Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

MARTIN WEISER

Interior Point Methods in Function Space

ZIB-Report 03-35 (November 2003)

Interior point methods in function space*

Martin Weiser

November 4, 2003

Abstract

A primal-dual interior point method for optimal control problems is considered. The algorithm is directly applied to the infinite dimensional problem. Existence and convergence of the central path are analyzed, and linear convergence of a short step pathfollowing method is established.

Keywords: interior point methods in function space, optimal control, complementarity functions

MSC (2000): 49M15, 90C48, 90C51

*Supported by the DFG Research Center "Mathematics for key technologies" (FZT 86) in Berlin.

1 Introduction

Numerical methods for solving optimal control problems governed by ODEs fall into two categories, the indirect methods [2, 3, 4, 6, 14, 15, 30] relying on Pontryagin's maximum principle, and the direct methods [7, 17, 20, 29, 36] based on the Karush-Kuhn-Tucker necessary conditions. Direct methods can be characterized by several features. Among them are

- position of discretization: Discretize-then-optimize approaches use an a priori parameterization of the control and possibly the state variables to reduce the optimal control problem to a finite dimensional nonlinear program. These large nonlinear programs can then be solved by standard NLP solvers. Adaptive mesh refinement can be performed after the finite dimensional optimum has been reached. On the other hand, optimize-then-discretize approaches formulate the optimization algorithms directly in the infinite dimensional function space, employing discretization only for solving linear operator equations. Adaptive mesh refinement is used to meet the accuracy requirements imposed on the solution of the linear equations by the optimization algorithm.

  Somewhere in between are function space SQP methods where linear-quadratic programs are discretized.

- type of optimization algorithm: Among the most popular algorithms employed for solving the optimization problems arising in optimal control are sequential quadratic programming (SQP) and interior point methods (IPM). A recent alternative are semismooth Newton methods [5, 33].

Discretize-then-optimize methods are covered by a vast amount of published literature using almost any available algorithm for solving the finite dimensional NLPs. Solutions on consecutive mesh refinement levels or in consecutive SQP steps often exhibit pronounced similarities. These redundancy can be directly exploited by active set type methods. In contrast, IPMs are considered to benefit less from this redundancy [19, 39]. Nevertheless, IPMs are reported to be very efficient for solving optimal control problems — a fact which is not well explained by straightforward application of finite dimensional IP convergence theory to the discretized problems. The best currently known convergence rates of $1 - \text{const}/\sqrt{n}$ would instead predict a pronounced mesh dependence of the convergence.

Among the optimize-then-discretize approaches, the SQP methods dominate the published material [1, 17, 21, 22, 26, 31, 32]. Here, Robinson's theory of generalized equations [28] can be used to analyze the function space methods, which leaves, however, the question of how to solve the infinite dimensional linear-quadratic programs. This is implicitly addressed by infinite dimensional interior point methods, which have nevertheless attracted less attention [34, 35, 23].

The present paper presents an infinite dimensional interior point method directly applied to optimal control problems in function space in Section 2. Existence and convergence of the central path are analyzed in Section 3. Finally, linear convergence of a theoretical short-step pathfollowing algorithm with classical predictor is shown in Section 4. In particular, the rate of convergence does not depend on the size of any discretization.

Notation. The Lebesgue spaces and Sobolev spaces of functions with values in $\mathbb{R}^n$ are denoted by L_p^n and $(W_p^m)^n$, respectively. $S(x,\rho)$ is the open ball around x with radius ρ.

Some variables and operators are constructed such that they have a natural block partitioning corresponding to the components u and y of x. The individual blocks are denoted by the corresponding component as a superscript, e.g.

$$g(x) = \begin{bmatrix} g^u(u) \\ g^y(y) \end{bmatrix} \quad \text{and} \quad \Psi(g(x),\eta) = \begin{bmatrix} \Psi^u(g^u(u),\eta^u) \\ \Psi^y(g^y(y),\eta^y) \end{bmatrix}.$$

2 Problem Setting

On the time interval $\Omega = [0,1]$ we consider the optimal control problem

$$\begin{aligned} \min J(x) \quad \text{subject to} \quad & c(x) = 0 \quad \text{a.e.} \\ & r(x) = 0 \\ & g(x) \geq 0 \quad \text{a.e.} \end{aligned} \tag{1}$$

with a partitioning of the variable $x = (u,y) \in X = L_\infty^{n_u}(\Omega) \times (W_\infty^1)^{n_y}(\Omega)$ into controls and states, a Lagrange type cost functional

$$J(x) = \int_0^1 \tilde{f}(u(t),y(t))\,dt\,,$$

ordinary differential equations with boundary conditions

$$c(x) = \begin{bmatrix} \bar{c}(x) \\ y(0) - y_0 \end{bmatrix}, \quad \bar{c}(x)(t) = \tilde{c}(x(t)) - \dot{y}(t)\,, \tag{2}$$

$$r(x) = \tilde{r}(y(1)) \tag{3}$$

as equality constraints, and pointwise state and control constraints

$$g(x)(t) = \begin{bmatrix} \tilde{g}^u(u(t)) \\ \tilde{g}^y(y(t)) \end{bmatrix}.$$

For the whole paper we will restrict the discussion to the fixed time interval Ω and hence simplify the notation by omitting it from the function spaces. We assume all the functions $\tilde{f} : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \to \mathbb{R}$, $\tilde{c} : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \to \mathbb{R}^{n_y}$, $\tilde{r} : \mathbb{R}^{n_y} \to \mathbb{R}^{n_r}$, $\tilde{g}^u : \mathbb{R}^{n_u} \to \mathbb{R}^{n_{\eta_u}}$, and $\tilde{g}^y : \mathbb{R}^{n_y} \to \mathbb{R}^{n_{\eta_y}}$ to be twice Lipschitz continuously differentiable.

For convenience, we give here a theorem on Nemyckii operators in L_∞, the straightforward proof of which can be found in [37].

Theorem 2.1. *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is k times differentiable and its k-th derivative satisfies the Lipschitz condition*

$$|f^{(k)}(x) - f^{(k)}(y)| \leq \kappa |x-y|\,, \tag{4}$$

the corresponding Nemyckii operator $\mathbf{f}$ defined by $\mathbf{f}(u)(t) = f(u(t))$ maps L_∞^n into L_∞^m and is k times Fréchet differentiable. For $1 \leq p \leq \infty$ its k-th derivative

can be continuously extended to an operator $\mathbf{f}^{(k)}(u) : (\prod_{j=1}^{k} L^n_{pk}) \to L^m_p$ *that inherits boundedness and Lipschitz continuity from* $f^{(k)}$*:*

$$\left\|\mathbf{f}^{(k)}(u)\right\|_{(\prod_{j=1}^{k} L^n_{pk})\to L^m_p} \leq \sup_{|x|\leq\|u\|_{L^n_\infty}} |f^{(k)}(x)| \tag{5}$$

$$\left\|\mathbf{f}^{(k)}(u+\delta u) - \mathbf{f}^{(k)}(u)\right\|_{(\prod_{j=1}^{k} L^n_{pk})\to L^m_p} \leq \kappa \left\|\delta u\right\|_{L^n_\infty} \tag{6}$$

