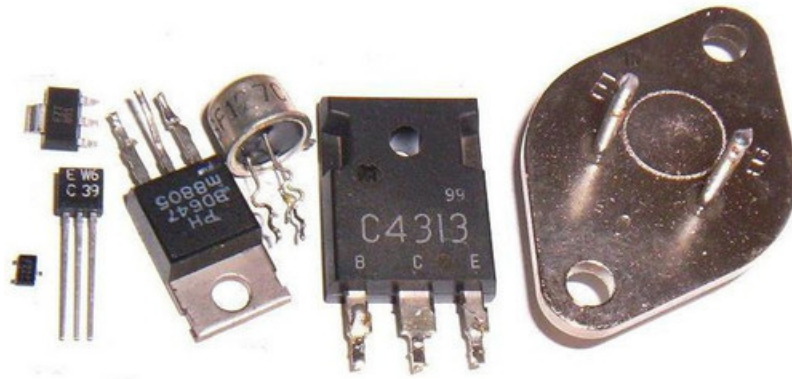


## Transistors 101

Created by Dave Astels

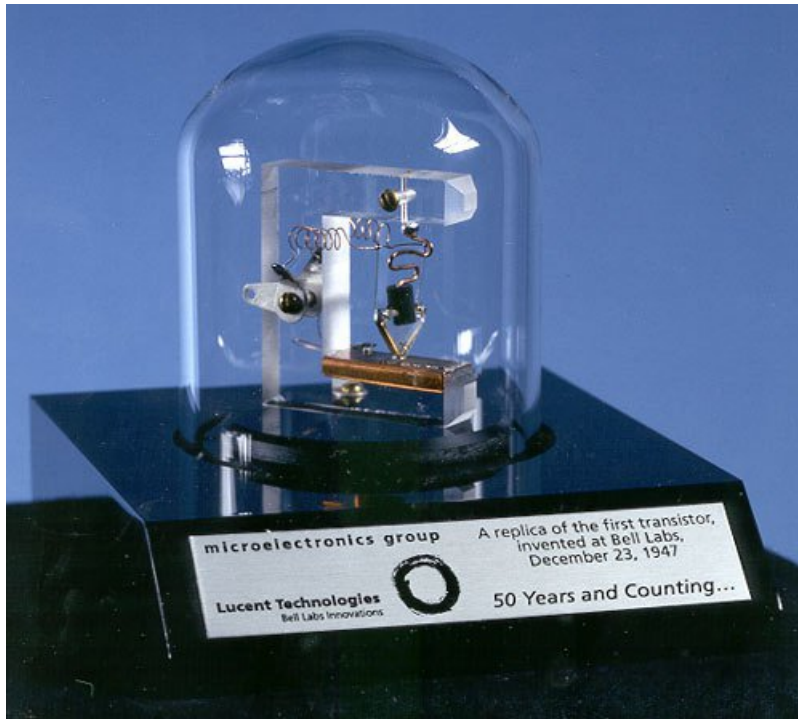


Last updated on 2018-09-12 08:31:52 PM UTC

## Guide Contents

Guide Contents	2
Overview	3
Basic Operation	4
NPN and PNP: What?!	6
Example uses	7
Darlington Pairs	10
TIP120 Power Darlington Transistors - 3 pack	11
ULN2803: 8 Channel Darlington Driver (Solenoid/Unipolar Stepper)	12
Open Collector	13
FETs	15
Examples	16
Driving a LED Strip	16
Flexible Silicone Neon-Like LED Strip in Various Colors	16
Electromagnets, Solenoids, and Motors	17
1N4001 Diode - 10 pack	18
Motor H bridge	19
FAQ	21
How do I pick a transistor for the job?	21
How do I choose a bias resistor(s)?	21
How do I figure out how much heat will be generated?	21
How do I figure out what heat sink I'll need (if any)?	21
TO-220 Clip-On Heatsink	22

## Overview



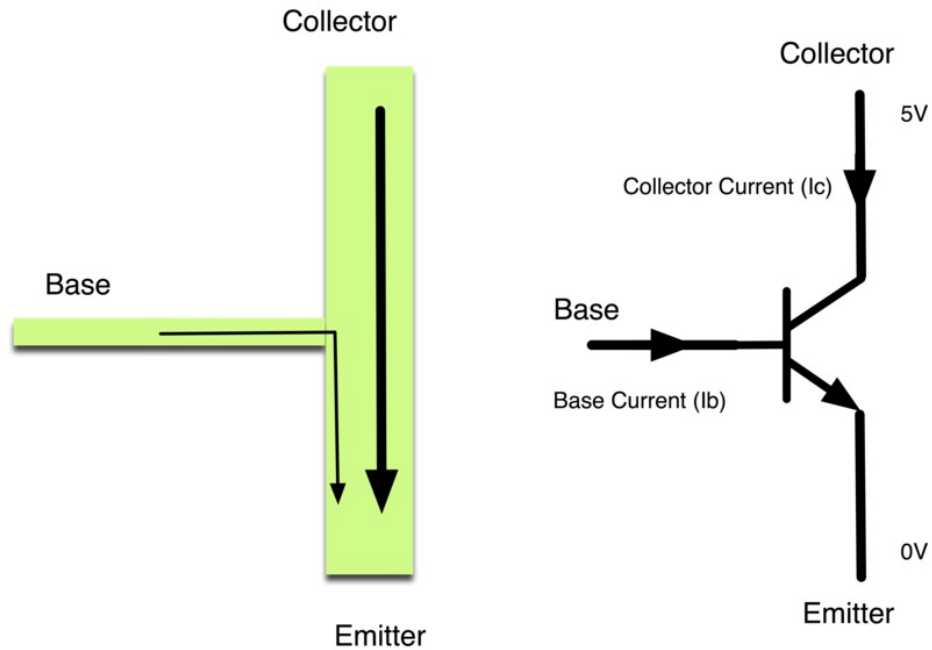
At their core, transistors amplify power. Note that: power. You can use a transformer to amplify either current or voltage, but you have to trade off the other. i.e. you can get more voltage but at a lower current, or more current, but at a lower voltage. Transistors can give you an increase in both.

