Chapter 6

Real-World Component Characteristics

When is an inductor not an inductor? When it's a capacitor! This statement may seem odd, but it suggests the main message of this chapter. In the earlier chapter about **Electrical Fundamentals**, the basic components of electronic circuits were introduced. You saw that each has its own unique function to perform. For example, a capacitor stores energy in an electric field, a diode rectifies current and a battery provides voltage. All of these unique and different functions are necessary in order to build large circuits that perform useful tasks. The first part of this chapter, up to Low-Frequency Transistor Models, was written by Leonard Kay, K1NU.

As you may know from experience, these component pictures are *ideal*. That is, they are perfect mathematical pictures. An ideal component (or *element*) by definition behaves exactly like the mathematical equations that describe it, and *only* in that fashion. For example, an ideal capacitor passes a current that is equal to the capacitance C times the rate of change of the voltage across it. Period, end of sentence.

We call any other exhibited behavior either *nonideal*, *nonlinear* or *parasitic*. Nonideal behavior is a general term that covers any deviation from the theoretical picture. Ideal circuit elements are often linear: The graphs of their current versus voltage characteristics are straight lines when plotted on a suitable (Cartesian, rectangular) set of axes. We therefore call deviation from this behavior nonlinear: For example, as current through a resistor exceeds its power rating, the resistor heats up and its resistance changes. As a result, a graph of current versus voltage is no longer a straight line.

When a component begins to exhibit properties of a different component, as when a capacitor allows a dc current to pass through it, we call this behavior *parasitic*. Nonlinear and parasitic behavior are both examples of nonideal behavior.

Much to the bane of experimenters and design engineers, ideal components exist only in electronics textbooks and computer programs. Real components, the ones we use, only approximate ideal components (albeit very closely in most cases).

Real diodes store minuscule energy in electric fields (junction capacitance) and magnetic fields (lead inductance); *real* capacitors conduct some dc (modeled by a parallel resistance), and real battery voltage is not *perfectly* constant (it may even decrease nonlinearly during discharge).

Knowing to what extent and under what conditions real components cease to behave like their ideal counterparts, and what can be done to account for these behaviors, is the subject of component or circuit *modeling*. In this chapter, we will explore how and why the real components behave differently from ideal components, how we can account for those differences when analyzing circuits, how to select components to minimize, or exploit, nonideal behaviors and give a brief introduction to computer-aided circuit modeling.

Much of this chapter may seem intuitive. For example, you probably know that #20 hookup wire works just fine for wiring many experimental circuits. When testing a circuit after construction, you probably don't even think about the voltage drops across those pieces of wire (you inherently assume that they have zero resistance). But, connect that same #20 wire directly across a car battery—with no *other* series resistance—and suddenly, the resistance of the wire *does* matter, and it changes—as the wire heats, melts and breaks—from a very small value to infinity!

LUMPED VS DISTRIBUTED ELEMENTS

Most electronic circuits that we use everyday are inherently and math-ematically considered to be composed of *lumped elements*. That is, we assume each component acts at a single point in space, and the wires that connect these lumped elements are assumed to be *perfect conductors* (with zero resistance and insignificant length). This concept is illustrated in **Fig 6.1**. These assumptions are perfectly reasonable for many applications, but they have limits. Lumped element models break down when:

- Circuit impedance is so low that the small, but non-zero, resistance in the wires is important. (A significant portion of the circuit power may be lost to heat in the conductors.)
- Operating frequency (f_0) is high enough that the length of the connecting wires is a significant fraction (>0.1) of the wavelength. (Radiation from the conductors may be significant.)
- Transmission lines are used as conductors. (Their characteristic impedance is usually significant, and impedances connected to them are transformed as a function of the line length. See the **Transmission Lines** chapter for more information.)

Effects such as these are called *distributed*, and we talk of *distributed* elements or effects to contrast them to lumped elements.

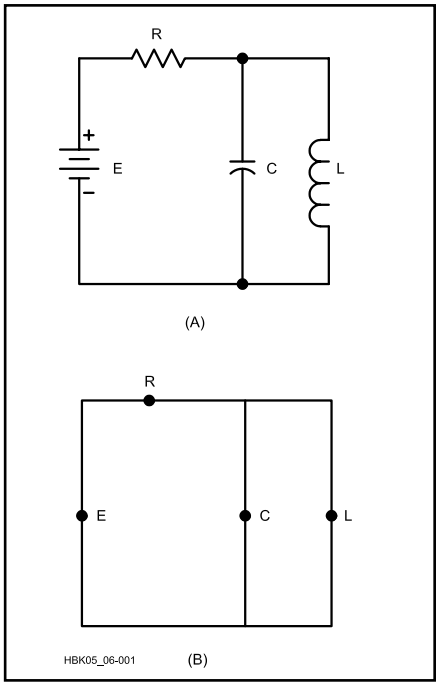


Fig 6.1—The lumped element concept. Ideally, the circuit at A is assumed to be as shown at B, where the components are isolated points connected by perfect conductors. Many components exhibit nonideal behavior when these assumptions no longer hold.

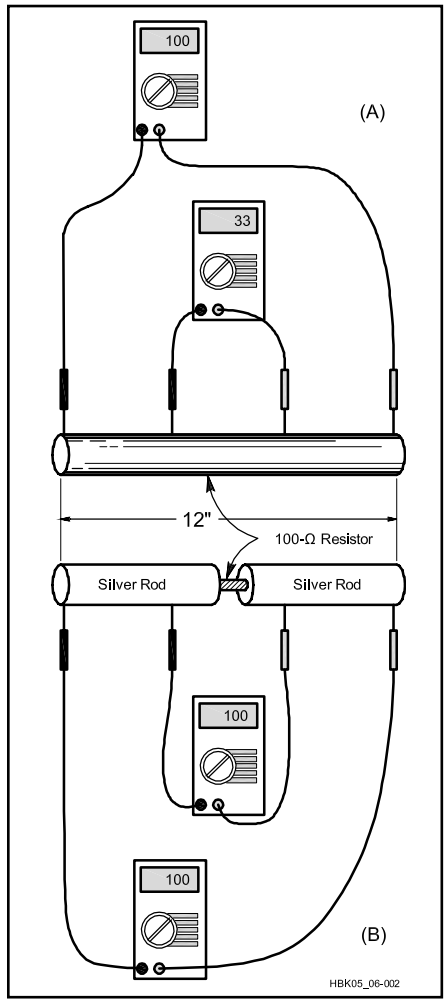


Fig 6.2—A, distributed and B, lumped resistances. See text for discussion.

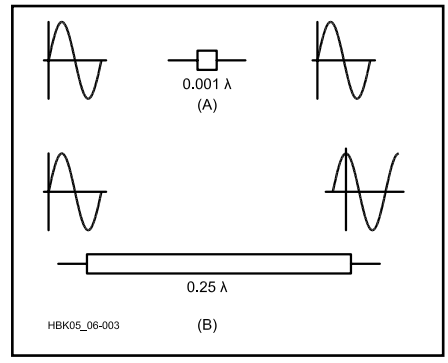


Fig 6.3—The effects of distributed resistance on the phase of a sinusoidal current. There is no phase delay between ends of a lumped element.

To illustrate the differences between lumped and distributed elements, consider the two resistors in **Fig 6.2**, which are both 12-inches long. The resistor at A is a uniform rod of carbon—a battery anode, for example. The second "resistor" B is made of two 6-inch pieces of silver rod (or other highly conductive material), with a small resistor soldered between them. Now imagine connecting the two probes of an ohmmeter to each of the two resistors, as in the figure. Starting with the probes at the far ends, as we slide the probes toward the center, the carbon rod will display a constantly decreasing resistance on the ohmmeter. This represents a *distributed* resistance. On the other hand, the ohmmeter connected to the other 12-inch "resistor" will display a constant resistance as long as one probe remains on each side of the small resistance. (Oh yes, as long as we neglect the resistance of the silver rods!) This represents a *lumped* resistance connected by perfect conductors.

Lumped elements also have the very desirable property that they introduce no phase shift resulting from propagation delay through the element. (Although combinations of lumped elements can produce phase shifts by virtue of their R, L and C properties.) Consider a lumped element that is carrying a sinusoidal current, as in **Fig 6.3A**. Since the element has negligible length, there is no phase difference in the current between the two sides of the element—*no matter how high the frequency*—precisely *because* the element length is negligible. If the physical length of the element were long, say 0.25 λ as shown in Fig 6.3B, the current phase would *not* be the same from end to end. In this instance, the current is delayed 90 electrical degrees. The amount of phase difference depends on the circuit's electrical length.

Because the relationship between the physical size of a circuit and the wavelength of an ac current present in the circuit will vary as the frequency of the ac signal varies, the ideas of lumped and distributed effects actually occupy two ends of a spectrum. At HF (30 MHz and below), where $\lambda \geq 10$ m, the lumped element concept is almost always valid. In the UHF region and above, where $\lambda \leq 1$ m and physical component size can represent a significant fraction of a wavelength, everything shows distributed effects to one degree or another. From roughly 30 to 300 MHz, problems are usually examined on a case-by-case basis.

Of course, if we could make resistors, capacitors, inductors and so on, very small, we could treat them as lumped elements at much higher frequencies.

Thanks to the advances constantly being made in microelectronic circuit fabrication, this is in fact possible. Commercial monolithic microwave integrated circuits (MMICs) often use lumped elements that are valid up to 50 GHz. Since this frequency represents a wavelength of roughly 5 mm, this implies a component size of less than 0.5 mm.

LOW-FREQUENCY COMPONENT MODELS

Every circuit element behaves nonideally in some respect and under some conditions. It helps to know the most common types of nonideal behavior so we can design and build circuits that will perform as intended under expected ranges of operating conditions. In the sections below, we will discuss the basic components in order of increasing common nonideal behavior. Please note that much of this section applies only through HF; the peculiarities of VHF frequencies and above are discussed later in the chapter.

First, remember that some of the common limitations associated with real components are manufacturing concerns: tolerance and standard values. Tolerance is the measure of how much the actual value of a component may differ from its labeled value; it is usually expressed in percent. By convention, a "*higher*" (actually *closer*) tolerance component indicates

a *lower* (lesser) percentage deviation and will usually cost more. For example, a 1-kΩ resistor with a 5% tolerance has an actual resistance of anywhere from 950 to 1050 Ω. When designing circuits be careful to include the effects of tolerance. More specific examples will be discussed below.

Keep standard values in mind because not every conceivable component value is available. For instance, if circuit calculations yield a 4,932-Ω resistor for a demanding application, there may be trouble. The nearest commonly available values are 4700 Ω and 5600 Ω, and they'll be rated at 5% tolerance!

These two constraints prevent us from building circuits that do precisely what we wish. In fact, the measured performance of any circuit is the summation of the tolerances and temperature characteristics of all the components in both the circuit under test and the test equipment itself. As a result, most circuits are designed to operate within tolerances such as "up to a certain power" or "within this frequency range." Some circuits have one or more adjustable components that can compensate for variations in others. These limitations are then further complicated by other problems.

RESISTORS

Resistors are made in several different ways: carbon composition, carbon film, metal film, and wire wound. Carbon composition resistors are simply small cylinders of carbon mixed with various binding agents. Carbon is technically a semiconductor and can be *doped* with various impurities to produce any desired resistance. Most common everyday 1/2- and 1/4-W resistors are of this sort. They are moderately stable from 0 to 60 °C (their resistance increases above and below this temperature range). They are not inductive, but they are relatively noisy, and have relatively wide tolerances.

The other resistors exploit the fact that resistance is proportional to the length of the resistor and inversely proportional to its cross-sectional area:

Wire-wound resistors are made from wire, which is cut to the proper length and wound on a coil form (usually ceramic). They are capable of handling high power; their values are very stable, and they are manufactured to close tolerances.

Metal-film resistors are made by depositing a thin film of aluminum, tungsten or other metal on an insulating substrate. Their resistances are controlled by careful adjustments of the width, length and depth of the film. As a result, they have very close tolerances. They are used extensively in surface-mount technology. As might be expected, their power handling capability is somewhat limited. They also produce very little electrical noise.

Carbon film resistors use a film of doped carbon instead of metal. They are not quite as stable as other film resistors and have wider tolerances than metal-film resistors, but they are still as good as (or better than) composition resistors.

Resistors behave much like their ideal through AF; lead inductance becomes a problem only at higher frequencies. The major departure from ideal behavior is their temperature coefficient (TC). The resistivity of most materials changes with temperature, and typical TC values for resistor materials are given in **Table 6.1**. TC values are usually expressed in parts-per-million (PPM) for each degree (centigrade) change from some nominal temperature, usually room temperature (77 °F/27 °C). A positive TC indicates an increase in resistance with increasing temperature while a negative TC indicates a decreasing resistance. For example, if a 1000-Ω resistor with a TC of +300 PPM/°C is heated to 50°C, the *change* in resistance is 300(50 – 27) = 6900 PPM, yielding a new resistance of

$$1000\left(1+\frac{6900}{1000000}\right)=1006.9\ \Omega$$

Carbon-film resistors are unique among the major resistor families because they alone have a negative temperature coefficient. They are often used to "offset" the thermal effects of the other components (see Thermal Considerations, below).

If the temperature increase is small (less than 30-40 °C), the resistance change with temperature is nondestructive—the resistor will return to normal when the temperature returns to its nominal value. Resistors that get too hot to touch, however, may be permanently damaged even if they appear normal. For this reason, be conservative when specifying power ratings for resistors. It's common to specify a resistor rated at 200% to 400% of the expected dissipation.

Table 6.1

Temperature Coefficients for Various Resistor Compositions

1 PPM = 1 part per million = 0.0001%

Type	*TC (PPM/°C)*
Wire wound	±(30 - 50)
Metal Film	±(100 - 200)
Carbon Film	+350 to –800
Carbon composition	±800

Wire-wound resistors are essentially inductors used as resistors. Their use is therefore limited to dc or low-frequency ac applications where their reactance is negligible. Remember that this inductance will also affect switching transient waveforms, even at dc, because the component will act as an RL circuit.

As a rough example, consider a 1-Ω, 5-W wire-wound resistor that is formed from #24 wire on a 0.5-inch diameter form. What is the approximate associated inductance? First, we calculate the length of wire using the wire tables:

$$L=\frac{R}{\Omega/\text{ ft for \#24 wire}}$$

$$=\frac{1}{25.7\ \Omega/1000\text{ ft}}=39\text{ ft}$$

This yields a total of (39 × 12 inches) / (0.5 π inch/turn) ≈ 298 turns, which further yields a coil length (for #24 wire close-wound at 46.9 turns per inch) of 6.3 inches, assuming a single-layer winding. Then, from the inductance formula for air coils in the **Electrical Fundamentals** chapter, calculate

$$L=\frac{(0.5)^2\times 298^2}{18\times 0.5+40\times 6.3}=85\ \mu\text{H}$$

Real wire-wound resistors have multiple windings layered over each other to minimize both size and parasitic inductance (by winding each layer in opposite directions, much of the inductance is canceled). If we assume a five-layer winding, the length is reduced to 1.8 inches and the inductance to approximately 17 µH. If we want the inductive reactance to stay below 10% of the resistor value, then this resistor cannot be used above f = 0.1 / (2π 17 µH) = 937 Hz, or roughly 1 kHz.

Another exception to the rule that resistors are, in general, fairly ideal has to do with skin effect at RF. This will be discussed later in the chapter. **Fig 6.4** shows some more accurate circuit models for resistors at low frequencies. For a treatment of pure resistance theory, look at the **Electrical Fundamentals** chapter.

VOLTAGE AND CURRENT SOURCES

An ideal voltage source maintains a constant voltage across its terminals no matter how much current is drawn. Consequently, it is capable of providing infinite power, for an infinite period of time. Similarly, an ideal current source provides a

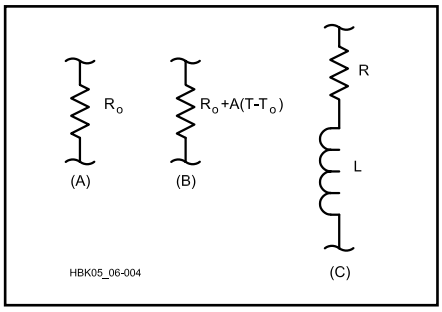


Fig 6.4—Circuit models for resistors. A is the ideal element. At B is the simple temperature-varying model for noninductive resistors. The wire-wound model with associated inductance is shown at C. For UHF or microwave designs, the model at C could be used with L representing lead inductance.

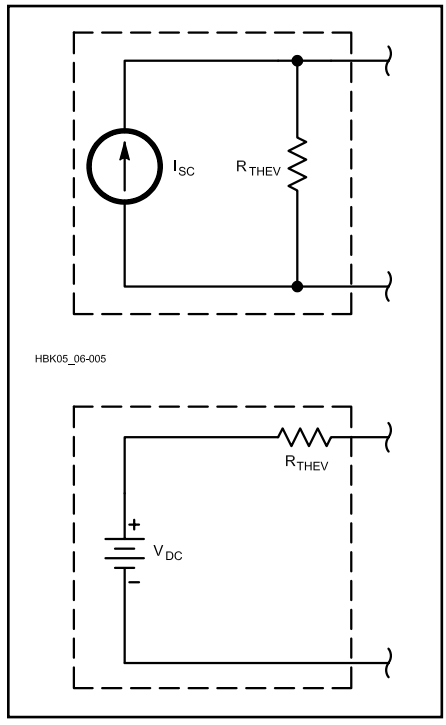


Fig 6.5—Norton and Thevenin equivalent circuits for real voltage and current sources.

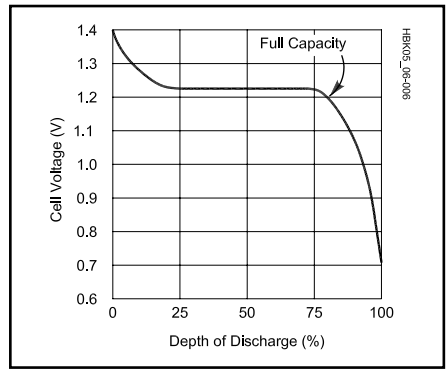


Fig 6.6—Discharge curve (terminal voltage vs percent capacity remaining) for a typical NiCd rechargeable battery. Note the dramatic voltage drop near the end.

constant current through its terminals no matter what voltage appears across it. It, too, can deliver an infinite amount of power for an infinite time.

Internal Resistance

As you may have learned through experience, *real* voltage and current sources—batteries and power supplies—do not meet these expectations. All real power sources have a finite *internal resistance* associated with them that limits the maximum power they can deliver.

We can model a real dc voltage source as a Thevenin-equivalent circuit of an *ideal* source in series with a resistance R_{thev} that is equal to the source's internal resistance. Similarly, we can model a real dc current source as a Norton-equivalent circuit: an ideal current source in *parallel* with R_{thev}. These two circuits shown in **Fig 6.5** are interchangeable through the relation

$$V_{OC} = I_{SC} \times R_{thev} \qquad (1)$$

Using these more realistic models, the maximum current that a real voltage source can deliver is seen to be I_{sc} and the maximum voltage is V_{oc}.

We can model sinusoidal voltage or current sources in much the same way, keeping in mind that the internal *impedance*, Z_{thev}, for such a source may not be purely resistive, but may have a reactive component that varies with frequency.

Battery Capacity and Discharge Curves

Batteries have a finite energy capacity as well as an internal resistance. Because of the physical size of most batteries, a convenient unit for measuring capacity is the *milliampere hour* (mAh) or, for larger units, the *ampere hour* (Ah). Note that these are units of charge (coulombs per second × seconds) and thus they measure the total net charge that the battery holds on its plates when full.

A capacity of 1 mAh indicates that a battery, when fully charged, holds sufficient electrochemical energy to supply a steady current of 1 mA for 1 hour. If a battery discharges at a constant rate, you could estimate the useful time of a battery charge by the simple formula

$$\text{Time (hr)} = \frac{\text{capacity (mAh)}}{\text{current (mA)}} \qquad (2)$$

In practice, however, the usable capacity of a rechargeable battery depends on the discharge *rate*, decreasing slightly as the current increases. For example, a 500 mAh nickel-cadmium (NiCd) battery may supply a current of 50 mA for almost 10 hours while maintaining a current of 500 mA for only 45 minutes.

The finite capacity of a battery does not *in itself* change the circuit model for a real source as shown in Fig 6.5. However, the chemistry of electrolytic cells, from which batteries are made, does introduce a minor, but significant, change. The voltage of a discharging battery does not stay constant, but slowly drops as time goes on. **Fig 6.6** illustrates the *discharge curve* for a typical NiCd rechargeable battery. For a given battery design, such curves are fairly reproducible and often used by battery-charging circuits to sense when a battery is fully charged or discharged.

CAPACITORS

The ideal capacitor is a pair of infinitely large parallel metal plates separated by an insulating or *dielectric* layer, ideally a vacuum (see **Fig 6.7).** For a discussion of pure capacitance see the **Electrical Fundamentals** chapter. For this case, the capacitance is given by

$$C = \frac{A\,\varepsilon_r\,\varepsilon_0}{d} \qquad (3)$$

where

C = capacitance, in farads
A = area of plates, cm
d = spacing of the plates, cm
ε_r = dielectric constant of the insulating material
ε_0 = permittivity of free space, 8.85×10^{-14} F/cm.

If the plates are not infinite, the actual capacitance is somewhat higher due to *end effect*. This is the same phenomenon that causes a dipole to resonate at a lower frequency if you place large insulators (which add capacitance) on the ends.

If we built such a capacitor, we would find that (neglecting the effects of quantum mechanics) it would be a perfect open circuit at dc, and it could be operated at whatever voltage we desire without breakdown.

Leakage Conductance and Breakdown

If we use anything other than a vacuum for the insulating layer, even air, we introduce two problems. Because there are

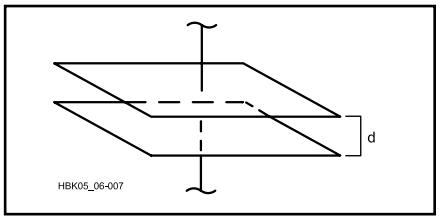


Fig 6.7—The ideal parallel-plate capacitor.

atoms between the plates, the capacitor will now be able to conduct a dc current. The magnitude of this *leakage current* will depend on the insulator quality, and the current is usually very small. Leakage current can be modeled by a resistance R_L in parallel with the capacitance (in the ideal case, this resistance is infinite).

In addition, when a high enough voltage is applied to the capacitor, the atoms of the dielectric will ionize due to the extremely high electric field and cause a large dc current to flow. This is *dielectric breakdown,* and it is destructive to the capacitor if the dielectric is ruined. To avoid dielectric breakdown, a capacitor has a *working voltage* rating, which represents the maximum voltage that can be permitted to develop across it.

