Transistor Basics

Lab 1: The Bipolar (Junction) Transistor

ECE 327: *Electronic Devices and Circuits Laboratory I*

Abstract

In the lab, we explore several common transistor circuits; we build a common-emitter amplifier, a *pnp* current source, an *npn* emitter follower, and class B and class AB (i.e., biased) push–pull amplifiers. Here, we introduce basic operating guidelines for bipolar transistors and outlines of these circuits.

Contents

Note about diode/bandgap conventions: We use the convention that a typical silicon BJT base–emitter diode drop is 0.65 V and a standard general purpose silicon diode drop is 0.6 V. Other conventions use 0.6 V or 0.7 V for one or both. Measured laboratory results will most likely be between these two values.

Diodes and BJTs implemented on the same *integrated circuit* (i.e., on the same piece of silicon) may have equivalent characteristics. That is, the diodes and transistors can be *matched*. Matched components are convenient to use in many circuit designs. We use *discrete* elements in this laboratory, and so it is not possible to match components unless they're all implemented within the one part. In the laboratory, a "diode-connected" transistor (i.e., shorted base and collector) may match the base–emitter characteristics of another transistor better than a diode.

Diode drops are strongly temperature dependent. Room-temperature transistors have base–emitter drops around 0.65 V, but hot transistors have drops near 0.5 V because they need less excitation for conduction. So temperature matching is just as important as component matching. Internal temperature compensation in "bandgap voltage references" lets them provide a temperature-independent voltage reference. Their output reference of ∼1.22 V is the Silicon diode drop at absolute zero (i.e., 0 K or −273.15°C). It is not a coincidence that the Silicon bandgap (i.e., the energy separating valence and conduction electron bands) is ∼1.22 eV.

Temperature dependence and manufacturing variations (and the "Early effect") are always a concern.

1 Bipolar Junction Transistor Model

A bipolar junction transistor (BJT) can be in three modes:

Trivial "switch" modes:

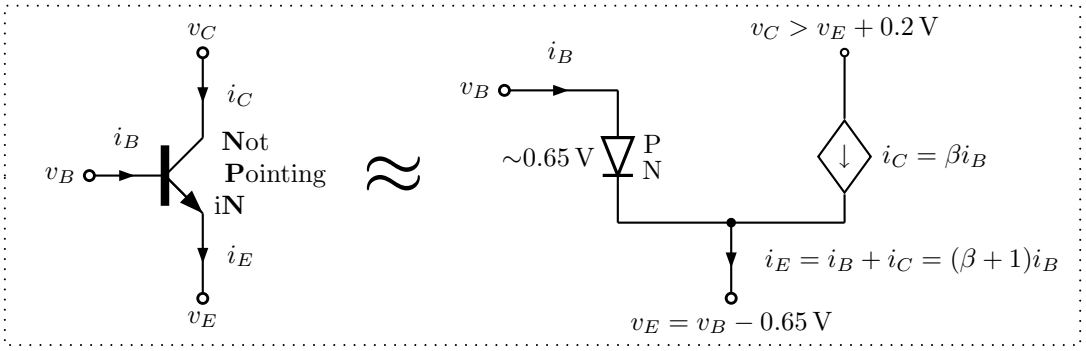
- *cutoff mode*: Transistor acts like an *open switch* between collector and emitter (i.e., collector–emitter "resistance" is infinite).
- *active mode*: Transistor acts like a *dynamic resistor* between collector and emitter that adjusts its resistance in order to keep collector current at a set level (i.e., collector–emitter resistance is finite and positive).
- *saturation mode*: Transistor acts like a *closed switch* between collector and emitter (i.e., collector–emitter "resistance" is very low).

In the *active mode*, the transistor adjusts the *collector* current to be a version of the *base* current amplified by some constant $\beta > 0$. If the base current falls to 0, the transistor enters *cutoff mode* and shuts off. When the base current rises too far, the transistor loses its ability to decrease the collector–emitter resistance to linearly increase the collector current. In this case, the transistor enters *saturation mode*. To keep the transistor out of saturation mode, the collector and emitter should be separated by at least 0.2 V.

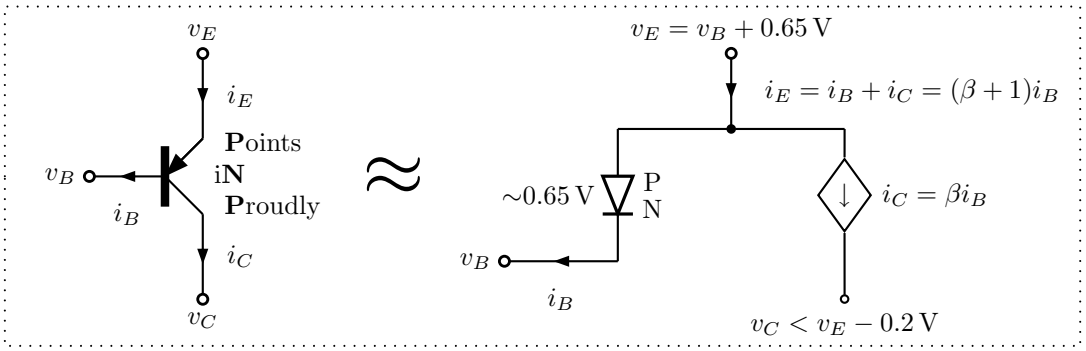
A simple model for the operation of a transistor in *active mode* is shown in Figure 1.1. It requires knowing the current gain β in order to design the circuit. In both of these models,

$$i_C = \beta i_B \quad \text{and} \quad i_E = (\beta + 1) i_B,$$

and the emitter is separated from the base by a diode. In order for this diode to conduct current, it must be forward biased with 0.65 V[1].



(a) Active mode *npn* BJT.



(b) Active mode *pnp* BJT.

Figure 1.1: Simple models of active mode bipolar junction transistors (BJTs).

[1] This is approximate. In general, the value will be between 0.6 V and 0.7 V, and may rise to 0.8 V in saturation mode.

