Physics 364, Fall 2014, reading due 2014-10-12.

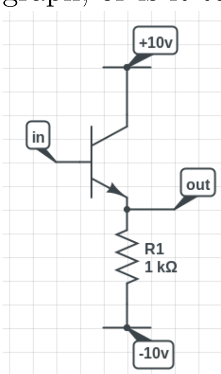
Email your answers to `ashmansk@hep.upenn.edu` by 11pm on Sunday

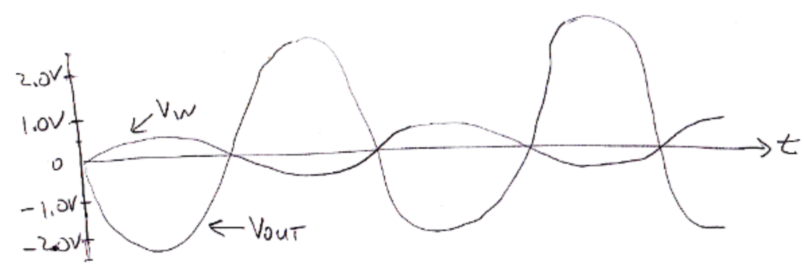
Course materials and schedule are at `http://positron.hep.upenn.edu/p364`

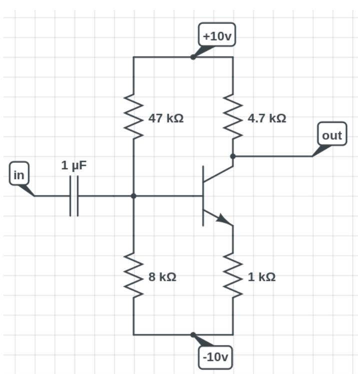
Assignment: (a) First read through my notes (starting on next page), which directly relate to what we will do in the labs. (b) Then read Eggleston's chapter 4 (Bipolar Junction Transistors, pp 104–130). You can spread chapter 4 over two weeks if you wish, as next week's reading will be from my notes only, no textbook. (c) Then email me your answers to the questions below.

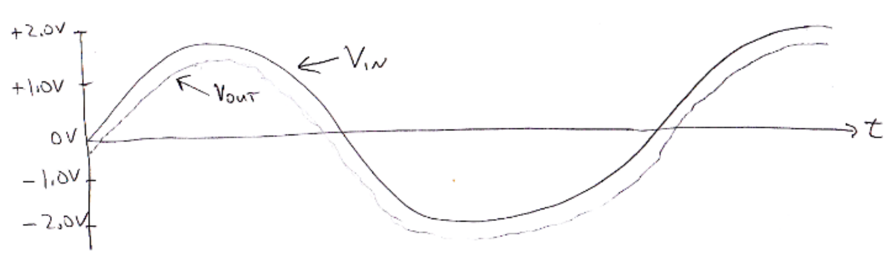
1. (a) In your own words, define *output impedance.* (b) In your own words, define *input impedance.* (c) When sending voltage signals around, which of these do you want to keep as large as possible, and which do you want to keep as small as possible?

2. The figure below shows two circuits (top circuit and bottom circuit) and two graphs (top and bottom) of $V_{\text{in}}(t)$ and $V_{\text{out}}(t)$. (a) What is the name of the top circuit? (b) What is the name of the bottom circuit? (c) Which $V(t)$ graph goes with which circuit? (Is it top circuit with top graph and bottom circuit with bottom graph, or is it top circuit with bottom graph and bottom circuit with top graph?)







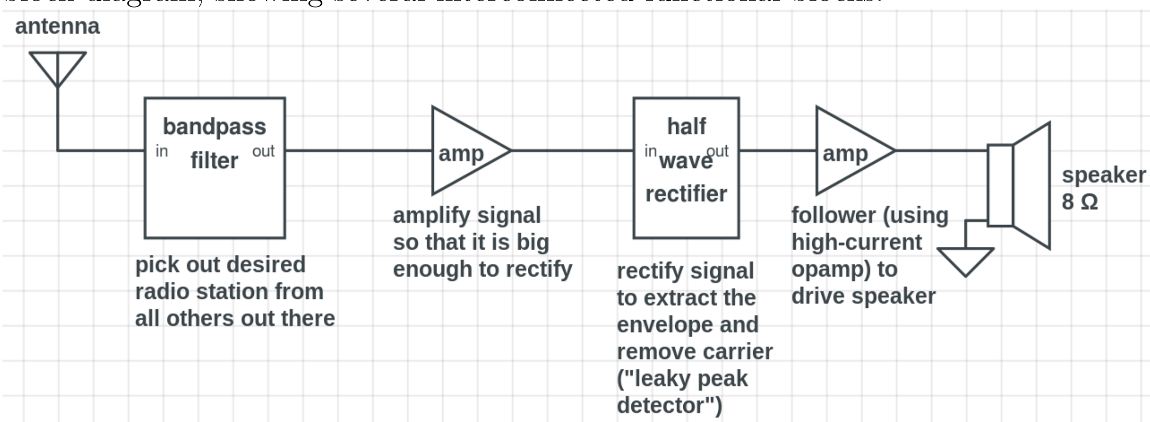


3. Is there anything from this reading that you would like me to try to clarify? If you didn't find anything confusing, what topic did you find most interesting?

4. How much time did it take you to complete this assignment?

We're going to spend the next couple of weeks on *transistor* circuits. You will probably find transistors to be the most conceptually difficult part of the course: once we get through transistors, you should find the rest of the course relatively easy to digest. And what we cover toward the end of the term should leave you equipped with a handy toolkit for taking on your own electronics projects beyond Physics 364.

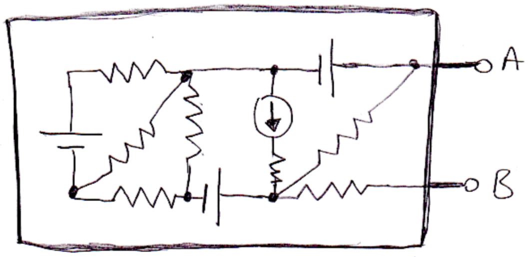
Before we introduce transistors, let's step back to review (or restate, from a new point of view) the most important concepts that the course has covered so far. The first key idea is that circuits to a large degree can be decomposed into functional blocks (voltage dividers, frequency-selective filters, rectifiers, amplifiers, etc., etc.), and that each functional block typically has some *input* connections and some *output* connections. For example, your Lab 10 AM radio receiver was decomposed into this block diagram, showing several interconnected functional blocks:

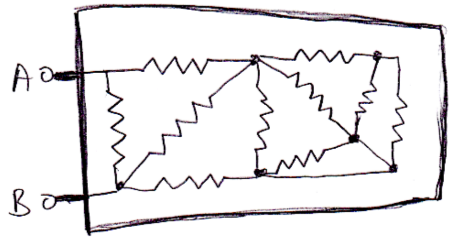


The radio signal was processed by several successive circuit fragments, stage-by-stage. The boundary between adjacent stages consists of an upstream circuit fragment's output, sending a voltage[1] signal to a downstream circuit fragment's input: at each boundary, an upstream *voltage source* is driving a downstream *load*.

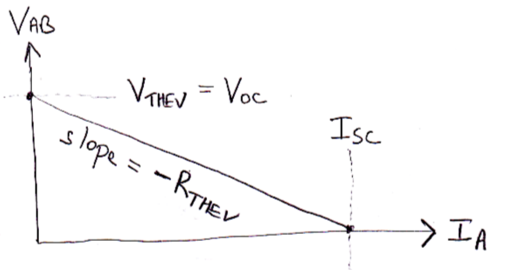
Not every voltage source is effective at driving every load. Some voltage sources are more capable than others, and some loads are more challenging to drive than others. Also, the voltage source might internally be very complicated, like the below (left) combination of (ideal) batteries, resistors, and (ideal) current sources. The load might internally be very complicated, like the below (right) rat's nest of resistors. Is there a simple way to represent how capable a given source is of driving a given load?

