Optimal Control: Theory, Algorithms, and Applications, pp. 1-2
W. W. Hager and P. M. Pardalos, Editors

On a SQP-Multigrid Technique for Nonlinear Parabolic Boundary Control Problems *

Helmuth Goldberg
Technical University of Chemnitz–Zwickau, Faculty of Mathematics, D–09107 Chemnitz, Germany
goldberg@mathematik.tu-chemnitz.de

Fredi Tröltzsch
Technical University of Chemnitz–Zwickau, Faculty of Mathematics, D–09107 Chemnitz, Germany
f.troeltzsch@mathematik.tu-chemnitz.de

Received April 22, 1997; Accepted October 14, 1997

Abstract

We consider the application of an SQP method to an optimal control problem governed by the heat equation with nonlinear boundary conditions. The objective functional consists of a quadratic terminal part and a quadratic regularization term. To handle the quadratic optimal control subproblems with high precision, very large scale mathematical programming problems have to be treated. The solution of the constrained problem is computed by solving a sequence of unconstrained ones by a method due to Bertsekas. A multigrid approach developed by Hackbusch is applied to solve the unconstrained problems. Some numerical examples illustrate the behaviour of the method.

Keywords: optimal control, semilinear parabolic equation, multigrid method, SQP method.

*Supported by Deutsche Forschungsgemeinschaft (DFG), under grant "Tr 302/3–1".

1 Introduction

The behaviour of Lagrange–Newton–SQP methods for solving nonlinear optimal control problems has been the subject of several recent publications. For instance, their application to the control of ordinary differential equations was discussed by Alt [1], [2], [3], Alt and Malanowski [5], Machielsen [28]. The case of weakly singular integral equations was considered by Alt, Sontag and Tröltzsch [6]. Control problems for nonlinear partial differential equations were studied by Heinkenschloss [18], Heinkenschloss and Sachs [19], Heinkenschloss and Tröltzsch [20], Ito and Kunisch [21], [22], Kelley and Sachs [23], [24], Kupfer and Sachs [26], Goldberg and Tröltzsch [14], and Tröltzsch [33], [34]. We refer also to a recent paper by Gill and others [11]. It is meanwhile shown in most of the cases mentioned above that the (continuous) SQP method exhibits the expected local q–quadratic convergence in spaces of type L^∞. We refer to [6], [20], [33], [34] for the proof under strong second order sufficient optimality conditions. A detailed convergence analysis assuming weaker second order conditions is contained for a simplified model in [14] and for a general class of control problems governed by semilinear parabolic equations in [35]. Owing to their quadratic convergence, these methods appear to be promising for a high precision numerical solution of control problems.

In the applications to function spaces, the method has to be linked with a discretization. It may appear on a different level. The simplest way is that of discretizing the optimal control problem as a whole, to obtain a large scale finite–dimensional optimization problem. Then the SQP method is applied in finite dimensions. This direct approach was successfully tested for many control problems governed by ordinary differential equations and for some parabolic control problems with moderate discretization.

However, the use of direct methods is limited due to the appearing large dimensions. Even a moderate accuracy for solving the state equation may lead to a huge number of variables. For instance, discretizing a parabolic equation in a rectangular 2D–domain with uniform meshes of 100 node points for the time and the two space variables leads to 10^6 state variables. Note that this large number is already needed to solve the state equation with a moderate precision of order 10^{-2}. Moreover, the analytic scheme of the discretization should be available to establish the discretized problem. Very efficient solvers for partial differential equations may use grid generators and time-dependent adaptive grids, for instance in domains with curved boundaries. This is another reason to avoid direct methods in some

situations.

In this paper, we pursue a different strategy. Considering the SQP method in an infinite-dimensional setting, a sequence of constrained linear-quadratic optimal control problems has to be solved. The solution of these sub-problems is determined iteratively through unconstrained control problems using a projection method due to Bertsekas. On this level, necessary and sufficient optimality conditions are expressed by a fixed-point equation, which can be solved by a multigrid technique owing to Hackbusch [15]. We refer also to Hackbusch and Will [17]. Similar fixed point techniques were used by Kelley and Sachs [23], [24], too. In this way, the discretization enters through the solution of the fixed-point equations. It is one advantage of the Hackbusch multigrid idea that any good solver for the partial differential equation is useful for this purpose, independently on how the equation is discretized.

Our main task is to investigate a method close to the infinite-dimensional version of the SQP method. Moreover, the technique should be able to compute optimal controls with a satisfactory precision by using the best available solvers for the PDE.

We rely on the numerical analysis of [14], where the convergence was shown for a simplified n-dimensional model under weak second order assumptions. Numerical examples were presented there for the one–dimensional heat equation with nonlinear boundary condition. Here, we concentrate on the computational aspects which are worked out in more detail. Moreover, we regard examples in a domain $\Omega \subset I\!R^2$ where the dimension of the discretized problems is already very large.

It is evident that the precision of computed optimal controls cannot be better than that for solving the state equation. The (continuous) SQP method will converge quadratically as long as the precision for solving the linear-quadratic subproblems is compatible with the distance of the current iterate to the exact one. Hence the discretization level has to be increased from step to step. This is the point where the dimension becomes soon astronomical. Due to this reason, we are not able to report on a sequence of accuracy 10^{-1}, 10^{-2}, 10^{-4}, 10^{-8}, as the reader might expect.

We focus our attention on a satisfactory graphical accuracy of the computated optimal control. In 2D-domains, this moderate precision leads already to more than $4 \cdot 10^6$ state variables. Undoubtedly, such high accuracy will not be needed in many practical applications. On the other hand, 3D domains will lead to the same large dimensions for a much lower precision. In our test examples, the storage capacity of the computer has not yet been

exhausted. The main reason to avoid a further refinement was the long running time.

For testing the SQP method we consider the optimal boundary control problem to *minimize*

$$\varphi(y,u) = \frac{1}{2}\|y(\cdot,T) - y_T\|^2_{L^2(\Omega)} + \frac{\lambda}{2}\|u\|^2_{L^2(\Sigma)} \tag{1.1}$$

subject to

$$\begin{array}{rcll} y_t(x,t) & = & \Delta_x y(x,t) & \text{in } Q \\ y(x,0) & = & y_I(x) & \text{in } \Omega \\ \partial_\nu y(x,t) & = & b(y(x,t)) + u(x,t) & \text{on } \Sigma \end{array} \tag{1.2}$$

and

$$u_a \le u(x,t) \le u_b \qquad \text{a.e. on } \Sigma. \tag{1.3}$$

In this setting, $\Omega \subset I\!R^n$ is a bounded domain with boundary Γ of class $C^{2,\alpha}$, $(0 < \alpha < 1)$ such that Ω is locally at one side of Γ. We put $Q = \Omega \times (0,T)$, $\Sigma = \Gamma \times (0,T)$; $T > 0$, $\lambda > 0$, $u_a < u_b$ are fixed real numbers, and y_I, $y_T \in C(\overline{\Omega})$ are given functions. By ∂_ν the (outward) normal derivative on Γ is denoted. We assume that $b = b(y)$ belongs to $C^{2,1}(I\!R)$ and is monotone non–increasing. The *control* function $u = u(x,t)$ is looked upon in $L^\infty(\Sigma)$, while the *state* $y = y(x,t)$ is defined as weak solution of (1.2) in $Y = W(0,T) \cap C(\overline{Q})$, where $W(0,T) = \{y \in L^2(0,T;H^1(\Omega)) |\, y_t \in L^2(0,T;H^1(\Omega)')\}$ (cf. Lions and Magenes [27]).

