JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 74, No. 1, JULY 1992

On the Problem of Optimizing Contact Force Distributions[1]

A. KLARBRING[2]

Communicated by E. J. Haug

Abstract. The problem of optimizing the distribution of contact forces between a rigid obstacle and a discretized linear elastic body is considered. The design variables are the initial gaps between the potential contact nodal points and the obstacle. Two different cost functionals are investigated: the first reflects the objective of minimizing the maximum contact force; the second is the equilibrium potential energy. Contrary to what has been claimed in the literature, it is shown that these cost functionals do not give, in general, the same optimal design. However, it is also shown that, if a certain frequently realized assumption is met by the system flexibility matrix, then this equality does hold.

The min-max cost functional is nonconvex and nondifferentiable, and Clarke's theory of nonsmooth optimization is used to establish a sufficient optimality condition. Investigating its consequences, both necessary and sufficient optimality conditions can be given. The equilibrium potential energy cost functional, on the other hand, turns out to have the remarkable porperties of differentiability and convexity.

Key Words. Contact problems, optimal shape design, optimality conditions, nondifferentiable optimization.

1. Introduction

Consider a linear elastic body, fixed over a part of its boundary, while another part may come into frictionless contact with a rigid support. This paper is concerned with the optimal shape design problem of finding the best shape of the obstacle. Similar problems have been considered in Conry and Seireg (Ref. 1), Haug and Kwak (Ref. 2), Benedict and Taylor

[1]This work was supported by The Center for Industrial Information Technology (CENIIT), Linköping Institute of Technology, Linköping, Sweden.

[2]Professor, Department of Mechanical Engineering, Division of Solid Mechanics, Linköping Institute of Technology, Linköping, Sweden.

(Ref. 3), Benedict (Ref. 4), Kikuchi and Taylor (Ref. 5), Benedict, Sokolowski and Zolesio (Ref. 6), Bendsøe and Sokolowski (Ref. 7), and in the recent monograph by Haslinger and Neittaanmäki (Ref. 8). The definition of "best," i.e., the choice of cost functional, differs between these publications. In the earlier works (Refs. 1 and 2), where discrete or discretized bodies are treated, the objective was to minimize the maximum contact force (or stress). However, such cost functional is nondifferentiable, and mainly because of this difficulty, Benedict and Taylor (Ref. 3) suggested to minimize the equilibrium potential energy of the system. Intuitively, they argued that such cost functional should give the same result as the min-max one. This conjecture seems to have been accepted without further investigation. In this paper, we treat this issue in some detail for the case of discrete or discretized bodies, when the design constraint is of volume or isoperimetric type. It is shown that in general the conjecture of Benedict and Taylor is false. However, under a certain condition on the flexibility matrix of the elastic body or structure, it can be shown to hold. The mechanical interpretation of this condition is that, when equal contact forces are applied to all contact nodes, no associated (work conjugate) displacement is negative. Such a condition can be expected to hold for many practical situations, notably, for Hertz-type contact problems (the elastic body can be well approximated by a half-space).

Following this introduction, the discrete contact problem is introduced as in Klarbring (Ref. 9) and some useful quadratic programming formulations are given. Next, the two optimal design problems arising from the use of the two different cost functionals are treated separately.

Section 3 is concerned with the problem of minimizing the maximum contact force [Problem (M)]. The existence of a solution is first shown. Then the theory of nonsmooth and nonconvex optimization developed by Clarke (Ref. 10) is used to establish a sufficient optimality condition. Interpreting this condition, a useful result is established: for an optimal design, the equilibrium configuration is such that all potential contact nodes are actually in contact with the obstacle. Using the restriction inferred by this result, the problem can be restated as a linear programming problem, and necessary and sufficient optimality conditions can be established.

In Section 4, the problem of minimizing the equilibrium potential energy is treated. It is first noted that it can be equivalently stated as the problem of maximizing the equilibrium reciprocal energy [Problem (P)]. The reciprocal energy is the energy associated with a formulation of contact problems in terms of contact forces; see Ref. 9. This problem turns out to have two remarkable properties: firstly, the cost functional is differentiable as a function of the design, despite the fact that the solution of the state problem is not; secondly, it is concave. The latter property enables us to set up a necessary and sufficient optimality condition that, in fact, can be solved in closed form for the optimal design.

In Section 5, the proof of equivalence of the two problems is given; in Section 6, a modification of the design constraint is discussed. In Section 7, some conclusions and comments end the paper.

Notation. The paper uses the following notations for subvectors and submatrices: if a vector x is of order n, $I \subset \{1, \ldots, n\}$, then x_I represents the subvector of x consisting of all elements x_i for $i \in I$; if a matrix A is of order $n \times m$, $I \subset \{1, \ldots, n\}$ and $J \subset \{1, \ldots, m\}$, then $A_{I\cdot}$ represents the submatrix of A consisting of all the rows $A_{i\cdot}$ for $i \in I$, while $A_{\cdot J}$ represents the submatrix of A consisting of all the columns $A_{\cdot j}$ for $j \in I$; finally, $A_{IJ} = (A_{I\cdot})_{\cdot J}$.