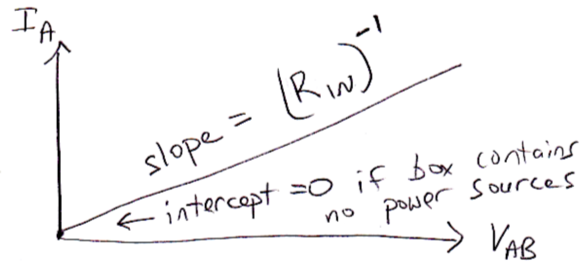
[1]Sometimes it is convenient to represent a signal as a current rather than as a voltage (e.g. antenna, photodiode) but most signals are conveniently modeled as voltage sources.





When we connect the source to the load (terminal A to terminal A, and B to B), the voltage V_{AB} between the two terminals of the source will cause a current I to flow from the source's terminal A to the load's terminal A. (Most of the time, we replace both B terminals with a *ground* connection, and just talk about V_A instead of V_{AB}. But we should keep in mind when we do this that we are measuring V_A with respect to ground.) For an ideal source, V_{AB} will be independent of I; but no source is truly ideal. So the most convenient way to characterize the source is to draw its V_{AB}-vs.-I_A curve, and the most convenient way to characterize the load is to draw its I_A-vs.-V_{AB} curve. If the source's V_{AB}-vs.-I_A curve is linear, then it will look like the graph below (left); and if the load's I_A-vs.-V_{AB} curve is linear (and in addition if the load contains no power sources), then it will look like the graph below (right).





The left curve has two key parameters: an intercept, called V_{thevenin}, and a slope, called $(-R_{\text{thevenin}})$. The right curve has one key parameter: a slope, called $(1/R_{\text{in}})$. It can be proven[2] that for a black box having two terminals (A and B) and containing only (ideal and independent) voltage sources, (ideal and independent) current sources, and resistors, the I-V curves truly are linear. When the load is connected to the source, V_{AB} is reduced from its *open-circuit* value (known as V_{thevenin}) to the value[3]

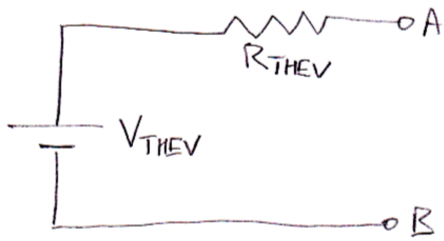
$$V_{AB} = V_{\text{thev}} \frac{R_{\text{in}}}{R_{\text{in}} + R_{\text{thev}}}.$$

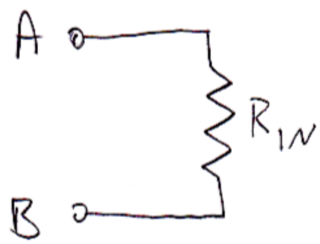
If $R_{\text{thev}} \ll R_{\text{in}}$, the source barely notices the presence of the load; but if R_{thev} is comparable to or larger than R_{in}, then V_{AB} is reduced non-negligibly from its open-circuit value.

[2] This called *Thévenin's theorem.*

[3] Not only is this "the voltage divider equation;" it is also the solution to the simultaneous equations (represented by the two graphs) $V_{AB} = V_{\text{thev}} - R_{\text{thev}} I_A$ and $I_A = V_{AB}/R_{\text{in}}$.

So when talking about how the upstream and downstream circuit fragments interact with one another, we simplify the problem by representing the upstream fragment by the below (left) circuit and the downstream fragment by the below (right) circuit, which have the same I-V curves as the original circuit fragments but are far easier to analyze.





It turns out that even in cases in which the I-V curves are not perfectly linear, the Thévenin resistance concept is still useful. Then instead of talking about R_{thev} (which strictly applies only in the linear case), we can define the **output resistance** (of the black-box upstream voltage source) to be

$$R_{\text{out}} = -\frac{\mathrm{d}V_{AB}}{\mathrm{d}I_A}$$

and we can generalize the **input resistance** (of the black-box downstream load) by writing

$$\frac{1}{R_{\text{in}}} = \frac{\mathrm{d}I_A}{\mathrm{d}V_{AB}}.$$

These last two definitions turn out to be especially helpful when the voltage sent from upstream to downstream is the superposition of a small time-varying signal with a large DC offset: then the DC solution tells you which region of a non-linear I-V curve to look at, and the slopes of the curves (R_{out} and R_{in} as defined above) give you the factor $(R_{\text{in}})/(R_{\text{in}} + R_{\text{out}})$ by which the small time-varying signal is reduced from its ideal ("open circuit") value.

We can also generalize the input and output resistance concepts to include time-varying voltage and current sources, capacitors, inductors, etc., by including phase shifts (via the complex-number formalism described in reading #2) and defining **output impedance**, $\mathbf{Z}_{\text{out}}$, (for the upstream circuit fragment) as

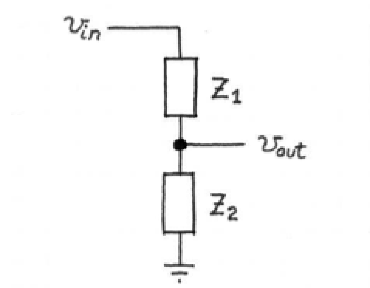
$$\mathbf{Z}_{\text{out}} = -\frac{\mathrm{d}\mathbf{V}_{AB}}{\mathrm{d}\mathbf{I}_A}$$

and **input impedance**, $\mathbf{Z}_{\text{in}}$, (for the downstream circuit fragment) as

$$\frac{1}{\mathbf{Z}_{\text{in}}} = \frac{\mathrm{d}\mathbf{I}_A}{\mathrm{d}\mathbf{V}_{AB}}\ .$$

Remember that an ideal voltage source has $\mathbf{Z}_{\text{out}} = 0$ and that the ideal load for a voltage source has $\mathbf{Z}_{\text{in}} = \infty$. Also keep in mind that people use the phrases "output impedance," "source impedance," "output resistance," "source resistance," and "Thévenin resistance" more or less interchangeably; and that people use the phrases "input impedance," "input resistance," and "load resistance" more or less interchangeably. Also, people sometimes are sloppy and use the phrase "internal resistance" to mean either Z_{in} or Z_{out}, depending on the context.[4] So we have just recapped (and generalized somewhat) the **first very important concept** from Physics 364: that of **input and output resistances**.

The **second very important concept** from Physics 364 is **impedance** (symbol $\mathbf{Z}$, which is a complex quantity), which mathematically unifies resistors, capacitors, and inductors. Using $\mathbf{Z}_R = R$, $\mathbf{Z}_L = j\omega L$, and $\mathbf{Z}_C = 1/(j\omega C)$, where $j = \sqrt{-1}$, we generalized Ohm's law to $\mathbf{V} = \mathbf{IZ}$. From there, we generalized our humble voltage divider into several different styles of frequency-selective **filters**: low-pass, high-pass, and band-pass, using the generalized voltage-divider equation: $\mathbf{V}_{\text{out}} = \frac{\mathbf{Z}_2}{\mathbf{Z}_1+\mathbf{Z}_2}\mathbf{V}_{\text{in}}$.



generalized voltage divider

The **third very important concept** from Physics 364 up to this point is the **opamp golden rules**. As long as negative feedback is present from V_{out} back to V_-, and V_{out} is not saturated, then for an ideal opamp (the $A \to \infty$, $R_{\text{in}} \to \infty$ limit), **(Rule 1)** V_{out} will adjust itself as needed such that $V_+ = V_-$, and **(Rule 2)** the opamp's inputs draw negligible current. Using these simple rules, we designed circuits that carried out all sorts of useful operations: follower, inverting and non-inverting amplifier, weighted sum, integrator, logarithm, etc. The non-ideal features of opamps are important to be aware of, so that you know how to read the manufacturers' data sheets when selecting an opamp, but the golden rules alone will get you quite far.

I know that we cover a lot of material in this course. The above recap was to help you to distinguish between the small number of ideas that I think are so important that I hope everyone remembers them well after the course is over, vs. the larger number of ideas that I want you to see once or twice, so that someday you can remember to look them up if or when you need them.

OK, now on to **transistors.** The transistor is the basis of nearly all modern electronics. A transistor enables a low-power signal to control a much higher-power signal (higher current, higher voltage, or both). Thus, transistors can form switches, amplifiers, digital logic gates (we'll study these later), and much more. Transistors are the building blocks from which opamps like the '741 are made.[5] My aim is for you to learn

[4] I sometimes deliberately switch from one synonymous phrase to another, to expose you to as much as possible of the commonly used language, but I try to avoid using ambiguous language.

