3 Monotone Multigrid Methods

In this chapter, we introduce a monotone multigrid method for solving Signorini's problem in linear elasticity. We proof the global convergence of the method. The method itself is defined in terms *truncated functions* depending on the solution and *monotone restrictions*. In combination with a projected block Gauß–Seidel method, these form the main ingredients of the monotone multigrid method for Signorini's problem. Before describing the method in detail, let us note some remarkable features of the resulting iteration scheme.

- It is a globally convergent nonlinear method.
- The *global* nonlinearity of the problem is resolved *locally*.
- It reduces to a linear subspace correction method once the actual zone of contact has been found.
- It is of optimal complexity per iteration step.
- It does not use any regularization.

By design, our iterative method does not distinguish between a linear and a nonlinear iteration step. Rather, it acts as a *locally* nonlinear solver, i.e., the *global* nonlinearity of the problem is resolved by solving *local* nonlinear subproblems. Thus, our method can be viewed as an iterative scheme which sits on top of an "adaptive" local linearization strategy. In other words, our method always uses as much information on the nonlinearity as it is available. The global control of this mechanism is provided by the minimization of the total energy. This leads to an efficient and easy to handle nonlinear method.

Let us recall that the nondifferentiablity of the energy function $\mathcal{J}$ at the contact boundary constitutes the main difficulty in constructing an efficient solver. Newton methods cannot be applied directly. To overcome this difficulty, often regularization techniques, i.e., penalty methods, are used, see [CSW99, ESW99, Glo84, GLT81]. These require the careful handling of regularization parameters in order to find a reasonable compromise between efficiency and accuracy. In case of frictional contact, the accuracy provided by penalty method may even not be sufficient, see [ARS99].

Dual techniques, see, e.g.,[Ber82, Glo84, GLT81, GL89, DV97, PC99]) are based on saddle point formulations. The constraints are incorporated by means of Lagrange multipliers. Active set strategies [BH93, HM83, Hop87, HK94, Dos97] iteratively provide approximations of the contact set. A linear subproblem with given contact set has to be solved in each iteration step and multigrid methods are typically used for this purpose. An active set strategy with inexact linear solver has been proposed by Dostál [Dos97]. Recently, Schöberl [Sch98a, Sch98b] has developed an approximate variant of the projection method (cf. e.g. [Glo84, p. 5]) using a domain decomposition preconditioner and a linear multigrid solver on the interior nodes.

We emphasize that the algorithm presented herein does not involve any regularization or dual formulation. Particularly, it should be considered as a descent method rather than an active set strategy. The central idea of the monotone multigrid method is to minimize the functional $\bar{\mathcal{J}}$ of total energy successively in direction of suitable chosen d–dimensional

subspaces. By this means, we obtain nonlinear variants of successive subspace correction in the sense of Xu [Xu92]. See e.g., [HD97, TX] for a similar approach to smooth variational problems.

In its simplest form, minimization can be obtained by using well known projected block Gauß–Seidel relaxation, corresponding to the subspaces spanned by some Lagrangian finite element basis functions. Unfortunately, the convergence speed of Gauß–Seidel relaxation decreases with decreasing mesh size. In the context of linear multigrid method, convergence speed is increased by using additional coarse grid spaces, corresponding to some frequency like decomposition of the error. Here, we cannot follow the standard multigrid approach, since any coarse grid correction has to be admissible *with respect to the fine grid constraints*. Any correction associated with a geometric object being in contact with the obstacle requires zero correction in normal direction. This leads to coarse grid corrections which are bound to the corresponding $(d-1)$–dimensional space of tangential displacements. In case of different normals, this seems to be hardly realizable for any non trivial correction, since different fine grid subspaces share the same coarse grid correction. The remedy is using sophisticated modifications of the multilevel nodal basis, giving rise to nonlinear truncated coarse grid functions. The corrections originating from these modified functions find their natural interpretation as *local projections* onto the local spaces of tangential displacements. These truncated coarse grid functions are introduced in Section 3.2.

For problems with the outer normal being constant along the Signorini boundary, also block versions of suitable scalar methods can be applied. Multigrid methods for scalar obstacle problems have been considered by several authors, see, e.g., [Bel93, BC83, HM83, Hop90, Kor94, Man84]).As mentioned above, there is no obvious generelization to Signorini's problem with varying normals.

Even in case of no contact, the coarse grid corrections have to be admissible with respect to the fine grid constraints. Straightforward implementation leads to additional prolongations in order to check the constraints. As a consequence, the complexity of one iteration step on level J is $\mathcal{O}(n_J \log n_J)$ for uniformly refined triangulations and might be even $\mathcal{O}(n_J^2)$ in the adaptive case. Optimal complexity of multigrid $\mathcal{V}$–cycle is preserved by approximating fine grid constraints on coarser grids using so called *monotone restrictions*, see Section 3.3. This modification may slow down convergence, as long as the algebraic error is too large. In our numerical experiments, see Chapter 5, we observe that initial iterates as provided by nested iteration are usually accurate enough to provide fast convergence throughout the whole iteration.

This chapter is organized as follows: A general framework of our method together with basic convergence results is presented in Section 3.1. In particular, it turns out that the discrete contact set is detected in a finite number of steps, if the given discrete problem is non-degenerate. Then, our nonlinear iteration automatically becomes a linear subspace correction method for the resulting linear problem.

A suitable multilevel splitting and the monotone restrictions are described in Section 3.3. In case of spatially varying normals, various fine grid directions are incorporated in each coarse grid space. Similar techniques provide appropriate monotone restrictions of fine grid constraints. The resulting *truncated monotone multigrid method* can be arranged as a multigrid $\mathcal{V}$-cycle with projected block Gauß–Seidel smoothing and sophisticated

restriction and prolongation. It is described in Section 3.2.

For the ease of presentation, we restrict ourselves to the case of constraints given only in *normal* direction. The generalization of the results given in this chapter to the general case of box constraints is straightforward. Our implementation, which is described in Chapter 4, is capable of handling box constraints, i.e., constraints in both, normal and tangential direction. Let us finally note that this chapter is mostly borrowed from [KK00].

3.1 Minimization of Energy

The basic idea of monotone multigrid methods is to minimize the energy functional $\bar{\mathcal{J}}$ in direction of suitably chosen functions $\boldsymbol{\mu}_p$. For example, straightforward implementation of this idea leads to nonlinear Gauß–Seidel relaxation, if minimizing is done only in direction of the basis functions of the fine grid space $\boldsymbol{S}^{(J)}$. Since relaxation methods as Gauß–Seidel relaxation are known to decrease in convergence speed for increasing number of unknowns, in this section we consider so called *extended relaxations*, see [Kor94]. The aim is, to accelerate the convergence using corrections originating from minimization of $\bar{\mathcal{J}}$ in direction of functions with large support, which are called *coarse grid corrections*. Obviously, the proper choice of the functions used for the extended relaxation is crucial for the convergence speed of the resulting method.

Since minimization of the functional $\bar{\mathcal{J}}$ has to be carried out with respect to the constraints of the finest grid, we have to ensure that the coarse grid corrections are admissible with respect to the fine grid constraints. This is done by using so called *monotone restrictions*, which give rise to interior approximations of the convex set $\mathcal{K}$.

In this section, we introduce monotone multigrid methods as subspace correction methods. We define monotone restrictions and proof the global convergence of the method. For the construction of suitable coarse grid functions for Signorini's problem in linear elasticity, we refer the reader to Section 3.2.

Being concerned with linear elasticity, we carefully have to distinguish between any generic object as, e.g., the the elasticity tensor, and its representation with respect to a family of given local coordinate system. In particular, the proper choice of the local coordinate systems used is of crucial importance for the considerations made in the next Section. Let us recall, that the summation convention is enforced on small Latin indices i, j, s, which always range from 1 to d. Moreover, by $\boldsymbol{E}^i$ we denote the canonical basis vectors of the Euclidian vector space $\mathbb{R}^d$. As in the previous chapters, we use boldfaces letters for vectors and tensors and normal typeface letters for scalar objects. For the ease of presentation, all convex spaces ares denoted by calligraphic letters, as in $\mathcal{D}$ and $\mathcal{K}$. Note, that calligraphic letters are also used for the triangulation $\mathcal{T}$, the set of nodes $\mathcal{N}$ and the energy functional $\mathcal{J}$.

Now, we are in a position to define the finite element spaces associated with our subspace correction method. Let $\boldsymbol{S}^{(J)}$ be the finite element space spanned by the nodal basis functions

$$\lambda_p^{(J)} \boldsymbol{E}^i, \quad i = 1, \dots, d, \quad p \in \mathcal{N}^{(J)}$$

with piecewise linear, scalar functions $\lambda_p^{(J)}$ satisfying

$$\lambda_p^{(J)}(q) = \delta_{pq}, \quad p, q \in \mathcal{N}^{(J)},$$

see also (2.27). Here, δ_{pq} denotes the Kronecker delta. The nodes $p \in \mathcal{N}^{(J)}$ are assumed to be the vertices of a suitable triangulation $\mathcal{T}_J$, see also Section 3.2. As mentioned in the introduction, a nonlinear version of successive subspace correction, cf. Xu [Xu92], based on the splitting $\boldsymbol{S}^{(J)} = \boldsymbol{V}_1 + \ldots + \boldsymbol{V}_{n_J}$ with d–dimensional subspaces

$$\boldsymbol{V}_l = \operatorname{span}\{\lambda_{p_l}^{(J)} \boldsymbol{E}^1, \ldots, \lambda_{p_l}^{(J)} \boldsymbol{E}^d\} \qquad l = 1, \ldots, n_J = \#\mathcal{N}^{(J)}$$

gives rise to the well known projected block Gauß–Seidel relaxation $\mathcal{M}_J$, see.,e.g., [Glo84, p. 151]. As the convergence speed of this method usually deteriorates rapidly with increasing refinement level, we consider the *extended splitting*

$$\boldsymbol{S}^{(J)} = \boldsymbol{V}_1 + \ldots + \boldsymbol{V}_{n_J} + \boldsymbol{V}_{n_J+1}^{\nu} + \ldots + \boldsymbol{V}_m^{\nu}. \tag{3.1}$$

The additional d–dimensional subspaces $\boldsymbol{V}_l^{\nu}$, $\nu = n_J + 1, \ldots, m$, are intended to improve the low-frequency contributions of the error. Hence, basis functions $\boldsymbol{\mu}_l^{1,\nu}, \ldots, \boldsymbol{\mu}_l^{d,\nu}$ of $\boldsymbol{V}_l^{\nu}$ should have large support. In order to allow a stepwise adaptation of the splitting (3.1) to the actual contact zone, the spaces $\boldsymbol{V}_l^{\nu}$ may vary in each iteration step.