2. State Problem

Consider a discrete or discretized elastic body. Its configurations are represented by a vector u of nodal displacements and the forces acting on it by a vector F. Assume that the structure is free of mechanisms so that u and F are related by a symmetric positive-definite stiffness matrix K, i.e.,

$$F = Ku. \tag{1}$$

Furthermore, assume that a vector w of contact displacements may be derived from the nodal displacements as

$$w = C_1 u, \tag{2}$$

where C_1 is a kinematic transformation matrix. Introduce also a matrix C_2 such that $[C_1^t, C_2^t]^t$ is nonsingular. The superscript t means transpose of vector or matrix. A new displacement vector u_2 is defined by

$$u_2 = C_2 u. \tag{3}$$

For vectors P of contact forces and F_2 of prescribed forces, work conjugate with w and u_2, we then obtain

$$F = C_1^t P + C_2^t F_2. \tag{4}$$

Next, frictionless contact conditions are introduced. Let g be a vector of gaps between contact boundary nodes and the rigid obstacle. We then have the following complementarity condition

$$w \leq g, \qquad P \leq 0, \qquad P^t (w - g) = 0. \tag{5}$$

It can be seen (Ref. 9) that (1), (2), (4), and (5) represent the necessary and sufficient Kuhn-Tucker conditions of the following quadratic program (QP):

$$\text{(P)} \qquad \Pi_p(u) = \min\{\Pi_p(u^*) \mid u^* \in K_p\}, \tag{6}$$

where

$$\Pi_p(u) = 1/2u^t Ku - F_2^t C_2 u \tag{7}$$

is the potential energy function and

$$K_p := \{u \mid C_1 u \leq g\} \tag{8}$$

is the set of kinematically admissible displacements.

Problem (P) is a formulation of the frictionless contact problem in terms of displacements. It will next be shown that a QP problem in terms of contact forces may be derived. From (1), (2), and (4), with

$$A = C_1 K^{-1} C_1^t, \qquad w^* = C_1 K^{-1} C_2^t F_2,$$

we have

$$AP = w - w^*. \tag{9}$$

Then, (5) and (9) represent the Kuhn-Tucker conditions of the following QP:

$$\text{(R)} \qquad \Pi_R(P) = \min\{\Pi_R(P^*) \mid P^* \in K_R\}, \tag{10}$$

where

$$\Pi_R(P) = 1/2P^t AP + P^t(w^* - g), \tag{11}$$

$$K_R := \{P \mid P \leq 0\}. \tag{12}$$

This is the so-called reciprocal formulation of the contact problem; see Refs. 9 and 11-13.

It may be seen (Ref. 9) that (P) and (R) are dual problems in the sense of the duality theory of QPs. We have the following result.

Proposition 2.1. Problems (P) and (R) have unique solutions u and P which are related by the equation

$$u = K^{-1}(C_1^t P + C_2^t F_2). \tag{13}$$

For the solutions u and P, it further holds that

$$\Pi_p(u) + \Pi_R(P) + 1/2F_2^t C_2 K^{-1} C_2^t F_2 = 0. \tag{14}$$

For completeness we write down explicitly the Kuhn-Tucker conditions of problem (R),

$$AP + w^* - g \leq 0, \qquad P \leq 0, \qquad (AP + w^* - g)^t P = 0. \tag{15}$$

This is a linear complementary problem (LCP), so the existence of a unique solution can also be inferred from the theory of such problems.

3. Problem (M): Contact Force Minimization

A natural objective when choosing the shape of contacting bodies is to minimize the peak contact force (or stress in the continuous case, see Section 7). This problem was considered in Refs. 1, 2, and 6. Here, we will consider the case where the shape of the rigid obstacle is altered while the deformable structure is fixed, meaning that the design variable is g. Since (R) is uniquely solvable, the function $g \mapsto P(g)$ is well defined. As cost functional, we consider the max function

$$J_M(g) = \max\{-P_i(g) \mid 1 \le i \le n\}, \tag{16}$$

where n is the number of contact points and P_i is the ith element of P. Obviously, if the design is not subjected to restrictions, the minimum value of J_M will be zero. Relevant design conditions are

$$g_1 \le g \le g_2, \tag{17}$$

$$\mathrm{U}^t g = \sum_{i=1}^{n} g_i \le V, \tag{18}$$

where g_1 and g_2 are fixed vectors, V represents the volume of the gap between the structure and the rigid obstacle, and $\mathrm{U}^t = (1, \dots, 1)$ is a vector of ones. Note that V may be negative.

In the present study, our main concern is with the constraint (18). In Section 6, a discussion is given concerning modification of the results when (17) is imposed.

Let K be the closed convex set defined by (18). We then state our problem (M) formally: find $g \in K$ such that

$$\text{(M)} \qquad J_M(g) = \min\{J_M(g^*) \mid g^* \in K\}, \tag{19}$$

The functions $g \mapsto P_i(g)$, $1 \le i \le n$, are not differentiable due to the unilateral constraint $P \in K_R$. However, they are directionally differentiable and Lipschitz continuous (Refs. 14 and 15). The latter implies that also J_M is Lipschitz continuous (Ref. 10). Furthermore, $P_i(g)$ is in general nonconvex (even nonmonotone) as it is understood from an elementary example given by Cottle (Ref. 16). Thus, J_M will in general be nonconvex, as well as nonsmooth. However, it is possible to assure the existence of a minimum value.

Theorem 3.1. Problem (M) has a solution.

