Advances in computational contact mechanics

S. A. Meguid · A. Czekanski

Abstract Great strides have recently been made in the application of computational mechanics to the design of highly complex engineering systems. It has now become abundantly clear that advanced modelling techniques are central to the competitiveness of the industrialised nations. Excellent examples of this assertion are the computer-integrated design of the recent Boeing 777 aircraft, the collapsible foam-filled structures for the car of the next century and new prosthetic implants for Rheumatoid Arthritis. It is with this in mind that the author focuses his attention to a class of problems where contact mechanics plays a major role in dictating the mechanical integrity of the component/system. Three aspects of the current study are accordingly examined. The first is concerned with the development of the appropriate dynamic variational inequalities expressions, which are capable of the accurate and consistent representation of contact problems. The second is concerned with the development of robust solution algorithms that guarantee the accurate imposition of the kinematic contact constraint and avoid interpenetration. The third is concerned with the application of the developed algorithms to realistic design problems involving intricate mechanical and biomechanical systems.

1 Introduction

Contact problems play an important role in dictating the integrity, performance and safety of many engineering systems/components involved in vehicle design, armament and ballistics, metal forming/cutting, and surface treatments, just to name a few. Despite their importance to the mechanical integrity of the systems examined, contact effects are frequently treated using oversimplifying assumptions, which neglect the main feature of the problem. The reason is that modelling dynamic contact in solids poses mathematical and computational difficulties. With the application of loads to the bodies in contact, the actual surface on which these bodies meet, change with time, and the stresses at the surfaces are generally unknown and complex to determine.

The subject of contact mechanics may be said to have started in 1882 with the pioneering publication of Hertz (Hertz 1882). His theory of contact was developed for elastic smooth frictionless bodies, with the contact region being small compared with the

S. A. Meguid (✉) · A. Czekanski
Mechanics and Aerospace Design Laboratory,
Department of Mechanical and Industrial Engineering,
University of Toronto, 5 King's College Road, Toronto,
ON, Canada M5S 3G8
e-mail: meguid@mie.utoronto.ca

A. Czekanski
e-mail: Czekanski@mie.utoronto.ca

dimensions of the bodies. Hertz formulated the conditions of contact by considering two elastic bodies. In order to obtain expressions for the size of the contact zone and the value of the contact pressure, several simplifications and restrictions were imposed. These include:

(i)   the contacting surfaces must be smooth,
(ii) the strains and displacements adhere to the small strain theory of elasticity,
(iii) the stress, in each of the half-space bodies considered, is related to the strain by Hooke's law,
(iv) the surfaces of the elastic half-space bodies considered are frictionless, and
(v)   the bodies are in static contact.

Despite its limitations, Hertz's theory of contact has stood the test of time and has been a landmark in applied mechanics for many decades. More representative analytical solutions of contact problems, mostly developed in the second half of the twentieth century, have been largely associated with the removal of some of these restrictions (see e.g. Johnson 1985, Gladwell 1980). Expressions for contact stresses, size of contact region and displacements at the contact surface were derived for a large number of simplified cases.

In spite of the progress so far made in analytical contact mechanics, the proper treatment of the general dynamic frictional contact problem is still beyond the reach of these analytical solutions. In order to overcome these limitations, most contact problems are currently being treated using computational techniques, with the finite element method being the most appealing.

2 Computational contact mechanics

With the rapid development in the capabilities of digital computers, numerical methods became the prime interest of researchers. Several formulations have been proposed to treat frictional contact problems using the Boundary Element Method (BEM) (Aliabadi and Brebbia 1993) and the Finite Element Method (FEM) (Bathe 1996; Crisfield 1997; Hughes 1987; Zienkiewicz and Taylor 1988; Zhong 1993). However, the FEM is now considered the most favourable, because of its proven success in treating a wide range of engineering problems in areas of mechanics of solids, fluid flow, heat transfer, electromagnetism and coupled field problems. In addition, the BEM becomes computationally inefficient when treating nonlinear problems requiring the calculation of the field variables inside the domain (Aliabadi and Brebbia 1993).

Original FE formulations did not accommodate the treatment of contact and ad hoc techniques were used. In these cases, simplifying assumptions regarding the actual contact surface and the distribution of the contact stress were made. These simplifying assumptions enabled the treatment of each individual body as a separate problem, see for example (Kenny et al. 1991). This ad hoc technique proved inadequate in many cases where neither the contact surface nor the stresses on it could be easily estimated. This has prompted the development of contact elements.

FE methods treat contact problems by extending the variational formulation upon which the FE method is based. Contact elements are formulated and assembled into the original FE code in order to enforce the contact conditions (Böhm 1987). The solution is then obtained by solving the resulting set of nonlinear equations. A large number of contact element formulations has appeared in the literature and has been implemented in a number of commercial FE packages over the last two decades (see for example Chaudhary and Bathe 1985). The solution techniques adopted in these formulations were based upon using either function method or Lagrange multipliers in identifying the contact surface and imposing the contact constraints.

Figure 1 shows a typical example of such contact elements. This particular element is adopted in several FE packages including ANSYS (ANSYS 1999) and MARC (MARC 1993). In this case, the element is based on two stiffness values. They are the normal contact stiffness K_N and the tangential contact stiffness K_T.

The normal stiffness K_N is used to penalize interpenetration between the two bodies, while the tangential stiffness K_T is used to approximate the sudden jump in the tangential force, as represented by Coulomb's friction law when sliding is detected between the two contacting nodes.

Advantages of contact elements
Because of the simplicity of their formulation, contact elements enjoy the following advantages:
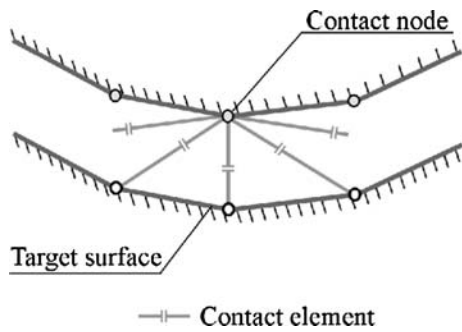
Fig. 1 A typical contact element

(i) they are simple to formulate,
(ii) they are easily accommodated into existing FE codes, and
(iii) they are easy to use.

However, experience with contact elements indicates that they suffer from several difficulties as indicated below.

Disadvantages of contact elements

Figure 2a, b show the variation of the normal and tangential stiffness, K_N and K_T, with displacement. A major problem in the implementation of contact element is the assignment of values to K_N and K_T.

The values of K_N and K_T are required to be very large. However, the use of excessively high values of K_N and K_T results in ill-conditioned global stiffness matrices, leading to numerical errors and divergence. On the other hand, the use of smaller values of K_N and K_T, results in convergence to the wrong solution allowing for interpenetration and wrong estimates of the stick and slip regions. In addition, they lack in mathematical rigor as evidenced by Böhm (1987) as they ignore the non-differentiability of Coulomb's friction law. Therefore, the difficulties with contact elements can be summarized as follows:
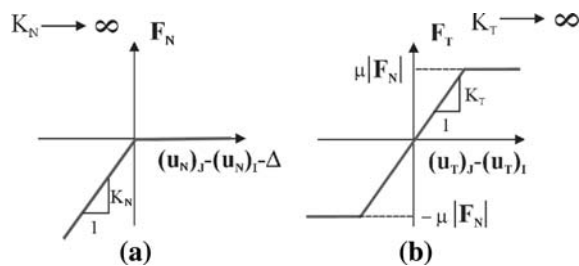
Fig. 2 Variation of contact stiffness with relative nodal displacement: (a) normal stiffness K_N and (b) tangential stiffness K_T

(i) their performance, in terms of convergence and accuracy, depends on user defined parameters,
(ii) they increase the size of the problem, and
(iii) they do not deal with the non-differentiability issue of Coulomb's friction forces.

In an attempt to overcome some of these difficulties, several formulations combined the method of Lagrange multipliers with that of the penalty function method. In this case, the active contact constraints are first identified by a penalty-type contact element and then the frictional forces are estimated and used with Lagrange multipliers to enforce contact conditions. Although these formulations are better than the above two approaches, they are still dependent upon user defined parameters which govern the convergence and accuracy of the solution. In addition, mathematical inconsistencies associated with the use of Coulomb friction forces are not resolved by combining the two techniques.

Because of the difficulties associated with traditional contact elements, current research efforts on the subject are focused on the formulation and implementation of the newly developed Variational Inequalities (VI) formulations.

2.1 Solution techniques

The exact variational representation of frictional contact problems results in a variational inequality. However, most FE methods treat contact problems by extending the variational formulation, which involve integrals over unknown contact surfaces (e.g., Hughes 1987; Zienkiewicz and Taylor 1988; Zhong 1993). Hughes et al. were the first to provide a detailed analysis of a class of impact frictionless problems (Hughes et al. 1976). In their method, Lagrange multipliers together with Newton-Raphson iterative procedure were utilised to solve the resulting dynamic expressions. Chaudhary and Bathe used Lagrange formulation to solve the static and dynamic contact problems accounting for friction (Chaudhary and Bathe 1986). Wriggers and Simo developed consistently linearised contact formulations for static problems (Wriggers and Simo 1985), while Wriggers et al. (1990) used a penalty formulation to solve these dynamic contact problems. Parisch developed consistent tangent stiffness matrices for treating quasi-static large deformation problems (Parisch 1989). Laursen and Simo developed an approach for treating

dynamic frictional contact problems which experience large deformations (Laursen and Simo 1993). In order to impose the kinematic contact conditions, they employed the penalty and augmented Lagrangian regularisation approaches.

The use of the variational method to formulate quasi-static or dynamic contact problems lacks in mathematical rigour, especially when frictional effects are taken into account. This is partly due to the non-differentiability of Coulomb's friction law, which is not properly addressed in the variational formulations. Furthermore, it usually results in the introduction of user-defined parameters, which influence the accuracy of the solution and the rate of convergence (see e.g., Böhm 1987; Refaat and Meguid 1994).

3 Variational inequalities formulations

The theory of variational inequalities is a relatively young mathematical discipline. It can be considered as an alternative mathematical description of physical problems, which is useful in problems involving constraints. A wide range of applications ranging from network economics to contact mechanics, are ideally formulated using variational inequalities (see e.g., Ferris and Pang 1997). However, the VI approach has not gained popularity in the engineering community because most of the work has appeared in the mathematical literature. The focus of the work was to examine the existence and uniqueness of the resulting inequalities (Adams 1975; Duvaut and Lions 1976; Glowinski et al. 1981; Hlavacek et al. 1988; Rabier et al. 1986; Kinderlehrer and Stampacchia 1980). Only in the last two decades interesting results related to contact have appeared in the engineering literature (see e.g., Kikuchi and Oden 1988; Oden and Martins 1983, 1985; Refaat and Meguid 1996).

The main basis for the development of VI formulations for contact problems was the works of Duvaut (Duvaut and Lions 1976) on the solution of the Signorini problem. Later on, Stampacchia laid the foundation of the theory itself (Kinderlehrer and Stampacchia 1980). Several authors examined the existence and uniqueness of the VI formulation of contact problems (see e.g., Kikuchi and Oden 1988; Klarbring 1988).

However, the literature indicates that little work has been carried out to develop suitable computational techniques to make use of these theoretical results. In this regard, dynamic elastic contact for small deformation was presented by Kikuchi and Oden (Kikuchi and Oden 1988; Oden and Martins 1983, 1985). They devoted their efforts to the mathematical questions concerning existence and uniqueness of the variational inequalities representing different contact problems. They also presented a solution technique based on the use of the penalty method, regularisation, and implicit Newmark time integration. Unfortunately, the resulting solution algorithm suffers from the same disadvantages as those outlined in the traditional penalty approach. A few other publications have devoted attention to the practical implementation of variational inequalities in static contact problems (see e.g., Bogomolny 1984; Bischoff 1984; Qin and He 1995; El-Abbasi 1999).

The current research study overcomes these difficulties by developing and implementing a new variational inequalities methodology to treat dynamic frictional contact in engineering structures.

In the unilateral (Signorini-type) problem, the contact constraints are shown in Fig. 3, and can be expressed as follows:

$$\text{before/after contact } u_N - g < 0 \Rightarrow \sigma_N = 0 \text{ \& } \dot{u}_N \text{ is unconstrained} \tag{1a}$$

$$\text{during contact } u_N - g = 0 \Rightarrow \sigma_N \leq 0 \text{ \& } \dot{u}_N \leq 0 \tag{1b}$$

These contact constraints state that: (i) the magnitude of the normal contact stress is less than or equal to zero, and (ii) the displacement of the contacting surfaces must not allow for interpenetration. Note that no constants are required to describe the contacting surfaces in the normal direction.

Coulomb's friction condition can be expressed as follows:

$$u_N < g \Rightarrow \sigma_T = \mathbf{0} \tag{2}$$

$$u_N = g \Rightarrow \begin{cases} |\sigma_T| \leq -\mu\sigma_N \\ |\sigma_T| < -\mu\sigma_N \Rightarrow \dot{\mathbf{u}}_T = \mathbf{0}, \\ |\sigma_T| = -\mu\sigma_N \Rightarrow \exists\lambda \geq 0, \quad \dot{\mathbf{u}}_T = -\lambda\boldsymbol{\sigma}_T. \end{cases} \tag{3}$$

where μ is the coefficient of friction and is assumed to be independent of velocity. This means that tangential stresses are only present if there is contact.
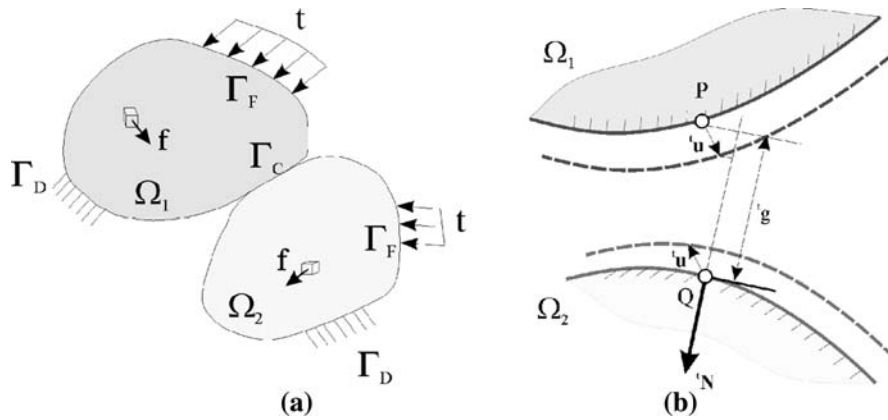
Fig. 3 Two bodies in contact: (**a**) unilateral contact model, and (**b**) schematic of kinematic contact condition

3.1 Dynamic VI formulation for elasto-plastic media

The dynamic contact problem for elastic-plastic media must satisfy the following equations:

(i) The equation of motion:

$$\frac{\partial^t \sigma_{ij}}{\partial^t x_j} + {}^t f_i = {}^t \rho \frac{\partial^2 ({}^t u_i)}{\partial t^2} \tag{4}$$

(ii) The constitutive law governing the behaviour of the material:

$$\sigma_{ij}(\mathbf{u}) = C_{ijkl} \varepsilon_{kl}(\mathbf{u}) \tag{5}$$

(iii) The initial conditions:

$$\mathbf{u}(0) = \mathbf{u}_o, \quad \dot{\mathbf{u}}(0) = \dot{\mathbf{u}}_o \tag{6}$$

(iv) The displacement and traction boundary conditions:

$$\mathbf{u}_i = \bar{\mathbf{u}}_o \quad \text{on } \Gamma_D \tag{7}$$

$$\sigma_i n_i = t_i \quad \text{on } \Gamma_F \tag{8}$$

(v) The kinematic contact and frictional conditions on Γ_C.

