Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

UWE PRÜFERT FREDI TRÖLTZSCH MARTIN WEISER

The convergence of an interior point method for an elliptic control problem with mixed control-state constraints.[1]

[1]Supported by the DFG Research Center MATHEON "Mathematics for key technologies" in Berlin.

ZIB-Report 04-47 (November 2004)

The convergence of an interior point method for an elliptic control problem with mixed control-state constraints.[†]

Uwe Prüfert[*] Fredi Tröltzsch[*] Martin Weiser

December 1, 2004

Abstract

The paper addresses primal interior point method for state constrained PDE optimal control problems. By a Lavrentiev regularization, the state constraint is transformed to a mixed control-state constraint with bounded Lagrange multiplier. Existence and convergence of the central path are established, and linear convergence of a short-step pathfollowing method is shown. The behaviour of the regularizations are demonstrated by numerical examples.

AMS MSC 2000: 90C51, 49J20, 65M15

Keywords: interior point methods in function space, optimal control, mixed control-state constraints, Lavrentiev regularization

1 Introduction

The application of interior point methods to optimal control problems has received a good deal of interest in the past years. This parallels the fast development of numerical methods in large scale optimization where interior point methods play an important role. In the context of PDE control, their performance was carefully tested by Haddoux et al. [5] for discretized versions of elliptic control problems. Similarly, Grund and Rösch [4] considered different codes of interior point methods for elliptic control problems with pointwise state-constraints. Trust-region interior point techniques have been considered by M. Ulbrich, S. Ulbrich and M. Heinkenschloss in [11] for the optimal control of semilinear parabolic equations in a function space setting. Moreover, affine-scaling interior-point methods were presented for semilinear parabolic boundary control in [10].

In [13, 12] primal-dual interior point methods in the infinite dimensional function space setting for ODE problems have been analyzed and their computational realization by inexact pathfollowing methods has been suggested. In [14] this method has been enhanced on the control of elliptic PDE problems with control constrains.

A satisfactory convergence theory, however, had only been obtained for control constraints, whereas results for state constraints are scarce. The difficulty arises from the fact that Lagrange multipliers for state constraints are usually only measures, which hampers theoretical convergence analysis and affects the numerical solution.

[†]Supported by the DFG Research Center MATHEON "Mathematics for key technologies" in Berlin.

[*]Institut für Mathematik, Technische Universität Berlin, Straße des 17. Juni 136, D-10623, Germany

As concerns the regularity of Lagrange multipliers, the situation changes for mixed control-state constraints such as constraints of bottleneck type. Under natural assumptions, their multipliers can shown to be functions in certain L^p-spaces, we only mention [9, 2, 1]. In [6], the idea came up to add a tiny fraction of the control to the state constraint such that a mixed control-state constraint results. The Lagrange multiplier to this mixed constraint is a bounded and measurable function. This Lavrentiev-regularization for state constraints has been analyzed in the context of primal-dual active set methods for elliptic control problems.

In the current paper, both ideas are combined. We analyze a primal interior point method applied to a Lavrentiev regularized state constrained optimal control problem defined in §2. We show existence and convergence of the central path defined by the interior point method in §3 and §4, respectively. In §5, we turn to the linear convergence of an implementable short-step pathfollowing method. The paper is concluded with a set of numerical examples in §6.

2 Problem setting

In this paper we consider the optimal control problem

$$\text{(P)} \qquad \min J(y,u) := \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} \tag{1}$$

subject to the elliptic boundary value problem

$$\begin{aligned} Ay &= u && \text{in } \Omega && (2)\\ \partial_n y + \alpha y &= 0 && \text{on } \Gamma && (3) \end{aligned}$$

and to the pointwise mixed control-state constraints

$$y + \lambda u \geq y_c \qquad \text{a.e. in } \Omega. \tag{4}$$

In this setting, $\Omega \subset \mathbb{R}^N$, $N \in \{2,3\}$, is a bounded domain with $C^{0,1}-$boundary Γ, y_c, $y_d \in L^\infty(\Omega)$ and $\alpha \in L^\infty(\Gamma)$ are fixed functions, and ν, $\lambda \in \mathbb{R}$, $\lambda > 0$, are given constants. By A we denote the differential operator

$$(Ay)(x) = -\sum_{i,j=1}^{N} \frac{\partial}{\partial x_i}\left(a_{ij}(x)\frac{\partial}{\partial x_j}y(x)\right) + c_0(x)y(x)$$

with coefficients $a_{ij} \in C^{1,1}(\Omega)$, $c_0 \in L^\infty(\Omega)$ satisfying $a_{ij}(x) = a_{ji}(x)$ and the condition of uniform ellipticity

$$\sum_{i,j=1}^{N} a_{ij}(x)\xi_i\xi_j \geq \delta|\xi|^2 \quad \forall \xi \in \mathbb{R}^N.$$

Moreover, we require $c_0(x) \geq 0, \alpha(x) \geq 0$ and assume that one of these two functions is not vanishing identically. We refer to problem (1)–(4) as problem (P). Let us introduce the following

Notations. By $\|\cdot\| = \|\cdot\|_{L^2(\Omega)}$ and $(\cdot,\cdot)$ we denote the natural norm and the associated inner product of $L^2(\Omega)$, respectively. We use $\|A\|_{L^p \to L^q}$ to denote the norm of a linear continuous operator $A : L^p(\Omega) \to L^q(\Omega)$. In the case $p = q = 2$, this norm is just denoted by $\|A\|$. For $\|A\|_{L^p \to L^p}$ we write $\|A\|_{L^p}$. Throughout the paper, c is a generic constant.

Moreover we write L^p for $L^p(\Omega)$ to shorten the notation. If no confusion is possible, we write $S+v$ instead of $S+vI$, although S is an operator and v a function.

If $v \in L^2(\Omega)$ is a given function, then $v \le 0$ means $v(x) \le 0$ for a.a. $x \in \Omega$. By ∂_n the co-normal derivative

$$\partial_n u = \sum_{ij=1}^{N} n_i a_{ij} D_j u$$

is denoted.

The main scope of our paper is to discuss the convergence of the standard interior point method for the problem (P). The simplest and well known idea of introducing this method is the elimination of the mixed control-state constraint $y+\lambda u \ge y_c$ by a logarithmic barrier function. We substitute (P) by the problem

$$(\mathrm{P}_\mu) \qquad \min J_\mu(y,u) := \frac{1}{2}\|y-y_d\|^2 + \frac{\nu}{2}\|u\|^2 - \mu \int_\Omega \ln\left((y+\lambda u - y_c)(x)\right) dx \tag{5}$$

subject to

$$Ay = u \qquad \text{in} \quad \Omega \tag{6}$$
$$\partial_n y + \alpha y = 0 \qquad \text{on} \quad \Gamma \tag{7}$$

with $u \in L^2$.

In our analysis, we transform the state-constrained problem (P) to the problem (11)–(12) with control constraints. We have two reasons for: The analysis of this transformed problem is simpler than that for (P), since we are able to prove the necessary regularity of Lagrange multipliers. Moreover, it is easier to show the existence of the central path for (11)– (12).

3 Existence of the central path

In this section we establish the existence of unique minima v_μ of (P_μ) for all $\mu > 0$. We refer to the mapping $\mu \mapsto v_\mu$ as the central path, even though continuity is proved only in Section 4. First we recall some known facts about the state-equation (2)–(3).

Theorem 3.1. *Under our assumptions, for all $u \in L^r(\Omega)$ with $r > \frac{N}{2}$, equation* (2) *has a unique solution $y \in H^1(\Omega) \cap C(\bar\Omega)$. There is a constant $c(\Omega, r)$ such that*

$$\|y\|_{H^1(\Omega)} + \|y\|_{C(\bar\Omega)} \le c\,\|u\|_{L^r(\Omega)}.$$

The theorem was shown by Casas [3]. It ensures that, for $N \le 3$, the mapping $G : u \mapsto y$ is continuous from L^2 to $H^1(\Omega) \cap C(\bar\Omega)$. In particular, it is continuous in L^2. We denote the associated mapping by $S = EG$, where $E : H^1(\Omega) \to L^2$ is the embedding operator from $H^1 \cap C(\bar\Omega)$ in L^2. Therefore, we have $S : L^2 \to L^2$, continuously.

By S, problem (P) becomes equivalent to

$$\min \frac{1}{2}\|Su - y_d\|^2 + \frac{\nu}{2}\|u\|^2 \tag{8}$$

subject to

$$\lambda u + Su - y_c \ge 0 \quad \text{a.e. in } \Omega. \tag{9}$$

REMARK. S is known to be a compact operator. By $\lambda > 0$, $-\lambda$ is not an eigenvalue of S, see the discussion below.

To transform (8)–(9) into a control-constrained problem, we substitute

$$v := Su + \lambda u.$$

By our assumption,

$$D := (S + \lambda I)^{-1} \tag{10}$$

exists as a continuous operator in L^2. In fact, since $\lambda > 0$, we have $\lambda u + Su = 0 \Leftrightarrow \lambda u + y = 0 \Leftrightarrow u = -\frac{1}{\lambda} y$. This means $Ay = -\frac{1}{\lambda} y$, hence $Ay + \lambda u = 0$ and $\partial_n y + \alpha y = 0$. Clearly, by coercivity this equation has only the trivial solution. After this substitution, (8)–(9) is equivalent to

$$\min f(v) := \frac{1}{2}\|SDv - y_d\|^2 + \frac{\nu}{2}\|Dv\|^2 \tag{11}$$

subject to

$$v - y_c \geq 0, \tag{12}$$

where $v \in L^2$. This is a control-constrained problem for the new control v that is interesting in itself. For the special choice $D = I$ our analysis covers problems with simple bounds on the control $v = u$. The interior point method for (11) and (12) is equivalent to solving

$$\min \left\{ f(v) - \mu \int_\Omega \ln\left(v(x) - y_c(x)\right) dx \right\}. \tag{13}$$

Obviously, the quadratic functional f is continuously differentiable in L^2. Its derivative is given by

$$f'(v_\varepsilon)\, v = (\tilde{p} + \nu D^* D v_\varepsilon \,,\, v)$$

with $\tilde{p} = D^* S^* (SDv - y_d)$. Here, S^*, $D^* : L^2 \to L^2$ are the Hilbert space adjoints to S, D, respectively. If $v_\varepsilon(x) - y_c(x) \geq \varepsilon > 0$ holds a.e. on Ω, then the functional

$$\phi(v) = \mu \int_\Omega \ln\left(v(x) - y_c(x)\right) dx$$

is differentiable at $v_\varepsilon \in L^2$ in any direction $v = \tilde{v} - v_\varepsilon$, where $\tilde{v}(x) - y_c(x) \geq \varepsilon$ a.e. in Ω. Moreover, it is differentiable in any direction $h \in L^\infty(\Omega)$, since $v + t\,h - y_c \geq \varepsilon/2$ for sufficiently small t.

