12

Thermoelastic contact

12.1 Introduction

Classical elastic contact stress theory concerns bodies whose temperature is uniform. Variation in temperature within the bodies may, of itself, give rise to thermal stresses but may also change the contact conditions through thermal distortion of their surface profiles. For example if two non-conforming bodies, in contact over a small area, are maintained at different temperatures, heat will flow from the hot body to the cold one through the 'constriction' presented by their contact area. The gap between their surfaces where they do not touch will act more or less as an insulator. The interface will develop an intermediate temperature which will lie above that of the cold body, so that thermal expansion will cause its profile to become more convex in the contact region. Conversely the interface temperature will lie below that of the hot body, so that thermal contraction will lead to a less convex or a concave profile. Only if the material of the two bodies is similar, both elastically and thermally, will the expansion of the one exactly match the contraction of the other; otherwise the thermal distortion will lead to a change in the contact area and contact pressure distribution. This problem will be examined in §4 below.

A somewhat different situation arises when heat is generated at or near to the interface of bodies in contact. An obvious example of practical importance is provided by frictional heating at sliding contacts. Also inelastic deformation in rolling contact liberates heat beneath the surface which, with poorly conducting materials, can lead to severe thermal stresses. The passage of a heavy electric current between non-conforming surfaces in contact leads to a high current density and local heating at the contact constriction.

The analysis of a thermoelastic contact problem consists of three parts: (i) the analysis of heat conduction to determine the temperature distribution in the two contacting bodies; (ii) the analysis of the thermal expansion of the

https://doi.org/10.1017/CBO9781139171731.013 Published online by Cambridge University Press

bodies, to determine the thermal distortion of their surface profiles; (iii) the isothermal contact problem to find the contact stresses resulting from the deformed profiles. In the simplest cases these three aspects are uncoupled and the analysis can proceed in the above sequence. In many cases these aspects are not independent. Where heat is generated by sliding friction, for example, the distribution of heat liberated at the interface, which governs the temperature distribution, is proportional to the distribution of contact pressure, which itself depends on the thermoelastic distortion of the solids. Nevertheless our discussion of thermoelastic contact will follow that sequence.

Readers of this book will have already appreciated the advantages, when calculating elastic deformations, of representing bodies in contact, whatever their actual profile, by an elastic half-space bounded by a plane surface. The same idealisation is also helpful in calculating the temperature distributions in thermal problems. It may be justified in the same way; the temperature *gradients* which give rise to thermal stress and distortion are large only in the vicinity of the contact region where the actual surfaces of the bodies are approximately plane. Widespread changes in the temperature of the bulk of the bodies lead only to overall and approximately uniform expansion or contraction which neither introduces thermal stresses nor significantly changes the profile in the contact zone.

12.2 Temperature distributions in a conducting half-space

The theory of heat conduction in solids is not the concern of this book. A full account of the theory and the analysis of most of the problems we require are contained in the book by Carslaw & Jaeger (1959). Only the results will be summarised here. We are interested in the flow of heat into a half-space through a restricted area of the surface. We shall start with the temperature distribution in the half-space due to a 'point source' of heat located at the surface. Since the conduction equations are linear the temperature distribution due to any distribution of heat supplied to the surface can be found by the superposition of the solution for 'point sources' in the same way as elastic stress distributions due to surface tractions were found from 'point force' solutions in Chapters 2 and 3.

The half-space is taken to be uniform with conductivity k, density ρ, specific heat capacity c and thermal diffusivity $\kappa (= k/\rho c)$.

(a) Instantaneous point source

A quantity of heat H is liberated instantaneously at time $t = 0$ at the origin O on the surface of a half-space, whose temperature is initially uniform and equal to θ_0. The temperature at subsequent times at a point situated a radial

https://doi.org/10.1017/CBO9781139171731.013 Published online by Cambridge University Press

distance R from O is given by (C & J §10.2)†

$$\theta - \theta_0 = \frac{H}{4\rho c(\pi\kappa t)^{3/2}} \exp(-R^2/4\kappa t) \tag{12.1}$$

At any point in the solid the temperature rises rapidly from θ_0 to a maximum value when $t = R^2/6\kappa$ and slowly decays to θ_0 as the heat diffuses through the solid.

(b) Instantaneous line source

The treatment of two-dimensional problems is facilitated by the use of a line source in which H units of heat per unit length are instantaneously liberated on the surface of a half-space along the y-axis. The temperature distribution is cylindrical about the y-axis and at a distance R is given by (C & J §10.3)

$$\theta - \theta_0 = \left(\frac{H}{2\pi kt}\right) \exp(-R^2/4\kappa t) \tag{12.2}$$

(c) Continuous point source

If heat is supplied to the half-space at O at a steady rate $\dot{H}$, the temperature at a distance R from O may be found by integrating (12.1) with respect to time. It varies according to

$$\theta - \theta_0 = (\dot{H}/2\pi kR) \operatorname{erfc}(R/4\kappa t) \tag{12.3}$$

where $\operatorname{erfc}(\chi) = 1 - \operatorname{erf}(\chi) = 1 - (2/\pi^{1/2})\int_0^\chi \exp(-\xi^2)\, d\xi$.

After a sufficient time has elapsed, a steady state is reached in the neighbourhood of the source (i.e. where $R \ll 4\kappa t$) in which the temperature is given by

$$\theta - \theta_0 = \dot{H}/2\pi kR \tag{12.4}$$

The infinite temperature at $R = 0$ is a consequence of assuming that the heat is introduced at a point.

(d) Distributed heat sources (C & J §10.5)

In reality heat is introduced into the surface of a solid over a finite area. Assuming the remainder of the surface to be perfectly insulated, the temperature within the solid may be found by the superposition of point or line sources. If heat is supplied at a steady rate $\dot{h}$ per unit area, then equations (12.3) and (12.4) can be used.

Suppose we wish to find the steady-state distribution of temperature throughout the surface of a half-space when heat is supplied steadily to a small area A of the surface. For a single source equation (12.4) applies, in which we

† Carslaw & Jaeger (1959).

shall denote by $\bar{\theta}$ temperature at a point on the surface distance r from the source. The reader of this book will recognise that equation (12.4) is analogous to equation (3.22*b*) which expresses the normal displacement $\bar{u}_z$ of a point on the surface of an elastic half-space due to a point force P acting at a distance r, viz.

$$\bar{u}_z = \frac{1-\nu^2}{\pi E}\,\frac{P}{r}$$

This analogy may be used to determine the steady-state distribution of surface temperature due to a distributed heat supply. For example, the temperature due to uniform supply of heat to a circular area of radius a is analogous to the displacement produced by a uniform pressure. If $(1-\nu^2)/E$ is replaced by $\frac{1}{2}k$ and p by $\dot{h}$, the surface temperature distribution is given by equations (3.29). The maximum temperature at the centre of the circle is

$$\bar{\theta}_{\mathrm{max}} - \theta_0 = \dot{h}a/k \tag{12.5a}$$

and the average temperature over the heated circle is

$$\bar{\theta}_{\mathrm{max}} - \theta_0 = 8\dot{h}a/3\pi k \tag{12.5b}$$

Similarly the temperature at the surface of a uniformly heated polygonal region may be found from the results of §3.3.

