# 7

# Solution of a Coercive Variational Inequality by FETI–DP Method

Numerical experiments in Chap. 5 demonstrated the capability of algorithms with the rate of convergence in bounds on the spectrum to solve special classes of bound constrained problems with optimal, i.e., asymptotically linear, complexity. There is a natural question whether there are effective methods which can reduce the solution of some real-world problems to these special classes.

To give an example of such method, we present here the one which can be used to reduce the coercive variational inequality which describes the equilibrium of a system of 2D elastic bodies in mutual contact to the class of bound constrained QP problems with uniformly bounded spectrum of the Hessian matrix. Let us recall that a contact problem is called coercive if all the bodies are fixed along the part of the boundary in a way which excludes their rigid body motion. To simplify our exposition, we restricted our attention to the solution of a scalar variational inequality governed by the Laplace operator.

Our main tool is a variant of the *finite element tearing and interconnecting (FETI) method*, which was originally proposed by Farhat and Roux [86, 87] as a parallel solver for the problems described by elliptic partial differential equations. The basic idea of FETI is to decompose the domain into non-overlapping subdomains that are "glued" by equality constraints. The variant that we consider here is the FETI–DP method proposed for linear problems by Farhat et al. [83]; it assumes that the subdomains are not completely separated, but remain joined at some nodes that are called *corners* as in Fig. 7.2. After eliminating the primal variables from the KKT conditions for the minimum of the discretized energy function subject to the bound and equality constraints by solving nonsingular local problems, the original problem is reduced to a small, relatively well conditioned bound constrained quadratic programming problem in the Lagrange multipliers.

Though not discovered in this way, the FETI-based methods for linear elliptic problems can be considered as a successful application of the duality theory to the convex QP problems. Here we use the standard duality theory for coercive equality and inequality constrained problems as described in Sect. 2.6.4.

Zdeněk Dostál, *Optimal Quadratic Programming Algorithms*,
Springer Optimization and Its Applications, DOI 10.1007/978-0-387-84806-8_7,

## 7.1 Model Coercive Variational Inequality

Let $\Omega = (0,1) \times (0,1)$ denote an open domain with the boundary $\Gamma$ and its three parts $\Gamma_u = \{0\} \times [0,1]$, $\Gamma_f = [0,1] \times \{0,1\}$, and $\Gamma_c = \{1\} \times [0,1]$. The parts $\Gamma_u$, $\Gamma_f$, and $\Gamma_c$ are called respectively the Dirichlet boundary, the Neumann boundary, and the contact boundary. On the contact boundary $\Gamma_c$, let us define the obstacle $\ell$ by the upper part of the circle with the radius $R = 1$ and the center $S = (1, 0.5, -1.3)$.

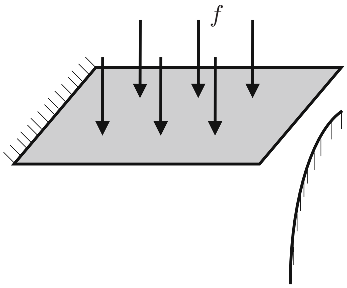


**Fig. 7.1.** Coercive model problem

Let $H^1(\Omega)$ denote the Sobolev space of the first order in the space $L^2(\Omega)$ of functions on $\Omega$ whose squares are integrable in the Lebesgue sense, let

$$\mathcal{K} = \{u \in H^1(\Omega) : \ u = 0 \ \text{ on } \ \Gamma_u \quad \text{and} \quad \ell \le u \ \text{ on } \ \Gamma_c\},$$

and let us define for any $u \in H^1(\Omega)$

$$f(u) = \frac{1}{2}\int_\Omega \|\nabla u(x)\|^2 \mathrm{d}\Omega + \int_\Omega u \mathrm{d}\Omega.$$

Thus we can define the continuous problem to find

$$\min_{u \in \mathcal{K}} f(u). \tag{7.1}$$

Since the Dirichlet conditions are prescribed on the part $\Gamma_u$ of the boundary with the positive measure, the cost function $f$ is coercive, which guarantees the existence and uniqueness of the solution by Proposition 2.5.

