Understanding the *p-n* Junction

by **Dr. Alistair Sproul**
Senior Lecturer in Photovoltaics
The Key Centre for Photovoltaic Engineering, UNSW

Dr. Alistair Sproul is a senior lecturer at the Key Centre for Photovoltaic Engineering, UNSW. He specialises in teaching Photovoltaics and Renewable Energy Engineering courses at the undergraduate and postgraduate levels. He is also a consultant to industry where he provides expert advice in the optimisation of renewable energy systems.

The *p-n* junction is the fundamental building block of the electronic age. Most electronic devices are made of silicon. By exploring the electrical properties of silicon, it is possible to acquire an understanding of the inner workings of the *p-n* junction.

Silicon

A single silicon atom consists of fourteen negatively charged electrons surrounding a nucleus of fourteen positively charged protons and fourteen electrically neutral neutrons. As there are an equal number of positive and negative charges, the silicon atom has no net electric charge. Of the fourteen electrons, only the four outer electrons (visualised as those on the outermost shell of the atom) are available for chemical bonding. The remaining 10 electrons are tightly bound to the nucleus and do not form bonds to other atoms.

In a crystal of silicon, each silicon atom bonds to four other silicon atoms. Each bond consists of two electrons, one electron from each of the silicon atoms involved in the bond. This type of bonding, where electrons are shared equally by the atoms involved, is called covalent bonding.

These four bonding electrons give silicon a very nice property: it can form a crystal involving all of the bonding electrons, with none left over. (Carbon, which also has four bonding electrons, can use all four of its electrons in covalent carbon-carbon bonds and form crystals of diamond.) A silicon crystal consists of atoms that have no net electric charge (having the same number of electrons as protons) and therefore the silicon crystal itself has no net electric charge.

Because of its electrical properties, silicon is called a semiconductor. If the silicon crystal is very pure, then the outer four electrons occupy all of the covalent bonds in the crystal. At temperatures near absolute zero the electrons remain in their bonding positions. In this case, a silicon crystal is nearly a perfect electrical insulator. If, however, sufficient energy is supplied to the crystal, by heat or light, it is possible to break some of these covalent bonds. In silicon, the bonding energy of the covalent bond is 1.1 electron volts. When a bond is supplied with this amount of energy (or more) an electron can break free of an atom and the electron is then free to move throughout the crystal.

An energy of 1.1 electron volts corresponds to a photon in the near infrared region of the electromagnetic spectrum. A photon with less than 1.1 electron volts of energy does not have enough energy to interact with a bond in the silicon crystal and will therefore pass right through the crystal without dislodging any electrons.

Excerpt from

Solar Cells: Resource for the Secondary Science Teacher

Available from the
UNSW Bookshop
http://www.bookshop.unsw.edu.au

Copyright UNSW ~ copying permitted for educational purposes.

IPR2025-00845
Apple-Sony EX1023 Page 1

Phosphorus

A single phosphorus atom also has no net electric charge. Phosphorus is the next element up from silicon in the periodic table with fifteen electrons and fifteen protons. Again, the negative electric charges from the electrons exactly counteract the positive electric charges from the protons resulting in an atom that has no net electric charge.

Of these fifteen electrons, only five of them (again visualised as those on the outermost shell of the atom) are available for chemical bonding. The configuration of these five available bonding electrons gives phosphorus its distinct chemical properties. For *p-n* junctions the important aspect is that phosphorus has one more electron available for bonding than silicon.

Substituting with Phosphorus

In a process called "doping", impurity atoms are placed in very pure silicon to change the electrical characteristics of the crystal. Phosphorus can replace some of the silicon atoms in the crystal and when this happens the phosphorus atom is called a dopant atom. The dopant phosphorus atoms each form four covalent bonds in the same way that a silicon atom does with its neighbours. The fifth of the phosphorus bonding electrons, the one that is not used in covalent bonding, now plays an important role. This electron is only weakly attached to the phosphorus atom: so weakly (0.045 electron volts bonding energy) that at normal temperatures the thermal energy within the crystal is sufficient to free it from the phosphorus atom. When this happens the resulting untethered electron is free to travel around the crystal.

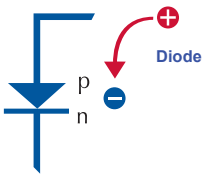
An electric field.

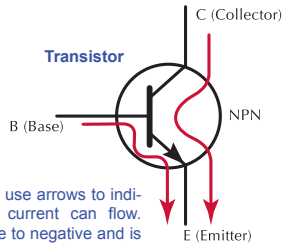
The positive point charge.

The direction of the electric field.

The Direction of the Electric Field

By convention, the direction of the electric field is the direction that a positive point charge would move if placed in an area of electric charge. If the positive point charge were placed near a positive charge then the like charges would cause the point charge to move away. The electric field is therefore pointing out of a positive charge. A negative charge has the electric field lines point into it.





The symbols for diodes and transistors use arrows to indicate the direction that conventional current can flow. Conventional current flows from positive to negative and is opposite to the direction that electron current flows. If a battery is connected to a diode, current will flow only when the positive terminal of the battery is connected to the start (the not pointed end) of the arrow.

Conventional current has the current going from positive to negative, whereas, in terms of electron flow the electrons go from negative to positive. Most of electronics and electrical engineering use conventional current, unless otherwise specified. This is an artefact from the 18th century when scientists did not know that electrons were the charge carriers in wires. They guessed the sign of the charge carriers and unfortunately got it wrong.

Copyright UNSW ~ copying permitted for educational purposes.

IPR2025-00845
Apple-Sony EX1023 Page 2

The electrical conductivity of a silicon crystal can be exquisitely controlled by carefully adjusting the number of dopant atoms in the crystal. In typical solar cell applications there is about 1 dopant atom for every 5,000,000 silicon atoms. When an atom like phosphorus, with more than four bonding electrons, is used to dope silicon, the resulting crystal material is called n-type silicon. This is because the untethered electrons available from the dopant atoms each have a negative electric charge.

The Charge

As stated above, when silicon is doped with phosphorus the phosphorus substitutes for silicon atoms in the crystal and forms four covalent bonds, just like silicon does with itself. The thermal energy in the crystal (just from being at room temperature) breaks the weakly bonded fifth electron in phosphorus away from the phosphorus atom and that electron is now available to move freely around the crystal.

A phosphorus atom that is substituted in the silicon crystal, however, cannot move because it is covalently bonded and is therefore fixed in place in the crystal. With only fourteen electrons (the weakly bound fifth electron is now drifting freely around the crystal) but fifteen protons, the stationary phosphorus atom now exhibits a positive electric charge.

The sum of all of the electric charges in the doped material, however, is zero (the crystal has no net electric charge) because all of the free electrons in the material exactly match the number of positive charges from the phosphorus atoms in the crystal.

If it were possible to remove or drain some of the free (mobile) electrons from the n-type crystal the loss of the negatively charged electrons would leave the crystal with a positive electric charge. (The concept is similar to removing electrons from the top terminal of a Van de Graaff electrostatic generator by spraying them onto a moving belt. By removing electrons from it, the top terminal acquires a positive charge.)

If some of the mobile electrons in an n-type silicon crystal were removed from the crystal, the crystal would now have a positive electric charge. As will be stated later, in forming a *p-n* junction some of the free electrons in the n-type crystal can drift away into the p-type material thus forming a local positive electric charge in a region in the n-type silicon crystal.

