# CHAPTER 5
# THERMAL CONSIDERATIONS

# Thermal considerations Chapter 5

## INTRODUCTION

The perfect power switch is not yet available. All power semiconductors dissipate power internally both during the on-state and during the transition between the on and off states. The amount of power dissipated internally generally speaking increases in line with the power being switched by the semiconductor. The capability of a switch to operate in a particular circuit will therefore depend upon the amount of power dissipated internally and the rise in the operating temperature of the silicon junction that this power dissipation causes. It is therefore important that circuit designers are familiar with the thermal characteristics of power semiconductors and are able to calculate power dissipation limits and junction operating temperatures.

This chapter is divided into three parts. Part One describes the essential thermal properties of semiconductors and explains the concept of a limit, in terms of continuous mode and pulse mode operation. Part Two gives worked examples showing junction temperature calculations for a variety of applied power pulse waveforms. Part Three discusses component heat dissipation and heatsink design.

## PART ONE: THERMAL PROPERTIES

### The power dissipation limit

The maximum allowable power dissipation forms a limit to the safe operating area of power transistors. Power dissipation causes a rise in junction temperature which will, in turn, start chemical and metallurgical changes. The rate at which these changes proceed is exponentially related to temperature, and thus prolonged operation of a power transistor above its junction temperature rating is liable to result in reduced life. Operation of a device at, or below, its power dissipation rating (together with careful consideration of thermal resistances associated with the device) ensures that the junction temperature rating is not exceeded.

All power semiconductors have a power dissipation limitation. For rectifier products such as diodes, thyristors and triacs, the power dissipation rating can be easily translated in terms of current ratings; in the on-state the voltage drop is well defined. Transistors are, however, somewhat more complicated. A transistor, be it a power MOSFET or a bipolar, can operate in its on-state at any voltage up to its maximum rating depending on the circuit conditions. It is therefore necessary to specify a Safe Operating Area (SOA) for transistors which specifies the power dissipation limit in terms of a series of boundaries in the current and voltage plane. These operating areas are usually presented for mounting base temperatures of 25 °C. At higher temperatures, operating conditions must be checked to ensure that junction temperatures are not exceeding the desired operating level.

### Continuous power dissipation

The total power dissipation in a semiconductor may be calculated from the product of the on-state voltage and the forward conduction current. The heat dissipated in the junction of the device flows through the thermal resistance between the junction and the mounting base, $R_{thj\text{-}mb}$. The thermal equivalent circuit of Fig.1 illustrates this heat flow; $P_{tot}$ can be regarded as a thermal current, and the temperature difference between the junction and mounting base $\Delta T_{j\text{-}mb}$ as a thermal voltage. By analogy with Ohm's law, it follows that:

$$P_{tot} = \frac{T_j - T_{mb}}{R_{thj-mb}} \qquad (1)$$

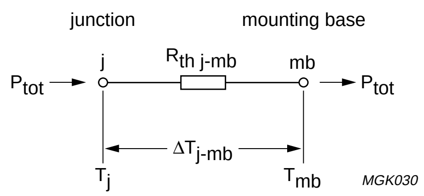


Fig.1 Heat transport in a transistor with power dissipation constant with respect to time.

Figure 2 shows the dependence of the maximum power dissipation on the temperature of the mounting base. $P_{totmax}$ is limited either by a maximum temperature difference:

$$\Delta T_{j-mbmax} = T_{jmax} - T_{mbK} \qquad (2)$$

or by the maximum junction temperature $T_{jmax}$ ($T_{mb\,K}$ is usually 25 °C and is the value of $T_{mb}$ above which the maximum power dissipation must be reduced to maintain the operating point within the safe operating area).

In the first case, $T_{mb} \le T_{mb\,K}$:

$$P_{totmaxK} = \frac{\Delta T_{j-mbmax}}{R_{thj-mb}} \qquad (3)$$

that is, the power dissipation has a fixed limit value ($P_{tot\ max\ K}$ is the maximum DC power dissipation below $T_{mb\ K}$). If the transistor is subjected to a mounting-base temperature $T_{mb\ 1}$, its junction temperature will be less than $T_{jmax}$ by an amount ($T_{mb\ K} - T_{mb\ 1}$), as shown by the broken line in Fig.2.

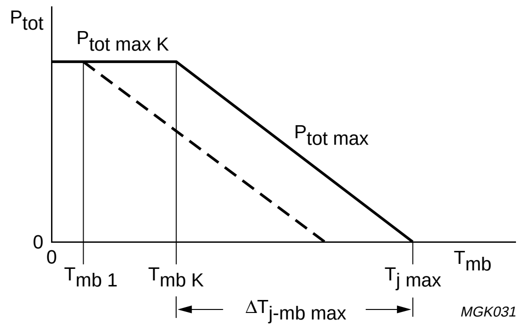


Fig.2 Maximum DC power dissipation in a transistor as a function of the mounting-base temperature.

In the second case, $T_{mb} > T_{mb\ K}$:

$$P_{totmax} = \frac{T_{jmax} - T_{mb}}{R_{thj-mb}} \quad (4)$$

that is, the power dissipation must be reduced as the mounting base temperature increases along the sloping straight line in Fig.2. Equation (4) shows that the lower the thermal resistance $R_{thj\text{-}mb}$, the steeper is the slope of the line. In this case, $T_{mb}$ is the maximum mounting-base temperature that can occur in operation.

### Example

The following data is provided for a particular transistor.

$P_{tot\ max\ K} = 75$ W

$T_{jmax} = 175$ °C

$R_{thj\text{-}mb} \leq 2$ K/W

The maximum permissible power dissipation for continuous operation at a maximum mounting-base temperature of $T_{mb} = 80$ °C is required.

Note that the maximum value of $T_{mb}$ is chosen to be significantly higher than the maximum ambient temperature to prevent an excessively large heatsink being required.

From Equation (4) we obtain:

$$P_{totmax} = \frac{175 - 80}{2}\ \text{W} = 47.5\ \text{W}$$

Provided that the transistor is operated within SOA limits, this value is permissible since it is below $P_{tot\ max\ K}$. The same result can be obtained graphically from the $P_{tot\ max}$ diagram (Fig.3) for the relevant transistor.

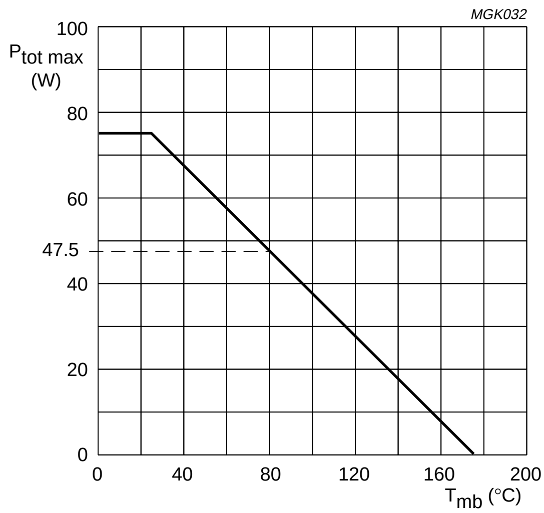


Fig.3 Example of the determination of maximum power dissipation.

### Pulse power operation

When a power transistor is subjected to a pulsed load, higher peak power dissipation is permitted. The materials in a power transistor have a definite thermal capacity, and thus the critical junction temperature will not be reached instantaneously, even when excessive power is being dissipated in the device. The power dissipation limit may be extended for intermittent operation. The size of the extension will depend on the duration of the operation period (that is, pulse duration) and the frequency with which operation occurs (that is, duty factor).

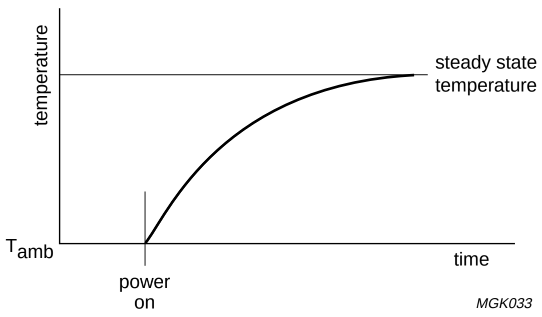


Fig.4 Heating of a transistor chip.

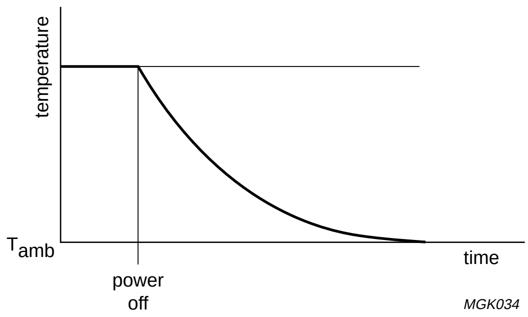


Fig.5 Heating and cooling follow the same law.

If power is applied to a transistor, the device will immediately start to warm up (see Fig.4). If the power dissipation continues, a balance will be struck between heat generation and removal resulting in the stabilization of $T_j$ and $\Delta T_{j\text{-}mb}$. Some heat energy will be stored by the thermal capacity of the device, and the stable conditions will be determined by the thermal resistances associated with the transistor and its thermal environment. When the power dissipation ceases, the device will cool (the heating and cooling laws will be identical, see Fig.5). However, if the power dissipation ceases before the temperature of the transistor stabilizes, the peak values of $T_j$ and $\Delta T_{j\text{-}mb}$ will be less than the values reached for the same level of continuous power dissipation (see Fig.6). If the second pulse is identical to the first, the peak temperature attained by the device at the end of the second pulse will be greater than that at the end of the first pulse. Further pulses will build up the temperature until some new stable situation is attained (see Fig.7). The temperature of the device in this stable condition will fluctuate above and below the mean. If the upward excursions extend into the region of excessive $T_j$ then the life expectancy of the device may be reduced. This can happen with high-power low-duty-factor pulses, even though the average power is below the DC rating of the device.