The solution can be interpreted as the displacement of the membrane under the traction defined by the unit density. The membrane is fixed on $\Gamma_u$, not allowed to penetrate the obstacle on $\Gamma_c$, and pulled horizontally in the direction of the outer normal by the forces with the unit density along $\Gamma_f$. See also Fig. 7.1. We used the discretized problem (7.1) as a benchmark in Sect. 5.11.1.

## 7.2 FETI–DP Domain Decomposition and Discretization

The first step in our domain decomposition method is to partition the domain $\Omega$ into $p$ square subdomains with the sides $H = 1/q$, $q > 1$, $p = q^2$. We call $H$ the *decomposition parameter*. The continuity of the global solution in $\Omega$ is enforced by the "gluing" conditions $u^i(\mathbf{X}) = u^j(\mathbf{X})$ that should be satisfied for any point $\mathbf{X}$ on the interface $\Gamma^{ij}$ of $\Omega^i$ and $\Omega^j$ except crosspoints. We call a common *crosspoint* either a corner that belongs to four subdomains, or a corner that belongs to two subdomains and is located on $\Gamma$. An important feature for developing FETI–DP type algorithms is that a single degree of freedom is considered at each crosspoint, while two degrees of freedom are introduced at all the other matching nodes across subdomain edges. Thus the body is decomposed into the subdomains that are joined in the corners as in Fig. 7.2.

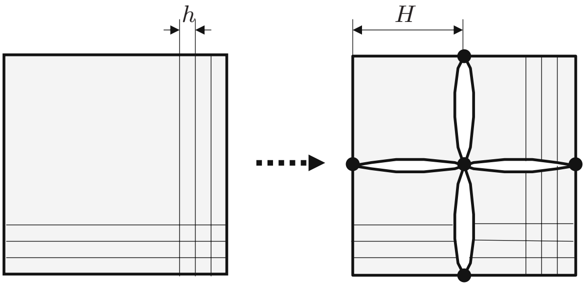


**Fig. 7.2.** FETI–DP domain decomposition and crosspoints

After modifying appropriately the definition of problem (7.1), introducing regular grids in the subdomains $\Omega^i$ with the *discretization parameter* $h$ that match across the interfaces $\Gamma^{ij}$ of $\Omega^i$ and $\Omega^j$, keeping in mind that the crosspoints are global, and using the Lagrangian finite element discretization, we get the discretized version of problem (7.1) with auxiliary domain decomposition in the form

$$\min \frac{1}{2}\mathbf{x}^T \mathsf{A}\mathbf{x} - \mathbf{b}^T\mathbf{x} \quad \text{s.t.} \quad \mathsf{B}_{\mathcal{I}*}\mathbf{x} \le \mathbf{c}_{\mathcal{I}} \quad \text{and} \quad \mathsf{B}_{\mathcal{E}*}\mathbf{x} = \mathbf{o}. \tag{7.2}$$

We assume that the nodes that are not the crosspoints are indexed contiguously in the subdomains, so that Hessian matrix $\mathsf{A} \in \mathbb{R}^{n\times n}$ in (7.2) has the form

$$\mathsf{A} = \begin{bmatrix} \mathsf{A}_{\mathrm{r}1} & \mathsf{O} & \dots & \mathsf{O} & \mathsf{A}_{\mathrm{c}1} \\ \mathsf{O} & \mathsf{A}_{\mathrm{r}2} & \dots & \mathsf{O} & \mathsf{A}_{\mathrm{c}2} \\ . & . & \dots & . & . \\ \mathsf{O} & \mathsf{O} & \dots & \mathsf{A}_{\mathrm{r}p} & \mathsf{A}_{\mathrm{c}p} \\ \mathsf{A}_{\mathrm{c}1}^T & \mathsf{A}_{\mathrm{c}2}^T & \dots & \mathsf{A}_{\mathrm{c}p}^T & \mathsf{A}_{\mathrm{cc}} \end{bmatrix}$$

with the band matrices $\mathsf{A}_{\mathrm{r}i}$. Since the diagonal blocks can be interpreted as the stiffness matrices of the subdomains that are fixed at least in corners, $\mathsf{A}_{\mathrm{r}i}$ are positive definite. We refer to the points that are not crosspoints as *reminders*; the subscripts c and r refer to the crosspoints and reminders, respectively. We assume that the Dirichlet conditions are enhanced in $\mathsf{A}$ by deleting the corresponding rows and columns. The vector $\mathbf{b} \in \mathbb{R}^n$ represents the discrete analog of the linear term $b(u)$.