2 The Ideal Bipolar Junction Transistor

Because the current gain β is typically unknown or varies greatly with temperature, time, collector–emitter potential, and other factors, good designs should not depend on it. In this laboratory, we assume that β is sufficiently large (i.e., $\beta \gg 1$, where $\beta \approx 100$ in our laboratory) so that

$$i_B \approx 0 \qquad \text{and} \qquad i_C \approx i_E.$$

These simple rules are similar to the rules we use with operational amplifiers. The analysis approach usually follows these steps:

1. Calculate the transistor base potential v_B by assuming that no current enters the base (i.e., $i_B \approx 0$).
2. Calculate the potential v_E at the emitter of the transistor using v_B. For an *npn* transistor,
$$v_E = v_B - 0.65\,\text{V},$$
and for a *pnp* transistor,
$$v_E = v_B + 0.65\,\text{V}.$$
3. Calculate the emitter current i_E using the emitter voltage v_E and the rest of the circuit.
4. Assume that $i_C \approx i_E$ and analyze the rest of the circuit.
   - Because we know v_E, we usually know i_E as well. So our i_E dictates what i_C should be.

However, keep these notes in mind.

- For an *npn* transistor, active mode requires $v_C - v_E > 0.2\,\text{V}$. For a *pnp* transistor, active mode requires $v_E - v_C > 0.2\,\text{V}$. If this condition is violated, the transistor is saturated, and the analysis cannot continue using these simple rules. In *design* problems, change parameters (e.g., resistors, supply rails, etc.) to prevent saturation.
- Sometimes it's easier to find v_E first and use it to calculate v_B.
- How "small" i_B must be to neglect its effect depends on the circuit. In particular, $i_B \times R_B$ must be very small, where R_B is the the Thévenin equivalent resistance looking *out* of the transistor base.

The Base-Emitter Diode: Always keep Figure 1.1 in mind. The Ebers-Moll model of a BJT treats the current-voltage relationship of the base–emitter junction just like a Shockley ideal diode whos current is mirrored by the collector with gain β. When v_B and v_E are not obvious, remember the base–emitter *diode*.

Saturation Mode and Compliance

These simple rules are only valid when the transistor is in *active mode*. In this mode, the transistor is able to dynamically adjust its collector–emitter resistance from near $0\,\Omega$ (e.g., a closed switch with some small finite resistance) to ∞ (e.g., an open switch). These simple rules fail when they imply that the collector–emitter resistance is *negative* (i.e., when they imply that the transistor is acting like a battery).

- When the collector–emitter potential in an *npn* transistor drops to below 0.2 V, the transistor *saturates* and leaves active mode.
- When the emitter–collector potential in a *pnp* transistor drops to below 0.2 V, the transistor *saturates* and leaves active mode.

In nearly every design in this laboratory, we will keep the transistors in *active mode*. The only exception to this is when we use the transistor as a *switch*. For the *switch* case, the transistor base current is made sufficiently high to drive it into saturation, which brings its collector–emitter potential as close to 0 as possible.

Compliance: A circuit's *compliance* refers to the output voltages possible that keep its active components (i.e., its transistors) all in their active modes. If the output voltage is outside of the compliance range, the device will not operate correctly (e.g., a current source will stop providing constant current).

3 Common-Emitter Amplifier

An *npn* common-emitter amplifier (with *emitter degeneration*) is shown in Figure 3.1 for some time $t \geq 0$.

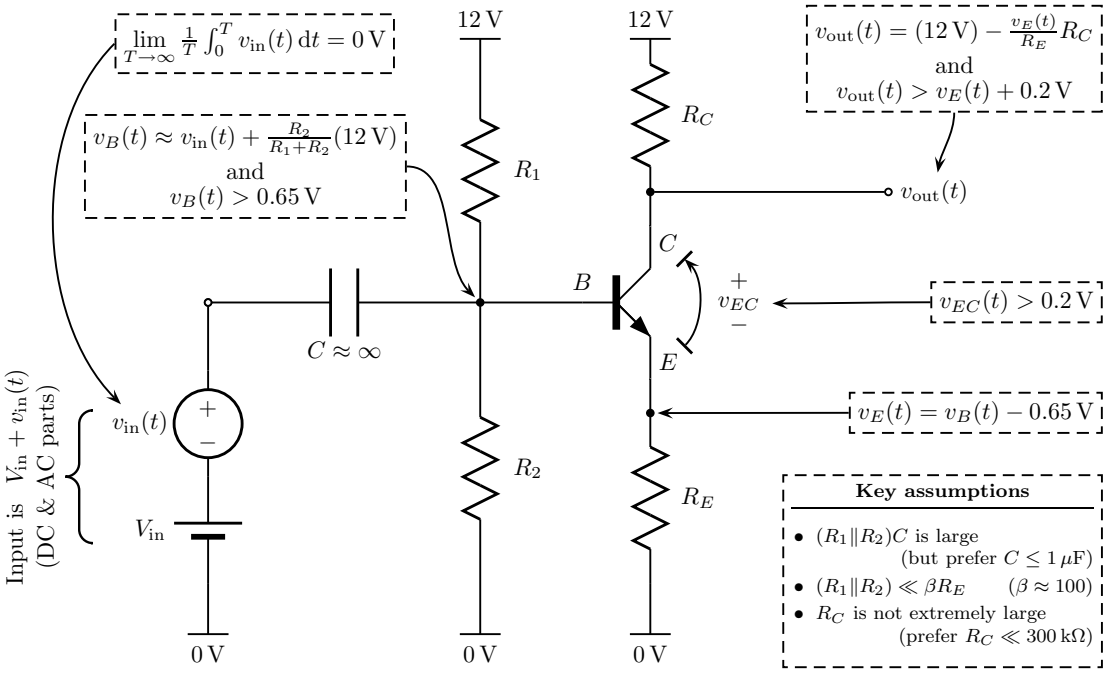


Figure 3.1: Single-ended *npn* common-emitter amplifier (a level-shifter amplifier).

