A computational contact formulation based on surface potentials

Roger A. Sauer \(^1\) and Laura De Lorenzis \(^2\)

\(^1\)Graduate School AICES, RWTH Aachen University, Templergraben 55, 52056 Aachen, Germany; sauer@aices.rwth-aachen.de

\(^2\)Università del Salento; via per Monteroni, 73100 Lecce, Italy; laura.delorenzis@unisalento.it


Submitted on 21 March 2012, Revised on 20 July 2012, Accepted on 2 September 2012

Abstract: This work presents the theory and numerical implementation of a contact formulation based on surface potentials. The new theory formulates contact based on distance-dependent surface interaction potentials and distinguishes between three interaction classes: Point interaction, short-range surface interaction and long-range surface interaction. Here the focus is placed on frictionless contact, although the first class readily admits frictional sticking contact as is also shown. The proposed contact theory provides a unified framework for various contact formulations, ranging from numerical constraint formulations, like penalty, barrier, cross-constrained and augmented Lagrange multiplier methods, to physical interaction formulations, like cohesive zone models, as well as electrostatic, gravitational and van-der-Waals interaction. Apart from recovering classical penalty and barrier formulations, the new theory also naturally leads to a modified penalty and barrier method. The formulation also recovers classical one-pass contact algorithms, however the real advantages lie in a novel two-pass contact algorithm, denoted the two-half-pass contact algorithm, since each pass only accounts for the contact forces acting on the slave body. This implies that traction continuity is only satisfied in theory, but not imposed a priori in the algorithm. Instead, it is obtained naturally to high accuracy as is demonstrated by several 2D and 3D numerical examples. These include sliding contact, peeling contact and electrostatic attraction between deformable solids. Among the examples is a detailed comparison between the new formulation and classical one-pass approaches. It is further shown that the new contact formulation passes the contact patch test.

Keywords: computational contact mechanics, surface potentials, adhesion, electrostatics, nonlinear finite element methods, contact patch test

List of important symbols

\(a_{\alpha}\) co-variant surface tangent vectors in the current configuration; \(\alpha = 1, 2\)

\(a_{\alpha}^p\) \(a_{\alpha}\) evaluated at \(x_p\)

\(\hat{a}_{\alpha}^p\) normalized tangent vector \(a_{\alpha}^p\)

\(\hat{a}_p^\alpha\) orthonormal tangent vectors along the directions of the principal surface curvatures

\(A_{\alpha}\) co-variant surface tangent vectors in the reference configuration

\(\beta_{\alpha}\) density of surface interactions points in the current configuration

\(\beta_{\alpha}^0\) density of surface interactions points in the reference configuration

\(^1\)Corresponding author, email: sauer@aices.rwth-aachen.de

\(^2\)This pdf is the personal version of an article whose final publication is available at www.sciencedirect.com
\(B_k\) current configuration of body \(k\)
\(B_{0k}\) reference configuration of body \(k\)
\(\partial B_k\) surface of \(B_k\)
\(B_k^h\) discretization of body \(B_k\)
\(\Gamma^e_k\) finite element on \(\partial B_k^h\)
\(d\alpha_k\) differential surface element on \(\partial B_k\)
\(dA_k\) differential surface element on \(\partial B_{0k}\)
\(\delta\) Dirac delta function
\(\delta\varphi_k\) virtual deformation of body \(k\)
\(\delta\Pi_{c_k}\) virtual contact work of body \(k\)
\(E\) Young’s modulus
\(E_k\) electric field at \(x_k\) due to \(\Phi_\ell\)
\(\epsilon_n\) penalty parameter
\(f_s\) magnitude of the interaction force at \(x_k\) due to \(\phi_s\)
\(F_v\) (volume) deformation gradient
\(F_s\) surface deformation gradient
\(F_k\) interaction force at \(x_k\) due to \(\Phi_\ell\)
\(f_c\) contact force vector of finite element \(\Gamma^e_k\)
\(f_c\) contact force vector for the entire system
\(g_k\) signed distance from \(x_\ell\) to \(x_k\)
\(g_n\) signed normal gap from \(\partial B_k\) to \(x_k\)
\(J_{sk}\) surface stretch of body \(k\)
\(J_p\) \(J_{sd}\) evaluated at \(x_p \in \partial B_\ell\)
\(k\) index for the body under consideration, i.e. the slave body; \(k = 1, 2\)
\(\kappa_\alpha\) principal surface curvatures
\(k_c\) FE tangent matrix associated with \(f_c\)
\(\ell\) index for the neighboring body, i.e. the master body; \(\ell = 1, 2; \ell \neq k\)
\(\lambda\) stretch measure
\(n_k\) direction vector associated with \(g_k\)
\(n_p\) surface normal of \(\partial B_\ell\) at \(x_p\)
\(n_{el}\) number of finite elements
\(n_{el}^c\) number of nodes per surface element \(\Gamma^e_k\)
\(N_I\) shape function of finite element node \(I\)
\(N_k\) array of the shape functions for element \(\Gamma^e_k\)
\(\nu\) Poisson’s ratio
\(\xi^\alpha\) convective surface coordinates of \(\partial B_\ell\)
\(\xi_p^\alpha\) value of \(\xi^\alpha\) at \(x_p\)
\(p\) contact pressure
\(P_k\) net interaction force acting on \(B_k\) due to \(\Pi_c\)
\(q_k\) electrical charge at \(x_k\)
\(r_k\) distance vector from \(x_\ell\) to \(x_k\)
\(r_k\) absolute distance between \(x_\ell\) and \(x_k\)
\(\bar{r}_k\) normalization of \(r_k\)
\(\bar{r}_n\) absolute distance between \(x_p\) and \(x_k\)
\(\bar{r}_n\) normalization of \(r_n\)
\(\sigma\) Cauchy stress tensor
\(t_k\) physical contact surface traction acting at \(x_k\)
\(T_k\) traction \(t_k\) scaled by \(J_{sk}\)
\(t_k, T_k\) magnitudes of \(t_k\) and \(T_k\)
1 Introduction

This work presents a new contact formulation that unifies a large range of different interaction models. The new formulation is based on surface potentials and is therefore denoted the surface-potential-based-contact (SPBC) formulation. Contact formulations based on potentials have a long tradition in contact mechanics, e.g. see Laursen (2002); Wriggers (2006). The usual approach is to formulate the contact energy as an integral over the common contact surface. In a computational setting, the surfaces in contact are, in general, not identical due to discretization error. Therefore the question arises which surface to consider. Choosing either surface of the two contacting bodies – the classical one-pass approach – leads to a biased, unsymmetrical formulation. To circumvent this bias, we propose a potential-based formulation that uses the double integration over both surfaces. This approach is motivated by van-der-Waals adhesion and electrostatic interaction, which appear as double integrals.

The new model is formulated in the general framework of large deformations and arbitrary surface geometries. Further, the new approach provides a unified framework for numerical constraint formulations – like penalty, barrier, cross-constrained and augmented Lagrange multiplier methods (Wriggers, 2006) – and physical interaction phenomena – like adhesion (Raous et al., 1999; Sauer and Li, 2007b,a; Del Piero and Raous, 2010), cohesive-zone models (Xu and Needleman, 1993; Tvergaard, 2003) and electrostatic interactions (Shadowitz, 1988). In order to incorporate all these cases into the SPBC formulation, we distinguish between three interaction classes: long-range surface-, short-range surface- and point interactions. The proposed approach is related to the formulation of Argento et al. (1997), which uses double surface integration to describe molecular interaction in the context of small deformations.

All classes have in common that they lead naturally to an unbiased two-pass contact algorithm. Since each pass naturally only accounts for half of the contact forces (i.e. only for those acting on one of the two surfaces) the proposed algorithm is termed the two-half-pass contact algorithm. In the case of class P interaction, a conceptually similar contact algorithm was proposed by Papadopoulos and co-workers (Papadopoulos et al., 1995; Papadopoulos and Solberg, 1998; Jones and Papadopoulos, 2001; Solberg et al., 2007). The major differences are that our algorithm is derived from a contact potential and is more general, as it encompasses a large range of physical and numerical contact-interaction models. Also, traction continuity is not imposed a priori. It is rather obtained automatically as is shown. A further difference is that in our approach a different definition of the gap between the interacting bodies is considered. In the
current form, our approach is purely displacement based.
The current approach is based on a variational structure and is derived from an (elastic) potential. Therefore, the contact formulation is fully reversible and no dissipation, due to sliding friction or contact wear, is considered. Currently, the point interaction formulation contains frictional sticking contact, while the formulations for short- and long-range interactions are frictionless. The extension to general, frictional contact is subject of future work.

The remainder of this work is organized as follows. Sec. 2 presents the theory of the SPBC formulation for the three interaction classes. It is shown that the formulation satisfies contact equilibrium. In Sec. 3, eleven different contact-interaction formulations are discussed that fit into the framework of the SBPC formulation. In Sec. 4 the proposed formulation is further illustrated by a simple thought experiment and an analytically solved contact example. In a first reading, the theory of Secs. 2, 3 and 4 may be skipped in favor of the summary in Sec. 2.6. Sec. 5 presents the two-half-pass contact algorithm for the three interaction classes. The proposed algorithm satisfies the contact patch test as is shown in Sec. 6. Numerical examples then follow in Sec. 7. Sec. 8 concludes the paper.

2 A surface potential formulation for contact

2.1 Variational framework for frictionless contact

In this work, we consider conservative systems. Their mechanical behavior can be derived from a potential invoking the principle of stationary potential energy for quasi-static problems, or Hamilton’s principle for time-dependent problems (Sauer and Li, 2007b). For a system involving contact, the potential energy can be expressed as

$$\Pi = \Pi_{\text{int}} + \Pi_{\text{c}} - \Pi_{\text{ext}},$$

(1)

where $\Pi_{\text{int}}$ and $\Pi_{\text{ext}}$ are potentials associated with the internal and external forces of the system, and $\Pi_{\text{c}}$ is a potential that is associated with the contact forces. The weak form governing the contact problem is found by computing the variation of $\Pi$, denoted $\delta \Pi$. For quasi-static problems the weak form follows from setting $\delta \Pi = 0$ for all variations of the mechanical deformation $\varphi$, i.e. we have

$$\delta \Pi = \delta \Pi_{\text{int}} + \delta \Pi_{\text{c}} - \delta \Pi_{\text{ext}} = 0 \quad \forall \delta \varphi \in \mathcal{V},$$

(2)

where $\mathcal{V}$ is a suitable function space for $\delta \varphi$. In the following section an expression for $\Pi_{\text{c}}$ is formulated that is based on a surface potential. The approach is motivated by the description of van-der-Waals interaction (Sauer and Li, 2007b; Sauer and Wriggers, 2009).

2.2 Surface potential

Consider the configuration of two bodies, $B_1$ and $B_2$,\(^3\) where all surface elements $da_1 \subset \partial B_1$ and $da_2 \subset \partial B_2$, separated by the distance $r$, interact with each other via the surface interaction potential $\phi_s(r)$, such that the global interaction energy is given by the integration

$$\Pi_{\text{c}} = \int_{\partial B_1} \int_{\partial B_2} \beta_1^s \beta_2^s \phi_s da_2 da_1,$$

(3)

see Fig. 1. The potential $\phi_s$ may be defined from the physics of the interaction or, as will

\(^3\)Following Truesdell and Noll (2004), for cases where no confusion should occur, we simply write ‘body $B_k$’ instead of ‘the configuration of body $k$’.
also be considered here, as a purely numerical quantity. Three classes of interactions can be derived from $\Pi_c$, as will be discussed in detail in Sec. 2.3. In eqs (3), $\beta^s_k (k = 1, 2)$ denotes the current surface density of body $B_k$, measured in the units [length$^{-2}$]. We assume that the surface material in any surface region is conserved during deformation, i.e. the surface material is assumed to remain on the surface during deformation, while the bulk material is assumed to remain in the bulk. In other words, we consider the deformation of the surface but not its creation or consumption. This is a reasonable assumption for solids, but it is not realistic for fluids, which require a more general framework. Under the above assumption the current surface density $\beta^s_k$ is related to the reference surface density $\beta^s_{0k}$ by the relation

$$\beta^s_k = J^{-1\,s}_{sk} \beta^s_{0k} \tag{4}$$

where

$$J_{sk} := \frac{\partial a_k}{\partial A_k}, \tag{5}$$

denotes the surface stretch during deformation, i.e. it defines the relation

$$da_k = J_{sk} dA_k, \tag{6}$$

between the area elements $dA_k \subset \partial B_k$ and $dA_k \subset \partial B_{0k}$. The deformation of the two bodies is described by the mapping $x_k = \varphi_k(X_k)$ of a material point originally located at $X_k$ within the (undeformed) reference configuration $B_{0k}$ to the new position $x_k$ within the (deformed) current configuration $B_k$. According to eqs (4) and (6) the quantity $\beta^s_{sk} da_k$ is conserved during deformation, i.e.

$$\beta^s_k da_k = \beta^s_{0k} dA_k = \text{const.}, \tag{7}$$

which is an important relation for rewriting integrals and obtaining derivatives as will be seen later. This implies that the surface interaction $\phi_s$ between $d\alpha_1$ and $d\alpha_2$ is invariant under the deformation (but not the separation!) of the surfaces. It then follows that expression (3) is equivalent to

$$\Pi_c = \int_{\partial B_{01}} \int_{\partial B_{02}} \beta^s_{01} \beta^s_{02} \phi_s dA_2 dA_1. \tag{8}$$

The potential $\phi_s(r)$ depends on the position $x_1$ and $x_2$ of surface elements $d\alpha_1$ and $d\alpha_2$, so that we can also write $\phi_s = \phi_s(x_1, x_2)$. If the deformations $\varphi_1$ and $\varphi_2$ are varied by $\delta \varphi_1$ and $\delta \varphi_2$,

---

4See Appendix D for further details.
potential $\phi_s$ varies by\textsuperscript{5}

$$\delta \phi_s = \frac{\partial \phi_s}{\partial x_1} \cdot \delta \varphi_1 + \frac{\partial \phi_s}{\partial x_2} \cdot \delta \varphi_2 = \delta \phi_{s1} + \delta \phi_{s2}.$$  

\text{(9)}

The variation of the contact energy $\Pi_c$ thus becomes

$$\delta \Pi_c = \delta \Pi_{c1} + \delta \Pi_{c2},$$  

\text{(10)}

where the two contributions are given by

$$\delta \Pi_{ck} = \int_{\partial B_01} \int_{\partial B_{02}} \beta_{01}^{s} \beta_{02}^{s} \frac{\partial \phi_s}{\partial x_k} \cdot \delta \varphi_k \, dA_2 \, dA_1,$$  

\text{(11)}

with $k = 1, 2$. No summation is implied on repeated index $k$. Due to eq. (7) we can equivalently write

$$\delta \Pi_{ck} = \int_{\partial B_1} \int_{\partial B_2} \beta_{1}^{s} \beta_{2}^{s} \frac{\partial \phi_s}{\partial x_k} \cdot \delta \varphi_k \, da_2 \, da_1.$$  

\text{(12)}

In order to characterize the gradient of $\phi_s$, we define the distance vector, as well as its magnitude and direction, between a point $x_k$ on $\partial B_k$ and a point $x_\ell$ on the neighboring surface $\partial B_\ell$ ($\ell = 1, 2; \ell \neq k$) as

$$r_k := x_k - x_\ell, \quad r_k := \| r_k \|, \quad \bar{r}_k := \frac{r_k}{r_k}.$$  

\text{(13)}

To distinguish between the cases of penetration and separation, we introduce the signed distance

$$g_k := \begin{cases} -r_k & \text{if } x_k \in B_\ell \text{ and } x_\ell \in B_k, \\ r_k & \text{else}, \end{cases}$$  

\text{(14)}

and the vector

$$n_k := \begin{cases} -\bar{r}_k & \text{if } x_k \in B_\ell \text{ and } x_\ell \in B_k, \\ \bar{r}_k & \text{else}, \end{cases}$$  

\text{(15)}

as they are also shown in Fig. 2. The distance is always taken positive if at least one of the interacting points $x_1$ and $x_2$ is outside the neighboring body.\textsuperscript{6} Note, that we can also write

$$n_k = (x_k - x_\ell) / g_k,$$  

\text{(16)}

for all cases. According to definitions (13)–(15) we can write

$$r_\ell = -r_k, \quad r_\ell = r_k, \quad \bar{r}_\ell = -\bar{r}_k, \quad g_\ell = g_k, \quad n_\ell = -n_k,$$  

\text{(17)}

\textsuperscript{5}The reference configurations $B_{01}$ and $B_{02}$ are considered fixed.

\textsuperscript{6}Note that the usage of $g_k$ and $n_k$ only makes sense in a sufficiently close neighborhood of $x_k$ and $B_\ell$. 

Figure 2: Signed distance definition for the case that point $x_k \in \partial B_k$ has penetrated into body $B_\ell$. 

6
and obtain the gradients

\[ \frac{\partial r_k}{\partial x_k} = \bar{r}_k \quad \text{and} \quad \frac{\partial g_k}{\partial x_k} = n_k . \]  

(18)

Denoting

\[ f_s(g_k) := -\frac{\partial \phi_s}{\partial g_k}, \]  

(19)

the negative gradient of \( \phi_s \) can thus be written as

\[ -\frac{\partial \phi_s}{\partial x_k} = -\frac{\partial \phi_s}{\partial g_k} \frac{\partial g_k}{\partial x_k} = f_s(g_k) n_k , \]  

(20)

which corresponds to a force with magnitude \( |f_s(g_k)| \), acting at \( x_k \) in the direction \( n_k \). If \( f_s(g_k) > 0 \), \(-\partial \phi_s/\partial x_k\) is a force repelling \( x_k \) away from \( x_\ell \). Further, note that

\[ \frac{\partial \phi_s}{\partial x_\ell} = -\frac{\partial \phi_s}{\partial x_k} . \]  

(21)

### 2.3 Contact forces

Variation \( \delta \Pi_c \) denotes the virtual work of the contact forces. This can be seen if we rewrite eq. (12) as

\[ \delta \Pi_{ck} = -\int_{\partial B_k} \beta^s_k \mathbf{F}_k \cdot \delta \varphi_k \, da_k \]  

(22)

where we have defined the force

\[ \mathbf{F}_k := -\int_{\partial B_\ell} \beta^s_\ell \frac{\partial \phi_s}{\partial x_k} \, da_\ell = -\frac{\partial}{\partial x_k} \int_{\partial B_\ell} \beta^s_\ell \phi_s \, da_\ell , \]  

(23)

or equivalently

\[ \mathbf{F}_k := \int_{\partial B_\ell} \beta^s_\ell f_s n_k \, da_\ell , \]  

(24)

due to (20). \( \mathbf{F}_k \) corresponds to the force acting on \( x_k \) due to the combined interaction forces from all \( x_\ell \in \partial B_\ell \). The quantity

\[ t_k := \beta^s_k \mathbf{F}_k , \]  

(25)

appearing in integral (22), corresponds to the surface traction acting on the surface element \( da_k \) due to the interaction with the neighboring solid. Note that with the help of eq. (7), both the integrals in eqs. (22) and (23) can also be written as integrals over the reference surfaces. In that case we can identify the surface traction

\[ \mathbf{T}_k := \beta^{s*}_{0k} \mathbf{F}_k , \]  

(26)

which corresponds to a simple scaling of \( t_k \) by \( J_{sk} = \partial a_k / \partial A_k \).

In order to evaluate contact formulation (22)–(23), we distinguish between the following three interaction classes:

**P Point interaction.** Surface integration (23) can be degenerated into a single point evaluation. In this manner we obtain a framework as it appears in classical contact formulations (Wriggers, 2006), see Sec. 2.3.1. An example for class P interaction are classical penalty formulations and cohesive zone models, see Sec. 3.1.
**Short-range surface interaction.** Short-range interactions, due to decaying $\phi_s$, imply that only those parts of $\partial B_\ell$ closest to point $x_k$ contribute significantly to $F_k$. Under reasonable assumptions, integral (23) can then be integrated analytically, see Sec. 2.3.2. Examples for class S interaction are van-der-Waals adhesion and a modified penalty method, see Sec. 3.2.

