THE CONVERGENCE OF AN INTERIOR POINT METHOD FOR AN ELLIPTIC CONTROL PROBLEM WITH MIXED CONTROL-STATE CONSTRAINTS.*

UWE PRÜFERT†, FREDI TRÖLTZSCH†, AND MARTIN WEISER‡

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 method is are demonstrated by numerical examples.

Key words. interior point method, function space, optimal control, mixed control-state constraints, Lavrentiev regularization

AMS subject classifications. 90C51, 49J20, 65M15

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. [8] for discretized versions of elliptic control problems. Similarly, Grund and Rösch [7] 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 [15] 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 [14].

In [17, 16] primal-dual interior point methods have been analysed for ODE problems in the infinite dimensional function space setting and their computational realization by inexact pathfollowing methods has been suggested. In [18] 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.

Concerning 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 [12, 4, 3]. In [9], 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. Some results concerning the convergence of the solutions of the regularized problem to that of the origine state constrained can be found in [9].

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

‡Konrad-Zuse-Zentrum für Informationstechnik Berlin, Takustraße 7, D-14195 Berlin-Dahlem, Germany

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 are able to 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

$$\min J(y, u) := \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} \qquad (2.1)$$

subject to the elliptic boundary value problem

$$Ay = u \qquad \text{in } \Omega \qquad (2.2)$$

$$\partial_n y + \alpha y = 0 \qquad \text{on } \Gamma \qquad (2.3)$$

and to the pointwise mixed control-state constraints

$$y + \lambda u \geq y_c \qquad \text{a.e. in } \Omega. \qquad (2.4)$$

In an appendix, we briefly discus the pass to the limit $\lambda \downarrow 0$. 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 $\nu, \lambda \in \mathbb{R}$, $\lambda > 0$, 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. 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 $\|B\|_{L^p \to L^q}$ to denote the norm of a linear continuous operator $B : L^p(\Omega) \to L^q(\Omega)$. In the case $p = q = 2$, this norm is just denoted by $\|B\|$. For $\|B\|_{L^p \to L^p}$ we write $\|B\|_{L^p}$. 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 $v \leq 0$ means $v(x) \leq 0$ for a.a. $x \in \Omega$. In (2.3), ∂_n denotes the outward co-normal derivative at Γ. The main scope of our paper is to discuss the convergence of the standard interior point method for the problem (2.1)–(2.4) in function space. The simplest and well known idea of introducing this method is the elimination of the mixed control-state constraint $y + \lambda u \geq y_c$ by a logarithmic barrier function. We substitute (2.1)–(2.4) by the problem

$$\min J_\mu(y, u) := \frac{1}{2}\|y - y_d\|^2 + \frac{\nu}{2}\|u\|^2 - \mu \int_\Omega \ln\left((y + \lambda u - y_c)(x)\right) dx \qquad (2.5)$$

subject to $u \in L^2$ and

$$Ay = u \qquad \text{in} \quad \Omega \tag{2.6}$$

$$\partial_n y + \alpha y = 0 \qquad \text{on} \quad \Gamma. \tag{2.7}$$

In our analysis, we shall transform the state-constrained problem (2.1)–(2.4) to the problem (3.4)–(3.5) with control constraints. We have two reasons for this transformation: The analysis of this transformed problem is simpler than that for (2.1)–(2.4), 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 the transformed problem.

3. Existence of the central path. In this section we establish the existence of unique minima u_μ of (2.4)–(2.7) for all $\mu > 0$. To do this, we show the existence of a unique solution v_μ of the transformed problem (P$_\mu$) below. We refer to the mappings $\mu \mapsto u_\mu$ and $\mu \mapsto v_\mu$ as the central path, even though continuity is proved only in Section 4. First we recall some known facts about the state-equation (2.2)–(2.3).

THEOREM 3.1. *Under our assumptions, for all $u \in L^r(\Omega)$ with $r > \frac{N}{2}$, equation (2.2) has a unique solution $y \in H^1(\Omega) \cap C(\bar{\Omega})$. There is a constant $c(\Omega, r)$ such that*

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

We refer to Casas [5], Alibert and Raymond [1]. The reader might also consult the presentation of these results in the monography [13]. 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 (2.1)–(2.4) becomes equivalent to

$$\min \frac{1}{2}\|Su - y_d\|^2 + \frac{\nu}{2}\|u\|^2 \tag{3.1}$$

subject to

$$\lambda u + Su - y_c \ge 0 \quad \text{a.e. in } \Omega. \tag{3.2}$$

REMARK. S is known to be compact. By $\lambda > 0$, $-\lambda$ is not an eigenvalue of S. 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 $Ay + \frac{1}{\lambda}y = 0$ and $\partial_n y + \alpha y = 0$. By coercivity, this equation has only the trivial solution.

To transform (3.1)–(3.2) into a control-constrained problem, we substitute

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

By our assumption,

$$D := (S + \lambda I)^{-1} \tag{3.3}$$

exists as a continuous linear operator in L^2. After this substitution, (3.1)–(3.2) is equivalent to

$$(\text{P}) \qquad \min f(v) := \frac{1}{2}\|SDv - y_d\|^2 + \frac{\nu}{2}\|Dv\|^2 \tag{3.4}$$

3

subject to the constraints on the new control $v \in L^2$,

$$v - y_c \geq 0. \tag{3.5}$$

This simplification to a control-constrained problem can be made more explicit: By $v = Su + \lambda u = y + \lambda u$, we have $u = \lambda^{-1}(v - y)$. Inserting this in the state equation and in J, we see that (3.4)–(3.5) is equivalent to

$$\min \tilde{J}(y, v) = \frac{1}{2}\|y - y_d\|^2 + \frac{\nu}{2\lambda^2}\|v - y\|^2 \tag{3.6}$$

subject to

$$\begin{aligned} Ay + \frac{1}{\lambda}y &= \frac{1}{\lambda}v \\ \partial_n y + \alpha y &= 0 \end{aligned} \tag{3.7}$$

and $v - y_c \geq 0$.

For the special choice $D = I$, our analysis of (3.4)–(3.5) covers problems with simple bounds on the control v. The interior point method for (3.4)–(3.5) (or (2.1)–(2.3), respectively) is equivalent to solving

$$\min f_\mu(v) := f(v) - \mu \int_\Omega \ln(v(x) - y_c(x))\, dx. \tag{3.8}$$

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

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

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

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

is directionally differentiable in any direction $h \in L^\infty(\Omega)$, since $v_\varepsilon + t\, h - y_c \geq \varepsilon/2$ for sufficiently small t. In this case,

$$\phi'(v_\varepsilon)h = \int \frac{\mu}{v_\varepsilon(x) - y_c} h\, dx.$$

Suppose now that (3.8) admits a solution $\tilde{v}_\varepsilon = v_\varepsilon(\mu) \in L^2$ satisfying $v_\varepsilon(x) - y_c(x) \geq \varepsilon > 0$. Then we get from the differentiability properties mentioned above

$$f_\mu(v_\varepsilon) = f'(v_\varepsilon) - \phi'(v_\varepsilon) = 0, \tag{3.9}$$

since in this case $\tilde{v} + th, h \in L^\infty$, has distance $\varepsilon/2$ for all small t. Therefore, it holds

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

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

$$\eta(x) := \frac{\mu}{v_\varepsilon(x) - y_c(x)}. \tag{3.10}$$

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 (2.1)–(2.4) as $\mu \downarrow 0$. However, we have to show that (3.8) is solvable, i.e. that the central path exists. Notice that, by $u = D\, v$, (3.8) and (2.4)–(2.7) are equivalent.

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

$$(\mathrm{P}_\mu^\varepsilon) \qquad \min_{v(x) - y_c(x) \geq \varepsilon} f_\mu(v),$$

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 $u_\mu = (\lambda I + S)^{-1} v_\mu$ of (2.4)–(2.7) 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

$$
\begin{aligned}
f_\mu(v) &= \frac{1}{2} \|SDv - y_d\|_{L^2}^2 + \frac{\nu}{2} \|Dv\|_{L^2}^2 - \mu \int_\Omega \ln(v - y_c)\, dx \\
&\geq \frac{\nu}{2} \|Dv\|_{L^2}^2 - \mu \int_\Omega (v - y_c)\, dx \qquad\qquad (3.11) \\
&\geq \frac{\nu \delta_0}{2} \|v\|_{L^2}^2 - \mu \|v - y_c\|_{L^1} \geq \frac{\nu \delta_0}{2} \|v\|_{L^2}^2 - \mu\, c \, \|v - y_c\|_{L^2}. \qquad (3.12)
\end{aligned}
$$

In (3.11), we have used $\ln(x) < x$ for all $x > 0$. With $\delta_0 = \|S + \lambda I\|^{-2} > 0$ we observe that $\|v\| \to \infty$ 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 $C_\varepsilon \subset L^2$ defined by

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

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 infimum 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 $C_\varepsilon \cap \{v \in L^2 \,|\, \|v\| \leq c_v\}$, the functional f_μ is bounded, hence

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

is finite. Here and in what follows, we suppress for a while the dependence of the problem and its solutions on μ.