The same analogy is useful when it is required to find the distribution of heat supplied to a small area of the surface which would maintain a steady prescribed temperature distribution in that area. For example, consider a half-space in which a circular area of the surface, of radius a, is maintained at a steady uniform temperature $\bar{\theta}_{\mathrm{c}}$. The temperature far away is θ_0 and the surface of the half-space outside the circle is insulated. The analogous elastic problem concerns the indentation of an elastic half-space by a rigid circular punch which imposes a uniform displacement of the surface $\bar{u}_z$. The pressure under the punch is given by equations (3.34) and (3.36), from which the required distribution of heat supply may be deduced to be

$$\dot{h} = \frac{2k(\bar{\theta}_{\mathrm{c}} - \theta_0)}{\pi(a^2 - r^2)^{1/2}} \tag{12.6}$$

(e) Moving heat sources (C & J §10.7)

In order to investigate the temperature produced by frictional heating in sliding contact we need to examine the temperature produced in a half-space by a heat source which moves on the surface. If we are dealing with the steady state it is convenient to fix the heat source and imagine the half-space moving beneath it with a steady velocity V parallel to the x-axis. The temperature field is then a function of position but not of time.

We shall examine the two-dimensional case of an infinitely long source of heat parallel to the y-axis, uniformly distributed over the strip $-a \leqslant x \leqslant a$. Referring to Fig. 12.1, the distributed source is regarded as an array of line sources of strength $\dot{h}$ ds. The element of material at (x, z) at time t was located at $(x - Vt', z)$ at an earlier instant $(t - t')$. The heat liberated by the line source at s in the time interval dt' is $\dot{h}$ ds dt', whereupon the steady temperature of an element located instantaneously at x is found by the integration of equation (12.2) from $t = -\infty$ to the current instant $t = 0$.

$$\theta(x,z) - \theta_0 = \frac{\dot{h}}{2\pi k} \int_{-a}^{a} \mathrm{d}s \int_{-\infty}^{0} \left[\exp\left\{-\frac{(x-s-Vt')^2 + z^2}{4\kappa t'}\right\} \frac{\mathrm{d}t'}{t'}\right] \tag{12.7}$$

The maximum temperature occurs on the surface ($z = 0$) for which equation (12.7) can be written in the form

$$\bar{\theta} - \theta_0 = \frac{\dot{h}a}{kL^{1/2}} F(L, X) \tag{12.8}$$

where $L = (Va/2\kappa)$ and $X = (Vx/2\kappa)$. The integrals have been evaluated by Jaeger (1942). Surface temperature distributions are shown in Fig. 12.2(a). The maximum temperature occurs towards the rear of the heated zone which has had the longest exposure to heat. Maximum and average temperatures for the heated zone are plotted against the speed parameter L in Fig. 12.2(b). The parameter L, known as the Peclet number, may be interpreted as the ratio of the speed of the surface to the rate of diffusion of heat into the solid. At large values of L (>5), the heat will diffuse only a short distance into the solid in the time taken for the surface to move through the heated zone. The heat flow will then be approximately perpendicular to the surface at all points. The

Fig. 12.1

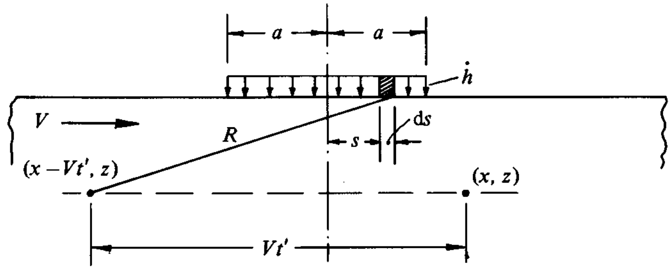
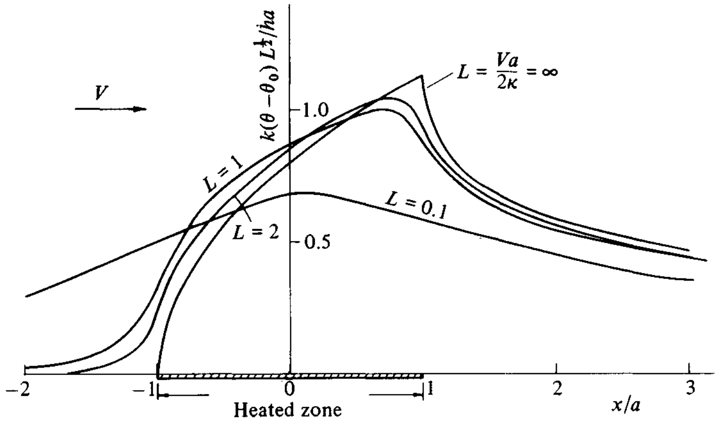
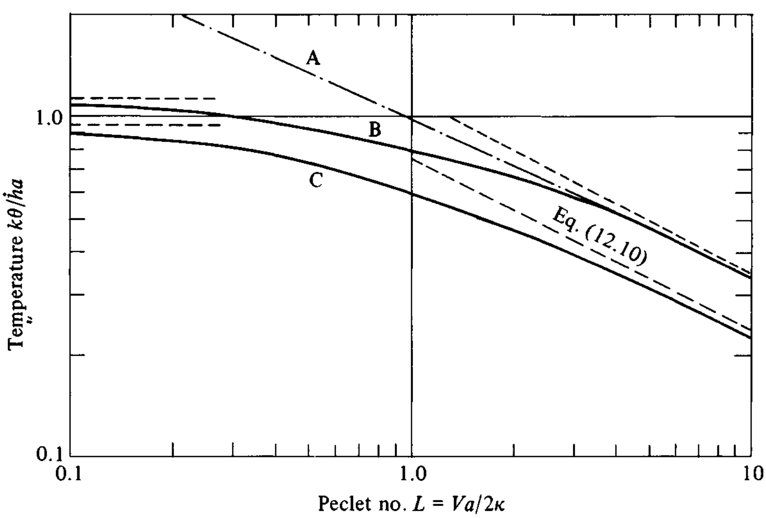


Fig. 12.2. Surface temperature rise due to a uniform moving line heat source. (*a*) Temperature distribution; (*b*) Maximum and average temperatures as a function of speed. A – Band source (max); B – square source (max); C – square source (mean).



(*a*)



(*b*)

https://doi.org/10.1017/CBO9781139171731.013 Published online by Cambridge University Press

temperature of a surface point is then given by (C & J §2.9)

$$\bar{\theta} - \theta_0 = \frac{2\dot{h}(\kappa t)^{1/2}}{\pi^{1/2}k} = \frac{\dot{h}a}{k}\left\{\frac{2}{\pi}\left(\frac{2\kappa}{Va}\right)(1+x)\right\}^{1/2}, \quad -a \leqslant x \leqslant a \tag{12.9}$$

The mean temperature for a band source is then

$$\bar{\theta}_{\text{mean}} - \theta_0 \approx \frac{4}{3\pi^{1/2}}\frac{\dot{h}a}{k}L^{-1/2} \tag{12.10}$$

Since equation (12.9) is based on one-dimensional heat flow into the solid, it applies to a uniform source of any planform. Thus the mean temperature for a square source of side $2a$ is also given by expression (12.10).

