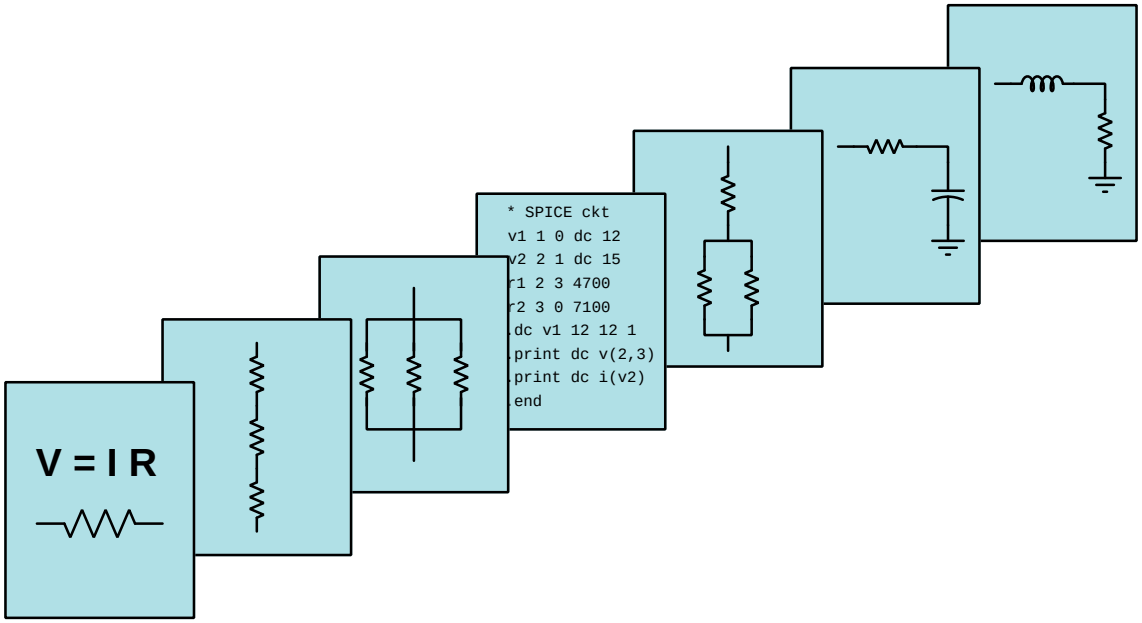
Modular Electronics Learning (ModEL) project



Transistor Switching Circuits

© 2019-2026 by Tony R. Kuphaldt – under the terms and conditions of the Creative Commons Attribution 4.0 International Public License

Last update = 5 May 2026

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

Contents

1 Introduction **3**
- 1.1 Recommendations for students . . . 3
- 1.2 Challenging concepts related to transistor switching circuits . . . 5
- 1.3 Recommendations for instructors . . . 7

2 Case Tutorial **9**
- 2.1 Example: low-side NPN switch . . . 10
- 2.2 Example: high-side NPN switch . . . 11
- 2.3 Example: low-side N-channel switch . . . 12
- 2.4 Example: high-side N-channel switch . . . 13
- 2.5 Example: interposing MCU to a heavy load . . . 14
- 2.6 Example: designing a BJT switching circuit . . . 17
- 2.7 Example: designing a MOSFET switching circuit . . . 19

3 Tutorial **21**
- 3.1 BJT and FET review . . . 21
- 3.2 Transistor saturation . . . 26
- 3.3 High-side and low-side switching . . . 31
- 3.4 BJT switching . . . 32
- 3.5 MOSFET switching . . . 34
- 3.6 Switching inductive loads . . . 36
- 3.7 Paralleling transistors . . . 38
- 3.8 H-bridge networks . . . 40

4 Historical References **43**
- 4.1 Schottky-augmented BJT switches . . . 44

5 Derivations and Technical References **49**
- 5.1 Alternative MOSFET symbols . . . 50
- 5.2 Gauss' Law of Electric Fields . . . 51

6 Animations **57**
- 6.1 Animation of an NPN BJT turning on . . . 58
- 6.2 Animation of an E-type MOSFET turning on . . . 73

7 Questions **87**
7.1 Conceptual reasoning . . . 91
7.1.1 Reading outline and reflections . . . 92
7.1.2 Foundational concepts . . . 93
7.1.3 Transistor summary . . . 95
7.1.4 Simple PNP lamp control . . . 96
7.1.5 Resistor-transistor-switch circuit . . . 97
7.1.6 Identifying BJT states . . . 98
7.1.7 Identifying MOSFET states . . . 100
7.1.8 Stepper motor drive circuit . . . 103
7.1.9 BJT datasheet research . . . 104
7.1.10 MOSFET datasheet research . . . 105
7.1.11 Proper BJT types . . . 106
7.1.12 Electronic ignition circuit . . . 107
7.1.13 Design a touch-controlled lamp . . . 109
7.1.14 Darlington and Sziklai pair connections . . . 110
7.1.15 ULN2068B Darlington array IC . . . 111
7.1.16 Logic gate voltage states . . . 113
7.1.17 Electron versus hole mobility . . . 114
7.1.18 Explaining the meaning of calculations . . . 115
7.2 Quantitative reasoning . . . 116
7.2.1 Miscellaneous physical constants . . . 117
7.2.2 Introduction to spreadsheets . . . 118
7.2.3 Voltage calculations in a simple lamp circuit . . . 121
7.2.4 Darlington pair current calculations . . . 122
7.2.5 Over- and under-driving a BJT switch . . . 123
7.2.6 Saturated versus unsaturated BJT switching circuits . . . 125
7.2.7 Calculations in TIP31C transistor circuit . . . 126
7.2.8 Calculations in IRF9520 transistor circuit . . . 127
7.2.9 BJT switching circuit simulation program . . . 128
7.3 Diagnostic reasoning . . . 129
7.3.1 BJT switching circuit configurations . . . 130
7.3.2 MOSFET switching circuit configurations . . . 133
7.3.3 Reliable and unreliable lamp control circuits . . . 134
7.3.4 Fixing a mis-designed BJT motor control circuit . . . 135
7.3.5 Fixing a mis-designed MOSFET motor control circuit . . . 136
7.3.6 Failed electronic ignition system . . . 137
7.3.7 Failed double-transistor ignition system . . . 139

A Problem-Solving Strategies **141**

B Instructional philosophy **143**
B.1 First principles of learning . . . 144
B.2 Proven strategies for instructors . . . 145
B.3 Proven strategies for students . . . 147
B.4 Design of these learning modules . . . 148

C Tools used **151**

D Creative Commons License **155**

E References **163**

F Version history **165**

Index **168**

Chapter 1

Introduction

1.1 Recommendations for students

This module focuses on the use of semiconductor transistors as on/off switching elements. It is an excellent prelude to an exploration of transistors as linear (i.e. analog) amplification devices. The Tutorial begins with a review of BJT and MOSFET operating principles, then progresses to the most common switching circuit configurations.

Important concepts related to transistor switching circuits include **bipolar junction transistor function**, **field-effect transistor function**, **saturation**, **majority** versus **minority** charge carriers, **sourcing** versus **sinking** current, PN junction **forward voltage**, **threshold voltage**, **sources** versus **loads**, **commutating diodes**, **paralleling** transistors, **swamping**, **ballasting**, **negative feedback**, and **H-bridge** circuits.

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

- How might an experiment be designed and conducted to measure the beta ratio for a bipolar junction transistor? What hypothesis (i.e. prediction) might you pose for that experiment, and what result(s) would either support or disprove that hypothesis?
- How might an experiment be designed and conducted to measure the threshold voltage for a field-effect transistor? What hypothesis (i.e. prediction) might you pose for that experiment, and what result(s) would either support or disprove that hypothesis?
- How might an experiment be designed and conducted to determine the point at which a transistor becomes saturated? What hypothesis (i.e. prediction) might you pose for that experiment, and what result(s) would either support or disprove that hypothesis?
- What is the fundamental operating principle of a BJT?
- What is the fundamental operating principle of a JFET?
- What is the fundamental operating principle of a MOSFET?
- What factors limit the current-handling capability of a transistor?

- What factors limit the switching speed of a transistor?
- How do we quantify *gain* for a transistor?
- What is high-side switching, and how do we do this effectively using a BJT versus using a MOSFET?
- What is low-side switching, and how do we do this effectively using a BJT versus using a MOSFET?
- How do NPN- and PNP-type transistors compare with one another?
- How do N-channel and P-channel transistors compare with one another?
- What does it mean for a component to "source" or "sink" current?
- What does "saturation" mean for a transistor when operated as an on/off switch?
- Why are commutating diodes necessary when switching inductive loads?
- How may we safely parallel transistors for switching high amounts of current?
- What purpose do the resistors serve in a BJT H-bridge circuit?
- Which ways do currents flow in a BJT H-bridge circuit during each of the two polarities of load energization?
- What purpose do the resistors serve in a MOSFET H-bridge circuit?
- Which ways do currents flow in a MOSFET H-bridge circuit during each of the two polarities of load energization?
- How might an experiment be designed and conducted to measure the beta ratio for a bipolar junction transistor?
- How might an experiment be designed and conducted to measure the threshold voltage for a field-effect transistor?

Central to this topic is the matter of how each type of transistor must be stimulated in order to ensure full-on and full-off operation. This, perhaps more than any other transistor-related topic, causes more confusion among students first learning about transistor switching circuits. While it is possible to memorize criteria for all types of transistors, it is preferable to memorize just the fundamental principles and reason from there. Understanding what, exactly, makes a BJT or a MOSFET turn on is key to being able to deduce the necessary stimuli to force the desired transistor states in any given application.

1.2 Challenging concepts related to transistor switching circuits

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

- **Necessary conditions for transistor operation** – bipolar junction transistors require a base current in the proper direction to inject minority carriers into the base layer (i.e. PNP transistors require hole injection from emitter to base, and so the emitter must be made positive and the base made negative, relative to each other; NPN transistors require electron injection from emitter to base, and so the emitter must be negative and the base positive.), and the collector-to-emitter voltage must be of the correct polarity (i.e. so that collector terminal current goes in the same direction as base terminal current, either *out* or *in*). Both currents join at the emitter terminal, making the emitter current the sum of the base and collector currents.

- **Sinking versus Sourcing currents** – this concept is easier to grasp when the student is accustomed to conventional flow notation for current direction, because then the terms "source" and "sink" make intuitive sense as though describing a fluid's direction of flow.

- **BJT beta** – the "beta" (β or h_{FE}) value for a bipolar junction transistor predicts the number of electric charge carriers that are allowed through its collector terminal for every one charge carrier passing through its base terminal. While not particularly stable over different operating conditions, this amplification factor for a BJT is nevertheless useful in making approximate current calculations for BJT-based circuits. One important caveat of beta is that it represents an *upper limit* (i.e. *maximum*) collector-to-base current ratio. If the collector-emitter voltage drop is low enough, it is possible for the transistor to pass less current through its collector terminal than beta would predict, in which case we say the transistor is "saturated".

- **Transistor saturation** – saturation in general refers to any condition where a dependent variable no longer varies significantly when the independent variable has reached a certain value. Part of the confusion surrounding the concept of saturation is that one must precisely define *which two* parameters are associated with each other. For example, when examining transistor characteristic curves, some will define "saturation" in terms of controlled signal versus controlling signal (i.e. the region where the curve rises steeply from zero), while others will define "saturation" in terms of controlled signal versus available source voltage (i.e. where the curve flattens out as supply voltage increases). Both are examples of saturation, but with respect to different independent variables. If the independent variable in question is not clearly defined, the student is left to wonder what "saturation" really means when they encounter characteristic curves having completely different regions labeled "saturation".

- **Relevance of foundational concepts** – this is not a new concept in itself, but rather an approach to learning and to problem-solving that many students struggle with. Students must be encouraged to continually apply trusted concepts such as Ohm's Law, Kirchhoff's Voltage Law, and Kirchhoff's Current Law to every new subject they encounter in Electronics, rather than to ignore these first principles in favor of memorizing new rules. Memorization without understanding is a recipe for failure. Logically reasoning from trusted laws and principles,

however, strengthens recall and hones the ability to tackle novel problems. Within the subject of transistor switching circuits, Ohm's Law, KVL, and KCL are ever-relevant and need only be augmented by a few new facts (e.g. BJT charge carrier injection, PN junction depletion regions, beta, FET channel modulation, FET drain-source resistance, etc.) to fully equip oneself for analyzing most any transistor circuit.

The *Case Tutorial* chapter contains multiple example scenarios of both saturated and unsaturated transistor switching circuits.

1.3 Recommendations for instructors

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

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

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

  Assessment – Students show how quantitative results were obtained by the author in the Tutorial chapter's examples.

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

- **Outcome** – Apply the concept of saturation to transistor switching circuit analysis

  Assessment – Determine whether or not a transistor will be operating in its saturated mode in a circuit given component values; e.g. pose problems in the form of the "Saturated versus unsaturated BJT switching circuits" Quantitative Reasoning question.

- **Outcome** – Design a transistor switching circuit

  Assessment – Choose appropriate transistor types/orientations to complete a schematic diagram for a circuit employing a transistor to switch power to a high-current load; e.g. pose problems in the form of the "Proper BJT types" Conceptual Reasoning question.

  Assessment – Sketch wire placement in a pictorial diagram for a circuit employing a transistor to switch power to a high-current load; e.g. pose problems in the form of the "Design a touch-controlled lamp" Conceptual Reasoning question.

  Assessment – Insert a transistor into an existing schematic diagram to augment that circuit to switch power to a high-current load; e.g. pose problems in the form of the "Electronic ignition circuit" Conceptual Reasoning question.

- **Outcome** – Evaluate transistor switching circuit designs

  Assessment – Identify whether or not a given transistor switching circuit is properly designed; e.g. pose problems in the form of the "BJT switching circuit configurations" and "Reliable and unreliable lamp control circuits" Diagnostic Reasoning questions.

- **Outcome** – Independent research

Assessment – Locate bipolar transistor datasheets and properly interpret some of the information contained in those documents including maximum power dissipation, maximum base current, range of beta values (often shown as h_{FE}), etc.

Assessment – Locate MOSFET datasheets and properly interpret some of the information contained in those documents including maximum power dissipation, threshold voltage (gate-to-source), and maximum drain-source resistance when on, etc.

Chapter 2

Case Tutorial

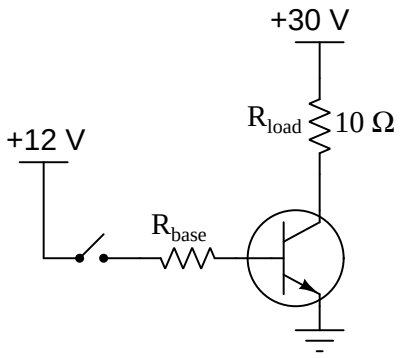
The idea behind a *Case Tutorial* is to explore new concepts by way of example. In this chapter you will read less presentation of theory compared to other Tutorial chapters, but by close observation and comparison of the given examples be able to discern patterns and principles much the same way as a scientific experimenter. Hopefully you will find these cases illuminating, and a good supplement to text-based tutorials.

These examples also serve well as challenges following your reading of the other Tutorial(s) in this module – can you explain *why* the circuits behave as they do?

2.1 Example: low-side NPN switch

Assumptions:

- $\beta = 25$
- $V_{CE} = 0.3$ Volts when saturated "on"
- $V_{BE} = 0.7$ Volts when "on"



Calculated parameters:

	Switch off	**Switch on**
V_{load}	0 V	29.7 V
I_{load}	0 A	2.97 A
I_{switch}	0 A	$>$ 118.8 mA
P_Q	0 W	0.974 W

Advertised β ratios assume the transistor to be in the "active" or "throttling" mode of operation, not at or near saturation which is the goal of a switching circuit. Therefore, the calculated switch current (equal to I_B) must actually be greater than that predicted by β and load current (I_C) in order to ensure a low collector-emitter voltage drop.

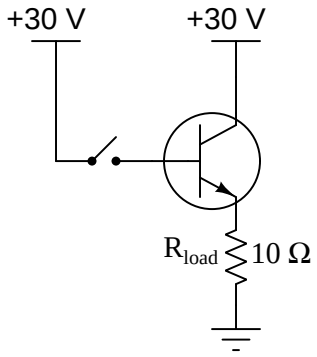
A base resistor value of 95.12 Ω would yield this exact amount of switch (base) current, and so to ensure saturation we would want to choose an R_{base} value less than this.

Analysis for a high-side PNP transistor switching circuit will proceed similarly to this circuit, with a base resistor necessary to establish the proper amount of I_B to ensure saturation.

2.2 Example: high-side NPN switch

Assumptions:

- $\beta = 25$
- $V_{BE} = 0.7$ Volts when "on"



Calculated parameters:

	Switch off	**Switch on**
V_{load}	0 V	29.3 V
I_{load}	0 A	2.93 A
I_{switch}	0 A	112.7 mA
P_Q	0 W	2.051 W

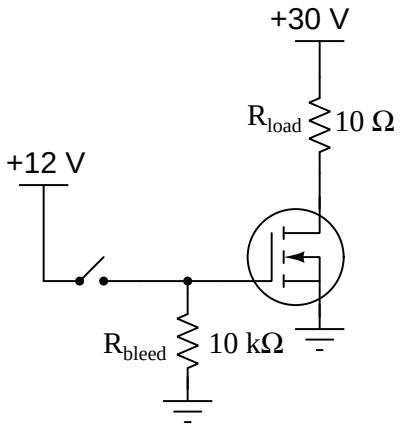
Closing the switch makes the collector and base terminals electrically common to each other, thereby ensuring $V_{CE} = V_{BE}$. At 0.7 Volts this is too great a V_{CE} voltage drop for the transistor to be saturated on, and therefore the load receives less voltage than it ideally could. No base resistor is necessary, because the base current self-regulates with the load being positioned between the emitter terminal and the DC source (ground).

Analysis for a low-side PNP transistor switching circuit will proceed similarly to this circuit.

2.3 Example: low-side N-channel switch

Assumptions:

- $V_{GS(on)} = 5$ Volts
- $R_{DS(on)} = 0.4\ \Omega$ when saturated "on"



Calculated parameters:

	Switch off	**Switch on**
V_{load}	0 V	28.85 V
I_{load}	0 A	2.885 A
I_{switch}	0 A	1.2 mA
P_Q	0 W	3.328 W

The "bleed" resistor exists solely to discharge the MOSFET's gate-source capacitance which might otherwise retain a voltage capable of keeping the transistor turned on after the switch opens. No resistor is necessary in series with the transistor's gate terminal because the gate of a MOSFET is insulated (inside the transistor, by the oxide layer) and therefore cannot pass a continuous current.

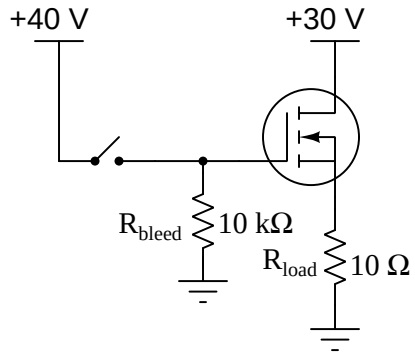
With the 12 Volt source easily exceeding the $V_{GS(on)}$ rating necessary to ensure the transistor's saturation in the "on" state, we may assume its channel resistance (R_{DS}) is at a minimum.

Analysis for a high-side P-channel transistor switching circuit will proceed similarly to this circuit.

2.4 Example: high-side N-channel switch

Assumptions:

- $V_{GS(on)} = 5$ Volts
- $R_{DS(on)} = 0.4\ \Omega$ when saturated "on"



Calculated parameters:

	Switch off	**Switch on**
V_{load}	0 V	28.85 V
I_{load}	0 A	2.885 A
I_{switch}	0 A	4.0 mA
P_Q	0 W	3.328 W

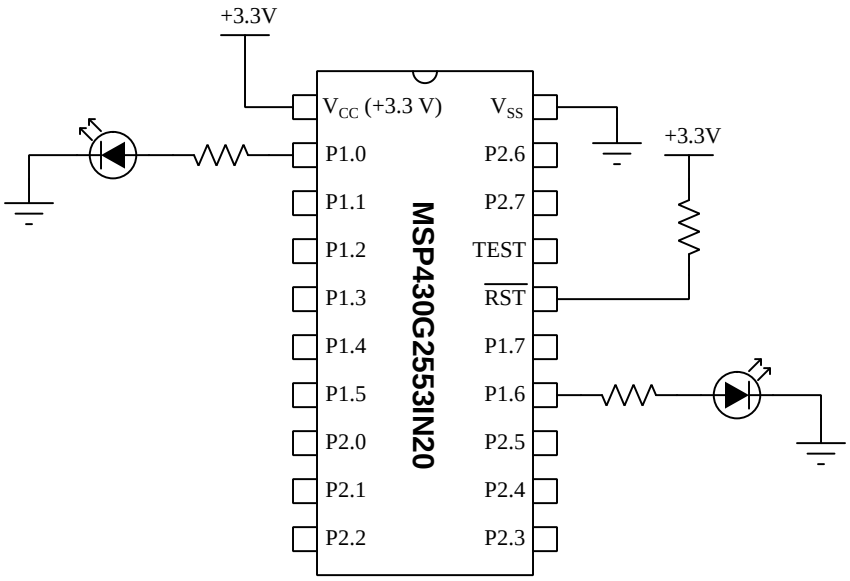
The "bleed" resistor exists solely to discharge the MOSFET's gate-source capacitance which might otherwise retain a voltage capable of keeping the transistor turned on after the switch opens. No resistor is necessary in series with the transistor's gate terminal because the gate of a MOSFET is insulated (inside the transistor, by the oxide layer) and therefore cannot pass a continuous current.

However, we must use a voltage source greater than that powering the load to successfully drive this MOSFET to saturation, because the $V_{GS(on)}$ rating demands a gate potential at least 5 Volts *greater* than the potential at the source terminal. This is why N-channel MOSFETs are seldom used as high-side switches, and when they are they require a supplementary voltage source to achieve the necessary gate-driving signal.

Analysis for a low-side P-channel transistor switching circuit will proceed similarly to this circuit.

2.5 Example: interposing MCU to a heavy load

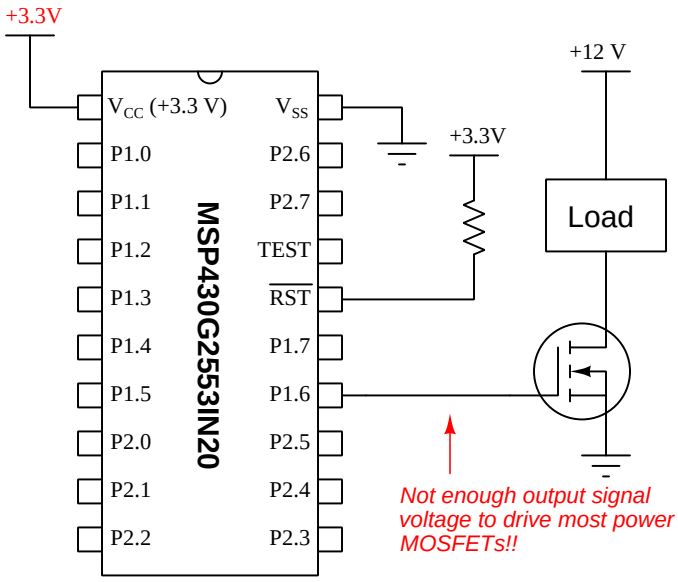
A *microcontroller* (or "MCU") is a single-IC computer designed to interface directly with switched devices. They are used extensively for what are called "embedded control systems" where they implement control functions in vehicles, appliances, etc., driving on/off discrete signals through sets of "output pins" on the device. For example, this MCU controls the illumination of two LEDs, one connected to output pin `P1.0` and the other connected to output pin `P1.6`:



A typical demonstration program stored to the MCU's memory might alternately blink these two LEDs on and off. Most MCU programs (and circuits!) are much more complicated than the one shown here, but this is good enough for our purposes.

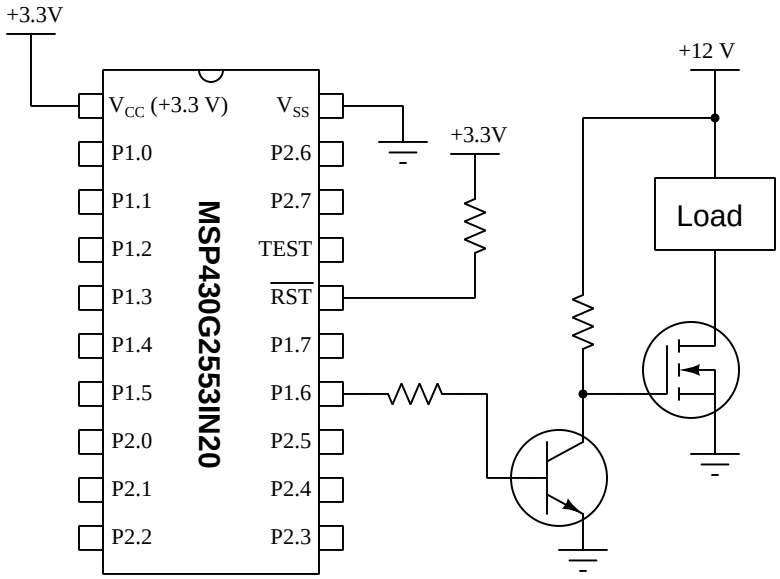
Suppose we wanted to control a load with much higher current requirements than a simple LED – something requiring far more current than the MCU can source or sink through any of its output pins. This will require some electrically-controlled switching device to "interpose" between the MCU's output and the high-current load. A MOSFET would be a natural choice for this, since when turned on it presents very little resistance to load current and therefore makes a very power-efficient switch in its "on" state.

It may seem as though something this simple might work, but it's quite possible it will not:



Many power MOSFETs have $V_{GS(on)}$ minimum requirements greater than 3.3 Volts, which is all this MCU is able to output with its 3.3 Volt power supply. Thus, a circuit like this would be unreliable. While the MOSFET will certainly be off whenever pin `P1.6` is in a "low" state (0 Volts to ground, logical "0"), the MOSFET will very likely only turn *partially* on whenever pin `P1.6` goes to a "high" state, resulting in less-than-full voltage across the load and overheating of the MOSFET due to it exhibiting substantial $R_{DS(on)}$ resistance.

The following solution is simple, using a BJT directly driven by the MCU to provide a higher-voltage gate-to-source signal for the MOSFET to operate on. The MCU will have no trouble activating the BJT since bipolar junction transistors only require 0.7 Volts base-to-emitter to turn on, which is far less than the several Volts of gate-source voltage typically required to activate a power MOSFET:



The resistor between pin `P1.6` and the BJT's base terminal simply limits base current to a level acceptable to the MCU. The resistor between the +12 Volt power supply terminal and the BJT's collector terminal simply acts as a "pull-up" to provide a "high" voltage to the MOSFET's gate terminal whenever the BJT is off, so its value may be quite large[1].

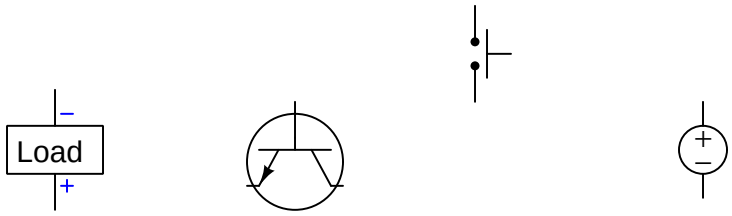
A caveat to be aware of with this added BJT stage is that the logic of the MCU's output logic state versus the load's energization state will be inverted: the load will now be energized when MCU pin `P1.6` is "low" (0 Volts to ground, which is a "0" digital state), and de-energized when pin `P1.6` is "high" (+3.3 Volts to ground, a "1" digital state). However, it is generally easy to alter the programming of the MCU to re-invert the logic so that the load energizes with an internal "1" digital state and turns off with a "0" digital state if that is what you desire.

Another caveat applies to loads that are inductive. When switching power to any inductive load (e.g. solenoid actuator, relay coil, electric motor), you will need to include a commutating diode in parallel with that load to provide a safe pathway for current whenever the MOSFET turns off and the load's magnetic field collapses.

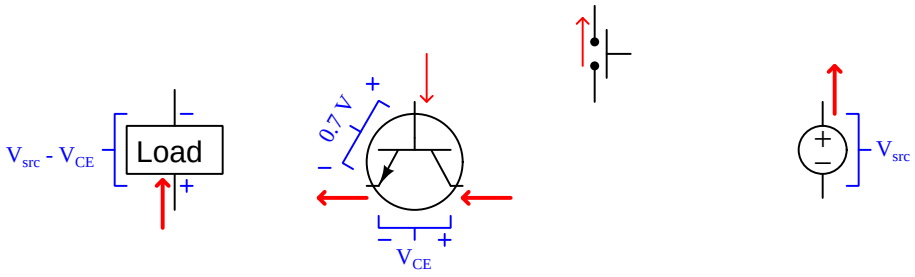
[1] The only real limit here is that this pull-up resistor must "charge" up the MOSFET's gate-to-source capacitance in a reasonably short amount of time, but since C_{GS} is typically very small for most MOSFETs this is generally not a concern. Only if the switching frequency is very high would we concern ourselves about sizing that pull-up resistor small enough to charge up C_{GS} quickly enough. Otherwise, a relatively large pull-up resistor value – say, 10 kΩ or more – will minimize power dissipation when the BJT is "on" which is a good thing.

2.6 Example: designing a BJT switching circuit

Suppose we needed to connect this NPN bipolar transistor in such a way that the pushbutton switch would turn it on, causing the transistor to energize the load:

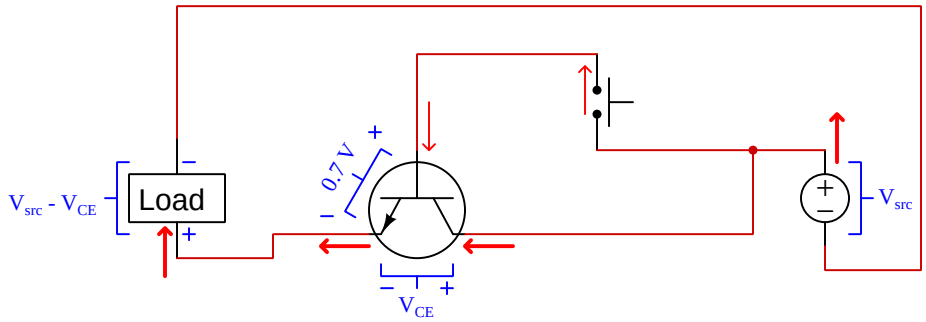


A practical first step in solving this problem is to correctly annotate all components with current arrows and voltage polarity symbols we would expect when the load is energized. Here I will use both thick and thin red arrows to denote large and small currents, respectively, so that I may clearly differentiate between the small *controlling* current carried by the switch versus the large *controlled* current provided by the source and energizing the load. Since the switch is a non-polarized component, the direction I chose for its current (up) is arbitrary. These annotations will help guide our wire connections, making sure all current directions and voltages are compatible:



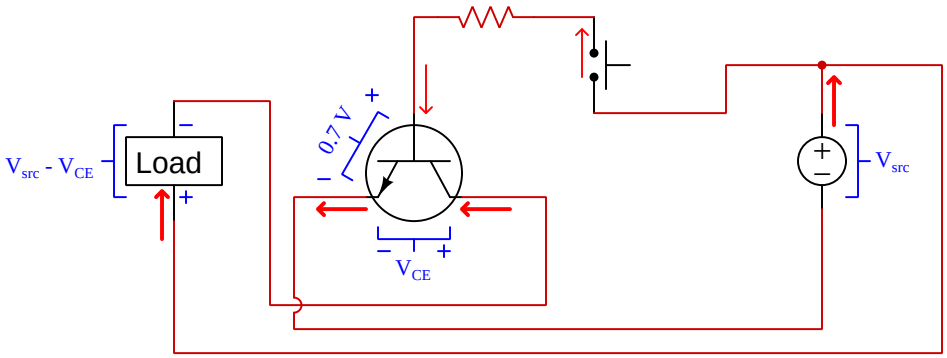
Note how the transistor's base-emitter junction voltage is assumed to be 0.7 Volts since we know that is typical for a forward-biased PN junction, and also how the load's voltage will be the difference between source voltage and the transistor's collector-emitter voltage drop. As we will see, multiple solutions are possible.

Here is a *high-side* solution where the transistor is positioned on the load's positive terminal:



V_{CE} must be equal to 0.7 Volts when the transistor is on, because the closed switch makes base and collector terminals electrically common with each other, and this in turn makes V_{CE} parallel with V_{BE}. This means the transistor will not be saturated, and its power dissipation will be equal to $0.7I_{load}$.

Here is a *low-side* solution where the transistor is positioned on the load's negative terminal:



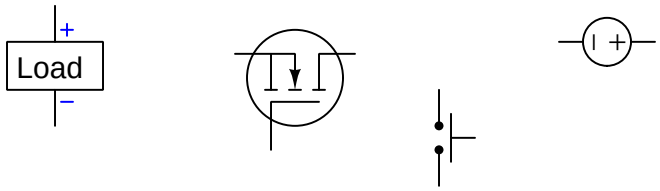
In this configuration a resistor must be placed in series with the base terminal in order to prevent full source voltage from being applied across the base-emitter junction when the switch closes. The value of this resistor must be chosen such that the transistor is able to pass full load current ($I_C = \beta I_B$). If we under-size this resistor so as to prepare the transistor to pass more current than the load will draw, we may ensure the transistor will be *saturated*[2] and only drop 0.3 Volts between collector and emitter when turned on. This decreases its power dissipation ($0.7I_B + 0.3I_C$) while also increasing load voltage compared to the high-side design.

Designing a transistor switching circuit using a PNP bipolar transistor rather than an NPN follows the same principles, the only significant difference being the current directions and voltage polarities at the transistor.

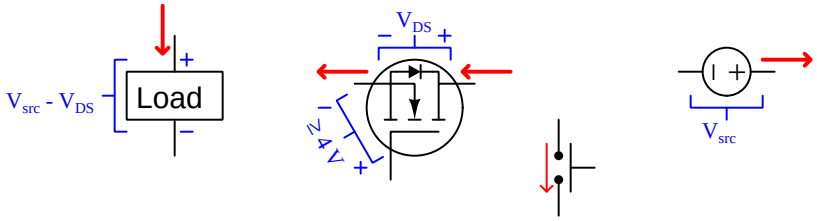
[2] This refers to the transistor being turned on to a degree beyond what is strictly necessary to pass the expected amount of load current. In other words, a transistor is "saturated on" when stronger activation of its controlling voltage or current fails to yield substantial increases in controlled current. To put this in colloquial terms, *the pedal is all the way to the floor.*

2.7 Example: designing a MOSFET switching circuit

Suppose we needed to connect this N-channel, E-type field-effect transistor in such a way that the pushbutton switch would turn it on, causing the transistor to energize the load:

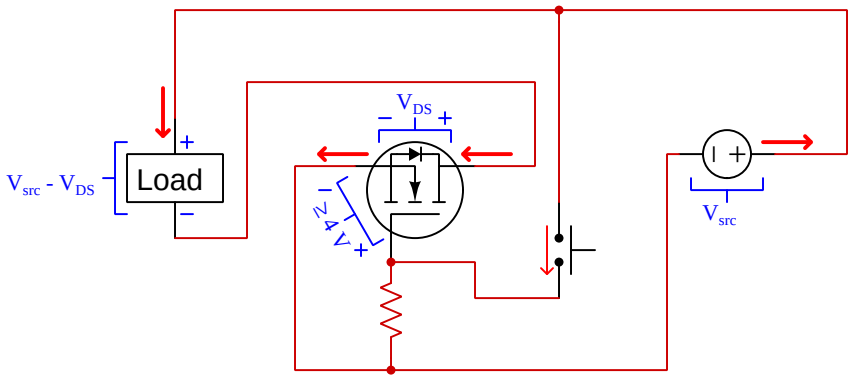


A practical first step in solving this problem is to correctly annotate all components with current arrows and voltage polarity symbols we would expect when the load is energized. Here I will use both thick and thin red arrows to denote large and small currents, respectively, so that I may clearly differentiate between the small *controlling* current carried by the switch versus the large *controlled* current provided by the source and energizing the load. Since the switch is a non-polarized component, the direction I chose for its current (down) is arbitrary. The MOSFET's parasitic diode has been added to aid in the proper annotation of drain-source current. These annotations will help guide our wire connections, making sure all current directions and voltages are compatible:



Here we assume a turn-on voltage ($V_{GS(on)}$, also known as *threshold voltage*) of 4 Volts, which is a reasonable figure for many power MOSFETs. This means the gate-source voltage (V_{GS}) must be equal to or greater than 4 Volts in order to assure the transistor's fully-on state. The load's voltage will be the difference between source voltage and the transistor's drain-source voltage drop.

