Subsurface stresses in an elliptical Hertzian contact

James Anthony Greenwood

Abstract
The traditional solution for the stresses below an elliptical Hertzian contact expresses the results in terms of incomplete Legendre elliptic integrals, so are necessarily based on the length of the semi-major axis \( a \) and the axis ratio \( k \). The result is to produce completely different equations for the stresses in the \( x \) and \( y \) directions; and although these equations are now well-known, their derivation from the fundamental, symmetric, integrals is far from simple. When instead Carlson elliptic integrals are used, they immediately match the fundamental integrals, allowing the equations for the stresses to treat the two semi-axes equally, and so providing a single equation where two were needed before. The numerical evaluation of the Carlson integrals is simple and rapid, so the result is that more convenient answers are obtained more conveniently. A bonus is that the temptation to record the depth of the critical stresses as a fraction of the length of the semi-major axis is removed. Thomas and Hoersch’s method of finding all the stresses along the axis of symmetry has been extended to determine the full set of stresses in a principal plane. The stress patterns are displayed, and a comparison between the answers for the planes of the major and minor semiaxes is made. The results are unchanged from those found from equations given by Sackfield and Hills, but not previously evaluated. The present equations are simpler, not only in the simpler elliptic integrals, but also for the “tail” of elementary functions.

Keywords
Subsurface stresses, elliptical contact, Hertzian contact, Carlson elliptic integrals

Introduction
The analysis of elliptical Hertzian contacts is most conveniently done using Love’s potential function

\[
\varphi = \int_y^0 \left( 1 - \frac{x^2}{a^2 + s} - \frac{y^2}{b^2 + s} \right) ds,
\]

where \( y \) is the positive root of

\[
\frac{x^2}{a^2 + s} + \frac{y^2}{b^2 + s} + \frac{s^2}{k} = 0.
\]

Thus, from Johnson’s Contact Mechanics, the displacements due to a Hertzian pressure distribution

\[
p = p_0 [1 - x^2/a^2 - y^2/b^2]^{1/2},
\]

are

\[
w = \frac{abp_0}{2E} \int_0^y \left( 1 - \frac{x^2}{a^2 + s} - \frac{y^2}{b^2 + s} \right) \frac{ds}{[a^2 + s(b^2 + s)]^{1/2}}.
\]

Within the loaded region we require

\[
w = x^2/2R_1 + y^2/2R_2,
\]

so that (since within the loaded region \( y = 0 \)), we have

\[
\frac{1}{2R_1} = \frac{abp_0}{2E} \int_0^\infty \frac{ds}{s^{1/2}(a^2 + s)^{3/2}(b^2 + s)^{1/2}}
\]

and

\[
\frac{1}{2R_2} = \frac{abp_0}{2E} \int_0^\infty \frac{ds}{s^{1/2}(a^2 + s)^{1/2}(b^2 + s)^{3/2}}.
\]

Those with considerable mathematical dexterity can transform these integrals into Legendre elliptic integrals

\[
(K(e) - E(e)) \quad \text{and} \quad E(e) - (1 - e^2)K(e)
\]

(where \( e^2 = 1 - b^2/a^2 \)); but an easier alternative is to use Carlson elliptic integrals, defined as

\[
R_F(p, q, r) = \frac{1}{2} \int_0^\infty \frac{dt}{(t + p)^{1/2}(t + q)^{1/2}(t + r)^{1/2}}
\]

and

\[
R_D^*(p, q; r) = \frac{1}{2} \int_0^\infty \frac{dt}{(t + p)^{1/2}(t + q)^{1/2}(t + r)^{3/2}}
\]

(Carlson\textsuperscript{3} or Greenwood\textsuperscript{4} (Note that all three parameters in \( R_F \) are interchangeable, but only the first two in \( R_D^* \). The spot * indicates that this is one-third of Carlson’s \( R_D \)). Now no mathematical dexterity is required to see that

Department of Engineering, University of Cambridge, Cambridge, UK

Corresponding author:
JA Greenwood, Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, UK.
Email: jag@eng.cam.ac.uk

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/03093247211060288
journals.sagepub.com/home/sdj

journals.sagepub.com/home/sdj

J Strain Analysis
2022, Vol. 57(8) 677–687
© IMechE 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/03093247211060288
journals.sagepub.com/home/sdj

J Strain Analysis
2022, Vol. 57(8) 677–687
© IMechE 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/03093247211060288
journals.sagepub.com/home/sdj

Original Article
\[
\frac{1}{2R_1} = \frac{abp_0}{E} R_D^*(0, b^2 ; a^2); \quad \frac{1}{2R_2} = \frac{abp_0}{E} R_D^*(0, a^2 ; b^2).
\]

(Equally readily, the displacement is found to be \(\delta = \frac{abp_0}{E} R_p(0, a^2 , b^2)\).)

Thus, using the natural affinity between Love’s potential function and Carlson elliptic integrals, we have replaced a pair of strongly differing equations by a pair of simple equations. Actually, we have done better than that: we have replaced the pair by a single equation. For the curvature in the \(x\)-direction, the square of the semi-axis in the \(x\)-direction is the third argument: for the curvature in the \(y\)-direction, the third argument is the square of the semi-axis in the \(y\)-direction. We have a single equation, readily obtained, with no mention of major and minor axes.

True, there are no tables of Carlson integrals, and it appears that the two functions \(K(e)\) and \(E(e)\) of a single variable have been replaced by two functions of two variables. But here the Carlson scaling laws come in. A variable have been replaced by two functions of major and minor axes. It shows that \((c/\sqrt{2})\) and \((c/\sqrt{2})\) are \(R_D^*(p, q, r)\) and \(R_D^*(p, q, r)\), so that one argument can simply be reduced to 1. Thus the complete Legendre integrals of a single variable can be replaced by Carlson integrals with two passive arguments (0, 1) and calculated as functions of the single remaining variable. The incomplete Legendre integrals, functions of \(e\) (more often denoted by \(k\) or \(m=k^2\)) and the modular angle could be replaced by Carlson integrals with just two active arguments (i.e. differing from unity). In the computer age, this will probably never be done: a very simple, very rapid algorithm for calculating Carlson integrals is available. It is as simple as the recommended algorithm for the complete Legendre integrals, and since the same algorithm is used for all arguments, is distinctly preferable to the usual method for incomplete Legendre integrals.

