Rolling Contact Phenomena: Linear Elasticity

J.J. Kalker*

*Department of Applied Mathematical Analysis, Delft University of Technology, The Netherlands

Abstract. In this paper, we treat the rolling contact phenomena of linear elasticity, with special emphasis on the elastic half-space.
Section 1 treats the basics; rolling is defined, the distance between the deformable bodies is calculated, the slip velocity between the bodies is defined and calculated; a very brief recapitulation of the theory of elasticity follows, and the boundary conditions are formulated.
Section 2 treats the half-space approximation. The formulae of Boussinesq-Cerruti are given, and the concept of quasiidentity is introduced. Then follows a brief description of the linear theory of rolling contact for Hertzian contacts, with numerical results, and of the theory of Vermeulen-Johnson for steady-state rolling. Finally, some examples are given.
Section 3 is devoted to the simplified theory of rolling contact.
In Section 4, the variational, or weak theory of contact is considered. First, we set up the virtual work inequality, and it is shown that it is implied by the boundary conditions of contact. Then the complementary virtual work inequality is postulated, and it is shown that it implies the boundary conditions of contact. Elasticity is introduced into both inequalities, and the potential energy and the complementary energy follow. Finally, surface mechanical principles are derived.
In Section 5, we return to the exact half-space theory. The problem is discretized, and solved by means of the CONTACT algorithm. Finally, results are shown in Section 6.

1 Basics

In this Section, rolling is defined. Also, the distance between the bodies is calculated, the slip velocity between the bodies is defined and computed; a very brief recapitulation of the theory of elasticity follows, and the boundary conditions are formulated.

1.1 Definition of rolling

Consider two bodies of revolution. They are distinguished from each other by attaching to them the numbers 1 and 2. They are, for the time being considered as rigid. They

are pressed together so that they touch in a point; line contact will be treated presently. They are rotated, so that the contact point moves over the bodies. Then there are two possibilities: either the velocity $\mathbf{v}^1$ of the contact point over body 1 equals the velocity $\mathbf{v}^2$ of the contact point over body 2, or this is not so. In the former case (equal velocities) one speaks of *rolling*, in the latter case one speaks of *sliding* or *rolling with sliding*.
We consider the case that the bodies contact each other along a line. Again, the bodies are rotated, and the line moves over the bodies. Rolling occurs when the velocities of the contacting line over the bodies are equal at each point of the line, otherwise we speak of sliding or rolling with sliding.
We now consider that the bodies are deformable. First we have to define what we mean by that. We assume that each body is made up of particles that are glued together to form a continuum. Stresses and strains may be present in such a body. We count the displacement from an unstressed state in a manner which will be explained presently. The bodies are pressed together so that a contact patch forms between them, and they are rotated. When the velocity $\mathbf{v}^1$ of the contact patch over body 1 almost equals the velocity $\mathbf{v}^2$ of the contact patch over body 2 we speak of *rolling*, otherwise of *sliding* or *rolling with sliding*:

$$|\mathbf{v}^1 - \mathbf{v}^2| \ll |\mathbf{v}^1 + \mathbf{v}^2| \tag{1}$$

(rolling, otherwise rolling with sliding).
As we said, we count the displacement from the unstressed state, that is, the state in which there are no stresses acting in the bodies. A local Cartesian coordinate system is attached to each body. Each particle corresponds to a point $\mathbf{y}$ of the local coordinate system. The body is deformed; the displacement of the point $\mathbf{y}$ is denoted by $\mathbf{w}$, function of $\mathbf{y}$ and the time t. In the deformed state the particle lies in

$$\mathbf{y} + \mathbf{w}, \qquad \mathbf{w} = \mathbf{w}(\mathbf{y}, t) \tag{2}$$

We assume $\mathbf{w}$ small as well as its gradients:

$$|\mathbf{w}| \ll |\mathbf{y}|; \qquad |\partial \mathbf{w}/\partial \mathbf{y}| \ll 1 \tag{3}$$

We assume that inertia terms may be neglected and that the deformation is elastic: then we arrive at a linearly elastostatic theory.
We have two bodies; then we call the position $\mathbf{y}^a$ in the a-th coordinate system, and the displacement $\mathbf{w}^a$.
We want to compare the quantities of the bodies. To that end, we must refer them to a single coordinate system. So we introduce a third, global coordinate system in which the particles in the undeformed state are given by $\mathbf{x}^a$, and in the deformed state by $\mathbf{x}^a + \mathbf{u}^a$, that is, in the global coordinate system we also distinguish between the particles of body 1 and body 2.
The global coordinate system is connected to the two local systems by rotation matrices $A(t)$ and the distance between the origins $\mathbf{R}(t)$ which are functions of the time. The rotation matrices are orthogonal matrices. Their columns are denoted by $\mathbf{n}^a, \mathbf{t}^a, \mathbf{b}^a$. Let $(1,0,0)^T$ be given in the local system; then $\mathbf{n}^a$ is that same vector in the global system. Similarly, $\mathbf{t}^a$ is the global representation of the vector $(0,1,0)^T$ in the local system, and $\mathbf{b}^a$ is the global representation of $(0,0,1)^T$. It is easy to see that $\mathbf{n}^a, \mathbf{t}^a, \mathbf{b}^a$ indeed form

an orthonormal system.
The following connection exists between the $\mathbf{y}^a + \mathbf{w}^a$ and the $\mathbf{x}^a + \mathbf{u}^a$:

$$\mathbf{x}^a + \mathbf{u}^a \;=\; A^a(\mathbf{y}^a + \mathbf{R}^a) + A^a\mathbf{w}^a \tag{4}$$

We identify

$$\mathbf{x}^a \;=\; (A^a\mathbf{y}^a + \mathbf{R}^a) \qquad \Rightarrow \qquad \mathbf{y}^a \;=\; A^{aT}\,(\mathbf{x}^a - \mathbf{R}^a) \tag{5}$$

$$\mathbf{u}^a \;=\; A^a\mathbf{w}^a(\mathbf{y}^a, t) \qquad \Rightarrow \qquad \mathbf{w}^a \;=\; A^a\mathbf{u}^a \tag{6}$$

In particular, (5) gives the global coordinate system in its dependence on the local coordinate system, and vice versa, and (6) defines the global displacement $\mathbf{u}^a$ in terms of the local displacement $\mathbf{w}^a$, and vice versa.

It is clear from (5) that a variable belonging to body a can be written as a function of $\mathbf{x}^a$ and t or of $\mathbf{y}^a$ and t.

Let $\mathbf{m}^a$ be the outer normal on body a at $\mathbf{y}^a$; since $\mathbf{w}^a$ is small with small gradients, $\mathbf{m}^a$ is also the normal on the deformed body a at $\mathbf{y}^a + \mathbf{w}^a$. To see this, we assume that the surface of the body is given by $F(\mathbf{y}) = 0$. Let $\mathbf{z} = \mathbf{y} + \mathbf{w}$ be the position of $\mathbf{y}$ in the deformed state; that is, $\mathbf{y} = \mathbf{z} - \mathbf{w}$. The deformed surface is then given by $F(\mathbf{z} - \mathbf{w}) = 0$. The normal $\mathbf{m}$ on the undeformed body at $\mathbf{y}$ is given by

$$\mathbf{m} = \partial F/\partial \mathbf{y} \tag{7}$$

The normal on the deformed body is

$\mathbf{m}' = \partial F/\partial \mathbf{z} = (\partial F/\partial \mathbf{y})(\partial \mathbf{y}/\partial \mathbf{z})$.

But

$\partial \mathbf{y}/\partial \mathbf{z} = I - \partial \mathbf{w}/\partial \mathbf{z} \approx I$,

where I is the 3×3 unit matrix. From this proposition it follows that $\mathbf{m} \approx \mathbf{m}'$. In the global coordinates, $\mathbf{m}^a$ becomes $\mathbf{n}^a$, with

$$\mathbf{n}^a = A^a\mathbf{m}^a \tag{8}$$

The points $\mathbf{y}^1 + \mathbf{w}^1$ and $\mathbf{y}^2 + \mathbf{w}^2$ are in contact with one another when

$$\mathbf{x}^1 + \mathbf{u}^1 = \mathbf{x}^2 + \mathbf{u}^2, \;\text{ and }\; \mathbf{n}^1 = -\mathbf{n}^2 \tag{9}$$

Example: A rail surface consists of the union of a number of circular cylinders with parallel axes that touch one another. That is, the cylinders intersect and at the intersection the first derivatives are continuous.

The equation of one circular cylinder is, in a Cartesian coordinate system (O;x,y,z)

$$(x-a)^2 + (z-b)^2 - R^2 = 0, \quad \mathbf{n} = q(2(x-a), 0, 2(z-b))$$

with the scalar q chosen so that $\mathbf{n}$ is a unit vector; then $\mathbf{n}$ is the unit normal on the rail.

1.2 The Distance

We assume that the bodies are not in contact at the point $\mathbf{x}$, but that they are $O(\mathbf{u})$ apart, and that the surface is smooth, all in the neighborhood of the point $\mathbf{x}$, while the relation $\mathbf{n}^1 = \mathbf{n}^2$ is approximately valid. In a small neighborhood of $\mathbf{x}$ the bodies may be visualized as two parallel slabs, see Fig. 1.

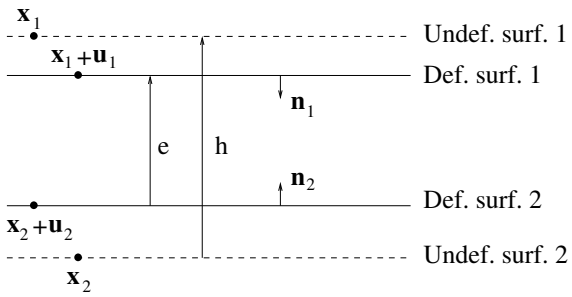


Figure 1: The Distance

Let the distance between the slabs, (2) to (1), be h in the unstressed state, and e in the deformed state. Then we have

$$h = \mathbf{n}^{2T}(\mathbf{x}^1 - \mathbf{x}^2), \quad e = \mathbf{n}^{2T}\{(\mathbf{x}^1 + \mathbf{u}^1) - (\mathbf{x}^2 + \mathbf{u}^2)\} \tag{10}$$

We simplify the expression for e. As we will see later on, the displacement occurs in a linearized contact problem such as we have here (small displacements, small displacement gradients) only in the form $\mathbf{u}^1 - \mathbf{u}^2$. We call this the *displacement difference*, and we denote it by $\mathbf{u}$. Introduction of the displacement difference into the expression for e, we find

$$e = h + \mathbf{n}^{2T}\mathbf{u} = h - \mathbf{n}^{1T}\mathbf{u} \tag{11}$$

e: deformed, h: undeformed distance, (2) to (1).
We analyze the deformed distance e.

- $e > 0$: there is a gap between the bodies at $\mathbf{x}^1$, $\mathbf{x}^2$; $\mathbf{x} = (\mathbf{x}^1 + \mathbf{x}^2)/2$
- $e = 0$: the bodies are in contact at $\mathbf{x}$;
- $e < 0$: the bodies penetrate at $\mathbf{x}$. Impossible.

In sum, only $e(\mathbf{x}) \geq 0$ is possible.

In addition to the distance there is the play of forces. Contact tractions $\mathbf{p}^1$, $\mathbf{p}^2$ (dimension: N/m^2) act on bodies (1) and (2). According to Newton's Third Law,

$$\mathbf{p}^1 = -\mathbf{p}^2 \doteq \mathbf{p} \tag{12}$$

Then the normally directed traction on (1) at $\mathbf{x}$, positive if compressive, is given by

$$p_N = \mathbf{n}^{2T}(\mathbf{x})\mathbf{p} \qquad \text{Normal component} \tag{13}$$

and the tangential traction is given by

$$\mathbf{p}_T = \mathbf{p} - \mathbf{n}^2 p_N(\mathbf{x}) \quad \text{Tangential component} \tag{14}$$

When the bodies do not attract each other, the normal component of the traction vanishes outside contact and is positive (=compressive) inside, while the deformed distance is positive outside contact and vanishes inside. We can summarize this as follows:

$$e(\mathbf{x}) \geq 0: \quad \text{either a gap or contact;} \tag{15}$$

$$p_N(\mathbf{x}) \geq 0: \quad \text{either no, or a compressive normal traction;} \tag{16}$$

$$e p_N = 0 \tag{17}$$

In words, in contact, p_N may be positive, and $e = 0$;
outside contact, the normal traction vanishes, and $e > 0$.

We formalize this as follows. We choose a potential contact zone (also called the potential contact area, pot.con.) which is such that

- The potential contact zone encompasses the real contact zone completely;
- In the potential contact zone $p_N \geq 0$, $e \geq 0$, $p_N e = 0$.

Note that the pot.con. may be chosen freely, as long as the above is satisfied.

Example: Consider a so-called Winkler bedding, i.e. a rigid flat plate upon which are mounted springs. The springs are tangentially unconnected; they are equally long and have the same positive spring constant per unit area $k\ m^3/N$ in the normal direction:

$\mathbf{n}^{2T}\mathbf{u}^1 = kp_N$, $\mathbf{n}^{2T}\mathbf{u}^2 = -kp_N$, with $k > 0$
$\Rightarrow e = h + \mathbf{n}^{2T}\mathbf{u} = h + \mathbf{n}^{2T}(\mathbf{u}^1 - \mathbf{u}^2) = h + 2kp_N$

Suppose that $h > 0$. Since $p_N \geq 0$ we must have that $e \geq h > 0$, and hence $p_N = 0$. Then $e = h > 0$, and $\mathbf{n}^{2T}\mathbf{u}^a = 0$, $a = 1, 2$.
Suppose that $h < 0$. Since $e \geq 0$ we must have that $2kp_N = e - h > 0 \Rightarrow e = 0$ and $2kp_N = -h > 0$. Hence $\mathbf{n}^{2T}\mathbf{u}^1 = -\mathbf{n}^{2T}\mathbf{u}^2$. The solution is shown in Fig. 2.

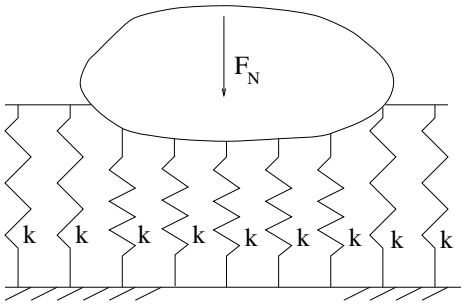


Figure 2: Compression of a Winkler bedding (normal contact)

It is seen that the inequalities play an crucial role in determining the elastic field. This is typical of a contact problem.

1.3 The Slip

We determine the relative velocity of a particle of body 1 with respect to body 2.

Suppose that the particle $\mathbf{y}^1$ is in contact with the particle $\mathbf{y}^2$ at the time t, that is

$$\mathbf{x}^1(t) + \mathbf{u}^1(\mathbf{x}^1, t) = \mathbf{x}^2(t) + \mathbf{u}^2(\mathbf{x}^2, t)$$

If we denote time differentiation by a high dot, the velocity of particle 1 with respect to the global coordinate system is

$$\mathbf{v}^1 = \dot{\mathbf{x}}^1(t) + \dot{\mathbf{u}}^1(\mathbf{x}^1, t) = \dot{\mathbf{x}}^1(t) + (\partial \mathbf{u}^1/\partial \mathbf{x}^1)\dot{\mathbf{x}}^1 + \partial \mathbf{u}^1/\partial t \tag{18}$$

$\mathbf{v}^2$ is similarly defined.

The slip $\mathbf{s}$ is the relative velocity of two particles in contact, that is,

$$\mathbf{s} = \mathbf{v}^1 - \mathbf{v}^2 = \{\dot{\mathbf{x}}^1(t) - \dot{\mathbf{x}}^2(t)\} + \{\dot{\mathbf{u}}^1(\mathbf{x}^1, t) - \dot{\mathbf{u}}^2(\mathbf{x}^2, t)\} \tag{19}$$

Now,

$$\mathbf{x}^1 = \mathbf{x}^2 + \mathbf{u}^2 - \mathbf{u}^2 \quad \Rightarrow \quad \mathbf{x}^1 \approx \mathbf{x}^2 \approx \mathbf{x} \doteq (\mathbf{x}^1 + \mathbf{x}^2)/2 \tag{20}$$

so that, owing to the smallness of $\mathbf{u}^a$, and if $\dot{\mathbf{x}}^1 \approx \dot{\mathbf{x}}^2$

$$\mathbf{s} = \dot{\mathbf{x}}^1 - \dot{\mathbf{x}}^2 + (\partial \mathbf{u}^1/\partial \mathbf{x} - \partial \mathbf{u}^2/\partial \mathbf{x})\dot{\mathbf{x}} + \partial \mathbf{u}^1/\partial t - \partial \mathbf{u}^2/\partial t \tag{21}$$

If $\dot{\mathbf{x}}^1$ is not approximately equal to $\dot{\mathbf{x}}_2$, then $(\dot{\mathbf{x}}^1 - \dot{\mathbf{x}}^2)$ is large with respect to $\{(\partial \mathbf{u}^a/\partial \mathbf{x}^a)\dot{\mathbf{x}}^a + \partial \mathbf{u}^a/\partial t\}$, and $\mathbf{s}$ is also given by (21), albeit that the second and third terms are negligible with respect to the first.

We call

$$\mathbf{c} = \dot{\mathbf{x}}^1 - \dot{\mathbf{x}}^2 \qquad \text{creep} \tag{22}$$

$$\mathbf{v} = -\dot{\mathbf{x}} \qquad \text{rolling velocity} \tag{23}$$

$$\mathbf{u} = \mathbf{u}^1 - \mathbf{u}^2 \qquad \text{displacement difference} \tag{24}$$

$$\mathbf{s} = \mathbf{c} - (\partial \mathbf{u}/\partial \mathbf{x})\mathbf{v} + \partial \mathbf{u}/\partial t \qquad \text{slip} \tag{25}$$

The minus sign in the definition of the rolling velocity calls for comment, see Fig. 3:

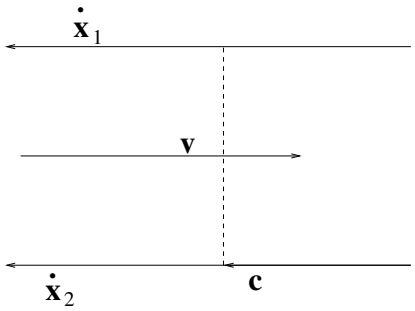


Figure 3: The definitions of creep and rolling velocity

Place the origin of the global coordinate system temporarily in $\mathbf{x}$. Then $\dot{\mathbf{x}}$ is the velocity of the particles near $\mathbf{x}$, and it is seen that the material flows backwards through the coordinate system, counter to the rolling direction. Hence the minus sign in the definition of the rolling velocity.

If a coordinate system may be found in which all quantities $\mathbf{c}$, $\mathbf{u}$, $\mathbf{v}$ are independent of the time t, one speaks of steady state rolling, otherwise of non-steady state rolling:

$$\mathbf{s} = \mathbf{c} - (\partial \mathbf{u}/\partial \mathbf{x})\mathbf{v} \qquad \text{steady state rolling} \tag{26}$$

$$\mathbf{s} = \mathbf{c} - (\partial \mathbf{u}/\partial \mathbf{x})\mathbf{v} + \partial \mathbf{u}/\partial t \qquad \text{non-steady state (or transient) rolling} \tag{27}$$

We finish this section on the slip by analyzing the creep for bodies of revolution that are rotated about their axes which are almost in the same plane. A number of interesting technological problems fall into this category.

1. Problems in which the contact area is almost flat. Examples:
   (a) A ball rolling over a plane;
   (b) An offset printing press: contact short in the rolling direction;
   (c) An automotive wheel rolling over a road.
2. Problems with contact short in the rolling direction, curved in the lateral. Examples:

(a) A railway wheel rolling over a rail;

(b) A ball rolling in a deep groove, as in ball bearings.

3. Problems in which the contact area is curved in the rolling direction, and conforming in the lateral direction. Example: a pin rolling in a hole.

1.4 Leading Edge, Trailing Edge

The contact region C has an edge. This edge consists of three parts: the leading edge, the trailing edge, and, possibly, a neutral edge. When the rolling velocity points outside the contact area at a point of the edge, this points belongs to the LEADING EDGE: particles move into the contact area with rolling velocity.

When the rolling velocity points into the contact region C at a point of the edge, the point belongs to the TRAILING EDGE: and particles leave the contact area at macroscopic rolling velocity.

When the rolling velocity vanishes, or is parallel to the edge, we speak of a NEUTRAL EDGE: particles move only into or outside the contact area through the elastic deformation, not through the rigid body motion.

1.5 Example

We give an example of a simplified wheel-rail system. see Fig. 4.

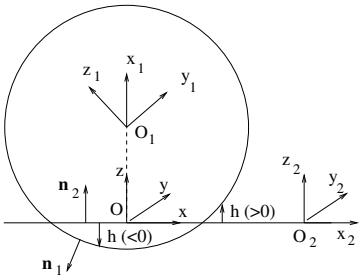


Figure 4: A simplified wheel-rail system

The origin O^1 of the wheel lies on the axis of the cylinder (body 1). The origin O^2 of the rail (body 2) lies on its surface. The origin O of the global coordinate system lies on the surface of the rail, perpendicularly below O^1. It is contact fixed. The y-axis of the wheel coincides with the axis of the cylinder; it is the rotation axis of the wheel. The y-axis of the global system coincides with the projection of the y-axis of the rail. The x-axis is in the same sense as the rolling direction. The z-axis of the global system points vertically upwards. The sense of the y-axis is so that the global system is right-handed. The wheel rotates about its y-axis, clockwise, so that the wheel rolls over the rail to the

right.

We have

$\mathbf{x}^a = A^a \mathbf{y}^a + \mathbf{R}^a \quad \longleftrightarrow \quad \mathbf{y}^a = A^{aT}(\mathbf{x}^a - \mathbf{R}^a)$.
$\mathbf{x}^a$ is the global coordinate of the particle $\mathbf{y}^a$. $\mathbf{R}^a$ is the position of O^a in the global coordinates. The particle-fixed velocity at $\mathbf{x}$ is

$\dot{\mathbf{x}}^a = \dot{A}^a \mathbf{y}^a + \dot{\mathbf{R}}^a = \dot{A}^a A^{aT}(\mathbf{x}^a - \mathbf{R}^a) + \dot{\mathbf{R}}^a$.
This means: substitute $\mathbf{x}^a$,$\mathbf{R}^a$,$\dot{\mathbf{R}}^a$ in the right-hand side, and you find $\dot{\mathbf{x}}^a$ without calculating $\mathbf{y}^a$.

Indeed we have:

BODY 1

$$\mathbf{R}^1 = (0,0,r)^T \quad \Longrightarrow \quad \dot{\mathbf{R}}^1 = 0$$

$$A^1 = \begin{pmatrix} +\cos wt & 0 & +\sin wt \\ 0 & 1 & 0 \\ -\sin wt & 0 & +\cos wt \end{pmatrix};$$

$$\dot{A}^1 = w \begin{pmatrix} -\sin wt & 0 & +\cos wt \\ 0 & 0 & 0 \\ -\cos wt & 0 & -\sin wt \end{pmatrix};$$

$$\dot{A}^1 A^{1T} = w \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix};$$

$$\dot{\mathbf{x}}^1 = \dot{A}^1 A^{1T}(\mathbf{x}^1 - \mathbf{R}^1) = \begin{pmatrix} wz - wr \\ 0 \\ -wx \end{pmatrix}.$$

BODY 2

$$\mathbf{x}^2 = \mathbf{y}^2 + \begin{pmatrix} -vt \\ 0 \\ 0 \end{pmatrix} \Longrightarrow \dot{\mathbf{x}}^2 = \begin{pmatrix} -v \\ 0 \\ 0 \end{pmatrix}$$

When $\mathbf{x} = 0$:

$$\mathbf{c} = \dot{\mathbf{x}}^1 - \dot{\mathbf{x}}^2 = \begin{pmatrix} v - wr \\ 0 \\ 0 \end{pmatrix}; \quad \mathbf{v} = -(\dot{\mathbf{x}}^1 + \dot{\mathbf{x}}^2)/2 = \begin{pmatrix} (v + wr)/2 \\ 0 \\ 0 \end{pmatrix}$$

$\mathbf{c}$: creep at the origin, $\mathbf{v}$: rolling velocity at the origin.

1.6 Recapitulation of the linear theory of elasticity

In the linear theory of elasticity it is more convenient to work with index notation than with matrix-vector notation. In index notation, the displacement $\mathbf{u}^a$ ($a = 1, 2$: body number) is denoted by its components u_i^a, $i = 1, 2, 3$, in a Cartesian coordinate system. The body number a is omitted when no confusion may arise. Partial differentiation with respect to the coordinate x_j is denoted by $_{,j}$:

$$u_{i,j} \doteq \partial u_i / \partial x_j, \quad \text{the displacement gradient} \tag{28}$$

summation over 1,2,3 of repeated subscripts is understood. This summation convention is undone by placing one or more subscripts between brackets.

As usual, we work in a small displacement, small displacement gradient theory. The linearized stresses are defined as

$$e_{ij} = (u_{i,j} + u_{j,i})/2 = e_{ji}, \tag{29}$$

and the stress is given by σ_{ij}. The equations of equilibrium are

$$\sigma_{ij} = \sigma_{ji} \quad \text{and} \quad \sigma_{ij,j} = 0 \tag{30}$$

where we disregard inertial forces such as gravity in the rightmost equation. The surface load, also called the surface traction, is $p_i = \sigma_{ij} n_j$, where n_j is the outer normal on the surface.

According to Hooke's Law there is a linear relationship between the stresses σ_{ij} and the strains e_{hk}. In the general case, this relation reads

$$\sigma_{ij} = E_{ijhk} e_{hk}, \quad E_{ijhk}\text{: elastic constants} \tag{31}$$

and in the case of an isotropic material

$$e_{ij} = \frac{1+\nu}{E}\sigma_{ij} - \frac{\nu}{E}\delta_{ij}\sigma_{kk}, \quad \sigma_{ij} = \frac{E e_{ij}}{1+\nu} + \frac{E\delta_{ij}e_{kk}}{(1+\nu)(1-2\nu)} \tag{32}$$

with

E: Young's Modulus, ν: Poisson's Ratio, (elastic constants)

δ_{ij}: Kronecker delta, =0 when $i \neq j$, =1 when $i = j$.

