# Variational Principles of Contact Elastostatics

J. J. KALKER

Department of Mathematics, Delft University of Technology, Delft 8, The Netherlands

[Received 5 December 1975 and in revised form 5 July 1976]

In this paper, the variational principles of contact elastostatics, which were proposed and proved by Fichera (1964) and Duvaut & Lions (1972) are developed in an engineering fashion without use of functional analysis. The theory contains a number of new elements, but the elegant existence proofs of the French-Italian school are missing. A start is made by extending the principle of virtual work to normal and frictional contact in such a manner that it needs no longer be known beforehand whether-in the case of normal contact-actual contact is or is not established, or-in the case of frictional contact-slip does or does not occur. Then the principle of minimal potential energy is set up for a non-linear elastic body in contact with a rigid base. Uniqueness and minimality of the solution are proved under certain conditions, the Reissner principle is established, and the principle of minimal complementary energy is derived. Finally the principles are cast in what is termed surface mechanical form, and two examples are given: the variational principle for normal half-space contact problems, and a new principle for time-dependent frictional half-space contact. Upon discretization, these principles provide a quadratic object function to be minimized under linear or quadratic inequality constraints. The positions of the contact area and of the regions of slip and adhesion appear as by-products of the calculation.

## 1. Introduction

ATTEMPTS to solve contact problems by the adaptation of standard analytical tools of classical mechanics often make it necessary to restrict consideration to two-dimensional or axially symmetric geometries. With these simplifications, it is impossible to cover the entire domain of contact mechanics, and a mathematical description is needed which is specific for contact problems. Such a description is to be found in the theory of variational inequalities developed in Italy and France in the last decade, by Signorini (1959), Fichera (1964), Stampacchia (1967), and Duvaut & Lions (1972). This school uses advanced functional analytic tools and lays emphasis on theoretical development: for instance Duvaut & Lions (1972) confine themselves to existence and uniqueness theorems.

By contrast, investigators in the Netherlands and the United States (Kalker, 1966, 1967, 1971*a*; Kalker & van Randen, 1972; Singh & Paul, 1974; Conry & Seireg, 1971; Engel, 1975; Johns & Leissa, 1975) turned to the variational formulation of contact problems with the purpose of arriving at a numerical treatment. Although such numerical methods are successful and compare with experiments and with solutions obtained by other means, the variational formulation is mostly of an ad-hoc character. A middle position is occupied by investigators in the Soviet Union and Hungary (Fridman & Chernina, 1967; Páczelt, 1974, 1977) who developed finite

element methods for the frictionless contact problem on the basis of the variational principles described in the present paper.

The aim of the present paper is to treat the variational principles of the French-Italian school in a manner in which functional analysis is avoided, while the proofs that will be given are of an "engineering" character, i.e. they are rigorous if the functions that occur are sufficiently often piecewise differentiable. The elegant existence proofs of the French-Italian school do not admit such a simplified treatment; but for this existence there is abundant proof in the sense of the engineer in view of the great number of contact problems that have been solved in the 96 years that have elapsed since Hertz initiated the theory of contact elastostatics.

#### Notations

The coordinate system is Cartesian with coordinates  $\xi_i$ , i = 1, 2, 3. Subscripts i, j, h, k range through these same values, except in Section 2. A tensor is denoted by its basic symbol; these basic symbols can be combined through addition and scalar multiplication. A tensor V may be a function of a variable W; notation: V(W). Most tensors V are functions of the variable W indicated in the Appendix; then V denotes V(W). The length of a vector V (= tensor of rank 1) is denoted by |V|. The components of a tensor V of rank 2, say, are given in index notation as  $V_{ij}$ . When two equal indices occur in an expression, summation over them is understood.

When it leads to clearer expressions, matrix-vector notation is used: in Section 6.1 matrices and vectors are partitioned, and in Section 7, Remark 2, rigid displacements are considered. Then, vectors are denoted by small, bold symbols, and matrices by bold capitals, which are mostly identical with the basic symbols of the corresponding tensors. In such a case,  $(V_i)$  is sometimes used for V.

When V is prescribed, the prescribed value is indicated by  $\tilde{V}$ . Differentiation with respect to  $x_k$  is denoted by  $_{,k}$  (see equation (17)); differentiation with respect to time is denoted by a dot over the symbol.

#### 2. Extension of the Principle of Virtual Work

In statics, a system may be characterized by fixing its position by means of the generalized coordinates  $\{q_i\}$  and the generalized forces  $\{Q_i\}$  associated with them. The virtual work  $\delta W$  is the work done by the generalized forces when the system undergoes a virtual displacement  $\{\delta q_i\}$ :

$$\delta W = Q_i \delta q_i, \tag{1}$$

 $\{\delta q_i\}$ : arbitrary virtual displacement; summation over repeated indices.

When the system is constrained, that is, if a relationship exists between the coordinates  $q_i$  or their variations  $\delta q_i$ , for instance

$$f_j(q_i) = 0 \Rightarrow \frac{\partial f_j}{\partial q_i} \delta q_i = 0$$
; or, more generally,  $a_{ji} \delta q_i = 0$ , (2)

a kinematically admissible virtual displacement is defined as a virtual displacement  $\{\delta q_i\}$  satisfying (2). Then, certain combinations of the generalized forces do no work; in particular, when (2) are the constraints, then these combinations are of the form

$$R_{ji} = \lambda a_{ji}, \quad \lambda: \text{ a multiplier.}$$
 (3)

The  $R_{ji}$  are called *reaction forces* connected with the *j*th constraint. These reaction forces may be left out of consideration in the virtual work (1); they do not influence it. The remaining forces are called *non-reaction forces*; they will, again, be denoted by  $Q_i$ . In many cases the non-reaction forces may be derived from a potential V, in the following manner,

$$V = V(q_i), Q_i = -\partial V / \partial q_i.$$
<sup>(4)</sup>

Under these circumstances we speak of monogenic forces.

The variations that are kinematically admissible (kin. adm.) in view of (2) have the property that

$$\{\delta q_i\}$$
 is kin. adm.  $\Rightarrow \{-\delta q_i\}$  is kin. adm. (5)

Constraints which possess the property (5) are called *bilateral*. If the constraints are so that they may be expressed as functions of the coordinates they are called holonomic. The principle of virtual work enunciates that in equilibrium the virtual work of the non-reaction forces vanishes for all bilateral kin. adm. virtual displacements

$$\delta W \equiv Q_i \delta q_i = 0, \{\delta q_i\} \text{ kin. adm.} \Leftrightarrow \text{equilibrium}$$
(6)

or, in other terms, that in equilibrium the non-reaction forces are expressible in the reactions. In the case when we have a monogenic system with holonomic constraints, (6) reduces to the demand that V is stationary under the equality constraints (2).

Apart from equality constraints such as (2) there also may be inequality constraints, such as

$$f_i(q_i) \ge 0. \tag{7}$$

Such an inequality constraint will be investigated with the aid of an example, for which we take a mass point m in a gravity field (see Fig. 1).



FIG. 1. A mass point in a gravity field.

Let the potential of the field be given by V = mgZ, and the inequality constraint is that  $Z \ge 0$ . It is known that equilibrium occurs when the mass point lies on the plane Z = 0. The contact force is acting upwards, but ceases to exist when contact is broken. Hence the contact force performs no work in a kinematically admissible variation of the coordinate Z, so that we will leave it out of the consideration. The virtual work becomes:

$$\delta W = -mg\delta Z; \forall \delta Z: \exists e_0 > 0: 0 < \varepsilon < e_0 \Rightarrow (Z + \varepsilon \delta Z) \ge 0;$$
  

$$\Leftrightarrow Z = 0 \Rightarrow \delta Z \ge 0; Z > 0 \Rightarrow \delta Z \text{ is arbitrary}$$
(8)  

$$\delta W \le 0 \text{ in equilibrium, since } Z = 0.$$

If, on the other hand, we start from the demand that  $\delta W \leq 0$  and  $Z \geq 0$ , we obtain

$$\delta W = -mg\delta Z \leq 0, Z \geq 0 \Rightarrow \exists \varepsilon_0 : (Z + \varepsilon \delta Z) \geq 0 \text{ if } 0 < \varepsilon < \varepsilon_0.$$
(9a)

(1) Let 
$$Z > 0$$
:  $\delta Z$  is arbitrary; take  $\delta Z < 0 \Rightarrow$  violation of (9a). (9b)

(2) Let 
$$Z = 0$$
:  $\delta Z \ge 0 \Rightarrow \delta W = -mg\delta Z \le 0$  since  $mg > 0$ : equilibrium. (9c)

In this case the equilibrium is described by

$$\delta W \equiv Q \delta Z \leq 0, Q: \text{ non-reaction force}; \quad Z \ge 0.$$
(10)

It is seen from (8), (9c) that the constraint  $Z \ge 0$  is not bilateral when Z = 0; it will be called unilateral. In general, conditions (7) are also unilateral. We arrive, if only from one simple case, at the conclusion that  $\delta W$  is non-positive for all kinematically admissible variations of the generalized coordinates. Similar considerations led Fourier (1971) to the conclusion that the principle of virtual work should be extended to the pronouncement that  $\delta W \le 0$  for all kinematically admissible virtual displacements. We observe that we saw in (9) that for bilateral constraints  $\delta W \le 0 \Rightarrow \delta W = 0$ ; indeed,

$$\delta W = Q_i \delta q_i \leq 0 \forall \text{ kin. adm. } \{\delta q_i\} \Rightarrow \{-\delta q_i\} \text{ kin. adm.} 
Q_i(-\delta q_i) \leq 0 \Rightarrow \delta W = Q_i \delta q_i \geq 0 \Rightarrow Q_i \delta q_i = \delta W = 0$$
(11)

For monogenic forces and holonomic constraints, the principle reads

Equilibrium  $\Leftrightarrow \delta V \ge 0$  subject to  $f_j = 0, j = 1(1)m; f_j \ge 0, j = m + 1(1)n$ . (12)



FIG. 2. An example involving friction.