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

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

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

We recall problem (P_μ^ε),

$$\min f_\mu(v) := \frac{1}{2}\|SDv - y_d\|^2 + \frac{\nu}{2}\|Dv\|^2 - \mu \int_\Omega \ln(v - y_c)\,dx$$

$$v(x) - y_c(x) \geq \varepsilon \quad \text{a.e. in } \Omega.$$

To shorten the notation, we continue to denote the optimal solution $v_\varepsilon(\mu)$ of (P_μ^ε) by v_ε. 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

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

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

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

In terms of our transformation, this can be written as

$$\left(D^*S^*(SDv_\varepsilon - y_d) + \nu D^*Dv_\varepsilon - \frac{\mu}{v_\varepsilon - y_c}\,,\, v - v_\varepsilon\right) \geq 0 \quad \forall v \in C_\varepsilon. \qquad (3.13)$$

Define $p_\varepsilon := D^*S^*(SDv_\varepsilon - y_d)$. Then we can re-write (3.13) as

$$\left(p_\varepsilon + \nu D^*Dv_\varepsilon - \frac{\mu}{v_\varepsilon - y_c}\,,\, v - v_\varepsilon\right) \geq 0 \quad \forall v \in C_\varepsilon. \qquad (3.14)$$

In terms of a PDE, p_ε is given by

$$Ap_\varepsilon + \frac{1}{\lambda}p_\varepsilon = y_\varepsilon - y_d - \frac{\nu}{\lambda^2}(v_\varepsilon - y_\varepsilon)$$

$$\partial_n p_\varepsilon + \alpha p_\varepsilon = 0,$$

where y_ε solves (3.7). Note, this is the adjoint to the problem (3.6)–(3.7). We shall show that $\|p_\varepsilon\|_\infty$ is bounded, independently of ε:
The operator S is 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 $z = Sw + \lambda Iw$. It follows $\lambda w = z - Sw = z - SDz$ and $w = \lambda^{-1}z - \lambda^{-1}SDz$. Therefore D admits the form

$$D = \lambda^{-1}(I - SD). \qquad (3.15)$$

From this representation we easily get the additional regularity property $D : L^\infty \to L^\infty$, continuously. This is also visible from (3.6)–(3.7) since $v \mapsto y$ is continuous from L^∞ to L^∞ and therefore also $v \mapsto u = \lambda^{-1}(v - y(v)) : L^\infty \to L^\infty$.

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

$$\|p_\varepsilon\|_\infty = \|D^*S^*(SDv_\varepsilon - y_d)\|_\infty$$
$$\leq \|D^*\|_{L^\infty}\|S^*\|_{L^2 \to L^\infty}\|SDv_\varepsilon - y_d\| \leq c_p, \qquad (3.16)$$

6

where c_p does not depend on ε, since $\|SDv_\varepsilon - y_d\| \le \|S\|_{L^2}\|D\|_{L^2}\|c_v\| + \|y_d\|$. Next we evaluate (3.13). Let us define the sets

$$M_+(\varepsilon) := \left\{ x \in \Omega \,\Big|\, p_\varepsilon(x) + \nu(D^*Dv_\varepsilon)(x) - \frac{\mu}{v_\varepsilon(x) - y_c(x)} > 0 \right\}$$

$$M_0(\varepsilon) := \left\{ x \in \Omega \,\Big|\, p_\varepsilon(x) + \nu(D^*Dv_\varepsilon)(x) - \frac{\mu}{v_\varepsilon(x) - y_c(x)} = 0 \right\}.$$

Due to (3.13), $M_+(\varepsilon) \cup M_0(\varepsilon)$ cover Ω up to a set of measure zero. Clearly, the variational inequality (3.13) implies $v_\varepsilon(x) - y_c = \varepsilon$ for almost all $x \in M_+(\varepsilon)$.

THEOREM 3.4. *There exist constants $a, b > 0, \varepsilon_M$ such that the set $M_+(\varepsilon)$ has measure zero for all $0 < \varepsilon < \varepsilon_M$.*

Proof. For almost all $x \in M_+(\varepsilon)$, the constraint is active, i.e. $v_\varepsilon(x) - y_c(x) = \varepsilon$. Thus, by (3.16), we have for almost all $x \in M_+(\varepsilon)$

$$c_p + \nu(D^*Dv_\varepsilon)(x) - \frac{\mu}{\varepsilon} \ge p_\varepsilon(x) + \nu(D^*Dv_\varepsilon)(x) - \frac{\mu}{v_\varepsilon(x) - y_c(x)} > 0. \qquad (3.17)$$

By (3.15),

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

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

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

Almost everywhere on $M_+(\varepsilon)$ it holds $v_\varepsilon(x) = y_c(x) + \varepsilon$, 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 (3.17), Theorem 3.3 yields

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

Clearly, the right hand side tends to infinity as $\varepsilon \downarrow 0$ while the left hand side remains bounded. Therefore, the inequality cannot be satisfied for small ε.

Solving this quadratic inequality for ε, we get from

$$\varepsilon^2 + \left(c_p \frac{\lambda^2}{\nu} + \|y_c\|_\infty + c_v\lambda^2 \cdot \|K\|_{L^2 \to L^\infty} \right)\varepsilon - \frac{\lambda^2}{\nu}\mu > 0$$

$$\varepsilon > -\frac{1}{2}\left(c_p \frac{\lambda^2}{\nu} + \|y_c\|_\infty + \|K\|_{L^2 \to L^\infty} c_v\lambda^2 \right)$$

$$+ \sqrt{\frac{1}{4}\left(c_p \frac{\lambda^2}{\nu} + \|y_c\|_\infty + \|K\|_{L^2 \to L^\infty} c_v\lambda^2 \right)^2 + \frac{\lambda^2}{\nu}\mu}.$$

Setting

$$a = \frac{1}{2}\left(c_p \frac{\lambda^2}{\nu} + \|y_c\|_\infty + \|K\|_{L^2 \to L^\infty} c_v\lambda^2 \right) \qquad (3.18)$$

and

$$b = \frac{\lambda^2}{\nu} a^{-2} \qquad (3.19)$$

we can write this as $\varepsilon > a(\sqrt{1 + b\mu} - 1)$ with $a, b > 0$. For smaller ε, $M_+(\varepsilon)$ must therefore have measure zero. $\qquad\blacksquare$

COROLLARY 3.5. *For all $0 < \varepsilon < \varepsilon_M$, the solution $v_\varepsilon(\mu)$ 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 $v_\varepsilon(\mu)$ 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, $v_\varepsilon(\mu)$ is the unique solution $v(\mu)$ of (P_μ). $\qquad\blacksquare$

COROLLARY 3.6. *There exists a constant $c_\mu > 0$ such that for $\mu \leq 1$ the unique solution v_μ of (3.8) satisfies $v_\mu \geq y_c + c_\mu \mu$ a.e. on Ω.*

Proof. Let v_μ be the solution of (P_μ). Then an $\varepsilon > 0$ exists such that v_μ is a solution of $(\mathrm{P}_\mu^\varepsilon)$ too. For that ε it holds $v_\mu \geq y_c + \varepsilon$. Choosing $c_\mu = \frac{\varepsilon}{\mu}$ yields $v_\mu \geq y_c + c_\mu \mu$. $\qquad\blacksquare$

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 (3.8) can be characterized by (3.9) as

$$F(v_\mu; \mu) = (D^*S^*SD + \nu D^*D)v_\mu - D^*S^*y_d - \frac{\mu}{v_\mu - y_c} = 0 \quad \text{a.e. on } \Omega \quad (4.1)$$

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

$$\partial_v F(v; \mu) = (D^*S^*SD + \nu D^*D) + \frac{\mu}{(v - y_c)^2} \qquad (4.2)$$

$$= (D^*S^*SD + \nu K) + \left(\frac{\nu}{\lambda^2} + \frac{\mu}{(v - y_c)^2} \right)$$

$$= \bar{K} + \left(\frac{\nu}{\lambda^2} + \frac{\mu}{(v - y_c)^2} \right), \qquad (4.3)$$

where

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

is a bounded operator from L^2 to L^∞. From (4.2) and (3.3) 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

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

8

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

$$\|\partial_v F(v; \mu)^{-1}\| \leq \frac{1}{\nu}(\|S\| + |\lambda|)^2. \tag{4.4}$$

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 (4.4), 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

$$\|\xi\| \leq \frac{1}{\nu}(\|S\| + |\lambda|)^2 \|z\| \leq \frac{\sqrt{|\Omega|}}{\nu}(\|S\| + |\lambda|)^2 \|z\|_\infty. \tag{4.5}$$

Now we have by (4.3)

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

and hence by (4.5)

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

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 $\|\partial_v F(v; \mu)^{-1}\|_{L^\infty \to L^\infty} \leq c_i$. $\square$

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 (3.1).

