Vol. **2**, **4**, 1–49 (1995)

Archives of Computational Methods in Engineering
State of the art reviews

Finite Element Algorithms for Contact Problems

P. Wriggers
Institut für Mechanik, TH Darmstadt
D-64289 Darmstadt, Hochschulstrasse 1, Germany

Summary

The numerical treatment of contact problems involves the formulation of the geometry, the statement of interface laws, the variational formulation and the development of algorithms. In this paper we give an overview with regard to the different topics which are involved when contact problems have to be simulated. To be most general we will derive a geometrical model for contact which is valid for large deformations. Furthermore interface laws will be discussed for the normal and tangential stress components in the contact area. Different variational formulations can be applied to treat the variational inequalities due to contact. Several of these different techniques will be presented. Furthermore the discretization of a contact problem in time and space is of great importance and has to be chosen with regard to the nature of the contact problem. Thus the standard discretization schemes will be discussed as well as techiques to search for contact in case of large deformations.

1. INTRODUCTION

Boundary value problems involving contact are of great importance in industrial applications in mechanical and civil engineering. The range of application includes metal forming processes, drilling problems, bearings, crash analysis of cars, car tires or cooling of electronic devices. Other applications are related to biomechanics where human joints, implantats or teeth are of consideration. Due to this variety contact problems are today combined either with large elastic or inelastic deformations including time dependent responses. Thermal coupling might have to be considered, see the cooling of electronic devices, the heat removal within nuclear power plant vessels or thermal insulation of astronautic vehicles. Even stability behaviour has to be linked to contact, like wrinkling arising in metal forming problems.

Due to this technical importance a great number of researchers have investigated contact problems. In the ancient egypt people needed to move large stone blocks to build the pyramids and thus had to overcome the frictional force associated with it. Thus many known researchers in the past have investigated frictional contact problems, amongst them were Da Vinci, Amontons, Newton, Coulomb. Their investigations were based on the assumption of rigid bodies. Starting with the classical analytical work of Hertz (1882) on the elastic contact of two spheres the deformation of the bodies being in contact has been taken into account. However only very few problems involving contact can be solved analytically. Thus for most industrial applications numerical methods have to be applied when the contacting bodies have complex geometries . Due to that the solution of contact problems with finite element methods has a relatively long history, see Wilson, Parsons (1970) or Chan, Tuba (1971) for early treatments.

In this overview article we will restrict ourselves mainly to finite element techniques for the treatment of contact problems despite many other numerical schemes and analytical approaches could be discussed as well. Furthermore we like to note that the description of the mechanical behaviour of the bodies coming into contact will not be investigated in detail, although this is of great importance. This article thus concentrates on the behaviour in the contact interface. The associated formulation and discretization within the finite element method will be considered as well as the development of algorithms.

©1995 by CIMNE, Barcelona (Spain). ISSN: 1134–3060 Received: March 1995

The following introductory remarks are related to the steps which have to be followed when treating contact problems within the finite element method.

1. Continuum based contact kinematics,
2. Constitutive equations for contact interfaces,
3. Weak form of contact contributions and overall solution strategies for contact problems,
4. Discretization of contact surfaces,
5. Algorithms for the integration of constitutive equations in the contact area,
6. Contact search algorithms,
7. Adaptive methods for contact problems.

1. Contact kinematics. Since the contact area is not known a priori, and depends in a nonlinear way on the loading, contact problems are nonlinear even for linear elastic solids. Furthermore many technical contact problems involve also large deformations of the bodies being in contact; thus we will formulate all contact relations for finite deformations. In general two steps have to be followed to set up the contact geometry: the search for contact and development of the local kinematical relations. Here we will focus on the local kinematical relations, searching algorithms are discussed later, see part 6. In a large deformation, continuum based formulation of contact kinematics the distance between the bodies being in contact is minimized as can be found for the classical non–penetration condition in e.g. Alart, Curnier (1988). In case that a small penetration due to the approach of the two bodies in contact has to be allowed the contact kinematics are developed in Wriggers, Miehe (1992). This non–penetration function plays also a significant role for the definition of the tangential velocity in the contact interface which is needed to formulate frictional problems, see e.g. Simo, Laursen (1992), Wriggers, Miehe (1992), Laursen, Simo (1993), Wriggers, Miehe (1994) or Klarbring (1994).

2. Constitutive equations for contact interfaces. Due to the precision which is needed to resolve the mechanical behaviour in the contact interface, different approaches have been used in the literature to model the mechanical behaviour in contact area. Two main lines can be followed within the finite element method to impose contact conditions in normal direction. These are the formulation of the non–penetration condition as a purely geometrical constraint and the development of constitutive laws for the micromechanical approach within the contact area.

The first formulation is in general used for problems with “low contact precision” where the most essential necessity is the correct enforcement of the geometrical constraints like in crash or forming simulations. In this case it is not possible to specify constitutive relations in the contact interface. Here the normal contact pressure is related to the reaction in the contact area and can be deduced from the constraint equations. This procedure is the classical way to formulate contact constraints; thus numerous researcher have used this strategy. For applications using the finite element method we like to mention the early work by Wilson, Parsons (1970) or Chan, Tuba (1971) for small deformation problems or the work by Alart, Curnier (1988) for large deformations.

However, as discussed before, there exist also contact problems where the knowledge of the micromechanical approach is essential for a proper treatment of the physical phenomena. Then interface compliances are needed for these problems with “high contact precision”. Constitutive equations for the normal contact can be developed by investigating the micomechanical behaviour within the contact surface. Associated models have been developed based on experiments, see e.g. Greenwood, Williamson (1966) or Kragelsky, Dobychin, Kombalov (1982). The micromechanical behaviour depends in general on material parameters like hardness and on geometrical parameters like surface roughness. It should be noted that the real micromechanical phenomena are extremly complex: due to very high local pressures, e.g. in the case of an impact, chemical reactions can be initiated in the interface by the mechanical forces. The models which are used try only to capture the most important phenomena and assume either an elastic or a plastic deformation of the asperities being in real contact in the interface.

The interfacial behaviour in the tangential direction (frictional response) is even more complicated. Some researchers try to formulate a third body in the interface which has special properties and is only present in the moment of the tangential mechanical loading, see Kragelsky (1956). We will here restrict ourselves to more simple formulations which yield constitutive equations for frictional contact as long as one does not assume perfect stick within contact area. The most frequently used constitutive equation is the classical law of Coulomb. However, other frictional laws are available which take into account local, micromechanical phenomena within the contact interface, see e.g. Woo, Thomas (1980). An extensive overview may be found in Oden, Martins (1986). The main governing phenomena are adhesion of the surfaces and ploughing of the asperities. For the physical background see e.g. Tabor (1981). During the last years frictional phenomena have also been considered within the framework of the theory of plasticity. This leads to non–associative slip rules, different relations have been proposed for frictional problems by e.g. Bowden and Tabor (1964) and Michalowski, Mroz (1978). Further discussion is contained in Curnier (1984). The application of constitutive equations for friction within finite element calculation can be found in e.g. Fredriksson (1976), Wriggers, Vu Van, Stein (1990).

In cases where thermomechanical contact has to be considered, a "high contact precision" formulation must be applied to account correctly for the pressure dependency of the heat conduction in the contact area. This is due to the fact that the heat conduction depends on the approach of the two rough surfaces being in contact, see section 3.3. In this context models have been discussed for the onstitutive behaviour in normal direction on the basis of statistical methods, see e.g. Cooper, Mikic, Yovanovich (1969) or Song, Yovanovich (1987). A finite element treatment for thermomechanical contact problems can be found in Zavarise (1991), Zavarise, Wriggers, Stein, Schrefler (1992a,b) and in combination with frictional heating in Wriggers, Miehe (1994). Also other contact phenomena like wear, see e.g. Johannson, Klarbring (1992), need special constitutive laws which have to be developed in the interface.

3. Weak form of contact contributions and overall solution strategies. The weak formulation of contact problems leads to variational inequalities, see Duvant, Lions (1976) since contact conditions are represented as inequality constraints. Different possibilities exist for the numerical solution of these problems. Among them are the so called active set strategies which are applied in combination with Lagrangian multiplier or penalty techniques, see e.g. the text books of Bertsekas (1984) or Luenberger (1984); these methods are well known in optimization theory. Other solution schemes are based on mathematical programming, see e.g. Conry, Seireg (1971) or Klarbring (1986), who applied this method to frictional contact problems.

Most standard finite element codes which are able to handle contact problems use either the penalty or the Lagrangian multiplier method, for an overview and the mathematical framework, see e.g. Kikuchi, Oden (1988). Each of the methods has its own advantages and disadvantages which will be discussed in detail in the following. The methods are designed to fulfill the constraint equations in normal direction in the contact interface. For the tangential part we need in general constitutive relations; associated techniques will be discussed later. A combination of the penalty and the Lagrangian multiplier technique leads to the so called augmented Lagrangian methods which try to combine the merits of both approaches. A general discussion of these techniques can be found in Glowinski, Le Tallec (1984) and with special attention also to inequality constraints in Bertsekas (1984). However this technique requires an algorithmic treatment, the Uzawa method, which increases the total number of iterations. For applications of augmented Lagrangian techniques to contact problems within the finite element method, see e. g. Wriggers, Simo, Taylor (1985), Simo, Laursen (1992) or for a symmetrical treatment of the frictional part Laursen, Simo (1993b) or Zavarise, Wriggers, Schrefler (1995). In the case of "high precision contact", when constitutive interface laws are employed, special augmented Lagrangian techniques are needed, since often ill–conditioning of the problem may occur, see Wriggers, Zavarise (1993).

In case of thermomechanical contact problems two fields – deformation and temperature – interact and thus have to be considered within the formulation. In the general setting these fields are coupled since the constitutive parameters depend on the temperature, the evolution of the thermal field is related to the deformation and heat can be generated by dissipative mechanisms like plastic deformations or frictional forces. The technical importance of these models has lately raised some interest in these phenomena, thus many contributions can be found in the literature. Here we discuss only the research which is directly related to the numerical treatment of contact problems within the finite element method. A finite element model based on micromechanical interface laws is derived in e.g. Zavarise, Wriggers, Stein, Schrefler (1992) for finite deformations. Wriggers, Zavarise (1993) have developed a simple element for the case of small deformations. In both approaches a global iterative procedure has been used for a stationary process. Staggered schemes, which treat the deformation and temperature fields seperately can be computationally more advantageous, see Wriggers, Miehe (1994) for thermomechanical contact and Simo, Miehe (1992) for thermomechanical problems without contact.

4. Discretization of the contact surfaces. When the discretization of contact surfaces is concerned one has to distinguish between the contact of two deformable bodies or the contact of a deformable body with a rigid obstacle. On a first glance it seems that the latter case is simply a special case of the first problem, which is true. But due to the fact that the surface description of a rigid obstacle can be given once and for all by the correct geometrical model this knowledge can be used within the discretization process. Hansson, Klarbring (1990) have developed a formulation based on CAD–surfaces, Williams, Pentland (1992) considered so called superquadrics to specify the geometry of contacting objects and Wriggers, Imhof (1993) formulated the contact problem with splines.

In the first applications of finite elements to contact problems of two deformable bodies only small changes in the geometry were assumed so that the geometrically linear theory could be applied. Then it is possible to incorporate the contact constraints on a purely nodal basis, see e.g. Francavilla, Zienkiewicz (1975). Later also contact elements were developed which resulted from a degenerated solid element, see e.g. Stadter, Weiss (1979) or the textbook of Kikuchi, Oden (1988). A mathematical study of these classes of elements which also accounts for the correct integration rules can be found in Oden (1981). All of the above mentioned elements need a discretization in which the element nodes match each other in the contact interface. For the general case of nodes being arbitrary distributed along the possible contact interface between two bodies, which can occur when automatic meshing is used for two different bodies, Simo, Wriggers, Taylor (1984) developed a segment approach to discretize the contact interface.

For the general case of contact including large deformations the most frequently used discretization is the so called node–to–segment approach. Here arbitrary sliding of a node over the entire contact area is allowed. Early implementations can be found in Hallquist (1976) or Hughes, Taylor, Kanuknukulchai (1978) which have been developed for more and more general cases, Hallquist, Goudreau, Benson (1983), Bathe, Chaudary (1985) and Wriggers, Vu Van, Stein (1990). Now some finite element codes include also self–contact, see Hallquist, Schweizerhof, Stillman (1992). Also the idea of contact segments proposed by Simo, Wriggers, Taylor (1984), has been followed up and applied to problems involving large deformations, see Papadopoulos, Taylor (1992).

A consistent linearization is needed within Newton procedures to solve the nonlinear contact problems incrementally. For a discretization using the node–to–segment approach Wriggers, Simo (1985) derived for two–dimensional large deformation problems the needed matrix formulation under the assumption of frictionless contact. The formulation for frictional contact can be found in Wriggers, Vu Van, Stein (1990). The associated tangent matrices for the frictionless three–dimensional case of a node–to–surface discretizations is developed in Parisch (1989). The special case of the contact of a body with a rigid obstacle is treated in Hansson, Klarbring (1990), Wriggers, Imhof (1993) or Heegard, Curnier (1993). The consistent linearization for a continuum based approach to contact problems has been derived in Laursen, Simo (1993a).

5. Algorithms for the integration of constitutive equations in the contact area. In general we have to distinguish three cases of constitutive equations in the contact interface. These are related to the normal, the tangential and the thermal part of the contact.

For the normal contact a mere function evaluation –like for finite elasticity– can be used to obtain for a given approach the contact pressure; even if the micromechanical derivation of the contact compliance involves plastic deformations. This is theoretically not satisfactory but up to now – due to the extremely complex behaviour in the contact interface – the only possible method for the macroscopic description of normal contact compliance.

The situation is different for friction. Then one has to solve an evolution equation for the frictional slip which needs special algorithms. In early finite element applications often so–called ”trial–and–error” algorithms have been applied, see e.g. (19), which might not converge in some cases. More reliable methods are provided by the mathematical programming approach, Klarbring (1986). Another way which is now becoming more and more standard for numerical simulations involving friction is related to the possibility to recast the fricional interface laws in terms of non–associated plasticity. First formulations and applications in finite element analysis are found in Fredriksson (1976). A theoretical basis was also provided by Michalowski, Mroz (1978). The major break through in terms of convergence behaviour and reliability of the solution algorithms came with the application of the return mapping schemes to frictional problems. Its application can be found in Wriggers (1987) or Giannokopoulos (1990) for geometrically linear problems. This approach provides the possibility to develop algorithmic tangent matrices which are needed to achieve quadratic convergence within Newton–type iterative schemes. Due to the non–associativity of the frictional slip these matrices are non–symmetrical. For the case of large deformations associated formulations have been developed in Ju, Taylor (1988) for a regularized Coulomb friction law and in Wriggers, Vu Van, Stein (1990) for different frictional laws formulated in terms of non–associated plasticity. A three–dimensional formulation can be found in Laursen, Simo (1993a) who also developed an algorithmic symmetrization (Laursen, Simo (1993b)), see also Zavarise, Schrefler, Wriggers (1995).

6. Contact search algorithms. The search for the active set of contact constraints is not trivial in case of large deformations since a surface point of a body may contact any portion of the surface of another body. Such point can even come into contact with a part of the surface of its own body. Thus the search for the correct contact location needs, depending on the problem, eventually considerable effort. An implementation where each node of a surface is checked against each element surface in the mesh is too exhaustive and thus computationally inefficient and refined algortihms have to be constructed. This especially true when the contact of more then two bodies has to be considered or when self–contact is possible.

The development of search algorithms can be split in two general approaches. The first is connected with the contact between a deformable and a rigid body. In this case the rigid body can be described by implicit functions such as superquadrics or hyperquadrics, see

Williams, Pentland (1992). This leads to a simple and efficient contact check for points lying on the surface of the deformable body. For the special case of cylinders or ellipses see also e.g. Hallquist, Schweizerhof, Stillman (1992).

In case that two or more deformable bodies contact each other or that self–contact of one body occurs the search algorithms are more complex and normally split into a global and a local search. Within the global search a hierarchical structure can be set up to find out which bodies, parts of the bodies, surfaces or parts of the surfaces are able to come into contact within a given time step or displacement increment, see e.g. Zhong, Nilsson (1989), Zhong (1993) or Williams, O'Connor (1995). Different methods can be applied to determine the possible contact partners. Lately a considerable impact has come from discrete finite element methods where several thousand particles have to be included in the contact search. Methods like space cell decomposition have been considered by Belytschko, Neal (1989) a combination with binary tree search can be found in Munjiza, Owen, Bicanic (1995); whereas Williams, O'Connor (1995) rely on heapsort algorithms for the global search.

Once the possible contactors are known the local search is needed to check whether a penetration has occured and to determine its exact location. Different possibilities exist to find the correct finite element surface which is associated with a node that might penetrate through this surface. Here the node–to–segment algorithm Hallquist (1978), the pinball technique, Belytschko, Neal (1991), or methods based on discrete function representation Williams, O'Connor (1995) can be applied among other possibilities.

7. Adaptive methods for contact problems. Since numerical methods for contact problems yield approximate solutions it is necessary to control the errors inherited in the method. During the last ten years research activities have been focused on adaptive techniques providing automatically a numerical modell which is accurate and reliable. The objective of adaptive techniques is to obtain a mesh which is optimal in the sense that the computational costs involved are minimal under the constraint that the error in the finite element solution is beyond a certain limit. Since the computational effort can be linked to the number of unknowns of the finite element mesh the task is to find a mesh with minimum number of unknowns or nodes for a given error tolerance. In general, adaptive methods rely on error indicators and error estimators which can be computed *a priori* or *a posteriori*. For an overview over different techniques, see e.g. Johnson (1987) and references therein. Based on the error distribution a new partially refined mesh can be constructed which yields a better approximate solution. To obtain an optimal mesh in the sense of an equal solution quality it is desirable to design the mesh such that the error contributions of the elements are equidistributed over the mesh. During the last years a growing number of papers has been devoted to this topic and applied to problems of solid and fluid mechanics, see e.g. Zienkiewicz, Taylor (1989).

The methods rely on error estimators which have been developed so far in different versions. The estimators which are most frequently used in solid mechanics for elastic problems are residual based error estimators, see e.g. Babuska, Rheinboldt (1978) or Johnson, Hansbo (1992), or error estimators which use superconvergence properties, see e.g. Zienkiewicz, Zhu (1987).

For frictionless contact problems *a priori* error estimators have been derived for linear elastic bodies, see e.g. Kikuchi, Oden (1988) or Hlavacek, Haslinger, Necas, Lovisek (1988). An adaptive method for problems with unilateral constraints has been developed by Lee, Oden, Ainsworth (1991) who treated as an example a free surface flow problem. In Wriggers, Scherf, Carstensen (1994) a residual based error estimator has been developed following an approach persued by Johnson, Hansbo (1992) for unilateral membrane problems. But also the Z^2 error estimators, due to Zienkiewicz, Zhu (1987), can be applied to contact problems, see Wriggers, Scherf (1995).

2. CONTACT GEOMETRY

This section summarizes relations which are necessary to formulate the geometrical contact conditions. In detail the penetration and the relative slip in the contact area are discussed. The first condition also includes the non–penetration condition which is used classically in contact mechanics. The derivation presented here can be used for frictional or frictionless problems. It follows closely the approach discussed in Wriggers, Miehe (1992) and Wriggers, Miehe (1994). Similar ideas may be found in Laursen, Simo (1993) or for the frictionless case in Curnier, Alart (1992).

We assume that two bodies which undergo large deformations can come into contact. Let $\mathcal{B}^\gamma$, $\gamma = 1, 2$, denote the two bodies of interest and $\boldsymbol{\varphi}_t^\gamma : \mathcal{B}^\gamma \rightarrow \mathbb{R}^3$ the associated deformation maps at time $t \in \mathbb{R}_+$. $\boldsymbol{\varphi}_t^\gamma$ maps points $\mathbf{X}^\gamma \in \mathcal{B}^\gamma$ of the reference configuration onto points $\mathbf{x}^\gamma = \boldsymbol{\varphi}_t^\gamma(\mathbf{X}^\gamma)$ of the current configuration.

Motivated by micromechanical investigations of contact problems we view the mechanical approach of the two contact surfaces as a microscopical penetration of the current mathematical boundaries $\boldsymbol{\varphi}_t^\gamma(\Gamma_c^\gamma)$. Note that we can recover the non–penetration condition as a limiting case. In this formulation $\Gamma_c^\gamma \subset \partial\mathcal{B}^\gamma$ are possible contact surfaces of the bodies $\mathcal{B}^\gamma$, see Figure 1 for an illustration of this concept. In what follows we denote $\boldsymbol{\varphi}_t^1(\Gamma_c^1)$ as the current slave surface which penetrates in the case of contact into the current master surface $\boldsymbol{\varphi}_t^2(\Gamma_c^2)$. The latter one plays within our formulation of the contact geometry the role of a (moving) reference surface. We parametrise the master surface in its reference and current configuration by the natural parameters ξ^1, ξ^2, i.e. we consider material curves $\mathbf{X}^2 = \hat{\mathbf{X}}^2(\xi^1, \xi^2) \subset \Gamma_c^2$ and $\mathbf{x}^2 = \hat{\mathbf{x}}_t^2(\xi^1, \xi^2) \subset \boldsymbol{\varphi}_t^2(\Gamma_c^2)$. Then the local deformation gradient of the master surface is given by $\mathbf{F}_t^2 := \mathbf{a}_\alpha^2 \otimes \mathbf{A}^{2\alpha}$ based on the tangent vectors of the contact surface $\mathbf{a}_\alpha^2 := \hat{\mathbf{x}}_{t,\alpha}^2(\xi^1, \xi^2)$ and $\mathbf{A}_\alpha^2 := \hat{\mathbf{X}}_{,\alpha}^2(\xi^1, \xi^2)$ with the standard relations $\mathbf{a}_\alpha^2 \cdot \mathbf{a}^{2\beta} = \delta_\alpha^\beta$ and $\mathbf{A}_\alpha^2 \cdot \mathbf{A}^{2\beta} = \delta_\alpha^\beta$ and $(\ )_{,\alpha}$ denotes differentiation with respect to ξ^α.

