A Numerical Approach to the Design of Masonry Structures *

Michal Kočvara and Jiří V. Outrata **

1 Introduction

Assume that U and Y are Banach spaces, $\Gamma[U \rightsquigarrow Y^*]$ is a proper convex- and closed-valued multifunction and $F[U \times Y \to Y^*]$ is a continuously differentiable operator. Consider a class of optimum design problems

$$\begin{array}{ll} & f(u,y) \to \inf \\ \text{subject to} & \\ & 0 \in F(u,y) + N_{\Gamma(u)}(y) \\ & u \in U_{ad} \\ & y \in V, \end{array} \tag{1.1}$$

where u is the *control* variable, y is the *state* variable, $f[U \times Y \to \mathbb{R}]$ is a continuously differentiable objective function, and $U_{ad} \subset U$, $V \subset Y$ are the sets of admissible controls and states, respectively. $N_{\Gamma(u)}(y)$ is the normal cone to $\Gamma(u)$ at y provided $y \in \Gamma(u)$ and it is the empty set otherwise. In various shape optimization problems in mechanics, Γ is given in the form

$$\Gamma(u) = \{y \in Y \mid A(u)y = b(u), -q(y) \in K\}, \tag{1.2}$$

where the map A assigns to controls u linear continuous operators, K is the cone of nonnegative functions in a *constraint* space Z and the maps A, b and q are continuously differentiable. The equation in (1.2) expresses the so-called equilibirum condition and the "inequality" $-q(y) \in K$ some physical requirement (an obstacle, plastic yielding, ...).

Assume that V consists of the states y for which $q(y) \in \operatorname{int} K$ in a prescribed subdomain of the designed domain. This special requirement can be respected by an exact penalty dependent on the Karush–Kuhn–Tucker (K.K.T.) vector λ assigned to the inequality $-q(y) \in K$. After an appropriate discretization we arrive at a

*Partly supported by the German Scientific Foundation (DFG).

**Institute of Information Theory and Automation, Czech Academy of Sciences, 18208 Praha 8, Pod vodárenskou věží 4, Czech Republic

finite-dimensional optimization problem

$$
\begin{aligned}
& \phi(u,y,\lambda) \to \inf \\
\text{subject to} \quad & \\
& 0 \in \begin{pmatrix} F(u,y) + A^T(u)\mu + (\nabla q(y))^T\lambda \\ A(u)y - b(u) \\ -q(y) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ N_{\mathbb{R}^s_+}(\lambda) \end{pmatrix} \qquad (1.3) \\
& u \in U_{ad},
\end{aligned}
$$

in which U is replaced by $\mathbb{R}^n$, Y by $\mathbb{R}^m$, Z by $\mathbb{R}^s$ and the range space of $A(u)$ by $\mathbb{R}^p$; for the sake of simplicity we use the same symbols for the original and the discretized sets, maps and variables. $\varphi[\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^s \to \mathbb{R}]$ is the discretized "augmented" objective, respecting the state constraints of the above type and $\mu \in \mathbb{R}^p$ is the multiplier vector assigned to the (discretized) equality constraint $A(u)y = b(u)$.

Main difficulties in solving (1.3) are due to its first constraint — so called generalized equation ([11]). However, on the basis of the results from [11, 7, 12], necessary optimality conditions for rather similar problems have already been derived in [10] and [9] and in [10] also a numerical approach has been proposed under the restriction that $\nabla_y F(u,y)$ is symmetric over $U \times Y$ and φ does not depend on λ. In fact, this numerical approach has already been applied to shape optimization problems of a simpler structure in [5, 6] and it represents a useful alternative to the classical regularization techniques, cf. [4]. The aim of this contribution is

(i) to extend the numerical approach of [5, 10, 6] to problems of the type (1.3), and

(ii) to apply it to the design of masonry-like structures analyzed in [2].

For the reader's convenience we recall still the definition of the generalized gradient and generalized Jacobian used throughout the paper.