If in addition f *is* $k+1$ *times differentiable and its* $k+1$*-st derivative satisfies the Lipschitz condition*

$$|f^{(k+1)}(x) - f^{(k+1)}(y)| \leq \kappa|x-y|\,,$$

then $\mathbf{f}$ *maps* $(W^1_\infty)^n$ *into* $(W^1_\infty)^m$ *and is* k *times differentiable. For* $p \geq 1$ *its* k*-th derivative can be continuously extended to an operator* $\mathbf{f}^{(k)}(u) : (\prod_{j=1}^{k}(W^1_{pk})^n) \to (W^1_p)^m$ *that inherits boundedness and Lipschitz continuity from* $f^{(k)}$ *and* $f^{(k+1)}$*:*

$$\left\|\mathbf{f}^{(k)}(u)\right\|_{(\prod_{j=1}^{k}(W^1_{pk})^n)\to(W^1_p)^m} \leq \sup_{|x|\leq\|u\|_{L^n_\infty}} (k+1)|f^{(k)}(x)| + |f^{(k+1)}(x)| \tag{7}$$

$$\left\|\mathbf{f}^{(k)}(u+\delta u) - \mathbf{f}^{(k)}(u)\right\|_{(\prod_{j=1}^{k}(W^1_{pk})^n)\to(W^1_p)^m} \leq (k+2)\kappa \left\|\delta u\right\|_{(W^1_\infty)^n}$$

If the derivatives of $f : \mathbb{R}^n \to \mathbb{R}^n$ *and* $g : \mathbb{R}^n \to \mathbb{R}^n$ *commute, then so do the derivatives of the corresponding Nemyckii operators* $\mathbf{f}'$ *and* $\mathbf{g}'$*.*

With Theorem 2.1 above we conclude that

$$\begin{aligned} J :& L^{n_u}_\infty \times (W^1_\infty)^{n_y} \to \mathbb{R}\,, \\ c :& L^{n_u}_\infty \times (W^1_\infty)^{n_y} \to L^{n_y}_\infty\,, \text{ and} \\ g :& L^{n_u}_\infty \times (W^1_\infty)^{n_y} \to L^{n_{\eta u}}_\infty \times L^{n_{\eta y}}_\infty \end{aligned}$$

are twice Lipschitz-continuously differentiable operators.

The aim of the interior point method discussed here is to approximate Kuhn-Tucker points x_*. These are feasible points characterized by the existence of Lagrange multipliers $\lambda_c \in \mathbb{R}^{n_y} \times (L^{n_y}_\infty)^*$, $\lambda_r \in \mathbb{R}^{n_r}$, and $\eta \in (L^{n_{\eta u}}_\infty)^* \times ((W^1_\infty)^{n_{\eta y}})^*$, such that the following conditions are satisfied:

$$\begin{gathered} J'(x_*) - c'(x_*)^*\lambda_c - r'(x_*)^*\lambda_r - g'(x_*)^*\eta = 0 \\ c(x_*) = 0\,, \quad r(x_*) = 0 \\ g(x_*) \geq 0\,, \quad \eta \geq 0\,, \quad \langle \eta, g(x_*)\rangle = 0 \end{gathered} \tag{8}$$

Under certain assumptions (see e.g. [25, 27]) these conditions are necessary for x_* to be a local solution of (1). Thus, Kuhn-Tucker points are promising candidates for solutions.

Unfortunately, the unwieldy complementarity condition (8) is difficult to handle numerically. The idea of primal-dual interior point methods is to relax the complementarity condition by

$$\eta \cdot g(x) = \mu\,, \quad \eta \geq 0\,, \quad g(x) \geq 0 \tag{9}$$

and to consider the homotopy $\mu \to 0$. Alternatively, complementarity functions $\psi(a, b; \mu) : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$ can be used to construct Nemyckii operators Ψ, such that

$$\Psi(g(x), \eta; \mu) = 0$$

is more or less equivalent to the classical interior point relaxation (9).

These relaxations, however, are only well defined if $\eta \in L_1$, and continuously differentiable only in case $\eta \in L_\infty$. Note that this is only required to hold during the homotopy for $\mu > 0$, not at the Kuhn-Tucker point itself. We will prove in Theorem 3.4 that the homotopy can indeed be performed in the more regular setting of $\eta \in L_\infty^{n_{\eta u}} \times L_\infty^{n_{\eta y}} \subset (L_\infty^{n_{\eta u}})^* \times ((W_\infty^1)^{n_{\eta y}})^*$ for $\mu > 0$.

Define the Lagrangian

$$L(x, \lambda_c, \lambda_r, \eta) = J(x) - \langle \lambda_c, c(x) \rangle - \langle \lambda_r, r(x) \rangle - \langle \eta, g(x) \rangle.$$

Let

$$F(x, \lambda_c, \lambda_r, \eta; \mu) = \begin{bmatrix} \partial_x L(x, \lambda_c, \lambda_r, \eta) \\ -c(x) \\ -r(x) \\ \Psi(\eta, g(x); \mu) \end{bmatrix}. \tag{10}$$

As will be shown in Theorem 3.2 below, F maps

$$V \times \mathbb{R}_+ = (L_\infty^{n_u} \times (W_\infty^1)^{n_y}) \times (\mathbb{R}^{n_y} \times L_\infty^{n_y}) \times \mathbb{R}^{n_r} \times (L_\infty^{n_{\eta u}} \times L_\infty^{n_{\eta y}}) \times \mathbb{R}_+ \tag{11}$$

into

$$Z = (L_\infty^{n_u} \times (W_1^1)^{n_y *}) \times (\mathbb{R}^{n_y} \times L_\infty^{n_y}) \times \mathbb{R}^{n_r} \times (L_\infty^{n_{\eta u}} \times L_\infty^{n_{\eta y}}).$$

3 The central path

The main object of analytical interest is the *central path* defined by the homotopy (9) in μ. First we consider its actual existence in the regular setting given by (11) before discussing convergence.

Throughout the paper, we will use the Fischer-Burmeister function [18]

$$\psi(a, b; \mu) = a + b - \sqrt{a^2 + b^2 + 2\mu} \tag{12}$$

as an example from a large class of different complementarity functions (see [11, 12, 13, 24]).

3.1 Existence

We begin with establishing some bounds on derivatives of the complementarity function and their inverses.

