PERGAMON

International Journal of Solids and Structures 36 (1999) 2075–2090

INTERNATIONAL JOURNAL OF SOLIDS and STRUCTURES

A Newton method for three-dimensional fretting problems

Niclas Strömberg

Division of Mechanics, Department of Mechanical Engineering, Linköping Institute of Technology, S-581 83 Linköping, Sweden

Received 6 March 1997; in revised form 16 February 1998

Abstract

The present paper concerns the numerical treatment of fretting problems using a finite element analysis. The governing equations resulting from a formal finite element discretization of an elastic body with a potential contact surface are considered in a quasi-static setting. The constitutive equations of the potential contact surface are Signorini's contact conditions, Coulomb's law of friction and Archard's law of wear. Using a backward Euler time discretization and an approach based on projections, the governing equations are written as an augmented Lagrangian formulation which is implemented and solved using a Newton algorithm for three-dimensional fretting problems of didactic nature. Details concerning the implementation are provided.

1. Introduction

Fretting is a wear phenomenon occurring in contacts of machine elements subjected to oscillatory movements with small amplitudes. In steel assemblies, this phenomenon can be identified as oxide debris ($\alpha - Fe_2O_3$) called 'cocoa', 'red-mud' or 'blood', resulting from the particle detachment from the contact surfaces. This might lead to loss of clearance or cause jamming in assemblies such as splines, shrink-fits, bolted joints etc., see e.g. the reviews by Hurricks (1970), Waterhouse (1984) and Vincent et al. (1992). In severe cases, fretting might cause initial cracks leading to fretting fatigue failures in machine components, see e.g. the reviews by Waterhouse (1992) and Hills (1994).

Calculation of contact stresses in fretting is usually performed by applying a frictional analysis. Unfortunately, such an analysis does not include possible changes in the contact geometry caused by the particle detachment during the wear process. Since the distribution of contact stresses is very sensitive to the shape of the contact geometry, it might be of importance, especially when analysing fretting problems, to incorporate the material removal in the frictional analysis. This is done in the present paper; the change in contact geometry due to wear is included in the frictional analysis by defining a certain internal state variable. By using this internal variable approach, an elastic body with a potential contact surface possessing constitutive properties that describe contact, friction and wear is considered. For this system, quasi-static fretting problems of didactic nature

0020–7683/99/$ - see front matter
PII: S0020–7683(98)00080–8

are solved using a Newton algorithm. The structure of the algorithm is fitted for implementation in commercial finite element codes.

The constitutive equations of the potential contact surface are represented by a coupling of Signorini's contact conditions, Coulomb's law of friction and Archard's law of wear. The coupling of these tribological laws was given as a generalized standard material within a framework of continuum thermodynamics in Strömberg et al. (1996). In Strömberg (1997), inspired by Alart and Curnier (1991) and Klarbring (1992), this generalized standard material was rewritten in the form of an augmented Lagrangian formulation, which was solved for two-dimensional fretting problems using a Newton algorithm suggested by Pang (1990). In this work an improved version, from a numerical point of view, of this augmented Lagrangian formulation is considered for three-dimensional fretting problems. Due to the strong non-linearities of the tribological laws, the augmented Lagrangian formulation is not differentiable in the classical sense. However, it is Lipschitz continuous and directionally differentiable, which makes Pang's Newton algorithm for solving Bouligand differentiable equation systems applicable. Numerically, it is important to investigate the algorithm for three-dimensional problems and not only for the two-dimensional case. For instance, the directional derivative of the augmented Lagrangian formulation appearing in the algorithm simplifies a lot in two-dimensions compared to the three-dimensional case. Recently, augmented Lagrangian techniques for solving frictional contact problems were discussed by Wriggers (1996).

The essential idea of the fretting model utilized in this paper is to include wear in frictional contact models by defining a certain internal state variable for the contact interface (Strömberg et al., 1996). This internal state variable measures the wear developed in the contact interface as an increase in the gap between the bodies. The evolution law for the internal state variable is taken to be Archard's wear law (Archard, 1953), i.e. the increase in wear gap is proportional to the contact pressure and the amount of sliding between the bodies. This constitutive assumption is in agreement with e.g. Stowers and Rabinowicz (1973), who suggest that Archard's law of wear can be used in order to predict the volume worn away in fretting. Thus, compared to pure frictional contact models, this internal state variable induces an additional coupling between contact and friction owing to the fact that the contact geometry is changed when slip is developed. Recently, Faanes (1996) discussed how fretting surfaces might be modified by wear in order to improve the analysis of fretting fatigue cracks.

The Newton method discussed above is implemented and used to solve three-dimensional fretting problems of didactic nature. In particular, a punch subjected to a cyclic indentation into an elastic half-plane and a unilaterally constrained elastic block subjected to cyclic loads are considered. In contrast to the solutions of these problems using a pure frictional contact model, the contact pressures do not reach a steady state after the first cycle. But instead one gets, using this fretting model, an evolution in the contact pressures during the cycles depending on increases in the wear gap which are governed by Archard's wear law. On the contrary to intuition, depending on the stick-slip conditions for the particular problems considered, the wear gaps increase not in such a manner that uniform contact pressures are obtained; instead large contact stresses are developed in the sticking area (no sliding implies no wear, i.e. the wear gap is zero here) and zero contact stresses are obtained in the slipping area after many cycles (sliding implies wear, i.e. the wear gap is greater than zero here). In addition, large gradients of the contact stresses are produced in the intermediate location of stick and slip. This stress distribution is in major contrast to the uniform

distribution obtained for fretting with global sliding conditions. For two-dimensional problems, the contact tractions dependency on the removal of material in fretting was considered by Johansson (1992) and Strömberg (1997).

The non-uniform contact stresses, which are obtained in general for fretting problems with stick-slip conditions, might explain why initial cracks in fretting fatigue sometimes are observed in regions between stick and slip, see e.g. Waterhouse (1981) and the review by Mutoh (1995). A combination of large values and large gradients of the contact tractions might cause damages in the bulk material near the contact surface, which in turn might lead to initial cracks. In conclusion, concerning the fatigue life, fretting with stick-slip conditions might be a much more severe mechanism than fretting with global sliding conditions. Similar observations have been found experimentally by Vingsbo and Söderberg (1988).

The contents of this study is as follows: in Section 2 the quasi-static fretting problem is given. The governing equations defining the problem were derived from a continuum thermodynamical model using a formal space discretization in Strömberg (1997). In Section 3 the main steps of the derivation of the augmented Lagrangian formulation from the governing equations are briefly presented. The Newton algorithm used for solving this system is also given. In Section 4 details about the numerical implementation of the algorithm are provided. In Section 5 the problems discussed above are considered as well as a comparison of three different punches pressed into an elastic half-plane showing the sensitivity in the contact pressure with respect to changes in the contact geometry. Finally, in Section 6, concluding remarks are given.

2. The quasi-static fretting problem

Let us consider an elastic body for which a potential contact surface is unilaterally constrained by a rigid foundation. The potential contact surface is treated as a material surface which possesses constitutive properties that describe contact, friction and wear. Besides the well-known internal state variable used in friction problems for measuring slip (see e.g. Michalowski and Mróz, 1978; Curnier, 1984; Klarbring, 1990), an additional internal state variable is defined for this material surface which measures the wear gap at the contact interface. Within the framework of this internal state variable approach (see Strömberg et al., 1996), the fretting laws are given by Signorini's contact conditions, Coulomb's law of friction and Archard's law of wear. Concerning the first formulation of a contact interface within the framework of continuum thermodynamics, at least to my knowledge, see Frémond (1987).

