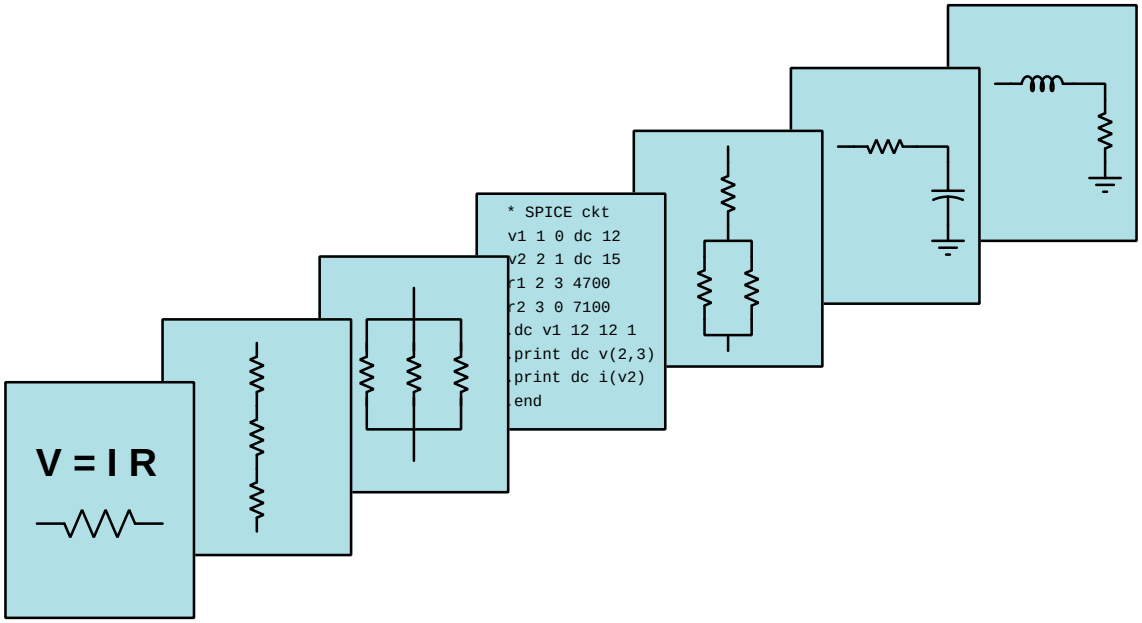
Modular Electronics Learning (ModEL) project



Conductors, Insulators, and Semiconductors

© 2020-2025 by Tony R. Kuphaldt – under the terms and conditions of the Creative Commons Attribution 4.0 International Public License

Last update = 22 October 2025

This is a copyrighted work, but licensed under the Creative Commons Attribution 4.0 International Public License. A copy of this license is found in the last Appendix of this document. Alternatively, you may visit `http://creativecommons.org/licenses/by/4.0/` or send a letter to Creative Commons: 171 Second Street, Suite 300, San Francisco, California, 94105, USA. The terms and conditions of this license allow for free copying, distribution, and/or modification of all licensed works by the general public.

Contents

1 Introduction **3**
- 1.1 Recommendations for students . . . 3
- 1.2 Challenging concepts related to semiconductors . . . 5
- 1.3 Recommendations for instructors . . . 6

2 Tutorial **7**
- 2.1 Atomic structure . . . 8
- 2.2 Energy bands and conductivity . . . 11
- 2.3 Energy bands and electric fields . . . 13
- 2.4 Doping . . . 16
- 2.5 The PN junction . . . 19
- 2.6 Semiconductor production . . . 25

3 Animations **29**
- 3.1 Animation of a forward-biased PN diode junction . . . 30

4 Questions **55**
- 4.1 Conceptual reasoning . . . 59
  - 4.1.1 Reading outline and reflections . . . 59
  - 4.1.2 Foundational concepts . . . 61
  - 4.1.3 Solar panel output . . . 64
  - 4.1.4 Photodiode biasing . . . 64
  - 4.1.5 Light-emitting diodes . . . 65
  - 4.1.6 Hayes-Shockley experiment . . . 66
  - 4.1.7 Hall effect on electrons versus holes . . . 68
  - 4.1.8 Loaded versus unloaded solar panel . . . 69
- 4.2 Quantitative reasoning . . . 70
  - 4.2.1 Miscellaneous physical constants . . . 71
  - 4.2.2 Introduction to spreadsheets . . . 72
  - 4.2.3 Predicting diode current . . . 75
- 4.3 Diagnostic reasoning . . . 76
  - 4.3.1 Tilted band diagrams . . . 77

A Problem-Solving Strategies **79**

B Instructional philosophy **81**
B.1 First principles of learning 82
B.2 Proven strategies for instructors 83
B.3 Proven strategies for students 85
B.4 Design of these learning modules 86

C Tools used **89**

D Creative Commons License **93**

E References **101**

F Version history **103**

Index **104**

Chapter 1

Introduction

1.1 Recommendations for students

This module marks the transition of study from *electric* circuits to *electronic* circuits, the distinction being that "electronic" circuits incorporate devices designed to actively[1] control the motion of electrons passing through. Such devices permit electricity to be used for much more than just energy conveyance, for example communication, data recording, and computation.

Important concepts related to conductors, insulators, and semiconductors include the **Conservation of Energy**, the **Conservation of Electric Charge**, charge carrier **mobility**, **subatomic particles** (electrons, protons, neutrons), electron **shells**, **quantum states**, **valence electrons**, electron **bands**, **allotropes**, **bandgap**, **energy band diagrams**, **electric fields**, **free electrons**, **holes**, **doping**, **majority carrier**, **minority carrier**, **acceptor** versus **donor** atoms, **Fermi level**, **depletion region**, **solar cell**, **biasing**, **Shockley diode equation**, and **diodes**.

Here are some good questions to ask of yourself while studying this subject:

- How might an experiment be designed and conducted to test the voltage and current output of a solar cell under different lighting conditions? What kind of graphs might you plot with the collected data?
- How might an experiment be designed and conducted to explore the characteristics of a PN junction? What hypothesis (i.e. prediction) might you pose for that experiment, and what result(s) would either support or disprove that hypothesis?
- What conservation law(s) is Kirchhoff's Voltage Law based on?
- What conservation law(s) is Kirchhoff's Current Law based on?
- In what ways do subatomic particles behave in *quantum* fashion as opposed to "classical" motion as we experience with macroscopic objects?

[1] Here, the word "active" refers to control via electrical energy. That is to say, an electronic device allows electricity to control other electricity!

- What is a *quantum leap*?
- What is an electron *shell*?
- What does a band diagram show us?
- Why do electron bands form in solid materials?
- What condition(s) is/are necessary for electricity to conduct through a solid substance?
- What are "holes" and how does their behavior contrast with free electrons?
- In what type(s) of substance are we likely to find free electrons?
- In what type(s) of substance are we likely to find holes?
- What property of light allows it to create an electron/hole pair within a semiconductor?
- What happens to a photon after it strikes a semiconductor to create an electron/hole pair?
- What is *doping* in a semiconductor material?
- Why are dopants intentionally added to otherwise pure semiconducting materials?
- What are some ways used to dope semiconducting materials?
- How does the water-in-a-tilted-chamber help explain electrical conduction through different substances?
- What does the Fermi level represent in an energy band diagram?
- What happens to the energy band diagram when a DC voltage is applied across the specimen?
- What happens to the energy band diagram when a P-type semiconductor joins an N-type semiconductor?
- What factor(s) other than applied voltage affect the amount of current passing through a PN junction?
- How does a photovoltaic solar cell function to convert light into electricity?
- How pure does silicon have to be in order to be useful as a raw material for semiconductor manufacture?
- Why must semiconductor material begin in such a pure state prior to the chip-making process?
- What is unique about the manufacturing environment where ultra-pure silicon is made into functioning electronic circuits?

1.2 Challenging concepts related to semiconductors

The following list cites concepts related to this module's topic that are easily misunderstood, along with suggestions for properly understanding them:

- **Electrical conduction in semiconductors** – what differentiates semiconductors from regular (Ohmic) conductors is the relative scarcity of free charge carriers, and how the concentration of charge carriers may be externally controlled in semiconductors by light and/or by strong electric fields. Current through a PN junction does not follow Ohm's Law (which is a great illustration of how Ohm's "Law" is not really a Law in the strict sense of the word, but rather a definition of resistance), but rather grows exponentially with applied voltage as described by the Shockley diode equation.

- **Electron versus Conventional flow notation** – so-called "Conventional flow" notation for electric current is exactly what the name suggests: nothing more than an accepted convention for labeling the direction of current. With semiconductors we must consider two different types of mobile electric charge carriers: free electrons which inhabit the conduction band like metallic conductors, and "holes" which are electron-vacancies within the valence band, holes behaving more like positively-charged particles in their own right. Likening electrons to water droplets and holes to air bubbles in water is a useful analogy to explain the difference.

- **Doping versus charge** – P-type semiconductor has been doped with impurities in order to create holes in the valence band, while N-type semiconductor has been doped with impurities to create free electrons in the conduction band, but this does not mean that P-type is positively-charged or that N-type is negatively charged. In both cases the doped semiconductor materials are *electrically neutral* because the proton-to-electron ratio of 1:1 has not been disturbed. The dopants simply have different atomic numbers (i.e. different number of protons than the semiconductor substrate, along with an equal number of electrons), and so the "excess" electrons to make N-type or the "electron deficiency" to make P-type really refers to the valence number of the dopant elements compared to that of the semiconductor substrate (usually silicon) and how those differing numbers of valence electrons do not neatly fit that of their neighboring atoms in the semiconductor crystal lattice.

- **Energy band diagrams** – these represent the potential energy levels of electrons and holes within semiconductors, with "up" referring to higher potential energy levels for electrons as well as lower potential energy levels for holes (i.e. electrons "want" to fall down to lower energy levels like water droplets fall, and holes "want" to rise up to lower energy levels, both under the influence of gravity).

1.3 Recommendations for instructors

This section lists realistic student learning outcomes supported by the content of the module as well as suggested means of assessing (measuring) student learning. The outcomes state what learners should be able to do, and the assessments are specific challenges to prove students have learned.

- **Outcome** – Demonstrate effective technical reading and writing

  Assessment – Students present their outlines of this module's instructional chapters (e.g. Case Tutorial, Tutorial, Historical References, etc.) ideally as an entry to a larger Journal document chronicling their learning. These outlines should exhibit good-faith effort at summarizing major concepts explained in the text.

  Assessment – Students explain how to interpret semiconductor energy band diagrams.

  Assessment – Students present their "Conceptual Toolbox" Theory Journal entries listing important concepts applied in the Tutorial text(s). A "Conceptual Toolbox" serves as inspiration when solving problems, each concept suggesting a potential problem-solving action. Challenge students to collaboratively generate a Conceptual Toolbox for the day's topic during the class session and then use it when solving problems together, thus fostering the habit of reasoning from principles rather than recalling procedures!

- **Outcome** – Predict semiconductor device behavior

  Assessment – Predict the effect of varying lightfall on a photovoltaic cell or array; e.g. pose problems in the form of the "Solar panel output" Conceptual Reasoning question.

  Assessment – Predict the effect of varying electrical load on energy transfer within a photovoltaic cell or array; e.g. pose problems in the form of the "Loaded versus unloaded solar panel" Conceptual Reasoning question.

- **Outcome** – Apply the Shockley diode equation

  Assessment – Calculate forward current for a PN junction given necessary parameters; e.g. pose problems in the form of the "Predicting diode current" Quantitative Reasoning question.

- **Outcome** – Independent research

  Assessment – Read and summarize in your own words reliable source documents on the subject of semiconductor physics. Recommended readings include patents filed by early researchers (especially those owned by Bell Laboratories).

Chapter 2

Tutorial

Our study of electric circuits began in this series of learning modules with an exploration of *energy*: the ability to set matter into motion, and specifically how electric charges could act as couriers of energy, absorbing energy from *sources* and delivering energy to *loads* while traveling through materials called *conductors* and being prohibited from traveling through *insulators*. We learned that the measure of energy lost or gained by electric charges between two different points was called *voltage*, that the rate of electric charge motion was called *current*, and that any opposition to current resulting in the conversion of charges' energy into heat was called *resistance*. Combined with certain conservation laws, including the Conservation of Energy (that energy is eternal, impossible to create or destroy) and Conservation of Electric Charge (that likewise electric charges are eternal, impossible to create or destroy), we developed and applied a variety of laws describing the behavior of electric circuits, including Ohm's Law (voltage is the product of current and resistance, $V = IR$), Joule's Law (power is the product of voltage and current, $P = IV$), Kirchhoff's Voltage Law (the algebraic sum of all voltages in any closed path is equal to zero), and Kirchhoff's Current Law (the algebraic sum of currents entering and exiting a point is likewise equal to zero). Energy and its conversion from one form to another has been the common thread in this exploration of electricity and electric circuits.

It should come as no surprise that energy is also a foundational concept in the study of atoms and subatomic particles, some of which possess electric charge and may under certain circumstances participate in the formation of an electric current. In particular, the energy level of electrons within atoms determines those electrons' "freedom" to move about, and as such determines whether or not that substance is electricity conductive in its solid form. We will begin this tutorial with an exploration of electron energy states within atoms, and how this applies conductors, insulators, and a new class of substances called *semiconductors*.

2.1 Atomic structure

We know that most matter is comprised of atoms, and atoms in turn are comprised of three basic subatomic particle types: *electrons*, *protons*, and *neutrons*. Protons possess a positive electric charge while electrons possess an equal-magnitude but opposite-sign (negative) electric charge. Neutrons are uncharged particles, neither positive nor negative. Both protons and neutrons are massive in comparison to electrons, protons and neutrons residing in the center of the atom (called the *nucleus*) while electrons inhabit prescribed "orbital" spaces around the nucleus.

Perhaps the most important new foundational concept in this tutorial is the fact that electrons can only orbit an atom in set spaces, referred to as *shells* and *subshells*. Each shell and subshell represents a different level of potential energy for the electron: the closer the electron orbits the atom's nucleus, the less energy it possesses. This is not unlike the potential energy of a mass orbiting a planet: the closer the mass orbits the planet, the less gravitational potential energy it possesses. Unlike masses in planetary orbits which are free to orbit at any arbitrary altitude (as determined by their energies), electrons are not free to assume arbitrary energy levels around the nucleus of an atom[1]. Instead, **the gaps existing between shells and subshells represent uninhabitable zones for electrons**. Electrons may "jump" from one subshell to another but not exist in-between. In other words, the energy possessed by an orbiting electron is a *quantum* property[2] rather than a *continuous* property. Another important concept to grasp is that the subshells themselves possess a limited capacity for electrons. The first shell of an atom can only hold two electrons, while the second shell has a maximum capacity of eight electrons. This phenomenon is akin to different levels and rows of seats in an amphitheater: people may sit in any of those seats (those closest to the action being the most prized, and therefore costing the most energy to occupy), but between rows and between levels there is no space for anyone to sit. As the inner-most rows fill up, addition spectators must sit in higher-level rows.

While all of this may seem very strange to someone accustomed to thinking of electrons as nothing more than tiny satellites whirling around an atom's nucleus, the quantum nature of atomic electrons is central to the understanding of chemistry as well as the conduction of electricity in solid substances. If it were not for the quantized nature of electron placement around each atom's nucleus, atoms would not have stable and predictable chemical properties. As it is, the highly ordered nature of atomic electrons is what gives each element its unique chemical properties, and explains why some atoms bond readily to other atoms to form molecules while other atom combinations do not bond at all.

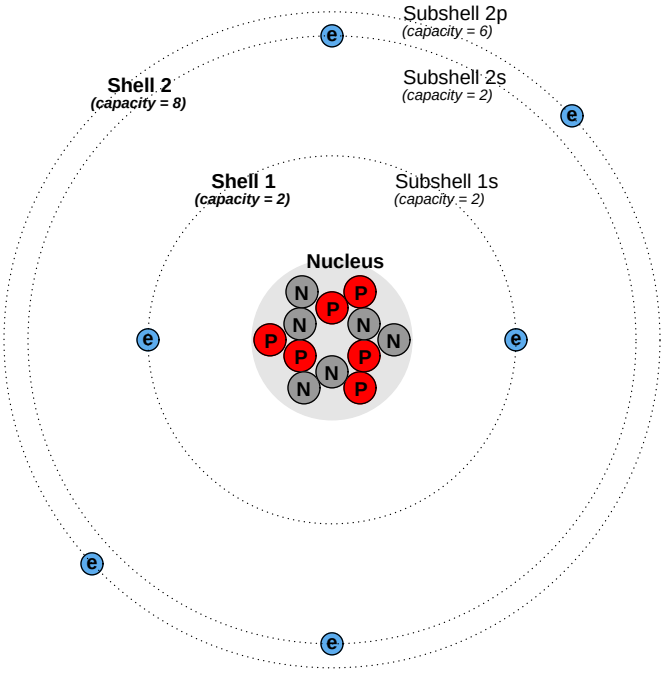
[1] It should be noted that electrons are free to possess any level of energy above the *vacuum level*, which is the amount of energy required for an electron to entirely escape an atom and go its own way through the vacuum outside of the atom. We might think of this "vacuum energy level" as being analogous to the *escape velocity* needed for an object to leave the gravitational influence of a planet.

[2] The so-called *quantum revolution* in physics during the early 20th century was marked by a number of realizations, including the particle-wave duality of matter and the exclusiveness of certain particle properties. Electrons, like photons and other particles of matter, exhibit wave-like behavior as well as particle-like behavior. The tendency for electrons to only exhibit certain discrete energy levels while orbiting an atom's nucleus is explained by the electron existing as a "standing wave" around the nucleus, with only a select set of energy levels (wave frequencies) permitting resonance of that wave. When an electron gains or loses energy, its frequency changes, and it therefore "stands" in a different shell or subshell within the atom. Furthermore, the number of electrons able to inhabit a single subshell is limited by *quantum states*, which means it is possible for shells and subshells to be completely "filled" and therefore unable to host additional electrons.

A full presentation of quantum theory is far beyond the scope of this text, and would not be necessary to understand the function of electronic devices at the circuit level anyway. Suffice it to say that the following quantum principles are important to grasp for our scope of understanding:

- Each electron possesses a set of quantum states (energy level and "spin") that are discrete rather than continuous
- No two electrons in close vicinity (i.e. within the same atom or molecule) may share the exact same set of quantum states
- Each electron shell has a limited capacity (i.e. limited available quantum states)
- The lower an electron's energy level, the closer it resides to the nucleus of an atom
- An electron will "jump up" to a higher shell if it receives sufficient energy from an external source, and it must relinquish energy if it "falls down" to a lower shell

The following illustration shows a pictorial representation of a single atom of *carbon*, possessing six protons, six neutrons, and six electrons. Two electrons completely occupy the lowest-energy Shell 1 (having only one subshell, 1s) while four more electrons occupy Shell 2 at higher energy states, two in the 2s subshell and two in the 2p subshell. Higher-level shells also exist for a carbon atom, but are not shown here because they are typically unoccupied by any electrons:



As crude[3] as it is, the model does help us envision the energy gaps between shells and subshells, and the relative placement of the particles.

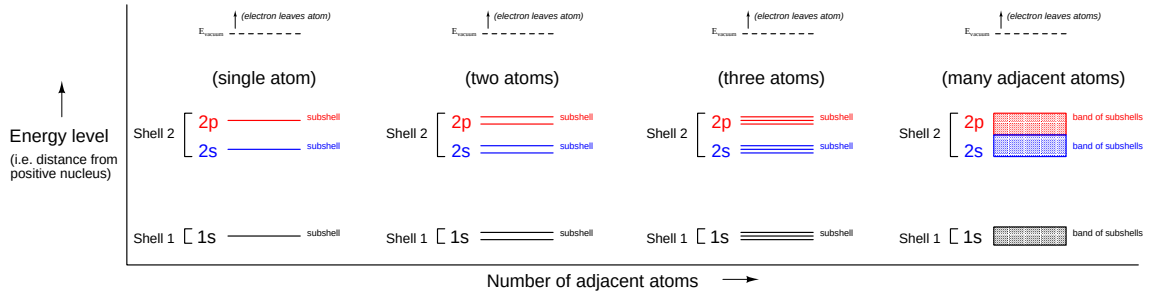
The outer-most (highest-energy) subshell of the carbon atom containing electrons (2p) has a maximum electron capacity of six. Since a carbon atom only requires six electrons total to electrically balance the six protons in its nucleus, and there are already two electrons in Shell 1 and two electrons in subshell 2s, the next subshell (2p) need only carry two additional electrons which leaves that outer subshell with four vacancies. These vacancies are important because they permit occupancy by electrons *from neighboring atoms*, which is how atoms join[4] together to form molecules. These outer-most shells of an atom are called *valence* shells, and the electrons inhabiting them are called *valence electrons*.

[3] The actual shapes of atomic orbitals are highly non-intuitive, most of them resembling dumbbells, teardrops, and halos. What is more, the three-dimensional shapes of these orbitals represent *zones of probability* where electrons *might* be found. One aspect of "quantum weirdness" is that particles lack definite positions.

[4] Atomic bonds may be grouped into two major categories: *ionic* and *covalent*. An ionic bond is formed when one atom "steals" an electron from another atom by means of providing a more favorable energy state for that electron than its parent atom. The atom gaining the electron now becomes a negatively charged ion, while the atom robbed of its electron becomes a positively charged ion. Those two ions now become electrostatically attracted toward each other. Table salt (sodium chloride, $NaCl$), is an example of an ionically-bonded substance. A covalent bond is formed when two or more atoms mutually share electrons: those electrons find favorable energy states in *both* atoms simultaneously, and in doing so link those atoms together to form a molecule. Most hydrocarbon compounds (e.g. methane, CH_4) are examples of covalently-bound substances. A simple empirical test of bonding type is whether or not the substance is electrically conductive in its liquid form. Ionically-bonded liquids are conductors of electricity, because the electrically-imbalanced ions are function as mobile charge carriers. Covalently-bonded liquids are insulators, because their atoms are all electrically balanced.

2.2 Energy bands and conductivity

In a solid material, atoms are linked to each other by these electron bonds to form a rigid structure, but the available quantum states of electrons happens to be far more complex than for a single atom. The valence subshells of adjacent atoms overlap, and in doing so must assume slightly different energy levels in order to provide unique quantum states for their electrons. When very large numbers of atoms reside close together in a solid, the abundance of valence subshells overlapping each other results in a relatively wide *band* of possible energy states at slightly different levels that could be inhabited by electrons[5]. The following illustration compares the energy levels of electron subshells for an individual atom (highest occupied shell = 2) versus multiple (bonded) atoms of the same type:



An electrical *conductor* is a substance allowing free motion of charge carriers within in. Atoms within a solid material are physically bound and cannot drift, and so the only way for a solid to be electrically conductive is if some of the atoms' electrons are able to easily leave their parent atoms and move into vacancies within the subshells of adjacent atoms. If the outer-most energy band of a group of atoms is *partially full* of electrons, with room for additional electrons, it means electrons may travel from atom to atom within the unfilled portion[6] of that band with only a slight investment of energy. This is the defining characteristic of a *conductor*: a band of unoccupied quantum electron states stretching across the entire bulk of the specimen permitting mobility of electrons throughout.

An *insulator* by contrast exhibits a large energy gap between a filled energy band and the next available (empty) energy band. Electrons within the filled band[7] are invested in binding the atoms together, and cannot move from their positions. Those electrons could, in principle, "jump" up to the next empty band to travel there, but only if sufficient energy enters from an external source to boost those electrons to that necessary energy level[8]. Without that energy boost, the electrons are "locked" into place and cannot drift.

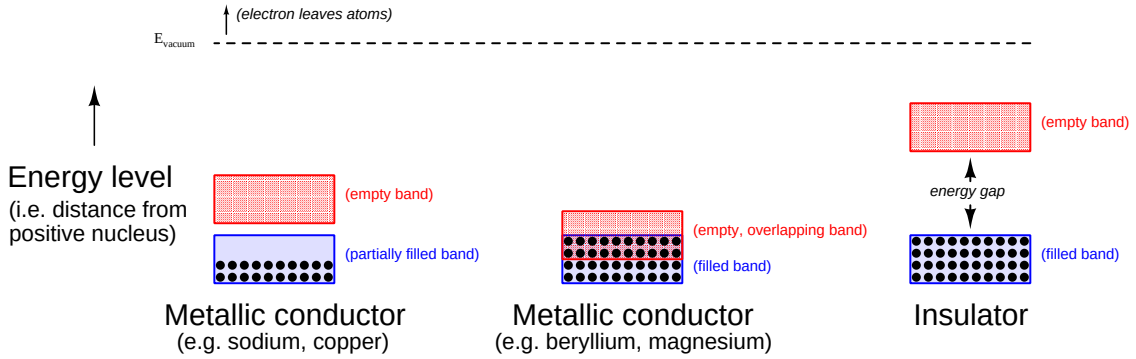
[5] This phenomenon is analogous to multiple single-lane roads joining to form a multi-lane highway, the single-lane roads representing subshells of individual atoms, and the multi-lane highway representing the combination of many atoms' outer subshells in close proximity with each other.

[6] Returning to the road/highway analogy of energy bands, a partially-filled band is akin to a highway with some empty lanes. Any vehicle may leave the full lane and travel freely in an adjacent empty lane.

[7] Using our road/highway analogy again, a completely-filled band is akin to a highway with all lanes so completely filled with vehicles that it has become a parking lot, where no vehicles may move from their positions.

[8] Imagine a highway completely full of vehicles (i.e. a completely filled band) unable to move, and then a crane picking up one of those vehicles to place it on an adjacent highway with empty lanes (i.e. an external energy source boosting an electron up to an empty band).

A few *energy band diagrams* illustrate the difference between conductors and insulators. The fundamental difference between an insulating solid and a conducting solid is whether an energy gap exists which electrons must cross in order to become mobile. In conductors, no such energy gap exists for valence electrons, and they are able to move freely through the typically unoccupied layers of the band (called the *conduction band*). In insulators, a large energy gap separates the filled valence band from the next available (empty) band, which means those valence electrons are immobile unless and until some sufficiently large energy source[9] propels them upward into the empty band:



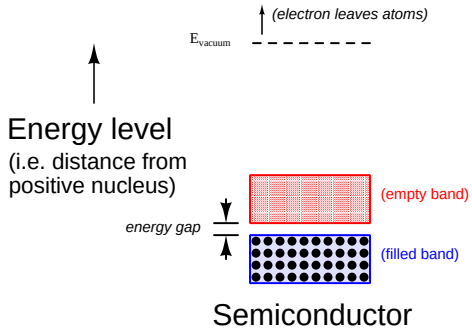
Remember that in order for an electron to drift within a solid substance, there must be unoccupied subshells in adjacent atoms for it to move into. For metals such as sodium and copper the upper levels within the partially-filled valence band provide that mobility. Metals such as beryllium and magnesium possess an empty band overlapping the full valence band, those unoccupied energy states similarly offering places for electrons to move from atom to atom. Insulators, by contrast, exhibit a large energy gap between a full valence band and the next band of available energy states. No electron can break away from its parent atom without having its energy boosted by some external source to get into that much-higher band. In the absence of such an energy source, an insulating material therefore is non-conducting.

The placement of these energy bands, and therefore the ability of any solid substance to conduct electricity, depends as much on the configuration of atoms comprising molecules as much as it does on the individual atoms themselves[10]. Substances comprised of identical atoms, bound together in different ways, are called *allotropes*. The different allotropes of carbon illustrate how the molecular bonds – and not just the atoms themselves – determine the substance's electrical characteristics. *Graphite* is an allotrope of carbon whereby the individual carbon atoms link together to form flat "sheets" of molecules, and happens to be a fair conductor of electricity. *Diamond* is a different allotrope of carbon whereby the individual carbon atoms link together to form tetrahedral pyramid shapes, and happens to be an insulator at room temperature. Yet another allotrope of carbon is *Fullerene* whereby individual carbon atoms link together to form hollow polyhedral spheres or tubes, the degree of electrical conductivity varying with the specific geometry of each molecule.

[9] It is worth noting that *temperature* is an expression of energy capable of "promoting" fixed valence electrons into the conduction band of a material. This is why some insulating materials such as glass become electrically conductive when heated.

[10] Many introductory textbooks fail to make this point clear, instead presenting conductivity as being a simple function of each atom's electron shells rather than also being a function of how those atoms bond with each other to form molecules and larger structures.