Definition 1.1 Let the operator $F[\mathsf{R}^m \to \mathsf{R}^n]$ be Lipschitz near $x_0 \in \mathsf{R}^m$ and let σ_F denote the set of points at which F fails to be differentiable. The *generalized Jacobian* of F at x_0, denoted $\partial F(x_0)$, is the set of $[n \times m]$ matrices, given by

$$\partial F(x_0) = \ \text{conv} \left\{ \lim_{i\to\infty} \nabla F(x_i) \,|\, x_i \to x_0,\ x_i \notin \sigma_F \right\}.$$

For $n = 1$, $\partial F(x_0)$ is termed the *generalized gradient* of F at x_0. Its elements are called *subgradients.*

Remark 1.1 The term subgradient is borrowed from convex analysis, where it is used for the vectors belonging to the subdifferential. Recently, however, it is frequently used also in a nonconvex setting for the elements of various generalized gradients.

The following notation is employed: $\mathcal{N}(C)$ is the null space of a linear operator C, $S'(x_0; d)$ is the directional derivative of an operator S at a point x_0 in the direction d, E is the unit matrix, P_1 denotes the subspace of the polynomials of order one and x^i is the $i-$th coordinate of a vector $x \in \mathbb{R}^m$. For $x, y \in \mathbb{R}^n$ the inequalities $x \geq y$ $(x > y)$ mean $x^i \geq y^i$ $(x^i > y^i)$ for all i. H^1, H^1_0 are the usual Sobolev spaces $W^{1,2}, W^{1,2}_0$ and $\overline{A}$ denotes the closure of a set A.

2 Numerical approach

Suppose that φ is continuously differentiable, U_{ad} is a nonempty compact subset of $\mathbb{R}^n$ and for any u from an open set $\mathcal{A} \supset U_{ad}$, the following assumptions hold:

(A1) $\Gamma(u) \neq \emptyset$;

(A2) $F(u, \cdot)$ is strongly monotone over $\mathbb{R}^m$, uniformly with respect to $u \in \mathcal{A}$, i.e., there exists a constant $\gamma > 0$ such that

$$\langle F(u, y_1) - F(u, y_2), y_1 - y_2 \rangle \geq \gamma \|y_1 - y_2\|^2 \quad \text{for all } y_1, y_2 \in \mathbb{R}^m, u \in \mathcal{A}.$$

Under (A1), (A2), the generalized equation in (1.3) defines an operator S which assignes to each $u \in \mathcal{A}$ a unique vector $y \in \mathbb{R}^m$ such that with a suitable pair $(\mu, \lambda) \in \mathbb{R}^p \times \mathbb{R}^s_+$, (y, μ, λ) is a solution of this generalized equation. To ensure the unicity of the multipliers μ, λ as well, we impose still the well-known linear independency constraint qualification, namely

(LI) for each $u \in \mathcal{A}$, $y = S(u)$, the vectors $A^i(u), i = 1, 2, \dots, p$, and $\nabla q^j(y)$, $j \in I(u) := \{i \in \{1, 2, \dots, s\} \mid q^i(y) = 0\}$, are linearly independent.

Then, as shown e.g. in [11], the operator S is locally Lipschitz over $\mathcal{A}$ and also the maps Π, Λ assigning to $u \in \mathcal{A}$ the respective multipliers μ and λ are single-valued and locally Lipschitz. Therefore, problem (1.3) may be rewritten to the form

$$\begin{aligned} & \Theta(u) := \varphi(u, S(u), \Lambda(u)) \to \inf \\ \text{subject to} & \\ & u \in U_{ad}. \end{aligned} \tag{2.1}$$

This problem may be numerically solved by available nondifferentiable optimization (NDO) methods, provided Θ is at least directionally differentiable and we are able to compute subgradients from $\partial\Theta(u)$ for arbitrary $u \in U_{ad}$. Fortunatelly, under (A1), (A2) and (LI), both maps S and Λ are directionally differentiable, whereby these derivatives satisfy a certain system of equations and inequalities, cf. [3, Thm 5.2]. In these relations an important role is played by the index set of *strongly active inequalities*

$$J(u) := \{i \in I(u) \mid \lambda^i > 0\}$$

which will be used also in the further development.