**Long-range surface interaction.** Long-range interactions require the full integration of both (22) and (23). For deformable bodies numerical integration is generally needed. Only special cases can be integrated analytically, see Sec. 4.2. Examples for class L interaction are electrostatic or gravitational interactions, see Sec. 3.3. A general computational formulation for class L interaction is presented in Sec. 5.3.

In the following we discuss the analytical integration used in classes P and S. Physical examples for all classes follow in Sec. 3. Formulation (22)–(23) leads naturally to a special contact algorithm that is denoted the *two-half-pass algorithm*. This algorithm, as well as numerical examples for all classes, are presented in Secs. 5 and 7.

**2.3.1 Point interaction (class P)**

For class P we suppose that point $x_k \in \partial B_k$ only interacts with one particular point on $\partial B_\ell$ that depends on $x_k$ and is denoted $x_0^\ell$. For example, $x_0^\ell$ can be taken as the closest projection point of $x_k$ on $\partial B_k$, which is discussed in Sec. 2.3.3. $x_0^\ell$ can also be considered as the current position of the initial projection point of $x_k$, e.g. to model the effect of a connected filament, as it can be considered in a cohesive zone formulation. In consequence the integration over $\partial B_\ell$ vanishes, such that the interaction force at $x_k$ becomes

$$F_k = -\frac{\partial \phi_s(x_k, x_0^\ell)}{\partial x_k},$$  

(27)

Mathematically, this expression can be obtained from eq. (23) if the interaction density of $\partial B_\ell$ is replaced by a Dirac delta function, i.e.

$$\beta_\ell^s = \delta(x_\ell - x_0^\ell).$$  

(28)

Since this leads to a different treatment of surfaces $\partial B_k$ and $\partial B_\ell$, the formulation tends to be biased (see cases a. and b. below). Due to eq. (20), $F_k$ can be written as

$$F_k = f_s(g_k^0) n_k^0,$$  

(29)

where $g_k^0$ and $n_k^0$ now refer to the signed distance and direction, as introduced in eqs. (13)–(15), between $x_k$ and $x_0^\ell$. This setup is similar to the formulation of class S, see eq. (34) in Sec. 2.3.2. But contrary to class S, the contact force of class P does not depend explicitly on the surface stretch $J_s^\ell$ of $\partial B_\ell$. Also, depending on the chosen point $x_0^\ell$, $F_k$ is not necessarily normal to $\partial B_\ell$, as it is for class S. For the evaluation of the contact force $F_k$ and the virtual contact work $\delta \Pi_c$ from eqs. (10) and (12), several options can be considered:

a. $\beta_1^s = \delta$, i.e. $\phi_s = \phi_s(x_0^1, x_2)$: This corresponds to a one-pass contact algorithm where $\partial B_1$ is the master and $\partial B_2$ is the slave surface, i.e.

$$\delta \Pi_c = -\int_{\partial B_2} \beta_2^s F_1 \cdot \delta \varphi_1 \, da_2 - \int_{\partial B_2} \beta_2^s F_2 \cdot \delta \varphi_2 \, da_2,$$  

(30)

with $F_1 = -\partial \phi_s/\partial x_1^0$ and $F_2 = -\partial \phi_s/\partial x_2$. 

8
b. $\beta_2^s = \delta$, i.e. $\phi_s = \phi_s(x_1, x_2^0)$: this corresponds to a one-pass contact algorithm where $\partial B_2$ is the master and $\partial B_1$ is the slave surface, i.e.

$$
\delta \Pi_c = - \int_{\partial B_1} \beta_1^s F_1 \cdot \delta \phi_1 \, da_1 - \int_{\partial B_2} \beta_2^s F_2 \cdot \delta \phi_2 \, da_2 ,
$$

with $F_1 = -\partial \phi_s/\partial x_1$ and $F_2 = -\partial \phi_s/\partial x_2^0$.

c. Alternating $\beta_1^s = \delta$ and $\beta_2^s = \delta$ corresponds to a two-pass algorithm where $\partial B_1$ and $\partial B_2$ alternate as master and slave surface, i.e.

$$
\delta \Pi_c = - \int_{\partial B_1} \beta_1^s F_1 \cdot \delta \phi_1 \, da_1 - \int_{\partial B_2} \beta_2^s F_2 \cdot \delta \phi_2 \, da_2 ,
$$

with $F_1 = -\partial \phi_s(x_1, x_2^0)/\partial x_1$ and $F_2 = -\partial \phi_s(x_1^0, x_2)/\partial x_2$.

Cases a. and b. are classical contact descriptions that have been considered by many researchers before. In a discrete setting they can, for example, be evaluated by a NTS (Node-to-Surface) formulation. Case c., on the other hand, yields a new contact description that is unbiased w.r.t. the surfaces and preserves the general structure of eq. (22), which is also common to the other two classes. Formulation c. is similar to the one proposed by Papadopoulos et al. (1995). Examples for class P interaction potentials are discussed in Sec. 3.1. It is shown there, that they can be used to define both normal and tangential contact models.

### 2.3.2 Short-range surface interaction (class S)

For class S we suppose that potential $\phi_s$ decays with distance $g_k$. In this case the surface integration (22) can be carried out analytically over a (flat) plane.\(^7\) The contact force at $x_k$ can then be expressed in terms of the minimum distance vector $x_k - x_p$, where $x_p \in \partial B_\ell$ is the point closest to $x_k$. Considering constant density within the neighborhood of $x_p$,\(^8\) eq. (24) becomes

$$
F_k := \beta_k^s \int_{\partial B_\ell} f_s n_k \, da_\ell .
$$

This can be integrated analytically to obtain

$$
F_k = 2\pi \beta_k^s g_n \int_{\partial B_\ell} \phi_s(g_n) n_p ,
$$

see Appendix A. Here $g_n$ denotes the signed normal distance formally defined in Sec. 2.3.3, eq. (39).

Introducing the force constant $F_0 := 2\pi \beta_0^s r_0 g_0$, where $r_0$ is some length constant and $g_0$ is some energy constant, and using eq. (4), eq. (34) can be rewritten in the normalized form

$$
\frac{F_k}{F_0} := \frac{g_n \phi_s(g_n)}{r_0} \frac{\phi_s(g_n)}{\phi_0} J_0^{-1} n_p ,
$$

where $J_p$ denotes the surface stretch of $\partial B_\ell$, evaluated at $x_p$. Note, that according to footnote 6, $J_p$ is considered constant in the neighborhood of $x_p$ that is relevant for integration.

\(^7\)Here the decay of $\phi_s$ is considered such that the neighboring surface $\partial B_\ell$ appears flat within the cutoff radius of $\phi_s$. This is usually not the case for long-range interactions, which therefore cannot be treated by class S. As a first order improvement, the curvature of the surface can be taken into account (Sauer and Wriggers, 2009).

\(^8\)This is satisfied if one assumes that the reference density $\beta_0^s$ and the deformation are nearly homogeneous within the cutoff radius of $\phi_s$. Note that $\beta_0^s$ can be considered as a purely numerical, and thus chosen, parameter.
As seen, $F_k$ is an explicit function of $g_n$, $n_p$ and $J_p$. Depending on the signs of $g_n$ and $\phi_s$, we have the four cases

1. $\phi_s > 0$ for $g_n > 0$ : $x_k \not\in B_\ell$ is repelled away from $\partial B_\ell$ ,
2. $\phi_s < 0$ for $g_n > 0$ : $x_k \not\in B_\ell$ is attracted toward $\partial B_\ell$ ,
3. $\phi_s > 0$ for $g_n < 0$ : $x_k \in B_\ell$ is pushed inward from $\partial B_\ell$ ,
4. $\phi_s < 0$ for $g_n < 0$ : $x_k \in B_\ell$ is pushed out of $B_\ell$ .

We finally note that the class S interaction formulation is a frictionless contact model, since the contact force only depends on $n_p$.

### 2.3.3 Closest distance computation for class P and class S

In order to evaluate $F_k$ according to eq. (34) and (29) – if $x^0_k = x_p$ such that $g_k^0 = g_n$ and $n_k^0 = n_p$ – the signed surface distance $g_n$ is needed. In practice the (absolute) surface distance $r_n$ is computed from the minimum distance problem

$$r_n(x_k) := \min r_k \quad \forall \, x_\ell \in \partial B_\ell .$$

The point $x_\ell$ that solves this minimization problem is denoted as $x_p$; the outward surface normal at $x_p$ is denoted as $n_p$ (as shown in Fig. 23, Appendix A). Note that $x_p$ and $n_p$, as well as $r_n$, are functions of $x_k$. Analogous to eq. (13) we thus have

$$r_n := x_k - x_p , \quad r_n = \| r_n \| , \quad \bar{r}_n := \frac{r_n}{r_n} .$$

Note that in general, eq. (37) can have several local minima. The signed distance $g_n$ then follows as

$$g_n := \begin{cases} \quad r_n & \text{if } x_k \not\in B_\ell , \\ -r_n & \text{if } x_k \in B_\ell , \end{cases}$$

and the surface normal of $\partial B_\ell$ at $x_p$ can be determined from

$$n_p := \begin{cases} \quad \bar{r}_n & \text{if } x_k \not\in B_\ell , \\ -\bar{r}_n & \text{if } x_k \in B_\ell , \end{cases}$$

or equivalently

$$n_p = \frac{x_k - x_p}{g_n} .$$

The normal $n_p$ can then be used to define a set of tangent vectors for $\partial B_\ell$.

### 2.4 Contact energy

Since a mixture of numerical and analytical integration is used in the previously described contact formulation, the question arises on how to compute the energy $\Pi_c$ consistently. This is discussed next.

Since the formulation of the contact forces is based on a potential, i.e. eq. (3), it is fairly straightforward to evaluate the contact energy. Therefore consider the deformation states $\varphi_1^a$ and $\varphi_2^a$ of the two bodies at time $t_a$ and the states $\varphi_1^b$ and $\varphi_2^b$ at time $t_b$. The difference in the contact energy between these two states is simply given by the difference in $\Pi_c$ evaluated at those states, i.e.

$$\Delta \Pi_c = \Pi_c(\varphi_1^b, \varphi_2^b) - \Pi_c(\varphi_1^a, \varphi_2^a) .$$
Note, that according to the surface potential formulation, eq. (3), the contact energy is stored in the four-dimensional volume $\partial B_1 \times \partial B_2$.

The contact energy of eq. (3) can be rewritten as

$$\Pi_c = \int_{\partial B_1} \beta_1^s \Phi_2 \, da_1 ,$$

or equivalently

$$\Pi_c = \int_{\partial B_2} \beta_2^s \Phi_1 \, da_2 ,$$

where we have defined

$$\Phi_{\ell} := \int_{\partial B_{\ell}} \beta_{\ell}^s \phi_{\ell} \, da_{\ell} .$$

These expressions are general and apply to all three classes introduced in Sec. 2.3. For class L, no further simplification is possible in general, and eq. (45) needs to be evaluated numerically. For class P, we simply find $\Phi_{\ell} = \phi_{\ell}$ due to eq. (28). For class S, integration (45) can be carried out analytically according to the procedure of Sec. 2.3.2. According to Appendix A (see Fig. 23 and Eq. (149)) we have $da_{\ell} = 2 \pi g_k \, dg_k$, so that eq. (45) then becomes

$$\Phi_{\ell}(g_n) = 2 \pi \beta_{\ell}^s \int_{g_n}^{\infty} g_k \, \phi_{\ell} \, dg_k ,$$

with $r_{\ell}$ according to eq. (156). Due to the different integration limits for $g_k < 0$ and $g_k > 0$, the function $\Phi(g_n)$ has a jump at $g_n = 0$. The jump can be removed if we simply set $r_{\ell} = +\infty$. This does not affect the contact forces, since it only adds a constant (the value of the jump) to $\Phi_{\ell}$ for the case $g_n < 0$. Together with eq. (34) we thus have

$$\Phi_{\ell}(g_n) = \int_{g_n}^{\infty} \| F_k \| \, dg_k .$$

With the definition of $\Phi_{\ell}$ for the three classes, the variation $\delta \Pi_{ck}$, introduced in eq. (12), can then also be written as

$$\delta \Pi_{ck} = \int_{\partial B_k} \beta_{\ell}^s \frac{\partial \Phi_{\ell}}{\partial x_k} \cdot \delta \varphi_{\ell} \, da_k ,$$

so that, due to eqs. (22) and (23), the contact force $F_k$ can also be defined as

$$F_k := - \frac{\partial \Phi_{\ell}}{\partial x_k} .$$

Inserting expression (46) reproduces formula (34) correctly.

The difference in the contact energy, given in eq. (42), can now be expressed as

$$\Delta \Pi_c = \int_{\partial B_1} \beta_1^s \Delta \Phi_2(g_{na}, g_{nb}) \, da_1 ,$$

or equivalently

$$\Delta \Pi_c = \int_{\partial B_2} \beta_2^s \Delta \Phi_1(g_{na}, g_{nb}) \, da_2 ,$$

with

$$\Delta \Phi_{\ell}(g_{na}, g_{nb}) = \Phi_{\ell}(g_{nb}) - \Phi_{\ell}(g_{na}) .$$
Here \( g_{na} \) (\( g_{nb} \) respectively) denotes the signed distance between \( B_\ell \) and \( x_k \) at the deformation states \( \varphi^a_\ell \) and \( \varphi^b_\ell \) (\( \varphi^a_k \) and \( \varphi^b_k \) respectively). Note that in the evaluation of eq. (52), the four cases

1. \( g_{na} > 0 \), \( g_{nb} > 0 \),
2. \( g_{na} > 0 \), \( g_{nb} \leq 0 \),
3. \( g_{na} \leq 0 \), \( g_{nb} \leq 0 \),
4. \( g_{na} \leq 0 \), \( g_{nb} > 0 \),

need to be distinguished.

In a finite element setting, expressions (50) and (51), just as the virtual work (22), are computed by numerical integration. The fact that two alternative expressions for \( \Delta \Pi_c \) exist can be used to measure the integration error present in the contact formulation.

Remark 1: It is important to note that potential \( \phi_s \) and force \( f_s n_k \) describe the interaction between surface points \( x_k \in \partial B_k \) and \( x_\ell \in \partial B_\ell \), while \( \Phi_\ell \) and \( F_k \) correspond to the potential and force at \( x_k \) due to the interaction of \( x_k \) with the entire surface \( \partial B_\ell \). For class P, \( \phi_s \) and \( \Phi_\ell \) as well as \( f_s n_k \) and \( F_k \) are identical.

Remark 2: Using eq. (7), the integration in eqs. (43)–(45), (48) and (49)–(51) can be easily transformed to the reference surfaces.

### 2.5 Contact equilibrium

Here we show that the contact formulations outlined in Sec. 2.3 satisfy local and global contact equilibrium. We first consider surface interaction (class S and class L), and then point interaction (class P).

#### 2.5.1 Surface interaction (class S and class L)

1. **Global equilibrium:** Global contact equilibrium is satisfied if

\[
P_1 + P_2 = 0 \, ,
\]

where \( P_k \) denotes the net contact force

\[
P_k := \int_{\partial B_k} \beta^a_k F_k \, da_k \,'.
\]

Proof: According to eq. (23)

\[
P_1 = -\int_{\partial B_1} \beta^a_1 \int_{\partial B_2} \beta^b_2 \partial \phi_s \frac{\partial x_1}{\partial x_2} \, da_2 \, da_1 \, .
\]

Pulling the density \( \beta^a_1 \) inside the second integral and using eq. (21), we obtain

\[
P_1 = \int_{\partial B_1} \int_{\partial B_2} \beta^a_1 \beta^b_2 \frac{\partial \phi_s}{\partial x_2} \, da_2 \, da_1 \, .
\]

Changing the order of integration, we get

\[
P_1 = \int_{\partial B_2} \beta^b_2 \int_{\partial B_1} \beta^a_1 \frac{\partial \phi_s}{\partial x_2} \, da_1 \, da_2 = -P_2 \, .
\]
2. Local equilibrium: Local equilibrium is only satisfied if the interaction can be formulated locally, i.e. the interacting surfaces actually touch or are at least very close. That is not the case for long-range surface interactions (class L). For class S local equilibrium is satisfied if the neighboring surfaces are locally parallel: Consider the subregions of the contact surfaces $\partial P_k \subset \partial B_k (k = 1, 2)$, such that $\partial P_1$ and $\partial P_2$ are parallel ($\partial P_1 \parallel \partial P_2$) and have equal area ($|\partial P_1| = |\partial P_2|$). Then the local contact forces

$$P_k := \int_{\partial P_k} \beta_k^n F_k \, da_k ,$$

(59)

satisfy local contact equilibrium

$$P_1 + P_2 = 0 \quad \forall \partial P_k .$$

(60)

Proof:

According to eq. (34), the contact traction $t_k = \beta_k^n F_k$ acting at $x_k \in \partial P_k$ is given by

$$t_k(x_k) = 2\pi \beta_1^n \beta_2^n g_n \phi_s n_p .$$

(61)

Here, $n_p$ corresponds to the normal of $\partial P_k$. Since $\partial P_1 \parallel \partial P_2$, we have $n_p(x_1) = -n_p(x_2)$ and $g_n(x_1) = g_n(x_2)$. Since $|\partial P_1| = |\partial P_2|$ we thus have $t_1 = -t_2$. From eq. (59) then follows that $P_1 = -P_2$ for all $\partial P_k$.

Due to the equivalence of the traction across the contact surface, both linear and angular momentum are conserved for the two bodies in contact.

Note that for all classes global contact equilibrium is always fulfilled. Local contact equilibrium, on the other hand, is only fulfilled if $\partial P_1 = \partial P_2$. This is typically not exactly satisfied in a numerical setting. There, however, we can still obtain an excellent agreement of the local contact tractions, as is demonstrated by the numerical examples in Sec. 7.

2.5.2 Point interaction (class P)

The three versions given in Sec. 2.3.1 yield the following definitions of the global contact interaction forces

a. $P_1 = \int_{\partial B_2} \beta_2^n F_1 \, da_2 , \quad P_2 = \int_{\partial B_2} \beta_2^n F_2 \, da_2 ,$

b. $P_1 = \int_{\partial B_1} \beta_1^n F_1 \, da_1 , \quad P_2 = \int_{\partial B_1} \beta_1^n F_2 \, da_1 ,$

c. $P_1 = \int_{\partial B_1} \beta_1^n F_1 \, da_1 , \quad P_2 = \int_{\partial B_2} \beta_2^n F_2 \, da_2 .$

(62)

It can be seen that version c. is the most consistent formulation, since here the force $P_k$ acting on body $B_k$ is obtained from the integration over the surface of that body. In version a. and b. this consistency is missing. Global equilibrium is satisfied if $P_2 = -P_1$. According to Sec. 2.3.1,

$$F_1 = f_s(g_0^0) n_1^0 , \quad g_1^0 = \pm \|x_1 - x_2^0\| , \quad n_1^0 = (x_1 - x_2^0)/g_1^0 ,$$

$$F_2 = f_s(g_0^0) n_2^0 , \quad g_2^0 = \pm \|x_2 - x_1^0\| , \quad n_2^0 = (x_2 - x_1^0)/g_2^0 ,$$

(63)

where $x_1 \in \partial B_1$ and $x_2 \in \partial B_2$ are general points and $x_1^0 \in \partial B_1$ and $x_2^0 \in \partial B_2$ are special points that depend on $x_1$ and $x_2$, respectively.
Local equilibrium, and thus also global equilibrium, are satisfied if we consider the subregions \( \partial P_k \subset \partial B_k \) such that \( \partial P_1 \parallel \partial P_2 \) and \( |\partial P_1| = |\partial P_2| \). In this case we have \( \mathbf{n}_1^0 = -\mathbf{n}_1^0 \) and \( g_2^0 = g_1^0 \). so that \( \mathbf{F}_2 = -\mathbf{F}_1 \). With this it can be easily seen that versions a. and b. satisfy local, and thus also global, equilibrium. Version c. can also be formulated such that it satisfies equilibrium. This is shown in the context of the specific examples in the following section.

