



Transistors

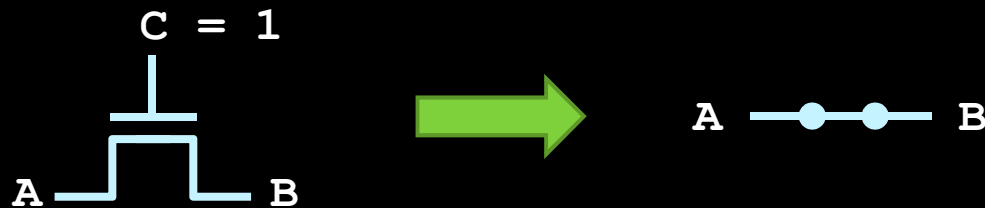
Introduction to Transistors

- **Transistors** form the basic building blocks of all computer hardware.
- Invented by William Shockley, John Bardeen and Walter in 1947, replacing previous vacuum-tube technology.
 - Won Nobel Prize for Physics in 1956.
- Used for applications such as amplification, switching and digital logic design.

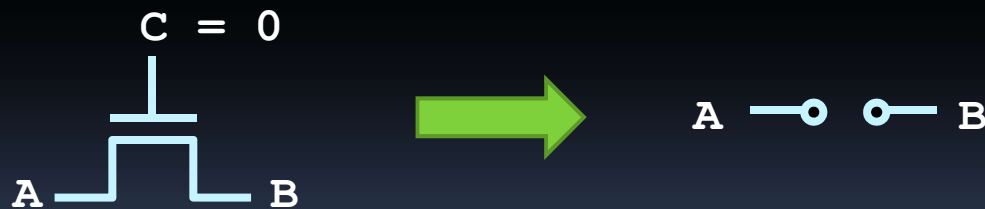


What do transistors do?

- Transistors connect Point A to Point B, based on the value at Point C.
 - If the value at Point C is high, A & B are connected.

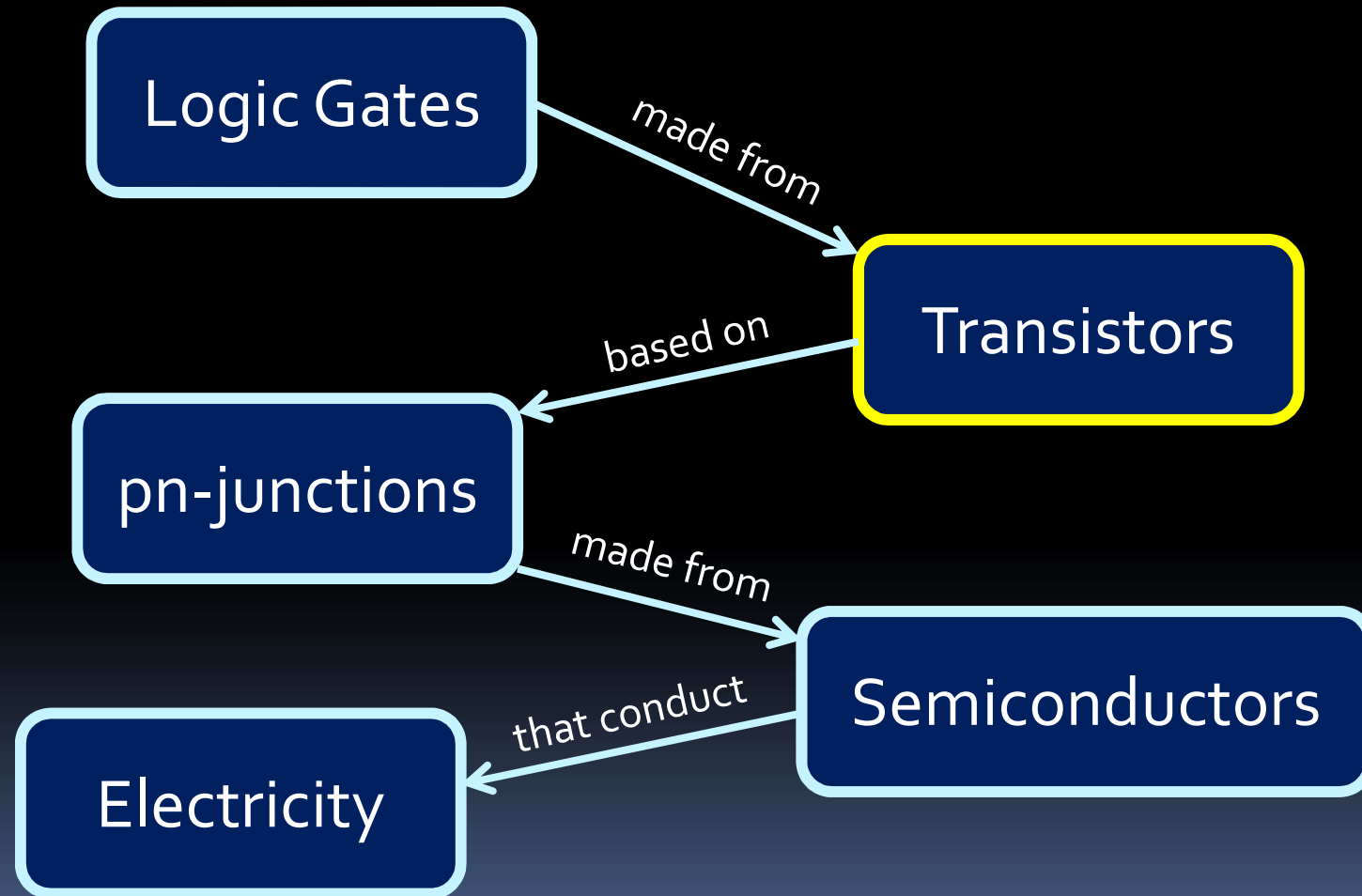


- And if the value at Point C is low, A & B are not.

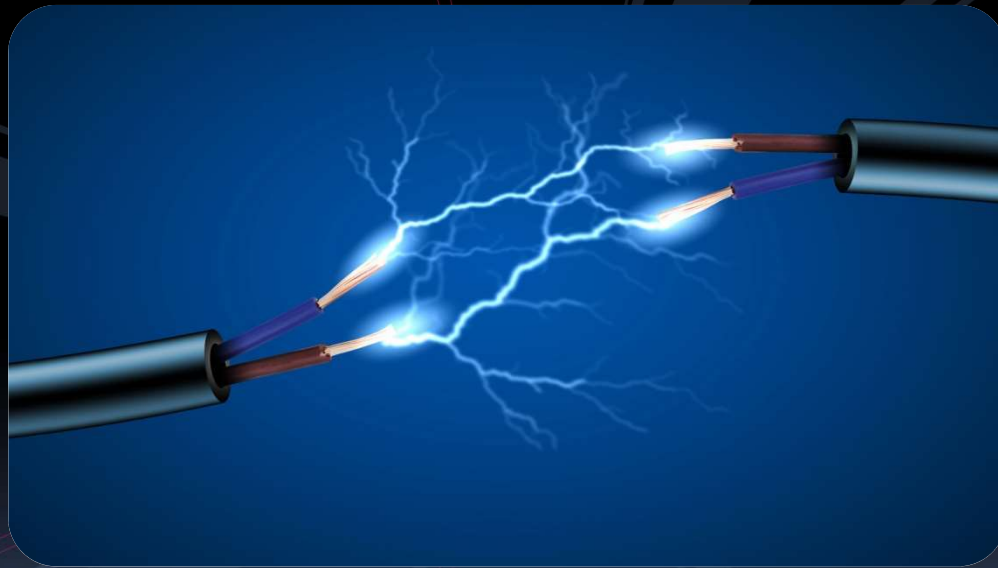


- Need to know a little about electricity now....