To simplify the notation, for an index set $M \subset \{1, 2, \dots, s\}$ and a vector $d \in \mathbb{R}^s$, d_M denotes the subvector composed from the components $d^i, i \in M$. Analogously, for a matrix D with s rows, D_K denotes a submatrix composed from the rows $D^i, i \in K$. Further, the arguments at I and J will be omitted whenever it cannot lead to any confusion.

Let us turn our attention to the computation of subgradients. For this purpose we employ the generalized Jacobian Chain Rule of [1], according to which at a fixed $u_0 \in \mathcal{A}$, $y_0 = S(u_0)$ and $\lambda_0 = \Lambda(u_0)$ one has

$$\begin{aligned} \partial\Theta(u_0) = \nabla_u\varphi(u_0, y_0, \lambda_0) \quad &+ \quad \{P^T(u_0)\nabla_y\varphi(u_0, y_0, \lambda_0) \,|\, P(u_0) \in \partial S(u_0)\} \\ &+ \quad \{Q^T(u_0)\nabla_\lambda\varphi(u_0, y_0, \lambda_0) \,|\, Q(u_0) \in \partial\Lambda(u_0)\}. \end{aligned} \tag{2.2}$$

The matrices $P(u_0), Q(u_0)$ will be computed exactly according to Definition 1.1 Assume first that at u_0 the scrict complementarity holds, i.e. $I(u_0) = J(u_0)$. Then the maps S, Π, Λ are differentiable at u_0 ([7]) and the respective derivatives solve the matrix linear equation

$$D_I(u_0, y_0, \mu_0, \lambda_0) \begin{pmatrix} \nabla S(u_0) \\ \nabla\Pi(u_0) \\ \nabla\Lambda_I(u_0) \end{pmatrix} = \begin{pmatrix} -\nabla_u\mathcal{L}(u_0, y_0, \mu_0, \lambda_0) \\ \nabla_u(A(u_0)y_0) - \nabla b(u_0) \\ 0 \end{pmatrix}, \tag{2.3}$$

where $\mu_0 = \Pi(u_0)$, $\mathcal{L}(u, y, \mu, \lambda) = F(u, y) + (A(u))^T\mu + (\nabla q(y))^T\lambda$ is the *Lagrangian* and

$$D_I(u_0, y_0, \mu_0, \lambda_0) = \begin{pmatrix} \nabla_y\mathcal{L}(u_0, y_0, \mu_0, \lambda_0) & (A(u_0))^T & (\nabla q_I(y_0))^T \\ -A(u_0) & 0 & 0 \\ -\nabla q_I(y_0) & 0 & 0 \end{pmatrix}. \tag{2.4}$$

The derivatives of Λ^i for $i \notin I(u_0)$ are zero. To analyze the general case $I(u_0) \neq J(u_0)$, we fix an index set $\widetilde{J}$ satisfying the inclusions

$$J(u_0) \subset \widetilde{J} \subset I(u_0) \tag{2.5}$$

and denote $\mathcal{D} = I(u_0) \setminus \widetilde{J}, \mathcal{B} = \widetilde{J} \setminus J(u_0)$. Let o, o_1, o_2, o_3 be the cardinalities of $I(y_0), J(y_0), \mathcal{D}$ and $\mathcal{B}$, respectively. Matrices $P(u_0) \in \partial S(u_0)$, $Q(u_0) \in \partial S(u_0)$ can be computed according to the following assertion.

Theorem 2.1 *Let assumptions (A1),(A2) and (LI) hold and consider the point* $(u_0, y_0, \mu_0, \lambda_0)$, *where* $u_0 \in \mathcal{A}$, $y_0 = S(u_0)$, $\mu_0 = \Pi(u_0)$ *and* $\lambda_0 = \Lambda(u_0)$. *Assume that* $\widetilde{J}$ *fulfills inclusions (2.5) and the linear system*