At very low speeds ($L < 0.1$) the temperature distribution becomes symmetrical and similar to that for a stationary source. In the case of an infinitely long band source no steady-state temperature is reached, but the square source reaches a maximum temperature of $\theta_0 + 1.12\dot{h}a/k$ at the centre and a mean temperature in the heated zone of $\theta_0 + 0.946\dot{h}a/k$.

12.3 Steady thermoelastic distortion of a half-space

The equations which express the stresses and displacements in an elastic half-space due to an arbitrary steady distribution of temperature have been derived by various authors (e.g. Boley & Weiner, 1960). If the surface $z = 0$ is stress-free, then it can be shown that all parallel planes are stress-free, i.e. $\sigma_z = \tau_{yz} = \tau_{xz} = 0$ throughout. The normal displacement at depth z is given by (see Williams, 1961)

$$u_z = -(1+\nu)\alpha \int \theta \, \mathrm{d}z \tag{12.11}$$

Williams has expressed the thermoelastic stresses and displacements in an elastic half-space in terms of two harmonic functions and Barber (1975) has used this formulation to derive some general results which are useful in thermoelastic contact problems:

(i) If heat is supplied at a rate $\dot{h}$ per unit area to the free surface of a half-space, the surface distorts according to:

$$\frac{\partial^2 \bar{u}_z}{\partial x^2} + \frac{\partial^2 \bar{u}_z}{\partial y^2} = c\dot{h}(x,y) \tag{12.12a}$$

With circular symmetry this equation becomes

$$\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\left(r\frac{\mathrm{d}\bar{u}_z}{\mathrm{d}r}\right) = c\dot{h}(r) \tag{12.12b}$$

where $c = (1+\nu)\alpha/k$ is referred to as the 'distortivity' of the material.

https://doi.org/10.1017/CBO9781139171731.013 Published online by Cambridge University Press

In a two-dimensional situation $\partial^2 \bar{u}_z/\partial y^2 = 0$, so that equation (12.12*a*) implies that the curvature of the surface at any point is directly proportional to the rate of heat flow at that point; convex when heat is flowing into the surface and concave when it is flowing out. An insulated surface which is initially flat will remain flat. Thus uniform heating of a half-space over a long narrow strip will distort the surface such that it has a constant convex curvature within the strip and has plane inclined surfaces outside the strip. This general theorem is due to Dundurs (1974) (see also Barber, 1980*a*).

(ii) If the surface of a half-space, $z = 0$, is heated so that it has a distribution of surface temperature $\bar{\theta}(x, y)$, then the surface stress required to maintain the surface flat, such that $\bar{u}_z = 0$, is given by

$$\bar{\sigma}_z = -\tfrac{1}{2}\{cE/(1-\nu^2)\}\bar{\theta}(x,y) \tag{12.13}$$

The application of an equal and opposite traction will free the surface from traction and allow it to distort. It follows from equation (12.13) that the surface displacements are those which would be caused by a surface pressure $p(x, y)$ proportional to the distribution of surface temperature $\bar{\theta}(x, y)$. In this way the steady thermal distortion of a half-space can be found by the methods discussed earlier in this book if the temperature distribution over the whole surface is known. In most contact situations, however, it is usual to assume that no heat is transferred across a non-contacting surface; the boundary conditions, therefore, are more appropriately expressed in terms of heat flux rather than temperature.

We will now look at a number of particular cases.

(a) Point source of heat

The temperature distribution due to a continuous point source of heat at point O on the surface is given by equation (12.4). The normal displacement of a point on the surface distance r from O then follows from equation (12.11)

$$u_z = -(1+\nu)\alpha \int \dot{H}/\{2\pi k(r + z^2)^{1/2}\}\,\mathrm{d}z$$

$$\bar{u}_z = -(c\dot{H}/2\pi)\ln(r_0/r) \tag{12.14}$$

where r_0 is the position on the surface where $\bar{u}_z = 0$. Since heat is being injected continuously into the solid and the surface is assumed to be insulated except at O, the expansion of the surface given by equation (12.14) increases without limit as the datum for displacements is taken at an increasing distance from the source.

(b) Uniform heating of a circular region

When heat is supplied steadily over a small area of the surface, the surface distortion can be found by superposition of the point source solutions (Barber, 1971*a*). For example, heating a thin annulus of radius a gives rise to surface displacements:

$$\bar{u}_z = -(c\dot{H}/2\pi)\ln(r_0/a), \quad r \leqslant a \tag{12.15a}$$

which is constant and

$$\bar{u}_z = -(c\dot{H}/2\pi)\ln(r_0/r), \quad r > a \tag{12.15b}$$

which is the same as for a point source at the centre. From these results we may easily proceed to uniform heating of a circular area, radius a, whence

$$\bar{u}_z = \begin{cases} -(c\dot{H}/4\pi)\{2\ln(r_0/a) + (1 - r^2/a^2)\}, & r \leqslant a \quad (12.16a) \\ -(c\dot{H}/2\pi)\ln(r_0/r), & r > a \quad (12.16b) \end{cases}$$

This distorted surface is shown in Fig. 12.3. Alternatively the results expressed by equations (12.16) could have been obtained directly by integrating equation (12.12*b*) with $\dot{h} = \dot{H}/\pi a^2$ for $r \leqslant a$ and $\dot{h} = 0$ for $r > a$.

(c) Circular region at uniform temperature

To maintain a circular region of the surface at a uniform temperature θ_c which is different from the temperature at a distance θ_0 requires a supply of heat per unit area $\dot{h}(r)$ distributed according to equation (12.6). Hence substituting $2\pi r\dot{h}(r)\,dr$ for $\dot{H}$ in equation (12.13) and integrating from $r = 0$ to $r = a$ gives (Barber, 1971*b*)

$$\bar{u}_z = \begin{cases} -\dfrac{2}{\pi}cka(\theta_c - \theta_0)\left[\ln\left(\dfrac{r_0}{a}\right) - \ln\{1 + (1 - r^2/a^2)^{1/2}\}\right. \\ \left. \quad + \left(1 - \dfrac{r^2}{a^2}\right)^{1/2}\right], \quad r \leqslant a & (12.17a) \\ -\dfrac{2}{\pi}cka(\theta_c - \theta_0)\ln(r_0/r), \quad r > a & (12.17b) \end{cases}$$

This distortion is also illustrated in Fig. 12.3.

(d) Moving heat source

We wish to find the thermal distortion of the surface due to the moving heat source shown in Fig. 12.1. Barber has shown that a concentrated line source $\dot{H}$, moving with velocity V causes a displacement of the surface at a point distance ξ ahead of the source given by

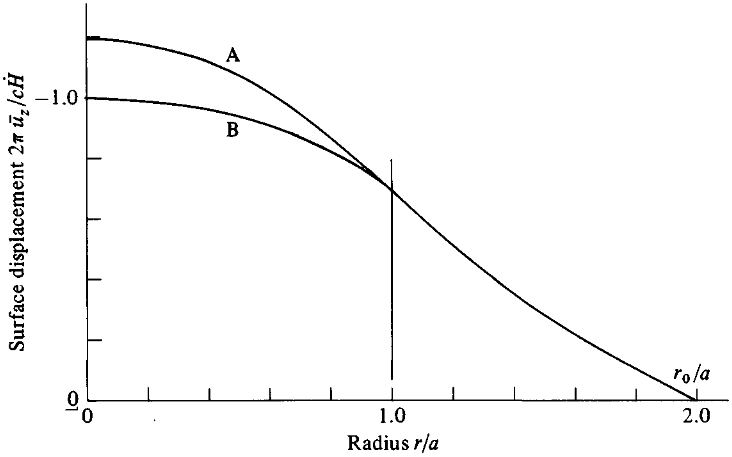
$$\bar{u}_z = -(2c\kappa\dot{H}/V)\exp(-X^2)I_0(X^2) \tag{12.18a}$$

where $X = (V\xi/2\kappa)^{1/2}$ and I_0 is a modified Bessel function.† At all points

† See Abramowitz & Stegun, *Handbook of Mathematical Functions*, Dover (1965) for definitions and tabulation of modified Bessel functions I_0 and I_1, and also for integration of equation (12.18*a*).

https://doi.org/10.1017/CBO9781139171731.013 Published online by Cambridge University Press

Fig. 12.3. Thermoelastic distortion of a half-space heated over a circular area radius a. Curve A – uniform-heat input, eqs. (12.16); curve B – uniform temperature, eqs. (12.17).