1.7 Friction

When two bodies slide over each other, it will often be observed that this motion is opposed by a force. This is the phenomenon of friction; the force is called the force of friction. Usually a finite compensating force is needed to set a body sliding, while in many experiments the friction force remains constant during sliding. So it is assumed

that the shearing force is bounded by a force bound g, which depends on the normal force F_z, the magnitude of the sliding velocity V and other parameters, thus

$$g = g(F_z, V, ...), \quad ...: \text{ other parameters; } F_z: \text{ normal comp. of total contact force} \tag{33}$$

When the sliding velocity (the slip, in fact) vanishes, the tangential force may fall below the traction bound g in absolute value; when sliding occurs, the tangential force is at the traction bound, and it opposes the slip:

$$|F_\tau| < g(F_z, V, ...) \tag{34}$$

F_τ: tangential component of total contact force, $\tau = 1, 2$,

$|F_\tau| = \sqrt{F_1^2 + F_2^2}$

$$\text{if } V \neq 0: \quad F_\tau = -g v_\tau / V, \quad \text{Greek index: tangential comp.} \tag{35}$$

v_τ: tangential component of sliding velocity; $V = |v_\tau|$.

Coulomb [2] proposed that g is proportional to F_z with a constant of proportionality called the coefficient of friction:

$$g(F_z, V, ...) = f F_z \quad \text{Coulomb, 1785} \tag{36}$$

In order to interpret (36) Archard proposed in [1] that the friction was primarily caused by the adhesion of the bodies to each other. The adhesion takes place at the tips of the roughnesses of the bodies, which are called the asperities. At these tips the bodies are in contact and they form junctions, the real contact surface C_r, as opposed to the apparent contact area C which consists of the junctions and the region in between them. Archard showed that the area of the real contact surface is proportional to F_z, the normal compressive force.

At the real contact surface the bodies are welded together by interatomic forces. Owing to the sliding motion the welded asperities shear; eventually the welds break, and the asperities form new welds with different partners. The shearing of the asperities is accompanied by micro-plastic deformation, and also by the detachment of material particles from the bodies, thus leading to wear. From this it is seen that friction and wear are closely connected phenomena.

Despite this bolstering of Coulomb's Law, it is generally agreed by tribologists that (36) must be modified. The simplest modification is the introduction of two coefficients of friction, one for non-sliding (f_{stat}) and one for sliding (f_{kin}), with

$$f_{stat} > f_{kin} \tag{37}$$

This did not suffice for many researchers, and they proposed more complicated relationships for the coefficient of friction.

So far we considered the total contact force, and the global velocity in sliding. In contact mechanics, however, we need a local form of the friction law. A very simple theory suggests itself: instead of global quantities use local quantities, that is instead of $\mathbf{F}$ use $\mathbf{p}$, the local traction, and instead of the global sliding velocity V use the slip $\mathbf{s}$:

$|p_\tau| \leq g(p_z, |s_\tau|, ...)$

$$\text{if} \quad |s_\tau| \neq 0 \qquad \Longrightarrow \qquad p_\tau = -g s_\tau / |s_\tau| \tag{38}$$

p_τ: tangential traction component

s_τ: slip component, $\tau = 1, 2$

This law was stated and experimentally confirmed by Rabinowicz [7]. It was used earlier in theoretical work see [4], [6], [3], and [5]. The traction bound is taken as

$$g(p_z, |s_\tau|, ...) = f(|s_\tau|, ...) p_z \tag{39}$$

and f is usually taken constant.

1.8 Boundary conditions

We recall Hooke's Law:

$$\sigma_{ij} = E_{ijhk} e_{hk} \tag{40}$$

It is valid for all types of bodies. Sometimes it is possible to bring it in surface mechanical form:

$$u_i(\mathbf{x}) = \int_{\partial V} A_{ik}(\mathbf{x}, \mathbf{y}) p_k(\mathbf{y}) dS(\mathbf{y} \tag{41}$$

where $A(\mathbf{x}, \mathbf{y})$ is the displacement at $\mathbf{x}$ due to a point load at $\mathbf{y}$; it is called the Influence Function. The influence function depends strongly on the form of the body. In 3D elasticity it has been calculated for a few forms, in particular the half-space.

The advantage of (41) over (40) resides in the fact that for a 3D body (41) is taken over its 2D boundary only, while (40) extends over its entire 3D iterior.

The surface of body a is divided into three parts.

- In A_p the surface load is prescribed as $\overline{\mathbf{p}}$.
- In A_u the surface displacement is prescribed as $\overline{\mathbf{u}}$
- A_c is the potential contact zone. It was already defined below eq. (17).

We give here a slightly different definition of the potential contact zone, which is roughly equivalent to the one given before. It reads as follows:

The potential contact zone can be freely chosen under the following conditions:

1. It must completely encompass the contact region;
2. In it, $\mathbf{x}^1 - \mathbf{x}^2 = O(\mathbf{u})$;
3. In it, $\mathbf{n}^1 \approx -\mathbf{n}^2$.

Then the following relations hold in the potential contact:

$$e(\mathbf{x}) = h(\mathbf{x}) + \mathbf{n}^{2T}\mathbf{u}(\mathbf{x}), \quad \mathbf{u}(\mathbf{x}) = \mathbf{u}^1(\mathbf{x}) - \mathbf{u}^2(\mathbf{x}), \quad e \geq 0 \tag{42}$$

$$p_N(\mathbf{x}) = \mathbf{n}^{2T}\mathbf{p}^{\mathbf{1}}, \quad p_N \geq 0, \quad p_N e = 0 \tag{43}$$

where the fields in $\mathbf{x}^1$ and $\mathbf{x}^2$ are suitably extended to

$$\mathbf{x} = (\mathbf{x}_1 + \mathbf{x}_2)/2. \tag{44}$$

The region where $e = 0$ is called the contact zone (contact region, contact area, contact patch). It is denoted by C. In it, $p_N \geq 0$.
The region where $e > 0$ is called the exterior zone (.. region, .. area). It is denoted by E. In it, $p_N = 0$. (42),(43) are the mathematical description of $C \cup E = A_c$. As we saw in the Example of the Winkler bedding, (42),(43) determine the solution of the frictionless problem completely and uniquely. The inequalities of (42),(43) play a crucial role in the determination of the contact area and the elastic field.
The equation $p_N e = 0$ is very important. In conjunction with the inequalities it means that when $e > 0 \implies p_N = 0$ and vice versa.

We turn to the boudary conditions of friction.
We had defined the tangential component of the slip in the contact zone as $\mathbf{s} = \dot{\mathbf{x}}^1 - \dot{\mathbf{x}}^2 + \dot{\mathbf{u}}^1 - \dot{\mathbf{u}}^2$.
The contact area is divided into two parts, *viz.* the stick area (area of adhesion, ..zone, ..region) H, where the tangential component of the slip vanishes, and the area (zone, region) of slip S, where this is not so. We have for the total boundary conditions in the potential contact area A_c:

$$\text{in } H: \quad |\mathbf{s}_T| = 0, \qquad |\mathbf{p}_T| \leq g \tag{45}$$

$$\text{in } S: \quad |\mathbf{s}_T| \neq 0, \quad \mathbf{p}_T = -g\mathbf{s}_T/|\mathbf{s}_T \tag{46}$$

$$S \cup H = C, \qquad S \cap H = \oslash \tag{47}$$

$$\text{in } C: \quad e = 0, \qquad p_N \geq 0 \tag{48}$$

$$\text{in } E: \quad \mathbf{p} = 0, \qquad e > 0 \tag{49}$$

$$C \cup E = A_c, \qquad C \cap E = \oslash \tag{50}$$

References of Section 1

[1] J.F. Archard (1957), *Elastic deformation and the laws of friction.* Proceedings of the Royal Society of London, Ser. **A 243**, p. 190-205.

[2] C.A. Coulomb (1785), *Théorie des machines simples*. Mémoire de Mathématique et de Physique de l'Académie Royale, p. 161-342.

[3] F.W. Carter (1926), *On the Action of a Locomotive Driving Wheel*. Proceedings of the Royal Society of London **A 112** p. 151-157.

[4] C. Cattaneo (1938), *Sul contatto di due corpi elastici: distribuzione locale degli sforzi*. Accademia Nazionale Lincei Rendiconti Ser 6 **XXVII** p. 342-348, 434-436, 474-478.

[5] H. Fromm (1927), *Berechnung des Schlupfes beim Rollen deformierbaren Scheiben*. Zeitschrift für angewandte Mathematik und Mechanik **7** p. 27-58.

[6] R.D. Mindlin (1949), *Compliance of elastic bodies in contact*. Journal of Applied Mechanics **16**, p. 353-383.

[7] E. Rabinowicz, *Friction and Wear of Materials*. John Wiley and Sons, New York.

2 The half-space approximation

One of the ways to attack a contact problem is by means of the FEM. Especially when we deal with 3D problems this is very time consuming. The BEM of (1.41) is a much more promising option, if it is at all feasible.

This is so when we deal with a homogeneous, isotropic elastic 3D half-space. This is of great technological importance, as many problems may be approximately solved by using a half-space.

A half-space consists of all points on one side of a plane, the bounding plane. For instance, in a Cartesian coordinate system $(O; x, y, z)$ a half-space may be defined by $\{(x, y, z) : z \geq 0\}$. The contact field of two elastic bodies may be calculated by means of half-spaces when the maximum diameter of the contact area is small with respect to a typical dimension of the bodies, such as their diameter or the minimum radius of curvature near the contact. Then, the elastic field in the contact part may be calculated by replacing the body locally by a half-space. The boundary conditions are then of the real body, but the elasticity equations are solved for the half-space. The situation is shown in Fig. 5.
Examples of the cases where the half-space approximation is valid:

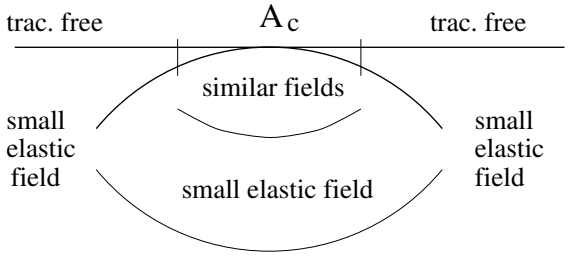


Figure 5: The half-space approximation

A ball pressed onto a thick slab, tread contact in wheels and rails, see Fig. 6.

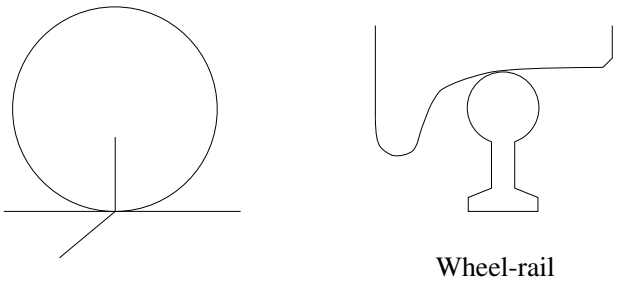


Figure 6: Examples of the half-space approximation

Properties of the half-space approximation are:

1. Many geometries are elastically alike.
   This is most important as it renders software written for half-spaces applicable to many geometries. The relative ease of the half-space approximation leads one to use it even where one is apt to make serious errors.

2. The influence numbers $A_{ik}(\mathbf{x}, \mathbf{y})$ can be calculated exactly by means of the formulae of Boussinesq [1] and Cerruti [2].
   Derivations of these formulae may be found in [6] and in [3].

   We give these formulae. We denote the global coordinate system by $(O; x, y, z) = (O; x_i), i = 1, 2, 3$. O lies in the surface of the half-space. The axes of x and y lie in the surface of the half-space. The axis if $x = x_1$ points in the rolling direction, the axis of $z = x_3$ points vertically upwards into body 1, and $y = x_2$ completes the right-handed coordinate system. Moreover, we denote the components of $\mathbf{u}^a$ by $u_i^a = (u^a, v^a, w^a)$ and the components of the surface traction by $\mathbf{p}^a$ by $p_i^a = (p_x^a, p_y^a, p_z^a)$.
   Clearly, by Newton's Third Law,

$$p_i^1 = -p_i^2 \doteq p_i \tag{51}$$

   (41) reads

$$u_i^a(\mathbf{x}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} A_{ij}^a \; p_j^a(x', y') \; dx'dy' \tag{52}$$

   In body $a, a = 1, 2$ we have

$$A_{1,3}^a = \frac{1}{4\pi G^a}\{\frac{(x - x')|z|}{r^3} - \frac{(1 - 2\nu^a)(x - x')}{r(|z| + r)}\} \tag{53}$$

$$A_{2,3}^a = \frac{1}{4\pi G^a}\{\frac{(y - y')|z|}{r^3} - \frac{(1 - 2\nu^a)(y - y')}{r(|z| + r)}\} \tag{54}$$

$$A_{3,3}^a = \frac{-(-1)^a}{4\pi G^a}\{\frac{z^2}{r^3} + \frac{2(1 - \nu^a)}{r}\} \tag{55}$$

$$r = \sqrt{(x - x')^2 + (y - y')^2 + z^2} \tag{56}$$

The factor $-(-1)^a$ in (55) calls for comment. When $a = 1$, it equals unity. Then the z-axis points into the body, and the component w^1 has the same direction as the concentrated load $p_3^1 = p_3$. When $a = 2$, we must have the same formulae but in a coordinate system in which the z-axis points vertically downwards, while the axes of x and y remain the same. Then one must invert the sign of z in all formulae (53)-(55), and one must reverse the sign of w as well.
So we see that the displacements in bodies 1 and 2 due to a normal loading obey the law

$$\begin{aligned} u^1(x,y,z) &= u^2(x,y,-z) \\ v^1(x,y,z) &= v^2(x,y,-z) \\ w^1(x,y,z) &= -w^2(x,y,-z) \\ \text{if} \quad p_x &= p_y = 0 \end{aligned}$$

The displacement differences, which are prescribed in the normal and tangential problems, are

$$\begin{aligned} u(x,y) &= u^1(x,y,0) - u^2(x,y,0) = \\ &= \frac{1}{4\pi}\{\frac{1-2\nu^1}{G^1} - \frac{1-2\nu^2}{G^2}\}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_z(x',y')\frac{x'-x}{R^2}dx'dy' \\ v(x,y) &= v^1(x,y,0) - v^2(x,y,0) = \\ &= \frac{1}{4\pi}\{\frac{1-2\nu^1}{G^1} - \frac{1-2\nu^2}{G^2}\}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_z(x',y')\frac{y'-y}{R^2}dx'dy' \\ w(x,y) &= w^1(x,y,0) - w^2(x,y,0) = \\ &= \frac{1}{2\pi}\{\frac{1-\nu^1}{G^1} + \frac{1-\nu^2}{G^2}\}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_z(x',y')/Rdx'dy' \qquad (57) \\ p_x &= p_y = 0, \qquad R = \sqrt{(x'-x)^2 + (y'-y)^2} \\ G^a: &\quad \text{modulus of rigidity,} \quad \nu^a: \quad \text{Poisson's Ratio:} \\ \text{Elastic} &\quad \text{constants of body } a \end{aligned}$$

We combine ν^1, ν^2 and G^1, G^2 in the following manner:

$$1/G = \{(1/G^1) + (1/G^2)\}/2 \qquad (58)$$

$$\nu/G = \{(\nu^1/G^1) + (\nu^2/G^2)\}/2 \qquad (59)$$

$$K = (G/4)\{(1-2\nu^1)/G^1 - (1-2\nu^2)/G^2\} \qquad (60)$$

It is easy to see that G lies between G^1 and G^2. In the case of elastic symmetry (= both bodies having the same elastic constants),

$$G = G^1 = G^2, \quad \nu = \nu^1 = \nu^2, \quad K = 0 \qquad (61)$$

The constant K vanishes when there is elastic symmetry, and also when both bodies are incompressible: Steel on Steel, Rubber on Steel. Its maximum is 0.5

but in practice it is mostly small, e.g. 0.03 for steel on brass, and 0.09 for steel on aluminium. In terms of the constants of (58)-(60), the displacement difference becomes:

$$\begin{aligned} u(x,y) &= \frac{K}{\pi G}\int_{-\infty}^{+\infty} p_z(x',y')\frac{x'-x}{R}dx'dy' && (62)\\ v(x,y) &= \frac{K}{\pi G}\int_{-\infty}^{+\infty} p_z(x',y')\frac{y'-y}{R}dx'dy' && (63)\\ w(x,y) &= \frac{1-\nu}{\pi G}\int_{-\infty}^{+\infty} p_z(x',y')/Rdx'dy' && (64)\\ p_x &= p_y = 0 && (65)\\ R &= \sqrt{(x'-x)^2+(y'-y)^2} && (66) \end{aligned}$$

The procedure for the tangential traction/displacement is very nearly the same. We have for the displacement due to the load p_x, second index of A^a_{ij} equal to 1:

$$\begin{aligned} A^a_{11} &= (-(-1)^a)\frac{1}{4\pi G^a}\{(1/r)+\\ &+ \frac{1-2\nu^a}{|z|+r}+\frac{(x'-x)^2}{r^3}-\frac{(1-2\nu^a)(x'-x)^2}{r[|z|+r]^2}\} && (67)\\ A^a_{21} &= (-(-1)^a)\frac{1}{4\pi G^a}\{\frac{(x'-x)(y'-y)}{r^3}+\\ &- \frac{(1-2\nu^a)(x'-x)(y'-y)}{r[|z|+r]^2}\} && (68)\\ A^a_{31} &= -\frac{1}{4\pi G^a}\{\frac{(x'-x)|z|}{r^3}+\frac{(1-2\nu^a)(x'-x)}{r[|z|+r]}\} && (69) \end{aligned}$$

The displacement due to a load p_y in the y direction, second index of A^a_{ij} equal to 2 is found by the interchange of x and y, u and v, p_x and p_y. The factor $-(-1)^a$ calls for comment: we must take into account that the shearing tractions on the bodies have different signs, and that therefore u, v, w have different signs in the two bodies, but that w is taken in a coordinate system that has a z-axis with the other sign, so that the factor $-(-1)^a$ is neutralized for the vertical displacement w.

The total surface displacement differences $u(x,y), v(x,y), w(x,y)$ become (note that $z=0$!)

$$\begin{aligned} u(x,y) &= \frac{1}{\pi G}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{p_x(x',y')[\frac{1-\nu}{R}+\frac{(x'-x)^2}{R^3}]+\\ &+ p_y(x',y')\frac{(x'-x)(y'-y)}{R^3}+Kp_z(x',y')\frac{x'-x}{R^2}\}dx'dy' && (70)\\ v(x,y) &= \frac{1}{\pi G}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{p_x(x',y')\frac{(x'-x)(y'-y)\nu}{R^3}+ \end{aligned}$$

$$
\begin{aligned}
+ \quad & p_y(x',y')[\frac{1-\nu}{R}+\frac{(y'-y)^2}{R^3}]+Kp_z(x',y')\frac{y'-y}{R^2}\}dx'dy' \\
w(x,y) \quad & = \frac{1}{\pi G}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{\ -Kp_x(x',y')\frac{x'-x}{R^2}+ \\
- \quad & Kp_y(x',y')\frac{y'-y}{R^2}+p_z(x',y')\frac{1-\nu}{R}\}dx'dy' \qquad (71)
\end{aligned}
$$

3. Quasiidentity.
The half-spaces in contact are called quasiidentical, when $K=0$:

$$\text{Quasiidentity} \quad \Longleftrightarrow K=0 \qquad (72)$$

When $K=0$, the displacement differences u and v are not influenced by the normal traction p_z, and the displacement diference w is not influenced by the tangential tractions p_x and p_y. Yet there is a coupling through the traction bound, *viz.*

$$|(p_x,p_y)| \le fp_z, \quad f\text{: coefficient of friction.} \qquad (73)$$

Thus we must act as follows:

$$
\begin{aligned}
&\text{Determine} \quad p_z \quad \text{when } p_x=p_y=0 \quad \text{(Normal problem)} \\
&\text{Determine} \quad p_x, \quad p_y \quad \text{when } g=fp_z \quad \text{(Tangential problem)} \qquad (74)
\end{aligned}
$$

This is called the Johnson process. It is approximately valid when K is small. When $K \neq 0$, we must act differently. We use the so-called Panagiotopoulos process:

(a) Set $I=0$, and $p_x^I=p_y^I=0$

(b) Determine the compressive traction p_z^I with (p_x^I,p_y^I) as the tangential traction, by means of a normal (= compressive traction) contact algorithm.

(c) Determine the next tangential traction (p_x^{I+1},p_y^{I+1}) with p_z^I as the compressive traction, by means of a tangential contact algorithm.

(d) If (p_x^{I+1},p_y^{I+1}) is close enough to (p_x^I,p_y^I), we stop, else we set $I=I+1$, and we restart at b.

The Panagiotopoulos process was used by Oden and Pires [7] to prove the existence of the elastic field for elastostatic contact. I use it myself for 3D half-space contact, and for 2D contact of elastic and viscoelastic layered media. It can, apparently, always be made to converge, but sometimes one must perturb the discretization of the contact problem.

Remark: The Johnson process is a one-step Panagiotopoulos process.

4. Slip in the Half-Space.
We had seen in Section 1 that the slip was given by

$$\mathbf{s} \quad = \mathbf{c} - (\partial \mathbf{u}/\partial \mathbf{x})\mathbf{v} + \partial \mathbf{u}/\partial t$$

with	$\mathbf{s}$	the slip
	$\mathbf{c}$	the creep
	$\mathbf{u}$	the displacement difference
	$\mathbf{v}$	the rolling velocity
	$\mathbf{x}$	the position
	t	the time.

In half-space rolling, we assume that the rolling takes place along the positive x-axis with constant velocity V, and that the motion takes place in the contact plane, so that

$$\mathbf{c} = V(\upsilon_x - \varphi y, \upsilon_y + \varphi x)^T \tag{75}$$
$$\mathbf{v} = (V, 0)^T \tag{76}$$

and

$$s_x = V(\upsilon_x - \varphi y - \partial u/\partial x + \partial u/\partial(Vt)) \tag{77}$$
$$s_y = V(\upsilon_y + \varphi x - \partial v/\partial x + \partial v/\partial(Vt)) \tag{78}$$

υ_x is called the longitudinal creepage, υ_y is called the lateral creepage, and φ is called the spin. Another word for creepage is creep ratio.
We identify

$$\int_0^t Vt'dt' = q, \quad \text{distance traversed} \tag{79}$$

and divide (77) and (78) by V, where we call

$$\mathbf{S} = \mathbf{s}/V, \quad \text{the relative slip} \tag{80}$$

When we replace the time variable t by the distance traversed q, and the slip $\mathbf{s}$ by the relative slip $\mathbf{S}$, then we see that the relative slip is independent of the rolling velocity V. Since the relative slip has the same direction as the real slip, and the formula for the tangential traction in sliding depends only on that direction when the coefficient of friction is independent of the rolling velocity (as we will assume), the entire problem becomes independent of the rolling velocity, or, more precisely, depends on the rolling velocity only through the coefficient of friction.

2.1 The Hertz Problem

Consider two paraboloids with parallel axes, numbered 1 and 2, see Fig. 7.

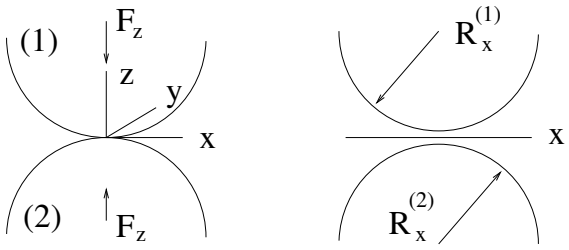


Figure 7: The Hertz problem

They are brought together so that their tips touch. A Cartesian coordinate system is introduced in which the plane of x and y is the common tangent plane, the origin O lies in the tangent point, and the z-axis points vertically upwards into paraboloid 1. We assume for simplicity that the planes of principal curvature coincide with one another and with the planes of x and y, as is often the case in rolling. The radii of curvature in the plane of x are $R_x^{(1)}$ and $R_x^{(2)}$ and those in the plane of y are $R_y^{(1)}$, $R_y^{(2)}$. We count them positive when the centre of curvature lies inside the corresponding paraboloid.
The bodies are compressed over a distance q. The problem is:
FIND THE CONTACT AREA AND THE PRESSURE DISTRIBUTION WHEN FRICTION IS ABSENT OR IN THE CASE OF QUASIIDENTITY.