Where do transistors fit?

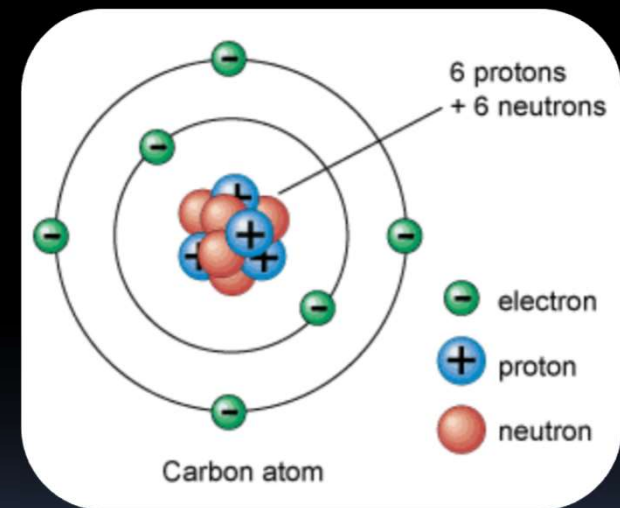


Electricity Basics



Intro to Electricity

- Electricity is the the flow of charged particles (usually electrons) through a material.
- These charged particles come from atoms, which are made up of **protons** (positive charge), **neutrons** (no charge) and **electrons** (negative charge)
 - Electricity stems from electron movement.



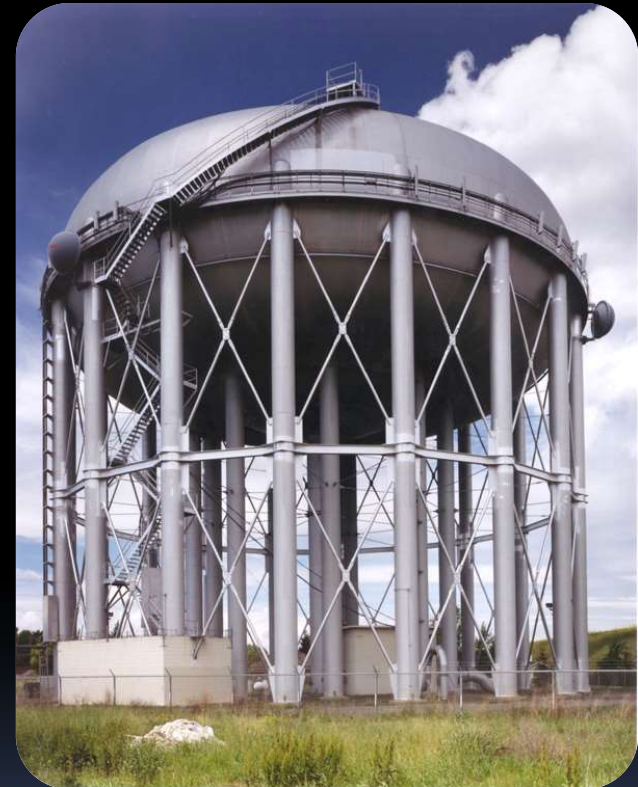
Electricity = electrons

- Electrical particles (like electrons) want to flow from regions of **high electrical potential** (many electrons) to regions of **low electrical potential** (not as many electrons).
 - Similar to gravitational potential
- This potential is referred to as **voltage**.
- The rate of electron flow is called the **current**.



Water Analogy

- To help picture this concept of voltage and current, imagine a reservoir:
 - Electrons flow from high to low potential like water would flow from the reservoir to the ground.
 - **Voltage** is like the elevation of the water above the ground.
 - **Current** is the rate at which the water flows.
- The relationship between voltage (**V**) and current (**I**) is called **resistance**: $R = V/I$

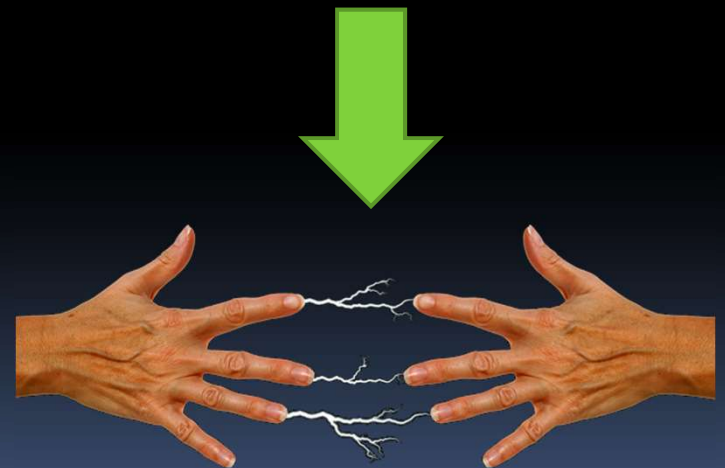


A Note about Current

- Even though current is caused by electrons flowing through a material, the convention is to measure current as the **movement of positive charges**.
 - *Protons don't actually move.* When electrons move from point A to point B, the result is that B becomes more negative and A becomes more positive.
 - Scientists historically viewed current in terms of this creation of positive charge in a material.
 - It's not completely clear why scientists decided this. Just go with it 😊

Static electricity example

- When you shuffle your feet back and forth on a carpet, you pick up extra electrons in your body and develop an electrical imbalance, relative to the ground.
- When you touch an object or person who is electrically balanced, those extra electrons transfer over to that object or person.



Van de Graaff Generator



Sources of electricity

- Where do these electrons (and this electricity) come from?
- Two common sources:
 - **Batteries** have a concentration of particles stored inside them up that will run out eventually (like water reservoirs).
 - Most electricity that we use comes from electrical outlets, that are constantly being supplied with electric particles that never run out (like waterfalls).
 - Discussion point: power bars.



The path of electricity

- A few things to note about the path that electric particles like to take:
 - Current always flows toward the zero voltage point of a circuit.
 - Commonly referred to as **ground**.
 - Electricity always like to take the path of least resistance, from source to ground.
 - Even though electrical current is the flow of electrons through a medium, its direction is typically expressed as *the movement of the positive charges*.

