

CHAPTER 3

THE CONTACT MECHANICS OF SOLID BODIES SUCH AS WHEEL ON RAIL.

3.1 Nomenclature

A, B, C	Coefficients in equations 3.1 to 3.9.
a	Radius of a circular Hertzian area of contact.
a	Half the contact width of the Hertzian contact strip for cylindrical (line) contact. (This contact can be regarded as a contact ellipse with a major axis "a" of infinite length; i.e., one convention has half the contact width, the ellipse minor axis, as "b". However when considering purely elliptical contact, it is commonly taken as a".)
a, b	Major and minor semi-axes of the contact ellipse of two bodies loaded against each other.
a, b	Half lengths of approximated contact rectangle (cf. Equ. 3.32).
a', b'	Major and minor semi-axes of the elliptical contours of separation of two bodies.
C_{11} , C_{22} etc.	Non dimensional creep coefficients.
c	$c = \sqrt{ab}$ where "a and b" are the contact ellipse semi-axes.
D(e), K(e)	Complete integrals of argument, $e = \sqrt{1-b^2/a^2}$, $a > b$. (Equations 3.8 to 3.13.)

Nomenclature continued...

E	The elastic modulus.
E*	The "effective elastic modulus" where, $1/E^* = (1-\nu_1^2)/E_1 + (1-\nu_2^2)/E_2$
F	Force to overcome rolling resistance in free rolling.
F ₁ , F ₂	Functions of the contact ellipse shape, related to D(e) and K(e). (Equations 3.14 to 3.16.)
G	Modulus of rigidity.
h	Separation between two bodies just touching.
h	Hydrostatic component of stress for material in compression where, $h = 1/3.(\sigma_1 + \sigma_2 + \sigma_3)$
k	Yield stress in simple shear.
k _e	Initial yield stress of material in shear.
L	Length of cylindrical (line) contact.
M _z	Spin traction.
O _{xyz}	Cartesian coordinate system.
P	Normal load or force.

Nomenclature continued

PS	Proof stress.
P_s	Normal load for shakedown.
P_y	Normal load for first yield.
p	Normal contact stress.
p_m	Mean contact stress.
p_o	Maximum contact stress.
p_o^s or $(p_o)^s$	Maximum contact stress at a shakedown limit.
$(p_o)_Y$	Maximum contact stress to initiate yield.
Q	Tangential load or force.
Q/P	"Traction coefficient" where $Q < \mu P$, (i.e. where there is "static friction").
q	Tangential stress.
q_o	Maximum tangential stress.
R'_1, R'_2	Maximum principal radii of curvature for bodies 1 and 2.
R''_1, R''_2	Minimum principal radii of curvature for bodies 1 and 2.

Nomenclature continued

R', R''	Principal "relative" maximum and minimum radii of curvature where, $1/R' = 1/R'_1 + 1/R'_2$ and $1/R'' = 1/R''_1 + 1/R''_2$
R_e	The "equivalent relative radius of curvature" for bodies 1 and 2 in loaded contact where, $R_e = \sqrt{(R'R'')}$.
R_w	Radius of railway wheel.
S	Surface point on body in contact.
s_x, s_y	Velocities of microslip between contacting points in steady rolling. (Put dots above "s"? ask James)
T	Internal distance point of body in contact.
u_z	Elastic displacement of surface point on body in the z direction.
u_x, u_y	Components of tangential elastic displacement at a surface point. (Put dots above "u"?)
V	Velocity of rolling body along the "x" axis. When considering creepage, $V = \frac{1}{2} \cdot (V_T + V_C)$.
V_C	Circumferential velocity of wheel at its yawed angle.
V_T	Translational velocity of wheel along the rail.
$V\phi$	Velocity due to yaw of wheel on rail.

Nomenclature continued

Y	Yield stress in simple tension.
z_1, z_2	Surface profile functions for bodies 1 & 2 in contact. (Equations 3.1a, 3.1b).
δ	(Bulk) approach of two bodies along the centre line O_z .
δ_1, δ_2	Maximum elastic displacements of bodies 1 and 2 along the centre line O_z .
δ_v	Creep velocities.
ξ_r	Relative rigid slip where " $\xi_r V$ " is "rigid slip".
ξ_x, ξ_y	Creep ratios (longitudinal and lateral creepage).
θ	Railway wheel cone angle (cf. Figure 3.8).
λ	Coefficient of rolling resistance (F/P).
λ	Angle between vertical axes of wheel and rail.
μ	Coefficient of (kinetic) friction.
ν	Poisson's ratio.
$\sigma_1, \sigma_2, \sigma_3$	Principal direct stresses.
$\sigma_x, \sigma_y, \sigma_z$	Principal direct stresses acting along x, y, z axes.

Nomenclature continued

τ	Shear stress.
τ_1	Principal shear stress.
τ_{\max}	Maximum shear stress.
τ_{zx}	Orthogonal shear stress.
$(\tau_{zx})_r$	Residual shear stress.
ϕ	Yaw angle of wheel on rail.
ψ	Spin parameter/creepage.
ω	Angle between the principal planes of curvature of two bodies . (Equations 3.2 to 3.5).
ω_z	Angular velocity due to spin.

3.2 Introduction

This chapter examines the mechanics of loaded contact between solid, non-conforming bodies such as a wheel on a rail. Initially the stresses and deformations of static contact, generated from a normal force, are considered. Then further consideration is given to the effect of relative motions between the bodies and the effect of applying tangential forces.

Even at the highest axle loads on the British Rail network (10 tonnes) the contact area between the wheel and rail is approximately 1 cm^2 ; i.e. the system can be viewed as a large bearing race and, as in bearing contact, very small, highly compressed regions are constrained within a bulk of relatively unstrained material. This small area of contact, with its relatively low frictional resistance, means that rail vehicles are highly efficient movers of mass compared to road vehicles. In this work only the dry contact of the wheel on rail is considered, whereas on Japanese and European rail networks there is usually a light slurry of water, oil droplets, wear and chemically induced ferrous oxides and biological products, i.e., a "boundary lubricant". This can slightly reduce friction. Rain can further reduce friction and the consequent reduction in wear can accelerate other problems such as rolling contact fatigue*. (* This has been separately studied by the author^[Garnham & Beynon, 1990 (see Appendix II); Beynon et al, 1994].)

The quoted mechanical properties of materials, as determined in laboratory tests (including some fatigue and stress corrosion tests), are almost always based upon their strength in uniaxial tension. Such testpieces are not constrained by adjacent material and are free to thin in cross-sectional area whilst being stretched. The behaviour of small areas of material under compressive loads, constrained within a relatively unstressed matrix, differs. Loading above the tensile yield point is commonplace. As a general guide^[Halling 1975], metals have compressive elastic behaviour up to mean contact pressures of $1.5Y$ or $3k$ (where "Y" is the uniaxial tensile yield stress and "k" is the yield stress in shear). From $1.5Y$ to $3Y$, or $3k$ to $6k$, a sub-surface element of material behaves in a plastic manner but it is constrained by surrounding elastic material. This zone increases

in size with increasing load and it will finally spread to the surface of the body such that, above 3Y or 6k, the material behaves in a fully plastic manner. This is akin to the hardness test where, for metals, $H \approx 3Y \approx 6k$ (with an assumption of zero work-hardening). There is no accepted practice for determining the mechanical properties of materials in compression due to the influence of friction on testpiece holding and the variability of testpiece geometry; test areas are not constrained within an unloaded matrix.

For an initial analysis, estimations of contact areas and material deformations are still based upon the theories of Hertz^[1882]. These covered the frictionless, elastic, compressive contact of non-conforming solid bodies. His basic formulae are still used for rapid determinations of contact stress, area and deformation in engineering contacts including rail-wheel. More sophisticated computerised systems have recently been developed to give more precise rail-wheel data^[Kalker 1979,1990]. These must account for the relative motion of the bodies (roll, slide and spin) in macro and micro movements with respect to differing wheel-rail geometries and wheelset-rail alignments.

This chapter primarily addresses the "bulk" consequences of loaded contact between solid bodies and the effects within the area of contact formed. For the most part, it will assume that the bodies have perfectly smooth surfaces. However, at a microscopic level, engineered metal surfaces are never smooth and hence real initial contact occurs between the peaks of asperities. The loads are such that these will plastically deform until the load can be supported by a network of semi-flattened asperities; i.e. the real area of contact. The effect of such surface profiles and the generation of high stresses at a near-surface microscopic level is of significance with respect to static and sliding friction and the generation of surface initiated micro-cracks. These factors are discussed (cf. section 3.9).

3.3 Normal loading of elastic, isotropic bodies without other movements.

This situation was investigated by Hertz^[1882]. His work has been described by K.L. Johnson in various papers and in his comprehensive book on contact mechanics^[1985]. Some of Johnson's approach is summarised here. Two non-conforming solid bodies approaching each other must initially touch at a point, except in the case of aligned cylindrical contact, where initial contact is along a line. The surfaces will elastically compress and form an area of contact.

Hertzian Theory makes the following assumptions:

- * The surfaces are frictionless.
- * The surfaces are continuous and non-conforming.
- * The contact area is very small compared with the bulk volume of the two bodies.
- * The contact stresses are large compared with other stresses in the two bodies.
- * Subsequent strains are small.
- * Each body can be regarded as an elastic half-space bounded by the plane $z = 0$.

The approach of two bodies.

For the initial treatment of this approach let two, curved, irregularly shaped objects meet at a point O, approaching along a common coordinate z , such that they form a contact area in a tangent plane described by the coordinates x and y , both of which are at 90° to each other and to z (Figure 3.1). Near to the point of contact, the shape of the bodies can be approximately,

$$z_1 = A_1x^2 + B_1y^2 + C_1xy + \dots \quad (3.1a)$$

with a similar expression for z_2 when higher order terms have been neglected. The "containing" shape of two unloaded bodies, 1 and 2, can be described by their respective principal ("maximum and minimum") *radii* of curvature at the point of contact, R'_1 , R''_1 and R'_2 and R''_2 . (Note, concave surfaces will be negative.) If the orientation of the x and y axes are chosen such that the term in " C_{xy} " is zero, the

surface profiles, z_1 and z_2 , can be expressed as,

$$z_1 = \frac{1}{R'_1} x_1^2 + \frac{1}{R''_2} y_1^2 \quad (3.1b)$$

$$z_2 = - \left(\frac{1}{R'_2} x_2^2 + \frac{1}{R''_2} y_2^2 \right) \quad (3.1c)$$

If they just touch without deformation then the small separation between them, parallel to the z axis, can be described by h , where $h = z_1 - z_2$. Therefore for a common set of suitably orientated x and y axes, from Equation 3.1,

$$h = Ax^2 + By^2 = (\text{for } \omega = 0) \frac{1}{2R'} x^2 + \frac{1}{2R''} y^2 \quad (3.2)$$

where "A" and "B" (with $B \gg A$) are constants, whose value depends upon the magnitude of the principal curvatures, and R' and R'' are defined as the principal *relative* radii of curvature where $1/R' = 1/R'_1 + 1/R'_2$ and $1/R'' = 1/R''_1 + 1/R''_2$. Only where the planes of principal curvature of the two bodies coincide (i.e. $\omega = 0$) does,

$$A = \frac{1}{R'} ; \quad B = \frac{1}{R''} \quad (3.3)$$

When the axes are not common ($\omega \neq 0$ as in Figure 3.1), it can be shown that^[Johnson 1985],

$$A + B = \frac{1}{2} \left(\frac{1}{R'} + \frac{1}{R''} \right) = \frac{1}{2} \left(\frac{1}{R'_1} + \frac{1}{R''_1} + \frac{1}{R'_2} + \frac{1}{R''_2} \right) \quad (3.4)$$

and

$$B - A = \frac{1}{2} \sqrt{\left[\left(\frac{1}{R'_1} - \frac{1}{R''_1} \right)^2 + \left(\frac{1}{R'_2} - \frac{1}{R''_2} \right)^2 + 2 \left(\frac{1}{R'_1} - \frac{1}{R''_1} \right) \left(\frac{1}{R'_2} - \frac{1}{R''_2} \right) \cos 2\omega \right]} \quad (3.5)$$

The separation of the two bodies can be described by elliptical contours whose ratio of semi-axes $a'/b' = \sqrt{(B/A)}$. Examples of the inter-relationships of the "R" terms, with the contact of some regularly shaped objects, are given in Figure 3.2.

The compression of two bodies.

It is assumed here that the two bodies do not greatly differ in elasticity and that ω is zero; a situation applicable to most engineering contacts including wheel-rail contact. Figure 3.3 shows the compressive contact of two bodies. A normal load has been applied such that they have both elastically compressed and formed an area of contact. Remote parts of the bodies (T_1 and T_2) have approached each other by a distance δ . Points on the approaching contact surfaces (S_1 and S_2) have been elastically displaced by u_{z1} and u_{z2} . Therefore, from Equation 3.2, inside the contact area,

$$u_{z1} + u_{z2} = \delta - Ax^2 - By^2 \quad (3.6)$$

and outside the contact area,

$$u_{z1} + u_{z2} > \delta - Ax^2 - By^2 \quad (3.7)$$

At the contact centre ($x, y = 0$),

$$\delta = u_{z1} + u_{z2} = \delta_1 + \delta_2 \text{ (the maximum elastic displacements)}$$

Hertz Theory of Elasticity explains the pressure distributions required to satisfy Equations 3.6 and 3.7. Within the contact patch it shows that;

$$A = \frac{P_o}{E^*} \frac{b}{e^2 a^2} [K(e) - D(e)] \quad (3.8)$$

$$B = \frac{P_o}{E^*} \frac{b}{e^2 a^2} \left[\left(\frac{a}{b} \right)^2 D(e) - K(e) \right] \quad (3.9)$$

$$\delta = \frac{P_o}{E^*} [b K(e)] \quad (3.10)$$

where p_o is the maximum contact stress; a and b are the respective major and minor semi-axes of the contact ellipse and E^* is the "effective elastic modulus" with E^* given by,

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

where E_1 , E_2 and ν_1 , ν_2 are the respective Young's Moduli and Poisson's ratios for the two bodies. $K(e)$ and $D(e)$ are complete integrals of argument, $e = \sqrt{1-(b^2/a^2)}$, $a > b$.

The pressure distribution across the contact is semi-ellipsoidal therefore, for a total load P , the mean contact stress, p_m , is given by $p_m = P/\pi ab$ and the maximum contact stress, p_o , is given by,

$$P_o = \frac{3}{2} P_m = \frac{3}{2} \frac{P}{\pi ab} \quad (3.11)$$

The pressure distribution within the contact ellipse will be given by,

$$p(x,y) = p_o \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}} \quad (3.11a)$$

Determination of elliptical major and minor semi-axes, "a" and "b".

From Equations 3.3, 3.8 and 3.9,

$$\frac{B}{A} = \left(\frac{R'}{R''} \right) = \frac{(a/b)^2 \cdot D(e) - K(e)}{K(e) - D(e)} \quad (3.12)$$

and

$$\begin{aligned} \sqrt{(AB)} &= \frac{1}{2} \cdot \sqrt{(1/R'R'')} = \frac{1}{2} \cdot R_e \\ &= \frac{p_o}{E^*} \cdot \frac{b}{a^2 e^2} \cdot \sqrt{[(a/b)^2 \cdot [D(e) - K(e)] \cdot [K(e) - D(e)]]} \end{aligned} \quad (3.13)$$

where R_e {i.e. $\sqrt{(R'R'')}$ } is called the "equivalent relative radius of curvature".

Substituting p_o from Equations 3.11 into Equation 3.13,

$$(a/b)^{3/2} = \left(\frac{3PR_e}{4E^*} \right) \cdot \frac{4}{\pi e^2} \cdot (b/a)^{3/2} \cdot \sqrt{[(a/b)^2 D(e) - K(e)][K(e) - D(e)]}$$

From this, (ab) can be expressed thus,

$$\sqrt{(ab)} = \left(\frac{3PR_e}{4E^*} \right)^{1/3} \cdot F_1(e) \quad (3.14)$$

p_o thus,

$$p_o = \frac{3P}{2\pi ab} = \left(\frac{6PE'^2}{\pi^3 R_e^2} \right) [F_1(e)]^{-2} \quad (3.15)$$

and displacement δ (from Equations 3.10 and 3.11) thus,

$$\delta = \frac{3P}{2\pi abE'} \cdot bK(e) = \left(\frac{9P^2}{16E'^2 R_e^2} \right)^{1/3} \cdot F_2(e) \quad (3.16)$$

where F_1 and F_2 are functions of ellipse shape, related to $K(e)$ and $D(e)$. The values of these functions and the shape of the contact ellipse (b/a) are given against values of $\sqrt{(R'R'')}$ in Figure 3.4^[from Johnson 1982, 1985].

Equations 3.14, 3.15 and 3.16 can be viewed in the manner of the first terms representing a sphere of radius R_e in contact with a plane (or the contact of spheres of radii R_1, R_2 , if R is substituted for R_e and $1/R = 1/R_1 + 1/R_2$) whilst the second terms, containing functions F_1 and F_2 , are the correction factors for elliptical contact. Apart from Johnson's figure, in a different manner, other publications present charts of similar elliptical functions to help the engineer. For example, ESDU^[1978] present the above Hertzian analysis and equations in a different manner for easier engineering reference. Their function / coefficients can be read off charts against values of A/B . Dyson et al^[1992] have also produced a simplified method for determining Hertzian elliptical contact data.

If the ratio R'/R'' is small (i.e. approaching spherical contact) then the correction factors are close to unity, as seen in the top left corner of Figure 3.4.

Before contact, the contours of separation $(b'/a') = \sqrt{(A/B)} = \sqrt{(R''/R')}$. The ellipse shape is changed by loaded contact. This shape of the contact ellipse (b/a) is part of

Equation 3.12. This can be re-arranged to read:

$$b/a = \sqrt[3]{(R''/R')} \cdot \text{a function of } (e) \equiv \sqrt[3]{(R''/R')} \cdot \text{a function of } (R'/R'')$$

Therefore,

$$b/a \cdot \sqrt[3]{(R'/R'')} = \text{a function of } \sqrt[3]{(R'/R'')} \quad (3.17)$$

This relationship has been plotted in Figure 3.4, hence the relationship (b/a) can be determined. It can be seen from Figure 3.4 that a good approximation is,

$$b/a \cdot \sqrt[3]{(R'/R'')} \approx (R'/R'')^{-1/6}, \quad \text{i.e.} \quad b/a \approx (R'/R'')^{-2/3} = (B/A)^{-2/3} \quad (3.18)$$

Therefore the values of a and b can be determined from Equations 3.14 and 3.17, or approximated from 3.14 and 3.18.

Spherical contact

As the two bodies approach spherical shapes so the functions of (e), F_1 and F_2 (the right hand terms in Equations 3.14, 3.15, 3.16 and 3.17) approach unity, as shown in Figure 3.4. At unity, $R' = R''$, therefore these terms disappear and R replaces R_c in the left hand part of the equations, where $1/R = 1/R_1 + 1/R_2$ (R_1, R_2 now being the sphere radii). Now $P = \pi a^2 p_m$ where "a" is the radius of the contact circle, therefore,

$$p_o = \frac{3}{2} p_m = \frac{3P}{2\pi a^2} = \left(\frac{6PE^*}{\pi^3 R^2} \right)^{1/3} \quad (3.19)$$

$$a = \left(\frac{3PR}{4E^*} \right)^{1/3} \quad \left(= \frac{\pi p_o R}{2E^*} \right) \quad (3.20)$$

$$\delta = \frac{a^2}{R} = \left(\frac{9P^2}{16RE^*} \right)^{1/3} \quad \left(= \frac{\pi a p_o}{2E^*} \right) \quad (3.21)$$

Cylindrical (line or rectangular) contact.

[NOTE: With elliptical contact, "2a" is taken as the length of the major axis and "2b" as the length of the minor axis, therefore as the ellipse elongates towards line contact, "a" tends towards infinity and, one convention has it, "2b" is the subsequent contact width, for line contacts in cylindrical rolling movement in the "y" direction (eg, ESDU publications). However, for the sake of simplicity, many authors, including KL Johnson, use "2a" for the line contact width, with rolling and sliding in the "x" direction, and this practice is followed in this text.]

Wide cylindrical contact was the form of contact for the test disc results reported in the present work.

For the analysis of this condition, R'_1 and R'_2 are the respective cylinder radii with R''_1 and R''_2 becoming infinite. Therefore $1/R = 1/R'_1 + 1/R'_2$. (For a cylinder on a flat plane $R''_2 = \infty$ {Figure 3.2}.) Here P' is the load per unit axial length such that $p_m = P'/2a$ where "a" is half the contact width. Now,

$$p_o = \frac{4p_m}{\pi} = \frac{2P'}{\pi a} = \sqrt{\left(\frac{P'E}{\pi R} \right)} \quad (3.22)$$

$$a = \frac{2P'}{\pi p_o} = \sqrt{\left(\frac{4P'R}{\pi E^*} \right)} \quad (3.23)$$

For the line contact of two steel cylinders, where $\nu_1 \approx \nu_2 \approx 0.3$, Timoshenko and Goodier^[3rd Ed., 1951] have expressed the above formulae thus,

$$p_o = 0.418 \sqrt{\left(\frac{P'E}{R}\right)} \quad (3.24)$$

$$a = 1.52 \sqrt{\left(\frac{P'R}{E}\right)} \quad (3.25)$$

where $1/E = 1/E_1 + 1/E_2$. These formulae have been used throughout the experimentation described in the present work and E_1 and E_2 have been taken as 210 GPa. (Measured values for R52 and W64 were 209 and 199 GPa respectively.)

ESDU^[1978] give the approach of cylinders of "infinite" length (δ_c) which have the same values of "E" and "ν" as,

$$\delta_c = 2P'\pi E^* \cdot [2/3 + \log_e(4R'_1/a) + \log_e(4R'_2/a)] \quad (3.27)$$

3.4 Stress distributions for the normal compression of two bodies.

Derivations of the distributions are explained in Johnson^[1985]; summaries are given here.

Normal pressure distributions (i.e. in the "z" direction) across the contact areas for the three forms of contact are shown in Figure 3.5.

