I_{\mathrm{OUT}} = \frac{V_{-} - V_{\mathrm{OUT}}}{R}$$

$$C\frac{dV_{\mathrm{IN}}}{dt} = \frac{-V_{\mathrm{OUT}}}{R}$$

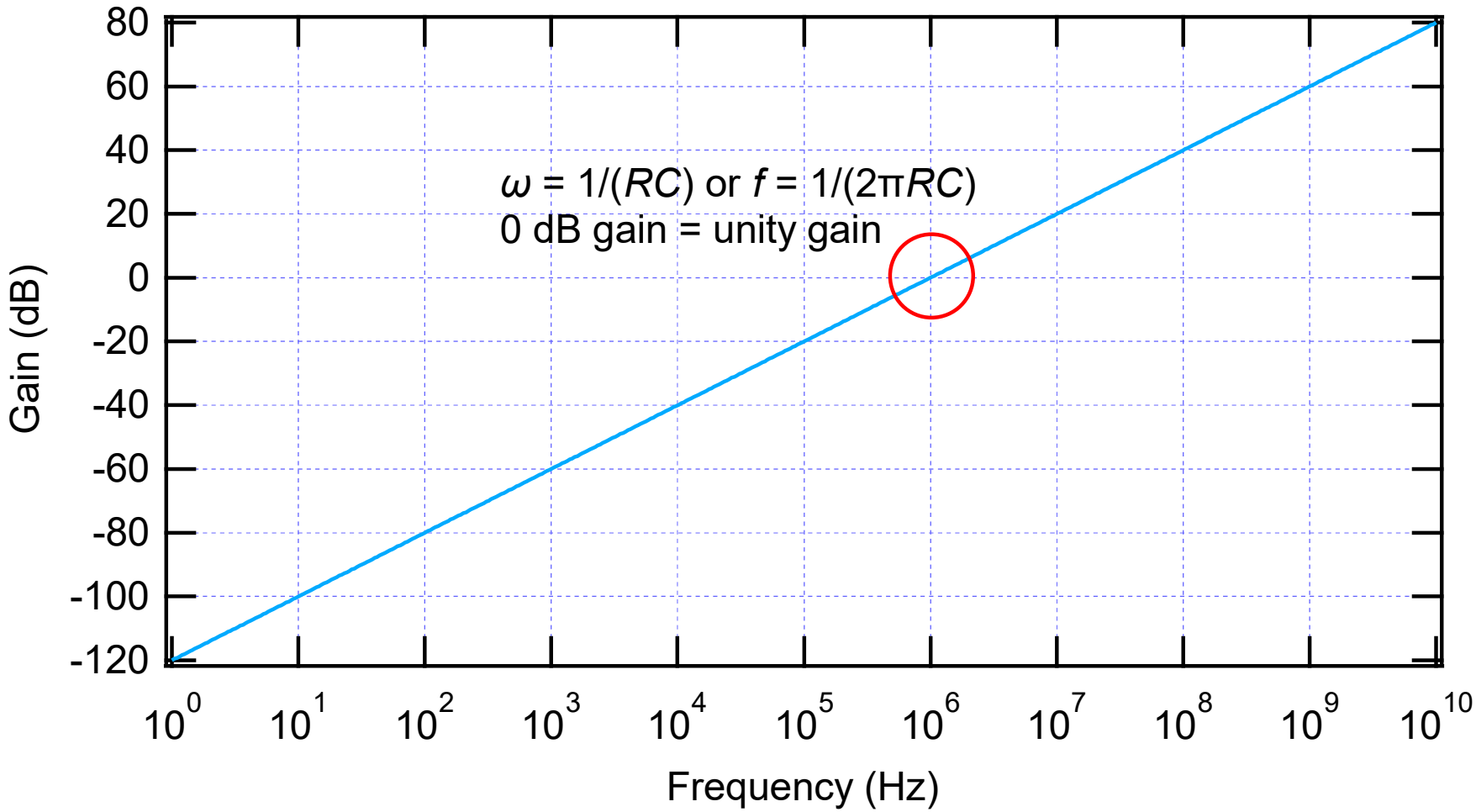
$$V_{\mathrm{OUT}} = -(RC)\frac{dV_{\mathrm{IN}}}{dt}$$

- V_{OUT} is the DIFFERENTIAL of V_{IN}, this gives us a DIFFERENTIATOR function.

$$A = -j\omega RC$$

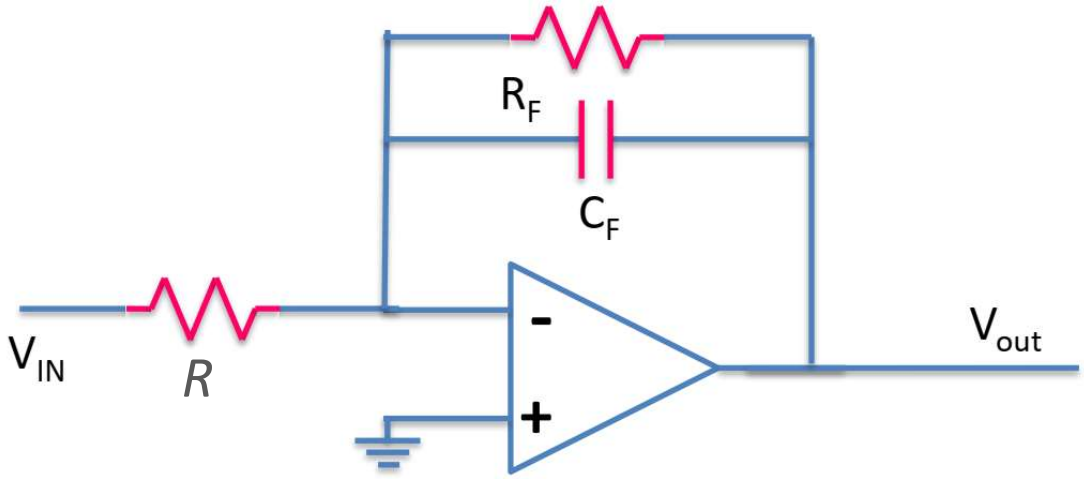
A Differentiator

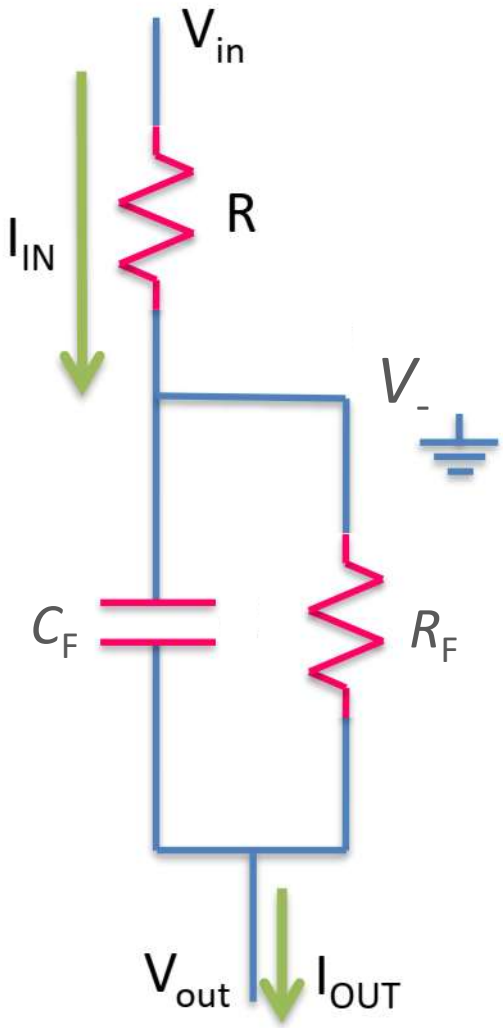
$$|A| = \omega RC$$



Attention! The plot is in the log-log scale.

More Complex Structures



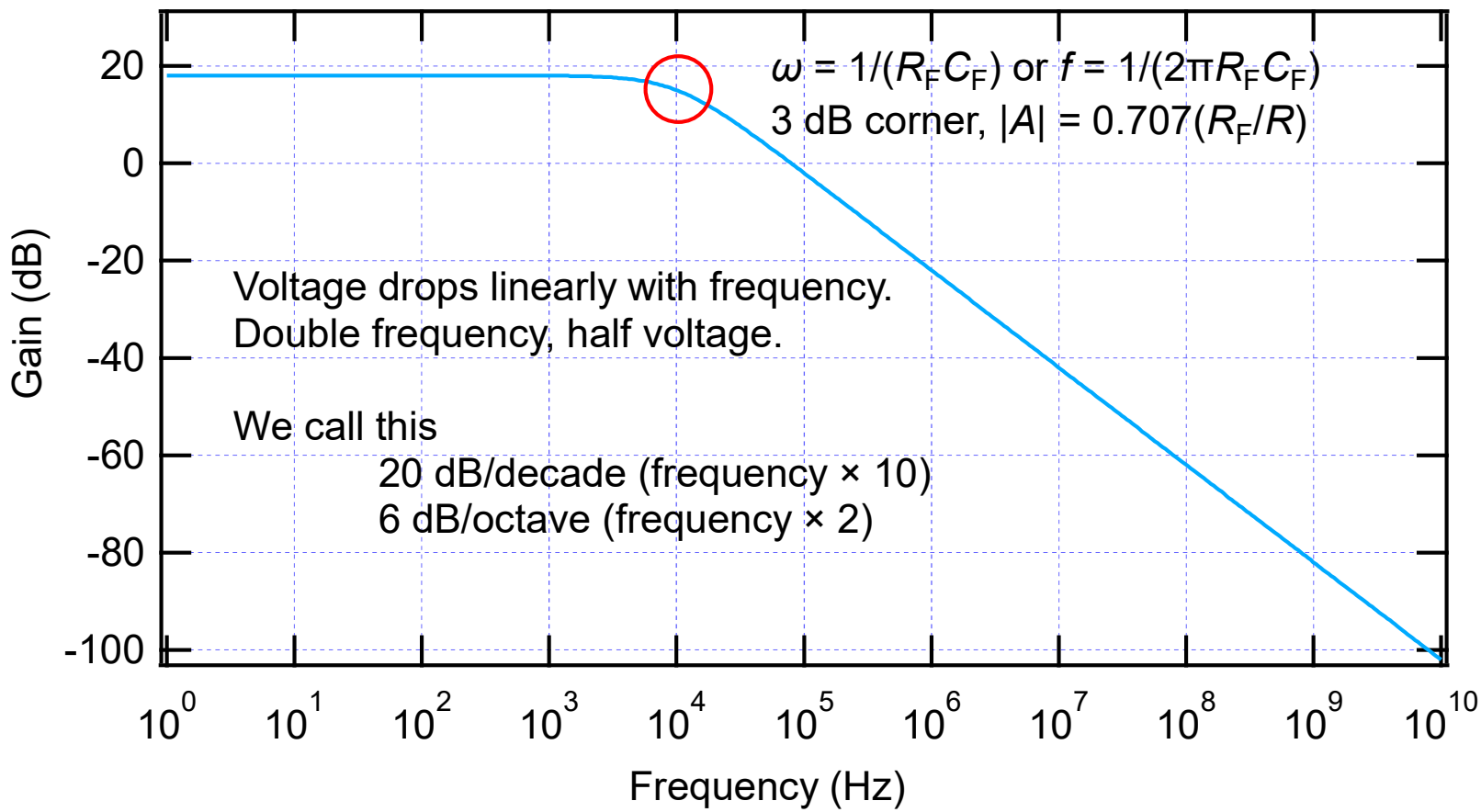


- Let's look at the gain:

$$A = -\frac{Z_F \parallel R_F}{R} = -\frac{1}{R}\left(\frac{1}{\frac{1}{R_F} + j\omega C_F}\right)$$

$$A = -\left(\frac{R_F}{R}\right)\left(\frac{1}{1 + j\omega R_F C_F}\right)$$

Low-Pass Filter With Corner at $1/(2\pi RC)$



- We call filters with a single "corner" a single-pole filter. You can get multiple pole filters (higher-order filters). You get a pole every time you hit a frequency point that halves the power. Normally it's one pole per capacitor.