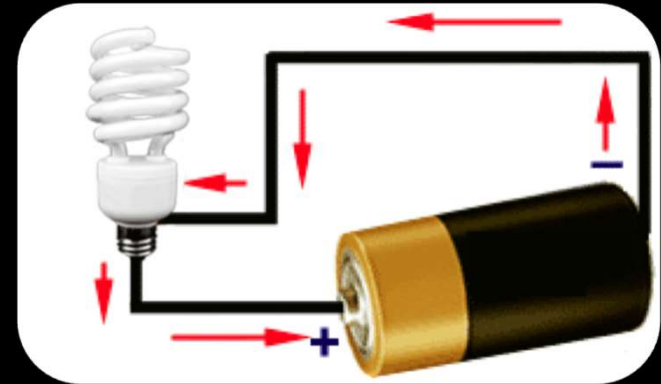


Electricity Example



Using electricity

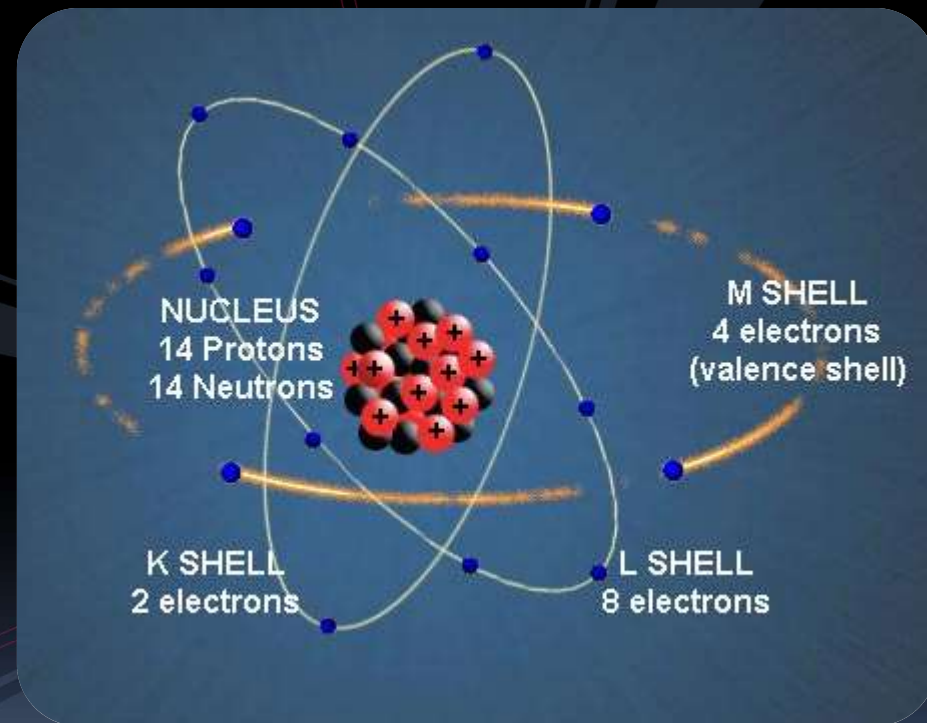
- Knowing that electric particles want to travel from areas of high concentration to areas of low concentration, we can use this to drive **circuits**.
- Each of these circuits has a **source** of electrical particles, some **path** between this source and the ground, and some **resistance** along this path that dissipates these electrons.



Resistance is Futile

- In the water analogy, **resistance** would be measure how restrictive the pipe is that connects the reservoir to the ground.
 - Wide, smooth pipe = low resistance
 - Narrow, twisty pipe = high resistance
- Electrical resistance indicates how well a material allows electricity to flow through it:
 - High resistance (aka **insulators**) don't conduct electricity at all, or only under special circumstances.
 - Low resistance (aka **conductors**) conduct electricity well, and are generally used for wires.
 - These are largely determined by the position on the element on the periodic table.
 - Measured in ohms (Ω). More ohms, more resistance.
- **Semiconductors** are somewhere in between conductors and insulators.

Semiconductors



What are semiconductors?

- Electricity can flow freely through a solid if there are free valence electrons in the outer layer after the solid is formed.
- Semiconductor materials (silicon, germanium) straddle the boundary between conductors and **insulators**, behaving like one or the other, depending on factors like temperature and **impurities** in the material.

Periodic Table of the Elements

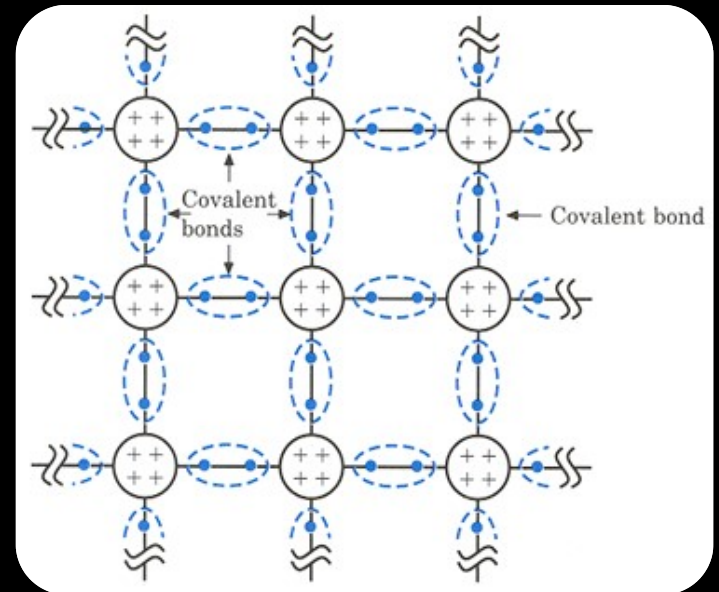
Legend:

- Alkali metals
- Alkaline earth metals
- Transition metals
- Lanthanide series
- Actinide series
- Poor metals
- Nonmetals
- Noble gases
- Solid
- Liquid
- Gas
- Synthetic

Atomic masses in parentheses are those of the most stable or common isotope.

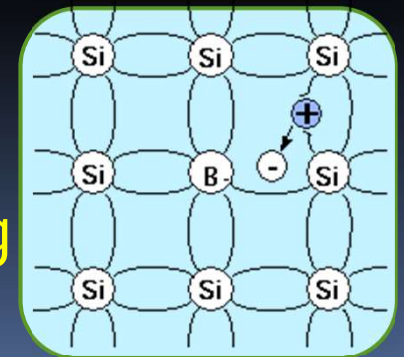
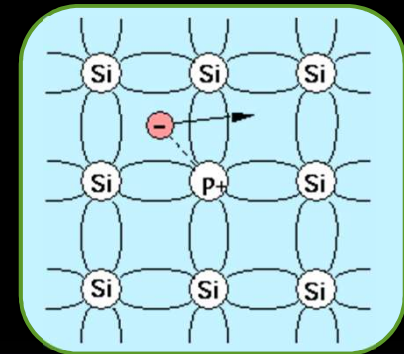
Semiconductor Conductivity

- Semiconductors are solid and stable at room temperature, but energy can make electrons from the valence layer become loose.
- At room temperature, a weak current will flow through the material, much less than that of a conductor.



