TRANSISTORS

INTRODUCTION

The transistor is our most important example of an "active" component, a device that can amplify, producing an output signal with more power in it than the input signal. The additional power comes from an external source of power (the power supply, to be exact). Note that voltage amplification isn't what matters, since, for example, a step-up transformer, a "passive" component just like a resistor or capacitor, has voltage gain but no power gain. Devices with power gain are distinguishable by their ability to make oscillators, by feeding some output signal back into the input.

It is interesting to note that the property of power amplification seemed very important to the inventors of the transistor. Almost the first thing they did to convince themselves that they had really invented something was to power a loudspeaker from a transistor, observing that the output signal sounded louder than the input signal.

The transistor is the essential ingredient of every electronic circuit, from the simplest amplifier or oscillator to the most elaborate digital computer. Integrated circuits (ICs), which have largely replaced circuits constructed from discrete transistors, are themselves merely arrays of transistors and other components built from a single chip of semiconductor material.

A good understanding of transistors is very important, even if most of your circuits are made from ICs, because you need to understand the input and output properties of the IC in order to connect it to the rest of your circuit and to the outside world. In addition, the transistor is the single most powerful resource for interfacing, whether between ICs and other circuitry or between one subcircuit and another. Finally, there are frequent (some might say too frequent) situations where the right IC just doesn't exist, and you have to rely on discrete transistor circuitry to do the job. As you will see, transistors have an excitement all their own. Learning how they work can be great fun.

Our treatment of transistors is going to be quite different from that of many other books. It is common practice to use the h-parameter model and equivalent circuit. In our opinion that is unnecessarily complicated and unintuitive. Not only does circuit behavior tend to be revealed to you as something that drops out of elaborate equations, rather than deriving from a clear understanding in your own mind as to how the circuit functions; you also have the tendency to lose sight of which parameters of transistor behavior you can count on and, more important, which ones can vary over large ranges.

In this chapter we will build up instead a very simple introductory transistor model and immediately work out some circuits with it. Soon its limitations will become apparent; then we will expand the model to include the respected Ebers-Moll conventions. With the Ebers-Moll equations and a simple 3-terminal model, you will have a good understanding of transistors; you won't need to do a lot of calculations, and your designs will be first-rate. In particular, they will be largely independent of the poorly controlled transistor parameters such as current gain.

Some important engineering notation should be mentioned. Voltage at a transistor terminal (relative to ground) is indicated by a single subscript (C, B, or E): V_C is the collector voltage, for instance. Voltage between two terminals is indicated by a double subscript: V_{BE} is the base-to-emitter voltage drop, for instance. If the same letter is repeated, that means a power-supply voltage: V_{CC} is the (positive) power-supply voltage associated with the collector, and V_{EE} is the (negative) supply voltage associated with the emitter.

2.01 First transistor model: current amplifier

Let's begin. A transistor is a 3-terminal device (Fig. 2.1) available in 2 flavors (*npn* and *pnp*), with properties that meet the following rules for *npn* transistors (for *pnp* simply reverse all polarities):

1. The collector must be more positive than the emitter.

2. The base-emitter and base-collector circuits behave like diodes (Fig. 2.2). Normally the base-emitter diode is conducting and the base-collector diode is reverse-biased, i.e., the applied voltage is in the opposite direction to easy current flow.

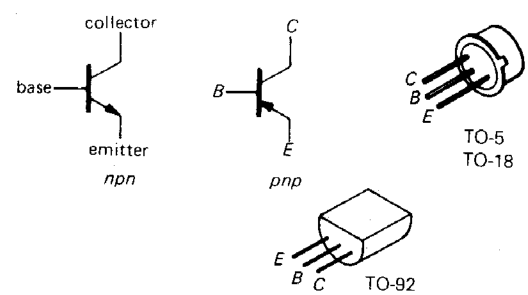

Figure 2.1. Transistor symbols, and small transistor packages.

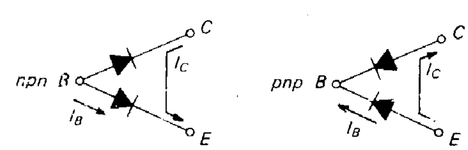

Figure 2.2. An ohmmeter's view of a transistor's terminals.

3. Any given transistor has maximum values of I_C, I_B, and V_{CE} that cannot be exceeded without costing the exceeder the price of a new transistor (for typical values, see Table 2.1). There are also other limits, such as power dissipation ($I_C V_{CE}$), temperature, V_{BE}, etc., that you must keep in mind.

4. When rules 1–3 are obeyed, I_C is roughly proportional to I_B and can be written as

$$I_C = h_{FE} I_B = \beta I_B$$

where h_{FE}, the current gain (also called beta), is typically about 100. Both I_C and I_E flow to the emitter. Note: The collector current is not due to forward conduction of the base-collector diode;

that diode is reverse-biased. Just think of it as "transistor action."

Property 4 gives the transistor its usefulness: A small current flowing into the base controls a much larger current flowing into the collector.

Warning: h_{FE} is not a "good" transistor parameter; for instance, its value can vary from 50 to 250 for different specimens of a given transistor type. It also depends upon the collector current, collector-to-emitter voltage, and temperature. *A circuit that depends on a particular value for h_{FE} is a bad circuit.*

Note particularly the effect of property 2. This means you can't go sticking a voltage across the base-emitter terminals, because an enormous current will flow if the base is more positive than the emitter by more than about 0.6 to 0.8 volt (forward diode drop). This rule also implies that an operating transistor has $V_B \approx V_E + 0.6$ volt ($V_B = V_E + V_{BE}$). Again, polarities are normally given for *npn* transistors; reverse them for *pnp*.

Let us emphasize again that you should not try to think of the collector current as diode conduction. It isn't, because the collector-base diode normally has voltages applied across it in the reverse direction. Furthermore, collector current varies very little with collector voltage (it behaves like a not-too-great current source), unlike forward diode conduction, where the current rises very rapidly with applied voltage.

SOME BASIC TRANSISTOR CIRCUITS

2.02 Transistor switch

Look at the circuit in Figure 2.3. This application, in which a small control current enables a much larger current to flow in another circuit, is called a transistor switch. From the preceding rules it is easy to understand. When the mechanical switch is open, there is no base current. So, from rule 4, there is no collector current. The lamp is off.

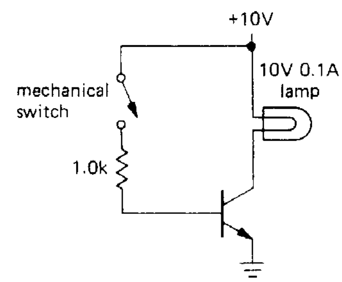

Figure 2.3. Transistor switch example.

When the switch is closed, the base rises to 0.6 volt (base-emitter diode is in forward conduction). The drop across the base resistor is 9.4 volts, so the base current is 9.4mA. Blind application of rule 4 gives $I_C = 940$mA (for a typical beta of 100). That is wrong. Why? Because rule 4 holds only if rule 1 is obeyed; at a collector current of 100mA the lamp has 10 volts across it. To get a higher current you would have to pull the collector below ground. A transistor can't do this, and the result is what's called saturation – the collector goes as close to ground as it can (typical saturation voltages are about 0.05–0.2V, see Appendix G) and stays there. In this case, the lamp goes on, with its rated 10 volts across it.