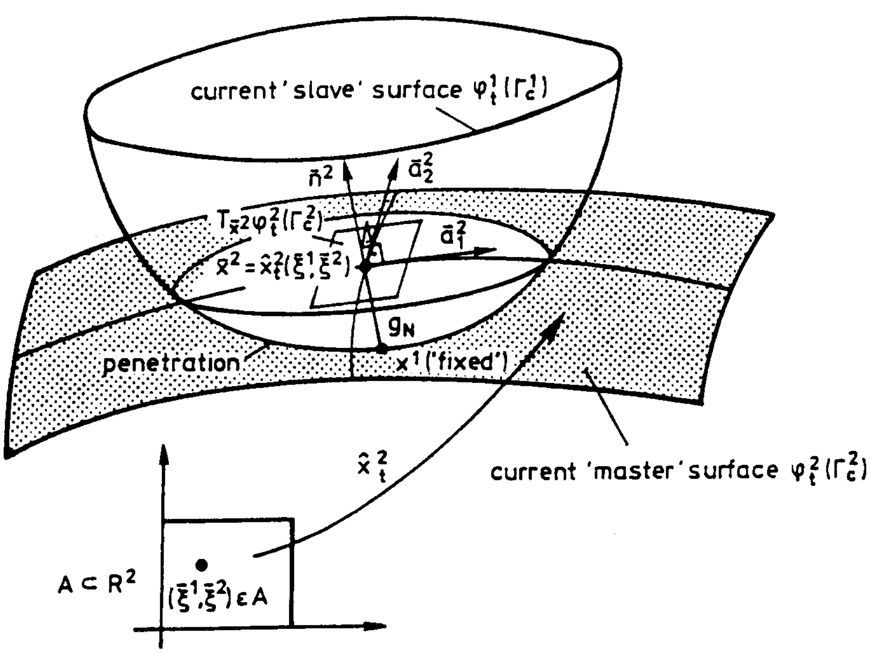


Figure 1. Contact geometry and geometrical approach

2.1 Penetration

As the first relevant function for the contact geometry we define a penetration function on the current slave surface $\boldsymbol{\varphi}_t^1(\Gamma_c^1)$ by setting, see Wriggers, Miehe (1992)

$$g_{N+} \;=\; \begin{cases} \| \,\mathbf{x}^1 - \hat{\mathbf{x}}_t^2\,(\bar{\xi}^1,\,\bar{\xi}^2)\, \| & \text{for } [\,\mathbf{x}^1 - \hat{\mathbf{x}}_t^2(\bar{\xi}^1,\,\bar{\xi}^2)\,]\cdot\bar{\mathbf{n}}^2 < 0 \\ 0 & \text{otherwise} \end{cases} \;. \tag{1}$$

Here $(\bar{\xi}^1,\,\bar{\xi}^2)$ is the minimizer of the distance function for a given slave point $\mathbf{x}^1$

$$\hat{d}^1(\xi^1,\,\xi^2) = \| \,\mathbf{x}^1 - \hat{\mathbf{x}}_t^2(\xi^1,\,\xi^2)\, \| \longrightarrow \text{MIN}\,. \tag{2}$$

The values $(\bar{\xi}^1,\,\bar{\xi}^2)$ are obtained by writing the necessary condition for the minimum of the distance function (2)

$$\frac{d}{d\xi^\alpha}\,\hat{d}^1(\xi^1,\,\xi^2) = \frac{\mathbf{x}^1 - \hat{\mathbf{x}}_t^2(\xi^1,\,\xi^2)}{\| \,\mathbf{x}^1 - \hat{\mathbf{x}}_t^2(\xi^1,\,\xi^2)\, \|}\cdot\hat{\mathbf{x}}_{t,\alpha}^2(\xi^1,\,\xi^2) = 0\,. \tag{3}$$

The solution of (3) requires the orthogonality of the first and second term. Since $\hat{\mathbf{x}}_{t,\alpha}^2(\xi^1,\,\xi^2)$ is the tangent vector $\mathbf{a}_\alpha^2$ the first term must denote the normal $\mathbf{n}^2$. Thus we have the condition $-\mathbf{n}^2\cdot\mathbf{a}_\alpha^2 = 0$ which means that the current master point $\hat{\mathbf{x}}_t^2(\xi^1,\,\xi^2)$ is the orthogonal projection of a given slave point $\mathbf{x}^1$ onto the current master surface $\boldsymbol{\varphi}_t^2(\Gamma_c^2)$.

Here and in the following we will denote by a bar over a quantity its evaluation at the minimal distance point $(\bar{\xi}^1,\,\bar{\xi}^2)$ which means that these values denote the solution point of (3). Thus $\bar{\mathbf{n}}^2 := (\bar{\mathbf{a}}_1^2\times\bar{\mathbf{a}}_2^2)\,/\,\|\bar{\mathbf{a}}_1^2\times\bar{\mathbf{a}}_2^2\|$ is the outward unit normal on the current master surface at the master point where $\bar{\mathbf{a}}_\alpha^2$ are tangent vectors at $\hat{\mathbf{x}}_t^2(\bar{\xi}^1,\,\bar{\xi}^2)$.
The penetration function (1) contains two informations:

1. g_{N+} serves as a local contact check, i.e. we set: *contact* $\Leftrightarrow g_{N+} > 0$
2. g_{N+} enters for $g_{N+} > 0$ as a local kinematical variable the constitutive function for the contact pressure.

   By taking the time derivative of (2) at the minimal distance point $(\bar{\xi}^1,\,\bar{\xi}^2)$ one obtains, in the case of contact, the rate of penetration

$$\dot{g}_{N+} \;=\; [\,\mathbf{v}_t^1 - \hat{\mathbf{v}}_t^2(\bar{\xi}^1,\,\bar{\xi}^2)\,]\cdot\bar{\mathbf{n}}^2 \tag{4}$$

   for given spatial velocities $\mathbf{v}_t^1$ and $\hat{\mathbf{v}}_t^2(\bar{\xi}^1,\,\bar{\xi}^2)$ at the slave and master points.

REMARK I:

1. $g_{N+}^L = [\,\mathbf{x}^1 - \hat{\mathbf{x}}_t^2(\bar{\xi}^1,\,\bar{\xi}^2)\,]\cdot\bar{\mathbf{n}}^2 \geq 0$ represents the classical non–penetration condition for finite deformation.
2. The time derivative of the penetration function (4) can be viewed as the variation of (2) when the velocities are exchanged by the associated variations leading to

$$\delta g_{N+} \;=\; [\,\boldsymbol{\eta}^1 - \hat{\boldsymbol{\eta}}^2(\bar{\xi}^1,\,\bar{\xi}^2)\,]\cdot\bar{\mathbf{n}}^2 \tag{5}$$

   with $\boldsymbol{\eta}$ being the virtual displacement or test function.
3. For the analysis of small deformation problems the kinematical relation (1) or the non–penetration condition from Remark I.1 can be linearized which yields

$$\Delta g_{N+} \;=\; [\,\mathbf{u}^1 - \hat{\mathbf{u}}^2(\bar{\xi}^1,\,\bar{\xi}^2)\,]\cdot\bar{\mathbf{N}}^2 + g_0 \tag{6}$$

   $\mathbf{u}^\gamma$ represents the displacement field which is introduced in the kinematically linear case to connect the current and the reference configuration via: $\mathbf{x}^\gamma = \mathbf{X}^\gamma + \mathbf{u}^\gamma$. The variable g_0

denotes the initial gap between the two bodies which is given by $g_0 = [\mathbf{X}^1 - \hat{\mathbf{X}}^2(\bar{\xi}^1, \bar{\xi}^2)] \cdot \bar{\mathbf{N}}^2$ and the normal $\bar{\mathbf{N}}^2 = (\bar{\mathbf{A}}_1^2 \times \bar{\mathbf{A}}_2^2) \,/\, \| \bar{\mathbf{A}}_1^2 \times \bar{\mathbf{A}}_2^2 \|$ is related to the reference configuration.

2.2 Tangential Relative Velocity and Tangential Relative Slip

The tangential relative slip between two bodies is related to the change of the solution point $(\bar{\xi}^1, \bar{\xi}^2)$ of the minimal distance problem. Thus we can compute the time derivative of ξ^α from (3). This yields the following result

$$\frac{d}{dt} \{ [\mathbf{x}_t^1 - \hat{\mathbf{x}}_t^2(\bar{\xi}^1, \bar{\xi}^2)] \cdot \bar{\mathbf{a}}_\alpha^2 \} = [\mathbf{v}_t^1 - \hat{\mathbf{v}}_t^2(\bar{\xi}^1, \bar{\xi}^2) - \bar{\mathbf{a}}_\beta^2 \, \dot{\bar{\xi}}^\beta] \cdot \bar{\mathbf{a}}_\alpha^2 + [\mathbf{x}_t^1 - \hat{\mathbf{x}}_t^2(\bar{\xi}^1, \bar{\xi}^2)] \cdot \dot{\bar{\mathbf{a}}}_\alpha^2 = 0 \quad (7)$$

with $\dot{\bar{\mathbf{a}}}_\alpha^2 = \hat{\mathbf{v}}_{t,\alpha}^2(\bar{\xi}^1, \bar{\xi}^2) + \hat{\mathbf{x}}_{t,\alpha\beta}^2(\bar{\xi}^1, \bar{\xi}^2) \, \dot{\bar{\xi}}^\beta$ we obtain $\dot{\bar{\xi}}^\beta$ from the following system of equations

$$\bar{H}_{\alpha\beta} \, \dot{\bar{\xi}}^\beta = \bar{R}_\alpha \tag{8}$$

with

$$\begin{aligned} \bar{H}_{\alpha\beta} &= [\bar{a}_{\alpha\beta} + g_{N+} \, \bar{b}_{\alpha\beta}] \,, \\ \bar{R}_\alpha &= [\mathbf{v}_t^1 - \hat{\mathbf{v}}_t^2(\bar{\xi}^1, \bar{\xi}^2)] \cdot \bar{\mathbf{a}}_\alpha^2 + g_{N+} \, \bar{\mathbf{n}}^2 \cdot \hat{\mathbf{v}}_{t,\alpha}^2(\bar{\xi}^1, \bar{\xi}^2) \,. \end{aligned} \tag{9}$$

$\bar{a}_{\alpha\beta}$ and $\bar{b}_{\alpha\beta}$ are the first and second fundamental form of the deformed surface, well known from differential geometry.

Let us now define the tangential relative velocity function on the current slave surface $\boldsymbol{\varphi}_t^1(\Gamma_c^1)$ by setting

$$\mathcal{L}_v \, \mathbf{g}_T \ := \ \dot{\bar{\xi}}^\alpha \, \bar{\mathbf{a}}_\alpha^2 \,. \tag{10}$$

Equation (10) determines per definition the evolution of the tangential slip $\mathbf{g}_T$ which enters as a local kinematical variable the constitutive function for the contact tangential stress, see next section. The rate $\dot{\bar{\xi}}^\alpha$ in (10) at the solution point $(\bar{\xi}^1, \bar{\xi}^2)$ has been already computed in (8).

REMARK II:

1. Note that the last terms in $\bar{H}_{\alpha\beta}$ and $\bar{R}_\alpha$ of (9) depend on the penetration g_{N+}. In the case of a strong enforcement of the non–penetration condition ($g_{N+} = 0$) with Lagrangian multipliers these terms vanish. Then the evolution $\mathcal{L}_v \, \mathbf{g}_T$ in (10) is given by the projection of the spatial velocities $\mathbf{v}_t^1$ and $\hat{\mathbf{v}}_t^2(\bar{\xi})$ evaluated at the slave and master points onto the tangential direction of the master surface at the master point:
$$\mathcal{L}_v \, \mathbf{g}_T \ := \bar{\mathbf{P}}_T \, [\mathbf{v}_t^1 - \hat{\mathbf{v}}_t^2(\bar{\xi}^1, \bar{\xi}^2)] \,, \quad \text{with} \quad \bar{\mathbf{P}}_T = \bar{\mathbf{a}}_\alpha^2 \otimes \bar{\mathbf{a}}^{2\alpha} \,.$$
2. If the deformed contact surface is flat then the curvature tensor $\bar{b}_{\alpha\beta}$ is zero. This is always the case for a surface discretization by three node triangular elements.
3. Note that the (a priori objective) Lie derivative of the tangential vector $\mathbf{g}_T$ has the representation $\mathcal{L}_v \, \mathbf{g}_T = \mathbf{F}_t^2 \{ \frac{d}{dt} [\mathbf{F}_t^{2-1}(\mathbf{g}_T)] \} = \dot{\bar{\xi}}^\alpha \, \bar{\mathbf{a}}_\alpha$ based on the deformation gradient $\mathbf{F}_t^2$ of the master surface defined above. Thus (10) represents an evolution equation for the objective rate $\mathcal{L}_v \, \mathbf{g}_T$ of the tangential vector introduced above.
4. In case of no relative movement in tangential direction (stick condition) we have $\mathcal{L}_v \, \mathbf{g}_T = \mathbf{g}_T = \mathbf{0}$.

5. In the geometrically linear case we obtain from (7) and (6)

$$\frac{d}{dt}\left\{\,[\,\mathbf{x}^1_t - \hat{\mathbf{x}}^2_t(\bar{\xi}^1,\,\bar{\xi}^2)\,]\cdot\bar{\mathbf{A}}^2_\alpha\,\right\} = [\,\mathbf{v}^1_t - \hat{\mathbf{v}}^2_t(\bar{\xi}^1,\,\bar{\xi}^2) - \bar{\mathbf{A}}^2_\beta\,\dot{\bar{\xi}}^\beta\,]\cdot\bar{\mathbf{A}}^2_\alpha$$

which yields

$$\bar{A}^2_{\alpha\beta}\,\dot{\bar{\xi}}^\beta = [\,\mathbf{v}^1_t - \hat{\mathbf{v}}^2_t(\bar{\xi}^1,\,\bar{\xi}^2)\,]\cdot\bar{\mathbf{A}}^2_\alpha$$

The terms multiplied by g_{N+} can be neglected. Thus $\dot{\bar{\xi}}^\beta$ is given by the projection of the difference velocity of the two bodies at the contact point on the tangent direction of the undeformed surface. From the last equation we can deduce the the relative tangential velocity at the contact point: $\dot{\mathbf{g}}_T = \dot{\bar{\xi}}^\beta\,\bar{\mathbf{A}}^2_\alpha$.

3. CONSTITUTIVE EQUATIONS FOR CONTACT INTERFACES

As discussed in the introductory remarks, the normal contact stresses can be obtained in two generally different ways. On one hand the contact stresses follow from the constraint equations. On the other hand an approach of both bodies is observed in the contact area which then leads to the formulation of associated constitutive interface equations.

3.1 Normal Stress in the Contact Area

In the first case the mathematical condition for non–penetration is stated in remark I.1 as $g^L_{N+} \geq 0$ which precludes the penetration of one body into another. Then contact takes place when g^L_{N+} is equal to zero. In this case the associated normal components p_N of the stress vector $\mathbf{t} = p_N\,\bar{\mathbf{n}} + t^\beta\,\bar{\mathbf{a}}_\beta$ in the contact interface must be non–zero. The stress vector acts on both surfaces, obeying the action–reaction principle: $\mathbf{t}^2(\bar{\xi}^1,\,\bar{\xi}^2) = -\mathbf{t}^1$ in the contact point $\mathbf{x}^1$. We have $p_N = p^1_N = p^2_N < 0$ since adhesive stresses will not be allowed in the contact interface. This leads to the statement

$$g^L_{N+} \geq 0\,, \qquad p_N \leq 0\,, \qquad p_N\, g^L_{N+} = 0 \tag{11}$$

which is well known as the Kuhn–Tucker condition for frictionless contact problems. These conditions provide the basis to treat contact problem in the context of constraint optimization. For further details see the next section.

When the micromechanical behaviour of the contact area is studied different phenomena have to be considered for the mechanical interface description. Here we restrict ourselves to constitutive models which have been derived based on micromechanical observations of physical contact surfaces. These models can be related to formulations relative to mathematical contact surfaces by an averaging process as symbolically indicated in Figure 2. Goal of this section is to formulate local constitutive equations for the pressure and the tangential stress on the slave surface at point $\mathbf{x}^1$ relative to the bases $\{\bar{\mathbf{a}}^2_\alpha, \bar{\mathbf{n}}^2\}$ acting on body $\mathcal{B}^1$.

It is well known that the contact pressure is related to the approach of the physical surfaces which come into contact, i.e. the penetration of the mathematical surfaces results from the deformation of the micro–asperities, see Figure 3. Let us assume the following general form of the constitutive law

$$p_N = f\,(d) \quad \text{or} \quad d = h\,(p_N) \tag{12}$$

where f and h are nonlinear functions of the current mean plane distance d or the contact pressure p_N, respectively.

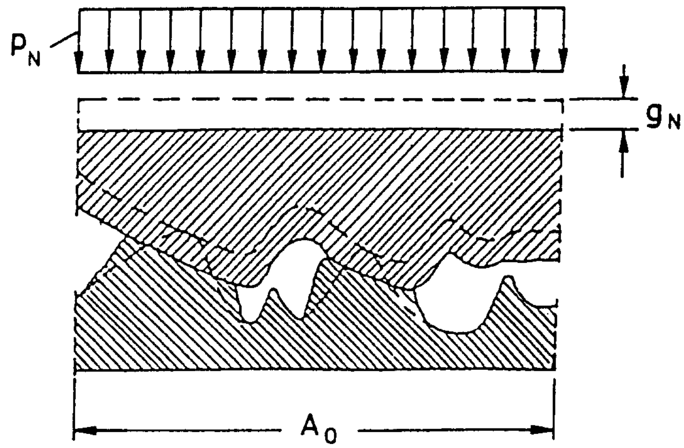


Figure 2. Averaging of micromechanical contact relations

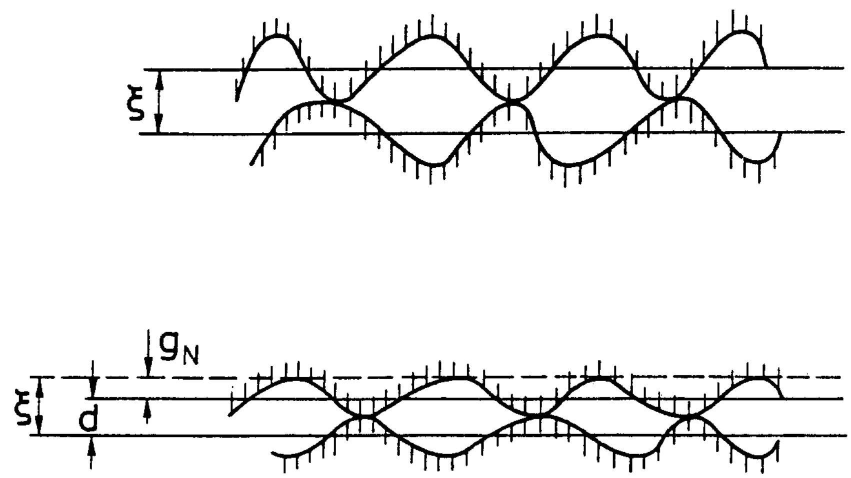


Figure 3. Physical approach in Γ_c

Most of the interface laws can be written in the form (12). Out of many different possibilities two constitutive equations for normal pressure in the contact area will be stated. The first was developed in Zavarise (1991), Zavarise, Schrefler, Wriggers (1992), and is based on a statistical model of the microgeometry proposed by Cooper, Mikic, Yovanovich (1969), recently revisited in Song, Yovanovich (1987).

$$p_N = \frac{c_1 \left(1617646.152 \, \frac{\sigma}{m}\right)^{c_2}}{5.589^{1+0.0711 \, c_2}} \exp\left[-\frac{1+0.0711 \, c_2}{(1.363\sigma)^2} \, d^2\right] . \tag{13}$$

Here c_1 and c_2 are mechanical constants expressing the nonlinear distribution of the surface hardness, σ and m are statistical parameters of the surface profile, representing the RMS surface roughness and the mean absolute asperity slope. Thus we have an exponential law of the form $p_N = c_3 \, e^{-c_4 \, d^2}$.

In case of contact the current mean plane distance is related to the geometrical approach g_{N+} (1) as follows

$$g_{N+} = \zeta - d \tag{14}$$

where ζ is the initial mean plane distance in the contact area Γ_c.

Another law for the contact pressure has been given, based on experimental investigations, by Kragelsky, Dobychin, Kombalov (1982). These authors formulated the following nonlinear elastic constitutive equation for the contact pressure

$$p_N = c_N \, (\, g_{N+} \,)^n \tag{15}$$

in terms of the penetration g_{N+} defined in (1). Here c_N and n are material parameters which have to be determined by experiments.

3.2 Tangential Contact Stress and Tangential Frictional Slip

Many different constitutive models have been developed to formulate the interfacial behaviour due to friction. In this overview we restrict our consideration to two models for frictional behaviour; one being the classical Coulomb model. The response in tangential direction can be divided in two different actions. In the first no tangential relative displacement of the two bodies occurs which is the so–called stick condition. The second action is associated with a relative tangential movement in the contact interface which denotes the so–called slip.

The stick condition can be formulated with remark II.4 simply as

$$\mathbf{g}_T^L = \mathbf{0} \tag{16}$$

which imposes in general a nonlinear constraint equation on the motion in the contact interface. Associated with this constraint is a Lagrangian multiplier, $\boldsymbol{\lambda}_T$, which denotes the reaction due to (16).

In case of sliding the law of Coulomb yields

$$\mathbf{t}_T = -\mu \, | \, p_N \, | \, \frac{\mathcal{L}_v \, \mathbf{g}_T}{\| \, \mathcal{L}_v \, \mathbf{g}_T \, \|} \tag{17}$$

where μ is the sliding frictional coefficient which depends on the surface roughness and may also depend on the sliding velocity $\mathcal{L}_v \, \mathbf{g}_T$, the pressure p_N or the temperature.

These and other constitutive equations for friction can be formulated in the framework of elastoplasticity. This has been investigated by several authors who also developed different constitutive equations for frictional problems, see e.g. Michalowski, Mroz (1978) or Curnier (1984). A treatment of frictional interface laws in terms of non–associated plasticity has been considered within a finite element formulation by Wriggers (1987), Giannokopoulos (1989), Wriggers, Vu Van, Stein (1990) or Laursen, Simo (1993).

Classically one has to distinguish between stick which means no relative tangential movement in the contact interface and slip which is associated with relative tangential movement. The key idea of the elasto–plastic approach is a split of the tangential slip $\mathbf{g}_T$ into an elastic part $\mathbf{g}_T^e$ and a plastic (slip) part $\mathbf{g}_T^s$, see equation $(18)_2$ below. The elastic part describes the micro displacement which can be regarded as stick behaviour since the associated deformations vanish once the loading is removed from the system. The constitutive behaviour for the tangential elastic micro–displacements can be deduced from experiments and is related to the elastic deformation of the asperities due to tangential loading. Here we assume, as the simplest possible model, an isotropic linear elastic constitutive equation for the tangential contact stress

$$\mathbf{t}_T = c_T \, \mathbf{g}_T^e \qquad \text{with} \qquad \mathbf{g}_T^e := \mathbf{g}_T - \mathbf{g}_T^s \tag{18}$$

where c_T is a material parameter.