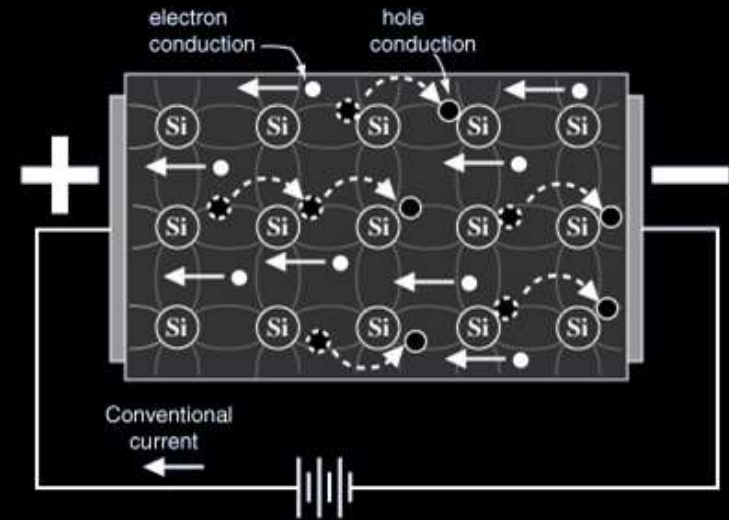
Adding Impurities

- Semiconductors don't conduct electricity naturally.
 - Silicon has 4 electrons in its outer (valence) electron layer.
 - Each atom wants 8, forms a lattice with its neighbours.
- To encourage the semiconductor's conductivity, **impurities** are introduced the fabrication process, to increase the number of free charge **carriers**.
 - **n-type**: adding elements from group 15, which have 5 electrons in its valence layer (e.g. phosphorus, arsenic).
 - **p-type**: adding elements from group 13, which have 3 electrons in its valence layer (e.g. boron).
- This process is also referred to as **doping** the semiconductor.



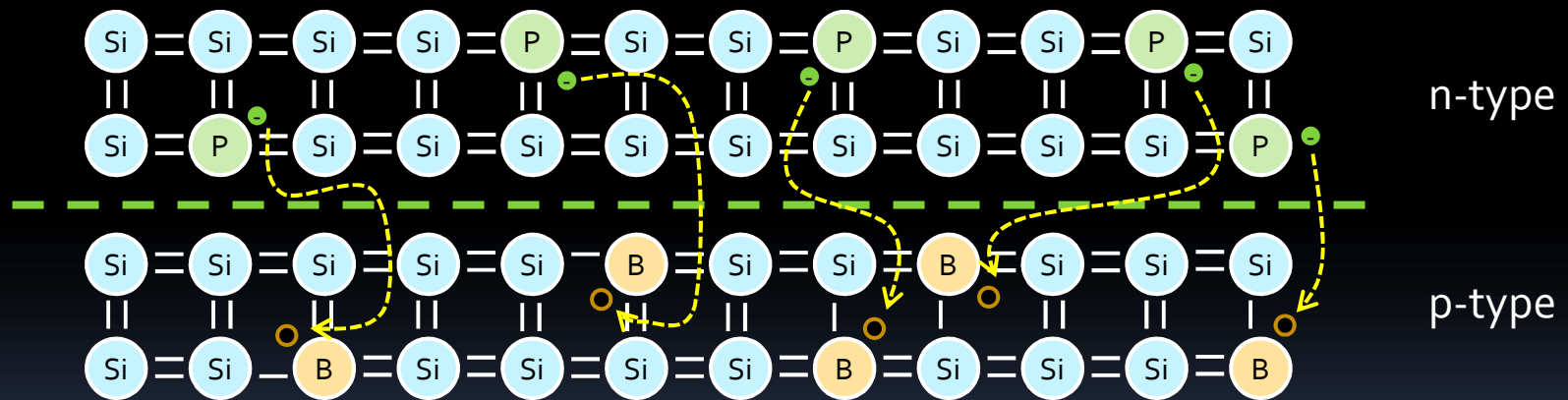
Impurities

- In the case of n-type semiconductors, the carriers are electrons that are not bound to the solid, and can flow more freely through the material.
- For p-type semiconductors, the carriers are called **holes**, to represent the electron gap as a particle as well.



Bringing p and n together

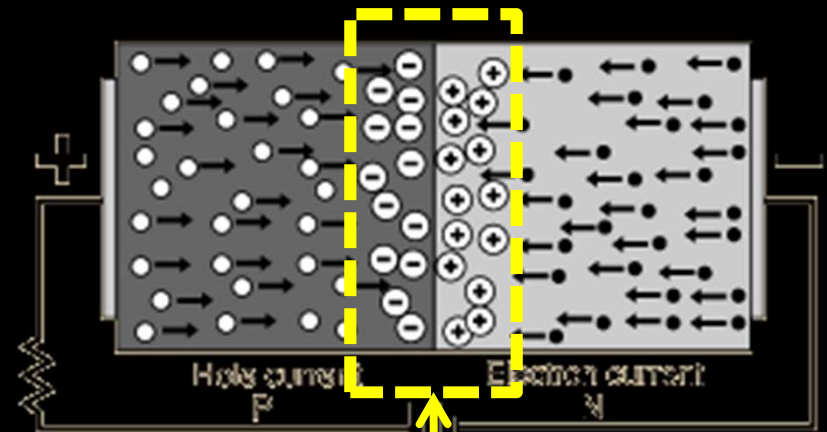
- What would happen if you brought some p-type material into contact with some n-type material?



- The electrons at the surface of the n-type material are drawn to the holes in the p-type.

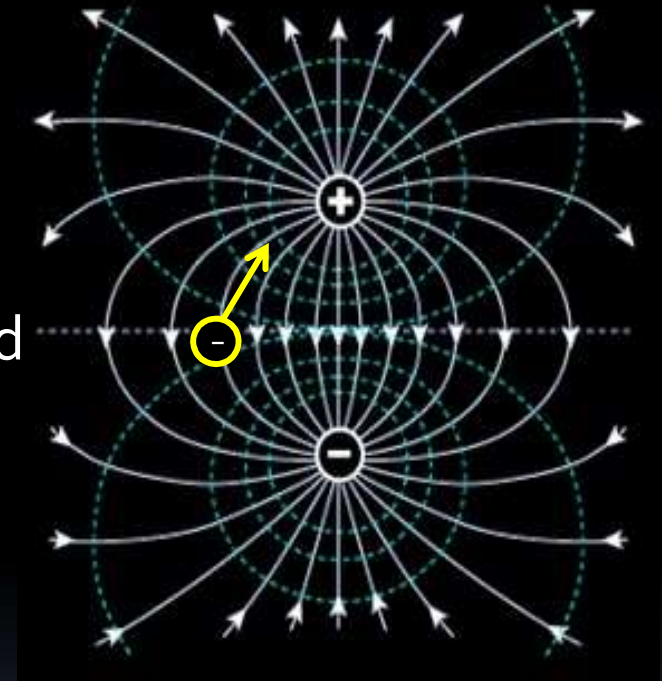
p-n Junctions

- When left alone, the electrons from the n section of the junction will mix with the holes of the p section, cancelling each other out, and creating a particle-free section called the **depletion layer**.
- Once this depletion layer is wide enough, the doping atoms that remain will create an **electric field** in that region.