### 2.6 Summary of important relations

Table 1 provides a summary of the important relations of the preceding sections.

<table>
<thead>
<tr>
<th>interaction</th>
<th>class L</th>
<th>class S</th>
<th>class P</th>
</tr>
</thead>
<tbody>
<tr>
<td>total energy</td>
<td>( \Pi_k = \int_{\partial B_k} \beta^\ell_k \Phi_\ell , da_k = \int_{\partial B_0k} \beta^\ell_{0k} \Phi_\ell , dA_k ), since ( \beta^\ell_k , da_k = \beta^\ell_{0k} , dA_k )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>energy due to ( \partial B_\ell )</td>
<td>( \Phi_\ell = \int_{\partial B_\ell} \beta^\ell_s \phi_s , da_\ell )</td>
<td>( \Phi_\ell = \int_{g_n}</td>
<td></td>
</tr>
<tr>
<td>force at ( x_k \in \partial B_k )</td>
<td>( F_k = -\frac{\partial \Phi_\ell}{\partial x_k} )</td>
<td>( F_k = 2\pi \beta^\ell_s g_n \phi_s(g_n) \mathbf{n}_p )</td>
<td>( F_k = -\frac{\partial \phi_s(x_k, x_\ell)}{\partial x_k} )</td>
</tr>
<tr>
<td>net force on ( \partial B_k )</td>
<td>( P_k = \int_{\partial B_k} t_k , da_k = \int_{\partial B_0k} T_k , dA_k ), ( t_k = \beta^\ell_k F_k ), ( T_k = \beta^\ell_{0k} F_k )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>equilibrium</td>
<td>only satisfied globally</td>
<td>satisfied globally &amp; locally</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Interaction forces between surfaces \( \partial B_k \), \( k = 1, 2 \), according to the SPBC formulation. Index \( \ell \neq k \) denotes the body neighboring body \( k \). No sum is implied on \( k \) or \( \ell \)

### 3 Examples for potential \( \phi_s \)

In order to illustrate the SPBC formulation, several examples are considered in the following. They show that the new formulation unifies many different contact approaches.

#### 3.1 Class P: Point interaction

For the following examples we consider version c. as is introduced in Sec. 2.3.1. Contrary to a. and b., version c. is symmetric with respect to the two contact partners. Point interaction formulations according to version a. and b. are the dominating formulations in the present literature, e.g. see Wriggers (2006), Zavarise and De Lorenzis (2009b). As discussed in Sec. 2.3.1, the interaction potential \( \phi_s = \phi_s(x_k, x_\ell) \) is now only evaluated at the special point \( x_\ell = x^0_\ell \).

1. **Class P penalty method** (classical penalty method): **Contact force proportional to the normal penetration** \(-g_n\). Point interaction according to the classical penalty method is formulated in terms of the signed surface distance \( g_n \). The interaction potential is now given as

   \[
   \phi_s(x_k, x_\ell) = \frac{\epsilon g_k^2}{2\beta^\ell_k}, \quad \text{for } g_k < 0, \quad \text{else } \phi_s = 0.
   \]
Evaluated at \( x^0 = x_p \) (i.e. \( g_k = g^0_k = g_n \)) this yields
\[
\phi_s(x^k, x^0) = \frac{\epsilon g_n^2}{2\beta^0_k}, \quad \text{for} \quad g_n < 0, \quad \text{else} \quad \phi_s = 0.
\]
(65)
e.g. see Fischer and Wriggers (2005). Here \( \epsilon \) is the penalty parameter. This corresponds to a penalty formulation w.r.t. the current configuration. Replacing \( \beta^0_k \) by \( \beta^0_{0k} \) yields a penalty formulation w.r.t. the reference configuration. Since \( \beta^0_k \, da_k = \text{const.} \), the contact force, given by eq. (27), can be found from the expression
\[
\beta^0_k \, F_k \, da_k = -\frac{\partial(\beta^0_k \, \phi_s \, da_k)}{\partial x_k} = -\frac{\epsilon}{2} \frac{\partial g_n^2}{\partial x_k} \, da_k - \frac{\epsilon}{2} \frac{\partial g_n}{\partial x_k} \, da_k, \quad \text{for} \quad g_n < 0,
\]
(66)
which is needed for the evaluation of \( \delta \Pi_{ck} \), see eq. (22). The second term is sometimes neglected since \( g_n \) goes to zero as \( \epsilon g_n \) approaches the correct contact pressure. In this case we have
\[
F_k = -\frac{\epsilon g_n^2}{\beta^0_k} \, n_k^0, \quad \text{for} \quad g_n < 0,
\]
(67)
due to eq. (18); note that \( n_k^0 = n_p \). It can be confirmed that this \( F_k \) also satisfies local and global contact equilibrium for version c. Following the arguments of Sec. 2.5.2 we now find \( \beta^1_k \, F_1 = -\beta^2_k \, F_2 \), which implies traction continuity \( t_1 = -t_2 \) (i.e. local equilibrium) and thus global equilibrium.

2. Exponential cohesive zone model (Xu and Needleman, 1993): Exponentially decaying contact force: Consider two bodies that are bonded along the common surface \( \partial_c B_{01} = \partial_c B_{02} \) in the reference configuration such that \( \beta^0_{01} = \beta^0_{02} \). This can be understood as bonding the two surfaces at a number of common points \( X_k \in \partial_c B_{0k} \). According to version c. (see Sec. 2.3.1), we consider that during deformation the point \( x_k = \phi_k(X_k) \) of surface \( \partial_c B_{0k} \) interacts with the point \( x^0_\ell = \phi_\ell(X_k) \) of the neighboring surface \( \partial_c B_\ell \). The exponential cohesive zone model is obtained by defining the surface potential as
\[
\phi_s(r^0_k) = -\phi_0 \left( 1 + \frac{r^0_k}{r_0} \right) \exp \left( -\frac{r^0_k}{r_0} \right), \quad r^0_k > 0.
\]
(68)
The contact force then becomes
\[
F_k(r^0_k) = -\frac{\phi_0 \, r^0_k}{r_0 \, r^0} \, \exp \left( -\frac{r^0_k}{r_0} \right) \, r^0_k, \quad r^0_k > 0.
\]
(69)
A graph of these functions is shown in Figure 3.

Figure 3: Exponential cohesive zone model: surface potential \( \phi_s \) (left); contact force \( F_k \) (right).
The cohesive zone model given here satisfies contact equilibrium. According to Sec. 2.5.2 the global interaction forces are given by eq. (62c) for version c. Transforming the integrals to the reference configuration gives

\[ P_1 = \int_{\partial c_{B_0}} \beta_{01}^s F_1 \, dA_1, \quad P_2 = \int_{\partial c_{B_0}} \beta_{02}^s F_2 \, dA_2, \]  

(70)

where \( \partial c_{B_0} = \partial c_{B_0} \) and \( \beta_{02}^s = \beta_{01}^s \), as initially stated. The cohesion forces \( F_1 \) and \( F_2 \) are in the form of eq. (63). Now, if \( x_2^0 \) coincides with \( x_2 \), then \( x_1^0 \) coincides with \( x_1 \). Therefore we have \( r_0^1 = r_0^2 \) and \( n_0^1 = -n_0^2 \) for this pairing. This means that \( F_2 = -F_1 \) and thus \( P_2 = -P_1 \). Therefore global contact equilibrium is satisfied. Local equilibrium is further satisfied if \( \beta_{01}^s = \beta_{02}^s \), since then the tractions \( t_1 = \beta_{01}^s F_1 \) and \( t_2 = \beta_{02}^s F_2 \) are equal. This will be the case for the surface regions that are still closely bonded, but not for those that have already separated, as is seen in the example of Sec. 7.4 (see Fig. 22c).

### 3.2 Class S: Short-range surface interaction

#### 3. Class S barrier method

(Modified barrier method): \textit{Contact force proportional to} \( r_n^{-1} \).

Considering the surface potential

\[ \phi_s(r_k) = \phi_0 \left( \frac{r_0}{r_k} \right)^2, \]  

(71)

defined for positive distances, we obtain the contact force

\[ F_k = F_0 \frac{r_0}{r_n} n_p, \]  

(72)

from eq. (35) with \( F_0 := 2\pi\beta_0^s r_0 \phi_0 \), and the contact energy difference

\[ \Delta \Phi_f(r_{na}, r_{nb}) = \Phi_0 \left( J_p^{-1} \ln \frac{r_{na}}{r_0} - J_p^{-1} \ln \frac{r_{nb}}{r_0} \right), \]  

(73)

from eqs. (52) and (46) with \( \Phi_0 := F_0 r_0 \). Since \( \|F_k\| \to \infty \) for \( r_n \to 0 \), two bodies approaching each other can not penetrate in theory, so that the signed distance would always remain positive. In a numerical setting a failsafe has to be included in order to treat penetration, e.g. by considering a cross-constrained potential (see example 9). The plots of \( \phi_s \), \( \Delta \Phi_f \) and \( F_k := \|F_k\| \), according to the barrier method, are shown in Fig. 4 for \( J_p = 1 \). Note that no computational

![Figure 4](image-url)

Figure 4: Barrier method: surface potential \( \phi_s \) (left); contact force \( F_k := \|F_k\| \) (middle); contact potential \( \Delta \Phi_f \) for \( r_{na} = r_0 \) (right).

active set strategy is needed for barrier methods, since \( F_k \) is defined smoothly for all distances.
Further, note that the barrier method can also be formulated in the framework of class P, see Wriggers (2006).

4. **Constant resistance to penetration:** *Contact force constant.* This is achieved by setting

\[
\phi_s(g_k) := \phi_0 \begin{cases} 
    r_0/g_k & \text{if } r_{\text{lim}} < g_k \leq 0, \\
    0 & \text{else},
\end{cases}
\]

so that

\[
F_k(g_n) := J_p^{-1} n_p \begin{cases} 
    F_0 & \text{if } r_{\text{lim}} < g_n \leq 0, \\
    0 & \text{else},
\end{cases}
\]

according to eq. (34). The contact potential \(\Phi^\ell\) now becomes

\[
\Phi^\ell(g_n) := \frac{\Phi_0}{J_p r_0} \begin{cases} 
    -r_{\text{lim}} & \text{if } g_n < r_{\text{lim}}, \\
    -g_n & \text{if } r_{\text{lim}} < g_n \leq 0, \\
    0 & \text{if } g_n > 0,
\end{cases}
\]

according to eq. (46). Here the constant \(-\Phi_0 r_{\text{lim}}/r_0/J_p\) has been added for \(g_n < 0\) to avoid a jump in \(\Phi^\ell\). In principle \(r_{\text{lim}}\) can approach \(-\infty\). The three functions are shown in Fig. 5. The usefulness of this contact model is seen in example 7 below.

5. **Class S penalty method** (modified penalty method): *Contact force proportional to the penetration \(-g_n\).* This is achieved by setting the surface potential to

\[
\phi_s(g_k) := \begin{cases} 
    -\phi_0 & \text{if } r_{\text{lim}} < g_k < 0, \\
    0 & \text{else},
\end{cases}
\]

which can equivalently be expressed with the help of the Heaviside function \(H\) as

\[
\phi_s(g_k) = \phi_0 [H(g_k) - H(g_k - r_{\text{lim}})].
\]

Here \(r_{\text{lim}} < 0\) is a constant. The contact force then becomes

\[
F_k(g_n) := -\frac{F_0}{J_p} n_p \begin{cases} 
    g_n & \text{if } r_{\text{lim}} < g_n < 0, \\
    0 & \text{else},
\end{cases}
\]

with \(F_0 = 2\pi\beta_0^s r_0 \phi_0\), as before. With this, the contact traction \(T_k\), defined in eq. (26), becomes

\[
T_k = -\epsilon_n g_n J_p^{-1} n_p,
\]

where \(\epsilon_n = 2\pi\beta_0^s \beta_0^s \phi_0\) denotes the penalty parameter. The appearance of \(J_p\) means that the traction \(T_k\) depends on the surface deformation of the neighboring body. This formulation gives
a symmetric tangent for conforming FE meshes as is shown in Appendix F. The evaluation of eq. (46) now gives
\[
\Phi_\ell(g_n) := \begin{cases} 
\frac{\Phi_0}{2J_p r_0^2} \frac{r_{\text{lim}}^2}{r_0^2} & \text{if } g_n < r_{\text{lim}}, \\
\frac{g_n^2}{2J_p r_0^2} & \text{if } r_{\text{lim}} < g_n < 0, \\
0 & \text{if } g_n > 0,
\end{cases}
(81)
\]
with \(\Phi_0 = F_0r_0\), as before. The constant \(2\Phi_0 r_{\text{lim}}^2/r_0^2/J_p\) is added to \(\Phi_\ell\) for \(g_n < 0\). A graph of these functions is shown in Fig. 6. The limit value \(r_{\text{lim}} < 0\) is necessary to ensure that \(\phi_s\) approaches zero as \(g_k \to -\infty\), and to ensure the consistency with eqs. (156) and (34). Note that \(-r_{\text{lim}}\) may not be considered too large, as otherwise the assumption in footnote 6 is violated. Also note that \(r_{\text{lim}}\) is irrelevant for a numerical implementation considered in Sec. 5.

Remark 1: From this approach one can also easily construct penalty formulations where the contact force is proportional to \(g_n^p\), where \(p > 1\) is some chosen power.

Remark 2: There is an important conceptual difference to classical penalty formulations discussed in example 1. Classically the penalty traction depends on the deformation of the surface it acts on. Here, on the other hand, the penalty traction depends on the deformation of the neighboring body. Both formulations satisfy contact equilibrium as is seen in Sec. 2.5. Further details in the context of a computational implementation are discussed in Sec. 5 and 7.

6. Penalty method with slight adhesion: Extension of the penalty force law into the regime of small separations. This is achieved if
\[
\phi_s(g_k) := \begin{cases} 
-\phi_0 & \text{if } r_{\text{lim}^-} < g_k < r_{\text{lim}^+}, \\
0 & \text{else},
\end{cases}
(82)
\]
with the constants \(r_{\text{lim}^-} < 0\) and \(r_{\text{lim}^+} > 0\) so that
\[
F_k(g_n) := -\frac{F_0 n_p}{J_p r_0} \begin{cases} 
g_n & \text{if } r_{\text{lim}^-} < g_n < r_{\text{lim}^+}, \\
0 & \text{else},
\end{cases}
(83)
\]
and
\[
\Phi_\ell(g_n) := \frac{\Phi_0}{2J_p r_0^2} \begin{cases} 
\frac{r_{\text{lim}^-}^2 - r_{\text{lim}^+}^2}{r_0^2} & \text{if } g_n < r_{\text{lim}^-}, \\
\frac{r_{\text{lim}^-}^2}{r_0^2} - \frac{r_{\text{lim}^+}^2}{r_0^2} & \text{if } r_{\text{lim}^-} < g_n < r_{\text{lim}^+}, \\
0 & \text{if } g_n > r_{\text{lim}^+}.
\end{cases}
(84)
\]

7. Offset penalty method: Linearly increasing contact force starting at \(r_m > 0\). By setting
\[
\phi_s(g_k) := \phi_0 \begin{cases} 
\frac{r_m}{g_k} - 1 & \text{if } r_{\text{lim}} < g_k < r_m, \\
\frac{r_m}{g_k} & \text{else},
\end{cases}
(85)
\]
such that

\[ F_k(g_n) := \frac{F_0}{J_p r_0} \left\{ \begin{array}{ll}
r_m - g_n & \text{if } r_{\lim} < g_n < r_m, \\
0 & \text{else},
\end{array} \right. \]  

(86)

and

\[ \Phi_\ell(g_n) := \frac{\Phi_0}{2 J_p r_0^2} \left\{ \begin{array}{ll}
(r_{\lim} - r_m)^2 & \text{if } g_n < r_{\lim}, \\
(g_n - r_m)^2 & \text{if } r_{\lim} < g_n < r_m, \\
0 & \text{if } g_n > r_m,
\end{array} \right. \]  

(87)

with \( r_{\lim} < 0 \) and \( r_m > 0 \), the penalty contact force law is offset by the distance \( r_m \). This case can be viewed as a superposition of examples 4 and 6. It is useful in order to define cross-constrained contact laws (see example 9).

8. Adhesive contact: Contact force increasingly repulsive for small distances and attractive and decaying for large distances. Considering for example

\[ \phi_s(r_k) = \phi_0 \left[ \frac{1}{45} \left( \frac{r_0}{r_k} \right)^{10} - \frac{1}{3} \left( \frac{r_0}{r_k} \right)^{4} \right], \quad r_k > 0, \]  

(88)

we find

\[ F_k(r_n) = \frac{F_0}{J_p} \left[ \frac{1}{45} \left( \frac{r_0}{r_n} \right)^{9} - \frac{1}{3} \left( \frac{r_0}{r_n} \right)^{3} \right] n_p, \quad r_n > 0, \]  

(89)

with \( F_0 = 2 \pi \beta_0 r_0 \phi_0 \), again. This particular force law can also be derived from the volume interaction

\[ \Pi_\ell = \int_{B_1} \int_{B_2} \beta_1 \beta_2 \phi \, dv_2 \, dv_1, \]  

(90)

where \( \phi \) is the Lennard-Jones potential (Sauer and Li, 2007b; Sauer and Wriggers, 2009). According to eq. (46), the contact energy \( \Phi_\ell \) now becomes

\[ \Phi_\ell(r_n) = \frac{\Phi_0}{360} \left[ \frac{1}{6} \left( \frac{r_0}{r_n} \right)^{6} - \frac{1}{360} \left( \frac{r_0}{r_n} \right)^{2} \right], \quad r_n > 0. \]  

(91)

Remark: For very small bodies with large surface curvatures, eq. (88) corresponds to long-range surface interaction, and should be treated as class L interaction.

9. Cross-constrained potential: Regularization for the barrier and adhesion formulation by combining them with the penalty method. Examples 3 and 8 are only defined for non-penetration \( (g_n > 0) \). In order to extend the formulation to \( g_n < 0 \), e.g. to provide a treatment for numerical penetration, let us consider

\[ \phi_s = \left\{ \begin{array}{ll}
0 & \text{if } g_k < r_{\lim}, \\
\phi_0 \frac{F_0}{r_0} \left( \frac{F^{\ast}}{g_k} + F'_a \right) & \text{if } r_{\lim} < g_k < r_m, \\
\phi_a(g_k) & \text{if } g_k \geq r_m,
\end{array} \right. \]  

(92)

where \( F^{\ast} = F_a(r_m) - r_m F'_a(r_m) \), and where \( \phi_a, F_a \) and \( F'_a = \partial F_a / \partial g_n \) denote the surface potential, contact force magnitude and corresponding derivative defined by the barrier or adhesion contact formulation in examples 3 and 8 above. It then follows that

\[ F_k = n_p \left\{ \begin{array}{ll}
0 & \text{if } g_n < r_{\lim}, \\
F^{\ast} + F'_a g_n & \text{if } r_{\lim} < g_n < r_m, \\
F_a(g_n) & \text{if } g_n \geq r_m,
\end{array} \right. \]  

(93)
as shown in Fig. 7. In the figure the adhesion formulation of example 6 is used. The parameter $r_m > 0$ denotes the distance where the transition from barrier or adhesion function to penalty method occurs. To ensure a repulsive force during penetration, the stiffness $F'_{ab}(r_m)$ must be negative. The contact energy $\Phi_\ell$ is now given by

$$\Phi_\ell = \begin{cases} F^*(r_m - r_{\text{lim}}) + \frac{1}{2} F'_{ab}(r_m^2 - r_{\text{lim}}^2) + \Phi_{ab}(r_m) & \text{if } g_n < r_{\text{lim}} , \\ F^*(r_m - g_n) + \frac{1}{2} F'_{ab}(r_m^2 - g_n^2) + \Phi_{ab}(r_m) & \text{if } r_{\text{lim}} < g_n < r_m , \\ \Phi_{ab}(g_n) & \text{if } g_n \geq r_m , \end{cases}$$

(94)

see Fig. 7.