A weak solution y of (1.2) is defined by $y(x,0) = y_0$ and

$$(y_t(t), v)_{(H^1)',H^1} + \int_\Omega \nabla y(t) \nabla v \, dx = \int_\Gamma (b(y(\cdot,t)) + u(\cdot,t)) v \, dS \tag{1.4}$$

for a.e. $t \in (0,T)$ and all $v \in H^1(\Omega)$ (dS: surface measure on Γ).

Let $U_{ad} = \{u \in L^\infty(\Sigma) : u_a \le u(x,t) \le u_b \text{ a. e. on } \Sigma\}$ be the *admissible set.*

2 Necessary and Sufficient Optimality Conditions

First of all we should mention the following result on existence and uniqueness for the state equation (1.2). It is due to Raymond and Zidani [31]:

Theorem 2.1 *For each control $u \in L^\infty(\Sigma)$ the equation (1.2) has a unique weak solution $y \in C(\bar{Q}) \cap W(0,T)$.*

(cf. [31], Theorem 3.1 and Proposition 3.1).

As an immediate conclusion we obtain by standard methods the existence of at least one optimal control, as u is appearing linearly, φ is convex and continuous and U_{ad} is weakly–star compact. However, we do not focus our method on (globally) optimal controls only. The SQP method will converge in a neighbourhood of any locally optimal control, provided that some natural assumptions are satisfied. To make them precise we state at first a set of standard first and second order optimality conditions.

The *first order necessary optimality conditions* for a pair $(\bar{y}, \bar{u})$ consists of the state equation (1.2), the constraint $u \in U_{ad}$, the *adjoint equation*

$$\begin{array}{rcll} -p_t(x,t) & = & \Delta_x p(x,t) & \text{in } Q \\ p(x,T) & = & \bar{y}(x,T) - y_T(x) & \text{in } \Omega \\ \partial_n p(x,t) & = & b'(\bar{y}) p(x,t) & \text{on } \Sigma \end{array} \tag{2.5}$$

and of the *variational inequality*

$$\int_\Sigma (\bar{p}(x,t) + \lambda \bar{u}(x,t))(u(x,t) - \bar{u}(x,t))\, dSdt \geq 0. \tag{2.6}$$

The adjoint equation fits in the theory for the state equation by the transformation $t' := T - t$. Moreover, we shall assume that $(\bar{y}, \bar{u})$ satisfies second order sufficient optimality conditions. Following Dontchev, Hager, Poore and Yang [9] we introduce for arbitrarily small (but fixed) $\sigma > 0$ the set

$$I_\sigma = \{(x,t) \in \Sigma \mid \quad |\lambda \bar{u}(x,t) + \bar{p}(x,t)| \geq \sigma\} \tag{2.7}$$

of sufficiently strong active inequalities.

To formulate associated second order sufficient optimality conditions, we introduce the *Lagrange function* $\mathcal{L}$,

$$\begin{array}{rcl} \mathcal{L}(y,u,p) & = & \varphi(y,u) - \int_Q \{y_t\, p + < \nabla y\,,\, \nabla p >\} dxdt \\ & & + \int_\Sigma p\,(b(y) + u)\, dSdt. \end{array}$$

$\mathcal{L}$ is defined on $Y \times L^\infty(\Sigma) \times Y$. It is twice continuously differentiable w.r. to (y,u) in $Y \times L^\infty(\Sigma)$. Note that this is not true in $W(0,T) \times L^\infty(\Sigma)$. The product $\int_Q y_t\, p\, dxdt$ in the definition of $\mathcal{L}$ is defined in the sense of the pairing between $L^2(0,T; H^1(\Omega)')$ and $L^2(0,T; H^1(\Omega))$. However, this will not be needed in this paper. The second order derivative of $\mathcal{L}$ w.r. to (y,u) is

$$\begin{array}{rcl} \mathcal{L}''(y,u,p)[(y_1,u_1),(y_2,u_2)] & = & \int_\Omega y_1(T) y_2(T)\, dx + \\ & & + \int_\Sigma (\lambda u_1 u_2 + p\, b''(y) y_1 y_2)\, dSdt. \end{array}$$

Second order sufficient optimality conditions for $(\bar{y}, \bar{u}, \bar{p})$ are formulated as follows:
(SSC): There exist $\delta > 0$ and $\sigma > 0$ such that

$$\mathcal{L}''(\bar{y}, \bar{u}, \bar{p})[(y, u), (y, u)] \geq \delta \|u\|^2_{L^2(\Sigma)} \tag{2.8}$$

for all $(y, u) \in W(0, T) \times L^2(\Sigma)$ such that $u = 0$ on I_σ and

$$\begin{aligned} y_t &= \Delta_x y \\ y(0) &= 0 \\ \partial_\nu y &= b'(\bar{y})y + u. \end{aligned} \tag{2.9}$$

Now we assume once and for all that a *reference pair* $(\bar{y}, \bar{u})$ is given, which satisfies together with an associated adjoint state $\bar{p}$ the optimality system and the second order sufficient optimality condition.

It can be shown that $\bar{u}$ is under these assumptions *locally optimal* in the sense of $L^\infty(\Sigma)$ (this can be even proved in $L^p(\Sigma)$ for $p > N + 1$): For a weaker version of second order sufficient conditions we refer to the proof in Goldberg and Tröltzsch [13]. In the case of an elliptic equation of state this is shown in Casas, Tröltzsch and Unger [8]. Their technique can easily be transferred to the parabolic case considered here.