The tangential plastic slip $\mathbf{g}_T^s$ is governed by a constitutive evolution equation which can be derived by using standard concepts of the theory of elastoplasticity. Within this framework we can formulate a plastic slip criterion function for a given contact pressure p_N with material parameter μ which determines the frictional sliding, see Figure 4a,

$$\hat{f}_s(\mathbf{t}_T) = \| \mathbf{t}_T \| - \mu \, p_N \leq 0 \,. \tag{19}$$

Another slip criterion function has been formulated in Wriggers, Vu Van, Stein (1990) which additionally takes into account the pressure dependency of the tangential response. Here the form $\mu = \tau_0 \,/\, p_r + \beta$ proposed by Tabor (1981) for most solids is assumed, where τ_0 and β are constitutive parameters and describe a model with linear varying shear strength of the interfacial material due to the true contact pressure. The true pressure p_r is related to the true contact area A_r (real contact area due to the contact of the asperities in the contact interface) whereas the pressure p_N is associated with the nominal contact area A. Woo, Thomas (1980) have formulated a relation between the true and the nominal area based on experimental observations

$$\frac{A_r}{A} = \left(\frac{|\, p_N \,|}{A \, H} \right)^n , \qquad n = \frac{5}{6} \,. \tag{20}$$

with the hardness H of the material. With these relations Wriggers, Vu Van, Stein (1990) arrived at the following slip criterion, see Figure 4b,

$$\hat{f}_s(\mathbf{t}_T, \, p_N) = \| \, \mathbf{t}_T \, \| - \alpha \ \ |\, p_N \,|^n - \beta \ \ |\, p_N \,| \leq 0 \,, \qquad \alpha = \frac{A \, \tau_0}{(A \, H)^n} \,. \tag{21}$$

Note that the choice of one of the slip criteria (19) or (21) has to be made with regard to experimental data within the contact interface; there are of course other slip criteria possible.

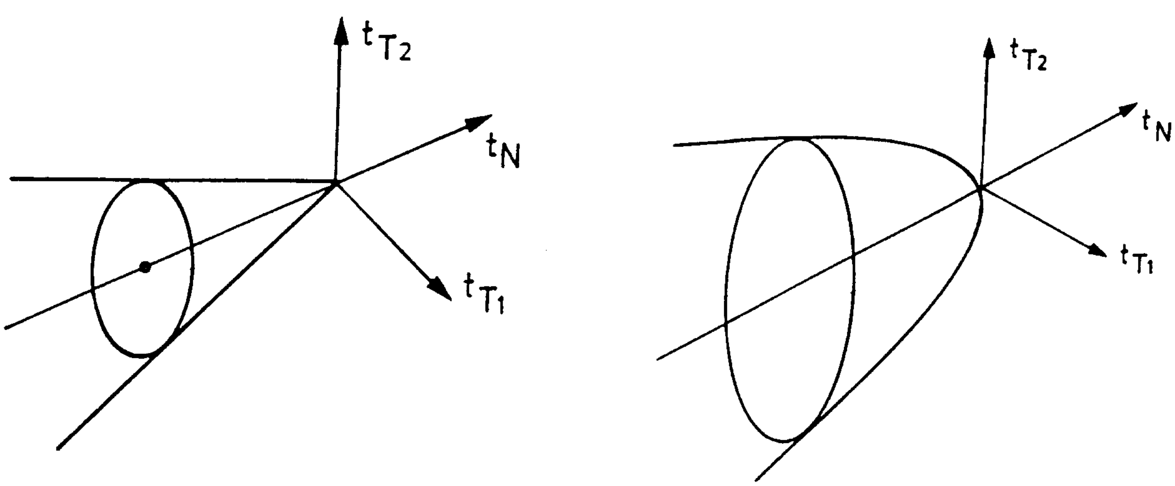


Figure 4. a) Coulomb frictional cone; b) Parabolic slip surface

The constitutive evolution equation for the plastic or frictional slip can be stated in form of a slip rule for large deformations in the contact zone as follows

$$\mathcal{L}_v \, \mathbf{g}_T^s = \lambda \, \frac{\partial \hat{f}_s(\mathbf{t}_T)}{\partial \mathbf{t}_T} = \lambda \, \mathbf{n}_T \,, \qquad \text{with} \quad \mathbf{n}_T = \frac{\mathbf{t}_T}{\|\mathbf{t}_T\|} \tag{22}$$

which denotes a normality rule for a fixed contact pressure p_N. Here λ is the plastic parameter which describes the magnitude of the plastic slip. Equations (18), (19) or (20) and (22), along with the loading–unloading conditions in Kuhn–Tucker form

$$\lambda \geq 0\,, \quad \hat{f}_s(\mathbf{t}_T) \leq 0\,, \quad \lambda\,\hat{f}_s(\mathbf{t}_T) = 0\,, \tag{23}$$

establish the constitutive framework for the tangential slip–stick behaviour. The algorithmic treatment will be discussed in section 6.

4. BOUNDARY VALUE PROBLEM, GLOBAL SOLUTION STRATEGIES

For the formulation of the boundary value problem we have to discuss only the additional terms due to contact in detail. The equations describing the behaviour of the bodies coming into contact do not change. However, for completeness, the balance equations and a simple constitutive model are stated for elastic solids undergoing finite deformation.

4.1 Local Balance Equations for the Solid

We can formulate the local momentum equation for a body $\mathcal{B}^\gamma$ as

$$\text{DIV}\ \mathbf{P}^\gamma + \bar{\mathbf{f}}^\gamma = \mathbf{0} \tag{24}$$

in case that inertia terms are neglected. $\mathbf{P}^\gamma$ denotes the first Piola–Kirchhoff stress tensor acting in the body γ, $\bar{\mathbf{f}}^\gamma$ are the body forces. Next we formulate the boundary conditions for the deformation and the stress field

$$\begin{aligned} \boldsymbol{\varphi}^\gamma &= \bar{\boldsymbol{\varphi}}^\gamma \qquad \text{on} \quad \Gamma_\varphi^\gamma\,, \\ \mathbf{t}^\gamma &= \bar{\mathbf{t}}^\gamma \qquad \text{on} \quad \Gamma_\sigma^\gamma \end{aligned} \tag{25}$$

where $\bar{\boldsymbol{\varphi}}^\gamma$ and $\bar{\mathbf{t}}^\gamma$ are described quantities. Furthermore we have to account for the contact condition which is given by equation $(12)_2$ with the definition of the gap function (1) when an approach of the bodies in the contact interface is allowed or by the condition defined in remark I.1 which yields the inequality

$$g_{N+}^L \geq 0 \qquad \text{on} \quad \Gamma_c\,. \tag{26}$$

4.2 Constitutive Relations

As a model for non–linear constitutive equations we use a form valid for finite elasticity which leads to a non–linear relation between the Kirchhoff stress $\boldsymbol{\tau}$ and the left Cauchy Green tensor $\mathbf{b} = \mathbf{F}\,\mathbf{F}^T$: $\boldsymbol{\tau} = \mathbf{f}\,(\,\mathbf{b}\,)$. The Kirchhoff stress is related to the first Piola–Kirchhoff stress via $\boldsymbol{\tau} = \mathbf{P}\,\mathbf{F}^T$, with $\mathbf{F}$ being the deformation gradient. The simplest constitutive equation for hyperelasticity is known as the Neo–Hookian model and can e. g. be applied for rubber materials undergoing moderately large strains, see e.g. Ogden (1984). It is stated below for the body $\mathcal{B}^\gamma$ with the Jacobian of the deformation $J^\gamma = \det \mathbf{F}^\gamma$

$$\boldsymbol{\tau}^\gamma = \Lambda^\gamma\,(\,J^\gamma - 1\,)\,\mathbf{1} + \mu^\gamma\,(\,\mathbf{b}^\gamma - \mathbf{1}\,)\,. \tag{27}$$

Material parameters for the body $\mathcal{B}^\gamma$ are the Lamè constants Λ^γ and μ^γ. The material model is valid for finite elastic deformations. Of course we can consider more complicated constitutive relations which can also be of inelastic nature. It should be noted that since the contact has to be formulated only within the interface, the constitutive laws for the bodies coming into contact can be arbitrary and do not affect the main algorithmic treatment of the contact problem. However it is clear that the physical properties of the surfaces of the bodies are influenced by the general constitutive behaviour.

4.3 Weak Formulation

For a numerical solution of the nonlinear boundary value problem summarized above we will use the finite element method. Thus we need the weak form of equations (24) to (27). Due to the fact that the constraint condition (24) is represented by an inequality we obtain in general a variational inequality. The general form can be written as

$$\sum_{\gamma=1}^{2} \int_{\Omega^\gamma} \boldsymbol{\tau}^\gamma \cdot \operatorname{grad}(\boldsymbol{\eta}^\gamma - \boldsymbol{\varphi}^\gamma)\, dV \geq \sum_{\gamma=1}^{2} \int_{\Omega^\gamma} \bar{\mathbf{f}}^\gamma \cdot (\boldsymbol{\eta}^\gamma - \boldsymbol{\varphi}^\gamma)\, dV - \int_{\Gamma_\sigma{}^\gamma} \bar{\mathbf{t}}^\gamma \cdot (\boldsymbol{\eta}^\gamma - \boldsymbol{\varphi}^\gamma)\, dA \tag{28}$$

where the integration is performed with respect to the domain Ω^γ occupied by the body $\mathcal{B}^\gamma$ in the reference configuration. The stress tensor and the gradient operator "grad" are evaluated with respect to the current coordinates.

We now have to find the deformation $(\boldsymbol{\varphi}^1, \boldsymbol{\varphi}^2) \in \mathbf{K}$ such that (26a) is fulfilled for all $(\boldsymbol{\eta}^1, \boldsymbol{\eta}^2) \in \mathbf{K}$ with

$$\mathbf{K} = \{ (\boldsymbol{\eta}^1, \boldsymbol{\eta}^2) \in \mathbf{V} \,|\, [\boldsymbol{\eta}^1 - \hat{\boldsymbol{\eta}}^2(\bar{\xi}^1, \bar{\xi}^2)] \cdot \bar{\mathbf{n}}^2 \geq 0 \},$$

see also section 2. In case of finite elasticity the existence of the solution of (28) can be proved, see e.g. Ciarlet (1988) or Curnier, He, Telega (1992). For this, the strain energy function has to be polyconvex and the solution lies in the usual Sobolev space $W^{1,p}$. The space $\mathbf{V}$ is defined as $\mathbf{V} = \{ \boldsymbol{\eta}^\gamma \in [W^{1,p}(\Omega^\gamma)]^{dim} \,|\, \boldsymbol{\eta}^\gamma = \mathbf{0} \quad \text{on } \Gamma_u \}$, dim denotes the dimension of the problem at hand.

REMARK III:

1. In the geometrically linear case these equations can be recast in a weak or variational formulation as follows

$$a(\mathbf{u}, \mathbf{v} - \mathbf{u}) \geq f(\mathbf{v} - \mathbf{u}), \tag{29}$$

   with

$$a(\mathbf{u}, \mathbf{w}) = \int_\Omega \boldsymbol{\varepsilon}(\mathbf{u}) : \boldsymbol{\mathcal{C}}_0 : \boldsymbol{\varepsilon}(\mathbf{w})\, d\Omega,$$

$$f(\mathbf{w}) = \int_\Omega \hat{\mathbf{b}} \cdot \mathbf{w}\, d\Omega + \int_{\Gamma_\sigma} \hat{\mathbf{t}} \cdot \mathbf{w}\, d\Gamma$$

   and $\Omega = \cup_\gamma \mathcal{B}^\gamma$. $\boldsymbol{\mathcal{C}}_0$ is the elasticity matrix due to the classical constitutive law of Hooke. The linear strain tensor is defined by $\boldsymbol{\varepsilon}(\mathbf{u}) = \frac{1}{2}(\nabla \mathbf{u} + \nabla^T \mathbf{u})$.
   The problem is now, as in the nonlinear case, to find $\mathbf{u} \in \mathbf{K}$ such that (29) is fulfilled for all $\mathbf{v} \in \mathbf{K}$ with, see equation (6),

$$\mathbf{K} = \{ \mathbf{v} \in \mathbf{V} \,|\, (\mathbf{v}^1 - \bar{\mathbf{v}}^2) \cdot \bar{\mathbf{n}}^2 + g_0 \geq 0 \text{ on } \Gamma_c \}$$

   and

$$\mathbf{V} = \{ \mathbf{v} \in [H^1(\Omega)]^{dim} \,|\, \mathbf{v} = \mathbf{0} \text{ on } \Gamma_u \}.$$

   The mathematical structure of the variational inequality (29) is discused in detail in e. g. Duvaut, Lions (1976) or Kikuchi, Oden (1988). Due to the inequality constraint on the deformation field the contact problem is nonlinear even for the linear elastic case (29).

Algorithms for solving variational inequalities are given by mathematical programming, active set strategies or sequential quadratic programming methods, to name only a few. Each of these methods which are well known from optimization theory, see e. g. Luenberger (1984), has been applied to contact problems. For mathematical programming see e. g. Conry, Seireg (1971) or Klarbring (1986). The sequential quadratic programming approach has been considered by e.g. Barthold, Bischoff (1988) and lately with application to large strain elasticity by Björkman, Klarbring, Sjödin, Larsson, Rönnqvist (1995).

Here we will investigate in more detail the active set strategies which are applied in many existing finite element codes. Within this method the contact constraints can be introduced via Lagrangian multipliers or penalty terms. Furthermore we will observe that the introduction of a constitutive equation for the normal approach in the contact interface, see section 3.1, yields a formulation which is like a nonlinear penalty method.

Within an active set strategy we can write the weak form as an equality since we know the active set within an incremetal solution step. Then equations (24) to (27) yield

$$\begin{aligned}\sum_{\gamma=1}^{2}\{ \int_{\Omega^\gamma} \boldsymbol{\tau}^\gamma \cdot \operatorname{grad} \boldsymbol{\eta}^\gamma \, dV - \int_{\Omega^\gamma} \bar{\mathbf{f}}^\gamma \cdot \boldsymbol{\eta}^\gamma \, dV - \int_{\Gamma_\sigma{}^\gamma} \bar{\mathbf{t}}^\gamma \cdot \boldsymbol{\eta}^\gamma \, dA \} \\ + \, ''Contact \;\; Contributions'' = 0\end{aligned} \tag{30}$$

Note that the integration is performed with regard to the reference configuration but the stress tensor and the gradients are evaluated with respect to current configuration. $\boldsymbol{\eta}^\gamma \in V$ is the so called test function or virtual displacement which is zero at the boundary Γ_φ^γ where the deformations are prescribed.

For two bodies being in contact we obtain the weak form of the interface by assuming that contact is active at the surface Γ_c. Then the formulation follows for the three different cases as given below.

1. **Lagrangian multiplier method:**

$$\int_{\Gamma_c} (\lambda_N \, \delta g_{N+}^L + \boldsymbol{\lambda}_T \cdot \delta \mathbf{g}_T \,) \, dA \tag{31}$$

Here λ_N denotes the Lagrangian multiplier which can be identified as the contact pressure p_N. δg_{N+}^L is the variation of the normal gap, defined in remark I.1, which yields the same result as given in equation (5). The term $\boldsymbol{\lambda}_T \cdot \delta \mathbf{g}_T$ is associated with the tangential stick or slip motion and needs further discussion. In case of pure stick the relative tangential slip $\mathbf{g}_T$ is zero which yields a constraint equation from which $\boldsymbol{\lambda}_T$ follows as a reaction. In case of sliding the tangential stress vector $\mathbf{t}_T$ is determined by the constitutive law for frictional slip, see section 3.2 and thus we should write instead of $\boldsymbol{\lambda}_T \cdot \delta \mathbf{g}_T \longrightarrow \mathbf{t}_T \cdot \delta \mathbf{g}_T$.

2. **Penalty method:**
In this formulation a penalty term due to the constraint condition is added to the weak form (26). This means that once the constraint equation for g_{N+}^L is violated

$$\int_{\Gamma_c} \epsilon_N \, g_{N+}^L \, \delta g_{N+}^L \, dA \,, \quad \epsilon_N > 0 \tag{32}$$

has to be considered for normal contact. It can be shown, see e.g. Luenberger (1984), that the solution of the Lagrangian multiplier method can be recovered from this formulation for $\epsilon_N \to \infty$, however this will lead to an ill–conditioned problem, see next section. As in the Lagrangian multiplier method we have to distinguish between pure stick in the contact interface which produces a penalty term also for the tangential direction

$$\int_{\Gamma_c} (\epsilon_N \, g_{N+}^L \, \delta g_{N+}^L + \epsilon_T \, \mathbf{g}_T \cdot \delta \mathbf{g}_T \,) \, dA \,, \quad \epsilon_N > 0 \;, \epsilon_T > 0 \tag{33}$$

and the slip condition which leads to

$$\int_{\Gamma_c} (\epsilon_N \, g_{N+}^L \, \delta g_{N+}^L + \mathbf{t}_T \cdot \delta \mathbf{g}_T \,) \, dA \,, \quad \epsilon > 0 \tag{34}$$

In the latter equation one of the frictional laws from section 3.2 has to be applied.

3. **Constitutive equation in the interface:**

$$\int_{\Gamma_c} (\, p_N \, \delta g_N + \mathbf{t}_T \cdot \delta \mathbf{g}_T \,) \, dA \tag{35}$$

In this case the constitutive equation which have been discussed in section 3.1 and 3.2 have to applied for the determination of p_N and $\mathbf{t}_T$. One can easily see, that the introduction of the constitutive equation for the normal pressure (15) yields a nonlinear penalty functional for the normal contact. The standard penalty method can be recovered from this relation by using $n = 1$. However this choice is somehow artificial since the usual range of the constitutive paramter n, stemming from experiments, is in the range $2 \leq n \leq 3.33$.

In equations (31) to (35) the variation of the normal gap function g_{N+} is needed which yields, see equation (5):

$$\delta g_{N+} = [\, \boldsymbol{\eta}^1 - \boldsymbol{\eta}^2(\bar{\xi}_1, \bar{\xi}_2)] \cdot \bar{\mathbf{n}}^2 \,. \tag{36}$$

Furthermore the variation of the tangential slip can be stated as

$$\delta \mathbf{g}_T = \delta \bar{\xi}^\alpha \, \bar{\mathbf{a}}_\alpha^2 \,. \tag{37}$$

The latter relation follows simply from (10) by replacing the velocities $\mathbf{v}$ by the test function $\boldsymbol{\eta}$ in (9).

REMARK IV:

1. When in the constitutive law for the elastic micro displacements (18) the constitutive parameter c_T is exchanged by the penalty parameter ϵ_T then we can think of the resulting equations as a penalty regularization of the frictional interface law (17), see also Ju, Taylor (1988) or Curnier, Alart (1989).
2. A further possibility to incorporate constraint equations is provided by a direct elimination of the variables in the contact interface. In this case we can write on Γ_c, see section 2.1, $g_{N+} = 0 \longrightarrow \mathbf{x}^1 \cdot \bar{\mathbf{n}}^2 = \hat{\mathbf{x}}_t^2 \cdot \bar{\mathbf{n}}^2$ and thus eliminate in Γ_c either the displacements related to $\mathcal{B}^1$ or to $\mathcal{B}^2$. Since this method is associated in practical applications with a constant change of the number of unknowns in the global system of equations, it is not so attractive.

3. Another technique for problems with inequality constraints is the so called barrier method. It adds a constraint functional of the type $\int_{\Gamma_c} \epsilon_N \, / \, g_{N+}^2 \, \delta g_{N+} \, d\Gamma$ to equation (30) which is always active for all possible contact nodes. However, due to the construction of the constraint functional the solution has always to stay in the feasable region which means that no penetration is allowed in any intermediate iteration step. To ensure this special safe guard algorithms are needed, see e.g. Bazaraa, Sherali, Shetty (1993).
4. A technique based on a new constraint functional which includes the penalty and the barrier formulation as limit cases has been developed lately and named methods of cross constraints, see Zavarise, Wriggers, Schrefler (1995b). Due to its construction the functional is also active when the gap function is open as in the barrier method, however a safe guard algorithm has not to be applied since the solution is not restricted to the feasable region.
5. Perturbed Lagrangian formulations can be used to combine both penalty and Lagrangian multiplier methods in a mixed formulation, see e.g. Oden (1981) or Simo, Wriggers, Taylor (1985). In this case the following functional

$$\Pi_p = \Pi + \int_{\Gamma_c} [\, \lambda_N \, g_{N+} \, - \frac{1}{2\,\epsilon_N} \, \lambda_N^2 \,] \, d\Gamma \longrightarrow STAT$$

is defined where Π denotes the total energy of the two bodies. The Lagrangian multiplier term is regularized by the second term in the integral which can be view as the complementary energy due to the Lagrangian multiplier. The variation leads to

$$\delta\Pi_p = \delta\Pi + \int_{\Gamma_c} [\, \lambda_N \, \delta g_{N+} \, + \delta\lambda_N \, (\, g_{N+} - \frac{1}{\epsilon_N} \, \lambda_N \,) \,] \, d\Gamma = 0 \tag{38}$$

The first term is again associated with the Lagrangian multiplier formulation (31) whereas the second term yields the "constitutive law": $\lambda_N = \epsilon_N \, g_{N+}$ if evaluated locally. If we insert this result for λ_N in the first term of (38) we obtain the standard penalty formulation (32). However, equation (38) can also be a starting point for special mixed formulations, see section 5.4.

4.4 Augmented Lagrangian Formulation

A major problem associated with the numerical treatment of the penalty method and the contact interface laws is ill–conditioning which arises when the penalty parameter ϵ_N or the stiffnesses due to laws (13) or (15) are combined with stiffnesses of the bodies within the finite element formulation. One way to overcome the problem of ill–conditioning is the use of very high precision arithmetic throughout the computation, see Zavarise, Schrefler, Wriggers (1992). This approach is motivated by the fact that there exist estimations for the magnitude of the penalty parameter ϵ_N, see Nour–Omid, Wriggers (1987). These estimates lead the penalty parameter which avoids ill–conditioning as follows

$$\epsilon_N = \frac{k}{\sqrt{N\,t}} \tag{39}$$

k is a characteristic stiffness parameter of the adjoint elements (e. g. the modulus of compression), N is the total number of unknowns and t denotes the computer precision. The penalty parameter is directly limited by the latter quantity. Since however the precision which is needed, either to fulfil the constraint equation $(11)_1$ or to evaluate a technical relevant interface law, is higher than the precision standardly used for finite element computations this approach is not very advantageous.

Another method to overcome the problem of ill–conditioning is based on the augmented Lagrangian technique, well known in optimization theory. This technique has been considered extensively within the context of incompressibility constraints in e. g. Glowniski, Le Tallec (1984) and was also applied to contact problems, see Wriggers, Simo, Taylor (1985) or Kikuchi, Oden (1988) for frictionless contact. Recently this approach has been extended successfully also to large displacement contact problems including friction, see Alart, Curnier (1991) or Laursen, Simo (1991). A formulation which accounts for micromechanical interface laws can be found in Wriggers, Zavarise (1993).

The main idea is to combine either the penalty method or the constitutive interface laws with Lagrangian multiplier methods. This works in the way that in augmented Lagrangian techniques a Lagrangian multiplier $\bar{\lambda}_N$ is introduced and held constant during an iteration loop to solve (30) which is nonlinear with respect to the deformation $\boldsymbol{\varphi}^\gamma$.