A *low-side* solution is really the only viable configuration for the N-channel MOSFET given the relatively high V_{GS} voltage necessary to turn it on. If we were to attempt a high-side design we would require a second voltage source with a potential 4 Volts greater than V_{src} in order to adequately drive the MOSFET's gate. A low-side solution doesn't require this because the transistor's source terminal connects to the negative pole of the power supply rather than to the load:



A resistor is needed between the MOSFET's gate and source terminals to drain off the stored electrostatic charge inside the capacitance that is intrinsic to the dielectric (insulation) separating the gate terminal from the MOSFET's channel. If it were not for this resistor, the transistor would turn on just fine when the switch closes but likely fail to turn off once the switch opens. This resistor value may be quite large, because the natural gate-channel capacitance is extremely small[3] which means a resistor of tens or hundreds of kilo-Ohms in size will still create a quick time constant ($\tau = RC$) capable of bleeding off that stored charge and promptly returning the transistor to an "off" state.

Transistor saturation[4] is assured when $V_{GS} \geq V_{GS(on)}$, which means the only voltage dropped between drain and source will be due to the transistor's "on" channel resistance, typically in the milli-Ohm range for power MOSFETs. Power dissipation for the transistor is calculated just as though the transistor's conductive channel were a simple resistor: $I_{load}^2 R_{DS(on)}$, or $I_{load} V_{DS}$, or $\frac{V_{DS}^2}{R_{DS(on)}}$.

Designing a transistor switching circuit using a P-channel, E-type field-effect transistor rather than an N-channel follows the same principles, the only significant difference being the current directions and voltage polarities at the transistor.

[3] Typically in the range of tens or hundreds of pico-Farads, which means a resistor sized at 1 Mega-Ohm would create a time constant (τ) in the range of tens or hundreds of micro-seconds. This brief "turn-off" time delay is truly negligible for a transistor controlled by a pushbutton switch!

[4] This refers to the transistor being turned on to a degree beyond what is strictly necessary to pass the expected amount of load current. In other words, a transistor is "saturated on" when stronger activation of its controlling voltage or current fails to yield substantial increases in controlled current. To put this in colloquial terms, *the pedal is all the way to the floor.*

Chapter 3

Tutorial

3.1 BJT and FET review

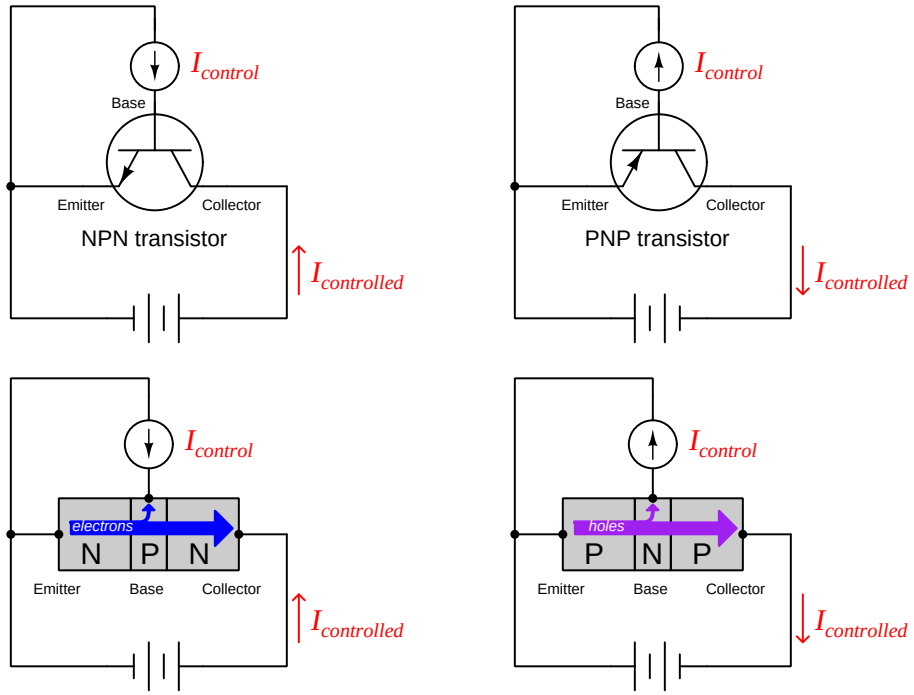
Transistors are semiconductor devices allowing one electrical signal to control another. Specifically, transistors work by allowing a relatively weak electrical signal to control the electrical conductivity of a solid material, and this modulated conductivity permits or denies passage of a much stronger electrical signal. This is the basic principle of *amplification*: where a smaller signal exerts control over a larger signal. Two major types of transistors dominate modern electronic circuits, *bipolar junction transistors* (BJTs) and *field-effect transistors* (FETs). Each has its own unique principle of operation[1], and although these principles differ markedly, both types of transistors perform similar functions within similar circuits.

Bipolar junction transistors function on the basis of *charge carrier injection*, where a weak electrical signal injects charge carriers into a region of semiconducting material previously lacking plentiful carriers. This injection of foreign charge carriers turns a formerly non-conductive material into a conductive material, allowing the transistor to pass a stronger (i.e. amplified) signal.

Field-effect transistors function on the basis of *channel width modulation*, where the effective width of an electrically conductive channel (i.e. pathway) varies with the application of an electric field from a weak voltage signal. This channel may be "pinched off" or "turned on" by this controlling voltage signal, allowing the transistor to conduct a stronger (i.e. amplified) signal.

[1] Most functional characteristics of transistors may be deduced from a proper understanding of these operating principles. Important behaviors such as the transistor's default state (i.e. whether or not it conducts current in the absence of a "driving" signal), as well as the necessary drive signal types to turn the transistor on (or off) also logically follow from these principles. When first learning about transistors, try to follow the logic connecting each transistor's operating principles with its functional characteristics. Avoid making the common mistake of merely committing facts to memory, devoid of understanding.

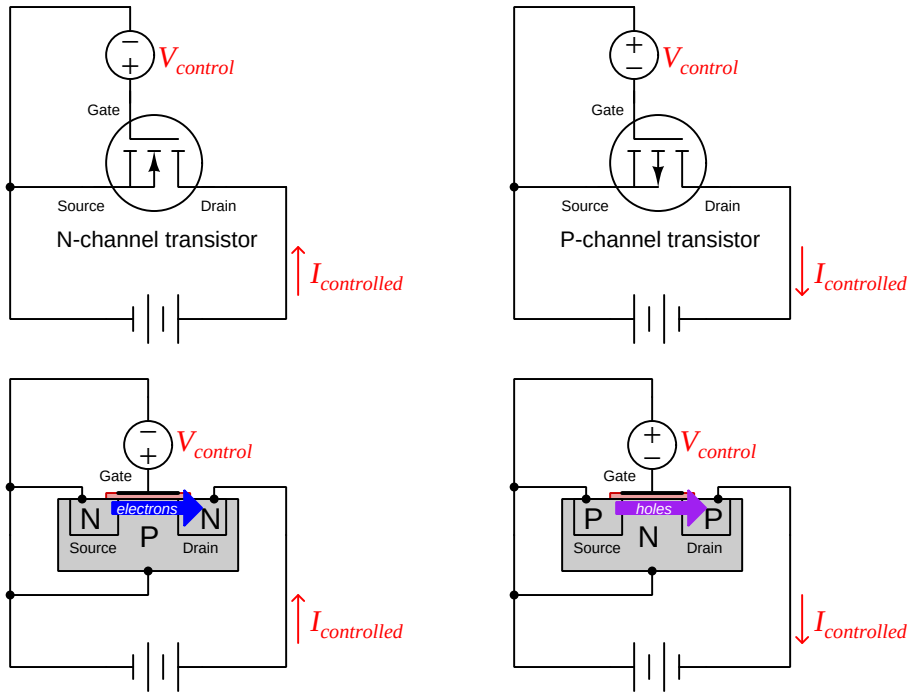
Bipolar junction transistors consist of a three-layer "sandwich" of differently-doped semiconductor materials, denoted either NPN or PNP. These three layers form two dissimilar "PN" junctions, hence the name *bipolar*. The middle layer (called the *base*) acts as a controlled-conductivity bridge between the other two layers (called the *emitter* and *collector*) conducting the controlled current. When unstimulated, base-emitter and base-collector depletion regions together prohibit current between collector and emitter. However, if an external source forward-biases the base-emitter junction to collapse that depletion region, charge carriers will be "emitted" by the emitter into the extremely thin base layer where most of them coast through the other side of the base and become "injected" into the depletion region of the base-collector junction, thus making the transistor conductive from end to end. Bipolar junction transistors are *minority-carrier* devices because the majority-type charge carriers emitted by the emitter layer are "minority" type once they enter the base (e.g. holes emitted from a P-type emitter, being injected in the N-type base).



The key to turning on a BJT is to forward-bias the base-emitter junction in order to initiate this charge-injection process. Without this injection of charge carriers into the depleted base-collector region, the transistor defaults to its "off" state because there is no continuous region of charge carriers naturally stretching from emitter to collector.

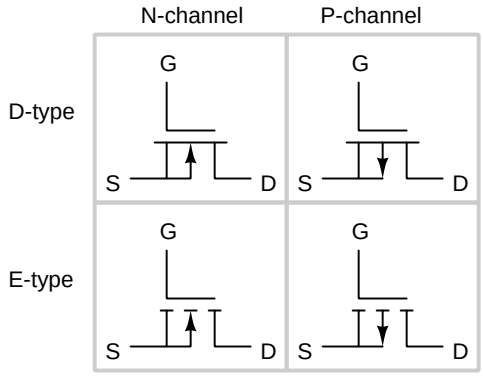
Field-effect transistors consist of a channel formed of one type of semiconductor material (i.e. either P-type or N-type) that is either flanked or interrupted by a region of the opposite doping type. An electric field applied transverse to this channel affects the quantities and types of charge carriers residing within the channel. Like bipolar junction transistors which exist in NPN and PNP forms, field-effect transistors may be manufactured with either N-type channels or P-type channels. However, unlike BJTs, there are several distinct design variations for FETs: *junction* FETs and two different forms of *insulated-gate* FETs (commonly called *MOSFETs* for their metal-oxide-semiconductor construction). For the sake of simplicity, we will show an *enhancement-mode* MOSFET in the following illustration because it behaves most similarly to a BJT.

Instead of emitter, base, and collector terminals, FETs have *source*, *gate*, and *drain* terminals. The controlled current flows between drain and source, while the gate terminal provides a point to apply the controlling voltage signal. Field-effect transistors are *majority-carrier* devices because the charge carriers leaving the source to pass through the channel (e.g. holes through a P-type channel, electrons through an N-type channel) are native to that channel's doping, never traversing a region of opposite doping:

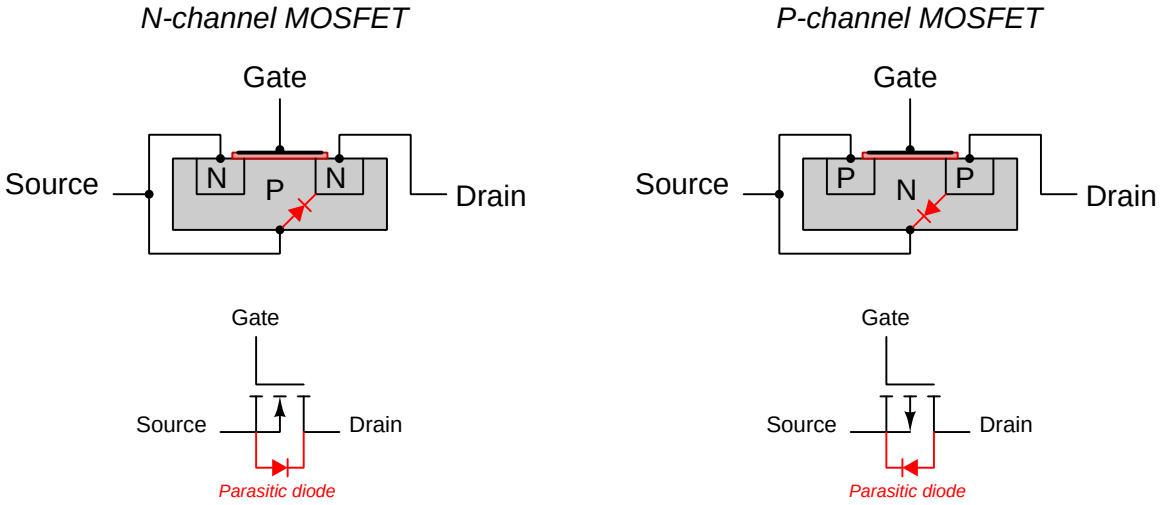


The key to turning on a FET is to electrostatically attract more majority carriers into the channel by giving the gate terminal a polarity opposite the channel type. Reversing this polarity forces the FET to turn off. Different forms of FET (JFET, depletion-mode, enhancement-mode) possess different natural states. JFETs are normally-on devices which may only be turned off (i.e. depleted). D-type MOSFETs are normally-half-on devices which may be turned on more (i.e. enhanced) or turned off more (i.e. depleted). E-type MOSFETs – the type shown above – are normally-off devices which may only be turned on (i.e. enhanced) similar to the behavior of a BJT.

Although the MOSFET is fundamentally a four-terminal device, they are often constructed with the substrate and source terminals connected by wire, which means the controlling voltage must be impressed between the gate and source terminals:



An interesting consequence of this internal connection is that it results in a single PN diode-like parasitic junction connected between the source and drain terminals:



When the source and drain terminals are properly biased[2], this parasitic diode will be reverse-biased and have no effect on the transistor's operation. However, if source-drain polarity ever reverses, this diode will become forward-biased and short past the transistor's channel!

[2]The arrow inside of a MOSFET symbol actually shows you which way this parasitic diode points! For any PN junction the tip of the arrowhead points toward the "N" side and the wide end of the arrowhead is on the "P" side. Note how the wide end of the N-channel MOSFET's arrow connects to the source terminal, as does the anode of the parasitic diode; likewise the tip of the P-channel MOSFET's arrowhead connects to the source terminal, as does the cathode of the parasitic diode. Your task in deciding how to connect any MOSFET into a circuit is to keep this parasitic diode reverse-biased so that the transistor will not be bypassed by it.

Field-effect transistor (FET) and bipolar junction transistor (BJT) behavior differs in several important ways:

Characteristic	**FET**	**BJT**
Charge carrier type	Majority	Minority
Principle of operation	Effective width of a continuous channel modulated by voltage	Charge carriers injected into a depleted region
Drive signal	Voltage only (no DC current)	Voltage and current both necessary
Switching time limit	RC time constant of gate capacitance and drive resistance	Stored charges in base waiting to be swept out
Temperature effect	Greater temperature increases channel resistance	Greater temperature decreases effective resistance
Resting status	Normally-on and normally-off types	Normally-off (only)
Gain	Transconductance $g = \frac{\Delta I_{DS}}{\Delta V_{GS}}$	Beta ratio $\beta = \frac{I_C}{I_B}$
Current regulation	Depletion region "pinch" effect intensifies with increasing V_{DS}	Rate of charge injection into base sets a maximum limit on I_C

Generally speaking, FETs are capable of switching on and off faster[3] than BJTs, and FETs generally offer less resistance[4] to the controlled current when fully turned on than BJTs which means FETs generally have higher current ratings than comparable BJTs. However, comparable BJTs generally have higher voltage ratings than FETs. Perhaps the starkest contrast between the two transistor technologies is their drive requirements: FETs never require a constant current at their gate terminals, whereas BJTs *do* require base current to continuously activate.

[3]The natural resistive-capacitive time delay of an FET's gate may be minimized through external circuit design, but the stored charge within a BJT's base region is largely a function of a transistor's internal construction.

[4]Charge carriers crossing the base-emitter junction of a BJT require an energy investment to do so, and much of this energy later dissipates in the form of heat. No such intrinsic barrier exists within a FET, which means conductance is limited only by charge carrier density and channel area.

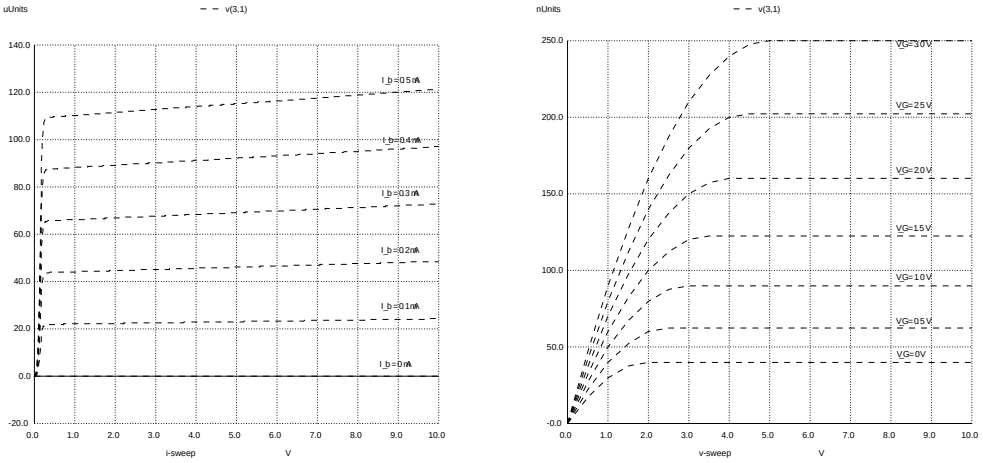
3.2 Transistor saturation

Semiconductors differ from conductors principally in the prevalence of mobile charge carriers. Conductive materials harbor a vast multitude of charge carriers free to move, just waiting for an applied electric field to prompt their motion. The rate of charge carrier drift through a conductor is directly proportional to applied voltage, which is the basis of Ohm's Law ($I = \frac{V}{R}$). In contrast, semiconductors offer very few free charge carriers unless and until stimulated in specific ways, which means voltage applied across a length of pure (intrinsic) semiconducting material results in practically zero current unless that semiconductor material is properly stimulated. A useful analogy for conductors is that of a sandy desert plain where grains of sand drift along by the force of the wind: there is no shortage of sand here, and so the rate of sand displacement across the plain is a direct function of wind speed. A companion analogy for intrinsic semiconductors is that of a rocky plateau devoid of sand: the absence of sand means no amount of wind will cause sand to drift across the plateau unless somehow sand is introduced into that environment.

Bipolar junction transistors (BJTs) are stimulated into a state of conduction by forward-biasing their base-emitter junction, causing charge carriers to be injected into the base-collector depletion region where few charge carriers existed before, and thus permitting conduction between the emitter and collector terminals. Enhancement-type MOSFETs are stimulated into a state of conduction by applying a transverse electric field to displace charge carriers into a region where they did not exist before, and thereby form a conductive bridge between the source and drain terminals. Both types of devices "turn on" by moving charge carriers into interior regions of the transistor formerly lacking those carriers, electrically bridging what is normally a non-conducting gap in the solid material. Without those respective stimuli, both the BJT and the E-type MOSFET become non-conductive and "turn off" due to a lack of available charge carriers within their interiors. Whereas conductors such as copper naturally present an abundance of free charge carriers, free charge carrier prevalence within a semiconductor is something we can control at will, and transistors exploit this property of controllable conductivity inherent to semiconductors.

If we *mildly* stimulate a transistor, we find that the controlled current is more or less fixed regardless of how much voltage we apply in the load circuit. This is why the current/voltage characteristic curves tend to flatten with increasing source voltage: whether BJT or MOSFET, the few charge carriers existing within the transistor to support this current are fixed in rate by the controlling stimulus. The situation here is analogous to spilling a trickle of sand onto a rocky plateau on a windy day: the number of sand grains drifting across the plateau depends almost entirely on the rate we spill the sand and not on the speed of the wind, simply because the only sand there to drift is the sand we intentionally spill – increasing wind speed raises the *velocity* of each sand grain but does not increase the total *number of sand grains drifting by per second*.

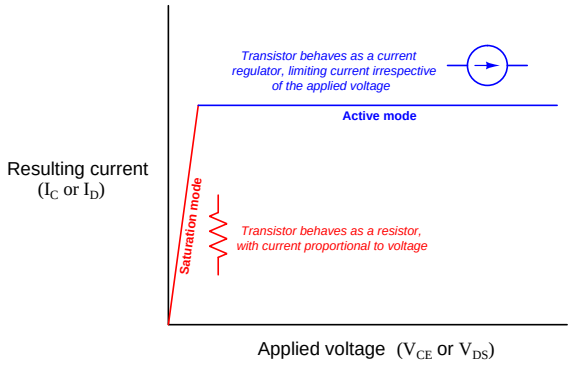
We see this current-limiting behavior for both BJTs (left) and MOSFETs (right) in the near-level portions of the characteristic curves shown below. For the BJT we plot collector current (I_C) on the vertical axis and collector-emitter voltage (V_{CE}) on the horizontal. For the MOSFET we plot drain current (I_D) on the vertical axis and drain-source voltage (V_{DS}) on the horizontal. Each curve represents a different degree of fixed stimulus, I_B for the BJT and V_G for the MOSFET:



As we progress from left to right on the horizontal axis of each graph, increasing applied voltage across each transistor, we see a dramatic increase of current at first but then a "leveling" of current to a relatively fixed value, that value determined by the degree of stimulus. We call these (nearly) level portions of the graph the transistor's *active mode* where it more or less regulates current. An "active" transistor does not obey Ohm's Law because the rate of charge carrier introduction is fixed by that stimulus, like the rate of sand drifting across a rocky plateau being fixed by the rate at which we spill new sand onto the plateau. Where Ohm's Law predicts current through a resistor increasing proportionately to applied voltage, here we see a transistor defiant in the face of a rising voltage as it holds current (relatively) constant.

If our intention, however, is to use a BJT or a MOSFET as an *on/off switch*, we do not want the transistor to be limiting current in any manner when we turn it on. What we'd rather have an "on" transistor do in any switching circuit is behave as similarly to a conductor as possible. Since we know transistors work by introducing charge carriers into formerly-depleted spaces, making a transistor behave as a good conductor means flooding its interior space with a plentiful supply of charge carriers. In other words, we need to *saturate* the transistor's interior with as many charge carriers as we can to turn it fully "on".

We see this *saturation mode* behavior in the near-vertical sections of the characteristic curves shown above: where the rise in current traces a steep upwards slope as voltage increases from zero. Appealing to our sand analogy again, a saturated transistor is akin to *dumping vast quantities of sand onto the rocky plateau*, temporarily turning that plateau into a sandy plain and permitting as much sand-particle drift as there is wind to drive it. Contrasting active and saturated modes for a generic transistor on a characteristic curve:

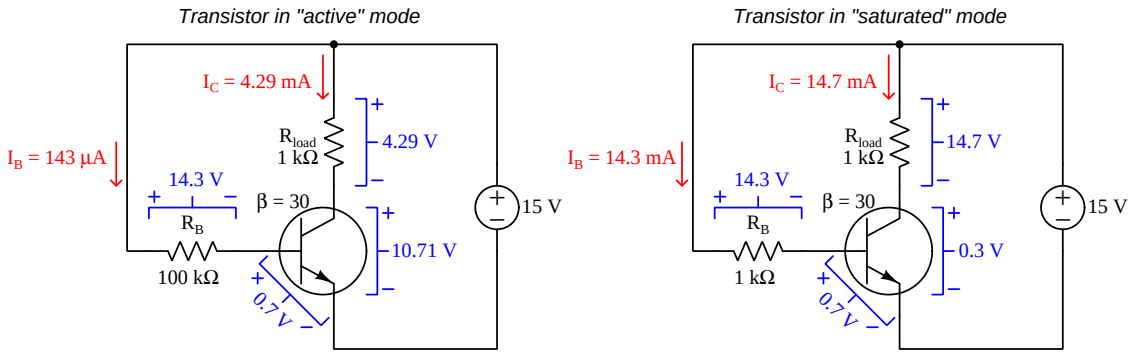


Predicting voltages and currents within "saturated" transistor switching circuits differs somewhat from predictions made in circuits where the transistor in question is in its "active" mode. For active-mode transistors, we assume the controlled current will be limited by the stimulating quantity (i.e. base current for a BJT, gate-source voltage for a FET), which is to say we assume the transistor *regulates* load current. Controlled-current prediction for a BJT uses the *beta* parameter, where $I_C = \beta I_B$, collector current limited to a multiple of base current. Controlled-current prediction for a FET uses the square-law proportion, where $I_D = k(V_{GS} - V_{th})^2$, drain current being limited to a proportion of the square of the difference between gate-source voltage and the FET's threshold voltage. Each of these formulae would be more truthful, however, if we replaced the "equals" symbol with a "less-than-or-equals" symbol, indicating that the predicted current is really a *maximum* value, and that the actual controlled current could be less if limited by something external to the transistor:

$$I_C \leq \beta I_B \qquad\qquad I_D \leq k(V_{GS} - V_{th})^2$$

Such is the case with saturated-mode transistors: we intentionally over-drive the transistor when turning it "on" in order to flood its interior with a super-abundance of free charge carriers to make it as good an electrical conductor as it is capable of being. This means we must design a saturated-BJT switching circuit to force a base current significantly larger than what the beta ratio would ordinarily require for the expected load current. Similarly, for a saturated-MOSFET switching circuit we would design it to ensure we had more gate-source voltage than would ordinarily be necessary for the expected drain current.

Compare the following BJT circuits, the left-hand circuit operating the transistor in its active mode and the right-hand circuit operating the transistor in its saturated mode, each with the same source voltage and load resistance values:



In the active-mode circuit we find the collector-to-base current ratio agreeing with the BJT's advertised beta value: with a base current of 143 microAmperes and a beta value of 30, this transistor cannot allow more than 4.29 milliAmperes through its collector because that is the limit set by the rate of charge carrier injection. With the load resistance being 1000 Ohms and the supply voltage 15 Volts, this forces the transistor to drop most of that source voltage between its collector and emitter terminals. In the saturation-mode circuit the advertised beta value multiplied by the increased base current would ordinarily predict a collector current of 429 milliAmperes, but this is impossible[5] with a 15 Volt source and a 1000 Ohm load. So, instead we find the saturated transistor dropping a negligible amount of voltage between its collector and emitter terminals as it passes nearly all the current the source voltage and load resistance will allow. To summarize, the "active" transistor functions as a current-regulating load while the "saturated" transistor functions as a load with a very small voltage drop.

In a perfect world a saturated BJT would drop zero voltage between its collector and emitter terminals regardless of load current, but in reality there will always be some voltage dropped. A commonly-used approximation for saturated V_{CE} is 0.3 Volts, but the real value varies with base current overdrive, collector current magnitude, device temperature, the cross-sectional area of the emitter and collector regions, etc. Similarly, a perfect MOSFET in its saturated state would act like a zero-Ohm resistor, but instead it offers a finite amount of resistance commonly referred to as $R_{DS(on)}$ subject to many of the same parameters as the BJT.

[5] Even if the transistor were able to act as a perfect short between collector and emitter, 15 Volts across 1000 Ohms would only yield 15 milliAmperes of collector current.

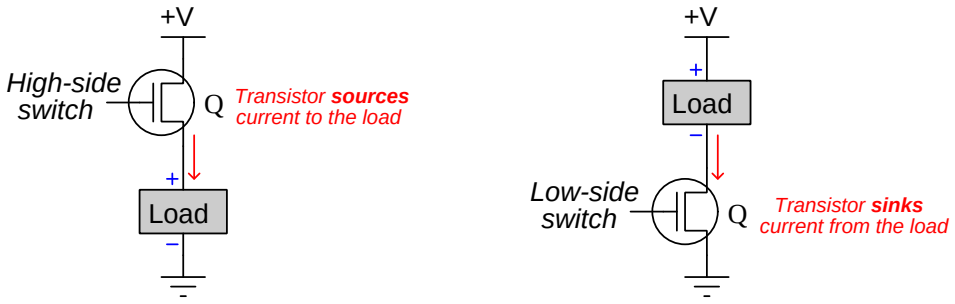
I recommend the following procedure when designing a transistor switching circuit with intent to saturate the transistor in its "on" mode:

- First, calculate maximum possible load current assuming the transistor is fully "on" (i.e. V_{CE} = 0 Volts for a BJT and V_{DS} = 0 Volts for a MOSFET). You may imagine the transistor being shorted (from collector to emitter for a BJT, from drain to source for a MOSFET) and then calculate how much current the load will allow for the given power supply voltage.
- **For a BJT**, use the *lowest* advertised beta value (β, or h_{FE}) shown in the transistor's datasheet to calculate the amount of base current necessary to ensure at least that much load current. Your circuit's base current should be made *greater* than this to guarantee saturation just in case the real beta value is less than advertised, so consider this a low design limit. Then, research the transistor's maximum allowable base current (again, from its datasheet). Your circuit's base current will need to be *less* than this so as to not harm the transistor, so consider this a high design limit. Choose a base current value for your circuit that is somewhere between those two design limits, then select components (e.g. base current-limiting resistor) necessary to establish the base current at your chosen value. Finally, calculate transistor power dissipation based on an assumed collector-emitter voltage drop of 0.3 Volts (or any saturated V_{CE} value shown in the datasheet for a comparable amount of load current), and determine how much heat-sinking the transistor will require based on datasheet thermal ratings.
- **For a MOSFET**, research the transistor's turn-on or threshold gate-to-source voltage ($V_{GS(on)}$). Your circuit's gate-source voltage will need to be *more* than this to guarantee saturation, so consider this a low design limit. Then, research the transistor's maximum allowable gate-source voltage. Your circuit's gate-source voltage will need to be *less* than this so as to not harm the transistor, so consider this a high design limit. Design your circuit so that the amount of gate-source voltage you impress upon the transistor is between these two design limits. Finally, calculate transistor power dissipation based on the maximum possible load current and the transistor's $R_{DS(on)}$ parameter specified in the datasheet, and determine how much heat-sinking the transistor will require (if any).

A reasonable question to ask at this point is, *"How can I empirically determine if a transistor is in its active or saturated mode?"* In other words, it's one thing to *predict* necessary parameters to force a transistor into saturation, but how may we tell we've actually achieved saturation when testing a real circuit? Empirically this may be determined by measuring V_{CE} for a BJT and V_{DS} for a MOSFET while the transistor is conducting load current – this measured voltage should be negligible if the transistor's interior is actually saturated with free charge carriers, because the load terminals of the transistor (collector and emitter for a BJT; drain and source for a MOSFET) should pose little opposition to through-current in a saturated state.

3.3 High-side and low-side switching

When used as an on/off switch for a DC load, a transistor[6] may either be placed in series on the load's positive-terminal side or on its negative-terminal side. In honor of the positive pole of a DC power source being considered a higher potential than the negative pole, these two transistor placements are referred to as *high-side switching* and *low-side switching*, respectively:



Another useful convention for specifying transistor and load placement in DC circuits is to describe the current's direction relative to each component. Following "conventional flow" notation for current, we would say that the high-side transistor *sources* current to the load, and that the load *sinks* current from the transistor. For the low-side switching circuit, we would say the opposite: the load *sources* current to the transistor, and the transistor *sinks* current from the load. Here, the term "source" is equivalent to *give* while the term "sink" means *receive*, again assuming current direction represented using conventional flow notation[7].

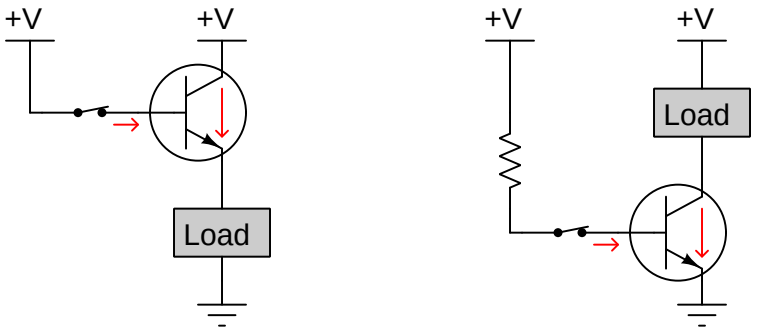
Control of electric power delivered to a load is possible in either circuit configuration, whether high-side (transistor sourcing) or low-side (transistor sinking), and so at first this circuit design choice may seem arbitrary. However, there is an important detail omitted from the simple examples previously shown: how we must "drive" the transistor's controlling terminal (i.e. the *base* terminal for a BJT and the *gate* terminal for a FET) in order to achieve fully-on and fully-off transistor states. Like any good switch, an "off" transistor should exhibit infinite resistance between its controlled terminals, and an "on" transistor should exhibit negligible resistance between its controlled terminals. If we were to "drive" the transistor to a partially-on state where it presents moderate resistance to load current, the result will be limited power delivered to the load and energy wasted at the transistor (in the form of dissipated heat). Therefore, the transistor's driving signal must be such that it drives the transistor between two extreme states rather than in-between "on" and "off". Designing the circuit in such a way that the transistor receives the necessary driving-signal states is not necessarily simple, as we will soon see.

[6] In the following diagram we see a generic FET symbol used for the transistor (Q), but this symbol is really intended to represent *any* type of transistor.

[7] This is one of several reasons why these tutorials exclusively use conventional flow notation rather than electron flow notation. Tracing the motion of electrons in circuits makes perfect sense for vacuum tubes, and is helpful when tracing conduction in N-doped semiconductors, but it fails miserably for other uses. The terms "source" and "sink" make no sense when viewed from the perspective of electron flow. Also, the arrows implicit in semiconductor device symbols all assume conventional flow to show the direction of forward current through PN junctions. Finally, the complementary notation of voltage polarity (+ and −) makes sense of energy gains and losses for sources and loads *only* if current is traced in the "conventional" direction.

3.4 BJT switching

Consider the proper use of an NPN bipolar junction transistor in both high-side and low-side switching configurations. Since BJTs are minority-carrier devices, this means *electrons* internally emitted from the N-type emitter region pass through the P-type base to become minority carriers and make the transistor conductive, with a small percentage of these electrons exiting the base terminal and the majority being received by the collector and exiting the transistor's collector terminal. Therefore, we need the emitter to be more negative than the collector, and also more negative than the base, in order to properly bias the transistor. Current, labeled according to conventional flow, points opposite the actual motion of electrons, which means controlled (load) current from collector to emitter and controlling (drive) current from base to emitter:



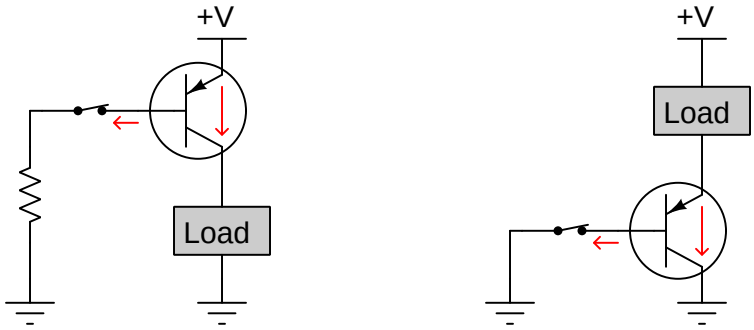
In either circuit, closing the mechanical switch causes the transistor to turn on and energize the load. However, the load will not energize to the same degree in these two circuits. In the high-side circuit the load will receive approximately 0.7 Volts less than +V when energized, but in the low-side circuit the load will receive slightly more: about 0.3 Volts less than +V when energized. Recall that the base-emitter junction of a conducting transistor always drops approximately 0.7 Volts which is standard for a forward-biased silicon PN junction (V_{BE}). When a BJT's driving (base) current is high enough to *saturate* the base layer with minority carriers for maximum conductivity, there will actually be less than 0.7 Volts dropped between collector and emitter (V_{CE}), usually closer to 0.3 Volts. Turning our attention back to the schematic diagrams we see why the low-side switching configuration is able to achieve this saturated-on state but the high-side circuit cannot. In the high-side circuit the closed mechanical switch makes the base and collector terminals electrically common to each other, which forces V_{CE} to be exactly equal to V_{BE}, and thus $V_{CE} \approx 0.7$ Volts. However, in the low-side circuit V_{BE} and V_{CE} are not forced into equality, and so with proper sizing of the resistor[8] to saturate the transistor's base we may have $V_{BE} \approx 0.7$ Volts and $V_{CE} \approx 0.3$ Volts or less.

The freedom to saturate the low-side NPN transistor but not the high-side NPN transistor means the low-side-switched load can receive more power and also that the low-side transistor will dissipate less power compared to the high-side circuit. Thus, the low-side switching configuration is more energy-efficient than the high-side using NPN transistors. An advantage of the high-side NPN circuit, though, is the ability to turn off faster than the low-side circuit for the same reason: a

[8] This resistor is necessary in the low-side circuit to avoid over-powering the transistor's base-emitter junction which of course requires only 0.7 Volts to forward-bias. If the base were directly connected to +V through the switch, the transistor would be destroyed!

saturated BJT requires slightly more time[9] to clear its base layer of all injected charge carriers than a non-saturated BJT once current ceases through the base terminal.