Lemma 3.1. *The complementarity function Ψ defined via (12) maps $L_\infty^n \times L_\infty^n \times \mathbb{R}$ continuously into L_∞^n. Its derivative $\partial_g \Psi(g, \eta; \mu)$ is symmetric positive semidefinite, bounded by*

$$\|\partial_g \Psi\|_{L_\infty \to L_\infty} \leq 2, \tag{13}$$

$$\left\|(\partial_g \Psi)^{-1}\right\|_{L_\infty \to L_\infty} \leq \max\left(3, \frac{2}{\mu} \|g\|_{L_\infty}^2\right), \tag{14}$$

and Lipschitz continuous with a Lipschitz constant of $\mu^{-1/2}$. The corresponding holds for $\partial_\eta \Psi(g, \eta; \mu)$. Furthermore, the derivatives commute.

Proof. The claimed properties of the Nemyckii opterator Ψ are directly inherited from ψ due to Theorem 2.1. From $(1+\phi)^{-1/2} \leq \max(1-\phi/4, 2/3)$ for $\phi > 0$ we infer

$$\begin{aligned}
\min\left(\frac{\mu}{2a^2}, \frac{1}{3}\right) &= 1 - \max\left(1 - \frac{\mu}{2a^2}, \frac{2}{3}\right) \leq 1 - \frac{1}{\sqrt{1+\frac{2\mu}{a^2}}} \\
&\leq 1 - \frac{1}{\sqrt{1+\frac{b^2}{a^2}+\frac{2\mu}{a^2}}} = 1 - \frac{|a|}{\sqrt{a^2+b^2+2\mu}} \\
&\leq \partial_a \psi(a, b; \mu) \\
&\leq 1 + \frac{|a|}{\sqrt{a^2+b^2+2\mu}} \leq 2\,.
\end{aligned} \tag{15}$$

Thus, $\partial_a \psi$ is uniformly positive definite. Due to Theorem 2.1, the derivative $\partial_g \Psi(g, \eta; \mu)$ of the Nemyckii operator Ψ is bounded by (13) and has an inverse which is bounded by (14).

As for the Lipschitz continuity, we estimate

$$|\partial_a^2 \psi| = \left| \frac{\sqrt{a^2+b^2+2\mu} - \frac{a^2}{\sqrt{a^2+b^2+2\mu}}}{a^2+b^2+2\mu} \right| \leq \frac{1 - \frac{a^2}{a^2+b^2+2\mu}}{\sqrt{a^2+b^2+2\mu}} \leq \frac{1}{\sqrt{2\mu}}$$

and

$$|\partial_{ab} \psi| = \left| \frac{ab}{(a^2+b^2+2\mu)^{3/2}} \right| \leq \frac{|ab|}{(2|ab|+2\mu)^{3/2}} \leq \frac{2}{3\sqrt{6\mu}}\,,$$

such that $\|\psi''\| \leq \mu^{-1/2}$. This Lipschitz constant for $\partial_a \psi$ is inherited by $\partial_g \Psi$. Because of symmetry, the same holds for $\partial_\eta \Psi$, which commutes with $\partial_g \Psi$. □

Theorem 3.2. *The complementarity formulation* (10) *is a continuously differentiable mapping from $V \times \mathbb{R}_+$ to Z. Moreover, for any bounded set $D \subset V$ there is a constant $c(D)$ such that the derivative $\partial_v F$ satisfies the Lipschitz condition*

$$\|\partial_v F(v + \delta v; \mu) - \partial_v F(v; \mu)\|_{V \to Z} \leq c(1 + \mu^{-1/2}) \, \|\delta v\|_V \tag{16}$$

on D.

Proof. The image spaces and differentiability of the second to fourth component of F have already been established in Section 2 and Lemma 3.1. Only the adjoint expression

$$J'(x) - c'(x)^* \lambda_c - r'(x)^* \lambda_r - g'(x)^* \eta$$

remains to be discussed. We consider the terms separately.

First we write $J(x) = \langle \mathbb{1}, \tilde{\mathbf{f}}(x) \rangle$ with $\tilde{\mathbf{f}}'(x) \in \mathcal{L}(L_1^{n_u} \times (W_1^1)^{n_y}, L_1)$ due to Theorem 2.1 and thus obtain

$$J'(x) = \tilde{\mathbf{f}}'(x)^* \mathbb{1} \in \left(L_1^{n_u} \times (W_1^1)^{n_y}\right)^*. \tag{17}$$

With δ_0 denoting the point evaluation of the y component at $t = 0$ we have

$$c'(x) = \begin{bmatrix} \bar{\mathbf{c}}'(x) - \partial_t \\ \delta_0 \end{bmatrix} \in \mathcal{L}\big(L_1^{n_u} \times (W_1^1)^{n_y} \to L_1^{n_y} \times \mathbb{R}^{n_r}\big)$$

again by Theorem 2.1, such that

$$c'(x)^*\lambda_c \in \left(L_1^{n_u} \times (W_1^1)^{n_y}\right)^*. \tag{18}$$

Similarly we obtain

$$r'(x)^*\lambda_r \in \left(L_1^{n_u} \times (W_1^1)^{n_y}\right)^* \quad \text{and} \quad g'(x)^*\eta \in \left(L_1^{n_u} \times (W_1^1)^{n_y}\right)^*. \tag{19}$$

Collecting (17)–(19), $F(v;\mu) \in Z$ is verified. Continuous differentiability is inherited from J, c, g, and ψ.

As for the Lipschitz continuity of the derivative, we have to estimate differences of

$$\partial_v F(v;\mu) = \begin{bmatrix} \partial_x^2 L(v) & -c'(x)^* & -r'(x)^* & -g'(x)^* \\ -c'(x) & & & \\ -r'(x) & & & \\ \partial_g \Psi(g(x),\eta;\mu)g'(x) & & & \partial_\eta \Psi(g(x),\eta;\mu) \end{bmatrix}$$

for arguments v_1 and v_2. We cover the the blocks separately. First we see that

$$c'(x_1) - c'(x_2) = \tilde{\mathbf{c}}'(x_1) - \tilde{\mathbf{c}}'(x_2).$$

Since x_1 and x_2 are bounded in terms of D, the derivative of the Nemyckii operator $\tilde{\mathbf{c}}$ inherits the Lipschitz constant $\kappa_c(D)$ of $\tilde{c}'$ due to (6) of Theorem 2.1 with $p = \infty$. Thus we conclude

$$\|c'(x_1) - c'(x_2)\|_{X \to L_\infty^{n_y} \times \mathbb{R}^{n_y}} \le \kappa_c(D)\|x_1 - x_2\|_X.$$

Analogously, we obtain

$$\|g'(x_1) - g'(x_2)\|_{X \to L_\infty^{n_\eta}} \le \kappa_g(D)\|x_1 - x_2\|_X.$$

Concerning the dual operators $c'(x)^*$ and $g'(x)^*$, we apply Theorem 2.1 with $p = 1$ in (6) and obtain

$$\|c'(x_1)^* - c'(x_2)^*\|_{L_\infty^{n_y} \times \mathbb{R}^{n_y} \to L_\infty^{n_u} \times ((W_1^1)^{n_y})^*} \le \kappa_c(D)\|x_1 - x_2\|_X$$

and

$$\|g'(x_1)^* - g'(x_2)^*\|_{L_\infty^{n_\eta} \to L_\infty^{n_u} \times ((W_1^1)^{n_y})^*} \le \kappa_g(D)\|x_1 - x_2\|_X.$$