3 The SQP Method

In this section we recall the (continuous) SQP method. Let $w_0 = (y_0, p_0, u_0)$ be a starting triplet (we shall assume that w_0 is close to the reference triplet $\bar{w} = (\bar{y}, \bar{u}, \bar{p})$). Then the Sequential Quadratic Programming (SQP) method determines a sequence $w_k = (y_k, p_k, u_k)$ as follows. Let $\varepsilon > 0$ be given. Initiating from w_k, the next iterate w_{k+1} is obtained from solving the linear–quadratic control problem:
(QP^k) Minimize

$$\varphi'(y_k, u_k)(y - y_k, u - u_k) + \frac{1}{2}\mathcal{L}''(y_k, u_k, p_k)[(y - y_k, u - u_k)]^2 \tag{3.10}$$

subject to

$$\begin{aligned} y_t &= \Delta y && \text{in} \quad Q \\ y(0) &= y_I && \text{in} \quad \Omega \\ \partial_\nu y &= b(y_k) + b'(y_k)(y - y_k) + u && \text{on} \quad \Sigma, \\ u &\in U_{ad}. \end{aligned} \tag{3.11}$$

In (3.10) the expression $[(y-y_k, u-u_k)]^2$ stands for $[(y-y_k, u-u_k), (y-y_k, u-u_k)]$. The solution is (y_{k+1}, u_{k+1}), while the next iterate p_{k+1} of the adjoint state is obtained from the adjoint equation

$$\begin{array}{rcll} -p_t & = & \Delta p & \text{in } Q \\ p(T) & = & y_{k+1}(T) - y_T & \text{in } \Omega \\ \partial_\nu p & = & b'(y_k)p + p_k b''(y_k)(y_{k+1}-y_k) & \text{on } \Sigma. \end{array} \tag{3.12}$$

For convenience we indicate the explicit expressions of φ' and $\mathcal{L}''$:

$$\begin{array}{rcl} \varphi'(y_k,u_k)(y-y_k,u-u_k) & = & \int_\Omega y_k(T)(y(T)-y_k(T))\,dx + \\ & & + \int_\Sigma \lambda\, u_k\,(u-u_k)\,dSdt \\ \mathcal{L}''(y_k,u_k,p_k)[(y-y_k,u-u_k)]^2 & = & \int_\Omega (y(T)-y_k(T))^2\,dx + \\ & & + \int_\Sigma \lambda (u-u_k)^2\,dSdt \\ & & + \int_\Sigma p_k\, b''(y_k)(y-y_k)^2\,dSdt. \end{array}$$

Unfortunately, the linear-quadratic optimal control problem above is not necessarily convex. Our second order sufficient optimality condition imposed on $\bar{w}$ is too weak to guarantee convexity in a neighbourhood of $\bar{u}$ (this is explained below by a simple example adopted from [14]). Therefore, we cannot expect that the SQP method converges locally to $\bar{w}$, unless $\bar{w}$ belongs to a unique global minimum. Note that our method determines global minima of the quadratic sub-problems. This is the reason to restrict the optimization in (3.10)–(3.11) to a neighbourhood U_{ad}^r of the starting element u_o (containing $\bar{u}$ in its interior), where

$$U_{ad}^r = \{u \in U_{ad} \mid \|u-u_o\|_{L^\infty} \le r\}.$$

The necessity of this restriction is illustrated by the example

$$\begin{array}{rcl} min & -x^2 & \\ x & \in & [-2,1]. \end{array} \tag{3.13}$$

This *nonconvex* quadratic problem has stationary solutions at -2, 0, and 1. The points -2, and 1 are strict local minima at which first order sufficient conditions are satisfied. Therefore, (SSC) is trivially fulfilled. Choose $\bar{x}=1$ to be our reference solution. The quadratic sub-problems are identical to (3.13) and will always deliver the global minimum at $x_{n+1}=-2$, independently on how close x_n is taken to $\bar{x}=1$. Convergence to $\bar{x}$ can only be guaranteed by restriction to a neighbourhood of $\bar{x}=1$. We cannot do better in our framework.

The following convergence result can be shown, cf. Goldberg and Tröltzsch [14] for the simplified problem discussed here and to Tröltzsch [35] for a detailed analysis in the case of a general class of nonlinear parabolic control problems. Let $B_\rho(\bar{w})$ denote the open ball around $\bar{w}$ in the natural norm of $C(\bar{Q})^2 \times L^\infty(\Sigma)$.

Theorem 3.2 *Let $\rho > 0$ be sufficiently small and $r := 2\rho$. If the search in (3.10)–(3.11) is restricted to U_{ad}^r, then the SQP method generates for any starting point $w_0 = (y_0, p_0, u_0) \in B_\rho(\bar{w})$ a unique sequence $\{(y_k, p_k, u_k)\}$ such that*

$$\|(y_{k+1}, p_{k+1}, u_{k+1}) - (\bar{y}, \bar{p}, \bar{u})\| \le c_\rho \|(y_k, p_k, u_k) - (\bar{y}, \bar{p}, \bar{u})\|^2, \qquad (3.14)$$

$(k = 0, 1, ...)$, holds with a certain positive constant c_ρ, where $\|\cdot\|$ denotes the norm of $C(\bar{Q})^2 \times L^\infty(\Sigma)$.

Remark: In [14], r corresponds to $2/3\varepsilon$, ρ to $\varepsilon/3$. If the second order sufficient optimality condition is required for all $(x,t) \in \Sigma$, then the quadratic sub-problems are convex and the restriction to U_{ad}^r is not necessary. In our test examples, we did not use U_{ad}^r. A different method of Newton type, presented by Kelley and Sachs [25] for the control of ordinary differential equations, is able to avoid this restriction to a neighbourhood. However, the authors have to impose some structural assumptions on the active set and conditions on the slope of the switching function at the junction points.

This convergence result remains true for a very general class of parabolic control problems. We refer to [35]. However, it is more or less of theoretical value only. Any implementation has to be linked with some discretization. We discretize (QP^k) and solve its discretized version. This result $(y_{k+1}, p_{k+1}, u_{k+1})$ is taken to define $(QP^{(k+1)})$, which is discretized again. In this way the accuracy of the SQP method depends on that for solving the quadratic subproblems. Theoretically, one might increase the level of discretization from step to step in order to maintain quadratic convergence. Without aiming to give a rigorous error analysis for an inexact solution of these problems, we briefly sketch the following estimate:

Let h denote a mesh size parameter describing the discretization of the quadratic subproblems and let w_k^h be the current solution obtained from $(\mathrm{QP}_h^{(k-1)})$, being the discretization of $(\mathrm{QP}^{(k-1)})$ with mesh size h. Let us assume that w_k^h belongs to the region of quadratic convergence to $\bar{w}$. Then

$$\|w_{k+1} - \bar{w}\| \le c\, \|w_k^h - \bar{w}\|^2$$

holds for the *exact* solution w_{k+1} of (QP^k). However, we solve the discretized version of (QP^k) with another mesh size h^+, i.e. we solve $(\mathrm{QP}^k_{h^+})$ (say exactly). Then

$$\begin{aligned} \|w^{h^+}_{k+1} - \bar{w}\| &\leq \|w^{h^+}_{k+1} - w_{k+1}\| + \|w_{k+1} - \bar{w}\| \\ &\leq \alpha(h^+) + c\,\|w^h_k - \bar{w}\|^2, \end{aligned}$$

holds, where $\alpha(h^+)$ is an error estimate for the distance between exact and approximate solution of (QP^k). Recent publications on error estimates show that in many cases $\alpha(h^+) = O(h)$ can be expected. Now let us adapt h^+ according to the rule

$$\alpha(h^+) \leq c\,\|w^h_k - \bar{w}\|^2.$$

Then

$$\|w^{h^+}_{k+1} - \bar{w}\| \leq 2\,c\,\|w^h_k - \bar{w}\|^2$$

is obtained for all steps of the SQP method with the same constant c. This gives a rule for adapting the precision for the solution of the quadratic subproblems: The method continues to converge quadratically as long as the current mesh size h^+ is compatible with the reached accuracy, i.e. if $\alpha(h^+) \leq c\,\|w^h_k - \bar{w}\|^2$. Certainly, this successive refinement leads after a few steps to astronomical dimensions of the discretized problems. We did not try to do this. In this paper, we solve the quadratic subproblems with quite high precision, i.e. a fixed mesh-size leading to results which seemed to us graphically acceptable . In this case, after a few steps the speed of convergence will mainly be limited by $\alpha(h)$.