The principal direct stresses acting along the z,x,y axes, at, or below, the contact centre, are σ_z , σ_x and σ_y . The shear stresses acting along these axes at, or below, the contact centre are zero. The principal shear stresses act on the planes bisecting the angle between the z,x,y planes. Their magnitude (by Tresca's criterion) is given by,

$$(\sigma_z - \sigma_x)/2 \text{ in the zx plane} \quad (3.28a)$$

$$(\sigma_x - \sigma_y)/2 \text{ in the xy plane} \quad (3.28b)$$

$$(\sigma_y - \sigma_z)/2 \text{ in the yz plane} \quad (3.28c)$$

The distributions of the principal direct and shear stresses, at the centre of these types of contact (i.e. along the z-axis, the axis of symmetry $\{x,y=0\}$), are shown in Figure 3.6; i.e. the normalised stress magnitudes (divided by p_o) have been plotted against normalised sub-surface depth down the z axis (divided by the contact width "a"). Figure 3.6 shows that the overall maximum shear stress is *sub-surface* and lies between the zy planes, whereas the highest shear stress *at* the surface lies between the zx planes. (ESDU 78935^[1978] has a table showing the magnitude and location of maximum shear stress for changes in contact ellipse shape and Poisson's ratio.) The top plot of Figure 3.13a shows the contours of principal shear stress beneath such a contact.

The depth $[z/a]$ and value $[\tau_{\max}/p_o]$ of maximum shear for the symmetric configurations, with $\nu = 0.3$, are,

	z/a	τ_{\max}/p_o
Spherical contact :	0.48	0.31
Cylindrical contact :	0.78	0.30

For elliptical contact the magnitude and depth of maximum shear are variable with ellipse shape. An example (shown below) is given by Johnson^[1985] for two similar steels with $\nu = 0.25$:

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
τ_{\max}/p_o	0.300	0.322	0.325	0.323	0.317	0.310

For elliptical contact, direct stresses along the principal planes/axes, at the centre of contact SURFACE, are;

$$\sigma_z = p_o \quad (3.29)$$

$$\sigma_x = 2\nu.p_o + (1-2\nu).(b/\{a + b\}) \quad (3.30)$$

$$\sigma_y = 2\nu.p_o + (1-2\nu).(a/\{a + b\}) \quad (3.31)$$

3.5 The hertzian contact of wheel on rail.

In a simplified sense, wheel/rail contact can be represented by a conical surface (the wheel) on a hemi-spherical cylindrical surface (the rail head) as shown in Figure 3.7. The area of such a contact is elliptical. The true plane of loading (shown in Figure 3.8), along the line UT, is not normal to the wheel axis and the section of the wheel cone along that plane will be elliptical in shape.

With reference to Figure 3.8, from curvature theory it can be shown^[ESDU 78035, 1978] that the relevant principle radius of curvature of the wheel, R''_1 , is related to the ellipse by, $R''_1 = a^2/b$, and from geometry, the ellipse semi-axes are given by, $a = r \cdot \sin(90 + \theta) / \sin(90 - 2\theta)$ and $b = r + n \cdot \tan\theta$, where $n = a \cdot \sin\theta$. R''_2 is given by the rail geometry, R'_1 and R'_2 are both infinity and the angle between the principle planes of curvature, ω , is 90° . Therefore, from the approach described in Section 3.3, the dimensions of the contact ellipse for a given load and wheel cone angle (θ) can be calculated.

The ESDU example shows that for a cone angle of 2.5° and a wheel load of 100kN (10 tons force), the ellipse semi-axes will be 7.38 and 5.24mm, thus the contact area will be 1.2 cm^2 and the maximum contact stress, 1.24 GPa.

More detailed aspects of the practical considerations of rail based transport, including wheel/rail contact, can be found in Esveld's^[1989] comprehensive book. For elliptical contact, he shows that the maximum shear stress from hertzian contact, with a centred wheelset, will lie at a depth of $0.64a$, where "a" is the contact ellipse semi-axis along the rail (Figure 3.9). In tangent (i.e. straight) track this normally lies between 4 and 6mm beneath the rail head. The maximum shear stress occurs across the rail. In contrast to the hertzian values of maximum shear stress given in Section 3.3, he states that the Theory of Elasticity is not applicable for all cases, and that for practical considerations, the maximum shear stress, τ_{\max} , can be taken as $\approx 0.3p_m$, rather than $0.3p_o$. (Note, $p_o = 1.5p_m$ for elliptical contact.) He presents a simplified method [after Eisenmann^[1977],

of calculating wheel/rail contact stresses (Figure 3.9), where the contact is taken as cylindrical with all principle radii infinitely large except the wheel radius. Thus the contact can be approximated to a rectangle of dimensions $2a$ and $2b$ with a semi-elliptical axial pressure distribution.

From Equ. 3.22, assuming the steel and rail have the same values of " E " and " ν ", it can be shown that,

$$p_m = \sqrt{\left(\frac{\pi E}{64(1-\nu^2)} \frac{P}{rb} \right)} \quad (3.32)$$

thus for a standard situation where $E = 210 \text{ GPa}$, $\nu = 0.3$ and $b \approx 6\text{mm}$,

$$p_m = 1374 \cdot \sqrt{(P/r)} \text{ MPa} \quad (P [\text{kN}], r [\text{mm}]) \quad (3.33)$$

where $\tau_{\max} \approx 412 \cdot \sqrt{(P/r)}$.

Esveld qualifies these observations by saying that in tight curves, with rail lubrication, the Herizian elliptical stress distribution should be used and here τ_{\max} will lie between 2 and 4mm beneath the rail head, with a value approximately 50% higher than that given above. Note, Esveld's approximation is for elastic contact. Of parallel interest is the work of Kapoor and Johnson^[1992] where they show that plastic material flow within an elliptical contact will eventually alter the shape from elliptical to near-rectangular.

In reality, wheel and rail profiles are more complex than single cones or cylinders. For steady state curving, Cheesewright^[1981] has calculated/computed contact patch dimensions, shapes and maximum contact pressures for lateral displacements of various types of wheel profile, using basic hertzian analysis and two other, more sophisticated computer programs. Figure 3.10 shows the complex patterns generated by a high speed locomotive wheel (on this figure, 36mm represents the lateral position for a symmetrically aligned wheelset). Note the very high stresses with flange contact. Within a few months of service, "running-in" wear of the wheel and rail will trim down the highest stresses. Further complications are introduced by the yawing of rigid wheelsets

and independent movements of non-rigid wheelsets; a numerical solution for this problem has been devised by Duffek^[1982]. Kalker^[1990] has developed a program to simulate wheel wear, backed by track studies, and this has shown how wear effects wheel-rail contact positions, and thus, the contact stresses.

For some wheel profiles on European rail networks, there is a non-herzian form of contact with two separate elliptical contact zones. Pascal and Sauvage^[1993] have developed a simplified method of calculating forces for such contacts by approximating to a single ellipse. Their model also purports to explain the alignment of surface fatigue cracks.

3.6 The effect of tangential loading in combination with normal loading.

Friction

If a force is applied to the two non-conforming bodies so as to generate a pressure for one to slide against the other, this is opposed by a tangential force of friction "Q" (Figure 3.11). If the force is increased until sliding across the complete contact just occurs, then that is the point of "limiting friction" from where Amonton's Law applies, i.e. Q represents the force of "*kinetic friction*" such that $Q = \mu P$, where " μ " is the coefficient of (kinetic) friction and P is the applied normal load. If the force is insufficient to cause complete sliding (as in tractive propulsion of a locomotive for example), the average value of Q for the complete contact patch is variable and it is less than μP , i.e. a force of "*static friction*". Q/P ($< \mu$ for sliding) is often termed the "coefficient of traction".

Displacements:

The normal component of displacement, δ_z , due to the tangential force, is proportional to the elastic constant $(1-\nu)/G$ (where G is the Modulus of Rigidity / Bulk Modulus)^[Johnson 1985]. As the tangential tractions at the contact are equal and opposite for the two bodies, if they have similar elastic properties, displacements will be equal and opposite, i.e. the shape of the contact will still be determined, from hertzian theory, by

the normal force and the shape of the profiles. Such is the case in this work.

Stress distribution:

For simplicity the two dimensional case of cylindrical contact will be considered. The principle applies to spherical and elliptical contact. Let the bodies shown in Figure 3.11 represent a cylinder sliding on a plane perpendicular to its axis, where $2a$ is the width of the contact strip. The distribution of normal loading will then be herzian as shown in Figure 3.12a. If $q_0 = \mu p_0$ is the tangential traction at $x = 0$, it can be shown that^[Johnson, 1985] the direct stress in the y direction reaches a maximum of $-2q_0$ at the leading edge of the contact and a maximum tension of $2q_0$ at the trailing edge (Figure 3.12b). The combined effect of normal and tangential loading on direct stress in the y direction is shown in Figure 3.12c. Representative values are given for $Q/P = 1/3$.

The effect of tangential loading on the shape of the stress field are shown in Figures 3.13a and 3.13b for Tresca and Von Mises stress criterion, respectively. As the coefficient of traction increases, the maximum shear stress moves nearer the surface and increases in value. The start of plastic yielding will be determined by the maximum value of this stress. For simple shear, this will equal the yield stress " k " and, from the established yield criteria, the value of maximum contact pressure p_0 for first yield can be found. As the principal shear stress in the plane of deformation is $\tau_1 = (Q/P)p_0$, material throughout the width of the contact surface will begin to yield when,

$$k = (Q/P)p_0 \quad \text{or} \quad p_0/k = 1/(Q/P) \quad (3.34)$$

The contact pressures required for the onset of yield with tractive contact, as determined by both the Tresca and Von Mises yield criteria, are shown in Figure 3.14^[after Johnson, 1985]. This figure also shows the theoretical traction coefficient level at which the principal yield stress should reach the surface. Practical guides on how to estimate stresses with tangential loading are given in ESDU 84017^[1984] and ESDU 85007^[1985].

3.7 Contact effects during rolling.

Two non-conformal contacting bodies can move against each other in three ways: rolling, sliding and spinning. Even with just "free rolling", where the initial tractive force has been withdrawn, there is a resistance to the passing of the material through the contact, i.e. "rolling resistance". This is obviously also present with tractive rolling. The situation where there is a tractive force, but of insufficient magnitude to cause gross sliding, is very common; for example, driving road and rail wheels, tightened bolts and nuts, etc. Such contact situations can be complex, with areas of both "stick" and "slip" across the contact. With wheel/rail contact, there is rolling, sliding and spin with elements of micro-slippage, "creepage", generated in the longitudinal, transverse and spin directions. A static situation, where micro-slippage can occur, is found in the normal compressive contact of materials which have different elastic properties where different deformation strains have to be accommodated. Contacting materials with similar properties are assumed throughout the present work.

Rolling resistance.

The energy required to deform the material(s) constitutes a "rolling resistance" to the movement in both tractive and free rolling. Rolling will be resisted by the work of compression ahead of the contact and of tension behind the contact, with micro-slippage accommodating the elastic material movements as the surfaces pass through the contact. There will be energy dissipation due to elastic hysteresis loss, ϵ , as material is compressed through the contact.

If "F" is the force to overcome this rolling resistance, then a coefficient of rolling resistance can be termed λ , where $\lambda = F/P$, analogous to the friction coefficient, μ . For example, it can be shown for cylindrical contact that^[Halling, 1975],

$$\lambda = F/P = 2\epsilon a/3\mu R \quad (3.35)$$

where "a" is the half the contact width (Equ. 3.25) and "R" is the principle relative

radius of curvature (Section 3.3). Similarly, for spherical contact,

$$\lambda = F/P = 3\epsilon a/16R \quad (3.36)$$

where "a" is the contact radius (Equ. 3.20) and for elliptical contact (with a small rolling element moving over a large profiled surface),

$$\lambda = F/P = 3\epsilon a/16R' \quad (3.37)$$

Here "a" is the semi-axis in the direction of motion, as determined from Equations 3.14 and 3.17 / 3.18, and R' is the rolling radius of curvature of the rolling element.

Loads may be such that the sub-surface region of maximum stress exceeds the (compressive) elastic limit, consequently material will pass through an elastic - plastic - elastic cycle. With consequent work-hardening, rolling resistance is initially increased by this plastic work dissipation.

In every case it is apparent that rolling resistance is reduced by minimising the contact area; i.e. by the use of small contact radii and materials with high moduli of elasticity. Such is the case with counterformal bearing type contacts, such as wheel/rail, and the wear machine cylindrical disc contacts described in this work, particularly if these contacts are dry. Under these conditions, rolling resistance is relatively insignificant compared to frictional resistance.

Effect of static application of tangential force.

Where the overall level of friction is below the limiting level of friction (i.e. $Q < \mu P$), part of the contact "sticks" and part "slips". Thus the distribution of shear stresses shown in Figure 3.12b would be amended by an opposing distribution within the "stick" region, as shown in Figure 3.15 (tangential stress is "T" in this figure; normal stresses are not shown). Variations in the size of the "slip" regions, with oscillating or variable

loads, can result the wear and fatigue phenomena known as "fretting" within bolted assemblies, wire ropes, etc.^[Halling 1975].

Effect of tractive rolling; the concept of creepage.

Where there is a rolling movement, the distribution of shear stresses from tangential traction (cf. Figure 3.12b), shown in Figure 3.15e, is amended; the area of "stick" moves to the leading edge of the contact. The effect on the distribution of tangential force and strain for the two bodies is shown in Figure 3.16. For tractive rolling, the "stick" zone has to be situated at the leading edge of the contact otherwise a region of "slip" at the leading edge would have traction in the same direction as slippage, as shown in the "not possible" part of Figure 3.16. (Note, Figure 3.16 shows tangential stress only, not combined normal and tangential stress as in Figure 3.12c.)

The concept that there must be some micro-slippage between the contact of rolling objects, where there is traction, was first described in the last century with studies of rubber rolling on steel^[Johnson, 1985]. The conditions within the contact patch for the hertzian contact of steel bodies, using wheel-rail contact as an example, were first addressed by Carter^[1926]. He showed that zones of stick and slip must exist within the contact patch as the velocity of the driving roller must be fractionally faster than that of the driven roller in order to accommodate the tangential strains of the materials across the contact. The difference in tangential strains within the "stick" part of the contact must be accommodated by micro-slippage or "creep" in the balance of the contact area. The "creep ratio" has been defined as the distance difference between one revolution of a driving, or driven, roller and that of the unstrained circumference (positive for the driving roller, negative for the driven roller). "Creepage" is defined as the difference in the circumferential speeds, divided by the mean of the two speeds.

For "steady state rolling" (i.e. uniform motion under constant forces) it can be shown^[Johnson 1985] that for the general case of elliptical contact:

$$\dot{s}_x/V = \xi_x - \psi_y/c + (\partial \bar{u}_{x1}/\partial x - \partial \bar{u}_{x2}/\partial x) \quad (3.38)$$

$$\dot{s}_y/V = \xi_y - \psi_x/c + (\partial \bar{u}_{y1}/\partial x - \partial \bar{u}_{y2}/\partial x) \quad (3.39)$$

where:

- * \dot{s}_x and \dot{s}_y are the velocities of micro-slip between contacting points in steady rolling.
- * V is the rolling velocity along the x axis.
- * $\xi_x = (\delta V_{x1} - \delta V_{x2})/V$ and $\xi_y = (\delta V_{y1} - \delta V_{y2})$ are the longitudinal and transverse "creep ratios" where the " δV " expressions are the "creep velocities".
- * $\psi \equiv (\omega_{z1} - \omega_{z2})c/V$ is the non-dimensional "spin parameter or spin creepage", where ω_{z1} and ω_{z2} are the angular velocities due to spin.
- * $c = \sqrt{ab}$ where a and b are the contact ellipse semi-axes.
- * \bar{u}_x and \bar{u}_y are the components of tangential elastic displacement at a surface point.

It can be shown from the boundary conditions of these equations^[Johnson 1985, Halling 1975] that the "stick" zone must be situated at the leading edge of the contact.

Carter^[1926] observed that, although new wheel/rail profiles contacted over an elliptical area, once worn, the shape of the contact tended toward rectangular and he therefore addressed creepage two dimensionally, as for cylindrical contact. For the contact of similar elastic cylinders, using the hertzian determination for maximum contact stress, p_0 , it can be shown^[Carter, 1926; Johnson, 1985; Halling, 1975] that the longitudinal creep ratio is given by,

$$\xi_x = -\mu a/R \cdot (1 - \sqrt{1 - Q_x/\mu P}) \quad (3.40a)$$

and in non-dimensional terms,

$$\xi_x(R/\mu a) = -(1 - \sqrt{1 - Q_x/\mu P}) \quad (3.40b)$$

This relationship is known as a "creep curve" and it is shown in Figure 3.17. With

reference to this figure, and the "possible" traction curve shown in Figure 3.16, if any tangential force is applied to a contact ($Q/P < \mu$) there will be an element of micro-slip at the contact trailing edge to accommodate the elastic strains of the stressed material surfaces. As the tractive force is increased, the proportional area of "stick" will decrease and that of microslip will increase until it reaches the leading edge, at which point $Q/P = \mu$ and there is complete sliding over all the contact area (the "sliding" part of the curve shown in Figure 3.17). With the initial condition of tractive rolling with high friction (i.e. where $Q/P \ll \mu$), the slip region at the trailing edge all but disappears and the distribution of tangential traction approaches the limiting form known as the "creep coefficient", given by the dashed line in Figure 3.17. Here the creep ratio is given by,

$$\xi_x = aQ_x/2RP \quad (3.40c)$$

The situation for the contact of three dimensional bodies with longitudinal, transverse and spin creepage is complex. Johnson^[1958a & b] has shown that pure spin introduces an additional element of transverse creepage due to the tangential elastic compliance of the surface. For the situation where tangential tractions approach their limits, i.e. the slip zone approaches zero, Kalker^[1968] further developed Johnson's work into the "linear creep theory" which could be summarised into three linear creep equations,

$$Q_x/Gab = C_{11}\xi_x \quad (3.41a)$$

$$Q_y/Gab = C_{22}\xi_y + C_{23}\psi \quad (3.41b)$$

$$M_z/G(ab)^{3/2} = C_{32}\xi_y + C_{33}\psi \quad (3.41c)$$

where M_z is the spin traction and C_{11} , C_{22} , C_{23} , C_{32} and C_{33} are non-dimensional creep coefficients found from theory. Note that the creepage coefficients rely solely on material properties and contact areas.

For the opposite situation (with all three forms of creepage), where there is low friction

such that the situation approaches complete slippage, and elastic displacement due to tangential forces can be ignored, creepage relationships can be described by,

$$x_p/a = -\xi_y/\psi \quad \text{and} \quad y_p/a = \xi_x/\psi \quad (3.42)$$

For this case, there is a point $P(x_p, y_p)$, which does not necessarily have to be in the contact, where there is no slip (see Figure 3.18). At any point $A(x, y)$ the resultant tangential traction $q(x, y)$ has the magnitude $\mu p(x, y)$ in a direction perpendicular to the line PA.

The more realistic situation with partial slippage lies between the determinations described above and is more complex. One form of analysis has been the "strip theory". This was suggested by Haines and Ollerton^[1963] for solely longitudinal creepage and then developed by Kalker^[1967a] for all forms of creepage. It establishes conditions for cylindrical (rectangular) contact and then this is transposed to elliptical and circular contact by viewing those contact areas as a series of thin independent rectangular strips. Results are summarised in Kalker^[1969]. However it had severe limitations^[Kalker 1979]. Based on this theory, Figure 3.19 shows how different combinations of creepage affect the distribution of stick and slip areas.

Figure 3.20 shows the wheel on rail situation^[after Johnson, 1985]. Since wheels are coned, longitudinal creepage can arise (for a rigid wheelset) when the two wheels run at different radii, in addition to that generated by driving traction. Transverse creepage can be generated when the wheels are yawed at an angle ϕ to the running direction (as in curving) and spin creepage must be generated as the rail vertical axis is at an angle (λ) to the wheelset axis; the wheel has an angular velocity of spin ($\omega_z = \omega \sin \lambda$) relative to the rail.

In his examinations of wheel-rail contact, Kalker^[1979, 1990] differentiates between the creepages due to the accommodation of material strain ("elastic slip") and that due to

driving/braking traction, when viewing the wheel and rail as rigid non-elastic bodies with motion in their common tangent plane. The latter is termed "rigid slip" and is defined by,

$$\xi_r V = (\xi_x - \psi_y, \xi_y + \psi_x) V \quad (3.43)$$

where ξ_r is termed "relative rigid slip", V is the rolling velocity $\{V = (V_1 + V_2)/2\}$, cf. Figure 3.20} and spin creepage ψ is defined, for this situation, by,

$$\psi = -V_\phi/V + (\sin\lambda)/R, \text{ i.e.,}$$

$$\psi = (\text{spin velocity due to yaw}) + (\text{spin velocity due to wheel conicity and/or camber})$$

Typical values for straight track rolling are $\xi_x, \xi_y \approx 0.5\%$ and $\psi \approx 0.01\%$ [Kalker, 1990].

Rigid slip of a driving wheel is usually around 0.1% of the vehicle velocity [Kalker 1979].

"True slip" consists of "rigid slip" plus "an accommodation of material elasticity with respect to time".

Kalker also notes that his contact theories are "quasi-static" as inertial effects, requiring a "dynamic" theory, would only become significant at speeds around 500 km/hr. In his 1990 paper, he lists the six theories in addition to Hertz, which are still of significance to the European railway industry and he discusses their applicability, versus computing time, for different contact situation, together with their inter-relationships.

Under actual track conditions creep coefficients are often observed to be less than predicted due to the influence of surface (semi-lubricating) contaminants, surface roughness and vibration [Johnson, 1985].

Some laboratory testing of wheel-rail contact has involved full size contact tests [McEwen and Harvey, 1983] or scaled down contacts [Kalousek, 1982]. All three modes of creepage would be present and their distribution would be affected by wear of the wheel and rail profiles.

For the present work, it was felt that material behaviour could be characterised simply by testing cylindrical discs, i.e. with solely longitudinal creepage. Discs were made equal in width to avoid the end effects described by Johnson^[1982] (Figure 3.21) and of sufficient width to give an even axial distribution of contact stress over most of the contact.

Creepage considerations for such a contact have been described in the beginning of this section. On passenger rail networks such as British Rail, during heavy steady state rail curving without wheel flange contact, the overall creepage will be around 3%, most of which consists of transverse creepage. Rail wear and fatigue problems are concentrated at such areas, with failure reflecting the strain generated by transverse creepage. The situation thus differs from twin disc laboratory tests in that the major creepage component is not in the direction of rolling.

It should be noted that occasionally tractive contacts can change from "steady state" contact conditions (with stick and slip zones in the contact) to evolving zones of predominant stick and predominant slip, thus giving surface undulations ("corrugations and facets"). Such a phenomenon has been observed on road and rail and they were encountered during some of the tests described in this work (cf. Chapters 6 and 7).