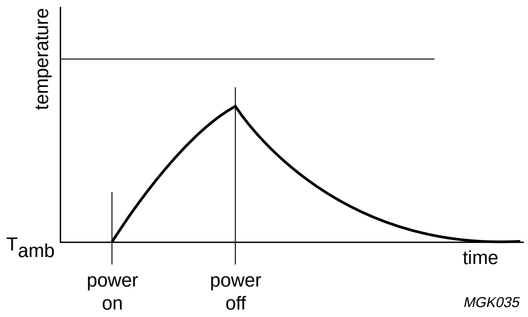


Fig.6 The peak temperature caused by a short power pulse can be less than steady-state temperature resulting from the same power.

Figure 8 shows a typical safe operating area for DC operation of a power MOSFET. The corresponding rectangular- pulse operating areas with a fixed duty factor, $\delta = 0$, and the pulse time $t_p$ as a parameter, are also shown. These boundaries represent the largest possible extension of the operating area for particular pulse times. When the pulse time becomes very short, the power dissipation does not have a limiting action and the pulse current and maximum voltage form the only limits. This rectangle represents the largest possible pulse operating area.

# Thermal considerations Chapter 5

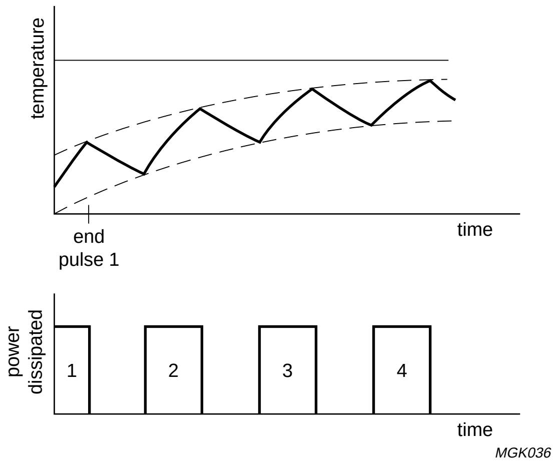


Fig.7 A train of power pulses increases the average temperature if the device does not have time to cool between pulses.

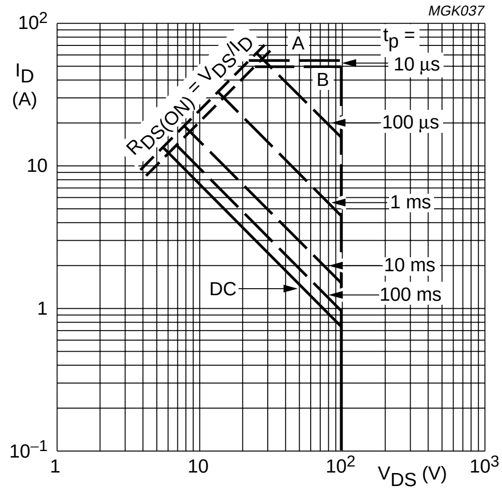


Fig.8 DC and rectangular pulse operating areas with fixed parameters $\delta = 0$, $t_p$ and $T_{mb} = 25$ °C.

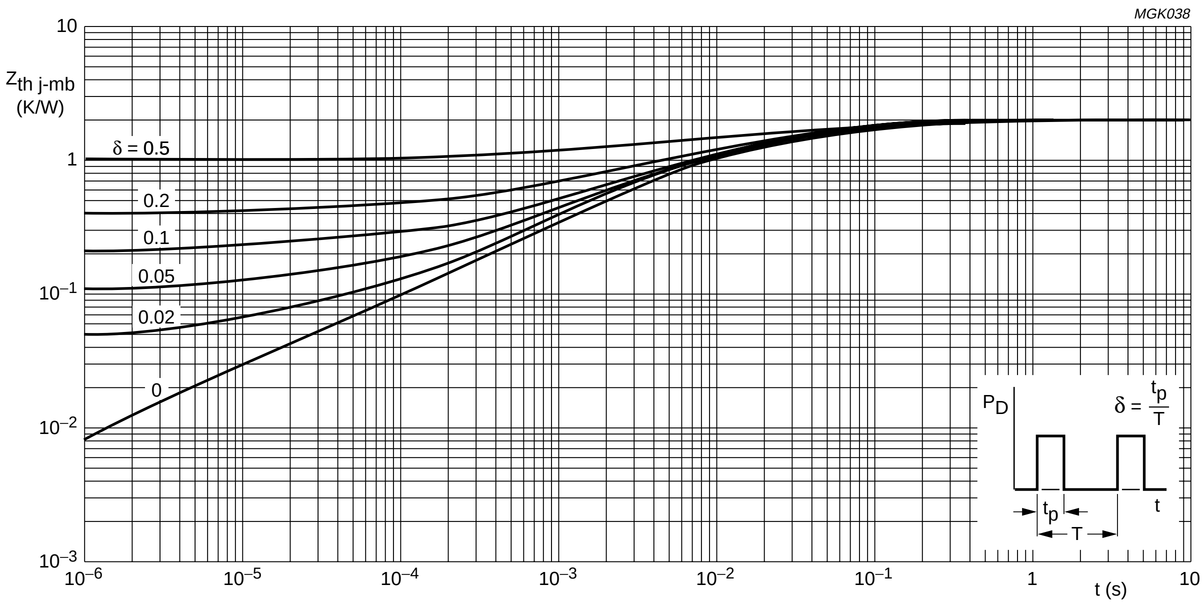


Fig.9 Examples of the transient thermal impedance as a function of the pulse time with duty factor as parameter.

In general, the shorter the pulse and the lower the frequency, the lower the temperature that the junction reaches. By analogy with Equation (3), it follows that:

$$P_{tot\ M} = \frac{T_j - T_{mb}}{Z_{thj-mb}} \quad (5)$$

where $Z_{thj\text{-}mb}$ is the transient thermal impedance between the junction and mounting base of the device. It depends on the pulse duration $t_p$, and the duty factor $\delta$, where:

$$\delta = \frac{t_p}{T} \quad (6)$$

and T is the pulse period. Figure 9 shows a typical family of curves for thermal impedance against pulse duration, with duty factor as a parameter.

Again, the maximum pulse power dissipation is limited either by the maximum temperature difference $\Delta T_{j\text{-}mb\ max}$ (Equation (2)), or by the maximum junction temperature $T_{jmax}$, and so by analogy with Equations (3)and (4):

$$P_{tot\ max\ K} = \frac{\Delta T_{j-mbmax}}{Z_{thj-mb}} \quad (7)$$

when $T_{mb} \le T_{mb\ K}$, and:

$$P_{tot\ M\ max} = \frac{T_{jmax} - T_{mb}}{Z_{thj-mb}} \quad (8)$$

when $T_{mb} > T_{mb\ K}$. That is, below a mounting-base temperature of $T_{mb\ K}$, the maximum power dissipation has a fixed limit value; and above $T_{mb\ K}$, the power dissipation must be reduced linearly with increasing mounting-base temperature.

**Short pulse duration (Fig.10a)**

As the pulse duration becomes very short, the fluctuations of junction temperature become negligible, owing to the internal thermal capacity of the transistor. Consequently, the only factor to be considered is the heating of the junction by the average power dissipation; that is:

$$P_{tot(av)} = \delta P_{totM} \quad (9)$$

The transient thermal impedance becomes:

$$\lim_{t_p \to 0} Z_{thj-mb} = \delta R_{thj-mb} \quad (10)$$

The $Z_{thj\text{-}mb}$ curves approach this value asymptotically as $t_p$ decreases. Figure 9 shows that, for duty factors in the range 0.1 to 0.5, the limit values given by Equation (10) have virtually been reached at $t_p = 10^{-6}$ s.

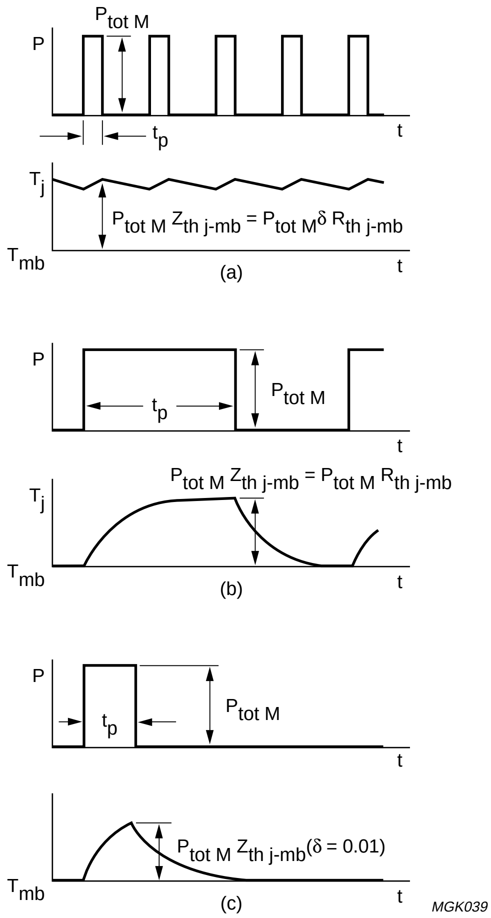


Fig.10 Three limit cases of rectangular pulse loads: (a) short pulse duration, (b) long pulse duration, (c) single-shot pulse.

**Long pulse duration (Fig.10b)**

As the pulse duration increases, the junction temperature approaches a stationary value towards the end of a pulse. The transient thermal impedance tends to the thermal resistance for continuous power dissipation; that is:

$$\lim_{t_p \to \infty} Z_{thj-mb} = R_{thj-mb} \quad (11)$$

Figure 9 shows that $Z_{thj\text{-}mb}$ approaches this value as $t_p$ becomes large. In general, transient thermal effects die out in most power transistors within 0.1 to 1.0 seconds. This time depends on the material and construction of the case, the size of the chip, the way it is mounted, and other factors. Power pulses with a duration in excess of this time have approximately the same effect as a continuous load.

### Single-short pulse (Fig.10c)

As the duty factor becomes very small, the junction tends to cool down completely between pulses so that each pulse can be treated individually. When considering single pulses, the $Z_{thj\text{-}mb}$ values for $\delta = 0$ (Fig.9) give sufficiently accurate results.

## PART TWO: WORKED EXAMPLES

### Calculating junction temperatures

Most applications which include power semiconductors usually involve some form of pulse mode operation. This section gives several worked examples showing how junction temperatures can be simply calculated. Examples are given for a variety of waveforms:

1. periodic waveforms
2. single-shot waveforms
3. composite waveforms
4. a pulse burst
5. non-rectangular pluses.

