Journal of Engineering Mathematics, Vol. 6, No. 2, April 1972
Wolters-Noordhoff Publishing–Groningen
Printed in the Netherlands

A Minimum Principle for Frictionless Elastic Contact with Application to Non-Hertzian Half-Space Contact Problems

J. J. KALKER

Institute of Mathematics, Delft University of Technology, Netherlands

Y. VAN RANDEN

SKF European Research Centre, Jutphaas, Netherlands

(Received September 10, 1971)

SUMMARY

A variational principle governing the frictionless contact between two elastic bodies is established, which is valid both for linear and for non-linear elasticity. In the case of linear elasticity it appears to lead to an infinite dimensional convex quadratic programming problem. It is applied to the half-space geometry in linear elasticity and it is established that non-Hertzian normal half-space contact problems are physically meaningful.

A Hertzian and a non-Hertzian normal contact problem are investigated numerically, to which end the principle is discretised on a triangular network. In the case of the Hertz problem it is found that the exact relationships between penetration, maximum pressure, and total normal force are well satisfied. The form of the contact area is given only crudely, unless the discretisation network is considerably refined. It appeared that such a refinement is only necessary close to the edge, in which case passable results will be obtained.

1. Introduction

When two elastic bodies are pressed together, then, as a consequence of their elastic properties, an area will be formed where they are in contact. If there is no friction between the bodies in this contact area, the contact stress must be normal to the boundary of the bodies. This type of contact, which is dealt with in this paper, is called *normal contact.*

When the geometry of the surfaces of both the bodies, their elastic properties, and the total compressive force or the undeformed penetration of both the bodies into each other are given, the objective is to find the contact area and the pressure acting in it.

The uniqueness of the solution of this problem in the case of linear elasticity has been proved by Kalker [1].

Even in the simplest case, that is, when the bodies are considered to be elastic half spaces, the solution of this problem has only been given for special geometries of the surfaces of both the bodies.

Hertz (1881) gave an analytical solution for quadratic surfaces (see Love [2] p. 193–198), and Galin [3] described solutions, obtained by various authors, for more complicated surfaces, but with a given boundary of the contact area.

In the present paper a numerical method will be given which, in principle, is applicable to any kind of geometry of the contacting bodies.

However, the method utilises the normal surface displacement field due to a concentrated normal load, which is known in readily computable form only for a limited class of geometries of which the half space is a pre-eminent member. Indeed, the examples treated in this paper (sect. 3 and 4) are half space problems in the linear theory of elasticity.

The method is based upon mathematical programming theory and practice. The connection between one-sided constraints in mechanics and mathematical programming was pointed out earlier by Moreau, see, e.g., [8].

The connection between mathematical programming and the normal contact problem is the theme of a recent paper by Conry and Seireg [9], which appeared after we finished our researches. Their point of departure is eq. (24a) of the present paper,

$$P \geqq 0, \quad \rho \geqq 0, \quad P\rho = 0 \text{ in } D \tag{a}$$

D: region of the surface of the lower body containing the contact area;
P: pressure exerted by the bodies on each other at the surface;
ρ: distance of a point of the lower body to the upper, taken positive when no penetration takes place.

Note that in (a) no distinction is made between the contact area and the surface of the bodies outside it: the contact area is a result of the calculation. Both our method and the method of Conry and Seireg possess this characteristic feature.

In [9], the problem (a) is transformed to a quadratic programming problem in the following manner. Equivalent to (a) is the system

$$P \geqq 0, \quad \rho \geqq 0, \quad \int_D P\rho \, dA = 0, \tag{b}$$

and equivalent to system (b) is, under the supposition that a solution of (a) exists,

$$P \geqq 0, \quad \rho \geqq 0, \quad \int_D P\rho \, dA = \text{minimal} \tag{c}$$

An argument of this type was given before in [10] p. 32. Since in the case of linear elasticity, to which Conry and Seireg confine themselves, the distance ρ is a first-degree functional of the pressure P, (c) is a quadratic programming problem of which they prove strict convexity, and hence uniqueness of the solution.

Viewed as a quadratic programming problem the system (c) has the disadvantage that it contains many constraints, for, apart from the constraints $P \geqq 0$, there are the constraints that $\rho \geqq 0$ in every point under consideration. Consequently Conry and Seireg successfully develop a modification of the simplex algorithm of linear programming by means of which the relations (a) can be solved directly.

In the present paper the point of departure is that the internal energy of a system of elastic bodies under isentropic conditions is minimal, together with the non-penetration property $\rho \geqq 0$.

This leads to a mathematical programming problem of which the relations (a), (24a) are the Kuhn–Tucker relations characterising a minimum.

When the constitutive equations of the bodies are linear, a quadratic programme is obtained which is very similar to (c), but in which the relations $\rho \geqq 0$ do not figure as additional constraints. As a consequence, it becomes attractive to solve the quadratic programme, after discretisation, by standard methods such as Wolfe [6] and Beale [7].

In practice we preferred to modify Wolfe's method slightly so that in each iteration step of the simplex method of linear programming upon which Wolfe's method is based, a meaningful normal contact problem is solved.

It appears from the examples that the pressure distribution, which is discretised as a piecewise linear function on a network of triangles (a facette function), is given with good accuracy in points which are not too close to the boundary of the contact area, even when the network is crude. The boundary of the contact area, however, needs a very fine mesh of triangles to be ascertained with any confidence.

A preliminary report was given in the 1970 GAMM Tagung at Delft [5].

2. The Minimum Principle

Consider two elastic bodies in contact, see fig. 1, in which the lower body (1) and part of the upper body (2) are shown. The distance of a point $\boldsymbol{x}$ of the surface of (1) to body (2) in the deformed state is denoted by $|\bar{\rho}|$. $\bar{\rho}$ is taken positive when the bodies do not penetrate at $\boldsymbol{x}$, and negative when they do. The function $\rho(\bar{\rho}, x)$ satisfies the following requirements:

ρ is once differentiable with respect to x and twice in $\bar{\rho}$;

$\bar{\rho}=0 \Leftrightarrow \rho=0$;

$$\frac{\partial\rho}{\partial\bar{\rho}} \to 1 \text{ as } \bar{\rho} \to 0 \tag{1}$$

In other terms, when the bodies are close together, ρ coincides with $\bar{\rho}$, but apart from that ρ is practically arbitrary. Since the actual value of $\bar{\rho}$ is here of importance only when the bodies are close together, we are justified in using the term "distance" when we refer to ρ.

