Open access • Journal Article • DOI:10.1016/S0997-7538(02)01208-1

Finite element algorithms for thermoelastic wear problems — Source link

Peter Ireman, Anders Klarbring, Niclas Strömberg

Institutions: Linköping University

Published on: 01 Jan 2002 - European Journal of Mechanics A-solids (Elsevier)

Topics: Numerical analysis, Gauss–Seidel method, Finite element method, Thermoelastic damping and Newton's method

Finite element algorithms for thermoelastic wear problems

Peter Ireman, Anders Klarbring, Niclas Strömberg

Finite element algorithms for thermoelastic wear problems

Peter Ireman, Anders Klarbring, Niclas Strömberg

Department of Mechanical Engineering, Linköping University, SE-581 83 Linköping, Sweden

Received 14 May 2001; received in revised form 12 December 2001

In the present paper three algorithms are applied to a finite element model of two thermoelastic bodies in frictional wearing contact. All three algorithms utilize a modification of a Newton method for B-differentiable equations as non-linear equation solver. In the first algorithm the fully-coupled system of thermomechanical equations is solved directly using the modified method, while in the other two algorithms the equation system is decoupled in one mechanical part and another thermal part which are solved using an iterative strategy of Gauss–Seidel type. The two iterative algorithms differ in which order the parts are solved. The numerical performance of the algorithms are investigated for two two-dimensional examples. Based on these numerical results, the behaviour of the model is also discussed. It is found that the iterative approach where the thermal subproblem is solved first is slightly more efficient for both examples. Furthermore, it is shown numerically how the predicted wear gap is influenced by the bulk properties of the contacting bodies, in particular how it is influenced by thermal dilatation.

Keywords: Thermomechanical contact; Numerical methods; Friction; Wear; Semi-smooth equations; Pang's Newton method

1. Introduction

The present paper concerns the numerical treatment of the thermoelastic model of contact, friction and wear developed in (Strömberg et al., 1996). A first attempt in this direction is reported in (Strömberg, 1998), where a nonsmooth Newton method was used to solve the discretized problem directly without any splitting of the problem in mechanical and thermal parts. However, that work was restricted to one body only in frictional wearing contact with a rigid support. The present work extends this approach to the case of *two-body contact*, which introduces additional considerations concerning heat flow across the contact interface and how heat generated by friction and wear is divided between the two bodies. In addition, the direct use of the Newton method to solve the fully-coupled thermoelastic problem is compared to two strategies of Gauss–Seidel type, where the mechanical and thermal problems are solved uncoupled from each other. The two iterative strategies differ in which order the mechanical and the thermal problems are solved. The efficiency of the three approaches, i.e. the direct approach and the two iterative strategies, is investigated. For the two examples presented in this work it is found that the performance of the algorithms is very similar. The iterative approach where the thermal subproblem is solved first is slightly more efficient than the other two and the ranking between the direct approach and the iterative strategy where the mechanical subproblem is solved first seems to depend on the type of boundary conditions.

A possible drawback with the direct use of the Newton method is the numerical difficulties reported in (Strömberg, 1998). In that paper, these problems were solved by eliminating the temperature algebraically before the Newton step was performed. However, in order to extend the method to contact between two thermoelastic bodies, involving heat flow across the contact interface, these numerical difficulties have to be overcome in a different way, since elimination of the temperature is not possible when two-body contact is treated. Fortunately, this was resolved simply by using a different solver for the system of linear

equations appearing in the Newton method. By solving the same problem as in (Strömberg, 1998), with the temperature as an explicit variable, it was concluded that there were no numerical problems in a different implementation of the algorithm.

The numerical methods suggested in this work are very much based on the papers by Alart and Curnier (1991) and Klarbring (1992). Alart and Curnier derived an equation system for frictional contact from an augmented Lagrangian which in turn was solved using a Newton method without line-search. Klarbring derived the same equation system by using the approach of projections, i.e. deriving equivalent equations from the variational inequalities of Signorini and Coulomb. Furthermore, Klarbring also suggested that this equation system can be solved using a Newton method for B-differentiable equations (B for Bouligand) proposed by Pang (1990). A function which is Lipschitz continuous and directionally differentiable is B-differentiable. Both Alart and Curnier, and Klarbring obtained the Lipschitz continuity of the projection corresponding to Coulomb's law, needed for establishing the B-differentiability, by replacing the normal force with the projection corresponding to Signorini's condition. In (Strömberg, 1997) this was instead obtained by simply replacing the normal force, N, by $\max(0, N)$. This approach simplify the frictional formulation, especially when wear is included in the formulation.

A modification of Pang's Newton method was introduced in (Strömberg, 1997) to solve the augmented Lagrangian formulation of frictional contact (including wear). A draw-back of Pang's method is the non-linearity introduced by the directional derivative at non-differentiable states. This draw-back was overcome in (Strömberg, 1997) by just determining the directional derivative for one arbitrary direction at non-differentiable states, and then letting this particular derivative represent the directional derivative in all the other directions at the corresponding state. In such manner the step of solving the search direction is linearized. This approach of modifying Pang's Newton method has proven numerically to be very successful for solving both friction problems as well as wear problems. For instance in (Christensen et al., 1998) it was proven that this method is superior to an interior point method; in (Strömberg, 1999) three-dimensional wear problems were solved successfully; and, recently, the transmission error in spur gears was studied using this method (Lundvall and Klarbring, 2000). Furthermore, in (Johansson and Klarbring, 2000) impact problems were solved using the same idea, and most recently in (Christensen, 2000) elastoplastic friction problems were solved using this approach.

In this work, the augmented Lagrangian approach outlined above is adopted for solving two-dimensional thermoelastic wear problems. Similar problems were solved by Johansson and Klarbring (1993) by splitting the problem into one mechanical part and one thermal part. Wear was not included in their work. Another example of work concerning finite element treatment of thermomechanical friction problems, also excluding wear, is by Wriggers and Miehe (1994). In their work, the problem was also decoupled in one mechanical part and one thermal part.

The contents of the present study is as follows: in Section 2 the governing equations for the unilateral frictional wearing contact between two thermoelastic bodies are given. In Section 3 the governing equations are put together to an initial boundary value problem, which is discretized in space using finite element approximations and in time using Euler backward finite differences. Furthermore, the contact conditions are rewritten as equations by means of projections. The result is a system of semismooth equations. A function is said to be semismooth if a certain limit of the B-derivative is satisfied. Thus, a semismooth function is B-differentiable. Recently, Christensen and Pang (1998) proved that the nonsmooth projections of Signorini and Coulomb are semismooth. This is also true when wear is included in the projections. In Section 4, three algorithms for treating the resulting system of equations are presented. The first algorithm is a direct approach for solving the system of semismooth equations, while the two other algorithms utilizes an iterative approach of Gauss–Seidel type. All algorithms involve the use of a modification of Pang's Newton method, which is described in general terms. Furthermore modifified directional derivatives, leading to the modification of Pang's method, are presented. In Section 5, two example problems are defined and solved using the different algorithms and with different constitutive settings. The performance of the algorithms are discussed as well as the behaviour of the model based on the numerical results.

2. Governing equations

In this section we present the governing equations of a model for thermoelastic frictional wearing contact. The equations are: (i) for the bulk material of the two bodies: the equilibrium equations, the constitutive law of isotropic linear elasticity including thermal expansion and Fourier's law of heat conduction; and (ii) for the contact interface: the law of action and reaction, Signorini's law of contact, Coulomb's law of friction, Archard's law of wear and the energy balance combined with constitutive equations of contact heat transfer.

The model has been derived in detail from a thermodynamic generalized standard model concept in (Johansson and Klarbring, 1993; Strömberg et al., 1996).

Let us consider the two bodies Ω^l $(l = 1, 2)$ shown in Fig. 1. The boundary of each body $\partial\Omega^l$ is divided into three disjoint parts: $\Gamma_t^l \subset \partial\Omega^l$ on which tractions $\hat{\boldsymbol{t}}^l$ are prescribed, $\Gamma_u^l \subset \partial\Omega^l$ with prescribed displacements $\hat{\boldsymbol{u}}$ and $\Gamma_c^l \subset \partial\Omega^l$ representing the potential contact surfaces. It is possible for heat to flow through these contact surfaces and, in addition, there may be a part, Γ_T^l of the boundary of each body, respectively, with $\Gamma_T^l \cap \Gamma_c^l = \emptyset$, where the temperature $\widehat{T}$ is prescribed.

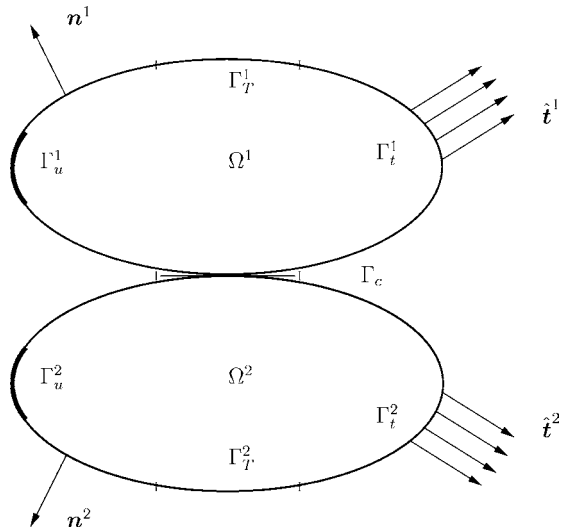

Fig. 1. Two thermoelastic bodies in frictional wearing contact.

The boundaries Γ_c^l have outward unit normal vectors $\boldsymbol{n}_c^l$. For the case of small displacement problems we are considering, these boundaries are almost coinciding and consequently the corresponding normal vectors are almost identical, i.e. $\Gamma_c^1 \simeq \Gamma_c^2$ and $\boldsymbol{n}_c^1 \simeq -\boldsymbol{n}_c^2$. This is utilized to define a common contact surface $\Gamma_c \simeq \Gamma_c^1 \simeq \Gamma_c^2$ with a normal vector $\boldsymbol{n}_c \simeq \boldsymbol{n}_c^1 \simeq -\boldsymbol{n}_c^2$.

The equilibrium equations for the bodies and the law of action and reaction read

$$\operatorname{div}\boldsymbol{\sigma} = \boldsymbol{0} \quad \text{in } \Omega^1 \cup \Omega^2, \tag{1}$$

$$\boldsymbol{\sigma}\boldsymbol{n} = \hat{\boldsymbol{t}} \quad \text{on } \Gamma_t^1 \cup \Gamma_t^2, \tag{2}$$

$$\boldsymbol{\sigma}^1\boldsymbol{n}_c = -\boldsymbol{\sigma}^2\boldsymbol{n}_c = -\boldsymbol{p} \quad \text{on } \Gamma_c, \tag{3}$$

where $\boldsymbol{\sigma}$ is the Cauchy stress tensor, div represents the divergence of a tensor, $\boldsymbol{n}$ is the unit normal vector on $\Gamma_t^1 \cup \Gamma_t^2$ and $\boldsymbol{p}$ is the contact traction vector. Superscripts 1 and 2 are used to denote quantities related to Ω^1 and Ω^2, respectively.

Established constitutive assumptions of isotropic thermoelasticity are utilized for the bodies. These constitutive assumptions are formulated as

$$\boldsymbol{\sigma} = \tilde{\lambda}\,\mathrm{tr}(\boldsymbol{\epsilon})\boldsymbol{I} + 2\tilde{\mu}\boldsymbol{\epsilon} - \big(3\tilde{\lambda} + 2\tilde{\mu}\big)\alpha(T - T_0)\boldsymbol{I}, \tag{4}$$

where $\boldsymbol{I}$ is the identity tensor, $\tilde{\lambda}$ and $\tilde{\mu}$ are the Lamé coefficients, α is the coefficient of thermal dilatation, T_0 is a reference temperature, tr stands for the trace of a tensor[1] and $\boldsymbol{\epsilon} = \boldsymbol{\epsilon}(\boldsymbol{u})$ is the infinitesimal strain tensor defined by

$$\boldsymbol{\epsilon}(\boldsymbol{u}) = \frac{1}{2}\big(\nabla\boldsymbol{u} + \nabla\boldsymbol{u}^{\mathrm{T}}\big),$$

where $\boldsymbol{u}$ is the displacement vector and ∇ denotes the gradient of a scalar or vector.