Suppose now that (11)–(12) admits a solution $v_\varepsilon = v_\varepsilon(\mu) \in L^2$ satisfying $v_\varepsilon(x) - y_c(x) \geq \varepsilon > 0$. Then we get from the differentiability properties mentioned above

$$f'(v_\varepsilon) - \phi'(v_\varepsilon) = 0, \tag{14}$$

since in this case v_ε belongs to the L^∞-interior of the admissible set. Therefore, it holds

$$\tilde{p} + \nu D^* D v_\varepsilon - \frac{\mu}{v - y_c} \quad = \quad 0 \quad \text{a.e. in } \Omega.$$

Define $\eta \in L^\infty(\Omega)$ by

$$\eta(x) := \frac{\mu}{v(x) - y_c(x)}. \tag{15}$$

Then we have $\eta \geq 0$, $v_\varepsilon - y_c \geq 0$ and $\eta(v_\varepsilon - y_c) = \mu$ for almost all $x \in \Omega$. This function η will tend to a Lagrange multiplier for (P) as $\mu \downarrow 0$. However, we have to show that (11)–(12) is solvable, i.e. that the central path exists.

To verify this, we consider for fixed $\mu > 0$, $\varepsilon > 0$ the auxiliary problem

$$(\mathrm{P}_\mu^\varepsilon) \qquad \min_{v(x)-y_c(x)\geq\varepsilon} f_\mu(v) = f(v) - \mu \int_\Omega \ln\left(v(x) - y_c(x)\right) dx$$

where $v \in L^2$. We first prove that this problem is solvable. Next we show that the solution is not active for all sufficiently small $\varepsilon > 0$. In this way, finally a solution of (P_μ) is found.

Lemma 3.2. *For all $\mu \geq 0$, it holds that $f_\mu(v) \to \infty$ if $\|v\|_{L^2} \to \infty$ and $v(x) \geq y_c(x) + \varepsilon$.*

Proof. Since $\|v\| = \|D^{-1}Dv\| \leq \|S + \lambda I\| \, \|Dv\|$, we have

$$\begin{aligned} f_\mu(v) &= \frac{1}{2}\,\|SDv - y_d\|^2 + \frac{\nu}{2}\,\|Dv\|^2 - \mu \int_\Omega \ln\left(v - y_c\right) dx \\ &\geq \frac{\nu}{2}\,\|Dv\|^2 - \mu \int_\Omega (v - y_c)\, dx \\ &\geq \frac{\nu\delta_0}{2}\,\|v\|^2 - \mu\,\|v - y_c\|_{L^1} \geq \frac{\nu\delta_0}{2}\,\|v\|^2 - \mu\, c\,\|v - y_c\| \end{aligned} \tag{16}$$

with $\delta_0 = \|S + \lambda I\|^{-1} > 0$. Obviously, $\|v\| \to \infty$ implies $f_\mu(v) \to \infty$. □

Theorem 3.3. *For all $\mu \geq 0$ and $0 < \varepsilon \leq 1$, problem $(\mathrm{P}_\mu^\varepsilon)$ has a unique solution $v_\varepsilon(\mu)$. There is a constant $c_v < \infty$ independent of μ and ε such that $\|v_\varepsilon(\mu)\| \leq c_v$.*

Proof. Obviously, f_μ is convex and continuous on the convex and closed subset $C_\varepsilon \subset L^2$,

$$C_\varepsilon = \left\{v \in L_2(\Omega) \,|\, v(x) - y_c(x) \geq \varepsilon > 0 \text{ for a.a. } x \in \Omega\right\}.$$

Therefore, f_μ is lower semicontinuous on C_ε. Lemma 3.2 yields the existence of $c_v > 0$ such that all $v \in C_\varepsilon$ with $\|v\| > c_v$ can be neglected for the search of the infinimum of f_μ: We take $\tilde{v} := y_c + 1$, then the logarithmic term vanishes and

$$f_\mu(v) \geq f_\mu(y_c + 1) = \frac{1}{2}\,\|SD\tilde{v} - y_d\|^2 + \frac{\nu}{2}\,\|D\tilde{v}\|^2$$

for all sufficiently large $\|v\|$. On $C_\varepsilon \cap \left\{v \in L^2 \,|\, \|v\| \leq c_v\right\}$, the functional f_μ is bounded, hence

$$j(\varepsilon) := \inf_{v \in C_\varepsilon} f_\mu(v)$$

if finite. Let $v_n \in C_\varepsilon$, $\|v_n\| \leq c_v$, be an infimal sequence, i.e. $f_\mu(v_n) \to j$ for $n \to \infty$.

We can assume w.l.o.g. weak convergence in L^2, $v_n \rightharpoonup v_\varepsilon \in C_\varepsilon$. By lower semicontinuity, a standard argument yields

$$f_\mu(v_\varepsilon) = j,$$

hence v_ε is the solution of $(\mathrm{P}_\mu^\varepsilon)$. □

We recall problem $(\mathrm{P}_\mu^\varepsilon)$,

$$\min f_\mu(v) := \frac{1}{2}\|SDv - y_d\|^2 + \frac{\nu}{2}\|Dv\|^2 - \mu \int_\Omega \ln\left(v - y_c\right) dx$$
$$v(x) - y_c(x) \geq \varepsilon \quad \text{a.e. in } \Omega.$$

As in the theorem above, we denote the solution of $(\mathrm{P}_\mu^\varepsilon)$ by v_ε, since μ is taken fixed for a while. Take any other $v \in C_\varepsilon$ and $t \in [0,1]$. Then $v_\varepsilon + t(v - v_\varepsilon) \in C_\varepsilon$, hence $f_\mu(v_\varepsilon + t(v - v_\varepsilon))$ is defined. Note that f_μ is not Gâteaux-differentiable in L^2, since $f_\mu(v_\varepsilon + ht)$ may be undefined for $h \in L^2$. However, it is directionally differentiable in the direction $v - v_\varepsilon$. From

$$0 \le \frac{f_\mu(v_\varepsilon + t(v - v_\varepsilon)) - f_\mu(v_\varepsilon)}{t}$$

we find by $t \downarrow 0$ for the directional derivative

$$f_\mu'(v_\varepsilon)(v - v_\varepsilon) \ge 0 \quad \forall v \in C_\epsilon.$$

In terms of our transformation, this can be written as

$$\Big(D^* S^*(SDv_\varepsilon - y_d) + \nu D^* D v_\varepsilon - \frac{\mu}{v_\varepsilon - y_c}\,,\, v - v_\varepsilon\Big) \ge 0 \quad \forall v \in C_\varepsilon. \tag{17}$$

Define $p_\varepsilon := D^* S^*(SDv_\varepsilon - y_d)$. Then we can re-write (17) as

$$\Big(p_\varepsilon + \nu D^* D v_\varepsilon - \frac{\mu}{v_\varepsilon - y_c}\,,\, v - v_\varepsilon\Big) \ge 0 \quad \forall v \in C_\varepsilon. \tag{18}$$

We shall show that $\|p_\varepsilon\|_\infty$ is bounded, independently of ε: The operator S is known to be self-adjoint, $S = S^*$. Moreover, as $S = EG$, S is even linear and continuous from L^2 to L^∞. The same holds for S^*.

Let us discuss the form and the regularity properties of the operator D. We have $D = (S + \lambda I)^{-1}$. Put $w = Dz$. Then $z = Sw + \lambda I w$. It follows $\lambda w = z - Sw = z - SDz$ and $w = \lambda^{-1} z - \lambda^{-1} SDz$. Therefore D admits the form

$$D = \lambda^{-1}(I - SD). \tag{19}$$

From this representation we get the additional regularity property $D : L^\infty \to L^\infty$, continuously. This follows from $D : L^2 \to L^2$ and $S : L^2 \to L^\infty$. Moreover, we have $D^* = (\lambda I + S^*)^{-1}$. With the same argument, $D^* = \lambda^{-1}(I - S^* D^*)$, hence also $D^* : L^\infty \to L^\infty$ since $S^* = S : L^2 \to L^\infty$ as well. Notice, that S and D commute, S^* and D^* as well.

We know from Lemma 3.2 that $\|v_\varepsilon\|$ is bounded by a constant c_v that does not depend on ε. Now we estimate $\|p_\varepsilon\|_\infty$ by

$$\begin{aligned} \|p_\varepsilon\|_\infty &= \|D^* S^*(SDv_\varepsilon - y_d)\|_\infty \\ &\le \|D^*\|_{L^\infty \to L^\infty} \|S^*\|_{L^2 \to L^\infty} \|SDv_\varepsilon - y_d\| \le c_p, \end{aligned} \tag{20}$$

where c_p does not depend on ε, since $\|SDv_\varepsilon - y_d\| \le \|S\|_{L^2\to L^2} \|D\|_{L^2 \to L^2} \|c_v\| + \|y_d\|$. Next we evaluate (17). Let us define the sets

$$\begin{aligned} M_+(\varepsilon) &:= \Big\{x \in \Omega \,\Big|\, p_\varepsilon(x) + \nu(D^* D v_\varepsilon)(x) - \frac{\mu}{v_\varepsilon(x) - y_c(x)} > 0\Big\} \\ M_0(\varepsilon) &:= \Big\{x \in \Omega \,\Big|\, p_\varepsilon(x) + \nu(D^* D v_\varepsilon)(x) - \frac{\mu}{v_\varepsilon(x) - y_c(x)} = 0\Big\}. \end{aligned}$$

Due to (17), $M_+(\varepsilon) \cup M_0(\varepsilon)$ cover Ω up to a set of measure zero. Clearly, the variational inequality (17) implies $v_\varepsilon(x) - y_c = \varepsilon$ for almost all $x \in M_+(\varepsilon)$.