A previous article\(^4\) shows the advantages to be gained by the use of Carlson integrals in the analysis of the surface properties of Hertzian contacts; the determination of the contact geometry as cited above, the determination of the surface stresses under normal pressures as discussed by Hertz, and those caused by tangential loads as studied by Mindlin and others. Here we consider to what extent the analysis of the subsurface stresses could equally benefit.

The prospective advantages are of course the same: (1) Natural affinity between Love’s potentials and Carlson integrals (2) No need to distinguish between major and minor semi-axes: to assume \(a \equiv b\) and focus on the ratio \(b/a\); (3) Treating the two directions equally promotes a beneficent mindset; (4) The convenience of ending with a single equation instead of two, usually very different, ones. And to the author, perhaps alone, the advantage of dimensioned equations, providing an immediate check for errors; others may not need this!

The convenience of Carlson integrals over complete Legendre integrals is perhaps minor, the real gain is in the easier derivation: but see Discussion for what happens when we encounter incomplete Legendre integrals.

### Analysis of Elliptical Hertzian Contact

The basis of the solution is spelt out by Love,\(^1\) and taken on from there by Thomas and Hoersch,\(^5\) [“T\&H”] and it is hard to improve on their presentation, up to the point where they are reduced to converting their answers into Legendre integrals, and so forcing their admirable, symmetric, equations into the inevitable contorted form.

Love introduced two potential functions \(\phi\) and \(\chi\), defined as

\[
\phi = \int_y^x \left[ \frac{1 - \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{t}}{\sqrt{\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{t}\right)}} \right] \frac{dt}{\gamma} \quad \text{and} \quad \chi = \int_z^x \varrho \, dz,
\]

where \(\gamma\) is the largest root of

\[
f = \frac{x^2}{a^2 + \gamma} + \frac{y^2}{b^2 + \gamma} + \frac{z^2}{\gamma} - 1 = 0
\]

Love gives equations only for the displacements: it seems that T\&H were the first to continue and find the strains and hence obtain the basic equations for the stresses (see also Johnson\(^2\) or Hills et al.\(^6\) (HNS)).

*(A factor \((\pi/2)abp_0\) will subsequently be inserted for a Hertzian pressure distribution.)*

\[
2\pi \sigma_x = 2 \nu \frac{\partial \phi}{\partial x} - z \frac{\partial^2 \phi}{\partial x^2} - (1 - 2
\nu) \frac{\partial^2 \chi}{\partial x^2} \quad \text{(3a)}
\]

\[
2\pi \sigma_y = 2 \nu \frac{\partial \phi}{\partial y} - z \frac{\partial^2 \phi}{\partial y^2} - (1 - 2\nu) \frac{\partial^2 \chi}{\partial y^2} \quad \text{(3b)}
\]

\[
2\pi \sigma_z = \frac{\partial \phi}{\partial z} - z \frac{\partial^2 \phi}{\partial z^2} \quad \text{(3c)}
\]

Subsequently Fessler and Ollerton\(^7\)

\[
2\pi \tau_{xz} = -z \frac{\partial^2 \phi}{\partial x \partial z} \quad 2\pi \tau_{yz} = -z \frac{\partial^2 \phi}{\partial y \partial z} \quad \text{(3d)}
\]

The equation for \(\tau_{yz}\), which we do not need, is given by Johnson\(^2\) and Hills et al.\(^6\) added.

Note the complete symmetry between the pairs \((\sigma_x, \tau_{xz})\) and \((\sigma_y, \tau_{yz})\); this is lost in the final answers.
when Legendre integrals are used, but is preserved when we use Carlson integrals.

Since \( \phi \) and \( \chi \) are both harmonic, and \( \varphi = \frac{\partial \phi}{\partial z} \) so that

\[
\frac{\partial \varphi}{\partial z} = \frac{\partial \varphi}{\partial z}
\]

then from equations (3)

\[
2\pi(\sigma_x + \sigma_y + \sigma_z) = (1 + 4\nu) \frac{\partial \varphi}{\partial z} + (1 - 2\nu) \frac{\partial \varphi}{\partial z}
\]

\[
= 2(1 + \nu) \frac{\partial \varphi}{\partial z}
\]

(4)

We note that \( (\sigma_{x+y} + \sigma_{y+z}) \) is independent of the second potential function \( \chi \). Indeed, not all the stresses involve \( \chi \), or elliptic integrals; for \( (\sigma_z, \tau_{xz}, \tau_{yz}) \) are all elementary functions, independent of \( \nu \).