Figure 7: Cross-constrained contact formulation for $r_{\text{lim}} = -1.5 r_0$ and $r_m = 0.7 r_0$: surface potential $\phi_s$ (left); contact force $F_k$ (middle); contact potential $\Phi_\ell$ (right). The dashed line at $r_m$ corresponds to the adhesion formulation of example 8.

10. Class S augmented Lagrange multiplier method: Exact enforcement of the impenetrability constraint. We consider the augmented Lagrange multiplier method based on Uzawa’s algorithm (Laursen, 2002; Wriggers, 2006). This algorithm iterates on the contact force $F_k$ according to

$$F_k^{\text{new}} = F_k^{\text{old}} + \Delta F_k$$

(95)

where $F_k^{\text{old}}$ is treated as a constant force and where $\Delta F_k$ is obtained from the penalty method. Hence, the Augmented Lagrange multiplier method is a simple combination of examples 4 and 5 above.

Finally, we note that one can also construct other contact force laws, based on these examples. In order to employ the analytical integration used to define these forces (see Sec. 2.3) it is important to ensure that $\phi_s$ is decaying. We also note that formulations 3 to 7, 9 and 10 can also be formulated as point interaction (class P) formulations. In particular, approach 9 can for example be used to combine the cohesive zone and class P penalty formulation.

3.3 Class 3: Long-range surface interaction

11. Electrostatic interaction: An example for long-range attraction or repulsion is Coulomb interaction between charged surfaces. In this case we have

$$\phi_s(r_k) = k_c \frac{q_1 q_2}{r_k} ,$$

(96)

where $k_c$ is Coulomb’s constant, and where $q_k$ denotes the charge of $da_k$ (i.e. $\beta^*_k g_k$ denotes the charge density at $da_k$). According to eq. (19) and (20), the electrostatic force at $x_k \in \partial B_k$, due
to $d\ell \subset \partial B_\ell$, is
\[ \frac{\partial \phi_k}{\partial x_k} = k_e \frac{q_1 q_2}{r_k^2} n_k, \tag{97} \]
The electrostatic force due to the entire surface $\partial B_\ell$ is then given by integration (23). The electromagnetic field at $x_k$, due to $\partial B_\ell$, then simply is
\[ E_k = \frac{F_k}{q_k}, \tag{98} \]
which satisfies the electrostatic Maxwell equations since it is derived from a potential (Shadowitz, 1988). Once the mechanical deformation is known, the electromechanical field surrounding $\partial B_\ell$ can be computed by a simple post-processing step, see Sec. 7.3. As pointed out before, in general, both $\phi_s$ and $\partial \phi_s/\partial x_k$ cannot be integrated analytically over $\partial B_\ell$ and numerical integration has to be considered, see Sec. 5.3. The integration can be carried out for special surfaces, like cylinders and spheres. An example is examined in Sec. 4.

Remark 1: If $\partial B_\ell$ is a flat homogeneously charged plane, the integration scheme of Sec. 2.3.2 can be used. According to example 4 above, the interaction force $F_k$ will then be constant. This corresponds to the known result for parallel-plate capacitors.

Remark 2: Assumption (7) implies that the charges don’t redistribute themselves during deformation. This is reasonable for non-conducting solids. For conducting solids, eq. (7) needs to be rewritten.

Remark 3: Potential (96) also describes gravitational attraction between hollow bodies. In this case $-k_e$ corresponds to the constant of gravity and $\beta_s q_k$ denotes the mass density at $d\ell_k$.

4 Analytical examples

This section presents two contact examples that demonstrate the formalism of the surface-potential-based contact model.

4.1 An illustrative thought experiment

The first example considers a simple thought experiment that illustrates the difference between short-range and long-range surface interaction, as they are described in Sec. 2. We therefore examine the contact between the two deformable bodies shown in Fig. 8. Body $B_1$, which has a circular surface between points $A$ and $B$, is considered to be rigid in the radial direction, but deformable in the circumferential direction. Body $B_2$, initially a rectangular block, is fully deformable and adapts to the shape of body $B_1$ during contact. The contact is supposed to be frictionless.

The thought experiment starts with the configuration shown in Fig. 8a, where the surface of $B_1$ is defined by the points $A$, $B$ and $C$. This system is supposed to be in static equilibrium. We now consider deforming $B_1$ in circumferential direction so that $A \rightarrow A'$ and $B \rightarrow B'$ while point $C$ remains fixed. This deformation is considered to be applied instantaneously. Body $B_2$ is supposed to react slowly, eventually reaching static equilibrium again. Two cases are considered:

1. The shape of $B_2$ does not change, i.e any further deformation of $B_2$ is negligible. This situation is shown Fig. 8a. In this case, the range of the interaction is short (class S), and the internal energy of $B_2$ as well as the contact energy in the system do not change.
Figure 8: Contact between a deformable block and a radially rigid circle: a. repulsive contact due to short-range surface interaction (class S); b. repulsive contact due to long-range surface interaction (class L).

2. $B_2$ deforms further, as shown in Fig. 8b, e.g. due to increased long-range repulsion (class L).

In response to the instantaneous deformation of $B_1$ the surface density $\beta_{s1}$ increases instantaneously. Let us now examine how this affects the two cases:

**Case 2: Long-range surface interaction** (Fig. 8b). At the first instance, $B_2$ remains unaffected. Therefore

- the surface density $\beta_{s2}$ remains unchanged
- the surface distance $g_k$ between $\partial B_1$ and $\partial B_2$ thus remains unchanged
- the potential $\phi_s$ remains unchanged
- the interaction energy $\Phi_2 = \int_{\partial B_2} \beta_{s2} \phi_s da_2$ remains unchanged
- the interaction energy $\Phi_1 = \int_{\partial B_1} \beta_{s1} \phi_s da_1$ increases along with $\beta_{s1}$
- the contact energy $\Pi_c = \int_{\partial B_1} \beta_{s1} \Phi_2 da_1 = \int_{\partial B_2} \beta_{s2} \Phi_1 da_2$ increases
- the force $F_1(x_1) = -\partial \Phi_2 / \partial x_1$ acting on $da_1 \subset \partial B_1$ remains unchanged
- the force $F_2(x_2) = -\partial \Phi_1 / \partial x_2$ acting on $da_2 \subset \partial B_2$ increases
- the increased contact force $F_2$ will push the surface $\partial B_2$ away so that the gap increases
- this, in turn, leads to a decrease of the magnitude of potential $\phi_s$ and a subsequent adjustment of $\Phi_1$, $\Phi_2$, $F_1$ and $F_2$
- for strong repulsion huge changes in $B_2$ can occur as indicated in Fig. 8b
- therefore, in this case, the work $\Pi_{ext}$ applied to $B_1$ is not only required to deform $B_1$, but also to deform $B_2$ and change $\Pi_c$. 
Case 1: Short range surface interaction (Fig. 8a). For case 1, the behavior is quite different. Now, $B_2$ remains completely unaffected. The first few steps are the same as for case 2:

- the surface density $\beta_2^s$ remains unchanged
- the distance $g_k$, potential $\phi_s$ and interaction energy $\Phi_2$ remain unchanged; interaction energy $\Phi_1$ increases along with $\beta_1^s$
- the force
  \[ F_1(x_1) = 2\pi \beta_2^s g_n \phi_s n_p(x_1) \]
  acting on $da_1 \subset \partial B_1$ remains unchanged
- the force
  \[ F_2(x_1) = 2\pi \beta_1^s g_n \phi_s n_p(x_2) \]
  acting on $da_2 \subset \partial B_2$ increases
- the contact traction
  \[ t_1 = \beta_1^s F_1 \]
  acting on $da_1$ increases and is equal to
  \[ t_2 = \beta_2^s F_2 \]
  acting on $da_2$
- but since $B_2$ is not deforming further, $t_1$ and $t_2$ cannot have changed; the tractions must remain constant as $B_1$ deforms; thus, an increase of $\beta_1^s$ must lead to a corresponding decrease in $g_n \phi_s(g_n)$
- this is achieved by changes in $g_n$; these changes must be very small, so that the changes in the deformation of $B_2$ are negligible;
- in that case, the changes in $\Pi_c$ are also negligible, and all external work is spent on the deformation of $B_1$.

4.2 A solved contact example

As a second example we present a simple analytical contact problem, which is solved with three of the contact models outlined in Sec. 3 – barrier method, penalty method and electrostatic interaction. The example consists of a solid rigid cylinder $B_1$ with radius $R_0$ and a deformable thin tube $B_2$ with initial (undeformed) mean radius $R$, where $R < R_0$. The surface densities $\beta_{01}^s$ and $\beta_{02}^s$ are considered constant across the bodies. The tube is stretched, such that its radius increases to $r > R_0$, and is then placed over the rigid cylinder as is shown in Fig. 9. The tube

![Figure 9: Contact interaction between a rigid cylinder and a deformable tube](image)

is supposed to obey the incompressible Neo-Hookean material model

\[ \sigma = \mu F_{\nu} F_{\nu}^T + qI, \]  

where $\sigma$ is the Cauchy stress, $\mu$ is the shear modulus, $F_{\nu}$ is the deformation gradient and $q$ is a Lagrange multiplier that is to be determined. The stretching of the tube radius is described by the parameter $\lambda > 1$, such that $r = \lambda R$. We consider a state of plane strain ($\lambda_z = 1$). Due to incompressibility, the stretches of the tube wall in circumferential and radial direction
then follow as \( \lambda_\theta = \lambda \) and \( \lambda_r = \lambda^{-1} \), i.e. the wall thickness \( T \) decreases to \( t = \lambda^{-1}T \). The deformation gradient for the tube thus is

\[
F_v = \begin{bmatrix}
\lambda^{-1} & 0 & 0 \\
0 & \lambda & 0 \\
0 & 0 & 1
\end{bmatrix}
\]  \hspace{1cm} (100)

in the \( \{r, \theta, z\} \) basis. The surface deformation is characterized by \( J_{s1} = 1 \) (rigid cylinder) and \( J_{s2} = \lambda \) (stretched tube). Thus the current surface densities are \( \beta^s_1 = \beta^s_{01} \) and \( \beta^s_2 = \beta^s_{02}/\lambda \). The normal gap between the solid cylinder and the hollow tube is \( g_n = r - R_0 \), or

\[
g_n/R_0 = \lambda R - 1 \ , \hspace{1cm} (101)
\]

where we have introduced \( \bar{R} := R/R_0 \).

Considering the standard membrane assumption \( \sigma_{rr} = 0 \), the Lagrange multiplier becomes \( q = -\mu \lambda^{-2} \). The pressure \( p \) acting on the inside of the tube, due to the interaction with the cylinder is related to the stress \( \sigma_{\theta\theta} \) according to the well known formula

\[
\sigma_{\theta\theta} = \frac{p r}{t} , \hspace{1cm} (102)
\]

where \( \sigma_{\theta\theta} = \mu \lambda^2 - \lambda^{-2} \) according to eq. (99). Combining the last two equations, we can rewrite the equilibrium equation in the dimensionless form

\[
\bar{p} (\lambda^2 - \lambda^{-2}) = \bar{p} \lambda^2 \ , \hspace{1cm} (103)
\]

with \( \bar{p} := p/p_0 \) and \( \bar{\mu} := \mu T/p_0 R \). This equation is now solved for the following interaction models (a) penalty method (class S), (b) barrier method (class S) and (c) electrostatic repulsion (class L).

(a) **Penalty Method**: The pressure between cylinder and tube, according to class S penalty interaction (see eq. (61) and (77)), is \( p = -2 \pi \beta_1^s \beta_2^s g_n \phi_0 \) which can be simply written as

\[
\lambda \bar{p} = 1 - \lambda \bar{R} \ , \hspace{1cm} (104)
\]

using eq. (101) and \( p_0 = 2 \pi \beta_{01}^s \beta_{02}^s \phi_0 R_0 \). Combining eq. (103) and (104) we find the equation

\[
(\bar{R} + \bar{\mu}) \lambda^4 - \lambda^3 - \bar{\mu} = 0 \ . \hspace{1cm} (105)
\]

The parameter \( \bar{\mu} \) is inversely proportional to the penalty parameter \( p_0 \). Taking \( \bar{\mu} \to 0 \) yields the exact limit value \( \lambda = 1/\bar{R} = R_0/R \). Table 2 below shows the physical solution of (105) for several other values of \( \bar{\mu} \).

<table>
<thead>
<tr>
<th>( \bar{\mu} )</th>
<th>penalty method</th>
<th>barrier method</th>
<th>electrostatics</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.001 ( \lambda = 1.996, \ p = 0.0469 \mu )</td>
<td>( \lambda = 45.73, \ p = 0.0500 \mu )</td>
<td>( \lambda = 22.36, \ p_2 = 0.0500 \mu )</td>
<td></td>
</tr>
<tr>
<td>0.01 ( \lambda = 1.963, \ p = 0.0466 \mu )</td>
<td>( \lambda = 15.18, \ p = 0.0500 \mu )</td>
<td>( \lambda = 7.073, \ p_2 = 0.0500 \mu )</td>
<td></td>
</tr>
<tr>
<td>0.1 ( \lambda = 1.701, \ p = 0.0440 \mu )</td>
<td>( \lambda = 5.585, \ p = 0.0499 \mu )</td>
<td>( \lambda = 2.279, \ p_2 = 0.0481 \mu )</td>
<td></td>
</tr>
<tr>
<td>1 ( \lambda = 1.219, \ p = 0.0193 \mu )</td>
<td>( \lambda = 2.742, \ p = 0.0491 \mu )</td>
<td>N.A.</td>
<td></td>
</tr>
<tr>
<td>10 ( \lambda = 1.013, \ p = 0.0024 \mu )</td>
<td>( \lambda = 2.100, \ p = 0.0474 \mu )</td>
<td>N.A.</td>
<td></td>
</tr>
<tr>
<td>100 ( \lambda = 1.001, \ p = 0.0002 \mu )</td>
<td>( \lambda = 2.011, \ p = 0.0469 \mu )</td>
<td>N.A.</td>
<td></td>
</tr>
<tr>
<td>1000 ( \lambda = 1.000, \ p = 0.0000 \mu )</td>
<td>( \lambda = 2.001, \ p = 0.0469 \mu )</td>
<td>N.A.</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Stretch \( \lambda \) and internal tube pressure \( p \) for various interaction models considering \( \bar{R} = 1/2 \) and \( R/T = 20 \).
(b) **Barrier Method**: The contact pressure according to class S barrier interaction (see eq. (61) and (71)) is

\[ p = 2\pi \beta_1 \beta_2 r_0^2 \phi_0 / g_n, \]

which can be simply written as

\[ \lambda \bar{p} = (\lambda \bar{R} - 1)^{-1}, \]

(106)

using eq. (101) and \( p_0 = 2\pi \beta_{01} \beta_{02}^2 \phi_0 R_0 \) and \( r_0 = R_0 \). Combining (106) with (103) leads to the equation

\[ (\lambda^4 - 1)(\lambda \bar{R} - 1) - \lambda^3 / \bar{\mu} = 0, \]

(107)

which can be solved for \( \lambda \) (see table 2). In the limit \( \bar{\mu} \to \infty \) we approach the exact answer \( \lambda = R_0 / R \).

(c) **Electrostatic interaction**: The electrical field \( E_\ell \) inside a charged hollow cylinder \( B_\ell \) is zero \( (E_\ell = 0) \), and the electric field outside a cylinder \( B_\ell \), charged on the surface by \( \beta_\ell^e q_\ell \), is inversely proportional to the distance from the center of the cylinder\(^9\), i.e.

\[ E_\ell = 4k_e \beta_\ell^e q_\ell / r \]

(108)

(Shadowitz, 1988). The force acting on the area element \( d a_k \), which has the charge \( q_k \), is \( F_k = q_k E_\ell \) (\( \ell \neq k \)), while the traction is \( t_k = \beta_k^e F_k \). The surface pressures on the cylinder \( B_1 \) and tube \( B_2 \) are thus given by \( p_1 = 0 \) and \( p_2 = 4k_e \beta_1^e \beta_2^e q_1 q_2 / r \). The last equation can be rewritten as

\[ \lambda^2 \bar{p} = \bar{R}, \]

(109)

with \( \bar{p} = p_2 / p_0 \) and \( p_0 = 4k_e \beta_{01} \beta_{02}^e q_1 q_2 / R_0 \). Combining eqs. (109) and (103) we find the (physical) solution

\[ \lambda = \sqrt{\bar{R} + \sqrt{\bar{R}^2 + 4\bar{\mu}^2}} / 2\bar{\mu}, \]

(110)

which is included in table 2. Since we are not allowing for contact in case (c) we must restrict \( \bar{\mu} \) such that \( \lambda > R_0 / R \). In the case of electrostatic interaction, the pressure on both surfaces are not equal as is the case for the penalty and barrier method (as long as \( g_n \) is small). This reiterates that class S, but not class L, satisfy local equilibrium.

5 The two-half-pass contact algorithm

This section presents suitable contact algorithms for the three interaction classes mentioned above. These algorithms fall into a special category, which is termed the two-half-pass contact algorithm. We first examine the general idea of this approach and then focus on the particular expressions for the different interaction classes.