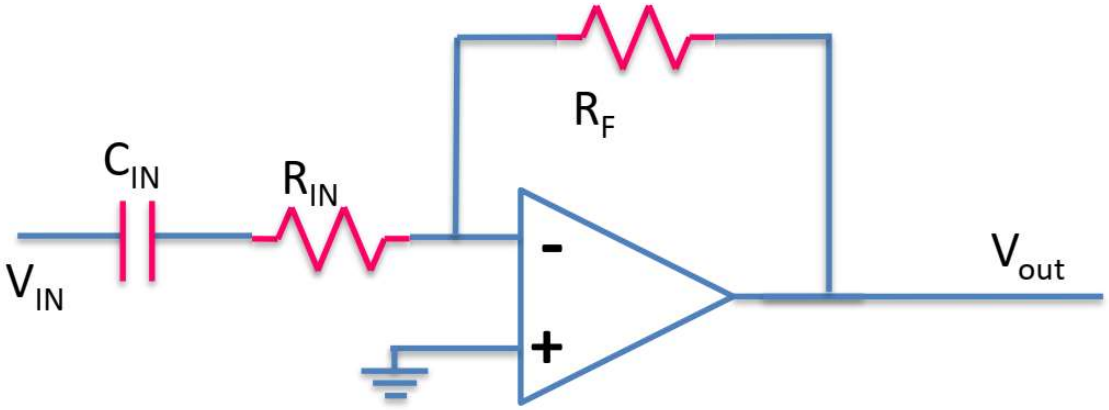
Other Characteristics: Unity Gain Bandwidth

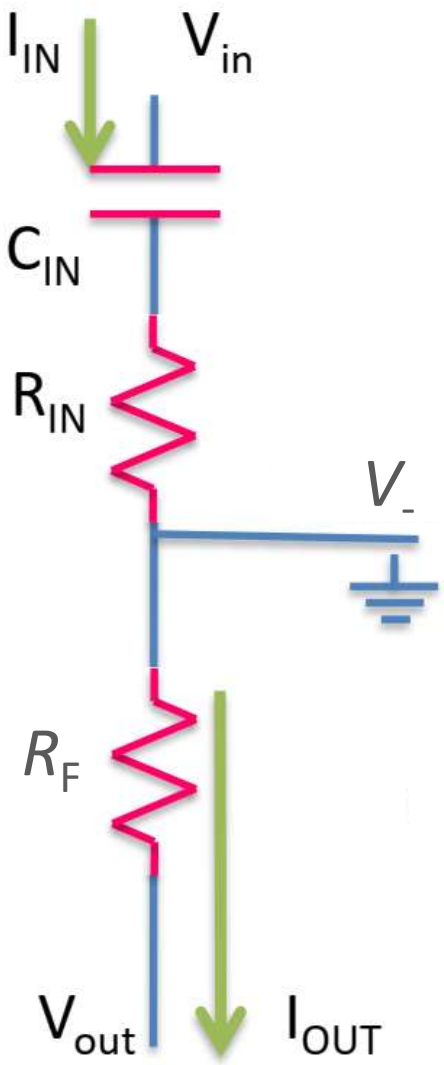
- One characteristic often quoted for op-amps is unity gain bandwidth.
- This works on the assumption that even with the best of intentions, there is parasitic capacitances in the circuit, when nothing else arises from the transistors. This will cause the roll-off in gain that you see.
- Unity gain bandwidth assumes that we have first-order roll-off (a single aggregate capacitor effect) and it is the bandwidth for which gain is > 1. Or more simply, the frequency at which gain = 1.
- MORE IMPORTANTLY. If we have a single capacitor effect (first-order filter effect), then the gain will change by a factor of 10 if the frequency changes by a factor of 10. The product (so-called gain-bandwidth product or GBW) is a constant. So

If	$G = 1$	at BW = 100 MHz (GBW = 100M)
then	$G = 10$	at BW = 10 MHz (GBW = 100M)
then	$G = 100$	at BW = 1 MHz (GBW = 100M)

- This is not perfectly true in all cases, but it is a useful guide.

More Complex Structures



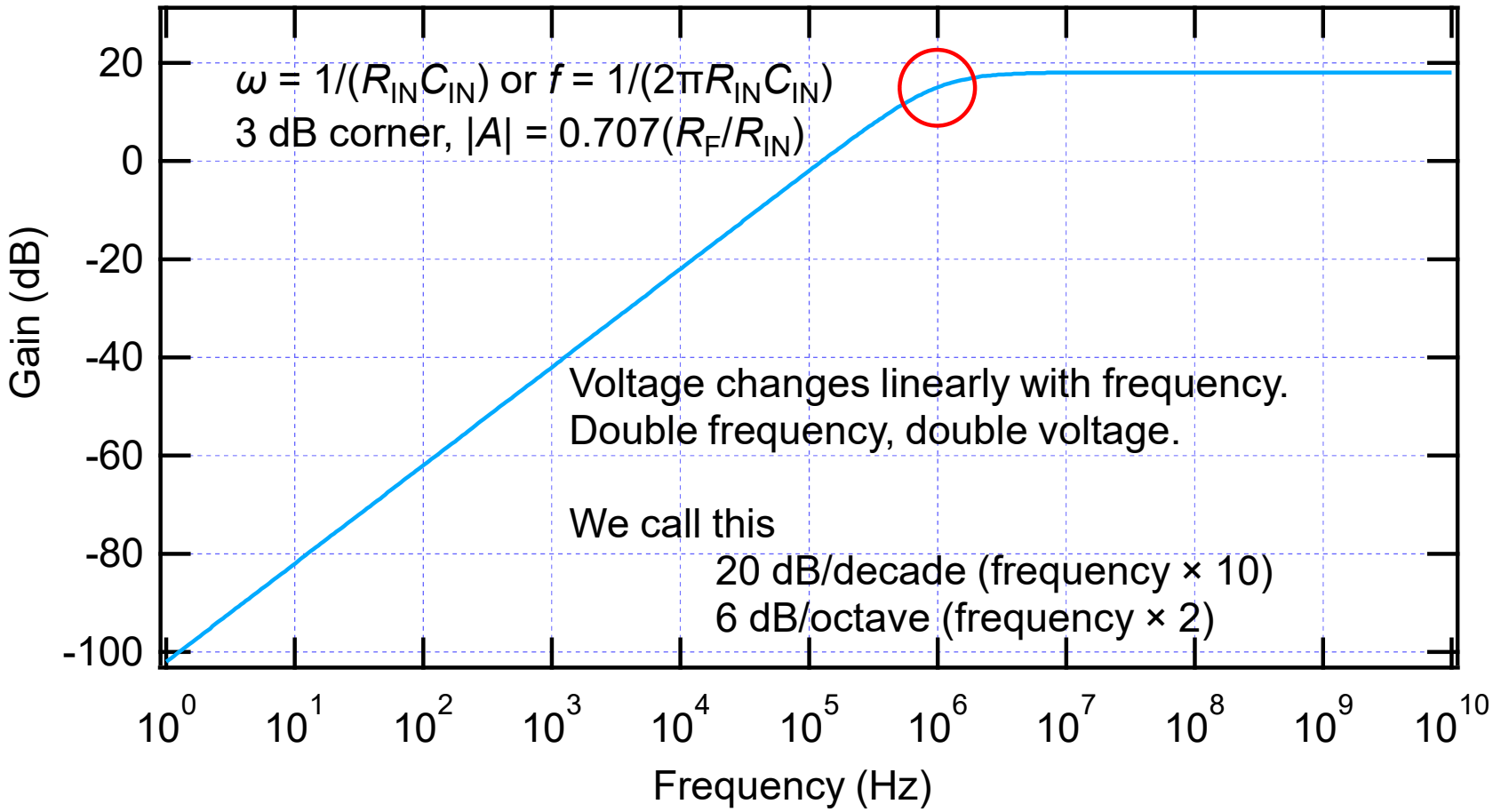


- Let's look at the gain:

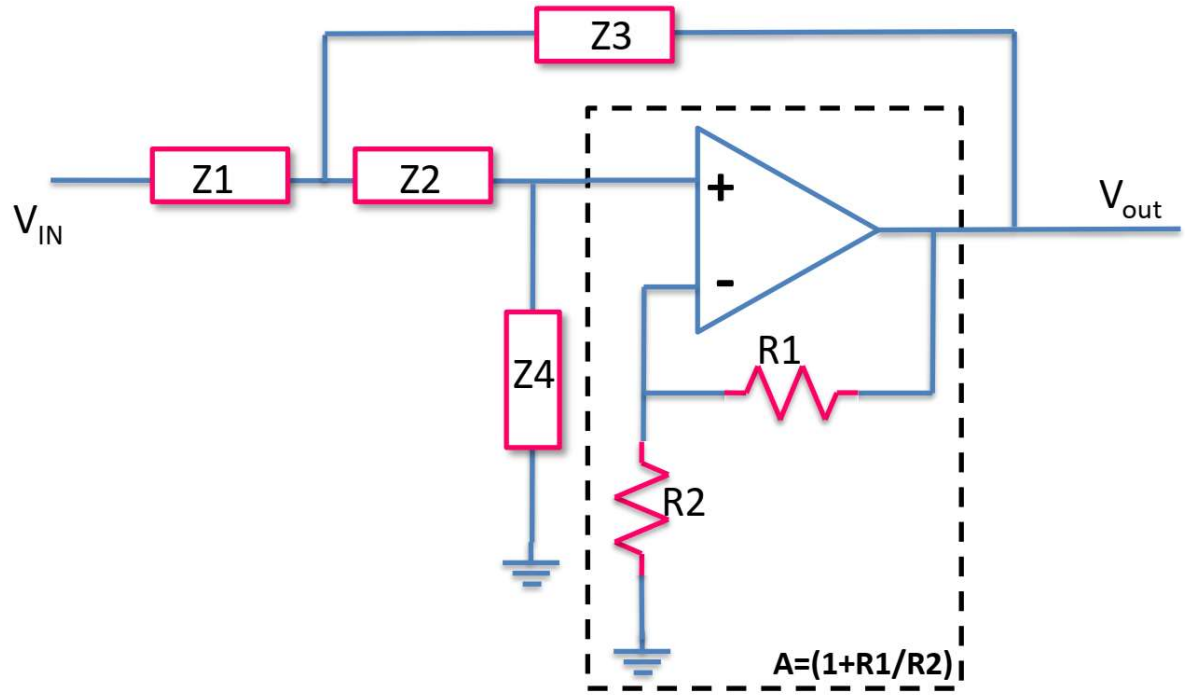
$$A = -\frac{R_F}{R_{IN} + Z_{IN}} = -R_F\left(\frac{1}{R_{IN} + \frac{1}{j\omega C_{IN}}}\right)$$

$$A = -\left(\frac{j\omega R_F C_{IN}}{1 + j\omega R_{IN} C_{IN}}\right)$$

High-Pass Filter With Corner at $1/(2\pi RC)$



The Sallen-Key Filter Structure



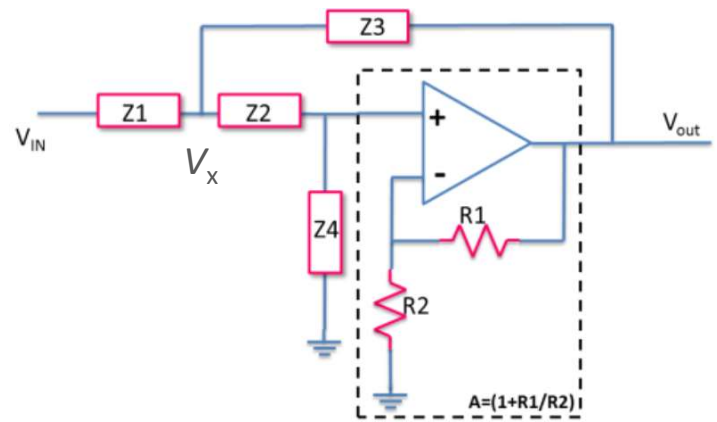
- This structure is a general-purpose structure for constructing filters. It is well understood. We have to look at what the gain equation is. The important first step is that the dashed-box is a standard non-inverting amplifier where we control gain. We're going to set R_1 = 0 and assume unity gain for the moment.
- What we need to figure out is what the value of V_+ is.

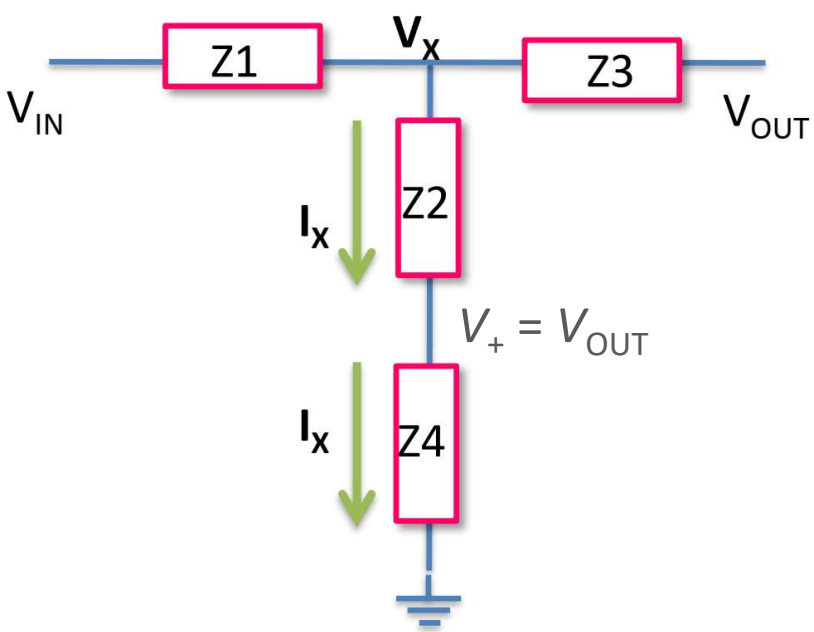
The Sallen-Key Filter Structure

- We're only after V_+. Let's re-draw the circuit and accept the following criteria. There is a non-inverting amplification of A = 1.0 (we can change the value of gain by changing R_1 and R_2 if we want). This means

$$V_{\mathrm{OUT}} = A \cdot V_+ = V_+$$

- We're going to find V_x, get the current I_x, and then work out $V_+ = I_x Z_4$.