Following Strömberg (1997), the above system is discretized in space. The total number of kinematic freedoms is η_u and the number of potential contact nodes is η_c. The displacement vector is represented by $\mathbf{u} \in \Re^{\eta_u}$. For each contact node $i \in \{1, \ldots, \eta_c\}$, three unit directions are defined: one normal direction and two tangential directions with respect to the support. Using these unit directions, it is possible to express the normal displacement and the tangential displacements at each contact node in the following way:

$$u_{in} = \mathbf{C}_{in}\mathbf{u} \in \Re, \quad u_{it} = \mathbf{C}_{it}\mathbf{u} \in \Re, \quad u_{io} = \mathbf{C}_{io}\mathbf{u} \in \Re,$$

where subscripts n, t and o stand for normal, tangential and orthogonal direction, respectively,

and $\mathbf{C}_{in}$, $\mathbf{C}_{it}$ and $\mathbf{C}_{io}$ are the ith rows of the transformation matrices $\mathbf{C}_n \in \Re^{\eta_c \times \eta_u}$, $\mathbf{C}_t \in \Re^{\eta_c \times \eta_u}$ and $\mathbf{C}_o \in \Re^{\eta_c \times \eta_u}$, respectively.

After a formal discretization, the equilibrium equations can be written as

$$\mathbf{K}\mathbf{u} = \mathbf{F} - \mathbf{C}_n^T \mathbf{P}_n - \mathbf{C}_t^T \mathbf{P}_t - \mathbf{C}_o^T \mathbf{P}_o, \tag{1}$$

where $\mathbf{K} \in \Re^{\eta_u \times \eta_u}$ is the stiffness matrix, $\mathbf{F} \in \Re^{\eta_u}$ is the external nodal force vector, $\mathbf{P}_n \in \Re^{\eta_c}$ is the normal contact force vector, and $\mathbf{P}_t \in \Re^{\eta_c}$ and $\mathbf{P}_o \in \Re^{\eta_c}$ are the tangential contact forces. In this work, it is assumed that the stiffness matrix $\mathbf{K}$ is positive semi-definite, i.e. the body is not necessarily fixed at the boundary such that rigid body motions are prevented.

In order to account for wear in the normal contact law, Signorini's unilaterally contact conditions are extended in the following manner for $i \in \{1, \ldots, \eta_c\}$ (Strömberg et al., 1996):

$$P_{in} \in \Re_+ : \quad (u_{in} - w_i - g_i)(P'_{in} - P_{in}) \leqslant 0 \quad \forall P'_{in} \in \Re_+, \tag{2}$$

where w_i is the nodal value of the internal state variable measuring the wear gap, g_i is the initial gap between the contact node and the support, $\Re_+$ is the non-negative orthant of $\Re$ (i.e. $\Re_+ = \{x \in \Re : x \geqslant 0\}$) and P'_{in} stands for a 'trial' contact force.

For each contact node, Coulomb's law of friction is given as the following well-known maximum dissipation principle: find $(P_{it}, P_{io}) \in \mathscr{F}(P_{in})$ such that

$$\dot{u}_{it}(P'_{it} - P_{it}) + \dot{u}_{io}(P'_{io} - P_{io}) \leqslant 0 \quad \forall (P'_{it}, P'_{io}) \in \mathscr{F}(P_{in}), \tag{3}$$

where

$$\mathscr{F}(P_{in}) = \{(P_{it}, P_{io}) : \sqrt{(P_{it})^2 + (P_{io})^2} \leqslant \mu (P_{in})_+\}$$

is Coulomb's set of admissible tangential contact forces. Here, it is assumed that the friction coefficient μ is constant over the potential contact surface. Furthermore, a superimposed dot stands for the time derivative and $(x)_+ = \max(0, x)$, where the 'max'-operator denotes maximum of the two arguments. By using this latter function in $\mathscr{F}(P_{in})$, one obtains a set of admissible tangential contact forces which is defined not only for $P_{in} \geqslant 0$ but also for $P_{in} < 0$.

Finally, the evolution of the internal state variable w_i is governed by Archard's law of wear. That is, assuming that k is a constant wear coefficient,

$$\dot{w}_i = \frac{k}{I_i} P_{in} \sqrt{(\dot{u}_{it})^2 + (\dot{u}_{io})^2}, \quad i \in \{1, \ldots, \eta_c\}, \tag{4}$$

where I_i is a weighting factor depending on what quadrature rule has been used in the space discretization of the contact surface. This weighting factor appears naturally in Archard's wear law when the discretization is carried out, see Strömberg (1997). If Archard's wear coefficient k_a and the penetration hardness p_s are defined, then the following relationship holds:

$$k = \frac{k_a}{3p_s}.$$

The fretting problem has now been stated in eqns (1)–(4). That is, for a given load history $\mathbf{F}(t)$, $t \in [0, T]$, on a time interval $[0, T]$, the following quasi-static problem is to be solved: find $t \mapsto \mathbf{u}(t)$,

$t \mapsto \mathbf{P}_n(t)$, $t \mapsto \mathbf{P}_t(t)$ and $t \mapsto \mathbf{P}_o(t)$ such that the functions $\mathbf{u}(t)$, $\mathbf{P}_n(t)$ and $\mathbf{P}_o(t)$ satisfy conditions (1)–(4).

3. The Newton algorithm

In order to solve the quasi-static problem stated above, we start to approximate the time rates appearing in (3) and (4), at a time t_{k+1}, by the following implicit rule:

$$(\dot{u}_{it}(t_{k+1}), \dot{u}_{io}(t_{k+1})) \approx \left(\frac{u_{it}(t_{k+1}) - u_{it}(t_k)}{t_{k+1} - t_k}, \frac{u_{io}(t_{k+1}) - u_{io}(t_k)}{t_{k+1} - t_k} \right), \tag{5}$$

$$\dot{w}_i(t_{k+1}) \approx \frac{w_i(t_{k+1}) - w_i(t_k)}{t_{k+1} - t_k}, \tag{6}$$

i.e. a backward Euler time discretization is applied. Using (5) and (6), and defining $(\bar{u}_{it}, \bar{u}_{io}) = (u_{it}(t_{k+1}) - u_{it}(t_k), u_{io}(t_{k+1}) - u_{io}(t_k))$, one can approximate (4) by

$$w_i(t_{k+1}) \approx \bar{w}_i + \frac{k}{I_i} P_{in}(t_{k+1}) \sqrt{(\bar{u}_{it})^2 + (\bar{u}_{io})^2}, \tag{7}$$

where $\bar{w}_i = w_i(t_k)$.

Next, multiplying (2) by any $r_i > 0$, and adding and subtracting this by P_{in}, it is recognised that P_{in} is the projection of $(P_{in} + r_i(u_{in} - w_i - g_i))$ onto $\mathfrak{R}_+$. In a similar way, it can be verified from (3), on using (5), that (P_{it}, P_{io}) can be expressed as the projection of $(P_{it} + r_i\bar{u}_{it}, P_{io} + r_i\bar{u}_{io})$ onto $\mathscr{F}(P_{in})$ (concerning the approach of projections, see De Saxce and Feng, 1991 and Klarbring, 1992).