Dielectrics

The leakage conductance and breakdown voltage characteristics of a capacitor are strongly dependent on the composition and quality of the dielectric. Various materials are used for different reasons such as availability, cost, and desired capacitance range. In rough order of "best" to "worst" they are:

Vacuum Both fixed and variable vacuum capacitors are available. They are rated by their maximum working voltages (3 to 60 kV) and currents. Losses are specified as negligible for most applications.

Air An *air-spaced* capacitor provides the best commonly available approximation to the ideal picture. Since $e_r = 1$ for air, air-dielectric capacitors are large when compared to those of the same value using other dielectrics. Their capacitance is very stable over a wide temperature range, leakage losses are low, and therefore a high Q can be obtained. They also can withstand high voltages. For these reasons (and ease of construction) most variable capacitors in tuning circuits are air-spaced.

Plastic film Capacitors with plastic film (polystyrene, polyethylene or Mylar) dielectrics are more expensive than paper capacitors, but have much lower leakage rates (even at high temperatures) and low TCs. Capacitance values are more stable than those of paper capacitors. In other respects, they have much the same characteristics as paper capacitors. Plastic-film variable capacitors are available.

Mica The capacitance of mica capacitors is very stable with respect to time, temperature and electrical stress. Leakage and losses are very low. Values range from 1 pF to 0.1 μF, with tolerances from 1 to 20%. High working voltages are possible, but they must be derated severely as operating frequency increases.

Silver mica capacitors are made by depositing a thin layer of silver on the mica dielectric. This makes the value even more stable, but it presents the possibility of silver migration through the dielectric. The migration problem worsens with increased dc voltage, temperature and humidity. Avoid using silver-mica capacitors under such conditions.

Ceramic There are two kinds of ceramic capacitors. Those with a low dielectric constant are relatively large, but very stable and nearly as good as mica capacitors at HF. High dielectric constant ceramic capacitors are physically small for their capacitance, but their value is not as stable. Their dielectric properties vary with temperature, applied voltage and operating frequency. They also exhibit piezoelectric behavior. Use them only in coupling and bypass roles. Tolerances are usually +100% and –20%. Ceramic capacitors are available in a wide range of values: 10 pF to 1 μF. Some variable units are available.

Electrolytic These capacitors have the space between their foil plates filled with a chemical paste. When voltage is applied, a chemical reaction forms a layer of insulating material on the foil.

Electrolytic capacitors are popular because they provide high capacitance values in small packages at a reasonable cost. Leakage is high, as is inductance, and they are polarized—there is a definite positive and negative plate, due to the chemical reaction that provides the dielectric. Internal inductance restricts aluminum-foil electrolytics to low-frequency applications. They are available with values from 1 to 500,000 μF.

Tantalum electrolytic capacitors perform better than aluminum units but their cost is higher. They are smaller, lighter and more stable, with less leakage and inductance than their aluminum counterparts. Reformation problems are less frequent, but working voltages are not as high as with aluminum units.

Electrolytics should not be used if the dc potential is well below the capacitor working voltage.

Paper Paper capacitors are inexpensive; capacitances from 500 pF to 50 μF are available. High working voltages are possible, but paper-dielectric capacitors have high leakage rates and tolerances are no better than 10 to 20%. Paper-dielectric capacitors are not polarized; however, the body of the capacitor is usually marked with a color band at one end. The band indicates the terminal that is connected to the outermost plate of the capacitor. This terminal should be connected to the side of the circuit at the lower potential as a safety precaution.

Loss Angle

For ac signals (even at low frequencies), capacitors exhibit an additional parasitic resistance that is due to the electromagnetic properties of dielectric materials. This resistance is often quantified in catalogs as *loss angle,* θ, because it represents the angle, in the complex impedance plane, between $Z_C = R + jX_C$ and X_C. This angle is usually quite small, and would be zero for an ideal capacitor.

Loss angle is normally specified as $\tan \theta$ at a certain frequency, which is simply the ratio R/X_C. The loss angle of a given capacitor is relatively constant over frequency, which means the *effective series resistance* or $ESR = (\tan \theta) / (2 \pi f C)$ goes down as frequency goes up. This resistance is placed in *series* with the capacitor because it came (mathematically) from the equation for Z_C above. It can always be converted into a parallel resistance if desired. To summarize, **Figs 6.8** and **6.9** show reasonable models for the capacitor that are good up to VHF.

Temperature Coefficients and Tolerances

As with resistors, capacitor values vary in production, and most capacitors have a tolerance rating either printed on them or listed on a data sheet. Capacitance varies with temperature, and this is important to consider when constructing a circuit that will carry high power levels or operate in a hot environment. Also, as just described, not all capacitors are available in all ranges of values due to inherent differences in material properties. Typical values, temperature coefficients and leakage conductances for several capacitor types are given in **Table 6.2.**

DIODES

An ideal diode acts as a rectifying switch—it is a short circuit when forward biased and an open circuit when reverse biased. Many circuit components can be completely described in terms of current-voltage (or *I-V*) characteristics, and to

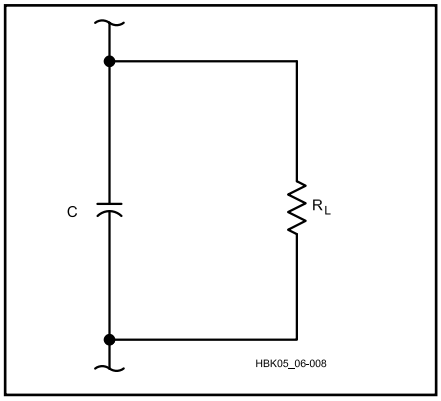


Fig 6.8—A simple capacitor model for frequencies well below self-resonance.

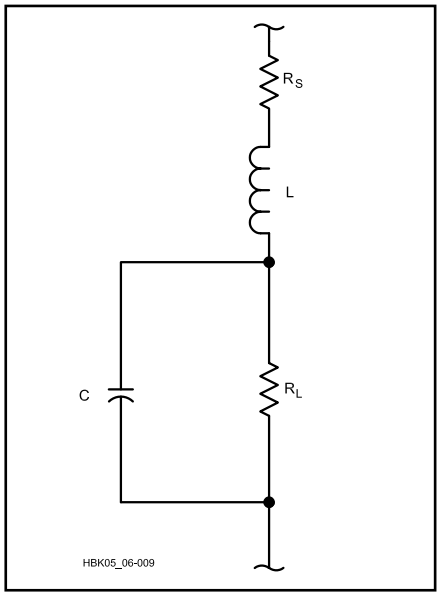


Fig 6.9—A capacitor model for VHF and above including series resistance and distributed inductance.

Table 6.2

Typical Temperature Coefficients and Leakage Conductances for Various Capacitor Constructions

Type	*TC @ 20°C (PPM/°C)*	*DC Leakage Conductance (Ω)*
Ceramic Disc	±300(NP0)	> 10 M
	+150/–1500(GP)	> 10 M
Mica	–20 to +100	> 100,000 M
Polyester	±500	> 10 M
Tantalum Electrolytic	±1500	> 10 MΩ
Small Al Electrolytic(≈ 100 µF)	–20,000	500 k - 1 M
Large Al Electrolytic(≈ 10 mF)	–100,000	10 k
Vacuum (glass)	+100	≈ ∞
Vacuum (ceramic)	+50	≈ ∞

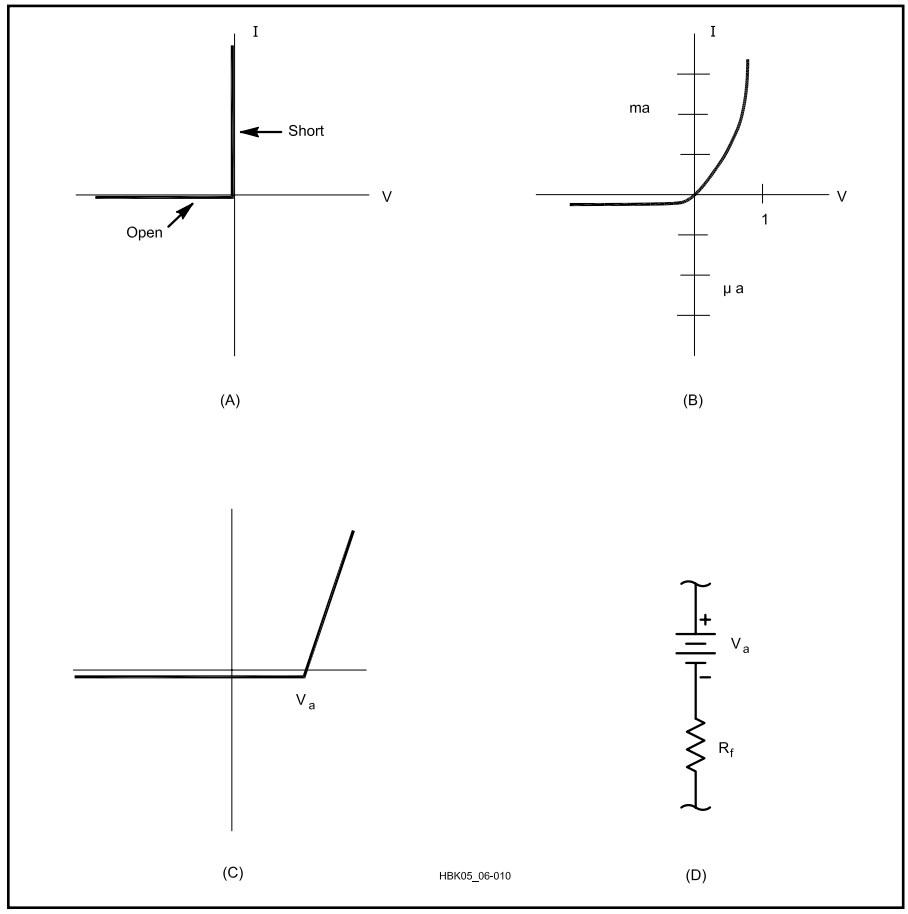


Fig 6.10—Circuit models for rectifying switches (diodes). A: I-V curve of the ideal rectifier. B: I-V curve of a typical semiconductor diode. Note the different scales for + and – current. C shows a simplified diode I-V curve for dc-circuit calculations. D is an equivalent circuit for C.

discuss the diode, this approach is especially helpful. **Fig 6.10A** shows the I-V curve for an ideal rectifier.

In contrast, the I-V curve for a semiconductor diode junction is given by the following equation (slightly simplified).

$$I = I_S \left(e^{\frac{V}{V_t}} \right) \qquad (4)$$

where

I = diode current
V = diode voltage
I_S = reverse-bias saturation current
V_t = kT/q, the thermal equivalent of voltage (about 25 mV at room temperature).

This curve is shown in Fig 6.10B.

The obvious differences between Fig 6.10A and B are that the semiconductor diode has a finite *turn-on* voltage—it requires a small but nonzero forward bias voltage before it begins conducting. Furthermore, once conducting, the diode voltage continues to increase very slowly with increasing current, unlike a true short circuit. Finally, when the applied voltage is negative, the current is not exactly zero but very small (microamperes).

For bias (dc) circuit calculations, a useful model for the diode that takes these two effects into account is shown by the artificial I-V curve in Fig 6.10C. The small reverse bias current I_s is assumed to be completely negligible.

When converted into an equivalent circuit, the model in Fig 6.10C yields the picture in Fig 6.10D. The ideal voltage source

V_a represents the turn-on voltage and R_f represents the effective resistance caused by the small increase in diode voltage as the diode current increases. The turn-on voltage is material-dependent: approximately 0.3 V for germanium diodes and 0.7 for silicon. R_f is typically on the order of 10 Ω, but it can vary according to the specific component. R_f can often be completely neglected in comparison to the other resistances in the circuit. This very common simplification leaves only a pure voltage drop for the diode model.

Temperature Bias Dependence

The reverse saturation current I_s is not constant but is itself a complicated function of temperature. For silicon diodes (and transistors) near room temperature, I_s increases by a factor of 2 every 4.8 °C. This means that for every 4.8 °C rise in temperature, either the diode current doubles (if the voltage across it is constant), or if the current is held constant by other resistances in the circuit, the diode voltage will *decrease* by $V_t \times \ln 2 =$ 18 mV. For germanium, the current doubles every 8 °C and for gallium arsenide (GaAs), 3.7 °C. This dependence is highly reproducible and may actually be exploited to produce temperature-measuring circuits.

While the change resulting from a rise of several degrees may be tolerable in a circuit design, that from 20 or 30 degrees may not. Therefore it's a good idea with diodes, just as with other components, to specify power ratings conservatively (2 to 4 times over) to prevent self-heating.

While component derating does reduce self-heating effects, circuits must be designed for the expected operating environment. For example, mobile radios may face temperatures from –20° to +140°F (–29° to 60°C).

Junction Capacitance

Immediately surrounding a PN junction is a *depletion layer*. This is an electrically charged region consisting primarily of ionized atoms with relatively few elec-trons and holes (see **Fig 6.11**). Outside the depletion layer are the remainder of the P and N regions, which do not contribute to diode operation but behave primarily as parasitic series resistances.

We can treat the depletion layer as a tiny capacitor consisting of two parallel plates. As the reverse bias applied to a diode changes, the width of the depletion layer, and therefore the capacitance, also changes. There is an additional *diffusion capacitance* that appears under forward bias due to electron and hole storage in the bulk regions, but we will not discuss this here because diodes are usually reverse biased when their capacitance is exploited. The diode junction capacitance under a reverse bias of V volts is given by

$$C_j = C_{j0} \Big/ \sqrt{V_{on} - V} \qquad (5)$$

where C_{j0} = measured capacitance with zero applied voltage.

Note that the quantity under the radical is a large *positive* quantity for reverse bias.

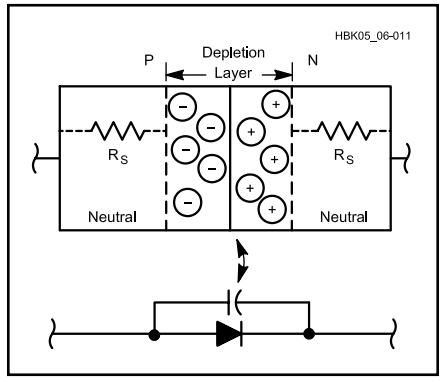


Fig 6.11—A more detailed picture of the PN junction, showing depletion layer, bulk regions and junction capacitance.

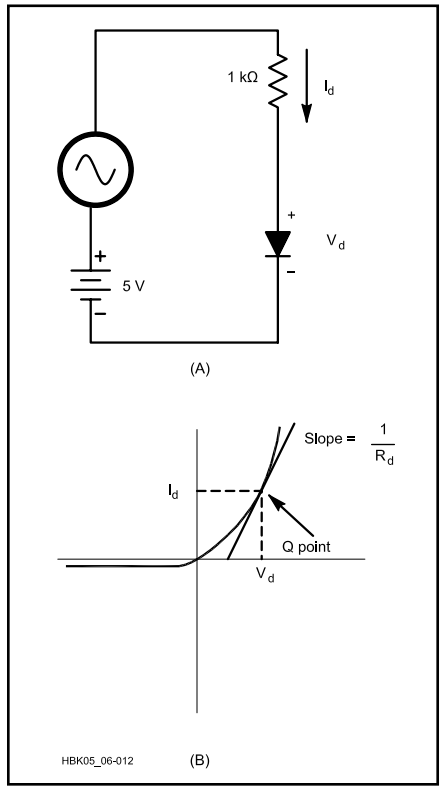


Fig 6.12—A simple resistor-diode circuit used to illustrate dynamic resistance. The ac input voltage "sees" a diode resistance whose value is the slope of the line at the Q-point, shown in B.

As seen from the equation, for large reverse biases C_j is inversely proportional to the square root of the voltage.

Junction capacitances are small, on the order of pF. They become important, however, for diode circuits at RF, as they can affect the resonant frequency. In fact, *varactors* are diodes used for just this purpose.

Reverse Breakdown and Zener Diodes

If a sufficiently large reverse bias is applied to a real diode, the internal electric field becomes so strong that electrons and holes are ripped from their atoms, and a large reverse current begins to flow. This is called *breakdown* or *avalanche*. Unless the current is so large that the diode fails from overheating, breakdown is not destructive and the diode will again behave normally when the bias is removed.

Once a diode breaks down, the voltage across it remains relatively constant regardless of the level of breakdown current. *Zener* diodes are diodes that are specially made to operate in this region with a very constant voltage, called the *Zener voltage*. They are used primarily as voltage regulators. When operating in this region, they can be modeled as a simple voltage source. Zener diodes are rated by their Zener voltage and power dissipation.

A Diode AC Model

Fig 6.12A shows a simple resistor-diode circuit to which is applied a dc bias voltage plus an ac signal. Assuming that the voltage drop across the diode is 0.6 V and R_f is negligible, we can calculate the bias current to be I = (5 – 0.6) / 1 kΩ = 4.4 mA. This point is marked on the diodes I-V curve in Fig 6.12B. If we draw a line tangent to this point, as shown, the slope of this line represents the *dynamic* resistance R_d of the diode seen by a small ac signal, which at room temperature can be approximated by

$$R_d = \frac{25}{I}\,\Omega \qquad (6)$$

where I is the diode current in mA. Note that this resistance changes with bias current and should not be confused with the dc forward resistance in the previous section, which has a similar value but represents a different concept. **Fig 6.13** shows a low-frequency ac model for the diode, including the dynamic resistance and junction capacitance.

Switching Time

If you change the polarity of a signal applied to the ideal switch whose I-V curve appears in Fig 6.10A, the switch

turns on or off instantaneously. A real diode cannot do this, as a finite amount of time is required to move electrons and holes in or out of the diode as it changes states (effectively, the diode capacitances must be charged or discharged). As a result, diodes have a maximum useful frequency when used in switching applications. The operation of diode switching circuits can often be modeled by the picture in **Fig 6.14.** The approximate switching time (in seconds) for this circuit is given by

$$t_s = \tau_p \frac{\left(\frac{V_1}{R_1}\right)}{\left(\frac{V_2}{R_2}\right)} = \tau_p \frac{I_1}{I_2} \qquad (7)$$

where τ_p is the minority carrier lifetime of the diode (a material constant determined during manufacture, on the order of 1 ms). I_1 and I_2 are currents that flow during the switching process. The minimum time in which a diode can switch from one state to the other and back again is therefore 2 t_s, and thus the maximum usable switching frequency is f_{sw} (Hz) = $^1/_2$ t_s. It is usually a good idea to stay below this by a factor of two. Diode data sheets usually give typical switching times and show the circuit used to measure them.

Note that f_{sw} depends on the forward and reverse currents, determined by I_1 and I_2 (or equivalently V_1, V_2, R_1, and R_2). Within a reasonable range, the switching time can be reduced by manipulating these currents. Of course, the maximum power that other circuit elements can handle places an upper limit on switching currents.

Schottky Diodes

Schottky diodes are made from metal-semiconductor junctions rather than PN junctions. They store less charge internally and as a result, have shorter switching times and junction capacitances than standard PN junction diodes.

Their turn-on voltage is also less, typically 0.3 to 0.4 V. In most other respects they behave similarly to PN diodes.

INDUCTORS

Inductors are the problem children of the component world. (Fundamental inductance is discussed in the **Electrical Fundamentals** chapter). Besides being difficult to fabricate on integrated circuits, they are perhaps the most nonideal of real-world components. While the leakage conductance of a capacitor is usually negligible, the series resistance of an inductor often is not. This is basically because an inductor is made of a long piece of relatively thin wire wound into a small coil. As an example, consider a typical air-core tuning coil from a component catalog, with L = 33 mH and a minimum Q of 30 measured at 2.5 MHz. This would indicate a series resistance of $R_s = 2 \pi f L / Q = 17 \Omega$. This R_s could significantly alter the resonant frequency of a circuit. For frequencies up to HF this is the only significant nonlinearity, and a low-frequency circuit model for the inductor is shown in **Fig 6.15.** Many of the problems associated with inductors are actually due to core materials, and not the coil itself.

Magnetic Materials

Many discrete inductors and transformers are wound on a core of iron or other magnetic material. As discussed in the **Electrical Fundamentals** chapter, the inductance value of a coil is proportional to the density of magnetic field lines that pass through it. A piece of magnetic material placed inside a coil will concentrate the field lines inside itself. This permits a higher number of field lines to exist inside a coil of a given cross-sectional area, thus allowing a much higher inductance than what would be possible with an air core. This is especially important for transformers that operate at low frequencies (such as 60 Hz) to ensure the reactance of the windings is high.

Core Saturation

Magnetic (more precisely, ferro-magnetic) materials exhibit two kinds of nonlinear behavior that are important to circuit design. The first is the phenomenon of *saturation*. A magnetic core increases the magnetic flux density of a coil because the current passing through the coil forces the atoms of the iron (or other material) to line up, just like many small compass needles, and the magnetic field that results from the atomic alignment is *much* larger than that produced by the current with no core. As coil current increases, more and more atoms line up. At some high current, all of the atoms will be aligned and the core is *saturated*. Any further increase in current can't increase the core alignment any further. Of course, added current generates its own magnetic flux, but it is very small compared to the magnetic field contributed by the aligned atoms in the magnetic core.

The important concept in terms of circuit design is that as long as the coil current remains below saturation, the inductance of the coil is essentially constant. **Fig 6.17** shows graphs of magnetic flux linkage (N f) and inductance (L) vs current (i) for a typical iron-core inductor. These quantities are related by the equation

$$N\phi = Li \qquad (8)$$

where
N = number of turns,
ϕ = flux density
L = inductance
i = current.

In the lower graph, a line drawn from any point on the curve to the (0,0) point will show the effective inductance, $L = N \phi / i$,

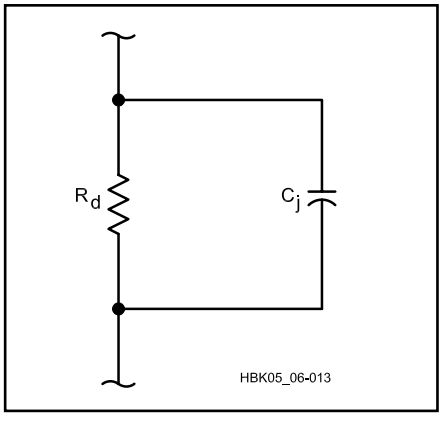


Fig 6.13—An ac model for diodes. R_d is the dynamic resistance and C_j is the junction capacitance.