Overdriving the base (we used 9.4mA when 1.0mA would have barely sufficed) makes the circuit conservative; in this particular case it is a good idea, since a lamp draws more current when cold (the resistance of a lamp when cold is 5 to 10 times lower than its resistance at operating current). Also transistor beta drops at low collector-to-base voltages, so some extra base current is necessary to bring a transistor into full saturation (see Appendix G). Incidentally, in a real circuit you would probably put a resistor from base to ground (perhaps 10k in this case) to make sure the base is at ground with the switch open. It wouldn't affect the "on" operation, because it would sink only 0.06mA from the base circuit.

There are certain cautions to be observed when designing transistor switches:
1. Choose the base resistor conservatively to get plenty of excess base current, especially when driving lamps, because of the reduced beta at low V_{CE}. This is also a good idea for high-speed switching, because of capacitive effects and reduced beta at very high frequencies (many megahertz). A small "speedup" capacitor is often connected across the base resistor to improve high-speed performance.
2. If the load swings below ground for some reason (e.g., it is driven from ac, or it is inductive), use a diode in series with the collector (or a diode in the reverse direction to ground) to prevent collector-base conduction on negative swings.
3. For inductive loads, protect the transistor with a diode across the load, as shown in Figure 2.4. Without the diode the inductor will swing the collector to a large positive voltage when the switch is opened, most likely exceeding the collector-emitter breakdown voltage, as the inductor tries to maintain its "on" current from V_{CC} to the collector (see the discussion of inductors in Section 1.31).

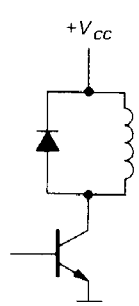

Figure 2.4. Always use a suppression diode when switching an inductive load.

Transistor switches enable you to switch very rapidly, typically in a small fraction of a microsecond. Also, you can switch many different circuits with a single control signal. One further advantage is the possibility of remote *cold switching*, in which only dc control voltages snake around through cables to reach front-panel switches, rather than the electronically inferior approach of having the signals themselves traveling through cables and switches (if you run lots of signals through cables, you're likely to get capacitive pickup as well as some signal degradation).

"Transistor man"

Figure 2.5 presents a cartoon that will help you understand some limits of transistor

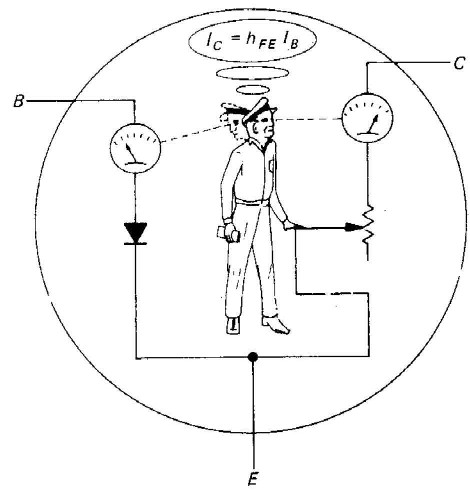

Figure 2.5. "Transistor man" observes the base current, and adjusts the output rheostat in an attempt to maintain the output current h_{FE} times larger.

behavior. The little man's perpetual task in life is to try to keep $I_C = h_{FE}I_B$; however, he is only allowed to turn the knob on the variable resistor. Thus he can go from a short circuit (saturation) to an open circuit (transistor in the "off" state), or anything in between, but he isn't allowed to use batteries, current sources, etc. One warning is in order here: Don't think that the collector of a transistor looks like a resistor. It doesn't. Rather, it looks approximately like a poor-quality constant-current sink (the value of current depending on the signal applied to the base), primarily because of this little man's efforts.

Another thing to keep in mind is that, at any given time, a transistor may be (a) cut off (no collector current), (b) in the active region (some collector current, and collector voltage more than a few tenths of a volt above the emitter), or (c) in saturation (collector within a few tenths of a volt of the emitter). See Appendix G on transistor saturation for more details.

2.03 Emitter follower

Figure 2.6 shows an example of an *emitter follower*. It is called that because the output terminal is the emitter, which follows the input (the base), less one diode drop:

$V_E \approx V_B - 0.6$ volt

The output is a replica of the input, but 0.6 to 0.7 volt less positive. For this circuit, V_{in} must stay at +0.6 volt or more, or else the output will sit at ground. By returning the emitter resistor to a negative supply voltage, you can permit negative voltage swings as well. Note that there is no collector resistor in an emitter follower.

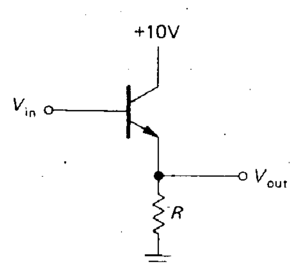

Figure 2.6. Emitter follower.

At first glance this circuit may appear useless, until you realize that the input impedance is much larger than the output impedance, as will be demonstrated shortly. This means that the circuit requires less power from the signal source to drive a given load than would be the case if the signal source were to drive the load directly. Or a signal of some internal impedance (in the Thévenin sense) can now drive a load of comparable or even lower impedance without loss of amplitude (from the usual voltage-divider effect). In other words, an emitter follower has current gain, even though it has no voltage gain. It has power gain. Voltage gain isn't everything!

Impedances of sources and loads

This last point is very important and is worth some more discussion before we calculate in detail the beneficial effects of emitter followers. In electronic circuits, you're always hooking the output of something to the input of something else, as suggested in Figure 2.7. The signal source might be the output of an amplifier stage (with Thévenin equivalent series impedance $\mathbf{Z}_{out}$), driving the next stage or perhaps a load (of some input impedance $\mathbf{Z}_{in}$). In general, the loading effect of the following stage causes a reduction of signal, as we discussed earlier in Section 1.05. For this reason it is usually best to keep $Z_{out} \ll Z_{in}$ (a factor of 10 is a comfortable rule of thumb).

In some situations it is OK to forgo this general goal of making the source stiff compared with the load. In particular, if the load is always connected (e.g., within a circuit) and if it presents a known and constant $\mathbf{Z}_{in}$, it is not too serious if it "loads" the source. However, it is always nicer if signal levels don't change when a load is connected. Also, if $\mathbf{Z}_{in}$ varies with signal level, then having a stiff source ($Z_{out} \ll Z_{in}$) assures linearity, where otherwise the level-dependent voltage divider would cause distortion.

