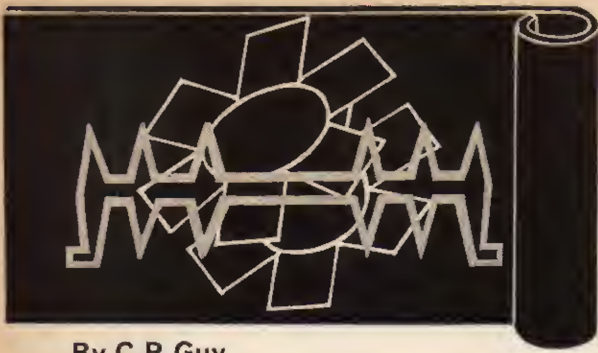


HEATSINKS FOR SEMICONDUCTORS

By C.P. Guy

It is not so long ago since transistors were regarded primarily as low power devices—transistor amplifiers for audio necessitated expensive devices if their output power was to be much more than one watt. The general availability of high power devices now means that the amateur can entertain building high power amplifiers at moderate cost.

The main drawback of power transistors, however, is that it is able to generate more heat than its own mass can dissipate. This can have damaging consequences, more so for germanium devices than silicon. It is often essential to fit heat sinks and radiators, these being available commercially in a wide range of shapes and sizes, but it is a simple matter to make one's own heat sinks. Design or selection of a suitable heat sink is quite simple and a few minutes' calculation may save a couple of expensive output transistors from destruction.

WHENCE DOES THE HEAT COME?

The amount of heat generated by a transistor depends largely on four factors: mode of operation; signal amplitude; bias level; waveform of applied signal.

When referring to a transistor data sheet, figures are quoted for the maximum collector voltage and current and maximum power dissipation. If the transistor is used as a square wave generator or class B square wave amplifier, very little power is dissipated by the transistor compared with an amplifier operating in class A with the same signal output power. The reason for this is that in class B, when driven by a square wave input signal, the collector voltage will be high with low collector current (i.e. transistor "cut off") or low collector voltage with high current (i.e. transistor is "saturated"). The net result is a very low average power dissipation.

If the transistor is driven too fast by the square wave input, then the time taken for the transistor to switch from cut-off to saturation will become an appreciable fraction of the pulse time and the average power will rise. Similarly, if the transistor is loaded too heavily, the output square wave will develop rounded corners, due to the internal resistance. Consequently, internal heating occurs.

Two interesting points arise from this: firstly, this partially explains the reason why the designer of digital equipment (in which transistors are used as on/off switches) is so interested in the cut-off frequency of the device he uses, since a lower frequency device may generate considerable internal heat when driven by too fast a train of pulses. He is, of course, also interested in

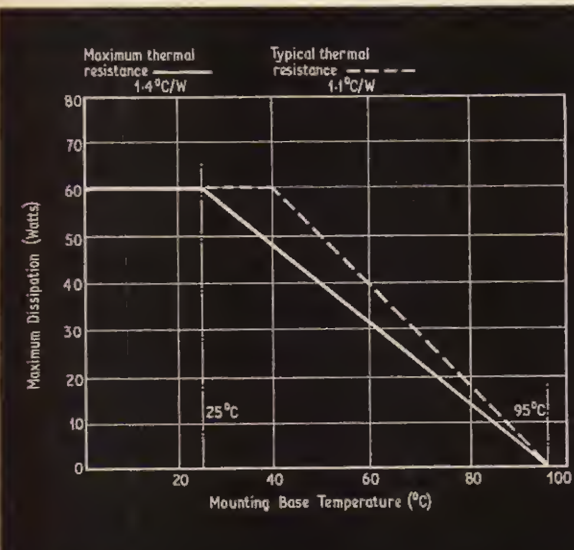


Fig. 1a. Transistor Derating Curve for germanium

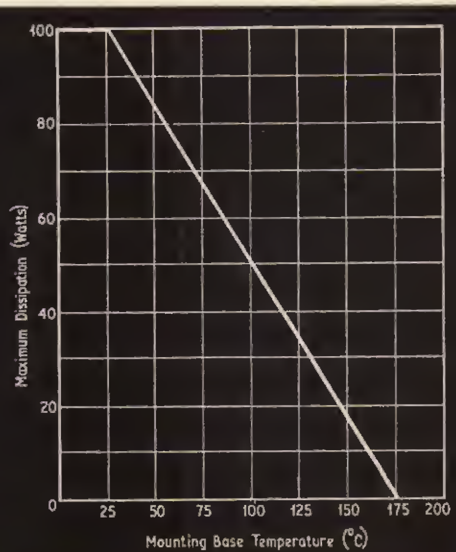


Fig. 1b. Transistor Derating Curve for silicon

whether or not the device will "follow" the input waveform reliably.

