8

Transistor Switches

Now we begin the study of transistor circuits. Transistors have three terminals. Usually one of the terminals is the input, another is the output, and the third is a common connection that is shared between the input and the output. Transistor circuits can increase the power of a signal. For this they require an additional DC power source. Circuits that increase power are called *active* circuits. By comparison, a *passive* circuit has loss. The filters we covered in the earlier chapters are examples of passive circuits. We will study several different active circuits. An amplifier increases the power of a signal without changing the frequency. In an *oscillator*, an output sine wave is generated without any input signal. Transistors can also be used in passive circuits. In Problem 5, we saw that a transistor could act as a fast switch, with either a low resistance between the output terminals or a high resistance, depending on the input voltage. We will also use a transistor as a variable attenuator to control the signal level.

Manufacturers can combine many transistors on one chip of silicon. These circuits are called *integrated circuits*, or ICs. Many thousands of different integrated circuits are available. One common type of IC includes several amplifiers cascaded one after another, so that the output signal is much larger than the input. These circuits are called *op amps*, short for *operational amplifiers*. Our audio amplifier is a specialized kind of op amp. We also use integrated-circuit mixers to shift the frequencies in our transceiver. In addition, there is a *regulator* IC that provides a stable DC supply voltage that is close to 8 V over a wide range of currents.

A key component within transistors is a silicon diode (Figure 8.1a). The two ends of a diode are traditionally labeled *p* and *n*. The *p* region is the anode, and the *n* region is the cathode. In the *n* region, current is carried primarily by electrons. Electrons have a negative charge (hence the label *n*). The *p* region is more difficult to describe. Physically, the current results from places in the silicon crystal where electrons are missing. These vacancies are called *holes*. Holes respond to an electric field just like electrons, except they act as if they have a positive charge rather than a negative charge. In addition, holes are usually slower. Whether a material is *p* type or *n* type depends on the kind of impurity introduced into the silicon during fabrication. Elements from the third column of the periodic table, such as boron or aluminum, make p-type material, whereas elements from column five, such as phosphorus or arsenic, produce material of n type. A diode is formed by a junction of p-type and n-type materials.

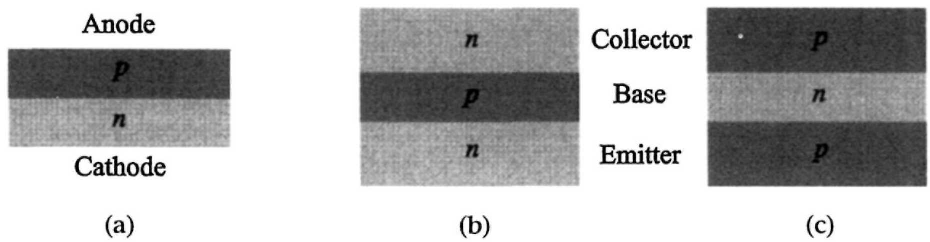


Figure 8.1. p and n regions in a pn diode (a). Layers for an npn transistor (b) and a pnp transistor (c).

There are two major families of transistors, *bipolar-junction* transistors, or BJTs, and *field-effect* transistors, or FETs. There are many different kinds of each transistor, and I defer to a book on solid-state devices for the details. We will think of a bipolar transistor as being controlled by an input current and a field-effect transistor as being controlled by an input voltage. Both kinds of transistors are found throughout electronics. We begin with BJTs.

8.1 Bipolar Transistors

A bipolar transistor is a three-layer sandwich and is called either *npn* (Figure 8.1b) or *pnp* (Figure 8.1c), depending on how the layers are arranged. The center layer is called the *base*, and the outer layers are the *collector* and *emitter*. In some respects, a transistor is a back-to-back connection of diodes. We can test a transistor like we test a diode by putting an ohmmeter across the terminals. We will find a diode between the base and the collector and between the base and the emitter. In practice, this is a good way to see if there is a problem with a transistor. If we connect the ohmmeter terminals to the collector and emitter, it will show an open circuit, because the diodes are connected back-to-back. This means that whichever way the voltage is applied, one of the diodes will be on, and one will be off, so that the current is blocked.