We choose the unstressed state so, that the displacements and the stresses vanish at infinity. Then the equation of the surface of body a is

$$z^{(a)} = -(-1)^a \{ \frac{x^2}{2R_x^{(a)}} + \frac{y^2}{2R_y^{(a)}} \} - q^{(a)} \tag{81}$$

In the contact region, we have that

$$h = z^{(1)} - z^{(2)} = Ax^2 + By^2 - q \tag{82}$$

with

$$A = \frac{1}{2R_x^{(1)}} + \frac{1}{2R_x^{(2)}} \tag{83}$$

$$B = \frac{1}{2R_y^{(1)}} + \frac{1}{2R_y^{(2)}} \tag{84}$$

$$q = q^{(1)} - q^{(2)} \tag{85}$$

$$\frac{1}{R} = \frac{A+B}{2} \tag{86}$$

Hertz [9], [10] assumed that the contact area was elliptic in the half-space approximation, with

$$C = \{(x, y, 0) | (x/a)^2 + (y/b)^2 \leq 1\} \tag{87}$$

Then it may be shown that the normal pressure is semi-ellipsoidal

$$p_z(x,y) = Gf_{00}\sqrt{1-(x/a)^2-(y/b)^2} \tag{88}$$

where G and ν are the combined modulus of rigidity and the Poisson's ratio, see (58), (59).
The total normal force is found by integration of (88)

$$N = \int\int_C p_z dxdy = (2/3)\pi abGf_{00}, \quad f_{00} = \frac{3N}{2\pi abG} \tag{89}$$

The minor and the major semi-axes of the contact ellipse are denoted by S and L, and the axial ratio by $m' = S/L$:

$$m' = S/L, \quad S = \min\,(a,b), \quad L = \max\,(a,b) \tag{90}$$

The excentricity is signed. We have

$$|m| = \sqrt{1-m'^2}, \quad m > 0 \quad \text{if} \;\; a < b, \quad m < 0 \quad \text{if} \quad a > b \tag{91}$$

Then q, A and B are given by

$$q = \frac{3N(1-\nu)S\mathbf{K}}{2\pi abG} \tag{92}$$

$$A(|m|) = B(-|m|) = \frac{3N(1-\nu)(\mathbf{D}-m^2\mathbf{C})}{2\pi abSG} \tag{93}$$

$$B(|m|) = A(-|m|) = \frac{3N(1-\nu)(1-m^2)\mathbf{D}}{2\pi abSG} \tag{94}$$

where $\mathbf{K}$, $\mathbf{B}$ and $\mathbf{D}$ are complete elliptic integrals. We will encounter more complete elliptic integrals. They can all be expressed as a linear combination of two of them, for which we take $\mathbf{C}$ and $\mathbf{D}$:

$$\mathbf{B} = \int_0^{\pi/2} \frac{\cos^2 t}{\sqrt{1-m^2\sin^2 t}}dt, \mathbf{B} = \mathbf{D} - m^2\mathbf{C} \tag{95}$$

$$\mathbf{C} = \int_0^{\pi/2} \frac{\sin^2 t \;\cos^2 t}{\sqrt{1-m^2\sin^2 t}^3}dt \tag{96}$$

$$\mathbf{D} = \int_0^{\pi/2} \frac{\sin^2 t}{\sqrt{1-m^2\sin^2 t}}dt \tag{97}$$

$$\mathbf{E} = \int_0^{\pi/2} \sqrt{1-m^2\sin^2 t}\,dt, \quad \mathbf{E} = (2-m^2)\mathbf{D} - m^2\mathbf{C} \tag{98}$$

$$\mathbf{K} = \int_0^{\pi/2} \frac{1}{\sqrt{1-m^2\sin^2 t}}dt, \quad \mathbf{K} = 2\mathbf{D} - m^2\mathbf{C} \tag{99}$$

We give these functions in their dependence on m' in Table 1, taken from [4].

m'	**B**	**C**	**D**	**E**	**K**	m^2
$\downarrow 0$	1	-2+ln 4/m'	-1+ln 4/m'	1	+ln 4/m'	1.00
0.1	0.9889	1.7351	2.7067	1.0160	3.6956	0.99
0.2	0.9686	1.1239	2.0475	1.0505	3.0161	0.96
0.3	0.9451	0.8107	1.6827	1.0965	2.6278	0.91
0.4	0.9205	0.6171	1.4388	1.1507	2.3593	0.84
0.5	0.8959	0.4863	1.2606	1.2111	2.1565	0.75
0.6	0.8719	0.3929	1.1234	1.2763	1.9953	0.64
0.7	0.8488	0.3235	1.0138	1.3456	1.8626	0.51
0.8	0.8267	0.2706	0.9241	1.4181	1.7508	0.36
0.9	0.8055	0.2292	0.8491	1.4933	1.6546	0.19
1.0	.7864=$\frac{\pi}{4}$	.1964=$\frac{\pi}{16}$	.7854=$\frac{\pi}{4}$	1.571=$\frac{\pi}{2}$	1.571=$\frac{\pi}{2}$	0.00

Table 1: Complete elliptic integrals (Jahnke-Emde, 1943)[4]

It is seen from Table 1 that $\mathbf{D} > \mathbf{C}$, and it follows from (93) and (94) that $A(|m| = B(-|m|) \geq A(-|m|) = B(|m|)$, so that we have

$$A \geq B \implies m \geq 0, \quad a \leq b \tag{100}$$

$$A \leq B \implies m \leq 0, \quad a \geq b \tag{101}$$

$A, B:$ see (86)

In order to find the excentricity of the contact ellipse, we set with Hertz [9]

$$\cos t = \frac{|A-B|}{A+B}, \quad \text{with } R = \frac{2}{A+B} \tag{102}$$

and it follows from (93),(94), the fact that $\mathbf{D} > \mathbf{C}$ and the expression of $\mathbf{E}$ in $\mathbf{D}$ and $\mathbf{C}$ that

$$\cos t = m^2(\mathbf{D} - \mathbf{C})/\mathbf{E} \tag{103}$$

We give the axial ratio m' as a function of t in Table 2.

t	90^o	80^o	70^o	60^o	50^o	40^o	30^o	20^o	10^o	0^o
$m' = S/L$	1.00	.79	.62	.47	.36	.26	.18	.10	.05	0

Table 2: The axial ratio of the contact ellipse as a function of t.

Table 2 is taken from [6] (page 197). We see from (86) to (103) that the shape of the contact ellipse depends only on the radii of curvature of the bodies, and not on the applied load nor the elastic properties. The size of the contact area does, however:

$$A + B = \frac{2}{R} = \frac{3N(1-\nu)\mathbf{E}}{2G\pi abS} \tag{104}$$

or

$$3N(1-\nu)R\mathbf{E} = 4\pi abSG \tag{105}$$

A frequently used quantity is f_{00}, see (88). Gf_{00} is the maximum value of the ellipsoidal pressure distribution $p_z(x, y)$. It is

$$f_{00} = \frac{3N}{2\pi abG} = \frac{2S}{(1-\nu)R\mathbf{E}} \tag{106}$$

Finally we determine the deepest penetration of the bodies, see (82):

$$q = (1-\nu)\mathbf{K}Sf_{00} = \frac{2S^2\mathbf{K}}{R\mathbf{E}} \tag{107}$$

2.1.1 The linear theory of rolling contact for Hertzian contacts

We recall the formulae for the relative slip $\mathbf{s}/V$ in steady state rolling:

$$S_x = s_x/V = \upsilon_x - \varphi y - \partial u(x,y)/\partial x \tag{108}$$

$$S_y = s_y/V = \upsilon_y + \varphi x - \partial v(x,y)/\partial x \tag{109}$$

with

V	:	the rolling velocity
$\upsilon_x, \upsilon_y, \varphi$	:	longitudinal, lateral, spin creepage
(u, v)	:	tangential displacement difference, body (1)-(2)
x-axis	:	rolling direction
z-axis	:	vertical direction pointing into body (1)
y-axis	:	completes right-handed system
O	:	origin in centre of contact patch (ellipse)

When the creepages are very small, it is easy for (u, v) to compensate them, without violating the traction bound. That is, almost the entire contact patch will be covered by the stick region. Instead of saying that the creepages are small, one can also say that the coefficient of friction goes up to infinity. Then the traction bound will not be violated as long as (p_x, p_y) is finite. Now for finite (u, v) the traction is finite except maybe on the edge of the elliptic contact area, as it may be shown that the traction has inverse square root behaviour near the edge:

$$(p_x(x,y), p_y(x,y)) = O(\sqrt{1-(x/a)^2-(y/b)^2}^{-1}) \tag{110}$$

so that the slip is confined to part of the edge of the contact area.

Hence inside the contact patch we have

$$\partial u(x,y)/\partial x = \upsilon_x - \varphi y, \quad \partial v(x,y)/\partial x = \upsilon_y + \varphi x \tag{111}$$

Integrating with respect to x we find

$$u(x,y) = \upsilon_x x - \varphi xy + k(y) \quad \text{in } C \tag{112}$$

$$v(x,y) = \upsilon_y x + \varphi x^2/2 + l(y) \quad \text{in } C \tag{113}$$

$$p_x(x,y) = p_y(x,y) = 0 \quad \text{in } E \tag{114}$$

where k and l are arbitrary functions of y.

The question arises, how to determine k and l. We observe that (112) to (114) fully determine the contact problem together with the proper behaviour at infinity, where the elastic field has died out according to the half-space hypothesis.
To that end, we observe that at the leading edge of the contact ellipse the traction must vanish. For at the leading edge unloaded material flows into the contact patch. During transit of the contact patch, traction builds up, which is suddenly released at the trailing edge. So at the leading edge the traction must vanish, but it need not do so at the trailing edge. Hence k and l must be determined so that the traction vanishes at the leading edge, which is given by

$$\{(x,y,z)|x \geq 0, z = 0, |y| = b\sqrt{1-(x/a)^2}\} \quad \text{leading edge} \tag{115}$$

We explain why this theory is called the Linear Theory. When we consider (112)-(114), and multiply $\upsilon_x, \upsilon_y, \varphi, k, l$ with the same constant D, then it is clear that $D(p_x, p_y)$ satisfy the no-slip contact conditions with $D(p_x, p_y) = (0,0)$ on the leading edge. So the no-slip tangential traction corresponding to $D(\upsilon_x, \upsilon_y, \varphi)$ is $D(p_x, p_y)$ – a linear relationship.

At this place, we will not enter into the details of the calculation, but merely mention the result. We tabulate the total tangential force (F_x, F_y) and the twisting moment M_z in terms of the creepages, as follows:

$$F_x = \int\int_C p_x dxdy, F_y = \int\int_C p_y dxdy, \tag{116}$$

$$M_z = \int\int_C (xp_y - yp_x)dxdy \tag{117}$$

$$F_x = -GabC_{11}\upsilon_x, \quad F_y = -GabC_{22}\upsilon_y - G(ab)^{1.5}C_{23}\varphi, \tag{118}$$

$$M_z = -G(ab)^{1.5}C_{32}\upsilon_y - G(ab)^2 C_{33}\varphi \tag{119}$$

with

$$F_x, F_y, M_z \quad : \quad \text{total force and moment on body 1} \tag{120}$$

$$\upsilon_x, \upsilon_y, \varphi \quad : \quad \text{creepage (relative rigid slip) of body 1 wrt.body 2} \tag{121}$$

This form of the total forces and the twisting moment follows from symmetries, while it is an empirical fact that

$$C_{11} > 0, C_{22} > 0, C_{23} = -C_{32} > 0, C_{33} > 0 \tag{122}$$

These creepage and spin coefficients are tabulated in Table 3. The creepage and spin coefficients can be calculated in a purely numerical way. They can also be calculated more accurately in a semi-analytical way; these more accurate results are shown here.

a/b	C_{11}		C_{22}		$C_{23} =$	$-C_{32}$	C_{33}	
	$\nu = 0$	0.5	$\nu = 0$	0.5	$\nu = 0$	0.5	$\nu = 0$	0.5
0.1	2.51	4.85	2.51	2.53	.334	.731	6.42	11.7
0.2	2.59	4.81	2.59	2.66	.483	.809	3.46	5.66
0.3	2.68	4.80	2.68	2.81	.607	.889	2.49	3.72
0.4	2.78	4.82	2.78	2.98	.720	.977	2.02	2.77
0.5	2.88	4.83	2.88	3.14	.827	1.07	1.74	2.22
0.6	2.98	4.91	2.98	3.31	.930	1.18	1.56	1.86
0.7	3.09	4.97	3.09	3.48	1.03	1.29	1.43	1.60
0.8	3.19	5.05	3.19	3.65	1.13	1.40	1.34	1.42
0.9	3.29	5.12	3.29	3.82	1.23	1.51	1.27	1.27
1.0	3.40	5.20	3.40	3.98	1.33	1.63	1.21	1.16
1/.9	3.51	5.30	3.51	4.16	1.44	1.77	1.16	1.06
1/.8	3.65	5.42	3.65	4.39	1.58	1.94	1.10	.954
1/.7	3.82	5.58	3.82	4.67	1.76	2.18	1.05	.852
1/.6	4.06	5.80	4.06	5.04	2.01	2.50	1.01	.751
1/.5	4.37	6.11	4.37	5.56	2.35	2.96	.958	.650
1/.4	4.84	6.57	4.84	6.31	2.88	3.70	.912	.549
1/.3	5.57	7.34	5.57	7.51	3.79	5.01	.868	.446
1/.2	6.96	8.82	6.96	9.79	5.72	7.89	.828	.341
1/.1	10.7	12.9	10.7	16.0	12.2	18.0	.795	.228

Table 3: Table of the creepage and spin coefficients

2.1.2 The theory of Vermeulen and Johnson for steady state rolling

In 1964, Vermeulen and Johnson [8] gave a theory for steady-state rolling for pure creepage ($\varphi = 0$). They assumed that the stick area was also bounded by an ellipse, with the same axial ratio as the contact ellipse and the same orientation, see Fig. 8, but touching the contact ellipse at its foremost point.

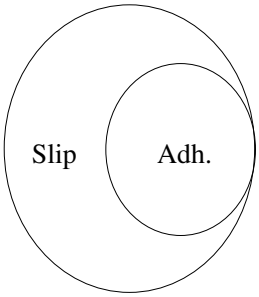


Figure 8: Areas and slip and adhesion according to Vermeulen and Johnson

Based upon that, they derived the creepage-force law shown in Fig. 9, broken line. In the left figure, the dots and crosses are the experimental values of Vermeulen and Johnson. Kalker improved the theoretical curve a little by making use of the creepage coefficients, and the result is shown in the left and right figure 9 by a drawn line. In the left figure, the dots and crosses are the experimental results of Vermeulen and Johnson, and in the right figure they are a comparison with the results of the numerical programs Fastsim (see Section 3) and Contact (see Section 4).

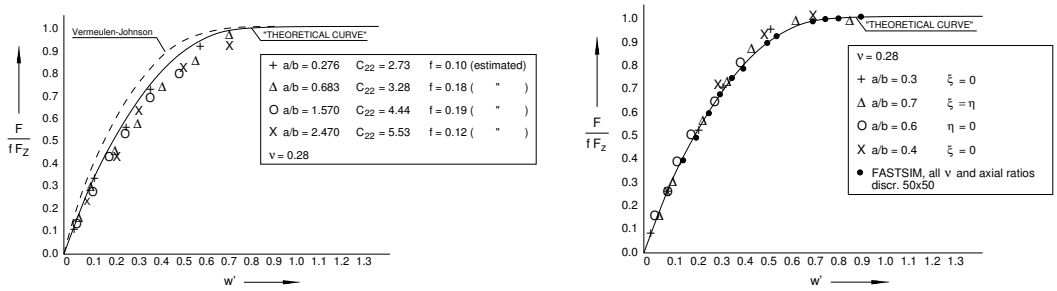


Figure 9: Results of Vermeulen and Johnson. Left, comparison with the experiment; right, comparison with Fastsim and Contact, Kalker (1990)

It is seen that the drawn line is a near perfect fit to the experiment, Contact and Fastsim.
It is of interest to have the equation of the drawn line.
Let

$$\xi' \doteq -\frac{abGC_{11}\upsilon_x}{3fF_z} \tag{123}$$

$$\eta' \doteq -\frac{abGC_{22}\upsilon_y}{3fF_z} \quad \text{with} \tag{124}$$

$$C_{ii} : \text{Creepage coefficient of the linear theory, see Table 3} \tag{125}$$

$$a, b : \text{Semiaxes of contact ellipse, a in rolling direction} \tag{126}$$

$$
\begin{aligned}
G &: \text{Modulus of rigidity} && (127)\\
f &: \text{Coefficient of friction} && (128)\\
F_z &: \text{Total compressive normal force} && (129)
\end{aligned}
$$

We set

$$w' = |(\xi', \eta')| \tag{130}$$

and it equals

$$
\begin{aligned}
w' &= 1 - [1 - (F/fF_z)]^{1/3} \quad \text{if } F \leq fF_z, \quad F = |(F_x, F_y)| \\
&\geq 1 \qquad\qquad\qquad\qquad\quad \text{if } F > fF_z
\end{aligned} \tag{131}
$$

We have for the "theoretical line":

$$
\begin{aligned}
(F_x, F_y) &= (F/w')(\xi', \eta') && (132)\\
F &= fF_z[1 - (1 - w')^3] \quad \text{if } w' \leq 1 && (133)\\
&= fF_z \qquad\qquad\qquad\quad \text{if } w' \geq 1 && (134)
\end{aligned}
$$

The only difference between the definition of Kalker's line and the line of Vermeulen-Johnson is that Vermeulen-Johnson derive their line analytically on the basis of their division of the contact area in regions of stick and slip, and arrive then at a different set of creepage coefficients, whereas Kalker maintains the form of the line of Vermeulen and Johnson, but gives it the correct slope at the origin $w' = 0$.
Indeed we have for the linear theory:

$$
\begin{aligned}
F &= \frac{\partial F}{\partial w'}|_{w'=0} w' = 3fF_z w' \Longrightarrow \\
(F_x, F_y) &= (F/w')(\xi', \eta') = 3fF_z(\xi', \eta') = -abG(C_{11}\upsilon_x, C_{22}\upsilon_y)
\end{aligned}
$$

as it should be.

2.2 Examples

2.2.1 Klingel rolling

Klingel rolling is a phenomenon that takes place in railway wheel sets when the wheels are conical. Then the wheel set may be modelled by a rigid double cone, the rails are modelled by two parallel straight rigid knives $2b$ apart, upon which the wheel set is moving. The theory can be extended to more realistic cases.
It is observed that the wheel set is not moving in a straight line, but has rather a sinusoidal motion. The situation is shown in Fig. 10. IT IS REQUIRED TO GIVE AN EXPLANATION OF THIS PHENOMENON, WHICH IS CALLED THE KLINGEL MOTION.

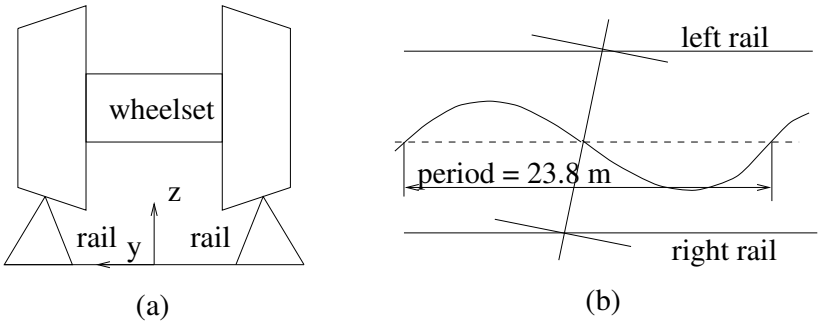


Figure 10: a: The wheel set as a double cone on knife-like rails b: The sinusoidal motion of the centre of the wheel set

A coordinate system is introduced in which the origin lies in the middle between of the rails, the x-axis points in the rolling direction along the rails, the z-axis points vertically upwards, and the y-axis completes the right-handed system, *i.e.* it points in the lateral, left direction if one stands looking in the direction of x. Let the centre of the wheel set have coordinates (x, y, R), where we neglect the height difference of the centre of the wheel set due to the shift of the mid position over a distance y with respect to the rails. The apex angle of the double cone is 2γ, a small angle, the radius of the double cone in the mid-position $y = 0$ is R, so that the radius of the wheels are

$$\begin{aligned} \text{radius of left wheel on rail} &= R + \gamma y \\ \text{radius of right wheel on rail} &= R - \gamma y \end{aligned}$$

$2b$ is the width of the track.
The angular velocity of the wheel set about the y-axis is $\omega > 0$. The angle that the axis of the wheel set makes with the y-axis is α, a small angle. We have, when there is no slip between the wheel set and the rails:

$$\dot{y}/V = \alpha; \quad \dot{\alpha} = -\omega\gamma y/b \quad \Rightarrow \quad \ddot{y} = -\omega\gamma y V/b = -\omega^2\gamma y R/b$$

with

$$V = \omega R = \text{rolling velocity}$$

Consequently,

$$y = y_0 \cos(\omega t\sqrt{\gamma R/b})$$

where y_0 is the amplitude (integration constant), the phase is arbitrary, and the angular frequency is $\omega\sqrt{\gamma R/b}$, hence the frequency is $(\omega\sqrt{\gamma R/b})/(2\pi)\ Hz$. The period on the rails is $2\pi R/\sqrt{\gamma R/b}$, and therefore independent of the angular velocity of the wheel set. A pilot numerical value of the period on the rails is found by setting

$$\begin{aligned} \gamma = 1/40, \ \ R = 0.5 \ \ m, \ \ b = 0.715 \ \ m \quad \Longrightarrow \quad & \text{period} = 23.8 \ \ m \\ & \text{Frequency at 100 km/h} = 1.17 \ \ Hz \end{aligned}$$

2.2.2 A ball between two surfaces

Consider two steel surfaces that are mounted on an axle so that they make an angle of 2γ with one another. A steel ball with radius R is placed between them. The flat surfaces are rotated in such a way that the angle 2γ is conserved; a coordinate system is introduced in which the z-axis points vertically upwards, the y-axis points horizontally outwards, the x-axis completes the right-handed system, and the origin O lies in the centre of the ball. The motion of the planes is so that the origin of the coordinate system is stationary.
IT IS REQUIRED TO FIND THE TOTAL FORCE IN Y-DIRECTION, WHEN THE RADIAL POSITION OF THE ORIGIN OF THE COORDINATE SYSTEM IS GIVEN. The situation is sketched in Fig. 11.

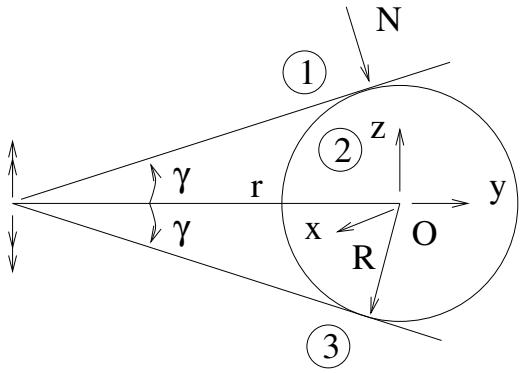


Figure 11: A ball between two surfaces

This problem was investigated by Johnson [5] in 1959.
The angular velocity of the upper plane is ω and that of the lower plane is $-\omega$. The ball does not rotate about the vertical axis z, by the symmetry of the two contacts. The spin creepage at the upper contact, (2)-(1), is $-\omega/V = -1/r$, where r is the distance between the axis of rotation of the two surfaces and the centre of the ball, see Fig. 11 and $\gamma = R/r$ is taken small. V is the velocity of rolling at the upper contact. The spin creepage, (2)-(3), at the lower contact is the same. The radius of the circular contact patch is a, the total normal force is $N = 16a^3G/(3R(1-\nu))$, where R is the radius of the ball. So the total force in y-direction is, according to the linear theory,

$$\begin{aligned} F &= 2N\gamma + 2Ga^3C_{23}/r = N\gamma[2 - 2C_{23}a^3G/(NR)] = \\ &= N\gamma[2 - 2 * 3(1-\nu)/16C_{23}] = 1.6N\gamma \end{aligned}$$

when we take care to stop the lateral creepage.

2.2.3 Creepage and spin for a railway wheel set

A railway wheel set consists of two wheels mounted on an axle. The wheel surfaces consist of a flange, a tread, and a throat in between; the tread has (almost) the form of

a cone, as well as the flange; the throat is more or less circular, see Fig. 12. The rail consists of foot, a web and a rail head, see also Fig. 12.

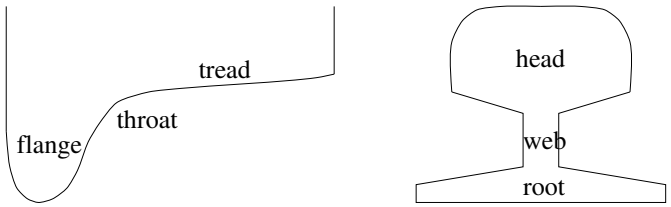


Figure 12: A railway wheel and rail surface

Fig. 13 schematically shows the wheel set standing on the rails. A coordinate system is introduced in which the origin lies in the centre of the wheel set, the z-axis points vertically upwards, the x-axis points along the rails in the rolling direction from left to right, and the y-axis completes the right-handed coordinate system, pointing to the left if one is facing the rolling direction.

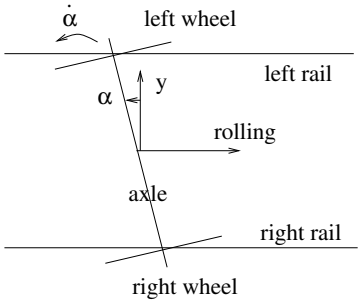


Figure 13: The wheel set standing on the rails

We start with the lateral creepage. Let α be the small angle between the y-axis and the centre line of the wheel set, shown positive in Fig. 13. The lateral creepage consists of two components, one due to α and one due to $\dot{y}$, the lateral velocity of the centre of the wheel set. We have that $\dot{y}/V$ is compensated in part by α:

$$\upsilon_y = \dot{y}/V - \alpha \tag{135}$$

Note that in Klingel rolling, the lateral creepage vanishes. Next we consider the longitudinal creepage. It consists of three parts: the creepage due to the difference of the rolling radii of the wheels, and the rotation of the wheel set about a vertical axis through its centre, and the braking or accelerating motion of the wheel set as a whole. The longitudinal creepage depends on whether we consider the left or the right wheel. The first component reads $\omega(R_{left} - R_{right})/V$, where R_{left} and R_{right} are the radii of left and right wheel, which depend on the lateral coordinate y, and V is the rolling velocity of the centre of the wheel set. It compensates the longitudinal creepage, and it changes sign

with the wheel.
The second component reads $-\dot{\alpha}b/V$; it changes sign with the wheel.
The third component reads $(V-\omega R_{mean})/V$; it does not change sign with the wheel. In total, we have for the longitudinal creepage in the left wheel

$$v_{x,left} = \frac{-\omega(R_{left}-R_{right}) - \dot{\alpha}b + (V-\omega R_{mean})}{V} \tag{136}$$

and in the right wheel

$$v_{x,right} = \frac{\omega(R_{left}-R_{right}) + \dot{\alpha}b + (V-\omega R_{mean})}{V} \tag{137}$$

Finally we consider the spin. It consists of two components, the kinematic spin and the geometric spin.
The kinematic spin is due to the rotation of the wheel set about the vertical axis. It equals $\dot{\alpha}/V$, and it is the same for both wheels.
The geometric spin is due to the conicity of the wheels, see Fig. 14. The conicity of the wheels is $\gamma_{left},\ \gamma_{right}$. The normal at the contact point of the right wheel makes an angle γ_{right} with the z-axis, hence the angular velocity about this normal is $\omega\gamma_{right}$, and the contribution to the spin is $\omega\gamma_{right}/V = \gamma_{right}/R$. Similarly, the angular velocity about the normal at the contact point of the left wheel is $-\omega\gamma_{left}$, and the contribution to the spin on the left wheel is $-\gamma_{left}/R$. The total spin on the wheels is

$$\varphi_{left} = \dot{\alpha}/V - \gamma_{left}/R \tag{138}$$
$$\varphi_{right} = \dot{\alpha}/V + \gamma_{right}/R \tag{139}$$

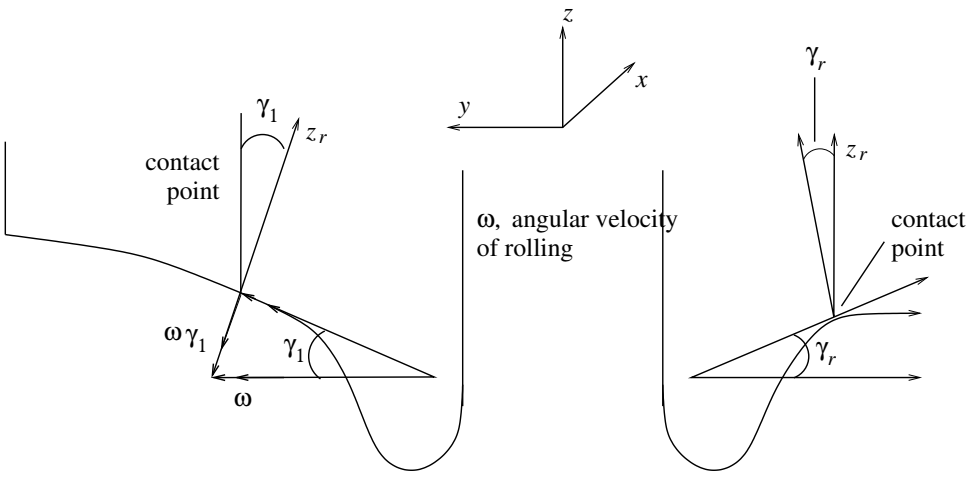


Figure 14: The geometric spin

References of Section 2

[1] J. Boussinesq (1885), *Application des potentiels à l'équilibre et du mouvement des solides élastiques*. Paris, Gauthier-Villars.