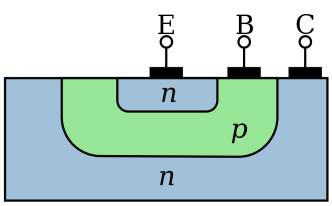
[5] You can see a schematic diagram of the '741 opamp's internal circuitry at `en.wikipedia.org/wiki/Operational_amplifier#Internal_circuitry_of_741_type_op-amp` .

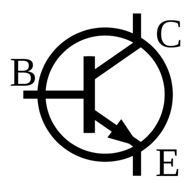
just enough about transistors that transistor-based devices (such as opamps) are no longer opaque to you — so that you see that their operation does not require any magic. In real life, you will probably never design a circuit transistor-by-transistor, though you may occasionally use a transistor as a quick fix to interface two circuits to one another or to switch on and off the flow of a large current.

There are two widely used varieties of transistor — the Bipolar Junction Transistor (BJT) and the Field Effect Transistor (FET). We will study BJTs for a week or two, and then look at FETs for about a week after that.

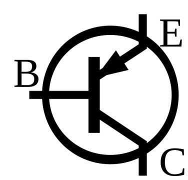
BJTs come in two flavors: NPN and PNP. An NPN transistor is a 3-terminal device consisting of two layers of n-type material separated by a layer of p-type material. By contrast, a PNP transistor consists of two layers of p-type material separated by a layer of n-type material. The physics of semiconductors, doping, n-type and p-type material, etc. is covered in section 3.1 of Eggleston's textbook. You can understand how to use transistors without knowing anything about semiconductor physics, but if you are a physics or EE major, then you are probably curious about the underlying physics. (I hope eventually to work some of the physics into these notes.)

The figure below shows (left) the typical geometry of an NPN transistor and (center) its schematic symbol. The PNP schematic symbol is shown on the right, for comparison, but for this first week, we will only discuss NPN transistors.[6]





(NPN)



(PNP)

The three terminals of a BJT are called **E**mitter, **B**ase, and **C**ollector. A transistor is roughly analogous to a valve on a water pipe, with the water supply above **C**, the knob at **B**, and the downstream pipe below **E**. By manipulating the base (like the knob on a valve), you can control the downward flow of current from collector to emitter. To say it again, the dominant current in an NPN transistor flows **from C**ollector **to E**mitter; this current is primarily controlled by changing V_{BE}, the potential of the base with respect to the emitter.

There are several ways to remember this. First, the little arrow is always at the emitter,[7] and the arrow points in the direction in which conventional (positive) charge flows. Second, if you think of the flow of electrons (opposite the flow of conventional current) inside the transistor, nearly all of the electrons emitted by the emitter get

[6] When using a PNP transistor, it is usually easiest first to imagine how the NPN case would work and then to reverse all of the polarities.

[7] I find it a useful mnemonic to imagine arrows emitting things.

collected by the collector. Third, you can just somehow memorize the NPN schematic symbol and then think of the current flowing downward.

In describing NPN transistor operation, we will refer often to currents I_C (current flowing *into* the transistor at the collector), I_B (current flowing *into* the transistor at the base), and I_E (current flowing *out of* the transistor at the emitter). K.C.L. requires $I_E = I_C + I_B$. We will also refer to potential differences V_{BE} (base voltage minus emitter voltage) and V_{CE} (collector voltage minus emitter voltage).

A BJT has **three modes of operation**. In **cutoff** mode, $I_C \approx 0$ (the "valve" is turned off). In **active** mode (sometimes called "linear active" mode), the collector current I_C is proportional to the much smaller base current I_B (the "valve" is partially open, so that wiggling the knob changes the flow); the constant of proportionality is called β (typically $50 \lesssim \beta \lesssim 300$), so that $I_C = \beta I_B$. In **saturation** mode, I_C is maximal (the "valve" is fully on — so far open that wiggling the knob doesn't affect the flow).

To keep things simple this first week, we will focus our attention mainly on (linear) active mode, which is useful for making transistor-based *amplfiers*. (Cutoff and saturation modes are useful for making transitor-based *switches*.)

The simplest way to understand the operation of an NPN transistor is the following.

(1) The base→emitter junction behaves as a diode. So I_B increases exponentially with V_{BE}:

$$I_B = I_0 \cdot \left(\exp(\frac{V_{BE}}{V_T}) - 1\right)$$

where $V_T = kT/e \approx 25$ mV at room temperature, and the (usually neglected) -1 term after the exponential is just there so that $I_B = 0$ when $V_{BE} = 0$. The factor I_0 is a miniscule current, typically in the femtoampere range.[8] The diode nature of the B→E junction also implies that when the transistor is active, $V_{BE} \approx 0.6$ V: the base voltage must be roughly a *diode drop* above the emitter voltage to allow an appreciable current to flow at the emitter.

(2) The collector current I_C is proportional to the base current I_B by a factor $\beta \sim 100$. In practice, $50 \lesssim \beta \lesssim 300$, and β varies considerably from transistor to transistor.

(3) To keep the transistor in the *active* region, where $I_C = \beta I_B$, the collector must be at least a few tenths of a volt above the emitter: $V_{CE} \gtrsim 1$ V. If V_{CE} becomes too small, the transistor enters *saturation* mode.[9]

[8] Warning: it turns out that I_0 is not really constant: it *increases* with temperature.

[9] The usefulness of saturation mode is that you can switch a large current with relatively small power $I_C V_{CE}$ dissipated in the transistor.

(4) Warning: letting V_{BE} go negative by more than a few volts will cook the transistor.[10] Each model of transistor also quotes a maximum allowable value for V_{CE}, which is typically several tens of volts.

The diode-like behavior of the B→E junction is illustrated in the right-hand graph[11] below: you can see that V_{BE} is the main knob for controlling the current through the transistor. The three modes of transistor operation are illustrated in the left-hand graph below, which shows I_C as a function of V_{CE}, for a variety of different I_B values: you can see that V_{CE} is what separates the active region (where $I_C \propto I_B$) from the saturation region (where I_C can be very large but is no longer proportional to I_B). Again, we will mostly consider *active* mode, where $I_C = \beta I_B$, and $I_E = I_B + I_C = (\beta + 1) I_B$, with $\beta \sim 100$.

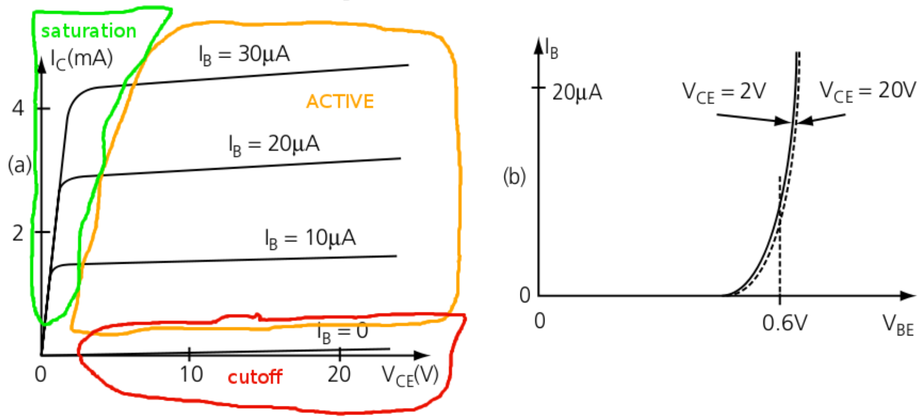
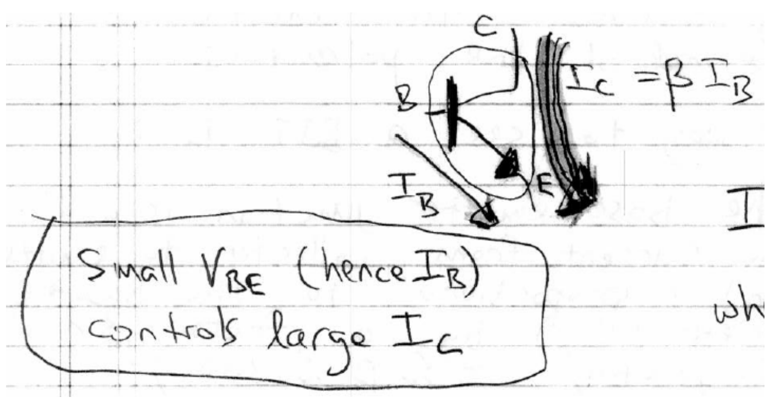


Fig. 9.26. Characteristics of the bipolar transistor.