The full rank submatrices $\mathsf{B}_{\mathcal{I}*}$ and $\mathsf{B}_{\mathcal{E}*}$ of a matrix $\mathsf{B} \in \mathbb{R}^{m\times n}$ describe the discretized nonpenetration and gluing conditions, respectively. The rows of $\mathsf{B}_{\mathcal{E}*}$ are filled with zeros except 1 and $-1$ in positions that correspond to the nodes with the same coordinates on the subdomain interfaces. If $\mathbf{b}_i$ denotes a row of $\mathsf{B}_{\mathcal{E}*}$, then $\mathbf{b}_i$ has just two nonzero entries, 1 and $-1$. The continuity of the solution across the interface in the nodes with indices $i, j$ (see Fig. 7.3) is enforced by the equalities

$$x_i = x_j.$$

Denoting

$$\mathbf{b}_k = (\mathbf{s}_i - \mathbf{s}_j)^T,$$

where $\mathbf{s}_i$ denotes the $i$th column of the identity matrix $\mathsf{I}_n$, we can write the "gluing" equalities conveniently in the form

$$\mathbf{b}_k \mathbf{x} = 0,$$

so that $\mathbf{b}_k\mathbf{x}$ denotes the jump across the boundary. The nonpenetration condition $x_i \geq \ell_i$ that should be satisfied for the variables corresponding to the nodes on $\Gamma_c$, is implemented by $\mathbf{b}_i\mathbf{x} \leq -\ell_i$ with $\mathbf{b}_i = -\mathbf{s}_i^T$. The coordinates $-\ell_i$ are assembled into the vector $\mathbf{c}_{\mathcal{I}}$.

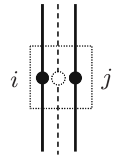


**Fig. 7.3.** "Gluing" along subdomain interface

Our next step is to reduce the problem to the subdomain interfaces and $\Gamma_c$ by the duality theory. To this end, let us denote the Lagrange multipliers associated with the inequality and equality constraints of problem (7.2) by $\lambda_{\mathcal{E}}$ and $\lambda_{\mathcal{I}}$, respectively, and assume that the rows of $\mathsf{B}$ are ordered in such a way that

$$\boldsymbol{\lambda} = \begin{bmatrix} \boldsymbol{\lambda}_{\mathcal{I}} \\ \boldsymbol{\lambda}_{\mathcal{E}} \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} \mathbf{c}_{\mathcal{I}} \\ \mathbf{o}_{\mathcal{E}} \end{bmatrix}, \quad \text{and} \quad \mathsf{B} = \begin{bmatrix} \mathsf{B}_{\mathcal{I}} \\ \mathsf{B}_{\mathcal{E}} \end{bmatrix}.$$

Since we formed $\mathsf{B}$ in such a way that it is a full rank matrix with orthogonal rows, we can use Proposition 2.21 to get that the Lagrange multipliers $\boldsymbol{\lambda}$ for problem (7.1) solve the dual problem

$$\max\ \Theta(\boldsymbol{\lambda}) \quad \text{s.t.} \quad \boldsymbol{\lambda}_{\mathcal{I}} \geq \mathbf{o},$$

where $\Theta(\boldsymbol{\lambda})$ is the dual function. Changing the signs of $\Theta$ and discarding the constant term, we get that the Lagrange multipliers $\boldsymbol{\lambda}$ solve the bound constrained problem

$$\min\ \theta(\boldsymbol{\lambda}) \quad \text{s.t.} \quad \boldsymbol{\lambda}_{\mathcal{I}} \geq \mathbf{o}, \tag{7.3}$$

where $\theta$ and the standard FETI notation are defined by

$$\theta(\boldsymbol{\lambda}) = \frac{1}{2}\boldsymbol{\lambda}^T\mathsf{F}\boldsymbol{\lambda} - \boldsymbol{\lambda}^T\mathbf{d}, \quad \mathsf{F} = \mathsf{B}\mathsf{A}^{-1}\mathsf{B}^T, \quad \mathbf{d} = \boldsymbol{\lambda}^T\mathsf{B}\mathsf{A}^{-1}\mathbf{b} - \mathbf{c}.$$