The example given above was a frictionless contact problem. We will now apply the principle (11)-(12) to an elementary problem involving friction (see Fig. 2). It will appear as an example in which the virtual work itself is unilateral in character, i.e.

$$\delta W(\delta q_i) \neq -\delta W(-\delta q_i).$$

Consider a particle *m* subject to the inequality constraint  $Z \ge 0$ . An external force Q acts on the particle, with Z-component, -N, and tangential component T. Coulomb's friction law reads, in terms of the external force Q:

$$|T| \le \mu N, \mu$$
: coefficient of friction. (13a)

If the particle is moved on the plane Z = 0 over a small shift s, Coulomb states that the external force T must be given by

$$T_i = \mu N s_i / |s|. \tag{13b}$$

The work done by the force T is  $T_i s_i$ ; the work done by the generated frictional force (no reaction force!) is  $(-\mu N|s|)$ . The virtual work and the equilibrium condition

become

$$\delta W = -N\delta Z + T_i \delta s_i - \mu N\delta |s| \leq 0, \text{ subject to } (Z + \varepsilon \delta Z) \geq 0, 0 < \varepsilon < \varepsilon_0; \delta s$$
  
arbitrary. (14)

We see whether (14) indeed describes the equilibrium condition by calculating the external force corresponding to all possible configurations.

(1) 
$$Z > 0 \Rightarrow \delta Z$$
,  $\delta s$  arbitrary. Take  $\delta s = 0 \Rightarrow \delta |s| = 0 \Rightarrow N = 0$  (15a)  
 $\Rightarrow \delta W = T_t \delta s_t \leq 0$ ,  $\delta s$  arbitrary  $\Rightarrow T = 0$ .

That means that equilibrium in free space can only occur when Q = (T, -N) = 0.

(2) 
$$Z = 0 \Rightarrow \delta Z \ge 0$$
,  $\delta s$  is arbitrary. Take  $\delta s = 0 \Rightarrow \delta |s| = 0 \Rightarrow N \ge 0$ ; (15b)  
take  $\delta Z = 0 \Rightarrow \delta W = T_t \delta s_t - \mu N \delta |s| \le 0$ .

(a) Let s = 0. Then  $\delta |s| = |\delta s|$ ,  $\delta s$  arbitrary  $\Rightarrow$  since  $N \ge 0$ : $|T| \le \mu N$ , cf. (13a)

(b) Let  $s \neq 0$ . Then  $\delta |s| = s_i \delta s_i / |s|$ , and  $\delta W = \{T_i - \mu N s_i / |s|\} \delta s_i \leq 0$ , (15c)

 $\delta s$  arbitrary; W has now a bilateral character  $\Rightarrow T_i = \mu N s_i / |s|$ , cf. (13b).

Hence (13) follows from (14), and vice versa.

#### 3. The Principle of Virtual Work in Contact Elastostatics

We wish to find the analogue of the virtual work equilibrium conditions (14) for elastostatics. We consider an elastic body which in the undeformed state (particle x, Cartesian coordinates:  $x_i$ ) occupies the region G. Since the law of friction acts on the velocity or the shift, an incremental theory is indicated. To construct it, we start from an equilibrium reference state in which the particle x has the coordinates  $y_i$ ; y = y(x). This reference state is supposed to be the equilibrium state of the body at the time t, and is assumed to be known. Ordinarily the displacement (y-x) is small; but since little complication is met if we regard  $x \to y$  as a large deformation, we will assume the latter.

At the time  $(t+\tau)$  ( $\tau$ : an increment of time) the body will be in the deformed state in which the particle x has the coordinates  $(y_t+\eta_i)$ ;  $\eta = \eta(y)$ . The deformation  $y \to y+\eta$  is assumed to be small enough so that the small displacement gradient theory of linear elasticity is applicable. In particular this means that the inner normal N = N(y) on the body at the time t almost coincides with the normal  $N(y+\eta)$  at the time  $(t+\tau)$ .

The time increment  $\tau$  is assumed to be large with respect to a typical response time of the elastic body, but on the other hand small enough for the displacement with respect to the reference state to be linear in time:

At the time  $t + \theta \tau (0 \le \theta \le 1)$  the deformation is  $y \to y + \theta \eta$ . (16) This implies that inertia is neglected.

The body may be inhomogeneous and anisotropic, but its specific energy E must depend on the metric tensor in the deformed state only:

$$E = E([y_{i,j} + \eta_{i,j}][y_{i,k} + \eta_{i,k}], x); i, j, k:1, 2, 3; k = \partial/\partial x_k.$$
(17)

At the time  $(t+\tau)$  the displacement  $\eta$  is prescribed on the part of the surface of the body which in the undeformed state is given by  $S_u \subset \partial G$ . It is assumed that  $S_u$  has a non-zero area, which means that the body is properly supported:

$$\eta = \tilde{\eta} = \text{prescribed for } x \in S_u \subset \partial G; \text{ Area } (S_u) \neq 0.$$
 (18)

On another part  $S_p^* \subset \partial G$  the surface load  $(p+\pi) = (p(x) + \pi(x))$  is prescribed, which, owing to the fact that p, the surface load in the reference state, is known, amounts to

$$\pi_i = \tilde{\pi}_i = \text{prescribed on } S_p^* \subset \partial G. \tag{19}$$

On the remainder of the boundary  $C \subset \partial G$ , called the contact area, the body is in contact with a smooth surfaced rigid body which at the time  $(t+\tau)$  occupies the region

$$h(\xi, t+\tau) \leq 0, \xi_i:$$
coordinates;  $x \in C \Leftrightarrow h(y+\eta, t+\tau) = 0.$  (20)

The space derivatives of h are assumed to be of order unity; the time derivative is assumed to be small with respect to the velocity of sound in the elastic body. The contact area is not necessarily connected; it may depend on time, and is one of the unknowns of the problem. Near but outside the contact area the surface of the body will be assumed free of traction, which means that the region  $S_p^*$  completely surrounds the contact area.



FIG. 3. An elastic body in contact with a rigid base.

It was stated that the contact area is unknown. It may, however, be embedded in a larger region, called the potential contact area  $S_c$ , which consists of the contact area C and part of the region  $S_p^*$ . The remainder of  $S_p^*$  will be called  $S_p$ . This potential contact area is chosen *a priori* by the investigator; it is assumed to have the following properties necessary for contact:

(a) 
$$C \subset S_c \subset S_p^* \cup C \subset \partial G; \quad S_c \cup S_p \cup S_q = \partial G.$$
 (21a)

(b) y' is the point of the surface of the rigid body nearest to y (see Fig. 3). The distance y-y' is assumed to be O(η) when x ∈ S<sub>c</sub>. In addition, the inner normal N(y) on the particle x in the deformed and reference states is assumed to coincide almost with the outer normal on the rigid body at y':

$$N_i(y,t) \simeq N_i(y+\eta,t+\tau) \simeq h_{y_i}(y',t+\tau)/|h_y| \simeq h_{y_i}(y,t)/|h_y|,$$
  
with  $h_{y_i} \equiv \partial h/\partial y_i$ , so that  $h_y = |h_y|N.$ 

(c) The area  $S_c$  carries a contact load only.

In  $S_c$ , the conditions that govern the contact are:

 $h(y+\eta, t+\tau) \ge 0$  (no penetration condition) (22a)

(21c)

with 
$$h(y+\eta, t+\tau) = h(y, t+\tau) + \eta_i N_i |h_y|.$$
(22b)

In order to define the shift which occurs in the frictional part of the virtual work equation (14), we introduce the following notations. Let V be a vector defined at y or  $(y+\eta)$  on the surface of the deformed body, and let N be the inner normal on the body at y in the reference state. Then:

$$V_{N} \equiv \text{normal component of } V = V_{i}N_{i} (N: \text{ inner normal, } V_{N}: \text{ a scalar}) \\V_{T} \equiv \text{ tangential component of } V; V_{T_{i}} = V_{i} - V_{N}N_{i}$$

$$(23)$$

If  $\phi = \phi(y')$  is the displacement from t to  $(t+\tau)$  of the point y' (see (21b)) of the rigid surface, then

the shift of the particle x with respect to the rigid surface  $\equiv \eta_T - \phi_T$ . (24)

We also need an expression for the frictional traction. If  $(p+\pi)$  is the traction acting on the surface particle of x of the elastic body at  $S_c$ , then

traction bound = 
$$\mu(p_N + \pi_N) \equiv g; \mu$$
: coefficient of friction. (25)

The well-known principle of virtual work in elastostatics, which is conventionally written in terms of  $\delta V = -\delta W \ge 0$ , is modified by two elements, see (14):

(1) (22) is added as an auxiliary condition.

(2) A term

$$\int_{S_o} g\delta |\eta_T - \phi_T| \ dS$$

is added to the conventional variation  $\delta V$ .