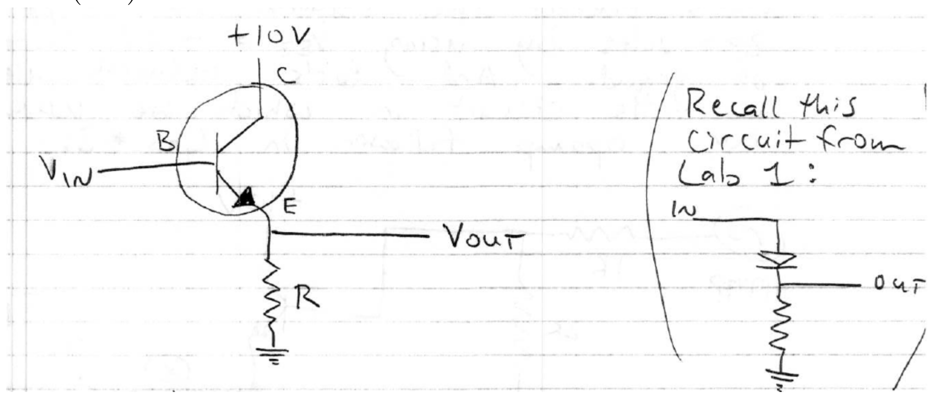
Let's use the above picture of transistor operation to analyze a few of the most common transistor circuits. You will quickly appreciate just how easy opamps are to use, in comparison with discrete transistors!

The first circuit example is called the **emitter follower**. Eggleston's textbook describes the same circuit, but calls it the *common-collector amplifier.* The emitter

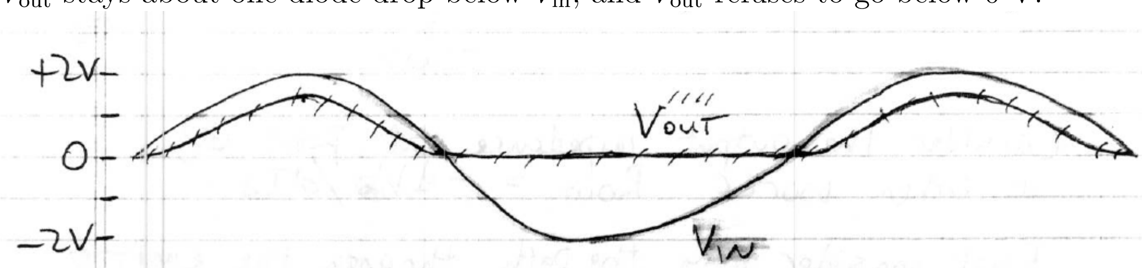
[10] As you have seen in the lab, some of our components appear to be powered by smoke: once the internal smoke escapes from a component, the component no longer functions! (Just kidding.)

[11] Figure 9.26 from D.V. Bugg's electronics textbook.

follower is intended to serve the same purpose as an opamp follower (a.k.a. buffer amplifier): to have V_{out} reproduce ("follow") V_{in}, but with the capability of supplying a larger current to the downstream circuit. So a follower (whether opamp-based or transistor-based) should have a voltage gain of 1, a large input impedance, and a small output impedance.[12] In its simplest form, an emitter follower looks like the figure below (left).



Treating the base→emitter junction as a diode, we have $V_{\text{out}} \approx V_{\text{in}} - 0.7$ V, as long as $V_{\text{in}} \gtrsim 0.7$ V. If V_{in} is too small (or is negative), the diode-like B→E junction turns off, which then cuts off I_C. If you let $V_{\text{in}}(t)$ be a sine wave of 2 V amplitude, the result (as shown below) looks a lot like the output of the half-wave rectifier (shown above right) from Lab 6. So we see that V_{out} "follows" V_{in}, but with two limitations: V_{out} stays about one diode drop below V_{in}, and V_{out} refuses to go below 0 V.[13]



Since the whole point of a follower is to have a much higher input impedance than that of its downstream load and a much smaller output impedance than that of its upstream source, let's evaluate the emitter follower's input impedance. If we change V_{in} by ΔV, then V_{out} changes by the same ΔV (as long as $V_{BE} \gtrsim 0.7$ V), so $\Delta I_E = \Delta V/R$. Thus

$$\Delta I_B = \frac{\Delta I_E}{\beta+1} = \frac{\Delta V}{(\beta+1)R} \quad \Rightarrow \quad R_{\text{in}} = \frac{\mathrm{d}V_{\text{in}}}{\mathrm{d}I_{\text{in}}} = (\beta+1)R.$$

The follower makes the load appear (to the source) to be a factor $\beta+1 \sim 100$ larger;

[12]If you don't understand why the ideal input impedance is large and the ideal output impedance is small, go back and re-read the first section of this week's notes, or bug one of us for a better explanation.

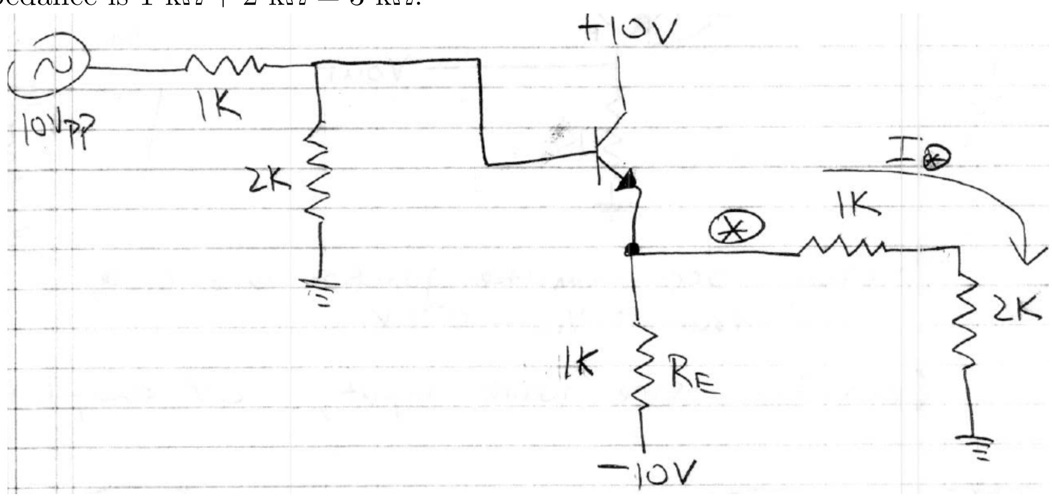
[13]We'll see in just a minute that the second problem is easily cured by replacing the ground connection beneath the resistor with a −10 V power supply.

and for a voltage source, larger R_{in} is easier to drive. By the way, you might have asked, "where is the load?" for the emitter follower. If a load were connected at V_{out}, it would appear in parallel with resistor R. So really we just showed that for the emitter follower, $R_{\text{in}} = (\beta+1)(R \parallel R_{\text{load}})$. The next example will show the upstream and downstream circuit fragments explicitly.

Now let's prevent the follower from "clipping" at zero volts by using $V_{EE} = -10$ V instead of ground. And let's actually use the emitter follower in the circuit in which we used an opamp follower in Lab 7.

First, let me digress for more nomenclature. The potential difference between the base and ground is called V_B, and similarly the collector voltage is V_C and emitter voltage is V_E, with respect to ground. We already introduced $V_{BE} = V_B - V_E$ and $V_{CE} = V_C - V_E$. Two other commonly used variables are V_{CC} for the positive power supply voltage that sits up above the collector and V_{EE} for the negative power supply voltage that sits down below the emitter. Also, the resistor that sits between the emitter and V_{EE} is normally labeled R_E.

Now look at the circuit below. A 10 V_{pp} (5 V amplitude) sine wave drives a voltage divider. The output impedance (a.k.a. Thévenin resistance) of the voltage divider is 1 kΩ $\parallel$ 2 kΩ = 667 Ω. Then we have an emitter follower with $V_{CC} = +10$ V, $V_{EE} = -10$ V, and $R_E = 1$ kΩ. Finally, we have another voltage divider, whose input impedance is $1\ \text{k}\Omega + 2\ \text{k}\Omega = 3\ \text{k}\Omega$.



