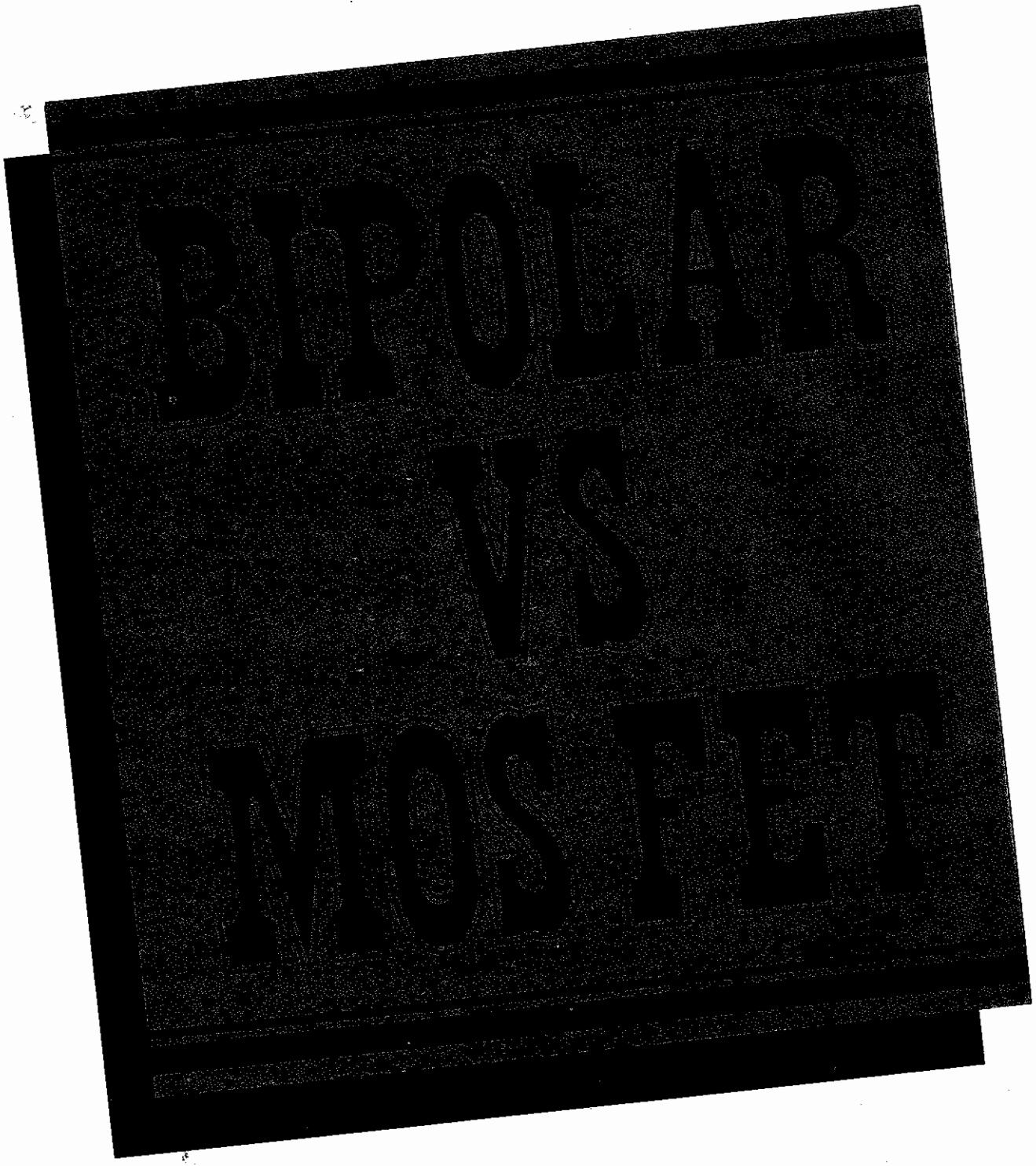


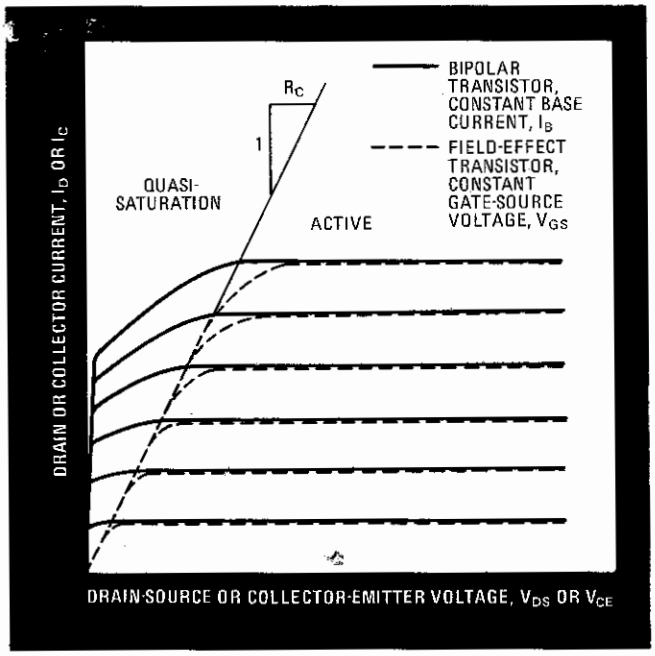
Seeing where the power lies

For switching applications requiring minimum device size, bipolar transistors are more efficient below 15 kHz

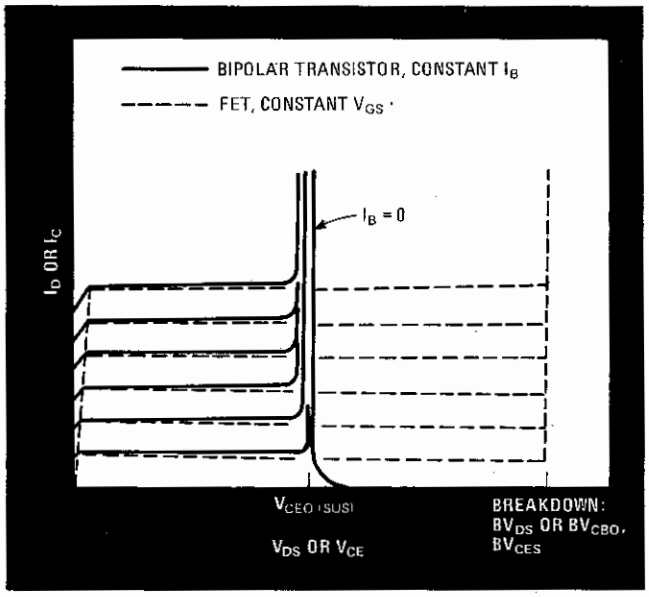
by Philip L. Hower, *Westinghouse Electric Corp., Research and Development Center, Pittsburgh, Pa.*

☐ Much has been written about the advantages of MOS field-effect power transistors over bipolar devices, leaving the impression that MOS FETs have few disadvantages other than a comparatively higher price that is on the decline. A comparison of both devices, however, shows that bipolar power transistors can outshine MOS FETs in areas other than price. For example, where minimum device size is important and power is being switched at a frequency of 15 kilohertz or less, bipolar power transistors excel.

A valid comparison of both devices, which have as much in common as they have differences, can be made with regard to switching frequency, junction temperature, and current density, provided that the devices have the same die area. In addition, the costs per assembled



1. Comparison. The forward voltage drop of a bipolar power transistor is lower than that of a power MOS FET, as can be seen from the collector and drain characteristics of the devices. The difference is a result of the additional charge a bipolar device stores during turn-on.



2. Limits. At high voltage, two potentials come into play for a bipolar power transistor, whereas only one is of interest for a MOS FET. The two are the open-circuit collector-emitter (sustained) and collector-base voltages, the higher of which is the breakdown voltage.

die are assumed to be approximately the same.

Besides being fabricated using the same silicon semiconductor technology, both bipolar transistors and MOS FETs are charge-controlled components. Charge is supplied to and removed from each device to achieve switching action. Both have lightly doped collector or drain regions, which are needed to block voltages while in their off-states. Another similarity of the two is that each has the same "ideal" maximum operating temperature, which is determined by thermal runaway of leakage current generated within the lightly doped region. This maximum operating temperature can be decreased in either type of device depending on the junction passivation scheme used.