To an ohmmeter, there is no difference between a transistor and a pair of diodes connected back-to-back. But things become quite different if we have a current between the base and the emitter and a voltage applied to the collector. We

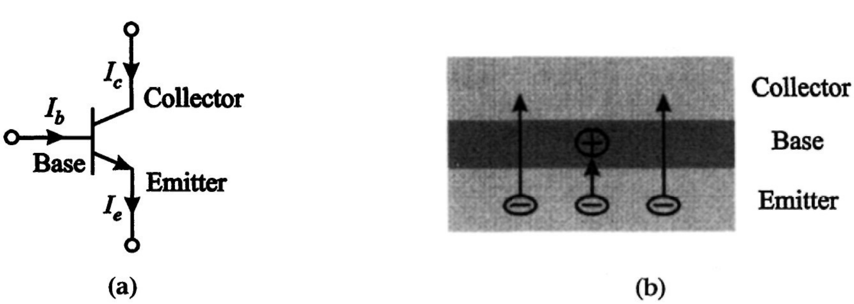


Figure 8.2. Schematic symbol and current directions (a) and electron flow (b) for an npn transistor.

start with npn transistors, which are more common than pnp transistors because they are faster, reflecting the fact that electrons generally move faster than holes. Figure 8.2a shows the schematic symbol for an npn transistor. The current flow in the base–emitter diode comes primarily from the flow of electrons from the emitter to the base (Figure 8.2b). Ordinarily, when current flows in a diode, an electron that crosses from the n region to the p region would meet with a hole and fill the vacancy. This is called *recombination*. However, in a transistor, the base region is made quite thin, a micron or less, so that the electron usually continues on to the collector, attracted by the voltage there. If the electron gets through the base without meeting up with a hole then it contributes to the collector current rather than to the base current.

The proportion of electrons from the emitter that make it to the collector is called the *collection efficiency*, and it is traditionally written as α. We write

$$I_c = \alpha I_e, \tag{8.1}$$

where I_c is the collector current and I_e is the emitter current. A typical value for α is 0.99, indicating that almost all of the electrons from the emitter end up in the collector rather than in the base. We can use Kirchhoff's current law to find the base current I_b by subtraction as

$$I_b = I_e - I_c = (1 - \alpha)I_e. \tag{8.2}$$

If α is close to one, then the base current is much smaller than the collector current. The ratio of collector current to base current is called the *current gain*, and we write it as β:

$$\beta = I_c/I_b. \tag{8.3}$$

We can divide Equation 8.1 by Equation 8.2 to get

$$\beta = \frac{\alpha}{1-\alpha}. \tag{8.4}$$

If $\alpha = 0.99$, then β is about 100. In manufacturers' data sheets, you will usually see h_{FE} instead of β. The letter "h," short for *hybrid*, refers to a particular equivalent circuit model. So far, we have been thinking of the base current as a kind of leakage. However, we can turn this picture around and think of the base current as controlling the collector current. If we change the base current, the collector current follows. Because β is a big number, the collector current changes by a much larger amount. This is the basis for an amplifier with the base as the input and the collector as the output.

8.2 Transistor Models

We give an equivalent circuit for the transistor in Figure 8.3a. This circuit is not complete because only one diode is shown. The diamond denotes a current