From the point of view of reliability, it is most important to know what the peak junction temperature will be when the power waveform is applied and also what the average junction temperature is going to be.

Peak junction temperature will usually occur at the end of an applied pulse and its calculation will involve transient thermal impedance. The average junction temperature (where applicable) is calculated by working out the average power dissipation using the DC thermal resistance.

When considering the junction temperature in a device, the following formula is used:

$$T_j = T_{mb} + \Delta T_{j-mb} \qquad (12)$$

where $\Delta T_{j\text{-}mb}$ is found from a rearrangement of Equation (7). In all the following examples the mounting base temperature ($T_{mb}$) is assumed to be 75 °C.

### Periodic rectangular pulse

Figure 11 shows an example of a periodic rectangular pulse. This type of pulse is commonly found in switching applications. 100 W is dissipated every 400 µs for a period of 20 µs, representing a duty cycle (δ) of 0.05.

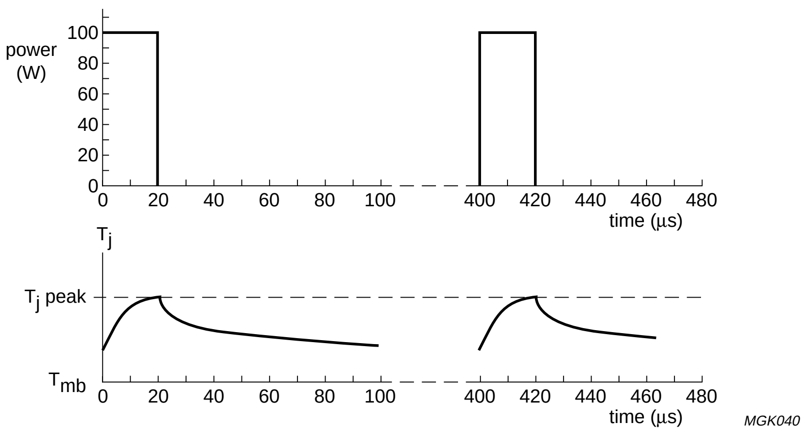


Fig.11 Periodic rectangular pulse.

The peak junction temperature is calculated as follows:

**Peak T $_j$:**

$t = 2\times10^{-5}\ s$

$P = 100\ W$

$\delta = \frac{20}{100} = 0.05$

$Z_{thj-mb} = 0.12\ K/W$

$\Delta T_{j-mb} = P \times Z_{thj-mb} = 100 \times 0.12 = 12\ ^oC$

$T_j = T_{mb} + \Delta T_{j-mb} = 75 + 12 = 87\ ^oC$

**Average T$_j$:**

$P_{av} = P \times \delta = 100 \times 0.05 = 5\ W$

$\Delta T_{j-mb(av)} = P_{av} \times Z_{thj-mb(\delta = 1)} = 5 \times 2 = 10\ ^oC$

$T_{j(av)} = T_{mb} + \Delta T_{j-mb(av)} = 75 + 10 = 85\ ^oC$

The value for $Z_{th\ j\text{-}mb}$ is taken from the δ = 0.05 curve shown in Fig.12 (This diagram repeats Fig.9 but has been simplified for clarity). The above calculation shows that the peak junction temperature will be 85 °C.

**Single shot rectangular pulse**

Figure 13 shows an example of a single shot rectangular pulse. The pulse used is the same as in the previous example, which should highlight the differences between periodic and single shot thermal calculations. For a single shot pulse, the time period between pulses is infinity, i.e. the duty cycle δ = 0. In this example 100 W is dissipated for a period of 20 µs. To work out the peak junction temperature the following steps are used:

$t = 2\times10^{-5}\ s$

$P = 100\ W$

$\delta = 0$

$Z_{thj-mb} = 0.04\ K/W$

$\Delta T_{j-mb} = P \times Z_{thj-mb} = 100 \times 0.04 = 4\ °C$

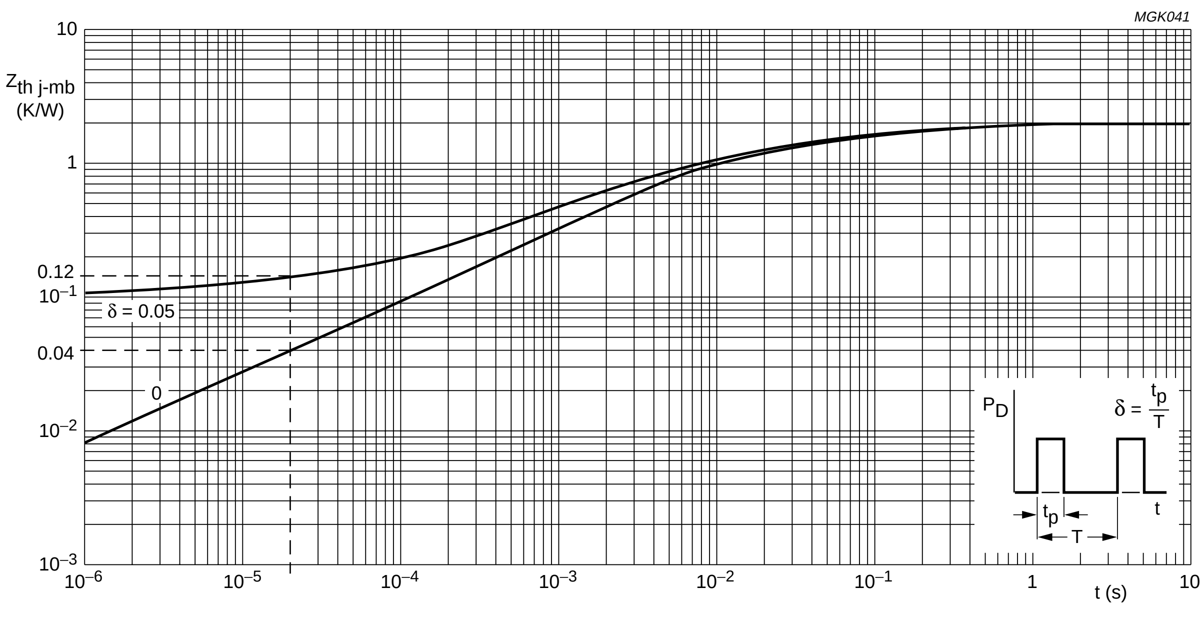


Fig.12 Thermal impedance curves for δ = 0.05 and δ = 0.

The value for $Z_{th\ j\text{-}mb}$ is taken from the $\delta = 0$ curve shown in Fig.12. The above calculation shows that the peak junction temperature will be 4 °C above the mounting base temperature.

For a single shot pulse, the average power dissipated and average junction temperature are not relevant.

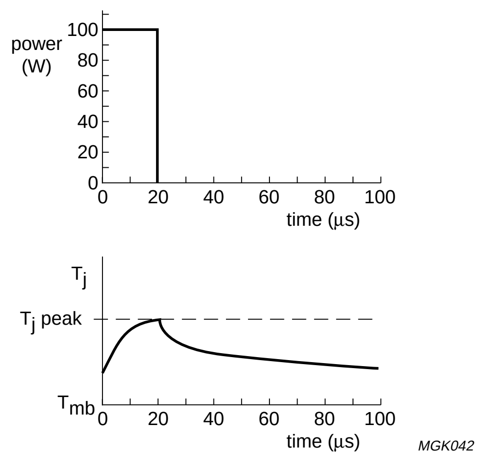


Fig.13 Single shot pulse.

### Composite rectangular pulse

In practice, a power device frequently has to handle composite waveforms, rather than the simple rectangular pulses shown so far. This type of signal can be simulated by superimposing several rectangular pulses which have a common period, but both positive and negative amplitudes, in addition to suitable values of $t_p$ and $\delta$.

By way of an example, consider the composite waveform shown in Fig.14. To show the way in which the method used for periodic rectangular pulses is extended to cover composite waveforms, the waveform shown has been chosen to be an extension of the periodic rectangular pulse example. The period is 400 µs, and the waveform consists of three rectangular pulses, namely 40 W for 10 µs, 20 W for 150 µs and 100 W for 20 µs. The peak junction temperature may be calculated at any point in the cycle. To be able to add the various effects of the pulses at this time, all the pulses, both positive and negative, must end at time $t_x$ in the first calculation and $t_y$ in the second calculation. Positive pulses increase the junction temperature, while negative pulses decrease it.

### Calculation for time $t_x$

$$\Delta T_{j+mb@x} = P_1 \times Z_{thj-mb(t1)} + P_2 \times Z_{thj-mb(t3)} + P_3 \times Z_{thj-mb(t4)} - P_1 \times Z_{thj-mb(t2)} - P_2 \times Z_{thj-mb(t4)} \quad (13)$$

In Equation (15), the values for $P_1$, $P_2$ and $P_3$ are known: $P_1$ = 40 W, $P_2$ = 20 W and $P_3$ = 100 W. The $Z_{th}$ values are taken from Fig.9. For each term in the equation, the equivalent duty cycle must be worked out. For instance the first superimposed pulse in Fig.14 lasts for a time $t_1$ = 180 µs, representing a duty cycle of 180/400 = 0.45 = $\delta$. These values can then be used in conjunction with Fig.9 to find a value for $Z_{th}$, which in this case is 0.9 K/W. Table 1 gives the values calculated for this example.

**Table 1** Composite pulse parameters for time $t_x$

| | | t1<br>180 µs | t2<br>170 µs | t3<br>150 µs | t4<br>20 µs |
|---|---|---|---|---|---|
| Repetitive | $\delta$ | 0.450 | 0.425 | 0.375 | 0.050 |
| T = 400 µs | $Z_{th}$ | 0.900 | 0.850 | 0.800 | 0.130 |
| Single shot | $\delta$ | 0.000 | 0.000 | 0.000 | 0.000 |
| T = ∞ | $Z_{th}$ | 0.130 | 0.125 | 0.120 | 0.040 |

Substituting these values into Equation (15) for $T_{j\text{-}mb@x}$ gives:

Repetitive:

$$\Delta T_{j-mb@x} = 40 \times 0.9 + 20 \times 0.85 + 100 \times 0.13 - 40 \times 0.85 - 20 \times 0.13 = 29.4\ °C$$

$$T_j = T_{mb} + \Delta T_{j-mb} = 75 + 29.4 = 104.4\ °C$$

Single shot:

$$\Delta T_{j-mb@x} = 40 \times 0.13 + 20 \times 0.125 + 100 \times 0.04 - 40 \times 0.125 - 20 \times 0.04 = 5.9\ °C$$

$$T_j = T_{mb} + \Delta T_{j-mb} = 75 + 5.9 = 80.9\ °C$$

Hence the peak values of $T_j$ are 104.4 °C for the repetitive case, and 80.9 °C for the single shot case.