Notice that using the block and band structure of $\mathsf{A}$, we can effectively evaluate $\mathsf{A}^{-1}\mathbf{y}$ for any $\mathbf{y} \in \mathbb{R}^n$ in two steps. Indeed, using the Cholesky decomposition described in Sect. 1.5, we can eliminate the reminders, reducing the unknowns to the corners. In the next step, we decompose the small Schur complement matrix which is associated with the crosspoint variables. However, the implementation of this procedure is a bit tricky and not directly related to the quadratic programming, the main topic of this book. We refer interested readers to Dostál, Horák, and Stefanica [70] or to the Ph.D. thesis of Horák [121]. If the dimension of the blocks $\mathsf{A}_{\mathrm{r}i}$ is uniformly bounded, then the computational cost increases nearly proportionally with $p$. Moreover, the time that is necessary for the decomposition $\mathsf{A} = \mathsf{L}\mathsf{L}^T$ and evaluation of $(\mathsf{L}^{-1})^T\mathsf{L}^{-1}\mathbf{y}$ can be reduced nearly proportionally by parallel implementation.

The preconditioning effect of the FETI–DP duality transformation is formulated in the following proposition.

**Proposition 7.1.** *Let $\mathsf{F}_{H,h}$ denote the Hessian of the reduced dual function $\theta$ of* (7.3) *defined by the decomposition parameter $H$ and the discretization parameter $h$.*

*Then there are constants $C_1 > 0$ and $C_2 > 0$ independent of $h$ and $H$ such that*

$$C_1 \;\leq\; \lambda_{\min}(\mathsf{F}_{H,h}) \quad \textit{and} \quad \lambda_{\max}(\mathsf{F}_{H,h}) = \|\mathsf{F}_{H,h}\| \;\leq\; C_2\left(\frac{H}{h}\right)^2. \tag{7.4}$$

*Proof.* See [70]. □

Proposition 7.1 shows that the FETI–DP procedure reduces the conditioning of the Hessian of discretized energy from $O(h^{-2})$ to $O(H^2/h^2)$.

## 7.3 Optimality

To show that Algorithm 5.8 is optimal for the solution of problem (or a class of problems) (7.3), let us introduce new notation that complies with that used to define the class of problems (5.117) introduced in Sect. 5.8.4.

We use

$$\mathcal{T} = \{(H,h) \in \mathbb{R}^2 : H \leq 1,\ 0 < 2h \leq H, \quad \text{and} \quad H/h \in \mathbb{N}\}$$

as the set of indices, where $\mathbb{N}$ denotes the set of all positive integers. Given a constant $C \geq 2$, we define a subset $\mathcal{T}_C$ of $\mathcal{T}$ by

$$\mathcal{T}_C = \{(H,h) \in \mathcal{T} :\ H/h \leq C\}.$$

For any $t \in \mathcal{T}$, we define

$$\mathsf{A}_t = \mathsf{F}, \quad \mathbf{b}_t = \mathbf{d}, \quad \boldsymbol{\ell}_{t,\mathcal{I}} = \mathbf{o}_{\mathcal{I}}, \quad \text{and} \quad \boldsymbol{\ell}_{t,\mathcal{E}} = -\infty$$

by the vectors and matrices generated with the discretization and decomposition parameters $H$ and $h$, respectively, so problem (7.3) with the fixed discretization and decomposition parameters $h$ and $H$ is equivalent to the problem

$$\text{minimize} \ \ f_t(\boldsymbol{\lambda}_t) \ \ \text{s.t.} \ \ \boldsymbol{\lambda}_t \geq \boldsymbol{\ell}_t \tag{7.5}$$