3.8 Plastic deformation with the contact of solid bodies.

"Rail wear", as shown in Figure 1.1, includes not only wear by material removal as debris, but also gradual plastic flow of material out of the contact zone to form a lip on the flange. Such plastic flow would have been progressive over many rolling contact cycles. Plastic deformation within the wheel/rail contact zones will "feed" the wear mechanisms where particles are removed as debris.

Contact mechanicians view the plastic behaviour of material (i.e. "inelastic contact") within certain behavioural modes. Initially, yielding within complex stress systems where there is no tangential force, just normal contact, will be considered.

Yield with normal contact.

In uniaxial tension the elastic behaviour of metals either ceases at a specific yield point, or else the elastic-plastic transition is less obvious and a 0.2% proof stress is taken as the transition point. With complex stress systems, the state of stress can always be described in terms of the three principal stresses, σ_1 , σ_2 and σ_3 . Yield criteria then define conditions where plastic behaviour is initiated for any combination of these stresses; only sudden initiation of yield is considered. Two criteria are commonly used, Tresca and von Mises. The descriptions here are those given by Parkins^[1968]. Johnson^[1985] mentions a third, the "maximum reduced stress criterion".

The Tresca criterion states that plastic behaviour occurs when the maximum shear stresses reach a critical value. The shear stresses were given in equation 3.28. If $\sigma_1 > \sigma_2 > \sigma_3$ then the maximum shear stress must be $(\sigma_1 - \sigma_3)/2$. For uniaxial tension, $\sigma_1 = Y$, the yield stress in simple tension, and $\sigma_2 = \sigma_3 = 0$, therefore the maximum shear stress, $\sigma_1/2 = Y/2$. Thus the Tresca criterion is $(\sigma_1 - \sigma_3) = Y$. For pure shear, where k is the yield stress in simple shear, $\sigma_2 = 0$ and $\sigma_1 = -\sigma_3 = k$. Therefore,

$$k = Y/2 = 0.50Y \quad (3.44)$$

and the Tresca criterion can thus be written,

$$\{|\sigma_1 - \sigma_2|, |\sigma_2 - \sigma_3|, |\sigma_3 - \sigma_1|\}_{\max} = 2k = Y \quad (3.45)$$

Von Mises' criterion takes into account the fact that metal cannot be compressed in all directions and therefore the hydrostatic component "h", {where $h = (\sigma_1 + \sigma_2 + \sigma_3)/3$ }, must be considered. Thus the effective principal stresses are $(\sigma_1 - h)$, $(\sigma_2 - h)$ and $(\sigma_3 - h)$. The criterion states that deformation occurs when these reach a critical value, i.e.,

$$(\sigma_1 - h)^2 + (\sigma_2 - h)^2 + (\sigma_3 - h)^2 = [(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2]/3 = \text{a constant.}$$

Again, considering the case of uniaxial tension, where $\sigma_1 = Y$ and $\sigma_2 = \sigma_3 = 0$, from this the constant $= \frac{2}{3}Y^2$, therefore now,

$$(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2 = 2Y^2 \quad (3.46)$$

From the case of pure shear ($\sigma_2 = 0$, $\sigma_1 = -\sigma_3 = k$) this gives under von Mises,

$$k = Y/\sqrt{3} = 0.58Y \quad (3.47)$$

This differs slightly in value from Tresca (Eq. 3.44). Although refined experiments with ductile metals tend to support von Mises, as the difference is small and most metals are not fully isotropic, Tresca is usually used for algebraic simplicity^[Johnson, 1985].

The "maximum reduced stress criterion" also considers the hydrostatic component, $\{|\sigma_1-h|, |\sigma_2-h|, |\sigma_3-h|\}_{\max} = k = \frac{2}{3}Y = 0.67Y$ (3.48) Johnson^[1985] states that, for a stable plastic material, this criterion and Tresca's provide limits between which an acceptable yield criterion must lie. He has examined the relationship of these criteria with respect to the normal contact of solids.

Cylindrical contact (two dimensional).

Here yield by Tresca is determined by the maximum shear stress, τ_{\max} , in just the x-z plane. Contours of $\tau_1 = (\sigma_z - \sigma_x)/2$ are shown in Figure 3.13a where $\tau_{\max} = 0.30p_o$ at $z = 0.78a$ sub-surface. Substituting *Tresca* gives,

$$0.60p_o = 2k = Y \quad (3.49)$$

therefore yield begins at 0.78a below the surface when p_o reaches the value,

$$(p_o)_Y = 4p_m/\pi = 3.3k = 1.67Y \quad (3.50)$$

The other criteria both involve the third principal stress and take account of Poisson's Ratio ν . If $\nu = 0.3$ then, by substituting into the *von Mises* equation (3.46) it can be shown that yield begins at 0.70a below the surface when,

$$(p_o)_Y = 3.1k = 1.79Y \quad (3.51)$$

and by the *reduced stress criterion*,

$$(p_o)_Y = 2.7k = 1.80Y \quad (3.52)$$

Therefore the value of $(p_o)_Y$ is not greatly influenced by the stress criterion used. Maximum shear stress profiles for normal contact, and for increasing traction, as determined by Tresca and von Mises' criteria, are shown in Figures 3.13a & b, respectively. The load for initial yield can be obtained by substituting $(p_o)_Y$ in equation 3.22 or equation 3.24.

Substituting into von Mises equation (3.51) the 0.2% proof stress values for the wear

test materials described in this work (given in Table 2.3) gives,

		----- Pearlite -----		----- Bainite -----		
(MPa)		<u>R52</u>	<u>W64</u>	<u>B04</u>	<u>B20</u>	<u>B52</u>
0.2% PS	:	443	367	638	750	853
$(p_o)_Y$:	793	646	1142	1343	1527

The value of $(p_o)_Y$ would be lower with the influence of surface friction, however these figures do not account for the work-hardening of the materials nor for the directionality of their work-hardening with respect to the predominant strain directions.

Spherical contact.

With spherical contact, Johnson^[1985] similarly shows that, for $\nu = 0.3$, the maximum shear stress at $0.48a$ below the surface is $0.62p_o$. Thus,

$$\text{by Tresca,} \quad (p_o)_Y = 3(p_m)_Y/2 = 3.2k = 1.60Y \quad (3.53a)$$

$$\text{by von Mises,} \quad (p_o)_Y = 3(p_m)_Y/2 = 2.8k = 1.60Y \quad (3.53b)$$

The requisite loads can be calculated from equation 3.19.

Elliptical contact.

As shown in the table at the end of Section 3.4, as the ratio of the semi-axes, b/a , moves from zero (cylindrical) to unity (spherical) the maximum yield stress only slightly increases as b/a approaches 0.5. Therefore $(p_o)_Y$ will not vary greatly from the values given in Equations 3.50 and 3.53. Of more significance is the movement of the position of maximum yield toward from 0.78 to $0.48(z/b)$ as the ellipse becomes less eccentric.

From the typical p_o values generated by new rail profiles (Figure 3.10) it can be expected that a pearlitic rail will begin to yield under the higher contact stresses induced by wheelset transverse movements during curving. As shown in Figure 3.13, the addition of tangential forces will increase this tendency.

Yield with rolling contact and tractive force.

Just as hertzian analysis is valid for free rolling as well as normal static contact, so can the onset of yield with free rolling be considered in the manner described above.

With tractive rolling, the magnitude of the traction coefficient influences the point of first yield (Figure 3.13). When there is not complete sliding (i.e. $Q < \mu P$), the contact has zones of stick and slip. Figure 3.22^[after Johnson 1985] shows the stress distribution across a cylindrical contact which has a fixed tangential tractive force ($Q = 0.2P$). As increasing values of kinetic friction coefficient are applied to the contact, the tangential surface shear stress (i.e. traction) increases and the size of the microslip zone decreases (Figure 3.22a). The shape of the tangential surface traction curve becomes significant when, as the friction coefficient is increased, shear stress increases in relation to sub-surface shear stresses, i.e. the principal shear stress moves from sub-surface to surface. Thus the location of first yield will be at the surface stress peak, i.e. at the boundary of the stick and slip zones (Figure 3.22b).

For normal contact and free rolling, p_0/k values for first yield were given earlier in this section. For tractive rolling, the relationship is shown in Figure 3.22c. As the friction coefficient increases, the value of p_0 required for first yield diminishes.

Repeated rolling - work hardening and shakedown.

Material response to repeated rolling contact loads has been categorised into the four modes of behaviour shown in Figure 3.23^[after Johnson 1989]:

- * *Perfectly elastic* - where initial yield is not exceeded.
- * *"Elastic shakedown"* - where the first load cycle exceeds the elastic limit but, due to changes caused by the resultant plastic flow, the steady cyclic state lies within the altered elastic limit.
- * *"Plastic shakedown" ("cyclic plasticity")* - where the steady state eventually attains a closed loop of plastic stress-strain with no net accumulation of uni-directional plastic strain.

- * "*Ratchetting*" ("*incremental collapse*") - where, at steady state, there is an open cycle of plastic stress and strain such that an increment of uni-directional strain is generated with each cycle.

Three partially inter-linked mechanisms cause material to "shakedown":

- * Residual stresses generated by the initial cycles of plastic deformation. These are generally protective with hertzian type contacts, i.e. they act against further deformation.
- * Strain (or work) hardening of the material during the early loading cycles increasing the yield strength and thus reducing the degree of plastic deformation.
- * Surface movement of material due to the initial plastic deformation resulting in a more conformal contact and thus a lowering of contact stresses.

The mechanism of shakedown can be examined with reference to Figure 3.24a & b, which show the stress distribution and the effect of rolling, respectively. With reference to Figure 3.24b, first yield will occur at element "C" by shear on planes at 45° to the axes. Lateral expansion of the element will be stopped by the development of residual stresses parallel to the surface, hence yield will cease. The alternating "orthogonal" shear, τ_{zx} , of the elements at "B" and "D" cannot be reduced by the introduction of residual shear stress $(\tau_{zx})_r$ and hence "orthogonal" shear at these locations controls the shakedown limit and it is here that *repeated* plastic deformation will first occur when the shakedown limit is exceeded^[Johnson and Jeffries, 1963]. Variations in orthogonal shear stress distribution and magnitude, with increasing traction coefficient, are shown in Figure 3.24c. Around the levels found in the present work, maximum stresses would be at the surface.

For the initial modelling of plastic contact, it is convenient to consider the material behaviour as elastic and then perfectly plastic, and to consider plane strain conditions, i.e. line (cylindrical) contact.

* "Free rolling" - cylindrical line contact.

It can be shown^[Johnson, 1985] that residual stresses will lie parallel to the surface and vary with depth only. They will be compressive and restrict plastic deformation. Following Melan's shakedown theorem for elastic-plastic materials^[Melan, 1938], and based on both Tresca and von Mises yield criteria, Johnson shows that the limiting condition for shakedown is when the value of the shear stress, τ_{zx} , reaches k (the yield stress in simple shear). This maximum value, which is found to be $0.25p_o$ (by von Mises), is located at $\pm 0.87a$ on the x axis and $0.5a$ on the z axis (where $2a$ is the contact width).

$$\text{i.e. for shakedown, } p_o \leq 4k \quad (3.54)$$

The requisite residual stresses at $0.5a$ for shakedown are,

$$(\sigma_x)_r = -0.134p_o \text{ and } (\sigma_y)_r = -0.213p_o \quad (3.55)$$

From Equ. 3.51, p_o for first yield is $3.1k$, therefore the ratio of loads required for shakedown (P_s) and first yield (P_y) is,

$$P_s/P_y = (p_o \text{ for shakedown})^2 / (p_o \text{ for yield})^2 = 1.66 \quad (3.56)$$

i.e. the load must be 66% higher than the yield load to give continuous deformation with cyclic loading. The modified stress state with residual stresses is shown in Figure 3.25.

* "Tractive rolling" with cylindrical line contact.

For tractive rolling the analysis is similar. Although there is an asymmetric cycle of shear stress, material response, as shown in Figure 3.26, is similar to that shown in Figure 3.23. The shakedown limit is still determined by the maximum value of the shear stress, τ_{zx} ^[Johnson 1985]. This lies below the surface for $Q/P \leq 0.367$ and at the surface above this. For tractive rolling ($Q/P < \mu$) with line contact, the shakedown limits for a

perfectly plastic material (from Melan's theory^[1938]) and those for a kinematically hardening material (from Ponter's theory^[1976]) are shown in Figure 3.27.

** "Free rolling" - three dimensional contact.*

For elliptical and spherical contact, all components of residual stress have to be considered, plus the reduction in contact stress due to plastic deformation. If the latter is ignored, for spherical contact (eg, a ball rolling on a half space) it can be shown^[Johnson, 1985] that shakedown is dependent on the maximum residual shear stress, $(\tau_{zx})_r$ and that this is $0.21p_o$. Therefore for shakedown,

$$p_o \leq 4.7k \quad (3.57)$$

From Equ. 3.53, p_o for first yield (by von Mises) is $2.8k$, therefore the ratio of the requisite loads for shakedown and first yield is,

$$P_s/P_y = (4.7/2.8)^3 = 4.7 \quad (3.58)$$

A significant difference from the line contact case given in Equ. 3.56. Plastic deformation will further increase the requisite load for shakedown with 3 dimensional contact, however this is affected by strain hardening as shown below.

** "Tractive rolling" - three dimensional contact.*

This situation has been addressed by Ponter et al^[1985]. They produced a shakedown map for spherical contact of elastic - perfectly plastic materials and, for elliptical contact, the variations of shakedown limits, etc., for changes in the ellipse shape with two given traction coefficients (Figure 3.28).

** Strain hardening*

Steels, as with most metals, are not perfectly plastic and they strain harden under cyclic loading; i.e. the value of "k" will increase. Hardening can be expressed in three models

(Figure 3.26). With isotropic hardening, material will shakedown to the elastic state in one quarter cycle. Ponter^[1976] extended his shakedown theory to incorporate "kinematic" hardening; here, material will shakedown under a closed cycle of plastic strain under all conditions and a linear hardening law applies, where the hardening rate is constant and independent of the direction of strain or mean stress^[Bower and Johnson, 1990]. Such is the case where the zone of plasticity is sub-surface and surrounded by an elastic matrix. For such materials in two dimensional contact, Equ. 3.54 remains valid, whereas Equ. 3.57 for the three dimensional case does not.

This can be viewed with reference to Figure 3.27, the shakedown map for tractive rolling with cylindrical contact. Kinematic hardening raises the shakedown limit in the area where yield is sub-surface; this closed loop of sub-surface plasticity results finally in crack initiation by low cycle fatigue, whereas for a perfectly plastic material or non-linear kinematic hardening, shear strain accumulation often leads to ductile crack initiation^[Johnson 1988].

"Ratchetting" - cumulative plastic flow and incremental collapse.

No incremental, directional plastic flow can occur with kinematic hardening and a linear hardening law. Material must either exhibit elastic - perfectly plastic behaviour or follow a "non-linear" hardening law, where the hardening rate can vary and is a function of mean stress during the cycle. Rail steel exhibits such non-linear behaviour and its ratchetting behaviour can be predicted^[Bower and Johnson, 1990].

Dependent upon whether the traction coefficient is (theoretically) below or above 0.367, yield will be initiated sub-surface or at the surface, respectively. (The case where loads are so high that sub-surface zones spread to the surface will not be considered.) Where ratchetting can occur at the surface, the reduction in surrounding material constraint results in comparatively higher strains.

In the sub-surface case, with reference to Figure 3.24b, once shakedown is exceeded

orthogonal plastic shear will occur at elements "B" and "D". Although elastic stresses and strains at these locations are equal, above shakedown, even with free rolling, there will be an increment of permanent shear at the "exit" location "D". The equilibrium of residual stresses is maintained^[Johnson 1985].

Bower^[1989] and Bower and Johnson^[1990] have modelled the ratchetting behaviour of hard drawn copper and standard pearlitic (BS11) rail steel, for both sub-surface^[1989] and surface cases^[1990], and they have supported the model experimentally. Non-proportional cycles, such as occur with rolling-sliding contact, were tested using solid cylindrical specimens in cyclic tension-compression tests, with non-zero mean stress. To represent sliding contact, specimens were subjected to a cycle of tension, torque and then compression. Ratchetting rates for copper (where the shakedown limit is not greatly exceeded), as a function of load factor (applied p_o/p_s for shakedown) and traction coefficient, are shown in Figure 3.29. The significant increases in strain per cycle when the strain field moves to the surface, and when the load factor is increased, are clearly evident.

Whereas copper ratchetted up (accumulated strain) at an even rate per cycle (following a cycle like that shown in Figure 3.23), with pearlitic rail steel the rate of strain decreases with continuous cycling and eventually the material settles to a closed cycle of strain^[Bower 1989]. Bower and Johnson's model^[1991] accounts for this material behavioural difference. Their comparison between the experimental response of pearlitic rail steel to their non-linear hardening model is shown in Figure 3.30. Here the loading cycle represents rolling and sliding contact; a similar correlation was obtained where the loading cycle represents sliding contact. Further load cycles on the strained pearlitic structure will eventually initiate fatigue failure of the strained element.

Shakedown maps determined from their calculations are shown in Figure 3.31 for conditions of full slip; in Figure 3.31a for line and point contacts (p_o/k against Q/P where $Q/P = \mu$) and in Figure 3.31b for conditions of partial slip with line contacts

(p_o/k against Q/P , where $Q/P < \mu$). In the latter figure, the reduction in the shakedown limit with an increase in the friction coefficient (μ) is shown. It can be seen from these figures that, for a fixed traction (Q/P), plastic flow will occur more readily for partial slip with a high friction coefficient than for full slip with a low friction coefficient. It should be noted that with conditions of full slip for line contact, shown in Figure 3.31a, at the shakedown limit, yield is reached across the entire contact width simultaneously.

On a rail, both longitudinal and lateral tractions are present and the continuous lateral movement of the wheelsets across the track prevents a hertzian contact groove forming, with the consequent contact stress reduction. Therefore, the four shakedown lines shown in Figure 3.31a can be taken as valid. As will be seen in Chapter 5, rail material is laterally strained out of the contact zone with rail-wheel contact and this plastic movement does gradually effect the rail profile, consequently contact stresses are reduced. Bhagava et al^[1988] have analysed rolling contact for rail steel and have derived values of cyclic strain for perfectly plastic hardening and kinematic hardening; the latter being an order of magnitude smaller. They mention that, for heavy haul rail, the p_o/k values for new profiles are around 6, well above the shakedown limit. This infers that such rails will rapidly "run-in" to a more conformal shape. Kapoor and Johnson^[1992] have studied progressive ratchetting of elliptical contacts and have noted that the contact deforms to a near-rectangular shape (discussed in Section 3.5).

No studies appear to have yet been made on the ratchetting response of bainitic and martensitic steels (for rail track usage).

It should be mentioned that some metals structures can kinematically soften, rather than harden, under cyclic stressing. For example, heavily cold worked steel where the high density of dislocations can gradually re-align ("untangle or self cancel") under repeated cyclic loading ^[Landgraf, 1969].

3.9 Surface roughness.

All considerations so far have made the assumption that contacting surfaces are perfectly smooth, however most engineering surfaces have both micro-roughness and some degree of waviness. This results in asperity tip contact under load and the contact mechanics situations described in this chapter are set up in a series of micro-scales. This situation can also arise from the entrapment of debris particulates in the contact. It is a common experience that rough surfaces have more resistance to rolling and sliding than smooth surfaces. Even where bulk stress levels are below the elastic limit, there will be plastic deformation of asperity tips; such deformation will progressively decrease as asperity contact areas increase with every loading cycle; i.e. as the surfaces wear-in or run-in.

The implications of surface roughness, particularly on lubricated contacts, has been extensively studied recently by workers at Imperial College (London, UK). Poon and Sayles^[1991] have classified contacts into three forms; plastic asperity contact, elastic asperity contact and near-smooth elastic contact. The effects (with lubrication) on friction, adhesion, wear and boundary lubrication have been examined. Bailey and Sayles^[1991] have calculated contact stress levels on a micro-scale from asperity contacts (Figure 3.32a). From their figure, it is clear that such contacts must plastically deform and that local high shear stresses are generated near the surface. It should also be noted that, once plastically deformed, the loaded asperities also elastically deform and the "loaded" surface roughness, of significance for lubricated contact, will be less than that measured by profilometry. Snidle and Evans^[1994] have recently produced similar results from a different approach (Figure 3.32b).

The as-ground disc surfaces of the rolling-sliding dry contact tests described in this work would have run-in, with respect to asperity flattening, within a few cycles. The decrease in friction with such running-in was observed on the (torque monitoring) chart recorder. However, accumulation of plastic strain (work-hardening) at the surface, and the establishment of steady state wear, took far longer. Rails wear-in to a "worn -profile"

relatively quickly (within a few months) particularly where there are high creepages as on curved track.

Another effect of surface roughness is to increase friction due to the energy absorbed by asperity impacts and the dynamic movements of contacting masses "riding over" surface irregularities^[Johnson, 1985]. This resistance force is related to rolling speed. As will be seen in Chapters 6 and 7, regular surface plastic deformations were slowly generated on some test disc surfaces, under certain test conditions, well after running-in. This created a form of "surface roughness" which generated its own friction oscillations.

3.10 Summary.

The highly loaded contact of solid non-conforming bodies, such as wheel/rail contact, has been examined. The formulae generated from the theories of Hertz^[1881] are still widely used in engineering and they give a good first approximation for wheel/rail contact forces. Where this contact is dry, the coefficient of traction normally exceeds 0.2 and is usually around 0.6, therefore stress maxima are usually at, or very near, the contact surfaces. The coefficients of traction, with the dry twin disc tests described in this work, ranged between 0.3 at high contact stresses to around 0.8 at low contact stresses, and maximum matrix deformation was observed to be at the surface. In later work on the LEROS machine, with the same type of test discs^[Garnham & Beynon, 1990, Beynon et al, 1994], the introduction of water as a lubricant (for rolling contact fatigue tests) reduced the traction coefficient to between 0.05 and 0.25 and zones of matrix plastic deformation were observed primarily sub-surface, as predicted by theory, and also to a lesser degree at the surface, due to surface roughness effects (see Section 3.9 above).

The wheel-rail contact area has been shown to be elliptical and the ellipse shape can greatly vary with wheelset movements across the track (Figure 3.10), as can arise during curving. For straight track, particularly when worn-in, contact can be approximated to cylindrical (cf. Section 3.5, Esvelde^[1989] and Section 3.8, Kapoor and Johnson^[1992]). Exact analyses of contact have been developed by Kalker in the form of

computer programs of varying complexity (including Hertz). He has recently comparatively reviewed these programs^[Kalker, 1990].

