ISSN 1995-0802, Lobachevskii Journal of Mathematics, 2016, Vol. 37, No. 5, pp. 561–569. © Pleiades Publishing, Ltd., 2016.

Preconditioned Uzawa-Type Method for a State Constrained Parabolic Optimal Control Problem with Boundary Control

A. Lapin* and **E. Laitinen****

(Submitted by A. M. Elizarov)

Kazan (Volga Region) Federal University, Institute of Computational Mathematics and Information Technologies, Department of Mathematical Statistics, Kremlevskaya ul. 35, Kazan, 420008 Russia
University of Oulu, P.O. Box 3000, 90014 Oulu , Finland

Received August 7, 2015

Abstract—Iterative solution method for mesh approximation of an optimal control problem of a system governed by a linear parabolic equation is constructed and investigated. Control functions of the problem are in the right-hand side of the equation and in Neumann boundary condition, observation is in a part of the domain. Constraints on the control functions, state function and its time derivative are imposed. A mesh saddle point problem is constructed and preconditioned Uzawa-type method is applied to its solution. The main advantage of the iterative method is its effective implementation: every iteration step consists of the pointwise projections onto the segments and solving the linear mesh parabolic equations.

DOI: 10.1134/S1995080216050085

Keywords and phrases: *Parabolic optimal control problem, mesh approximation, state constraints, saddle point problem, iterative method.*

1. INTRODUCTION

State constrained parabolic optimal control problems arise when solving real world applications (see [1, 2] and bibliography therein). While state constrained elliptic optimal control problems are thoroughly investigated, only a few contributions are known on numerical analysis of state constrained parabolic optimal control problems [3–6]. In [3, 4] the problems with point-wise constraints for the state function are investigated. In particular, Lavrentiev-type regularization is applied to the problems with distributed and boundary control in [3], and error bounds for control and state mesh functions are obtained in [4] when approximating the state equation by linear finite elements in space and a discontinuous Galerkin scheme in time. In [5, 6] new iterative solution methods are proposed for finite-dimensional approximations of the problems with point-wise bounds on time derivative of the state. In our knowledge the convergence of mesh approximations of the parabolic optimal problems with constraints for time derivative of the state is not investigated.

A common way to solve optimal control problems consists of using Lagrange functions and constructing the iterative solution methods for the corresponding saddle point problems. Unconstrained saddle point problems are thoroughly investigated (see survey paper [7] containing exhaustive list of references on this subject and recent articles [8, 9]). The development of the efficient numerical methods to solve large scale constrained saddle-point problems is too far from complete. In this way, Uzawa, Arrow–Hurwitz and operator-splitting iterative methods for the constrained saddle point problems arising from augmented Lagrangian approach are investigated in monographs [10, 11]. Solution methods for different classes of the constrained saddle point problems can be found in [12–18].

In this paper we consider a parabolic optimal control problem with distributed and boundary control and with observation in a part of the domain. Constraints on the control, state and on time derivative of

*E-mail: avlapine@mail.ru
**E-mail: erkki.laitinen@oulu.fi

state are imposed. We approximate this problem by finite element in space and weighted finite difference in time scheme, prove the existence of a solution and construct iterative solution method.

We construct preconditioned Uzawa-type iterative solution method with block diagonal preconditioner for the corresponding saddle point problem. The preconditioner is energy equivalent to the "main" matrix of the problem with the constants of the equivalence which don't depend on mesh parameters. The crucial point in constructing the effectively implemented Uzawa type methods is an equivalent transformation of the original saddle point problem following [14].

2. FORMULATION OF THE PROBLEM AND ITS APPROXIMATION

Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with the boundary $\partial\Omega = \Gamma_D \cup \Gamma_N$, meas $\Gamma_D > 0$ and let $\Omega_1 \subseteqq \Omega$ be its subdomain. Let further $Q_T = \Omega \times (0,T]$, $Q_1 = \Omega_1 \times (0,T]$, $\Sigma_D = \Gamma_D \times (0,T]$ and $\Sigma_N = \Gamma_N \times (0,T]$. Denote by $V = \{u \in H^1(\Omega) : u(x) = 0 \text{ on } \Gamma_D\}$ Sobolev space with inner product $(u,v) = \int_\Omega \nabla u \cdot \nabla v \, dx$ and norm $||u|| = (u,u)^{1/2}$.

We consider a parabolic initial-boundary value problem

$$\frac{\partial y}{\partial t} - \Delta y = u \quad \text{in} \quad Q_T,$$
$$y = 0 \quad \text{on} \quad \Sigma_D, \quad \frac{\partial y}{\partial n} = q \quad \text{on} \quad \Sigma_N,$$
$$y = 0 \quad \text{for} \quad t = 0, \quad x \in \Omega, \tag{1}$$

which will be a state equation. The functions $u = u(x,t)$ and $q = q(x,t)$ are variable control functions, and the solution $y(x,t)$ of (1) is a state function.