Since a large number of transistors are required in most digital equipment, coupled with the fact that equipment becomes physically smaller and smaller, careful thermal analysis is necessary and forced airflow is often employed.

The other interesting point to be deduced from the power transfer efficiency of class B amplifiers is the apparently large output power of pulse-width modulated amplifiers. In these, the output transistors are used as on/off switches and, consequently, dissipate very little power. Therefore, low power devices can be used for high output powers.

Much audio equipment is operated in either class AB or class A mode. Class AB operation eliminates crossover distortion inherent in pure class B amplifiers by the application of a slight forward bias which produces a relatively small collector current (known as the quiescent current). Whenever current flows through a resistor (the junction in this case), heat is generated, so that even the quiescent current will cause the junction temperature to rise to a steady state temperature level.

When the transistor is driven by a sine wave input signal or the complex waveforms encountered in audio working, the average time that current is on is increased; power dissipation in each transistor is increased until it reaches approximately 20 per cent of the maximum stage output. This figure will vary according to the type of input waveform (i.e. programme material).

Class A operation results in a much higher amount of heat dissipation. The quiescent current is set at approximately the mid-point of the I_C/I_B transfer characteristic. There is a relatively large "no-signal" current (about half of the peak collector current) and, hence, the power dissipated at the junction is considerable.

GERMANIUM OR SILICON

Germanium transistors are cheaper than silicon and possess more restrictive characteristics. In general, the maximum junction temperature is in the order of 85 to 100 degrees C, whereas silicon yields figures in the order of 175 to 200 degrees C. Germanium transistors also have the added disadvantage that the leakage current is generally some orders of magnitude higher than in silicon types.

It is possible, in certain circuit configurations, for a transistor to destroy itself by a process known as "thermal runaway". This is caused by the fact that leakage current increases with temperature.

It is possible, under certain circumstances, for the increase in leakage current to cause a further rise in junction temperature which, in turn, causes more leakage current; both current and temperature interact and increase until the maximum junction temperature is exceeded. Quite soon after this occurrence, it becomes necessary to replace the transistor and re-design the circuit to prevent a recurrence.

Maximum ratings on a data sheet are generally quoted at a certain temperature, usually 25 degrees C. It is common practice to derate the power dissipation linearly from the maximum power quoted at 25 degrees C to zero dissipation at the maximum operational temperature. Typical graphs from a manufacturer's data sheets are shown in Figs. 1a and 1b.

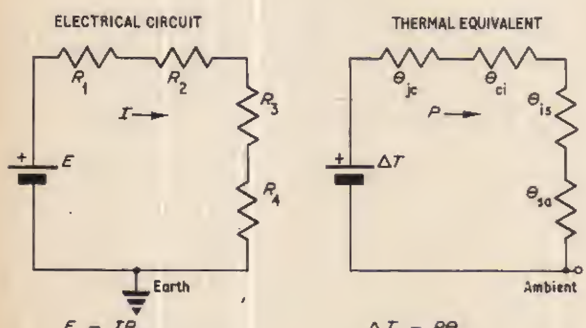
Alternatively, it may be specified that the power dissipation should be reduced so much for each degree rise in temperature. This figure is the thermal conductance of the device which is the inverse of the thermal resistance.

ELECTRICAL ANALOGUE TO THE OPERATION OF A HEAT SINK

It is perhaps helpful to create an analogy between heat power and electrical power: it should be remembered that the heat sink attempts to lower the temperature of the device junction to the ambient temperature (i.e. the temperature surrounding the complete assembly). The junction temperature is usually referred to as T_j and that of ambient temperature as T_{amb}.