with $t = (H,h)$, $f_t(\boldsymbol{\lambda}) = \frac{1}{2}\boldsymbol{\lambda}^T \mathbf{A}_t \boldsymbol{\lambda} - \mathbf{b}_t^T \boldsymbol{\lambda}$. Using these definitions, we obtain

$$\|\boldsymbol{\ell}_t^+\| = 0, \tag{7.6}$$

where for any vector $\mathbf{v} = [v_i]$, $\mathbf{v}^+$ denotes the vector with the entries $v_i^+ = \max\{v_i, 0\}$. Moreover, it follows by Proposition 7.1 that for any $C \geq 2$, there are the constants $a_{\max}^C > a_{\min}^C > 0$ such that for any $t \in \mathcal{T}_C$

$$a_{\min}^C \leq \lambda_{\min}(\mathsf{A}_t) \leq \lambda_{\max}(\mathsf{A}_t) \leq a_{\max}^C, \tag{7.7}$$

where $\lambda_{\min}(\mathsf{A}_t)$ and $\lambda_{\max}(\mathsf{A}_t)$ denote the extreme eigenvalues of $\mathsf{A}_t$.

Our optimality result for a model coercive boundary variational inequality then reads as follows.

**Theorem 7.2.** *Let $C \geq 2$ and $\varepsilon > 0$ denote given constants, let $\{\boldsymbol{\lambda}_t^k\}$ be generated by Algorithm 5.8 (MPRGP) for the solution of (7.5) with the parameters $\Gamma > 0$ and $\overline{\alpha} \in (0, a_{\max}^{-1}]$, starting from $\boldsymbol{\lambda}_t^0 = \max\{\mathbf{o}, \boldsymbol{\ell}_t\}$.*

*Then an approximate solution $\lambda_t^{k_t}$ of any problem (7.5) which satisfies*

$$\|\mathbf{g}_t^P(\boldsymbol{\lambda}^{k_t})\| \leq \varepsilon \|\mathbf{g}_t^P(\boldsymbol{\lambda}_t^0)\|$$

*and*

$$a_{\min}^C \|\boldsymbol{\lambda}^{k_t} - \widehat{\boldsymbol{\lambda}}_t\| \leq f_t(\boldsymbol{\lambda}_t^{\overline{\ell}}) - f_t(\widehat{\boldsymbol{\lambda}}_t) \leq \varepsilon \left( f_t(\boldsymbol{\lambda}_t^0) - f(\widehat{\boldsymbol{\lambda}}_t) \right)$$

*is generated at $O(1)$ matrix–vector multiplications by $\mathsf{A}_t$ for any $t \in \mathcal{T}_C$.*

*Proof.* The class of problems (7.5) with $t \in \mathcal{T}_C$ satisfies the assumptions of Theorem 5.16. $\square$

## 7.4 Numerical Experiments

In this section we illustrate numerical scalability of MPRGP Algorithm 5.8 on the class of problems arising from application of the FETI–DP method to our boundary variational inequality (7.1). The domain $\Omega$ was partitioned into identical squares with the side $H \in \{1/2, 1/4, 1/8\}$. The squares were then discretized by the regular grid with the stepsize $h$. The solution for $H = 1/4$ and $h = 1/4$ is in Fig. 7.4.

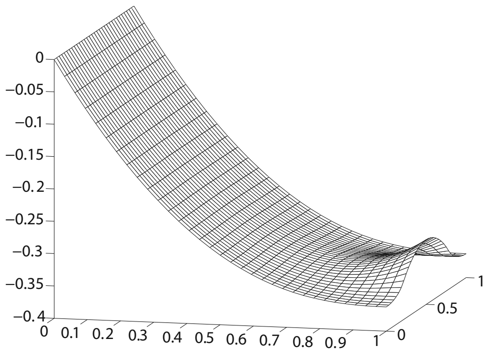


**Fig. 7.4.** Solution of the coercive model problem (7.1)

The computations were performed with parameters $\Gamma = 1$, $\overline{\alpha} \approx 1/\|\mathsf{A}\|$, and $\boldsymbol{\lambda}^0 = \mathbf{o}$. The stopping criterion in the conjugate gradient iteration was

$$\|\mathbf{g}^P(\boldsymbol{\lambda}^k)\|/\|\mathbf{g}^P(\boldsymbol{\lambda}^0)\| < 10^{-6}.$$