Proposition 1. *Let $\partial\Omega \in C^2$, $u \in L_2(Q_T)$ and $q \in W = L_2(0,T;H^{1/2}(\Gamma_N)) \cap H^{1/4}(0,T;L_2(\Gamma_N))$. Then there exists a unique solution y of problem (1), such that $y \in L_\infty(0,T;V) \cap H^1(0,T;L_2(\Omega))$ and the following stability inequality takes place:*

$$\sup_{0\leqslant t\leqslant T} ||y(t)||_V + \left|\left|\frac{\partial y(t)}{\partial t}\right|\right|_{L_2(Q_T)} \leqslant C_a \left(||u(t)||_{L_2(Q_T)} + ||q(t)||_W\right), \quad C_a = \text{const}. \tag{2}$$

The proof of the proposition above can be found in [19] on page 34.

The mentioned regularity properties of state function y allow to define, in particular, the point-wise constraints for its time derivative. Define the following sets of constraints:

$$U_{ad} = \{u \in L_2(Q_T) : \ |u(x,t)| \leqslant u_{\max} \quad \text{a.e.} \quad (x,t) \in Q_T\},$$
$$Q_{ad} = \{q \in W : \ |q| \leqslant \bar{q} \quad \text{a.e.} \quad \Sigma_N\},$$
$$Y_{ad} = \left\{ y \in L_2(0,T;H_0^1(\Omega)) : \ \frac{\partial y}{\partial t} \in L_2(Q_T), \quad y_{\min} \leqslant y(x,t) \leqslant y_{\max} \right.$$
$$\left. \text{and} \quad dy_{\min} \leqslant \frac{\partial y}{\partial t} \leqslant dy_{\max} \quad \text{a.e.} \quad Q_T \right\}.$$

Above constants $\bar{u} > 0$, $\bar{q} > 0$ and $-\infty \leqslant y_{\min}, dy_{\min} < 0 < y_{\max}, dy_{\max} \leqslant \infty$.

Let an objective function be defined by the equality

$$J(y,u,q) = \frac{1}{2}\int\limits_{Q_1} (y(x,t) - y_d(x,t))^2 dxdt + \frac{1}{2}\int\limits_{Q_T} u^2 dxdt + \frac{1}{2}\int\limits_{\Sigma_N} q^2 d\Gamma dt$$

with a given observation function $y_d(x,t) \in L_2(Q_1)$.

We will solve the following optimal control problem:

$$\min_{(y,u,q)\in K} J(y,u,q),$$
$$K = \{(y,u,q) \in Y_{ad} \times U_{ad} \times Q_{ad} : \quad \text{equation (1) holds}\}. \tag{3}$$

Lemma 1. *Problem* (3) *has a unique solution* (y, u, q).

Proof. The sets of constraints U_{ad}, Q_{ad} and Y_{ad} are convex, closed and contain zero elements, moreover U_{ad} and Q_{ad} are bounded. These properties together with linearity of state equation and stability inequality (2) ensure that the set K is convex, closed, bounded and nonempty. Functional $J = J(y, u, q)$ is continuous. The established properties of J and K ensure the existence of a solution to problem (3). Its uniqueness follows from the strict convexity of the functional J on the set K. To prove this property of J we observe that it is convex in y and strictly convex in u and q, and the equalities $u_1 = u_2$ and $q_1 = q_2$ imply $y_1 = y_2$ for the solutions of problem (1). □

We construct an approximation of problem (3) supposing for the simplicity that the domains Ω and Ω_1 have polygonal boundaries and that the function y_d is continuous.

Let a family T_h of nonoverlapping closed triangles e (finite elements) with maximal diameter h compose a conforming and regular triangulation $\overline{\Omega} = \bigcup_{e\in T_h} e$ of $\overline{\Omega}$ ([20], p. 124). We suppose that T_h generates the triangulations T_h^1 on $\overline{\Omega}_1$ and ∂T_h on $\overline{\Gamma}_N$, i.e. $\overline{\Omega}_1$ consists of integer number of $e \in T_h$ and $\overline{\Gamma}_N$ consists of integer number of sides ∂e of elements $e \in T_h$. Define the finite element space $V_h \subset V$ of the continuous and piecewise linear functions (linear on each e) which vanish on the boundary Γ_D and the finite element space $U_h \in L_2(\Gamma_N)$ of the piecewise linear functions on Γ_N (linear on each $\partial e \in \Gamma_N$), which are the traces on Γ_N of the functions from V_h.

To approximate the integrals of a continuous function $g(x)$ over a finite element $e \in T_h$ or its side ∂e we use the quadrature formulas

$$\int_e g(x)dx \approx S_e(g) = \frac{1}{3}\text{meas}(e)\sum_{\alpha=1}^{3} g(x_\alpha), \quad x_\alpha \text{ are the vertices of } e,$$

$$\int_{\partial e} g(x)d\Gamma \approx S_{\partial e}(g) = \frac{1}{2}\text{meas}(\partial e)\sum_{\alpha=1}^{2} g(x_\alpha), \quad x_\alpha \text{ are the vertices of } \partial e.$$

The corresponding composite quadrature formulas are

$$S_\Omega(g) = \sum_{e\in T_h} S_e(g), \quad S_{\Omega_1}(g) = \sum_{e\in T_h^1} S_e(g), \quad S_\Gamma(g) = \sum_{\partial e\in \partial T_h} S_{\partial e}(g).$$

Let further $\omega_t = \{t_j = j\tau,\ j = 0, 1, \dots N_t;\ N_t\tau = T\}$ be a uniform mesh on the segment $[0, T]$. We denote by y_h with subscript h a mesh function from the space V_h or U_h and by y_h^j a time depending mesh function at a time level $t_j \in \omega_t$. Let also y_{dh}^j be the continuous and piecewise linear in space variables function which coincides with $y_d(x, t_j)$ at the nodes of the triangulation T_h^1.