Substituting with Boron

Boron, with five protons and five electrons, can also substitute in a silicon crystal where a silicon atom would normally reside. Because a boron atom has only three electrons available in its bonding shell, only three covalent bonds can be formed between a boron atom and the silicon atoms in a crystal. This is because it is necessary to share two electrons, one from each atom, in order to form a covalent bond. Without a fourth electron in the bonding shell boron can form only three covalent bonds.

Copyright UNSW ~ copying permitted for educational purposes.

IPR2025-00845
Apple-Sony EX1023 Page 3

At temperatures near absolute zero this boron atom with its missing bond is stable. At room temperature, there is sufficient thermal energy to push a nearby electron into this vacancy. When this happens the atom that supplied the electron to the boron atom now has an electron vacancy that can be filled by an electron from another atom in the crystal. In this way, the vacancy or "hole" can move from atom to atom. This can be viewed as negatively charged electrons moving around filling the holes (thereby leaving a new hole), or alternatively, simply as positive charges moving through the material (as moving holes).

At room temperatures, a boron atom in the crystal has been forced to have one more electron than the number of protons in its nucleus, and, in this configuration, the boron atom has acquired a negative electric charge.

When an atom like boron (with fewer bonding electrons than silicon) is used to dope silicon, the resulting material is called p-type silicon. This is because these types of dopant atoms generate mobile holes in the crystal with each hole having a positive electric charge.

Motion of electrons and holes

A silicon crystal will have mobile electrons or holes depending upon the doping. As electrons and holes each have an electric charge, they are of course affected by electric fields. Another important effect is the random thermal motion of the electrons and holes. This type of motion is called diffusive flow and is important in understanding *p-n* junctions and solar cells.

Diffusion

If a teacher places a few drops of perfume at the front of a classroom it won't be long until students at the back of the room can detect its odour. Over a long period of time the classroom will eventually contain an even distribution of perfume molecules in the air. This is a purely random phenomenon and is not related to any local concentration of perfume in the air: perfume atoms will drift randomly until eventually there is an even distribution of perfume molecules in the classroom's air.

In an n-type phosphorus doped silicon crystal the free electrons will diffuse, like perfume in a classroom, throughout the crystal in a purely random fashion until there is an equal distribution of free electrons throughout the volume of the n-type silicon crystal. In a p-type boron doped silicon crystal the corresponding holes will become equally distributed throughout the p-type crystal's volume.

Forming a *p-n* Junction

Doping one side of a piece of silicon with boron (a p-type dopant) and the other side with phosphorus (an n-type dopant) forms a *p-n* junction. First, however, consider two separate pieces of silicon - one being n-type, the other being p-type (see Figure 1).

Copyright UNSW ~ copying permitted for educational purposes.

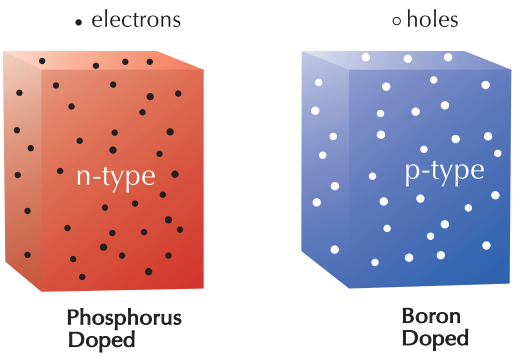


Figure 1 Doped silicon.

The n-type material has large numbers of free electrons (negatively charged) that can move through the material. The number of positively charged phosphorus atoms (called positive ions), which are not free to move, exactly balance the number and charge of these negative free electrons. Similarly, for the p-type material, there are large numbers of free holes (positively charged) that can move through the material. Their number and positive charge is exactly counter-balanced by the number of negatively charged boron atoms (called negative ions). Now imagine that the n-type and the p-type materials are brought together (see Figure 2).

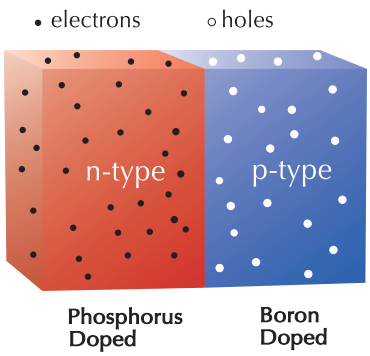


Figure 2 n-type and p-type materials brought together.

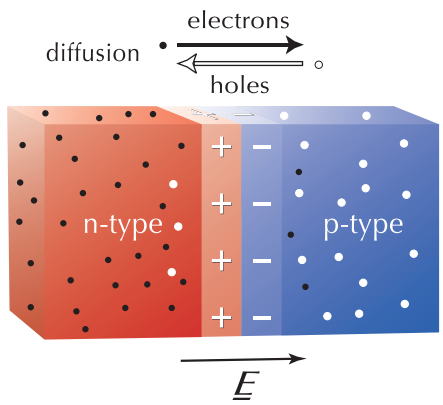
It is interesting to see what happens to the electrons and holes once these two pieces of silicon are joined. Due to the doping of the silicon crystal, there are large numbers of mobile electrons on the n-type side, but very few mobile electrons on the p-type side. Because of the random thermal motion of the free electrons, electrons from the n-type side start to diffuse into the p-type side. Similarly, due to the doping of the silicon, there are large numbers of mobile holes on the p-type side, but very few mobile holes on the n-type side. Holes in the p-type side, therefore, start to diffuse across into the n-type side.

Copyright UNSW ~ copying permitted for educational purposes.

IPR2025-00845
Apple-Sony EX1023 Page 5

Now, if the electrons and holes had no electric charge, this diffusion process would eventually result in the electrons and holes being uniformly distributed throughout the entire volume. They do, however, have an electric charge and this causes something interesting to happen!

As the electrons in the n-type material diffuse across towards the p-type side, they leave behind positively charged phosphorus ions, near the interface between the n and p regions. Similarly, the positive holes in the p-type region diffuse towards the n-type side and leave behind negatively charged boron ions (see Figure 3).



In a process called "diffusion", the random motions of mobile electrons and holes over time cause them to attempt to distribute equally within the total volume. As they cross the junction, the fixed ions they leave behind establish a "built-in" electric field at the junction.

Figure 3 Diffusion establishes “built-in” electric field.

These fixed ions set up an electric field right at the junction between the n-type and p-type material. This electric field points from the positively charged ions in the n-type material to the negatively charged ions in the p-type material. The free electrons and holes are influenced by this "built-in" electric field with the electrons being attracted towards the positive phosphorus ions and the holes being attracted towards the negative boron ions.

Thus, the “built-in” electric field causes some of the electrons and holes to flow in the opposite direction to the flow caused by diffusion (see Figure 4).

A mobile electron or hole near the "built-in" electric field will be attracted and swept back into its original volume. At the junction there are two effects occurring (1) diffusion with electrons moving from n-type to p-type and, (2) the "built in" electric field sweeping locally affected electrons back into the n-type volume. The holes are affected similarly but in opposite directions.