Consideration of the electrical circuit in Fig. 2 reveals that, in order to produce a flow of electrical charge (coulombs) from one point to another, a difference in electrostatic pressure (voltage) must exist. The rate of flow of charge may be given in coulombs per second, or amperes. Whatever impedes this flow is called electrical resistance (R) and is measured in ohms (Ω).

Similarly, in order to produce a flow of heat energy (joules) from one point to another, a difference in heat "pressure" (temperature) must exist. The rate of flow may be measured in joules per second, or watts. Whatever impedes this flow is known as thermal resistance (θ) and is measured in thermal ohms (degrees centigrade per watt). Table 1 shows the analogy between electrical and thermal terms.



$E = IR_{tot}$

$R_{tot} = R_1 + R_2 + R_3 + R_4$

$\Delta T = P\theta_{tot}$

$\Delta T = T_j - T_{amb.}$

$\theta_{tot} = \theta_{jc} + \theta_{ci} + \theta_{is} + \theta_{sa}$

Fig. 2. Electrical analogue to thermal circuit

Table 1: COMPARISON OF ELECTRICAL AND THERMAL TERMS

Electrical Term			Thermal Term
EMF	V	volts	Temperature Differential ΔT (degrees centigrade)
Charge	Q	coulombs	Energy J joules
Current	I	amperes	Power P watts
Resistance	R	ohms	Thermal Resistance θ (degrees centigrade per watt)
Conductance	G	mhos	Thermal Conductance (watts per degree centigrade)

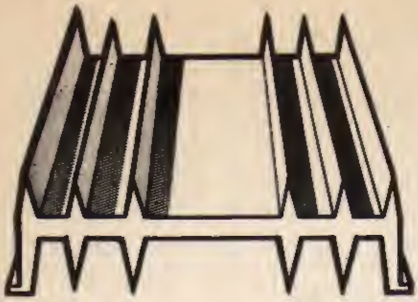


Fig. 3. Typical finned commercial heat sink made of extruded aluminium

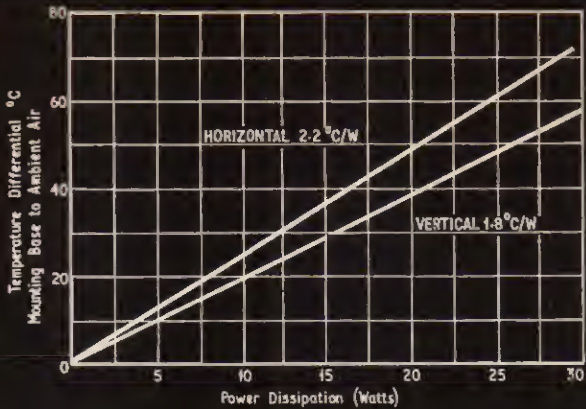


Fig. 4. Thermal characteristics of a finned heat sink

The thermal conductivity path from the transistor junction to the ambient air contains thermal resistances between the junction and the case (θ_{jc}), between the case and heat sink (θ_{cs}), also through an insulator in some cases (θ_{ci}), and finally from the heat sink to the ambient air (θ_{sa}). Due to these resistances, there will always be a temperature differential (ΔT) between junction and ambient and this is the quantity that must be controlled and kept to a minimum.

It is possible to treat these thermal resistances in the same way as electrical resistances and obtain the equation:

$$\theta_t = \theta_{jc} + \theta_{cs} + \theta_{sa} \qquad \ldots (1)$$

or, if an insulating washer is used between the case of the device and the heat sink:

$$\theta_t = \theta_{jc} + \theta_{ci} + \theta_{is} + \theta_{sa} \qquad \ldots (2)$$

where

θ_{tot} = total thermal resistance

θ_{jc} = thermal resistance, junction to case

θ_{cs} = thermal resistance, case to sink

θ_{ci} = thermal resistance, case to insulator

θ_{is} = thermal resistance, insulator to sink

θ_{sa} = thermal resistance, sink to ambient

As a general rule, $\theta_{cs} = \theta_{ci}$. For transistors without a heat sink, θ_{cs} and θ_{sa} combine and become a single quantity, θ_{ca}, the thermal resistance from the case to ambient.