Approximation of state problem (1) is the following weighted finite-difference in time and finite element in space problem:

$$\begin{aligned} &S_\Omega\left(\frac{y_h^j - y_h^{j-1}}{\tau} z_h\right) + S_\Omega\left(\nabla(\sigma y_h^j + (1-\sigma)y_h^{j-1}) \cdot \nabla z_h\right) \\ &= S_\Omega(u_h^j\, z_h) + S_\Gamma(q_h^j\, z_h) \quad \forall z_h \in V_h, \quad j = 1, 2, \dots, N_t, \end{aligned} \tag{4}$$

with initial value $y_h^0 = 0$ and a weight $\sigma \in [0, 1]$. This scheme includes: forward Euler ($\sigma = 0$), backward Euler ($\sigma = 1$) and Crank–Nicolson ($\sigma = 1/2$) schemes.

Define the approximations of the objective function and the sets of constraints by the following equalities:

$$J_h(y_h, u_h, q_h) = \frac{\tau}{2}\sum_{j=1}^{N_t}\left(S_{\Omega_1}((y_h^j - y_{d\,h}^j)^2) + S_\Omega(u_h^j)^2 + S_\Gamma(q_h^j)^2\right), \tag{5}$$

$$U_{ad}^h = \{|u_h^j| \leqslant \bar{u} \quad \forall x \in \Omega, \quad j = 1, 2, \dots, N_t\},$$

$$Q_{ad}^h = \{|q_h^j| \leqslant \bar{q} \quad \forall x \in \Omega, \quad j = 1,2,\dots,N_t\},$$

$$Y_{ad}^h = Y_0^h \bigcap Y_1^h, \quad Y_0^h = \{y_h^j : y_{\min} \leqslant y_h^j \leqslant y_{\max}, \quad \forall x \in \Omega, \quad j = 1,2,\dots,N_t\},$$

$$Y_1^h = \{y : \tau dy_{\min} \leqslant y_h^j - y_h^{j-1} \leqslant \tau dy_{\max} \quad \forall x \in \ \Omega, \quad j = 1,2,\dots,N_t, \quad (y_h^0 = 0)\}. \tag{6}$$

Approximation procedures result to the following mesh optimal control problem:

$$\text{find} \quad \min_{(y_h,u_h,q_h)\in K_h} J_h(y_h,u_h,q_h),$$

$$K_h = \{(y_h,u_h,q_h) : \ y_h \in Y_{ad}^h, u_h \in U_{ad}^h, \quad q_h \in Q_{ad}^h, \ \text{equation (4) holds}\}. \tag{7}$$

Lemma 2. *Mesh optimal control problem* (7) *has a unique solution* (y_h, u_h, q_h).

Proof. Similar to lemma 1 the result follows from the facts that the set K_h is nonempty, closed, convex and bounded, while the function J_h is continuous and strictly convex on K_h. □

It is well-known that problem (4) is unconditionally stable for $\sigma \geqslant 1/2$. In the case $0 \leqslant \sigma < 1/2$ it is stable under the additional condition for time-step $\tau < \tau_0(h) \simeq h^2$. More precisely, the following statement takes place (this is a slightly modified result of [22], p. 391):

Proposition 2. Let $0 \leqslant \sigma < 1/2$ and $\tau \leqslant 2(\nu_{\max}(1-2\sigma))^{-1}$, where $\nu_{\max}$ is the maximal eigenvalue of the following eigenvalue problem:

$$(y_h, \nu) : \ S_\Omega\left(\nabla y_h \cdot \nabla z_h\right) = \nu S_\Omega(y_h\, z_h) \quad \forall z_h \in V_h.$$

Then for a solution to problem (4) *the following stability inequality holds:*

$$\sum_{j=1}^{N_t} S_\Omega\left(|y_h^j|^2\right) \leqslant C_T \left(\sum_{j=1}^{N_t} S_\Omega\left(|u_h^j|^2\right) + \sum_{j=1}^{N_t} S_\Gamma\left(|q_h^j|^2\right)\right), \quad C_T = \text{const}. \tag{8}$$

3. ALGEBRAIC FORM OF PROBLEM (7) AND SADDLE POINT PROBLEM

Denote by $y \in \mathbb{R}^{N_y}$ the vector of nodal values of a function $y_h \in V_h$ ($N_y = \dim . V_h$) Then we get the "onto" correspondence $y \Leftrightarrow y_h$. Similarly a vector $q \in \mathbb{R}^{N_q}$ corresponds to $q_h \in Q_h$.[1)] By $(.,.)_y$ ($(.,.)_q$) and $||.||_y$ ($||.||_q$) we denote the inner product and euclidian norm in $\mathbb{R}^{N_y}$ ($\mathbb{R}^{N_q}$) and by $(.,.)$ and $||.||$ as the inner product and euclidian norm in $\mathbb{R}^{N_tN_y}$ and $\mathbb{R}^{N_tN_q}$ (concrete case will be obvious from the context).