5.1 The general idea

Both bodies \( B_k \) (\( k = 1, 2 \)) and their surfaces \( \partial B_k \) are partitioned into a set of volume elements \( \Omega_k^e \) and surface elements \( \Gamma_k^e \). Each of those is supposed to have \( n_k^e \) nodes. Within each finite

\(^9\)The electric fields are obtained from integrating (96) over the cylinder surface, see Shadowitz (1988).
element \((x_k \in \Omega_k^e \text{ or } x_k \in \Gamma_k^e)\), the displacement field \(u_k\) of body \(B_k\) and its variation \(\delta \varphi_k\) are approximated by the standard FE interpolation\(^{10}\)

\[
    u_k^e(x_k) = N_k(x_k) u_k^e, \quad \delta \varphi_k^e(x_k) = N_k(x_k) v_k^e, \quad (111)
\]

where \(u_k^e\) and \(v_k^e\) are arrays with size\(^{11}\) \(\left(n_{\text{dim}} n_k^e \times 1\right)\) that denote the displacements and variations of the elemental nodes, and where

\[
N_k = \begin{bmatrix} N_1 \mathbf{I} & N_2 \mathbf{I} & \ldots & N_{n_k^e} \mathbf{I} \end{bmatrix} \quad (112)
\]

is a \(\left(n_{\text{dim}} \times n_{\text{dim}} n_k^e\right)\) matrix formed by the \(n_k^e\) shape functions \(N_I\) \((I = 1, 2, \ldots, n_k^e)\) of the element. With interpolation (111), applied to each finite element, the weak form (2) can be rewritten as

\[
v^T [f_{\text{int}} + f_c - f_{\text{ext}}] = 0, \quad \forall v \in V^h. \quad (113)
\]

Here the vector \(v\) contains the virtual displacements of all the finite element nodes and the vectors \(f_{\text{int}}, f_c\) and \(f_{\text{ext}}\) contain the internal forces, contact-interaction forces and external forces acting on all the finite element nodes. Further, \(V^h\) denotes the space of admissible virtual displacements. Expressions for \(f_{\text{int}}\) and \(f_{\text{ext}}\), as well as an extension to the dynamic case, can be found in standard nonlinear FEM textbooks, e.g. see Wriggers (2008). The contact problem is solved by solving the nonlinear equation

\[
f(u) = f_{\text{int}} + f_c - f_{\text{ext}} = 0. \quad (114)
\]

The force contributions appearing here are assembled from the corresponding forces acting on the individual finite elements. According to eqs. (7), (22) and definitions (25), (26), the contact-interaction force acting on surface element \(\Gamma_k^e \in \partial B_k^h\) is given by

\[
f_{ck}^e := - \int_{\Gamma_k^e_{\partial B_k}} N_k^T T_k \, dA_k = - \int_{\Gamma_k^e} N_k^T t_k \, d\alpha_k. \quad (115)
\]

This expression is evaluated equivalently for both surfaces \((k = 1, 2)\). This procedure is referred to as \textit{two half-passes}. A pass over the two contact surfaces is used, but in each case only the contact traction \(t_k\) is considered. This is different from conventional one- and two-pass contact algorithms. The standard procedure for one-pass approaches is to account for \(t_k\) on one surface, say \(\partial B_k = \partial B_1\), and then account for \(-t_k\) on the neighboring surface \(\partial B_2\). This procedure involves additional effort in the integration of neighboring surface quantities. Further, it creates a bias of one surface over the other. A classical two-pass approach considers using the one-pass approach twice, once for each surface. This removes the bias, but doubles the amount of computational work that is needed to evaluate \(f_c\), and can also lead to the over-constraining of the contact problem. In case of the proposed two-half-pass approach the computational work is similar to that of the classical one-pass approach, while there is no bias and the integration procedure is less complicated. The absence of \(-t_k\) applied to \(\partial B_1\) implies that no numerical difficulties can arise from this. The two-half pass approach therefore turns out to be more robust than classical one-pass approaches, as the examples in Sec. 7 show. An apparent drawback of the two-half-pass algorithm, in case of classes S and P, is that the local balance of linear momentum, and thus also local angular momentum, are not satisfied exactly in the discrete

\(^{10}\)In the following description, \(u_k^e\) and \(u_k^h\) are used to denote the displacement field and its FE approximation, \(u_k^e\) is used to denote the stacked vector of all nodal displacements of element \(e\), and \(u\) is used to denote the stacked vector of all nodal displacements of the two discretized bodies \(B_1^h\) and \(B_2^h\). Analogous definitions are used for the reference configuration and the current configuration (characterized by the vector fields \(X\) and \(x\)). Note that italic font is used for field variables and normal font is used for discrete variables.

\(^{11}\)\(n_{\text{dim}} = 2\) in 2D and \(n_{\text{dim}} = 3\) in 3D.
Further, we find \( \partial \) or (26). Since eq. (34) or (27). The finite element force vector is then given by expression (115), using eq. (25) as

As is shown in eqs. (39) and (38), the distance \( g \) linear, 2D elements simplified expressions follow, see Appendix E.

The gradients of \( \partial \) quantities depend on point \( x \) which characterize the change in \( \partial \). In eq. (118b) we have

For both class S and class P, the contact-interaction force \( F_k \) is given analytically, either through eq. (34) or (27). The finite element force vector is then given by expression (115), using eq. (25) or (26). Since \( f^c_{ek} \) is defined on surface element \( \Gamma^e_k \), it depends on the nodal displacements \( u^e_k \) of this element. Due to the dependency on the contact gap between \( \partial B_k \) and \( \partial B_{\ell} \), the force vector \( f^c_{ek} \) also depends on the displacements \( u^e_{\ell} \) of the surface element \( \Gamma^e_{\ell} \). The linearization of \( f^c_{ek} \) w.r.t. these displacements thus follows as

\[
f^c_{ek}(u^e_k + \Delta u^e_k, u^e_{\ell}) \approx f^c_{ek}(u^e_k, u^e_{\ell}) + \Delta f^c_{ek}(u^e_k, u^e_{\ell}) ,
\]

with

\[
\Delta f^c_{ek}(u^e_k, u^e_{\ell}) = \frac{\partial f^c_{ek}}{\partial u^e_k} \Delta u^e_k + \frac{\partial f^c_{ek}}{\partial u^e_{\ell}} \Delta u^e_{\ell} .
\]

In here we can identify the two stiffness matrices

\[
k^c_{kk} := \frac{\partial f^c_{ek}}{\partial u^e_k} = - \int_{\Gamma^e_k} N^e_k \frac{\partial T^e_k}{\partial u^e_k} dA_k ,
\]

\[
k^c_{\ell\ell} := \frac{\partial f^c_{ek}}{\partial u^e_{\ell}} = - \int_{\Gamma^e_{\ell}} N^e_{\ell} \frac{\partial T^e_{\ell}}{\partial u^e_{\ell}} dA_{\ell} ,
\]

which characterize the change in \( f^c_{ek} \) due to changes of surfaces \( \partial B^L_k \) and \( \partial B^L_{\ell} \). If \( B_k \) is considered rigid and immobile, matrix \( k^c_{kk} \) is zero. In general \( T_k \) can depend on the distance \( g_n \), the direction \( n_p \) and the surface stretch \( J_p := \beta_{nk}/\beta_{nk} = \partial a_k/\partial A_{\ell} \) according to eq. (34). Since these quantities depend on point \( x_k \in \partial B_k \) and surface \( \partial B_{\ell} \), which is characterized by the nodal points \( x_{\ell} \), we can write \( T_k = T_k(g_n, n_p, J_p) = T_k(x_k, x_{\ell}) \). In eq. (118a), we thus have

\[
\frac{\partial T_k}{\partial u^e_k} = \frac{\partial T_k}{\partial x_k} \frac{\partial x_k}{\partial u^e_k} = \frac{\partial T_k}{\partial x_k} N_k ,
\]

due to eq. (111). In eq. (118b) we have

\[
\frac{\partial T_k}{\partial x_k} = \frac{\partial T_k}{\partial x_k} = \left[ \frac{\partial T_k}{\partial x_1}, \frac{\partial T_k}{\partial x_2}, \ldots, \frac{\partial T_k}{\partial x_n} \right] .
\]

Further, we find

\[
\frac{\partial T_k}{\partial g_n} = \frac{\partial T_k}{\partial g_n} \frac{\partial g_n}{\partial x_k} + T_k \frac{\partial n_p}{\partial x_k} + \frac{\partial T_k}{\partial J_p} \frac{\partial J_p}{\partial x_k} ,
\]

\[
\frac{\partial T_k}{\partial x_{\ell}} = \frac{\partial T_k}{\partial x_{\ell}} \frac{\partial x_{\ell}}{\partial x_k} + T_k \frac{\partial n_p}{\partial x_{\ell}} + \frac{\partial T_k}{\partial J_p} \frac{\partial J_p}{\partial x_{\ell}} .
\]

The gradients of \( g_n \), \( n_p \) and \( J_p \) w.r.t. \( x_k \) and \( x_{\ell} \) are derived in Appendices B, C and D. For linear, 2D elements simplified expressions follow, see Appendix E.

As is shown in eqs. (39) and (38), the distance \( g_n \) and the surface normal \( n_p \) can be expressed as

\[
g_n = \pm \sqrt{(x_k - x_p) \cdot (x_k - x_p)} ,
\]
and

\[ n_p = \frac{x_k - x_p}{g_n} . \]  

(123)

From this, one can find

\[ \frac{\partial g_n}{\partial x_k} = n_p \]  

(124)

and

\[ \frac{\partial g_n}{\partial x_\ell} = -n_p N_\ell(x_p) , \]  

(125)

see Appendix C. For class P, \( T_k \) is independent of \( J_p \), so that the last term in eq. (121) vanishes. For class S, the contact traction \( T_k = \beta_{0k} F_k \) is proportional to \( J_p^{-1} \) (see eq. (35)). Therefore

\[ \frac{\partial T_k}{\partial J_p} = -\frac{T_k}{J_p} , \]  

(126)

By substitution of eqs. (119), (120) and (124)–(126), eq. (121) can be rewritten as

\[ \frac{\partial T_k}{\partial u_k} = T_k' n_p \otimes n_p N_k + T_k \frac{\partial n_p}{\partial x_k} N_k - \frac{T_k}{J_p} \otimes \frac{\partial J_p}{\partial x_k} N_k , \]  

\[ \frac{\partial T_k}{\partial u_\ell} = -T_k' n_p \otimes n_p N_\ell + T_k \frac{\partial n_p}{\partial u_\ell} - \frac{T_k}{J_p} \otimes \frac{\partial J_p}{\partial u_\ell} , \]  

(127)

with \( T_k' := \partial T_k/\partial g_n \) and where \( N_\ell \), also defined by eq. (112), is evaluated at \( x_p \). Here \( \partial n_p/\partial x_k \) and \( \partial J_p/\partial x_k \) are given in eqs. (180) and (194). They are tensors with dimensions \( n_{\text{dim}} \times n_{\text{dim}} \) and \( n_{\text{dim}} \times 1 \), respectively. Since \( N_\ell \) is an array with dimension \( n_{\text{dim}} \times n_{\text{dim}} n_k^\ell \), the dimension of \( \partial T_k/\partial u_\ell \) also becomes \( n_{\text{dim}} \times n_{\text{dim}} n_k^\ell \).

Derivatives \( \partial n_p/\partial u_\ell \) and \( \partial J_p/\partial u_\ell \) are given in eqs. (186) and (204). Their dimensions are such that \( \partial T_k/\partial u_\ell \) has dimension \( n_{\text{dim}} \times n_{\text{dim}} n_k^\ell \).

The projection direction \( n_p \) and its derivatives depend on the location of the closest projection point \( x_p \). In some cases \( x_p \) falls on an element boundary. This is then typically such that direction \( n_p \) given in eq. (123) is not normal to \( \partial B_k^\alpha \). In 3D, two such cases need to be distinguished: (i) projection onto an element corner, and (ii) projection onto the element edge. The last case reduces to the projection onto a line element. Both cases are contained as special cases in the formulas derived in Appendix C. For example, for case (i), the surface curvatures \( \kappa_\alpha \) approach infinity, such that

\[ \frac{\partial n_p}{\partial x_k} = \frac{1}{g_n} (I - n_p \otimes n_p) , \quad \text{for} \quad \kappa_\alpha = \infty , \]  

(128)

and

\[ \frac{\partial n_p}{\partial u_\ell} = -\frac{1}{g_n} (I - n_p \otimes n_p) N_\ell , \quad \text{for} \quad \kappa_\alpha = \infty , \]  

(129)

since \( c^{\alpha\beta} = 0 \) in eqs. (180) and (186). Note that \( N_\ell \) now only contains one non-zero entry. At the element boundaries, the surface stretch \( J_p \) is typically not unique either: the stretch jumps across the boundary. In this case one can consider averaging \( J_p \) between adjacent elements. Note, that all these distinctions for \( n_p \) and \( J_p \) are unnecessary for \( C^1 \)-smooth surface discretizations, such as those achieved through smoothing techniques (Krstulovic-Opara et al., 2002), Hermite-based surface enrichment (Sauer, 2011) and NURBS-based discretizations (Temizer et al., 2011; De Lorenzis et al., 2011, 2012).

\footnotetext[12]{Note that for \( C^0 \) FE discretizations the surface normal at the element boundaries is not even defined uniquely.}
For class P interaction, the traction \( T_k \) depends on the points \( x_k \in \partial B_k \) and \( x_\ell^0 \in \partial B_\ell \) but is not a function of the surface stretch \( J_p \), i.e. \( T_k = T_k(x_k, x_\ell^0) \) according to eqs. (26) and (27). The last terms in eq. (127) thus vanish. In the case where \( x_\ell^0 \) is taken as the closest projection point of \( x_k \) (e.g. for class P penalty method), the other terms in eq. (127) are evaluated as specified in eqs. (180) and (186). In the case where \( x_\ell^0 \) is a material point of \( \partial B_\ell \) (e.g. for class P cohesive zone formulation), the dependency on the surface curvatures \( \kappa_\alpha \) disappears, i.e. we have the case given in eqs. (128) and (129).

Table 3 gives a summary of the two-half-pass contact algorithm for class S and class P interaction.\(^\text{13}\) The algorithm works both for the cases where \( B_\ell \) is deformable or rigid. In the latter

---

Table 3: The two-half-pass contact algorithm for class P and class S interaction.

---

\(^\text{13}\)In the table, czm stands for the class P cohesive zone model.
For each discretized body, the net contact-interaction forces can be computed from

\[ P_h^k = \sum_{\Gamma^e_k \in \partial B_h^k} f_{ce}^k. \] (130)

Contrary to the continuum case (reported in Sec. 2.5), these forces do not satisfy global contact equilibrium exactly, i.e. \( \Delta P_h := P_h^1 + P_h^2 \) does not fully vanish. In fact, the error \( \Delta P_h \) can be used as an error measure of the accuracy of the two-half-pass formulation, see Sec. 7.1. \( \Delta P_h \) must vanish as the FE mesh size goes to zero. It can also decrease with increasing penalty parameter, as is seen in Sec. 7.1.

### 5.3 Class L interactions

For class L surface interaction, the force \( F_k \) is defined through integration (23), which, in general, needs to be carried out numerically. The procedure for this is analogous to the procedure used for volume interactions (Sauer and Li, 2007b). Combining eq. (23) with eq. (22) and applying eq. (111), we find the FE force vector of surface element \( \Gamma^e_k \in \partial B_k \) as

\[ f_{ce}^k = \sum_{\Gamma^e_\ell \in \partial B^\ell} f_{ce}^{\ell k}, \] (131)

with

\[ f_{ce}^{\ell k} := \int_{\Gamma^e_1} \int_{\Gamma^e_2} {\mathbf{N}}_k^T \beta_1 \beta_2 \frac{\partial \phi_s}{\partial \mathbf{x}_k} \partial \phi_s \partial \mathbf{x}_\ell \partial \mathbf{x}_\ell \partial \mathbf{a}_2 \partial \mathbf{a}_1. \] (132)

Here the gradient of \( \phi_s \) can be written as

\[ \frac{\partial \phi_s}{\partial \mathbf{x}_k} = -f_s(g_k) \mathbf{n}_k, \quad \text{for} \quad f_s = -\frac{\partial \phi_s}{\partial g_k}, \] (133)

where \( g_k \) and \( \mathbf{n}_k \) denote the signed distance, and corresponding direction, between points \( \mathbf{x}_k \in B_k \) and \( \mathbf{x}_\ell \in B_\ell \) (see eqs. (14) and (15)). The integration in (132) can be easily rewritten with respect to the reference element domains \( \Gamma^e_0 \) and \( \Gamma^e_0 \) using eq. (7). But note, that \( \phi_s(g_k) \) is always evaluated for the distance \( g_k \) in the current configuration. The force \( f_{ce}^{\ell k} \) (131) depends on the deformation of both surface elements \( \Gamma^e_k \in \partial B_k \) and \( \Gamma^e_\ell \in \partial B_\ell \) (\( \ell \neq k \)), so that we find the two tangent contributions

\[ k_{ce}^{\ell k} = \frac{\partial f_{ce}^{\ell k}}{\partial \mathbf{u}_k} = \sum_{\Gamma^e_\ell \in \partial B^\ell} k_{ce}^{\ell k}, \]

\[ k_{ce}^{\ell k} = \frac{\partial f_{ce}^{\ell k}}{\partial \mathbf{u}_\ell} = k_{ce}^{\ell k}, \] (134)

with

\[ k_{ce}^{\ell k} := \int_{\Gamma^e_1} \int_{\Gamma^e_2} {\mathbf{N}}_k^T \beta_1 \beta_2 \frac{\partial^2 \phi_s}{\partial \mathbf{x}_k \partial \mathbf{x}_k} \mathbf{N}_k \partial \mathbf{a}_2 \partial \mathbf{a}_1, \]

\[ k_{ce}^{\ell k} := \int_{\Gamma^e_1} \int_{\Gamma^e_2} {\mathbf{N}}_k^T \beta_1 \beta_2 \frac{\partial^2 \phi_s}{\partial \mathbf{x}_k \partial \mathbf{x}_\ell} \mathbf{N}_\ell \partial \mathbf{a}_2 \partial \mathbf{a}_1, \] (135)

and

\[ \frac{\partial^2 \phi_s}{\partial \mathbf{x}_k \partial \mathbf{x}_k} = \frac{f_s(g_k)}{g_k} I + \left[ f'_s(g_k) - \frac{f_s(g_k)}{g_k} \right] \mathbf{n}_k \otimes \mathbf{n}_k, \]

\[ \frac{\partial^2 \phi_s}{\partial \mathbf{x}_k \partial \mathbf{x}_\ell} = -\frac{\partial^2 \phi_s}{\partial \mathbf{x}_k \partial \mathbf{x}_\ell}. \] (136)
The algorithm for class L interaction is similar to the algorithm in Table 3. For step 2, we now simply need a double loop over the surface elements $\Gamma_e^1$ and $\Gamma_e^2$ to evaluate the expressions above. This is straightforward to implement, but is less efficient than the formulation of Sec. 5.2. Careful attention must be paid to problems that exploit symmetries (Sauer and Li, 2007b). A numerical example on long-range electrostatic interaction is discussed in Sec. 7.3.

It can be easily seen that the above formulation satisfies global contact equilibrium exactly: Since $\partial \phi_s/\partial x_\ell = -\partial \phi_s/\partial x_k$, the global contact forces $P_h^1$ and $P_h^2$, defined by eq. (130), satisfy $P_h^1 + P_h^2 = 0$ exactly, even for very coarse meshes.

6 The contact patch test

This section shows that the proposed contact algorithm based on the two half-passes satisfies the contact patch test to machine precision. The contact patch test was first proposed by Taylor and Papadopoulos (1991) and later revisited by Crisfield (2000), El-Abbasi and Bathe (2001) and Zavarise and De Lorenzis (2009a). Here a further modification is considered that generalizes the test.

The test consists of a smaller block $B_1$ in frictionless contact with a larger block $B_2$, see Fig. 10. The original test considered equal material behavior; here we will generalize to the case of different materials. A non-linear, isotropic, Neo-Hookean material model is considered in the form (Zienkiewicz and Taylor, 2005)

$$\sigma = \frac{\Lambda}{J_v} (\ln J_v) I + \frac{\mu}{J_v} (F_v F_v^T - I),$$

(137)

where $J_v$ is the determinant of the deformation gradient $F_v = \text{Grad} \varphi$. $\Lambda = 2\mu\nu/(1 - 2\nu)$ and $\mu = E/(1 + \nu)$ correspond to the bulk and shear moduli (in the linear elastic regime). Here $E$ is Young’s modulus and $\nu$ is Poisson’s ratio. The values $E_1 = E_0$ and $\nu_1 = 0.4$ for $B_1$, and $E_2 = 2E_0$ and $\nu_2 = 0.1$ for $B_2$ are chosen. Due to these elastic constants, relative tangential sliding motion occurs during contact, which leads to a change of the contact area.
during computation. Contact occurs due to a pressure $p$ that is applied to the free parts of the upper surfaces of both blocks. The quantity $p$ refers to the actual pressure on the deformed surfaces. 2D, bi-linear finite elements are considered.

Fig. 11 shows the solution obtained from the two-half-pass algorithm. First the class P penalty

![Image](image_url)

Figure 11: Patch test solution for an applied pressure of $p = 0.8E_0$ according to the two-half-pass contact algorithm: a. Using the penalty method with $\epsilon_n = 100 \frac{E_0}{L_0}$ and 3 Gauss points per contact element; b. Using one additional augmentation step, during which $B_2$ is considered frozen. The color shows the numerical error in $\sigma_{yy}/E_0$ compared to the exact solution.

method is used and a small penetration appears (about $0.01L_0$), but the stress field (and thus the relative deformation) is exact within machine precision. Then a single augmentation step is applied using the Uzawa algorithm (Wriggers (2006); Laursen (2002), see also Sec. 3.2.10) which brings the penetration down to about $2 \cdot 10^{-14}L_0$, while the stress error increases to the level of $10^{-13}$. This solution is still highly accurate but does not quite reach machine precision. A further reduction of the penetration and the stress error could not be achieved by applying further augmentation steps, possibly due to overconstraining present in the two half-passes. To circumvent this problem, block $B_2$ can be frozen during the augmentation steps. This essentially corresponds to only considering the constraints on $\partial B_1$ active while ignoring the constraints of $\partial B_2$. With only one augmentation step we thus achieve accuracy within machine precision, both for the stress (see Fig. 11b) and for the penetration ($\sim 2 \cdot 10^{-16}L_0$).