Introducing charge

However, the manner in which charge is introduced and its location within each device differs. In bipolar transistors, turn-on is achieved by supplying excess minority carriers (holes in an npn device) to the base—but what constitutes the base in a bipolar transistor is subject to the operating point of the transistor's collector. When a bipolar transistor is turned on at its lowest voltage state (classical saturation), charge must be supplied in three different regions. Excess charge is first built up in the metallurgical base region. It is then supplied to the collector n region to form a current-induced base. Finally, the base-collector junction, known as a remote base, becomes externally forward-biased, thereby allowing the parasitic base-collector diode to store excess charge.

Most of the switching losses incurred in a bipolar power transistor during turn-on occur during the time needed to store charge in the metallurgical base and collector n regions.

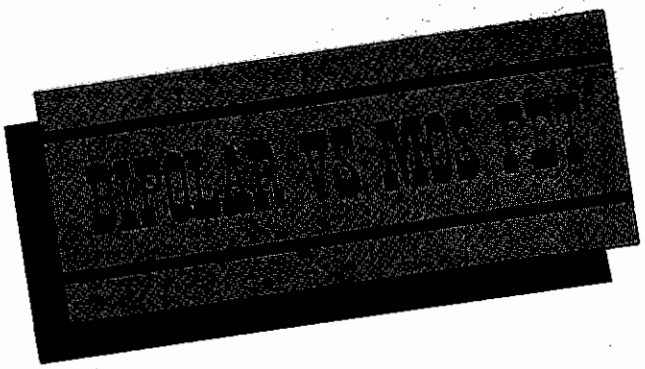
Turning off, or removing charge from, a bipolar power transistor involves a reversal of the sequence of events described for turn-on. Charge is first removed from the remote base in what is a relatively lossless process. It is then removed from the collector n region via two different mechanisms: by a vertical back-injection of holes into the emitter from the current-induced base and by charge removal from the lateral edges of the current-induced base. This lateral sweep-out leads to a reduction in the effective conduction area and to an increase in current density.

Removal of charge from the collector n region completes the storage-time phase. During this interval, the collector current remains relatively constant, except near the end of the storage time, when current begins to decrease and the collector voltage begins to rise. When the charge from this region is nearly all removed, the collector junction begins its voltage-blocking action.

At this point, the bipolar power transistor is operating within the active region, and the applied reverse current quickly removes charges from the metallurgical base, producing a correspondingly rapid decrease in collector current. The accompanying current fall time that completes the turn-off process depends on the base transit time, the current gain, and the reverse base current. Charge removal from the metallurgical base and collector n regions is a relatively high-loss process and accounts for most of the switching losses in a bipolar power transistor.

Small turn-on charge

In contrast to bipolar power transistors, MOS FETs require only a small amount of charge for turning on. This charge is used to form the channel, reduce the width of the depletion layer (a layer that supports the off-state voltages), and charge parasitic gate-drain and gate-source capacitances. For small values of parasitic



capacitances, turn-on time is rapid, because of the relatively short transit times associated with the channel and the depletion layer. Thus switching losses in a MOS FET can be considerably lower than those of comparable-sized bipolar power transistors.

However, the fact that a bipolar power transistor stores more charge than a MOS FET can be an advantage. The additional charge reduces forward-voltage drop (Fig. 1), assuming that the resistivity and thickness of the n layer for both a bipolar device and a MOS FET are the same. Furthermore, it is assumed that both devices have the same current-carrying cross-sectional areas. The latter assumption neglects a MOS FET's channel voltage drop and any resistance associated with current-spreading or junction-FET action in the drain region of the device.

