4

Normal contact of elastic solids: Hertz theory†

4.1 Geometry of smooth, non-conforming surfaces in contact

When two non-conforming solids are brought into contact they touch initially at a single point or along a line. Under the action of the slightest load they deform in the vicinity of their point of first contact so that they touch over an area which is finite though small compared with the dimensions of the two bodies. A theory of contact is required to predict the shape of this area of contact and how it grows in size with increasing load; the magnitude and distribution of surface tractions, normal and possibly tangential, transmitted across the interface. Finally it should enable the components of deformation and stress in both bodies to be calculated in the vicinity of the contact region.

Before the problem in elasticity can be formulated, a description of the geometry of the contacting surfaces is necessary. In Chapter 1 we agreed to take the point of first contact as the origin of a rectangular coordinate system in which the x-y plane is the common tangent plane to the two surfaces and the z-axis lies along the common normal directed positively into the lower solid (see Fig. 1.1). Each surface is considered to be topographically smooth on both micro and macro scale. On the micro scale this implies the absence or disregard of small surface irregularities which would lead to discontinuous contact or highly local variations in contact pressure. On the macro scale the profiles of the surfaces are continuous up to their second derivative in the contact region. Thus we may express the profile of each surface in the region close to the origin approximately by an expression of the form

$$z_1 = A_1 x^2 + B_1 y^2 + C_1 xy + \ldots \tag{4.1}$$

where higher order terms in x and y are neglected. By choosing the orientation of the x and y axes, x_1 and y_1, so that the term in xy vanishes, (4.1) may be written:

† A summary of Hertz elastic contact stress formulae is given in Appendix 3, p. 427.

https://doi.org/10.1017/CBO9781139171731.005 Published online by Cambridge University Press

$$z_1 = \frac{1}{2R_1'} x_1^2 + \frac{1}{2R_1''} y_1^2 \tag{4.2a}$$

where R_1' and R_1'' are the principal radii of curvature of the surface at the origin. They are the maximum and minimum values of the radius of curvature of all possible cross-sections of the profile. If a cross-sectional plane of symmetry exists one principal radius lies in that plane. A similar expression may be written for the second surface:

$$z_2 = -\left(\frac{1}{2R_2'} x_2^2 + \frac{1}{2R_2''} y_2^2\right) \tag{4.2b}$$

The separation between the two surfaces is then given by $h = z_1 - z_2$. We now transpose equation (4.1) and its counterpart to a common set of axes x and y, whereby

$$h = Ax^2 + By^2 + Cxy$$

By a suitable choice of axes we can make C zero, hence

$$h = Ax^2 + By^2 = \frac{1}{2R'} x^2 + \frac{1}{2R''} y^2 \tag{4.3}$$

where A and B are positive constants and R' and R'' are defined as the principal *relative* radii of curvature. If the axes of principal curvature of each surface, i.e. the x_1 axis and the x_2 axis, are inclined to each other by an angle α, then it is shown in Appendix 2 that

$$(A + B) = \tfrac{1}{2}\left(\frac{1}{R'} + \frac{1}{R''}\right) = \tfrac{1}{2}\left(\frac{1}{R_1'} + \frac{1}{R_1''} + \frac{1}{R_2'} + \frac{1}{R_2''}\right) \tag{4.4}$$

and

$$|B - A| = \tfrac{1}{2}\left\{\left(\frac{1}{R_1'} - \frac{1}{R_1''}\right)^2 + \left(\frac{1}{R_2'} - \frac{1}{R_2''}\right)^2 + 2\left(\frac{1}{R_1'} - \frac{1}{R_1''}\right)\left(\frac{1}{R_2'} - \frac{1}{R_2''}\right)\cos 2\alpha\right\}^{1/2} \tag{4.5}$$

We introduce an equivalent radius R_e defined by

$$R_e = (R'R'')^{1/2} = \tfrac{1}{2}(AB)^{-1/2}$$

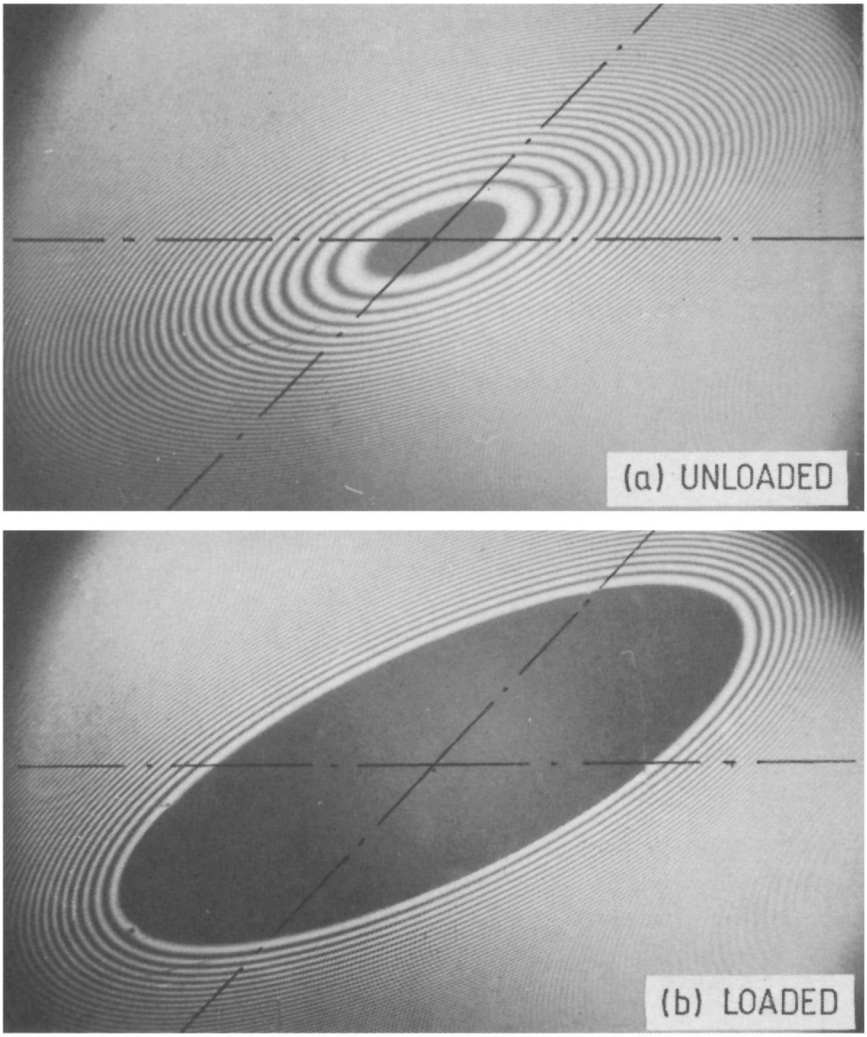
In this description of the initial separation between the two surfaces in terms of their principal radii of curvature we have taken a *convex* surface to have a *positive* radius. Equations (4.4) and (4.5) apply equally to concave or saddle-shaped surfaces by ascribing a negative sign to the concave curvatures.

It is evident from equation (4.3) that contours of constant gap h between the undeformed surfaces are ellipses the length of whose axes are in the ratio

https://doi.org/10.1017/CBO9781139171731.005 Published online by Cambridge University Press

$(B/A)^{1/2} = (R'/R'')^{1/2}$. Such elliptical contours are displayed by the interference fringes between two cylindrical lenses, each of radius R, with their axes inclined at 45°, shown in Fig. 4.1(*a*). In this example $R_1' = R_2' = R$; $R_1'' = R_2'' = \infty$; $\alpha = 45°$, for which equations (4.4) and (4.5) give $A + B = 1/R$ and $B - A = 1/\sqrt{(2)}R$, i.e. $A = (1 - 1/\sqrt{2})/2R$ and $B = (1 + 1/\sqrt{2})/2R$. The relative radii of curvature are thus: $R' = 1/2A = 3.42R$ and $R'' = 1/2B = 0.585R$. The equivalent radius $R_e = (R'R'')^{1/2} = \sqrt{(2)}R$ and $(R'/R'')^{1/2} =$

Fig. 4.1. Interference fringes at the contact of two equal cylindrical lenses with their axes inclined at 45°: (*a*) unloaded, (*b*) loaded.



https://doi.org/10.1017/CBO9781139171731.005 Published online by Cambridge University Press

$(B/A)^{1/2} = 2.41$. This is the ratio of the major to minor axes of the contours of constant separation shown in Fig. 4.1(*a*).