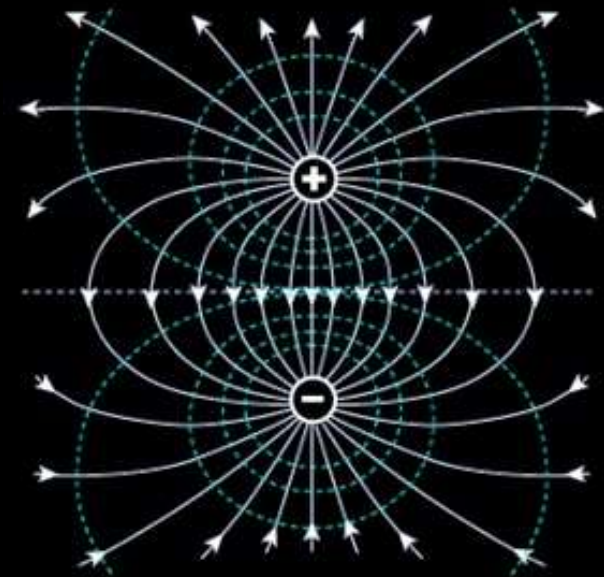
Electric fields

- What is an **electric field**?
 - An electric field is created when a charge difference exists between two regions.
 - Any electrons in the middle would be attracted to the positive side and repelled by the negative side.
- Example: depletion layer
 - When a phosphorus atom loses its electron, that atom develops an overall positive charge.
 - Similarly, when a boron atom takes on an extra electron, that atom develops an overall negative charge.
 - This creates **an electric field in the depletion layer**.



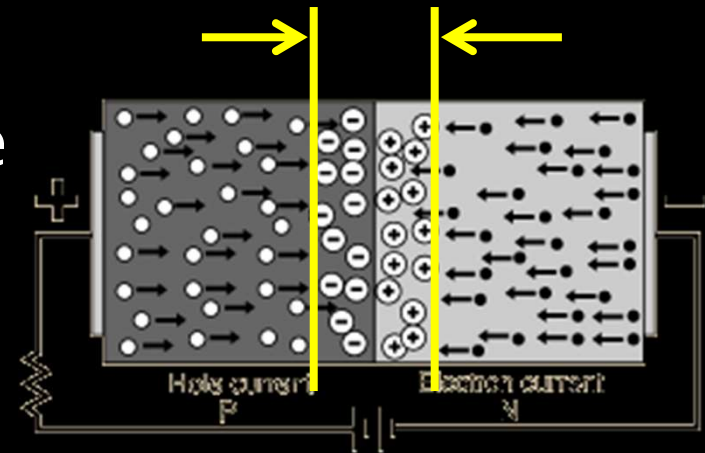
Electric fields

- A depletion layer is made up of many of these electrically imbalanced phosphorus and boron atoms.
- The electric field caused by these atoms will cause holes to flow back to the p section, and electrons to flow back to the n section.
 - The current caused by this electric field is called **drift**.
 - The current caused by the initial electron/hole recombination is called **diffusion**.
- At rest, these two currents reach **equilibrium**.



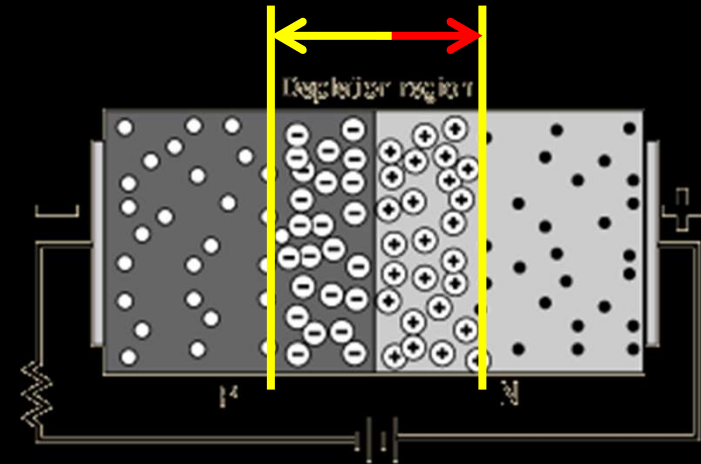
Forward Bias

- What happens when a voltage is applied to this junction?
 - It depends on the direction in which the voltage is applied.
- **Forward bias:**
 - When a positive voltage is applied to the p end of this junction, electrons are injected into the n-type section.
 - This narrows the depletion layer and increasing the electron diffusion rate.
- With a smaller depletion layer, the electrons travel more easily through to the p-type section, and back into the other terminal of the voltage source

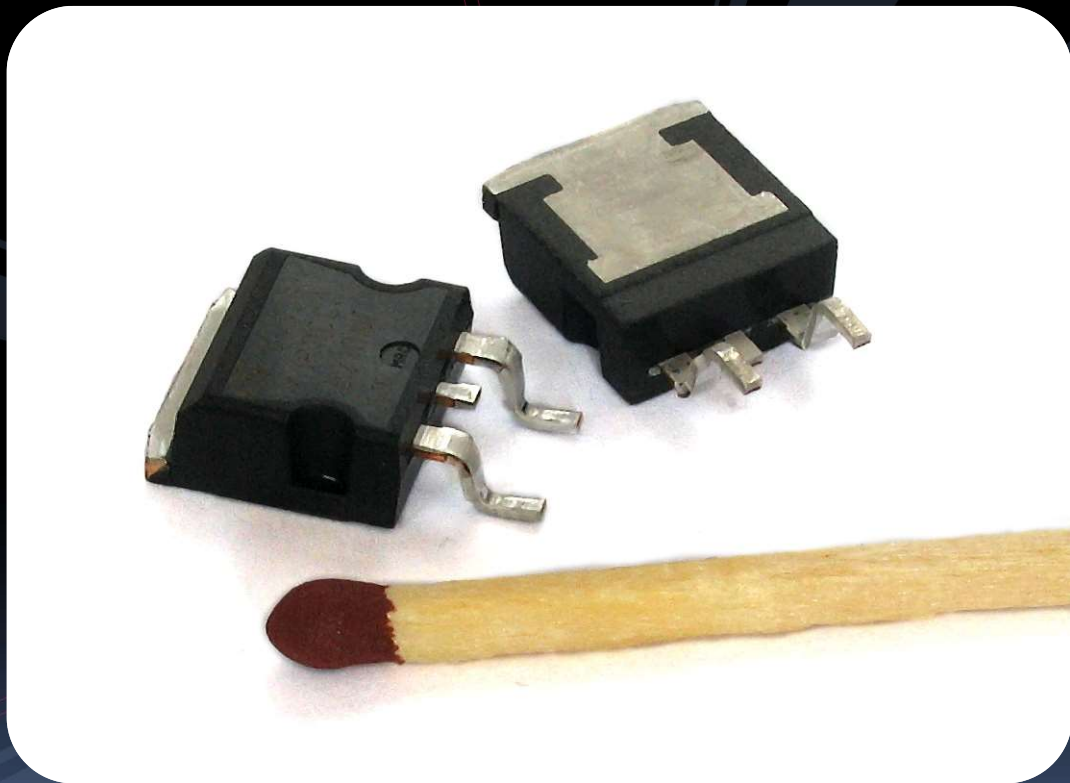


Reverse Bias

- **Reverse bias:**
 - When a positive voltage is applied to the n side of the junction, the depletion region at the junction becomes wider, preventing the carriers from passing.
 - a small current still flows through the circuit, but it is weak and does not increase with an increase in the applied voltage.
- So when a junction is forward biased, it becomes like a virtual **short-circuit**, and when the junction is reverse biased, it becomes like a virtual **open-circuit**.
 - This is the basis of **transistors**!