THEOREM 4.3. *For $\mu \to 0$, the central path converges towards a KKT point v_0 of (3.1). There exists a constant $c_0 < \infty$ such that the following error estimate holds for all $\mu \leq 1$:*

$$\|v_0 - v_\mu\|_{L^\infty} \leq c_0\sqrt{\mu} \tag{4.6}$$

Proof. First we will establish an L^2-bound on $v'_\mu = -\partial_v F(v_\mu; \mu)^{-1}\partial_\mu F(v_\mu; \mu)$ and infer an L^∞-bound from that. From this we will determine the existence 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 introduce the diagonal preconditioner

$$z_\mu = \sqrt{\frac{\nu}{\lambda^2} + \frac{\mu}{(v_\mu - y_c)^2}} \tag{4.7}$$

9

and write (4.7) as

$$z_\mu v_\mu' = \left(z_\mu^{-1} \left(\bar{K} + \frac{\nu}{\lambda^2} + \frac{\mu}{(v_\mu - y_c)^2} \right) z_\mu^{-1} \right)^{-1} z_\mu^{-1} \frac{1}{v_\mu - y_c}$$

$$= \left(z_\mu^{-1} \bar{K} z_\mu^{-1} + I \right)^{-1} \left(\frac{\nu(v_\mu - y_c)^2}{\lambda^2} + \mu \right)^{-1/2}.$$

Since $z_\mu^{-1} \bar{K} z_\mu^{-1}$ is positive semidefinite, we may conclude that $\|z_\mu v_\mu'\|_{L^2} \leq \sqrt{|\Omega|/\mu}$. From $z_\mu \geq \sqrt{\nu}/\lambda$ a.e. we finally obtain

$$\|v_\mu'\|_{L^2} \leq \sqrt{\frac{\nu |\Omega|}{\mu \lambda}}.$$

(ii) L^∞-estimates. Using the splitting (4.3) to move the coupling term $\bar{K} v_\mu'$ in $\partial_v F(v_\mu; \mu) v_\mu' = -\partial_\mu F(v_\mu; \mu)$ to the right hand side, and the fact that

$$ax + \frac{b}{x} \geq 2\sqrt{ab}$$

holds for arbitrary $a, b, x > 0$, we obtain

$$\|v_\mu'\|_{L^\infty} \leq \left\| \left(\frac{\nu}{\lambda^2} + \frac{\mu}{(v_\mu - y_c)^2} \right)^{-1} \frac{1}{v_\mu - y_c} \right\|_{L^\infty} + \left\| \left(\frac{\nu}{\lambda^2} + \frac{\mu}{(v_\mu - y_c)^2} \right)^{-1} \bar{K} v_\mu' \right\|_{L^\infty}$$

$$\leq \left\| \left(2\sqrt{\frac{\nu\mu}{\lambda^2}} \right)^{-1} \right\|_{L^\infty} + \left\| \left(\frac{\nu}{\lambda^2} + \frac{\mu}{(v_\mu - y_c)^2} \right)^{-1} \right\|_{L^\infty} \|\bar{K}\|_{L^2 \to L^\infty} \|v_\mu'\|$$

and infer

$$\|v_\mu'\|_{L^\infty} \leq \left\| \frac{\lambda}{2\sqrt{\nu\mu}} \right\|_{L^\infty} + \frac{\lambda^2}{\nu} \|\bar{K}\|_{L^2 \to L^\infty} \sqrt{\frac{\nu|\Omega|}{\mu\lambda}} \leq \frac{c_0}{\sqrt{\mu}}$$

for some $c_0 < \infty$.

(iii) *Distance to the limit point.* The distance between two points on the central path is therefore bounded by

$$\|v_{\mu_1} - v_{\mu_2}\|_{L^\infty} \leq \int_{\mu_1}^{\mu_2} \|v_\mu'\|_{L^\infty} \, d\mu \leq \frac{c_0}{2} (\sqrt{\mu_2} - \sqrt{\mu_1}). \tag{4.8}$$

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 (4.6).

(iv) *First order necessary conditions.* Recalling the Lagrange multiplier approximations η_μ from (3.10) we write (3.9) 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 $\eta_\mu \geq 0$ 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. $\eta_0 \geq 0$ 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 (3.4). □

10

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_μ defined in (4.7), which is closely connected to the energy norms used in the theory of self-concordant barrier functionals [10, 11].

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 $\|\Delta v_h^k - \Delta v^k\|_{\mu^{k+1}} \leq \delta \|\Delta v^k\|_{\mu^{k+1}}$
    $v^{k+1} = v^k + \Delta v_h^k$

Note that the accuracy matching in Algorithm 5.1 will require mesh refinement as $\mu \to 0$. Alternatively, on a fixed discretization the algorithm can be performed only up to some $\mu_{\min} > 0$ while still meeting the accuracy requirement.

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 formulate the main result, the proof of which is postponed to the end of this section.

THEOREM 5.2. *Let a tolerance* $\vartheta < 1/(18c_z)$, $\mu^0 > 0$, *and an initial iterate* v^0 *with* $\|v^0 - v_{\mu^0}\|_{\mu^0} \leq 2\vartheta \sqrt{\mu^0}$ *be given. Choose* $\delta \leq 1/4$ *and a reduction factor satisfying*

$$1 - \sigma \leq \frac{\vartheta}{3\vartheta(c_z + 1/2) + c_z}.$$

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,*

$$\left\|v^k - v_{\mu^k}\right\|_{\mu^k} \leq 2\vartheta \sqrt{\mu^k} \quad and \quad \left\|v^k - v_0\right\|_{\mu^k} \leq (c_0 + 2\vartheta)\sigma^{k/2}\sqrt{\mu^0}.$$

We stress that this conceptual algorithm is deliberately designed to be simplistic in order to facilitate convergence analysis. We do not recommend to use it for actual computation. First, the admissible choice of parameters σ and v_0 depends on the problem specific constant c_z, which will usually be unavailable in actual computation. Second, the bounds given for ϑ and σ are global worst-case bounds that will be

11

unnecessarily restrictive locally. Adaptive stepsize and accuracy selection will result in a far more efficient algorithm.

The proof of Theorem 5.2 will require the usual ingredients for Newton convergence theorems which we will provide now: boundedness of $\partial_v F^{-1}$, Lipschitz continuity of the local norms, and Lipschitz continuity of $\partial_v F$. First we turn to $\partial_v F^{-1}$ and derive the analogue of Lemma 4.1 for the scaled norm.

LEMMA 5.3. *There is some constant $1 \leq c_z < \infty$ independent of μ, such that*

$$\|\partial_v F(v; \mu)^{-1}\zeta\|_\mu \leq c_z \|z_\mu^{-1}\zeta\|_{L^\infty} \tag{5.1}$$

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

Proof. From $z_\mu \geq \sqrt{\mu}/(v_\mu - y_c)$ we see that

$$\left\|\frac{v - v_\mu}{v_\mu - y_c}\right\|_{L^\infty} = \left\|\frac{v - v_\mu}{\sqrt{\mu}}\sqrt{\frac{\mu}{(v_\mu - y_c)^2}}\right\|_{L^\infty} \leq \left\|z_\mu\frac{v - v_\mu}{\sqrt{\mu}}\right\|_{L^\infty} = \frac{\|v - v_\mu\|_\mu}{\sqrt{\mu}} \leq \vartheta$$

for $v \in B_\mu(v_\mu; \vartheta\sqrt{\mu})$. For almost all $x \in \Omega$ we therefore have $-(v - v_\mu) \leq \vartheta(v_\mu - y_c)$, which implies

$$v \geq (1 - \vartheta)v_\mu + \vartheta y_c = (1 - \vartheta)(v_\mu - y_c) + y_c \geq (1 - \vartheta)c_\mu\mu + y_c > y_c \tag{5.2}$$

due to Corollary 3.6. Lemma 4.1 now provides the invertibility of $\partial_v F(v; \mu)$. As in the proof of Theorem 4.3, we have $z_\mu \partial_v F(v; \mu)^{-1} z_\mu = (z_\mu^{-1}\bar{K}z_\mu^{-1} + I)^{-1}$ and hence the L^2-estimate

$$\|z_\mu \partial_v F(v; \mu)^{-1}\zeta\|_{L^2} = \|(z_\mu^{-1}\bar{K}z_\mu^{-1} + I)^{-1}z_\mu^{-1}\zeta\|_{L^2} \leq \|z_\mu^{-1}\zeta\|_{L^2}. \tag{5.3}$$

Defining $\phi = \partial_v F(v; \mu)^{-1}\zeta$ we have $(\bar{K} + z_\mu^2)\phi = \zeta$ and

$$\|z_\mu \partial_v F(v; \mu)^{-1}\zeta\|_{L^\infty} = \|z_\mu\phi\|_{L^\infty} = \|z_\mu^{-1}(\zeta - \bar{K}\phi)\|_{L^\infty}$$

$$\leq \|z_\mu^{-1}\zeta\|_{L^\infty} + \frac{\lambda}{\sqrt{\nu}}\|\bar{K}\|_{L^2 \to L^\infty}\|\phi\|_{L^2} \leq \|z_\mu^{-1}\zeta\|_{L^\infty} + \frac{\lambda^2}{\nu}\|\bar{K}\|_{L^2 \to L^\infty}\|z_\mu\phi\|_{L^2}.$$

Using (5.3) yields

$$\|z_\mu \partial_v F(v; \mu)^{-1}\zeta\|_{L^\infty} \leq \left(1 + \frac{\lambda^2\|\bar{K}\|_{L^2 \to L^\infty}}{\nu\sqrt{|\Omega|}}\right)\|z_\mu^{-1}\zeta\|_{L^\infty}$$

and establishes the constant c_z. □

Next we prove Lipschitz continuity of the scaled norms.

LEMMA 5.4. *For all $v \in L^\infty$ and $0 < \sigma \leq 1$,*

$$\|v\|_{\sigma\mu} \leq \sigma^{-c_z}\|v\|_\mu \tag{5.4}$$

holds. Moreover, the derivative of the central path is bounded by $\|v_\mu'\|_\mu \leq c_z/\sqrt{\mu}$.