[2] V. Cerruti (1882), *Accademia dei Lincei, Roma. Mem. fis. mat.*

[3] G.M.L. Gladwell (1980), *Contact problems in the classical theory of elasticity*. Sijthoff and Noordhoff, Alphen a/d Rijn, the Netherlands.

[4] E. Jahnke, F. Emde (1943), *Tables of functions*. Dover publications, New York.

[5] K.L. Johnson (1959), *The influence of elastic deformation upon the motion of a ball rolling between two surfaces*. Proceedings of the Institution of Mechanical Engineers **173**, p. 795-810.

[6] A.E.H. Love (1926), *A treatise on the theory of elasticity*. 4th Ed. Cambridge University Press.

[7] J.T. Oden, E.B. Pires (1983), *Nonlocal and nonlinear friction laws and variational principles for contact problems in elasticity*. Journal of Applied Mechanics **50**, p. 67-76.

[8] P.J. Vermeulen, K.L. Johnson (1964), *Contact of nonspherical bodies transmitting tangential forces*. Journal of Applied Mechanics **31**, p. 338-340.

[9] H. Hertz (1882a), *Über die Berührung fester elastischer Körper*. In: H. Hertz, Gesammelte Werke, Band 1. Leipzig: J.A. Barth, (1895), p. 155-173.

[10] H. Hertz (1882b), *Über die Berührung fester elastischer Körper und über die Härte*. In: H. Hertz, Gesammelte Werke, Band 1. Leipzig: J.A. Barth, (1895), p. 174-196.

3 The simplified theory of rolling contact

3.1 Discretization of the slip

We recall the definition of the slip, see (25):

$$\begin{aligned} \mathbf{s} &= \mathbf{c} - (\partial \mathbf{u}/\partial \mathbf{x})\mathbf{v} + \partial \mathbf{u}/\partial t \qquad (140)\\ \text{with} & \\ \mathbf{s} &= \text{the slip: the velocity of (1) over (2)} \\ \mathbf{c} &= \text{the creep: the rigid velocity of (1) over (2)} \ = \dot{\mathbf{x}}^1 - \dot{\mathbf{x}}^2 \\ \mathbf{u} &= \text{the surface displacement difference} \ = \mathbf{u}^1 - \mathbf{u}^2 \\ \mathbf{x} &= \text{the position} \\ \mathbf{v} &= \text{the rolling velocity} \ = -(\dot{\mathbf{x}}^1 + \dot{\mathbf{x}}^2) \\ t &= \text{the time} \\ \dot{\mathbf{x}}^a &= \text{velocity of particle } \mathbf{y}^a \text{ with respect to contact patch} \end{aligned}$$

We consider $\mathbf{u}(\mathbf{x} + k\mathbf{v}, t - k), k > 0$ which we expand about $k = 0$, taking along only the first two terms:

$$\mathbf{u}(\mathbf{x} + k\mathbf{v}, t - k) = \mathbf{u}(\mathbf{x}, t) + k\{(\partial \mathbf{u}/\partial \mathbf{x})\mathbf{v} - \partial \mathbf{u}/\partial t\} + O(k^2)$$

If we neglect the $O(k^2)$, then we obtain

$$\begin{aligned} \mathbf{s} &= \mathbf{c} + (\mathbf{u} - \mathbf{u}')/k \qquad (141)\\ \text{where} & \\ \mathbf{u} &= \mathbf{u}(\mathbf{x}, t) \\ \mathbf{u}' &= \mathbf{u}(\mathbf{x} + k\mathbf{v}, t - k) \end{aligned}$$

In principle, $\mathbf{u}'$ is known in a non-steady state, where $\mathbf{u}$ evolves in time under a non-constant creep $\mathbf{c}(t)$.
In a steady state, $\mathbf{u}$ and $\mathbf{u}'$ are independent of explicit time, and

$$\mathbf{u} = \mathbf{u}(\mathbf{x}), \mathbf{u}' = \mathbf{u}(\mathbf{x} + k\mathbf{v}), \qquad (142)$$

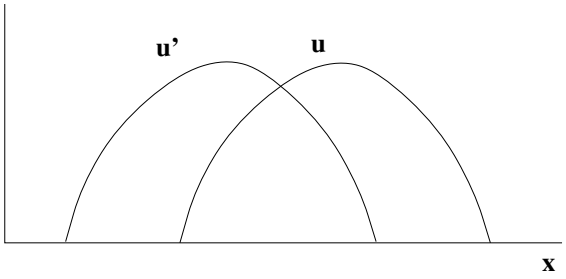


Figure 15: $\mathbf{u}$ and $\mathbf{u}'$ in a steady state

see Fig. 15. (141) is the promised discretization of the slip. It holds for both steady and non-steady rolling. In the steady state, $\mathbf{u}'$ is known and $\mathbf{u}$ is unknown, in the steady state both are unknown, but the solution is independent of explicit time.

3.2 Simplified theory

In the simplified theory of rolling contact we approximate the relation between the tangential surface displacement (u^a, v^a) of body a and the tangential surface traction on body a, *viz.* (p_x^a, p_y^a) by

$$(u^a, v^a) = L^a(p_x^a, p_y^a) \tag{143}$$

where L^a is a parameter called the FLEXIBILITY. The flexibility is comparable to $1/E^a$, where E^a is the modulus of elasticity of body a.

To imagine this load displacement law, we think of a bed of springs, see Fig. 16.

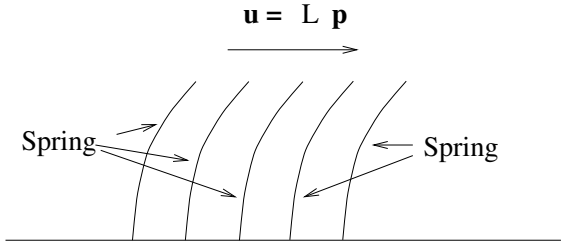


Figure 16: The body as a bed of springs

This is the basic hypothesis of the Simplified Theory.

Important remark. We note that the normal displacement is not approximated by a relationship like (143). This is because the lack of accuracy in the normal simplified relationship. Instead, we use the Hertz theory, whereby the contact becomes elliptical and the normal traction distribution semi-ellipsoidal.
Later, we will go further and approximate the semi-ellipsoidal normal traction by a

paraboloidal one, on the same axes, and with the same total normal force.

We consider the relation between the surface traction (p_x^1, p_y^1) and (p_x^2, p_y^2), and we prove that they are opposite:

$$(p_x^1, p_y^1) = -(p_x^2, p_y^2) \doteq (p_x, p_y) \tag{144}$$

Indeed, outside contact $(p_x^1, p_y^1) = -(p_x^2, p_y^2) = (0, 0)$, while inside contact, by Newton's Third Law, $(p_x^1, p_y^1) = -(p_x^2, p_y^2)$. This proves proposition (144).

As we saw before, a key role is played by the displacement difference. It is

$$(u, v) \doteq (u^1 - u^2, v^1 - v^2) = (L^1 + L^2)(p_x, p_y) \doteq L(p_x, p_y) \tag{145}$$

We consider the flexibility L. It is comparable to $1/E$, where E is Young's modulus. Like E, it depends on the material of the bodies (steel, *e.g.*) but unlike E it depends on the form of the bodies, and also on the loading:

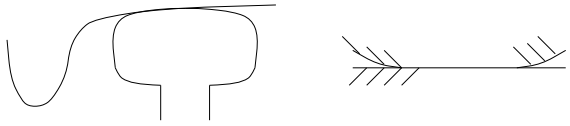


Figure 17: Flat bodies of various forms

$$\begin{aligned} L &= L(\text{depth/contact length},\ (p_x, p_y),\ \text{material}) \\ E &= E(\text{material}),\ \text{by contrast} \end{aligned}$$

Flat bodies of various forms are shown in Fig. 17. We consider the wheel-rail system. The contact area is about 1 cm long; the wheel is about 10 cm wide, the rail head is about 5 cm thick; depth/contact length $> 5 \approx \infty$ in elasticity! The material is always steel, so $L = L[(p_x, p_y)]$. Analysis of (p_x, p_y): p_x and p_y depend strongly on the form and size of the contact patch, that is on the semi-axes of the contact ellipse. Also, we will consider three special loadings, (1),(2),(3), with flexibilities $L_1(a, b)$,$L_2(a, b)$, $L_3(a, b)$. From these we will later make one single $L(a, b, (1), (2), (3))$.

Summarized so far. We consider two bodies in contact. Only surface quantities are of interest. We have a normal problem $(w^a; p_z^a)$ and a tangential problem $(u^a, v^a; p_x^a, p_y^a)$, The normal problem is solved by Hertz: the contact patch is elliptical, with semi-axes a, b; the pressure distribution is semi-ellipsoidal.

The tangential problem approximately satisfies the hypothesis of the simplified theory:

$$(u^a, v^a) = L^a(p_x^a, p_y^a)$$

but

$$\begin{aligned}(p_x^1, p_y^1) &= -(p_x^2, p_y^2) = (p_x, p_y) \Longrightarrow \\ (u^1, v^1) &= L^1(p_x, p_y) \\ (u^2, v^2) &= -L^2(p_x, p_y)\end{aligned}$$

and the displacement difference is

$$\begin{aligned}(u, v) &\doteq (u^1 - u^2, v^1 - v^2) = (L^1 + L^2)(p_x, p_y) \doteq L(p_x, p_y) \\ L &= L(a, b, (1), (2), (3))\end{aligned}$$

with (1),(2),(3) three special loadings.

3.2.1 Coulomb's Law

We recall Coulomb's Law:

$$\begin{aligned}g &= fp_z; \quad |(p_x, p_y)| < g \quad \Longrightarrow \\ (S_x, S_y) &= \text{local velocity of (1) over (2)} = (0, 0); \qquad (146) \\ |(S_x, S_y)| &> 0 \quad \Longrightarrow \\ (p_x, p_y) &= -g(S_x, S_y)/|(S_x, S_y)| \qquad (147)\end{aligned}$$

with

$$\begin{aligned}(S_x, S_y) &= (\upsilon_x - \varphi y, \upsilon_y + \varphi x) - \partial(u, v)/\partial x + \partial(u, v)/\partial t \\ &= (\upsilon_x - \varphi y, \upsilon_y + \varphi x) + [(u, v) - (u', v')]/k \qquad (148) \\ u &= u((x, y), t); u' = u((x + k, y), t - k); (\mathbf{v}^T = (1, 0, 0)). \qquad (149)\end{aligned}$$

3.2.2 The linear theory

In the linear theory, the stick area H covers the entire contact area C; as then $|(p_x, p_y)| < g$, this corresponds to small tractions. In steady state rolling we have

$$\begin{aligned}(0, 0) &= (S_x, S_y) = (\upsilon_x - \varphi y, \upsilon_y + \varphi x) - \partial(u, v)/\partial x \quad \Longrightarrow \\ u &= \upsilon_x x - \varphi xy + k(y) \qquad (150) \\ v &= \upsilon_y x + \varphi x^2/2 + l(y) \qquad (151)\end{aligned}$$

where k and l are arbitrary functions of y which have the character of integration constants: their derivative with respect to x vanishes.
With the simplified theoretic hypothesis $(u, v) = L(p_x, p_y)$ we obtain

$$\begin{aligned}(p_x, p_y) &= (\upsilon_x x - \varphi xy + k(y), \upsilon_y x + \varphi x^2/2 + l(y))/L, \ \mathbf{x} \in C \\ &= (0, 0), \ \mathbf{x} \notin C \qquad (152)\end{aligned}$$

The question arises how to determine k and l. It is answered as follows.
A particle lies in front of the leading edge of the contact patch, and is unloaded. It moves until it reaches the leading edge, still unloaded. It enters the contact patch, and traction starts building up from zero. Traction keeps building up until the traction bound is reached. Then the traction remains on the traction bound, and slip sets in. At the trailing edge, slip prevails and the particle leaves the contact patch, again unloaded. In linear theory, the traction bound is never reached, but at the trailing edge the traction is suddenly released till zero.

We denote the leading edge by $x_L = a(y) > 0$, and the trailing edge is then $x_T = -a(y) < 0$. At the leading edge we have

$$\begin{aligned} p_x(a(y), y) &= \upsilon_x a(y) - \varphi a(y) y + k(y) = 0 \Longrightarrow \\ k(y) &= -\upsilon_x a(y) + \varphi a(y) y \Longrightarrow \\ p_x(x, y) &= \upsilon_x [x - a(y)] - \varphi [x - a(y)] y \ \text{ inside } C \\ &= 0 \ \text{ outside } C \end{aligned} \tag{153}$$

Similarly

$$\begin{aligned} p_y(x, y) &= \upsilon_y [x - a(y)] + \varphi [x^2 - a(y)^2]/2 \ \text{ inside } C \\ &= 0 \ \text{ outside } C \end{aligned} \tag{154}$$

where

$$C = \{(x, y, z) | z = 0, (x/a)^2 + (y/b)^2 \leq 1\} \tag{155}$$

$$a(y) = a\sqrt{1 - (y/b)^2} \tag{156}$$

We can calculate the total force due to these loadings:

$$F_x = \int_{-b}^{b} \int_{-a(y)}^{a(y)} p_x(x, y) dx dy = -8a^2 b \upsilon_x / (3L) \tag{157}$$

$$\begin{aligned} F_y &= \int_{-b}^{b} \int_{-a(y)}^{a(y)} p_y(x, y) dx dy = \\ &= -8a^2 b \upsilon_y / (3L) - \pi a^3 b \varphi / (4L) \end{aligned} \tag{158}$$

This is the force calculated by the simplified theory. On the other hand, F_x and F_y have been calculated by the true theory of elasticity. Such a theory will be termed the Exact or Complete theory. We had found

$$\begin{aligned} F_x &= -abGC_{11}\upsilon_x \\ F_y &= -abGC_{22}\upsilon_y - (ab)^{1.5}GC_{23}\varphi \end{aligned} \tag{159}$$

where the C_{ij} are the creepage and spin coefficients tabulated in Table 3 of Section 2 which depend only on (a/b) and Poisson's ratio ν. G is the modulus of rigidity which can be expressed in Young's modulus E and ν:

$$G = \frac{E}{2(1+\nu)} \tag{160}$$

3.2.3 The flexibility parameter

We had seen that the linear theory provides a link between the simplified and the exact theories. We can use that link to calculate three values of the flexibility parameter L. Indeed,

$$\begin{aligned} \text{Simplified theory: } F_x &= -8a^2b\upsilon_x/(3L) \\ F_y &= -8a^2b\upsilon_y/(3L) - \pi a^3 b\varphi/(4L) \\ \text{Exact theory: } F_x &= -abGC_{11}\upsilon_x \\ F_y &= -abGC_{22}\upsilon_y - (ab)^{1.5}GC_{23}\varphi \end{aligned}$$

Equating the coefficients of $\upsilon_x, \upsilon_y, \varphi$ in simplified and exact theory gives three values of L:

$$(\upsilon_x) \quad : \quad L_1 = \frac{8a}{3GC_{11}} \tag{161}$$

$$(\upsilon_y) \quad : \quad L_2 = \frac{8a}{3GC_{22}} \tag{162}$$

$$(\varphi) \quad : \quad L_3 = \frac{\pi a^2}{4G\sqrt{ab}C_{23}} \tag{163}$$

We add a small table valid for $\nu = 0.25$; we see that these values differ considerably.

a/b	0.1	0.3	1.0	1/.3	1/.1
$GL_1/a = 8/(3C_{11})$	0.806	0.775	0.647	0.421	0.228
$GL_2/a = 8/(3C_{22})$	1.06	0.970	0.784	0.417	0.208
$GL_3/a = \pi a/[4(ab)^{0.5}C_{23}]$	0.525	0.602	0.534	0.352	0.170

Table 4: $L(a, b, (1), (2), (3))$

The dependence on a/b is very marked; but between the L_i for constant a/b there are also large differences, especially between L_2 and L_3 which both refer to F_y. We form

a single value of L as a weighted mean of the L_i:

$$L = \frac{L_1|v_x| + L_2|v_y| + L_3|\varphi|\sqrt{ab}}{\sqrt{v_x^2 + v_y^2 + ab\varphi^2}} \tag{164}$$

Clearly, when

$$\begin{aligned} v_x = v_y = 0 &\Rightarrow L = L_3 \\ v_y = \varphi = 0 &\Rightarrow L = L_1 \\ \varphi = v_x = 0 &\Rightarrow L = L_2 \end{aligned}$$

as it should be.

3.2.4 The traction bound

According to the Hertz theory, the normal traction has the following form:

$$p_z(x,y) = Z_0\sqrt{1-(x/a)^2-(y/b)^2}$$

with Z_0 constant; the pressure is semi-ellipsoidal.

The simplified theoretic analogue of the Hertz theory exists, but it has grave defects; yet its normal pressure is extremely interesting, because, with it, the simplified theory becomes a consistent whole. To find this adapted pressure, we take a and b from the Hertz theory, and the form of the normal pressure from simplified theory:

$$p'_z(x,y) = Z'_0\{1-(x/a)^2-(y/b)^2\} \quad \text{(no root!)} \tag{165}$$

Z_0 is known from the Hertz theory; Z'_0 must be adapted. We do this so that the total compressive forces F_z and F'_z are equal:

$$\begin{aligned} F_z &= \int\int_C Z_0\sqrt{1-(x/a)^2-(y/b)^2}dxdy = \\ &= 2\pi ab Z_0/3 \Longrightarrow \\ Z_0 &= \frac{3F_z}{2\pi ab} \\ F'_z &= \int\int_C Z'_0\{1-(x/a)^2-(y/b)^2\}dxdy = \\ &= \pi ab Z'_0/2 \longrightarrow \\ Z'_0 &= \frac{2F_z}{\pi ab} \end{aligned}$$

So

$$p_z(x,y) = \frac{3F_z}{2\pi ab}\sqrt{1-(x/a)^2-(y/b)^2} \tag{166}$$

$$p'_z(x,y) = \frac{2F_z}{\pi ab}\{1-(x/a)^2-(y/b)^2\} \tag{167}$$

Both can be used in the traction bound $g = fp_z$. We choose p'_z for reasons of consistency – once simplified theory, always simplified theory.

3.2.5 An analytical solution

As we saw, the tangential traction due to pure longitudinal creepage ($\upsilon_y = \varphi = 0$) has the following form in no-slip theory:

$$\begin{aligned} p_x(x,y) &= (x - a(y))\upsilon_x / L \\ a(y) &= a\sqrt{1-(y/b)^2} \end{aligned} \tag{168}$$

The traction bound reads

$$\begin{aligned} g &= \frac{2fF_z}{\pi ab}\{[1-(y/b)^2] - (x/a)^2\} = \\ &= \frac{2fF_z}{\pi a^3 b}\{a(y)^2 - x^2\} \end{aligned} \tag{169}$$

So, the picture is as shown in Fig. 18:

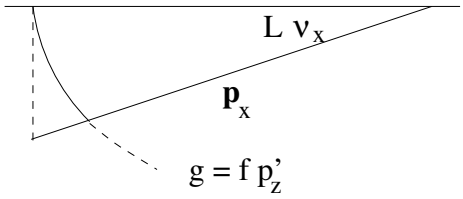


Figure 18: Traction distribution for pure longitudinal creepage in simplified theory

Near the trailing edge $-a(y)$ Coulomb's Law is broken by the linear theory. So the linear theory is never really valid, but only approximately so when $\upsilon_x \to 0$. Also shown in the figure is the nonlinear simplified theory. The exact solution is shown in Fig. 19. The solution for pure and combined lateral and longitudinal creepage is analogous.

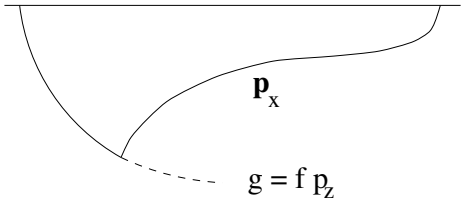


Figure 19: Exact theoretic traction distribution for pure longitudinal creepage

Summarizing, the simplified theoretic solution for the case $\varphi = 0$ is analytically known. For the case of nonvanishing φ we need a numerical theory, F A S T S I M.

3.2.6 Fastsim

FASTSIM is an algorithm to determine the traction in the general case (notably non-vanishing φ). We recall Eq. (141), where we insert the simplified theoretic hypothesis:

$$\begin{aligned} \mathbf{s} &= \mathbf{c} + (\mathbf{u} - \mathbf{u}')/k \\ &= \mathbf{c} + L(\mathbf{p} - \mathbf{p}')/k, \end{aligned}$$

where

$$\begin{aligned} \mathbf{c} &= \text{creep} \quad \dot{\mathbf{x}}_1 - \dot{\mathbf{x}}_2 \\ \mathbf{p} &= \text{present traction} \quad = (p_x[(x,y),t], p_y[(x,y),t])^T \\ \mathbf{p}' &= \text{previous traction} \quad = \mathbf{p}(\mathbf{x} + \mathbf{v}k), t - k) \end{aligned}$$

We follow a particle on its path. At the leading edge $a(y,t)$, $\mathbf{p} = 0$, that is, known. We work by induction. Suppose $\mathbf{p}'$ is known. What is $\mathbf{p}$?

$$\mathbf{p} = \mathbf{p}' + (k/L)[\mathbf{s} - \mathbf{c}]$$

On the right-hand side, only the slip $\mathbf{s}$ is unknown. The rigid slip $\mathbf{c}$ is, of course, known. Tentatively we set $\mathbf{s} = 0$.

$$\mathbf{p}_H \doteq \mathbf{p}(\mathbf{x})_{s=0} = \mathbf{p}' - (k/L)\mathbf{c}$$

We determine $|\mathbf{p}_H|$, and compare with (known)g.

- $\mathbf{p}_H \leq g$: set $\mathbf{p} = \mathbf{p}_H$;
  then $|\mathbf{p}| \leq g$, and $\mathbf{s} = 0$: COULOMB.
- $\mathbf{p}_H > g$: set $\mathbf{p} = (g/|\mathbf{p}_H|)\mathbf{p}_H$. Then $|\mathbf{p}| = g$, and

$$\begin{aligned} \mathbf{s} &= \mathbf{c} + (L/k)(\mathbf{p} - \mathbf{p}') = \\ &= (L/k)(-\mathbf{p}_H + \mathbf{p}) \\ &= -(L/k)\mathbf{p}_H(1 - [g/|\mathbf{p}_H|]) \end{aligned}$$

  so that $\mathbf{s}$ is exactly opposite $\mathbf{p}$: COULOMB.

SUMMARY OF THE ALGORITHM for steady-state rolling:
$\mathbf{p}(\mathbf{x}, t) = \mathbf{p}(\mathbf{x})$.

$$\begin{aligned} \text{given} \quad &: \\ \mathbf{x} &= (x, y) \\ \mathbf{x} + k\mathbf{v} &= (x + q, y) \\ m &= \text{number of } x\text{-intervals} \\ n &= \text{number of } y\text{-intervals} \\ V &= |\mathbf{v}| = \text{rolling velocity}, > 0 \\ L &= \text{flexibility}, > 0 \end{aligned}$$

$$
\begin{aligned}
g &= g(x,y) \text{ traction bound at } (x,y) \in C \\
\mathbf{c}(x,y) &= \text{rigid slip at } (x,y) \in C \\
a(y) &= \text{leading edge} \\
-a(y) &= \text{trailing edge} \\
a, b &: \text{semi-axes of contact ellipse} \\
\mathbf{p}_H, \mathbf{p} &: \text{tangential traction, to be calculated} \\
\mathbf{F} &: \text{total tangential force, to be calculated.}
\end{aligned}
$$

THE ALGORITHM:

1. set $r = 2b/n$; $y = b - r/2$; $\mathbf{F} = 0$ (Initiation of programme and y-loop)
2. $q = 2a(y)/m$; $x = a(y) - q$; $p = p(x + q, y) = 0$ (Initiation of x-loop).
3. $\mathbf{p}' = \mathbf{p}$ ($\mathbf{p}'$ is $\mathbf{p}$ just found).
4. $$\mathbf{p} = \mathbf{p}' - [q/VL]\mathbf{c}(x + q/2, y)$$
   (Form $\mathbf{p}_H$; $(x + q/2)$ is point between x and $x + q$)
5. If $|\mathbf{p}| > g$ then $\mathbf{p} = (g/|\mathbf{p}|)\mathbf{p}$
   (Form $\mathbf{p}$ if the traction bound is exceeded).
   Otherwise, $|\mathbf{p}| \leq g$, we are in the stick area, and
   $\mathbf{p}(x, y) = \mathbf{p} = \mathbf{p}_H$.
6. $\mathbf{p}(x, y) = \mathbf{p}$; $\mathbf{F} = \mathbf{F} + qr\mathbf{p}$
   (Fill in $\mathbf{p}(x, y)$; update $\mathbf{F}$).
7. $x = x - q$; if $x > -a(y)$ go to 3 (Test of x-loop)
8. $y = y - r$; if $y > -b$ go to 2 (Test of y-loop)
9. READY

3.2.7 Results

We show some results.