To present the main idea, we apply this technique here only for the normal direction and combine it with the penalty method leading to the weak form

$$\begin{aligned}\sum_{\gamma=1}^{2}\{ \int_{\mathcal{B}^\gamma} \boldsymbol{\tau}^\gamma \cdot \operatorname{grad} \boldsymbol{\eta}^\gamma \, dV - \int_{\mathcal{B}^\gamma} \bar{\mathbf{f}}^\gamma \cdot \boldsymbol{\eta}^\gamma \, dV - \int_{\Gamma_\sigma{}^\gamma} \bar{\mathbf{t}}^\gamma \cdot \boldsymbol{\eta}^\gamma \, dA \} \\ + \int_{\Gamma_c} [\, \bar{\lambda}_N + \epsilon_N \, g^L_{N+} \,) \, \delta g_{N+} + \mathbf{t}_T \cdot \delta \mathbf{g}_T \,] \, dA = 0\end{aligned} \tag{40}$$

Since $\bar{\lambda}_N$ is unknown an update procedure for the Lagrangian multiplier has to be constructed within an iteration loop. The simplest update is: $\bar{\lambda}_{N_{new}} = \bar{\lambda}_{N_{old}} + \epsilon_N \, g^L_{N+_{new}}$ which is only of first order accuracy. For other possibilities, see e.g. Bertsekas (1983) or in the context of finite element contact problems Alart, Curnier (1991).

In case that the augmented Lagrangian technique is employed for constitutive equations in the contact interface Wriggers, Zavarise (1993) developed the following weak form for the augmented Lagrangian method

$$\begin{aligned}\sum_{\gamma=1}^{2}\{ \int_{\mathcal{B}^\gamma} \boldsymbol{\tau}^\gamma \cdot \operatorname{grad} \boldsymbol{\eta}^\gamma \, dV - \int_{\mathcal{B}^\gamma} \bar{\mathbf{f}}^\gamma \cdot \boldsymbol{\eta}^\gamma \, dV - \int_{\Gamma_\sigma{}^\gamma} \bar{\mathbf{t}}^\gamma \cdot \boldsymbol{\eta}^\gamma \, dA \} \\ + \int_{\Gamma_c} \{ \, [\, \bar{p}_N + \epsilon_N \, c_+(\varphi^\gamma, \bar{p}_N) \,] \, \delta g_{N_+} + \mathbf{t}_T \cdot \delta \mathbf{g}_T \, \} \, dA = 0\end{aligned} \tag{41}$$

$$\text{subject to} \quad c_+ \, (\boldsymbol{\varphi}^\gamma, \, p_N) = g_{N_+} - [\, \zeta - d \, (p_N) \,] = 0 \, .$$

This equation is nonlinear in the contact pressure $\bar{p}_N$ but since this quantity is fixed we do not have to consider this dependancy. Note that we use here a linear penalty law even in the presence of a nonlinear relation (13) for the approach. Thus the fulfillment of the nonlinear interface law (14) will be practically accounted for by the update of the Lagrangian multiplier $\bar{p}_N$

$$\bar{p}_{N_{new}} = \bar{p}_{N_{old}} + \epsilon_N \, c_+ \, (\boldsymbol{\varphi}^\gamma_{new}, \bar{p}_{N_{old}}) \tag{42}$$

with the known quantities $\{..\}_{old}$ from the previous state. Due to the appearance of the nonlinear function c_+ the update is related, but different, to the standard update procedure for the Lagrangian multipliers, see equation (40).

Augmented Lagrangian techniques have also be applied to frictional problems, see Alart, Curnier (1991), Laursen, Simo (1993b) or Zavarise, Wriggers, Schrefler (1995). The last two papers investigate also the possibility of an algorithmic symmetrization of the frictional part of the tangent matrix, see also next section.

5. DISCRETIZATION TECHNIQUES WITHIN THE CONTACT AREA

The discretization of the domain contributions of the bodies being in contact in (30) is not objective of this work. Within this context we refer to the finite element implementations of boundary-value-problems regarding finite elasticity, see e.g. Wriggers (1993) and references therein. This leads to the following matrix formulation for the weak form (30)

$$\mathbf{G}(\mathbf{v}) = \sum_{\gamma=1}^{2} \{ \int_{\mathcal{B}^\gamma} \mathbf{B}^T \boldsymbol{\tau}^\gamma \, dV - \int_{\mathcal{B}^\gamma} \mathbf{N}^T \bar{\mathbf{f}}^\gamma \, dV - \int_{{\Gamma_\sigma}^\gamma} \mathbf{N}^T \bar{\mathbf{t}}^\gamma \, dA \} \tag{43}$$

where the matrix $\mathbf{N}$ contains the shape functions and the so–called $\mathbf{B}$–matrix contains the derivatives of the shape functions. Any standard finite element book can be used, for details, see e.g. Zienkiewicz, Taylor (1988).

Here we focus on the contact constraints. For reasons of simplicity we will restrict ourselves here to two dimensional formulations. Three dimensional contact discretizations can be found in e.g. Hallquist, Goudreau, Benson (1985), Laursen, Simo (1993) or Heegaard, Curnier (1993). We like now to discuss different possibilities to discretize the contact contributions (31) to (35) and the variations in normal and tangential direction (36) and (37).

The basic difference between the Lagrangian method (31) and the penalty approach (32) lies in the fact that the Lagrangian multiplier formulation is a mixed method which means that both variables λ_N and δg_N have to be discretized

$$\int_{\Gamma_c} \lambda_N \, \delta g_N \, d\Gamma \longrightarrow \int_{\Gamma_c^h} \lambda_N^h \, \delta g_N^h \, d\Gamma \tag{44}$$

with the interpolations for λ_N^h and δg_N^h

$$\lambda_N^h = \sum_K M_K(\xi) \, \lambda_{NK} \quad \text{and} \quad \delta g_N^h = \sum_I N_I(\xi) \, \delta g_{NI}$$

Note that the interpolations have to be chosen in such a way that they fulfil the LBB condition for this mixed formulation, see e.g. Kikuchi, Oden (1988). Contrary, the penalty method needs only the discretization of the displacement variables

$$\int_{\Gamma_c} \epsilon_N \, g_N \, \delta g_N \, d\Gamma \longrightarrow \int_{\Gamma_c^h} \epsilon_N \, g_N^h \, \delta g_N^h \, d\Gamma \tag{45}$$

with the interpolation

$$g_N^h = \sum_I N_I(\xi) \, g_{NI} \quad \text{and} \quad \delta g_N^h = \sum_I N_I(\xi) \, \delta g_{NI}$$

In the following we will only discuss discretizations related to the penalty method and to the formulation using constitutive equations in the contact interface where the contact pressure follows e.g. via $(10)_1$ and (12) from the displacement variables.

In general there are different discretizations of Γ_c possible which depend on the problem (linear or nonlinear kinematics), on the discretization of the bodies in contact and on the type of constitutive interface law. Some discretizations are depicted in Figure 5 a) to d).

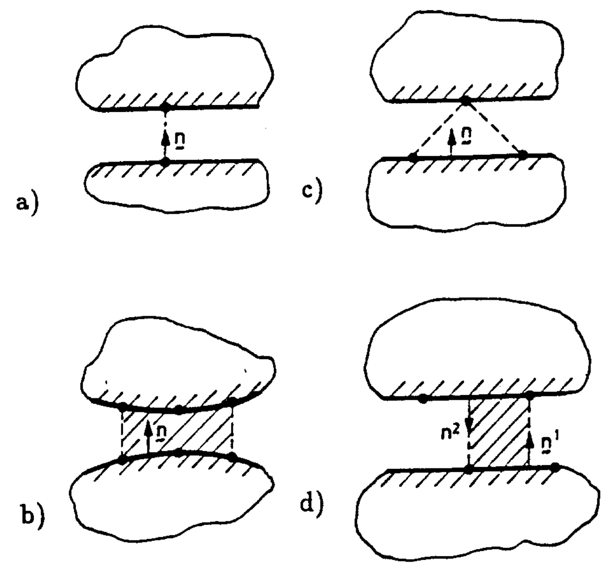


Figure 5. Different contact discretizations

5.1 Node–to–node Contact Element

Figure 5 a) shows the so-called node–to–node contact which can only be applied to geometrically linear problems since a relative tangential movement of the nodes is not allowed in the contact area. Due to its simplicity it resolves the integral (32) to

$$\int_{\Gamma_c} \epsilon_N \, g_N \, \delta g_N \, d\Gamma \longrightarrow \sum_{i=1}^{n_c} \epsilon_N \, g_{N\,i} \, \delta g_{N\,i} \, A_i = \sum_{i=1}^{n_c} \epsilon_N \, g_{N\,i} \, (\boldsymbol{\eta}_i^1 - \boldsymbol{\eta}_i^2) \cdot \mathbf{n}_i^2 \, A_i \tag{46}$$

where n_c are the contact nodes in Γ_c^h. The test function $\boldsymbol{\eta}_i^\alpha$ and the normal vector $\mathbf{n}_i^2$ is defined for the node i, see e.g. Wriggers, Zavarise (1993). Often the area A_i is neglected (or "hidden" in the penalty parameter ϵ_N) in the node–to–node contact formulation which means that the contact stress $p_N = \epsilon_N \, g_N$ becomes a contact (nodal) force $f_{N\,i} = \epsilon_N \, g_{N\,i}$. Then an evaluation of a contact interface law like (12) is not possible with discretization (46). The associated matrix formulation leads in the geometrically linear case for the contact element i to the definition of the contact residual $G_i^c = \boldsymbol{\eta}^T \mathbf{G}_i^c$ and its associated tangent matrix $\mathbf{K}_i^c$ with

$$\mathbf{G}_i^c = \epsilon_N \, g_{N\,i} \, \mathbf{N}_i \,, \quad \mathbf{K}_i^c = \epsilon_N \, \mathbf{N}_i \, \mathbf{N}_i^T \,, \quad \text{with} \quad \mathbf{N}_i = \begin{Bmatrix} \mathbf{n}_i^2 \\ -\mathbf{n}_i^2 \end{Bmatrix} \tag{47}$$

5.2 Isoparametric Discretization of the Contact Contribution

In Figure 5 b) a contact element is shown which also does not allow a relative tangential movement in the contact area and thus is only valid for geometrically linear applications. Within this element the gap function g_{N+} is discretized by an isoparametric interpolations leading also to a well defined contact pressure. We obtain with the interpolation

$$g_{N+}^h = \sum_I N_I(\xi) \, g_{NI} \quad \text{and} \quad \delta g_{N+}^h = \sum_I N_I(\xi) \, (\boldsymbol{\eta}_I^1 - \boldsymbol{\eta}_I^2) \cdot \mathbf{n}^2$$

the discretization of the contact integral (32)

$$\int_{\Gamma_c} \epsilon_N \, g_N \, \delta g_N \, d\Gamma \longrightarrow \int_{-1}^{1} \epsilon_N \, g_{N+}^h(\xi) \, [\sum_I N_I(\xi) \, (\boldsymbol{\eta}_I^1 - \boldsymbol{\eta}_I^2) \cdot \mathbf{n}^2] \, \| \frac{dx^h}{d\xi} \| \, d\xi \tag{48}$$

Finally numerical integration can be applied to evaluate (48). For the proper choice of the numerical integration rule, see e.g. Oden (1981) who has discussed this topic in the context of perturbed Lagrangian formulations. This discretization leads to a contact element which can be applied together with four or nine node quadilaterals for the continuum problem. Due to the smooth discretization a good approximation of the contact pressure is obtained.

5.3 Node–to–segment Contact Discretization

A more general discretization of the contact interface which allows also for large tangential sliding is given by the setup depicted in Figure 5 c). This discretization is named node–to–segment contact element and is widely used in nonlinear finite element simulations of contact problems.

Due to its importance we like to consider this contact element in more detail. Assume that the discrete slave point (s) comes into contact with the master segment (1)–(2), see Figure 6. The kinematical relations can be directly computed using the equations stated in section 2. With the interpolation for the master segment

$$\hat{\mathbf{x}}^2(\xi) = \mathbf{x}_1^2 + (\mathbf{x}_2^2 - \mathbf{x}_1^2) \, \xi \tag{49}$$

one can easily compute the tangent vector of the segment leading to

$$\bar{\mathbf{a}}_1^2 = \hat{\mathbf{x}}^2(\xi)_{,1} = (\mathbf{x}_2^2 - \mathbf{x}_1^2) \tag{50}$$

It is connected to an orthonormal base vector $\mathbf{a}_1^2$ by $\mathbf{a}_1^2 = \bar{\mathbf{a}}_1^2 \, / \, l$ with $l = \| \, \mathbf{x}_2^2 - \mathbf{x}_1^2 \, \|$ being the current length of the master segment. With the unit tangent vector $\mathbf{a}_1^2$ the unit normal to the segment (1)–(2) can be defined as $\mathbf{n}^2 = \mathbf{e}_3 \times \mathbf{a}_1^2$.

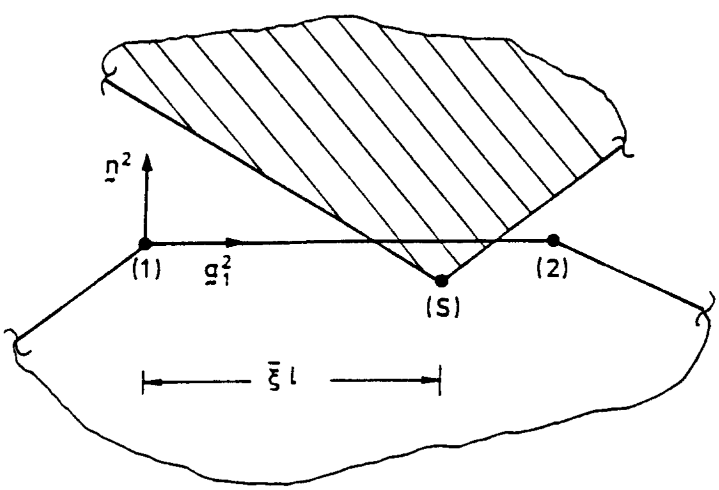


Figure 6. Node–to–segment element

$\bar{\xi}$ and g_N are given by the solution of the minimal distance problem, i.e. by the projection of the slave node $\mathbf{x}_s$ in (s) onto the master segment (1)–(2)

$$\bar{\xi} = \frac{1}{l}(\mathbf{x}_s^1 - \mathbf{x}_1^2) \cdot \mathbf{a}_1^2 \qquad \text{and} \qquad g_{N\,s} = \| \; \mathbf{x}_s^1 - (1-\bar{\xi})\,\mathbf{x}_1^2 - \bar{\xi}\,\mathbf{x}_2^2 \; \| \; . \tag{51}$$

From these equations and the local formulation (4) we compute directly the variation of the gap function δg_{N+} on the straight master segment (1)–(2).

$$\delta g_{N\,s} \;=\; [\,\boldsymbol{\eta}_s^1 - (1-\bar{\xi})\,\boldsymbol{\eta}_1^2 - \bar{\xi}\,\boldsymbol{\eta}_2^2\,] \cdot \mathbf{n}^2 \,. \tag{52}$$

The local equation (9) yields the expression for $\delta\bar{\xi}$. With the interpolation for the variation $\hat{\boldsymbol{\eta}}^2(\xi) = \boldsymbol{\eta}_1^2 + \xi\,(\boldsymbol{\eta}_2^2 - \boldsymbol{\eta}_1^2)$ on the straight master segment (1)–(2) we specialize

$$\begin{aligned} \bar{H}_{\alpha\beta} &= (a_{\alpha\beta} + g_N\,b_{\alpha\beta}) \Longrightarrow \bar{H}_{1\,1} = a_{1\,1} = l^2 \\ \bar{R}_1 &= [\,\boldsymbol{\eta}^1 - \hat{\boldsymbol{\eta}}^2(\bar{\xi})\,] \cdot \bar{\mathbf{a}}_1^2 + g_{N\,s}\,\bar{\mathbf{n}}^2 \cdot \hat{\boldsymbol{\eta}}_1^2(\bar{\xi}) \end{aligned}$$

which leads to

$$\delta g_T = l\,\delta\bar{\xi} \;=\; [\,\boldsymbol{\eta}_s^1 - (1-\bar{\xi})\,\boldsymbol{\eta}_1^2 - \bar{\xi}\,\boldsymbol{\eta}_2^2\,] \cdot \mathbf{a}_1^2 \;+\; \frac{g_{N\,s}}{l}\,[\,\boldsymbol{\eta}_2^2 - \boldsymbol{\eta}_1^2\,] \cdot \mathbf{n}^2 \,. \tag{53}$$

Equations (51), (52) and (53) characterize the main kinematical relations of the contact element in Figure 5c).

In what follows we compute the contribution of the node–to–segment element to the weak form (30). The basic formulation for this discretization is analogous to (46). Thus we assume that we know the normal force $P_{Ns} = p_{Ns}\,A_s$ and the tangential force $T_{Ts} = t_{Ts}\,A_s$ at the discrete contact point (s) of the contact element under consideration where A_s denotes the area of the contact element. Both forces, $P_{N\,s}$ and $T_{T\,s}$, can be obtained from the relations discussed in section 3. This leads to

$$\int_{\Gamma_c} (\,p_N\,\delta g_N + t_T\,\delta g_T\,)\,d\Gamma \longrightarrow \sum_{s=1}^{n_c} (\,P_{N\,s}\,\delta g_{N\,s} + T_{T\,s}\,\delta g_{T\,s}\,) \tag{54}$$

In practice we compute the normal force $P_{N\,s}$ either from equation (13) or (15) multiplied by the area of the contact element. For the tangential force $T_{T\,s}$ we have to perform an algorithmic update which is described in section 6.

Thus the contributions of one contact element in (54) takes the form

$$\delta g_{N\,s}\,P_{N\,s} \;+\; \delta g_{T\,s}\,T_{T\,s} \tag{55}$$

for the discrete contact point (s) with the mechanical relative (Lie–type) variations analogous to (52) and (53). This equations can now be cast into a matrix formulation. For the normal part $(54)_1$ we set for the variation (52) of the penetration

$$\delta g_{N\,s} = \boldsymbol{\eta}^T \mathbf{N}_s \,. \tag{56}$$

With the same notation we can express the variation (53) of the tangential gap

$$\delta g_{T\,s} = \boldsymbol{\eta}^T \;\left(\mathbf{T}_s + \frac{g_{N\,s}}{l}\,\mathbf{N}_{0\,s}\right) \,. \tag{57}$$

In (56) and (57) the following vectors have been used

$$\boldsymbol{\eta} = (\,\boldsymbol{\eta}_s^1 \quad \boldsymbol{\eta}_1^2 \quad \boldsymbol{\eta}_2^2\,)^T \;, \tag{58}$$

$$\mathbf{N}_s = \left\{ \begin{array}{c} \mathbf{n}^2 \\ -(1-\bar{\xi})\,\mathbf{n}^2 \\ -\bar{\xi}\,\mathbf{n}^2 \end{array} \right\}_s , \qquad \mathbf{N}_{0\,s} = \left\{ \begin{array}{c} \mathbf{0} \\ -\mathbf{n}^2 \\ \mathbf{n}^2 \end{array} \right\}_s , \tag{59}$$

and

$$\mathbf{T}_s = \left\{ \begin{array}{c} \mathbf{a}_1^2 \\ -(1-\bar{\xi})\,\mathbf{a}_1^2 \\ -\bar{\xi}\,\mathbf{a}_1^2 \end{array} \right\}_s , \qquad \mathbf{T}_{0\,s} = \left\{ \begin{array}{c} \mathbf{0} \\ -\mathbf{a}_1^2 \\ \mathbf{a}_1^2 \end{array} \right\}_s . \tag{60}$$

Thus the virtual mechanical work (55) of the contact element can be written in the matrix formulation $\boldsymbol{\eta}^T \mathbf{G}_s^c$ with the contact element residual

$$\mathbf{G}_s^c = P_{N\,s}\,\mathbf{N}_s + T_{T\,s}\,(\mathbf{T}_s + \frac{g_{N\,s}}{l}\mathbf{N}_{0\,s}) . \tag{61}$$

Due to this approach a pure displacement formulation of the contact problem is possible by expressing $P_{N\,s}$ either through (13) or (15) or by the penalty relation $P_{N\,s} = \epsilon_N\, g_{N\,s}$. This is in contrast to the Lagrangian multiplier technique, where $P_{N\,s} = \lambda_{N\,s}$. But we observe that this discretization can be applied to both methods. In case of the augmented Lagrangian method we have to replace $P_{N\,s}$ in (61) by

$$P_{N\,s}^{new} = \bar{P}_{N\,s}^{old} + \epsilon_N \left\{ g_{N\,s}^{new} - [\,\zeta - d(P_{N\,s}^{old})\,] \right\} \tag{62}$$

according to (41) where $g_{N\,s}$ is given by (51).

Often a Newton–Raphson iteration is used to solve the global set of equations. Then the linearization of (61) is needed to achieve quadratic convergence near the solution point. The associated derivation is a little bit cumbersome and thus only the final results will be summarized for this discretization. Details of the frictionless case can be found in Wriggers, Simo (1985) and for contact including friction in Wriggers, Vu Van, Stein (1990).

The tangent matrix for the normal contact is derived from the term $\delta g_{N\,s}\, P_{N\,s}$ in (55). Note that in (52) the change in $\bar{\xi}$ has be considered as well as the change of the normal $\mathbf{n}^2$. For the penalty approach with $P_{N\,s} = \epsilon_N\, g_{N\,s}$ we obtain the tangent matrix

$$\mathbf{K}_{N\,s}^c = \epsilon_N \left[\mathbf{N}_s\,\mathbf{N}_s^T - \frac{g_{N\,s}}{l} \left(\mathbf{N}_{0\,s}\,\mathbf{T}_s^T + \mathbf{T}_s\,\mathbf{N}_{0\,s}^T + \frac{g_{N\,s}}{l}\,\mathbf{N}_{0\,s}\,\mathbf{N}_{0\,s}^T \right) \right] \tag{63}$$

The used matrices have been defined in (59) and (60). Note that in a geometrically linear case all terms vanish which are multiplied by $g_{N\,s}$. This gives the simple matrix $\mathbf{K}_{N\,s}^{L\,c} = \epsilon_N\,\mathbf{N}_s\,\mathbf{N}_s^T$.

For the tangential contributions in the contact area we have to linearize the term $\delta g_{T\,s}\, T_{T\,s}$ in (55) which yields for the pure stick condition using equation $(18)_1$

$$\begin{aligned} \mathbf{K}_{T\,s}^c = c_T \Big\{ & (\mathbf{T}_s + \frac{g_{N\,s}}{l}\,\mathbf{N}_{0\,s})\,(\mathbf{T}_s + \frac{g_{N\,s}}{l}\,\mathbf{N}_{0\,s})^T \\ & + \frac{g_{N\,s}}{l} \big[\mathbf{N}_{0\,s}\,\mathbf{N}_s^T + \mathbf{N}_s\,\mathbf{N}_{0\,s}^T - \mathbf{T}_{0\,s}\,\mathbf{T}_s^T - \mathbf{T}_s\,\mathbf{T}_{0\,s}^T \\ & -2\,\frac{g_{N\,s}}{l}\,(\mathbf{N}_{0\,s}\,\mathbf{T}_{0\,s}^T + \mathbf{T}_{0\,s}\,\mathbf{N}_{0\,s}^T) \big] \Big\} \end{aligned} \tag{64}$$

Also in this case all terms containing $g_{N\,s}$ disappear in a geometrically linear situation which yields $\mathbf{K}_{T\,s}^{L\,c} = c_T\,\mathbf{T}_s\,\mathbf{T}_s^T$. The case of frictional slip leads to an additional contribution in (64) which will be discussed in section 6.2.