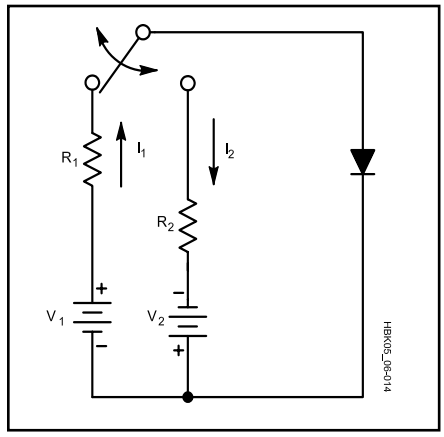


Fig 6.14—Circuit used for computation of diode switching time.

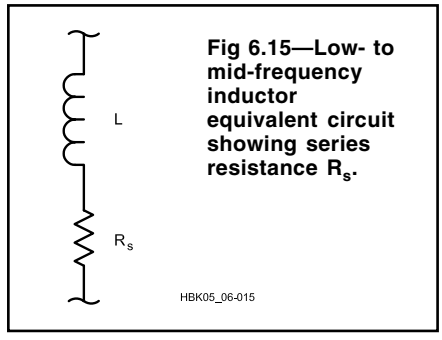


Fig 6.15—Low- to mid-frequency inductor equivalent circuit showing series resistance R_s.

at that current. These results are plotted on the upper graph.

Note that below saturation, the inductance is constant because both N ϕ and i are increasing at a steady rate. Once the saturation current is reached, the inductance decreases because N ϕ does not increase anymore (except for the tiny additional magnetic field the current itself provides). This may render some coils useless at VHF and higher frequencies. Air-coil inductors do not suffer from saturation.

One common method of increasing the saturation current level is to cut a small air gap in the core (see **Fig 6.16**). This gap forces the flux lines to travel through air for a short distance. Since the saturation flux linkage of the core is unchanged, this method works by requiring a higher current to achieve saturation. The price that is paid is a reduced inductance below saturation. Curves B in Fig 6.17 show the result of an air gap added to that inductor.

Manufacturer's data sheets for magnetic cores usually specify the saturation flux density. Saturation flux density, ϕ, in gauss can be calculated for ac and dc currents from the following equations:

$$\phi_{ac} = \frac{3.49\ V}{fNA} \quad (9)$$

$$\phi_{dc} = \frac{NIA_L}{10\ A} \quad (10)$$

where

V = RMS ac voltage

f = frequency, in MHz

N = number of turns

A = equivalent area of the magnetic path in square inches (from the data sheet)

I = dc current, in A

A_L = inductance index (also from the data sheet).

Hysteresis

Consider **Fig 6.18**. If the current passed through a magnetic-core inductor is increased from zero (point a) to near the point of saturation (point b) and then decreased back to zero, we find that a magnetic field remains, because some of the core atoms retain their alignment. If we then increase the current in the opposite direction and again return to zero (through points c, d and e), the curve does *not* retrace itself. This is the property of *hysteresis*.

If a circuit carries a large ac current (that is, equal to or larger than saturation), the path shown in Fig 6.18 (from b to e and back again) will be retraced many times each second. Since the curve is nonlinear, this will introduce distortion in the resulting waveform. Where linear circuit operation is crucial, it is important to restrict the operation of magnetic-core inductors well below saturation.

Eddy Currents

The changing magnetic field produced by an ac current generates a "back voltage" in the core as well as the coil itself. Since magnetic core material is usually conductive, this voltage causes a current to flow in the core. This eddy current serves no useful purpose and represents power lost to heat. Eddy currents can be substantially reduced by laminating the core—slicing the core into thin sheets and placing a suitable insulating material (such as varnish) between them (see Fig 6.16).

Transformers

In an ideal transformer (described in the **Electrical Fundamentals** chapter), *all* the power supplied to its primary terminals is available at the secondary terminals. An air-core transformer provides a very good approximation to the ideal case, the major loss being the series resistances of the windings. As you might guess, there are several ways that a magnetic-core transformer can lose power. For example, useless heat can be generated through hysteresis and eddy currents; power may be lost to harmonic generation through nonlinear saturation effects.

Another source of loss in transformers (or actually, any pair of mutual inductances) is *leakage reactance*. While the main purpose of a magnetic core is to concentrate the magnetic flux entirely within itself, in a real-life device there is always some small amount of flux that does not pass through both windings. This leakage reactance can be pictured as a small amount of self-inductance appearing on

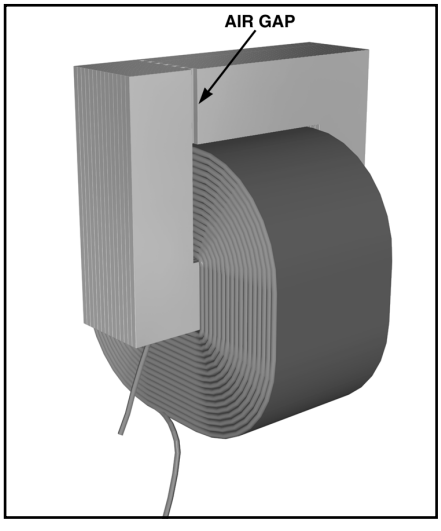


Fig 6.16—Typical construction of a magnetic-core inductor. The air gap greatly reduces core saturation at the expense of some inductance. The insulating laminations between the core layers help to minimize eddy currents.

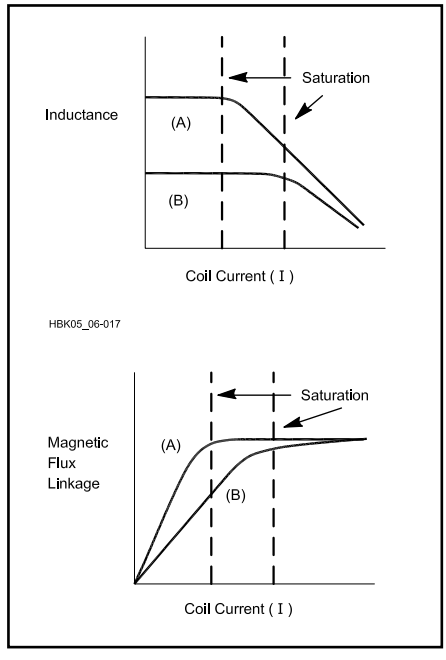


Fig 6.17—Magnetic flux linkage and inductance plotted versus coil current for (A) a typical iron-core inductor. As the flux linkage N ϕ in the coil saturates, the inductance begins to decrease since inductance = flux linkage / current. The curves marked B show the effect of adding an air gap to the core. The current handling capability has increased but at the expense of reduced inductance.

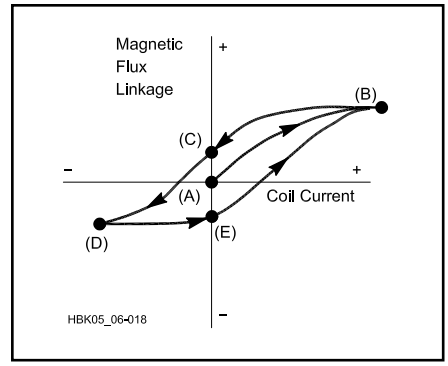


Fig 6.18—Hysteresis loop. A large current (larger than saturation current) passed through a magnetic-core coil will cause some permanent magnetization of the core. This results in a different path of flux linkage vs current to be traced as the current decreases.

each winding. **Fig 6.19** shows a fairly complete circuit model for a transformer at low to medium frequencies.

COMPONENTS AT RF

The models described in the previous section are good for dc, and ac up through AF and low RF. At HF and above (where we do much of our circuit design) several other considerations become very important, in some cases dominant, in our component models. To understand what happens to circuits at RF (a good cutoff is 30 MHz and above) we turn to a brief discussion of some electromagnetic and microwave theory concepts.

Parasitic Inductance

Maxwell's equations—the basic laws of electromagnetics that govern the propagation of electromagnetic waves and the operation of all electronic components—tell us that any wire carrying a current that changes with time (one example is a sine wave) develops a changing magnetic field around it. This changing magnetic field in turn induces an opposing voltage, or back EMF, on the wire. The back EMF is proportional to how fast the current changes (see **Fig 6.20**).

We exploit this phenomenon when we make an inductor. The reason we typically form inductors in the shape of a coil is to concentrate the magnetic field lines and thereby maximize the inductance for a given physical size. However, *all* wires carrying varying currents have these inductive properties. This includes the wires we use to connect our circuits, and even the *leads* of capacitors, resistors and so on.

The inductance of a straight, round, nonmagnetic wire in free space is given by:

$$L = 0.00508\,b\left[\ln\left(\frac{2b}{a}\right) - 0.75\right] \qquad (11)$$

where

L = inductance, in µH
a = wire radius, in inches
b = wire length, in inches
ln = natural logarithm ($2.303 \times \log_{10}$).

Skin effect (see below) changes this formula slightly at VHF and above. As the frequency approaches infinity, the constant 0.75 in the above equation approaches 1. This effect usually presents no more than a few percent change.

As an example, let's find the inductance of a #18 wire (diam = 0.0403 inch) that is 4 inches long (a typical wire in a circuit). Then a = 0.0201 and b = 4:

$$L = 0.00508(4)\left[\ln\left(\frac{8}{0.0201}\right) - 0.75\right]$$

$$= 0.0203\left[5.98 - 0.75\right] = 0.106\ \mu H$$

It is obvious that this *parasitic* inductance is usually very small. It becomes important *only at very high frequencies*; at AF or LF, the parasitic inductive reactance is practically zero. To use this example, the reactance of a 0.106 µH inductor even at 10 MHz is only 6.6 Ω. **Fig 6.21** shows a graph of the inductance for wires of various gauges (radii) as a function of length.

Any circuit component that has wires attached to it, or is fabricated from wire, will have a parasitic inductance associated with it. We can treat this parasitic inductance in component models by adding an inductor of appropriate value in *series* with the component (since the wire lengths are in series with the element). This (among other reasons) is why minimizing lead lengths and interconnecting wires becomes very important when designing circuits for VHF and above.

Parasitic Capacitance

Maxwell's equations also tell us that if the voltage between any two points changes with time, a *displacement* current is generated between these points. See **Fig 6.22.** This displacement current results from the propagation, at the speed of light, of the electromagnetic field between the two points and is not to be confused with conduction current, which is caused by the movement of electrons. This displacement current is directly proportional to how fast the voltage is changing.

A capacitor takes advantage of this consequence of the laws of electromagnetics. When a capacitor is connected to an ac voltage source, a steady ac current can flow because taken together, conduction current and displacement current "complete the loop" from the positive source terminal, across the plates of the capacitor, and back to the negative terminal.

In general, parasitic capacitance shows up *wherever* the voltage between two points is changing with time, because the laws of electromagnetics require a displacement current to flow. Since this phenomenon represents an *additional* current path from one point in space to another, we can add this parasitic capacitance to our component models by adding a capacitor of appropriate value in *parallel* with the component. These parasitic capacitances are typically less than 1 pF, so that below VHF they can be treated as open circuits (infinite resistances) and thus neglected.

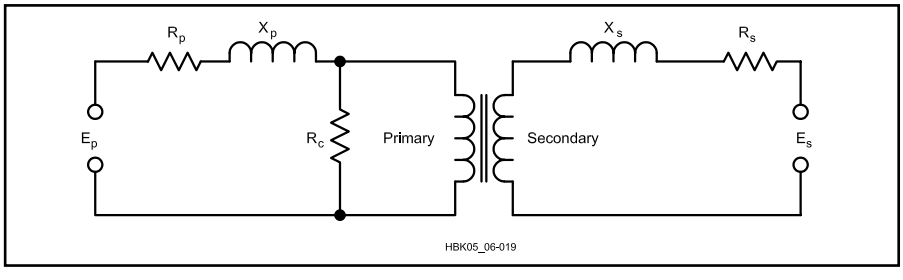


Fig 6.19—An equivalent circuit for a transformer at low to medium frequencies. R_p and R_s represent the series resistances of the windings, X_p and X_s are the leakage inductances, and R_c represents the dissipative losses in the core due to eddy currents and so on.

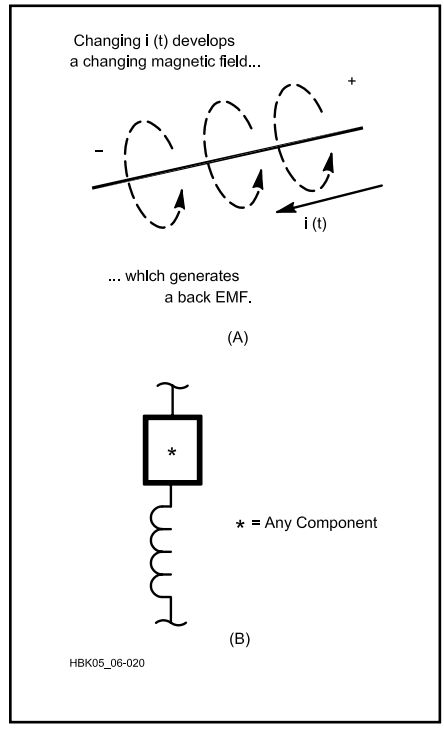


Fig 6.20—Inductive consequences of Maxwell's equations. At A, any wire carrying a changing current develops a voltage difference along it. This can be mathematically described as an effective inductance. B adds parasitic inductance to a generic component model.

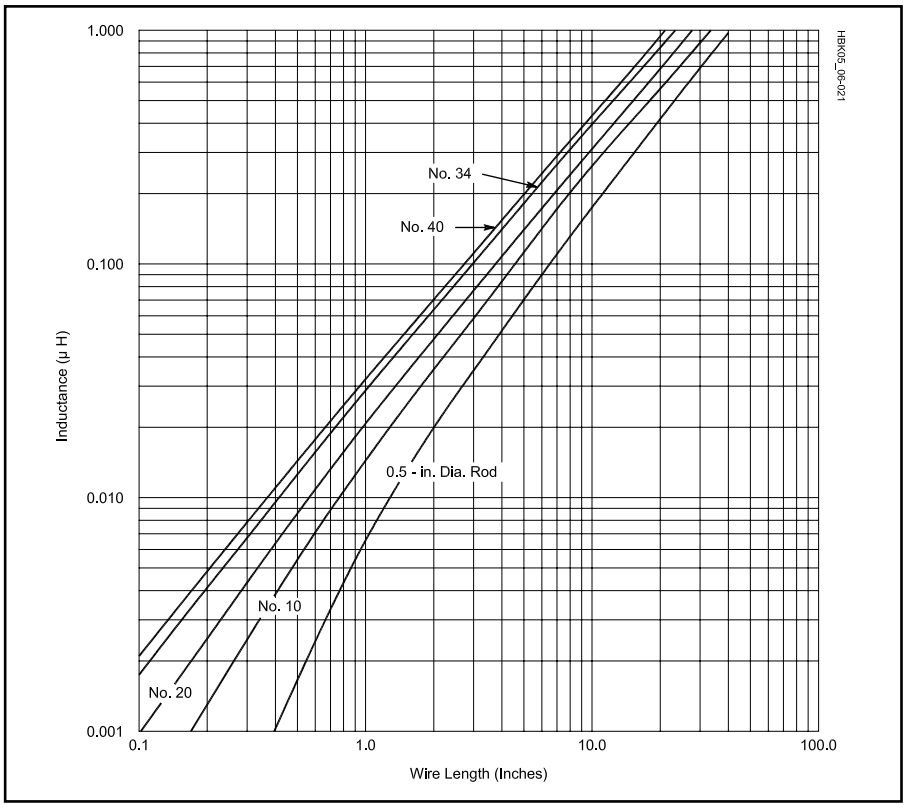


Fig 6.21—A plot of inductance vs length for straight conductors in several wire sizes.

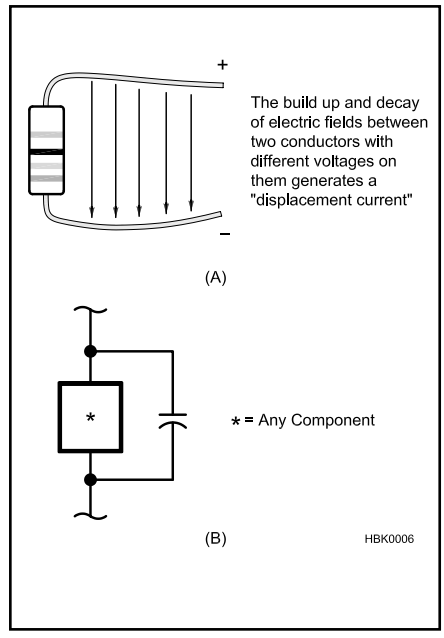


Fig 6.22—Capacitive consequences of Maxwell's equations. A: Any changing voltage between two points, for example along a bent wire, generates a displacement current running between them. This can be mathematically described as an effective capacitance. B adds parasitic capacitance to a generic component model.

Consider the inductor in **Fig 6.23**. If this coil has n turns, then the ac voltage between identical points of two neighboring turns is 1/n times the ac voltage across the entire coil. When this voltage changes due to an ac current passing through the coil, the effect is that of many small capacitors acting in parallel with the inductance of the coil. Thus, in addition to the capacitance resulting from the leads, inductors have higher parasitic capacitance due to their physical shape.

Package Capacitance

Another source of capacitance, also in the 1-pF range and therefore important only at VHF and above, is the packaging of the component itself. For example, a power transistor packaged in a TO-220 case (see **Fig 6.24**), often has either the emitter or collector connected to the metal tab itself. This introduces an extra *inter-electrode capacitance* across the junctions.

The copper traces on a PC board also present capacitance to the circuit it holds. Double-sided PC boards have a certain capacitance per square inch. It is possible to create capacitors by leaving unetched areas of copper on both sides of the board. The capacitance is not well controlled on inexpensive low-frequency boards, however. For this reason, the copper on one side of a double-sided board should be completely removed under frequency-determining circuits such as VFOs. Board capacitance is *exploited* to make microwave transmission lines (microstrip lines). The capacitance of boards for microwave use is better controlled than that of less expensive board material.

Stray capacitance (a general term used for any "extra" capacitance that exists due to physical construction) appears in any circuit where two metal surfaces exist at different voltages. Such effects can be modeled as an extra capacitor in parallel with the given points in the circuit. A rough value can be obtained with the parallel-plate formula given above.

Thus, similar to inductance, *any* circuit component that has wires attached to it, or is fabricated from wire, or is near or attached to metal, will have a parasitic capacitance associated with it, which again, becomes important only at RF.

A GENERAL MODEL

The parasitic problems due to component leads, packaging, leakage and so on are relatively common to all components. When working at frequencies where many or all of the parasitics become important, a complex but completely general model such as that in **Fig 6.25** can be used for just about any component, with the actual component placed in the box marked "*". Parasitic capacitance C_p and leakage conductance G_L appear in parallel across the device, while series resistance R_s and parasitic inductance L_p appear in series with it. Package capacitance C_{pkg} appears as an additional capacitance in parallel across the whole device. This maze of effects may seem overwhelming, but remember that it is very seldom necessary to consider all parasitics. In the Computer-Aided-Design section of this chapter, we will use the power of the computer to examine the combined effects of these multiple parasitics on circuit performance.

Self-Resonance

Because of the effects just discussed, a capacitor or inductor—all by itself—exhibits the properties of a resonant RLC circuit as we increase the applied frequency. **Fig 6.26** illustrates RF models for the capacitor and inductor, which are based on the general model in Fig 6.25, leaving out the packaging capacitance. Note the slight difference in configuration; the pairs C_p, R_p and L_s, R_s are in series in the capacitor but in parallel in the

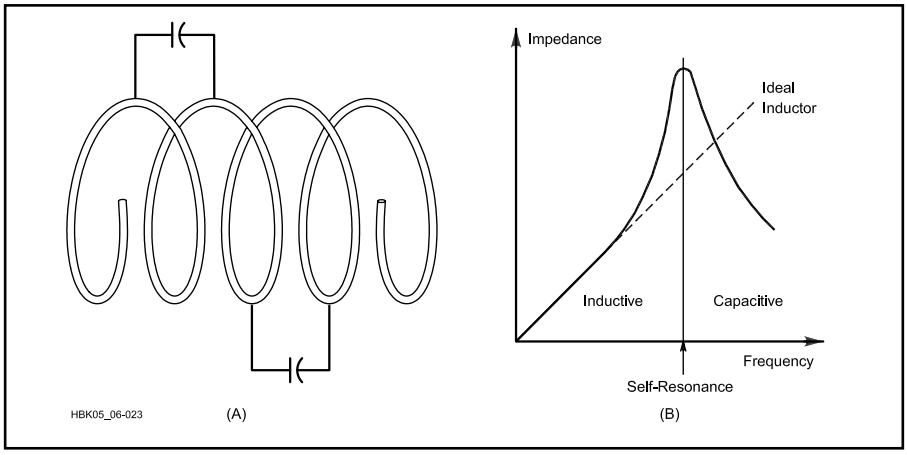


Fig 6.23—Coils exhibit distributed capacitance as explained in the text. The graph at B shows how distributed capacitance resonates with the inductance. Below resonance, the reactance is predominantly inductive which increases as frequency increases. However, above resonance, the reactance becomes predominantly capacitive which decreases as frequency increases.

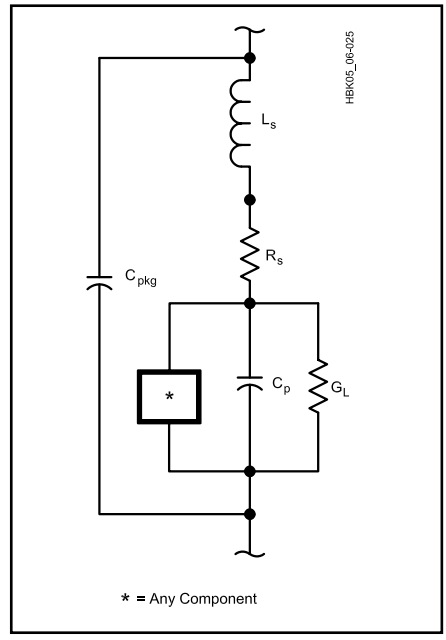


Fig 6.25—A general model for electrical components at VHF frequencies and above. The box marked "*" represents the component itself. See text for discussion.

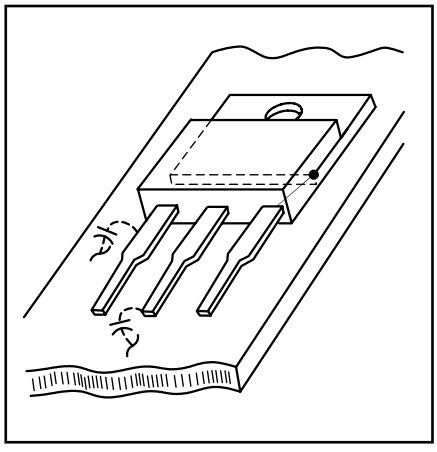


Fig 6.24—Unexpected stray capacitance. The mounting tab of TO-220 transistors is often connected to one of the device leads. Because one lead is connected to the chassis, small capacitances from the other lead to the chassis appear as additional package capacitance at the device. Similar capacitance can appear at any device with a conductive package.