First: The regions of stick (A) and slip (S) for various creepages and spins (qualitative). These regions were obtained with a parabolic traction bound (Fig. 20). They are quite close to the areas of slip and stick obtained with the exact theory. Using an ellipsoidal traction bound gives bad results.

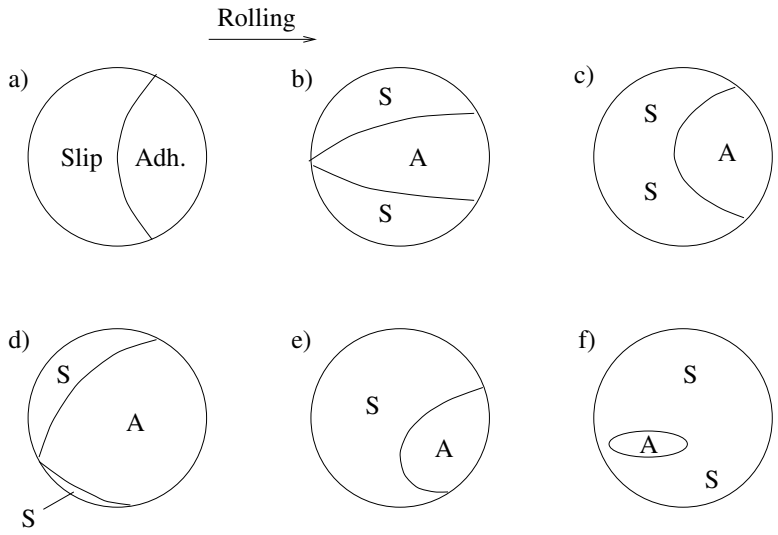


Figure 20: Areas of stick (A) and slip (S). a): Pure creepage ($\varphi = 0$); b): Pure spin, small; ($v_x = v_y = 0$); c): Lateral creepage with spin ($v_x = 0$); d): Longitudinal creepage with spin ($v_y = 0$); e): General case; f): Large pure spin. (Source: Kalker [1])

Next, we show a specific case of a calculation by Fastsim. We see areas of slip (shaded) and stick, an element division, and the traction (arrows). It is a case of pure spin (Fig. 21).

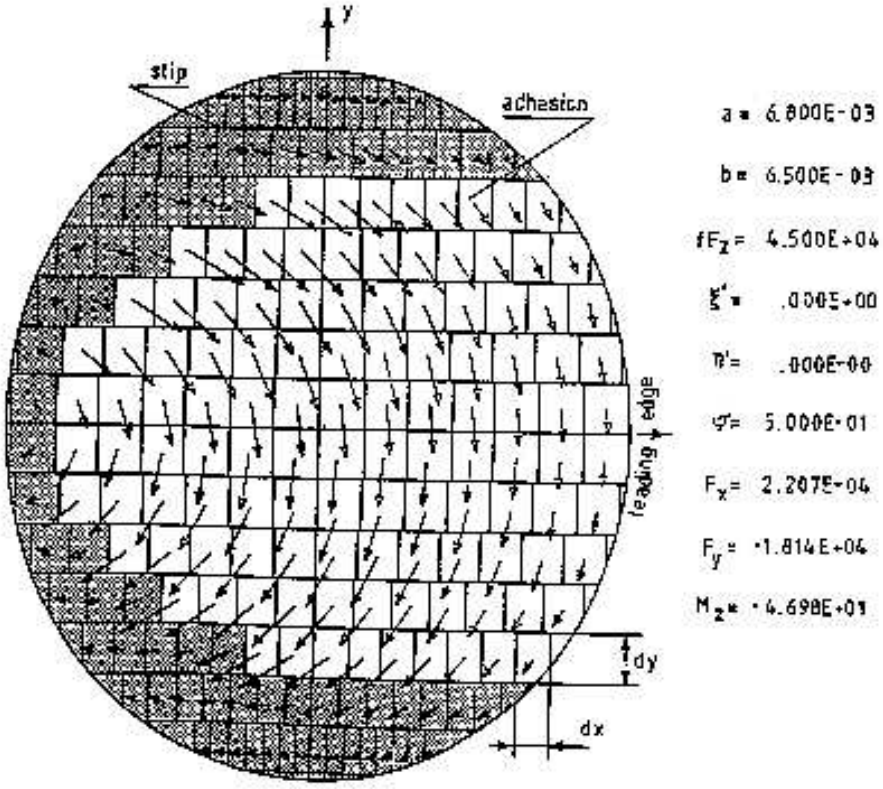


Figure 21: Result of a calculation by Fastsim. a=6.8 mm. b=6.5 mm. $fF_z = 45000$ N. $v_x = v_y = 0$; $\varphi = 0.00054mm^{-1}$. $F_x = 0$, $F_y = -1.812e4N$. (Source: Kalker and Piotrowski [2])

Next, we show the total force due to longitudinal creepage and lateral creepage without spin in comparison with the exact theory (Fig. 22), and the same for pure spin (Fig. 23). The curves have been scaled so that in each figure the slope in the origin is the same. Indeed we have:

$$\begin{aligned} \xi' &= -\frac{abGC_{11}v_x}{3fF_z} \\ \eta' &= -\frac{abGC_{22}v_y}{3fF_z} \\ w' &= \sqrt{\xi'^2 + \eta'^2} \\ \psi &= -\frac{(ab)^{1.5}GC_{23}\varphi}{fF_z} \end{aligned}$$

The correspondence is striking, especially in the pure creepage case (Fig. 22).

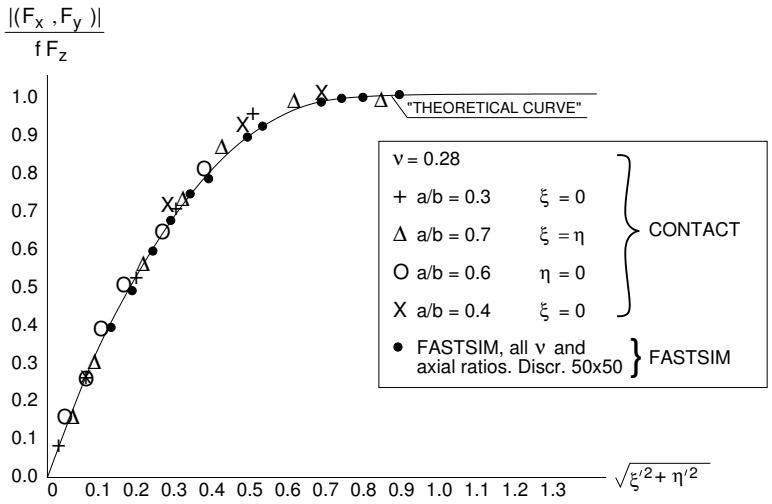


Figure 22: The total force due to pure creepage. Comparison of Fastsim and exact theory. (Source: Kalker [1]).

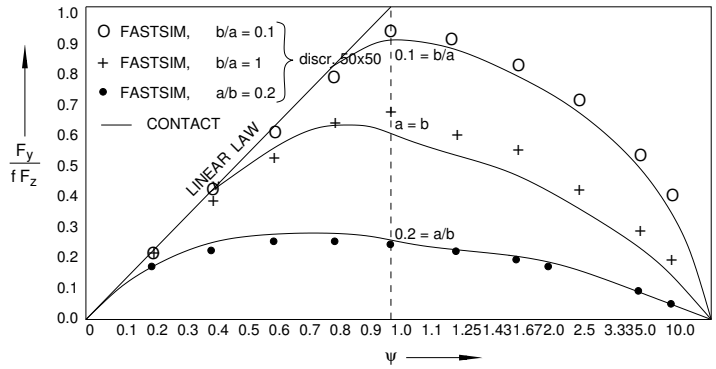


Figure 23: The total force due to pure spin. Comparison of Fastsim (dots) and exact theory (Contact) (Source: Kalker [1]).

Finally we show the total force in some cases where only one of the creepages vanishes, Fig. 24. It is seen that $v_y = -\varphi$, $v_x = 0$ is pretty bad, but we think that it is one of the worst cases. The others are quite good.

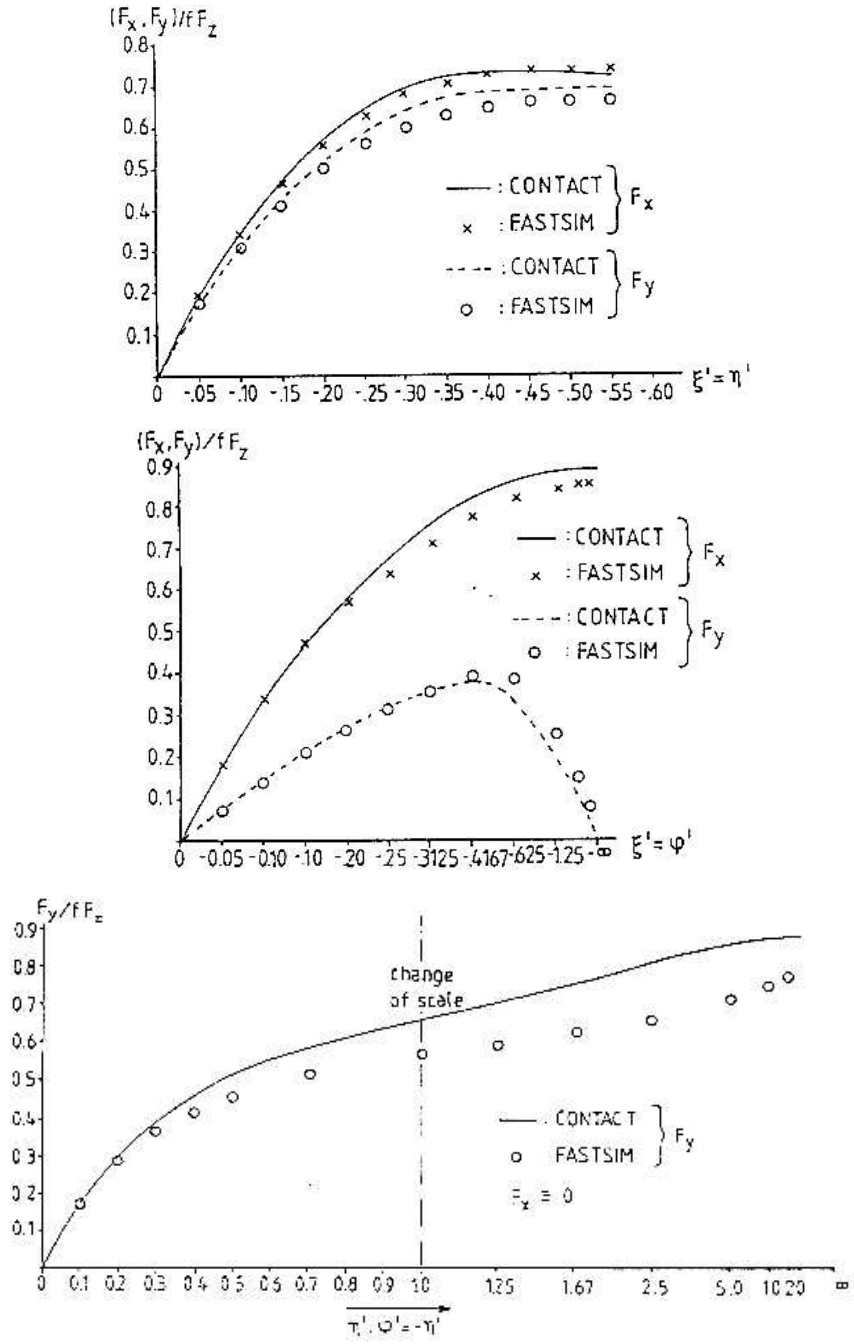


Figure 24: Some curves when only one of the creepages vanishes. The total force for $f = 1$, $F_z = 1$, $a = b = 1$, $G = 1$, $\nu = 0.25$, and (1) $\xi' = \eta'$, $\psi = 0$; (2) $\xi' = \psi$, $\eta' = 0$; (3) $\eta' = -\psi$, $\xi' = 0$. (Source: Kalker [1]).

3.2.8 Summary of Simplified Theory

We split the problem into two subproblems, *viz.* the normal problem and the tangential problem. The normal problem is solved by Hertz, the tangential problem by simplified theory.

In Simplified Theory, we set $\mathbf{u} = L\mathbf{p}$, pictured by a bed of springs (brush model, mattress model). Coulomb's Law is adopted.

Linear theory, *i.e.* the theory in which the slip vanishes, was studied as an example of simplified theory. Its development closely parallels the exact theoretic model. The comparison of simplified and exact linear theories yields the value of L, the flexibility parameter, in its dependence on the modulus of rigidity G, the semi-axes of the contact ellipse a and b, and the creepages v_x, v_y and φ:

$$L = L(G, a, b : v_x, v_y, \varphi).$$

In fact, the linear theory is only approximate, it holds for small creepages.

When the spin vanishes, one can develop an analytical solution of simplified theory, indicated in the text. When the spin does not vanish, one needs a numerical method: FASTSIM.

Results show the division of the contact area into regions of stick and slip; the behaviour of the total force when only one creepage component is non-zero, and some cases when only one component of the creepage vanishes. We see that Simplified Theory may contain an error of roughly 10-15 percent of the maximum force fF_z, accurate enough for many needs.

IT NEEDS TO BE EMPHASIZED THAT SIMPLIFIED THEORY CAN ONLY BE USED FOR QUASIIDENTICAL BODIES.

References of Section 3

[1] J.J. Kalker (1990), *Three-dimensional elastic bodies in rolling contact*. Kluwer, Dordrecht.

[2] J.J. Kalker, J. Piotrowski (1989), *Some new results in rolling contact*. Vehicle System Dynamics **18**, p. 223-242.

4 The variational theory of contact

In this Section, we treat the variational theory of contact. It was introduced by Fichera [2] and Duvaut and Lions [1]. We follow the description of Kalker's book [4]. The variational theory of contact is an alternative to the conventional boundary condition description of contact of Section 1. In fact, the conventional description follows from it, and vice versa. We base ourselves on the Principle of Virtual Work and its dual, the Principle of Complementary Virtual Work from which we derive variational inequalities. The point of departure for both principles are the equations of equilibrium:

$$\sigma_{ij,j} + f_i - \rho\ddot{u}_i = 0, i, j = 1, 2, 3, \text{ summation convention} \tag{170}$$

where

$$\begin{array}{rcl} u_i & : & \text{the displacement} \\ \sigma_{ij} & : & \text{the linearized stress} \\ f_i & : & \text{the body force per unit volume} \\ (\dot{}) & = & d/dt, \quad t\text{: time; particle fixed time differentiation} \\ {}_{,i} & = & \partial/\partial x_i \end{array}$$

in two bodies that occupy the volumes V^1 and V^2.

4.1 The Variation of a Function

Consider a function f of one or more variables. f is subject to certain constraints. If it satisfies these constraints, it is called feasible. Consider another function δf; this is a function called the variation of f. It may be arbitrary (bilateral variation), or arbitrary nonnegative or nonpositive (nonnegative or nonpositive unilateral variation). At any rate, when f is feasible, $f + \epsilon.\delta f$ must be feasible also for ϵ small enough, positive.

Example. Let f be subject to the constraint $f \geq 0$ and let actually $f > 0$. Then for any δf, $f + \epsilon.\delta f > 0$ for ϵ small enough. So the variation δf is bilateral. Now let actually $f = 0$. Then for ϵ small enough, positive δf must be nonnegative: *i.e.* δf is nonnegative, unilateral.

Example. Let f be fully prescribed. Then $\delta f = 0$.

4.2 Virtual Work

We multiply the equation (170) with minus the variation of u_i: $-\delta u_i$ and integrate over V^1, V^2:

$$0 = -\sum_{a=1,2} \int_{V^a} (\sigma_{ij,j} + f_i - \rho\ddot{u}_i)\delta u_i dV \tag{171}$$

The integrand on the right-hand side is the virtual work done on a particle of volume dV, so that (171) is the virtual work equation. This is equivalent to

$$0 = \sum_{a=1,2} [- \int_{V^a} (\sigma_{ij,j} + f_i - \rho \ddot{u}_i) \delta u_i dV + \int_{\partial V^a} p_i \delta u_i dS]$$
$$- \sum_{a=1,2} \int_{\partial V^a} p_i \delta u_i dS \tag{172}$$

with

dV	:	element of volume
dS	:	element of surface
p_i	=	$\sigma_{ij} n_j$, surface load on ∂V, body number omitted
n_j	:	outer normal on V at ∂V.

In the third term of (172) we introduce the boundary conditions of Section 1, subsection 1.7,

$$\begin{aligned} u_i &= \overline{u}_i, \text{ prescribed displacement in region } A_u \subset \partial V \\ &\Rightarrow \delta u_i = 0 \text{ on } A_u \qquad (173) \\ p_i &= \overline{p}_i, \text{ prescribed load in region } A_p \subset \partial V \qquad (174) \end{aligned}$$

In the potential contact area $A_c^1 \approx A_c^2 \approx A_c$ we have:

$$\begin{aligned} p_i^1 &= -p_i^2 \doteq p_i \text{ in } A_c, \text{ (Newton's Third Law)} \\ &\Rightarrow p_i^1 \delta u_i^1 + p_i^2 \delta u_i^2 = p_i \delta u_i \qquad (175) \end{aligned}$$

with $u_i \doteq (u_i^2 - u_i^2)$, displacement difference.

This gives for (172):

$$\begin{aligned} 0 &= \sum_{a=1,2} [- \int_{V^a} (\sigma_{ij,j} + f_i - \rho \ddot{u}_i) \delta u_i dV + \int_{\partial V^a} p_i \delta u_i dS] + \\ &- \sum_{a=1,2} [\int_{A_p^a} \overline{p}_i \delta u_i dS] - \int_{A_c} p_i \delta u_i dS \qquad (176) \end{aligned}$$

In the potential contact A_c we introduce a right-handed orthogonal curvilinear coordinate system x, y: they are represented by Greek indices which run through the values x, y. We introduce a coordinate z along the inner normal to body 1 at (x, y). dS is the element of area at (x, y). Then we may write

$$p_i \delta u_i = p_z \delta u_z + p_\tau \delta u_\tau \tag{177}$$

with

p_z	:	normal pressure, positive if compressive
p_τ	:	tangential traction

We consider the deformed distance and the slip.

The deformed distance $e = h + u_z$; h is prescribed, so $p_z \delta u_z = p_z \delta e$. Now, as we saw in Section 1, if $e > 0$ then $p_z = 0$ (outside contact). If $e = 0$ (inside contact), then $p_z \geq 0$ (compression). e cannot be negative, so, if $e = 0$ then $\delta e \geq 0$, since varied quantities must be feasible. Thus if the contact conditions are satisfied, then

$$p_z \delta e \geq 0 \quad \text{in } A_c \quad \text{sub } e \geq 0 \tag{178}$$

where 'sub' means 'subject to the auxiliary condition(s)'.
The contact formation can be summarized as follows:

$$p_z \geq 0, \quad p_z e = 0, \quad e \geq 0 \tag{179}$$

The contact area does not occur explicitly in the formulations (178) and (179).

The slip (*i.e.* the velocity of body 1 over body 2) is given by

$$s_\tau = w_\tau + \dot{u}_\tau \tag{180}$$

with

$$u_\tau = u_\tau^1 - u_\tau^2 \tag{181}$$
$$w_\tau = \dot{x}_\tau^1 - \dot{x}_\tau^2 \tag{182}$$

u_τ is called the displacement difference, and w_τ is the rigid slip, which is defined as the local velocity of body 1 over body 2 when both are regarded as rigid.

We integrate (180) from time t' to time t, where $t' < t$. We call

$$\int_{t'}^{t} s_\tau(x_\alpha; q) dq = E_\tau \approx (t - t') s_\tau \quad \text{(local) shift} \tag{183}$$
$$\int_{t'}^{t} w_\tau(x_\alpha; q) dq = W_\tau \approx (t - t') w_\tau \quad \text{(local) rigid shift} \tag{184}$$

and we denote the displacement difference

$$u_\tau = u_\tau(x_\alpha; t) \tag{185}$$
$$u'_\tau = u_\tau(x_\alpha; t') \tag{186}$$

Note that u'_τ is the displacement difference at time t', not a derivative, and that the coordinate system is particle fixed.
The integral of (180) is

$$E_\tau = W_\tau + u_\tau - u'_\tau \tag{187}$$

We consider an evolution: we proceed stepwise, from t' to t. That means: u'_τ, the rigid slip w_τ and the rigid shift E_τ are all known: they control the evolution from t' to t. So we have

$$\delta E_\tau = \delta u_\tau \quad \Rightarrow \quad p_\tau \delta u_\tau = p_\tau \delta E_\tau \tag{188}$$

Assume there is slip ($|s_\tau| \neq 0$ so that $|E_\tau| \neq 0$), and

$$p_\tau = -gE_\tau/|E_\tau|, \quad |E_\tau| = \sqrt{E_1^2 + E_2^2} \tag{189}$$

where g is the traction bound, and we have adapted Coulomb's law to shifts. So:

$$|E_\tau| \neq 0 \;\Rightarrow p_\tau \delta E_\tau = -gE_\tau \delta E_\tau/|E_\alpha| = -g\delta|E_\tau| \tag{190}$$

Assume that there is no slip:

$$|E_\tau| = 0 \;\Rightarrow |p_\tau| \leq g \tag{191}$$

By Schwartz's inequality and (191) we have

$$p_\tau \delta E_\tau \geq -|p_\alpha||\delta E_\tau| \geq -g|\delta E_\tau| \tag{192}$$

Since $E_\tau = 0$, $\tau = 1, 2$, we have

$$|\delta E_\tau| = |E_\tau + \delta E_\tau| - |E_\tau| = \delta|E_\tau| \;\text{ if } (|E_\tau| = 0) \tag{193}$$

so that by (188), (192, 193, 190)

$$\begin{aligned} p_\tau \delta u_\tau &= p_\tau \delta E_\tau \geq -g\delta|E_\tau| \\ &\Leftrightarrow \; p_\tau \delta u_\tau = -g|\delta E_\tau| + \eta \\ \eta &\geq 0 \;\text{ for slip and no-slip} \end{aligned} \tag{194}$$

We note that (178) holds both inside and outside the contact, while (194) holds both in the slip area ($|E_\tau| \neq 0$) and in the adhesion zone ($|E_\tau| = 0$). So (178) and (194) will lead to a uniform formulation of the contact conditions on A_c in which neither the unknown contact area nor the unknown areas of slip and adhesion are mentioned explicitly, but note that A_c is known *a priori*.

We conclude from (176),(178),(194) that a necessary condition for contact is

$$\begin{aligned} 0 \;=\; & \sum_{a=1,2} [-\int_{V^a} (\sigma_{ij,j} + f_i - \rho\ddot{u}_i) dV + \int_{\partial V^a} p_i \delta u_i dS] + \\ - \; & \sum_{a=1,2} \int_{A_p^a} \overline{p}_i \delta u_i dS + \int_{A_c} g\delta|E_\tau| dS - \text{ nonnegative} \\ \forall \quad & \delta u_i \;\text{ sub } u_i = \overline{u}_i \text{ in } A_u^a;\; e \geq 0 \text{ in } A_c \end{aligned} \tag{195}$$

Hence, by Gauss's Theorem

$$\begin{aligned} 0 \;\leq\; & \delta U \doteq \sum_{a=1,2} [-\int_{V^a} (\sigma_{ij,j} - \rho\ddot{u}_i + f_i)\delta u_i) dV + \\ + \; & \int_{A_p^a} (p_i - \overline{p}_i)\delta u_i dS] + \int_{A_c} [p_i \delta u_i + g\delta|E_\tau|] dS = \end{aligned} \tag{196}$$

$$
\begin{aligned}
&= \sum_{a=1,2} [\int_{V^a} (\sigma_{ij}\delta u_{i,j} + \rho\ddot{u}_i - f_i\delta u_i dS) + \\
&- \int_{A_p^a} \overline{p}_i \delta u_i dS] + \int_{A_c} g\delta|E_\tau| dS \qquad (197)\\
&\forall \quad \delta u_i \ \text{ sub } u_i = \overline{u}_i \text{ in } A_u^a;\ e = h + u_z \geq 0 \text{ in } A_c \qquad (198)
\end{aligned}
$$

with h, e: distance between opposing points in the undeformed and deformed state.

It may be shown that the conditions (196) or (197) sub (198) are not only necessary, but also sufficient for the contact conditions. We will not prove that here; we will give a similar proof when we consider the complementary virtual work inequality.

4.3 Complementary Virtual Work

We start from the equilibrium equations (170), which we take as auxiliary conditions which must always be satisfied. We consider the quasistatic case that the density $\rho = 0$, that is, accelerations are not taken into account. Also, the body force f_i is prescribed. Together with (170) $\rho = 0$ this implies that

$$\delta(\sigma_{ij,j} + f_i - \rho\ddot{u}_i) = \delta\sigma_{ij,j} = 0 \qquad (199)$$

We multiply (199) by the displacement u_i, and integrate:

$$0 = \sum_{a=1,2} \int_{V^a} u_i \delta\sigma_{ij,j} dV \qquad (200)$$

and we find in much the same manner as before, after some calculation

$$
\begin{aligned}
0 \quad &\geq \quad \delta C \doteq \sum_{a=1,2} [\int_{V^a} u_i \delta\sigma_{ij,j} dV + \\
&- \int_{\partial V^a} u_i \delta p_i dS + \int_{A_u^a} \overline{u}_i \delta p_i dS] + \\
&- \int_{A_c} [h\delta p_z + |E_\tau|\delta g + (W_\tau - u'_\tau)\delta p_\tau] dS \qquad (201)\\
&= \sum_{a=1,2} [\int_{V_a} -e_{ij}\delta\sigma_{ij} dV + \int_{A_u^a} \overline{u}_i \delta p_i dS] \\
&- \int_{A_c} [h\delta p_z + |E_\tau|\delta g + (W_\tau - u'_\tau)\delta p_\tau] dS \qquad (202)\\
&\forall \quad \delta p_i, \delta\sigma_{ij} \text{ sub } \sigma_{ij,j} + f_i = 0 \text{ in } V_a \\
&p_i \quad = \overline{p}_i \text{ in } A_p^a;\ p_z \geq 0,\ |p_\tau| \leq g \text{ in } A_c. \qquad (203)
\end{aligned}
$$

In (202) the term $-e_{ij}\delta\sigma_{ij}$ appears instead of $-u_{i,j}\delta\sigma_{ij}$. These expressions are equal because $\sigma_{ij} = \sigma_{ji}$ and $e_{ij} = (u_{i,j} + u_{j,i})/2$.