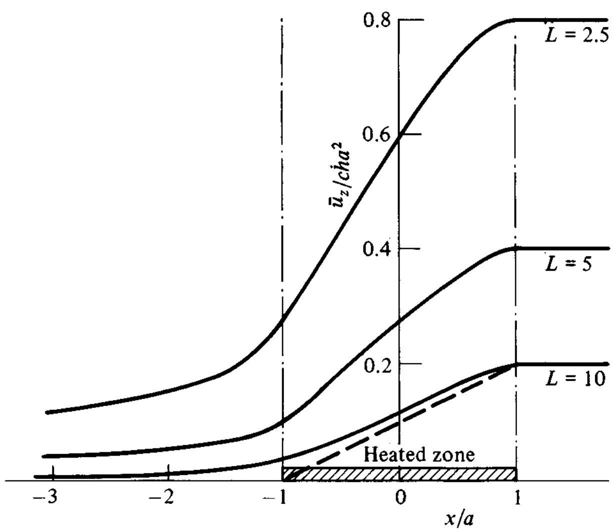


behind the source the displacement is constant and given by

$$\bar{u}_z = -2c\kappa\dot{H}/V \tag{12.18b}$$

The thermal distortion due to a uniform rate of heating $\dot{h}$ over the strip $-a \leqslant x \leqslant a$ can be found by superposition. The displacement $\bar{u}_z(x)$ of a surface point is found by putting $\xi = s - x$ in equation (12.18*a*) when ξ is positive,

Fig. 12.4. Thermoelastic distortion of a half-space caused by a moving heat source.



i.e. when the point in question is ahead of the element $\dot{h}$ ds of the heat source, and by using equation (12.18b) when ξ is negative. In this way the distorted surface shapes shown in Fig. 12.4 have been found as a function of the Peclet number L defined above. At the leading edge of the heat source ($x = -a$) the thermal displacement varies with speed according to

$$\bar{u}_z(-a) = -2c\dot{h}a^2\, \mathrm{e}^{-2L}\{I_0(2L) + I_1(2L)\}/L \tag{12.19a}$$

$$\approx -2c\dot{h}a^2(\pi L^3)^{-1/2} \tag{12.19b}$$

for large values of L. At the trailing edge the displacement is

$$\bar{u}_z(a) = -2c\dot{h}a^2/L \tag{12.19c}$$

With increasing speed conduction ahead of the source becomes less effective so that the displacement at the leading edge falls. At high speed the displacement increases almost in proportion to the distance from the leading edge.

12.4 Contact between bodies at different temperatures

We shall consider first the situation in which a hot body, at temperature θ_1, is pressed into contact with a cooler one at temperature θ_2, and will restrict the discussion to frictionless surfaces and to a circular contact area of radius a.

If we assume, in the first instance, that the bodies make perfect thermal contact at their interface and that each body is effectively a half-space, then the heat conduction problem is straightforward. The temperature of the interface θ_c is uniform and the distribution of heat flux across the interface is given by equation (12.6), i.e.

$$\dot{h}(r) = \frac{2k_1(\theta_1 - \theta_c)}{\pi(a^2 - r^2)^{1/2}} = \frac{2k_2(\theta_c - \theta_2)}{\pi(a^2 - r^2)^{1/2}}$$

from which it may be seen that θ_c divides the difference in temperature between the two bodies in the inverse proportion to their thermal conductivities k_1 and k_2. The total heat flux is thus

$$\dot{H} = 4k_1 a(\theta_1 - \theta_c) = 4k_2 a(\theta_c - \theta_2) = 4ka(\theta_1 - \theta_2) \tag{12.20}$$

where $k = k_1k_2/(k_1 + k_2)$.

The distortion of the surface of an elastic half-space by this distribution of heat flux is given by equations (12.17). Since heat is flowing into the cooler body it develops a bulge in which the displacements are proportional to the value of its distortivity c_2 $(= \alpha_2(1 + \nu_2)/k_2)$, while the warmer body develops a hollow of similar shape whose depth is proportional to c_1. If $c_2 = c_1$ the bulge in the cooler body will just fit the hollow in the warmer body and the contact stresses due to the external load will not be influenced by the existence of the heat flux.

This is clearly the case with identical materials. When the two materials are different the heat flux will give rise to an additional ('thermal') contact pressure such as to suppress the mismatch in the distorted profiles of the two surfaces. The required pressure, acting over the circle $r \leqslant a$, is that which (by equation (12.17)) would produce combined displacements of both surfaces given by

$$(\bar{u}_z)_1 + (\bar{u}_z)_2 = \frac{1}{2\pi}(c_2 - c_1)\dot{H} \times [\ln(r_0/a) - \ln\{1 + (1 - r^2/a^2)^{1/2}\} + (1 - r^2/a^2)^{1/2}] \tag{12.21}$$

It can be found by the methods of Chapter 3, with the result (Barber, 1973)

$$p'(r) = (c_2 - c_1)\frac{\dot{H}E^*}{2\pi^2 a}\left[\frac{\pi^2}{8} - \chi_2\left\{\frac{a - (a^2 - r^2)^{1/2}}{a + (a^2 - r^2)^{1/2}}\right\}\right] \tag{12.22}$$

where $\chi_2(x)$ is Legendre's chi function†, which is defined by

$$\chi_2(x) = \tfrac{1}{2}\int_0^x \ln\left(\frac{1+s}{1-s}\right)\frac{ds}{s} = \sum_{m=1}^{m=\infty} \frac{x^{2m-1}}{(2m-1)^2}$$

This pressure distribution is plotted in Fig. 12.5. It corresponds to a total load

$$P' = (1/2\pi)(c_2 - c_1)\dot{H}E^*a = (2/\pi)k(c_2 - c_1)(\theta_1 - \theta_2)E^*a^2 \tag{12.23}$$

The differential thermal expansion of the two bodies (12.21) has been annulled by the application of the 'thermal' pressure distribution given by equation (12.22). To find the net pressure we must add a Hertz contact pressure to account for the isothermal elastic deformation, viz.:

$$p''(r) = (2E^*a/\pi R)\{1 - (r/a)^2\}^{1/2} \tag{12.24}$$

and the corresponding load

$$P'' = 4E^*a^3/3R \tag{12.25}$$

The total load P, thermal plus isothermal, is thus $P' + P''$, which leads to the relationship:

$$\beta(a/a_0)^2 + (a/a_0)^3 = 1 \tag{12.26}$$

where $a_0 = (3RP/4E^*)^{1/3}$ is the contact radius under isothermal (Hertz) conditions and $\beta = (3kR/2\pi a_0)(c_2 - c_1)(\theta_1 - \theta_2)$. This relationship for $\beta > 0$ is plotted on the right-hand side of Fig. 12.6. As expected an increase in temperature difference or in differential distortivity causes the relative curvature of the two surfaces to increase and the contact area to decrease. In the case of two

† L. Lewin, *Dilogarithms and Associated Functions*, MacDonald, London, 1958.