In principle, it is possible to choose the basis functions $\boldsymbol{\mu}_l^{1,\nu}, \ldots, \boldsymbol{\mu}_l^{d,\nu}$ as the nodal basis functions of the coarse grid spaces $\boldsymbol{S}^{(j)}$, $j \leq J$. Unfortunately, this choice might lead to vanishing coarse grid corrections at the actual contact zone, spoiling the convergence speed of the method. For an explanation of this behavior, see Section 3.2. Consequently, we need to construct extended search directions, which take the actual contact zone set into account. Since the actual shape of the extended search directions does not exert influence on the global convergence proof to be given in this section, let us assume the functions $\boldsymbol{\mu}_l^{1,\nu}, \ldots, \boldsymbol{\mu}_l^{d,\nu}$ are given. Their detailed construction is carried out in Section 3.2.

Adapting multigrid terminology, the leading projected block Gauß–Seidel relaxation $\mathcal{M}_J$ plays the role of a *fine grid smoother*, $\bar{\boldsymbol{u}}_J^{\nu} = \boldsymbol{w}_{n_J}^{\nu} = \mathcal{M}_J(\boldsymbol{u}_J^{\nu})$ is the *smoothed iterate* and subsequent corrections $\boldsymbol{v}_l$, $l = n_J + 1, \ldots, m$, are called *coarse grid corrections*. Note, that $\bar{\boldsymbol{u}}_J^{\nu} \in \mathcal{K}_J$ holds for all $\boldsymbol{u}_J^0 \in \boldsymbol{S}^{(J)}$.

Remark 3.1 *The subspaces $\boldsymbol{V}_l^{\nu}$, $\nu = n_J+1, \ldots, m$ are inspired by coarse grid corrections from linear multigrid theory. In contrast to linear multigrid methods, the shape of the "coarse grid" functions $\boldsymbol{\mu}_l^{1,\nu}, \ldots, \boldsymbol{\mu}_l^{d,\nu}$ depends on the actual iterate, i.e., on the actual guess of the discrete contact zone. Even in case of a known contact zone, they might differ from those used in linear multigrid methods. In Section 3.2, we give an interpretation of the extended search directions as local projections onto the set of admissible displacements.*

The *extended relaxation* based on the splitting (3.1) now reads as follows. Starting with the given ν–th iterate $\boldsymbol{u}_J^{\nu} = \boldsymbol{w}_0^{\nu} \in \mathcal{K}_J$, we compute a sequence of intermediate iterates $\boldsymbol{w}_l^{\nu} = \boldsymbol{w}_{l-1}^{\nu} + \boldsymbol{v}_l^{*,\nu}$, $l = 1, \ldots, m$. The new iterate is $\boldsymbol{u}_J^{\nu} = \boldsymbol{w}_m^{\nu}$. The corrections $\boldsymbol{v}_l^{*,\nu}$ are the unique solutions of the local subproblems

$$\boldsymbol{v}_l^{*,\nu} \in \mathcal{D}_l{}^{*,\nu}: \quad \mathcal{J}(\boldsymbol{w}_{l-1}^{\nu} + \boldsymbol{v}_l^{*,\nu}) \leq \mathcal{J}(\boldsymbol{w}_{l-1}^{\nu} + \boldsymbol{v}) \quad \boldsymbol{v} \in \mathcal{D}_l{}^{*,\nu}, \tag{3.2}$$

with closed, convex sets $\mathcal{D}_l{}^{*,\nu}$ defined by

$$\mathcal{D}_l{}^{*,\nu} = \{\boldsymbol{v} \in \boldsymbol{V}_l^{\nu} \mid \boldsymbol{w}_{l-1}^{\nu} + \boldsymbol{v} \in \mathcal{K}_J\} \subset \boldsymbol{V}_l^{\nu}. \tag{3.3}$$

For $\nu = n_J + 1, \ldots, m$, it might be too costly to check whether some $\boldsymbol{v} \in \boldsymbol{V}_l^\nu$ is contained in $\mathcal{D}_l^{*,\nu}$ or not. To see this, let us consider the coarse grid functions $\boldsymbol{\mu}_l^{1,\nu}, \ldots, \boldsymbol{\mu}_l^{d,\nu}$ spanning the subspace $\boldsymbol{V}_l^\nu$. Then by definition, the solution $\boldsymbol{v}_J^{*,\nu}$ of the local subproblem (3.2) is admissible with respect to all local constraints associated with nodes $p \in \operatorname{supp} \boldsymbol{\mu}_l^{i,\nu}$. Checking for these constraints would require additional prolongations, resulting in an overall complexity of $\mathcal{O}(n_J \log n_J)$ per iteration step. Hence, optimal corrections $\boldsymbol{v}_l^{*,\nu}$, $l = n_J + 1, \ldots, m$, are replaced by approximations $\boldsymbol{v}_l^\nu \in \boldsymbol{V}_l^\nu$ provided by approximate local problems

$$\boldsymbol{v}_l^\nu \in \mathcal{D}_l{}^\nu : \quad \mathcal{J}(\boldsymbol{w}_{l-1} + \boldsymbol{v}_l^\nu) \leq \mathcal{J}(\boldsymbol{w}_{l-1} + \boldsymbol{v}) \quad \boldsymbol{v} \in \mathcal{D}_l{}^\nu. \tag{3.4}$$

Since the constraints at the contact boundary Γ_S are given in *normal direction*, let us introduce for any $p \in \mathcal{N}^{(J)} \cap \Gamma_S$ the rotated local coordinate system $\{\boldsymbol{e}^i(p_l)\}$. We choose $\boldsymbol{e}^1(p_l) = \boldsymbol{n}(p_l)$ and extend $\boldsymbol{e}^1(p_l)$ by Householder reflection to an orthonormal system in $\mathbb{R}^d$ in the following way: As it is well known, any orthogonal transformation in $\mathbb{R}^d$ may be decomposed into a rotation R and a reflection with respect to some hypersurface H. Choosing R the identity on $\mathbb{R}^d$ and H the hypersurface with $0 \in H$ and H being normal to $\frac{1}{2}(\boldsymbol{e}^1(p_l) - \boldsymbol{E}^1(p_l))$ for given $\boldsymbol{e}^1(p_l)$, we can uniquely define the local unit vectors $\boldsymbol{e}^i(p_l)$. Denoting the Householder reflection at p_l defined above by $\boldsymbol{O}(p_l)$, we can write

$$\boldsymbol{e}^i(p_l) = \boldsymbol{O}(p_l)\boldsymbol{E}^i(p_l)\,, \qquad p \in \mathcal{N}^{(J)}\,. \tag{3.5}$$

Bearing in mind that $\boldsymbol{e}^1(p_l) = \boldsymbol{n}(p_l)$ for $1 \leq l \leq n_J$, i.e., $p \in \mathcal{N}^{(J)}$, it is clear that the non penetration condition (2.4) can be reformulated as, $1 \leq l \leq n_J$,

$$\mathcal{D}_l = \{\boldsymbol{v} \in \boldsymbol{V}_l \mid \boldsymbol{v}(p_l) \cdot \boldsymbol{e}^1(p_l) \in [\underline{\psi}_l^1, \overline{\psi}_l^1]\}$$

with $\overline{\psi}_l^1 = g(p_l)$ and $\underline{\psi}_l^1 = -\infty$. More generally, this gives rise to the following definition of the closed convex subsets $\mathcal{D}_l{}^\nu \subset \boldsymbol{V}_l^\nu$

$$\mathcal{D}_l{}^\nu = \{\boldsymbol{v} \in \boldsymbol{V}_l^\nu \mid \boldsymbol{v}(p_l) \cdot \boldsymbol{e}^i(p_l) \in [\underline{\psi}_l^{i,\nu}, \overline{\psi}_l^{i,\nu}] \text{ for all } i = 1, \ldots, d\}\,, \tag{3.6}$$

which are intended to approximate $\mathcal{D}_l{}^{*,\nu}$. Here, we have $\underline{\psi}_l^{i,\nu}, \overline{\psi}_l^{i,\nu} \in \mathbb{R}$. For completeness for any $p_l \in \mathcal{N}^{(J)} \cap \operatorname{int} \operatorname{supp} \boldsymbol{\mu}_l^{i,\nu}$, we set $\boldsymbol{e}^1(p_l) = \boldsymbol{E}^1(p_l)$ and $\overline{\psi}_l^{i,\nu} = \infty$ and $\underline{\psi}_l^{i,\nu} = -\infty$. It is clear that $\mathcal{D}_l{}^\nu = \boldsymbol{V}_l^\nu$ or, equivalently, $\underline{\psi}_l^{i,\nu} = -\infty$, $\overline{\psi}_l^{i,\nu} = +\infty$, if all $\boldsymbol{v} \in \boldsymbol{V}_l^\nu$ satisfy $\boldsymbol{v}|_{\Gamma_S} = 0$. In the sequel, we suppress the index ν if it is clear from the context.

Remark 3.2 *For any $p_l \in \mathcal{N}^{(J)} \cap \operatorname{int} \operatorname{supp} \boldsymbol{\mu}_l^{i,\nu}$, the choice of the local unit vector $\boldsymbol{e}^1(p_l)$ is arbitrary. In particular, in our implementation of the method all extended relaxation steps are carried out with respect to the canonical local coordinate system $\boldsymbol{E}^i$.*

In the following theorem, we prove the global convergence of the extended relaxation method.

Theorem 3.3 *Assume that*

$$\mathbf{0} \in \mathcal{D}_l \subset \mathcal{D}_l{}^*. \tag{3.7}$$

Then the approximate extended relaxation

$$\boldsymbol{u}_J^{\nu+1} = \boldsymbol{u}_J^{\nu} + \sum_{l=1}^{n_J} \boldsymbol{v}_l^* + \sum_{l=n_J+1}^{m} \boldsymbol{v}_l \tag{3.8}$$

with $\boldsymbol{v}_l^*$ and $\boldsymbol{v}_l$ computed from (3.2) and (3.4), respectively, is globally convergent.

Proof The sequence of iterates $\boldsymbol{u}_J^{\nu}$, $\nu = 0, 1, \ldots$, is bounded because our scheme (3.2) is *monotone* in the sense that

$$\mathcal{J}(\boldsymbol{u}_J^{\nu+1}) \leq \mathcal{J}(\boldsymbol{w}_{l+1}^{\nu}) \leq \mathcal{J}(\boldsymbol{w}_l^{\nu}) \leq \mathcal{J}(\bar{\boldsymbol{u}}_J^0) < \infty\ , \quad \nu = 1, 2, \ldots\ ,$$

and we have $\mathcal{J}(\boldsymbol{v}^{\nu}) \to \infty$ for any unbounded sequence $\boldsymbol{v}^{\nu} \in \boldsymbol{S}^{(J)}$.

