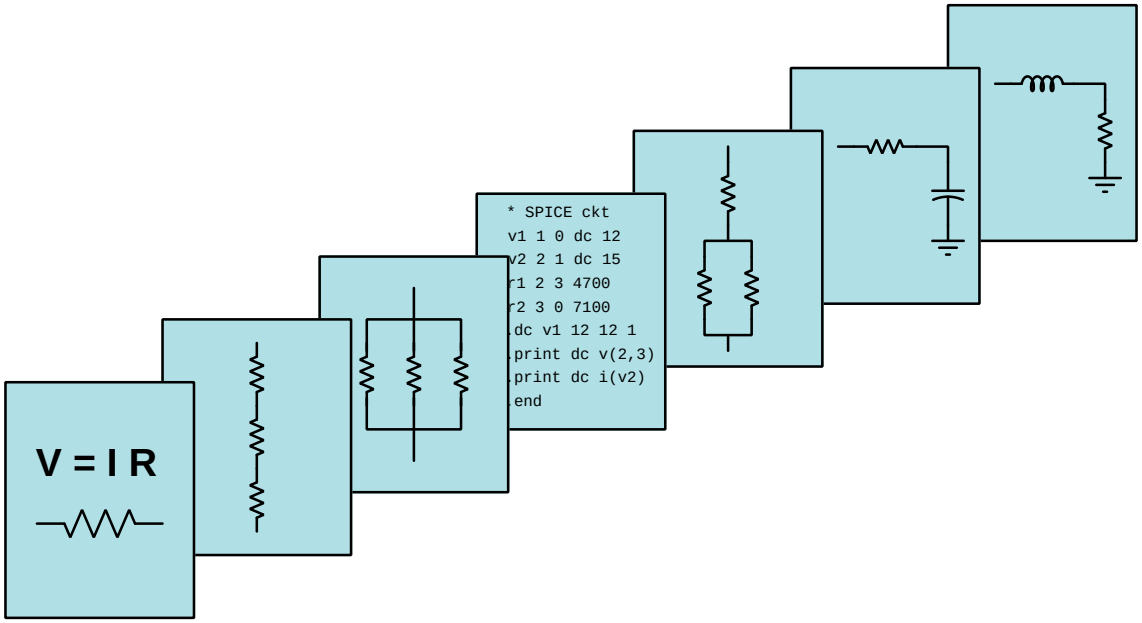
Modular Electronics Learning (ModEL) project



Semiconducting Electronic Devices

© 2016-2021 by Tony R. Kuphaldt – under the terms and conditions of the Creative Commons Attribution 4.0 International Public License

Last update = 9 May 2021

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**

2 Simplified Tutorial **5**

3 Full Tutorial **13**

4 Questions **33**

4.1 Conceptual reasoning . . . 37

4.1.1 Reading outline and reflections . . . 38

4.1.2 Foundational concepts . . . 39

4.1.3 BJT switching circuit configurations . . . 42

4.1.4 MOSFET switching circuit configurations . . . 45

4.2 Quantitative reasoning . . . 46

4.2.1 Miscellaneous physical constants . . . 47

4.2.2 Introduction to spreadsheets . . . 48

4.2.3 First quantitative problem . . . 51

4.2.4 Second quantitative problem . . . 51

4.3 Diagnostic reasoning . . . 52

4.3.1 Faults in a PNP-controlled MOSFET switching circuit . . . 53

5 Projects and Experiments **55**

5.1 Recommended practices . . . 55

5.1.1 Safety first! . . . 56

5.1.2 Other helpful tips . . . 58

5.1.3 Terminal blocks for circuit construction . . . 59

5.1.4 Conducting experiments . . . 63

5.1.5 Constructing projects . . . 67

5.2 Experiment: Driving a heavy load with a bipolar transistor . . . 68

5.3 Experiment: Driving a heavy load with a power MOSFET transistor . . . 70

A Problem-Solving Strategies **73**

B Instructional philosophy **75**
B.1 First principles of learning 76
B.2 Proven strategies for instructors 77
B.3 Proven strategies for students 79
B.4 Design of these learning modules 80

C Tools used **83**

D Creative Commons License **87**

E References **95**

F Version history **97**

Index **97**

Chapter 1

Introduction

This module introduces some of the elementary semiconductor devices and their operating characteristics, without delving much into solid-state physics. As such, this module strives to be very practical for students new to the subject, providing a simple overview of electronic device functionality.

A *semiconducting* solid material is one that is an electrical insulator for all practical purposes in its pure form under ordinary conditions, but which may be made to conduct electricity either by adding impurities and/or altering its environmental conditions. In other words, a semiconductor is a substance whose electrons may be stimulated to provide *controllable electrical conductivity*. This property makes semiconductors exceptionally useful for the construction of switching devices having no moving parts. The existence of controllable-conductivity devices also marks the threshold between *electric* and *electronic* circuits.

The simplest semiconducting electronic device is a two-terminal component called a *diode*, which functions as a one-way valve for electric current: allowing electricity to pass in one direction but not in the other.

A class of three-terminal semiconducting devices called *transistors* function as electrically-controlled switches, the conductivity between two terminals being adjusted by the presence of potential and/or current at the third terminal. Transistors may be broadly grouped into two categories: *bipolar junction transistors* (BJTs) and *field-effect transistors* (FETs), the former category of devices being controlled by the injection of charge carriers into the device from an external electric current and the latter category being controlled by the impression of an electric field from an external voltage. In either type of device, the controlling signal is much smaller in magnitude than the electrical power being switched on or off by the transistor to an external load. Therefore, transistors represent an example of *amplification*, where a low-power signal exerts control over a higher-powered signal.

Another class of semiconducting devices called *thyristors* behave similarly to transistors, except for their self-latching ability, where they remain in the "on" (conductive) state even after the controlling voltage or current signal ceases. *Silicon-controlled rectifiers* (SCRs) and *TRIACs* are two categories of thyristor commonly used.

Optoelectronic semiconducting devices interact with light, some of them controlled by another's

emitting light. *Light-emitting diodes* (LEDs) are a popular example of the latter while *solar cells* and *phototransistors* are an example of the former.

Chapter 2

Simplified Tutorial

An *electronic* device is one whose electrical conductivity may be controlled by manipulating the energy states of electrons inside of it. The first electronic devices to find popular application were *vacuum tubes* (alternatively called *electron tubes*), which used heated metal plates housed inside of glass envelopes from which atmospheric air had been pumped out. Electrical conduction occurred as a result of electrons leaving the heated metal plate and entering the vacuum space, their travel to a (cold) metal plate controlled by electric fields impressed between grids of wires placed in between the two metal plates.

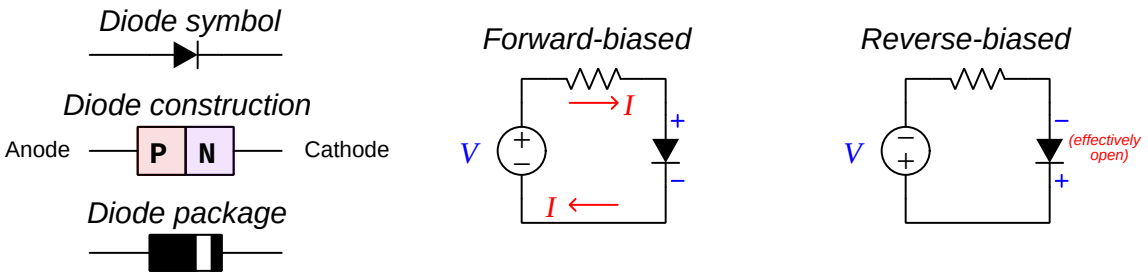
Vacuum tubes represented a major technical innovation, making possible such media advances such as radio and television, and even the first electronic computers. However, these devices were very energy-inefficient due to their need of being heated to high temperatures in order to liberate electrons from the hot metal plate surface into the vacuum space where their travel could be electrically controlled. Vacuum tubes, being made of glass and requiring the maintenance of high-vacuum states for proper operation, tended to be fragile. The development of *solid-state*[1] electronic devices based on *semiconductor* materials was pursued as an alternative to vacuum tube technology, and the resulting inventions paved the way for the modern electronic era.

A "semiconducting" material is a solid substance containing electrons that are *very nearly* in a state of free motion. Conductors, by contrast, contain electrons of sufficiently high energy to be unbound from their constituent atoms and able to freely drift throughout that solid's volume. The electrons within insulating substances are bound to their constituent atoms and cannot drift apart. The nearly-conductive nature of semiconducting substances permits their higher-energy electrons to be manipulated to produce electrical conductivity on demand, much like vacuum tubes but without the high power requirements of heating and without the mechanical challenge of maintaining vacuum-tight seals.

The conductivity of a semiconducting material may be influenced by a number of different factors including temperature, light, addition of impurities, and the presence of electric fields. Devices exist to exploit all of these factors for various practical purposes.

[1] "Solid-state" refers to a device where electrical charge carrier conduction occurs within a solid material, as opposed to a stream of electrons hurtling through a vacuum as was the case with electron tubes.

The simplest semiconductor device is called the *diode*, which functions as a one-way valve for electric current. Its schematic symbol is an arrowhead touching a perpendicular line, with the permitted direction of current (conventional-flow notation) following the arrowhead. When current flows through a diode, it is said to be *forward-biased*; when a diode blocks current due to the voltage polarity being "backwards" it is said to be *reverse-biased*:



Semiconductor diodes are manufactured from a sandwich of two differently-alloyed layers of semiconducting material (called *P-type* and *N-type*), and their function is based on a thin region at the junction of those layers being normally depleted of free charge carriers (i.e. electrically insulating). Applying a "forward-bias" voltage across the diode's terminals has the effect of collapsing this "depletion region" to permit conduction from one layer to the other. If a reverse-bias voltage is applied to the diode, its internal depletion region grows larger and the diode becomes a better insulator, thereby blocking passage of current. The energy required to collapse this depletion region is small, but results in a fairly constant forward-bias voltage across the diode necessary for conduction. In other words, a forward-biased diode does not conduct as well as a metal wire (which poses negligible energy loss to passing charge carriers) does, but demands a certain amount of energy from each passing charge carrier for passage across the depletion region. This energy demand is not a linear function of current, and so diodes do *not* follow Ohm's Law. Instead, the forward voltage drop of a semiconductor diode is primarily a function of the type of semiconducting material it is made of, and secondarily a function of temperature[2]. Silicon diodes typically exhibit 0.5 to 0.7 Volts of forward-bias drop, while germanium diodes typically drop 0.3 Volts or less.

Many different types of diodes exist, some designed to function as *rectifiers* (acting as one-way valves for electric power), others designed to produce very constant amounts of voltage drop, and yet others designed to emit light (i.e. *light-emitting diodes*, or LEDs). A *photovoltaic cell* or *solar cell* is a special-purpose diode with its PN junction exposed to sunlight to produce electricity when illuminated.

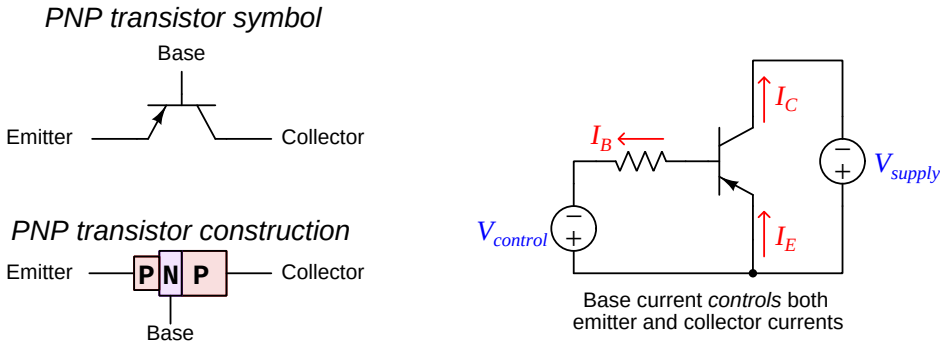
Diodes, although useful in their own ways, do not allow one electrical signal to control another, which is the ultimate goal of an *electronic* device. For this function we need a different kind of electronic device called a *transistor*.

[2]This forward-bias voltage drop becomes smaller as temperature rises, because thermal energy assists charge carriers crossing the depletion zone and therefore less electrical energy is needed. Interestingly, this relationship between forward voltage drop and temperature for a semiconductor diode is precise enough to permit the use of diodes as electronic temperature sensors!

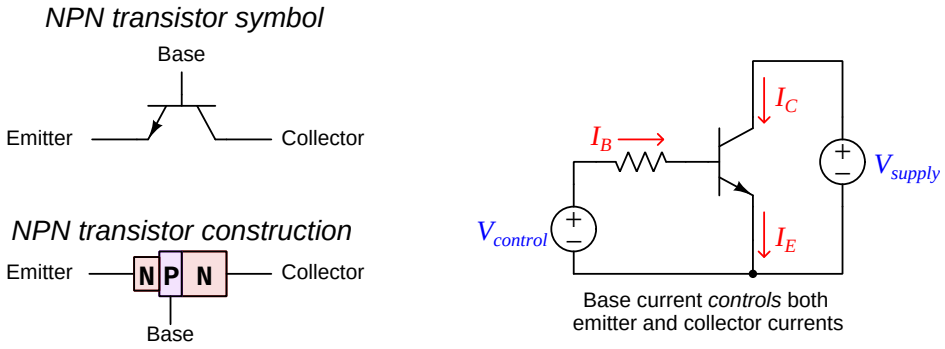
One category of transistor is the so-called *bipolar junction transistor*, or BJT, comprised of a three-layer sandwich of differently-alloyed semiconductor material. With two different types of alloying (often referred to as *doping*), P-type and N-type, bipolar junction transistors may be made from PNP layers or from NPN layers. The most significant difference between these two types of BJT is the required polarities to make them function.

Either type of BJT works on the principle of *charge carrier injection*. Like PN junction diodes, non-conducting depletion regions exist within the sandwich structure of the transistor, prohibiting conduction in its normal state. In order to make a BJT conduct current between its outer two terminals (called the "collector" and "emitter"), a small amount of current must pass through the middle terminal (called the "base"). From the controlling source's perspective, the base-emitter junction of a BJT resembles a diode, and requires the same amount of forward-bias voltage to collapse that depletion region to permit current. The base layer of a bipolar transistor, however, is manufactured to be extremely thin, and so the vast majority of the charge carriers liberated by the controlling source become "swept" into the base-collector depletion region which makes the transistor become electrically conductive from end to end.

A PNP transistor is shown here, where I_B is a very small controlling current and I_C is a much larger controlled current:

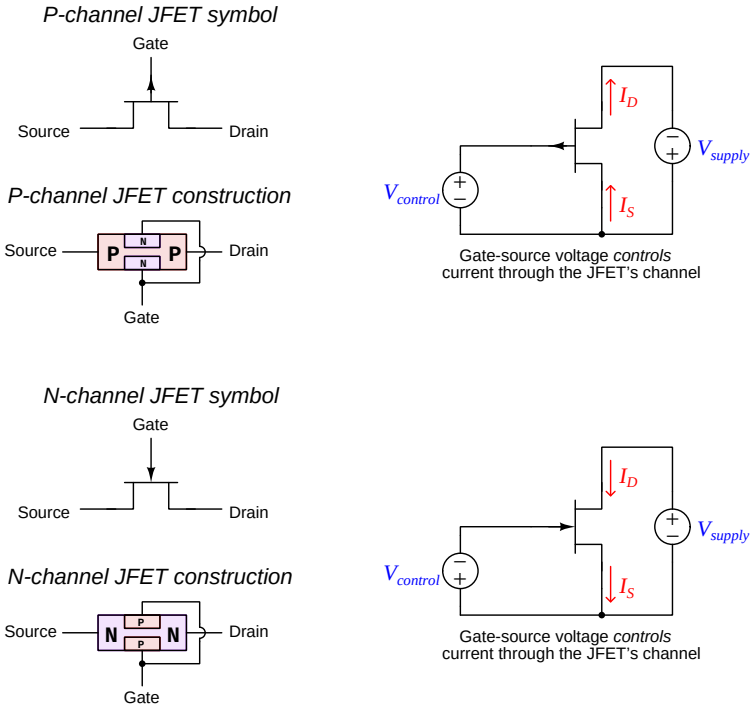


An NPN transistor is shown in the following illustration. Once again, I_B is a very small controlling current and I_C is a much larger controlled current, the principal difference between the PNP and NPN transistor functions being the directions of current. Neither transistor will function if the polarities are incorrect:



Another category of transistor is the so-called *field-effect transistor*, or FET. These transistors, while also made of P-type and N-type semiconducting materials, operate on an entirely different principle than charge carrier injection. In a field-effect transistor, the end terminals (called *source* and *drain* rather than *emitter* and *collector*) attach to an uninterrupted length of alloyed semiconductor, either P-type or N-type. The alloying, or "doping" of this material makes it electrically conductive by default, a narrow *channel* through which electric charge carriers may flow with relatively little resistance. Adjacent to the conductive channel is a piece of semiconducting material with the opposite alloying, creating a non-conducting depletion region at the PN junction (as in a diode). if a voltage is impressed across this PN junction between the *gate* terminal and the conductive channel, it will cause the depletion region to expand or collapse (depending on the polarity of that voltage in relation to the P- and N-types). An expanding depletion region will "pinch off" the conductive channel between source and drain terminals, turning the transistor off, while a collapsed depletion region will "open up" the channel to turn the transistor on.