Restrictions imposed upon admissible displacement, velocity and stress field are not, in general, equalities but rather inequalities. Accordingly, the variational inequalities approach can be viewed as being the formulation of the principle of rate of virtual work in an inequality form. In the following, we will derive the variational inequality for dynamic frictional contact problems. The basic steps of the derivation are summarised in Fig. 4.

The elasto-plastic material considered here is assumed to experience large displacements and large strains. In this case, it is necessary to adopt the
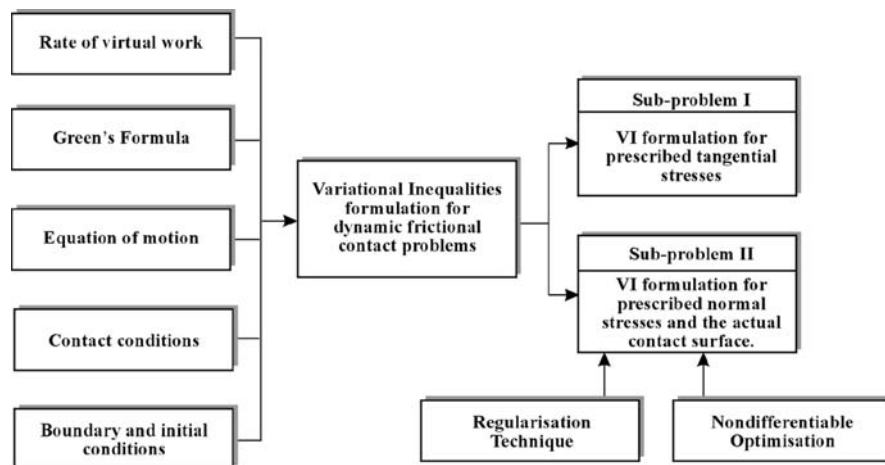
Fig. 4 Basic steps of derivation of dynamic VI formulations for elastic media

Green-Lagrange strain tensor and its energy conjugate, the second Piola-Kirchhoff stress tensor, to account for the necessary kinematic relationships. These formulations are developed in the updated Lagrangian framework. The incremental dynamic variational inequality takes the form:

$$\langle \rho \ddot{\mathbf{u}}, \dot{\mathbf{v}} - \dot{\mathbf{u}} \rangle + A(\mathbf{u}, \dot{\mathbf{v}} - \dot{\mathbf{u}}) + j(\mathbf{u}, \dot{\mathbf{v}}) - j(\mathbf{u}, \dot{\mathbf{u}}) \geq R(\dot{\mathbf{v}} - \dot{\mathbf{u}}),$$
$$\dot{\mathbf{v}} \in \mathbf{K} \tag{9}$$

$$\mathbf{u}(0) = \mathbf{u}_o \quad \text{and} \quad \dot{\mathbf{u}}(0) = \dot{\mathbf{u}}_o$$

where,

(i)   the first term represents the rate of virtual work done by the inertia forces:

$$\langle \rho \ddot{\mathbf{u}}, \dot{\mathbf{v}} \rangle = \int\limits_{t+\Delta t\Omega} \rho \ddot{\mathbf{u}} \cdot \dot{\mathbf{v}} d^{\,t+\Delta t}\Omega \tag{10}$$

(ii) the second represents the rate of virtual work done by the internal forces:

$$A(\mathbf{u}, \dot{\mathbf{v}}) = \int\limits_{t\Omega} {}_t C_{ijrs}^{ep} {}_t e_{ij}(\mathbf{u}) {}_t e_{ij}(\dot{\mathbf{v}}) d^{\,t}\Omega$$
$$+ \int\limits_{t\Omega} {}^t \sigma_{ij}(\mathbf{u}) {}^t \eta_{ij}(\dot{\mathbf{v}}) d^{\,t}\Omega \tag{11}$$

(iii) the fourth and fifth terms represent the rate of virtual dissipation by the frictional forces, and can be represented as follows:

$$j(\mathbf{u}, \dot{\mathbf{v}}) = \int\limits_{t+\Delta t\Gamma_C} -\mu^{t+\Delta t} \sigma_N(\mathbf{u}) |\dot{\mathbf{v}}_T| d^{\,t+\Delta t}S \tag{12}$$

(iv) the term on the right-hand-side represents the rate of virtual work done by the residual external forces (unbalance force between the external loading and internal stresses):

$$R(\dot{\mathbf{v}}) = \int\limits_{t+\Delta t\Omega} {}^{t+\Delta t}\mathbf{f} \cdot (\dot{\mathbf{v}}) d^{\,t+\Delta t}\Omega$$
$$+ \int\limits_{t+\Delta t\Gamma_F} {}^{t+\Delta t}\mathbf{t} \cdot (\dot{\mathbf{v}}) d^{\,t+\Delta t}S$$
$$- \int\limits_{t\Omega} {}^t \sigma_{ij}(\mathbf{u}(t)) {}_t e_{ij}(\dot{\mathbf{v}}) d^{\,t}\Omega \tag{13}$$

In addition to the above variational inequality, the elasto-plastic behaviour of the material will be characterised by a yield function F, a strain hardening function ${}^t\kappa$ and the appropriate flow rule. In this case, the constitutive relations take the following form:

$$F({}^t\sigma_{ij}, {}^t\kappa) = 0, \, d_{\,t}e_{ij}^p = {}^t d\lambda \frac{\partial F}{\partial^{\,t}\sigma_{ij}} \tag{14}$$

$$\overset{\triangledown}{{}^t\sigma_{ij}} dt = {}_t C_{ijrs}^e (d_{\,t}e_{rs} - d_{\,t}e_{rs}^p) \tag{15}$$

with C_{ijrs}^e being the material elastic constitutive tensor.

3.2 Special sub-problems of VI formulations

Currently, there is no direct method available for the solution of the general frictional contact problem of Eq. 9 with all the terms included. In our approach, we divide the general elasto-dynamic variational inequality into two consistent sub-problems. Details are provided below.

3.2.1 Sub-problem I: prescribed tangential stresses

If the tangential stresses $\boldsymbol{\sigma}_T$ are assumed to be known everywhere on the boundaries of the two bodies at time $t + \Delta t$, then variational inequality (9) reduces to (Fig. 5a):
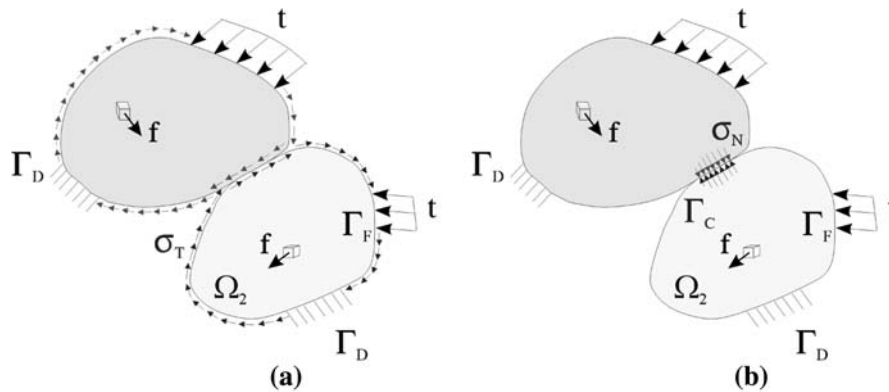
$$\langle \rho \ddot{\mathbf{u}}, \dot{\mathbf{v}} - \dot{\mathbf{u}} \rangle + A(\mathbf{u}, \dot{\mathbf{v}} - \dot{\mathbf{u}}) \geq R_1(\dot{\mathbf{v}} - \dot{\mathbf{u}}) \quad \dot{\mathbf{v}} \in \mathbf{K} \tag{16}$$

where

$$R_1(\dot{\mathbf{v}}) = R(\dot{\mathbf{v}}) + \int\limits_{t+\Delta t\Gamma_C} \boldsymbol{\sigma}_T \cdot \dot{\mathbf{v}} d^{\,t+\Delta t}S \tag{17}$$

The effect of friction is now included in the expression of $R_1(\dot{\mathbf{v}})$ as known tractions over the boundary Γ_C. Note that for the unilateral contact model, the rate of work done by the normal contact stresses $\langle P(\mathbf{u}), \dot{\mathbf{v}} - \dot{\mathbf{u}} \rangle$ is always negative or equal to zero. Therefore, variational inequality Eq. 16 is reduced to:

$$\langle \rho \ddot{\mathbf{u}}, \dot{\mathbf{v}} - \dot{\mathbf{u}} \rangle + A(\mathbf{u}, \dot{\mathbf{v}} - \dot{\mathbf{u}}) \geq R_1(\dot{\mathbf{v}} - \dot{\mathbf{u}}) \tag{18}$$

Fig. 5 (**a**) Prescribed tangential contact stresses on the boundaries (sub-problem I), and (**b**) prescribed normal contact stresses together with the actual contact surface Γ_C (sub-problem II)

3.2.2 Sub-problem II: prescribed normal stresses and contact surface

In sub-problem II, where the normal stresses σ_N together with the actual contact surface Γ_C are known (Fig. 5b), the variational inequality (9) reduces to:

$$\langle \rho \ddot{\mathbf{u}}, \dot{\mathbf{v}} - \dot{\mathbf{u}} \rangle + A(\mathbf{u}, \dot{\mathbf{v}} - \dot{\mathbf{u}}) + j(\dot{\mathbf{v}}) - j(\dot{\mathbf{u}}) \geq R_2(\dot{\mathbf{v}} - \dot{\mathbf{u}}),$$
$$\dot{\mathbf{v}} \in \mathbf{K} \tag{19}$$

where

$$j(\dot{\mathbf{v}}) = \int\limits_{^{t+\Delta t}\Gamma_C} -^{t+\Delta t}\hat{\sigma}_N |\dot{\mathbf{v}}_T| d\,^{t+\Delta t}S \tag{20a}$$

$$R_2(\dot{\mathbf{v}}) = R(\dot{\mathbf{v}}) + \int\limits_{^{t+\Delta t}\Gamma_C} {}^{t+\Delta t}\sigma_N \dot{\mathbf{v}}_N d\,^{t+\Delta t}S \tag{20b}$$

The above VI formulation (19) has a non-differentiable frictional term $j(\cdot)$ (Eq. 20a). In order to overcome this difficulty, two approaches, *regularisation* and *nondifferentiable optimisation*, were adopted (Czekanski and Meguid 2001).

3.3 Modelling of friction

The above VI formulation (19) has a non-differentiable frictional term $j(\cdot)$ (Eq. 20a). In order to overcome this difficulty, two approaches, *regularisation* and *nondifferentiable optimisation*, were adopted (Czekanski and Meguid 2001).

3.3.1 Regularisation technique

In the first method, we adopt the following regularisation approximation and its derivative:

$$|\dot{\mathbf{u}}| \rightarrow \psi_\varepsilon(\dot{\mathbf{u}}) = \begin{cases} |\dot{\mathbf{u}}| - \frac{\varepsilon}{2}, & |\dot{\mathbf{u}}| > \varepsilon \\ \frac{1}{2\varepsilon}\dot{\mathbf{u}} \cdot \mathbf{u}, & |\dot{\mathbf{u}}| \leq \varepsilon \end{cases} \tag{21}$$

$$\phi_\varepsilon(\dot{\mathbf{u}}) = \begin{cases} \frac{\dot{\mathbf{u}}}{|\dot{\mathbf{u}}|}, & |\dot{\mathbf{u}}| > \varepsilon \\ \frac{\dot{\mathbf{u}}}{\varepsilon}, & |\dot{\mathbf{u}}| \leq \varepsilon \end{cases} \tag{22}$$

In this case, the regularised friction conditions reduce to:

$$u_N \geq g \Rightarrow \boldsymbol{\sigma}_T \leq \hat{\sigma}_T \begin{cases} \frac{\dot{\mathbf{u}}_T}{|\dot{\mathbf{u}}_T|}, & |\dot{\mathbf{u}}_T| > \varepsilon \\ \frac{\dot{\mathbf{u}}_T}{\varepsilon}, & |\dot{\mathbf{u}}| \leq \varepsilon \end{cases} \tag{23a}$$

$$u_N < g \Rightarrow \boldsymbol{\sigma}_T = \mathbf{0} \tag{23b}$$

Accordingly, the regularised form of $j(.)$ can be expressed as:

$$j_\varepsilon(\dot{\mathbf{v}}) = \int\limits_{^{t+\Delta t}\Gamma_C} {}^{t+\Delta t}\hat{\sigma}_T \psi_\varepsilon(\dot{\mathbf{v}})\,d\,^{t+\Delta t}S \tag{24}$$

In this case, the regularised frictional term is convex and differentiable with respect to the second argument (Kikuchi and Oden 1988). The partial derivatives of $j_\varepsilon(\dot{\mathbf{w}})$ in the direction $\dot{\mathbf{v}}$, are given by:

$$\langle J_\varepsilon(\dot{\mathbf{w}}), \dot{\mathbf{v}} \rangle \equiv \langle \partial j_\varepsilon(\dot{\mathbf{w}}), \dot{\mathbf{v}} \rangle$$
$$= \int\limits_{^{t+\Delta t}\Gamma_C} {}^{t+\Delta t}\hat{\sigma}_N \phi_\varepsilon(\dot{\mathbf{w}})\dot{\mathbf{v}}_T d\,^{t+\Delta t}S \tag{25}$$

Based on this regularisation technique, the following formulation is obtained:

$$\langle \rho \ddot{\mathbf{u}}, \dot{\mathbf{v}} - \dot{\mathbf{u}} \rangle + A(\mathbf{u}, \dot{\mathbf{v}} - \dot{\mathbf{u}}) + \langle J_\varepsilon(\dot{\mathbf{u}}), \dot{\mathbf{v}} - \dot{\mathbf{u}} \rangle$$
$$= R_2(\dot{\mathbf{v}} - \dot{\mathbf{u}}) \tag{26}$$

It is appropriate to identify the role of the regularisation parameter ε. This parameter determines the region where the tangential component of the relative sliding velocity on the contact surface is small enough to represent sticking. Experimental studies concerning friction between metallic surfaces indicate the existence of a gradual transition from sticking to slipping, which can be represented using the parameter ε (Oden and Martins 1985).