The following direct method works very well for the solution of the quadratic sub–problems, if the requirement of precision is quite low: Let $h > 0$ stand for a certain mesh size characterizing the discretization of the parabolic PDE and the discretization of the control u. Let the boundary domain Σ be subdivided into m parts Σ_j, $j = 1, \ldots, m$, where $m \in I\!N$. We are looking for the control u as a piecewise constant function

$$u_h = \sum_{j=1}^{m} u_j\, e_j, \tag{3.15}$$

where $u_j \in I\!R$ and e_j is the characteristic function of Σ_j. After having solved the state equation (3.11) for all e_j, $j = 1, ..., m$, the solution y_h associated to u_h is given by superposition. Inserting y_h in the objective functional, a quadratic optimization problem with bound constraints $u_a \leq u_i \leq u_b$, $i = 1, ..., m$, is obtained. If m is not too large (say some hundred), this problem

can be solved efficiently by standard software packages. For instance, we made good experiences with E04NAF (NAG library) developed by Gill and Murray [12]. We refer also to Alt and Mackenroth [4] or Mackenroth [29], who reported on the same positive experience with this technique. For large m, the storage capacity of the computer may soon be exceeded, as C is very large. Moreover, C has to be computed (e.g. the state equation has to be solved for all basic functions e_j, the occuring basic states y_j have to be inserted into the objective functional and the coefficients corresponding to the quadratic parts of $\bar{u}$ are finally the entries of C) and stored in each SQP-step. For that reason, we have decided to choose another approach to solve the quadratic subproblems.

4 A Multigrid Approach

The essential difficulty for solving the linear-quadratic subproblems is not connected with the presence of the control-constraint $u \in U_{ad}$. It appears also in the unconstrained case, where a large-scale backward-forward system of two coupled parabolic equations has to be solved. A way to solve unconstrained optimal control problems was presented by Hackbusch [15]. Let us give a brief sketch of this idea. We consider for simplicity the unconstrained quadratic optimal control problem to *minimize*

$$\frac{1}{2}\|y(T,\cdot) - y_T\|^2_{L^2(\Omega)} + \frac{\lambda}{2}\|u\|^2_{L^2(\Sigma)} \tag{4.16}$$

subject to

$$\begin{array}{rcll} y_t(x,t) & = & \Delta y(x,t) & \text{in } Q \\ y(0,x) & = & y_I(x) & \text{in } \Omega \\ \partial_n y(x,t) & = & u(x,t) - y(x,t) & \text{on } \Sigma. \end{array} \tag{4.17}$$

Suppose that $\bar{u}$ is the optimal solution of problem (4.16)–(4.17). Then the optimal triplet $(\bar{u}, \bar{y}, \bar{p})$ has to fulfill the state equation (4.17) and the first order necessary optimality conditions including the adjoint equation

$$\begin{array}{rcll} -p_t(x,t) & = & \Delta p(x,t) & \text{in } Q \\ p(x,T) & = & y(x,T) - y_T(x) & \text{in } \Omega \\ \partial_n p(x,t) & = & -p(x,t) & \text{on } \Sigma \end{array} \tag{4.18}$$

and

$$u(x,t) = -\frac{p(x,t)}{\lambda} \quad \text{on} \quad \Sigma. \tag{4.19}$$

Introduce now an operator T mapping the control space $U = L^2(\Sigma)$ into itself by

$$(Tu)(x,t) = -\frac{p(x,t)}{\lambda}.$$

Please note that the chain $u \mapsto y \mapsto p \mapsto Tu$ defined by (4.17), (4.18), and (4.19) is behind this construction. The operator T is well defined because state and adjoint equations have unique weak solutions. An optimal solution $\bar{u}$ has to be a fixed point of T,

$$\bar{u} = T\bar{u}. \tag{4.20}$$

It is obvious that T is affine linear

$$Tu = Ku + f,$$

with a compact operator K in $L^2(\Sigma)$ and a fixed $f \in L^2(\Sigma)$. K is of *Fredholm type* with nonnegative kernel. This can be illustrated most easily by the Green's function $G = G(x,\xi,t)$ for (4.17): Then $G \geq 0$ and

$$\begin{aligned}
y(x,t) &= \int_\Omega G(x,\xi,t) y_I(\xi)\, d\xi + \int_0^t \int_\Gamma G(x,\xi,t-s) u(\xi,s)\, dS_\xi ds \\
p(x,t) &= \int_\Omega G(x,\xi,T-t)(y(\xi,T) - y_T(\xi))\, d\xi \\
&= \int_0^T \int_\Omega \int_\Gamma G(x,\xi,T-t) G(\xi,\eta,T-s) u(\eta,s)\, dS_\eta d\xi ds \\
&\quad + \int_\Omega \int_\Omega G(x,\xi,T-t) G(\xi,\eta,T) y_I(\eta)\, d\eta d\xi \\
&\quad - \int_\Omega G(x,\xi,T-t) y_T(\xi)\, d\xi \\
&=: -\lambda(Ku+f).
\end{aligned} \tag{4.21}$$

This property can be shown also in the framework of weak solutions without relying on Green's functions. However, the representation (4.21) shows best the nature of K as integral operator. It stands behind the multigrid strategy to determine $u(x,t)$. Although this cannot be expected in real computations we tacitly assume that $G(x,\xi,t)$ is exactly known. In other words, we assume to solve the PDEs (4.17)–(4.18) exactly. The multigrid strategy refers to a

discretization of u and to an associated collocation method applied to the equation $u = Ku + f$. This means for a fixed grid that Σ is subdivided into $\bar{\Sigma} = \bar{\Sigma}_1 \cup \bar{\Sigma}_2 \cup \ldots \cup \bar{\Sigma}_m$, u_h is taken constant u_j on Σ_j, and $y(x,t)$, $p(x,t)$ are evaluated at prescribed points $(t_j, x_j) \in \Sigma_j$, $j = 1, \ldots, m$ (for instance, at certain "midpoints" of Σ_j) by solving their PDEs exactly (in practice this means solving the PDE by a sufficiently high precision).

The main steps of such a multigrid algorithm, described here for two grids, are well known. We refer, for instance, to Hackbusch [16].

Let $\bar{\Sigma} = \bar{\Sigma}^h_1 \cup \ldots \cup \bar{\Sigma}^h_{m(h)}$ denote the partition of Σ associated to the mesh size parameter h. Then one multigrid–step (**MG**) is performed as follows.