The surface of the bodies is assumed to consist of four regious, denoted by I to IV, see fig. 1.

I: There is no contact between the bodies, hence ρ, the distance in the deformed state of a point of the surface of the lower body to the upper body is positive:

$$\rho>0 \text{ in I} \tag{2}$$

II: The bodies make frictionless contact with each other, and no adhesion of the upper surface to the lower surface is present. The distance vanishes:

$$\rho=0\,, \text{ no friction or adhesion in II} \tag{3}$$

III: The bodies adhere to each other. The distance vanishes:

$$\rho=0\,, \text{ adhesion in III} \tag{4}$$

IV: In this region, the body is clamped, that is, no displacement can occur:

$$\boldsymbol{u} = \text{prescribed in IV}\,, \tag{5}$$

$\boldsymbol{u}$: displacement of the bodies in accordance with the laws of elasticity and the surface loads (6)

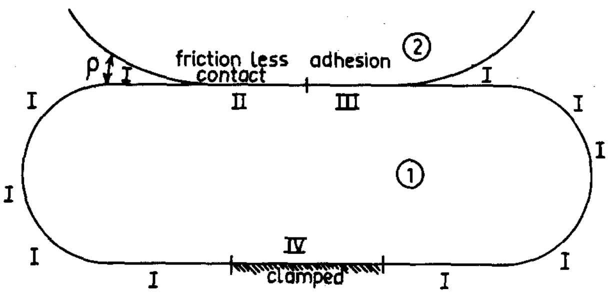

Figure 1. Two bodies in contact.

We will denote the displacement of the surface of body (i) by $\boldsymbol{u}_i$, and the traction exerted on the surface by $\boldsymbol{p}_i$.

A variation the internal energy U of the system is given by

$$\delta U = \delta Q + \delta W \tag{7}$$

where W is the work done on the system and Q is the heat supplied to the system.

In the adiabatic case, $\delta Q=0$ so that

$$\delta U = \delta W\,, \text{ adiabatic case} \tag{8}$$

which means that the virtual work done on the system is a total variation. When all changes take place isentropically, the condition of equilibrium is that U is minimal, or, in variational terms:

$$\delta S=0 \;;\; \delta W = \delta U \geqq 0\,, \text{ for all feasible variations of the energy} \tag{9}$$

where S means entropy.

Now,

$$\delta U = \delta W = \int_{\mathrm{I}\cup\mathrm{II}\cup\mathrm{III}\cup\mathrm{IV}} (\boldsymbol{p}_1 \delta \boldsymbol{u}_1 + \boldsymbol{p}_2 \delta \boldsymbol{u}_2) dA \geqq 0 \tag{10}$$

or, since the variations in one region are independent of those in another,

$$\int_i (\boldsymbol{p}_1 \delta \boldsymbol{u}_1 + \boldsymbol{p}_2 \delta \boldsymbol{u}_2) dA \geqq 0 \,. \tag{11}$$

Each region is considered in turn.

I: (see (2)). Here, the distance ρ is non-zero, and consequently $\delta \boldsymbol{u}_1$ and $\delta \boldsymbol{u}_2$ are arbitrary. So,

$$-P = \boldsymbol{p}_1 = \boldsymbol{p}_2 = 0 \; ; \int_{\mathrm{I}} = 0 \; ; \tag{12}$$

where P is the pressure exerted on boundary.

II: (see (3)). Let $\boldsymbol{n}_i$ be the outer normal to body i and let $\boldsymbol{t}_i$ be the direction of the projection of the surface traction $\boldsymbol{p}_i$ on the surface. Let u_{ni}, p_{ni}, u_{ti}, p_{ti} be the components of the displacements and tractions in the $\boldsymbol{n}_i$, $\boldsymbol{t}_i$ directions respectively. Then:

$$\boldsymbol{p}_i \delta \boldsymbol{u}_i = p_{ni} \delta u_{ni} + p_{ti} \delta u_{ti} \tag{13}$$

δu_{ti} is arbitrary owing to the absence of friction, hence

$$p_{ti} = 0 \,. \tag{14}$$

Also, in II, $\boldsymbol{n}_2 = -\boldsymbol{n}_1$, so that

$$-\delta\rho \equiv \delta u_{n1} + \delta u_{n2} \leqq 0 \text{ in II} \,, \tag{15}$$

since $\rho = 0$, and the bodies cannot penetrate. First we take

$$\delta u_{n1} = -\delta u_{n2} = \delta u = \text{arbitrary in II} \,. \tag{16}$$

Then it is found from (11), (13), and (14), that

$$p_{n1} = p_{n2} = -P \,, \; P\text{: pressure} \;;\; \text{II} \,. \tag{17}$$

So we must have

$$\int_{\mathrm{II}} = \int_{\mathrm{II}} -P(\delta u_{n1} + \delta u_{n2}) dA \geqq 0 \,, \; (15) \Rightarrow P \geqq 0 \,. \tag{18}$$

III: (see (4)). Owing to adhesion, $\delta \boldsymbol{u}_1 = \delta \boldsymbol{u}_2 =$ arbitrary, and hence $\boldsymbol{p}_1 = -\boldsymbol{p}_2$:

$$\boldsymbol{p}_1 = -\boldsymbol{p}_2 \; ; \; \int_{\mathrm{III}} = 0 \; ; \; \rho = 0 (\mathrm{III}) \,. \tag{19}$$

IV: As a consequence of (5), $\delta \boldsymbol{u}_i = 0$, and hence

$$\boldsymbol{p}_i = \text{arbitrary} \,, \; \delta \boldsymbol{u}_i = 0 \,, \; \int_{\mathrm{IV}} = 0 \,. \tag{20}$$

Since by (12) and (15)

$$P(\delta u_{n1} + \delta u_{n2}) = -P\,\delta\rho \text{ in I and II} \tag{21}$$

the variation of the energy can according to (12), (18), (19), (20) be written in the form

$$0 \leqq \delta U = \delta w = \int_{\mathrm{I}\cup\mathrm{II}} -P(\delta u_{n1} + \delta u_{n2}) dA \; = \; + \int_{\mathrm{I}\cup\mathrm{II}} P\,\delta\rho\, dA \tag{22}$$

for all feasible variations of ρ.