Conduction may differ

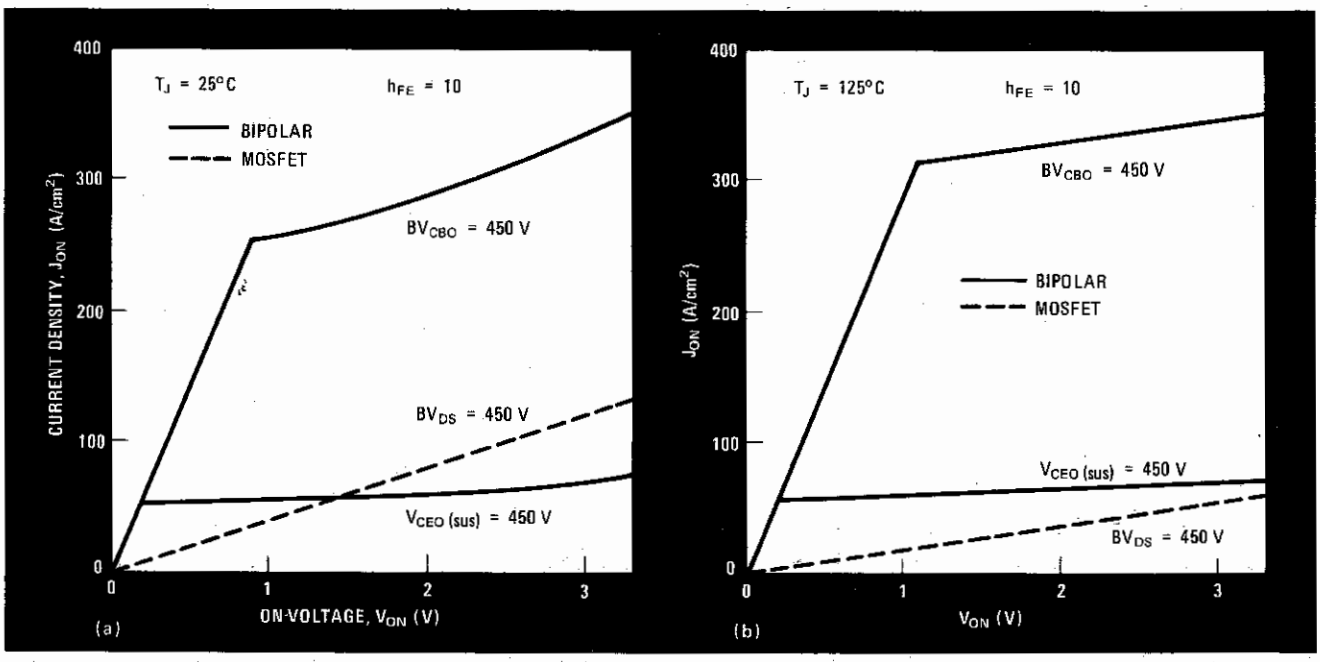
An assumption of equal conduction areas for both types of device does not necessarily mean that individual die areas will be equal. The effective emitter area in a typical bipolar transistor design is about one half that of its die area. Until recently, such a ratio has not been typical for MOS FETs, being usually smaller. Nevertheless, recent MOS FET design advances have made possible similar ratios. Consequently, for the purpose of the thermal calculations here, the one-half ratio of conduction area to die area will be assumed for both bipolar devices and MOS FETs.

As can be seen in Fig. 1, the quasi-saturation region of a bipolar power transistor permits operation at a lower forward-voltage drop than does that of the MOS FET. The reduction in collector-emitter voltage (V_{CE}) depends in each case on the values of the base current (I_B) and the collector current (I_C).

Figure 2 shows the collector and drain characteristics for both bipolar devices and MOS FETs at higher voltages. The major difference from the characteristics shown in the graph in Fig. 1 is that there are now two voltages of interest for the bipolar power transistor: an open-base sustaining collector-emitter voltage ($V_{CEO(sus)}$) and an open-emitter collector-base breakdown voltage (BV_{CBO}). The higher voltage corresponds to the breakdown voltage for both device types, a breakdown that can only be achieved in bipolar power transistors when the emitter is not injecting.

Don't cross

Crossing the $V_{CEO(sus)}$ line at high currents during the inductive turn-off of a bipolar power transistor is likely to initiate secondary breakdown. To avoid that, the bipolar power transistor would have to operate with a reduced off-voltage or with a snubber circuit to improve the I_C and V_{CE} turn-off loci. An alternative is to increase collector thickness and resistivity until $V_{CEO(sus)}$ is equal to the drain-source breakdown voltage (BV_{DS}) of a MOS FET. This increase in resistivity, however, brings with it an undesirable reduction in current gain (h_{FE}) within the quasi-saturation region.



3. Density. At a junction temperature of 25°C (a), large current densities arise in a bipolar power transistor if the open-emitter collector-base breakdown voltage, BV_{CBO}, is the limiting factor. If the open-base sustaining collector-emitter voltage, $V_{CEO(sus)}$, is the limit, current density is reduced by a factor of five. At 125°C (b), a bipolar transistor's behavior changes little, but a MOS FET's changes by a factor of two.