Kirchhoff's current law at point x:

$$\frac{V_{\mathrm{x}} - V_+}{Z_2} = \frac{V_{\mathrm{IN}} - V_{\mathrm{x}}}{Z_1} + \frac{V_{\mathrm{OUT}} - V_{\mathrm{x}}}{Z_3}$$

$$\frac{V_{\mathrm{x}} - V_{\mathrm{OUT}}}{Z_2} = \frac{V_{\mathrm{IN}} - V_{\mathrm{x}}}{Z_1} + \frac{V_{\mathrm{OUT}} - V_{\mathrm{x}}}{Z_3}$$

Kirchhoff's current law at V_+:

$$I_{\mathrm{x}} = \frac{V_{\mathrm{x}} - V_{\mathrm{OUT}}}{Z_2} = \frac{V_{\mathrm{OUT}} - 0}{Z_4}$$

$$V_{\mathrm{x}} = (1 + \frac{Z_2}{Z_4})V_{\mathrm{OUT}}$$

The Sallen-Key Filter Structure

- We insert $V_{\mathrm{x}} = (1 + \frac{Z_2}{Z_4})V_{\mathrm{OUT}}$ into equation $\frac{V_{\mathrm{x}} - V_{\mathrm{OUT}}}{Z_2} = \frac{V_{\mathrm{IN}} - V_{\mathrm{x}}}{Z_1} + \frac{V_{\mathrm{OUT}} - V_{\mathrm{x}}}{Z_3}$:

$$\frac{(1 + \frac{Z_2}{Z_4})V_{\mathrm{OUT}} - V_{\mathrm{OUT}}}{Z_2} = \frac{V_{\mathrm{IN}} - (1 + \frac{Z_2}{Z_4})V_{\mathrm{OUT}}}{Z_1} + \frac{V_{\mathrm{OUT}} - (1 + \frac{Z_2}{Z_4})V_{\mathrm{OUT}}}{Z_3}$$

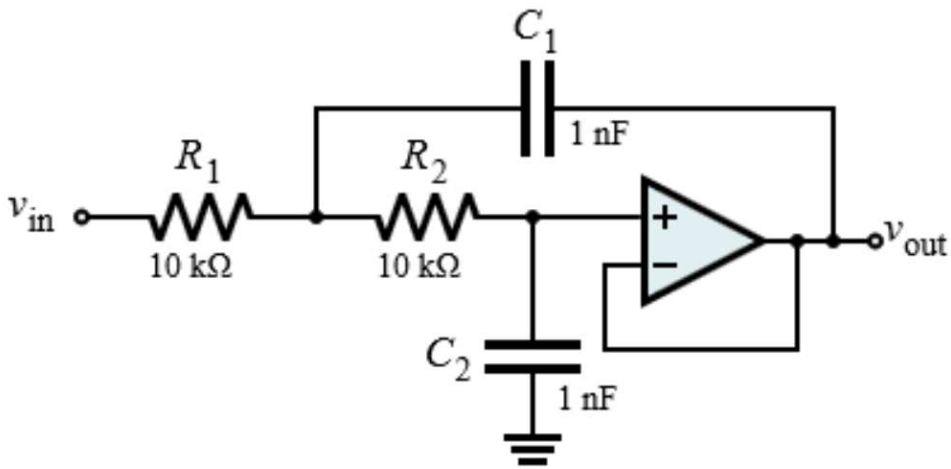
- This equation now contains only V_{IN} and V_{OUT}. Solving this equation gives the "gain" of the circuit:

$$\frac{V_{\mathrm{OUT}}}{V_{\mathrm{IN}}} = \frac{Z_3 Z_4}{Z_1 Z_2 + Z_1 Z_3 + Z_2 Z_3 + Z_3 Z_4}$$

- If all Z's are complex, then this is a second-order filter (In fact, this requirement is usually too strong. We only need either Z_1 and Z_2 or Z_3 and Z_4 to be imaginary).

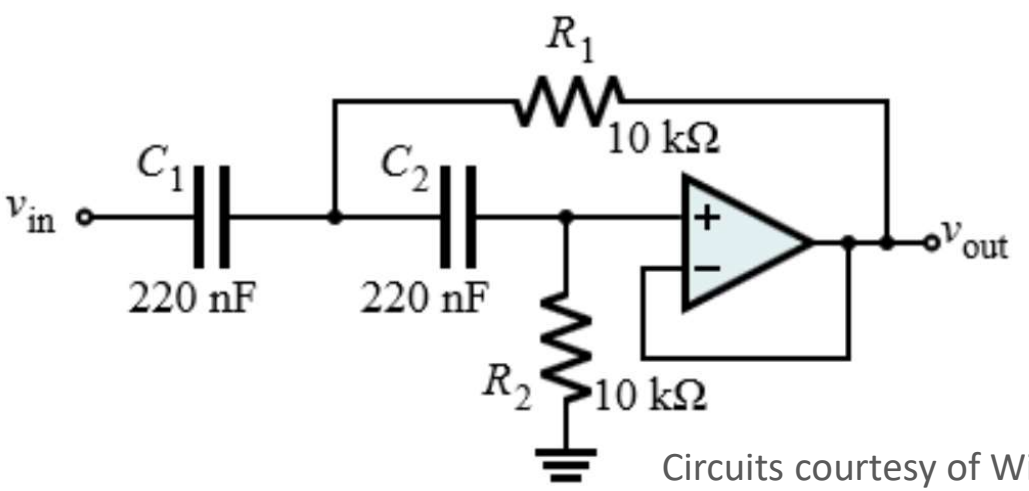
Sallen-Key Filter Examples

- Low-Pass Filter (LPF):



$$f_c = \frac{1}{2\pi\sqrt{R_1 R_2 C_1 C_2}}$$

- High-Pass Filter (HPF):

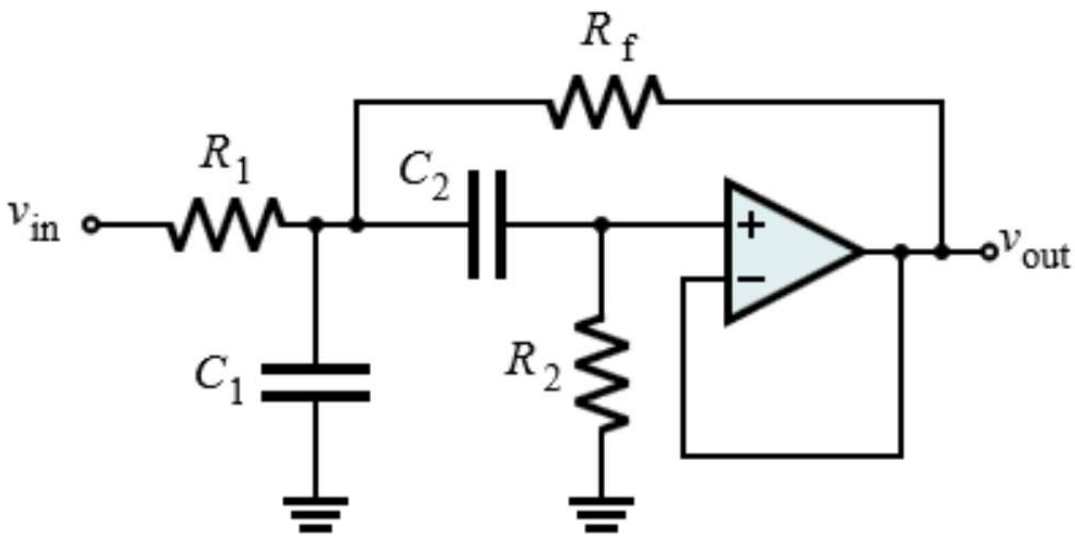


$$f_c = \frac{1}{2\pi\sqrt{R_1 R_2 C_1 C_2}}$$

Circuits courtesy of Wikipedia – creative commons license.

Sallen-Key Filter Examples

- Band-Pass Filter (BPF):



$$f_0 = \frac{1}{2\pi}\sqrt{\frac{R_f + R_1}{C_1 C_2 R_1 R_2 R_f}}$$

- In these three circuits, and in the analysis, we assumed gain = 1 by setting R_1 = 0 and R_2 = ∞. However, we can add gain by the equation $A = 1 + (R_1/R_2)$. This is possible for all circuits and just adds amplification to the filter function.
- More gain tends to widen the bandwidth of our filter as the passband starts higher and needs longer to decay before it is a small number (< 1).

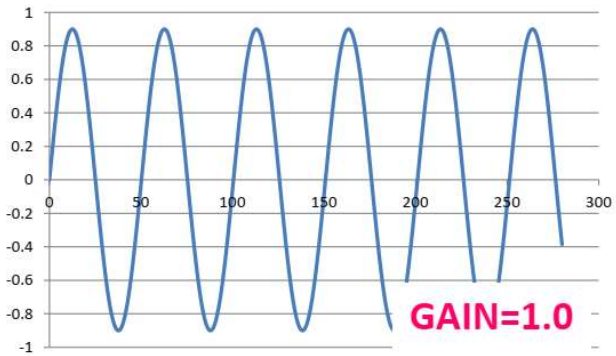
Op-Amp Oscillators

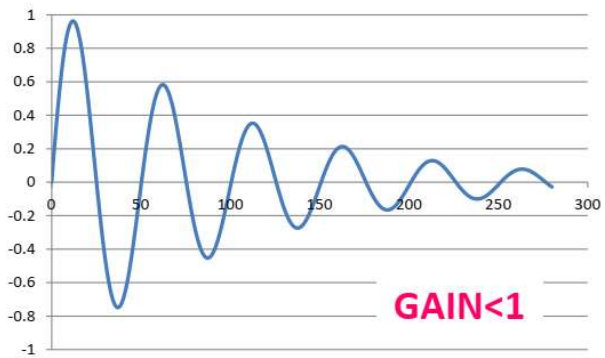
- Oscillations occur when the output of a system is swinging wildly from one extreme to another.
- It occurs when signals chase each other instead of being in negative feedback. For example, subtracting a sine wave when it's in the positive half, it arrives too late (now it's in the negative half) and ends up giving positive feedback.
- There is a set of rules for this called the Barkhausen criterion:

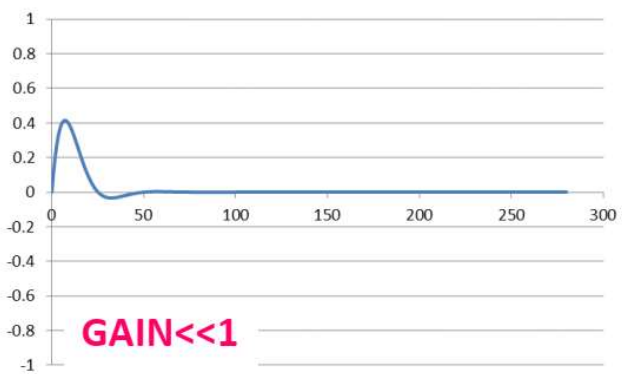
  A practical steady-state oscillator requires an amplification ($A > 1$) and a positive feedback.
- Op-amps produce gain, lots of it. Parasitic and intentional capacitors will introduce phase shifts.
- It is VERY EASY to accidentally satisfy this criterion when producing a filter. Ironically, it can be VERY HARD to achieve it on purpose.

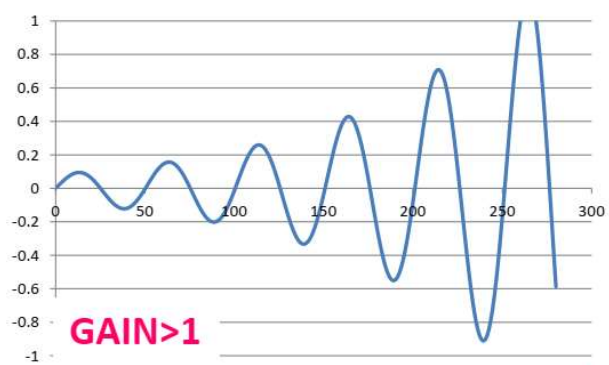
Op-Amp Oscillators

- Under negative feedback a phase shift of 180° means that instead of subtracting, one is now adding (minus minus).
- The question is, when one is adding the output to the signal at a frequency with this phase shift, does the combination get bigger or smaller? This is the gain criterion.