5.4 Discretization with Contact Segments

The discretization of the contact interface by segments as described for the linear case in Simo, Wriggers, Taylor (1985) or for large deformations in Papadopoulos, Taylor (1992) leads to a special mixed formulation. Following Simo, Wriggers, Taylor (1985) we state the interpolation of the gap function g_N and its variation δg_N for a geometrically linear setting as follows, see also remark I.3,

$$g_N = [\,\mathbf{u}^1(\xi) - \mathbf{u}^2(\xi)\,] \cdot \mathbf{n}(\xi) \qquad \delta g_N = [\,\boldsymbol{\eta}^1(\xi) - \boldsymbol{\eta}^2(\xi)\,] \cdot \mathbf{n}(\xi) \tag{65}$$

These interpolations are applied within a segement which is defined by the edge nodes $\mathbf{x}_2^A$ and the projections onto the other surface $\bar{\mathbf{x}}^A$, see Figure 7.

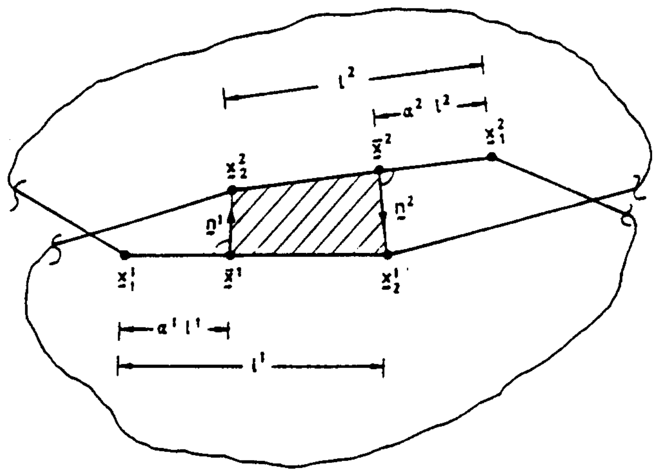


Figure 7. Contact segment element

Within this segment the displacement field and its variation is given as

$$\mathbf{u}^\gamma(\xi) = (1-\xi)\,\bar{\mathbf{u}}^\gamma + \xi\,\mathbf{u}_2^\gamma \qquad \boldsymbol{\eta}^\gamma(\xi) = (1-\xi)\,\bar{\boldsymbol{\eta}}^\gamma + \xi\,\boldsymbol{\eta}_2^\gamma \tag{66}$$

or the surface of the body $\mathcal{B}^\gamma$, $\gamma = 1\,,2$. For the perturbed Lagrangian approach (38) the contact contributions take the form

$$\begin{aligned} \int_{\Gamma_c} \lambda_N\,\delta g_N\,d\Gamma &= \sum_{s=1}^{n_{seg}} \int_{\Gamma_s} \lambda_N\,\delta g_N\,d\Gamma \\ \int_{\Gamma_c} (-\frac{\lambda_N}{\epsilon_N} + g_N)\,\delta\lambda_N\,d\Gamma &= \sum_{s=1}^{n_{seg}} \int_{\Gamma_s} (-\frac{\lambda_N}{\epsilon_N} + g_N)\,\delta\lambda_N\,d\Gamma = 0 \end{aligned} \tag{67}$$

where the latter equation can be solved for λ_N directly. With the interpolations (65) and assuming a constant contact pressure λ_N within the segment, $\lambda_N = \bar{\lambda}_N = CONST.$, we obtain for the segment Γ_s

$$\begin{aligned} \int_{\Gamma_s} \lambda_N\,\delta g_N\,d\Gamma &= \bar{\lambda}_N \int_0^1 [\,\boldsymbol{\eta}^1(\xi) - \boldsymbol{\eta}^2(\xi)\,] \cdot \mathbf{n}(\xi)\,\|\frac{d\Gamma}{d\xi}\|\,d\xi \\ \int_{\Gamma_s} (-\frac{\lambda_N}{\epsilon_N} + g_N)\,\delta\lambda_N\,d\Gamma \Longrightarrow \bar{\lambda}_N &= \frac{\epsilon_N}{L_s} \int_0^1 [\,\mathbf{u}^1(\xi) - \mathbf{u}^2(\xi)\,] \cdot \mathbf{n}(\xi)\,\|\frac{d\Gamma}{d\xi}\|\,d\xi \end{aligned} \tag{68}$$

As has been shown in Simo, Wriggers, Taylor (1985), the evaluation of these integrals by the trapezoidal rule yields the simple formulas

$$
\begin{aligned}
\bar{\lambda}_N \int_0^1 [\,\boldsymbol{\eta}^1(\xi) - \boldsymbol{\eta}^2(\xi)\,] \cdot \mathbf{n}(\xi)\, \| \frac{d\Gamma}{d\xi} \|\, d\xi \approx \frac{1}{2}\, \bar{\lambda}_N\, (\,\delta g_N|_{\xi=0} + \delta g_N|_{\xi=1}\,) \\
\bar{\lambda}_N \approx \frac{\epsilon_N}{2}\, (\,g_N|_{\xi=0} + g_N|_{\xi=1}\,)
\end{aligned}
\tag{69}
$$

where g_N and δg_N can be expressed by the quantities in equations (65) and (66). This completes the discretization for contact segments. For more details, see Simo, Wriggers, Taylor (1985), and for its nonlinear extension, see Papdopoulos, Taylor (1992).

5.5 Global Set of Equations

For a global algorithmic treatment we have to state the discrete set of equations. This leads for the penalty method to the general matrix formulation of the weak form

$$
\mathbf{G}_c^p(\mathbf{v}) = \mathbf{G}(\mathbf{v}) + \cup_{s=1}^{n_c} \mathbf{G}_s^c(\mathbf{v}) = \mathbf{0} \tag{70}
$$

where $\mathbf{G}(\mathbf{v})$ denotes the contributions of the bodies due to the weak form (43). In the second term s is associated with the active contact element, node or segment and $\mathbf{G}_s^c(\varphi)$ has to be computed according to the chosen discretization, see e.g. sections 5.1 to 5.4.

For the Lagrangian multiplier method the set of equations yields

$$
\begin{aligned}
\mathbf{G}_c^1(\mathbf{v}\,,\boldsymbol{\lambda}) &= \mathbf{G}(\mathbf{v}) + \cup_{s=1}^{n_c} \mathbf{C}_s^l(\mathbf{v})^T \lambda_s = \mathbf{0} \\
\mathbf{G}_c^2(\mathbf{v}\,,\boldsymbol{\lambda}) &= \cup_{s=1}^{n_c} \mathbf{C}_s^g(\mathbf{v}) = \mathbf{0}
\end{aligned}
\tag{71}
$$

Here the matrix $C_s^l(\mathbf{v})$ is related to the variation of δg_s, see e.g. $(68)_1$, and $\mathbf{C}_s^g(\mathbf{v})$ denotes the matrix formulation of the gap function g_s itself, see e.g. $(68)_2$. These matrices also depend on the chosen discretization and are ment to contain not only the terms of the normal contact as indicated in (68) but also the terms due to friction.

In case that Newton type methods are employed to solve (70) or (71) a linearization of the discrete set of equations has to be performed. Especially in the large deformation case the change in the normal has to be taken into account. The resulting expressions can be found for the two dimensional frictionless case in Wriggers, Simo (1985) an extention for finite frictional slip has been derived in Curnier, Alart (1988) and Wriggers, Vu Van, Stein (1990). Three dimensional discretizations and linearizations have been obtained by Parisch (1988) for the frictionless and by Laursen, Simo (1993) for the frictional contact.

6. ALGORITHMS FOR CONTACT PROBLEMS

In this section we consider the algorithms which are essential for the treatment of contact problems. In general we have to distinguish between global algorithms which are necessary to find the correct number of active constraint equations and local algorithms which are needed to update contact stresses within the constitutive equations in the interface. Furthermore, also algorithms have to be deviced for coupled problems which may be necessary in case of thermomechanical coupling or for fluid–structure interaction problems.

The bandwidth of the global algorithms for constraint optimization is very broad. We like to mention, see also the introductory remarks, the simplex method, active set strategies, sequential quadratic programming, penalty and augmented Lagrangian techniques as well as barrier methods. All these techniques have advantages and disadvantages concerning efficiency, accuracy or robustness and thus have to be applied according to the problem at hand. Algorithms for coupled problems, like staggered schemes, depend on the type of coupling and thus have to be designed with special care regarding robustness and efficiency. In the following we will sketch some of the global algorithms which are mainly applied to contact problems.

The update algorithms for the contact stresses, especially the tangential stresses due to friction, have been settled. In this case the so called projection methods or return mapping schemes yield the most efficient and robust treatment. Due to the fact that a algorithmic tangent operator can be constructed this technique can be incorporated in a Newton–Raphson scheme.

6.1 Global Algorithms

The algorithm which is applied in many standard finite element programs is related to the penalty method. This is mainly due to its simplicity and furthermore it yields for many applications a robust algorithm. The penalty method is mostly combined with an active set strategy. The global set of equations is given in (70). Now the algorithm for the penalty method can be summarized in Box 1.

Initialize algorithm
set:$\mathbf{v}_1 = \mathbf{0}\,, \quad \epsilon_N = \epsilon_0$
LOOP over iterations : $i = 1, .., $ convergence
Check for contact: $g_{N\,si} \leq 0 \rightarrow$ active node, segment or element
Solve: $\mathbf{G}_c(\mathbf{v}_i) = \mathbf{G}(\mathbf{v}_i) + \cup_{s=1}^{n_c} \mathbf{G}_s^c(\mathbf{v}_i) = \mathbf{0}$
Check for convergence: $\|\mathbf{G}_c(\mathbf{v}_i)\| \leq TOL \Rightarrow$ END LOOP
END LOOP
Eventually update penalty parameter: ϵ_N

Box 1. Contact algorithm using the penalty method

Usually the solution of $\mathbf{G}_c(\mathbf{v}) = \mathbf{0}$ is performed by a Newton–Raphson iteration leading to

$$\begin{aligned} D\,\mathbf{G}_c(\mathbf{v}_i^n)\,\Delta\mathbf{v}_i^{n+1} &= -\mathbf{G}_c(\mathbf{v}_i^n) \\ \mathbf{v}_i^{n+1} &= \mathbf{v}_i^n + \Delta\mathbf{v}_i^{n+1} \end{aligned} \tag{72}$$

where the operator D denotes the directional derivative of the vector $\mathbf{G}_c(\mathbf{v}_i^n)$ which results in the tangent matrix $\mathbf{K}_T(\mathbf{v}_i^n) = D\,\mathbf{G}_c(\mathbf{v}_i^n)$. The iteration index n is related to the Newton loop to solve $\mathbf{G}_c(\mathbf{v}_i) = \mathbf{0}$ in Box 1. Often the active set strategy, stated in Box 1, is accelerated in such a way that the update of the active set of contact constraints is performed within each step in the Newton iteration. Then the iteration (72) yields

$$\begin{aligned} D\,\mathbf{G}_c(\mathbf{v}_i)\,\Delta\mathbf{v}_{i+1} &= -\mathbf{G}_c(\mathbf{v}_i) \\ \mathbf{v}_{i+1} &= \mathbf{v}_i + \Delta\mathbf{v}_{i+1} \end{aligned} \tag{73}$$

which is considerably faster. However this procedure might not converge for all cases and thus has to be applied with care.

Within the algorithm of Box 1 an increase of the penalty parameter is necessary when the final result shows visible penetrations and thus does not fulfill the constraint equation $g_{n+} = 0$ in a correct way. On the other hand a penalty parameter which has been chosen too large can lead to ill–conditioning of the equation system and thus has to be reduced to avoid this. On possibility for the choice of ϵ_N is to relate the penalty parameter to the bulk modulus of the contacting bodies. However, since it is quite hard to estimate the penalty parameter for all cases it makes sense to apply the augmented Lagrangian technique.

Augmented Lagrangian technique are usually applied together with Uzawa type algorithms, see Bertsekas (1984), Glowinski, Le Tallec (1984) or Laursen, Simo (1991), which lead to an inner loop for the contact and an outer loop for the update of the Lagrangian parameters.

Let us remark that it is standard practice in augmented Lagrangian iterations also to update the penalty number ϵ_N in order to obtain good convergence, see Bertsekas (1984). This is due to the fact that a small penalty parameter leads to very slow convergence since the update formula (42) is of first order and the contact forces due to the penalty are small. Thus it makes sense to increase the penalty parameter within a contact element s according to an update scheme, see Bertsekas (1984). Here we like to show this approach for the augmented Lagrangian scheme in combination with constitutive interface laws like (13). The update scheme yields

$$\epsilon_{N\,s\,n+1} = \begin{cases} 10 \cdot \epsilon_{N\,s\,n} & \text{for } [c_+(\mathbf{V}_s, \bar{P}_{N\,s})]_{n+1} > \frac{1}{4} \cdot [c_+(\mathbf{V}_s, \bar{P}_{N\,s})]_n \text{ and } \epsilon_{N\,s\,n} \le \frac{k}{\sqrt{N\,t}} \\ \epsilon_{N\,s\,n} & \text{for } [c_+(\mathbf{V}_s, \bar{P}_{N\,s})]_{n+1} \le \frac{1}{4} \cdot [c_+(\mathbf{V}_s, \bar{P}_{N\,s})]_n \end{cases} . \quad (74)$$

In relation (74) also a stopping criterion for the update of the penalty parameter has been introduced to avoid ill–conditioning. This is given by the estimate (39). The global augmented Lagrangian algorithm is shown in Box 2. Here we use again the discrete formulation (70) which has to be adjusted to incorporate the fixed Lagrangian parameters $\bar{P}_{N\,s}$, see (62) for the node–to–segment discretization. By $\cup_{s=1}^{n_c} \mathbf{G}^a_{s\,n+1}(\mathbf{v},\, \bar{P}_{N\,s})$ we denote the contribution of the fourth term in (41) for an active contact element s.

Initialize algorithm
set: $d_0 = \xi\,, \quad \mathbf{v} = \mathbf{0}\,, \quad \bar{P}_0 = 0\,, \quad \epsilon_N = \epsilon_{N\,0}$
LOOP over augmentations: $n = 1, .., $ convergence
  LOOP over iterations : $i = 1, .., $ convergence
    Solve: $\mathbf{G}_c(\mathbf{v}_i,\, \bar{P}_{N_n}) = \mathbf{G}(\mathbf{v}_i) + \cup_{s=1}^{n_c} \mathbf{G}^a_{s\,n+1} = \mathbf{0}$
    Check for convergence: $\|\mathbf{G}_c(\mathbf{v}_i,\, \bar{P}_{N_n})\| \le TOL \Rightarrow$ END LOOP
  END LOOP
  LOOP over contact nodes : $s = 1, .., n_c$
    Update: $\bar{P}_{N_{s\,n+1}}$ according to (62)
    Update: $d_{s\,n+1} = h(\bar{P}_{N_{s\,n+1}})$ according to (12)
    Update: $\epsilon_{N\,s\,n+1}$ according to (74)
    Check for convergence: $\frac{1}{\zeta}\,\|g_{N_+}(\mathbf{V}_{s\,i}) - (\zeta - d_{n+1})\| \le TOL \Rightarrow STOP$
  END LOOP
END LOOP

Box 2. Augmented Lagrangian algorithm

Additional algorithms for contact problems can be found in the literature, for references see section 1.3.

6.2 Local Update Algorithm for Tangential Contact Stress

The algorithmic update of the tangential stress $\mathbf{t}_{T\,n+1}$ due to friction is performed by the return algorithm based on an objective (backward Euler) integration of the evolution equation (22) for the plastic slip, see e.g. Wriggers (1987), Ju & Taylor (1988), Giannokopoulos (1989), Wriggers, Vu Van, Stein. (1990). The results can be summarized as follows: Integration of (10) gives the increment of the total slip within the time step Δt_{n+1}

$$\Delta \mathbf{g}_{T\,n+1} = (\,\bar{\xi}^{\alpha}_{n+1} - \bar{\xi}^{\alpha}_{n}\,)\bar{\mathbf{a}}_{\alpha\,n+1}\,. \tag{75}$$

The total slip has to be decomposed into an elastic and an plastic part, see $(18)_2$. Thus in the case of contact, i.e. for $g_{N+\,n+1} > 0$, we know the contact pressure $p_{N_{n+1}}$. Then we can compute the elastic trial state $(18)_1$ and evaluate the slip criterion (19) or (21) for that state

$$\begin{aligned} \mathbf{t}^{tr}_{t\,n+1} &:= c_T\,(\,\mathbf{g}_{T\,n+1} - \mathbf{g}^{s}_{T\,n}\,) = \mathbf{t}_{T\,n} + c_T\,\Delta\mathbf{g}_{T\,n+1}\,, \\ f^{tr}_{s\,n+1} &:= \|\mathbf{t}^{tr}_{T\,n+1}\| - \mu\,p_{N\,n+1}\,. \end{aligned} \tag{76}$$

If this state is elastic ($f^{tr}_{s\,n+1} \le 0$) then no friction takes place and we have to use the elastic relation $(18)_1$. In case that $f^{tr}_{s\,n+1} > 0$ then we have to perform the return algorithm. Using the implicit Euler scheme, (22) yields

$$\mathbf{g}^{s}_{T\,n+1} = \mathbf{g}^{s}_{T\,n} + \Delta\lambda\,\mathbf{n}_{T\,n+1}\,. \tag{77}$$

With the standard arguments regarding the projection schemes, see e.g. Simo, Taylor (1985), we obtain

$$\begin{aligned} \mathbf{t}_{T\,n+1} &= \mathbf{t}^{tr}_{t\,n+1} - \Delta\lambda\,c_T\,\mathbf{n}_{T\,n+1}\,, \\ \mathbf{n}_{T\,n+1} &= \mathbf{n}^{tr}_{T\,n+1}\,, \\ \Delta\lambda &= \frac{1}{c_T}\,(\,\|\,\mathbf{t}^{tr}_{t\,n+1}\,\| - \mu\,p_{N\,n+1}\,)\,. \end{aligned} \tag{78}$$

From these relation we can compute the stress update $\mathbf{t}_{T\,n+1}$ and the frictional slip $\mathbf{g}^{s}_{T\,n+1}$:

$$\begin{aligned} \mathbf{t}_{T\,n+1} &= \mu\,p_{N\,n+1}\,\mathbf{n}^{tr}_{T\,n+1}\,, \\ \mathbf{g}^{s}_{T\,n+1} &= \mathbf{g}^{s}_{T\,n} + \frac{1}{c_T}\,(\,\|\,\mathbf{t}^{tr}_{t\,n+1}\,\| - \mu\,p_{N\,n+1}\,)\,\mathbf{n}^{tr}_{T\,n+1}\,. \end{aligned} \tag{79}$$

which completes the algorithm for the frictional interface law.

The tangent matrix which is needed within a Newton iteration can be derived by linearizing the term which appears in the weak form with respect to the displcaement field. In the two dimensional case of the node–to–segement contact element, see section 5.3, the explicit matrix form results from the term $\delta g_{T\,s\,n+1}\,T_{T\,s\,n+1}$ and can be stated for a contacting node (s), with $\mathbf{K}^{c}_{T\,s}$ from (64), as

$$\mathbf{K}^{S\,c}_{T\,s} = \mathbf{K}^{c}_{T\,s} + \mu\,\epsilon_N\,\left(\mathbf{T}_s + \frac{g_{N\,s}}{l}\,\mathbf{N}_{0\,s}\right)\mathbf{N}^{T}_{s} \tag{80}$$

Note that this matrix is unsymmetric which corresponds to the non–associativity of Coulomb's frictional law.

7. ADAPTIVE METHODS FOR CONTACT PROBLEMS

In this section we like to discuss adaptive finite element methods which include also error estimates for contact problems. Mathematically sound error indicators have been derived so far only for small strain problems due to the analytical complexity. Here we like to summarize some results which can be applied to contact problems. Basically one has two different possibilities to derive error estimators which can be applied within adaptive methods to refine the finite element mesh. These are the residual based error estimators and the projection methods which rely on superconvergence properties. Both techniques will be discussed for geometrically linear problems.

Let $\mathbf{u}$ denote the exact solution of (29) and let $\mathbf{u}_h$ denote the discrete FEM–solution of (70). With

$$\mathbf{e} = \mathbf{u} - \mathbf{u}_h \tag{81}$$

we define the error in the displacement field.

7.1 Residual Based Error Estimator for Contact

For the linear elastic problem a residual based error estimator can be found in Johnson, Hansbo (1992) for the stresses as follows

$$\| \boldsymbol{\sigma} - \boldsymbol{\sigma}_h \|^2_{E^{-1}} \leq \| h\, C_1\, R_1(\boldsymbol{\sigma}_h) \|^2_{L_2(\Omega)} + \| h\, C_2\, R_2(\boldsymbol{\sigma}_h) \|^2_{L_2(\Omega)} \tag{82}$$

where the quantities are defined on the finite element as follows

$$\begin{aligned} R_1(\boldsymbol{\sigma}_h) &= |\, R_1(\boldsymbol{\sigma}_h) \,| = |\, \operatorname{div} \boldsymbol{\sigma}_h + \hat{\mathbf{b}} \,| && \text{on } T \\ R_2(\boldsymbol{\sigma}_h) &= \max_{S \in \partial T} \sup_S \frac{1}{2\, h_T} \,|\, [\, \boldsymbol{\sigma}_h\, \mathbf{n}_S \,] \,| && \text{on } \partial T \\ \text{or } R_2(\boldsymbol{\sigma}_h) &= \frac{1}{h_T} \,(\, \hat{\mathbf{t}} - \boldsymbol{\sigma}_h\, \mathbf{n} \,) && \text{on } \partial T \cap \Gamma_\sigma \end{aligned} \tag{83}$$

Here Ω denotes the discretized region, h_T is a characteristic length of an element, T is the area of a finite element and ∂T its surface. The norm $\| \cdot \|_{E^{-1}}$ in (82) is the complementary energy norm (written in stress space)

$$\| \boldsymbol{\sigma} - \boldsymbol{\sigma}_h \|^2_{E^{-1}} = \left\{ \int_\Omega (\, \boldsymbol{\sigma} - \boldsymbol{\sigma}_h \,) \, : \, \boldsymbol{\mathcal{C}}_0^{-1} \, : \, (\, \boldsymbol{\sigma} - \boldsymbol{\sigma}_h \,)\, d\Omega \right\}^{1/2} \tag{84}$$

The equivalence of this norm to the energy norm can be shown easily by inserting the constitutive equations for elasticity into the last expression leading to $\| \boldsymbol{\sigma} - \boldsymbol{\sigma}_h \|^2_{E^{-1}} = \| \mathbf{e} \|^2_E$.