The additional effect of tangential loading, in combination with normal loading, has been shown not to affect normal compressive stress distributions and hence the shape of the contact area, however it does affect the location and magnitude of shear stresses. The maximum shear stress both increases and moves towards the surface with an increase in the traction/friction coefficient. If the tangential force is sufficient to cause complete sliding across the contact, there is *kinetic friction*, i.e. $Q = \mu P$. With a wheel tractively driving on a rail, there is obviously not complete sliding and here Q is variable and less than μP , a condition of *static friction*. With tractive rolling, there is small creepage of one surface relative to the other and areas of stick and slip are present in the contact, within which the local friction is static and kinetic, respectively. Where possible, these *local* variations in tangential force should be taken into account when considering the effects of tractive rolling on material behaviour, rather than the average tractive force over the complete contact.

It has been shown that with tractive driving of a railway wheelset, the geometry of the (conical) wheel on rail, together with wheelset yawing and transverse shift during curving, generate contact creepages in the longitudinal, transverse and spin directions. With heavy curving, there is significant wheelset shift which generates significant, extra transverse creepages, particularly on the high rail from the leading wheelset. This last movement contributes most to the composite creepage of around 3% (for the British rail system) which is generated during curving and thereby causes most wear. It is the resistance of rail steel to these wear forces which this work primarily addresses. Such multi-directional creepage strains are difficult to control and analyse with laboratory tests, therefore uni-directional creepage, via twin disc cylindrical contact, was used for this experimentation. This enabled the response of the various test materials to cyclic rolling and sliding stresses to be more simply examined. Additionally, these test discs could be simply directionally sectioned for micro-examination. One major difference

from the rail situation was that here the primary creepage was in the direction of rolling and across it. During the experimentation, test rails of similar steels were assessed on track by British Rail and similar results were obtained, as will be described.

It has been mentioned that descriptions of "rail wear" often encompass the composite loss of material from the loaded areas of the original rail cross-section; this consists of loss by plastic flow from the contact area, in addition to loss via wear debris. Such bulk plastic flow has not been considered in this work; the "mushrooming" of the cylindrical test discs under heavy loads did not significantly affect the *loaded* contact width. However the plastic flow at the contact surfaces, in the direction of creepage, was of great importance as it was the wear and surface fatigue of these strained surface layers that were actually being measured.

Plastic behaviour is governed by the yield stress in shear. The requisite values of maximum contact stress to give yield with normal contact, by the various yield criteria, have been shown to be similar. The movement of the location of maximum shear, from sub-surface to surface, with an increase in the ratio of tangential force over normal force (Q/P), has been shown. The concept of "shakedown" has been described, as has the concept of incremental plasticity beyond shakedown. The effect of the Q/P value on shakedown has been mapped, thus for a given condition, the presence, severity and location of plasticity can be predicted. For the dry wear tests described in this work, first yield was always at the disc surfaces, and under most conditions, incremental (cyclic) plasticity was present.

Finally, the differentiation has been mentioned, between bulk effects which determine the general Hertzian analysis and smaller scale surface effects due to surface roughness, which set up a series of smaller Hertzian contacts. This is more significant for the analysis of lubricated contacts; with these dry wear tests, surfaces had conformed (i.e. run-in) within a few cycles. However, during many of the dry wear tests, regular surface plastic deformations (corrugations) were generated and these would have

affected the stress distributions. Their effect is considered in later chapters.

3.11 References.

- Bailey, D.M. and Sayles, R.S (1991). "Effect of roughness and sliding friction on contact stresses." *Trans. ASME, J.Trib.*, **113**, pp. 729-738.
- Bhagava, V., Hain, G.T. and Rubin, C.A. (1988). "Analysis of rolling contact with kinematic hardening for rail steel properties." *Wear* **122**, pp. 267-283.
- Beynon, J.H., Garnham, J.E. and Sawley, K.J. (1994). "Rolling contact fatigue of four pearlitic rail steels." Initially submitted to *Wear*, August 1994. Re-submitted following review, June 1995.
- Böhm, K., Schlicht, H., Zwirlein, O. and Eberhard, R. (1974). "Non-metallic inclusions and rolling contact fatigue." Proc. Symp. on "Bearing Steels: The rating of nonmetallic inclusions", 22-24/5/74, Boston, Mass., USA. ASTM S.T.P. 575; ed. JCC Hoo.
- Bower, A.F. (1989). "Cyclic properties of hard drawn copper and rail steel." *J. Mech. Phys. Solids* **37** (4), pp. 455-470.
- Bower, A.F. and Johnson, K.L. (1990). "Plastic flow and shakedown of the rail surface in repeated rail-wheel contact." *Proc. 3rd. Int. Symp. on "Contact mechanics and wear in rail-wheel systems"*, 22-26/7/90, Univ. of Cambridge, (UK). Pub. *Wear* **144** (1991), pp. 1-18.
- Carter, F.W. (1926). "On the action of a locomotive driving wheel." *Proc. Royal Soc. A* **112**, pp.151-157.
- Cheesewright, P.R. (1981). "Wheel-rail contact pressure and geometry for various vehicle and tyre profiles." *British Rail Research Report TM MF 3* (File 263/384/3).
- Cole, S.J. (1994). Personal communication. Shear stress plots based on equations in Johnson^[1985].
- Duffek, W. (1982). "Contact geometry of wheel rail vehicles." *Proc. 1st. Int. Symp. on "Contact mechanics and wear of rail-wheel systems"*, Univ. of B.C., Canada, 6-9/7/82. Pub. Univ. of Waterloo Press (1983), eds. J Kalousek et al.

- Dyson, A., Evans, H.P. and Snidle, R.W. (1992). "A simple accurate method for calculation of stresses and deformation in elliptical hertzian contacts." Technical Note, *Proc. Inst. Mech, Engrs.; Part C - J. Mech. Eng. Science* 206, pp. 139-141.
- ESDU 78035 (1978). "Contact phenomena I: Stresses, deflections and contact dimensions for normally loaded unlubricated elastic components." *Engineering Sciences Data Unit*, London, UK.
- ESDU 84017 (1984). "Contact phenomena II: Stress fields and failure criteria in concentrated elastic contacts under combined normal and tangential loading." Ibid.
- ESDU 85007 (1985). "Contact phenomena III: Calculation of individual stress components in concentrated elastic contacts under combined normal and tangential loading." Ibid.
- Eisenmann, J. (1977). "Die Schiene als Träger und Fahrbahn." Chapter 2 of *Die Eisenbahnschiene*, ed. F. Fastenrath, Verlag W Ernst, e.a., Berlin.
- Esveld, C. (1989). *Modern Railway Track*, pub. MRT-Productions, Duisburg, Germany.
- Garnham, J.E. and Beynon, J.H. (1990). "The early detection of rolling-sliding contact fatigue cracks." *Proc. 3rd. Int. Symp. on "Contact mechanics and wear in rail-wheel systems"*, 22-26/7/90, Univ. of Cambridge, (UK). Pub. *Wear* 144 (1991), pp. 103-116.
- Haines, D.J. and Ollerton, E. (1963). "Contact stress distributions on elliptical contact surfaces subjected to radial and tangential forces." *Proc. Inst. Mech, Engrs.* 177, p. 95.
- Halling, J. (1975). Ed. *Principles of Tribology*, Macmillan Press, London.
- Hearle, A.D. (1984). PhD Thesis, Univ. of Cambridge (UK).

- Hertz, H. (1882). "Über die Berührung fester elastischer Körper" (On the contact of elastic solids). *J. reine und angewandte Mathematik* 92, pp. 156-171. (For english translation - Miscellaneous Papers by H. Hertz, eds. Jones and Schott, pub. Macmillan, London, 1896.)
- Hills D.A. and Sackfield, A. (1984). *J. Strain Anal. Eng. Des.* 19 (1), pp. 9-14.
- Johnson, K.L. (1958). "The effect of spin upon the rolling motion of an elastic sphere on a plane." *Trans. ASME, J. App. Mechanics* 25, p. 339.
- Johnson, K.L. and Jeffries, J.A. (1963). "Plastic flow and residual stress in rolling and sliding contact." *Proc. Symp. on 'Fatigue in rolling contact'*, 28/3/63, Inst. Mech. Engrs., London, pp. 55-65.
- Johnson, K.L. (1982). "One hundred years of Hertz." *Proc. Inst. Mech. Engrs. Tribology Group* 196, p. 39.
- Johnson, K.L. (1985). *Contact Mechanics*. Cambridge University Press, Cambridge, UK.
- Johnson, K.L. (1988). "The strength of surfaces in rolling contact." Proc. 3rd. BP lecture to Inst. Mech. Engrs. Trib. Group, London, 8/12/88. Pub. *Proc. Inst. Mech. Engrs.* 203-C3, pp. 151-163.
- Kalker, J.J. (1967a). "A strip theory for rolling with spin and slip." *Proc. Kon. Ned. Akad. van Wetenschappen* B70, p. 10.
- Kalker, J.J. (1967b). "On the rolling contact of two elastic bodies in the presence of dry friction." PhD Thesis, University of Delft.
- Kalker, J.J. (1968). "The tangential force transmitted by two elastic bodies rolling over each other with pure creepage." *Wear* 11, pp. 421-430.
- Kalker, J.J. (1969). "Transient phenomena in two elastically similar rolling cylinders in the presence of dry friction." Lab. v. Tech. Mech., Univ. of Delft, Report No. 11.
- Kalker, J.J. (1973). "Simplified theory of rolling contact." Univ. of Delft Progress Report 1, pp. 1-10.

- Kalker, J.J. (1979). "Survey of wheel-rail rolling contact theory." *Vehicle System Dynamics* 5, pp. 317-358.
- Kalker, J.J. (1982). "Two algorithms for the contact problem in elastostatics." *Proc. 1st. Int. Symp. on "Contact mechanics and wear of rail-wheel systems"*, Univ. of B.C., Canada, 6-9/7/82. Pub. Univ. of Waterloo Press (1983), eds. J Kalousek et al., pp. 102-120.
- Kalker, J.J. (1985). "On the contact problem in elastostatics." *Universal Problems in Structural Analysis. CISM Courses and Lectures No. 288*. Eds. G del Piero et al. Pub. Springer (Vienna, New York).
- Kalker, J.J. (1988). "Contact mechanics algorithms." *Comm. Appl. Num. Meth.* 4, pp. 25-32.
- Kalker, J.J. (1990). "Wheel-rail rolling contact theory." *Proc. 3rd. Int. Symp. on "Contact mechanics and wear in rail-wheel systems"*, 22-26/7/90, Univ. of Cambridge, (UK). Pub. *Wear* 144 (1991), pp. 243-262.
- Kapoor, A. and Johnson, K.L. (1992). "Effect of changes in contact geometry on shakedown of surfaces in rolling/sliding contact." *Int. J. Mech. Sci.* 34 (3), pp. 223-239.
- Landgraf, R.W. (1969). "The resistance of metals to cyclic deformation." *Proc. Symp. on "High fatigue resistance in metals"*, 22-27/6/69, Atlantic City, NJ (USA). Pub. ASTM STP467 (1970), pp. 3-36.
- Melan, E. (1938). *Sitzungsberichte der Ak. Wissenschaft., Wien, Ser. 2A* 147, p. 73.
- Pascal J.P. and Sauvage, G. (1993). "The available methods to calculate the wheel-rail forces in non-hertzian contact patches and rail damaging." *Vehicle System Dynamics* 22, pp. 263-275.
- Parkins, R.N. (1968). *Mechanical Treatment of Metals*. Pub. George, Allen & Unwin (London).
- Ponter, A.R.S. (1976). "A general shakedown theorem for inelastic materials." *Proc. 3rd. Int. Conf. on "Struct. Mech. in Reactor Tech."*, Section L, Imperial College, London.

- Ponter, A.R.S., Hearle, A.D. and Johnson, K.L. (1985). "Application of the kinematical shakedown theorem to rolling and sliding point contacts." *J. Mech. Phys. Solids* 33 (4), pp. 339-362.
- Poon, C.Y. and Sayles, R.S. (1991). "The classification of rough surface contacts in relation to tribology." *Proc. Int. Conf. on "Frontiers of Tribology"*, 15-17/4/91, Stratford upon Avon (UK). Pub. *J. Phys. Part D (Appl. Phys.)* 25 (1A), pp. A334-A339).
- Shen, Z.Y., Hendrick, J.K. and Elkins, J.A. (1984). "A comparison of alternate creep-force models for rail vehicle dynamic analysis." *Proc. 8th IAVSD Symp. on "The Dynamics of Vehicles"*, MIT, Cambridge (USA). Ed. J.K. Hendrick, pub. Swets & Zeitlinger, Lisse (The Netherlands).
- Snidle, R.W. and Evans, H.P. (1994). "A simple method of elastic contact simulation." *Proc. Instn. Mech. Engrs. (Part J: J Eng. Trib.)* 208 pp. 291-293.
- Timoshenko, S.P. and Goodier, J.N. (1951). *Theory of Elasticity*, 3rd Edition. Pub. McGraw Hill (N.York, London), pp. 403-420.
-

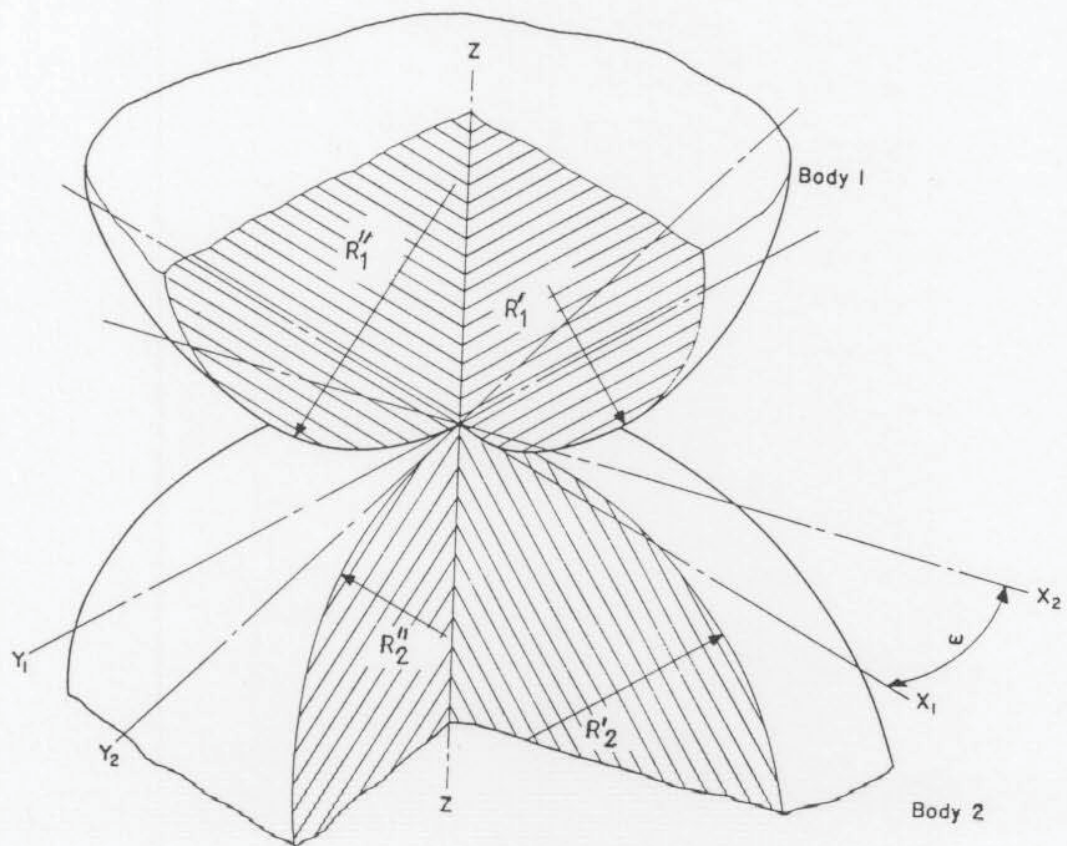


Figure 3.1 The contact of two non-conforming bodies. "Z" is the common normal. "X₁, X₂, Y₁ and Y₂" all lie in the tangent plane. R'₁ and R'₂ are the principal curvature radii of bodies 1 and 2, respectively, in one principal plane of curvature and R''₁ and R''₂ are the principal curvature radii in the other principal plane^[after ESDU 78035, 1978].


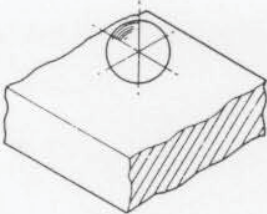
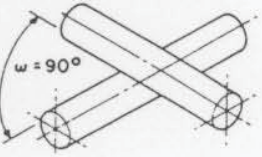
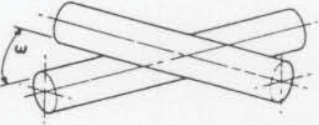
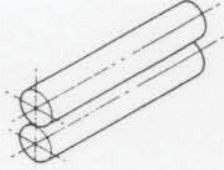
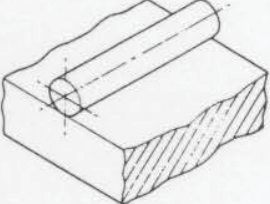
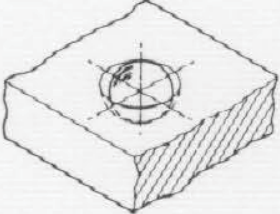
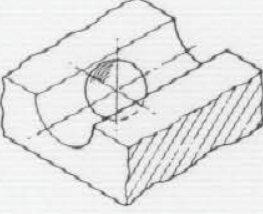
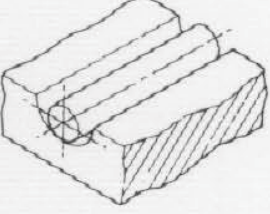
 <p>(a) Sphere on sphere (non-conformal)</p> $R'_1 = R''_1, R'_2 = R''_2$ <p>Contact shape is circular</p>	 <p>(b) Sphere on plate (non-conformal)</p> $R'_1 = R''_1, R'_2 = R''_2 = \infty$ <p>Contact shape is circular</p>	 <p>(c) Cylinder on cylinder (non-conformal)</p> $R''_1 = R''_2 = \infty$ <p>Contact shape is circular when $\omega = 90^\circ$ and cylinders are of equal diameter.</p>
 <p>(d) Cylinder on cylinder (non-conformal)</p> $R''_1 = R''_2 = \infty$ <p>Contact shape is elliptical when $0^\circ < \omega < 90^\circ$ and $R'_1 \neq R'_2$</p>	 <p>(e) Cylinder on cylinder (non-conformal)</p> $R''_1 = R''_2 = \infty$ <p>Contact shape is an ellipse of infinite length (i.e. rectangular) when $\omega = 0^\circ$</p>	 <p>(f) Cylinder on plate (non-conformal)</p> $R''_1 = R'_2 = R''_2 = \infty$ <p>Contact shape is an ellipse of infinite length (i.e. rectangular)</p>
 <p>(g) Sphere in socket (conformal)</p> $R'_1 = R''_1, R'_2 = R''_2 \text{ (negative)}$ <p>Contact shape is circular</p>	 <p>(h) Sphere in groove (conformal)</p> $R'_1 = R''_1, R'_2 = \infty, R''_2 \text{ is negative}$ <p>Contact shape is elliptical</p>	 <p>(i) Cylinder in groove (conformal)</p> $R''_1 = R''_2 = \infty, R'_2 \text{ is negative}$ <p>Contact shape is an ellipse of infinite length (i.e. rectangular)</p>

Figure 3.2 Various types of hertzian contact for solids of regular shape^[after ESDU 78035, 1978].

Only in sketches *c* and *d* are the principal planes of curvature not coincident. Subscripts for infinite radii are denoted as R''_1 , R''_2 in sketch *c* and R'_2 in sketch *h*; they have been chosen thus to ensure $B > A$ (cf. Equations 3.2 to 3.5).

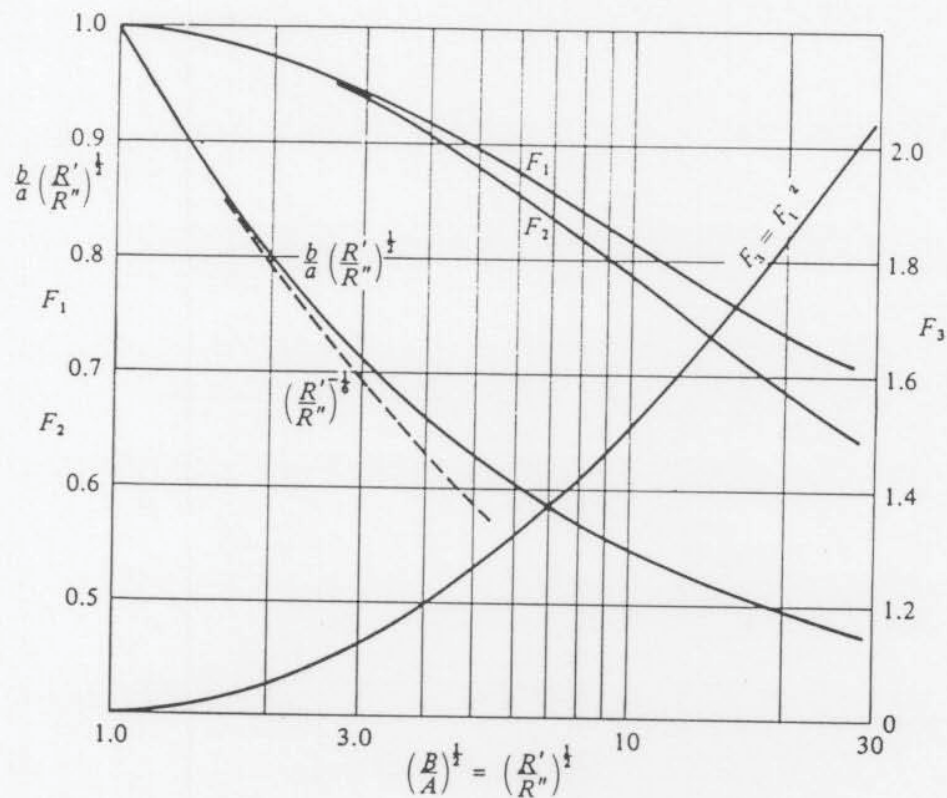
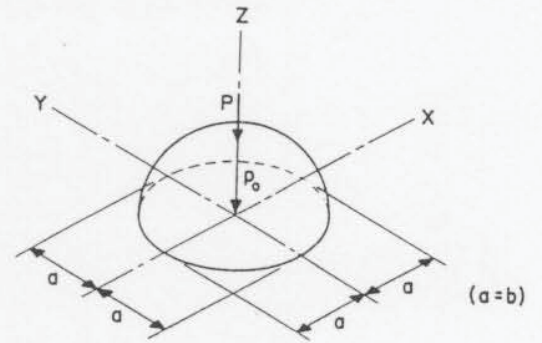


Figure 3.4 The compressive contact of bodies with general profiles^[from Johnson, 1985]. The shape of the contact ellipse b/a and the functions F_1 , F_2 and $F_3 (= F_1^{-2})$ in terms of the ratio of relative curvatures, R'/R'' . (For use in Equations 3.14 to 3.18.)