A third category of solid material is the *semiconductor*, a substance having a narrow energy gap (also called the *bandgap*) separating a filled valence band from an empty conduction band. In a semiconductive substance the valence electrons are very nearly (but not quite) at the necessary energy level needed to move about freely within the substance. This means semiconductors are electrically insulating until they experience the proper stimulus – namely, some stimulating event granting enough energy for electrons to make the modest "leap" across that short energy gap and thereby become free to drift through the material – at which point they conduct electricity. This phenomenon of *controlled conductivity* makes semiconducting substances very useful in constructing electronic devices:

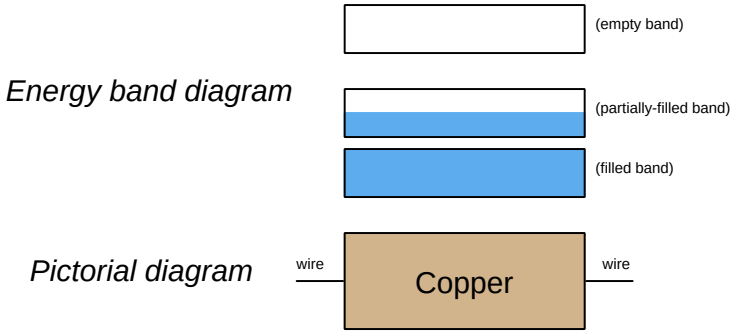


Temperature is an important factor in semiconductivity, just as it is a critical factor in the behavior of electrical insulators. Substances such as gases may be electrically insulating at room temperature and conducting at very high temperatures, while a semiconductor may be insulating at room temperature and conducting at slightly elevated temperatures. This means semiconductors are of no use beyond certain temperatures, because the ambient temperature forces them to be conductive when we need their conductivity to be controllable. Early semiconductor devices were manufactured from the element *germanium*, but this element proved to be unreliable for the purpose because its upper temperature limit was impractically low for applications where the device had to dissipate significant power. The element *silicon* has much better temperature characteristics for practical applications and is the dominant semiconducting substance used today.

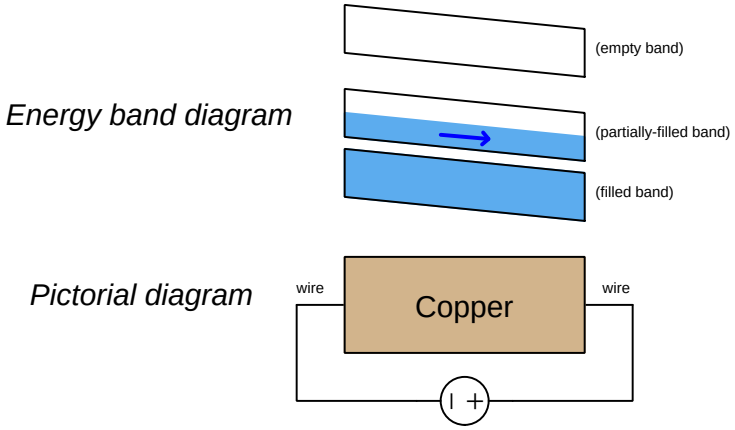
2.3 Energy bands and electric fields

The vertical axis of an energy-band diagram represents potential energy for electrons, a direct function of distance from the positively-charged nucleus of an atom. Electrons tend to settle into the lowest-possible energy states just as pebbles tend to settle into the lowest height in a pile of rocks, or as droplets of water settle into the lowest height in a pool. Any unoccupied regions will be at the highest energy levels within a band. This is why the energy band diagram for the metal copper shows the partially-filled band with electrons in the bottom portion only. Those valence electrons don't spread out over the entire band, but rather exist at the lowest portions only.

A helpful analogy for the behavior of electrons within energy bands is to envision electrons as molecules of water, and each band as a hollow chamber. Gravity plays the role of the electric field existing between the electrons (water molecules) and the earth (atomic nucleus), accounting for their default positions at the bottom of the chamber. Here we show a three-band diagram for a metal such as copper, with the lowest band(s) completely full, the middle band partially full, and the upper band completely empty. A pictorial representation located below the energy band diagram shows a metal specimen with metal wires attached at either end:



If we attach a voltage source to these two wires, we impose a *horizontal* electric field upon these electrons. This potential difference across the specimen *tilts*[11] the energy band diagram, causing some of the electrons to drift:

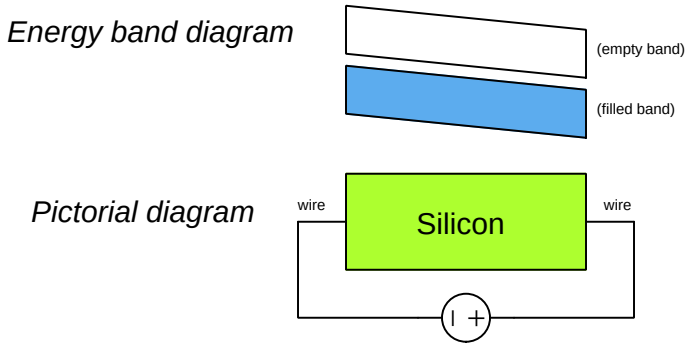


Note how the only band with *moving* electrons inside of it is the partially-filled band[12]. The empty upper band of course has no electrons in it to move, and the full lower band has no unoccupied space for its electrons to move into.

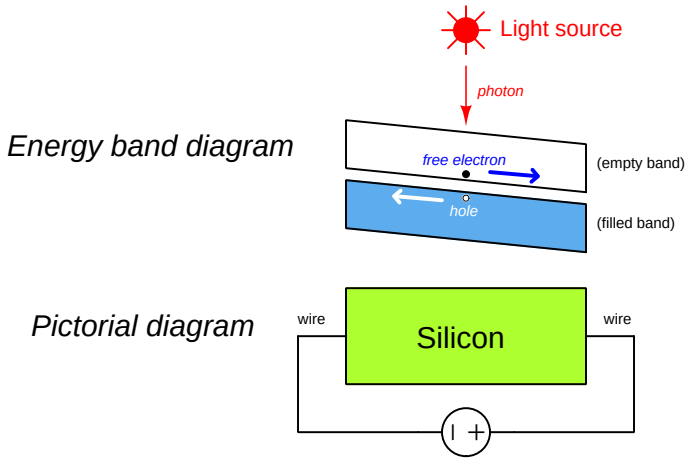
[11] Since we define energy in a band diagram as potential energy of negative electrons separated from positive nuclei, a negative ($-$) potential is higher and a positive ($+$) potential lower on the vertical axis. The direction of electron flow, of course, is from $-$ to $+$.

[12] Applying the highway analogy, it is as if there are three highways: one with all lanes empty, one with empty lanes adjacent to lanes filled with parked vehicles, and another with every lane blocked by parked vehicles. Only the partially-full highway supports traffic.

Now let us consider a different test specimen, this time made of a pure semiconductor material (e.g. silicon) rather than a metal such as copper. Applying a voltage across the ends of this specimen once again "tilts" the energy band diagrams, but no flow occurs because there are no partially-filled bands within the semiconductor:



However, something very interesting happens when we shine light at this semiconducting specimen: if a photon (a particle of light) possessing energy equal to the energy gap between the filled and empty bands (i.e. the material's *bandgap*) happens to enter this material, that photon's energy will transfer to an electron within the filled band, "boosting" it into the higher-level band. This creates a *free electron* in the upper band and a corresponding *hole* (i.e. electron vacancy) in the lower band. With the electric field provided by the voltage source "tilting" the energy bands, the free electron drifts from − to + while the hole drifts from + to −:



This counter-flowing motion of electrons and holes in an electric field is very much like the motions of *water droplets* and *air bubbles*, respectively, in a gravitational field: a droplet of water in an otherwise air-filled chamber will roll downhill by the force of gravity; an air bubble in an otherwise water-filled chamber will float uphill, also by gravity. It is worth noting that free electrons and holes are not equally mobile: all other factors being equal, free electrons move over twice as fast as holes!

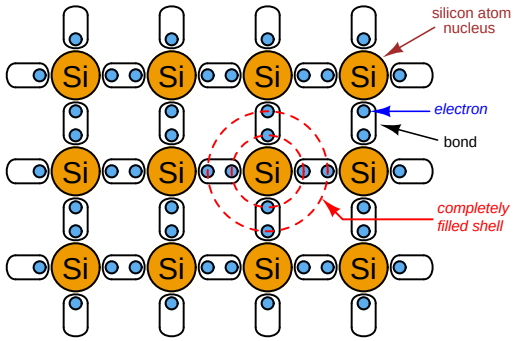
2.4 Doping

The number of electron-hole pairs created is a direct function of how many photons of appropriate energy level strike the specimen. If the light source is continuous, there will be a continuous flow of charge carriers (electrons and holes, moving in opposite directions with the source's electric field) inside the specimen, and the circuit will exhibit a current. High temperature will do the same to a semiconductor: if hot enough, thermal agitation of the atoms will promote some electrons into the upper energy band, creating free electrons and also holes where the electrons left the lower band. Thus, we see how semiconductive substances behave as *controlled conductors*, their electrical conductivity subject to external stimuli.

Light and heat are not the only ways to stimulate a semiconductor to conduct electricity, though. Another method is to purposely insert impurities into the specimen. Specifically, if we alloy the otherwise pure semiconductor with elements whose atoms have a different number of valence electrons in their outer-most electron shells, it is possible to create conditions inside the semiconductor where either free electrons or holes naturally exist. This intentional mixing of substances is called *doping*.

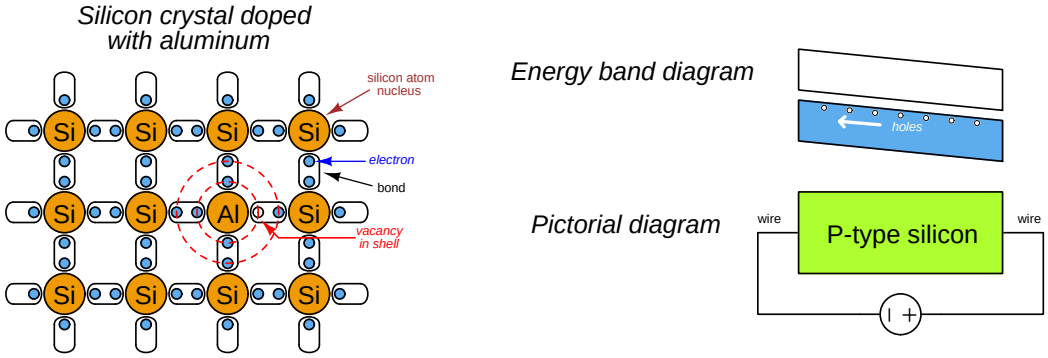
Silicon is the most common semiconductive element, and it is *tetravalent* (i.e. each atom of silicon possesses *four* valence electrons in its outer-most shell). Each atom of silicon in a pure specimen bonds with four other atoms of silicon, forming a pyramidal crystal structure. In this bonded structure, each atom effectively has eight outer electrons: four of its own and one shared from each of the four neighboring atoms of silicon. Since eight is the maximum number of electrons for the outer-most shell, the energy band formed by all the silicon atoms' valence electrons is completely full.

Silicon crystal lattice

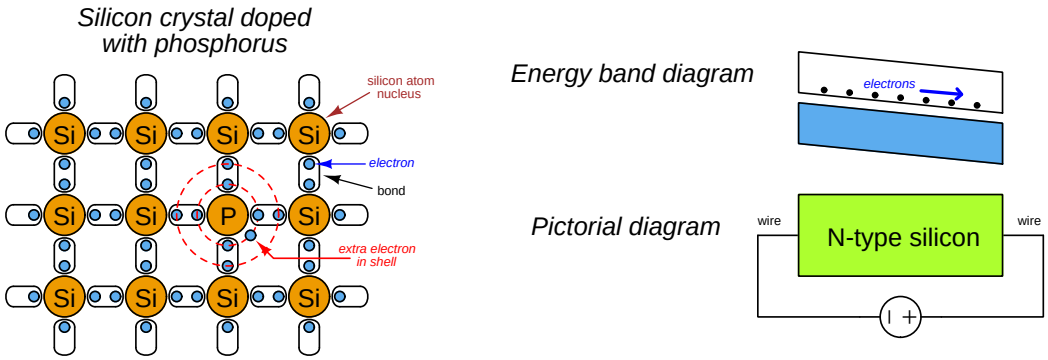


We may, however, disrupt this perfect filling of the silicon crystal's valence energy band by intentionally mixing in some atoms of other elements. Any atom that isn't tetravalent like silicon will interrupt the perfectly-filled condition of the valence band, resulting in different electrical properties.

For example, if we "dope" silicon with a *trivalent* (i.e. three valence electrons) element such as aluminum, boron, indium, or gallium, these atoms only have three electrons each to share with their silicon neighbors, each one creating a vacancy (i.e. hole) in the crystal that an electron could occupy. We refer to a trivalent-doped semiconductor as *P-type* in honor of the "positive" holes[13] existing in the valence band. If we connect a P-type specimen to a voltage source, these holes constitute the majority of charge carriers for the circuit's current:



A similar phenomenon occurs if we "dope" silicon with a *pentavalent* (i.e. five valence electrons) element such as phosphorus, arsenic, antimony, or bismuth. Those atoms have five electrons to share, but only have four silicon neighbors each. This creates an extra electron with no space to fit in the filled energy band, and so it exists in the next higher band as a free electron. We refer to pentavalent-doped semiconductor as *N-type* in honor of the negative holes existing in the energy band above the valence band (called the "conduction band"). If we connect an N-type specimen to a voltage source, these electrons constitute the majority of charge carriers for the circuit's current:



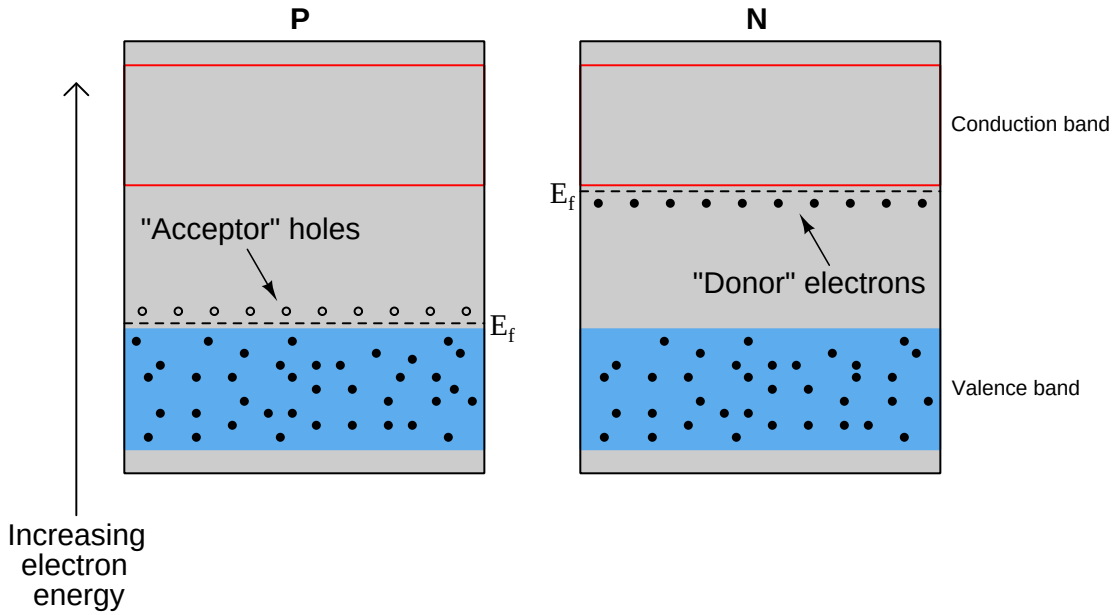
[13] Strange as it may seem, it actually makes a lot of sense to treat these electron vacancies as though they were positively-charged particles, just as it makes intuitive sense to describe the motion of an air bubble in water (floating up) rather than think of that motion in terms of the surrounding water molecules falling down. You will also note that these holes move in the direction of *conventional flow*, the way engineers commonly denote direction of electric current in any circuit. The rationale for "conventional flow" is grounded in historical assumptions about electricity, long before people knew what electrons were. Some call this "hole flow" because it matches the direction of hole motion, but it should be understood that holes do not exist in metals but only in semiconductors.

Thus, by judiciously alloying silicon (or other semiconducting substance) with certain other substances, we are not only able to make it conductive at room temperature but also dictate the *type of charge carrier* comprising that conductivity. Low doping concentrations result in low conductivity (i.e. high resistivity), and vice-versa. In fact, this is how resistors are fabricated in semiconductor integrated circuits: by carefully controlling the amount of doping applied to the silicon between two specific points.

A sample of absolutely pure silicon is called an *intrinsic* semiconductor, while a sample of doped semiconductor (either N-type or P-type) is called *extrinsic*. For an N-type semiconductor, the vast majority of charge carriers are free electrons in the conduction band and so we refer to the electrons as the *majority carriers*, but inevitably there will also be a relatively small number of holes in the valence band as well which we call *minority carriers*. For a P-type semiconductor the roles are reversed: holes are the dominant type of charge carrier in this substance and so they are the majority carriers while the inevitable free electrons are the minority carriers. Metallic substances, by contrast, only contain free electrons and no holes whatsoever.

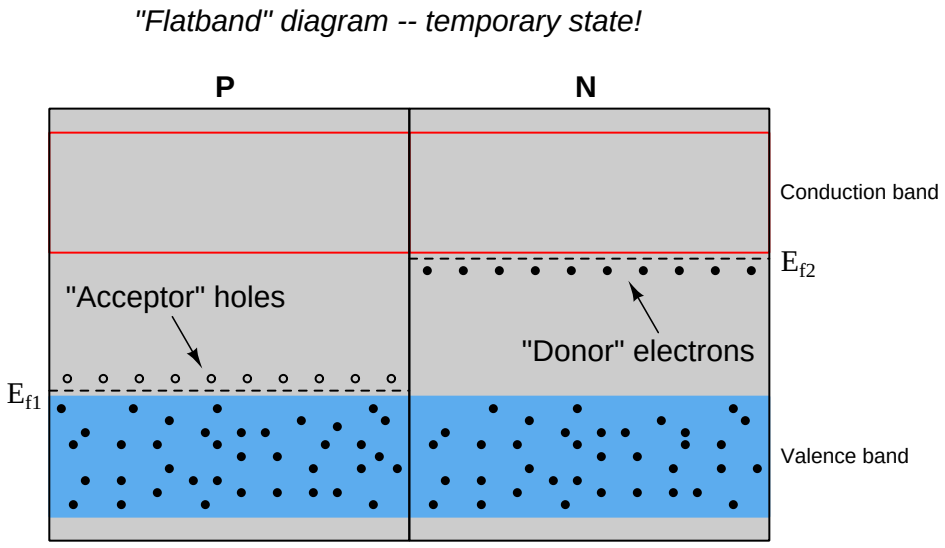
2.5 The PN junction

By far, the most interesting behaviors of semiconducting substances arise when we form junctions of dissimilarly-doped material, i.e. when we join P-type to N-type. To begin our exploration, we will take a closer look at the energy band diagrams of both P-type and N-type extrinsic semiconductors:



Trivalent dopants added to silicon produce "acceptor" sites in the crystalline structure just above the valence band for pure silicon. With very little energy (room temperature is enough), electrons from the valence band may move up into these acceptor sites to leave mobile holes within the valence band, making it P-type material. Conversely, pentavalent dopants added to silicon produce "donor" sites just below the conduction band for pure silicon. Again, room temperature is sufficient thermal energy to boost electrons from the donor atoms into the conduction band where they are mobile, making it N-type material. The energy level marked *Fermi level* in each half of the diagram shows the highest energy level of any electron at a temperature of absolute zero where there is no atomic motion.

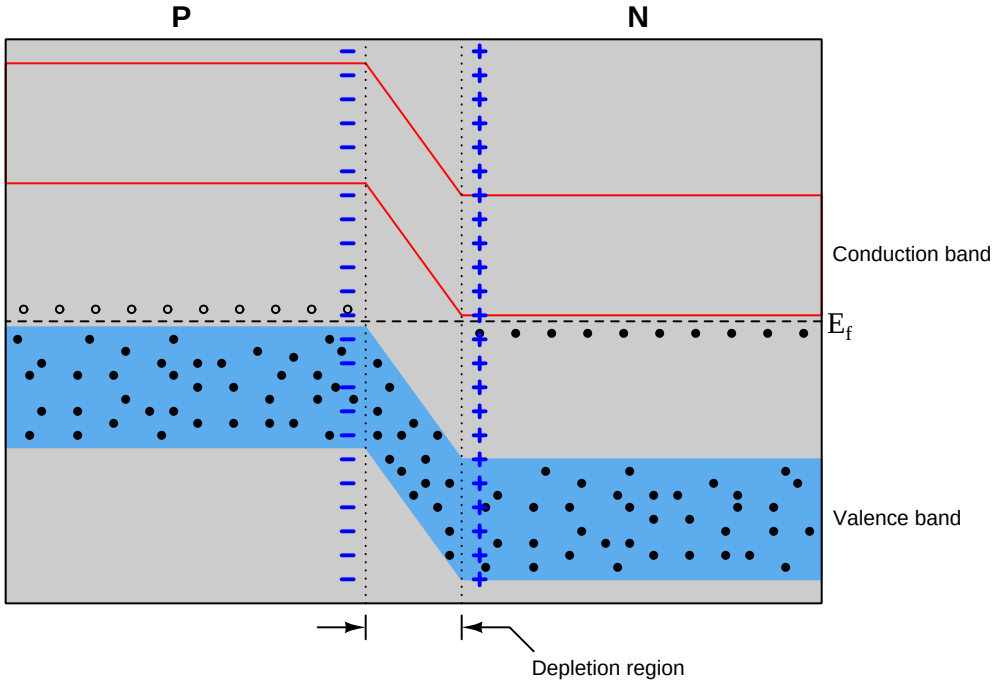
If we join the P-type and N-type materials together, the energy states at the time of initial contact are unchanged, but this state of affairs is only temporary:



Very shortly after joining, electrons from the N-type side "fall" into holes on the P-type side as they naturally seek lower energy states. Just as an elevated mass "seeks" ground level by the force of gravity, any high-energy electron "elevated" above positively-charged atomic nuclei will "fall" into the lowest-energy region it may find. Joining the P-type and N-type materials together provide new low-energy spaces for the N-type's electrons to fall into.

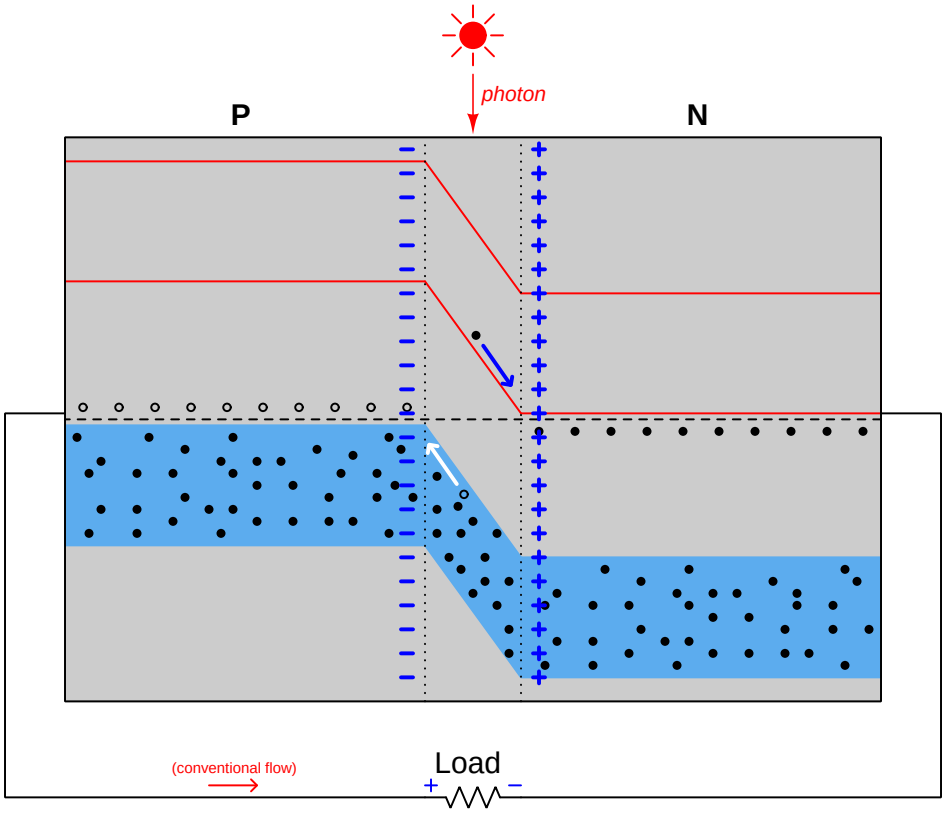
Recall that the *Fermi level* represents the highest-level energy condition for any electrons in the substance. If we envision the electrons as water molecules in a large body of water, the Fermi level represents the height of the water's surface on a calm day. Joining the P- and N-type semiconductors together is akin to opening a connecting channel between two bodies of water at different surface heights: as one empties, the other fills.

As electrons from the N side fall into holes in the P side, an electrical charge imbalance develops across the junction. This creates an internal electric field (− on the P side, + on the N side) tilting the energy bands until the two sides' Fermi levels equalize:



Eventually, this electric field grows strong enough to repel additional electrons on the N side from falling into holes in the P side, the result being a stable region of material surrounding the physical junction devoid of either electrons or holes but possessing an electric field. This is called the *depletion region*. As the name suggests, its lack of mobile charge carriers means that it is electrically insulating, just like a region of pure (intrinsic) semiconductor.

One practical application for a PN semiconductor is as a source of electricity when exposed to light. If a photon of appropriate energy enters the depletion region, it may generate an electron-hole pair by boosting[14] the energy level of a valence-band electron into the conduction band:

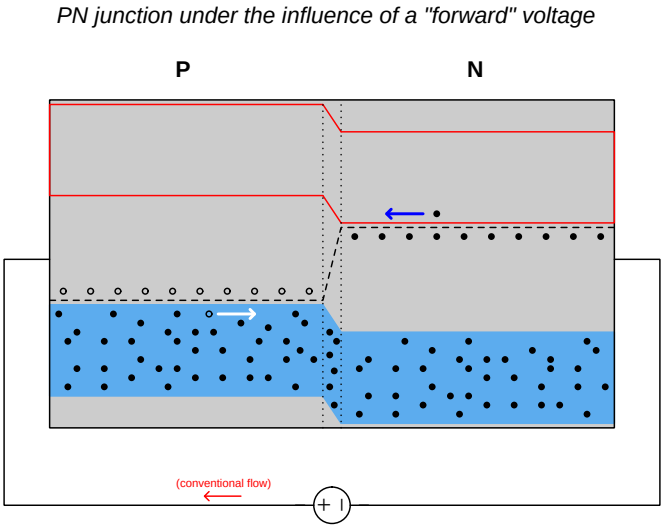


The negatively-charged electron "falls downhill" following the electric field's influence in the depletion region, while the positively-charged hole "rises uphill" following that same field. If a pair of wires connected to the N and P sides provide a complete path for load current, the PN junction will serve as an electrical source to a load. This is the basis for a *photovoltaic solar cell*.

Interestingly, the output voltage of a solar cell is determined by the strength of the electric field across the depletion region, which in turn is a function of the *bandgap* energy for the semiconducting material. For a silicon PN junction that voltage is approximately 0.7 Volts. A solar cell's output current is a function of the rate at which photons strike the PN junction (i.e. photons per second), since every electron-hole pair arises from the annihilation of a photon, and current is defined as the rate of charge carrier flow (i.e. charge carriers per second). This means a solar cell's open-circuit voltage output is relatively constant regardless of light intensity, but its short-circuit current varies in direct proportion to the intensity of light exposure.

[14] A photon of the right energy striking the PN junction causes one of the electrons residing in the lower (valence) energy band of the depletion region to "jump" up to the empty (conduction) energy band. The result is *two* new charge carriers: a free electron in the conduction band, and a hole in the valence band, where formerly no charge carriers existed in the depletion region.