As $\boldsymbol{u}_J^{\nu}$ is bounded and $\boldsymbol{S}^{(J)}$ has finite dimension, it is sufficient to show that each convergent subsequence of $\boldsymbol{u}_J^{\nu}$ converges to $\boldsymbol{u}_J$. Let $\boldsymbol{u}_J^{\nu_k}$ be an arbitrary, convergent subsequence of $\boldsymbol{u}_J^{\nu}$, with some limit $\boldsymbol{u}^* \in \boldsymbol{S}^{(J)}$,

$$\boldsymbol{u}_J^{\nu_k} \to \boldsymbol{u}^*, \quad k \to \infty. \tag{3.9}$$

It can easily be verified that $\mathcal{M}_J$ is continuous so that

$$\mathcal{M}_J(\boldsymbol{u}_J^{\nu_k}) \to \mathcal{M}_J(\boldsymbol{u}^*), \quad k \to \infty. \tag{3.10}$$

Again, monotonicity of the iteration implies

$$\mathcal{J}(\boldsymbol{u}_J^{\nu_{k+1}}) \leq \mathcal{J}(\boldsymbol{u}_J^{\nu_k+1}) \leq \mathcal{J}(\mathcal{M}_J(\boldsymbol{u}_J^{\nu_k})) \leq \mathcal{J}(\boldsymbol{u}_J^{\nu_k}).$$

In virtue of the convergence (3.9), (3.10) and the continuity of $\mathcal{J}$ on $\mathcal{K}_J$, this leads to

$$\mathcal{J}(\mathcal{M}_J(\boldsymbol{u}^*)) = \mathcal{J}(\boldsymbol{u}^*). \tag{3.11}$$

It can easily be seen that (3.11) holds, if and only if all local corrections of the projected block Gauß–Seidel relaxation applied to $\boldsymbol{u}^*$ are zero, i.e., $\mathcal{M}_J(\boldsymbol{u}^*) = \boldsymbol{u}^*$. As the finite element solution $\boldsymbol{u}_J$ is the only fixed point of the projected block Gauß–Seidel relaxation (cf. e.g. [Glo84, pp. 152]), we have shown $\boldsymbol{u}^* = \boldsymbol{u}_J$. This completes the proof.

□

Let us discuss the above result in some detail. It can easily be seen, that any intermediate iterate has to be admissible with respect to the fine grid constraints. If not, monotonicity of the method is lost and convergence cannot be guaranteed. In particular, the coarse grid corrections *must not violate the fine grid constraints.* During the nonlinear phase, this restriction might slow down the convergence of the method, but it is necessary for the global convergence. From numerical experiments, it can be seen that the non penetration condition must not be relaxed on the coarse grids. Even projecting the coarse grid corrections back onto the set $\mathcal{K}_j$, $1 \leq\leq J$ after prolongation, see [BC83], does not ensure global convergence of the method. The second important item to be mentioned is the *locality* of the constraints. The method relies heavily on the constraints being decoupled with respect to the nonlinearity. This is a necessary condition for the convergence of the fine grid Gauß–Seidel relaxation, see, e.g., [Glo84].

As a corollary, we also obtain convergence of the intermediate iterates

$$\boldsymbol{w}_l^\nu \to \boldsymbol{u}_J \qquad \nu \to \infty. \tag{3.12}$$

Indeed, the sequence $\boldsymbol{w}_l^\nu$, $\nu = 0, 1, \ldots$, is bounded and due to the monotonicity

$$\mathcal{J}(\boldsymbol{u}_J^{\nu_k+1}) \le \mathcal{J}(\boldsymbol{w}_l^{\nu_k}) \le \mathcal{J}(\boldsymbol{u}_J^{\nu_k})$$

and the continuity of $\mathcal{J}$ on $\mathcal{K}_J$ the limit $\boldsymbol{w}^*$ of an arbitrary convergent subsequence $\boldsymbol{w}_l^{\nu_k}$, $k = 0, 1, \ldots$, must satisfy $\mathcal{J}(\boldsymbol{w}^*) = \mathcal{J}(\boldsymbol{u}_J)$, giving $\boldsymbol{w}^* = \boldsymbol{u}_J$.

Up to now, we have been proving global convergence of the method. We note that if the contact boundary is known, Signorini's problem is linear. Does our method reduce to a linear subspace correction method in that case? And if so, how long does it take, to identify the discrete contact zone?

The answer to the first question is affirmative and is given in Theorem 3.5. The answer to the second question is much more difficult to find and is partly given in Lemma 3.4 and Theorem 3.5. It is shown, that the discrete contact zone is found after a finite number of iteration steps, but no estimate of the number of nonlinear iteration steps is given. On the other hand, during the search phase for the actual zone of contact, the energy is already reduced monotonously and the gained solution of the iteration process might be satisfactory, even if the discrete zone of contact has not been completely identified.

Let us now state more precisely what we understand to be the *discrete contact zone* or *coincidence set*. For given $\boldsymbol{w} \in \mathcal{K}_J$, the coincidence set is defined by

$$\mathcal{N}^{(J)^\bullet}(\boldsymbol{w}) = \{p \in \mathcal{N}^{(J)} \cap \Gamma_S \mid \boldsymbol{w}(p) \cdot \boldsymbol{n}(p) = g(p)\}.$$

No contact occurs at $\mathcal{N}^{(J)^\circ}(\boldsymbol{w}) = \mathcal{N}^{(J)} \setminus \mathcal{N}^{(J)^\bullet}(\boldsymbol{w})$. Note, that $\bar{\boldsymbol{u}}_J^\nu \in \mathcal{K}_J$ holds for all $\boldsymbol{u}_J^0 \in \boldsymbol{S}^{(J)}$.

Once the coincidence set $\mathcal{N}^{(J)^\bullet}(\boldsymbol{u}_J)$ is known, the minimization problem (2.28) can be rewritten as a *reduced linear problem*

$$a(\boldsymbol{u}_J, \boldsymbol{v}) = f(\boldsymbol{v}) \qquad \boldsymbol{v} \in \boldsymbol{S}^{(J)^\circ}, \tag{3.13}$$

where the subspace $\boldsymbol{S}^{(J)^\circ} \subset \boldsymbol{S}^{(J)}$ is defined by

$$\boldsymbol{S}^{(J)^\circ} = \{\boldsymbol{v} \in \boldsymbol{S}^{(J)} \mid \boldsymbol{v}(p) \cdot \boldsymbol{n}(p) = 0 \quad \text{for all } p \in \mathcal{N}^{(J)^\bullet}(\boldsymbol{u}_J)\}.$$

In the remainder of this section, we will show that the iteration (3.8) reduces asymptotically to a linear subspace correction method for the linear problem (3.13).

Lemma 3.4 *Assume that the discrete problem (2.28) is non-degenerate in the sense that*

$$f(\lambda_p^{(J)}\boldsymbol{n}(p)) - a(\boldsymbol{u}_J, \lambda_p^{(J)}\boldsymbol{n}(p)) > 0 \qquad \textit{for all } p \in \mathcal{N}^{(J)^\bullet}(\boldsymbol{u}_J) \tag{3.14}$$

and that the coarse grid spaces $\boldsymbol{V}_l^\nu$ in (3.1) are chosen such that

$$\boldsymbol{\mu}_l^{i,\nu}(p) \cdot \boldsymbol{n}(p) = 0 \qquad \textit{for all } p \in \mathcal{N}^{(J)^\bullet}(\bar{\boldsymbol{u}}_J^\nu) \tag{3.15}$$

holds for all $\nu \ge 0$.

Then there is a $\nu_0 \ge 0$ such that

$$\mathcal{N}^{(J)^\bullet}(\boldsymbol{u}_J^\nu) = \mathcal{N}^{(J)^\bullet}(\boldsymbol{u}_J) \qquad \textit{for all } \nu \ge \nu_0. \tag{3.16}$$

Proof Let $p \in \mathcal{N}^{(J)^\circ}(\boldsymbol{u}_J) \cap \Gamma_S$ or, equivalently, $\boldsymbol{u}_J(p) \cdot \boldsymbol{n}(p) < g(p)$. Convergence of $\boldsymbol{u}_J^\nu$ implies $\boldsymbol{u}_J^\nu(p) \cdot \boldsymbol{n}(p) < g(p)$ for all $\nu \geq \nu_0$ with sufficiently large ν_0. Hence,

$$\mathcal{N}^{(J)^\circ}(\boldsymbol{u}_J) \subset \mathcal{N}^{(J)^\circ}(\boldsymbol{u}_l^\nu) \qquad \text{for all } \nu \geq \nu_0.$$

Now, let $p_l \in \mathcal{N}^{(J)^\bullet}(\boldsymbol{u}_J)$. As a consequence of (3.14) and the convergence of the intermediate iterates (3.12), we obtain

$$f(\lambda_{p_l}^{(J)}\boldsymbol{n}(p_l)) - a(\boldsymbol{w}_l^\nu, \lambda_{p_l}^{(J)}\boldsymbol{n}(p_l)) > 0 \qquad \text{for all } \nu \geq \nu_0 \tag{3.17}$$

for $l = 1, \ldots, n_J$ and sufficiently large ν_0. Now assume that $p_l \notin \mathcal{N}^{(J)^\bullet}(\bar{\boldsymbol{u}}_J^\nu)$ or, equivalently,

$$\bar{\boldsymbol{u}}_J^\nu(p_l) \cdot \boldsymbol{n}(p_l) = \boldsymbol{w}_{l-1}^\nu(p_l) \cdot \boldsymbol{n}(p_l) + \boldsymbol{v}_l^{*,\nu}(p_l) \cdot \boldsymbol{n}(p_l) < g(p_l).$$

In the light of (3.2), the correction $\boldsymbol{v}_l^{*,\nu}$ then satisfies the variational equality

$$a(\boldsymbol{v}_l^{*,\nu}, \boldsymbol{v}) = f(\boldsymbol{v}) - a(\boldsymbol{w}_{l-1}^\nu, \boldsymbol{v}) \qquad \boldsymbol{v} \in \boldsymbol{V}_l.$$

Hence, $f(\boldsymbol{v}) - a(\boldsymbol{w}_l^\nu, v) = 0$ $\boldsymbol{v} \in \boldsymbol{V}_l$ in contradiction to (3.17). We have shown

$$\mathcal{N}^{(J)^\bullet}(\boldsymbol{u}_J) \subset \mathcal{N}^{(J)^\bullet}(\bar{\boldsymbol{u}}_J^\nu) \qquad \text{for all } \nu \geq \nu_0.$$

It is clear from (3.15) that $\mathcal{N}^{(J)^\bullet}(\bar{\boldsymbol{u}}_J^\nu) \subset \mathcal{N}^{(J)^\bullet}(\boldsymbol{u}_J^{\nu+1})$, yields $\mathcal{N}^{(J)^\bullet}(\boldsymbol{u}_J) \subset \mathcal{N}^{(J)^\bullet}(\boldsymbol{u}_J^\nu)$ for all $\nu \geq \nu_0 + 1$. This completes the proof.