The implications of eq. (22) are considered:

$$0 \leqq \delta U = \int_{\mathrm{I}\cup\mathrm{II}} P\,\delta\rho\, dA \,, \quad \rho \geqq 0 \,, \quad P \text{ and } \rho \text{ connected by laws of elasticity} \tag{23}$$

(1) I: $\rho > 0 \Rightarrow \delta\rho$ arbitrary $\Rightarrow P = 0$

(2) II: $\rho = 0 \Rightarrow \delta\rho \geqq 0$ (no penetration) $\Rightarrow P \geqq 0$ (no friction or adhesion). (24)

In I and II: $P \geqq 0\,,\ \rho \geqq 0\,,\ P\rho = 0\,.$ (24a)

(3) δU is a total variation, hence the right-hand side of (22) is also the total variation of the internal energy. In fact,

$$\frac{\delta U}{\delta\rho} \stackrel{\text{def}}{=} \lim_{\delta A \to 0} \frac{\int_{\delta A} P\,\delta\rho\, dA}{\int_{\delta A} \delta\rho\, dA} = P\,. \tag{25}$$

where P is the pressure exerted by one body on the other.

A Legendre transform is applied to U. Its result is called the contact enthalpy H:

$$H \stackrel{\text{def}}{=} \int_{\mathrm{I} \cup \mathrm{II}} P\rho\, dA - U\,; \tag{26}$$

in variational form it reads

$$\delta H = \int_{\mathrm{I} \cup \mathrm{II}} \delta(P\rho) dA - \int_{\mathrm{I} \cup \mathrm{II}} P\,\delta\rho\, dA = \int_{\mathrm{I} \cup \mathrm{II}} \rho\,\delta P\, dA\,, \tag{27}$$

$$P \geqq 0\,.$$

Evidently, it follows that $\int \rho\,\delta P$ is a total variation. Also, in equilibrium, we have by (24)

$$\left.\begin{array}{l} P = 0 \Rightarrow \rho \geqq 0 \text{ and } \delta P \geqq 0 \Rightarrow \rho\,\delta P \geqq 0\,; \\ P > 0 \Rightarrow \delta P \text{ is arbitrary}\,,\ \rho = 0 \Rightarrow \rho\,\delta P = 0 \end{array}\right\} \quad \text{in } \mathrm{I} \cup \mathrm{II} \tag{28}$$

so that in equilibrium

$$\delta H = \int_{\mathrm{I} \cup \mathrm{II}} \rho\,\delta P \quad dA \geqq 0 \Rightarrow H \text{ is minimal} \tag{29}$$

with the auxiliary condition that $P \geqq 0$, ρ and P connected by laws of elasticity.

Analogous to (25) we have

$$\rho = \frac{\delta H}{\delta P}\,. \tag{30}$$

So:

$$\begin{array}{l} H = \int P\rho\, dA - U\,; \quad \delta H = \int \rho\,\delta P\, dA\,, \quad P \geqq 0\,; \quad \rho = \dfrac{\delta H}{\delta P} \\ \qquad\qquad\qquad\qquad\quad \delta U = \int P\,\delta\rho\, dA\,, \quad \rho \geqq 0\,; \quad P = \dfrac{\delta U}{\delta\rho}\,. \end{array} \tag{31}$$

If both U and H are strictly convex, it follows that

minimise U under the auxiliary condition that $\rho \geqq 0$ (32a)

minimise H under the auxiliary condition that $P \geqq 0$ (32b)

are equivalent. H and U are strictly convex in linear elasticity.

The above theory is perfectly general and applies to any elastic system under the conditions of isentropy and absence of dissipation. The enthalpy minimum principle appears to be somewhat more fruitful in contact theory than the energy minimum principle.

2.1. *Specialisation to Linear Elasticity*

The enthalpy minimum principle will be formulated in the case of the linear, small displacement, small displacement gradient theory of elasticity. Owing to the smallness of displacements and gradients, the distance ρ can be written in the form of the sum of the distance $R(x)$ in the state when $P=\boldsymbol{u}=0$, and the normal displacement difference $w(x)=-u_{n1}-u_{n2}$:

$$\begin{aligned} &\rho(\boldsymbol{x}) = R(\boldsymbol{x})+w(\boldsymbol{x}) && \boldsymbol{x} \text{ surface point} \\ &w(\boldsymbol{x}) = -u_{n1}(\boldsymbol{x})-u_{n2}(\boldsymbol{x}) && R(\boldsymbol{x})\text{: distance in undeformed state}. \\ &R(\boldsymbol{x}) = \rho(\boldsymbol{x})|_{w=0}\,. \end{aligned} \tag{33}$$

Now, in linear elasticity we have, if $u=0$ in IV (see (5)), and if there is no adhesion (III $=\varnothing$):

$$U=\tfrac{1}{2}\int PwdA\,, \quad \boldsymbol{u}=0 \text{ in IV}\,, \ \text{III}=\varnothing \tag{34}$$

$$\begin{aligned} H &= \int P\rho\, dA-U = \int(P\rho-\tfrac{1}{2}Pw)dA = \int(PR+Pw-\tfrac{1}{2}Pw)dA \\ H &= \int(PR+\tfrac{1}{2}Pw)dA\,. \end{aligned} \tag{35}$$

Therefore, the enthalpy minimum principle becomes

$$\begin{aligned} &\min!\ H=\int(PR+\tfrac{1}{2}Pw)dA = U+\int_{\text{surface}} PR\,dA \\ &\text{auxiliary conditions: } P\geqq 0\,,\ R \text{ given}\ ; \\ &P \text{ and } w \text{ connected by laws of elasticity}\,. \end{aligned} \tag{36}$$