Proof. We begin with estimating the derivative of the central path in the scaled norm. Lemma 5.3 applied to $v_\mu' = -\partial_v F(v_\mu; \mu)^{-1}\partial_\mu F(v_\mu; \mu)$ results in

$$\|v_\mu'\|_\mu \leq c_z\|z_\mu^{-1}\partial_\mu F(v_\mu; \mu)\|_{L^\infty} = c_z\|z_\mu^{-1}(v_\mu - y_c)^{-1}\|_{L^\infty} \leq \frac{c_z}{\sqrt{\mu}}. \tag{5.5}$$

We proceed with bounding the expression

$$\phi(\sigma) = \frac{v_\mu - y_c}{v_{\sigma\mu} - y_c}.$$

Note that since $\mu \mapsto v_\mu$ is a differentiable mapping from $\mathbb{R}_+$ to L^∞, ϕ maps $]0,1]$ differentiably into L_∞. Using $z_{\tau\mu} \geq \sqrt{\tau\mu}/(v_{\tau\mu} - y_c)$ and (5.5), we start with

$$\|\phi'(\tau)\|_{L_\infty} = \left\|\frac{v_\mu - y_c}{(v_{\tau\mu} - y_c)^2} v'_{\tau\mu}\mu\right\|_{L^\infty} \leq \left\|\frac{v_\mu - y_c}{v_{\tau\mu} - y_c}\right\|_{L^\infty} \left\|\frac{\sqrt{\tau\mu}}{v_{\tau\mu} - y_c} v'_{\tau\mu}\right\|_{L^\infty} \frac{\mu}{\sqrt{\tau\mu}}$$

$$\leq \|\phi(\tau)\|_{L_\infty} \|v'_{\tau\mu}\|_{\tau\mu} \frac{\mu}{\sqrt{\tau\mu}} \leq \|\phi(\tau)\|_{L_\infty} \frac{c_z}{\sqrt{\tau\mu}} \frac{\mu}{\sqrt{\tau\mu}} \leq \|\phi(\tau)\|_{L_\infty} \frac{c_z}{\tau}$$

for $\sigma \leq \tau \leq 1$. From this we infer

$$\|\phi(\sigma)\|_{L_\infty} \leq \|\phi(1)\|_{L_\infty} + \int_\sigma^1 \|\phi'(\tau)\|_{L_\infty}\, d\tau \leq 1 + \int_\sigma^1 \|\phi(\tau)\|_{L_\infty} \frac{c_z}{\tau}\, d\tau.$$

The Bellmann-Gronwall lemma now yields $\|\phi(\sigma)\|_{L_\infty} \leq \sigma^{-c_z}$ for $0 < \sigma \leq 1$. Next we estimate

$$\left\|\frac{z_{\sigma\mu}}{z_\mu}\right\|_{L^\infty} = \left\|\frac{\frac{\nu}{\lambda^2} + \frac{\sigma\mu}{(v_{\sigma\mu} - y_c)^2}}{\frac{\nu}{\lambda^2} + \frac{\mu}{(v_\mu - y_c)^2}}\right\|_{L^\infty}^{1/2} \leq \max\left(1, \sqrt{\sigma}\left\|\frac{v_\mu - y_c}{v_{\sigma\mu} - y_c}\right\|_{L^\infty}\right), \qquad (5.6)$$

the case depending on whether

$$\frac{\sigma\mu}{(v_{\sigma\mu} - y_c)^2} \leq \frac{\mu}{(v_\mu - y_c)^2}$$

holds. Dropping the factor $\sqrt{\sigma} \leq 1$ for simplicity,

$$\|v\|_{\sigma\mu} = \left\|z_\mu \frac{z_{\sigma\mu}}{z_\mu} v\right\|_{L^\infty} \leq \max\left(1, \sigma^{-c_z}\right)\|v\|_\mu$$

proves the claim. $\square$

Finally, we prove Lipschitz continuity of $\partial_v F$.

LEMMA 5.5. *For all $0 < \vartheta < 1$, the Lipschitz condition*

$$\left\|\partial_v F(v;\mu)^{-1}(\partial_v F(v;\mu) - \partial_v F(\hat{v};\mu))(v - \hat{v})\right\|_\mu \leq \frac{2c_z}{(1-\vartheta)^3 \sqrt{\mu}}\|v - \hat{v}\|_\mu^2 \qquad (5.7)$$

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

Proof. Using Lemma 5.3 and in view of the representation (4.2) of $\partial_v F$, we have

$$\left\|\partial_v F(v;\mu)^{-1}(\partial_v F(v;\mu) - \partial_v F(\hat{v};\mu))(v - \hat{v})\right\|_\mu$$

$$\leq c_z\left\|z_\mu^{-1}(\partial_v F(v;\mu) - \partial_v F(\hat{v};\mu))(v - \hat{v})\right\|_{L^\infty}$$

$$= c_z\left\|z_\mu^{-1}\left(\frac{\mu}{(v - y_c)^2} - \frac{\mu}{(\hat{v} - y_c)^2}\right)(v - \hat{v})\right\|_{L^\infty}.$$

13
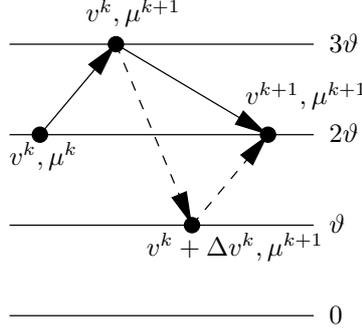
FIG. 5.1. *Errors of the iterates during the proof of Theorem 5.2.*

Using the fact that the Lipschitz constant of x^{-2} for $x \geq a > 0$ is given by $2a^{-3}$, and that $v - y_c \geq (1 - \vartheta)(v_\mu - y_c)$ for $v \in B_\mu(v_\mu, \vartheta\sqrt{\mu})$ due to (5.2), we can proceed with

$$\left\| \partial_v F(v; \mu)^{-1}(\partial_v F(v; \mu) - \partial_v F(\hat{v}; \mu))(v - \hat{v}) \right\|_\mu$$

$$\leq c_z \left\| z_\mu^{-1} \mu \frac{2(v - \hat{v})}{((1 - \vartheta)(v_\mu - y_c))^3}(v - \hat{v}) \right\|_{L^\infty}$$

$$= \frac{2c_z}{(1 - \vartheta)^3} \left\| \frac{\mu}{z_\mu^3(v_\mu - y_c)^3} z_\mu^2 (v - \hat{v})^2 \right\|_{L^\infty}$$

$$\leq \frac{2c_z}{(1 - \vartheta)^3} \left\| \frac{\mu}{z_\mu^3(v_\mu - y_c)^3} \right\|_{L^\infty} \|v - \hat{v}\|_\mu^2$$

$$\leq \frac{2c_z}{\sqrt{\mu}(1 - \vartheta)^3} \|v - \hat{v}\|_\mu^2,$$

where the last inequality is a direct consequence of (4.7). □

Now we are prepared to prove Theorem 5.2.

Proof. First we give an outline of the proof (see Fig. 5.1). We use induction and assume that $\|v^k - v_{\mu^k}\|_{\mu^k} \leq 2\vartheta\sqrt{\mu^k}$. Decreasing the homotopy parameter μ^k by a factor of σ will lead to an error bound $\|v^k - v_{\mu^{k+1}}\|_{\mu^{k+1}} \leq 3\vartheta\sqrt{\mu^{k+1}}$. Then the inexact Newton corrector reduces the error again to $\|v^k - v_{\mu^{k+1}}\|_{\mu^{k+1}} \leq 2\vartheta\sqrt{\mu^{k+1}}$. We show this by deriving an error bound $\|v^k - v_{\mu^{k+1}}\|_{\mu^{k+1}} \leq \vartheta\sqrt{\mu^{k+1}}$ for the exact Newton corrector and adding an inexactness bounded by $\vartheta\sqrt{\mu^{k+1}}$.