From the boxed expressions in Figure 3.1, it should be clear that at any time t,

$$v_{\text{out}}(t) = \underbrace{(12\,\text{V}) + (0.65\,\text{V})\,\frac{R_C}{R_E} - \frac{R_C}{R_E}\,\frac{R_2}{R_1+R_2}\,(12\,\text{V})}_{\text{DC offset of output signal}} - \underbrace{\frac{R_C}{R_E}}_{|\text{Gain}|}\,v_{\text{in}}(t). \tag{3.1}$$

Pay attention to the inequalities; they keep the transistor out of cutoff and saturation modes. So

$$(0.65\,\text{V}) - \frac{R_2}{R_1+R_2}\,(12\,\text{V}) < \min_t\{v_{\text{in}}\} \quad \text{and} \quad \max_t\{v_{\text{in}}\} < (11.8\,\text{V})\,\frac{R_E}{R_E+R_C} - (12\,\text{V})\,\frac{R_2}{R_1+R_2} + (0.65\,\text{V}). \tag{3.2}$$

For maximal bandwidth and minimal current draw, it's best to make R_1 and R_2 **as large as possible**. However, to keep the circuit as insensitive to variations in β as possible, **make sure that**

$$(R_1 \| R_2) \ll \beta R_E \qquad \text{where} \qquad R_1 \| R_2 \triangleq \frac{R_1 R_2}{R_1+R_2} \qquad \text{and} \qquad \beta \approx 100. \tag{3.3}$$

That is, keep the parallel combination $R_1 \| R_2$ less than the approximate input impedance of the BJT. Otherwise, even a small base current will cause a significant drop across $R_1 \| R_2$.

The capacitor C is needed to AC couple the signal to the base of the transistor. In particular, C, R_1, and R_2 form a high-pass filter with single time constant $(R_1 \| R_2)C$, and so $C \gg 0$. If f is the lowest frequency of interest in the input (i.e., all signals reach the amplifier with at least half power above this frequency), then

$$C \geq \frac{1}{2\pi f\,(R_1 \| R_2)} \qquad \text{where} \qquad R_1 \| R_2 \triangleq \frac{R_1 R_2}{R_1+R_2}. \tag{3.4}$$

At signal frequencies, the circuit's input impedance may be very low. If the circuit *loads* the input source (i.e., input signal amplitude drops when connected to your circuit) and you cannot increase R_1 and R_2 any more, then try the configuration discussed in Appendix A.

The output impedance of this circuit can be prohibitively high. Its use often necessitates also using a buffer like the one in section 6. To prevent attenuation or low-pass effects, keep R_C low (e.g., $R_C \ll 300\,\text{k}\Omega$).

4 *pnp* Current Sources

For Figure 4.1, pick v_B and R_E for current i_{out}. Device in compliance when $\frac{v_{R_E} \times L}{R_E} < v_{\text{comp}} < 11.8\,\text{V}$.

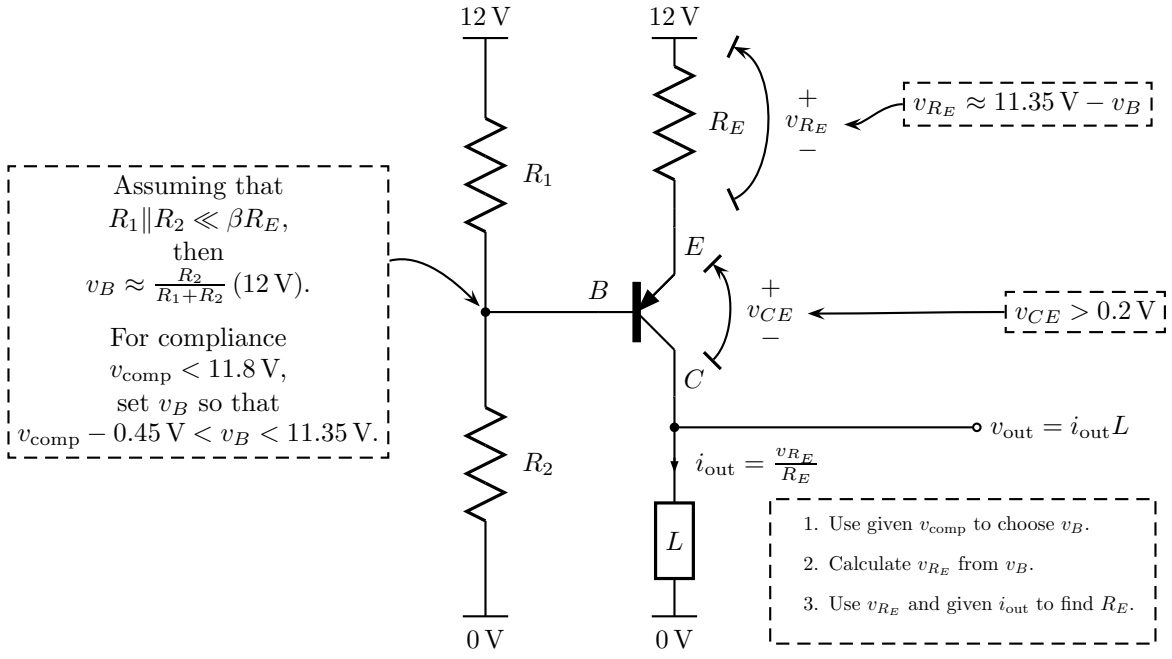


Figure 4.1: Resistor-biased *pnp* current source—output i_{out}, v_{comp} compliance, load L.

For Figure 4.2, pick v_B for current i_{out}. Device is in compliance when $\frac{v_{R_E} \times L}{R_E} < 10.65\,\text{V}$.