Now let us consider the use of PNP bipolar junction transistors in both high-side and low-side switching applications. Again, since BJTs are minority-carrier devices, this means for a PNP transistor that *holes* will internally leave the P-type emitter to enter the N-type base as minority carriers, after which some will exit the transistor's base terminal while most pass through the base layer on to the P-type collector and exit out of the collector terminal. Proper biasing of a PNP transistor, therefore, consists of having the emitter terminal more positive than either the base terminal or the collector terminal in order to send those holes in the right directions. Currents drawn in the directions of conventional flow notation will follow the paths of these holes: into the emitter terminal, out the base terminal, and out the collector terminal:



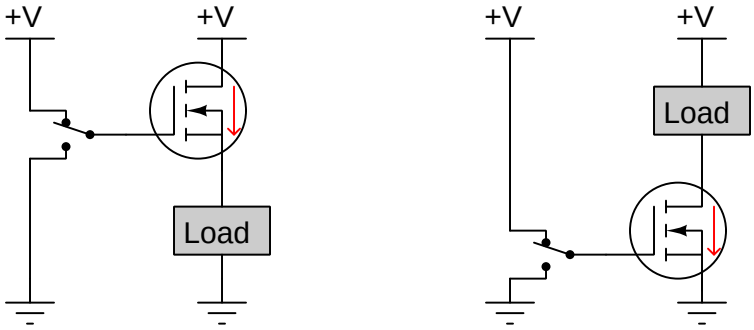
Comparing the PNP versus NPN versions of these switching circuits, you will no doubt notice a few important differences. In the NPN versions we saw the mechanical switch connecting the transistor's base terminal upward to +V (either directly or through a resistor), but here in the PNP versions we see the switch routing current downward from the base terminal toward ground. This, of course, is necessary for proper biasing of the base-emitter junction. With the mechanical switch *sinking* current from the transistor's base rather than *sourcing* current as was the case with the NPN transistors, we see the necessity of a base resistor move from the low-side to the high-side configuration. Now, with PNP transistors, it is the high-side circuit requiring a resistance to avoid over-powering the base-emitter junction while the low-side circuit tolerates a direct connection from base to ground. Similarly, the high-side circuit is now the most energy-efficient because the high-side transistor is able to fully saturate[10] while the low-side transistor cannot.

[9] This "storage time" is typically in the order of fractions of a microsecond.

[10] In either NPN or PNP circuits using a resistor to limit base current, this resistor's size is relatively easy to calculate. By Kirchhoff's Voltage Law we know that this resistor must drop all but 0.7 Volts' worth of the source voltage (e.g. with a 12 Volt source, the resistor will drop approximately 11.3 Volts in the "on" state). The transistor's datasheet should specify the necessary base current to achieve saturation for a desired amount of load current, and with these two parameters (resistor voltage drop and resistor current) we may apply Ohm's Law ($R = \frac{V}{I}$) to calculate the necessary resistor size. In the absence of a specified saturation I_B, one may simply estimate the required base current to conduct load current based on β and then be sure to make the base current significantly greater than that. Note that saturation is guaranteed *only* with a sufficiently low-valued resistor, which makes the resistor's selection extremely important for efficient circuit operation. Too much resistance and the transistor will fail to saturate, making its V_{CE} higher than necessary and wasting energy. Too little resistance and the transistor may become damaged by overpowering the base-emitter junction.

3.5 MOSFET switching

Now we will explore the use of insulated-gate (MOSFET) transistors for switching purposes, beginning with N-channel enhancement-mode MOSFETs[11]. Field-effect transistors are majority-carrier devices, which for an N-channel FET means *electrons* must be internally emitted from the source to pass through the N-type channel and be received at the drain, and that in order to turn the transistor on we must apply a *positive* polarity to the gate terminal with respect to the substrate to attract electrons to the channel. Likewise, the source must be more negative than the drain in order to motivate negatively-charged electrons to flow from source to drain. Current drawn in the direction of conventional flow notation points opposite the actual flow of electrons through the transistor, and so our load current arrow points from drain to source. The gate of a MOSFET is electrically insulated and therefore *cannot* pass a continuous current:



In either circuit the transistor will be "on" with the toggle switch in the upward-facing direction (as shown) because the gate terminal is made positive with respect to the substrate[12] which enhances the N-type channel to energize the load. However, as it so happens the low-side switching circuit will function quite well but the high-side circuit will function very poorly. The MOSFET with its source terminal connected to ground will enjoy the full source voltage (+V) between its gate and substrate when the switch flips "up" and this will make its channel strongly conductive. The MOSFET with its drain connected to +V will have an elevated substrate potential once the load drops voltage, thereby severely limiting the amount of voltage available between gate and substrate to turn it on. Recall that a BJT only requires approximately 0.7 Volts to forward-bias its base-emitter junction, but a MOSFET typically has a gate-substrate "threshold" voltage rating of *several Volts* necessary for turn-on. What will happen in the high-side circuit when activated is that both the transistor and the load will only be partially energized, with the transistor dissipating a great deal of heat due to its large drain-source voltage drop.

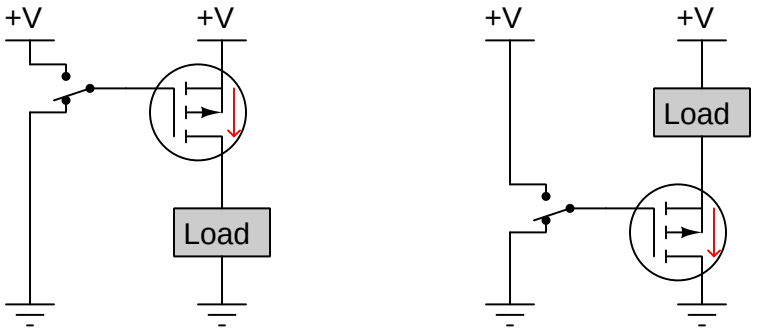
This problem – of using an N-channel FET in a high-side switching role – is so severe that the circuit is simply impractical as drawn. The only way to achieve the saturated "on" state necessary for efficient operation is to apply a gate voltage *greater than +V* by a margin exceeding the transistor's threshold voltage[13].

[11] These are the most similar in biasing requirements to NPN-type BJTs which was the first type of BJT we explored as a switching element.

[12] Like many MOSFETs, these transistors have their substrate terminals internally connected to their source terminals.

[13] It is not uncommon to see a gate driving voltage of 12 Volts when the load power supply is only 5 Volts, giving the

Exploring the use of P-channel MOSFETs for high-side and low-side switching, we must consider (again) the fact that FETs are majority-carrier devices and therefore the source terminal of a P-channel FET must always be more positive than the drain terminal in order to allow positively-charged *holes* to flow from source to drain. Accordingly, the gate terminal must be made *negative* with respect to the substrate in order to attract holes to the channel and enhance the transistor into a conducting state. Conventional-flow current notation points in the same direction as the actual hole motion:



In either circuit the "down" position of the toggle switch turns the transistor on, because that is the position applying the negative polarity to the gate terminal necessary to attract positively-charged holes to the channel inside the transistor. The effectiveness of each configuration is reversed from what we saw with N-channel FETs, though: now it is the high-side circuit that is able to function as intended (because its transistor sees full source voltage between gate and substrate with the switch in the "down-on" position) while the low-side circuit struggles to turn on completely. As with the high-side N-channel circuit, a greater voltage level is required for the gate on the low-side P-channel transistor in order to achieve the required V_{GS} threshold voltage while the load is fully energized[14]. In this case, the gate terminal of the low-side MOSFET must be brought to a potential that is *below ground* (i.e. a negative voltage with respect to ground) in order to saturate it.

A less apparent problem with the use of P-channel FETs for switching purposes is the fact that holes are far less mobile than electrons[15], all other factors being equal. This means any given P-channel FET will exhibit greater channel resistance when "on" (and have a correspondingly lower current limit) than an N-channel FET of the same dimensions, doping concentration, heat-shedding ability, and operating conditions. For this reason, N-channel FETs are much preferred over P-channel FETs for power-switching applications. In fact, you will likely never see a P-channel FET used for switching except in high-side configurations where the simpler driving requirements of the P-channel FET may outweigh the disadvantage of lower conductivity.

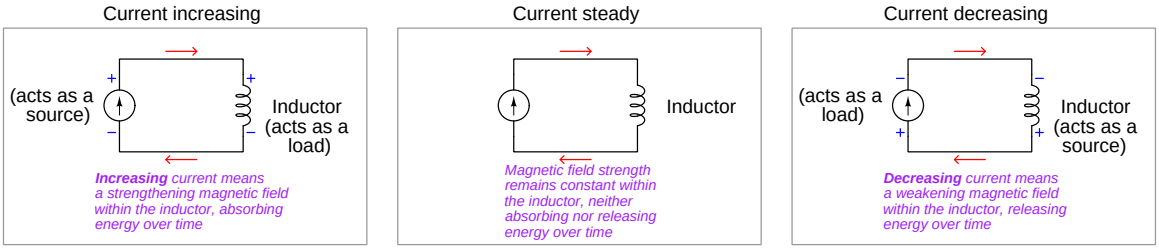
FET 7 Volts of gate-to-substrate voltage to ensure a fully-on state. This requires either a separate (higher-voltage) source just to drive the gate, or a *charge pump* circuit to boost the power supply voltage to a level high enough to suffice.

[14] If this is not clear to see, just imagine a perfect "on" state where V_{DS} is zero and the load drops full source voltage. This would put the transistor's drain *and* source terminals both at ground potential. Since the gate of a P-channel enhancement-mode MOSFET must be negative with respect to its substrate (source) in order to turn on, this means the gate potential must be negative with respect to ground.

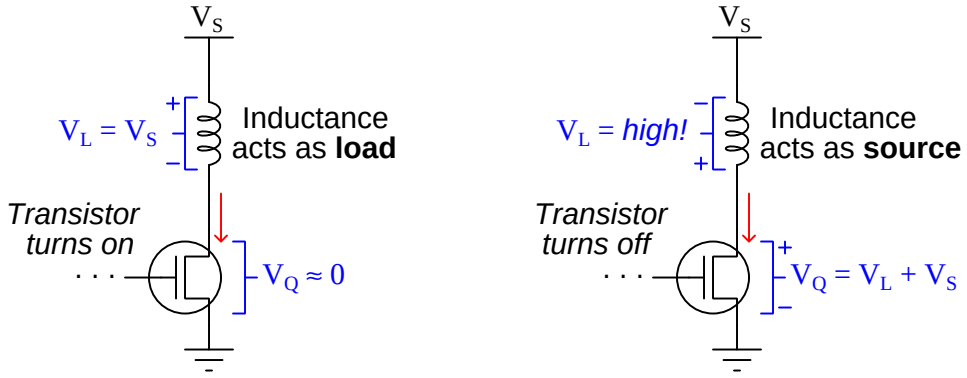
[15] The ratio between electron mobility and hole mobility typically exceeds 2:1, making the difference in effectiveness between N-channel and P-channel FETs significant indeed!

3.6 Switching inductive loads

An important consideration when using any type of transistor to interrupt current to an inductive load is that the inductance of the load will behave (momentarily) as an electrical *source* during the interruption. Inductors, you will recall, store energy in their magnetic fields, and the strength of any electromagnetic field is proportional to current through the conductor. Since energy is always conserved, and inductor current directly relates to its stored energy, both the direction and magnitude of an inductor's current must therefore be conserved, which simply means neither is able to change without a transfer of energy in or out of the inductor. As current increases through an inductor, its magnetic field grows which means it absorbs energy from the rest of the circuit as a *load*. Conversely, as current decreases through an inductor, its magnetic field weakens which means it releases its stored energy to the rest of the circuit as a *source*:

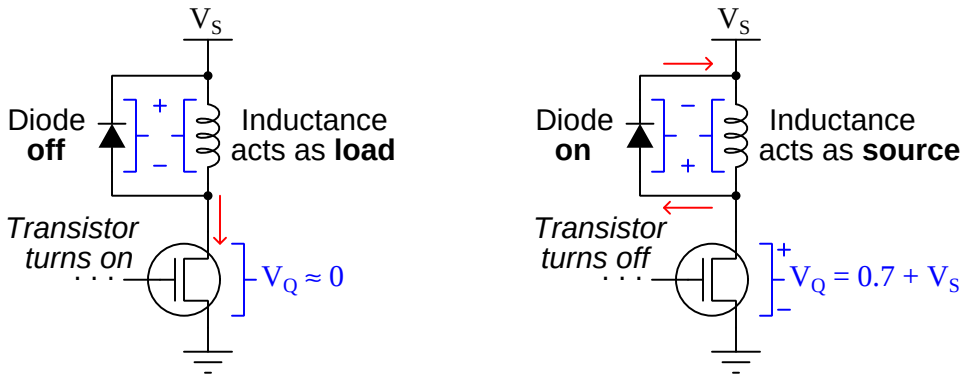


The quicker current is forced to decrease through an inductor, the greater its generated voltage as it releases its stored energy at a more rapid rate. Transistors are able to turn off *very rapidly*, which will result in any series-connected inductance generating *high voltage*. This voltage output by the de-energizing inductance will be dropped across the transistor's power terminals, possibly enough to destroy the transistor!



Note how the turned-off transistor must drop the *sum* of the inductor's voltage and the DC source voltage, given the reversed polarity of the inductor's voltage.

A simple and very common solution to this problem consists of connecting a diode in parallel with the inductive load, oriented such that it will become forward-biased only when the inductance begins to act as a source. This is typically known as a *commutating diode*[16] or a *free-wheeling diode*[17] or a *flyback diode*[18], because it switches the path of the inductance's current during the de-energization period and allows the inductor to gradually de-energize like a flywheel slowly losing speed to friction:



The result of this diode is that the transistor need only withstand a voltage equal to the DC source voltage (V_S) plus the diode's forward voltage drop (approximately 0.7 Volts for a silicon diode) when it turns off, rather than the hundreds or even thousands of Volts that might be output by the collapsing magnetic field of a rapidly-de-energizing inductance.

Power transistor arrays manufactured as integrated circuits often have built-in diodes for this very purpose. The anode of each commutating diode is connected to the same terminal of the IC as the current-sinking terminal of its respective power transistor, while the cathode goes to a dedicated IC terminal usually labeled V_S or some equivalent suggesting external connection to the positive rail of the inductive load's DC power source.

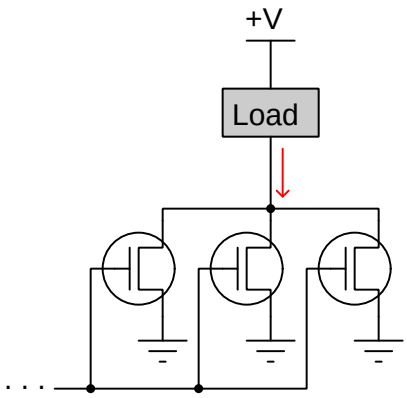
[16] "Commutate" simply means "to switch".

[17] In this case, the energized inductor is akin to a mechanical flywheel still possessing rotational inertia which must be dissipated.

[18] The origin of this term is anchored in the design of antique television monitors using cathode-ray tubes (CRTs), the electron beam's horizontal deflection being controlled by a special oscillator circuit that would sweep the beam from left to right and then very quickly "fly back" to the left-hand side of the screen to sweep left-to-right again in order to illuminate another row on the screen. These became known as "flyback" circuits. This same oscillator (typically operating at a frequency of 15.75 kHz) was also employed in these old television receiver units to energize a step-up transformer to generate high voltage necessary for proper operation of (other) parts of the cathode ray tube (CRT), simply because this high frequency allowed the use of a light-weight transformer with a minimal iron core. A step-up transformer operating at the typical line frequency of 50 Hz or 60 Hz would have required much more iron in its core to function properly. This voltage-boosting circuit generated high voltages by energizing a winding (coil) of this transformer and then suddenly halting current through it in order to force the core's magnetic field to suddenly collapse and induce a high voltage across one of the other windings, kind of like how our example inductive load is able to generate transistor-destroying voltage levels. Therefore, the legacy of these television "flyback" circuits is invoked in applications where a high voltage can potentially develop from an inductor's current being forcibly halted by a transistor switch suddenly turning off.

3.7 Paralleling transistors

In applications where load current exceeds the current-handling ability of any single transistor, multiple transistors must be paralleled:

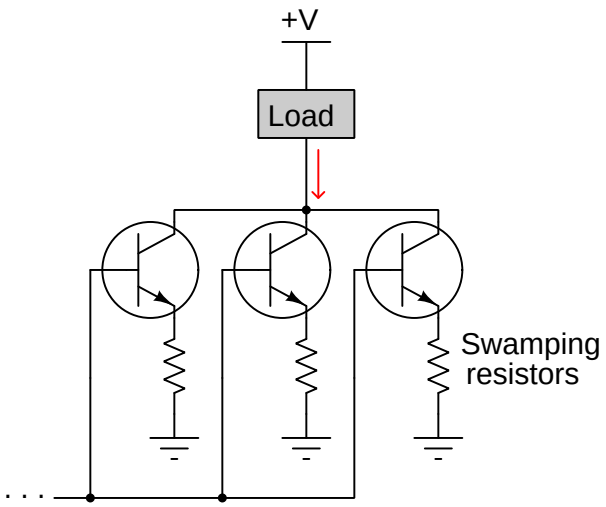


Ideally, load current will be split into exactly equal proportions according to the number of identical transistors connected in parallel with each other. However, no two transistors are ever *exactly* the same and so we must always expect some current imbalance. A prudent design choice is to parallel more transistors than strictly necessary assuming equal division of current, just in case one or more of them end up carrying more than their fair share.

However, the problem may be more complicated than imperfect matches between paralleled transistors. If bipolar junction transistors are used, a phenomenon called *thermal runaway* may occur. The basis of thermal runaway is that BJTs are "minority carrier" devices and therefore drop less voltage (i.e. allow easier passage of current) between collector and emitter as their temperature increases[19]. Any imbalance between paralleled BJTs will result in the hotter transistor bearing more of the load's current, which of course increases its power dissipation and makes it hotter yet. The end-result is one overheating transistor conducting most of the load current while the other transistors run cool.

[19] Majority carriers are those charge carriers predominating within a doped semiconductor: electrons within N-type materials and holes within P-type materials. Minority carriers are the opposite: holes within N-type materials and electrons within P-type. Bipolar junction transistors work on the basis of charge carriers being *injected* into a region of doped material in which they are not native. For NPN transistors this means "majority" electrons native to the N-type emitter being injected as "minority" carriers into the P-type base, most of those electrons carrying over into the N-type collector's depletion region where they are swept through by the reverse-bias base-collector voltage. The hotter the transistor, the longer the those minority carriers will persist within the base without being "recombined" with majority carriers there, which means more will become injected into the base-collector depletion region and be swept through the transistor. So, higher temperature essentially increases the transistor's beta ratio and makes it more conductive from emitter to collector, all other factors being equal. This makes a hot BJT amplify current to a greater degree than a cold BJT, and therefore drop less collector-emitter voltage when turned on.

A clever solution to the problem of thermal runaway is to add resistance in series with the emitter terminal of each transistor so as to introduce a (small) voltage drop proportional to its current. This additional resistance helps solve the problem in two different ways. First, the presence of additional resistance makes the transistor's effective collector-emitter resistance seem smaller by proportion, so that changes in the transistor's characteristics due to heat will have less effect than before when the *only* factor determining current division was the effective collector-emitter resistance of each transistor. This general strategy is known by the colorful name of *swamping*[20] or *ballasting*[21]. Second, the placement of said resistor on the *emitter* terminal of each transistor adds some *negative feedback*[22] which will act to decrease current if other factors make that current increase.



Interestingly, FETs do not suffer from the problem of thermal runaway because their drain-source channel resistance naturally *increases* with temperature, owing to the fact FETs are "majority carrier" devices[23]. If one FET among a paralleled bank of FETs happens to become hotter than the others, it automatically restricts current to force the others to carry more, and thereby redistributes power dissipation to the other transistors.

[20] The notion here is that the added resistance "swamps out" variations in resistance between transistors so those variations are no longer significant.

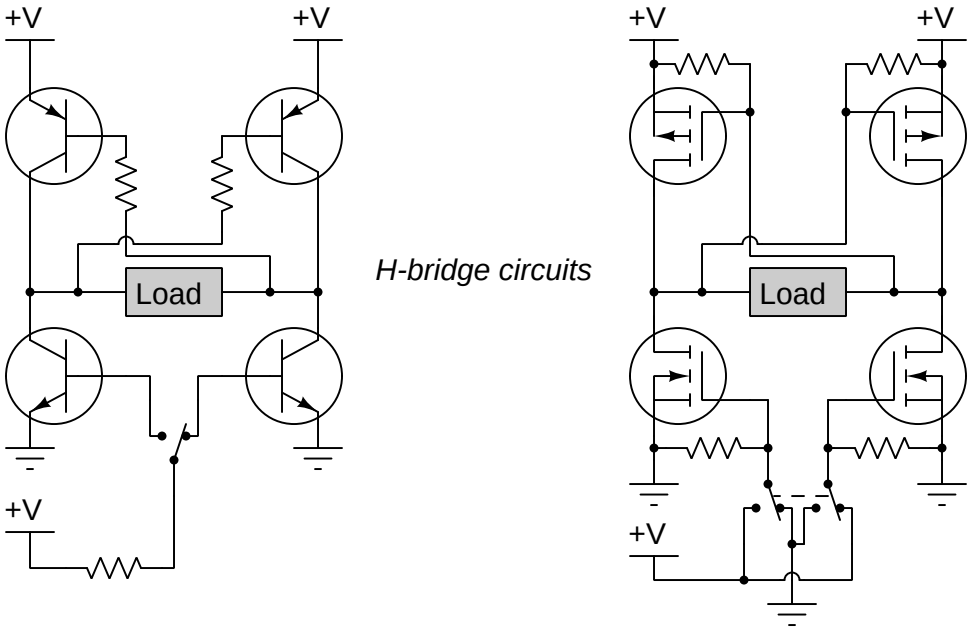
[21] "Ballasting" is a general term referring to any method limiting current.

[22] Negative feedback is a complicated concept, and really impossible to fully explain in a footnote. A simple explanation is that for any given base-terminal voltage (measured with respect to either power supply terminal), an additional resistive voltage drop between the emitter terminal and the nearest power supply terminal acts to reduce the base-emitter voltage drop by Kirchhoff's Voltage Law. This directs the current-heavy transistor to conduct a bit less current, forcing the other paralleled transistors to bear more of the burden and thereby dissipate more heat. The hallmark of negative feedback is *self-correction*, in this case the offending transistor self-correcting (i.e. forcing itself not to saturate as hard) in order to make the balance of currents more even.

[23] FETs rely on majority carriers (i.e. charge carriers native to a region of semiconductor material) to conduct the controlled current. Majority carriers behave like electrons within metal: there are so many of them already present that the major factor impeding current is charge-carrier collision driven by temperature rather than charge carrier scarcity as is the case for minority carriers. Higher temperatures cause charge carrier motion to be more erratic, increasing collisions with the atomic lattice and with each other, creating more electrical resistance between an FET's drain and source than if the device were cold.

3.8 H-bridge networks

A very useful form of transistor switching circuit with the unique capability to reverse load polarity is the so-called *H-bridge*. These are commonly used to switch power on and off, as well as control the direction, of DC electric motors. Both BJT and MOSFET H-bridge circuits are shown in the following schematic diagrams:



The general design philosophy of an H-bridge circuit is the same regardless of transistor type. A double-throw switch activates one of the two lower transistors, which in turn provides the means to activate the upper transistor in the opposite corner. With two transistors in opposite corners of the bridge turned on, and the other two turned off, the upper "on" transistor sources current to the load while the lower "on" transistor sinks current to ground. Flipping the double-pole switch causes all four transistor states to reverse, sourcing and sinking current through the load in the other direction and thereby reversing the load's polarity.

Resistors in the BJT H-bridge serve to limit base current to a level sufficient to saturate the transistors. Resistors in the MOSFET H-bridge serve a different purpose, and that is to provide a definite electrical potential[24] at each transistor's gate terminal so it cannot "float" (i.e. have no definite connection to either ground or to +V). It is always important to design FET circuits such

[24]This exact same technique of using a resistor to provide a default +V state or ground state in the absence of another connection is commonly used in digital electronic circuits utilizing FETs, and for precisely the same reason. FETs are controlled by nothing more than electric fields, and because of this require absolutely no current through their gate terminals to maintain a given state. A FET with a "floating" gate will probably just remain in the last state it was commanded to assume by the last applied gate signal, but it could also switch state if a stray electric field from an external object happens to influence the floating gate terminal.

that the gate terminals never "float" because that may lead to erratic transistor behavior, allowing the transistor(s) to turn on or off at the whim of random static-electricity charges.

Chapter 4

Historical References

This chapter is where you will find references to historical texts and technologies related to the module's topic.

Readers may wonder why historical references might be included in any modern lesson on a subject. Why dwell on old ideas and obsolete technologies? One answer to this question is that the initial discoveries and early applications of scientific principles typically present those principles in forms that are unusually easy to grasp. Anyone who first discovers a new principle must necessarily do so from a perspective of ignorance (i.e. if you truly *discover* something yourself, it means you must have come to that discovery with no prior knowledge of it and no hints from others knowledgeable in it), and in so doing the discoverer lacks any hindsight or advantage that might have otherwise come from a more advanced perspective. Thus, discoverers are forced to think and express themselves in less-advanced terms, and this often makes their explanations more readily accessible to others who, like the discoverer, comes to this idea with no prior knowledge. Furthermore, early discoverers often faced the daunting challenge of explaining their new and complex ideas to a naturally skeptical scientific community, and this pressure incentivized clear and compelling communication. As James Clerk Maxwell eloquently stated in the Preface to his book *A Treatise on Electricity and Magnetism* written in 1873,

> It is of great advantage to the student of any subject to read the original memoirs on that subject, for science is always most completely assimilated when it is in its nascent state . . . `[page xi]`

Furthermore, grasping the historical context of technological discoveries is important for understanding how science intersects with culture and civilization, which is ever important because new discoveries and new applications of existing discoveries will always continue to impact our lives. One will often find themselves impressed by the ingenuity of previous generations, and by the high degree of refinement to which now-obsolete technologies were once raised. There is much to learn and much inspiration to be drawn from the technological past, and to the inquisitive mind these historical references are treasures waiting to be (re)-discovered.

4.1 Schottky-augmented BJT switches

BJTs operated in a "saturated" mode where more minority charge carriers are injected into the base layer than strictly necessary, for the purpose of ensuring a solid "on" state for the transistor, suffer the limitation of relatively slow turn-off time. The primary mechanism of this time delay is *stored charge* within the base layer. As base current is interrupted, some injected charges still remain within the base layer which permit continued collector current until they are swept out of the base at which point the transistor returns to its default "off" state.

James Robert Biard, an engineer working for Texas Instruments, devised a clever solution to the stored charge delay problem by preventing the transistor from ever operating in its saturated mode. His United States patent filed on 31 December 1964 and granted 26 August 1969, entitled *Unitary Semiconductor High Speed Switching Device Utilizing a Barrier Diode* (number 3,463,975), describes the problem of transistor delay times in high-speed digital logic circuits on page 1:

> In electronic circuitry of the type used to perform logic operations in digital computing systems one of the main factors influencing the design is the Switching speed, i.e., the time required to change the conductive state of the functional elements of the circuitry. When transistors are used as the active elements in such circuits, the primary limiting factor with respect to the switching speed is the time required to turn on and turn off the collector current in these semiconductor devices. This invention is primarily concerned with this speed factor.
>
> The switching speed of transistors can be optimized by judicious selection of the device geometry, i.e., the size and shape of the active regions and junctions, and of the electrical characteristics of the regions of the device, particularly the resistivity, impurity gradient, carrier life time, etc., as well as selection of the semiconductor material itself. Once the speed of the device is increased to the maximum permitted by materials technology and production feasibility, the operating speed of the system in which such devices are used can be greatly affected by the circuitry surrounding the devices.
>
> The period of time required to turn on a transistor, referred to as "rise time," can be minimized by driving the base of the device with fairly large electrical signals. Unfortunately, this large input tends to drive the transistor into a saturated condition, to the detriment of another factor affecting the switching speed, this being the "storage time." When the transistor is saturated, the collector-base junction is forward biased and the base region stores a large concentration of minority carriers. Before the transistor can be considered turned off, the collector-base junction has to be returned to the usual reverse bias state, and to do this the stored carriers must be swept out of the base region. The time period required to do this, referred to as the storage time, is often the primary limiting factor on the switching speed. Thus, the operating speed of the logic system is often a compromise between using sufficient driving voltage to obtain a fairly short rise time for the transistors, but yet keeping the devices out of saturation to avoid unduly long storage time.

Biard's patent then goes on to describe existing methods of improving transistor delay in high-speed circuits, all of which defied economical construction in an integrated circuit which was the challenge facing engineers at the time who desired to build complex digital logic circuits on single

wafers of silicon rather than construct them from discrete components. Such *integrated* circuits were necessary to build compact electronic computers.

> Several techniques have been developed for increasing the switching speed of such circuits, including the use of resistance-capacitance coupling so that voltage spikes appear on the transistor base to turn on and turn off the unit while the hold-on voltage is much lower, holding the transistor out of saturation. The disadvantages of R-C coupling are the necessity of additional components and the inherent delay introduced by the R-C combination itself.
>
> Another technique which has been quite successful is the use of a diode in a transistor switching circuit shunting the collector-base junction and biased or otherwise adapted to conduct in the forward direction at a slightly lower voltage than the collector-base diode. This prevents the collector-base from becoming forward biased, a condition necessary for saturation, and therefore avoids storage of a high concentration of carriers in the base. With this arrangement, the base can be driven with large signals, but yet storage time is quite short. The forward voltage necessary for conduction of the shunt diode may be reduced below that of the collector-base junction by adding a low level voltage source in series with the diode, but this is inconvenient in some circuits, especially integrated semiconductor networks. Another technique is to use a different material for the diode, i.e., use a germanium diode with a silicon transistor, but again this is impossible in a monolithic integrated circuit where all of the components are formed in a unitary body of semiconductor material.

Biard's invention was to construct a special type of diode within the integrated-circuit transistor, one made of a metal-semiconductor junction rather than a semiconductor-semiconductor junction. He explains it thusly:

> In accordance with this invention, a P-N junction in a semiconductor device is prevented from becoming forward biased by shunting the junction with a metal-semiconductor diode, commonly referred to as a Schottky barrier diode. Particularly, the collector-base junction of a transistor has a metal-semiconductor diode connected across it and poled for conduction in the forward direction the same way as the collector-base. The metal-semiconductor diode may be fabricated upon the same semiconductor wafer as the transistor, or each transistor in an integrated semiconductor circuit may have one of the shunt diodes connected therewith on the monolithic semiconductor bar. A unitary semiconductor element is thereby provided which can be operated at high switching speeds.

Schottky diodes consist of a junction between a metal and an N-type semiconductor. Both materials conduct electricity using free electrons (and not holes), and so there are no minority carriers anywhere within a Schottky diode that require "sweeping out" of the junction when turning off. This makes Schottky diodes fast-switching devices, and furthermore they only happen to drop about 0.4 Volts when forward-biased compared with 0.7 Volts for standard silicon PN junction diodes. This low forward voltage drop would allow the diode to "turn on" before the transistor's base-collector junction could become forward-biased, thereby clamping the base-collector voltage drop at 0.4 Volts and preventing saturation. When the transistor is commanded to turn off, the

Schottky diode becomes reverse-biased and immediately ceasing conduction while the transistor's base layer returns to a non-conducting state much faster than usual because it was never fully saturated when conducting.

Figure 3 of Biard's patent shows an equivalent schematic diagram of a single transistor equipped with this "shunting" diode (numbered 22 in the diagram). Please note that this diagram contains an error in that there should be a connecting "dot" joining the diode's cathode terminal to the transistor's collector terminal:

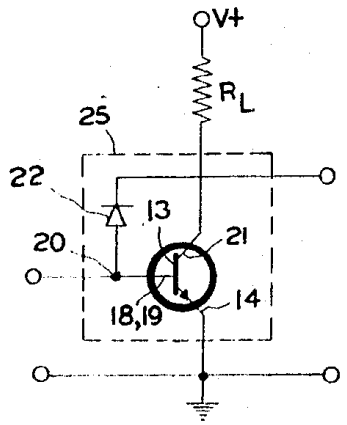


FIG. 3

Figure 1 of the patent shows the physical construction of this transistor with its auxiliary diode. This vertically-layered construction is typical of the *planar* style of transistor where each layer of material is thin and relatively wide:

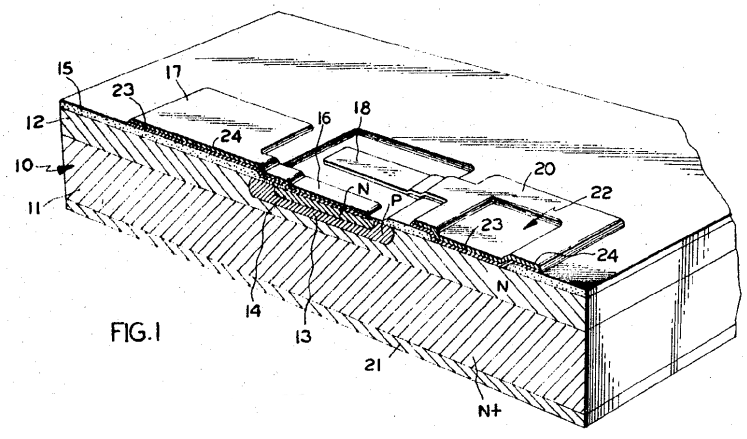
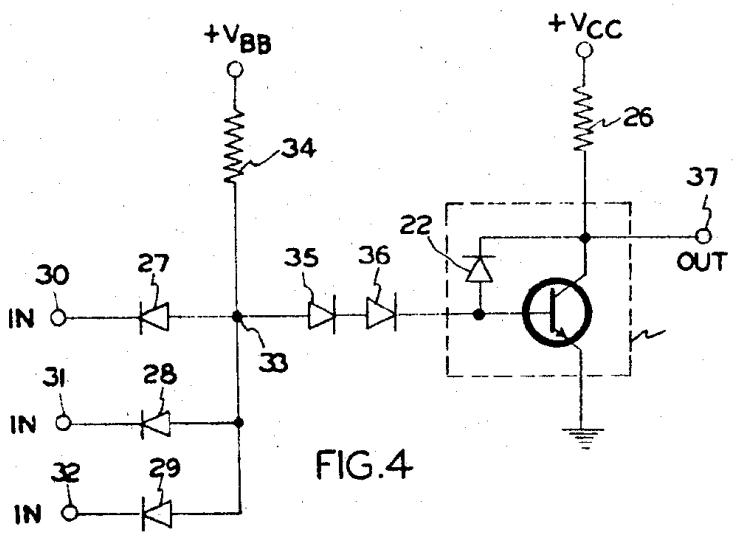


FIG. 1

As an example of how this high-speed transistor might be applied in a practical logic "gate" circuit, Biard showed a diode-transistor logic (DTL) version of a three-input NAND gate in Figure 4 of his patent:



The passive network comprised of diodes 27, 28, and 29 together with resistor 34 forms what is known as a *passive-AND* gate, where any "low" voltage input signal results in a "low" voltage state at node 33. The AND gate's signal then drives the transistor, which "inverts" that signal so that the result is a NAND function: any "low" input state results in a "high" output state, and the only way to obtain a "low" output state is if *every* input is "high".

Biard's innovation paved the way for the popular 74S00 series of Transistor-to-Transistor Logic (TTL) integrated digital circuits, where the letter "S" in the part number signified the presence of these Schottky diodes.

Chapter 5

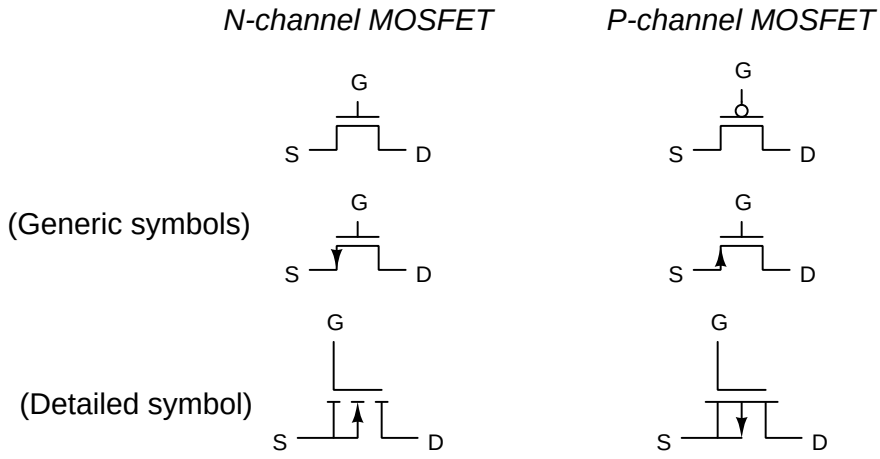
Derivations and Technical References

This chapter is where you will find mathematical derivations too detailed to include in the tutorial, and/or tables and other technical reference material.