This term is comparable to the term  $-\mu N\delta |s|$  of (14), with the sign reversed since we use  $\delta V = -\delta W$  rather than  $\delta W$  as was done in (14). We have

$$\delta V = -\delta W \equiv \delta \left\{ \int_{G} (E - X_{i} \eta_{i}) \, dG - \int_{S_{p}} (p_{i} + \tilde{\pi}_{i}) \eta_{i} \, dS \right\} + \int_{S_{o}} g \delta |\eta_{T} - \phi_{T}| \, dS,$$

$$(26)$$

X = X(y) is the body force, independent of time;  $\delta V \ge 0 \forall \delta \eta$  compatible with (27):

subject to 
$$\eta_i = \tilde{\eta}_i$$
 on  $S_u$ ;  
 $h(y+\eta, t+\tau) = h(y, t+\tau) + \eta_N |h_y| \ge 0$  in  $S_c$ .
$$\left. \right\} (27)$$

Since the  $\eta$  are small, the variational inequality (26) may be simplified by expanding the energy E in powers of  $\eta_{i,j}$ ; constant terms are dropped, and only linear and quadratic terms are retained. We define:

 $s_{jl}z = \partial E(y_{h,k})/\partial y_{i,j}$ : the Piola stress tensor in the reference state;  $s_{jl} + \sigma_{jl} = \partial E(y_{h,k} + \eta_{h,k})/\partial y_{l,j}$ : the Piola stress tensor in the deformed state;  $\sigma_{jl}$ : increment of the Piola stress; (28a)

$$E_{ijhk} = \partial^2 E(y_{i,m}) / \partial y_{i,j} \partial y_{h,k}; \sigma_{ji} = E_{ijhk} \eta_{h,k}.$$
(28b)

When the reference state is so near the undeformed state that linear elasticity is valid, then  $s_{jl}$  reduces to the classical stress and  $E_{ljhk}$  to the elastic constants. The elastic

energy becomes, apart from a term depending on y only:

$$E = s_{jl}\eta_{l,j} + \frac{1}{2}E_{ljhk}\eta_{l,j}\eta_{h,k}.$$
 (29)

We introduce (29) into (26):

$$\delta V = \delta \left\{ \int_{G} \left[ s_{ji} \eta_{i,j} + \frac{1}{2} E_{ijhk} \eta_{i,j} \eta_{h,k} - X_{i} \eta_{i} \right] dG - \int_{S_{r}} (p_{i} + \tilde{\pi}_{i}) \eta_{i} dS \right\} + \int_{S_{c}} g\delta |\eta_{T} - \phi_{T}| dS \ge 0$$
subject to  $\eta_{i} = \tilde{\eta}_{i}$  on  $S_{r}$ ;  $h(v, t+\tau) + \eta_{v} |h_{i}| \ge 0$  on  $S_{r}$ .
$$(30)$$

subject to  $\eta_i = \tilde{\eta}_i$  on  $S_{\mu}$ ;  $h(y, t+\tau) + \eta_N |h_y| \ge 0$  on  $S_c$ .

The variation is performed with the proviso that  $\delta \eta_i = 0$  on  $\delta G$ . All such  $\delta \eta$  are variations compatible with (27), and the Euler-Lagrange equations are

> $s_{ii,i} + \sigma_{ii,i} + X_i = 0$  equilibrium equations, time:  $t + \tau$ . (31a)

Now the reference state was assumed to be in equilibrium. Hence

$$s_{jl,j} + X_l = 0$$
 equilibrium, time: t. (31b)

Since X is independent of time, see (26), the  $X_i$  in (31a, b) denote one and the same quantity. Hence

$$\sigma_{jl,j} = 0. \tag{31c}$$

J

We identify

$$p_i = s_{ji}n_j, \quad \pi_i = \sigma_{ji}n_j, \quad n: \text{ outer normal on } \partial G \text{ at } x$$
 (32)

with the surface load in the reference state at the particle x, and the increment of this load respectively. The loads and stresses are measured with respect to surfaces in the undeformed state.

The term  $s_{il}\eta_{l,i}$  in (30) is partially integrated. We find, if (31, 32) are taken into account, and a constant term

$$\int_{S_u} p_i \tilde{\eta}_i \, dS$$

is dropped:

$$\delta V \equiv \delta \left\{ \int_{G} \frac{1}{2} E_{ijkk} \eta_{i,j} \eta_{k,k} \, dG - \int_{S_{p}} \tilde{\pi}_{i} \eta_{i} \, dS \right\} + \int_{S_{o}} \left\{ p_{i} \delta \eta_{i} + g \delta |\eta_{T} - \phi_{T}| \right\} \, dS,$$

$$(33)$$

subject to (27);  $\delta V \ge 0 \forall \delta \eta$  compatible with (27).

We recall that g was defined in (25) as  $\mu(p_N + \pi_N)$ , where  $\mu$  is the (position dependent) coefficient of friction. However, up to now it has not been proved possible to establish existence and uniqueness of the solution of (33) when g depends on the deformation  $\eta$ . When q depends on position and time only, existence and uniqueness have been established by Fichera (1964) for the case that  $g = \mu = 0$  (the "Signorini Problem", also called "The normal contact problem"), and for g > 0 by Duvaut & Lions (1972),

who, in addition, confine themselves to the case when the non-penetration condition  $h \ge 0$  is replaced by the condition that  $(p_N + \pi_N)$  is prescribed in S<sub>c</sub>; this leads to a variational principle which is given below in (41). Both proofs are restricted to coercivity of the energy, see (39), which holds in the case of classical elasticity. In addition. Duvaut & Lions gave a proof for the case of elastodynamics and viscoelasticity. Regarding uniqueness, an elementary proof, based on Kirchhoff's uniqueness theorem for classical elasticity has been given by Kalker (1971b). In the present paper, existence will not be proved; a uniqueness proof for the case that  $g = g(x, t+\tau)$ will be given in the next section. Simultaneously it will be shown that the potential energy V (which then exists) attains its unique minimum at the solution  $\eta$ .

#### 4. The Uniqueness and Minimality Property of the Solution

In order to prove uniqueness it is assumed that the traction bound q is a function of position and time only:

$$g = g(x, t+\tau). \tag{34}$$

Then the variational inequality (33) can be written as

$$V = \int_{G} \frac{1}{2} E_{ijkk} \eta_{i,j} \eta_{k,k} \, dG - \int_{S_{p}} \tilde{\pi}_{i} \eta_{i} \, dS + \int_{S_{o}} \{ p_{i} \eta_{i} + g | \eta_{T} - \phi_{T} | \} \, dS$$

$$V: \text{ potential energy; subject to (27); } \delta V \ge 0 \, \forall \delta \eta \text{ compatible with (27).}$$

$$(35)$$

V: potential energy; subject to (27);  $\delta V \ge 0 \forall \delta \eta$  compatible with (27).

Let  $\eta$  be a solution of this problem and let  $\eta'$  satisfy the auxiliary conditions (27). V' is the potential energy connected with  $\eta'$ . Then

$$V' = \int_{G} \frac{1}{2} E_{ijhk} \eta'_{i,j} \eta'_{h,k} \, dG - \int_{S_{p}} \tilde{\pi}_{i} \eta'_{i} \, dS + \int_{S_{o}} \{p_{i} \eta'_{i} + g | \eta'_{T} - \phi_{T} | \} \, dS \qquad (36a)$$

$$= \int_{G} \frac{1}{2} E_{ijhk} \eta_{i,j} \eta_{h,k} \, dG - \int_{S_{p}} \tilde{\pi}_{i} \eta_{i} \, dS + \int_{S_{o}} \{p_{i} \eta_{i} + g | \eta_{T} - \phi_{T} | \} \, dS + \dots V \qquad \int_{G} E_{ijhk} \eta_{h,k} (\eta'_{i,j} - \eta_{i,j}) \, dG - \int_{S_{p}} \tilde{\pi}_{i} (\eta'_{i} - \eta_{i}) \, dS + \qquad \dots \ge \delta V \ge 0 \qquad (36b)$$

$$\int_{S_{o}} \{p_{i} (\eta'_{i} - \eta_{i}) + g[|\eta'_{T} - \phi_{T}| - |\eta_{T} - \phi_{T}|]\} \, dS + \qquad \dots \ge 0 \qquad (36b)$$

The first line of (36b) we identify with V. Regarding the second and third lines we observe that  $(\eta' - \eta)$  is an admissible variation of  $\eta$  except possibly on  $S_{\mu}$  and  $S_{c}$ . But on  $S_{\mu}$  we have

$$\eta' = \tilde{\eta} = \eta \quad \text{on} \quad S_u \Rightarrow \eta' - \eta = 0;$$
 (37a)

further we have on  $S_c$  that

$$h(y+\eta', t+\tau) \ge 0, \quad h(y+\eta, t+\tau) \ge 0$$

so that by (22b)

$$\begin{array}{l} h(y+\eta+\varepsilon(\eta'-\eta), t+\tau) = h(y, t+\tau) + (1-\varepsilon)\eta_N |h_y| + \varepsilon \eta'_N |h_y| \\ = \varepsilon h(y+\eta', t+\tau) + (1-\varepsilon)h(y+\eta, t+\tau) \ge 0 \text{ if } 0 \le \varepsilon \le 1. \end{array}$$

$$\begin{array}{l} (37b) \end{array}$$

(37a, b) establish that  $(\eta' - \eta)$  is a feasible variation of  $\eta$  when (27) is taken into account. Finally,

$$\delta |\eta_{T} - \phi_{T}| = \lim_{\varepsilon \downarrow 0} \{ |\eta_{T} + \varepsilon(\eta_{T}' - \eta_{T}) - \phi_{T}| - |\eta_{T} - \phi_{T}| \} / \varepsilon$$

$$= \lim_{\varepsilon \downarrow 0} \{ |(1 - \varepsilon)(\eta_{T} - \phi_{T}) + \varepsilon(\eta_{T}' - \phi_{T})| - |\eta_{T} - \phi_{T}| \} / \varepsilon$$

$$\leq |\eta_{T}' - \phi_{T}| - |\eta_{T} - \phi_{T}|$$

$$(37c)$$

while  $g \ge 0$ , so that the second and third lines of (36b) are  $\ge \delta V \ge 0$ . In the classical theory of elasticity,  $(E_{ijkk}\gamma_{ij}\gamma_{kk})$  ( $\gamma$ : linearized strain) is positive definite, so that the fourth line of (36b) is non-negative. Hence