Another practical application for a PN junction arises from our ability to expand or collapse the depletion region by applying an external voltage. If we connect a voltage source across the PN junction with the source's polarity arranged such that it opposes the polarity of the electric field inside the depletion region, that field will become nullified. If the applied voltage is large enough, the depletion region will shrink to the point where electrical conduction occurs through the junction, electrons from the N side and holes from the P side both approaching the junction where they combine:



This is called *forward-biasing* the PN junction: connecting the P side to the + of the external source and the N side to the − of the source. The *Shockley diode equation*, named after William Shockley, predicts forward-bias current as a highly non-linear function of applied voltage and temperature:

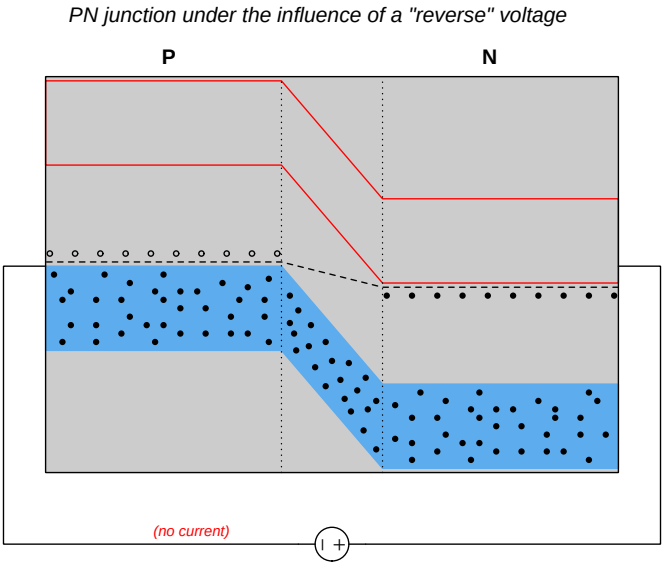
$$I = I_S \left(e^{\frac{qV}{nkT}} - 1 \right)$$

Where,

I = Forward-bias current through the diode, Amperes
I_S = Reverse-bias saturation current[15] through the diode, Amperes
e = Euler's constant (≈ 2.71828)
V = Voltage applied to the PN junction externally, Volts
q = Elementary charge of an electron (1.602×10^{-19} Coulombs)
n = Ideality factor (1 for a perfect junction)
k = Boltzmann's constant ($1.3806504 \times 10^{-23}$ J / K)
T = Absolute temperature (Kelvin), 273.15 more than degrees Celsius

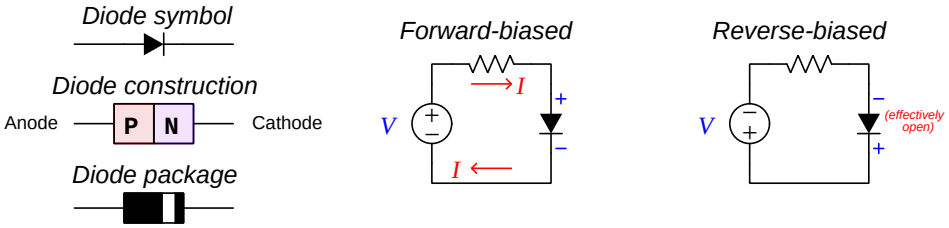
[15] A very small amount of current will still flow in the reverse-biased condition, due to so-called *minority carriers* in the P and N halves of the diode. This tiny current, usually in the range of *nano-Amperes*, is referred to as the *reverse saturation current* because its value does not increase appreciably with greater reverse-bias voltage but rather "saturates" or "plateaus" at a constant value. This saturation current, while fairly independent of applied voltage, varies greatly with changes in device *temperature*.

Reversing the polarity of the external source causes the depletion region to widen, which in turn ensures no conductivity through the PN junction:



Naturally, this is called *reverse-biasing* the PN junction: connecting the P side to the − of the external source and the N side to the + of the source.

A very useful electronic component called a *diode* works on this principle, allowing current to pass through it in one direction but not in the other:



2.6 Semiconductor production

The element silicon, as mentioned previously, is the dominant semiconductor material in use at the time of this writing. It is quite abundant on Earth and well-behaved as an electrical substance over a wide range of practical temperatures. Not only may it be "doped" to precisely control the type and quantity of electrical charge carriers, but it may also be intentionally oxidized in order to produce an insulator with a high dielectric strength. Silicon is most often found as a mineral compound rather than in pure elemental form, the most common natural compound of silicon being quartz (SiO_2) or silica ($Si(OH)_4$). In fact, a substantial percentage of the Earth's crust is comprised of silicon-based compounds.

Silicon feedstock for electronic applications, however, must be exceptionally pure, elemental silicon, ideally cast in a single-crystal (monocrystalline) form so as to avoid inter-crystal boundaries that might impede charge carrier motion. Through a long series of chemical reduction and refining stages, ultra-pure silicon is cast into long ingots known in the industry as *boules* weighing several hundred kilograms each. These cylindrical ingots of pure silicon are then cut into thin *wafers* using diamond saw tools, and those wafers trimmed to precise diameter and their surfaces polished to a mirror finish.

A photograph of the top of a silicon boule is shown here, this piece of solid silicon weighing approximately 18 kg and having a diameter slightly greater than 300 millimeters:

A full silicon boule takes multiple days to manufacture with current technology, with silicon purities ranging from 1 part per million (i.e. 99.9999% pure) to 0.01 parts per billion (i.e. 99.999999999% pure) depending on the application. The manufacture of photovoltaic solar cells which are essentially just a single PN junction, for example, can tolerate a higher level of impurities than state-of-the-art microcircuitry where millions or even billions of transistors are etched onto a single *chip* of silicon. High purity is important for any semiconductor device so as to eliminate the potential for unwanted electric charge carriers within the material, so that charge carrier type and concentration may be intentionally controlled by doping the silicon with either trivalent or tetravalent elements.

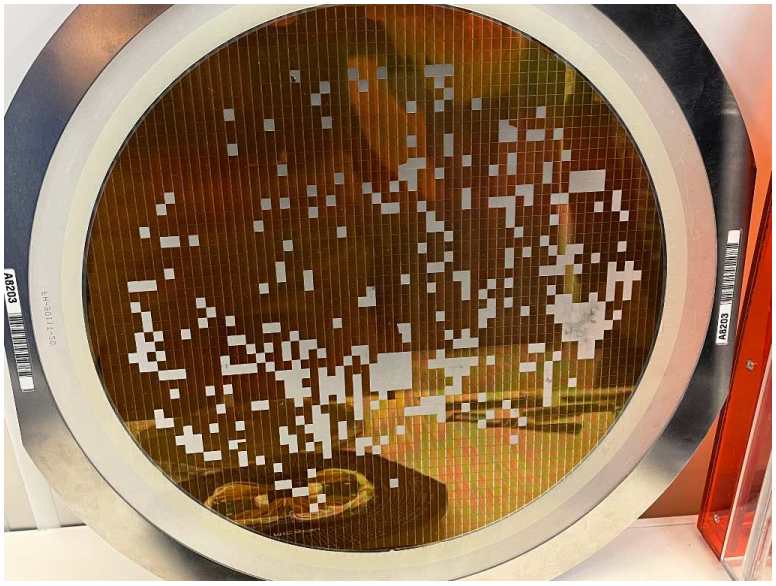
Doping is intentional addition of specific impurities (called *dopants*) into ultra-pure silicon in order to introduce mobile charge carriers, either free electrons for N-type doped material or valence-band electron vacancies called "holes" for P-type doped material. Dopant concentrations[16] vary considerably according to the needs of the particular device(s) being manufactured, but range from parts-per-billion (ppb) concentrations for lightly-doped silicon to parts-per-thousand for heavily-doped ("degenerate")[17] silicon.

Since most semiconductor devices require multiple types of doping to construct, the most basic being a simple *diode* with a single PN junction, doping is a process that must occur during the fabrication of each device rather than being a pre-treatment of the semiconductor materials. The trade term for such a semiconductor manufacturing facility is *fab*, and inside of a fab there are complex machines simply called *tools* that perform tasks such as etching and doping of the silicon on microscopic scales. These tools will use techniques such as gaseous, liquid, or solid *diffusion* to disperse dopant material onto exposed silicon at elevated temperatures, or alternatively *ion implantation* using electrically-accelerated dopant ions fired directly into the silicon surface. Either way, what was once intrinsic (pure) silicon will become either P-type of N-type extrinsic silicon per the requirements of the device(s) being manufactured.

[16] It is common in semiconductor literature for dopant concentrations to be specified in units of atoms per cubic centimeter. It is helpful to know that in one cubic centimeter of pure silicon material there exist approximately 5×10^{22} atoms of silicon, based on silicon's atomic weight of 28 (grams per mole) and its mass density of 2.329085 grams per cm^3. Thus, a dopant concentration of 10^{15} cm^{-3} equates to a unitless concentration ratio of $\frac{10^{15}}{5 \times 10^{22}}$ or 20 ppb. A convention used by semiconductor engineers to describe relative dopant concentrations within the layers of a particular semiconductor device is to add $+$ or $-$ superscripts to the letters p and n. For example, p^+ is P-type material with a heavier doping than p^- material. Similarly, n^+ is N-type material that is more heavily doped than n^-. Note how the "+" and "−" superscripts refer to relative dopant concentration levels rather than electrical polarity.

[17] So-called "degenerate" semiconductor material is that which has been so heavily doped that it behaves like a metallic conductor.

Fab tools work on disk-shaped silicon *wafers* that have been sliced from a boule of silicon. In a modern fab these wafers are almost entirely handled by machines rather than by human hands in order to reduce opportunities for contamination of the wafers by bodily residue such as dead skin cells and hair follicles. All persons entering a fab must wear whole-body "bunny suits" covering all but their faces in order to minimize contamination of the fab space by these bodily residues. The entire fab is purged with ultra-pure air filtered to remove the vast majority of particular matter (e.g. dust). For the manufacture of *integrated circuits* where entire functioning circuits consisting of multiple semiconductor devices are etched onto a small area of silicon, the wafer is divided into many rectangular *chips* for each of these circuits. After electrical testing, the chips are cut apart from each other on the wafer using a diamond saw where any rejected chips may be scrapped and the accepted chips harvested for packaging. Below we see a wafer after the rejection process, where a rectangular void marks the former place of each rejected memory-circuit chip on the 300 millimeter diameter wafer:



Chapter 3

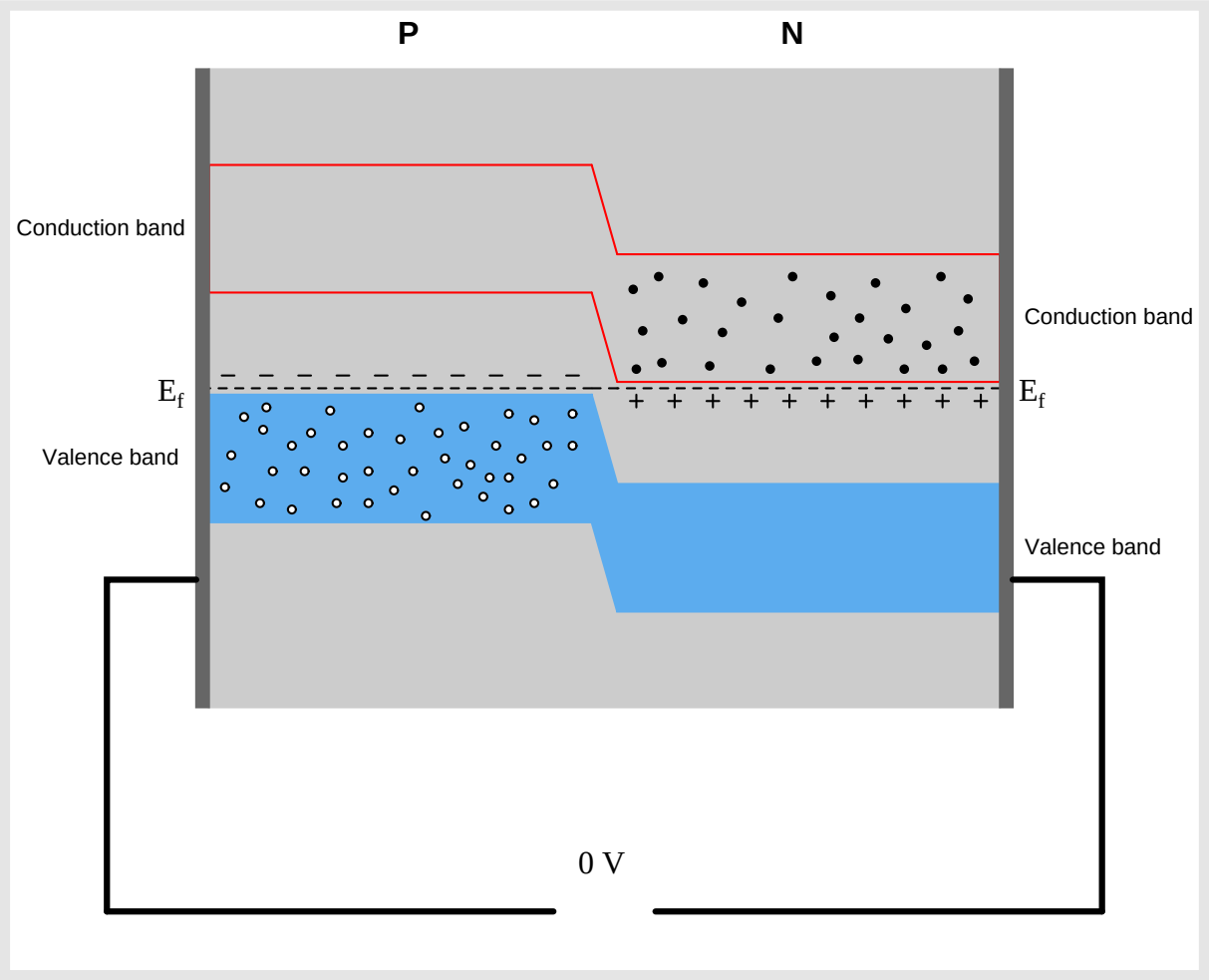
Animations

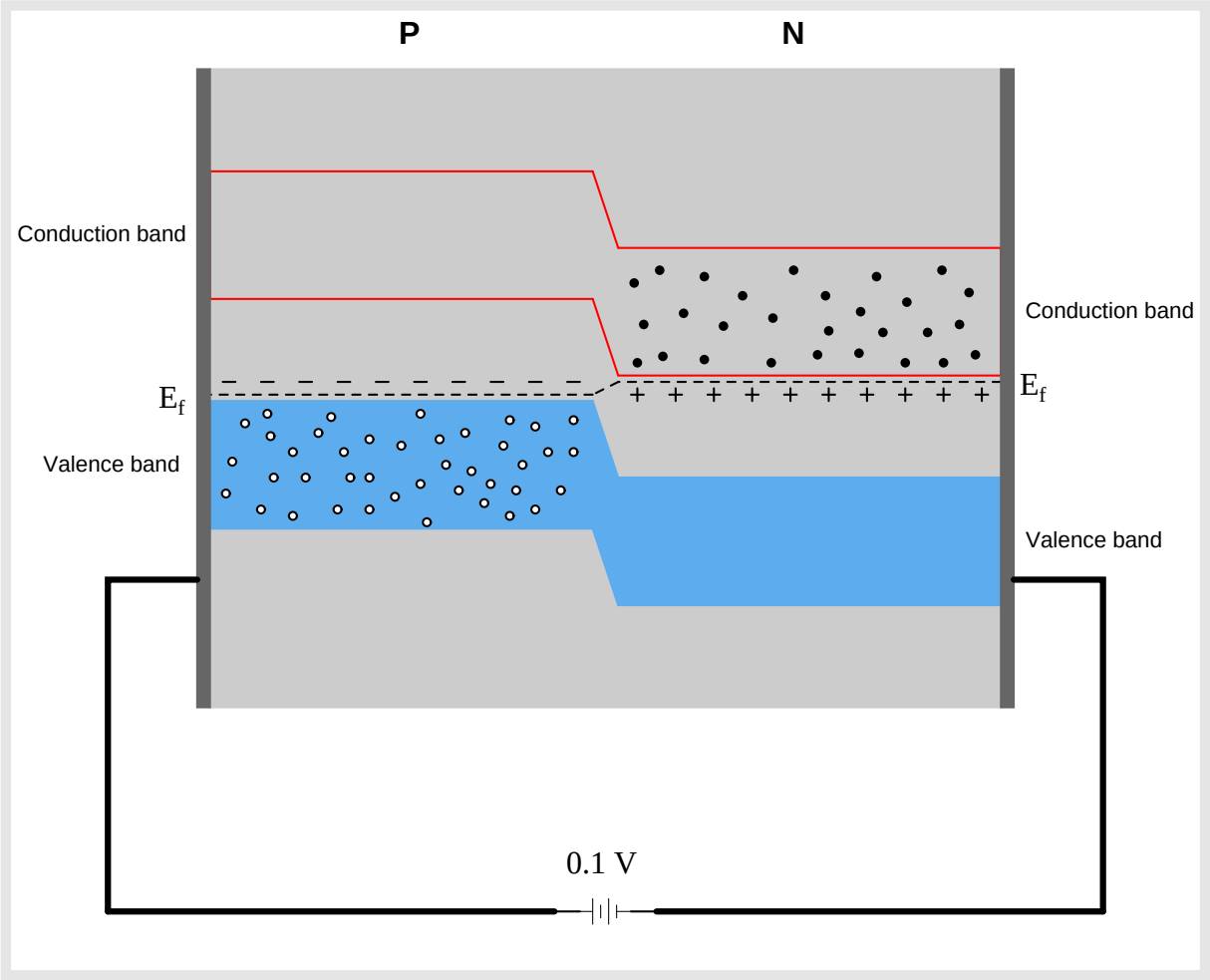
Some concepts are much easier to grasp when seen in *action*. A simple yet effective form of animation suitable to an electronic document such as this is a "flip-book" animation where a set of pages in the document show successive frames of a simple animation. Such "flip-book" animations are designed to be viewed by paging forward (and/or back) with the document-reading software application, watching it frame-by-frame. Unlike video which may be difficult to pause at certain moments, "flip-book" animations lend themselves very well to individual frame viewing.

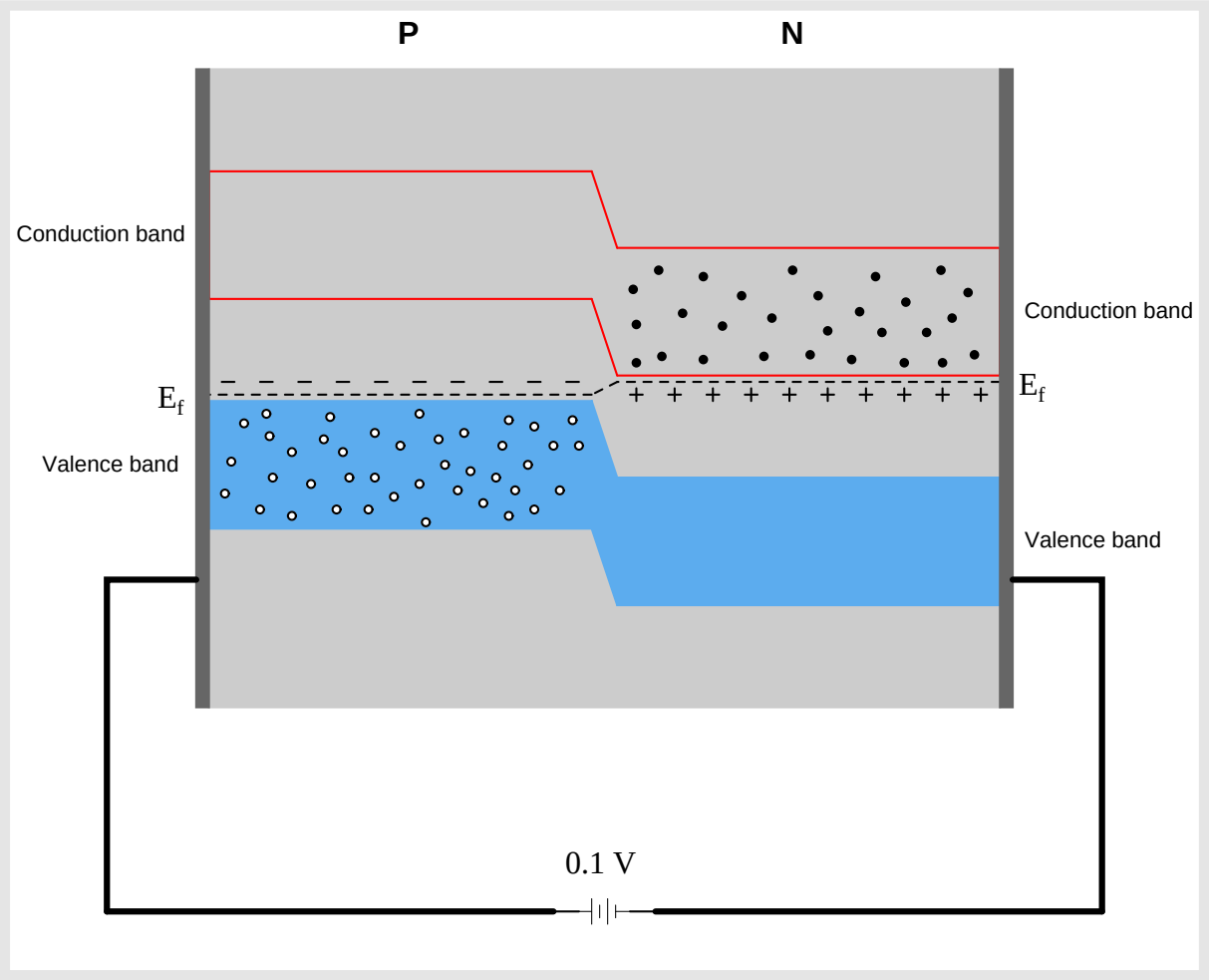
3.1 Animation of a forward-biased PN diode junction

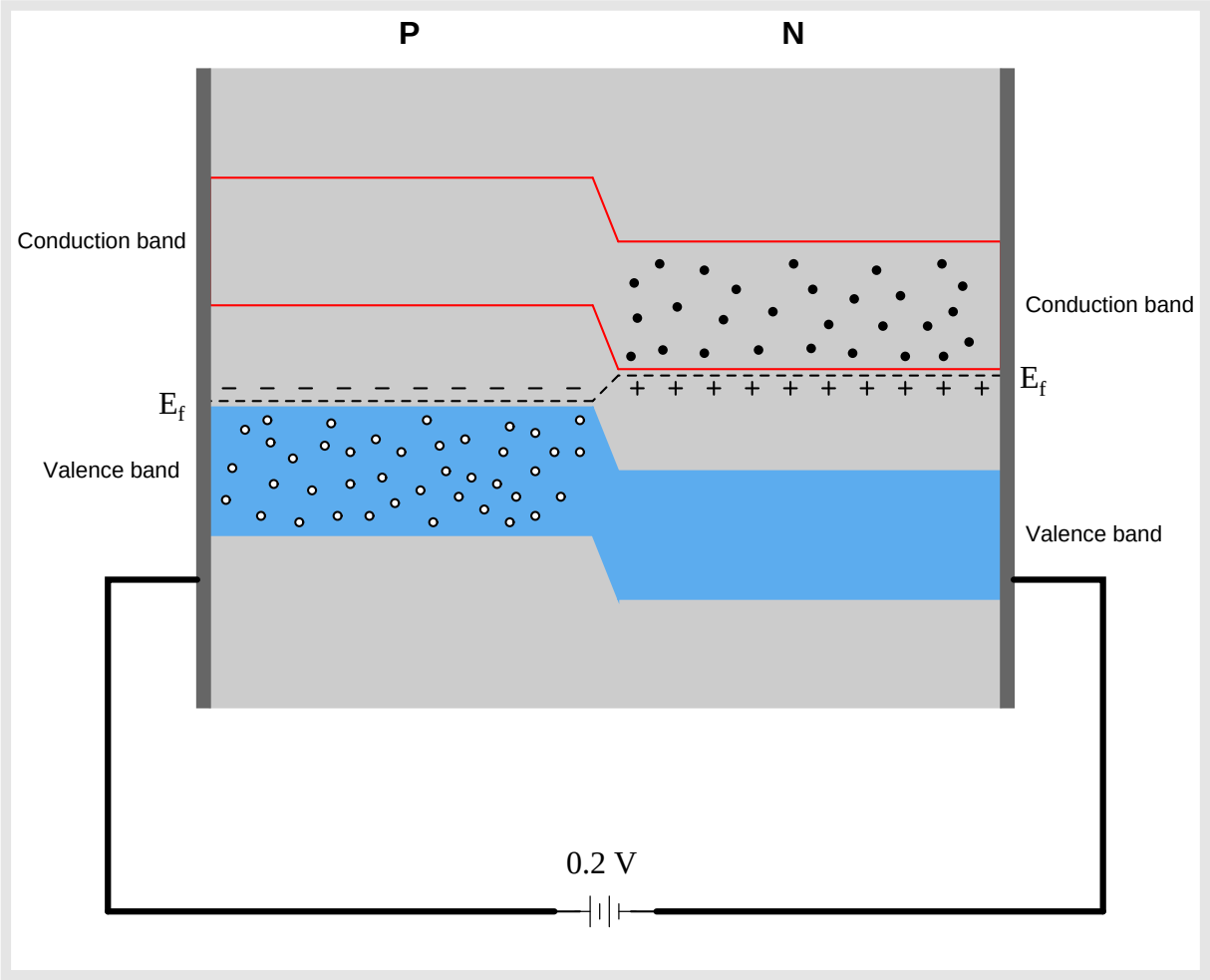
The following animation illustrates the behavior of mobile[1] charge carriers inside a semiconductor diode junction. A DC voltage source provides forward-biasing voltage to the PN junction, and this voltage source's value is progressively raised until the diode begins to conduct. Energy levels within the diode's semiconductor halves are shown relative to one another by means of height in the illustration.

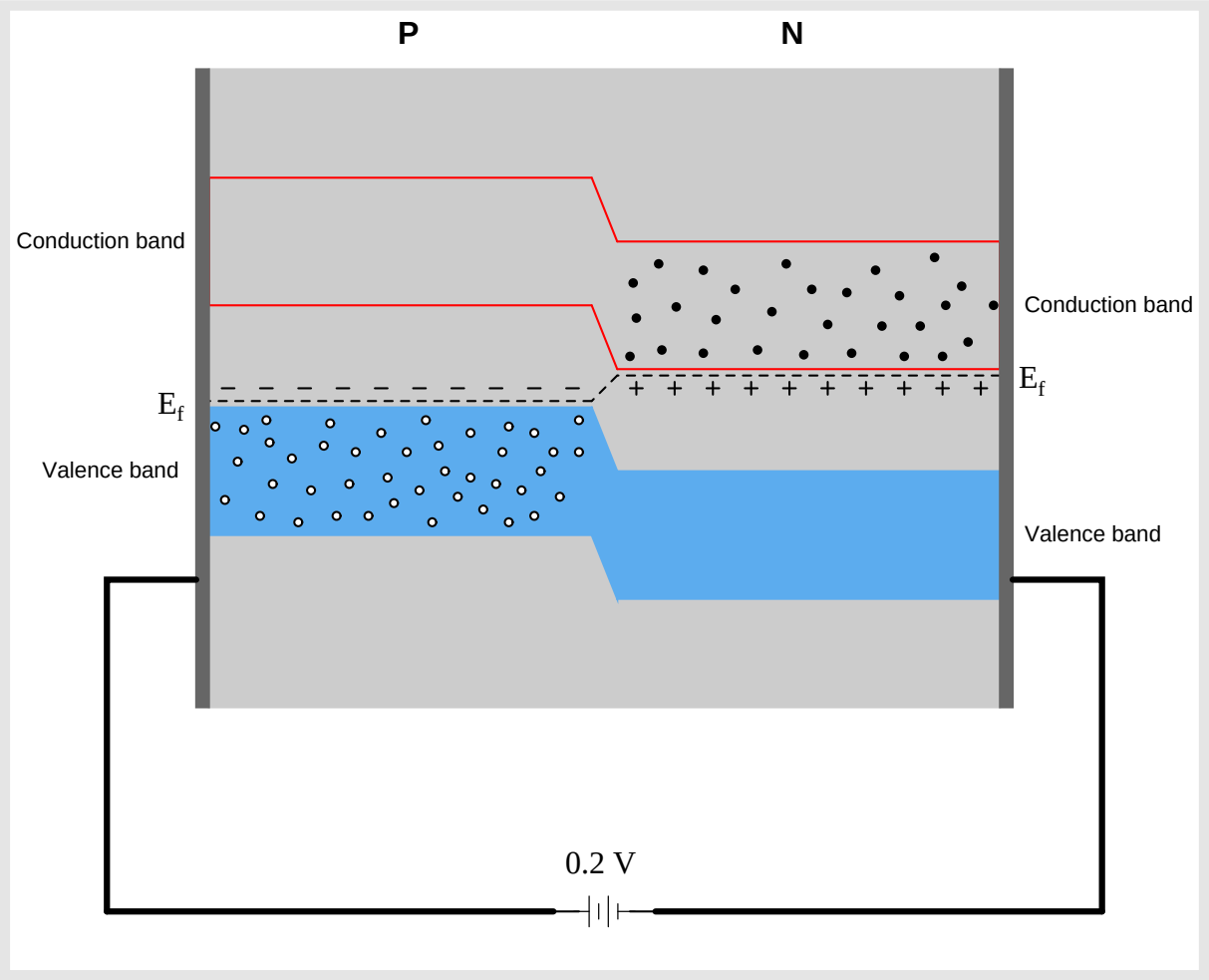
[1] In this animation, immobile electrons in the valence band are not shown – only the holes in that band.