Finally, there are two situations where $Z_{out} \ll Z_{in}$ is actually the wrong thing to

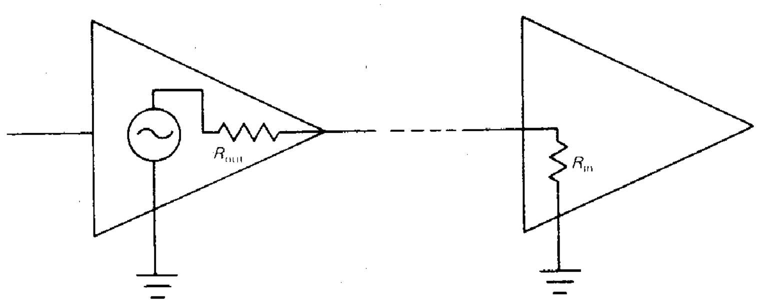

Figure 2.7. Illustrating circuit "loading" as a voltage divider.

do: In radiofrequency circuits we usually *match* impedances ($Z_{out} = Z_{in}$), for reasons we'll describe in Chapter 14. A second exception applies if the signal being coupled is a *current* rather than a voltage. In that case the situation is reversed, and one strives to make $Z_{in} \ll Z_{out}$ ($Z_{out} = \infty$, for a current source).

Input and output impedances of emitter followers

As you have just seen, the emitter follower is useful for changing impedances of signals or loads. To put it bluntly, that's the whole point of an emitter follower.

Let's calculate the input and output impedances of the emitter follower. In the preceding circuit we will consider R to be the load (in practice it sometimes is the load; otherwise the load is in parallel with R, but with R dominating the parallel resistance anyway). Make a voltage change ΔV_B at the base; the corresponding change at the emitter is $\Delta V_E = \Delta V_B$. Then the change in emitter current is

$$\Delta I_E = \Delta V_B / R$$

so

$$\Delta I_B = \frac{1}{h_{fe}+1}\Delta I_E = \frac{\Delta V_B}{R(h_{fe}+1)}$$

(using $I_E = I_C + I_B$). The input resistance is $\Delta V_B/\Delta I_B$. Therefore

$$r_{in} = (h_{fe}+1)R$$

The transistor beta (h_{fe}) is typically about 100, so a low-impedance load looks like a much higher impedance at the base; it is easier to drive.

In the preceding calculation, as in Chapter 1, we have used lower-case symbols such as h_{fe} to signify small-signal (incremental) quantities. Frequently one concentrates on the *changes* in voltages (or currents) in a circuit, rather than the steady (dc) values of those voltages (or currents). This is most common when these "small-signal" variations represent a possible signal, as in an audio amplifier, riding on a steady dc "bias" (see Section 2.05). The distinction between dc current gain (h_{FE}) and small-signal current gain (h_{fe} isn't always made clear, and the term beta is used for both. That's alright, since $h_{fe} \approx h_{FE}$ (except at very high frequencies), and you never assume you know them accurately, anyway.

Although we used resistances in the preceding derivation, we could generalize to complex impedances by allowing ΔV_B, ΔI_B, etc., to become complex numbers. We would find that the same transformation rule applies for impedances: $\mathbf{Z}_{\text{in}} = (h_{fe} + 1)\mathbf{Z}_{\text{load}}$.

We could do a similar calculation to find that the output impedance $\mathbf{Z}_{\text{out}}$ of an emitter follower (the impedance looking into the emitter) driven from a source of internal impedance $\mathbf{Z}_{\text{source}}$ is given by

$$\mathbf{Z}_{\text{out}} = \frac{\mathbf{Z}_{\text{source}}}{h_{fe} + 1}$$

Strictly speaking, the output impedance of the circuit should also include the parallel resistance of R, but in practice $\mathbf{Z}_{\text{out}}$ (the impedance looking into the emitter) dominates.

EXERCISE 2.1

Show that the preceding relationship is correct. Hint: Hold the source voltage fixed, and find the change in output current for a given change in output voltage. Remember that the source voltage is connected to the base through a series resistor.

Because of these nice properties, emitter followers find application in many situations, e.g., making low-impedance signal sources within a circuit (or at outputs), making stiff voltage references from higher-impedance references (formed from voltage dividers, say), and generally isolating signal sources from the loading effects of subsequent stages.

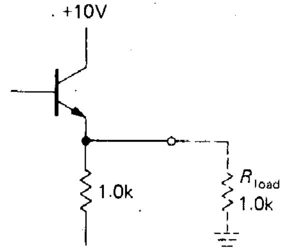

Figure 2.8. An *npn* emitter follower can source plenty of current through the transistor, but can sink limited current only through its emitter resistor.

EXERCISE 2.2

Use a follower with base driven from a voltage divider to provide a stiff source of +5 volts from an available regulated +15 volt supply. Load current (max) = 25mA. Choose your resistor values so that the output voltage doesn't drop more than 5% under full load.

Important points about followers

1. Notice (Section 2.01, rule 4) that in an emitter follower the *npn* transistor can only "source" current. For instance, in the loaded circuit shown in Figure 2.8 the output can swing to within a transistor saturation voltage drop of V_{CC} (about +9.9V), but it cannot go more negative than −5 volts. That is because on the extreme negative swing, the transistor can do no more than turn off, which it does at −4.4 volts input (−5V output). Further negative swing at the input results in backbiasing of the base-emitter junction, but no further change in output. The output, for a 10 volt amplitude sine-wave input, looks as shown in Figure 2.9.

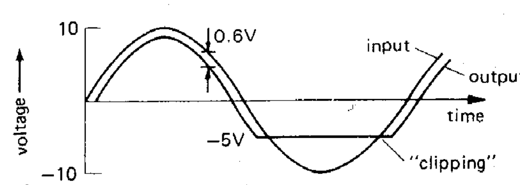

Figure 2.9. Illustrating the asymmetrical current drive capability of the *npn* emitter follower.

Another way to view the problem is to say that the emitter follower has low small-signal output impedance. Its large-signal output impedance is much larger (as large as R_E). The output impedance changes over from its small-signal value to its large-signal value at the point where the transistor goes out of the active region (in this case at an output voltage of -5V). To put this point another way, a low value of small-signal output impedance doesn't necessarily mean that the circuit can generate large signal swings into a low-resistance load. Low small-signal output impedance doesn't imply large output current capability.

Possible solutions to this problem involve either decreasing the value of the emitter resistor (with greater power dissipation in resistor and transistor), using a *pnp* transistor (if all signals are negative only), or using a "push-pull" configuration, in which two complementary transistors (one *npn*, one *pnp*), are used (Section 2.15). This sort of problem can also come up when the load of an emitter follower contains voltage or current sources of its own. This happens most often with regulated power supplies (the output is usually an emitter follower) driving a circuit that has other power supplies.

2. Always remember that the base-emitter reverse breakdown voltage for silicon transistors is small, quite often as little as 6 volts. Input swings large enough to take the transistor out of conduction can easily result in breakdown (with consequent degradation of h_{FE}) unless a protective diode is added (Fig. 2.10).

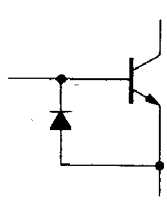

Figure 2.10. A diode prevents base-emitter reverse voltage breakdown.

3. The voltage gain of an emitter follower is actually slightly less than 1.0, because the base-emitter voltage drop is not really constant, but depends slightly on collector current. You will see how to handle that later in the chapter, when we have the Ebers-Moll equation.

2.04 Emitter followers as voltage regulators

The simplest regulated supply of voltage is simply a zener (Fig. 2.11). Some current must flow through the zener, so you choose

$$\frac{V_{in} - V_{out}}{R} > I_{out}\ (\text{max})$$

Because V_{in} isn't regulated, you use the lowest value of V_{in} that might occur for this formula. This is called worst-case design. In practice, you would also worry about component tolerances, line-voltage limits, etc., designing to accommodate the worst possible combination that would ever occur.