3.3.2 Nondifferentiable optimisation

Mathematical formulation and the numerical computations required as well as the accuracy of the solution obtained using the above technique is dependent on the form of the regularisation function employed and the value of the user-defined regularisation parameter. Several NDO algorithms have been proposed (Balinski and Wolfe 1975; Fletcher 1982; Lemarechal 1974; Mifflin 1982; Shimizu et al. 1997; Shor 1985; Wierzbicki 1982). However, they are inappropriate for dealing with the current VI problem. These difficulties have been overcome in the current solution strategy. The solution of variational inequality (19) is equivalent to the minimisation of the following function π:

$$\pi = \langle \rho \ddot{\mathbf{u}}, \dot{\mathbf{u}} \rangle + \frac{1}{2} A(\mathbf{u}, \dot{\mathbf{u}}) + j(\dot{\mathbf{u}}) - R(\dot{\mathbf{u}}) | \mathbf{N} \cdot \mathbf{u}$$
$$- \mathrm{g} \le 0 \text{ on } \Gamma_C \tag{27}$$

The function π is nondifferentiable, because of the existence of the absolute value function $|\cdot|$ in the frictional term $j(\cdot)$. The frictionless VI formulation is strictly convex and assumes a unique minimum (Hlavacek et al. 1988, Kikuchi and Oden 1988). By considering the frictional formulation, the problem remains convex and quadratic, if the line representing $\dot{\mathbf{u}}_T = \mathbf{0}$ is not crossed. In other words, the function is a convex and quadratic on either side of the $\dot{\mathbf{u}}_T = \mathbf{0}$. Therefore, it is necessary to create an artificial cutting line located at vector $\dot{\mathbf{u}}_T = \mathbf{0}$, which we will denote *artificial nondifferentiable optimisation (ANDO)* constraint. Therefore, the frictional term $j(\cdot)$ represented in Eq. 20, can be described by two complementary subproblems with additional ANDO constraints (Fig. 6):

$$j(\dot{\mathbf{u}}) = \int_{t+\Delta t \Gamma_C} {}^{t+\Delta t}\hat{\sigma}_T \dot{\mathbf{u}}_T \mathrm{d}\, {}^{t+\Delta t}S | \dot{\mathbf{u}}_T \ge \mathbf{0} \quad \text{on } \Gamma_C \tag{28a}$$

$$j(\dot{\mathbf{u}}) = -\int_{t+\Delta t \Gamma_C} {}^{t+\Delta t}\hat{\sigma}_T \dot{\mathbf{u}}_T \mathrm{d}\, {}^{t+\Delta t}S | \dot{\mathbf{u}}_T \le \mathbf{0} \quad \text{on } \Gamma_C$$
$$\tag{28b}$$

In this case, each of these complementary two subproblems is strictly convex and assumes a unique minimum.
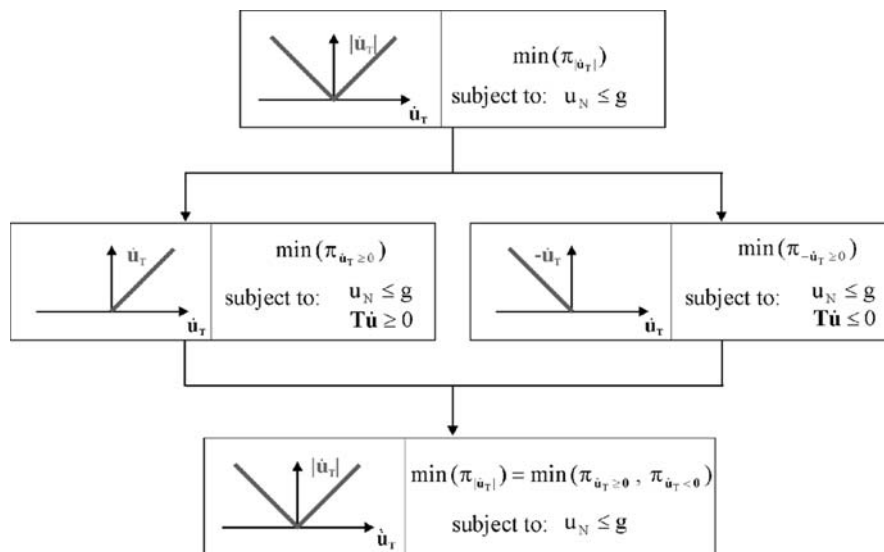
Fig. 6 Solution of nondifferentiable VI function using complementary sub-problems

The frictional contact problem involves two distinct states: *slip and stick*. The two complementary NDO sub-problems of Eqs. 27 and 28 lead to two solutions. The ANDO constraint is always active ($\dot{\mathbf{u}}_T = \mathbf{0}$) in at least one of the two sub-problems. For the second sub-problem, the ANDO constraint could also be active. In this case, both sub-problems must have the same solution, which represents a state of stick. On the other hand, the constraint in the second sub-problem could be inactive. In this case, the solution of the second sub-problem is always the global minimum, which represents a state of slip.

4 Representation of kinematic contact conditions

The accurate and efficient representation of contact surfaces has received little attention in the scientific literature. Most formulations rely on the use of element interpolation functions to describe the contact surface and to impose the kinematic contact conditions. In most contact algorithms, the contact surfaces are defined as a sequence of lines (or curves) connecting the FE nodes with only C^0-continuity. In this case, the normal vector is not uniquely defined at element boundaries (Fig. 7).

Even when higher order elements, such as the 8 and 9 noded 2D elements, are used, the contact surfaces are still non-smooth at the exterior nod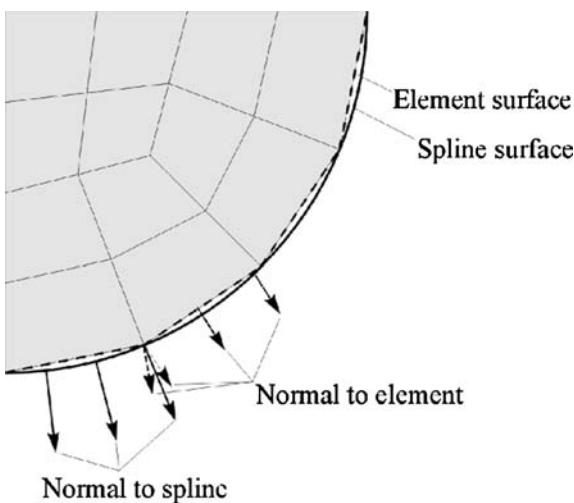es. In fact, higher-order smoothness of the contact surface cannot be achieved through appropriate element selection, since the use of C^1 or higher-order continuous elements in mechanics problems is undesirable (Zienkiewicz and Taylor 1988). On the other hand, only C^0-continuity of the contact surface violates the assumptions of smoothness, which are essential to the proofs of uniqueness and convexity of contact problems (Kikuchi and Oden 1988). To overcome these inconsistencies, contact problems commonly involve a disproportionately fine mesh at the vicinity of the contact region.

This has motivated the use of splines passing through the FE nodes and possessing at least C^1-continuity. The resulting normal vector is uniquely defined at all point belonging to the selected surface (Fig. 7).

4.1 Cubic spline interpolation of surfaces

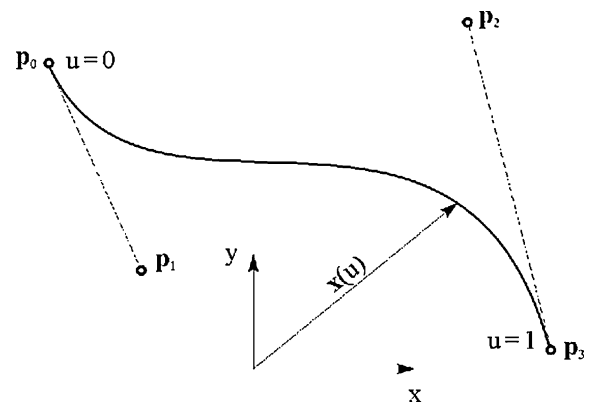Figure 8 shows a smooth cubic spline segment connecting two points in a 2D space. Assuming a

Fig. 7 Contact surface and normal vector based on spline and traditional element interpolations

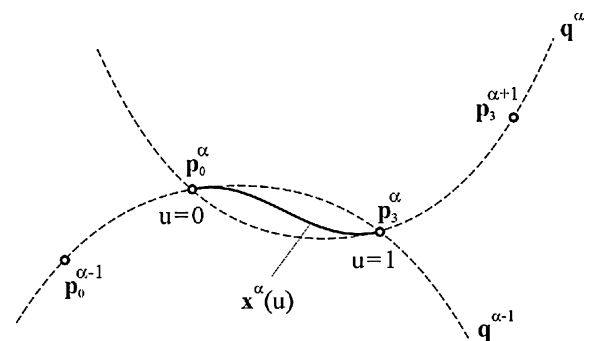**Fig. 8** Characteristics of a cubic spline segment

Fig. 9 Overhauser splines and associated control points

parametric representation, the coordinates of any point on this curve can be expressed in terms of an independent variable u, such that:

$$\mathbf{x}(u) = \begin{Bmatrix} x(u) \\ y(u) \end{Bmatrix}, \quad 0 \le u \le 1 \tag{29}$$

For each curve segment α, the interpolation function passes through the end points $\mathbf{p}_0$ and $\mathbf{p}_3$, while the intermediate points $\mathbf{p}_1$ and $\mathbf{p}_2$ dictate the shape of the curve (Fig. 9). The segment can be described as follows:

$$\mathbf{x}^{\alpha}(u) = \sum_{i=0}^{3} B_i^3(u)\, \mathbf{p}_i^{\alpha} \tag{30}$$

where

$$B_i^m(u) = \binom{m}{i} u^i (1-u)^{m-i} \tag{31}$$

are the Bernstein polynomials.

Each spline segment interpolates between two neighbouring points on the contact surface. The end control points $\mathbf{p}_0$ and $\mathbf{p}_3$ are located at these points. This guarantees that the interpolated surface passes through all the external points. The intermediate points $\mathbf{p}_1$ and $\mathbf{p}_2$ of each segment dictate the shape of the curve and their location is selected based on the specific spline form adopted. In this case, the location of the intermediate control points for all the spline segments are coupled. They can be obtained by solving a predominantly tri-diagonal matrix expressing the C^2 continuity equations (Hughes et al. 1976).

When the location of the FE nodes changes, the matrix must be solved again for the new location of the intermediate points. The overhead associated with this process does not offset the advantages of second order continuity. Note, however, that if the change in the nodal locations is governed by an affine transformation, such as rigid body translation, rotation or shearing, the modified spline needed not be evaluated again (Hughes et al. 1976). Accordingly, the use of C^2 cubic splines is only justified for rigid surfaces.

By requiring just C^1 continuity, simpler and faster interpolation functions can be constructed where the intermediate control points can be obtained without resorting to matrix solution. In this case, the location of the intermediate control points of a segment is only governed by a few neighboring nodes. This property is known as local support. A detailed review of C^1-continuous cubic splines is provided in Ref. (Hughes

et al. 1976). Interpolation functions can be constructed which satisfy a prescribed tangential vector, a prescribed tangential direction or a prescribed normal direction (Hughes et al. 1976; Wriggers et al. 1990). However, these vectors are generally not available in standard FE meshes. Overhauser splines offer an alternative approach that ensures C^1-continuity without requiring prescribed tangential or normal vectors (Parisch 1989; Laursen and Simo 1993). Accordingly, they are the most suitable interpolation form for the finite element contact problem involving flexible bodies. These splines have also been used to develop smooth boundary elements (Refaat and Meguid 1994) and to evaluate contact between gear teeth (Ferris and Pang 1997). Other cubic spline interpolation schemes possessing the same properties as Overhauser splines were also developed (Duvaut and Lions 1976; Adams 1975). However, not all of them are unique. The Catmull-Rom spline (Duvaut and Lions 1976), for example, is identical to the Overhauser spline.

Each Overhauser curve segment can be considered a linear blend of two parabolas $\mathbf{q}^{\alpha-1}$ and $\mathbf{q}^{\alpha}$ (Fig. 9):

$$\mathbf{x}^{\alpha}(u) = (1-u)\,\mathbf{q}^{\alpha-1}(u+1) + u\,\mathbf{q}^{\alpha}(u), \quad 0 \le u \le 1 \tag{32}$$

Each parabola passes through the two surface nodes $\mathbf{p}_0^{\alpha}$ and $\mathbf{p}_3^{\alpha}$ as well as a neighboring surface node (one from each side). Consequently, the Overhauser spline can be expressed directly in terms of the coordinates of the two points defining the segment and their two adjacent points:

$$\mathbf{x}^{\alpha}(u) = \begin{bmatrix} \mathbf{p}_0^{\alpha-1} & \mathbf{p}_0^{\alpha} & \mathbf{p}_3^{\alpha} & \mathbf{p}_3^{\alpha+1} \end{bmatrix} \begin{bmatrix} -\frac{1}{2}u^3 + u^2 - \frac{1}{2}u \\ \frac{3}{2}u^3 - \frac{5}{2}u^2 + 1 \\ -\frac{3}{2}u^3 + 2u^2 + \frac{1}{2}u \\ \frac{1}{2}u^3 - \frac{1}{2}u^2 \end{bmatrix}, \quad 0 \le u \le 1 \tag{33}$$

It is convenient to use the Bernstein polynomial form of Eq. 30 to represent all forms of cubic splines. In this case, only the location of the four control points should be stored for each spline segment. The first and last control points of each segment coincide with the finite element nodes, while the intermediate control points can generally be obtained using the following relationship:

$$\mathbf{p}_1^\alpha = \mathbf{p}_0^\alpha + \tfrac{1}{3}\mathbf{m}^\alpha(0) \tag{34}$$
$$\mathbf{p}_2^\alpha = \mathbf{p}_3^\alpha - \tfrac{1}{3}\mathbf{m}^\alpha(1)$$

where $\mathbf{m}$ is the vector tangent to the spline curve such that:

$$\mathbf{m}(u) = \frac{d}{du}\mathbf{x}(u) \tag{35}$$

Splines possess several properties which are fundamental to their use for modeling contact surfaces. According to the convex hull property, the spline curve cannot exceed the geometric bounds of the control polygon (Hughes et al. 1976). This polygon is formed from the spline's control points. This feature is extremely useful for contact detection, since it is much simpler and faster to check interference between a master node and the control polygon rather than the actual spline curve. Most cubic spline interpolation schemes are invariant under affine transformations. This property is generally useful for modeling rigid surfaces. Local support is another important feature that should be present in the splines defining the contact surface. This implies that a change in the location of one of the contact nodes only affects a few neighboring splines and not the entire contact surface. Note, however, that C^2-continuous cubic splines do not possess this property.

4.2 Node-to-surface spline slave-master algorithm

In cases involving contact between flexible bodies, analytical surface profiles that describe the initial geometry cannot be used to describe the deformed one. As a result, spline-based surface interpolation should be used.