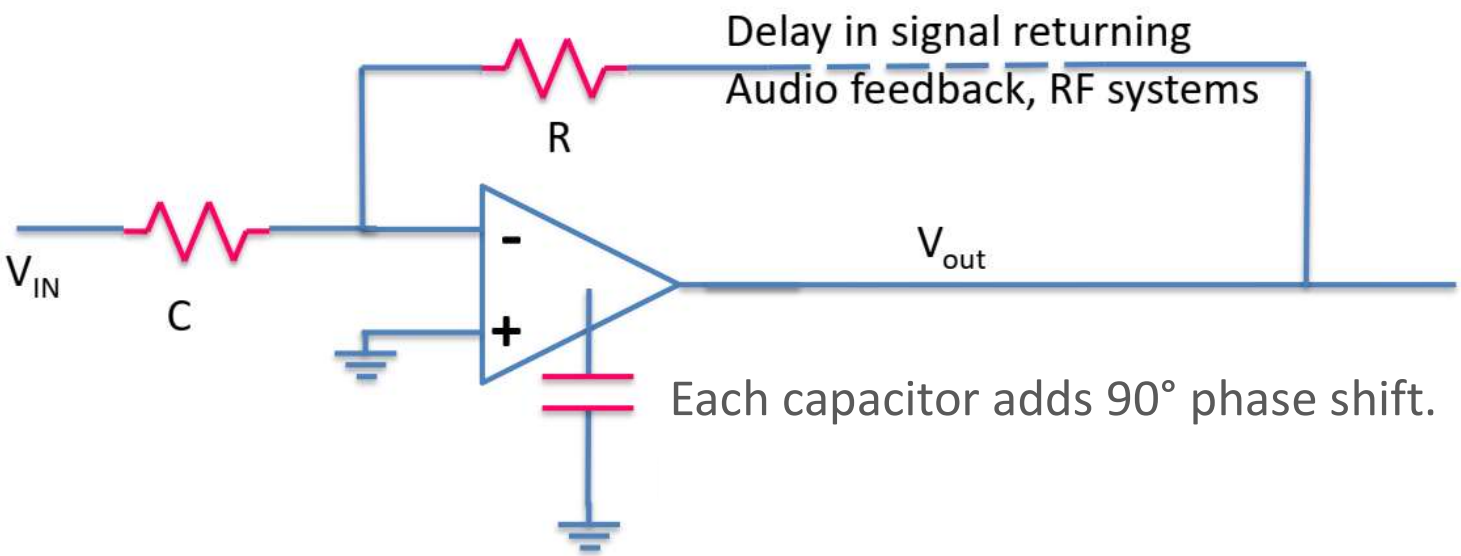






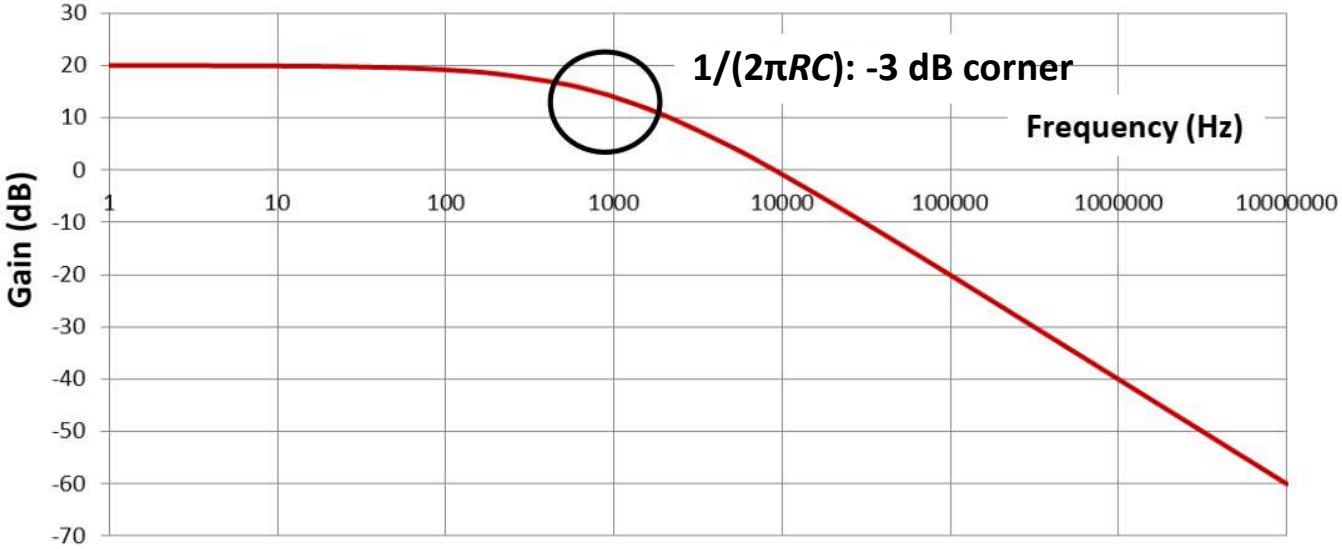


Phase Shifts: How Do They Arise?

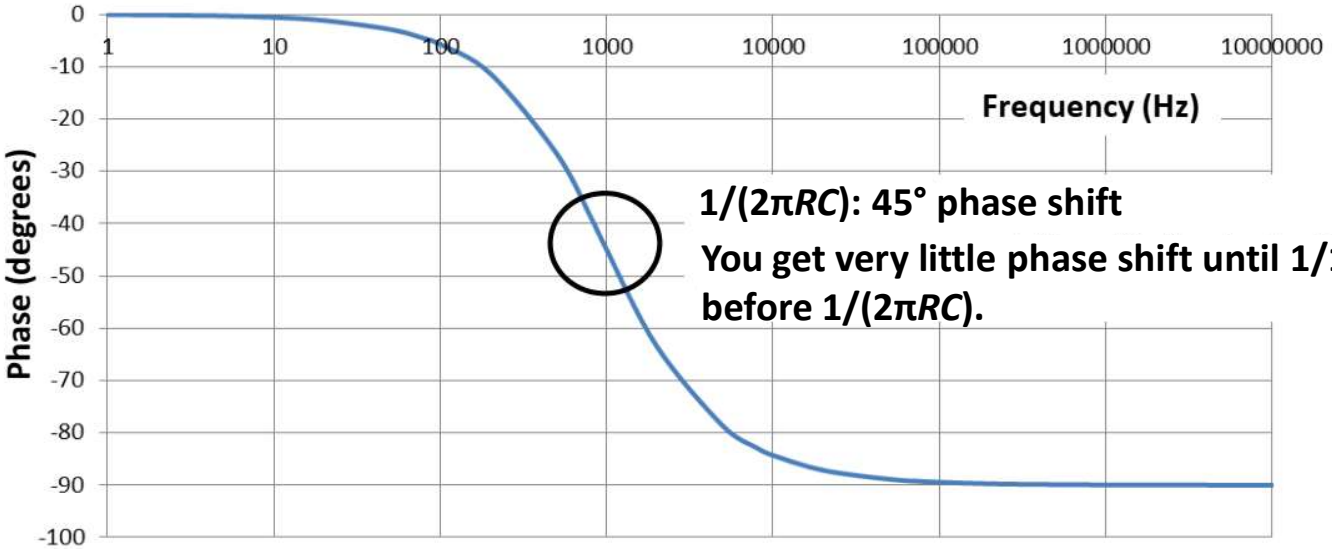


A capacitor to ground INSIDE THE OP-AMP is almost always present due to parasitic capacitances in the transistors. You will normally get at least a 90° phase shift out of it. But the real question is at what frequency.

Phase Shift: A Single Capacitor



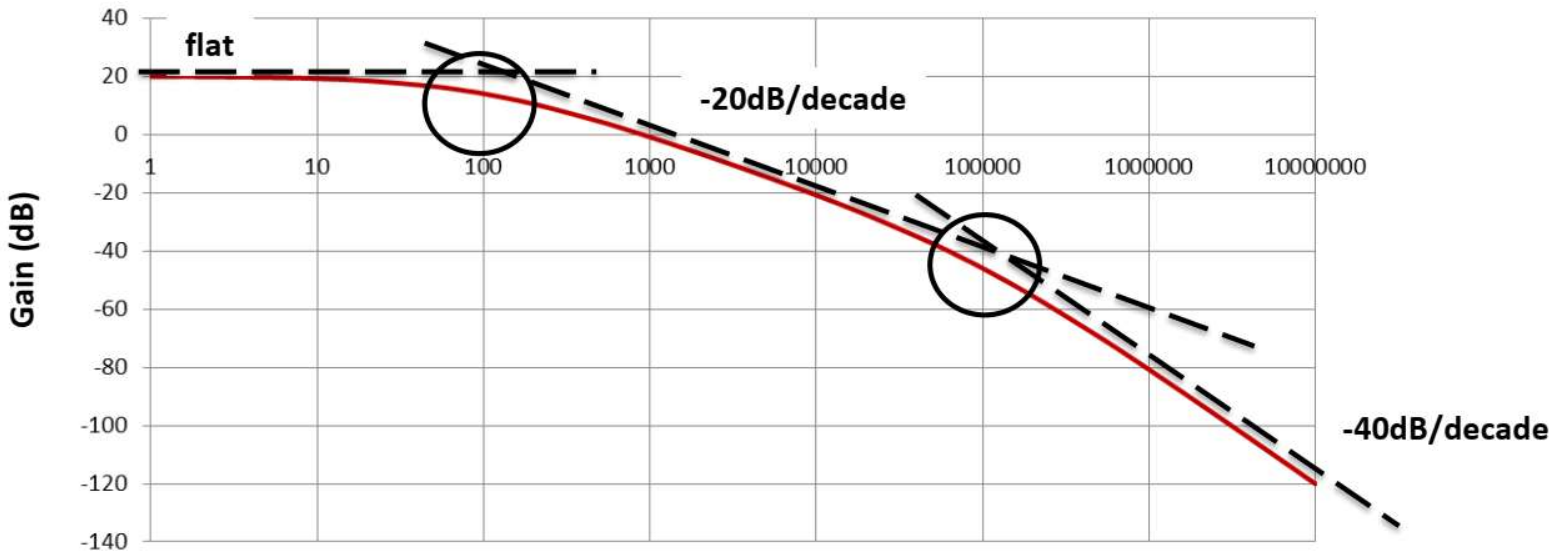
Each capacitor adds a 90° phase shift and a reduction of 20 dB/decade



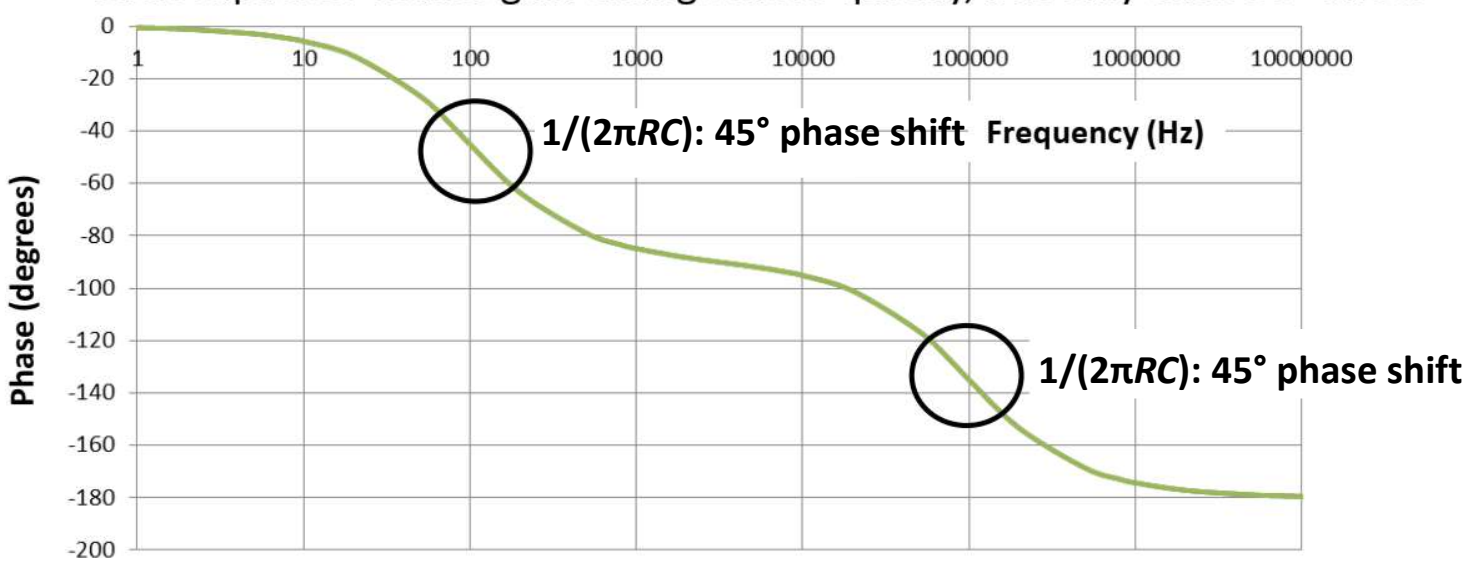
1/(2π*RC*): 45° phase shift

You get very little phase shift until 1/10 of the frequency before 1/(2π*RC*).