$$\begin{aligned} -(\nabla q_{\mathcal{D}}(y_0))^T y_1^* + (\nabla_y\mathcal{L}(u_0, y_0, \mu_0, \lambda_0))^T y_3^* + (A(u_0))^T y_4^* + (\nabla q_{\widetilde{J}}(y_0))^T y_5^* &= 0 \\ y_2^* + \nabla q_{\mathcal{B}}(y_0) y_3^* &= 0 \\ (\nabla_u\mathcal{L}(u_0, y_0, \mu_0, \lambda_0))^T y_3^* + (\nabla_u(A(u_0)y_0) - \nabla b(u_0))^T y_4^* &= 0 \end{aligned} \tag{2.6}$$

does not possess a solution $(y_1^*, y_2^*, y_3^*, y_4^*, y_5^*) \in \mathbb{R}^{o_2} \times \mathbb{R}^{o_3} \times \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^{o_1+o_3}$ *satisfying the conditions*

$$(y_1^*, y_2^*) \geq 0, \quad (y_1^*, y_2^*) \neq 0, \quad y_3^* \in \mathcal{N}(A(u_0)) \cap \mathcal{N}(\nabla q_J(y_0)). \tag{2.7}$$

Then one has

$$(D_{\widetilde{J}}(u_0, y_0, \mu_0, \lambda_0))^{-1} \begin{pmatrix} -\nabla_u \mathcal{L}(u_0, y_0, \mu_0, \lambda_0) \\ \nabla_u(A(u_0)y_0) - \nabla b(u_0) \\ 0 \end{pmatrix} \in \begin{pmatrix} \partial S(u_0) \\ \partial \Pi(u_0) \\ \partial \Lambda_{\widetilde{J}}(u_0) \end{pmatrix}, \tag{2.8}$$

where $D_{\widetilde{J}}(u_0, y_0, \mu_0, \lambda_0)$ *is given by (2.4) with* $I(u_0)$ *being replaced by* $\widetilde{J}$.

Proof. We first show that under the assumptions imposed there exists a direction $h \in \mathbb{R}^n$ such that

$$J(u_0 + \vartheta h) = I(u_0 + \vartheta h) = \widetilde{J} \tag{2.9}$$

for all sufficiently small positive ϑ. By using of directional derivatives of S and Λ, (2.9) may be rewritten to the form

$$\begin{array}{rcl} \nabla q_{\mathcal{D}}(y_0)S'(u_0; h) & < & 0 \\ (\Lambda^i)'(u_0; h) & > & 0 \quad \text{for all } i \in \mathcal{B}. \end{array}$$

Taking into account the characterization of $S'(u_0; h)$, $\Lambda'(u_0; h)$ from [3, Thm. 5.2], we observe that the desired direction exists, whenever the following linear system of equations and inequalities is consistent:

$$\begin{array}{rcl} \nabla_y \mathcal{L}(u_0, y_0, \mu_0, \lambda_0)S'(u_0; h) + (A(u_0))^T \Pi'(u_0; h) & + & \\ (\nabla q_{\widetilde{J}}(y_0))^T \Lambda'_{\widetilde{J}}(u_0; h) & = & -\nabla_u \mathcal{L}(u_0, y_0, \mu_0, \lambda_0)h \\ A(u_0)S'(u_0; h) & = & -\nabla_u(A(u_0)y)h + \nabla b(u_0)h \\ \nabla q_{\widetilde{J}}(y_0)S'(u_0; h) & = & 0 \\ \nabla q_{\mathcal{D}}(y_0)S'(u_0; h) & < & 0 \\ -\Lambda'_{\mathcal{B}}(u_0; h) & < & 0. \end{array}$$

However, its consistency is according to the well-known Motzkin theorem of the alternative ([8]) equivalent to the inconsistency of (2.6), (2.7), and thus relation (2.9) is fulfilled. Consequently, the maps S, Π, Λ are differentiable at the points $u_0 + \vartheta h$ for all sufficiently small $\vartheta > 0$. With respect to Definition 1.1 it remains to prove that

$$\lim_{\vartheta \searrow 0} \begin{pmatrix} \nabla S(u_0 + \vartheta h) \\ \nabla \Pi(u_0 + \vartheta h) \\ \nabla \Lambda_{\widetilde{J}}(u_0 + \vartheta h) \end{pmatrix} = (D_{\widetilde{J}}(u_0, y_0, \mu_0, \lambda_0))^{-1} \begin{pmatrix} -\nabla_u \mathcal{L}(u_0, y_0, \mu_0, \lambda_0) \\ \nabla_u(A(u_0)y_0) - \nabla b(u_0) \\ 0 \end{pmatrix},$$

which may be done easily by employing the classical implicit function argument. □

Remark 2.1 An appropriate completion of a matrix from $\partial \Lambda_{\widetilde{J}}(u_0)$ by zero elements provides us with a matrix from $\partial \Lambda(u_0)$.