We can now say more precisely what we mean by non-conforming surfaces: the relative curvatures $1/R'$ and $1/R''$ must be sufficiently large for the terms Ax^2 and By^2 on the right-hand side of equation (4.3) to be large compared with the higher order terms which have been neglected. The question of conforming surfaces is considered in §5.3.

A normal compressive load is now applied to the two solids and the point of contact spreads into an area. If the two bodies are solids of revolution, then $R_1' = R_1'' = R_1$ and $R_2' = R_2'' = R_2$, whereupon $A = B = \frac{1}{2}(1/R_1 + 1/R_2)$. Thus contours of constant separation between the surfaces before loading are circles centred at O. After loading, it is evident from the circular symmetry that the contact area will also be circular. Two cylindrical bodies of radii R_1 and R_2 in contact with their axes parallel to the y-axis have $R_1' = R_1, R_1'' = \infty$, $R_2' = R_2, R_2'' = \infty$ and $\alpha = 0$, so that $A = \frac{1}{2}(1/R_1 + 1/R_2), B = 0$. Contours of constant separation are straight lines parallel to the y-axis and, when loaded, the surfaces will make contact over a narrow strip parallel to the y-axis. In the case of general profiles it follows from equation (4.3) that contours of constant separation are ellipses in plan. We might expect, therefore, that under load the contact surface would be elliptical in shape. It will be shown in due course that this is in fact so. A special case arises when two equal cylinders both of radius R are in contact with their axes perpendicular. Here $R_1' = R, R_1'' = \infty, R_2' = R$, $R_2'' = \infty, \alpha = \pi/2$, from which $A = B = \frac{1}{2}R$. In this case the contours of constant separation are circles and identical to those due to a sphere of the same radius R in contact with a plane surface ($R_2' = R_2'' = \infty$).

We shall now consider the deformation as a normal load P is applied. Two solids of general shape (but chosen convex for convenience) are shown in cross-section after deformation in Fig. 4.2. Before deformation the separation between two corresponding surface points $S_1(x, y, z_1)$ and $S_2(x, y, z_2)$ is given by equation (4.3). From the symmetry of this expression about O the contact region must extend an equal distance on either side of O. During the compression distant points in the two bodies T_1 and T_2 move towards O, parallel to the z-axis, by displacements δ_1 and δ_2 respectively. If the solids did not deform their profiles would overlap as shown by the dotted lines in Fig. 4.2. Due to the contact pressure the surface of each body is displaced parallel to Oz by an amount $\bar{u}_{z1}$ and $\bar{u}_{z2}$ (measured positive into each body) relative to the distant points T_1 and T_2. If, after deformation, the points S_1 and S_2 are coincident within the contact surface then

$$\bar{u}_{z1} + \bar{u}_{z2} + h = \delta_1 + \delta_2 \tag{4.6}$$

Writing $\delta = \delta_1 + \delta_2$ and making use of (4.3) we obtain an expression for the

https://doi.org/10.1017/CBO9781139171731.005 Published online by Cambridge University Press

elastic displacements:

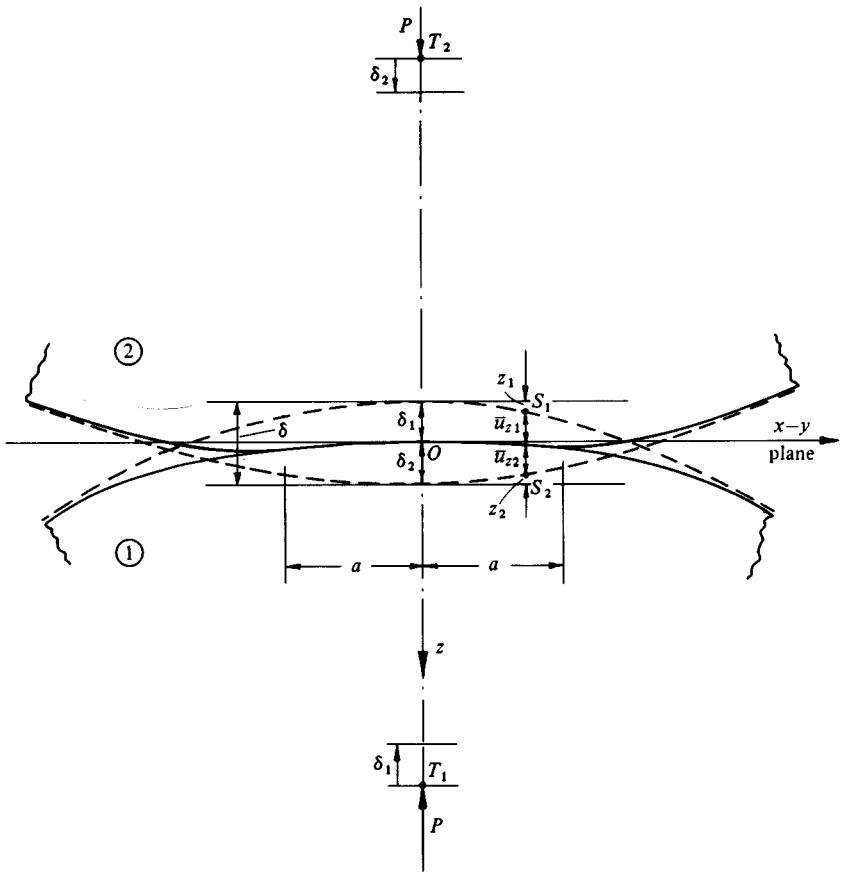
$$\bar{u}_{z1} + \bar{u}_{z2} = \delta - Ax^2 - By^2 \tag{4.7}$$

where x and y are the common coordinates of S_1 and S_2 projected onto the x-y plane. If S_1 and S_2 lie outside the contact area so that they do not touch it follows that

$$\bar{u}_{z1} + \bar{u}_{z2} > \delta - Ax^2 - By^2 \tag{4.8}$$

To solve the problem, it is necessary to find the distribution of pressure transmitted between the two bodies at their surface of contact, such that the resulting elastic displacements normal to that surface satisfy equation (4.7) within the contact area and equation (4.8) outside it.

Fig. 4.2.



https://doi.org/10.1017/CBO9781139171731.005 Published online by Cambridge University Press

Before proceeding to examine the problem in elasticity, however, it is instructive to see how the deformations and stresses grow as the load is applied on the basis of elementary dimensional reasoning. For simplicity we shall restrict the discussion to (*a*) solids of revolution in which the contact area is a circle of radius a and (*b*) two-dimensional bodies in which the contact area is an infinite strip of width $2a$.