 $V' \ge V$ ; V is the (minimal) potential energy attained at  $\eta$ . (38) Now, if  $\eta'$  is likewise a solution, we have that V' = V, and hence, by the above,

 $E_{ijhk}(\eta'_{i,j} - \eta_{i,j})(\eta'_{h,k} - \eta_{h,k}) = E_{ijhk}\gamma_{ij}\gamma_{hk} = 0,$ with  $\gamma$ : linearized strain associated with  $(\eta' - \eta)$ . (39)

It follows from the positive definiteness of  $(E_{ijkk}\gamma_{ij}\gamma_{kk})$  that  $\gamma = 0$ , and then from the fact that  $\eta' - \eta = 0$  on  $S_u$ , which has a non-vanishing area, that  $\eta' = \eta$  throughout the body, which establishes uniqueness.

It is noted that the assumption of the validity of classical elasticity has been used for its property of coercivity only, i.e. the property

$$E_{ijkk}(\eta'_{i,j} - \eta_{i,j})(\eta'_{h,k} - \eta_{h,k}) \ge a\gamma_{ij}\gamma_{ij}, a > 0, \text{ constant}$$

 $\gamma =$  linearized strain associated with  $(\eta' - \eta)$ .

If this property is assumed also in the non-classical case, the uniqueness-minimality proof given above continues to hold.

Summarizing, we have found that if the traction bound g is a function of position and time only, see (34), then there exists a potential energy V, defined by (35), which upon minimization, yields the unique displacement  $\eta$ , solution of the problem, if, at any rate the coercivity relation (39) is satisfied.

## 5. Discussion of the Restriction (34) on the Traction Bound

The demand that g is a function of position and time only does not, in many cases, hinder the calculation of contact problems by variational means. Such a situation occurs in the classical theory of elasticity when an incompressible elastic half-space is in contact with a rigid punch, and also when two bodies with equal elastic constants, which, geometrically, are each other's mirror image with respect to a plane through the contact area, are in contact with each other. In this connection it is observed that although in classical elasticity the variational description of two bodies in contact is more complicated than the one which we consider, this complication is not essential.

In the two cases mentioned it may be shown (see, e.g., de Pater, 1962, p. 33) that the normal pressure  $(p_N + \pi_N)$  is not affected by the value of g. Hence  $\pi_N$  may be calculated with g = 0 (the normal contact problem); the variational principle is:

min! 
$$V_N \equiv \int_G \frac{1}{2} E_{ijkk} \eta_{i,j} \eta_{h,k} \, dG - \int_{S_p} \tilde{\pi}_i \eta_i \, dS + \int_{S_o} p_i \eta_i \, dS$$
, subject to (27). (40)

Next, the value of g is set equal to  $g = \eta(p_N + \pi_N)$ , and the final field is calculated with the aid of the variational principle of Duvaut & Lions (1972) in which the no-penetration condition (22) is replaced by prescribing in  $S_c \pi_N = \tilde{\pi}_N$ , the increment of the normal pressure, calculated from (40):

$$\min ! V_{T} \equiv \int_{G} \frac{1}{2} E_{ijhk} \eta_{i,j} \eta_{h,k} \, dG - \int_{S_{p}} \tilde{\pi}_{i} \eta_{i} \, dS + \int_{S_{a}} \{p_{T_{i}} \eta_{T_{i}} - \tilde{\pi}_{N} \eta_{N} + g | \eta_{T} - \phi_{T} | \} \, dS$$

$$(41)$$

$$(41)$$

subject to  $\eta_i = \tilde{\eta}_i$  in  $S_u$ ,  $g = \mu(p_N + \tilde{\pi}_N)$ .

In the case of asymmetric half-space contact the restriction (34) on g is not very restrictive. The normal contact problem, in which  $\mu = q = 0$ , is calculated first from (40) or otherwise; from that result an approximation of the traction bound q is taken. Then, two courses are open to the investigator. Either he calculates an approximation of the final field by minimizing  $V_T$ , (see 41); this is Johnson's approximation (1962), which is sufficiently accurate when the coefficient of friction is small enough (say  $\mu = 0.3$  in steel on aluminium contact). Or the problem (35) is solved with  $g = \mu(p_N + \tilde{\pi}_N)$ , where  $\tilde{\pi}_N$  is the value of  $\pi_N$  just calculated. Then a new value of  $\pi_N$ is found, which is used to correct g, etc. The question of the convergence of this process, however, is open. It would seem that in half-space contact, which yields already a good approximation of the contact elastic field when the contact is almost flat and the diameter of the contact area is less than one fourth of the diameter of the body (e.g. wheel-rail contact), convergence is assured when, say,  $\mu < 1$ , but there is no direct evidence to support this. In many technological applications the symmetry mentioned in the first paragraph of this section is present, as in the contact between a steel rail and a steel wheel (half-space contact, which implies the geometric symmetry, and steel-on-steel, which implies the material symmetry) or when a massive rubber body is in contact with a concrete foundation (again half-space contact, while rubber is almost incompressible, and the foundation is relatively rigid).

It is concluded that for most technological applications the restriction (34) on g leads to acceptable calculating schemes, but it is of great theoretical significance to clarify the general case.

# 6. Dual Theorems: the Variational Principles of Reissner and of Minimum Complementary Energy

We wish to dualize the problem (35). This is done according to the pattern set by Noble & Sewell (1972) and Arthurs (1970). A difficulty arises when the variables  $\eta$  must be eliminated from the form  $E_{ijkk}\eta_{i,j}\eta_{k,k}$  with the aid of the dual variables

 $\sigma_{jl} = E_{ijkk}\eta_{k,k}$ , owing to the fact that the  $(9 \times 9)$  matrix  $(E_{(ij)(kk)})$  may be singular. This elimination is studied in Section 6.1. In 6.2, Reissner's principle is proposed and established through the identity of its equivalent local equations with those of (35). Also, a mechanical interpretation is given to these local equations. Finally, in 6.3, the principle of minimum complementary energy is established.

6.1. The Dualization of a Quadratic Form

We define:

$$\mathbf{d} = \mathbf{E}\mathbf{e}, \text{ with } \mathbf{E} = \begin{pmatrix} \mathbf{B} & \mathbf{E}_{12} \\ \mathbf{E}_{12}^T & \mathbf{E}_{22} \end{pmatrix},$$

a symmetric matrix. We assume that the last columns, transposed  $(E_{12}^T E_{22})$  depend upon the first, and that **B** is a regular symmetric matrix. Owing to this linear dependence there exists a matrix **A**:

$$\mathbf{A}: \begin{pmatrix} \mathbf{E}_{12} \\ \mathbf{E}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{B} \\ \mathbf{E}_{12}^T \end{pmatrix} \mathbf{A} \Rightarrow \begin{cases} \mathbf{E}_{12} = \mathbf{B}\mathbf{A} \Rightarrow \mathbf{A} = \mathbf{B}^{-1}\mathbf{E}_{12} \\ \mathbf{E}_{22} = \mathbf{E}_{12}^T\mathbf{A} = \mathbf{A}^T\mathbf{B}\mathbf{A}. \end{cases}$$
(42)

Hence the matrix E becomes

$$\mathbf{E} = \begin{pmatrix} \mathbf{B} & \mathbf{B}\mathbf{A} \\ \mathbf{A}^{T}\mathbf{B} & \mathbf{A}^{T}\mathbf{B}\mathbf{A} \end{pmatrix} = \begin{pmatrix} \mathbf{I} \\ \mathbf{A}^{T} \end{pmatrix} \mathbf{B} \ (\mathbf{I}, \mathbf{A}). \tag{43}$$

We wish to express the quadratic form  $e^{T}Ee$  as a quadratic form involving the **d** variables (dualization); we will show that if

$$\mathbf{C} \equiv \begin{pmatrix} \mathbf{B}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{I} \\ \mathbf{0} \end{pmatrix} \mathbf{B}^{-1} \ (\mathbf{I}, \mathbf{0}) \Rightarrow \mathbf{d}^{T} \mathbf{C} \mathbf{d} = (\mathbf{d}_{I}^{T}, \mathbf{d}_{A}^{T}) \mathbf{C} \begin{pmatrix} \mathbf{d}_{I} \\ \mathbf{d}_{A} \end{pmatrix} = \mathbf{d}_{I}^{T} \mathbf{B}^{-1} \mathbf{d}_{I}$$
$$= \mathbf{e}^{T} \mathbf{E} \mathbf{e}.$$
(44)