In Wriggers, Scherf, Carstensen (1994) an additional term for the error associated with contact has been derived:

$$R_3(\boldsymbol{\sigma}_h, \mathbf{u}) = |\, \epsilon_N\, g_{N+}\, \mathbf{n}^2 - \mathbf{t}_h | \quad \text{on } \partial T \cap \Gamma_c \tag{85}$$

where the term on the right side corresponds to the local equilibrium in the contact interface. The term $\epsilon_N\, g_{N+}\, \mathbf{n}^2$ can be interpreted as the contact pressure on Γ_c. Adding (85) to equation (82) leads for the linear elastic contact problem to the following *a posteriori* error estimate

$$\| \boldsymbol{\sigma} - \boldsymbol{\sigma}_h \|^2_{E^{-1}} \leq \sum_{k=1}^{3} \| h\, C_k\, R_k(\boldsymbol{\sigma}_h) \|^2_{L_2(\Omega)} \tag{86}$$

A thorough mathematical derivation of the *a posteriori* error estimator can be found in Carstensen, Scherf, Wriggers (1995). Within the finite element discretization equation (86) has to be evaluated on the element domain which yields

$$\| \boldsymbol{\sigma} - \boldsymbol{\sigma}_h \|^2_{E^{-1}} \leq C \sum_T [E_T(h_T, \mathbf{u}_h, \hat{\mathbf{b}}_T)]^2 \tag{87}$$

E_T can be computed for each element in the finite element mesh as follows, see Wriggers, Scherf, Carstensen (1994)

$$\begin{aligned} E_T^2 = & h_T^2 \int_T |\,\mathrm{div}\boldsymbol{\sigma}_h + \hat{\mathbf{b}}\,|^2 d\Omega \;+\; h_T \int_{\partial T \cap \Omega} 1/2\,|\,[\mathbf{t}_h]\,|^2 d\Gamma \;+ \\ & h_T \int_{\partial T \cap \Gamma_\sigma} |\,\hat{\mathbf{t}} - \mathbf{t}_h\,|^2\, d\Gamma \;+\; h_T \int_{\partial T \cap \Gamma_c} |\,\epsilon_N\, g_{N+}\, \mathbf{n} - \mathbf{t}_h|^2\, d\Gamma \end{aligned} \tag{88}$$

Inequality (87) yields an upper bound for the error which is bounded by the deviation of the discrete solution from equilibrium and the element size. The first and the third term of the right hand side contribute to the error bound if the local equilibrium and the traction boundary condition, respectively, are violated. In (88) we have introduced the stress vector $\mathbf{t}_h = \boldsymbol{\sigma}_h\,\mathbf{n}$. Local equilibrium requires that $[\mathbf{t}_h] = \mathbf{0}$ which is associated with the second term where $[\mathbf{t}_h]$ describes the jumps of the tractions over the interface. The fourth term has already been discussed above.

The error estimator described above yields a measure between the exact penalty solution of (30) with (33) and its finite element approximation (43) with (45). What really is needed is the error between the exact solution of (29) and the approximate finite element solution (43) with (45). So far there do not exist computable error bounds for contact problems in elasticity. But we can make use of a result derived by Kikuchi, Oden (1988) to change the penalty parameter in such a way that an optimal convergence rate of the method is achieved. To this purpose we state the result of Kikuchi, Oden (1988) which was derived for a perturbed Lagrangian formulation of the contact problem

$$\| \mathbf{u} - \mathbf{u}_{\varepsilon h} \|_1 + |\, p_N - p_{N\,\varepsilon h}\,|^* \leq C_3\, h + C_4\, \epsilon_N^{-1}\, h^{-1/2} \tag{89}$$

From this equation it is clear that an optimal convergence rate can be obtained for $\epsilon_N \approx h^{-3/2}$, h being the characteristic length of an element. According to this relation we develop now the following update at iteration $k+1$ for the penalty parameter in the contact interface

$$\epsilon_{N\,k+1} = \epsilon_{N\,0} \left(\frac{h_{k+1}}{h_0} \right)^{-\frac{3}{2}} ,$$

where $\epsilon_{N\,0}$ and h_0 are the starting values at the beginning of the adaptive iteration.

7.2 Error Estimator for Contact Based on Projection

Another possibility to derive an error estimator for elastic contact problems starts directly from the complementary elastic energy norm (84). A simple but in many cases efficient error estimator is now provided by the the superconvergent–stress–recovery technique which is due to Zienkiewicz, Zhu (1987). The equivalence of such error measures with the residual based error estimators of the last section has been shown in Verführt (1993). The idea to derive these error estimators is based on the fact that many finite element meshes have superconvergence properties which means that there exist points in which the stresses are approximated with higher accuracy. By using projection procedure the stresses $\boldsymbol{\sigma}^*$ can be computed from the superconvergent points, see e.g. Zienkiewicz, Taylor (1988). It should be noted in passing that the stress–recovery error estimators work also well in if the sampling points are not superconvergence points, see Babuska, Strouboulis, Upadhyay, Gangaraj, Copps (1994). An especially efficient projection technique is provided by the lumped L^2–projection which is described in detail in Zienkiewicz, Taylor (1988). In general the projection procedures assume that the projected stresses do not have jumps which needs some special considerations in the contact interface, see below. Denoting by $\mathbb{P}$ a projection operator we obtain $\boldsymbol{\sigma}^*$ from

$$\int_\Omega \mathbb{P}\left[\,\boldsymbol{\sigma}^* - \boldsymbol{\sigma}_h\,\right] d\Omega = \mathbf{0} \tag{90}$$

and can then compute an approximation of the error using (84)

$$\|\,\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\,\|^2_{E^{-1}} \leq \int_\Omega (\,\boldsymbol{\sigma}^* - \boldsymbol{\sigma}_h\,) \,:\, \boldsymbol{\mathcal{C}}_0^{-1} \,:\, (\boldsymbol{\sigma}^* - \boldsymbol{\sigma}_h)\, d\Omega \tag{91}$$

This error estimator can be evaluated in an efficient way and has been shown to be robust, see Babuska, Strouboulis, Upadhyay, Gangaraj, Copps (1994). Equation (91) does not include special boundary terms for the contact contributions. The estimate for the contact area is in this case included implicitly since the evaluation of the integral in (91) has to be done with respect to $\Omega = \Omega^1 \cup \Omega^2$ and thus includes also the contact interface. However a special projection has to be performed since one has to treat normal and tangential stress components in the contact interface differently. Thus in the frictionless contact the normal component, given by $p_N = \mathbf{n}^2 \cdot \boldsymbol{\sigma} \cdot \mathbf{n}^2$, has to be projected using all elements connected to a point in the contact interface, e. g. from both bodies. For the tangential stresses the projection $\mathbb{P}$ can only be applied within the body Ω^γ to evaluate the error estimator. Thus first this special projection scheme has to be used and then the normal and tangential components of the stresses are transformed back to $\boldsymbol{\sigma}^*$ and evaluated with (91). This technique has been applied in Wriggers, Scherf (1995) for frictionless elasto–plastic contact problems.

The error within the whole domain is computed by the sum over all elements T with Ω being the union of all elements. Thus we have

$$\|\,\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\,\|_{E^{-1}} = \|\,\mathbf{e}\,\|_E = \sum_T \|\,\mathbf{e}\,\|_T \tag{92}$$

The technique described above has the disadvantage that the special projection for the contact stresses is time consuming since additional search procedures are necessary to find the elements related to a node on Γ_c. Thus a direct application of the Z^2–projection to the contact stresses on each surface $\partial\Omega^\gamma \cap \Gamma_c$ is preferable. This leads to the additional error in $\partial\Omega^\gamma \cap \Gamma_c = \sum_T \partial T \cap \Gamma_c$

$$\|\,\mathbf{e}\,\|^{C\,\gamma} = \sum_T \|\,\mathbf{e}\,\|^{C\,\gamma}_{\partial T} \tag{93}$$

with

$$\|\,\mathbf{e}\,\|^{C\,\gamma}_{\partial T} = \int_{\partial T \cap \Gamma_c} \frac{1}{\epsilon_N}\,(p_N^* - p_{N\,h})^2\, d\Gamma + \int_{\partial T \cap \Gamma_c} \frac{1}{c_T}\,(\mathbf{t}_T^* - \mathbf{t}_{T\,h}) \cdot (\mathbf{t}_T^* - \mathbf{t}_{T\,h})\, d\Gamma \tag{94}$$

where p_N^* and $\mathbf{t}_T^*$ are obtained by a projection along the surface $\partial\Omega^\gamma \cap \Gamma_c$. In (94) $\|\,\mathbf{e}\,\|^{C\,\gamma}_{\partial T}$ is computed for the case of tangential stick. For frictionless contact the second term in (94) has to be omitted. Note that for frictional contact problems so far no sound mathematical error estimators exist and thus could not be included in this overview.

Equation (94) provides the additional term due to contact which has to be included in the error $\|\mathbf{e}\|_T^\gamma$ within an element T in Ω^γ which is connected to Γ_c

$$\| \mathbf{e} \|_T^C = \| \mathbf{e} \|_T^\gamma + \| \mathbf{e} \|_{\partial T}^{C\,\gamma} \tag{95}$$

REMARK V:

So far the error estimators have been developed for small elastic deformations. In case of large elastic strains there is no mathematically sound basis. However it should be noted that existence results exist for polyconvex materials, see e.g. the overview in Ciarlet (1988). These results have been extended to contact problems in Ciarlet (1988), Oden, Kikuchi (1988) and Curnier, He and Telega (1992). Thus there is at least an existence result available for contact problems. The question of uniqueness, also needed for the derivation of error estimators, can of course not be solved since problems undergoing large elastic deformations may exhibit as well material as geometrical instabilities (e.g. limit points or bifurcations).

If we now formulate all equations associated with the variational inequality (28) in the tangent space of a given deformation map and exclude within this configuration instabilities, then the information from these incremental equations can be used for an error estimate within the incremental step. This means that we can exchange the stresses in (88) by the appropriate nonlinear stress measures when using the residual based error estimator. If we employ the error estimator based on the superconvergent recovery technique then additionally the incremental constitutive tensors have to be used in (91) and (94) to compute the error (95).

7.3 Adaptive Mesh Refinement Strategy

An adaptive algorithm is usually stated as a nonlinear optimization problem: contruct a mesh such that the associated FEM–solution satisfies

$$\| \boldsymbol{\sigma} - \boldsymbol{\sigma}_h \|_{E^{-1}} = \|\mathbf{u} - \mathbf{u}_h\|_E \le C \sum_T [E_T(h_T, \mathbf{u}_h, \hat{\mathbf{b}}_T)]^2 \le TOL\,, \tag{96}$$

with TOL being a given tolerance. Furthermore the expense to compute $\mathbf{u}_h$ or $\boldsymbol{\sigma}_h$ should be nearly minimal. The measure E_T can be either

$$\begin{aligned}
E_{T\,1}^2 &= E_T^2 \ \text{ from equation (88) or} \\
E_{T\,2}^2 &= \int_T (\,\boldsymbol{\sigma}^* - \boldsymbol{\sigma}_h\,) \,:\, \boldsymbol{\mathcal{C}}_0^{-1} \,:\, (\boldsymbol{\sigma}^* - \boldsymbol{\sigma}_h)\, d\Omega \ \text{ or} \\
E_{T\,3}^2 &= \|\mathbf{e}\|_T^C \ \text{ from equation (95)}
\end{aligned}$$

As a measure of computational work the total number of degrees of freedom is chosen. Since the exact solution $\mathbf{u}$ is not known we demand that the error contributions of all elements yields

$$\sum_T E_{T_k}^2 \le TOL\,, \tag{97}$$

which guarantees that (96) is fulfilled, k being 1, 2 or 3, depending on the choice of error estimator. Here the constant C appearing in (87) has been included in TOL for convenience when the error measure $E_{T\,1}$ is used. (97) serves as a stopping criterion in the adaptive process. To minimize the number of degrees of freedom during refinement, we require that the mesh is an optimal mesh, i.e. that the error E_T^2 is equally distributed between elements:

$$\sum_T E_{T_k}^2 = N\, E_{T_k}^2\,. \tag{98}$$

N denotes the number of elements in the mesh. Finally (95) yields together with (96) the refinement criterion

$$E_{T_k}^2 \leq \frac{TOL}{N}\,.$$

Now we state the overall algorithm of the h–adaptive method for contact problems. The algorithm includes the following steps:

1. Set initial values: $l = 0$, $\lambda_0 = 0$, $\Delta\lambda$, $i = 0$
2. Generation of start mesh: $\mathcal{M}_i$
3. Loop over load increments : $\lambda_{l+1} = \lambda_l + \Delta\lambda$
   3.1 IF $\lambda_{l+1} > \lambda_{max} \Longrightarrow$ STOP
   3.2 Iteration loop to solve contact problem
   3.3 Mesh optimization
      - Compute $E_{T_k}^2$
      - IF $\sum E_{T_k}^2 < TOL \Longrightarrow$ GOTO 3.
      - IF $E_{T_k}^2 > TOL\,/\,N \Longrightarrow$ refine element T
      - Set $i = i + 1$
      - Generate new mesh $\mathcal{M}_i$
        Delaunay triangularization
        Smoothing, if necessary
      - Interpolate displacement and history variables the new mesh
      - GOTO 3.2

The new mesh is assumed to be generated by a Delaunay triangularization, but also different generation techniques like the advancing front method or others can be applied. Smoothing is used when the form of the element deteriorates too much, e. g. an inner angle becomes too small. This procedure can be quite costly and has to be implemented with care. For cases with large deformations or involving inelastic materials the displacement and history variables have to be transferred to the new mesh, see e.g. Ortiz, Quigley (1991).

8. NUMERICAL EXAMPLES

Numerical examples are presented in this section to show some results of large deformation contact processes and of the adaptive methods. The first two examples depict the behavior of large sliding within the contact interface for frictionless contact and contact with friction. The discretization is based on the formulation given in section 5.3 which is used in many finite element codes. The last two examples exploit the adaptive finite element scheme discussed in section 7. Here the classical Hertz problem is solved as well as a finite deformation problem of a rubber sealing. All problems have been simulated using an extended version of the finite element program FEAP developed by R. L. Taylor, see Zienkiewicz, Taylor (1988).

8.1 Large Frictionless Sliding of a Rubber Blade

In this example large sliding of the cross section of an elastic rubber blade is considered. The rubber blade, as shown in Figure 8, is subjected to a vertical prescribed displacement and pressed against a rubber block. The contact is assumed to be frictionless. In this example the penalty method is applied to fulfill the contact constraint conditions. The hyperelastic material response is described by a one term strain energy function derived by Ogden (1984) for compressible Neo–hookian materials: $W = \frac{\mu_1}{\alpha_1}\,(\lambda_1^\alpha + \lambda_2^\alpha + \lambda_3^\alpha - 3) - \mu_1 \ln J + \frac{\Lambda}{2}\,(J-1)^2$. The constitutive parameters for the block and the blade are given in the Table 1.

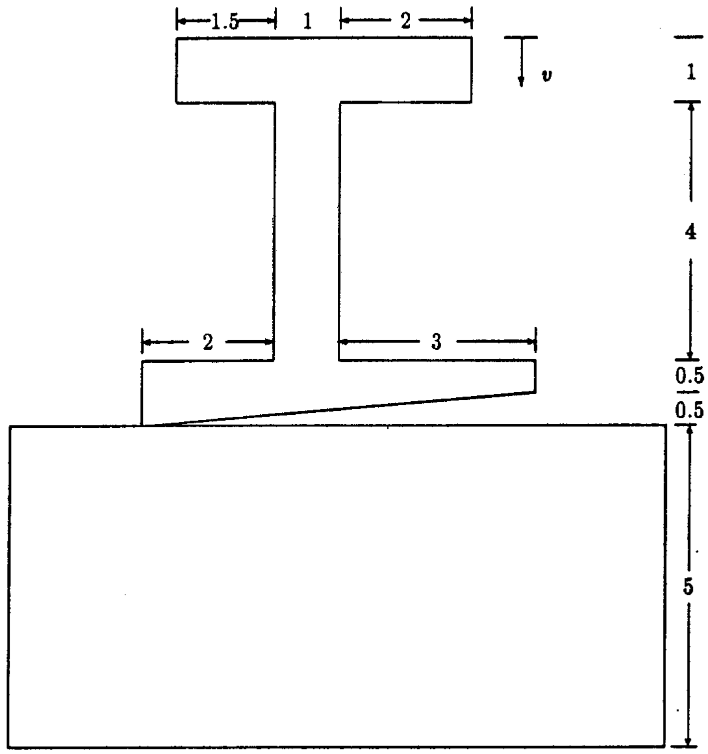


Figure 8. Rubber blade contacting an elastic block

Part	Λ	μ_1	α_1
Blade	4	6.3	1.3
Block	10	6.3	1.3

Table 1. Constitutive parameters

It.	1	2	3	4	5
$\|\mathbf{G}\|$	$4.1 \cdot 10^{0}$	$7.3 \cdot 10^{-1}$	$1.2 \cdot 10^{-1}$	$2.3 \cdot 10^{-1}$	$1.7 \cdot 10^{-2}$
It.	6	7	8	9	
$\|\mathbf{G}\|$	$8.5 \cdot 10^{-3}$	$1.0 \cdot 10^{-3}$	$1.8 \cdot 10^{-7}$	$8.8 \cdot 10^{-12}$	

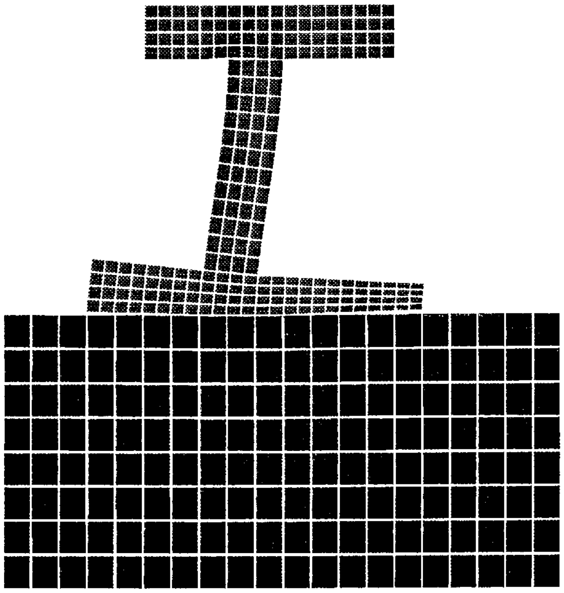
Table 2. Convergence rate at step 8

Λ and μ_1 are the Lame constants and α_1 denotes a dimensionless parameter. The problem is discretized by 4–node continuum elements which are based on the formulation given in Wriggers, Hueck (1995).

The total displacement is prescribed within 20 steps of $\Delta v = 0.1$. The typical convergence behaviour is reported in Table 2 for step number 8 ($v = 0.8$). One observes first no reduction in the norm of the residual $\|\mathbf{G}\|$ which is due to the search for the nodes being in contact. Once the correct set of contact nodes is found the algorithm convergences quadratically as depicted in the last three iteration steps.

Figure 9 shows different intermediate states of the deformation. Large sliding occurs in the contact interface which is captured here by the node–to–segment discretization.

The load deflection curve is depicted in Figure 10. First the load increases which is due to the stiffening of the system when the entire surface of the blade comes into contact with



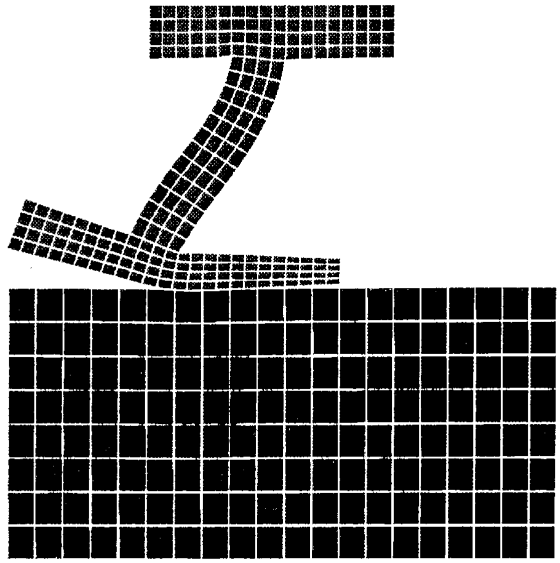


Figure 9. a) Deformed mesh at $v = 0.4$ b) Deformed mesh at $v = 0.8$

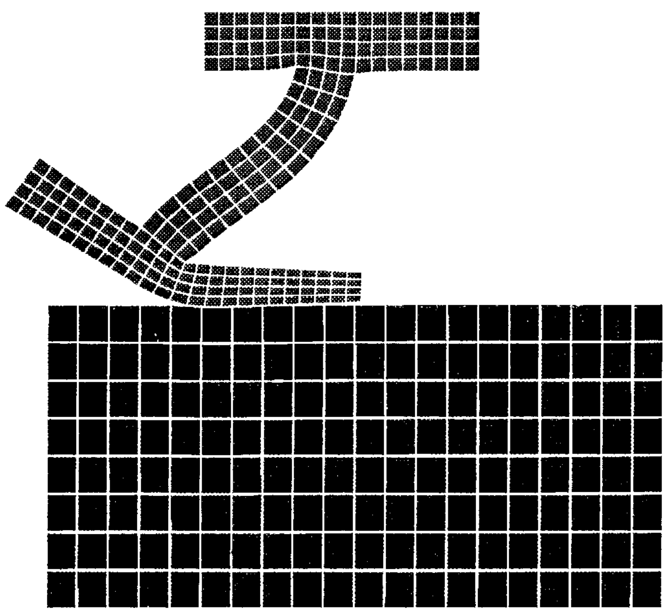


Figure 9. c) Deformed mesh at $v = 1.2$

the block, see Figure 9 a). Then a limit point occurs and after that the load decreases slightly which is associated with considerable sliding in the contact interface, see Figures 9 b) – d). Figure 9 d) also shows the distribution of the stresses σ_{yy} in vertical direction. Even with this relatively coarse mesh we observe a smooth distribution of the contact stresses.