We note that in Fig. 4.2, $\delta_1 = \bar{u}_{z1}(0)$ and $\delta_2 = \bar{u}_{z2}(0)$, so that equation (4.6) can be written in non-dimensional form

$$\left\{\frac{\bar{u}_{z1}(0)}{a} - \frac{\bar{u}_{z1}(x)}{a}\right\} + \left\{\frac{\bar{u}_{z2}(0)}{a} - \frac{\bar{u}_{z2}(x)}{a}\right\} = \tfrac{1}{2}(1/R_1 + 1/R_2)x^2/a \qquad (4.9)$$

Putting $x = a$ and writing $\bar{u}_z(0) - u_z(a) = d$, the 'deformation' within the contact zone, (4.9) becomes

$$\frac{d_1}{a} + \frac{d_2}{a} = \frac{a}{2}\left(\frac{1}{R_1} + \frac{1}{R_2}\right) \qquad (4.10)$$

Provided that the deformation is small, i.e. $d \ll a$, the state of *strain* in each solid is characterised by the ratio d/a. Now the magnitude of the strain will be proportional to the contact pressure divided by the elastic modulus; therefore, if p_m is the average contact pressure acting mutually on each solid, (4.10) becomes†

$$p_m/E_1 + p_m/E_2 \propto a(1/R_1 + 1/R_2)$$

i.e.

$$p_m \propto \frac{a(1/R_1 + 1/R_2)}{1/E_1 + 1/E_2} \qquad (4.11)$$

Thus, for a given geometry and materials, the contact pressure and the associated stresses increase in direct proportion to the linear dimension of the contact area. To relate the growth of the contact to the load, two and three-dimensional contacts must be examined separately.

(*a*) In the contact of cylinders, the load per unit axial length $P = 2ap_m$, whence from (4.11)

$$a \propto \{P(1/E_1 + 1/E_2)/(1/R_1 + 1/R_2)\}^{1/2} \qquad (4.12)$$

and

$$p_m \propto \{P(1/R_1 + 1/R_2)/(1/E_1 + 1/E_2)\}^{1/2} \qquad (4.13)$$

† It transpires that the 'plane-strain modulus' $E/(1-\nu^2)$ is the correct elastic modulus to use in contact problems, but Young's modulus E is used here to retain the simplicity of the argument.

https://doi.org/10.1017/CBO9781139171731.005 Published online by Cambridge University Press

from which we see that the contact width and contact pressure increase as the square root of the applied load.

(*b*) In the contact of spheres, or other solids of revolution, the compressive load $P = \pi a^2 p_m$. Hence from (4.11)

$$a \propto \{P(1/E_1 + 1/E_2)/(1/R_1 + 1/R_2)\}^{1/3} \tag{4.14}$$

and

$$p_m \propto \{P(1/R_1 + 1/R_2)^2/(1/E_1 + 1/E_2)^2\}^{1/3} \tag{4.15}$$

In this case the radius of the contact circle and the contact pressure increase as the cube root of the load.

In the case of three-dimensional contact the compressions of each solid δ_1 and δ_2 are proportional to the local indentations d_1 and d_2, hence the approach of distant points

$$\begin{aligned} \delta = \delta_1 + \delta_2 &\propto d_1 + d_2 \\ &\propto \{P^2(1/E_1 + 1/E_2)^2(1/R_1 + 1/R_2)\}^{1/3} \end{aligned} \tag{4.16}$$

The approach of two bodies due to elastic compression in the contact region is thus proportional to $(\text{load})^{2/3}$.

In the case of two-dimensional contact the displacements δ_1 and δ_2 are not proportional to d_1 and d_2 but depend upon the arbitrarily chosen datum for elastic displacements. An expression similar to (4.16) cannot be found in this case.

We have shown how the contact area, stress and deformation might be expected to grow with increasing load and have also found the influence of curvature and elastic moduli by simple dimensional reasoning. To obtain absolute values for these quantities we must turn to the theory of elasticity.

4.2 Hertz theory of elastic contact

The first satisfactory analysis of the stresses at the contact of two elastic solids is due to Hertz (1882*a*). He was studying Newton's optical interference fringes in the gap between two glass lenses and was concerned at the possible influence of elastic deformation of the surfaces of the lenses due to the contact pressure between them. His theory, worked out during the Christmas vacation 1880 at the age of twenty-three, aroused considerable interest when it was first published and has stood the test of time. In addition to static loading he also investigated the quasi-static impacts of spheres (see §11.4). Hertz (1882*b*) also attempted to use his theory to give a precise definition of hardness of a solid in terms of the contact pressure to *initiate* plastic yield in the solid by pressing a harder body into contact with it. This definition has proved unsatisfactory because of the difficulty of detecting the point of first yield under the action of contact stress. A satisfactory theory of hardness had to wait for the development of the theory of plasticity. This question is considered in Chapter 6.

Hertz formulated the conditions expressed by equations (4.7) and (4.8) which must be satisfied by the normal displacements on the surface of the solids. He first made the hypothesis that the contact area is, in general, elliptical, guided no doubt by his observations of interference fringes such as those shown in Fig. 4.1(*b*). He then introduced the simplification that, for the purpose of calculating the local deformations, each body can be regarded as an elastic half-space loaded over a small elliptical region of its plane surface. By this simplification, generally followed in contact stress theory, the highly concentrated contact stresses are treated separately from the general distribution of stress in the two bodies which arises from their shape and the way in which they are supported. In addition, the well developed methods for solving boundary-value problems for the elastic half-space are available for the solution of contact problems. In order for this simplification to be justifiable two conditions must be satisfied: the significant dimensions of the contact area must be small compared (*a*) with the dimensions of each body and (*b*) with the relative radii of curvature of the surfaces. The first condition is obviously necessary to ensure that the stress field calculated on the basis of a solid which is infinite in extent is not seriously influenced by the proximity of its boundaries to the highly stressed region. The second condition is necessary to ensure firstly that the surfaces just outside the contact region approximate roughly to the plane surface of a half-space, and secondly that the strains in the contact region are sufficiently small to lie within the scope of the linear theory of elasticity. Metallic solids loaded within their elastic limit inevitably comply with this latter restriction. However, caution must be used in applying the results of the theory to low modulus materials like rubber where it is easy to produce deformations which exceed the restriction to small strains.

Finally, the surfaces are assumed to be frictionless so that only a normal pressure is transmitted between them. Although physically the contact pressure must act perpendicular to the interface which will not necessarily be planar, the linear theory of elasticity does not account for changes in the boundary forces arising from the deformation they produce (with certain special exceptions). Hence, in view of the idealisation of each body as a half-space with a plane surface, normal tractions at the interface are taken to act parallel to the z-axis and tangential tractions to act in the x-y plane.

Denoting the significant dimension of the contact area by a, the relative radius of curvature by R, the significant radii of each body by R_1 and R_2 and the significant dimensions of the bodies both laterally and in depth by l, we may summarise the assumptions made in the Hertz theory as follows:

(i) The surfaces are continuous and non-conforming: $a \ll R$;

(ii) The strains are small: $a \ll R$;

https://doi.org/10.1017/CBO9781139171731.005 Published online by Cambridge University Press

(iii) Each solid can be considered as an elastic half-space: $a \ll R_{1,2}, a \ll l$;
(iv) The surfaces are frictionless: $q_x = q_y = 0$.

The problem in elasticity can now be stated: the distribution of mutual pressure $p(x, y)$ acting over an area S on the surface of two elastic half-spaces is required which will produce normal displacements of the surfaces $\bar{u}_{z1}$ and $\bar{u}_{z2}$ satisfying equation (4.7) within S and (4.8) outside it.

(a) Solids of revolution

We will consider first the simpler case of solids of revolution ($R_1' = R_1'' = R_1; R_2' = R_2'' = R_2$). The contact area will be circular, having a radius a, say. From equations (4.4) and (4.5) it is clear that $A = B = \frac{1}{2}(1/R_1 + 1/R_2)$, so that the boundary condition for displacements within the contact expressed in (4.7) can be written

$$\bar{u}_{z1} + \bar{u}_{z2} = \delta - (1/2R)r^2 \tag{4.17}$$

where $(1/R) = (1/R_1 + 1/R_2)$ is the relative curvature.