Indeed,

$$\mathbf{d}^{T}\mathbf{C}\mathbf{d} = \mathbf{e}^{T}\mathbf{E}\mathbf{C}\mathbf{E}\mathbf{e} = \mathbf{e}^{T}\begin{pmatrix}\mathbf{I}\\\mathbf{A}^{T}\end{pmatrix}\mathbf{B}(\mathbf{I}, \mathbf{A})\begin{pmatrix}\mathbf{I}\\\mathbf{0}\end{pmatrix}\mathbf{B}^{-1}(\mathbf{I}, \mathbf{0})\begin{pmatrix}\mathbf{I}\\\mathbf{A}^{T}\end{pmatrix}\mathbf{B}(\mathbf{I}, \mathbf{A})\mathbf{e}$$
$$= \mathbf{e}^{T}\begin{pmatrix}\mathbf{I}\\\mathbf{A}^{T}\end{pmatrix}\mathbf{B}\mathbf{B}^{-1}\mathbf{B}(\mathbf{I}, \mathbf{A})\mathbf{e} = \mathbf{e}^{T}\mathbf{E}\mathbf{e}.$$

Now, the d's are no longer arbitrary, since the system d = Ee may not be contradictory. We have:

$$\mathbf{d} = \begin{pmatrix} \mathbf{d}_I \\ \mathbf{d}_A \end{pmatrix} = \begin{pmatrix} \mathbf{I} \\ \mathbf{A}^T \end{pmatrix} \mathbf{B}(\mathbf{I}, \mathbf{A}) \mathbf{e} = \begin{pmatrix} \mathbf{I} \\ \mathbf{A}^T \end{pmatrix} \mathbf{B} \mathbf{v} = \begin{pmatrix} \mathbf{d}_I \\ \mathbf{A}^T \mathbf{d}_I \end{pmatrix} \Rightarrow \mathbf{d}_A = \mathbf{A}^T \mathbf{d}_I = \mathbf{E}_{12}^T \mathbf{B}^{-1} \mathbf{d}_I.$$
(45)

If the components of e are arbitrary, then also the components of v are arbitrary which means that one can freely choose the components of  $d_I = Bv$ , owing to the regularity of B. Then  $d_A$  is completely fixed, according to (45); that is, (45) is equivalent to the fact that E is singular. Hence

$$\mathbf{e}^{T}\mathbf{E}\mathbf{e} = \mathbf{d}^{T}\mathbf{C}\mathbf{d} \text{ subject to } \mathbf{d}_{A} = \mathbf{E}_{12}^{T}\mathbf{B}^{-1}\mathbf{d}_{I} = \mathbf{A}^{T}\mathbf{d}_{I}$$
$$\mathbf{e}^{T}\mathbf{E}\mathbf{e} = \mathbf{d}_{I}^{T}\mathbf{B}^{-1}\mathbf{d}_{I}.$$
(46)

Example-Classical two-dimensional elasticity.

$$E_{ijhk} = E_{jlhk} = E_{hkij}.$$

$$E = \begin{pmatrix} E_{11 \ 11} & E_{11 \ 12} & E_{11 \ 22} & E_{11 \ 21} \\ E_{12 \ 11} & E_{12 \ 12} & E_{12 \ 22} & E_{12 \ 21} \\ E_{22 \ 11} & E_{22 \ 12} & E_{22 \ 22} & E_{22 \ 21} \\ E_{21 \ 11} & E_{21 \ 12} & E_{21 \ 22} & E_{21 \ 21} \end{pmatrix}; B = \begin{pmatrix} E_{11 \ 11} & E_{11 \ 12} & E_{11 \ 22} \\ E_{11 \ 12} & E_{12 \ 22} & E_{12 \ 22} \\ E_{11 \ 22} & E_{12 \ 22} & E_{22 \ 22} \end{pmatrix};$$

$$E_{12} = \begin{pmatrix} E_{11 \ 11} & E_{11 \ 12} & E_{12 \ 22} \\ E_{11 \ 22} & E_{12 \ 22} & E_{22 \ 22} \\ E_{12 \ 22} & E_{22 \ 22} \end{pmatrix}; E_{22 \ 22}$$

The matrix  $E_{12}$  equals the second column of **B**, hence

$$\mathbf{A}^{T} = (0, 1, 0); \quad \mathbf{E}_{22} = \mathbf{A}^{T} \mathbf{B} \mathbf{A} = (0, 1, 0) \mathbf{B} \begin{bmatrix} 0\\ 1\\ 0 \end{bmatrix} = (E_{12 \ 12}).$$

If  $\sigma_{ij} = E_{ijhk}\eta_{h,k}$ , and  $\sigma_I^T = (\sigma_{11}\sigma_{12}\sigma_{22})$ ,  $\sigma_A = (\sigma_{21})$ , we have that  $E_{ijhk}\eta_{h,k}\eta_{i,j} = \sigma_I^T \mathbf{B}^{-1}\sigma_I$ , subject to  $(\sigma_{21}) \equiv \sigma_A = \mathbf{A}^T \sigma_I = (\sigma_{12})$ , the familiar symmetry of the stress tensor.

## 6.2. Reissner's Variational Principle. The Equivalent Local Equations

The specific complementary energy is introduced:

$$H = \sigma_{ji}\eta_{i,j} - \frac{1}{2}E_{ijkk}\eta_{i,j}\eta_{k,k} = \frac{1}{2}S_{ijkk}(x)\sigma_{j,i}\sigma_{k,k},$$

$$Z\sigma_{ji} = E_{ijkk}\eta_{k,k}, \text{ the Piola incremental stress tensor;}$$

$$(47)$$

 $S_{ijkk}$  is determined with the aid of the theory of Section 6.1. The  $\sigma_{ji}$  are subject to the relations

$$\boldsymbol{\sigma}_{A} = \mathbf{A}^{T} \boldsymbol{\sigma}_{I}, \text{ see (46).}$$

1 . 5

It will be established that the problem (35) is described by the Reissner principle

$$\delta_{\eta} R \leq 0, \forall \, \delta_{\eta',\sigma} R \geq 0, \forall \, \delta\lambda', \, \delta\sigma \text{ compatible with the equations (51)} \\ \delta_{\cdot}: \text{ variation with respect to the variable (.)}$$

$$(49)$$

$$R = \int_{G} \left[ -\sigma_{ji} \eta_{i,j} + \frac{1}{2} S_{ijkk} \sigma_{ji} \sigma_{kh} \right] dG + \int_{S_p} \tilde{\pi}_i \eta_i \, dS + \int_{S_u} \lambda'_i (\eta_i - \tilde{\eta}_i) \, dS + \int_{S_u} \left[ \lambda'_4 h(y+\eta, t+\tau) - p_i \eta_i - g |\eta_T - \phi_T| \right] dS; \quad g: \text{see } (34);$$

$$(50)$$

subject to

$$\boldsymbol{\sigma}_{A} = \mathbf{A}^{T} \boldsymbol{\sigma}_{I}^{\mathbf{I}}, \qquad \lambda_{4}^{\prime} \ge 0.$$
 (51)

This is done by identifying the local equations equivalent to (49, 50, 51) with those equivalent to (35).

The equations equivalent to  $\delta_{\lambda'} R \ge 0$  are:

$$\eta_i = \tilde{\eta}_i \text{ in } S_u \tag{52a}$$

$$h(y+\eta, t+\tau) = 0 \text{ if } \lambda'_4 > 0 \Big\} \underset{\longleftrightarrow}{\leftarrow} \Big\{ h(y+\eta, t+\tau) \ge 0$$
 (52b)

$$h(y+\eta, t+\tau) \ge 0 \text{ if } \lambda'_4 = 0 \int \left( \lambda'_4 h(y+\eta, t+\tau) = 0 \right), \quad (52c)$$

(52a, b) coincide with (27), the auxiliary conditions of problem (35). As to the equations equivalent to  $\delta_{\sigma}R \ge 0$  we introduce the notation of 6.1, that is

$$(\eta_{i,j}) = (\mathbf{e}_I^T, \mathbf{e}_A^T), (\sigma_{ji}) = (\boldsymbol{\sigma}_I^T, \boldsymbol{\sigma}_A^T); \boldsymbol{\sigma}_A = \mathbf{A}^T \boldsymbol{\sigma}_I.$$
(53)