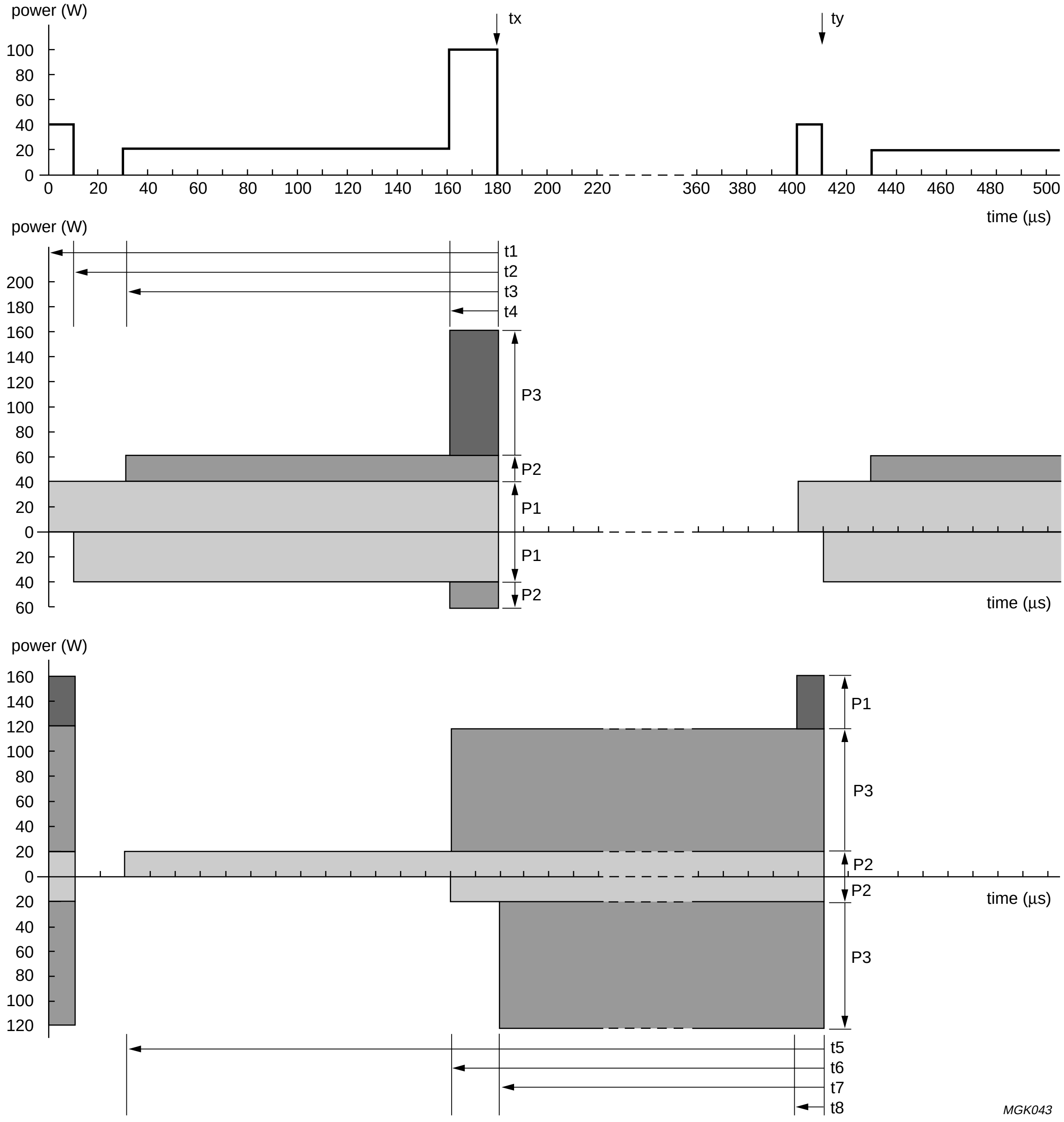


Fig.14 Periodic composite waveform.

**Calculation for time $t_y$**

$$\Delta T_{j+mb@y} = P_2 \times Z_{thj-mb(t5)} + P_3 \times Z_{thj-mb(t6)} + P_1 \times Z_{thj-mb(t8)} - P_2 \times Z_{thj-mb(t6)} - P_3 \times Z_{thj-mb(t7)} \quad (14)$$

Where $Z_{th\text{-}mb(t)}$ is the transient thermal impedance for a pulse time t.

**Table 2** Composite pulse parameters for time $t_y$

| | | t5<br>380 µs | t6<br>250 µs | t7<br>230 µs | t8<br>10 µs |
|---|---|---|---|---|---|
| Repetitive | δ | 0.950 | 0.625 | 0.575 | 0.025 |
| T = 400 µs | $Z_{th}$ | 1.950 | 1.300 | 1.250 | 0.080 |
| Single shot | δ | 0.000 | 0.000 | 0.000 | 0.000 |
| T = ∞ | $Z_{th}$ | 0.200 | 0.160 | 0.150 | 0.030 |

Substituting these values into Equation (15) for $T_{j\text{-}mb(y)}$ gives:

Repetitive:

$$\Delta T_{j-mb(y)} = 20 \times 1.95 + 100 \times 1.3 + 40 \times 0.08 - 20 \times 1.3 - 100 \times 1.25 = 21.2\ ^{o}C$$

$$T_j = T_{mb} + \Delta T_{j-mb} = 75 + 21.2 = 96.2\ ^{o}C$$

Single shot:

$$\Delta T_{j-mb@y} = 20 \times 0.2 + 100 \times 0.16 + 40 \times 0.03 - 20 \times 0.16 - 100 \times 0.15 = 3\ ^{o}C$$

$$T_j = T_{mb} + \Delta T_{j-mb} = 75 + 3 = 78\ ^{o}C$$

Hence the peak values of $T_j$ are 96.2 °C for the repetitive case, and 78 °C for the single shot case.

The average power dissipation and the average junction temperature can be calculated as follows:

$$P_{av} = \frac{25 \times 10 + 5 \times 130 + 20 \times 100}{400} = 7.25\ \text{W}$$

$$\Delta T_{j-mb(av)} = P_{av} \times Z_{th-mb(\delta=1)} = 7.25 \times 2 = 14.5\ ^{\circ}C$$

$$\Delta T_{j(av)} = T_{mb} + \Delta T_{j-mb(av)} = 75 + 14.5 = 89.5\ ^{\circ}C$$

Clearly, the junction temperature at time $t_x$ should be higher than that at time $t_y$, and this is proven in the above calculations.

### Burst pulses

Power devices are frequently subjected to a burst of pulses. This type of signal can be treated as a composite waveform and as in the previous example simulated by superimposing several rectangular pulses which have a common period, but both positive and negative amplitudes, in addition to suitable values of $t_p$ and $\delta$.

Consider the waveform shown in Fig.15. The period is 240 µs, and the burst consists of three rectangular pulses of 100 W power and 20 ms duration, separated by 30 ms. The peak junction temperature will occur at the end of each burst at time $t = t_x = 140$ µs. To be able to add the various effects of the pulses at this time, all the pulses, both positive and negative, must end at time $t_x$. Positive pulses increase the junction temperature, while negative pulses decrease it.

$$\Delta T_{j+mb@x} = P \times Z_{thj-mb(t1)} + P \times Z_{thj-mb(t3)} + P \times Z_{thj-mb(t5)} - P \times Z_{thj-mb(t2)} - P \times Z_{thj-mb(t4)} \quad (15)$$

where $Z_{thj\text{-}mb(t)}$ is the transient thermal impedance for a pulse time t.

The $Z_{th}$ values are taken from Fig.9. For each term in the equation, the equivalent duty cycle must be worked out. These values can then be used in conjunction with Fig.9 to find a value for $Z_{th}$. Table 3 gives the values calculated for this example.

**Table 3** Burst Mode pulse parameters

| | | t1<br>120 µs | t2<br>100 µs | t3<br>70 µs | t4<br>50 µs | t5<br>20 µs |
|---|---|---|---|---|---|---|
| Repetitive | $\delta$ | 0.500 | 0.420 | 0.290 | 0.210 | 0.083 |
| T = 240 µs | $Z_{th}$ | 1.100 | 0.800 | 0.600 | 0.430 | 0.210 |
| Single shot | $\delta$ | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| T = ∞ | $Z_{th}$ | 0.100 | 0.090 | 0.075 | 0.060 | 0.040 |

Substituting these values into Equation (17) gives:

Repetitive:

$$\Delta T_{j-mb@x} = 100 \times 1.10 + 100 \times 0.60 + 100 \times 0.21 - 100 \times 0.80 - 100 \times 0.43 = 68\ °C$$

$$T_j = 75 + 68 = 143\ °C$$

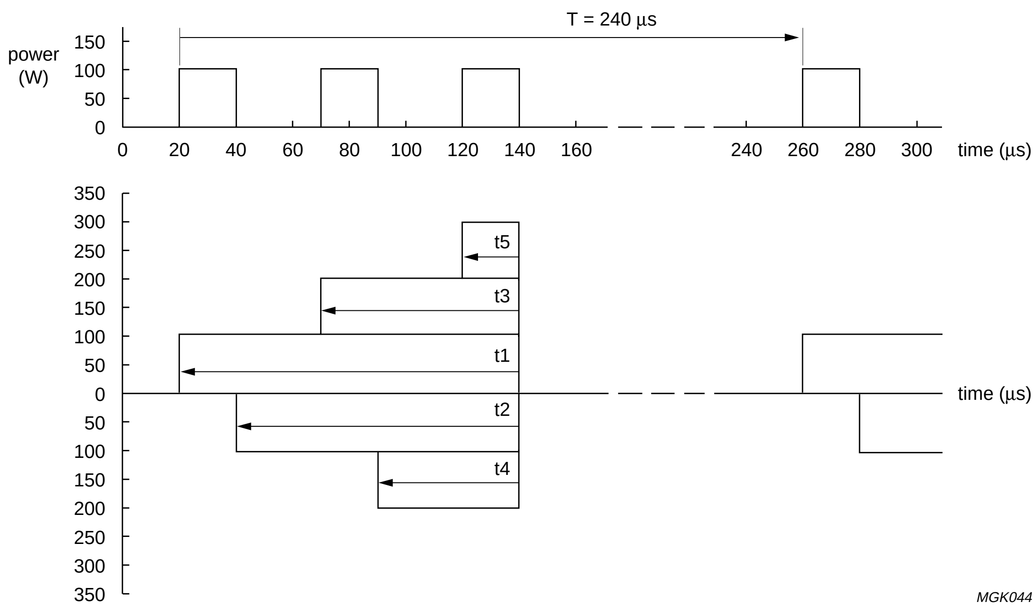


Fig.15 Burst mode waveform.