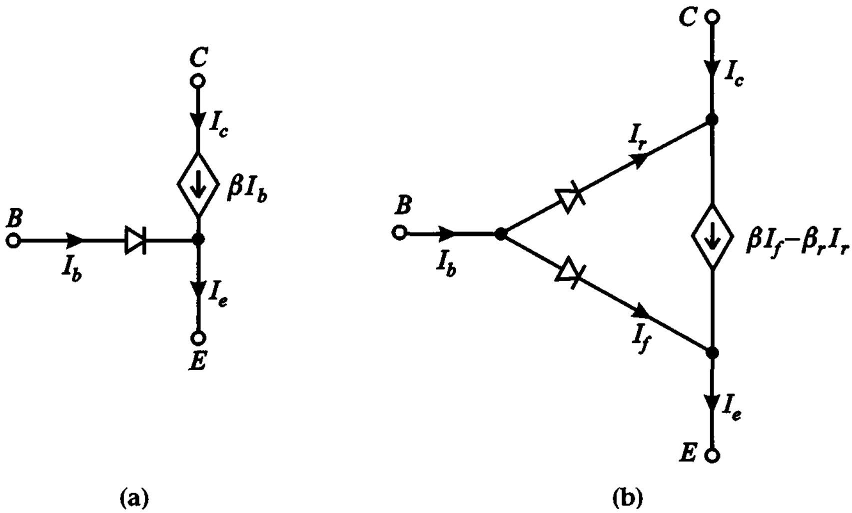


Figure 8.3. Transistor model with one diode (a), and the complete model with two diodes (b).

source βI_b. It is different from the current sources we have considered so far because it depends on another current in the circuit. It is called a *dependent* current source to contrast it with the previous *independent* sources. We will use diamonds for dependent sources and circles for independent sources.

We have considered the action of the base–emitter diode, but a similar picture holds for the base–collector diode. For this diode, we can define a reverse current gain β_r. Typically transistors are optimized for a large value of the normal current gain, and for this reason, β_r is much smaller, 10 or less. Figure 8.3b adds this effect to the circuit. Now I_b has two components, the forward base–emitter diode current I_f and the reverse base–collector diode I_r. The current generator has an additional component $-\beta_r I_r$.

We can use this model to understand how the collector current I_c behaves when we vary the collector–emitter voltage V_{ce} (Figure 8.4a). We will assume a constant

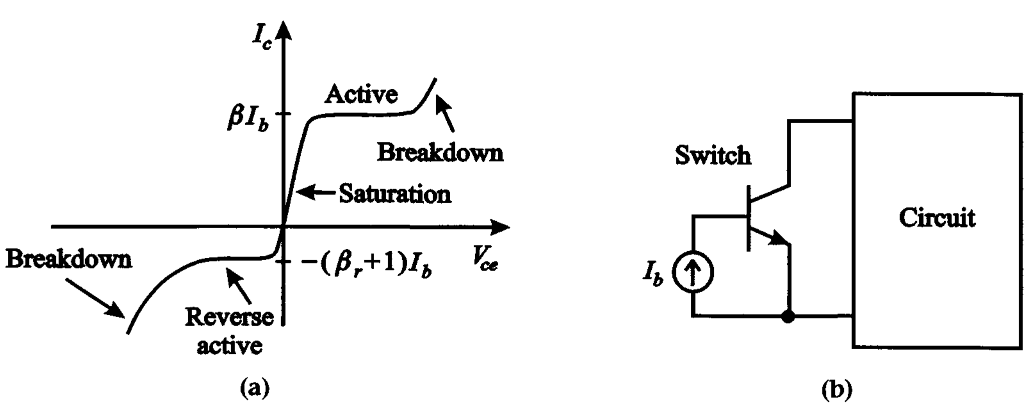


Figure 8.4. Collector current I_c for an npn transistor as a function of the collector–emitter voltage V_{ce}, with I_b constant (a). Transistor switch (b).

Table 8.1. Regions of Operation for an npn Transistor.

Region	V_{be}	V_{bc}	V_{ce}	I_c
Active	V_f	$<V_f$	$>V_s$	βI_b
Reverse active	$<V_f$	V_f	$<-V_s$	$-(\beta_r+1)I_b$
On (saturated)	V_f	V_f	$>-V_s$, $<V_s$	$>-(\beta_r+1)I_b$, $<\beta I_b$
Off	$<V_f$	$<V_f$	Any	0

Note: V_f is the forward voltage of the base–emitter and base–collector diodes. This is about 100 mV larger than in ordinary diodes. It is typically 0.7 to 0.8 V. V_s is the boundary between the saturation and active regions, typically 0.1 to 0.2 V.

base current I_b. When V_{ce} is positive, the base–emitter diode is conducting and the base–collector diode is off. The collector current is given by

$$I_c = \beta I_b. \tag{8.5}$$

We say the transistor is *active*. The active region is used for amplifiers. When V_{ce} is negative, the base–collector diode is on, and the base–emitter diode is off. This is the *reverse active region*. We can write the collector current as

$$I_c = -(\beta_r + 1)I_b. \tag{8.6}$$

The collector voltage that can be applied is limited by the reverse breakdown voltage of the diodes. In particular, one needs to be careful operating in the reverse active region, because breakdown usually occurs at only a few volts.