This proves to be extremely useful - who *doesn't* want more power? Transistors are very truly the basis of modern electronics, but we'll be looking at a specific use in this guide: controlling high power circuits/devices with a low power microcontroller.

This guide will focus on Bipolar Junction Transistors, aka BJTs. Bipolar because they use two kinds of silicon. Junction because pieces of those two different kinds of silicon are up against each other.

## Basic Operation

The central idea to the operation of a transistor is that a little bit of current flowing between the base and emitter causes a larger current to flow between the collector and emitter.



A transistor is essentially an amplifier. In fact the initial tests of the first transistor involved hooking a speaker to its output and hearing that it was louder than the input. Every modern piece of audio circuitry has transistors at its heart amplifying signals.

A transistor will act linearly if the base-emitter current is within a specified range (depending on the model of transistor). That means that the collector-emitter current will be some multiple of the base-emitter current. That multiplier is called the transistor's gain. For example, the gain of a 2N3904 (a very common small transistor) is typically around 100. The gain can vary between specific parts even of the same part number. Additionally the exact value of the gain depends on the collector-emitter current. The result is that the value of the gain isn't all that important. The key thing is that it's there, and it is often a couple orders of magnitude, especially with fairly high collector-emitter currents.

Another nice feature of transistors is that they can be driven into *cutoff* and *saturation*. When this happens they stop acting linearly and act non-linearly: in cutoff they act like an open switch, and in saturation they act like a closed switch. This switching behavior is incredibly useful, and is the foundation of digital electronics.

Using a transistor as a switch has some advantages;

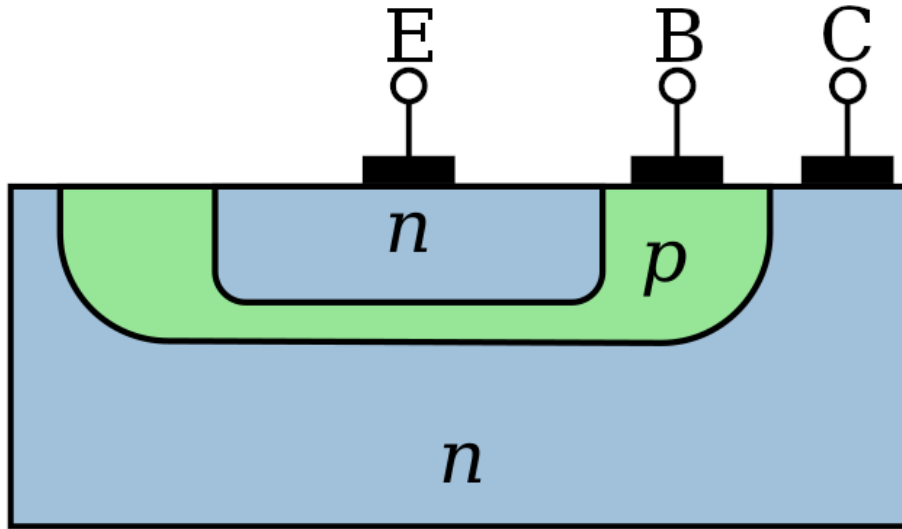
- A circuit that can provide only a small current (e.g. from a GPIO pin) can control a much larger current (e.g. a motor).
- Switching state happens very quickly (the 2N3904 takes at most 50 ns - about 5 hundred-thousandths of a millisecond). That's a general purpose part; some are much faster.
- Because of the above point, transistors can be switched off & on at a very high frequency... high enough to make modern computer speeds possible.
- A transistor switch can be controlled by a pulse-width modulated (PWM) signal, something that a mechanical switch can't do.

- Transistors don't bounce like a mechanical switch.
- A single control signal can control multiple circuits as each requires only a small current.

## NPN and PNP: What?!

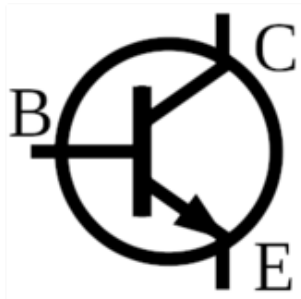
NPN and PNP refer to the arrangement of the pieces that make up the transistor. The practical result is the direction of current flow.

A bipolar junction transistor is made up of three pieces of silicon. Depending on what is added to the silicon, it will be either N-type or P-type. An NPN transistor has a piece of P-type silicon (the base) sandwiched between two pieces of N-type (the collector and emitter). In a PNP transistor, the type of the layers are reversed. Below is a typical cross section of a transistor.

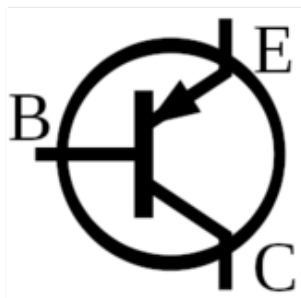


---

NPN and PNP transistors have very similar schematic symbols. The only difference is the direction of the arrow on the emitter. In a NPN (on the left) it points outward, for a PNP (on the right) it points inward.



You can interpret that arrow in a couple ways: it's the direction of positive current flow (opposite of electron flow), and also as pointing toward the lower voltage when the transistor is switched on, i.e. for an NPN the emitter has to be at a lower voltage than the base/collector in order for it to conduct, whereas for a PNP it has to be higher.