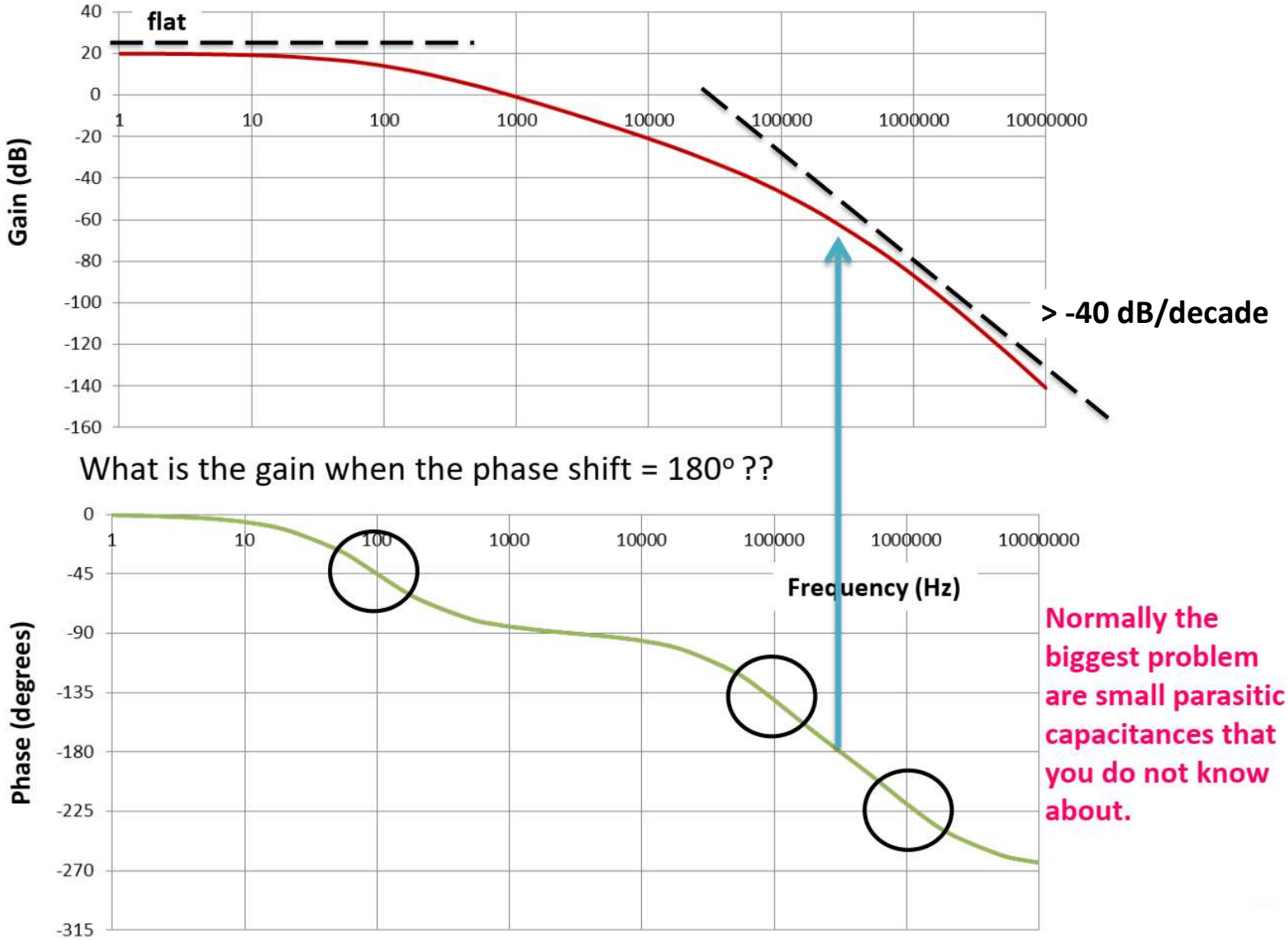
Phase Shifts: Two Capacitors



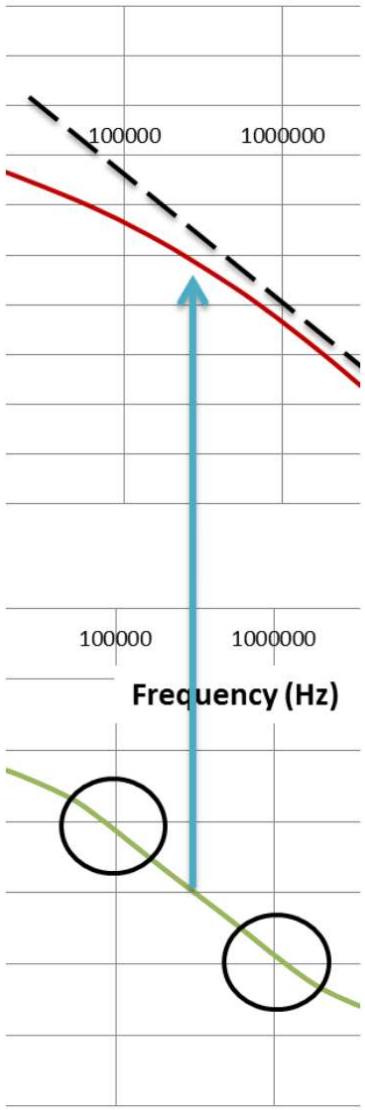
Each capacitor makes gain change more quickly, but only adds 90° once.



Phase Shifts: Three Capacitors

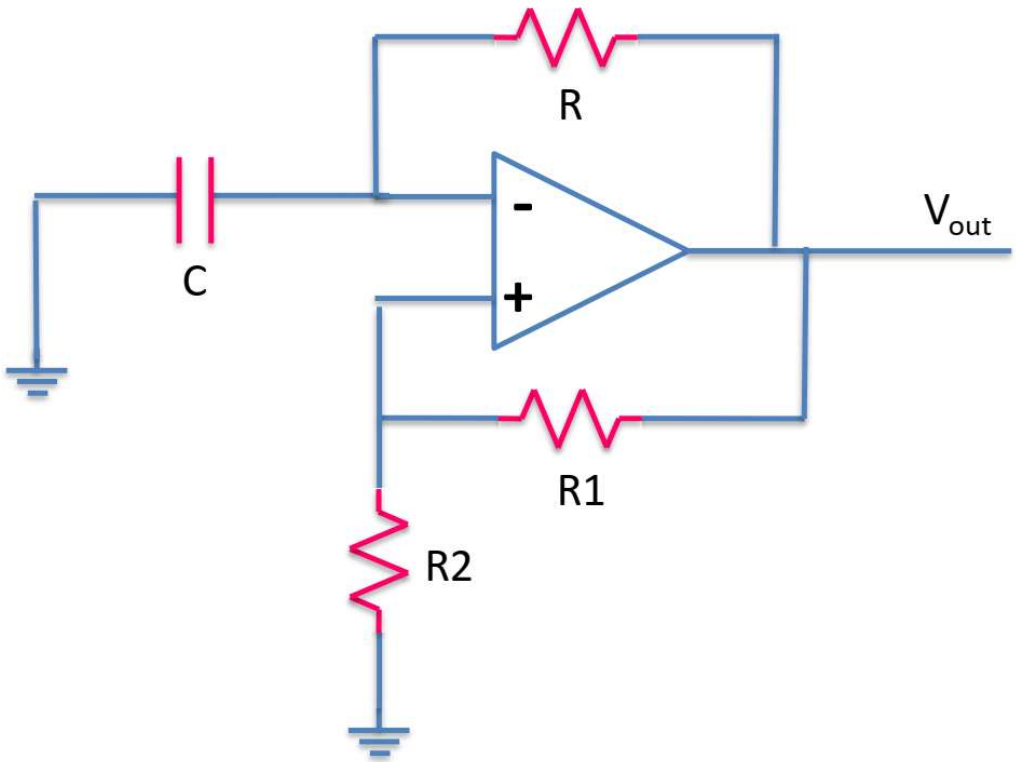


Gain & Phase Shifts



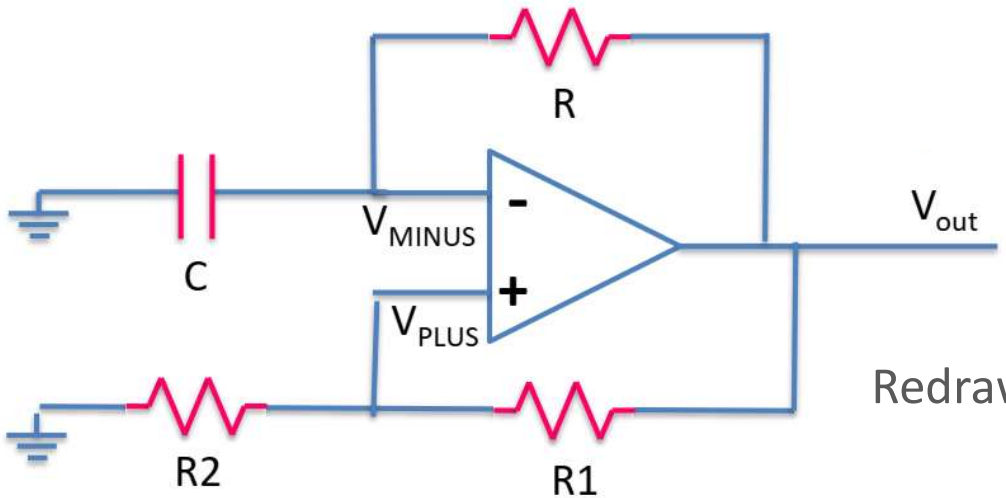
- In this case, gain = -60 dB at phase = 180°.
- Barkhausen criterion for oscillation says that gain must be > 1 (0 dB) at phase = 180°. This means we get no oscillations.
- Imagine now if we make the gain of the op-amp bigger. If we add no extra capacitors, we don't get phase shifts. In this scenario, we increase gain at the same phase. At some point we could get oscillations.
- We often like a lot of gain in our op-amps. However, sneaky accidental parasitic capacitors could cause us a lot of problems.

Self-Starting Oscillators



- This is a self-starting oscillator of the "relaxation" type.
- R_1 and R_2 look similar to a non-inverting amplifier option, but the top two look like a differentiator. A bit confusing. It is easier to understand once you've seen one working in your lab. But let's also try to explain it.

Relaxation Oscillator



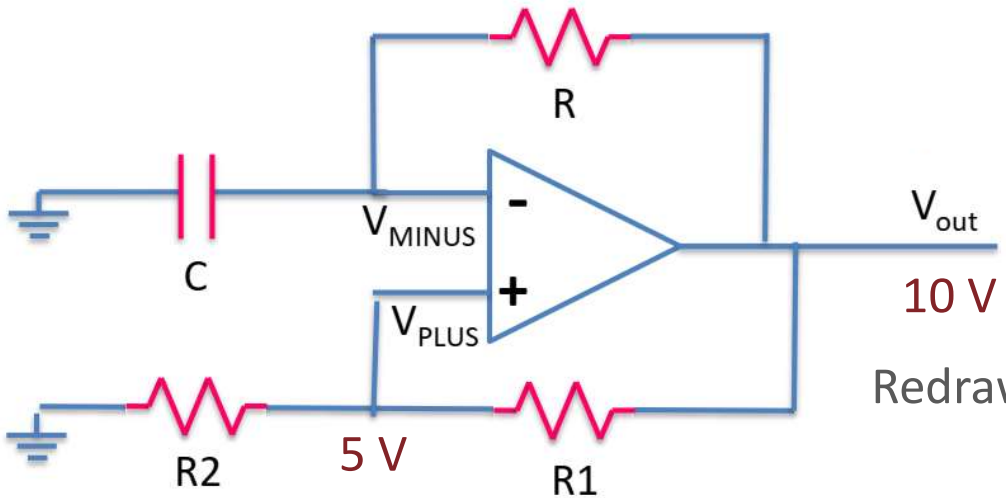
Redrawn for clarity.

- First, the basic rule remains the same.

$$V_{\mathrm{OUT}} = A(V_{+} - V_{-})$$

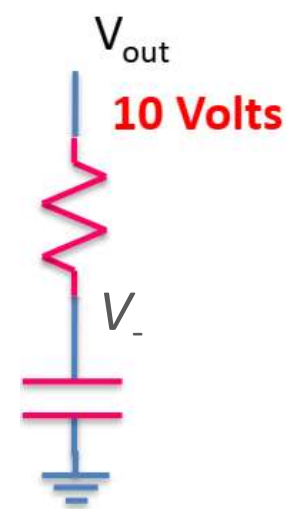
- Let's imagine the capacitor is in zero volts but V_{OUT} for some reason is positive (e.g., a little bit of noise or an internal input offset).
- In this scenario, V_{+} would be some fraction of a positive number. Therefore, V_{OUT} would grow more positive, and hit the max voltage (say 10 Volts).
- V_{+} would hit a maximum value given by R_1 and R_2 (say, 5 V if $R_1 = R_2$).

Relaxation Oscillator

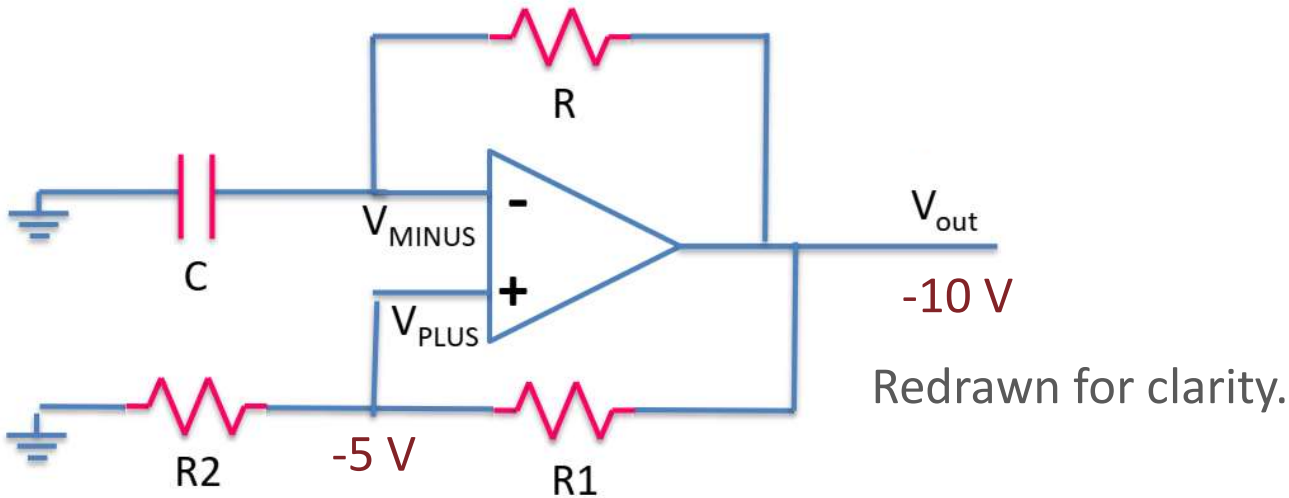


Redrawn for clarity.

- We've been ignoring what has been happening on the *RC* path. As V_{OUT} has grown bigger, the capacitor has been slowly charging up. Gradually the capacitor will charge up to the maximum voltage V_{OUT}.
- Before it gets there though, V_- which is defined by the charge on the capacitor will be bigger than V_+. Then the op-amp will obey the rule, $V_{\mathrm{OUT}} = A(V_+ - V_-)$.
- V_{OUT} will become negative and hit the minimum voltage of -10 V.