Of course, the satisfaction of the above conditions can hardly be tested in the presented form. It is clear that if $n \geq o_2 + o_3$ (which is almost always satisfied in practical situations), then the inconsistency assumption holds for any choice of $\widetilde{J}$ whenever the matrix

$$\begin{pmatrix} -(\nabla q_{\mathcal{D}}(y_0))^T & 0 & (\nabla_y \mathcal{L}(u_0, y_0, \mu_0, \lambda_0))^T & (A(u_0))^T & (\nabla q_{\widetilde{J}}(y_0))^T \\ 0 & E & \nabla q_{\mathcal{B}}(y_0) & 0 & 0 \\ 0 & 0 & (\nabla_u \mathcal{L}(u_0, y_0, \mu_0, \lambda_0))^T & (\nabla_u (A(u_0)y_0) - \nabla b(u_0))^T & 0 \\ 0 & 0 & A(u_0) & 0 & 0 \\ 0 & 0 & \nabla q_J(y_0) & 0 & 0 \end{pmatrix}$$

has maximal rank. However, this test may also be rather cumbersome. Fortunately, in the solved example the proposed approach worked well with $\widetilde{J} = J(y_0)$ and so we were not forced to perform any testing during the iteration process.

On the basis of (2.2) and (2.8), the computation of subgradients of Θ can proceed according to the following statement:

Proposition 2.1 *Let all assumptions of Theorem 2.1 be fulfilled and assume that the triple* $(p_0, q_0, r_0) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^{o_1+o_3}$ *solves the linear system*

$$(D_{\widetilde{J}}(u_0, y_0, \mu_0, \lambda_0))^T \begin{pmatrix} p \\ q \\ r \end{pmatrix} + \begin{pmatrix} \nabla_y \varphi(u_0, y_0, \lambda_0) \\ 0 \\ (\nabla_\lambda \varphi(u_0, y_0, \lambda_0))_{\widetilde{J}} \end{pmatrix} = 0.$$

Then one has

$$\xi := \nabla_u \varphi(u_0, y_0, \lambda_0) + (\nabla_u \mathcal{L}(u_0, y_0, \mu_0, \lambda_0))^T p_0 - (\nabla_u (A(u_0)y_0) - \nabla b(u_0))^T q_0 \in \partial\Theta(u_0).$$

The proof is based on the standard "adjoint equation technique", known from optimal control theory. This way of computing subgradients has been applied in the design problem investigated in the next section.

3 Design of masonry structures

We shall investigate masonry-like material that behaves elastically if subject to pressure forces, but is extremely weak if we try to pull it apart. In general, rather small tension suffices to produce a fracture: the material breaks down and shows lines of fracture. However, the appearence of fracture may not be destructive; in most cases it is compatible with global equilibrium of the structure. The mathematical description of the equilibrium of such structures has been done in [2]. Our goal in this paper is to optimize the shape of such structures, i.e., to find such a shape (from a given set of shapes) that some cost functional is minimized.

Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with a Lipschitz boundary $\partial\Omega$. Let $\partial\Omega = \overline{\Gamma}_u \cup \overline{\Gamma}_g$, $\mathrm{meas}_2(\overline{\Gamma}_u \cap \overline{\Gamma}_g) = \emptyset$, Γ_u of positive measure. Denote by $\mathbb{R}^{2\times 2}_s$ the set of symmetric 2×2 matrices. We introduce the set of *stress fields*

$$S(\Omega) = \{\tau : \Omega \to \mathbb{R}^{2\times 2}_s \,|\, \tau_{ij} \in L^2(\Omega), \quad i, j = 1, 2\}$$

and the isomorphism

$$\beta : S(\Omega) \to S(\Omega)$$

defined by the generalized inverse Hooke's law and such that

$$\beta_0\|\sigma\|_\Omega \leq (\beta\sigma,\sigma)_\Omega \leq \beta_1\|\sigma\|_\Omega \quad \forall\sigma \in S(\Omega).$$

Here, $(\cdot,\cdot)_\Omega$ and $\|\cdot\|_\Omega$ denote $L^2(\Omega)$ inner product and norm, respectively, and β_0, β_1 are some positive constants.

Assume that a body force $F = (F_1, F_2) \in [L^2(\Omega)]^2$ and a surface traction $g = (g_1, g_2) \in [L^2(\Gamma_g)]^2$ are given. We introduce the space of virtual displacements

$$H = \{w \in [H^1(\Omega)]^2 \;\Big|\; w = 0 \text{ on } \Gamma_u\}$$

and define the set of *statically admissible* stress fields satisfying the *weak equilibrium conditions*

$$E(\Omega) = \{\tau \in S(\Omega) \;\Big|\; (\tau, e(w))_\Omega = \sum_{i=1}^{2}\int_\Omega F_i w_i \,\mathrm{d}x + \sum_{i=1}^{2}\int_{\Gamma_g} g_i w_i \,\mathrm{d}\Gamma \quad \forall w \in H\},$$

where

$$(e(w))_{ij} = \frac{1}{2}\left(\frac{\partial w_i}{\partial x_j} + \frac{\partial w_j}{\partial x_i}\right), \quad i, j = 1, 2,$$

is the *strain field*.

The masonry-like materials can be characterized in different ways. The most restrictive one is the requirement that the stress tensor $\tau(x), x \in \Omega$, cannot have negative eigenvalues ([2]). We consider, however, a more simple situation which can be used, e.g., in modelling of supporting pillars in cathedrals: we require that the 'vertical' stress component is nonnegative, i.e., $(\tau)_{22} \geq 0$ a.e. in Ω. In pillars, typically, the external force coming from the weight of the supported roof causes a bending which is compensated by the self-weight of the pillar. Our task is to find the lightest pillar which is still able to support the roof. We introduce the set of admissible stress fields as

$$M(\Omega) = \{\tau \in S(\Omega) \;\Big|\; (\tau)_{22} \geq 0 \text{ a.e. in } \Omega\}.$$

The problem of masonry structures can be formulated as a variational inequality (VI):

$$\begin{aligned} &\text{Find } \sigma \in E(\Omega) \cap M(\Omega) \text{ such that} \\ &(\beta\sigma, \tau - \sigma)_\Omega \geq 0 \quad \forall\tau \in E(\Omega) \cap M(\Omega). \end{aligned} \tag{3.1}$$

It can be shown that, under some additional conditions on the pressure g, problem (3.1) has a unique solution.

Now, let us introduce the set of *admissible design variables*

$$U_{ad} = \{u \in C^{(0),1}([0,1]), \underline{u} \leq u \leq \overline{u}, \left|\frac{\mathrm{d}u}{\mathrm{d}x_2}\right| \leq c_1, \left|\frac{\mathrm{d}^2u}{\mathrm{d}x_2^2}\right| \leq c_2\},$$

where $\underline{u}, \overline{u}, c_1$ and c_2 are given positive numbers. Consider a family of admissible domains $\Omega(u)$ with variable right "vertical" part of the boundary

$$\Omega(u) = \{(x_1, x_2) \;\Big|\; 0 < x_1 < u(x_2), 0 < x_2 < 1\}.$$

The goal of the optimum design problem is to minimize the volume of the structure, denoted by $\operatorname{meas}\Omega(u)$. However, to prevent the dectruction, we require that the inequality constraint in the definition of $M(\Omega)$ is strongly active only in a given subset $\Omega_0(u) \subset \Omega(u)$. In other words, denoting by λ the Lagrange multiplier corresponding to this constraint, we require $\lambda = 0$ in $\Omega(u)\backslash\Omega_0(u)$. Of course, in this way one cannot prevent some constraints in $\Omega(u) \setminus \Omega_0(u)$ to be active provided the correpsponding multipliers are zero. However, in the examples solved such phenomenon has been observed merely very closely to the boundary of $\Omega(u)$. The optimum design problem reads as