Similar estimates for $r'(x)$ and $r'(x)^*$ are straightforward. As for $\partial_x^2 L(v)$, we estimate

$$\begin{aligned} \|J''(x_1) - J''(x_2)\|_{X \to L_\infty^{n_u} \times ((W_1^1)^{n_y})^*} &\le \kappa_f(D)\|x_1 - x_2\|_X \\ \|c''(x_1)^* - c''(x_2)^*\|_{X \times L_\infty^{n_y} \times \mathbb{R}^{n_y} \to L_\infty^{n_u} \times ((W_1^1)^{n_y})^*} &\le \kappa_c(D)\|x_1 - x_2\|_X \\ \|g''(x_1)^* - g''(x_2)^*\|_{X \times L_\infty^{n_\eta} \to L_\infty^{n_u} \times ((W_1^1)^{n_y})^*} &\le \kappa_g(D)\|x_1 - x_2\|_X \end{aligned}$$

as before. In view of

$$c''(x_1)^*\lambda_{c1} - c''(x_2)^*\lambda_{c2} = c''(x_1)^*(\lambda_{c1} - \lambda_{c2}) + (c''(x_1)^* - c''(x_2)^*)\lambda_{c2}$$

and the boundedness of $c''(x_1)^*$ due to (5) of Theorem 2.1, we derive a constant $\kappa(D)$ for

$$\|c''(x_1)^*\lambda_{c1} - c''(x_2)^*\lambda_{c2}\|_{X \to L_\infty^{n_u} \times ((W_1^1)^{n_y})^*} \le \bar{\kappa}_c(D)\|v_1 - v_2\|_X.$$

Treating $r''(x)^*\lambda_r$ and $g''(x)^*\eta$ similarly, we obtain the desired estimate

$$\|\partial_x^2 L(v_1) - \partial_x^2 L(v_2)\|_{X \to L_\infty^{n_u} \times ((W_1^1)^{n_y})^*} \leq \kappa_L(D)\|v_1 - v_2\|_X.$$

Up to now, the Lipschitz constants have been completely independent of μ. For the blocks $\partial_g \Psi(v)g'(x)$ and $\partial_\eta \Psi(v)$ we obtain a Lipschitz constant of $\kappa_\Psi \leq \text{const}\,(1+\mu^{-1/2})$. Combining the Lipschitz constants of the individual blocks finally verifies (16). □

In order to prove the existence of the central path via an implicit function theorem, we first have to establish bounds on the inverse of $\partial_v F$.

Theorem 3.3. *Suppose there are an open bounded set $D \subset V$ and constants $\beta > 0$ and $\alpha > 0$, such that the following conditions hold uniformly for all $v \in D$ and $\mu > 0$.*

1. *The state equation satisfies the following inf-sup-condition:*

$$\inf_{\xi \in \mathbb{R}^{n_r}} \sup_{\delta u \in L_2^{n_u}} \frac{\xi^T \partial_y r(x) \partial_y c(x)^{-1} \partial_u c(x) \delta u}{|\xi| \, \|\delta u\|_{L_2^{n_u}}} \geq \beta$$

   (The linearized state equation is controllable.)

2. *A strengthened Legendre-Clebsch type condition holds:*

$$\xi^T M_u(t) \xi \geq \alpha |\xi|^2$$

   for all $\xi \in \mathbb{R}^{n_u}$ and almost all $t \in \Omega$. Here,

$$\begin{aligned} M_u(t) &:= \partial_u^2 \tilde{f}(x(t)) - \partial_u^2 \tilde{c}(x(t))^T \lambda_c(t) - (\tilde{g}^u)''(u(t))^T \eta^u(t) \\ &+ (\tilde{g}^u)'(u(t))^T \partial_\eta \psi(\tilde{g}^u(u(t)), \eta^u(t); \mu)^{-1} \partial_g \psi(\tilde{g}^u(u(t)), \eta^u(t); \mu)(\tilde{g}^u)'(u(t)) \,. \end{aligned}$$

3. *The augmented second derivative of the Lagrangian is positive definite on the nullspace of the state equation:*

$$\langle \xi, (\partial_x^2 L(v) + g'(x)^* \partial_\eta \Psi(g(x), \eta)^{-1} \partial_g \Psi(g(x), \eta) g'(x)) \xi \rangle \geq \alpha \|\xi\|^2_{L_2^{n_u} \times (W_2^1)^{n_y}}$$

   for all $\xi \in \ker c'(x)$.

Then $\partial_v F(v; \mu)$ has an inverse which is bounded by

$$\left\| \partial_v F(v; \mu)^{-1} \right\|_{Z \to V} \leq \text{const}\,(1 + \mu^{-3}) \tag{20}$$

uniformly for $v \in D$.

Proof. We show that there is a unique solution of $\partial_v F(v; \mu)\Delta v = z$ with $\|\Delta v\|_V \leq \text{const}\,(1+\mu^{-3})\,\|z\|_Z$.

In order to simplify the notation, let $C = -c'(x)$, $C_u = -\partial_u c(x)$, $C_y = -\partial_y c(x)$, and analogously G, G_u, G_y, R, and R_y. Define $\Psi_\eta = \partial_\eta \Psi(g(x), \eta)$, $\Psi_g = \partial_g \Psi(g(x), \eta)$, and $\Psi_\eta^u = \partial_{\eta^u} \Psi^u(g^u(u), \eta^u)$, $\Psi_g^u = \partial_{g^u} \Psi^u(g^u(u), \eta^u)$, and analogously Ψ_η^y and Ψ_g^y. Moreover, let $M_u = \partial_u^2 L(v) + G_u^*(\Psi_\eta^u)^{-1}\Psi_g^u G_u$, and analogously M_y. Finally, let $M_{uy} = \partial_{uy} L(v)$ and $M_{yu} = \partial_{yu} L(v)$.

The state derivative C_y represents the linearization of the initial value problem (2) and has a bounded solution for any right hand side. Thus, C_y has a bounded inverse. More precisely, for any $p \geq 1$,

$$C_y^{-1} : L_p^{n_y} \to (W_p^1)^{n_y} \text{ is bounded uniformly for } v \in D. \tag{21}$$

Therefore we can define the solution operator $S = C_y^{-1} C_u$.

In the following, we will refrain from writing the number of components of the function spaces, which should be clear from context.