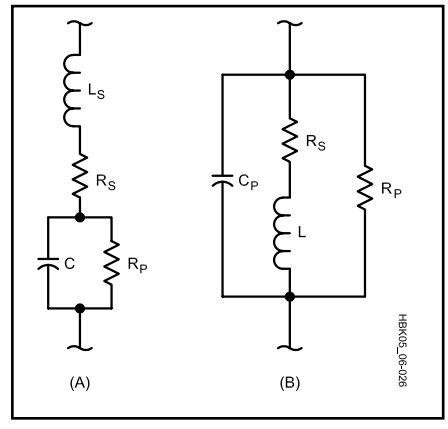


Fig 6.26—Capacitor (A) and inductor (B) models for RF frequencies.

inductor. This is because in the inductor, C_p and R_p are the parasitics, while in the capacitor, L_s and R_s are the added effects.

At some sufficiently high frequency, both inductors and capacitors become *self-resonant*. Just like a tuned circuit, above that frequency the capacitor will appear inductive, and the inductor will appear capacitive.

For an example, let's calculate the approximate self-resonant frequency of a 470-pF capacitor whose leads are made from #20 wire (diam 0.032 inch), with a total length of 1 inch. From the formula above, we calculate the approximate parasitic inductance

$$L(\mu H) = 0.00508(1)\left[\ln\left(\frac{2(1)}{(0.032/2)}\right) - 0.75\right]$$
$$= 0.021\,\mu H$$

and then the self-resonant frequency is roughly

$$f = \frac{1}{2\pi\sqrt{LC}} = 50.6\ MHz$$

The purpose of making these calculations is to give you a rough feel for actual component values. They could be used as a rough design guideline, but should not be used quantitatively. Other factors such as lead orientation, shielding and so on, can alter the parasitic effects to a large extent. Large-value capacitors tend to have higher parasitic inductances (and therefore a lower self-resonant frequency) than small-value ones.

Self-resonance becomes critically important at VHF and UHF because the self-resonant frequency of many common components is at or below the frequency where the component will be used. In this case, either special techniques can be used to construct components to operate at these frequencies, by reducing the parasitic effects, or else the idea of lumped elements must be abandoned altogether in favor of microwave techniques such as striplines and waveguides.

Skin Effect

The resistance of a conductor to ac is different than its value for dc. A consequence of Maxwell's equations is that thick, near-perfect conductors (such as metals) conduct ac only to a certain depth that is proportional to the wavelength of the signal. This decreases the effective cross-section of the conductor at high frequencies and thus increases its resistance.

This resistance increase, called *skin effect*, is insignificant at low (audio) frequencies, but beginning around 1 MHz (depending on the size of the conductor) it is so pronounced that practically all the current flows in a very thin layer near the conductor's surface. For this reason, at RF a hollow tube and a solid rod of the same diameter and made of the same metal will have the same resistance. The depth of this skin layer decreases by a factor of 10 for every 100 × increase in frequency. Consequently, the RF resistance is often much

higher than the dc resistance. Also, a thin highly conductive layer, such as silver plating, can lower resistance for UHF or microwaves, but does little to improve HF conductivity.

A rough estimate of the cutoff frequency where a nonferrous wire will begin to show skin effect can be calculated from

$$f = \frac{124}{d^2} \qquad (12)$$

where

f = frequency, in MHz

d = diam, in mils (a mil is 0.001 inch).

Above this frequency, increase the resistance of the wire by 10 × for every 2 decades of frequency (roughly 3.2 × for every decade). For example, say we wish to find the RF resistance of a 2-inch length of #18 copper wire at 100 MHz. From the wire tables, we see that this wire has a dc resistance of (2 in.) (6.386 Ω/1000 ft) = 1.06 milliohms. From the above formula, the cutoff frequency is found to be $124 / 40.3^2$ = 76 kHz. Since 100 MHz is roughly three decades above this (100 kHz to 100 MHz), the RF resistance will be approximately (1.06 mΩ) (10 × 3.2) = 34 mΩ. Again, values calculated in this manner are approximate and should be used qualitatively—that is, when you want an answer to a question such as, "Can I neglect the RF resistance of this length of connecting wire at 100 MHz?"

For additional information, see *Reference Data for Engineers*, Howard W. Sams & Co, Indianapolis, IN 46268. Chapter 6 contains a discussion and several design charts.

Effects on Q

Recall from the **AC Theory** portion of Chapter 4 that circuit Q, a useful figure of merit for tuned RLC circuits, can be defined in several ways:

$$Q = \frac{X_L \text{ or } X_C \text{ (at resonance)}}{R} \qquad (13)$$

$$= \frac{\text{energy stored per cycle}}{\text{energy dissipated per cycle}}$$

Q is also related to the bandwidth of a tuned circuit's response by

$$Q = \frac{f_0}{BW_{3\,dB}} \qquad (14)$$

Parasitic inductance, capacitance and resistance can significantly alter the performance and characteristics of a tuned circuit if the design frequency is anywhere near the self-resonant frequencies of the components.

As an example, consider the resonant circuit of **Fig 6.27A**, which could represent the input tank circuit of an oscillator. Neglecting any parasitics, $f_0 = 1 / (2 \pi (LC)^{0.5}) = 10.06$ MHz. As in many real cases, assume the resistance arises entirely from the inductor series resistance. The data sheet for the inductor specified a minimum Q of 30, so assuming Q = 30 yields an R value of $X_L / Q = 2 \pi$ (10.06 MHz) (5 μH) / 30 = 10.5 Ω.

Next, let's include the parasitic inductance of the capacitor (Fig 6.27B). A reasonable assumption is that this capacitor has the same physical size as the example from the Parasitic Inductance discussion above, for which we calculated L_s = 0.106 μH. This would give the capacitor a self-resonant frequency of 434 MHz—well above our area of interest. However, the added parasitic inductance does account for an extra 0.106/5.00 = 2% inductance. Since this circuit is no longer strictly series or parallel, we must convert it to an equivalent form before calculating the new f_0.

An easier and faster way is to *simulate* the altered circuit by computer. This analysis was performed on a desktop computer using *SPICE*, a standard circuit simulation program; for more details, see the Computer-Aided Design section below. The voltage response of the circuit (given an input current of 1 mA) was calculated as a function of frequency for both cases, with and without parasitics. The results are shown in the plot in Fig 6.27C, where we can see that the parasitic circuit has an f_0 of 9.96 MHz (a shift of 1%) and a Q (measured from the –3 dB points) of 31.5. For comparison, the simulation of the unaltered circuit does in fact show f_0 = 10.06 MHz and Q = 30.

Inductor Coupling

Mutual inductance will also have an effect on the resonant frequency and Q of the involved circuits. For this reason, inductors in frequency-critical circuits should always be shielded, either by constructing compartments for each circuit block or through the use of "can"-mounted coils. Another helpful technique is to mount nearby coils with their axes perpendicular as in **Fig 6.28**. This will minimize coupling.

As an example, assume we build an oscillator circuit that has both input and output filters similar to the resonant circuit in Fig 6.27A. If we are careful to keep the two coils in these circuits uncoupled, the frequency response of either of the two circuits is that of the solid line in **Fig 6.29**, reproduced from Fig 6.27C.

If the two coils are coupled either through careless placement or improper shielding, the resonant frequency and Q will be affected. The dashed line in Fig 6.29 shows the frequency response that results from a coupling coefficient of k = 0.05, a reasonable value for air-wound inductors mounted perpendicularly in close proximity on a circuit chassis. Note the resonant frequency shifted from 10.06 to 9.82 MHz, or 2.4%. The Q has gone up slightly from 30.0 to 30.8 as a result of the slightly higher inductive reactance at the resonant frequency.

To summarize, even small parasitics can significantly affect frequency responses of RF circuits. Either take steps to minimize or eliminate them, or use simple circuit theory to predict and anticipate changes.

Dielectric Breakdown and Arcing

Anyone who has ever watched a capacitor burn out, or heard the hiss of an arc across the inductor of an antenna tuner, while loading the 2-m vertical on 160 m, or touched a doorknob on a cold winter day has seen the effects of dielectric breakdown. When the dielectric is gaseous, especially air, we often call this *arcing*.

In the ideal world, we could take any two conductors and put as large a voltage as we want across them, no matter how close together they are. In the real world, there is a voltage limit (*dielectric strength*, measured in kV/cm and determined by the insulator between the two conductors) above which the insulator will break down.

Because they are charged particles, the electrons in the atoms of a dielectric material feel an attractive force when placed in an electric field. If the field is sufficiently strong, the force will strip the electron from the atom. This electron is available to conduct current, and furthermore, it is traveling at an extremely high velocity. It is very likely that this electron will hit another atom, and free another electron. Before long, there are many stripped electrons producing a large current. When this happens, we say the dielectric has suffered *breakdown*.

If the dielectric is liquid or gas, it can heal when the applied voltage is removed. A solid dielectric, however, cannot repair itself. A good example of this is a CMOS integrated circuit. When exposed to the very high voltages associated with static electricity, the electric field across the very thin gate oxide layer exceeds the dielectric strength of silicon dioxide, and the device is permanently damaged.

Capacitors are, by nature, perhaps the component most often associated with dielectric failure. To prevent damage, the working voltage of a capacitor—and there

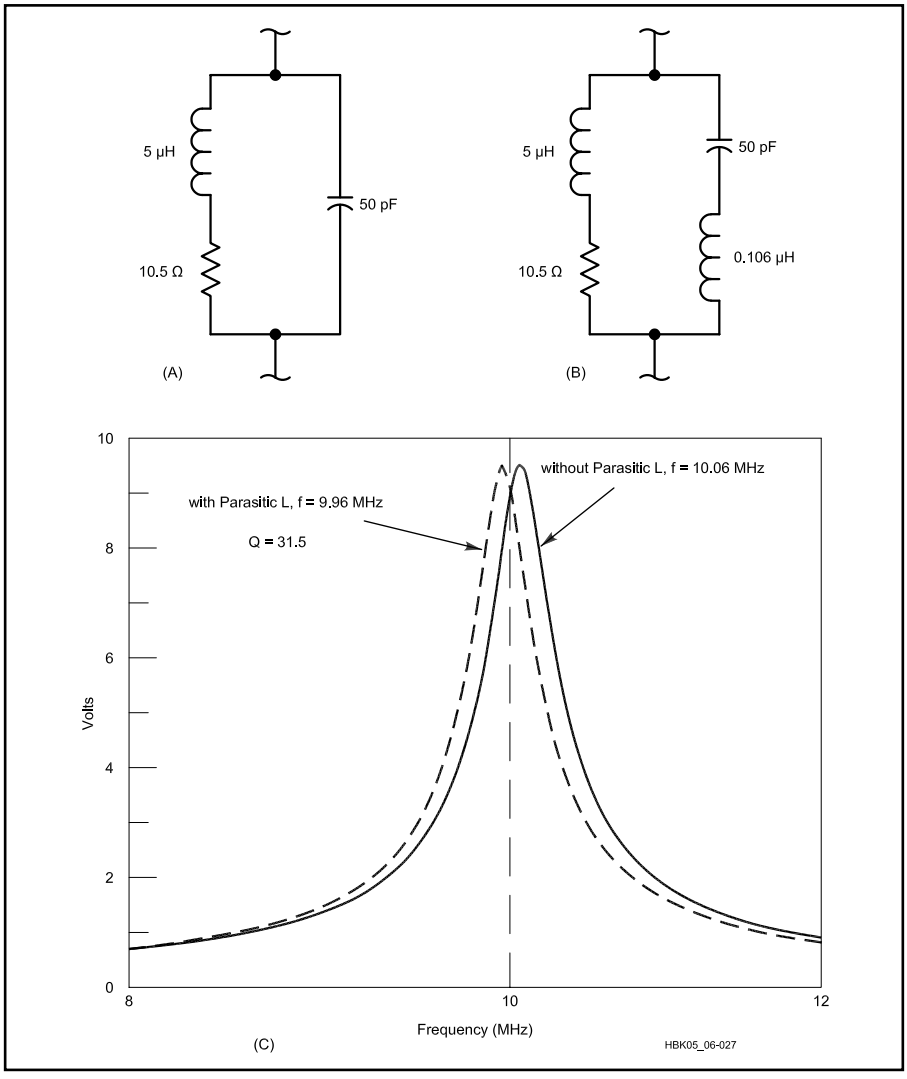


Fig 6.27—A is a tank circuit, neglecting parasitics. B same circuit including Lp on capacitor. C frequency response curves for A and B. The solid line represents the unaltered circuit (also see Fig 6.29) while the dashed line shows the effects of adding parasitic inductance.

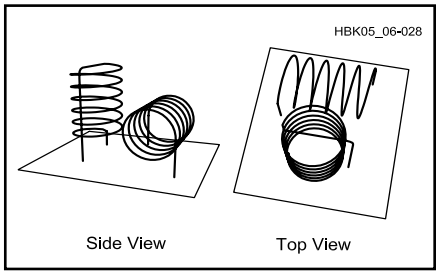


Fig 6.28—Unshielded coils in close proximity should be mounted perpendicular to each other to minimize coupling.

are separate dc and ac ratings—should ideally be 2 or 3 times the expected maximum voltage in the circuit.

Arcing is most often seen in RF circuits where the voltages are normally high, but it is possible anywhere two components at significantly different voltage levels are closely spaced.

The breakdown voltage of a dielectric layer depends on its composition and thickness (see **Table 6.3**). The variation with thickness is not linear; doubling the thickness does not quite double the breakdown voltage. Breakdown voltage is also a function of geometry: Because of electromagnetic considerations, the breakdown voltage between two conductors separated by a fixed distance is less if the surfaces are pointed or sharp-edged than if they are smooth or rounded. Therefore, a simple way to help prevent breakdown in many projects is to file and smooth the edges of conductors.

Radiative Losses and Coupling

Another consequence of Maxwell's equations states that any conductor placed in an electromagnetic field will have a current induced in it. We put this principle to good use when we make an antenna. The unwelcome side of this law of nature is the phrase "any conductor"; even conductors we don't intend as antennas will act this way.

Fortunately, the *efficiency* of such "antennas" varies with conductor length. They will be of importance only if their length is a significant fraction of a wavelength. When we make an antenna, we usually choose a length on the order of λ/2. Therefore, when we *don't* want an antenna, we should be sure that the conductor length is *much less* than λ/2, no more than 0.1 λ. This will ensure a very low-efficiency antenna. This is why 60-Hz power lines do not lose a significant fraction of the power they carry over long distances—at 60 Hz, 0.1 λ is about 300 miles!

In addition, we can use shielded cables. Such cables do allow some penetration of EM fields if the shield is not solid, but even 95% coverage is usually sufficient, especially if some sort of RF choke is used to reduce shield current.

Radiative losses and coupling can also be reduced by using twisted pairs of conductors—the fields tend to cancel. In some applications, such as audio cables, this may work better than shielding.

This argument also applies to large components, and remember that a com-ponent or long wire can *radiate* RF, as well as receive them. Critical stages such as tuned circuits should be placed in shielded compartments where possible. See the **EMI** and **Transmission Lines** chapters for more information.

REMEDIES FOR PARASITICS

The most common effect (always the most annoying) of parasitics is to influence the resonant frequency of a tuned circuit. This shift could cause an oscillator to fail, or more commonly, could cause a stable circuit to oscillate. It can also degrade filter performance (more on this later) and basically causing any number of frequency-related problems.

We can often reduce parasitic effects in discrete-component circuits by simply exploiting the models in Fig 6.26. Since parasitic inductance and loss resistance appear in series with a capacitor, we can reduce both by using several smaller capacitances in parallel, rather than one large one.

An example is shown in the circuit block in **Fig 6.30**, which is representative of the input tank circuit used in many HF VFOs. C_{main}, C_{trim}, C1 and C3 act with L to set the oscillator frequency. Therefore, temperature effects are critical in these components. By using several capacitors in

Table 6.3
Dielectric Constants and Breakdown Voltages

Material	*Dielectric Constant**	*Puncture Voltage***
Aisimag 196	5.7	240
Bakelite	4.4-5.4	240
Bakelite, mica filled	4.7	325-375
Cellulose acetate	3.3-3.9	250-600
Fiber	5-7.5	150-180
Formica	4.6-4.9	450
Glass, window	7.6-8	200-250
Glass, Pyrex	4.8	335
Mica, ruby	5.4	3800-5600
Mycalex	7.4	250
Paper, Royalgrey	3.0	200
Plexiglas	2.8	990
Polyethylene	2.3	1200
Polystyrene	2.6	500-700
Porcelain	5.1-5.9	40-100
Quartz, fused	3.8	1000
Steatite, low loss	5.8	150-315
Teflon	2.1	1000-2000

*At 1 MHz
**In volts per mil (0.001 inch)

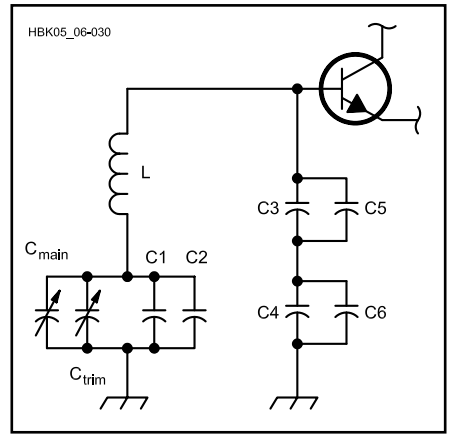


Fig 6.30—A tank circuit of the type commonly used in VFOs. Several capacitors are used in parallel to distribute the RF current, which reduces temperature effects.

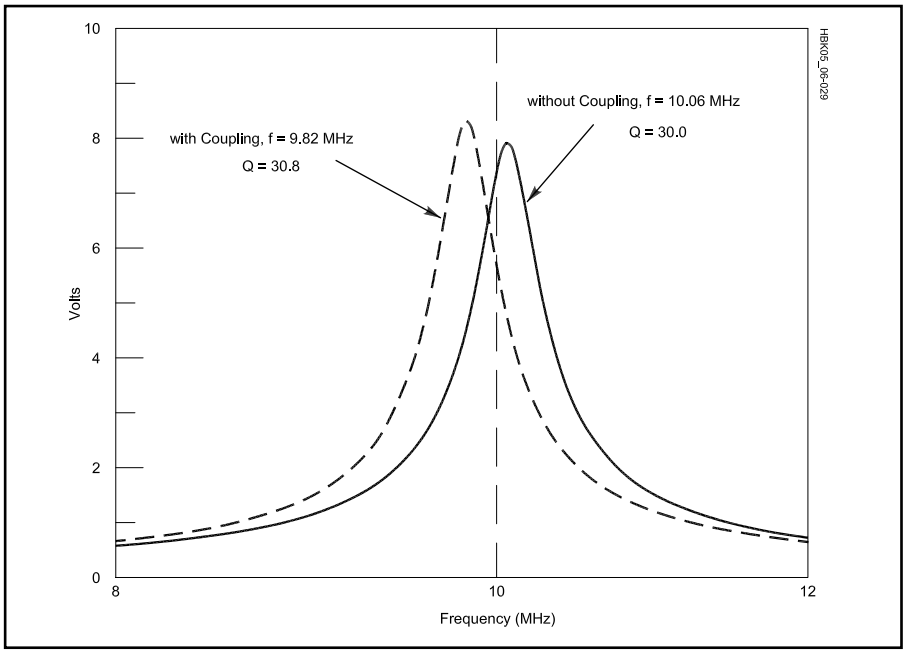


Fig 6.29—Result of light coupling (k=0.05) between two identical circuits of Fig 6.27A on their frequency responses.

parallel, the RF current (and resultant heat) is reduced in each component. Parallel combinations are used at the feedback capacitors for the same reason.

Another example of this technique is the common capacitor bypass arrangement shown in **Fig 6.31**. C1 provides a bypass path at audio frequencies, but it has a low self-resonant frequency due to its large capacitance. How low? Even if we assume the 0.02-µH value shown previously in this chapter, f_{sc} = 355 kHz. Adding C2 (with a smaller capacitance but much higher self-resonant frequency) in parallel provides bypass at high frequencies where C1 appears inductive.

Construction Techniques

These concepts also help explain why so many different components exist with similar values. As an example, assume you're working on a project that requires you to wind a 5 µH inductor. Looking at the coil inductance formula in the **AC Theory** section of Chapter 4, it comes to mind that many combinations of length and diameter could yield the desired inductance. If you happen to have both 0.5 and 1-inch coil forms, why should you select one over the other? To eliminate some other variables, let's make both coils 1 inch long, close-wound, and give them 1-inch leads on each end.

Let's calculate the number of turns required for each. On a 0.5-inch-diameter form:

$$n = \frac{\sqrt{5\left[(18\times 0.5)+(40\times 1)\right]}}{0.5} = 31.3 \text{ turns}$$

This means coil 1 will be made from #20 wire (29.9 turns per inch). Coil 2, on the 1-inch form, yields

$$n = \frac{\sqrt{5\left[(18\times 1)+(40\times 1)\right]}}{1} = 17.0 \text{ turns}$$

which requires #15 wire in order to be close-wound.