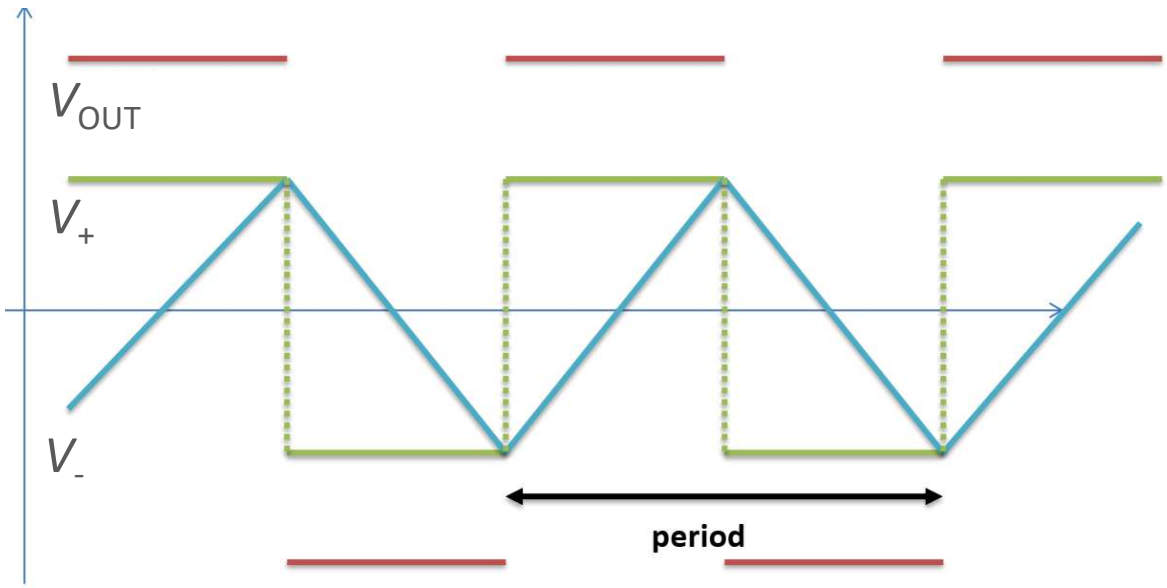


Relaxation Oscillator



- As V_+ is set by resistors not capacitors, it instantly changes value to -5 V. But this means V_- is very much bigger than V_+, so V_{OUT} stays at -10 V.
- While V_{OUT} is at -10 V, the capacitor discharges and loses voltage until it hits -10 V. When it crosses below the value of V_+, then (V_+ - V_-) becomes positive, and V_{OUT} jumps up to 10 V.
- We start all over again.

Relaxation Oscillator



- The output is a square wave and V_- is a triangular sawtooth waveform.
- The period depends on a number of factors but basically it is the rate of change of the voltage across the capacitor which is related to $1/(2\pi RC)$ and the trigger threshold that is set by R_1 and R_2 (in another word, V_+).
- Make the trigger level smaller, the distance between the V_+ levels will be smaller, and triggering will happen faster.
- Make the capacitor smaller (C) or the resistor (R) smaller, and the capacitor will charge up faster, reducing the travel time again.

Op-Amp Oscillators

- There are many types of oscillators which we will not go into. However, just as amplifiers can oscillate because of parasitic capacitances, oscillators can also act strangely because of them. The relaxation oscillator is pretty robust and is the basis of the famous 555 timer.
- An old saying from an unknown engineer:

  My amplifiers oscillate

  and my oscillators amplify.
- Care is always needed, even with op-amps.

EE204FZ: Analogue Electronics 1

Introduction to Bipolar Junction Transistors (BJTs)

What is a transistor?

The field-effect transistor (FET) uses electric fields, and operates as

A voltage controlled variable "resistor" or "current source" depending on the configuration.

The bipolar-junction transistor (BJT) is also about the movement of charge carriers and will operate as

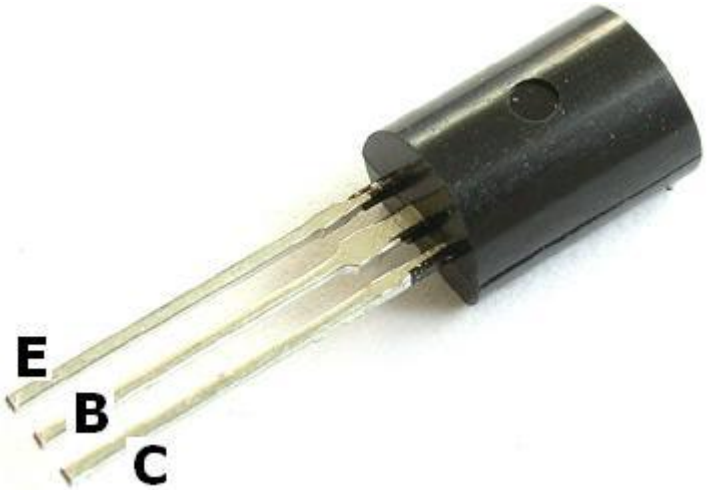
A current controlled "current source".

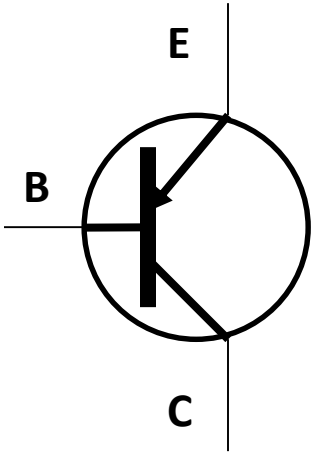
What type of transistor is best??

FETs	BJTs
Most modern designs use them	First type of transistor, so older designs use them
A bit noise	Best for low noise performance
Run slower, but now good up to about 10 Ghz	Run the fastest, up to 50 GHz or higher, material depending
Scales very well with new fab technology, can pack these transistors in very tightly	Do not use the modern very small nanoscale very well
FET circuits can be very power efficient	Not power efficient
Voltage controlled	Current controlled

As usual, take your pick, but mostly we use FETs

What is a *Bipolar Junction* Transistor?





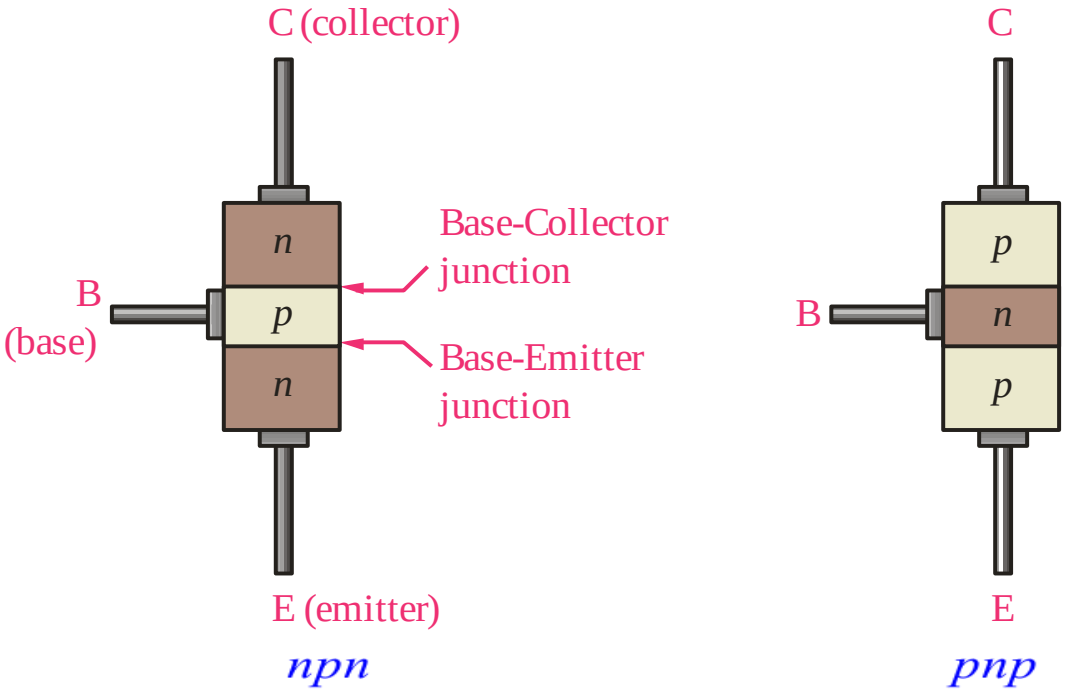
arrow points to N

The current that flows from the collector to the emitter is dependent on the current that flows through the base. There are two types of BJT, PNP and NPN which depends on the base semiconductor type.

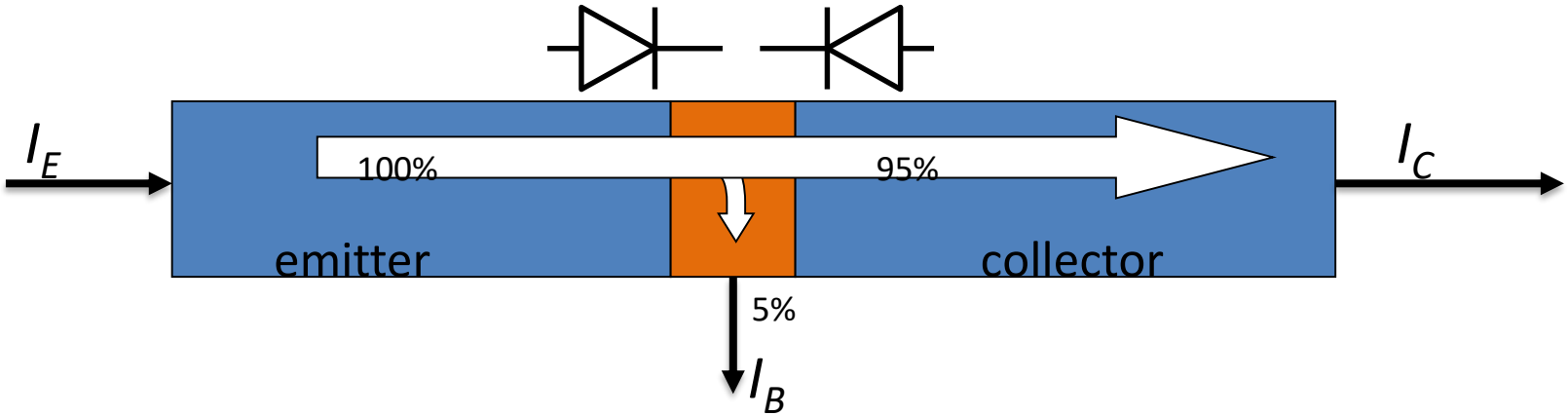
BJT Structure

The BJT has three regions called the emitter, base, and collector. Between the regions are junctions as indicated.

The base is a thin lightly doped region compared to the heavily doped emitter and moderately doped collector regions.

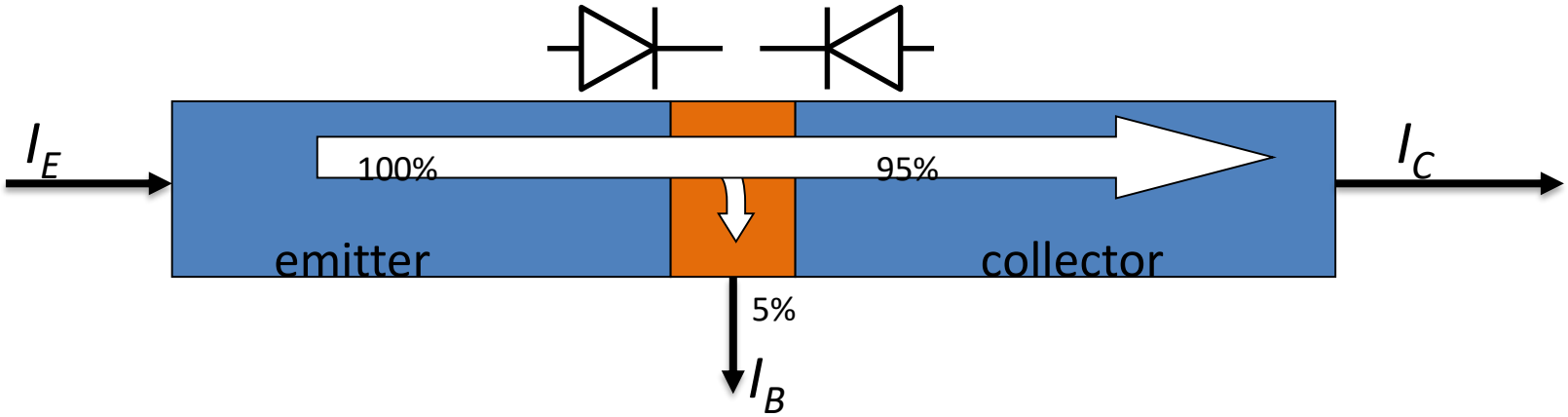


BJT Operation - Qualitative



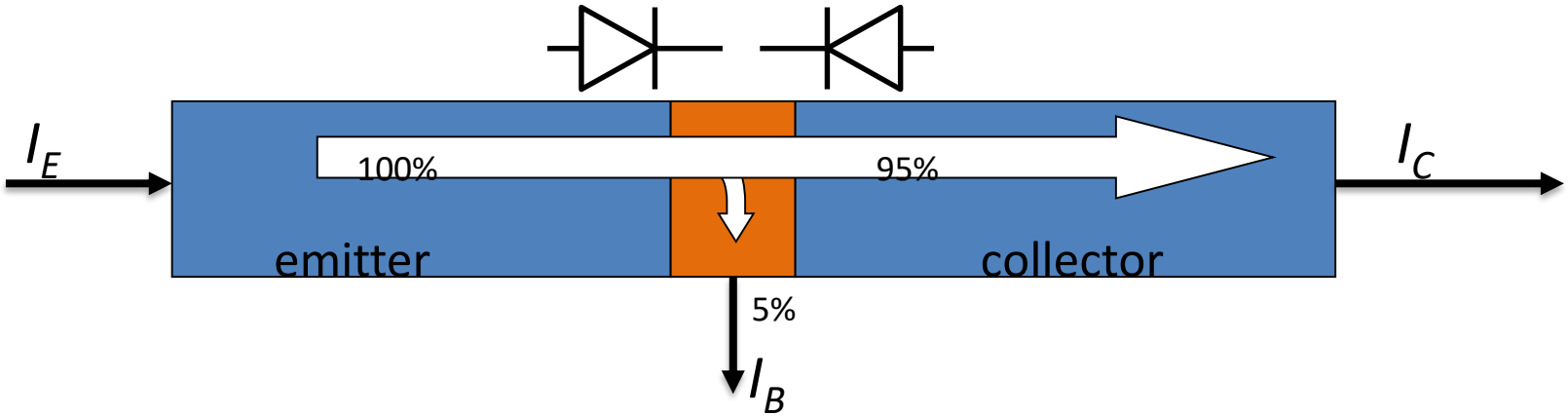
- The left hand (EB) junction is forward biased, the other (BC) is reverse biased.
- A current I_B is pulled from base which forces majority carriers from emitter into the base where they are **minority carriers**.
- Minority carriers near a **reverse biased junction** get swept across the junction.
- When the **base is narrow**, most carriers get caught by the reverse biased junction without recombining. The small percentage that gets drawn out of the base forms I_B.