1. Let an iterate u^k_h be given on the fine grid associated to parameter h .
2. Determine the residual $r_h = u^k_h - Tu^k_h$.
3. Reduce the residual to the coarse grid associated to l: $r_l = Rr_h$.
4. Compute a correction by solving $(I-K)d_l = r_l$, where I is the identity.
5. Prolongate the correction to the fine grid, $d_h = Pd_l$.
6. Determine the new iterate $u^{k+1}_h = u^k_h + d_h$.
7. If $\|u^k_h - u^{k+1}_h\|$ is not small enough, then set $u^k_h := u^{k+1}_h$, $k := k+1$, go to 2. Otherwise stop the algorithm.

Here, P and R are some prolongation and restriction operators respectively. $\|u^k_h - u^{k+1}_h\|$ is considered in an appropriate norm, l denotes the mesh size of the coarse grid.

Step 2 requires the application of the operator T. In theory, we get Tu by (exactly) solving the parabolic equation for y, inserting y into the adjoint equation, determining p and using finally the representation (4.19). In real computations, the parabolic equations have to be somehow discretized. Any sufficiently accurate solver can be used for this purpose. We do not consider the difficult error analysis connected with the approximate solution of the partial differential equations. In our further presentation we continue to regard the mapping $u \mapsto Tu$ as given exactly.

The bottleneck of the multigrid method is hidden in establishing the coarse–grid system

$$(I - K)d_l = r_l$$

or its discretized version

$$(I - K_l)\underline{d}_l = \underline{r}_l$$

in step 4. The vectors $\underline{d}_l$, $\underline{r}_l$ representing d_l, r_l belong to the finite–dimensional space $I\!R^{m_l}$. We need the associated matrix representation K_l of K. Let $\{e_1, \ldots, e_{m_l}\}$ be a basis for this space. K_l can be obtained by m_l–times applying the operator T

$$Ke_i = Te_i - f, \qquad i = 1, \ldots, m_l$$

on the coarse grid. This is computationally expensive, as the coarse grid system has to be solved very often during an SQP method. Therefore, we applied the *conjugate gradient method* to the coarse grid system. Here, we need only $K_l \underline{u}_l$ for certain vectors $\underline{u}_l$ occuring in the iteration process. If the iteration stops after a number of steps less than the dimension of $\underline{u}_l$, then the effort occuring in step 4 of the multigrid algorithm decreases.

Another promising approach for handling the optimality system (4.17)–(4.18) is to solve directly the occuring coupled forward–backward system. However, this leads to a system of much higher dimension than in the multigrid method proposed before. Moreover, an effective numerical technique for parabolic backward–forward systems has still to be developed. Therefore, we decided to use the multigrid approach, where the state and adjoint state equations are decoupled.

Remark: In principle, the multigrid strategy is able to handle any mesh-size, which is useful to solve the partial differential equations on the available computer. However, the method is quite slow.

5 Control Constraints

If the pointwise constraints $u \in U_{ad}$ are not imposed on u in (3.11), then the SQP method is nothing else than the known Newton method for solving the equations of the optimality system (consisting of (1.2), (2.5), and $\bar{u} = -\lambda^{-1}\bar{p}$). One step of the method can be performed by the technique described in the preceding section.

Let us now take into account the restrictions (1.3) on the control u. Once again we explain the technique for the simplified linear–quadratic problem (4.16)–(4.17), now with the additional constraint (1.3). We use an projection method due to Bertsekas [7] which was already successfully applied by other authors, for instance by Heinkenschloss and Sachs [19]. Let us first formulate this algorithm (**B**) and then discuss its steps.

1. Denote by $\underline{u}^k = (u_1^k, \ldots, u_m^k)^T$ be the vector representing the iterate

u_h^k (h fixed), fix positive numbers ε and σ and let $I = \{1, \dots, m\}$ be the index set associated to u_h^k.

2. Solve (4.17), (4.18) and get p_h^k with the same discretization as u_h^k. Let $\underline{p}^k = (p_1^k, \dots, p_m^k)^T$ be the representation of p_h^k.

3. Define the sets of strongly active inequalities by $I_a^\varepsilon = \{i \in I : u_i^k = u_a \text{ and } \lambda u_i^k + p_i^k > \varepsilon\}$ and $I_b^\varepsilon = \{i \in I : u_i^k = u_b \text{ and } \lambda u_i^k + p_i^k < -\varepsilon\}$.

4. Set $\hat{u}_i = u_i^k$ for all $i \in I_a^\varepsilon \cup I_b^\varepsilon$.

5. Solve the unconstrained problem (4.16)–(4.17) for u_i^k, $i \in I \setminus (I_a^\varepsilon \cup I_b^\varepsilon)$, while the remaining components of $\underline{u}^k$ are fixed due to 4, denote the solution by v_h^k with vector representation $\underline{v}^k$.

6. Set $u_h^{k+1} = \Pi v_h^k$, where Π is the projection onto $[u_a, u_b]^m$.

7. If $\|u_h^{k+1} - u_h^k\| \geq \sigma$ then put $u_h^k := u_h^{k+1}$, $k := k + 1$ and go to 2. Otherwise stop the iteration.

To illustrate the idea of this technique we consider the first order necessary optimality conditions for problem (4.16), (4.17), (1.3). The optimal triplet $(\bar{u}, \bar{y}, \bar{p})$ has to satisfy the adjoint equation (4.18) together with the variational inequality

$$\int_\Sigma (\bar{u} + \lambda \bar{p})(u - \bar{u})\, dSdt \geq 0 \quad \forall u \in U_{ad}. \tag{5.22}$$

A standard discussion of this inequality shows that

$$\bar{u} = \begin{cases} u_a, & \text{if } \bar{u} + \lambda \bar{p} > 0 \\ u_b, & \text{if } \bar{u} + \lambda \bar{p} < 0 \\ -\dfrac{\bar{p}}{\lambda}, & \text{if } \bar{u} + \lambda \bar{p} = 0 \end{cases} \tag{5.23}$$

(see, for instance, [10]) . These three possible cases for $\bar{u}$ are reflected by step 3 and 4 of the algorithm. If $|u_i^k + \lambda p_i^k| > \varepsilon$, then we can expect that this index i belongs to an active inequality. Therefore, we keep this value u_i^k fixed at the boundary in the next step.

6 Numerical Tests

6.1 The One-Dimensional Case

We have reported on our 1D–computational experience for $\Omega = (0,1)$ in the paper [14]. Let us recall the results for comparison. In our test examples, the control $u = u(t)$ is acting on the right end of $\Omega = (0,1)$. $\Sigma = \{0,1\} \times (0,T)$ splits into 2 parts and $b = 0$, $u = 0$ is kept fixed on the left part $x = 0$. However, we do not need this formal expression of the setting to make the problem comparable to the general problem of section 2. We just formulate the state equation as

$$\begin{array}{rcll} y_t(x,t) & = & y_{xx}(x,t) & \text{in } Q \\ y(0,x) & = & y_I(x) & \text{in } \Omega \\ y_x(t,0) & = & 0 & \text{on } (0,T] \\ y_x(t,1) & = & b(y(t,1)) + u(t) & \text{on } (0,T]. \end{array} \tag{6.24}$$

Let the interval $[0,T]$ be divided by the equidistant grid $0 = t_0 < t_1 < \ldots < t_{n_t} = T$, where n_t is a given integer. Thus, the subdomains Σ_k are given here by the intervals (t_{k-1}, t_k), $k = 1, \ldots, n_t$. The discretization of u is performed according to (3.15). We considered the following test example (going back to Schittkowski [32]).