$$\begin{array}{ll} & \varphi(u,\lambda) := \operatorname{meas}\Omega(u) + r\int_{\Omega(u)\backslash\Omega_0(u)}(\lambda(u))\,\mathrm{d}x \to \inf \\ \text{subject to} & \\ & \sigma(u) \text{ solves (3.1) with } \Omega := \Omega(u) \\ & \lambda(u) \text{ is the multiplier corresponding to the inequality in } M(\Omega(u)) \\ & u \in U_{ad}, \end{array} \tag{3.2}$$

where $r > 0$ is the penalty parameter.

Now we discretize problems (3.1) and (3.2) by the finite element method. In particular, we approximate the design function u by a piece-wise linear function u_h, introduce the set

$$U_{ad}^h = \{u_h \in U_{ad} \;\Big|\; u_h \text{ piece-wise linear}\}$$

and construct the polygonal computational domain $\Omega(u_h)$ such that the positions of all "inner" nodes are derived from the positions of the principal moving nodes lying on u_h. Let N be the number of principal moving nodes and n the number of all nodes in the mesh.

To the approximation of the stress field we use piece-wise constant elements, i.e., with each element we associate a vector from $\mathbb{R}^3$ (three entries of a symmetric 2×2 matrix). We denote the global vector of approximated element stresses by $y \in \mathbb{R}^{3p}$, p being the number of elements. Further, we denote by u the vector from $\mathbb{R}^N$ of x_1-coordinates of the principal moving nodes. Again, we use the same symbols for the original and discretized sets, maps and variables. So U_{ad} is the set associated with U_{ad}^h,

$$E(u) = \{v \in \mathbb{R}^{3p} \;\Big|\; A(u)v = b(u)\}$$

and

$$M(u) = M = \{v \in \mathbb{R}^{3p} \;\Big|\; \Upsilon^i(v) \leq 0, \quad i = 1, 2, \ldots, p\};$$

here $A(u)$ is the equilibrium matrix, b the right-hand side vector and

$$\Upsilon^i(v) = v^{3i-1}.$$

Let $B(u)$ be the flexibility (symmetric, positive definite) matrix associated with the inverse Hooke's law β. The discretized state problem can be the written as follows:

$$\begin{aligned} &\text{Find } v \in E(u) \cap M \text{ such that} \\ &\langle B(u)v, w - v \rangle \geq 0 \quad \forall w \in E(u) \cap M. \end{aligned} \tag{3.3}$$

Denoting by $\Lambda(u)$ the map assigning to u the Lagrangian multipliers $\lambda^i, i = 1, \ldots, p$, corresponding to inequalities in M, we can write the discretized optimum design problem as

$$\begin{aligned} &\varphi(u, \lambda) \equiv \operatorname{meas} \Omega_h + \frac{r}{h^2} \sum_{i \in \mathcal{D}_0} \lambda^i \to \inf \\ \text{subject to} \quad & \\ &\lambda = \Lambda(u) \\ &u \in U_{ad} \end{aligned} \tag{3.4}$$

where $\mathcal{D}_0$ contains indices of elements lying in $\Omega(u_h) \setminus \Omega_0(u)$ and h is the discretization parameter. This is a problem of a type (1.3). Hence, to its numerical solution, we can use the theory from Section 2, provided that assumptions (A1), (A2) and (LI) are fulfilled. Fortunately, (A1) holds for a proper choice of U_{ad}, which has been observed numerically. (A2) follows from the fact that the controls $u \in U_{ad}$ are uniformly Lipschitz continuous, the triangulation cannot degenerate and thus the flexibility matrices $B(u)$ are uniformly positive definite. (LI) is trivially satisfied.