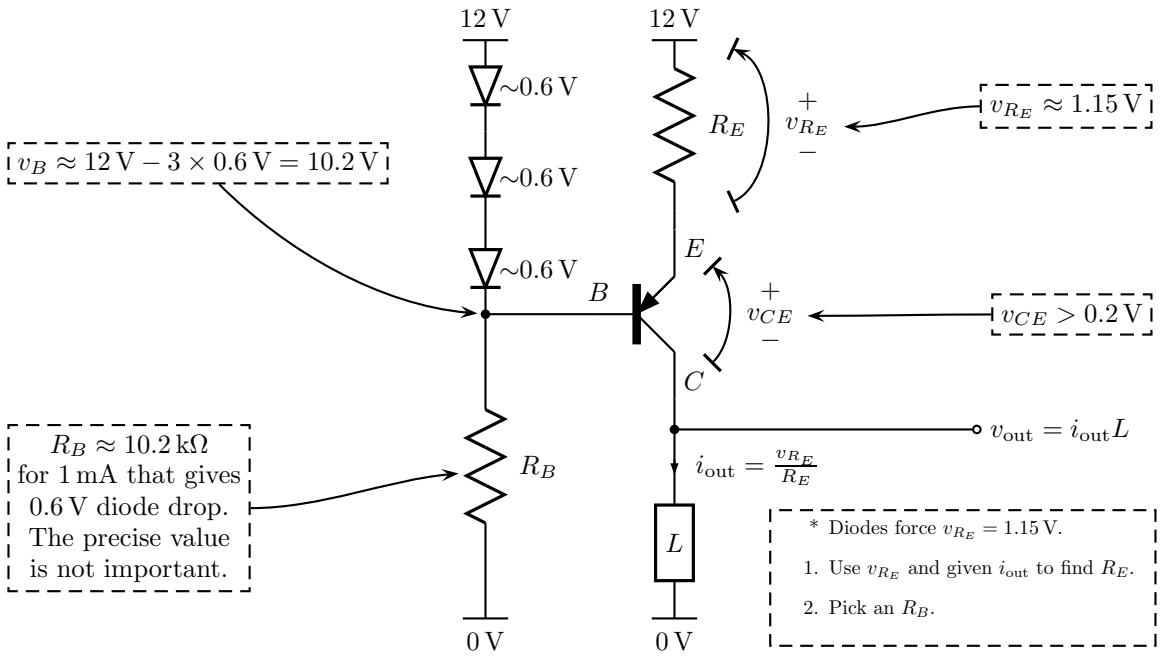


Figure 4.2: Diode-biased *pnp* current source—output i_{out}, 10.65 V compliance, load L.

Notice the simplicity of diode bias in Figure 4.2. For greater compliance, use two diodes instead of three.

5 *pnp* Current Mirror

Instead of using one of the *pnp* current sources in section 4, you may want to try a *pnp* current mirror. *See below for caveats.* As shown in Figure 5.1, a current mirror uses two transistors and a resistor. The transistor on the left sets up a constant current, and the transistor on the right mirrors that current.

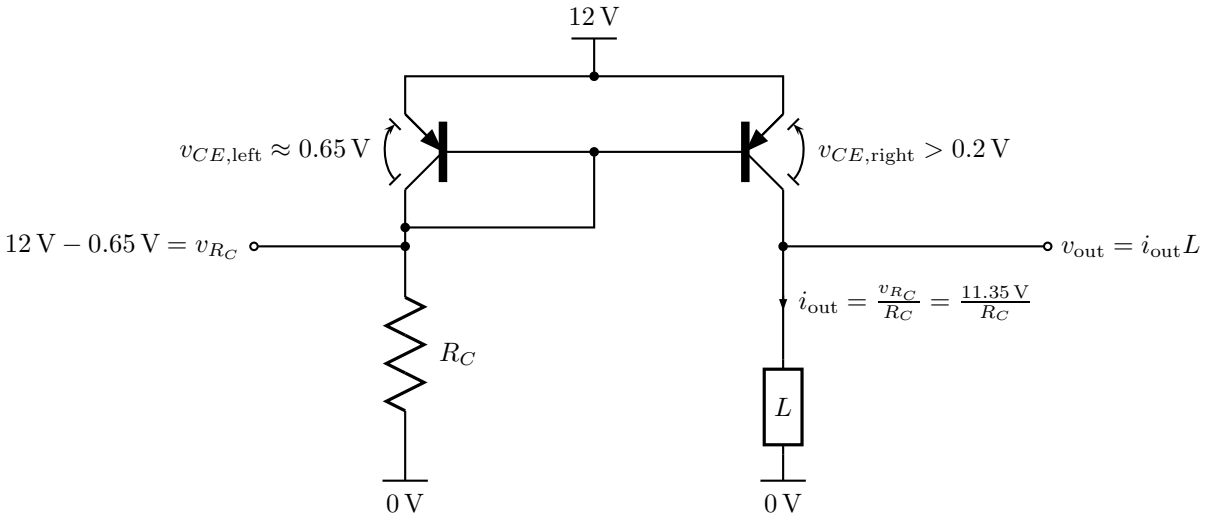


Figure 5.1: *pnp* current mirror—output i_{out}, 11.8 V compliance, load L.

The *pnp* current mirror is relatively simpler than other current sources and has high compliance.

1. The left *pnp* transistor is "*diode-connected*," which means that its collector and base are shorted together. Because the base is 0.65 V below the emitter, the collector is held at 0.65 V below the emitter as well. Therefore, the transistor looks like a diode.
2. Potential across resistor R_C is 11.35 V, and so the collector current on the left transistor is $(11.35\,\text{V})/R_C$.
3. The left and right *pnp* transistors share the same base and emitter nodes. If we assume that they are perfectly matched transistors, they should have the same collector current. Therefore,
$$i_{\text{out}} = \frac{v_{R_C}}{R_C} = \frac{11.35\,\text{V}}{R_C}.$$
4. In order to keep $v_{CE,\text{right}} > 0.2\,\text{V}$, $v_{\text{out}} < 11.8\,\text{V}$. Therefore, the device has a compliance of 11.8 V.

Additional transistors can be added in the same way as the right transistor in order to have multiple outputs all with the same current.