For comparison, we discuss the performance of the conventional full-pass contact algorithm. The generalized patch test (with $\nu_1 \neq \nu_2$) is not passed by the full-pass contact algorithm. This is in agreement with the initial observations of Taylor and Papadopoulos (1991). The two-half-pass algorithm, on the other hand, passes the generalized patch test, even when the constraint is not enforced exactly, e.g. when the penalty method is used. Also, we have observed that the full-pass algorithm (now for $\nu_1 = \nu_2$) behaves much less robust for large $p$ than the two-half-pass algorithm. Further we have observed that for the full-pass algorithm many more augmentation steps, even for large $\epsilon_n$, are required to bring the stress error and the penetration down to machine precision.

In summary, we can conclude that the unbiased, two-half-pass approach passes the generalized patch test whereas the classic full-pass approach does not. For the two-half-pass algorithm the relative deformation and stress are exact even when using the penalty method. A critical aspect that remains is the issue of overconstraining. Jones and Papadopoulos (2001) and Solberg et al.
(2007) suggest a patterning of the constraint points to remove the problem for their two-pass approach (which also satisfies the patch test). Alternative venues may be found by using mortar approaches. This aspect will be considered in future work.

7 Numerical examples

This section presents several numerical examples to show the versatility of the proposed formulation. The first example considers 2D sliding contact between deformable bodies and reports a detailed analysis of the new algorithm. The other examples consider 3D normal contact, electrostatic attraction and cohesive debonding. The examples demonstrate the fulfillment of traction continuity across the contact interface, a property that is not in-build a priori in the algorithm.

7.1 Ironing problem

7.1.1 Problem setup

We consider the ironing problem (Yang et al., 2005; Fischer and Wriggers, 2006) with frictionless contact using the setup of Sauer (2012). The problem consists of a deformable half-cylinder $B_1$, with initial radius $L_0$, pressed into a deformable block $B_2$, with initial dimension $2L_0 \times 10L_0$, and then moved across tangentially, see Fig. 12. Here, the bottom is fixed and periodic boundary conditions are applied on the sides. The Neo-Hookean material model (137) is used with the parameters $E_1 = 3E_0$, $E_2 = E_0$ and $\nu_1 = \nu_2 = 0.3$ for the two bodies. Parameters $E_0$ and $L_0$ are used for normalization and remain unspecified. The class S penalty method is chosen to model contact and the contact algorithm given in table 3 is used for the computations. 50 equidistant quadrature points are used for integrating (115). According to the class S penalty method, the contact traction is given by

\[ T_k = \begin{cases} -\epsilon_n g_n J_p^{-1} n_p, & g_n < 0, \\ 0, & g_n \geq 0, \end{cases} \tag{138} \]

Figure 12: Ironing problem: frictionless sliding contact between a deformable cylinder and a block. The coloring shows the stress $I_1 = \text{tr} \sigma$ normalized by $E_0$. 

33
where $\epsilon_n$ is the penalty parameter. Values between $\epsilon_n = 10 E_0 / L_0$ and $\epsilon_n = 1000 E_0 / L_0$ are considered. For the finite element refinement, two discretizations are considered: $n_{el} / L_0 = 4$ (4 elements per $L_0$ on $\partial B_2$), as in Fig. 12 and $n_{el} / L_0 = 8$.

Remark: Large penalty parameters combined with a large number of quadrature points can lead to severe ill-conditioning of the contact problem. For such parameters, the use of $C^1$-smooth surface representations can substantially reduce ill-conditioning, as is shown in Sauer (2012).

### 7.1.2 Contact pressure

We first examine the contact pressure and its continuity across the contact surface. This is shown in Fig. 13. The pressure field $p_k$ along the surface coordinate $s_k$ is obtained from linear interpolation of the nodal pressure values (Sauer, 2012)

$$
p_{kI} := \int_{\partial_k B_{0k}^h} N_{kI} t_k \cdot \mathbf{n}_p \, dA_k \over \int_{\partial_k B_{0k}^h} N_{kI} \, dA_k,
$$

(139)

where $t_k = T_k / J_{sk}$, as before. Here, $\partial_k B_{0k}^h$ denotes the discretized surface region where contact forces are present. From the two passes ($k = 1, 2$) we thus obtain the pressure on both surfaces. As Fig. 13.b shows, the pressure fields $p_1$ and $p_2$ are in excellent agreement. This property confirms that the two-half-pass formulation satisfies local contact equilibrium. Note that the pressure field defined by eq. (139) is smoother than the raw pressure at the quadrature points, which can show strong oscillations (De Lorenzis et al., 2011; Sauer, 2011; Temizer et al., 2011). This is due to the averaging that is used in definition (139). A comparison between the raw and smoothed pressure on both surfaces is shown in Fig. 14. Note that the raw pressure on both surfaces are almost identical, even though they are obtained from two different passes. The

---

14 The averaging can also be defined over the current surface $\partial_k B_{0k}^h$ which leads to slightly different nodal pressures due to the surface stretch.
Figure 14: Comparison of the raw and smoothed contact pressures obtained on the cylinder ($k = 1$, left) and block ($k = 2$, right). The smoothed pressure curves are the same as shown in Fig. 13b.

post-processed pressure (139) becomes even smoother if $C^1$-continuous surfaces interpolation is used (Sauer, 2012).

The unbalance between the pressures $p_1$ and $p_2$ can be examined by looking at the net contact forces $P_k = \| P_k \|$ acting on $B_k$, which are computed according to eq. (130). This is shown in table 4. As seen, the sum $\Sigma P$ approaches zero as $\epsilon_n$ increases. Further note that the equilibrium error $\Sigma P$ is smaller than the inaccuracy of the penalty method, which is expressed by $\Delta P_k$.

<table>
<thead>
<tr>
<th>$\epsilon_n$</th>
<th>$P_1$</th>
<th>$P_2$</th>
<th>$\Sigma P$</th>
<th>$\Delta P_1$</th>
<th>$\Delta P_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.6523</td>
<td>-0.6224</td>
<td>0.0298</td>
<td>0.0357</td>
<td>0.0649</td>
</tr>
<tr>
<td>30</td>
<td>0.6750</td>
<td>-0.6640</td>
<td>0.0110</td>
<td>0.0130</td>
<td>0.0233</td>
</tr>
<tr>
<td>100</td>
<td>0.6833</td>
<td>-0.6799</td>
<td>0.0034</td>
<td>0.0047</td>
<td>0.0074</td>
</tr>
<tr>
<td>300</td>
<td>0.6863</td>
<td>-0.6848</td>
<td>0.0015</td>
<td>0.0017</td>
<td>0.0025</td>
</tr>
<tr>
<td>1000</td>
<td>0.6880</td>
<td>-0.6873</td>
<td>0.0007</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Normal contact forces of the two-half-pass contact algorithm for increasing penalty parameter $\epsilon_n$ using $n_{el}/L_0 = 8$. $P_k$ denotes the net contact force acting on $B_k$. $\Sigma P$ denotes the equilibrium error between $P_1$ and $P_2$; $\Delta P_k$ denotes the difference between $P_k$ and the values of the last row. All forces are normalized by $E_0L_0$; $\epsilon_n$ is normalized by $E_0/L_0$.

7.1.3 Comparison with conventional full-pass: Pressure

Next, we compare the new contact formulation with the full-pass formulation of Fischer and Wriggers (2005). We first focus on the contact forces and pressure. Section 7.1.4 examines robustness and efficiency. For the full-pass approach the FE contact forces on slave, $\partial_c B_k$, and master surface, $\partial_c B_\ell$, are obtained as

$$ f_{ck}^e = - \int_{\Gamma_k^e} N_k^T t_k \, da_k , \quad f_{\ell}^e = \int_{\Gamma_\ell^e} N_\ell^T t_k \, da_k , $$

where the shape function matrix $N_\ell$ is evaluated at the projection point $x_p \in \partial_c B_\ell$. $f_{ck}^e$ and $f_{\ell}^e$ are computed from a single pass over the contact elements $\Gamma_k^e \subset \partial_c B_k^e$. The traction $t_k$ is
obtained from the classical penalty method (class P) as

$$t_k = \begin{cases} 
-\epsilon_n g_n n_p, & g_n < 0, \\
0, & g_n \geq 0.
\end{cases} \quad (141)$$

Fig. 15 shows the contact pressure according to the full-pass approach. As before, the pressure

![Figure 15: Contact pressures](image)

fields $p_1$ and $p_2$, acting on surfaces $\partial B_1$ and $\partial B_2$, are obtained from linear interpolation of nodal pressures. For the slave surface, the nodal pressures are obtained from definition (139). For the neighboring master surface, the nodal pressures are computed, in accordance to eq. (140b), from

$$p_{II} := \frac{\int_{\partial_c B_{h_k}} N_k t_k \cdot n_p \, dA_k}{\int_{\partial_c B_{h_k}} N_k \, dA_k} \quad . \quad (142)$$

Figs. 13 and 15 show that there is hardly any difference between the pressure fields obtained from the one-half-pass algorithm and those obtained by the two-half-pass algorithm. The similarity of the smoothed pressure comes from the similarity of the raw pressures: Those are obtained similarly when either projection onto the cylinder or onto the block surface is considered. For a more differentiated picture, we compare the net contact forces. Table 5 shows these contact forces according to the full-pass contact algorithm. As seen, the contact forces for the full-pass algorithm lie in-between those of the two-half-pass algorithm (table 4). Compared to the exact result ($P = 0.688$) attained for high $\epsilon_n$ the two-half-pass converges faster than the full-pass results.

### 7.1.4 Comparison with conventional full-pass: Robustness and efficiency

We now examine the robustness and efficiency during indentation (downward motion of the cylinder) and sliding (tangential motion of the cylinder). We first consider the indentation phase. Table 6 shows that the two-half-pass algorithm is more robust and efficient than the full-pass algorithm. The table gives the computational speed and the number of Newton iteration
Table 5: Normal contact forces of the full-pass contact algorithm for increasing penalty parameter $\epsilon_n$ using $n_{el}/L_0 = 8$. $P_k$ denotes the force acting on $B_k$, considering $B_k$ as the slave body; $\Delta P_k$ denotes the difference between $P_k$ and the values of the last row. All forces are normalized by $E_0 L_0$; $\epsilon_n$ is normalized by $E_0/L_0$.

<table>
<thead>
<tr>
<th>$\epsilon_n$</th>
<th>$P_1$</th>
<th>$P_2$</th>
<th>$\Delta P_1$</th>
<th>$\Delta P_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.6314</td>
<td>0.6251</td>
<td>0.0561</td>
<td>0.0623</td>
</tr>
<tr>
<td>30</td>
<td>0.6668</td>
<td>0.6649</td>
<td>0.0207</td>
<td>0.0225</td>
</tr>
<tr>
<td>100</td>
<td>0.6807</td>
<td>0.6802</td>
<td>0.0068</td>
<td>0.0072</td>
</tr>
<tr>
<td>300</td>
<td>0.6852</td>
<td>0.6849</td>
<td>0.0023</td>
<td>0.0025</td>
</tr>
<tr>
<td>1000</td>
<td>0.6875</td>
<td>0.6874</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 6: Normal contact computation: Robustness and efficiency of the full-pass and the two-half-pass contact algorithms. The three values in each field give the total number of Newton iteration steps, the total computation time and the total time spent in the contact algorithm. $Xn$ means that the algorithm failed in the $n$th displacement step due to loss of convergence. Here, $n_{el}/L_0 = 8$ and $\epsilon_n = 10 E_0/L_0$.

<table>
<thead>
<tr>
<th>number of displ. steps</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>full-pass, $B_1$ slave</td>
<td>X1</td>
<td>X1</td>
<td>X1</td>
<td>X1</td>
<td>X3</td>
<td>96</td>
<td>140</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>195s</td>
<td>286s</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>37s</td>
<td>54s</td>
<td></td>
</tr>
<tr>
<td>full-pass, $B_2$ slave</td>
<td>X1</td>
<td>X2</td>
<td>X1</td>
<td>69</td>
<td>72</td>
<td>80</td>
<td>88</td>
<td>125</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>77s</td>
<td>81s</td>
<td>90s</td>
<td>99s</td>
<td>140s</td>
</tr>
<tr>
<td>two-half-pass</td>
<td>X4</td>
<td>139s</td>
<td>151s</td>
<td>64</td>
<td>73</td>
<td>80</td>
<td>88</td>
<td>129</td>
</tr>
<tr>
<td></td>
<td>55s</td>
<td>59s</td>
<td>68s</td>
<td>78s</td>
<td>85s</td>
<td>94s</td>
<td>137s</td>
<td></td>
</tr>
</tbody>
</table>

steps for an imposed vertical displacement of $2L_0/3$ using the specified number of displacement steps. The full-pass algorithm with $B_1$ as slave body only runs robustly with 10 steps or more. For this case it is fastest, since here the slave surface only contains 24 contact elements, whereas it contains 80 elements if $B_2$ is slave and 24+80 if the two-half-pass algorithm is considered. The two-half-pass algorithm is faster than the full-pass algorithm with $B_2$ as slave, even for the same number of iteration steps. The reason for this, is that the contact routine is evaluated faster, even though two passes are taken. We note that in all cases, the initial guess for each displacement step is obtained from displacing the current cylinder configuration downward uniformly. We further note that the full-pass results are obtained from the classical penalty formulation (class P) while the two-half-pass results are obtained from the new, modified penalty formulation (class S) using the averaging of $J_p$ at the corner nodes. (These results are not much different from those obtained with class P penalty method.)

Next, we analyze the sliding performance. This is shown in table 7. As seen, the convergence behavior during sliding is similar to that during normal contact (table 6): For most cases the two-half pass approach is either more robust or faster than the full-pass approach. The two-half-pass is the only approach that is able to converge in one displacement step. In the case where the cylinder is taken as slave, the full-pass algorithm performs very poorly: For many step sizes the algorithm does not converge to the specified tolerance.\textsuperscript{15} In the case where the

\textsuperscript{15}One reason for this problem is that the full-pass-algorithm, with $B_1$ as slave, frequently gets stuck between alternating solutions, e.g. projection points jumping between elements. This did not occur for the two-half-pass algorithm or the full-pass algorithm with $B_2$ as slave.
<table>
<thead>
<tr>
<th>number of displ. steps</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>10</th>
<th>15</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>full-pass, $B_1$ slave</td>
<td>X1</td>
<td>20</td>
<td>32</td>
<td>nc3</td>
<td>50</td>
<td>nc5</td>
<td>nc4</td>
<td>nc6</td>
<td>nc9</td>
<td>nc13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>11.1s</td>
<td>17.7s</td>
<td></td>
<td>28s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.1s</td>
<td>6.4s</td>
<td></td>
<td>10.0s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>full-pass, $B_2$ slave</td>
<td>X1</td>
<td>42</td>
<td>53</td>
<td>X4</td>
<td>86</td>
<td>84</td>
<td>90</td>
<td>117</td>
<td>141</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td></td>
<td>38s</td>
<td>47s</td>
<td></td>
<td>77s</td>
<td>75s</td>
<td>81s</td>
<td>105s</td>
<td>128s</td>
<td>162s</td>
</tr>
<tr>
<td></td>
<td></td>
<td>24s</td>
<td>30s</td>
<td></td>
<td>48s</td>
<td>47s</td>
<td>50s</td>
<td>65s</td>
<td>79s</td>
<td>100s</td>
</tr>
<tr>
<td>two-half-pass</td>
<td>23</td>
<td>38</td>
<td>50</td>
<td>60</td>
<td>71</td>
<td>75</td>
<td>83</td>
<td>104</td>
<td>139</td>
<td>173</td>
</tr>
<tr>
<td></td>
<td>20s</td>
<td>31s</td>
<td>42s</td>
<td>51s</td>
<td>60s</td>
<td>64s</td>
<td>71s</td>
<td>88s</td>
<td>118s</td>
<td>148s</td>
</tr>
<tr>
<td></td>
<td>12s</td>
<td>19s</td>
<td>26s</td>
<td>31s</td>
<td>37s</td>
<td>39s</td>
<td>43s</td>
<td>54s</td>
<td>72s</td>
<td>90s</td>
</tr>
</tbody>
</table>

Table 7: Sliding contact computation: Robustness and efficiency of the full-pass and the two-half-pass contact algorithms. Data arranged as in Tab. 6. ncn means that the algorithm did not converge in the nth step, but continued converging in subsequent steps and finally succeeded to obtain the imposed displacement. Here, $n_{el}/L_0 = 4$ and $\epsilon_n = 100 E_0/L_0$.

block is taken as slave, the full-pass algorithm is slowest, due to the additional numerical work that is required for the evaluation of the contact forces.

Finally, Fig. 16 shows the net contact forces during sliding. As noted above, the full-pass

results lie between those of the two-half-pass algorithm. The fluctuations of $P_x$ and $P_y$ are errors caused by the FE discretization. These errors decrease with increasing FE refinement (both $h$-refinement and $p$-refinement) as is shown in (Sauer, 2012). During computation, the active contact set is only updated at a certain accuracy threshold. For all cases considered here, between 1 and 5 such updates occur, regardless of the considered contact algorithm.

In summary, we conclude that for most of the cases considered here, the two-half-pass algorithm is more accurate, robust and efficient than the full-pass contact algorithm.

Figure 16: Net contact forces during sliding according to the various contact algorithms: horizontal force component $P_x$ (right); vertical force component $P_y$ (left). Here, $n_{el}/L_0 = 4$, $\epsilon_n = 100$, $n_{qp} = 50$.
7.2 3D contact between a ring and a slab

The second example considers 3D contact between an elastic ring (outer radius: 0.5$L_0$, width: 0.8$L_0$, thickness: 0.05$L_0$, enclosed angle: 160 degrees) and an elastic slab (dimensions: 1.2$L_0 \times 1.2$L_0 \times 0.3$L_0$). Exploiting double symmetry, only one quarter of the geometry is modeled (see Fig.17). The slab is fixed at its lower side and a uniform downward displacement $\bar{u}_z = -0.16$L_0 is applied to the upper surface of the ring in 16 displacement increments, causing both ring and slab to undergo large deformations. Hyperelastic, neo-Hookean behavior is considered for both bodies, now according to

$$\sigma = \frac{\Lambda}{2J_v}(J_v^2 - 1) I + \frac{\mu}{J_v}(F_v F_v^T - I), \quad (143)$$

with material properties $E = 10 E_0$ and $\nu = 0.3$ for the ring, and $E = 0.03 E_0$ and $\nu = 0.3$ for the slab. For the integration of the contact forces, $4 \times 4$ Gauss points are used for each contact element $\Gamma^e_k$. The class S penalty method is used as described by eq. (138). The penalty parameter is taken as $\epsilon_n = 100 E_0 / L_0$.

Fig. 17 shows the deformed shape as it is computed by the two-half-pass algorithm of table 3. The slab deformation is particularly large along the indenting edge of the ring. Along this edge large contact pressures can be observed.