Using these projections together with (1) and (7), and defining $\mathbf{y} = (\mathbf{u}, \mathbf{P}_n, \mathbf{P}_t, \mathbf{P}_o)$, the fretting problem can now be treated by solving the following system of equations for each time increment:

$$\mathbf{H}(\mathbf{y}) = \begin{bmatrix} \mathbf{K}\mathbf{u} - \mathbf{F} + \mathbf{C}_n^T\mathbf{P}_n + \mathbf{C}_t^T\mathbf{P}_t + \mathbf{C}_o^T\mathbf{P}_0 \\ (-P_{in} + \Pi_{in}(\mathbf{y}))_{i=1}^{i=\eta_c} \\ \left(-\begin{pmatrix} P_{it} \\ P_{io} \end{pmatrix} + (\mathbf{\Pi}_{ito}(\mathbf{y}))^T \right)_{i=1}^{i=\eta_c} \end{bmatrix} = \mathbf{0}, \tag{8}$$

where, on using the definition $(P_{it}(r_i), P_{io}(r_i)) = (P_{it} + r_i\bar{u}_{it}, P_{io} + r_i\bar{u}_{io})$,

$$\Pi_{in}(\mathbf{y}) = \left(P_{in} + r_i \left(u_{in} - \bar{w}_i - \frac{k}{I_i} P_{in} \sqrt{(\bar{u}_{it})^2 + (\bar{u}_{io})^2} - g_i \right) \right)_+, \tag{9}$$

$$\mathbf{\Pi}_{ito}(\mathbf{y}) = \begin{cases} (P_{it}(r_i), P_{io}(r_i)) & \text{if } (P_{it}(r_i), P_{io}(r_i)) \in \mathscr{F}(P_{in}) \\ \dfrac{\mu(P_{in})_+ (P_{it}(r_i), P_{io}(r_i))}{\sqrt{(P_{it}(r_i))^2 + (P_{io}(r_i))^2}} & \text{otherwise} \end{cases} \tag{10}$$

are the explicit expressions of the projections of $(P_{in} + r_i(u_{in} - w_i - g_i))$ onto $\mathfrak{R}_+$ and $(P_{it}(r_i), P_{io}(r_i))$ onto $\mathscr{F}(P_{in})$, respectively. Notice also here that the approximation of Archard's law of wear given in (7) has been inserted into (9). The functions given in (9) and (10) are Lipschitz continuous and

directionally differentiable, see Strömberg (1997). Consequently, this is also true for $\mathbf{H}(\mathbf{y})$ in (8). The Lipschitz continuity and the directionally differentiability imply that $\mathbf{H}(\mathbf{y})$ is Bouligand differentiable. For solving such a system of equations, Pang (1990) suggested the Newton algorithm presented below.

A more detailed derivation of the augmented Lagrangian system in (8) and the projections in (9) and (10) can be found in Strömberg (1997), where an almost similar augmented Lagrangian system was solved for two-dimensional fretting problems using Pang's Newton algorithm. The difference between this augmented Lagrangian system and the one suggested previously concerns the equilibrium equations in (8). In the previous work (Strömberg, 1997), inspired by Alart and Curnier (1991), the contact forces appearing in the equilibrium equations were replaced by the projections in (9) and (10). However, the formulation used in (8) is a numerical improvement since the equilibrium equations remain linear in the contact forces. This is of advantage when the search direction is found in Pang's algorithm, see (11) and also Section 4.2.

So far, it has been numerically proven, for the two-dimensional case, that Pang's algorithm works well for solving augmented Lagrangian formulations of friction and fretting problems. An extensive study of this Newton algorithm for two-dimensional friction problems can be found in Christensen et al. (1997), where also friction problems were solved using an interior point method; it was found that the Newton method is superior both in speed and robustness. In this paper, the augmented Lagrangian formulation in (8) is solved for three-dimensional fretting problems using Pang's algorithm, i.e.

Algorithm: Let β, γ and ε be given scalars with $\beta\in(0,1)$, $\gamma\in(0,1/2)$ and ε small. Repeat the following steps for each time increment $k+1$:

(0) Let $\mathbf{y}^0$ be given from the previous time step k and set $l=0$.
(1) Find a search direction $\mathbf{z}=(\mathbf{z}_d,\mathbf{z}_n,\mathbf{z}_t,\mathbf{z}_o)$ such that

$$\mathbf{H}(\mathbf{y}^l)+\mathbf{H}'(\mathbf{y}^l;\mathbf{z})=\mathbf{0}, \tag{11}$$

where $\mathbf{H}'(\mathbf{y}^l;\mathbf{z})$ is the directional derivative.
(2) Let $\alpha^l=\beta^{m_l}$, where m_l is the first non-negative integer m for which the following decrease criterion holds:

$$g(\mathbf{y}^l+\beta^m\mathbf{z})\leqslant(1-2\gamma\beta^m)g(\mathbf{y}^l),\quad g(\mathbf{y})=\tfrac{1}{2}\mathbf{H}^T(\mathbf{y})\mathbf{H}(\mathbf{y}).$$

(3) Set $\mathbf{y}^{l+1}=\mathbf{y}^l+\alpha^l\mathbf{z}$.
(4) If $g(\mathbf{y}^{l+1})\leqslant\varepsilon$, then terminate with $\mathbf{y}^{l+1}$ as an approximate zero of $\mathbf{H}(\mathbf{y})$. Otherwise, replace l by $l+1$ and return to step 1.

After convergence in each time step is achieved, the internal state variables u_{it}, u_{io} and w_i are updated.

The algorithm's rate of convergence cannot be established, nor can it be shown that the stepsize α^l eventually will become 1. However, if $\mathbf{H}(\mathbf{y})$ is differentiable in the classical sense, i.e. Fréchet differentiable, at the solution point, then it is possible to show quadratic convergence. Furthermore, the line-search in step two of the algorithm, an Armijo procedure, makes the algorithm globally convergent. Concerning convergence results of the algorithm see Pang (1990) and Christensen et al. (1997). A practical drawback of the algorithm is the non-linearity of $\mathbf{H}'(\mathbf{y};\mathbf{z})$ in $\mathbf{z}$, which might

limit the effectiveness of finding search directions, at least for large scale problems. However, the absence of linearity is overcome by introducing an appropriate substitution of $\mathbf{H}'(\mathbf{y};\mathbf{z})$ such that (11) results in a system of linear equations, see Section 4.1.

4. Implementation

The Newton method presented is implemented in Fortran 77 for the three-dimensional case. Two different types of examples, given in Section 5, are solved on an HP 9000 Model 712 work station with 60 MHz clock speed using double precision arithmetic. In this section important information concerning the implementation of the algorithm is provided.

4.1. Substitution of $\boldsymbol{H'(y;z)}$

The essential task of the algorithm is to find the search direction, i.e. solving (11) in step one of the algorithm which generally is a system of non-linear equations. In order to get an efficient code, this task is treated by introducing an appropriate substitution of $\mathbf{H}'(\mathbf{y};\mathbf{z})$ such that (11) results in a system of linear equations which still yield an effective search direction.