The conditions (201)-(203) are implied by the contact conditions. Conversely, the contact problem is implied by those equations. We prove this.

To that end we restrict ourselves to the conditions in A_c. The other conditions are classical.
We start from (201), which is equivalent to (202). (203) is also valid. In (201) we note that $\delta\sigma_{ij,j} = 0$ in V^a; if we set $\delta p_i = 0$ on ∂V^a outside A_c, then

$$0 \leq \int_{A_c} [u_z \delta p_z + u_\tau \delta p_\tau + h\delta p_z + |E_\tau|\delta g + (W_\tau - u'_\tau)\delta p_\tau] dS$$

1. Set $\delta p_\tau = \delta g = 0$; this is the normal contact problem. We obtain, by the independence of the δp_z
$$0 \leq (u_z + h)\delta p_z = e\delta p_z \quad \text{sub } p_z \geq 0.$$
If $p_z > 0 \Rightarrow \delta p_z$ is bilateral, and $e = 0$ (contact).
If $p_z = 0 \Rightarrow \delta p_z \geq 0$, and $e \geq 0$ (no contact).
Here we define the contact ares as the region where $p_z > 0$ ("Force" definition). It then appears that the deformed distance e is nonnegative outside contact, and $= 0$ inside.

2. Now we set $\delta p_z = 0$. We are left with
$$\begin{aligned} 0 &\leq (u_\tau + W_\tau - u'_\tau)\delta p_\tau + |E_\tau|\delta g = \\ &= E_\tau \delta p_\tau + |E_\tau|\delta g \end{aligned}$$
sub $g - |p_\tau| \geq 0$.

   - If $|p_\tau| < g$ ("force" definition of the area of adhesion), then δp_τ and δg are independent and bilateral, so that
   $$E_\tau = |E_\tau| = 0$$
   - If $|p_\tau| = g$ ("force" definition of the area of slip), then $\delta g - \delta|p_\tau| \geq 0$. We decompose E_τ and δp_τ into components E^p_τ, δp^p_τ parallel to the vector (p_τ), and into components E^o_τ, p^o_τ orthogonal to that vector.

     We set $\delta p^p_\tau = \delta g = 0$. Then $E^o_\tau \delta p^o_\tau \geq 0$.
     Now δp^o_τ is bilateral, since in first order it does not contribute to
     $$\epsilon\delta|p_\tau| = \sqrt{|p_\tau + \epsilon\delta p^p_\tau|^2 + |\epsilon\delta p^o_\tau|^2} - |p_\tau|$$
     Thus $E^o_\tau = 0$, that is, the slip is parallel to the tangential traction:
     $$E_\tau = \pm|E_\tau|p_\tau/|p_\tau|.$$

Suppose $E_\tau = +|E_\tau|p_\tau/|p_\tau|$. Then

$$\begin{aligned} 0 &\leq |E_\tau|(p_\tau \delta p_\tau/|p_\tau|) + |E_\tau|\delta g = \\ &= |E_\tau|(\delta|p_\tau| + \delta g) \end{aligned}$$

Now take $\delta g = 0$. As $|p_\tau| = g$, $\delta|p_\tau| \leq 0$, and

$$0 \leq |E_\tau|\delta|p_\tau| \leq 0 \Rightarrow |E_\tau| = 0$$

Evidently this does not correspond to an area of slip, and anyway this situation (and much more) is also described by

$$E_\tau = -|E_\tau|p_\tau/|p_\tau|, \quad \Leftrightarrow \quad p_\tau = -gE_\tau/|E_\tau|$$

Then we have

$$0 \geq |E_\tau|(\delta g - \delta|p_\tau|) = |E_\tau|\delta(g - |p_\tau|)$$

Now, $\delta(g - |p_\tau|) \geq 0$, unilateral, hence $|E_\tau| \geq 0$, which corresponds to slip opposite the traction p_τ when it is at the traction bound.

We have established the normal contact formation conditions, *viz.* $p_z \geq 0$, hence:

1. If $p_z > 0$ then $e = 0$ (contact)
2. If $p_z = 0$ then $e \geq 0$ (no contact)

and we have established Coulomb's Law, *viz.* $|p_\tau| \leq g$, hence:

1. If $|p_\tau| < g$ then $E_\tau = 0$
2. If $|p_\tau| = g$ then $E_\tau = -|E_\tau|p_\tau/|p_\tau|$, or, equivalently, $p_\tau = -gE_\tau/|E_\tau|$.

These constitute the contact conditions.

4.4 Application to elasticity

We assume elasticity:

$$\begin{aligned} \frac{1}{2}(u_{i,j} + u_{j.i}) &= e_{ij} = e_{ji}, \text{ linearized strain;} \\ \sigma_{ij} &= E_{ijhk}e_{hk}, \text{ stress-strain relations; } \sigma_{ij} = \sigma_{ji}, \text{ stress;} \\ E_{ijhk}(\mathbf{y}) &= E_{jihk} = E_{hkij}, \text{ elastic constants;} \\ \frac{1}{2}E_{ijhk}e_{ij}e_{hk} &= \text{ elastic energy/unit volume } > 0 \\ &\quad \text{unless } e_{ij}e_{ij} = 0. \end{aligned} \tag{204}$$

The elastic constants are prescribed, but they may be position dependent. We can invert the stress-strain relations:

$$\begin{aligned} e_{hk} &= S_{hkij}(\mathbf{y})\sigma_{ij} \\ S_{hkij} &= S_{ijhk} = S_{khij} \\ S_{ijhk}\sigma_{ij}\sigma_{hk} &> 0 \quad \text{if } \sigma_{ij}\sigma_{ij} \neq 0 \end{aligned} \tag{205}$$

WE SET

$$\begin{aligned} \rho &= 0: \quad \text{elastostatics} && (206)\\ \delta g &= 0, \quad \text{that is, } g \text{ is given } a\ priori && (207) \end{aligned}$$

in order to be able to define a potential energy and a complementary energy of the system. We have, by (197)

$$\begin{aligned} 0 \leq\ & \delta U = \sum_{a=1,2} [\int_{V^a} (E_{ijhk} u_{h,k} \delta u_{i,j} - f_i \delta u_i) dV - \int_{A_p^a} \overline{p}_i \delta u_i] dS + \\ +\ & \int_{A_c} g \delta |E_\tau| dS = \\ =\ & \delta [\sum_{a=1,2} (\int_{V^a} (\frac{1}{2} E_{ijhk} u_{i,j} u_{h,k} - f_i u_i) dV \\ -\ & \int_{A_p^a} \overline{p}_i u_i dS) + \int_{A_c} g |E_\tau| dS] \end{aligned}$$

sub (206), (207), and (198).

This is equivalent to

$$\begin{aligned} \delta U &\geq 0, \quad \text{sub(206),(207),(198), with} \\ U &= \sum_{a=1,2} [-\int_{V^a} (\frac{1}{2} E_{ijhk} u_{h,k} u_{i,j} - f_i u_i) dV + \\ &- \int_{A_p^a} \overline{p}_i u_i dS] + \int_{A_c} g |E_\tau| dS && (208)\\ & && (209) \end{aligned}$$

U is called the potential energy of the system. In the same way,

$$\begin{aligned} \delta C &\leq 0 \ \text{ sub (203), (206), (207), with} \\ C &= \sum_{a=1,2} [-\int_{V^a} \frac{1}{2} S_{ijhk} \sigma_{ij} \sigma_{hk} dV + \int_{A_u^a} \overline{u}_i p_i dS] + \\ &- \int_{A_c} [h p_z + (W_\tau - u'_\tau) p_\tau] dS && (210) \end{aligned}$$

C is called the complementary energy.

It may be shown (see [4] (Section 4.2.1) that these conditions characterize

1. The global minimality of U at the solution;

2. The global maximality of C at the solution;

3. The equality of U and C at the solution;

4. The uniqueness of the solution,

all under the rather restrictive conditions (206) and (207).

4.5 The case that $\delta g \neq 0$

According to (207) the theory of the previous subsections does not seem to apply when δg is not constrained to be zero, that is, when g is not prescribed beforehand. We saw in Section 2 that the normal pressure is independent of the tangential tractions for symmetry of all 3D bodies, and for quasiidentical half-spaces. As the normal problem is not influenced by g, we can determine the normal problem regardless of g in those cases; thereafter $g = fp_z$, g fixed, f: coefficient of friction, and we have $\delta g = 0$. So in these cases the theory is exactly verified.

A process to deal with $\delta g \neq 0$ is the Panagiotopoulos process [7], see Section 2. This is an iterative method, and it results in the exact solution when it converges, which is not always. It is designed in such a way that the theories of minimal potential energy and maximal complementary energy can be used.

The Panagiotopoulos process works as follows:

1. Set $m = 0$. Assume that $p_\tau^{(0)} = 0$.

2. Determine $p_z^{(m)}$ with $p_\tau^{(m)}$ as tangential traction.
   This is a normal contact problem with fixed tangential traction, so that g does not play a role.

3. Determine $p_\tau^{(m+1)}$ with $p_z^{(m)}$ as normal traction, and $g = fp_z^{(m)}$ as traction bound.
   Now g is given, and hence $\delta g = 0$, and we can apply our theories.

4. When $p_\tau^{(m+1)}$ is close enough to $p_\tau^{(m)}$, stop; else reiterate.

4.6 Existence-uniqueness theory

The principle of virtual work has been used to establish existence and uniqueness theorems of the contact mechanical field for several types of bulk material:

- In 1964 Fichera [2] established the existence and uniqueness of the normal contact problem of frictionless contact $g = 0$.

- In 1972 Duvaut and Lions [1] established existence-uniqueness for the tangential field for linear viscoelastic and dynamic fields when g is a function of position only, independent of time and other quantities.

- In 1983 Oden and Pires [6] proved the existence of the linear elastic field due to normal contact and friction.
- In the foregoing analysis we have considered contact problems in which a single step is taken from a "previous" instant t' to the present time t.
  When we have a contact evolution, it is not *a priori* clear whether the solution exists and is continuous as a function of time. Under certain conditions an affirmative answer was given by Klarbring *e.a.* [5].
- Another problem is the uniqueness-existence of a steady state in a continuous evolution. Kalker [3] proved this for quasiidentical, 2D no-slip half-space rolling contact under the condition that the normal compressive force and the creepage were constant from a certain instant of time onwards.

4.7 Surface mechanical principles

We express the principles in surface mechanical form, *i.e.* a form in which the volume integral is absent. To that end we take test functions in the principles of minimum potential and maximum complementary energies which satisfy all elasticity equations as well as the homogeneous boundary conditions:

$$u_i = \overline{u}_i = 0 \text{ in } A_u^a, \text{ and } p_i = \overline{p}_i = 0 \text{ in } A_p^a. \tag{211}$$

We assume no body force, and elastostatics, so that the equilibrium conditions are

$$\sigma_{ij,j} = 0 \tag{212}$$

Finally,

$$\sigma_{ij} = E_{ijhk}(\mathbf{y})u_{i,j} \tag{213}$$

Then we obtain for the two principles, after some calculation

$$min_{u,p}U = \int_{A_c} [\frac{1}{2}p_z u_z + (\frac{1}{2}p_\tau u_\tau + g|E_\tau|)]dS \tag{214}$$

sub (211)-(213), and $e = h + u_z \geq 0$,

$$max_{u,p}C = -\int_{A_c} [(h + \frac{1}{2}u_z)p_z + (W_\tau + \frac{1}{2}u_\tau - u'_\tau)p_\tau]dS \tag{215}$$

sub (211)-(213), and $p_z \geq 0$, $|p_\tau| \leq g$,
which lack volume integrals.
Note that they are valid only when $\rho = 0$, and $\delta g = 0$. When $\delta g \neq 0$, one can use the Panagiotopoulos process. The principle (215) has been used extensively in numerical work (CONTACT, since 1982).

4.8 Complementary or potential energy in numerical work?

A disadvantage of the method of potential energy is that the integral over A_c at one stage or another contains the variation $\delta|s_\tau|$, while the derivative of $|s_\tau|$ is discontinuous when $|s_\tau| = 0$.

The method of maximum complementary energy does not have this disadvantage, but it can only be used in statics, and the equations of equilibrium have to be satisfied inside the bodies. This is no problem when one can use the surface mechanical method, as its test functions are required to do just that. So in that case the method of maximum complementary energy is to be preferred.

4.9 Conclusion

The variational, or weak formulation of the contact problem has been presented. It consists of two variational inequalities, one derived from the principle of virtual work and one from the principle of complementary virtual work. These principles are very fruitful guidances for deriving algorithms for the contact problem.

References of Section 4

[1] G. Duvaut, J.-L. Lions (1972), *Les inéquations en mécanique et en physique.* Paris, Dunod.

[2] G. Fichera (1964), *Problemi elastostatici con vincoli unilaterale: il probleme di Signorini con ambigue condizioni al contorno.* Memorie Accademie Nazionale dei Lincei Ser. 8, 7, p. 91-140.

[3] J.J. Kalker (1970), *Transient phenomena in two elastic cylinders rolling over each other with dry friction.* Journal of Applied Mechanics **37**, p. 677-688.

[4] J.J. Kalker (1990), *Three-dimensional, elastic bodies in rolling contact.* Kluwer Academic Publishers Dordrecht, Boston, London.

[5] A. Klarbring, A. Mikelič, M. Shillor (1990), *A global existence result for the quasistatic frictional contact problem with normal compliance.* Proc. IV. Meeting on Unilateral Problems in Structural Analysis, Ed. F. Maceri *et al.*, Birkhăuser Verlag, Zŭrich.

[6] J.T. Oden, E.B. Pires (1983), *Nonlocal and nonlinear friction laws and variational principles for contact problems in elasticity.* Journal of Applied Mechanics **50** p. 67-76.

[7] P.D. Panagiotopoulos (1975), *A nonlinear programming approach to the unilateral contact and friction -boundary value problem in the theory of elasticity.* Ingenieur Archiv **44**, p. 421-432.

5 Numerical analysis: Exact theory

5.1 Discretization

We choose the potential contact region as a rectangle with sides parallel to the $x-$ and y-axis in the surface of the half-space. The x-axis is in the direction of rolling. The z-axis points vertically upwards into body 1, and the y-axis completes the right-handed Cartesian coordinate system. The vertex of the potential contact with the lowest x and y values has the coordinates $(x_0,\ y_0, 0)$. The potential contact is subdivided into N equal and equally oriented subrectangles (elements) The elements are numbered from 1 to N; they have height $2\Delta y$ and width $2\Delta x$. Their centroids are denoted by

$$\mathbf{x}_I \ = \ (x_I, y_I, 0) \tag{216}$$

where a capital latin index runs from 1 to N.

The origin of the potential contact does not lie necessarily in its centre. The situation is shown in Figure 25.

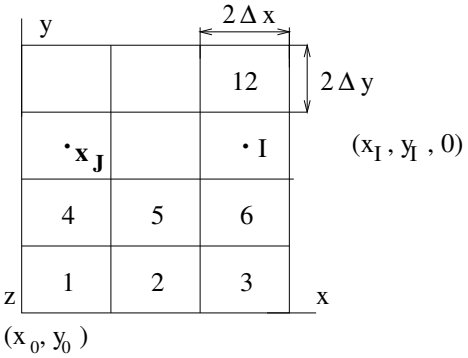


Figure 25: The potential contact area (pot.con.)

We take the traction $\mathbf{p} = (p_j)$ constant in each element, so that the traction is piecewise constant. This traction would seem to be representative for the force/unit area at the centroid of the element in question.
At the time t it is denoted by (p_{Jj}), and at the time t' it is denoted by (p'_{Jj}).
We are interested in the surface displacement difference $\mathbf{u}(x, y, 0)$ in the surface point $\mathbf{x} = (x, y, 0)$. To that end we integrate the representation of Boussinesq-Cerruti over each element. We use the following notation: $(B_{iJj})(x, y)$ is a 3-dimensional array. i corresponds to the component of the displacement difference $u_i(x, y)$, and Jj to the component of the traction p_{Jj}. $(x, y, 0)$ is the position where the displacement difference is taken. Then

$$(B_{iJj}(x, y)) \quad = \quad \frac{1}{\pi G} \int_{x_J-\Delta x}^{x_J+\Delta x} dx' \int_{y_J-\Delta y}^{y_J+\Delta y} dy' \cdot$$

$$\cdot \quad \begin{pmatrix} \frac{1-\nu}{R} + \frac{(x'-x)^2}{R^3} & \frac{\nu(x'-x)(y'-y)}{R^3} & \frac{K(x'-x)}{R^2} \\ \frac{\nu(x'-x)(y'-y)}{R^3} & \frac{(1-\nu)}{R} + \frac{(y'-y)^2}{R^3} & \frac{K(y'-y)}{R^2} \\ -\frac{K(x'-x)}{R^2} & -\frac{K(y'-y)}{R^2} & \frac{1-\nu}{R} \end{pmatrix} \tag{217}$$

with $R = \sqrt{(x'-x)^2 + (y'-y)^2}$.

There are six different integrands in this formula, *viz.*

$$\begin{aligned} J_1 &= \frac{1}{R} & J_4 &= \frac{(y'-y)^2}{R^3} \\ J_2 &= \frac{(x'-x)^2}{R^3} & J_5 &= \frac{x'-x}{R^2} \\ J_3 &= \frac{(x'-x)(y'-y)}{R^3} & J_6 &= \frac{y'-y}{R^2} \end{aligned}$$

Note that $J_1 = J_2 + J_4$, that the integration over J_2 can be derived from the integration over J_4, and that the integration over J_5 can be derived from the integration over J_6, all by symmetry. So we need only integrate over J_3, J_4 and J_6. We write the result using the notation

$$[[f(x,y)]] = [[f(x,y)]]_{x-x_J-\Delta x}^{x-x_J+\Delta x}{}_{y-y_J-\Delta y}^{y-y_J+\Delta y} \tag{218}$$

The result is:

$$(\pi G)B_{1J1}(x,y) = [[y\ln(x+r) + (1-\nu)x\ln(y+r)]] \tag{219}$$
$$(\pi G)B_{1J2}(x,y) = (\pi G)B_{2J1} = [[-\nu r]] \tag{220}$$
$$(\pi G)B_{1J3}(x,y) = -(\pi G)B_{3J1} = [[K(y\ln r + x\arctan(y/x))]] \tag{221}$$
$$(\pi G)B_{2J2}(x,y) = [[x\ln(y+r) + (1-\nu)y\ln(x+r)]] \tag{222}$$
$$(\pi G)B_{2J3}(x,y) = -(\pi G)B_{3J2} = [[K(x\ln r + y\arctan(x/y))]] \tag{223}$$
$$(\pi G)B_{3J3}(x,y) = [[(1-\nu)(x\ln(y+r) + y\ln(x+r))]], \tag{224}$$

with $r = \sqrt{x^2+y^2}$, and

$$\frac{1}{G} = \frac{1}{2}\left(\frac{1}{G^1} + \frac{1}{G^2}\right) \quad \text{1,2: body numbers} \tag{225}$$
$$\frac{\nu}{G} = \frac{1}{2}\left(\frac{\nu^1}{G^1} + \frac{\nu^2}{G^2}\right) \tag{226}$$
$$K = \frac{G}{4}\left(\frac{1 - 2\nu^1}{G^1} - \frac{1 - 2\nu^2}{G^2}\right) \tag{227}$$

We sample the displacement difference in $(\mathbf{x}_I, t)$: $(\mathbf{u})$ and in $(\mathbf{x}_I + \mathbf{v}(t-t'), t')$: $(\mathbf{u}')$, present and previous displacement differences:

$$\mathbf{u}_I = (u_{Ii}) = (u,v,w) \quad \text{at} \quad (\mathbf{x}_I, t) \tag{228}$$
$$\mathbf{u}'_I = (u'_{Ii}) = (u,v,w) \quad \text{at} \quad (\mathbf{x}_I + \mathbf{v}(t-t'), t') \tag{229}$$

To that end we define

$$\begin{aligned} A_{IiJj} &= B_{iJj}(\mathbf{x}_I) && (230)\\ C_{IiJj} &= B_{iJj}(\mathbf{x}_I + \mathbf{v}(t-t')) && (231)\\ \mathbf{p} &= (p_{Jj}) \;=\; \mathbf{p}(\mathbf{x}_J, t) && (232)\\ \mathbf{p}' &= (p'_{Jj}) \;=\; \mathbf{p}(\mathbf{x}_J, t') && (233) \end{aligned}$$

NOTE that in non-steady rolling $p'_{Jj} \neq p_{Jj}$, while p'_{Jj} is known; on the other hand, in steady state rolling, $p'_{Jj} = p_{Jj}$, so that p'_{Jj} is unknown.

We have:

$$\begin{aligned} u_{Ii} &= \sum_{J=1}^{N}\sum_{j=1}^{3} A_{IiJj}p_{Jj} && (234)\\ u'_{Ii} &= \sum_{J=1}^{N}\sum_{j=1}^{3} C_{IiJj}p'_{Jj} && (235) \end{aligned}$$

NOTE that in non-steady rolling u'_{Ii} is known, while in steady-state rolling it is unknown, but it equals u_{Ii} shifted over a distance $\mathbf{v}(t-t')$.

5.2 The NORM algorithm

We will now treat the algorithm for the normal contact. We assume that the tangential traction is prescribed, as well as the undeformed distance h: $p_{J\tau}$, $\tau = 1, 2$ is the tangential traction, function of the position $\mathbf{x}_J$; $p_{Jz} = p_z(\mathbf{x}_J)$ is the normal pressure, both on body 1. The normal pressure is positive in the contact region, and vanishes outside it.

The displacement difference, defined as $\mathbf{u}_I = \mathbf{u}^1(\mathbf{x}_I) - \mathbf{u}^2(\mathbf{x}_I)$ is, in components, $\mathbf{u}_I = (u_I, v_I, w_I)$. The undeformed distance h_I is the scalar distance between body 1 and body 2 at $\mathbf{x}_I$ in the undeformed state, measured from body 2 to body 1. The deformed distance e_I is similarly defined, but in the deformed state. We have:

$$e_I = h_I + w_I \tag{236}$$

The deformed distance is positive outside the contact area, and vanishes inside it. We take the displacement fixed at infinity, at zero. So the boundary conditions become:

$$\begin{aligned} &p_{Iz} \geq 0, \qquad w_I \geq 0, \qquad p_{Iz}w_I = 0 && (237)\\ &\mathbf{u}^a \to 0 \quad \text{if} \;\; |\mathbf{x}| \to \infty, \quad a = 1, 2: \quad \text{body number} && (238) \end{aligned}$$

The condition at infinity is automatically satisfied if we use the Boussinesq-Cerruti formulae of the previous subsection. The condition $p_z = 0$ outside the potential contact

is satisfied if we simply set the normal traction 0 where there are no elements, i.e. on $z = 0$, outside the pot.con. Inside the pot.con. we act as follows.

1. We set
$$h_I^* = h_I + \sum_{J,\tau} A_{I3J\tau} p_{J\tau} \qquad \forall \mathbf{x}_I \in Q$$

   Explanation. The given tangential tractions are taken into account by modifying the undeformed distance (h^* instead of h).
   Q is the entire pot.con.

2. Set $p_{Jz} = 0 \quad \forall J$: Clear all normal tractions.

3. Initially, the exterior $E = Q$, and the contact area $C = \emptyset$. During the process, E and C are modified until they correspond to the solution.

4.
$$e_I = h_I^* + \sum_J A_{I3J3} p_{Jz} \qquad \forall \mathbf{x}_I \in Q \tag{239}$$

   Equations: $e_I = 0 \;\; \forall \mathbf{x}_I \in C, \;\; p_{Jz} = 0 \;\; \forall \mathbf{x}_J \in E$.
   These are N linear equations for the N p_{Jz} which can be solved.

5. Are all pressures $p_{Jz} \geq 0$ in C? If "no", go to Point 6. If "yes", go to Point 7.

   Explanation. It is checked whether the pressures are all positive in C.
   If "yes", go to Point 7.
   If "no", place all elements with negative pressure in the Exterior E, where the pressure will be annihilated according to the equations of Point 4.

   Note. If we place only one such element in the exterior E, the algorithm can be proved mathematically by means of variational calculus. We have not succeeded in doing so when we place all such elements in the exterior E. Yet we never have had a failure!

6. Restore. If $p_{Jz} < 0$, element J is placed in Exterior E, and p_{Jz} is set equal to zero. Go to Point 4.

7. Are all deformed distances $e_I > 0$ in E? If "yes", WE ARE READY.
   If "no", place all elements with deformed distance < 0 in C; go to Point 4.

   Explanation. It is checked whether all deformed distances are positive in E.
   If "yes", all pressures are positive in C and vanish in E, while all deformed distances are positive in E and vanish in C. All contact conditions are met; we are READY.
   If "no", we place all elements with negative deformed distance in C and go to Point 4, where care is taken that the deformed distances in the new C are annihilated. Here also we should reform only one element in order to get a proof of the algorithm.

5.2.1 Normal force prescribed

There is an important variant of the NORM algorithm as we described it here. To state it, we first introduce the notion of the approach of the bodies.
Let h_I be the undeformed distance between the bodies; then $h_I - d$ is the distance, but with the bodies moved a distance d towards each other. d is called the approach of the bodies. In the algorithm above we set $d = 0$, or better, we give it a fixed value.
If we want to prescribe the total normal force, we regard d as a variable of the problem, but instead we demand that that the total normal force

$$F_z = 4\Delta x \Delta y \sum_{J \in Q} p_{Jz} = F'_z \tag{240}$$

is prescribed as $F'_z > 0$.