Example 1: This example is a linear–quadratic control problem of the type discussed in section 5. It is included here to stress that a very fine discretization of the control u and the PDE is needed to compute a sufficiently precise optimal control. "Sufficiently precise" means in this test example that a further refinement of the underlying grid did not change the graphical plot of the control (graphical precision).

We took $T = 1.58$, $\lambda = 0.001$, $y_T = 0.5(1 - x^2)$, $y_I \equiv 0$, $u_a = -1$, $u_b = 1$ and $b(y) = -y$. The state and adjoint equations are solved on $Q = (0,T) \times (0,1)$ by a Crank–Nicholson type finite difference method. Denote by n_t the discretization parameter of u and by n_{y_t}, n_{y_x} the parameters of y (i.e., the discretization with respect to time and space used in the PDE). Optimal controls were determined for the following triplets of (n_t, n_{y_t}, n_{y_x}): $(50, 100, 100)$, $(200, 400, 400)$, $(400, 800, 800)$. The results showed that the mesh-size $(400, 800, 800)$ was necessary to obtain graphically exact controls. A further refinement did not change the computer plot of the optimal control. 5 iterations (**B**) were needed to get the result for the finest discretization. These steps required 56, 34, 30, 14 and 1 multigrid iterations (**MG**), respectively.

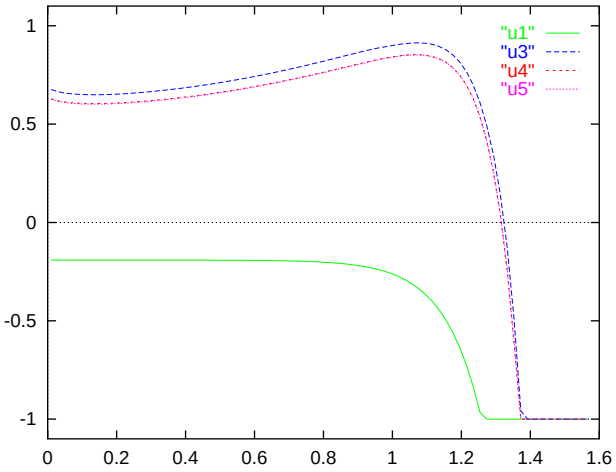

Figure 1: SQP iteration for the nonlinear problem

Example 2: This is a nonlinear problem with almost the same data as above, but with the nonlinear boundary condition

$$y_x(t,1) = u(t) - y(t,1)^2$$

solved by the SQP method. The iteration was started at $(y_0, u_0, p_0) = (0,0,0)$ with mesh size $(n_t, n_{y_t}, n_{y_x}) = (400, 800, 800)$, still quite moderate in view of the experience with example 1. We avoid any table of numbers for the progress of computation. It would pretend a high accuracy, which cannot be justified by the precision of the discretization. Once the SQP method is in the region of quadratic convergence, after a few iterations the reached precision is not compatible with the solution of the PDE. Please note that the mesh size 400 for the control will at most ensure a precision of the order 10^{-2} for the linear–quadratic sub–problems. This is just graphical precision. In Figure 1 some iterates are represented.

6.2 The Two-Dimensional Case

Next, we consider our problem (4.16)–(4.17) in $\Omega = (0,1) \times (0,1) \subset I\!R^2$. The control is acting on $\Gamma_1 = \{(x_1, x_2) \in \bar{\Omega} : x_2 = 1\}$, define $\Gamma_2 = \Gamma \setminus \Gamma_1$.

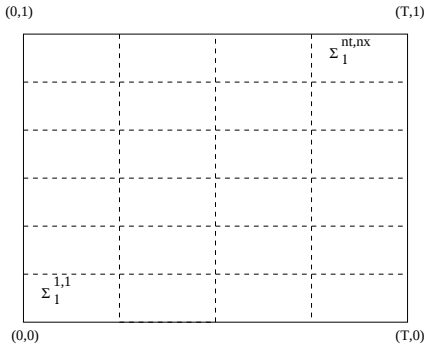

Figure 2: Domain of $\boldsymbol{u_h}$

The boundary condition is slightly changed,

$$\begin{array}{lcll} \partial_\nu y & = & b(y)+u & \text{on} \quad \Gamma_1 \\ \partial_\nu y & = & -y & \text{on} \quad \Gamma_2. \end{array} \tag{6.25}$$

As before, the intervals $[0,T]$ and $[0,1]$ are split into equidistant subintervals, $0 = t_0 < t_1 < \ldots < t_{n_t} = T$, $0 = x_0 < x_1 < \ldots < x_{n_x} = 1$, where n_t and n_x are positive integers. We split the control domain $\Sigma_1 = \Gamma_1 \times (0,T)$ for the control u into subdomains $\Sigma_1^{ki} = (t_{k-1}, t_k) \times (x_{i-1}, x_i)$, $k = 0, \ldots, n_t$, $i = 0, \ldots, n_x$. The partition of Σ_1 is shown in Figure 2.

To solve the underlying parabolic differential equations, the domain Ω was divided into equidistant subdomains $\Omega_{ij} = \{(x_{i-1}^1, x_i^1) \times (x_{j-1}^2, x_j^2) : i = 0, \ldots, n_{x^1}, j = 0, \ldots, n_{x^2}\}$, where n_{x^i} is the number of equidistant subintervals of $[0,1]$ in direction x_i, $i = 1,2$. The time interval $[0,T]$ is split into n_τ equidistant subparts. We put $h_\tau = \frac{T}{n_\tau}$, $h_{x^1} = \frac{1}{n_{x^1}}$ and $h_{x^2} = \frac{1}{n_{x^2}}$.

Remark: The number of subintervals of $[0,1]$ in x_i–direction was related to the number of subintervals of $[0,1]$ in x–direction for u_h: $n_{x^i} = Mn_x$, $i = 1,2$, where $M \in I\!N$. Analogously, $n_\tau = Mn_t$. We used $M = 2$ in the computations. Owing to the simple geometry of Ω, a finite difference splitting-up method was selected, since it was faster than available finite element codes.

The examples below are computed with the Bertsekas projection method and the multigrid method for the unconstrained quadratic subproblems. In the multigrid algorithm, the coarse–grid systems of linear equations were solved by a *conjugate gradient* method.