(i) To begin with, we split the error as follows:

$$\|v^k - v_{\mu^{k+1}}\|_{\mu^{k+1}} \leq \|v^k - v_{\mu^k}\|_{\mu^{k+1}} + \|v_{\mu^k} - v_{\mu^{k+1}}\|_{\mu^{k+1}}$$

$$\leq \|v^k - v_{\mu^k}\|_{\mu^{k+1}} + \int_{\mu^{k+1}}^{\mu^k} \|v_\tau'\|_{\mu^{k+1}} \, d\tau$$

In view of $\mu^{k+1} = \sigma\mu^k$, application of Lemma 5.4 and the induction assumption leads to

$$\|v^k - v_{\mu^{k+1}}\|_{\mu^{k+1}} \leq \sigma^{-c_z} \|v^k - v_{\mu^k}\|_{\mu^k} + \int_{\mu^{k+1}}^{\mu^k} (\mu^{k+1}/\tau)^{-c_z} \|v_\tau'\|_\tau \, d\tau$$

$$\leq \sigma^{-c_z} 2\vartheta\sqrt{\mu^k} + \int_{\mu^{k+1}}^{\mu^k} (\mu^{k+1}/\tau)^{-c_z} \frac{c_z}{\sqrt{\tau}} \, d\tau. \qquad (5.8)$$

14

The integral evaluates to

$$\frac{c_z}{(\mu^{k+1})^{c_z}} \int_{\sigma\mu^k}^{\mu^k} \tau^{c_z-1/2}\,d\tau = \frac{c_z}{(\sigma\mu^k)^{c_z}} \frac{(\mu^k)^{c_z+1/2} - (\sigma\mu^k)^{c_z+1/2}}{c_z + 1/2}$$

$$= \frac{c_z}{c_z + 1/2}\sigma^{-c_z}(1 - \sigma^{c_z+1/2})\sqrt{\mu^k}$$

$$\le c_z\sigma^{-c_z}(1 - \sigma)\sqrt{\mu^k}. \tag{5.9}$$

Inserting (5.8) into (5.9) leads to the estimate

$$\|v^k - v_{\mu^{k+1}}\|_{\mu^{k+1}} \le \sigma^{-c_z}(2\vartheta + c_z(1 - \sigma))\sqrt{\mu^k}$$

$$= \sigma^{-(c_z+1/2)}(2\vartheta + c_z(1 - \sigma))\sqrt{\mu^{k+1}}. \tag{5.10}$$

Next we rearrange the assumption on σ:

$$1 - \sigma \le \frac{\vartheta}{3\vartheta(c_z + 1/2) + c_z} \Leftrightarrow (3\vartheta(c_z + 1/2) + c_z)(1 - \sigma) \le \vartheta$$

$$\Leftrightarrow c_z(1 - \sigma) \le \vartheta - 3\vartheta(c + 1/2)(1 - \sigma)$$

Adding 2ϑ on both sides leads to

$$2\vartheta + c_z(1 - \sigma) \le 3\vartheta\big(1 + (c_z + 1/2)(\sigma - 1)\big) \le 3\vartheta\sigma^{c_z+1/2}. \tag{5.11}$$

Combining (5.10) and (5.11) yields $\|v^k - v_{\mu^{k+1}}\|_{\mu^{k+1}} \le 3\vartheta\sqrt{\mu^{k+1}}$.

(ii) Since $c_z \ge 1$, the assumption $\vartheta \le 1/(18c_z)$ implies

$$\frac{3\vartheta}{(1 - 3\vartheta)^3} \le 6\vartheta \le \frac{1}{3c_z}$$

and thus

$$\|v^k - v_{\mu^{k+1}}\|_{\mu^{k+1}} \le \frac{2}{3}\frac{(1 - 3\vartheta)^3}{2c_z}\sqrt{\mu^{k+1}}.$$

This is two third of the Lipschitz constant provided by Lemma 5.5 for the ball $B_{\mu^{k+1}}(v_{\mu^{k+1}}, 3\vartheta\sqrt{\mu^{k+1}})$. Thus, the conditions for local convergence of the *exact* Newton corrector as considered in the refined Newton-Mysovskii theorem [6] are satisfied. The error of the next iterate $v^k + \Delta v^k$ is bounded by

$$\|v^k + \Delta v^k - v_{\mu^{k+1}}\|_{\mu^{k+1}} \le \frac{1}{2}\frac{2c_z}{(1 - 3\vartheta)^3\sqrt{\mu^{k+1}}}\|v^k - v_{\mu^{k+1}}\|_{\mu^{k+1}}^2$$

$$\le \frac{1}{3}\|v^k - v_{\mu^{k+1}}\|_{\mu^{k+1}} \le \vartheta\sqrt{\mu^{k+1}}.$$

(iii) The length of the Newton step Δv^k by

$$\|\Delta v^k\|_{\mu^{k+1}} \le \|v^k + \Delta v^k - v_{\mu^{k+1}}\|_{\mu^{k+1}} + \|v^k - v_{\mu^{k+1}}\|_{\mu^{k+1}} \le 4\vartheta\sqrt{\mu^{k+1}},$$

and the error of the *inexact* iterate v^{k+1} is bounded by

$$\|v^{k+1} - v_{\mu^{k+1}}\|_{\mu^{k+1}} \le \|v^k + \Delta v^k - v_{\mu^{k+1}}\|_{\mu^{k+1}} + \delta\|\Delta v^k\|_{\mu^{k+1}}$$

$$\le (1 + 4\delta)\vartheta\sqrt{\mu^{k+1}}.$$

With the accuracy matching $\delta \leq 1/4$ we obtain $\|v^{k+1} - v_{\mu^{k+1}}\|_{\mu^{k+1}} \leq 2\vartheta \sqrt{\mu^{k+1}}$, which completes the induction.

(iv) Moreover, together with Theorem 4.3, we obtain

$$\|v_0 - v^k\|_{L^\infty} \leq \|v_0 - v_{\mu^k}\|_{L^\infty} + \frac{\lambda}{\sqrt{\nu}}\|v_{\mu^k} - v^k\|_{\mu^k}$$

$$\leq c_0 \sqrt{\mu^k} + 2\vartheta \sqrt{\mu^k} \leq (c_0 + 2\vartheta)\sigma^{k/2} \sqrt{\mu^0},$$

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

6. Numerical tests. In §5, we have formulated our algorithm in a fairly abstract way. Now we reformulate it in terms of PDEs. First we consider the numerical approximation of our state equation

$$-\Delta y + y = u \qquad (6.1)$$
$$\partial_n p = 0. \qquad (6.2)$$

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 discretized by standart linear finite elements based on a regular triangulation of Ω.

Let $V_h \subset H^1(\Omega)$ be the finite dimensional space of ansatz functions and let K and M the associated stiffness- and mass matricies. The control u_h and the state are taken from the same space V_h:

$$u_h = \sum_{i=1}^n u_i \phi_i \text{ and } y_h = \sum_{i=1}^n y_i \phi_i$$

where ϕ is a basis of V_h with $\phi_i(x_j) = \delta_{ij}$. Then the functions u_h and y_h can be identified by theirs coeffcient vectors $\mathbf{u} = (\mathbf{u_1}, \mathbf{u_2}, ..., \mathbf{u_n})^{\mathbf{T}}$ and $\mathbf{y} = (\mathbf{y_1}, \mathbf{y_2}, ..., \mathbf{y_n})^{\mathbf{T}}$ respectively. The coincidence of the notation for these vectors with the original functions $u = u(x)$, $y = y(x)$ will not cause confusion. Then,

$$(K + M)\mathbf{y} = \mathbf{Mu} \qquad (6.3)$$

is a finite element approximation of (6.1)–(6.2).

To simplify the contruction of test examples we introduce a desired control u_d,(cf. (6.8) below) which does not change the validity of our theorems. Further, we drop the transformation $D^{-1} = S + \lambda I$, which we established to abbreviate the formulation of our problem. Written in terms of u, the necessary optimality condition is

$$S^* \left(Su - y_d - \frac{\mu}{Su + \lambda u - y_d} \right) + \nu(u - u_d) - \frac{\mu\lambda}{Su + \lambda u - y_d} = 0 \,.$$

Now we are able to formulate a discrete version of our solution operator S:

$$S_h := (K + M)^{-1}M. \qquad (6.4)$$

From §3 we know $S = S^*$ (and $D^* = D$). From that, we get $S_h^* = S_h$.

The discrete optimality condition is now

$$F_h(u) = S_h \left(S_h u - y_d - \frac{\mu}{(S_h + \lambda I)u - y_c} \right) + \nu(u - u_d) - \left(\frac{\mu\lambda I}{(S_h + \lambda I)u - y_c} \right) \quad (6.5)$$

The matrix S_h is not sparse and we would have to compute the inverse of $K + M$, so that this approach is not practicable. For practical use, we retransform the formulation (4.1) in the following way: We use $S = S^*$ and $D^* = D$. Setting $p = S^* \left(Su - y_d - \dfrac{\mu}{Su + \lambda u - y_d} \right)$, using $Su = y$ and the definition of S we get the system of necessary optimality conditions

$$
\begin{array}{rclccrcll}
-\Delta p + p & = & y - y_d - \dfrac{\mu}{y + \lambda u - y_d} & \text{in } \Omega & & -\Delta y + y & = & u & \text{in } \Omega \\[2mm]
\partial_n p & = & 0 & \text{on } \partial\Omega & & \partial_n y & = & 0 & \text{on } \partial\Omega
\end{array}
$$

$$
p + \nu(u - u_d) - \frac{\lambda\mu}{y + \lambda u - y_d} \quad \text{in } \Omega.
$$

Using the discretization (6.3), in view of (6.5) we arrive at

$$
\tilde{F}_h(\mathbf{y}, \mathbf{u}, \mathbf{p}; \mu) = \begin{pmatrix} K + M & -M & 0 \\ -M & 0 & K + M \\ 0 & \nu I & I \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{p} \end{pmatrix} \tag{6.6}
$$

$$
+ \begin{pmatrix} 0 \\ M\mathbf{y_d} \\ -\nu\mathbf{u_d} \end{pmatrix} + \begin{pmatrix} 0 \\ M\left(\dfrac{\mu}{\mathbf{y} + \lambda\mathbf{u} - \mathbf{y_c}} \right) \\ -\dfrac{\lambda\mu}{\mathbf{y} + \lambda\mathbf{u} - \mathbf{y_c}} \end{pmatrix}. \tag{6.7}
$$

This is the discrete optimality system we use for our computations.