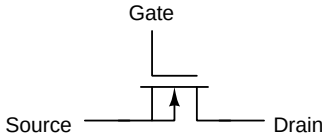
Junction Field-Effect Transistors (JFETs) have their gate terminals directly connected to the oppositely-doped semiconductor regions. The first type of JFET shown here is an P-channel, and the next is an N-channel type. Note as with the PNP and NPN varieties of bipolar transistor, the only difference in usage for these two JFET types is polarity:

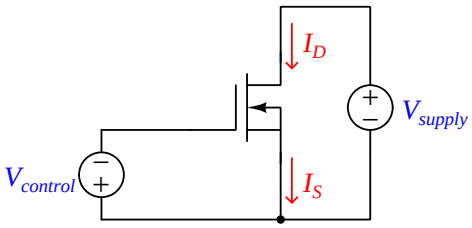


Insulated-Gate Field-Effect Transistors (IGFETs), more commonly known as *Metal-Oxide-Semiconductor Field-Effect Transistors* (MOSFETs) differ from JFETs in the addition of an insulating layer of metal oxide between the gate terminal and the semiconductor material. Besides P-channel and N-channel types, MOSFETs are also manufactured in two sub-categories: *enhancement* and *depletion*. An "enhancement" MOSFET is a normally-off device, constructed so that the depletion region fully cuts off the channel's conductivity with no gate voltage applied. A "depletion" mode MOSFET is much like a JFET, being a normally-on device which may be made more or less conductive by the application of a voltage between the gate terminal and the channel.

Rather than show all four possible combinations of MOSFET[3], only the N-channel variety will be shown for simplicity. As with other N- versus P-type transistor constructions, the significant difference is in the polarities necessary to proper operation. First, the "depletion" type of MOSFET:

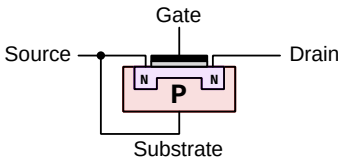
N-channel D-type MOSFET symbol

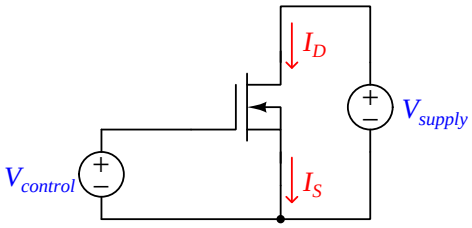




Gate-source voltage *restricts*
current through the MOSFET's channel

N-channel D-type MOSFET construction

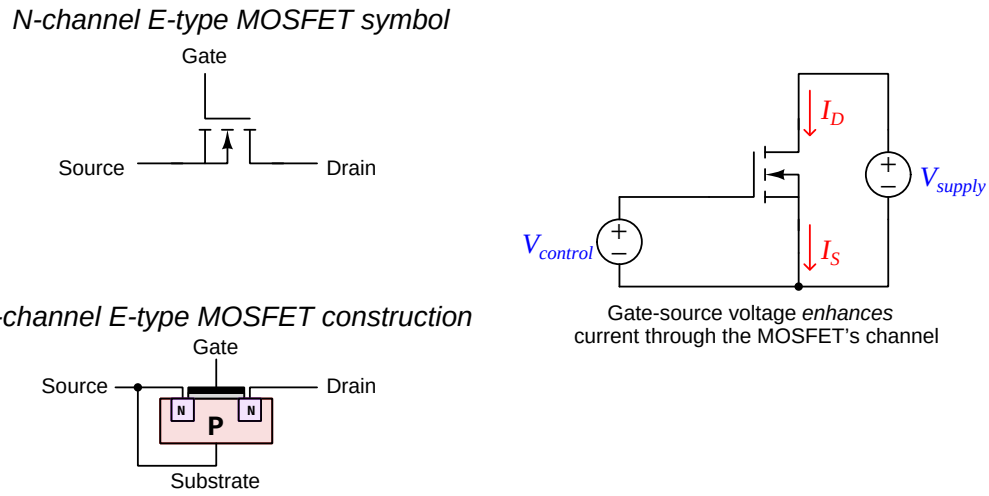




Gate-source voltage *enhances*
current through the MOSFET's channel

[3] N-channel enhancement-type, P-channel enhancement type, N-channel depletion type, N-channel enhancement type

Next, the "enhancement" type of MOSFET:

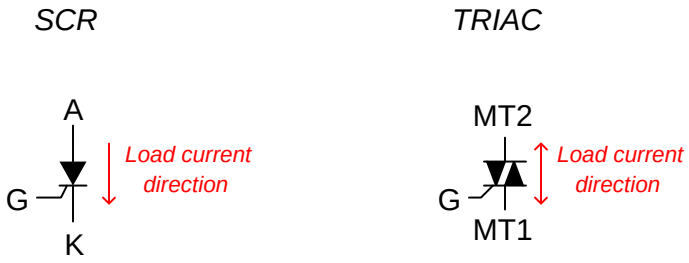


With either depletion- or enhancement-type MOSFETs, the gate-source voltage polarity necessary to make a MOSFET turn on (or turn on *more*) depends solely on its channel doping. For an N-channel MOSFET, the gate must be positive relative to the source in order to enhance the channel. For a P-channel MOSFET, the gate must be negative relative to the source in order to enhance. Reversing this polarity in either case restricts current through the MOSFET.

A hybrid style of transistor integrating BJT and MOSFET technologies is called the *Insulated Gate Bipolar Transistor*, or *IGBT*, combining the high current capacity of a BJT with the low controlling-current demands of a MOSFET.

Another class of semiconductor devices called *thyristors* also allows a small electrical signal to control a larger one, but what distinguishes thyristors from transistors is that a thyristor will remain in the "on" state after the initial triggering signal is removed, whereas a transistor returns to its previous state when the controlling signal ceases. That is to say, a thyristor internally provides its own maintaining-on stimulus once triggered by an external signal. This makes thyristors entirely on-off devices, whereas transistors which do not "latch" in their last state may be finely adjusted between conducting and non-conducting states.

Two popular types of thyristors include *Silicon-Controlled Rectifiers* (SCRs) and *TRIACs*, the fundamental difference between these two being that TRIACs are designed to carry load current in either direction while SCRs are designed to carry load current in one direction only, as shown in the following illustration:



Chapter 3

Full Tutorial

Semiconductors are materials capable of *controlled electrical conduction*: that is, a semiconductor may become electrically conductive under certain conditions and become (nearly) electrically insulating under other conditions[1]. This property makes semiconductors extremely useful as the basis for electronic components. The purpose of this tutorial is to give a broad overview of different semiconductor device characteristics without delving too deeply into their inner workings. Other tutorials in this series will devote attention to those inner details.

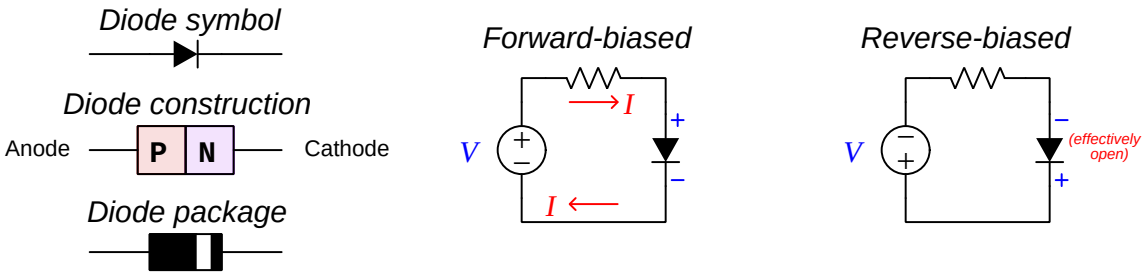
The most popular material used for the fabrication of semiconductor devices at the time of this writing is *silicon*, and so this will be the assumed substrate for all devices outlined in this tutorial. Like all semiconductor materials, silicon must be refined to an extremely high degree of purity to make it suitable for this application. In its pure, crystalline state it is referred to as an *intrinsic* semiconductor, with very few available charge carriers at room temperature, making it an electrical insulator. Carefully selected impurities called *dopants* may be added to pure silicon to create free charge carriers within the crystalline lattice of the material, making it an *extrinsic* semiconductor.

An *N-type* extrinsic semiconductor is formed by the addition of a dopant with atomic electrons at relatively high energy levels compared to the valence electrons of the pure silicon. These high-energy electrons are able to move freely through the "conduction band" of the material as charge carriers, making the material electrically conductive. A *P-type* extrinsic semiconductor is formed by the addition of a dopant creating electron vacancies in the "valence band" of the material. These vacancies (called *holes*) permit the motion of valence-level electrons from atom to atom. The holes may be viewed as charge carriers themselves (since the hole "moves" as soon as an electron vacates its former position to fill it), making the material electrically conductive.

By itself, a specimen of N-type or a P-type material is merely a conductor, the degree of electrical conductivity determined by the concentration of the doping. More doping equals more charge carriers, equating to better conductivity (i.e. less resistance), all other factors being equal. True semiconducting behavior does not emerge until we join N-type and P-type materials together to form a *PN junction*. The interactions between different types of charge carriers (electrons versus holes) in the junction of dissimilarly-doped materials gives semiconductors their uniquely useful properties.

[1] The conditions referred to include electrical and optical excitation. That is to say, we may control the electrical conductivity of a semiconductor by exciting its atoms with an electrical signal or with photons of light.

The simplest semiconductor device is the *diode*, made of a single PN junction. Its primary purpose is to serve as a one-way "valve" for electric current, permitting current in one direction but blocking it in the other direction:



P-type material forms the anode, while N-type material forms the cathode. At the boundary between these two halves of the diode, free electrons from the N-side fill holes in the P-side, creating a narrow zone near the P-N material interface devoid of charge carriers called the *depletion region*. Being depleted of all but a very few charge carriers, this zone inside the diode is electrically non-conductive. When the diode is *forward-biased*, this depletion region is forced to collapse, allowing free electrons from the "N" side to cross the junction and fall into holes moving toward the junction from the "P" side, permitting conduction through the diode. When the diode is reverse-biased, the depletion region expands, with holes and free electrons each moving away from the junction to maintain the diode in its default non-conducting[2] state.

PN diode junctions do not obey Ohm's Law, but rather exhibit a nonlinear response to applied voltage. The *Shockley diode equation*, named after William Shockley, predicts forward-bias current as a function of applied voltage and temperature:

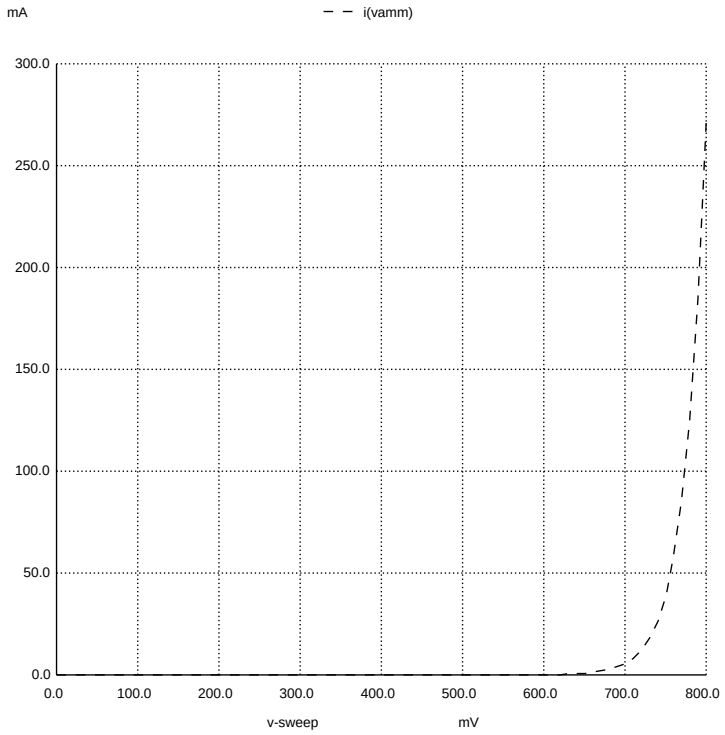
$$I = I_S \left(e^{\frac{qV}{nKT}} - 1 \right)$$

Where,

I = Forward-bias current through the diode, Amperes
I_S = Reverse-bias saturation current 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

[2] 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*.

The following graph shows calculated current values for a typical diode in forward-bias mode, for an applied voltage varying from 0 Volts to 800 milliVolts (0.8 Volts):

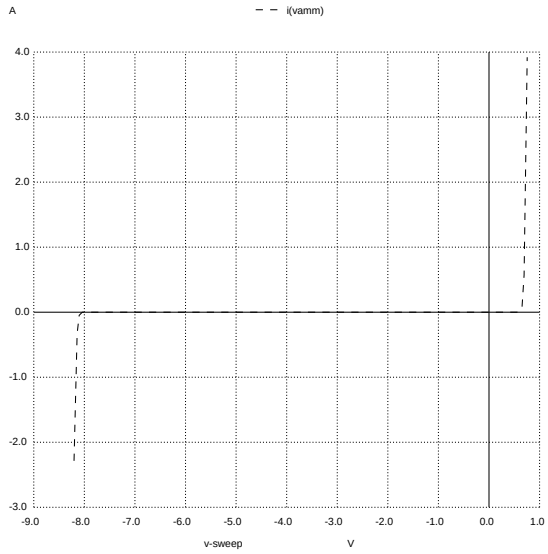


As shown by the graph, the diode's forward current does not become significant until the applied voltage is well in excess of 0.5 Volt. Calculated values for this particular diode include 567.7×10^{-15} Amperes at 0.1 Volt, 52.06×10^{-9} Amperes at 0.4 Volts, 0.1188×10^{-3} Amperes at 0.6 Volts, 5.675 mA at 0.7 Volts, and 271.1 mA at 0.8 Volts. For reasons of simplicity, a constant forward voltage drop of 0.5 Volts to 0.7 Volts is often assumed for silicon diodes at modest current values. This voltage drop means that diodes behave as electrical *loads*, with charge carriers losing some energy in the form of heat and light as they pass through the diode in its conductive state.

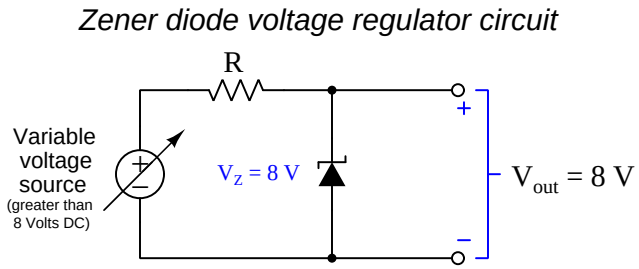
In reverse-bias mode, current through a diode saturates at an extremely low value called the *reverse saturation current* (I_S in the Shockley diode equation). For most practical purposes we may assume this value to be negligible.