In a first step, we reduce the system

$$\partial_v F(v;\mu)(\Delta x, \Delta\lambda_c, \Delta\lambda_r, \Delta\eta)^T = [z_a, z_c, z_r, z_p]^T ,$$

to a simple saddle point problem. Elimination of the inequality constraints' multipliers $\Delta\eta = \Psi_\eta^{-1}(z_p - \Psi_g G \Delta x)$ by Lemma 3.1 yields the equivalent system

$$\begin{bmatrix} M_u & M_{uy} & C_u^* & \\ M_{yu} & M_y & C_y^* & R_y^* \\ C_u & C_y & & \\ & R_y & & \end{bmatrix} \begin{bmatrix} \Delta u \\ \Delta y \\ \Delta\lambda_c \\ \Delta\lambda_r \end{bmatrix} = \begin{bmatrix} \bar{z}_a^u \\ \bar{z}_a^y \\ z_c \\ z_r \end{bmatrix} ,$$

where $(\bar{z}_a^u, \bar{z}_a^y)^T = \bar{z}_a = z_a - G^* \Psi_\eta^{-1} z_p$. Then, $\Delta y = C_y^{-1} z_c - S\Delta u$ and $\Delta\lambda_c = C_y^{-*}(\bar{z}_a^y - M_y C_y^{-1} z_c - (M_{yu} - M_y S)\Delta u - R_y^* \Delta\lambda_r)$ can be eliminated, which yields

$$\begin{bmatrix} M_u + S^* M_y S - (M_{uy} S + S^* M_{yu}) & -S^* R_y^* \\ -R_y S & \end{bmatrix} \begin{bmatrix} \Delta u \\ \Delta\lambda_r \end{bmatrix} = \begin{bmatrix} \hat{z}_a^u \\ \hat{z}_r \end{bmatrix} . \tag{22}$$

Here we set $\hat{z}_a^u = \bar{z}_a^u - M_{uy} C_y^{-1} z_c - S^*(\bar{z}_a^y - M_y C_y^{-1} z_c)$ and $\hat{z}_r = z_r - R_y C_y^{-1} z_c$.

In the second step, we establish the existence of a bounded solution of (22), first in $L_2^{n_u} \times \mathbb{R}^{n_r}$ and then in $L_\infty^{n_u} \times \mathbb{R}^{n_r}$. Due to Theorem 2.1 and the observation (21), M_u, $S^* M_y S$, $M_{uy} S$, and $S^* M_{yu}$ can all be continuously extended to L_2. Then, $M_u + S^* M_y S - (M_{uy} S + S^* M_{yu}) : L_2^{n_u} \to L_2^{n_u}$ is positive definite due to assumption 3. Moreover, $R_y S$ satisfies the inf-sup-condition of assumption 1. Therefore, Brezzi's splitting theorem [10, 8] guarantees the existence of a solution $(\Delta u, \Delta\lambda_r) \in L_2^{n_u} \times \mathbb{R}^{n_r}$ of (22) with

$$\begin{aligned} \|\Delta u\|_{L_2} &\leq \text{const} \left(\|\hat{z}_a^u\|_{L_2} + \kappa |\hat{z}_r| \right) \quad \text{and} \\ |\Delta\lambda_r| &\leq \text{const} \left(\kappa \|\hat{z}_a^u\|_{L_2} + \kappa^2 |\hat{z}_r| \right) , \end{aligned} \tag{23}$$

where

$$\kappa = 1 + \|M_u + S^* M_y S - (M_{uy} S + S^* M_{yu})\|_{L_2 \to L_2} ,$$

and the constants depend on α and β. Using Lemma 3.1 and, again, the extension of Nemyckii operators to L_2 provided by Theorem 2.1, we obtain the following dependencies on μ:

$$\begin{aligned} \|M_u\|_{L_2 \to L_2} &= \|\partial_u^2 L\|_{L_2 \to L_2} + \|G_u^* (\Psi_\eta^u)^{-1} \Psi_g^u G_u\|_{L_2 \to L_2} \\ &\leq \text{const} + \|G_u^*\|_{L_2 \to L_2} \|(\Psi_\eta^u)^{-1} \Psi_g^u\|_{L_2 \to L_2} \|G_u\|_{L_2 \to L_2} \\ &\leq \text{const} \, (1 + \|((\Psi_\eta^u)^{-1} \Psi_g^u\|_{L_2 \to L_2}) \\ &\leq \text{const} \, (1 + \mu^{-1}) \end{aligned}$$

$$\begin{aligned}\|M_y\|_{W_2^1\to(W_2^1)^*} &= \|\partial_y^2 L\|_{W_2^1\to(W_2^1)^*} + \|G_y^*(\Psi_\eta^y)^{-1}\Psi_g^y G_y\|_{W_2^1\to(W_2^1)^*}\\ &\le \text{const} + \|G_y^*\|_{L_2\to(W_2^1)^*}\|\Psi(_\eta^y)^{-1}\Psi_g^y\|_{L_2\to L_2}\|G_y\|_{W_2^1\to L_2}\\ &\le \text{const}\,(1+\|(\Psi_\eta^y)^{-1}\Psi_g^y\|_{L_2\to L_2})\\ &\le \text{const}\,(1+\mu^{-1}) \end{aligned} \tag{24}$$

$$\begin{aligned}\kappa &\le 1 + \|M_u\|_{L_2\to L_2} + \text{const}\,\|M_y\|_{W_2^1\to(W_2^1)^*} + \text{const}\\ &\le \text{const}\,(1+\mu^{-1})\end{aligned}$$

As for Δu and $\Delta\lambda_r$, we first observe

$$\|\tilde z_a^u\|_{L_2} \le \|z_a\|_{L_2} + \|G_u^*(\Psi_\eta^u)^{-1}z_p^u\|_{L_2} \le \text{const}\,(1+\mu^{-1})\|z\|_Z\,,$$

$$\begin{aligned}\|S^*M_yC_y^{-1}z_c\|_{L_2} &\le \|S^*\|_{(W_2^1)^*\to L_2}\|M_y\|_{W_2^1\to(W_2^1)^*}\|C_y^{-1}z_c\|_{W_2^1}\\ &\le \text{const}\,(1+\mu^{-1})\|z_c\|_{L_2} \le \text{const}\,(1+\mu^{-1})\|z\|_Z\,,\end{aligned}$$

and hence

$$\|\hat z_a^u\|_{L_2} \le \text{const}\,(1+\mu^{-1})\|z\|_Z\,. \tag{25}$$

From this we conclude that

$$\|\Delta u\|_{L_2} \le \text{const}\,(1+\mu^{-1})\|z\|_Z \quad\text{and}\quad |\Delta\lambda_r| \le \text{const}\,(1+\mu^{-2}).$$

Moreover, $|\hat z_r| \le \text{const}\,\|z\|_Z$ is evident from (21). Observing that $S : L_2^{n_u} \to (W_2^1)^{n_y}$ and $S^* : (W_1^1)^{n_y*} \to L_\infty^{n_u}$ due to (21), and additionally $R_y^* : \mathbb{R}^{n_r} \to (W_1^1)^{n_y*}$, we infer

$$(S^*M_yS - M_{uy}S - S^*M_{yu}) : L_2^{n_u} \to L_\infty^{n_u} \quad\text{and}\quad S^*R_y^* : \mathbb{R}^{n_r} \to L_\infty^{n_u}\,,$$

such that (22) implies

$$M_u\Delta u = \hat z_a^u - (S^*M_yS - M_{uy}S - S^*M_{yu})\Delta u + S^*R_y^*\Delta\lambda_r \in L_\infty^{n_u}\,.$$