Based on a master-slave technique, for every point on the master contact surface Γ_m a corresponding closest point on the slave surfaces Γ_s is determined from the kinematics of the deformation (Fig. 10). This is defined for $\mathbf{x} \in \Gamma_m$ as being

$$\|\mathbf{x} - \mathbf{y}^*({}^t\mathbf{x})\| = \min_{\mathbf{y} \in \Gamma_s} \|{}^t\mathbf{x} - {}^t\mathbf{y}\| \tag{36}$$

The unit normal vector to the slave surface can be defined as:

$${}^t\mathbf{N} = \frac{\partial^t\mathbf{y}^*/\partial\xi \times \partial^t\mathbf{y}^*/\partial\eta}{\|\partial^t\mathbf{y}^*/\partial\xi \times \partial^t\mathbf{y}^*/\partial\eta\|} \tag{37}$$

where ξ and η are two orthogonal directions tangential to the contact surface. If the contact surfaces are
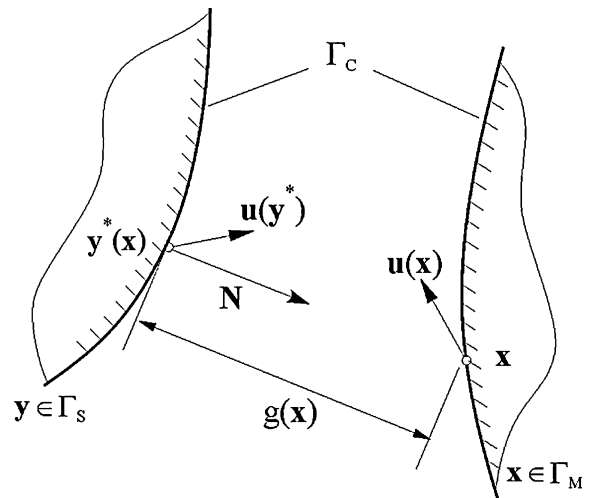

Fig. 10 Kinematic contact constraints

sufficiently smooth, the normal can alternatively be expressed as:

$${}^t\mathbf{N} = \frac{{}^t\mathbf{y}^*({}^t\mathbf{x}) - {}^t\mathbf{x}}{\|{}^t\mathbf{y}^*({}^t\mathbf{x}) - {}^t\mathbf{x}\|} \tag{38}$$

The unit outward normal vector for a 2-D spline segment α can be expressed as:

$${}^t\mathbf{N}^\alpha(u) = \frac{[-{}^tm_y^\alpha(u) \quad {}^tm_x^\alpha(u)]^T}{\|{}^t\mathbf{m}^\alpha(u)\|} \tag{39}$$

The gap function between two contact surfaces at unknown configuration $t + \Delta t$ can be expressed as (Fig. 10):

$${}^{t+\Delta t}g(\mathbf{x}) = \left[{}^{t+\Delta t}\mathbf{y}^*(\mathbf{x}) - {}^{t+\Delta t}\mathbf{x}\right] \cdot {}^{t+\Delta t}\mathbf{N} \geq 0 \tag{40}$$

This gap function is generally nonlinear. However, for incremental finite element analysis, a linearised form yields:

$${}^{t+\Delta t}g(\mathbf{x}) = {}^tg(\mathbf{x}) + \left[\Delta^t\mathbf{u}(\mathbf{y}^*) - \Delta^t\mathbf{u}(\mathbf{x})\right] \cdot {}^t\mathbf{N} \geq 0 \tag{41}$$

where $\Delta\mathbf{u}$ is the incremental displacement vector.

Contact search

Using nonlinear relationships, such as splines, to represent the contact surface can slow down the contact search procedure. To overcome this, the contact search is divided into two stages. The purpose of the first stage is to obtain a quick estimate of the proximity of a master node to a specific spline

segment. In this stage, interference is checked between the master node and the control polygon of the spline segment, as shown in Fig. 11. According to the convex hull property, the spline curve cannot exceed the geometric bounds of the control polygon (Hughes et al. 1976). For rapid penetration problems, it is sometimes useful to search within a wider envelope surrounding the control polygon. If the master node is inside the search region, an accurate iterative contact check is performed in the second stage of the search. In this stage, the exact target point and gap/penetration are determined.

4.3 Contact constraints

The discretisation of a contact constraint α (Eq. 40 or 41) can be represented as:

$$(\mathbf{A}_\alpha \simeq \boldsymbol{\delta}_\alpha - G_\alpha) \leq 0 \tag{42}$$

where G_α is the gap and $\boldsymbol{\delta}_\alpha$ is the vector of nodal degrees of freedom associated with this constraint.

$$\delta_\alpha = \{ \Delta \mathbf{U}_\mathrm{M} \quad \Delta \mathbf{U}_{\mathrm{S}(1)} \quad \Delta \mathbf{U}_{\mathrm{S}(2)} \quad \cdots \quad \Delta \mathbf{U}_{\mathrm{S}(L)} \}^\mathrm{T} \tag{43}$$

where $\Delta \mathbf{U}_\mathrm{i}$ represents all degrees of freedom associated with node i and L is the number of nodes on the slave surface that affect the displacement of the target node.

Matrix $\mathbf{A}_\alpha$, which represents the contact constraint, can be expressed as follows:

$$\{\mathbf{A}_\alpha\} = [\mathbf{N}_\alpha^\mathrm{T} \quad -\phi_{\mathrm{S}(1)}\mathbf{N}_\alpha^\mathrm{T} \quad -\phi_{\mathrm{S}(2)}\mathbf{N}_\alpha^\mathrm{T} \quad \cdots \quad -\Phi_{\mathrm{S}(L)}\mathbf{N}_\alpha^\mathrm{T}] \tag{44}$$

where the weight function $\phi_{\mathrm{S}(i)}$ is the contribution of each of the surface nodes to the normal displacement at the target point on the slave contact surface. These
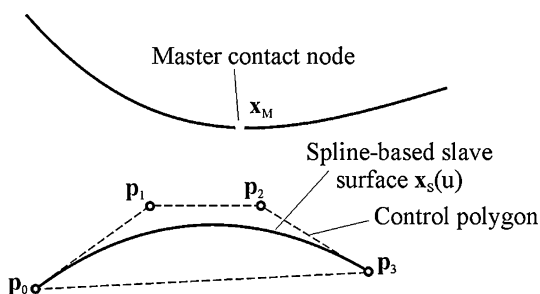
Fig. 11 Control polygon used for preliminary contact search

values are determined based on the contact search and are subject to the following restriction:

$$\sum_{\mathrm{i}=1}^{\mathrm{L}} \phi_{\mathrm{S}(i)} = 1 \tag{45}$$

For Overhauser splines, the four weight functions can easily be evaluated based in Eq. 33.

Finally, the assembly of all contact constraints yields a set of inequalities of the form:

$$\mathbf{AU} \leq \mathbf{G} \tag{46}$$

where $\mathbf{G}$ is the global vector of the gap functions, $\mathbf{U}$ is the assembled displacement vector and the $\mathbf{A}$-matrix represents the standard finite element assembly of all individual $\mathbf{A}_\alpha$ constraint matrices.

5 Finite element implementation

5.1 Time integration schemes

In order to solve the above VI formulations (Eqs. 16 and 19), the finite difference method is employed in the time domain to establish a relationship between the acceleration, velocity and displacement fields. The Newmark method (Newmark et al. 1959), the most frequently used implicit time integration scheme, allows us to approximate the velocities and displacements at time $t + \Delta t$ in terms of acceleration at time $t + \Delta t$ and the field variables at time t, as follows:

$$^{t+\Delta t}\dot{\mathbf{U}} = {}^t\dot{\mathbf{U}} + [(1 - \gamma)^t\ddot{\mathbf{U}} + \gamma\,^{t+\Delta t}\ddot{\mathbf{U}}]\Delta t \tag{47a}$$

$$^{t+\Delta t}\mathbf{U} = {}^t\mathbf{U} + {}^t\dot{\mathbf{U}}\Delta t + [(0.5 - \beta)^t\ddot{\mathbf{U}} + \beta\,^{t+\Delta t}\ddot{\mathbf{U}}]\Delta t^2 \tag{47b}$$

It has been shown that the use of the trapezoidal rule ($\gamma = 0.5$ and $\beta = 0.25$) with a fully implicit treatment of the contact constraints produces oscillations, which can be significant as the time steps and spatial discretizations are refined (Chaudhary and Bathe 1986). Recently, the generalized-α method was developed for solving structural dynamics problems with second order accuracy (Chung and Hulbert 1993). In this method, the equation of motion is modified as follows:

$$\mathbf{M}\,^{(t+\Delta t)-\alpha_\mathrm{B}}\ddot{\mathbf{U}} + \mathbf{C}\,^{(t+\Delta t)-\alpha_\mathrm{H}}\dot{\mathbf{U}} + \mathbf{K}\,^{(t+\Delta t)-\alpha_\mathrm{H}}\mathbf{U}$$
$$= {}^{(t+\Delta t)-\alpha_\mathrm{H}}\mathbf{F} \tag{48}$$

where

$$
\begin{aligned}
{}^{(t+\Delta t)-\alpha_H}\mathbf{U} &= (1-\alpha_H)^{t+\Delta t}\mathbf{U} + \alpha_H{}^t\mathbf{U} \\
{}^{(t+\Delta t)-\alpha_H}\dot{\mathbf{U}} &= (1-\alpha_H)^{t+\Delta t}\dot{\mathbf{U}} + \alpha_H{}^t\dot{\mathbf{U}} \\
{}^{(t+\Delta t)-\alpha_B}\ddot{\mathbf{U}} &= (1-\alpha_B)^{t+\Delta t}\ddot{\mathbf{U}} + \alpha_B{}^t\ddot{\mathbf{U}} \\
{}^{(t+\Delta t)-\alpha_H}\mathbf{F} &= (1-\alpha_H)^{t+\Delta t}\mathbf{F} + \alpha_H{}^t\mathbf{F}
\end{aligned}
\tag{49}
$$

The Newmark time-integration scheme (Eq. 47) is used to solve the above equation of motion.

5.2 General algorithm for dynamic contact problems

The reduced VI sub-problems are integrated into two-step and one-step algorithms to provide the complete solution. For each time increment, the solution algorithm is summarised as follows (Fig. 12a):

(i)   the effective stiffness matrix and load vector are calculated using generalized-α time integration,
(ii)   the reduced VI problem(s) is solved iteratively to obtain the displacement, velocity and acceleration fields as well as the current contact surface, contact forces, and
(iii)   steps (i) and (ii) are repeated until the final time is reached.

Step (ii) of the above algorithm rely on the decomposition of the physical problem into two and one sub-problem(s). The first two-step approach integrates the two sub-problems in one algorithm to provide the complete solution (Fig. 12b). For each time instant, first, sub-problem I is solved to obtain the current contact surface and normal contact stresses. Then, sub-problem II is solved to obtain the displacement, velocity, acceleration as well as the frictional forces. Those two steps are repeated until convergence is reached. The second one-step approach is based on the iterative solution of sub-problem II (Fig. 12c). It is worth noting that both proposed algorithms are identical in the frictionless cases.

Generally, convergence in contact problems is measured by the change in the contact and frictional states as well as an energy norm. In this study, the solution is deemed converged once the change in the normal and tangential contact forces between two consecutive iterations is less then 0.01%.

5.3 Two step algorithm

The proposed VI algorithm relies on the decomposition of the physical problem into two sub-problems to identify the candidate contact surface and contact stresses.

Fig. 12 Flow charts of: (a) general algorithm for dynamic contact problems, (b) two-step algorithm, and (c) one-step algorithm
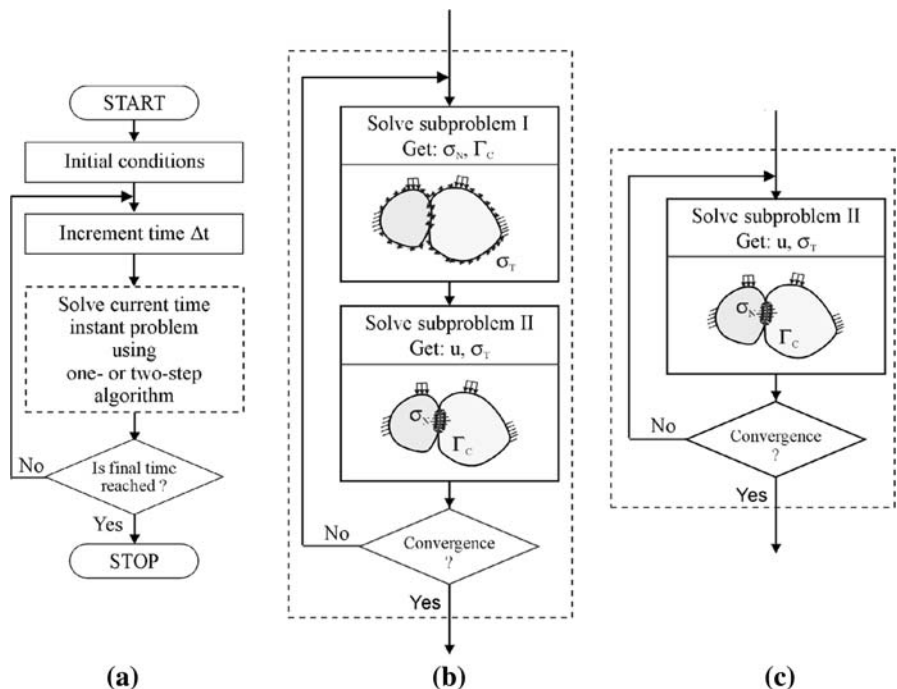
5.3.1 Solution of sub-problem I

Using expression (18), in conjunction with the generalized-α method, the reduced VI formulation is equivalent to solving the following minimisation problem (Czekanski et al. 2001):

$$\min\left\{\frac{1}{2}\,^{t+\Delta t}\mathbf{U}^T\,\hat{\mathbf{K}}\,^{t+\Delta t}\mathbf{U} - \,^{t+\Delta t}\mathbf{U}^{T\,t+\Delta t}\hat{\mathbf{F}}_1\right\} \tag{50}$$
$$\forall\,^{t+\Delta t}\mathbf{U}|\mathbf{A}\,^{t+\Delta t}\mathbf{U} \le \mathbf{G}$$

where

$$\hat{\mathbf{K}} = (1 - \alpha_H)\mathbf{K} + (1 - \alpha_B)\frac{1}{\beta\Delta t^2}\mathbf{M} \tag{51a}$$

$$
\begin{aligned}
^{t+\Delta t}\hat{\mathbf{F}}_1 = &(1 - \alpha_H)^{t+\Delta t}\mathbf{F} + \alpha_H\,^t\mathbf{F} + \,^{t+\Delta t}\mathbf{F}_T \\
&- \alpha_B\mathbf{M}\,^t\ddot{\mathbf{U}} - \alpha_H\mathbf{K}\,^t\mathbf{U} + (1 - \alpha_B)\mathbf{M} \\
&\times \left[\frac{1}{\beta\Delta t^2}\,^t\mathbf{U} + \frac{1}{\beta\Delta t}\,^t\dot{\mathbf{U}} + \left(\frac{1}{2\beta} - 1\right)\,^t\ddot{\mathbf{U}}\right]
\end{aligned} \tag{51b}
$$

where β and γ are Newmark time integration parameters, α_H is the generalized-α time integration parameter which modifies all terms in the equation of motion with the exception of the inertia term. The generalized-α time integration α_B modifies only the inertia term. Selection of those four parameters is crucial to the stability and accuracy of the solution. The details regarding the selection of the time integration scheme and parameters are provided in the chapter five.