Typical diode ratings include maximum forward current (I_F, based on the diode's ability to dissipate heat, since the forward voltage drop means it operates as a load when conducting), forward voltage drop at some specified current (V_F), maximum reverse voltage (V_R, also called *peak inverse voltage* or *blocking voltage*), and reverse recovery time (t_{rr}). The reverse recovery time for a diode is the amount of time required for the diode to transition from forward (conducting) mode to reverse (blocking) mode.

If we exceed the maximum reverse voltage rating of a diode, it will experience *breakdown* and permit a reverse current to flow. The amount of current through the diode increases exponentially as voltage exceeds this breakdown limit, similarly to how current increases exponentially as voltage exceeds the typical forward voltage (0.7 Volts for a silicon diode). So long as this current is limited by some other means, the diode will not be damaged. A SPICE computer analysis shows the behavior of a diode with a breakdown voltage rating of 8 Volts, as the applied voltage varies from −8.2 Volts to +0.75 Volts:



The breakdown voltage of a diode may be engineered by controlling the doping concentration within the semiconducting material. This allows construction of diodes with specific breakdown voltage ratings, useful as voltage-regulating devices. Such diodes are called *Zener diodes*, their breakdown voltage referred to as the *Zener voltage* (V_Z). A simple voltage regulator circuit shows how a variable voltage source may be regulated down to a constant voltage value:

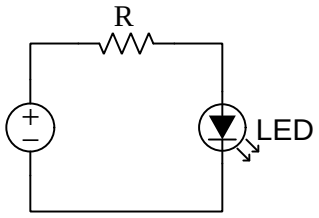


Current through the Zener diode is limited by the resistor (R), its value chosen so that the greatest expected voltage output by the variable source will not exceed the Zener diode's current rating. Note how the Zener diode is connected with its cathode positive and its anode negative, which is "backward" compared to normal diode operation but necessary to exploit breakdown.

Another useful characteristic of PN junction diodes is the emission of *light* from the junction when passing a forward current. We have already learned that diodes function as electrical loads, which means some energy is lost by charge carriers moving through. This energy transfer takes the form of both heat and light, the color (frequency) of that light determined by the energy shift experienced by charge carriers traversing the PN junction. If that energy shift corresponds[3] to a photon of light in the visible frequency range and the PN junction is given a transparent window for viewing, we will be able to see light emitted by the PN junction when it is forward-biased. This is the basis of a *Light-Emitting Diode*, or LED.

The fact that an LED is a diode, and passes current in only one direction, is usually incidental to its purpose as a light source. That is to say, in most cases we don't care that an LED only functions with one applied polarity. LEDs tend to be highly efficient as light sources, and are capable of extremely fast switching times. An idiosyncrasy of LEDs is their very narrow operating voltage range, synonymous with their forward voltage rating.

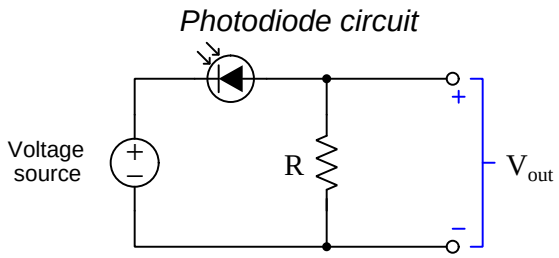
A typical LED circuit is shown here, with current-limiting resistor (R) to allow reliable operation despite variations in power supply voltage. If it were not for the resistor, the power source would have to be precisely matched[4] to the LED to permit ample light generation without damage to the LED. Note the pair of arrows pointing away from the LED, indicating light emission:



[3] The emitted photon's frequency (color) is a direct function of the *band gap* within the PN junction, and this is determined by the chemical composition of the junction's semiconducting material. Silicon's band gap is too small to produce visible photons, and so visible-light LEDs utilize other semiconducting materials such as Gallium Arsenide (GaAs) having larger band gaps.

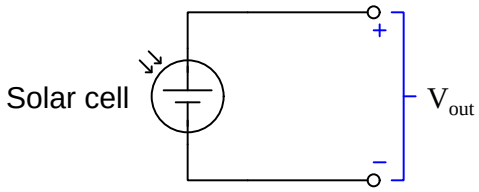
[4] To understand this problem, examine the graph of current versus voltage for any forward-biased PN junction diode: you will note that current through the diode is negligible until a certain minimum voltage is applied (approximately 0.7 Volts for a silicon diode). However, just a little bit more voltage than that results in excessive current which will destroy the diode! Adding a series current-limiting resistor to the LED circuit allows the power supply's voltage to vary a bit more without falling outside the diode's operating range, because the resistor drops any difference in voltage between the LED's forward rating and the power supply's voltage.

Light is capable of creating electron-hole pairs when striking a PN junction, making that junction electrically conductive in proportion to the intensity of light falling on it. If a reverse-bias voltage is applied to the PN junction, it will behave as an insulator (passing no current) until light falls on it. In this usage, the PN junction is called a *photodiode*:

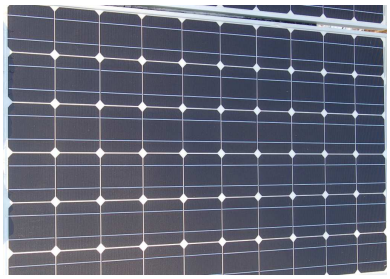


Note the pair of arrows facing toward the photodiode, indicating light reception.

This same phenomenon of light-induced charge carrier pair production may be used to generate a voltage from light striking a PN junction. This is the basis of a *photovoltaic cell*, otherwise known as a *solar cell*:



Each solar cell outputs a voltage equivalent to the forward voltage of its PN junction, with current output being a function of the number of photons of light collected by the junction. Therefore, solar cells are built as wide, thin wafers to optimize their light-collecting potential. These cells are connected in *series* arrays called *solar panels* in order to generate usable voltages.

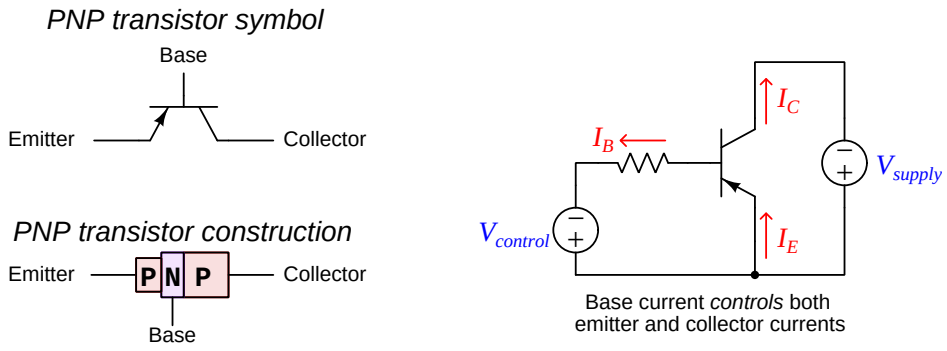


The solar panel shown above contains 60 square cells, all connected in series. Together, these cells output a maximum power of 250 Watts when exposed to bright sunlight and connected to a suitable load (i.e. a load with an impedance equal to the Thévenin equivalent impedance of the solar panel, as predicted by the Maximum Power Transfer Theorem).

Perhaps the most useful class of semiconductor device is the *transistor*. Whereas diodes switch between conducting and insulating based on the polarity of the applied voltage, transistors switch between conducting and insulating based on the application of a separate electrical signal. In other words, *transistors allow one electrical signal to control another*, much like an electromechanical relay but without any moving components. Furthermore, transistors are able to *regulate* electrical current somewhat like a resistor, neither fully on nor fully off. This allows transistors to *amplify* signals: allowing a relatively small electrical signal to modulate a relatively large electrical signal.

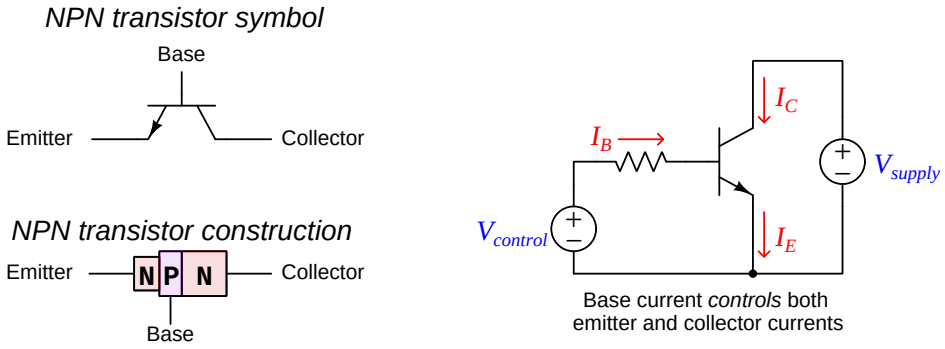
An important concept when dealing with any amplifying device is that of *gain*: the ratio between the controlled signal and the controlling signal. Ideal amplifying devices have gain values larger than one (1), which means the controlled signal is larger than the controlling signal.

The first commercially viable form of transistor was the *bipolar junction transistor*, or *BJT*, so named because it is a "sandwich" structure of three semiconductor material layers with alternating doping, either NPN or PNP. The following illustration shows a PNP transistor:

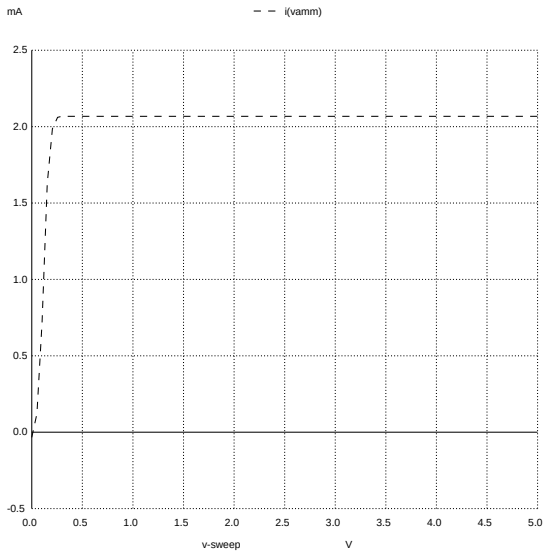


The NPN or PNP "sandwich" structure of a bipolar junction transistor creates two depletion regions: one at each interface of P-to-N material. This fact makes bipolar transistors "normally off" devices: in the absence of any external stimulus there is no conductive path between any of the transistor's three terminals. As shown in the above diagram, a controlling signal ($V_{control}$) is used to forward-bias the emitter-base PN junction, which collapses that depletion region and allows conduction of charge carriers between those two layers of the transistor. A depletion region still exists between the collector and base layers, but due to the thinness of the base layer some of the charge carriers flowing into it from the emitter layer end up becoming injected into that base-collector depletion region, making it electrically conductive. In a well-designed bipolar junction transistor a *majority* of charge carriers flowing into the base layer from the emitter will find themselves "swept" through this depletion region and into the collector by the electric field of the supply voltage, with only a small fraction of the emitter current exiting out the base terminal. The rate of charge carriers available to be swept into the collector is primarily determined by the control signal applied to the base terminal, which means a small controlling signal is able to influence a much larger controlled signal. This is the essence of *amplification*: a relatively small signal exerts control over a relatively large signal.

NPN transistors are identical in function to PNP transistors, except with reversed voltage polarities and current directions[5]:



An interesting characteristic of bipolar junction transistors is that the amount of current through the collector is primarily a function of the controlling signal at the base and not the power supply voltage between the collector and emitter terminals (V_{CE}). Here we see a SPICE computer analysis showing collector current (I_C) quickly rising and then leveling off as V_{CE} is increased from 0 Volts to 5 Volts, with the controlling base current (I_B) held at a constant value:

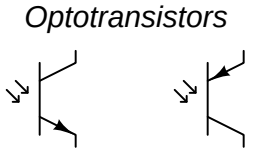


[5]The types of charge carriers injected vary as well. In a PNP transistor, the charge carrier type being injected into the N-type base layer are *holes* from the P-type emitter, which are swept into the P-type collector where they natively flow. In an NPN transistor, it is *electrons* which are injected from the N-type emitter through the P-type base and into the N-type collector. This is why bipolar junction transistors are classified as *minority carrier* devices: conduction is based on the flow of charge carriers through a "foreign" base region (i.e. holes injected into an N-type base for a PNP, and electrons injected into a P-type base for an NPN).

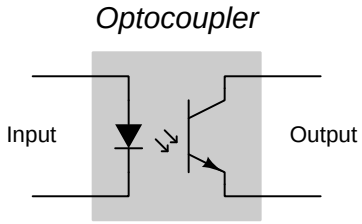
The reason for this current-limiting behavior is rooted in the action of charge carriers within the base layer of the transistor. When there is no base current, the base layer of the transistor acts as an insulator due to the depletion regions of both PN junctions. As soon as the base-emitter junction becomes forward-biased by the controlling source, charge carriers get injected into the base layer from the emitter. The power supply voltage between collector and emitter "sweeps" most of these charge carriers through the base layer and into[6] the collector to form a current, but the magnitude of this collector current is set by the controlling source and not the power supply. An increased power supply voltage will cause charge carriers to drift at a higher velocity (i.e. each one passing through the base layer at a higher speed), but it does little to increase the *number of charge carriers passing through per second of time* which is the definition of electric current.

The ratio of collector current to base current is called *beta* ($\beta = \frac{I_C}{I_B}$), and represents the current amplification factor – or current *gain* – of the transistor: the degree to which a small signal is able to control a larger signal. Beta varies widely with temperature and with collector-to-emitter voltage (V_{CE}), but is generally in the order of tens for high-power transistors to hundreds for small-signal transistors. Again, the reason why the collector current is so much larger than the base current is the relatively large collector-to-emitter voltage "sweeping" those charge carriers into the collector before they have a chance to exit the base terminal.

Just as diode PN junctions may be influenced by light, the base-emitter PN junction of a bipolar transistor may be similarly activated by light. This fact makes possible the construction of *optotransistors* – light-activated transistors:



Optotransistors may be equipped with transparent windows allowing external light to fall on the PN junction, or they may be packaged inside an opaque structure alongside an LED light source as an *optocoupler*:



Optocouplers are very useful when complete isolation must exist between two circuits. As you can see from the diagram, the input circuit's influence over the output circuit occurs via *light* rather

[6] For NPN transistors, where the emitter is made of N-type semiconductor material with the majority of charge carriers being free electrons, those *free electrons* are the charge carriers being injected into the base and then swept into the collector. For PNP transistors, *holes* are being injected into the base from the emitter and then swept into the collector by V_{CE}.

than an electrical connection which means the two circuits are electrically insulated from each other. This means the input and output circuits need not share any common reference points, common voltage levels, or any other shared electrical parameter. Typical breakdown voltage ratings between input and output for an optocoupler are in the thousands of Volts.

An entirely different form of transistor developed years after the BJT is the *field-effect* transistor[7]. Whereas bipolar junction transistors operate by injecting foreign charge carriers into a normally non-conducting "base" layer, field effect transistors utilize the effect of an electric field on a current-carrying semiconducting "channel". The intensity of the electric field controls the effective width of that conductive channel, thereby controlling the channel's ability to conduct a current.

As you should recall from your study of capacitors, the intensity of any electric field is directly proportional to the amount of voltage between two points and inversely proportional to the distance separating those two points. In a FET, the distance separating the various terminals is fixed, which means the electric field intensity is strictly a function of applied voltage. Therefore, field-effect transistors are voltage-controlled devices.