Define stiffness matrix $A \in \mathbb{R}^{N_y\times N_y}$, diagonal matrices $\tilde{M}, \tilde{M}_y \in \mathbb{R}^{N_y\times N_y}$ and $\tilde{M}_q \in \mathbb{R}^{N_q\times N_q}$, and rectangular matrix $\tilde{S}_q \in \mathbb{R}^{N_y\times N_q}$, by the following equalities:

$$(Ay,z)_y = S_\Omega\left(\nabla y_h \cdot \nabla z_h\right), \quad (\tilde{M}y,z)_y = S_\Omega(y_h z_h),$$

$$(\tilde{M}_y y,z)_y = S_{\Omega_1}(y_h z_h), \quad (\tilde{M}_q q,p)_q = S_\Gamma(q_h p_h), \quad (\tilde{S}_q q,z)_y = S_\Gamma(q_h z_h).$$

Above $y \Leftrightarrow y_h \in V_h$, $z \Leftrightarrow z_h \in V_h$ and $q \Leftrightarrow q_h \in Q_h$, $p \Leftrightarrow p_h \in Q_h$. With these notations mesh state equation (4) and objective function (5) can be written for the vectors of nodal values of mesh functions:

$$\tilde{M}\frac{y^j - y^{j-1}}{\tau} + A(\sigma y^j + (1-\sigma)y^{j-1}) = \tilde{M}u_j + \tilde{S}_q q^j, \quad j = 1,2,\dots,N_t, \quad y^0 = 0, \tag{9}$$

$$I(y,u,q) = \frac{\tau}{2}\sum_{j=1}^{N_t}\left((\tilde{M}_y(y^j - y_d^j), y^j - y_d^j)_y + (\tilde{M}u^j,u^j)_y + (M_q q^j, q^j)_q\right). \tag{10}$$

Further we use also the block diagonal matrices with constant blocks, namely, $M = \mathrm{diag}(\tilde{M},\tilde{M},\dots,\tilde{M}) \in \mathbb{R}^{N_tN_y\times N_tN_y}$ and similarly defined $M_y \in \mathbb{R}^{N_tN_y\times N_tN_y}$, $M_q \in \mathbb{R}^{N_tN_q\times N_tN_q}$ and $S_q \in \mathbb{R}^{N_tN_y\times N_tN_q}$. Define also matrix $L \in \mathbb{R}^{N_tN_y\times N_tN_y}$:

$$(Ly)_j = \{\tilde{M}\frac{y^j - y^{j-1}}{\tau} + A(\sigma y^j + (1-\sigma)y^{j-1}), \quad j = 1,2,\dots,N_t\},$$

[1)] Since hereafter we consider only finite dimensional problems, we use the same notations for the vectors as previously for the functions.

with formal component $y^0 = 0$.

Now we can rewrite mesh state equation (9) and objective function (10) in the following short manner (we scale objective function by dividing to τ):

$$Ly = Mu + S_q q,$$
$$I(y,u,q) = \frac{1}{2}(M_y(y-y_d), y-y_d) + \frac{1}{2}(Mu,u) + \frac{1}{2}(M_q q, q).$$

Note, that stability inequality (8) implies the estimate

$$(My, y) \leqslant C_T\big((Mu,u) + (M_q q, q)\big). \tag{11}$$

Point-wise constraints (6) can be obviously rewritten for the vectors of nodal values of mesh functions. Let further

$$R \in \mathbb{R}^{N_t N_y \times N_t N_y}, \quad (Ry)^j = \{y^j - y^{j-1} \text{ for } j = 2,\dots,M;\ y^1 \quad \text{for} \quad j = 1\}.$$

Then we can replace the constraint $y \in Y_{ad}^{1h}$ in the discrete optimal control problem by the following constraints:

$$p \in P_{ad}^h = \{\tau dy_{\min} \leqslant p_j \leqslant \tau dy_{\max},\ j = 1,2,\dots M\}, \quad Ry - p = 0.$$

At last, denote by ψ, θ, φ_u and φ_q the indicator functions of the sets Y_0^h, P_{ad}^h, U_{ad}^h and Q_{ad}^h, respectively.