A distribution of pressure which gives rise to displacements which satisfy (4.17) has been found in §3.4, where the pressure distribution proposed by Hertz (equation (3.39))

$$p = p_0\{1 - (r/a)^2\}^{1/2}$$

was shown to give normal displacements (equation (3.41*a*))

$$\bar{u}_z = \frac{1 - \nu^2}{E} \frac{\pi p_0}{4a} (2a^2 - r^2), \quad r \leqslant a$$

The pressure acting on the second body is equal to that on the first, so that by writing

$$\frac{1}{E^*} = \frac{1 - \nu_1^2}{E_1} + \frac{1 - \nu_2^2}{E_2}$$

and substituting the expressions for $\bar{u}_{z1}$ and $\bar{u}_{z2}$ into equation (4.17) we get

$$\frac{\pi p_0}{4aE^*} (2a^2 - r^2) = \delta - (1/2R)r^2 \tag{4.18}$$

from which the radius of the contact circle is given by

$$a = \pi p_0 R/2E^* \tag{4.19}$$

and the mutual approach of distant points in the two solids is given by

$$\delta = \pi a p_0/2E^* \tag{4.20}$$

The total load compressing the solids is related to the pressure by

$$P = \int_0^a p(r) 2\pi r \, dr = \tfrac{2}{3} p_0 \pi a^2 \tag{4.21}$$

https://doi.org/10.1017/CBO9781139171731.005 Published online by Cambridge University Press

Hence the maximum pressure p_0 is 3/2 times the mean pressure p_m. In a practical problem, it is usually the total load which is specified, so that it is convenient to use (4.21) in combination with (4.19) and (4.20) to write

$$a = \left(\frac{3PR}{4E^*}\right)^{1/3} \tag{4.22}$$

$$\delta = \frac{a^2}{R} = \left(\frac{9P^2}{16RE^{*2}}\right)^{1/3} \tag{4.23}$$

$$p_0 = \frac{3P}{2\pi a^2} = \left(\frac{6PE^{*2}}{\pi^3 R^2}\right)^{1/3} \tag{4.24}$$

These expressions have the same form as (4.14), (4.15) and (4.16) which were obtained by dimensional reasoning. However they also provide absolute values for the contact size, compression and maximum pressure.

Before this solution to the problem can be accepted, we must ask whether (4.17) is satisfied uniquely by the assumed pressure distribution and also check whether condition (4.8) is satisfied to ensure that the two surfaces do not touch or interfere *outside* the loaded circle. By substituting equation (3.42*a*) for the normal displacement ($r > a$) into equation (4.8) and making use of (4.19), it may be verified that the Hertz distribution of pressure does not lead to contact outside the circle $r = a$.

On the question of uniqueness, we note from §3.4 that a pressure distribution of the form (equation (3.34))

$$p = p_0' \{1 - (r/a)^2\}^{-1/2}$$

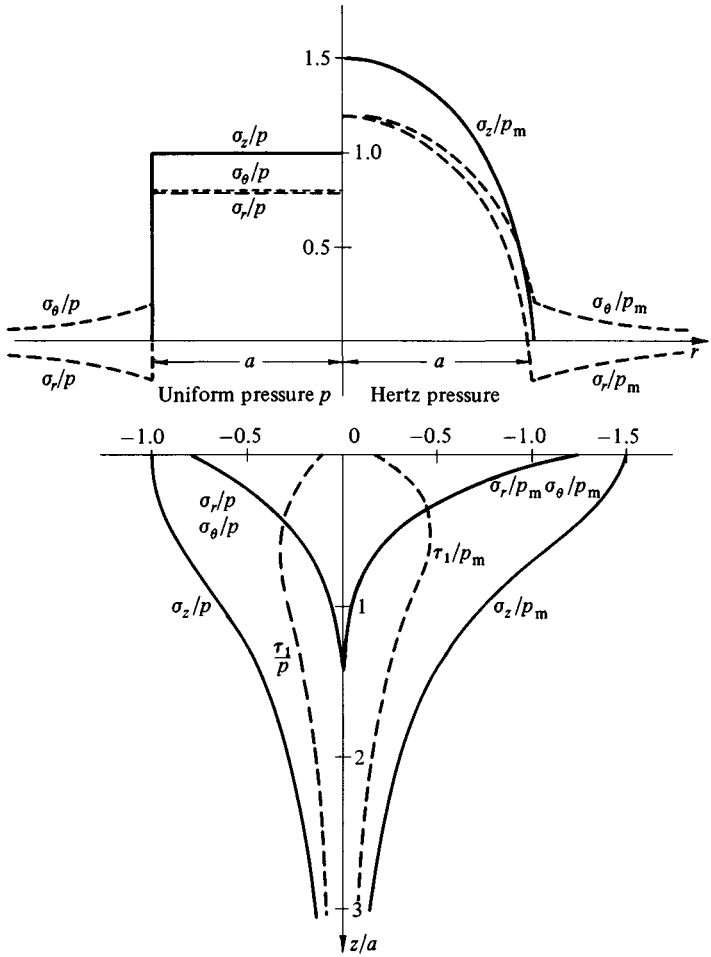
produces a uniform normal displacement within the loaded circle. Thus such a pressure could be added to or subtracted from the Hertz pressure while still satisfying the condition for normal displacements given by (4.17). However, this pressure distribution also gives rise to an infinite gradient of the surface immediately outside the loaded circle in the manner of a rigid cylindrical punch. Clearly two elastic bodies having smooth continuous profiles could not develop a pressure distribution of this form without interference outside the circle $r = a$. On the other hand, if such a pressure distribution were subtracted from the Hertz pressure, the normal traction just inside the loaded circle would be *tensile* and of infinite magnitude. In the absence of adhesion between the two surfaces, they cannot sustain tension, so that both positive and negative tractions of the form given above are excluded. No other distribution of normal traction produces displacements which satisfy (4.17) so that we conclude that the Hertz pressure distribution is the unique solution to the problem.

The stresses within the two solids due to this pressure distribution have been found in §3.4, and are shown in Fig. 4.3. At the surface, within the

https://doi.org/10.1017/CBO9781139171731.005 Published online by Cambridge University Press

contact area, the stress components are given by equation (3.43); they are all compressive except at the very edge of contact where the radial stress is tensile having a maximum value $(1-2\nu)p_0/3$. This is the greatest tensile stress anywhere and it is held responsible for the ring cracks which are observed to form when brittle materials like glass are pressed into contact. At the centre the radial stress is compressive and of value $(1+2\nu)p_0/2$. Thus for an incompressible material ($\nu = 0.5$) the stress at the origin is hydrostatic. Outside the contact area the radial and circumferential stresses are of equal magnitude and are tensile and compressive respectively (equation (3.44)).

Fig. 4.3. Stress distributions at the surface and along the axis of symmetry caused by (left) uniform pressure and (right) Hertz pressure acting on a circular area radius a.



https://doi.org/10.1017/CBO9781139171731.005 Published online by Cambridge University Press

Expressions for the stresses beneath the surface along the z-axis are given in equations (3.45). They are principal stresses and the principal shear stress ($\tau_1 = \frac{1}{2}$ (principal stress difference)) has a value of approximately $0.31p_0$ at a depth of $0.48a$ (for $\nu = 0.3$). This is the maximum shear stress in the field, exceeding the shear stress at the origin $= \frac{1}{2}|\sigma_z - \sigma_r| = 0.10p_0$, and also the shear stress in the surface at the edge of the contact $= \frac{1}{2}|\sigma_r - \sigma_\theta| = 0.13p_0$. Hence plastic yielding would be expected to initiate beneath the surface. This question is considered in detail in the next chapter.