JUNCTION POWER DISSIPATION

The temperature differential depends on the amount of power that the junction is dissipating. The average power dissipation may be approximated as

$$P_d = I_C \times V_{CE} \qquad \ldots (3)$$

where

P_d = average power dissipation in watts

I_C = collector current

V_{CE} = collector-to-emitter voltage

In a single ended class A output stage, the maximum output, P_{tot} may be deduced from the following equation:

$$P_{tot} = \frac{(V_{CE})^2}{2R_L} \qquad \ldots (4)$$

where R_L is the load resistance.

It has already been stated that the class A quiescent current results in a dissipation of approximately half the maximum power output. Under quiescent conditions, the dissipation is maximum since a signal will swing the operating point, and the product of current and voltage on either side of this line will result in less power dissipation. An equation giving an approximation of the maximum power dissipation can be derived from Equation 1.

$$P_d = 0{\cdot}5 \times P_{tot} \qquad \ldots (5)$$

In class AB, the maximum output power is

$$P_{tot} = \frac{2(V_{CC})^2}{2R_{CC}} \qquad \ldots (6)$$

where

V_{CC} = collector-to-collector voltage

R_{CC} = collector-to-collector load

Power dissipation for transistors operated in class B or AB varies according to signal, and it is necessary to resort to integral calculus for accurate results. However, if a sine wave input is assumed, a reasonable approximation is given by

$$P_d = 0{\cdot}4 \times P_{av} \qquad \ldots (7)$$

This is for two transistors in push-pull; therefore, each transistor dissipates half this power, i.e.:

$$P_d = 0{\cdot}2 \times P_{av} \qquad \ldots (8)$$

Notice that both these last two equations only refer to the average output power and not the maximum output power.

It is also necessary to know the new junction temperature (T_j) once the power dissipation (P_d) and total thermal resistance (θ_{tot}) have been calculated. This will be greater than the ambient temperature (T_{amb}) and is given by

$$T_j = P_d(\theta_{tot}) + T_{amb} \qquad \ldots (9)$$

These are the basic equations necessary for the design or selection of a heat sink, and also for the selection of a suitable transistor type. They will be referred to in a typical design procedure described later in this article.

HEAT SINKS

The simplest form of heat sink is a sheet of metal, usually mounted vertically, with the device mounted in the centre. As the amount of power dissipation in the increased in order to expose more surface area to the device is increased, so the size of the heat sink must be surrounding air.

The heat sink material is an important consideration —copper is somewhat better than aluminium but costs more. In fact, the difference in cost outweighs that in performance and aluminium is probably found more frequently in most applications. With any given heat sink, three factors affect its performance as a heat dissipator. These are effective surface area, position, and surface finish.

Commercially available heat sinks generally have fins so that a greater effective surface area is contained in a smaller volume. With finned heat sinks, the effective surface area may be less than the actual surface area but, as a rule, this is not important since the manufacturers invariably quote figures or a graph giving the thermal resistance of the sink to ambient.

A typical graph showing the dissipating characteristics of a commercially available unit (such as that shown in Fig. 3) is shown in Fig. 4. Sometimes figures are given showing the dissipation capability for certain temperature differentials. These reveal that the thermal resistance becomes slightly higher as the temperature differential increases. For instance, the thermal resistance of a finned heatsink was found to be 8 per cent worse for a temperature differential of 60 degrees C, as opposed to the figure for $\Delta T = 20$ degrees C.

For home made heat sinks, the heat transfer curve shown in Fig. 5 may be used. This relates the thermal resistance (θ_{sa}) against the area of one side of an aluminium sheet, $\frac{1}{8}$in thick, mounted vertically. It can be seen that when the area reaches about 140 square inches the thermal resistance reaches a minimum, and increasing the area beyond this limit does not appreciably change the thermal characteristics.