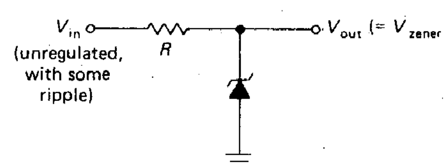

Figure 2.11. Simple zener voltage regulator.

The zener must be able to dissipate

$$P_{zener} = \left(\frac{V_{in} - V_{out}}{R} - I_{out}\right) V_{zener}$$

Again, for worst-case design, you would use V_{in} (max), R_{min}, and I_{out} (min).

EXERCISE 2.3

Design a +10 volt regulated supply for load currents from 0 to 100mA; the input voltage is +20 to +25 volts. Allow at least 10mA zener current under all (worst-case) conditions. What power rating must the zener have?

This simple zener-regulated supply is sometimes used for noncritical circuits, or circuits using little supply current. However, it has limited usefulness, for several reasons:

1. V_{out} isn't adjustable, or settable to a precise value.

2. Zener diodes give only moderate ripple rejection and regulation against changes of input or load, owing to their finite dynamic impedance.

3. For widely varying load currents a high-power zener is often necessary to handle the dissipation at low load current.

By using an emitter follower to isolate the zener, you get the improved circuit shown in Figure 2.12. Now the situation is much better. Zener current can be made relatively independent of load current, since the transistor base current is small, and far lower zener power dissipation is possible (reduced by as much as $1/h_{FE}$). The collector resistor R_C can be added to protect the transistor from momentary output short circuits by limiting the current, even though it is not essential to the emitter follower function. Choose R_C so that the voltage drop across it is less than the drop across R for the highest normal load current.

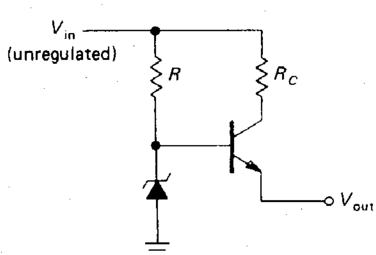

Figure 2.12. Zener regulator with follower, for increased output current. R_C protects the transistor by limiting maximum output current.

EXERCISE 2.4

Design a +10 volt supply with the same specifications as in Exercise 2.3. Use a zener and emitter follower. Calculate worst-case dissipation in transistor and zener. What is the percentage change in zener current from the no-load condition to full load? Compare with your previous circuit.

A nice variation of this circuit aims to eliminate the effect of ripple current (through R) on the zener voltage by supplying the zener current from a current source, which is the subject of Section 2.06. An alternative method uses a low-pass filter in the zener bias circuit (Fig. 2.13). R is chosen to provide sufficient zener current. Then C is chosen large enough so that $RC \gg 1/f_{ripple}$. (In a variation of this circuit, the upper resistor is replaced by a diode.)

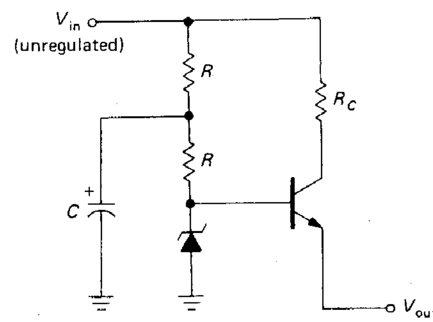

Figure 2.13. Reducing ripple in the zener regulator.

Later you will see better voltage regulators, ones in which you can vary the output easily and continuously, using feedback. They are also better voltage sources, with output impedances measured in milliohms, temperature coefficients of a few parts per million per degree centigrade, etc.

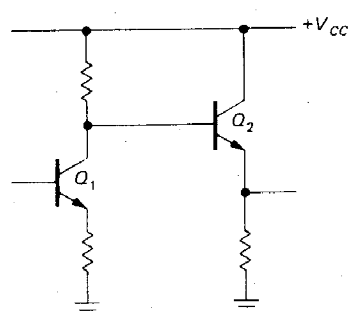

Figure 2.14

2.05 Emitter follower biasing

When an emitter follower is driven from a preceding stage in a circuit, it is usually OK to connect its base directly to the previous stage's output, as shown in Figure 2.14.

Because the signal on Q_1's collector is always within the range of the power supplies, Q_2's base will be between V_{CC} and ground, and therefore Q_2 is in the active region (neither cut off nor saturated), with its base-emitter diode in conduction and its collector at least a few tenths of a volt more positive than its emitter. Sometimes, though, the input to a follower may not be so conveniently situated with respect to the supply voltages. A typical example is a capacitively coupled (or ac-coupled) signal from some external source (e.g., an audio signal input to a high-fidelity amplifier). In that case the signal's average voltage is zero, and direct coupling to an emitter follower will give an output like that in Figure 2.15.

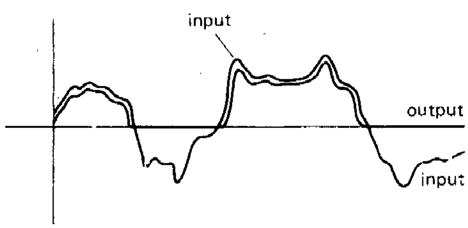

Figure 2.15. A transistor amplifier powered from a single positive supply cannot generate negative voltage swings at the transistor output terminal.

It is necessary to *bias* the follower (in fact, any transistor amplifier) so that collector current flows during the entire signal swing. In this case a voltage divider is the simplest way (Fig. 2.16). R_1 and R_2 are chosen to put the base halfway between ground and V_{CC} with no input signal, i.e., R_1 and R_2 are approximately equal. The process of selecting the operating voltages in a circuit, in the absence of applied signals, is known as setting the *quiescent point.* In this case, as in most cases, the quiescent point is chosen to allow maximum symmetrical signal swing of the output waveform without *clipping* (flattening of the top or bottom of the waveform). What values should R_1 and R_2 have? Applying our general principle (Section 1.05), we make the impedance of the dc bias source (the impedance looking into the voltage divider) small compared with the load it drives (the dc impedance looking into the base of the follower). In this case,

$$R_1 \| R_2 \ll h_{FE} R_E$$

This is approximately equivalent to saying that the current flowing in the voltage divider should be large compared with the current drawn by the base.

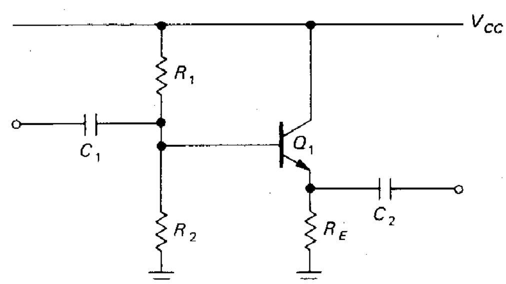

Figure 2.16. An ac-coupled emitter follower. Note base bias voltage divider.

Emitter follower design example

As an actual design example, let's make an emitter follower for audio signals (20Hz to 20kHz). V_{CC} is +15 volts, and quiescent current is to be 1mA.