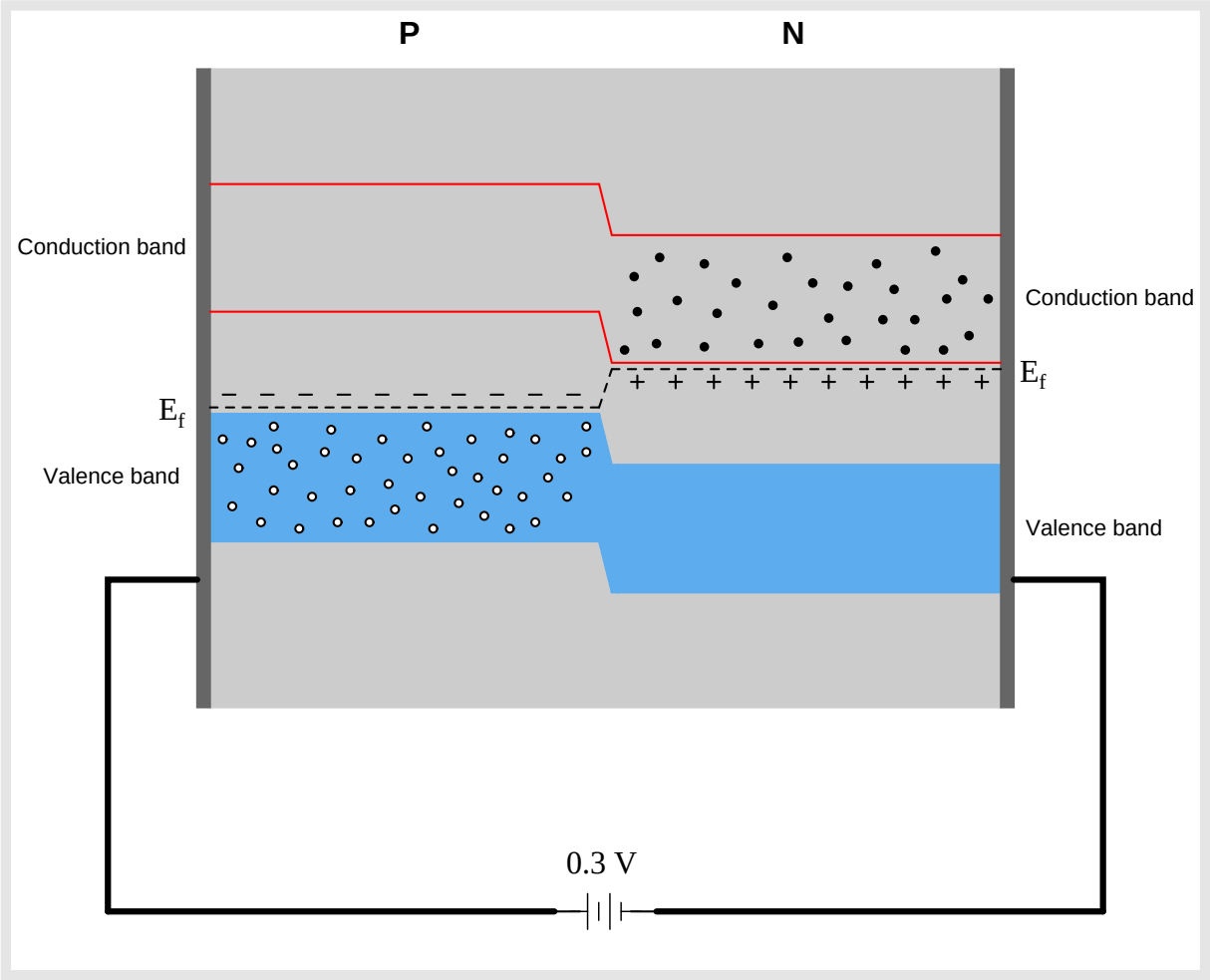

P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0 V

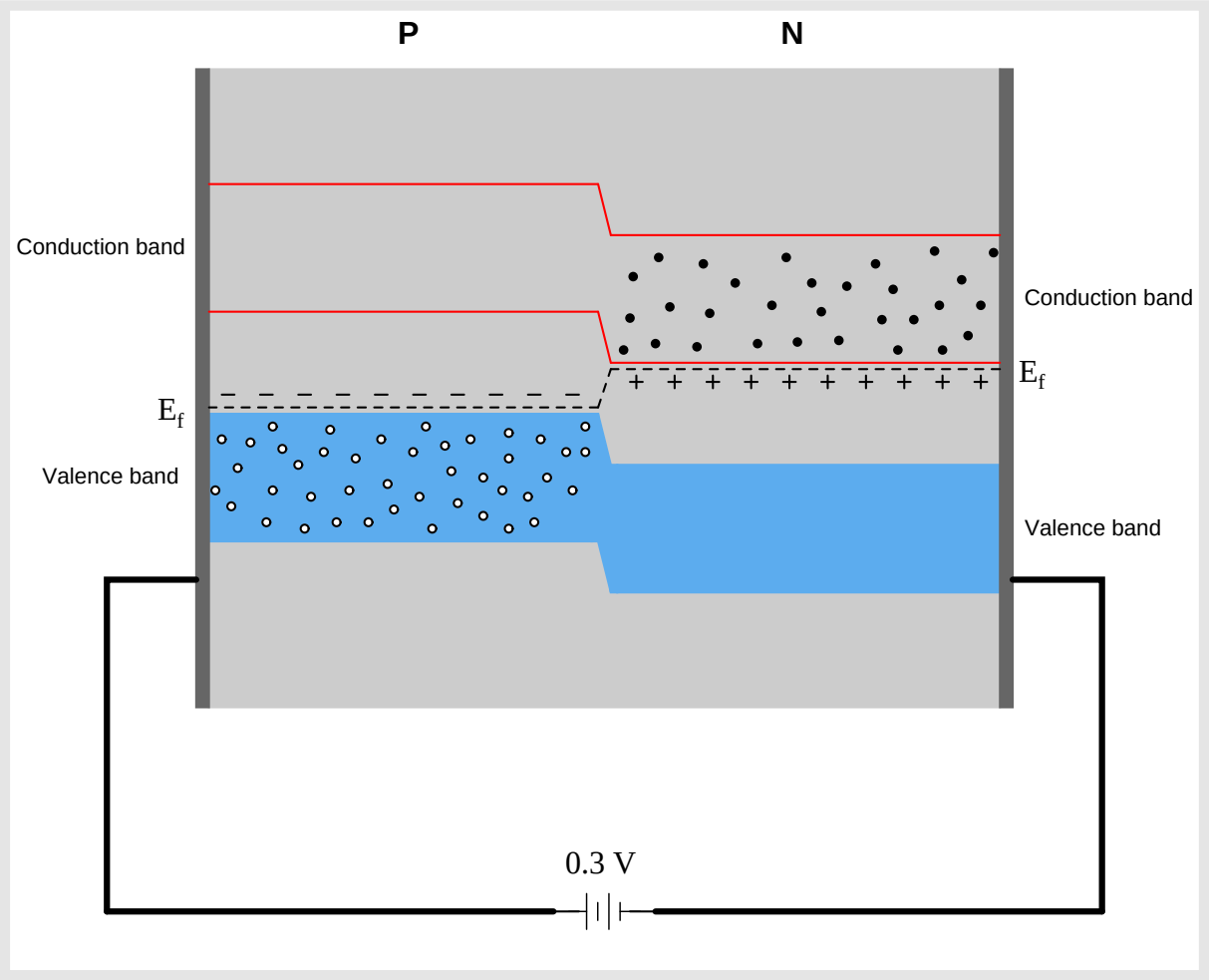

P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.1 V

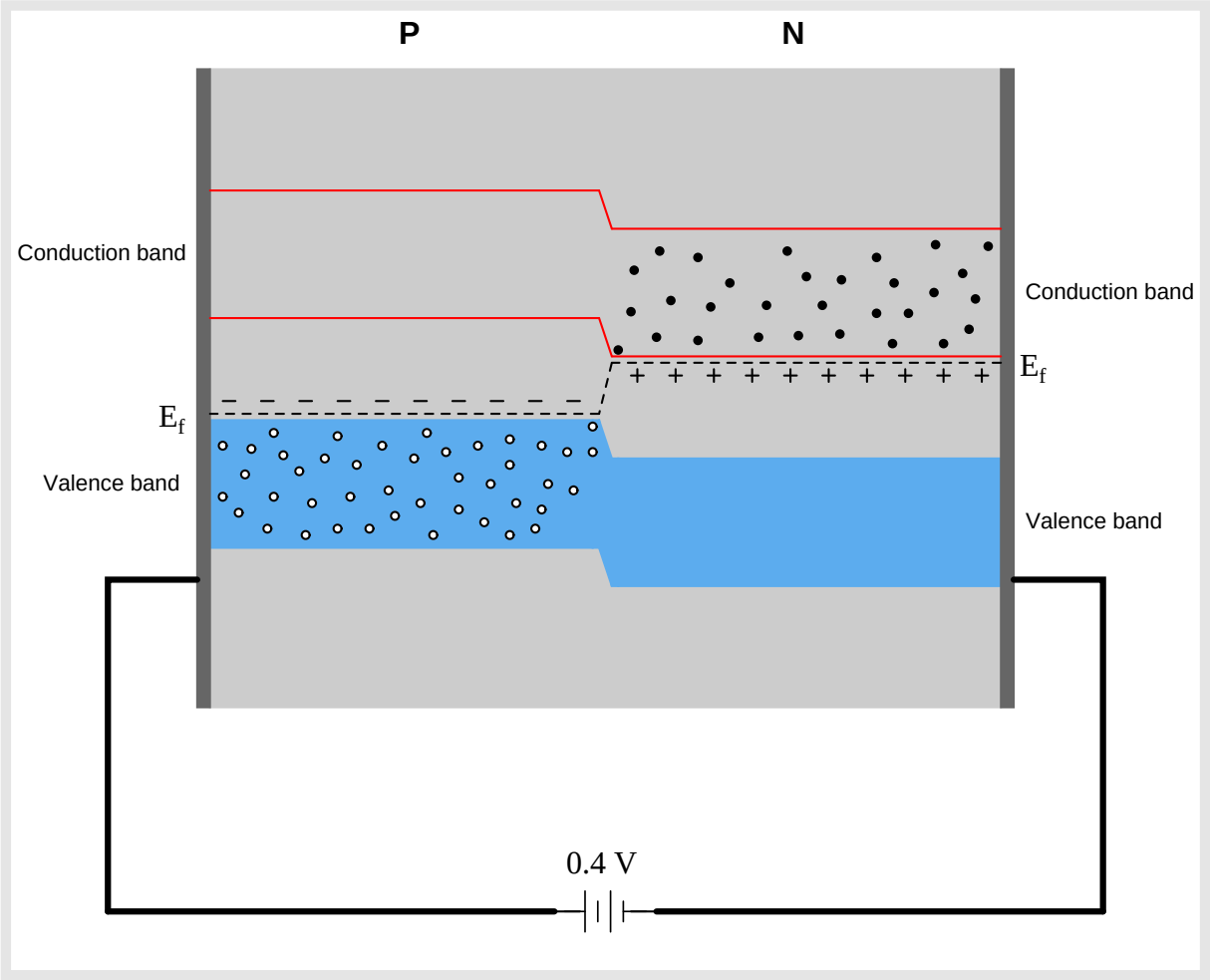

P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.1 V

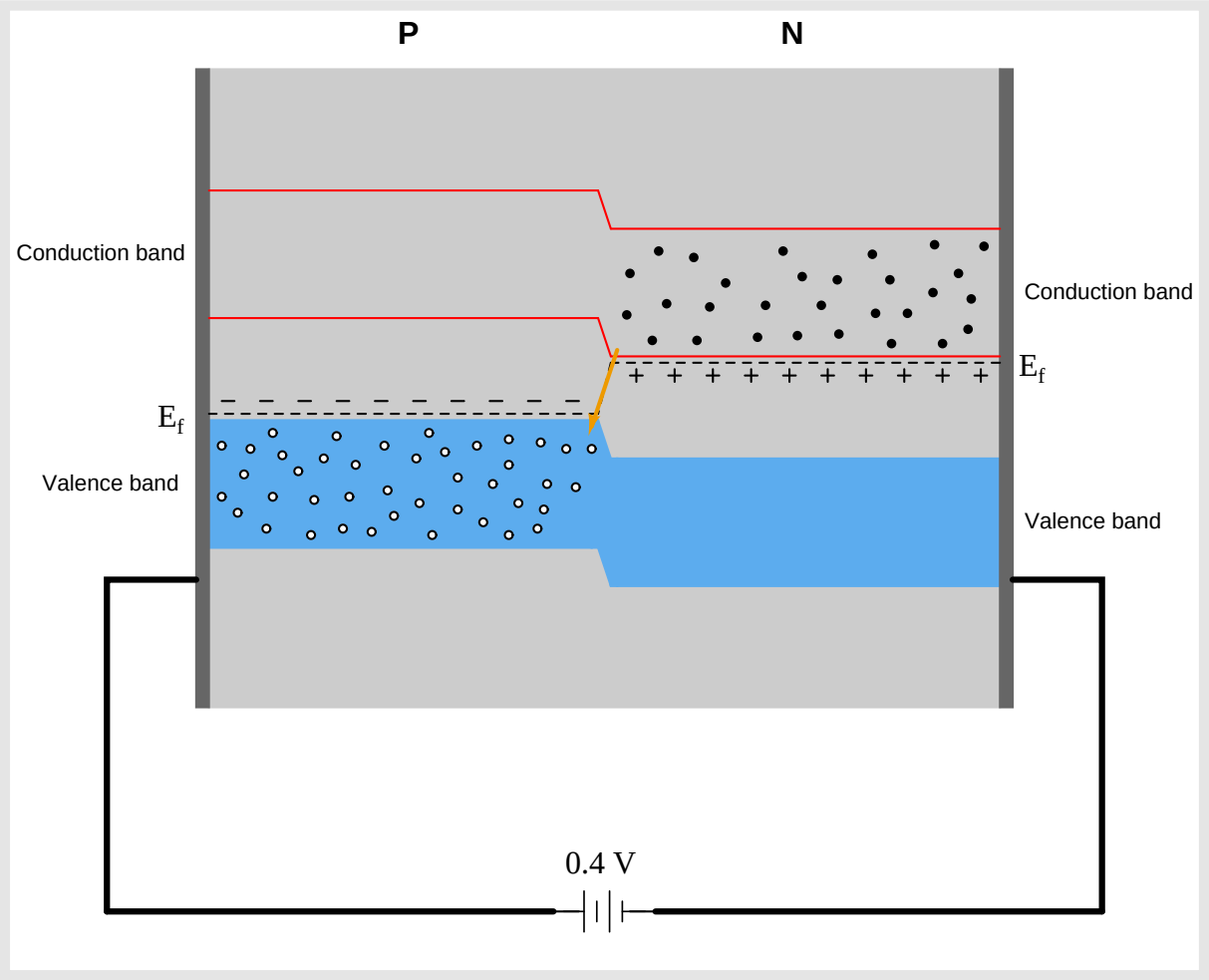

P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.2 V

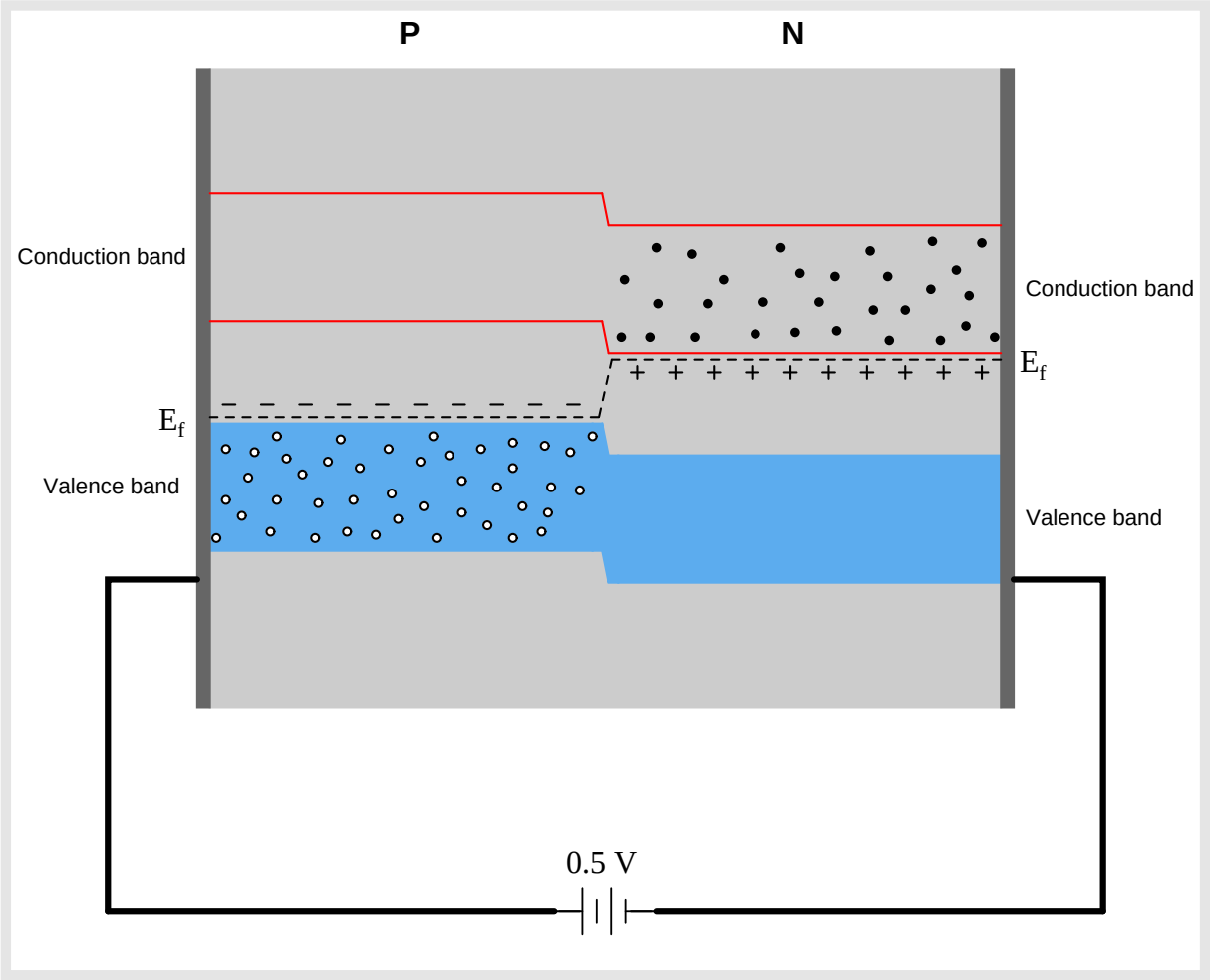

P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.2 V

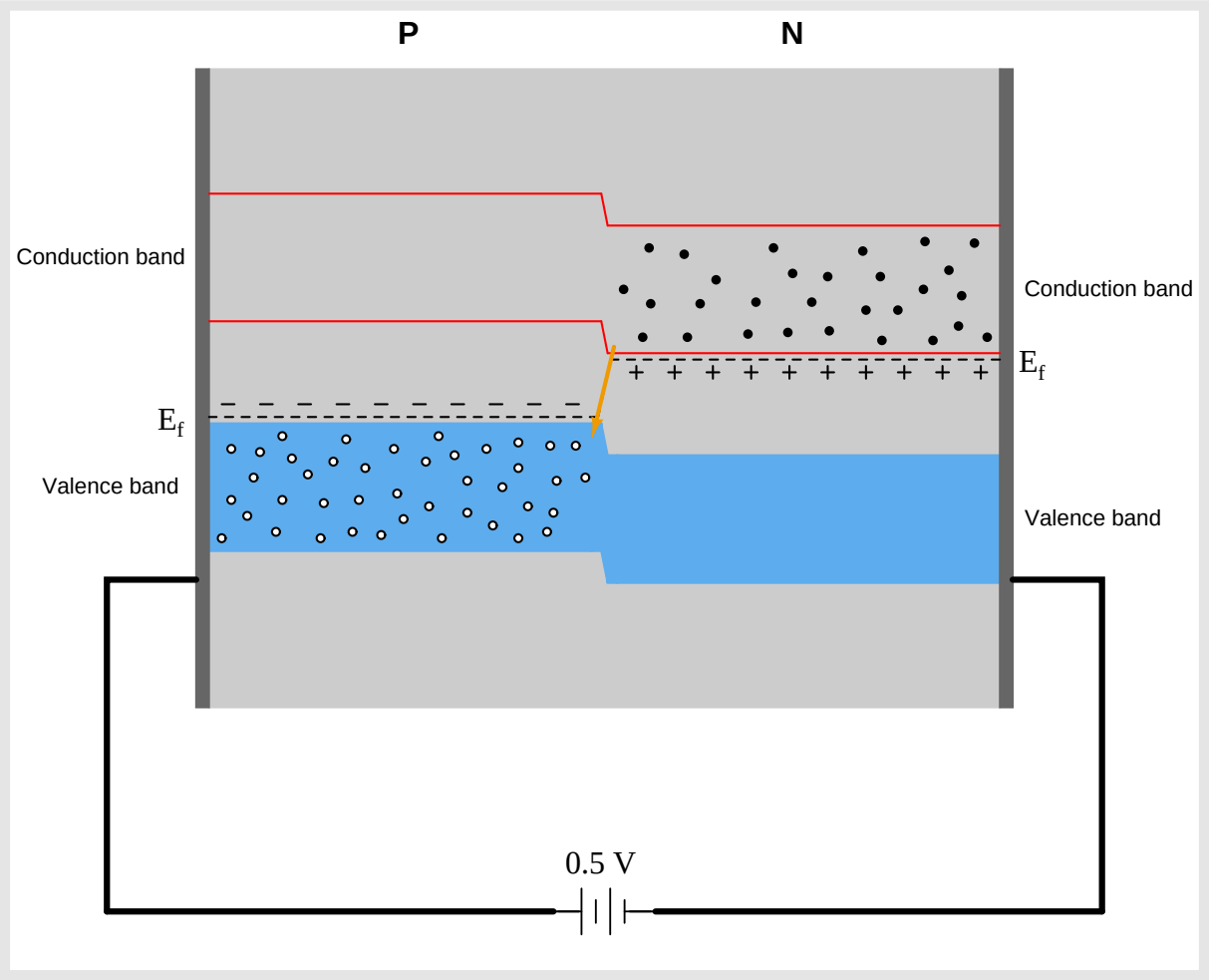

P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.3 V

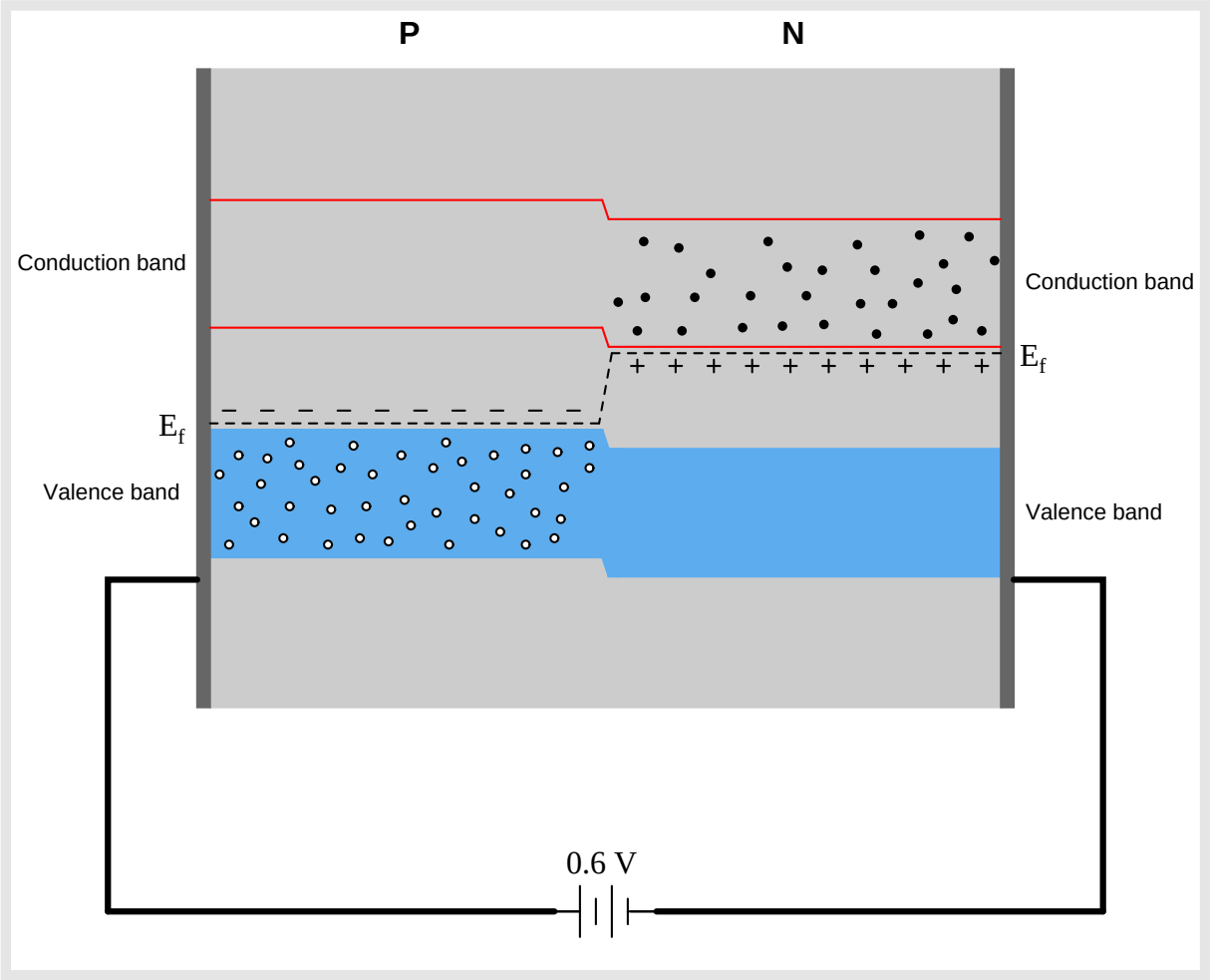

P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.3 V

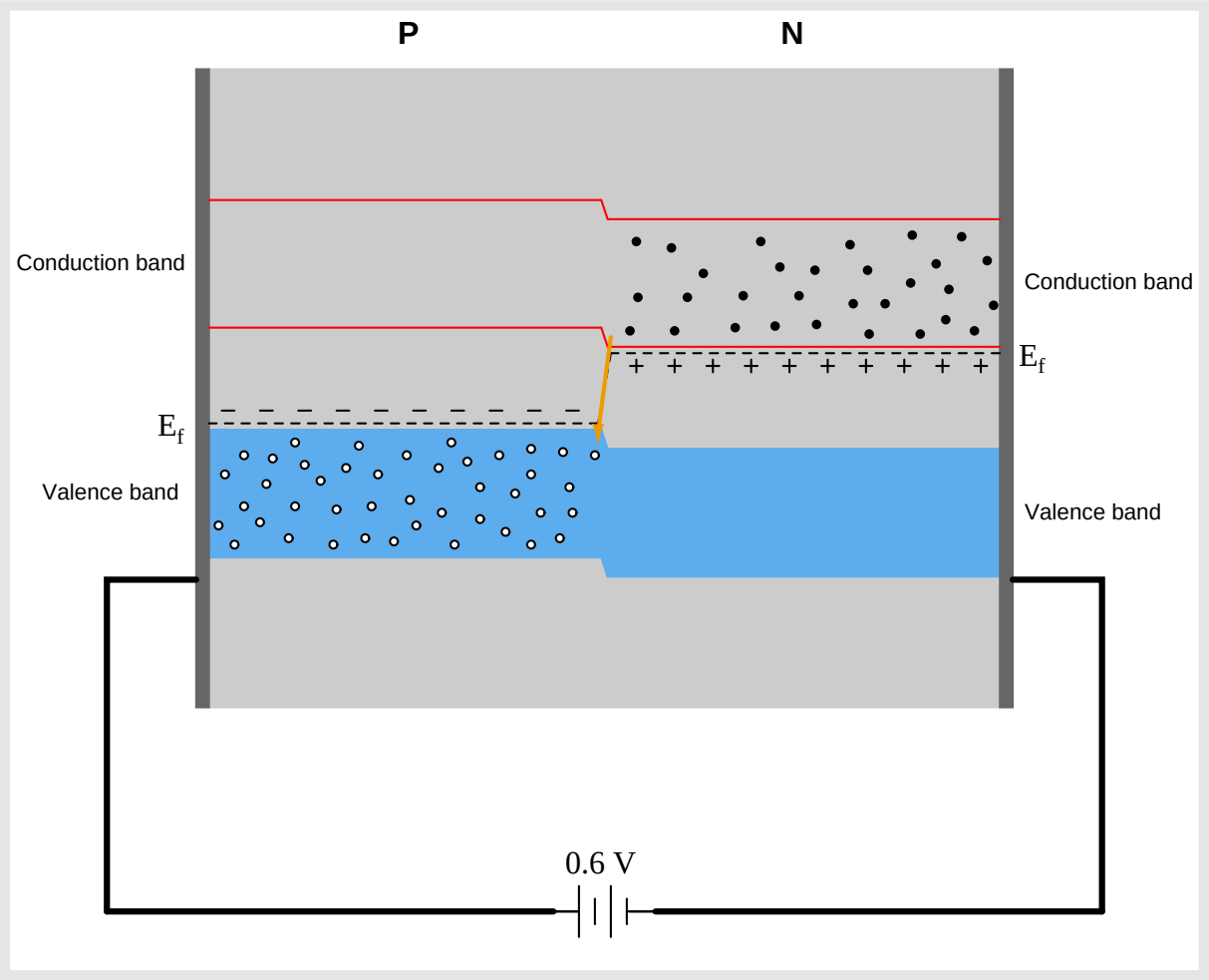

P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.4 V

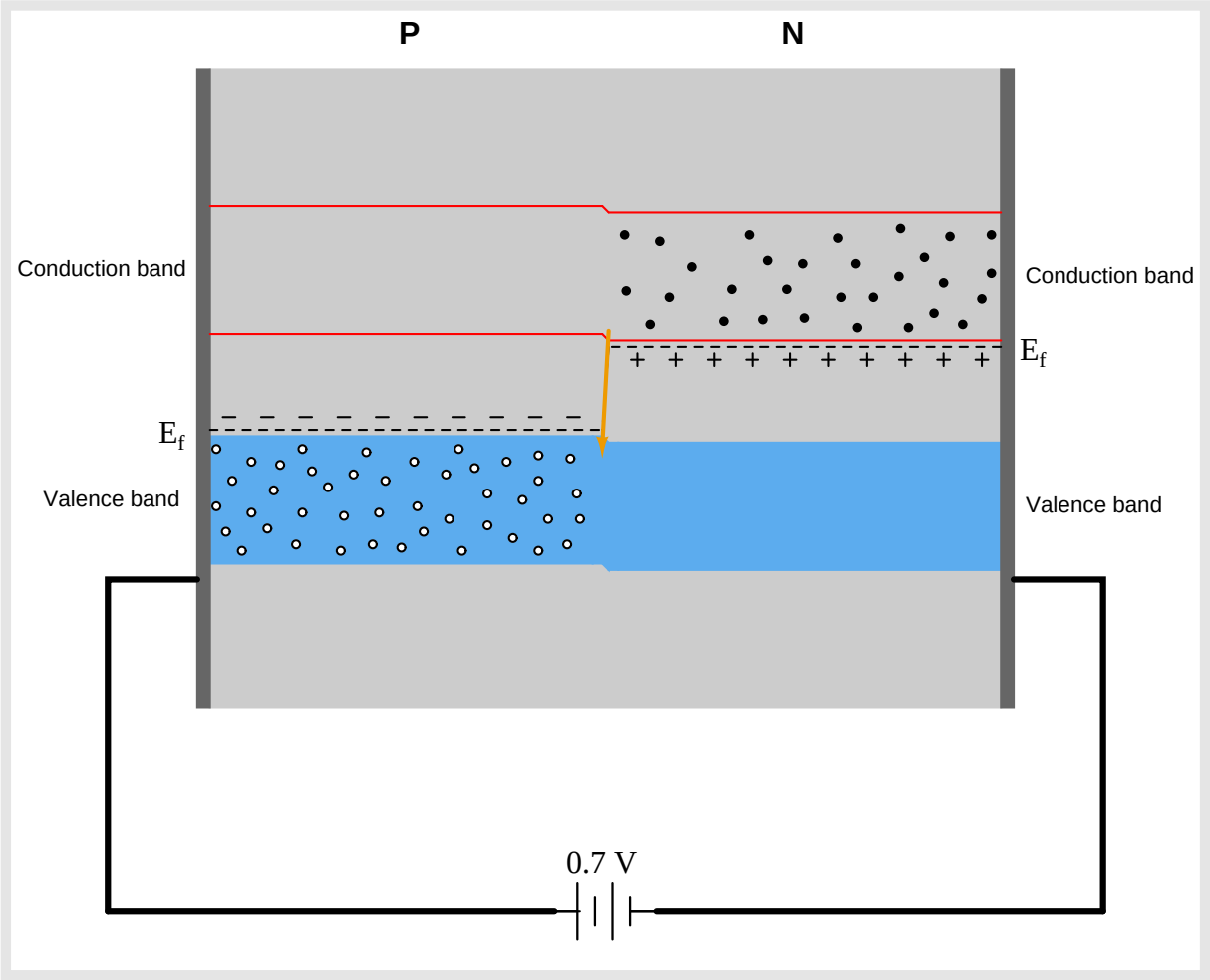

P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.4 V

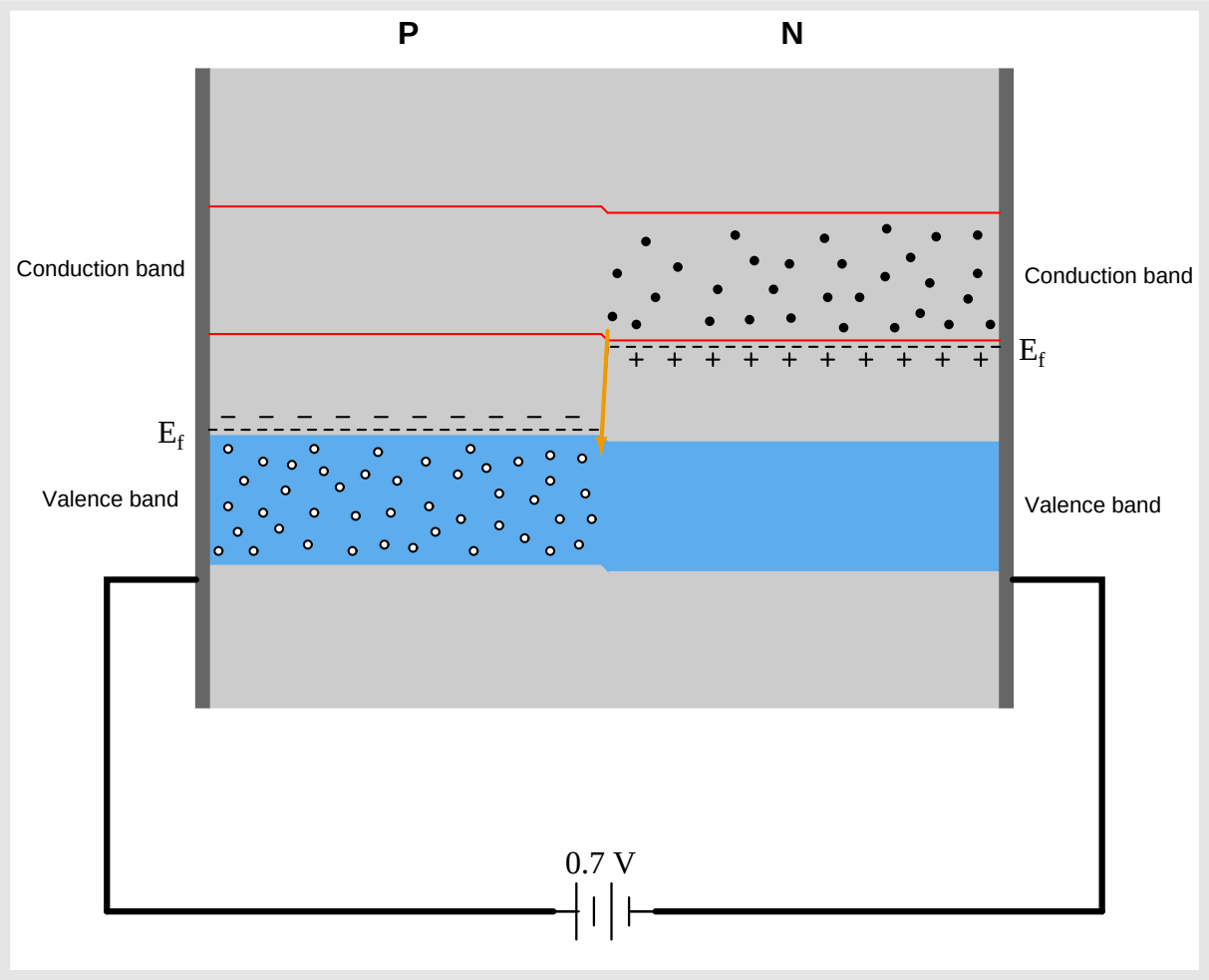

P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.5 V

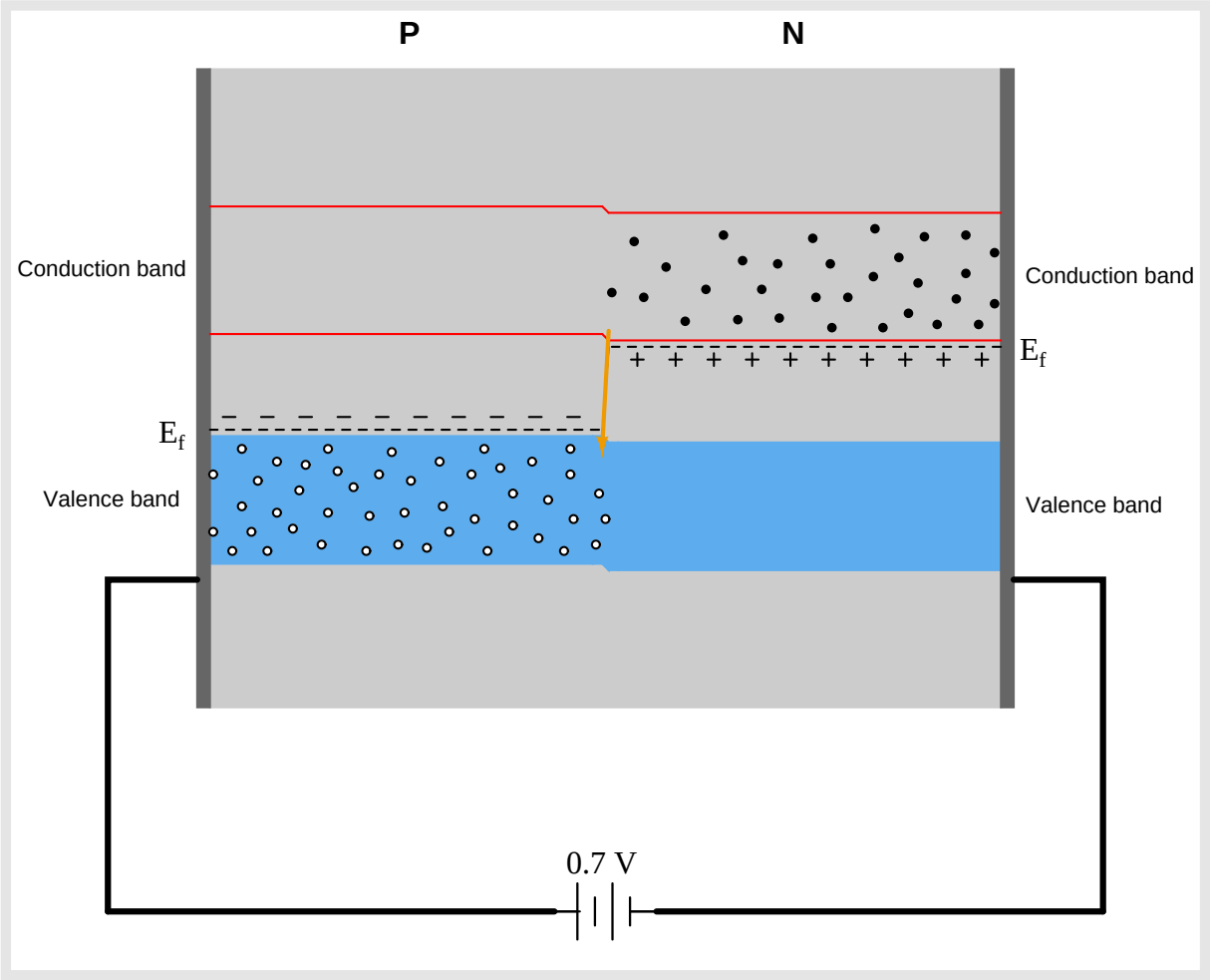

P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.5 V

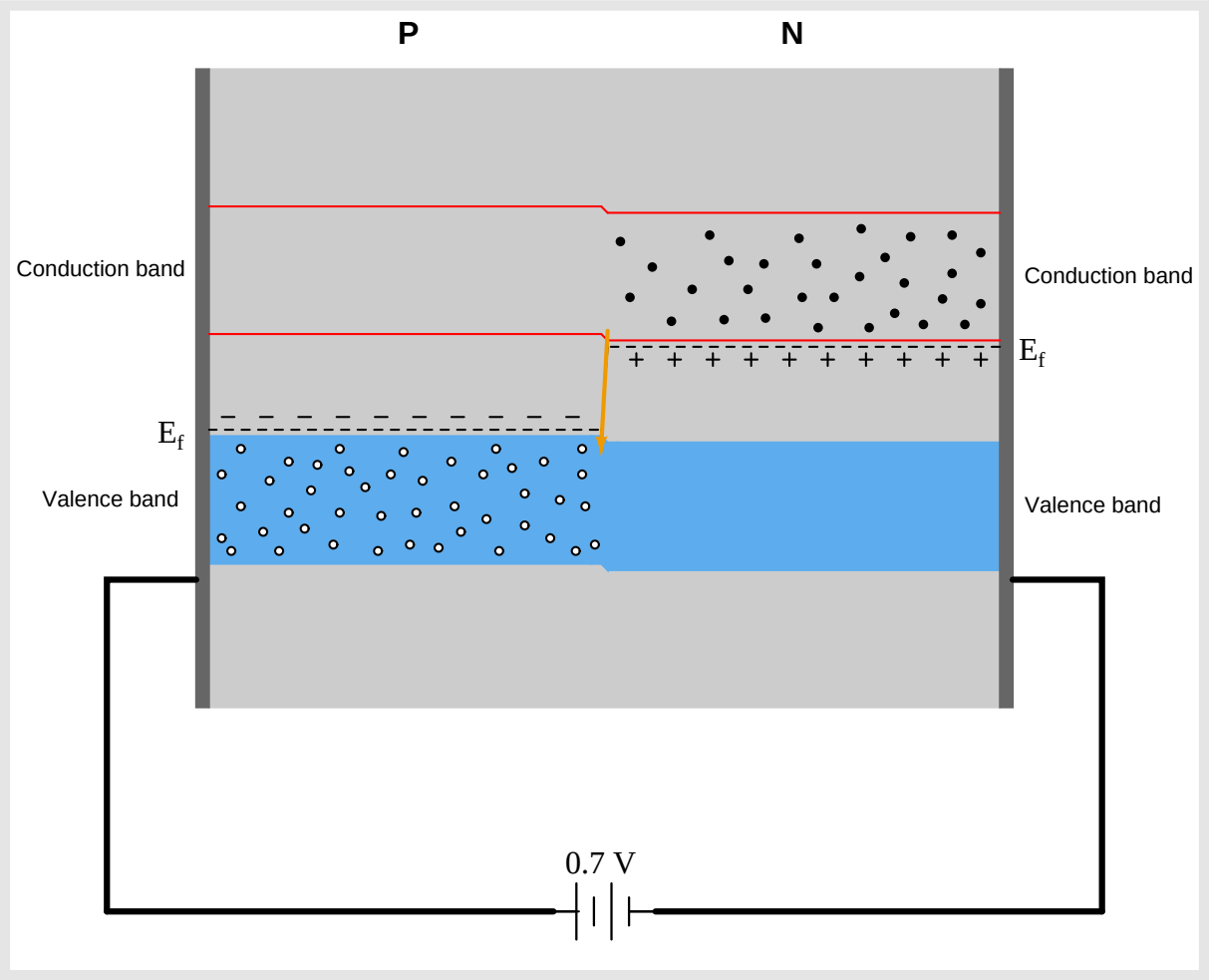

P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.6 V

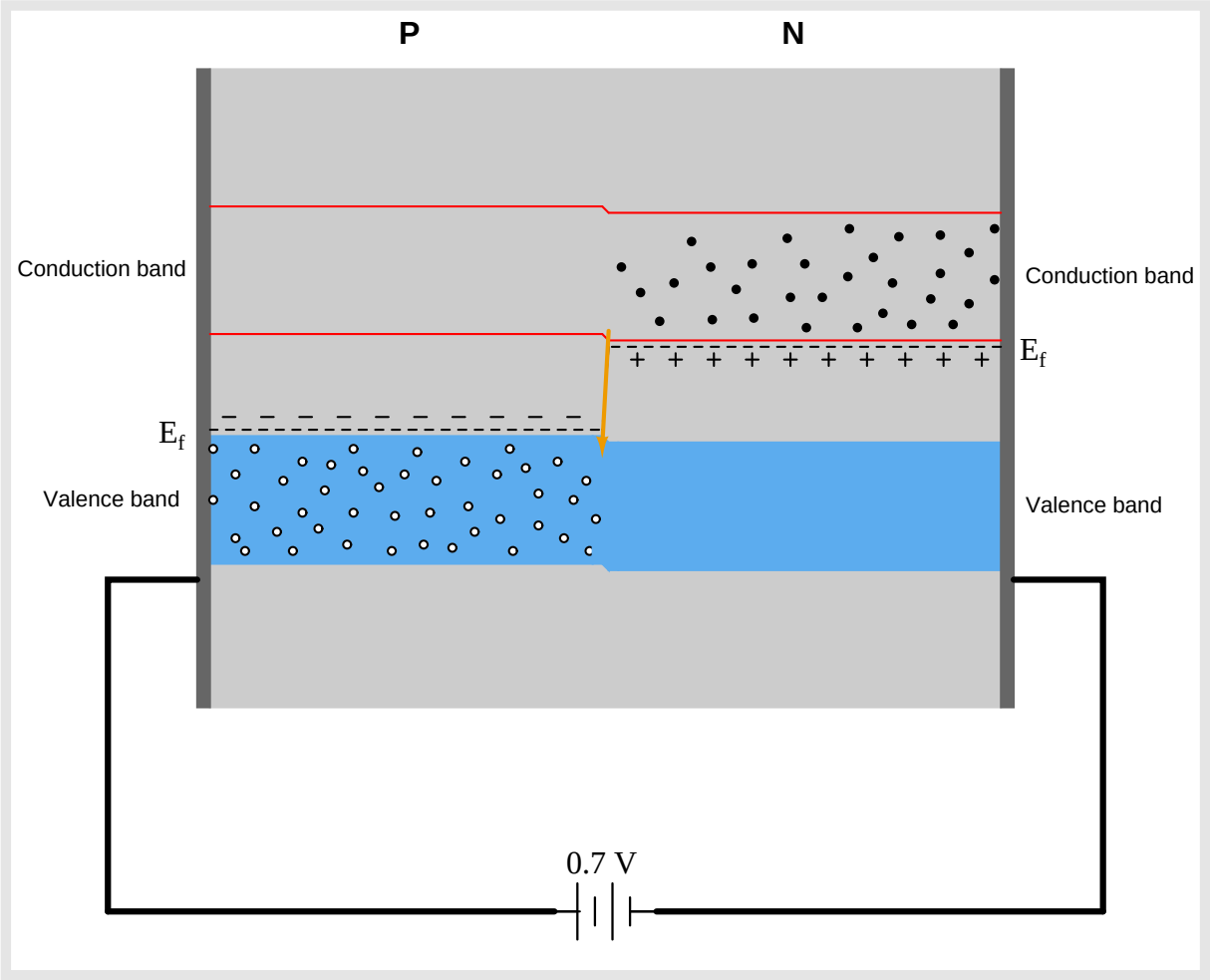

P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.6 V

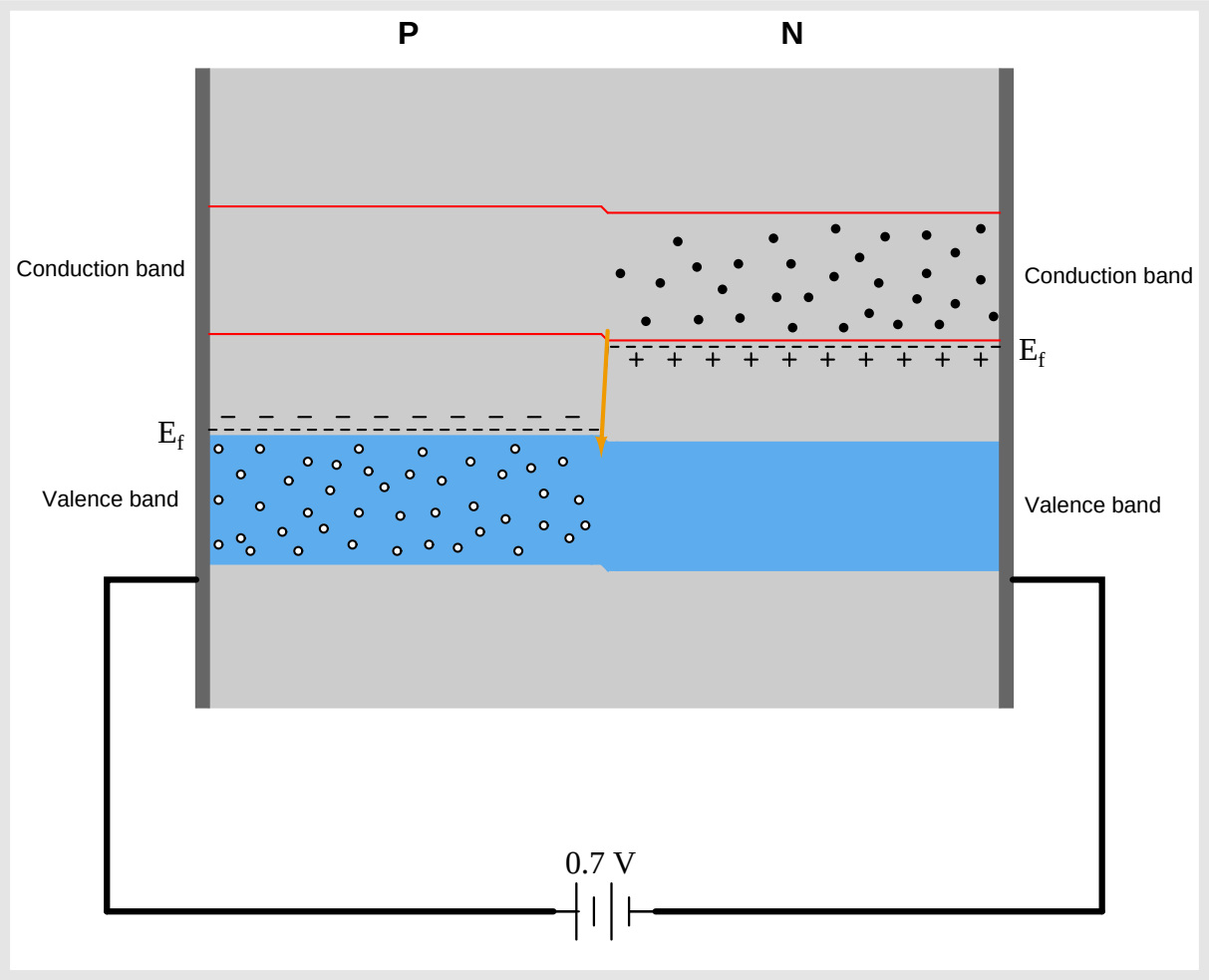

P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.7 V

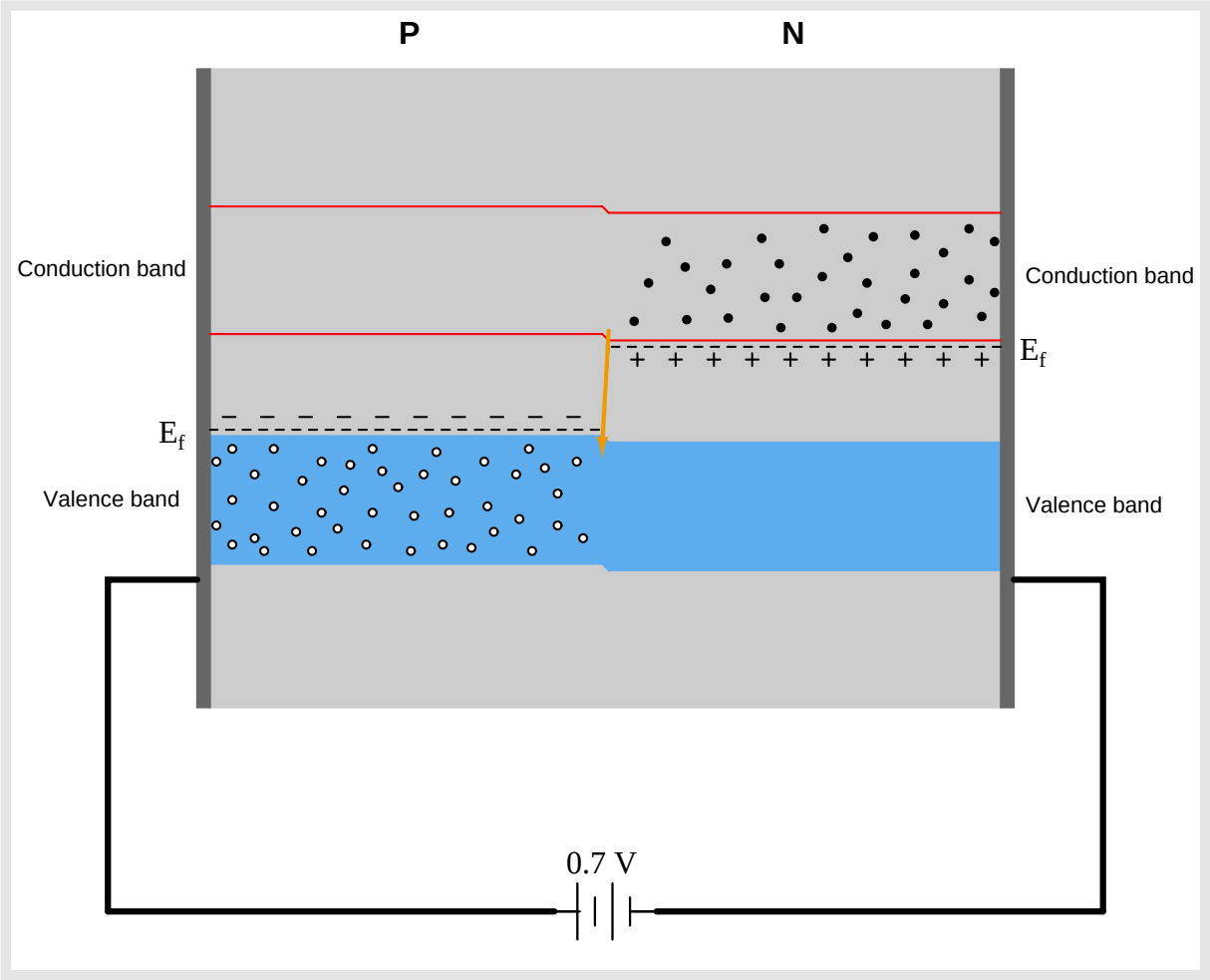

P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.7 V

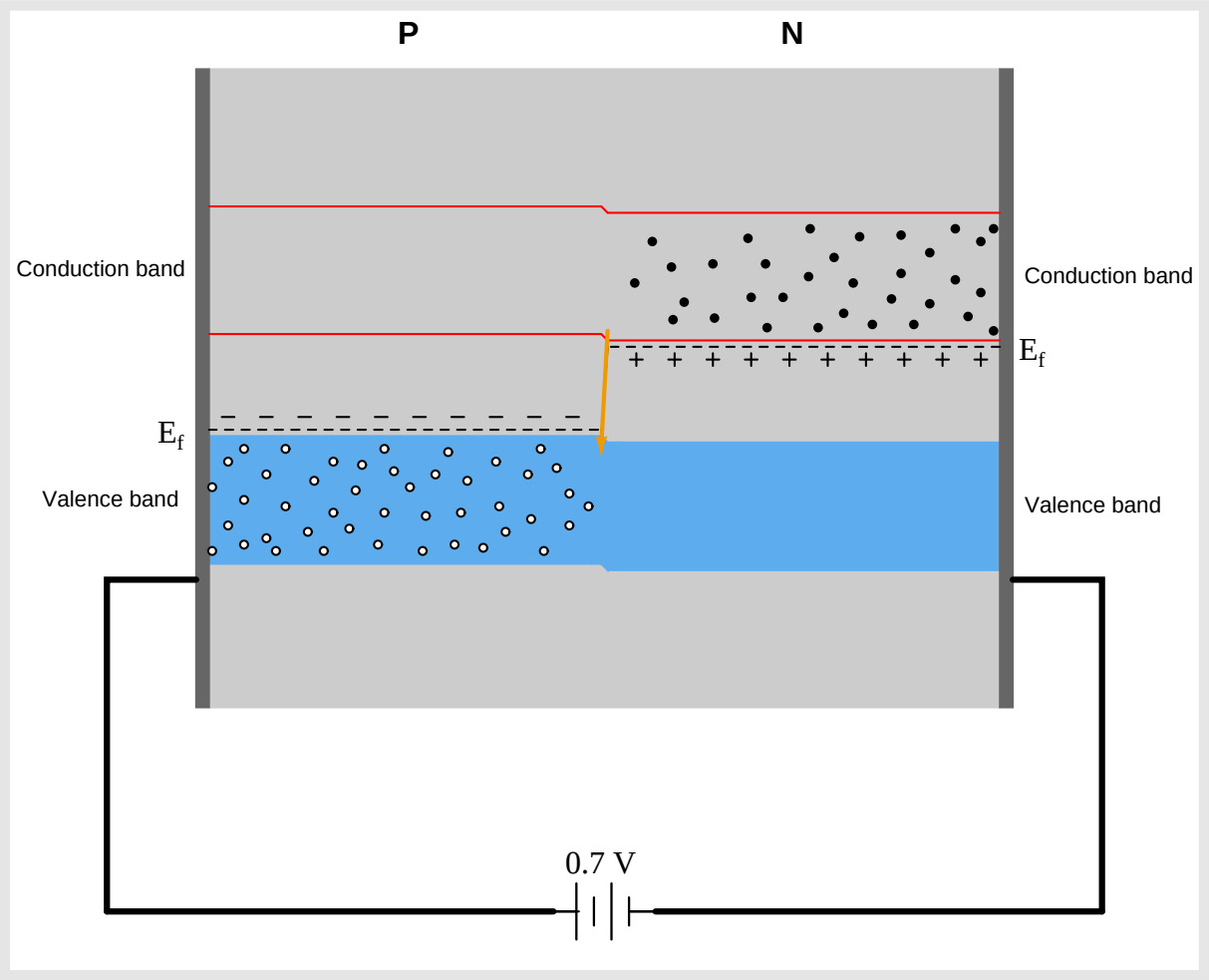

P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.7 V

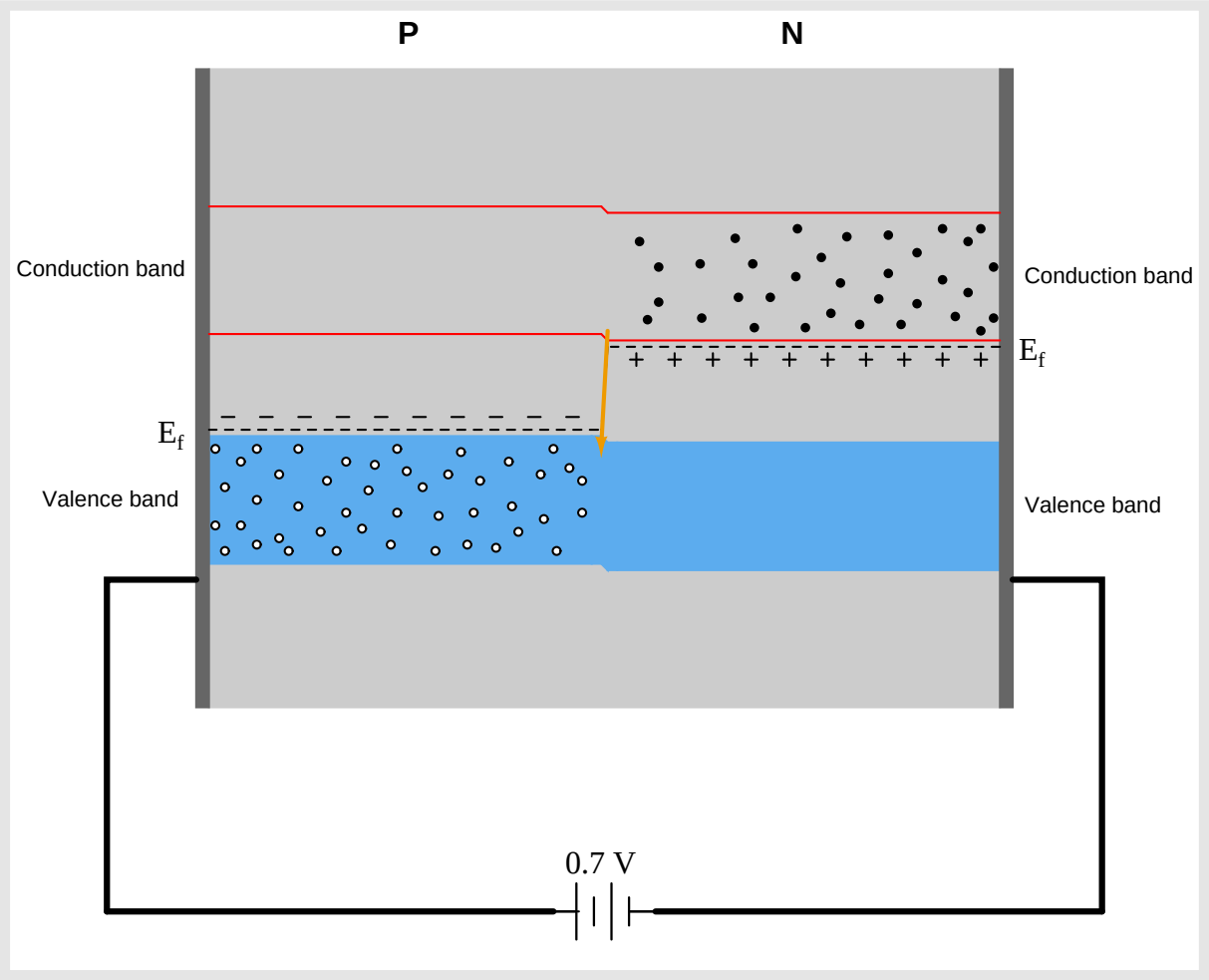

P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.7 V

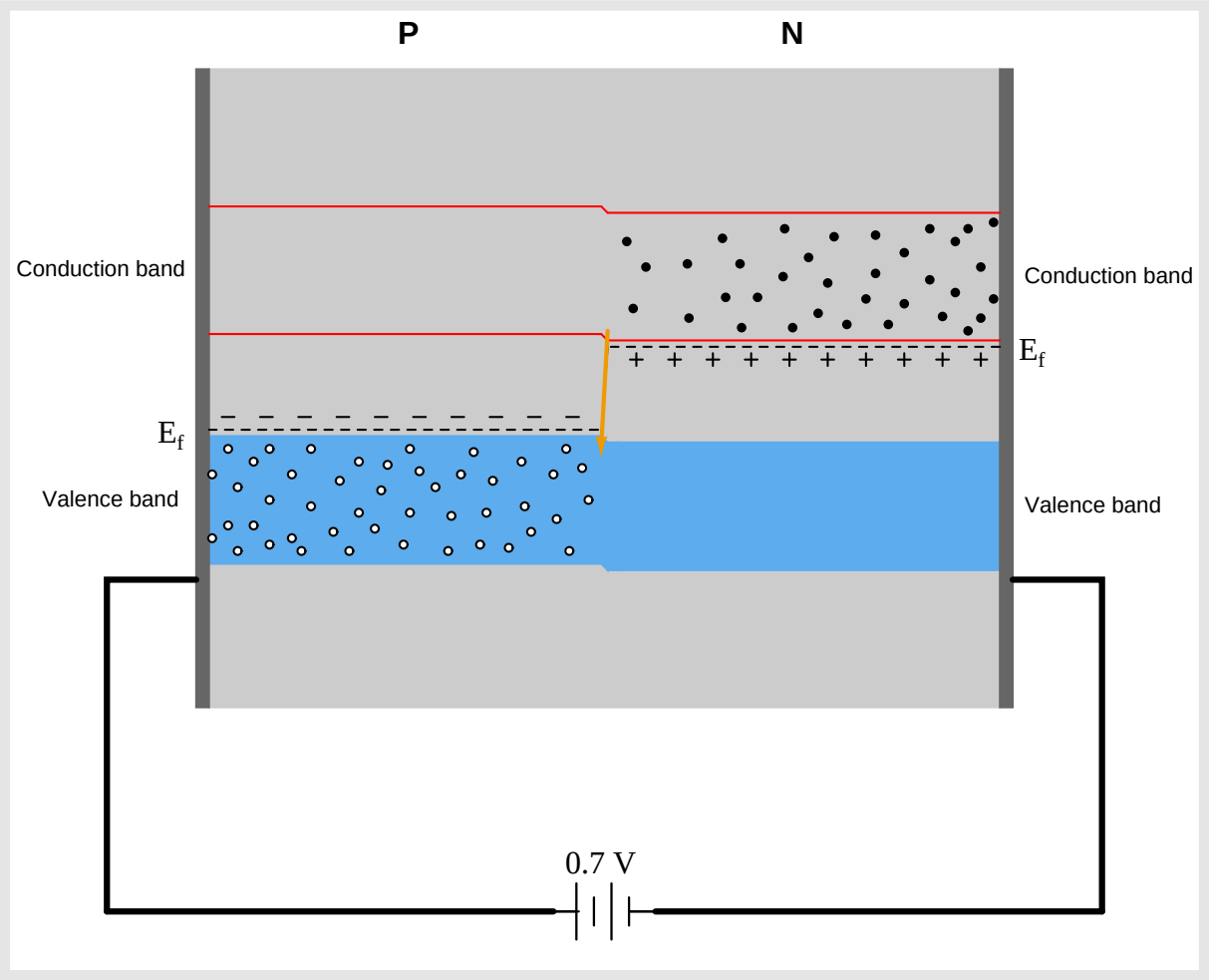

P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.7 V


P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.7 V


P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.7 V


P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.7 V


P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.7 V


P
N
Conduction band
Conduction band
E_f
E_f
Valence band
Valence band
0.7 V

Chapter 4

Questions

This learning module, along with all others in the ModEL collection, is designed to be used in an inverted instructional environment where students independently read[1] the tutorials and attempt to answer questions on their own *prior* to the instructor's interaction with them. In place of lecture[2], the instructor engages with students in Socratic-style dialogue, probing and challenging their understanding of the subject matter through inquiry.

Answers are not provided for questions within this chapter, and this is by design. Solved problems may be found in the Tutorial and Derivation chapters, instead. The goal here is *independence*, and this requires students to be challenged in ways where others cannot think for them. Remember that you always have the tools of *experimentation* and *computer simulation* (e.g. SPICE) to explore concepts!

The following lists contain ideas for Socratic-style questions and challenges. Upon inspection, one will notice a strong theme of *metacognition* within these statements: they are designed to foster a regular habit of examining one's own thoughts as a means toward clearer thinking. As such these sample questions are useful both for instructor-led discussions as well as for self-study.

[1] Technical reading is an essential academic skill for any technical practitioner to possess for the simple reason that the most comprehensive, accurate, and useful information to be found for developing technical competence is in textual form. Technical careers in general are characterized by the need for continuous learning to remain current with standards and technology, and therefore any technical practitioner who cannot read well is handicapped in their professional development. An excellent resource for educators on improving students' reading prowess through intentional effort and strategy is the book textitReading For Understanding – How Reading Apprenticeship Improves Disciplinary Learning in Secondary and College Classrooms by Ruth Schoenbach, Cynthia Greenleaf, and Lynn Murphy.

[2] Lecture is popular as a teaching method because it is easy to implement: any reasonably articulate subject matter expert can talk to students, even with little preparation. However, it is also quite problematic. A good lecture always makes complicated concepts seem easier than they are, which is bad for students because it instills a false sense of confidence in their own understanding; reading and re-articulation requires more cognitive effort and serves to verify comprehension. A culture of teaching-by-lecture fosters a debilitating dependence upon direct personal instruction, whereas the challenges of modern life demand independent and critical thought made possible only by gathering information and perspectives from afar. Information presented in a lecture is ephemeral, easily lost to failures of memory and dictation; text is forever, and may be referenced at any time.

General challenges following tutorial reading

- Summarize as much of the text as you can in one paragraph of your own words. A helpful strategy is to explain ideas as you would for an intelligent child: as simple as you can without compromising too much accuracy.
- Simplify a particular section of the text, for example a paragraph or even a single sentence, so as to capture the same fundamental idea in fewer words.
- Where did the text make the most sense to you? What was it about the text's presentation that made it clear?
- Identify where it might be easy for someone to misunderstand the text, and explain why you think it could be confusing.
- Identify any new concept(s) presented in the text, and explain in your own words.
- Identify any familiar concept(s) such as physical laws or principles applied or referenced in the text.
- Devise a proof of concept experiment demonstrating an important principle, physical law, or technical innovation represented in the text.
- Devise an experiment to disprove a plausible misconception.
- Did the text reveal any misconceptions you might have harbored? If so, describe the misconception(s) and the reason(s) why you now know them to be incorrect.
- Describe any useful problem-solving strategies applied in the text.