Now, strict convexity of the principle and uniqueness of the minimum will be established in the case of linear elasticity. Again, III is assumed to vanish. Let $\boldsymbol{K}_i(\boldsymbol{x}, \boldsymbol{x}')$ be the displacement in the point $\boldsymbol{x}$ of body i due to a concentrated unit load acting in the surface point $\boldsymbol{x}'$ of (i) in the direction of $\boldsymbol{n}_i$ under the auxiliary condition that $\boldsymbol{u}_i=\boldsymbol{0}$ in IV. When the pressure P vanishes, the equations of elasticity are satisfied by $\boldsymbol{u}_i\equiv\boldsymbol{0}$. This displacement field is unique, owing to Kirchhoff's uniqueness theorem of small displacement, small displacement gradient theory (see, e.g. Love [2], p. 170). The total displacement is hence given by

$$\boldsymbol{u}_i(\boldsymbol{x}) = \int_{\text{I}\cup\text{II}} \boldsymbol{K}_i(\boldsymbol{x}, \boldsymbol{x}')\{-P(\boldsymbol{x}')\}dA'\,, \tag{37}$$

and the elastic energy U_i, given by

$$\begin{aligned} U_i &= \int_{\text{I}\cup\text{II}} -\tfrac{1}{2}P(\boldsymbol{x})\boldsymbol{n}_i\cdot \boldsymbol{u}_i(\boldsymbol{x})dA = -\tfrac{1}{2}\int_{\text{I}\cup\text{II}} P(\boldsymbol{x})u_{ni}(\boldsymbol{x})dA \\ &= \int_{\text{I}\cup\text{II}} \tfrac{1}{2}\boldsymbol{n}_i\cdot \boldsymbol{K}_i(\boldsymbol{x}, \boldsymbol{x}')P(\boldsymbol{x})P(\boldsymbol{x}')dA\,dA' \end{aligned} \tag{38}$$

is a homogeneous quadratic form in $P(x)$; since U_i is strictly positive definite, the quadratic form is a strictly convex quadratic functional of the pressure distribution P. The displacement difference reads

$$w(\boldsymbol{x}) = -u_{n1}-u_{n2} = \int_{\text{I}\cup\text{II}} \sum_{i=1,2} \boldsymbol{n}_i\cdot \boldsymbol{K}_i(\boldsymbol{x}, \boldsymbol{x}')P(\boldsymbol{x}')dA' \tag{39}$$

so that the total internal energy

$$U = \int_{\text{I}\cup\text{II}} \tfrac{1}{2}w(\boldsymbol{x})P(\boldsymbol{x})dA = \int_{\text{I}\cup\text{II}} \frac{1}{2}\left[\sum_{1,2} \boldsymbol{n}_i\cdot \boldsymbol{K}_i(\boldsymbol{x}, \boldsymbol{x}')\right] P(\boldsymbol{x})P(\boldsymbol{x}')dA\,dA' \tag{40}$$

is likewise a strictly convex quadratic functional of the pressure distribution P. Since $\int PR\,dA$ is linear in P and hence also convex, the contact enthalpy

$$H=U+\int_{\text{I}\cup\text{II}} PR\,dA = \int_{\text{I}\cup\text{II}} P(\boldsymbol{x})R(\boldsymbol{x})dA + \int_{\text{I}\cup\text{II}} \frac{1}{2}\left[\sum_{1,2} \boldsymbol{n}_i\boldsymbol{K}_i(\boldsymbol{x}, \boldsymbol{x}')\right] P(\boldsymbol{x})P(\boldsymbol{x}')dA\,dA' \tag{41}$$

is a strictly convex functional of P. The auxiliary linear inequality condition $P \geqq 0$ defines a convex region of feasibility, so that the problem becomes an infinite dimensional strictly convex quadratic programming problem, of which the solution is unique, if we admit only continuous solutions.

3. Three-Dimensional, Non-Hertzian Half-Space Contact Problems

In the remainder of this paper the case will be considered that the lower body (1) is an elastic half-space made of a purely elastic material satisfying Hooke's law with modulus of rigidity G and Poisson's ratio σ, while the upper body (2) is a rigid indenter of which the edges are never in contact with the half-space. It is observed that the theory undergoes only very minor changes if the upper body (2) is of a form that can be suitably approximated by an elastic half-space with possibly different elastic constants (see [2] p. 193).

A coordinate system (x, y, z) is introduced in which the bounding plane of the half-space is the plane of x and y, and the positive z-axis has the direction of the inner normal to the half-space. The origin lies on the surface of the half-space, but is further unspecified.

It is well-known ([2], p. 191) that the normal component $\boldsymbol{n}_1\boldsymbol{K}_1$, of the surface displacement at (x, y) due to a unit concentrated load of magnitude 1 in the z-direction acting at (x', y') on the half-space is given by

$$\boldsymbol{n}_1\boldsymbol{K}_1(x, y, x', y') = (1-\sigma)/(2\pi G r)\,, \quad r = \{(x-x')^2+(y-y')^2\}^{\frac{1}{2}}\,. \tag{42}$$

The surface of the rigid indenter is given by

$$z = z(x, y) \quad \text{surface of indenter}\,. \tag{43}$$

Then, $-z(x, y)$ may be identified with the distance R introduced in sect. 2.1 eq. (33). For, when the rigid indenter is in contact with the half-space, the normal of the deformed half-space coincides with the normal of the indenter, and since linear elasticity is assumed to be valid, also almost with the normal on the undeformed half-space, which is the negative z-axis. So,

$$R(x, y) = -z(x, y) \quad \text{undeformed distance} \tag{44}$$

Of course, $R(x, y)$ must be somewhere negative, otherwise the bodies are not in contact and the problem is trivial.

The contact enthalpy minimum principle becomes, by (41), (44), (42)

$$\begin{aligned} &\min: H(z, P) = \int_D -P(x, y)z(x, y)dxdy + \int_D \int_D \frac{1-\sigma}{4\pi r G} P(x, y)P(x', y')dxdydx'dy'\,, \\ &r = \{(x-x')^2+(y-y')^2\}^{\frac{1}{2}}\ ; \ \text{auxiliary condition}: P(x, y) \geqq 0 \\ &D \text{ contains the contact area}\,. \end{aligned} \tag{45}$$

3.1. *The Existence of Non-Hertzian Half-Space Contact Problems*

The objection may be raised that the problem (45) has no physical significance except in the Hertzian case when $R(x, y)$ is quadratic; on the ground that else the almost parallelity of the normal on the indenter in the contact area to the z-axis cannot be guaranteed. It will now be shown that this is not so.