An interesting thing happens when V_{ce} is small, a few tenths of a volt or less. Current flows through both diodes at the same time. Small changes in the voltage shift the balance of current in the two diodes, and this causes large changes in the collector current. People call this the *saturation region* (Figure 8.4a), or they say the transistor is *on*. Saying that the collector current changes rapidly is another way of saying that the resistance is small. How small depends on the base current. The larger the base current, the steeper the slope in the saturation region, and the smaller the saturation resistance. Table 8.1 summarizes the regions of operation for an npn transistor.

8.3 Transistor Switches

We can use our model to understand how a transistor switch works. In Figure 8.4b, we connect a transistor to a circuit, and apply a control current I_b to the base. If the base current is zero, there is no collector current. We say the switch is *off*, and the transistor is effectively disconnected from the circuit. Now if we apply a base current, the resistance between the collector and emitter will be small, provided that the transistor stays in the saturation region. We say the switch is *on*. The transistor effectively shorts out the circuit.

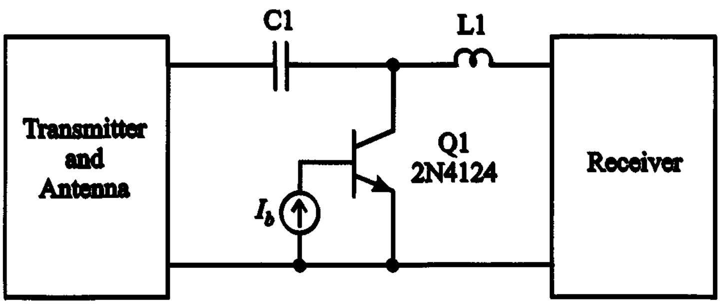


Figure 8.5. Operation of the Receiver Switch (Problem 19).

In the NorCal 40A, the Receiver Switch shorts out the RF Filter when we are transmitting (Figure 8.5). The transmitter and the antenna are connected to the receiver by a series resonant circuit. The transistor Q1 is the switch. In reception, the switch is off. Signals at the resonant frequency pass directly through from the antenna to the receiver. The switch is turned on when transmitting. This gives the switch a low resistance and blocks signals from the transmitter to protect the receiver.

To choose the base current to operate this switch, we need to know the relation between the saturation resistance and the base current. Figure 8.6 shows the saturation resistance for the transistor used in the Receiver Switch. Making this measurement is trickier than it might appear, because the collector current is

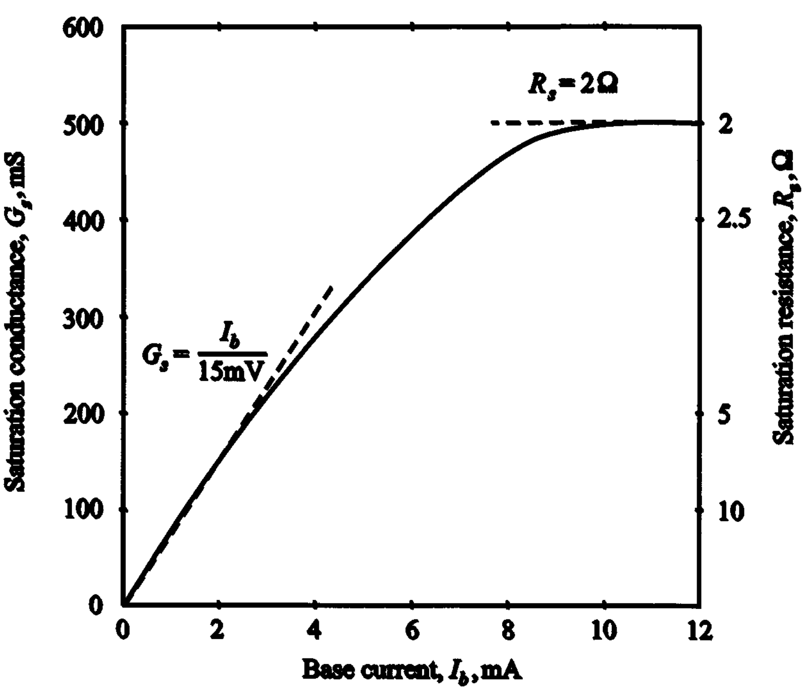


Figure 8.6. Measured saturation conductance G_s versus base current I_b for the 2N4124 transistor. The equivalent saturation resistance R_s is shown on the right.