Example 3: This is a convex linear–quadratic control problem, used to compare the precision of our computations for a single linear-quadratic subproblem with known results. We take $T = 1$, $\lambda = 0.001$, $y_I \equiv 0$, $y_T = 0.5x_1x_2 + 0.25$, $b(y) = -y$, $u_a = 0$, $u_b = 1$. This problem was considered by Mackenroth [29] with $\lambda = 0$. Therefore, the result was a control of bang–bang type. Our control is close to that of Mackenroth, but continuous, as $\lambda > 0$. Further parameters are $n_t = 10$, $n_x = 10$, $M = 2$ and $n_{x^2} = n_{x^1}$, thus we have

mesh size (n_x, n_t) for the control on the coarsest grid:	$(5, 5)$,
mesh size (n_x, n_t) on the finest grid:	$(10, 10)$,
mesh size (n_{x^1}, n_{x^2}, n_t) for the state:	$(20, 20, 20)$,
maximal number of state variables:	8000.

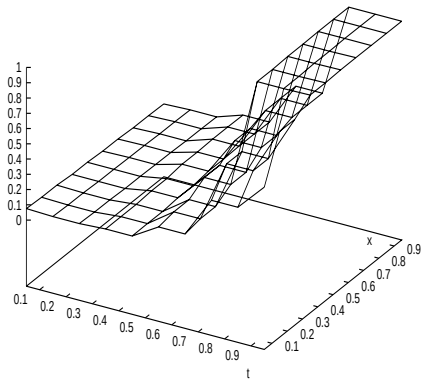

Figure 3: Optimal control of example 3

The result is shown in Figure 3. The approach considered here allows to solve the problem with a finer discretization.

Example 4: Regard the same problem as above with finer discretization: $n_t = 80$, $n_x = 80$, $M = 2$ and $n_{x^2} = n_{x^1} = 160$. Here we have

mesh size (n_x, n_t) for the control (coarsest grid):	$(5, 5)$,
mesh size (n_x, n_t) (finest grid):	$(80, 80)$,
mesh size (n_{x^1}, n_{x^2}, n_t) for the state:	$(160, 160, 160)$,
maximal number of state variables:	4.096.000.

7 iterations (**B**) were needed to get the result for the finest discretization, requiring 125, 67, 45, 32, 23, 13, and 1 multigrid steps (**MG**), respectively. The result is shown in Figure 4.

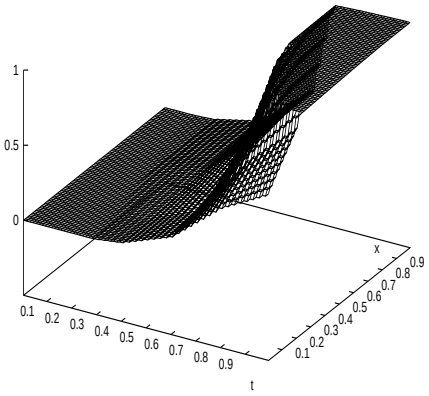

Figure 4: Optimal control of example 4

Finally, we discuss a nonlinear test example solved by the SQP method.

Example 5: Here, the boundary condition is replaced by a non–linear one. We put $T = 1$, $\lambda = 0.001$, $y_I \equiv 0$, $y_T = 0.5x_1x_2 + 0.25$, $b(y) = -y^2$, $u_a = 0$, $u_b = 0.2$, $n_t = 80$, $n_x = 80$, $M = 2$, and $n_{x^2} = n_{x^1} = 160$. Initial iterate of the SQP algorithm was the triplet $(0, 0, 0)$. The progress of iteration is shown in Figures 5 – 9. Our initial iterate was outside the convergence region of the SQP method. This region was hit by chance in the third step by u_3. A usual globalization technique would avoid this behaviour.

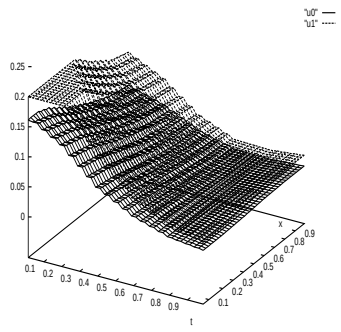

Figure 5: Example 5, $\boldsymbol{u_0}$ and $\boldsymbol{u_1}$

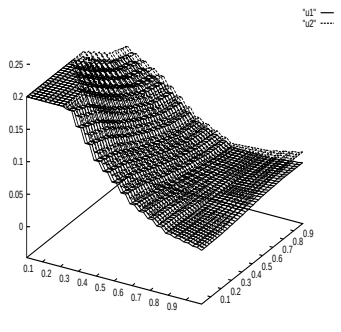

Figure 6: Example 5, $\boldsymbol{u_1}$ and $\boldsymbol{u_2}$

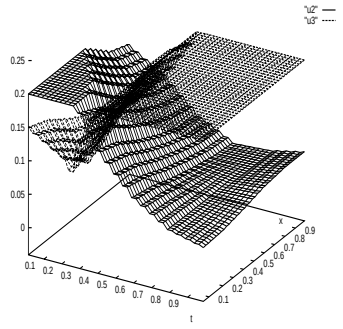

Figure 7: Example 5, $\boldsymbol{u_2}$ and $\boldsymbol{u_3}$

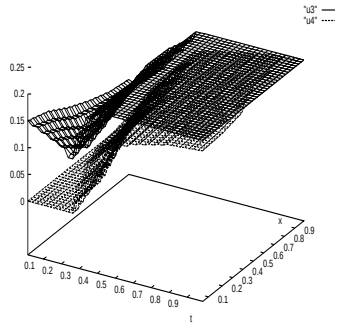

Figure 8: Example 5, $\boldsymbol{u_3}$ and $\boldsymbol{u_4}$

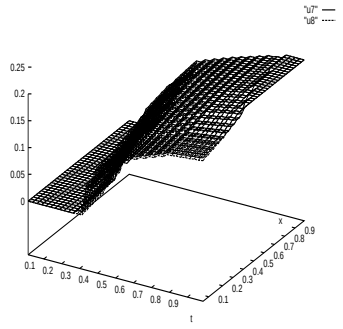

Figure 9: Example 5, $\boldsymbol{u_7}$ and $\boldsymbol{u_8}$

We observed the same fast convergence for a problem containing a nonlinear boundary condition of **Stefan Boltzmann** type

$$\partial_\nu y = u - y^4 \text{ on } \Gamma_1.$$

7 Final Comment

The method presented in this paper is not yet effective. SQP method, Bertsekas projection method and multigrid technique form a chain of 3 nested iteration schemes. Although the SQP method itself exhibits the expected fast convergence, the other inherent iteration procedures are slower and lead to long running times. Certainly, this long time is mainly connected with the high precision of the computation. The same effect was mentioned by Gill and others [11] who applied the software system DASOPT to solving

a parabolic optimal control problem for a quasilinear heat equation in a two-dimensional rectangular domain Ω with a moderate discretization. We confirmed their estimate $(n_{x_1} \cdot n_{x_2})^2$ for the order of the computational time in our own tests.

Nevertheless, it is obvious that our procedure is not optimal and can be improved, for instance by using an appropriate pre-conditioning in the multigrid steps. This was not our primary intention. We aimed to complete the theory of the standard (continuous) SQP method for parabolic control problems by associated numerical test examples. The method presented here is very close to the one discussed in our convergence analysis. Owing to its fast convergence, there was a need to solve the quadratic subproblems with comparably high precision. This paper shows one way leading to acceptable results.