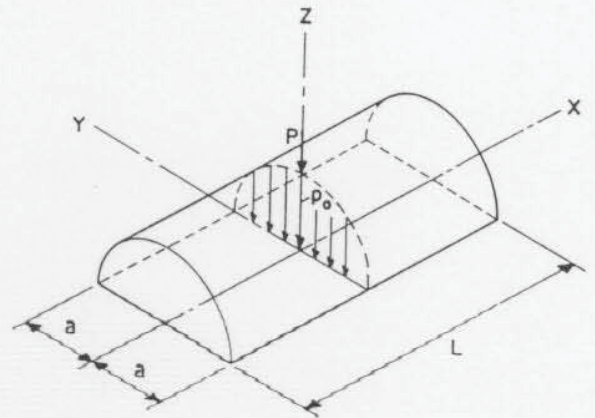
Circular contact

- Projected area : πa^2
 Average pressure : $P/\pi a^2$
 Maximum contact pressure : $p_o = 3P/2\pi a^2$
 Pressure distribution : $p_{(x,y)} = p_o \sqrt{1 - (x^2 + y^2)/a^2}$



Cylindrical (line) contact

- Projected area : $2aL$
 Average pressure : $P/2aL$
 Maximum contact pressure : $p_o = 2P/\pi aL$
 Pressure distribution : $p_{(y)} = p_o \sqrt{1 - (y^2/a^2)}$



Elliptical contact

- Projected area : πab
 Average pressure : $P/\pi ab$
 Maximum contact pressure : $p_o = 3P/2\pi ab$
 Pressure distribution : $p_{(x,y)} = p_o \sqrt{1 - (x^2/a^2) - (y^2/b^2)}$

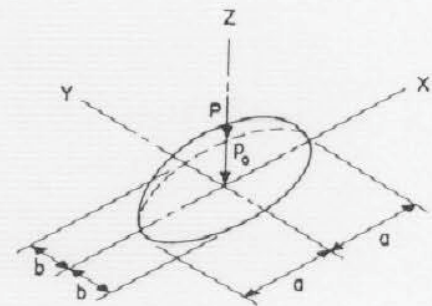


Figure 3.5 Types of hertzian contact^[after ESDU 78035, 1978]

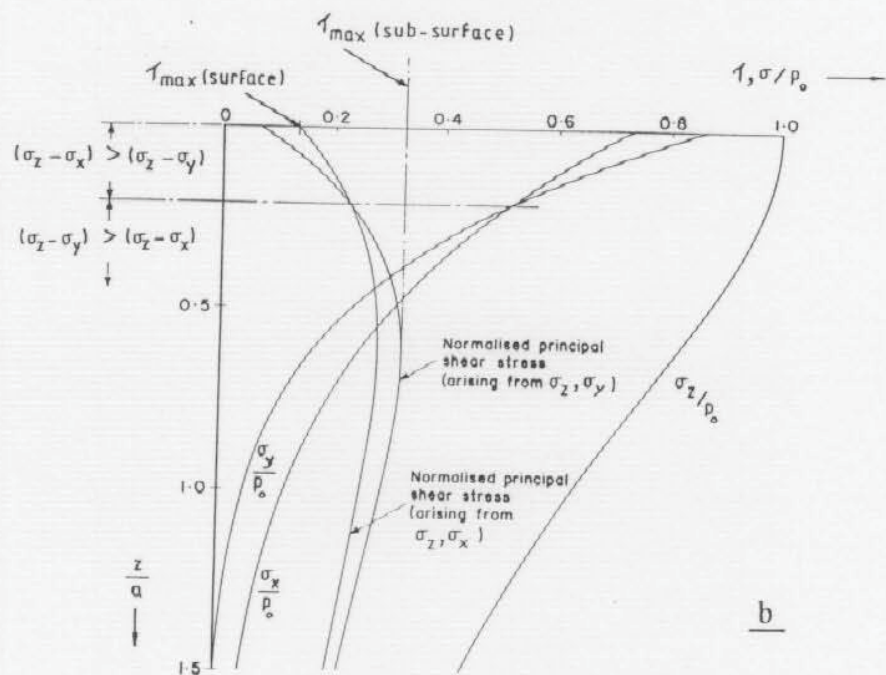
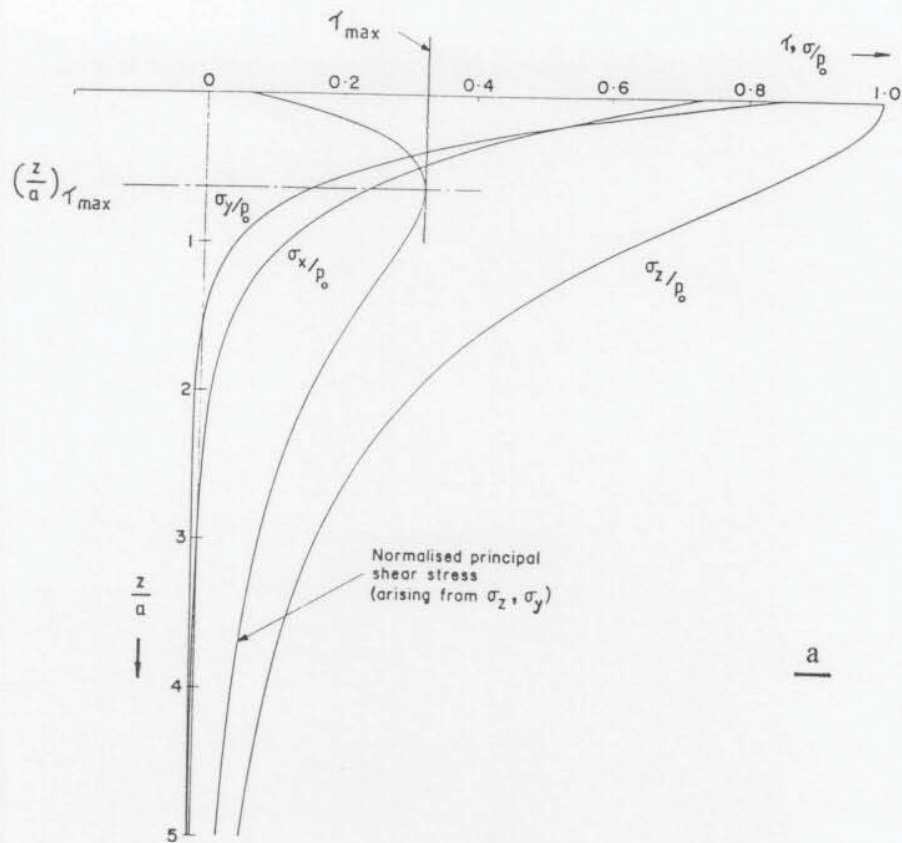


Figure 3.6 Principle sub-surface stresses along the axis of symmetry for elliptical hertzian contact^[after ESDU 78935, 1978]

(a) Overview.

(b) Detail of near surface stresses.

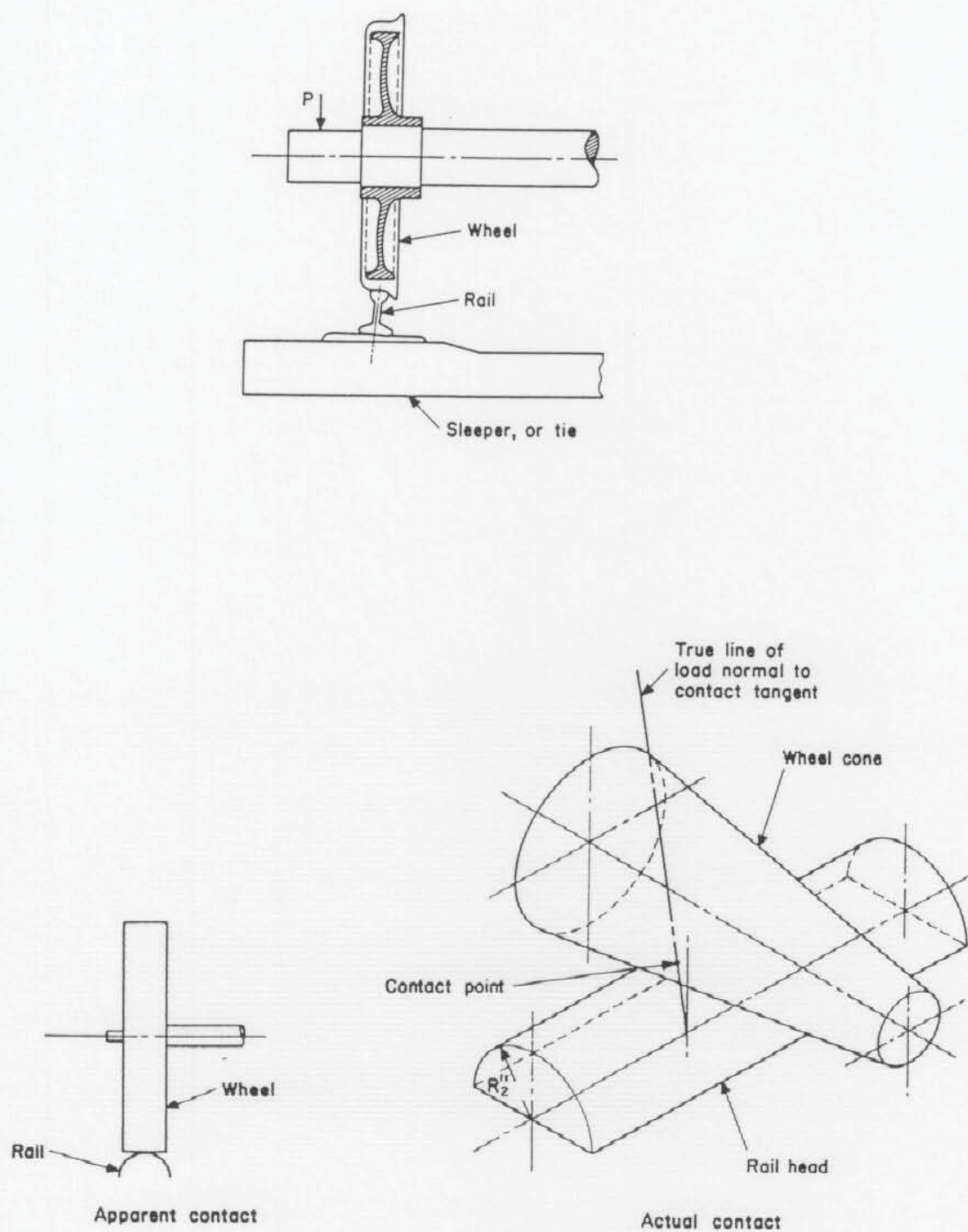
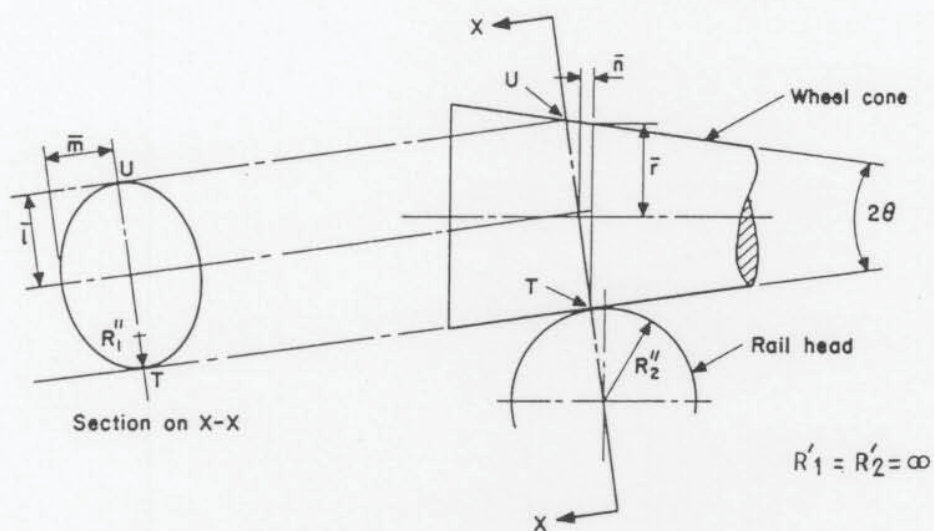


Figure 3.7 Hertzian wheel-rail contact; idealised cylindrical and conic contacting bodies^[after ESDU 78035, 1978].



From geometrical considerations it can be seen that

$$\bar{l} = \frac{r \sin (90 + \theta)}{\sin (90 - 2\theta)},$$

and $\bar{m} = \bar{r} + \bar{n} \tan \theta,$

where $\bar{n} = \bar{l} \sin \theta.$

Figure 3.8 Contact geometry of idealised hertzian wheel-rail contact.

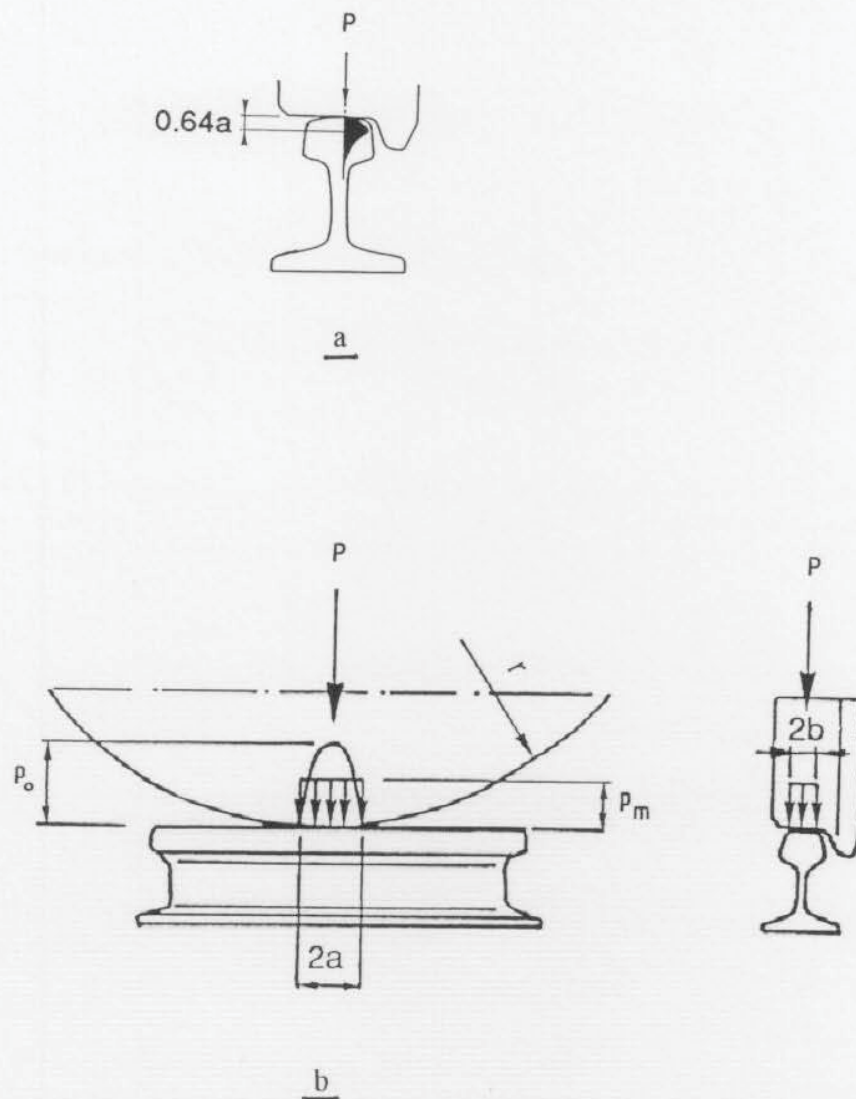


Figure 3.9 Simplified wheel-rail contact stresses^[after Esveld, 1989 and Eisenmann, 1977].

- (a) Shear stress distribution in the rail head at the contact centre.
- (b) Simplified wheel-rail contact distribution for straight track.

WHEEL LOAD = 10t
 SCALE = 2 × ACTUAL
 WHEEL DIAMETER = 853 mm
 FIGURES GIVE MAX. NORMAL CONTACT PRESSURE (MPa)

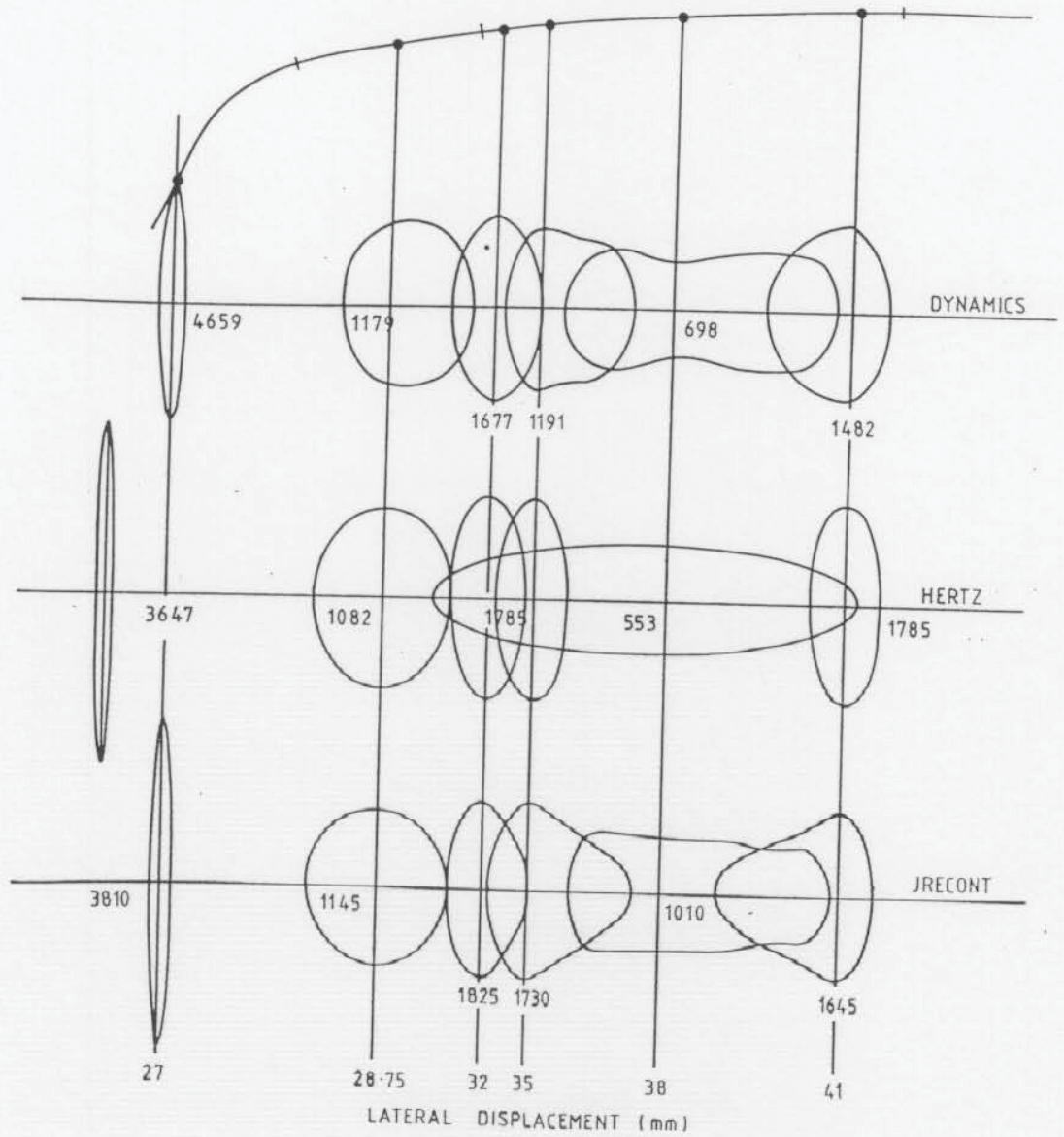


Figure 3.10 Wheel-rail contact areas, with respective values of maximum contact stress, calculated by three methods (including hertzian) for the steady state (non-dynamic) contact of one type of high speed locomotive wheel on a modern rail profile^[from Cheesewright, 1981].

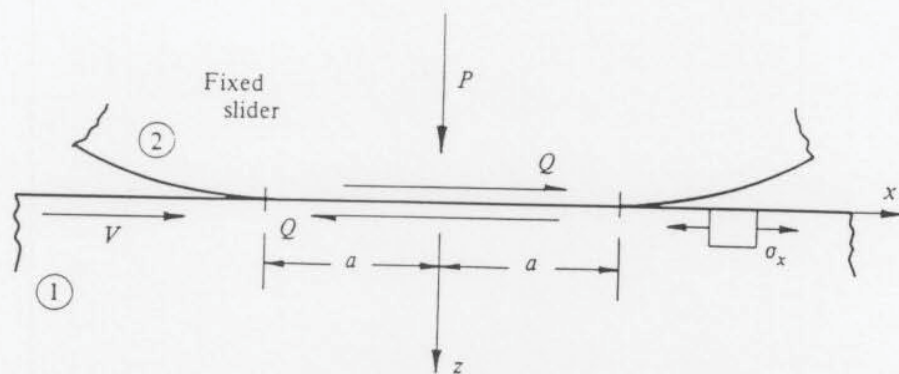
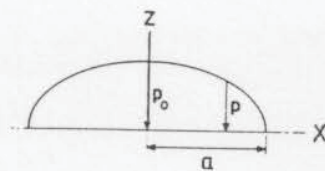
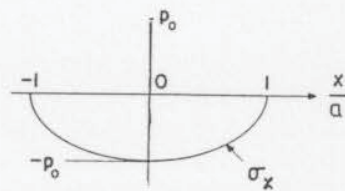


Figure 3.11 Tangential loading and sliding contact^[from Johnson, 1985].