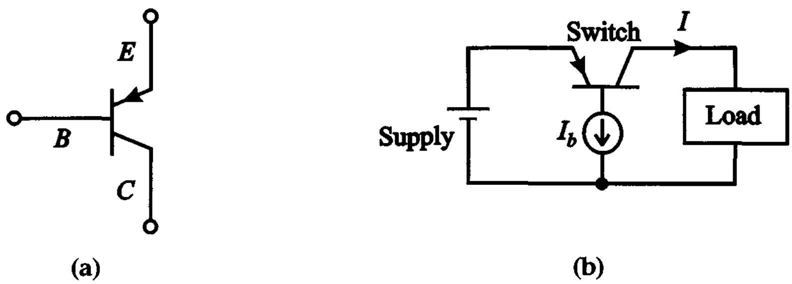


Figure 8.7. Circuit symbol for a pnp transistor (a), and a pnp switch for making a connection to a supply (b).

curved in the saturation region. The effective resistance depends on the value of the switch voltage and whether it is positive or negative. In the transmitter, the voltage is sinusoidal: It is positive half the time and negative the rest of the time. The plot shows the slope of the collector current versus collector voltage, evaluated at zero collector current. People call this a *small-signal conductance*, because it is appropriate for small voltages. There are two distinct regions in the plot. At small base currents, the saturation conductance G_s is proportional to I_b. We can write this relation as

$$G_s = I_b/15\,\mathrm{mV}. \tag{8.7}$$

For large base currents, the conductance approaches a maximum value of 500 mS. Expressed as a resistance, we write

$$R_s = 2\,\Omega. \tag{8.8}$$

This is a *parasitic resistance* that is associated with the silicon itself. The Receiver Switch in Problem 19 is designed with a base current of about 5 mA, which gives a saturation resistance of about 3 Ω. Because this is much smaller than the reactances of C1 and L1, the transistor shorts out the signal.

An npn switch is convenient for providing a short to ground. A pnp transistor can connect a load to a voltage source. Figure 8.7a shows the circuit symbol for a pnp transistor. The arrow points *in*, in contrast to the npn symbol, where the arrow points out. For the pnp transistor, the models are identical except that the diodes and current generators change directions. It is a good idea to sketch these models to make sure that you understand them. Figure 8.7b shows a pnp switch for controlling the voltage applied to a load. A voltage source is connected to the emitter and a load is attached to the collector. The switch is controlled by base current. When the base current is zero, the transistor is off, and the load current and voltage are zero. We apply a base current to turn the transistor on, and to supply current to the load. In our transceiver, a pnp switch controls current to the transmitter circuits. This is the Transmitter Switch, and we build it in Problem 20.

Choosing the base current for this switch is simpler than for the previous one. The collector current does not change sign, and we can use the ordinary current gain β, provided we include a safety factor to make sure that the transistor saturates. We want the transistor to be saturated, so that the voltage drop is small and most of the supply voltage appears across the load. In the NorCal 40A, the Transmitter Switch load current is 7 mA. We need to look at the manufacturer's specification for the minimum current gain. The transistor in this switch is the 2N3906, and its data sheet in Appendix D specifies that the minimum value of h_{FE} (β) for a collector current of 10 mA is 100. Now we can calculate the required base current as

$$I_b = 2I/100 = 140\,\mu\text{A}, \tag{8.9}$$

where I is the load current and 2 is a safety factor. As a check, we can consult Figure 14 in Motorola's data sheet, which shows that the voltage drop between the collector and emitter should be about 200 mV. This would only knock the 8-V supply down to 7.8 V, which is acceptable. The figure also shows that this drop can be reduced by increasing the base current.