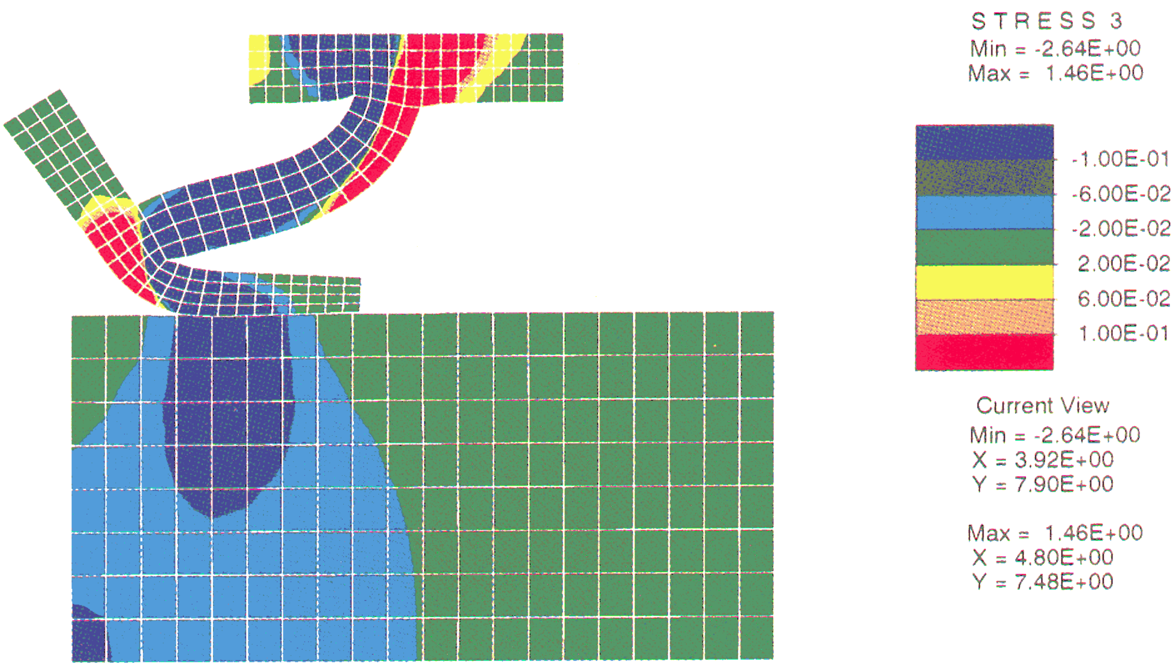


Figure 9. d) σ_{yy}–stresses at $v = 2.0$

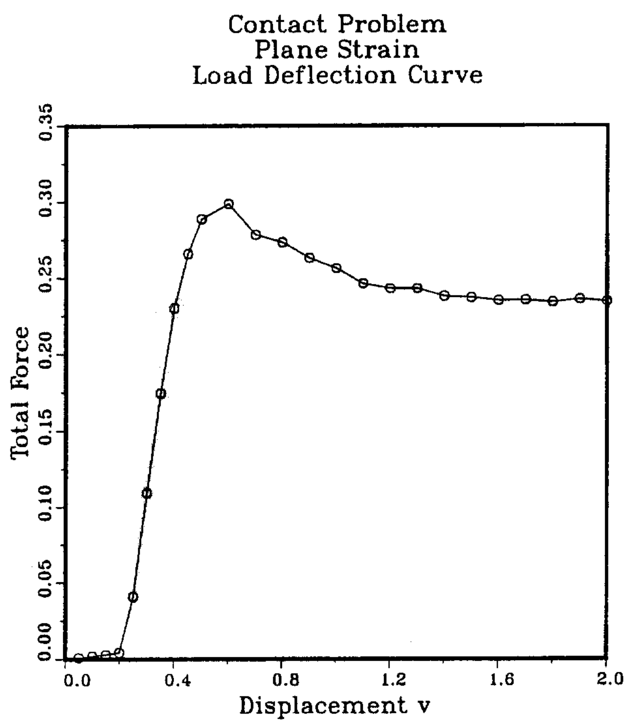


Figure 10. Load deflection curve

8.2 Frictional Contact of the Rubber Blade

To show the influence of the frictional behaviour in the contact interface the last example is now investigated using Coulomb's frictional law (19) with a frictional coeffiecient of $\mu = 0.4$. Again the load is applied in 20 steps of $\Delta v = 0.1$. The behaviour of the blade is now completely different from the frictionless solution, see Figure 11, where the deformations of several steps are shown. First a short tangential sliding takes place during the first deformation stages then blade sticks to the surface of the block. After that no essential horizontal movement occurs and finally buckling of the web is the result of the stick behaviour. The global behaviour of the contact process can also be seen by looking at the load–deflection curve in Figure 12. Up to the prescribed displacement $v = 0.4$ only slight differences to the frictionless solution can be observed. Then stiffening due to stick in the contact interface and the following buckling of the web, associated with the limit point and the decreasing load, are clearly visible.

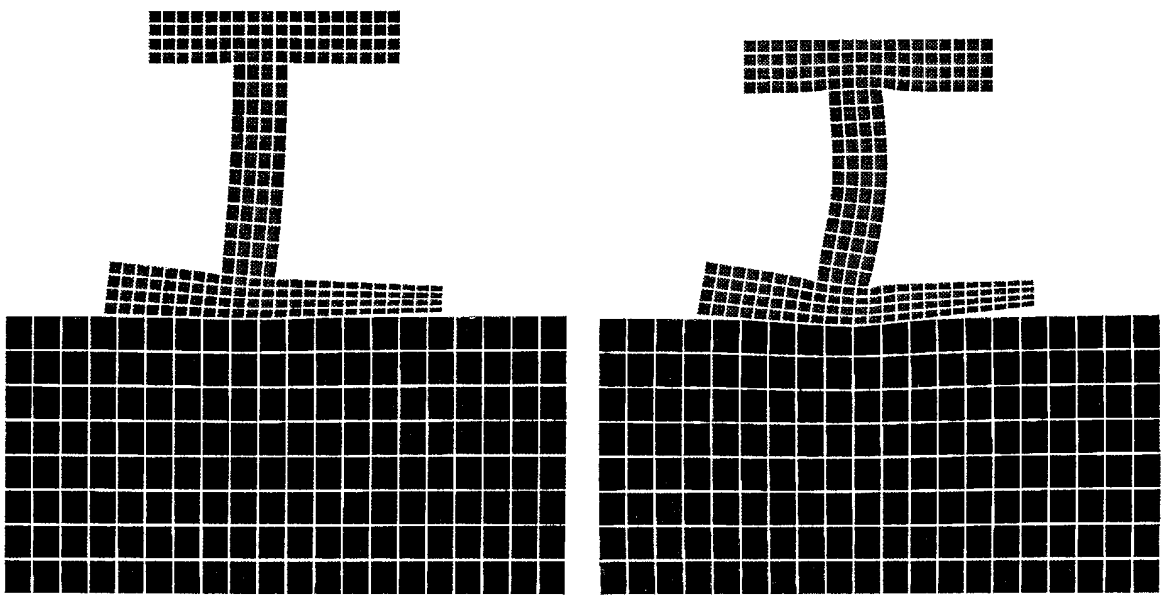


Figure 11. a) Deformed mesh at $v = 0.4$ b) Deformed mesh at $v = 0.8$

8.3 Adaptive Finite Element Solution of Hertzian Contact

In this example we apply the residual error estimator defined in equation (88) to solve the well known Hertzian problem of a disc contacting a planar surface. Since the analytical solution is known for this problem we can compare the results of the adaptive method directly to this solution. The problem is defined in Figure 13. To omit problems related to a point load in elasticity the load F is distributed over a small surface on top of the disc in the discrete finite element model.

Figure 14 depicts a series of meshes which are generated by the adaptive method described in section 7.3. The sixth mesh represents the converged solution. Finally the solution using the error estimator (88) are compared with results produced by the classical error estimator of Zienkiewicz, Zhu (1987) (without special treatment of the contact interface) and the analytical solution of the contact pressure. The associated numbers are contained in Table 3.

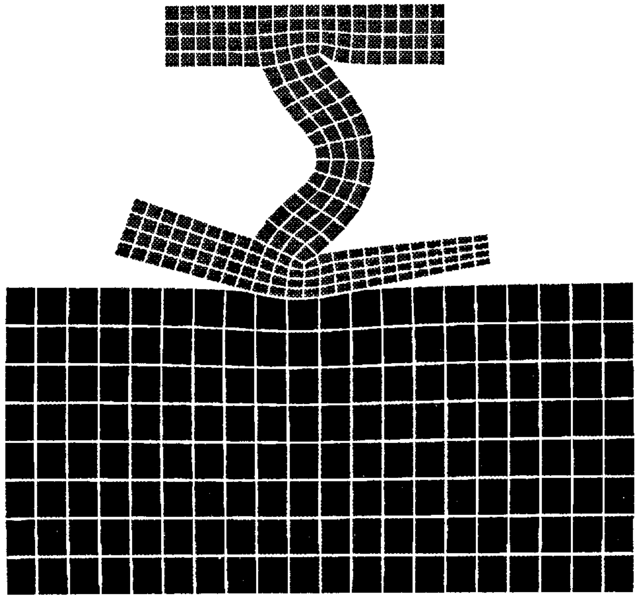


Figure 11. c) Deformed mesh at $v = 1.2$

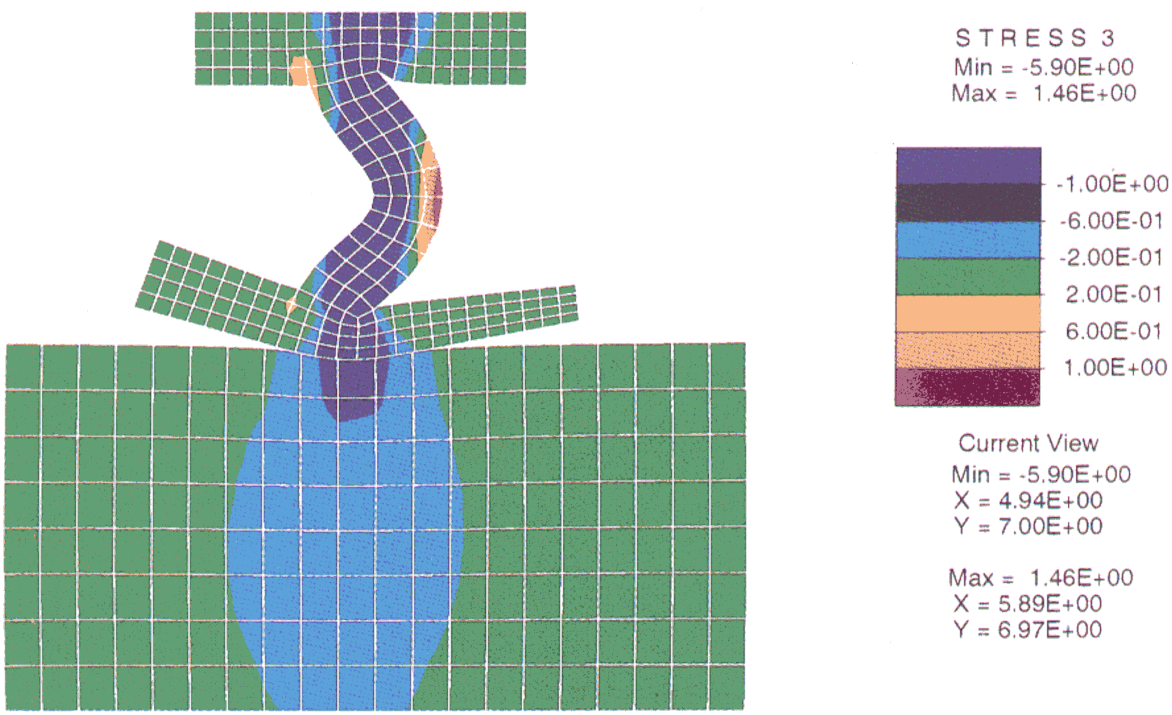


Figure 11. d) σ_{yy}–stresses at $v = 1.6$

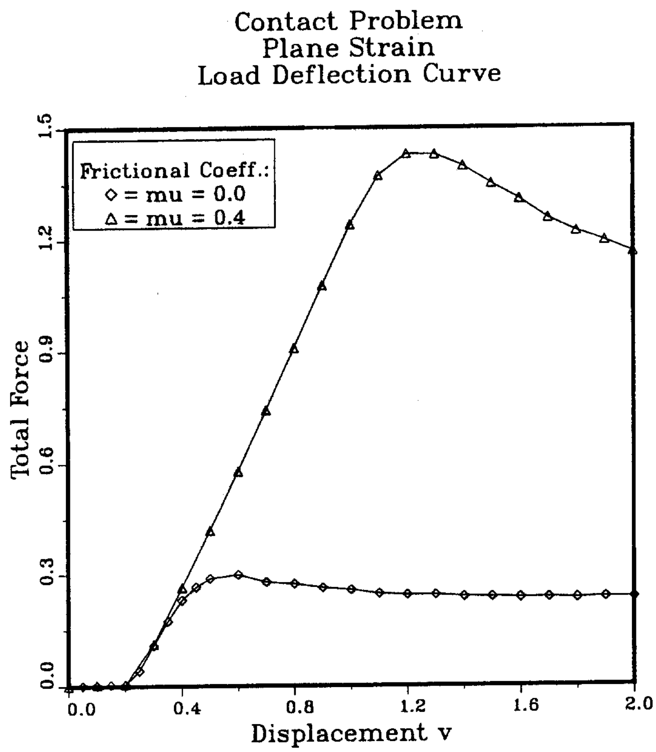


Figure 12. Load deflection curve for $\mu = 0$ and $\mu = 0.4$

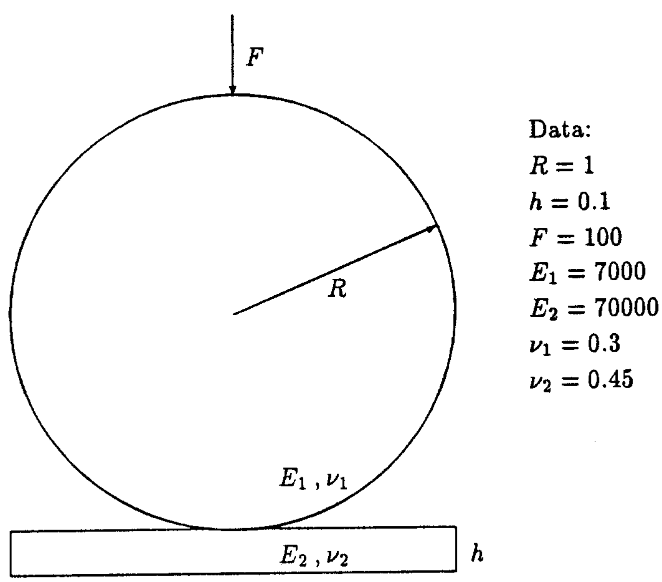


Figure 13. Hertzian contact problem

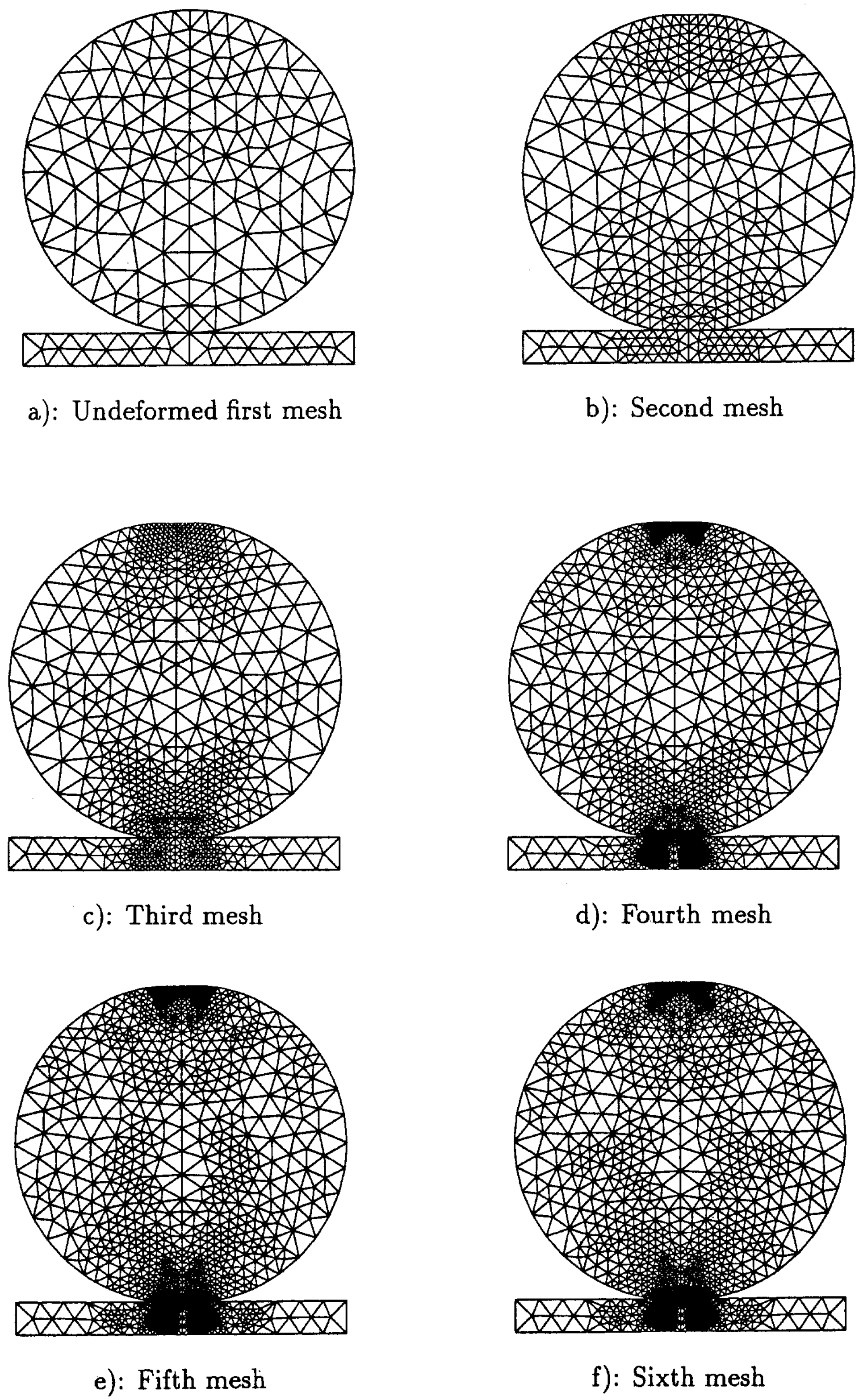


Figure 14. Series of meshes until convergence

	Error estimator (88) for contact problems	Error estimator Zienkiewicz, Zhu	Analytical solution
Final number of nodes	2035	2658	
Max. contact pressure	495	494	495

Table 3. Comparison of adaptive strategies

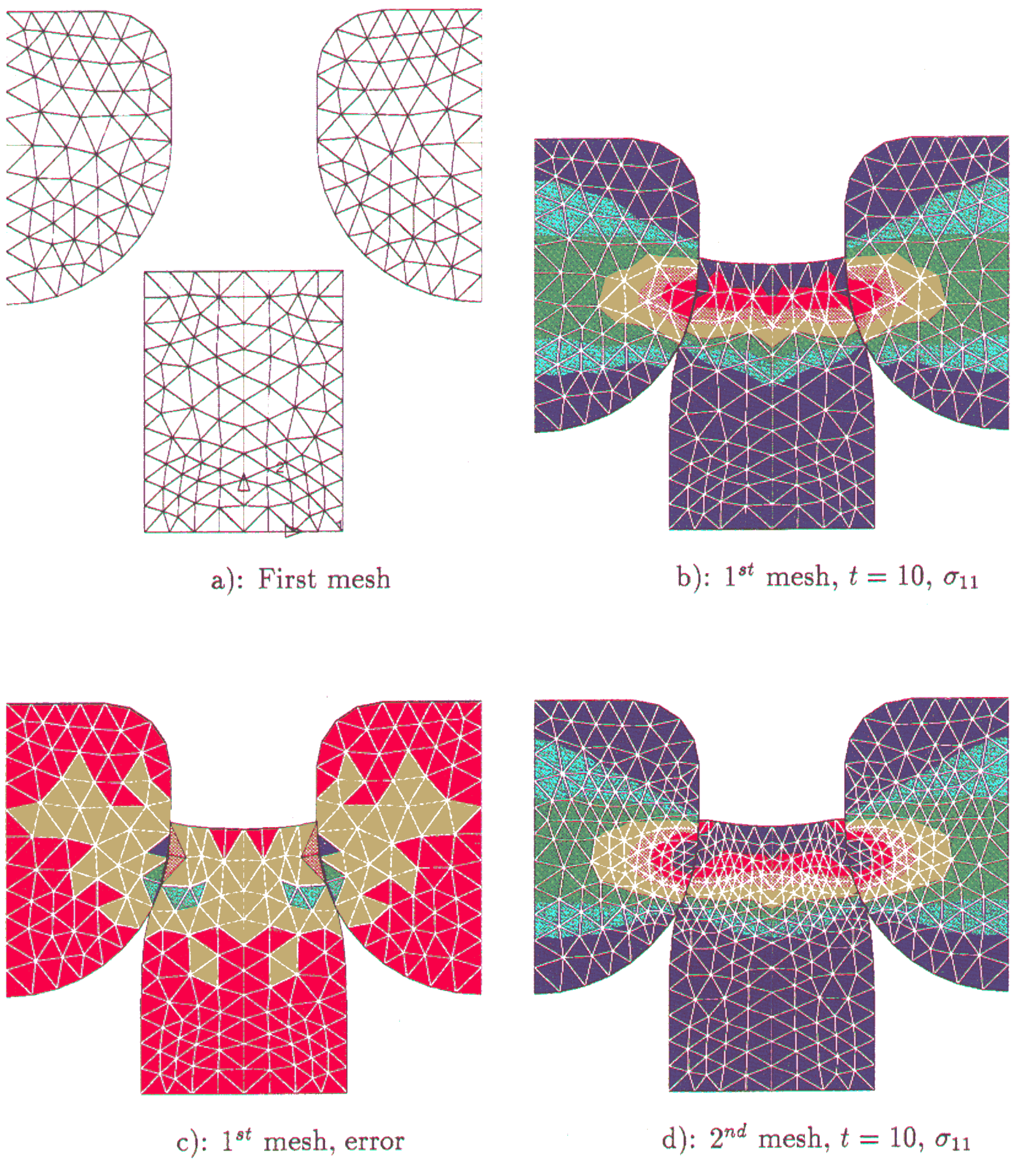
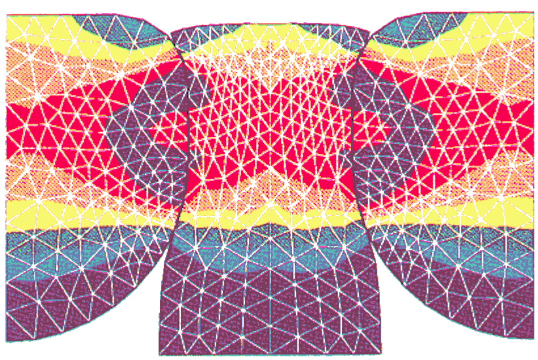


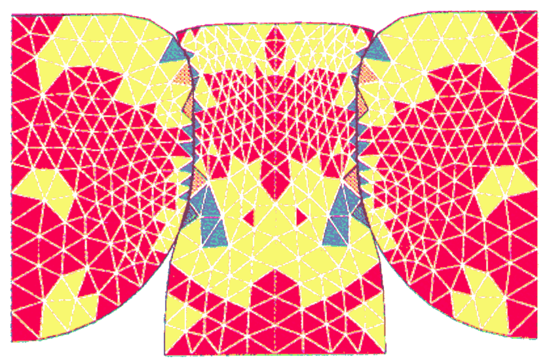
Figure 15. System, error distributions and stress contours

8.4 Adaptive Finite Element Solution of a Rubber Sealing

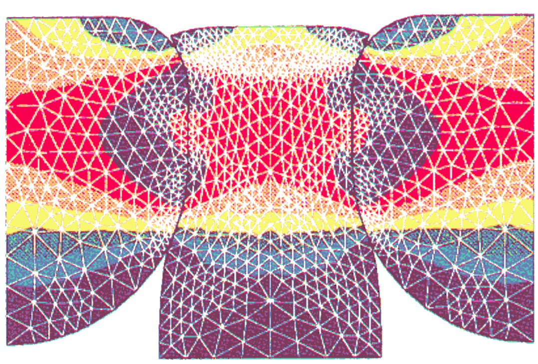
The final example is associated with large elastic strains. Here the same strain energy function as in example 8.1 describes the constitutive behaviour with $\alpha = 2$. Again the error estimator (88) is employed to solve the problem of a rubber seal. The adaptive technique is applied according following the discussion in remark V. The starting mesh is shown in Figure 15 a which depicts a rubber block (material data: $\Lambda_B = 100$, $\mu_B = 10$) which is pressed into the tool (material data: $\Lambda_T = 1000$, $\mu_T = 500$). The load is applied in a total of 20 steps. At load step 10 the first adaptive step is initiated. The error is shown in Figure 15 c and the resulting new mesh together with the contours of the stresses σ_{11} in direction 1 is depicted in Figure 15 d. One observes the evenly distributed mesh size due to the error estimator (88) within the contact area. Then this mesh is used to compute the solution until the load step of 16.5, see Figure 15 e. Again the error is computed, Figure 15 f, and the resulting mesh is shown in Figure 15 g. Despite relatively large errors near the contact zone in Figure 15 h, the stress distribution associated with this mesh, Figure 15 g, seems to be acceptable from an engineering point of view.