The non-linearity of $\mathbf{H}'(\mathbf{y};\mathbf{z})$ in $\mathbf{z}$ depends on the non-differentiability of the functions given in (9) and (10). However, these functions are only non-differentiable at states on certain boundaries; otherwise they are differentiable. The substitution of $\mathbf{H}'(\mathbf{y};\mathbf{z})$ concerns states on such boundaries of non-differentiability. In order to deal with the substitution of $\mathbf{H}'(\mathbf{y};\mathbf{z})$ on these boundaries, the following index sets are defined:

$$\mathscr{I}_< = \left\{i: P_{in} + r_i\left(u_{in} - \bar{w}_i - \frac{k}{I_i}P_{in}\sqrt{(\bar{u}_{it})^2 + (\bar{u}_{io})^2} - g_i\right) < 0\right\},$$

$$\mathscr{I}_= = \left\{i: P_{in} + r_i\left(u_{in} - \bar{w}_i - \frac{k}{I_i}P_{in}\sqrt{(\bar{u}_{it})^2 + (\bar{u}_{io})^2} - g_i\right) = 0\right\},$$

$$\mathscr{I}_> = \left\{i: P_{in} + r_i\left(u_{in} - \bar{w}_i - \frac{k}{I_i}P_{in}\sqrt{(\bar{u}_{it})^2 + (\bar{u}_{io})^2} - g_i\right) > 0\right\},$$

$$\mathscr{J} = \{i: P_{in} < 0\},$$

$$\mathscr{J}_< = \{i: P_{in} > 0, \sqrt{(P_{it}(r_i))^2 + (P_{io}(r_i))^2} < \mu P_{in}\},$$

$$\mathscr{J}_= = \{i: P_{in} > 0, \sqrt{(P_{it}(r_i))^2 + (P_{io}(r_i))^2} = \mu P_{in}\},$$

$$\mathscr{J}_> = \{i: P_{in} > 0, \sqrt{(P_{it}(r_i))^2 + (P_{io}(r_i))^2} > \mu P_{in}\},$$

$$\mathscr{K}_+ = \{i: P_{in} = 0 < \sqrt{(P_{it}(r_i))^2 + (P_{io}(r_i))^2}\},$$

$$\mathscr{K}_0 = \{i: P_{in} = P_{it}(r_i) = P_{io}(r_i) = 0\},$$

$$\mathscr{W}_+ = \{i: \sqrt{(\bar{u}_{it})^2 + (\bar{u}_{io})^2} > 0\},$$

$$\mathscr{W}_0 = \{i: \sqrt{(\bar{u}_{it})^2 + (\bar{u}_{io})^2} = 0\}.$$

That is, although the directional derivatives $\Pi'_{in}(\mathbf{y};\mathbf{z})$ and $\mathbf{\Pi}'_{ito}(\mathbf{y};\mathbf{z})$ always exist, $\Pi_{in}(\mathbf{y})$ is not differentiable for $i\in\mathscr{I}_{=}\cup(\mathscr{I}_{>}\cap\mathscr{W}_0)$ and $\mathbf{\Pi}_{ito}(\mathbf{y})$ is not differentiable for $i\in\mathscr{J}_{=}\cup\mathscr{K}_{+}\cup\mathscr{K}_0$. Consequently, $\mathbf{H}(\mathbf{y})$ is not differentiable when the index set $\mathscr{L}=\mathscr{I}_{=}\cup\mathscr{J}_{=}\cup\mathscr{K}_{+}\cup\mathscr{K}_0\cup(\mathscr{I}_{>}\cap\mathscr{W}_0)$ is non-empty; otherwise $\mathbf{H}'(\mathbf{y};\mathbf{z})=\nabla\mathbf{H}(\mathbf{y})\mathbf{z}$. Furthermore, if $i\notin\mathscr{L}$, then $\Pi'_{in}(\mathbf{y};\mathbf{z})=\nabla\Pi_{in}(\mathbf{y})\mathbf{z}$ and $\mathbf{\Pi}'_{ito}(\mathbf{y};\mathbf{z})=\nabla\mathbf{\Pi}_{ito}(\mathbf{y})\mathbf{z}$. Still, $\nabla\Pi_{in}(\mathbf{y})$ and $\nabla\mathbf{\Pi}_{ito}(\mathbf{y})$ depend on the state $\mathbf{y}$. This is dealt with the remaining sets given above, apart from the sets defining $\mathscr{L}$.

Using this approach of index sets, the following substitution is introduced:

$$\mathbf{H}'(\mathbf{y};\mathbf{z})\approx\mathbf{J}(\mathbf{y})\mathbf{z}=\begin{bmatrix}\mathbf{K}\mathbf{z}_d+\mathbf{C}_n^T\mathbf{z}_n+\mathbf{C}_t^T\mathbf{z}_t+\mathbf{C}_o^T\mathbf{z}_o\\(-z_{in})_{i\in\mathscr{I}_{<}\cup\mathscr{I}_{=}}\\(r_i\mathbf{C}_{in}\mathbf{z}_d)_{i\in\mathscr{I}_{>}\cap\mathscr{W}_0}\\\left(r_i\left(\mathbf{C}_{in}\mathbf{z}_d-\dfrac{k\|\mathbf{u}\|_i}{I_i}\left(z_{in}+\dfrac{P_{in}(\bar{u}_{it}\mathbf{C}_{it}\mathbf{z}_d+\bar{u}_{io}\mathbf{C}_{io}\mathbf{z}_d)}{\|\mathbf{u}\|_i^2}\right)\right)\right)_{i\in\mathscr{I}_{>}\cap\mathscr{W}_{+}}\\\begin{pmatrix}-z_{it}\\-z_{io}\end{pmatrix}_{i\in\mathscr{J}\cup\mathscr{K}_{+}\cup\mathscr{K}_0}\\\begin{pmatrix}r_i\mathbf{C}_{it}\mathbf{z}_d\\r_i\mathbf{C}_{io}\mathbf{z}_d\end{pmatrix}_{i\in\mathscr{J}_{<}\cup\mathscr{J}_{=}}\\\left(-\begin{pmatrix}z_{it}\\z_{io}\end{pmatrix}+\dfrac{\mu z_{in}(P_{it}(r_i),P_{io}(r_i))^T}{\sqrt{(P_{it}(r_i))^2+(P_{io}(r_i))^2}}+\mathbf{R}_i\begin{pmatrix}z_{it}+r_i\mathbf{C}_{it}\mathbf{z}_d\\z_{io}+r_i\mathbf{C}_{io}\mathbf{z}_d\end{pmatrix}\right)_{i\in\mathscr{J}_{>}}\end{bmatrix},\tag{12}$$

where $\|\mathbf{u}\|_i=\sqrt{(\bar{u}_{it})^2+(\bar{u}_{io})^2}$ and

$$\mathbf{R}_i=\frac{\mu P_{in}}{((P_{it}(r_i))^2+(P_{io}(r_i))^2)^{3/2}}\begin{pmatrix}(P_{io}(r_i))^2 & -P_{it}(r_i)P_{io}(r_i)\\-P_{it}(r_i)P_{io}(r_i) & (P_{it}(r_i))^2\end{pmatrix}.$$

Thus, the substitution concerns $i\in\mathscr{L}$; otherwise exact equality holds in the above expression. The correct expression of the directional derivative $\mathbf{H}'(\mathbf{y};\mathbf{z})$ for the pure frictional case, i.e. $k=0$ in (12), can be found in Christensen et al. (1997).

The basic idea with the above substitution is that the index set $\mathscr{L}$ is numerically almost always empty and therefore the above substitution is sufficient to get a vigorous search direction. Accordingly, (11) reduces to

$$\mathbf{z}=-(\mathbf{J}(\mathbf{y}^l))^{-1}\mathbf{H}(\mathbf{y}^l),\tag{13}$$

where the Jacobian $\mathbf{J}(\mathbf{y}^l)$ is defined by (12). In particular, if $\mathscr{L}$ is empty, then $\mathbf{J}(\mathbf{y}^l)=\nabla\mathbf{H}(\mathbf{y}^l)$. In conclusion, by introducing the substitution given in (12), (11) becomes a system of linear equations instead of a system of non-linear equations.