The contact pressure is shown in Fig. 18 along the two axes of symmetry. The plotted pressure values are obtained from the interpolation of the nodal pressures defined in eqs. (139) and (142). The following cases are shown in Fig. 18:

- two-half-pass result
- full-pass result with the ring surface taken as slave
- full-pass result with the slab surface taken as slave

For the full-pass result, only one curve is plotted which corresponds to the contact pressure on the surface taken as slave. The two-half-pass formulation yields two pressure curves: one for each surface. There is a non-negligible discrepancy between the two curves obtained from the
two full-pass approaches. The contact pressure curves of the two-half-pass approach at \( y = 0 \) (Fig. 18a) are virtually coincident, and fall within the range of the full-pass curves. The same trend is observed for \( x = 0 \) (Fig. 18b). At the end of the contact region (\( y \approx 0.4 L_0 \)) large contact pressures appear, due to the sharp indenter edge.

### 7.3 Electrostatic interaction of deformable rings

The third example considers long-range electrostatic attraction between two deformable rings as shown in Fig. 19a. The rings have the radii \( R_1 = R_0, R_2 = 2R_0 \), the center distance \( D = 5R_0 \), the common wall thickness \( T = R_0/100 \) and the common height \( H \ll R_0 \). They are considered fixed at points \( A \) and \( B \). Linear elastic material behavior, with Young’s modulus

\[ E_0 \] and Poisson’s ratio \( \nu = 0.2 \), is assumed for both. The rings are modeled by geometrically exact beam theory (Reissner, 1972; Wriggers, 2008). 100 and 200 two-noded finite elements are used for ring 1 and ring 2, respectively. The electrostatic interaction is described by Coulomb’s
potential
\[ \phi_s(g_k) = k_e \frac{q_1 q_2}{g_k}, \quad g_k > 0. \]  
(144)

where \( k_e \) is Coulomb’s constant\(^{16} \). The interaction is integrated according to eqs. (22) and (23) using the algorithm presented in Sec. 5.3. Therefore we need \( f_s = -\phi'_s = k_e q_1 q_2 / g_k^2 \) and \( f'_s = -2 k_e q_1 q_2 / g_k^3 \). In eq. (144), \( q_k \) denotes an electrical charge so that \( \beta^g_k q_k \) denotes the charge density of surface \( \partial B_k. \)\(^{17} \) We consider a constant charge density on both surfaces. Hence, the quantity \( k_e \beta^g_0 q_1 \beta^g_2 q_2 \) is a constant that has the unit energy times distance per area squared. For normalization we introduce the dimensionless constant
\[ \gamma := -k_e \beta^g_0 q_1 \beta^g_2 q_2 / E_0, \]  
(145)

which expresses the strength of attraction relative to the ring stiffness. \( \gamma > 0 \) implies that the rings have opposite charges. Figure 19b shows the attractive force between the rings for increasing \( \gamma \). At \( \gamma \approx 2.62 \cdot 10^{-8} \) the system becomes unstable: the attraction is so strong that the rings are overpowered and would snap into contact. This is not considered here. At this stage the rings have approached each other by \( 1.553 R_0 \), leading to the large deformation of the rings shown in figure 19a. Considering \( E_0 = 117 \) GPa (Copper\(^{18} \)), \( R_0 = 0.01 \) m and \( H = 0.001 \) m, the force for \( \gamma = 2.62 \cdot 10^{-8} \) equals 0.319 N. Considering equal ring charges \( (Q_2 = -Q_1) \), this \( \gamma \) corresponds to a charge of \( Q_1 = 51.9 \) nC. The dashed line in figure 19b shows the interaction force assuming rigid rings. The difference between the dashed and solid curves marks the additional force due to the deformation of the rings. At \( \gamma = 2.62 \cdot 10^{-8} \) the rigidity assumption underestimates the force by a factor of about 2.7.

Given the deformed shape of the rings, one can easily compute the electric field \( E \) and the electric potential \( \Phi \) surrounding the rings. For a charged particle \( q_3 \) influenced by surfaces \( \partial B_1 \) and \( \partial B_2 \) we find (see Sec. 3.3)
\[ E = \frac{F_1 + F_2}{q_3}, \quad \Phi = \frac{\Phi_1 + \Phi_2}{q_3} \]  
(146)

where \( F_k \) and \( \Phi_k \) follow from eqs. (23) and (45) with \( \phi_s = k_e q_1 q_3 / g_k \) or \( \phi_s = k_e q_2 q_3 / g_k \) depending on the integration domain. Fig. 20 shows a plot of the electric potential \( \Phi \) and the field lines, i.e. the direction of \( E \).

### 7.4 Double strip peeling

The last example considers the peeling of two elastic strips (Sauer, 2012). The problem geometry in the undeformed configuration is shown in Fig. 21. The strips are bonded together along their length, except for a short debonded length \( L_d \) on the right hand side. Nonlinear elastic material behavior is considered according to the Neo-Hooke law (137). We consider \( L = 20 \) \( L_0 \), \( L_d = 5 \) \( L_0 \), \( h_1 = h_2 = L_0 \), \( E_2 = 3 E_0 \), \( E_1 = E_0 \), and \( \nu_1 = \nu_2 = 0.3 \). The strips are separated by applying the displacement \( u_y \) to the upper strip as shown in Fig. 21. The debonding is described by the cohesive zone model
\[ T_k = -T_0 \frac{r_k}{r_0} \exp \left( 1 - \frac{r_k}{r_0} \right), \]  
(147)

where \( r_k = ||r_k|| \) denotes the length of the distance vector
\[ r_k = x_k - x_\ell. \]  
(148)

\(^{16} k_e = (4\pi \epsilon_0)^{-1}, \) where \( \epsilon_0 = 8.85 \cdot 10^{-12} \) C\(^2\)/(Nm\(^2\)) is the permittivity of free space.

\(^{17} \) The rings are thus charged by \( Q_1 := 2\pi R_0 H \beta^g_0 q_1 \) and \( Q_2 := 4\pi R_0 H \beta^g_2 q_2 \).

\(^{18} \) As noted in Sec. 3.3, a charge redistribution due to the conductivity of the rings is not considered here.
between the points \( \mathbf{x}_k \in \partial B_k \) and \( \mathbf{x}_\ell \in \partial B_\ell \). This model is a class P interaction model. \( T_0 \) and \( r_0 \) are model parameters. This model accounts for both mode I and mode II debonding. Compressive contact occurs ahead of the delamination front and is described by the class P penalty formulation; see Sec. 3.1. The chosen model parameters are \( T_0 = 0.02 \div 0.2 \, E_0 \), \( r_0 = 0.05 \, L_0 \) and \( \epsilon_n = 50 \, E_0 / L_0 \). Since the lower strip is stiffer than the upper strip \( (E_2 > E_1) \) the left section will rotate upward during debonding. This is shown in Fig. 22a. As seen, the two strips are discretized by a non-conforming finite element mesh: the upper strip contains \( 96 \times 4 \) elements, the lower strip contains \( 80 \times 4 \) elements. Fig. 22b shows the peeling force \( P_y \) as a function of \( u_y \) for various values of \( T_0 \). For \( T_0 = 0.2 \, E_0 \) large oscillations appear. They are an artefact of the discretization and disappear for finer discretizations (Sauer, 2012). Fig. 22c shows the magnitude of the surface traction \( t_1 \) and \( t_2 \). They illustrate nicely that local contact equilibrium is only satisfied in the compression zone but not in the separation zone. This property was proven in Sec. 3.2.2. Global contact equilibrium, on the other hand, is satisfied to high accuracy. This is shown in table 8. Compared to the magnitude of the peeling forces

<table>
<thead>
<tr>
<th>( T_0 )</th>
<th>( \Delta P_x )</th>
<th>( \Delta P_y )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.02</td>
<td>6.604 (10^{-7} )</td>
<td>2.819 (10^{-7} )</td>
</tr>
<tr>
<td>0.05</td>
<td>1.526 (10^{-6} )</td>
<td>7.100 (10^{-7} )</td>
</tr>
<tr>
<td>0.10</td>
<td>3.363 (10^{-6} )</td>
<td>2.173 (10^{-6} )</td>
</tr>
<tr>
<td>0.20</td>
<td>8.245 (10^{-6} )</td>
<td>1.015 (10^{-5} )</td>
</tr>
</tbody>
</table>

Table 8: Maximum unbalance in the two components of the net peeling forces \( P_1 \) and \( P_2 \) for various values of \( T_0 \). Forces normalized by \( E_0 L_0 \), \( T_0 \) normalized by \( E_0 \).
Figure 22: Double strip peeling: a. deformation and stress $I_1 = \text{tr} \, \sigma$ (normalized by $E_0$) at $u_y = 20L_0$ and $T_0 = 0.2E_0$; b. load-deformation curve $P_y(u_y)$; c. cohesive zone traction for $u_y = 20L_0$ and $T_0 = 0.2E_0$.

(see Fig. 22b) these errors are negligible.

8 Conclusion

This paper presents a contact-interaction formulation that is based on surface pair-potentials and is accordingly termed the surface-potential-based-contact (SPBC) formulation. The formulation is derived from a variational principle. Starting from this, expressions for the contact force and contact energy are derived. The derivation distinguishes between long-range surface interactions (class L), short-range surface interactions (class S) and point interactions (class P). The latter contains the framework that is currently most commonly used for contact computations. The other two classes are suitable to describe physical surface interactions, like electrostatic, van-der-Waals or gravitational interactions, as well as to formulate contact constraint formulations, like barrier, penalty or cross-constrained methods.

The SPBC model leads naturally to a contact-interaction formulation that is unbiased w.r.t. the two interacting surfaces. It is shown that the model satisfies global contact equilibrium and, in case of class P and class S interaction, also local contact equilibrium. A large number of surface potentials are presented that fall into the framework of the SPBC model. Two analytical examples are presented to illustrate the formalism of the model.
The paper further presents a computational framework for the three classes of the SPBC formulation and presents several numerical examples that illustrate the generality of the formulations. Shown here are examples on the penalty method (Sec. 6, Sec. 7.1 and Sec. 7.2), the augmented Lagrange multiplier method (Sec. 6), electrostatic attraction (Sec. 7.3) and cohesive zone debonding (Sec. 7.4).

The new formulation leads naturally to a special contact algorithm, termed the two-half-pass contact algorithm as it only accounts for 'half' of the contact tractions compared to classical full-pass approaches. Traction continuity is therefore not imposed a-priori in the algorithm, although it is satisfied in theory. The numerical examples in Sec. 7 demonstrate that the new formulation satisfies local contact equilibrium (i.e. traction continuity) and global contact equilibrium (i.e. balance of net contact forces) to high accuracy. The approximation errors that appear in these equilibrium checks are on the same order than those stemming from the discretization and the constraint enforcement (e.g. due to finite penalty parameters). Additionally, the new formulation passes the contact patch test, as shown in Sec. 6. The analysis in Sec. 7.1.4 further shows that the new contact algorithm tends to be more accurate, robust and efficient than classic full-pass approaches.

In the present form, the SPBC formulation does not account for sliding friction. This extension is currently being considered by the authors. It is expected that the advantages of the new formulation carry over to frictional sliding contact. Even though the new contact formulation is derived from a displacement-dependent potential, the resulting two-half-pass algorithm can also be used independently of this potential. In this fashion also a Lagrange multiplier version of the two-half-pass algorithm can be developed, in order to satisfy contact constraints exactly. Therefore it is necessary to address possible issues of over-constraining. Another interesting extension is the combination of the SPBC formulation with the mortar method of Puso and Laursen (2004) and Yang et al. (2005).

A Analytical integration for class S

This appendix provides a proof for eq. (34).

Let us first examine \( n_k \, da_\ell \). According to Fig. 23 we find for a plane

\[
da_\ell = 2\pi r_a \, dr_a = 2\pi r_k \, dr_k ,
\]

(149)
since \( r_a = r_k \sin \theta \) and \( dr_a = dr_k / \sin \theta \). In here, distance \( r_k \) can be related to the surface distance \( r_n \), given by

\[
r_n = r_k \cos \theta ,
\]

(150)
see Fig. 23. The direction \( n_k \) can be decomposed into the components parallel and perpendicular to the surface, i.e.

\[
n_k = \begin{cases}
\cos \theta \, n_p - \sin \theta \, \bar{a}_p & \text{if } x_k \notin B_\ell , \\
-\cos \theta \, n_p - \sin \theta \, \bar{a}_p & \text{if } x_k \in B_\ell \text{ and } x_\ell \notin B_k , \\
\cos \theta \, n_p + \sin \theta \, \bar{a}_p & \text{if } x_k \in B_\ell \text{ and } x_\ell \in B_k ,
\end{cases}
\]

(151)
according to definition (15). The unit vectors \( n_p \) and \( \bar{a}_p \) denote the outward normal and the tangent vector of surface \( \partial B_\ell \) at \( x_p \), as shown in Fig. 23. Due to symmetry, the second term integrates to zero in eq. (33). To get rid of the sign in the first term, we define the signed surface distance

\[
g_n := \begin{cases}
r_n & \text{if } x_k \notin B_\ell , \\
-r_n & \text{if } x_k \in B_\ell .
\end{cases}
\]

(152)
Together with definition (14) and eq. (150), we can then rewrite eq. (149) as

$$\frac{d\alpha_{\ell}}{\cos \theta} = \begin{cases} g_n \, dg_k & \text{if } x_k \notin B_\ell, \\ -g_n \, dg_k & \text{if } x_k \in B_\ell \text{ and } x_{\ell} \notin B_k, \\ g_n \, dg_k & \text{if } x_k \in B_\ell \text{ and } x_{\ell} \in B_k, \end{cases}$$

(153)

such that the normal component of $n_k \, da_\ell$ simply becomes

$$\left( n_k \, da_\ell \right)_\perp = 2\pi g_n \, dg_k \, n_p,$$

(154)

for all cases. As the tangential component integrates to zero, we thus have

$$F_k = 2\pi \beta_s \int_{\partial B_\ell} f_s(g_k) \, dg_k \, n_p.$$

(155)

The entire surface $\partial B_\ell$ is covered when integrating from $r_k = r_n$ to $r_k = \infty$. For $x_k \notin B_\ell$ this implies an integration domain of $[g_n, \infty)$ for variable $g_k$. For $x_k \in B_\ell$, a sign change of $g_k$ occurs at the transition from $x_{\ell} \in B_k$ to $x_{\ell} \notin B_k$ (see Eq. (14) and Fig. 2), so that the integration needs to be carried out over the two separate domains $[g_n, -r_{in}]$ and $[r_{in}, \infty)$, where $r_{in}$ denotes the distance from $x_k$ to the intersection of both surfaces.\(^{19}\) It is reasonable to consider that $\phi_s$ has decayed at the distance $r_{in}$.\(^{20}\) It then follows that

$$F_k = 2\pi \beta_s \int_{g_n}^{r_e} f_s(g_k) \, dg_k \, n_p, \quad \text{with } r_e := \begin{cases} \infty & \text{if } x_k \notin B_\ell, \\ -r_{in} & \text{if } x_k \in B_\ell. \end{cases}$$

(156)

Due to eq. (19), this integrates to expression (34) if $\phi_s$ is a decaying potential. \(\square\)

\(^{19}\)Note that, due to the sign change of $g_k$ at $r_{in}$, the two integration domains are not connected.

\(^{20}\)For very small penetrations this procedure is inaccurate in theory, which is not a problem since penetrations will only be permitted for numerical regularization.
B Computation of the projection point $x_p$ and its derivatives

For class P and S interaction, the contact tractions $T_k$ can depend on the distance $g_n$, the surface normal $n_p$ and the surface stretch $J_p$, which are all evaluated at a certain point $x_p = x_k(\xi_p)$ can be expressed in terms of the surface parameterization coordinate $\xi = (\xi^1, \xi^2)$. Due to the nonlinearity of the projection, these coordinates may need to be obtained by a local search based e.g. on Newton-Raphson iteration. For a given, isoparametric FE discretization of surface $\partial B_\ell$ we have

$$x_p = \sum_l N_l(\xi_p) x_l = N(\xi_p) x^p_\ell,$$

(157)

where $x_l$ denotes the location of the $l$th surface node of element $\Gamma_\ell^k \in \partial B_\ell^k$ and where $N(\xi_p)$ refers to the array of shape functions introduced in Sec. 5.1. The co-variant tangent vectors $a^p_\alpha$ and $a^\alpha_\beta$ at $x_p$ follow from

$$a^p_\alpha = \frac{\partial x_p}{\partial \xi^\alpha} = \sum_l N_l,\alpha(\xi_p) x_l = N(\xi_p) x^p_\ell,$$

(158)

where $N_l,\alpha := \partial N_l/\partial \xi^\alpha$. The surface coordinates $\xi^1_p$ and $\xi^2_p$ specifying $x_p$ are found from solving the two equations

$$f_\alpha = (x_p - x_k) \cdot a^p_\alpha = 0, \quad \alpha = 1, 2.$$

(159)

Since these are nonlinear equations in general, a procedure like Newton’s method must be used, i.e we iterate for $i \rightarrow i + 1$

$$\xi_{p,i+1}^\alpha = \xi_{p,i}^\alpha + \Delta \xi_{p,i+1}^\alpha,$$

(160)

where the update $\Delta \xi_{p,i+1}^\alpha$ is found from the matrix equation

$$\left[ \frac{\partial f_\alpha}{\partial \xi^\beta} \right]_{i} \Delta \xi_{p,i+1}^\beta = -f_{\alpha,i}, \quad \text{(sum in $\beta$)},$$

(161)

Here, and in the following repeated Greek indices imply summation. The gradient of $f_\alpha$ simply follows as

$$\frac{\partial f_\alpha}{\partial \xi^\beta} = a^p_\alpha \cdot a^\beta_\beta + (x_p - x_k) \cdot a^\alpha_\beta,$$

(162)

with

$$a^\alpha_\beta = \sum_l N_l,\alpha(\xi_p) x_l = N(\xi_p) x^\beta_\ell.$$  

(163)

With the definition of $n_p$ in eq. (123), eq. (162) can be rewritten into

$$\frac{\partial f_\alpha}{\partial \xi^\beta} = a_{\alpha\beta} - g_n a_{\alpha\beta},$$

(164)

where

$$a_{\alpha\beta} := a^p_\alpha \cdot a^\beta_\beta,$$

$$b_{\alpha\beta} := n_p \cdot a^\alpha_\beta.$$  

(165)

are the co-variant components of the metric tensor and the curvature tensor at $x_p \in \partial B_\ell$. Note that in general $a_{\alpha\beta} \neq \delta_{\alpha\beta}$ since the tangent vectors $a^\alpha_\alpha (\alpha = 1, 2)$ are neither orthogonal nor normalized.
The derivatives \( \frac{\partial \xi^\alpha_p}{\partial x_k} \) and \( \frac{\partial \xi^\alpha_p}{\partial x_I} \) are obtained from the residual equation (159). Setting \( f_\alpha \equiv 0 \) and differentiating w.r.t. \( x_k \) we find after some rearranging

\[
\frac{\partial \xi^\alpha_p}{\partial x_k} = c^{\alpha\beta} a^p_\beta .
\] (166)

where \( c^{\alpha\beta} \) are the contra-variant components of the tensor \( c = d^{-1} \), where \( d \) is defined by the co-variant components

\[
d_{\alpha\beta} := a_{\alpha\beta} - g_n b_{\alpha\beta} .
\] (167)

Differentiating \( f_\alpha \equiv 0 \) w.r.t. \( x_I \) gives

\[
\frac{\partial \xi^\alpha_p}{\partial x_I} = -c^{\alpha\beta} \left( N_I(\xi_p) a^p_\beta - g_n N_{I,\beta}(\xi_p) n_p \right) .
\] (168)

For the subsequent appendices, we require the derivatives of \( x_p \) w.r.t. \( x_k \) and \( x_I \). From expressions (158) and (166) we can find

\[
\frac{\partial x_p}{\partial x_k} = \frac{\partial x_p}{\partial \xi^\alpha_p} \otimes \frac{\partial \xi^\alpha_p}{\partial x_k} = c^{\alpha\beta} a^p_\alpha \otimes a^p_\beta .
\] (169)

Since \( x_p = x_p(x_I, \xi_p(x_k, x_I)) \) we further have

\[
\frac{\partial x_p}{\partial x_k} = \frac{\partial x_p}{\partial x_k} \bigg|_{\xi^\alpha_p} + \frac{\partial x_p}{\partial \xi^\alpha_p} \otimes \frac{\partial \xi^\alpha_p}{\partial x_I} .
\] (170)

Evaluation gives

\[
\frac{\partial x_p}{\partial x_I} = N_I(\xi_p) I - c^{\alpha\beta} a^p_\alpha \otimes \left( N_I(\xi_p) a^p_\beta - g_n N_{I,\beta}(\xi_p) n_p \right) .
\] (171)

The expressions given above simplify considerably for linear 2D finite elements as is shown in Appendix E.