FURTHER READING

It is a good idea to study the chapters on transistor circuits in the *The Art of Electronics*, by Horowitz and Hill, published by Cambridge University Press. Horowitz and Hill have an excellent chapter on regulators, which we do not cover. *Device Electronics for Integrated Circuits*, by Muller and Kamins, published by Wiley, is a good book for information on transistors themselves. For an exhaustive survey at an advanced level on the behavior of different kinds of transistors, see *Physics of Semiconductor Devices*, by S. M. Sze, published by Wiley.

PROBLEM 19 – RECEIVER SWITCH

Transmitters produce much more power than receivers can handle. The NorCal 40A has an output power of about 2 W, which would destroy the Receive Mixer. The Receiver Switch (Figure 8.5) keeps the transmitter power out of the receiver. The receiver could be switched out by hand, but it is more reliable to have a transistor do this automatically. Figure 8.8 shows the detailed circuit. The transistor Q1 has its collector attached

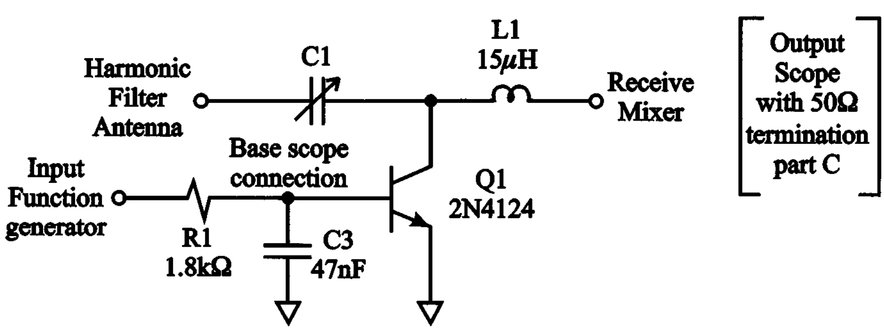


Figure 8.8. Details of the Receiver Switch and its connections. The triangles denote ground connections.

between the capacitor C1 and the inductor L1, and the emitter goes to ground. The base is connected to an RC delay circuit (R1 and C3). When transmitting, 8 V is applied to the delay circuit. This connection is called 8 V TX in the transceiver schematic. (TX is an abbreviation from telegraphy for "transmit.") The voltage produces a current in the base and turns the transistor on. This shorts out the filter, blocking the transmitter signal. In receiving, the input goes to zero volts. This stops base current and turns the transistor off, effectively removing it from the circuit. The filter can now operate normally.

The Receiver Switch prevents the receiver from being destroyed by the transmitter, but even more blocking is needed. The transmitter would still produce loud, annoying tones. In early radios, operators slipped off their headphones when they wanted to transmit. Modern transceivers have attenuators to reduce the signal before it gets to the audio amplifier. In the NorCal 40A, this job is handled by the AGC circuit.

Solder R1, C3, and Q1 into the circuit. For R1, leave enough room to connect the scope and the function generator. For the transistor, use the white outline to orient the package.

Set the function generator to give a 1-kHz square wave with an open-circuit high voltage of 8 V and low voltage of 0 V. This means that a 50-Ω function generator should be set to 4 Vpp with a DC offset of 2 V. These settings work because the open-circuit voltage is twice the indicated voltage. It is a good idea to check the voltage on an oscilloscope.

Connect the function generator to the input end of R1 and the scope to the other end (Figure 8.8). The base voltage should alternate between zero and the forward voltage of the base–emitter diode, with rising and falling transitions in between.

A. Consider the rising part of the base voltage waveform. Measure the initial slope. You will find it convenient to set the scope trigger for a positive slope, so that you can zoom in on the rising part of the waveform. Calculate approximately what the slope should be.

B. Now consider the falling part. It is best to set the slope trigger for a negative slope. Measure the time t_2 that it takes for the voltage to drop by a factor of two. At first the base–emitter diode will be on, and this causes the voltage to drop much faster than it does later. For this reason, you should make the measurement over a part of the curve where the voltage is below 0.6 V. Calculate what t_2 should be.