4.2. Linear equation solving

When implementing the algorithm, assuming that $\mathbf{J}(\mathbf{y}^l)$ always is non-singular, eqn (13) is solved by a lower and upper factorization, i.e.

$$\mathbf{J}(\mathbf{y}^l) = \mathbf{L}(\mathbf{y}^l)\mathbf{U}(\mathbf{y}^l), \tag{14}$$

where $\mathbf{L}(\mathbf{y}^l)$ is a lower triangular matrix and $\mathbf{U}(\mathbf{y}^l)$ is an upper triangular matrix.

Generally, this factorization step is performed with a pivoting technique using a package from Netlib.[1] In particular, if the stiffness matrix $\mathbf{K}$ is positive definite, then it is utilized that some parts of $\mathbf{J}(\mathbf{y}^l)$ are constant. That is, letting $\mathbf{z}_c = (\mathbf{z}_n, \mathbf{z}_t, \mathbf{z}_o)$, (14) is decomposed in the following manner:

$$\begin{bmatrix} \mathbf{J}_{dd} & \mathbf{J}_{dc} \\ \mathbf{J}_{cd}(\mathbf{y}^l) & \mathbf{J}_{cc}(\mathbf{y}^l) \end{bmatrix} = \begin{bmatrix} \mathbf{L}_{dd} & \mathbf{0} \\ \mathbf{L}_{cd}(\mathbf{y}^l) & \mathbf{L}_{cc}(\mathbf{y}^l) \end{bmatrix} \begin{bmatrix} \mathbf{U}_{dd} & \mathbf{U}_{dc} \\ \mathbf{0} & \mathbf{U}_{cc}(\mathbf{y}^l) \end{bmatrix},$$

where $\mathbf{J}_{dd}$ and $\mathbf{J}_{dc}$ are constant parts of $\mathbf{J}(\mathbf{y}^l)$ belonging to the equilibrium equations in (8), and $\mathbf{J}_{cd}(\mathbf{y}^l)$ and $\mathbf{J}_{cc}(\mathbf{y}^l)$ are parts depending on the state $\mathbf{y}^l$ from the tribological laws given in (9) and (10). Thus, in the initialization of the algorithm, $\mathbf{L}_{dd}$, $\mathbf{U}_{dd}$ and $\mathbf{U}_{dc}$ are determined from $\mathbf{J}_{dd} = \mathbf{L}_{dd}\mathbf{U}_{dd}$ and $\mathbf{J}_{dc} = \mathbf{L}_{dd}\mathbf{U}_{dc}$. Then, in each iteration, (14) is solved by using the following steps:

$$\mathbf{J}_{cd}(\mathbf{y}^l) = \mathbf{L}_{cd}(\mathbf{y}^l)\mathbf{U}_{dd} \quad \text{and} \quad \mathbf{J}_{cc}(\mathbf{y}^l) = \mathbf{L}_{cd}(\mathbf{y}^l)\mathbf{U}_{dc} + \mathbf{L}_{cc}(\mathbf{y}^l)\mathbf{U}_{cc}(\mathbf{y}^l).$$

In addition, in the initialization step of the algorithm, the equilibrium equations are statically condensed to the contact nodes using a Cholesky decomposition, such that, assuming that no contact nodes are partially constrained, $\mathbf{J}_{dd}$, $\mathbf{J}_{dc}$, $\mathbf{J}_{cd}(\mathbf{y}^l)$, $\mathbf{J}_{cc}(\mathbf{y}^l) \in \Re^{3\eta_c \times 3\eta_c}$. The above scheme is owing to the improved augmented Lagrangian formulation given in this paper. Previously, in Strömberg (1997), $\mathbf{J}_{dd}$ and $\mathbf{J}_{dc}$ were not constant but depended instead on the tribological laws given in (9) and (10). Finally, another improvement for solving (13) can be achieved by utilizing the structure of (12) when $i \in \mathscr{I}_< \cup \mathscr{I}_=$ and/or $i \in \mathscr{J} \cup \mathscr{K}_+ \cup \mathscr{K}_0$.

4.3. Miscellaneous

Parameters entered into the line-search procedure have been set to $\beta = 0.9$ and $\gamma = 0.1$. The performance of the line search procedure turns out well with this setting. However, it is of major importance to define an additional upper bound on m such that the line search is aborted if β^m gets numerically too small. Otherwise, it is likely that the algorithm will stall. In the implementation, this upper bound is set to 22, i.e. β^m never becomes less than 0.1. Concerning the termination check, ε is set equal to 10^{-2}. The remaining parameters to be set in the algorithm are the penalty terms r_i appearing in the tribological laws given by (9) and (10). In this work, each r_i is set equal to $10^{12}I_i$. Still, r_i can be chosen in a range between $10^9 I_i - 10^{13} I_i$ without any major differences in performance of the algorithm. However, if r_i is set to a lower value, then convergence difficulties may appear. On the contrary, if r_i is set larger, then the algorithm might be unstable.

5. Numerical examples

Two different types of three-dimensional problems are considered. First rigid punches with different shapes are pressed into an elastic half-plane and then a unilaterally constrained elastic

[1] http://www.netlib.org/ieeecss/cascade/dgefam.f and /dgeslm.f

block is subjected to cyclic loads. In each problem, the elastic structure is approximated by trilinear hexahedral elements (see e.g. Hughes, 1987) with Young's modulus and Poisson's ratio taken to be 210 GPa and 0.3, respectively. The potential contact surface is, in all examples, approximated by 100 contact elements using the trapezoidal integration rule in such a manner that the integration points coincide with the nodal displacement points of the contact surface. It is a matter of course to choose other integration rules such that the integration points not necessary will coincide with the nodal displacements. A study of different integration rules for the contact surface can be found in Johansson (1986).

5.1. Punch problems

A rigid punch with three different geometries is pressed into an elastic half-plane. The symmetry is utilized such that only a quarter of the punch and the half-plane is modelled. The elastic half-plane is approximated by a finite structure according to Fig. 1. In order to get a good approximation of an elastic half-plane, the finite structure is taken sufficiently large with respect to the potential contact surface. That is, the dimension of the finite structure is $1.0 \times 1.0 \times 1.0$ m^3 and the area of the potential contact surface is 0.1×0.1 m^2. The finite structure is modelled by $20 \times 20 \times 20$ finite elements and for the potential contact surface 10×10 elements are used. The friction coefficient is set to $\mu = 0.3$ and the wear coefficient is taken to be $k = 1.0 \cdot 10^{-11}$ Pa^{-1}. In steel assemblies, this is typical wear constant for adhesive wear which normally is the most dominating wear mechanism in fretting situations (see e.g. Rabinowicz, 1965).