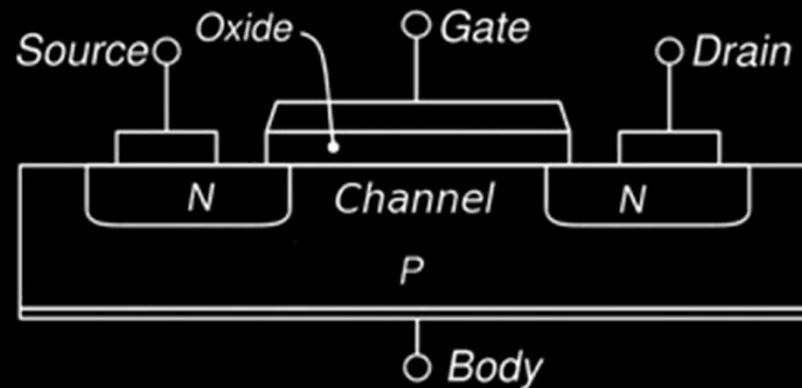
Transistors, finally



Creating transistors

- Transistors use the characteristics of p-n junctions to create more interesting behaviour.
- Three main types:
 - Bipolar Junction Transistors (BJTs)
 - Metal Oxide Semiconductor Field Effect Transistor (MOSFET)
 - Junction Field Effect Transistor (JFET)
- The last two are part of the same family, but we'll only look at the MOSFET for now.

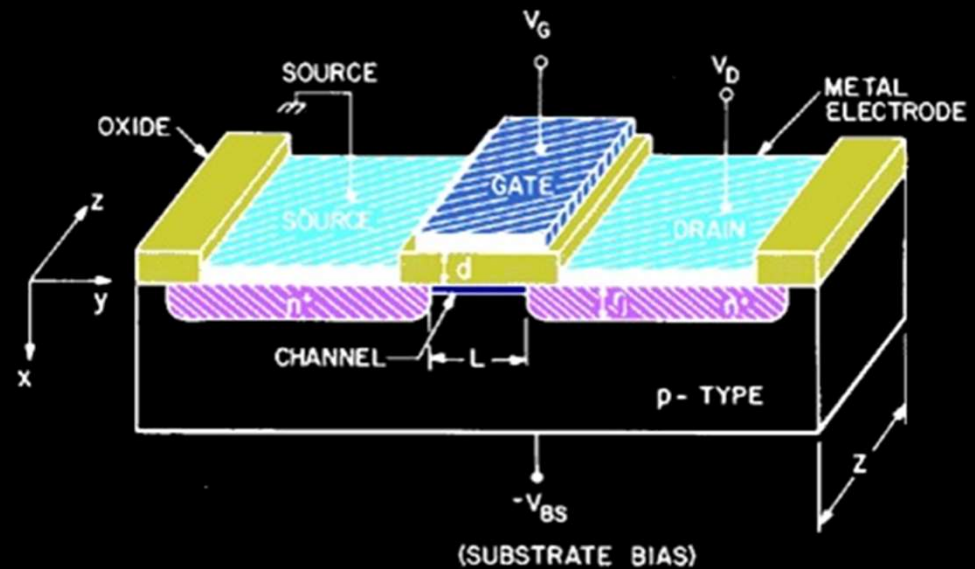
The MO of MOSFETs



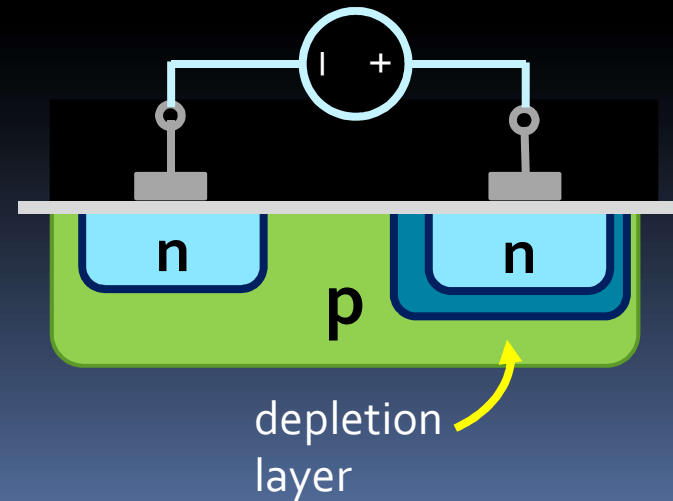
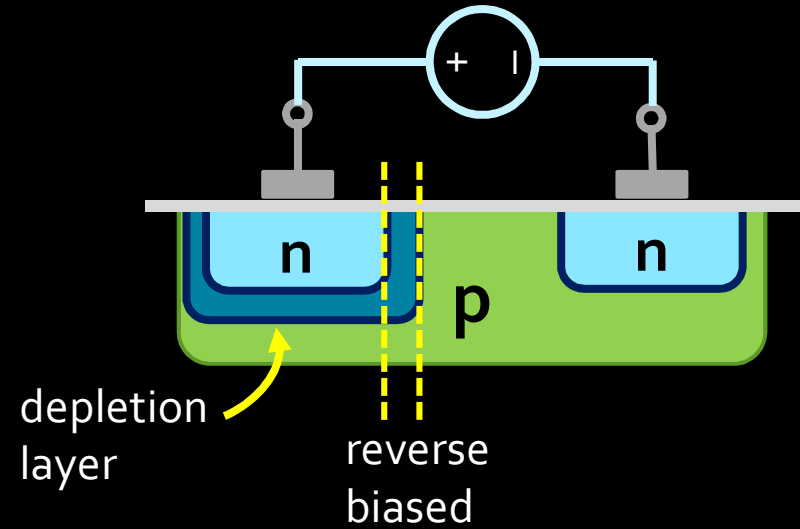
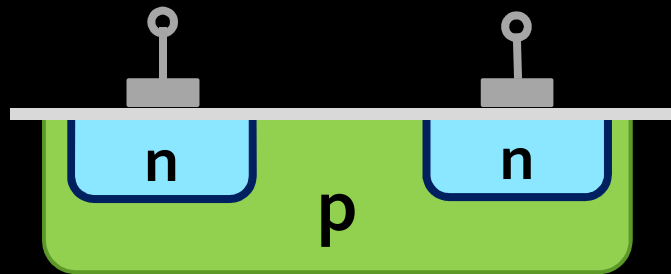
- Metal Oxide Semiconductor Field Effect Transistors are composed of a layer of semiconductor material, with two layers on top of the semiconductor:
 - An oxide layer that doesn't conduct electricity,
 - A metal layer (called the gate), that can have an electric charge applied to it
 - These are the M and O components of MOSFETs.

The S of MOSFETs

- The semiconductor sections are two pockets of n-type material, resting on a **substrate** layer of p-type material.
- A voltage is applied across the two n-type sections, called the **drain** and the **source**. No current will pass between them though, because the p section in between creates at least one reverse-biased pn junction.