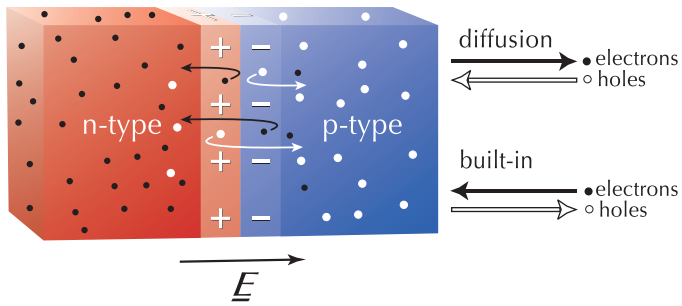


Figure 4 Motion of mobile electrons and holes due to diffusion and the “built-in” electric field.

Copyright UNSW ~ copying permitted for educational purposes.

These opposing flows eventually reach a stable equilibrium with the number of electrons flowing due to diffusion exactly balancing the number of electrons flowing back due to the electric field. The net flow of electrons across the junction is zero and the net flow of holes across the junction is also zero.

This begs the question, "If there is no net current flowing, of what use is it?" Although there is no net flow of current across the junction there has been established an electric field at the junction and it is this electric field that is the basis of the operation of diodes, transistors and solar cells.

Depletion Region

Within the depletion region, there are very few mobile electrons and holes. It is "depleted" of mobile charges, leaving only the fixed charges associated with the dopant atoms. As a result, the depletion region is highly resistive and now behaves as if it were pure crystalline silicon: as a nearly perfect insulator.

The resistance of the depletion region can be modified by "adding" an external electric field to the "built-in" electric field. If the "added" electric field is in the same direction as the "built-in" electric field, the depletion region's resistance will become greater. If the "added" electric field is opposite in direction to the "built-in" electric field, the depletion region's resistance will become smaller. The depletion region can therefore be considered to operate as a voltage-controlled resistor.

Forward Bias

If a positive voltage is applied to the p-type side and a negative voltage to the n-type side, current can flow (depending upon the magnitude of the applied voltage). This configuration is called "Forward Biased" (see Figure 5).

At the *p-n* junction, the "built-in" electric field and the applied electric field are in opposite directions. When these two fields add, the resultant field at the junction is smaller in magnitude than the magnitude of the original "built-in" electric field. This results in a thinner, less resistive depletion region. If the applied voltage is large enough, the depletion region's resistance becomes negligible. In silicon, this occurs at about 0.6 volts forward bias. From 0 to 0.6 volts, there is still considerable resistance due to the depletion region. Above 0.6 volts, the depletion region's resistance is very small and current flows virtually unimpeded.

Copyright UNSW ~ copying permitted for educational purposes.

IPR2025-00845
Apple-Sony EX1023 Page 7

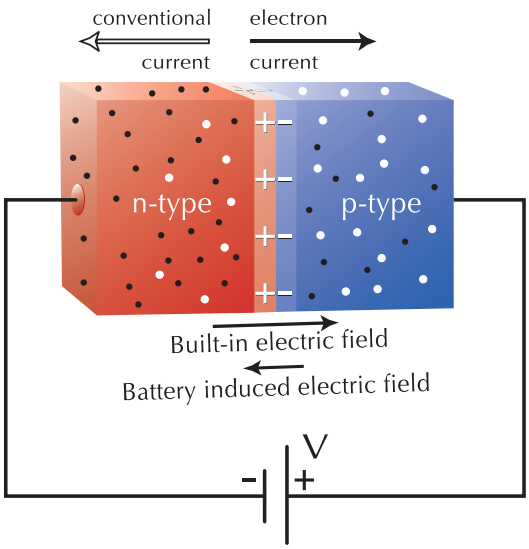


Figure 5 Forward bias of the *p-n* junction.

Reverse Bias

If a negative voltage is applied to the p-type side and a positive voltage to the n-type side, no (or exceptionally small) current flows. This configuration is called "Reverse Biased" (see Figure 6).

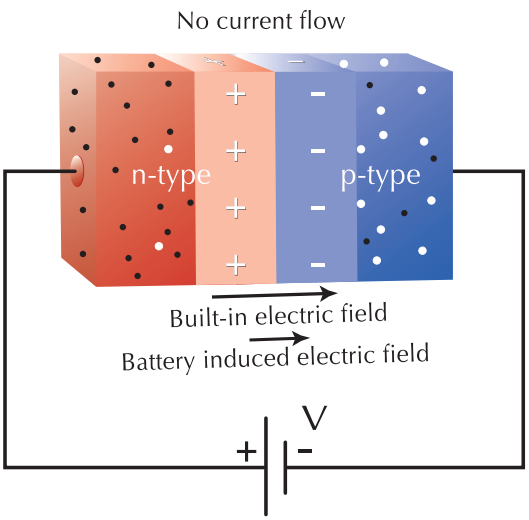


Figure 6 Reverse bias of the *p-n* Junction.

Copyright UNSW ~ copying permitted for educational purposes.

At the *p-n* junction, the "built-in" electric field and the applied electric field are in the same direction. When these two fields add, the resultant larger electric field is in the same direction as the "built in" electric field and this creates a thicker, more resistive depletion region. If the applied voltage becomes larger, the depletion region becomes thicker and more resistive.

In reality, some current will still flow through this resistance, but the resistance is so high that the current may be considered to be zero. As the applied reverse bias voltage becomes larger, the current flow will saturate at a constant but very small value. This is approximately 10^{-12} amperes per cm^2 of *p-n* junction area.

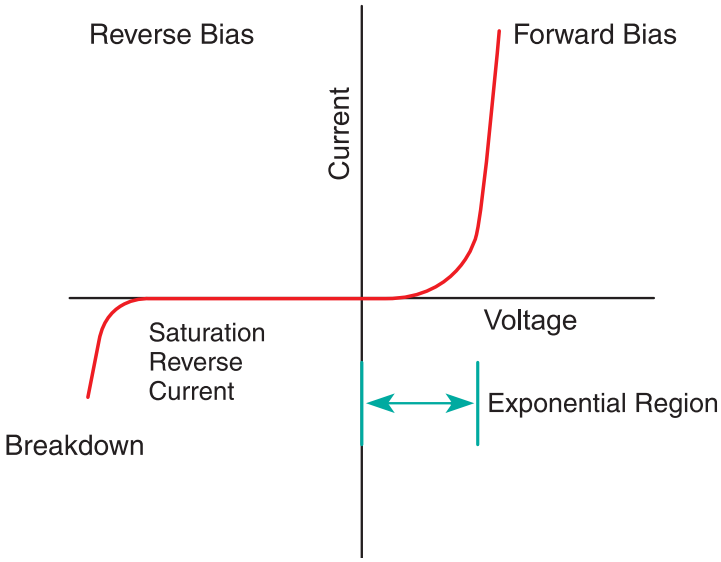


Figure 7 IV curve of the silicon *p-n* junction diode.

The current-voltage relation or IV curve of a silicon *p-n* junction is shown in Figure 7. In reverse bias, the *p-n* junction exhibits extreme electrical resistance and only a very small current flows. If the reverse bias voltage becomes too large then the junction will breakdown and current will flow. It is possible to design silicon *p-n* junctions in such a way that the breakdown voltage is at a specific desired value. Such *p-n* junctions are called zener diodes and are used as voltage references or overvoltage protectors in electrical cir-

In forward bias, the *p-n* junction exhibits an exponential lowering of resistance with applied voltage. From 0 to 0.5 volts, the silicon *p-n* junction is quite resistive. When the applied voltage approaches 0.6 volts, the exponential nature of the junction causes the resistance to drop dramatically. The silicon *p-n* junction diode appears to be an electrical switch that "turns on" when 0.6 volts is applied. Because electrical current flows only when it is forward biased, the diode appears to be an electrical one-way valve.