(b) General profiles

In the general case, where the separation is given by equation (4.3), the shape of the contact area is not known with certainty in advance. However we assume tentatively that S is elliptical in shape, having semi-axes a and b. Hertz recognised that the problem in elasticity is analogous to one of electrostatic potential. He noted that a charge, whose intensity over an elliptical region on the surface of a conductor varies as the ordinate of a semi-ellipsoid, gives rise to a variation in potential throughout that surface which is parabolic. By analogy, the pressure distribution given by equation (3.58)

$$p = p_0\{1 - (x/a)^2 - (y/b)^2\}^{1/2}$$

produces displacements within the ellipse given by equation (3.61):

$$\bar{u}_z = \frac{1-\nu^2}{\pi E}(L - Mx^2 - Ny^2)$$

Thus for both bodies,

$$\bar{u}_{z1} + \bar{u}_{z2} = (L - Mx^2 - Ny^2)/\pi E^* \tag{4.25}$$

which satisfies the condition (4.7): $\bar{u}_{z1} + \bar{u}_{z2} = \delta - Ax^2 - By^2$ provided that (from equations (3.62))

$$A = M/\pi E^* = (p_0/E^*)(b/e^2a^2)\{\mathbf{K}(e) - \mathbf{E}(e)\} \tag{4.26a}$$

$$B = N/\pi E^* = (p_0/E^*)(b/a^2e^2)\{(a^2/b^2)\mathbf{E}(e) - \mathbf{K}(e)\} \tag{4.26b}$$

$$\delta = L/\pi E^* = (p_0/E^*)b\mathbf{K}(e) \tag{4.26c}$$

where $\mathbf{E}(e)$ and $\mathbf{K}(e)$ are complete elliptic integrals of argument $e = (1 - b^2/a^2)^{1/2}$, $b < a$. The pressure distribution is semi-ellipsoidal and, from the known volume of an ellipsoid, we conclude that the total load P is given by

$$P = (2/3)p_0\pi ab \tag{4.27}$$

from which the average pressure $p_m = (2/3)p_0$.

To find the shape and size of the ellipse of contact, we write

$$\frac{B}{A} = \left(\frac{R'}{R''}\right) = \frac{(a/b)^2\mathbf{E}(e) - \mathbf{K}(e)}{\mathbf{K}(e) - \mathbf{E}(e)} \tag{4.28}$$

https://doi.org/10.1017/CBO9781139171731.005 Published online by Cambridge University Press

and

$$(AB)^{1/2} = \tfrac{1}{2}(1/R'R'')^{1/2} = 1/2R_e$$

$$= \frac{p_0}{E^*}\frac{b}{a^2e^2}\left[\{(a/b)^2\mathbf{E}(e) - \mathbf{K}(e)\}\{\mathbf{K}(e) - \mathbf{E}(e)\}\right]^{1/2} \tag{4.29}$$

We now write $c = (ab)^{1/2}$ and substitute for p_0 from (4.27) into (4.29) to obtain

$$c^3 \equiv (ab)^{3/2} = \left(\frac{3PR_e}{4E^*}\right)\frac{4}{\pi e^2}(b/a)^{3/2}$$

$$\times \left[\{(a/b)^2\mathbf{E}(e) - \mathbf{K}(e)\}\{\mathbf{K}(e) - \mathbf{E}(e)\}\right]^{1/2}$$

i.e.

$$c = (ab)^{1/2} = \left(\frac{3PR_e}{4E^*}\right)^{1/3} F_1(e) \tag{4.30}$$

The compression is found from equations (4.26*c*) and (4.27):

$$\delta = \frac{3P}{2\pi abE^*}b\mathbf{K}(e)$$

$$= \left(\frac{9P^2}{16E^{*2}R_e}\right)^{1/3}\frac{2}{\pi}\left(\frac{b}{a}\right)^{1/2}\{F_1(e)\}^{-1}\mathbf{K}(e)$$

$$= \left(\frac{9P^2}{16E^{*2}R_e}\right)^{1/3} F_2(e) \tag{4.31}$$

and the maximum pressure is given by

$$p_0 = \frac{3P}{2\pi ab} = \left(\frac{6PE^{*2}}{\pi^3 R_e^{\,2}}\right)^{1/3}\{F_1(e)\}^{-2} \tag{4.32}$$

The eccentricity of the contact ellipse, which is independent of the load and depends only on the ratio of the relative curvatures (R'/R''), is given by equation (4.28). It is apparent from equation (4.3) that, before deformation, contours of constant separation h are ellipses in which $(b/a) = (A/B)^{1/2} = (R''/R')^{1/2}$. Equation (4.28) has been used to plot the variation of $(b/a)(B/A)^{1/2}$ as a function of $(B/A)^{1/2}$ in Fig. 4.4. If the contact ellipse had the same shape as contours of equal separation when the surfaces just touch, $(b/a)(B/A)^{1/2}$ would always have the value 1.0. It may be seen from the figure that $(b/a)(B/A)^{1/2}$ decreases from unity as the ratio of relative curvatures (R'/R'') increases. Thus the contact ellipse is somewhat more slender than the ellipse of constant separation. The broken line in Fig. 4.4 shows that $(b/a)(B/A)^{1/2} \approx (B/A)^{-1/6}$, i.e.

$$b/a \approx (B/A)^{-2/3} = (R'/R'')^{-2/3} \tag{4.33}$$

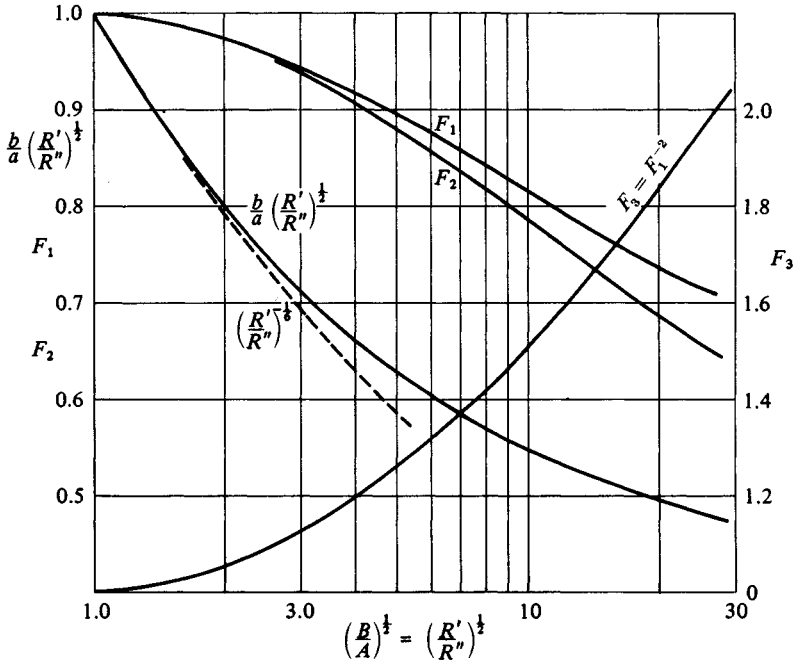
We have introduced an equivalent contact radius $c\ (= (ab)^{1/2})$ and an equivalent

https://doi.org/10.1017/CBO9781139171731.005 Published online by Cambridge University Press

relative curvature R_e $(= (R'R'')^{1/2})$ and obtained expressions for c, the maximum contact pressure p_0 and the compression δ in equations (4.30), (4.31) and (4.32). Comparison with the corresponding equations (4.22), (4.23) and (4.24) for solids of revolution shows that the first term is the same in each case; the second term may be regarded as a 'correction factor' to allow for the eccentricity of the ellipse. These correction factors – $F_1(e)$, $\{F_1(e)\}^{-2}$ and $F_2(e)$ – are also plotted against $(R'/R'')^{1/2}$ in Fig. 4.4; they depart rather slowly from unity with increasing ellipticity.