Theorem 3.4. *There exist constants $a, b > 0$ such that the set $M_+(\varepsilon)$ has measure zero for all $\varepsilon < a(\sqrt{1 + b\mu} - 1)$.*

Proof. For almost all $x \in M_+(\varepsilon)$, the constraint is active, i.e. $v_\varepsilon(x) - y_c(x) = \varepsilon$. Thus by (20) we have for almost all $x \in M_+(\varepsilon)$

$$c_p + \nu\,(D^*Dv_\varepsilon)(x) - \frac{\mu}{\varepsilon} \geq p_\varepsilon(x) + \nu(D^*Dv_\varepsilon)(x) - \frac{\mu}{v_\varepsilon(x) - y_c(x)} > 0. \tag{21}$$

By (19),

$$D^*D = \lambda^{-2}(I - S^*D^*)(I - SD) = \lambda^{-2}I + K$$

with $K : L^2 \to L^\infty$

$$K = A^{-2}\left\{-(S^*D^* + SD) + S^*D^*DS\right\}$$

bounded. Moreover, we know almost everywhere on $M_+(\varepsilon)$ that $v_\varepsilon(x) = y_c(x) + \varepsilon$, hence

$$c_p + \nu\,(D^*Dv_\varepsilon)(x) = c_p + \nu\,(\lambda^{-2}(y_c(x) + \varepsilon) + (K\,v_\varepsilon)(x)).$$

With the left-hand side of (21), Theorem 3.3 yields

$$c_p + \nu(\lambda^{-2}(y_c(x) + \varepsilon) + \|K\|_{L^2 \to L^\infty} c_v) > \frac{\mu}{\varepsilon}.$$

It is visible that the right hand side tends to zero as $\varepsilon \downarrow 0$ while the left hand side remains bounded. Therefore, the inequation can not be satisfied for small ε.

Solving this quadratic inequality for ε establishes the existence of constants $a, b > 0$ such that

$$\varepsilon > a(\sqrt{1 + b\mu} - 1).$$

For smaller ε, $M_+(\varepsilon)$ must therefore have measure zero. □

Corollary 3.5. *For all $\varepsilon < a(\sqrt{1 + b\mu} - 1)$, the solution v_ε of $(\mathrm{P}_\mu^\varepsilon)$ is the unique solution to (P_μ).*

Proof. For these ε, the set $M_+(\varepsilon)$ has measure zero. Therefore, it holds

$$p_\varepsilon(x) + \nu(D^*Dv_\varepsilon)(x) - \frac{\mu}{v_\varepsilon(x) - y_c(x)} = 0$$

almost everywhere on Ω, hence v_ε satisfies the first order necessary optimality conditions for the optimization problem (P_μ). This is a problem with convex objective functional; the necessary conditions are sufficient for optimality. Strong convexity yields uniqueness (notice that $\nu > 0$). Therefore, v_ε is the unique solution of (P_μ). □

Corollary 3.6. *There exists a constant $c_\mu > 0$ such that for $\mu \leq 1$ the unique solution v_μ of* (13) *satisfies $v_\mu \geq y_c + c_\mu \mu$ a.e. on Ω.*

4 Convergence of the central path

Having established the existence of the central path $\mu \mapsto v_\mu$ for all $\mu > 0$, we can proceed with proving continuity of the path and convergence towards a solution.

The unique minimizer of (13) can be characterized by (14) as

$$F(v_\mu; \mu) = (D^*S^*SD + \nu D^*D)v_\mu - \frac{\mu}{v_\mu - y_c} = 0 \quad \text{a.e.}$$

Since $v_\mu - y_c \geq c_\mu \mu$ holds for $\mu \leq 1$ by Corollary 3.6, F is directionally differentiable in all directions $v \in L^\infty$. We denote the partial derivatives w.r.t. v and μ by $\partial_v F$ and $\partial_\mu F$, respectively. The derivative $\partial_v F$ is

$$\begin{aligned}
\partial_v F(v;\mu) &= (D^*S^*SD + \nu D^*D) + \frac{\mu}{(v-y_c)^2} &(22)\\
&= (D^*S^*SD + \nu K) + \left(\frac{\nu}{\lambda^2} + \frac{\mu}{(v-y_c)^2}\right)\\
&= \bar{K} + \left(\frac{\nu}{\lambda^2} + \frac{\mu}{(v-y_c)^2}\right), &(23)
\end{aligned}$$

where

$$\bar{K} = D^*S^*SD + \nu K$$

is a bounded operator from L^2 to L^∞. Let $\gamma = \|\bar{K}\|_{L^2\to L^\infty}$. From (22) and (10) we see immediately that, for all $v \geq y_c + \epsilon$, $\partial_v F(v;\mu) \in \mathcal{L}(L^2, L^2)$ is a symmetric positive definite operator with

$$\langle \xi, \partial_v F(v;\mu)\xi\rangle \geq \nu\langle D\xi, D\xi\rangle \geq \nu\|S+\lambda I\|^{-2}\|\xi\|^2.$$

The Lax-Milgram theorem guarantees the existence of a bounded inverse $\partial_v F(v;\mu)^{-1} : L^2 \to L^2$ with

$$\|\partial_v F(v;\mu)^{-1}\| \leq \frac{1}{\nu}(\|S\| + |\lambda|)^2. \tag{24}$$

In the next lemma we prove a further regularity property of $\partial_v F$.

Lemma 4.1. *The derivative $\partial_v F(v;\mu) : L^\infty \to L^\infty$ with $v > y_c$ is a bijective operator with bounded inverse $\partial_v F(v;\mu)^{-1} : L^\infty \to L^\infty$, where $\|\partial_v F(v;\mu)^{-1}\|_{L^\infty\to L^\infty} \leq c_i$ is bounded independently of μ.*

Proof. Due to (24), for each $z \in L^\infty \subset L^2$ there is a solution $\xi \in L^2$ to $\partial_v F(v;\mu)\xi = z$ with

$$\|\xi\| \leq \frac{1}{\nu}(\|S\| + |\lambda|)^2\|z\| \leq \frac{\sqrt{|\Omega|}}{\nu}(\|S\| + |\lambda|)^2\|z\|_\infty. \tag{25}$$

Now we have by (23)

$$\left(\frac{\nu}{\lambda^2} + \frac{\mu}{(v-y_c)^2}\right)\xi = z - \bar{K}\xi$$

and hence by (25)

$$\begin{aligned}
\|\xi\|_\infty &\leq \frac{\lambda^2}{\nu}\left(\|z\|_\infty + \|\bar{K}\|_{L^2\to L^\infty}\|\xi\|\right)\\
&\leq \frac{\lambda^2}{\nu}\left(1 + \gamma\frac{\sqrt{|\Omega|}}{\nu}(\|S\| + |\lambda|)^2\right)\|z\|_\infty\\
&=: c_i\|z\|_\infty.
\end{aligned}$$

Thus, $\xi \in L^\infty$ holds, such that $\partial_v F(v;\mu) : L^\infty \to L^\infty$ is bijective and has a bounded inverse $\|\partial_v F(v;\mu)^{-1}\|_{L^\infty\to L^\infty} \leq c_i$.

□

With the invertibility of $\partial_v F$ at hand we make use of the implicit function theorem in order to justify the notion of a central path. We obtain the

Corollary 4.2. *The mapping $\mu \mapsto v_\mu$ is continuously differentiable from $\mathbb{R}_+$ to L^∞.*

Now we turn to convergence of the central path towards a solution of (8).

Theorem 4.3. *For $\mu \to 0$, the central path converges towards a KKT point v_0 of (8). There exists a constant $c_0 < \infty$ such that the following error estimate holds for all $\mu \le 1$:*

$$\|v_0 - v_\mu\|_{L^\infty} \le c_0\sqrt{\mu} \tag{26}$$

The proof is somewhat technical, for which we give a sketch of its main ideas beforehand. For this purpose we assume for now that $\bar{K} = 0$, such that $\partial_v F(v;\mu) = \nu/\lambda^2 + \mu/(v - y_c)^2$ is a Nemyckii operator. By the implicit function theorem, the derivative v'_μ of the central path is given by

$$\begin{aligned} v'_\mu &= -\partial_v F(v_\mu;\mu)^{-1}\partial_\mu F(v_\mu;\mu) \\ &= \left(\frac{\nu}{\lambda^2} + \frac{\mu}{(v_\mu - y_c)^2}\right)^{-1} \frac{1}{v_\mu - y_c} \\ &= \left(\frac{\nu(v_\mu - y_c)}{\lambda^2} + \frac{\mu}{v_\mu - y_c}\right)^{-1} \end{aligned} \tag{27}$$

Using the fact that

$$ax + \frac{b}{x} \ge 2\sqrt{ab} \tag{28}$$

holds for arbitrary $a, b, x > 0$, we see immediately that

$$v'_\mu \le \left(2\sqrt{\frac{\nu\mu}{\lambda^2}}\right)^{-1} \le \frac{c}{\sqrt{\mu}}.$$

Integrating the slope of the central path from 0 to μ yields the length of the central path and therefore an error bound of

$$\|v_\mu - v_0\|_\infty \le c\sqrt{\mu}.$$

However, the operator $\bar{K}$ is compact but nonzero, and introduces a nonlocal coupling across the domain Ω. Bounding this coupling requires a more involved proof as given below.

Proof. First we will establish an L^2-bound on v'_μ and infer an L^∞-bound from that. From this we will determine the existance of and distance to the limit point v_0, and finally check the first order necessary conditions for v_0.
(i) *L^2-estimate.* We set out to construct a splitting of the domain Ω into two different regions, such that the nonlocal coupling introduced by $\bar{K}$ is dominated by purely local effects in each subdomain and is in a certain sense sufficiently small across the subdomains. To this extend we define $T = D^*S^*SD + \nu D^*D = \bar{K} + \nu/\lambda^2$ and the characteristic functions of the almost active and almost inactive sets by

$$\chi_A = \begin{cases} 1, & v_\mu - y_c \le C \\ 0, & \text{otherwise} \end{cases} \qquad \text{and} \qquad \chi_I = 1 - \chi_A, \tag{29}$$

respectively, with

$$C = \sqrt{\frac{\mu}{2\|T\|(1 + \|T\|\,\|T^{-1}\|)}}. \tag{30}$$

Notice that multiplication by χ_A and χ_I acts as a projection onto two orthogonal subspaces of L^2 with disjoint support.