In the following sections, we need to measure small signals, and you will find it convenient to turn on the scope's low-pass filter, if one is available. This reduces high-frequency noise and makes the traces sharper. Attach the function generator to the Antenna jack J1 and the scope at the output as shown in Figure 8.8. Use a 50-Ω termination on the scope. The function generator should be set for a 1-Vpp, 7-MHz sine wave.

C. Now we measure the attenuation of the switch. First consider the signal with the Receiver Switch off. Adjust C1 for maximum scope voltage and record the voltage.

D. Now measure the voltage with the Receiver Switch on. You can turn on the switch by connecting a 12-V power supply to the input at R1. You should increase the function-generator amplitude setting to 10 Vpp, to make it easier to see the signal.

You will need to account for this voltage setting in the calculations. Measure the output voltage, and calculate the on–off rejection ratio R in dB from the expression

$$R = 20\log(V_{off}/V_{on}). \tag{8.10}$$

E. Find an approximate formula for the attenuation in terms of the saturation resistance R_s. The easiest way to do this is to think of the circuit as a pair of cascaded voltage dividers.

F. Now calculate the attenuation that we would expect. To start, find the base current from the voltage drop across R1. Then find R_s from Figure 8.6, and apply your attenuation formula.

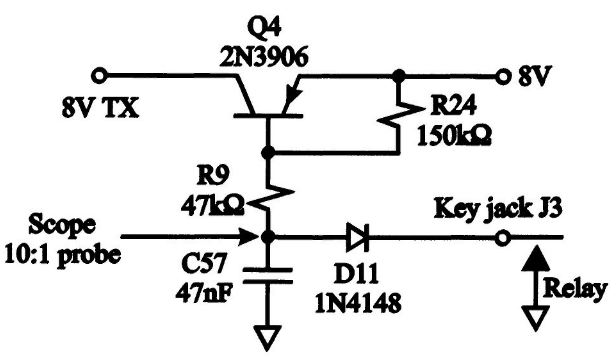
G. Simulate the attenuator with *Puff* from 0 MHz to 14 MHz.

H. The designer chose the 2N4124 for this switch because of its low off capacitance. This capacitance causes loss even when the transistor is off. Use your *Puff* simulation to find the loss. For this calculation use $C_{obo} = 3.5$ pF (Figure 1 in the data sheet). You will need to adjust C1 slightly for best transmission, just as in your measurements. For comparison, repeat the calculation for the 2N2222A transistor that you used in Problem 5, using $C_{cb} = 10$ pF (Figure 9 in the data sheet).

PROBLEM 20 – TRANSMITTER SWITCH

In the NorCal 40A, the Transmitter Switch uses a pnp transistor to provide the 8 V TX line that drives the Receiver Switch (Figure 8.9a). The Transmitter Switch also supplies the Transmit Mixer, the Buffer Amplifier, and the RIT circuit that shifts the frequency of the VFO during transmission. We will discuss these circuits later. The switch is controlled by a line to the Key jack J3. In operation, the jack is connected to a telegraph key. When the key is down, the line shorts to ground, and the radio transmits. When the key is up, the line opens, and the radio receives. We will use a relay in place of the key.

First consider what happens when the key is down. The capacitor C57 discharges through the diode D11, dropping the capacitor voltage down to the forward voltage



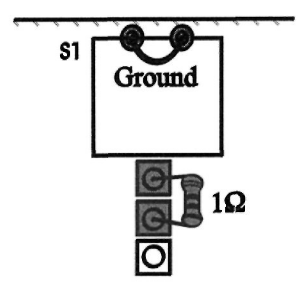


Figure 8.9. Connections to the Transmitter Switch (a), and installing a 1-Ω resistor in the S1 holes (b).

of the diode. A current flows in R9, and this pulls current through the base to turn on the transistor Q4, providing 8 V to the Receiver Switch. When the key is up, the base current stops flowing in the diode and begins to charge C57. As the capacitor charges, the base current drops. For a while nothing happens to the collector voltage, because the base current is still sufficient to turn the transistor on. However, eventually the base current drops enough to make the transistor *active*, where the collector current is given by $I_c = \beta I_b$. The collector voltage begins to fall now. Finally the capacitor voltage gets so high that the base–emitter diode turns off, and the collector current stops. The resistor R24 keeps the capacitor charging even after the base–emitter diode is off.