$\mathbf{F}$ represents the vector of external nodal forces, while vector $\mathbf{F}_T$ represents the frictional forces acting on Γ_C. The Hessian matrix $\hat{\mathbf{K}}$ in Eq. 50 is the effective stiffness matrix, which is positive definite. Accordingly, expression (50) represents a convex Quadratic Programming problem. The current contact surface is directly defined by the active set of constraints. In this study, we adopt the technique of Quadratic Programming to solve this minimisation problem based on the projected gradients search method. The flow chart for the algorithm of the sub-problem I for unilateral contact model is depicted in Fig. 13.

5.3.2 Solution of sub-problem II

The solution of sub-problem II is equivalent to the minimisation of the functional $^{t+\Delta t}\mathbf{U}$, subject to the equality constraint $\mathbf{B}\,^{t+\Delta t}\mathbf{U} = \mathbf{G}$ resulting from the known contact surface Γ_C and the prescribed displacements on Γ_D (Czekanski et al. 2001), i.e.,
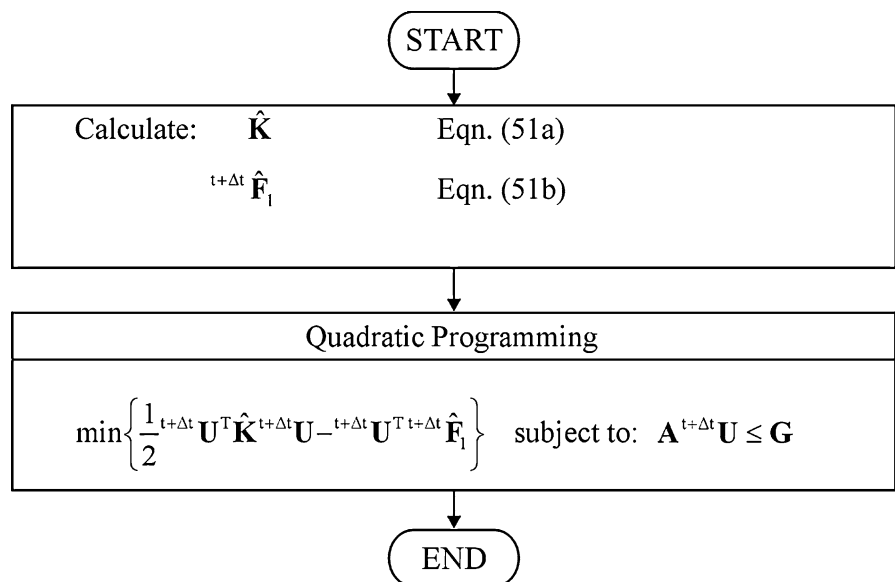
$$\min(\pi(^{t+\Delta t}\mathbf{U}))|\mathbf{B}\,^{t+\Delta t}\mathbf{U} = \mathbf{G} \quad \text{on } \Gamma_C \cup \Gamma_D \tag{52}$$

where

$$
\begin{aligned}
\pi(^{t+\Delta t}\mathbf{U}) = &\frac{1}{2}\,^{t+\Delta t}\mathbf{U}^T(1 - \alpha_H)\mathbf{K}\,^{t+\Delta t}\mathbf{U} - \,^{t+\Delta t}\mathbf{F}_\alpha^{T\,t+\Delta t}\mathbf{U} \\
&+ \,^{t+\Delta t}\ddot{\mathbf{U}}^T(1 - \alpha_B)\mathbf{M}\,^{t+\Delta t}\mathbf{U} + \,^{t+\Delta t}\mathbf{F}_T^{T\,t+\Delta t}\Psi_\varepsilon
\end{aligned} \tag{53}
$$

The modified external force vector is defined as:

Fig. 13 Flow chart of algorithm for sub-problem I

$$^{t+\Delta t}\mathbf{F}_\alpha = (1-\alpha_H)^{t+\Delta t}\mathbf{F} + \alpha_H{}^t\mathbf{F} - {}^t\mathbf{U}^T\alpha_H\mathbf{K}$$
$$- {}^t\ddot{\mathbf{U}}^T\alpha_B\mathbf{M} \qquad (54)$$

The last term in Eq. 53 represents the regularised discrete frictional forces. The solution to Eq. 53 can be obtained using Lagrange multipliers. In this case, the equivalent unconstrained functional $\chi\big(^{t+\Delta t}\mathbf{U}, {}^{t+\Delta t}\Lambda\big)$ is as follows:

$$\chi\big(^{t+\Delta t}\mathbf{U}, {}^{t+\Delta t}\Lambda\big) = \frac{1}{2}{}^{t+\Delta t}\mathbf{U}^T(1-\alpha_H)\mathbf{K}^{t+\Delta t}\mathbf{U}$$
$$+ {}^{t+\Delta t}\ddot{\mathbf{U}}^T(1-\alpha_B)\mathbf{M}^{t+\Delta t}\mathbf{U}$$
$$- {}^{t+\Delta t}\mathbf{F}_\alpha^{T\,t+\Delta t}\mathbf{U} + {}^{t+\Delta t}\mathbf{F}_T^T\Psi_\varepsilon$$
$$+ {}^{t+\Delta t}\Lambda(\mathbf{B}^{t+\Delta t}\mathbf{U} - \mathbf{G}) \qquad (55)$$

which attains its extremum when:

$$\mathbf{L}_1 = \frac{\partial\chi(^{t+\Delta t}\mathbf{U}, {}^{t+\Delta t}\Lambda)}{\partial^{t+\Delta t}\mathbf{U}} = \mathbf{0} \text{ and } \mathbf{L}_2$$
$$= \frac{\partial\chi(^{t+\Delta t}\mathbf{U}, {}^{t+\Delta t}\Lambda)}{\partial^{t+\Delta t}\Lambda} = \mathbf{0} \qquad (56)$$

where $\mathbf{L}_1$ and $\mathbf{L}_2$ are defined as:

$$\mathbf{L}_1\big(^{t+\Delta t}\mathbf{U}, {}^{t+\Delta t}\Lambda\big) = \hat{\mathbf{K}}^{t+\Delta t}\mathbf{U} - {}^{t+\Delta t}\hat{\mathbf{F}} + \Phi_\varepsilon^{T\,t+\Delta t}\mathbf{F}_T$$
$$+ \mathbf{B}^{t+\Delta t}\Lambda$$
$$= 0$$
$$(57a)$$

$$\mathbf{L}_2\big(^{t+\Delta t}\mathbf{U}, {}^{t+\Delta t}\Lambda\big) = \mathbf{B}^{t+\Delta t}\mathbf{U} - \mathbf{G} = \mathbf{0} \qquad (57b)$$

$\hat{\mathbf{K}}$ is the effective stiffness matrix defined in Eq. 51a, Λ represents the vector of nodal normal forces, and $\hat{\mathbf{F}}$ is the effective load vector, defined as:

$$^{t+\Delta t}\hat{\mathbf{F}} = (1-\alpha_H)^{t+\Delta t}\mathbf{F} + \alpha_H{}^{t+\Delta t}\mathbf{F} - \alpha_B\mathbf{M}^t\ddot{\mathbf{U}} - \alpha_H\mathbf{K}^t\mathbf{U}$$
$$- (1-\alpha_B)\mathbf{M}\left[\frac{1}{\beta\Delta t^2}{}^t\mathbf{U} + \frac{1}{\beta\Delta t}{}^t\dot{\mathbf{U}} + \left(\frac{1}{2\beta}-1\right){}^t\ddot{\mathbf{U}}\right]$$
$$(58)$$

The elements of the matrix Φ_ε of Eq. 57 (dimension $m \times n$), which represents the regularisation function $\phi_\varepsilon(\dot{\mathbf{v}})$, can be calculated as:

where m is the number of active contact constraints and n is the total number of degree of freedom.

In the above expression, the discretised tangential velocity component $\dot{\mathbf{U}}_T$ is determined by multiplying the global velocity vector $\dot{\mathbf{U}}$ with the $\mathbf{T}$ matrix, as follows:

$$^{t+\Delta t}\dot{\mathbf{U}}_T = \mathbf{T}^{t+\Delta t}\dot{\mathbf{U}} \qquad (60)$$

Equations 57a and 57b can be represented as follows:

$$\begin{bmatrix} \mathbf{L}_1\big(^{t+\Delta t}\mathbf{U}, {}^{t+\Delta t}\Lambda\big) \\ \mathbf{L}_2\big(^{t+\Delta t}\mathbf{U}, {}^{t+\Delta t}\Lambda\big) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \qquad (61)$$

or in the equivalent compact form $\mathbf{L}\big(^{t+\Delta t}\mathbf{U}, {}^{t+\Delta t}\Lambda\big) = \mathbf{0}$. The vector functions $\mathbf{L}_1$ and $\mathbf{L}_2$ are of dimensions n and m respectively. To solve this system of equations, a Newton–Raphson technique is used. The solution for iteration i is as follows:

$$\begin{bmatrix} ^{t+\Delta t}\mathbf{U} \\ ^{t+\Delta t}\Lambda \end{bmatrix}^i = \begin{bmatrix} ^{t+\Delta t}\mathbf{U} \\ ^{t+\Delta t}\Lambda \end{bmatrix}^{i-1} - \mathbf{L}'(^{t+\Delta t}\mathbf{U}^{i-1}, {}^{t+\Delta t}\Lambda^{i-1})^{-1}$$
$$\times \mathbf{L}(^{t+\Delta t}\mathbf{U}^{i-1}, {}^{t+\Delta t}\Lambda^{i-1}), \quad i > 1$$

The corresponding matrix of first derivatives $\mathbf{L}'$ can be expressed as:

$$\mathbf{L}'(^{t+\Delta t}\mathbf{U}, {}^{t+\Delta t}\Lambda)$$
$$= \begin{bmatrix} \dfrac{\partial\mathbf{L}_1(^{t+\Delta t}\mathbf{U}, {}^{t+\Delta t}\Lambda)}{\partial^{t+\Delta t}\mathbf{U}} & \dfrac{\partial\mathbf{L}_1(^{t+\Delta t}\mathbf{U}, {}^{t+\Delta t}\Lambda)}{\partial^{t+\Delta t}\Lambda} \\ \dfrac{\partial\mathbf{L}_2(^{t+\Delta t}\mathbf{U}, {}^{t+\Delta t}\Lambda)}{\partial^{t+\Delta t}\mathbf{U}} & \dfrac{\partial\mathbf{L}_2(^{t+\Delta t}\mathbf{U}, {}^{t+\Delta t}\Lambda)}{\partial^{t+\Delta t}\Lambda} \end{bmatrix} \qquad (63)$$

where

$$\frac{\partial\mathbf{L}_1(^{t+\Delta t}\mathbf{U}, {}^{t+\Delta t}\Lambda)}{\partial^{t+\Delta t}\mathbf{U}} = \hat{\mathbf{K}} + \frac{\gamma}{\beta\Delta t}{}^{t+\Delta t}\mathbf{C}_T,$$
$$\frac{\partial\mathbf{L}_1(^{t+\Delta t}\mathbf{U}, {}^{t+\Delta t}\Lambda)}{\partial^{t+\Delta t}\Lambda} = \mathbf{B}^T \qquad (64a)$$

$$\big\{\Phi_\varepsilon\big\}_{ij} = T_{ij}\begin{cases} 1, & \sum\limits_{k=1}^n T_{ik}{}^{t+\Delta t}\dot{U}_k > \varepsilon \\ \frac{1}{\varepsilon}\sum\limits_{k=1}^n T_{ik}{}^{t+\Delta t}\dot{U}_k, & \left|\sum\limits_{k=1}^n T_{ik}{}^{t+\Delta t}\dot{U}_k\right| \le \varepsilon \quad i=1,\ldots,m, \quad j=1,\ldots,n \\ -1, & \sum\limits_{k=1}^n T_{ik}{}^{t+\Delta t}\dot{U}_k < -\varepsilon \end{cases} \qquad (59)$$

$$\frac{\partial \mathbf{L}_2(^{t+\Delta t}\mathbf{U}, {}^{t+\Delta t}\Lambda)}{\partial^{t+\Delta t}\mathbf{U}} = \mathbf{B}, \text{ and } \frac{\partial \mathbf{L}_2(^{t+\Delta t}\mathbf{U}, {}^{t+\Delta t}\Lambda)}{\partial^{t+\Delta t}\Lambda} = \mathbf{0}$$
(64b)