Step 1. Choose V_E. For the largest possible symmetrical swing without clipping, $V_E = 0.5V_{CC}$, or +7.5 volts.

Step 2. Choose R_E. For a quiescent current of 1mA, $R_E = 7.5\text{k}$.

Step 3. Choose R_1 and R_2. V_B is V_E+ 0.6, or 8.1 volts. This determines the ratio of R_1 to R_2 as 1:1.17. The preceding loading criterion requires that the parallel resistance of R_1 and R_2 be about 75k or less (one-tenth of 7.5k times h_{FE}).

Suitable standard values are $R_1 = 130\text{k}$, $R_2 = 150\text{k}$.

Step 4. Choose C_1. C_1 forms a high-pass filter with the impedance it sees as a load, namely the impedance looking into the base in parallel with the impedance looking into the base voltage divider. If we assume that the load this circuit will drive is large compared with the emitter resistor, then the impedance looking into the base is $h_{FE}R_E$, about 750k. The divider looks like 70k. So the capacitor sees a load of about 63k, and it should have a value of at least 0.15μF so that the 3dB point will be below the lowest frequency of interest, 20Hz.

Step 5. Choose C_2. C_2 forms a high-pass filter in combination with the load impedance, which is unknown. However, it is safe to assume that the load impedance won't be smaller than R_E, which gives a value for C_2 of at least 1.0μF to put the 3dB point below 20Hz. Because there are now two cascaded high-pass filter sections, the capacitor values should be increased somewhat to prevent large attenuation (reduction of signal amplitude, in this case 6dB) at the lowest frequency of interest. $C_1 = 0.5\mu$F and $C_2 = 3.3\mu$F might be good choices.

Followers with split supplies

Because signals often are "near ground," it is convenient to use symmetrical positive and negative supplies. This simplifies biasing and eliminates coupling capacitors (Fig. 2.17).

Warning: You must always provide a dc path for base bias current, even if it goes only to ground. In the preceding circuit it is assumed that the signal source has a dc path to ground. If not (e.g., if the signal is capacitively coupled), you must provide a resistor to ground (Fig. 2.18). R_B could be about one-tenth of $h_{FE}R_E$, as before.

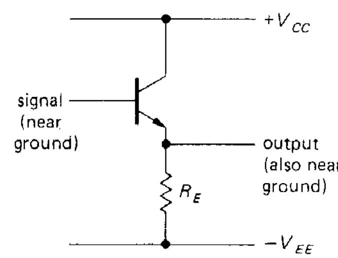

Figure 2.17. A dc-coupled emitter follower with split supply.

EXERCISE 2.5

Design an emitter follower with ±15 volt supplies to operate over the audio range (20Hz–20kHz). Use 5mA quiescent current and capacitive input coupling.

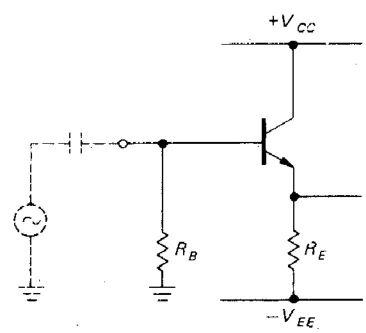

Figure 2.18

Bad biasing

Unfortunately, you sometimes see circuits like the disaster shown in Figure 2.19. R_B was chosen by assuming a particular value for h_{FE} (100), estimating the base current, and then hoping for a 7 volt drop across R_B. This is a bad design; h_{FE} is not a good parameter and will vary considerably. By using voltage biasing with a stiff voltage divider, as in the detailed example presented earlier, the quiescent point is insensitive to variations in transistor beta. For instance, in the previous design example the emitter voltage will increase by only 0.35 volt (5%) for a transistor with $h_{FE} = 200$ instead of the nominal

$h_{FE} = 100$. As with this emitter follower example, it is just as easy to fall into this trap and design bad transistor circuits in the other transistor configurations (e.g., the common-emitter amplifier, which we will treat later in this chapter).

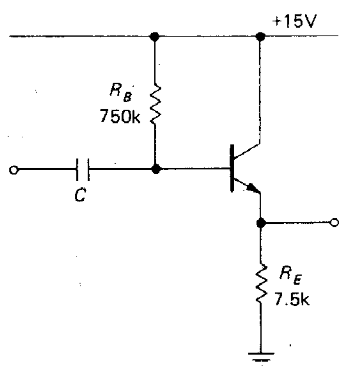

Figure 2.19. Don't do this!

2.06 Transistor current source

Current sources, although often neglected, are as important and as useful as voltage sources. They often provide an excellent way to bias transistors, and they are unequaled as "active loads" for super-gain amplifier stages and as emitter sources for differential amplifiers. Integrators, sawtooth generators, and ramp generators need current sources. They provide wide-voltage-range pull-ups within amplifier and regulator circuits. And, finally, there are applications in the outside world that require constant current sources, e.g., electrophoresis or electrochemistry.

Resistor plus voltage source

The simplest approximation to a current source is shown in Figure 2.20. As long as $R_{\text{load}} \ll R$ (in other words, $V_{\text{load}} \ll V$), the current is nearly constant and is approximately

$$I = V/R$$

The load doesn't have to be resistive. A capacitor will charge at a constant rate, as long as $V_{\text{capacitor}} \ll V$; this is just the first part of the exponential charging curve of an RC.

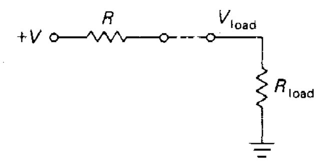

Figure 2.20

There are several drawbacks to a simple resistor current source. In order to make a good approximation to a current source, you must use large voltages, with lots of power dissipation in the resistor. In addition, the current isn't easily *programmable*, i.e., controllable over a large range via a voltage somewhere else in the circuit.

EXERCISE 2.6

If you want a current source constant to 1% over a load voltage range of 0 to +10 volts, how large a voltage source must you use in series with a single resistor?

EXERCISE 2.7

Suppose you want a 10mA current in the preceding problem. How much power is dissipated in the series resistor? How much gets to the load?

Transistor current source

Fortunately, it is possible to make a very good current source with a transistor (Fig. 2.21). It works like this : Applying V_B to the base, with $V_B > 0.6$ volt, ensures that the emitter is always conducting:

$$V_E = V_B - 0.6 \text{ volt}$$

So

$$I_E = V_E/R_E = (V_B - 0.6 \text{ volt})/R_E$$

But, since $I_E \approx I_C$ for large h_{FE},

$$I_C \approx (V_B - 0.6 \text{ volt})/R_E$$

independent of V_C, as long as the transistor is not saturated ($V_C > V_E + 0.2$ volt).

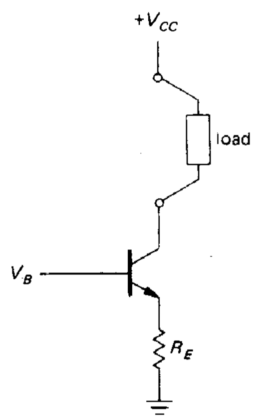

Figure 2.21. Transistor current source: basic concept.