The variation of R becomes

$$\delta_{\sigma}R = \delta_{\sigma_{I}} \int_{G} \left[ -\sigma_{ji}\eta_{i,j} + \frac{1}{2}S_{ijkk}\sigma_{ji}\sigma_{kk} \right] dG$$
  
$$= \delta_{\sigma_{I}} \int_{G} \left[ -\sigma_{I}^{T}\mathbf{e}_{I} - \sigma_{I}^{T}\mathbf{A}\mathbf{e}_{A} + \frac{1}{2}\sigma_{I}^{T}\mathbf{B}^{-1}\sigma_{I} \right] dG \ge 0, \forall \delta\sigma_{I}$$
  
$$\Rightarrow \mathbf{e}_{I} + \mathbf{A}\mathbf{e}_{A} = \mathbf{B}^{-1}\sigma_{I} \Rightarrow \sigma_{I} = \mathbf{B}(\mathbf{e}_{I} + \mathbf{A}\mathbf{e}_{A}) = \mathbf{B}(\mathbf{I}, \mathbf{A}) \begin{pmatrix} \mathbf{e}_{I} \\ \mathbf{e}_{A} \end{pmatrix}$$
  
$$(\sigma_{ji}) = \begin{pmatrix} \sigma_{I} \\ \sigma_{A} \end{pmatrix} = \begin{pmatrix} \mathbf{I} \\ \mathbf{A}^{T} \end{pmatrix} \mathbf{B}(\mathbf{I}, \mathbf{A}) \begin{pmatrix} \mathbf{e}_{I} \\ \mathbf{e}_{A} \end{pmatrix} = \mathbf{E}(\eta_{i,j})$$

or, in index notation,

 $\sigma_{ji} = E_{ijkk}\eta_{k,k}$ : the stress-strain relations (28b) follow from  $\delta_{\sigma}R \leq 0.$  (54) Hence, by (44), when  $\delta_{\sigma}R \geq 0$  subject to (51),

$$-\sigma_{ji}\eta_{i,j} + \frac{1}{2}S_{ijhk}\sigma_{ji}\sigma_{kh} = -E_{ijhk}\eta_{i,j}\eta_{h,k} + \frac{1}{2}E_{ijhk}\eta_{i,j}\eta_{h,k} = -\frac{1}{2}E_{ijhk}\eta_{i,j}\eta_{h,k}.$$

$$(55)$$

The equations equivalent to  $\delta R_{q} \leq 0$  are, as regards G,  $S_{p}$ ,  $S_{u}$ :

$$G: \quad \sigma_{jl,j} = 0 \quad (cf. (31)) \tag{56a}$$
  
$$S_{a}: \quad \pi_{i} = \tilde{\pi}_{i} = \text{prescribed} \tag{56b}$$

$$S_{\mu}: \quad \pi_{i} = \lambda_{i}' = \text{free.}$$
(56c)

Regarding  $S_c$  we have, if use is made of the representation (22b) of h:

$$\{-\pi_{l}+\lambda_{4}'N_{l}|h_{r}|-p_{l}\}\delta\eta_{l}-g\delta|\eta_{T}-\phi_{T}| \leq 0$$

If we set  $\delta \eta_T = 0$ , then  $\delta |\eta_T - \phi_T| = 0$ ,  $\delta \eta_i = N_i \delta \eta_N$ , with  $\delta \eta_N$  free. Then

$$-\pi_i N_i + \lambda'_4 |h_p| - p_i N_i = 0 \Rightarrow$$

 $S_c: p_N + \pi_N = \lambda'_4 |h_y| \ge 0, \text{ (by (51))}; \ \lambda'_4 h = 0 \ (\equiv (52c)) \Leftrightarrow (p_N + \pi_N)h = 0.$ (56d) We are then left with

$$\begin{aligned} (\pi_{T_i} + p_{T_i})\delta\eta_{T_i} + g\delta|\eta_T - \phi_T| &\ge 0. \\ \text{If } (\eta_T - \phi_T) \neq 0 \Rightarrow \delta|\eta_T - \phi_T| &= (\eta_{T_i} - \phi_{T_i})\delta\eta_{T_i}/|\eta_T - \phi_T| \forall \delta\eta_T \\ \Rightarrow (\pi_{T_i} + p_{T_i}) = -g(\eta_{T_i} - \phi_{T_i})/|\eta_T - \phi_T|, \end{aligned}$$

which means that  $(p_T + \pi_T)$ , the shear traction, has magnitude g and direction opposite to the slip  $(\eta_T - \phi_T)$ .

If 
$$(\eta_T - \phi_T) = 0 \Rightarrow \delta |\eta_T - \phi_T| = |\delta(\eta_T - \phi_T)| = |\delta\eta_T|$$
, so that  
 $(\pi_{T_1} + p_{T_1})\delta\eta_{T_1} + g|\delta\eta_T| \ge 0 \Rightarrow |\pi_T + p_T| \le g$ ,

which latter inequality is valid throughout  $S_c$ . Equivalent to the above is

$$S_c: |p_T + \pi_T| \le g, \ (p_{T_i} + \pi_{T_i})(\eta_{T_i} - \phi_{T_i}) + g|\eta_T - \phi_T| = 0.$$
(56e)

We will now show that the local equations (52, 54, 56) which are equivalent to (49, 50, 51) are identical with the local equations equivalent to (35).

*Proof.* (1) (52a, b) are identical with (27), the auxiliary conditions in (35); (52c) is discussed in point (5).

(2) (54) is identical with (28) as far as  $\sigma$  is concerned, while s enters neither in (49, 50, 51), nor in (35).

(3) (56a, b) also follow from (35).

(4) (56c) is the identification of a quantity not occurring in (35), and constitutes no restriction.

(5) (56d), which includes (52c), also follows from (35), as we will show now.

According to (35),

$$\begin{array}{c} (\pi_i + p_i)\delta\eta_i + g\delta|\eta_T - \phi_T| \ge 0 \text{ in } S_c \\ \text{subject to } h(y + \eta, t + \tau) = h(y, t + \tau) + \eta_N |h_y| \ge 0. \end{array} \right\}$$
(57)

We set  $\delta \eta_T = 0$ , so that  $\delta \eta_i = N_i \delta \eta_N$ , and we obtain

$$(\pi_N + p_N)\delta\eta_N \ge 0$$
 subject to  $h(y+\eta, t+\tau) = h(y, t+\tau) + \eta_N |h_y| \ge 0$ .

If  $h(y+\eta, t+\tau) > 0 \Rightarrow \delta\eta_N$  is free, and  $\pi_N + p_N = 0$ .

If  $h(y+\eta, t+\tau) = 0 \Rightarrow \delta \eta_N \ge 0$ , and  $\pi_N + p_N \ge 0$ 

 $\Rightarrow (\pi_N + p_N) \ge 0, (\pi_N + p_N)h = 0$ , i.e. (56d).

(6) (56e) follows from (57) if (56d) is taken account in the same way as (56e) was established above. Q.E.D.

Hence (35) and (49, 50, 51) are equivalent. Finally we have by (52c), (52a), and (55) that

> -V = R at the common solution of (35) and (49, 50, 51). (58)

It is of interest to state the equivalent equations in mechanical terms, from which it is seen that the local equivalent equations coincide with those used in classical contact elastostatics:

((52a): displacement prescribed on  $S_{\mu}$ ;

 $\begin{cases} (52b): \text{ no-penetration condition;} \\ (52c) \equiv \text{ the second equation (56d);} \end{cases}$  $\delta_{1'}$ 

δ. { (54): stress-strain relations for the incremental field;

- (56a): equations of equilibrium;
- (56b): surface load prescribed on  $S_p$ ;
- (56c): where the surface displacement is prescribed, the surface load is free;
- δ, (56d): in contact, the normal component of the load is compressive; (56e): Coulomb's friction law if  $g = \mu(p_N + \pi_N)$ ;

(56d, e): outside contact,  $S_c$  is free of traction.

#### 6.3. The Principle of Minimum Complementary Energy

In order to obtain the principle of minimum complementary energy, we integrate the term  $(-\sigma_{ji}\eta_{i,j})$  in (50) partially, and we take the equations equivalent to  $\delta_{\eta}R \leq 0$  (i.e. (56)) as auxiliary conditions, so that we do not have to vary  $\eta$ . In so doing we eliminate  $\lambda'$  by (56c), (56d). We obtain:

$$\delta_{\sigma}C \ge 0, \text{ subject to } \sigma_{A} = \mathbf{A}^{T}\sigma_{I}; \ \sigma_{Ji,J} = 0 \text{ in } G; \ \sigma_{Ji}n_{J} \equiv \pi_{i} = \tilde{\pi}_{i} \text{ in } S_{p}; \\ |p_{T} + \pi_{T}| \le g, \ (p_{N} + \pi_{N}) \ge 0, \ (p_{Ti} + \pi_{Ti})(\eta_{Ti} - \phi_{Ti}) + g|\eta_{T} - \phi_{T}| = 0 \text{ in } S_{c}; \end{cases}$$

$$R = C = \int_{G} \frac{1}{2} S_{ijkk} \sigma_{ji} \sigma_{kh} dG - \int_{Su} \pi_{i} \tilde{\eta}_{i} dS + \int_{S_{\sigma}} \{(p_{N} + \pi_{N})h(y + \eta, t + \tau)/|h_{y}| - (p_{i} + \pi_{i})\eta_{i} - g|\eta_{T} - \phi_{T}|\} dS.$$

$$(60)$$

We consider the integral over  $S_c$ , and take into account the second relation of (56e), and (22b):

$$\int_{S_{\sigma}} \{(p_{N} + \pi_{N})h(y + \eta, t + \tau)/|h_{y}| - (p_{i} + \pi_{i})\eta_{i} - g|\eta_{T} - \phi_{T}|\} dS$$
  
= 
$$\int_{S_{\sigma}} \{(p_{N} + \pi_{N})h(y, t + \tau)/|h_{y}| + (p_{N} + \pi_{N})\eta_{N} - (p_{N} + \pi_{N})\eta_{N} - (p_{T_{i}} + \pi_{T_{i}})\eta_{T_{i}} - g|\eta_{T} - \phi_{T}|\} dS$$
  
= 
$$\int_{S_{\sigma}} \{(p_{N} + \pi_{N})h(y, t + \tau)/|h_{y}| - (p_{T_{i}} + \pi_{T_{i}})\phi_{T_{i}}\} dS$$

so that (60) becomes

$$C = \int_{G} \frac{1}{2} S_{ijhk} \sigma_{ji} \sigma_{kk} \, dG - \int_{S_u} \pi_i \tilde{\eta}_i \, dS + \int_{S_u} \{(p_N + \pi_N)h(y, t+\tau)/|h_y| - (p_{T_i} + \pi_{T_i})\phi_{T_i}\} \, dS$$

R = C subject to (59).