- Devise a question of your own to challenge a reader's comprehension of the text.

General follow-up challenges for assigned problems

- Identify where any fundamental laws or principles apply to the solution of this problem, especially before applying any mathematical techniques.
- Devise a thought experiment to explore the characteristics of the problem scenario, applying known laws and principles to mentally model its behavior.
- Describe in detail your own strategy for solving this problem. How did you identify and organized the given information? Did you sketch any diagrams to help frame the problem?
- Is there more than one way to solve this problem? Which method seems best to you?
- Show the work you did in solving this problem, even if the solution is incomplete or incorrect.
- What would you say was the most challenging part of this problem, and why was it so?
- Was any important information missing from the problem which you had to research or recall?
- Was there any extraneous information presented within this problem? If so, what was it and why did it not matter?
- Examine someone else's solution to identify where they applied fundamental laws or principles.
- Simplify the problem from its given form and show how to solve this simpler version of it. Examples include eliminating certain variables or conditions, altering values to simpler (usually whole) numbers, applying a limiting case (i.e. altering a variable to some extreme or ultimate value).
- For quantitative problems, identify the real-world meaning of all intermediate calculations: their units of measurement, where they fit into the scenario at hand. Annotate any diagrams or illustrations with these calculated values.
- For quantitative problems, try approaching it qualitatively instead, thinking in terms of "increase" and "decrease" rather than definite values.
- For qualitative problems, try approaching it quantitatively instead, proposing simple numerical values for the variables.
- Were there any assumptions you made while solving this problem? Would your solution change if one of those assumptions were altered?
- Identify where it would be easy for someone to go astray in attempting to solve this problem.
- Formulate your own problem based on what you learned solving this one.

General follow-up challenges for experiments or projects

- In what way(s) was this experiment or project easy to complete?
- Identify some of the challenges you faced in completing this experiment or project.

- Show how thorough documentation assisted in the completion of this experiment or project.
- Which fundamental laws or principles are key to this system's function?
- Identify any way(s) in which one might obtain false or otherwise misleading measurements from test equipment in this system.
- What will happen if (component X) fails (open/shorted/etc.)?
- What would have to occur to make this system unsafe?

4.1 Conceptual reasoning

These questions are designed to stimulate your analytic and synthetic thinking[3]. In a Socratic discussion with your instructor, the goal is for these questions to prompt an extended dialogue where assumptions are revealed, conclusions are tested, and understanding is sharpened. Your instructor may also pose additional questions based on those assigned, in order to further probe and refine your conceptual understanding.

Questions that follow are presented to challenge and probe your understanding of various concepts presented in the tutorial. These questions are intended to serve as a guide for the Socratic dialogue between yourself and the instructor. Your instructor's task is to ensure you have a sound grasp of these concepts, and the questions contained in this document are merely a means to this end. Your instructor may, at his or her discretion, alter or substitute questions for the benefit of tailoring the discussion to each student's needs. The only absolute requirement is that each student is challenged and assessed at a level equal to or greater than that represented by the documented questions.

It is far more important that you convey your *reasoning* than it is to simply convey a correct answer. For this reason, you should refrain from researching other information sources to answer questions. What matters here is that *you* are doing the thinking. If the answer is incorrect, your instructor will work with you to correct it through proper reasoning. A correct answer without an adequate explanation of how you derived that answer is unacceptable, as it does not aid the learning or assessment process.

You will note a conspicuous lack of answers given for these conceptual questions. Unlike standard textbooks where answers to every other question are given somewhere toward the back of the book, here in these learning modules students must rely on other means to check their work. The best way by far is to debate the answers with fellow students and also with the instructor during the Socratic dialogue sessions intended to be used with these learning modules. Reasoning through challenging questions with other people is an excellent tool for developing strong reasoning skills.

Another means of checking your conceptual answers, where applicable, is to use circuit simulation software to explore the effects of changes made to circuits. For example, if one of these conceptual questions challenges you to predict the effects of altering some component parameter in a circuit, you may check the validity of your work by simulating that same parameter change within software and seeing if the results agree.

4.1.1 Reading outline and reflections

"Reading maketh a full man; conference a ready man; and writing an exact man" – Francis Bacon

Francis Bacon's advice is a blueprint for effective education: reading provides the learner with knowledge, writing focuses the learner's thoughts, and critical dialogue equips the learner to confidently communicate and apply their learning. Independent acquisition and application of knowledge is a powerful skill, well worth the effort to cultivate. To this end, students should read

[3] *Analytical* thinking involves the "disassembly" of an idea into its constituent parts, analogous to dissection. *Synthetic* thinking involves the "assembly" of a new idea comprised of multiple concepts, analogous to construction. Both activities are high-level cognitive skills, extremely important for effective problem-solving, necessitating frequent challenge and regular practice to fully develop.