A comparison of current density (J_{on}) for MOS FETs and bipolar power devices can be made by calculating the average current density in the on-state for different on-state voltages. For the MOS FET, the calculation is relatively simple. For each BV_{DS}, an optimum n-region thickness and doping concentration exist for a minimum on-resistance. A plot of current density versus voltage for BV_{DS} equal to 450 volts for a MOS FET is shown in Fig. 3. The slope of the line corresponds to a minimum on-resistance, which in this case is 27 milliohms for an area of 1 square centimeter. The plot is taken with the MOS FET's junction temperature at 25°C.

For bipolar power transistors, calculations involve assumptions of current gain. For the curve shown in Fig. 3, an h_{FE} of 10 at a transistor junction temperature of 25°C is assumed. Since the product of h_{FE} and I_C is constant at large currents, adjustments to other h_{FE} values can be made by appropriately scaling the vertical axis of the curve in Fig. 3.

The bipolar curves are calculated for two different designs: with $V_{CEO(sus)}$ equal to 450 V and with BV_{CBO} equal to 450 V. At low voltages, J_{on} is largely determined by the contact resistance of the emitter metalization, whose resistivity value of 3.5 mΩ-cm² is used here.

Note that large values of J_{on} (more than 300 amperes per square centimeter) are possible if BV_{CBO} is the limiting voltage. On the other hand, should $V_{CEO(sus)}$ be the limit, J_{on} is reduced by a factor of five. In that case, a higher current density can be achieved for MOS FETs with on-voltages of 1.5 V or more.

Taking the heat

Increasing a power transistor's junction temperature from 25° to 125°C results in a minor change of the bipolar devices' current-density curves and greater changes for the MOS FET's J_{on} curve. The reason is that the MOS FET's on-resistance increases by a factor of two as a result of the temperature dependence of electron mobility; this behavior means that MOS FET conduction losses increase with temperature, whereas those of bipolar power transistors remain approximately constant—a fact that is often overlooked when comparing the two types of transistor.

Although the example in Fig. 3 is shown for a blocking potential of 450 V, the relative spacing of the curves for the MOS FET and the bipolar transistors remains fixed for higher blocking voltages, save for the classical saturation voltage. This is because the respective on-current density for both devices increases approximately as $V_{CEO(sus)}$ and BV_{DS} raised to the power of 2.3.

Figure 4 shows power loss as a function of switching frequency for both types of device at a transistor junction temperature of 25°C. Note that at low frequencies, where conduction losses are dominant, a bipolar transistor has about one third the power loss of a MOS FET. This ratio corresponds to the ratio of a bipolar power transistor's on-resistivity to a MOS FET's. The base input power is included in these plots.

As switching frequency is increased, the switching losses of a bipolar power transistor increase at a faster rate than those of a MOS FET, crossing at about 15 kHz. Beyond this frequency, where switching losses dominate, a MOS FET has about a 3:1 advantage over a bipolar power transistor in terms of the amount of power dissipation per unit area.