Single Shot:

$$\Delta T_{j-mb@x} = 100\times0.10+100\times0.075$$
$$+100\times0.04-100\times0.09-100\times0.06$$
$$= 6.5\ ^\circ C$$

$$T_j = 75+6.5 = 81.5\ ^\circ C$$

Hence the peak value of $T_j$ is 143 °C for the repetitive case and 81.5 °C for the single shot case. To calculate the average junction temperature $T_{j(av)}$:

$$P_{av} = \frac{3\times100\times20}{240}$$
$$= 25\ W$$

$$\Delta T_{j-mb(av)} = P_{av}\times Z_{th-mb(\delta=1)}$$
$$= 25\times2 = 50\ ^\circ C$$

$$\Delta T_{j(av)} = 75+50 = 125\ ^\circ C$$

The above example for the repetitive waveform highlights a case where the average junction temperature (125 °C) is well within limits but the composite pulse calculation shows the peak junction temperature to be significantly higher. For reasons of improved long term reliability it is usual to operate devices with a peak junction temperature below 125 °C.

**Non-rectangular pulses**

So far, the worked examples have only covered rectangular waveforms. However, triangular, trapezoidal and sinusoidal waveforms are also common. In order to apply the above thermal calculations to non rectangular waveforms, the waveform is approximated by a series of rectangles. Each rectangle represents part of the waveform. The equivalent rectangle must be equal in area to the section of the waveform it represents (i.e. the same energy) and also be of the same peak power. With reference to Fig.16, a triangular waveform has been approximated to one rectangle in the first example, and two rectangles in the second. Obviously, increasing the number of sections the waveform is split into will improve the accuracy of the thermal calculations.

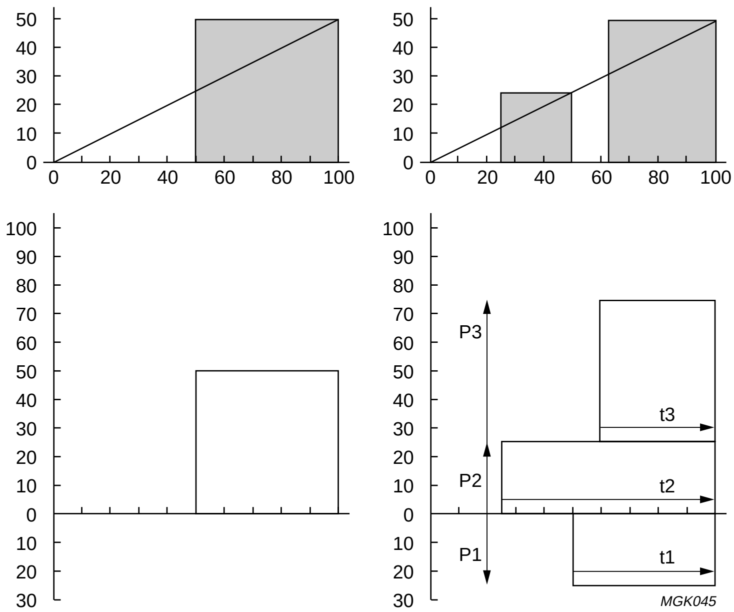


Fig.16 Non-rectangular waveform.

In the first example, there is only one rectangular pulse, of duration 50 µs, dissipating 50 W. So again using Equation (14) and a rearrangement of Equation (7):

$$\Delta T_{j-mb} = P_{tot\,M} \times Z_{thj-mb}$$

Single shot:

$$\Delta T_{j-mb} = 50 \times 0.065 = 3.25\ ^\circ C$$

$$\Delta T_{jpeak} = 75 + 3.25 = 78.5\ ^\circ C$$

10% duty cycle:

$$\Delta T_{j-mb} = 50 \times 0.230 = 11.5\ ^\circ C$$

$$\Delta T_{jpeak} = 75 + 11.5 = 86.5\ ^\circ C$$

50% duty cycle:

$$\Delta T_{j-mb} = 50 \times 1.000 = 50\ ^\circ C$$

$$\Delta T_{jpeak} = 75 + 50 = 125\ ^\circ C$$

When the waveform is split into two rectangular pulses:

$$\Delta T_{j\text{–}mb} = P_3 \times Z_{thj-mb(t3)} + P_1 \times Z_{thj-mb(t2)} - P_2 \times Z_{thj-mb(t2)} \qquad (16)$$

For this example $P_1$ = 25 W, $P_2$ = 25 W, $P_3$ = 50 W. Table 4 shows the rest of the parameters.

**Table 4** Non-rectangular pulse calculations

| | | t1<br>75 µs | t2<br>50 µs | t3<br>37.5 µs |
|---|---|---|---|---|
| Single shot | $\delta$ | 0.000 | 0.000 | 0.000 |
| T = ∞ | $Z_{th}$ | 0.085 | 0.065 | 0.055 |
| 10% duty cycle | $\delta$ | 0.075 | 0.050 | 0.037 |
| T = 1000 µs | $Z_{th}$ | 0.210 | 0.140 | 0.120 |
| 50% duty cycle | $\delta$ | 0.375 | 0.250 | 0.188 |
| T = 200 µs | $Z_{th}$ | 0.700 | 0.500 | 0.420 |

Substituting these values into Equation (18) gives:

Single shot:

$$\Delta T_{j-mb} = 50 \times 0.055 + 25 \times 0.85 - 25 \times 0.65 = 3.25\ ^\circ C$$

$$\Delta T_{jpeak} = 75 + 3.25 = 78.5\ ^\circ C$$

10% Duty cycle

$$\Delta T_{j-mb} = 50 \times 0.12 + 25 \times 0.21 - 25 \times 0.14 = 7.75\ ^\circ C$$

$$\Delta T_{jpeak} = 75 + 7.75 = 82.5\ ^\circ C$$

50% Duty cycle

$$\Delta T_{j-mb} = 50 \times 0.42 + 25 \times 0.7 - 25 \times 0.5 = 26\ ^\circ C$$

$$\Delta T_{jpeak} = 75 + 26 = 101\ ^\circ C$$

To calculate the average junction temperature:

$$P_{av} = \frac{50 \times 50}{1000} = 2.5\ W$$

$$\Delta T_{j-mb(av)} = P_{av} \times Z_{th-mb(\delta = 1)} = 2.5 \times 2 = 5\ ^\circ C$$

$$\Delta T_{j(av)} = 75 + 5 = 80\ ^\circ C$$

**Conclusion to part two**

A method has been presented to allow the calculation of average and peak junction temperatures for a variety of pulse types. Several worked examples have shown calculations for various common waveforms. The method for non-rectangular pulses can be applied to any wave shape, allowing temperature calculations for waveforms such as exponential and sinusoidal power pulses. For pulses such as these, care must be taken to ensure that the calculation gives the peak junction temperature, as it may not occur at the end of the pulse. In this instance several calculations must be performed with different endpoints to find the maximum junction temperature.

### PART 3: HEAT DISSIPATION

All semiconductor failure mechanisms are temperature dependent and so the lower the junction temperature, the higher the reliability of the circuit. Thus our data specifies a maximum junction temperature which should not be exceeded under the worst probable conditions. However, derating the operating temperature from $T_{jmax}$ is always desirable to improve the reliability still further. The junction temperature depends on both the power dissipated in the device and the thermal resistances (or impedances) associated with the device. Thus careful consideration of these thermal resistances (or impedances) allows the user to calculate the maximum power dissipation that will keep the junction temperature below a chosen value.

The formulae and diagrams given in this part can only be considered as a guide for determining the nature of a heatsink. This is because the thermal resistance of a heatsink depends on numerous parameters which cannot be predetermined. They include the position of the transistor on the heatsink, the extent to which air can flow unhindered, the ratio of the lengths of the sides of the heatsink, the screening effect of nearby components, and heating from these components. It is always advisable to check important temperatures in the finished equipment under the worst probable operating conditions. The more complex the heat dissipation conditions, the more important it becomes to carry out such checks.

#### Heat flow path

The heat generated in a semiconductor chip flows by various paths to the surroundings. Small signal devices do not usually require heatsinking; the heat flows from the junction to the mounting base which is in close contact with the case. Heat is then lost by the case to the surroundings by convection and radiation (Fig.17a). Power transistors, however, are usually mounted on heatsinks because of the higher power dissipation they experience. Heat flows from the transistor case to the heatsink by way of contact pressure, and the heatsink loses heat to the surroundings by convection and radiation, or by conduction to cooling water (Fig.17b). Generally air cooling is used so that the ambient referred to in Fig.17 is usually the surrounding air. Note that if this is the air inside an equipment case, the additional thermal resistance between the inside and outside of the equipment case should be taken into account.