If the coercivity relation (39) holds, then also when  $\sigma_A = \mathbf{A}^T \sigma_I$ ,  $S_{ijkk} \sigma_{ji} \sigma_{kk} \ge b \sigma_{ji} \sigma_{ji}$ , b > 0, constant, and it may be established as in Section 4 that C has a unique global minimum at the solution. Since then  $\sigma$  is unique it follows that the linearized strains  $\gamma$  are unique; and since  $\eta$  equals  $\tilde{\eta}$  on  $S_u$  it follows that  $\eta$  is unique, see Section 4. Thus we have, by (58) and (61):

(61)

$$-V(admiss. \eta) \leq -V(solution) = R(solution) = C(solution) \leq C(admiss. \sigma).$$
 (62)

## 7. Surface Mechanical Formulation

In surface mechanics we are exclusively interested in the stress and strain on the surface of the elastic body. The literature on contact mechanics is for the most part surface mechanical in nature; for, if the stresses and strains on the surface are known, the determination of the displacement-stress field inside the body is a matter which lacks contact mechanical characteristics. In accordance with this we will formulate the principles of minimum potential and complementary energy (35) and (61, 59) in surface mechanical form.

To that end we observe that in the solution  $\eta$  and  $\sigma$  are an equilibrium displacementstress field which satisfies the stress-strain relations:

$$\sigma_{ji,j} = 0 \text{ in } G, \, \sigma_{ji} = E_{ijhk} \eta_{h,k}, \, \sigma_A = \mathbf{A}^T \sigma_I \text{ in } G. \tag{63}$$

So in order to find the solution of (35) and (61, 59) we may confine ourselves to fields which satisfy (63), but perhaps not the boundary conditions (27), (59). Consider (35). The volume integral is integrated partially, and (63) is taken into account. This gives

$$\min_{\pi, \eta \text{ together}} V = \int_{S_p} (\frac{1}{2}\pi_i - \tilde{\pi}_i)\eta_i \, dS + \int_{S_u} \frac{1}{2}\pi_i \eta_i \, dS + \int_{S_o} \{(p_i + \frac{1}{2}\pi_i)\eta_i + g|\eta_T - \phi_T|\} \, dS$$
(64a)  
subject to  $\pi = \tilde{\pi} \text{ in } S + h(u + \pi t + \tau) > 0 \text{ in } S + (63)$ 

subject to  $\eta_i = \tilde{\eta}_i$  in  $S_u$ ,  $h(y+\eta, t+\tau) \ge 0$  in  $S_c$ ; (63). (64b)

In (61) we replace  $S_{ijhk}\sigma_{ji}\sigma_{kh}$  by its equivalent  $E_{ijhk}\eta_{i,j}\eta_{h,k}$ . Then it follows from (61, 59, 63) in the same manner

$$\min_{x, \eta \text{ together}} C = \int_{S_u} \pi_i (\frac{1}{2}\eta_i - \tilde{\eta}_i) \, dS + \int_{S_p} \frac{1}{2}\pi_i \eta_i \, dS + \int_{S_c} \frac{1}{2}\pi_i \eta_i + (p_N + \pi_N)h(y, t + \tau)/|h_y| - (p_{T_i} + \pi_{T_i})\phi_{T_i} \, dS \qquad (65a)$$

subject to 
$$\pi_i = \tilde{\pi}_i$$
 in  $S_p$ ,  $(p_N + \pi_N) \ge 0$ ,  $|p_T + \pi_T| \le g$  in  $S_c$ . (65b)

It should be noted that

١

$$\delta \int_{\partial G} \frac{1}{2} \pi_i \eta_i \, dS = \int_{\partial G} \pi_i \delta \eta_i \, dS = \int_{\partial G} \eta_i \delta \pi_i \, dS. \tag{66}$$

Remark 1. If we compare (64) with (65), or (35) with (61), it is seen that (64) and (35) contain a term  $g|\eta_T - \phi_T|$ , which is not differentiable if  $(\eta_T - \phi_T) = 0$ . This causes difficulties when the principles (64)-(35) are implemented numerically by discretizing the integrals and by then applying an optimization algorithm (but see Kalker, 1971*a*). In the principle of stationary complementary energy such difficulties are conspicuously absent, which renders this principle, stated in (61, 59) and (65), a most fruitful principle upon which to base a numerical implementation.

Remark 2. In the derivation of the principles (35), (61, 59), (64), (65) it has been assumed that h and  $\phi_T$  are given. This means in particular, that quantities like the depth of penetration of the rigid body and its rigid motion with respect to the elastic body are given. In many cases this is not realistic: instead of these kinematical quantities the force and the moment that the rigid body exerts on the elastic body are known, that is, in vector notation,

$$\int_{S_c} (\mathbf{p} + \pi) \, dS = \mathbf{f} = \text{given}, \qquad \int_{S_c} \mathbf{y} \times (\mathbf{p} + \pi) \, dS = \mathbf{m} = \text{given} \tag{67}$$

with 
$$\mathbf{p} = (p_i), \pi = (\pi_i), \mathbf{y} = (y_i)$$
, vectors;  $\times$  : cross product. (68)

The conditions (67) are isoperimetric in character, and may be entered in the minimum principles of complementary energy with the aid of constant Lagrange multipliers  $\lambda^{f}$  and  $\lambda^{m}$ . We will do so in (65). The last integral becomes

$$\begin{cases} \int_{S_c} \{\frac{1}{2}\pi \cdot \mathbf{\eta} + (p_N + \pi_N)[h(y, t+\tau)/|h_y| - \lambda_N^f - (\lambda^m \times \mathbf{y})_N] - (p_T + \pi_T) \cdot [\mathbf{\phi}_T + \lambda_T^f + (\lambda^m \times \mathbf{y})_T] \} dS; \end{cases}$$

$$\tag{69}$$

(70)

with  $\eta = (\eta_i), \phi = (\phi_i); .:$  inner product; subscript N: normal component (scalar); subscript T: tangential component (vector).

This means, in effect, that the surface of the rigid body is now given up to an additive constant  $\lambda_N^f$  and the normal component of a rigid rotation about the origin  $(\lambda^m \times y)_N$ . The shift  $\phi$  is likewise modified by the arbitrary rigid displacement  $\lambda^f + (\lambda^m \times y)$ . The conditions (67) now figure as auxiliary conditions.

One can also require that only the total force in a certain direction is prescribed while in the other two directions orthogonal to it the displacement is given. This means that the complementary energy principle is extremely flexible as regards the global conditions that prevail in a contact.

#### 7. Examples

### 7.1. Kalker-van Randen

The first example is the variational principle of Kalker-van Randen (1972) for the normal half-space contact problem. Take

$$S_{\mu} = \infty, \tilde{\eta} = 0; \tilde{\pi} = 0 \text{ on } S_{\nu}, y = x \Rightarrow p_{i} = s_{ii} = 0.$$

The problem is frictionless, hence g = 0. Surface displacement and normal load are connected by the Boussinesq-Cerrutti integral representation (Love, 1952),

$$\eta_i(x) = \int_{S_c} K_{ij}(q-x)\pi_j(q) \, dq, \qquad K: \text{ matrix kernel.}$$
(71)

In this integral representation, the displacement  $\eta$  vanishes at infinity, and the traction vanishes on  $S_p$ , the surface of the half-space outside  $S_c$ . The problem becomes

$$\min_{i, \eta \text{ together}} C \equiv \int_{S_o} \left\{ \frac{1}{2} \pi_i \eta_i + \pi_N h(x) / |h_x| \right\} dS, \qquad \pi_T = 0.$$

We introduce (71) in this problem, and find

$$\min \left\{ C \equiv \iint_{S_{\delta}^{2}} \frac{1}{2} \pi_{N}(x) K_{NN}(q-x) \pi_{N}(q) \, dq \, dx + \int_{S_{c}} \pi_{N}(x) h(x) / |h_{x}| \, dx \\ \text{subject to } \pi_{N} \ge 0 \text{ in } S_{c}; \, K_{NN} = N_{l} K_{ij} N_{j}.$$

$$(72)$$

# 7.2. Frictional Contact of a Rigid Body with an Elastic Half-space

The half-space  $x_N \ge 0$  is fixed at infinity; its surface outside contact is free of traction. Classical elasticity is taken to be valid. The surface traction is denoted by X, the displacement with respect to the undeformed state x is denoted by u. We have:

$$X_{i} = X_{i}(x, t); \quad X_{i} = 0 \text{ in } S_{p}; \\ u_{i} = u_{i}(x, t); \quad u_{i} = 0 \text{ in } S_{w} = \infty; \quad u_{i}, u_{i,j} \text{ small}; \quad y = x + u. \end{cases}$$
(73)

The equation of the surface of the rigid body is given by

$$h(\xi, t+\tau) = \xi_N - H(\xi_T, t+\tau), \xi_i \text{ coordinates, } H = O(u) \text{ in } S_c;$$
  

$$\xi_N = 0: \text{ undeformed half-space surface;}$$
  

$$h_{pi} = N_i, |h_p| = 1.$$
(74)