Install the parts in Figure 8.9: Q4, R24, R9, D11, C57, and J3. Leave a couple of millimeters of lead length on the diode so that you can attach a probe. Also, install the regulator circuit consisting of U5 (78L08), C42 (10 μF), and C43 (47 nF). Refer to the endpaper to see how these components fit into the circuit. The regulator provides a nearly ideal 8-V source for currents up to 100 mA. The capacitors prevent the regulators from oscillating and help filter out signals from other parts of the circuit and noise from the power-supply connection to J2. The large capacitor provides protection at low frequencies, and the small capacitor at high frequencies. We will not study regulators but detailed information about this regulator is given in Appendix D. You do need to be careful with the electrolytic capacitor C42. This type of capacitor can provide very large capacitances from 1 μF to 100 mF. However, they have limitations. The capacitance values are not precise. Typically the tolerance is only ±20%. The standard values are usually limited to multiples of 10, 22, 33, 47, and 68. They are also *polarized*, which means that they cannot take large negative voltages. You must get the wires in the right holes. You can recognize the positive lead because it is longer than the negative lead. The positive hole is marked on the circuit board with a plus sign. In addition, the minus lead is marked on the can with a minus sign. If you install the capacitor incorrectly, there will be a large current, and the capacitor will heat up. Often you can smell that something is wrong. Eventually there will be a small explosion, or a big explosion if it is a large capacitor, and material will pop out of the capacitor can.

In addition, we need to install the Supply jack J2 and the Schottky diode D7 (1N5817). Refer to the endpaper schematic for this. The diode prevents damage if the power supply is hooked up backwards. Install a 1-Ω resistor in the S1 holes as shown in Figure 8.9b. This is a good place to attach probes to measure the supply voltage and current. The third hole farthest from the edge of the board is not connected to the circuit; it should be left open. You should add a loop of wire between the two holes in the S1 outline on the edge of the board to provide a ground connection.

Attach a relay to the Key jack J3. A relay is a switch that is operated by an electromagnet. See Appendix A for more information. We use the Magnecraft W171DIP-7, which requires 5 V to close the relay. The switching time is 200 μs. It has a coil resistance of 500 Ω and includes a snubber diode.

Connect the function generator to the relay, and use a 20-Hz square wave with an amplitude of 5 Vpp. This gives an open-circuit positive voltage of 5 V, which is the proper voltage for the relay. This speed of 20 pulses per second is about as fast as the best operators can receive Morse code. A more typical speed would be two to three times slower. Slower speeds are not convenient to observe on an oscilloscope without storage because the trace fades. A 10:1 scope probe should be attached at the anode of D11 to monitor the voltage of the capacitor C57. We do need to use the 10:1 probe, for otherwise the scope resistance drains too much charge from the capacitor.

A. Sketch the voltage on C57, indicating the key-down time when the relay is closed and the key-up time when the relay is open. Measure the time that it takes the capacitor to charge halfway from its minimum voltage to its maximum voltage. Now calculate approximately what this time should be.

B. Calculate approximately the collector current I_c when Q4 is on. Assume that the base–emitter voltage of Q1 is 700 mV. You can neglect the saturation voltage of Q4. Calculate the base current I_b that is required to produce this collector current, assuming the manufacturer's minimum β value of 100.

C. For comparison, calculate I_b at key down, assuming a 700-mV drop in the base–emitter diode of Q4 and a 600-mV drop in D11.

D. Now sketch the collector voltage for Q4, showing where the transistor is saturated. What is the delay in going active? This delay is useful because it allows the Power Amplifier to shut down gradually over a time period of 1 to 2 ms. If a transmitter turns off too quickly, it causes an annoying clicking sound on nearby frequencies that interferes with other operators.

E. The time at which the transistor goes active is interesting because we can use it to infer β. Measure the voltage across R9 at this time, and use it to calculate I_b. Compare this with the collector current I_c that you calculated previously to find β.