Example 3.1 Let $\Gamma_u = \{(x_1, x_2) \mid 0 \leq x_1 \leq \beta, x_2 = 0\}$ be the part of the boundary $\partial\Omega$ with prescribed zero displacements, and let the constant surface traction $g = (5, 0)$ is prescribed on the left vertical part of $\partial\Omega$, i.e., on the segment $\{(x_1, x_2) \mid x_1 = 0, 0 \leq x_2 \leq 1\}$. The surface traction on the rest of $\partial\Omega_0$ is prescribed as zero and the body force $F = (0, -10)$ coresponds to the self-weight. The penalty parameter is $r = 1$. The domain is discretized by triangles defined by 7×7 grid of nodes. The set of admissible controls U_{ad} is characterized by numbers

$$\underline{u} = 0.2, \ \overline{u} = 1.0, \ c_1 = 5.0, \ c_2 = 4.0.$$

Thus we have 216 state variables (72 elements $\times$ 3 unknowns per element) and 7 design variables (x_1-coordinates of principal moving nodes). The index set $\mathcal{D}_0$ associated with the set Ω_0 contains indices of elements lying in the left half of Ω. The discretized state (nonlinear programming) problems (3.3) were solved by the SQP (Sequential Quadratic Programming) code NLPQLD due to K. Schittkowski [13], while the master optimization problems (3.4) by the NDO code BT (Bundle–Trust) due to H. Schramm and J. Zowe [14]; in particular, by its nonconvex version BTNCLC which can handle linear constraints. Figure 1 shows the optimum shape design obtained after 61 BT iterations together with the values of the 'vertical' components v^{3i-1} of the state vector (left figure) and the corresponding multipliers λ^i (right figure).

Example 3.2 All data are the same as in the previous example, but the surface traction g, which applies now only on the upper half of the left vertical part of $\partial\Omega_0$. The results are shown in Figure 2.

References

[1] F. H. Clarke. *Optimization and nonsmooth analysis.* J. Wiley & Sons, New York, 1983.

[2] M. Giaquinta and E. Giusti. Research on the equilibrium of masonry structures. *Arch. Rational Mech. Anal.*, 88:359–392, 1985.

[3] P. T. Harker and J. S. Pang. Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications. *Math. Prog.*, 48:161–220, 1990.

[4] J. Haslinger and P. Neittaanmaki. *Finite element approximation for optimal shape design: theory and applications.* J. Wiley & Sons, Chichester, 1988.

[5] M. Kočvara and J. V. Outrata. A nondifferentiable approach to the solution of optimum design problems with variational inequalities. In P. Kall, editor, *Proc. of the 15th IFIP Conf. on System Modelling and Optimization*, Lecture Notes in Control Inf. Sci. 180, pages 364–373, Zurich, Sept. 2–6 1991.

[6] M. Kočvara and J. V. Outrata. Shape optimization of elasto-plastic bodies governed by variational inequalities. In *Boundary Control and Boundary Variation*, Sophia Antipolis, 1992. To appear in Lecture Notes in Control Inf. Sci.

[7] J. Kyparisis. Solution differentiability for variational inequalities. *Math. Prog.*, 48:285–301, 1990.

[8] O. L. Mangasarian. *Nonlinear Programming.* McGraw Hill, New York, 1969.

[9] J. V. Outrata. On optimization problems with variational inequality constraints. To appear in *SIAM J. Optimization.*

[10] J. V. Outrata. On necessary optimality conditions for Stackelberg problems. *J. Optim. Theory Appl.*, 76:305–320, 1993.

[11] S. M. Robinson. Strongly regular generalized equations. *Math. Oper. Res.*, 5:43–62, 1980.

[12] S. M. Robinson. An implicit-function theorem for a class of nonsmooth functions. *Math. Oper. Res.*, 16:282–309, 1991.

[13] K. Schittkowski. NLPQL: a FORTRAN subroutine solving constrained nonlinear programming problems. *Annals Oper. Research*, 5:485–500, 1985/86.

[14] H. Schramm and J. Zowe. A version of the bundle idea for minimizing a nonsmooth function: conceptual idea, convergence analysis, numerical results. *SIAM J. Optimization*, 2:121–152, 1992.