Proof. Problem (M) is equivalent to the problem of finding $\beta \in R$ such that

$$\beta = \min\{\beta^* \mid \beta^* + P_i(g^*) \ge 0,\ 1 \le i \le n,\ g^* \in K\}.$$

Owing to (15), this problem can be rewritten as a complementary programming (CP) problem, Refs. 17 and 18:

$$\begin{aligned} &\min && \beta, \\ &\text{s.t.} && \beta + P_i \geq 0, \ 1 \leq i \leq n, \\ &&& U^t g \leq V, \\ &&& AP + w^* - g \leq 0, \ P \leq 0, \\ &&& P^t(AP + w^* - g) = 0. \end{aligned}$$

Next, let L and M be disjoint index subsets of the set $\{1, \ldots, n\}$. Then, the CP problem can be represented as 2^n auxiliary linear programming (LP) problems, defined as

$$\begin{aligned} &\min && \beta, \\ &\text{s.t.} && \beta + P_i \geq 0, \ 1 \leq i \leq n, \\ &&& U^t g \leq V, \\ &&& P_L \leq 0, \ A_L.P - g_L + w_L^* = 0, \\ &&& P_M = 0, \ A_M.P - g_M + w_M^* \leq 0. \end{aligned}$$

Not all the auxiliary problems have feasible solutions. However, at least one such problem has a solution: fix an admissible g and solve (15) to get the corresponding sets L and M. Furthermore, all optimum values are bounded below by zero. Thus, the solution av(M) follows by comparing the finite solutions of a finite number of LP problems. □

The idea of representing problem (M) by means of auxiliary LP problems can be made the basis of algorithms for its solution (Ref. 2).

The solution need not be unique, as is demonstrated by the following example.

Example 3.1. Let

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}, \qquad w^* = 0,$$

and

$$\text{(a)} \qquad g_1 + g_2 \leq V,$$

where $V < 0$. Lemma 3.1 will show that, if strict inequality holds in (a) at a minimum point of J_M, then $P = 0$ at that point. On the other hand, from (15) it is seen that $P = 0$ contradicts $V < 0$. Thus, we can consider only the

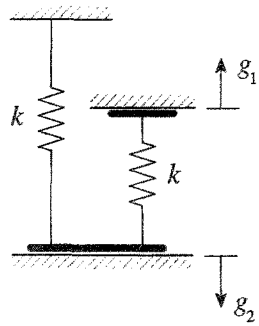

Fig. 1. Spring model representing the flexibility matrix of Example 3.1. The spring constant is $k=1$. The structure is shown in the unstressed configuration.

case where (a) holds as an equality. The functions $g \mapsto P_i(g)$ are then given by

$$P(g)^t = (V, V+g_2), \qquad \text{if } V+g_2 \leq 0,$$
$$P(g)^t = ((V-g_2)/2, 0), \qquad \text{if } V+g_2 > 0.$$

Then, any $g_1 \equiv V-g_2$ and g_2 with $0 \leq g_2 \leq -V$ is an optimal solution such that $J_M = -V$. This problem may be represented by the spring model of Fig. 1.

Next, we want to find necessary conditions for a point to be a local or global minimum of J_M. Since both the max function and P_i are nondifferentiable functions, this is not a standard problem. However, the theory developed by Clarke and summarized in Ref. 10 provides a useful tool for its solution: in our notations, the corollary of Proposition 2.4.2 in Ref. 10 reads as follows.

Corollary 3.1. Suppose that J_M is Lipschitz (as was already shown above) and attains a minimum (local or global) over K at g. Then,

$$0 \in \partial J_M(g) + N_K(g). \tag{20}$$

Here, $N_K(g)$ is the normal cone operator defined by

$$N_K(g) := \begin{cases} \{x \in R^n \mid x^t(f-g) \leq 0, \forall f \in K\}, & \text{if } g \in K, \\ \varnothing, & \text{if } g \notin K, \end{cases} \tag{21}$$

and $\partial J_M(g)$ is the generalized gradient of Clarke (Ref. 10). If J_M is convex, the generalized gradient coincides with the subdifferential of convex analysis (Ref. 10). To present the generalized gradient, we first note that Clarke's Proposition 2.3.12 gives

$$\partial J_M(g) \subset \text{co}\{-\partial P_i(g) \mid i \in M(g)\}, \tag{22}$$

where co stands for the convex hull and $M(g)$ is the set of indices i for which $J_M(g) = -P_i(g)$. Next, we will use Theorem 2.5.1 in Ref. 10. Let Ω_f be the set of points where P_i fails to be differentiable (which can be shown to be of measure zero when P_i is Lipschitz), and let S be any set of measure zero (which can of course be the empty set). Consider any sequence g_r converging to g while avoiding points belonging to $\Omega_f \cap S$ and such that the sequence $\nabla P_i(g_r)$ converges. Denote by $\lim \nabla P_i$ the set of all such limits. Then,

$$\partial P_i(g) = \text{co}\{\lim \nabla P_i\}. \tag{23}$$

To give a more explicit expression for (23), we will utilize the formulation (15) of Problem (R). For a given g, the unique solution P introduces the three index sets A, B, C defined by

$$i \in A \Leftrightarrow P_i < 0, \qquad A_i.P - g_i + w_i^* = 0, \tag{24a}$$

$$i \in B \Leftrightarrow P_i = 0, \qquad A_i.P - g_i + w_i^* = 0, \tag{24b}$$

$$i \in C \Leftrightarrow P_i = 0, \qquad A_i.P - g_i + w_i^* < 0. \tag{24c}$$

A sensitivity analysis (Ref. 14) shows that P_i are differentiable when $B = \varnothing$. Furthermore, it turns out that $\lim \nabla P_i$ includes only a finite number of elements which can be produced by including in all possible ways the indices of B in A and C; see Haslinger and Roubíček (Ref. 19). That is, letting

$$I = A \cup \bar{B} \quad \text{and} \quad J = C \cup B \backslash \bar{B},$$

where $\bar{B} \subset B$, and requiring that

$$A_I.P - g_I + w_I^* = 0, \qquad P_J = 0,$$

gives

$$P_I = (A_{II})^{-1}(g_I - w_I^*),$$

which means that

$$\lim \nabla P_i = \{x \,|\, x' = ((A_{II})_{i.}^{-1}, 0'), \text{ if } i \in I;\ x = 0, \text{ if } i \in J\}, \tag{25}$$

where I and J include complementary subsets of B in all possible ways as indicated above.