Current-source biasing

The base voltage can be provided in a number of ways. A voltage divider is OK, as long as it is stiff enough. As before, the criterion is that its impedance should be much less than the dc impedance looking into the base ($h_{FE}R_E$). Or you can use a zener diode, biased from V_{CC}, or even a few forward-biased diodes in series from base to the corresponding emitter supply. Figure 2.22 shows some examples. In the last example (Fig. 2.22C), a *pnp* transistor *sources* current to a load returned to ground. The other examples (using *npn* transistors) should properly be called current *sinks*, but the usual practice is to call all of them current sources. ["Sink" and "source" simply refer to the direction of current flow: If a circuit *supplies* (positive) current to a point, it is a *source*, and vice versa.] In the first circuit, the voltage-divider impedance of ~1.3k is very stiff compared with the impedance looking into the base of about 100k (for $h_{FE} = 100$), so any changes in beta with collector voltage will not much affect the output current by causing the base voltage to change. In the other two circuits the biasing resistors are chosen to provide several milliamps to bring the diodes into conduction.

Compliance

A current source can provide constant current to the load only over some finite range of load voltage. To do otherwise would be equivalent to providing infinite power. The output voltage range over which a current source behaves well is called its output *compliance*. For the preceding transistor current sources, the compliance is set by the requirement that

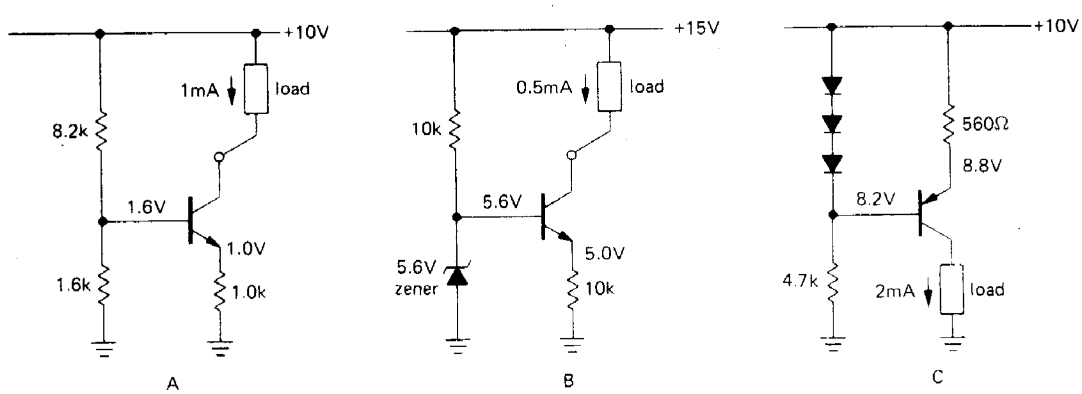

Figure 2.22. Transistor–current–source circuits, illustrating three methods of base biasing; *npn* transistors *sink* current, whereas *pnp* transistors *source* current. The circuit in C illustrates a load returned to ground.

the transistors stay in the active region. Thus in the first circuit the voltage at the collector can go down until the transistor is almost in saturation, perhaps +1.2 volts at the collector. The second circuit, with its higher emitter voltage, can sink current down to a collector voltage of about +5.2 volts.

In all cases the collector voltage can range from a value near saturation all the way up to the supply voltage. For example, the last circuit can source current to the load for any voltage between zero and about +8.6 volts across the load. In fact, the load might even contain batteries or power supplies of its own, carrying the collector beyond the supply voltage. That's OK, but you must watch out for transistor breakdown (V_{CE} must not exceed BV_{CEO}, the specified collector-emitter breakdown voltage) and also for excessive power dissipation (set by $I_C V_{CE}$). As you will see in Section 6.07, there is an additional safe-operating-area constraint on power transistors.

EXERCISE 2.8

You have +5 and +15 volt regulated supplies available in a circuit. Design a 5mA *npn* current source (sink) using the +5 volts on the base. What is the output compliance?

A current source doesn't have to have a fixed voltage at the base. By varying V_B you get a voltage-programmable current source. The input signal swing v_{in} (remember, lower-case symbols mean *variations*) must stay small enough so that the emitter voltage never drops to zero, if the output current is to reflect input voltage variations smoothly. The result will be a current source with variations in output current proportional to the variations in input voltage, $i_{\text{out}} = v_{\text{in}}/R_E$.

□ *Deficiencies of current sources*

To what extent does this kind of current source depart from the ideal? In other words, does the load current vary with voltage, i.e., have a finite ($R_{Th} < \infty$) Thévenin equivalent resistance, and if so why? There are two kinds of effects:

1. Both V_{BE} (Early effect) and h_{FE} vary slightly with collector-to-emitter voltage at a given collector current. The changes in V_{BE} produced by voltage swings across the load cause the output current to change, because the emitter voltage (and therefore the emitter current) changes, even with a fixed applied base voltage. Changes in h_{FE} produce small changes in output (collector) current for fixed emitter current, since $I_C = I_E - I_B$; in addition, there are small changes in applied base voltage produced by the variable loading of the nonzero bias source impedance as h_{FE} (and therefore the base current) changes. These effects are small. For instance, the current from the circuit in Figure 2.22A varied about 0.5% in actual measurements with a 2N3565 transistor. In particular, for load voltages varying from zero to 8 volts, the Early effect contributed 0.5%, and transistor heating effects contributed 0.2%. In addition, variations in h_{FE} contributed 0.05% (note the stiff divider). Thus these variations result in a less-than-perfect current source: The output current depends slightly on voltage and therefore has less than infinite impedance. Later you will see methods that get around this difficulty.

2. V_{BE} and also h_{FE} depend on temperature. This causes drifts in output current with changes in ambient temperature; in addition, the transistor junction temperature varies as the load voltage is changed (because of variation in transistor dissipation), resulting in departure from ideal current source behavior. The change of V_{BE} with ambient temperature can be compensated with a circuit like that shown in Figure 2.23, in which Q_2's base-emitter drop is compensated by the drop in emitter follower Q_1, with similar temperature dependence. R_3, incidentally, is a

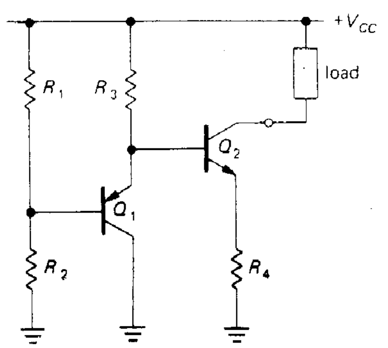

Figure 2.23. One method of temperature-compensating a current source.

pull-up resistor for Q_1, since Q_2's base sinks current, which Q_1 cannot source.