5.1 Alternative MOSFET symbols

As it so happens, E-type MOSFETs are far more popular than D-type MOSFETs, so much so that MOSFETs are sometimes assumed to be E-type unless otherwise specified. This assumption even extends to symbols used to represent MOSFETs in some schematic diagrams. Confusingly, the E-type is so common that sometimes the broken-line convention is dropped entirely.

Shown below is a set of symbols commonly used to represent E-type MOSFETs. The "generic" symbols are intended to be simple to draw, while the "detailed" symbols are more faithful to the actual construction of the device:

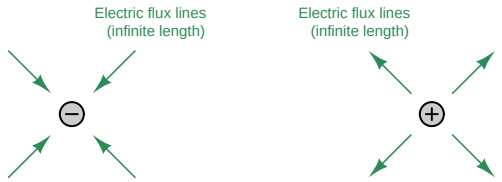


5.2 Gauss' Law of Electric Fields

One of fundamental principles relating to electric fields is *Gauss' Law of Electric Fields*. A common mathematical expression for this law is as follows:

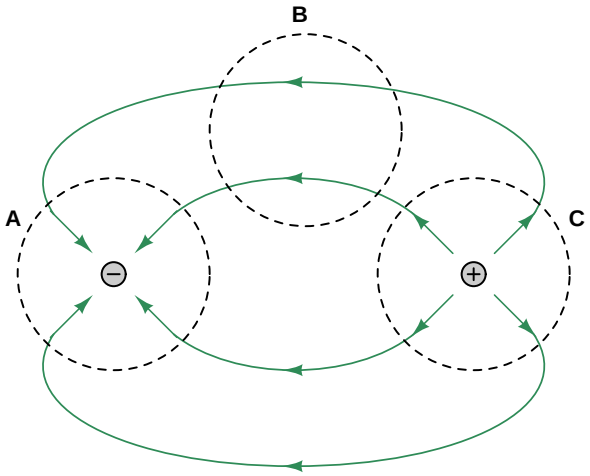
$$\oint \mathbf{E} \cdot d\mathbf{A} = \frac{Q}{\epsilon_0}$$

This equation simply describes electric fields (E) as lines of flux spreading out from a quantity of electric charge (Q). By convention we draw these electric flux lines as radiating away from positive charges and heading toward negative charges as shown below:



If we imagine a three-dimensional surface completely enclosing any electric charge, like a soap bubble surrounding that charge, the total quantity of electric flux lines passing through that surface will be the product of the electric field (also known as the electric flux density, **E**) and the enclosing area (**A**). This "soap bubble" is formally known as a *Gaussian surface*, and we may arbitrarily place that surface anywhere within space for the sake of solving an electric field problem. The major point to understand here is that *any* Gaussian surface drawn will experience an amount of electric flux penetrating its surface proportional to the amount of electric charge enclosed within that surface regardless of the size or shape.

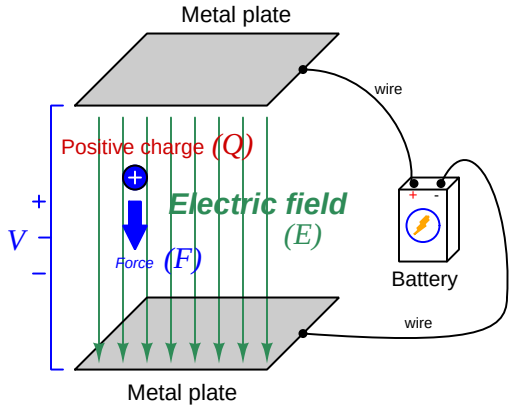
Observe what happens if we connect the electric flux lines together from two opposite charges, and then draw Gaussian surfaces in different regions of space labeled **A**, **B**, and **C**:



Gaussian surface **A** encloses a single negative charge and we see inward-pointing electric flux lines penetrating that surface. Gaussian surface **C** encloses a single positive charge and we see outward-

pointing electric flux lines penetrating that surface. Gaussian surface **B** contains no electric charges at all because it encloses empty space, and we see the number of electric flux lines entering the enclosed volume exactly equaling the number of flux lines leaving it. In other words, for Gaussian surface **B** there is zero net electric flux enclosed within it and also zero *net* penetrations since the number going in balances the number going out.

If we apply an electric field across open space containing a single positive charge, that charge will experience a force proportional to its charge-quantity and the strength of the electric field. The direction of this force will be downward to the negative plate and away from the positive as shown in the following illustration:



$$\vec{F} = Q\vec{E}$$

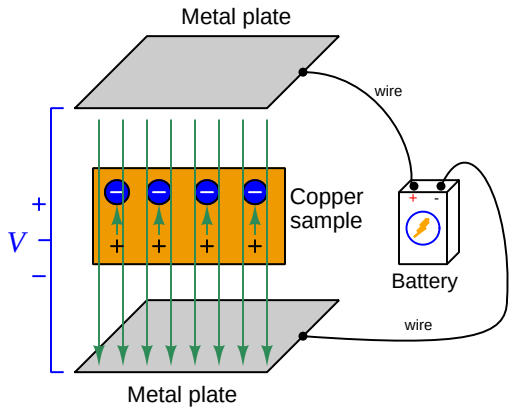
Where,

$\vec{F}$ = Force exerted on the charge (Newtons)
Q = Charge quantity (Coulombs[1])
$\vec{E}$ = Electric field (Newtons per Coulomb)

Unsurprisingly, a negative charge in that same space will experience a force in the opposite direction, upward to the positive plate and away from the negative. With nothing to impede the motion of these charges, they will accelerate as a result of the applied force from the electric field, their acceleration predicted by the familiar equation $F = ma$.

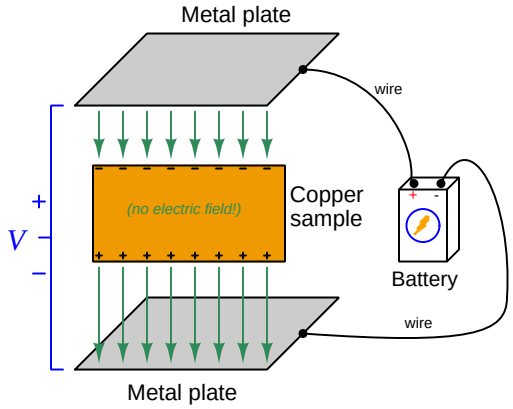
[1] One Coulomb of electric charge is equal to 6.2415×10^{18} electrons.

Now imagine a solid piece of electrically conductive material such as copper metal suspended between the same two metal plates. The outer-most electrons of the copper atoms are free to drift throughout the metal, which is the fact that makes copper metal a conductive material. The effect of the two plates' electric field on the free electrons within the copper sample will be much the same as upon free electrons in a vacuum space between two metal plates: those free electrons will accelerate upward in the direction of the positive plate and away from the negative. This migration of free electrons toward the upper surface of the copper sample leaves positively-charged regions near the lower surface of the copper:



The electrons will migrate as far as they are able, which in this case is the upper surface of the conductive copper sample. The separation of negative and positive charges within the copper sample – the negative charges being mobile electrons and the positive charges being immobile protons in the copper atom nuclei – itself forms an electric field within the copper, pointed in the direction opposite that of the battery's external field. These short vector-arrows pointed upward in the above illustration represent the flux lines from the positive stationary nuclei to the displaced negative electrons.

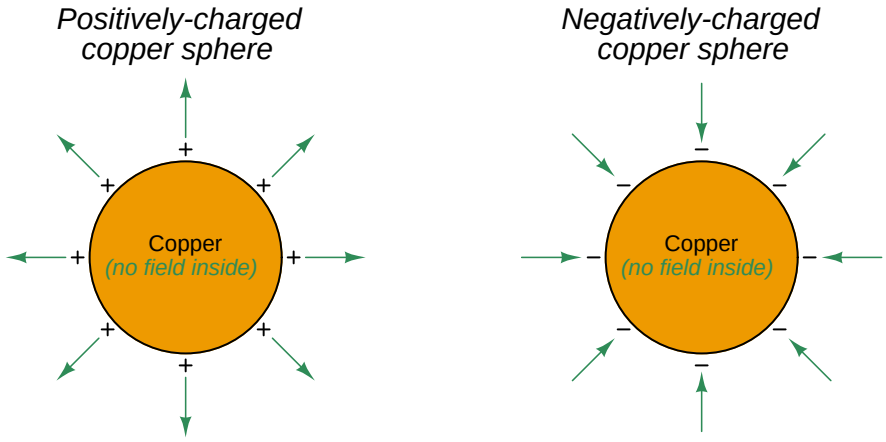
Since these two sets of electric flux lines point in opposite directions, their net effect is to cancel each other out. As more electrons drift upward this cancellation continues until there is zero net electric field within the copper sample. When the net electric field reaches zero, the electrons no longer experience any driving force upward. In this *equilibrium* condition of zero net force on all electric charges, the interior of the copper sample contains no net electric field:



It should not come as a surprise to anyone familiar with basic concepts of electric circuits that this sample of copper naturally excludes any electric field imposed externally upon it, since the mass of copper may be thought of as a large collection of electrically common points, and we know one of the characteristics of electrically common points is equipotentiality. *And, where there is no difference of electrical potential between two points (i.e. no voltage), there can be no electric field between those points.* A conductive object in a state of electrical equilibrium, therefore, contains no internal electric field(s).

This phenomenon of electrical conductors naturally cancelling out internal electric fields finds wide application in electrical and electronic systems. One such application is the *shielding* of sensitive components by surrounding them with a continuous conductive surface. Another application is in insulated-gate field-effect transistors (also known as MOSFETs) where an externally-applied electric field causes electron and hole charge carriers to reposition themselves and in so doing cause a conductive channel transverse to that field to either grow wider or narrower.

An interesting consequence of Gauss' Law of Electric Fields is that any net gain or loss of electric charge by a conductive object results in the unbalanced charges moving to the very exterior surface of the object. For example, if we inject extra electrons into a copper block, those electrons quickly settle into positions on the very outer skin of that block where they remain. Likewise, if electrons are extracted from a copper block, the remaining electrons will migrate within the block so as to leave the resulting positive charges (i.e. electron deficiencies) at the outer surface of that block:



All unbalanced charges migrate to the outside of the object

We may understand this by appealing to the proven fact that a conductive object in a state of electrical equilibrium must have no electric field within its interior. If we imagine drawing a Gaussian surface just beneath the skin of that conductive object, we may conclude from Gauss' Law that the interior of that Gaussian surface must contain no net electric charge (i.e. all positive and negative charges within it are perfectly balanced) because zero net electric field penetrating that Gaussian surface implies zero net electric charge within. This must mean that any and all unbalanced electric charges in that conductive object must reside *outside* the area of no electric field; i.e. outside the Gaussian surface on the conductive object's outer surface. Thus, a Gaussian surface drawn anywhere inside a charged conductive object experiences no electric flux lines but a Gaussian surface drawn outside that same object does.

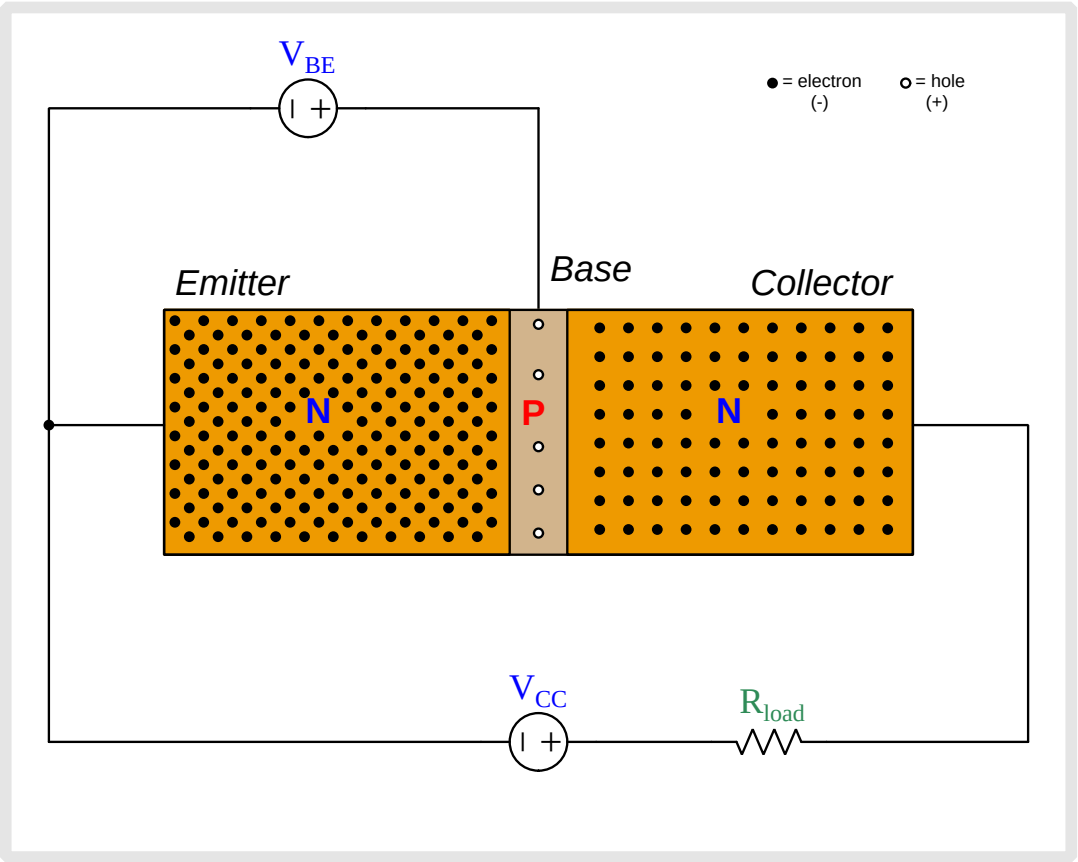
Chapter 6

Animations

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

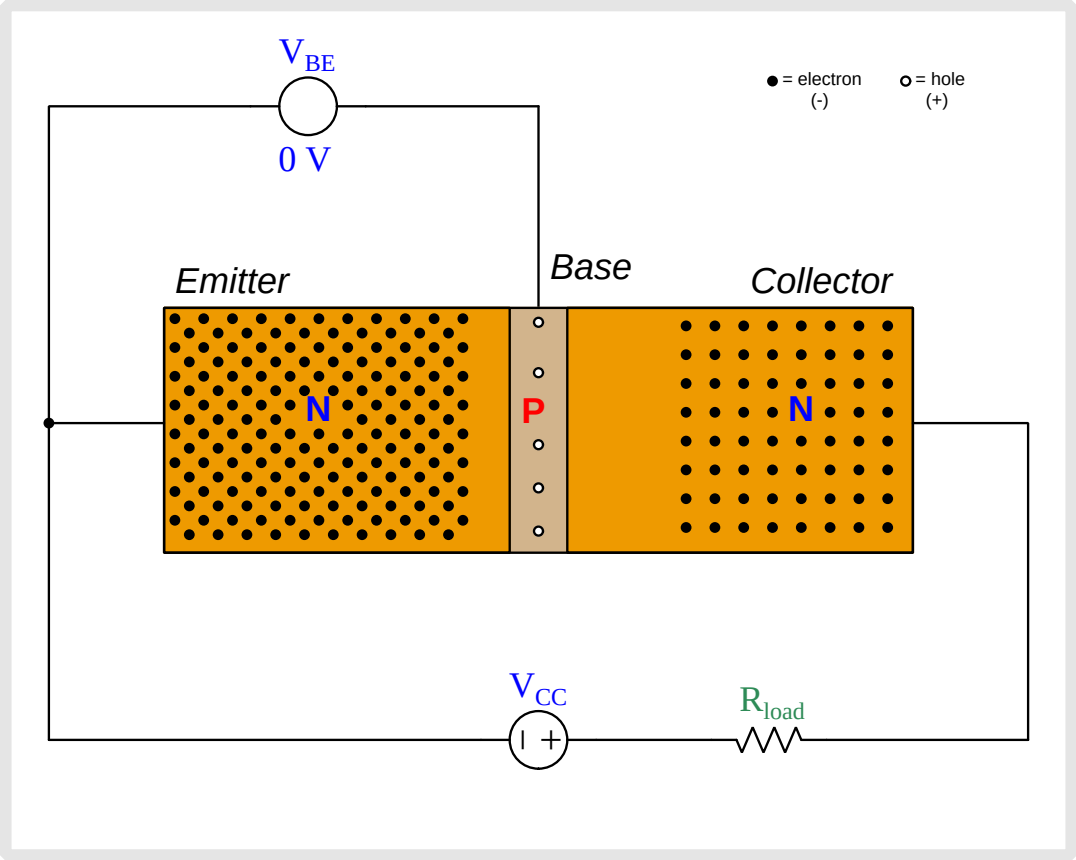
6.1 Animation of an NPN BJT turning on

Bipolar junction transistors are naturally "off" devices, having no continuous channel of conductive semiconductor material joining emitter and collector terminals. Instead the NPN or PNP sandwich interrupts that path with two series-opposed PN junctions. However, when the base-emitter PN junction is forward-biased by an external source, majority charge carriers from the emitter region are drawn into the thin and lightly-doped base region where some of them travel out the base but most carry past into the depletion region of the base-collector PN junction to be swept into the collector region. For an NPN transistor this means electrons from the N-type emitter region will be injected into the P-type base region (as "minority" type carriers there) before most of them get swept into the N-type collector region.

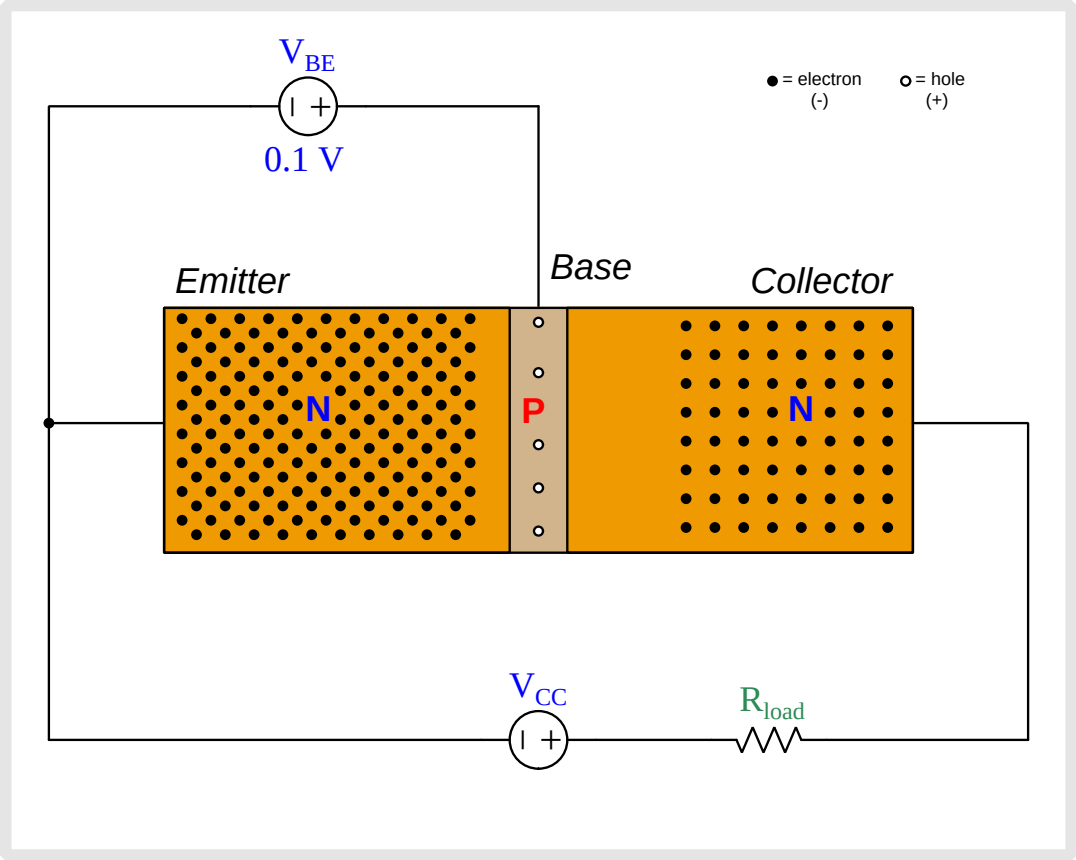


Above we see an NPN bipolar junction transistor connected to two sources. V_{BE} applies a controlling voltage between the base and emitter. V_{CC} is the voltage source which will drive current through the load resistor once the transistor begins to conduct.

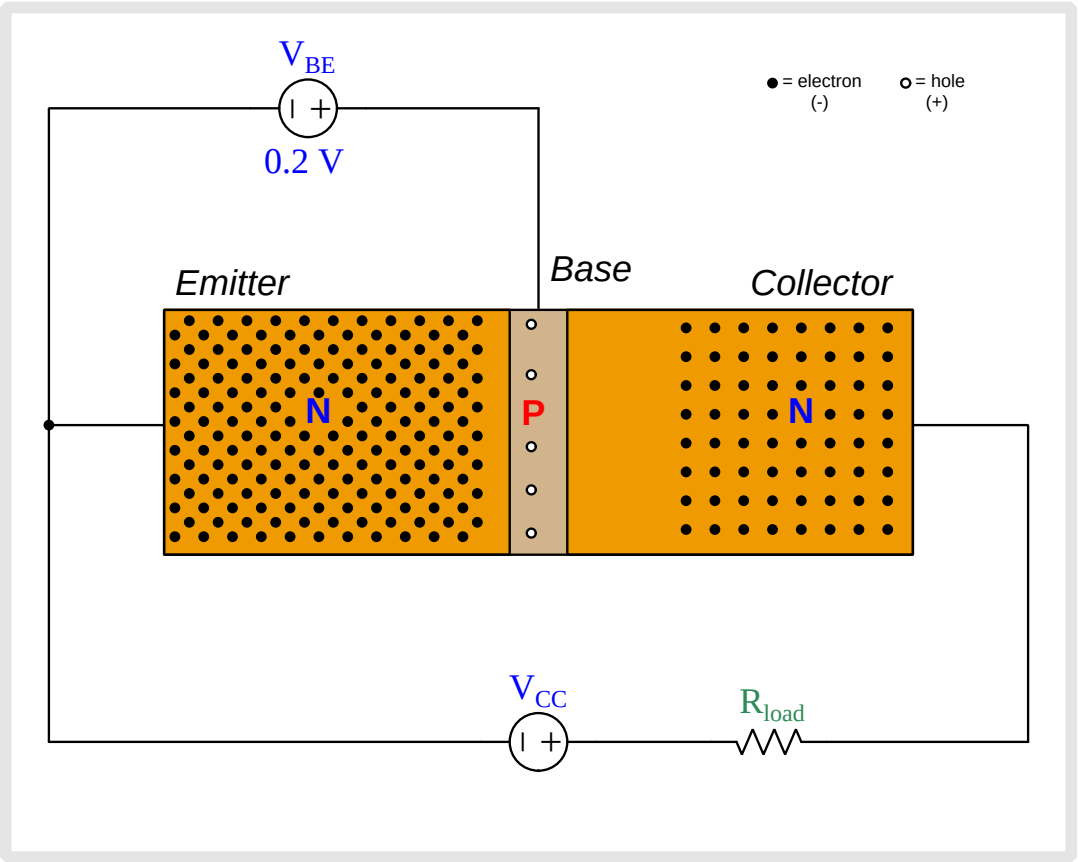
The N-doped emitter region is heavily doped and therefore contains many free electrons (negatively-charged) as mobile charge carriers. The collector region is also doped as an N-type semiconductor but the dopant concentration is less than that of the emitter. The P-doped base region is the least-doped of them all with just a few holes (positively-charged) as mobile charge carriers. Note that when load current begins to flow, its direction will be shown in *conventional flow notation* which is the same direction in which holes move but is opposite to the direction free electrons move.



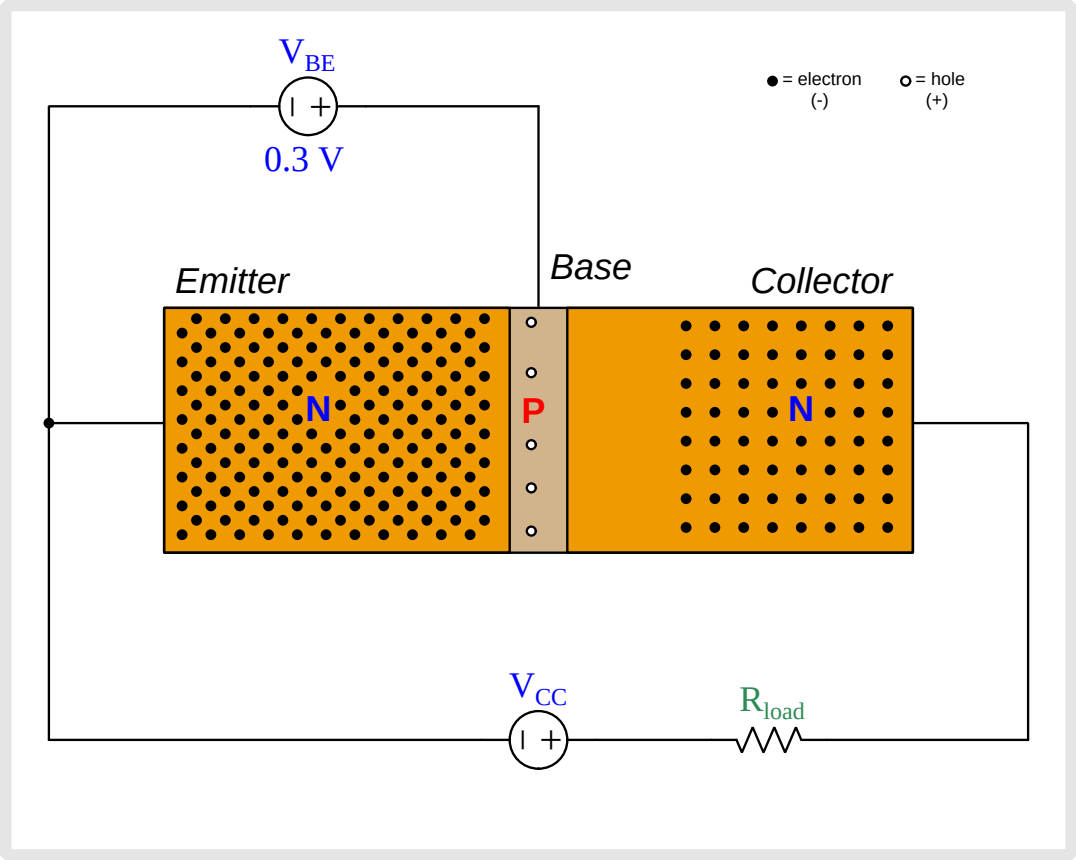
Note the depletion regions shown as areas with no free charge carriers, the base-collector depletion region being wider than the base-emitter depletion region due to the applied reverse-bias potential of V_{CC}.



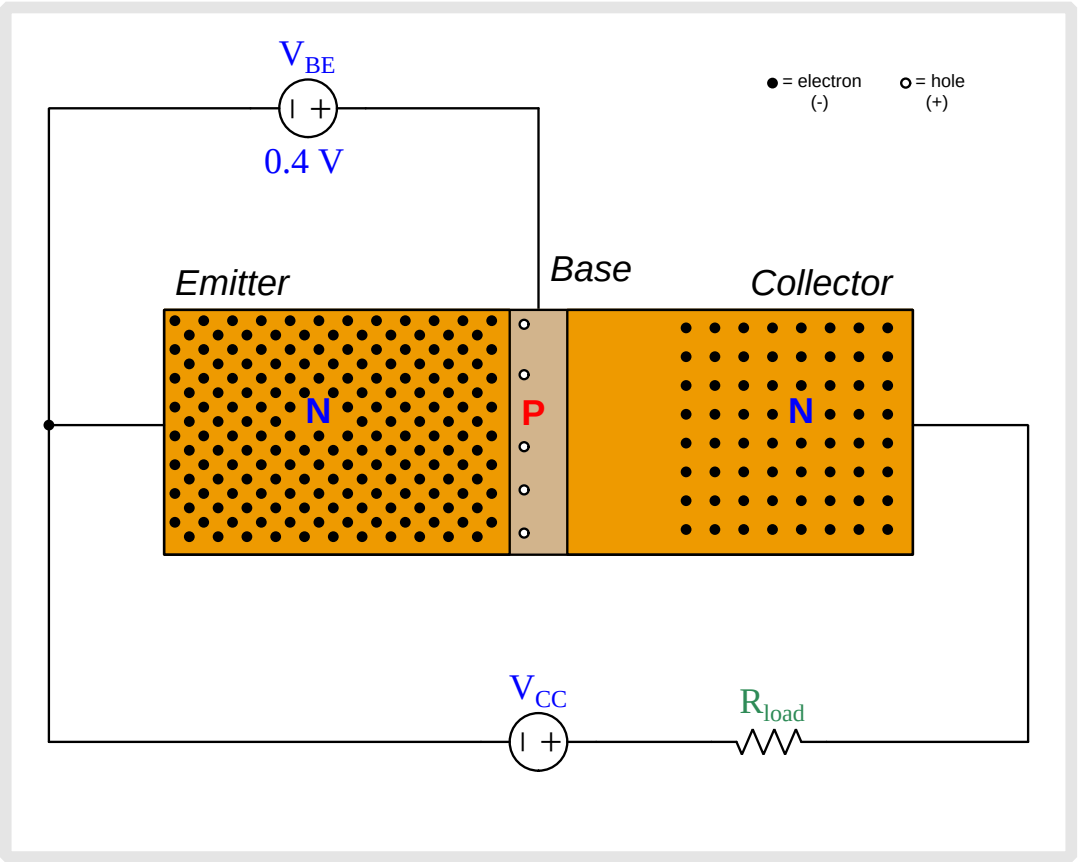
As we begin to apply a base-to-emitter voltage with the base positive and emitter negative, that applied potential difference drives the emitter and base regions' respective charge carriers toward each other, collapsing the base-emitter depletion region.



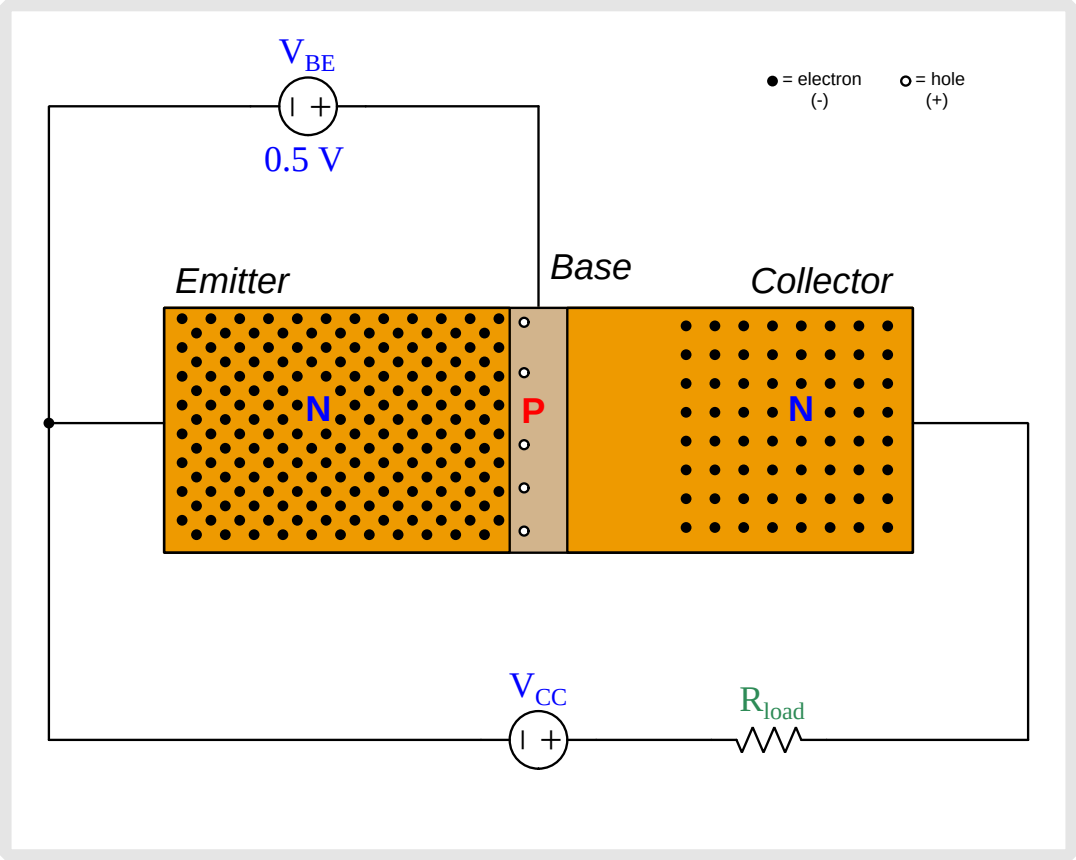
As we begin to apply a base-to-emitter voltage with the base positive and emitter negative, that applied potential difference drives the emitter and base regions' respective charge carriers toward each other, collapsing the base-emitter depletion region.



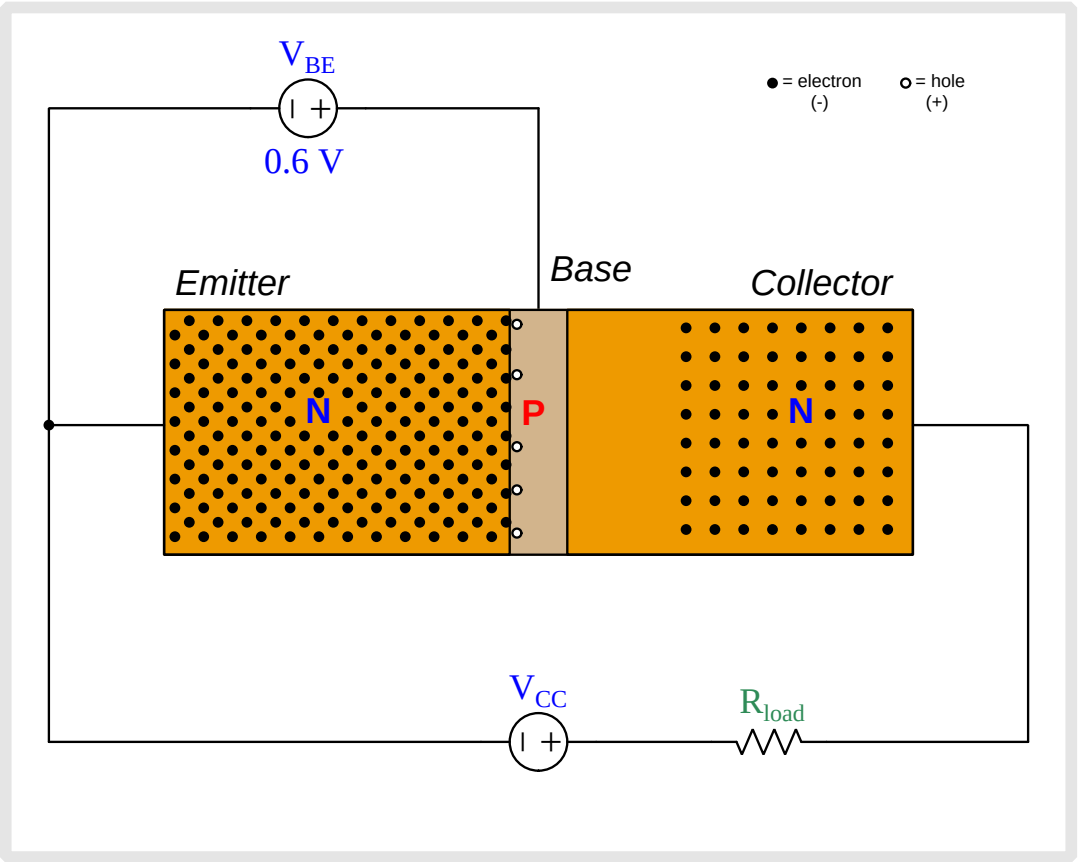
As we begin to apply a base-to-emitter voltage with the base positive and emitter negative, that applied potential difference drives the emitter and base regions' respective charge carriers toward each other, collapsing the base-emitter depletion region.