Finally, the damping matrix $\mathbf{C}_T$, which results from the regularisation of the frictional term, can be expressed as follow:

$$\{^{t+\Delta t}\mathbf{C}_T\}_{ij}$$

$$\equiv \sum_{l=1}^{m} \{^{t+\Delta t}\mathbf{F}_T\}_l \{\mathbf{T}\}_{ij} \begin{cases} 0, & \left|\sum_{k=1}^{n} T_{ij}{}^{t+\Delta t}\dot{U}_k\right| > \varepsilon \\ \frac{T_{ij}}{\varepsilon}, & \left|\sum_{k=1}^{n} T_{ij}{}^{t+\Delta t}\dot{U}_k\right| \leq \varepsilon \end{cases}$$
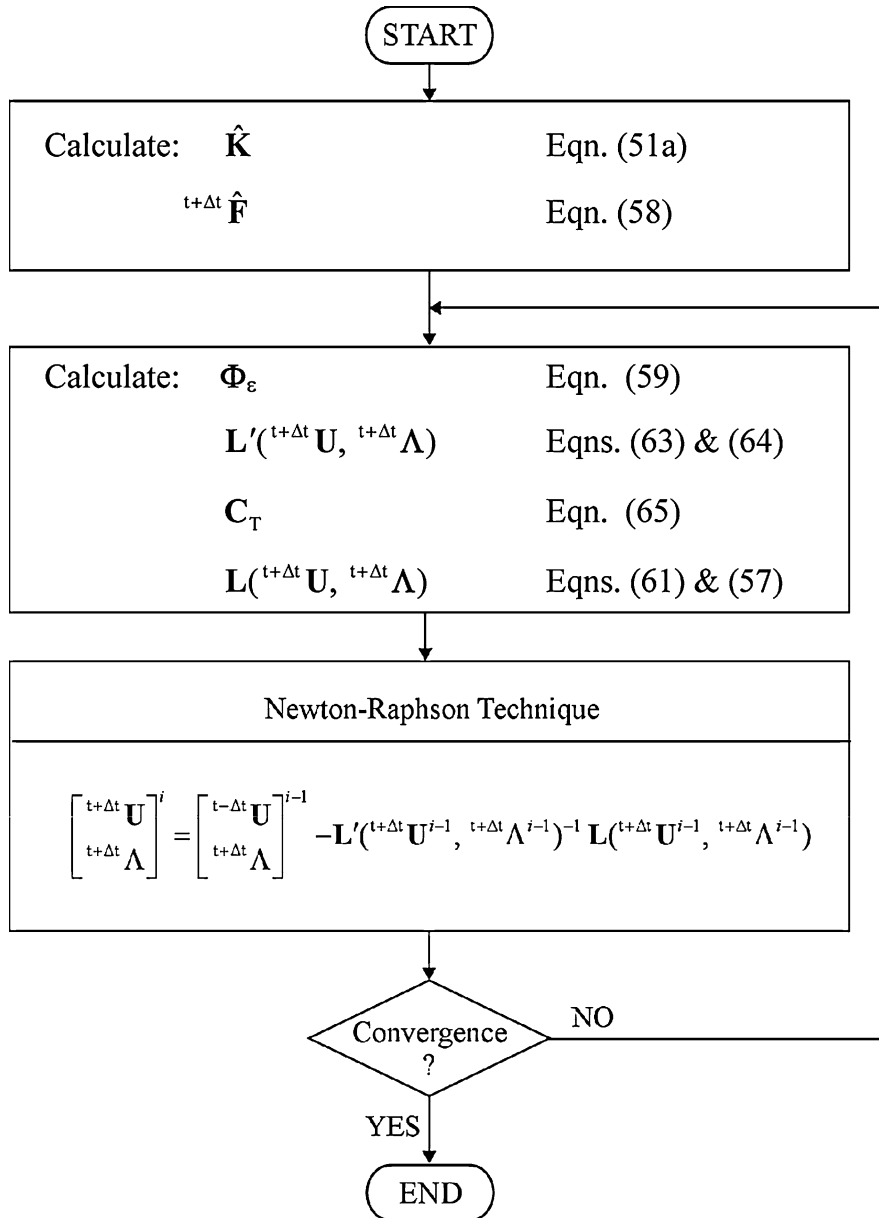(65)

Fig. 14 Flow chart of algorithm of sub-problem II

The flow chart for the algorithm for the sub-problem-II for unilateral contact model is depicted in Fig. 14.

5.4 One-step algorithm using nondifferentiable optimisation

The second approach, which treats the VI formulation in one-step approach, is based on the iterative solution of sub-problem II discussed in chapter three. Using Eqs. 27 and 28, in conjunction with the generalized-α method, the discretised reduced incremental VI formulation is equivalent to solving the following minimisation problem (Czekanski et al. 2001):

$$\min\left\{\frac{1}{2}{}^{t+\Delta t}\Delta\mathbf{U}^{i+1^T}\hat{\mathbf{K}}{}^{t+\Delta t}\Delta\mathbf{U}^{i+1} - {}^{t+\Delta t}\Delta\mathbf{U}^{i+1^T t+\Delta t}\Delta\hat{\mathbf{F}}^{i+1}\right.$$
$$\left. + {}^{t+\Delta t}\Delta\mathbf{U}^{i+1^T}\mathbf{TS}^{t+\Delta t}\mathbf{F}_T^i\right\} \forall \, {}^{t+\Delta t}\Delta\mathbf{U}^{i+1} \qquad (66)$$

subject to the following kinematic contact condition and ANDO constraint:

$$\mathbf{A}^{t+\Delta t}\Delta\mathbf{U}^{i+1} \leq \mathbf{G} \qquad (67a)$$

$$\mathbf{ST}^{t+\Delta t}\Delta\mathbf{U}^{i+1} \leq \mathbf{ST}\left({}^{t+\Delta t}\mathbf{U}^i - {}^t\mathbf{U}\right) \qquad (67b)$$

where

$$\hat{\mathbf{K}} = (1 - \alpha_H)\mathbf{K} + (1 - \alpha_B)\frac{1}{\beta\Delta t^2}\mathbf{M} \qquad (68)$$

$${}^{t+\Delta t}\Delta\hat{\mathbf{F}}^{i+1} = (1 - \alpha_H)^{t+\Delta t}\mathbf{F} + \alpha_H{}^t\mathbf{F} - \alpha_B\mathbf{M}^t\ddot{\mathbf{U}}$$
$$- \alpha_H\mathbf{K}^t\mathbf{U} + (1 - \alpha_B)\mathbf{M}$$
$$\times \left[\frac{1}{\beta\Delta t^2}{}^t\mathbf{U} + \frac{1}{\beta\Delta t}{}^t\dot{\mathbf{U}} + \left(\frac{1}{2\beta} - 1\right){}^t\ddot{\mathbf{U}}\right]$$
$$- {}^{t+\Delta t}\mathbf{F}_{int}^i \qquad (69)$$

This incremental solution can be represented as follows:

$${}^{t+\Delta t}\mathbf{U}^{i+1} = {}^{t+\Delta t}\mathbf{U}^i + {}^{t+\Delta t}\Delta\mathbf{U}^{i+1} \quad \text{for i > 1} \qquad (70)$$

The first constraint (Eq. 67a) represents the assembly of the kinematic contact conditions of the nodes on the candidate contact surface Γ_C. The second constraint (Eq. 67b) represents the assembly of the ANDO constraint ($s_k{}^{t+\Delta t}\dot{\mathbf{U}}_{Tk} \geq 0$) of the nodes on the candidate contact surface Γ_C. The sign matrix **S**, which switches between the two complementary sub-problems, is unknown in advance, and is part of

the solution. Therefore, the new algorithm employs a sequence of convex Quadratic Programming problems.

The flow chart for the algorithm is depicted in Fig. 15 and can be stated in the following steps:

(i) Calculate the tangential velocity vector $\dot{\mathbf{U}}_{Tk}$ and the signs of each of its terms s_k, for k = 1,..., m, where m is the number of contact constraints. The variable s_k can take only -1 or $+1$ for nodes in contact and 0 for nodes not in contact.

(ii) Create the ANDO constraint set: $s_k\dot{\mathbf{U}}_{Tk} \geq 0$ for k = 1,..., m.

(iii) Solve the current Quadratic Programming problem subject to the appropriate kinematic contact condition and ANDO constraints, and then find the minimum of the current objective function.

(iv) Identify the current active ANDO constraints set ($\dot{\mathbf{U}}_{Tk} = 0$ for k = 1,..., m). If this set is empty, then we obtain the optimal solution.

(v) For each active ANDO constraint k reverses the sign of the ANDO constraint, i.e. $s_k = -s_k$, and solves the resulting complementary Quadratic Programming problem and computes the minimum value of the objective function.

(vi) If all the elements of the active ANDO constraints are considered and cannot produce an improved value of the objective function, then **U** is the optimal solution. Exit.

The above algorithm is guaranteed to converge in a finite number of steps, because of the condition of strict convexity assumed and because the number of the cutting planes (i.e. the planes given by $\dot{\mathbf{U}}_{Tk} = 0$, k = 1,...,$m$) is finite.

6 Numerical examples

Four examples have been selected to demonstrate the flexibility, robustness and accuracy of the newly developed approach. The first involves dynamic frictionless contact between a cantilever beam and a rigid foundation. In the second example, we analyze contact in gears. The third example is devoted to the analysis of the fir-tree contact in aeroengine turbine disc assembly. Finally, we analyze the residual stress
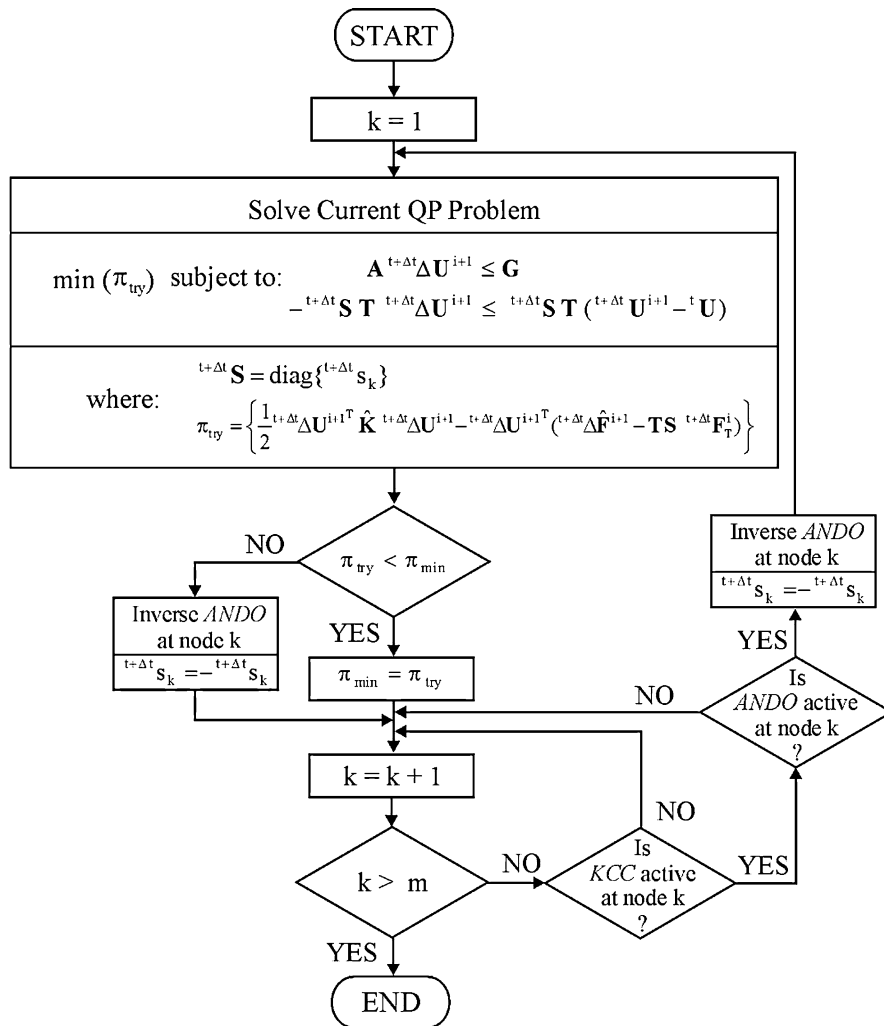
START

k = 1

Solve Current QP Problem

$\min (\pi_{\text{try}})$ subject to:
$$\mathbf{A}\,^{t+\Delta t}\Delta\mathbf{U}^{i+1} \leq \mathbf{G}$$
$$-\,^{t+\Delta t}\mathbf{S}\,\mathbf{T}\,^{t+\Delta t}\Delta\mathbf{U}^{i+1} \leq \,^{t+\Delta t}\mathbf{S}\,\mathbf{T}\,(^{t+\Delta t}\mathbf{U}^{i+1} - \,^{t}\mathbf{U})$$

where:
$$^{t+\Delta t}\mathbf{S} = \text{diag}\{^{t+\Delta t}s_k\}$$
$$\pi_{\text{try}} = \left\{ \frac{1}{2}\,^{t+\Delta t}\Delta\mathbf{U}^{i+1\,T}\,\hat{\mathbf{K}}\,^{t+\Delta t}\Delta\mathbf{U}^{i+1} - \,^{t+\Delta t}\Delta\mathbf{U}^{i+1\,T}(^{t+\Delta t}\Delta\hat{\mathbf{F}}^{i+1} - \mathbf{TS}\,^{t+\Delta t}\mathbf{F}_T^i) \right\}$$

$\pi_{\text{try}} < \pi_{\text{min}}$

NO

Inverse *ANDO* at node k
$$^{t+\Delta t}s_k = -\,^{t+\Delta t}s_k$$

YES

$\pi_{\text{min}} = \pi_{\text{try}}$

Inverse *ANDO* at node k
$$^{t+\Delta t}s_k = -\,^{t+\Delta t}s_k$$

YES

Is *ANDO* active at node k ?

NO

k = k + 1

k > m

NO

Is *KCC* active at node k ?

NO   YES

YES

END

Fig. 15 Details of nondifferentiable optimisation algorithm

field developed in the shot peening process as a result of inhomogeneous elasto-plastic deformation.

### 6.1 Contact between a cantilever beam and a flat rigid surface

The first example involves the dynamic frictionless contact between a cantilever beam and a flat rigid surface (Fig. 16). A beam of length $l = 0.2$ m and height $h = 0.02$ m is made from aluminum with the following properties: E = 73 GPa, $v = 0.33$, and $\rho = 2770$ kg/m³. A step load of F = 10^3 N was applied. The gap between the beam and the rigid foundation was selected to be equal to the maximum static deflection of the beam. A time step of $\Delta t = 5 \times 10^{-6}$ s was used. The results were obtained using the new VI approach and compared with the traditional penalty method. The displacements were normalized by the maximum static deflection of the beam, while the user-defined penalty stiffness parameters were normalized by the stiffness of the beam at the point of application of the load.

Figure 17 shows the normalized tip displacement predicted using the newly developed algorithm as well as the traditional penalty method. The results show that the VI does not result in any interpenetration. For low values of k_N, the beam experiences excessive unrealistic interpenetration. The penalty-based results approach the VI solution only for k_N/k_{beam} values approaching 10^2, or higher.
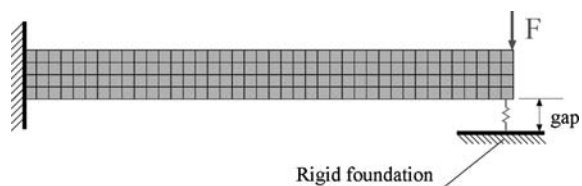
Fig. 16 Dynamic contact between a cantilever beam and a rigid foundation
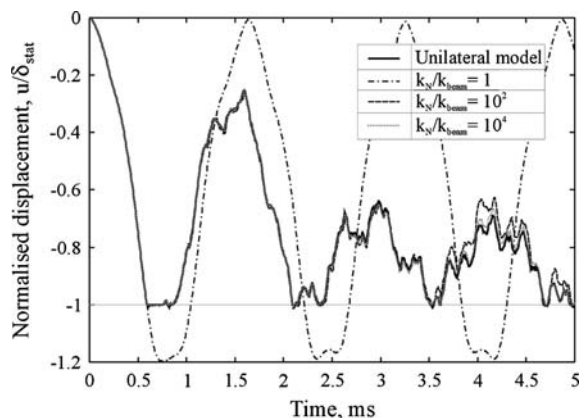
Fig. 17 Time variation of normalized tip displacement: comparison between VI and penalty based solution

6.2 Analysis of spur gear contact

Gearing is one of the most effective methods of transmitting rotary motion and torque from one shaft to another, with or without a change of speed or direction of motion. The design of efficient and reliable gearing systems is governed, amongst other factors, by their ability to withstand the frictional contact conditions and the high bending stresses experienced at the tooth root.