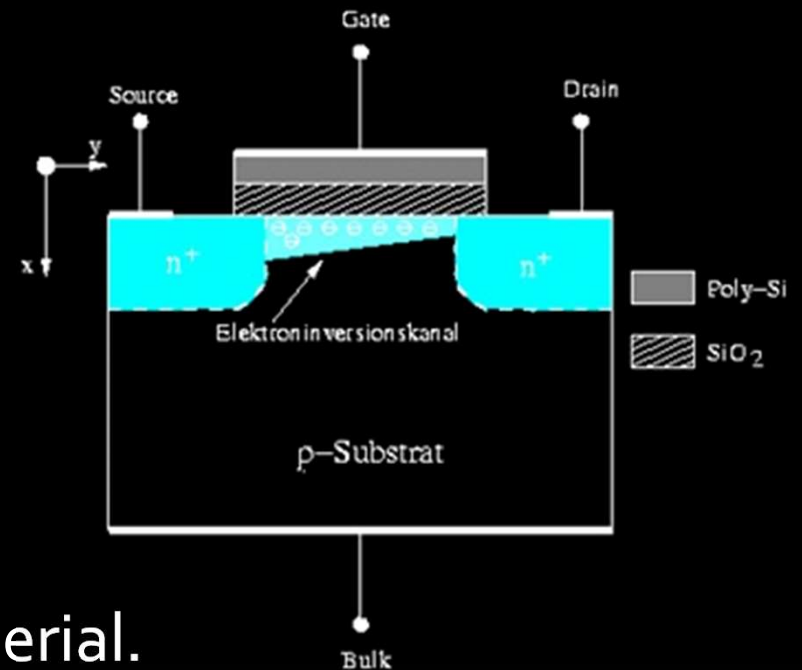


Applying voltage to NPN

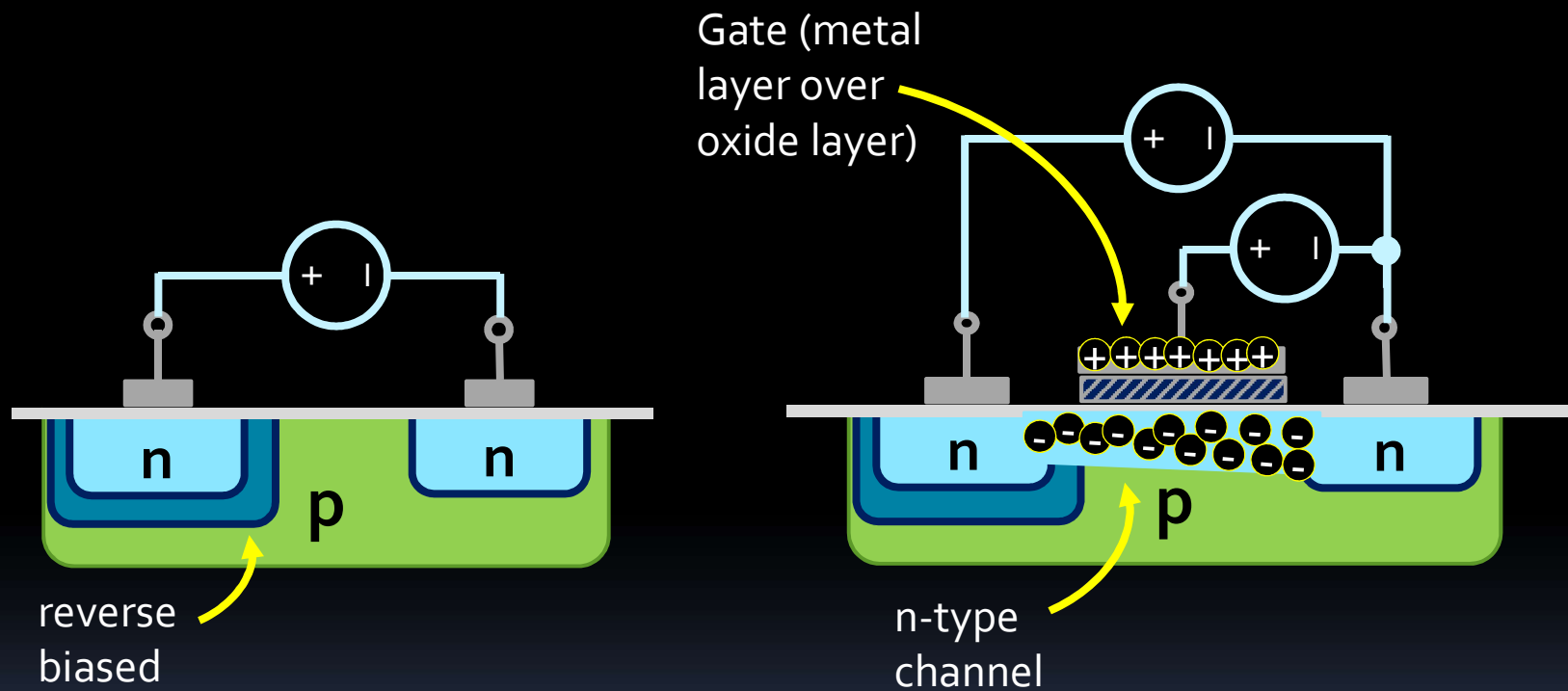


n-channel MOSFETs

- However, when a voltage is applied between the source and the metal plate (the **gate**), positive charges are built up in the metal layer, which attracts a layer of negative charge to the surface of the p-type material.
- This layer of electrons creates an n-type channel between the drain and the source, connecting the two and allowing current to flow between them.
 - the wider the channel, the higher the current



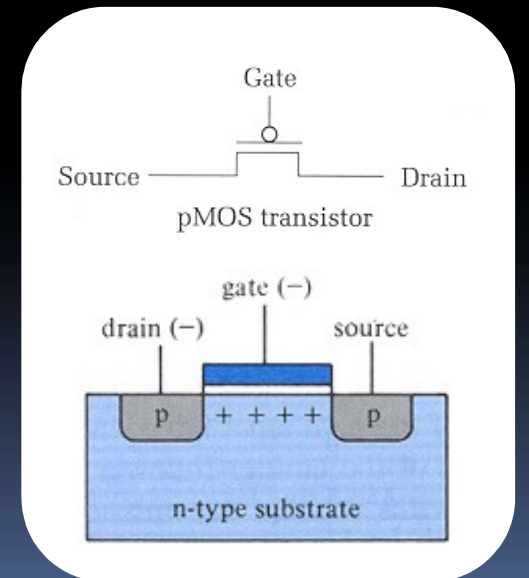
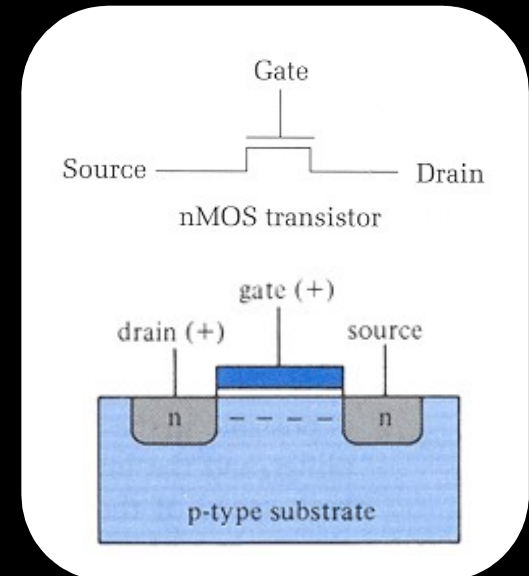
Applying voltage to NPN



n-type channel creates path between source and drain for current to conduct!

nMOS vs pMOS

- Two types of MOSFETs exist, based on the semiconductor type in the drain and source, and the channel formed.
 - **nMOS transistors** (the design described so far) conduct electricity when a positive voltage (5V) is applied to the gate.
 - **pMOS transistors** (indicated by a small circle above the gate) conduct electricity (i.e., act as a closed switch) when the gate voltage is logic-zero.