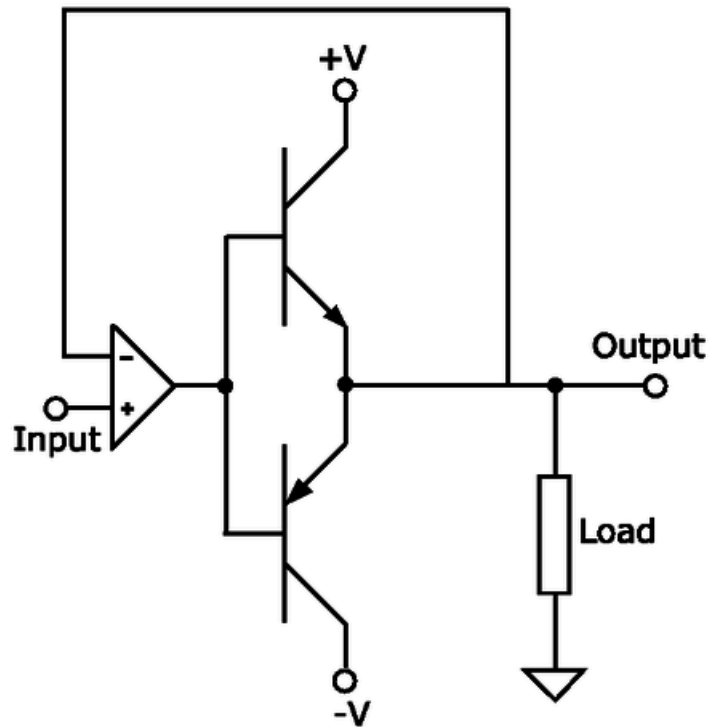


---

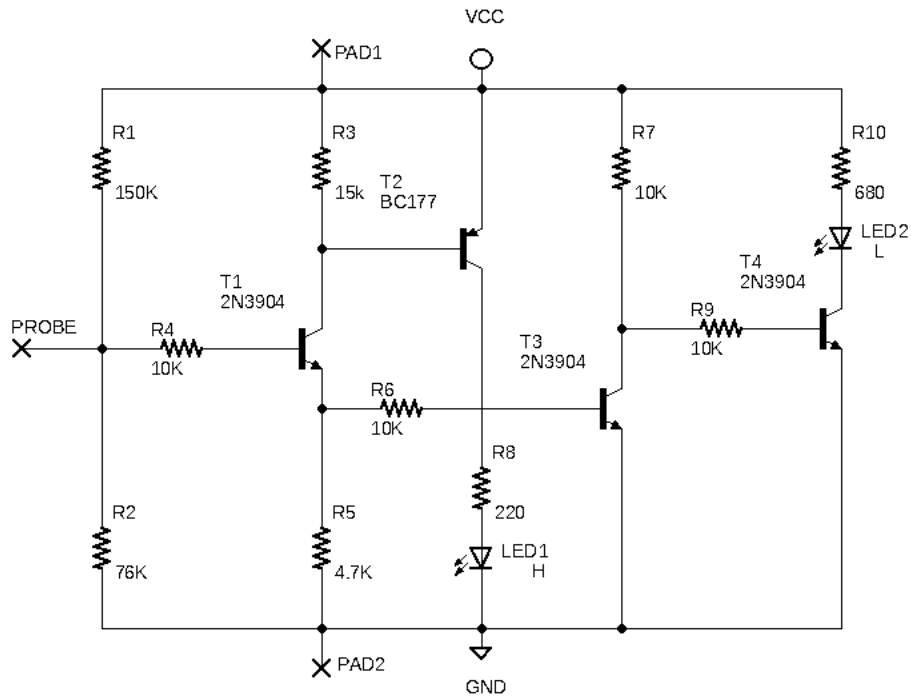
## Example uses

Since a PNP transistor works more or less the opposite of an NPN, it can be used where negative voltages are involved. One such example is in amplifying an AC signal\* for driving a speaker. A speaker moves air to create sound. While you can use a speaker with a DC signal (that varies between 0 and some positive voltage) that pushes the speaker cone outward (and relax to it's neutral position) to create a sound wave, you can create louder sound if you also pull the cone back.

Below is a simplified example of just that. when the input is positive, the NPN transistor conducts and drives the output toward +V. When the input is negative, the PNP transistor conducts, driving the output toward -V.



In my digital electronics tools learning guide is a [DIY logic probe](https://adafru.it/Cox) that uses a PNP transistor to drive one of the LEDs. When the **PROBE** input is high, **T1** switches on which, because of the resistors making up the voltage divider on its collector and emitter, drops the voltage at the base of **T2** enough below its emitter (at **VCC**) to cause it to switch on, thus turning on **LED1**.



\* An AC signal is one where the voltage can be positive and negative, and is typically constantly changing.



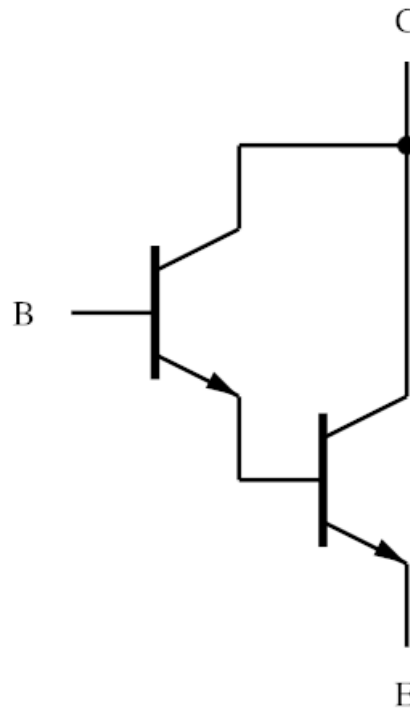


## Darlington Pairs

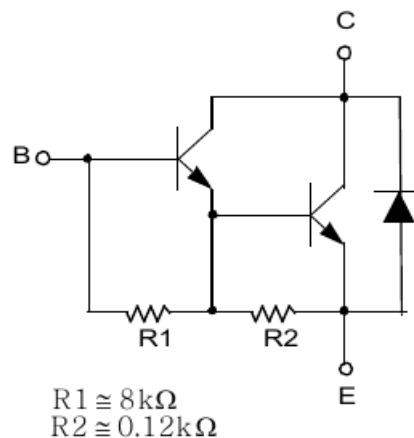
Sometimes you want to control a really large amount of current; more than a single transistor's gain can provide. I.e. the output current has to be more than the input current multiplied by the transistor's gain.