The different punch geometries considered are given by the following initial gaps: $g = 5.0 \cdot 10^{-3}(o^2 + t^2)$ m, $g = 0.75(o^2 + t^2)^2$ m and $g = 0.75(o^4 + t^4)$ m, where the tangential coor-

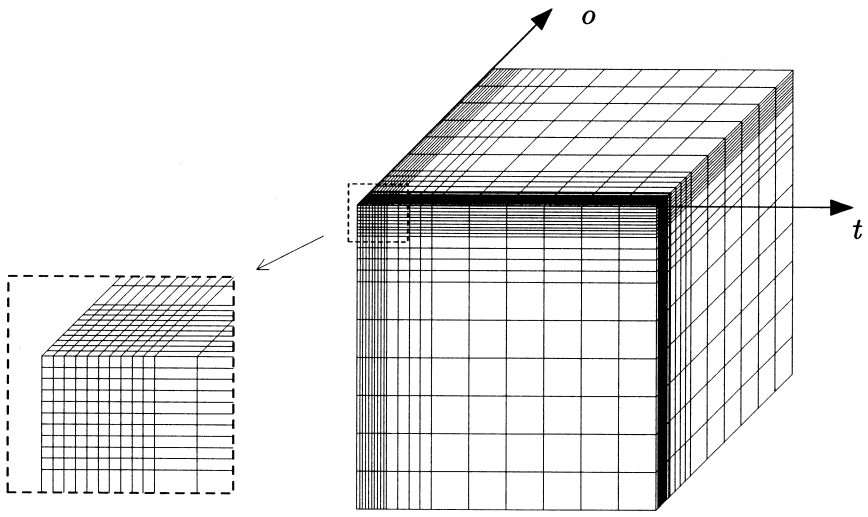

Fig. 1. The finite element approximation of an elastic half-plane, showing the mesh and the symmetry used. The potential contact surface, approximated by 100 contact elements, is zoomed in.

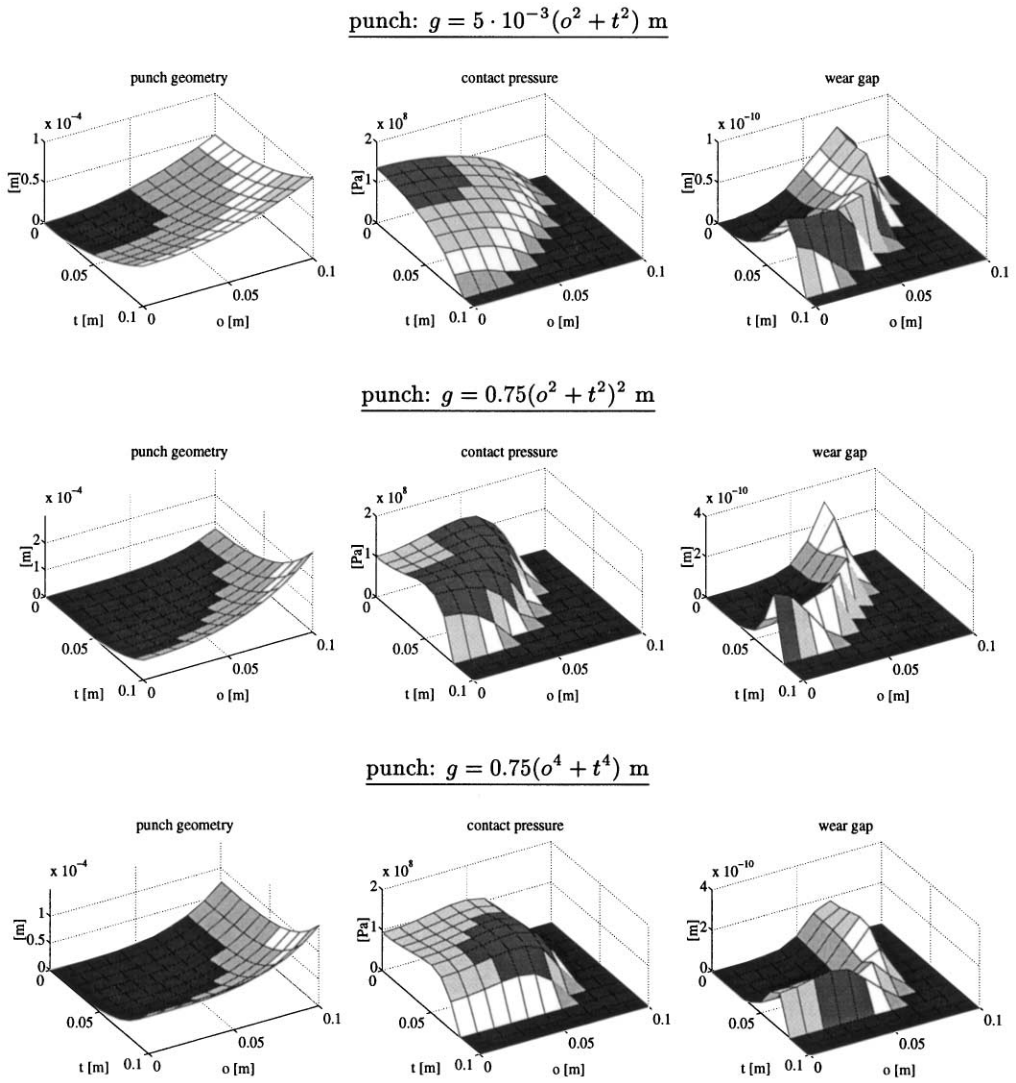

Fig. 2. The punch geometry, the contact pressure and the wear gap for three different punches pressed into an elastic half-plane.

dinates o and t are defined in Fig. 1. These geometries are compared in Fig. 2. The first punch geometry represents a Hertz punch with a radius of 100 m, while the two other punch geometries are only slight modifications of this one. In Fig. 2, the contact pressure and the wear gap are plotted after each punch is indented 0.07 mm. Each indentation is treated using 40 time increments. A comparison indeed emphasizes the sensitivity in the contact pressure with respect to changes in the structure, boundary conditions and initial conditions. It is obvious how small changes from the first punch geometry strongly affect the shape of the contact pressure. Furthermore, it is also obvious that the wear gap strongly depends on the punch geometry (look at the scales on the axis).

This sensitivity depends not only on the differences obtained in the contact pressure but also on the differences in the amount of slip, as explained by Archard's wear law in (4). It is important to notice that the sliding has appeared solely because of the 0.07 mm indentation and not because of any prescribed global tangential movement, i.e. the amount of slip is the measured micro-slip of Coulomb depending on the elasticity of the half-plane. In conclusion, when studying a fretting phenomenon it is of importance to define the geometry, the boundary conditions and the initial conditions of the problem carefully, and not only to choose the friction and wear coefficients with great precision.

Thus, from the discussion above, it is clear that in a fretting process the evolution in wear gap might change the surface geometry so that large changes in the contact stresses as well as in the amount of slip are induced, which in turn influence the wear rate. This fact is illustrated in Fig. 3, where the third punch, i.e. $g = 0.75(o^4 + t^4)$ m, is pressed cyclic into the elastic half-plane. The punch is repeatedly indented 0.07 mm during 500 cycles using 20 time increments for each cycle. The friction coefficient is still $\mu = 0.3$ but the coefficient of wear is taken to be $k = 1.0 \cdot 10^{-8}$ Pa^{-1}. After 500 cycles, the solution reaches a steady state. In the slipping region, i.e. where wear is developed, the contact pressure has turned toward a value of zero. On the contrary, a hump in the contact pressure has been developed in the sticking area with almost a four times greater maximum value than the maximum after the first cycle. The non-uniform distribution of the contact pressure depends on the stick-slip conditions.