Transistors to Gates

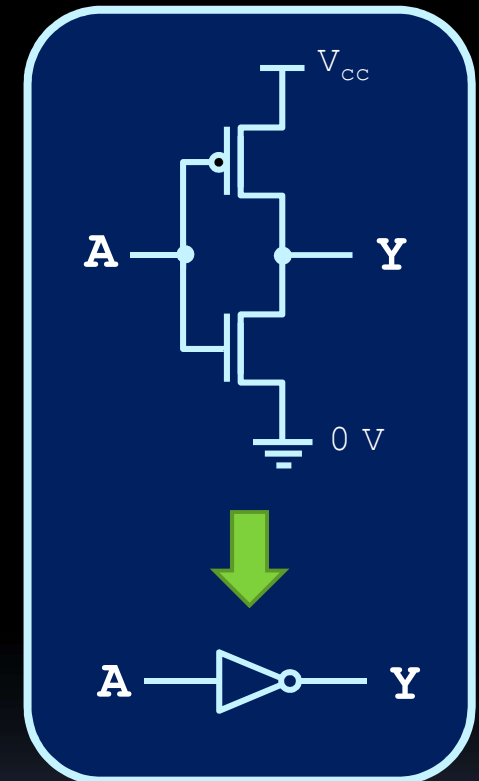
- MOSFETs can make current flow, based on the voltage values in the gate and drain.
 - i.e. the truth table on the right.
- One final step: combining MOSFETS to create high and low voltage outputs, based on high and low voltage inputs.
 - General approach: create transistor circuits that make current flow to outputs from high or low voltage, based on transistor input values.

MOSFET Truth Table

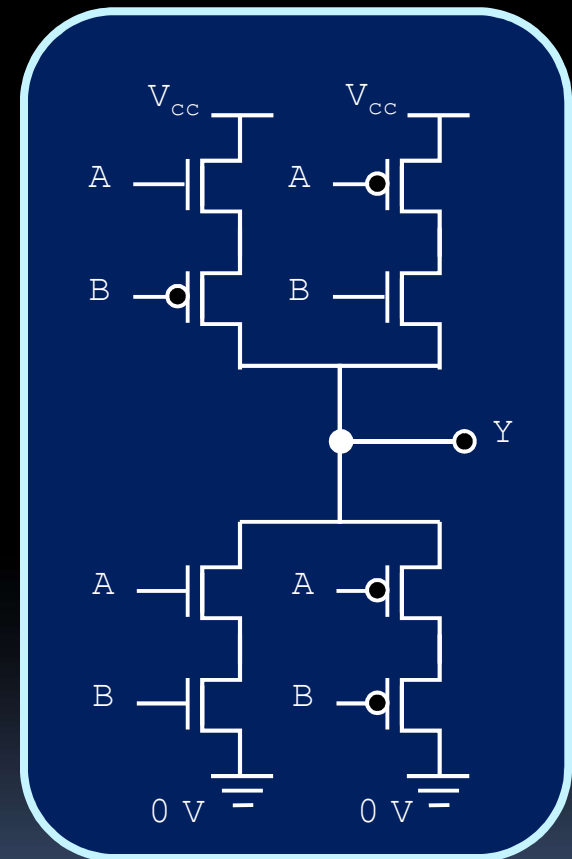
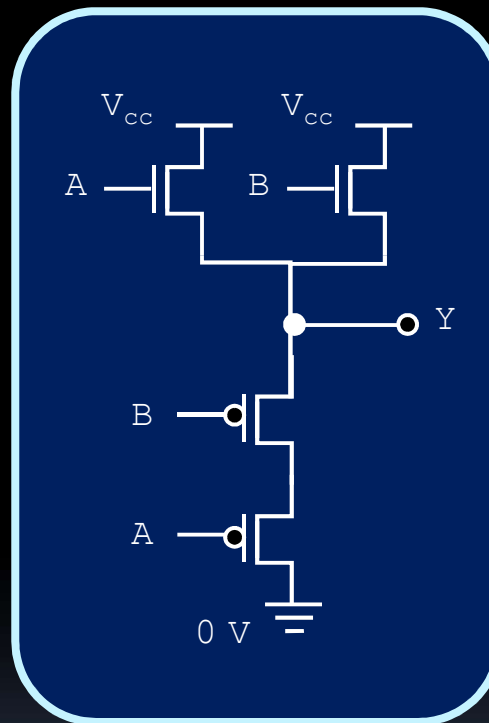
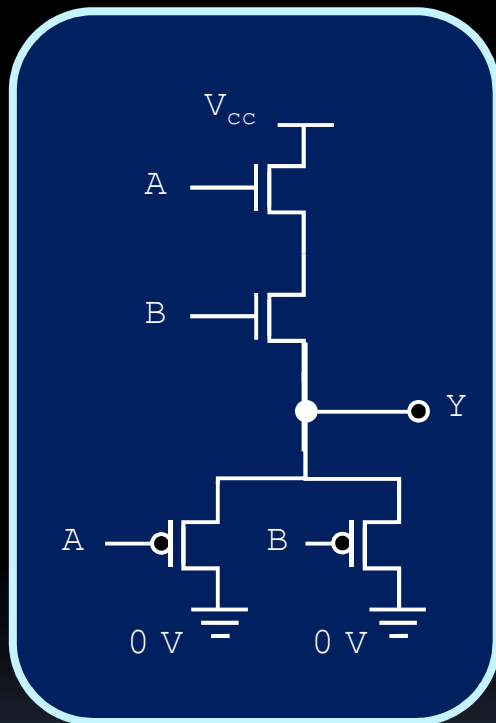
V_{DS}	V_{GS}	I_{DS}
Low	Low	Low
Low	High	Low
High	Low	Low
High	High	High

Making gates

- Since these transistors aren't simply on/off switches, digital logic gates (**AND, OR, NOT**) are created by a combination of transistors
 - Examples: NOT gate circuit in diagram.
- Physical data:
 - "High" input (aka V_{cc}) = 5V
 - "Low" input (aka Ground) = 0V
 - Switching time ≈ 120 picoseconds
 - Switching interval ≈ 10 ns
 - NAND is most common logic gate



Transistors into gates



Note: V_{CC} = "Common Collector Voltage"
= high voltage (5 V)

Transistor Fabrication

- Transistors are not formed by pushing large chunks of n- and p-type semiconductors together.
- Transistors are now made by bombarding silicon with doping substances to create the layers for each junction
 - Surface is oxidized in between stages to ensure that only the necessary sections are doped.



Fabrication Process