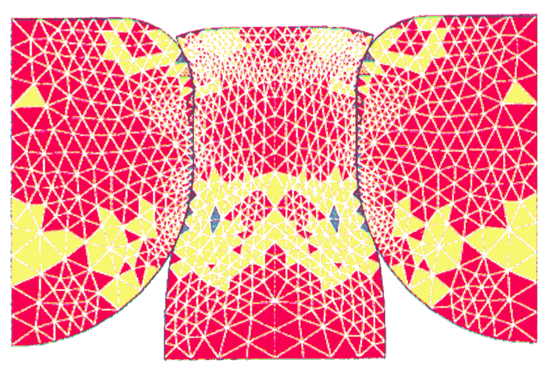
e): 2^{nd} mesh, $t = 16.5$, σ_{11}



f): 2^{nd} mesh, error



g): 3^{rd} mesh, $t = 16.5$, σ_{11}



h): 3^{rd} mesh, error

Figure 15. (*cont.*)

CONCLUSION

This overview summarizes some of the current research topics in computational contact mechanics. Due to the broadness of contact formulations and algorithms and the limitation of space not all promising new approaches have been discussed in detail. Also the reference list is far from being complete, but we hope to have included some of the main contributions during the last years. Numerical examples could have been presented for all mentioned topics. They have been omitted on purpose since this overview was aimed to provide the reader with the underlying theoretical derivations. These examples can be found in many of the cited papers.

ACKNOWLEDGEMENT

Part of this work was supported by the DFG (Deutsche Forschungsgemeinschaft) under contracts Wr19/4 and Wr19/8.

REFERENCES

1 Alart, P. and Curnier, A. (1991), "A Mixed Formulation for Frictional Contact Problems prone to Newton like Solution Methods", *Comp. Meth. Appl. Mech. Engng.*, **92**, 353–375.

2 Babuska, I. and Rheinboldt, W. (1978), "Error Estimates for Adaptive Finite Element Computations", *J. Num. Analysis*, **15**, 736–754.

3 Babuska, I., Strouboulis, T. Upadhyay, C.S., Gangaraj, S.K. and Copps, K. (1994), "Validation of A Posteriori Error Estimators by Numerical Approach", *Int. J. Num. Meth. Engng.*, **37**, 1073–1123.

4 Barthold, F.J. and Bischoff, D. (1988), "Generalization of Newton type Methods to Contact Problems with Friction", *J. Mec. Theor. Appl., Special Issue: Numerical Methods in Mechanics of Contact Involving Friction*, 97–110.

5 Bathe, K.J. and Chaudhary, A. B. (1985), "A Solution Method for Planar and Axisymmetric Contact Problems", *Int. J. Num. Meth. Engng.*, **21**, 65–88.

6 Bazaraa, M.S., Sherali, H.D. and Shetty, C.M. (1993), *Nonlinear Programming, Theory and Algorithms*, J. Wiley, New York.

7 Belytschko, T. and Neal, M.O. (1991), "Contact–Impact by the Pinball Algorithm with Penalty and Lagrangian Methods", *Int. J. Num. Meth. Engng.*, **31**, 547–572.

8 Bertsekas. D.P. (1984), *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York.

9 Björkman, G., Klarbring, A., Sjödin, B., Larsson, T. and Rönnqvist, M. (1995), "Sequential Quadratic Programming for Non–Linear Elastic Contact Problems", *Int. J. Num. Meth. Engng.*, **38**, 137–165.

10 Bowden, F.P. and Tabor, D. (1964), *The Friction and Lubrication of Solids*, Part II, Clarendon Press, Oxford.

11 Campos, L.T., Oden, J.T. and Kikuchi, N. (1982), "A Numerical Analysis of a Class of Contact Problems with Friction in Elastostatics", *Comp. Meth. Appl. Mech. Engng.*, **34**, 821–845.

12 Chan, S.H. and Tuba, I.S. (1971), "A Finite Element Method for Contact Problems in Solid Bodies", *Int. J. Mech. Sci*, **13**, 615–639.

13 Chaudhary, A.B. and Bathe, K.J. (1986), "A Solution Method for Static and Dynamic Analysis of Three-Dimensional Contact Problems with Friction", *Computer & Structures*, **24**, 137–147.

14 Conry, T.F. and Seireg, A. (1971), "A Mathematical Programming Method for Design of Elastic Bodies in Contact", *J. Appl. Mech.*, **38**, 1293–1307.

15 Ciarlet, P.G. (1988), *Mathematical Elasticity*, North Holland, Amsterdam.

16 Cooper, M.G., Mikic, B.B. and Yovanovich, M.M. (1969), "Thermal Contact Conductance.", *Int. J. of Heat and Mass Transfer*, **12**, 279-300.

17 Curnier, A. (1984), "A Theory of Friction", *Int. J. Solids Structures*, **20**, 637–647.

18 Curnier, A. and Alart, P. (1988), "A Generalized Newton Method for Contact Problems with Friction", *J. Mec. Theor. Appl., Special Issue: Numerical Methods in Mechanics of Contact Involving Friction*, 67–82.

19 Curnier, A., He, Q.C. and Telega, J.J. (1992), "Formulation of Unilateral Contact between two Elastic Bodies undergoing Finite Deformation", *C. R. Acad. Sci. Paris*, **314**, 1–6.

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

21 Eterovic, A. L. and Bathe, K.J. (1991), "An Interface Interpolation Scheme for Quadratic Convergence in the Finite Element Analysis of Contact Problems", in *Computational Methods in Nonlinear Mechanics*, eds. P. Wriggers, W. Wagner, Springer, Berlin.

22 Francavilla, A. and Zienkiewicz, O.C. (1975), "A Note on Numerical Computation of Elastic Contact Problems", *Int. J. Num. Meth. Engng.*, **9**, 913–924.

23 Fredriksson, B. (1976), "Finite Element Solution od Surface Nonlinearities in Structural Mechanics with Special Emphasis to Contact and Fracture Mechanics Problems", *Computer & Structures*, **6**, 281–290.

24 Giannokopoulos, A.E. (1989), "The Return Mapping Method for the Integration of Friction Constitutive Relations", *Computers & Structures*, **32**, 157–168.

25 Glowinski, R. and Le Tallec, P. (1984), "Finite Element Analysis in Nonlinear Incompressible Elasticity", in *Finite Element, Vol. V: Special Problems in Solid Mechanics*, eds. J. T. Oden and G. F. Carey, Prentice–Hall, Englewood Cliffs, New Jersey.

26 Hallquist, J.O., Goudreau, G.L. and Benson, D.J. (1985), "Sliding Interfaces with Contact–Impact in Large–Scale Lagrangian Computations", *Comp. Meth. Appl. Mech. Engng.*, **51**, 107–137.

27 Hallquist, J.O., Schweizerhof, K. and Stillman, D. (1992), "Efficiency Refinements of Contact Strategies and Algorithms in Explicit FE Programming", in *Proceedings of COMPLAS III*, eds. D.R.J. Owen, E. Hinton, E. E. Oñate, Pineridge Press.

28 Hansson, E. and Klarbring, A. (1990), "Rigid Contact Modelled by CAD Surface", *Eng. Computations*, **7**, 344–348.

29 Heegaard, J.–H. and Curnier, A. (1993), "An Augmented Lagrangian Method for Discrete Large–Slip Contact Problems", *Int. J. Num. Meth. Engng.*, **36**, 569–593.

30 Hertz, H. (1882), "Study on the Contact of Elastic Bodies", *J. Reine Angew. Math.*, **29**, 156–171.

31 Hlavacek, I., Haslinger, J., Necas, J. and Lovisek, J. (1988), *Solution of variational inequalities in mechanics*, Springer, New York.

32 Hughes, T.R.J., Taylor, R.L., Sackman, J.L., Curnier, A. and Kanoknukulchai, W. (1976), "A Finite Element Method for a Class of Contact–Impact Problems", *Comp. Meth. Appl. Mech. Engng.*, **8**, 249–276.

33 Hughes, T.R.J., Taylor, R.L. and Kanoknukulchai, W. (1977), "A Finite Element Method for Large Displacement Contact and Impact Problems", in *Formulations and Computational Algorithms in FE Analysis*, ed. K.J. Bathe, MIT–Press, Boston, 468–495.

34 Johannson, L. and Klarbring, A. (1992), "Thermoelastic Frictional Contact Problems: Modelling, Finite Element Approximation and Numerical Realization", preprint.

35 Johnson, C. (1987), *Numerical solutions of partial differential equations by the finite element method*, Cambridge Press, New York.

36 Johnson, C. and Hansbo, P. (1992), "Adaptive finite element methods in computational mechanics", *Comput. Meth. Appl. Mech. Engrg.*, **101**, 143–181.

37 Ju, W. and Taylor, R.L. (1988), "A Perturbed Lagrangian Formulation for the Finite Element Solution of Nonlinear Frictional Contact Problems", *Journal of Theoretical and Applied Mechanics*, **7**, 1–14.

38 Kikuchi, N. (1982), "A Smoothing Technique for Reduced Integration Penalty Methods in Contact Problems", *Int. J. Num. Meth. Engng.*, **18**, 343–350.

39 Kikuchi, N. and Oden, J.T. (1988), *Contact Problems in Elasticity: A Study of Variational Inequalities and Finite element Methods*, SIAM, Philadelphia.

40 Klarbring, A. (1986), "A Mathematical Programming Approach to Three–dimensional Contact Problems with Friction", *Comp. Meth. Appl. Mech. Engng.*, **58**, 175–200.

41 Klarbring, A. and Björkman, G. (1992), "Solution of Large Displacement Contact Problems with Friction using Newton's Method for Generalized Equations", *Int. J. Num. Meth. Engng.*, **34**, 249–269.

42 Kragelsky, I.V. (1956), *Die Entwicklung der Wissenschaft von der Reibung*, Verlag der Akademie der Wissenschaften der UdSSR, Moskau.

43 Kragelsky, I.V., Dobychin, M.N. and Kombalov, V.S. (1982), *Friction and Wear — Calculation Methods*, (Translated from The Russian by N. Standen), Pergamon Press.

44 Laursen, T.A. and Simo, J.C. (1991), "On the Formulation and Numerical Treatment of Finite Deformation Frictional Contact Problems", in *Computational Methods in Nonlinear Mechanics*, eds. P. Wriggers, W. Wagner, Springer, Berlin.

45 Laursen, T.A. and Simo, J.C. (1993a), "A Continuum–Based Finite Element Formulation for the Implicit Solution of Multibody, Large Deformation Frictional Contact Problems", *Int. J. Num. Meth. Engng.*, **36**, 3451–3485.

46 Laursen, T.A. and Simo, J.C. (1993b), "Algorithmic Symmetrization of Coulomb Frictional Problems using Augmented Lagrangians", *Comp. Meth. Appl. Mech. Engng.*, **108**, 133–146.

47 Lee, C.Y., Oden, J.T. and Ainsworth, M. (1991), "Local A Posteriori Error Estimates and Numerical Results for Contact Problems and Problems of Flow through Porous Media", in: *Nonlinear Computational Mechanics*, eds. P. Wriggers and W. Wagner, 671–689, Springer, Berlin.

48 Luenberger, D.G. (1984), *Linear and Nonlinear Programming*, Addison–Wesley, Massachusetts.

49 Michalowski, R. and Mroz, Z. (1978), "Associated and Non–associated Sliding Rules in Contact Friction Problems", *Arch. of Mechanics*, **30**, 259–276.

50 Munjiza, A., Owen, D.R.J. and Bicanic, N. (1995), "A Combined Finite–Discrete Element Method in Transient Dynamics of Fracturing Solids", *Eng. Computations* **12**, 145–174.

51 Nour-Omid, B. and Wriggers, P. (1986), "A Two–Level Iterative Method for the Solution of Contact Problems", *Comp. Meth. Appl. Mech. Engng.*, **54**, 131–144.

52 Nour-Omid, B. and Wriggers, P. (1987), "A Note on the Optimum Choice for Penalty Parameters", *Comm. Appl. Num. Meth*, **3**, 581–585.

53 Oden, J.T. (1981), "Exterior Penalty Methods for Contact Problems in Elasticity", in *Nonlinear Finite Element Analysis in Structural Mechanics*, eds, W. Wunderlich, E. Stein, K. J. Bathe, Springer, Berlin.

54 Oden, J.T. and Martins, J.A.C. (1986), "Models and Computational Methods for Dynamic Friction Phenomena", *Comp. Meth. Appl. Mech. Engng.*, **52**, 527–634.

55 Oden, J.T. and Pires, E.B. (1983a), "Algorithms and Numerical Results for Finite Element Approximations of Contact Problems with Non-Classical Friction Laws", *Computer & Structures*, **19**, 137–147.

56 Oden, J.T., Pires, E.B. (1983b), "Nonlocal and Nonlinear Friction Laws and Variational Priciples for Contact Problems in Elasticity", *J. Appl. Mech.*, **50**, 67–76.

57 Papadopoulos, P. and Taylor, R.L. (1992), "A Mixed Formulation for the Finite Element Solution of Contact Problems", *Comp. Meth. Appl. Mech. Engng.*, **94**, 373–389.

58 Parisch, H. (1989), "A Consistent Tangent Stiffness Matrix for Three–Dimensional Non–Linear Contact Analysis", *Int. J. Num. Meth. Engng.*, **28**, 1803–1812.

59 Simo, J.C., Wriggers, P. and Taylor, R.L. (1985), "A Perturbed Lagrangian Formulation for the Finite Element Solution of Contact Problems", *Comp. Meth. Appl. Mech. Engng.*, **50**, 163–180.

60 Simo, J.C. and Taylor, R.L. (1985), "Consistent Tangent Operators for Rate-independant Elastoplasticity", *Comp. Meth. Appl. Mech. Engng.*, **48**, 101–118.

61 Simo, J.C. and Laursen, T.A. (1992), "An Augmented Lagrangian Treatment of Contact Problems involving Friction", *Computers & Structures*, **42**, 97–116.

62 Simo, J.C. and Miehe, C. (1992), "Associative Coupled Thermoplasticity at Finite Strains: Formulation, Numerical Analysis and Implementation", *Comp. Meth. Appl. Mech. Engng.*, **98**, 41–104.

63 Song, S. and Yovanovich, M.M. (1987), "Explicit Relative Contact Pressure Expression: Dependence upon Surface Roughness Parameters and Vickers Microhardness Coefficients.", *AIAA* Paper 87-0152.

64 Stadter, J.T. and Weiss, R.O. (1979), "Analysis of Contact through Finite Element Gaps", *Computers & Structures*, **10**, 867–873.

65 Tabor, D. (1981), "Friction — The Present State of Our Understanding", *J. Lubr. Technology*, **103**, 169–179.

66 Verführt, R. (1993), "A Review of a posteriori error estimation and adaptive mesh–refinement techniques", Technical Report, Institut für Angewandte Mathematik, Universität Zürich.

67 Woo, K.L. and Thomas, T.R. (1980), "Contact of Rough Surfaces : A Review of Experimental Works", Wear, **58**, 331–340.

68 Williams, J.R. and Pentland, A. P. (1992), "Superquadrics and Modal Dynamics for Discrete Elements in Interactive Design", *Eng. Computations*, **9**, 115–127.

69 Williams, J.R. and O'Connor R. (1995), "A Linear Complexity Intersection Algorithm for Discrete Element Simulation of Arbitrary Geometries", *Eng. Computations*, **12**, 185–201.

70 Wilson, E.A. and Parsons, B. (1970), "Finite Element Analysis of Elastic Contact Problems using Differential Displacements", *Int. J. Num. Meth. Engng.*, **2**, 387–395.

71 Wriggers, P. and Simo, J.C. (1985), "A Note on Tangent Stiffness for Fully Nonlinear Contact Problems", *Communications in Applied Numerical Methods*, **1** , 199–203.

72 Wriggers, P., Simo, J.C. and Taylor, R.L. (1985), "Penalty and Augmented Lagrangian Formulations for Contact Problems", in *Proceedings of NUMETA 85 Conference*, eds. J. Middleton & G.N. Pande, Balkema, Rotterdam.

73 Wriggers, P., Wagner, W. and Stein, E. (1987), "Algorithms for Nonlinear Contact Constraints with Application to Stability Problems of Rods and Shells", *Computational Mechanics*, **2**, 215–230.

74 Wriggers, P. (1987), "On Consistent Tangent Matrices for Frictional Contact Problems", in *Proceedings of NUMETA 87 Conference*, eds. J. Middleton, G.N. Pande, Nijhoff, Dorbrecht.

75 Wriggers, P. and Wagner, W. (1988), "A Solution Method for the Postcritical Analysis of Contact Problems", in *The Mathematics of Finite Elements and Applications VI*, Proceedings of MAFLEAP 87, ed. J. Whiteman, Academic Press, London.

76 Wriggers, P., Vu Van, T. and Stein, E. (1990), "Finite Element Formulation of Large Deformation Impact-Contact Problems with Friction", *Computers & Structures*, **37**, 319–331.

77 Wriggers, P. and Miehe, C. (1992), "Recent Advances in the Simulation of Thermomechanical Contact Processes", in *Proceedings of COMPLAS III*, eds. D.R.J. Owen, E. Hinton, E. E. Onate, Pineridge Press.

78 Wriggers, P. (1993), "Continuum Mechanics, Nonlinear Finite Element Techniques and Computational Stability", in *Progress in Computational Analysis of Inelastic Structures*, ed. E. Stein, Springer, Wien.

79 Wriggers, P., Zavarise, G. (1993), "On the Application of Augmented Lagrangian Techniques for Nonlinear Constitutive Laws in Contact Interfaces", *Comm. Num. Meth. Engng.* **9**, 815–824, 1993.

80 Wriggers, P. and Imhof, M. (1993), "On the Treatment of Nonlinear Unilateral Contact Problems", *Ing. Archiv*, **63**, 116–129.

81 Wriggers, P. and Zavarise, G. (1993), "Thermomechanical Contact – A Rigorous but Simple Numerical Approach", *Computers & Structures*, **46**, 47–53.

82 Wriggers, P. and Miehe C. (1994), "Contact Constraints within Coupled Thermomechanical Analysis – A Finite Element Model", *Comp. Meth. Appl. Mech. Engng.*, **113**, 301–319.

83 Wriggers, P., Scherf, O. and Carstensen, C. (1994), "Adaptive Techniques for the Contact of Elastic Bodies", in *Recent Developments in Finite Element Analysis*, eds. T.J.R. Hughes, E. Oñate, O.C. Zienkiewicz, CIMNE, Barcelona.

84 Wriggers, P. and Scherf, O. (1995), "An adaptive finite element method for elastoplastic contact problems", in *Proceedings of COMPLAS 4*, eds. R.D. Owen, E. Hinton, E. Oñate, Pineridge Press, Swansea.

85 Zavarise, G. (1991), "Problemi termomeccanici di contatto – aspetti fisici e computazionali", *Ph.D. Thesis*, Ist. di Scienza e Tecnica delle Costruzioni, Univ. of Padua, Italy.

86 Zavarise, G., Schrefler, B.A. and Wriggers, P. (1992), "Consistent Formulation for Thermomechanical Contact based on Microscopic Interface Laws", in *Proceedings of COMPLAS III*, eds. D.R.J. Owen, E. Hinton, E. E. Oñate, Pineridge Press.

87 Zavarise, G., Wriggers, P., Stein, E. and Schrefler, B.A. (1992a), "Real Contact Mechanisms and Finite Element Formulation – A Coupled Thermomechanical Approach", *Int. J. Num. Meth. Engng.*, **35**, 767–786, 1992.

88 Zavarise, G., Wriggers, P., Stein, E. and Schrefler, B.A. (1992b), "A Numerical Model for Thermomechanical Contact based on Microscopic Interface Laws", *Mech. Res. Comm.*, **19**, 173–182.

89 Zavarise, G. and Wriggers, P. (1995), "Elastoplastic Contact Problems Solved by the Cross–Constraint Method", in *Proceedings of COMPLAS IV*, eds. D.R.J. Owen, E. Hinton, E. E. Oñate, Pineridge Press.

90 Zavarise, G., Wriggers, P. and Schrefler, B.A. (1995a), "On Augmented Lagrangian Algorithms for Thermomechanical Contact Problems with Friction", to appear in IJNME.

91 Zavarise, G., Wriggers, P. and Schrefler, B.A. (1995b), "A New Method for Solving Contact Problems", submitted to IJNME.

92 Zienkiewicz, O.C. and Zhu, J.Z. (1987), "A Simple Error Estimator and Adaptive Procedure for Practical Engineering Analysis", *Int. J. Num. Meth. Engrg.*, **24**, 337–357.

93 Zienkiewicz, O.C. and Taylor, R.L. (1989), *The Finite Element Method*, 4th edn., Mc Graw–Hill, London.

94 Zhong, Z.-H. and Nilsson, L. (1989), "A Contact Searching Algorithm for General Contact Problems", *Computers & Structures*, **33**, 197–209.

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

Please address your comments or questions on this paper to:
International Center for Numerical Methods in Engineering
Edificio C-1, Campus Norte UPC
Gran Capitán s/n
08034 Barcelona, Spain
Phone: 34-3-4016035; Fax: 34-3-4016517