For the potential contact interface, Signorini's contact conditions, Coulomb's law of friction and Archard's law of wear are assumed to be valid locally at each point $\boldsymbol{x} \in \Gamma_c$.

Signorini's contact conditions may be stated as the following variational inequality:

$$p_n \in \mathcal{K}_n: \quad \big(p_n' - p_n\big)(w_n - \omega - g) \leqslant 0, \quad \forall p_n' \in \mathcal{K}_n, \tag{5}$$

where $p_n = \boldsymbol{p}\cdot\boldsymbol{n}_c$, $\mathcal{K}_n = \{p_n\colon\ p_n \geqslant 0\}$, $w_n = \boldsymbol{w}\cdot\boldsymbol{n}_c$, g is the initial gap between the bodies, ω is the additional gap due to wear and $\boldsymbol{w} = \boldsymbol{u}^1 - \boldsymbol{u}^2$ is the relative displacement of points on Γ_c. The variational inequality (5) is equivalent to the complementarity conditions

$$p_n \geqslant 0, \qquad w_n - \omega - g \leqslant 0, \qquad p_n(w_n - \omega - g) = 0,$$

which indicate that the contact pressure is restricted in sign, the bodies can not penetrate and action at a distance is not allowed.

[1] $\mathrm{tr}(\boldsymbol{a}\otimes\boldsymbol{b}) = \boldsymbol{a}\cdot\boldsymbol{b}$, $(\boldsymbol{a}\otimes\boldsymbol{b})\boldsymbol{c} = (\boldsymbol{b}\cdot\boldsymbol{c})\boldsymbol{a}$, $\forall\boldsymbol{c}$ and $\boldsymbol{a}\cdot\boldsymbol{b}$ is the scalar product of two geometric vectors.

Coulomb's law of friction is formulated as the following principle of maximum dissipation:

$$\boldsymbol{p}_t \in \mathcal{F}(p_n): \quad \dot{\boldsymbol{w}}_t \cdot \big(\boldsymbol{p}_t' - \boldsymbol{p}_t\big) \leqslant 0, \quad \forall \boldsymbol{p}_t' \in \mathcal{F}(p_n), \tag{6}$$

where $\mathcal{F}(p_n) = \{\boldsymbol{p}_t\colon |\boldsymbol{p}_t| \leqslant \mu(p_n)_+\}$, μ is the coefficient of friction, $(x)_+ = \max(0, x)$ and $\boldsymbol{w}_t$ and $\boldsymbol{p}_t$ are the projections of $\boldsymbol{w}$ and $\boldsymbol{p}$ onto the contact tangent plane, i.e. $\boldsymbol{w}_t = (\boldsymbol{I} - \boldsymbol{n}_c \otimes \boldsymbol{n}_c)\boldsymbol{w}$ and $\boldsymbol{p}_t = (\boldsymbol{I} - \boldsymbol{n}_c \otimes \boldsymbol{n}_c)\boldsymbol{p}$. A superposed dot denotes the time derivative. The variational inequality (6) implies that if $\dot{\boldsymbol{w}}_t = \boldsymbol{0}$, then $|\boldsymbol{p}_t| \leqslant \mu(p_n)_+$, and if $\dot{\boldsymbol{w}}_t \neq \boldsymbol{0}$, then $\boldsymbol{p}_t = \mu(p_n)_+ \dot{\boldsymbol{w}}_t / |\dot{\boldsymbol{w}}_t|$.

The evolution of the wear gap is assumed to be governed by Archard's law of wear:

$$\dot{\omega} = k(p_n)_+ |\dot{\boldsymbol{w}}_t|, \tag{7}$$

where k is the wear coefficient.

Furthermore, the temperature fields inside the bodies are governed by the classical equation of heat conduction, representing the balance of energy and Fourier's law of heat diffusion:

$$\rho c \dot{T} = k_t \operatorname{div} \nabla T \quad \text{in } \Omega^1 \cup \Omega^2, \tag{8}$$

where ρ is the density, c is the specific heat capacity and k_t is the thermal conductivity.

The following expression, derived from the general energy balance of the contact interface, an assumption of zero heat capacity of the interface, Coulombs law of friction and Archard's law of wear, represents the balance of energy for the contact interface:

$$\big(k(p_n)_+^2 + \mu(p_n)_+\big)|\dot{\boldsymbol{w}}_t| + \sum_{l=1}^{2} \boldsymbol{q}^l \cdot \boldsymbol{n}_c^l = 0 \quad \text{on } \Gamma_c, \tag{9}$$

where $\boldsymbol{q}^l$ are the heat flux vectors. The thermal contact conditions are given by

$$\boldsymbol{q}^l \cdot \boldsymbol{n}_c^l = \vartheta^l (p_n)_+ \big(T^l - T^0\big), \quad l = 1, 2, \tag{10}$$

where ϑ^l are thermal contact conductances, see, e.g., Fried (1969), and T^0 is the intrinsic temperature of Γ_c. Notice, that this constitutive setting also includes an assumption that the overall contact conductance depends linearly on the contact pressure. The relations in (9) and (10) were derived in a framework of continuum thermodynamics in Strömberg et al. (1996).

By combining (9) and (10) it is possible to eliminate the intrinsic temperature

$$T^0 = \frac{\vartheta^1 T^1 + \vartheta^2 T^2}{\vartheta^1 + \vartheta^2} + \frac{1}{\vartheta^1 + \vartheta^2}\big(k(p_n)_+ + \mu\big)|\dot{\boldsymbol{w}}_t| \tag{11}$$

and obtaining the following expressions for the heat flux across the contact interface:

$$\begin{cases} \boldsymbol{q}^1 \cdot \boldsymbol{n}_c^1 = \dfrac{\vartheta^1 \vartheta^2}{\vartheta^1 + \vartheta^2}(p_n)_+ \big(T^1 - T^2\big) - \dfrac{\vartheta^1}{\vartheta^1 + \vartheta^2}\big(k(p_n)_+^2 + \mu(p_n)_+\big)|\dot{\boldsymbol{w}}_t|, \\ \boldsymbol{q}^2 \cdot \boldsymbol{n}_c^2 = \dfrac{\vartheta^1 \vartheta^2}{\vartheta^1 + \vartheta^2}(p_n)_+ \big(T^2 - T^1\big) - \dfrac{\vartheta^2}{\vartheta^1 + \vartheta^2}\big(k(p_n)_+^2 + \mu(p_n)_+\big)|\dot{\boldsymbol{w}}_t|. \end{cases} \tag{12}$$

The first term of (12) corresponds to the heat conduction across the interface and the second term corresponds to the amount of heat generated by friction and wear that flows into each body respectively. Expressions similar to (12) for the frictional heating problem without wear were derived in (Johansson and Klarbring, 1993).

The problem to be solved can now be summarized as follows: for a given time history of boundary conditions and permissible initial conditions, find the time history of displacement, stresses, contact tractions and temperature such that (1)–(8) and (12) are satisfied.

3. The system of discrete equations

In this section we derive a discrete approximation to the thermoelastic wear problem presented above. The derivation may be outlined as follows: (i) weak forms of the governing equations are introduced and put together to form an initial boundary value problem (IBVP); (ii) the IBVP is discretizied in space using finite element approximations and in time using Euler backward finite differences; and (iii) the discrete versions of the variational inequalities representing Signorini's contact conditions and Coulomb's law of friction are stated as equations by using the approach of projections (Klarbring, 1992).

The final result is a system of semismooth equations, which is solved in the next sections.

3.1. Variational formulation

The IBVP reads: given proper initial conditions and a history of boundary conditions $\hat{\boldsymbol{t}}(t)$, $\hat{\boldsymbol{u}}(t)$ and $\widehat{T}(t)$ on a time interval $[0,\tau]$, find $\boldsymbol{u}:[0,\tau]\to\mathcal{V}$, $T:[0,\tau]\to\mathcal{T}$, $p_n\in\{p_n\colon\ p_n(\boldsymbol{x})\in\mathcal{K}_n,\ \boldsymbol{x}\in\Gamma_c\}$ and $\boldsymbol{p}_t\in\{\boldsymbol{p}_t\colon\ \boldsymbol{p}_t(\boldsymbol{x})\in\mathcal{F}(p_n),\ \boldsymbol{x}\in\Gamma_c\}$ such that for each $t\in[0,\tau]$

$$\int_{\Omega^1\cup\Omega^2}\big(\tilde{\lambda}\,\mathrm{tr}\,\boldsymbol{\epsilon}(\boldsymbol{u})\boldsymbol{I}+2\tilde{\mu}\boldsymbol{\epsilon}(\boldsymbol{u})\big):\boldsymbol{\epsilon}(\boldsymbol{v})\,\mathrm{d}V-\int_{\Omega^1\cup\Omega^2}\big(3\tilde{\lambda}+2\tilde{\mu}\big)\alpha(T-T_0)\boldsymbol{I}:\boldsymbol{\epsilon}(\boldsymbol{v})\,\mathrm{d}V$$
$$=\int_{\Gamma_t^1\cup\Gamma_t^2}\hat{\boldsymbol{t}}\cdot\boldsymbol{v}\,\mathrm{d}A-\int_{\Gamma_c}\boldsymbol{p}\cdot\big(\boldsymbol{v}^1-\boldsymbol{v}^2\big)\,\mathrm{d}A\quad\forall\boldsymbol{v}\in\mathcal{V}_0,\tag{13}$$

$$\int_{\Omega^1\cup\Omega^2}\rho c\dot{T}\varphi\,\mathrm{d}V+\int_{\Omega^1\cup\Omega^2}k_t\nabla T\cdot\nabla\varphi\,\mathrm{d}V=-\int_{\Gamma_c}\frac{\vartheta^1\vartheta^2}{\vartheta^1+\vartheta^2}(p_n)_+\big(T^1-T^2\big)\big(\varphi^1-\varphi^2\big)\,\mathrm{d}A$$
$$+\int_{\Gamma_c}\left(\frac{\vartheta^1\varphi^1}{\vartheta^1+\vartheta^2}+\frac{\vartheta^2\varphi^2}{\vartheta^1+\vartheta^2}\right)\big(k(p_n)_+^2+\mu(p_n)_+\big)|\dot{\boldsymbol{w}}_t|\,\mathrm{d}A\quad\forall\varphi\in\mathcal{T}_0,\tag{14}$$

and weak forms of the tribological laws in (5)–(7) are satisfied, see (Strömberg, 1997). Here we use the following notations for function spaces:

$$\mathcal{V}=\big\{\boldsymbol{v}\colon\ \boldsymbol{v}(\boldsymbol{x})=\hat{\boldsymbol{u}},\ \boldsymbol{x}\in\Gamma_u\big\},\qquad\mathcal{V}_0=\big\{\boldsymbol{v}\colon\ \boldsymbol{v}(\boldsymbol{x})=\boldsymbol{0},\ \boldsymbol{x}\in\Gamma_u\big\},$$
$$\mathcal{T}=\big\{T\colon\ T(\boldsymbol{x})=\widehat{T},\ \boldsymbol{x}\in\Gamma_T\big\},\qquad\mathcal{T}_0=\big\{T\colon\ T(\boldsymbol{x})=0,\ \boldsymbol{x}\in\Gamma_T\big\}.$$

Equation (13) is the weak form of the equilibrium equations (1)–(3) with the assumption of isotropic thermoelasticity (4) inserted. The inner product of two tensors is defined by $\boldsymbol{A}:\boldsymbol{B}=\mathrm{tr}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B})$. Equation (14) is the weak form of the balance of energy for the bodies (8) and for the interface (12) put together using Fourier's law of heat diffusion. Here, it is also observed that (14) remains valid for T substituted by $T-T_0$, i.e. for the deviation from the reference temperature.

3.2. Space discretization

In the following we specialize the treatment to two-dimensional problems. The IBVP stated above is discretizied by introducing finite-dimensional approximations to the function spaces. Integrals over the contact interface Γ_c are evaluated by using an appropriate quadrature rule of the form

$$\int_{\Gamma_c}f(\boldsymbol{x})\,\mathrm{d}A\simeq\sum_{M\in\eta_c}I^M f\big(\boldsymbol{x}^M\big),\tag{15}$$

where I^M are weight factors and η_c is the set of integration points, denoted $\boldsymbol{x}^M$, on Γ_c. In case of bilinear elements, the trapezoidal rule is a possible realization of (15) that experience has shown to work well. In the following, in agreement with the assumption of small displacements, we assume that the space discretization is performed such that node to node contact is obtained, i.e. each node on one side of the potential contact interface forms a pair with one almost coinciding node on the other side of the potential contact interface. Then, when using the trapezoidal rule and bilinear elements, it is possible to choose the integration points on Γ_c such that they coincide with nodal pairs of the displacement elements.