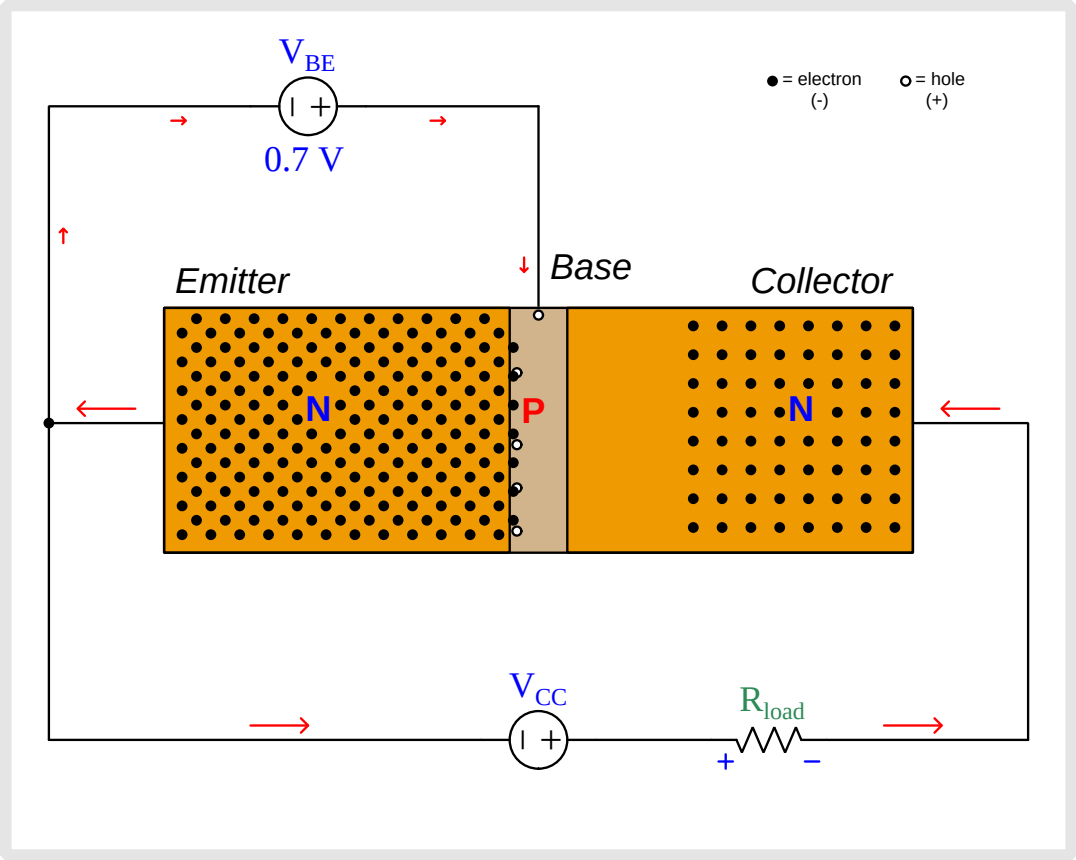
As we begin to apply a base-to-emitter voltage with the base positive and emitter negative, that applied potential difference drives the emitter and base regions' respective charge carriers toward each other, collapsing the base-emitter depletion region.



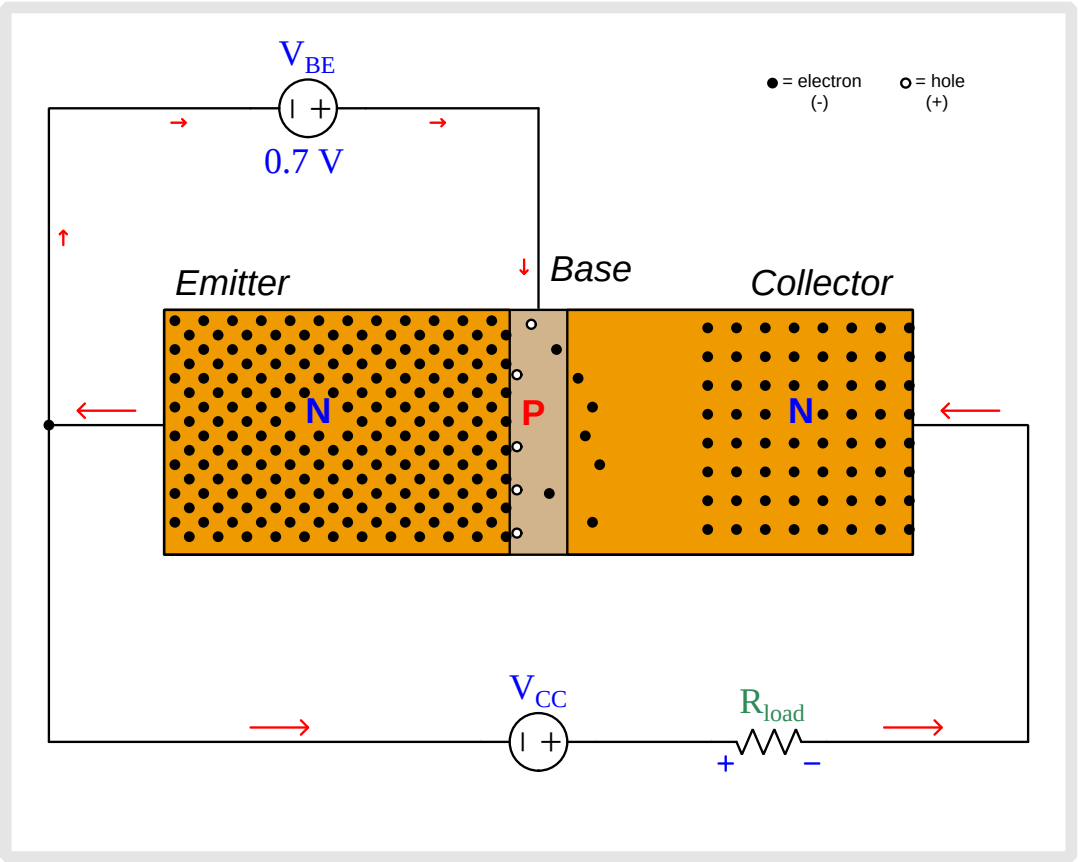
As we begin to apply a base-to-emitter voltage with the base positive and emitter negative, that applied potential difference drives the emitter and base regions' respective charge carriers toward each other, collapsing the base-emitter depletion region.



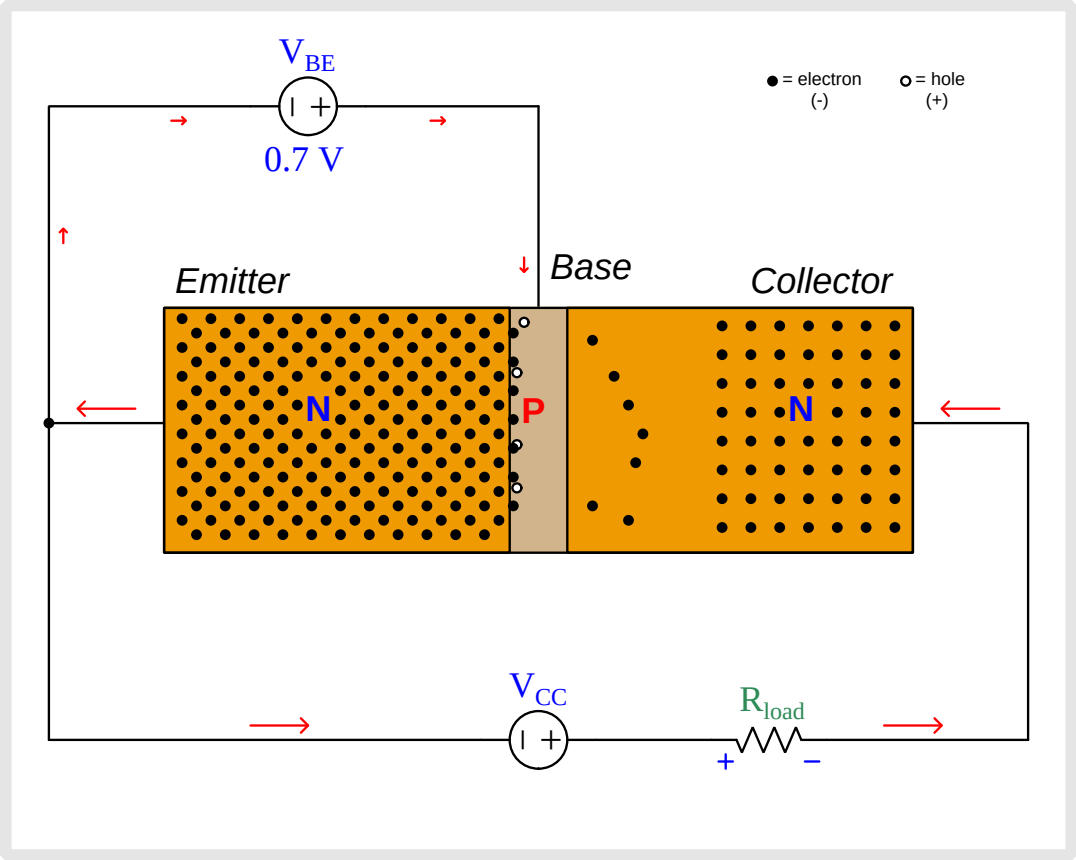
As we begin to apply a base-to-emitter voltage with the base positive and emitter negative, that applied potential difference drives the emitter and base regions' respective charge carriers toward each other, collapsing the base-emitter depletion region.



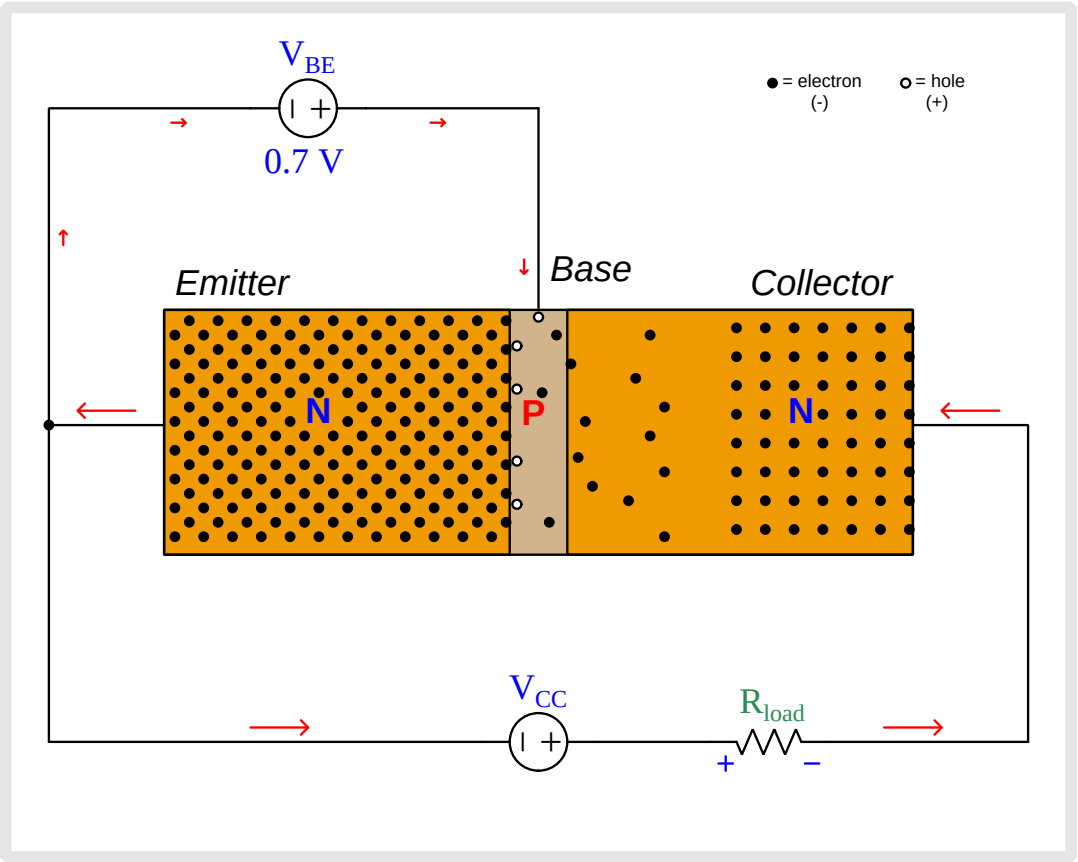
Now there is enough base-to-emitter voltage to fully collapse the base-emitter depletion region which allows base current to flow through the V_{BE} source. New holes enter into the base region from the base terminal to replenish any holes that have "re-combined" with electrons from the emitter region. Most of the emitter-injected electrons do *not* combine with holes, though, because the lightly-doped base region provides few holes to combine with. Therefore, most of the electrons injected into the base region from the emitter region begin to drift toward the collector.



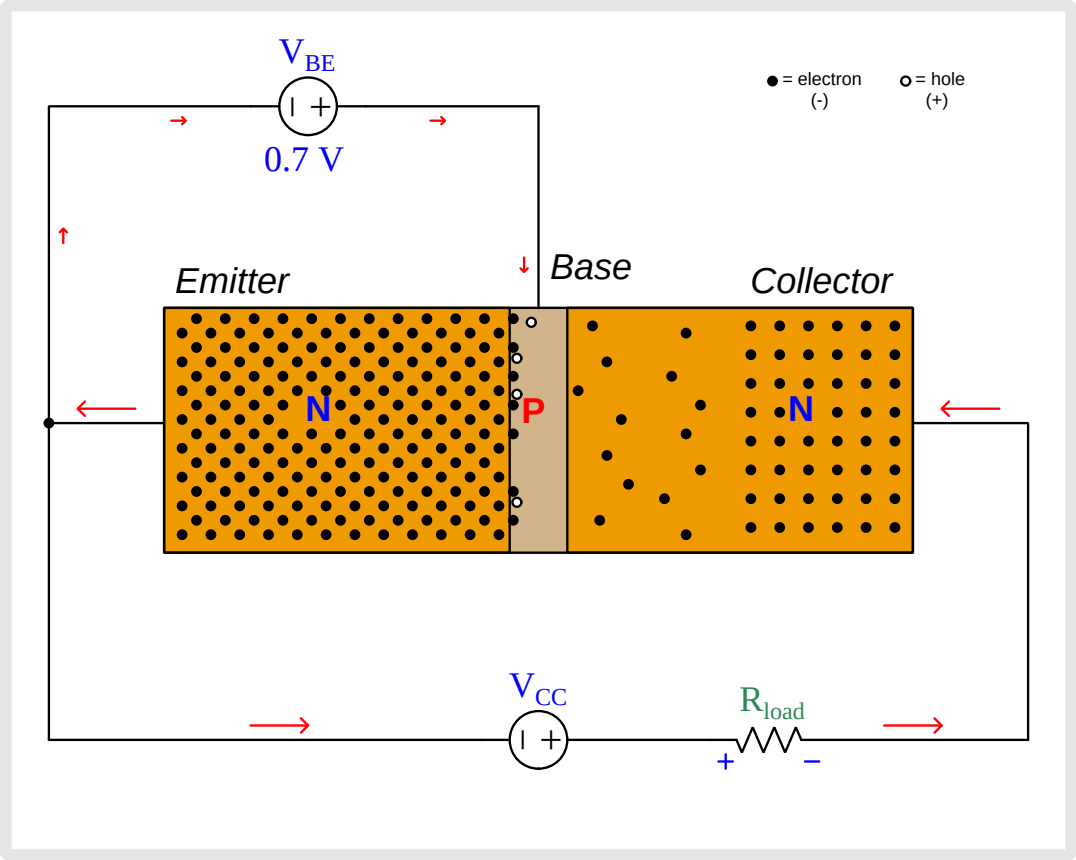
The electric field imposed by V_{CC} across the base-collector PN junction quickly "sweeps" electrons toward the collector terminal, constituting current through the load resistor. A small number of electrons injected into the base "recombined" with holes there to constitute base current rather than collector current.



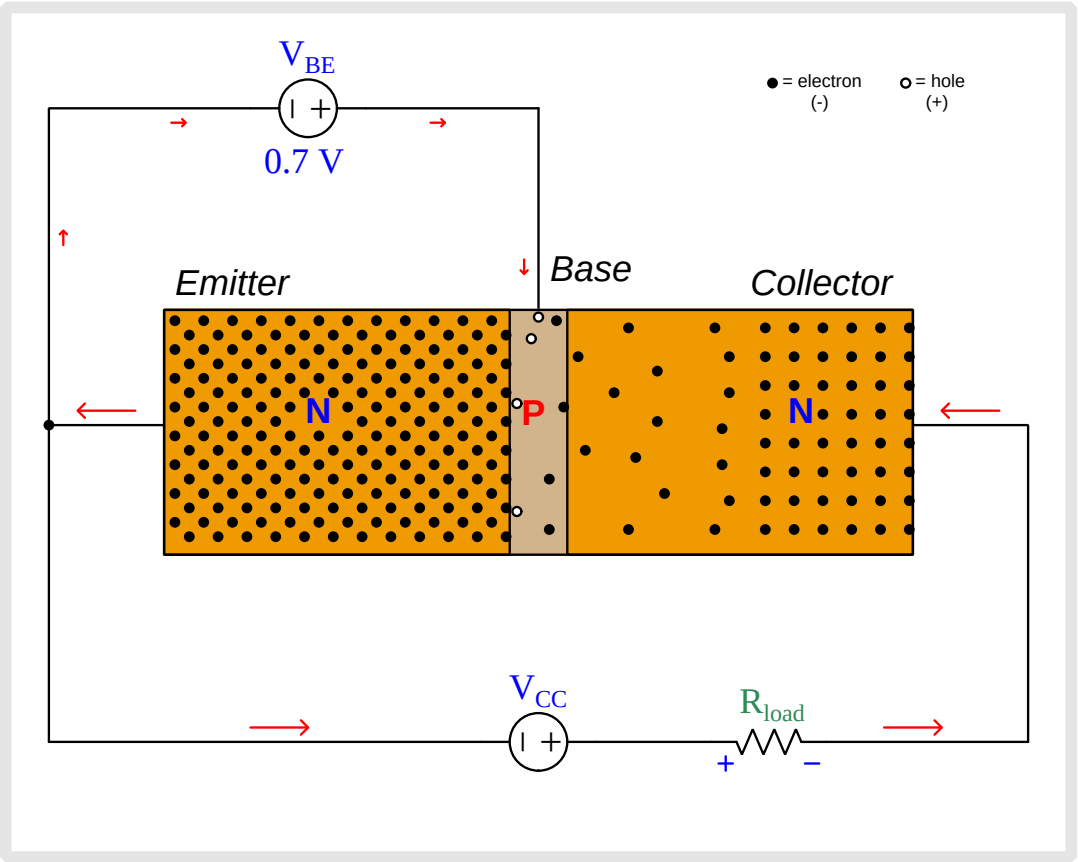
The electric field imposed by V_{CC} across the base-collector PN junction quickly "sweeps" electrons toward the collector terminal, constituting current through the load resistor. A small number of electrons injected into the base "recombined" with holes there to constitute base current rather than collector current.



The electric field imposed by V_{CC} across the base-collector PN junction quickly "sweeps" electrons toward the collector terminal, constituting current through the load resistor. A small number of electrons injected into the base "recombined" with holes there to constitute base current rather than collector current.



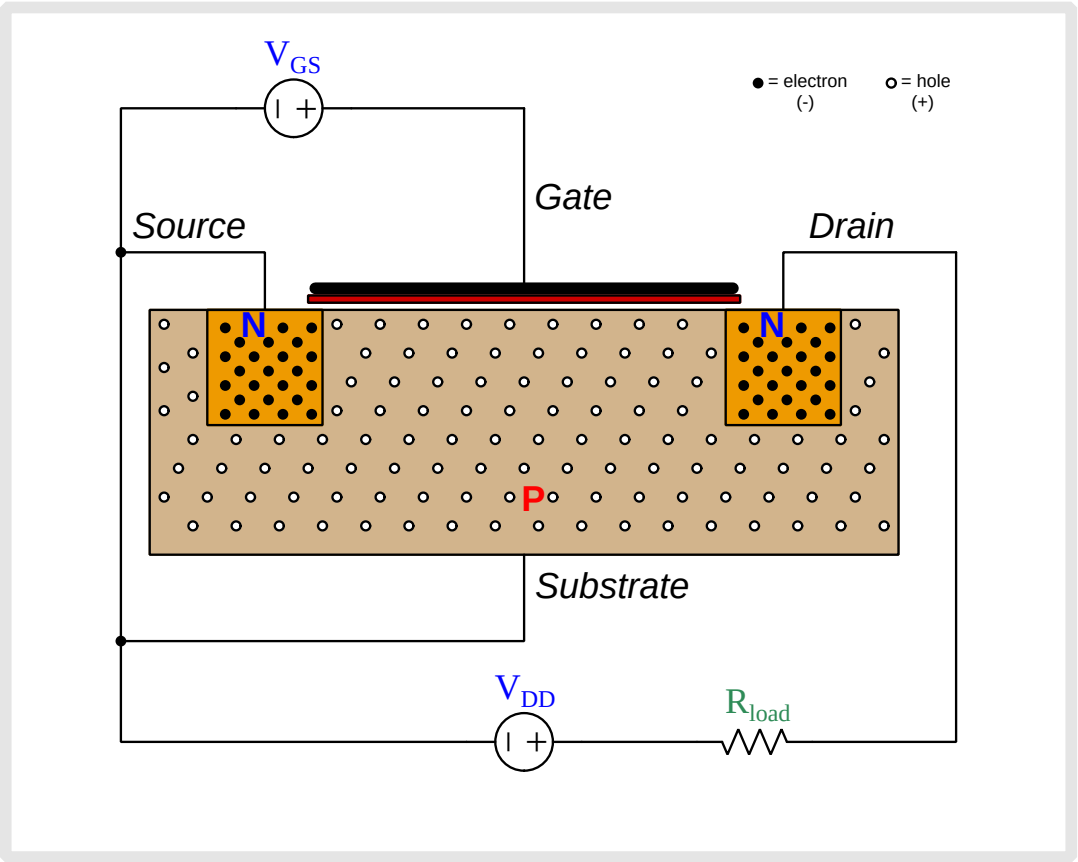
The electric field imposed by V_{CC} across the base-collector PN junction quickly "sweeps" electrons toward the collector terminal, constituting current through the load resistor. A small number of electrons injected into the base "recombined" with holes there to constitute base current rather than collector current.



The electric field imposed by V_{CC} across the base-collector PN junction quickly "sweeps" electrons toward the collector terminal, constituting current through the load resistor. A small number of electrons injected into the base "recombined" with holes there to constitute base current rather than collector current.

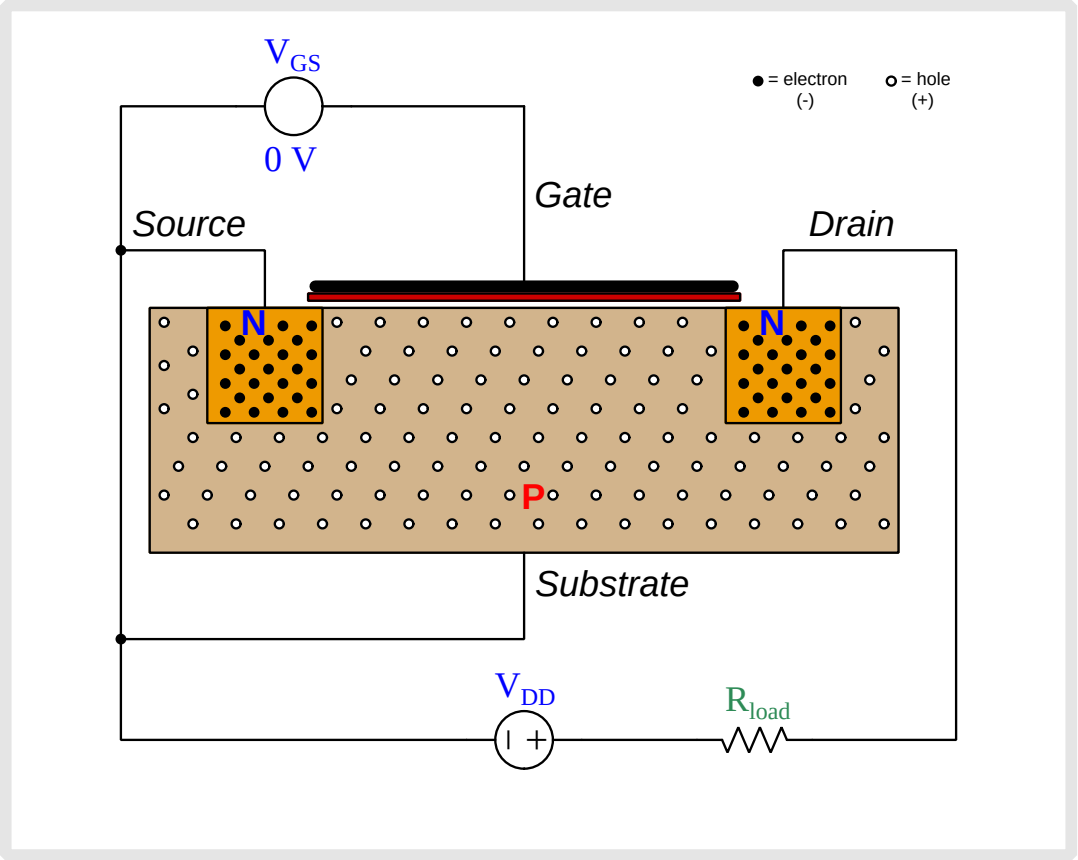
6.2 Animation of an E-type MOSFET turning on

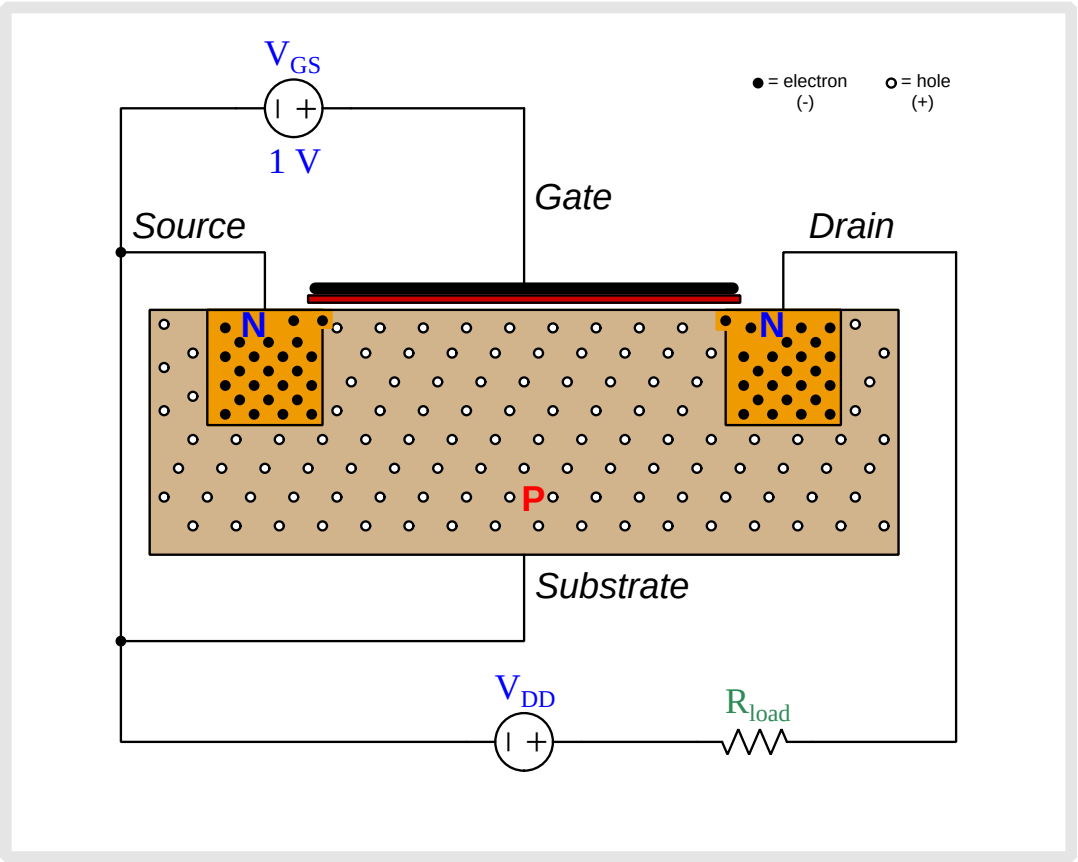
Enhancement-type (E-type) MOSFET transistors are naturally "off" devices, having no continuous channel of conductive semiconductor material joining source and drain by default the way D-type MOSFETs do. Instead the substrate semiconductor material, which is doped oppositely of the source and drain regions, interrupts that path with two series-opposed PN junctions. However, when a transverse electric field is applied between gate and substrate by an external voltage source, the surface-layer of the substrate in direct contact with the insulating metal-oxide layer converts to the opposite type of semiconductor. For an N-channel MOSFET with N-doped source and drain regions and with P-doped substrate, this means the surface of the P-doped substrate just beneath the gate effectively turns into N-type by the force of the electric field drawing free electrons from source and drain (and a few minority electrons existing within the substrate) to that surface region to form a thin film of free electrons capable of supporting source-drain conduction.



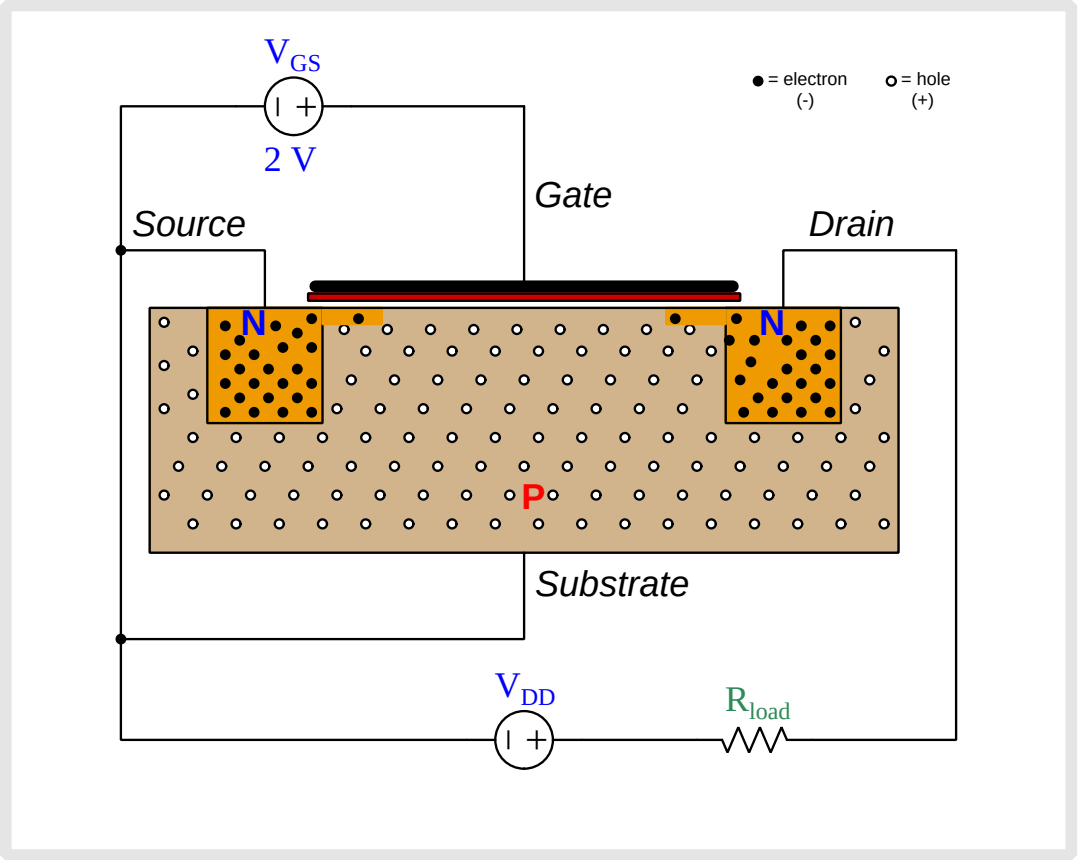
Above we see an N-channel, E-type MOSFET connected to two sources. V_{GS} applies a controlling voltage between the gate and the substrate, with the substrate connected to the source terminal to be one and the same. V_{DD} is the voltage source which will drive current through the load resistor once the transistor begins to conduct.

The N-doped source and drain regions have numerous free electrons (negatively-charged) as mobile charge carriers, while the P-doped substrate has holes (positively-charged). Note that when load current begins to flow, its direction will be shown in *conventional flow notation* which is the same direction in which holes move but is opposite to the direction free electrons move.

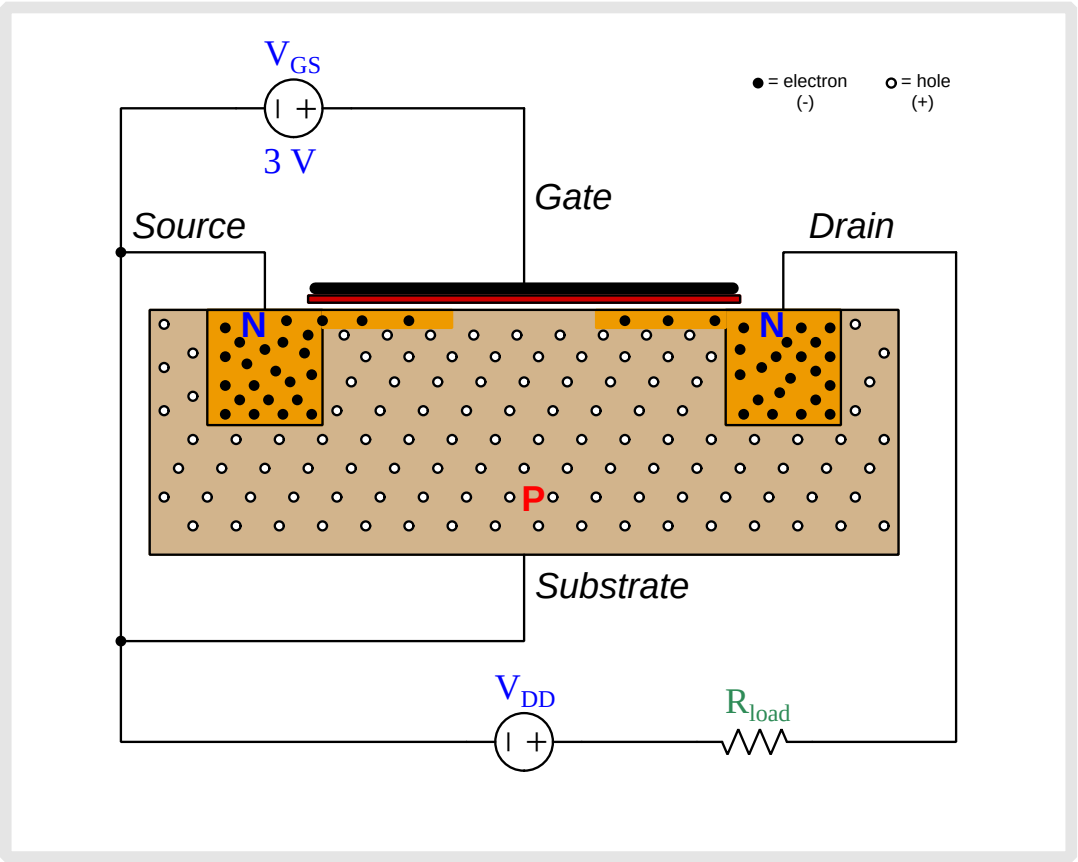

V_{GS}
0 V
● = electron (-)
○ = hole (+)
Gate
Source
Drain
N
N
P
Substrate
V_{DD}
R_{load}



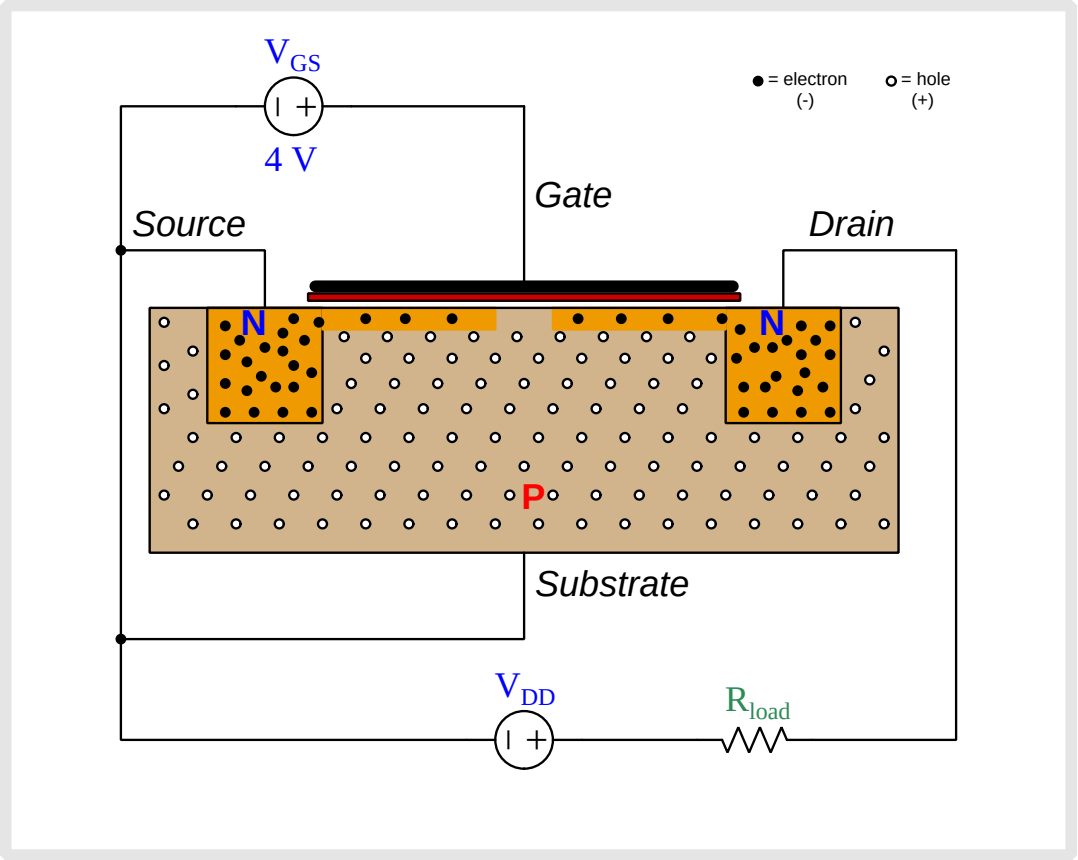
As we begin to apply a gate-to-substrate voltage with the gate positive and substrate negative, negatively-charged free electrons from the source and drain regions are attracted toward the metal gate while positively-charged holes in the substrate are repelled from the gate and driven downward.