We may reformulate (27) as

$$\partial_v F(v_\mu;\mu)v'_\mu = \left(T + \frac{\mu}{(v_\mu - y_c)^2}\right) v'_\mu = \frac{\chi_A + \chi_I}{v_\mu - y_c} \quad \text{a.e.}, \tag{31}$$

such that we can obtain individual bounds for each summand of the right hand side. First we consider

$$\partial_v F(v_\mu;\mu)v_I = \frac{\chi_I}{v_\mu - y_c}.$$

By (24) we readily obtain some constant $c_I < \infty$ such that

$$\|v_I\| = \left\|\partial_v F(v_\mu;\mu)^{-1}\frac{\chi_I}{v_\mu - y_c}\right\| \leq \frac{1}{\nu}(\|S\| + |\lambda|)^2\frac{1}{C} \leq \frac{c_I}{\sqrt{\mu}}. \tag{32}$$

Now we turn to the remaining part of (31) on the almost active set, which we write as

$$(\chi_A + \chi_I)\partial_v F(v_\mu;\mu)(\chi_A + \chi_I)v_A = \frac{\chi_A}{v_\mu - y_c}.$$

Expanding the left hand side and separating the terms according to the subspaces $L^2(\operatorname{supp}\chi_A)$ and $L^2(\operatorname{supp}\chi_I)$ generated by the projections χ_A and χ_I, respectively, yields

$$\begin{aligned}\chi_A\partial_v F(v_\mu;\mu)\chi_A v_A + \chi_A\partial_v F(v_\mu;\mu)\chi_I v_A &= \frac{\chi_A}{v_\mu - y_c}\\ \chi_I\partial_v F(v_\mu;\mu)\chi_A v_A + \chi_I\partial_v F(v_\mu;\mu)\chi_I v_A &= 0.\end{aligned}$$

In the upper left block of the equation system, completely defined on the almost active set, the interior point regularization dominates, such that we shift the remaining parts to the right hand side. The antidiagonal blocks contain only the nonlocal coupling introduced by $\bar{K}$ and are moved to the right hand side in both equations. We end up with

$$\left[\chi_A\frac{\mu}{(v_\mu - y_c)^2}\chi_A\right]\ \chi_A v_A = \frac{\chi_A}{v_\mu - y_c} - \chi_A T\chi_A v_A - \chi_A T\chi_I v_A \tag{33}$$

$$\left[\chi_I\left(T + \frac{\mu}{(v_\mu - y_c)^2}\right)\chi_I\right]\ \chi_I v_A = -\chi_I T\chi_A v_A. \tag{34}$$

Notice that the restriction

$$\chi_I\left(T + \frac{\mu}{(v_\mu - y_c)^2}\right)\chi_I \in \mathcal{L}(L^2(\operatorname{supp}\chi_I), L^2(\operatorname{supp}\chi_I))$$

is a symmetric positive definite operator with bounded inverse

$$\left\|\left(\chi_I\left(T + \frac{\mu}{(v_\mu - y_c)^2}\right)\chi_I\right)^{-1}\right\| \leq \|T^{-1}\|.$$

Hence, (34) has a unique solution which is bounded in terms of the right hand side. On the almost active set, (33) can be solved pointwise. Solving both equations yields

$$\begin{aligned}\|\chi_A v_A\| \leq{}& \left\|\chi_A\frac{v_\mu - y_c}{\mu}\right\| + \left\|\chi_A\frac{(v_\mu - y_c)^2}{\mu}\right\|\|T\|\|\chi_A v_A\|\\ &+ \left\|\chi_A\frac{(v_\mu - y_c)^2}{\mu}\right\|\|T\|\|\chi_I v_A\|\end{aligned} \tag{35}$$

and

$$\begin{aligned}\|\chi_I v_A\| &\le \left\| \left(\chi_I \left(T + \frac{\mu}{(v_\mu - y_c)^2} \right) \chi_I \right)^{-1} \right\| \|\chi_I T \chi_A v_A\| \\ &\le \|T^{-1}\| \|T\| \|\chi_A v_A\|. \end{aligned} \tag{36}$$

Inserting (36) into (35) and using (29) and (30) we obtain

$$\|\chi_A v_A\| \le \frac{C\sqrt{|\Omega|}}{\mu} + \frac{C^2}{\mu} \|T\| (1 + \|T^{-1}\| \|T\|) \|\chi_A v_A\| = \frac{C\sqrt{|\Omega|}}{\mu} + \frac{1}{2} \|\chi_A v_A\|.$$

By (30) this verifies the existence of some constant $\bar{c} < \infty$ such that $\|\chi_A v_A\| \le \frac{\bar{c}}{\sqrt{\mu}}$. Finally, $\|\chi_I v_A\| \le \|T^{-1}\| \|T\| \frac{\bar{c}}{\sqrt{\mu}}$ holds, such that by (32) there is a constant $\hat{c} < \infty$ with

$$\|v'_\mu\| \le \|\chi_A v_A\| + \|\chi_I v_A\| + \|v_I\| \le \frac{\hat{c}}{\sqrt{\mu}}. \tag{37}$$

(ii) *L^∞-estimates.* Returning to (31) we obtain

$$\begin{aligned}\|v'_\mu\|_{L^\infty} &\le \left\| \left(\frac{\nu}{\lambda^2} + \frac{\mu}{(v_\mu - y_c)^2} \right)^{-1} \frac{1}{v_\mu - y_c} \right\|_{L^\infty} + \left\| \left(\frac{\nu}{\lambda^2} + \frac{\mu}{(v_\mu - y_c)^2} \right)^{-1} \bar{K} v'_\mu \right\|_{L^\infty} \\ &\le \left\| \left(\frac{\nu(v_\mu - y_c)}{\lambda^2} + \frac{\mu}{v_\mu - y_c} \right)^{-1} \right\|_{L^\infty} + \left\| \left(\frac{\nu}{\lambda^2} + \frac{\mu}{(v_\mu - y_c)^2} \right)^{-1} \right\|_{L^\infty} \gamma \|v'_\mu\|. \end{aligned}$$

Using (28) we proceed with

$$\|v'_\mu\|_{L^\infty} \le \left\| \left(2\sqrt{\frac{\nu\mu}{\lambda^2}} \right)^{-1} \right\|_{L^\infty} + \frac{\lambda^2}{\nu} \gamma \frac{\hat{c}}{\sqrt{\mu}} \le \frac{c_0}{\sqrt{\mu}}$$

for some $c_0 < \infty$.
(iii) *Distance to the limit point.* The distance between two points on the central path is therefore bounded by

$$\|v_{\mu_1} - v_{\mu_2}\|_{L^\infty} \le \int_{\mu_1}^{\mu_2} \|v'_\mu\|_{L^\infty} \, d\mu \le \frac{c_0}{2} (\sqrt{\mu_2} - \sqrt{\mu_1}). \tag{38}$$

Since for any sequence $\mu_k \to 0$ the corresponding sequence v_{μ_k} of central path points forms a Cauchy sequence, the path converges towards some limit point v_0. Performing the limit process $\mu_1 \to 0$ verifies the error bound (26).
(iv)*First order necessary conditions.* Recalling the Lagrange multiplier approximations η_μ from (15) we write (14) as $f'(v_\mu) = \eta_\mu$. Due to the continuity of f' and the convergence of $v_\mu \to v_0$ in L^2, the multiplier approximations converge towards $\eta_0 = f'(v_0)$ in L^2. Since $\eta_\mu \ge 0$ and $\eta_\mu (v_\mu - y_c) = \mu$ for almost all $x \in \Omega$ and therefore $\langle \eta_\mu, v_\mu - y_c \rangle = \mu |\Omega|$, the same holds by continuity for η_0, i.e. $\eta_0 \ge 0$ and $\langle \eta_0, v_0 - y_c \rangle = 0$. Since the first order necessary conditions are satisfied, v_0 is a KKT point for (11). □

5 Convergence of a short step pathfollowing method

For the analysis of interior point methods, local norms are an invaluable tool. Here we use the scaled norm

$$\|v\|_\mu = \|z_\mu v\|_{L^\infty},$$

with the scaling

$$z_\mu = \sqrt{\frac{\nu}{\lambda^2} + \frac{\mu}{(v_\mu - y_c)^2}},$$

which is closely connected to the energy norms used in the theory of self-concordant barrier functionals [7, 8].

We consider a short-step pathfollowing method with classical predictor. Since we are interested in actually implementable algorithms, we have to use an inexact Newton corrector, which replaces the infinite dimensional Newton equation

$$\partial_v F(v^k; \mu^{k+1})\Delta v^k = -F(v^k; \mu^{k+1})$$

for the exact correction Δv^k by a suitably discretized finite dimensional counterpart

$$\partial_v F(v^k; \mu^{k+1})\Delta v_h^k = -F(v^k; \mu^{k+1}) + r^k.$$

for the inexact correction Δv_h^k, such that an inner residual r^k remains. The iteration index is denoted by a superscript. Another source of inexactness is e.g. the iterative solution of the state equation. The algorithm reads as follows.

Algorithm 5.1.
Choose $0 < \sigma < 1$*,* $\delta > 0$*,* $\mu^0 > 0$*, and* $v^0 > y_c$
For $k = 0, \ldots$
$\mu^{k+1} = \sigma \mu^k$
solve $\partial_v F(v^k; \mu^{k+1})\Delta v_h^k = -F(v^k; \mu^{k+1})$
up to a relative accuracy of $\|\Delta v_h^k - \Delta v^k\|_{\mu^{k+1}} \le \delta \|\Delta v^k\|_{\mu^{k+1}}$
$v^{k+1} = v^k + \Delta v_h^k$

The remainder of the section is devoted to proving that for suitable choices of σ, δ, μ^0, and v^0, all iterates of this algorithm are well defined and converge towards the solution point v_0. First we derive the analogue of Lemma 4.1 for the scaled norm.