Whilst the frictional contact conditions govern the pitting and spalling resistance of a given gear, the high stress concentration present at its root determines the root fillet bending stresses and ultimately its fatigue life. Recent advances in turbomachinery have resulted in increasing the load carrying capacity and speeds of gear trains which may, however, make them vulnerable to failure, especially if the designer is unable to accurately and reliably predict the states of contact and root stresses of the mating gears.

Unfortunately, conventional gear strength equations are inaccurate and inadequate to treat modern high speed high load gearing, since they model the gear tooth as a cantilever beam and provide formulae

which relate the maximum root bending stress to the design parameters. To overcome the difficulties and limitations associated with the use of these semi-empirical formulae, the newly developed algorithms were use to predict root stresses in gears.

Analysis of a rack-pinion assembly

The present rack-pinion contact problem can be approached from a number of different directions according to the objectives and level of accuracy desired. The gear assembly considered conforms to AGMA standards and is shown schematically in Fig. 18. It comprises a 20 teeth pinion with a rack. Both components have involute teeth profiles of 20° pressure angle, a module of 5 mm, an addendum of one module and dedendum of 1.25 module. The components are assumed to be made of steel with the following elastic properties: Young's modulus $E = 210$ GPa and Poisson's ratio $v = 0.3$. Furthermore, the load is represented by an upward force per unit thickness of the rack $P_0 = 90$ kN/m under the assumptions of plane stress conditions. Several cases were considered in this study, such as: (i) one-pair, (ii) two-pairs, and (iii) three-pairs of teeth in contact.

The objective of considering two-pairs of contacting teeth are: (i) to predict the effect of load-sharing upon the resulting stresses, (ii) to examine possible interaction effects between the adjacent teeth and (iii) to examine the ability of the proposed method to treat problems involving multiple contact surfaces
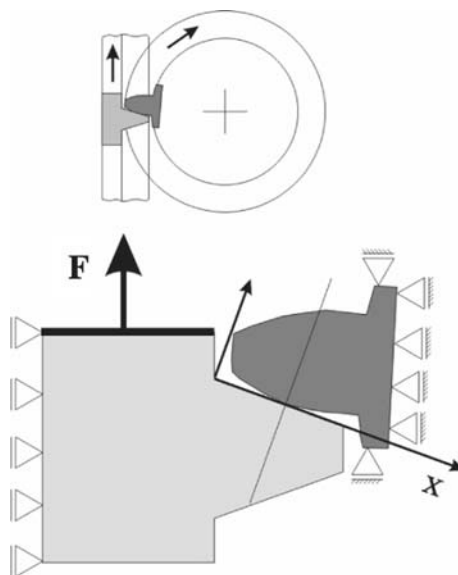
Fig. 18 Modelling of a pinion-rack problem
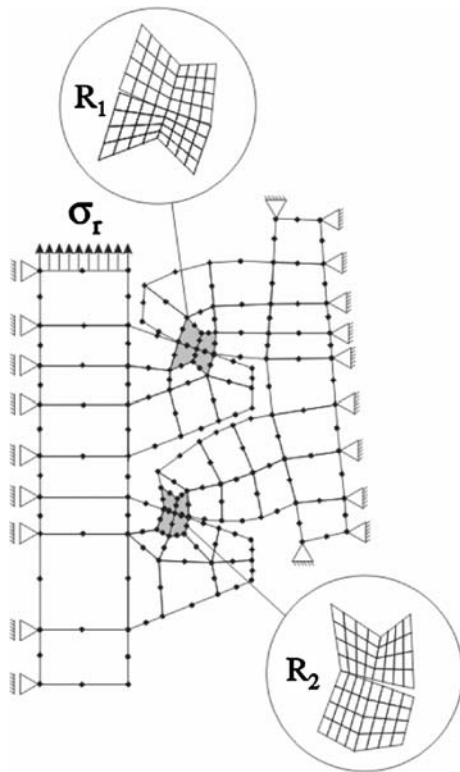
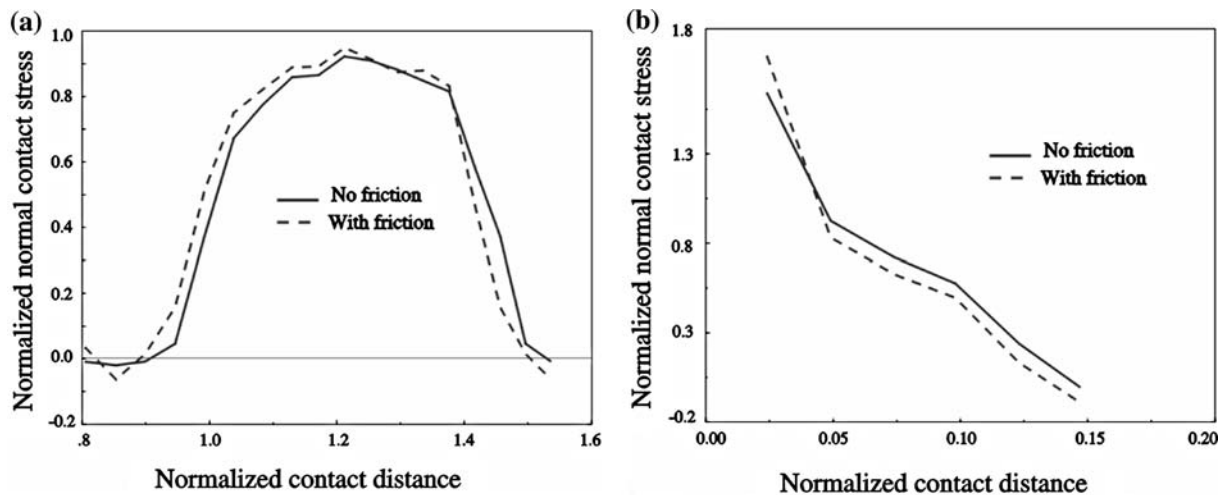**Fig. 19** Coarse and refined mesh for the case of two-pairs of teeth in contact

refined mesh was used to evaluate the contact stresses at the desired areas. It is worth pointing out that transferring the boundary conditions from the coarse to the refined model is the most intricate part of using refined mesh modelling. In this work, the boundary conditions were transferred as interpolated displacements, from the initial coarse model, using the appropriate shape functions.

Both frictionless and frictional contact problems were considered. Figure 20 shows the distribution of the resulting normal contact stress over the contact regions R_1 and R_2. In this case, contact region R_1 represents a pitch point contact position and contact region R_2 represents a point past the start of contact. The results reveal a reduction in the normal contact stress as a result of load sharing between the contacting teeth. In contact region R_2, the contact stress decreases monotonically along the contact length.

6.3 Analysis of fir-tree region in aeroengine turbine discs

Contact stresses, interface conditions (friction, surface roughness, residual stress), and the detailed geometry of the fir-tree joint determine the severity of the resulting stress field. These contact stresses do not remain constant but vary during operation. It is generally accepted that fretting initiated fatigue cracks are due to intermittent high frequency engine resonance and load fluctuations resulting from a change in engine speed or power requirements.

simultaneously. Figure 19 shows both the coarse and refined mesh of the examined model. In this case, the coarse mesh was used to characterise the overall structural response of the contacting teeth, while the

Fig. 20 Distribution of normal contact stresses in contact zones: (**a**) region R_1 and (**b**) region R_2

It was therefore the objective of this study to evaluate the contact behaviour of turbine disc assemblies using the finite element method. Specifically, our attention was devoted to examining the effect of the critical geometric features upon the contact stress distribution at the different teeth of a fir-tree joint. These features, shown in Fig. 21, include the number of fir-tree teeth n, flank length l, contact angle α, flank angles β and γ which define the tooth pitch.

The fir-tree models were meshed with four-noded quadrilateral elements (Fig. 22). Plane stress conditions were assumed. No attempt was made to model the blade in details, except insofar as providing the necessary centrifugal loading at the interface. All models were subjected to centrifugal loadings.

The material properties used for modelling the blade and the disc were that of a typical Nickel alloy used in disc design; namely, INCONEL 720. This material is creep and fatigue resistant. In this work, the following values were used: Young's modulus $E = 220$ GPa, 0.2% Proof stress $\sigma_{0.2} = 635$ MPa, Poisson's ratio $v_{0.2} = 0.29$, density $\rho_{0.2} = .8510$ kg/m^3 and coefficient of friction $\mu = 0.0, 0.1$ and 0.5.

Figure 23a represents a composite image of the photoelastic maximum shear stress contours and the FE predictions. The isochromatic fringe patterns compare well with the maximum shear stress contour plots obtained with FE analysis. In particular, the fringe patterns depict similar contact stresses and stress concentration around the different fillets. The results also reveal that peak stresses occur within the contact region at the bottom and top of each flank. Figure 23b shows the normalised maximum shear stress
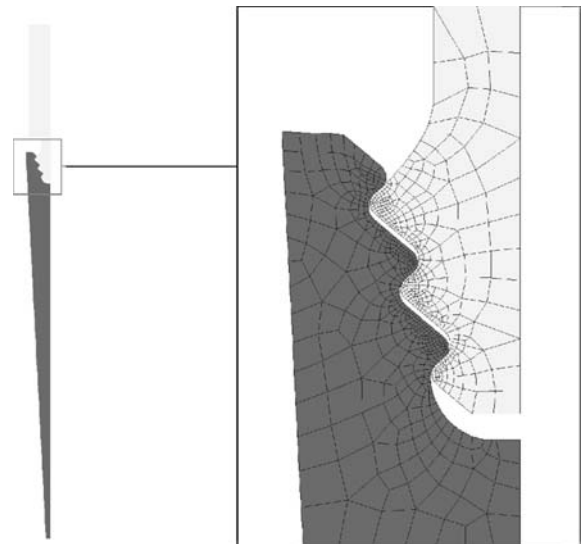

Fig. 22 Two-dimensional discretized geometry

trajectories along the bottom tooth interface. A maximum discrepancy of about 15% exists between the two-dimensional FE predictions and the photoelastic results.

The study also examined the effect of the critical geometrical features, such as the number of teeth, flank length and flank angle upon the stress field in the disc. It was further extended to account for the effect the interfacial friction between the disc and attached blades upon the resulting contact stresses at the interface. The results revealed the following:

(i) the maximum stress occurs at and just below the lower contact point along the bottom tooth of a turbine disc,
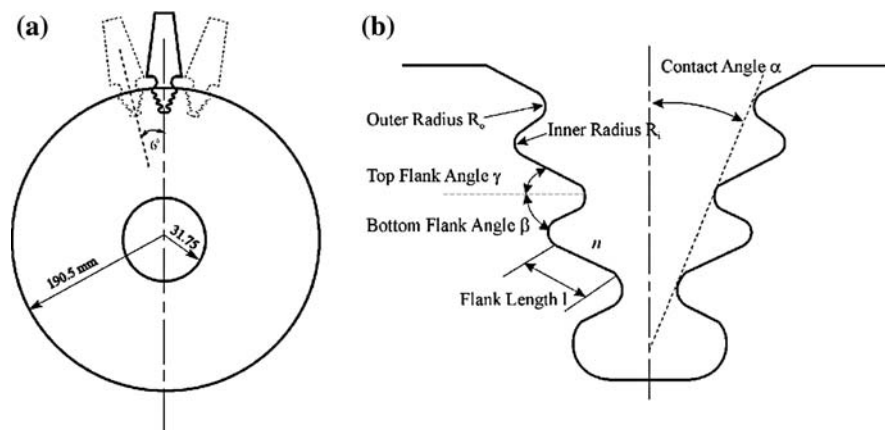
Fig. 21 Schematic of (**a**) disc and (**b**) fir-tree geometry
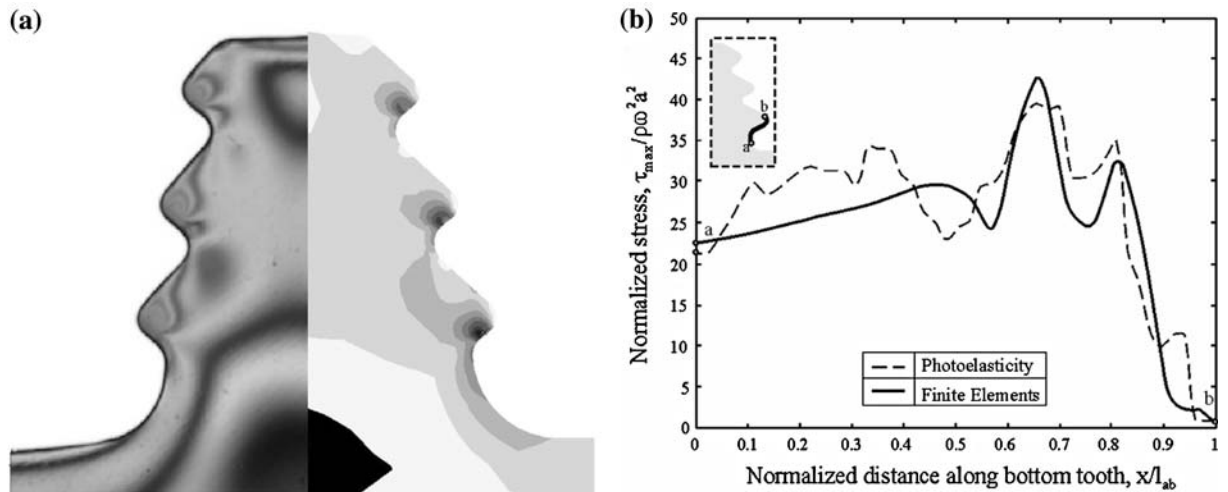
Fig. 23 (a) Composite image of the photoelastic maximum shear stress contours and FE predictions and (b) normalised maximum shear stress along the bottom contact region

(ii)   the position and magnitude of the maximum stress is dramatically influenced by the number of teeth, the contact angle, the flank length and flank angles,

(iii)   the coefficient of friction influences the peak stress value.