these educational resources closely, journal their own reflections on the reading, and discuss in detail their findings with classmates and instructor(s). You should be able to do all of the following after reading any instructional text:

☑ Briefly SUMMARIZE THE TEXT in the form of a journal entry documenting your learning as you progress through the course of study. Share this summary in dialogue with your classmates and instructor. Journaling is an excellent self-test of thorough reading because you cannot clearly express what you have not read or did not comprehend.

☑ Demonstrate ACTIVE READING STRATEGIES, including verbalizing your impressions as you read, simplifying long passages to convey the same ideas using fewer words, annotating text and illustrations with your own interpretations, working through mathematical examples shown in the text, cross-referencing passages with relevant illustrations and/or other passages, identifying problem-solving strategies applied by the author, etc. Technical reading is a special case of problem-solving, and so these strategies work precisely because they help solve any problem: paying attention to your own thoughts (metacognition), eliminating unnecessary complexities, identifying what makes sense, paying close attention to details, drawing connections between separated facts, and noting the successful strategies of others.

☑ Identify IMPORTANT THEMES, especially GENERAL LAWS and PRINCIPLES, expounded in the text and express them in the simplest of terms as though you were teaching an intelligent child. This emphasizes connections between related topics and develops your ability to communicate complex ideas to anyone.

☑ Form YOUR OWN QUESTIONS based on the reading, and then pose them to your instructor and classmates for their consideration. Anticipate both correct and incorrect answers, the incorrect answer(s) assuming one or more plausible misconceptions. This helps you view the subject from different perspectives to grasp it more fully.

☑ Devise EXPERIMENTS to test claims presented in the reading, or to disprove misconceptions. Predict possible outcomes of these experiments, and evaluate their meanings: what result(s) would confirm, and what would constitute disproof? Running mental simulations and evaluating results is essential to scientific and diagnostic reasoning.

☑ Specifically identify any points you found CONFUSING. The reason for doing this is to help diagnose misconceptions and overcome barriers to learning.

4.1.2 Foundational concepts

Correct analysis and diagnosis of electric circuits begins with a proper understanding of some basic concepts. The following is a list of some important concepts referenced in this module's tutorial. Define each of them in your own words, and be prepared to illustrate each of these concepts with a description of a practical example and/or a live demonstration.

Energy

Potential energy

Conservation of Energy

Electrons

Protons

Neutrons

Conductors versus insulators

Allotrope

Quantum theory

Electron shell

Valence

Atomic bond

Energy band

Energy band diagram

Conductors versus Insulators

Semiconductor

Bandgap

Doping

Tetravalence

Trivalence

Pentavalence

Hole

Electric field

N-type semiconductor

P-type semiconductor

Conventional versus electron flow

Intrinsic semiconductor

Extrinsic semiconductor

Majority carrier

Minority carrier

Absolute zero

Fermi level

Depletion region

Photon

Photovoltaic cell

Diode

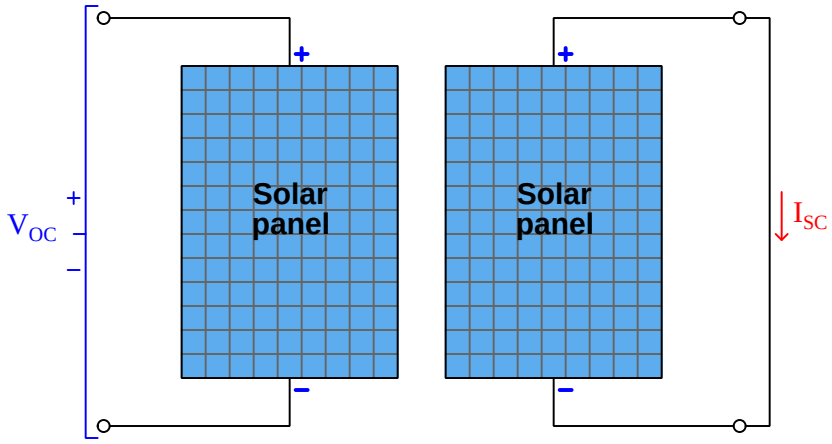
Forward-biasing

Reverse-biasing

Shockley diode equation

4.1.3 Solar panel output

Imagine setting two identical solar panels side-by-side so as to experience the exact same amount of illumination. One of these panels is open-circuited, while the other is short-circuited:



If some clouds begin to obscure the sunlight falling upon these two panels, how will this diminution of light affect the open-circuit voltage (V_{OC}) and the short-circuit current (I_{SC}), and most importantly *why*?

Challenges

- Suppose we need to re-design a solar panel to output greater voltage. How could it be designed differently to achieve this goal?
- Suppose we need to re-design a solar panel to output greater current. How could it be designed differently to achieve this goal?

4.1.4 Photodiode biasing

If an exposed PN junction is properly biased, a current will pass through it proportional to the amount of light falling on the junction. Should this external voltage bias be forward or reverse, and why?

Challenges

- Explain how a photodiode needs to be manufactured in order to be sensitive to a particular color of light.

4.1.5 Light-emitting diodes

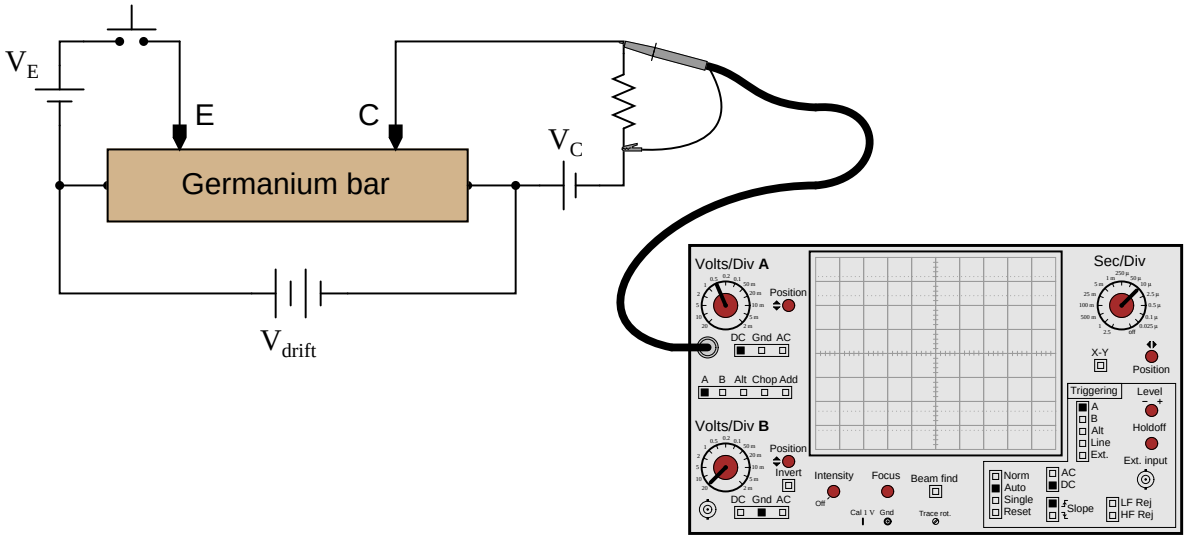
Every PN junction conducting a "forward" current emits light. Reference an energy-band diagram for a forward-biased PN junction and using that diagram explain why this is true.

Challenges

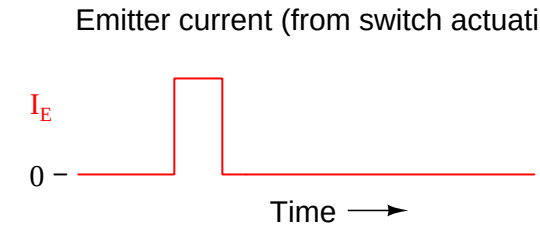
- What determines the color of the light emitted?
- Would the amount of voltage needed to energize a blue LED be different from the amount of voltage needed to energize a red LED? If so, would the blue LED require more or less voltage than the red while conducting the same amount of current?

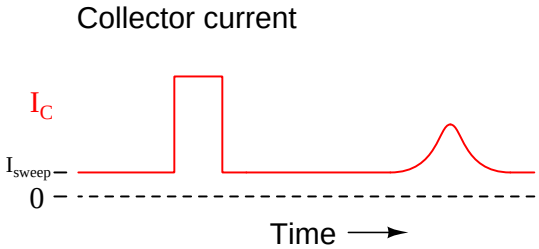
4.1.6 Hayes-Shockley experiment

A fascinating experiment carried out by J. R. Hayes and W. Shockley in the early 1950's involved a bar of N-doped germanium with two metal point contacts labeled "E" and "C", for "Emitter" and "Collector", respectively:

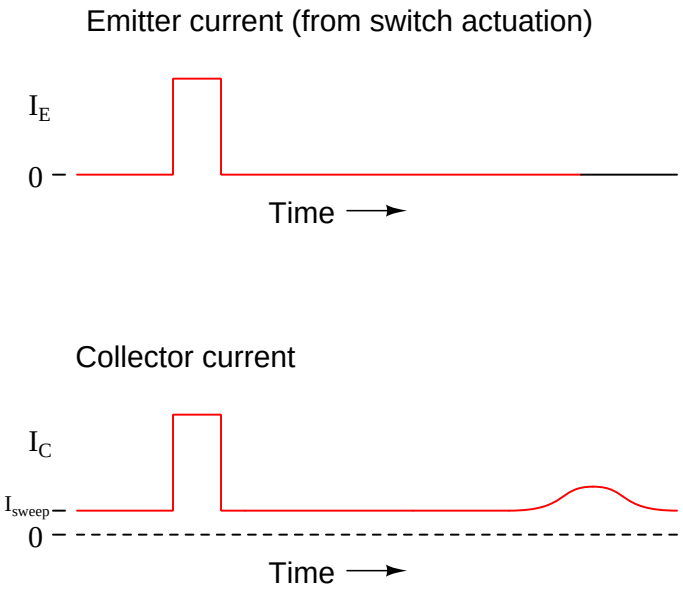


Upon actuating the switch, two distinct pulses were noted on the oscilloscope display:





With less drift voltage (V_{drift}) applied across the length of the bar, the second pulse was seen to be further delayed and more diffused:



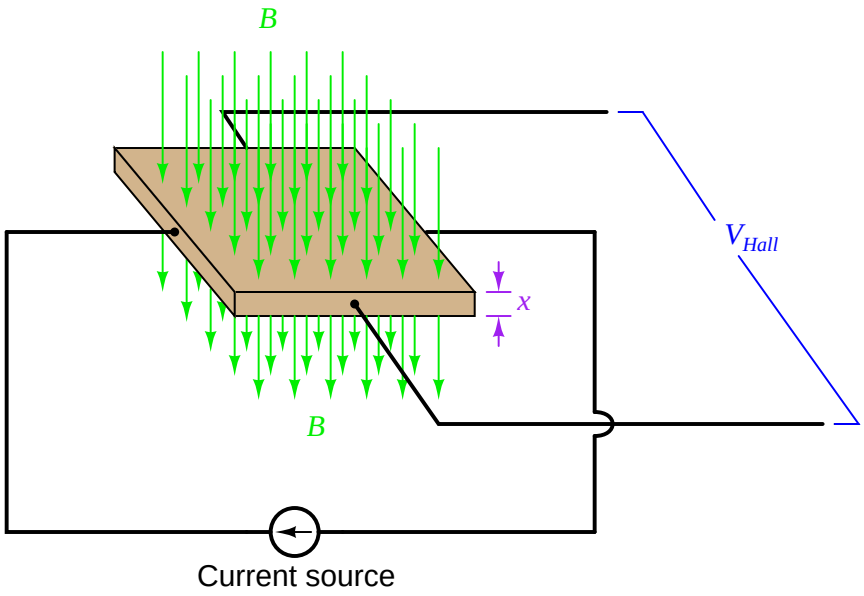
The instantaneous effect of the first pulse (precisely timed with the closure of the switch) is not the most interesting facet of this experiment. Rather, the second (delayed) pulse is. Explain what caused this second pulse, and why its shape depended on V_{drift}.

Challenges

- Identify a different way to stimulate this germanium bar other than by using a small voltage source V_E.