Consider any smooth function with bounded derivatives $z(x, y)$, and solve Problem (45) with it. Let $|t|$ be the maximum of the absolute value of the tangent of the angle of an outer normal to $z(x, y)$ and the z-axis for all points of the contact area. If $|t|$ is less than the maximum angle t_m for which the linear theory is acceptable, our case is proved. If not, let

$$\lambda = |t_m/t| > 0 \tag{46}$$

t: tangent of maximum angle occurring,
t_m: tangent of maximum angle acceptable,

and consider another contact problem of which the distance function z is given by

$$\bar{z}(x, y) = \lambda z(x, y) \tag{47}$$

and the pressure $\bar{P}$ is denoted by

$$\bar{P} = \lambda P \,. \tag{48}$$

It is easy to see from (45) that the contact enthalpy $H(\bar{z}, \bar{P})$ is connected with the contact enthalpy $H(z, P)$ of the original problem by

$$H(\bar{z}, \bar{P}) = \lambda^2 H(z, P) \,, \tag{49}$$

while it follows from the positivity of λ that $\bar{P} \geqq 0 \Leftrightarrow P \geqq 0$, so that the feasible regions of both problems (45) coincide. Hence, if $H(\bar{z}, \bar{P})$ is at a minimum, so is $H(z, P)$, and *vice versa*, and then the contact areas also coincide. But in that contact area the tangent of the maximum angle $\bar{t}$ of the new problem is given by

$$\bar{t} = \lambda t \Rightarrow |\bar{t}| = t_m \,. \tag{50}$$

In a realistic problem, let $z(x, y)$ be the equation of the surface of the rigid indenter, when it just touches the half-space.

The surface is given by

$$z(x, y) = z(x, y) + E \,, \quad E\text{: penetration} \tag{51}$$

when it is pressed into the half-space. Let D be the region of the (x, y) plane in which $|t| \leqq |t_m|$ on the indenter. This region contains the tangent point of the indenter with the half-space when $E=0$. The contact area grows around this point as E increases; so, when the penetration E is small enough, the contact area will lie entirely within the region D.

It remains to show that there exist smooth surfaces, which, however small the contact area, cannot be suitably approximated by a paraboloid.

Such a surface is, for example,

$$z(x, y) = -x^4 - y^4 + E \tag{52}$$

which has zero curvature at $x = y = 0$. Another example is a surface with shallow bumps, which will be treated as a numerical example, see sect. 4.2.

3.2. *Discretisation. Method of Solution. Implementation.*

The integral (45) is discretised on a mesh of triangles, of which an example is shown in fig. 2. Inside a triangle, the pressure P is assumed to be linear, so that the surface spanned by the pressure is diamond-like, with many triangular facets. The total normal displacement at a point (x, y) can be considered as the sum of the normal displacements at (x, y) due to loads in the form of a tetrahedron, see fig. 3. Three pressure tetrahedrons on the same triangles as base constitute one facet of the facet function. In the base of each tetrahedron, a local coordinate system (s, t) is introduced, see fig. 3. The vertex of the pressure lies at $(0, \gamma)$, and the basis $(\alpha, 0)$–$(\beta, 0)$ is free of pressure. γ is taken positive. The global coordinate system (x, y) is connected to the local system by a known translation and rotation. So the pressure on the triangular base of the tetrahedron is

$$P(s, t) = P_j t/\gamma \,, \quad P_j\text{: pressure at } (0, \gamma)\,, \text{ i.e. } (x_j, y_j)\,. \tag{53}$$

When the pressure P_j is unity, the normal displacement at (x_i, y_i) (local coordinates: $(\bar{x}_i, \bar{y}_i)$) due to the tetrahedron is

$$\bar{w}_{ijn} = \frac{(1-\sigma)}{2\pi G} \iint_{(\text{triangle } n)} \frac{t\,ds\,dt}{\gamma\{(\bar{x}_i - s)^2 + (\bar{y}_i - t)^2\}^{\frac{1}{2}}} \tag{54}$$

triangle n: base $[(\alpha, 0), (\beta, 0)]$, apex: $(0, \gamma)$ (i.e. (x_j, y_j)).

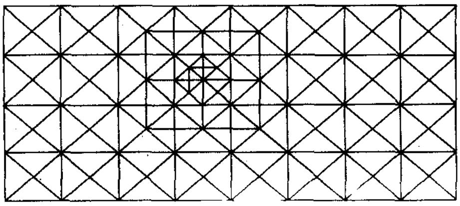

Figure 2. A triangular mesh with refinements.

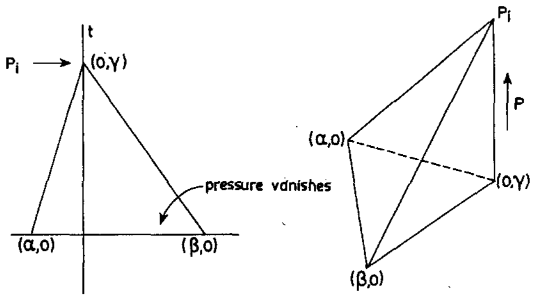

Figure 3. An element of pressure distribution.

The integral (54) can be evaluated analytically [4]. Let

$$I(k, l, t, y) = \tfrac{1}{2}(t^2-y^2)\,\mathrm{arcsinh}\left(\frac{kt+l}{|y-t|}\right) + \frac{ky+l}{2(k^2+1)^{\frac{1}{2}}}$$

$$\times\left[\left\{\left(t+\frac{kl-y}{k^2+1}\right)^2 + \left(\frac{ky+1}{k^2+1}\right)^2\right\}^{\frac{1}{2}} + \frac{y(k^2+2)-kl}{k^2+1}\,\mathrm{arcsinh}\left(\frac{(k^2+1)t+kl-y}{ky+l}\right)\right] \tag{55}$$

where $\mathrm{arcsinh}(x)$ is the inverse of the hyperbolic sine; then

$$\gamma\frac{2\pi G}{1-\sigma}\,\bar{w}_{ijn} = I(-\beta/\gamma, \beta-\bar{x}_i, \gamma, \bar{y}_i) - I(-\beta/\gamma, \beta-\bar{x}_i, 0, \bar{y}_i)$$

$$-I(-\alpha/\gamma, \alpha-\bar{x}_i, \gamma, \bar{y}_i) + I(-\alpha/\gamma, \alpha-\bar{x}_i, 0, \bar{y}_i) \tag{56}$$

The contribution of a unit pressure at the point (x_j, y_j) to the normal displacement in (x_i, y_i) is the sum of the integrals (54) over all triangles having (x_j, y_j) as a vertex:

$$w_{ij} = \sum_{n,\ \text{all triangles bordering on } (x_j, y_j)} w_{ijn}\ ; \tag{57}$$

the total displacement at (x_i, y_i) is then given by

$$w(x_i, y_i) = \sum_{j:\ \text{all } (x_j, y_j)} w_{ij} P_j\,. \tag{58}$$

Let A_j be the area of the union of all triangles bordering on (x_i, y_i). Then the total normal force is given by

$$N = \frac{1}{3}\sum_{i:\ \text{all } (x_i, y_i)} A_i P_i\ ; \tag{59}$$

where A_i is an area of all triangles bordering on (x_i, y_i).