Lemma 5.2. *There is some constant* $c_z < \infty$ *independent of* μ*, such that*

$$\|z_\mu \partial_v F(v; \mu)^{-1}\zeta\|_{L^\infty} \le c_z (1+\vartheta)^2 \|z_\mu^{-1}\zeta\|_{L^\infty}$$

for all $v \in B_\mu(v_\mu; \vartheta\sqrt{\mu}) = \{v \in L^\infty : \|v - v_\mu\|_\mu \le \vartheta\sqrt{\mu}\}$ *with* $\vartheta < 1$.

Proof. First we see that

$$\left\| \frac{v - v_\mu}{v_\mu - y_c} \right\|_{L^\infty} \le \left\| z_\mu \frac{v - v_\mu}{\sqrt{\mu}} \right\|_{L^\infty} \le \frac{\|v - v_\mu\|}{\sqrt{\mu}} \le \vartheta < 1$$

and therefore

$$v - y_c \ge (1-\vartheta)(v_\mu - y_c) \quad \text{and} \quad v - y_c \le (1+\vartheta)(v_\mu - y_c), \tag{39}$$

such that $\partial_v F(v_\mu; \mu)$ is invertible. Define $\xi = \partial_v F(v_\mu; \mu)^{-1}\zeta$.
Analogously to Lemma 4.1 we distinguish two cases and first assume that

$$\|z_\mu^{-1}\|_{L^\infty} \|\xi\|_{L^2} \ge \alpha \|z_\mu \xi\|_{L^\infty}$$

for some arbitrary $\alpha > 0$. Then we obtain

$$
\begin{aligned}
\|z_\mu^{-1}\partial_v F(v)\xi\|_{L^\infty} &\geq \|z_\mu^{-1}\|_{L^\infty}\|\partial_v F(v)\xi\|_{L^\infty} \\
&\geq \frac{1}{\sqrt{|\Omega|}}\|z_\mu^{-1}\|_{L^\infty}\|\partial_v F(v)\xi\|_{L^2} \\
&\geq \frac{1}{\sqrt{|\Omega|}}\|z_\mu^{-1}\|_{L^\infty}\nu(\|S\|+|\lambda|)^{-2}\|\xi\|_{L^2} \\
&\geq \frac{\nu\alpha}{\sqrt{|\Omega|}}(\|S\|+|\lambda|)^{-2}\|z_\mu\xi\|_{L^\infty}. \qquad (40)
\end{aligned}
$$

Otherwise we have by (39)

$$
\begin{aligned}
\|z_\mu^{-1}\partial_v F(v)\xi\|_{L^\infty} &= \left\| z_\mu^{-1}\bar{K}\xi + z_\mu^{-1}\left(\frac{\nu}{\lambda^2}+\frac{\mu}{(v-y_c)^2}\right)\xi\right\|_{L^\infty} \\
&\geq -\|z_\mu^{-1}\|_{L^\infty}\|\bar{K}\xi\|_{L^\infty} + \left\| z_\mu^{-1}\frac{\mu}{(1+\vartheta)^2(v_\mu-y_c)^2}\xi\right\|_{L^\infty} \\
&\geq -\|z_\mu^{-1}\|_{L^\infty}\gamma\|\xi\|_{L^2} + (1+\vartheta)^{-2}\|z_\mu\xi\|_{L^\infty} \\
&\geq -\gamma\alpha\|z_\mu\xi\|_{L^\infty} + (1+\vartheta)^{-2}\|z_\mu\xi\|_{L^\infty} \\
&= ((1+\vartheta)^{-2}-\gamma\alpha)\|z_\mu\xi\|_{L^\infty}. \qquad (41)
\end{aligned}
$$

Choosing

$$
\rho = \left(\frac{\nu}{\sqrt{|\Omega|}(\|S\|+|\lambda|)^2}+\gamma\right)^{-1} \quad \text{and} \quad \alpha = \rho(1+\vartheta)^{-2},
$$

the claim is verified for $c_z = (1-\gamma\rho)^{-1} < \infty$. □

Next we prove a continuity result for the scaled norm.

Lemma 5.3. *There is a constant $c_\sigma < \infty$ independent of μ such that*

$$
\|v\|_{\sigma\mu} \leq (1 + c_\sigma(1-\sigma)\|v\|_\mu \qquad (42)
$$

holds for all $v \in L^\infty$ and

$$
\frac{c_z}{c_z+1/2} \leq \sigma \leq 1.
$$

Proof. We begin with estimating the derivative of the central path in the scaled norm. Lemma 5.2 applied to (27) results in

$$
\|v'_\mu\|_\mu \leq c_z\|z_\mu^{-1}\partial_v F(v_\mu;\mu)v'_\mu\|_{L^\infty} = c_z\|z_\mu^{-1}(v_\mu-y_c)^{-1}\|_{L^\infty} \leq \frac{c_z}{\sqrt{\mu}}. \qquad (43)
$$

We proceed with introducing the monotonically decreasing majorant $\Theta(\sigma)$ for the expression

$$
f(\sigma) = \left\|\frac{v_\mu-y_c}{v_{\sigma\mu}-y_c}\right\|_{L^\infty} \leq \Theta(\sigma)
$$

by

$$\begin{aligned}
\Theta(\sigma) &= f(1) + \int_\sigma^1 f'(\tau)\, d\tau \\
&\le 1 + \int_\sigma^1 \left\| \frac{v_\mu - y_c}{(v_{\tau\mu} - y_c)^2} v'_{\tau\mu}\mu \right\|_{L^\infty} d\tau \\
&\le 1 + \int_\sigma^1 \left\| \frac{v_\mu - y_c}{v_{\tau\mu} - y_c} \right\|_{L^\infty} \left\| \frac{\sqrt{\tau\mu}}{v_{\tau\mu} - y_c} v'_{\tau\mu} \right\|_{L^\infty} \frac{\mu}{\sqrt{\tau\mu}}\, d\tau \\
&\le 1 + \int_\sigma^1 \Theta(\tau) \|v'_{\tau\mu}\|_{\tau\mu} \frac{\mu}{\sqrt{\tau\mu}}\, d\tau \\
&\le 1 + \int_\sigma^1 \Theta(\sigma) \frac{c_z}{\sqrt{\tau\mu}} \frac{\mu}{\sqrt{\tau\mu}}\, d\tau \\
&\le 1 + \Theta(\sigma) \frac{c_z}{\sigma}(1-\sigma).
\end{aligned}$$

Solving for Θ yields

$$\left\| \frac{v_\mu - y_c}{v_{\sigma\mu} - y_c} \right\|_{L^\infty} \le \left(1 - \frac{c_z}{\sigma}(1-\sigma)\right)^{-1}.$$

Now from

$$\|v\|_{\sigma\mu} = \left\| z_\mu \frac{z_{\sigma\mu}}{z_\mu} v \right\|_{L^\infty} \le \|v\|_\mu \sqrt{\sigma} \left\| \frac{v_\mu - y_c}{v_{\sigma\mu} - y_c} \right\|_{L^\infty} \le \sqrt{\sigma} \left(1 - \frac{c_z}{\sigma}(1-\sigma)\right)^{-1} \|v\|_\mu$$

the constant c_σ is easily established. □

Lemma 5.4. *There exists some constant $\omega < \infty$ such that the Lipschitz condition*

$$\left\| \partial_v F(v;\mu)^{-1} (\partial_v F(v;\mu) - \partial_v F(\hat{v};\mu))(v - \hat{v}) \right\|_\mu \le \frac{\omega}{\sqrt{\mu}} \|v - \hat{v}\|_\mu^2 \tag{44}$$

holds for all $v, \hat{v} \in B_\mu(v_\mu, \vartheta\sqrt{\mu})$ with $\vartheta < 1$.

Proof. Using Lemma 5.2 we have

$$\begin{aligned}
&\left\| \partial_v F(v;\mu)^{-1} (\partial_v F(v;\mu) - \partial_v F(\hat{v};\mu))(v - \hat{v}) \right\|_\mu \\
&\qquad \le c_z (1+\vartheta)^2 \left\| z_\mu^{-1} (\partial_v F(v;\mu) - \partial_v F(\hat{v};\mu))(v - \hat{v}) \right\|_{L^\infty} \\
&\qquad = c_z (1+\vartheta)^2 \left\| z_\mu^{-1} \left(\frac{\mu}{(v - y_c)^2} - \frac{\mu}{(\hat{v} - y_c)^2} \right) (v - \hat{v}) \right\|_{L^\infty} \\
&\qquad \le c_z (1+\vartheta)^2 \left\| z_\mu^{-1} \mu \frac{v - \hat{v}}{((1-\vartheta)(v_\mu - y_c))^3} (v - \hat{v}) \right\|_{L^\infty} \\
&\qquad = c_z \frac{(1+\vartheta)^2}{(1-\vartheta)^3} \left\| \frac{\mu}{z_\mu^3 (v_\mu - y_c)^3}\ z_\mu^2 (v - \hat{v})^2 \right\|_{L^\infty} \\
&\qquad \le c_z \frac{(1+\vartheta)^2}{(1-\vartheta)^3} \left\| \frac{\mu}{z_\mu^3 (v_\mu - y_c)^3} \right\|_{L^\infty} \|v - \hat{v}\|_\mu^2 \\
&\qquad \le \frac{c_z (1+\vartheta)^2}{\sqrt{\mu}(1-\vartheta)^3} \|v - \hat{v}\|_\mu^2,
\end{aligned}$$

which proves the claim for $\omega = c_z \frac{(1+\vartheta)^2}{(1-\vartheta)^3}$. □

We can now prove the convergence of the pathfollowing method.