□

We recall that continuous versions of the non-degeneracy condition (3.14) provide stability of the free boundary (cf. e.g. [Rod87, pp. 198]). As a by product of the proof, we also get

$$\mathcal{N}^{(J)^\bullet}(\bar{\boldsymbol{u}}_J^\nu) = \mathcal{N}^{(J)^\bullet}(\boldsymbol{u}_J) \qquad \nu \geq \nu_0. \tag{3.18}$$

The non-degeneracy condition (3.14) can be interpreted in terms of contact stresses, since, on the basis of Green's formula, $a(\boldsymbol{u}_J, \lambda_p^{(J)}\boldsymbol{n}(p)) - f(\lambda_p^{(J)}\boldsymbol{n}(p))$ can be seen to be the discrete normal stresses at p. Consequently, if $a(\boldsymbol{u}_J, \lambda_p^{(J)}\boldsymbol{n}(p)) = f(\lambda_p^{(J)}\boldsymbol{n}(p))$, the body hits the obstacle at p but the contact does not cause any deformation. Thus, the coincidence set may change without changing the solution, which is an obvious instability. The second condition, i.e., condition (3.15), guarantees that correction in the direction of $\boldsymbol{\mu}_l^{i,\nu}$ does not affect the actual guess of the coincidence set $\mathcal{N}^{(J)^\bullet}(\bar{\boldsymbol{u}}_J^\nu)$. This is in general not true, if the extended search directions are chosen as the coarse grid functions of the standard multilevel basis. As a consequence of (3.15), the coarse grid corrections from $\boldsymbol{V}_l^\nu$, $\nu = n_{J+1}, \ldots, m$, asymptotically reduce to a linear subspace correction for the reduced problem (3.13) provided that no constraints are active. This can be ensured by appropriate choice of local obstacles $\underline{\psi}_l^{i,\nu}$, $\overline{\psi}_l^{i,\nu}$.

In the following theorem, we show that the extended splitting (3.1) degenerates to a linear subspace correction method. To this end, let us assume, the coincidence set has been identified already and that $\boldsymbol{V}_l^\nu$ solely depends on $\mathcal{N}^{(J)^\bullet}(\bar{\boldsymbol{u}}_J^\nu)$, i.e. $\boldsymbol{V}_l^\nu = \boldsymbol{V}_l(\mathcal{N}^{(J)^\bullet}(\bar{\boldsymbol{u}}_J^\nu))$.

Then, we expect linear behavior of our method. In particular, any coarse grid correction should be admissible with respect to the local constraints. This is only possible, if the inner approximation (3.6) do not degenerate. Summing up, we call a sequence of local obstacles $\underline{\psi}_l^{i,\nu}$, $\overline{\psi}_l^{i,\nu}$, $\nu \geq 0$, *quasioptimal* (cf. [Kor94]), if convergence of the intermediate iterates $\boldsymbol{w}_l^\nu$ (see (3.12)) and convergence of the coincidence sets $\mathcal{N}^{(J)^\bullet}(\bar{\boldsymbol{u}}_J^\nu)$ (see (3.18)) implies that there is a positive number ψ^*, independent of ν, and some $\nu_0 \geq 0$, such that

$$\underline{\psi}_l^{i,\nu} \leq -\psi^* < 0 < \psi^* \leq \overline{\psi}_l^{i,\nu} \qquad \text{for all } \nu \geq \nu_0. \tag{3.19}$$

holds for all $i = 1, \ldots, d$ and $l = n_J + 1, \ldots, m$. For construction of such quasioptimal obstacles, we refer to Section 3.3 Now, we can give the following

Theorem 3.5 *Assume that the assumptions of Lemma 3.4 are satisfied. Assume further that the coarse grid spaces only depend on the actual guess of the coincidence set, i.e., $\boldsymbol{V}_l^\nu = \boldsymbol{V}_l(\mathcal{N}^{(J)^\bullet}(\bar{\boldsymbol{u}}_J^\nu))$, and that the local obstacles $\underline{\psi}_l^i$, $\overline{\psi}_l^i$ are quasioptimal in the sense of (3.19).*

Then there is a $\nu_0 > 0$ such that the approximate extended relaxation (3.8) is reducing to the linear successive subspace correction induced by the splitting

$$\boldsymbol{S}^{(J)^\circ} = \boldsymbol{V}_1^\circ + \cdots + \boldsymbol{V}_m^\circ$$

with

$$\boldsymbol{V}_l^\circ = \begin{cases} \boldsymbol{V}_l \cap \boldsymbol{S}^{(J)^\circ}, & l = 1, \ldots, n_J \\ \boldsymbol{V}_l(\mathcal{N}^{(J)^\bullet}(\boldsymbol{u}_J)), & l = n_J + 1 \ldots, m \end{cases}$$

as applied to the reduced linear problem (3.13).

Proof Let $l = 1, \ldots, n_J$. There is nothing to show, if $p_l \notin \Gamma_S$. Let $p_l \in \mathcal{N}^{(J)^\bullet}(\boldsymbol{u}_J)$. As a consequence of Lemma 3.4 and (3.18), the normal components of $\boldsymbol{v}_l^{*,\nu}$ are zero, if ν is sufficiently large. In this case, we have $\boldsymbol{v}_l^{*,\nu} \in \boldsymbol{V}_l \cap \boldsymbol{S}^{(J)^\circ} = \boldsymbol{V}_l^\circ$ and $\boldsymbol{V}_l^\circ = \mathcal{D}_l^{*,\nu}$ so that

$$\mathcal{J}(\boldsymbol{w}_{l-1} + \boldsymbol{v}_l^{*,\nu}) \leq \mathcal{J}(\boldsymbol{w}_{l-1} + \boldsymbol{v}) \qquad \boldsymbol{v} \in \boldsymbol{V}_l^\circ. \tag{3.20}$$

In the remaining case $p_l \in \mathcal{N}^{(J)^\circ}(\boldsymbol{u}_J) \cap \Gamma_S$, it follows directly from (3.18) that $\boldsymbol{v}_l^{*,\nu}$ must satisfy (3.20) with $\boldsymbol{V}_l^\circ = \boldsymbol{V}_l$ for sufficiently large ν.

Now, let $l = n_J + 1, \ldots, m$. Then, exploiting (3.18) we get $\boldsymbol{V}_l^\nu = \boldsymbol{V}_l(\mathcal{N}^{(J)^\bullet}(\boldsymbol{u}_J))$, i.e., $\boldsymbol{V}_l^\nu = \boldsymbol{V}_l^\circ$, for sufficiently large ν. Convergence of the intermediate iterates $\boldsymbol{w}_l^\nu$ (3.12) implies that the corrections $\boldsymbol{v}_l^\nu$ must tend to zero. Utilizing (3.19), it follows that

$$\boldsymbol{v}_l^\nu(p_l) \cdot \boldsymbol{e}^i(p_l) \in [-\psi^*, \psi^*] \subset (\underline{\psi}_l^{i,\nu}, \overline{\psi}_l^{i,\nu})$$

holds for sufficiently large ν. In this case,

$$\mathcal{J}(\boldsymbol{w}_{l-1} + \boldsymbol{v}_l^\nu) \leq \mathcal{J}(\boldsymbol{w}_{l-1} + \boldsymbol{v}) \qquad \boldsymbol{v} \in \boldsymbol{V}_l^\circ$$

and the assertion follows.

□

From Theorem 3.5, the differences between leading and extended relaxation can be seen. The actual discrete contact zone, or coincidence set, is fixed by the leading fine grid relaxation $\mathcal{M}_J$, whereas the coarse grid corrections are used for acceleration of the convergence only. Unfortunately, the coarse grid corrections must not contribute corrections in normal directions at critical nodes, leading to functions with large support but zero value at isolated points in the interior of their support. Such functions incorporate the boundary conditions associated with the actual coincidence set. Note, that the construction of the coarse grid functions and of the quasioptimal obstacles is crucial for the convergence speed, since omitting only a single coarse grid correction can spoil the convergence of the multigrid method, see also Chapter 5

Using optimal corrections from (3.2) instead of quasioptimal approximations (cf. Theorem 3.5), we asymptotically get the same linear subspace correction method for (3.13). However, using quasioptimal obstacles gives optimal complexity per iteration step.

3.2 Truncated Coarse Grid Functions

In this section, we explicitly construct extended search directions $\boldsymbol{\mu}_l^{i,\nu}$ with property (3.15). We give an interpretation of the resulting coarse grid corrections as *local projections* onto the set $\mathcal{K}_J$ of admissible displacements and show, how the constructed search directions can be obtained from the nodal basis functions of the fine grid space $\boldsymbol{S}^{(J)}$ by careful linear combination and *truncation*. This has been proposed for *scalar* obstacle problems in [Kor97a].

Before describing the truncated basis functions in detail, we give an example illustrating the main difficulty when using extended relaxation for Signorini's problem in linear elasticity. Let us recall, that any non admissible coarse grid correction can spoil the nonlinear convergence of the method and that a vanishing coarse grid correction can spoil the linear convergence speed. Thus, we need suitable coarse grid corrections corresponding to some *admissible* low frequency sliding of the body along the possibly curvilinear contact boundary. Clearly, functions representing such kind of displacement cannot be linear. Let us consider a simple two dimensional example of a curvilinear contact boundary as

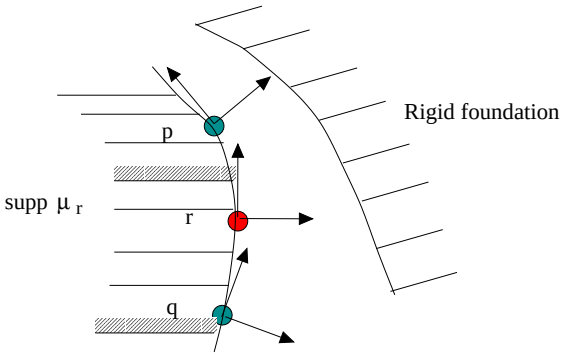

Figure 3.1: Coarse grid correction at curvilinear boundary

depicted in Figure 3.1. Here, we assume contact at the two different nodes p and q with

different outer normals $\boldsymbol{n}(p) \neq \boldsymbol{n}(q)$. Then, it is clear from (3.15), that any coarse grid correction $\boldsymbol{v}_r$ in direction of $\boldsymbol{\mu}_r$ with $p, q \in \operatorname{int} \operatorname{supp} \boldsymbol{\mu}_r$ must not contribute corrections in direction of $\boldsymbol{n}(p)$ and $\boldsymbol{n}(q)$, respectively. That is, the correction $\boldsymbol{v}_r$ has to satisfy $\boldsymbol{v}_r \in \operatorname{span}\{\boldsymbol{t}_p\}$ and $\boldsymbol{v}_r \in \operatorname{span}\{\boldsymbol{t}_p\}$, with $\boldsymbol{t}_p \neq \boldsymbol{t}_q$ the tangential vectors. This rules out a linear search direction $\boldsymbol{\mu}_r$. The basic idea is now to use *truncated* search directions, being piecewise linear functions only with respect to the finest grid $\mathcal{T}_J$. Note, that this problem does not show up if $\boldsymbol{n}(p) = \boldsymbol{n}(q)$. Then, the tangent vector $\boldsymbol{t}(p)$ and $\boldsymbol{t}(q)$ are collinear and simply setting the coarse grid correction in normal direction to zero would give an admissible coarse grid correction $\boldsymbol{v}_r$.