**Extracts from Thomas and Hoersch’s analysis**

We introduce a spurious \( \alpha \), which is identically zero, in order to make clearer the permutations to obtain \( \frac{\partial f}{\partial z} \), \( \frac{\partial f}{\partial z} \) from \( \frac{\partial g}{\partial z} \) thus we write the potential (1) as

\[
\varphi = \int_{\gamma}^{\gamma} \left[ 1 - \frac{x^2}{a^2 + t} + \frac{y^2}{b^2 + t} - \frac{z^2}{c^2 + t} \right] dt
\]

\[
\sqrt{(a^2 + t)(b^2 + t)(c^2 + t)}
\]

(1')

Differentiating \( \varphi \) is straightforward, but lengthy. The derivatives of both integrand and lower limit are needed, (though not for the first derivative, for there the integrand vanishes) and T&H show that

\[
\frac{\partial \gamma}{\partial x} = \frac{2x}{a^2 + \gamma} \left( \frac{x^2}{a^2 + \gamma} + \frac{y^2}{b^2 + \gamma} + \frac{z^2}{c^2 + \gamma} \right)^{-1}
\]

\[
= \frac{\partial f}{\partial x} / H
\]

(5)

\[
f = \frac{x^2}{a^2 + \gamma} + \frac{y^2}{b^2 + \gamma} + \frac{z^2}{c^2 + \gamma} - 1
\]

\[
H = \frac{x^2}{(a^2 + \gamma)^2} + \frac{y^2}{(b^2 + \gamma)^2} + \frac{z^2}{(c^2 + \gamma)^2}.
\]

Thus

\[
\frac{\partial \varphi}{\partial x} = \int_{\gamma}^{\gamma} \frac{2x}{a^2 + t} \cdot \frac{dt}{\sqrt{(a^2 + t)(b^2 + t)(c^2 + t)}}
\]

and permuting,

\[
\frac{\partial \varphi}{\partial z} = \int_{\gamma}^{\gamma} \frac{2z}{c^2 + t} \cdot \frac{dt}{\sqrt{(a^2 + t)(b^2 + t)(c^2 + t)}}
\]

(6)

This is readily seen to be

\[
\frac{\partial \varphi}{\partial z} = -4z R^*_{x-y}(a^2 + \gamma, b^2 + \gamma, c^2 + \gamma)
\]

Hence, for our Hertzian contact,

\[
(\sigma_x + \sigma_y + \sigma_z)/\rho_0 = -ab(1 + \nu) \cdot z R^*_{x-y}
\]

\[
(a^2 + \gamma, b^2 + \gamma; c^2 + \gamma)
\]

The second derivatives involve both terms: thus

\[
\frac{d^2 \varphi}{dx^2} = -2 \int_{\gamma}^{\gamma} \frac{dt}{\sqrt{(a^2 + t)(b^2 + t)(c^2 + t)}} + \frac{1}{\sqrt{(a^2 + \gamma)(b^2 + \gamma)(c^2 + \gamma)}} \cdot \frac{4x^2}{H}
\]

\[
(7a)
\]

which, by writing \( p = a^2 + \gamma; q = b^2 + \gamma; r = c^2 + \gamma \) (= \( \gamma \)), becomes

\[
\frac{d^2 \varphi}{dx^2} = -2 \int_{\gamma}^{\gamma} \frac{dt}{\sqrt{(a^2 + t)(b^2 + t)(c^2 + t)}} + \frac{1}{\sqrt{(p)(q)(r)}} \cdot \frac{4x^2}{H}
\]

Using Carlson integrals this is

\[
\frac{d^2 \varphi}{dx^2} = -4 R^*_{x-y}(a^2 + \gamma, b^2 + \gamma, c^2 + \gamma) + \frac{1}{\sqrt{(p)(q)(r)}} \cdot \frac{4x}{H}
\]

\[
(7b)
\]

Permuting the symbols gives

\[
\frac{d^2 \varphi}{dz^2} = -2 \int_{\gamma}^{\gamma} \frac{dt}{\sqrt{(a^2 + t)(b^2 + t)(c^2 + t)}} + \frac{1}{\sqrt{(p)(q)(r)}} \cdot \frac{4z^2}{H}
\]

\[
(7c)
\]

Note that \( z \frac{d \varphi}{dz} = \frac{\partial f}{\partial z} + \frac{1}{\sqrt{(p)(q)(r)}} \cdot \frac{4z^3}{H} \).

For \( \frac{\partial \varphi}{\partial \alpha} \) the derivative of the integrand vanishes, leaving (using (5)) just

\[
\frac{d^2 \varphi}{d\alpha^2} = \frac{2x}{\sqrt{p^2 q r}} \cdot \frac{2z}{r} \cdot \frac{1}{H} = \frac{4xz}{\sqrt{p^2 q r}} \cdot \frac{1}{H}
\]

From the basic equation (3) it will be clear that for an incompressible solid (\( \nu = 1/2 \)) these are all we need: \( 2\pi \sigma_x = 2\pi \frac{\partial \varphi}{\partial x} - \frac{\partial f}{\partial x} - (1 - 2\nu) \frac{\partial f}{\partial x} \) reduces to \( 2\pi \sigma_x = \frac{\partial f}{\partial x} \), while for all values of \( \nu \), \( 2\pi \sigma_x = \frac{\partial f}{\partial x} \), and equally from (3d) we can obtain \( \tau_{xz}, \tau_{yz} \).

Otherwise we need the second potential \( \chi \). T&H begin by showing
\frac{\partial^2\chi}{\partial x^2} = z \frac{\partial^2\phi}{\partial x^2} + F_z(z) \text{ where } F_z(z) = \int_z^\infty \frac{\partial^3\phi}{\partial z \partial x^2} \, dz \quad (8)