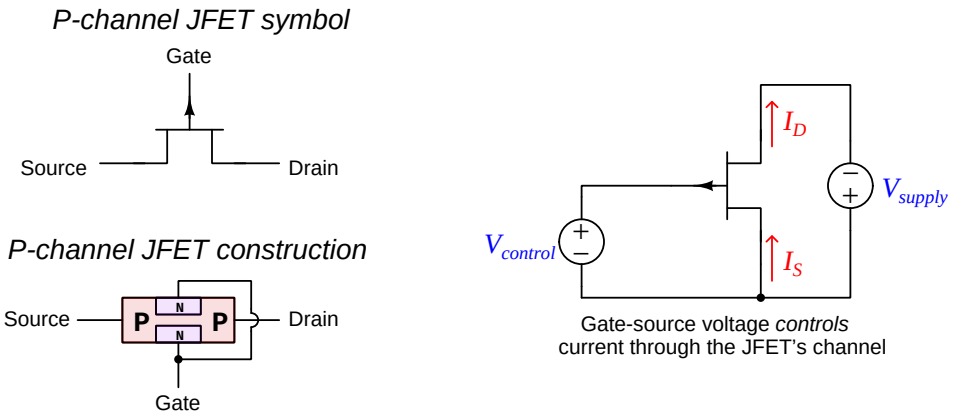
Like all transistors, field-effect transistors (FETs) *amplify* electrical signals. That is to say, they allow one electrical signal to exert control over another electrical signal. The ratio of the controlled signal to the controlling signal is called *gain*, and it is generally greater than one (1). The equivalent of beta for a FET is called *transconductance*: the ratio[8] of controlled current to controlling voltage ($G = \frac{I_{controlled}}{V_{controlling}}$).

Transconductance for a FET, like beta for a BJT, varies widely over different operating conditions. Amplifier designers use clever techniques to achieve stable, dependable gain values while using variable-gain transistors. These techniques will be discussed in detail in future modules.

[7] It is worth noting that field-effect transistors behave in remarkably similar ways to *vacuum tubes* which were the electronic device of choice prior to the advent of the transistor. Vacuum tubes also relied on a small controlling voltage to modulate a pathway for a controlled current, but they did so using metal plates and grids housed inside evacuated glass tubes. Whereas transistors control the flow of charge carriers within a crystalline solid (which is why they are referred to as "solid-state" devices), vacuum tubes controlled the flow of electrons through a small region of empty space. Tubes are a fascinating area of study unto themselves, but are considered obsolete for all but a few specialty applications due to the general superiority of solid-state transistor devices. In order for vacuum tubes to function, they required one of their internal metal surfaces to be heated red-hot in order to provide enough thermal energy for electrons within the metal to "boil off" into the surrounding space and exit the metal completely. Maintaining the temperature of a metal plate red-hot required substantial amounts of power. Furthermore, the high vacuum state necessary inside the tube's glass envelope to avoid collisions between the electron stream and gas molecules in that space was difficult to maintain for years on end due to the complexity of maintaining a perfect air-tight seal.

[8] Beta is a unitless ratio, being a ratio between two current values. Transconductance, however, actually has its own unit of measurement (Siemens) because it is a ratio of current (Amperes) to voltage (Volts).

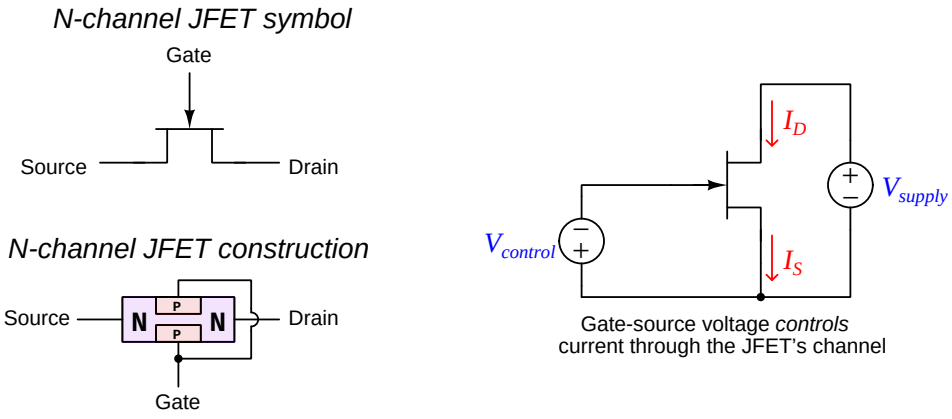
The first type of FET we will explore is called the *Junction Field-Effect Transistor*, or *JFET*. Like the BJT, the JFET is a three-terminal device. Unlike the BJT, the controlled current does not traverse any PN junctions but instead takes a straight path through an uninterrupted length of either N-type or P-type semiconductor material[9]. This make the JFET a *normally-on* device as opposed to the BJT which is *normally-off*. The following illustration shows a *P-channel* JFET:



Current through the JFET's channel is restricted by the narrowness of the channel. That channel is made even narrower by the application of a reverse-bias voltage across the PN junction formed between the gate and source terminals, which has the effect of expanding the depletion region and thereby narrowing the channel. Gate current is negligible, since the gate-source PN junction is reverse-biased and therefore non-conducting.

[9] This is why field-effect transistors are classified as *majority carrier* devices: conduction from one end to the other is based on charge carriers native to the material type. For a P-channel FET, holes are the charge carriers comprising the majority of current; for an N-channel FET it is electrons. Contrast this against bipolar junction transistors which are classified as *minority carrier* devices: current from one end of the device to the other is predicated on "foreign" charge carriers being injected into the depletion region of a different type of material (e.g. holes injected into an N-type base layer for a PNP, and electrons injected into a P-type base later for an NPN).

An *N-channel* JFET uses oppositely-doped semiconductor regions, and operates with a reverse polarity compared to the P-channel JFET:

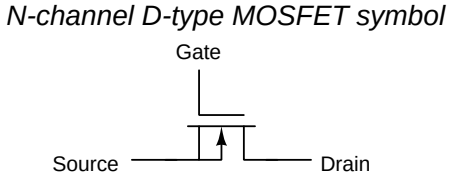


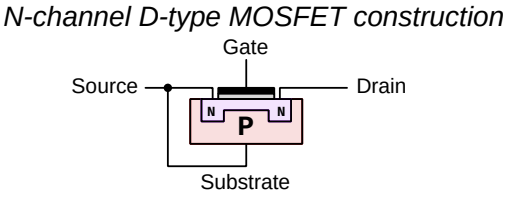
It should be noted that the gate-to-channel PN junction is rather fragile, not being designed to carry any substantial current, since JFETs are designed for their gate-channel junctions to always be operated in reverse-bias mode. Unlike BJTs where the transistor gets turned on by injecting charge carriers into the base region, JFETs already possess a conductive channel which gets pinched off by a widening depletion region when the gate-channel junction is reverse-biased.

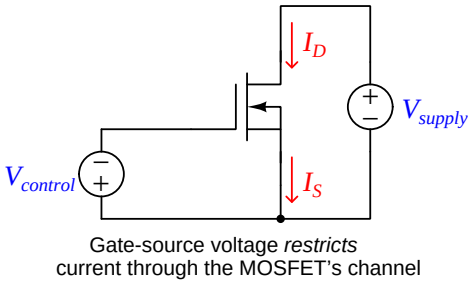
Another type of field-effect transistor uses a thin layer of insulating material (usually silicon dioxide, SiO_2) as a dielectric separating the metal gate terminal from the channel, and is called the *Metal Oxide Semiconductor Field-Effect Transistor*, or *MOSFET*[10]. Like JFETs, MOSFETs come in two different channel types: N-channel and P-channel. Unlike JFETs, MOSFETs are manufactured in two different configurations: *Depletion* and *Enhancement*. Thus, there are four different types of MOSFETs: N-channel depletion, P-channel depletion, N-channel enhancement, and P-channel enhancement.

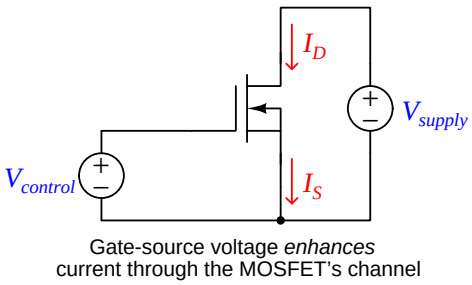
While this array of options may seem bewildering, the different types of MOSFETs are surprisingly easy to understand. They all work on the principle of controlling the effective width of a conductive channel by means of an applied gate-to-source voltage (V_{GS}). The insulating layer between the gate and channel means we are free to apply either polarity of V_{GS} without fear of forward-biasing a delicate PN junction. Depletion-type MOSFETs are normally-on devices which may be pinched off the same as a JFET, or caused to turn on more fully with the opposite V_{GS} polarity. Enhancement-type MOSFETs are normally-off devices which may be turned on by the application of V_{GS} of the correct polarity.

The following illustrations show N-channel depletion (D-type) MOSFETs. P-channel MOSFETs are not shown, but differ only in the material doping, voltage polarities, and current directions:





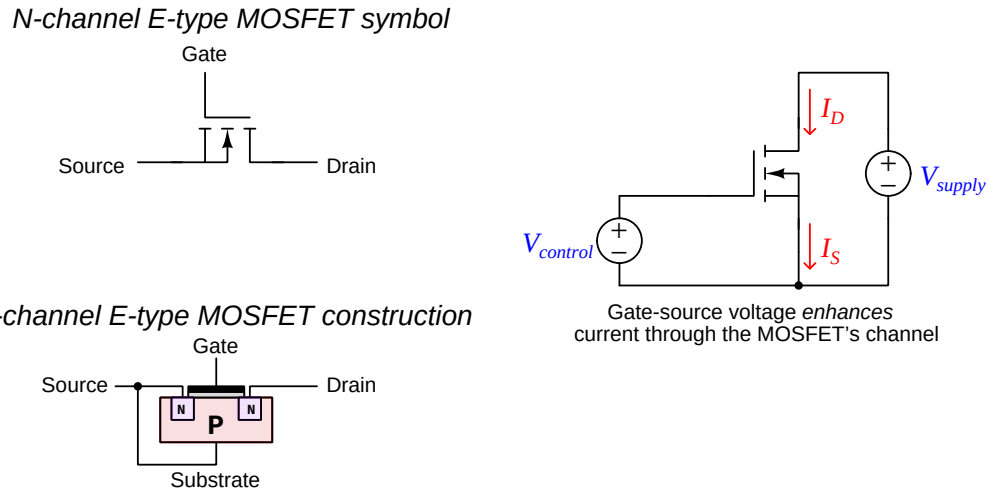




Whether an applied gate-source voltage tends to turn on or turn off a MOSFET depends on its polarity *and* the type of channel doping. For example, an N-channel MOSFET (where electrons are majority-type charge carrier within the channel) will be turned on more if the gate is made positive with respect to the substrate, because this attracts more electrons into and repels holes from the N-doped channel.

[10] An alternative name for this transistor is the *Insulated Gate Field-Effect Transistor*, or *IGFET*.

Since the E-type (enhancement) MOSFET is a normally-off[11] device, it is pointless to apply a *restricting* V_{GS} bias between the gate and source:



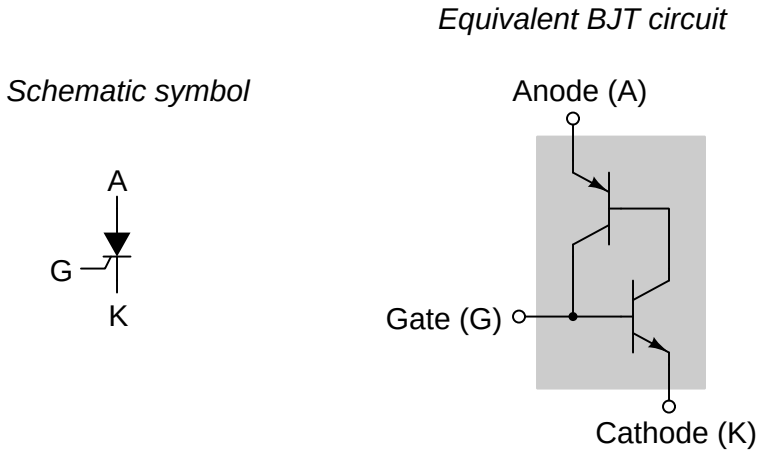
MOSFET gates, being insulated from the channel by a layer of silicon dioxide, cannot be damaged from forward-biasing as is the case with JFETs. However, the extreme thinness of that insulating layer makes it susceptible to damage by static electric voltages. It is helpful to mentally model that insulating layer as the dielectric of a capacitor, and just as capacitors have maximum voltage ratings (beyond which the potential energy difference of the charge carriers is enough to puncture that dielectric layer) so does the gate of a MOSFET. For this reason, precautions must be taken when handling MOSFETs so as to avoid damage by electrostatic charges. Workbenches connected to Earth ground and covered with conductive material present an equipotential surface for the transistor and assembled circuit to rest, while anti-static mats and grounded wristbands maintain the technician at the same (earth ground) potential as the workbench.

JFETs and MOSFETs alike are considered *high input impedance* transistors because their controlling (gate) terminals draw negligible current when operating (i.e. the impedance measured from the controlling terminal is extremely large). Unlike a BJT which requires some base current to inject charge carriers into the base layer to form a conductive path from emitter to collector, FETs are controlled solely by the *electric field* applied across the transistor's channel. This is a generally useful property for any amplifying device, as a high input impedance means little to no current will be demanded of the controlling signal source. Another way of stating this fact is to say field-effect transistors inherently exhibit an extremely large *current gain* value (the ratio between the amount of controlled current versus the amount of controlling current).

[11] Note how the symbol for an E-type MOSFET differs from that of a D-type MOSFET: the channel appears as a broken line rather than a solid line. This is how the "normally-off" nature of the E-type MOSFET is represented, as opposed to the "normally-on" nature of the D-type MOSFET.

A unique category of semiconductor device is the *thyristor*. A "thyristor[12]" is any semiconductor device exhibiting *hysteresis*[13], which is a difference between threshold conditions necessary to turn on versus turn off. Once triggered into its conductive state, a thyristor tends to remain in that "on" state until some other device interrupts the flow of current through it – simply removing the triggering influence is not enough to turn a thyristor off.

One of the most common types of thyristors is the *Silicon-Controlled Rectifier*, or *SCR*. This is a four-layer (PNPN) semiconductor device, which may be modeled as a pair of BJTs:



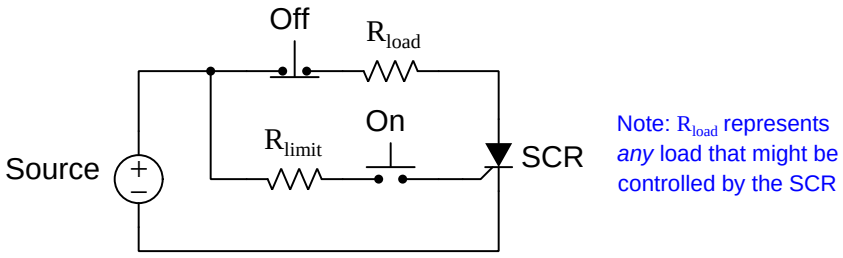
Recall that a BJT is a normally-off device, turned "on" by a through its base terminal. If we connect a voltage source between the anode (A) and cathode (K) terminals of an SCR, leaving the gate (G) terminal disconnected, it would appear the SCR will remain in the "off" state, as neither of its transistor can conduct until given a base current, and neither transistor will experience a base current until the other one conducts. However, all we need to get turn on the SCR is to inject a small "triggering" current through the gate-cathode PN junction, which will turn on the lower transistor, which in turn turns on the upper transistor, which in turn maintains the lower transistor in the "on" state and *latches* the SCR.

[12]Like the transistor, the thyristor was preceded by a tube-style electronic device called a *thyratron* which used a low-pressure gas inside of a glass tube to serve as the conductive medium for an electric current. A triggering electrode inside the thyratron tube served to ionize this gas at the command of an external control signal, but as soon as a large enough controlled current passed through the tube the gas would be maintained in its ionized state to allow conductivity even without the triggering signal. Thus, the thyratron exhibited the same hysteretic behavior as a thyristor: once triggered "on" by an external signal it would remain in that "on" state until something else interrupted the controlled current.