We have to adapt NORM to cope with this variant.

1. $h^*_I = h_I + \sum_{J,\tau}(A_{I3J\tau} p_{J\tau} - d$
   Initially, set $d = 0$. Then find the minimum of h^*_I over the entire pot.con Q: minimizer I'. Now, take d so that $h^*_{I'} = 0$.

2. Set $p_{I'z} = \frac{F'_z}{4\Delta x \Delta y} > 0$. Set the other $p_{Iz} = 0$, for all other $I \in Q$.

3. Set $C =$ Element I'. Set $E =$ all other elements of Q.
   E and C are modified until they correspond to the solution.

4. $e_I = h^*_I + \sum_J A_{I3J3} p_{Jz}$. Note that h^*_I contains d in a linear fashion. The equations are: $e_I = 0 \; \forall I \in C$; ($p_{Jz}$, d variable;) $p_{Jz} = 0$ in E; $4\Delta x \Delta y \sum_J p_{Jz} = F'_z$
   ($N + 1$ linear equations with $N + 1$ unknowns, *viz.* N p_{Iz}, and d.

From here on we follow the ordinary NORM (points 5. to 7.).
This algorithm will break down when $F'_z < 0$, as then at least one p_{Iz} must be negative.

5.3 The TANG algorithm

We will now treat the algorithm for the tangential contact. We assume that the normal traction is prescribed, as well as the creepages: $p_{Jz} = p_z(\mathbf{x}_J)$, $\mathbf{c}_I = \mathbf{c}(\mathbf{x}_I)$. In non-steady state rolling, the previous displacement difference $(u, v)'_I = (u, v)(\mathbf{x}_I + \mathbf{v}(t - t'), t')$ is known, in steady state rolling it is equal to $(u, v)'_I = (u, v)(\mathbf{x}_I + \mathbf{v}(t - t'))$, as the displacement difference $(u)_I$ is independent of explicit time in the contact fixed coordinate system. Therefore, $(u, v)'_I$ is "as unknown" as $(u, v)_I$ itself.

The displacement differences $\mathbf{u}$, $\mathbf{u}'$ depend on the surface traction field $\mathbf{p}_J = \mathbf{p}(\mathbf{x}_J, t)$ and the previous traction field $\mathbf{p}'_J = \mathbf{p}(\mathbf{x}_J, t')$ in the following manner:

$$u_I = \sum_J \sum_j A_{I1Jj} p_{Jj} \tag{241}$$
$$v_I = \sum_J \sum_j A_{I2Jj} p_{Jj} \quad \text{at } (\mathbf{x}_I, t), \tag{242}$$

$$u'_I = \sum_J \sum_j C_{I1Jj} p'_{Jj} \tag{243}$$
$$v'_I = \sum_J \sum_j C_{I2Jj} p'_J \quad \text{at } (\mathbf{x}_I + \mathbf{v}(t-t'), t') \tag{244}$$

The slip is defined as follows:

$$\mathbf{s} = \mathbf{c} - \frac{\partial \mathbf{u}}{\partial \mathbf{x}} \mathbf{v} + \frac{\partial \mathbf{u}}{\partial t}, \quad \text{or, discretized:} \tag{245}$$
$$\mathbf{s}_I = \mathbf{c}_I + \frac{(\mathbf{u}_I - \mathbf{u'}_I)\mathbf{v}}{t - t'} \tag{246}$$

The traction bound is given as $g_I = f p_{Iz}$, with f the coefficient of friction. We have:

$$|(p_{I1}, p_{I2})| \leq g_I; \tag{247}$$
$$\text{if the inequality holds: } \mathbf{s}_I = 0 \tag{248}$$
$$\text{if the equality holds: } \mathbf{s}_I = -|\mathbf{s}_I|(p_{I1}, p_{I2})/g_I \tag{249}$$

and the half-space is fixed at infinity.

The algorithm runs as follows:

1. Set $(p_{I1}, p_{I2}) = (0, 0)$: Clear the tangential tractions.

2. Initially, the slip area $S = \emptyset$, and the adhesion area H (= stick area) = the complete contact patch C. During the process, S and H are modified, but always in such a way that $H \bigcup S = C$, until they correspond to the solution.

3. Solve the following equations:

$$\mathbf{s}_I = 0 \text{ in Adhesion area } H\text{; these are linear equations} \tag{250}$$
$$|(p_{J1}, p_{J2})| = g_J \text{ in slip area } S\text{. (Non-linear equations)} \tag{251}$$

   Remark. These equations are nonlinear. They may be solved, e.g., by means of the Newton-Raphson method.

4. If $|p_{I1}, p_{I2}| > g_I$, for any element in H, place element I in Slip area S.
   If this has happened at least once, go back to Point 3.
   If it has not happened, go to Point 5.

5. If the slip $\mathbf{s}_I$ is in the same sense as the tangential traction (p_{I1}, p_{I2}), rather than opposite as it should be in the slip area, then place element I in the adhesion area H.
   If this has happened at least once, go back to Point 3, else we are READY.

5.3.1 Total force components prescribed

There is an important variant of the TANG algorithm as we described it here. It is that one or two of the total tangential force components are prescribed. In order to do that, either v_x and/or v_y must be left free, and the total tangential force F_x and/or F_y must be prescribed. This goes in much the same way as it was done in NORM, and we will not give any details.

Convergence becomes bad when $|(F_x, F_y)|$ approaches fF_z; the solution does not exist when this quantity exceeds fF_z. The convergence in notably threatened when two components of the total tangential force are prescribed.

6 Results

In the present Section we show some results of the numerical work.

In Subsection 1, we show numerical results of steady-state rolling in two dimensions, both for symmetrical and asymmetrical bodies. In Subsection 2, we show results on non-steady state rolling in 2D. In Subsection 3 we show results of 3D steady-state rolling, and in Subsection 4 results of non-steady state 3D rolling.

6.1 2D Steady-state rolling

6.1.1 Symmetrical bodies

Consider two infinitely long cylinders with parallel axes which are made of the same homogeneous, isotropic elastic material: modulus of rigity G, Poisson's ratio ν. The cylinders are pressed together so that a contact strip forms between them; the strip is bounded by two parallel lines. A coordinate system is introduced of the form $(O; x, y, z)$ in which the origin O lies on the centre line of the strip, the axis of x points in the direction perpendicular to the axis of the strip, in the plane of contact, and in the rolling direction, and the axis of z is also perpendicular to the centre line of the strip, but points upwards into the upper cylinder (body 1), while the axis of y lies along the centre line of the strip, in such a way that the coordinate system $(O; x, y, z)$ is right-handed.
An alternative notation for (x, y, z) is $\mathbf{x} = (x_1, x_2, x_3)$.

The rolling velocity and the creepage were introduced before. When the lateral creepage $v_y = 0$, and the spin also vanishes, then the motion and the elastic deformation are two-dimensional and in plane strain. Then the state of the traction distribution was determined analytically by Carter [6], and by Fromm [7]. Carter uses the half-space approximation, while Fromm does not. The normal traction distribution is Hertzian/semi-elliptical, and the tangential traction distribution is given in Fig. 26.
Shown are the numerical values obtained by CONTACT. The coincidence is perfect.

A particle enters the contact area from the right. It is unloaded at the time. It is seen that the stick area borders on the leading edge of the contact strip, and that the traction increases till it reaches the traction bound; there, slip sets in and the traction remains on the traction bound till the loaded particle leaves the contact area, and loses its loading.

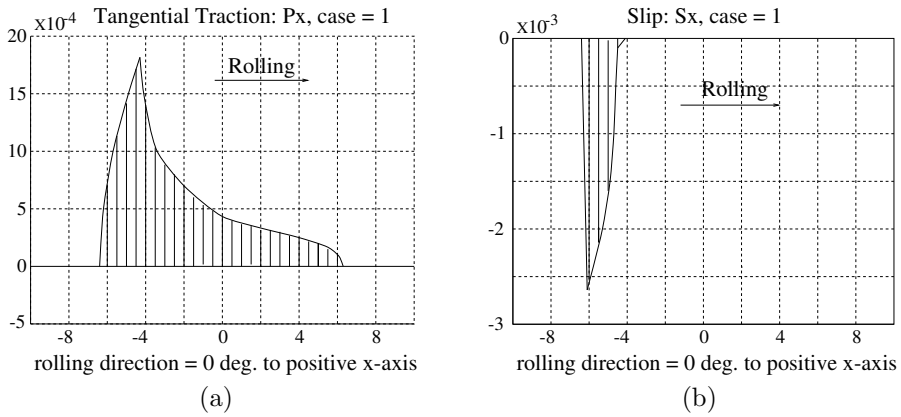
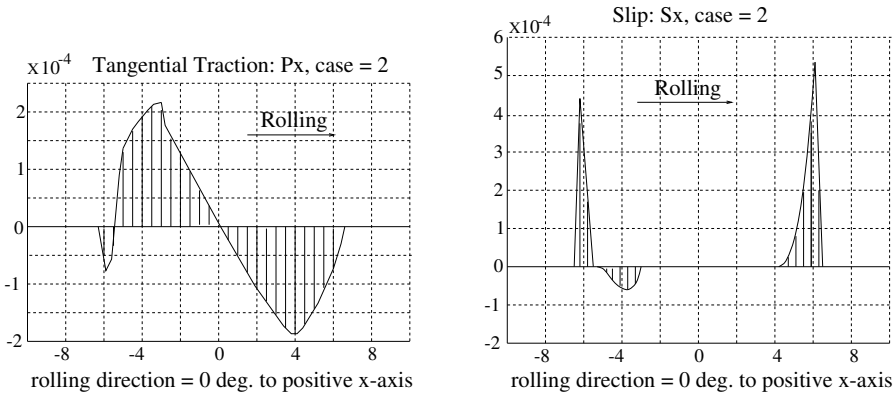


Figure 26: Symmetric rolling. Left: the traction distribution according to Carter (1926) and to CONTACT. Right: the slip distribution. Leading edge on the right, rolling from left to right.

6.1.2 2D asymmetric rolling

Now, we have the same geometry, but the elastic constants of the cylinders differ. Again, we set the lateral creepage and the spin equal to zero, and the motion and the elastic deformation are two-dimensional in plane strain. The state of traction was determined numerically by Bentall and Johnson [2] for so-called free rolling, that is, rolling with no net tangential force (dots), see Fig. 27. Also shown are the numerical values obtained by CONTACT (line). The coincidence is good, considering that both methods are numerical. The slip is shown in the figure on the right (CONTACT).

It is worth while to analyze the solution more deeply. Rolling takes place from left to right. The lower body is perfectly rigid, the upper body has a modulus of rigidity of unity. The Poisson ratio of the lower body (1) is irrelevant, as the body does not deform at all; the Poisson ratio of the upper body (2) is 0.286, a quite normal value. The total tangential force F_x vanishes.
Slip in the direction of rolling takes place in the interval $[0.410, 0.625]$ and in $[-0.625, -0.575]$. Slip counter to the direction of rolling takes place in $[-0.505, -0.300]$. In the remainder of the contact area sticking takes place. The traction is shown in the left figure, the slip in the right figure. Note the difference in the first derivative of the slip at the leading and the trailing edge of a slip area!



(a) (b)

Figure 27: Free rolling of asymmetric bodies, from left to right. Parameters: $G_1 = 1, G_2 = \infty, \nu_1 = 0.286, \nu_2 =$ irrelevant, $f = 0.1, F_x = 0$. Left figure: the traction. Right figure: the slip. Rolling is from left to right.

We also calculated some cases of tractive rolling. In Fig. 28 we compare two cases which are the same, except that the material constants of bodies 1 and 2 have been interchanged, while the tangential force has changed sign. The upper are the constants of Fig. 27. The right figures are the slip corresponding to the left ones.

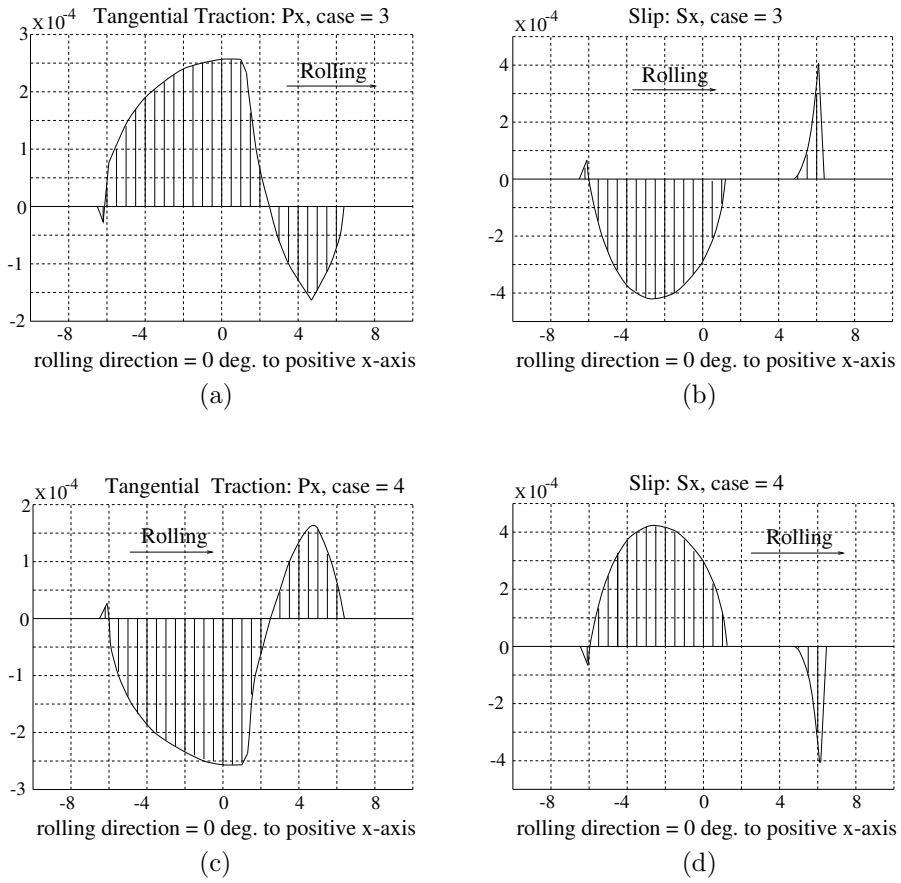


Figure 28: Tractive, asymmetric rolling. Comparison of interchanged material constants and inverse tangential force. p_x (and also the slip) merely change sign. Rolling is from left to right.

Another contrast is found in Fig. 29: The material constants are the same as in Fig. 27, but the total tangential force is $-0.5fF_z$. Also shown is the slip corresponding to the values of the left figure. Fig. 28 and Fig. 29 look totally different, although the absolute value of the total tangential force is exactly the same!

6.2 2D Transient rolling

The results from this section are taken from Kalker [11]. The notation of that paper differs slightly fom ours: in Kalker [11] we use X for p_x, and the slip is defined as the local velocity of body 2 over body 1, and is just the opposite from what we always have. We will show only traction distributions, as we are more interested in qualitative than in quantitative behaviour.

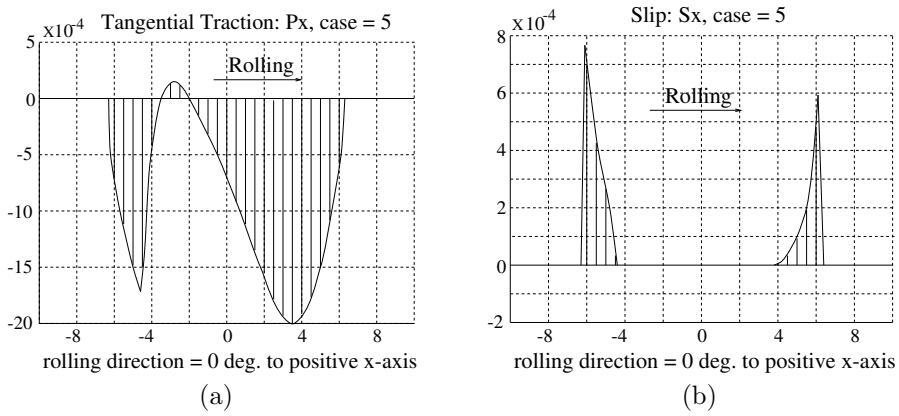


Figure 29: Traction and slip for the material constants of Fig. 27, but with total tangential force $-0.5fF_z$. Rolling is from left to right.

6.2.1 From Cattaneo to Carter

Two similar cylinders are pressed together. Then, opposing couples are applied to their axes, so that the cylinders start rolling over each other at time $t = 0$. The evolution is shown in Fig. 30. At the time $t = 0 + \epsilon$, $\epsilon \downarrow 0$ the traction distribution looks as as shown in Fig. 30(a). It is the so-called Cattaneo distribution, of shifting rather than rolling.

$F_x/fF_z = 0.75$; the half-circles represent the Coulomb traction bound. The dots are a comparison with a previous theory; it is seen that the coincidence is quite good. The contact width $2a = 2$. The rolling velocity $V = 1$.

Three phases can be distinguished in the evolution.

1. The first phase is the initiation, in this case the Cattaneo shift, Fig. 30(a).

2. In the second phase, the old traction distribution is shoved out of the contact area, and replaced with a traction distribution already akin to the steady state traction distribution. In the present case of Fig. 30, this phase takes place from $t = 0$ to $t = 1$ (Fig. 30(b) to Fig. 30(e)).
   Immediately after rolling starts, the traction at the point A (Fig. 30(a) and (b)) leaves the traction bound temporarily; it drops and rises again till it reaches the traction bound again at $t = 0.4$ (Fig. 30(b)). At that time, the stick area breaks into two. The left hand stick area vanishes fast; it has gone at $t = 0.58$.
   The vertical tangent at B moves inward with rolling velocity, together with the particle carrying it. It vanishes at $t = 1$, and this marks the end of phase 2.

3. In phase 3, the traction distribution adapts itself to the steady state, which is virtually reached when $t = 2$, $i.e >$ after one contact length has been traversed.

It is remarkable how fast the transience from a rather arbitrary initial traction distribution to that of steady state rolling takes place, once the total force is held constant.

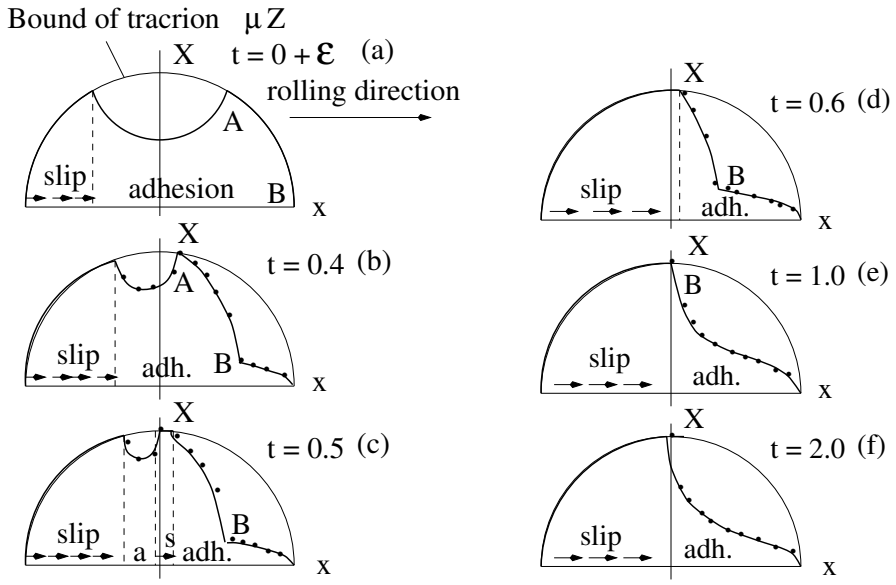


Figure 30: From Cattaneo to Carter (shift to steady state rolling). (Source: Kalker [11])

6.2.2 Rolling of similar cylinders under varying normal force

Two similar cylinders are rolled over each other under the action of a constant tangential force F_x. The normal force is so that the half contact width has the following behaviour:

$$a(t) = 1 + \frac{0.4}{\pi}\sin(2\pi t)$$

while

$V(t) = 1$, $K = 0$ (similarity), $f = 1$, $F_x = 0.255\pi$, $F_z = \frac{1}{2}\pi a^2$. In a quasi-steady state the initial traction distribution is immaterial.

The tangential tractions are shown in Fig. 31, in which are represented
$t = 1, 1.2, 1.4, 1.6, 1.8$, while $t = 2.0$ (full period after $t = 1$) is represented by dots in the figure corresponding to $t = 1$. The semi-circles represent the Coulomb traction. Slip takes place wherever the Coulomb traction is attained, opposite the traction.
From the coincidence of the traction at $t = 2$ with that at $t = 1$ it is seen that at $t = 1$ the quasi-steady state is already reached, from which it is seen that the transience proceeds extremely fast.

6.2.3 Harmonic rolling velocity and tangential force

Two similar cylinders are rolled over each other with a harmonic rolling velocity and a harmonic tangential force. The normal pressure and the friction coefficient remain constant, while a quasi-steady state is investigated in which the initial traction distribution

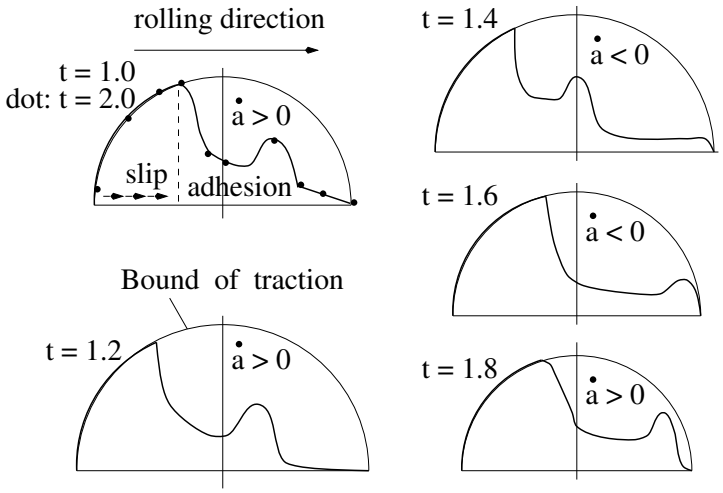


Figure 31: Rolling of similar cylinders under varying normal force. (Source: Kalker [11])

is immaterial.

The governing parameters of the program are:
$V(t) = \pi\cos(\pi t)$, $F_x = -0.355\pi\sin\pi t$,
$a = 1$, $K = 0$, $f = 1$, $F_z = \frac{1}{2}\pi$.

Half a period, from $t = 1$ to $t = 2$, is shown in Fig. 32. The leading edge at each instant is marked by L, the trailing edge by T. The traction distribution at $t = 2$ is the traction distribution at $t = 1$, multiplied geometrically by (-1) with the origin as centre of multiplication. The circles, as usual, represent the Coulomb traction; slip is present wherever the Coulomb traction is attained, and it is opposite the traction.

It is seen that at $t = 1.5$, when the rolling velocity changes sign, a vertical tangent is introduced at the edge $x = 1$, which moves inward with rolling velocity. At the same time the slip vanishes near the leading edge $x = 1$ and slip starts at the trailing edge $x = -1$.
Transience was completed in half a period from $t = 0$ to $t = 1$.

6.2.4 Frictional compression followed by transient rolling of dissimilar cylinders

Two dissimilar cylinders ($G_1 = 1$, $G_2 = \infty$, $\nu_1 = 0.286$, ν_2 immaterial, coefficient of friction $f = 0.15$) are brought into contact and compressed. The resulting traction distribution has been described by Spence [16]. The resulting tangential traction is shown in Fig. 33(a). Subsequently they are rolled; the semi-contact width $a = 1$. Here also, the three phases are clearly visible. The first phase, initiation, is given by Fig. 33(a); the second, that of the transience to almost the distribution of steady-state rolling is given

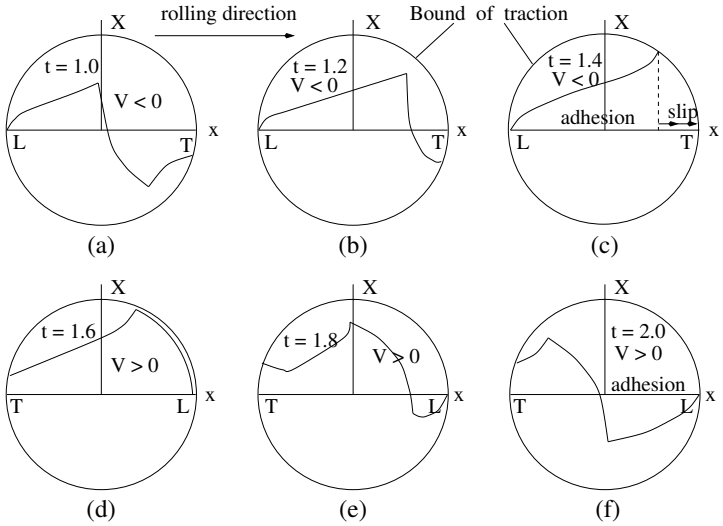


Figure 32: Harmonic rolling velocity and tangential force. (Source: Kalker [11])

by Figs. 33(b)-(d), while at $t = 1.8$ the steady state has virtually set in Fig. 33(e).

6.3 3D Steady state rolling

6.3.1 Quasiidentical steady state rolling: surface stresses

We recall the areas of slip and adhesion. They were described in the Section on simplified theory, *viz*. Section 3, and shown in Fig. 20. Here we will start with showing the contact stresses in quasiidentical steady state rolling with a circular contact area (Fig. 34). In Fig. 34, A signifies an adhesion area (stick area), and S a slip area. The arrows are in the direction of the traction, the circles are the boundary of the contact areas, the lines inside the contact areas are the slip-stick boundaries. In Figs. 34(a') and 34(d') are shown the absolute value of the tangential traction on the line $x - x$ in Figs. 34(a) and 34(d). The dot in the contact areas indicates the so-called spin pole: $(\upsilon_y/\varphi, -\upsilon_x/\varphi)$; the creep forms a rotating field around this point. This is the reason why it is lacking in Fig. 34(d) (pure longitudinal creepage). The rolling direction is throughout from left to right. Fig. 34 is Fig. 5.19 of Kalker [12], Chapter 5.