Using assumption 2, the desired regularity $\Delta u \in L_\infty^{n_u}$ is readily established:

$$\|\Delta u\|_{L_\infty} \le \text{const}\,\|\hat z_a^u - (S^*M_yS - M_{uy}S - S^*M_{yu})\Delta u + S^*R_y^*\Delta\lambda_r\|_{L_\infty}\,. \tag{26}$$

In order to estimate the right hand side of (26), we first note that since $\dot y$ appears linearly in c, M_y is a Nemyckii operator. We thus infer

$$\begin{aligned}\|M_y\|_{L_\infty\to L_\infty} &\le \|\partial_y^2 L\|_{L_\infty\to L_\infty} + \|G_y^*\|_{L_\infty\to L_\infty}\|(\Psi_\eta^y)^{-1}\Psi_g^y\|_{L_\infty\to L_\infty}\|G_y\|_{L_\infty\to L_\infty}\\ &\le \text{const}\,(1+\mu^{-1})\,,\end{aligned}$$

where we used Theorem 2.1 to obtain $G_y \in \mathcal{L}(L_1, L_1)$, which implies $G_y^* \in \mathcal{L}(L_\infty, L_\infty)$. Then we derive upper bounds for the individual terms in (26) as follows:

$$\begin{aligned}&\|S^*M_yS - M_{uy}S - S^*M_{yu}\|_{L_2\to L_\infty}\|\Delta u\|_{L_2}\\ &\quad\le \|S^*\|_{L_\infty\to L_\infty}\|M_y\|_{L_\infty\to L_\infty}\|S\|_{L_2\to L_\infty}\text{const}\,(1+\mu^{-1})\|z\|_Z\\ &\quad\le \text{const}\,(1+\mu^{-2})\|z\|_Z\,,\end{aligned}$$

$$\|S^* R_y^*\|_{\mathbb{R}^{n_r} \to L_\infty} |\Delta\lambda_r| \leq \text{const}\,(1+\mu^{-2})\|z\|\,,$$

and $\|\hat{z}_a^u\|_{L_\infty} \leq \text{const}\,(1+\mu^{-1})$ analogously to (25). Thus we conclude

$$\|\Delta u\|_{L_\infty} \leq \text{const}\,(1+\mu^{-2})\|z\|_Z\,. \tag{27}$$

In the final step of the proof, we will now trace back the elimination chain from the beginning. First we get

$$\begin{aligned}
\|\Delta\lambda_c\|_{\mathbb{R}^{n_r}\times L_\infty} &= \|C_y^{-*}(\bar{z}_a^y - M_y C_y^{-1} z_c - (M_{yu} - M_y S)\Delta u - R_y^*\Delta\lambda_r)\|_{\mathbb{R}^{n_r}\times L_\infty} \\
&\leq \text{const}\,\|\bar{z}_a^y - M_y C_y^{-1} z_c - (M_{yu} - M_y S)\Delta u - R_y^*\Delta\lambda_r\|_{(W_1^1)^*} \\
&\leq \text{const}\,\big(\,\|\bar{z}_a^y\|_{(W_1^1)^*} + \|M_y\|_{L_\infty\to L_\infty}\|C_y^{-1} z_c\|_{W_1^1} \\
&\qquad + \|M_{yu} - M_y S\|_{L_\infty\to L_\infty}\|\Delta u\|_{L_\infty} \\
&\qquad + \|R_y^*\|_{\mathbb{R}^{n_r}\to(W_1^1)^*}|\Delta\lambda_r|\big) \\
&\leq \text{const}\,\big(\,\|z_a^y - G_y^*(\Psi_\eta^y)^{-1}(z_p^y - \Psi_w^y z_s^y)\|_{(W_1^1)^*} + (1+\mu^{-1})\|z\|_Z \\
&\qquad + (1+\mu^{-1})\|\Delta u\|_{L_\infty} + |\Delta\lambda_r|\big) \\
&\leq \text{const}\,\big(\,\|z\|_Z + \|G_y^*\|_{L_\infty\to(W_1^1)^*}\|(\Psi_\eta^y)^{-1}\|_{L_\infty\to L_\infty}\|z_p^y - \Psi_w^y z_s^y\|_{L_\infty} \\
&\qquad + (1+\mu^{-3})\|z\|_Z\big) \\
&\leq \text{const}\,(1+\mu^{-3})\|z\|_Z\,.
\end{aligned} \tag{28}$$

The state Δy is bounded by

$$\|\Delta y\|_{W_\infty^1} \leq \|C_y^{-1} z_c\|_{W_\infty^1} + \|S\|_{L_\infty\to W_\infty^1}\|\Delta u\|_{L_\infty} \leq \text{const}\,(1+\mu^{-2})\|z\|_Z\,. \tag{29}$$

Finally, we obtain for the Lagrange multiplier $\Delta\eta$ the estimate

$$\begin{aligned}
\|\Delta\eta\|_{L_\infty} &\leq \|\Psi_\eta^{-1}\|_{L_\infty\to L_\infty}\big(\|z_p\|_{L_\infty} + \|\Psi_g G\Delta x\|_{L_\infty}\big) \\
&\leq \text{const}\,(1+\mu^{-3})\|z\|_Z\,.
\end{aligned} \tag{30}$$

Collecting equations (23) and (27)–(30) we obtain the claim (20). □

Now we are ready to prove that the central path exists locally, and that it can be continued up to $\mu = 0$ unless it leaves its bounded set of definition.

Corollary 3.4. *Suppose the assumptions of Theorem 3.3 are satisfied. If there are $v_0 \in D$ and $\mu_0 > 0$ with $F(v_0;\mu_0) = 0$, then there exists a maximal open interval $I \subset R_+$ around μ_0 and a continuously differentiable central path $v : I_\mu \to D$ with the following properties:*

1. *$v(\mu_0) = v_0$*
2. *$F(v(\mu);\mu) = 0$ for all $\mu \in I_\mu$*
3. *Either $\operatorname{dist}(v(I_\mu), \partial D) = 0$ or $\inf I_\mu = 0$ holds.*

Proof. Due to Theorems 3.2 and 3.3 there is an open neighborhood of (v_0, μ_0) on which F and $\partial_v F$ are continuous and $\partial_v F$ is bijective. The implicit function theorem (cf. [40, Theorem 4.B]) guarantees the existence of a continuously differentiable central path $v(\mu)$ with $F(v(\mu),\mu) = 0$ on an open interval around μ_0. A closer inspection of the proof of the implicit function theorem and using the bounds derived in Theorems 3.2 and 3.3 shows that there is a constant

$\epsilon = \epsilon(\text{dist}(v_0, D))$ independent of μ such that $v(\mu)$ exists on the open interval $]\mu_0 - \epsilon\mu^{-4}, \mu_0 + \epsilon\mu^{-4}[$.