In this case we can combine two transistors in what is called a Darlington connection. In this configuration the gain of the pair is very slightly more than the product of the gain on the individual transistors and can be 1000 or more. Specifically, its the product of the gains plus the sum of the gains, but the product is usually high enough that the sum can be ignored for calculations.

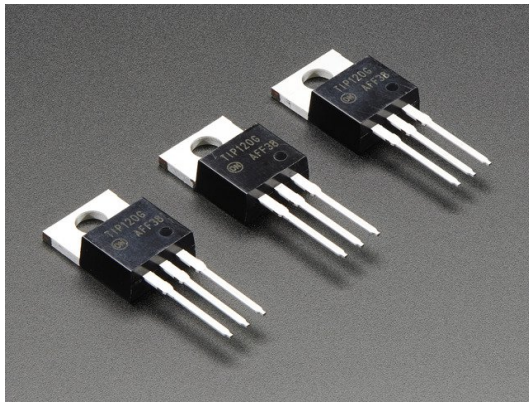
At its core, a Darlington pair is simply two transistors cascaded together:



A common Darlington pair is the TIP120 which is a great choice for the examples later in this guide. It can handle up to 5A continuously (with bursts of 8A) and up to 60V across the collector and emitter. It has a gain of ~1000.



A TIP120 is just the circuit above packaged in a convenient 3-pin package (base, collector, and emitter). Since it's meant for power switching/amplification it comes with a metal backplate to which a larger heatsink can be attached.

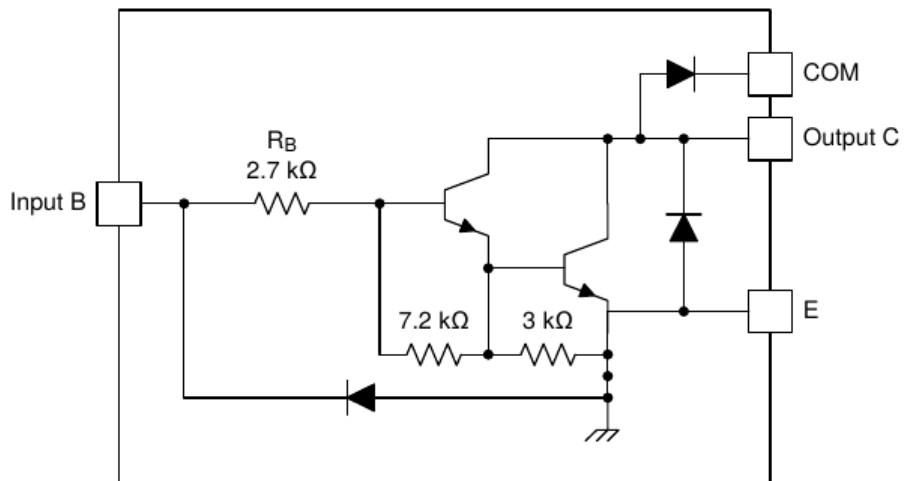


TIP120 Power Darlington Transistors - 3 pack

\$2.50  
OUT OF STOCK

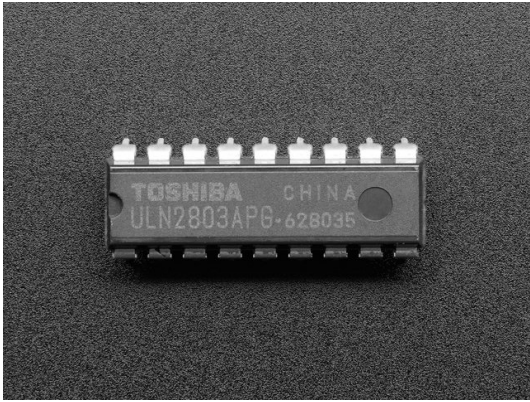
OUT OF STOCK

The "drive" connections of the CRICKIT are, in fact, darlington drivers. They are all on a single IC, called a **ULN2803**, with 7 or 8 Darlington's each with the following design:



This chip is very convenient to use as it provides everything needed: all the biasing resistors, and all the protection diodes needed for inductive loads such as electromagnets, solenoids, and relays.

The CRICKIT uses 4 of these drivers, though the chip contains 7. There are a variety of versions of this design, some with 8 drivers in a slightly larger chip.



ULN2803: 8 Channel Darlington Driver (Solenoid/Unipolar Stepper)