BJT Operation - Qualitative



- A certain percentage of carriers **(assume holes)** will recombine, holes with electrons in the n base. They never make it across .
- The electrons used up in the base by recombination with travelling holes, are supplied by the base current.
- If we don't, a charge imbalance will build up which will stop the emitter-collector current. So no base current, no full current.
- There is a fixed ratio between base and collector current. The main limit on the large current is how fast we can pump in replacement electrons, the faster we do it the better.
- The more base current, the bigger the electron current.

BJT Operation – Key Equations



From using KCL for the currents in the base, we get

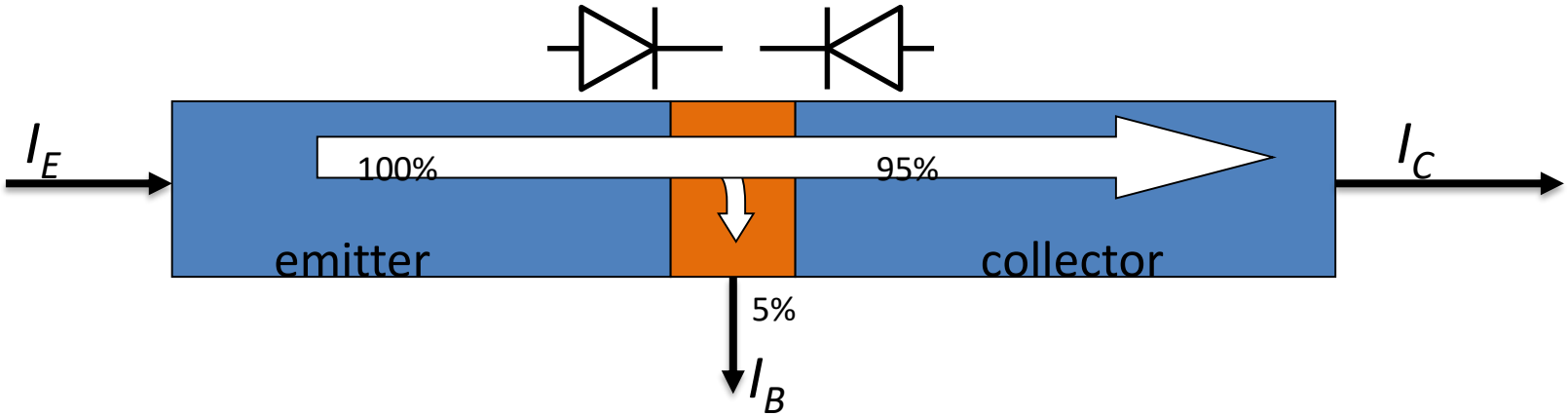
$$I_E = I_C + I_B$$

$$I_C = \alpha I_E$$

where α is always less than 1.0. It is called the common base current gain or **forward current transfer function**.

Hand-waving: Not all the emitter current gets to the collector, some always is lost in the base, has to be. But the closer to 1.0 the better.

BJT Operation – Key Equations



Reworking the equations a little, we get

$$I_C = \beta I_B$$

where β is typically large, > 50, called the **common-emitter, current gain**. When people mention BJT transistor current gain, they normally mean β.

Hand-waving: This tells us how powerful the base current is in modifying the main current, how big a change in the main current occurs for a little change in the base current. The higher the better.

More technically… the diode equation

Lets manipulate our equations a bit more.

$$I_C = I_E - I_B$$
$$I_C = \alpha I_E = \beta I_B$$

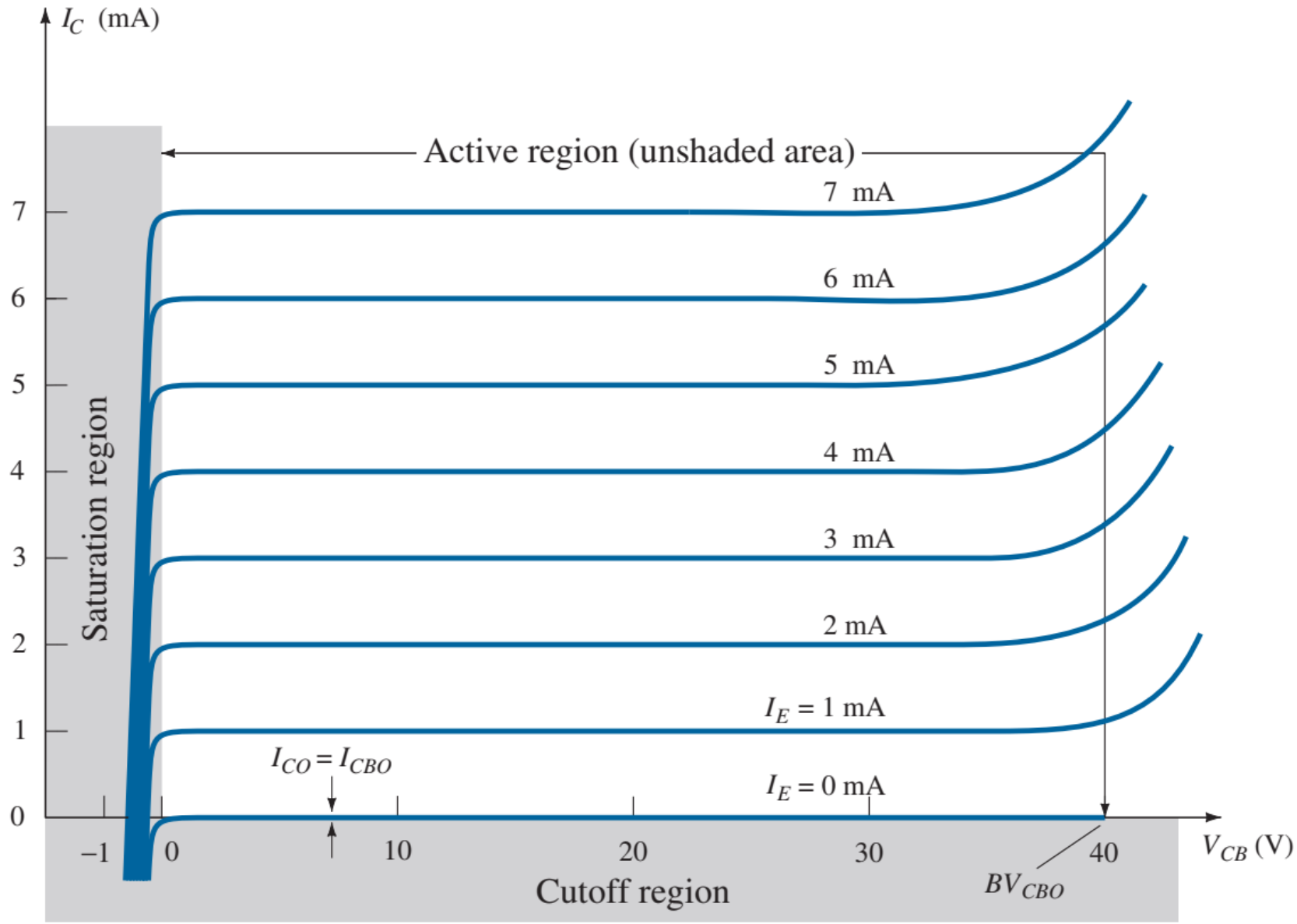
So

$$\alpha\left(I_C + I_B\right) = \beta I_B$$
$$\alpha\left(\beta I_B + I_B\right) = \beta I_B$$
$$\alpha\left(\beta + 1\right) I_B = \beta I_B$$
$$\alpha\left(\beta + 1\right) = \beta$$

Finally

$$\alpha = \frac{\beta}{\beta + 1} \qquad \beta = \frac{\alpha}{1 - \alpha}$$

BJT Operation – Plotting



BJT Operation – Key Equations

$$I_E = I_S\left(e^{(qV_{BE}/KT)} - 1\right) \quad *$$

$$I_C = \beta I_B \quad *$$

$$I_E = IC + I_B \quad *$$

$$\alpha = \frac{\beta}{\beta + 1}$$

$$\beta = \frac{\alpha}{1 - \alpha}$$

$$V_{\mathrm{CE}} > V_{\mathrm{CE(SAT)}}$$

$$V_{\mathrm{BE}} > V_{\mathrm{BE(SAT)}}$$

$V_{CE(SAT)}$ and $V_{BE(SAT)}$ are the **effective boundaries** of the active region of the BJT transistor and in the active region, we basically do not have the exponential, and just $I_C = \beta I_B$.

$V_{BE(SAT)}$ or $V_{BE(ON)}$ is the voltage required to make the forward biased diode work. If V_{BE} isn't bigger than this, no current flows.

$V_{CE(SAT)}$ is the minimum voltage required to make the transistor operate in the flat-region of the current, not the exponential area.

Using BJT's

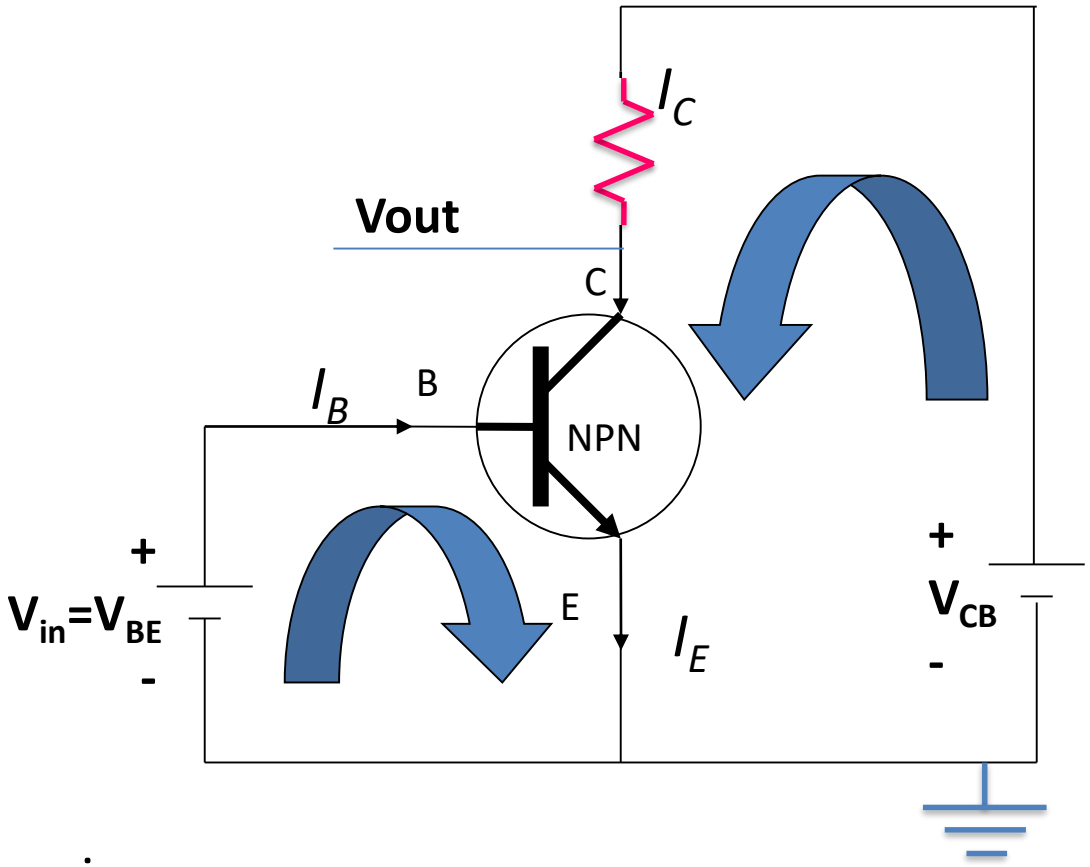
Like FETs there are two aspects to a circuit using BJT's:

- Large signal or dc operation – which we call **biasing**
- Small signal operation – or amplification

There are three modes of operation for a bipolar junction transistor:

- **Common base** – an input buffer and as an amplifier
- **Common collector** – an output buffer
- **Common emitter** – an amplifier

BJT – Common Emitter



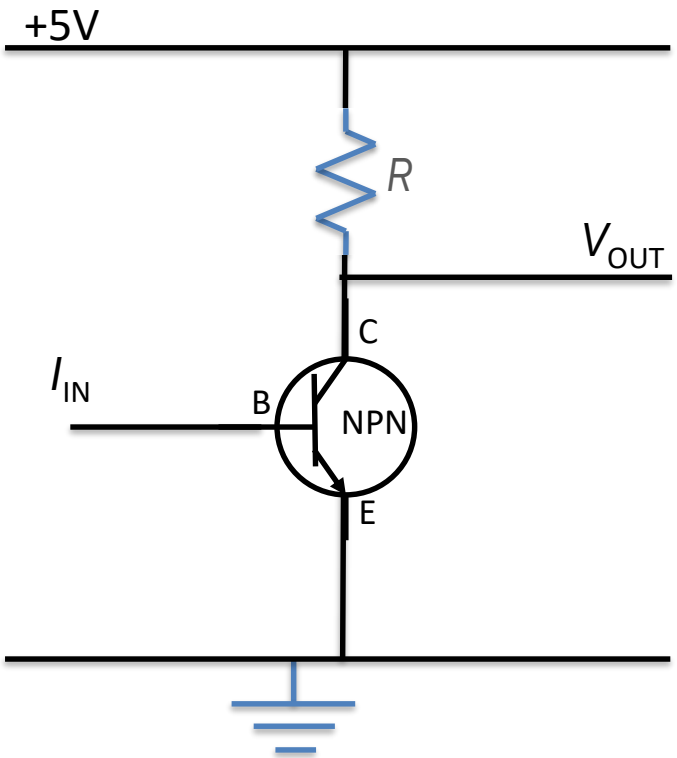
Phase inversion
current gain > 1 (10 to 200)
voltage gain > 1
Input impedance is higher than CB (GOOD for weak input signals)
Output impedance is lower than CB (GOOD for driving high-freq)
power gain > 1

Common-Base Configuration

BJT Biasing – BIG DIFFERENCE

Current flows through the base of a BJT transistor

A Biasing example



I have a small voltage signal that I am going to add on top of a DC average value. I will want some small-signal gain and a maximum output voltage swing of +/- 1.5 volts. Pick values of R, I_C and I_B for my design.