Let $I_\mu \subset \mathbb{R}_+$ be a maximal open interval around μ_0, such that property 2 holds. Now assume that property 3 does not hold, i.e. $\text{dist}(v(I_\mu), \partial D) \geq \varepsilon > 0$ and $\delta = \inf I_\mu > 0$. We consider $\mu = \delta + \epsilon\delta^{-4}/2$ with $\epsilon = \epsilon(\varepsilon)$. Again, due to the implicit function theorem, there is an open interval $J_\mu =]\mu - \epsilon\mu^{-4}, \mu + \epsilon\mu^{-4}[$ such that property 2 holds on J_μ and hence on $J_\mu \cup I_\mu$. Since $\mu - \epsilon\mu^{-4} < \delta$, this consequence contradicts the maximality of I_μ, and property 3 must be true. □

3.2 Convergence

Corollary 3.4 does not guarantee the existence of the central path for all $\mu > 0$, since the path may reach the boundary of D for some $\mu_{\text{lim}} > 0$. Moreover, the upper bound for $\|\partial_v F(v;\mu)^{-1}\|$ which has been established in Theorem 3.3 is useless for proving convergence of the path towards a Kuhn-Tucker limit point. The two reasons are the possible occurence of Dirac parts in the state constraints' multipliers at the beginning or end of constrained arcs, and the naive block elimination of the multipliers $\Delta\eta$ in the proof of Corollary 3.4.

Under more restrictive assumptions, in particular the restriction to purely control constrained problems, a splitting into nearly active and nearly inactive constraints can be used to show both, boundedness of the central path and independence of $\|\partial_v F(v;\mu)^{-1}\|$ with respect to μ.

Definition 3.5. For some $\rho > 0$ and functions $u \in L_\infty^{n_u}$ and $\eta \in L_\infty^{n_\eta^u}(\Omega)$, define the characteristic function $\chi^A = \chi^A(t; u, \eta, \mu)$ of the nearly active set vector Ω^A componentwise as

$$\chi_i^A(t) = \begin{cases} 1, & \tilde{g}_i^u(u_i(t)) \leq \rho\eta_i^u(t) \\ 0, & \text{otherwise.} \end{cases}$$

The corresponding characteristic function χ^I of the nearly inactive set vector Ω^I is defined as $\mathbf{1} - \chi^A$, where $\mathbf{1} \in L_\infty^{n_\eta^u}$ is the constant function with value 1.

Note that pointwise multiplication with χ^A defines an orthogonal projector onto the corresponding L_∞ space over the nearly active set vector Ω^A.

First we address the issue of the central path leaving a bounded domain of definition. Assuming a suitable constraint qualification for nearly active constraints of points on the central path, we establish a priori bounds for the central path.

Theorem 3.6. *Suppose $n_\eta^y = 0$, i.e. there are no state constraints. Assume that the following conditions are satisfied.*

(i) *The feasible region $D_u := \{u \in L_\infty^{n_u} : g(u) \geq 0\}$ is bounded.*

(ii) *The state contribution function in the state equation is linearly bounded:*

$$|\tilde{c}(u, y)| \leq \text{const}\,(1 + |y|) \quad \textit{for all } y \in \mathbb{R}^{n_y} \textit{ and } u \in D_u$$

Then there is a bounded set $D_y \subset (W_\infty^1)^{n_y}$ such that for all $\mu > 0$ every solution v of $F(v;\mu) = 0$ satisfies $u \in D_u$ and $y \in D_y$.

If, in addition, there is a constant $\beta > 0$ such that the equality constraints and nearly active control constraints satisfy the inf-sup-condition

$$\inf_{h\in\mathbb{R}^{n_r},\xi\in L_\infty^{n_\eta^u}} \sup_{\delta u\in L_1^{n_u}} \frac{h^T\partial_y r(x)\partial_y c(x)^{-1}\partial_u c(x)\delta u + \langle \chi^A\xi, \partial_u g'(u)\delta u\rangle}{(|h| + \|\chi^A\xi\|_{L_\infty^{n_\eta^u}})\,\|\delta u\|_{L_1^{n_u}}} \geq \beta \quad (31)$$

uniformly for central path solutions v with $x \in D_u \times D_y$, then there is a bounded set $D_0 \subset V$ such that $v \in D_0$.

Proof. Suppose $v = (u, y, \lambda_c, \lambda_r, \eta)$ is a central path solution of $F(v;\mu) = 0$ for some $\mu > 0$. Since $\Psi(g(u), \eta) = 0$ implies $g(u) \geq 0$, we have $u \in D_u$ by assumption (i). Assumption (ii) then guarantees the existence of a constant $\gamma_y < \infty$ such that $y \in S(0, \gamma_y) =: D_y$.

Now consider the state part of the adjoint equation

$$\partial_y J(x) - \partial_y c(x)^*\lambda_c - \partial_y r(y)^*\lambda_r = 0\,.$$

Due to the formulation of c as initial value problem, the inverse of $\partial_y c(x) : (W_1^1) \to L_1 \times \mathbb{R}^{n_y}$ is uniformly bounded on $D_u \times D_y$. Thus we can conclude that

$$\begin{aligned}\|\lambda_c\|_{L_\infty\times\mathbb{R}^{n_y}} &\leq \|\partial_y c(x)^{-*}\|_{(W_\infty^1)^*\to L_\infty\times\mathbb{R}^{n_y}}\,\|\partial_y J(x) - \partial_y r(y)^*\lambda_r\|_{(W_\infty^1)^*}\\ &\leq \text{const}\,\|\partial_y J(x) - \partial_y r(y)^*\lambda_r\|_{(W_\infty^1)^*}.\end{aligned}$$

Since $\partial_y\tilde f(x)$ is uniformly bounded in $L_\infty^{n_y}$ for $x \in D_u\times D_y$, so is $\|\partial_y J(x)\|_{(W_\infty^1)^*}$, and we obtain

$$\|\lambda_c\|_{L_\infty\times\mathbb{R}^{n_y}} \leq \text{const}\,(1 + |\lambda_r|). \quad (32)$$

Inserting $\lambda_c = \partial_y c(x)^{-*}(\partial_y J(x) - \partial_y r(y)^*\lambda_r)$ into the control part of the adjoint equation,

$$\partial_u J(x) - \partial_u c(x)^*\lambda_c - g'(u)^*\eta = 0,$$

and splitting the Lagrange multiplier η into nearly active and nearly inactive parts yields

$$\begin{aligned}\partial_u J(x) - \partial_u c(x)^*\partial_y c(x)^{-*}\partial_y J(x) - g'(u)^*\chi^I\eta&\\ = (\partial_y r(y)\partial_y c(x)^{-1}\partial_u c(x))^*\lambda_r + g'(u)^*\chi^A\eta\,.&\end{aligned}$$