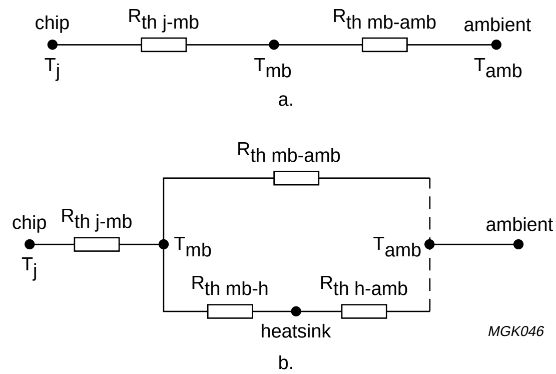


Fig.17 Thermal resistance in the heat flow process: (a) without heatsink, (b) with heatsink.

#### Contact thermal resistance $R_{th\ mb\text{-}h}$

The thermal resistance between the transistor mounting base and the heatsink depends on the quality and size of the contact areas, the type of any intermediate plates used, and the contact pressure. Care should be taken when drilling holes in heatsinks to avoid burring and distorting the metal, and both mating surfaces should be clean. Paint finishes of normal thickness, up to 50 µm (as a protection against electrolytic voltage corrosion), barely affect the thermal resistance. Transistor case and heatsink surfaces can never be perfectly flat, and so contact will take place on several points only, with a small air-gap over the rest of the area. The use of a soft substance to fill this gap lowers the contact thermal resistance. Normally, the gap is filled with a heatsinking compound which remains fairly viscous at normal transistor operating temperatures and has a high thermal conductivity. The use of such a compound also prevents moisture from penetrating between the contact surfaces. Proprietary heatsinking compounds are available which consist of a silicone grease loaded with some electrically insulating good thermally conducting powder such as alumina.The contact thermal resistance $R_{th\ mb\text{-}h}$ is usually small with respect to ($R_{th\ j\text{-}mb}$ +$R_{th\ h\text{-}amb}$) when cooling is by natural convection. However, the heatsink thermal resistance $R_{th\ h\text{-}amb}$ can be very small when either forced ventilation or water cooling are used, and thus a close thermal contact between the transistor case and heatsink becomes particularly important.

### Thermal resistance calculations

Fig.17a shows that, when a heatsink is not used, the total thermal resistance between junction and ambient is given by:

$$R_{th\ j\text{-}amb} = R_{th\ j\text{-}mb} + R_{th\ mb\text{-}amb} \quad (17)$$

However, power transistors are generally mounted on a heatsink since $R_{th\ j\text{-}amb}$ is not usually small enough to maintain temperatures within the chip below desired levels.

Fig.17b shows that, when a heatsink is used, the total thermal resistance is given by:

$$R_{th\ j\text{-}amb} = R_{th\ j\text{-}mb} + R_{th\ mb\text{-}h} + R_{th\ h\text{-}amb} \quad (18)$$

Note that the direct heat loss from the transistor case to the surroundings through $R_{th\ mb\text{-}amb}$ is negligibly small.

The first stage in determining the size and nature of the required heatsink is to calculate the maximum heatsink thermal resistance $R_{th\ h\text{-}amb}$ that will maintain the junction temperature below the desired value.

### Continuous operation

Under DC conditions, the maximum heatsink thermal resistance can be calculated directly from the maximum desired junction temperature.

$$R_{th\ j\text{-}amb} = \frac{T_j - T_{amb}}{P_{tot(av)}} \quad (19)$$

and

$$R_{th\ j\text{-}mb} = \frac{T_j - T_{mb}}{P_{tot(av)}} \quad (20)$$

Combining Equations (18) and (19) gives:

$$R_{th\ h\text{-}amb} = \frac{T_j - T_{amb}}{P_{tot(av)}} - R_{th\ j\text{-}mb} - R_{th\ mb\text{-}h} \quad (21)$$

and substituting Equation (20) into Equation (21) gives:

$$R_{th\ h\text{-}amb} = \frac{T_{mb} - T_{amb}}{P_{tot(av)}} - R_{th\ mb\text{-}h} \quad (22)$$

The values of $R_{th\ j\text{-}mb}$ and $R_{th\ mb\text{-}h}$ are given in the published data. Thus, either Equation (21) or Equation (22) can be used to find the maximum heatsink thermal resistance.

### Intermittent operation

The thermal equivalent circuits of Fig.17 are inappropriate for intermittent operation, and the thermal impedance $Z_{th\ j\text{-}mb}$ should be considered.

$$P_{totM} = \frac{T_j - T_{mb}}{Z_{th\ j\text{-}mb}}$$

thus:

$$T_{mb} = T_j - P_{totM} \times Z_{th\ j\text{-}mb} \quad (23)$$

The mounting-base temperature has always been assumed to remain constant under intermittent operation. This assumption is known to be valid in practice provided that the pulse time is less than about one second. The mounting-base temperature does not change significantly under these conditions as indicated in Fig.18. This is because heatsinks have a high thermal capacity and thus a high thermal time-constant.

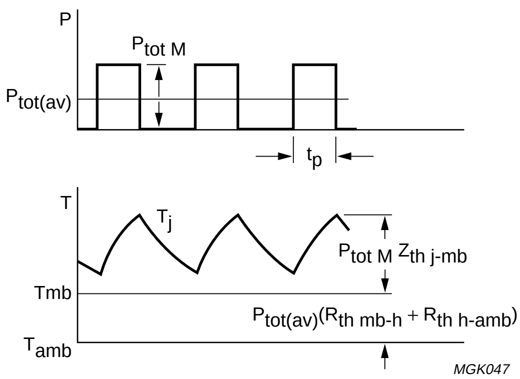


Fig.18 Variation of junction and mounting base temperature when the pulse time is small compared with the thermal time-constant of the heatsink.

Thus Equation (22) is valid for intermittent operation, provided that the pulse time is less than one second. The value of $T_{mb}$ can be calculated from Equation (23), and the heatsink thermal resistance can be obtained from Equation (22).

The thermal time constant of a transistor is defined as that time at which the junction temperature has reached 70% of its final value after being subjected to a constant power dissipation at a constant mounting base temperature.

Now, if the pulse duration $t_p$ exceeds one second, the transistor is temporarily in thermal equilibrium since such a pulse duration is significantly greater than the thermal time-constant of most transistors. Consequently, for pulse times of more than one second, the temperature difference $T_j$ -$T_{mb}$ reaches a stationary final value (Fig.19) and Equation (23) should be replaced by:

$$T_{mb} = T_j - P_{totM} \times R_{th\ j\text{-}mb} \qquad (24)$$

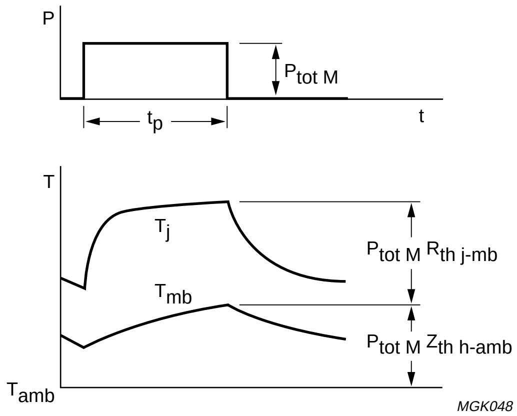


Fig.19 Variation of junction and mounting base temperature when the pulse time is not small compared with he thermal time-constant of the heatsink.

In addition, it is no longer valid to assume that the mounting base temperature is constant since the pulse time is also no longer small with respect to the thermal time constant of the heatsink.

### Smaller heatsinks for intermittent operation

In many instances, the thermal capacity of a heatsink can be utilized to design a smaller heatsink for intermittent operation than would be necessary for the same level of continuous power dissipation. The average power dissipation in Equation (22) is replaced by the peak power dissipation to obtain the value of the thermal impedance between the heatsink and the surroundings.

$$Z_{th\ h\text{-}amb} = \frac{T_{mb} - T_{amb}}{P_{totM}} - R_{th\ mb\text{-}h} \qquad (25)$$

The value of $Z_{th\ h\text{-}amb}$ will be less than the comparable thermal resistance and thus a smaller heatsink can be designed than that obtained using the too large value calculated from Equation (22).

### Heatsinks

Three varieties of heatsink are in common use: flat plates (including chassis), diecast finned heatsinks, and extruded finned heatsinks. The material normally used for heatsink construction is aluminium although copper may be used with advantage for flat-sheet heatsinks. Small finned clips are sometimes used to improve the dissipation of low-power transistors.

### Heatsink finish

Heatsink thermal resistance is a function of surface finish. A painted surface will have a greater emissivity than a bright unpainted one. The effect is most marked with flat plate heatsinks, where about one third of the heat is dissipated by radiation. The colour of the paint used is relatively unimportant, and the thermal resistance of a flat plate heatsink painted gloss white will be only about 3% higher than that of the same heatsink painted matt black. With finned heatsinks, painting is less effective since heat radiated from most fins will fall on adjacent fins but it is still worthwhile. Both anodising and etching will decrease the thermal resistivity. Metallic type paints, such as aluminium paint, have the lowest emissivities, although they are approximately ten times better than a bright aluminium metal finish.

### Flat-plate heatsinks

The simplest type of heatsink is a flat metal plate to which the transistor is attached. Such heatsinks are used both in the form of separate plates and as the equipment chassis itself. The thermal resistance obtained depends on the thickness, area and orientation of the plate, as well as on the finish and power dissipated. A plate mounted horizontally will have about twice the thermal resistance of a vertically mounted plate. This is particularly important where the equipment chassis itself is used as the heatsink. In Fig.20, the thermal resistance of a blackened heatsink is plotted against surface area (one side) with power dissipation as a parameter. The graph is accurate to within 25% for nearly square plates, where the ratio of the lengths of the sides is less than 1.25:1.

## Thermal considerations Chapter 5

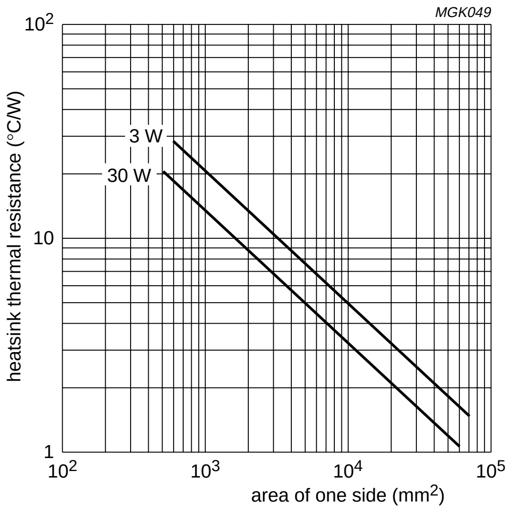


Fig.20 Generalized heatsink characteristics: flat vertical black aluminium, 3 mm thick, approximately square.