Copyright UNSW ~ copying permitted for educational purposes.

How a Solar Cell Works

Now it can be seen what happens inside a solar cell. Most solar cells are essentially large area *p-n* junctions. When light shines on them they can generate current and voltage. The reason this can happen is because of the "built-in" electric field at the junction of the p-type and n-type material.

First consider what happens if a silicon solar cell (which is a *p-n* junction) has a low resistance wire connected externally between the p and n contact. In the dark a solar cell will produce no current. If, however, light shines on the solar cell then current will flow through the wire, from the p-type side to the n-type side (conventional current).

How It Happens

The light has enough energy to break some of the chemical bonds of the silicon crystal. What that means is that electrons, which are normally involved in a silicon bond, are excited by the light into a higher energy state and the bond is broken. Not all of the bonds are broken otherwise the silicon would melt! The sun's light intensity on earth is strong enough to break about 1 bond for every 100 million silicon atoms in the solar cell. So the solar cell doesn't melt under normal conditions.

The excited electrons are now like the electrons from phosphorus dopant atoms - they are free to move through the material. Similarly, the broken bonds created by the light act as holes - just like the missing electrons in bonds between silicon and boron atoms - and these holes are also free to move throughout the material. What goes up, however, must eventually come down! Electrons and holes created in this way are physically near each other: for every electron excited by the light there is a corresponding hole generated. These electrons and holes can remain excited only for a short period of time. In a process called recombination, excited electrons stray too close to holes and the two fall back into bonded positions. When this happens the pair's electrical energy is lost as heat. If there is too much recombination, the solar cell won't work very well.

Current flow from a Solar Cell

The electrons and holes excited by light shining on the solar cell occur throughout the volume of the material, in the p-region, the n-region and at the junction region where there is a “built-in” electric field. It is easiest to understand the flow of current by first considering electrons and holes that are excited by the light in the junction region.

Because of the “built-in” electric field, the electrons are attracted towards the positive charge on the n-type material side. Similarly, the holes are attracted to the negative charge on the p-type material side. This separation of charges causes a current to flow across the junction. The direction of the current flow (conventional current) is the same as the motion of the holes (as they are positively charged). That is, the current flows across the junction from the n-type side to the p-type side.

Copyright UNSW ~ copying permitted for educational purposes.

Similarly, electrons and holes, which are far away from the junction, can be separated by the “built-in” electric field if they can find their way towards the junction before they recombine. This occurs simply because the electrons and holes will randomly diffuse throughout the material. If they manage to randomly wander into the “built-in” electric field region, before recombining, then they will contribute to the current flowing across the junction.

These charges will also cause flow of electrons through the external wire. Electric current can be measured by simply attaching a current meter in series with a solar cell. Because the resistance across the wire and the current meter is very low, there will be essentially no voltage across the solar cell, but current can flow. This is called the short-circuit current.

The Voltage of a Solar Cell

In the situation described above there is current flowing through the wire and meter but there isn't any voltage (yet!). Power can be extracted from the solar cell only if it can produce both voltage and current. What happens if the wire is cut?

In this case, there is no pathway for any light generated current to flow outside of the solar cell. Separation of electrons and holes, however, keeps happening inside the solar cell. Electrons are being pushed over toward the n-type side and holes are being pushed over towards the p-type side (if they manage to wander near the junction without first recombining). When the wire is cut, there is no external circuit for these charges to flow through. Instead, the flow of electrons into the n-type side and the flow of holes into the p-type side, causes a voltage to build up across the solar cell - positive voltage on the p-type side and negative voltage on the n-type side. There is now a voltage across the solar cell, however, there is no current flowing from the solar cell. The "open-circuit voltage" is measured by placing a voltmeter across the illuminated solar cell. For a silicon solar cell, it will measure about 0.6 volts in strong sunlight.

A Closer Look at the Open-Circuit voltage

If a *p-n* junction has an external forward bias voltage applied across it, a current will flow through the junction of the solar cell from the p-type side to the n-type side (conventional current).

When an open circuit solar cell is illuminated, the light generated electrons and holes that wander near the junction will be separated by the “built-in” electric field. The electrons will be pushed into the n-type region and the holes will be pushed into the p-type region. A light generated flow of current through the junction is opposite in direction to a forward bias flow of current.

In an illuminated open circuit solar cell there is a constant value of light generated electrons and holes flowing across the junction and into the corresponding p-type and n-type regions. The light generated current depends only on how intense the sunlight is and how much recombination occurs within the solar cell. This collection of charge creates a voltage across the solar cell. It could be imagined that this separation of charge would continue indefinitely with the result being that the voltage across the solar cell becomes infinite.

Copyright UNSW ~ copying permitted for educational purposes.

IPR2025-00845
Apple-Sony EX1023 Page 11

Something happens, however, to limit the voltage. The solar cell is now forward biased and just as for the *p-n* junction in the dark, current flows across the junction from the p-type side to the n-type side. This forward bias current is in the opposite direction to the light generated current described above. Both currents are flowing simultaneously. As the voltage across the solar cell builds up, it eventually reaches a point where the forward bias current exactly balances the light generated current. These two internal currents flow inside the solar cell cancelling each other out. The voltage at which this occurs is called the "open-circuit voltage". For a silicon solar cell it is typically 0.6 volts in bright sunshine.

Getting Power from a Solar Cell

For short-circuit conditions there is a low resistance wire connected between the solar cell contacts resulting in a current flowing but no generated voltage. If the wire is cut, the solar cell is now open-circuited resulting in a voltage being generated but no current flowing.

In order to get useful power out of a solar cell there needs to be a resistance between the two contacts that has just the right value. If the resistance is too low, then there will be mostly current coming from the solar cell and very little voltage generated. If too high a resistance is connected across the contacts there will be mostly voltage generated and very little current coming out of the solar cell.