Fig. 12.5. Contact of spherical bodies at different temperatures: body of lower distortivity at higher temperature: A – pressure distribution given by eq. (12.22); B – isothermal pressure distribution (Hertz).

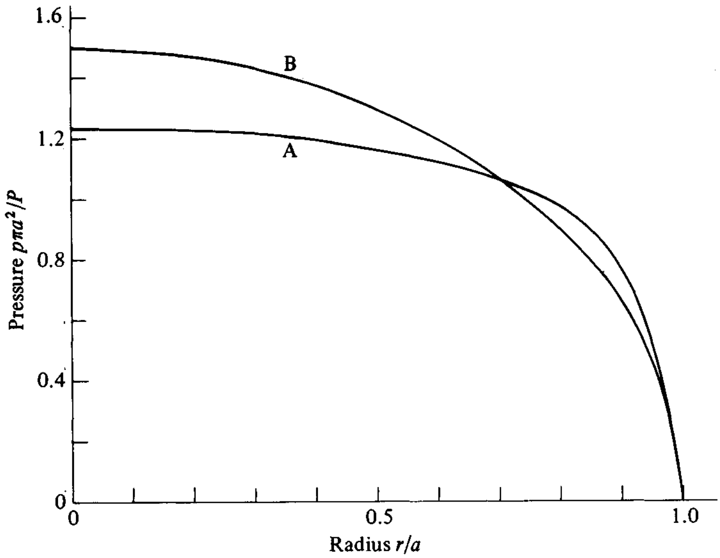
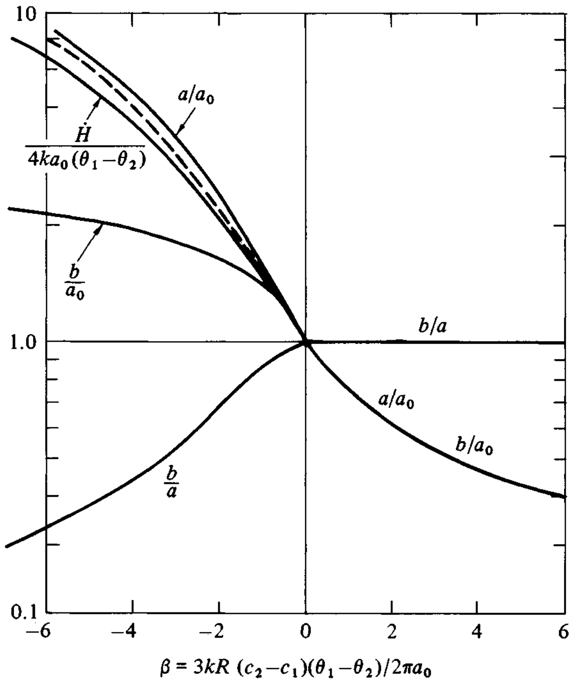


Fig. 12.6. Contact of spherical bodies at different temperatures. Exact solutions with annulus of imperfect contact $b < r < a$. Broken line – approximate solution (eq. (12.26)) assuming perfect contact throughout.



https://doi.org/10.1017/CBO9781139171731.013 Published online by Cambridge University Press

nominally flat surfaces ($R \to \infty$) the isothermal component of pressure is zero. Any slight departure from flatness will cause the heat transfer to be concentrated at the point of closest approach of the two surfaces. Consequent thermal expansion will cause a circular area of intimate contact to develop and the surrounding surfaces to separate. The size of the contact zone is then given by equation (12.23).

The above analysis of the contact of two dissimilar bodies whose temperatures are different is perfectly satisfactory provided that the product $(c_2 - c_1) \times (\theta_1 - \theta_2)$ is positive. If the situation is changed, such that the body of higher distortivity has the higher temperature, a new feature is introduced through the change in sign of the thermal pressure $p'(r)$. It may be seen from Fig. 12.5 that this pressure distribution falls to zero very steeply as $r \to a$ compared with the Hertz pressure. It can be shown that $p'(r)$ always exceeds $p''(r)$ in this region so that, when a negative (tensile) thermal pressure is added to the isothermal pressure, an annulus of *tensile* traction is found, however small the temperature difference between the surfaces. This suggests that the surfaces would peel apart at the edge, but it is not possible for them to do so and to still maintain equilibrium with the applied load. It must be concluded that there is no solution to the problem in the form posed above. Barber (1978) has shown that the paradox arises from the thermal boundary conditions which assume 'perfect contact' within the contact area, i.e. no discontinuity in temperature across the interface, and perfect insulation outside the contact. The difficulty can be removed by the introduction of an additional state of 'imperfect contact' in which the displacements (elastic and thermal) are such that the surfaces just touch and conduct some heat, but the contact pressure is zero. This state is achieved by a jump in temperature across the interface. These boundary conditions follow from the fact that, in reality, the change from perfect insulation to perfect conduction will not be discontinuous when surfaces come into contact. When the separation is sufficiently small heat can be transferred by radiation or conduction through the intervening gas; further, contaminant films and the inevitable roughness of real surfaces give rise to a thermal resistance at the interface which might be expected to depend inversely upon the contact pressure. The argument may be appreciated in its simplest form by reference to the one-dimensional model shown in Fig. 12.7(a). A rod of length l, Young's modulus E and coefficient of thermal expansion α is placed between two rigid conducting walls A and B, at temperatures θ_A and θ_B. The rod is attached to A and initially, when $\theta_A = \theta_B$, there is a small gap $g = g_0$ between the end of the rod and wall B. If θ_A is raised above θ_B, in the steady state the rod will acquire the temperature θ_A and the gap will close such that

$$g = g_0 - \alpha l(\theta_A - \theta_B) \tag{12.27}$$

This expression is only valid if $g > 0$, i.e. if

$$\theta_A - \theta_B \leqslant g_0/\alpha l \tag{12.28}$$

If the rod expands to make 'perfect contact' with the wall B, so that its temperature varies linearly from θ_A to θ_B, its unrestrained expansion would be $\frac{1}{2}\alpha l(\theta_A - \theta_B)$, but the actual expansion cannot exceed g_0, so that a pressure will develop on the end of the rod given by