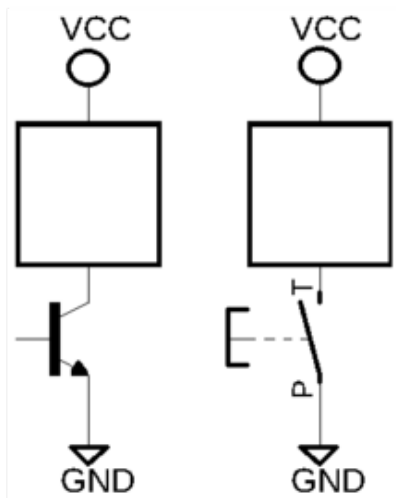
\$1.95  
IN STOCK

ADD TO CART

## Open Collector

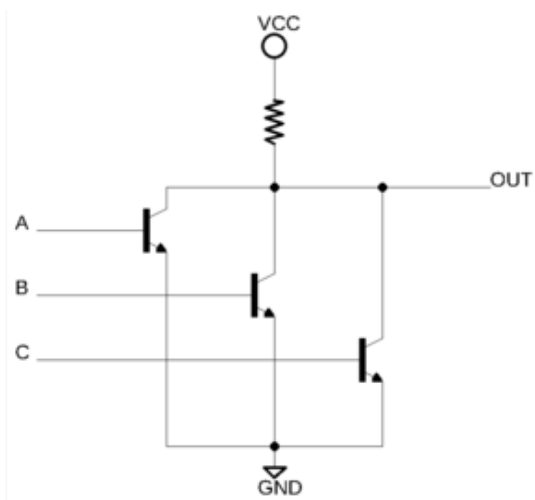
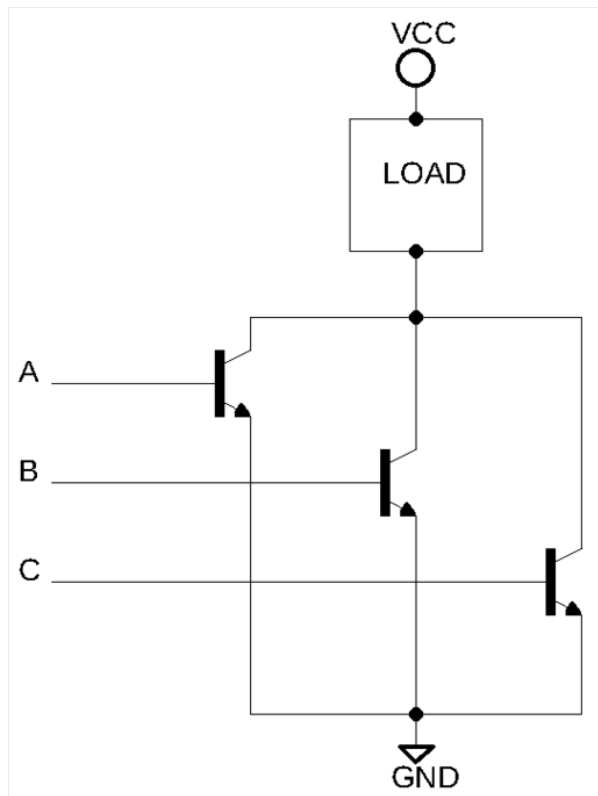
Rather than output a signal we often want to provide a current sink. This is common when controlling power to a circuit.

An open-collector output is used to connect one side of the device being controlled to ground. The other side of the device will be connected to power. If you look at the CRICKIT "drive" outputs you will see that this is exactly what they are. You connect the positive side of your device (a high power LED, solenoid, etc) to the 5v connection and the negative side to a drive output.



Electrically the transistor in an open collector configuration acts just like a mechanical on-off switch (except with the benefits listed previously).

In [another guide \(https://adafru.it/Coy\)](https://adafru.it/Coy) we see another use of open collector connections: when the transistor switch is in its off state, the load is disconnected. One advantage of this is that multiple open collector connections can be tied together and any one can turn on the load independent of the others, operating like an OR gate.



If we add a pull-up resistor instead of a load we get a logic circuit; each transistor functions as a tri-state inverter, together making up a (in this case) 3-input NOR gate. If A, B, and C are all low, none of the transistors are conducting (i.e. on) which means the output is high due to the pull-up resistor. If any on the inputs is high, the corresponding transistor is switched on, connecting the output to ground.

## FETs



No. Not that Fet.

FET as in Field Effect Transistor.

It's a different way to build a transistor. A way that has some distinct advantages. One major advantage for the usage we've been looking at is efficiency. FETs don't waste as much power, which means they don't heat up as much as bipolar transistors. Another is that to control them takes a voltage and not a current, unlike a BJT (which, if you recall, is controlled by the base-emitter current).

Because of their lower power/current characteristics, they are ideal to cram into a small space. This makes them perfect for building large digital chips.

We won't go into any more detail, but it's worth considering using a FET for high current applications or where you want to avoid the wasted power and/or generated heat.

## Examples

### Driving a LED Strip

A simple LED strip is a great example of a load. You simply control whether current is flowing through it. All you need to be concerned with is whether the control circuit can supply enough current.

These neon-like LED strips 'look' like one very big, high power LED - and they need 12V to light up!



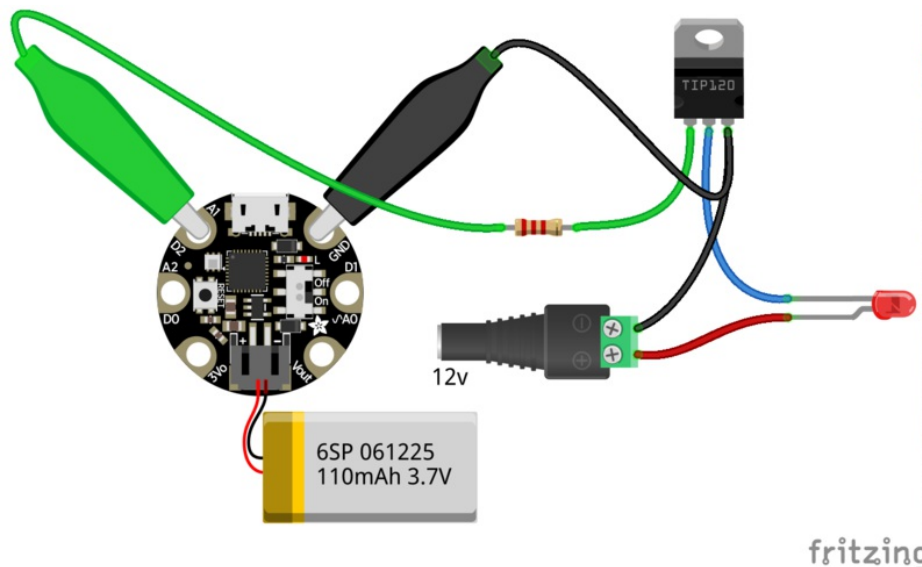
Flexible Silicone Neon-Like LED Strip in Various Colors