We have tested our method by the following example

$$
\text{(PT)} \qquad \min J(y, u; \mu) := \frac{1}{2}\|y - y_d\|_{L^2(\Omega)}^2 + \frac{\nu}{2}\|u - u_d\|_{L^2(\Omega)}^2 \tag{6.8}
$$

subject to

$$
-\Delta y + y = u \qquad \text{in } \Omega \tag{6.9}
$$
$$
\partial_n y = 0 \qquad \text{on } \Gamma \tag{6.10}
$$

and to the pointwise mixed control-state constraints

$$
y + \lambda u \geq y_c \qquad \text{a.e. in } \Omega. \tag{6.11}
$$

with $\Omega = (0,1) \times (0,1)$.

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 belongs to $L^2(\Omega)$ even for $\lambda = 0$.

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 its PDE-toolbox for mesh generation, matrix-assembling etc. The stopping criterion for the outer iteration was $\mu \leq \epsilon = 10^{-12}$. We used a Friedrichs-Keller triangulation with fixed mesh size $h = 0.025$. In the following, the numerical solutions are denoted by $(\cdot)_\mu$, the exact optimal control, optimal state, and the optimal adjoint state are denoted by $\bar{u}$, $\bar{y}$ and $\bar{p}$. In some figures, these functions are labeled as u etc. We observe that for fixed mesh size the numerical solutions converge to the projection of the exact solution onto the finite element space. All computations were performed on a Pentium IV/2.8GHz machine with 1GB RAM running under Linux.

6.1. Example 1. This example is a slight update of Example 1 in [9]. We choose $\bar{u} = 2$, $\bar{p} = -2$ and $\bar{y} = 2$. The desired state is given by

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

y_c is given by

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

and the Lagrange multiplier is

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

Moreover, we have chosen $u_d = -\lambda\eta(x_1, x_2)$. In (6.8) we choose $\nu = 1$ and $\lambda = 10^{-3}$. In our computations, we choose in Algorithm 5.1 $\sigma = 0.9$ and $\mu^0 = 1$.
The following figures show the functions y_d, y_c and the Lagrange multiplier η.
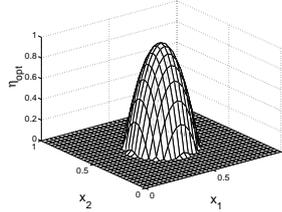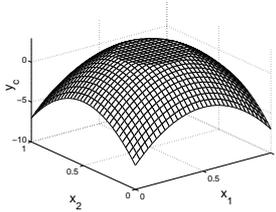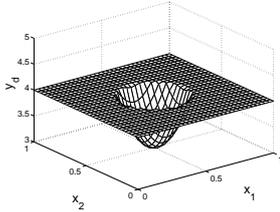
FIG. 6.1. *Desired state y_d*   FIG. 6.2. *State contraint y_c*   FIG. 6.3. *Multiplier η*

The next set of figures shows the numerical solutions y_h, u_h, p_h, and η_h of the problem regularized with $\lambda = 10^{-3}$.
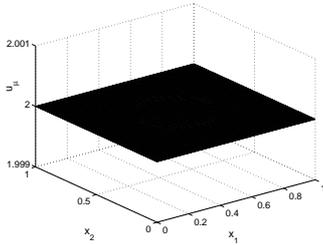
FIG. 6.4. *Control u_h*   FIG. 6.5. *State y_h*

18
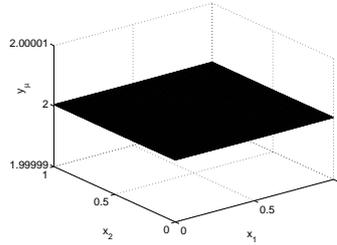
μ	err_u	err_y	err_p	err_η
1.0775e-02	8.9940e-03	8.9802e-03	8.9740e-03	5.9016e-01
1.0611e-03	1.1585e-03	1.1293e-03	1.1472e-03	4.9188e-01
1.0450e-04	3.0274e-04	1.2221e-04	2.6728e-04	4.6082e-01
1.0290e-05	1.7956e-04	1.7204e-05	1.5513e-04	3.0395e-01
1.0134e-06	5.2173e-05	3.0048e-06	4.2306e-05	1.4558e-01
1.1088e-07	1.4471e-05	5.6443e-07	9.8070e-06	6.4546e-02
1.0919e-08	3.5160e-06	9.4222e-08	1.8734e-06	2.2455e-02
1.0753e-09	7.3268e-07	1.4464e-08	3.1009e-07	1.0979e-02
1.0589e-10	1.2286e-07	1.9308e-09	4.2994e-08	1.0093e-02
1.0428e-11	1.5516e-08	2.1608e-10	4.8831e-09	1.0096e-02
1.0269e-12	1.6071e-09	2.1836e-11	4.9450e-10	1.0103e-02

TABLE 6.1
Relative errors for Example 1.
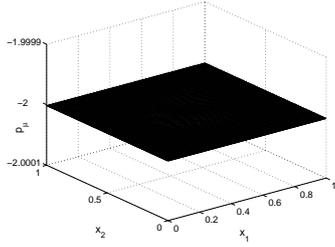
FIG. 6.6. *Adjoint state p_h*
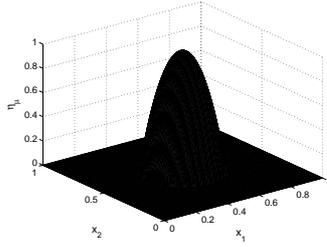
FIG. 6.7. *Lagrange multiplier η_h*

In the next tabular, we present the relative errors $err_u = \frac{\|u_\mu - \bar{u}\|_{l^2}}{\|\bar{u}\|_{l^2}}$, $err_y = \frac{\|y_\mu - \bar{y}\|_{l^2}}{\|\bar{y}\|_{l^2}}$, $err_p = \frac{\|p_\mu - \bar{p}\|_{l^2}}{\|\bar{p}\|_{l^2}}$, and $err_\eta = \frac{\|\eta_\mu - \bar{\eta}\|_{l^2}}{\|\bar{\eta}\|_{l^2}}$ for the regularized problem at $\lambda = 10^{-3}$ depending on μ. Table 6.1 shows the linear convergence in u, y, and p. This also reflected by the figures 6.8–6.11. For a comparison with results computed by a primal-dual active set strategy we refer to [9]. For $\mu < 10^{-10}$, the discretization error dominates the values of the error function $err(\eta)$.

The figures 6.8–6.11 show the differences between the numerical solutions u_μ, y_μ p_μ and η_μ and the exact solutions $\bar{u}$, $\bar{y}$, $\bar{p}$ and $\bar{\eta}$, measured in the L^2-norm at a regularization parameter $\lambda = 10^{-3}$ depending on the path-parameter μ. Both axes are scaled logarithmically. The behavior of the Lagrange multiplier for $\mu \to 0$ is remarkable, see also figures 6.12–6.15.
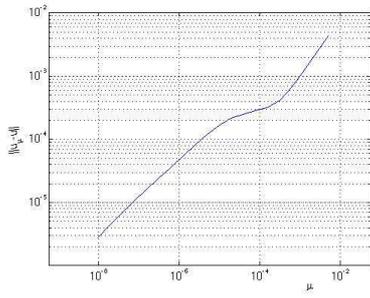
FIG. 6.8. *err(u)*
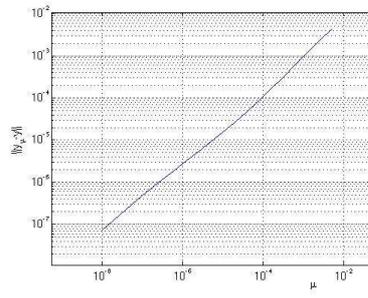
FIG. 6.9. *err(y)*
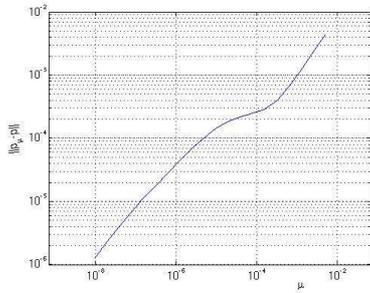
FIG. 6.10. *err(p)*
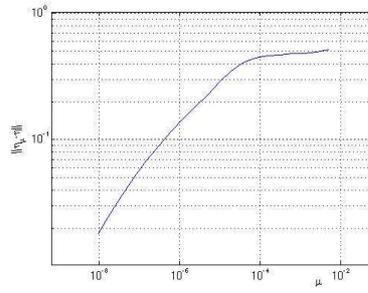
FIG. 6.11. *err(η)*

The next figures show the evolution of the Lagrange-multiplier η_μ along the central path.
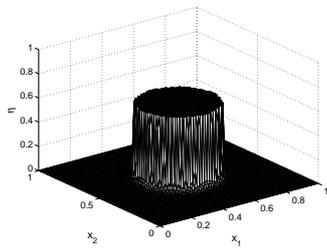
FIG. 6.12. *Multiplier η_h at $\mu = 10^{-2}$*
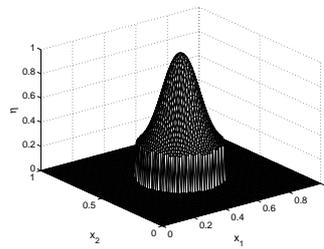
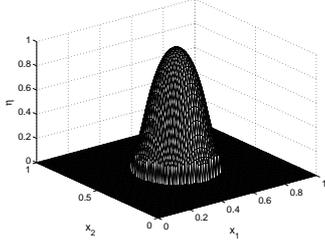FIG. 6.13. *Multiplier η_h at $\mu = 10^{-6}$*
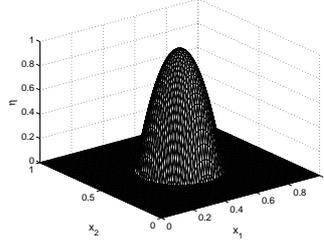
FIG. 6.14. *Multiplier η_h at $\mu = 10^{-7}$* FIG. 6.15. *Multiplier η_h at $\mu = 10^{-10}$*