As a result we obtain the following algebraic form of mesh optimal control problem (7):

$$\min_{Ly = Mu + S_q q,\ p = Ry} \{I(y,u,q) + \psi(y) + \theta(p) + \varphi_u(u) + \varphi_q(q)\}. \tag{12}$$

Construct Lagrange function for problem (12):

$$\mathcal{L}(y,u,q,p,\lambda,\mu) = I(y,u,q) + \psi(y) + \theta(p) + \varphi_u(u) + \varphi_q(q) + (\lambda, Ly - Mu - S_q q) + (\mu, Ry - p).$$

A saddle point of this Lagrangian satisfies the following saddle point problem (cf., e.g. [21], p. 169):

$$\begin{pmatrix} M_y & 0 & 0 & 0 & L^T & R^T \\ 0 & M & 0 & 0 & -M & 0 \\ 0 & 0 & M_q & 0 & -S_q^T & 0 \\ 0 & 0 & 0 & 0 & 0 & -E \\ L & -M & -S_q & 0 & 0 & 0 \\ R & 0 & 0 & -E & 0 & 0 \end{pmatrix} \begin{pmatrix} y \\ u \\ q \\ p \\ \lambda \\ \mu \end{pmatrix} + \begin{pmatrix} \partial\psi(y) \\ \partial\varphi_u(u) \\ \partial\varphi_q(q) \\ \partial\theta(p) \\ 0 \\ 0 \end{pmatrix} \ni \begin{pmatrix} M_y y_d \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{13}$$

where $\partial\psi$, $\partial\varphi_u$, $\partial\varphi_q$ and $\partial\theta$ are the subdifferentials of the corresponding functions and E is identity matrix. With the notations $z = (y,u,q,p)^T$, $\eta = (\lambda,\mu)^T$, $f = (M_y y_d, 0, 0, 0)^T$, $\Psi(z) = \psi(y) + \theta(p) + \varphi_u(u) + \varphi_q(q)$, and

$$\mathcal{A} = \begin{pmatrix} M_y & 0 & 0 & 0 \\ 0 & M & 0 & 0 \\ 0 & 0 & M_q & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} L & -M & -S_q & 0 \\ R & 0 & 0 & -E \end{pmatrix}$$

problem (13) can be rewritten in a compact form:

$$\begin{pmatrix} \mathcal{A} & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} z \\ \eta \end{pmatrix} + \begin{pmatrix} \partial\Psi(z) \\ 0 \end{pmatrix} \ni \begin{pmatrix} f \\ 0 \end{pmatrix}.$$

The degenerate matrix $\mathcal{A}$ is an obstacle to the application of Uzawa-type iterative methods for solving this saddle point problem. To overcome this deficiency we use two last equations of system (13) to transform it to the equivalent saddle point problem

$$\begin{pmatrix} \mathcal{A}_r & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} z \\ \eta \end{pmatrix} + \begin{pmatrix} \partial\Psi(z) \\ 0 \end{pmatrix} \ni \begin{pmatrix} f \\ 0 \end{pmatrix} \tag{14}$$

with

$$\mathcal{A}_r = \begin{pmatrix} M_y + r_1 M & -r_1 M L^{-1} M & -r_1 M L^{-1} S_q & 0 \\ 0 & M & 0 & 0 \\ 0 & 0 & M_q & 0 \\ -r_2 M R & 0 & 0 & r_2 M \end{pmatrix}, \quad r_1 > 0, \quad r_2 > 0,$$

instead of $\mathcal{A}$ and with the same matrix B and function Ψ.

Lemma 3. *Let* $(r_1, r_2) \in \omega = \{0 < r_1 < 2/C_T,\ 0 < r_2 < r_1 - r_1^2 C_T/2\}$. *Then matrix* $\mathcal{A}_r$ *is positive definite and energy equivalent to block-diagonal matrix*

$$\mathcal{A}_{00} = \mathrm{diag}\Big(M,\ M,\ M_q,\ M\Big)$$

with constants of the equivalence, which depend only on r_1, r_2:

$$c_0(r_1, r_2)(\mathcal{A}_{00} z, z) \leqslant (\mathcal{A}_r z, z) \leqslant c(r_1, r_2)(\mathcal{A}_{00} z, z), \quad z = (y, u, q, p)^T. \tag{15}$$

Proof. Let $z = (y, u, q, p)^T$, then

$$(\mathcal{A}_r z, z) = ((M_y + r_1 M) y, y) - r_1 (L^{-1}(Mu + S_q q), My) + (Mu, u) \\ + (M_q q, q) + r_2 (Mp, p) - r_2 (Mp, Ry).$$

Denote by $\tilde{y}$ the solution of the equation $L\tilde{y} = Mu + S_q q$. Then due to the stability inequality (11) $(M\tilde{y}, \tilde{y}) \leqslant C_T\big((Mu, u) + (M_q q, q)\big)$. Using this estimate and the inequality $|(Ry, Mp)| \leqslant 2(My, y)^{1/2}(Mp, p)^{1/2}$ we get

$$(\mathcal{A}_r z, z) \geqslant r_1 (My, y) + (Mu, u) + (M_q q, q) + r_2 (Mp, p) \\ - r_1 C_T^{1/2} (Mu, u)^{1/2} (My, y)^{1/2} - r_1 C_T^{1/2} (M_q q, q)^{1/2} (My, y)^{1/2} - 2 r_2 (My, y)^{1/2} (Mp, p)^{1/2}.$$