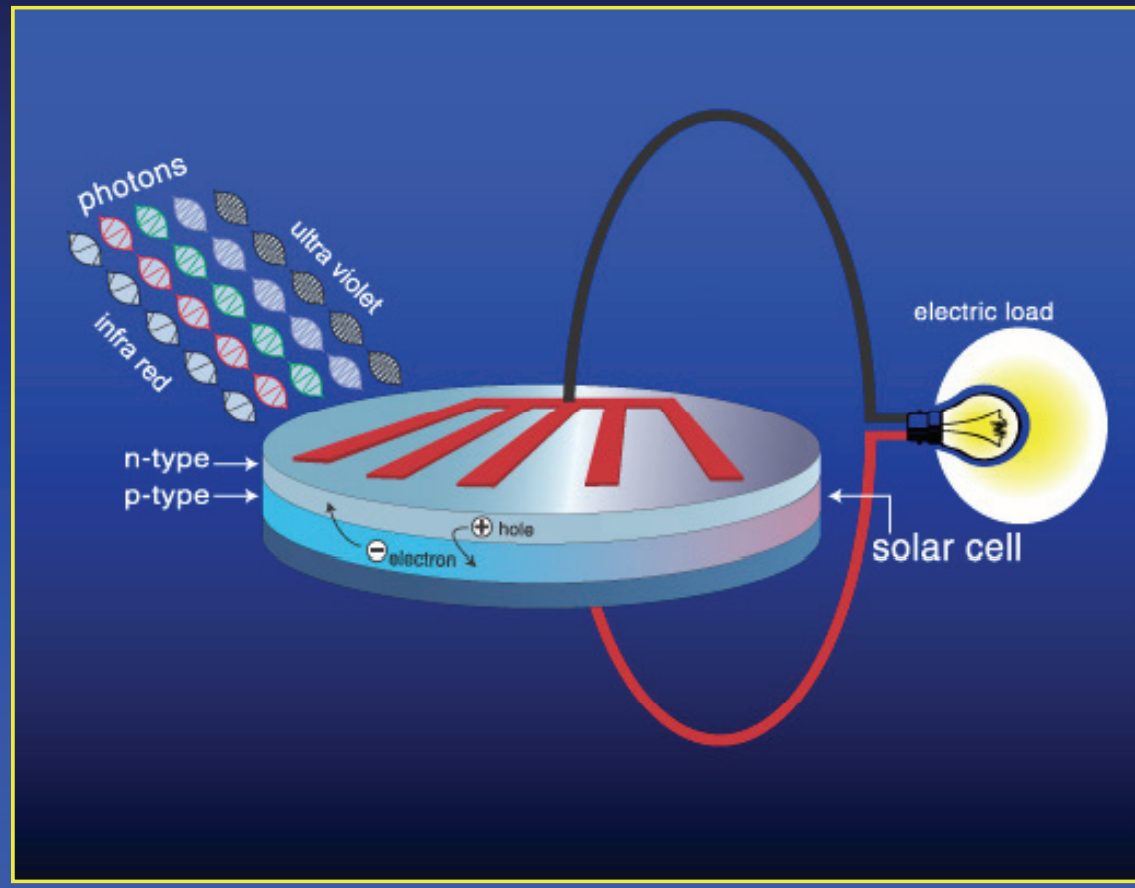
Only when the right resistance is connected across the solar cell will there be an optimal voltage generated with an optimal current flowing. With this value of resistance across the solar cell, the maximum electrical energy available from the solar cell is being delivered to the resistor. Under these conditions, the solar cell is said to be operating at its "Maximum Power Point".

For Large Amounts of Power

In order to generate large amounts of power from solar cells all that is needed is to wire many solar cells together. Collectively, the cells can generate enough power to do useful work. In fact, it is possible to wire together enough solar cells to form a great big solar cell power station. The electricity generated could then be used to power anything from a single light bulb to a large electric train, and with even more solar cells in the circuit, it would be possible to power an entire city!

Copyright UNSW ~ copying permitted for educational purposes.