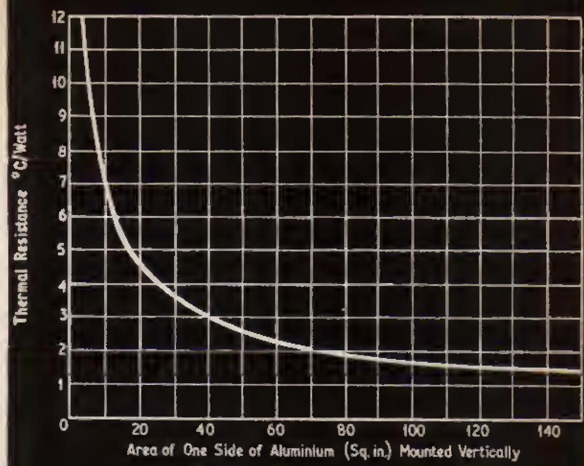


Fig. 5. Heat transfer curve for $\frac{1}{8}$in sheet aluminium

Table 2: TYPICAL APPROXIMATE THERMAL RESISTANCES

Device or Heat Sink	Approximate Thermal Resistance
Vertical Copper Sheet: 140 square inches $\frac{1}{8}$in thick	1·0°C/W
Vertical Aluminium Sheet: 140 square inches $\frac{1}{8}$in thick	1·4°C/W
Typical Finned Heat Sink (Figure 3)	
Mounted vertically	1·8°C/W
Mounted horizontally	2·2°C/W
Typical Power Transistor (TO3 Case or similar)	
Mounting base	0·9°C/W
Cap	6·5°C/W
Typical Low Power Transistor (TO5 Case)	0·2°C/mW
Typical Low Power Transistor (SO-2 Case)	0·3°C/mW
Typical Small Signal Transistor (TO18 Case)	0·3°C/mW
Mica Insulating Washer	
Dry	0·8°C/W to 0·5°C/W
With Silicon grease on both sides	0·4°C/W to 0·25°C/W

The position (i.e. vertical or horizontal) has a marked effect on the characteristics of the heat sink. It is usual to mount the sink vertically and the heat radiated from both sides is carried away by convection currents. When the heat sink is mounted horizontally, heat is dissipated normally from the upper surface but the air underneath is trapped by the sink itself. The graph of Fig. 4 illustrates the difference between horizontal and vertical mounting.

The heat sink is generally mounted externally to the equipment or, when mounted internally, in a ventilated position so that convection currents can carry the hot air upwards. If no ventilation is provided, the heat dissipated from the sink will raise the ambient temperature within the equipment, perhaps to a dangerous level.

The surface finish of the heat sink is not as important as one might suppose. Commercial heat sinks are usually supplied either in bright aluminium or an anodised finish. The best finish is matt black, but on a commercial heat sink a matt black finish may give only an 8 per cent improvement over the plain anodised finish.

It should be noted that the semiconductor device should be bolted as firmly as possible to the heat sink. Ideally, the pressure between the device and the heat sink should be specified (i.e. as fixing screw torque) but this is rarely done. The use of silicon grease between the device and sink will greatly improve the thermal resistance. In some cases, when the case of the transistor is at a different potential to the heat sink, it is necessary to use a mica insulating washer; this increases the thermal resistance between the case and sink (see Table 2) but is necessary in order to insulate electrically the heat sink from the transistor case.

TRANSISTOR ENCAPSULATIONS

Most power transistors have the collector bonded directly to a mounting base of substantial thickness.

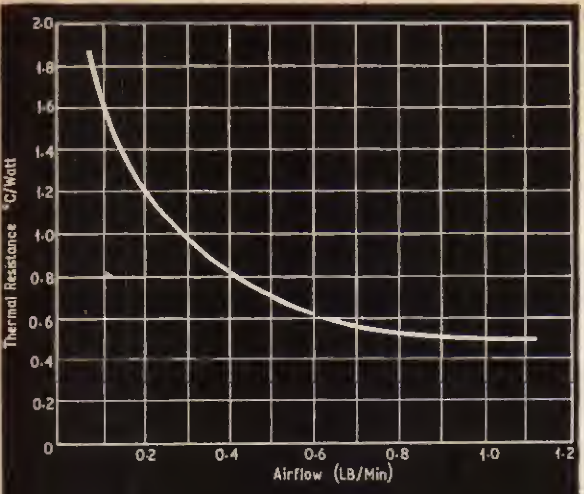


Fig. 6. Heat sink performance with forced airflow

The semiconductor wafer itself is protected by a sheet metal cover bonded to the mounting base. Sometimes the internal space may be filled with a dehydrating agent: in the smaller devices a plastics paste or silicon rubber is used to make the device more robust and improve the thermal conductance between the wafer of the semiconductor material and the case. For this reason, the figures quoted in Table 2 may vary between device types.