$$p = \tfrac{1}{2}E\alpha(\theta_A - \theta_B) - Eg_0/l \tag{12.29}$$

The contact pressure must be positive so that

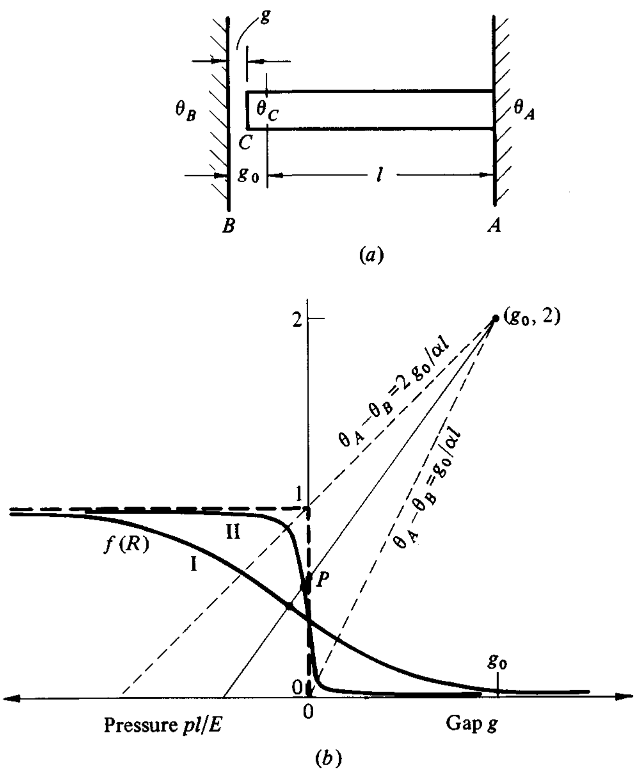
$$\theta_A - \theta_B \geqslant 2g_0/\alpha l \tag{12.30}$$

There is thus a range of temperature difference

$$g_0/\alpha l < (\theta_A - \theta_B) < 2g_0/\alpha l$$

in which no steady-state solution is possible. This is a similar state of affairs to

Fig. 12.7. Contact of an elastic rod between two rigid walls at different temperatures ($\theta_A > \theta_B$). (*a*) The system. (*b*) The thermal resistance $f(R)$ as a function of the gap g or the contact pressure p.



that encountered in the contact of spheres discussed above. To resolve the paradox we introduce a thermal resistance $R(g)$ which varies continuously with the gap, becoming very large as g becomes large. As we have seen negative gaps cannot exist and we should replace the unrestrained 'gap' $-g$ by a contact pressure $p = -Eg/l$. The resistance $R(p)$ will decrease as p increases. The temperature of the free end of the rod is denoted by θ_C so that, by equating the heat flux along the rod to that across the gap, we get

$$Sk(\theta_A - \theta_C)/l = (\theta_C - \theta_B)/R$$

where S is the cross-sectional area of the rod. Thus

$$\theta_A - \theta_C = (\theta_A - \theta_B)f(R) \tag{12.31}$$

where $f(R) = (1 + SkR/l)^{-1}$. This function is represented by curve I in Fig. 12.7(b). For large positive gaps R is large, hence $f(R)$ approaches zero; at high contact pressure (negative unrestrained gap), R is small and $f(R)$ approaches unity, but its precise form is unimportant.

The expression for the gap now becomes

$$\begin{aligned} g &= g_0 - \tfrac{1}{2}\alpha l(\theta_A + \theta_C - 2\theta_B) \\ &= g_0 - \alpha l(\theta_A - \theta_B) + \tfrac{1}{2}\alpha l(\theta_A - \theta_C) \end{aligned} \tag{12.32}$$

Eliminating $(\theta_A - \theta_C)$ from equations (12.31) and (12.32) we find

$$f(R) = 2 + 2(g - g_0)/\alpha l(\theta_A - \theta_B) \tag{12.33}$$

This equation plots in Fig. 12.7(b) as a straight line which passes through the point $(g_0, 2)$, and whose gradient is inversely proportional to the temperature difference $(\theta_A - \theta_B)$. Where the line intersects the curve of $f(R)$ gives the steady solution to the problem: it determines the gap g if the point of intersection is to the right of 0 and the pressure p if the point of intersection is to the left of 0. Note that a point of intersection exists for a line of any gradient, hence a solution can be found for all values of $(\theta_A - \theta_B)$.

If we now make the resistance curve $f(R)$ increasingly sensitive to the gap and the contact pressure, as shown by curve II in Fig. 12.7, in the limit it takes the form of a 'step', zero to the right of 0 and unity to the left. More significantly it has a vertical segment between 0 and 1 when $g = 0$. An intersection with the straight line given by equation (12.33) is still possible in the range $g_0/\alpha l < (\theta_A - \theta_B) < 2g_0/\alpha l$, as indicated by the point P in Fig. 12.7(b). Both the gap and the contact pressure are zero at this point; the temperature of the end of the rod θ_C is intermediate between θ_A and θ_B, given by putting $g = 0$ in equation (12.32), and some heat flows across the interface. These are the boundary conditions referred to by Barber (1978) as 'imperfect contact' and investigated further by Comninou & Dundurs (1979).

Returning to the contact of spheres when the heat flow is such that β is negative, the existence of tensile stresses as $r \to a$ when perfect contact is assumed suggests that the contact area will be divided into a central region ($r \leqslant b$) of perfect contact surrounded by an annulus ($b < r \leqslant a$) of imperfect contact. Barber (1978) has analysed this situation with results which are shown in Fig. 12.6 for negative values of β. The variation in contact radius (a/a_0) given by equation (12.26), which assumes perfect contact throughout, is also shown for comparison. With increasing (negative) temperature difference the contact size grows as the thermal distortion makes the surfaces more conforming. The exact variation is not very different from that predicted by equation (12.26). The radius b of the circle of perfect contact, within which the contact pressure is confined, also grows but more slowly. It is shown in Fig. 12.6. as a ratio of the isothermal contact radius a_0 and also as a ratio of the actual radius a. The mean contact pressure falls, therefore, but not to the extent which would be expected if perfect contact prevailed throughout. With perfect contact the heat flux through the contact $\dot{H}$ is proportional to the contact radius a, so that the influence of thermal distortion on $\dot{H}$ is expressed by the approximate curve of a/a_0 against β given by equation (12.26) and shown dotted in Fig. 12.6. The exact variation of heat flux is also shown. The effect of an annulus of imperfect contact upon the heat flux is not large; the reduction in conductivity of the interface is offset to some extent by the increase in the size of the contact. The analogous problems of two-dimensional contact of cylindrical bodies and of nominally flat wavy surfaces have been solved by Comninou *et al.* (1981) and Panek & Dundurs (1979).

When contact is made between a flat rigid punch and an elastic half-space which is hotter than the punch, at first sight a hollow would be expected to form in the half-space so that contact would be lost from the centre of the punch. This cannot happen, however, since by equation (12.12*b*) the surface can only become concave if heat is flowing from it, whereas no heat flows if there is no contact. This is another situation, investigated by Barber (1982), in which a state of imperfect contact exists, this time in a central region of the punch.

A basic feature of Fig. 12.6 calls for comment: for a given temperature difference between the bodies, the heat transfer from the body of lower distortivity into that of higher distortivity ($\beta < 0$) is greater than the heat transfer in the opposite direction ($\beta > 0$). This phenomenon has been called 'thermal rectification' and is frequently observed when heat is transferred between dissimilar solids in contact. The above theory, with modifications to allow for the geometry of the experimental arrangement, has shown reasonable agreement with measurements of heat transfer between rods having rounded ends in contact (see Barber, 1971*b*).