What are the series resistances associated with each? For coil 1, the total wire length is 2 inches + (31.3 × π × 0.5) = 51 inches, which at 10.1 Ω /1000 ft gives R_s = 0.043 Ω at dc. Coil 2 has a total wire length of 2 inches + (17.0 × π × 1) = 55 inches, which at 3.18 Ω /1000 ft gives a dc resistance of R_s = 0.015 Ω, or about 1/3 that of coil 1. Furthermore, at RF, coil 1 will begin to suffer from skin effect at a frequency about 3 times lower than coil 2 because of its smaller conductor diameter. Therefore, if Q were the sole consider-

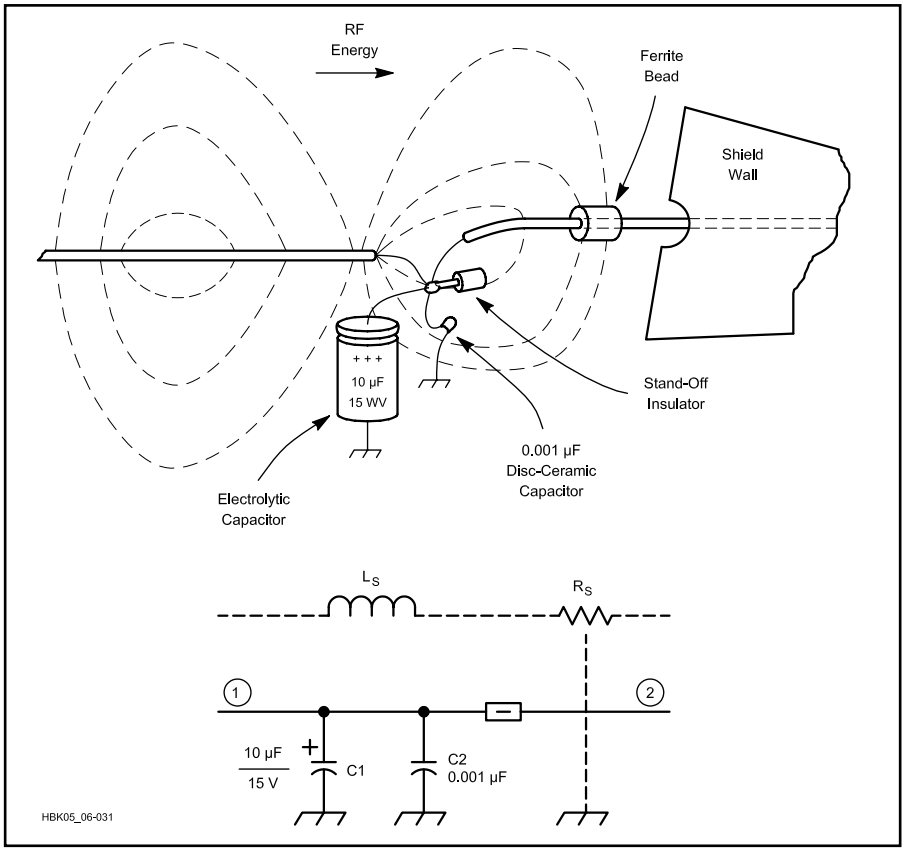


Fig 6.31—A typical method to provide bypassing at high frequencies when a large capacitor with a low-self-resonant frequency is required.

ation, it would be better to use the larger diameter coil.

Q is not the only concern, however. Such coils are often placed in shielded enclosures. Rule of thumb says the enclosure should be at least one coil diameter from the coil on all sides. That is, 3×3×2 inches for the large coil and 1.5×1.5×1.5 for the small coil, a volume difference of over 500%.

THERMAL CONSIDERATIONS

Any real energized circuit consumes electric power because any real circuit contains components that convert electricity into other forms of energy.[1] This *dissipated power* appears in many forms. For example, a loudspeaker converts electrical energy into the motion of air molecules we call sound. An antenna (or a light bulb) converts electricity into electromagnetic radiation. Charging a battery converts electrical energy into chemical energy (which is then converted back to electrical energy upon discharge). But the most common transformation by far is the conversion, through some form of *resistance*, of electricity into heat.

Sometimes the power lost to heat serves a useful purpose—toasters and hair dryers come to mind. But most of the time, this heat represents a power loss that is to be minimized wherever possible or at least taken into account. Since all real circuits contain resistance, even those circuits (such as a loudspeaker) whose primary purpose is to convert electricity to some *other* form of energy also convert some part of their input power to heat. Often, such losses are negligible, but sometimes they are not.

If unintended heat generation becomes significant, the involved components will get warm. Problems arise when the temperature increase affects circuit operation by either

- causing the component to fail, by explosion, melting, or other catastrophic event, or, more subtly,
- causing a slight change in the properties of the component, such as through a temperature coefficient (TC).

In the first case, we can design conservatively, ensuring that components are rated to safely handle two, three or more times the maximum power we expect them to dissipate. In the second case, we can specify components with low TCs, or we can design the circuit to minimize the effect of any one component. Occasionally we even exploit temperature effects (for example, using a resistor, capacitor or diode as a temperature sensor). Let's look more closely at the two main categories of thermal effects.

HEAT DISSIPATION

Not surprisingly, heat dissipation (more correctly, the efficient removal of generated heat) becomes important in medium- to high-power circuits: power supplies, transmitting circuits and so on. While these are not the only examples where elevated temperatures and related failures are of concern, the techniques we will discuss here are applicable to all circuits.

Thermal Resistance

The transfer of heat energy, and thus the change in temperature, between two ends of a block of material is governed by the following heat flow equation (see **Fig 6.32**):

$$P = \frac{k\,A}{L}\Delta T = \frac{\Delta T}{\theta} \qquad (15)$$

where

P = power (in the form of heat) conducted between the two points

k = *thermal conductivity*, measured in W/(m °C), of the material between the two points, which may be steel, silicon, copper, PC board material and so on

L = length of the block

A = area of the block

ΔT = *change* in temperature between the two points;

$\theta = \frac{L}{kA}$ is often called the *thermal resistance* and has units of °C/W.

Thermal conductivities of various common materials at room temperature are given in **Table 6.4.**

A very useful property of the above equation is that it is *exactly* analogous to Ohm's Law, and therefore the same principles and methods apply to heat flow problems as circuit problems. The following correspondences hold:

- Thermal conductivity W/(m °C) ↔ Electrical conductivity (S/m).
- Thermal resistance (°C/W) ↔ Electrical resistance (Ω).
- Thermal current (heat flow) (W) ↔ Electrical current (A).
- Thermal potential (T) ↔ Electrical potential (V).
- Heat source ↔ Current source.

For example, calculate the temperature of a 2-inch (0.05 m) long piece of #12 copper wire at the end that is being heated by a 25 W (input power) soldering iron, and whose other end is clamped to a large metal vise (assumed to be an infinite heat sink), if the ambient temperature is 25 °C (77 °F).

First, calculate the thermal resistance of the copper wire (diameter of #12 wire is 2.052 mm, cross-sectional area is 3.31×10^{-6} m^2)

$$\theta = \frac{L}{k\,A} = \frac{(0.05\ m)}{(390\ W/(m\,°C))(3.31\times10^{-6}\ m^2)}$$
$$= 38.7\ °C/W$$

Then, rearranging the heat flow equation above yields (after assuming the heat energy actually transferred to the wire is around 10 W)

$$\Delta T = P\,\theta = (10\ W)(38.7\ °C/W) = 387\ °C$$

So the wire temperature at the hot end is 25 C + ΔT = 412°C (or 774°F). If this sounds a little high, remember that this is for the steady state condition, where you've been holding the iron to the wire for a long time.

From this example, you can see that things can get very hot even with application of moderate power levels. In the case of a soldering iron, that's good, but in the case of a 25-W power transistor in a transmitter output stage, it's bad. For this reason, circuits that generate sufficient heat to alter, not necessarily damage, the components must employ some method of cooling, either active or passive. Passive methods include heat sinks or careful component layout for good ventilation. Active methods include forced air (fans) or some sort of liquid cooling (in some high-power transmitters).

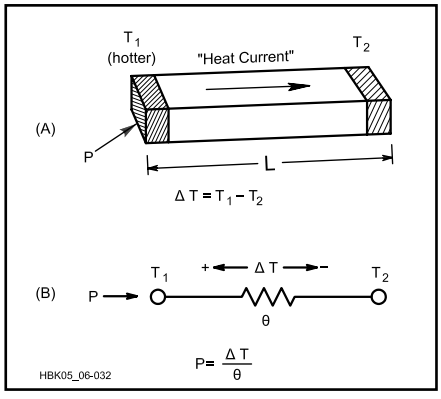


Fig 6.32—Physical and "circuit" models for the heat-flow equation.

Heat Sink Design and Use

The purpose of a heat sink is to provide a high-power component with a large surface area through which to dissipate heat. To use the models above, it provides a low thermal-resistance path to a cooler temperature, thus allowing the hot component to conduct a large "thermal current" away from itself.

Power supplies probably represent one of the most common high-power circuits amateurs are likely to encounter. Everyone has certainly noticed that power supplies get warm or even hot if not ventilated properly. Performing the thermal design for a properly cooled power supply is a very well-defined process and is a good illustration of heat-flow concepts.

A 28-V, 10-A power supply will be used for this design example. This material was originally prepared by ARRL Technical Advisor Dick Jansson, WD4FAB, and the steps described below were actually followed during the design of that supply.

An outline of the design procedure shows the logic applied:

1. Determine the expected power dissipation (P_{in}).
2. Identify the requirements for the dissipating elements (maximum component temperature).
3. Estimate heat-sink requirements.
4. Rework the electronic device (if necessary) to meet the thermal requirements.
5. Select the heat exchanger (from heat sink data sheets).

The first step is to estimate the filtered, unregulated supply voltage under full load. Since the transformer secondary output is 32 V ac (RMS) and feeds a full-wave bridge rectifier, let's estimate 40 V as the filtered dc output at a 10-A load.

Table 6.4

Thermal Conductivities of Various Materials

Gases at 0°C, Others at 25°C; from *Physics,* by Halliday and Resnick, 3rd Ed.

Material	$\left(\text{k in units of } \frac{W}{m\,°C}\right)$
Aluminum	200
Brass	110
Copper	390
Lead	35
Silver	410
Steel	46
Silicon	150
Air	0.024
Glass	0.8
Wood	0.08

The next step is to determine our critical components and estimate their power dissipations. In a regulated power supply, the pass transistors are responsible for nearly all the power lost to heat. Under full load, and allowing for some small voltage drops in the power-transistor emitter circuitry, the output of the series pass transistors is about 29 V for a delivered 28 V under a 10-A load. With an unregulated input voltage of 40 V, the total energy heat dissipated in the pass transistors is (40 V – 29 V) × 10 A = 110 W. The heat sink for this power supply must be able to handle that amount of dissipation and still keep the transistor junctions below the specified safe operating temperature limits. It is a good rule of thumb to select a transistor that has a maximum power dissipation of twice the desired output power.

Now, consider the ratings of the pass transistors to be used. This supply calls for 2N3055s as pass transistors. The data sheet shows that a 2N3055 is rated for 15-A service and 115-W dissipation. But the design uses *four* in parallel. Why? Here we must look past the big, bold type at the top of the data sheet to such subtle characteristics as the junction-to-case thermal resistance, q_{jc}, and the maximum allowable junction temperature, T_j.

The 2N3055 data sheet shows θ_{jc} = 1.52°C/W, and a maximum allowable case (and junction) temperature of 220°C. While it seems that one 2N3055 could barely, on paper at least, handle the electrical requirements, at what temperature would it operate?

To answer that, we must model the entire "thermal circuit" of operation, starting with the transistor junction on one end and ending at some point with the ambient air. A reasonable model is shown in **Fig 6.33**. The ambient air is considered here as an infinite heat sink; that is, its temperature is assumed to be a constant 25°C (77°F). θ_{jc} is the thermal resistance from the transistor junction to its case. θ_{cs} is the resistance of the mounting interface between the transistor case and the heat sink. θ_{sa} is the thermal resistance between the heat sink and the ambient air. In this "circuit," the generation of heat (the "thermal current source") occurs in the transistor at P_{in}.

Proper mounting of most TO-3 package power transistors such as the 2N3055 requires that they have an electrical insulator between the transistor case and the heat sink. However, this electrical insulator must at the same time exhibit a low thermal resistance. To achieve a quality mounting, use thin polyimid or mica formed washers and a suitable thermal compound to exclude air from the intersti-

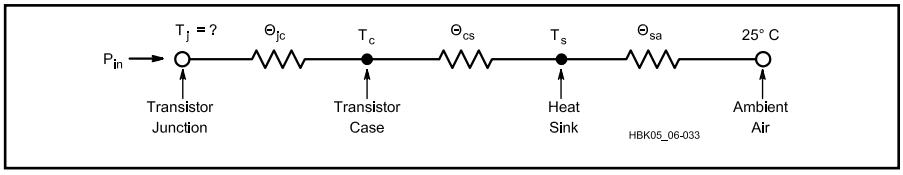


Fig 6.33—Resistive model of thermal conduction in a power transistor and associated heat sink. See text for calculations.

tial space. "Thermal greases" are commonly available for this function. Any silicone grease may be used, but filled silicone oils made specifically for this purpose are better.

Using such techniques, a conservatively high value for θ_{cs} is 0.50°C/W. Lower values are possible, but the techniques needed to achieve them are expensive and not generally available to the average amateur. Furthermore, this value of θ_{cs} is already much lower than θ_{jc}, which cannot be lowered without going to a somewhat more exotic pass transistor.

Finally, we need an estimate of θ_{sa}. **Fig 6.34** shows the relationship of heat-sink volume to thermal resistance for natural-convection cooling. This relationship presumes the use of suitably spaced fins (0.35 inch or greater) and provides a "rough order-of-magnitude" value for sizing a heat sink. For a first calculation, let's assume a heat sink of roughly 6×4×2 inch (48 cubic inches). From Fig 6.34, this yields a θ_{sa} of about 1°C/W.

Returning to Fig 6.33, we can now calculate the approximate temperature increase of a single 2N3055:

$$\delta T = P\,\theta_{total} = (110\,W) \times$$
$$(1.52\,°C/W + 0.5\,°C/W + 1.0\,°C/W) = 332\,°C$$

Given the ambient temperature of 25°C, this puts the junction temperature T_j of the 2N3055 at 25 + 332 = 357°C! This is clearly too high, so let's work backward from the air end and calculate just how many transistors we need to handle the heat.

First, putting more 2N3055s in parallel means that we will have the thermal model illustrated in **Fig 6.35**, with several identical θ_{jc} and θ_{cs} in parallel, all funneled through the same θ_{as} (we have one heat sink).

Keeping in mind the physical size of the project, we could comfortably fit a heat sink of approximately 120 cubic inches (6×5×4 inches), well within the range of commercially available heat sinks. Furthermore, this application can use a heat sink where only "wire access" to the transistor connections is required. This allows the selection of a more efficient design. In contrast, RF designs require the transistor mounting surface to be completely exposed so that the PC board can be mounted close to the transistors to minimize parasitics. Looking at Fig 6.34, we see that a 120-cubic-inch heat sink yields a θ_{sa} of 0.55°C/W. This means that the temperature of the heat sink when dissipating 110 W will be 25°C + (110 W) (0.55°C/W) = 85.5°C.

Industrial experience has shown that silicon transistors suffer substantial failure when junctions are operated at highly elevated temperatures. Most commercial and military specifications will usually not permit design junction temperatures to exceed 125°C. To arrive at a safe figure for our maximum allowed T_j, we must consider the intended use of the power supply. If we are using it in a 100% duty-cycle transmitting application such as RTTY or FM, the circuit will be dissipating 110 W continuously. For a lighter duty-cycle load such as CW or SSB, the "key-down" temperature can be slightly higher as long as the average is less than 125°C. In this intermittent type of service, a good conservative figure to use is T_j = 150°C.

Given this scenario, the temperature rise across each transistor can be 150 – 85.5 = 64.5°C. Now, referencing Fig 6.35, remembering the total θ for each 2N3055 is 1.52 + 0.5 = 2.02°C/W, we can calculate the maximum power each 2N3055 can safely dissipate:

$$P = \frac{\delta T}{\theta} = \frac{64.5\,°C}{2.02\,°C/W} = 31.9\,W$$

Thus, for 110 W full load, we need four 2N3055s to meet the thermal requirements of the design. Now comes the big question: What is the "right" heat sink to use? We have already established its requirements: it must be capable of dissipating 110 W, and have a θ_{sa} of 0.55°C/W (see above).

A quick consultation with several manufacturer's catalogs reveals that Wakefield Thermal Solutions, Inc. model nos. 441 and 435 heat sinks meet the needs of this application.[2] A Thermalloy model no. 6441 is suitable as well. Data published in the catalogs of these manufacturers show that in natural-convection service, the expected temperature rise for 100 W dissipation would be just under 60°C, an almost perfect fit for this application. Moreover, the no. 441 heat sink can easily mount four TO-3-style 2N3055 transistors. See **Fig 6.36**. Remember: heat sinks should be mounted with the fins and transistor mounting area vertical to promote convection cooling.

The design procedure just described is applicable to any circuit where heat buildup is a potential problem. By using the thermal-resistance model, we can easily calcu-

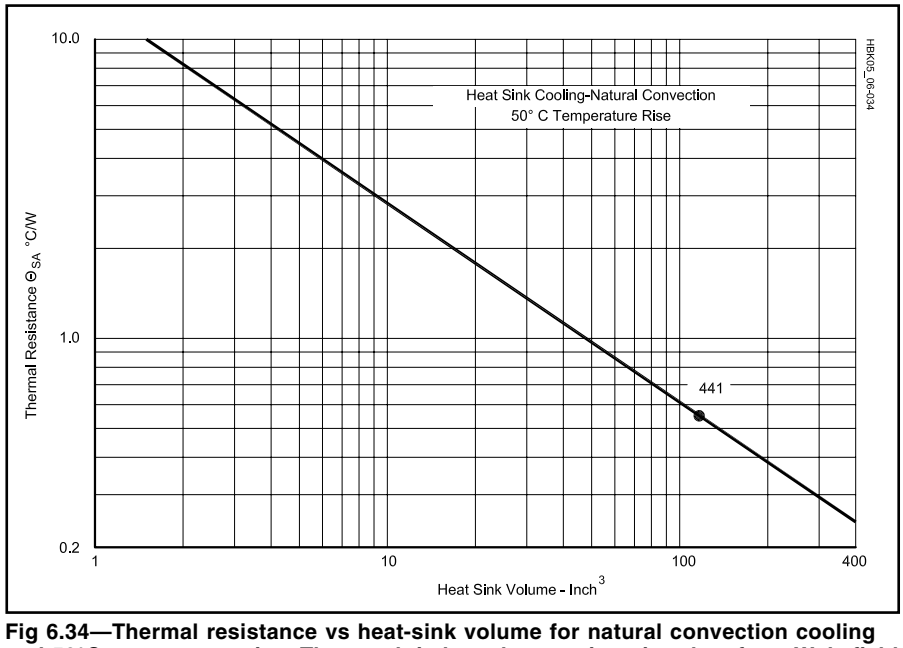


Fig 6.34—Thermal resistance vs heat-sink volume for natural convection cooling and 50°C temperature rise. The graph is based on engineering data from Wakefield Thermal Solutions, Inc.

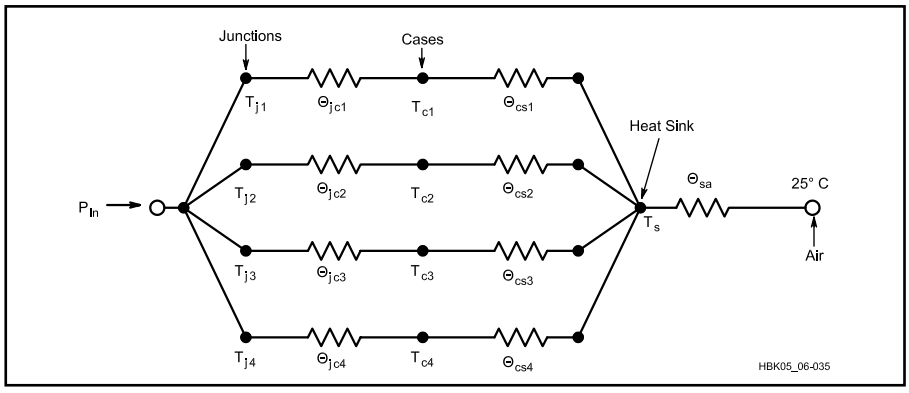


Fig 6.35—Thermal model for multiple power transistors mounted on a common heat sink.

late whether or not an external means of cooling is necessary, and if so, how to choose it. Aside from heat sinks, forced air cooling (fans) is another common method. In commercial transceivers, heat sinks with forced-air cooling are common.

Transistor Derating

Maximum ratings for power transistors are usually based on a case temperature of 25°C. These ratings will decrease with increasing operating temperature. Manufacturer's data sheets usually specify a *derating* figure or curve that indicates how the maximum ratings change per degree rise in temperature. If such information is not available (or even if it is!), it is a good rule of thumb to select a power transistor with a maximum power dissipation of at least twice the desired output power.

Rectifiers

Diodes are physically quite small, and they operate at high current densities. As a result their heat-handling capabilities are somewhat limited. Normally, this is not a problem in high-voltage, low-current supplies. The use of high-current (2 A or greater) rectifiers at or near their maximum ratings, however, requires some form of heat sinking. Frequently, mounting the rectifier on the main chassis (directly, or with thin mica insulating washers) will suffice. If the diode is insulated from the chassis, thin layers of silicone grease should be used to ensure good heat conduction. Large, high-current rectifiers often require special heat sinks to maintain a safe operating temperature. Forced-air cooling is sometimes used as a further aid.

Forced-Air Cooling

In Amateur Radio today, forced-air cooling is most commonly found in vacuum-tube circuits or in power supplies built in small enclosures, such as those in solid-state transceivers or computers. Fans or blowers are commonly specified in cubic feet per minute (CFM). While the nomenclature and specifications differ from those used for heat sinks, the idea remains the same: to offer a low thermal resistance between the inside of the enclosure and the (ambient) exterior.

For forced air cooling, we basically use the "one resistor" thermal model of Fig 6.32. The important quantity to be determined is heat generation, P_{in}. For a power supply, this can be easily estimated as the difference between the input power, measured at the transformer primary, and the output power at full load. For variable-voltage supplies, the worst-case output condition is minimum voltage with maximum current.

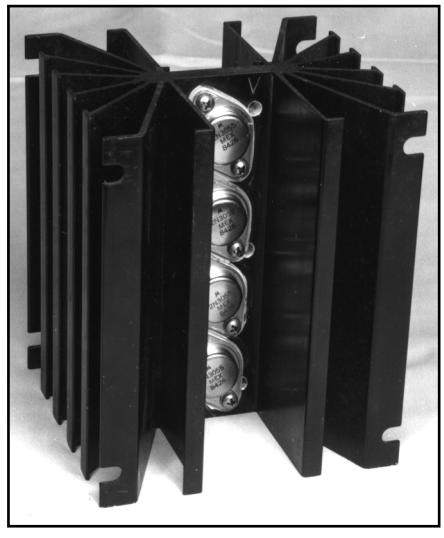
A discussion of forced-air tube cooling appears in the **RF Power Amplifiers** chapter.

TEMPERATURE STABILITY

Aside from catastrophic failure, temperature changes may also adversely affect circuits if the TCs of one or more components is too large. If the resultant change is not too critical, adequate temperature stability can often be achieved simply by using higher-precision components with low TCs (such as NP0/C0G capacitors or metal-film resistors). For applications where this is impractical or impossible (such as many solid-state circuits), we can minimize temperature sensitivity by *compensation* or *matching*—using temperature coefficients to our advantage.