6.4 Residual stress prediction in shot peening

In this case study, we examine the mechanically induced residual stress field developed in the outermost material layers due to the shot peening treatment. A finite element analysis was conducted to study the effect of the mechanical properties of the target material and the shot velocity of peeing media. The following dimensions were selected for the target: width $W = 5R$ and height $H = 4R$, where $R = 0.5$ mm is the radius of the shot. These dimensions were carefully selected to ensure that the boundary conditions do not affect the solution. The shot used was assumed to be rigid and modeled with the following properties: modulus of elasticity $E = 2 \times 10^4$ GPa, Poisson's ratio $v = 0.3$, and mass density $\rho = 7800$ kg/m^3. The following material properties were assumed for the high strength steel target considered: modulus of elasticity $E = 200$ GPa, Poisson's ratio $v = 0.3$, and mass density $\rho = 7800$ kg/m^3. An elasto-plastic rate independent bilinearly hardening material behavior was assumed

with an initial yield stress $\sigma_0 = 600$ MPa and strain-hardening coefficient $H = 20, 1000$ and 2000 MPa (corresponding to $H/E = 0.01, 0.5$ and 1.0%, respectively). Three different impact velocities were used: $v = 50, 75$ and 100 m/s. Coulomb's law, with friction coefficient $\mu = 0.25$, was assumed. The total integration time was $t_k = 10$ μs with a maximum time step of $\Delta t_{max} = 2 \times 10^{-8}$s. Generalized-$\alpha$ time integration scheme was utilized with the following parameters $\gamma = \sqrt{2} - 1/2, \beta = 1/2, \alpha_H = \sqrt{2} - 1$, and $\alpha_B = 0$. Four-noded axisymmetric finite elements, with large strain and displacement capabilities, were used to discretize the target (Fig. 24). A fine mesh was used in the impact region where high stress gradients are expected. The rest of the target was discretized using larger elements. Convergence tests were conducted revealing only minor variations.

Figure 25 shows the time variation of the velocity and total contact force during the entire impact process for an initial shot velocity $v = 75$ m/s and strain hardening $H = 1000$ MPa. The results show that contact with the target lasts for 1.70 μs. However, the shot rebound began at about 1.25 μs, at which time the shot approached zero velocity. The terminal shot velocity was 17.3 m/s leading to a coefficient of restitution of 23%. This effectively indicates that approximately 95% of the total initial kinetic energy of the shot was dissipated in plastic work, while the remainder was used in elastic rebound. Figure 25b shows that the maximum
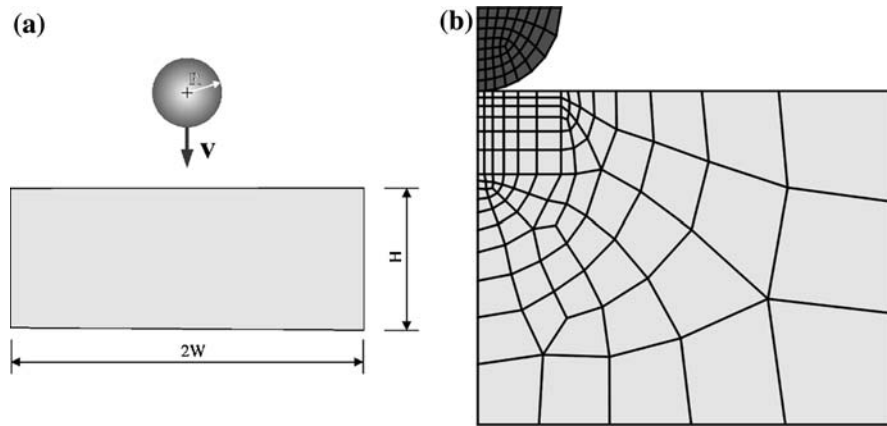
Fig. 24 Shot-peening problem: (**a**) geometry and (**b**) discretized model
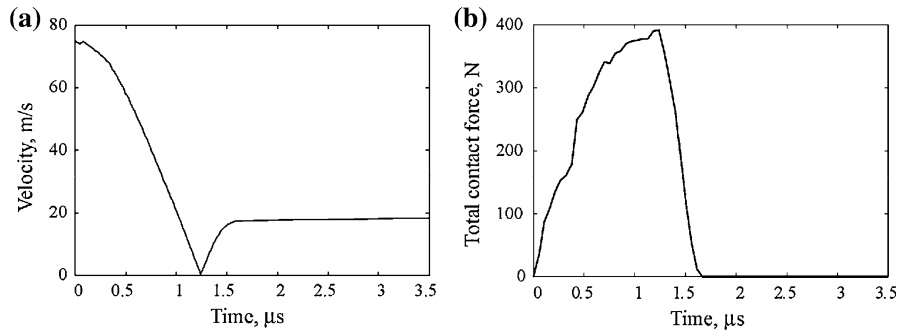
Fig. 25 Time history of: (**a**) the shot velocity, and (**b**) the total contact force

contact force is reached when the shot begins its rebound trip.

Figure 26a shows the variation of the normalized unloading residual stress σ_r^r/σ_0, beneath the centre line of the shot for the three selected impact speeds. The results show that a 100% increase in the shot velocity results in: (i) a significant increase in the depth of the compressed layer, and (ii) an insignificant change in the position and the magnitude of the maximum residual stress σ_r^r. It is interesting to note that the change in the residual stress from compression to tension coincides with the depth of the plastic zone.

Figure 26b shows the residual stress variation along the depth of the target for the three values of hardening coefficient and the shot velocity of 75 m/s. The results show that increasing the hardening coefficient of the material results in an increase in the depth of the compressed layer, little change in the magnitude of the maximum subsurface residual

stress σ_r^r, and a decrease in the surface residual stress.

7 Conclusions

In this paper, a new incremental variational inequality-based (VI) formulation is developed in an updated Lagrangian framework for the analysis of elasto-plastic dynamic frictional contact problems involving large deformation. The kinematic contact conditions are developed in a consistent manner based on the use of spline interpolation to represent contact surface. The developed finite element procedures are based on dividing the physical problem into sub-problems. These sub-problems are solved iteratively using Lagrange multipliers and/or mathematical programming in order to identify the candidate contact surface and contact stresses. In order to solve the dynamic expressions, we employ the generalized-α

(a)

(b)
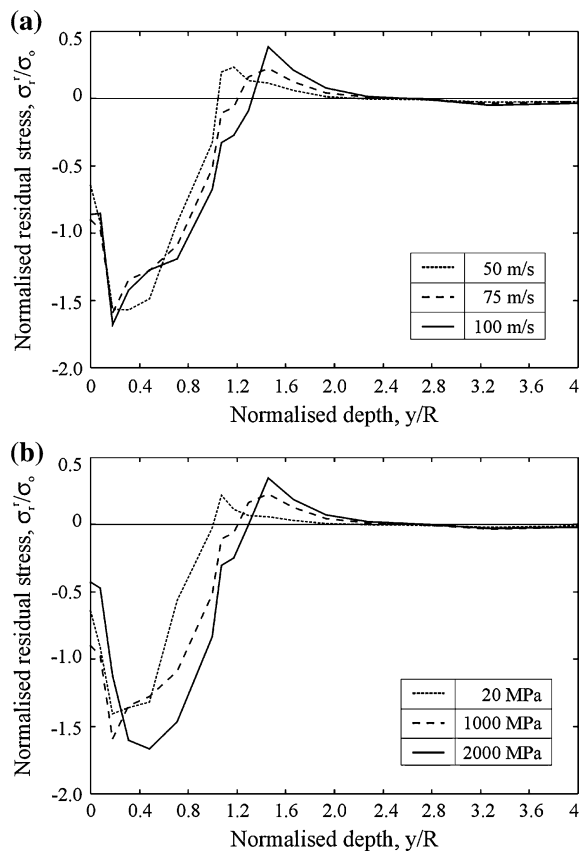
Fig. 26 Effect of: (**a**) shot velocity and (**b**) strain hardening rate upon normalized transverse residual stress

time integration scheme thus ensuring a convergent and accurate results. The newly developed approaches guarantee the accurate imposition of the active kinematic contact constraints.

References

Adams, R.A.: Sobolev Spaces. Academic Press, New York (1975)

Aliabadi, M.H., Brebbia, C.A.: Advanced Formulations in Boundary Element Methods. Southampton Computational Mechanics Publications (1993)

ANSYS version 5.5, User's Manual. Ansys Engineering Systems, Houston (1999)

Balinski, M.L., Wolfe, P.: Mathematical Programming Study 3: Nondifferentiable Optimization. North-Holland Publishing Company, Amsterdam (1975)

Bathe, K.J.: Finite Element Procedures. Prentice-Hall, New Jersey (1996)

Bischoff, D.: Indirect optimization algorithms for nonlinear contact problems. Proc. Mathematics of Finite Elements

and Applications V, pp. 533–545. Uxbridge, England (1984)

Bogomolny, A.: Variational formulation of the roller contact problem. Math. Methods Appl. Sci. **6**, 84–96 (1984)

Böhm, J.: A comparison of different contact algorithms with applications. Comput. Struct. **26**, 207–221 (1987)

Chaudhary, A.B., Bathe, K.J.: A solution method for planar and axisymmetric contact problems. Int. J. Numer. Methods Eng. **21**, 65–88 (1985)

Chaudhary, A.B., Bathe, K.J.: A solution method for static and dynamic analysis of three-dimensional contact problems with friction. Comput. Struct. **24**, 855–873 (1986)

Chung, J., Hulbert, G.M.: A time integration algorithm for structural dynamics with improved numerical dissipation: the generalized-α method. J. Appl. Mech. **60**, 371–375 (1993)

Crisfield, M.A.: Non-linear Finite Element Analysis of Solids and Structures. John Wiley & Sons, New York (1997)

Czekanski, A., Meguid, S.A.: Analysis of dynamic frictional contact problems using variational inequalities. Finite Elem. Anal. Design, **37**, 861–879 (2001)

Czekanski, A., Meguid, S.A.: Solution of dynamic frictional contact problems using nondifferentiable optimization. Int. J. Mech. Sci. **43**, 1369–1386 (2001)

Czekanski, A., Meguid, S.A., El-Abbasi, N., Refaat, M.H.: On the elastodynamic solution of frictional contact problems using variational inequalities. Int. J. Numer. Methods Eng. **50**, 611–627 (2001)

Duvaut, G., Lions, J.L.: Inequalities in Mechanics and Physics. Springer-Verlag, Berlin Heilberg (1976)

El-Abbasi, N.: A new strategy for treating frictional contact in shell structures using variational inequalities, Ph.D. Thesis, University of Toronto (1999)

Farin, G.: Curves and Surfaces for Computer-Aided Geometric Design—A Practical Guide. Academic Press, Toronto (1997)

Ferris, M.C., Pang, J.S.: Engineering and economic applications of complementarily problems. SIEM Rev. **39**, 669–713 (1997)

Fletcher, R.: Practical Methods of Optimization. John Wiley and Sons, New York (1982)

Gladwell, G.M.L.: Contact Problems in the Classical Theory of Elasticity. Sijhoff and Noordhoff, Alphen aan den Rijn (1980)

Glowinski, R., Lions, T.C., Tremolieres, R.: Numerical Analysis of Variational Inequalities. North Holland Publishing Company, Amsterdam (1981)

Hertz, H.: Über die Berührung Fester Elastischer Körper (On the Contact of Elastic Solids). Journal für die Reine und Angewandte Mathematik **92**, 156–171 (1882)

Hlavacek, I., Haslinger, J., Necas, J., Lovisek, J.: Solution of Variational Inequalities in Mechanics. Springer-Verlag, New York (1988)

Hughes, T.J.R.: The Finite Element Method: Linear Static and Dynamic Finite Element Analysis. Prentice-Hall, New Jersey (1987)

Hughes, T.J.R., Taylor, R.L. Sackman, J.L., Curnier, A., Kanoknukulchai, W.: A finite element method for a class of contact-impact problems. Comput. Methods Appl. Mech. Eng. **8**, 249–276 (1976)

Johnson, K.L.: Contact Mechanics. Prentice-Hall, Cambridge University Press, New York (1985)

Kenny, B., Patterson, E.A., Said, M., Aradhya, K.S.S.: Contact stress distributions in a turbine disc dovetail type joint—a comparison of photoelastic and finite element results. Strain **27**, 21–24 (1991)

Kikuchi, N., Oden, J.T.: Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Method. SIAM, Elsevier Science, Amsterdam (1988)

Kinderlehrer, D., Stampacchia, G.: An Introduction of Variational Inequalities and Their Applications. Academic Press, New York (1980)

Klarbring, A.: On discrete and discretized non-linear elastic structures in unilateral contact (stability, uniqueness and variational principles). Int. J. Solids Struct. **24**, 459–4789 (1988)

Laursen, T.A., Simo, J.C.: A continuum-based finite element formulation for the implicit solution of multibody, large deformation frictional contact problems. Int. J. Numer. Methods Eng. **36**, 3451–3485 (1993)

Lemarechal, C.: Nondifferentiable optimization; subgradient and ε-subgradient methods. Math. Program. Study **7**, 384–387 (1974)

MARC, Theoretical Manual. MARC Software International Inc., Palo Alto, California (1993)

Mifflin, R.: A modification and an extension of lemarechal's algorithm for nonsmooth minimization. Math. Program. Study **17**, 77–90 (1982)

Newmark, N.M.: A method of computation for structural dynamics. ASCE J. Eng. Mech. Div. **13**, 67–94 (1959)

Oden, J.T., Martins, J.A.C.: A numerical analysis of class of problems in elastodynamics with friction. Comput. Methods Appl. Mech. Eng. **40**, 327–360 (1983)

Oden, J.T., Martins, J.A.C.: Models and computational methods for dynamic friction phenomena. Comput. Methods Appl. Mech. Eng. **52**, 527–634 (1985)

Parisch, H.: Consistent tangent stiffness matrix for three-dimensional non-linear contact analysis. Int. J. Numer. Methods Eng. **28**, 1803–1812 (1989)

Qin, Q.H., He, X.Q.: Variational principles, FE and MPT for analysis of non-linear impact-contact problems. Comput. Methods Appl. Mech. Eng. **122**, 205–222 (1995)

Rabier, P., Martins, J.A.C., Oden, J.T., Campos, L.: Existence and local uniqueness of solutions to contact problems in elasticity with nonlinear friction law. Int. J. Mech. Sci. **24**, 1755–1768 (1986)

Refaat, M.H., Meguid, S.A.: On the elastic solution of frictional contact problems using variational inequalities. Int. J. Mech. Sci. **36**, 329–342 (1994)

Refaat, M.H., Meguid, S.A.: A novel finite element approach to frictional contact problems. Int. J. Numer. Methods Eng. **39**, 3889–3902 (1996)

Shimizu, K., Ishizuka, Y., Barol, J.F.: Nondifferentiable and Two-level Mathematical Programming. Kluwer Academic Publishers, Boston (1997)

Shor, N.Z.: Minimization Methods for Non-differentiable Functions. Springer-Verlag, New York (1985)

Wierzbicki, A.P.: Lagrangian functions and nondifferentiable optimization. In: Nurminski, E.A. (ed.) Progress in Non-differentiable Optimization, pp. 97–124. International Institute for Applied System Analysis, Luxemburg (1982)

Wriggers, P., Simo, J.C.: A note on tangent stiffness for fully nonlinear contact problems. Commun. Appl. Numer. Methods **1**, 199–203 (1985)

Wriggers, P., Vu Van, T., Stein, E.: Finite element formulation of large deformation impact-contact problems with friction. Comput. Struct. **37**, 319–331 (1990)

Zhong, Z.H.: Finite Element Procedures for Contact-Impact Problems. Oxford University Press, Oxford (1993)

Zienkiewicz, O.C., Taylor, R.L.: The Finite Element Method. McGraw-Hill, London (1988)