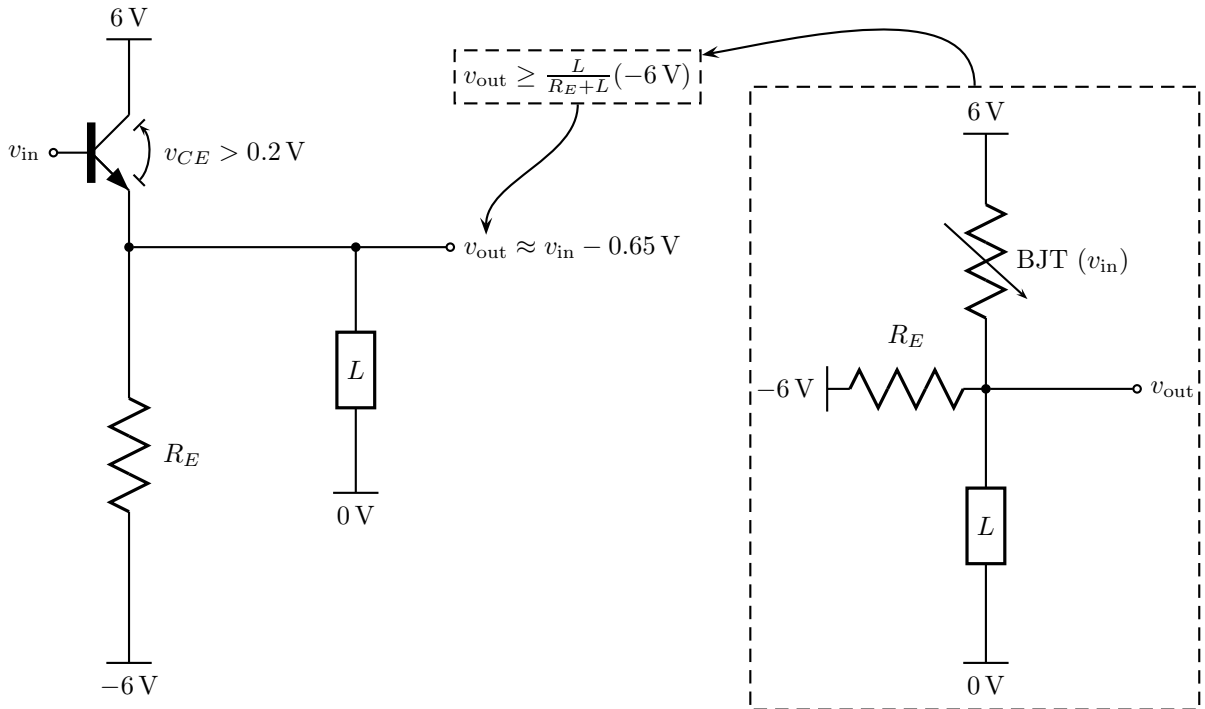
Caveats: For several reasons, this simple circuit may perform poorly in the laboratory.

- Because we are not using transistor pairs from a transistor array, the transistors we use are not matched. Hence, two transistors with the same base–emitter potential may not have the same collector current.
- The *Early effect* in BJTs causes collector current to depend on collector–emitter drop. As the two transistors will have different collector–emitter drops, the currents will not be matched.
- When the load resistor is small, the power dissipation in the right transistor will be high. As the BJT temperature rises, less base–emitter excitation is needed, and so the collector current will increase, causing further power dissipation. This *thermal runaway* can destroy the transistor (and quickly saturate the output). *Warning:* Transistors get hot to the touch during thermal runaway. To mitigate this problem, a source resistor at the 12 V supply can be added (to limit power at the cost of compliance).

Other current mirrors are (more) immune to these problems, but they usually require more complex configurations and/or tightly matched BJTs. **IF YOU USE THIS CURRENT MIRROR IN THE LAB, BE SURE TO MINIMIZE THE TIME THE CIRCUIT IS POWERED ON.**

6 *npn* Emitter Follower

The emitter follower in Figure 6.1(a) produces an output that follows the input, less one diode drop. However, it can *source* current to the output without loading the input. That is, it provides current/power gain.



(a) Implementation.

(b) Simplified using a "dynamic resistor."

Figure 6.1: *npn* emitter follower with load L and $-\frac{L}{R_E+L}(6\,\text{V}) + 0.65\,\text{V} < v_{\text{in}} < 6\,\text{V}$.

Emitter resistance: When output v_{out} is not clipped, the input current required is

$$\frac{1}{\beta}\left(\frac{v_{\text{in}} - 0.65\,\text{V} - 6\,\text{V}}{R_E} + \frac{v_{\text{in}} - 0.65\,\text{V}}{L}\right) = \frac{1}{\beta}\left(\frac{v_{\text{in}} - 0.65\,\text{V}}{R_E \| L} - \frac{6\,\text{V}}{R_E}\right)$$

where $\beta \geq 100$ for small-signal BJTs. That is, small-signal input impedance is approximately $\beta(R_E \| L)$. A low R_E increases transistor power dissipation and input current required, but it is required for clipping immunity, as shown below. High-β transistors reduce input current demand even for low R_E.

Clipping on negative swing: Consider the simplified circuit in Figure 6.1(b), which replaces the transistor with a "dynamic variable resistor" that "pulls up" the output toward 6 V. Increasing the size of this resistor moves the output lower. The lowest output occurs at transistor cutoff, as the pull-up resistor is *infinite* (i.e., an open circuit). Hence, the lowest output is $(-6\,\text{V}) \times L/(R_E + L)$. A small R_E reduces this distortion but also increases power dissipation. Not surprisingly, there is a power–distortion tradeoff.

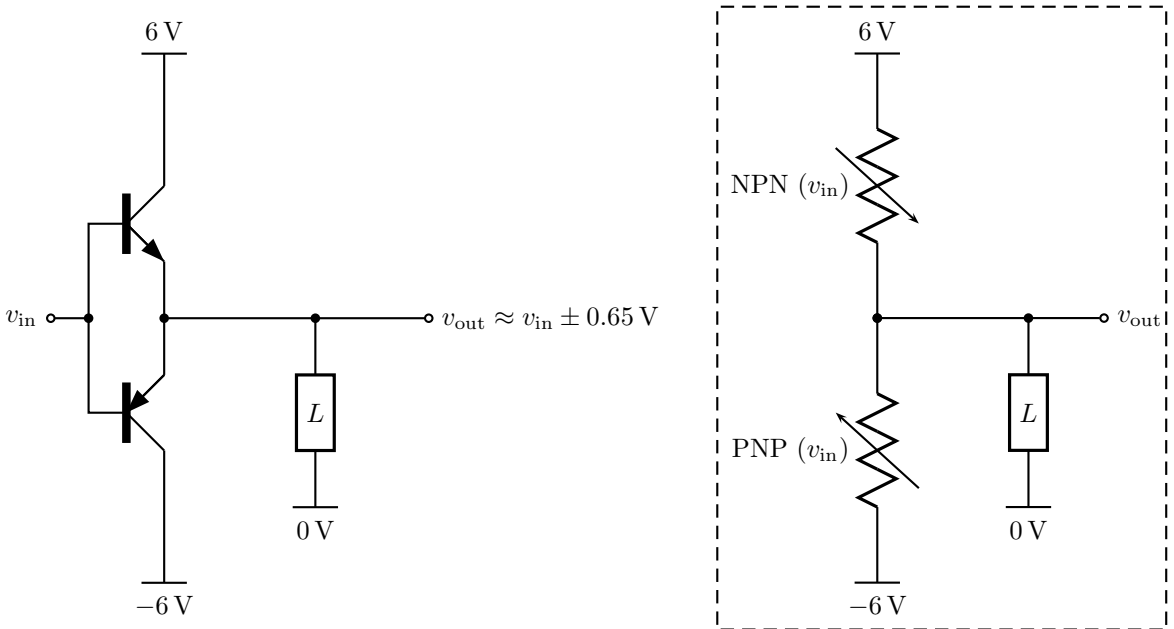
Unable to "sink" current: The output cannot swing lower than $(-6\,\text{V}) \times L/(R_E + L)$ because ground current from the load can only travel through the R_E resistor. The *npn* emitter diode cannot *sink* current, so it all must go through R_E. If R_E were replaced with another "dynamic resistor," more negative output swings could be produced. In other words, R_E should be replaced with another transistor, as in section 7.