Compensation is accomplished in one of two ways. If we wish to keep a certain circuit quantity constant, we can interconnect pairs of components that have equal but opposite TCs. For example, a resistor with a negative TC can be placed in series with a positive TC resistor to keep the total resistance constant. Conversely, if the important point is to keep the *difference* between two quantities constant, we can use components with the *same* TC so that the pair "tracks." That is, they both change by the same amount with temperature.

An example of this is a Zener reference circuit. Since the I-V equation for a diode is strongly affected by operating temperature, circuits that use diodes or transistors to generate stable reference voltages must use some form of temperature compensation. Since, for a constant current, a



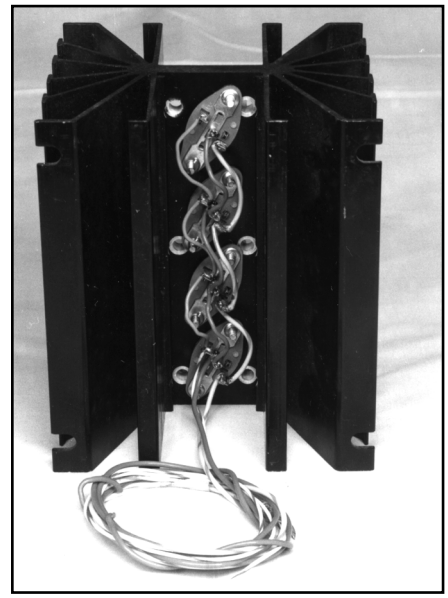


Fig 6.36—A Wakefield 441 heat sink with four 2N3055 transistors mounted.

The Thermistor in Homebrew Projects

Thermistors can be used to enhance project performance. Circuit temperature variations affect gain, distortion as well as control functions like receiver AGC or transmitter ALC. Thermistors can compensate for temperature changes. This greatly reduces the dangers of self destruction of overheated power transistors. It also can help reduce oscillator drift. In this sidebar, William Sabin, WØIYH describes how a simple temperature control circuit can be used to protect a MOSFET power amplifier or produce a temperature-stable heater to regulate the temperature of a VFO.

A thermistor is a small bit of intrinsic (no N or P doping) metal-oxide semiconductor compound material between two wire leads. As temperature increases, the number of liberated hole/electron pairs increases exponentially, causing the resistance to decrease exponentially. You can see this in the resistance equation:

$$R(T) = R(T0)e^{-\beta\left(\frac{1}{T0}-\frac{1}{T}\right)} \qquad \text{(Eq 1)}$$

where T is some temperature in kelvins and T0 is a reference temperature, usually 298 K (25°C), at which the manufacturer specifies R(T0). The constant β is experimentally determined by measuring resistance at various temperatures and finding the value of β that best agrees with the measurements. A simple way to get an approximate value of β is to make two measurements, one at room temperature, say 25°C (298 K) and one at 100°C (373 K) in boiling water. Suppose the resistances are 10 kΩ and 938 Ω. Eq 1 is solved for β

$$\beta = \frac{\ln\left(\frac{R(T)}{R(T0)}\right)}{\frac{1}{T}-\frac{1}{T0}} = \frac{\ln\left(\frac{938}{1000}\right)}{\frac{1}{373}-\frac{1}{298}} = 3507 \qquad \text{(Eq 2)}$$

For many electronics applications, the exact value of temperature is not as important as the ability to maintain that temperature. The following examples illustrate some thermistor circuit designs.

MOSFET Power Transistor Protector

It is very desirable to compensate the temperature sensitivity of power transistors. In bipolar (BJT) transistors, thermal runaway occurs because the dc current gain increases as the transistors get hotter. In MOSFET transistors, thermal runaway is less likely to occur. With excessive drain dissipation or inadequate cooling, however, the junction temperature may increase until its maximum allowable value is exceeded. You can place a thermistor on the heat sink close to the transistors so that the bias adjustment tracks the flange temperature. Another good idea is to attach a thermistor to the ceramic case of the FET with a small drop of epoxy. That way it will respond more quickly to a sudden temperature increase, possibly saving the transistor from destruction.

The circuit of **Fig A** uses a thermistor to detect a case temperature of about 93°C, with an ON/OFF operating range of about 0.3°C. A red LED warns of an overtemperature condition that requires attention. The LM339 comparator toggles when Rt = R5. Resistors R3, R4, R5 and the thermistor form a Wheatstone bridge. R3 and R4 are chosen to make the inputs to the LM339 about 0.5 V at the desired temperature. This greatly reduces self-heating of the thermistor, which could cause a substantial error in the circuit behavior.

The RadioShack Precision Thermistor RS 271-110 used in this circuit is rated at 10 kΩ ±1% at 25°C. It comes with a calibration chart from –50°C to +110°C that you can use to get an approximate resistance at any temperature in that range. For many temperature protection applications you don't have to know the exact temperature. You won't need an exact calibration, but you can verify that the thermistor is working properly by measuring its resistance at 20°C (68°F) and in boiling water (≈100°C).

As one example of how you can use this circuit, suppose you want to protect a power amplifier that uses

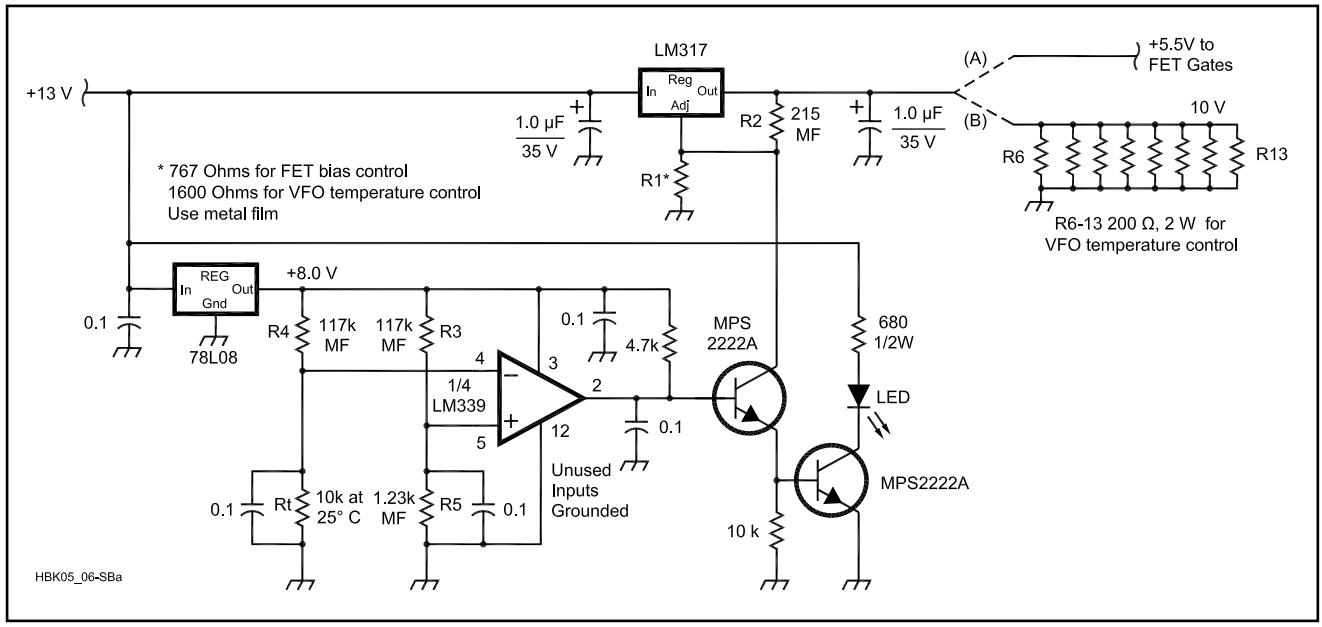


Fig A—Temperature controller for PA (A) or VFO (B).

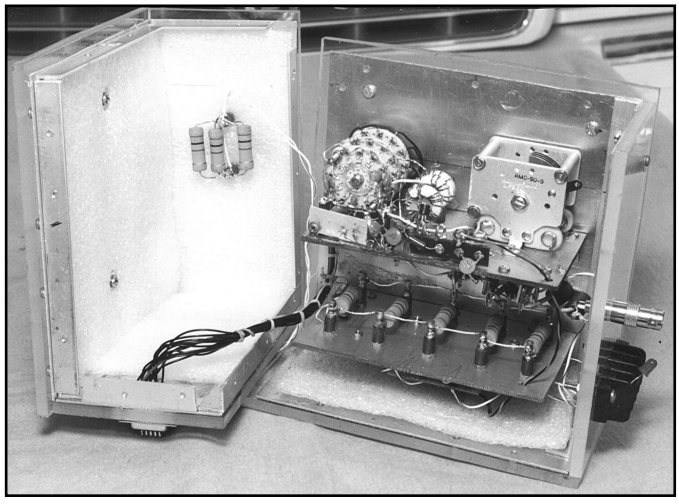


Fig B—Temperature controlled VFO cabinet. 1/4-inch plexiglass on the outside and 1/4-inch styrofoam panels on the inside are used to improve temperature stability.

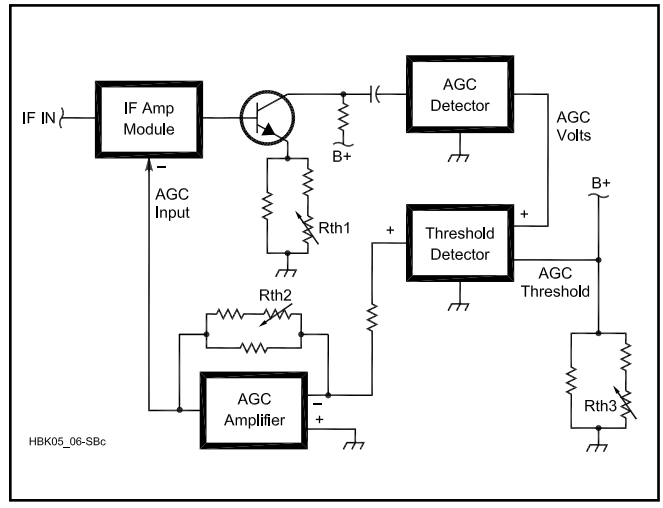


Fig C—Temperature compensation of IF amplifier and AGC circuitry.

a pair of MRF150 MOSFETs. Use the following procedure to get the desired temperature control:

The MRF150 MOSFET has a maximum allowed junction temperature of 200°C. The thermal resistance θ_{JC} from junction to case is 0.6°C per watt. The maximum expected dissipation of the FET in normal operation is 110 W. Select
a case temperature of 93°C. This makes the junction temperature 93°C + (0.6°C/W) × (110 W) =159°C, which is a safe 41°C below the maximum allowed temperature.

The FET has a rating of 300 W maximum dissipation at a case temperature of 25°C, derated at 1.71 W per °C. At 93°C case temperature the maximum allowed dissipation is 300 W – 1.71 W/°C × (93°C – 25°C) = 184 W. The safety margin *at that temperature* is 184 – 110 = 74 W.

A very simple way to determine the correct value of R5 is to put the thermistor in 93°C water (let it stabilize) and adjust R5 so that the circuit toggles. At 93°C, the measured value of the thermistor should be about 1230 Ω.

To use the circuit of Fig A to control the FET bias voltage, set R1 to be 767 Ω. This value, along with R2 adjusts the LM317 voltage regulator output to 5.5V for the FET gate bias. When the thermistor heats to the set point, the LM339 comparator toggles on. This brings the FET gate voltage to a low level and completely turns off the FETs until the temperature drops about 0.3°C. Use metal film resistors throughout the circuit.

Temperature controlled variable frequency oscillator (VFO)

The same circuit can be used to control the temperature of a VFO (variable frequency oscillator) with a few simple changes. Select R1 to be 1600 Ω (metal film) to adjust the LM317 for 10 V. Add resistors R6 through R13 as a heater element. Place the VFO in a thermally insulated enclosure such as shown in **Fig B**. The eight 200-Ω, 2-W, metal-oxide resistors at the output of the LM317 supply about 4 W to maintain a temperature of about 33°C inside the enclosure. The resistors are placed so that heat is distributed uniformly. Five are placed near the bottom and three near the top. The thermistor is mounted in the center of the box, close to the tuned circuit and in physical contact with the oscillator groundplane surface, using a small drop of epoxy.

Use a massive and well-insulated enclosure to slow the rate of temperature change. In one project, over a 0.1°C range, the frequency at 5.0 MHz varied up and down ±20 Hz or less, with a period of about five minutes. Superimposed is a very slow drift of average frequency that is due to settling down of components, including possibly the thermistor. These gradual changes became negligible after a few days of "burn-in." One problem that is virtually eliminated by the constant temperature is a small but significant "retrace" effect of the cores and capacitors, and perhaps also the thermistor, where a substantial temperature transient of some kind may take from minutes to hours to recover the previous L and C values.

For a much more detailed discussion of this material, see Sabin W.E., WØIYH "Thermistors in Homebrew Projects," *QEX* Nov/Dec 2000. This article is included (with all *QST*, *QEX* and *NCJ* articles from 2000) on the *2000 ARRL Periodicals* CD-ROM. (Order No. 8209.)

reverse-biased PN junction has a negative voltage TC while a forward-biased junction has a positive voltage TC, a good way to temperature-compensate a Zener reference diode is to place one or more forward-biased diodes in series with it.

RF Heating

RF current often causes component heating problems where the same level of dc current may not. An example is the tank circuit of an RF oscillator. If several small capacitors are connected in parallel to achieve a desired capacitance, skin effect will be reduced and the total surface area available for heat dissipation will be increased, thus significantly reducing the RF heating effects as compared to a single large capacitor. This technique can be applied to any similar situation; the general idea is to divide the heating among as many components as possible.

CAD TOOLS FOR CIRCUIT DESIGN

Hams today enjoy the easy availability of tremendous computing power to aid them in their hobby. A commercial appli-

cation that computers perform very well is circuit simulation. Today anyone with access to a computer software bulletin board is likely to find shareware or other inexpensive software to perform such analysis. ARRL used to offer *ARRL Radio Designer,* discussed in October 1994 *QST* and a subsequent *QST* column, "Exploring RF." Another good tool is *Ansoft Designer SV* (**www.ansoft.com**).

The advantages of computer circuit simulation over pencil-and-paper calculation are many. For example, while the analysis of a small circuit, say 5 to 10 components, may be recomputed without too much effort if the design changes, the same is not true for a 50 to 100-component circuit. Indeed, the larger circuit may require too many simplifying assumptions to calculate by hand at all. It is much faster to watch and adjust the response of a circuit on a computer screen than to breadboard the circuit in search of the same information.

All circuit simulators, including *SPICE*, work from a *netlist* that is simply a component-by-component list of the circuit that tells the program how components are interconnected. The program then uses standard techniques of numerical analysis and matrix mathematics to calculate the analysis you wish to perform. Common analyses include dc and ac bias and operating point calculations, frequency response curves, transient analysis (looking at waveforms in the circuit over a specified period of time), Fourier transforms, transfer functions, two-port parameter calculations, and pole-zero analysis. Many packages also perform sensitivity analyses: calculating what happens to the voltages and currents as you sweep the value(s) of an individual component (or group of components), and statistical analyses: determining the variational "window" of the system response that would result if all component values randomly varied by a given tolerance (also known as *Monte Carlo* analysis).

Beware!

Before we turn to a brief introduction to circuit simulation for the amateur, a *caveat* to the reader is definitely in order: However fast and powerful the computer may seem, keep in mind that it is only another tool in your workshop, just like your 'scope or soldering iron. Circuit simulations are only meaningful if you can interpret the results correctly, and a good initial circuit design based on real experience and common sense is mandatory. *SPICE* and other programs have no problem with a bench-top power supply delivering 10,000 V to a 5-W resistor because you made a mistake when specifying the circuit. Software also won't remind you that the resistor better be rated for 20 megawatts! Remember, the real power behind any simulation software is *your mind*.

SPICE

The electronics industry has been the main consumer of circuit simulation packages ever since the development of *SPICE* (or *S*imulation *P*rogram with *I*ntegrated *C*ircuit *E*mphasis) on mainframe systems in the mid 1970s. While many companies today produce other simulation software, *SPICE* remains the *de facto* industry standard. *SPICE* itself now exists in many "flavors", and specially tailored versions for desktop computers are sold by several companies, including *PSPICE* by Microsim Corp and *HSPICE* by Intusoft. In fact, an "evaluation" copy of *PSPICE* has been placed in the public domain by Microsim and is probably the easiest and most inexpensive way for amateurs to begin circuit simulation. It may be obtained (with a companion reference book) from technical bookstores, and also from many sites on the Internet.

A Basic Circuit

Consider the simple RC low-pass filter in **Fig 6.37A**. To simulate this circuit, we must first precisely describe it. Convention sets ground as *node* 0, and we label all other nodes (a node is a place where two or more components meet) with numbers, as shown in the figure. We give each element a unique name, and then prepare a *netlist*, listing each element with the numbers of its "+" and "–" nodes (*SPICE* assumes the ground node to be labeled 0), and its value, as shown at B in the figure.

The question we wish to answer by simulation is the following: what is the frequency response of this filter? Of course, for this simple problem we can calculate the answer by hand. The cutoff frequency of a simple RC filter is given by:

$$f_{co} = \frac{1}{2\pi R C} \qquad (17)$$

where

f_{co} = cutoff frequency (–3 dB)
R = resistance, in ohms
C = capacitance, in farads.

In this case f_{co} is 1.59 kHz. Beyond this, with one storage element (the capacitor) we expect the output to decrease by 20 dB for each decade of frequency.

Fig 6.38 shows the graph of the ratio of output to input voltages V_{out}/V_{in}, in dB for that frequency range as calculated by *SPICE*. Note the 3-dB point is indeed 1.59 kHz, and above that frequency the response drops by 20-dB/decade.

At this point we could use this circuit to ask some "What if?" questions, but for a circuit this simple we could answer those questions with pencil and paper. Below is a more extensive design to show the power of simulation, where "What if?" questions and the interaction of numerous components are much more easily and clearly examined with a computer.

A Case Study: Amplifier Distortion

An application where computer circuit simulation does a much faster, easier, and more accurate job than pencil and paper is Fourier analysis. Recall that any periodic signal can be broken down into a sum of sine waves of different amplitudes whose frequencies are multiples of the fundamental frequency of the signal. Performing such an analysis allows us to identify how much signal power is contained in the fundamental frequency and its harmonics. A Fourier *transform* is the representation of a signal by its frequency components, and resembles what you would see if you fed the signal into a spectrum analyzer.

Consider the amplifier circuit in **Fig 6.39**, which is a simple common-emitter audio amplifier. The frequency response of its voltage gain, calculated from *SPICE*, is shown in **Fig 6.40**. Pay particular attention to the voltage gain at 1 kHz, which is 148, or 43.4 dB. If we place a small ac signal on the input as shown in Fig 6.39, the output will be a faithfully amplified sine wave. However, the amplifier is powered from a 9-V supply, so when the output reaches this level, the waveform will begin to show signs of clipping. For a 1-kHz signal, this will occur at roughly V_{in} = 4.5/148 = 0.03 V amplitude.

Total Harmonic Distortion is a common figure of merit for audio amplifiers, and is defined as

$$THD = \frac{P_H}{P_F} \times 100 \qquad (18)$$

where

THD = total harmonic distortion, in percent
P_H = total power in all harmonics above the fundamental
P_F = power in fundamental.

SPICE readily calculates the Fourier content of a signal and also THD. Assume that we will tolerate 5% THD for our design. What is the maximum swing we can allow the input signal?

Fig 6.41 shows the output waveforms that *SPICE* calculates for our circuit with input signals of 0.001, 0.003, 0.01, 0.03

and 0.1-V amplitude. Note that, as expected, the last two input voltages show definite signs of clipping. What is not so evident is that the smaller inputs have some distortion as well, due to the small but finite nonlinearity of the amplifier even at those voltages.

Fig 6.42 shows the relative amplitudes of the harmonic content (Fourier transform) of the Fig 6.41 waveforms, all normalized to fundamental = 100%. You can see how the higher harmonics grow in strength as the signal increases, especially when the clipping starts. The THD for each case is also shown, and from this simulation we see we must limit our input to somewhere between 0.003 and 0.01 V. We would pin this down more closely by running another simulation around this "window."

A caution is in order here. Circuit models have limits just as do circuits. The models apply only over a limited dynamic range. That's why we have both small- and large-signal models. One challenge of CAD is determining the limits of your models. There is more information about this in the sections on modeling transistors.

Conclusion

Start using your computer for more challenging tasks than QSL bookkeeping! Just like any other piece of software, you'll find yourself using *SPICE* to do things you never thought of trying before. It can even do digital circuits, transmission lines and antennas. Remember that a computer is only as smart as the person using it! Have fun!

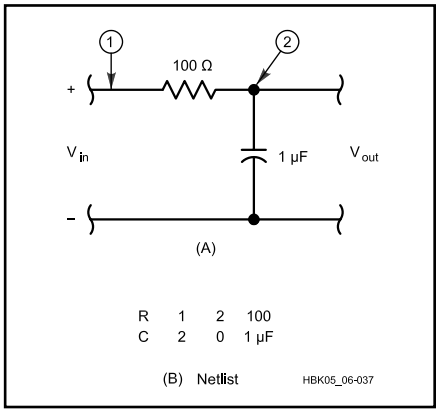


Fig 6.37—A, a simple RC low-pass filter.

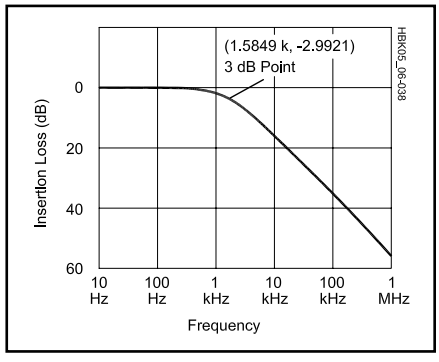


Fig 6.38—Low-pass filter output simulation.

References

J. Carr, *Secrets of RF Circuit Design*, Tab Books, 1991.

D. DeMaw, *Practical RF Design Manual*, Prentice-Hall, 1982.

R. Dorf, Ed., *The Electrical Engineering Handbook*, CRC Press and IEEE Press, 1993.

W. Hayward, *Introduction to RF Design*, ARRL, 1994.

P. Tuinenga, *SPICE: A Guide to Circuit Simulation and Analysis using PSPICE*, by Prentice-Hall, ISBN 0-13-747270-6. This is a wonderful PSPICE reference manual that comes with copy of the software. It is available at most university or technical bookstores.