An approximation of the first term of (45) is given by

$$\int_D P(x, y) R(x, y)\,dx\,dy \approx \frac{1}{3}\sum_i A_i P_i R_i\,, \qquad R_i = R(x_i, y_i) \tag{60}$$

and an approximation for the second term of (45) is

$$\int_D \int_D \frac{(1-\sigma)}{4\pi G r} P(x, y) P(x', y') dx dy dx' dy' = \frac{1}{6} \sum_{i,j} A_i w_{ij} P_i P_j . \tag{61}$$

Hence, problem (45) becomes

$$\min: \sum_i A_i P_i R_i + \frac{1}{6} \sum_{i,j} A_i w_{ij} P_i P_j$$
$$A_i, R_i, w_{ij} \text{ constant}, \quad P_i \geqq 0, \text{ all } i, \tag{62}$$

a quadratic programming problem, which, as the approximation of a convex mathematical programming problem, will, presumably, be convex.

It can be solved numerically by any quadratic programming routine, e.g. [6, 7]. In our implementation, Wolfe's method was used in a modified form described in [4]. This method generates by each successive iteration of the linear programming routine, upon which Wolfe's method is based, the solution of a normal contact problem corresponding to a rigid indenter which remains unaltered during the entire process, but in which the penetration E (see (51)) is different. In fact, the penetrations form an increasing sequence. The calculation time of each iteration is proportional to m^2, where m is the number of discretisation points (x_i, y_i) in D; each iteration yields a point of the penetration/total compressive force (E–N) curve.

When the matrix $A_i w_{ij}$ (see (62)) is given, the calculating time to obtain a contact area containing M points will, roughly speaking, be equal to M times the computer time needed for an iteration; that is, it is proportional to Mm^2. The computer time needed to calculate the matrix $A_i w_{ij}$ is proportional to $3m_2 m$, if m_2 is the number of triangles in the mesh. The constant of proportionality for the calculation of $A_i w_{ij}$ is, however, very much larger than the constant of proportionality of the iteration, owing to the complexity of the formulae (55) and (56). The calculating time needed for the matrix $A_i w_{ij}$ can be greatly reduced if, instead of an irregular mesh, a regular mesh is chosen, since in that case the coefficient $A_i w_{ij}$ depends only on $(x_i - x_j)$ and $(y_i - y_j)$. So, a factor of the order of m may be gained. In our implementation use was made of an irregular mesh, since we wished to have the possibility of refinement of the mesh, see fig. 2.

4. Numerical Results

The method outlined in sect. 3.2 was applied to the Hertz problem (sect. 4.1) and to a non-Hertzian problem (sect. 4.2).

4.1. *A Hertz Problem*

The form of the indenter (see (43), (44)) is given by

$$z(x, y) = -x^2/20 - y^2/5 + E, \quad E: \text{penetration}. \tag{63}$$

This Hertz problem was solved with a fairly regular triangular mesh of which a part is shown in fig. 4. In fig. 5, a contact area is shown containing 84 points with a non-zero pressure, i.e. $M = 84$. g, the ratio of the axes, should be 0.4 according to the theory. The cross in fig. 5 indicates the position of maximal pressure. The lines are isobars; the pressure difference between successive lines amounts to $\frac{1}{9}$ of the maximum pressure. The outermost line is the line of zero pressure, within which the contact area must lie. The exact boundary of the contact area is an ellipse. It is seen that the correspondence is bad. In fig. 6a is shown the pressure distribution

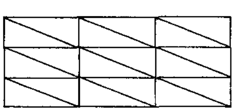

Figure 4. Part of the triangular mesh.

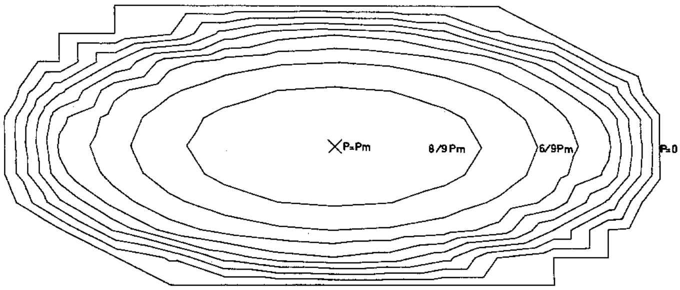

Figure 5. Contact area with isobars. P_m maximum pressure.

TABLE 1

Numerical and analytical values of N and P(0, 0) of a three dimensional Hertz-problem

$E \cdot 10$	M	$N_a \cdot 10^2$	$N_n \cdot 10^2$	perc. error N	$P(0,0)_a \cdot 10$	$P(0,0)_n \cdot 10$	perc. error $P(0,0)$
0.08697	10	0.0800	0.0820	2.5	0.0901	0.0963	6.88
0.1551	20	0.1906	0.1943	1.9	0.1203	0.1220	1.41
0.2384	30	0.3632	0.3674	1.16	0.1491	0.1510	1.27
0.3429	40	0.6265	0.6290	0.57	0.1789	0.1815	1.45
0.4000	50	0.7894	0.7951	0.72	0.1932	0.1952	1.04
0.5404	60	1.240	1.244	0.3	0.2245	0.2264	0.84
0.6010	70	1.454	1.460	0.4	0.2368	0.2384	0.68
0.7003	80	1.829	1.838	0.5	0.2556	0.2569	0.51
0.8006	87	2.235	2.241	0.3	0.2733	0.2747	0.51

over the x-axis, in the pressure distribution over the y-axis. The drawn circles represent the pressure according to the Hertz theory. It is seen that the correspondence is quite good. In table 1 is shown the penetration E, the total normal force obtained by analytic means (N_a), the total force as found numerically (N_n), the exact maximum pressure P_a, and the computed maximum pressure $P_n(0, 0)$. It is seen that the agreement of N_a and N_n is slightly better than that between P_a and P_n.

Since the contact area was so badly represented, it was decided to investigate the effect of local refinement of the net. The result is shown in fig. 7.

Fig. 7a shows the network, and fig. 7b the result. Again the isobars are at $\frac{1}{9}$th the maximum pressure. The outermost isobar is the numerically obtained boundary of the contact area, and the dots represent the exact contact ellipse, based upon the computed total force.