Theorem 5.5. *Assume that*

$$\delta \le \rho\frac{1-\rho}{1+\rho}, \quad \sigma \ge 1 - \left(1 - \frac{\rho\vartheta + 3c_\sigma + c_z}{\vartheta + 3c_\sigma + c_z}\right) \frac{\sqrt{\rho - \delta(1+\rho)} - \rho}{1-\rho}, \tag{45}$$

and $\|v^0 - v_{\mu^0}\|_{\mu^0} \le \rho\vartheta\sqrt{\mu^0}$ *for* $\vartheta \le (c_z+2)^{-1}$ *and some* $\rho < 1$. *Then the iterates* v^k *defined by Algorithm 5.1 are all well defined and converge linearly towards the limit point* v_0. *More precisely,*

$$\left\|v^k - v_{\mu^k}\right\|_{\mu^k} \le \rho\vartheta\sqrt{\mu^k}. \tag{46}$$

Proof. By induction, we assume that (46) holds for the current iterate v^k. For $\sigma > c_z/(c_z+2)$, Lemma 5.3 and (43) yield

$$\begin{aligned}
\|v^k - v_{\mu^{k+1}}\|_{\mu^{k+1}} &\le \|v^k - v_{\mu^k}\|_{\mu^{k+1}} + \|v_{\mu^k} - v_{\mu^{k+1}}\|_{\mu^{k+1}} \\
&\le (1 + c_\sigma(1-\sigma))\|v^k - v_{\mu^k}\|_{\mu^k} + \int_{\mu^{k+1}}^{\mu^k} \|v'_\tau\|_{\mu^{k+1}}\, d\tau \\
&\le (1 + c_\sigma(1-\sigma))\rho\vartheta\sqrt{\mu^k} + \int_{\mu^{k+1}}^{\mu^k} (1 + c_\sigma(1-\sigma))\|v'_\tau\|_\tau\, d\tau \\
&\le (1 + c_\sigma(1-\sigma)) \left(\rho\vartheta\sqrt{\mu^k} + \int_{\mu^{k+1}}^{\mu^k} \frac{c_z}{\sqrt{\tau}}\, d\tau\right) \\
&\le (1 + c_\sigma(1-\sigma)) \left(\rho\vartheta\sqrt{\mu^k} + \frac{c_z\mu^k(1-\sigma)}{\sqrt{\mu^{k+1}}}\right) \\
&\le (1 + c_\sigma(1-\sigma)) \left(\frac{\rho\vartheta}{\sqrt{\sigma}} + \frac{c_z(1-\sigma)}{\sigma}\right) \sqrt{\mu^{k+1}}.
\end{aligned}$$

Notice that

$$\sigma \ge \sigma_{\min} = \frac{\rho\vartheta + 3c_\sigma + c_z}{\vartheta + 3c_\sigma + c_z} > \frac{c_z}{c_z+2}$$

implies

$$\begin{aligned}
\|v^k - v_{\mu^{k+1}}\|_{\mu^{k+1}} &\le \frac{1}{\sigma}(1 + c_\sigma(1-\sigma))(\rho\vartheta + c_z(1-\sigma))\sqrt{\mu^{k+1}} \\
&= \frac{1}{\sigma}(\rho\vartheta + (c_\sigma\rho\vartheta + c_z)(1-\sigma) + c_\sigma c_z(1-\sigma)^2)\sqrt{\mu^{k+1}} \\
&< \frac{1}{\sigma}(\rho\vartheta + (c_\sigma\rho\vartheta + c_z + 2c_\sigma)(1-\sigma))\sqrt{\mu^{k+1}} \\
&\le \vartheta\sqrt{\mu^{k+1}}.
\end{aligned} \tag{47}$$

Since the estimate (47) is convex, we end up with the error bound

$$\|v^k - v_{\mu^{k+1}}\|_{\mu^{k+1}} < \left(\rho\vartheta + \frac{1-\sigma}{1-\sigma_{\min}}(1-\rho)\vartheta\right)\sqrt{\mu^{k+1}}.$$

Dropping the fixed argument μ^{k+1} from F we obtain the error of the exact Newton corrector result as

$$
\begin{aligned}
v_{\mu^{k+1}} - v^k - \Delta v^k &= v_{\mu^{k+1}} - v^k + \partial_v F(v^k)^{-1} F(v^k) \\
&= \partial_v F(v^k)^{-1} \left(F(v^k) + \partial_v F(v^k)(v_{\mu^{k+1}} - v^k) \right) \\
&= - \int_0^1 \partial_v F(v^k)^{-1} \left(\partial_v F(v^k + s(v_{\mu^{k+1}} - v^k)) - \partial_v F(v^k) \right) (v_{\mu^{k+1}} - v^k) \, ds
\end{aligned}
$$

and by Lemma 5.4

$$
\begin{aligned}
\| v_{\mu^{k+1}} - v^k - \Delta v^k \|_{\mu^{k+1}} &\le \int_0^1 \frac{\omega}{\sqrt{\mu^{k+1}}} s \| v_{\mu^{k+1}} - v^k \|_{\mu^{k+1}}^2 \, ds \\
&< \frac{\omega}{2\sqrt{\mu^{k+1}}} \left(\rho\vartheta + \frac{1-\sigma}{1-\sigma_{\min}} (1-\rho)\vartheta \right)^2 \mu^{k+1}.
\end{aligned}
$$

Since by Lemma 5.4 and the assumption on ϑ

$$
\frac{\omega}{2}\vartheta = \frac{c_z(1+\vartheta)^2}{2(1-\vartheta)^3}\vartheta \le \frac{c_z\vartheta}{1-2\vartheta} \le \frac{c_z \frac{1}{c_z+2}}{1-\frac{2}{c_z+2}} = 1
$$

holds, we can further estimate

$$
\| v_{\mu^{k+1}} - v^k - \Delta v^k \|_{\mu^{k+1}} < \left(\rho + \frac{1-\sigma}{1-\sigma_{\min}} (1-\rho) \right)^2 \vartheta \sqrt{\mu^{k+1}}.
$$

Here it is apparent that choosing $\sigma = \sigma_{\min}$ is just sufficient for an exact Newton corrector iteration to converge. However, we aim at restoring the tolerance $\rho\theta\sqrt{\mu^{k+1}}$ in a single Newton step. With the additional stepsize restriction

$$
\sigma \ge 1 - (1-\sigma_{\min}) \frac{\sqrt{\rho}-\rho}{1-\rho}
$$

we obtain

$$
\rho + \frac{1-\sigma}{1-\sigma_{\min}} (1-\rho) \le \sqrt{\rho}
$$

and thus

$$
\| v_{\mu^{k+1}} - v^k - \Delta v^k \|_{\mu^{k+1}} < \rho\vartheta\sqrt{\mu^{k+1}}.
$$

Up to now, we have considered the exact Newton correction with a length of

$$
\begin{aligned}
\| \Delta v^k \|_{\mu^{k+1}} &\le \| v_{\mu^{k+1}} - v^k - \Delta v^k \|_{\mu^{k+1}} + \| v_{\mu^{k+1}} - v^k \|_{\mu^{k+1}} \\
&\le \rho\vartheta\sqrt{\mu^{k+1}} + \vartheta\sqrt{\mu^{k+1}} \\
&= (1+\rho)\vartheta\sqrt{\mu^{k+1}}.
\end{aligned}
$$

The next iterate v^{k+1} given by the inexact Newton step has therefore an error bound of

$$
\begin{aligned}
\| v_{\mu^{k+1}} - v^{k+1} \|_{\mu^{k+1}} &\le \| v_{\mu^{k+1}} - v^k - \Delta v^k \|_{\mu^{k+1}} + \delta \| \Delta v^k \|_{\mu^{k+1}} \\
&\le \left[\left(\rho + \frac{1-\sigma}{1-\sigma_{\min}} (1-\rho) \right)^2 + \delta(1+\rho) \right] \vartheta \sqrt{\mu^{k+1}}.
\end{aligned}
$$

With the accuracy requirement and the final stepsize restriction given by (45), we obtain

$$\|v_{\mu^{k+1}} - v^{k+1}\|_{\mu^{k+1}} \leq \rho\vartheta\sqrt{\mu^{k+1}},$$

which completes the induction.

Moreover, together with Theorem 4.3, we obtain

$$\begin{aligned}\|v_0 - v^k\|_{L^\infty} \leq \|v_0 - v_{\mu^k}\|_{L^\infty} + \frac{\lambda}{\sqrt{\nu}}\|v_{\mu^k} - v^k\|_{\mu^{k+1}} \\ \leq c_0\sqrt{\mu^k} + \rho\vartheta\sqrt{\mu^k} \leq (c0 + \rho\vartheta)\sigma^{k/2}\sqrt{\mu^0},\end{aligned}$$

which proves r-linear convergence of v^k to the KKT point v_0. □

6 Numerical tests

We have tested our method by the following example

$$\min J(y,u) := \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u - u_d\|^2_{L^2(\Omega)} \tag{PT}$$

subject to

$$-\Delta y + y = u \quad \text{in}\,\Omega \tag{48}$$

$$\partial_n y = 0 \quad \text{on}\,\Gamma \tag{49}$$

and to the pointwise mixed control-state constraints

$$y + \lambda u \geq y_c \qquad \text{a.e. in}\,\Omega. \tag{50}$$

with $\Omega = (0,1) \times (0,1)$.
The function u_d is introduced for technical reasons.This does not change the validity of our theorems.
It is easy to verify, that (PT) fits into the setting of (P). For all $\lambda > 0$ the Lagrange multiplier η belongs to $L^2(\Omega)$. We consider three different examples. In example 1 and 2 the Lagrange multiplier is in $L^2(\Omega)$ also for $\lambda = 0$. In the third example $\eta \in \mathcal{B}(\Omega)$ for $\lambda = 0$.

We solved the regularized problems numerically by a short-step pathfollowing method, using a conform finite element method to solve the state and adjoint equation, where all variables were discretised by linear finite elements. Note that due to the linearity of the state equation, the computational all-at-once approach used here is indeed an implementation of the inexact Newton method described in §5. Using a primal algorithm, we have calculated the Lagrange multiplier η by the relation

$$\eta = \frac{\mu}{y - y_c + \varepsilon u}.$$

We implemented our method using Matlab and the PDE-toolbox for mesh generation, matrix-assembling etc. The stopping parameter for the outer iteration was $\mu \leq \epsilon = 10^{-6}$, except for the calculation of figures 27–30. For our computations we used a Friedrichs-Keller triangulation with fixed mesh size $h = 0.015625$. In the following, the numerical solutions are denoted by $(\cdot)_h$, the exact optimal control, optimal state resp. the optimal adjoint state are denoted by $\bar{u}$, $\bar{y}$ and $\bar{p}$, resp. In some figures these functions are labeled as u_{opt} etc. Notice, that for fixed mesh size the numerical solutions tend to the projection of the exact solution onto the finite element space. All computations were performed on a dual Pentium IV/2.8GHz machine with 1GB RAM running under Linux.