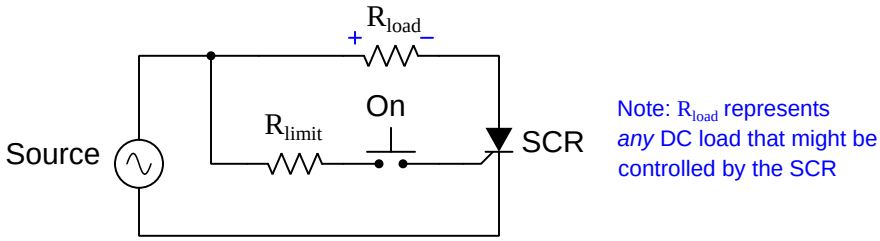
[13]Hysteresis is a general term applicable to many different physical phenomenon, some of them unrelated to electronics. A more general definition of hysteresis is *a lack of response from a device or system following a reversal of stimulus*. Ferromagnetism (the magnetization of iron-bearing materials) is a typical example of hysteresis because ferromagnetic materials tend to retain some magnetic force when removed from the influence of another magnet: the material becomes magnetized when exposed to an external magnetic field, and then retains some of that magnetization after the external field disappears, and must be exposed to a *reverse* field to be re-set back to zero magnetism. If hysteresis were absent, the material would return exactly to its previously unmagnetized state upon removal of the external field. Mechanisms with "loose" couplings tend to exhibit hysteresis, in that the motivating action must over-correct in order to bring the mechanism back to its original state. Anyone who has ever steered a vehicle with a "loose" steering mechanism knows this effect.

We see how the two equivalent BJTs inside of an SCR work together to keep each other in the "on" state. This mutually-controlling action gives the SCR its hysteretic behavior: once triggered "on" by a signal at the gate terminal, the SCR will remain in the "on" state even when the source of that triggering signal is removed. In fact, it is nearly impossible to command the SCR to turn back "off" by *any* signal applied to the gate terminal!

A design challenge when using SCRs is how to turn them off, given their tendency to latch in the "on" state following the slightest provocation at the gate terminal. Here is a simple DC-powered SCR circuit, using one normally-open pushbutton switch to trigger the SCR and one normally-closed pushbutton switch to interrupt current and thereby turn the SCR off:



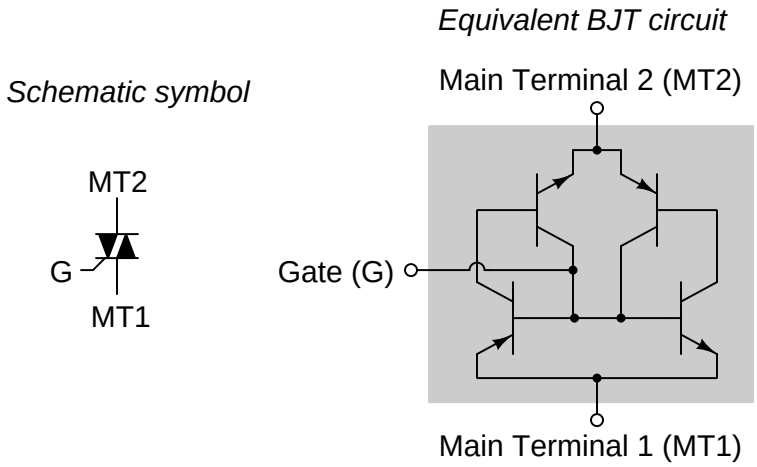
When the circuit's energy source is AC rather than DC, the problem of how to turn off the SCR meets a convenient solution: since the AC source's output repeatedly reverses polarity, and this fact guarantees two moments in time for every complete cycle when current decreases to zero, the SCR will naturally turn off if untriggered at any of these moments. Therefore, an AC version of the previous circuit has no need for an "Off" switch to unlatch the SCR:



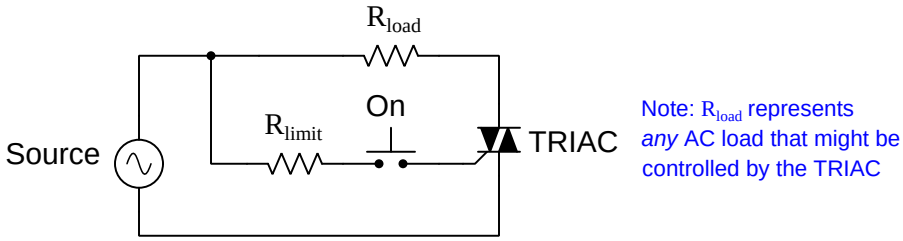
In this circuit, the SCR turns on whenever the "On" switch is pressed and the source polarity is correct (+ on top and − on bottom), and then turns off whenever the source crosses zero and reverses polarity. The SCR literally serves as a *controlled rectifier*: sending half-wave rectified power to the load whenever the switch is pressed, and turning off no more than one half-cycle after the switch is released.

The mutual-controlling action of an SCR's two transistors leads to another interesting consequence besides hysteresis: both of its transistors become *fully saturated* in their "on" states once triggered. In other words, once an SCR is turned on it fully commits to being in that "on" state with minimal voltage drop between the anode and cathode terminals. This effect minimizes the amount of energy lost to heat by charge carriers passing through the SCR, making SCRs very efficient switching devices for high-current power circuits.

A bidirectional version of the SCR exists, called the *TRIAC*. A TRIAC exhibits all the characteristics of an SCR, except with the ability to be triggered and control current in *both* directions:

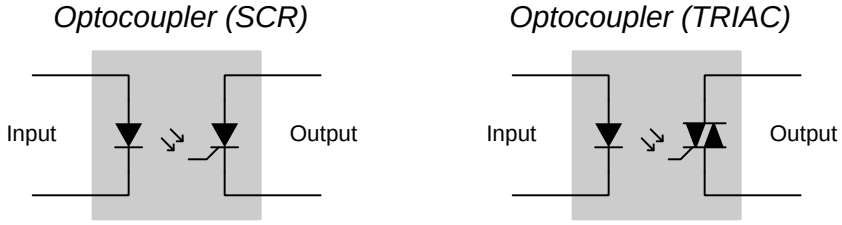


The bidirectional capability of a TRIAC allows it to be used as a full-wave AC switch (i.e. no rectification) to control power to an AC load:



In this circuit, the TRIAC turns on whenever the "On" switch is pressed and then turns off whenever the source crosses zero. Like the SCR, the TRIAC latches on even when triggered by a brief pulse at the gate terminal, and naturally turns off whenever the AC source reverses polarity. Like the SCR, the TRIAC requires very little triggering current to latch in its "on" state, and once turned on it saturates to the fully-conductive state, dropping minimal voltage between its two current-carrying terminals. These qualities make TRIACs very attractive switching devices for high-current AC power circuit applications.

Just as bipolar junction transistors (BJTs) may be optically triggered rather than electrically triggered, thyristors such as SCRs and TRIACs may also be manufactured with optical triggering: a brief pulse of light shining on the lower transistor base-emitter junction liberates electron-hole pairs which then conduct current and cause the device to latch in the "on" state. *Optocoupler* devices made of LEDs and thyristors serve as alternatives to optocoupled transistor switches:

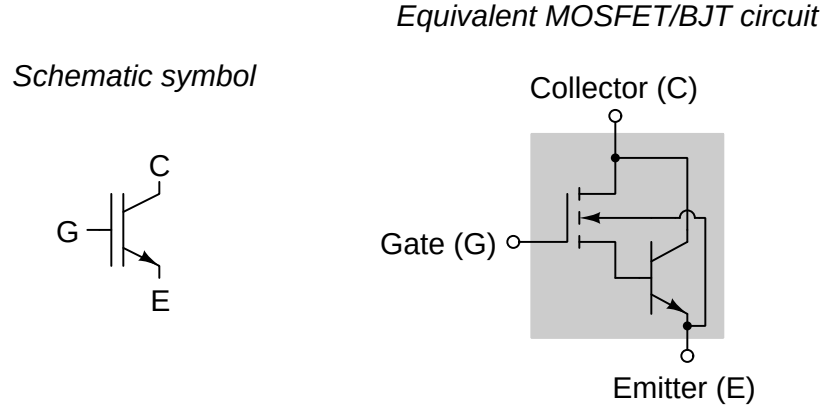


Opto-SCRs and opto-TRIACs, just like their gate-triggered equivalents, require some external means to halt current through them. Light shone on an opto-thyristor will turn it on, but it cannot turn it off.

Another variation on the themes of transistors and thyristors is to combine transistor types into hybrid devices in order to leverage the distinct advantages of each. Consider the following characteristics of bipolar and field-effect power transistors:

- BJTs are able to handle greater controlled currents than MOSFETS
- MOSFETs require far less control signal current than BJTs

Bipolar junction transistors may be constructed with current ratings exceeding those equivalently-sized MOSFETs, yet MOSFETs do not need nearly as much current to control them than bipolar transistors. A high-performance hybrid may be made of these two technologies, by situating a MOSFET at the gate of a BJT so that the easily-controlled MOSFET serves to control the heavier-duty BJT. The result is called an *Insulated Gate Bipolar Transistor*, or *IGBT*:



The MOSFET portion of an IGBT is turned on by applying the appropriate-polarity voltage between the gate and emitter terminals (thus creating an electric field between the gate terminal

and the MOSFET's substrate, enhancing its channel to make it conductive). When the MOSFET turns on, it conducts current through its channel to activate the base-emitter PN junction of the bipolar portion of the IGBT, thus turning on the transistor and allowing current to flow between the collector and emitter terminals.

IGBTs are found in high-power circuits, where a relatively weak controlling signal must exert control over a much larger controlled current. Solid-state electric motor controls are one common application of IGBTs.

Similar hybrids exist combining MOSFETs with thyristors, to create devices called *MOS-controlled thyristors*, or *MCTs*.

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.

[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.

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 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 full 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.

Conductors versus Insulators

Semiconductor

Resistance

Doping

N-type semiconductor

P-type semiconductor

Depletion region

Diode

Forward-biasing

Reverse-biasing

Shockley diode equation

Electrical source

Electrical load

Breakdown

Zener diode

Light-emitting diode

Photodiode

Photovoltaic cell

Maximum Power Transfer Theorem

Transistor

Amplification

Bipolar junction transistor (BJT)

Minority carrier device

Beta ratio

Optotransistor

Optocoupler

Field effect transistor (FET)

Majority carrier device

Junction field effect transistor (JFET)

Metal Oxide Semiconductor field effect transistor (MOSFET)

Depletion

Enhancement

Hysteresis

Thyristor

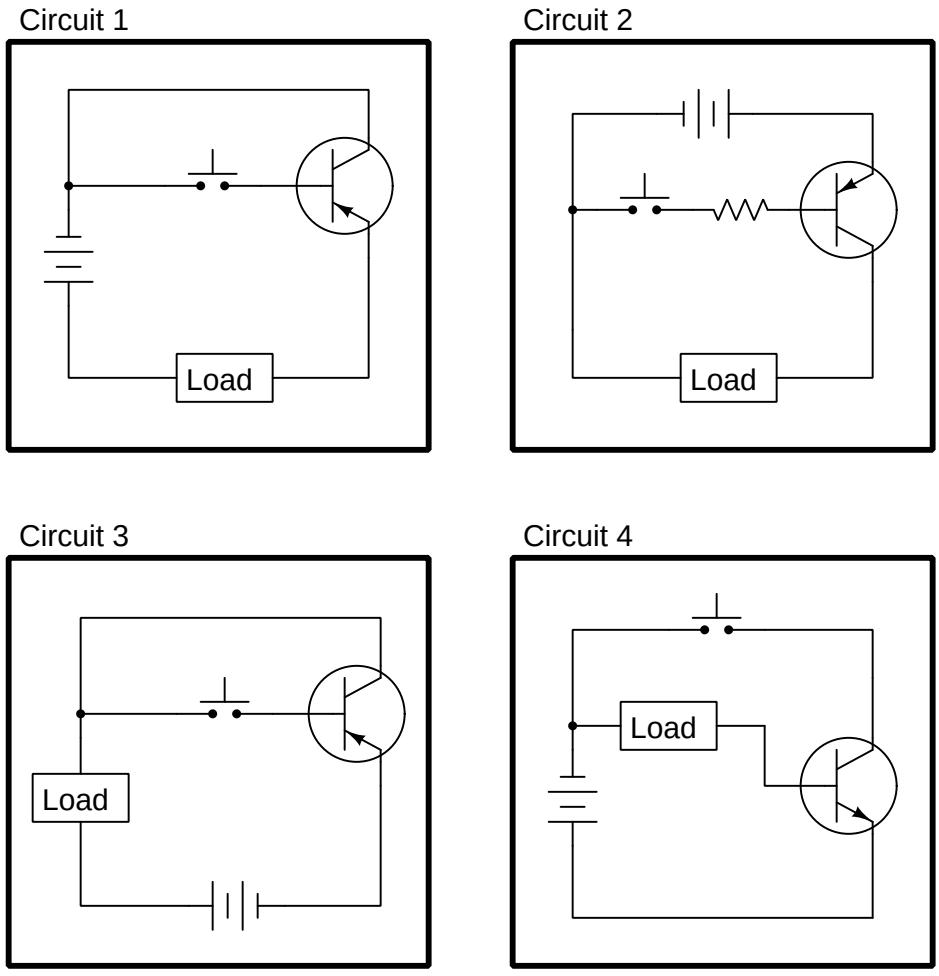
Silicon-controlled rectifier (SCR)

TRIAC

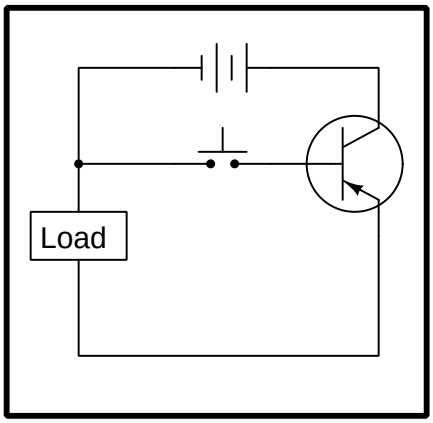
Insulated-gate bipolar transistor

4.1.3 BJT switching circuit configurations

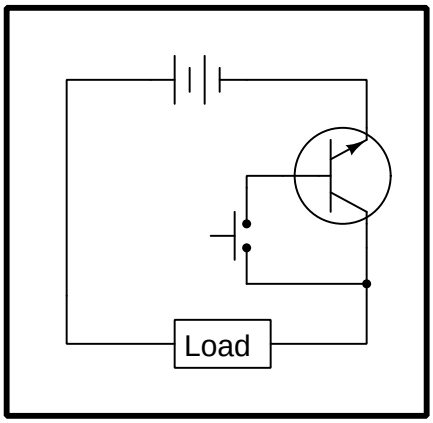
Some of the following bipolar junction transistor switch circuits are properly configured, and some are not. Identify which of these circuits will function properly (i.e. turn on the load when the switch closes) and which of these circuits are mis-wired:



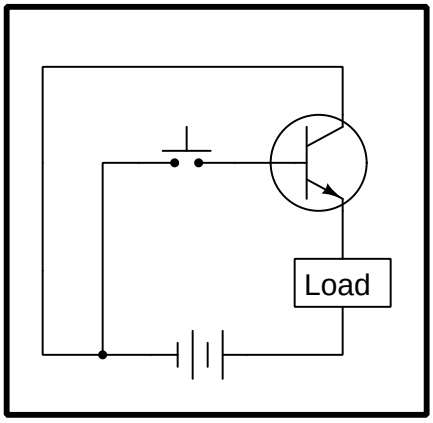
Circuit 5



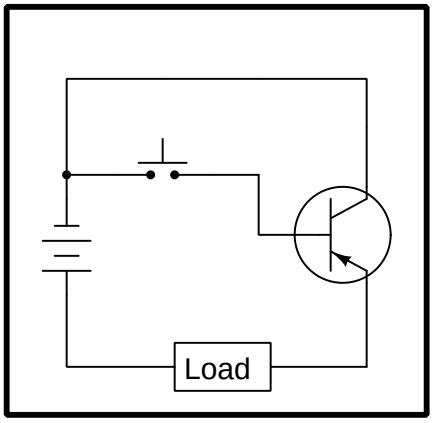
Circuit 6



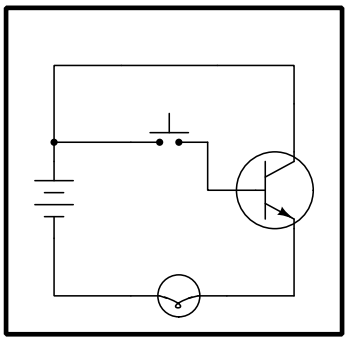
Circuit 7



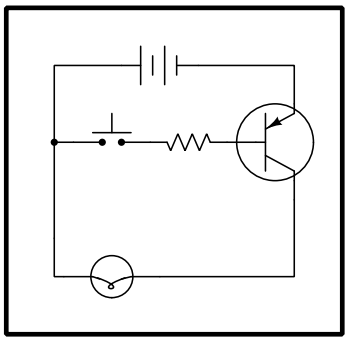
Circuit 8



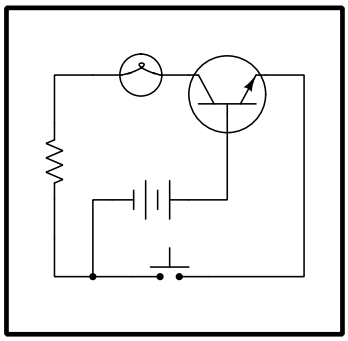
Circuit 9



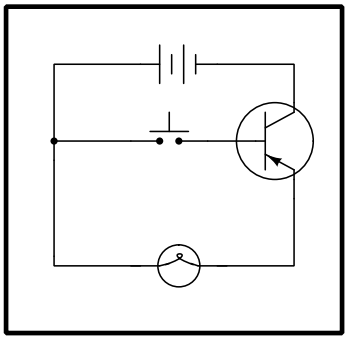
Circuit 10



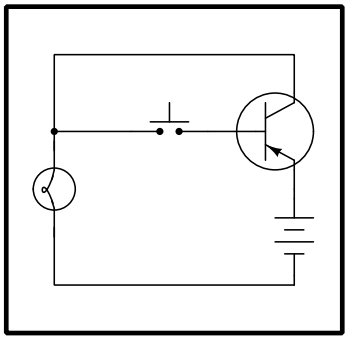
Circuit 11



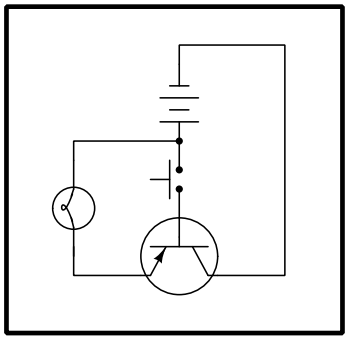
Circuit 12



Circuit 13



Circuit 14



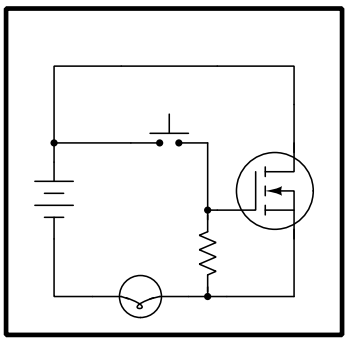
Challenges

- Identify all criteria necessary to activate a BJT.

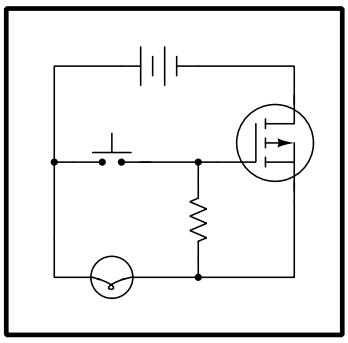
4.1.4 MOSFET switching circuit configurations

Some of the following metal-oxide-semiconductor field-effect transistor switch circuits are properly configured, and some are not. Identify which of these circuits will function properly (i.e. turn on the load when the switch closes) and which of these circuits are mis-wired:

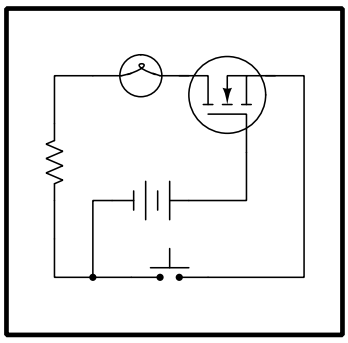
Circuit 1



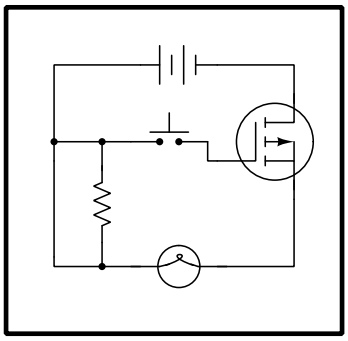
Circuit 2



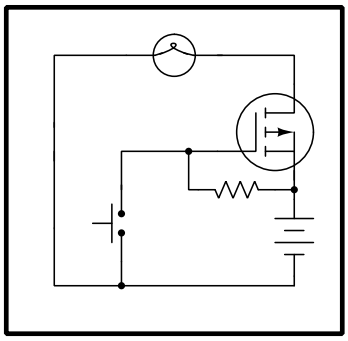
Circuit 3



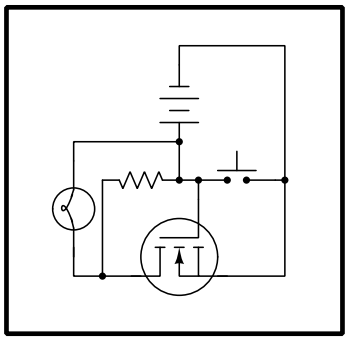
Circuit 4



Circuit 5



Circuit 6



Challenges

- Identify all criteria necessary to enhance a MOSFET.

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 First quantitative problem

This is a description of the problem!

Challenges

- ???.
- ???.
- ???.

4.2.4 Second quantitative problem

This is a description of the problem!

Challenges

- ???.
- ???.
- ???.

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.

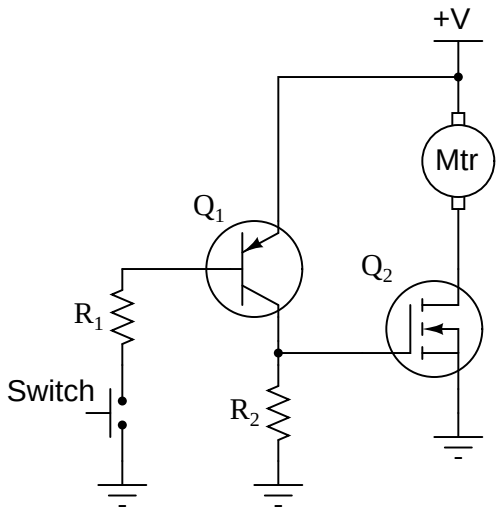
As always, your goal is to fully *explain* your analysis of each problem. Simply obtaining a correct answer is not good enough – you must also demonstrate sound reasoning in order to successfully complete the assignment. Your instructor's responsibility is to probe and challenge your understanding of 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 Faults in a PNP-controlled MOSFET switching circuit

Predict how this circuit will be affected as a result of the following faults. Consider each fault independently (i.e. one at a time, no coincidental faults):



- Transistor Q_1 fails shorted (collector-to-emitter):
- Transistor Q_2 fails open (drain-to-source):
- Resistor R_1 fails open:
- Resistor R_2 fails open:

Challenges

- Annotate this circuit with all voltage polarities and current directions for a condition where the pushbutton switch is released.
- Annotate this circuit with all voltage polarities and current directions for a condition where the pushbutton switch is pressed.

Chapter 5

Projects and Experiments

The following project and experiment descriptions outline things you can build to help you understand circuits. With any real-world project or experiment there exists the potential for physical harm. *Electricity can be very dangerous in certain circumstances, and you should follow proper safety precautions at all times!*

5.1 Recommended practices

This section outlines some recommended practices for all circuits you design and construct.

5.1.1 Safety first!

Electricity, when passed through the human body, causes uncomfortable sensations and in large enough measures[1] will cause muscles to involuntarily contract. The overriding of your nervous system by the passage of electrical current through your body is particularly dangerous in regard to your heart, which is a vital muscle. Very large amounts of current can produce serious internal burns in addition to all the other effects.

Cardio-pulmonary resuscitation (CPR) is the standard first-aid for any victim of electrical shock. This is a very good skill to acquire if you intend to work with others on dangerous electrical circuits. You should never perform tests or work on such circuits unless someone else is present who is proficient in CPR.

As a general rule, any voltage in excess of 30 Volts poses a definitive electric shock hazard, because beyond this level human skin does not have enough resistance to safely limit current through the body. "Live" work of any kind with circuits over 30 volts should be avoided, and if unavoidable should only be done using electrically insulated tools and other protective equipment (e.g. insulating shoes and gloves). If you are unsure of the hazards, or feel unsafe at any time, stop all work and distance yourself from the circuit!

A policy I strongly recommend for students learning about electricity is to *never come into electrical contact*[2] *with an energized conductor, no matter what the circuit's voltage*[3] *level!* Enforcing this policy may seem ridiculous when the circuit in question is powered by a single battery smaller than the palm of your hand, but it is precisely this instilled habit which will save a person from bodily harm when working with more dangerous circuits. Experience has taught me that students who learn early on to be careless with safe circuits have a tendency to be careless later with dangerous circuits!

In addition to the electrical hazards of shock and burns, the construction of projects and running of experiments often poses other hazards such as working with hand and power tools, potential

[1] Professor Charles Dalziel published a research paper in 1961 called "The Deleterious Effects of Electric Shock" detailing the results of electric shock experiments with both human and animal subjects. The threshold of perception for human subjects holding a conductor in their hand was in the range of 1 milliampere of current (less than this for alternating current, and generally less for female subjects than for male). Loss of muscular control was exhibited by half of Dalziel's subjects at less than 10 milliamperes alternating current. Extreme pain, difficulty breathing, and loss of all muscular control occurred for over 99% of his subjects at direct currents less than 100 milliamperes and alternating currents less than 30 milliamperes. In summary, it doesn't require much electric current to induce painful and even life-threatening effects in the human body! Your first and best protection against electric shock is maintaining an insulating barrier between your body and the circuit in question, such that current from that circuit will be unable to flow through your body.

[2] By "electrical contact" I mean either directly touching an energized conductor with any part of your body, or indirectly touching it through a conductive tool. The only physical contact you should ever make with an energized conductor is via an electrically insulated tool, for example a screwdriver with an electrically insulated handle, or an insulated test probe for some instrument.

[3] Another reason for consistently enforcing this policy, even on low-voltage circuits, is due to the dangers that even some low-voltage circuits harbor. A single 12 Volt automobile battery, for example, can cause a surprising amount of damage if short-circuited simply due to the high current levels (i.e. very low internal resistance) it is capable of, even though the voltage level is too low to cause a shock through the skin. Mechanics wearing metal rings, for example, are at risk from severe burns if their rings happen to short-circuit such a battery! Furthermore, even when working on circuits that are simply too low-power (low voltage and low current) to cause any bodily harm, touching them while energized can pose a threat to the circuit components themselves. In summary, it's generally wise (and *always* a good habit to build) to "power down" *any* circuit before making contact between it and your body.

contact with high temperatures, potential chemical exposure, etc. You should never proceed with a project or experiment if you are unaware of proper tool use or lack basic protective measures (e.g. personal protective equipment such as safety glasses) against such hazards.

Some other safety-related practices should be followed as well:

- All power conductors extending outward from the project must be *firmly* strain-relieved (e.g. "cord grips" used on line power cords), so that an accidental tug or drop will not compromise circuit integrity.
- All electrical connections must be sound and appropriately made (e.g. soldered wire joints rather than twisted-and-taped; terminal blocks rather than solderless breadboards for high-current or high-voltage circuits). Use "touch-safe" terminal connections with recessed metal parts to minimize risk of accidental contact.
- Always provide overcurrent protection in any circuit you build. *Always.* This may be in the form of a fuse, a circuit breaker, and/or an electronically current-limited power supply.
- Always ensure circuit conductors are rated for more current than the overcurrent protection limit. *Always.* A fuse does no good if the wire or printed circuit board trace will "blow" before it does!
- Always bond metal enclosures to Earth ground for any line-powered circuit. *Always.* Ensuring an equipotential state between the enclosure and Earth by making the enclosure electrically common with Earth ground ensures no electric shock can occur simply by one's body bridging between the Earth and the enclosure.
- Avoid building a high-energy circuit when a low-energy circuit will suffice. For example, I always recommend beginning students power their first DC resistor circuits using small batteries rather than with line-powered DC power supplies. The intrinsic energy limitations of a dry-cell battery make accidents highly unlikely.
- Use line power receptacles that are GFCI (Ground Fault Current Interrupting) to help avoid electric shock from making accidental contact with a "hot" line conductor.
- Always wear eye protection when working with tools or live systems having the potential to eject material into the air. Examples of such activities include soldering, drilling, grinding, cutting, wire stripping, working on or near energized circuits, etc.
- Always use a step-stool or stepladder to reach high places. Never stand on something not designed to support a human load.
- When in doubt, *ask an expert.* If anything even seems remotely unsafe to you, do not proceed without consulting a trusted person fully knowledgeable in electrical safety.

5.1.2 Other helpful tips

Experience has shown the following practices to be very helpful, especially when students make their own component selections, to ensure the circuits will be well-behaved:

- Avoid resistor values less than 1 kΩ or greater than 100 kΩ, unless such values are definitely necessary[4]. Resistances below 1 kΩ may draw excessive current if directly connected to a voltage source of significant magnitude, and may also complicate the task of accurately measuring current since any ammeter's non-zero resistance inserted in series with a low-value circuit resistor will significantly alter the total resistance and thereby skew the measurement. Resistances above 100 kΩ may complicate the task of measuring voltage since any voltmeter's finite resistance connected in parallel with a high-value circuit resistor will significantly alter the total resistance and thereby skew the measurement. Similarly, AC circuit impedance values should be between 1 kΩ and 100 kΩ, and for all the same reasons.

- Ensure all electrical connections are low-resistance and physically rugged. For this reason, one should avoid *compression splices* (e.g. "butt" connectors), solderless breadboards[5], and wires that are simply twisted together.

- Build your circuit with **testing** in mind. For example, provide convenient connection points for test equipment (e.g. multimeters, oscilloscopes, signal generators, logic probes).

- Design permanent projects with **maintenance** in mind. The more convenient you make maintenance tasks, the more likely they will get done.

- **Always document and save your work**. Circuits lacking schematic diagrams are more difficult to troubleshoot than documented circuits. Similarly, circuit construction is simpler when a schematic diagram precedes construction. Experimental results are easier to interpret when comprehensively recorded. Consider modern videorecording technology for this purpose where appropriate.

- **Record your steps** when troubleshooting. **Talk to yourself** when solving problems. These simple steps clarify thought and simplify identification of errors.