$V_{CE(SAT)}$= 0.3 V $V_{BE(ON)}$= 0.7 V β = 50

To get a $\pm$1.5 volts swing, pick a mid-point on the range, and a little bit on the high side to get the BJT some extra voltage. For me, I'll pick

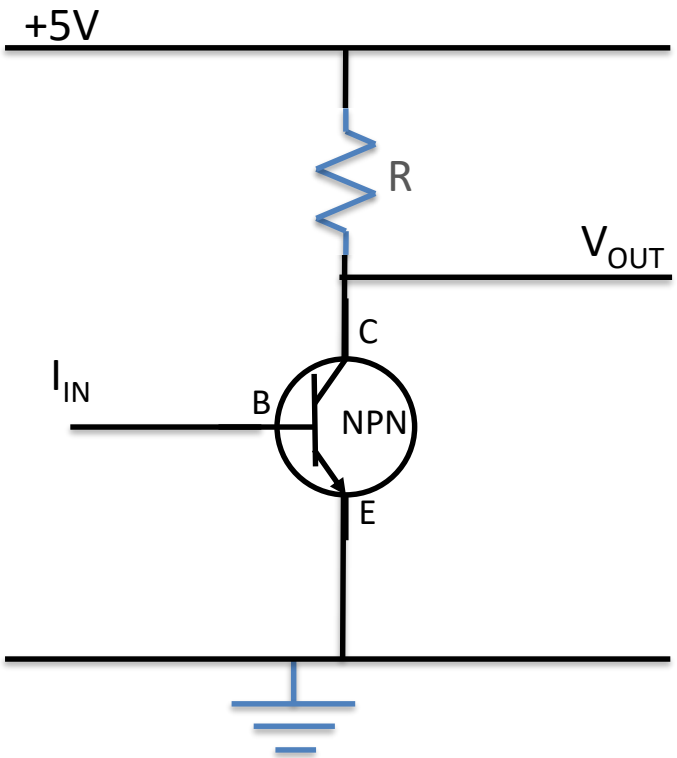
V_{OUT} to be 3 volts

$I_C \times R$ = 2 volts

V_{CE} = 3 volts >> $V_{CE(SAT)}$

A Biasing example

V_{OUT} to be 3 volts
$I_C \times R$ = 2 volts
V_{CE} = 3 volts >> $V_{CE(SAT)}$

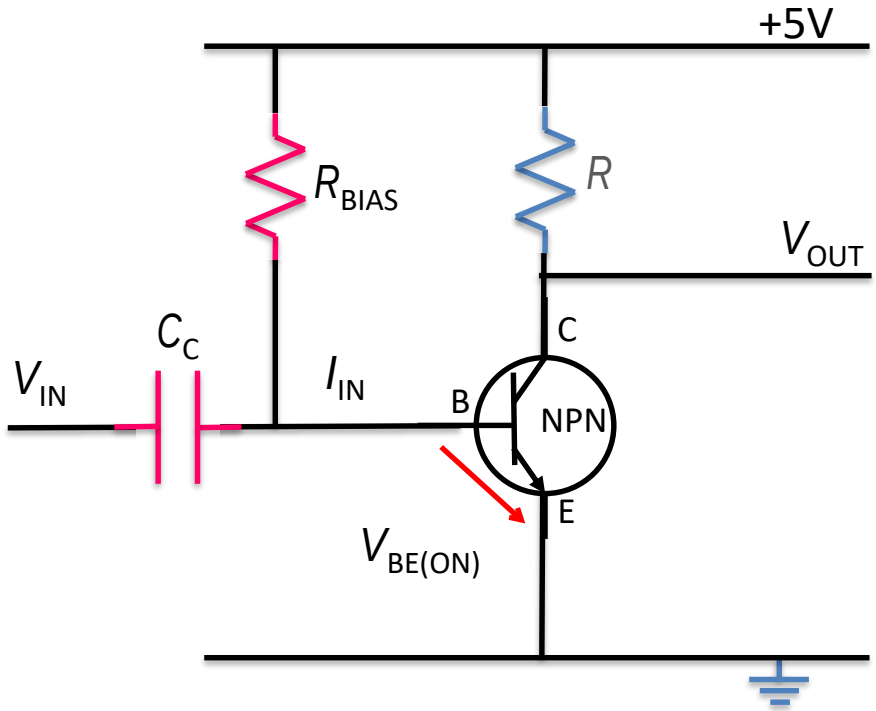


Normally you would need to look up a datasheet to find a value for I_C that gives you a good value of β.

Let's cheat, and say that I_C = 1 mA gives me a β of 50.

So I_C = 1 mA and $I_C \times R$ = 2 volts, so R = 2 kohm.

If I_C = 1 mA, then I_B must be equal to 1 mA/β or 20 μA.

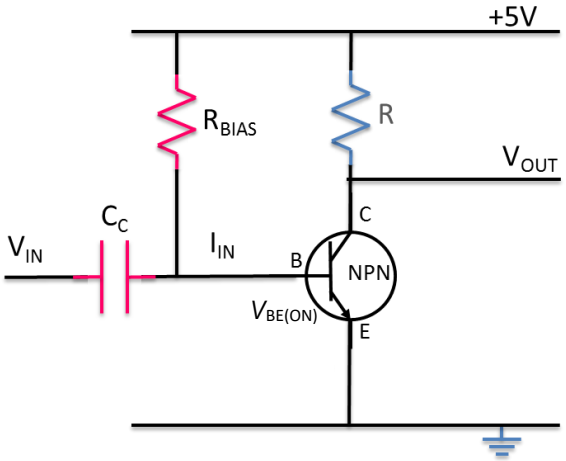


So we add a resistor to the V_{CC} (5 V) supply and let the current flow. Flowing from base to emitter, the current will experience a voltage drop called $V_{BE(ON)}$. We need to make sure it is always bigger than that, even with an input swing, so we add the amplitude of our input signal voltage to it:

$$I_{IN}\left(R_{BIAS}\right)+V_{BE(ON)}=5$$

$$R_{BIAS}=\frac{5-V_{BE(ON)}}{I_B}=\frac{\beta\left(5-V_{BE(ON)}\right)}{I_C}$$

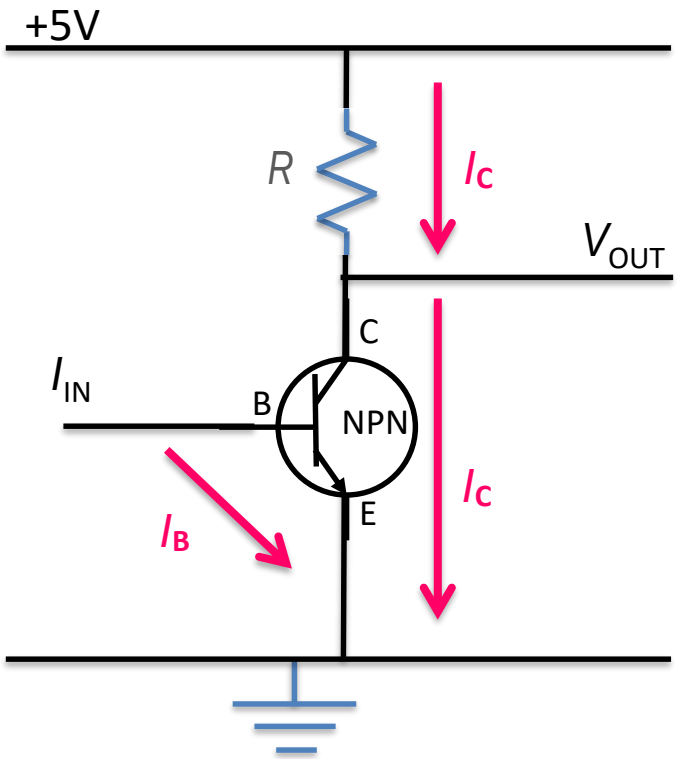
A Biasing example



Plugging in the numbers, and remembering we want V_{BE} to be always a little bigger than $V_{BE(ON)}$ so that our input signal won't turn it off when it goes a little low (say V_{IN} has amplitude 0.2). Altogether we get

$$
\begin{aligned}
R_{\mathrm{BIAS}} &= \frac{V_{\mathrm{BIAS}}}{I_{\mathrm{B}}} \\
&= \frac{\left(5 - V_{\mathrm{BE(ON)}} - |V_{\mathrm{IN}}|\right)}{I_{\mathrm{C}} / \beta} \\
&= \frac{(5 - 0.7 - 0.2)}{0.001/50} \\
&= \frac{50(4.1)}{0.001} = \frac{205}{0.001} \\
&= 205\,\mathrm{k}\Omega
\end{aligned}
$$

Amplification



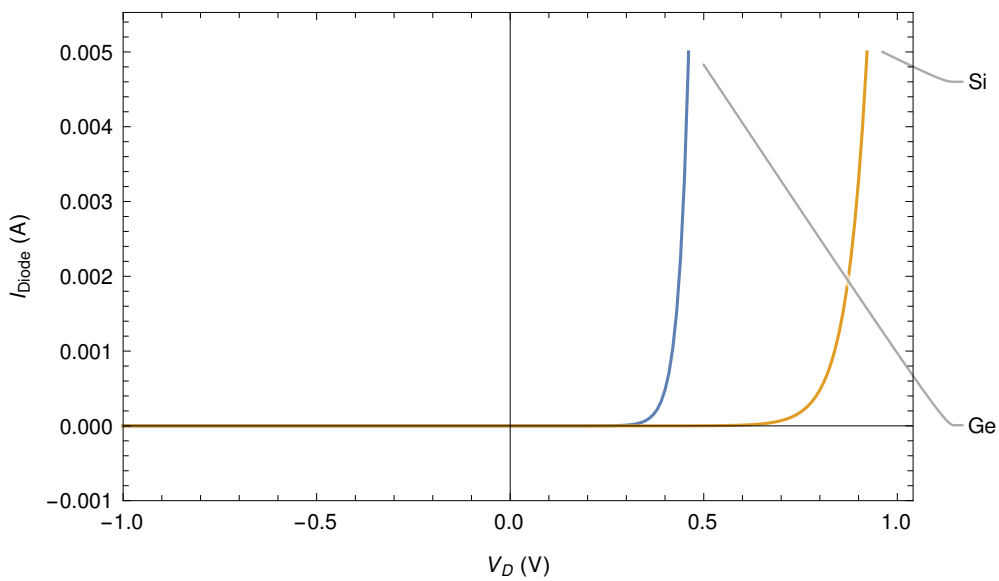
So very similar to the FET:

I_B increases, then $I_C = \beta I_B$ goes up.

I_C goes up, then $I_C \times R$ goes up.

$I_C \times R$ goes up, V_{OUT} goes down.

Thus we get amplification.

In[ ]:=
```
(*Diode current vs. electric potential difference across a p-n junction diode.*)
(*Assign the reverse saturation current of 100 pA.*)
(*The thermal voltage is ~26 mV at RT.*)
Is = 100 × 10^-12; VT = 26 × 10^-3;
(*Define the Shokley diode equation.*)
I_Diode[Is_, VD_, VT_, n_] := Is (Exp[VD/(n VT)] - 1);
(*Plot two curves, one with n = 1 (Ge) and the other with n = 2 (Si).*)
Curves = Table[Table[{VD, I_Diode[Is, VD, VT, n]}, {VD, -1.0, 1.0, 0.01}], {n, 1, 2}];
CurveLabels = {"Ge", "Si"};
(*One can clearly see from the plot that the value of n controls
 at which electric potential difference the diode turns "on".*)
ListPlot[Curves, Joined -> True, PlotRange -> {{-1, 1}, {-0.001, 0.005}}, Frame -> True,
 FrameLabel -> {"V_D (V)", "I_Diode (A)"}, LabelStyle -> Medium, PlotLabels -> CurveLabels]
```

Out[ ]=