A. Vladimirescu, *The SPICE Book*, Wiley & Sons, 1994.

D. Pederson and K. Mayaram, *Analog Integrated Circuits for Communication: Principles, Simulation and Design, 1991*, Kluwer Academic Publishers. Pederson is one of the inventors of SPICE. This book is about SPICE simulation.

J. White, Thermal Design of Transistor Circuits," April 1972 *QST*, pp 30-34.

LOW-FREQUENCY TRANSISTOR MODELS

The Fundamental Equations

Design models are based on the physics of the components we use. The complexity of models can vary widely though, and many times we can use very simple models to achieve our goals. Increasingly complex models are developed and used only when demanded by the circuit application.

In this discussion, we will focus on simple models for bipolar transistors (BJTs) and FETs. These models are reasonably accurate at low frequencies, and they are of some use at RF. For more sophisticated RF models, look to professional RF-design literature. This discussion is adapted from Wes Hayward's *Introduction to RF Design*. Derivations of the material shown here appear in that book, an excellent text for the beginning RF designer.

This discussion is centered on NPN BJTs and N-channel JFETs. The material here applies to PNPs and P-channel FETs when you simply change the bias polarities.

First, consider the bipolar transistor as a current controlled device. When the base current controls the collector current, this equation defines the transistor operation

$$I_c = \beta I_b \tag{19}$$

where β = common-emitter current gain.

When we consider a transistor as a voltage-controlled device, this equation describes it (emitter current, I_e, in terms of base-emitter voltage, V_{be}):

$$\begin{aligned} I_e &= I_{es}\left[\exp\left(qV/kT\right)-1\right] \\ &\approx I_{es}\exp\left(qV/kT\right) \end{aligned} \tag{20}$$

where

$V = V_{be}$
q = electronic charge
k = Boltzmann's constant
T = temperature in kelvins (K)
I_{es} = emitter saturation current, typically 1×10^{-13} A.

Both equations approximate models of more complex behavior. Equation 20 is a simplification of the first Ebers-Moll model (Ref 1). More sophisticated models for BJTs are described by Getreu in Ref 2. Equations 19 and 20 apply to both transistor dc biasing and signal design.

The operation of an N-channel JFET can be characterized by

$$I_D = I_{DSS}\left(1 - V_{sg}/V_p\right)^2 \tag{21}$$

where

I_{DSS} = drain saturation current
V_{sg} = the source-gate voltage
V_p = the pinch-off voltage.

This equation applies only so long as V_{sg} is between 0 and V_p. JFETs are seldom used with the gate-to-channel diode forward biased. Drain current, I_D, is 0 when V_{sg} exceeds V_p. This equation applies to both biasing and signal design.

Now that we have some basic equations, let's go on to some other areas:

- Small-signal amplifier design (and application limits).
- Large-signal amplifier design (distortion from nonlinearity).

BIPOLAR TRANSISTORS (SMALL SIGNALS)

Transistors are usually driven by both biasing and signal voltages. Small-signal models treat only the signal components. We will consider bias and nonlinear signal effects later.

A Basic Common-Emitter Model

Fig 6.43 shows a BJT amplifier. The circuit is adequately described by equation 19. A mathematical analysis of the control and output currents and voltages yields the small-signal common-emitter

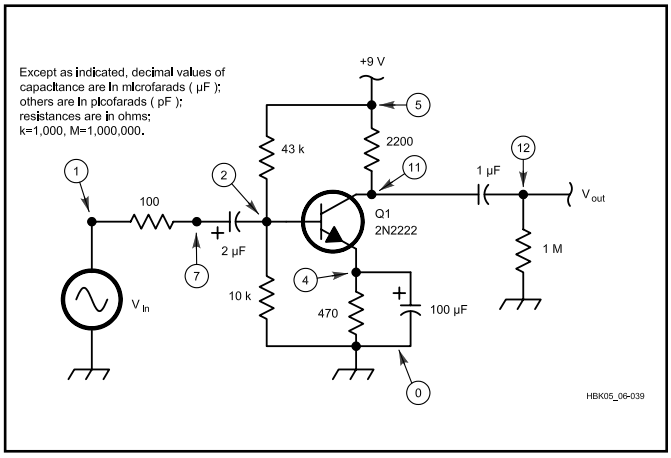


Fig 6.39—Common-emitter audio amplifier used in the text.

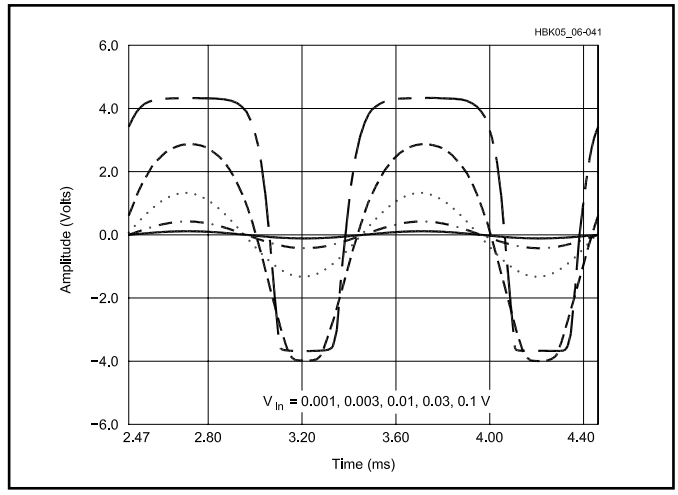


Fig 6.41—Output waveforms for 1-kHz inputs of different magnitudes. Note the clipping that begins to appear at approximately V_{in} = 0.03 V.

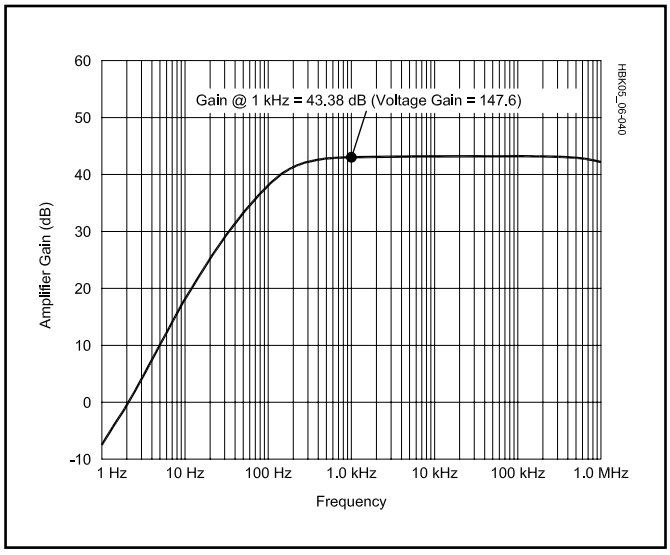


Fig 6.40—Frequency response of the voltage gain of the amplifier in Fig 6.39. The gain at 1 kHz is approximately 148 or 43 dB.

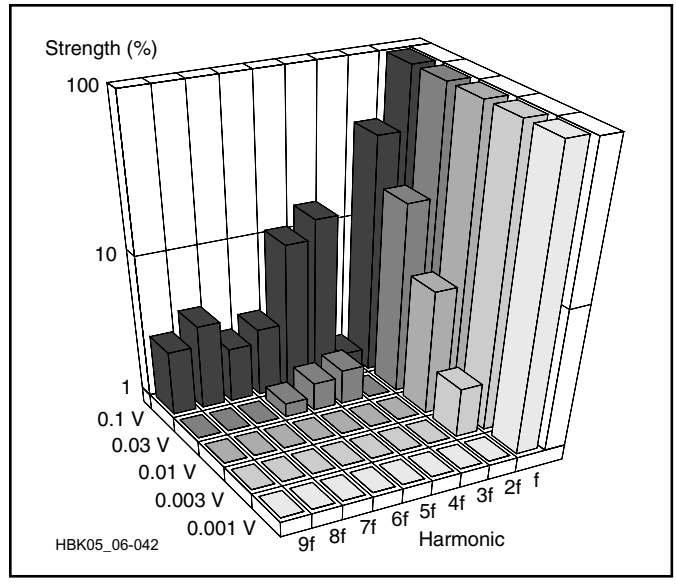


Fig 6.42—Fourier (harmonic) decomposition and Total Harmonic Distortion values for the waveforms in Fig 6.41.

amplifier model shown in **Fig 6.44**. This is the most common of all transistor small-signal models, a controlled current source with emitter resistance.

In order to use this model, however, we must have a value for r_e. That's no trouble, however, $r_e = kT / qI_0$, or $r_e = 26 / I_e$, where I_e is the dc bias current in milliamperes. This value applies at a typical ambient temperature of 300 K.

The device output resistance is infinite because it is a pure current source, which is a good approximation for most silicon transistors at low frequencies. In use, the collector lead would feed a load resistor, R_L. For this model, that resistance must be small enough so that the collector bias voltage is positive for the chosen bias current.

Gain vs Frequency

Fig 6.44 is a low frequency approximation. As signal frequency increases, however, current gain appears to decrease. The low-frequency current gain is β_0. β_0 is constant through the audio spectrum, but it eventually decreases, and at some high frequency it will drop by a factor of 2 for each doubling of signal frequency. A transistor's frequency vs current gain relationship is specified by its *gain-bandwidth product*, or F_T. F_T is the frequency at which the current gain is 1. Common transistors for lower RF applications might have $\beta_0 = 100$ and $F_T = 500$ MHz. The frequency at which current gain is β_0 is called F_b and related to F_T by $F_\beta = F_T / \beta_0$.

The frequency dependence of current gain is modeled by adding a capacitor across the base resistor of Fig 6.50A; **Fig 6.45**, the *hybrid-pi* model results. The capacitor reactance should equal the low-frequency input resistance, $(\beta + 1)r_e$, at F_b. This simulates a frequency-dependent current gain.

Three Simple Models

Even though transistor gain varies with frequency, the simple model is still useful under certain conditions. Calculations show that the simple model is valid, with $\beta = F_T / F$, for frequencies well above F_β.

APPENDIX: SPICE Input Files

CE Audio Amp (Case Study)

```
Amplifier with distortion test
*
VSIG  1 0  ac 0.01 SIN (0 0.001 1KHZ)
RIN   1 7  100
CIN   7 2  2UF
RB1   5 2  43K
RB2   2 0  10K
Q1    11 2 4  Q2N2222A
RC    5 11  2200
RE    4 0  470
CE    4 0  100UF
COUT  11 12   1UF
ROUT  12 0    1MEG
VBIAS 5 0  DC 9
*
..MODEL Q2N2222A NPN(Is=14.34f Xti=3 Eg=1.11
Vaf=74.03  Bf=255.9  Ne=1.307  +Ise=14.34f
Ikf=.2847 Xtb=1.5 Br=6.092 Nc=2 Isc=0 Ikr=0 Rc=1
+Cjc=7.306p Mjc=.3416 Vjc=.75 Fc=.5 Cje=22.01p
Mje=.377  Vje=.75  +Tr=46.91n  Tf=411.1p  Itf=.6
Vtf=1.7 Xtf=3 Rb=10)
*National   pid=19  case=TO18
*
..AC DEC 10 1HZ 1MEGHZ
..TRAN 1MS 10MS 2MS 20US
..FOUR 1KHZ V(12)
..PROBE v(12) v(1)
..END
```

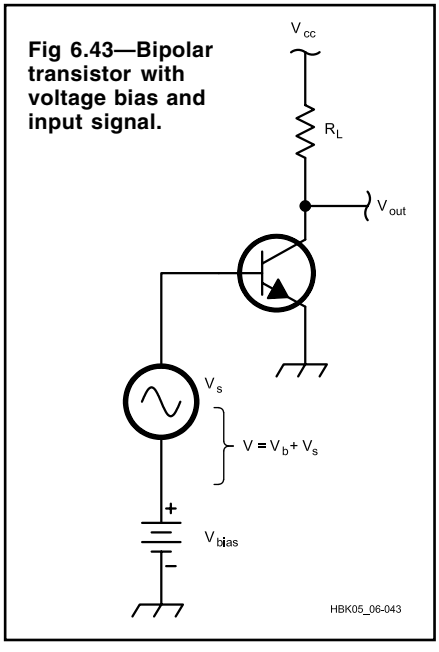


Fig 6.43—Bipolar transistor with voltage bias and input signal.

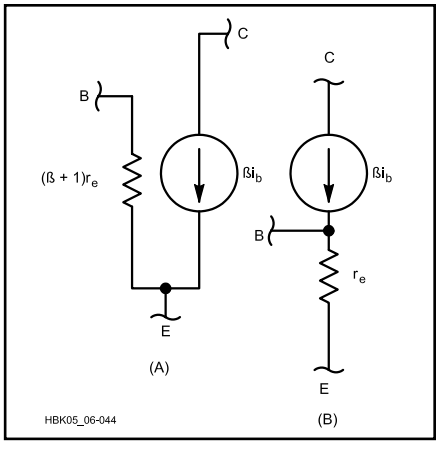


Fig 6.44—Simplified low-frequency model for the bipolar transistor, a "beta generator with emitter resistance." r_e=26 / I_e(mA dc).

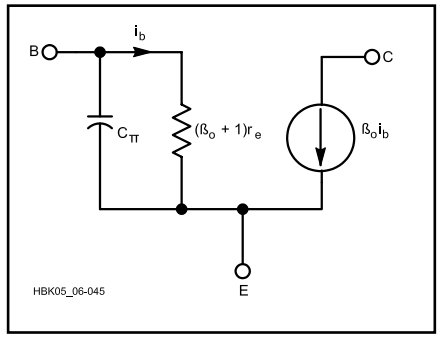


Fig 6.45—The hybrid-pi model for the bipolar transistor.

The approximation worsens, however, as the operating frequency (f_0) approaches F_T. **Fig 6.46** shows small-signal models for the three common amplifier configurations: common emitter (ce), common base (cb) and common collector (cc).

The common-collector amplifier, unlike the common-emitter or common-base configurations, has a finite output resistance. This resistance is calculated by short circuiting the input voltage source, V_s, and "driving" the output port with either a voltage or current source.

The common-collector example shows characteristics that are more typical of practical RF amplifiers than the idealized ce and cb amplifiers. Specifically, the input resistance is a function of both the device and the termination at the output. The output resistance is critically dependent upon the input driving source resistance.

These examples have used the simplest of models, the controlled current generator with an emitter resistance, r_e. Better models are necessary to design at high frequencies. The simplified hybrid-pi of Fig 6.45 is often suitable.

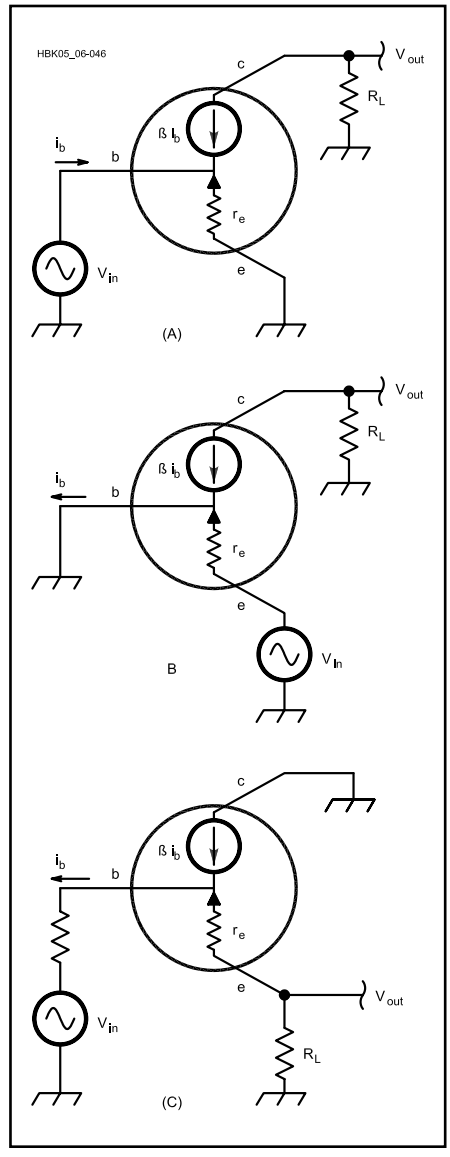


Fig 6.46—Application of small-signal models for analysis of (A) the ce amplifier, (B) the cb and (C) the cc bipolar transistor amplifiers.

Small-Signal Design at RF

Fig 6.47 shows a better small-signal model for RF design that expands on the hybrid-pi. Consider the physical aspects of a real transistor: There is some capacitance across each of the PN junctions (C_{cb} and C_p) and capacitance from collector to emitter (C_c). There are also capacitances between the device leads (C_e, C_b and C_c). There is a resistance in each current path, emitter to base and collector. From emitter to base, there is r_π from the hybrid-pi model and r´b, the "base spreading" resistance. From emitter to collector is R_o, the output resistance. The leads from the die to the circuit present three inductances.

Manual circuit analysis with this model doesn't look like much fun. It's best tackled with the aid of a computer and specialized software. Other methods are presented in Hayward's *Introduction to RF Design.*

Don't be intimidated by the complexity of the model, however. Surprisingly accurate results may be obtained, even at RF, from the simple models. Simple models also give a better "feel" for device characteristics that might be obscured by the mathematics of a more rigorous treatment. Use the simplest model that describes the important features of the device and circuit at hand.

Biasing Bipolar Transistors

Proper biasing of the bipolar transistor is more complicated than it might appear. The Ebers-Moll equation would suggest that a common-emitter amplifier could be built as shown in Fig 6.43, grounding the emitter and biasing the base with a constant voltage source. Further examination shows that this presents many problems. The worst is that constant-voltage bias ultimately leads to thermal *runaway*. Constant-voltage biasing applied to the base is almost never used.

Constant base-current biasing is shown in **Fig 6.48A**. This works reasonably well if the current gain is known, which is rarely true. A transistor with a typical β of 100 might actually have values ranging from 50 to 250. A slightly improved method is shown in Fig 6.48B, where the bias is derived from the collector. As current increases, collector voltage decreases, as does the bias current flowing through R_1. This ensures operation in the transistor active region.

The most common biasing method is shown in **Fig 6.49A**. The device model used, shown in Fig 6.49B, is based on the Ebers-Moll equation, which shows that virtually no transistor current flows until the base-emitter voltage reaches about 0.6. Then, current increases dramatically with small additional voltage change. The transistor is thus modeled as a current controlled generator with a battery in series with the base. The battery voltage is ΔV.

The circuit is analyzed with nodal equations. The collector resistance, R_5, is initially assumed to be zero. The analysis results in three equations for V_b, V_c´ and I_b:

$$V_b = \frac{V_{cc}\ R2R3 + \Delta V R2\left(\frac{R4 + R1}{\beta + 1}\right)}{R3R4 + R2R4 + R2R3 + R1R3 + \frac{R1R2}{\beta + 1}} \quad (22)$$

$$V_C' = \frac{R1V_{cc}R4V_b + \beta R1R4\left(\frac{\Delta V - V_b}{R3(\beta + 1)}\right)}{R1 + R4} \quad (23)$$

$$I_b = \frac{V_b - \Delta V}{R3(\beta + 1)} \quad (24)$$

The emitter current is then $I_b(\beta + 1)$. Once the circuit has been analyzed, R5 may be taken into account. The final collector voltage is

$$V_c = V_c' - \beta I_b\ R5 \quad (25)$$

The solution is valid so long as V_c exceeds V_b.

Analysis of these equations with a computer or hand-held programmable calculator shows that I_e is not a strong function of the transistor parameters, ΔV and β. In practice, the base biasing resistors, R1 and R2, should be chosen to draw a current

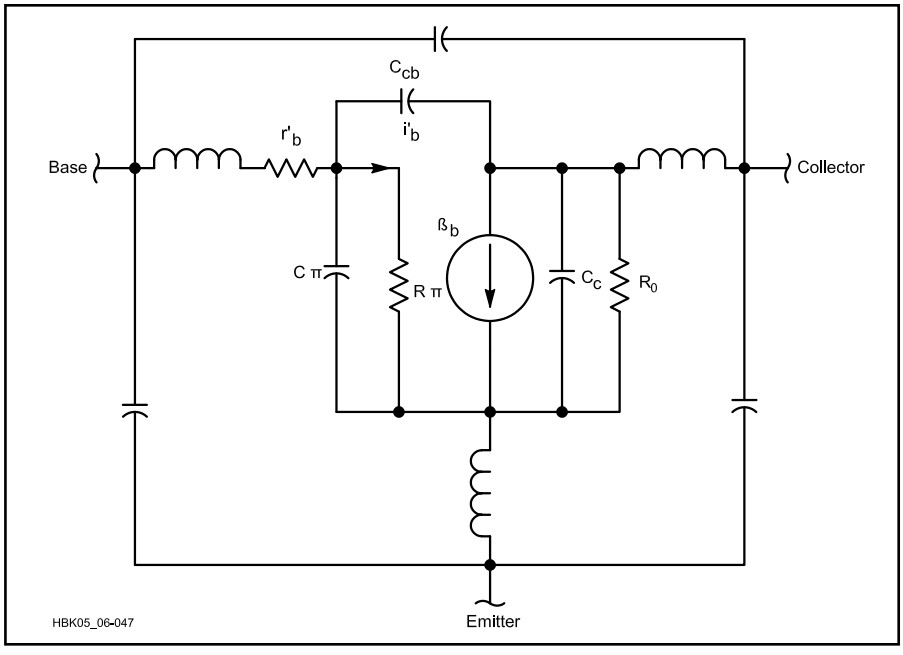


Fig 6.47—A more refined small-signal model for the bipolar transistor. Suitable for many applications near the transistor F_T.

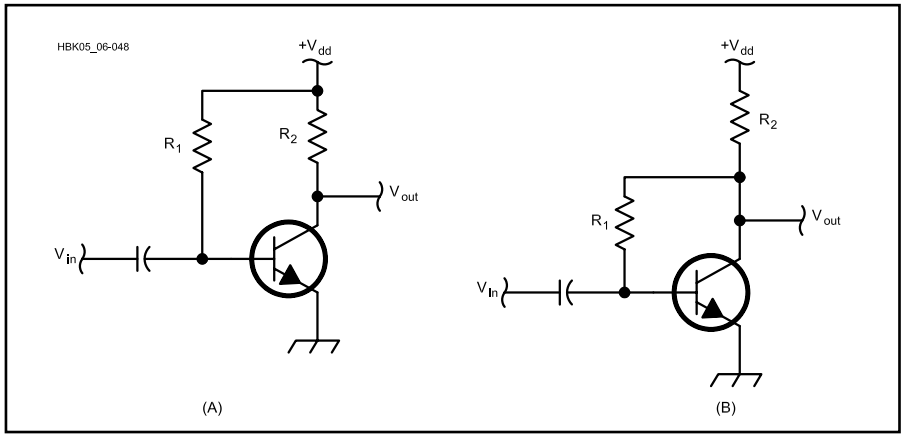


Fig 6.48—Simple biasing methods for ce amplifier. The scheme at (A) suffers if β is not well known. Negative feedback is used in (B).

that is much larger than I_b (to eliminate effects of β variation). V_b should be much larger than ΔV to reduce the effects of variations in ΔV.