The first question is, what is the input impedance of the emitter follower? From the argument we made above, it must be

$$R_{\text{in}}^{(\text{follower})} = (\beta+1)(R_E \parallel R_{\text{load}}) = (\beta+1)(1\ \text{k}\Omega \parallel 3\ \text{k}\Omega) \approx 100 \times 750\ \Omega \approx 75\ \text{k}\Omega.$$

Since 75 kΩ $\gg$ 667 Ω, the upstream voltage divider will have no trouble at all driving the emitter follower: V_B will be a sine wave of 3.33 V amplitude (6.67 V_{pp}).

Next, let's consider the output impedance of the emitter follower. We consider R_E to be part of the follower itself, while the downstream 1 kΩ and 2 kΩ resistors form an $R_{\rm load} = 3$ kΩ load. We want to know how small an $R_{\rm load}$ this emitter follower could drive without letting V_E droop. To evaluate this, we consider the voltage $V_* = V_E$ and the current I_* indicated on the figure above. The output impedance of the emitter follower will be $R_{\rm out} = -{\rm d}V_*/{\rm d}I_*$.

To compute $R_{\rm out}$, let's first consider only the path through the emitter (e.g. the case in which $R_E \to \infty$). Then if we wiggle the base voltage V_B, then

$$\Delta I_B = \Delta I_*/(\beta + 1)$$

and

$$\Delta V_* = \Delta V_B = \Delta I_B \cdot R_{\rm source} = \Delta I_B \cdot 667\ \Omega,$$

where $R_{\rm source}$ means the source impedance (a.k.a. Thévenin resistance) of the voltage divider to the left of the base. So far we have

$$R_{\rm out} \ = \ \frac{{\rm d}V_*}{{\rm d}I_*} \ = \ R_{\rm source}/(\beta+1) \ \approx \ 667\ \Omega/100 \ \approx \ 7\ \Omega.$$

Now let's also include the parallel resistance of R_E: as it turns out, R_E makes almost no difference in parallel with the very small impedance of the path through the emitter.

$$R_{\rm out} \ = \ \frac{R_{\rm source}}{\beta+1} \parallel R_E \ \approx \ 7\ \Omega \parallel 1\ {\rm k}\Omega \ \approx \ 7\ \Omega.$$

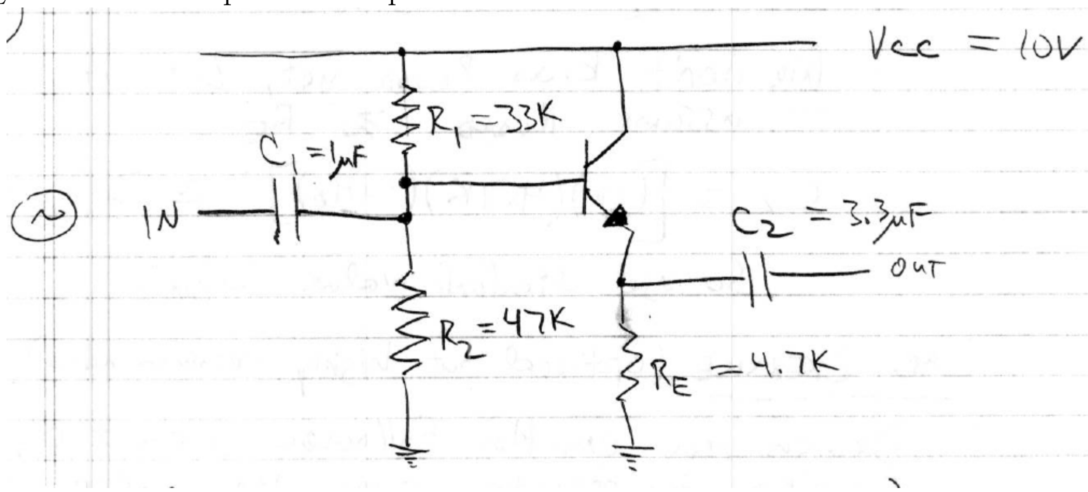
This is a good outcome. With $R_{\rm out} \approx 7\ \Omega$,[14] the follower will have absolutely no trouble driving the 3 kΩ load we have in the diagram above; in fact, it could easily drive a load resistance as small as a few hundred ohms. So the emitter follower is a good voltage source.

We have shown that the emitter follower does its job: the output (V_E) looks just like the input (V_B), except for the annoying ≈ 0.7 V offset. More important, the emitter follower makes the source think that it is driving a load that is $\mathcal{O}(10^2)$ times larger than the actual downstream load, and it makes the load think that it is being driven by a voltage source that is $\mathcal{O}(10^2)$ times stiffer (lower $R_{\rm thev}$) than the actual upstream voltage source. In other words, the upstream voltage divider sees a 75 kΩ load instead of the original 3 kΩ load, and the downstream voltage divider sees an upstream voltage source having $R_{\rm thev} = 7\ \Omega$, which is far better than the original 667 Ω. This is not nearly as large an improvement as you get from an opamp, and there is the annoying diode drop between $V_{\rm in}$ and $V_{\rm out}$, but it's not bad.

Next, let's **bias** the follower (and capacitively couple (a.k.a. "AC couple") the input and output) so that we can operate it from a single (positive) power supply. You

[14]It actually turns out that for this circuit $R_{\rm out} \approx 30\ \Omega$ — a fact that will make sense by the end of these notes when we have introduced "r_e." In any case, $R_{\rm out}$ is much smaller than the $R_{\rm thev} = 667\ \Omega$ of the upstream voltage divider.

will see this often with transistor amplifiers: the DC behavior is set up to put the transistor at a comfortable operating point, and the signal to be amplified is treated as an AC-only perturbation about this DC operating point. The input and output capacitors enable this separation. (You saw a preview of this AC-vs.-DC separation in the microphone amplifier from Lab 9.) The figure below shows an example of a biased emitter follower, with AC-coupled input and output. One handy side-effect of blocking DC signals at the input and output is that the 0.7 V offset between V_B and V_E doesn't show up at the output.



Here are the design steps for an AC-coupled emitter follower, as spelled out in section 2.05 of Horowitz & Hill's *The Art of Electronics* (2nd edition).

(1) Choose V_E to permit the largest symmetric signal amplitude: $V_E = \frac{1}{2}V_{CC}$.

(2) Choose R_E for a reasonable *quiescent* (a.k.a. DC, steady state) current. In this case, $R_E = 4.7\ \text{k}\Omega$, which gives $I_E = V_E/R_E \approx 1$ mA.[15]

(3) Choose R_1 and R_2 to arrange that $V_B = V_E + 0.7$ V: that implies that $R_2/R_1 = (5.7\ \text{V})/(4.3\ \text{V}) \approx 1.33$. Now that we know the ratio of R_1 and R_2, how do we decide how large to make them? Well, the rule of thumb for keeping the output of a voltage divider from drooping is $R_{\text{thev}} \ll R_{\text{load}}$. The Thévenin resistance of the voltage divider is $R_1 \parallel R_2$, and the load seen by the voltage divider is[16] $\beta R_E \approx 500\ \text{k}\Omega$. So we want $R_1 \parallel R_2 \ll 500\ \text{k}\Omega$. For example, we could choose $R_1 \parallel R_2 \approx 25\ \text{k}\Omega$. Then the desired values would be $R_1 = 38\ \text{k}\Omega$ and $R_2 = 50\ \text{k}\Omega$. We will instead choose $R_1 = 33\ \text{k}\Omega$ and $R_2 = 47\ \text{k}\Omega$, which are pretty close to the desired values and are standard 5% resistor values. Then we have

$$V_B = V_{CC} \cdot \frac{R_2}{R_1 + R_2} = (10\ \text{V}) \cdot \frac{47\ \text{k}\Omega}{33\ \text{k}\Omega + 47\ \text{k}\Omega} \approx 5.9\ \text{V}.$$

[15] I see now that I don't offer any explanation for why 1 mA is a "reasonable" DC current through the emitter.

[16] Approximately — we're neglecting whatever load might be present at the output.