Performing the space discretization outlined above, one obtains the following discrete equilibrium equations:

$$\overline{\boldsymbol{K}}_1\boldsymbol{d}_1-\widehat{\boldsymbol{K}}_1\boldsymbol{T}_1-\boldsymbol{F}_1+\boldsymbol{C}_n^{\mathrm{T}}\boldsymbol{P}_n+\boldsymbol{C}_t^{\mathrm{T}}\boldsymbol{P}_t=\boldsymbol{0},\tag{16}$$
$$\overline{\boldsymbol{K}}_2\boldsymbol{d}_2-\widehat{\boldsymbol{K}}_2\boldsymbol{T}_2-\boldsymbol{F}_2-\boldsymbol{C}_n^{\mathrm{T}}\boldsymbol{P}_n-\boldsymbol{C}_t^{\mathrm{T}}\boldsymbol{P}_t=\boldsymbol{0}\tag{17}$$

and discrete balance of energy for each body respectively:

$$\boldsymbol{M}_l\dot{\boldsymbol{T}}_l+\boldsymbol{O}_l\boldsymbol{T}_l-\boldsymbol{L}_l\big(\dot{\boldsymbol{d}}_1,\dot{\boldsymbol{d}}_2,\boldsymbol{P}_n,\boldsymbol{T}_1,\boldsymbol{T}_2\big)=\boldsymbol{0}\quad(l=1,2),\tag{18}$$

where subscript l indicates body 1 or 2. Matrices and vectors involved in the above expressions are

$$\overline{\boldsymbol{K}}_l = \big[\overline{K}_{ip}^{BA}\big]_l, \quad \overline{K}_{ip}^{BA} = \int_{\Omega^l} \big(\tilde{\lambda}\delta_{ij}\delta_{pq} + \tilde{\mu}(\delta_{ip}\delta_{jq} + \delta_{iq}\delta_{jp})\big)\frac{\partial N^A}{\partial x_q}\frac{\partial N^B}{\partial x_j}\,\mathrm{d}V,$$

$$\widehat{\boldsymbol{K}}_l = \big[\widehat{K}_i^{BA}\big]_l, \quad \widehat{K}_i^{BA} = \int_{\Omega^l} \alpha\big(3\tilde{\lambda} + 2\tilde{\mu}\big)N^A\frac{\partial N^B}{\partial x_i}\,\mathrm{d}V, \qquad \boldsymbol{F}_l = \big\{F_i^B\big\}_l, \quad F_i^B = \int_{\Gamma_t^l} t_i N^B\,\mathrm{d}A,$$

$$\boldsymbol{C}_n = \big[C_{ni}^{MB}\big], \quad C_{ni}^{MB} = N^B\big(\boldsymbol{x}^M\big)n_{ci}\big(\boldsymbol{x}^M\big), \qquad \boldsymbol{C}_t = \big[C_{ti}^{MB}\big], \quad C_{ti}^{MB} = N^B\big(\boldsymbol{x}^M\big)n_{ti}\big(\boldsymbol{x}^M\big),$$

$$\boldsymbol{P}_n = \big\{P_n^M\big\}, \quad P_n^M = I^M p_n\big(\boldsymbol{x}^M\big), \qquad \boldsymbol{P}_t = \big\{P_t^M\big\}, \quad P_t^M = I^M p_t\big(\boldsymbol{x}^M\big),$$

$$\boldsymbol{M}_l = \big[M^{BA}\big]_l, \quad M^{BA} = \int_{\Omega^l} \rho c N^A N^B\,\mathrm{d}V, \qquad \boldsymbol{O}_l = \big[O^{BA}\big]_l, \quad O^{BA} = \int_{\Omega^l} k_t\frac{\partial N^A}{\partial x_i}\frac{\partial N^B}{\partial x_i}\,\mathrm{d}V,$$

where vectors and tensors have been expressed in a Cartesian frame, N^A represents finite element shape functions, δ_{ij} is the Kronecker delta, $\boldsymbol{n}_t$ is the unit tangent vector of the potential contact surface and the summation convention is applied to repeated indices. Furthermore, $\boldsymbol{d}_l = \{d_k^A\}_l$ are vectors of nodal displacements and $\boldsymbol{T}_l = \{T^A\}_l$ are vectors of nodal temperatures, interpreted as the deviation from the reference temperature. Assuming that nodal values are ordered such that the indices corresponding to contact nodes are put first, the vectors $\boldsymbol{L}_l$, corresponding to the discretization of the right-hand side of (14), has the form $\boldsymbol{L}_l = (\boldsymbol{L}_l^c, \boldsymbol{0})$ and typical entries of $\boldsymbol{L}_1^c$ and $\boldsymbol{L}_2^c$, respectively, are given by

$$L_1^M = \frac{\vartheta^1\vartheta^2}{\vartheta^1+\vartheta^2}\big(P_n^M\big)_+\big(T_2^M - T_1^M\big) + \frac{\vartheta^1}{\vartheta^1+\vartheta^2}\left(\frac{k}{I^M}\big(P_n^M\big)_+^2 + \mu\big(P_n^M\big)_+\right)\big|\dot{w}_t^M\big| \quad \text{and} \tag{19}$$

$$L_2^M = \frac{\vartheta^1\vartheta^2}{\vartheta^1+\vartheta^2}\big(P_n^M\big)_+\big(T_1^M - T_2^M\big) + \frac{\vartheta^2}{\vartheta^1+\vartheta^2}\left(\frac{k}{I^M}\big(P_n^M\big)_+^2 + \mu\big(P_n^M\big)_+\right)\big|\dot{w}_t^M\big|, \tag{20}$$

where $\{w_t^M\} = \boldsymbol{C}_t(\boldsymbol{d}_1 - \boldsymbol{d}_2)$ and the fact that

$$N^B\big(\boldsymbol{x}^M\big) = \begin{cases} 1 & \text{if } B = M, \\ 0 & \text{if } B \neq M, \end{cases}$$

when using bilinear elements and the trapezoidal rule as outlined above is used.

The space discretization of the tribological laws are obtained by assuming that (5)–(7) are valid in each integration point $M \in \eta_c$, i.e. for the two-dimensional case:

$$P_n^M \in \mathcal{K}_n^h\colon \quad \big(P_n^{M\prime} - P_n^M\big)\big(w_n^M - \omega^M - g^M\big) \leqslant 0 \quad \forall P_n^{M\prime} \in \mathcal{K}_n^h, \tag{21}$$

$$P_t^M \in \mathcal{F}^h\big(P_n^M\big)\colon \quad \dot{w}_t^M\big(P_t^{M\prime} - P_t^M\big) \leqslant 0 \quad \forall P_t^{M\prime} \in \mathcal{F}^h\big(P_n^M\big), \tag{22}$$

$$\dot{\omega}^M = \frac{k}{I^M}\big(P_n^M\big)_+\big|\dot{w}_t^M\big|, \tag{23}$$

where $\mathcal{K}_n^h = \{P_n^M\colon\ P_n^M \geqslant 0\}$, $\mathcal{F}^h(P_n^M) = \{P_t^M\colon\ |P_t^M| \leqslant \mu(P_n^M)_+\}$, $\omega^M = \omega(\boldsymbol{x}^M)$, $g^M = g(\boldsymbol{x}^M)$ and $\{w_n^M\} = \boldsymbol{C}_n(\boldsymbol{d}_1 - \boldsymbol{d}_2)$.

The relations (21)–(23) were derived in (Strömberg, 1997) from weak forms of (5)–(7) using the integration rule in (15).

3.3. Time discretization

The time rates appearing in (18), (22) and (23) are approximated by an Euler backward discretization:

$$\dot{\boldsymbol{T}}_l(t_{k+1}) \simeq \frac{\boldsymbol{T}_l(t_{k+1}) - \boldsymbol{T}_l(t_k)}{t_{k+1} - t_k}, \tag{24}$$

$$\dot{w}_t^M(t_{k+1}) \simeq \frac{w_t^M(t_{k+1}) - w_t^M(t_k)}{t_{k+1} - t_k}, \tag{25}$$

$$\dot{\omega}^M(t_{k+1}) \simeq \frac{\omega^M(t_{k+1}) - \omega^M(t_k)}{t_{k+1} - t_k}. \tag{26}$$

By inserting (24) and (25) in (18)–(20), the following form of the equations governing the temperature fields are obtained for each body ($l = 1, 2$), respectively:

$$\big(\boldsymbol{M}_l + (t_{k+1} - t_k)\boldsymbol{O}_l\big)\boldsymbol{T}_l - \boldsymbol{Q}_l - \overline{\boldsymbol{L}}_l(\boldsymbol{d}_1, \boldsymbol{d}_2, \boldsymbol{P}_n, \boldsymbol{T}_1, \boldsymbol{T}_2) = \boldsymbol{0}, \tag{27}$$

where $\boldsymbol{Q}_l = \boldsymbol{M}_l \boldsymbol{T}_l(t_k)$ and typical entries of $\overline{\boldsymbol{L}}_l^c$ are given by

$$\overline{L}_1^M = \frac{\vartheta^1\vartheta^2}{\vartheta^1+\vartheta^2}\big(P_n^M\big)_+\big(T_2^M - T_1^M\big)(t_{k+1}-t_k) + \frac{\vartheta^1}{\vartheta^1+\vartheta^2}\left(\frac{k}{I^M}\big(P_n^M\big)_+^2 + \mu\big(P_n^M\big)_+\right)\big|\overline{w}_t^M\big| \quad \text{and} \tag{28}$$

$$\overline{L}_2^M = \frac{\vartheta^1\vartheta^2}{\vartheta^1+\vartheta^2}\big(P_n^M\big)_+\big(T_1^M - T_2^M\big)(t_{k+1}-t_k) + \frac{\vartheta^2}{\vartheta^1+\vartheta^2}\left(\frac{k}{I^M}\big(P_n^M\big)_+^2 + \mu\big(P_n^M\big)_+\right)\big|\overline{w}_t^M\big|, \tag{29}$$

where $\overline{w}_t^M = w_t^M(t_{k+1}) - w_t^M(t_k)$. Furthermore, inserting (25) in (22) gives

$$P_t^M \in \mathcal{F}^h\big(P_n^M\big): \quad \overline{w}_t^M\big(P_t^{M'} - P_t^M\big) \leqslant 0, \quad \forall P_t^{M'} \in \mathcal{F}^h\big(P_n^M\big). \tag{30}$$

Finally, inserting (25) and (26) in (23) gives

$$\omega^M(t_{k+1}) = \omega_0^M + \frac{k}{I^M}\big(P_n^M\big)_+\big|\overline{w}_t^M\big|, \quad \text{where } \omega_0^M = \omega^M(t_k). \tag{31}$$

3.4. Projections

The last step in order to obtain a formulation suitable for numerical treatment is to rewrite the variational inequalities in (21) and (30) as equations by means of projections (Klarbring, 1992).