As a first attempt to construct suitable functions $\boldsymbol{\mu}_l^{i,\nu}$, let us consider linear combinations of the truncated nodal basis functions $\boldsymbol{\mu}_s^{(J)}$, given with respect to the local coordinate system $\boldsymbol{e}_s^i$ for $s \in \{p, q, r\}$. In order to guarantee (3.15), we *truncate* the fine grid basis functions $\boldsymbol{\mu}_s^{(J)}$, $s \in \{p, q, r\}$ by

$$\left(\boldsymbol{\mu}_p^{(J)}\right)^i = \begin{cases} 0 & \text{if } i = 1 \\ \lambda_p^{(J)} \boldsymbol{e}^i(p) & \text{else} \end{cases},$$

Then, condition (3.15) is true for any linear combination of the truncated functions $\boldsymbol{\mu}_s^{(J)}$. Applying the standard restriction, we get

$$\left(\widetilde{\boldsymbol{\mu}}_r^{(J-1)}\right)^i = \sum_{s \in \{p,q,r\}} \lambda_s^{(J-1)}(r) \left(\boldsymbol{\mu}_q^{(J)}\right)^j \tag{3.21}$$

as a candidate coarse grid function. Unfortunately, extended relaxation in direction of $\widetilde{\boldsymbol{\mu}}_r^{(J-1)}$ does not improve convergence speed, since in case of no contact the extended relaxation using $\widetilde{\boldsymbol{\mu}}_r^{(J-1)}$ does not degenerate to a standard multigrid method. More precisely, the approximation property of the coarse grid spaces is lost.

Here, the remedy is to use appropriate *weights* in definition (3.21) instead of $\lambda_s^{(J-1)}(r)$, see Lemma 3.6. These weights take care of the extended relaxation degenerating to a standard multigrid method in case of no contact. Moreover, the convergence speed of the resulting method is comparable to linear multigrid methods applied to Signorini's problem with known contact stresses.

Let us now formulate the ideas given above more precisely. Assume that $\mathcal{T}_J$ is resulting from J refinements of a coarse triangulation $\mathcal{T}_0$. Though the algorithms and convergence results to be presented can be easily generalized to the non-uniform case, let us assume for the moment that the triangulations are uniformly refined. More precisely, for the two dimensional case, each triangle $t \in \mathcal{T}_k$ is subdivided into four congruent subtriangles in order to produce the next triangulation $\mathcal{T}_{k+1}$. For the three dimensional case, each tetrahedron is subdivided into eight congruent subtetrahedrons.

Using this hierarchy of grids and the corresponding hierarchy of finite element spaces, we now choose suitable spaces $\boldsymbol{V}_{n_J+1}, \ldots, \boldsymbol{V}_m$. Each space $\boldsymbol{V}_l = \boldsymbol{V}_{l(p,k)}$ is associated with a node $p \in \mathcal{N}_k$ on some refinement level $k \leq J-1$. We frequently use the notation $\boldsymbol{V}_p^{(k)} = \boldsymbol{V}_{l(p,k)}$, $\boldsymbol{\mu}_p^{(k)} = \boldsymbol{\mu}_{l(p,k)}$. The ordering $l = l(p,k)$ is taken from fine to coarse, i.e., $l(p,k) \leq l(q,j)$ implies $k \geq j$.

For each $p \in \mathcal{N}^{(J)}$ we choose the local orthonormal system $\boldsymbol{e}^i(p) \in \mathbb{R}^d,\ i = 1, \dots d$, see (3.5) with the property $\boldsymbol{e}^1(p) = \boldsymbol{n}(p)$ for all $p \in \mathcal{N}^{(J)} \cap \Gamma_S$. Starting with

$$\left(\boldsymbol{\mu}_p^{(J)}\right)^i = \begin{cases} 0 & \text{if } i = 1 \text{ and } p \in \mathcal{N}^{(J)\bullet}(\bar{\boldsymbol{u}}_J^\nu) \\ \lambda_p^{(J)} \boldsymbol{e}^i(p) & \text{else} \end{cases}, \tag{3.22}$$

we recursively define truncated basis functions

$$\left(\boldsymbol{\mu}_p^{(k-1)}\right)^i = \sum_{q \in \mathcal{N}_k} \lambda_p^{(k-1)}(q)\ \boldsymbol{e}^i(p) \cdot \boldsymbol{e}^j(q)\ \left(\boldsymbol{\mu}_q^{(k)}\right)^j \tag{3.23}$$

and we set

$$\boldsymbol{V}_p^{(k)} = \operatorname{span}\{\left(\boldsymbol{\mu}_p^{(k)}\right)^1, \dots, \left(\boldsymbol{\mu}_p^{(k)}\right)^d\}, \qquad k = J-1, \dots, 0. \tag{3.24}$$

Note, that $\operatorname{supp}\left(\boldsymbol{\mu}_p^{(k)}\right)^i = \operatorname{supp} \lambda_p^{(k)}$ and that $\left(\boldsymbol{V}_p^{(k)}\right)^\nu = \boldsymbol{V}_p^{(k)}(\mathcal{N}^{(J)\bullet}(\bar{\boldsymbol{u}}_J^\nu))$ only depends on $\mathcal{N}^{(J)\bullet}(\bar{\boldsymbol{u}}_J^\nu)$.

Lemma 3.6 *Let $0 \le k \le J$, $p \in \mathcal{N}_k$ and $1 \le i, j \le d$. Then*

$$\left(\boldsymbol{\mu}_p^{(k)}\right)^i(q) \cdot \boldsymbol{e}^j(q) = \begin{cases} 0 & \text{if } j = 1 \text{ and } q \in \mathcal{N}^{(J)\bullet}(\bar{\boldsymbol{u}}_J^\nu) \\ \lambda_p^{(k)}(q)\ \boldsymbol{e}^i(p) \cdot \boldsymbol{e}^j(q) & \text{else} \end{cases} \tag{3.25}$$

holds for all $q \in \mathcal{N}^{(J)}$.

Proof The assertion is clear for $k = J$. Assume that (3.25) holds for some $k \le J$. If $j \ne 1$ or $q \notin \mathcal{N}^{(J)\bullet}(\bar{\boldsymbol{u}}_J^\nu)$ we obtain (summation on s)

$$\begin{aligned} \left(\boldsymbol{\mu}_p^{(k-1)}\right)^i(q) \cdot \boldsymbol{e}^j(q) &= \sum_{r \in \mathcal{N}_k} \lambda_p^{(k-1)}(r)\ \boldsymbol{e}^i(p) \cdot \boldsymbol{e}^s(r)\ \left(\boldsymbol{\mu}_r^{(k)}\right)^s(q) \cdot \boldsymbol{e}^j(q) \\ &= \sum_{r \in \mathcal{N}_k} \lambda_p^{(k-1)}(r)\ \lambda_r^{(k)}(q)\ \boldsymbol{e}^i(p) \cdot \boldsymbol{e}^s(r)\ \boldsymbol{e}^s(r) \cdot \boldsymbol{e}^j(q) \\ &= \sum_{r \in \mathcal{N}_k} \lambda_p^{(k-1)}(r)\ \lambda_r^{(k)}(q)\ \boldsymbol{e}^i(p) \cdot \boldsymbol{e}^j(q) \\ &= \lambda_p^{(k-1)}(q)\ \boldsymbol{e}^i(p) \cdot \boldsymbol{e}^j(q) \end{aligned}$$

exploiting the identity

$$\boldsymbol{e}^i(p) \cdot \boldsymbol{e}^s(r)\ \boldsymbol{e}^s(r) \cdot \boldsymbol{e}^j(q) = \boldsymbol{e}^i(p) \cdot \boldsymbol{e}^j(q), \qquad p, q, r \in \mathcal{N}^{(J)}.$$

Now let $q \in \mathcal{N}^{(J)\bullet}(\bar{\boldsymbol{u}}_J^\nu)$. Then definition (3.22) yields

$$\left(\boldsymbol{\mu}_p^{(J)}\right)^i(q) \cdot \boldsymbol{n}(q) = 0, \qquad \text{for all } p \in \mathcal{N}^{(J)}.$$

Using (3.23), the assertion now follows by induction.

□

Lemma 3.6 reveals the construction principle of coarse grid spaces $V_p^{(k)}$. If $q \notin \mathcal{N}^{(J)\bullet}(\bar{\boldsymbol{u}}_J^\nu)$, we get

$$\big(\boldsymbol{\mu}_p^{(k)}\big)^i(q) = \lambda_p^{(k)}(q)\; \boldsymbol{e}^i(p)\,.$$

Hence, $\boldsymbol{V}_p^{(k)} = \operatorname{span}\{\lambda_p^{(k)}\boldsymbol{E}^1, \ldots, \lambda_p^{(k)}\boldsymbol{E}^d\}$, if there is no contact in the neighborhood of p or, more precisely, if $\operatorname{int}\operatorname{supp}\lambda_p^{(k)} \cap \mathcal{N}^{(J)\bullet}(\bar{\boldsymbol{u}}_J^\nu) = \emptyset$. In this case, we obtain the same local correction $\boldsymbol{v}_p^{(k)}$ as classical multigrid method with canonical Galerkin restriction and block Gauß–Seidel smoother. On the other hand, if $q \in \mathcal{N}^{(J)\bullet}(\bar{\boldsymbol{u}}_J^\nu)$, Lemma 3.6 provides

$$\big(\boldsymbol{\mu}_p^{(k)}\big)^i(q) \cdot \boldsymbol{n}(q) = 0\,.$$

Hence, the spaces $\boldsymbol{V}_p^{(k)}$ satisfy condition (3.15).