4.1.7 Hall effect on electrons versus holes

The *Hall effect* is based on the deflection of charge carriers in a current-carrying substance under the influence of a perpendicular magnetic field:



Charge carriers deflected by the magnetic field generate a voltage, called the *Hall voltage* (V_{Hall}), along an axis perpendicular to both the current and the magnetic field. The magnitude of this voltage is directly proportional to the strengths of the magnetic flux density (B) and current (I) and inversely proportional to the thickness of the material. The constant K in the following formula is determined by the material type:

$$V_{Hall} = K\frac{IB}{x}$$

Use the right-hand rule to trace the deflection path of holes (positive charge carriers) as they move through the material in the circuit shown above. Then use the left-hand rule to trace the deflection path of electrons (negative charge carriers) as they move through the material. Do you find a definite polarity emerging for V_{Hall}? If so, which direction is it?

Challenges

- Identify ways to maximize the Hall voltage for any given material specimen.

4.1.8 Loaded versus unloaded solar panel

Suppose we have a solar panel exposed to bright sunlight, connected to and powering a DC load. Energy from the sunlight is converted by the cells in the panel to electricity, and that electricity delivers energy to the load where the energy is used for some practical purpose.

Suppose now that the electrical circuit breaks, so that the load de-energizes. Meanwhile, sunlight still strikes the panel at full strength. What happens to the energy of the sunlight that was formerly being converted into electricity? Devise an experiment to test your hypothesis.

Challenges

- Explain how the Conservation of Energy relates to this scenario.

4.2 Quantitative reasoning

These questions are designed to stimulate your computational thinking. In a Socratic discussion with your instructor, the goal is for these questions to reveal your mathematical approach(es) to problem-solving so that good technique and sound reasoning may be reinforced. Your instructor may also pose additional questions based on those assigned, in order to observe your problem-solving firsthand.

Mental arithmetic and estimations are strongly encouraged for all calculations, because without these abilities you will be unable to readily detect errors caused by calculator misuse (e.g. keystroke errors).

You will note a conspicuous lack of answers given for these quantitative questions. Unlike standard textbooks where answers to every other question are given somewhere toward the back of the book, here in these learning modules students must rely on other means to check their work. My advice is to use circuit simulation software such as SPICE to check the correctness of quantitative answers. Refer to those learning modules within this collection focusing on SPICE to see worked examples which you may use directly as practice problems for your own study, and/or as templates you may modify to run your own analyses and generate your own practice problems.

Completely worked example problems found in the Tutorial may also serve as "test cases[4]" for gaining proficiency in the use of circuit simulation software, and then once that proficiency is gained you will never need to rely[5] on an answer key!

[4] In other words, set up the circuit simulation software to analyze the same circuit examples found in the Tutorial. If the simulated results match the answers shown in the Tutorial, it confirms the simulation has properly run. If the simulated results disagree with the Tutorial's answers, something has been set up incorrectly in the simulation software. Using every Tutorial as practice in this way will quickly develop proficiency in the use of circuit simulation software.

[5] This approach is perfectly in keeping with the instructional philosophy of these learning modules: *teaching students to be self-sufficient thinkers*. Answer keys can be useful, but it is even more useful to your long-term success to have a set of tools on hand for checking your own work, because once you have left school and are on your own, there will no longer be "answer keys" available for the problems you will have to solve.

4.2.1 Miscellaneous physical constants

Note: constants shown in **bold** type are *exact*, not approximations. Values inside of parentheses show one standard deviation (σ) of uncertainty in the final digits: for example, the magnetic permeability of free space value given as $1.25663706212(19) \times 10^{-6}$ H/m represents a center value (i.e. the location parameter) of $1.25663706212 \times 10^{-6}$ Henrys per meter with one standard deviation of uncertainty equal to $0.0000000000019 \times 10^{-6}$ Henrys per meter.

Avogadro's number (N_A) = **6.02214076** $\times 10^{23}$ **per mole** (mol^{-1})

Boltzmann's constant (k) = **1.380649** $\times 10^{-23}$ **Joules per Kelvin** (J/K)

Electronic charge (e) = **1.602176634** $\times 10^{-19}$ **Coulomb** (C)

Faraday constant (F) = **96,485.33212...** $\times 10^{4}$ **Coulombs per mole** (C/mol)

Magnetic permeability of free space (μ_0) = $1.25663706212(19) \times 10^{-6}$ Henrys per meter (H/m)

Electric permittivity of free space (ϵ_0) = $8.8541878128(13) \times 10^{-12}$ Farads per meter (F/m)

Characteristic impedance of free space (Z_0) = 376.730313668(57) Ohms (Ω)

Gravitational constant (G) = $6.67430(15) \times 10^{-11}$ cubic meters per kilogram-seconds squared ($m^3/kg\text{-}s^2$)

Molar gas constant (R) = **8.314462618... Joules per mole-Kelvin** (J/mol-K) = 0.08205746(14) liters-atmospheres per mole-Kelvin

Planck constant (h) = **6.62607015** $\times 10^{-34}$ **joule-seconds** (J-s)

Stefan-Boltzmann constant (σ) = **5.670374419...** $\times 10^{-8}$ **Watts per square meter-Kelvin**^{4} ($W/m^2 \cdot K^4$)

Speed of light in a vacuum (c) = **299,792,458 meters per second** (m/s) = 186282.4 miles per second (mi/s)

Note: All constants taken from NIST data "Fundamental Physical Constants – Complete Listing", from `http://physics.nist.gov/constants`, National Institute of Standards and Technology (NIST), 2018 CODATA Adjustment.

4.2.2 Introduction to spreadsheets

A powerful computational tool you are encouraged to use in your work is a *spreadsheet*. Available on most personal computers (e.g. Microsoft Excel), *spreadsheet* software performs numerical calculations based on number values and formulae entered into cells of a grid. This grid is typically arranged as lettered columns and numbered rows, with each cell of the grid identified by its column/row coordinates (e.g. cell `B3`, cell `A8`). Each cell may contain a string of text, a number value, or a mathematical formula. The spreadsheet automatically updates the results of all mathematical formulae whenever the entered number values are changed. This means it is possible to set up a spreadsheet to perform a series of calculations on entered data, and those calculations will be re-done by the computer any time the data points are edited in any way.

For example, the following spreadsheet calculates average speed based on entered values of distance traveled and time elapsed:

	A	B	C	D
1	Distance traveled	46.9	Kilometers	
2	Time elapsed	1.18	Hours	
3	Average speed	= B1 / B2	km/h	
4				
5				

Text labels contained in cells `A1` through `A3` and cells `C1` through `C3` exist solely for readability and are not involved in any calculations. Cell `B1` contains a sample distance value while cell `B2` contains a sample time value. The formula for computing speed is contained in cell `B3`. Note how this formula begins with an "equals" symbol (`=`), references the values for distance and speed by lettered column and numbered row coordinates (`B1` and `B2`), and uses a forward slash symbol for division (`/`). The coordinates `B1` and `B2` function as *variables*[6] would in an algebraic formula.

When this spreadsheet is executed, the numerical value `39.74576` will appear in cell `B3` rather than the formula `= B1 / B2`, because 39.74576 is the computed speed value given 46.9 kilometers traveled over a period of 1.18 hours. If a different numerical value for distance is entered into cell `B1` or a different value for time is entered into cell `B2`, cell `B3`'s value will automatically update. All you need to do is set up the given values and any formulae into the spreadsheet, and the computer will do all the calculations for you.

Cell `B3` may be referenced by other formulae in the spreadsheet if desired, since it is a variable just like the given values contained in `B1` and `B2`. This means it is possible to set up an entire chain of calculations, one dependent on the result of another, in order to arrive at a final value. The arrangement of the given data and formulae need not follow any pattern on the grid, which means you may place them anywhere.

[6] Spreadsheets may also provide means to attach text labels to cells for use as variable names (Microsoft Excel simply calls these labels "names"), but for simple spreadsheets such as those shown here it's usually easier just to use the standard coordinate naming for each cell.

Common[7] arithmetic operations available for your use in a spreadsheet include the following:

- Addition (`+`)
- Subtraction (`-`)
- Multiplication (`*`)
- Division (`/`)
- Powers (`^`)
- Square roots (`sqrt()`)
- Logarithms (`ln()` , `log10()`)

Parentheses may be used to ensure[8] proper order of operations within a complex formula. Consider this example of a spreadsheet implementing the *quadratic formula*, used to solve for roots of a polynomial expression in the form of $ax^2 + bx + c$:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

	A	B
1	x_1	= (-B4 + sqrt((B4^2) - (4*B3*B5))) / (2*B3)
2	x_2	= (-B4 - sqrt((B4^2) - (4*B3*B5))) / (2*B3)
3	a =	9
4	b =	5
5	c =	-2

This example is configured to compute roots[9] of the polynomial $9x^2 + 5x - 2$ because the values of 9, 5, and -2 have been inserted into cells `B3`, `B4`, and `B5`, respectively. Once this spreadsheet has been built, though, it may be used to calculate the roots of *any* second-degree polynomial expression simply by entering the new a, b, and c coefficients into cells `B3` through `B5`. The numerical values appearing in cells `B1` and `B2` will be automatically updated by the computer immediately following any changes made to the coefficients.

[7] Modern spreadsheet software offers a bewildering array of mathematical functions you may use in your computations. I recommend you consult the documentation for your particular spreadsheet for information on operations other than those listed here.

[8] Spreadsheet programs, like text-based programming languages, are designed to follow standard order of operations by default. However, my personal preference is to use parentheses even where strictly unnecessary just to make it clear to any other person viewing the formula what the intended order of operations is.

[9] Reviewing some algebra here, a *root* is a value for x that yields an overall value of zero for the polynomial. For this polynomial ($9x^2 + 5x - 2$) the two roots happen to be $x = 0.269381$ and $x = -0.82494$, with these values displayed in cells `B1` and `B2`, respectively upon execution of the spreadsheet.

Alternatively, one could break up the long quadratic formula into smaller pieces like this:

$$y = \sqrt{b^2 - 4ac} \qquad z = 2a$$

$$x = \frac{-b \pm y}{z}$$

	A	B	C
1	x_1	= (-B4 + C1) / C2	= sqrt((B4^2) - (4*B3*B5))
2	x_2	= (-B4 - C1) / C2	= 2*B3
3	a =	9	
4	b =	5	
5	c =	-2	

Note how the square-root term (y) is calculated in cell `C1`, and the denominator term (z) in cell `C2`. This makes the two final formulae (in cells `B1` and `B2`) simpler to interpret. The positioning of all these cells on the grid is completely arbitrary[10] – all that matters is that they properly reference each other in the formulae.

Spreadsheets are particularly useful for situations where the same set of calculations representing a circuit or other system must be repeated for different initial conditions. The power of a spreadsheet is that it automates what would otherwise be a tedious set of calculations. One specific application of this is to simulate the effects of various components within a circuit failing with abnormal values (e.g. a shorted resistor simulated by making its value nearly zero; an open resistor simulated by making its value extremely large). Another application is analyzing the behavior of a circuit design given new components that are out of specification, and/or aging components experiencing drift over time.

[10] My personal preference is to locate all the "given" data in the upper-left cells of the spreadsheet grid (each data point flanked by a sensible name in the cell to the left and units of measurement in the cell to the right as illustrated in the first distance/time spreadsheet example), sometimes coloring them in order to clearly distinguish which cells contain entered data versus which cells contain computed results from formulae. I like to place all formulae in cells below the given data, and try to arrange them in logical order so that anyone examining my spreadsheet will be able to figure out *how* I constructed a solution. This is a general principle I believe all computer programmers should follow: *document and arrange your code to make it easy for other people to learn from it.*

4.2.3 Predicting diode current

Calculate the amount of current that will flow through a diode at a forward voltage of 0.73 Volts at room temperature (20 ^{o}C), given the following parameters sampled from a SPICE model for a typical rectifying diode:

```
.MODEL 1N4001 D IS=1.22483E-8 N=1.8335 BV=50 IBV=5E-8 VJ=0.7
```

Challenges

- A commonly-taught rule is that a silicon diode will not begin to conduct current unless and until at least 0.7 Volts forward bias is applied. How accurate is this statement, based on the Shockley equation?

4.3 Diagnostic reasoning

These questions are designed to stimulate your deductive and inductive thinking, where you must apply general principles to specific scenarios (deductive) and also derive conclusions about the failed circuit from specific details (inductive). In a Socratic discussion with your instructor, the goal is for these questions to reinforce your recall and use of general circuit principles and also challenge your ability to integrate multiple symptoms into a sensible explanation of what's wrong in a circuit. Your instructor may also pose additional questions based on those assigned, in order to further challenge and sharpen your diagnostic abilities.

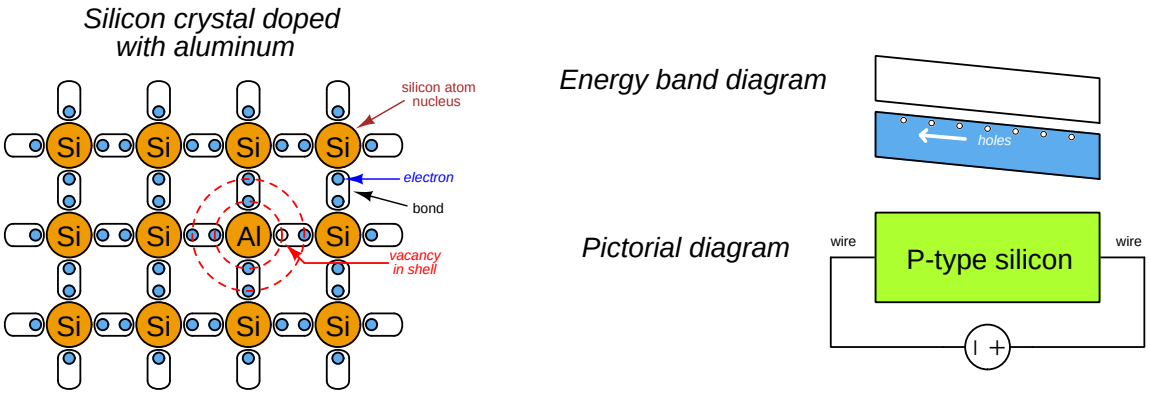
As always, your goal is to fully *explain* your analysis of each problem. Simply obtaining a correct answer is not good enough – you must also demonstrate sound reasoning in order to successfully complete the assignment. Your instructor's responsibility is to probe and challenge your understanding of the relevant principles and analytical processes in order to ensure you have a strong foundation upon which to build further understanding.

You will note a conspicuous lack of answers given for these diagnostic questions. Unlike standard textbooks where answers to every other question are given somewhere toward the back of the book, here in these learning modules students must rely on other means to check their work. The best way by far is to debate the answers with fellow students and also with the instructor during the Socratic dialogue sessions intended to be used with these learning modules. Reasoning through challenging questions with other people is an excellent tool for developing strong reasoning skills.

Another means of checking your diagnostic answers, where applicable, is to use circuit simulation software to explore the effects of faults placed in circuits. For example, if one of these diagnostic questions requires that you predict the effect of an open or a short in a circuit, you may check the validity of your work by simulating that same fault (substituting a very high resistance in place of that component for an open, and substituting a very low resistance for a short) within software and seeing if the results agree.

4.3.1 Tilted band diagrams

When a voltage is applied across the length of a uniformly doped semiconductor, it functions as a resistor and drops voltage. This may be represented in band-diagram form by showing the bands *tilted*:



However, when we apply voltage across a *junction* of P-type and N-type semiconductor, the only tilting effect we see appearing in the band diagram is across the depletion region (at the junction) and not within either the P- or N-type regions.

Why don't *all* regions of a PN junction experience a "tilt" to their energy bands as an external voltage is applied? What is incorrect about thinking that all the energy bands throughout the device will tilt with the applied voltage?

Challenges

- Under what condition – forward-biased or reverse-biased – do you think we might ever see the energy bands outside the depletion region begin to tilt, and why?

Appendix A

Problem-Solving Strategies

The ability to solve complex problems is arguably one of the most valuable skills one can possess, and this skill is particularly important in any science-based discipline.

- Study principles, not procedures. Don't be satisfied with merely knowing how to compute solutions – learn *why* those solutions work.
- Identify what it is you need to solve, identify all relevant data, identify all units of measurement, identify any general principles or formulae linking the given information to the solution, and then identify any "missing pieces" to a solution. Annotate all diagrams with this data.
- Sketch a diagram to help visualize the problem. When building a real system, always devise a plan for that system and analyze its function *before* constructing it.
- Follow the units of measurement and meaning of every calculation. If you are ever performing mathematical calculations as part of a problem-solving procedure, and you find yourself unable to apply each and every intermediate result to some aspect of the problem, it means you don't understand what you are doing. Properly done, every mathematical result should have practical meaning for the problem, and not just be an abstract number. You should be able to identify the proper units of measurement for each and every calculated result, and show where that result fits into the problem.
- Perform "thought experiments" to explore the effects of different conditions for theoretical problems. When troubleshooting real systems, perform *diagnostic tests* rather than visually inspecting for faults, the best diagnostic test being the one giving you the most information about the nature and/or location of the fault with the fewest steps.
- Simplify the problem until the solution becomes obvious, and then use that obvious case as a model to follow in solving the more complex version of the problem.
- Check for exceptions to see if your solution is incorrect or incomplete. A good solution will work for *all* known conditions and criteria. A good example of this is the process of testing scientific hypotheses: the task of a scientist is not to find support for a new idea, but rather to *challenge* that new idea to see if it holds up under a battery of tests. The philosophical

principle of *reductio ad absurdum* (i.e. disproving a general idea by finding a specific case where it fails) is useful here.

- Work "backward" from a hypothetical solution to a new set of given conditions.
- Add quantities to problems that are qualitative in nature, because sometimes a little math helps illuminate the scenario.
- Sketch graphs illustrating how variables relate to each other. These may be quantitative (i.e. with realistic number values) or qualitative (i.e. simply showing increases and decreases).
- Treat quantitative problems as qualitative in order to discern the relative magnitudes and/or directions of change of the relevant variables. For example, try determining what happens if a certain variable were to increase or decrease before attempting to precisely calculate quantities: how will each of the dependent variables respond, by increasing, decreasing, or remaining the same as before?
- Consider limiting cases. This works especially well for qualitative problems where you need to determine which direction a variable will change. Take the given condition and magnify that condition to an extreme degree as a way of simplifying the direction of the system's response.
- Check your work. This means regularly testing your conclusions to see if they make sense. This does *not* mean repeating the same steps originally used to obtain the conclusion(s), but rather to use some other means to check validity. Simply repeating procedures often leads to *repeating the same errors* if any were made, which is why alternative paths are better.

Appendix B

Instructional philosophy

B.1 First principles of learning

- **Anyone can learn anything** given appropriate time, effort, resources, challenges, encouragement, and expectations. Dedicating time and investing effort are the student's responsibility; providing resources, challenges, and encouragement are the teacher's responsibility; high expectations are a responsibility shared by both student and teacher.

- **Transfer is not automatic.** The human mind has a natural tendency to compartmentalize information, which means the process of taking knowledge learned in one context and applying it to another usually does not come easy and therefore should never be taken for granted.

- **Learning is iterative.** The human mind rarely learns anything perfectly on the first attempt. Anticipate mistakes and plan for multiple tries to achieve full understanding, using the lessons of those mistakes as feedback to guide future attempts.

- **Information is absorbed, but understanding is created.** Facts and procedures may be memorized easily enough by repeated exposure, but the ability to reliably apply principles to novel scenarios only comes through intense personal effort. This effort is fundamentally creative in nature: casting new concepts in one's own words, running mental experiments to perceive patterns and trends, etc. These acts of making knowledge "one's own" need not be perfect in order to be effective, as the value lies in the activity and not necessarily the finished product. Depth of understanding and permanence of memory are both directly related to mental effort during the learning process – i.e. *easy learning is shallow and fleeting.*

- **Education trumps training.** There is no such thing as an entirely isolated subject, as all fields of knowledge are connected. Training is narrowly-focused and task-oriented. Education is broad-based and principle-oriented. When preparing for a life-long technical career, education beats training every time.

- **Character matters.** Poor habits are more destructive than deficits of knowledge or skill. This is especially true in collective endeavors, where a team's ability to function depends on trust between its members. Simply put, no one wants an untrustworthy person on their team. An essential component of education then, is character development.

- **People learn to be responsible by bearing responsibility.** An irresponsible person is someone who has never *had* to be responsible for anything that mattered enough to them. Just as anyone can learn anything, anyone can become responsible if the personal cost of irresponsibility becomes high enough.

- **What gets measured, gets done.** Accurate and relevant assessment of learning is key to ensuring all students learn. Therefore, it is imperative to measure what matters.

- **Failure is nothing to fear.** Every human being fails, and fails in multiple ways at multiple times. Eventual success only happens when we don't stop trying.

B.2 Proven strategies for instructors

- Assume every student is capable of learning anything they desire given the proper conditions. Treat them as capable adults by granting real responsibility and avoiding artificial incentives such as merit or demerit points.
- Create a consistent culture of high expectations across the entire program of study. Demonstrate and encourage patience, persistence, and a healthy sense of self-skepticism. Anticipate and de-stigmatize error. Teach respect for the capabilities of others as well as respect for one's own fallibility.
- Replace lecture with "inverted" instruction, where students first encounter new concepts through reading and then spend class time in Socratic dialogue with the instructor exploring those concepts and solving problems individually. There is a world of difference between observing someone solve a problem versus actually solving a problem yourself, and so the point of this form of instruction is to place students in a position where they *cannot* passively observe.
- Require students to read extensively, write about what they learn, and dialogue with you and their peers to sharpen their understanding. Apply Francis Bacon's advice that "reading maketh a full man; conference a ready man; and writing an exact man". These are complementary activities helping students expand their confidence and abilities.
- Use artificial intelligence (AI) to challenge student understanding rather than merely provide information. Find productive ways for AI to critique students' clarity of thought and of expression, for example by employing AI as a Socratic-style interlocutor or as a reviewer of students' journals. Properly applied, AI has the ability to expand student access to critical review well outside the bounds of their instructor's reach.
- Build frequent and rapid feedback into the learning process so that students know at all times how well they are learning, to identify problems early and fix them before they grow. Model the intellectual habit of self-assessing and self-correcting your own understanding (i.e. a cognitive *feedback loop*), encouraging students to do the same.
- Use "mastery" as the standard for every assessment, which means the exam or experiment or project must be done with 100% competence in order to pass. Provide students with multiple opportunity for re-tries (different versions of the assessment every time).
- Require students to devise their own hypotheses and procedures on all experiments, so that the process is truly a scientific one. Have students assess their proposed experimental procedures for risk and devise mitigations for those risks. Let nothing be pre-designed about students' experiments other than a stated task (i.e. what principle the experiment shall test) at the start and a set of demonstrable knowledge and skill objectives at the end.
- Have students build as much of their lab equipment as possible: building power sources, building test assemblies[1], and building complete working systems (no kits!). In order to provide

[1] In the program I teach, every student builds their own "Development Board" consisting of a metal chassis with DIN rail, terminal blocks, and an AC-DC power supply of their own making which functions as a portable lab environment they can use at school as well as take home.

this same "ground-up" experience for every new student, this means either previous students take their creations with them, or the systems get disassembled in preparation for the new students, or the systems grow and evolve with each new student group.

- Incorporate external accountability for you and for your students, continuously improving the curriculum and your instructional methods based on proven results. Have students regularly network with active professionals through participation in advisory committee meetings, service projects, tours, jobshadows, internships, etc. Practical suggestions include requiring students to design and build projects for external clients (e.g. community groups, businesses, different departments within the institution), and also requiring students attend all technical advisory committee meetings and to dialogue with the industry representatives at those meetings.
- Repeatedly explore difficult-to-learn concepts across multiple courses, so that students have multiple opportunities to build their understanding.
- Relate all new concepts, whenever possible, to previous concepts and to relevant physical laws. Challenge each and every student, every day, to *reason* from concept to concept and to explain the logical connections between. Challenge students to verify their conclusions by multiple approaches (e.g. double-checking their work using different methods). Ask *"Why?"* often.
- Maintain detailed records on each student's performance and share these records privately with them. These records should include academic performance as well as professionally relevant behavioral tendencies.
- Hold mandatory "check-in" meetings between all program faculty and each new student during their first term. Offer these to all other students as an option, except for any students continuing to manifest unprofessional behaviors, poor academic performance, or who have some other need for a face-to-face meeting with faculty.
- Address problems while they are small, before they grow larger. This is equally true for tutoring technical concepts as it is for helping students build professional habits.
- Build rigorous quality control into the curriculum to ensure every student masters every important concept, and that the mastery is retained over time. This includes (1) review questions added to every exam to re-assess knowledge taught in previous terms, (2) cumulative exams at the end of every term to re-assess all important concepts back to the very beginning of the program, and (3) review assessments in practical (hands-on) coursework to ensure critically-important skills were indeed taught and are still retained. What you will find by doing this is that it actually boosts retention of students by ensuring that important knowledge gets taught and is retained over long spans of time. In the absence of such quality control, student learning and retention tends to be spotty and this contributes to drop-out and failure rates later in their education.
- Finally, *never rush learning.* Education is not a race. Give your students ample time to digest complex ideas, as you continually remind yourself of just how long it took you to achieve mastery! Long-term retention and the consistently correct application of concepts are always the result of *focused effort over long periods of time* which means there are no shortcuts to learning.

B.3 Proven strategies for students

The single most important piece of advice I have for any student of any subject is to take responsibility for your own development in all areas of life including mental development. Expecting others in your life to entirely guide your own development is a recipe for disappointment. This is just as true for students enrolled in formal learning institutions as it is for auto-didacts pursuing learning entirely on their own. Learning to think in new ways is key to being able to gainfully use information, to make informed decisions about your life, and to best serve those you care about. With this in mind, I offer the following advice to students:

- **Approach all learning as valuable**. No matter what course you take, no matter who you learn from, no matter the subject, there is something useful in every learning experience. If you don't see the value of every new experience, you are not looking closely enough!
- **Continually challenge yourself**. Let other people take shortcuts and find easy answers to easy problems. The purpose of education is to stretch your mind, in order to shape it into a more powerful tool. This doesn't come by taking the path of least resistance. An excellent analogy for an empowering education is productive physical exercise: becoming stronger, more flexible, and more persistent only comes through intense personal effort.
- **Master the use of language**. This includes reading extensively, writing every day, listening closely, and speaking articulately. To a great extent language channels and empowers thought, so the better you are at wielding language the better you will be at grasping abstract concepts and articulating them not only for your benefit but for others as well.
- **Do not limit yourself to the resources given to you**. Read books that are not on the reading list. Run experiments that aren't assigned to you. Form study groups outside of class. Take an entrepreneurial approach to your own education, as though it were a business you were building for your future benefit.
- **Express and share what you learn**. Take every opportunity to teach what you have learned to others, as this will not only help them but will also strengthen your own understanding[2].
- Realize that **no one can give you understanding**, just as no one can give you physical fitness. These both must be *built*.
- **Above all, recognize that learning is hard work, and that a certain level of frustration is unavoidable**. There are times when you will struggle to grasp some of these concepts, and that struggle is a natural thing. Take heart that it will yield with persistent and varied[3] effort, and never give up! That concepts don't immediately come to you is not a sign of something wrong, but rather of something right: that you have found a worthy challenge!

[2] On a personal note, I was surprised to learn just how much my own understanding of electronics and related subjects was strengthened by becoming a teacher. When you are tasked every day with helping other people grasp complex topics, it catalyzes your own learning by giving you powerful incentives to study, to articulate your thoughts, and to reflect deeply on the process of learning.

[3] As the old saying goes, "Insanity is trying the same thing over and over again, expecting different results." If you find yourself stumped by something in the text, you should attempt a different approach. Alter the thought experiment, change the mathematical parameters, do whatever you can to see the problem in a slightly different light, and then the solution will often present itself more readily.

B.4 Design of these learning modules

"The unexamined circuit is not worth energizing" – Socrates (if he had taught electricity)

These learning modules, although useful for self-study, were designed to be used in a formal learning environment where a subject-matter expert challenges students to digest the content and exercise their critical thinking abilities in the answering of questions and in the construction and testing of working circuits. Every effort has been made to embed the following instructional and assessment philosophies within:

- The first goal of education is to enhance clear and independent thought, in order that every student reach their fullest potential in a highly complex and inter-dependent world. Robust reasoning is *always* more important than particulars of any subject matter, because its application is universal.
- Literacy is fundamental to independent learning and thought because text continues to be the most efficient way to communicate complex ideas over space and time. Those who cannot read with ease are limited in their ability to acquire knowledge and perspective.
- Articulate communication is fundamental to work that is complex and interdisciplinary.
- Faulty assumptions and poor reasoning are best corrected through challenge, not presentation. The rhetorical technique of *reductio ad absurdum* (disproving an assertion by exposing an absurdity) works well to discipline student's minds, not only to correct the problem at hand but also to learn how to detect and correct future errors.
- Important principles should be repeatedly explored and widely applied throughout a course of study, not only to reinforce their importance and help ensure their mastery, but also to showcase the interconnectedness and utility of knowledge.