\$13.95  
OUT OF STOCK

OUT OF STOCK

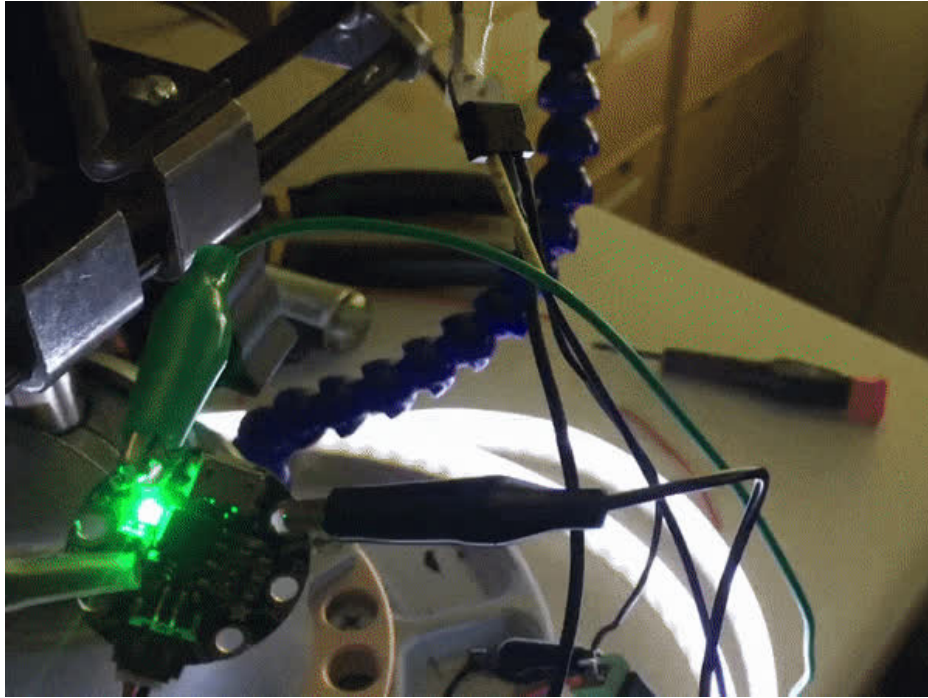
Here we have a Gemma M0 using a PWM output (a 3.3v signal that can provide a handful of milliamps) to drive a TIP120 Darlington Pair to control an led strip (no Fritzing part for it yet, so a single LED is used as a proxy). The strip takes 12v and 0.8A... aka 800 mA.

The gain of the TIP120 is about 1000, so to get 0.8A through the strip, we need a base current of 0.8mA. Erring on the side of being generous, a 2.2K bias resistor should give a bit over 1mA of base current. That's plenty to drive the TIP120 into saturation.



Here's the circuit in action. At the top you can see the TIP120. In the bottom-left is the GemmaM0 with the ground and control connected by the black and green clips, respectively. The other alligator clip is an insulated 3rd hand holding the Gemma. On the bench below is the LED strip which fades off and on due to the ramping PWM signal.





The code running this demo is:

```
import time
import board
import pulseio

led = pulseio.PWMOut(board.D2, frequency=5000, duty_cycle=0)

while True:
    for i in range(100):
        if i < 50:
            led.duty_cycle = int(i * 2 * 65535 / 100) # Up
        else:
            led.duty_cycle = 65535 - int((i - 50) * 2 * 65535 / 100) # Down
        time.sleep(0.05)
```

## Electromagnets, Solenoids, and Motors

These are electrically very similar: they are all inductive loads (made of a coil of wire) that can draw a moderate to high amount of current. As such they can't be controlled directly from a GPIO pin. This is where transistors come in.

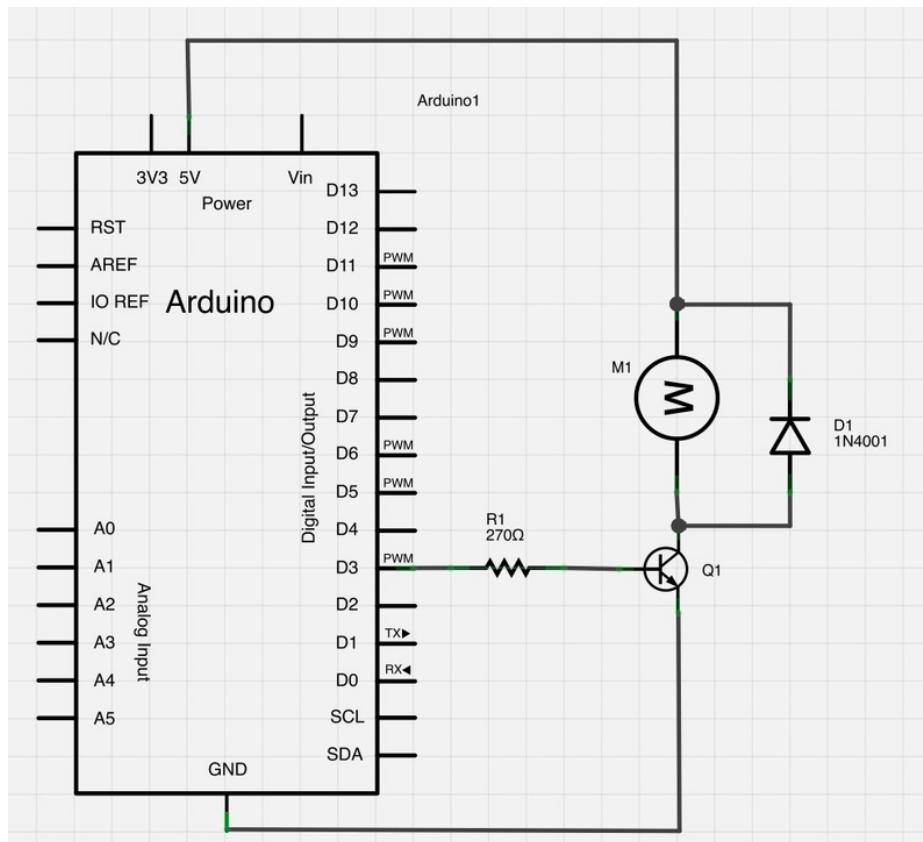
An electromagnet is a coil of wire around a core of magnetic material. When current flows through the coil, a strong magnetic field is generated. I.e. it becomes a magnet. When the current stops, it ceases to be a magnet.