12.5 Frictional heating and thermoelastic instability

In the sliding contact of nominally flat surfaces heat is liberated by friction at the interface at a rate

$$\dot{h} = \mu V p \tag{12.34}$$

where V is the sliding velocity and μ the coefficient of friction. If the pressure p is uniform then the heat conducted to the surfaces will be uniform and so will the surface temperature. It has been frequently observed with brake blocks, for example, that the stationary surface develops 'hot spots' where the temperature is much in excess of its expected mean value. This phenomenon was investigated by Barber (1969). He showed that initial small departures from perfect conformity concentrated the pressure and hence the frictional heating into particular regions of the interface. These regions expanded above the level of the surrounding surface and reduced the area of real contact, as described in the previous section, thereby concentrating the contact and elevating the local temperature still further. This process has come to be called 'thermoelastic instability' and has been studied in detail by Burton (1980). If sliding continues the expanded spots, where the pressure is concentrated, wear down until contact occurs elsewhere. The new contact spots proceed to heat, expand and carry the load; the old ones, relieved of load, cool, contract and separate. This cyclic process has been frequently observed in the sliding contact of conforming surfaces. The scale of the hot spots is large compared with the scale of surface roughness and the time of the cycle is long compared with the time of asperity interactions. The essential mechanism of thermoelastic instability may be appreciated by the simple example considered below.

Two semi-infinite sliding solids having nominally flat surfaces, which are pressed into contact with a mean pressure $\bar{p}$, are shown in Fig. 12.8. To avoid the transient nature of heat flow into a moving surface, the moving surface will be taken to be perfectly flat, and non-conducting. The stationary solid has a distortivity c and its surface has a *small* initial undulation of amplitude Δ and wavelength λ. In the present example, where the mating surface is non-conducting, it is immaterial whether the undulations are parallel or perpendicular to the direction of sliding. The isothermal pressure required to flatten this waviness is found in Chapter 13 (eq. (13.7)), to be

$$p'' = (\pi E^* \Delta/\lambda) \cos(2\pi x/\lambda) \tag{12.35}$$

The steady thermal distortion of the surface is given by

$$\frac{d^2 \bar{u}_z}{dx^2} = c\dot{h} = c\mu V p(x) \tag{12.36}$$

It is clear that the initial sinusoidal undulation of wavelength λ is going to result in a fluctuation of pressure at the same wavelength, which may be expressed by

$$p(x) = \bar{p} + p^* \cos(2\pi x/\lambda) \tag{12.37}$$

We are concerned here only with the fluctuating components of pressure and heat flux which, when substituted in equation (12.36) and integrated, give the thermal distortion of the surface to be

$$\bar{u}_z = -(c\mu V p^* \lambda^2/4\pi^2) \cos(2\pi x/\lambda) \tag{12.38}$$

The thermal pressure $p'(x)$ required to press this wave flat can now be added to the isothermal pressure given by (12.35) to obtain the relationship

$$p^* = \frac{\pi E^*}{\lambda}(\Delta + c\mu V p^* \lambda^2/4\pi^2)$$

whereupon

$$\frac{p^*}{\bar{p}} = \frac{\pi E^* \Delta/\lambda\bar{p}}{1 - c\mu V E^* \lambda/4\pi} \tag{12.39}$$

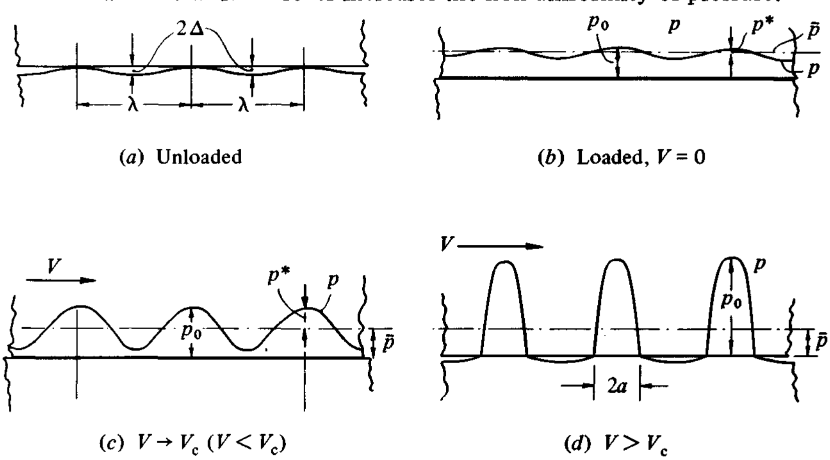
As the sliding velocity approaches a critical value V_c given by

$$V_c = 4\pi/c\mu E^* \lambda \tag{12.40}$$

the fluctuations in pressure given by equation (12.39) increase rapidly in magnitude (Fig. 12.8(c)).

When the fluctuation in pressure p^* reaches the mean pressure $\bar{p}$ the surfaces will separate in the hollows of the original undulations and the contact will concentrate at the crests (Fig. 12.8(d)). A simple treatment of this situation

Fig. 12.8. Mechanism of thermoelastic instability. Thermal expansion causes small initial pressure fluctuations to grow when the sliding speed approaches a critical value V_c. At high speed contact becomes discontinuous which further increases the non-uniformity of pressure.



(a) Unloaded

(b) Loaded, $V = 0$

(c) $V \to V_c$ ($V < V_c$)

(d) $V > V_c$

https://doi.org/10.1017/CBO9781139171731.013 Published online by Cambridge University Press

may be carried out by assuming that the pressure in the contact patch, $-a \leqslant x \leqslant +a$, is Hertzian, i.e. $p(x) = p_0\{1-(x/a)^2\}^{1/2}$, where

$$p_0 = aE^*/2R \tag{12.41}$$

The curvature of $1/R$ of the distorted surface at $x = 0$ is given by

$$1/R = c\mu V p_0 + 4\pi^2\Delta/\lambda^2 \tag{12.42a}$$

$$\approx c\mu V p_0 \tag{12.42b}$$

if the initial undulation is small compared with the subsequent thermal distortion. Thus equation (12.42) gives†

$$a \approx 2/c\mu V E^* \tag{12.43}$$

The transition from continuous to discontinuous contact at the interface takes place when V approaches V_c given by (12.40). Putting $V = V_c$ in (12.43) then gives an approximate expression for the critical contact size:

$$a_c \approx \lambda/2\pi \tag{12.44}$$

The non-uniform pressure distribution leads directly to non-uniform heat input and to a non-uniform distribution of surface temperature. The temperature distribution can be found using the analogy with the surface displacements produced by a pressure which is proportional to the heat flux at the surface. Below the critical speed, while the surfaces are in continuous contact, the pressure fluctuations are sinusoidal with an amplitude p^* given by equation (12.39). It follows that the fluctuations in heat flux and temperature will also be sinusoidal with amplitudes h^* and θ^*. From the analogy mentioned above we find

$$\theta^* = \lambda h^*/2\pi k = \mu V \lambda p^*/2\pi k \tag{12.45}$$

Above the critical speed, the surfaces are in discontinuous contact. The surface displacements and contact pressures where a wavy surface is in discontinuous contact with a plane are given in §13.2. From those results it may be deduced that the temperature difference between the centre of a contact patch and the centre of the trough is given by

$$\bar{\theta}(0) - \bar{\theta}(\lambda/2) \approx \frac{\mu VP}{\pi k}\,\frac{\sin\psi}{\psi}\left\{\frac{1-\cos\psi}{\sin^2\psi} + \ln\left(\frac{1+\cos\psi}{\sin\psi}\right)\right\} \tag{12.46}$$

where $\psi = \pi a/\lambda$.