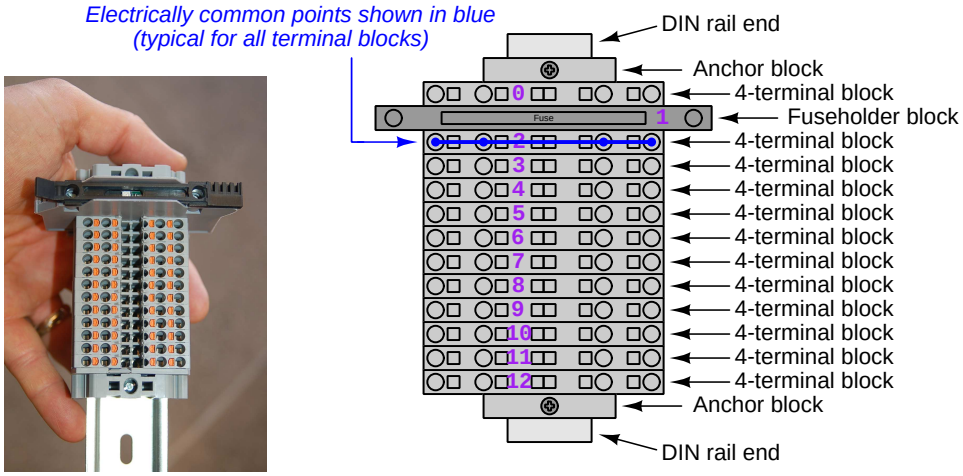
[4] An example of a necessary resistor value much less than 1 kΩ is a *shunt resistor* used to produce a small voltage drop for the purpose of sensing current in a circuit. Such shunt resistors must be low-value in order not to impose an undue load on the rest of the circuit. An example of a necessary resistor value much greater than 100 kΩ is an electrostatic *drain resistor* used to dissipate stored electric charges from body capacitance for the sake of preventing damage to sensitive semiconductor components, while also preventing a path for current that could be dangerous to the person (i.e. shock).

[5] Admittedly, solderless breadboards are very useful for constructing complex electronic circuits with many components, especially DIP-style integrated circuits (ICs), but they tend to give trouble with connection integrity after frequent use. An alternative for projects using low counts of ICs is to solder IC sockets into prototype printed circuit boards (PCBs) and run wires from the soldered pins of the IC sockets to terminal blocks where reliable temporary connections may be made.

5.1.3 Terminal blocks for circuit construction

Terminal blocks are the standard means for making electric circuit connections in industrial systems, and are also quite useful for student-built experiments, providing a robust alternative to solderless breadboards[6]. Terminal blocks provide highly reliable connections rated for significant voltage and current magnitudes, and they force the builder to think carefully about component placement which is an important mental practice. Terminal blocks designed to mount on standard 35 mm DIN rail[7] come in many types and capabilities.

Every student of electricity should build their own terminal block array consisting of several terminal blocks where each block has at least 4 connection points all electrically common to each other[8] and at least one terminal block that is a fuse holder for overcurrent protection. A pair of anchoring blocks hold all terminal blocks securely on the DIN rail, preventing them from sliding off the rail. Each of the terminals should bear a unique number starting from 0. An example is shown in the following photograph and illustration:



Screwless terminal blocks (using internal spring clips to clamp wire and component lead ends) are preferred over screw-based terminal blocks, as they reduce assembly and disassembly time, and also minimize repetitive wrist stress from twisting screwdrivers. Some screwless terminal blocks require the use of a special tool to release the spring clip, while others such as the Phoenix Contact model 3211019 blocks shown in the photograph provide buttons[9] for this task which may be pressed using the tip of any suitable tool such as a small screwdriver or multimeter test probe tip.

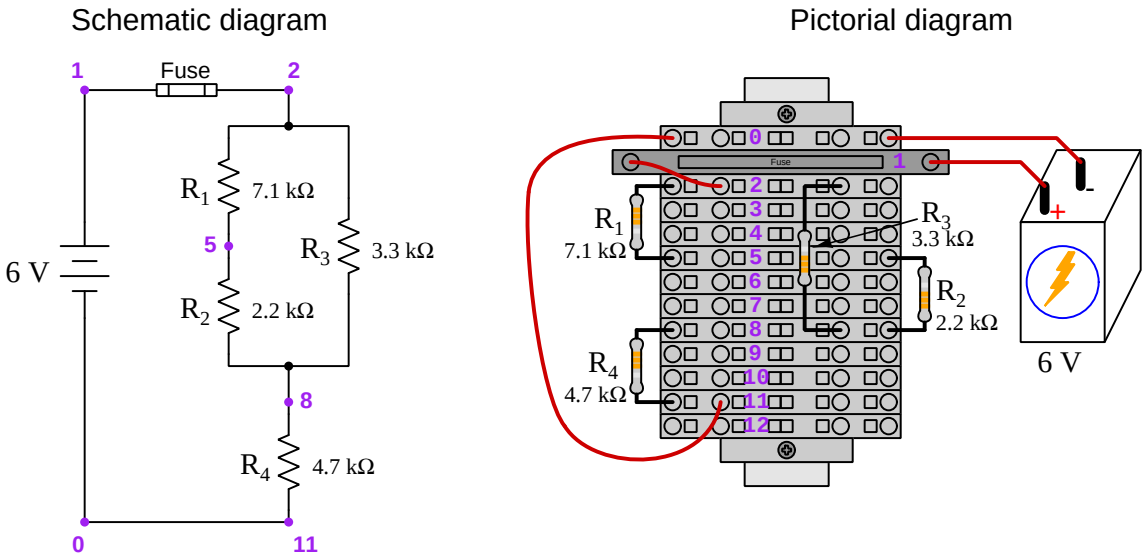
[6] Solderless breadboard are preferable for complicated electronic circuits with multiple integrated "chip" components, but for simpler circuits I find terminal blocks much more practical. An alternative to solderless breadboards for "chip" circuits is to solder chip sockets onto a PCB and then use wires to connect the socket pins to terminal blocks. This also accommodates *surface-mount* components, which solderless breadboards do not.

[7] DIN rail is a metal rail designed to serve as a mounting point for a wide range of electrical and electronic devices such as terminal blocks, fuses, circuit breakers, relay sockets, power supplies, data acquisition hardware, etc.

[8] Sometimes referred to as *equipotential*, *same-potential*, or *potential distribution* terminal blocks.

[9] The small orange-colored squares seen in the above photograph are buttons for this purpose, and may be actuated by pressing with any tool of suitable size.

The following example shows how such a terminal block array might be used to construct a series-parallel resistor circuit consisting of four resistors and a battery:



Numbering on the terminal blocks provides a very natural translation to SPICE[10] netlists, where component connections are identified by terminal number:

```
* Series-parallel resistor circuit
v1 1 0 dc 6
r1 2 5 7100
r2 5 8 2200
r3 2 8 3300
r4 8 11 4700
rfuse 1 2 0.01
rwire 0 11 0.01
.op
.end
```

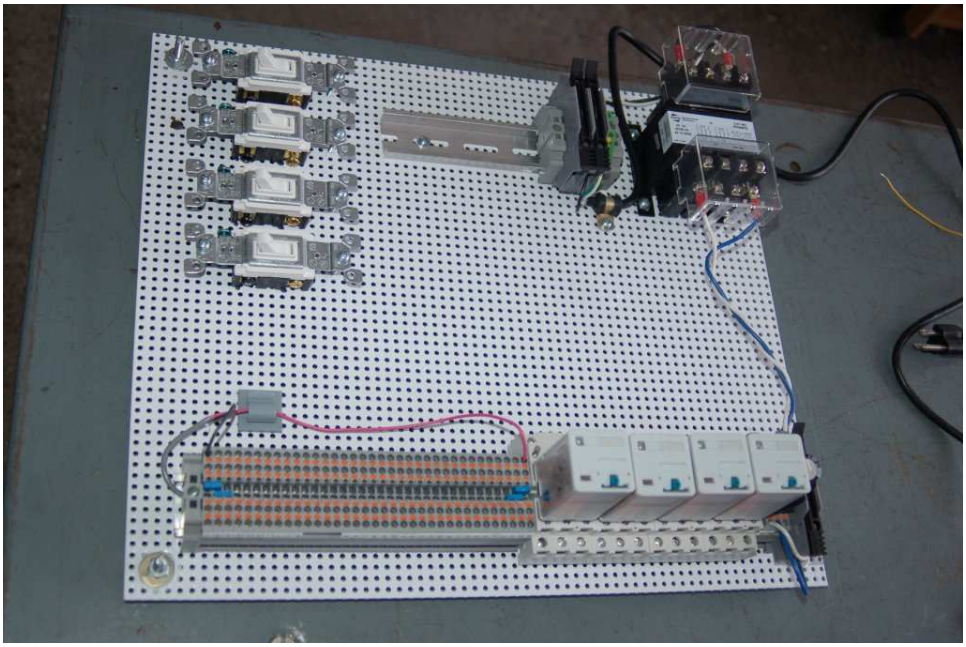
Note the use of two resistances `rfuse` and `rwire` to describe fuse resistance between terminals 1 and 2 and "jumper" wire resistance between terminals 0 and 11, respectively. Being resistances, SPICE requires a resistance value for each, and here we see they have both been set to an arbitrarily low value of 0.01 Ohm appropriate to their real-life resistance values.

Listing all components and wires along with their numbered terminals happens to be a useful documentation method for any circuit built on terminal blocks, independent of SPICE. Such a

[10] SPICE is computer software designed to analyze electrical and electronic circuits. Circuits are described for the computer in the form of *netlists* which are text files listing each component type, connection node numbers, and component values.

"wiring sequence" may be thought of as a *non-graphical description* of an electric circuit, and is exceptionally easy to follow.

An example of a more elaborate terminal block array is shown in the following photograph, with terminal blocks and "ice-cube" style electromechanical relays mounted to DIN rail, which is turn mounted to a perforated subpanel[11]. This "terminal block board" hosts an array of thirty five undedicated terminal block sections, four SPDT toggle switches, four DPDT "ice-cube" relays, a step-down control power transformer, bridge rectifier and filtering capacitor, and several fuses for overcurrent protection:



Four plastic-bottomed "feet" support the subpanel above the benchtop surface, and an unused section of DIN rail stands ready to accept other components. Safety features include electrical bonding of the AC line power cord's ground to the metal subpanel (and all metal DIN rails), mechanical strain relief for the power cord to isolate any cord tension from wire connections, clear plastic finger guards covering the transformer's screw terminals, as well as fused overcurrent protection for the 120 Volt AC line power and the transformer's 12 Volt AC output. The perforated holes happen to be on $\frac{1}{4}$ inch centers with a diameter suitable for tapping with 6-32 machine screw threads, their presence making it very easy to attach other sections of DIN rail, printed circuit boards, or specialized electrical components directly to the grounded metal subpanel. Such a "terminal block board" is an inexpensive[12] yet highly flexible means to construct physically robust circuits using industrial wiring practices.

[11] An electrical *subpanel* is a thin metal plate intended for mounting inside an electrical enclosure. Components are attached to the subpanel, and the subpanel in turn bolts inside the enclosure. Subpanels allow circuit construction outside the confines of the enclosure, which speeds assembly. In this particular usage there is no enclosure, as the subpanel is intended to be used as an open platform for the convenient construction of circuits on a benchtop by students. In essence, this is a modern version of the traditional *breadboard* which was literally a wooden board such as might be used for cutting loaves of bread, but which early electrical and electronic hobbyists used as platforms for the construction of circuits.

[12] At the time of this writing (2019) the cost to build this board is approximately $250 US dollars.

5.1.4 Conducting experiments

An *experiment* is an exploratory act, a test performed for the purpose of assessing some proposition or principle. Experiments are the foundation of the *scientific method*, a process by which careful observation helps guard against errors of speculation. All good experiments begin with an *hypothesis*, defined by the American Heritage Dictionary of the English Language as:

> An assertion subject to verification or proof, as (a) A proposition stated as a basis for argument or reasoning. (b) A premise from which a conclusion is drawn. (c) A conjecture that accounts, within a theory or ideational framework, for a set of facts and that can be used as a basis for further investigation.

Stated plainly, an hypothesis is an *educated guess* about cause and effect. The correctness of this initial guess matters little, because any well-designed experiment will reveal the truth of the matter. In fact, *incorrect* hypotheses are often the most valuable because the experiments they engender lead us to surprising discoveries. One of the beautiful aspects of science is that it is more focused on the process of *learning* than about the status of *being correct*[13]. In order for an hypothesis to be valid, it must be testable[14], which means it must be a claim possible to refute given the right data. Hypotheses impossible to critique are useless.

Once an hypothesis has been formulated, an experiment must be designed to test that hypothesis. A well-designed experiment requires careful regulation of all relevant variables, both for personal safety and for prompting the hypothesized results. If the effects of one particular variable are to be tested, the experiment must be run multiple times with different values of (only) that particular variable. The experiment set up with the "baseline" variable set is called the *control*, while the experiment set up with different value(s) is called the *test* or *experimental*.

For some hypotheses a viable alternative to a physical experiment is a *computer-simulated experiment* or even a *thought experiment*. Simulations performed on a computer test the hypothesis against the physical laws encoded within the computer simulation software, and are particularly useful for students learning new principles for which simulation software is readily available[15].

[13] Science is more about clarifying our view of the universe through a systematic process of error detection than it is about proving oneself to be right. Some scient*ists* may happen to have large egos – and this may have more to do with the ways in which large-scale scientific research is *funded* than anything else – but scientific *method* itself is devoid of ego, and if embraced as a practical philosophy is quite an effective stimulant for humility. Within the education system, scientific method is particularly valuable for helping students break free of the crippling fear of *being wrong*. So much emphasis is placed in formal education on assessing correct retention of facts that many students are fearful of saying or doing anything that might be perceived as a mistake, and of course making mistakes (i.e. having one's hypotheses disproven by experiment) is an indispensable tool for learning. Introducing science in the classroom – *real* science characterized by individuals forming actual hypotheses and testing those hypotheses by experiment – helps students become self-directed learners.

[14] This is the principle of *falsifiability*: that a scientific statement has value only insofar as it is liable to disproof given the requisite experimental evidence. Any claim that is unfalsifiable – that is, a claim which can *never* be disproven by any evidence whatsoever – could be completely wrong and we could never know it.

[15] A very pertinent example of this is learning how to analyze electric circuits using simulation software such as SPICE. A typical experimental cycle would proceed as follows: (1) Find or invent a circuit to analyze; (2) Apply your analytical knowledge to that circuit, predicting all voltages, currents, powers, etc. relevant to the concepts you are striving to master; (3) Run a simulation on that circuit, collecting "data" from the computer when complete; (4) Evaluate whether or not your hypotheses (i.e. predicted voltages, currents, etc.) agree with the computer-generated results; (5) If so, your analyses are (provisionally) correct – if not, examine your analyses and the computer simulation again to determine the source of error; (6) Repeat this process as many times as necessary until you achieve mastery.

Thought experiments are useful for detecting inconsistencies within your own understanding of some subject, rather than testing your understanding against physical reality.

Here are some general guidelines for conducting experiments:

- The clearer and more specific the hypothesis, the better. Vague or unfalsifiable hypotheses are useless because they will fit *any* experimental results, and therefore the experiment cannot teach you anything about the hypothesis.
- Collect as much data (i.e. information, measurements, sensory experiences) generated by an experiment as is practical. This includes the time and date of the experiment, too!
- *Never* discard or modify data gathered from an experiment. If you have reason to believe the data is unreliable, write notes to that effect, but never throw away data just because you think it is untrustworthy. It is quite possible that even "bad" data holds useful information, and that someone else may be able to uncover its value even if you do not.
- Prioritize *quantitative* data over *qualitative* data wherever practical. Quantitative data is more specific than qualitative, less prone to subjective interpretation on the part of the experimenter, and amenable to an arsenal of analytical methods (e.g. statistics).
- Guard against your own bias(es) by making your experimental results available to others. This allows other people to scrutinize your experimental design and collected data, for the purpose of detecting and correcting errors you may have missed. Document your experiment such that others may independently replicate it.
- Always be looking for sources of error. No physical measurement is perfect, and so it is impossible to achieve *exact* values for any variable. Quantify the amount of uncertainty (i.e. the "tolerance" of errors) whenever possible, and be sure your hypothesis does not depend on precision better than this!
- Always remember that scientific confirmation is provisional – no number of "successful" experiments will prove an hypothesis true for all time, but a single experiment can disprove it. Put into simpler terms, *truth is elusive but error is within reach.*
- Remember that scientific method is about *learning*, first and foremost. An unfortunate consequence of scientific triumph in modern society is that science is often viewed by non-practitioners as an unerring source of truth, when in fact science is an ongoing process of challenging existing ideas to probe for errors and oversights. This is why it is perfectly acceptable to have a failed hypothesis, and why the only truly failed experiment is one where nothing was learned.