For a better understanding of these truncated functions, let us consider a simple one-dimensional situation as depicted in Figure 3.2. Here, the truncated coarse grid function $\mu_r^{(J-1)}$ has zero value at the point $p \in \operatorname{int}\operatorname{supp}\mu_r^{(J-1)}$. There is no contribution of the

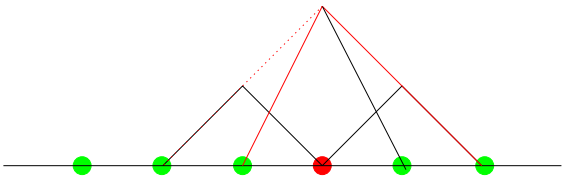

Figure 3.2: Truncated functions (red) for the scalar case

coarse grid correction in direction of the function $\mu_r^{(J)}$. In case of Signorini's problem, the situation is slightly more complicated. Again, let us consider the situation as depicted in Figure 3.1 and let the correction in direction of $\boldsymbol{\mu}_r^{(J-1)}$ be denoted by $\boldsymbol{c}_r$. The displacement $\boldsymbol{c}_r$ can be interpreted as a low frequency sliding of the body along the curvilinear contact boundary, that is, the whole body is assumed to move in direction of $\boldsymbol{c}_r$. In order to satisfy the constraints $\boldsymbol{c}_r(p) \in \operatorname{span}\{\boldsymbol{t}_p\}$ and $\boldsymbol{c}_r(q) \in \operatorname{span}\{\boldsymbol{t}_q\}$, at $s = p, q \in \mathcal{N}^{(J)}$, we take the weighted *projection* of $\boldsymbol{c}_r(r)$ onto $\operatorname{span}\{\boldsymbol{t}_s\}$ with respect to the Euclidian scalar product, see Figure 3.3. That is, for $\boldsymbol{c}_r(r) = \boldsymbol{e}^i(r)c_r^i$, we have

$$\big(\boldsymbol{c}_r(s)\big)^i \;=\; \lambda_s^{(J)}\,\boldsymbol{e}^i(s)\cdot\boldsymbol{c}_r^j(s), \qquad s = p, q \in \mathcal{N}^{(J)}\,.$$

Roughly speaking, coarse grid basis functions $\big(\boldsymbol{\mu}_p^{(k)}\big)^i$ are obtained by careful truncation and bending of nodal basis functions $\lambda_p^{(k)}\boldsymbol{e}^i(p)$. In case of constant normal directions we get $\boldsymbol{e}^i(p) = \boldsymbol{e}^i(q)$ for all $p, q \in \mathcal{N}^{(J)}$ and (3.23) reduces to the canonical restriction

$$\big(\boldsymbol{\mu}_p^{(k-1)}\big)^i = \sum_{q\in\mathcal{N}_k} \lambda_p^{(k-1)}(q)\,\big(\boldsymbol{\mu}_q^{(k)}\big)^i. \tag{3.26}$$

Note, that canonical restriction (3.26) could be used in the case of spatially varying normals as well, because the resulting coarse grid basis functions would still satisfy (3.15).

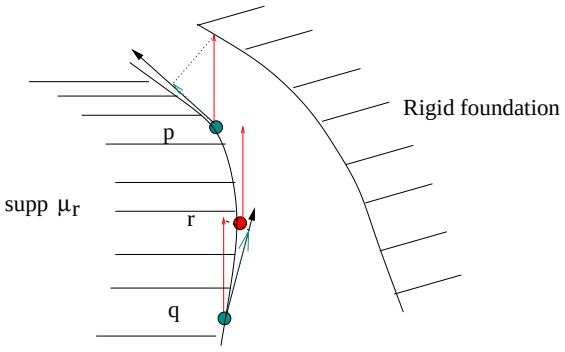

Figure 3.3: Interpretation of coarse gird correction $\boldsymbol{c}_r$ (red) as local projection (cyan)

However, varying normals result in a large energy of such coarse grid functions As a consequence poor convergence rates of the corresponding subspace correction are obtained. Similar effects caused by jumping coefficients have been discussed, e.g., by Wan et al. [WCS98].

In order to complete the construction of our multigrid method, we now describe the recursive construction of local obstacles $\underline{\psi}^i_{l(p,k)} = (\underline{\psi}^{(k)}_p)^i$, $\overline{\psi}^i_{l(p,k)} = (\overline{\psi}^{(k)}_p)^i$, $k = J-1, \ldots, 0$, occurring in (3.3). Starting with

$$\begin{array}{rclrcll} \left(\underline{\psi}^{(J)}_p\right)^1 & = & -\infty, & \left(\overline{\psi}^{(J)}_p\right)^1 & = & g(p) - \bar{\boldsymbol{u}}^\nu_J(p)\cdot\boldsymbol{n}(p) & \\ \left(\underline{\psi}^{(J)}_p\right)^i & = & -\infty & \left(\overline{\psi}^{(J)}_p\right)^i & = & +\infty & i = 2,\ldots,d \end{array} \qquad p \in \mathcal{N}^{(J)}, \tag{3.27}$$

we assume that local obstacles $(\underline{\psi}^{(k)}_p)^i$, $(\overline{\psi}^{(k)}_p)^i$ have been constructed for some $k \leq J$.

3.3 Monotone Restrictions

From Theorem 3.5 and Lemma 3.7, we know that on the discrete level we asymptotically have to solve a linear problem. Thus, it is desirable to "forget" about the contact condition (2.4) after the coincidence set has been determined. Or, in terms of computational effort, we require the extended relaxation (3.8) to be of the same complexity as the same extended relaxation *without* obstacles. As has been mentioned above, checking the coarse grid corrections for being admissible with respect to the fine grid constraints spoils the optimal complexity per iteration step. In particular, one step of the extended relaxation (3.8) for the asymptotic linear problem may be also interpreted as one step of a linear $\mathcal{V}$–cycle.

Now, to preserve optimal complexity of one iteration step, we approximate the fine grid constraints on the coarse grid by suitable *coarse grid obstacles*. These coarse grid obstacles ensure computational work of $\mathcal{O}(n_J)$ per iteration step and, by construction, take care of the admissibility of the solution. Construction of the coarse grid obstacles is done by means of *monotone restrictions*, which are given in this section.

For fixed $p \in \mathcal{N}_{k-1}$ and $i = 1, \dots d$ local obstacles on the next coarser level are now obtained by *monotone restriction* defined as follows

$$(\underline{\psi}_p^{(k-1)})^i = d_i^{-1} \max(\Psi_p^i)_- , \qquad (\overline{\psi}_p^{(k-1)})^i = d_i^{-1} \min(\Psi_p^i)_+ . \tag{3.28}$$

The factor $d_i \leq d$ denotes the number of non-zero entries in the i–th row of $\big(\boldsymbol{e}^i(p) \cdot \boldsymbol{e}^j(q)\big)_{i,j=1,\dots d}$. Further let Ψ_p^i be defined by

$$\Psi_p^i = \left\{ \frac{(\underline{\psi}_q^{(k)})^j}{\boldsymbol{e}^i(p) \cdot \boldsymbol{e}^j(q)}, \frac{(\overline{\psi}_q^{(k)})^j}{\boldsymbol{e}^i(p) \cdot \boldsymbol{e}^j(q)} \;\middle|\; q \in \text{int supp } \lambda_p^{(k-1)} \cap \mathcal{N}_k,\ j = 1, \dots, d, \right.$$
$$\left. j \neq 1 \text{ or } q \notin \mathcal{N}^{(J)^\bullet}(\bar{\boldsymbol{u}}_J^\nu),\ \boldsymbol{e}^i(p) \cdot \boldsymbol{e}^j(q) \neq 0 \right\}. \tag{3.29}$$

The sets $(\Psi_p^i)_-$ and $(\Psi_p^i)_+$ contain the non-positive and non-negative elements of Ψ_p^i, respectively. Observe that the weights $\boldsymbol{e}^i(p) \cdot \boldsymbol{e}^j(q)$ may be positive, negative or zero.

Once all local corrections $\boldsymbol{v}_p^{(k-1)} = \boldsymbol{v}_{l(p,k)}$ on level $k-1$ have been computed from (3.4), we update the obstacles according to

$$\begin{aligned} \big(\underline{\psi}_p^{(k-1)}\big)^i &\to \big(\underline{\psi}_p^{(k-1)}\big)^i - \boldsymbol{v}_p^{(k-1)}(p) \cdot \boldsymbol{e}^i(p) \\ \big(\overline{\psi}_p^{(k-1)}\big)^i &\to \big(\overline{\psi}_p^{(k-1)}\big)^i - \boldsymbol{v}_p^{(k-1)}(p) \cdot \boldsymbol{e}^i(p) \end{aligned} \qquad p \in \mathcal{N}_{k-1}. \tag{3.30}$$

Monotone restriction (3.28) and update (3.30) are repeated inductively until the coarsest level is reached. It is clear by construction that

$$(\underline{\psi}_p^{(k)})^i \leq 0 \leq (\overline{\psi}_p^{(k)})^i \qquad k = J-1, \dots, 0. \tag{3.31}$$

Lemma 3.7 *Let $\bar{\boldsymbol{u}}_J^\nu \in \mathcal{K}_J$. Then the subsets $\mathcal{D}_l = \mathcal{D}_p^{(k)}$,*

$$\mathcal{D}_p^{(k)} = \{\boldsymbol{v} \in \boldsymbol{V}_p^{(k)} \mid \boldsymbol{v}(p) \cdot \boldsymbol{e}^i(p) \in [(\underline{\psi}_p^{(k)})^i, (\overline{\psi}_p^{(k)})^i] \ \textit{ for all } i = 1, \dots, d\}$$

satisfy condition (3.7).

Proof From corrections on levels J to $k \leq J$, we obtain the intermediate iterate $\boldsymbol{w}^{(k)}$,

$$\boldsymbol{w}^{(k)} = \boldsymbol{u}_J^\nu + \sum_{j=k}^{J} \sum_{p \in \mathcal{N}_j} \boldsymbol{v}_p^{(k)}.$$

We show by induction that

$$\boldsymbol{w}^{(k)} + \sum_{p \in \mathcal{N}_{k-1}} \boldsymbol{z}_p^{(k-1)} \in \mathcal{K}_J \qquad \boldsymbol{z}_p^{(k-1)} \in \mathcal{D}_p^{(k-1)}, \qquad k = J, \dots, 1, \tag{3.32}$$

holds after the monotone restriction (3.28). Simultaneously, we prove the auxiliary result

$$\boldsymbol{w}^{(k)} + \sum_{p\in\mathcal{N}_k} \boldsymbol{z}_p^{(k)} \in \mathcal{K}_J \qquad \boldsymbol{z}_p^{(k)} \in \mathcal{D}_p^{(k)}, \qquad k = J,\dots,0, \tag{3.33}$$

where $\mathcal{D}_p^{(k)}$ is taken after update (3.30) for $k = J-1,\dots,0$.