For $(r_1, r_2) \in \omega$ the quadratical form $F(y, u, q, p) = r_1 y^2 + u^2 + q^2 + r_2 p^2 - r_1 C_T^{1/2} uy - r_1 C_T^{1/2} qy - 2 r_2 yp$ is positive definite and there exists $c_0 > 0$ such that $F(y, u, q, p) \geqslant c_0 (y^2 + u^2 + q^2 + p^2)$. As a consequence $(\mathcal{A}_r z, z) \geqslant c_0 (\mathcal{A}_{00} z, z)$. Since $(M_y y, y) \leqslant (My, y)$, we obtain

$$(\mathcal{A}_r z, z) \leqslant (1 + r_1)(My, y) + (Mu, u) + (M_q q, q) + r_2 (Mp, p) \\ + r_1 C_T^{1/2} (Mu, u)^{1/2} (My, y)^{1/2} + r_1 C_T^{1/2} (M_q q, q)^{1/2} (My, y)^{1/2} + 2 r_2 (My, y)^{1/2} (Mp, p)^{1/2},$$

whence $(\mathcal{A}_r z, z) \leqslant c_1 (\mathcal{A}_{00} z, z)$ with a constant c_1 depending on r_1, r_2. □

Theorem 1. *Problem* (13) *has a solution* $(y, u, q, p, \lambda, \mu)$ *with unique* y, u, q, p, *which coincide with the solution of problem* (12).

Proof. Matrix $\mathcal{A}_r$ is positive definite, matrix B has a full column rank and function Ψ is convex, proper and lower semicontinuous. Moreover, zero vector satisfies the equation $Bz = 0$ and belongs to int dom Ψ. All these properties ensure the result of the theorem (cf. [23]). □

4. ITERATIVE SOLUTION METHOD FOR SADDLE POINT PROBLEM

A preconditioned Uzawa-type iterative method for solving saddle point problem (14) reads as follows:

$$\mathcal{A}_r z^{k+1} + \partial\Psi(z^{k+1}) \ni B^T\eta^k + f,$$
$$\frac{1}{\rho}D(\eta^{k+1} - \eta^k) + Bz^{k+1} = 0, \tag{16}$$

where D is a symmetric and positive definite matrix (preconditioner), $\rho > 0$ is an iterative parameter. Iterative method (16) converges for any initial guess η^0 if the pair (D, ρ) satisfies the following assumption ([14]):

$$D \geqslant \frac{(1+\varepsilon)\rho}{2} B\mathcal{A}_{rs}^{-1}B^T, \quad \varepsilon > 0, \tag{17}$$

where $\mathcal{A}_{rs} = 0.5(\mathcal{A} + \mathcal{A}^T)$ is the symmetric part of $\mathcal{A}_r$.

Below we construct an easy invertible block diagonal preconditioner D which is spectrally equivalent to $B\mathcal{A}_{rs}^{-1}B^T$ with the constants, which don't depend on meshsizes h and τ.

Due to lemma 3 the matrix $B\mathcal{A}_{rs}^{-1}B^T$ is spectrally equivalent to

$$B\mathcal{A}_{00}^{-1}B^T = \begin{pmatrix} LM^{-1}L^T + M + S_qM_q^{-1}S_q^T & LM^{-1}R^T \\ RM^{-1}L^T & RM^{-1}R^T + M^{-1} \end{pmatrix}.$$

In turn, this matrix is spectrally equivalent to a block-diagonal one. More precisely, the following statement takes place for $K_T = 1 + C_T + \sqrt{C_T^2 + 4}$.

Lemma 4. *Matrix* $D = \begin{pmatrix} LM^{-1}L^T & 0 \\ 0 & M^{-1} \end{pmatrix}$ *is spectrally equivalent to* $B\mathcal{A}_{00}^{-1}B^T$ *with constants, which don't depend on meshsizes* h *and* τ. *In particular,*

$$(B\mathcal{A}_{00}^{-1}B^T\eta, \eta) \leqslant K_T(D\eta, \eta) \quad \forall \eta = (\lambda, \mu). \tag{18}$$

Proof. Using the inequalities

$$(M\lambda, \lambda) + (M_q^{-1}S_q^T\lambda, S_q^T\lambda) \geqslant 0 \quad \text{and} \quad (M^{-1}R^T\mu, R^T\mu) \leqslant 4(M^{-1}\mu, \mu)$$

we estimate the quadratical form $(B\mathcal{A}_{00}^{-1}B^T\eta, \eta)$ from below:

$$(B\mathcal{A}_{00}^{-1}B^T\eta, \eta) = (M^{-1}L^T\lambda, L^T\lambda) + (M\lambda, \lambda) + (M_q^{-1}S_q^T\lambda, S_q^T\lambda)$$
$$+ (M^{-1}R^T\mu, R^T\mu) + (M^{-1}\mu, \mu) + 2(M^{-1}L^T, R^T\mu)$$
$$\geqslant \left(1 - \frac{1}{\varepsilon}\right)(M^{-1}L^T\lambda, L^T\lambda) + (1-\varepsilon)(M^{-1}R^T\mu, R^T\mu)$$
$$+ (M^{-1}\mu, \mu) \geqslant \left(1 - \frac{1}{\varepsilon}\right)(M^{-1}L^T\lambda, L^T\lambda) + (1 - 4|1-\varepsilon|)(M^{-1}\mu, \mu).$$

For a fixed $0 < \varepsilon < (\sqrt{2} - 1)/2$ we get $(B\mathcal{A}_{00}^{-1}B^T\eta, \eta) \geqslant c(\varepsilon)(D\eta, \eta)$, $c(\varepsilon) > 0$.