6.2. Example 2. This example is constructed such that $\bar{y}$, $\bar{u}$ and $\bar{p}$ are trigono-metric functions of the form $\varphi(x_1, x_2) = c \cos{(\pi x_1)} \cos{(2\pi x_2)}$. 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 $\bar{u} = -\Delta\bar{y} + \bar{y} = (5\pi^2 + 1)\,\bar{y}$, and $u_d = \bar{u} + \frac{1}{\nu}\bar{p} - \frac{\lambda}{\nu}\eta = \bar{y} - \frac{\lambda}{\nu}\eta$, respectively. Choosing $\hat{y} = 2\sin{(2\pi x_1)} - 1.5$, $\bar{\eta} = \max{\{\hat{y} - \bar{y}, 0\}}$, and $y_c = \min{\{\hat{y}, \bar{y}\}} - \lambda\bar{u}$, 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 $y_d = \Delta\bar{p} - \bar{p} + \bar{y} - \bar{\eta} = ((5\nu\pi^2)(5\pi^2 + 1) + 1)\,\bar{y} - \bar{\eta}$. Figures 6.16–6.18 show the functions y_d y_c and η.
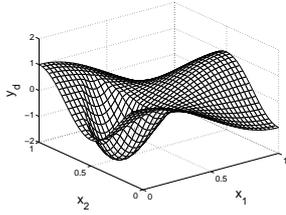
FIG. 6.16. *Desired state y_d* FIG. 6.17. *State constraints y_c* FIG. 6.18. *Multiplier $\bar{\eta}$*
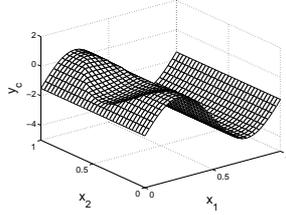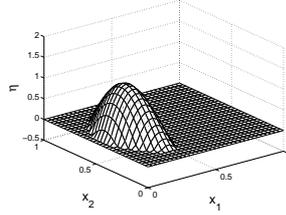
The following figures show the numerical solutions for $\nu = 10^{-3}$ and $\lambda = 10^{-3}$.

FIG. 6.19. *Control u_μ* FIG. 6.20. *State y_μ*

21
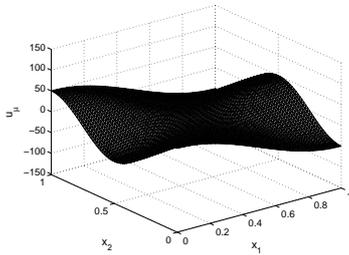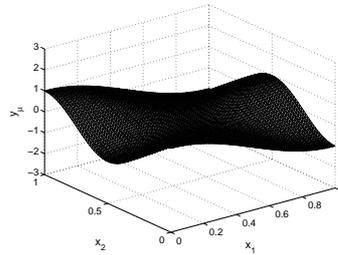
μ	err_u	err_y	err_p	err_η
1.0775e-02	1.7515e-02	4.2897e-02	1.4888e-02	3.2848e-01
1.0611e-03	5.0088e-03	6.7794e-03	3.8653e-03	1.4802e-01
1.0450e-04	1.3658e-03	1.1136e-03	9.0134e-04	5.9756e-02
1.0290e-05	4.4222e-04	3.9979e-04	3.0695e-04	2.1376e-02
1.0134e-06	2.8956e-04	3.7943e-04	2.4390e-04	1.1738e-02
1.1088e-07	2.7751e-04	3.8078e-04	2.3967e-04	1.0866e-02
1.0919e-08	2.7666e-04	3.8100e-04	2.3932e-04	1.0830e-02
1.0753e-09	2.7662e-04	3.8102e-04	2.3930e-04	1.0833e-02

TABLE 6.2

Relative errors for Example 2.
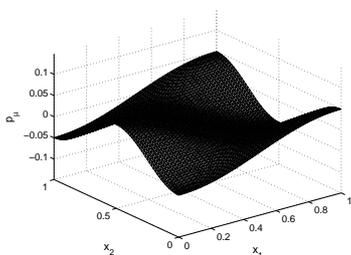
FIG. 6.21. *Adjoint state p_μ*
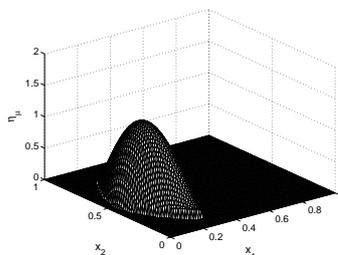
FIG. 6.22. *Multiplier η_μ*

Table 6.2 displays the relative error $err(\cdot)$ for u, y, p, and η. Here, the discretization error dominates $err(\cdot)$ for all values $\mu < 10^{-6}$.

Figures 6.23–6.26 present the differences between the numerical solutions and the optimal solutions at $\nu = 10^{-3}$ and $\lambda = 10^{-3}$.
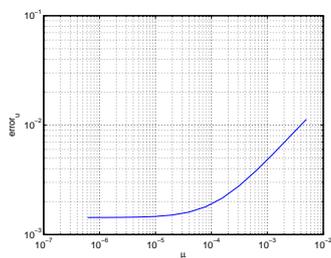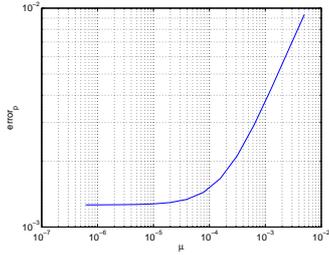

FIG. 6.23. *err_u*

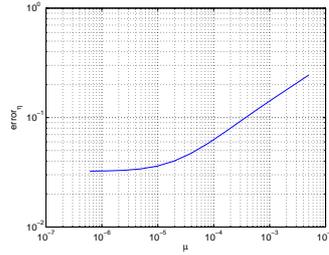

FIG. 6.24. *err_y*

22
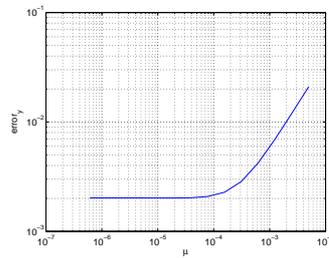
FIG. 6.25. err_p

FIG. 6.26. err_η

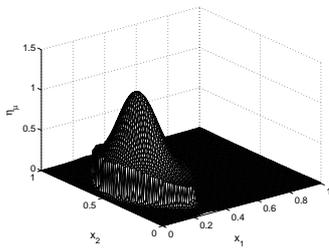The following set of figures shows the evolution of the Lagrange multiplier eta_μ along the central path.
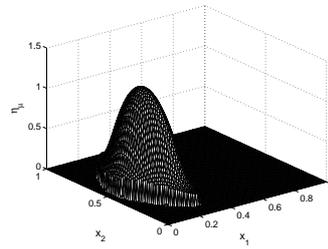

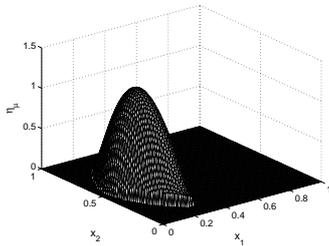FIG. 6.27. *Control* $\eta_m u$ *at* $\mu = 0.001$

FIG. 6.28. *Control* eta_μ *at* $\mu = 10^{-4}$

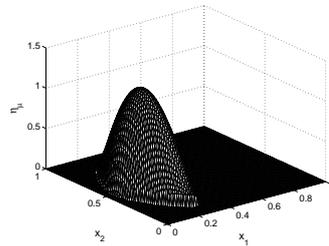
FIG. 6.29. *Control* η_μ *at* $\mu = 10^{-5}$

FIG. 6.30. *Control* η_μ *at* $\mu = 10^{-6}$

6.3. Example 3. In this example, we consider the problem (PT) in the following setting:

$$y_d = \cos(\pi x_1) \cos(2\pi x_2) \tag{6.12}$$

$$y_c = \min\{6 \sin(\pi x_1) \sin(\pi x_2) - 4, 1\} \tag{6.13}$$

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 6.31 and 6.32 we show the functions y_d and y_c.
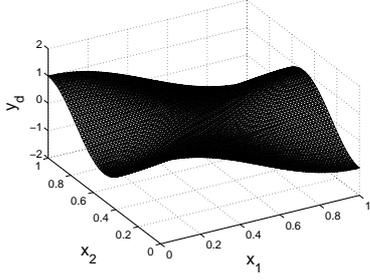
23