Assertion (3.33) is clear for $k = J$. Assuming that (3.33) holds for some $k = J,\dots,1$ we now prove (3.32). Let $\boldsymbol{z}_p^{(k-1)} = (\boldsymbol{z}_p^{(k-1)})_i\ (\boldsymbol{\mu}_p^{(k-1)})^i \in \mathcal{D}_p^{(k-1)}$. Inserting (3.23), we get the representation

$$\sum_{p\in\mathcal{N}_{k-1}} \boldsymbol{z}_p^{(k-1)} = \sum_{p\in\mathcal{N}_k} \boldsymbol{z}_p^{(k)}$$

with

$$\boldsymbol{z}_p^{(k)} = \sum_{q\in\mathcal{N}_{k-1}} \lambda_q^{(k-1)}(p)\ (\boldsymbol{z}_q^{(k-1)})_i\ \boldsymbol{e}^i(q)\cdot\boldsymbol{e}^j(p)\ \left(\boldsymbol{\mu}_p^{(k)}\right)^j \in \boldsymbol{V}_p^{(k)}.$$

Let $j \neq 1$ or $p \notin \mathcal{N}^{(J)^\bullet}(\bar{\boldsymbol{u}}_J^\nu)$. Exploiting (3.28), we get

$$\boldsymbol{z}_p^{(k)}(p)\cdot\boldsymbol{e}^j(p) = \sum_{q\in\mathcal{N}_{k-1}} \lambda_q^{(k-1)}(p)\ (\boldsymbol{z}_p^{(k-1)})_i\ \boldsymbol{e}^i(q)\cdot\boldsymbol{e}^j(p) \in [(\underline{\psi}_p^{(k)})^j, (\overline{\psi}_p^{(k)})^j]$$

for all $j = 1,\dots,d$. In the remaining case, $j = 1$ and $p \in \mathcal{N}^{(J)^\bullet}(\bar{\boldsymbol{u}}_J^\nu)$, Lemma 3.5 leads to

$$\boldsymbol{z}_p^{(k)}(p)\cdot\boldsymbol{e}^1(p) = 0 \in \left[(\underline{\psi}_p^{(k)})^1, (\overline{\psi}_p^{(k)})^1\right].$$

Hence, $\boldsymbol{z}_p^{(k)} \in \mathcal{D}_p^{(k)}$ for all $p \in \mathcal{N}_k$ and (3.32) follows from (3.33).

Finally, it is easily seen that the update (3.30) is performed in such a way that (3.33) holds for $k-1$.

□

Lemma 3.8 *The local obstacles $\underline{\psi}_{l(p,k)}^i = (\underline{\psi}_p^{(k)})^i$, $\overline{\psi}_{l(p,k)}^i = (\overline{\psi}_p^{(k)})^i$, $k = J-1,\dots,0$, as obtained from (3.28) and (3.30) are quasioptimal in the sense of (3.19).*

Proof The local obstacles as obtained from (3.28) and (3.30) depend continuously on the smoothed iterate $\bar{\boldsymbol{u}}_J^\nu$ and on the coarse grid corrections $\boldsymbol{v}_p^{(k)}$. As $\bar{\boldsymbol{u}}_J^\nu \to \boldsymbol{u}_J$ and $\boldsymbol{v}_p^{(k)} \to \boldsymbol{0}$, it is sufficient to show that (3.19) holds if $\bar{\boldsymbol{u}}_J^\nu$ and $\boldsymbol{v}_p^{(k-1)}$ are replaced by $\boldsymbol{u}_J$ and $\boldsymbol{0}$, respectively. This can be done by induction.

□

In case of constant normal directions, i.e., $\boldsymbol{e}^i(p) = \boldsymbol{e}^i(q)$ for all $p, q \in \mathcal{N}^{(J)}$, definition (3.28) reduces to the restriction

$$\left(\overline{\psi}_p^{(k-1)}\right)^1 = \min\left\{\left(\overline{\psi}_q^{(k)}\right)^1 \;\middle|\; q \in (\text{int supp}\ \lambda_p^{(k-1)} \cap \mathcal{N}_k) \setminus \mathcal{N}^{(J)^\bullet}(\bar{\boldsymbol{u}}_J^\nu)\right\}$$

as proposed by Mandel [Man84] in the scalar case (see also [Kor94]). No tangential constraints occur on coarse grids. However, in case of spatially varying normal directions monotone restriction (3.28) causes tangential constraints on coarse levels though no such constraints are present on the finest grid. This leads to more pessimistic coarse grid constraints in comparison with the scalar case. Improvements of (3.28) are possible by generalizing ideas from [Kor94].

Now we are ready to state the main result of this chapter.

Theorem 3.9 *The truncated monotone multigrid method based on local spaces $\boldsymbol{V}_p^{(k)}$ from (3.24) and on local obstacles $\underline{\psi}_p^{(k)}$, $\overline{\psi}_p^{(k)}$ as obtained from (3.28) and (3.30) is globally convergent. If the discrete problem (2.28) satisfies the non-degeneracy condition (3.14), then there is a $\nu_0 \geq 0$ such that the iteration reduces to the linear subspace correction method for the linear reduced problem (3.13) induced by the splitting*

$$\boldsymbol{S}^{(J)^\circ} = \sum_{k=0}^{J} \sum_{p \in \mathcal{N}_k} \left(\boldsymbol{V}_p^{(k)}\right)^\circ \tag{3.34}$$

with $\left(\boldsymbol{V}_p^{(J)}\right)^\circ = \boldsymbol{V}_p^{(J)} \cap \boldsymbol{S}^{(J)^\circ}$, $p \in \mathcal{N}^{(J)}$, and $\left(\boldsymbol{V}_p^{(k)}\right)^\circ = \boldsymbol{V}_p^{(k)}(\mathcal{N}^{(J)^\bullet}(\boldsymbol{u}_J))$, $p \in \mathcal{N}_k$, $k = J-1, \ldots, 0$.

Proof Utilizing (3.31) and Lemma 3.7, global convergence follows from Theorem 3.3. Asymptotic reduction to a linear iteration follows from Theorem 3.5 in combination with Lemma 3.7 and Lemma 3.8.

□

Note, that splitting (3.34) depends only on the choice of additional coarse grid spaces and not on the choice of quasioptimal restriction. In the light of Theorem 3.9, linear multigrid convergence theory can be applied in order to derive asymptotic convergence rates.

3.4 Algebraic Formulation

In this section, we shall derive an algebraic reformulation of the truncated monotone multigrid method considered in Theorem 3.9. Although the construction of the extended splitting (3.1) looks rather technical, implementation is straightforward, if Galerkin assembling of coarse grid matrices is used. Considering the truncated monotone multigrid method in it's algebraic formulation shows where the nonlinearity enters.

Let us recall, that the resulting algorithm can be implemented as a standard linear multigrid $\mathcal{V}$-cycle and is of optimal complexity per step. Denoting

$$\boldsymbol{a}_{pq} = \left(a(\lambda_p^{(J)}\boldsymbol{e}^i, \lambda_q^{(J)}\boldsymbol{e}^j)\right)_{i,j=1,\ldots,d}, \qquad \boldsymbol{b}_p = \left(f(\lambda_p^{(J)}\boldsymbol{e}^i)\right)_{i=1,\ldots,d},$$

we define the stiffness matrix and the right hand side by

$$\boldsymbol{A} = (\boldsymbol{a}_{pq})_{p,q \in \mathcal{N}^{(J)}}, \qquad \boldsymbol{b} = (b_p)_{p \in \mathcal{N}^{(J)}}.$$

The vector representation of the given iterate $\boldsymbol{u}_J^\nu$ is

$$\boldsymbol{u}_J^\nu = (\boldsymbol{u}_p)_{p\in\mathcal{N}^{(J)}}\,, \qquad \boldsymbol{u}_p = (u_p^i)_{i=1,\dots,d}\,, \qquad u_p^i = u_J^\nu(p)\cdot\boldsymbol{e}^i(p). \tag{3.35}$$

We shall use a similar partitioning of vectors $\boldsymbol{v} = (v_p^i)_{p\in\mathcal{N}_k, i=1,\dots,d} \in \mathbb{R}^{dn_k}$ on each level $0 \le k \le J$. The residual is given by

$$\boldsymbol{r} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{u}_J^\nu.$$

Solving the defect problem

$$\boldsymbol{v}^* \in \boldsymbol{\mathcal{D}}: \qquad \tfrac{1}{2}\boldsymbol{v}^*\cdot\boldsymbol{A}\boldsymbol{v}^* - \boldsymbol{r}\cdot\boldsymbol{v}^* \le \tfrac{1}{2}\boldsymbol{v}\cdot\boldsymbol{A}\boldsymbol{v} - \boldsymbol{r}\cdot\boldsymbol{v} \qquad \boldsymbol{v}\in\boldsymbol{\mathcal{D}} \tag{3.36}$$

with constraints

$$\boldsymbol{\mathcal{D}} = \{\boldsymbol{v}\in\mathbb{R}^{dn_J} \mid v_p^1 \le g(p) - u_p^1 \text{ for all } p\in\mathcal{N}^{(J)}\cap\Gamma_S\}$$

exactly, we would obtain the exact solution $\boldsymbol{u}_J = \boldsymbol{u}_J^\nu + \boldsymbol{v}^*$. The approximate correction as obtained by one step of the projected block Gauß–Seidel relaxation with $d\times d$ blocks $\boldsymbol{a}_{pq}$ is denoted by $\mathrm{GS}_J(\boldsymbol{A},\boldsymbol{r},\boldsymbol{\mathcal{D}})$. Hence, the vector representation of the smoothed iterate $\bar{\boldsymbol{u}}_J^\nu$ is given by

$$\bar{\boldsymbol{u}}_J^\nu = \boldsymbol{u}_J^\nu + \mathrm{GS}_J(\boldsymbol{A},\boldsymbol{r},\boldsymbol{\mathcal{D}}).$$

Here, for given matrix $\boldsymbol{A}^{(k)} = (\boldsymbol{a}_{pq}^{(k)})_{p,q\in\mathcal{N}_k}$, residual $\boldsymbol{r}^{(k)}$ and constraints $\boldsymbol{\mathcal{D}}^{(k)}$, the approximate correction resulting from one step of projected block Gauß–Seidel relaxation with $d\times d$ blocks $\boldsymbol{a}_{pq}^{(k)}$ is denoted by $\mathrm{GS}_k(\boldsymbol{A}^{(k)},\boldsymbol{r}^{(k)},\boldsymbol{\mathcal{D}}^{(k)})$.