The following is an example of a well-planned and executed experiment, in this case a physical experiment demonstrating Ohm's Law.

```
Planning Time/Date = 09:30 on 12 February 2019

HYPOTHESIS: the current through any resistor should be exactly proportional
to the voltage impressed across it.

PROCEDURE: connect a resistor rated 1 k Ohm and 1/4 Watt to a variable-voltage
DC power supply.  Use an ammeter in series to measure resistor current and
a voltmeter in parallel to measure resistor voltage.

RISKS AND MITIGATION: excessive power dissipation may harm the resistor and/
or pose a burn hazard, while excessive voltage poses an electric shock hazard.
30 Volts is a safe maximum voltage for laboratory practices, and according to
Joule's Law a 1000 Ohm resistor will dissipate 0.25 Watts at 15.81 Volts
(P = V^2 / R), so I will remain below 15 Volts just to be safe.

Experiment Time/Date = 10:15 on 12 February 2019

DATA COLLECTED:
  (Voltage)    (Current)         (Voltage)   (Current)
   0.000 V   =  0.000 mA          8.100    =  7.812 mA
   2.700 V   =  2.603 mA          10.00 V  =  9.643 mA
   5.400 V   =  5.206 mA          14.00 V  =  13.49 mA

Analysis Time/Date = 10:57 on 12 February 2019

ANALYSIS: current definitely increases with voltage, and although I expected
exactly one milliAmpere per Volt the actual current was usually less than
that.  The voltage/current ratios ranged from a low of 1036.87 (at 8.1 Volts)
to a high of 1037.81 (at 14 Volts), but this represents a variance of only
-0.0365% to +0.0541% from the average, indicating a very consistent
proportionality -- results consistent with Ohm's Law.

ERROR SOURCES: one major source of error is the resistor's value itself.  I
did not measure it, but simply assumed color bands of brown-black-red meant
exactly 1000 Ohms.  Based on the data I think the true resistance is closer
to 1037 Ohms.  Another possible explanation is multimeter calibration error.
However, neither explains the small positive and negative variances from the
average.  This might be due to electrical noise, a good test being to repeat
the same experiment to see if the variances are the same or different.  Noise
should generate slightly different results every time.
```

The following is an example of a well-planned and executed *virtual* experiment, in this case demonstrating Ohm's Law using a computer (SPICE) simulation.

```
Planning Time/Date = 12:32 on 14 February 2019

HYPOTHESIS: for any given resistor, the current through that resistor should be
exactly proportional to the voltage impressed across it.

PROCEDURE: write a SPICE netlist with a single DC voltage source and single
1000 Ohm resistor, then use NGSPICE version 26 to perform a "sweep" analysis
from 0 Volts to 25 Volts in 5 Volt increments.

    * SPICE circuit
    v1 1 0 dc
    r1 1 0 1000
    .dc v1 0 25 5
    .print dc v(1) i(v1)
    .end

RISKS AND MITIGATION: none.

DATA COLLECTED:
       DC transfer characteristic  Thu Feb 14 13:05:08  2019
    ------------------------------------------------------------
    Index   v-sweep         v(1)            v1#branch
    ------------------------------------------------------------
    0       0.000000e+00    0.000000e+00    0.000000e+00
    1       5.000000e+00    5.000000e+00    -5.00000e-03
    2       1.000000e+01    1.000000e+01    -1.00000e-02
    3       1.500000e+01    1.500000e+01    -1.50000e-02
    4       2.000000e+01    2.000000e+01    -2.00000e-02
    5       2.500000e+01    2.500000e+01    -2.50000e-02

Analysis Time/Date = 13:06 on 14 February 2019

ANALYSIS: perfect agreement between data and hypothesis -- current is precisely
1/1000 of the applied voltage for all values.  Anything other than perfect
agreement would have probably meant my netlist was incorrect.  The negative
current values surprised me, but it seems this is just how SPICE interprets
normal current through a DC voltage source.

ERROR SOURCES: none.
```

As gratuitous as it may seem to perform experiments on a physical law as well-established as Ohm's Law, even the examples listed previously demonstrate opportunity for real learning. In the physical experiment example, the student should identify and explain why their data does not perfectly agree with the hypothesis, and this leads them naturally to consider sources of error. In the computer-simulated experiment, the student is struck by SPICE's convention of denoting regular current through a DC voltage source as being *negative* in sign, and this is also useful knowledge for future simulations. Scientific experiments are most interesting when things *do not* go as planned!

Aside from verifying well-established physical laws, simple experiments are extremely useful as educational tools for a wide range of purposes, including:

- Component familiarization (e.g. *Which terminals of this switch connect to the NO versus NC contacts?*)
- System testing (e.g. *How heavy of a load can my AC-DC power supply source before the semiconductor components reach their thermal limits?*)
- Learning programming languages (e.g. *Let's try to set up an "up" counter function in this PLC!*)

Above all, the priority here is to inculcate the habit of hypothesizing, running experiments, and analyzing the results. This experimental cycle not only serves as an excellent method for self-directed learning, but it also works exceptionally well for troubleshooting faults in complex systems, and for these reasons should be a part of every technician's and every engineer's education.

5.1.5 Constructing projects

Designing, constructing, and testing projects is a very effective means of practical education. Within a formal educational setting, projects are generally chosen (or at least vetted) by an instructor to ensure they may be reasonably completed within the allotted time of a course or program of study, and that they sufficiently challenge the student to learn certain important principles. In a self-directed environment, projects are just as useful as a learning tool but there is some risk of unwittingly choosing a project beyond one's abilities, which can lead to frustration.

Here are some general guidelines for managing projects:

- Define your goal(s) before beginning a project: what do you wish to achieve in building it? What, exactly, should the completed project *do*?
- Analyze your project prior to construction. Document it in appropriate forms (e.g. schematic diagrams), predict its functionality, anticipate all associated risks. In other words, *plan ahead*.
- Set a reasonable budget for your project, and stay within it.
- Identify any deadlines, and set reasonable goals to meet those deadlines.
- Beware of *scope creep*: the tendency to modify the project's goals before it is complete.
- Document your progress! An easy way to do this is to use photography or videography: take photos and/or videos of your project as it progresses. Document failures as well as successes, because both are equally valuable from the perspective of learning.

5.2 Experiment: Driving a heavy load with a bipolar transistor

Conduct an experiment to show how a bipolar transistor (either NPN or PNP) is able to switch electrical power on and off to a heavy load at the command of an electrical signal of far less current, "heavy" being defined as a load whose current is nearly equal to the maximum collector current rating of the transistor. In order to avoid high-voltage transients that could destroy the transistor, the load should be resistive rather than inductive.

EXPERIMENT CHECKLIST:

- Prior to experimentation:
  - ☑ Write an hypothesis (i.e. a detailed description of what you expect will happen) unambiguous enough that it could be disproven given the right data.
  - ☑ Write a procedure to test the hypothesis, complete with adequate controls and documentation (e.g. schematic diagrams, programming code).
  - ☑ Identify any risks (e.g. shock hazard, component damage) and write a mitigation plan based on best practices and component ratings.
- During experimentation:
  - ☑ Safe practices followed at all times (e.g. no contact with energized circuit).
  - ☑ Correct equipment usage according to manufacturer's recommendations.
  - ☑ All data collected, ideally quantitative with full precision (i.e. no rounding).
- After each experimental run:
  - ☑ If the results fail to match the hypothesis, identify the error(s), correct the hypothesis and/or revise the procedure, and re-run the experiment.
  - ☑ Identify any uncontrolled sources of error in the experiment.
- After all experimental re-runs:
  - ☑ Save all data for future reference.
  - ☑ Write an analysis of experimental results and lessons learned.

Challenges

- Science is an *iterative* process, and for this reason is never complete. Following the results of your experiment, what would you propose for your *next* hypothesis and *next* experimental procedure? Hint: if your experiment produced any unexpected results, exploring those unexpected results is often a very good basis for the next experiment!

- ???.
- ???.
- ???.

5.3 Experiment: Driving a heavy load with a power MOSFET transistor

Conduct an experiment to show how a MOSFET type of transistor (either N-channel or P-channel enhancement type) is able to switch electrical power on and off to a heavy load at the command of an electrical signal of far less current, "heavy" being defined as a load whose current is nearly equal to the maximum drain current rating of the transistor. In order to avoid high-voltage transients that could destroy the transistor, the load should be resistive rather than inductive.

EXPERIMENT CHECKLIST:

- Prior to experimentation:
  - ☑ Write an hypothesis (i.e. a detailed description of what you expect will happen) unambiguous enough that it could be disproven given the right data.
  - ☑ Write a procedure to test the hypothesis, complete with adequate controls and documentation (e.g. schematic diagrams, programming code).
  - ☑ Identify any risks (e.g. shock hazard, component damage) and write a mitigation plan based on best practices and component ratings.
- During experimentation:
  - ☑ Safe practices followed at all times (e.g. no contact with energized circuit).
  - ☑ Correct equipment usage according to manufacturer's recommendations.
  - ☑ All data collected, ideally quantitative with full precision (i.e. no rounding).
- After each experimental run:
  - ☑ If the results fail to match the hypothesis, identify the error(s), correct the hypothesis and/or revise the procedure, and re-run the experiment.
  - ☑ Identify any uncontrolled sources of error in the experiment.
- After all experimental re-runs:
  - ☑ Save all data for future reference.
  - ☑ Write an analysis of experimental results and lessons learned.

Challenges

- Science is an *iterative* process, and for this reason is never complete. Following the results of your experiment, what would you propose for your *next* hypothesis and *next* experimental procedure? Hint: if your experiment produced any unexpected results, exploring those unexpected results is often a very good basis for the next experiment!

- ???.
- ???.
- ???.

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

"Understanding Power MOSFETs", Application Note AN-7500, revision A1, Fairchild Semiconductor, 2002.

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.

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.

6 July 2020 – added entry to Foundational Concepts list.

8,10 April 2020 – minor typographical corrections courtesy of Ron Felix.

18 October 2019 – minor typographical and grammatical corrections.

16 June 2019 – minor edits to diagnostic questions, replacing "no multiple faults" with "no coincidental faults".

2 April 2019 – added more Foundational Concepts to the list.

23 March 2019 – added Conceptual Reasoning questions on proper connections to transistors, a Diagnostic Reasoning question on faults within a BJT/MOSFET circuit, and also added thyristors to the Simplified Tutorial.

13 March 2019 – added experiments to use power transistors (both BJT and MOSFET) to switch power to a heavy load.

August 2018 – added Simplified Tutorial chapter.

October 2017 – document first created.

Index

Adding quantities to a qualitative problem, 74
Amplification, 19
Annotating diagrams, 73
Anode terminal, 14, 27

Base terminal, 7, 19
Beta, 21
Bias, forward, 6, 14
Bias, reverse, 6, 14
Bipolar junction transistor, 3, 7, 19
BJT, 3, 7
Blocking voltage, 15
Breadboard, solderless, 58, 59
Breadboard, traditional, 62
Breakdown, diode, 16

Cardio-Pulmonary Resuscitation, 56
Cathode terminal, 14, 27
Channel, FET, 8
Checking for exceptions, 74
Checking your work, 74
Code, computer, 83
Collector terminal, 7, 19
CPR, 56
Current gain, 21

Dalziel, Charles, 56
Depletion MOSFET, 9, 25
Depletion region, 6–8, 14, 21, 23
Dimensional analysis, 73
DIN rail, 59
Diode, 3, 6
Diode equation, 14
Diode, light-emitting, 6, 17
Diode, Zener, 16
DIP, 58
Dopant, 13
Doping, 13
Drain terminal, 8, 9, 23, 25

Edwards, Tim, 84
Electric shock, 56
Electrically common points, 57
Electron tube, 5
Electronic circuit, 3
Electrostatic discharge, 26
Emitter terminal, 7, 19
Enclosure, electrical, 62
Enhancement MOSFET, 9, 25
Equipotential points, 57, 59
ESD, 26
Experiment, 63
Experimental guidelines, 64
Extrinsic semiconductor, 13

FET, 3, 8, 22
Field-effect transistor, 3, 8, 22
Forward bias, 6, 14
Forward current, 15
Forward voltage, 15

Gain, 19
Gain, current, 21
Gate terminal, 8, 9, 23, 25, 27
Germanium, 6
Graph values to solve a problem, 74
Greenleaf, Cynthia, 33

How to teach with these modules, 81
Hwang, Andrew D., 85
Hysteresis, 27

IC, 58
Identify given data, 73
Identify relevant principles, 73

IGBT, 10, 30
IGFET, 9, 25
Insulated Gate Bipolar Transistor, 10, 30
Intermediate results, 73
Intrinsic semiconductor, 13
Inverted instruction, 81

JFET, 23
Junction field-effect transistor, 23
Junction, PN, 13

Knuth, Donald, 84

Lamport, Leslie, 84
Latching, SCR, 27
LED, 4, 6, 17
Light-emitting diode, 6, 17
Limiting cases, 74
Load, 15

Majority carrier device, 23
Maximum Power Transfer theorem, 18
MCT, 31
Metacognition, 38
Minority carrier device, 14, 23
Moolenaar, Bram, 83
MOS-controlled thyristor, 31
MOSFET, 9, 25
Murphy, Lynn, 33

N-channel JFET, 24
N-type semiconductor, 6, 13
NPN, 19

Ohm's Law, 6, 14
Open-source, 83
Opto-SCR, 30
Opto-TRIAC, 30
Optocoupler, 21, 30
Optotransistor, 21

P-channel JFET, 23
P-type semiconductor, 6, 13
Phototransistor, 4
Photovoltaic cell, 6, 18
PN junction, 13
PNP, 19
Potential distribution, 59
Problem-solving: annotate diagrams, 73
Problem-solving: check for exceptions, 74
Problem-solving: checking work, 74
Problem-solving: dimensional analysis, 73
Problem-solving: graph values, 74
Problem-solving: identify given data, 73
Problem-solving: identify relevant principles, 73
Problem-solving: interpret intermediate results, 73
Problem-solving: limiting cases, 74
Problem-solving: qualitative to quantitative, 74
Problem-solving: quantitative to qualitative, 74
Problem-solving: reductio ad absurdum, 74
Problem-solving: simplify the system, 73
Problem-solving: thought experiment, 64, 73
Problem-solving: track units of measurement, 73
Problem-solving: visually represent the system, 73
Problem-solving: work in reverse, 74
Project management guidelines, 67

Qualitatively approaching a quantitative problem, 74

Reading Apprenticeship, 33
Recovery time, 15
Reductio ad absurdum, 74, 80, 81
Reverse bias, 6, 14
Reverse current, 15
Reverse voltage, 15

Safety, electrical, 56
Saturation, 28
Schoenbach, Ruth, 33
Scientific method, 38, 63
Scope creep, 67
SCR, 3, 11, 27
Semiconductor, 3, 5
Shockley diode equation, 14
Shockley, William, 14
Shunt resistor, 58
Silicon, 6, 13
Silicon-controlled rectifier, 3, 11, 27
Simplifying a system, 73
Socrates, 80

Socratic dialogue, 81
Solar cell, 4, 6, 18
Solar panel, 18
Solderless breadboard, 58, 59
Solid state, 5, 22
Source terminal, 8, 9, 23, 25
SPICE, 33, 64
SPICE netlist, 60
Stallman, Richard, 83
Static electricity, 26
Subpanel, 62
Surface mount, 59

Terminal block, 57–60, 62
Thévenin impedance, 18
Theorem, Maximum Power Transfer, 18
Thought experiment, 64, 73
Thyratron, 27
Thyristor, 3, 11, 27
Torvalds, Linus, 83
Transistor, 3, 6, 19
Transistor, bipolar junction, 19
Transistor, field effect, 22
Transistor, IGFET, 9, 25
Transistor, JFET, 23
Transistor, MOSFET, 9, 25
TRIAC, 3, 11, 29
Tube, electron, 5
Tube, vacuum, 5

Units of measurement, 73

Vacuum tube, 5, 22
Visualizing a system, 73

Wiring sequence, 61
Work in reverse to solve a problem, 74
WYSIWYG, 83, 84

Zener diode, 16