A solenoid is basically an electromagnet in which the core can slide back and forth inside the coil. When current flows through the coil, the core moves. It can be used to create linear motion from a current. Ever have a electric model railroad? The switches points were moved by a solenoid.

A motor is an arrangement of coils and permanent magnets such that controlling the current through the coils (done inside the motor) causes the motor's shaft to spin, thus creating rotary motion from a current.

For our purpose of turning them on and off, we can think of them as all just being an inductive load.

Simon Monk did [a guide on controlling a DC motor with an Arduino](https://adafru.it/dtG) (https://adafru.it/dtG). Naturally, this used a transistor.



Since a transistor is a switch, it can be driven with a PWM signal. With a motor, a higher duty cycle results in a faster rotation, i.e. higher speed.

It's important to note the diode. When a motor stops, the magnetic field collapses and induces a reverse current. That can be enough to burn out the transistor. The diode serves to drain off that current harmlessly. Whenever you are controlling a motor or other inductive load, you must have a diode across it like this.



1N4001 Diode - 10 pack

\$1.50  
IN STOCK

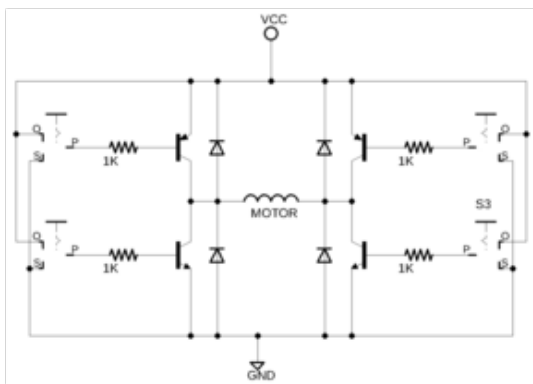
ADD TO CART

Not including the diode can destroy your microcontroller, not just the transistor!

## Motor H bridge

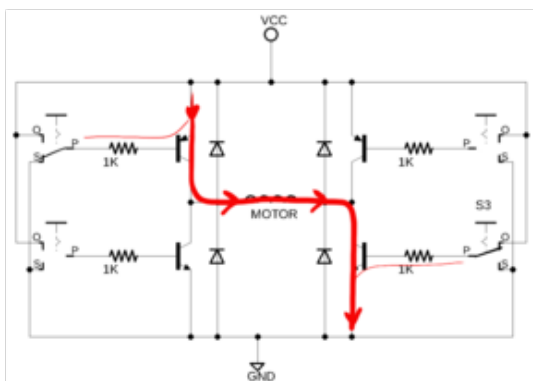
The above motor controller can turn the motor on and off, as well as control its speed, but it can only make it spin in one direction.

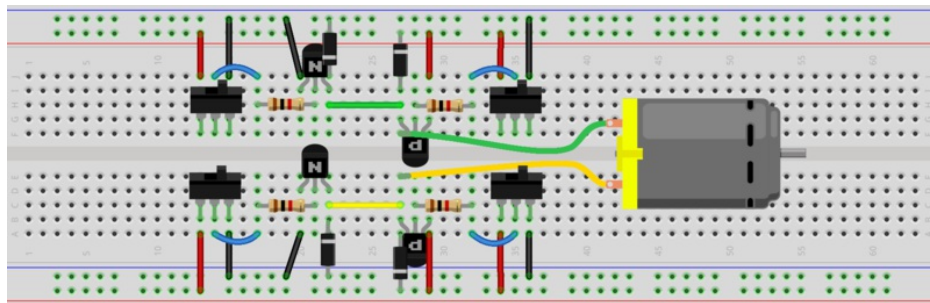
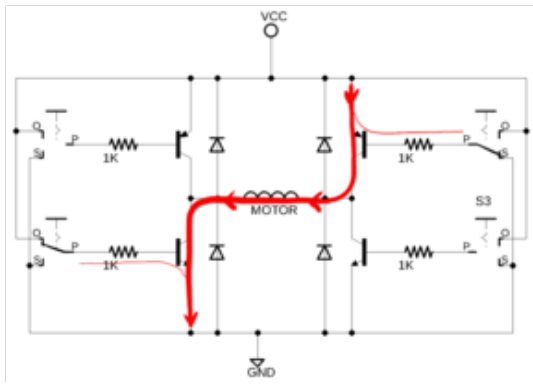
To control the direction of the motor as well, it can be driven using an arrangement of transistors called an H-bridge, shown below. Notice the resistors between each input switch and transistor base. These limit the current flowing through the base to just enough to drive it into saturation. Also, there is a diode between the pairs of collectors and power & ground. These allow the current generated when the motor stops to be drained away harmlessly. Without them the reverse current could blow out the transistors, as mentioned earlier.



Here's the schematic of an H-Bridge motor driver. Notice the use of PNP Transistors at the top (which connect the motor to Vcc) and NPN transistors at the bottom (which connect it to ground). By enabling diagonally opposite pairs of transistors the motor can be spun in either direction by controlling the direction which current flows through it.