□ *Improving current-source performance*

In general, the effects of variability in V_{BE}, whether caused by temperature dependence (approximately -2mV/°C) or by dependence on V_{CE} (the Early effect, given roughly by $\Delta V_{BE} \approx -0.0001\,\Delta V_{CE}$), can be minimized by choosing the emitter voltage to be large enough (at least 1V, say) so that changes in V_{BE} of tens of millivolts will not result in large fractional changes in the voltage across the emitter resistor (remember that the *base* voltage is what is held constant by your circuit). For instance, choosing $V_E = 0.1$ volt (i.e., applying about 0.7V to the base) would cause 10% variations in output current for 10mV changes in V_{BE}, whereas the choice $V_E = 1.0$ volt would result in 1% current variations for the same V_{BE} changes. Don't get carried away, though. Remember that the lower limit of output compliance is set by the emitter voltage. Using a 5 volt emitter voltage for a current source running from a +10 volt supply limits the output compliance to slightly less than 5 volts (the collector can go from about $V_E +$ 0.2V to V_{CC}, i.e., from 5.2V to 10V).

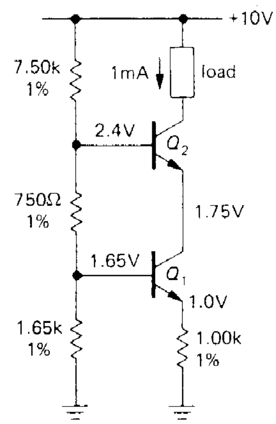

Figure 2.24. Cascode current source for improved current stability with load voltage variations.

Figure 2.24 shows a circuit modification that improves current-source performance significantly. Current source Q_1 functions as before, but with collector voltage held fixed by Q_2's emitter. The load sees the same current as before, since Q_2's collector and emitter currents are nearly equal (large h_{FE}). But with this circuit the V_{CE} of Q_1 doesn't change with load voltage, thus eliminating the small changes in V_{BE} from Early effect and dissipation-induced temperature changes. Measurements with 2N3565s gave 0.1% current variation for load voltages from 0 to 8 volts; to obtain performance of this accuracy it is important to use stable 1% resistors, as shown. (Incidentally, this circuit connection also finds use in high-frequency amplifiers, where it is known as the "cascode.") Later you will see current source techniques using op-amps and feedback that circumvent the problem of V_{BE} variation altogether.

The effects of variability of h_{FE} can be minimized by choosing transistors with large h_{FE}, so that the base current contribution to the emitter current is relatively small.

Figure 2.25 shows one last current source, whose output current doesn't depend on supply voltage. In this circuit, Q_1's V_{BE} across R_2 sets the output current, independent of V_{CC}:

$$I_{\text{out}} = V_{BE}/R_2$$

R_1 biases Q_2 and holds Q_1's collector at two diode drops below V_{CC}, eliminating Early effect as in the previous circuit. This circuit is not temperature-compensated; the voltage across R_2 decreases approximately 2.1mV/°C, causing the output current to decrease approximately 0.3%/°C.

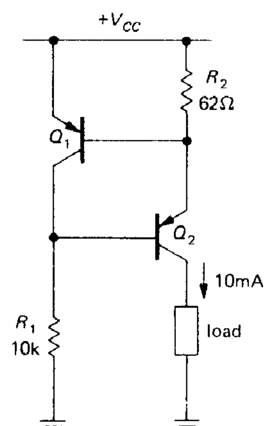

Figure 2.25. Transistor V_{BE}-referenced current source.

2.07 Common-emitter amplifier

Consider a current source with a resistor as load (Fig. 2.26). The collector voltage is

$$V_C = V_{CC} - I_C R_C$$

We could capacitively couple a signal to the base to cause the collector voltage to vary. Consider the example in Figure 2.27. C is chosen so that all frequencies of interest are passed by the high-pass filter it forms in combination with the parallel resistance of the base biasing resistors (the

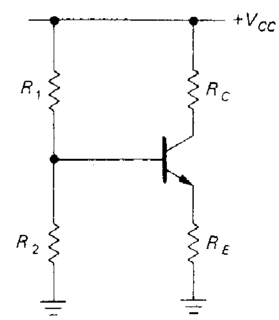

Figure 2.26

impedance looking into the base itself will usually be much larger because of the way the base resistors are chosen, and it can be ignored); that is,

$$C \geq \frac{1}{2\pi f(R_1 \| R_2)}$$

The quiescent collector current is 1.0mA because of the applied base bias and the 1.0k emitter resistor. That current puts the collector at +10 volts (+20V, minus 1.0mA through 10k). Now imagine an applied wiggle in base voltage v_B. The emitter follows with $v_E = v_B$, which causes a wiggle in emitter current

$$i_E = v_E/R_E = v_B/R_E$$

and nearly the same change in collector current (h_{fe} is large). So the initial wiggle in base voltage finally causes a collector voltage wiggle

$$v_C = -i_C R_C = -v_B(R_C/R_E)$$

Aha! It's a *voltage amplifier*, with a voltage amplification (or "gain") given by

$$\text{gain} = v_{\text{out}}/v_{\text{in}} = -R_C/R_E$$

In this case the gain is −10,000/1000, or −10. The minus sign means that a positive wiggle at the input gets turned into a negative wiggle (10 times as large) at the output. This is called a *common-emitter amplifier* with emitter degeneration.

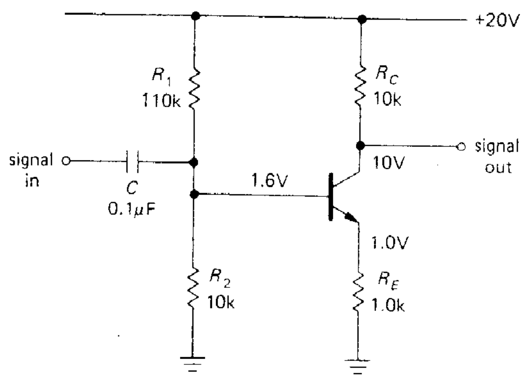

Figure 2.27. An ac common-emitter amplifier with emitter degeneration. Note that the output terminal is the collector, rather than the emitter.

Input and output impedance of the common-emitter amplifier

We can easily determine the input and output impedances of the amplifier. The input signal sees, in parallel, 110k, 10k, and the impedance looking into the base. The latter is about 100k (h_{fe} times R_E), so the input impedance (dominated by the 10k) is about 8k. The input coupling capacitor thus forms a high-pass filter, with the 3dB point at 200Hz. The signal driving the amplifier sees 0.1μF in series with 8k, which to signals of normal frequencies (well above the 3dB point) just looks like 8k.

The output impedance is 10k in parallel with the impedance looking into the collector. What is that? Well, remember that if you snip off the collector resistor, you're simply looking into a current source. The collector impedance is very large (measured in megohms), and so the output impedance is just the value of the collector resistor, 10k. It is worth remembering that the impedance looking into a transistor's collector is high, whereas the impedance looking into the emitter is low (as in the emitter follower). Although the output impedance of a common-emitter amplifier will be dominated by the collector load resistor, the output impedance of an emitter follower will not be dominated by the emitter load resistor, but rather by the impedance looking into the emitter.

2.08 Unity-gain phase splitter

Sometimes it is useful to generate a signal and its inverse, i.e., two signals 180° out of phase. That's easy to do – just use an emitter-degenerated amplifier with a gain of -1 (Fig. 2.28). The quiescent collector voltage is set to $0.75V_{CC}$, rather than the usual $0.5V_{CC}$, in order to achieve the same result – maximum symmetrical output swing without clipping at either output. The collector can swing from $0.5V_{CC}$ to V_{CC}, whereas the emitter can swing from ground to $0.5V_{CC}$.