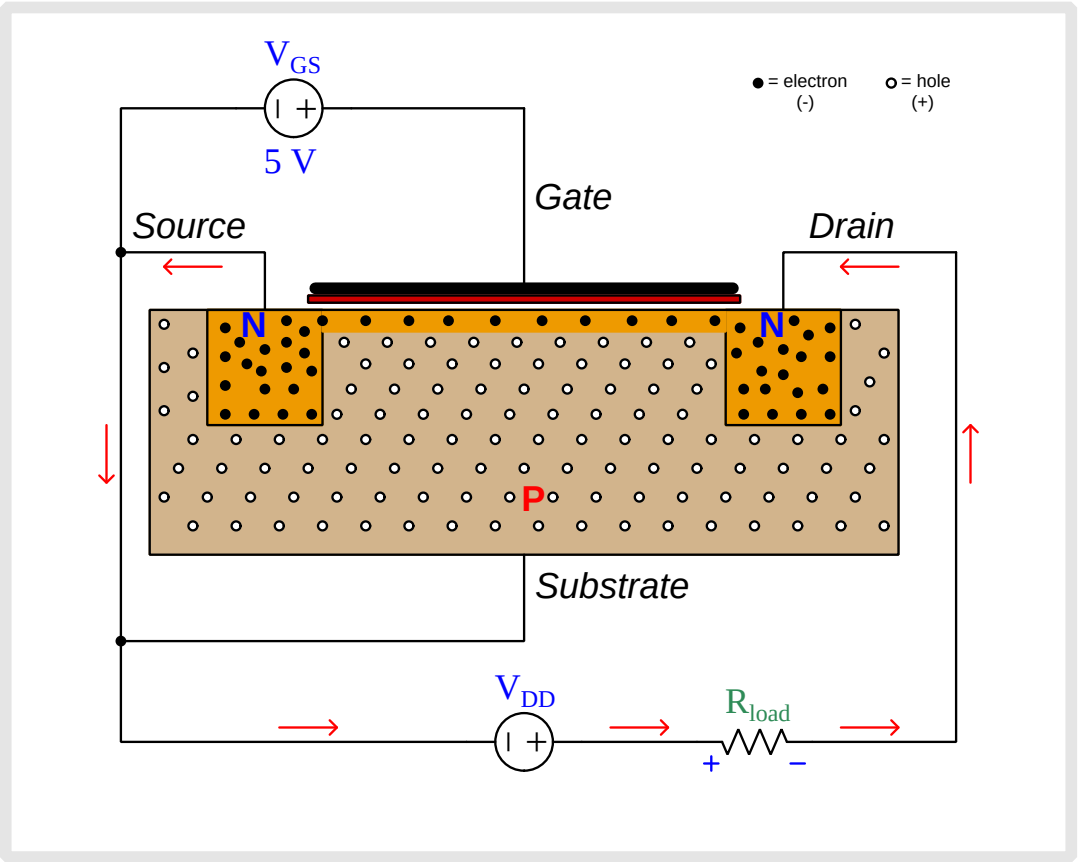
As this V_{GS} voltage increases, a thin film of free electrons begins to form beneath the gate, effectively turning what used to be P-type semiconductor material into N-type. This is called an *inversion layer*.



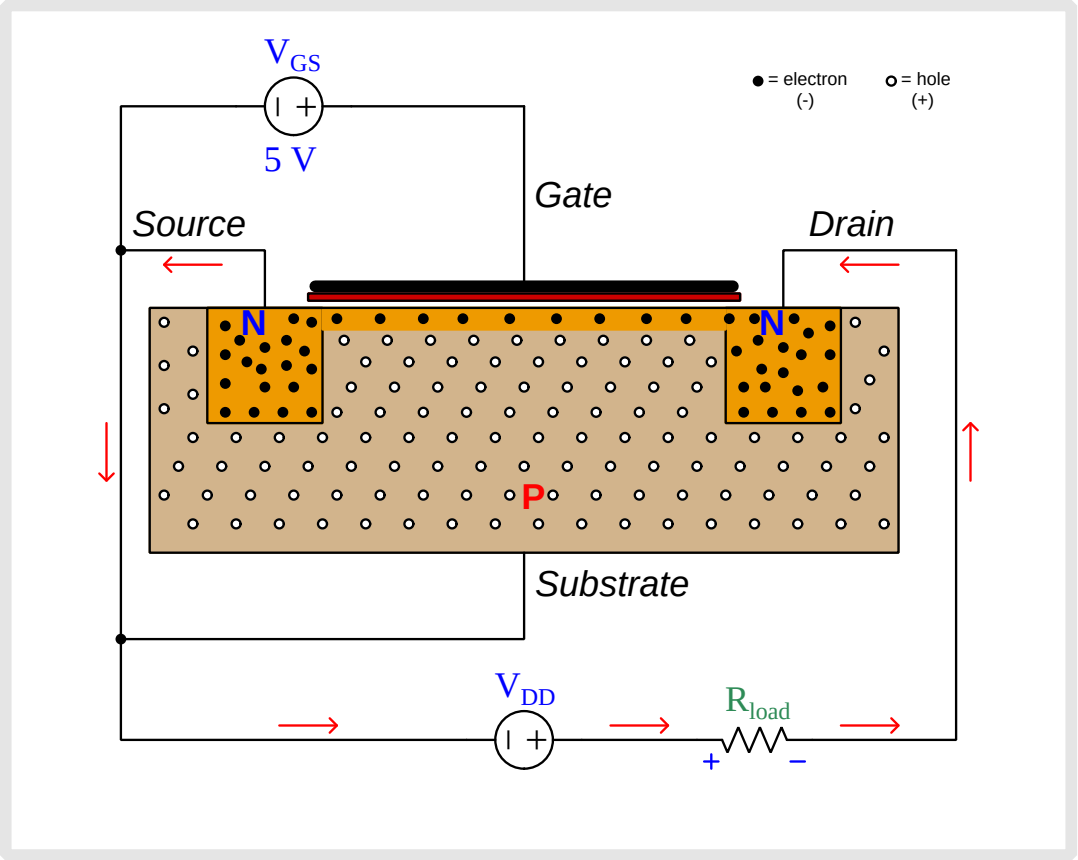
As this V_{GS} voltage increases, a thin film of free electrons begins to form beneath the gate, effectively turning what used to be P-type semiconductor material into N-type. This is called an *inversion layer*.



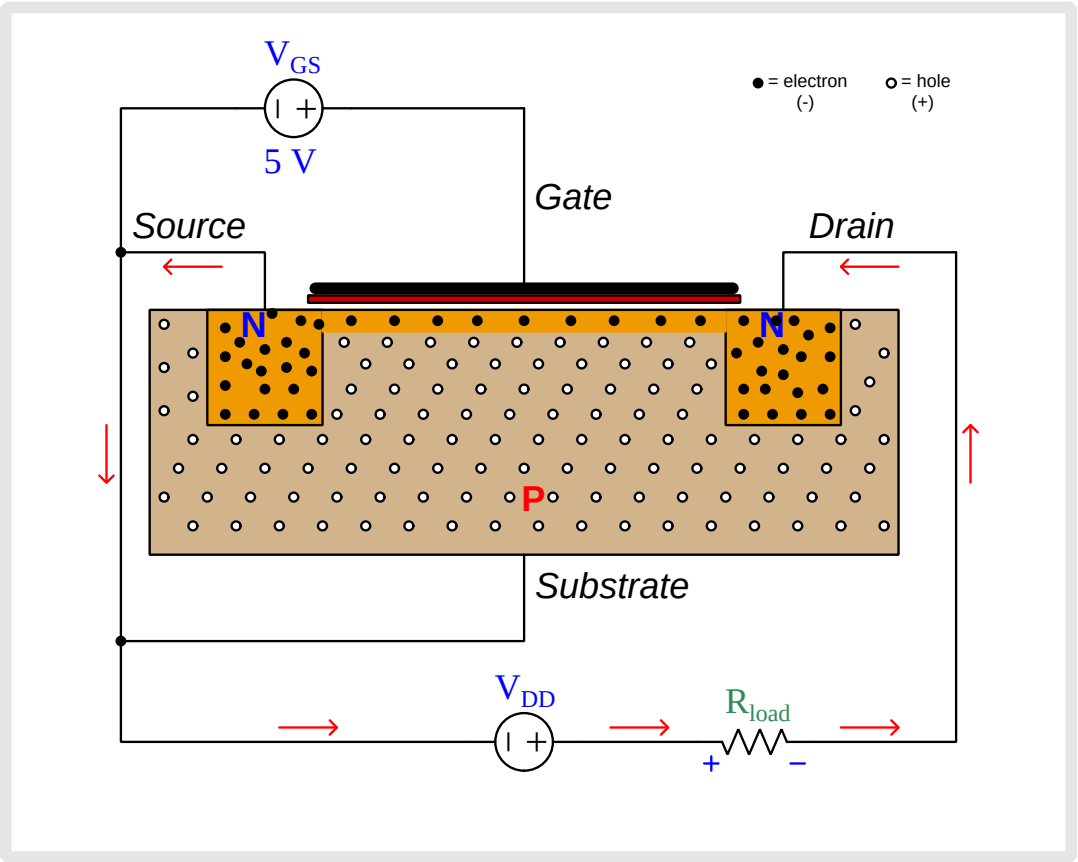
As this V_{GS} voltage increases, a thin film of free electrons begins to form beneath the gate, effectively turning what used to be P-type semiconductor material into N-type. This is called an *inversion layer*.



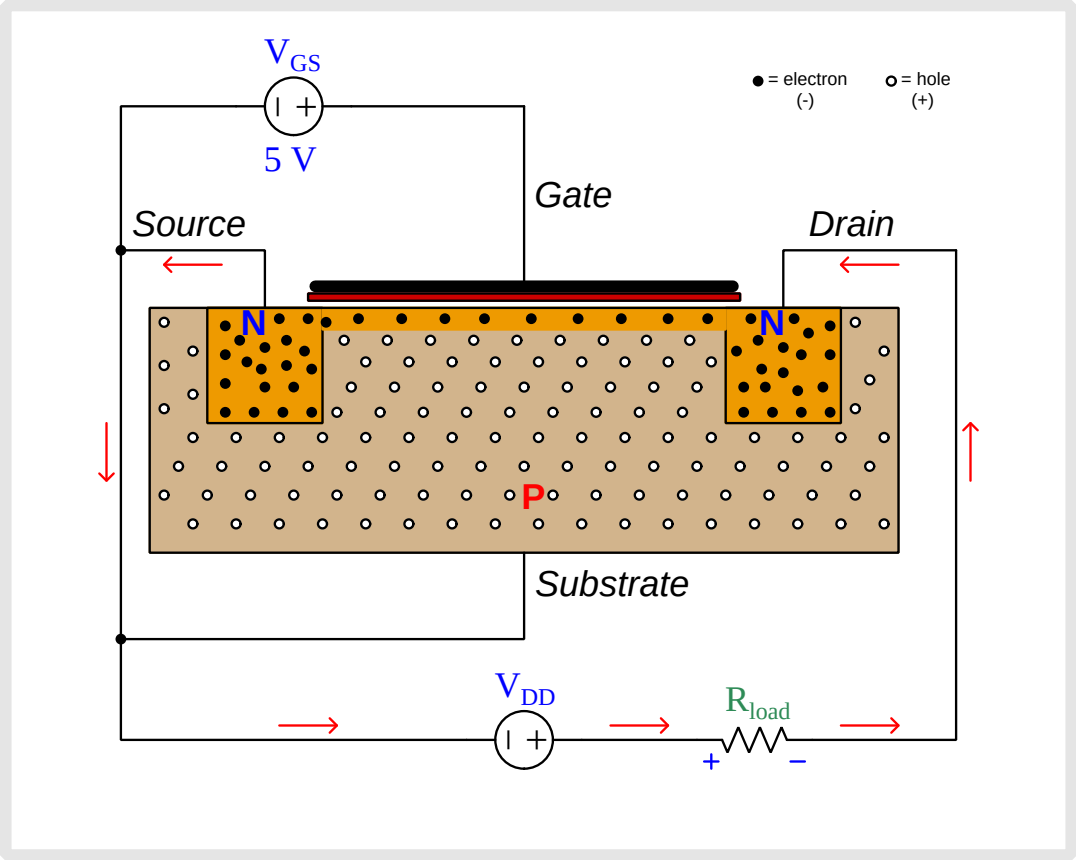
Once V_{GS} reaches the threshold or "turn-on" level, this inversion layer completely connects drain and source to form a new N-type channel permitting conduction between source and drain. The V_{DD} source is now able to promote a current through the MOSFET.



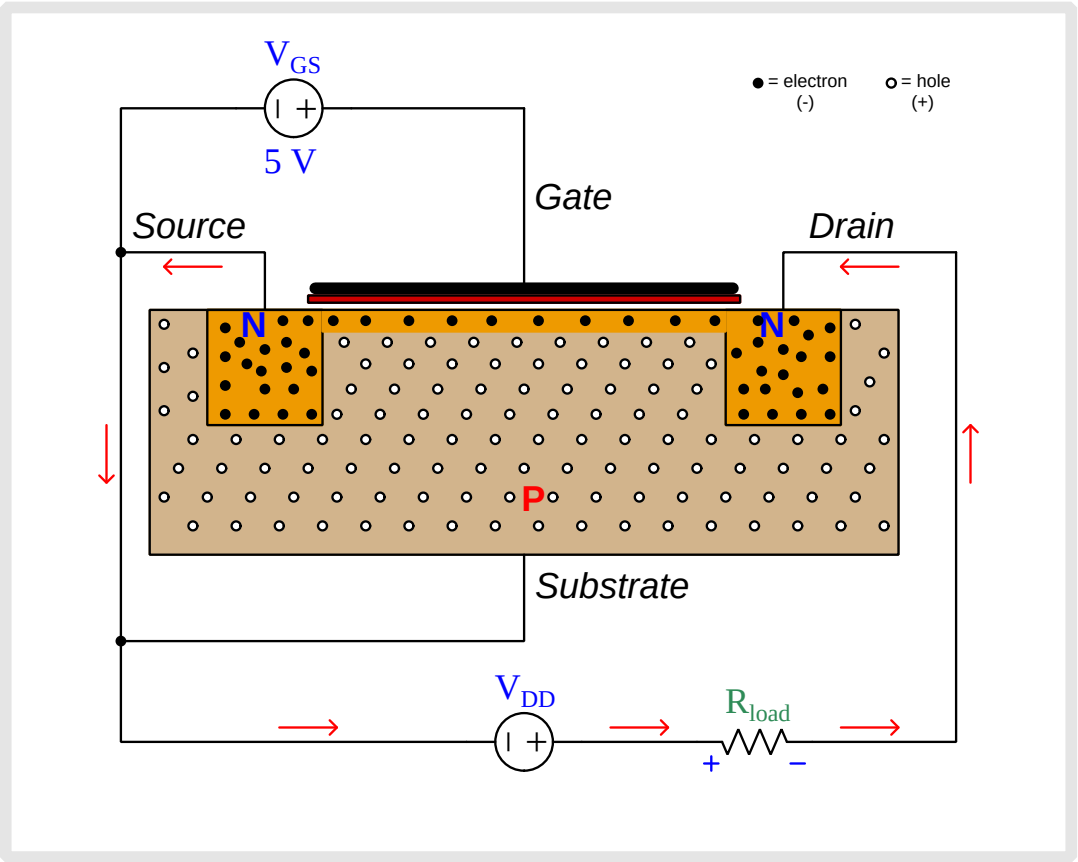
Note how the mobile charge carriers within the inversion layer are electrons, which flow opposite the direction of "conventional flow" as notated by the red arrows. Notice also how new electrons enter the inversion layer from the source region of the MOSFET, which is why that region is so-named (because that is the "source" of fresh charge carriers entering the inversion layer).



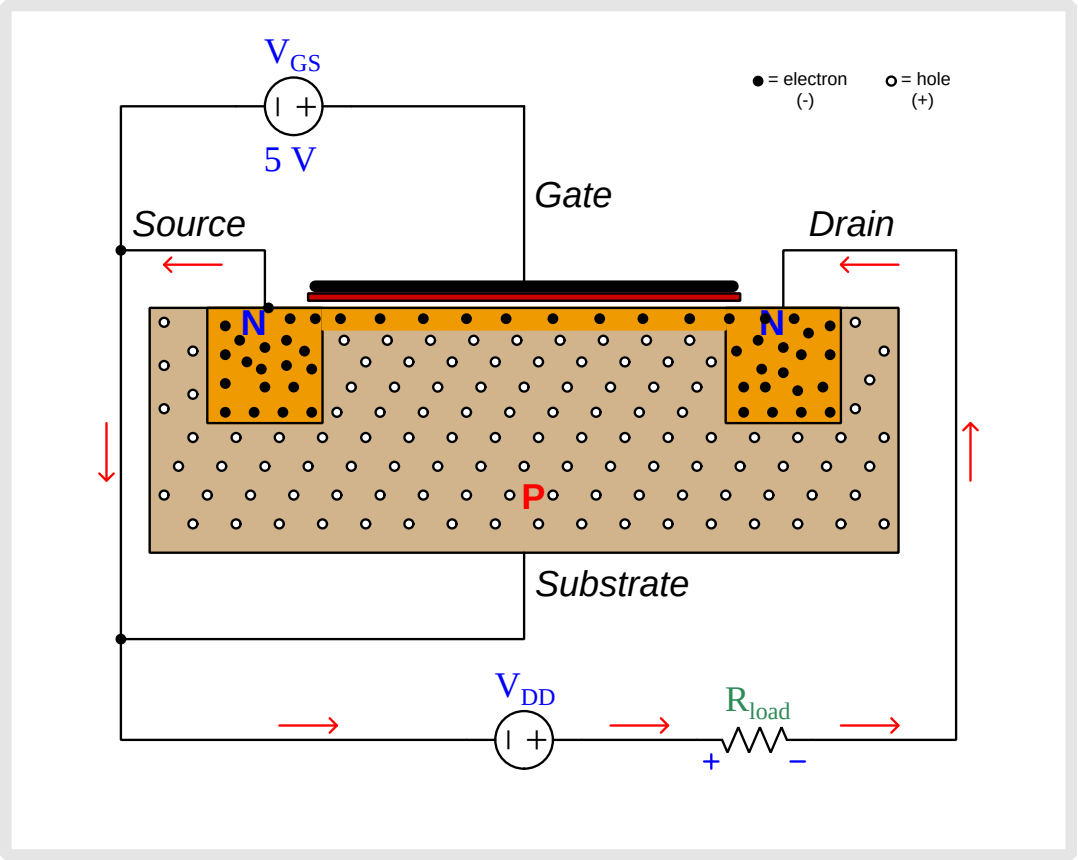
Note how the mobile charge carriers within the inversion layer are electrons, which flow opposite the direction of "conventional flow" as notated by the red arrows. Notice also how new electrons enter the inversion layer from the source region of the MOSFET, which is why that region is so-named (because that is the "source" of fresh charge carriers entering the inversion layer).



Note how the mobile charge carriers within the inversion layer are electrons, which flow opposite the direction of "conventional flow" as notated by the red arrows. Notice also how new electrons enter the inversion layer from the source region of the MOSFET, which is why that region is so-named (because that is the "source" of fresh charge carriers entering the inversion layer).



Note how the mobile charge carriers within the inversion layer are electrons, which flow opposite the direction of "conventional flow" as notated by the red arrows. Notice also how new electrons enter the inversion layer from the source region of the MOSFET, which is why that region is so-named (because that is the "source" of fresh charge carriers entering the inversion layer).



Note how the mobile charge carriers within the inversion layer are electrons, which flow opposite the direction of "conventional flow" as notated by the red arrows. Notice also how new electrons enter the inversion layer from the source region of the MOSFET, which is why that region is so-named (because that is the "source" of fresh charge carriers entering the inversion layer).

Chapter 7

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?

7.1 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.

7.1.1 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.

7.1.2 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.

Doping

N-type semiconductor

P-type semiconductor

Depletion region

BJT principles

FET principles

Depletion mode

Enhancement mode

Temperature effect on transistor action

High-side versus Low-side switching

Sourcing versus Sinking

Saturation

Threshold voltage

Conservation of Energy

Electrical source

Electrical load

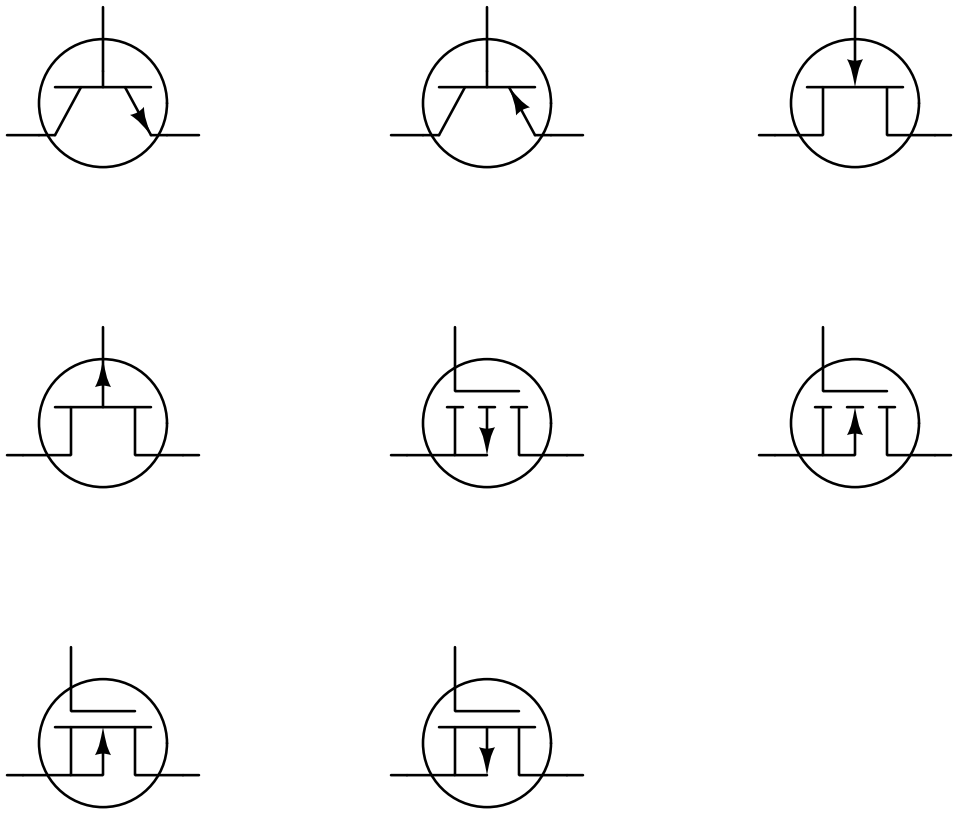
Commutating diode

Swamping

Floating FET gate

7.1.3 Transistor summary

When analyzing and designing transistor switching circuits, it can seem overwhelming at first to remember all the different types of transistors and what is necessary to activate (i.e. turn on) each type. A helpful *active reading* strategy is to make for yourself a reference page in your Journal showing all the different transistor types by schematic symbol, annotating each one with notes to help you remember how it works:



For each one of these transistors, identify its type, the condition (i.e. voltage/current) necessary to turn it on, the foundational principle of that transistor's operation, and the direction of controlled current once the transistor is on.

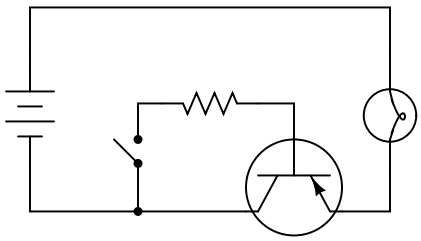
Challenges

- In this question an array of transistors was presented to you in schematic-symbol form. For future reference, how could you apply this same active reading strategy to other important and potentially hard-to-remember concepts, where no such question is assigned to you to answer? Identify some practical examples of this for past concepts you found challenging.

- Which of these transistor types is normally half-off?
- What is the operating principle of a BJT?
- What is the operating principle of a FET?

7.1.4 Simple PNP lamp control

Explain the function of this light-switching circuit, tracing the directions of all currents when the switch closes:

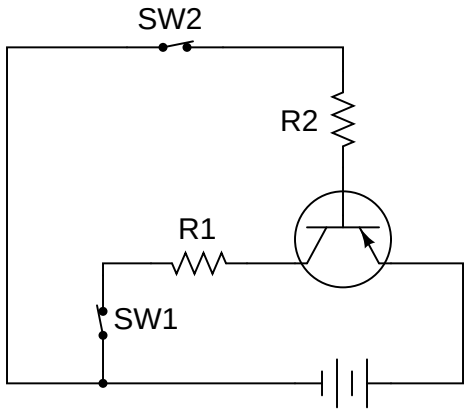


Challenges

- Identify the type of transistor used in this circuit, as well as its principle of operation.
- Is the transistor sourcing current to the lamp, or sinking current from it?
- Describe a procedure for properly sizing the resistor in this circuit.

7.1.5 Resistor-transistor-switch circuit

Trace the directions of all currents in this circuit, and determine which current is larger: the current through resistor R1 or the current through resistor R2, assuming equal resistor values.



If switch SW2 were opened (and switch SW1 remained closed), what would happen to the currents through R1 and R2?

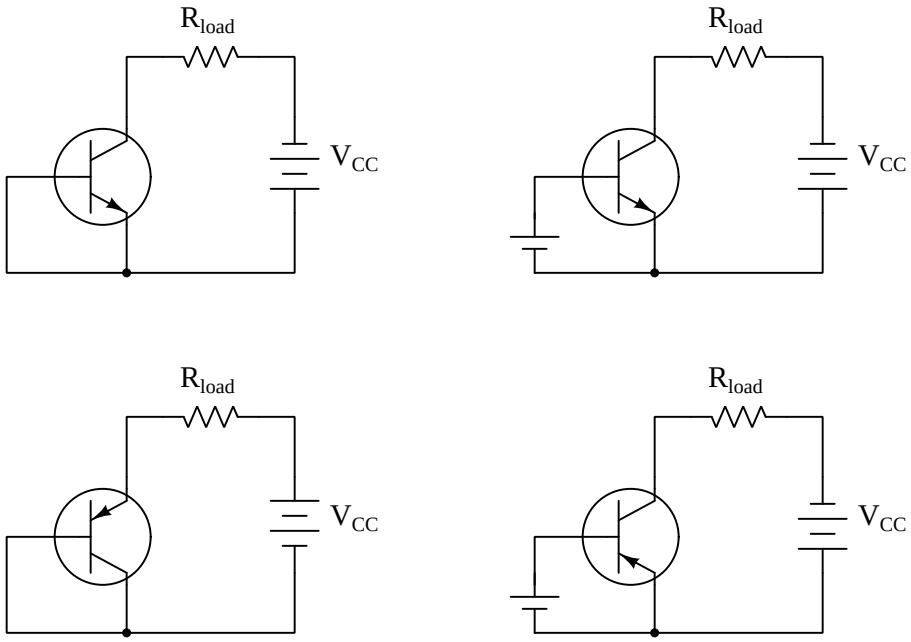
If switch SW1 were opened (and switch SW2 remained closed), what would happen to the currents through R1 and R2?

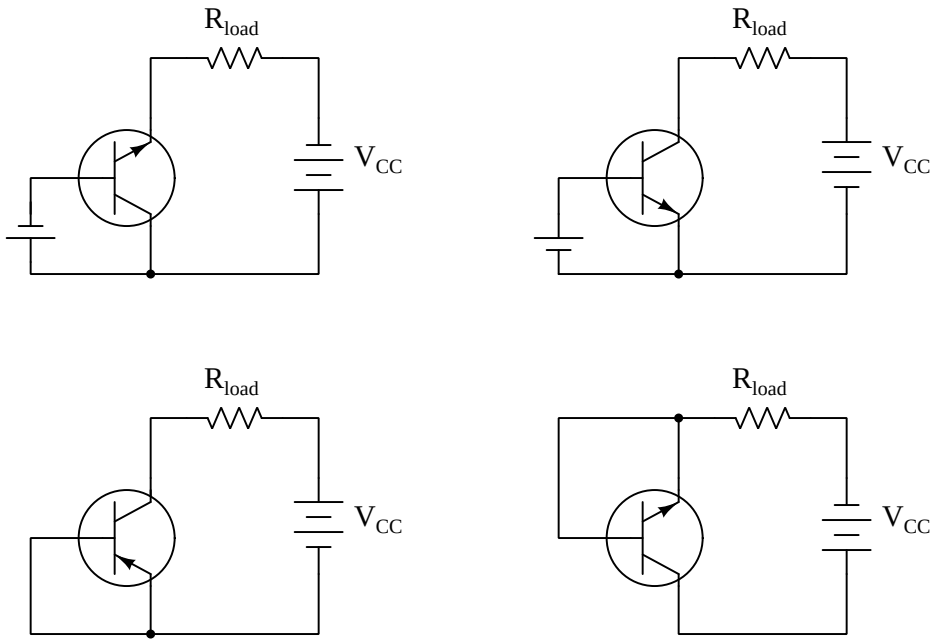
Challenges

- Identify the type of transistor used in this circuit, as well as its principle of operation.
- Is the transistor sourcing current to R1, or sinking current from it?

7.1.6 Identifying BJT states

Identify each type of BJT (whether it is NPN or PNP), label the terminals, and determine whether the BJT in each of these circuits will be turned *on* or *off*:



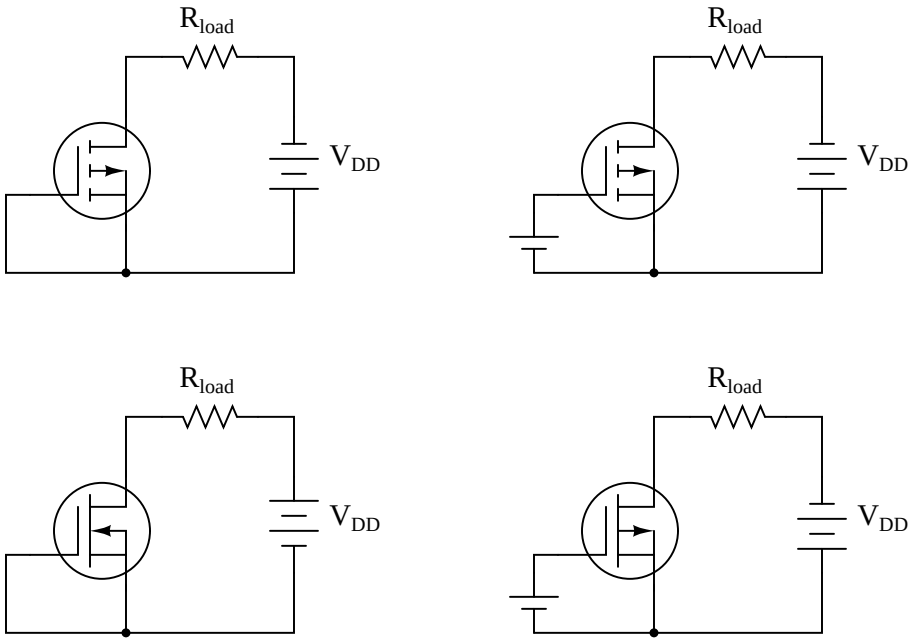


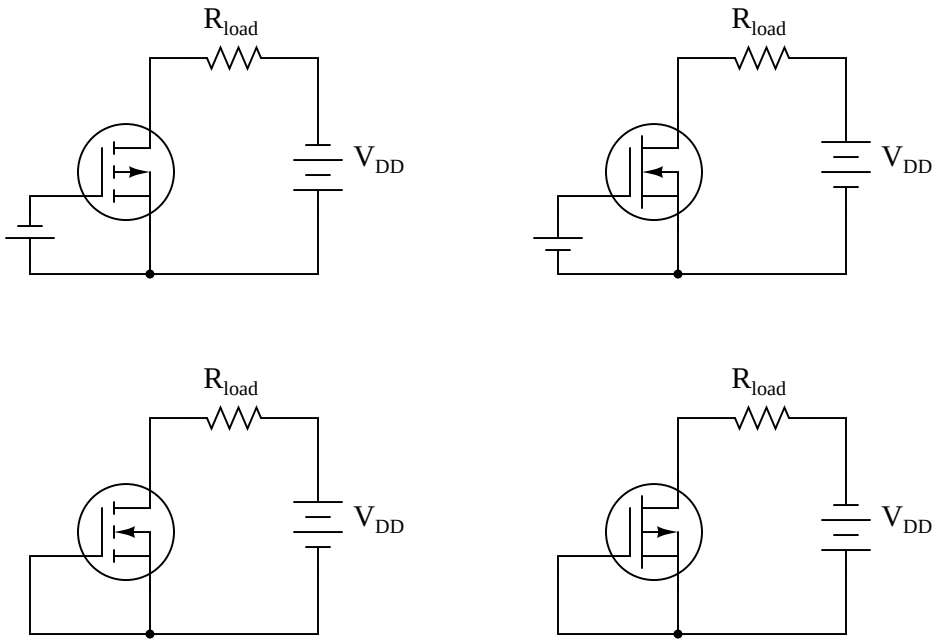
Challenges

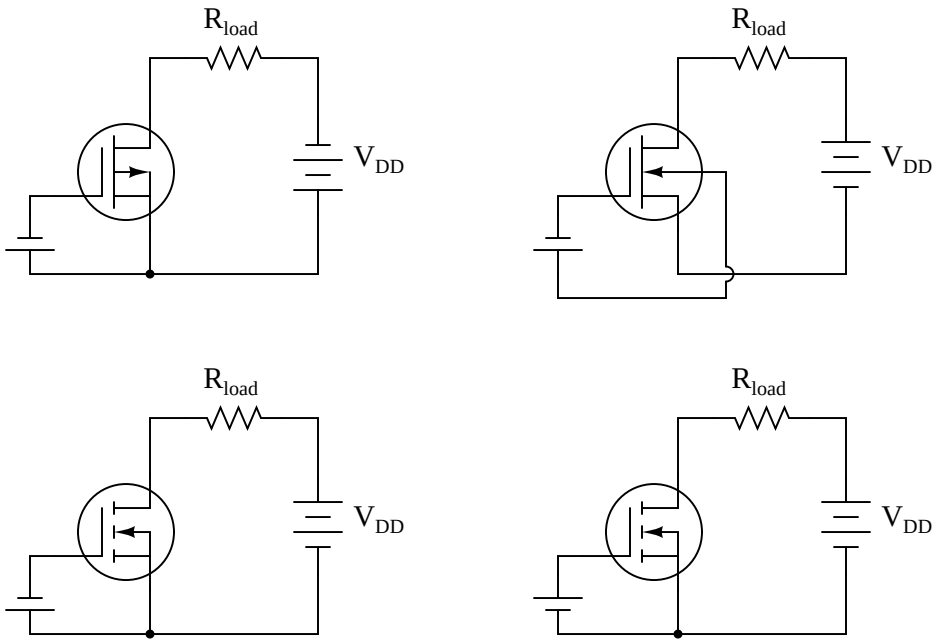
- Explain how to identify whether a BJT will be *saturated* on versus merely on.