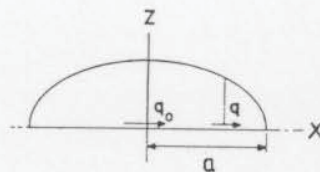


Elliptical distribution of normal surface loading intensity across contact width.

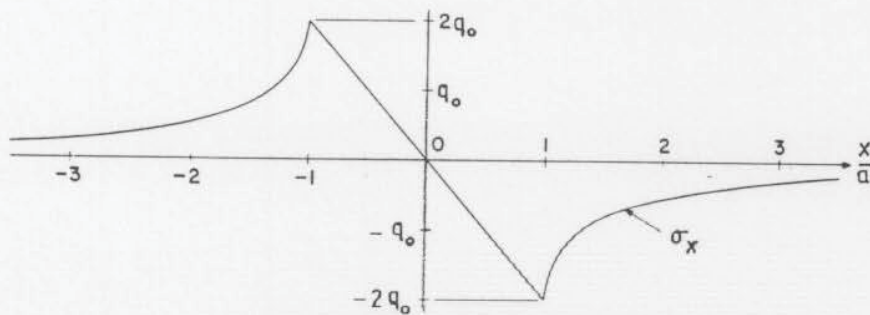


Distribution of direct stress at surface, in X-direction, resulting from purely normal load.

a. Normal loading

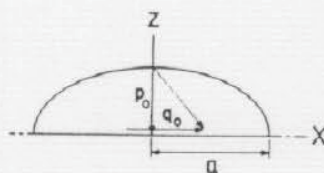


Elliptical distribution of tangential surface loading intensity across contact width.

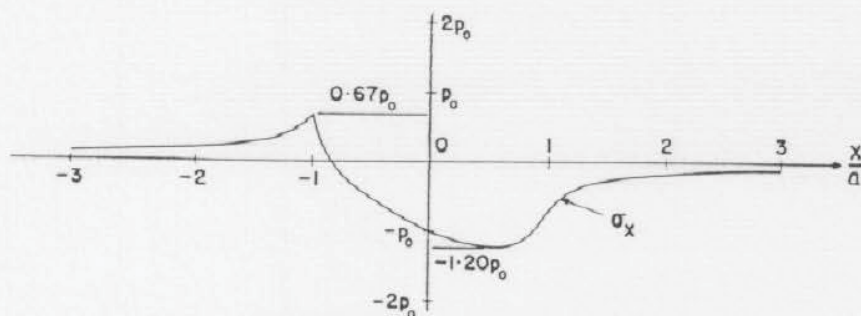


Distribution of direct stress at surface, in X-direction, resulting from purely tangential load.

b. Tangential loading



Elliptical distribution of combined normal and tangential loading intensity across contact width.



Distribution of direct stress at surface, in X-direction, resulting from combined normal and tangential load (coefficient of friction $\frac{1}{3}$)

c. Combined normal and tangential loading

Figure 3.12 Stress distributions for the sliding of a cylindrical contact [after ESDU 78035, 1978 and Johnson, 1985]

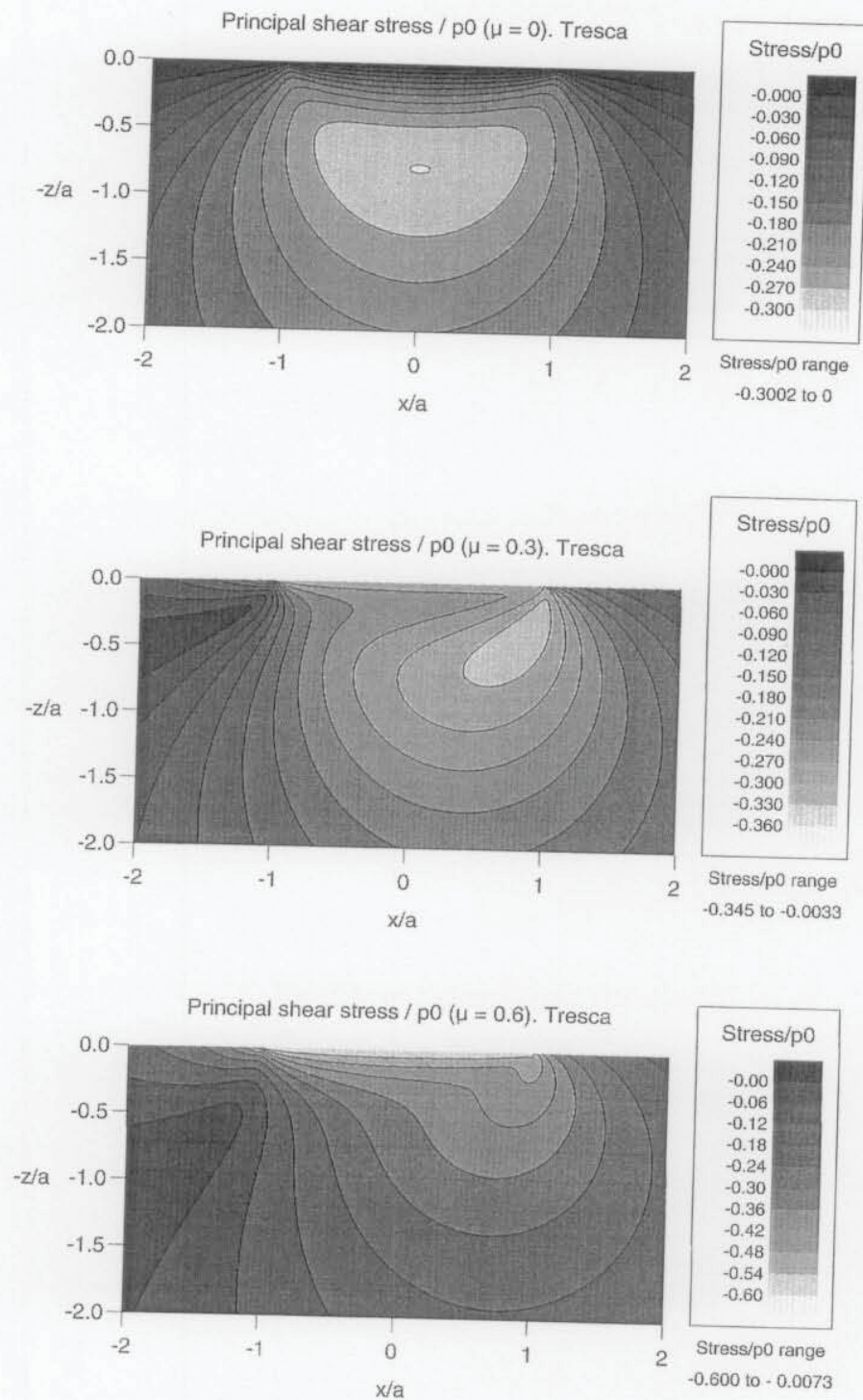


Figure 3.13 The effect of tangential loading on the size and distribution of shear stress with cylindrical contact, using different yield criteria^[from Cole, 1994].

(Note: μ is the traction coefficient, rather than the kinetic friction coefficient, in these figures.)

- (a) Principal shear stress (Tresca).
 (b) Von Mises (shear stress) (overleaf.....)

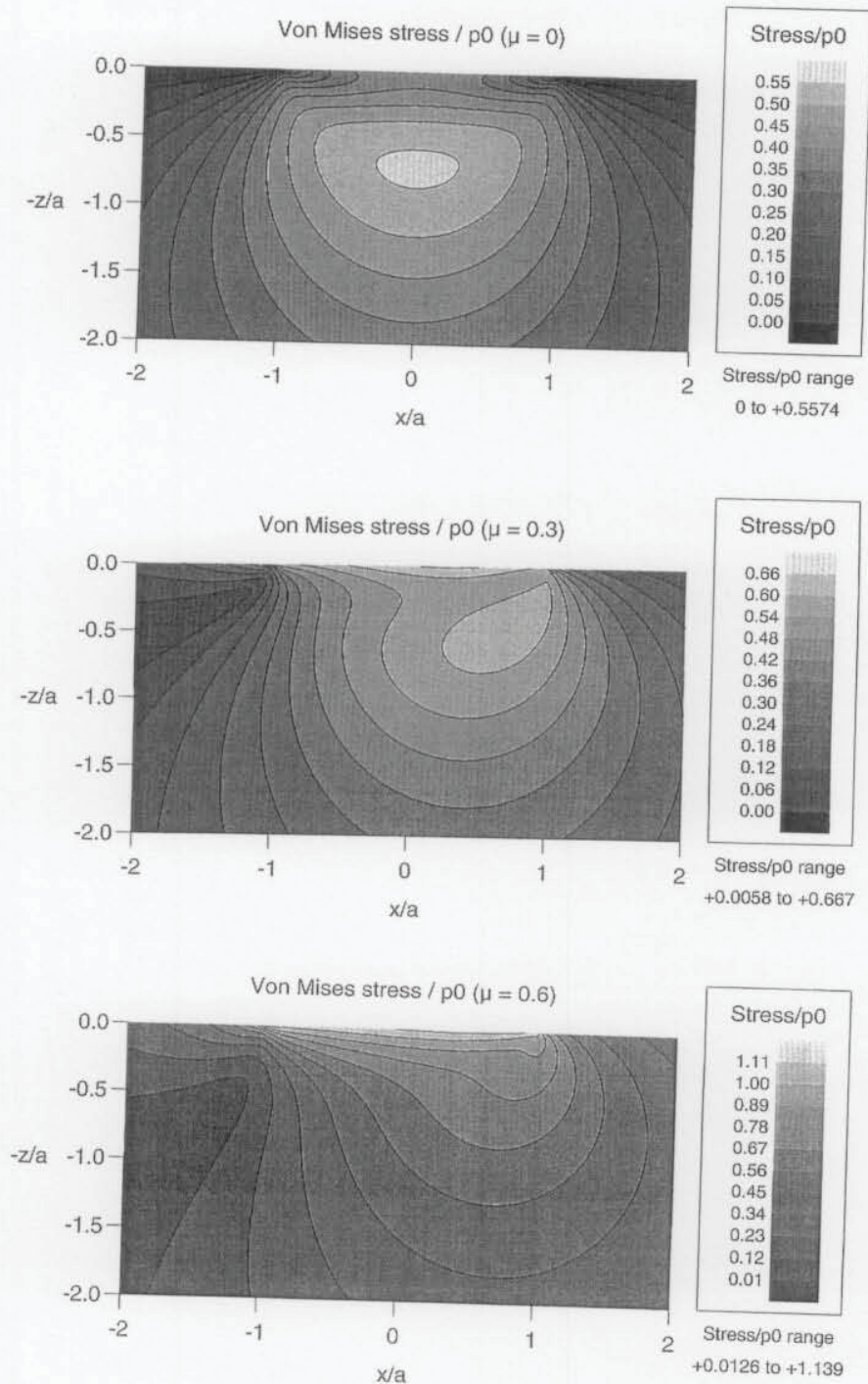


Figure 3.13b

Von Mises (shear) stress.

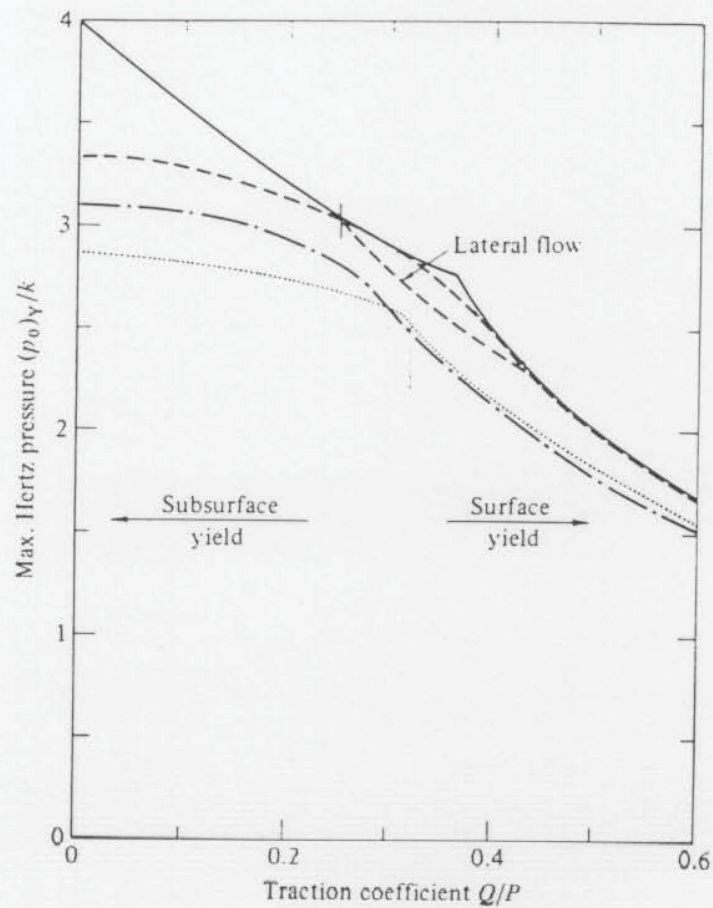


Figure 3.14 The effect of (sliding) friction on the contact pressures required for first yield and shakedown^[from Johnson, 1985].

- | | |
|-------------------|-----------------------------------------------|
| Large dashed line | : line contact, first yield (Tresca). |
| Chain line | : line contact, first yield (von Mises). |
| Small dashed line | : spherical contact, first yield (von Mises). |
| Solid line | : line contact, shakedown (Tresca). |

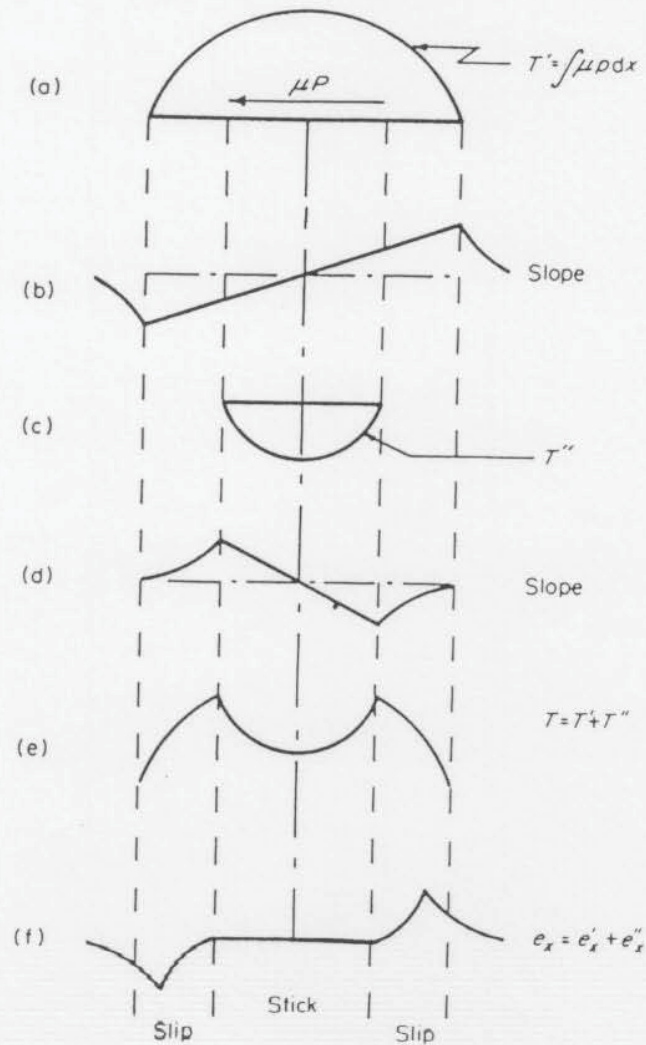


Figure 3.15 The distribution of tangential traction (T) and surface strains (ϵ_x) for a static contact where tangential force $Q < \mu P$. The distribution of traction (T') and surface strain (ϵ'_x) for sliding (cf. the direct stress at the surface from tangential loading shown in Figure 3.12b) are shown here in "a" and "b", respectively. These are opposed by traction (T'') and strain (ϵ''_x) in the central "stick region" of the contact, as shown in "c" and "d", respectively. The sum of the strains, representative of the contact, are shown in "e" and "f", respectively^[from Halling, 1975].
 (Note: The additional effect of normal loading (as in Figure 3.12c) is not shown here.)

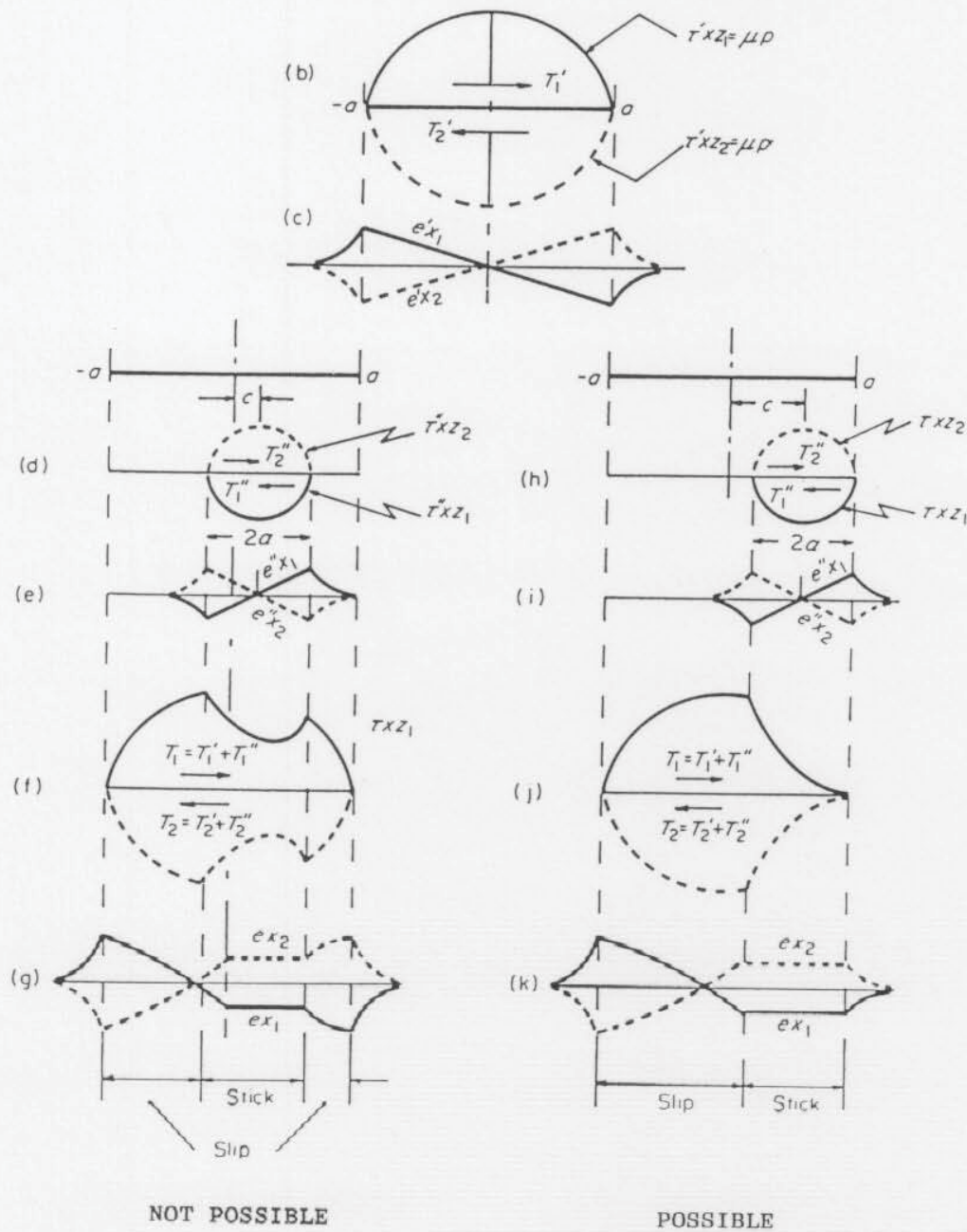


Figure 3.16 The distribution of tangential tractions for cylinders 1 and 2 (T_1 , T_2), and respective surface strains (ϵ_{x1} , ϵ_{x2}), for a rolling contact where tangential force $Q < \mu P$. With reference to Figure 3.15, the offset of the central stick region with rolling must be located at the leading edge of the contact (i.e. $c = a - \alpha$), as shown in sketches "h" to "k" of this figure. The situation shown in sketches "d" to "g" is not possible, as the forward area of slip would be in the same direction as traction^[after Halling, 1975].

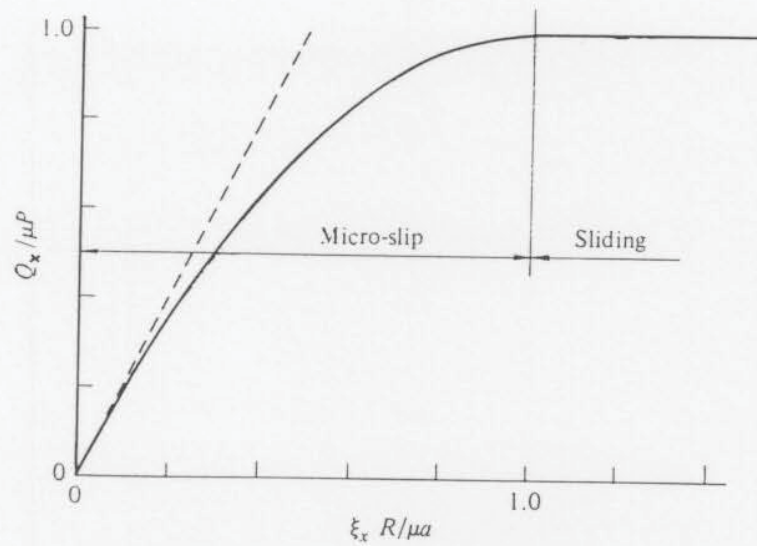


Figure 3.17 Creep curve for the tractive rolling of cylinders^[from Johnson 1985].

- Solid line : creep curve^[after Carter, 1926] cf. Equations 3.40a & b.
Dashed line : limiting curve for high friction with no slip,
cf. Equation 3.40c.