By multiplying (21) by $r > 0$ and adding and subtracting P_n^M, the variational inequality equivalent to Signorini's contact conditions takes the form: find $P_n^M \in \mathcal{K}_n^h$ such that

$$\big(P_n^M - \big(P_n^M + r\big(w_n^M - \omega^M - g^M\big)\big)\big)\big(P_n^{M'} - P_n^M\big) \geqslant 0, \quad \forall P_n^{M'} \in \mathcal{K}_n^h. \tag{32}$$

Equation (32) states that P_n^M is the projection of $P_n^M + r(w_n^M - \omega^M - g^M)$ onto $\mathcal{K}_n^h$, i.e., on using (31),

$$P_n^M = \left(P_n^M + r\left(w_n^M - \omega_0^M - \frac{k}{I^M}\big(P_n^M\big)_+\big|\overline{w}_t^M\big| - g^M\right)\right)_+. \tag{33}$$

In the same manner (30) is rephrased: find $P_t^M \in \mathcal{F}^h(P_n^M)$ such that

$$\big(P_t^M - \big(P_t^M + r\overline{w}_t^M\big)\big)\big(P_t^{M'} - P_t^M\big) \geqslant 0, \quad \forall P_t^{M'} \in \mathcal{F}^h\big(P_n^M\big). \tag{34}$$

Thus, P_t^M is the projection of $P_t^M + r\overline{w}_t^M$ onto $\mathcal{F}^h(P_n^M)$, i.e.

$$P_t^M = \begin{cases} P_t^M + r\overline{w}_t^M, & \text{if } \big|P_t^M + r\overline{w}_t^M\big| \leqslant \mu\big(P_n^M\big)_+, \\ \mu\big(P_n^M\big)_+ \operatorname{sgn}\big(P_t^M + r\overline{w}_t^M\big), & \text{otherwise.} \end{cases} \tag{35}$$

3.5. System of semismooth equations

The discrete IBVP to be solved is defined by Eqs. (16), (17), (27), (33) and (35). These equations form a system of semismooth equations which is solved using three different algorithms presented in the next section. Before applying the algorithms the equation system is rewritten by performing static condensation such that only variables on the contact surface is treated. This is presented in detail in the appendix, ending up in the following system of semismooth equations:

$$\boldsymbol{H}_C(\boldsymbol{y}_C) = \left\{ \begin{array}{c} \bar{\kappa}\bar{\boldsymbol{d}} - \hat{\kappa}_1\boldsymbol{T}_1^c - \hat{\kappa}_2\boldsymbol{T}_2^c - \boldsymbol{R} + \overline{\boldsymbol{C}}_n^{\mathrm{T}}\boldsymbol{P}_n + \overline{\boldsymbol{C}}_t^{\mathrm{T}}\boldsymbol{P}_t \\ \boldsymbol{A}_1^c\boldsymbol{T}_1^c - \overline{\boldsymbol{Q}}_1^c - \widehat{\boldsymbol{L}}_1^c\big(\bar{\boldsymbol{d}}, \boldsymbol{P}_n, \boldsymbol{P}_t, \boldsymbol{T}_1^c, \boldsymbol{T}_2^c\big) \\ \boldsymbol{A}_2^c\boldsymbol{T}_2^c - \overline{\boldsymbol{Q}}_2^c - \widehat{\boldsymbol{L}}_2^c\big(\bar{\boldsymbol{d}}, \boldsymbol{P}_n, \boldsymbol{P}_t, \boldsymbol{T}_1^c, \boldsymbol{T}_2^c\big) \\ -\boldsymbol{P}_n + \boldsymbol{\Pi}_n\big(\bar{\boldsymbol{d}}, \boldsymbol{P}_n\big) \\ -\boldsymbol{P}_t + \boldsymbol{\Pi}_t\big(\bar{\boldsymbol{d}}, \boldsymbol{P}_n, \boldsymbol{P}_t\big) \end{array} \right\} = \boldsymbol{0}, \tag{36}$$

where $\boldsymbol{y}_C = (\bar{\boldsymbol{d}}, \boldsymbol{P}_n, \boldsymbol{P}_t, \boldsymbol{T}_1^c, \boldsymbol{T}_2^c)$ and

$$\bar{\boldsymbol{d}} = \boldsymbol{d}_1^c - \boldsymbol{d}_2^c, \qquad \bar{\kappa} = \overline{\boldsymbol{K}}_1^c - \overline{\boldsymbol{K}}_1^c\boldsymbol{E}^{-1}\overline{\boldsymbol{K}}_1^c, \qquad \hat{\kappa}_1 = \widehat{\boldsymbol{K}}_1^c - \overline{\boldsymbol{K}}_1^c\boldsymbol{E}^{-1}\widehat{\boldsymbol{K}}_1^c, \qquad \hat{\kappa}_2 = -\overline{\boldsymbol{K}}_1^c\boldsymbol{E}^{-1}\widehat{\boldsymbol{K}}_2^c,$$
$$\boldsymbol{R} = \overline{\boldsymbol{F}}_1^c - \overline{\boldsymbol{K}}_1^c\boldsymbol{E}^{-1}\big(\overline{\boldsymbol{F}}_1^c + \overline{\boldsymbol{F}}_2^c\big), \qquad \boldsymbol{E} = \overline{\boldsymbol{K}}_1^c + \overline{\boldsymbol{K}}_2^c, \qquad \boldsymbol{A}_l = \boldsymbol{M}_l + (t_{k+1}-t_k)\boldsymbol{O}_l, \quad l = 1, 2.$$

Furthermore, $\boldsymbol{\Pi}_n$ and $\boldsymbol{\Pi}_t$ denote the projections defined by (33) and (35), respectively. Superscript c denotes a condensed entity. $\overline{\boldsymbol{C}}_n$ and $\overline{\boldsymbol{C}}_t$ is obtained by deleting columns with zeros corresponding to non-contact nodes in $\boldsymbol{C}_n$ and $\boldsymbol{C}_t$, respectively. The functions $\widehat{\boldsymbol{L}}_1$ and $\widehat{\boldsymbol{L}}_2$ are the same as $\overline{\boldsymbol{L}}_1$ and $\overline{\boldsymbol{L}}_2$ defined by (28) and (29), respectively, except that the dependency on $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ is replaced by a dependency on $\bar{\boldsymbol{d}}$.

The formulation derived in (Strömberg, 1998) is obtained as a special case of (36) by setting $\bar{\boldsymbol{\kappa}} = \overline{\boldsymbol{K}}^c$, $\hat{\boldsymbol{\kappa}}_1 = \widehat{\boldsymbol{K}}^c$, $\boldsymbol{R} = \overline{\boldsymbol{F}}^c$, $\vartheta^1 = 1$ and all quantities related to $l = 2$ to zero.

4. Algorithms

The IBVP defined in the previous section can now be treated by solving a sequence of discrete semismooth equations defined in (36). In this section we present Pang's Newton algorithm for B-differentiable equations in general terms, followed by two different approaches for using it to solve (36). These are: (i) to solve (36) directly using a modification of Pang's Newton method and (ii) to split (36) into a mechanical and a thermal part and to use an iterative technique where, in each iteration, the mechanical part is solved using the modification of Pang's Newton method introduced in approach (i) and the thermal part is simply a system of linear equations.

The second approach is in turn divided into two different strategies, differing in the order the mechanical and the thermal parts are solved.

4.1. Semismoothness

The directional derivative of a function f at $\boldsymbol{x}$ in a direction $\boldsymbol{h}$ is defined by

$$f'(\boldsymbol{x};\boldsymbol{h}) = \lim_{t\downarrow 0} \frac{f(\boldsymbol{x}+t\boldsymbol{h}) - f(\boldsymbol{x})}{t}.$$

A function f is said to be directionally differentiable at $\boldsymbol{x}$ if the above limit exists for all directions $\boldsymbol{h}$. A function f is said to be B-differentiable at $\boldsymbol{x}$ if it is directionally differentiable at $\boldsymbol{x}$ and Lipschitz continuous in a neighborhood of $\boldsymbol{x}$. A function f is said to be semismooth at $\boldsymbol{x}$ if it is B-differentiable at $\boldsymbol{x}$ and the following limit holds

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{f'(\boldsymbol{x}+\boldsymbol{h};\boldsymbol{h}) - f'(\boldsymbol{x};\boldsymbol{h})}{\|\boldsymbol{h}\|} = 0.$$

A continuous function f is said to be piecewise smooth at $\boldsymbol{x}$ if there exist an open neighborhood W of $\boldsymbol{x}$ and a finite family of smooth functions f_i on W such that $f(\tilde{\boldsymbol{x}}) = f_i(\tilde{\boldsymbol{x}})$ for each $\tilde{\boldsymbol{x}}$ on W and for some integer i. A piecewise smooth function is semismooth. Let f and g be piecewise smooth functions, then the following functions are also piecewise smooth: $f+g$, $f-g$, fg, $\max(f,g)$, $\min(f,g)$ and $|f|$ (Chaney, 1990).

Except for the functions $\widehat{\boldsymbol{L}}_l^c$, $\boldsymbol{\Pi}_n$ and $\boldsymbol{\Pi}_t$ the equation system in (36) is smooth. A smooth function f is indeed piecewise smooth, semismooth and B-differentiable, $f'(\boldsymbol{x};\boldsymbol{h}) = (\nabla f(\boldsymbol{x}))\boldsymbol{h}$. In Strömberg (1997) it was shown that $\boldsymbol{\Pi}_n$ and $\boldsymbol{\Pi}_t$ are B-differentiable. It can also be shown that these functions are semismooth. This was proved for the frictional case in (Christensen and Pang, 1998). The additional term appearing in $\boldsymbol{\Pi}_n$ when wear is included does not destroy this property of $\boldsymbol{\Pi}_n$. In fact, by applying the rules for piecewise smooth function presented above, it can be shown that $\boldsymbol{\Pi}_n$ is piecewise smooth. This is also true for $\widehat{\boldsymbol{L}}_l^c$. Thus, the system of equations in (36) is semismooth and consequently B-differentiable.

4.2. Pang's Newton method

The algorithm below for solving $\boldsymbol{H}(\boldsymbol{y}) = \boldsymbol{0}$, where $\boldsymbol{H}(\boldsymbol{y})$ is B-differentiable was suggested by (Pang, 1990).

Algorithm BN: 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 $\boldsymbol{y}^0$ be given from the previous time step k and set $j = 0$.
(1) Find a search direction $\boldsymbol{z}$ such that

$$\boldsymbol{H}(\boldsymbol{y}^j) + \boldsymbol{H}'(\boldsymbol{y}^j;\boldsymbol{z}) = \boldsymbol{0}, \tag{37}$$

where $\boldsymbol{H}'(\boldsymbol{y}^j;\boldsymbol{z})$ is the directional derivative.
(2) Let $\alpha^j = \beta^{m_j}$, where m_j is the smallest integer $m \geqslant 0$ for which the following decrease criterion holds:

$$\phi(\boldsymbol{y}^j + \beta^m \boldsymbol{z}) \leqslant (1 - 2\gamma\beta^m)\phi(\boldsymbol{y}^j), \qquad \phi(\boldsymbol{y}) = \frac{1}{2}\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{y})\boldsymbol{H}(\boldsymbol{y}).$$

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

4.3. Direct approach – modified Newton method

The system of equations to be solved in (37) for $\boldsymbol{H}(\boldsymbol{y})$ defined by (36), is non-linear due to the non-linearity of the directional derivative. In order to increase the effectiveness of solving this step an appropriate modification of the directional derivative $\boldsymbol{H}'(\boldsymbol{y};\boldsymbol{z})$ is introduced such that (37) becomes linear at non-differentiable states. At points where $\boldsymbol{H}(\boldsymbol{y})$ is smooth, Eq. (37) is of course linear as $\boldsymbol{H}'(\boldsymbol{y};\boldsymbol{z}) = \nabla \boldsymbol{H}(\boldsymbol{y})\boldsymbol{z}$.

At non-differentiable states the directional derivative is modified by choosing one of the possible directions, here expressed by $\bar{\boldsymbol{z}}$, and defining $\boldsymbol{J}(\boldsymbol{y})$ by the corresponding directional derivative in the following manner:

$$\boldsymbol{J}(\boldsymbol{y})\bar{\boldsymbol{z}} = \boldsymbol{H}'(\boldsymbol{y};\bar{\boldsymbol{z}})$$

and then replacing the directional derivative in all the other directions by

$$\boldsymbol{H}'(\boldsymbol{y};\boldsymbol{z}) \simeq \boldsymbol{J}(\boldsymbol{y})\boldsymbol{z}.$$

In such a manner (37) becomes linear for alla possible states. This approach was first adopted in (Strömberg, 1997) and has since then proven to work well in many other works, see, e.g., (Christensen et al., 1998; Strömberg, 1998, 1999; Johansson and Klarbring, 2000; Lundvall and Klarbring, 2000). Thus, following this approach, Algorithm BN is modified by replacing (37) with

$$\boldsymbol{H}\big(\boldsymbol{y}^j\big) + \boldsymbol{J}\big(\boldsymbol{y}^j\big)\boldsymbol{z} = \boldsymbol{0}, \tag{38}$$

where $\boldsymbol{J}(\boldsymbol{y}) = \nabla \boldsymbol{H}(\boldsymbol{y})$ at states where $\boldsymbol{H}(\boldsymbol{y})$ is smooth and at non-differentiable states $\boldsymbol{J}(\boldsymbol{y})$ is defined by using the approach outlined above. The modification of Algorithm BN done in (38), together with an additional upper bound $\overline{m}$ on m_j in step 2 of Algorithm BN, is denoted Algorithm MBN. The upper bound $\overline{m}$ prevents the algorithm from stalling.