7.1.7 Identifying MOSFET states

Identify each type of MOSFET (whether it is N-channel or P-channel, D-type or E-type), label the terminals, and determine whether the MOSFET in each of these circuits will be turned *on* or *off*:




R load
V DD
R load
V DD
R load
V DD
R load
V DD

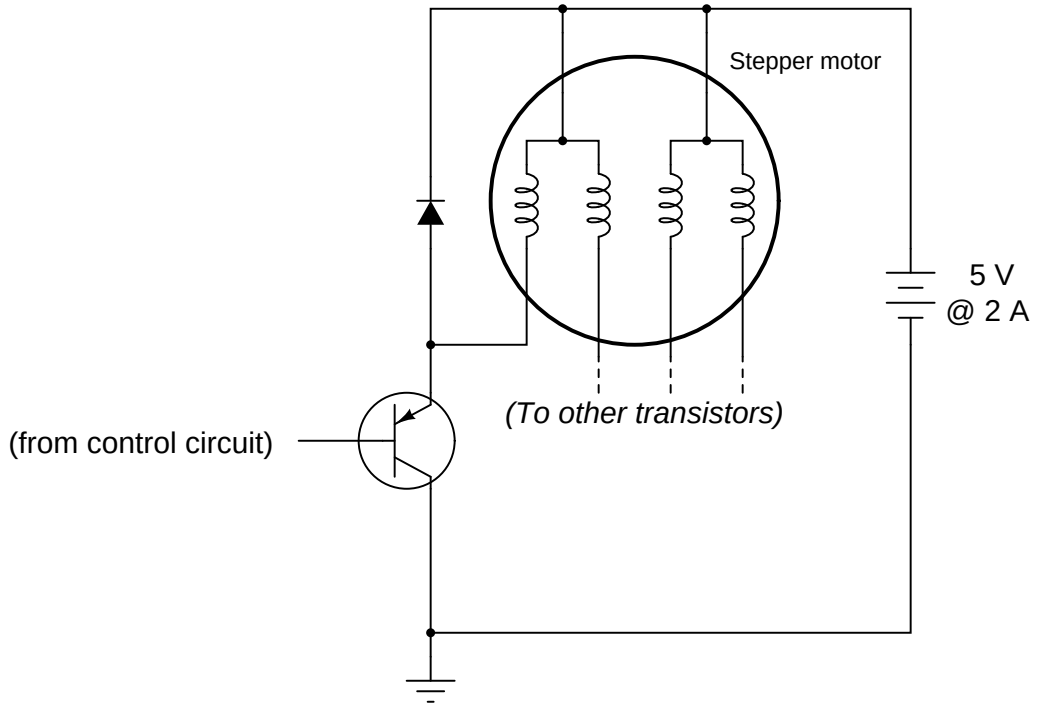


Challenges

- Which of these transistors is *depleted* and which is *enhanced*.

7.1.8 Stepper motor drive circuit

A *stepper motor* is a special type of electric motor typically used for digital positioning systems. The motor shaft rotates by alternately energizing its coils in a specific sequence. These electromagnet coils draw a fair amount of DC current (several amperes is not uncommon for heavy-duty stepper motors), and as such are usually triggered by power transistors:



The control circuit that usually sends pulse signals to the base of the power transistors is not shown in this diagram, for simplicity. Your task is to draw a pushbutton switch in this schematic diagram showing how the first motor coil could be manually energized and de-energized. Be sure to note the directions of currents through the transistor, so that your switch is installed correctly!

Also, explain the purpose of the diode connected in parallel with the motor coil. Actually, there will be one of these diodes for each of the motor coils, but the other three are not shown for the sake of simplicity.

Challenges

- Will your circuit's transistor operate in saturated or unsaturated mode?
- Re-design the circuit so that the transistor operates in its *other* mode (unsaturated or saturated) when turned on.

7.1.9 BJT datasheet research

Download a datasheet for a bipolar junction transistor, and from that datasheet identify and explain the meanings of the following parameters:

- Absolute maximum collector-emitter voltage:
- Absolute maximum base-emitter voltage:
- Absolute maximum continuous collector current:
- Absolute maximum power dissipation:
- DC current gain (also known as β or as h_{FE}):
- Storage time:
- Base saturation current:

Challenges

- What external factors might affect a transistor's maximum power dissipation?
- Devise a practical experiment to *test* any of these parameters.

7.1.10 MOSFET datasheet research

Download a datasheet for an enhancement-mode ("E-type") metal-oxide-semiconductor transistor, and from that datasheet identify and explain the meanings of the following parameters:

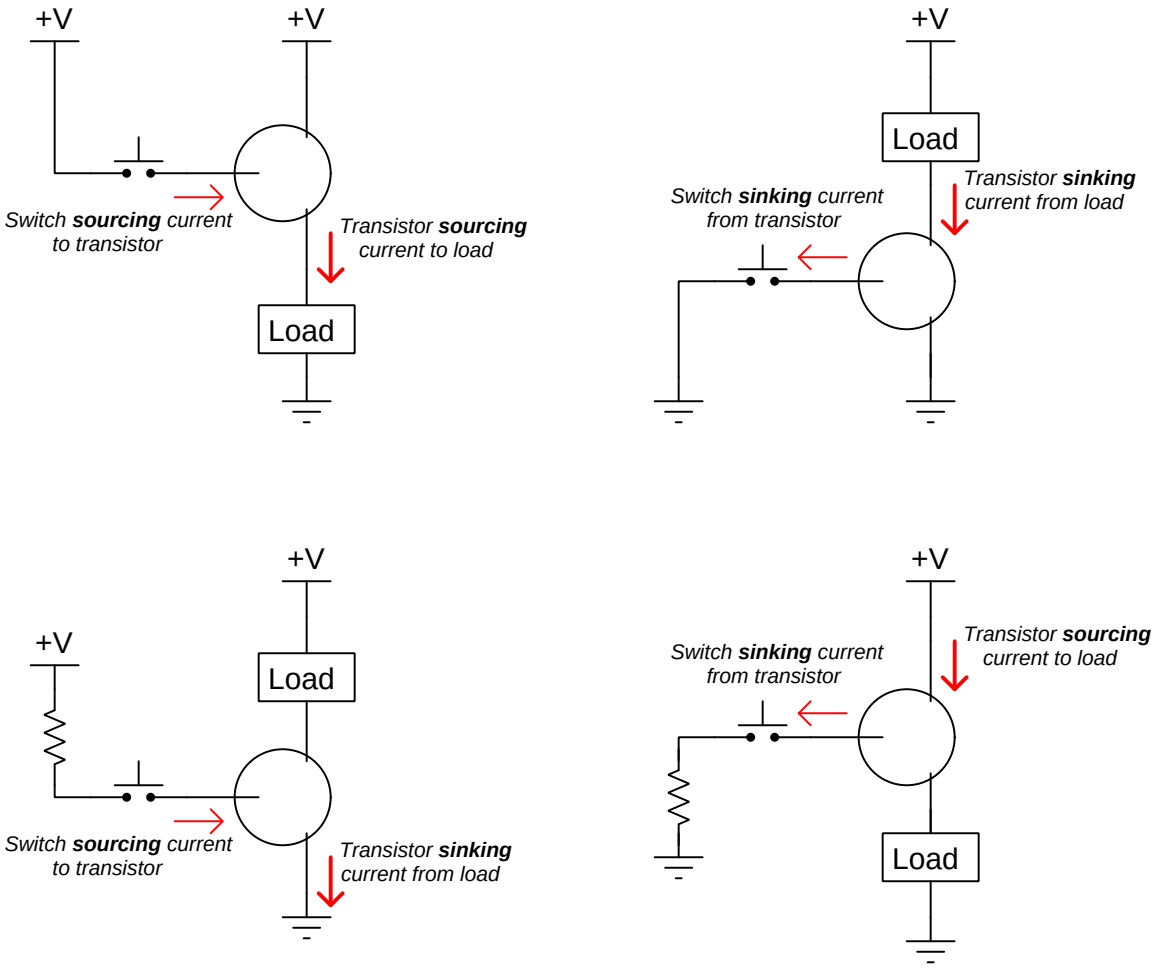
- Absolute maximum drain-source voltage:
- Absolute maximum gate-source voltage:
- Absolute maximum continuous drain current:
- Absolute maximum power dissipation:
- Forward transconductance (also known as g_fs):
- Turn-off delay time:
- Gate-source threshold (or "turn-on") voltage:

Challenges

- What external factors might affect a transistor's maximum power dissipation?
- Devise a practical experiment to *test* any of these parameters.

7.1.11 Proper BJT types

Choose the right type of bipolar junction transistor for each of these switching applications (i.e. either NPN or PNP), drawing the correct transistor symbol inside each circle:



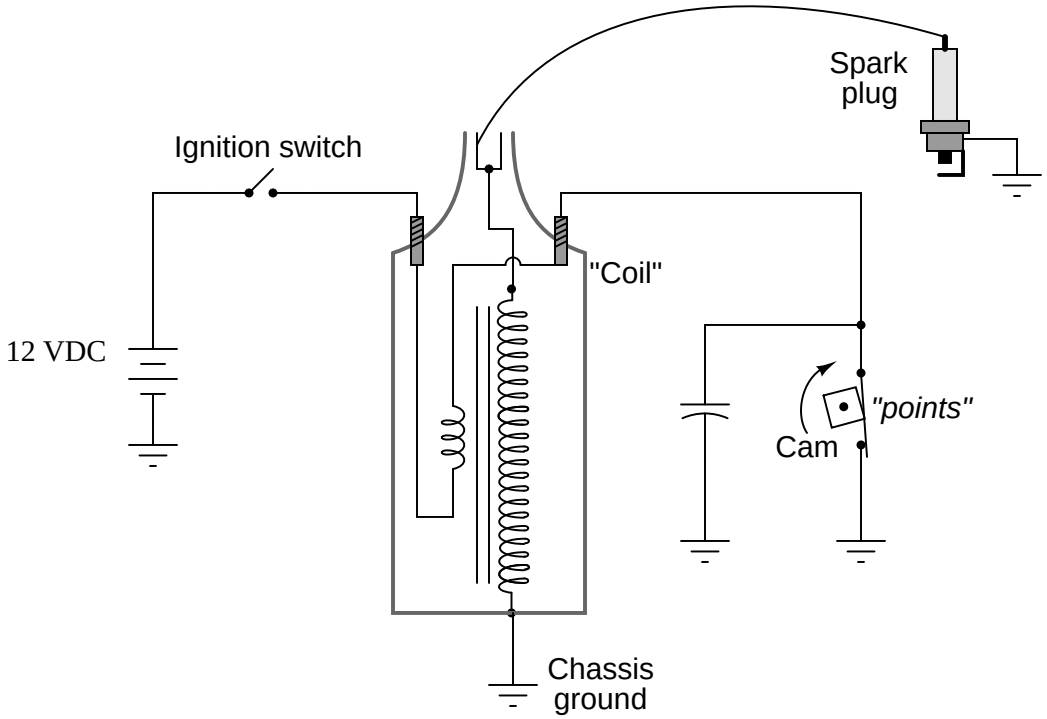
Also, determine for each of these switching circuits whether or not it is possible to saturate the transistor fully "on".

Challenges

- Define the terms "source" and "sink" as they apply to these circuits.

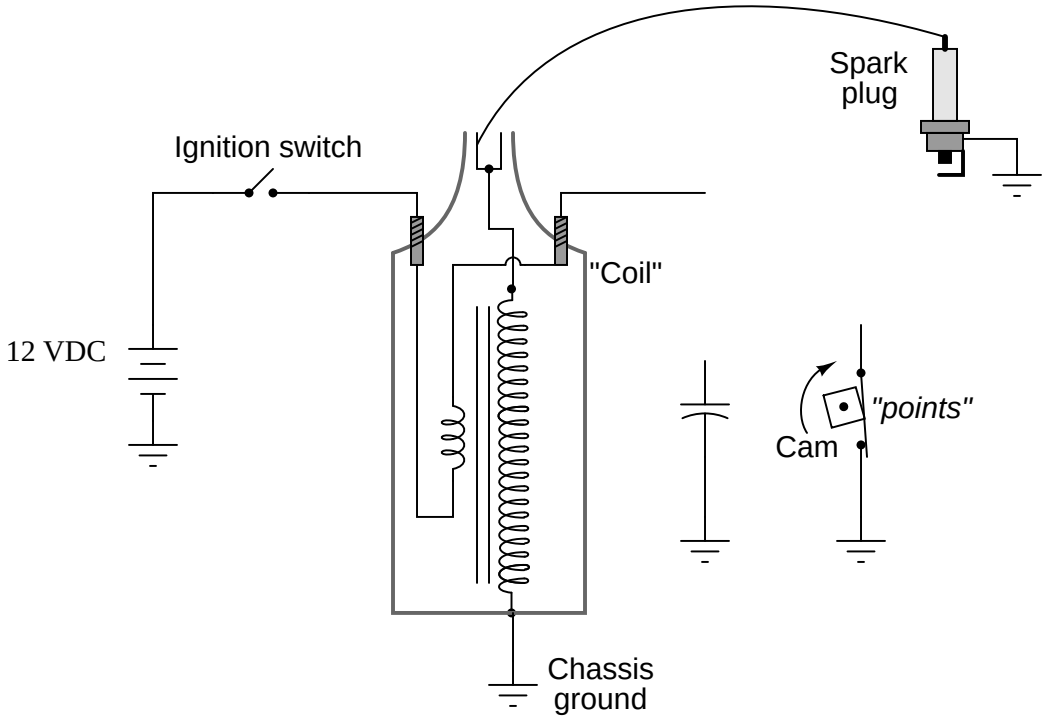
7.1.12 Electronic ignition circuit

The classic *ignition system* design for single-cylinder gasoline-fueled internal combustion engine is an example of a transformer operated on DC by means of an oscillating switch contact, commonly referred to as the contact "points":



The cam-actuated "point" switch contacts open every time a spark is needed to ignite the air-fuel mixture in the engine's cylinder. Naturally, these contacts suffer a substantial amount of wear over time due to the amount of current they must make and break, and the frequency of their cycling. This device was seen by engine designers as a prime candidate for replacement with solid-state technology (i.e. a transistor). If a transistor could take the place of mechanical "point" contacts for making and breaking the ignition coil's current, it should result in increased service life.

Insert a transistor into the following circuit in such a way that it controls the ignition coil's current, with the "point" switch contacts merely controlling the transistor's state (turning it on and off.

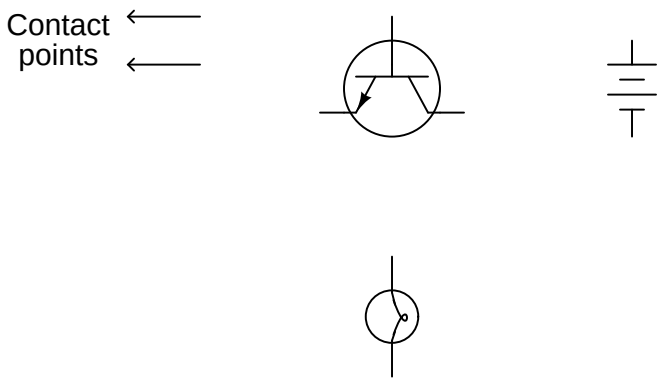


Challenges

- Do you think a MOSFET would function better or worse than a BJT for this application?
- Explain why a commutating diode would *not* be appropriate in this circuit to protect the "points" from arcing, or to protect the transistor from over-voltage failure.

7.1.13 Design a touch-controlled lamp

Draw the necessary wire connections so that bridging the two contact points with your finger (creating a high-resistance connection between those points) will turn the light bulb on:

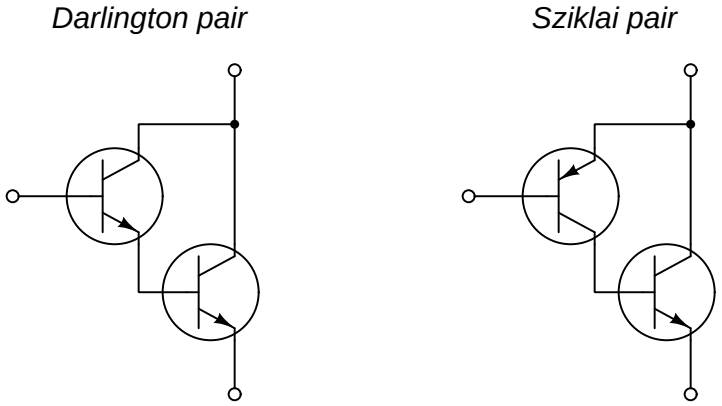


Challenges

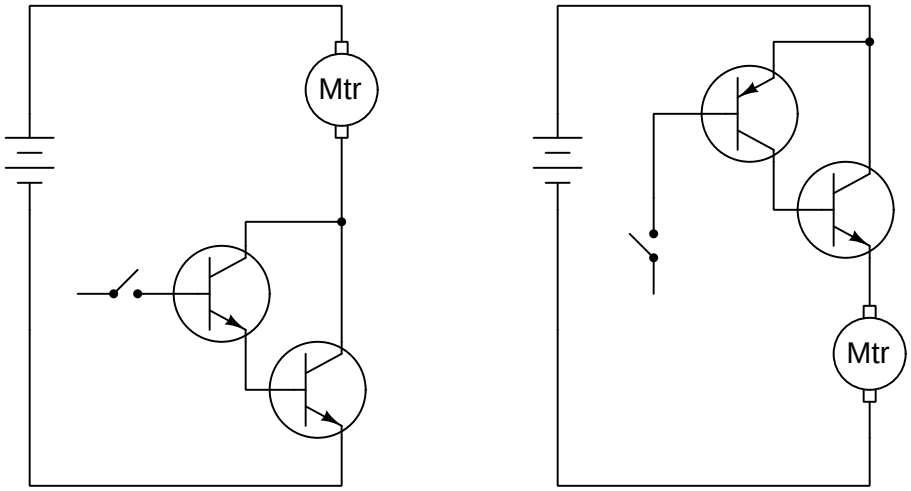
- Do you think a MOSFET would function better or worse than a BJT for this application?

7.1.14 Darlington and Sziklai pair connections

Two common dual-transistor networks used to achieve greater amplification are the *Darlington pair* and *Sziklai pair*. Each contains two bipolar junction transistors, but while the Darlington pair uses same-type transistors the Sziklai pair uses opposite-type transistors:



Complete the following circuits, showing how a switch would be connected to each of the transistor pairs to exert control over the electric motor:

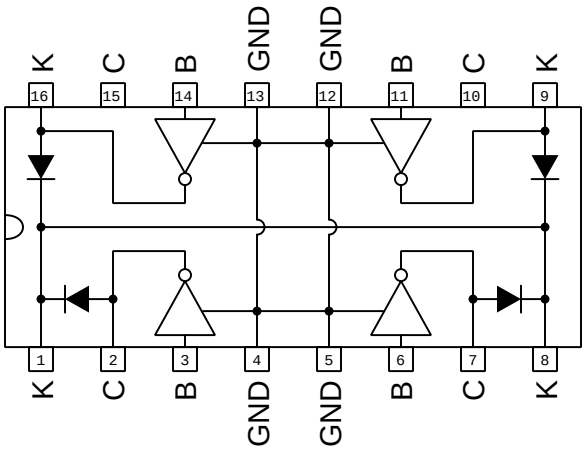


Challenges

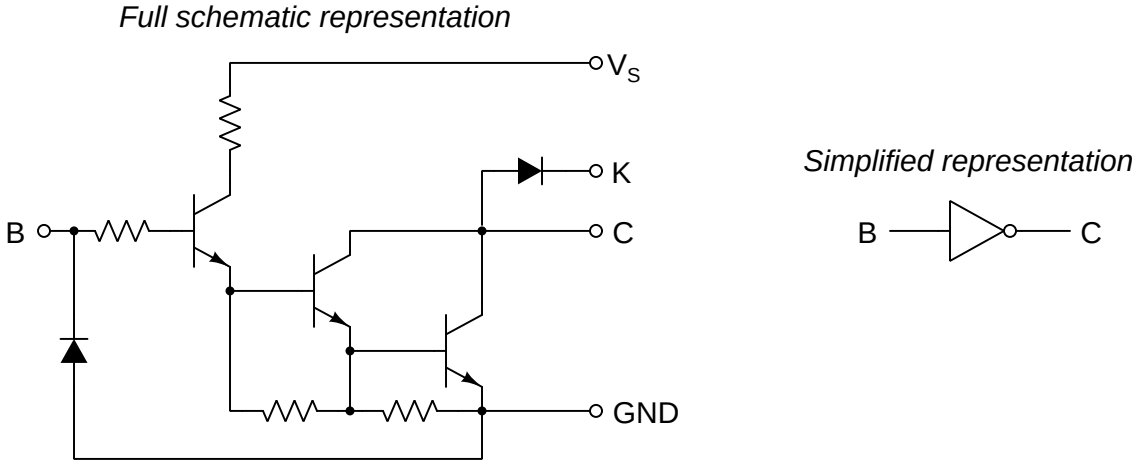
- Which configuration, Darlington or Sziklai, yields the greatest amount of amplification?

7.1.15 ULN2068B Darlington array IC

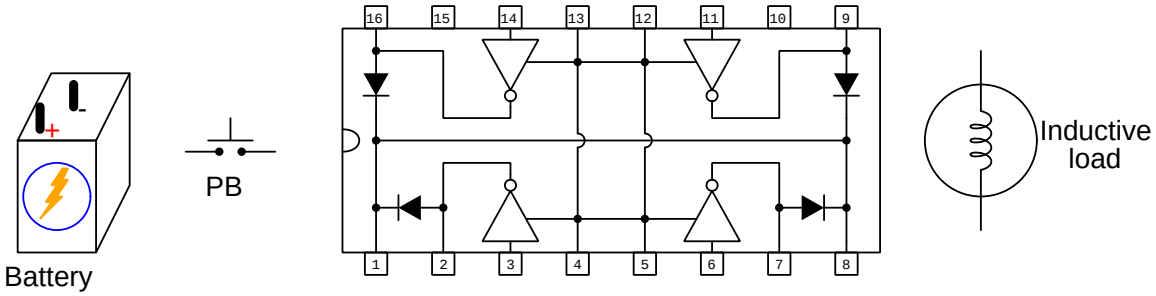
Darlington transistor networks are so useful for switching power to DC loads that they are manufactured in integrated circuit arrays such as the model ULN2068B. This particular IC sports four Darlington-triplet networks, each one shown in simplified form as a logical "NOT" or inverter gate:



Each of these Darlington "inverters" has its own resistors and protection diodes also built into the integrated circuit:



Sketch wires to form a switching circuit so that the pushbutton controls one of the Darlington networks in this IC to energize an inductive load (a solenoid coil), using the built-in commutation diode provided by the IC for protection against "flyback" or "kickback":

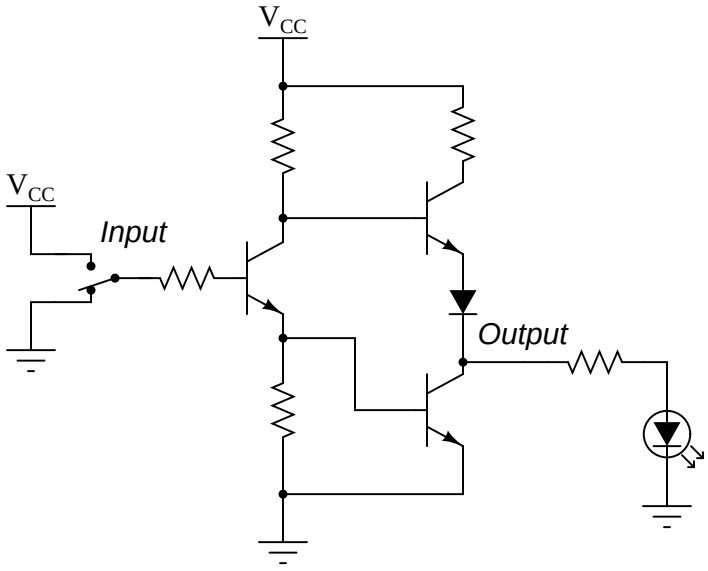


Challenges

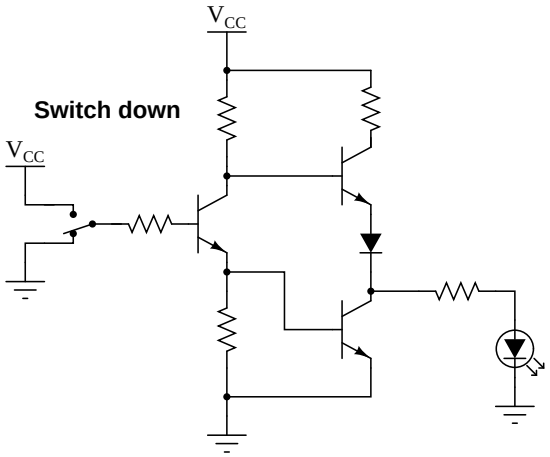
- What might happen if the commutating diode is not properly connected in this circuit?

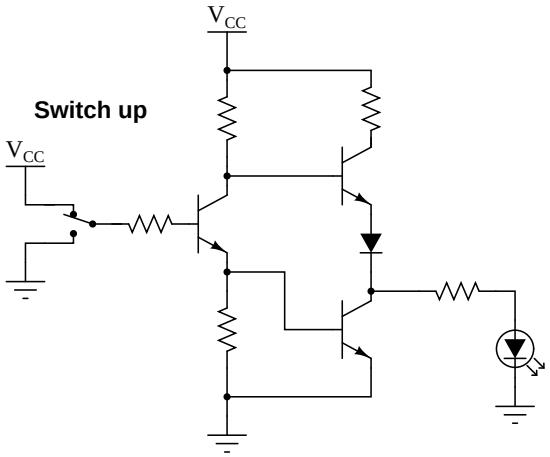
7.1.16 Logic gate voltage states

The circuit shown here is part of a digital *logic gate* circuit:



Logic circuits operate with their transistors either fully "on" or fully "off", never in-between. Determine which state the LED will be in (either on or off) for both switch positions. You may find it helpful to trace currents and label all voltage drops in this circuit for the two switch states:





For your voltage drop calculations, assume the following parameters:

- $V_{CC} = 5$ Volts

- V_{BE} (conducting) = 0.7 Volts
- V_{CE} (conducting) = 0.3 Volts
- V_f (regular diode conducting) = 0.7 Volts
- V_f (LED conducting) = 1.6 Volts

Challenges

- What type of logic gate circuit is this?

7.1.17 Electron versus hole mobility

All other factors being equal, N-type semiconductors have less resistance than P-type semiconductors. This is due to the fact that free electrons are naturally more mobile than holes, able to traverse a semiconductor crystal's length with fewer collisions that would rob energy from the electrons and result in a voltage drop from one end of the material to the other.

Given this fact, which type of MOSFET makes a better high-power switch, N-channel or P-channel?

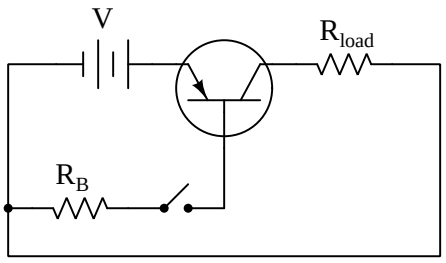
Given this fact, which type of BJT makes a better high-power switch, PNP or NPN?

Challenges

- Why are free electrons more mobile than holes?

7.1.18 Explaining the meaning of calculations

Examine the calculations performed in the analysis of this transistor circuit (in the switch's "on" state), then describe what each of the calculations represents, identify its proper unit of measurement and any appropriate metric prefix(es), and identify the foundational concept legitimizing each step:



Assume $V = 15$ Volts, $R_B = 1$ kΩ, $R_{load} = 40\ \Omega$, and $h_{FE} = 90$.

1. $\frac{15-0.7}{1000} = 14.3 \times 10^{-3}$
2. $(14.3 \times 10^{-3}) \times 90 = 1.287$
3. $\frac{15-0.3}{40} = 367.5 \times 10^{-3}$
4. $367.5 \times 10^{-3} < 1.287$
5. $(367.5 \times 10^{-3}) + (14.3 \times 10^{-3}) = 381.8 \times 10^{-3}$
6. $(14.3 \times 10^{-3})(0.7) + (367.5 \times 10^{-3})(0.3) = 120.26 \times 10^{-3}$

Is it possible to determine whether this transistor will be saturated or not when the switch is closed? Explain why.

Challenges

- What does it mean for a transistor to operate in a *saturated* condition?
- Why is saturation often desired for transistor switching circuits?

7.2 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.

7.2.1 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.

7.2.2 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	

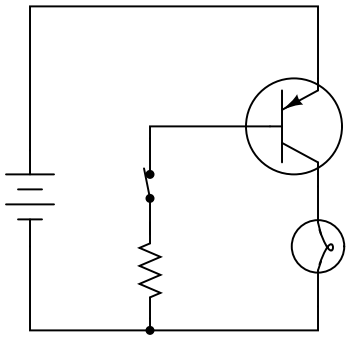
Note how the square-root term (y) is calculated in cell `C1`, and the denominator term (z) in cell `C2`. This makes the two final formulae (in cells `B1` and `B2`) simpler to interpret. The positioning of all these cells on the grid is completely arbitrary[10] – all that matters is that they properly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

7.2.3 Voltage calculations in a simple lamp circuit

Calculate all component voltage drops in this circuit, assuming a supply voltage of 15 Volts, an emitter-base forward voltage drop of 0.7 Volts, and a (saturated) emitter-collector voltage drop of 0.3 Volts:



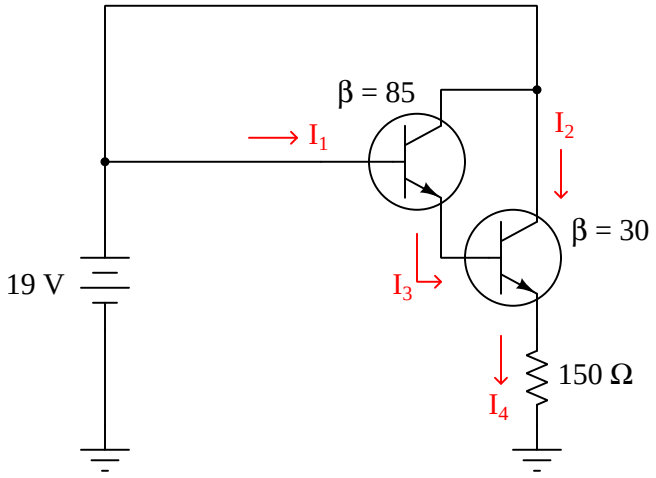
Also, explain the role β plays in this circuit, if any. Does β matter at all in this application, and if so how does it matter?

Challenges

- Calculate the necessary resistor value to provide the transistor's base terminal with 10 mA of current.
- What ill effects, if any, may result from selecting a value for R that is too small?
- What ill effects, if any, may result from selecting a value for R that is too large?

7.2.4 Darlington pair current calculations

A *Darlington pair* is a network of two transistors where the emitter current from one becomes the base current of the next, for the purpose of increased amplification. Calculate all labeled currents in this Darlington pair circuit assuming a typical forward base-emitter junction voltage drop of 0.7 Volts for each transistor:

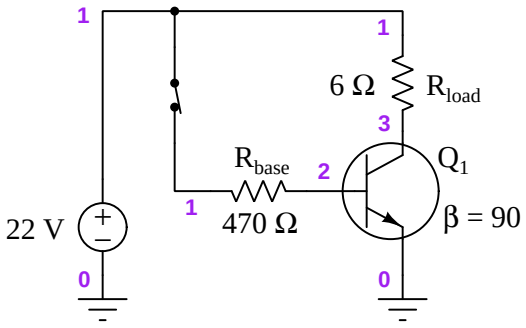


Challenges

- Where is the most logical location for a pushbutton switch in this circuit, to control the Darlington pair?
- What is the point of using a Darlington pair rather than a single transistor?

7.2.5 Over- and under-driving a BJT switch

A SPICE simulation computes voltages within this simple transistor circuit, illustrating a BJT switching circuit in the "on" state:



SPICE netlist:

```
* SPICE circuit
vsupply 1 0 dc
rload 1 3 6
rbase 1 2 470
q1 3 2 0 generic
.model generic npn bf=90
.dc vsupply 22 22 1
.print dc v(1,3) v(3) v(2)
.end
```

When run, the SPICE simulation provides the following results:

```
------------------------------------------------------------------------
Index   v-sweep         v(1)-v(3)       v(3)            v(2)
------------------------------------------------------------------------
0       2.200000e+01    2.182500e+01    1.750045e-01    9.863298e-01
```

These results show the transistor to be *fully on*, dropping only 0.175 Volts between collector and emitter terminals, and dropping 0.9863 Volts between base and emitter. Normally, we assume a V_{BE} value of 0.7 Volts to be fully forward-biased and a V_{CE} value of 0.3 Volts to be fully conducting for a silicon BJT, but here we are "driving" the transistor harder than the assumed values.

Identify how to increase the "drive" of this transistor even further than it is, and predict the effects of the harder-driving on those same voltages. Feel free to edit the SPICE netlist and re-run the simulation to check your predictions!

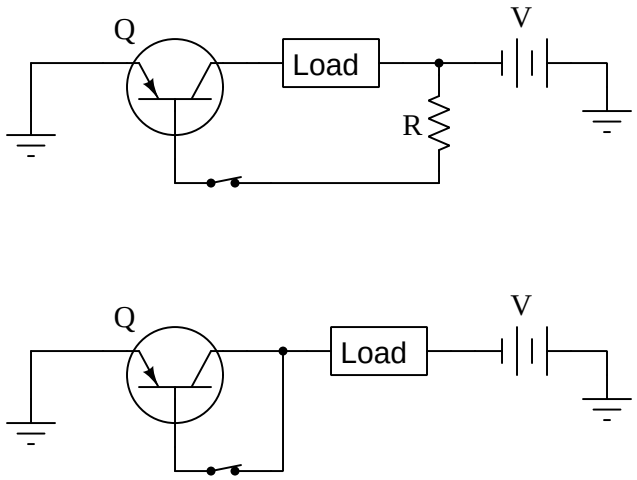
Next, modify the circuit simulation to "under-drive" the transistor so that it is less than fully-on, then predict the effects of under-driving on those same voltages. Once again, feel free to use SPICE to verify your predictions.

Challenges

- Identify *multiple* ways to either over- or under-drive this transistor.
- In which scenario – over-driven or under-driven – do you suspect the transistor will dissipate the most heat?

7.2.6 Saturated versus unsaturated BJT switching circuits

Below are schematic diagrams for two similar BJT switching circuits, one of them operating with the transistor in "saturated[11]" mode (i.e. *fully* on) and the other operating in "active" (unsaturated) mode:



Load resistance, source voltage, and transistor β ratings are identical for both circuits: $R_{load} = 25\ \Omega$, $V = 40$ V, and $\beta = 30$. In the upper circuit the resistor value is 470 Ω.

Calculate the following for both circuits and comment on the contrasts. Also, identify which circuit has the saturated transistor versus which one has the unsaturated (active) transistor:

- $V_{load} =$
- $I_{load} =$
- $V_{CE} =$
- $V_{BE} =$
- $P_{load} =$
- $P_Q =$

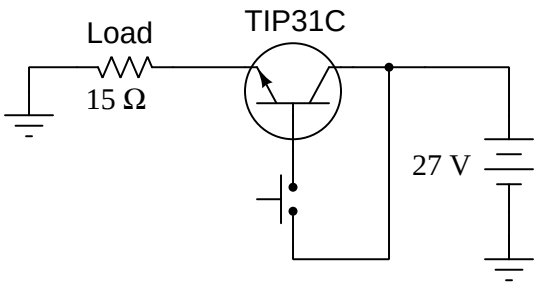
Challenges

- Identify for each circuit whether the transistor is *sourcing* or *sinking* current.
- Replace the BJTs with appropriate MOSFETs, and then comment on how well each circuit design functions.

[11] By "saturated" I mean in this context that the transistor is passing as much current through its collector terminal as it can given a limited supply voltage. In other words, the base is being "excited" more than enough for the I_C it's passing from the load.

7.2.7 Calculations in TIP31C transistor circuit

Determine the following in this switching circuit using a model TIP31C transistor (Maximum I_C = 3 Amperes, Maximum V_{BE} = 1.8 Volts, Maximum $V_{CE(off)}$ = 100 Volts, β = 25, Maximum $V_{CE(on)}$ = 1.2 Volts):



- P_Q when "on" =
- P_{load} when "on" =
- I_{switch} when "on" =
- V_{CE} when "on" =

Identify how some of the foundational concepts you've studied apply to this circuit: **behavior of sources vs. loads**, properties of **opens vs. shorts**, **Ohm's Law**, **Conservation of Energy**, **Joule's Law**, **Kirchhoff's Voltage Law**, **Kirchhoff's Current Law**, **Conservation of Electric Charge**, **Beta ratio**, **properties of PN junctions**. Feel free to include any other foundational concepts not listed here.