and we need to differentiate \( (7a): \)
\[
\frac{\partial^2\phi}{\partial x^2} = -2 \int_\gamma \frac{dy}{\sqrt{(a^2 + \gamma)(b^2 + \gamma)(c^2 + \gamma)}},
\]
where the \( z \) in the numerator is itself \( z(\gamma) \) and makes the integration difficult! But on the axis we have simply \( \gamma = z^2 \); so substituting (and dropping \( c^2 \)) we get:
\[
F_z(z) = 2 \int_\gamma \frac{dy}{\sqrt{(a^2 + \gamma)(b^2 + \gamma)}} = 2R_b(a \mid \gamma),
\]
and T & H show that the integral \( I(b; a \mid \gamma) \) equals
\[
\frac{2}{a^2 - b^2} \left[ 1 - \sqrt{\frac{b^2}{a^2} + 1} \right] \text{ as is readily checked by differentiation.}
\]
T&H stop there, with an equation that appears to fail for a circular contact, but a little algebra gives
\[
\int_\gamma \frac{dt}{\sqrt{(a^2 + t)(b^2 + t)(c^2 + t)}} = 2R_b^*(q, r; p) \quad (10)
\]
However T&H needed numerical answers for their stresses, and in 1930 the only elliptic integrals available were Legendre’s: so they transformed (spoilt?) their beautiful, symmetric, integrals by substituting \( t = \sqrt{a^2 \omega^2} \), \( \gamma = z^2 \), and writing \( k = b/a \), so that, for example, the \( \frac{\partial^2\phi}{\partial x^2} \) integral \( (7a) \)
\[
\int_\gamma \frac{dt}{\sqrt{(a^2 + \gamma)(b^2 + \gamma)(c^2 + \gamma)}} = 2R_b^*(q, r; p)
\]
became \( \frac{1}{a^2} \int_\gamma \frac{dt}{\sqrt{(1 + \omega^2)(k^2 + \omega^2)}} \),
which they identified as \( \frac{1}{a^2} \int \frac{1}{k^2} [F(\xi, k^2) - E(\xi, k^2)] \)
(\( k^2 = 1 - \xi^2 \) and \( \xi = \cot \phi \). (Inconsiderate of Love to use for his potential the standard symbol for the modular angle)).

But the \( \frac{\partial^2\phi}{\partial x^2} \) integral \( (7b) \), which we know from symmetry to be
\[
\int_\gamma \frac{dt}{\sqrt{(a^2 + t(b^2 + t)(c^2 + t)}} = 2R_b^*(p, r; q),
\]
was found to become \( \frac{1}{a^2} \int [E(\xi, k^2) / k^2 - F(\xi, k^2)] \) + algebraic terms (see Discussion).

The Carlson forms needed no changes of variable or careful mathematics, and retain the natural symmetry.

**Stresses on the axis of symmetry**

The stresses are conventionally given as the sum of the answer for \( \nu = 0 \) and a term proportional to \( \nu \), as
\[
\frac{\partial\phi}{\partial x} = \Omega + \nu \Omega', \text{ so we have (after restoring the factor } \frac{\pi}{2} ab p_0 \text{)}
\]
\[
\Omega = -\frac{(ab/2)[z \frac{\partial^2\phi}{\partial x^2} + 0.5F_z(z)]};
\]
\[
\text{and } \Omega' = \frac{(ab/2)[\frac{\partial\phi}{\partial x} + z \frac{\partial^2\phi}{\partial x^2} + F_z(z)]}.
\]

Using Carlson integrals, \( \frac{\partial\phi}{\partial x} = -4ab R_b^*(p, q; r) \) and \( \frac{\partial^2\phi}{\partial x^2} = -4R_b^*(p, q; r) \),

(11a)
\[
\Omega = 2abz R_b^*(q, r; p) + \frac{ab}{\sqrt{\beta} (\sqrt{\beta} + \sqrt{q})}
\]
\[
\Omega' = -\frac{(2abz)}{\sqrt{\beta} \sqrt{\beta} + \sqrt{q}) R_b^*(p, q; r) + R_b^*(p, q; r)}
\]
\[
\text{and } \Omega' = -\frac{(2abz)}{\sqrt{\beta} (\sqrt{\beta} + \sqrt{q}) R_b^*(p, q; r) + R_b^*(p, q; r)}
\]
\[
\text{There is a problem here, for at } z = 0, R_b^*(p, q; z^2) \text{ is infinite (though the product } z \cdot R_b^*(p, q; z^2) \text{ is finite): but we evade this by using the identity}
\]
\[
R_b^*(p, q; r) + R_b^*(q, r; p) + R_b^*(r, p; q) = \frac{1}{\sqrt{pq}} \quad (12)
\]

(proved in Greenwood$^4$ (Appendix A3); quoted (in a different notation) in Hills et al.$^5$ (10.52)).

Then equation (11b) can be converted into
\[
\Omega' = -\frac{(2ab)}{\sqrt{\beta} \sqrt{\beta} + \sqrt{q}) R_b^*(p, q; r)}
\]
\[
+ \frac{2ab}{\sqrt{\beta} \sqrt{\beta} + \sqrt{q})}
\]
or combining the first and last terms (using \( \gamma = z^2 \))
\[
\Omega' = (2abz) R_b^*(q, r; p) - \frac{2ab}{\sqrt{\beta} \sqrt{\beta} + \sqrt{q})} \quad (11c)
\]
Table 1. Stresses on the axis of symmetry.

| $\Omega_0$ // $\Omega_1$ // $\Omega_2$ | $\Omega_0'$ // $\Omega_1'$ // $\Omega_2'$ |
|-------------------------------------|-------------------------------------|
| **a dirn**                          | **b dirn**                          |
| 2z $R_0^q(p, q \cdot r)$ - $\frac{1}{\sqrt{p^2 + \sqrt{q^2}}}$ | 2z $R_0^p(q, p \cdot r)$ - $\frac{1}{\sqrt{p^2 + \sqrt{q^2}}}$ |
| 2z $R_0^r(p, q \cdot r)$ - $\frac{1}{\sqrt{q^2 + \sqrt{p^2}}}$ | 2z $R_0^q(q, p \cdot r)$ - $\frac{1}{\sqrt{p^2 + \sqrt{q^2}}}$ |

$\sigma_x/p_0 = -ab/\sqrt{pq}$ where $p = a^2 + z^2$, $q = b^2 + z^2$, $\gamma = z^2$.

To bring out that a simple permutation is all that is needed to give the stress in the $b$-direction, the two results (11a and 11c) are displayed as Table 1 (We have used the permuted $F_s(z) = \frac{4}{\sqrt{4(p^2 + \sqrt{q^2})}}$).