References

[1] Alt, W. (1990), "The Lagrange-Newton method for infinite-dimensional optimization problems," *Numer. Funct. Anal. and Optimiz. 11*, 201–224.

[2] Alt, W. (1994), "The Lagrange Newton method for infinite-dimensional optimization problems," *Control and Cybernetics 23*, 87–106.

[3] Alt, W., "Discretization and mesh independence of Newton's method for generalized equations."

[4] Alt, W. and Mackenroth, U. (1992), "On the numerical solution of state-constrained coercive parabolic optimal control problems," In *Optimal Control of Partial Differential Equations*, Volume 68 of *Lecture Notes ISNM*, pages 44–62, Birkhäuser Verlag. Hoffmann, K.H. and Krabs, W (eds.).

[5] Alt, W. and Malanowski, K. (1993), "The Lagrange-Newton method for nonlinear optimal control problems," *Comp. Opt. and Applic. 2*, 77–100.

[6] Alt, W., Sontag, R. and Tröltzsch, F. (1996), "An SQP Method for Optimal Control of a Weakly Singular Hammerstein Integral Equation," *Appl. Math. Opt. 33*, 227–252.

[7] Bertsekas, D.B. (1982), "Projected Newton methods for optimization problems with simple constraints," *SIAM J. Control and Optimization 20*, 221–246.

[8] Casas, E., Tröltzsch, F. and Unger,A (1996), "Second order sufficient optimality conditions for a nonlinear elliptic control problem," *J. of Analysis and Applications (ZAA) 15*, 687–707.

[9] Dontchev, A.L., Hager, W.W., Poore, A.B., and Yang, B. (1995), "Optimality, stability, and convergence in optimal control," *Appl. Math. Optim. 31*, 297–326.

[10] Eppler, K. and Tröltzsch, F. (1986), "On switching points of optimal controls for coercive parabolic boundary control problems," *Optimization 17*, 93–101.

[11] Gill, P.E., Park, K., Petzold, L., Rosen, J.B., and Jay, L.O. "Numerical optimal control of parabolic PDEs using DASOPT," Report NA 96-1, Dept. of Math., University of California, San Diego.

[12] Gill, P.E., Murray, W., Saunders, M.A., and Wright, M.H. (1982), "Users's guide for SOL/QPSOL: A Fortran package for quadratic programming," Technical Report No. 82–7, Department of Operations Research, Stanford University, Stanford, CA 94305.

[13] Goldberg, H. and Tröltzsch, F. (1993), "Second order sufficient optimality conditions for a class of nonlinear parabolic boundary control problems," *SIAM J. Control and Optimization 31*, 1007–1025.

[14] Goldberg, H. and Tröltzsch, F., "On a Lagrange-Newton method for a nonlinear parabolic boundary control problem," Preprint TU-Chemnitz, Fak. f. Mathematik, Report 96–8, to appear.

[15] Hackbusch, W. (1979), "On the fast solving of parabolic boundary control problems," *SIAM J. Control and Optimization 17, No. 2*, 231–244.

[16] Hackbusch, W. (1989), *Integralgleichungen*. B. G. Teubner, Stuttgart.

[17] Hackbusch, W. and Will, Th. (1983), "A numerical method for a parabolic bang–bang problem," *Control and Cybernetics 12, No. 3–4*, 100–116.

[18] Heinkenschloss, M., "The numerical solution of a control problem governed by a phase field model," *Optimization Methods and Software 7*, 1997, to appear.

[19] Heinkenschloss, M and Sachs, E.W. (1994), "Numerical solution of a constrained control problem for a phase field model," *Control and Estimation of Distributed Parameter Systems, Int. Ser. Num. Math. 118*, 171–188.

[20] Heinkenschloss, M. and Tröltzsch, F., "Analysis of the Lagrange-SQP-Newton method for the control of a Phase field equation," Virginia Polytechnic Institute and State Unversity, ICAM Report 95-03-01.

[21] Ito, K. and Kunisch, K. (1996), "Augmented Lagrangian-SQP methods for nonlinear optimal control problems of tracking type," *SIAM J. Control Optim. 34*, 874–891.

[22] Ito, K. and Kunisch, K. (1996), "Augmented Lagrangian-SQP methods in Hilbert spaces and application to control in the coefficients problems," *SIAM J. Optim. 6*, 96–125.

[23] Kelley, C.T. and Sachs, E.W. (1991), "Fast algorithms for compact fixed point problems with inexact function evaluations," *SIAM J. Scientific and Stat. Computing 12*, 725–742.

[24] Kelley, C.T. and Sachs, E.W. (1994), "Multilevel algorithms for constrained compact fixed point problems," *SIAM J. Scientific and Stat. Computing 15*, 645–667.

[25] Kelley, C.T. and Sachs, E.W. (1995), "Solution of optimal control problems by a pointwise projected Newton method," *SIAM J. Contr. Optimization 33*, 1731–1757.

[26] Kupfer, S.F. and Sachs, E.W. (1992), "Numerical solution of a nonlinear parabolic control problem by a reduced SQP method," *Computational Optimization and Applications 1*, 113–135.

[27] Lions, J.L. and Magenes, E. (1968), *Problèmes aux limites non homogènes et applications*, Volume 1–3. Dunod, Paris.

[28] Machielsen, K. (1987), "Numerical solution of optimal control problems with state constraints by sequential quadratic programming in function space," CWI Tract, 53, Amsterdam.

[29] Mackenroth, U. (1987), "Numerical solution of some parbolic boundary control problems by finite elements," In *Control Problems for Systems Described by Partial Differential Equations and Applications*, Volume 97 of *Lecture Notes Contr. Inf. Sci.*, pp 325–335, Berlin, Springer–Verlag.

[30] Maurer, H. (1981), "First and second order sufficient optimality conditions in mathematical programming and optimal control," *Math. Programming Study 14*, 163–177.

[31] Raymond, J.P. and Zidani, H., "Hamiltonian Pontryagin's principles for control problems governed by semilinear parabolic equations," Preprint 1995, to appear.

[32] Schittkowski, K. (1979), "Numerical solution of a time-optimal parabolic boundary-value control problem," *JOTA 27*, 271–290.

[33] Tröltzsch, F. (1994), "Convergence of an SQP–Method for a class of nonlinear parabolic boundary control problems," In W. Desch, F. Kappel, K. Kunisch, eds., *Control and Estimation of Distributed Parameter Systems. Nonlinear Phenomena,* Int. Series of Num. Mathematics, Vol. 118, Birkhäuser Verlag, Basel, pp. 343–358.

[34] Tröltzsch, F. (1994), "An SQP method for the optimal control of a nonlinear heat equation," *Control and Cybernetics, 23(1/2)*, 267–288.

[35] Tröltzsch, F., "On the Lagrange–Newton method for the optimal control of semilinear parabolic equations," 1997, to appear.