For each $H$, we chose $h = H/16$, so that ratio $H/h$ was fixed to $H/h = 16$ and the meshes matched across the interface of each couple of neighboring subdomains. Selected results of the computations for varying values of $H \in \{1/8, 1/32, 1/64\}$ and $h = H/16$ are in Fig. 7.5. The primal dimension $n$ is on the horizontal axis; the computation was carried out for primal dimension $n \in \{1156, 4624, 18496\}$ with corresponding dual dimensions $m \in \{93, 425, 1809\}$. The key point is that the number of the conjugate gradient iterations for a fixed ratio $H/h$ varies very moderately with the increasing number of subdomains. This indicates that the unspecified constants in Theorem 7.2 are not very large and we can observe numerical scalability in practical computations. For more numerical experiments with the solution of coercive problems see Dostál, Horák, and Stefanica [70].

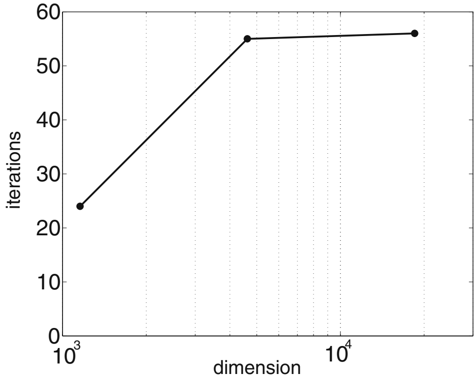


**Fig. 7.5.** Scalability of MPRGP with FETI–DP

## 7.5 Comments and References

More problems described by variational inequalities can be found in the book by Lions and Duvaut [143]. Solvability, approximation, and classical numerical methods for variational inequalities or contact problems are discussed in the books by Glowinski [99], Kinderlehrer and Stampaccia [128], Glowinski, Lions, and Trèmoliéres [101], Hlaváček et al. [120], or Eck, Jarušek, and Krbec [80]. The formulation and alternative algorithms for the solution contact problems of elasticity are in Kikuchi and Oden [127], Laursen [142], or Wriggers [181].

Probably the first theoretical results concerning development of scalable algorithms for coercive problems were proved by Schöberl [165, 166]. Our first proof of numerical scalability of an algorithm for the solution of a coercive variational inequality used optimal penalty in dual FETI problem [66]. The proof of Proposition 7.1 is due to D. Stefanica [70]. The optimality was proved also for multidomain coercive problems [70] and for the FETI–DP solution of coercive problems with nonpenetration mortar conditions on contact interface [71]. For more details of mortar implementation of constraints we refer to Wohlmuth [179]. Numerical evidence of scalability of a different approach combining FETI–DP with a Newton-type algorithm for 3D contact problems was given in Avery et al. [2]. See also Dostál et al. [76]. The performance of the method can be further improved by enforcing zero averages of primal variables on the interfaces of subdomains as used by Klawonn and Rheinbach [129] or by preconditioning of linear auxiliary problems by standard preconditioners described, e.g., in Tosseli and Widlund [175]. Preconditioning of linear step was successfully applied by Avery et al. [2].

It should be noted that the effort to develop scalable solvers for coercive variational inequalities was not restricted to FETI. For example, using the ideas related to Mandel [147], Kornhuber [132], Kornhuber and Krause [133], and Krause and Wohlmuth [135] gave an experimental evidence of numerical scalability and the convergence theory for the algorithm based on monotone multigrid. Badea, Tai, and Wang [6] proved linear rate of convergence in terms of the decomposition parameter and overlap for the Schwarz domain decomposition method which assumes exact solution of subdomain problems. See also Zeng and Zhou [185], Tai and Tseng [173], Tarvainen [174], and references therein.

A readable introduction into the formulation and implementation of the FETI methods, including FETI–DP, can be found in Kruis [136]. Let us stress that our goal here is only to illustrate the optimality of MPRGP Algorithm 5.8 on the problem whose structure is the same as that of important real-world problems.