Negative swing caveat: Pulling the input signal down very far will cause the transistor's base–emitter diode to breakdown. Care should be taken to prevent this from happening. Prevent v_{in} from swinging more than 6 V lower than the emitter. For protection, add a reverse-biased diode from the base to the emitter.

7 Simple Class-B Push–Pull Amplifier

In section 6, the *npn* emitter follower cannot track some negative signals because it lacks the ability to *sink* current away from its emitter resistor. If the emitter resistor is replaced with a "dynamic resistor" able to alter the current through it, then the output will be able to swing over a wider range of negative values.

The circuit in Figure 7.1(a) is a "*class-B push–pull amplifier*." It is a combination of an *npn* emitter follower and a *pnp* emitter follower. The *pnp* transistor is the "dynamic resistor" we needed in section 6. Thus, the push–pull amplifier is a kind of "two-tailed follower." Compare Figure 7.1(b) to Figure 6.1(b).



(a) Implementation.

(b) Simplified using "dynamic resistors."

Figure 7.1: Class-B Push–Pull Amplifier with load L and $-6\,\text{V} < v_{\text{in}} < 6\,\text{V}$.

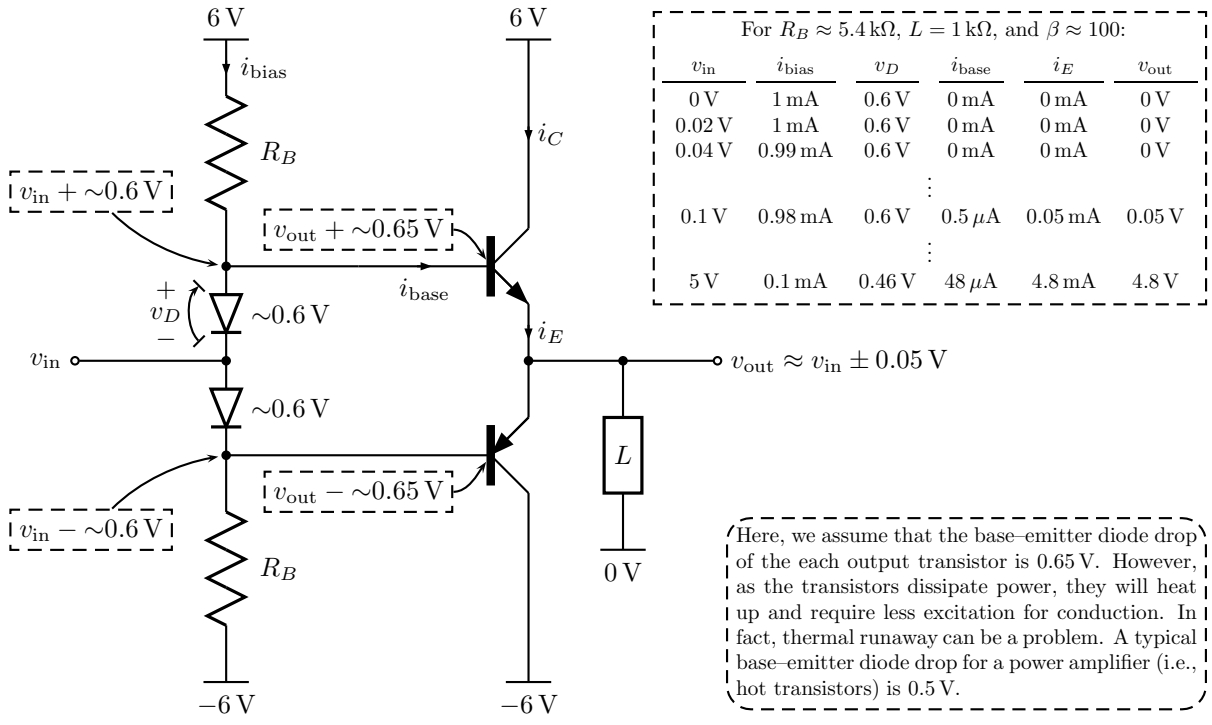
Pushing, Pulling, Sourcing, and Sinking: It is a "*push–pull*" amplifier because the upper *npn* transistor "pushes" (i.e., sources) current for positive input and the lower *pnp* transistor "pulls" (i.e., sinks) current for negative input. It is a "*class-B*" amplifier because only one transistor is active at a time. That is, for sufficiently positive signals, the upper transistor will be active and the lower transistor will be in cutoff mode (and vice versa). In Figure 7.1(b), this case corresponds to the PNP resistor being ∞ and the NPN resistor being some finite resistance that puts the desired v_{out} between 0 V (from ground) and 6 V. For the class-B amplifier, only one "dynamic resistor" has a finite resistance at a time.