As an example, consider the contact of the cylinders, each of radius R, with their axes inclined at 45°, illustrated by the interference fringes shown in Fig. 4.1(b). As shown in §4.1 the ratio of relative curvatures $(R'/R'')^{1/2} = (B/A)^{1/2} = 2.41$ and the equivalent radius $R_e = (R'R'')^{1/2} = \sqrt{(2)}R$. Under load, the ratio of major to minor axis $a/b = 3.18$ from the curve in Fig. 4.4 or ≈ 3.25 from equation (4.33). Also from Fig. 4.4, $F_1 \approx F_2 = 0.95$ and $F_1^{-2} \approx 1.08$ from which the contact size c $(= (ab)^{1/2})$, the compression δ and the contact pressure p_0 can be found using equations (4.30), (4.31) and (4.32) respectively. Even

Fig. 4.4. Contact of bodies with general profiles. The shape of the ellipse b/a and the functions F_1, F_2 and F_3 $(= F_1^{-2})$ in terms of the ratio (R'/R'') of relative curvatures, for use in equations (4.30), (4.31) and (4.32).



https://doi.org/10.1017/CBO9781139171731.005 Published online by Cambridge University Press

though the ellipse of contact has a 3:1 ratio of major to minor axes, taking $F_1 = F_2 = F_3 = 1.0$, i.e. using the formulae for circular contact with an equivalent radius R_e, leads to overestimating the contact size c and the compression δ by only 5% and to underestimating the contact pressure p_0 by 8%.

For ease of numerical computation various authors, e.g. Dyson (1965) and Brewe & Hamrock (1977) have produced approximate algebraic expressions in terms of the ratio (A/B) to replace the elliptic integrals in equations (4.30), (4.31) and (4.32). Tabular data have been published by Cooper (1969).

It has been shown that a semi-ellipsoidal pressure distribution acting over an elliptical region having the dimensions defined above satisfies the boundary conditions (4.7) within the ellipse. To confirm the hypothesis that the contact area is in fact elliptical it is necessary that condition (4.8) also be satisfied: that there should be no contact outside the prescribed ellipse. From equation (3.59), the displacements on the surface outside the loaded ellipse are given by

$$\bar{u}_z = \frac{1-\nu^2}{E}\,\frac{\pi ab}{2}\,p_0 \int_{\lambda_1}^{\infty} \left(1 - \frac{x^2}{a^2+w} - \frac{y^2}{b^2+w}\right) \frac{\mathrm{d}w}{\{(a^2+w)(b^2+w)w\}^{1/2}}$$

where λ_1 is the positive root of equation (3.53). We write $\int_{\lambda_1}^{\infty} \equiv [I]_{\lambda_1}^{\infty} = [I]_0^{\infty} - [I]_0^{\lambda_1}$. In the region in question: $z = 0, x^2/a^2 + y^2/b^2 > 1$, from which it appears that $[I]_0^{\lambda_1}$ is negative. But the pressure distribution and contact dimensions have been chosen such that

$$\frac{ab}{2E^*}\,p_0\,[I]_0^{\infty} = \delta - Ax^2 - By^2.$$

Hence in the region outside the contact

$$\bar{u}_{z1} + \bar{u}_{z2} > \delta - Ax^2 - By^2$$

i.e. condition (4.8) is satisfied and the assumption of an elliptical contact area is justified.

Expressions for the stresses within the solids are given by equations (3.64)-(3.69). The general form of the stress field is similar to that in which the contact region is circular. If a and b are taken in the x and y directions respectively with $a > b$, at the centre of the contact surface

$$\sigma_x = -p_0\{2\nu + (1-2\nu)b/(a+b)\} \tag{4.34a}$$

$$\sigma_y = -p_0\{2\nu + (1-2\nu)a/(a+b)\} \tag{4.34b}$$

At the ends of the major and minor axes, which coincide with the edge of the contact region, there is equal tension and compression in the radial and circumferential directions respectively, thus at $x = \pm a, y = 0$,

$$\sigma_x = -\sigma_y = p_0(1-2\nu)\,\frac{b}{ae^2}\left\{\frac{1}{e}\tanh^{-1} e - 1\right\} \tag{4.35a}$$

https://doi.org/10.1017/CBO9781139171731.005 Published online by Cambridge University Press

Table 4.1

b/a	0	0.2	0.4	0.6	0.8	1.0
z/b	0.785	0.745	0.665	0.590	0.530	0.480
$(\tau_1)_{max}/p_0$	0.300	0.322	0.325	0.323	0.317	0.310

and at $x = 0, y = \pm b$,

$$\sigma_y = -\sigma_x = p_0(1-2\nu)\frac{b}{ae^2}\left\{1 - \frac{b}{ae}\tan^{-1}\left(\frac{ea}{b}\right)\right\} \tag{4.35b}$$

The maximum shear stress occurs on the z-axis at a point beneath the surface whose depth depends upon the eccentricity of the ellipse as given in Table 4.1. Numerical values of the stresses along the z-axis have been evaluated by Thomas & Hoersch (1930) for $\nu = 0.25$ and by Lundberg & Sjövall (1958).

The simplest experimental check on the validity of the Hertz theory is to measure the growth in size of the contact ellipse with load which, by (4.30), is a cube-root relationship. Hertz performed this experiment using glass lenses coated with lampblack. A thorough experimental investigation has been carried out by Fessler & Ollerton (1957) in which the principal shear stresses on the plane of symmetry, given by equations (3.64) and (3.69) have been measured using the frozen stress method of photo-elasticity. The ratio of the major axis of the contact ellipse a to the minimum radius of curvature R was varied from 0.05 to 0.3 with araldite models having different combinations of positive and negative curvature. At the smallest values of a/R the measured contact size was somewhat greater than the theory predicts. This discrepancy is commonly observed at light loads and is most likely due to the topographical roughness of the experimental surfaces (see Chapter 13). At high loads there was good agreement with the theoretical predictions of both contact area and internal stress up to the maximum value of a/R used (= 0.3). This reassuring conclusion is rather surprising since this value of (a/R) corresponds to strains in the contact region rising to about 10%.

(c) Two-dimensional contact of cylindrical bodies

When two cylindrical bodies with their axes both lying parallel to the y-axis in our coordinate system are pressed in contact by a force P per unit length, the problem becomes a two-dimensional one. They make contact over a long strip of width $2a$ lying parallel to the y-axis. Hertz considered this case as the limit of an elliptical contact when b was allowed to become large compared with a. An alternative approach is to recognise the two-dimensional

https://doi.org/10.1017/CBO9781139171731.005 Published online by Cambridge University Press

nature of the problem from the outset and to make use of the results developed in Chapter 2 for line loading of a half-space.

Equation (4.3) for the separation between corresponding points on the unloaded surfaces of the cylinders becomes

$$h = z_1 + z_2 = Ax^2 = \tfrac{1}{2}(1/R_1 + 1/R_2)x^2 = \tfrac{1}{2}(1/R)x^2 \tag{4.36}$$

where the relative curvature $1/R = 1/R_1 + 1/R_2$. For points lying within the contact area after loading, equation (4.7) becomes

$$\bar{u}_{z1} + \bar{u}_{z2} = \delta - Ax^2 = \delta - \tfrac{1}{2}(1/R)x^2 \tag{4.37}$$

whilst for points outside the contact region

$$\bar{u}_{z1} + \bar{u}_{z2} > \delta - \tfrac{1}{2}(1/R)x^2 \tag{4.38}$$

We are going to use Hertz' approximation that the displacements $\bar{u}_{z1}$ and $\bar{u}_{z2}$ can be obtained by regarding each body as an elastic half-space, but a difficulty arises here which is absent in the three-dimensional cases discussed previously. We saw in Chapter 2 that the value of the displacement of a point in an elastic half-space loaded two-dimensionally could not be expressed relative to a datum located at infinity, in view of the fact that the displacements decrease with distance r from the loaded zone as $\ln r$. Thus $\bar{u}_{z1}$ and $\bar{u}_{z2}$ can only be defined relative to an arbitrarily chosen datum. The approach of distant points in the two cylinders, denoted by δ in equation (4.37), can take any value depending upon the choice of datum. In physical terms this means that the approach δ cannot be found by consideration of the local contact stresses alone; it is also necessary to consider the stress distribution within the bulk of each body. This is done for circular cylinders in §5.6.