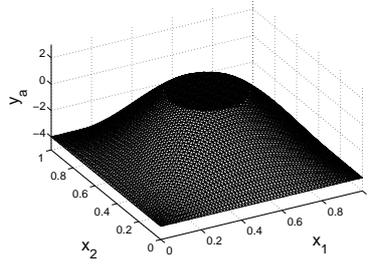FIG. 6.31. *Desired state* y_d

FIG. 6.32. *State contraints* y_c

For our computations we choose $\nu = 10^{-6}$ and $\lambda = 10^{-6}$. The following set of figures present the numerical solutions $u_h 1$, y_h, p_h and η_h.
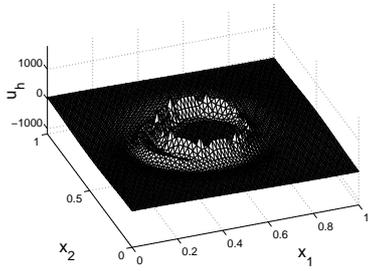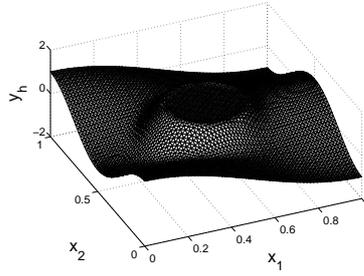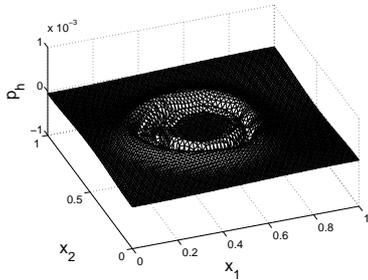

FIG. 6.33. *Control* u_h

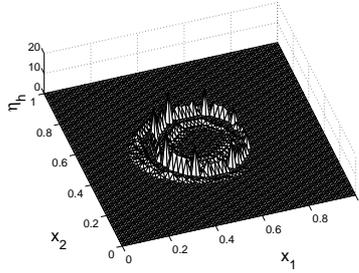FIG. 6.34. *State* y_h

FIG. 6.35. *Adjoint state* p_h

FIG. 6.36. *Lagrange multiplier* η_h *at* $\nu = 10^{-6}$ *and* $\lambda = 10^{-16}$

7. Pass to the limit $\lambda \downarrow 0$. We have considered our method for fixed $\lambda > 0$. Nevertheless, we briefly mention for convenience how close the solution $\bar{u}$ of the problem with pure state constraint $y \geq 0$ is approximated by $\bar{u}_\lambda$, the one associated to the Lavrentiev regularized constraint $\lambda u + y \geq 0$. It is known from [**?**] that $\|\bar{u} - \bar{u}_\lambda\| \to 0$ as $\lambda \downarrow 0$. To get a convergence rate, we require the following two assumptions:

Uniform boundedness: *There is* $M > 0$ *such that*

$$\|\bar{u}_\lambda\|_\infty \leq M \quad \forall\, \lambda > 0. \tag{7.1}$$

Notice that (7.1), together with $\bar{u}_\lambda \to \bar{u}$ in L^2, implies that $\|\bar{u}\|_\infty \leq M$.

24

Slater condition: *There is some $u_0 \in L^\infty$ and an $\varepsilon > 0$ such that*

$$y_0(x) \geq y_c(x) + \varepsilon \quad \forall x \in \bar{\Omega} \tag{7.2}$$

holds for the associated state $y_0 = S\,u_0$.

Under these assumptions, the first of them being quite strong, but often satesfied in concrete examples, the estimate

$$\|\bar{u} - \bar{u}_\lambda\| \leq C\sqrt{\lambda} \tag{7.3}$$

is obtained by a fairly standard technique, cf. for instance Alt [2]. We briefly sketch the main steps: In view of the assumptions (7.1) and (7.2), there exist positive constants c_0 and λ_0, and controls $\tilde{u}(\lambda)$, $\hat{u}(\lambda)$ with associated states $\tilde{y}(\lambda)$, $\hat{y}(\lambda)$ having the following properties: For all $\lambda \in (0, \lambda_0]$, it holds

$$\lambda\,\tilde{u}(\lambda) + \tilde{y}(\lambda) \geq y_c, \qquad \|\bar{u} - \tilde{u}(\lambda)\| \leq c_0\,\lambda \tag{7.4}$$

$$\hat{y}(\lambda) \geq y_c, \qquad \|\bar{u}_\lambda - \hat{u}(\lambda)\| \leq c_0\,\lambda. \tag{7.5}$$

The upper estimate is obtained as follows: We define $\tilde{u}(\lambda) = (1 - \rho)\bar{u} + \rho u_0$ with suitable $\rho > 0$. Then by $\bar{y} \geq y_c$ and (7.2)

$$
\begin{aligned}
\lambda\,\tilde{u}(\lambda) + \tilde{y}(\lambda) &= \lambda\,(1 - \rho)\bar{u} + \lambda\rho u_0 + (1 - \rho)\bar{y} + \rho y_0 \\
&\geq \lambda\,(1 - \rho)\bar{u} + \lambda\rho u_0 + (1 - \rho)y_c + \rho(y_c + \varepsilon) \\
&= \lambda\,(1 - \rho)\bar{u} + \lambda\rho u_0 + \rho\varepsilon + y_c.
\end{aligned}
$$

For $0 < \rho \leq 1$, we get $\|\lambda(1 - \rho)\bar{u} + \lambda\rho u_0\|_\infty \leq \lambda\,(M + \|u_0\|_\infty)$. Take

$$\rho = \frac{\lambda}{\varepsilon}(M + \|u_0\|_\infty) \tag{7.6}$$

and assume that λ is so small, say $\lambda \leq \lambda_0$, such that $\rho \leq 1$ holds. Then

$$\lambda(1 - \rho)\bar{u} + \lambda\rho u_0 + \rho\varepsilon + y_c \geq -\lambda\,(M + \|u_0\|_\infty) + \rho\varepsilon + y_c \geq y_c$$

so that $\lambda\,\tilde{u}(\lambda) + \tilde{y}(\lambda) \geq y_c$. Moreover,

$$\|\bar{u} - \tilde{u}(\lambda)\| = \|\bar{u} - (1 - \rho)\bar{u} - \rho u_0\| \leq \rho(M + \|u_0\|_\infty) \leq c_0\lambda$$

because of (7.6).

In the same way, the relations (7.5) are shown by the ansatz $\hat{u}(\lambda) = (1 - \rho)\bar{u}_\lambda + \rho u_0$ with certain $\rho \in (0, 1)$. We exploit $\lambda\bar{u}_\lambda + \bar{y}_\lambda \geq y_c$, hence $\bar{y}_\lambda \geq y_c - \lambda M$. The term $-\lambda M$ can be compensated by adding a small multiple of y_0.

Invoking (7.4), (7.5), the estimate (7.3) is now obtained immediately. The functional $f(u)$ is uniformly Lipschitz with constant L on the set of all u with $\|u\|_\infty \leq M$. We find by Taylor expansion

$$
\begin{aligned}
f(\bar{u}_\lambda) - f(\bar{u}) &\geq f'(\bar{u})(\bar{u}_\lambda - \bar{u}) + \frac{\kappa}{2}\|\bar{u}_\lambda - \bar{u}\|^2 \\
&= f'(\bar{u})(\hat{u}(\lambda) - \bar{u}) + \frac{\kappa}{2}\|\bar{u}_\lambda - \bar{u}\|^2 + f'(\bar{u})(\bar{u}_\lambda - \hat{u}(\lambda)) \\
&\geq \frac{\kappa}{2}\|\bar{u}_\lambda - \bar{u}\|^2 - c_1\,L\lambda
\end{aligned}
$$

since $\hat{u}(\lambda)$ satisfies the constraints of (P), and hence the variational inequality is fulfilled. Moreover, (7.5) was used. On the other hand,

$$f(\bar{u}_\lambda) - f(\bar{u}) = f(\bar{u}_\lambda) - f(\tilde{u}(\lambda)) + f(\tilde{u}(\lambda)) - f(\bar{u}) \leq 0 + c_2 L\lambda$$

is found. Altogether, $\kappa\|\bar{u}_\lambda - \bar{u}\|^2 \leq 2\,(c_1 + c_2)L\lambda$ follows from the last two inequations, implying the estimate (7.3).

REFERENCES

[1] J.-J. Alibert and J.-P. Raymond. Boundary control of semilinear elliptic equations with discontinuous leading coefficients and unbounded controls. *Numer. Funct. Anal. and Optimization*, 3 and 4:235–250, 1997.

[2] W. Alt. On the approximation of infinite optimization problems with an application to optimal control problems. *Appl. Math. Opt.*, 12:15–27, 1984.

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

[4] 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.

[5] E. Casas. Control of an elliptic problem with pointwise state constraints. *SIAM J. Control and Optimization*, 4:1309–1322, 1986.

[6] P. Deuflhard and F. Potra. Asymptotic mesh independence of Newton-Galerkin methods via a refined Mysovskii theorem. *SIAM J. Numer. Anal.*, 29:1395–1412, 1992.

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

[8] 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.

[9] 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.

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

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

[12] 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.

[13] F. Tröltzsch. *Optimale Steuerung partieller Differentialgleichungen. Theorie, Verfahren und Anwendungen.* Vieweg, 2005.

[14] 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.

[15] 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.

[16] M. Weiser. Interior point methods in function space. *SIAM J. Control Opt.*, to appear.

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

[18] M. Weiser and A. Schiela. Function space interior point methods for PDE constrained optimization. *PAMM*, 4:43–46, 2004.