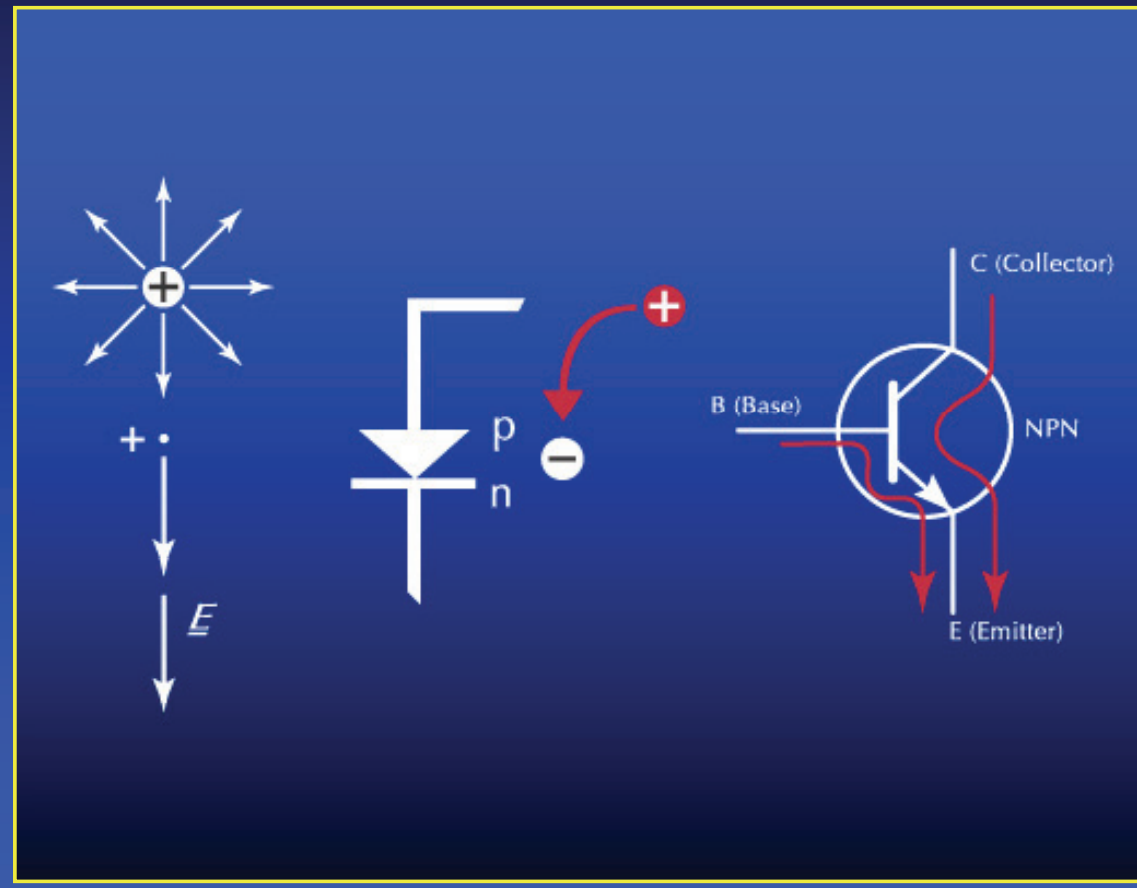
The Silicon Solar Cell



UNSW

Photovoltaics - Electricity from Sunlight

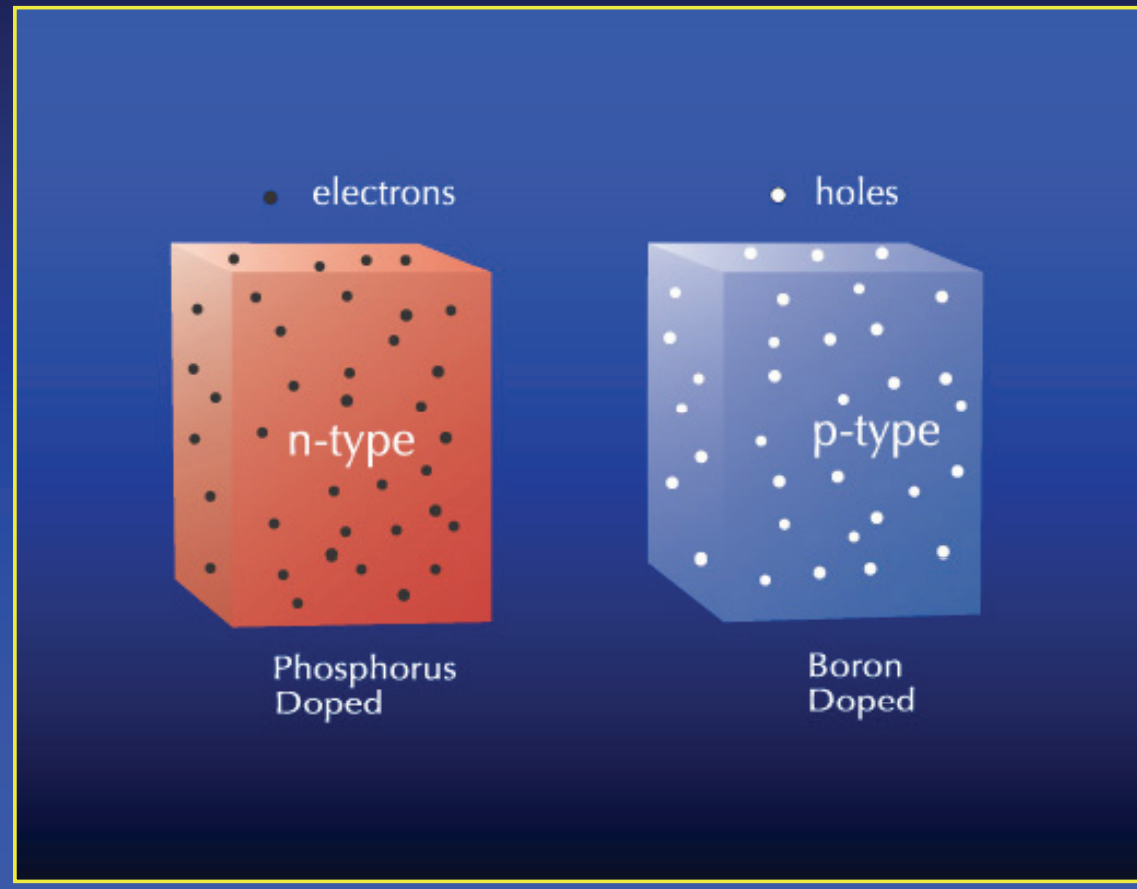
Direction of Current Flow



UNSW

Photovoltaics - Electricity from Sunlight

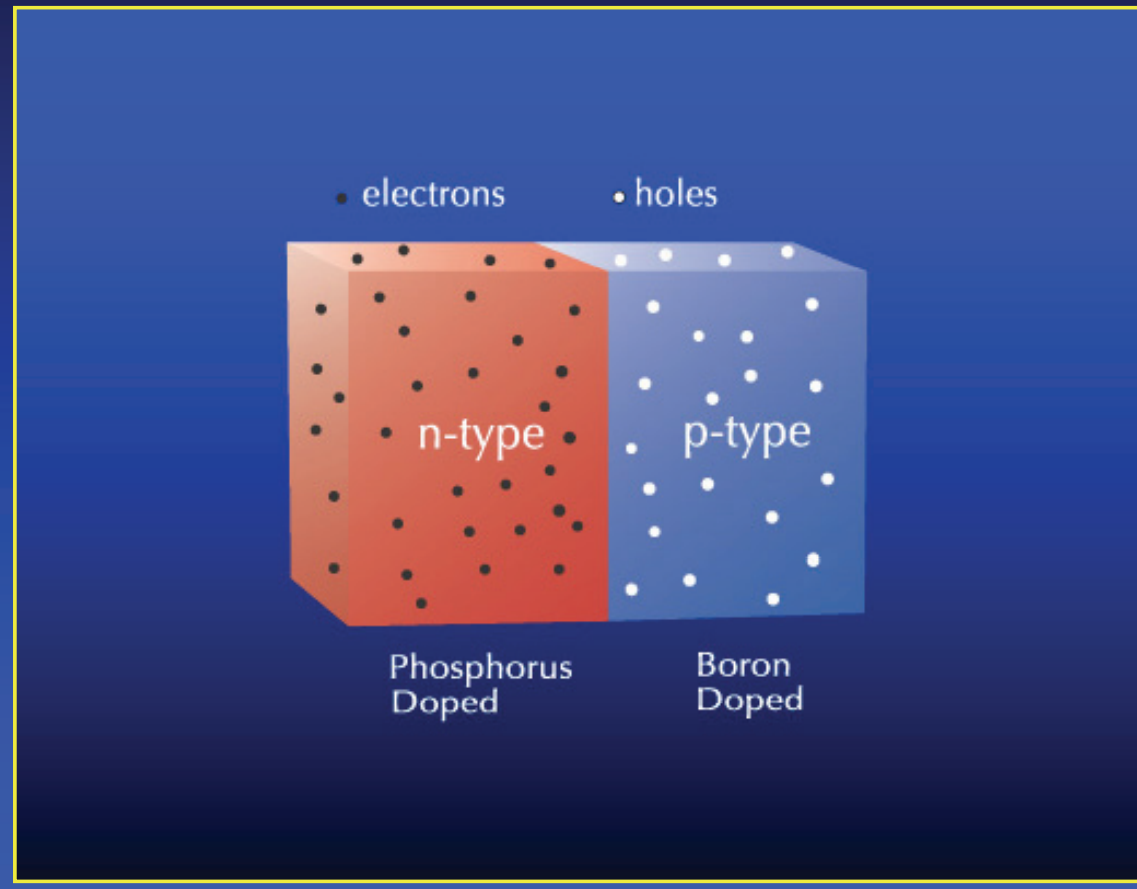
Two Pieces of Doped Silicon



UNSW

Photovoltaics - Electricity from Sunlight