If the figures in this table are examined, it can be seen that there is a considerable difference in thermal resistance when the transistor is mounted by its cap to the heat sink, and when it is mounted by its base. This is because the cap is made from thin sheet metal which offers a greater resistance to heat conduction, whereas the base is much more substantial and consequently a better conductor of heat.

Similarly, the smaller transistors have high thermal resistance since they are fabricated from thin sheet metal. It is apparent that there is little point in mounting the smaller transistors on a heat sink since it will do little more than add a safety margin to the device. It will not greatly increase the power dissipation capability from that quoted in free air.

FORCED AIRFLOW

Many commercially available instruments use forced airflow to aid the cooling of devices. This has the effect of helping the action of convection currents and means that a smaller heat sink can be used which can be of importance in miniaturised equipment.

The graph in Fig. 6 shows the improvement which can be obtained by an airflow of 1lb/min with a heat sink that has a still air thermal resistance of 1·7 degrees C per watt ; this figure is reduced by about 60 per cent to 0·6 degrees C per watt. As with the area of the heat sink, a limit is reached above which a higher airflow pressure does not have an appreciably greater effect on the thermal characteristics. This limit is about 1lb of air per minute, and above this figure increased airflow mainly increases the noise level.

Fans are more expensive than sheets of aluminium and are generally only used when there are many (hundreds or even thousands) of components, where it would be both uneconomical and impractical to provide heat sinks for each heat generating component.

DESIGN METHOD

Power Dissipation and Maximum Ratings

For amplifiers, decide on the output power and class of amplification (A, B, or AB) to be employed.

Select the transistor that would seem to be suitable. Use Equations 4 or 6 to check that the maximum collector voltage will not have to be exceeded (remember that the peak voltage in class A will be twice the average voltage).

Find the maximum power that the transistor will have to dissipate using Equations 5 or 8. Check that the maximum collector current rating is not exceeded by using Equation 3. (Again, remember that the peak current in class A will be twice the average current.)

Total Thermal Resistance—θ_{tot}

Using Equation 1, calculate the total thermal resistance. For the moment, use a convenient value for θ_{sa}—this may have to be changed later. θ_{jc} and θ_{cs} should be obtained from the manufacturer's data sheet; they are generally quoted as a combined value since they cannot be modified without redesigning the device. Do not forget to include θ_{ci} if an insulating washer is to be used (a value for this can be obtained from Fig. 3 if necessary) when Equation 2 is used.

Junction Temperature—T_j

Use Equation 9 to find the new junction temperature. The ambient temperature should be the highest that the equipment is likely to be subjected to: 50 degrees C is a realistic figure for domestic equipment.

Derating

Refer to the manufacturer's data sheet and use the figure calculated for the junction temperature to find how much the transistor should be derated at this temperature. If the transistor's maximum power rating, once it has been derated, is exceeded by power which will be dissipated (P_d), then the following courses of action can be taken.

1. Increase the size of the heat sink (i.e. decrease θ_{sa}).
2. If this is not practical, then forced airflow might be suitable for reducing θ_{sa}.
3. Supply voltage (and consequently maximum output power) can be reduced.
4. Select a device with a higher dissipation capability.

If the derated maximum transistor power rating is higher than the power dissipated, then the converse of the above procedure could be considered, i.e. reduce the size of the heat sink.

A suitable heat sink may now be constructed from the graph of Fig. 5 if required: otherwise a commercial heat sink with a thermal resistance equal to, or less than the calculated value may be selected.

CONCLUSION

Precise thermal analysis of a circuit would require a computer to solve all the possible factors involved. This article has attempted to describe a simple approach to the subject and, of necessity, a number of approximations have had to be made. Emphasis has been placed on classes A and AB amplifier circuits since it is here that the constructor will most need to use a heat sink. Calculations for other circuits (i.e. d.c. regulators) may be derived from the equations given. ★