In Fig. 34(a) and 34(a') we show the case of pure spin, $\upsilon_x = \upsilon_y = 0$. The stick area has its characteristic pointed form. The traction forms nearly a rotating field, but it is clear that a net lateral force results. On the line $x-x$, which is the path of a particle, the traction distribution $|(p_x, p_y)|$ looks very much like the 2D Carter distribution (Fig. 26 of this Section), with as vertical tangent at the stick-slip boundary, see Fig. 34(a').
Fig. 34(b) shows the case of combined longitudinal creepage and spin, $\upsilon_y = 0$. Here it is clear that there will be a net longitudinal and lateral component of the tangential force. Fig. 34(c) shows the case of combined lateral creepage and spin ($\upsilon_x = 0$); the traction becomes more directed along the y-axis, as compared to Fig. 34(a).

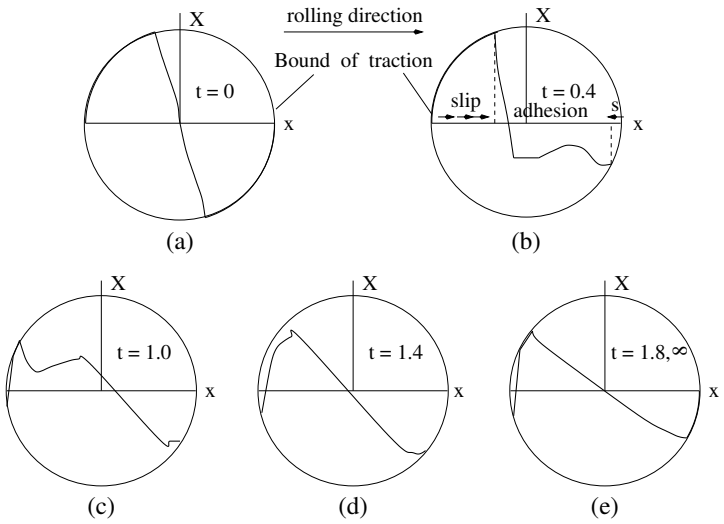


Figure 33: Frictional compression followed by rolling of dissimilar cylinders. (Source: Kalker [11])

Fig. 34(d) shows the case of pure longitudinal creepage, $v_y = \varphi = 0$; all tractions are almost parallel to the x-axis.
Fig. 34(d') shows $|p_x|$ along the line $x-x$; again the impression is one of Carter's traction distribution. Note that with almost the same stick-slip bound at $x-x$, the present graph is "thinner".
Fig. 34(e), finally, shows the case of large spin, with almost no stick area; the field is fully rotating, and the lateral force is lower than in Fig. 34(a), which also represents pure spin.

6.3.2 Quasiidentical steady state rolling: subsurface stresses

The subsurface stresses are important in strength and endurance calculations. They were first calculated by Ahmadi [1]. They can be computed by a special module of CONTACT. In Kalker [12], an algorithm is presented to calculate the displacements and the displacement gradients on and inside the elastic half-space $z \geq 0$ due to a uniform load of arbitrary direction acting on a rectangle on the surface of a half-space. The half-space is homogeneous and isotropic, with modulus of rigidity G and Poisson's ratio ν.

The most important quantities that determine the strength of the material are the first and second invariants of the stress, *viz.*

$$\sigma_{ii} \quad : \quad \text{first invariant} \tag{252}$$

$$s_{ij}s_{ij} \quad : \quad \text{second invariant, with}$$

$$s_{ij} = \sigma_{ij} - \sigma_{hh}\delta_{ij}/3 \quad : \quad \text{stress deviator} \tag{253}$$

The second invariant yield an ideal stress σ_I,

$$\sigma_I = \sqrt{s_{ij}s_{ij}} \tag{254}$$

The von Mises yield criterion of plasticity is a bound on the ideal stress:

$$\sigma_I \leq k, \quad \text{with } k \text{ the yield stress} \tag{255}$$

Fig. 35 shows σ_I and $-\sigma_{ii}$ on the z-axis for a uniform load on a square, centered at the origin, and loaded

1. by a uniform traction in the z-direction, of unit intensity;
2. by a purely shearing traction of unit intensity, in the direction of a side of the square.

Poisson's ratio $\nu = 0.28$, $G = 1$.

In case 1, the purely normal load, shown by the full lines in Fig. 35, the ideal stress σ_I has a maximum of 0.55 at about 0.4 side-length under the surface. This behaviour is well-known. $-\sigma_{ii}$ has a boundary maximum of 2.6, and drops rapidly to where its value meets the falling-off part of the curve of σ_I at $z = 1.40$.
In case 2 $\sigma_{ii} = 0$. σ_I (the broken line in Fig. 35) behaves like $-\sigma_{ii}$ in case 1, starting at $\sigma_I = 1.42$ on $z = 0$, but dropping much more rapidly so that it has virtually vanished at $z = 1.67$ side length.

Fig. 35 can be used to assess the quality of the half-space approximation. The stress behaviour is dominated by the normal pressure from a depth of about 1.40 times the contact area diameter. Looking at the drawn lines of Fig. 35 we see that at a depth of about three contact area diameters the the loads the ideal stress has dropped to 10 % of its maximum value, and $-\sigma_{ii}$ to about 2.5 % of its maximum value. So it seems safe to assume that the stresses have almost died out at that level. This supports the statement that the half-space approximation is justified when the diameter of contact is less than 1/3 of the diameter of the contacting bodies.
At a depth of 5 times the diameter of contact, the numbers are 1 % and 0.25 %, respectively.

6.3.3 The total force transmited in rolling

A very important quantity of rolling contact in technology is the total force transmitted. It finds application in vehicle dynamics, both in rail vehicles and in motor cars. As to motor cars, more will be heard of it in the lectures of Prof. Pacejka. The car tyre is, however, so complicated that a rigorous theory of it is still beyond our ken. In rail vehicles, we have the advantage that we can regard the contact as quasiidentical, and fortunately the theory for that is well developed.

Speed is of the essence in the total force theories. In vehicle dynamics, an something like a million steps must be taken in order to advance the real time by one second. And

in other applications speed is a pleasant advantage because one does not have to wait a long time before getting accurate results.

We have the following theories.

- **The linear theory** is valid for longitudinal, lateral and spin creepage which are infinitesimally small. The theory is described in Section 2. The quantities determining the total force components are called the creepage and spin coefficients; they are tabulated in Section 2, Table 3; they hold only for quasiidentical, Hertzian bodies.
- **The theory of Vermeulen and Johnson** is valid for unrestricted longitudinal and lateral creepage and zero spin. It is an empirical formula with a theoretical background. It is described in Section 2. It holds only for quasiidentical, Hertzian bodies.
- **FASTSIM** is an algorithm based on the simplified theory of rolling contact. It is very fast but approximate. It can be used for unrestricted creepage and spin, but is confined to quasiidentity, and to Hertzian bodies. It is described in Section 3.
- **CONTACT** is an algorithm based on the true theory of elasticity. Owing to the fact that a discretization is used it is approximate. It is about 2000 times as slow as FASTSIM. It is the only algorithm which can be used for non quasiidentical and non-Hertzian bodies. It is described in the present Section 5.
- **Table Book.** A book of tables has been constructed in which one can interpolate linearly to obtain the total force and the twisting moment from the longitudinal, lateral and spin creepages and the ratio of the axes of the Hertzian contact ellipse. The construction has taken place with CONTACT, and comprises 115,000 entries (4.5 MByte). Poisson's ratio has been fixed to 0.28, the value for steel, for the Table Book is intended for railway use. The programme using the Table Book is about 8 times as fast as FASTSIM, but it needs this enormous amount of storage space, while FASTSIM needs only 41 kB and can be used for unrestricted Poisson's ratio. But in many cases FASTSIM's error exceeds that of the Table Book. The Table Book is only for quasiidentical, Hertzian bodies with Poisson's ratio equal to 0.28.

6.4 3D Transient rolling

6.4.1 Quasiidentical bodies

The present section is based on Kalker (1990) sec. 5.2.2.5. We consider Kalker's Fig. 22, Section 5, (here Fig. 36): Quasiidentical transient rolling. The calculations have been made by CONTACT. Considerable smoothing and editing of the figures has taken place; the figure is an interpretation of the true result. Two identical spheres are compressed and rolled over each other.
Radius spheres: $R = 337.5$, $G_1 = G_2 = G = 1$, $\nu_1 = \nu_2 = \nu = 0.28$. F_z = Constant =

$0.4705 = (7/9)^3$; $a = b = 3.5$. $f = 0.4013$.
Radius stick area "Cattaneo": $0.7a = 2.45$. F_x = Constant throughout = $-f \times F_z \times 0.657 = -0.1240$, $v_y = \varphi = 0$.
After the Cattaneo shift the spheres roll with constant force $(F_x, 0, F_z)$ without spin, with velocity $V = 1$. Step is $Vt = 0.5$.
Elements: squares with sides 1.
Potential contact: 9×9, center in origin.

In the figures p_x is shown for various values of y (upper 4 rows). Full line: p_x, dotted line: traction bound fp_z.
In the lowest row, we show the areas of slip (S) and adhesion (A). The columns correspond to $Vt = 0, 1, 3, 3, 5, 7 = \infty$. The traction distributions are very much like those in the two-dimensional case (Fig. 30).

6.4.2 Non-quasiidentical bodies

We start from the Spence compression, *viz.* a sphere with radius 243, modulus of rigidity $G = 2$, and Poisson's ratio $\nu = 0$ pressed onto a flat, rigid slab. The friction coefficient is $f = 0.4013$. The final radius of contact is 3.5 units. Then rolling starts in the x-direction, with creepage and spin kept zero. The surface is discretized into squares with side 1; the potential contact is a square with side 7. The distance traversed $V(t - t')$ is discretized into steps of 0.2 units. The results are shown in Fig. 37. This figure is similar to Fig. 36. In the basic figures the absolute value of the tangential traction $|p_\tau|$ is shown drawn, together with the traction bound fp_z (broken line), as a function of the rolling coordinate x, with the lateral coordinate y as a parameter. The tangential traction is mirror symmetric about the x-axis. The arrows under the x-axis represent the direction of the traction. The lowest row represents the contact area which is taken circular, and its division into regions of Slip (S) and Adhesion (A). The columns depict the situation when the distance traversed is $Vt = 0, 1, 2, 3, 4, 5, 7, 10, 13$. At the final position, which represents almost 2 contact widths traversed, the steady state has been virtually attained.

Fig. 37 was made after considerable smoothing and editing.

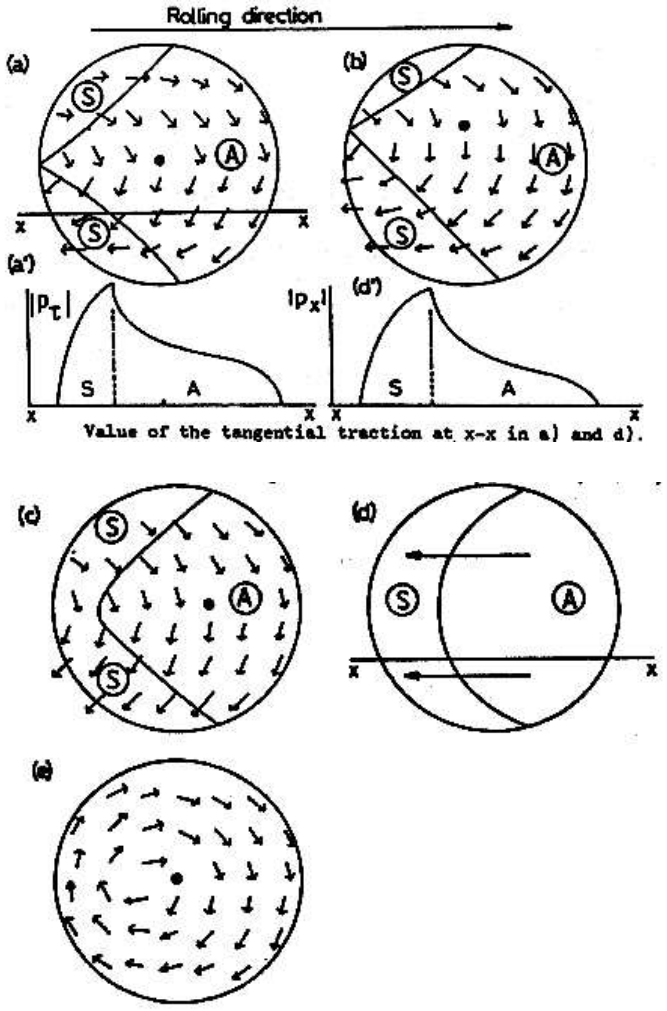


Figure 34: Contact stresses in quasiidentical steady state rolling with circular contact area. (Source: Kalker [12])

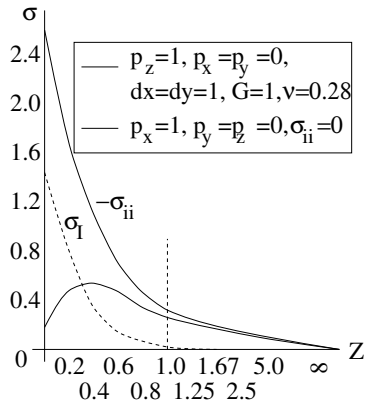


Figure 35: Subsurface stresses in the half-space $z \geq 0$. (Source: Kalker [12] Fig. 20, Section 5)

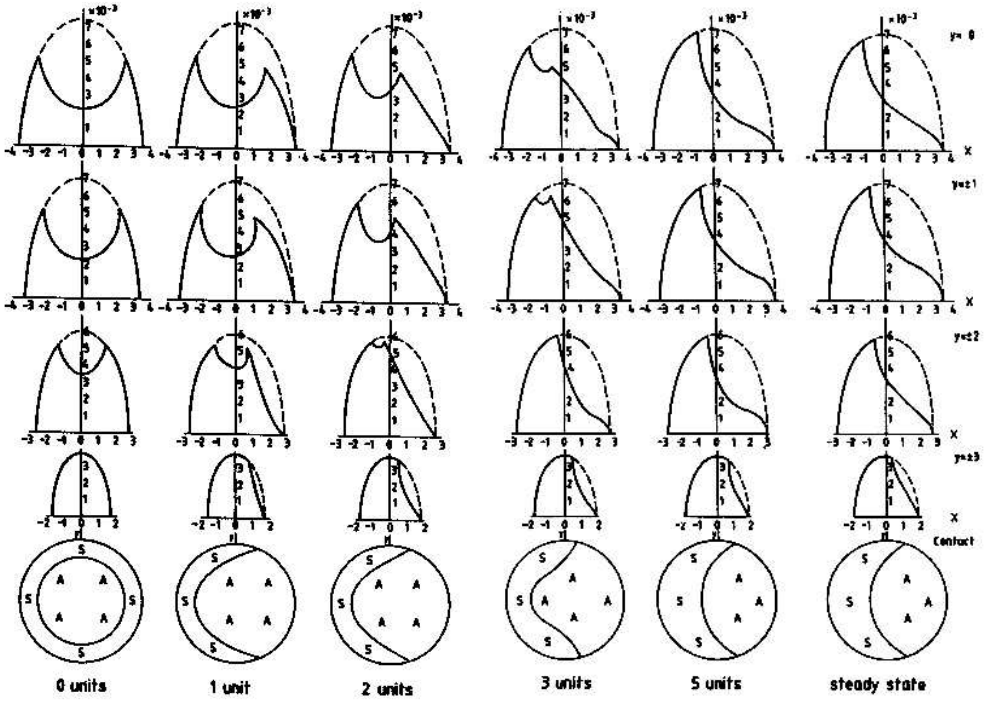


Figure 36: 3D quasiidentical tansient rolling

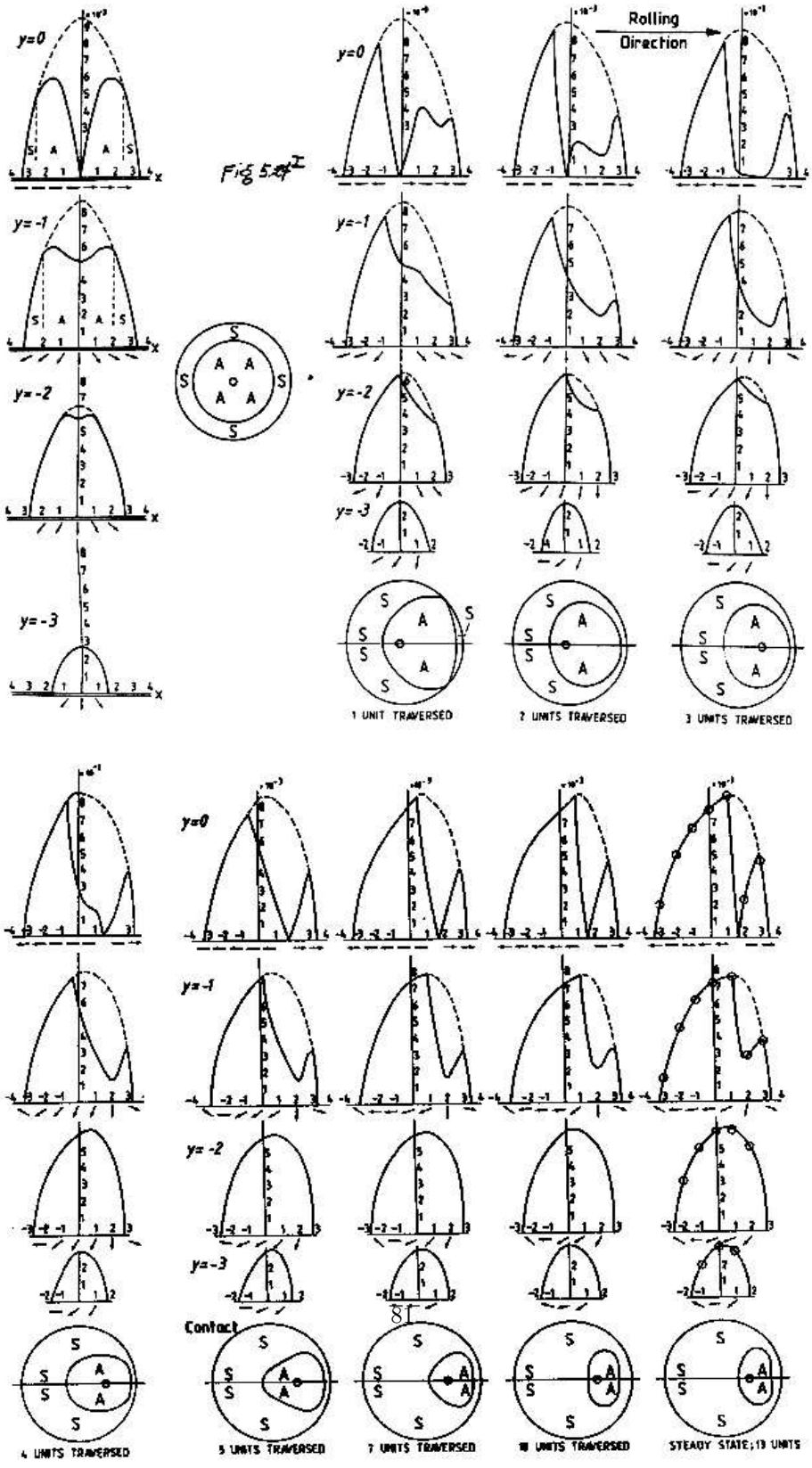


Figure 37: 3D non-quasiidentical transient rolling

7 Further and alternate research

1. The work of Bufler [5].
   In this work, Bufler solves the non-quasi-identical problem of two 2-D bodies rolling over each other with infinite friction (or infinitesimal creepage). An analytical, closed form solution is found.

2. The works of Bentall and Johnson ([2] and [3]).
   The work from 1967 considers two non-quasi-identical 2-D cylinders rolling over each other with finite creepage. This work is, in content, a generalisation of Bufler's work (see 1.). The method is a boundary element method in which the traction elements are continuous and piecewise linear. The work from 1968 concerns a strip between two rolling cylinders. Here also, the traction elements are continuous and piecewise linear. The solutions are numerical with a partly heuristic algorithm.

3. The work of Braat, Kalker and Saes.
   This work was published by Kalker[13] and Braat and Kalker [4]. It concerns rolling bodies consisting of different flat layers which are glued together. The geometry is 2-D. Examples are: Rubber on steel, steel on aluminium, various kinds of plastic glued onto each other, etc. The rolling bodies are cylinders which are approximated by layered half-spaces (2-D). Only steady-state rolling is considered, but the layers may be elastic or viscoeleastic.
   The solution is found by the boundary element method and the CONTACT algorithm. The Panagiotopoulos method is used throughout, as the rolling bodies are not symmetric. The element equations are found with the aid of an Airy function $H(x, z)$ which must satisfy a bipotential equation

   $$H_{,xxxx} + 2H_{,xxzz} + H_{,zzzz} = 0$$

   where x is the coordinate in rolling direction and the axis of z points vertically into the upper cylinder. The displacements and the stresses are derivatives of H with respect to x and z. A Fourier transform is performed in the x-direction, and the bipotential equation reduces to an ordinary differential eauation in z which can be solved. Inverse transformation then yields the element equations, which can be inserted into the CONTACT algorithm.

4. The work of Gross-Thebing [8].
   This work concerns the linear theory of rolling contact when the bodies are 3-D and the creepage is a harmonic function of the time and friction is infinite. Consequently, the tangential contact force is likewise harmonic, but the creepage and the force are not in phase. The theory will be covered extensively in a lecture by Knothe during this course.

5. Kalker's thesis work [10].
   This concerns 3-D rolling bodies, quasi-identical, and with an elliptic Hertzian contact area. As distinct from the works considered up to now, it is not a boundary

element method with rectangular elements which carry a constant or continuous, piecewise linear load, but inside the contact ellipse the displacements are polynomials; it can then be shown that the tractions which vanish outside the contact ellipse are polynomials of the same degree as the displacements, but multiplied with a certain function. This is applied to find the connection between the creepage and spin on the one hand and the total force and the twisting moment on the other hand (creepage-force law) for infinite friction, or in other terms infinitesimal creepage and spin. It is also applied to find an approximate creepage-force law for finite creepage and spin, but the process involved does not always converge.

6. Nielsen's thesis work [15].
   This concerns 2-D rolling bodies, quasi-identical, with a contact area in the form of one or more strips. It is a generalisation of Kalker's work, see 5. The importance of the work lies in the applications: apart from Carter's work and an elementary generalisation of it Nielsen treats corrugation (among other things the growth or diminishing of a corrugation field by abrasive wear), velocity dependent coefficient of friction, asperity fields (in which we have multiple contacts), and non-steady 2-D contact.

7. The strip theory (Haines/Ollerton [9]).
   The work of Carter [6] and Nielsen [**?**] may also be applied as a strip theory, in which 2-D solutions are placed beside one another to form a 3-D approximate solution. The idea to do this is due to Haines and Ollerton, who did experiments to verify their theory. The work of Nielsen and Haines/ Ollerton is confined to longitudinal creepage. Kalker [10] extended the work of Haines and Ollerton to lateral creepage and spin. Notably the spin solution is approximate. Note that the simplified theory, treated in Section 3, is also a strip theory.

8. Kalker and Piotrowski [14] investigated the case that the friction coefficient has two values, *viz.* $f = f_{kin}$ when there is slip, and $f = f_{stat}$ when there is no slip. Fastsim was used. The effect on the total tangential force appeared to be small ($f = f_{kin}$).

References of Section 6

[1] N. Ahmadi (1982), *Non-Hertzian normal and tangential loading of elastic bodies in contact*. Ph.D Thesis, Appendix C2. Northwestern University, Evanston IL, USA.

[2] R.H. Bentall, K.L. Johnson (1967), *Slip in the rolling contact of two dissimilar elastic cylinders*. International Journal of Mechanical Sciences **9**, p. 389-404.

[3] R.H. Bentall, K.L. Johnson (1968), *An elastic strip in plane rolling contact*. International Journal of Mechanical Sciences **10**, p. 637.

[4] G.F.M. Braat, J.J. Kalker (1993), *Theoretical and experimental analysis of the rolling contact between two cylinders coated with multilayered viscoelastic rub-*

ber, in: A.H. Aliabadi and C.A. Brebbia (Eds.) Contact Mechanics, Computational Techniques; Contact Mechanics Publications, pp. 119-126.

[5] H. Bufler (1959), *Zur Theorie der rollenden Reibung*. Ingenieur Archiv **27**, p. 137.

[6] F.C. Carter (1926), *On the action of a locomotive driving wheel*. Proceedings of the Royal Society of London, **A112**, p. 151-157.

[7] H. Fromm (1927), *Berechnung des Schlupfes beim Rollen deformierbarer Scheiben*. Zeitschrift für angewandte Mathematik und Mechanik **7**, p. 27-58.

[8] A. Gross-Thebing (1992), *Lineare Modellierung des instationären Rollkontaktes von Rad und Schiene*. Fortschritt Berichte VDI, Reihe 12, Nr. 199, VDI-Verlag, Düsseldorf.

[9] D.J. Haines, E. Ollerton (1963), *Contact stresses in flat elliptical contact surfaces which support radial and shearing forces during rolling*. Proceedings of the Institute of Mechanical Engineering Science, **179**, Part 3.

[10] J.J. Kalker (1967), *On the rolling contact of two elastic bodies in the presence of dry friction*. Thesis Delft.

[11] J.J. Kalker (1971), *A minimum principle of the law of dry friction with application to elastic cylinders in rolling contact*. Journal of Applied Mechanics, **38**, p. 875-880, 881-887.

[12] J.J. Kalker (1990), *Three-dimensional elastic bodies in rolling contact*. Kluwer Academic Publishers (Dordrecht) 314+XXVI pp.

[13] J.J. Kalker (1991), *Viscoelastic multilayered cylinders rolling with dry friction*. Journal of Applied Mechanics, **58** p. 666-679.

[14] J.J. Kalker, J. Piotrowski (1989), *Some new results in rolling contact*. Vihicle System Dynamics **18**, p. 223-242.

[15] J.B. Nielsen (1998), *New developments in the theory of wheel/rail contact mechanics*. Thesis Lyngby IMM-PHD-1998-51.

[16] D.A. Spence (1975), *The Hertz problem with finite friction*. Journal of Elasticity, **5**, p. 297-319.