Heating up

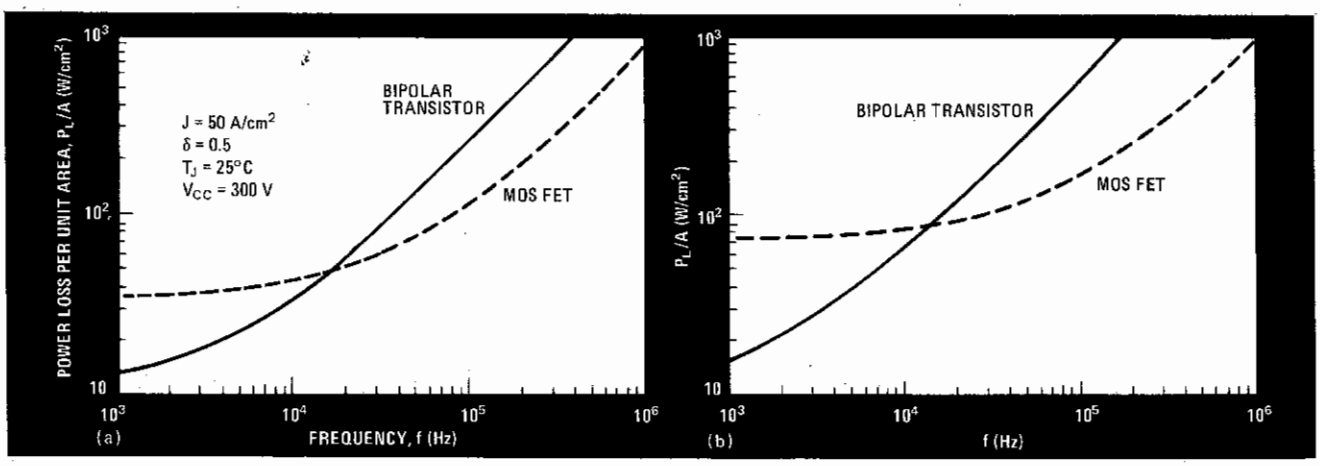
At a transistor junction temperature of 125°C, power-loss curves for both devices take on a different look. At low frequencies, the MOS FET's power losses increase even further than at 25°C, whereas those of the bipolar power transistor remain the same, further widening the bipolar transistor's advantage. However, switching losses for a bipolar power transistor increase even more rapidly with increasing frequency, again crossing over with the MOS FET's power-loss curve at about 15 kHz. At higher frequencies, the MOS FET's advantage climbs to a 6:1 ratio in terms of power loss per unit area.

The curves in Fig. 4 do not account for the practical case where heat removal becomes an important factor. In this case, the assumption of constant junction temperature may be misleading.

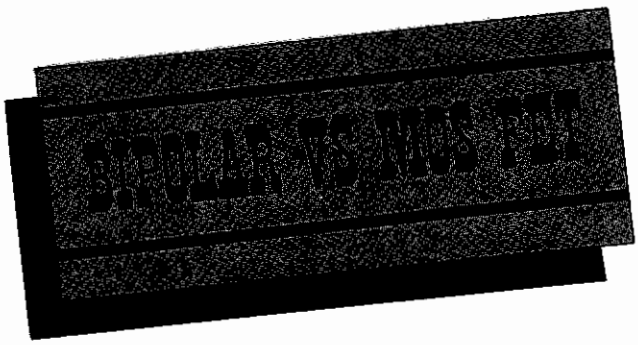
To account for heat flow to the ambient, a fixed thermal resistance can be assumed:

$$R_{\theta JA} = R_{\theta JC} + R_{\theta HS}$$

where $R_{\theta JA}$ is the fixed thermal resistance, $R_{\theta JC}$ is the transistor's junction-to-case thermal resistance, and $R_{\theta HS}$



4. Power loss. At a junction temperature of 25°C (a), power loss per unit area for a bipolar transistor is one third that of a MOS field-effect transistor's at under 15 kHz. Above that frequency, the ratio reverses. Above 125°C (b), the ratio between the two is even greater.



is the effective thermal resistance of the heat sink.

A reasonable value for $R_{\theta JC}$ is approximately 0.2°C/watt, assuming a 1-cm² area and a heat-flow cross-sectional area of 2 cm². This value can be verified with a conventional TO-3 package having sides of 236 mils each and a die 6 millimeters on a side.

Reasonable values

Such a package has a value for $R_{\theta JC}$ of 1.1°C/W, which is in line with the value of 0.2°C/W cited above. $R_{\theta HS}$ is assumed to have a unit area value of about 0.3°C/W, a reasonable value with moderate-sized air-cooled heat sinks.

Heat loss to the ambient, P_{JA}, obeys the equation:

$$T_J - T_A = P_{JA}(R_{\theta JC} + R_{\theta HS}) = P_{JA}R_{\theta JA}$$

where T_J equals the junction temperature and T_A the ambient temperature.

Figure 5 shows P_{JA} as a linear function of T_J. This heat-loss-to-the-ambient curve intersects the power-loss curve at two points, A and B. Two other possibilities exist: no intersecting of the two curves or a tangential solution, where the power loss (P_L) is equal to P_{JA} at only one point. Neither of these possibilities can be considered, however, since thermal runaway will occur in either case.

It is well known from a simple heat-flow argument that point A will be thermally stable, whereas point B will not. It should therefore be understood that calculations of maximum power limits for different values of $R_{\theta JA}$ and T_A using data from point B are impractical.