(4) Choose C_1 such that $f_{3\text{dB}}$ is less than any signal frequency of interest. In the Horowitz & Hill example, the goal is to build a follower for audio frequencies (20 Hz $<$ $f <$ 20 kHz), so they choose $f_{3\text{dB}} = 10$ Hz. Then

$$10 \text{ Hz} = f_{3\text{dB}} = \frac{1}{2\pi RC}$$

where

$$R \approx R_1 \parallel R_2 \parallel (\beta R_E) \approx 33 \text{ k}\Omega \parallel 47 \text{ k}\Omega \parallel 470 \text{ k}\Omega \approx 20 \text{ k}\Omega.$$

So then

$$C_1 = \frac{1}{2\pi R f_{3\text{dB}}} = \frac{1}{(2\pi)(20 \text{ k}\Omega)(10 \text{ Hz})} \approx 0.8 \ \mu\text{F}.$$

Rounding off to $C_1 = 1 \ \mu$F gives $f_{3\text{dB}} \approx 8$ Hz.

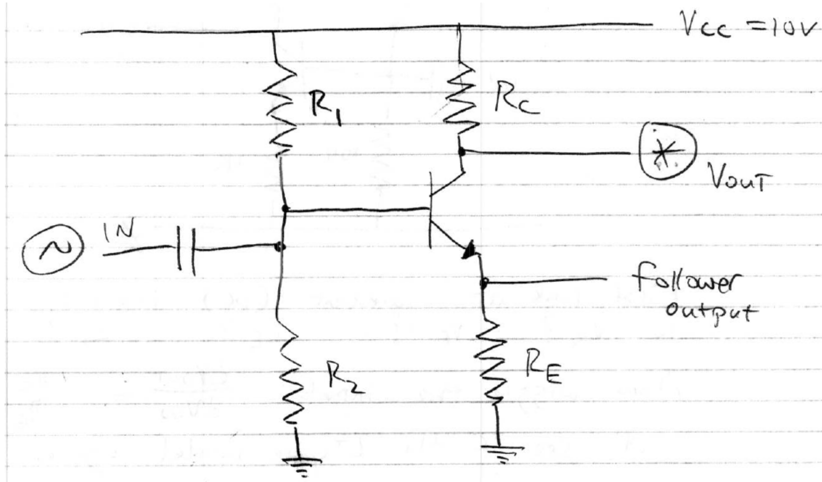
(5) Choose C_2 for $f_{3\text{dB}} \leq 10$ Hz (the minimal signal frequency of interest, which in this case is 10 Hz). To do this properly, we need to know the resistance of the load that would be downstream of C_2, since $f_{3\text{dB}} = 1/(2\pi R_{\text{load}} C_2)$. But it is safe to assume that $R_{\text{load}} \gtrsim R_E$ (since otherwise the emitter output would droop), so we choose $C_2 = [(2\pi)(4.7 \text{ k}\Omega)(10 \text{ Hz})]^{-1} \approx 3.2 \ \mu$F. In practice, the nearest standard capacitor value is 3.3 μF.

Wow! So you can see that designing a transistor-based follower is a lot more involved than designing an opamp-based follower! The transistor portion of the course is intended both to give you a sense of what is going on inside an opamp and to give you some appreciation for just how convenient it is to be able to work with opamps. If your head hurts after you first read about transistors, don't worry — this is perfectly normal!

Optional lab exercise for the ambitious student (from Horowitz & Hill): Design an emitter follower with ± 10 V power supplies to operate over the audio range (20 Hz – 20 kHz). Use 3 mA quiescent current. First assume that the input source provides a DC path to ground. Then modify your design to use a capacitively coupled input. Remember that there must be a DC path from base to ground.[17]

Our next circuit example is the **common emitter amplifier**, which is also described in the transistor chapter of Eggleston's textbook. We start from the emitter follower and simply add a resistor R_C at the collector, as shown in the figure below.

[17]If there is no DC path from base to ground (or to some other supply voltage), then there is no path for the base current I_B to flow, and the transistor then cannot operate (since $I_C = \beta I_B$ in the active mode).



Now what happens when we wiggle V_{in}? Well, as before, the emitter voltage follows the base voltage: $V_E \approx V_B - 0.7$ V. So then the current through the emitter is $I_E = V_E/R_E \approx (V_B - 0.7\text{ V})/R_E$. Now the collector current is

$$I_C = I_E - I_B = I_E - \frac{I_E}{\beta + 1} \approx I_E$$

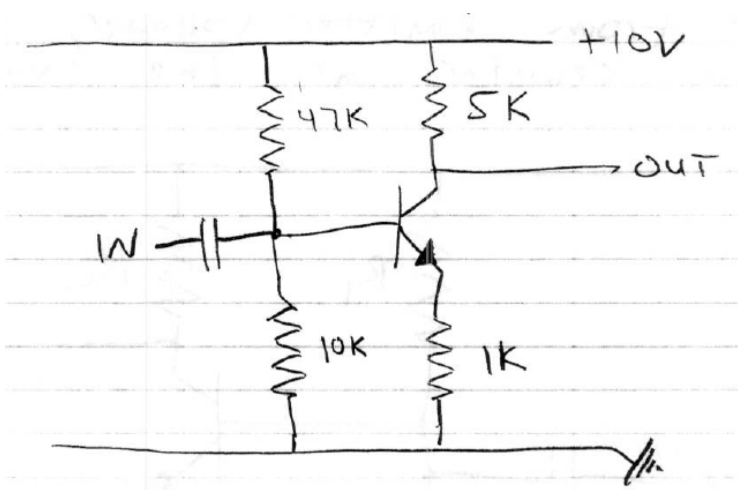
since $\beta \gg 1$. So then

$$V_* = V_{CC} - I_C R_C \approx V_{CC} - I_E R_C = V_{CC} - \frac{R_C}{R_E} \cdot (V_B - 0.7\text{ V}).$$

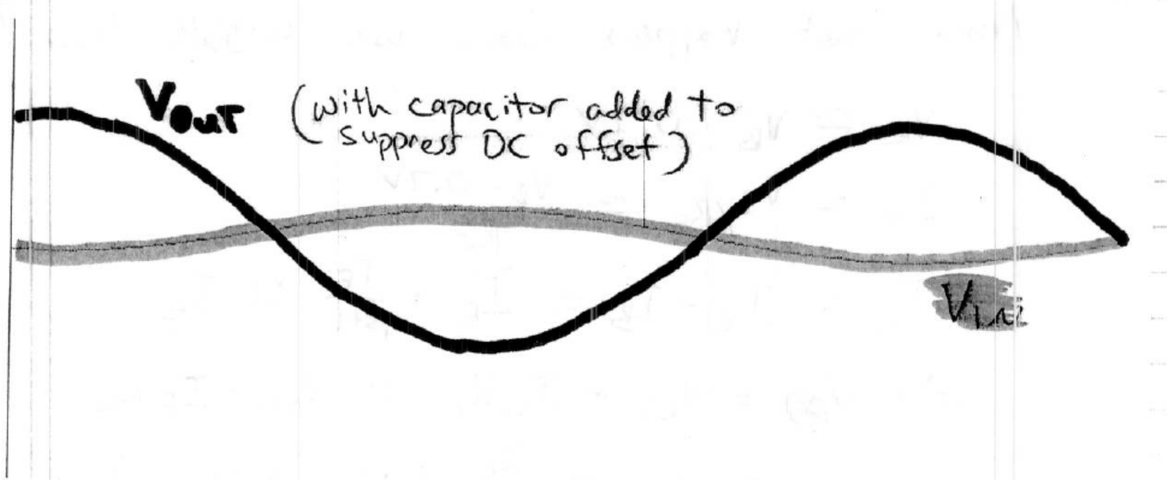
So

$$\frac{\mathrm{d}V_*}{\mathrm{d}V_{\mathrm{in}}} = -\frac{R_C}{R_E}.$$

We have an inverting amplifier, with gain $-R_C/R_E$. We can choose the gain by suitable component selection. Let's fill in some component values in the drawing below.