Let us prove estimate (18). For any $\varepsilon > 0$ we have

$$(B\mathcal{A}_{00}^{-1}B^T\eta, \eta) \leqslant \left(1 + \frac{1}{\varepsilon}\right)(M^{-1}L^T\lambda, L^T\lambda) + (M\lambda, \lambda)$$
$$+ (M_q^{-1}S_q^T\lambda, S_q^T\lambda) + (1 + 4\varepsilon)(M^{-1}\mu, \mu). \tag{19}$$

Due to Cauchy inequality and stability estimate (11) the following chain of the inequalities is true:

$$||M^{-1/2}L^T\lambda|| = \sup_v \frac{(M^{-1/2}L^T\lambda, v)}{||v||} = \sup_y \frac{(\lambda, Ly)}{||M^{1/2}y||}$$

$$\geqslant \sup_{u,q} \frac{(\lambda, Mu + S_q q)}{||M^{1/2}(L^{-1}(Mu + S_q q))||} \geqslant \frac{1}{C_T^{1/2}} \sup_{u,q} \frac{(\lambda, Mu + S_q q)}{||M^{1/2}u|| + ||S_q^{1/2} q||}.$$

Choosing subsequently $q = 0$ and $u = 0$ in this inequality we have

$$||M^{-1/2}L^T\lambda|| \geqslant \frac{1}{C_T^{1/2}} \sup_{u} \frac{(\lambda, Mu)}{||M^{1/2}u||} = \frac{1}{C_T^{1/2}} ||M^{1/2}\lambda||,$$

$$||M^{-1/2}L^T\lambda|| \geqslant \frac{1}{C_T^{1/2}} \sup_{q} \frac{(\lambda, S_q q)}{||S_q^{1/2} q||} = \frac{1}{C_T^{1/2}} ||M^{-1/2}S_q^T\lambda||. \tag{20}$$

Estimates (19) and (20) yield

$$(B\mathcal{A}_{00}^{-1}B^T\eta, \eta) \leqslant \left(1 + \frac{1}{\varepsilon} + 2C_T\right)(M^{-1}L^T\lambda, L^T\lambda) + (1 + 4\varepsilon)(M^{-1}\mu, \mu).$$

For $\varepsilon = (K_T - 1)/4$ we get estimate (18). □

Method (16) for problem (13) with preconditioner $D = \begin{pmatrix} LM^{-1}L^T & 0 \\ 0 & M^{-1} \end{pmatrix}$ reads as follows:

$$\begin{cases} Mu^{k+1} + \partial\varphi_u(u^{k+1}) \ni M\lambda^k, \\ M_q q^{k+1} + \partial\varphi_q(q^{k+1}) \ni S_q^T \lambda^k, \\ (M_y + r_1 M)y^{k+1} + \partial\psi(y^{k+1}) \ni My_d + r_1 ML^{-1}Mu^{k+1} \\ \; + r_1 ML^{-1}S_q q^{k+1} - L^T\lambda^k - R^T\mu^k, \\ r_2 Mp^{k+1} + \partial\theta(p^{k+1}) \ni r_2 MRy^{k+1} + \mu^k, \end{cases} \tag{21}$$

$$\begin{cases} LM^{-1}L^T \dfrac{\lambda^{k+1} - \lambda^k}{\rho} = Ly^{k+1} - Mu^{k+1} - S_q q^{k+1}, \\ \dfrac{\mu^{k+1} - \mu^k}{\rho} = MRy^{k+1} - Mp^{k+1}. \end{cases} \tag{22}$$

Theorem 2. *Method* (21), (22) *converges if* $(r_1, r_2) \in \omega$ *and*

$$0 < \rho < 2c_0(r_1, r_2)/K_T, \tag{23}$$

where the domain ω *and the constant* $c_0(r_1, r_2)$ *are defined in lemma* 3.

Proof. From inequality (15) we get the following estimate:

$$\mathcal{A}_{rs}^{-1} \leqslant c_0^{-1}(r_1, r_2)\mathcal{A}_{00}^{-1}.$$

This estimate and (18) yield

$$B\mathcal{A}_{rs}^{-1}B^T \leqslant c_0^{-1}(r_1, r_2)B\mathcal{A}_{00}^{-1}B^T \leqslant c_0^{-1}(r_1, r_2)K_T D.$$

Thus, assumption (17) is satisfied if the iterative parameter ρ satisfies the inequality $c_0^{-1}(r_1, r_2)K_T < 2/\rho$, which is just (23). □

Implementation. On every step of method (21), (22) we have to solve three inclusions (21) with diagonal matrices and diagonal operators. Solving the inclusions reduces to pointwise projections (for all coordinates of nodal vectors on every time level) on the corresponding sets of the constraints.