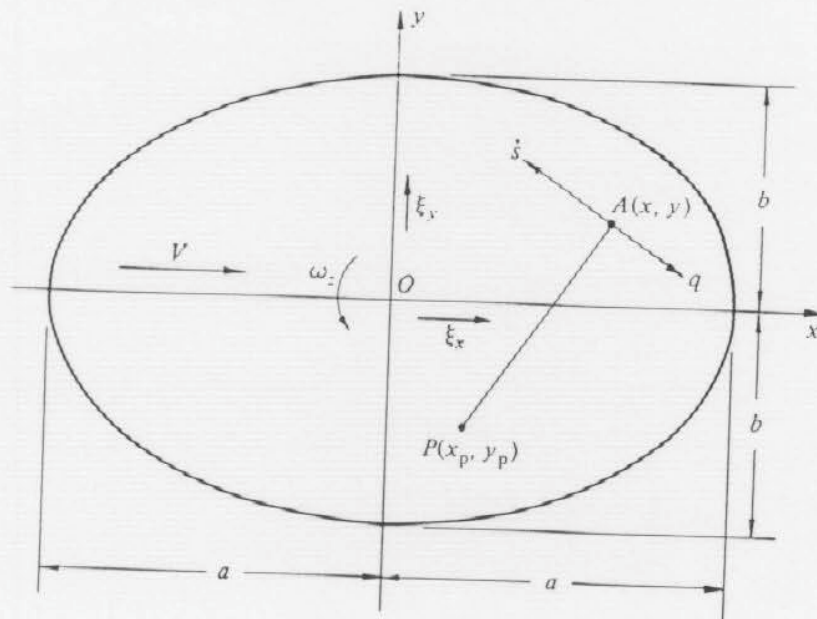


Figure 3.18 Tractive rolling with creep, spin and low friction; location of "spin pole", P^[from Johnson, 1985].

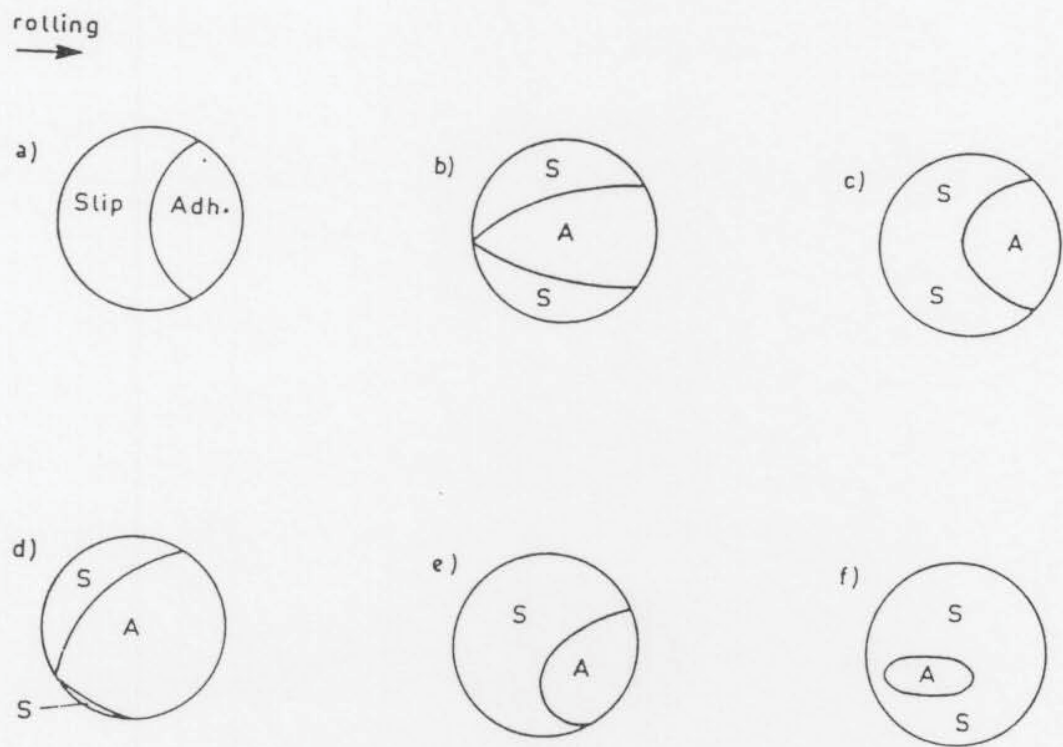


Figure 3.19 Areas of slip and adhesion for different combinations of longitudinal creepage, lateral creepage and spin^[from Kalker, 1979].

- (a) Pure longitudinal creepage.
- (b) Pure spin.
- (c) Lateral creepage with spin.
- (d) Longitudinal creepage with spin.
- (e) General mixed case.
- (f) Large amount of pure spin.

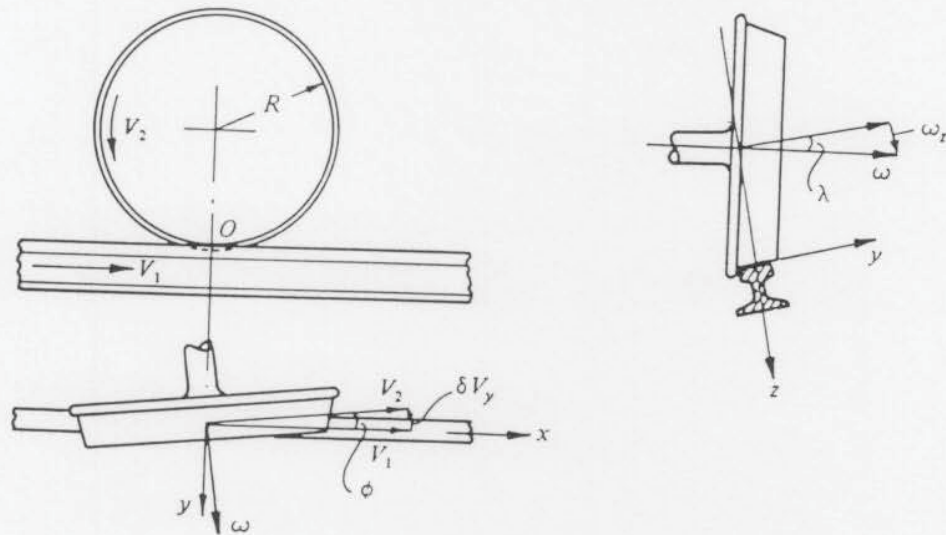


Figure 3.20 Creep motion of a railway wheel [from Johnson, 1985].

Longitudinal creep ratio: $\xi_x = (V_2 - V_1)/V_1$

Lateral creep ratio: $\xi_y = \delta V_y/V_1 = \tan \phi$

Spin parameter: $\psi = \omega(ab)^{1/2}/V_1 = \{(ab)^{1/2}/R\} \tan \lambda$

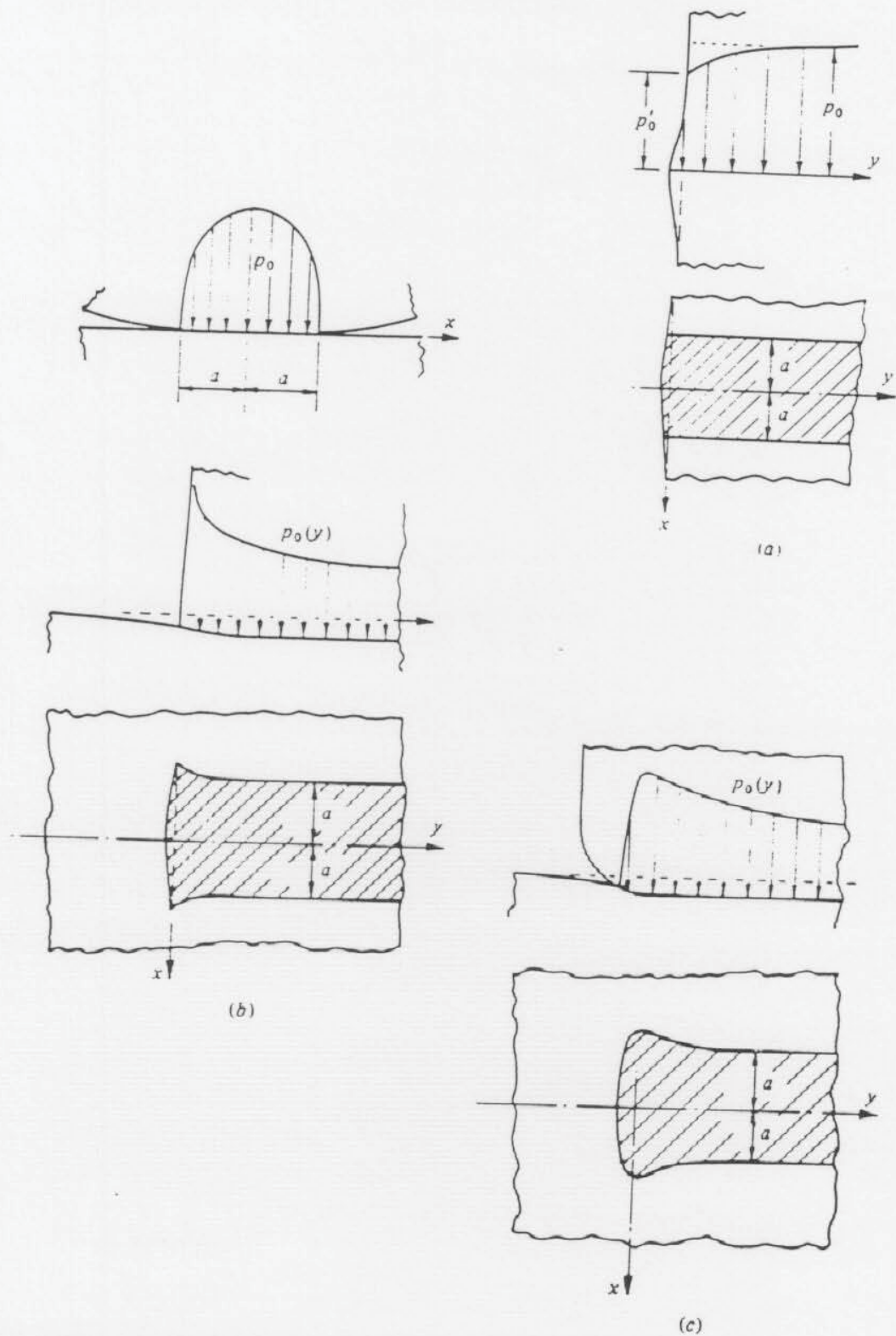


Figure 3.21 Effect of roller edge on maximum contact stress (p_o) distribution^[from Johnson 1985].

- (a) Two coincident sharp ends.
- (b) One ("embedded") sharp end.
- (c) One ("embedded") rounded end.

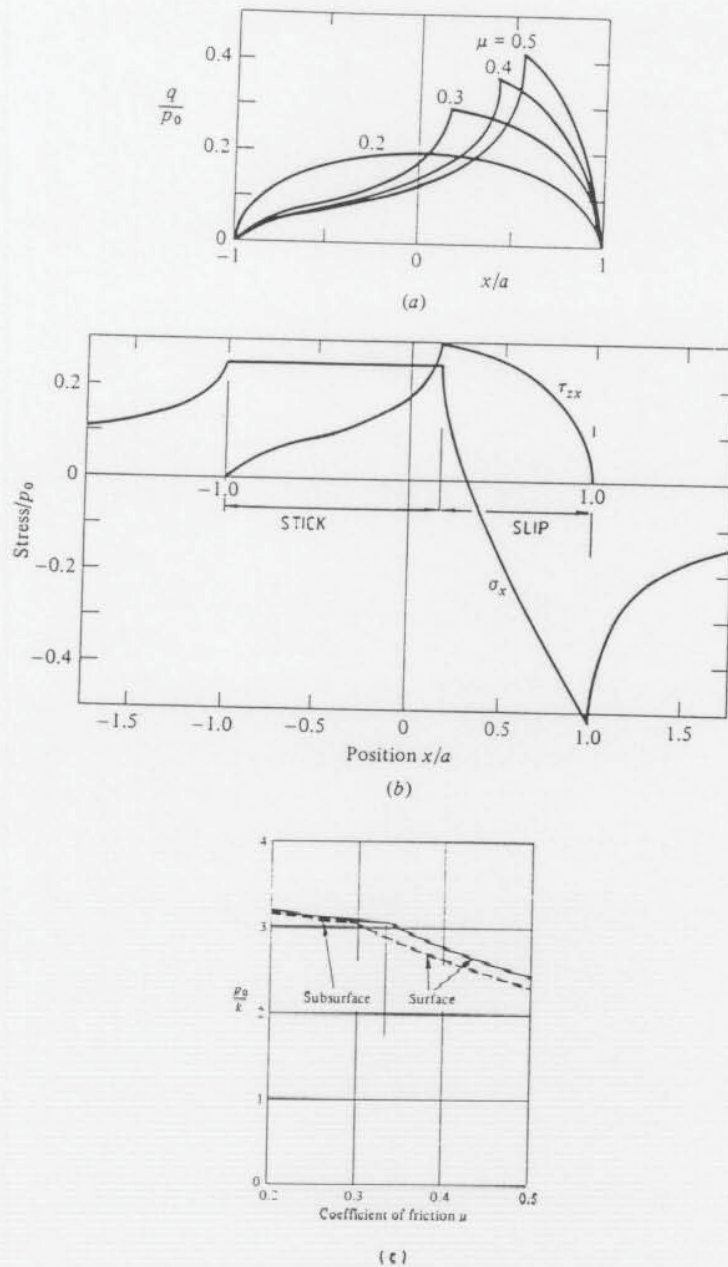


Figure 3.22 Tractive rolling (i.e. $Q < \mu P$) of cylinders, moving from right to left, with a tangential traction of $Q_x = 0.2P$ [from Johnson 1985].

- (a) The change in tangential surface traction distribution for increasing values of friction coefficient μ ; the area of slip is decreased.
- (b) The distribution of surface stresses, σ_x and τ_{zx} , where $\mu = 0.3$.
- (c) The relationship between the p_0/k values of first yield (dashed line) and shakedown (solid line) with increasing friction coefficient μ .

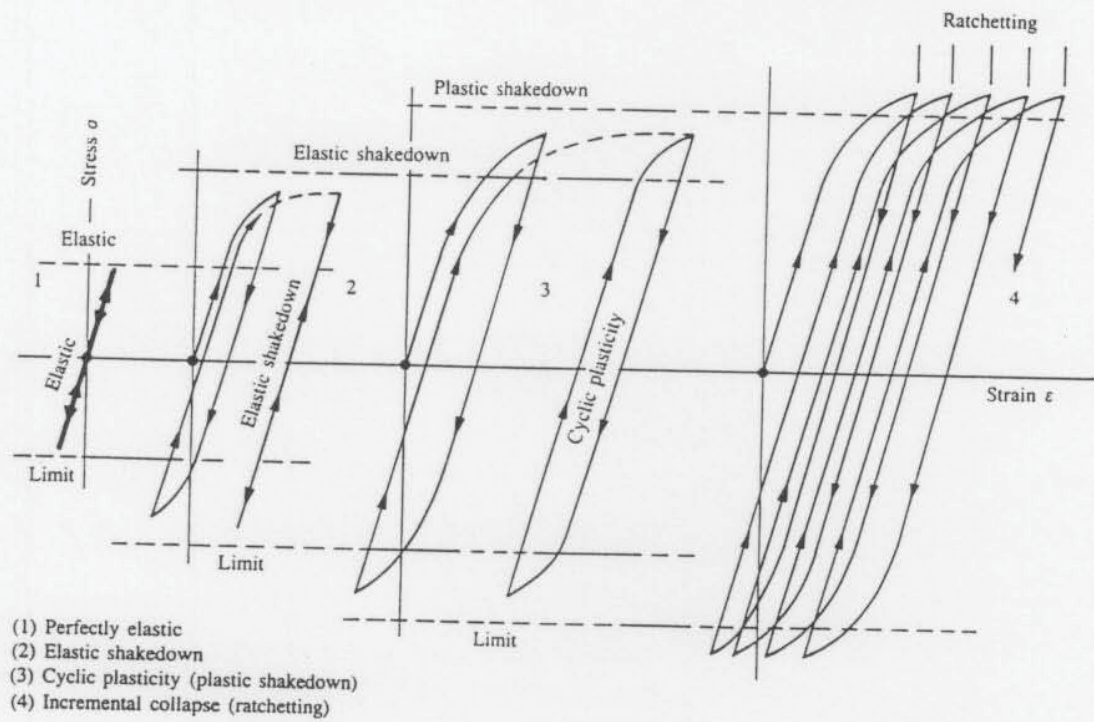


Figure 3.23 Material response to cyclic loading^[from Johnson 1988].

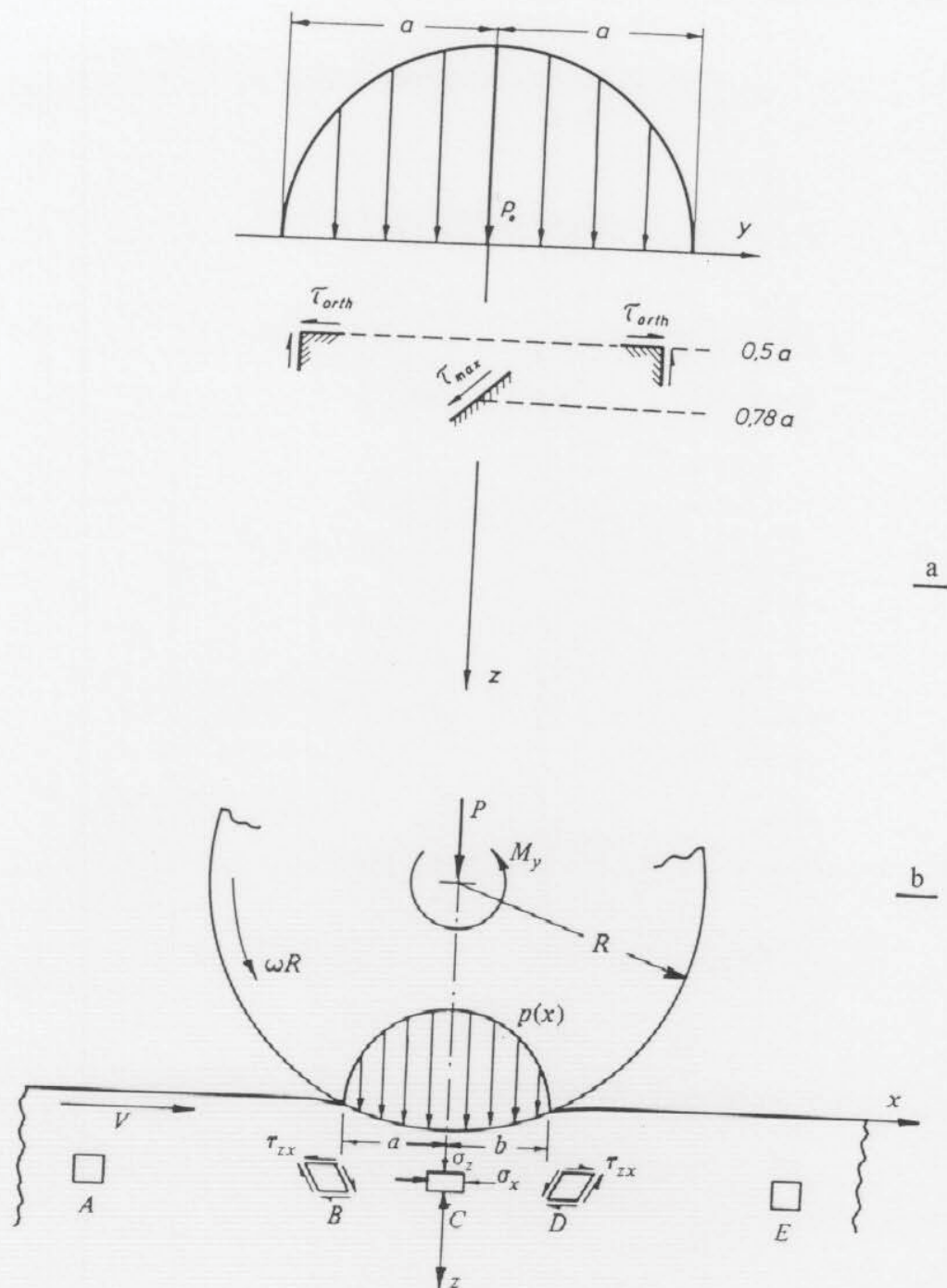


Figure 3.24 Sub-surface stress with cylindrical rolling contact.

(a) Location of principal shear stress and maximum orthogonal stress^[from Böhm et al, 1974].

(b) A material element experiences a cycle of reversed shear and compression, moving through A-B-C-D-E^[from Johnson, 1985].

overleaf.. (c) Orthogonal shear stress change with tractive coefficient^[from Cole, 1994].