Two observations may be made. In the first place it is seen that the contact boundary in the quadrant where the mesh is refined is quite adequately represented. Secondly it is observed that this is so notwithstanding the crudeness of the net in the remaining three quadrants.

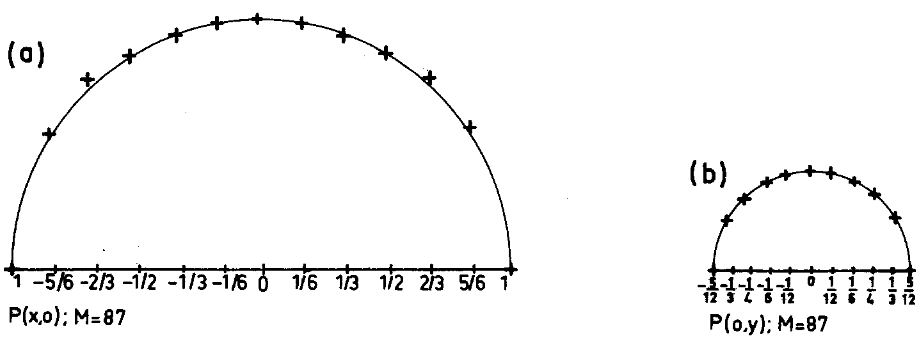

Figure 6. Pressure distributions.

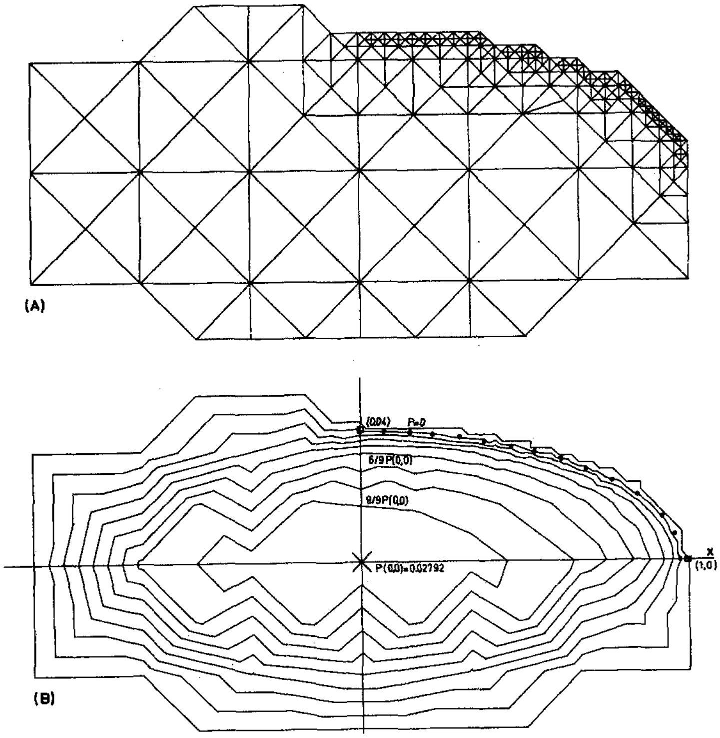

Figure 7. The Hertz problem with local refinements in the mesh. (a) The mesh (b) The result. Near $\frac{1}{9}P(0, 0)$ isobar in quadrant I: points of the exact boundary. $E=0.07886$. $N=0.02196$. $P(0, 0)=0.02792$ $R=x^2/20+y^2/5-E$. $W=\iint \{p(s, t)/r\}\, ds\, dt$. $r^2=(x-s)^2+(y-t)^2$.

4.2. *A Non-Hertzian Problem*

A non-Hertzian problem is treated in which

$$z(x, y)= -\tfrac{1}{10}(x^2-\tfrac{1}{4})^2-\tfrac{1}{10}(y^2-\tfrac{1}{4})^2+E\,. \tag{64}$$

This function has minima in the points $(\frac{1}{2}, \frac{1}{2})$, $(-\frac{1}{2}, -\frac{1}{2})$, $(-\frac{1}{2}, \frac{1}{2})$ and $(\frac{1}{2}, -\frac{1}{2})$, a maximum

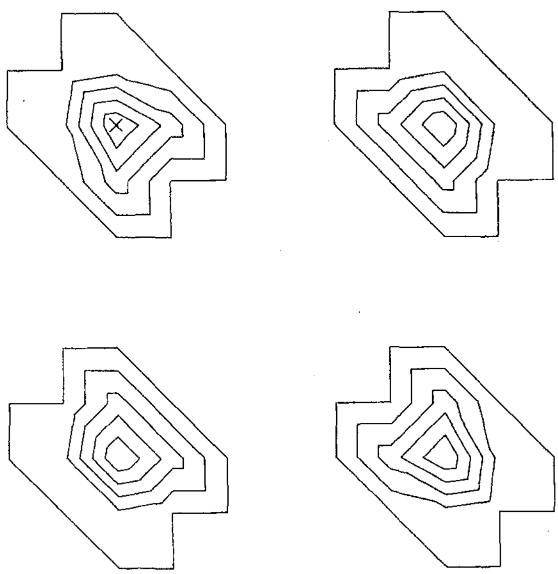

Figure 8a. Contact area with isobars $M=20$.

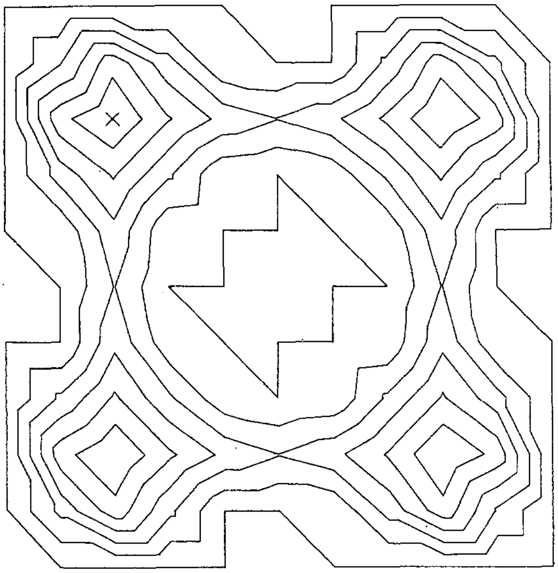

Figure 8b. Contact area with isobars. $M=60$.