After convergence in each time step of Algorithm MBN, ω^M is updated using (31) and the state of tangential slip w_t^M is stored. Furthermore, $\boldsymbol{T}_1$ and $\boldsymbol{T}_2$ are calculated using (A.3) in Appendix A in order to obtain new values of $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$ by using $\boldsymbol{Q}_l = \boldsymbol{M}_l \boldsymbol{T}_l(t_k)$.

4.4. Modified directional derivatives

The modification of the directional derivative $\boldsymbol{H}'_C(\boldsymbol{y}_C;\boldsymbol{z})$ defining $\boldsymbol{J}_C(\boldsymbol{y}_C)$ at non-differentiable states are presented here. Except for the functions $\widehat{\boldsymbol{L}}_l^c$, $\boldsymbol{\Pi}_n$ and $\boldsymbol{\Pi}_t$ the equation system in (36) is smooth. Thus, it is only the directional derivatives of these functions which are needed to be modified according to the approach outlined above.

The modification of the directional derivatives of $\widehat{\boldsymbol{L}}_1^c$ and $\widehat{\boldsymbol{L}}_2^c$ are given by

$$\begin{aligned}
&\widehat{\boldsymbol{L}}_1^{c\prime}\big(\bar{\boldsymbol{d}}, \boldsymbol{P}_n, \boldsymbol{T}_1^c, \boldsymbol{T}_2^c; \boldsymbol{z}_d, \boldsymbol{z}_n, \boldsymbol{z}_{T1}, \boldsymbol{z}_{T2}\big) \\
&\quad\simeq \left\{ \begin{array}{l} 0, \quad M \in \eta_{c1} \\ \dfrac{\vartheta^1\vartheta^2}{\vartheta^1+\vartheta^2}(t_{k+1}-t_k)\big(P_n^M\big(z_{T2}^M - z_{T1}^M\big) + \big(T_2^M - T_1^M\big)z_n^M\big), \quad M \in \eta_{c2} \end{array} \right\} \\
&\qquad + \left\{ \begin{array}{l} 0, \quad M \in \eta_{c3} \\ \dfrac{\vartheta^1}{\vartheta^1+\vartheta^2}\left(\left(\dfrac{2k}{I^M}P_n^M + \mu\right)|\overline{w}_t^M|z_n^M + \left(\dfrac{k}{I^M}P_n^M + \mu\right)P_n^M \operatorname{sgn}\big(\overline{w}_t^M\big)z_{wt}^M\right), \quad M \in \eta_{c4} \end{array} \right\}
\end{aligned}$$

and

$$\begin{aligned}
&\widehat{\boldsymbol{L}}_2^{c\prime}\big(\bar{\boldsymbol{d}}, \boldsymbol{P}_n, \boldsymbol{T}_1^c, \boldsymbol{T}_2^c; \boldsymbol{z}_d, \boldsymbol{z}_n, \boldsymbol{z}_{T1}, \boldsymbol{z}_{T2}\big) \\
&\quad\simeq \left\{ \begin{array}{l} 0, \quad M \in \eta_{c1} \\ \dfrac{\vartheta^1\vartheta^2}{\vartheta^1+\vartheta^2}(t_{k+1}-t_k)\big(P_n^M\big(z_{T1}^M - z_{T2}^M\big) + \big(T_1^M - T_2^M\big)z_n^M\big), \quad M \in \eta_{c2} \end{array} \right\} \\
&\qquad + \left\{ \begin{array}{l} 0, \quad M \in \eta_{c3} \\ \dfrac{\vartheta^2}{\vartheta^1+\vartheta^2}\left(\left(\dfrac{2k}{I^M}P_n^M + \mu\right)|\overline{w}_t^M|z_n^M + \left(\frac{k}{I^M}P_n^M + \mu\right)P_n^M \operatorname{sgn}\big(\overline{w}_t^M\big)z_{wt}^M\right), \quad M \in \eta_{c4} \end{array} \right\},
\end{aligned}$$

where

$$\eta_{c1} = \big\{M\colon\ P_n^M \leqslant 0\big\}, \qquad \eta_{c2} = \big\{M\colon\ P_n^M > 0\big\}, \qquad \eta_{c3} = \big\{M\colon\ P_n^M|\overline{w}_t^M| \leqslant 0\big\}, \qquad \eta_{c4} = \big\{M\colon\ P_n^M|\overline{w}_t^M| > 0\big\},$$

and $\boldsymbol{z}_{wt} = \overline{\boldsymbol{C}}_t \boldsymbol{z}_d$.

Modified directional derivatives of $\boldsymbol{\Pi}_n$ and $\boldsymbol{\Pi}_t$ are derived in (Strömberg, 1998). These are:

$$\boldsymbol{\Pi}_n'(\bar{\boldsymbol{d}}, \boldsymbol{P}_n; \boldsymbol{z}_d, \boldsymbol{z}_n) \simeq \left\{ \begin{array}{ll} 0, & M \in \eta_{c5} \\ z_n^M + r z_{wn}^M, & M \in \eta_{c6} \\ z_n^M + r\left(z_{wn}^M - \dfrac{k}{I^M} |\overline{w}_t^M| z_n^M - \dfrac{k}{I^M} P_n^M \operatorname{sgn}(\overline{w}_t^M) z_{wt}^M \right), & M \in \eta_{c7} \end{array} \right\} \quad \text{and}$$

$$\boldsymbol{\Pi}_t'(\bar{\boldsymbol{d}}, \boldsymbol{P}_n, \boldsymbol{P}_t; \boldsymbol{z}_d, \boldsymbol{z}_n, \boldsymbol{z}_t) \simeq \left\{ \begin{array}{ll} 0, & M \in \eta_{c1} \\ z_t^M + r z_{wt}^M, & M \in \eta_{c8} \\ \mu z_n^M \operatorname{sgn}(P_t^M + r\overline{w}_t^M), & M \in \eta_{c9} \end{array} \right\},$$

where

$$\eta_{c5} = \left\{ M\colon\ P_n^M + r\left(w_n^M - \omega_0^M - \frac{k}{I^M} (P_n^M)_+ |\overline{w}_t^M| - g^M \right) \leqslant 0 \right\},$$

$$\eta_{c6} = \left\{ M\colon\ P_n^M |\overline{w}_t^M| \leqslant 0,\ P_n^M + r\left(w_n^M - \omega_0^M - \frac{k}{I^M} (P_n^M)_+ |\overline{w}_t^M| - g^M \right) > 0 \right\},$$

$$\eta_{c7} = \left\{ M\colon\ P_n^M |\overline{w}_t^M| > 0,\ P_n^M + r\left(w_n^M - \omega_0^M - \frac{k}{I^M} (P_n^M)_+ |\overline{w}_t^M| - g^M \right) > 0 \right\},$$

$$\eta_{c8} = \left\{ M\colon\ P_n^M > 0, \left| P_t^M + r\overline{w}_t^M \right| \leqslant \mu P_n^M \right\}, \qquad \eta_{c9} = \left\{ M\colon\ P_n^M > 0, \left| P_t^M + r\overline{w}_t^M \right| > \mu P_n^M \right\},$$

and $\boldsymbol{z}_{wn} = \boldsymbol{C}_n \boldsymbol{z}_d$.

4.5. Iterative strategy

Instead of solving (36) directly using Algorithm MBN, one can adopt an iterative strategy, which from an algorithmic point of view is a method of Gauss–Seidel type, see, e.g., (Ortega and Reinboldt, 1970).

The strategy is adopted by first splitting (36) into the two parts:

$$\boldsymbol{H}_M(\boldsymbol{y}_M, \boldsymbol{y}_T) = \left\{ \begin{array}{c} \bar{\kappa}\bar{\boldsymbol{d}} - \hat{\kappa}_1 \boldsymbol{T}_1^c - \hat{\kappa}_2 \boldsymbol{T}_2^c - \boldsymbol{R} + \overline{\boldsymbol{C}}_n^{\mathrm{T}} \boldsymbol{P}_n + \overline{\boldsymbol{C}}_t^{\mathrm{T}} \boldsymbol{P}_t \\ -\boldsymbol{P}_n + \boldsymbol{\Pi}_n \\ -\boldsymbol{P}_t + \boldsymbol{\Pi}_t \end{array} \right\} = \boldsymbol{0} \quad \text{and}$$

$$\boldsymbol{H}_T(\boldsymbol{y}_M, \boldsymbol{y}_T) = \left\{ \begin{array}{l} \boldsymbol{A}_1^c \boldsymbol{T}_1^c - \overline{\boldsymbol{Q}}_1^c - \widehat{\boldsymbol{L}}_1^c \\ \boldsymbol{A}_2^c \boldsymbol{T}_2^c - \overline{\boldsymbol{Q}}_2^c - \widehat{\boldsymbol{L}}_2^c \end{array} \right\} = \boldsymbol{0},$$

where $\boldsymbol{y}_M = (\bar{\boldsymbol{d}}, \boldsymbol{P}_n, \boldsymbol{P}_t)$ and $\boldsymbol{y}_T = (\boldsymbol{T}_1^c, \boldsymbol{T}_2^c)$. These equations are then treated by either of the following two algorithms:

Algorithm GSMT: Repeat the following steps for each time increment $k+1$:

(0) Set $\boldsymbol{y}_T^{i=0} = \boldsymbol{y}_T(t_k)$.
(1) Solve $\widetilde{\boldsymbol{H}}_M(\boldsymbol{y}_M^{i+1}) = \boldsymbol{H}_M(\boldsymbol{y}_M^{i+1}, \boldsymbol{y}_T^i) = \boldsymbol{0}$ to get $\boldsymbol{y}_M^{i+1}$.
(2) Solve $\widetilde{\boldsymbol{H}}_{\mathrm{T}}(\boldsymbol{y}_T^{i+1}) = \boldsymbol{H}_T(\boldsymbol{y}_M^{i+1}, \boldsymbol{y}_T^{i+1}) = \boldsymbol{0}$ to get $\boldsymbol{y}_T^{i+1}$.
(3) Terminate if $(1/2)\boldsymbol{H}_C^{\mathrm{T}}(\boldsymbol{y}_C^{i+1})\boldsymbol{H}_C(\boldsymbol{y}_C^{i+1}) \leqslant \epsilon$. Otherwise replace i by $i+1$ and return to step 1.

Algorithm GSTM: Repeat the following steps for each time increment $k+1$:

(0) Set $\boldsymbol{y}_T^{i=0} = \boldsymbol{y}_T(t_k)$.
(1) Solve $\overline{\boldsymbol{H}}_T(\boldsymbol{y}_t^{i+1}) = \boldsymbol{H}_T(\boldsymbol{y}_M^i, \boldsymbol{y}_T^{i+1}) = \boldsymbol{0}$ to get $\boldsymbol{y}_T^{i+1}$.
(2) Solve $\overline{\boldsymbol{H}}_M(\boldsymbol{y}_M^{i+1}) = \boldsymbol{H}_M(\boldsymbol{y}_M^{i+1}, \boldsymbol{y}_T^{i+1}) = \boldsymbol{0}$ to get $\boldsymbol{y}_M^{i+1}$.
(3) Terminate if $(1/2)\boldsymbol{H}_C^{\mathrm{T}}(\boldsymbol{y}_C^{i+1})\boldsymbol{H}_C(\boldsymbol{y}_C^{i+1}) \leqslant \epsilon$. Otherwise replace i by $i+1$ and return to step 1.

Thus, the only difference between the algorithms GSMT and GSTM is in which order the mechanical and the thermal problems are solved. When using the algorithms above, $\widetilde{\boldsymbol{H}}_M(\boldsymbol{y}_M) = \boldsymbol{0}$ or $\overline{\boldsymbol{H}}_M(\boldsymbol{y}_M) = \boldsymbol{0}$ is solved by using algorithm MBN

and $\widetilde{\boldsymbol{H}}_T(\boldsymbol{y}_T) = \boldsymbol{0}$ or $\overline{\boldsymbol{H}}_T(\boldsymbol{y}_T) = \boldsymbol{0}$ is a system of linear equations. After convergence in each time step is achieved, ω^M, w_t^M, $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$ are updated.