Distortion and Power: The *class-B* amplifier does not "waste" power because transistors only dissipate power (i.e., activate) when absolutely necessary to produce output. This property is attractive in low-power applications, but it comes with a power–distortion tradeoff. For signals between −0.65 V and 0.65 V, both "dynamic resistors" are infinite (i.e., "off"), and so the output is 0 V. Larger signals that swing through this "dead zone" suffer significant distortion. We show in section 8 that some of this *crossover distortion* can be removed at the cost of additional power dissipation; we use extra biasing power so that the "dynamic resistors" turn "on" more easily.

Practical Applications: Circuits like this one (and the one described in section 8) are often used as output stages of buffering devices, like operational amplifiers. As the combination of two emitter followers, it provides very high power gain and is still be able to handle positive and negative signals (i.e., both sourcing and sinking load current). Because of the distortion it adds, this simple push–pull stage will only be used in low-power applications that sacrifice distortion for reduced power consumption.

8 Biased Class-B Push–Pull Amplifier

The circuit in Figure 8.1 is called a *class-B biased push–pull amplifier*. It is nearly identical to Figure 7.1, but a new biasing circuit keeps the transistors on for "longer" to reduce *crossover distortion*.



For $R_B \approx 5.4\,\text{k}\Omega$, $L = 1\,\text{k}\Omega$, and $\beta \approx 100$:

v_{in}	i_{bias}	v_D	i_{base}	i_E	v_{out}
0 V	1 mA	0.6 V	0 mA	0 mA	0 V
0.02 V	1 mA	0.6 V	0 mA	0 mA	0 V
0.04 V	0.99 mA	0.6 V	0 mA	0 mA	0 V
⋮					
0.1 V	0.98 mA	0.6 V	0.5 μA	0.05 mA	0.05 V
⋮					
5 V	0.1 mA	0.46 V	48 μA	4.8 mA	4.8 V

Here, we assume that the base–emitter diode drop of the each output transistor is 0.65 V. However, as the transistors dissipate power, they will heat up and require less excitation for conduction. In fact, thermal runaway can be a problem. A typical base–emitter diode drop for a power amplifier (i.e., hot transistors) is 0.5 V.

Figure 8.1: Class-B Biased Push–Pull Amplifier with load L and $|v_{\text{in}}| < \sim 5.4\,\text{V}$.

Biasing diodes: The diodes move some of the "dead zone" from the middle of the input range to the ends, which reduces distortion while decreasing maximum input swing. Ideally, the diodes would have the same potential drop as the base–emitter diode in each of the transistors, which is usually true for a prefabricated transistor/diode array (i.e., inside an *integrated* circuit). In our case, our *discrete* components are not matched, so the output still has crossover distortion (e.g., consider signals between $-0.05\,\text{V}$ and $0.05\,\text{V}$). Additionally, as v_{in} swings and i_{base} varies, v_D changes slightly, so there is a *new* source of distortion.

Class-B push–pull amplifier: The amplifier is still a *class-B* amplifier because only one transistor is active at a time. However, the two diodes make the transistors more sensitive to small signals, and so they activate more readily than in the case in section 7. The "dead zone" that causes output distortion is smaller.

Class-AB push–pull amplifier: The amplifier bridges on being a "*class-AB* amplifier." If additional biasing diodes and emitter resistance are added, *both* transistors will be active simultaneously for all small signals, which is like a *class-A* amplifier (i.e., both transistors are active even though only one is "needed"). However, if these signals swing out of a small region (determined by load and how many diodes are added), the amplifier behaves like a class-B amplifier (i.e., only one transistor is active). The class-A operation "wastes" power for less distortion, and the class-B operation introduces distortion to save power.

Power and distortion: The circuit "wastes" power in its biasing diodes in order to reduce output distortion. Increasing the biasing current will further reduce output distortion as the v_D drop will depend less on the load. There is always a tradeoff between distortion and power. Clean amplifiers use extra power for *low distortion* of *small* signals; the extra power is not simply for production of large (e.g., high volume) signals.

A Bootstrapping for Higher Input Impedance

Depending on your component choices and signal source, the circuit in Figure 3.1 may *load* the source so that the input signal is noticeably attenuated when connected to the circuit. That is, at signal frequencies, the input impedance of the circuit may be low compared to the output impedance of the signal source, and so dissipation in the signal source causes attenuation of the signal entering the circuit.

To ensure the current into the base of the transistor is negligible, the biasing network must have a relatively low equivalent resistance at DC when looking out of the base. However, there is a clever method we can use to raise the impedance of the network *at signal frequencies* when looking out of the capacitor. By *bootstrapping* some of the transistor's output signal back into the input, we can make the input impedance (at signal frequencies) very large (i.e., approximately βR_E, the input impedance of the transistor).

Consider the modified circuit in Figure A.1.