For the present purpose of finding the local contact stresses the difficulty is avoided by differentiating (4.37) to obtain a relation for the surface *gradients*. Thus

$$\frac{\partial \bar{u}_{z1}}{\partial x} + \frac{\partial \bar{u}_{z2}}{\partial x} = -(1/R)x \tag{4.39}$$

Referring to Chapter 2, we see that the surface gradient due to a pressure $p(x)$ acting on the strip $-a \leqslant x \leqslant a$ is given by equation (2.25*b*). The pressure on each surface is the same, so that

$$\frac{\partial \bar{u}_{z1}}{\partial x} + \frac{\partial \bar{u}_{z2}}{\partial x} = -\frac{2}{\pi E^*} \int_{-a}^{a} \frac{p(s)}{x - s}\, \mathrm{d}s$$

Substituting in equation (4.39)

$$\int_{-a}^{a} \frac{p(s)}{x - s}\, \mathrm{d}s = \frac{\pi E^*}{2R}\, x \tag{4.40}$$

https://doi.org/10.1017/CBO9781139171731.005 Published online by Cambridge University Press

This is an integral equation for the unknown pressure $p(x)$ of the type (2.39) in which the right-hand side $g(x)$ is a polynomial in x of first order. The solution of this type of equation is discussed in §2.7. If, in equation (2.45), we put $n = 1$ and write $\pi E(n+1)B/2(1-\nu^2) = \pi E^*/2R$, the required distribution of pressure is given by equation (2.48) in which, by (2.47),

$$I_n = I_1 = \pi(x^2/a^2 - \tfrac{1}{2})$$

Thus

$$p(x) = -\frac{\pi E^*}{2R}\frac{x^2 - a^2/2}{\pi(a^2-x^2)^{1/2}} + \frac{P}{\pi(a^2-x^2)^{1/2}} \tag{4.41}$$

This expression for the pressure is not uniquely defined until the semi-contact-width a is related to the load P. First we note that the pressure must be positive throughout the contact for which

$$P \geqslant \pi a^2 E^*/4R \tag{4.42}$$

If P exceeds the value given by the right-hand side of (4.42) then the pressure rises to an infinite value at $x = \pm a$. The profile of an elastic half-space which is loaded by a pressure distribution of the form $p_0(1-x^2/a^2)^{-1/2}$ is discussed in §2.7(*a*). The surface gradient just outside the loaded region is infinite (see Fig. 2.12). Such a deformed profile is clearly inconsistent with the condition of our present problem, expressed by equation (4.38), that contact should not occur outside the loaded area. We must conclude therefore that

$$P = \pi a^2 E^*/4R$$

i.e.

$$a^2 = \frac{4PR}{\pi E^*} \tag{4.43}$$

whereupon

$$p(x) = \frac{2P}{\pi a^2}(a^2 - x^2)^{1/2} \tag{4.44}$$

which falls to zero at the edge of the contact.

The maximum pressure

$$p_0 = \frac{2P}{\pi a} = \frac{4}{\pi}p_{\mathrm{m}} = \left(\frac{PE^*}{\pi R}\right)^{1/2} \tag{4.45}$$

where p_{m} is the mean pressure.

The stresses within the two solids can now be found by substituting the pressure distribution (4.44) into equation (2.23). At the contact interface $\sigma_x = \sigma_y = -p(x)$; outside the contact region all the stress components at the

https://doi.org/10.1017/CBO9781139171731.005 Published online by Cambridge University Press

surface are zero. Along the z-axis the integration is straightforward giving

$$\sigma_x = -\frac{p_0}{a}\{(a^2 + 2z^2)(a^2 + z^2)^{-1/2} - 2z\} \tag{4.46a}$$

$$\sigma_z = -\frac{p_0}{a}(a^2 + z^2)^{-1/2} \tag{4.46b}$$

These are principal stresses so that the principal shear stress is given by

$$\tau_1 = -\frac{p_0}{a}\{z - z^2(a^2 + z^2)^{-1/2}\}$$

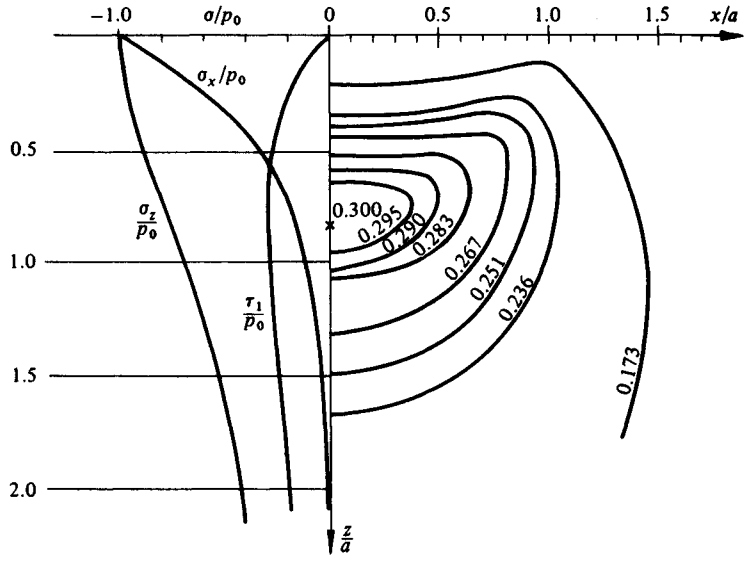
from which

$$(\tau_1)_{\max} = 0.30p_0, \quad \text{at } z = 0.78a \tag{4.47}$$

These stresses are all independent of Poisson's ratio although, for plane strain, the third principal stress $\sigma_y = \nu(\sigma_x + \sigma_z)$. The variations of σ_x, σ_z and τ_1 with depth below the surface given by equations (4.46) are plotted in Fig. 4.5(a), which may be seen to be similar to the variations of stress beneath the surface in a circular contact (Fig. 4.3). Contours of principal shear stress τ_1 are plotted in Fig. 4.5(b), which may be compared with the photo-elastic fringes shown in Fig. 4.6(d). McEwen (1949) expresses the stresses at a general point (x, z) in terms of m and n, defined by

$$m^2 = \tfrac{1}{2}[\{(a^2 - x^2 + z^2)^2 + 4x^2z^2\}^{1/2} + (a^2 - x^2 + z^2)] \tag{4.48a}$$

Fig. 4.5. Contact of cylinders: (a) subsurface stresses along the axis of symmetry, (b) contours of principal shear stress τ_1.



(a) (b)

https://doi.org/10.1017/CBO9781139171731.005 Published online by Cambridge University Press