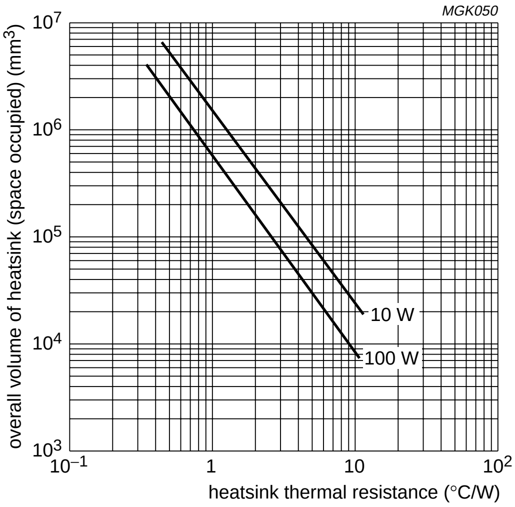


Fig.21 Generalized heatsink characteristics: blackened aluminium finned heatsinks.

### Finned heatsinks

Finned heatsinks may be made by stacking flat plates, although it is usually more economical to use ready made diecast or extruded heatsinks. Since most commercially available finned heatsinks are of reasonably optimum design, it is possible to compare them on the basis of the overall volume which they occupy. This comparison is made in Fig.21 for heatsinks with their fins mounted vertically; again, the graph is accurate to 25%.

### Heatsink dimensions

The maximum thermal resistance through which sufficient power can be dissipated without damaging the transistor can be calculated as discussed previously. This section explains how to arrive at a type and size of heatsink that gives a sufficiently low thermal resistance.

### Natural air cooling

The required size of aluminium heatsinks - whether flat or extruded (finned) can be derived from the nomogram in Fig.22. Like all heatsink diagrams, the nomogram does not give exact values for $R_{th\ h\text{-}amb}$ as a function of the dimensions since the practical conditions always deviate to some extent from those under which the nomogram was drawn up. The actual values for the heatsink thermal resistance may differ by up to 10% from the nomogram values. Consequently, it is advisable to take temperature measurements in the finished equipment, particularly where the thermal conditions are critical.

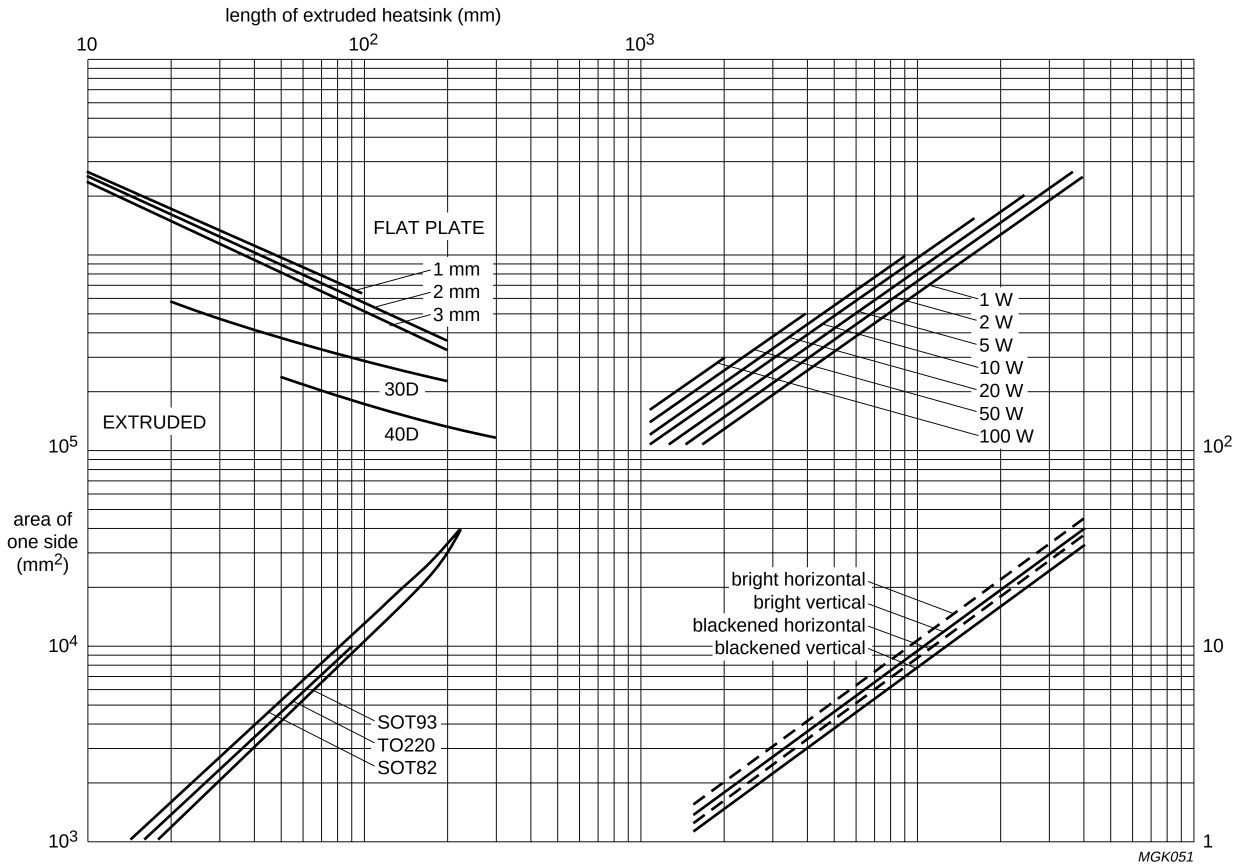


Fig.22 Heatsink nomogram.

The conditions to which the nomogram applies are as follows:

- natural air cooling (unimpeded natural convection with no build up of heat)
- ambient temperature about 25 °C, measured about 50 mm below the lower edge of the heatsink (see Fig.23)
- single mounting (that is, not affected by nearby heatsinks)
- atmospheric pressure about 10 N/m$^2$
- distance between the bottom of the heatsink and the base of a draught-free space about 100 mm (see Fig.23)
- transistor mounted roughly in the centre of the heatsink (this is not so important for finned heatsinks because of the good thermal conduction).

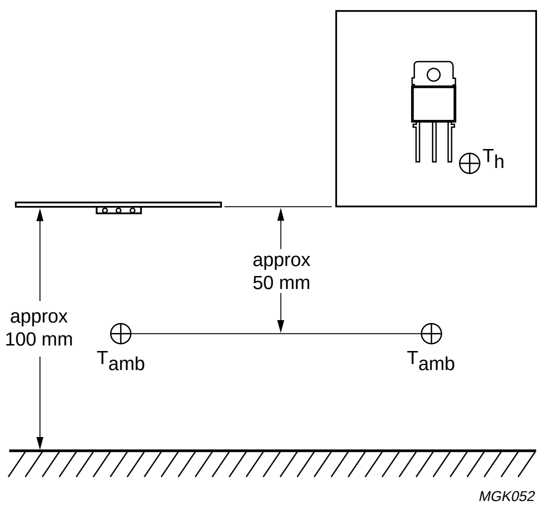


Fig.23 Conditions applicable to nomogram shown in Fig.22.

The appropriately-sized heatsink is found as follows.

1. Enter the nomogram from the right hand side of section 1 at the appropriate $R_{th\ h\text{-}amb}$ value (see Fig.24). Move horizontally to the left, until the appropriate curve for orientation and surface finish is reached.
2. Move vertically upwards to intersect the appropriate power dissipation curve in section 2.
3. Move horizontally to the left into section 3 for the desired thickness of a flat-plate heatsink, or the type of extrusion.
4. If an extruded heatsink is required, move vertically upwards to obtain its length (Figs 25a and 25b give the outlines of the extrusions).
5. If a flat-plate heatsink is to be used, move vertically downwards to intersect the appropriate curve for envelope type in section 4.
6. Move horizontally to the left to obtain heatsink area.
7. The heatsink dimensions should not exceed the ratio of 1.25:1.

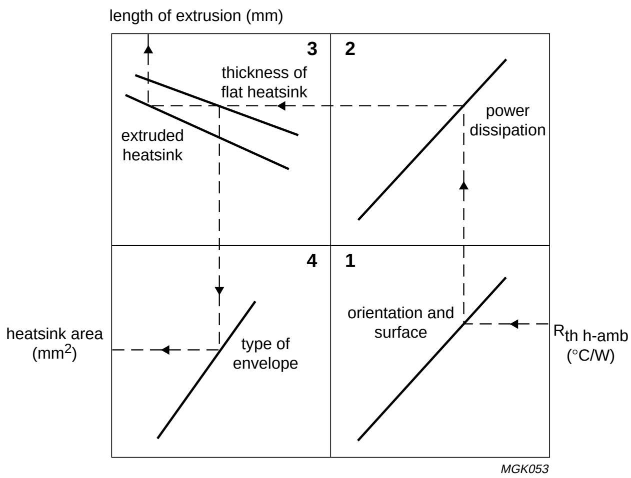


Fig.24 Use of the heatsink nomogram.

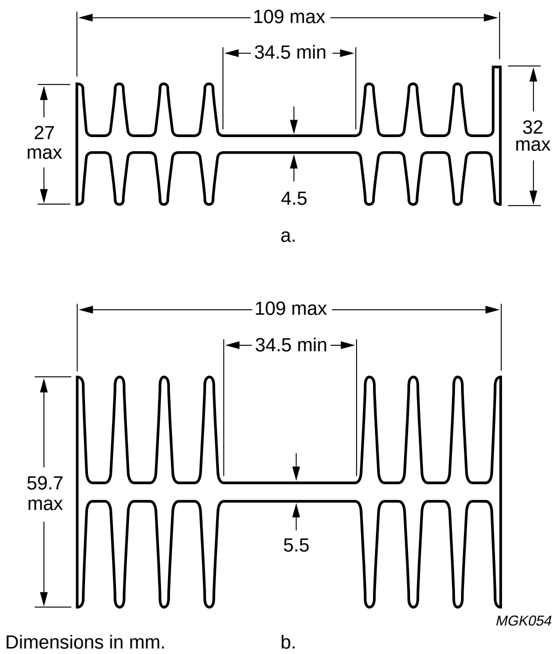


Fig.25 Outline of extrusion: a = 30D, b = 40D.