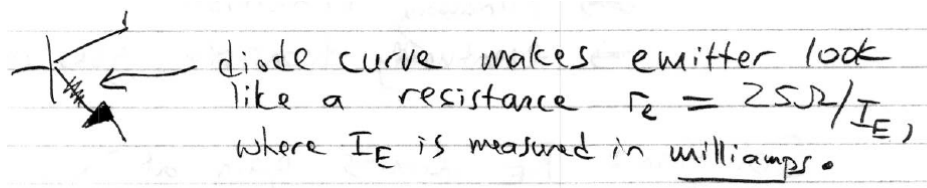
First look at the quiescent (DC) state: What is V_B? V_E? I_E? I_C? V_{out}? Think about these for yourself, then compare with my answer in the footnotes.[18] Now wiggle the input. $\mathrm{d}V_{\text{out}}/\mathrm{d}V_{\text{in}} = -R_C/R_E = -5$. Let's see if the CircuitLab model agrees (see graph below, which shows V_{in} and V_{out} vs. time, with DC offset of V_{out} suppressed).



Now suppose that we get greedy and try to increase the voltage gain of our common emitter amplifier by reducing R_E to a very small value. We encounter two problems.

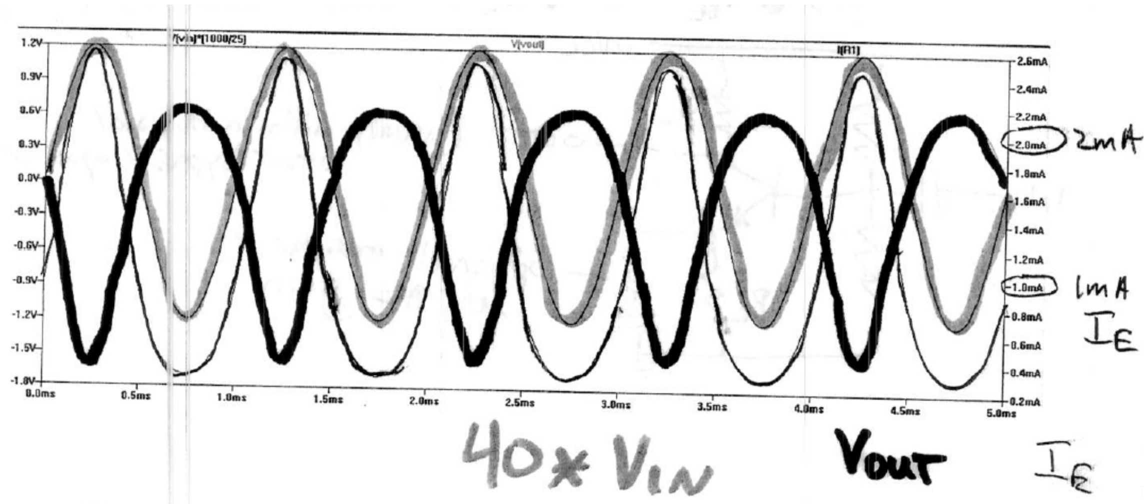
The first problem with using $R_E = 0$ arises from the non-linear diode-like nature of the base-emitter junction. Recall that $I_B = I_0 \cdot \left(\exp(\frac{V_{BE}}{V_T}) - 1\right) \approx I_0 \exp(V_{BE}/25\text{ mV})$ and that in active mode, $I_E = (\beta+1) I_B$. Differentiating, we have

$$\frac{\mathrm{d}V_{BE}}{\mathrm{d}I_E} \approx \frac{25\text{ mV}}{I_E} = \frac{25\ \Omega}{I_E[\text{mA}]}.$$

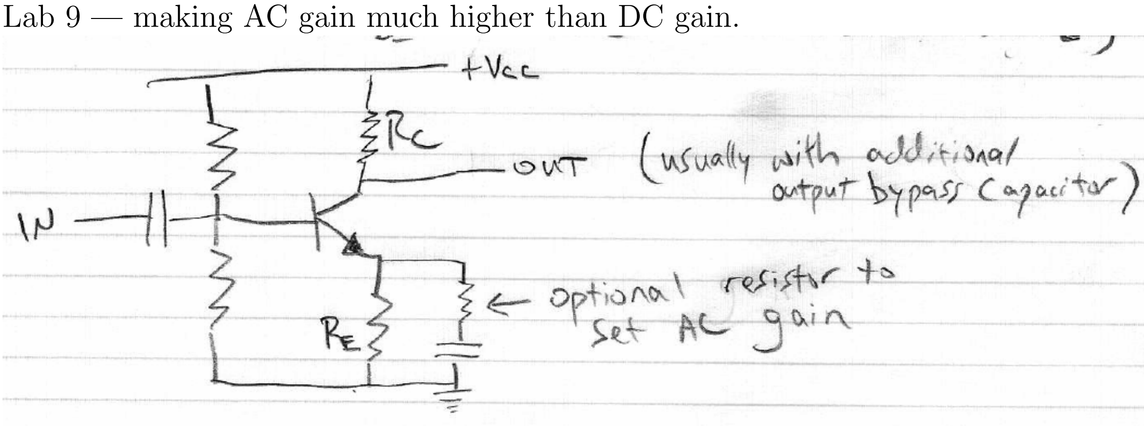


The diode-like relationship between V_{BE} and I_E makes the emitter look like a resistance $r_e = 25\ \Omega/I_E$, where I_E is measured in *milliamps.* If we don't keep $R_E \gg r_e$, we will see non-linear response: the "barn roof" distortion that we saw from the opamp logarithm circuit in Lab 8. The figure below illustrates the effect of using $R_C = 1\text{ k}\Omega$ and $R_E = 0\ \Omega$. With $I_E \sim 1$ mA (but varying between 0.2 mA and 2.5 mA during the cycle), we expect a gain around 40, which we see, but V_{out} is distorted by the fact that the gain varies as V_{in} changes.

[18] Don't peek until you've worked these numbers out for yourself! I find $V_B = 1.75$ V, $V_E \approx 1.05$ V, $I_E = V_E/R_E \approx 1$ mA, $I_C \approx I_E \approx 1$ mA, $V_{\text{out}} = V_{CC} - I_C R_C = 5$ V.



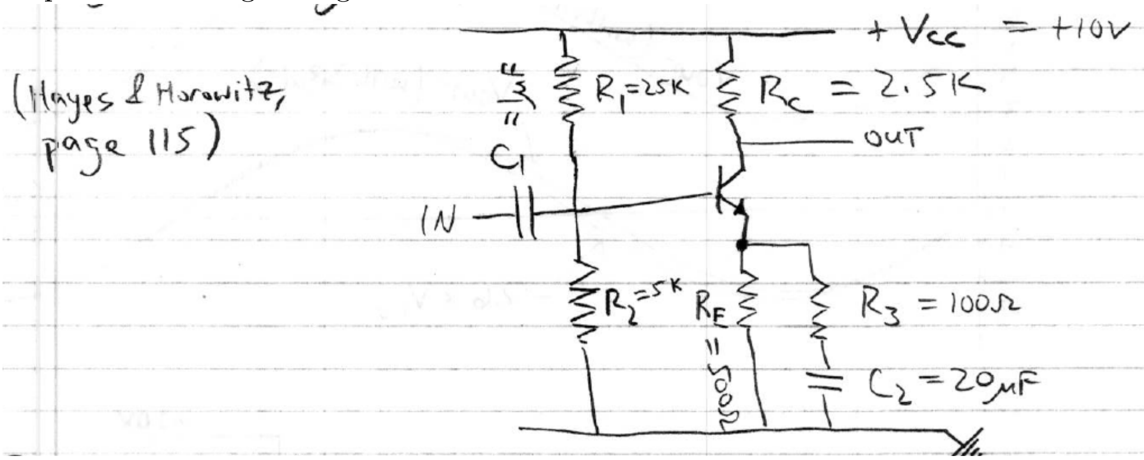
The second problem with letting $R_E \to 0$ is *thermal instability.* (This is pretty detailed, so I don't expect you to remember it.) It turns out that I_0 increases as temperature increases: at constant V_{BE}, I_C grows 9% per °C.[19] Omitting R_E gives you high gain, which then causes the transistor to dissipate a lot of power $V_{CE}I_C$, which then heats up the transistor, which then increases I_C, which then decreases r_e, which then allows I_C to increase even further. In this runaway condition, I_C grows until the transistor goes into saturation, at which point it is no longer behaving as a useful amplfier. The solution (shown in the figure below) to the thermal runaway problem is to choose R_E large enough to limit I_C to the milliamp range at DC; and then use a bypass capacitor to increase the AC gain. (Then if temperature increases, I_C increases, with $I_E \approx I_C$. Increasing the I_E flowing through R_E raises V_E, thus lowering V_{BE}. Reducing V_{BE} in turn lowers I_C. So we have used feedback to stabilize I_C against thermal runaway.) We used a similar trick in the microphone amplifier of Lab 9 — making AC gain much higher than DC gain.