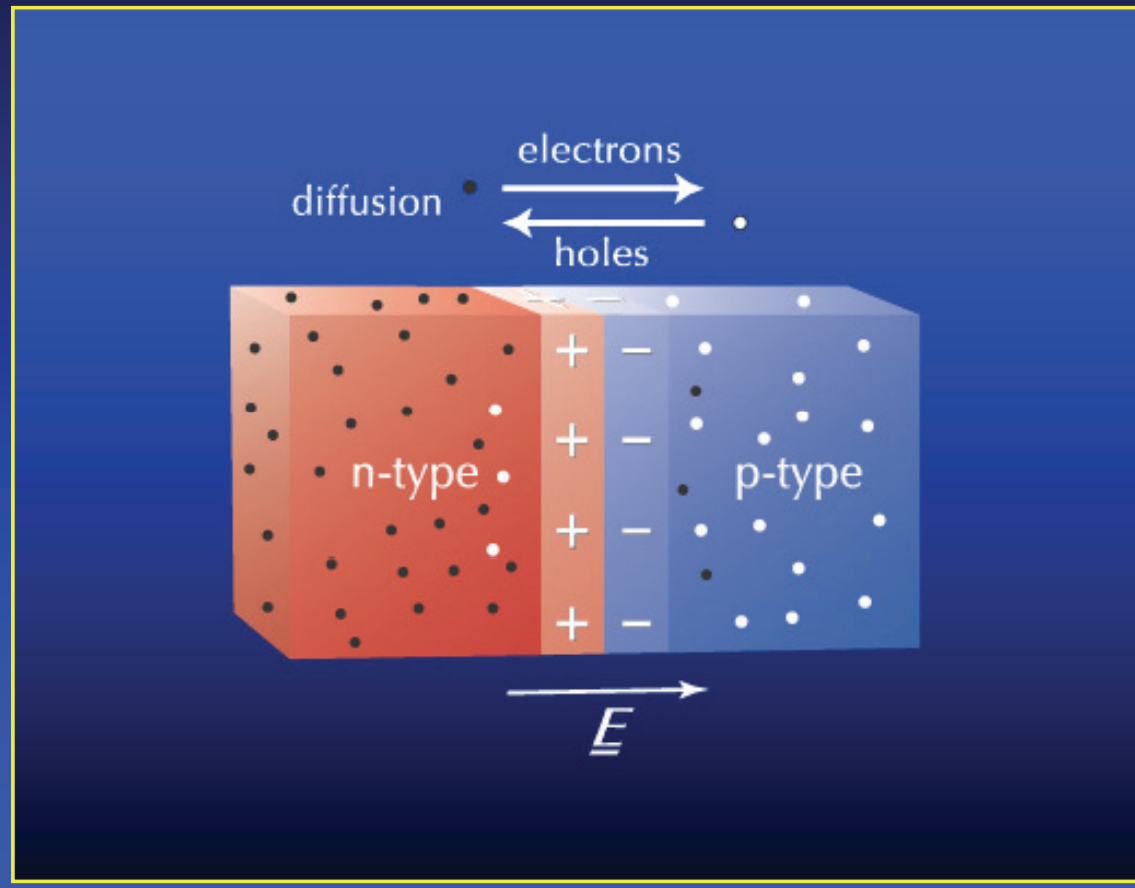
n-Type and *p*-Type Joined



UNSW

Photovoltaics - Electricity from Sunlight

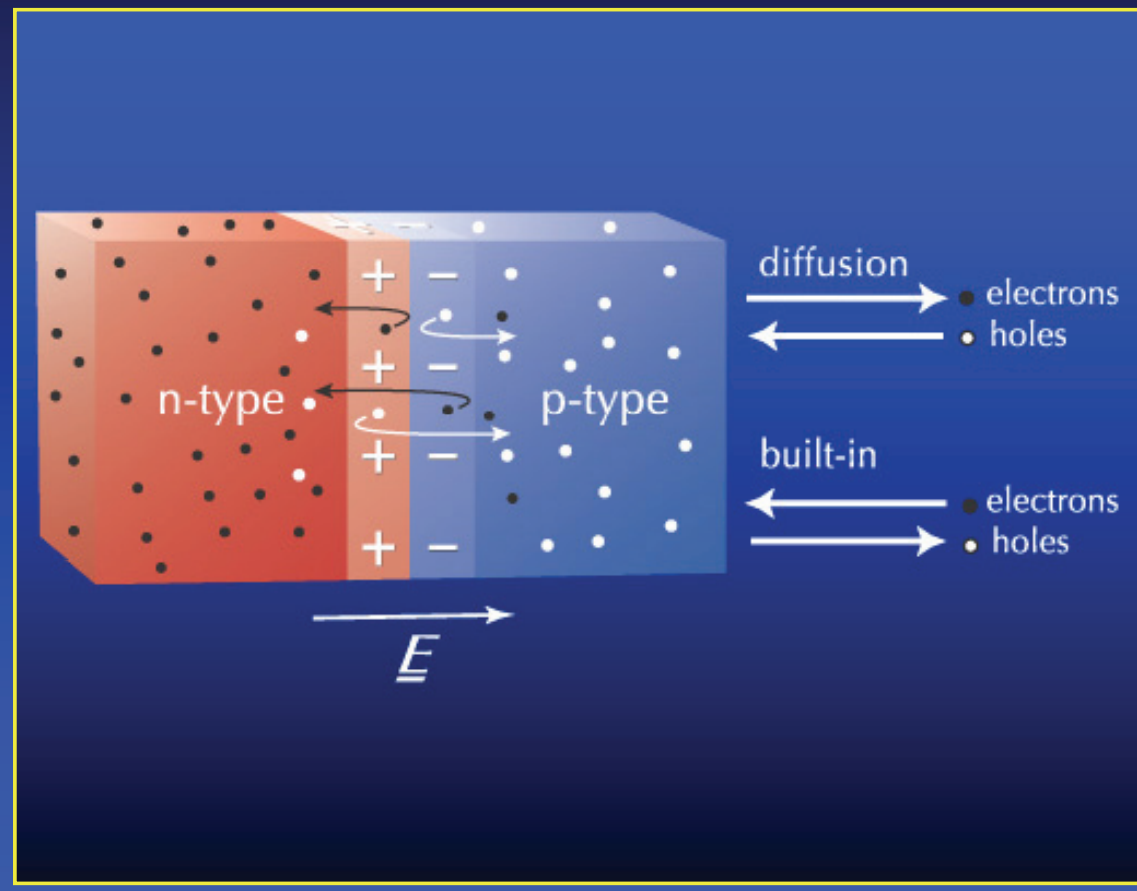
Diffusion via Random Motion



UNSW

Photovoltaics - Electricity from Sunlight

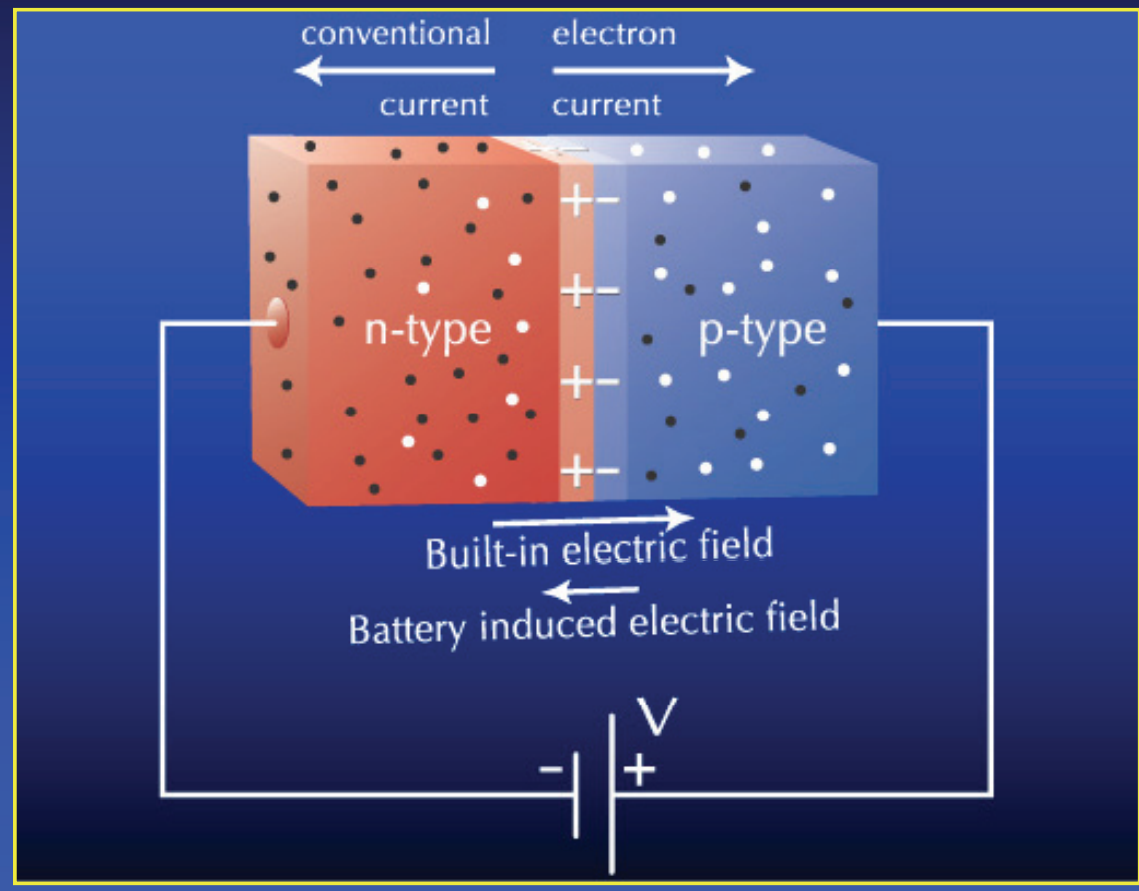
Diffusion & Built-In Currents



UNSW

Photovoltaics - Electricity from Sunlight