The safe operating frequency

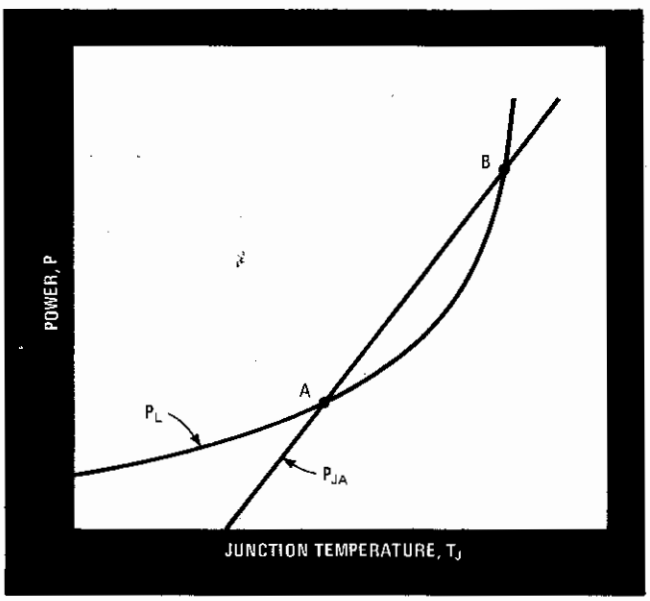
Figure 6 illustrates the case where the operating frequency (f) is selected to give a power loss equal to the heat loss to the ambient ($P_L = P_{JA}$) at a transistor junction temperature of 200°C for an assumed unit-area fixed thermal resistance ($AR_{\theta JA}$) of 0.5°C-cm²/W. The 200°C temperature is typically used as a minimum value for silicon devices having a leakage-current limitation. As can be seen from the graph, both bipolar devices and MOS FETs are thermally stable—that is, they are of the A type in Fig. 5.

The frequency for which the junction temperature is a maximum at $P_L = P_{JA}$ is defined as the maximum safe operating frequency. Plotting current density as a function of this safe operating frequency for both bipolar devices and MOS FETs would reveal a higher current-density value for the former at frequencies under 15 kHz and a lower one at frequencies above that. At 15 kHz, the curves for both devices intersect.

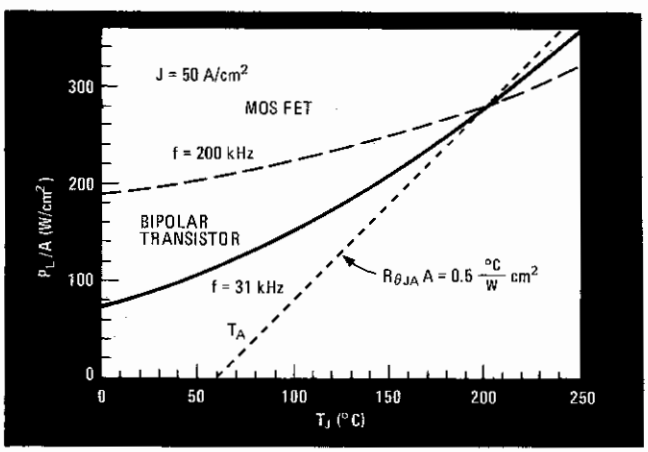
For frequencies below 1 kHz, bipolar power transistors have better than a 2:1 advantage in current- and power-handling capability over MOS FETs. On the other hand, at a frequency greater than 100 kHz, MOS FETs retain at least a 2:1 advantage in this capability.

Possible trends

These calculations have been based on theoretical estimates and measurements on actual devices. As device technologies improve, changes in the switching losses per unit area can be expected. However, conduction losses are already close to theoretical limits and are unlikely to change. Thus the high-frequency portions of Fig. 4 will shift with technological advances, whereas the left-hand portions will remain fixed. □



5. Heat loss. In any transistor, heat loss to the ambient is a linear function of the device's junction temperature. The heat-loss curve intersects the power-loss curve at two points, only one of which (point A) is useful for calculating total device power losses.



6. Safe. At junction temperatures of 200°C or less, the maximum safe operating switching frequency for a bipolar power transistor is 31 kHz, whereas that of a MOS FET is a much higher 200 kHz. This maximum occurs at the point at which power loss equals heat loss.