5. Numerical tests

The three algorithms presented above are implemented in MATLAB, version 5.3. The problems are scaled such that the length unit (mm) is used everywhere and $r = 10^5$ is chosen in the projections. The parameters β, γ and $\overline{m}$ appearing in Algorithm MBN are set to $\beta = 0.9$, $\gamma = 0.1$ and $\overline{m} = 22$. Furthermore, observe that $\frac{1}{2}\boldsymbol{H}_C^{\mathrm{T}}(\boldsymbol{y}_C)\boldsymbol{H}_C(\boldsymbol{y}_C) \leqslant \epsilon$ is used as a stopping criterion in all algorithms, where ϵ is set to 10^{-10}.

The algorithms have been used to solve a number of two-dimensional problems. For instance, the examples that were solved in (Strömberg, 1998) were solved here using Algorithm MBN, without any of the numerically difficulties that were reported by Strömberg, see the introduction.

In this section two other examples are defined and solved. The examples are referred to Examples 1 and 2, respectively. The examples differ in the boundary conditions and tribological parameters. The major difference is the type of load history; in Example 1 the load history is governed by prescribed displacements and in Example 2 by prescribed temperatures. Furthermore, the examples are solved for two different settings of bulk properties, corresponding to steel and aluminum. Results from numerical tests are presented and discussed from two different points of view: (i) the numerical performance of the three algorithms, and (ii) the behaviour of the model based on the numerical solution.

5.1. The two examples

Consider two thermoelastic bodies as shown in Fig. 2. The upper body is denoted punch and the lower body is denoted foundation. The potential contact surface is defined by the lower end of the punch.

The dimensions of the punch are 20×4 (mm^2) and the dimensions of the foundation are 40×20 (mm^2). The properties of the punch is approximated using 40×8 bilinear elements and the foundation using 60×20 bilinear elements. Plane strain is assumed and the thickness is set to 1 (mm). The potential contact surface is divided into 40 contact elements such that node to node contact is obtained between the punch and the foundation.

In both examples the upper end of the punch is subjected to uniformly distributed tractions $\hat{\boldsymbol{t}}(t) = -200\boldsymbol{e}_2$ (N/mm^2). The load $\hat{\boldsymbol{t}}$ is applied in an initial set of 20 time steps. Furthermore, the lower end of the foundation is locked in the $\boldsymbol{e}_2$-direction and the midpoint is also locked in the $\boldsymbol{e}_1$-direction, see Fig. 2.

Moreover, the thermal contact conductances are assumed to be equal for the two contact surfaces and set to be $\vartheta^1 = \vartheta^2 = 10^{-3}$ (W/N $\cdot$ K). This is a value for steel which was adapted from experimental curves given in (Fried, 1969) by (Johansson and Klarbring, 1993). A consequence of this assumption is that the heat generated by friction and wear is divided equally between the two bodies, see Eq. (12). The initial contact gap is set to $g(x) = 0.0005x^2$, where x is the coordinate along the contact surface with its origin at the midpoint of the contact surface.

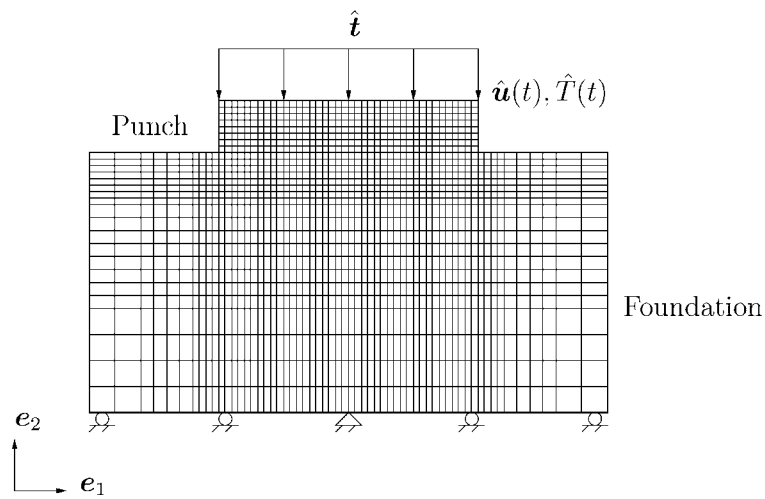

Fig. 2. Geometry and boundary conditions of Examples 1 and 2.

Table 1
Bulk properties of the two different materials

	Steel	Aluminum
$\tilde{\lambda}$ [Pa]	$1.1885 \cdot 10^{11}$	$4.0384 \cdot 10^{10}$
$\tilde{\mu}$ [Pa]	$7.9231 \cdot 10^{10}$	$2.6923 \cdot 10^{10}$
α [1/K]	$12 \cdot 10^{-6}$	$24 \cdot 10^{-6}$
ρ [kg/m^3]	7800	2750
c [J/kg $\cdot$ K]	460	960
k_t [W/m $\cdot$ K]	46	170

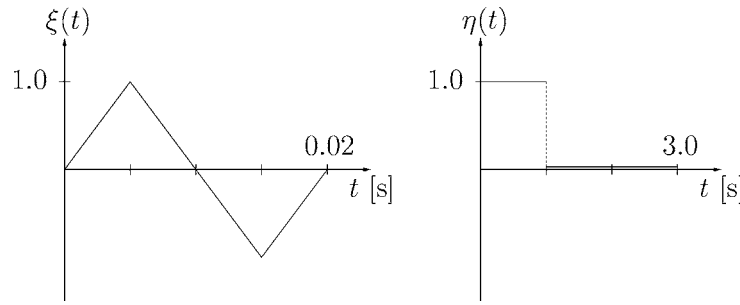

Fig. 3. One cycle of the prescribed displacement and temperature, respectively.

The two sets of constitutive parameters, defining the two different bulk materials, corresponding to steel and aluminum, respectively, are given in Table 1.

In Example 1, the upper end of the punch is in addition subjected to the prescribed displacement given by $\hat{\boldsymbol{u}}(t) = 0.15\xi(t)\boldsymbol{e}_1$ (mm), where one cycle of $\xi(t)$ is given in Fig. 3. The displacement cycle is divided into 8 time steps. Furthermore, all boundaries except for the contact surfaces are assumed to be thermally insulated. In this example the friction coefficient is set to $\mu = 0.3$ and the wear coefficient to $k = 10^{-11}$ (mm^2/N).

In Example 2, the upper end of the punch is fixed in the direction $\boldsymbol{e}_1$. In addition, the upper end is also subjected to a prescribed cyclic deviation from the reference temperature, $\widehat{T}(t) = 400\eta(t)$ (K), according to Fig. 3 and zero prescribed deviation from the reference temperature at the lower end of the foundation. All the other boundaries except the contact surfaces are thermally insulated. The temperature cycle is divided into 24 time steps. In this example the friction coefficient is set to $\mu = 0.2$ and the wear coefficient to $k = 10^{-9}$ (mm^2/N).

Although it could be argued that the assumptions of quasi-statics, small displacements, node to node contact and small deviations from a reference temperature are violated in the examples defined above, we believe that they serve the purpose to compare the performance of the different algorithms and to show some interesting aspects of the thermoelastic model of contact, friction and wear.

5.2. Comparison of the three algorithms

In this subsection we compare the numerical performance of the three algorithms based on execution statistics from 100 cycles for the two example problems defined above using steel as bulk properties.

Tables 2 and 3 show the average values over 100 cycles of the number of Gauss–Seidel iterations per time increment (GS/inc.), the number of Newton iterations per time increment (Newt./inc.), the number of line-searches per Newton iteration (Arm./Newt.) and the measured CPU-time per time increment (CPU/inc.) normalized to the CPU-time obtained using Algorithm MBN.

The execution statistics in Tables 2 and 3 indicate that algorithm GSTM is always more efficient than Algorithm MBN. This can be concluded partly from the fact that the size of the linear system of equations to be solved at each Newton iteration of the mechanical subproblem is $2/3$ of the size of the corresponding system of equations in the direct approach. Furthermore, due to the linearity of the thermal equations, these are only solved once in each Gauss–Seidel iteration and it is found that often only one Newton iteration is needed the second time the mechanical subproblem is solved. This means that efficiency is gained by using an iterative approach as long as the number of Gauss–Seidel iterations is small. When the number of Gauss–Seidel iterations becomes large compared to the number of Newton iterations in the direct approach this gain might be lost. This is the case for Algorithm GSMT in Example 2.

Table 2
Execution statistics for Example 1

Algorithm	GS/inc.	Newt./inc.	Arm./Newt.	CPU/inc.
MBN	–	5.27	1.46	1.00
GSMT	3.80	8.92	1.27	0.98
GSTM	3.85	8.92	1.27	0.97

Table 3
Execution statistics for Example 2

Algorithm	GS/inc.	Newt./inc.	Arm./Newt.	CPU/inc.
MBN	–	3.98	1.04	1.00
GSMT	3.81	7.53	1.14	1.13
GSTM	3.52	6.43	1.11	0.98

Table 4
Execution statistics for Example 1 without wear

Algorithm	GS/inc.	Newt./inc.	Arm./Newt.	CPU/inc.
MBN	–	6.08	1.40	1.00
GSMT	3.61	7.86	1.41	0.70
GSTM	3.61	7.87	1.37	0.69

Table 5
Execution statistics for Example 1 with 81 contact nodes

Algorithm	GS/inc.	Newt./inc.	Arm./Newt.	CPU/inc.
MBN	–	5.34	2.00	1.00
GSMT	3.84	9.17	1.58	0.84
GSTM	3.87	9.12	1.59	0.84

Table 6
Execution statistics for Example 2 with 81 contact nodes

Algorithm	GS/inc.	Newt./inc.	Arm./Newt.	CPU/inc.
MBN	–	4.75	1.15	1.00
GSMT	3.77	8.84	1.30	0.93
GSTM	3.54	7.41	1.21	0.80

One should also note that the performance of the two iterative strategies depends on the type of boundary conditions. For Example 1 GSMT and GSTM performs equally, while for Example 2 it is slightly better to solve the thermal subproblems first, i.e. using Algorithm GSTM.

The pattern indicated above is essentially the same when the number of time steps, the number of contact nodes or the tolerance in the stop criterion is changed. One might note that the difference between the direct approach (MBN) and the iterative approaches (GSMT and GSTM) is decreased when the number of time steps is increased while the difference is increased when the number of contact nodes is increased, when ϵ is assigned a higher value or when wear is excluded from the model. For instance, if wear is excluded from the model the execution satistics in Table 4 is obtained for Example 1.

Futhermore, if the number of contact nodes is doubled the execution statistics in Tables 5 and 6 are obtained for Example 1 and Example 2, respectively. By comparing these statistics by those presented in Tables 2 and 3, respectively, it is seen that this clearly follows the trend outlined above. Hence, it seems that algorithm GSTM is the preferable one with respect to computing time.

5.3. Results for different constitutive settings

In this subsection we present and discuss some numerical results for the example problems. Let us first consider Example 1. In Table 7 twelve different settings of the problem is defined, differing in bulk properties of the bodies and whether wear and thermal effects are included or not.

Table 7
Different constitutive settings

Punch	Found.	Thermal	Wear
Steel	Steel	yes	no
Steel	Steel	yes	yes
Steel	Steel	no	yes
Steel	Alum.	yes	no
Steel	Alum.	yes	yes
Steel	Alum.	no	yes
Alum.	Steel	yes	no
Alum.	Steel	yes	yes
Alum.	Steel	no	yes
Alum.	Alum.	yes	no
Alum.	Alum.	yes	yes
Alum.	Alum.	no	yes

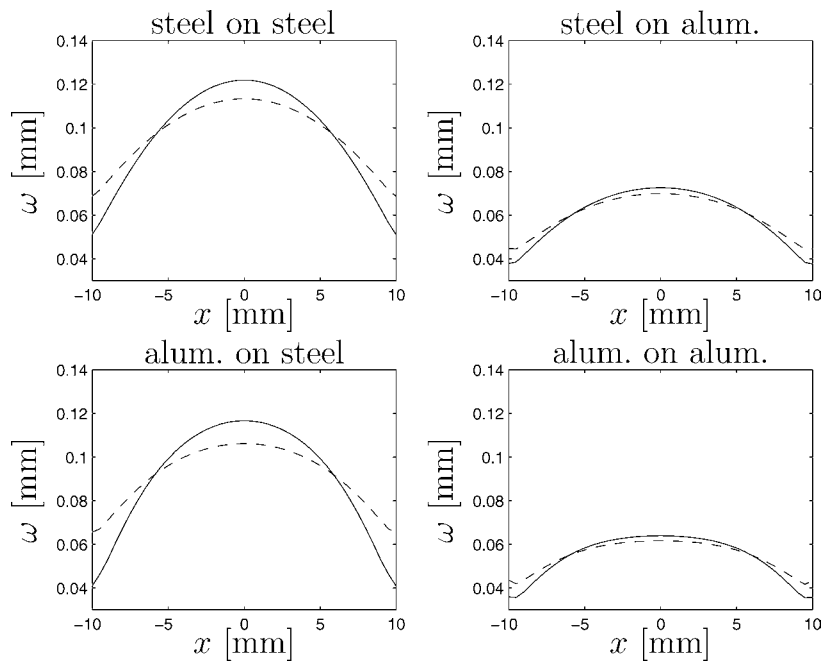

Fig. 4. Wear gap after 100 cycles when thermal effects are included (solid line) and excluded (dashed line).