Finally, to write condition (20) explicitly, we characterize the normal cone. According to well-known results of convex analysis, we have that, if $g \in K$,

$$N_K(g) = \{x = \text{U}\mu \,|\, \mu \geq 0,\ \mu(\text{U}'g - V) = 0\}. \tag{26}$$

If $g \notin K$, the normal cone is, according to (21), the empty set.

Condition (20) states that there exists an element $x \in \partial J_M(g)$ such that $-x \in N_K(g)$. This means that, at a minimum point of J, there exists an element of $\partial J_M(g)$ that is equal to $-\mu$U. We have the following necessary optimality condition.

Theorem 3.2. At a design where J_M attains a minimum over K, there exists a multiplier $\mu \geq 0$ such that

$$\mu \mathrm{U} \in \mathrm{co}\{x \,|\, x^t = ((A_{II})_{i\cdot}^{-1}, 0^t), \text{if } i \in I;\ x = 0, \text{if } i \in J;\ i \in M(g)\}, \tag{27}$$

where i, I, and J take all possible values.

Two distinct cases arise, as reflected by the following two corollaries.

Corollary 3.2. If (27) holds with $\mu = 0$, then $A = \varnothing$ in the corresponding equilibrium state.

Corollary 3.3. If (27) holds with $\mu > 0$, then $C = \varnothing$ in the corresponding equilibrium state.

Corollary 3.2 holds, since the diagonal elements of $(A_{II})^{-1}$ are nonzero. As a consequence of these results, we have that $A \neq \varnothing$ implies $C = \varnothing$, and $C \neq \varnothing$ implies $A = \varnothing$.

Note that $A = \varnothing$ corresponds to an equilibrium state with zero contact force. $C = \varnothing$ means that, for the optimal design, the equilibrium state is such that all contact nodes are in contact.

The next theorem for some cases limits the set of possible minimizers. We first state two lemmas.

Lemma 3.1. If J_M attains a minimum at an interior point g of K, then $P(g) = 0$.

Proof. That g is an interior point of K means that $\mathrm{U}^t g < V$, which by (26) implies $\mu = 0$. Then, the result follows from Corollary 3.2. □

Lemma 3.2. If $\mathrm{U}^t w^* > V$, then for all $g \in K$, $P(g) \neq 0$.

Proof. Assume that $P(g) = 0$, which by (15) implies $w^* \leq g$. That is,

$$\mathrm{U}^t w^* \leq \mathrm{U}^t g \leq V,$$

and we obtain a contradiction. □

Theorem 3.3. Under the assumption of Lemma 3.2, any minimizer of J_M is such that $A \neq \varnothing$ ($\Rightarrow C = \varnothing$), i.e., $P \neq 0$, and J_M attains no minimum at interior points.

Under the assumptions of Lemma 3.2, this theorem tells us which one of the 2^n LP problems mentioned in the proof of Theorem 3.1 actually gives the solution. Since LP problems are convex, from the standard theory (see, for instance, Ref. 20) one obtains necessary and sufficient conditions.

Theorem 3.4. Let the assumption of Lemma 3.2 hold. Then, g is a solution of Problem (M) if and only if there exist $\mu \geq 0$ and $\lambda_2 \geq 0$ such that

$$\mu \mathrm{U} \in \mathrm{co}\{A_{\cdot i}^{-1} \mid i \in M(g)\} - A^{-1}\lambda_2,$$

$$AP - g + w^* = 0, \qquad \mathrm{U}^t g - V = 0,$$

$$P \leq 0, \qquad \lambda_2^t P = 0.$$

Proof. Take $M = \emptyset$ in the proof of Theorem 3.1. The Kuhn–Tucker conditions of the obtained problem are

$$1 - \lambda_1^t \mathrm{U} = 0,$$

$$-\lambda_1 + \lambda_2 + A\lambda_3 = 0,$$

$$\mu \mathrm{U} - \lambda_3 = 0,$$

$$\mu \geq 0, \qquad \mathrm{U}^t g \leq V, \qquad \mu(\mathrm{U}^t g - V) = 0,$$

$$\lambda_1 \geq 0, \qquad U\beta + P \geq 0, \qquad \lambda_1^t(\mathrm{U}\beta + P) = 0,$$

$$\lambda_2 \geq 0, \qquad P \leq 0, \qquad \lambda_2^t P = 0,$$

$$AP - g + w^* = 0.$$

Taking account of Theorem 3.3, the theorem then follows. □

A related intuitively obvious result that nevertheless should be explicitly stated is the following theorem.

Theorem 3.5. If $\mathrm{U}^t w^* \leq V$, then any $g \in K$ such that $w^* \leq g$ is a solution of Problem (M) and corresponds to $P = 0$.