It must be emphasised that these are not the two distinct equations for the stresses along the directions of the major and minor axes: they are the stresses along the “$a$” and “$b$” directions, either of which might be the major semi-axis of the ellipse: **there is really only a single equation**.

**Verification**

Thomas and Hoersch, with the handicap (in 1930) of working from tables of incomplete Legendre functions at 5° intervals, tabulate the stresses on the axis for $\nu = 0.25$ and a number mildly elliptical geometries (appropriate for wheel on rail contacts). Their answers are related to the principal curvatures at the contact, but have been converted to values of $\sigma_x/p_0$ by taking the value they give for $\sigma_x$ at $z = 0$ to be $p_0$. For their moderate case, $k = \sin(35^\circ) = 0.57358$ there is almost perfect agreement with the present calculations (occasional $\pm 1$ in the fourth decimal). Figure 1 shows the (less good) comparison for $k = \sin(20^\circ) = 0.34202$, the markers being the T&H scaled values: the lines the present ones. It is clear that the present equations using Carlson integrals are completely equivalent to the T&H equations using Legendre integrals.

$[b/a = 0.342 : \nu = 0.25]$.

It is well-known that for a circular contact the value of the critical stress on the axis depends on Poisson’s ratio, and T&H’s equations have always allowed an investigation of how this result transfers to elliptical contacts, but seems not to have been done. It is very readily accomplished using the equations above, and Figure 2 shows the result.

Note that we plot the values of the greatest von Mises stress on the axis, defined as $\sqrt{0.5 \cdot [(\sigma_x - \sigma_y)^2 + (\sigma_x - \sigma_z)^2 + (\sigma_y - \sigma_z)^2]}$. The maximum is for a circular contact, and for $\nu = 0.3$ is 0.62$p_0$. Johnson quotes a critical stress of 0.31$p_0$, but this is for the maximum principal shear stress, here equal to $\frac{\sigma_x - \sigma_y}{2}$. For a circular contact where $\sigma_x = \sigma_y$ this is indeed just half our von Mises stress.

**Progress so far**

A single, simple equation has been derived from which both the stresses on the axis $\sigma_x$ and $\sigma_y$ can be found by permuting the variables. These replace the pair of more complicated equations derived by T&H. T&H regard converting their symmetric integrals into Legendre incomplete elliptic integrals as “the authors’ innovation”: and comment “The direct derivation of these results is rather tedious” (it is not given). Perhaps I
should write “the present innovation is converting the integrals into Carlson elliptic integrals” and “The direct derivation of these results is quick and straightforward”!

The full stress distribution on a principal plane

The general equations still hold, so combining (3a) with (8), we get

\[ 2\pi \sigma_x = 2v \frac{\partial \phi}{\partial z} - 2(1 - \nu)z \frac{\partial^2 \phi}{\partial x^2} - (1 - 2\nu)F_z(z) \]  

(12)

where \( F_z(z) = \int_{-\infty}^{\infty} z \frac{\partial^2 \phi}{\partial x \partial y} \, dz \)

provided \( x = 0 \).

In describing the evaluation of \( F_z(z) \) for points on the axis, the actual significance of \( F_z(z) \) has been glossed over. It is a contribution to \( \sigma_z \) from points in the \( y \)-plane (\( x = 0 \)). For points on the \( x \)-plane, we find the out-of-plane stress! The general equations still hold, so combining (3a) with (8), we get

\[ \frac{\partial \phi}{\partial z} = \frac{\partial \phi}{\partial x} \]  

provided \( x = 0 \).

The switch from \( F_z(z) \) to \( F_z(z) \) simply requires the interchange of \( a \) and \( b \), but the evaluation (of either) will now be harder, for again the integrand contains both explicitly and implicitly as \( \sigma_z \) and \( \sigma_x \) on the surface within the contact area, the greatest values are of course on the axis (for a similar geometry but see Appendix for \( d = b \), i.e. for \( x^2 = a^2 - b^2 \)).

To avoid the confusion of denoting the in-plane principal stress as \( \sigma_z \) when on the \( y = 0 \) plane but as \( \sigma_r \) for the \( x = 0 \) plane, we shall (as for a circular contact) use \( \sigma_r \) for the radial, in-plane, stress and \( \sigma_\theta \) for the out-of-plane stress. Thus, equations (11) are for \( \Omega_{\sigma_r} \) and \( \Omega_{\sigma_\theta} \) and give the hoop stress \( \sigma_\theta \).

The sum of the principal stresses is given by equation (4), with which the \( (\sigma/2)ab\theta_0 \) factor included becomes (using (6)):

\[ \sigma_r + \sigma_\theta + \sigma_z = \frac{ab}{4} \frac{\partial \phi}{\partial z} = \frac{ab}{4} \frac{x}{y} \]  

(16)

(Note that on the surface within the contact area, the product \( ab \cdot z R_s^* \) must be replaced by its limiting value \( 4(1 - x^2/a^2) \), the mathematical limit, not just copied from Hertz!).

Then, since for the normal stress \( \sigma_z \), we have (7c)

\[ \frac{\sigma_z}{\sigma_0} = \frac{-ab \frac{x}{y} \frac{d}{d} \frac{1}{b}}{b} \]  

(16)

At greater depths very small (hoop) tensile stresses are found. Note a problem of identification here: on the axis \( x = y = 0 \), which is the hoop stress?}

Figure 3 shows a typical hoop stress distribution. It is indeed much the same as for a circular hertz contact. In the region shown, the hoop stress is entirely compressive and almost entirely confined to the region below the contact area. The greatest values are of course on the axis, and the greatest value is (for \( \nu = 0.3 \)) close to 0.9\( \sigma_0 \). But as is known for a circular contact, and as found by Thomas and Hoersch (and here: see Figure 1), at greater depths very small (hoop) tensile stresses are found. Note a problem of identification here: on the axis \( x = y = 0 \), which is the hoop stress?}

Figure 4 shows the matching radial stress distribution. Some differences from the hoop stress distribution are clear. The radial stresses extend substantially further down than the hoop stresses; but this we have already seen from Figure 1 showing the stresses along the axis (for a similar geometry but \( \nu = 0.25 \)). For \( b/a = 0.3 \), \( \nu = 0.3 \), we find that at \( z = b \) on the axis,
\( \sigma_r = 0.187p_0 \) while the hoop stress has fallen to \( \sigma_\theta = 0.063p_0 \), so this is quite different from a circular contact, where on the axis, the two must be equal.