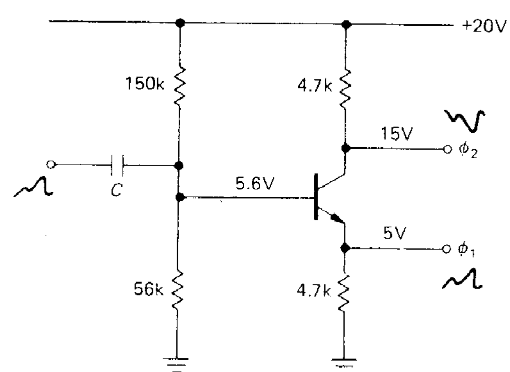

Figure 2.28. Unity-gain phase splitter.

Note that the phase-splitter outputs must be loaded with equal (or very high) impedances at the two outputs in order to maintain gain symmetry.

Phase shifter

A nice use of the phase splitter is shown in Figure 2.29. This circuit gives (for a sine wave input) an output sine wave of adjustable phase (from zero to 180°), but with constant amplitude. It can be best understood with a phasor diagram of voltages (see Chapter 1); representing the input signal by a unit vector along the real axis, the signals look as shown in Figure 2.30.

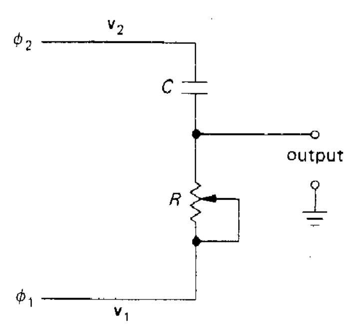

Figure 2.29. Constant-amplitude phase shifter.

Signal vectors v_R and v_C must be at right angles, and they must add to form a vector of constant length along the real axis. There is a theorem from geometry that says that the locus of such points is a circle. So the resultant vector (the output voltage) always has unit length, i.e., the same amplitude as the input, and its phase can vary from nearly zero to nearly 180° relative to the input wave as R is varied from nearly zero to a value much larger than Z_C at the operating frequency. However, note that the phase shift also depends on the frequency of the input signal for a given setting of the potentiometer R. It is worth noting that a simple RC high-pass (or low-pass) network could also be used as an adjustable phase shifter. However, its output amplitude would vary over an enormous range as the phase shift was adjusted.

An additional concern here is the ability of the phase-splitter circuit to drive the RC phase shifter as a load. Ideally, the load should present an impedance that is large compared with the collector and emitter resistors. As a result, this circuit is of limited utility where a wide range of phase shifts is required. You will see improved phase-splitter techniques in the next chapter.

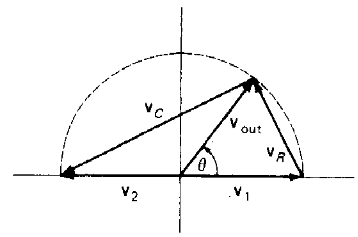

Figure 2.30. Phasor diagram for phase shifter.

2.09 Transconductance

In the preceding section we figured out the operation of the emitter-degenerated amplifier by (a) imagining an applied base voltage swing and seeing that the emitter voltage had the same swing, then (b) calculating the emitter current swing; then, ignoring the small base current contribution, we got the collector current swing and thus (c) the collector voltage swing. The voltage gain was then simply the ratio of collector (output) voltage swing to base (input) voltage swing.

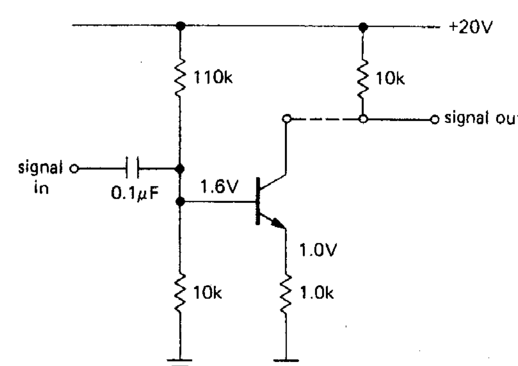

Figure 2.31. The common-emitter amplifier is a transconductance stage driving a (resistive) load.

There's another way to think about this kind of amplifier. Imagine breaking it apart, as in Figure 2.31. The first part is a voltage-controlled current source, with quiescent current of 1.0mA and gain of -1mA/V. Gain means the ratio output/input; in this case the gain has units of current/voltage, or 1/resistance. The inverse of resistance is called *conductance* (the inverse of reactance is *susceptance*, and the inverse of impedance is *admittance*) and has a special unit, the *siemens*, which used to be called the *mho* (ohm spelled backward). An amplifier whose gain has units of conductance is called a *transconductance* amplifier; the ratio $I_{\text{out}}/V_{\text{in}}$ is called the transconductance, g_m.

Think of the first part of the circuit as a transconductance amplifier, i.e., a voltage-to-current amplifier with transconductance g_m (gain) of 1mA/V (1000μS, or 1mS, which is just $1/R_E$). The second part of the circuit is the load resistor, an "amplifier" that converts current to voltage. This resistor could be called a *transresistance* amplifier, and its gain (r_m) has units of voltage/current, or resistance. In this case its quiescent voltage is V_{CC}, and its gain (transresistance) is 10kV/A (10kΩ), which is just R_C. Connecting the two parts together gives you a voltage amplifier. You get the overall gain by multiplying the two gains. In this case $G = g_m R_C = R_C/R_E$, or -10, a unitless number equal to the ratio (output voltage)/(input voltage).

This is a useful way to think about an amplifier, because you can analyze performance of the sections independently. For example, you can analyze the transconductance part of the amplifier by evaluating g_m for different circuit configurations or even different devices, such as field-effect transistors (FETs). Then you can analyze the transresistance (or load) part by considering gain versus voltage swing trade-offs. If you are interested in the overall voltage gain, it is given by $G_V = g_m r_m$, where r_m is the transresistance of the load. Ultimately the substitution of an active load (current source), with its extremely high transresistance, can yield one-stage voltage gains of 10,000 or more. The *cascode* configuration, which we will discuss later, is another example easily understood with this approach.

In Chapter 4, which deals with operational amplifiers, you will see further examples of amplifiers with voltages or currents as inputs or outputs; voltage amplifiers (voltage to voltage), current amplifiers (current to current), and transresistance amplifiers (current to voltage).

Turning up the gain: limitations of the simple model

The voltage gain of the emitter-degenerated amplifier is $-R_C/R_E$, according to our model. What happens as R_E is reduced toward zero? The equation predicts that the gain will rise without limit. But if we made actual measurements of the preceding circuit, keeping the quiescent current constant at 1mA, we would find that the gain would level off at about 400 when R_E is zero, i.e., with the emitter grounded. We would also find that the amplifier would become significantly nonlinear (the output would not be a faithful replica of the input), the input impedance would become small and nonlinear, and the biasing would become critical and unstable with temperature. Clearly our transistor model is incomplete and needs to be modified in order to handle this circuit situation, as well as others we will talk about shortly. Our fixed-up model, which we will call the transconductance model, will be accurate enough for the remainder of the book.