Finally, let's go through the Horowitz & Hill design example for a common emitter

[19]Alternatively, at constant I_C, V_{BE} falls ≈ 2 mV per °C. This is one reason why an opamp's V_{offset} varies with temperature.

amplifier with large AC gain.[20]



(1) Choose I_C (quiescent) to be a few milliamps; choose R_C to center V_{out} between V_{CC} and V_{EE}, given I_C. For example, if $V_{CC} = +10$ V, $I_C = 2$ mA, then choose $R_C = (5\text{ V})/(2\text{ mA}) = 2.5\text{ k}\Omega$.

(2) Choose R_E to put $V_E \approx 1$ V, for temperature stability. So $R_E = (1\text{ V})/(2\text{ mA}) = 500\ \Omega$.

(3) Choose $R_1 \parallel R_2$ such that $R_{\text{in}}^{(\text{transistor})} \approx (\beta R_E) \gg (R_1 \parallel R_2)$, so that the transistor is not too difficult a load for the bias network (i.e. for the voltage divider). So $R_1 \parallel R_2 \ll 50\text{ k}\Omega$.

(4) Choose R_1, R_2 to put V_B at $V_E + 0.7\text{ V} = 1.7$ V. So for instance choose $R_2 = 5\text{ k}\Omega$ and $R_1 = R_2 \cdot \frac{8.3\text{ V}}{1.7\text{ V}} = 25\text{ k}\Omega$. Then $R_1 \parallel R_2 \approx 5\text{ k}\Omega \ll 50\text{ k}\Omega$.

(5) Choose R_3 (if any) for AC gain. For example, if $R_3 = 100\ \Omega$ then gain is $-R_C/\left(r_e + (R_E \parallel (R_3 + Z_{C2}))\right)$. If C_2 is large enough to be neglected, we get

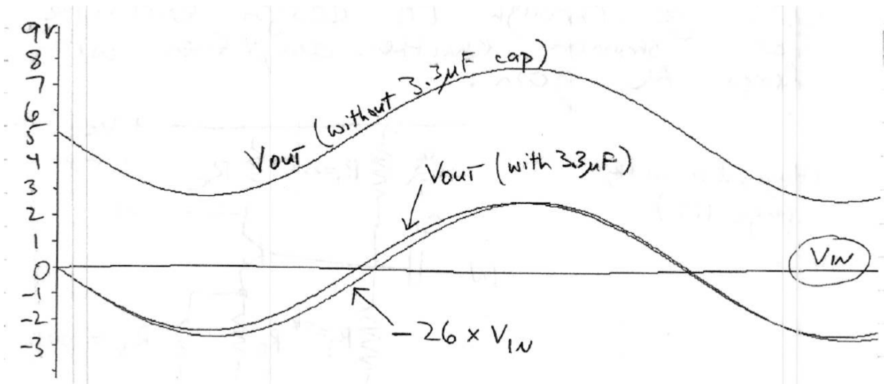
$$\text{gain} \approx -\frac{R_C}{r_e + (R_E \parallel R_3)} \approx -\frac{2.5\text{ k}\Omega}{\frac{25\text{ mV}}{2\text{ mA}} + (500\ \Omega \parallel 100\ \Omega)} \approx -\frac{2.5\text{ k}\Omega}{96\ \Omega} \approx -26.$$

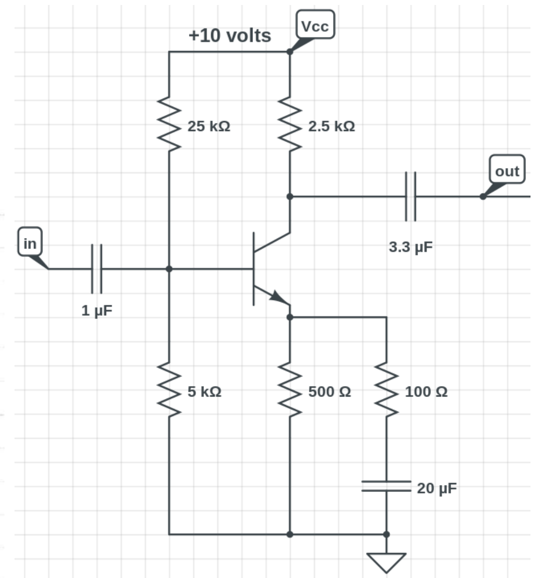
(6) Choose C_2 for $f_{3\text{dB}} = \frac{1}{2\pi R C_2}$, where $R = R_3 + r_e = 113\ \Omega$. For example, $C_2 = 20\ \mu$F gives $f_{3\text{dB}} = 70$ Hz.

(7) Choose C_1 for $f_{3\text{dB}}$, where $R = R_1 \parallel R_2 \parallel (\beta \cdot (R_E \parallel R_3))$, which is about $5\text{ k}\Omega \parallel 25\text{ k}\Omega \parallel 10\text{ k}\Omega \approx 3\text{ k}\Omega$. If we use 70 Hz to match (6), we find $C = 0.75\ \mu$F. So we'll round it to an available capacitor value and use 1 μF.

Now let's try it in CircuitLab! (See below.)

[20] Taken from Hayes & Horowitz, page 115.





The key points to take away from this reading are:

- A transistor allows a small current I_B (or analogously a small voltage V_{BE}) to control a large current I_C. A device that enables a small signal to control a much larger signal makes it possible to build amplifiers and switches — the two most common uses of transistors.
- The dominant current through an NPN transistor flows in at the collector and out at the emitter. A much smaller current flows in at the base. In active mode, I_C and I_B are proportional by a factor β, with $50 \lesssim \beta \lesssim 300$.
- The knob for controlling I_C is the diode-like base→emitter junction: in active mode, $I_C = \beta I_B \approx \beta I_0 \exp(V_{BE}/25 \text{ mV})$.
- Because the B→E junction looks like a diode, $V_{BE} \approx 0.7$ V when the transistor is active. This simple observation, that the emitter stays about a diode drop below the base, is the key to analyzing the emitter follower and the common emitter amplifier.
- These ideas are sufficient for understanding why an emitter follower works as a follower, and why a common emitter amplifier works as an inverting amplifier with gain $-R_C/R_E$.
- To analyze the case in which R_E becomes small, we need to consider more carefully the diode-like nature of the B→E junction to find the *dynamic resistance*[21] $r_e = 25\ \Omega/(I_E\ [\text{mA}])$ that behaves as if it were in series with the emitter.

[21] In general, dynamic resistance means $\mathrm{d}V/\mathrm{d}I$, the slope of the I-V curve.

- To balance large gain at signal (AC) frequencies with good stability at DC, we split the gain-setting emitter resistor R_E into two paths, thus reducing the gain at DC. Remember that we used a similar trick to amplify the microphone signal in Lab 9, with high gain at AC and low gain at DC; in that case, the motivation for low DC gain was to avoid amplifying the opamp's imperfections.

These basic rules let you understand the operation of an NPN Bipolar Junction Transistor in active mode. We used these rules to analyze two important circuits: the emitter follower, which is the BJT equivalent of the opamp follower (voltage gain = 1, but current out can be much larger than current in); and the common emitter amplifier, which is the BJT equivalent of the opamp inverting amplifier (voltage gain can be large).

Transistors are a very big deal (1956 Nobel Prize in physics). They are the basis of both the analog amplifiers that make your stereo audible and the digital computers that make your smartphone so capable.

The transistor is the key component of an opamp like the '741. (See 741 schematic below — we'll see next week how to recognize a few more fragments of this circuit.) But trying to understand transistor circuits can be much more involved than trying to understand opamp circuits. You can already see how much easier life is when someone else puts the transistors together for you into an opamp. I hope you see now why we chose to cover opamps before transistors in this course — so that you have a better sense of how amplifiers are used before we look into the details of how they are implemented using transistors.