In contrast, Figure 5 shows that the surface stresses do behave as for a circular contact, with the hoop stress not falling to zero at the contact edge, while the radial stress has become tensile there; and as for a circular contact becoming equal and opposite outside the contact. Figure 4 brings out that the tensile region is not purely at the surface, extending to perhaps 0.2\( \rightarrow b \) (though whether this is a sensible measure, here on the a-axis, is not clear for comparison, we need to examine the stresses on the principal plane along the minor axis.

Not surprisingly, significant hoop stresses exist much further beyond the edge of the contact area, although Figure 7 shows that actually on the surface, the behaviour is the same as for a circular contact. The radial stresses (Figure 6), somewhat to my surprise, show the same tendency more strongly.

Figure 7 compares the surface stresses along the major and minor axes. We see that they are what one might expect from a knowledge of the pattern for a circular contact. Here they decay from the values at the origin \( -\sigma_r/p_0 = a/(a + b) + 2\nu b/(a + b); -\sigma_\theta/p_0 = b/(a + b) + 2\nu a/(a + b) \) to the known values at the ends of the axis\(^2 \) (CM 3.68a), followed by a state of pure shear \( \sigma_r = -\sigma_\theta \). The surface stresses outside the contact area decay, but the variation \( (1 - 2\nu) ab/3 \pi^2 \) as for a circular contact is found only when \( b = a \) or \( r \) is large; otherwise we have (for \( x > a \))

\[
\frac{\sigma_r}{p_0} = \frac{ab(1-2\nu)}{a^2-b^2} \left\{ \left( \frac{x}{e} \right) \tanh^{-1}\left( \frac{e}{x} \right) - 1 \right\} \quad \text{or} \quad \frac{\sigma_\theta}{p_0} = \frac{ab(1-2\nu)}{a^2-b^2} \left\{ \left( \frac{x}{e} \right) \tan^{-1}\left( \frac{e}{x} \right) - 1 \right\}
\]

(\( e = \sqrt{a^2 - b^2} \)) depending on whether \( b < a \) or \( b > a \).

Note that in order to display the comparison, the “major” and “minor” axis results were obtained by setting \( a = 1 \) and taking “b” successively as 1/3 and 3; the different decay rates of the stresses along the two axes...
is partly due to this scaling. Note also that the equations (3.68a,b) in Hills et al.$^6$ are incorrect: the arguments of the tanh$^{-1}$ and tan$^{-1}$ terms should have been $ae/x$ and $b^2/ae$y, not as printed.

Discussion

Equations for the full stress distribution are not new; indeed Sackfield and Hills,$^8$ “S&H,” give equations for the complete subsurface stress distribution, not only, as here, for the stresses on the principal planes.

But they give no derivation. (The outline in HNS$^6$ (chapter 12) suggests that $\chi$ was obtained by field point integration over the ellipse as described by Barber,$^9$ but while this could certainly be used to obtain $\phi$, it is not clear how it could be used when the kernel $1/R$ is replaced by $\ln(R + z)$. S&H give a simpler form for the stresses on the principal plane of the major axis, but do not evaluate their equations, possibly dissuaded by the existence of the “tails,” although we would claim that our tail is the simpler.

For the S&H solution (and ours) depends on the Thomas & Hoersch integrals $I_1$, $I_2$, $I_3$, declared by S&H to be elliptic integrals, as indeed they are. Thus

$$I_2 = \int_0^\infty \frac{dt}{\gamma (a^2 + t)^{1/2} (b^2 + t)^{1/2}} = 2 K_0^*(\gamma, a^2 + \gamma; b^2 + \gamma);$$

$$I_3 = \int_0^\infty \frac{dt}{\gamma (a^2 + t)^{1/2} (b^2 + t)^{1/2}} = 2 K_0^*(a^2 + \gamma, b^2 + \gamma; \gamma)$$

and $I_1$ is just the third permutation of the same three parameters. The Carlson forms above are immediately found (just set $t = w + \gamma$) from the integrals in the form in which they first appear in T&H’s analysis.

But $I_1$, $I_2$, $I_3$ are not “Legendre elliptic integrals”: while $I_1$ is indeed just the difference of two Legendre integrals, $I_2$ and $I_3$ are each the sum of one or two Legendre integral integrals and algebraic terms. To derive them T&H begin by setting $t = a^2 \sigma^2$: then (dropping a factor $2a^{-1}) I_2 = \int_0^\infty \frac{d\sigma}{(1 + \sigma^2)^{1/2}(k^2 + \sigma^2)^{1/2}}$;

$$I_3 = \int_0^\infty \frac{d\sigma}{(1 + \sigma^2)^{1/2}(k^2 + \sigma^2)^{1/2}} = \frac{E(\varphi, k')}{k^2} \left(1 + s^2\right)$$

(relation to the Hertz problem is now lost). They then find (laboriously)

$$I_2 = \frac{1}{k^2} \left[\frac{E(\varphi, k')}{k^2} = F(\varphi, k') - 1 \frac{\left(k^2 + s^2\right)^{1/2}}{k^2} + 1 \frac{\left(1 + s^2\right)^{1/2}}{k^2}\right].$$

$$I_3 = \frac{1}{k^2} \frac{\left(k^2 + s^2\right)^{1/2}}{1 + s^2} \frac{E(\varphi, k')}{k^2}$$

two apparently unrelated equations, which cannot immediately be checked.