Figure 4 shows the wear gap after 100 cycles when thermal effects are included in the model (solid line) or not (dashed line), for each combination of bulk materials. In the same fashion Fig. 5 shows the intrinsic temperature, see Eq. (11), of the potential contact surface, when wear is included in the model (solid line) or not (dashed line). The major point illustrated by Fig. 4 is that the calculated wear gap might be quite different in shape as well as in maximum depth, depending only on whether thermal dilatation of the bodies are taken into account or not. One might also conclude from Fig. 4 that less material is worn away when the foundation is made of aluminum. This is due to larger elastic deformations of the foundation compared to the cases when the foundation is made of steel. Thus, the difference must not be explained by a different wear coefficient. This fact should be taken in account when wear coefficients are experimentally determined. Figure 5 shows that the temperature of the contact surface is much lower for the cases when the foundation is made of aluminum. This is due to smaller dissipation because of the larger elastic deformations but is also influenced by the faster transport of heat in aluminum than in steel. This also has the effect that the influence of wear on the temperature is less pronounced when the foundation is made of aluminum.

We finish this subsection by presenting some numerical results for Example 2, where the constitutive parameters for steel have been used for both the punch and the foundation. In this problem we have a situation of stick-slip fretting induced by

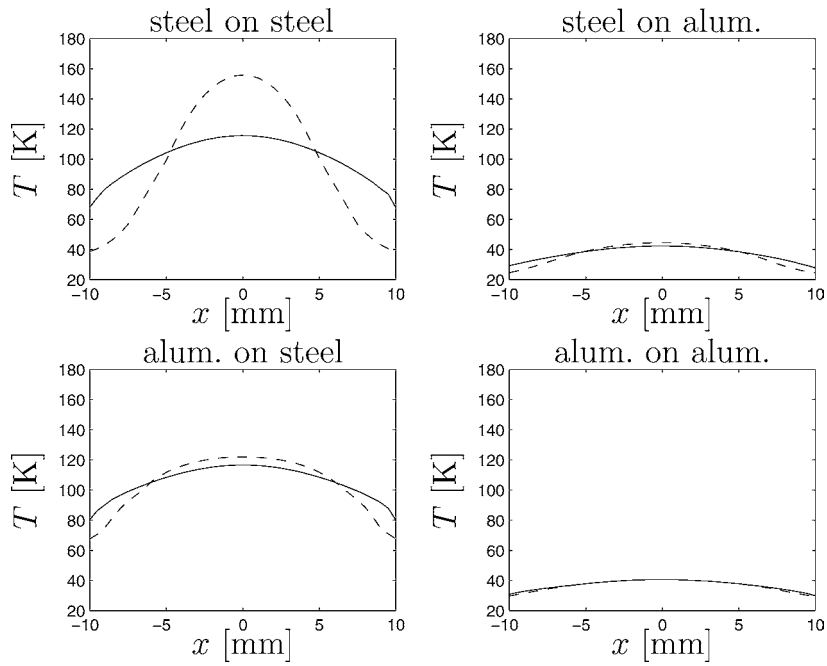

Fig. 5. The intrinsic temperature of the punch after 100 cycles when wear is included (solid line) and excluded (dashed line).

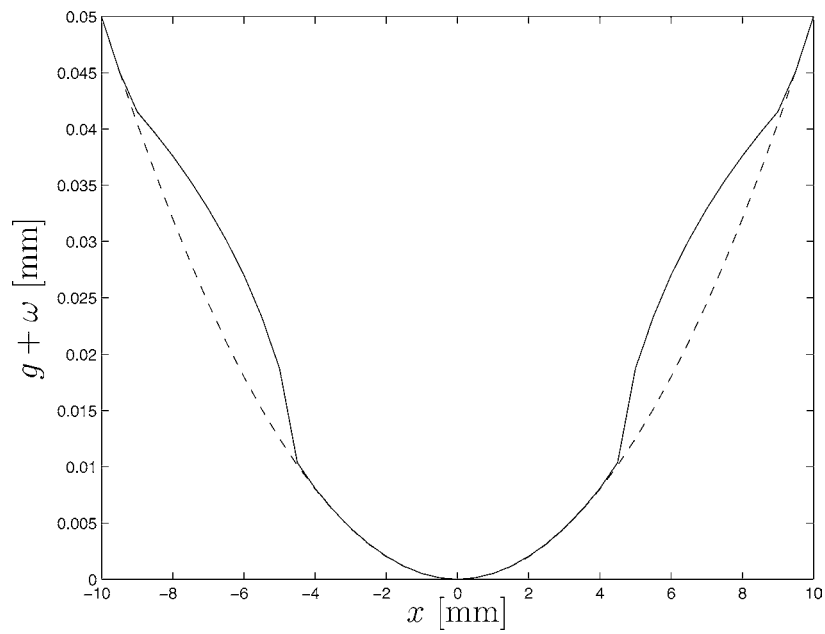

Fig. 6. Initial contact gap (dashed line) and contact gap after 100 cycles (solid line) for Example 2.

thermal boundary conditions. Figure 6 shows the initial contact gap (dashed line) and the contact gap after 100 cycles (solid line) for Example 2. Figure 7 shows the normal contact traction (solid line) and the tangential contact traction (dashed line) after 0, 1, 10 and 100 cycles, respectively. Although the amount of wear might seem small, it still has a major influence on the normal contact pressure.

In (Ciavarella and Hills, 1999) it was shown for similar geometries that the region of stick remains unchanged if a prescribed tangential load varies between two fixed limits, even when wear is developed in the slip region. They also showed that under

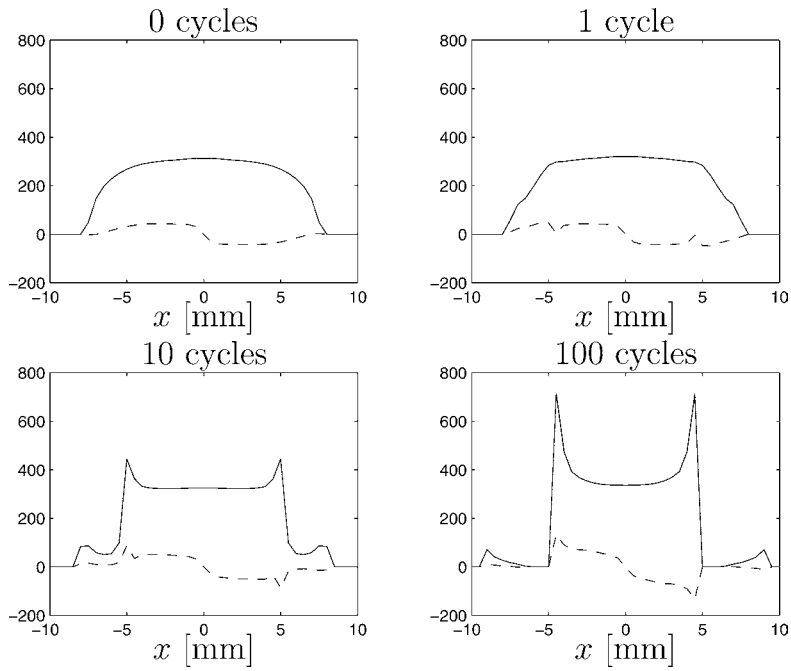

Fig. 7. Normal (solid line) and tangential (dashed line) contact traction after 0, 1, 10 and 100 cycles.

these conditions wear will develope such that a steady-state of full stick is reached. In our example we have no prescribed tangential load. However, one can see that the effect of prescribed temperature varying between two fixed limits is the same as described in (Ciavarella and Hills, 1999).

6. Concluding remarks

In the present work a finite-element model for treating unilateral frictional wearing contact between two isotropic linearly thermoelastic bodies is presented. Two approaches to handle the resulting system of semismooth equations are suggested. The first approach is a direct use of a modified Newton method for B-differentiable equations on the fully-coupled problem, while in the second approach the mechanical and thermal problems are solved uncoupled from each other using an iterative strategy. The latter approach might be viewed as a strategy of Gauss–Seidel type. This approach is applied in two different ways, either by solving the mechanical problem first or the thermal problem first.

It is found that the iterative approach where the thermal subproblem is solved first seems always more efficient than the direct approach and at least as efficient as the iterative approach where the mechanical subproblem is solved first, no matter if the problem is controlled by a varying prescribed displacement or a varying prescribed temperature. It is also found that the gain in efficiency by using an iterative approach is increased with increasing number of contact nodes. Based on these observations it is conluded that the iterative strategy where the thermal subproblem is solved first is prefeable, with respect to computing time.

Furthermore, some results that shows some interesting aspects of the behaviour of the model is presented. It is shown numerically that the wear gap might be quite different depending on whether thermal dilatation is taken into account or not. It is also shown how the wear gap and the temperature at the contact surface is influenced by the bulk properties of the contacting bodies. For instance, it is shown that if only the material in the support is switched from steel to aluminum, keeping all other parameters fixed, the amount of wear is decreased. Thus, the difference in wear rate is not explained by a difference in wear coefficient, but instead is explained by a difference in elastic response.

Acknowledgements

This work was supported by the Swedish Research Council (VR). We are grateful to Dr. Peter W. Christensen for his comments on the section concerning semismoothness.

Appendix A. Static condensation

The equilibrium equations (16), (17) and the balance of energy, Eq. (27), may be written, for each body, respectively, as

$$\begin{bmatrix} \overline{\boldsymbol{K}}_l^{cc} & \overline{\boldsymbol{K}}_l^{cr} \\ \overline{\boldsymbol{K}}_l^{rc} & \overline{\boldsymbol{K}}_l^{rr} \end{bmatrix} \begin{Bmatrix} \boldsymbol{d}_l^c \\ \boldsymbol{d}_l^r \end{Bmatrix} - \begin{bmatrix} \widehat{\boldsymbol{K}}_l^{cc} & \widehat{\boldsymbol{K}}_l^{cr} \\ \widehat{\boldsymbol{K}}_l^{rc} & \widehat{\boldsymbol{K}}_l^{rr} \end{bmatrix} \begin{Bmatrix} \boldsymbol{T}_l^c \\ \boldsymbol{T}_l^r \end{Bmatrix} - \begin{Bmatrix} \boldsymbol{F}_l^c \\ \boldsymbol{F}_l^r \end{Bmatrix} + \begin{Bmatrix} \boldsymbol{P}_l \\ \boldsymbol{0} \end{Bmatrix} = \begin{Bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{Bmatrix} \quad (l = 1, 2) \quad \text{and} \tag{A.1}$$

$$\begin{bmatrix} \boldsymbol{A}_l^{cc} & \boldsymbol{A}_l^{cr} \\ \boldsymbol{A}_l^{rc} & \boldsymbol{A}_l^{rr} \end{bmatrix} \begin{Bmatrix} \boldsymbol{T}_l^c \\ \boldsymbol{T}_l^r \end{Bmatrix} - \begin{Bmatrix} \boldsymbol{Q}_l^c \\ \boldsymbol{Q}_l^r \end{Bmatrix} - \begin{Bmatrix} \overline{\boldsymbol{L}}_l^c \\ \boldsymbol{0} \end{Bmatrix} = \begin{Bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{Bmatrix} \quad (l = 1, 2), \tag{A.2}$$

where $\boldsymbol{P}_1 = -\boldsymbol{P}_2 = \overline{\boldsymbol{C}}_n^{\mathrm{T}} \boldsymbol{P}_n + \overline{\boldsymbol{C}}_t^{\mathrm{T}} \boldsymbol{P}_t$, $\boldsymbol{A}_l = \boldsymbol{M}_l + (t_{k+1} - t_k)\boldsymbol{O}_l$, superscripts c denotes degrees of freedom on Γ_c and superscripts r denotes degrees of freedom to be reduced.