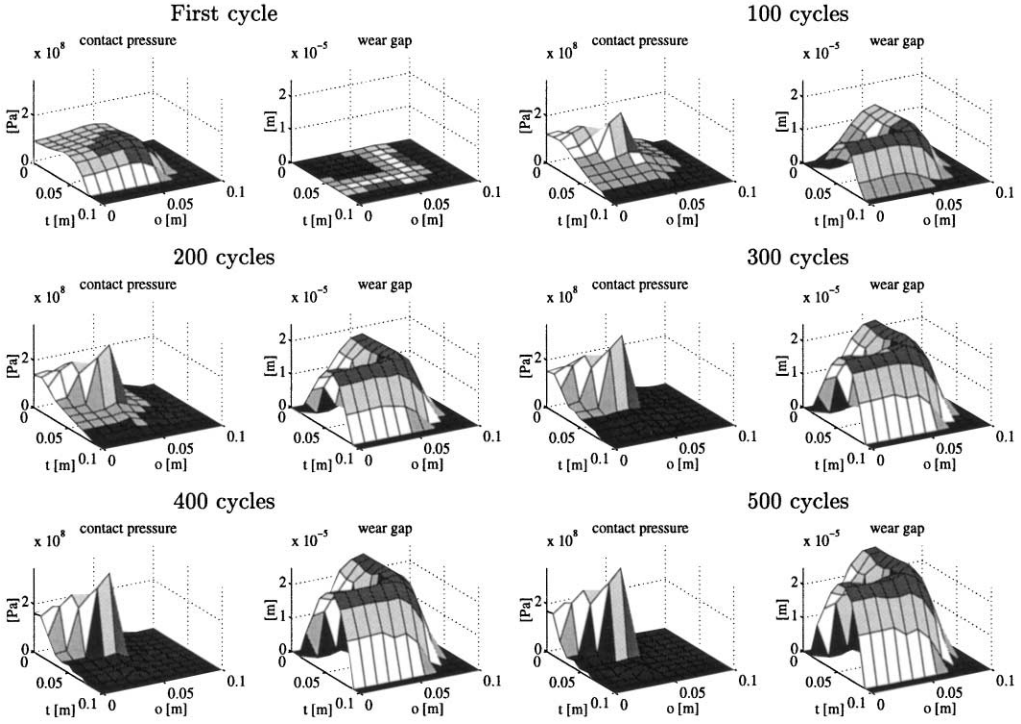

Fig. 3. The evolution in contact pressure and wear gap during repeated indentations of the punch represented by $g = 0.75(o^4 + t^4)$ m.

The wear coefficient used in the above example is unrealistically large from a physical point of view. However, this wear coefficient is used in order to numerically accelerate the fretting process. One might consider the result in Fig. 3 as a solution of a similar problem during 500,000 cycles with $k = 1.0 \cdot 10^{-11}$ Pa^{-1}. This with a small error compared to the uncertainties in the experimentally determined value of the wear coefficient. In such a manner, it is possible to solve real fretting problems within a reasonable time. The size of the error induced by decreasing the number of cycles and simultaneously increasing the wear coefficient was investigated in more detail by Johansson (1992) and Strömberg (1997).

5.2. Elastic block problem

The structures involved in real fretting problems rarely consist of elastic half-planes but rather of structures with finite extensions. Of course, using the elastic half-plane assumption is a first-rate approach for many problems, but not generally. This justifies the use of a finite element analysis for solving fretting problems.

A finite elastic block, with dimensions $0.2 \times 0.2 \times 0.1$ m^3, is considered in Fig. 4. It is unilaterally constrained by a rigid foundation at the bottom and subjected to surface loads Q_1 and Q_2. The surface load Q_1, at the top of the block, is held fixed at 100 MPa and the surface loads Q_2, applied at opposite sides of the block, are alternating between ± 100 MPa. Due to the symmetry only a quarter of this problem is considered, i.e. a cube of $0.1 \times 0.1 \times 0.1$ m^3. Thus, rigid body motion in the tangential directions are excluded. However, the body is not fixed perpendicular with respect to these tangential directions (only unilaterally constrained). Consequently, the resulting stiffness matrix $\mathbf{K}$ of the problem is positive semi-definite, implying that the scheme developed in Section 4.2 is not applicable. Instead (13) is solved straightforwardly by using the package from Netlib.

The above problem has been solved for 300 cycles using eight time increments for each cycle. The quarter of the structure, i.e. the cube which is considered, is modelled using $10 \times 10 \times 10$ elements. The friction coefficient is $\mu = 0.4$ and the wear coefficient is $k = 1.0 \cdot 10^{-10}$ Pa^{-1}. The contact pressure and the wear gap are plotted in Fig. 5 when Q_2 is passing -100 MPa after 1, 100, 200 and 300 cycles, respectively. After 300 cycles (or 3000 cycles with $k = 1.0 \cdot 10^{-11}$ Pa^{-1}), the solution has reached a steady state. That is, in the sticking area a plateau in the contact pressure has been developed which drops to a zero value in the slipping zone. Furthermore, the steady state

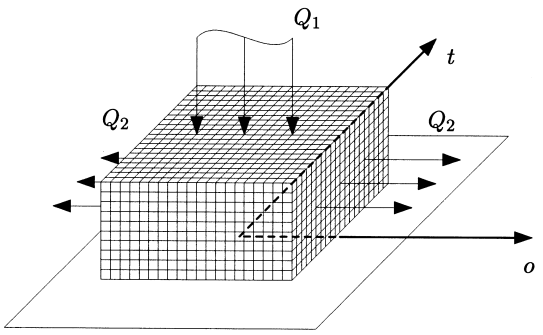

Fig. 4. An elastic block unilaterally constrained to a rigid support and subjected to surface loads Q_1 and Q_2.

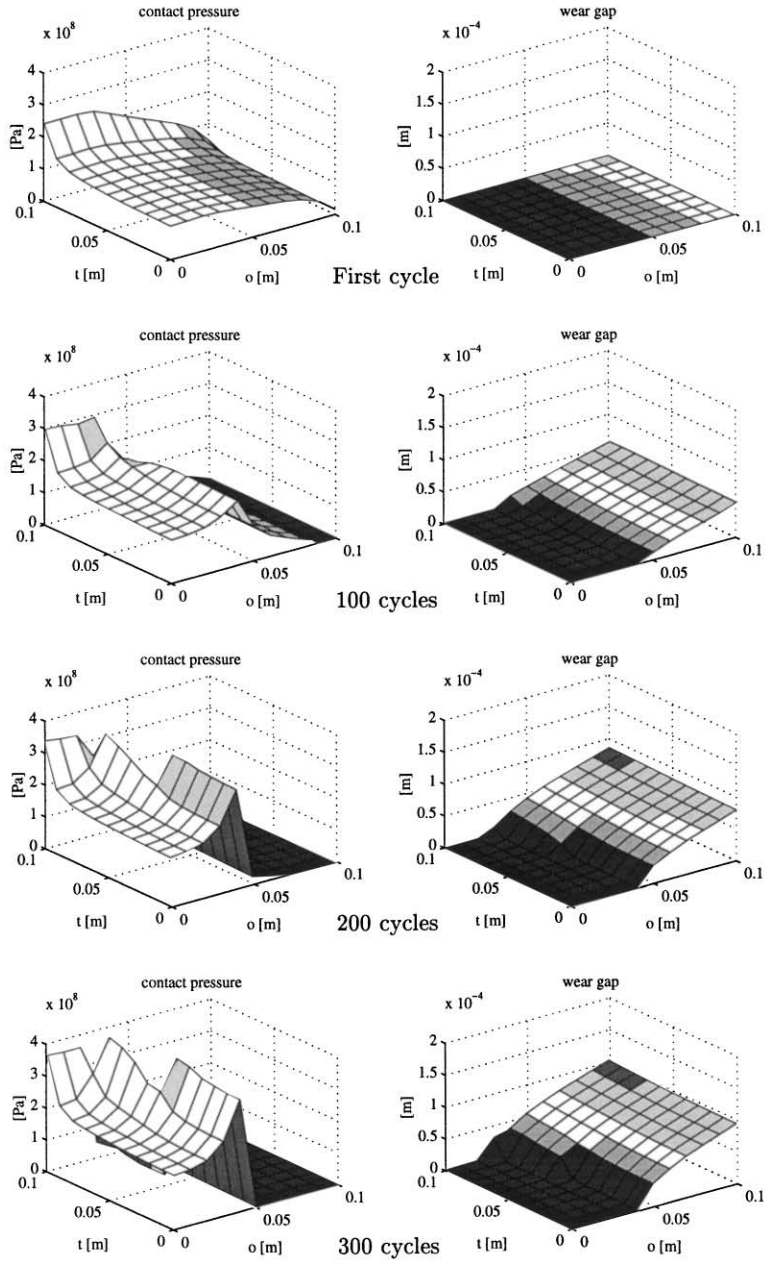

Fig. 5. The evolution in contact pressure and wear gap for the elastic block problem. Due to the symmetry, only a quarter of the problem is considered. The result is plotted when Q_2 is passing -100 MPa. The tangential coordinates o and t are defined in Fig. 4.