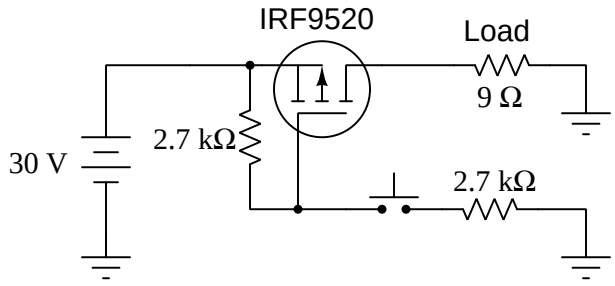
Also, determine whether or not this transistor will operate in a mode of *saturation* when on.

Challenges

- Why is no resistor necessary in series with the transistor's base terminal?
- Is this circuit using the TIP31C as a low-side switch or as a high-side switch?
- Modify this circuit to use the same transistor on the other side of the load.

7.2.8 Calculations in IRF9520 transistor circuit

Determine the following in this switching circuit using a model IRF9520 transistor (Maximum I_D = 6.8 Amperes, Maximum V_{GS} = 20 Volts, Maximum V_{DS} = 100 Volts, V_{th} = 2 to 4 Volts, Maximum $R_{DS(on)}$ = 0.60 Ω):



- P_Q when "on" =
- P_{load} when "on" =
- I_{switch} when "on" =
- V_{GS} when "on" =
- V_{DS} when "on" =

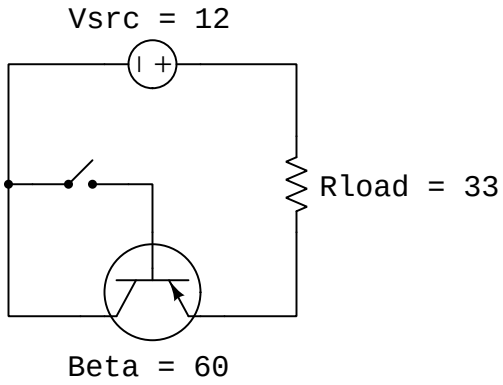
Also, determine whether or not this transistor will operate in a mode of *saturation* when on.

Challenges

- Why are the 2.7 kΩ resistors necessary?
- Suppose we assume $R_{DS(on)}$ to be zero. How will this affect our calculations?
- The model IRF510 transistor exhibits a lower $R_{DS(on)}$ with the same applied V_{GS}. Both the IRF510 and IRF9520 are physically similar, and both rated at 50 Watts dissipation. Why does the IRF510 have less "on" resistance?
- Is this circuit using the IRF9520 as a low-side switch or as a high-side switch?
- Modify this circuit to use the same transistor on the other side of the load.

7.2.9 BJT switching circuit simulation program

Write a text-based computer program (e.g. C, C++, Python) to calculate all currents in the following transistor switch circuit when the manual switch is turned "on", given the variable names and values shown for each component in the schematic diagram:



Challenges

- Will this transistor be in a saturated state when turned on? Why or why not?
- Modify the circuit to ensure the transistor will be in a different saturated state than it is now when turned on.

7.3 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.

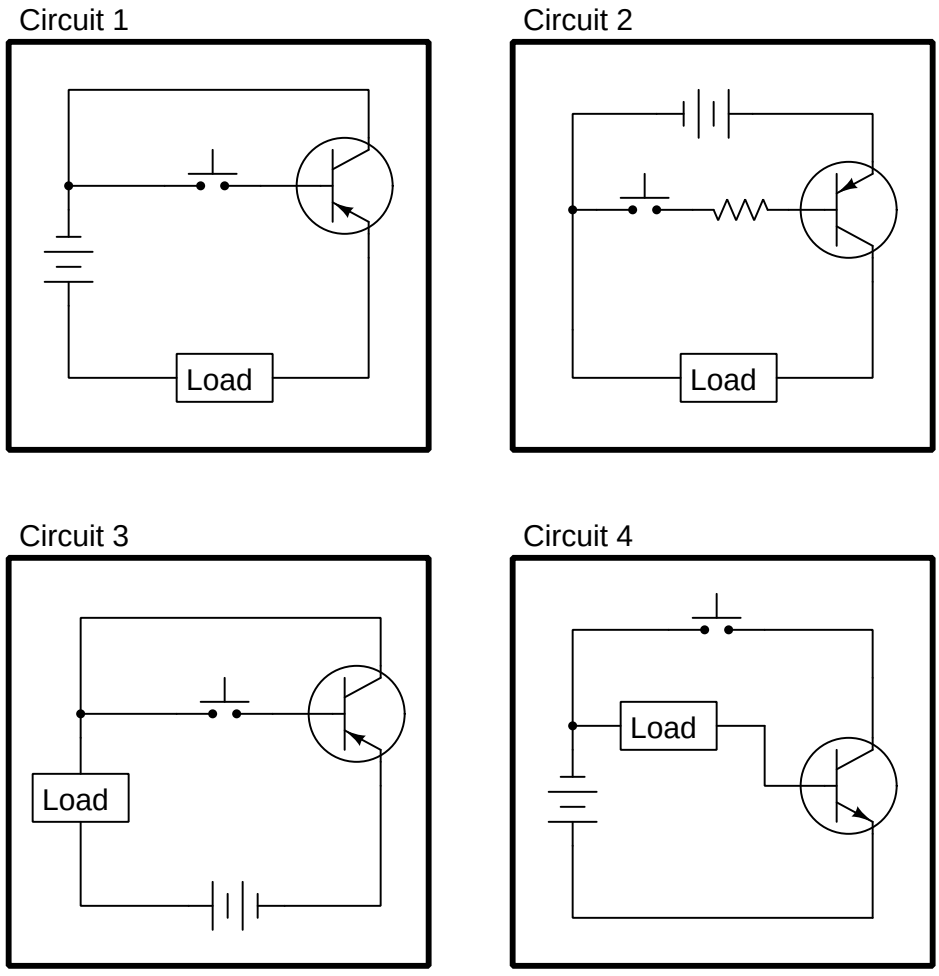
As always, your goal is to fully *explain* your analysis of each problem. Simply obtaining a correct answer is not good enough – you must also demonstrate sound reasoning in order to successfully complete the assignment. Your instructor's responsibility is to probe and challenge your understanding of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

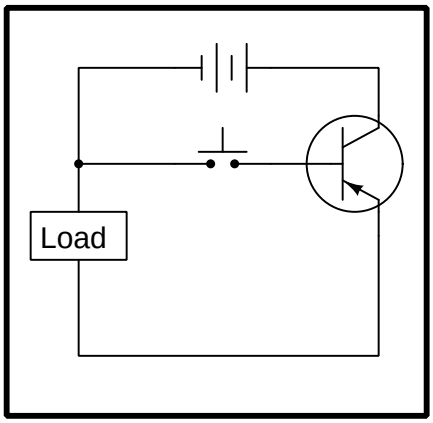
Another means of checking your diagnostic answers, where applicable, is to use circuit simulation software to explore the effects of faults placed in circuits. For example, if one of these diagnostic questions requires that you predict the effect of an open or a short in a circuit, you may check the 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.

7.3.1 BJT switching circuit configurations

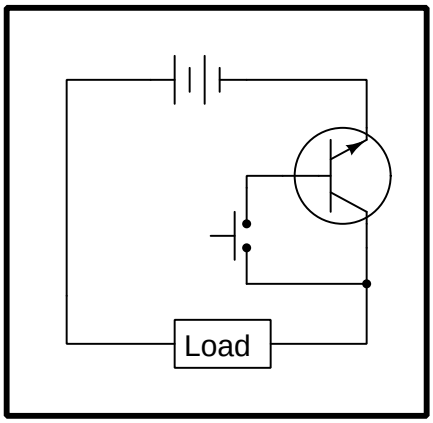
Some of the following bipolar junction transistor switch circuits are properly configured, and some are not. Identify which of these circuits will function properly (i.e. turn on the load when the switch closes) and which of these circuits are mis-wired:



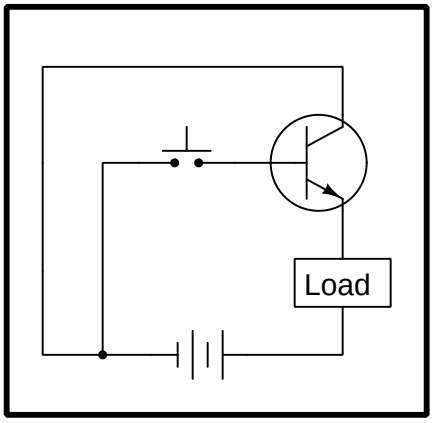
Circuit 5



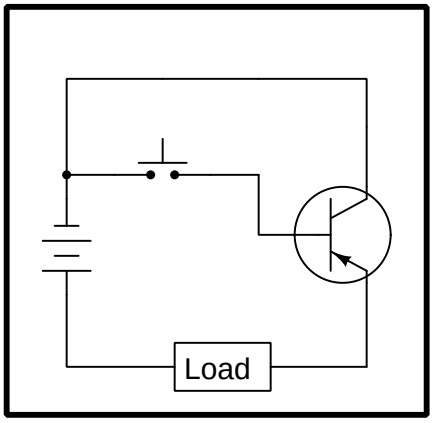
Circuit 6



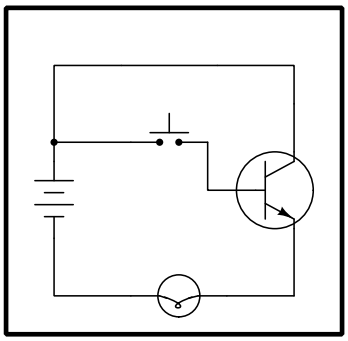
Circuit 7



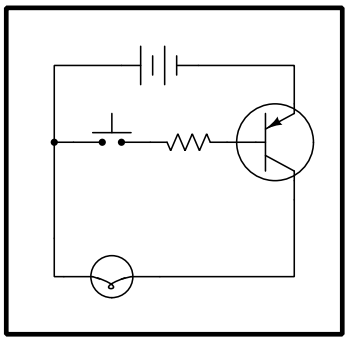
Circuit 8



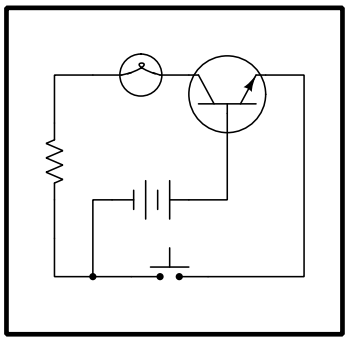
Circuit 9



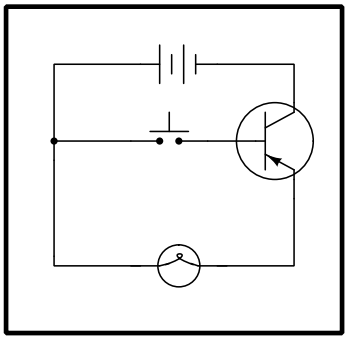
Circuit 10



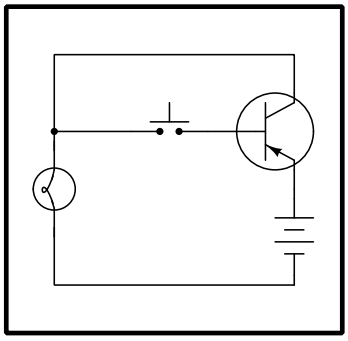
Circuit 11



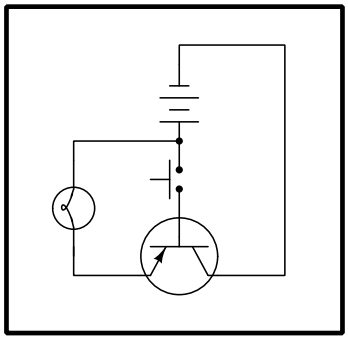
Circuit 12



Circuit 13



Circuit 14



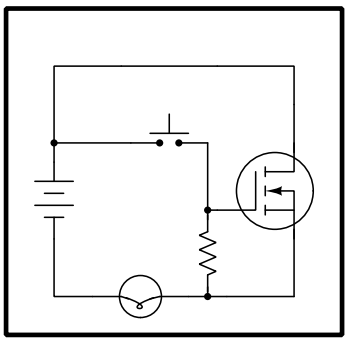
Challenges

- Identify all criteria necessary to activate a BJT.
- Of the functional transistor circuits, which of them are capable of saturating the transistor and which of them are not?

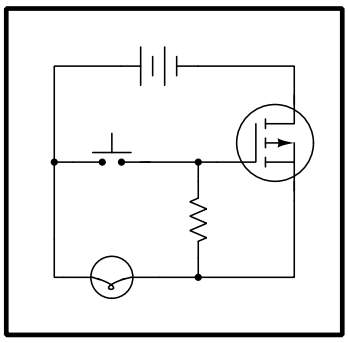
7.3.2 MOSFET switching circuit configurations

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:

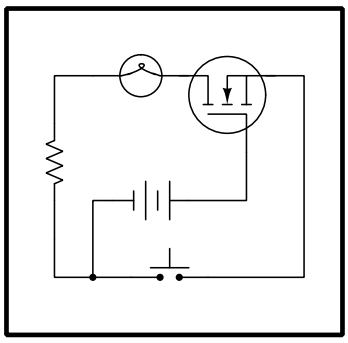
Circuit 1



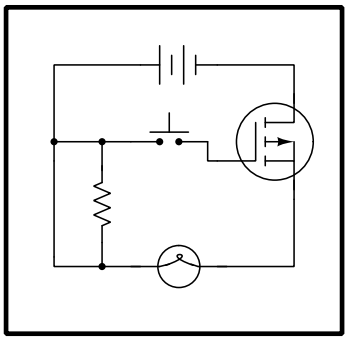
Circuit 2



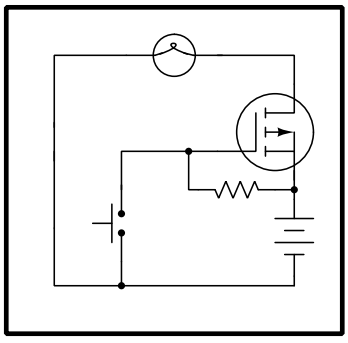
Circuit 3



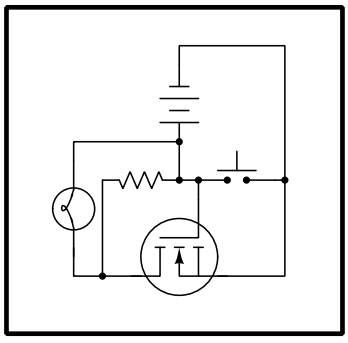
Circuit 4



Circuit 5



Circuit 6



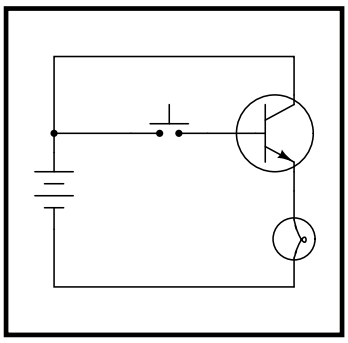
Challenges

- Identify all criteria necessary to enhance a MOSFET.

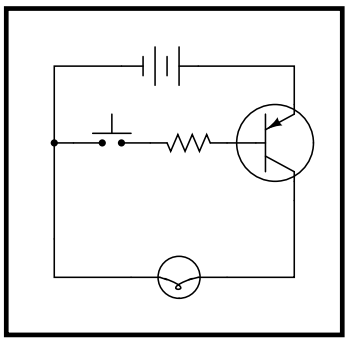
7.3.3 Reliable and unreliable lamp control circuits

In each of the following circuits, the lamp will energize when the pushbutton switch is actuated. Assume that the supply voltage in each case is somewhere between 5 and 30 Volts DC (with lamps and resistors appropriately sized):

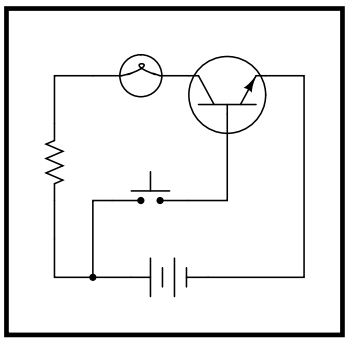
Circuit 1



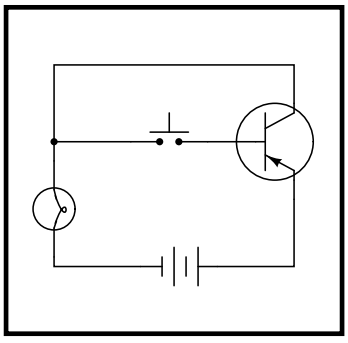
Circuit 2



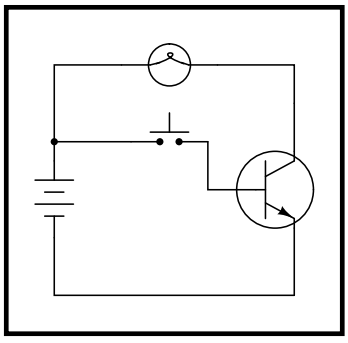
Circuit 3



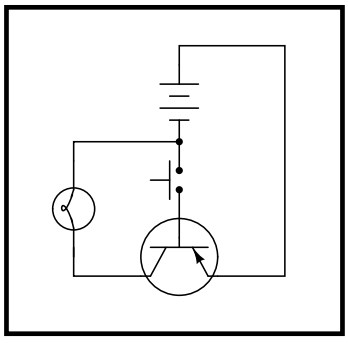
Circuit 4



Circuit 5



Circuit 6



However, not all of these circuits are properly designed. Some of them will function perfectly, but others will function only once or twice before their transistors fail. Identify the faulty circuits, and explain why they are flawed.

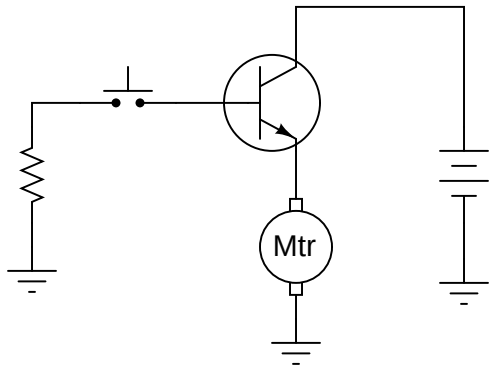
Challenges

- Sketch the respective paths of switch and lamp current for each circuit.
- For each faulty circuit design, re-draw it to correct the fault and make its operation reliable.

7.3.4 Fixing a mis-designed BJT motor control circuit

A student attempts to build a circuit that will turn a DC motor on and off with a very delicate (low current rating) pushbutton switch. Unfortunately, there is something wrong with the circuit, because the motor does not turn on no matter what is done with the switch:

This circuit does not work!



Correct the error(s) in this circuit, showing how it must be set up so that the transistor functions as intended.

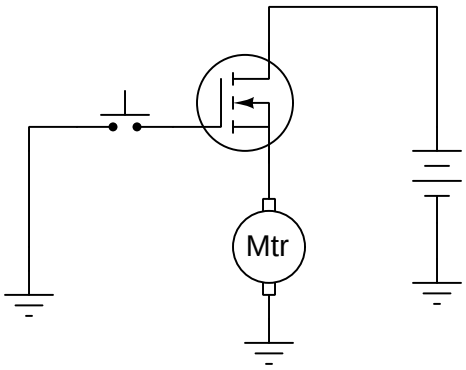
Challenges

- Will your redesigned circuit's transistor operate in saturated or unsaturated mode?
- Re-design the circuit so that the transistor operates in its *other* mode (unsaturated or saturated) when turned on.

7.3.5 Fixing a mis-designed MOSFET motor control circuit

A student attempts to build a circuit that will turn a DC motor on and off with a very delicate (low current rating) pushbutton switch. Unfortunately, there is something wrong with the circuit, because the motor does not turn on no matter what is done with the switch:

This circuit does not work!



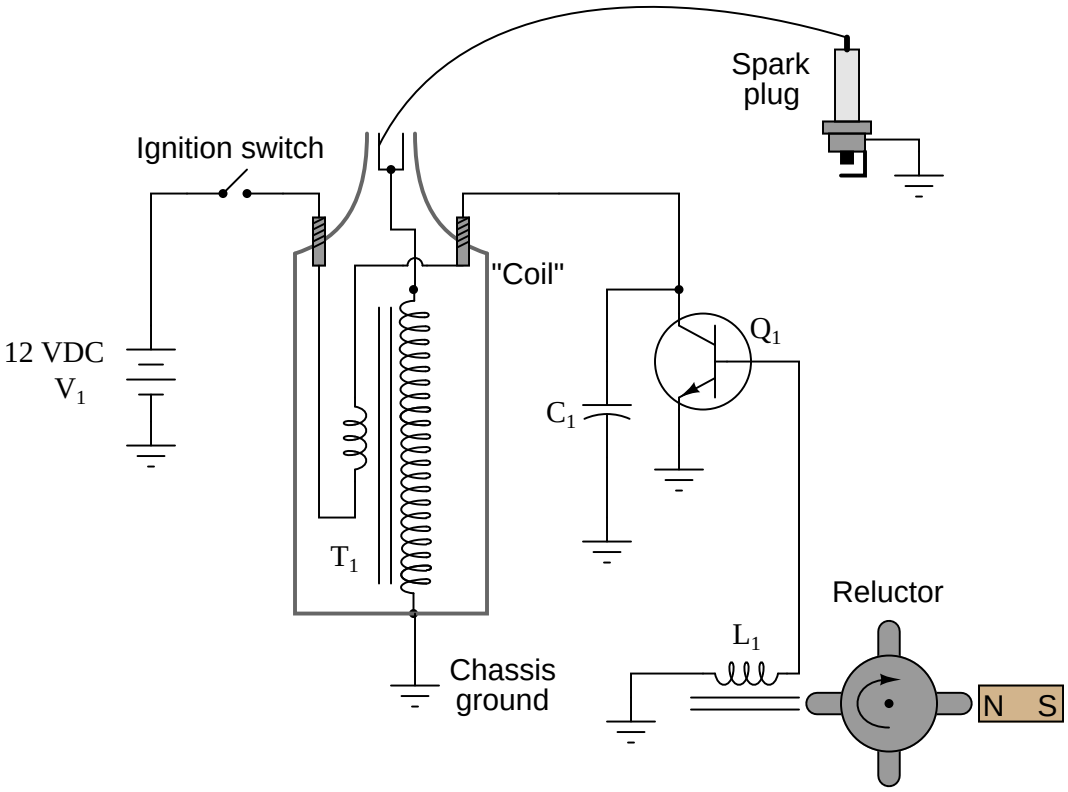
Correct the error(s) in this circuit, showing how it must be set up so that it functions as intended. Incidentally, this is a *very* common design error made by students new to the use of MOSFETs!

Challenges

- Will your redesigned circuit's transistor operate in saturated or unsaturated mode?
- Re-design the circuit so that the transistor operates in its *other* mode (unsaturated or saturated) when turned on.

7.3.6 Failed electronic ignition system

Predict how all component voltages and currents in 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 open (collector-to-emitter):
- Transistor Q_1 fails shorted (collector-to-emitter):
- Reluctor magnet weakens:
- Capacitor C_1 fails shorted:
- Capacitor C_1 fails open:
- Transformer ("coil") T_1 primary winding fails open:
- Transformer ("coil") T_1 secondary winding fails open:

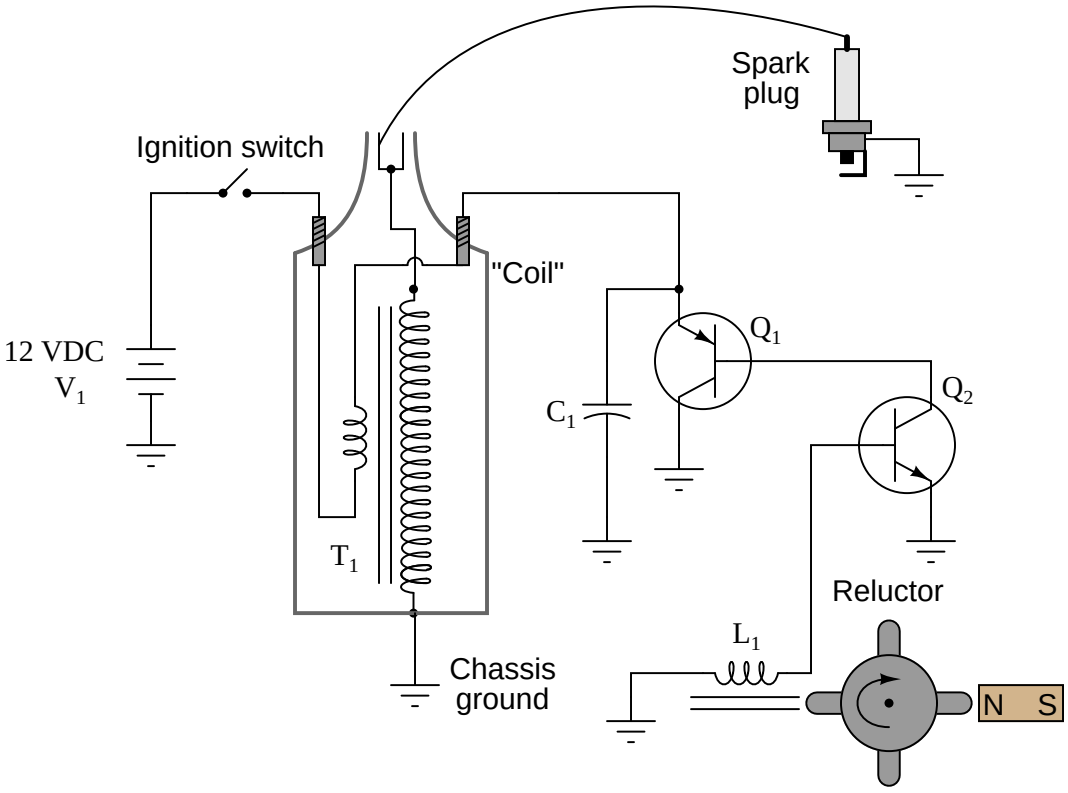
For each of these conditions, explain *why* the resulting effects will occur.

Challenges

- Explain the function of the *reluctor* assembly.

7.3.7 Failed double-transistor ignition system

Predict how all component voltages and currents in 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_2 fails open (collector-to-emitter):
- Transistor Q_1 fails shorted (collector-to-emitter):
- Capacitor C_1 fails shorted:
- Capacitor C_1 fails open:
- Reluctor coil L_1 fails open:

For each of these conditions, explain *why* the resulting effects will occur.

Challenges

- Explain the purpose for using multiple transistors in this circuit.

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

Biard, James R., *US Patent 3,463,975*, "Unitary Semiconductor High Speed Switching Device Utilizing a Barrier Diode", application 31 December 1964, patent granted 26 August 1969.

"IRF9610, SiHF9610 Power MOSFET datasheet", document number 91000, Vishay InterTechnology Incorporated, 8 February 2017.

Mays, Lonne, "FET vs. BJT vs. IGBT: What's the Right Choice for Your Power Stage Design?", AllAboutCircuits.com, 13 March 2017.

Pierret, Robert F., *Semiconductor Device Fundamentals*, Addison-Wesley Publishing Company, 1996.

Skovmand, Tim, "Micropower High Side MOSFET Drivers", Application Note 53, Linear Technology Corporation, Milpitas, CA, January 1993.

Watson, Joseph, *Analog & Switching Circuit Design*, 2nd Edition, John Wiley & Sons, Inc., 1989.

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.

5 May 2026 – added a new Conceptual Reasoning question on quad Darlington array ICs and their use with inductive loads, and also a comment in the "Switching inductive loads" section of the Tutorial chapter on commutating diodes built into power transistor array ICs.

4 November 2025 – multiple typographical error corrections, courtesy of Wyatt Gleason. Also edited an instructor note.

9 September 2025 – added more Challenge questions.

9 April 2025 – added a new Derivations and Technical References chapter.

30 March 2025 – added another bullet-point to the "Challenging concepts related to transistor switching circuits" section of the Introduction chapter.

19 November 2024 – edits made to the "Electronic ignition circuit" Conceptual Reasoning question. Also added a new Diagnostic Reasoning question exploring incorrect use of a MOSFET in a switching circuit, and re-named the BJT version of this same problem to include "BJT" so as to differentiate them from each other. Also edited multiple instructor notes and illustrations.

12 November 2024 – minor edits to the Tutorial.

5 November 2024 – minor edits to the Tutorial.

2 November 2024 – added a new Animations chapter with an animation showing the injection of charge carriers in an NPN transistor with applied base-emitter voltage.

1 November 2024 – added a new Animations chapter with an animation showing the development

of an inversion layer in an E-type MOSFET with applied gate-source voltage.

11 September 2024 – added a reference to the reader's Journal to the "Transistor summary" Conceptual Reasoning question. Also added "flyback diode" as a synonym for commutating diodes.

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

7 February 2024 – corrected a written typo ("Is is") and a graphical typo (image_2258 and image_2259).

12 November 2023 – elaborated on the mechanisms of thermal runaway for BJTs and thermal stability for FETs.

7 November 2023 – added polarity markings to image_6426 and also clarified the mechanism of charge carrier injection (for BJTs).

13 September 2023 – minor edits to the Tutorial elaborating on transistor saturation, namely the technique of flooding a transistor with a super-abundance of charge carrier so it becomes as good of a conductor as it possibly can.

22 June 2023 – added a requirement in the "Explaining the meaning of calculations" Conceptual Question to also express each quantity using appropriate metric prefixes.

11 May 2023 – added a new Quantitative Reasoning question, for students to simulate a BJT switching circuit using computer programming.

11-13 April 2023 – added questions to the Introduction chapter. Also edited image_5719 and image_5721 to include the MOSFET's parasitic diode as an aid to properly annotating the correct direction of drain-source current.

8 April 2023 – elaborated more on how to design a switching circuit to ensure transistor saturation.

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

9 November 2022 – minor edits to the Case Tutorial chapter.

7 November 2022 – minor clarifying edits to the Tutorial chapter, including a new characteristic curve showing saturated versus active modes for a generic transistor.

5 November 2022 – added a new section to the Tutorial chapter discussing saturation.

4 May 2022 – added "ballasting" as a term to describe the use of swamping resistors to equalize current in paralleled BJTs.

14 April 2022 – corrected some typos in the "Identifying BJT states" question where all the supply voltages were labeled V_{DD} rather than V_{CC}.

10 February 2022 – added review of MOSFET parasitic diodes to the Full Tutorial.

9 February 2022 – added more Challenge questions as well as additional sub-questions to the "Explaining the meaning of calculations" Conceptual Reasoning question.

11 November 2021 – added two new Case Tutorial sections showing how to sketch transistor switching circuits.

16 September 2021 – added a query for students in the "Proper BJT types" Conceptual Reasoning question, to determine which circuits were capable of transistor saturation.

28 August 2021 – challenged the reader to identify foundational concepts within some of the questions.

27 July 2021 – minor addition to the "Calculations in TIP31C transistor circuit" Quantitative Reasoning question.

13 May 2021 – divided Tutorial into sections, and edited image_2228.

9 May 2021 – commented out or deleted empty chapters.

12 March 2021 – added a new Case Tutorial section showing MCU driving BJT driving MOSFET.

11 November 2020 – added a Case Tutorial chapter.

10 November 2020 – added more Challenge questions.

9 November 2020 – minor additions to the Introduction chapter.

8 November 2020 – minor edits to the Tutorial, and additions made to the "Calculations in TIP31C transistor circuit" and "Calculations in IRF9520 transistor circuit" questions.

23 September 2020 – added instructor notes to some questions, and added a new Conceptual question.

22 September 2020 – edited Tutorial to make more clear the states of toggle switches in the example schematics, and made other minor edits to the text.

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

22 June 2020 – corrected error in Sziklai-pair connections question.

20 February 2020 – added another Quantitative Reasoning problem ("Saturated versus unsaturated BJT switch").

19 February 2020 – added some new Quantitative Reasoning problems ("Identifying BJT states", etc.).

15 February 2020 – added a new Conceptual Reasoning problem ("Electron versus hole mobility").

11 February 2020 – added a new Quantitative Reasoning problem ("Over- and under-driving a BJT switch").

10 February 2020 – added some more sub-questions to the "Voltage calculations in a simple lamp circuit" Quantitative Reasoning problem.

24 January 2020 – added Foundational Concepts to the list in the Conceptual Reasoning section.

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

20 October 2019 – added patent number to the Historical Reference section on James Robert Biard's invention of using a barrier diode to improve high-speed transistor switching.

9 October 2019 – added Questions.

8 October 2019 – finished writing Tutorial. Also added an entry to the Historical References chapter.

7 October 2019 – document first created.

Index

74S00 series digital logic, 47

Active mode, 27
Adding quantities to a qualitative problem, 142
Amplification, 21
AND logic gate, 47
Annotating diagrams, 141

Ballasting resistor, 39
Base, 22
Bipolar, 22
Bipolar junction transistor, 21
BJT, 21
Bridge, 40

Charge carrier injection, BJT, 22
Charge pump, 34
Checking for exceptions, 142
Checking your work, 142
Code, computer, 151
Collector, 22
Commutating diode, 37
Commutation diode, 112
Conservation of Energy, 36
Conventional flow notation, 31
Coulomb, 52

Darlington pair, 110, 122
Dimensional analysis, 141
Diode, commutating, 37
Diode, commutation, 112
Diode, flyback, 37
Diode, free-wheeling, 37
Diode-Transistor Logic, 47
Drain, 23
DTL, 47

Edwards, Tim, 152

Electric field, 51
Electrically common points, 32, 54
Electromagnet, 36
Electron flow notation, 31
Electrostatic shielding, 54
Emitter, 22
Equilibrium, 54
Equipotential, 54

Feedback, negative, 39
FET, 21
Field, electric, 51
Field-effect transistor, 21
Floating, 41
Flyback diode, 37, 112
Free-wheeling diode, 37

Gate, 23
Gauss' Law of Electric Fields, 51
Gaussian surface, 51
Graph values to solve a problem, 142
Greenleaf, Cynthia, 87

H-bridge, 40
High-side switching, 31
How to teach with these modules, 149
Hwang, Andrew D., 153

Identify given data, 141
Identify relevant principles, 141
Inductive kickback, 112
Inductive load, 36
Inductor, 36
Injection, charge carriers in BJT, 22
Insulated-gate field effect transistor, 23
Intermediate results, 141
Intrinsic semiconductor, 26
Inversion layer, 77

Inverted instruction, 149
Inverter gate, 111
Inverter logic gate, 47

JFET, 23
Junction field-effect transistor, 23

Kirchhoff's Voltage Law, 33
Knuth, Donald, 152

Lamport, Leslie, 152
Limiting cases, 142
Load, 36
Logic gate, 111
Low-side switching, 31

Magnetic field, 36
Majority charge carriers, 38, 39
Maxwell, James Clerk, 43
Metacognition, 92
Metal-oxide-semiconductor field-effect transistor, 23
Minority charge carriers, 32, 33, 38
Moolenaar, Bram, 151
MOSFET, 23, 54
Murphy, Lynn, 87

NAND logic gate, 47
Negative feedback, 39
Newton, 52
NOT gate, 111
NOT logic gate, 47

Ohm's Law, 26, 27, 33
Open-source, 151
Overdrive, transistor, 29

Pinch-off voltage, 21
Planar transistor, 46
Problem-solving: annotate diagrams, 141
Problem-solving: check for exceptions, 142
Problem-solving: checking work, 142
Problem-solving: dimensional analysis, 141
Problem-solving: graph values, 142
Problem-solving: identify given data, 141
Problem-solving: identify relevant principles, 141
Problem-solving: interpret intermediate results, 141
Problem-solving: limiting cases, 142
Problem-solving: qualitative to quantitative, 142
Problem-solving: quantitative to qualitative, 142
Problem-solving: reductio ad absurdum, 142
Problem-solving: simplify the system, 141
Problem-solving: thought experiment, 141
Problem-solving: track units of measurement, 141
Problem-solving: visually represent the system, 141
Problem-solving: work in reverse, 142

Qualitatively approaching a quantitative problem, 142

Reading Apprenticeship, 87
Reductio ad absurdum, 142, 148, 149
Resistor, ballasting, 39
Resistor, swamping, 39
Runaway, thermal, 38

Saturation, 10–13, 18, 20, 32–34, 39, 41, 44
Saturation mode, 28
Schoenbach, Ruth, 87
Scientific method, 92
Shielding, electrostatic, 54
Simplifying a system, 141
Sinking, 31
Socrates, 148
Socratic dialogue, 149
Source, 23, 36
Sourcing, 31
SPICE, 87
Stallman, Richard, 151
Storage time, BJT, 33
Swamping resistor, 39
Sziklai pair, 110

Thermal runaway, 38
Thought experiment, 141
Threshold voltage, MOSFET, 34
Time, storage, 33
Torvalds, Linus, 151
Transistor-to-Transistor Logic, 47

TTL, 47
Turn-on voltage, 21

Units of measurement, 141

Visualizing a system, 141
Voltage, pinch-off, 21
Voltage, turn-on, 21

Work in reverse to solve a problem, 142
WYSIWYG, 151, 152