The curves in section 2 take account of the non linear nature of the relationship between the temperature drop across the heatsink and the power dissipation loss. Thus, at a constant value of the heatsink thermal resistance, the greater the power dissipation, the smaller is the required size of heatsink. This is illustrated by the following example.

**Example**

An extruded heatsink mounted vertically and with a painted surface is required to have a maximum thermal resistance of $R_{th\ h\text{-}amb}$ = 2.6 °C/W at the following powers:

a) $P_{tot\ (av)}$ = 5 W

b) $P_{tot\ (av)}$ = 50 W

Enter the nomogram at the appropriate value of the thermal resistance in section 1, and via either the 50 W or 5 W line in section 2, the appropriate lengths of the extruded heatsink 30D are found to be:

a) length = 110 mm

b) length = 44 mm

Case (b) requires a shorter length since the temperature difference is ten times greater than in case (a).

As the ambient temperature increases beyond 25 °C, so does the temperature of the heatsink and thus the thermal resistance (at constant power) decreases owing to the increasing role of radiation in the heat removal process. Consequently, a heatsink with dimensions derived from Fig.22 at $T_{amb}$ > 25 °C will be more than adequate. If the maximum ambient temperature is less than 25 °C, then the thermal resistance will increase slightly. However, any increase will lie within the limits of accuracy of the nomogram and within the limits set by other uncertainties associated with heatsink calculations.

For heatsinks with relatively small areas, a considerable part of the heat is dissipated from the transistor case. This is why the curves in section 4 tend to flatten out with decreasing heatsink area. The area of extruded heatsinks is always large with respect to the surface of the transistor case, even when the length is small.

If several transistors are mounted on a common heatsink, each transistor should be associated with a particular section of the heatsink (either an area or length according to type) whose maximum thermal resistance is calculated from Equations (21) or (22); that is, without taking the heat produced by nearby transistors into account. From the sum of these areas or lengths, the size of the common heatsink can then be obtained. If a flat heatsink is used, the transistors are best arranged as shown in Fig.26. The maximum mounting base temperatures of transistors in such a grouping should always be checked once the equipment has been constructed.

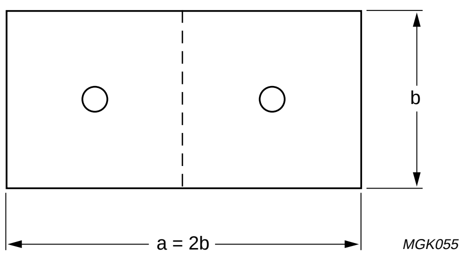


Fig.26 Arrangement of two equally loaded transistors mounted on a common heatsink.

# Thermal considerations Chapter 5

## Forced air cooling

If the thermal resistance needs to be much less than 1 °C/W, or the heatsink not too large, forced air cooling by means of fans can be provided. Apart from the size of the heatsink, the thermal resistance now only depends on the speed of the cooling air. Provided that the cooling air flows parallel to the fins and with sufficient speed (>0.5 m/s), the thermal resistance hardly depends on the power dissipation and the orientation of the heatsink. Note that turbulence in the air current can result in practical values deviating from theoretical values.

Figure 27 shows the form in which the thermal resistances for forced air cooling are given in the case of extruded heatsinks. It also shows the reduction in thermal resistance or length of heatsink which may be obtained with forced air cooling.

The effect of forced air cooling in the case of flat heatsinks is seen from Fig.28. Here, too, the dissipated power and the orientation of the heatsink have only a slight effect on the thermal resistance, provided that the air flow is sufficiently fast.

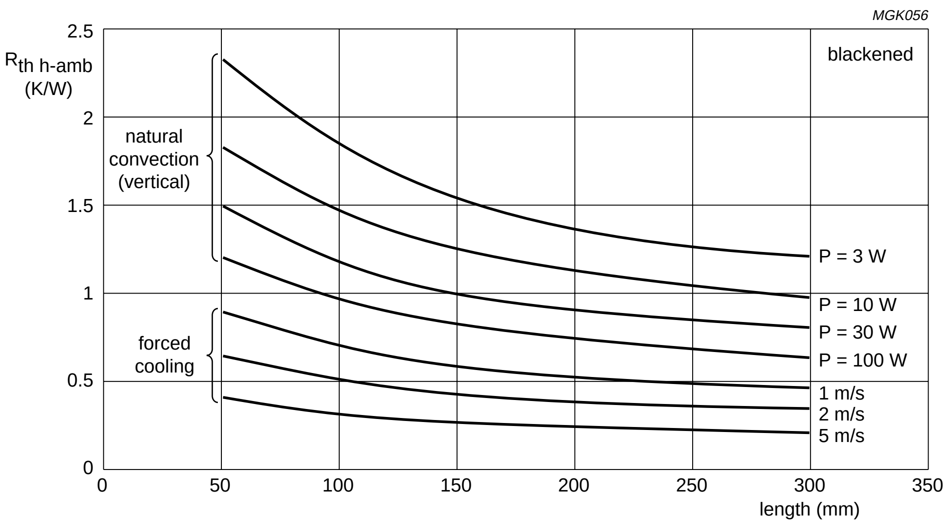


Fig.27 Thermal resistance of a finned heatsink (type 40D) as a function of the length, with natural and forced air cooling.

# Thermal considerations Chapter 5

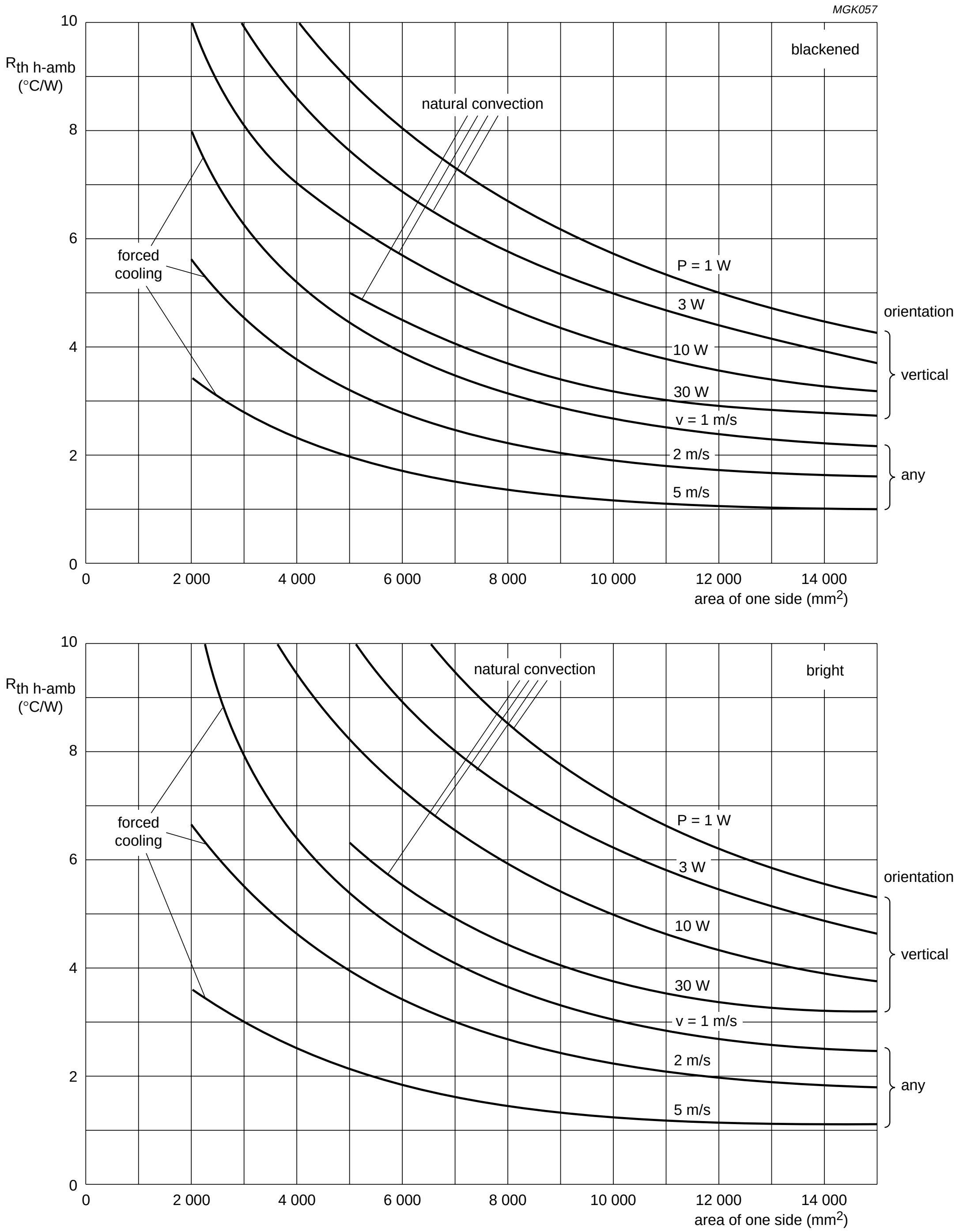


Fig.28 Thermal resistance of heatsinks (2 mm thick copper or 3 mm aluminium) under natural convection and forced cooling conditions, with a SOT93 package.

**Conclusion to part three**

The majority of power transistors require heatsinking, and when the maximum thermal resistance that will maintain the device's junction temperature below its rating has been calculated, a heatsink of appropriate type and size can be chosen. The practical conditions under which a transistor will be operated are likely to differ from the theoretical considerations used to determine the required heatsink, and so temperatures should always be checked in the finished equipment. Finally, some applications require a small heatsink, or one with a very low thermal resistance, in which case forced air cooling by means of fans should be provided.