Below is a breadboard layout.





fritzing

That's a fair number of components which take up a fair bit of space. Thankfully, you can get two of these packaged neatly onto a 16-pin chip: the DRV8833 which is used on the CRICKIT. Technically the DRV8833 uses FET transistors but the operation of the circuit is identical.

One handy feature of an H-bridge is that not only can it control forward and reverse operation of the motor, but it also provides coast (the motor spins freely) and brake (the motor resists rotation) modes.

See [this post on robotroom.com \(https://adafru.it/Coz\)](https://adafru.it/Coz) and the [DRV8833 datasheet \(https://adafru.it/CoA\)](https://adafru.it/CoA) for more details.

## FAQ

I let people know on the Adafruit Discord that I was working on this guide and asked for questions. here are some of them.

### How do I pick a transistor for the job?

The first thing to consider is the amount of current being controlled. If it's no more than a couple dozen mA (e.g. for a simple LED) then something like a 2N2222/2N3904 will be fine. For a higher current load like a LED strip or a solenoid, a Darlington pair like a TIP120 is a better choice. For high current control, consider using a FET to minimize the waste and heat.

### How do I choose a bias resistor(s)?

Assuming you want to use the transistor as a switch, then you want the bias resistor to ensure that the transistor is saturated when the input is high.

For an example, the load you are controlling requires 200mA; that's the collector current  $I_c$ . Further let's use a typical transistor (such as a PN2222) that has a gain of  $\sim 100$  at that collector current. Looking at the [datasheet \(https://adafru.it/CoB\)](https://adafru.it/CoB) we see that the *Base-Emitter Saturation Voltage* (on page 2) is 0.6v. That will cause the transistor to saturate. If we control this with a 3.3v microcontroller then the bias resistor needs to have  $3.3v - 0.6v = 2.7v$  across it. With an  $I_c$  of 200mA and a gain of 100, we have a base current  $I_b$  of  $I_c/gain = 200mA/100 = 2mA$ . This is within the range of a typical GPIO pin's current rating\*. We will need a resistor of  $2.7v/2mA = 2.7/0.002 = 1350$  ohms. 1.2K ohms is a common value, and using a slightly lower bias resistor than the calculation indicates will result in a slightly higher base-emitter current and guarantee that the transistor is saturated when the input is high.

If you're not sure, or you don't want to do the math, a 1K resistor is a great place to start. Reduce it if you need more current. Increase the resistor ohms if you want less current.

### How do I figure out how much heat will be generated?

Heat generated by a transistor is measured in watts (W) and is computed by multiplying the saturation voltage of the transistor (across the collector/emitter pins) by the collector-emitter current. The TIP120 used in the earlier example has a saturation voltage 2v when the current is 3A (and 4v when it's 5A). This information will be on the transistor's datasheet. So if we run 3A through it it will have a voltage drop of 2v. That's 6W of power being dissipated as heat.

In general, a TIP120 can dissipate about 2W without a heatsink but it will get *hot* - so pick up a heat sink!

### How do I figure out what heat sink I'll need (if any)?

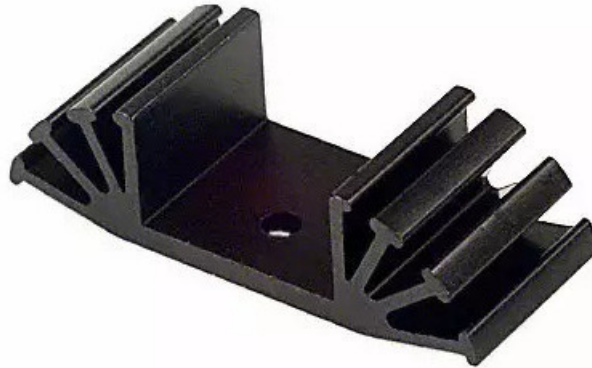
Let's consider the TIP120 again. It's maximum working case temperature is 150C. Aiming to keep the temperature below 100C is a reasonable goal. That's well below the limit.

To choose a heatsink we also need to know the ambient temperature where the transistor will be operating. 50C is a reasonable guess if the circuit is enclosed with a moderate amount of circuitry. If it's lower there isn't a problem (it's easier to get rid of excess heat). The difference between the transistor case temperature we are aiming for and the ambient is 50C. Again, reusing the above examples, the transistor dissipates 6W. To choose a heatsink, we divide the temperature difference by the dissipated power:  $50C/6W = 8.3$  C/W. This number is called thermal resistance. A heatsink needs to be selected that has a thermal resistance of about 8.3 C/W.

Since we chose 100C target which is well below the worst case, we can play around a bit with the exact thermal

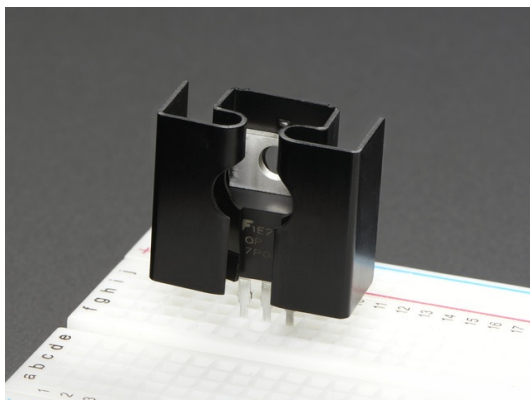
resistance of the heatsink to work with size and supply. This will also give us more options.

For example, we can find a TO220 specific heatsink at Digikey with a thermal resistance of 9.0, yet is still rated for 6W at 50C:



\* For example, a GPIO pin on the SAMD21 can provide up to 7mA.

These simple clip on heat-sinks will let you handle a few extra Watts, and are easy to use



TO-220 Clip-On Heatsink

\$0.75  
OUT OF STOCK

OUT OF STOCK