Unfortunately, the elliptic integrals are only a part of the solution, and are accompanied by a “tail”: our $F_1(z)$ or S&H’s $I_4$, which are complicated “elementary” quantities. It is not clear how the very different expressions are related, but numerical tests (but using our forms for $I_2$, $I_3$) establish that the answers are the same (A different form of the Sackfield and Hills equations for the full stress distribution is given by Hills et al.$^6$, from which the tail for the stresses on the $y = 0$ plane may be abstracted. But our attempt to evaluate this gave impossible answers). So equations for the subsurface stresses have been available since 1983. What is new, is that the derivation of our equations is given; and that for the first time the stresses have been evaluated and plotted, so that the (minor) distortion of the stresses under a circular contact by the ellipticity may be seen and evaluated. The undoubted easing of the calculation of the elliptic integrals is somewhat offset by the existence of the “tails,” although we would claim that our tail is the simpler.

And indeed, with a computer program capable of complex arithmetic, the need to distinguish between $a > b$ and $b > a$ finally disappears: we found (using MATLAB) that the single equation (A5) does produce the same answers as the alternative two real equations (A6) and (A5b).

Conclusion

The simple, direct, link between the equations for the stresses underneath an elliptical Hertzian contact and (readily calculable) Carlson elliptic integrals has been used to obtain equations for the stresses in the principal planes, while avoiding the distinction between major and minor axes. The stresses along the axis of symmetry, well known from the work of Thomas and Hoersch,$^5$ are now given by a single (simpler) equation, in which “a” and “b” appear on an equal basis. Away from the axis of symmetry, the equations for the stresses involve both elliptic integrals and a long “tail” of
complicated elementary functions. The considerable gain from using the immediately obvious, neatly matched, Carlson integrals instead of the rather messy transformation into non-matching incomplete Legendre integrals is diluted by the existence of the tails, though we would argue that ours is the simpler. The final numerical answers are the same as those obtained from the equations given in 1983 by Sackfield and Hills, now for the first time evaluated: but our equations are complete with a derivation.

Maps of the principal stresses \( \sigma_{r} \) and \( \sigma_{b} \) for a typical elliptical contact \( k = 1/3 \) are presented. These are much as might be guessed from a knowledge of the stresses under a circular contact, but some differences are noted, particularly when the stresses in the plane of the minor axis are compared to those in the plane of the major axis.

Sadly, the interchangeability of \( a \) and \( b \) in the elliptic integrals is not maintained in the tails.

The tail for the stresses in the plane of the major principal axis contains a term in \( \arctanh \) (or a logarithm): the corresponding term in the tail for stresses in the plane of the minor axis contains a term in \( \arctan \). But this was obtained from the equations given in 1983 by Sackfield and Hills, now for the first time evaluated: but our equations are complete with a derivation.

Appendix 1

**Evaluation of \( F_{y}(z) \)**

\[
F_{y}(z) = 2 \int_{0}^{\infty} \frac{(d^{2} + t)}{(a^{2} + t)(b^{2} + t)(d^{2} + t)} \, dt
\]

Changing the integration variable from \( y \) to \( t \) where \( d^{2} = a^{2} - x^{2} \) and \( y \) is the positive root of the quadratic \( y^{2} + \gamma(a^{2} - x^{2} - z^{2}) - d^{2}z^{2} = 0 \). Although we have used the symbol \( d^{2} \) (in the interests of dimensional homogeneity), it will often be negative: this causes no problem as we never need \( d \) itself.

It is easily confirmed that the \( \gamma(z) \) relation is monotonic, so that the original integration wrt \( z \) can properly be turned into integration wrt \( \gamma \) (or \( t \)).

Note that for \( x < a \) the curves start at \( \gamma|_{z=0} = 0 \); but for \( x > a \) at \( \gamma|_{z=0} = x^{2} - a^{2} = -d^{2} \).

We start the evaluation by bringing the surd down to the denominator and then splitting the integrand in two:

\[
F_{y}(z) = 2 \int_{0}^{\infty} \frac{(d^{2} + t)}{(a^{2} + t)(b^{2} + t)(d^{2} + t)} \, dt
\]

\[
\text{and} \quad a^{2} - d^{2} = x^{2} - ;
\]

\[
F_{y}(z) = \frac{2}{a^{2} - b^{2}} \left[ \int_{0}^{\infty} \frac{(d^{2} - b^{2})}{(b^{2} + t)^{3/2}} \, dt + \int_{0}^{\infty} \frac{x^{2}}{(a^{2} + t)(b^{2} + t)(d^{2} + t)} \, dt \right]
\]

Both integrals are elementary: we have

\[
\frac{d}{dt}[(d^{2} + t)^{1/2}(b^{2} + t)^{-1/2}] = \frac{1}{2}[(d^{2} + t)^{-1/2}(b^{2} + t)^{-1/2} - (d^{2} + t)^{1/2}(b^{2} + t)^{-3/2}] = \frac{1}{2}[(d^{2} + t)^{-1/2}(b^{2} + t)^{-3/2}] \cdot [(b^{2} + t) - (d^{2} + t)] = \frac{(b^{2} - d^{2})}{2}[(d^{2} + t)^{-1/2}(b^{2} + t)^{-3/2}] .
\]
so the first integral (courtesy of T&H) equals

\[
-2 \left[ \frac{d^2 + \gamma}{b^2 + \gamma} \right] = 2 \left[ \frac{d^2 + \gamma}{b^2 + \gamma} - 1 \right] \quad (A3)
\]

Note that when \( x^2 > a^2 \) the least value of \( \gamma \) (at \( z = 0 \)) is \( x^2 - a^2 \), so at the worst \( \sqrt{d^2 + \gamma} \) is zero: it can never be imaginary.