Solving the second equation of (A.2) for $\boldsymbol{T}_l^r$ yields

$$\boldsymbol{T}_l^r = \left(\boldsymbol{A}_l^{rr}\right)^{-1} \left(\boldsymbol{Q}_l^r - \boldsymbol{A}_l^{rc} \boldsymbol{T}_l^c\right). \tag{A.3}$$

By inserting (A.3) into the first equation of (A.2) one finds

$$\boldsymbol{A}_l^c \boldsymbol{T}_l^c - \overline{\boldsymbol{Q}}_l^c - \overline{\boldsymbol{L}}_l^c = \boldsymbol{0}, \quad \text{where} \quad \boldsymbol{A}_l^c = \boldsymbol{A}_l^{cc} - \boldsymbol{A}_l^{cr} \left(\boldsymbol{A}_l^{rr}\right)^{-1} \boldsymbol{A}_l^{rc} \quad \text{and} \quad \overline{\boldsymbol{Q}}_l^c = \boldsymbol{Q}_l^c - \boldsymbol{A}_l^{cr} \left(\boldsymbol{A}_l^{rr}\right)^{-1} \boldsymbol{Q}_l^r. \tag{A.4}$$

Solving the second equation of (A.1) for $\boldsymbol{d}_l^r$ yields

$$\boldsymbol{d}_l^r = \left(\overline{\boldsymbol{K}}_l^{rr}\right)^{-1} \left(-\overline{\boldsymbol{K}}^{rc} \boldsymbol{d}_l^c + \widehat{\boldsymbol{K}}_l^{rc} \boldsymbol{T}_l^c + \widehat{\boldsymbol{K}}_l^{rr} \boldsymbol{T}_l^r + \boldsymbol{F}_l^r\right). \tag{A.5}$$

By inserting (A.3) and (A.5) into the first equation of (A.1) one finds that

$$\begin{aligned} &\overline{\boldsymbol{K}}_l^c \boldsymbol{d}_l^c - \widehat{\boldsymbol{K}}_l^c \boldsymbol{T}_l^c - \overline{\boldsymbol{F}}_l^c + \boldsymbol{P}_l = \boldsymbol{0}, \quad \text{where} \\ &\overline{\boldsymbol{K}}_l^c = \overline{\boldsymbol{K}}_l^{cc} - \overline{\boldsymbol{K}}_l^{cr} \left(\overline{\boldsymbol{K}}_l^{rr}\right)^{-1} \overline{\boldsymbol{K}}_l^{rc}, \\ &\widehat{\boldsymbol{K}}_l^c = \widehat{\boldsymbol{K}}_l^{cc} - \overline{\boldsymbol{K}}_l^{cr} \left(\overline{\boldsymbol{K}}_l^{rr}\right)^{-1} \widehat{\boldsymbol{K}}_l^{rc} + \overline{\boldsymbol{K}}_l^{cr} \left(\overline{\boldsymbol{K}}_l^{rr}\right)^{-1} \widehat{\boldsymbol{K}}_l^{rr} \left(\boldsymbol{A}_l^{rr}\right)^{-1} \boldsymbol{A}_l^{rc} - \widehat{\boldsymbol{K}}_l^{rc} \left(\boldsymbol{A}_l^{rr}\right)^{-1} \boldsymbol{A}_l^{rc} \quad \text{and} \\ &\overline{\boldsymbol{F}}_l^c = \boldsymbol{F}_l^c - \overline{\boldsymbol{K}}_l^{cr} \left(\overline{\boldsymbol{K}}_l^{rr}\right)^{-1} \boldsymbol{F}_l^r + \widehat{\boldsymbol{K}}_l^{cr} \left(\boldsymbol{A}_l^{rr}\right)^{-1} \boldsymbol{Q}_l^r. \end{aligned} \tag{A.6}$$

In order to further reduce the number of unknowns we introduce the relative displacement $\bar{\boldsymbol{d}} = \boldsymbol{d}_1^c - \boldsymbol{d}_2^c$. Equation (A.6) may now be written as

$$\begin{bmatrix} \overline{\boldsymbol{K}}_1^c & \overline{\boldsymbol{K}}_1^c \\ \boldsymbol{0} & \overline{\boldsymbol{K}}_2^c \end{bmatrix} \begin{Bmatrix} \bar{\boldsymbol{d}} \\ \boldsymbol{d}_2^c \end{Bmatrix} - \begin{bmatrix} \widehat{\boldsymbol{K}}_1^c & \boldsymbol{0} \\ \boldsymbol{0} & \widehat{\boldsymbol{K}}_2^c \end{bmatrix} \begin{Bmatrix} \boldsymbol{T}_1^c \\ \boldsymbol{T}_2^c \end{Bmatrix} - \begin{Bmatrix} \overline{\boldsymbol{F}}_1^c \\ \overline{\boldsymbol{F}}_2^c \end{Bmatrix} + \begin{Bmatrix} \boldsymbol{P} \\ -\boldsymbol{P} \end{Bmatrix} = \begin{Bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{Bmatrix}, \quad \text{where} \tag{A.7}$$

$$\begin{Bmatrix} \boldsymbol{d}_1^c \\ \boldsymbol{d}_2^c \end{Bmatrix} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{I} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \begin{Bmatrix} \bar{\boldsymbol{d}} \\ \boldsymbol{d}_2^c \end{Bmatrix}$$

and $\boldsymbol{P} = \boldsymbol{P}_1 = -\boldsymbol{P}_2$ have been used, and, finally, $\boldsymbol{I}$ is an identity matrix of appropriate size.

By adding the first equation of (A.7) to the second, one finds that

$$\begin{bmatrix} \overline{\boldsymbol{K}}_1^c & \overline{\boldsymbol{K}}_1^c \\ \overline{\boldsymbol{K}}_1^c & \overline{\boldsymbol{K}}_1^c + \overline{\boldsymbol{K}}_2^c \end{bmatrix} \begin{Bmatrix} \bar{\boldsymbol{d}} \\ \boldsymbol{d}_2^c \end{Bmatrix} - \begin{bmatrix} \widehat{\boldsymbol{K}}_1^c & \boldsymbol{0} \\ \widehat{\boldsymbol{K}}_1^c & \widehat{\boldsymbol{K}}_2^c \end{bmatrix} \begin{Bmatrix} \boldsymbol{T}_1^c \\ \boldsymbol{T}_2^c \end{Bmatrix} + \begin{Bmatrix} \overline{\boldsymbol{F}}_1^c \\ \overline{\boldsymbol{F}}_1^c + \overline{\boldsymbol{F}}_2^c \end{Bmatrix} + \begin{Bmatrix} \boldsymbol{P} \\ \boldsymbol{0} \end{Bmatrix} = \begin{Bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{Bmatrix}. \tag{A.8}$$

Solving the second equation of (A.8) for $\boldsymbol{d}_2^c$ yields

$$\boldsymbol{d}_2^c = \left(\overline{\boldsymbol{K}}_1^c + \overline{\boldsymbol{K}}_2^c\right)^{-1} \left(-\overline{\boldsymbol{K}}_1^c \bar{\boldsymbol{d}} + \widehat{\boldsymbol{K}}_1^c \boldsymbol{T}_1^c + \widehat{\boldsymbol{K}}_2^c \boldsymbol{T}_2^c + \overline{\boldsymbol{F}}_1^c + \overline{\boldsymbol{F}}_1^c\right).$$

By inserting this into the first equation of (A.8) one finds that

$$\begin{aligned} &\bar{\boldsymbol{\kappa}} \bar{\boldsymbol{d}} - \hat{\boldsymbol{\kappa}}_1 \boldsymbol{T}_1^c - \hat{\boldsymbol{\kappa}}_2 \boldsymbol{T}_2^c - \boldsymbol{R} + \boldsymbol{P} = \boldsymbol{0}, \quad \text{where} \\ &\bar{\boldsymbol{\kappa}} = \overline{\boldsymbol{K}}_1^c - \overline{\boldsymbol{K}}_1^c (\overline{\boldsymbol{K}}_1^c + \overline{\boldsymbol{K}}_2^c)^{-1} \overline{\boldsymbol{K}}_1^c, \qquad \hat{\boldsymbol{\kappa}}_1 = \widehat{\boldsymbol{K}}_1^c - \overline{\boldsymbol{K}}_1^c (\overline{\boldsymbol{K}}_1^c + \overline{\boldsymbol{K}}_2^c)^{-1} \widehat{\boldsymbol{K}}_1^c, \qquad \hat{\boldsymbol{\kappa}}_2 = -\overline{\boldsymbol{K}}_1^c (\overline{\boldsymbol{K}}_1^c + \overline{\boldsymbol{K}}_2^c)^{-1} \widehat{\boldsymbol{K}}_2^c \quad \text{and} \\ &\boldsymbol{R} = \overline{\boldsymbol{F}}_1^c - \overline{\boldsymbol{K}}_1^c (\overline{\boldsymbol{K}}_1^c + \overline{\boldsymbol{K}}_2^c)^{-1} (\overline{\boldsymbol{F}}_1^c + \overline{\boldsymbol{F}}_2^c). \end{aligned} \tag{A.9}$$

References

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

Chaney, R.W., 1990. Piecewise C^k functions in nonsmooth analysis. Nonlinear Anal., Theory, Methods Appl. 15, 649–660.

Christensen, P.W., Klarbring, A., Pang, J.S., Strömberg, N., 1998. Formulation and comparison of algorithms for frictional contact problems. Int. J. Numer. Methods Engng. 42, 145–173.

Christensen, P.W., Pang, J.S., 1998. Frictional contact algorithms based on semismoooth Newton methods, in: Fukushima, M., Qi, L. (Eds.), Reformulation – Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods. Kluwer, Dordrecht, pp. 81–116.

Christensen, P.W., 2000. Computational nonsmooth mechanics – contact, friction and plasticity. Dissertations No. 657, Linköping University, pp. 174–198.

Ciavarella, M., Hills, D.A., 1999. Brief note: Some observations on oscillating tangential forces and wear in general plane contacts. Eur. J. Mech. A 18, 491–497.

Fried, E., 1969. Thermal conduction contribution to heat transfer at contacts, in: Tye, R.P. (Ed.), Thermal Conductivity, Vol. 2. Academic Press, London, pp. 253–274.

Johansson, L., Klarbring, A., 1993. Thermoelastic frictional contact problems: modelling, FE-approximation and numerical realization. Comput. Methods Appl. Mech. Engng. 105, 181–210.

Johansson, L., Klarbring, A., 2000. Study of frictional impact using a nonsmooth equations solver. J. Appl. Mech. 67, 267–273.

Klarbring, A., 1992. Mathematical programming and augmented Lagrangian methods for frictional contact problems, in: Proc. Contact Mechanics Int. Symp., pp. 409–422.

Lundvall, O., Klarbring, A., 2000. Prediction of transmission error in spur gears as a consequence of wear using FEM. Submitted for publication.

Ortega, J.M., Reinboldt, W.C., 1970. Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York.

Pang, J.S., 1990. Newton's method for B-differetiable equations. Math. Oper. Res. 15, 311–341.

Strömberg, N., Johansson, L., Klarbring, A., 1996. Derivation and analysis of a generalized standard model for contact, friction and wear. Internat. J. Solids Structures 33, 1817–1836.

Strömberg, N., 1997. An augmented Lagrangian method for fretting problems. Eur. J. Mech. A 16, 573–593.

Strömberg, N., 1998. Finite element treatment of two-dimensional thermoelastic wear problems. Comput. Methods Appl. Mech. Engng. 177, 441–455.

Strömberg, N., 1999. A Newton method for three-dimensional fretting problems. Internat. J. Solids Structures 36, 2075–2090.

Wriggers, P., Miehe, C., 1994. Contact constraints within coupled thermomechanical analysis – a finite element model. Comput. Methods Appl. Mech. Engng. 113, 301–319.