We write

$$\bullet = \partial/\partial t; \quad \eta_i = \tau \dot{u}_i, \, \pi_i = \tau \dot{X}_i. \tag{75}$$

As to the rigid body, let V = V(t) be the velocity of the rigid body at the origin and  $\omega = \omega(t)$  its angular velocity about the origin. The velocity of the rigid body  $\Phi = \Phi(y')$  at y' is given by

$$\Phi_i = V_i + \varepsilon_{ijk} y'_j \omega_k, \ \varepsilon_{ijk}: \text{ alternating symbol}$$
(76)

so that the shift  $\phi = \phi(y')$  becomes

$$\phi_i = (V_i + \varepsilon_{ijk} y'_j \omega_k) \tau$$

Since in the contact area  $y' = y \simeq x = x_T$ , the tangential component of the shift becomes

$$\phi_{T_i} = (V_i + \varepsilon_{ijk} x_{T_j} \omega_K)_T \tau. \tag{77}$$

In particular, if we take N as the  $x_3$ -direction we have

$$\begin{array}{c} x_{T} = (x_{1}, x_{2}, 0), \quad x_{N} = x_{3}; \\ \phi_{T_{1}} = \phi_{1} = (V_{1} + x_{2}\omega_{3})\tau, \ \phi_{T_{2}} = \phi_{2} = (V_{2} - x_{1}\omega_{3})\tau, \ \phi_{T_{3}} = 0. \end{array} \right\}$$
(78)

A number of constant terms are omitted in (65), which thus appears to be equivalent to

$$\min_{\mathbf{x}, \mathbf{y} \text{ together}} \left\{ C^{\bullet} = \int_{S_{\mu}} \pi_{i} (\frac{1}{2}\eta_{i} - \tilde{\eta}_{i}) \, dS + \int_{S_{\mu}} \frac{1}{2} \pi_{i} \eta_{i} \, dS + \int_{S_{\mu}} \frac{1}{2} \pi_{i} \eta_{i} + \pi_{N} h(y, t+\tau) / |h_{y}| - \pi_{T_{i}} \phi_{T_{i}} \right\} \, dS$$

$$\text{subject to } \pi = \tilde{\pi} \text{ in } S_{\mu}; \, p_{N} + \pi_{N} \ge 0, \, |p_{T} + \pi_{T}| \le q \text{ in } S_{c}.$$

$$(79)$$

subject to  $\pi = \tilde{\pi}$  in  $S_p$ ;  $p_N + \pi_N \ge 0$ ,  $|p_T + \pi_T| \le g$  in  $S_c$ . Introduction of (73)-(77) into (79) yields:

$$\min_{\substack{X_{i_{l}}, \, \mu_{l} \text{ together}}} \sum_{S_{c}} \{ \frac{1}{2} \dot{X}_{i} \dot{u}_{i} \tau^{2} + \dot{X}_{N} [u_{N} - H(x_{T}, \, t+\tau)] \tau - \dot{X}_{T_{l}} (V_{l} + \varepsilon_{ij_{k}} x_{T_{j}} \, \omega_{k})_{T} \tau^{2} \} \, dS$$
subject to  $X_{N} + \dot{X}_{N} \tau \ge 0, \, |X_{T} + \tau \dot{X}_{T}| \le g;$ 

$$\dot{u}_{l}(x) = \int_{S_{c}} K_{ij} (x-q) \dot{X}_{j}(q) \, dq.$$
(80)

Starting from known values of X at the time t,  $\dot{X}\tau$  and  $\dot{u}\tau$  are calculated according to a scheme analogous to the one given in Section 5. Then

$$X(t+\tau) = X(t) + \tau \dot{X}(t),$$

and the time is advanced by a step  $\tau$ , etc.

Note that a state of steady rolling can only be found with this principle by taking the limit as  $t \to \infty$ . Although a state of steady rolling is virtually achieved when the roller, under steady state conditions, has traversed approximately one contact width

(see Kalker, 1971*a*), the principle (80) does not seem very suitable for the problem of steady rolling. Variational principles which give the steady state directly are found in Kalker, 1966, 1967 (a non-convex principle) and 1971*a* (a convex, non-differentiable principle).

#### References

ARTHURS, A. M. 1970 Complementary variational principles. Oxford: Clarendon Press.

CONRY, T. F. & SEIREG, A 1971 J. Appl. Mech. 38, 387-392.

DE PATER, A. D. 1962 In Proc. Symp. Rolling contact phenomena (Ed. Bidwell). Elsevier.

DUVAUT, G. & LIONS, J. L. 1972 Les inequations en mécanique et en physique. Paris: Dunod.

ENGEL, P. A. 1975 In Proc. Symp. On the mechanics of contact between deformable bodies (Ed. A. D. de Pater & J. J. Kalker). Delft University Press.

FICHERA, G. 1964 Mem. Accad. Naz. Lincei ser. 8 7, 91-140.

FRIDMAN, V. M. & CHERNINA, V. S. 1967 (In Russian.) Mekhanika Tverdogo Tela AN USSR 1, 116-120.

JOHNS, T. G. & LEISSA, A. W. 1975 In Proc. Symp. On the mechanics of contact between deformable bodies (Ed. A. D. de Pater & J. J. Kalker). Delft University Press.

JOHNSON, K. L. 1962 In Proc. Symp. Rolling contact phenomena (Ed. Bidwell). Elsevier.

KALKER, J. J. 1966 Wear 9, 20-38.

KALKER, J. J. 1967 Thesis, Delft. Available from the author.

KALKER, J. J. 1971a J. Appl. Mech. 38, 875-887.

KALKER, J. J. 1971b In Kontaktnoe vzaimodeistvie twordikh tel Moscow: Izdatel'stwo Nauka (in Russian). English translation available from the author.

KALKER, J. J. & VAN RANDEN, Y. 1972 J. Engng Math 6, 193-206.

LANCZOS, C. 1971 The variational principles of mechanics. University of Toronto Press.

LIONS, J. L. & STAMPACCHIA, G. 1967 Comm. Pure Appl. Math. 20, 493-579.

Love, A. E. H. 1952 A treatise on the mathematical theory of elasticity. 4th Edn. Cambridge University Press.

Noble, B. & Sewell, M. J. 1972 J. Inst. Maths Applics 9, 123-193.

PACZELT, I. 1974 (In Russian.) Acta Techn. Acad. Sci. Hung. Tom 76, 217-241.

PACZELT, I. 1977 Acta Techn. Acad. Sci. Hung. (to be published).

SIGNORINI, A. 1959 Rc. Mat. Applic. 18.

SINGH, K. P. & PAUL, B. 1974 J. Appl. Mech. 41, 484-490.

Symbol	Defined in	Symbol	Defined in
a <sub>ji</sub>	(2)	Т	in Section 2: Fig. 2
Á	(42)	subscript T	(23)
subscript A	(44)	$u = u(x, t); u_i$	(73)
В	Section 6.1. above (42)	$V = V(t); V_i$	(75) to (76)
С	(20); (61)	V	(4), (35)
$E = E([y_{i}]_{i} +$		$V_N$	(40)
$\eta_{1,1}[y_{1,1}+\eta_{1,1}],x)$	(17), (29)	.,	
$E_{ijkk}$	(28)	VT	(41)
E	Section 6.1 above (42)	δΫ	(25) to (26)
f	(2)	δW	(1)
f	(67)	$x; x_i$	Section 3 above (16)
g	(25), (34)	$X = X(y); X_t$	(26), (31); in Section 7.2: (73)
G	Section 3 above (16)	$y = y(x); y_i$	Section 3 above (16)
$\overline{h}(\ldots)$	(20)	v': v'	(21b) Fig. 3
$h_{\rm m} = h_{\rm m}(y, t); h_{\rm m}$	(21b), (27)	Z, , , , , , , , , , , , , , , , , , ,	Fig. 1: (8)
н Н	(47); (74)	ν: ν,,	(39)
subscript I	(44)	δ	Variation symbol: (37c)
K. Ku. Kun	(71), (72)	8 <b>.</b>	(76)
,ıj, <sub>NN</sub>	(8)	$\lambda', \lambda', \lambda_{a'}$	(50), (51)
m	(67)	$\lambda^{f}, \lambda^{m}$	(69)
N	in Section 2. Fig. 2	11	(13), (25)
$N = N(y); N_i$	Section 3 above (16)	ξ; ξι	Cartesian
subscript N	(23)	$\pi = \pi(\mathbf{x}); \pi$	(18) to $(19)$ , $(32)$
p = p(x)	(18) to $(19)$ $(32)$	$\sigma = \sigma(\mathbf{x}); \sigma_{i}$	(28) $(31)$
$p = p(x); p_1$	Section 2	$\tau = \sigma(x), \sigma_{ff}$	Section 3 above (16)
96 ×1 R	(50)	$\dot{\phi} = \phi(v') \cdot \phi$	(23) to (24)
R.,	(3)	$\varphi = \varphi(y'), \varphi_i$ $\Phi = \Phi(y'), \Phi_i$	(76)
*)I	Section 2: Fig. 2 (3)	$\varphi = \varphi(t), \varphi_1$	(75) to $(76)$
$s = s(x) \cdot s$	(28) (31)	$\omega = \omega(r), \omega_l$	
$s = s(x), s_{jj}$	(20), (31)		
s, s	(21) (21a)		
Sp C#	(214)		
$S_p^+$	(17)		

# Appendix: List of Symbols

. .