Now, we describe the coarse grid correction of $\bar{\boldsymbol{u}}_J^\nu$. It is clear how to obtain the actual coincidence set $\mathcal{N}^{(J)\bullet}(\bar{\boldsymbol{u}}_J^\nu)$. We define the truncated stiffness matrix $\boldsymbol{A}^{(J)} = \mathrm{trc}(\boldsymbol{A})$ by setting those rows and columns of $\boldsymbol{A}$ to zero that are associated with basis functions $\lambda_p\boldsymbol{e}^1(p)$, $p\in\mathcal{N}^{(J)\bullet}(\bar{\boldsymbol{u}}_J^\nu)$. In practical implementation, this is realized by appropriate flags, see Chapter 4. Using the partitioning $\boldsymbol{v} = (v_p^i)_{p\in\mathcal{N}_k,i=1,\dots,d}$ of some vector $\boldsymbol{v}\in\mathbb{R}^{dn_k}$ on some level k, the vector $\mathrm{trc}(\boldsymbol{v})$ is obtained by annihilating all v_p^1 with $p\in\mathcal{N}^{(J)\bullet}(\bar{\boldsymbol{u}}_J^\nu)\cap\mathcal{N}_k$. The truncated residual is given by $\boldsymbol{r}^J = \mathrm{trc}(\boldsymbol{b} - \boldsymbol{A}\bar{\boldsymbol{u}}_J^\nu)$.

Recursive definition (3.23) of $\boldsymbol{\mu}_p^{(k)}$ gives rise to the restriction matrix $\boldsymbol{R}_k^{k-1}$,

$$\boldsymbol{R}_k^{k-1} = (\lambda_p^{(k)}(q)\boldsymbol{e}_{pq})_{p\in\mathcal{N}_{k-1},q\in\mathcal{N}_k} \qquad \boldsymbol{e}_{pq} = (\boldsymbol{e}^i(p)\cdot\boldsymbol{e}^j(q))_{i,j=1,\dots,d}.$$

As it is standard, we define our prolongation by

$$\boldsymbol{P}_{k-1}^k = \big(\boldsymbol{R}_k^{k-1}\big)^{\mathrm{T}}.$$

Local obstacles

$$\begin{aligned} \underline{\boldsymbol{\psi}}^{(k)} &= \big(\underline{\boldsymbol{\psi}}_p^{(k)}\big)_{p\in\mathcal{N}_k}, \quad \underline{\boldsymbol{\psi}}_p^{(k)} = \big((\underline{\psi}_p^{(k)})^i\big)_{i=1,\dots,d}, \\ \overline{\boldsymbol{\psi}}^{(k)} &= \big(\overline{\boldsymbol{\psi}}_p^{(k)}\big)_{p\in\mathcal{N}_k}, \quad \overline{\boldsymbol{\psi}}_p^{(k)} = \big((\overline{\psi}_p^{(k)})^i\big)_{i=1,\dots,d} \end{aligned}$$

are initialized according to (3.27) and the monotone restriction

$$\underline{\psi}^{(k-1)} = \underline{\mathcal{R}}_k^{k-1}(\underline{\psi}^{(k)}), \qquad \overline{\psi}^{(k-1)} = \overline{\mathcal{R}}_k^{k-1}(\overline{\psi}^{(k)}),$$

is defined according to (3.28). Local obstacles $\underline{\psi}^{(k)}$, $\overline{\psi}^{(k)}$ give rise to the constraints

$$\mathcal{D}^{(k)} = \{\boldsymbol{v} \in \mathbb{R}^{dn_k} \mid \underline{\psi}^{(k)} \leq \operatorname{trc}(\boldsymbol{v}) \leq \overline{\psi}^{(k)}\}.$$

Now we are ready to rewrite our algorithm as a multigrid $\mathcal{V}$-cycle.

Algorithm 1 (Monotone Multigrid Method for Signorini's Problem)

given: $\boldsymbol{u}_J^\nu \in \mathbb{R}^{n_k}$
compute: $\bar{\boldsymbol{u}}_J^\nu = \boldsymbol{u}_J^\nu + \mathrm{GS}_J(\boldsymbol{A}, \boldsymbol{r}, \mathcal{D})$ (fine grid smoothing)
initialize: $\boldsymbol{A}^{(J)} = \operatorname{trc}(\boldsymbol{A})$ $\boldsymbol{r}^{(J)} = \operatorname{trc}(\boldsymbol{b} - \boldsymbol{A}\bar{\boldsymbol{u}}_J^\nu)$ (truncate stiffness matrix and residual)
$\underline{\psi}^{(J)}$, $\overline{\psi}^{(J)}$ according to (3.27) (local obstacles)
for $k = J-1, \ldots, 1$ do
$\boldsymbol{v}^{(k)} = \mathrm{GS}_k(\boldsymbol{A}^{(k)}, \boldsymbol{r}^{(k)}, \mathcal{D}^{(k)})$ (projected block Gauß–Seidel smoothing)
$\boldsymbol{r}^{(k)} = \boldsymbol{r}^{(k)} - \boldsymbol{A}^{(k)}\boldsymbol{v}^{(k)}$ (update of residual)
$\underline{\psi}^{(k)} = \underline{\psi}^{(k)} - \boldsymbol{v}^{(k)}$ $\overline{\psi}^{(k)} = \overline{\psi}^{(k)} - \boldsymbol{v}^{(k)}$ (update of local obstacles cf. (3.30))
$\boldsymbol{A}^{(k-1)} = \boldsymbol{R}_k^{k-1}\boldsymbol{A}^{(k)}\boldsymbol{P}_{k-1}^k$ (Galerkin restriction of stiffness matrix)
$\boldsymbol{r}^{(k-1)} = \boldsymbol{R}_k^{k-1}\boldsymbol{r}^{(k)}$ (restriction of residual)
$\underline{\psi}^{(k-1)} = \underline{\mathcal{R}}_k^{k-1}(\underline{\psi}^{(k)})$, $\overline{\psi}^{(k-1)} = \overline{\mathcal{R}}_k^{k-1}(\overline{\psi}^{(k)})$ (monotone restriction of local obstacles cf. (3.28))

$\boldsymbol{v}^{(0)} = \mathrm{GS}_0(\boldsymbol{A}^{(0)}, \boldsymbol{r}^{(0)}, \mathcal{D}^{(0)})$ (approx. solution on $\mathcal{T}_0$)

for $k = 1, \ldots, J-1$ do
$\boldsymbol{v}^{(k)} = \boldsymbol{v}^{(k)} + \boldsymbol{P}_{k-1}^k\boldsymbol{v}^{(k-1)}$ (Interpolation)
new iterate: $\boldsymbol{u}_J^{\nu+1} = \bar{\boldsymbol{u}}_J^\nu + \boldsymbol{P}_{J-1}^J\boldsymbol{v}^{(J-1)}$

Our implementation of the Galerkin restriction takes advantage of the fact that the local updates of the coincidence set only cause local updates of the stiffness matrix, see Section 4.5.

Let us briefly consider some variants of the above multigrid algorithm which can be analyzed in the same general framework and which have convergence properties as stated in Theorem 3.9.

In order to further improve coarse grid transport in the transient phase, i.e., until the exact coincidence set is known, we consider a *fully truncated variant* performing truncation recursively on all levels in all directions $i = 1, \ldots, d$. More precisely, we introduce d different sets of critical nodes $\mathcal{N}_k^{\bullet,i}$, $i = 1, \ldots, d$, on each refinement level k. Starting with

$\mathcal{N}_J^{\bullet,1} = \mathcal{N}_J^{\bullet}(\bar{\boldsymbol{u}}_J^{\nu})$ and $\mathcal{N}_J^{\bullet,i} = \emptyset$, $i = 2, \ldots, d$, the update

$$\mathcal{N}_k^{\bullet,i} = (\mathcal{N}_{k+1}^{\bullet,i} \cap \mathcal{N}_k^{\bullet}) \cup \{p \in \mathcal{N}_k \mid \text{constraint in the direction of } \boldsymbol{e}^i(p) \text{ was activated when computing } \boldsymbol{v}_p^{(k)}\} \quad (3.37)$$

takes place after the correction on level k. Recursive truncation can be formulated algebraically by introducing an operator trc_k that annihilates coefficients associated with i and p, if $p \in \mathcal{N}_k^{\bullet,i}$. Then, the implementation is based on appropriate flags. Corresponding quasioptimal local obstacles are obtained by a similar modification of (3.28). We have already seen that no constraints are active in local coarse grid problems, if the exact coincidence set $\mathcal{N}_j^{\bullet}(u_j)$ has been detected and coarse grid corrections are small enough. Hence, in the non-degenerate case, fully truncated monotone multigrid still asymptotically reduces to the linear subspace correction generated by (3.34).

Multiple pre- and post-smoothing or $\mathcal{W}$-cycles are performed in the usual way. In terms of subspace corrections such algorithms can be formulated by multiple occurrence of the same coarse grid space $\boldsymbol{V}_p^{(k)}$.

Exact solution on the initial grid $\mathcal{T}_0$ corresponds to replacing the spaces $\boldsymbol{V}_p^{(0)}$, $p \in \mathcal{N}_0$, by $\boldsymbol{V}^{(0)} = \boldsymbol{S}^{(0)}$.

In case of adaptively refined grids, coarse grid smoothing is applied only at new nodes and their neighbors. Again, there is a corresponding interpretation in terms of subspaces $\boldsymbol{V}_p^{(k)}$. In the adaptive case, it may happen that the dimension of $\boldsymbol{V}_p^{(k)}$ is less than d, see also Section 4.5.

Other finite element discretizations like piecewise quadratics or bilinear elements on quadrilaterals can be treated in a similar way. We only have to plug in the appropriate nodal basis functions instead of $\boldsymbol{\lambda}_p^{(j)}$.

Remark 3.10 *The monotone multigrid method developed in this work is capable of handling* box-constraints*, i.e., constraints in normal* and *tangential direction. This turns out to be necessary on the coarser grids of the multigrid hierarchy, when constraints with respect to different normal directions have to be taken into account, see Chapter 5. In particular, even if the outer normals at the contact boundary do not differ, box constraints are required for solving contact problems involving friction, see Chapter 6. Then, the constraints in tangential direction depend on the friction functional and the actual guess of the stick and slip regions at the contact boundary.*