6.1 Example 1

This example is taken from [6]. We choose $\bar{u} = 2$, $\bar{p} = -2$ and $\bar{y} = 2$. The desired state is given by

$$y_d(x_1, x_2) = 4 - \max\left\{-20\left((x_1 - 0.5)^2 - (x_2 - 0.5)^2\right) + 1, 0\right\},$$

y_c is given by

$$y_c(x_1, x_2) = \min\left\{-20\left((x_1 - 0.5)^2 - (x_2 - 0.5)^2\right) + 3, 2\right\}$$

and the Lagrange multiplier is

$$\eta(x_1, x_2) = \max\left\{-20\left((x_1 - 0.5)^2 - (x_2 - 0.5)^2\right) + 1, 0\right\}.$$

Moreover, we have chosen $u_d = 0$. In (PT) we choose $\nu = 1$ and $\lambda = 10^{-16}$.
The following figures show the exact functions y_d, y_c and the Lagrange multiplier η.

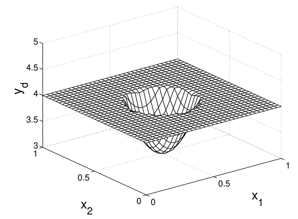


Figure 1: Desired state y_d

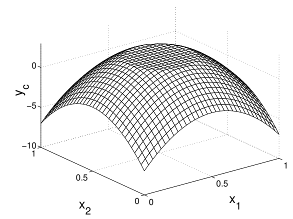


Figure 2: State contraint y_c

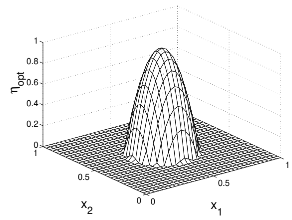


Figure 3: Multiplier η

The next set of figures show the numerical solutions y_h, u_h, p_h, and η_h of the problem regularized with $\lambda = 10^{-16}$.

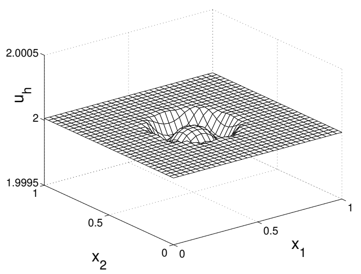


Figure 4: Control u_h

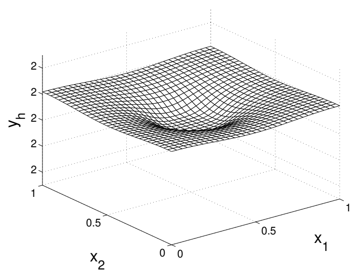


Figure 5: State y_h

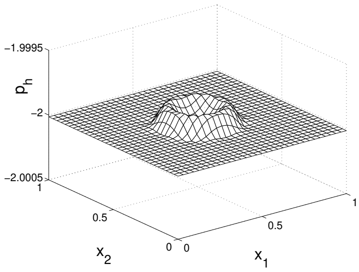


Figure 6: Adjoint state p_h

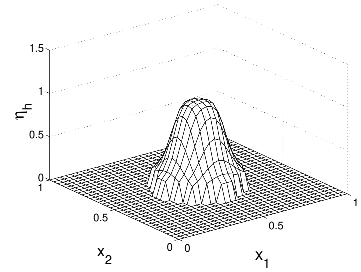


Figure 7: Lagrange multiplier η_h

For a comparsion with results computed by a primal-dual active set strategy we refer to [6]. Note that the scale for y, u and p is in the span of $[1.999999, 2.000001]$, $[1.9995, 2.0005]$ respectively $[-2.0005, -1.9995]$. In contrast to the primal-dual active set strategy in [6], small values of λ do not influence the convergence rate.

The following figures 8–11 show the differences between the numerical solutions u_h,y_h p_h and η_h and the exact solutions u, y, p and η, masured in the L^2-norm for regularization parameter $\lambda = 10^{-16}$. Both axes are scaled logarithmically. With this choice of the regularisation parameter, the convergence of the path is visible.The behavior of the Lagrange multiplier for $\mu \to 0$ is remarkable, see also figures 12–15.

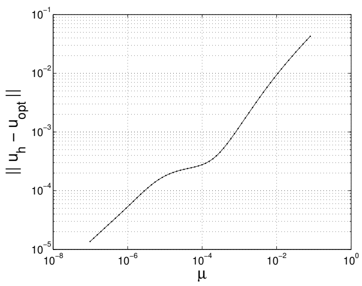


Figure 8: Error $\|u - u_{\text{opt}}\|$

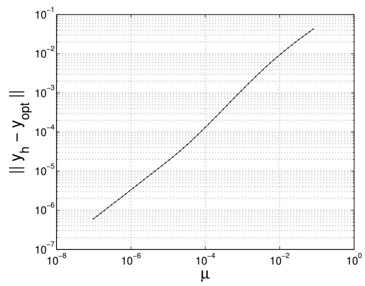


Figure 9: Error $\|y - y_{\text{opt}}\|$

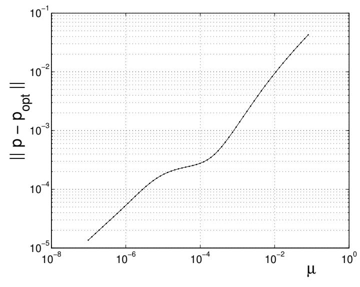


Figure 10: Error $\|p - p_{\text{opt}}\|$

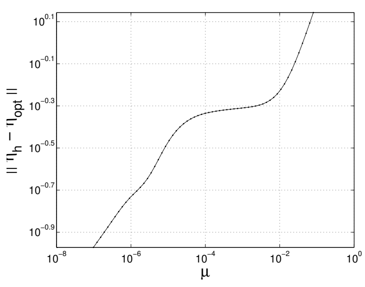


Figure 11: Error $\|\eta - \eta_{\text{opt}}\|$

The next figures show the evolution of the Lagrange-multiplier η_h along the central path.

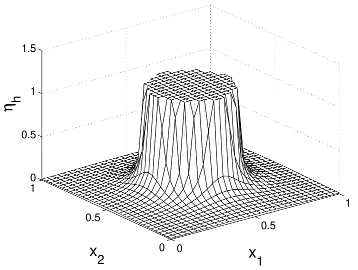


Figure 12: Multiplier η_h at $\mu = 0.01$

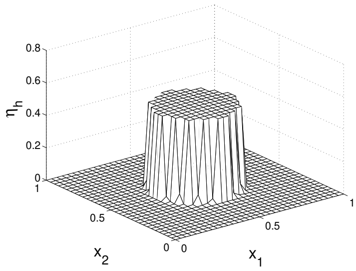


Figure 13: Multiplier η_h at $\mu = 10^{-4}$

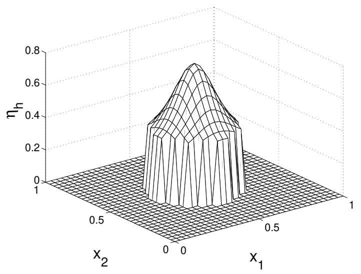


Figure 14: Multiplier η_h at $\mu = 10^{-6}$

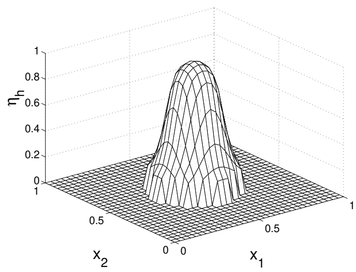
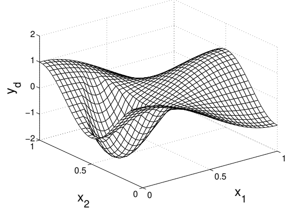
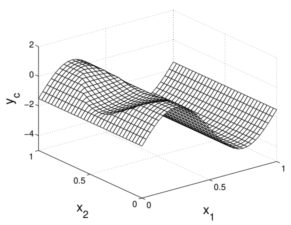


Figure 15: Multiplier η_h at $\mu = 10^{-8}$

6.2 Example 2

This example is constructed such that $\bar{y}$, $\bar{u}$ and $\bar{p}$ are trigonometric functions of the form $c\cos(\pi x_1)\cos(2\pi x_2)$. We choose $c = 1$ for $\bar{y}$ and $c = (-5\nu\pi^2)$ for $\bar{p}$. From the state equation and the optimality condition we get $\bar{u} = -\Delta\bar{y} + \bar{y} = \left(5\pi^2 + 1\right)\bar{y}$, and $u_d = \bar{u} + \frac{1}{\nu}\bar{p} = \bar{y}$, respectively.
Choosing $\hat{y} = 2\sin(2\pi x_1) - 1.5$, $\bar{\eta} = \max\{\hat{y} - \bar{y}, 0\}$, and $y_c = \min\{\hat{y}, \bar{y}\}$, the complementary slackness condition is fullfilled. All these functions are continuous. Therefore the adjoint equation can be treated in a classical way. From the adjoint equation we get $y_d = \Delta\bar{p} - \bar{p} + \bar{y} - \bar{\eta} = \left(\left(5\nu\pi^2\right)\left(5\pi^2+1\right)+1\right)\bar{y} - \bar{\eta}$. Figures 16–18 show the functions y_d y_c and η.





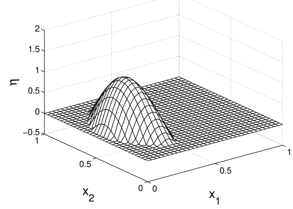


Figure 17: State constraints

Figure 16: Desired state y_d y_c

Figure 18: Multiplier η

The following figures show the numerical solutions for $\nu = 10^{-6}$ and $\lambda = 10^{-6}$.