Three additional biasing schemes are presented in **Fig 6.50**. All provide bias that is stable regardless of device parameter variations. A and B require a negative power supply. The circuit of Fig 6.50C uses a second, PNP, transistor for bias control. The PNP transistor may be replaced with an op amp if desired. All three circuits have the transistor emitter grounded directly. This is often of great importance in microwave amplifiers. These circuits may be analyzed using the simple model of Fig 6.49.

The biasing equations presented may be solved for the resistors in terms of desired operating conditions and device parameters. It is generally sufficient, however, to repetitively analyze the circuit, using standard resistor values.

The small-signal transconductance of a common-emitter amplifier was found in the previous section. If biased for constant current, the small-signal voltage gain will vary inversely with temperature. Gain may be stabilized against temperature variations with a biasing scheme that causes the bias current to vary in *proportion to absolute temperature*. Such methods, termed PTAT methods, are often used in modern integrated circuits and are finding increased application in circuits built from discrete components.

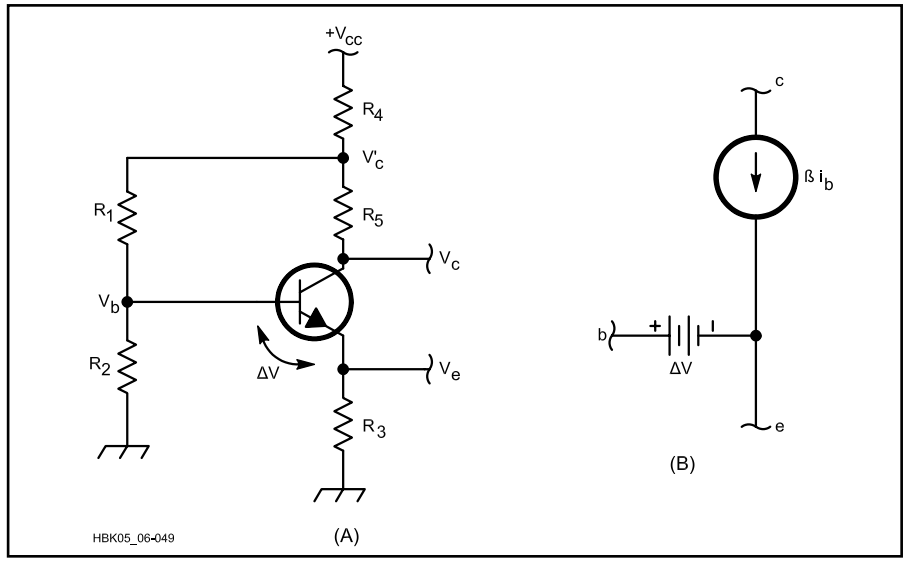


Fig 6.49—(A) Circuit used for evaluation of transistor biasing. (B) The model used for bias calculations.

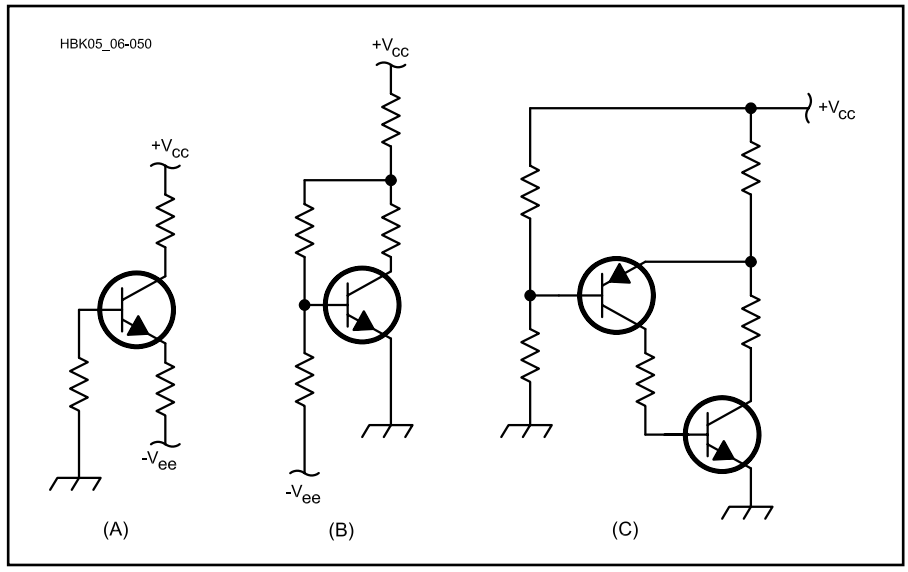


Fig 6.50—Alternative biasing methods. (A) and (B) use dual power supplies, (B) and (C) allow the emitter to be at ground while still providing temperature-stable operation.

Large-Signal Operation

The models presented in previous sections have dealt with small signals applied to a bipolar transistor. While small-signal design is exceedingly powerful, it is not sufficient for many designs. Large signals must also be processed with transistors. Two significant questions must be considered with regard to transistor modeling. First, what is a reasonable limit to accurate application of small-signal methods? Second, what are the consequences of exceeding these limits?

The same analysis of the Ebers-Moll model yields an equation for collector current. The mathematics show that current will vary in a complicated way, for the sinusoidal signal voltage is embedded within an exponential function. Nonetheless, the output is a sinusoidal current if the signal voltage is sufficiently low.

The current of the equation may be studied by normalizing the current to its peak value. The result is relative current, I_r, which is plotted in **Fig 6.51** for V_p values of 1, 10, 30, 100, and 300 mV. The 1-mV case is very sinusoidal. Similarly, the 10-mV curve is generally sinusoidal with only minor distortions. The higher amplitude cases show increasing distortion.

Constant base-voltage biasing is unusual. More often, a transistor is biased to produce nearly constant emitter current. When such an amplifier is driven by a large input signal, the average bias voltage will adjust itself until the time average of the nonlinear current equals the previous constant bias current. Hence, it is vital to consider the average relative current of the waveforms of Fig 6.51. This is evaluated through calculus.

The average relative currents for the cases analyzed occur at the intersection of the curves with the dotted lines of Fig 6.51.

Ready-Made Models

Many manufacturers provide computer models, S-parameter files or other data helpful when including their devices in circuits modeled with *SPICE* and other computer tools. These files are often available from component manufacturers' Web sites, from software providers, or from third parties. Because this information changes, use your favorite Internet search tool to look for information on components of interest.

For example, the dotted curve intersects the V_p = 300-mV waveform at an average relative current of 0.12. If an amplifier was biased to a constant current of 1 mA, but was driven with a 300-mV signal, the positive peak current would reach a value greater than the average by a factor of 1/(0.12). The average current would remain at 1 mA, but the positive peak would be 8 mA. The transistor would not conduct for most of the cycle.

The curves have presented data based upon the simplest of large-signal models, the Ebers-Moll equation. Still, the simple model has yielded considerable information. The analysis suggests that a reasonable upper limit for accurate small-signal analysis is a peak base signal of about 10 mV. The effect of emitter degeneration is also evident. Assume a transistor is biased for r_e = 5 Ω and an external emitter resistor of 10 Ω is used. Only the r_e portion of the 15 Ω total is nonlinear. Hence, this amplifier would tolerate a 30-mV signal while still being well described with a small-signal analysis.

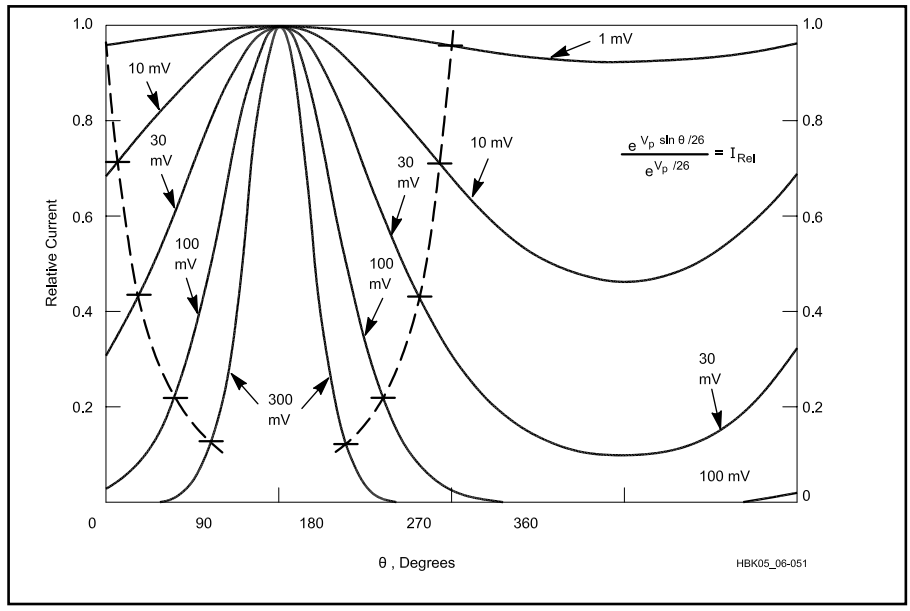


Fig 6.51—Normalized relative current of bipolar transistor under sinusoidal drive at the base.

FETS

An often used device in RF applications is the field-effect transistor (FET). There are many kinds: JFETs, MOSFETs and so on. Here we will discuss JFETs, with the understanding that other FETs are similar.

We viewed the bipolar transistor as controlled by either voltage or current. The JFET, however, is purely a voltage controlled element, at least at low frequencies. The input gate is usually a reverse biased diode junction with virtually no current flow. The drain current is related to the source-gate voltage by:

$$I_D = I_{DSS}\left(1 - \frac{V_{sg}}{V_p}\right)^2 \qquad 0 \le V_{sg} \le V_p$$
$$I_D = 0 \qquad V_{sg} > V_p \qquad (26)$$

where I_{DSS} is the drain saturation current and V_p is the pinch-off voltage. Operation is not defined when V_{sg} is less than zero because the gate diode is then forward biased. Equation 26 is a reasonable approximation as long as the drain bias voltage exceeds the magnitude of the pinch-off voltage.

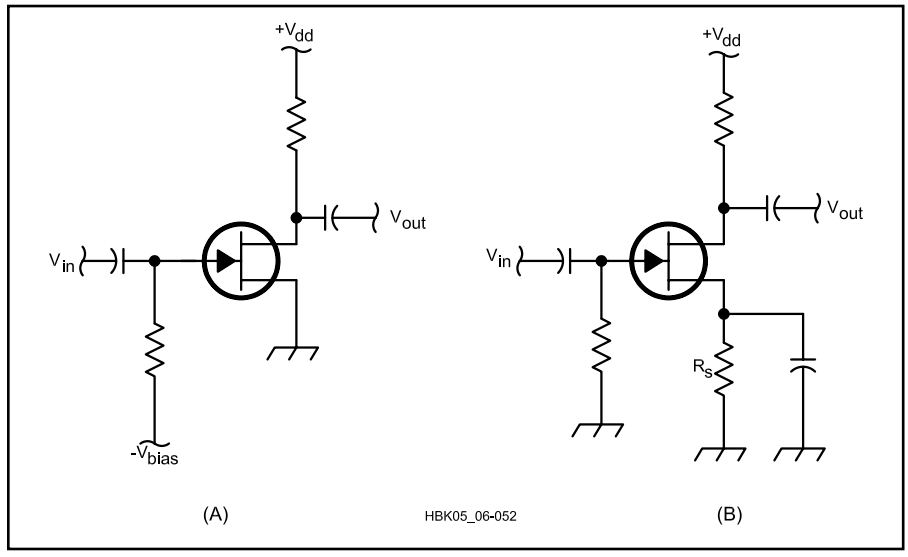


Fig 6.52—Biasing schemes for a common-source JFET amplifier. $-V_{bias}$ is normally adjusted to suit each device; there is a significant spread over a product run. Also note that some FETs can exhibit thermal runaway in some current/temperature ranges.

Biasing FETs

Two virtually identical amplifiers using N-channel JFETs are shown in **Fig 6.52**. The two circuits illustrate the two popular methods for biasing the JFET. Fixed gate-voltage bias, Fig 6.52A, is feasible for JFETs because of their favorable temperature characteristics. As the temperature of the usual FET increases, current decreases, avoiding the thermal-runaway problem of bipolar transistors.

A known source resistor, R_s in Fig 6.52B, will lead to a known source voltage. This is obtained from a solution of equation 26 (see **Eq 27**).

The drain current is then obtained by direct substitution.

$$V_{sg} = \frac{\left[\frac{1}{R_s I_{DSS}} + \frac{2}{V_p}\right] - \left[\left(\frac{1}{R_s I_{DSS}} + \frac{2}{V_p}\right)^2 - \left(\frac{2}{V_p}\right)^2\right]^{0.5}}{\frac{2}{V_p^2}} \qquad (27)$$

Alternatively, a desired drain current less than I_{DSS} may be achieved with a proper choice of source resistor

$$R_s = \frac{V_p\left(1 - \frac{I_D}{I_{DSS}}\right)^{0.5}}{I_D} \qquad (28)$$

The small-signal transconductance of the JFET is obtained by differentiating equation 26

$$g_m = \frac{dI_D}{dV_{sg}} = \frac{-2\,I_{DSS}}{V_p}\left(1 - \frac{V_{sg}}{V_p}\right) \qquad (29)$$

The minus sign indicates that the equation describes a common-gate configuration. The amplifiers of Fig 6.52 are both common-source types and are described by equation 29 except that g_m is now positive. Small-signal models for the JFET are shown in **Fig 6.53**. The simple model is that inferred from the equations while the model of Fig 6.53B contains capacitive elements that are effective in describing high-frequency behavior. Like the bipolar transistor, the JFET model will grow in complexity as more sophisticated applications are encountered.

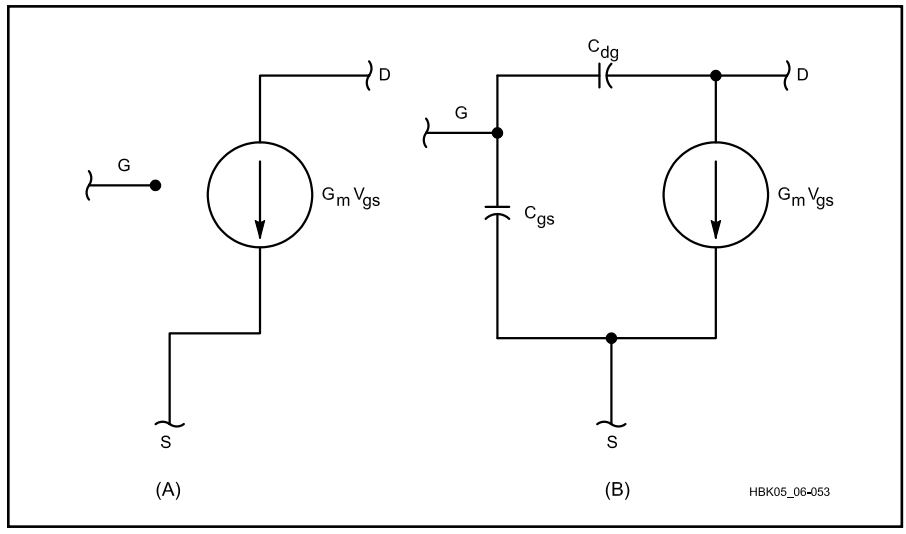


Fig 6.53—Small-signal models for the JFET. (A) is useful at low frequency, (B) is a modification to approximate high-frequency behavior.

Large-Signal Operation

Large-signal JFET operation is examined by normalizing the previous equation to $V_p = 1$ and $I_{DSS} = 1$ and injecting a sinusoidal signal. The circuit is shown in **Fig 6.54**. Also shown in the figure are examples for a variety of bias and sinusoid amplitude conditions. The main feature is the asymmetry of the curves. The positive portions of the oscillations are farther from the mean than are the negative excursions. This is especially dramatic when the bias, v_0, is large, which places the quiescent point close to pinch-off. With such bias and high-amplitude drive, conduction occurs only over a small fraction of the total input waveform period.

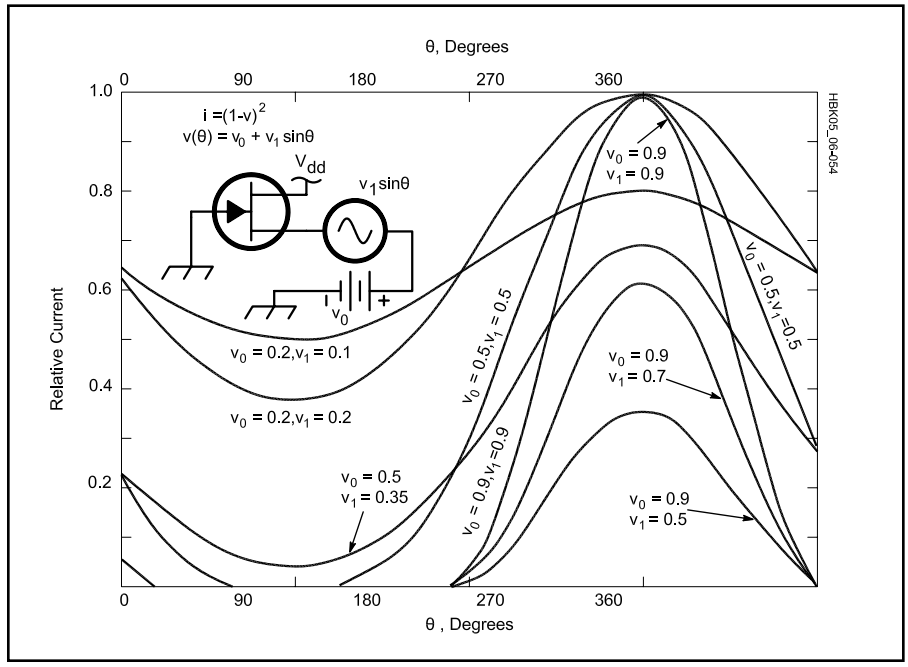


Fig 6.54—Relative normalized drain current for a JFET with constant voltage bias and sinusoidal signals. Relatively "clean" waveforms exist for low signals while large input amplitudes cause severe distortion.

The average current for these operating conditions can be determined by calculus. The average current values obtained may be further normalized by dividing by the corresponding dc bias current, $I_0 = (1 - v_0)^2$. The results are shown in **Fig 6.55**. The curves show that the average current increases as the amplitude of the drive increases. This, again, is most pronounced when the FET is biased close to pinch-off.

Although practical for the JFET, constant-voltage operation in the previous curves is not common. Instead, a resistive bias is usually employed, Fig 6.52B. With this form of bias, the increased current from high signal drive will cause the voltage drop across the bias resistor to increase. This will then move the quiescent operating level

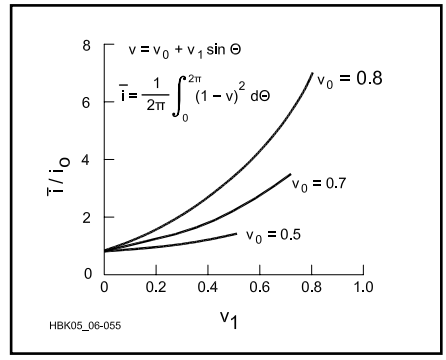


Fig 6.55—Change in average current of a JFET with increasing input signals. The average current with no input signals is I_0, while v_1 is the normalized drive amplitude, and v_0 is the bias voltage.

closer to pinch-off, accompanied by a reduced small-signal trans-conductance. This behavior is vital in describing the limiting found in FET oscillators.

The limits on small-signal operation are not as well defined for a FET as they were for the bipolar transistor. Generally, a maximum voltage of 50 to 100 mV is allowed at the input (normalized to a 1-V pinch-off) without severe distortion. The voltages are much higher than they were for the bipolar transistor. However, the input resistance of the usual common source amplifier is so high and the corresponding transconductance low enough that the available gain is no greater than could be obtained with a bipolar transistor. The distortion is generally less with FETs, owing to the lack of high-order curvature in the defining equations.

Many of the standard circuits used with bipolar transistors are also practical with FETs. Noting that the transconductance of a bipolar transistor is $g_m = I_e$ (dc mA) / 26, the previous equations may be applied directly. The "emitter current" is chosen to correspond with the FET transconductance. A very large value is used for current gain. The same calculator or computer program is then used directly. In practice, much higher terminating impedances are needed to obtain transducer gain values similar to those of bipolar transistors.

References

1. Ebers, J., and Moll, J., "Large-Signal Behavior of Junction Transistors," *Proceedings of the IRE*, 42, pp 1761-1772, December 1954.
2. Getreu, I., *Modeling the Bipolar Transistor*, Elsevier, New York, 1979. Also available from Tektronix, Inc, Beaverton, Oregon, in paperback form. Must be ordered as Part Number 062-2841-00.
3. Searle, C., Boothroyd, A., Angelo, E. Jr., Gray, P., and Pederson, D., *Elementary Circuit Properties of Transistors*, Semiconductor Electronics and Education Committee, Vol 3, John Wiley & Sons, New York, 1964.
4. Clarke, K. and Hess, D., *Communication Circuits: Analysis and Design*, Addison-Wesley, Reading, Massachusetts, 1971.
5. Middlebrook, R., *Design-Oriented Circuit Analysis and Measurement Techniques*. Copyrighted notes for a short course presented by Dr Middlebrook in 1978.

Suggested Additional Reading

Alley, C. and Atwood, K., *Electronic Engineering*, John Wiley & Sons, New York, 1962.

Terman, F., *Electronic and Radio Engineering*, McGraw-Hill, New York, 1955.

Gilbert, B., "A New Wide-Band Amplifier Technique," *IEEE Journal of Solid-State Circuits*, SC-3, 4, pp 353-365, December 1968.

Egenstafer, F., "Design Curves Simplify Amplifier Analysis," *Electronics*, pp 62-67, August 2, 1971.

Notes

[1]Superconducting circuits are the one exception; but they are outside the scope of this book.

[2]There are numerous manufacturers of excellent heat sinks. References to any one manufacturer are not intended to exclude the products of any others, nor to indicate any particular predisposition to one manufacturer. The catalogs and products referred to here are from: Wakefield Thermal Solutions, Inc., 33 Bridge Street, Pelham, NH 03076, Phone: (603) 635-2800.