Proof. That such g gives $P = 0$ is obvious from (15). Since $J_M \geq 0$, it is a solution if it exists. This, however, follows by taking $g = w^* \in K$. □

Note that Theorem 3.5 does not exclude the possibility of local minimizers such that $P(g) \neq 0$ even if the assumption is satisfied. To show such a property, we introduce an assumption on the matrix A, which is likely to hold in many practical situations.

Assumption A1. The sum of the elements of each column (or row) is positive; that is,

$$\sum_{i=1}^{n} A_{ij} > 0, \qquad \forall j \in (1, \ldots, n) \Leftrightarrow AU > 0 \Leftrightarrow U'A > 0'.$$

Theorem 3.6. Let Assumption A1 hold. Then, the following two situations never occur simultaneously:

(i) J_M has a minimizer g_1 such that $A \neq \emptyset$, i.e., $P(g_1) \neq 0$.
(ii) J_M has a minimizer g_2 such that $A = \emptyset$, i.e., $P(g_2) = 0$.

Proof. According to Lemma 3.1, situation (ii) implies that

$$U'w^* \leq V. \tag{28}$$

On the other hand, from (15), situation (i) means that, since $A \neq \emptyset$ implies $C = \emptyset$,

$$U'AP + U'w^* = U'g_1. \tag{29}$$

Also, Lemma 3.1 implies $U'g_1 = V$, which together with (28) and (29) gives

$$U'AP \geq 0$$

in situation (i). However, under Assumption A1, only $P = 0$ is compatible with this inequality, which contradicts $A \neq \emptyset$. □

In Section 5, it will be seen that Assumption A1 leads to even stronger results. In fact, the convex set of solutions of Theorem 3.4 turns out to consist of one point only, which can be explicitly given.

Theorem 3.7. Under Assumptions A and that of Lemma 3.2, Problem (M) has the unique solution g which corresponds to

$$P = U(V - U'w^*)/U'AU = A^{-1}(g - w^*).$$

Proof. This result follows from Corollary 4.2 and Theorem 5.1. □

4. Problem (P): Minimization of Potential Energy

Basically, since (16) is nondifferentiable, Benedict and Taylor (Ref. 3; see also Haslinger and Neittaanmäki, Ref. 8) have suggested to use as cost functional the potential energy in equilibrium instead of (16), i.e.,

$$\bar{J}_P(g) = \Pi_P(u(g), g), \tag{30}$$

where $g \mapsto u(g)$ is the well-defined function given by the unique solution of (6). Remarkably, $\bar{J}_P(g)$ turns out to be differentiable (Ref. 8) despite the fact that $g \mapsto u(g)$ is nondifferentiable.

Here, we will not consider (30) directly, but instead base our development on the reciprocal formulation (10). Due to relation (14), minimizing $\bar{J}_P(g)$ is equivalent to maximizing

$$J_P(g) = \Pi_R(P(g), g) \tag{31}$$

as was previously noted by Kikuchi and Taylor (Ref. 5). Thus, we consider the problem of finding $g \in K$ such that

$$\text{(P)} \qquad J_P(g) = \max\{J_P(g^*) \mid g^* \in K\}. \tag{32}$$

Note that, although (P) can be expressed as a saddle-point problem, the existence of a solution does not follow from the general theory of Ekeland and Temam (Ref. 21), since K is unbounded and the functional does not have the proper growth properties. Similarly, the theory in Ref. 8 demands that K be bounded. However, we will see that the solution can actually be constructed explicitly.

We now show that J_P is differentiable. Denote by

$$P' = \lim_{t \to 0^+} [P(g + \tilde{g}t) - P(g)]/t \tag{33}$$

the directional derivative of $P(g)$ at the point g in direction $\tilde{g}$. The directional derivative of J_P becomes

$$J'_P = P^t A P' + (w^* - g)^t P' - P^t \tilde{g} = (AP + w^* - g)^t P' - P^t \tilde{g}. \tag{34}$$

Now, if $i \in C$, then obviously $P'_i = 0$; and if $i \in A \cup B$, then from (24)

$$(AP + w^* - g)_i = 0.$$

Thus, we conclude that J_P is differentiable and that

$$\nabla J_P(g) = -P(g). \tag{35}$$

Standard theory of constraint maximization [or the equivalent of (20)] gives the following optimality condition.

Theorem 4.1. At a point $g \in K$ where J_P attains a maximum, there exists a multiplier $\bar{\mu} \geq 0$ such that

$$\bar{\mu} \mathrm{U} = -P(g). \tag{36}$$

Corollary 4.1. An optimal solution of (P) corresponds to a uniform distribution of contact forces.

Equation (35) is also the key to the proof of concavity of J_P.

Theorem 4.2. J_P is concave.

Proof. Set $P^* = P(g^*)$ and $P = P(g)$. Then,

$$\Pi_R(P, g^*) \geq \Pi_R(P^*, g^*).$$

Invoking the explicit form of Π_R, one gets

$$\Pi_R(P, g) - P^t(g^* - g) \geq \Pi_R(P^*, g^*),$$

and using (35) and the definition of J_P, one obtains

$$J_P(g) + \nabla J_P(g)^t(g^* - g) \geq J_P(g^*).$$

For a differentiable function, this inequality defines concavity (see, for instance, Ref. 20). □

Corollary 4.2. The following conditions are necessary and sufficient for a point $g \in K$ to be a solution of Problem (P):

$$\bar{\mu} \geq 0, \qquad \bar{\mu}(\mathrm{U}^t g - V) = 0, \tag{37a}$$

$$\bar{\mu}\mathrm{U} = -P(g). \tag{37b}$$

This result allows us to construct the solution explicitly.