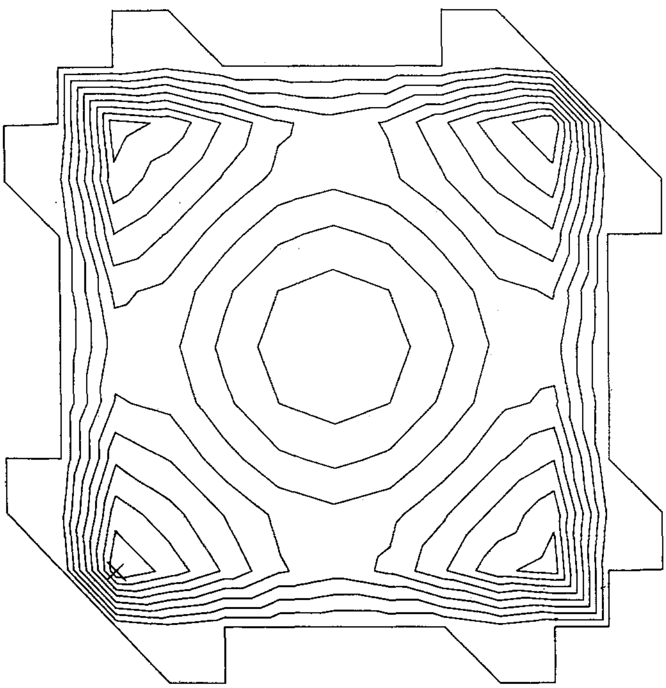

Figure 8c. Contact area with isobars $M=90$.

in (0, 0) and saddle points in $(\frac{1}{2}, 0)$, $(0, \frac{1}{2})$, $(-\frac{1}{2}, 0)$ and $(0, -\frac{1}{2})$.

In fig. 8 some results are shown. The contact area and several isobars are drawn at different stages of the computational process. The cross marks the maximum of the pressure distribution. Fig. 8a gives the result when 20 discretized points have a non-zero pressure. Four separated contact areas are formed. The four contact areas are not exactly congruent. The contact areas in the first and in the third quadrant are exactly congruent and so are the contact areas in the second and in the fourth quadrant. This is caused by the symmetry of the triangular network. The influence of this symmetry decreases when the total compressive force increases.

Fig. 8b gives the result when 60 discretized points have a non-zero pressure. Now there is one contact area. The origin still has no pressure, so the innermost line of fig. 8b is the boundary of the contact area. The form of this line shows the influence of the triangular network. Fig. 8c

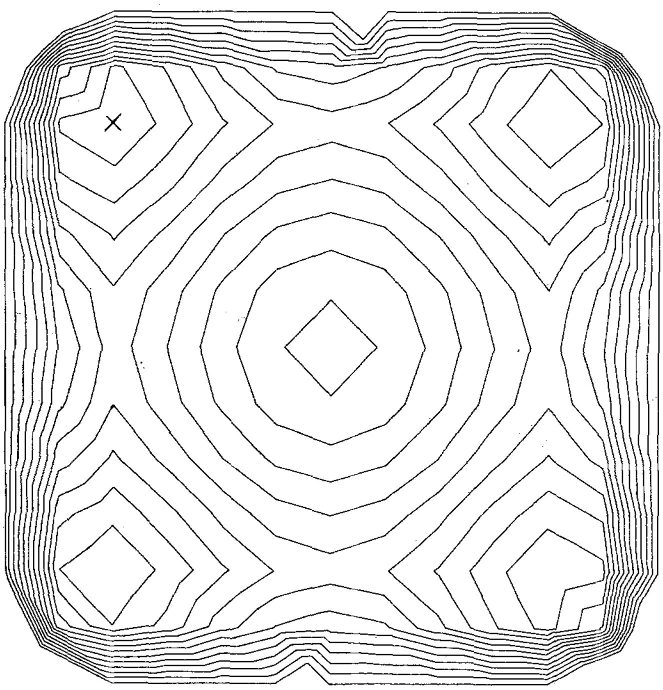

Figure 8d. Contact area with isobars $M=121$.

gives the result when 90 discretized points have a non-zero pressure. Here the origin also has a non-zero pressure. Fig. 8d gives the result after the final iteration, that is, when 121 discretized points have a non-zero pressure.

5. Conclusion

A variational principle governing the frictionless contact between two bodies has been established, which is valid both for linear and for non-linear elasticity. In the case of linear elasticity it appears to lead to an infinite dimensional convex quadratic programming problem. It is applied to the half-space geometry in linear elasticity, and it is established that non-Hertzian normal half-space contact problems are physically meaningful.

A Hertzian and a non-Hertzian normal contact problem were investigated numerically. In the case of the Hertzian problem, it was found that the exact relationships between penetration, maximum pressure, and total normal force are well satisfied. The form of the contact area is given only crudely, unless the mesh is considerably refined. This refinement is, however, only needed very close to the edge, since in that case passable results will be obtained.

REFERENCES

[1] L. A. Galin, *Contact problems in the theory of elasticity*, Moscow, 1953) (Russ.) Eng. translation: N. Carolina State College (1961).
[2] A. E. H. Love, *A treatise on the mathematical theory of elasticity*, 4th ed. Cambridge (1926).
[3] J. J. Kalker, *On the uniqueness of contact problems in the theory of elasticity*, can be obtained from the author. Publ. in Russ. in: Kontaktnoe vzaimodeistvie tverdikh tel, Isd "Nauka", Moscow (1971).
[4] Y. van Randen, *A numerical treatment of non-Hertzian frictionless contact problems*, (1970). Can be obtained from the author.
[5] J. J. Kalker and Y. van Randen, Non-Hertzian frictionless contact problems, *ZAMM* 51 (1971) T116–T117.
[6] P. Wolfe, The simplex method for quadratic programming, *Econometrica*, 27 (1959) 382–398.
[7] E. M. L. Beale, On quadratic programming, *Naval Research and Logistics Quarterly*, 6 (1959).
[8] J. J. Moreau, One-sided constraints in hydrodynamics, in: *Non-linear programming*, Ed. J. Abadie, North-Holland (1967) 261–279.
[9] T. F. Conry and A. Seireg, A mathematical programming method for design of elastic bodies in contact, *J. Appl. Mech.* (1971) 387–392.
[10] J. J. Kalker, Rolling with slip and spin in the presence of dry friction, *Wear* 9 (1966) 20–38.