Returning to (A2), the second integral is

\[
J = \int_0^\infty (x^2 + 2x \cdot \sqrt{b^2 + \gamma} + x) \, dt
\]

This (courtesy of Prof. J R Barber) is equal to

\[
J = \frac{2}{\sqrt{a^2 - d^2}(a^2 - b^2)} \times \ln \left[ \frac{(x + e) \sqrt{a^2 + \gamma} + e \sqrt{d^2 + \gamma}}{x \sqrt{b^2 + \gamma} + e \sqrt{d^2 + \gamma}} \right] \quad (A4)
\]

as again may be verified by differentiating wrt \( y \).

More conveniently, recalling that \( \sqrt{a^2 - d^2} = x \) and defining \( \sqrt{a^2 - b^2} = e \)

\[
J = \frac{2}{x \cdot e} \ln \left[ \frac{(x + e) \sqrt{a^2 + \gamma} + e \sqrt{d^2 + \gamma}}{x \sqrt{b^2 + \gamma} + e \sqrt{d^2 + \gamma}} \right] \quad (A5)
\]

To obtain the corresponding result for the minor axis

we need to take \( b > a \) so we have

\[
J = \frac{2}{x \cdot e} \ln \left[ \frac{(x + e) \sqrt{a^2 + \gamma} + e \sqrt{d^2 + \gamma}}{x \sqrt{b^2 + \gamma} + e \sqrt{d^2 + \gamma}} \right] \quad (A5b)
\]

Numerical evaluation showed that the two gave identical answers, except that for \( d = b \) (which implies that \( e = x \)), (A5b) reduces to the unhelpful \( J = \frac{\pi}{2} \ln \left( \frac{a^2 + \gamma}{b^2 + \gamma} \right) \) instead of the correct \( J = \frac{\pi}{2} \ln \left( \frac{a^2 + \gamma}{b^2 + \gamma} \right) \).

### Appendix 2 (crib)

\( I_2 = \int_0^\infty \frac{dt}{\gamma (a^2 + \gamma)^{3/2}(b^2 + \gamma)^{1/2}} \)

\[
= 2 R_0^\gamma(p, \ a^2 + \gamma; \ b^2 + \gamma)
\]

\( I_3 = \int_0^\infty \frac{dt}{\gamma (a^2 + \gamma)^{3/2}(b^2 + \gamma)^{1/2}} \)

\[
= 2 R_0^\gamma(a^2 + \gamma, \ b^2 + \gamma; \ y)
\]

For \( y = 0 \) plane

\( (1) \) Find: \( \gamma = \frac{e^2}{\sigma_0 + r} + \frac{e^2}{\sigma_0 + r} = 1 \)

\( (2) \) Calculate: \( p = \frac{a^2 + \gamma}{\sigma_0 + r}; q = \frac{b^2 + \gamma}{\sigma_0 + r}; r = \frac{e^2 + \gamma}{\sigma_0 + r}; s = \frac{d^2 + \gamma}{\sigma_0 + r}; \) where \( d^2 = a^2 - x^2 \)

\( (3) \) Then \( \frac{a^2}{\sigma_0 + r} = \frac{\cos \theta}{\sqrt{p + q}} \quad (A5b) \)

\( (4) \) Calculate the elliptic integrals \( I_2 = 2 R_0^\gamma(p, \ q; \ r); \quad I_3 = 2 R_0^\gamma(p, \ q; \ y) \)

\( (5) \) Find:

\[
(\sigma_0 + \sigma_0 + \sigma_0)p_0 = -ab (1 + n) \cdot z R_0^\gamma(p, \ q; \ r) - n \cdot z I_3
\]

\( (6) \) Find:

\[
F_s(z) = \frac{2 \cdot (a^2 - b^2)}{2 \sqrt{b^2 - a^2}} \cdot \left[ \frac{2 \sqrt{b^2 - a^2}}{\sqrt{b^2 - a^2}} \right] + x^2 \]

where, defining \( e^2 = \frac{(x + e)^2}{\sqrt{b^2 + \gamma}} \),

for \( a > b \quad J = \frac{2}{x \cdot e} \ln \left[ \frac{(x + e) \sqrt{a^2 + \gamma} + e \sqrt{d^2 + \gamma}}{x \sqrt{b^2 + \gamma} + e \sqrt{d^2 + \gamma}} \right] \)

but for \( b > a \quad J = \frac{2}{x \cdot e} \ln \left[ \frac{(x + e) \sqrt{a^2 + \gamma} + e \sqrt{d^2 + \gamma}}{x \sqrt{b^2 + \gamma} + e \sqrt{d^2 + \gamma}} \right] \)

\( (7) \) Find: \( \sigma_\gamma/p_0 = \Omega_\gamma + \nu \cdot \Omega_{\gamma} \), where

\[
\Omega_\gamma = - \frac{(ab)}{z} (I_2 - 0.25 F_s(z))
\]

\[
\Omega_{\gamma} = (ab) \left[ -4z R_0^\gamma(p, \ q; \ r) - 4z R_0^\gamma(p, \ q; \ r) + F_s(z) \right]
\]

\( (8) \) Finally

\[
\sigma_\gamma/p_0 = \frac{(\sigma_0 + \sigma_0 + \sigma_0)}{p_0} - \sigma_\gamma/p_0 - \sigma_0/p_0
\]
Limiting values as $z \to 0$

$z^2/\gamma \to 1 - x^2/a^2$; $\gamma H \to 1 - x^2/a^2$

hence $\sigma_z/p_0 \to \sqrt{1 - x^2/a^2}$

$z R_0^* (p, q; \gamma) \sim \frac{z}{\sqrt{pq\gamma}} \sim \frac{1}{ab\sqrt{\gamma}} \to \frac{1}{ab} \sqrt{1 - x^2/a^2}$