Corollary 4.3. If $\mathrm{U}^t w^* > V$, then J_P has only one maximizer g, which is such that

$$P = -\bar{\mu}\mathrm{U} = \mathrm{U}(V - \mathrm{U}^t w^*)/\mathrm{U}^t A \mathrm{U} = A^{-1}(g - w^*).$$

Proof. From Lemma 3.2, we see that $\bar{\mu} > 0$, meaning that $\mathrm{U}^t g = V$ at a maximizer. In a neighborhood of a solution where $\bar{\mu} > 0$, the function $g \mapsto P(g)$ can be represented as

$$P(g) = -A^{-1}(w^* - g). \tag{38}$$

Then, taking (36) into account, the result follows. □

Theorem 4.3. If $V \leq \mathrm{U}^t w^*$, then a solution of Problem (P) exists, corresponds to $P = 0$, and is given by any $g \in K$ such that $w^* \leq g$.

Proof. Assume that there are maximizers such that $P \neq 0$, i.e., $\bar{\mu} > 0$. From Corollary 4.3, we see that this leads to $V > \mathrm{U}^t w^*$, so we get a contradiction. Existence follows by taking $g = w^* \in K$ and the characterization $w^* \leq g$ from (15). □

Example 4.1. This is Example 3.1 revisited. With the data of Example 3.1, the unique maximum point of J_P is $g^t = (V, 0)$, corresponding to $P = \mathrm{U} V$.

5. Comparison of Problems (M) and (P)

When comparing the two cost functionals, one notices that by definition J_M cannot give an optimal design such that the related contact forces have greater magnitude than the ones given by J_P. On the other hand, J_P gives a uniform distribution of contact forces that might be preferable from the point of view of applications. In this section, it is shown that, in important cases, the two cost functionals actually give the same result, so the uniform force distribution is the one with lowest maximum. However, we also show that this equality does not hold generally. Denote by g_M an optimal design given by J_M; correspondingly, denote by g_P an optimal design given by J_P. By definition,

$$-P_i(g_M) \leq -P_i(g_P), \qquad 1 \leq i \leq n, \tag{39}$$

and from this it follows that

$$\max\{-P_i(g_M) + P_i(g_P) \,|\, 1 \leq i \leq n\} \leq 0. \tag{40}$$

Example 3.1 has shown that, in general, equality does not hold for all i in (39). Nevertheless, it seems to indicate the possibility of an equality for some i and thereby the conclusion that the inequality in (40) can be substituted for an equality. However, the following example shows that such a proposition is false.

Example 5.1. Let

$$A = \begin{bmatrix} 5 & -2 \\ -2 & 1 \end{bmatrix}, \qquad w^* = 0,$$

and

$$g_1 + g_2 \leq V < 0.$$

As concluded in Example 3.1, when $V < 0$ it is necessary to consider equality only in the design constraint. The functions $g \mapsto P_i(g)$ are given by

$$P(g)^t = (2V - g_1, 5V - 3g_1), \qquad \text{if } 5V/3 \leq g_1,$$

$$P(g)^t = (g_1/5, 0), \qquad \text{if } 5V/3 > g_1.$$

Thus, the optimal solution of Problem (M) is

$$g_1 = 5V/3, \qquad g_2 = -2V/3,$$

which gives $J_M = -V/3$.

On the other hand, the unique minimum point of J_P is

$$g_1 = 3V/2, \qquad g_2 = -V/2,$$

corresponding to $P = \mathrm{U}V/2$. The spring model of Fig. 2 represents this problem.

Subsequently, we will give restrictions on the matrix A such that inequality can be substituted for equality in (39) and (40), respectively. We first note that we can restrict the investigation to the case where $\mathrm{U}^t w^* > V$, since (as shown in previous sections) in the other case both problems have coinciding trivial solutions corresponding to $P = 0$.

Theorem 5.1. Let Assumption A1 hold. Then,

$$P_i(g_M) = P_i(g_P), \qquad 1 \leq i \leq n.$$

Proof. Since we may consider only the case $\mathrm{U}^t w^* > V$, we have

$$AP(g_M) + w^* - g_M = 0, \qquad AP(g_P) + w^* - g_P = 0.$$

Also, it holds that $\mathrm{U}^t g_M = V$ and $\mathrm{U}^t g_P = V$, so we have

$$\mathrm{U}^t AP(g_M) + \mathrm{U}^t w^* - V = 0, \qquad \mathrm{U}^t AP(g_P) + \mathrm{U}^t w^* - V = 0.$$

That is,

$$\mathrm{U}^t A(P(g_M) - P(g_P)) = 0.$$

Invoking Assumption A1, the theorem follows. □

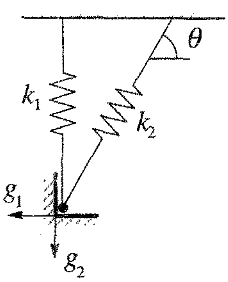

Fig. 2. Spring model representing the flexibility matrix of Example 5.1. The spring constants are $k_1 = 1$, $k_2 = 5$ and the angle $\theta = \arctan 2$. The structure is shown in the unstressed configuration.

As is obvious from the previous proof, a slightly weaker assumption leads to equality in (40).