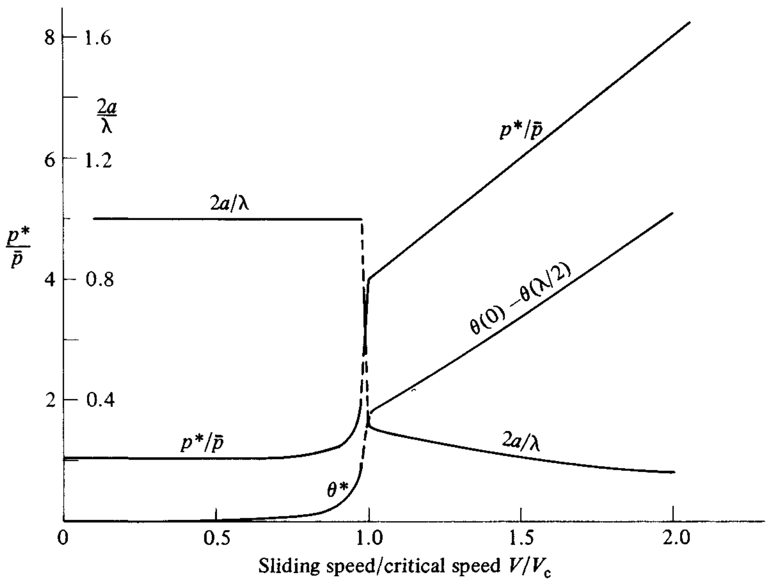
The mechanism of thermoelastic instability may now be described with reference to the above example (Figs. 12.8 and 12.9). In static contact any waviness of the surfaces in contact will give rise to a non-uniform distribution

† A more exact treatment which matches the pressure and distortion throughout the contact patch has been carried out by Burton & Nerlikar (1975) for multiple contacts; for a single contact patch Barber (1976) finds $a = 2.32/c\mu VE^*$.

of contact pressure. At low sliding speeds the variations in pressure from the steady mean value are augmented by thermoelastic distortion according to equation (12.38). When the velocity reaches a critical value V_c given by equation (12.40) the amplitude of the fluctuation increases very rapidly and, if they have not already done so, the surfaces separate at the positions of the initial hollows. Contact is discontinuous and the size of the contact patches shrinks to a width about 1/3 of the original wavelength (eq. (12.44)). Further increase in speed results in a stable decrease in the contact patch size according to equation (12.43). The sudden rise in pressure and drop in contact area at the critical speed are accompanied by a sharp rise in temperature fluctuation, as shown in Fig. 12.9.

Real surfaces, of course, will have a spectrum of initial undulations. Equation (12.39) suggests that the pressure variation grows in proportion to the ratio Δ/λ, i.e. to the slope of the undulations. The critical velocity V_c, however, is independent of the amplitude and inversely proportional to the wavelength. This suggests that long wavelength undulations will become unstable before the short ones and thereby dominate the process. The size of the body imposes an upper limit to the wavelength and hence a lower limit to the critical speed. The undulations in real surfaces are two-dimensional having curvature in both directions. Following the onset of instability, the same reasoning that led to

Fig. 12.9. Variation of the contact width a and the amplitude of pressure and temperature fluctuations with sliding speed.



https://doi.org/10.1017/CBO9781139171731.013 Published online by Cambridge University Press

equation (12.43) gives the radius of a discrete circular contact area to be†

$$a \approx \pi/c\mu VE^* \tag{12.47}$$

Another way of expressing the influence of the size of the body is to say that, if the nominal contact area has a diameter less than $2a$ given by (12.47), the situation will be stable.

The above analysis simplifies the real situation in two important ways: (i) the thermoelastic solutions employed refer to the steady state, whereas the unstable variation in contact pressure and area is essentially a transient process, and (ii) both surfaces will be conducting and deformable to a greater or lesser extent. To investigate these effects Dow & Burton (1972) and Burton *et al.* (1973) have studied the stability of small sinusoidal perturbations in pressure between two extended sliding surfaces in continuous contact. The equation of unsteady heat flow was used. They show first that a pair of identical materials is very stable; however high the sliding speed an impractical value of the coefficient of friction (>2) would be required to cause instability. When the two materials are different a thermal disturbance, comprising a fluctuation in pressure and temperature, moves along the interface at a velocity which is different from that of either surface. An appreciable difference in the thermal conductivities of the two materials, however, leads to the disturbance being effectively locked to the body of higher conductivity; most of the heat then passes into that surface. In the limit we have the situation analysed above where one surface is non-conducting. The critical velocity then approaches that given by equation (12.40). Some heat is, in fact, conducted to the mating surface, at a rate given by equation (12.10) which reduces the heat causing thermoelastic deformation of the more conducting surface and thereby increases the critical velocity above that given by (12.40).

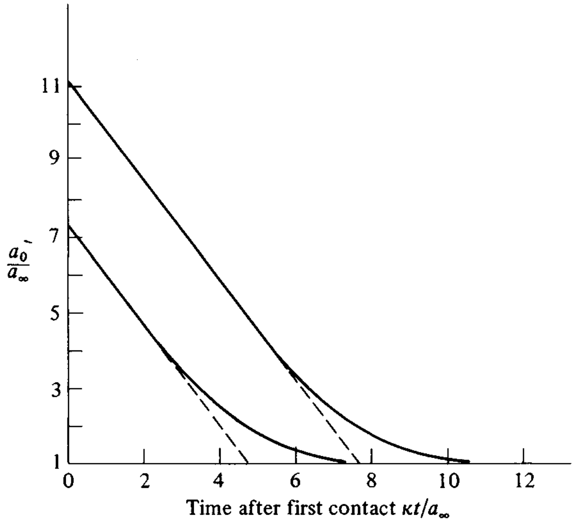
When the contact is discontinuous the analysis of transient thermoelasticity becomes more difficult. Some basic cases of the distortion of a half-space due to transient heating of a small area of the surface have been investigated by Barber (1972) and Barber & Martin-Moran (1982). These results have been used to investigate the transient shrinking of a circular contact area due to frictional heating when the moving surface is an insulator. The stationary conducting surface is assumed to have a slight crown so that before sliding begins there is an initial contact area of radius a_0. During sliding, in the steady state, the contact area shrinks to a radius a_∞. In this analysis the simplifying assumptions which we have used previously are applied: the pressure distribution is Hertzian and the curvature due to thermoelastic distortion is matched at the

† More exactly, for a single contact patch, Barber (1976) obtains $a = 1.28\pi/c\mu VE^*$

https://doi.org/10.1017/CBO9781139171731.013 Published online by Cambridge University Press

origin only. With these assumptions a_∞ is given by equation (12.47). Barber (1980*b*) shows that the contact radius shrinks initially at a uniform rate $1.34\kappa/a_\infty$. Only in the later stages is a_∞ approached asymptotically, as shown in Fig. 12.10.

Fig. 12.10. Transient thermoelastic variations of the radius of a circular area in sliding contact from its initial value a_0 to its steady-state value a_∞.



https://doi.org/10.1017/CBO9781139171731.013 Published online by Cambridge University Press