and

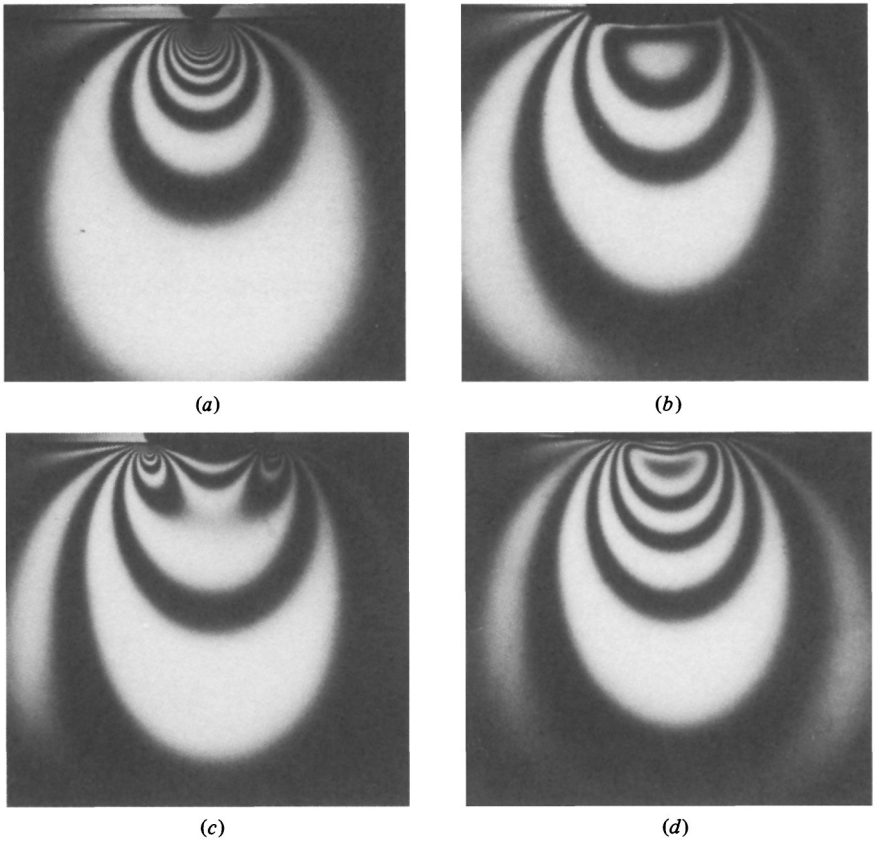
$$n^2 = \tfrac{1}{2}[\{(a^2 - x^2 + z^2)^2 + 4x^2z^2\}^{1/2} - (a^2 - x^2 + z^2)] \tag{4.48b}$$

where the signs of m and n are the same as the signs of z and x respectively. Whereupon

$$\sigma_x = -\frac{p_0}{a}\left\{m\left(1 + \frac{z^2 + n^2}{m^2 + n^2}\right) - 2z\right\} \tag{4.49a}$$

$$\sigma_z = -\frac{p_0}{a}\, m\left(1 - \frac{z^2 + n^2}{m^2 + n^2}\right) \tag{4.49b}$$

Fig. 4.6. Two-dimensional photo-elastic fringe patterns (contours of principal shear stress): (*a*) point load (§2.2); (*b*) uniform pressure (§2.5(*a*)); (*c*) rigid flat punch (§2.8); (*d*) contact of cylinders (§4.2(*c*)).



(*a*) (*b*)

(*c*) (*d*)

https://doi.org/10.1017/CBO9781139171731.005 Published online by Cambridge University Press

and

$$\tau_{xz} = -\frac{p_0}{a} n\left(\frac{m^2 - z^2}{m^2 + n^2}\right) \tag{4.49c}$$

Alternative expressions have been derived by Beeching & Nicholls (1948), Poritsky (1950), Sackfield & Hills (1983*a*). A short table of values is given in Appendix 4. The variation of stress with x at a constant depth $z = 0.5a$ is shown in Fig. 9.3.

4.3 Elastic foundation model

The difficulties of elastic contact stress theory arise because the displacement at any point in the contact surface depends upon the distribution of pressure throughout the whole contact. To find the pressure at any point in the contact of solids of given profile, therefore, requires the solution of an integral equation for the pressure. This difficulty is avoided if the solids can be modelled by a simple Winkler elastic foundation or 'mattress' rather than an elastic half-space. The model is illustrated in Fig. 4.7. The elastic foundation, of depth h, rests on a rigid base and is compressed by a rigid indenter. The profile of the indenter, $z(x, y)$, is taken as the sum of the profiles of the two bodies being modelled, i.e.

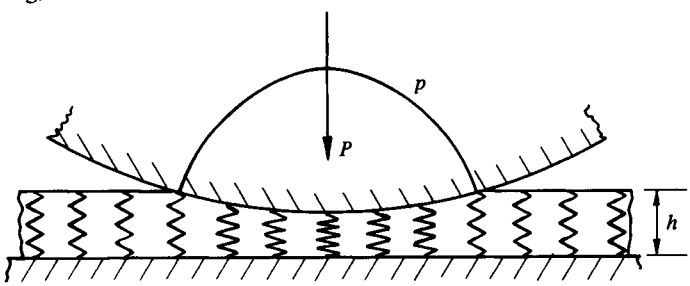
$$z(x, y) = z_1(x, y) + z_2(x, y) \tag{4.50}$$

There is no interaction between the springs of the model, i.e. shear between adjacent elements of the foundation is ignored. If the penetration at the origin is denoted by δ, then the normal elastic displacements of the foundation are given by

$$\bar{u}_z(x, y) = \begin{cases} \delta - z(x, y), & \delta > z \\ 0, & \delta \leqslant z \end{cases} \tag{4.51}$$

The contact pressure at any point depends only on the displacement at that

Fig. 4.7



https://doi.org/10.1017/CBO9781139171731.005 Published online by Cambridge University Press

point, thus

$$p(x,y) = (K/h)\bar{u}_z(x,y) \tag{4.52}$$

where K is the elastic modulus of the foundation.

For two bodies of curved profile having relative radii of curvature R' and R'', $z(x,y)$ is given by equation (4.3) so that we can write

$$\bar{u}_z = \delta - (x^2/2R') - (y^2/2R'') \tag{4.53}$$

inside the contact area. Since $\bar{u}_z = 0$ outside the contact, the boundary is an ellipse of semi-axes $a = (2\delta R')^{1/2}$ and $b = (2\delta R'')^{1/2}$. The contact pressure, by (4.52), is

$$p(x,y) = (K\delta/h)\{1 - (x^2/a^2) - (y^2/b^2)\} \tag{4.54}$$

which is paraboloidal rather than ellipsoidal as given by the Hertz theory. By integration the total load is

$$P = K\pi ab\delta/2h \tag{4.55}$$

In the axi-symmetric case $a = b = (2\delta R)^{1/2}$ and

$$P = \frac{\pi}{4}\left(\frac{Ka}{h}\right)\frac{a^3}{R} \tag{4.56}$$

For the two-dimensional contact of long cylinders, by equation (4.37)

$$\bar{u}_z = \delta - x^2/2R = (a^2 - x^2)/2R \tag{4.57}$$

so that

$$p(x) = (K/2Rh)(a^2 - x^2) \tag{4.58}$$

and the load

$$P = \tfrac{2}{3}\left(\frac{Ka}{h}\right)\frac{a^2}{R} \tag{4.59}$$

Equations (4.56) and (4.59) express the relationship between the load and the contact width. Comparing them with the corresponding Hertz equations (4.22) and (4.43), agreement can be obtained, if in the axi-symmetric case we choose $K/h = 1.70E^*/a$ and in the two-dimensional case we choose $K/h = 1.18E^*/a$. For K to be a material constant it is necessary to maintain geometrical similarity by increasing the depth of the foundation h in proportion to the contact width a. Alternatively, thinking of h as fixed requires K to be reduced in inverse proportion to a. It is a consequence of the approximate nature of the model that the values of K required to match the Hertz equations are different for the two configurations. However, if we take $K/h = 1.35E^*/a$, the value of a under a given load will not be in error by more than 7% for either line or point contact.

The compliance of a point contact is not so well modelled. Due to the neglect of surface displacements outside the contact, the foundation model gives

$\delta = a^2/2R$ which is half of that given by Hertz (equation (4.23)). If it were more important in a particular application to model the compliance accurately we should take $K/h = 0.60E^*/a$; the contact size a would then be too large by a factor of $\sqrt{2}$.

The purpose of the foundation model, of course, is to provide simple approximate solutions in complex situations where half-space theory would be very cumbersome. For example, the normal frictionless contact of bodies whose arbitrary profiles cannot be represented adequately by their radii of curvature at the point of first contact can be handled easily in this way (see §5.3). The contact area is determined directly in shape and size by the profiles $z(x, y)$ and the penetration δ. The pressure distribution is given by (4.52) and the corresponding load by straight summation of pressure. For a contact area of arbitrary shape a representative value of a must be chosen to determine (K/h).

The foundation model is easily adapted for tangential loading (see §8.7); also to viscoelastic solids (see §9.4).