Assumption A2. The sum of the elements of each column (or row) is nonnegative; that is,

$$\sum_{i=1}^{n} A_{ij} \geq 0, \qquad \forall j \in \{1, \ldots, n\} \Leftrightarrow A\mathrm{U} \geq 0 \Leftrightarrow \mathrm{U}^t A \geq 0^t.$$

Theorem 5.2. Let Assumption A2 hold. Then,

$$P_i(g_M) = P_i(g_P), \qquad \text{if } \mathrm{U}^t A_{\cdot i} > 0,$$

and there is always at least one such $i \in \{1, \ldots, n\}$.

Proof. The result follows from the proof of Theorem 5.1 and the fact that A is nonsingular. □

Note that Assumptions A1 and A2 are exact in the previous theorems. As shown by Examples 3.1, 4.1, and 5.1, when they are not satisfied, the conclusions of the theorems cannot be expected to hold in general.

6. Modified Design Constraint

In previous sections, the set of designs K was defined by the volume constraint (18). As noted in Section 3, a relevant constraint is also the box constraint (17).

When K is defined by both (17) and (18) and is nonempty, solutions of Problems (M) and (P) obviously exist, since K is then compact.

For such a new K, the optimality conditions (27) and (36) are modified only on their left-hand sides. That is, the normal cone needs to be differently expressed. If

$$K := \{g \mid \mathrm{U}^t g \leq V,\ g_1 \leq g \leq g_2\},$$

then

$$N_K(g) = \{x = \mathrm{U}\mu + \lambda_2 - \lambda_1 \mid \lambda_1, \lambda_2 \in R^n,\ \mu \geq 0,\ \mu(\mathrm{U}^t g - V) = 0,$$
$$\lambda_1 \geq 0,\ \lambda_1^t(g_1 - g) = 0,\ \boldsymbol{\lambda}_2 \geq 0,\ \boldsymbol{\lambda}_2^t(g - g_2) = 0\}.$$

The left-hand sides of (27) and (36) are changed for $\mathrm{U}\mu + \lambda_2 - \lambda_1$.

Going through the results of previous sections, we make the following conclusions for a changed K: only local versions of Corollaries 3.2 and 3.3 can be stated; i.e., if $\mu+\lambda_{2i}-\lambda_{1i}=0$, then $i\in B\cup C$; and if $\mu+\lambda_{2i}-\lambda_{1i}\neq 0$, then $i\in A\cup B$; Lemmas 3.1 and 3.2 are valid; Theorem 3.3 is valid, except for the implication within brackets, meaning that Theorem 3.4 is not valid; furthermore, all additional results of Section 3 are not valid or need modification; Theorem 4.2 is obviously not affected by a changed K; and Corollary 4.2 is modified as follows.

Theorem 6.1. A point $g\in K$ is a solution of Problem (P), with K given as above, if and only if

$$\mu\geq 0, \qquad \mu(\mathrm{U}^t g - V)=0,$$

$$\lambda_1\geq 0, \qquad \lambda_1^t(g_1-g)=0,$$

$$\lambda_2\geq 0, \qquad \lambda_2^t(g-g_2)=0,$$

$$\mathrm{U}\mu+\lambda_2-\lambda_1=-P(g).$$

However, it does not seem possible to solve these equations in order to extend Corollary 4.3.

A third type of constraint is considered by Haslinger and Neittaanmäki (Ref. 8). It can be inferred from the argument that the slope of the obstacle (measured from some reference line) should not be too large. In the discrete case, it is given by

$$-c\leq g_{i+1}-g_i\leq c, \qquad 1\leq i\leq n-1,$$

where $c>0$ is a constant. The introduction of such a constraint has no essential implications for the discussion in this section.

7. Conclusions and Comments

In relation to the problem of optimizing contact force distributions, two different cost functionals have been considered. The resulting problems are investigated in detail, and it is shown under what conditions they actually coincide (i.e., give the same optimal design). Some comments concerning extensions and interpretations of the results can be given.

The optimization process is viewed as one where the obstacle is changed and the deformable body is unchanged. However, as noted in Ref. 3 and used also in Refs. 1, 2, and 5, since we are dealing with small displacements, g must be small for the formulation to make physical sense. Thus, the optimization process can be reinterpreted as one where the obstacle is fixed and the boundary of the deformable body is changed, but without affecting the stiffness. This conclusion makes clear that the two-body problem, discussed in Ref. 3, is covered by the present theory.

Note that contact forces are treated as primary variables in this work, which in the case of a naturally discrete body is the obvious choice. However, in the case of a continuous body that is discretized (e.g., by finite elements), in order to get the state problem of Section 2, it may not be natural to optimize contact force distributions, but rather contact stress distributions. For discretizations that can be shown to converge to the continuous problem, contact forces are obtained from contact stresses simply by multiplication by weighting factors W_i (Ref. 22). These functions are given by the numerical integration rule employed and represent in the simple case of linear finite elements the area surrounding contact nodes. The cost functional (16) can be replaced by

$$\bar{J}_M(g) = \max\{-P_i(g)/W_i \mid 1 \leq i \leq n\},$$

and the constraint (18) by

$$\sum_{i=1}^{n} g_i W_i \leq V.$$

One then sees that, by redefining the elements of A as $W_i W_j A_{ij}$ and those of w^* as $W_i w_i^*$, problems formally identical to (M) and (P) are obtained, but with forces replaced by stresses and g replaced by $G = \{g_i W_i\}$. Thus mathematically, there is essentially no difference between optimizing contact forces and optimizing contact stresses, and there is no qualitative difference between the theory obtained using contact stresses and that obtained using contact forces.