### C Computation of gap \( g_n \), normal \( n_p \) and their derivatives

The normal gap can be expressed as

\[
g_n = \pm \sqrt{(x_k - x_p) \cdot (x_k - x_p)} ,
\] (172)

where the sign depends on wether \( x_k \in B_\ell \) or \( x_k \notin B_\ell \) (see eq. (39)). The functional dependence on the surface points \( x_k \in \partial B_k \) and \( x_I \in \partial B_\ell \) thus is

\[
g_n = g_n(x_k, x_p(x_I, \xi_p(x_k, x_I))) ,
\] (173)

and the gradient w.r.t. \( x_k \) becomes

\[
\frac{\partial g_n}{\partial x_k} = \frac{\partial g_n}{\partial x_k} \bigg|_{x_p} + \frac{\partial g_n}{\partial \xi^\alpha_p} \otimes \frac{\partial \xi^\alpha_p}{\partial x_k} .
\] (174)

It turns out that \( \frac{\partial g_n}{\partial \xi^\alpha_p} = 0 \), so that

\[
\frac{\partial g_n}{\partial x_k} = n_p .
\] (175)
According to (173), the gradient w.r.t. $x_I$ is
\[
\frac{\partial g_n}{\partial x_I} = \frac{\partial g_n}{\partial x_p} \cdot \frac{\partial x_p}{\partial x_I},
\]
(176)
where $\frac{\partial x_p}{\partial x_I}$ is given in eq. (171). With this we now simply find
\[
\frac{\partial g_n}{\partial x_I} = -n_p N_I(\xi_p),
\]
(177)
since the last term cancels.

The surface normal, expressed by eq. (123), depends on $x_k$ and $x_I$ as
\[
n_p = n_p(x_k, x_p(x_k, \xi_p(x_k, x_I)), g_n(x_k, x_I)),
\]
(178)
where $g_n$ is discussed above. According to this, the gradient w.r.t. $x_k$ is
\[
\frac{\partial n_p}{\partial x_k} = \left. \frac{\partial n_p}{\partial x_p} \right|_{x_p,g_n} + \frac{\partial n_p}{\partial x_p} \cdot \frac{\partial x_p}{\partial x_k} + \frac{\partial n_p}{\partial g_n} \otimes \frac{\partial g_n}{\partial x_k}.
\]
(179)
With the help of eq. (169), eq. (175) and $\frac{\partial n_p}{\partial g_n} = -n_p/g_n$, we find
\[
\frac{\partial n_p}{\partial x_k} = \frac{1}{g_n} \left[ I - n_p \otimes n_p - c^{\alpha\beta} a^{\alpha}_p \otimes a^{\beta}_p \right].
\]
(180)
This expression is equivalent to
\[
\frac{\partial n_p}{\partial x_k} = \frac{1}{\kappa_1 + g_n} \hat{a}^p_1 \otimes \hat{a}^p_1 + \frac{1}{\kappa_2 + g_n} a^p_2 \otimes a^p_2,
\]
(181)
given in Sauer and Wriggers (2009). Here $\kappa_1$ and $\kappa_2$ denote the principal curvatures\(^{21}\) of the master surface $\partial B_n$ at the projection point $x_p$, and $\hat{a}^p_1$ and $\hat{a}^p_2$ denote the corresponding normalized tangent vectors at $x_p$ along $\kappa_\alpha$.

According to eq. (178), the gradient of $n_p$ w.r.t. $x_I$ is
\[
\frac{\partial n_p}{\partial x_I} = \left. \frac{\partial n_p}{\partial x_p} \right|_{x_p,g_n} \cdot \left[ \frac{\partial x_p}{\partial x_I} \right]_{\xi_p} + \frac{\partial x_p}{\partial \xi_p} \otimes \frac{\partial \xi_p}{\partial x_I} + \frac{\partial n_p}{\partial g_n} \otimes \frac{\partial g_n}{\partial x_I}.
\]
(182)
The first and last terms are easy to evaluate, and are partly given by eqs. (171) and (177). For the middle term we require
\[
\frac{\partial n_p}{\partial \xi_p} = -\frac{1}{g_n} \left( a^\alpha_p - \frac{\partial g_n}{\partial \xi_p} n_p \right).
\]
(183)
Here the last term is zero as noted above. In view of (168) the middle term in eq. (182) now becomes
\[
\frac{\partial n_p}{\partial \xi_p} \otimes \frac{\partial \xi_p}{\partial x_I} = \frac{c^{\alpha\beta}}{g_n} a^\alpha_p \otimes \left[ N_I(\xi_p) a^{\beta}_p - g_n N_I,\beta(\xi_p) n_p \right].
\]
(184)
Thus
\[
\frac{\partial n_p}{\partial x_I} = -\frac{1}{g_n} \left[ I - n_p \otimes n_p \right] N_I(\xi_p) + \frac{c^{\alpha\beta}}{g_n} a^\alpha_p \otimes \left[ N_I(\xi_p) a^{\beta}_p - g_n N_I,\beta(\xi_p) n_p \right],
\]
(185)
and
\[
\frac{\partial n_p}{\partial u_\ell^c} = -\frac{1}{g_n} \left[ I - n_p \otimes n_p - c^{\alpha\beta} a^\alpha_p \otimes a^{\beta}_p \right] N_\ell(\xi_p) - c^{\alpha\beta} a^\alpha_p \otimes n_p N_\ell,\beta(\xi_p).
\]
(186)
\(^{21}\)Here the curvature is taken positive for convex bodies like a sphere.
If the curvature is zero (e.g. for linear triangular or linear 2D elements), the gradients of $n_p$ collapse to
\[
\frac{\partial n_p}{\partial x_k} = 0, \quad \frac{\partial n_p}{\partial x_I} = -a^{\alpha\beta} a_{\alpha}^p \otimes n_p N_{I,\beta}(\xi_p), \quad \text{(for } b_{\alpha\beta} = 0),
\]
(187)
since $n_p$ is constant across the element. This is in agreement with the 2D case reported in eqs. (207) and (208); see Appendix E. Here, $a^{\alpha\beta}$ are the contra-variant components of the metric tensor.

\section*{D Computation of the surface stretch $J_p$ and its derivatives}

Given the parameterizations $x_\ell = x_\ell(\xi^1, \xi^2)$ and $X_\ell = X_\ell(\xi^1, \xi^2)$ for the surface points $x_\ell \in \partial B_\ell$ and $X_\ell \in \partial B_{0\ell}$, the surface stretch is obtained from
\[
J_p = \frac{\| a_1^p \times a_2^p \|}{\| A_1^p \times A_2^p \|},
\]
(188)
where
\[
a_{\alpha}^p = \frac{\partial x_\ell}{\partial \xi^\alpha}(\xi_p), \quad A_{\alpha}^p = \frac{\partial X_\ell}{\partial \xi^\alpha}(\xi_p),
\]
(189)
are the surface tangent vectors at the projection point $x_p \in \partial B_\ell$ and its counterpart $X_p \in \partial B_{0\ell}$. Alternatively, the surface stretch can be obtained as the determinant of the surface deformation gradient $F_s = a_{\alpha} \otimes A_{\alpha}$, where $A_{\alpha}$ is the contra-variant counterpart of $A_{\alpha}$. In 2D the surface stretch reduces to
\[
J_p = \frac{\| a^p \|}{\| A^p \|}.
\]
(190)

For isoparametric finite elements the surfaces $\partial B_\ell$ and $\partial B_{0\ell}$ are described by
\[
x_\ell = \sum I N_I(\xi) \, x_I, \quad X_\ell = \sum I N_I(\xi) \, X_I,
\]
(191)
within each element. Here $x_I$ and $X_I$ denote the nodes of the surface element. The tangent vectors at $x_p = x_\ell(\xi_p)$ and $X_p = X_\ell(\xi_p)$ (where $\xi_p = \{ \xi^1_p, \xi^2_p \}$) thus follow as
\[
a_{\alpha}^p = \sum I N_{I,\alpha}(\xi_p) \, x_I, \quad A_{\alpha}^p = \sum I N_{I,\alpha}(\xi_p) \, X_I.
\]
(192)

According to the above equations, the surface stretch $J_p$ depends on the current nodal positions $x_I$ and on the location $\xi_p$, which in turn depends on the source point $x_k \in B_k$ and on the neighboring nodal points $x_I \in B_\ell$, i.e. we have
\[
J_p = J_p(x_I, \xi_p(x_k, x_I)).
\]
(193)

For the finite element stiffness matrix (see Sec. 5.2) we require the derivative $\partial J_p/\partial x_k$ and $\partial J_p/\partial x_I$. They are derived next.

1. Derivative $\partial J_p/\partial x_k$: According to chain rule, we have
\[
\frac{\partial J_p}{\partial x_k} = \frac{\partial J_p}{\partial \xi^\alpha_p} \frac{\partial \xi^\alpha_p}{\partial x_k}.
\]
(194)
The second part, \( \frac{\partial \xi_p^\alpha}{\partial \xi_k} \), is specified in eq. (166). The first part can be written as

\[
\frac{\partial J_p}{\partial \xi_k} = \frac{\partial J_p}{\partial a_p^\gamma} \frac{\partial a_p^\gamma}{\partial \xi_k} + \frac{\partial J_p}{\partial A_p^\gamma} \frac{\partial A_p^\gamma}{\partial \xi_k} \quad \text{(sum in } \gamma) .
\]  

(195)

Note that for linear elements (3-noded triangles or 2-noded line elements) the stretch \( J_p \) is constant, so that \( \frac{\partial J_p}{\partial \xi_k} = 0 \). According to eq. (192) we have

\[
\frac{\partial a_p^\gamma}{\partial \xi_k} = a_p^\gamma, \quad \frac{\partial A_p^\gamma}{\partial \xi_k} = A_p^\gamma,
\]

(196)

Note that \( A_p^\gamma, \alpha = 0 \) if the reference surface \( \partial B_0^\ell \) is flat. The remaining parts in eq. (195) are

\[
\frac{\partial J_p}{\partial a_p^\alpha} = J_p a_p^\alpha ,
\]

(197)

\[
\frac{\partial J_p}{\partial A_p^\alpha} = -J_p A_p^\alpha ,
\]

(198)

with \( A_p^\alpha = A^\alpha A_p^\beta, [A^\alpha] = [A_{\alpha\beta}]^{-1} \) and \( A_{\alpha\beta} = A_p^\alpha \cdot A_p^\beta \). In 2D the last two equations can be written as

\[
\frac{\partial J_p}{\partial a_p^\alpha} = \frac{a_p^\alpha}{||a_p^\alpha||},
\]

(199)

and

\[
\frac{\partial J_p}{\partial A_p^\alpha} = -\frac{J_p A_p^\alpha}{||A_p^\alpha||^2} .
\]

(200)

With this everything is specified to evaluate \( \frac{\partial J_p}{\partial \mathbf{x}_k} \).

2. Derivative \( \frac{\partial J_p}{\partial \mathbf{x}_I} \): Due to eq. (193), we can write

\[
\frac{\partial J_p}{\partial \mathbf{x}_I} = \left[ \frac{\partial J_p}{\partial \mathbf{x}_1} \bigg|_{\xi_p} + \frac{\partial J_p}{\partial \xi_k} \frac{\partial \xi_k^\alpha}{\partial \mathbf{x}_I} \right] .
\]

(201)

Contributions \( \frac{\partial J_p}{\partial \xi_k} \) and \( \frac{\partial \xi_k^\alpha}{\partial \mathbf{x}_I} \) are given in eq. (195) and eq. (168). The remaining term can written as

\[
\frac{\partial J_p}{\partial \mathbf{x}_I} \bigg|_{\xi_p} = N_{I,\alpha}(\xi_p) \frac{\partial J_p}{\partial a_p^\alpha} ,
\]

(202)

where \( \frac{\partial J_p}{\partial a_p^\alpha} \) is found in eq. (197), or (199) respectively. Given \( \frac{\partial J_p}{\partial \mathbf{x}_I} \), where \( I = 1, 2, ..., n_e^\ell \), we have

\[
\frac{\partial J_p}{\partial \mathbf{u}_I^\ell} = \left[ \frac{\partial J_p}{\partial \mathbf{x}_1}, \frac{\partial J_p}{\partial \mathbf{x}_2}, \ldots, \frac{\partial J_p}{\partial \mathbf{x}_{n_e^\ell}} \right] .
\]

(203)

as is needed in eq. (127). Evaluation gives

\[
\frac{\partial J_p}{\partial \mathbf{u}_I^\ell} = -\frac{\partial J_p}{\partial \xi_k} a_p^\alpha A_p^\gamma N_{I,\alpha}(\xi_p) + \left( \frac{\partial J_p}{\partial a_p^\alpha} + g_{n_e^\ell} \frac{\partial J_p}{\partial \xi_k} a_p^\alpha \right) N_{I,\alpha}(\xi_p) .
\]

(204)
E  Special case for linear, 2D finite elements

Here we particularize our results for the case that linear surface elements are used for 2D problems. In this case, interpolation (157) is characterized by the two shape functions

\[ N_1(\xi) = \frac{1}{2}(1 - \xi), \quad N_2(\xi) = \frac{1}{2}(1 + \xi), \quad \text{for} \ \xi \in [-1, 1]. \]  

(205)

The projection point \( x_p = x_\ell(\xi_p) \) can then be obtained analytically through (Sauer, 2006)

\[ \xi_p = \frac{(x_2 - x_1) \cdot (2x_k - x_1 - x_2)}{\ell_\ell^2}, \]  

(206)

where \( \ell_\ell = \sqrt{(x_2 - x_1) \cdot (x_2 - x_1)} \) denotes the current element length and \( x_I \) for \( I = 1, 2 \) denotes the two element nodes at \( \xi = \pm 1 \). If \( \xi_p \) lies outside the interval \([-1, 1]\), the neighboring element needs to be checked for a feasible solution. If \( \xi_p \) lies outside both, the projection is considered onto the common corner node.

The gradients of \( n_p \) w.r.t. \( x_k \) and \( x_I \) are discussed next. If the projection point \( x_k \) falls on a corner node, the gradients are given by eqs. (128) and (129). Otherwise, if \( \xi \in (-1, 1) \), we have, according to eq. (187),

\[ \frac{\partial n_p}{\partial x_k} = 0, \]  

(207)

and

\[ \frac{\partial n_p}{\partial u_\ell^e} = -\frac{2}{\ell_\ell} (\bar{a}^p \otimes n_p) N'_\ell(\xi_p), \]  

(208)

since \( a^p = \frac{1}{2}(x_2 - x_1) \) and \( a^{\alpha\beta} = \|a^p\|^{-2} \) according to eq. (158) and (165). Here, \( \bar{a}^p \) denotes the normalized tangent vector obtained as \( \bar{a}^p = 2a^p/\ell_\ell \), since \( \|a^p\| = \ell_\ell/2 \). Further,

\[ N'_1(\xi_p) = [N'_1(\xi_p) I, N'_2(\xi_p) I] = \frac{1}{2} \left[ -I, I \right]. \]  

(209)

Next we examine the surface stretch \( J_p \) defined by eq. (190). As noted in Appendix D, \( \partial J_p/\partial \xi_p = 0 \) for linear elements, so that we now have

\[ \frac{\partial J_p}{\partial x_k} = 0, \]  

(210)

according to eq (194), and

\[ \frac{\partial J_p}{\partial u_\ell^e} = \frac{\bar{a}^p}{\|A^p\|} N'_\ell(\xi_p), \]  

(211)

according to eqs. (199), (202) and (203). We thus obtain

\[ -\frac{T_k}{J_p} \otimes \frac{\partial J_p}{\partial u_\ell^e} = -\frac{T_k}{\|a^p\|} n_p \otimes \bar{a}^p N'_\ell(\xi_p), \]  

(212)

as is needed for eq. (127). For projections onto FE corner nodes, the stretch \( J_p \) can be averaged from the values of the adjacent elements \( \Gamma_\ell^e \) and \( \Gamma_{\ell+1}^e \). In this case \( J_p \) depends on the nodal displacements \( u_\ell^e \) and \( u_{\ell+1}^e \), and the FE tangent matrices have to be adjusted accordingly.

With these equations, everything is specified in order to evaluate the contact algorithm of table 3 for 2D, linear FE discretizations.
Symmetry proof for the FE tangent matrix

Here we show that the proposed contact formulation based on the two half-passes, leads to a symmetric FE tangent matrix for conforming contact meshes, i.e. we show that the tangent matrix defined in eq. (118) satisfies $k_{e\ell k}^c = k_{k\ell}^c T$ and $k_{e\ell k}^c = k_{k\ell}^c T$. The 2D case with linear elements is considered for this. The first case is easy to show: Combining eqs. (118) 1, (127) 1 and (210) we find

$$k_{e\ell k}^c = - \int_{\Gamma_k} T_k^T \frac{\partial n_p}{\partial x_k} N_k \cdot dA_k,$$  \hspace{1cm} (213)

which is symmetric according to eqs. (128) and (207).

To show the second case, we combine eqs. (118) 2, (127) 2, (208) and (212) to get

$$k_{e\ell k}^c = \int_{\Gamma_k} T_k^T N_k^T (n_p \otimes n_p) N_\ell + \frac{2T_k}{t_\ell} N_k^T (n_p \otimes \bar{a_\ell} + \bar{a_\ell} \otimes n_p) N_\ell^T \cdot dA_k.$$  \hspace{1cm} (214)

Here $N_\ell$ and $N_\ell^T$ are evaluated at the projection point $\xi^p$. According to this expression, $k_{e\ell k}^c$ is not symmetric in general. But it becomes fully symmetric for conforming, linear contact elements, as is shown next. If the discretized contact surfaces conform, the gap $g_n$ will be zero along the elements. Therefore $T_k$ will be constant within each $\Gamma_k$. Defining the symmetric tensor

$$M_k := n_p \otimes \bar{a_\ell} + \bar{a_\ell} \otimes n_p,$$ \hspace{1cm} (215)

we can then integrate the second term analytically, giving

$$k_{e\ell k}^c = \int_{\Gamma_k} t_\ell N_k^T n_p \otimes n_p N_\ell \cdot dA_k + t_k \begin{bmatrix} M_k & M_k \\ M_k & M_k \end{bmatrix},$$ \hspace{1cm} (216)

according to eqs. (112), (205) and (209). Here we have used the fact that $\Gamma_\ell$ and $\Gamma_k$ are equally long, due to the conforming meshes. The conforming meshes also imply that $M_\ell = M_k$. Further, $t_\ell = t_k$ as shown in Sec. 2.5. Therefore,

$$k_{e\ell k}^c T = k_{e\ell k}^c T.$$ \hspace{1cm} (217)

Due to this, the global contact stiffness matrix $k_c$, obtained from the assembly of $k_{e\ell k}^c$ and $k_{e\ell k}^c T$ for all contact elements $\Gamma_k$, is symmetric.

Acknowledgements

The authors are grateful for the support of the Alexander von Humboldt foundation, European Research Council (ERC Grant agreement 279439) and German Research Foundation (grants SA1822/5-1 and GSC111). The authors also thank Thang Xuan Duong for checking the manuscript carefully.

References