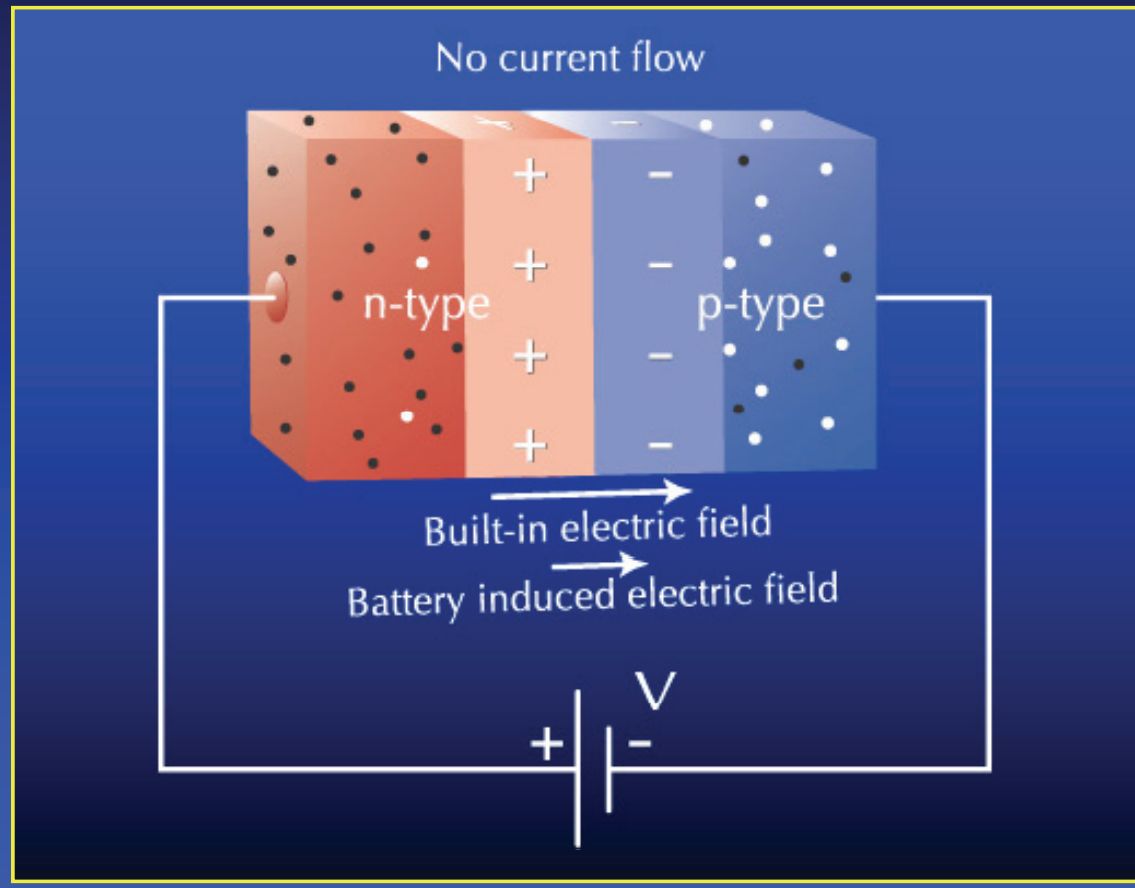
Forward Biased *pn*-Junction



UNSW

Photovoltaics - Electricity from Sunlight

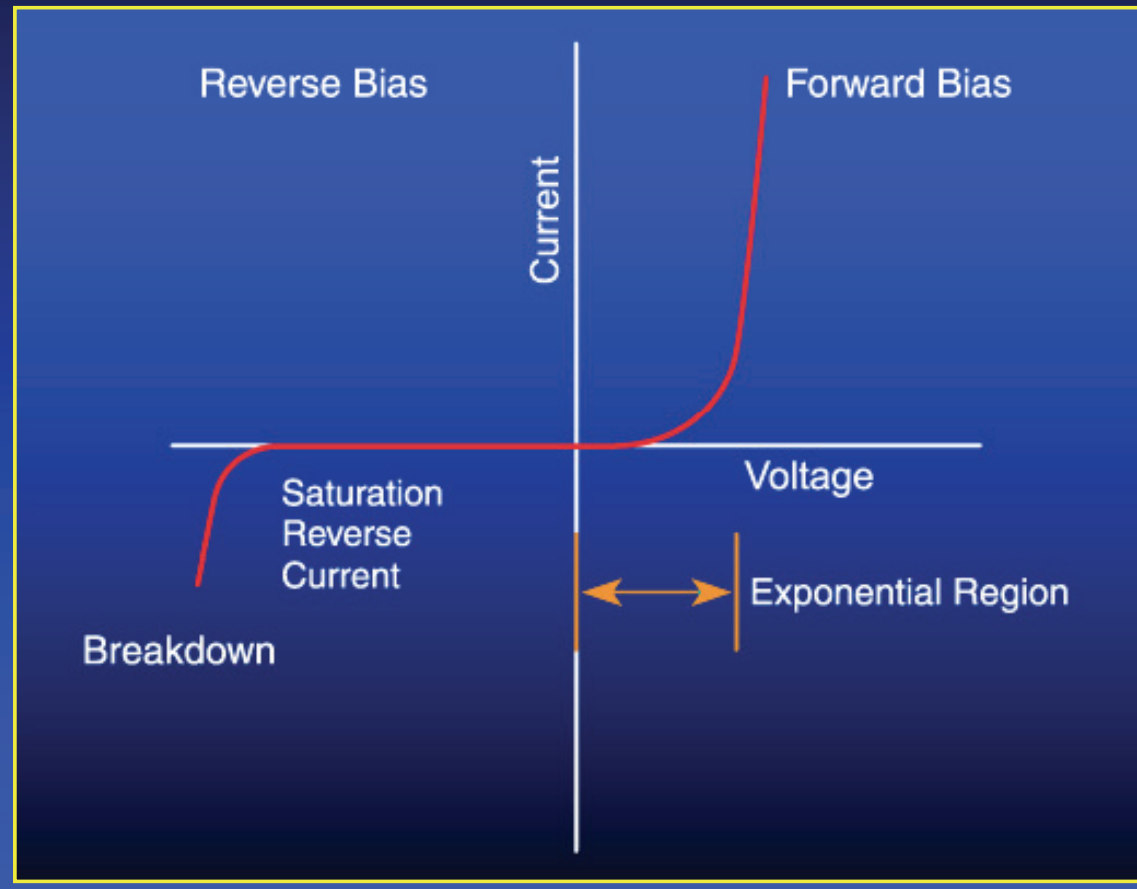
Reverse Biased *pn*-Junction



UNSW

Photovoltaics - Electricity from Sunlight

Characteristics of the *pn*-Junction



UNSW

Photovoltaics - Electricity from Sunlight