Solving a system of linear equations with the matrix $LM^{-1}L^T$ consists of sequential solution of the systems with the matrices L and L^T. In the particular case of the explicit finite difference approximation of state equation ($\sigma = 0$) these matrices are triangle ones and the solutions can be found by explicit calculations.

REFERENCES

1. R. Dautov, R. Kadyrov, E. Laitinen, A. Lapin, J. Pieskä, and V. Toivonen, "On 3D dynamic control of secondary cooling in continuous casting process," Lobachevskii J. Math. **13**, 3–13 (2003).
2. M. Gunzburger, E. Ozugurlu, J. Turner, and H. Zhang, "Controlling transport phenomena in the Czochralski crystal growth process," J. Cryst. Growth **234**, 47–62 (2002).
3. I. Neitzel and F. Troltzsch, "On regularization methods for the numerical solution of parabolic control problems with pointwise state constraints," in *ESAIM: Control, Optimisation and Calculus of Variations*. doi 10.1051/cocv:2008038 (2008).
4. K. Deckelnick and M. Hinze, "Variational discretization of parabolic control problems in the presence of pointwise state constraints," J. Comp. Math. **29**, 1–15 (2011).
5. A. Lapin and E. Laitinen, "Explicit algorithms to solve a class of state constrained parabolic optimal control problems," Russian J. Numer. Analysis Math. Modeling **30** (6), 351–362 (2015).
6. A. Lapin and E. Laitinen, "Iterative solution methods for parabolic optimal control problem with constraints on time derivative of state function," WSEAS Recent Advances in Mathematics: Mathematics and Computers in Science and Engineering **48**, 72–74 (2015).
7. M. Benzi, G. Golub, and J. Liesen, "Numerical solution of saddle point problems," Acta Numerica **14**, 1–137 (2005).
8. S. Schaible and J-C. Yao, "Recent developments in solution methods for variational inequalities and fixed point problems," in *Proceedings of the American Conference on Applied Mathematics (AMERICAN-MATH'10)* (Harvard University, Cambridge, 2010), pp. 142–145.
9. C. Pan, "On generalized preconditioned Hermitian and skew-Hermitian splitting methods for saddle point problems," WSEAS Transactions on Mathematics **11** (12), 1147–1156 (2012).
10. M. Fortin and R. Glowinski, *Augmented Lagrangian Methods* (North-Holland, Amsterdam, 1983).
11. R. Glowinski and P. LeTallec, *Augmented Lagrangian Operator-Splitting Methods in Nonlinear Mechanics* (SIAM, Philadelphia, PA, 1989).
12. N. Pop, "Saddle point formulation of the quasistatic contact problems with friction," in *Proceedings of the 7th WSEAS International Conference on Systems Theory and Scientific Computation* (Athens, 2007), pp. 250–254.
13. C. Graser and R. Kornhuber, "Nonsmooth Newton methods for set-valued saddle point problems," SIAM J. Numer. Anal. **47**, 1251–1273 (2009).
14. A. Lapin, "Preconditioned Uzawa type methods for finite-dimensional constrained saddle point problems," Lobachevskii J. Math. **31** (4), 309–322 (2010).
15. E. Laitinen, A. Lapin and S. Lapin, "On the iterative solution of finite-dimensional inclusions with applications to optimal control problems," Comp. Methods in Appl. Math. **10** (3), 283–301 (2010).
16. E. Laitinen, A. Lapin, "Iterative solution methods for a class of state constrained optimal control problems," Applied Mathematics **3** (12), 1862–1867 (2012).
17. E. Laitinen and A. Lapin, "Iterative solution methods for the large-scale constrained saddle point problems," Numerical methods for differential equations, optimization, and technological problems, Comp. Meth. Appl. Sci. **27**, 19–39 (2013).
18. A. Lapin, M. Khasanov, "Iterative solution methods for mesh approximation of control and state constrained optimal control problem with observation in a part of the domain," Lobachevskii J. Math. **35** (3), 241–258 (2014).
19. J.-L. Lions and E. Magenes, *Problemes aux limites non homogenes et applications* (Dunod, Paris, 1968).
20. Ph. G. Ciarlet, *The Finite Element Method for Elliptic Problems* (North-Holland, Amsterdam, 1978).
21. I. Ekeland and R. Temam, *Convex Analysis and Variational Problems* (North-Holland, Amsterdam, 1976).
22. A. Quarteroni and A. Valli, *Numerical Approximation of Partial Differential Equations* (Springer, 1997).
23. E. Laitinen, A. Lapin, and S. Lapin, "Iterative solution methods for variational inequalities with nonlinear main operator and constraints to gradient of solution," Lobachevskii J. Math. **33** (4), 364–371 (2012).