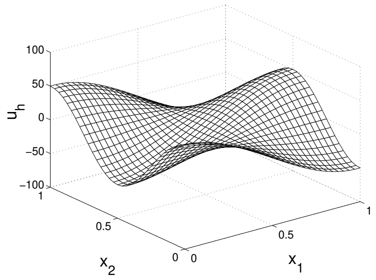


Figure 19: Control u_h

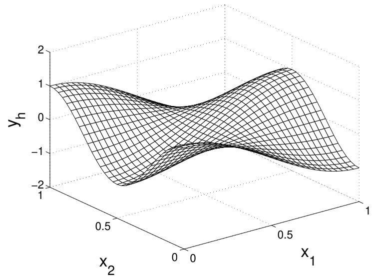


Figure 20: State y_h

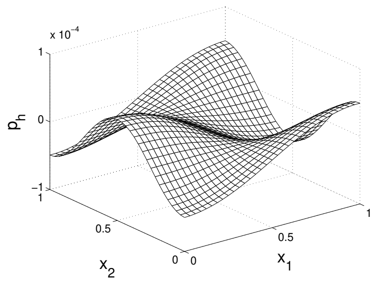


Figure 21: Adjoint state p_h

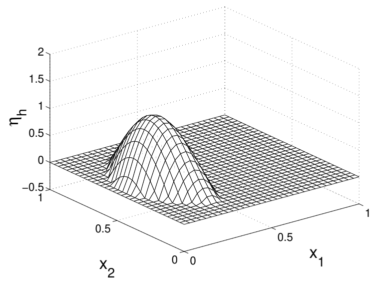


Figure 22: Multiplier η_h

Figures 23–26 show the differences $(\cdot)_h - (\cdot)_{\text{opt}}$ between the numerical solutions and the optimal solutions at $\nu = 10^{-6}$ and $\lambda = 10^{-6}$.

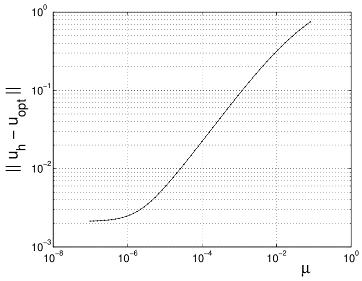


Figure 23: Error $\|u - u_{\text{opt}}\|$

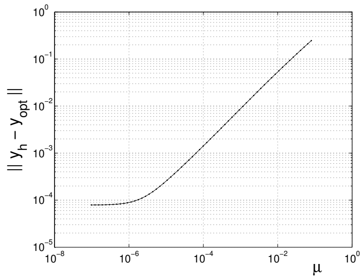


Figure 24: Error $\|y - y_{\text{opt}}\|$

u

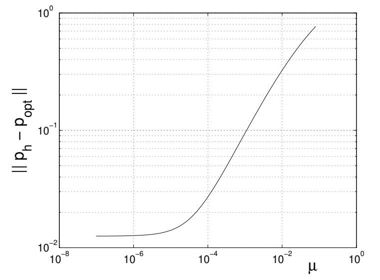


Figure 25: Error $\|p - p_{\mathrm{opt}}\|$

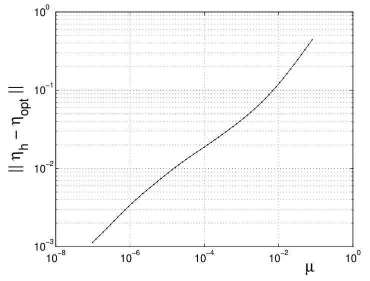


Figure 26: Error $\|\eta - \eta_{\mathrm{opt}}\|$

The following set of figueres shows the evolution of the control u_h along the central path.

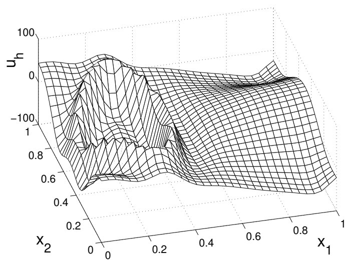


Figure 27: Control u_h at $\mu = 0.01$

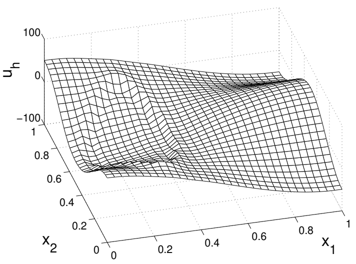


Figure 28: Control u_h at $\mu = 0.001$

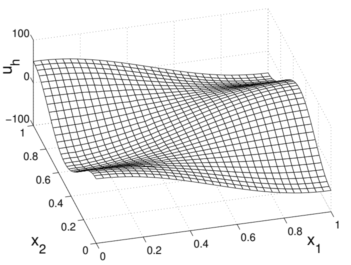


Figure 29: Control u_h at $\mu = 10^{-4}$

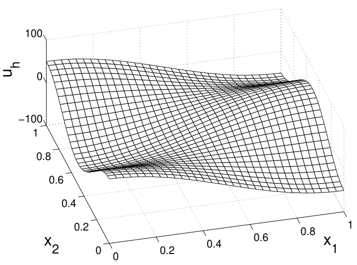


Figure 30: Control u_h at $\mu = 10^{-5}$

6.3 Example 3

In this example we consider the problem (PT) in the following setting:

$$y_d = \cos(\pi x_1)\cos(2\pi x_2) \tag{51}$$
$$y_c = \min\{6\sin(\pi x_1)\sin(\pi x_2) - 4, 1\} \tag{52}$$

and $u_d = 0$. Here, the optimal control $\bar{u}$ is unknown, just as the functions $\bar{y}$, $\bar{p}$ and the Lagrange-multiplier η. In figures 31 and 32 we show the functions y_d and y_c.

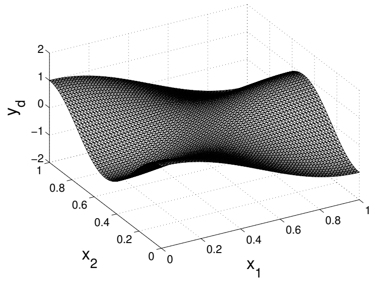


Figure 31: Desired state y_d

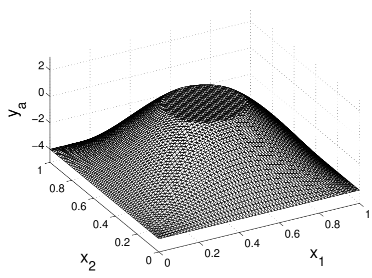


Figure 32: State contraints y_c

For our computations we choose $\nu = 10^{-6}$ and $\lambda = 10^{-16}$. The following set of figures shows the numerical solutions u_h, y_h, p_h and η_h.

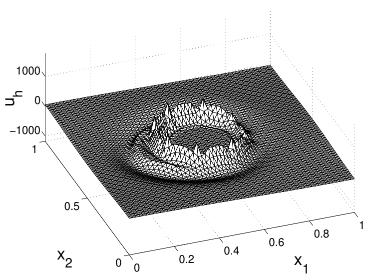


Figure 33: Control u_h

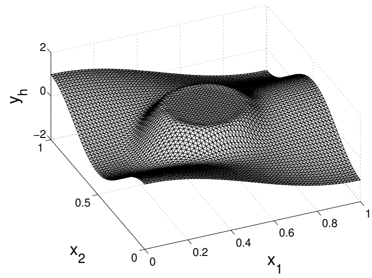


Figure 34: State y_h

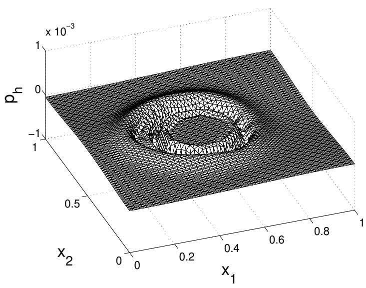


Figure 35: Adjoint state p_h

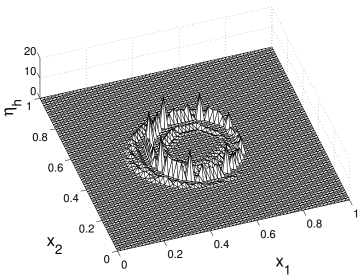


Figure 36: Lagrange multiplier η_h at $\nu = 10^{-6}$ and $\lambda = 10^{-16}$

Obviously the Lagrange-multiplier η_h shown in figure 36 tends to a measure with singular parts located on two circles in Ω, the points of nondifferentiability of $\bar{y}$.

References

[1] N. Arada and J. P. Raymond. Optimal control problems with mixed control-state constraints. *SIAM J. Control*, 39:1391–1407, 2000.

[2] M. Bergounioux and F. Tröltzsch. Optimal control of semilinear parabolic equations with state-constraints of bottleneck type. *ESAIM, Control, Optimisation and Calculus of Variations*, 4:595–608, 1999.

[3] E. Casas. Boundary control of semilinear elliptic equations with pointwise state constraints. *SIAM J. Control and Optimization*, 31:993–1006, 1993.

[4] T. Grund and A. Rösch. Optimal control of a linear elliptic equation with a supremum-norm functional. *Opt. Meth. Software*, 15:299–329, 2001.

[5] M. Haddou, M. Bergounioux, K. Kunisch, and M. Hintermüller. A comparison of a moreau–yosida-based active set strategy and interior point methods for constrained optimal control problems. *SIAM J. Optim.*, 11(2):495–521, 2000.

[6] C. Meyer, A. Rösch, and F. Tröltzsch. Optimal control of pdes with regularized pointwise state contraints. *Preprint-Reihe des Instituts für Mathematik, Technische Universität Berlin*, 2003.

[7] Y. Nesterov and A. Nemirovskii. *Interior-point polynomial methods in convex programming.* SIAM, Philadelphia, 1994.

[8] J. Renegar. *A mathematical view of interior-point methods in convex optimization.* SIAM, 2001.

[9] F. Tröltzsch. A minimum principle and a generalized bang-bang-principle for a distributed optimal control problem with constraints on the control and the state. *ZAMM*, 59:737–739, 1979.

[10] M. Ulbrich and S. Ulbrich. Superlinear convergence of affine-scaling interior-point newton methods for infinite-dimensional nonlinear problems with pointwise bounds. *SIAM J. Control Optim.*, 38:1938–1984, 2000.

[11] M. Ulbrich, S. Ulbrich, and M. Heinkenschloss. Global convergence of trust-region interior-point algorithms for infinite-dimensional nonconvex minimization subject to pointwise bounds. *SIAM J. Control and Optimization*, 37:731–764, 1999.

[12] M. Weiser. Interior point methods in function space. ZIB Report 03-35, 2003.

[13] M. Weiser and P. Deuflhard. The central path towards the numerical solution of optimal control problems. ZIB Report 01-12, 2001.

[14] M. Weiser and A. Schiela. Function space interior point methods for PDE constrained optimization. ZIB Report 04-27, 2004.