The results of this work reduce the computational effort needed to produce designs in practical situations to a minimum. If Assumption A is satisfied, both problems have coinciding solutions and the optimal gap is calculated simply by matrix multiplication. If Assumption A is not met, this is still true for Problem (P), while Problem (M) is solved by solution of a standard LP problem.

Related to problem (M) there are some open questions posed by this study. Firstly, what is the relation between the sufficient optimality condition of Theorem 3.2 and the necessary and sufficient one of Theorem 3.4? Secondly, does Theorem 3.6 hold even when Assumption A is not satisfied? By studying complementary cones (Ref. 15), one can convince oneself that, when $n = 2$, the answer to the latter question is in the affirmative. Since both of these questions are related to the existence of local minimizers, and since global minima can be determined without answering them, they are posed mainly from the point of view of theoretical curiosity.

References

1. CONRY, T. F., and SEIREG, A., *A Mathematical Programming Method for Design of Elastic Bodies in Contact*, Journal of Applied Mechanics, Vol. 2, pp. 387-392, 1971.
2. HAUG, E. J., and KWAK, B. M., *Contact Stress Minimization by Contour Design*, International Journal for Numerical Methods in Engineering, Vol. 12. pp. 917-930, 1978.
3. BENEDICT, R. L., and TAYLOR, J. E., *Optimal Design for Elastic Bodies in Contact*, Optimization of Distributed-Parameter Structures, Edited by E. J. Haug and J. Cea, Sijthoff and Noordhoff, Alphen aan den Rijn, Holland, pp. 1553-1599, 1981.
4. BENEDICT, R. L., *Maximum Stiffness Design for Elastic Bodies in Contact*, Journal of Mechanical Design, Vol. 104, pp. 825-830, 1982.
5. KIKUCHI, N., and TAYLOR, J. E., *Shape Optimization for Unilateral Elastic Contact Problems*, Numerical Methods in Coupled Problems, Proceeding of the International Conference at University College, Swansea, Wales, pp. 430-441, 1981.
6. BENEDICT, B., SOKOLOWSKI, J., and ZOLESIO, J. P., *Shape Optimization for Contact Problems*, System Modelling and Optimization, Edited by P. Thoft-Christensen, Springer-Verlag, Berlin, Germany, pp. 790-799, 1984.
7. BENDSØE, M. P., and SOKOLOWSKI, J., *Sensitivity Analysis and Optimal Design of Elastic Plates with Unilateral Point Supports*, Mechanics of Structures and Machines, Vol. 15, pp. 383-393, 1987.
8. HASLINGER, J., and NEITTAANMÄKI, P., *Finite-Element Approximation for Optimal Shape Design, Theory and Applications*, John Wiley and Sons, Chichester, England, 1988.
9. KLARBRING, A., *Quadratic Programs in Frictionless Contact Problems*, International Journal of Engineering Science, Vol. 24, pp. 1207-1217, 1986.
10. CLARKE, F. H., *Optimization and Nonsmooth Analysis*, John Wiley and Sons, New York, New York, 1983.
11. KIKUCHI, N., *A Class of Rigid Punch Problems Involving Forces and Moments by Reciprocal Variational Inequalities*, Journal of Structural Mechanics, Vol. 7, pp. 273-295, 1979.
12. DEMKOWICZ, L., *On Some Results Concerning the Reciprocal Formulation for the Signorini's Problem*, Computers and Mathematics with Applications, Vol. 8, pp. 57-75, 1982.
13. HASLINGER, J., and PANAGIOTOPOULOS, P. D., *The Reciprocal Variational Approach to the Signorini's Problem with Friction. Approximation Results*, Proceedings of the Royal Society of Edinburgh, Vol. 98A, pp. 365-383, 1984.

14. BENDSØE, M. P., and SOKOLOWSKI, J., *Design Sensitivity Analysis of Elastic-Plastic Analysis Problems*, Mechanics of Structures and Machines, Vol. 16, pp. 81-102, 1988.
15. MURTY, K. G., *Linear Complementarity, Linear and Nonlinear Programming*, Helderman-Verlag, Berlin, Germany, 1988.
16. COTTLE, R. W., *Monotone Solutions of the Parametric Complementarity Problem*, Mathematical Programming, Vol. 3, pp. 210-224, 1974.
17. IBARAKI, T., *Complementarity Programming*, Operations Research, Vol. 19, pp. 1523-1528, 1971.
18. KANEKO, I., *On Some Recent Engineering Applications of Complementarity Problems*, Mathematical Programming Study, Vol. 17, pp. 111-125, 1982.
19. HASLINGER, J., and ROUBÍČEK, T., *Optimal Control of Variational Inequalities. Approximation Theory and Numerical Realization*, Applied Mathematics and Optimization, Vol. 14, pp. 187-201, 1986.
20. CIARLET, P. G., *Introduction to Numerical Linear Algebra and Optimization*, Cambridge University Press, Cambridge, England, 1989.
21. EKELAND, I., and TEMAM, R., *Convex Analysis and Variational Problems*, North-Holland, Amsterdam, Holland, 1976.
22. KIKUCHI, N., and ODEN, J. T., *Contact Problems in Elasticity: A Study of Variational Inequalities and Finite-Element Methods*, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1988.