These learning modules were expressly designed to be used in an "inverted" teaching environment[4] where students first read the introductory and tutorial chapters on their own, then individually attempt to answer the questions and construct working circuits according to the experiment and project guidelines. The instructor never lectures, but instead meets regularly with each individual student to review their progress, answer questions, identify misconceptions, and challenge the student to new depths of understanding through further questioning. Regular meetings between instructor and student should resemble a Socratic[5] dialogue, where questions serve as scalpels to dissect topics and expose assumptions. The student passes each module only after consistently demonstrating their ability to logically analyze and correctly apply all major concepts in each question or project/experiment. The instructor must be vigilant in probing each student's understanding to ensure they are truly *reasoning* and not just *memorizing*. This is why "Challenge" points appear throughout, as prompts for students to think deeper about topics and as starting points for instructor queries. Sometimes these challenge points require additional knowledge that hasn't been covered in the series to answer in full. This is okay, as the major purpose of the Challenges is to stimulate analysis and synthesis on the part of each student.

The instructor must possess enough mastery of the subject matter and awareness of students' reasoning to generate their own follow-up questions to practically any student response. Even completely correct answers given by the student should be challenged by the instructor for the purpose of having students practice articulating their thoughts and defending their reasoning. Conceptual errors committed by the student should be exposed and corrected not by direct instruction, but rather by reducing the errors to an absurdity[6] through well-chosen questions and thought experiments posed by the instructor. Becoming proficient at this style of instruction requires time and dedication, but the positive effects on critical thinking for both student and instructor are spectacular.

An inspection of these learning modules reveals certain unique characteristics. One of these is a bias toward thorough explanations in the tutorial chapters. Without a live instructor to explain concepts and applications to students, the text itself must fulfill this role. This philosophy results in lengthier explanations than what you might typically find in a textbook, each step of the reasoning process fully explained, including footnotes addressing common questions and concerns students raise while learning these concepts. Each tutorial seeks to not only explain each major concept in sufficient detail, but also to explain the logic of each concept and how each may be developed

[4] In a traditional teaching environment, students first encounter new information via *lecture* from an expert, and then independently apply that information via *homework*. In an "inverted" course of study, students first encounter new information via *homework*, and then independently apply that information under the scrutiny of an expert. The expert's role in lecture is to simply *explain*, but the expert's role in an inverted session is to *challenge*, *critique*, and if necessary *explain* where gaps in understanding still exist.

[5] Socrates is a figure in ancient Greek philosophy famous for his unflinching style of questioning. Although he authored no texts, he appears as a character in Plato's many writings. The essence of Socratic philosophy is to leave no question unexamined and no point of view unchallenged. While purists may argue a topic such as electric circuits is too narrow for a true Socratic-style dialogue, I would argue that the essential thought processes involved with scientific reasoning on *any* topic are not far removed from the Socratic ideal, and that students of electricity and electronics would do very well to challenge assumptions, pose thought experiments, identify fallacies, and otherwise employ the arsenal of critical thinking skills modeled by Socrates.

[6] This rhetorical technique is known by the Latin phrase *reductio ad absurdum*. The concept is to expose errors by counter-example, since only one solid counter-example is necessary to disprove a universal claim. As an example of this, consider the common misconception among beginning students of electricity that voltage cannot exist without current. One way to apply *reductio ad absurdum* to this statement is to ask how much current passes through a fully-charged battery connected to nothing (i.e. a clear example of voltage existing without current).

from "first principles". Again, this reflects the goal of developing clear and independent thought in students' minds, by showing how clear and logical thought was used to forge each concept. Students benefit from witnessing a model of clear thinking in action, and these tutorials strive to be just that.

Another feature of these learning modules is that they do not treat topics in isolation. Rather, important concepts are introduced early in the series, and appear repeatedly as stepping-stones toward other concepts in subsequent modules. This helps to avoid the "compartmentalization" of knowledge, demonstrating the inter-connectedness of concepts and simultaneously reinforcing them. Each module is fairly complete in itself, reserving the beginning of its tutorial to a review of foundational concepts.

To high standards of education,

Tony R. Kuphaldt

Appendix C

Tools used

I am indebted to the developers of many open-source software applications in the creation of these learning modules. The following is a list of these applications with some commentary on each.

You will notice a theme common to many of these applications: a bias toward *code*. Although I am by no means an expert programmer in any computer language, I understand and appreciate the flexibility offered by code-based applications where the user (you) enters commands into a plain ASCII text file, which the software then reads and processes to create the final output. Code-based computer applications are by their very nature *extensible*, while WYSIWYG (What You See Is What You Get) applications are generally limited to whatever user interface the developer makes for you.

The `GNU/Linux` computer operating system

> There is so much to be said about Linus Torvalds' `Linux` and Richard Stallman's `GNU` project. First, to credit just these two individuals is to fail to do justice to the *mob* of passionate volunteers who contributed to make this amazing software a reality. I first learned of `Linux` back in 1996, and have been using this operating system on my personal computers almost exclusively since then. It is *free*, it is completely *configurable*, and it permits the continued use of highly efficient `Unix` applications and scripting languages (e.g. shell scripts, Makefiles, `sed`, `awk`) developed over many decades. `Linux` not only provided me with a powerful computing platform, but its open design served to inspire my life's work of creating open-source educational resources.

Bram Moolenaar's `Vim` text editor

> Writing code for any code-based computer application requires a *text editor*, which may be thought of as a word processor strictly limited to outputting plain-ASCII text files. Many good text editors exist, and one's choice of text editor seems to be a deeply personal matter within the programming world. I prefer `Vim` because it operates very similarly to `vi` which is ubiquitous on `Unix`/`Linux` operating systems, and because it may be entirely operated via keyboard (i.e. no mouse required) which makes it fast to use.

Donald Knuth's TEX typesetting system

Developed in the late 1970's and early 1980's by computer scientist extraordinaire Donald Knuth to typeset his multi-volume magnum opus *The Art of Computer Programming*, this software allows the production of formatted text for screen-viewing or paper printing, all by writing plain-text code to describe how the formatted text is supposed to appear. TEX is not just a markup language for documents, but it is also a Turing-complete programming language in and of itself, allowing useful algorithms to be created to control the production of documents. Simply put, TEX *is a programmer's approach to word processing*. Since TEX is controlled by code written in a plain-text file, this means anyone may read that plain-text file to see exactly how the document was created. This openness afforded by the code-based nature of TEX makes it relatively easy to learn how other people have created their own TEX documents. By contrast, examining a beautiful document created in a conventional WYSIWYG word processor such as Microsoft `Word` suggests nothing to the reader about *how* that document was created, or what the user might do to create something similar. As Mr. Knuth himself once quipped, conventional word processing applications should be called WYSIAYG (What You See Is *All* You Get).

Leslie Lamport's LATEX extensions to TEX

Like all true programming languages, TEX is inherently extensible. So, years after the release of TEX to the public, Leslie Lamport decided to create a massive extension allowing easier compilation of book-length documents. The result was LATEX, which is the markup language used to create all ModEL module documents. You could say that TEX is to LATEX as `C` is to `C++`. This means it is permissible to use any and all TEX commands within LATEX source code, and it all still works. Some of the features offered by LATEX that would be challenging to implement in TEX include automatic index and table-of-content creation.

Tim Edwards' `Xcircuit` drafting program

This wonderful program is what I use to create all the schematic diagrams and illustrations (but not photographic images or mathematical plots) throughout the ModEL project. It natively outputs PostScript format which is a true vector graphic format (this is why the images do not pixellate when you zoom in for a closer view), and it is so simple to use that I have never had to read the manual! Object libraries are easy to create for `Xcircuit`, being plain-text files using PostScript programming conventions. Over the years I have collected a large set of object libraries useful for drawing electrical and electronic schematics, pictorial diagrams, and other technical illustrations.

`Gimp` graphic image manipulation program

Essentially an open-source clone of Adobe's `PhotoShop`, I use `Gimp` to resize, crop, and convert file formats for all of the photographic images appearing in the ModEL modules. Although `Gimp` does offer its own scripting language (called `Script-Fu`), I have never had occasion to use it. Thus, my utilization of `Gimp` to merely crop, resize, and convert graphic images is akin to using a sword to slice bread.

`SPICE` circuit simulation program

`SPICE` is to circuit analysis as TEX is to document creation: it is a form of markup language designed to describe a certain object to be processed in plain-ASCII text. When the plain-text "source file" is compiled by the software, it outputs the final result. More modern circuit analysis tools certainly exist, but I prefer `SPICE` for the following reasons: it is *free*, it is *fast*, it is *reliable*, and it is a fantastic tool for *teaching* students of electricity and electronics how to write simple code. I happen to use rather old versions of `SPICE`, version 2g6 being my "go to" application when I only require text-based output. `NGSPICE` (version 26), which is based on Berkeley `SPICE` version 3f5, is used when I require graphical output for such things as time-domain waveforms and Bode plots. In all `SPICE` example netlists I strive to use coding conventions compatible with all `SPICE` versions.

Andrew D. Hwang's `ePiX` mathematical visualization programming library

This amazing project is a `C++` library you may link to any `C/C++` code for the purpose of generating PostScript graphic images of mathematical functions. As a completely free and open-source project, it does all the plotting I would otherwise use a Computer Algebra System (CAS) such as `Mathematica` or `Maple` to do. It should be said that `ePiX` is *not* a Computer Algebra System like `Mathematica` or `Maple`, but merely a mathematical *visualization* tool. In other words, it won't determine integrals for you (you'll have to implement that in your own `C/C++` code!), but it can graph the results, and it does so beautifully. What I really admire about `ePiX` is that it is a `C++` programming library, which means it builds on the existing power and toolset available with that programming language. Mr. Hwang could have probably developed his own stand-alone application for mathematical plotting, but by creating a `C++` library to do the same thing he accomplished something much greater.

`gnuplot` mathematical visualization software

Another open-source tool for mathematical visualization is `gnuplot`. Interestingly, this tool is *not* part of Richard Stallman's GNU project, its name being a coincidence. For this reason the authors prefer "gnu" *not* be capitalized at all to avoid confusion. This is a much "lighter-weight" alternative to a spreadsheet for plotting tabular data, and the fact that it easily outputs directly to an X11 console or a file in a number of different graphical formats (including PostScript) is very helpful. I typically set my `gnuplot` output format to default (X11 on my Linux PC) for quick viewing while I'm developing a visualization, then switch to PostScript file export once the visual is ready to include in the document(s) I'm writing. As with my use of `Gimp` to do rudimentary image editing, my use of `gnuplot` only scratches the surface of its capabilities, but the important points are that it's *free* and that it *works well*.

`Python` programming language

Both Python and C++ find extensive use in these modules as instructional aids and exercises, but I'm listing Python here as a *tool* for myself because I use it almost daily as a *calculator*. If you open a Python interpreter console and type `from math import *` you can type mathematical expressions and have it return results just as you would on a hand calculator. Complex-number (i.e. *phasor*) arithmetic is similarly supported if you include the complex-math library (`from cmath import *`). Examples of this are shown in the Programming References chapter (if included) in each module. Of course, being a fully-featured programming language, Python also supports conditionals, loops, and other structures useful for calculation of quantities. Also, running in a console environment where all entries and returned values show as text in a chronologically-ordered list makes it easy to copy-and-paste those calculations to document exactly how they were performed.

Appendix D

Creative Commons License

Creative Commons Attribution 4.0 International Public License

By exercising the Licensed Rights (defined below), You accept and agree to be bound by the terms and conditions of this Creative Commons Attribution 4.0 International Public License ("Public License"). To the extent this Public License may be interpreted as a contract, You are granted the Licensed Rights in consideration of Your acceptance of these terms and conditions, and the Licensor grants You such rights in consideration of benefits the Licensor receives from making the Licensed Material available under these terms and conditions.

Section 1 – Definitions.

a. **Adapted Material** means material subject to Copyright and Similar Rights that is derived from or based upon the Licensed Material and in which the Licensed Material is translated, altered, arranged, transformed, or otherwise modified in a manner requiring permission under the Copyright and Similar Rights held by the Licensor. For purposes of this Public License, where the Licensed Material is a musical work, performance, or sound recording, Adapted Material is always produced where the Licensed Material is synched in timed relation with a moving image.

b. **Adapter's License** means the license You apply to Your Copyright and Similar Rights in Your contributions to Adapted Material in accordance with the terms and conditions of this Public License.

c. **Copyright and Similar Rights** means copyright and/or similar rights closely related to copyright including, without limitation, performance, broadcast, sound recording, and Sui Generis Database Rights, without regard to how the rights are labeled or categorized. For purposes of this Public License, the rights specified in Section 2(b)(1)-(2) are not Copyright and Similar Rights.

d. **Effective Technological Measures** means those measures that, in the absence of proper authority, may not be circumvented under laws fulfilling obligations under Article 11 of the WIPO Copyright Treaty adopted on December 20, 1996, and/or similar international agreements.

e. **Exceptions and Limitations** means fair use, fair dealing, and/or any other exception or

limitation to Copyright and Similar Rights that applies to Your use of the Licensed Material.

f. **Licensed Material** means the artistic or literary work, database, or other material to which the Licensor applied this Public License.

g. **Licensed Rights** means the rights granted to You subject to the terms and conditions of this Public License, which are limited to all Copyright and Similar Rights that apply to Your use of the Licensed Material and that the Licensor has authority to license.

h. **Licensor** means the individual(s) or entity(ies) granting rights under this Public License.

i. **Share** means to provide material to the public by any means or process that requires permission under the Licensed Rights, such as reproduction, public display, public performance, distribution, dissemination, communication, or importation, and to make material available to the public including in ways that members of the public may access the material from a place and at a time individually chosen by them.

j. **Sui Generis Database Rights** means rights other than copyright resulting from Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, as amended and/or succeeded, as well as other essentially equivalent rights anywhere in the world.

k. **You** means the individual or entity exercising the Licensed Rights under this Public License. **Your** has a corresponding meaning.

Section 2 – Scope.

a. License grant.

1. Subject to the terms and conditions of this Public License, the Licensor hereby grants You a worldwide, royalty-free, non-sublicensable, non-exclusive, irrevocable license to exercise the Licensed Rights in the Licensed Material to:

A. reproduce and Share the Licensed Material, in whole or in part; and

B. produce, reproduce, and Share Adapted Material.

2. Exceptions and Limitations. For the avoidance of doubt, where Exceptions and Limitations apply to Your use, this Public License does not apply, and You do not need to comply with its terms and conditions.

3. Term. The term of this Public License is specified in Section 6(a).

4. Media and formats; technical modifications allowed. The Licensor authorizes You to exercise the Licensed Rights in all media and formats whether now known or hereafter created, and to make technical modifications necessary to do so. The Licensor waives and/or agrees not to assert any right or authority to forbid You from making technical modifications necessary to exercise the Licensed Rights, including technical modifications necessary to circumvent Effective Technological Measures.

For purposes of this Public License, simply making modifications authorized by this Section 2(a)(4) never produces Adapted Material.

5. Downstream recipients.

A. Offer from the Licensor – Licensed Material. Every recipient of the Licensed Material automatically receives an offer from the Licensor to exercise the Licensed Rights under the terms and conditions of this Public License.

B. No downstream restrictions. You may not offer or impose any additional or different terms or conditions on, or apply any Effective Technological Measures to, the Licensed Material if doing so restricts exercise of the Licensed Rights by any recipient of the Licensed Material.

6. No endorsement. Nothing in this Public License constitutes or may be construed as permission to assert or imply that You are, or that Your use of the Licensed Material is, connected with, or sponsored, endorsed, or granted official status by, the Licensor or others designated to receive attribution as provided in Section 3(a)(1)(A)(i).

b. Other rights.

1. Moral rights, such as the right of integrity, are not licensed under this Public License, nor are publicity, privacy, and/or other similar personality rights; however, to the extent possible, the Licensor waives and/or agrees not to assert any such rights held by the Licensor to the limited extent necessary to allow You to exercise the Licensed Rights, but not otherwise.

2. Patent and trademark rights are not licensed under this Public License.

3. To the extent possible, the Licensor waives any right to collect royalties from You for the exercise of the Licensed Rights, whether directly or through a collecting society under any voluntary or waivable statutory or compulsory licensing scheme. In all other cases the Licensor expressly reserves any right to collect such royalties.

Section 3 – License Conditions.

Your exercise of the Licensed Rights is expressly made subject to the following conditions.

a. Attribution.

1. If You Share the Licensed Material (including in modified form), You must:

A. retain the following if it is supplied by the Licensor with the Licensed Material:

i. identification of the creator(s) of the Licensed Material and any others designated to receive attribution, in any reasonable manner requested by the Licensor (including by pseudonym if designated);

ii. a copyright notice;

iii. a notice that refers to this Public License;

iv. a notice that refers to the disclaimer of warranties;

v. a URI or hyperlink to the Licensed Material to the extent reasonably practicable;

B. indicate if You modified the Licensed Material and retain an indication of any previous modifications; and

C. indicate the Licensed Material is licensed under this Public License, and include the text of, or the URI or hyperlink to, this Public License.

2. You may satisfy the conditions in Section 3(a)(1) in any reasonable manner based on the medium, means, and context in which You Share the Licensed Material. For example, it may be reasonable to satisfy the conditions by providing a URI or hyperlink to a resource that includes the required information.

3. If requested by the Licensor, You must remove any of the information required by Section 3(a)(1)(A) to the extent reasonably practicable.

4. If You Share Adapted Material You produce, the Adapter's License You apply must not prevent recipients of the Adapted Material from complying with this Public License.

Section 4 – Sui Generis Database Rights.

Where the Licensed Rights include Sui Generis Database Rights that apply to Your use of the Licensed Material:

a. for the avoidance of doubt, Section 2(a)(1) grants You the right to extract, reuse, reproduce, and Share all or a substantial portion of the contents of the database;

b. if You include all or a substantial portion of the database contents in a database in which You have Sui Generis Database Rights, then the database in which You have Sui Generis Database Rights (but not its individual contents) is Adapted Material; and

c. You must comply with the conditions in Section 3(a) if You Share all or a substantial portion of the contents of the database.

For the avoidance of doubt, this Section 4 supplements and does not replace Your obligations under this Public License where the Licensed Rights include other Copyright and Similar Rights.

Section 5 – Disclaimer of Warranties and Limitation of Liability.

a. Unless otherwise separately undertaken by the Licensor, to the extent possible, the Licensor offers the Licensed Material as-is and as-available, and makes no representations or warranties of any kind concerning the Licensed Material, whether express, implied, statutory, or other. This includes, without limitation, warranties of title, merchantability, fitness for a particular purpose, non-infringement, absence of latent or other defects, accuracy, or the presence or absence of errors,

whether or not known or discoverable. Where disclaimers of warranties are not allowed in full or in part, this disclaimer may not apply to You.

b. To the extent possible, in no event will the Licensor be liable to You on any legal theory (including, without limitation, negligence) or otherwise for any direct, special, indirect, incidental, consequential, punitive, exemplary, or other losses, costs, expenses, or damages arising out of this Public License or use of the Licensed Material, even if the Licensor has been advised of the possibility of such losses, costs, expenses, or damages. Where a limitation of liability is not allowed in full or in part, this limitation may not apply to You.

c. The disclaimer of warranties and limitation of liability provided above shall be interpreted in a manner that, to the extent possible, most closely approximates an absolute disclaimer and waiver of all liability.

Section 6 – Term and Termination.

a. This Public License applies for the term of the Copyright and Similar Rights licensed here. However, if You fail to comply with this Public License, then Your rights under this Public License terminate automatically.

b. Where Your right to use the Licensed Material has terminated under Section 6(a), it reinstates:

1. automatically as of the date the violation is cured, provided it is cured within 30 days of Your discovery of the violation; or

2. upon express reinstatement by the Licensor.

For the avoidance of doubt, this Section 6(b) does not affect any right the Licensor may have to seek remedies for Your violations of this Public License.

c. For the avoidance of doubt, the Licensor may also offer the Licensed Material under separate terms or conditions or stop distributing the Licensed Material at any time; however, doing so will not terminate this Public License.

d. Sections 1, 5, 6, 7, and 8 survive termination of this Public License.

Section 7 – Other Terms and Conditions.

a. The Licensor shall not be bound by any additional or different terms or conditions communicated by You unless expressly agreed.

b. Any arrangements, understandings, or agreements regarding the Licensed Material not stated herein are separate from and independent of the terms and conditions of this Public License.

Section 8 – Interpretation.

a. For the avoidance of doubt, this Public License does not, and shall not be interpreted to, reduce, limit, restrict, or impose conditions on any use of the Licensed Material that could lawfully

be made without permission under this Public License.

b. To the extent possible, if any provision of this Public License is deemed unenforceable, it shall be automatically reformed to the minimum extent necessary to make it enforceable. If the provision cannot be reformed, it shall be severed from this Public License without affecting the enforceability of the remaining terms and conditions.

c. No term or condition of this Public License will be waived and no failure to comply consented to unless expressly agreed to by the Licensor.

d. Nothing in this Public License constitutes or may be interpreted as a limitation upon, or waiver of, any privileges and immunities that apply to the Licensor or You, including from the legal processes of any jurisdiction or authority.

Creative Commons is not a party to its public licenses. Notwithstanding, Creative Commons may elect to apply one of its public licenses to material it publishes and in those instances will be considered the "Licensor." Except for the limited purpose of indicating that material is shared under a Creative Commons public license or as otherwise permitted by the Creative Commons policies published at `creativecommons.org/policies`, Creative Commons does not authorize the use of the trademark "Creative Commons" or any other trademark or logo of Creative Commons without its prior written consent including, without limitation, in connection with any unauthorized modifications to any of its public licenses or any other arrangements, understandings, or agreements concerning use of licensed material. For the avoidance of doubt, this paragraph does not form part of the public licenses.

Creative Commons may be contacted at `creativecommons.org`.

Appendix E

References

Benedict, R. Ralph, *Electronics for Scientists and Engineers*, Prentice-Hall Incorporated, Englewood Cliffs, NJ, 1967.

Evans, David R. "Lecture 1" notes for Physics 445/545 (Microelectronic Device Fabrication I) course, Portland State University, Fall 2016.

Evans, David R. "Notes" for Physics 445/545 (Microelectronic Device Fabrication I) course, Portland State University, Fall 2016.

Hall, E.H., "On a New Action of the Magnet on Electric Currents", *American Journal of Mathematics* Volume 2, Number 3, pages 287-292, September 1879.

Haynes, J.R. and Shockley, W., "The Mobility and Life of Injected Holes and Electrons in Germanium", *Physical Review* Volume 81, Number 5, pages 835-843, 1 March 1951.

Kroemer, Herbert, "Quasielectric fields and band offsets: teaching electrons new tricks", Nobel Lecture for the 2000 Nobel prize in Physics, The Nobel Foundation, 8 December 2000.

Luther, Bryan A. "Semiconductors" handout for Physics 225 (Modern Physics II) course, Concordia College, Spring 2014.

"Semiconductors" lecture notes for lecture 45 in *6.007 Electromagnetic Energy: From Motors to Lasers* course, Massachusetts Institute of Technology, Spring 2011.

"Silicon Ingot Production Process for Wafers" Application Note, MicroChemicals GmbH, Ulm, Germany.

Zong, Linqi; Zhu, Bin; Lu, Zhenda; Cui, Yi; "Nanopurification of silicon from 84% to 99.999% purity with a simple and scalable process", Proceedings of the National Academy of Sciences of the United States of America (PNAS), Volume 112, Number 44, 19 October 2015.

Appendix F

Version history

This is a list showing all significant additions, corrections, and other edits made to this learning module. Each entry is referenced by calendar date in reverse chronological order (newest version first), which appears on the front cover of every learning module for easy reference. Any contributors to this open-source document are listed here as well.

21-22 October 2025 – minor clarifying edits to the Tutorial, and also added some instructor notes. Also corrected grammatical errors, courtesy of Wyatt Gleason.

19 March 2025 – added more recommendations for instructors to the Introduction chapter.

28 October 2024 – added an Animations chapter showing conduction of current through a PN junction (identical to the Animation contained in the *PN Junctions and Diodes* module).

10 October 2024 – divided the Introduction chapter into sections, one with recommendations for students, one with a listing of challenging concepts, and one with recommendations for instructors.

18-24 March 2024 – minor addition to the question "Solar panel output" in the Conceptual Reasoning section. Also made a minor word change (from "vaunted" to "prized"). Also added photograph of a silicon boule piece, as well as made improvements to some of the illustrations showing doping (image_4188 and image_4187).

26 October 2023 – corrected spelling of the word "diminution".

22 March 2023 – added instructor comments.

28 November 2022 – placed questions at the top of the itemized list in the Introduction chapter prompting students to devise experiments related to the tutorial content.

26 October 2022 – minor edits to the Tutorial, and more questions added to the Introduction chapter. Also re-worded the "Light-Emitting Diodes" Conceptual Reasoning question for better readability.

23 March 2022 – minor edits to the Tutorial.

28 October 2021 – minor improvements to images 4194 through 4197 in the Tutorial (made shading the same as image_4198). Also modified the "Solar panel output" and "Light-emitting diodes" Conceptual Reasoning questions to be less vague.

9 May 2021 – commented out or deleted empty chapters.

22 March 2021 – added more digits to Boltzmann's constant, and clarified that Kelvin is 273.15 more than Celsius.

5 October 2020 – significantly edited the Introduction chapter to make it more suitable as a pre-study guide and to provide cues useful to instructors leading "inverted" teaching sessions.

18-22 July 2020 – added content to the Tutorial.

1 December 2016 – document first published.

Index

Acceptor site, 19
Adding quantities to a qualitative problem, 80
Allotrope, 12
Annotating diagrams, 79
Atom, 8

Band, conduction, 12
Bandgap, 13, 15
Boule, silicon, 25

Carbon, 12
Charge carrier, majority, 18
Charge carrier, minority, 18
Checking for exceptions, 80
Checking your work, 80
Chip, silicon, 25
Circuit, electric, 3
Circuit, electronic, 3
Code, computer, 89
Concentration, dopant, 26
Conduction band, 12
Conductor, 11
Conductor, band diagram, 12
Conventional flow notation, 17
Current, 7

Degenerate semiconductor, 26
Dimensional analysis, 79
Diode equation, 5, 23
Donor site, 19
Doping, 16, 26

Edwards, Tim, 90
Electric, 3
Electric field, 14
Electron, 8
Electron, valence, 10
Electronic, 3
Energy, 7
Energy band diagram, 12, 13
Energy band diagram, tilted, 14
Energy gap, 11
Escape velocity, 8
Extrinsic semiconductor, 18, 26

Fab, semiconductor, 26
Fermi level, 19
Fullerene, 12

Germanium, 13
Graph values to solve a problem, 80
Graphite, 12
Greenleaf, Cynthia, 55

How to teach with these modules, 87
Hwang, Andrew D., 91

Identify given data, 79
Identify relevant principles, 79
Insulator, 11
Insulator, band diagram, 12
Intermediate results, 79
Intrinsic semiconductor, 18, 26
Inverted instruction, 87

Joule's Law, 7

Kirchhoff's Current Law, 7
Kirchhoff's Voltage Law, 7
Knuth, Donald, 90

Lamport, Leslie, 90
Limiting cases, 80
Load, 7

Majority carrier, 18

Metacognition, 60
Minority carrier, 18
Moolenaar, Bram, 89
Murphy, Lynn, 55

Neutron, 8

Ohm's Law, 5, 7, 23
Open-source, 89
Orbital, 8

Problem-solving: annotate diagrams, 79
Problem-solving: check for exceptions, 80
Problem-solving: checking work, 80
Problem-solving: dimensional analysis, 79
Problem-solving: graph values, 80
Problem-solving: identify given data, 79
Problem-solving: identify relevant principles, 79
Problem-solving: interpret intermediate results, 79
Problem-solving: limiting cases, 80
Problem-solving: qualitative to quantitative, 80
Problem-solving: quantitative to qualitative, 80
Problem-solving: reductio ad absurdum, 80
Problem-solving: simplify the system, 79
Problem-solving: thought experiment, 79
Problem-solving: track units of measurement, 79
Problem-solving: visually represent the system, 79
Problem-solving: work in reverse, 80
Proton, 8

Qualitatively approaching a quantitative problem, 80
Quantum, 8
Quartz, 25

Reading Apprenticeship, 55
Reductio ad absurdum, 80, 86, 87
Resistance, 7

Schoenbach, Ruth, 55
Scientific method, 60
Semiconductor, 7, 13
Shell, 8
Shockley diode equation, 5, 23
Shockley, William, 5, 23
Silicon, 13
Simplifying a system, 79
Socrates, 86
Socratic dialogue, 87
Source, 7
SPICE, 55
Stallman, Richard, 89
Subshell, 8

Temperature, 12
Thought experiment, 79
Tool, semiconductor manufacturing, 26
Torvalds, Linus, 89

Units of measurement, 79

Vacuum energy, 8
Valence, 10
Visualizing a system, 79
Voltage, 7

Wafer, silicon, 25, 27
Work in reverse to solve a problem, 80
WYSIWYG, 89, 90