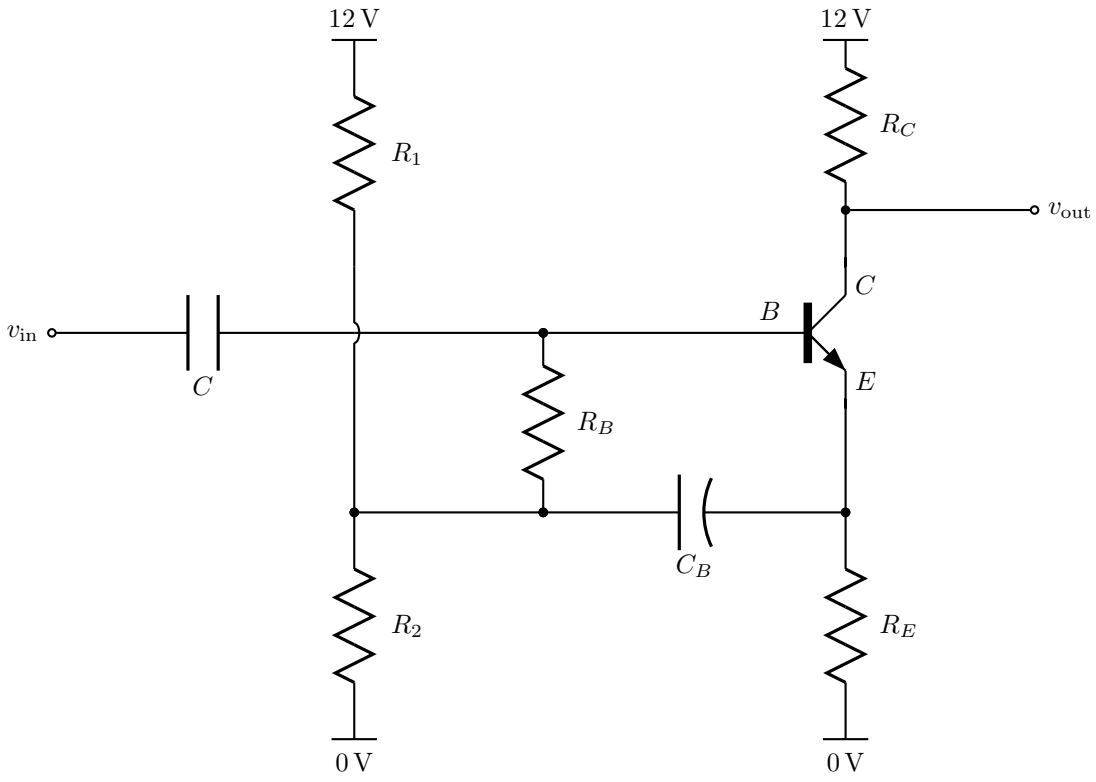


Figure A.1: *Bootstrapped* single-ended *npn* common-emitter amplifier.

Instead of Equation (3.3) and Equation (3.4), assume that

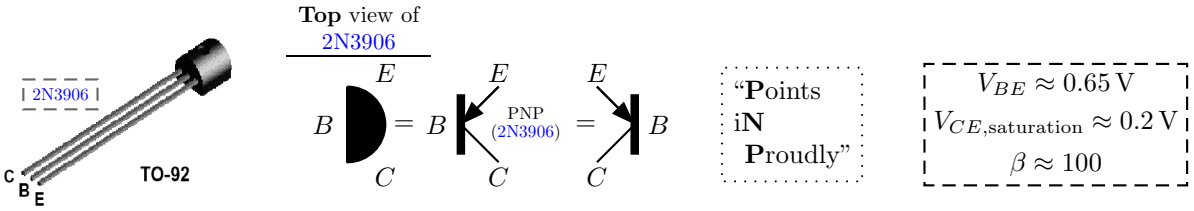
$$(R_1 \| R_2) + R_B \ll \beta R_E \qquad \text{and} \qquad \beta \approx 100 \qquad \text{and} \qquad C \geq \frac{1}{2\pi f\,(R_1 \| R_2 + R_B)}. \tag{A.1}$$

Otherwise, components can be chosen exactly as before. The *bootstrapping capacitor* C_B must be very large so that it looks like a short circuit to signal frequencies[2]. Theoretically, the resistor R_B can be chosen arbitrarily. As long as Equation (A.1) can be met, a high choice of R_B (e.g., $R_B > 1\,\text{k}\Omega$) is a good idea.

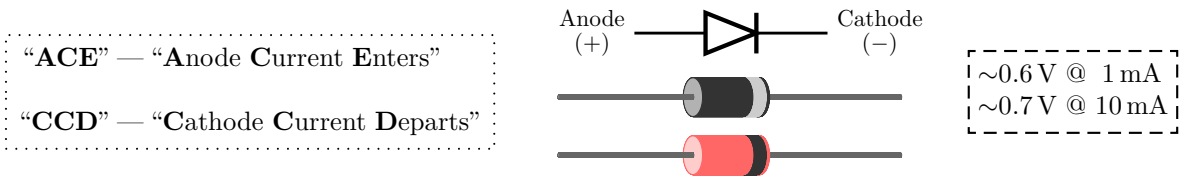
The signal at the transistor's emitter follows the signal at its base. At signal frequencies, C_B acts like a short circuit, and so both ends of R_B see the same potential. Hence, R_B *carries no current* at signal frequencies. Thus, the R_1–R_2 divider cannot load the input source because no current from the source makes its way across R_B (i.e., $R_B \approx \infty$ at signal frequencies). The current through R_1–R_2 that would normally come from the source comes from the *output* instead. This method is called *bootstrapping* because we use the circuit's *own output* to reduce current required from the input.

[2] A half-power expression for C_B is complicated; in the laboratory, it will usually be chosen heuristically.

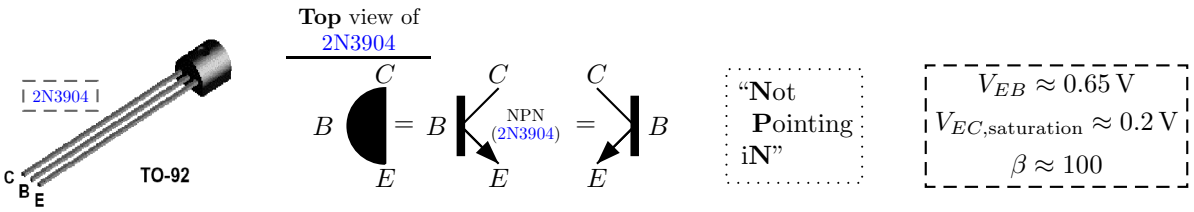
B Parts



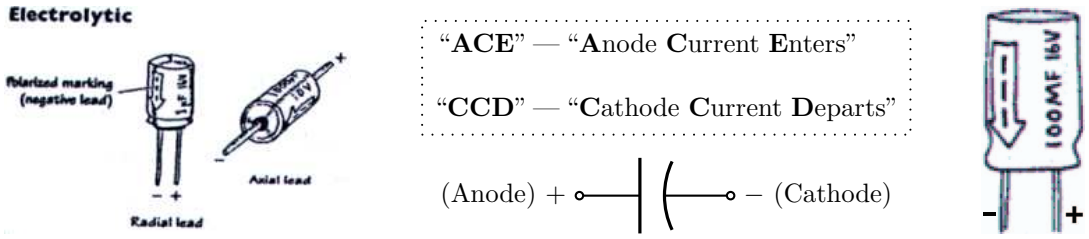
(a) 2N3906 PNP BJT transistor



(b) 1N914 silicon diode (generic small-signal diode)



(c) 2N3904 NPN BJT transistor



(d) Electrolytic capacitor

Figure B.1: Part pin-outs.