implies that the wear rate is almost zero. An important consequence of this is that the dissipation is also almost zero. This is in contrast to the result obtained from a pure frictional analysis, i.e. when the wear coefficient equals zero in our fretting model. For such an analysis, the dissipation would be constant for each cycle. This fact may be taken into consideration when designing frictional dampers (see e.g. Wang and Paul, 1993).

Finally, some execution statistics for the examples considered are provided. For the punch problems, the computing time for each time increment is about one minute on average; five Newton iterations per time increment are needed and three line searches per Newton iteration are performed. The same statistics are higher for the elastic block problem, probably depending on the fact that the stiffness matrix **K** is positive semi-definite for this problem. Here, the computing time for each time increment is about three minutes on average; in general ten Newton iterations per time increment are needed and 13 line searches per Newton iteration are performed.

6. Concluding remarks

In the present work a Newton algorithm for solving fretting problems has been developed and investigated. The algorithm has been implemented in Fortran 77 on a work station and two different types of three-dimensional fretting problems have been solved using the method. Details concerning the implementation are provided in the paper.

In a previous work by Christensen et al. (1997), it was shown that the developed Newton algorithm runs properly for two-dimensional friction problems. In this work it has been found that the algorithm also is attractive, both in robustness and speed, for three-dimensional fretting problems (consequently this is also true for three-dimensional friction problems). In addition, the structure of the algorithm is fitted for implementation in commercial finite element codes.

So far, the wear model adopted in this paper has been utilized to study the contact tractions dependency on the removal of material from fretting surfaces. Three-dimensional fretting problems with stick-slip conditions have been investigated. On the contrary to fretting problems with global sliding conditions, uniform contact pressures are not obtained for these problems. Instead the contact stresses evolve to strongly non-uniform fields. Such non-uniform contact stresses with large values and large gradients might cause crack initiation, leading to fretting fatigue failures. Thus, concerning the field of fretting fatigue, fretting with stick-slip conditions might be a much more severe mechanism than fretting with global sliding conditions. Therefore, for future plans, it would be of interest to extend the elastic properties of this wear model with an appropriate non-linear material model such as the Dang Van criterion (see Petiot et al., 1995) in order to study fretting fatigue.

Acknowledgement

This work was supported by the Swedish Research Council for Engineering Sciences (TFR).

References

Alart, P. and Curnier, A., 1991. A mixed formulation for frictional contact problems prone to Newton like solution methods. Comp. Meths. Appl. Mech. Engng 92, 353–375.

Archard, J.F., 1953. Contact and rubbing of flat surfaces. J. Appl. Phys. 24, 981–988.
Christensen, P.W., Klarbring, A., Pang, J.S., Strömberg, N., 1997. Formulation and comparison of algorithms for frictional contact problems. Accepted for publication in International Journal for Numerical Methods in Engineering.
Curnier, A., 1984. A theory of friction. Int. J. Solids Structures 20, 637–647.
De Saxce, G., Feng, Z.Q., 1991. New inequality and functional for contact with friction : the implicit standard material approach. Mech. Struct. and Mach. 19, 301–325.
Faanes, S., 1996. Distribution of contact stresses along a worn fretting surface. Int. J. Solids Structures 33, 3477–3489.
Frémond, M., 1987. Adhérence des solides. Journal de Mécanique Théorique et Appliquée 6, 383–407.
Hills, D.A., 1994. Mechanics of fretting fatigue. Wear 175, 107–113.
Hughes, T.J.R., 1987. The Finite Element Method, Linear Static and Dynamic Finite Element Analysis. Prentice-Hall, New Jersey.
Hurricks, P.L., 1970. The mechanism of fretting—a review. Wear 15, 389–409.
Johansson, L., 1986. A study on the discretization of a mixed formulation of the frictionless contact problem. M.Sc. thesis, LiTH-IKP-EX-583, Linköping Institute of Technology, Sweden.
Johansson, L., 1992. Elastic and thermoelastic contact problems with friction and wear. Ph.D. thesis, Dissertations No. 266, Linköping Institute of Technology, Sweden.
Klarbring, A., 1990. Derivation and analysis of rate boundary-value problems of frictional contact. Eur. J. Mech., A/Solids 9, 53–85.
Klarbring, A., 1992. Mathematical programming and augmented Lagrangian methods for frictional contact problems. In : Curnier, A. (Ed.), Proc. Contact Mechanics Int. Sypm. PPUR, pp. 409–422.
Michalowski, R., Mróz, Z., 1978. Associated and non-associated sliding rules in contact friction problems. Archives of Mechanics 30, 259–276.
Mutoh, Y., 1995. Mechanisms of fretting fatigue. JSME Int. J. 38, 405–415.
Pang, J.S., 1990. Newton's method for B-differentiable equations. Math. Operations Research 15, 311–341.
Petiot, C., Vincent, L., Dang Van, K., Maouche, N., Foulquier, J., Journet, B., 1995. An analysis of fretting-fatigue failure combined with numerical calculations to predict crack nucleation. Wear 181–183, 101–111.
Rabinowicz, E., 1965. Friction and Wear of Materials. John Wiley and Sons, New York.
Stowers, I.F., Rabinowicz, E., 1973. The mechanism of fretting wear. J. Lubrication Tech. 95, 65–70.
Strömberg, N., 1997. An augmented Lagrangian method for fretting problems. Eur. J. Mech., A/Solids 16, 573–593.
Strömberg, N., Johansson, L., Klarbring, A., 1996. Derivation and analysis of a generalized standard model for contact, friction and wear. Int. J. Solids Structures 83, 1817–1836.
Vincent, L., Berthier, Y., Dubourg, M.C., Godet, M., 1992. Mechanics and materials in fretting. Wear 153, 135–148.
Vingsbo, O., Söderberg, S., 1988. On fretting maps. Wear 126, 131–147.
Wang, Chau-Chang, Paul, B., 1993. Design of beams with Coulomb slip-damping at cover-plate interfaces. In : Attia, M.H., Komanduri, R. (Eds.), Contact Problems and Surface Interactions in Manufacturing and Tribological Systems. ASME, PED-Vol. 67, pp. 309–318.
Waterhouse, R.B., 1981. Theories of fretting processes. In : Waterhouse, R.B. (Ed.), Fretting Fatigue. Applied Science Publishers, Essex, U.K. pp. 203–219.
Waterhouse, R.B., 1984. Fretting wear. Wear 100, 107–118.
Waterhouse, R.B., 1992. Fretting fatigue. International Materials Review 37, 77–97.
Wriggers, P., 1996. Finite element methods for contact problems with friction. Tribology International 29, 651–658.