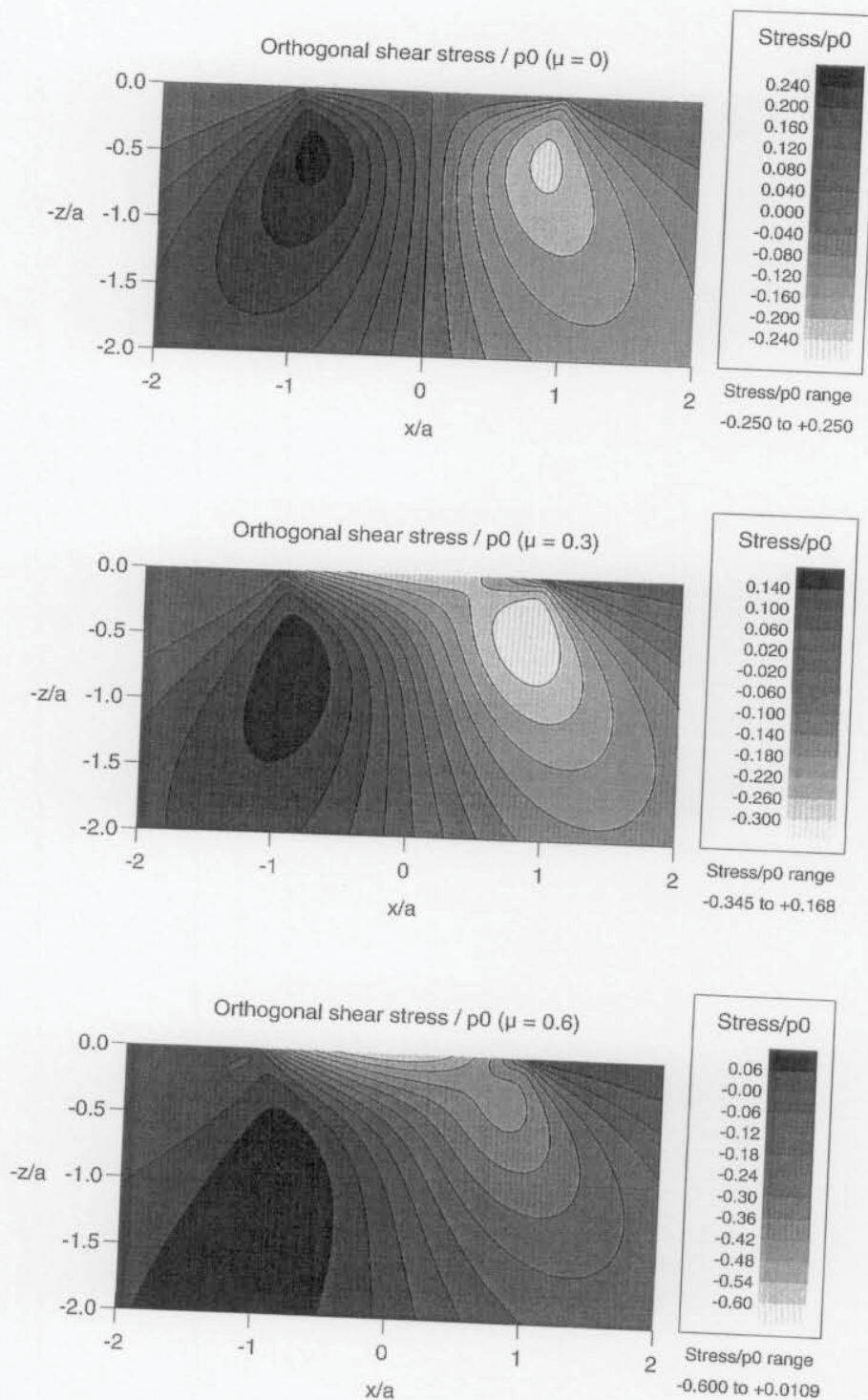


Figure 3.24c Orthogonal shear stress change with traction coefficient^[from Cole, 1994].
 (Note: μ is the traction coefficient, rather than the kinetic friction coefficient, in this figure.)

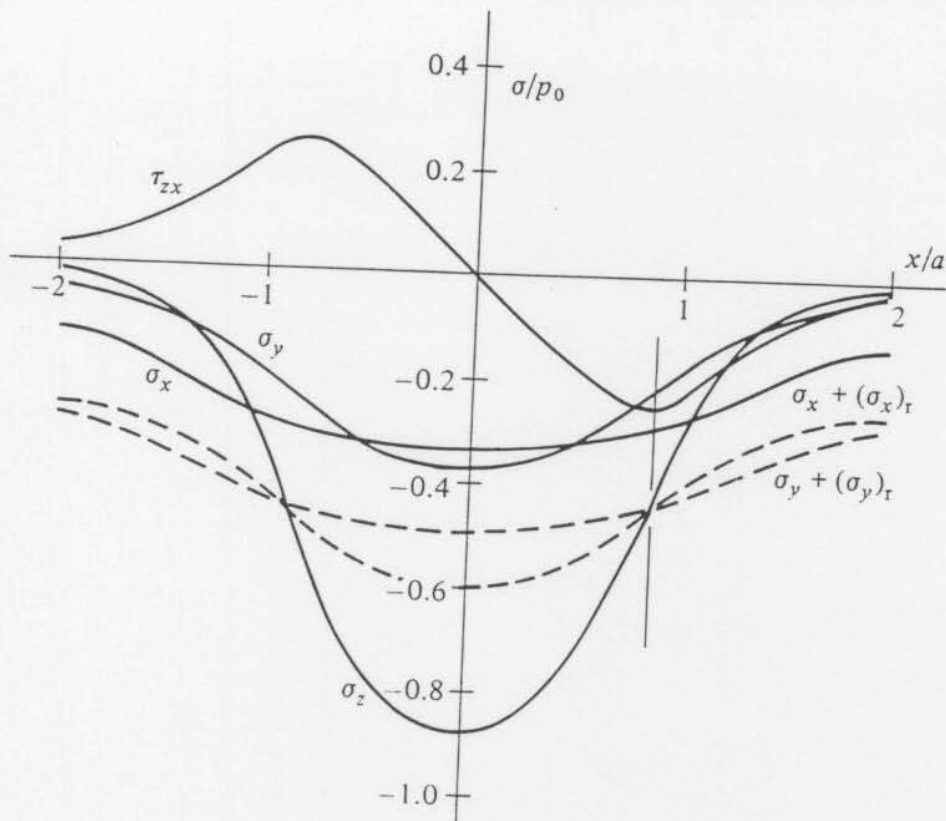


Figure 3.25 Pure rolling contact of elastic-plastic cylinders; stresses at depth $z = 0.5a$. Solid lines show elastic stresses; dashed lines show the addition of residual stresses $(\sigma_x)_r$ and $(\sigma_y)_r$ for shakedown^[from Johnson, 1985].

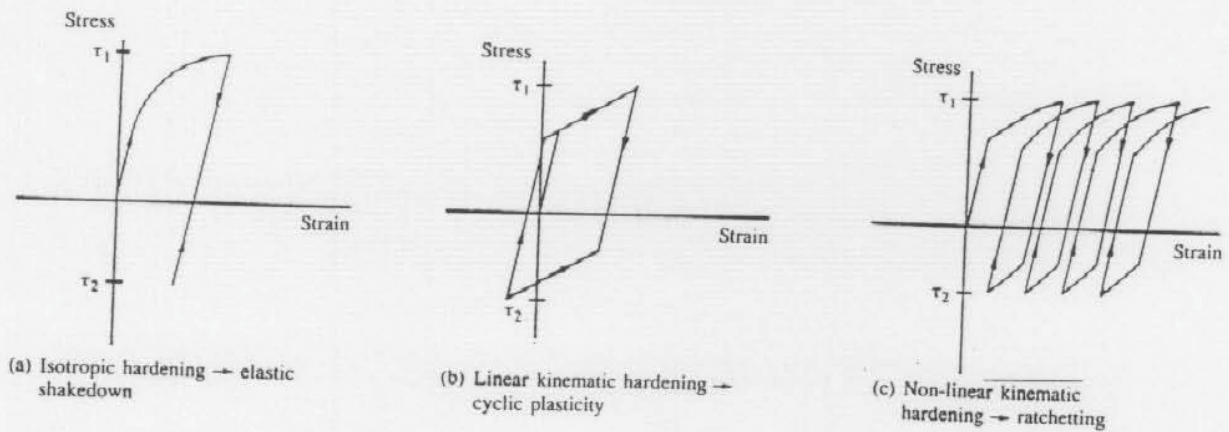


Figure 3.26 Response of material to unsymmetric cycles of shear stress ($\tau_1 > \tau_2$), as in tractive rolling^[from Johnson, 1988].

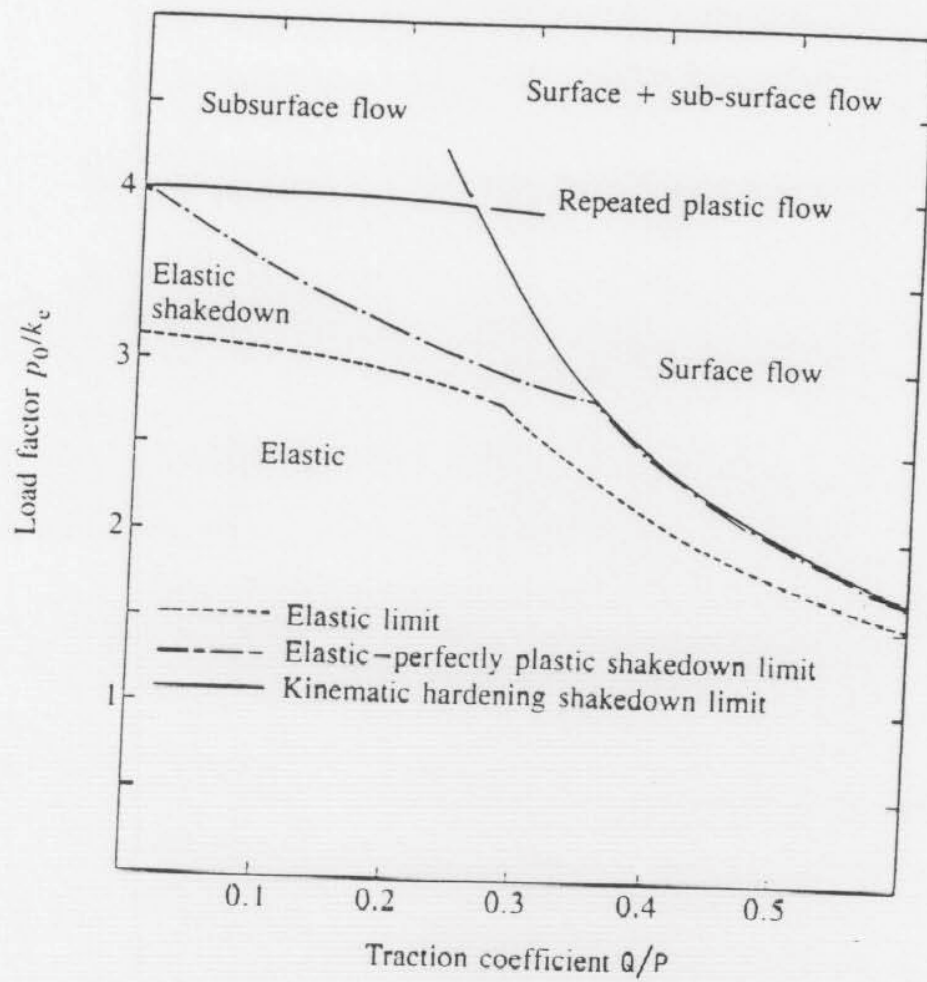


Figure 3.27 Shakedown map for the tractive rolling of cylinders, showing the relationship between the shakedown limits (for two types of material behaviour) and the traction coefficient^[from Johnson, 1988].

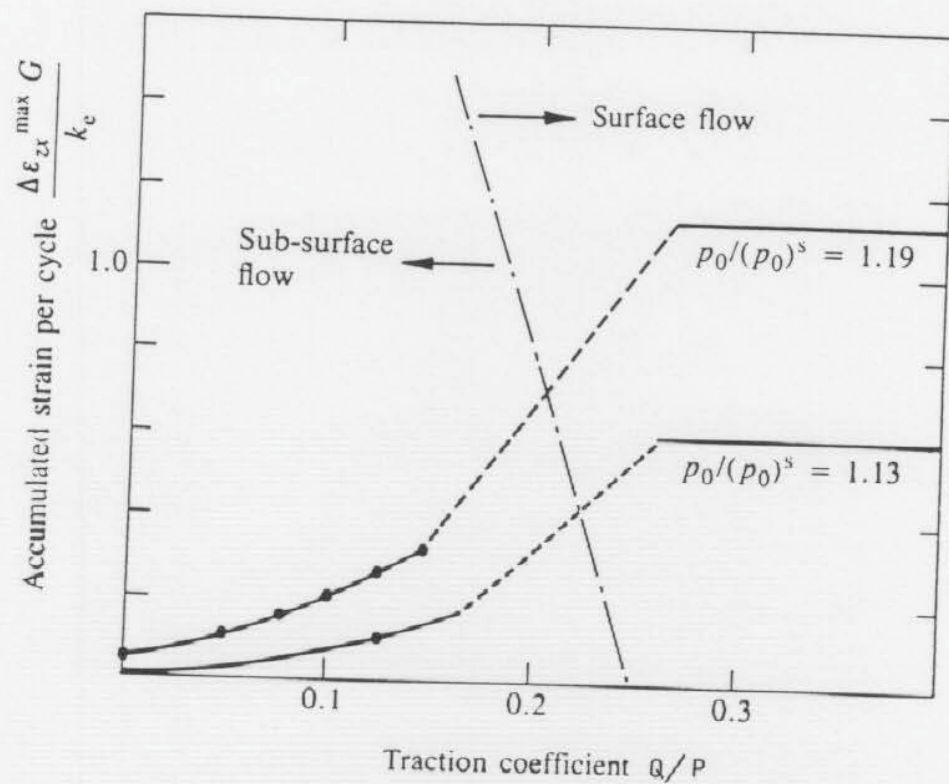
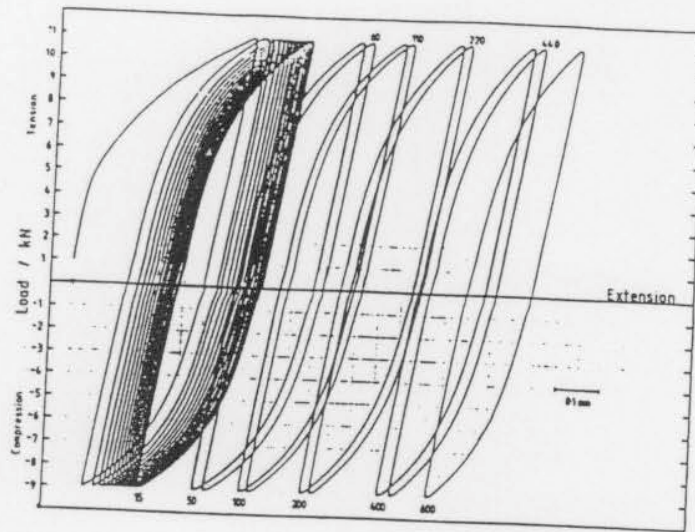
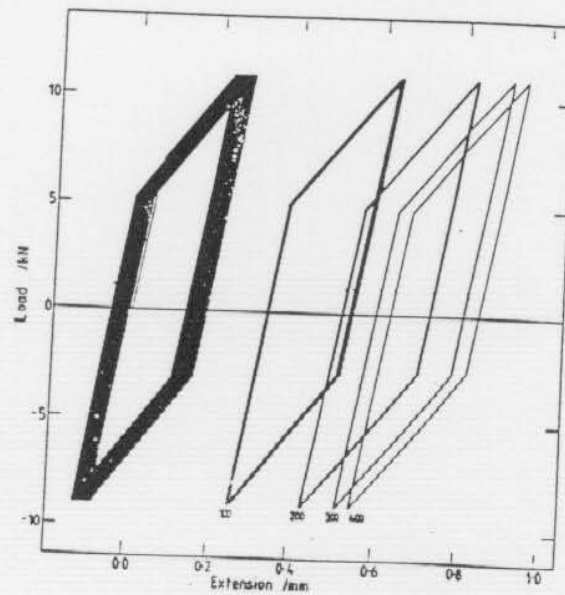


Figure 3.29 Ratchetting rate (of maximum strain accumulation) as a function of traction coefficient (Q/P) and load factor $p_0/(p_0)_s$, where $(p_0)_s$ is the shakedown value. Note the change in rate when the mechanism changes from sub-surface to surface^[from Johnson 1988].



a



b

Figure 3.30 Behaviour of a cylindrical specimen of BS11 rail steel under cyclic load-controlled tension-compression, with a mean tensile load [from Bower and Johnson, 1990].

(a) Measured response.

(b) Response predicted by non-linear kinematic hardening law.

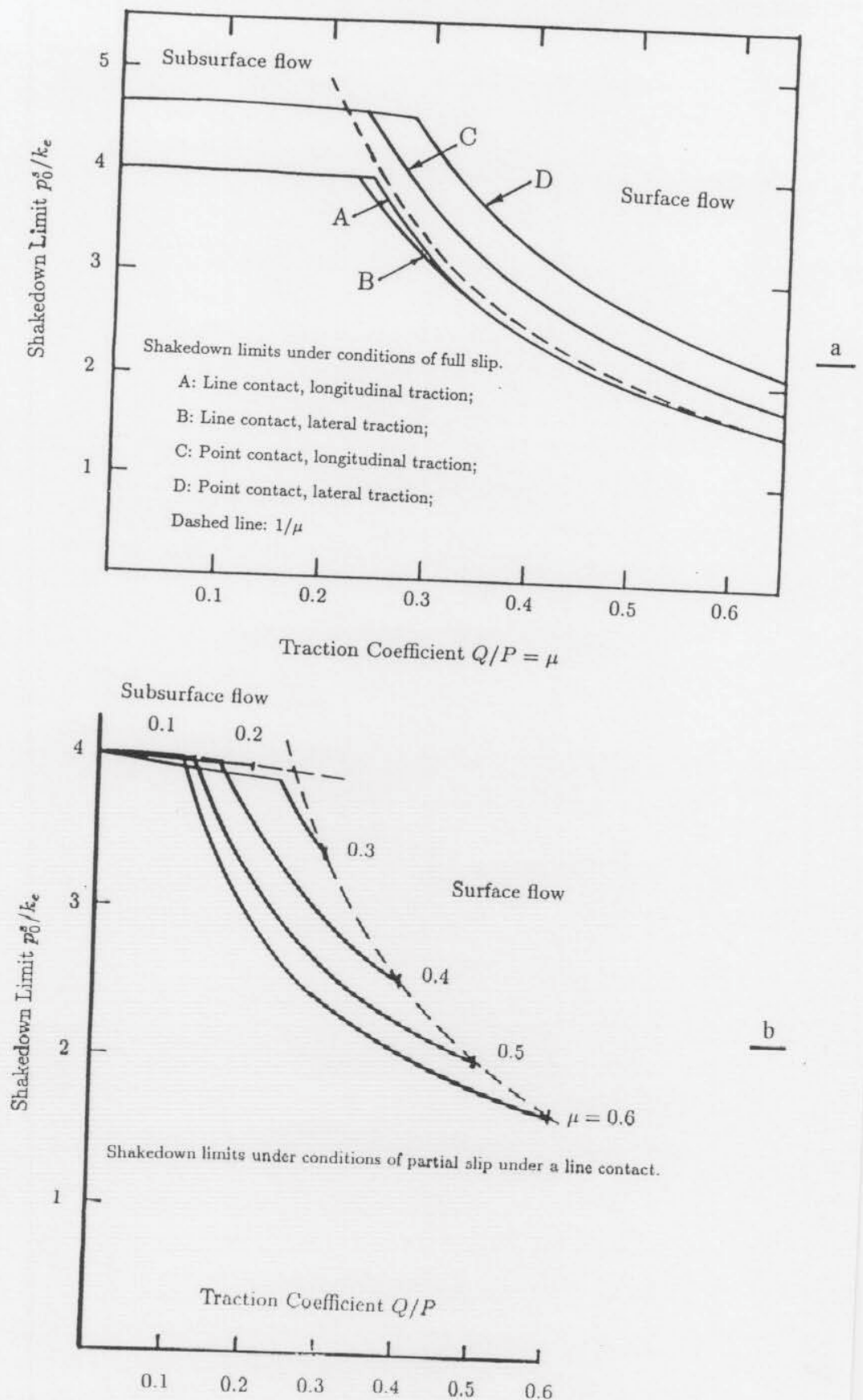
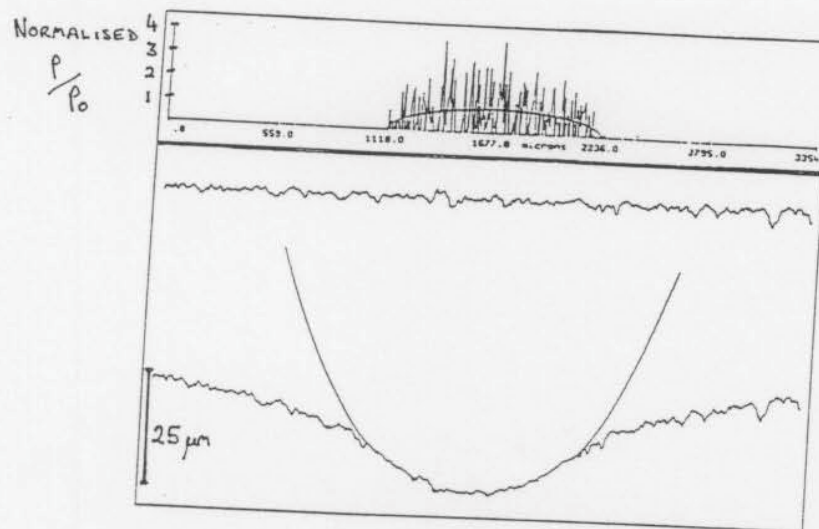
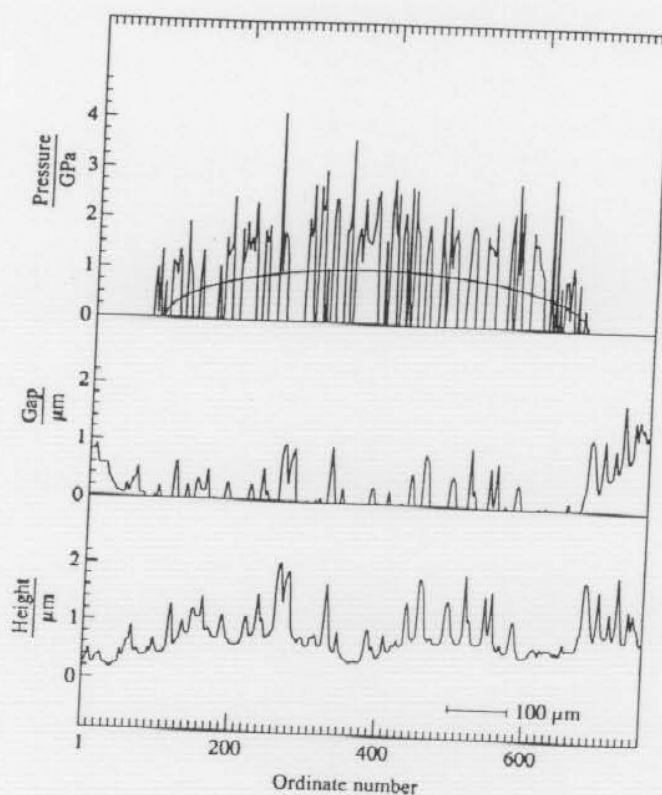


Figure 3.31 Shakedown maps calculated for conditions of contact loading relevant to railway practice^[from Bower and Johnson, 1990].

- (a) Under conditions of full slip for line and circular contacts.
 (b) Under conditions of partial slip for a line contact.



a



b

Figure 3.32 Simulated elastic contact of a ground surface with a smooth hard roller.

- (a) The distribution of actual maximum contact stress, due to the contact of discrete asperities, compared to the hertzian distribution. The elastic deformation is also shown^[from Bailey and Sayles, 1991].
- (b) Simulated contact between a run-in ground elastic cylinder and a rigid flat. The top graph shows variation from hertzian in (normalised) maximum contact stress^[from Snidle and Evans, 1994].