Then the inf-sup-condition of assumption (31) provides the estimate

$$\begin{aligned}\beta(|\lambda_r| + \|\chi^A\eta\|_{L_\infty}) &\leq \sup_{u\in L_1}\frac{\langle(\partial_y r(x)\partial_y c(x)^{-1}\partial_u c(x))^*\lambda_r + g'(u)^*\chi^A\eta, u\rangle}{\|u\|_{L_1}}\\ &\leq \|(\partial_y r(x)\partial_y c(x)^{-1}\partial_u c(x))^*\lambda_r + g'(u)^*\chi^A\eta\|_{L_\infty}\\ &= \|\partial_u J(x) - \partial_u c(x)^*\partial_y c(x)^{-*}\partial_y J(x) + g'(u)\chi^I\eta\|_{L_\infty}.\end{aligned}$$

Note that $\|\chi^I\eta\|_{L_\infty}$ is bounded by $\rho^{-1}\|g(u)\|_{L_\infty}$ and $\|\partial_y c(x)^{-*}\partial_y J(x)\|_{L_\infty\times\mathbb{R}^{n_y}}$ is bounded as shown above. Similarly, $\|\partial_u J(x)\|_{L_\infty}$ is bounded. $\|g'(u)\|_{L_\infty\to L_\infty}$ and $\|\partial_u c(x)\|_{L_1\to L_1\times\mathbb{R}^{n_y}}$ are bounded by Theorem 2.1. Thus we conclude that

$$|\lambda_r| + \|\chi^A\eta\|_{L_\infty} \leq \text{const}\,\beta^{-1}.$$

Combining this with $x \in D_u \times D_y$ verifies the boundedness of v. □

The splitting of the domain into nearly active and inactive regions leads also to improved estimates for the dependency of the complementarity function on the homotopy parameter μ.

The reason for the dependence of $\|\partial_v F(v;\mu)^{-1}\|$ on μ in Theorem 3.3 is the increase of $\|\partial_\eta \Psi^{-1}\|$ as $\mu \to 0$. This can be overcome by more sophisticated elimination of variables in the proof. As a preparation, we first prove a refinement of Lemma 3.1.

Lemma 3.7. *The Fischer-Burmeister complementarity function satisfies the following estimates:*

$$\|\chi^A \partial_g \Psi(g(u),\eta)^{-1}\|_{L_\infty \to L_\infty} \leq \left(1 - \frac{\rho}{\sqrt{1+\rho^2}}\right)^{-1} \tag{33}$$

$$\|\chi^I \partial_\eta \Psi(g(u),\eta)^{-1}\|_{L_\infty \to L_\infty} \leq \left(1 - \frac{1}{\sqrt{1+\rho^2}}\right)^{-1} \tag{34}$$

In particular, both bounds are independent of μ.

Proof. In the relevant inequality (15) we now assume that $a \leq \rho b$. This leads to

$$\partial_a \psi(a,b;\mu) \geq 1 - \frac{1}{\sqrt{1+\frac{b^2}{a^2}+\frac{2\mu}{a^2}}} \geq 1 - \frac{1}{\sqrt{1+\frac{1}{\rho^2}+\frac{2\mu}{a^2}}} \geq 1 - \frac{1}{\sqrt{1+\frac{1}{\rho^2}}}\,.$$

On the nearly active region, this assumption holds, such that due to the projection onto the nearly active region the estimate transfers to $\chi^A \partial_g \Psi(g(u),\eta)^{-1}$. Thus, (33) is verified. By symmetry, (34) is verified using the complementary assumption $a > \rho b$. □

Theorem 3.8. *Assume $n_\eta^y = 0$, i.e. only control constraints are present. Suppose there are a bounded set $D \subset V$ and constants $\beta > 0$ and $\alpha > 0$, such that the following conditions hold uniformly for all central path solutions $v = v(\mu) \in D$ with $F(v(\mu);\mu) = 0$ and $\mu > 0$.*

1. *State equation and nearly inactive control constraints satisfy the inf-sup-condition*

$$\inf_{h\in\mathbb{R}^{n_r},\xi\in L_p^{n_\eta^u}} \sup_{\delta u \in L_q^{n_u}} \frac{h^T \partial_y r(x) \partial_y c(x)^{-1} \partial_u c(x) \delta u + \langle \chi^A \xi, g'(u)\delta u\rangle}{(|h| + \|\chi^A \xi\|_{L_p^{n_\eta^u}})\,\|\delta u\|_{L_q^{n_u}}} \geq \beta$$

   for both $(p,q) = (\infty,1)$ and $(p,q) = (2,2)$.

2. *A strengthened Legendre-Clebsch type condition holds for almost all $t \in \Omega$:*

$$\begin{aligned} M_u(t) &:= \partial_u^2 \tilde f(x(t)) - \partial_u^2 \tilde c(x(t))^T \lambda_c(t) - \tilde g''(u(t))^T \eta(t) \\ &\quad + \tilde g'(u(t))^T \partial_\eta \psi(\tilde g(u(t)),\eta(t);\mu)^{-1} \chi^I \partial_g \psi(\tilde g(u(t)),\eta(t);\mu) \tilde g'(u(t)) \end{aligned}$$

   satisfies

$$\begin{aligned} &\xi^T M_u(t) \xi \geq \alpha |\xi|^2 && \text{for } \xi \in \ker \chi^A(t) \tilde g'(u(t)) \\ &\xi^T M_u(t) \xi \geq 0 && \text{otherwise.} \end{aligned}$$

3. The augmented second derivative of the Lagrangian,

$$M = \partial_x^2 L(v) + g'(x)^* \partial_\eta \Psi(g(x),\eta)^{-1} \partial_g \Psi(g(x),\eta) g'(x)\,,$$

is positive definite on the nullspace of the state equation:

$$\begin{aligned} &\langle \xi, M\xi\rangle \geq \alpha \|\xi\|^2_{L_2^{n_u}\times (W_2^1)^{n_y}} && \text{for } \xi \in \ker c'(x) \cap \ker \chi^A g'(u) \\ &\langle \xi, M\xi\rangle \geq 0 && \text{otherwise} \end{aligned}$$

Then $\partial_v F(v;\mu)$ has an inverse which is bounded uniformly for $(v,\mu) \in D \times \mathbb{R}_+$.

Proof. The structure and line of argument is similar to the proof of Theorem 3.3. We therefore concentrate on the differences and extensions. Define $C, C_u, C_y, R, R_y, M_{uy}, M_{yu}$, and S as before. Let $G = -g'(u)$. Define $\Psi_g = \partial_g \Psi(g(u),\eta)$ and analogously Ψ_η. Finally, let $M_u = \partial_u^2 L(v) + G^* \chi^I \Psi_\eta^{-1} \Psi_g G$ and $M_y = \partial_y^2 L(v)$.

As before, the first step consists of eliminating the Lagrange multiplier, but here only the nearly inactive part $\chi^I \eta = \chi^I \Psi_\eta^{-1}(z_p - \Psi_g G \Delta u)$. In order to symmetrize the remaining system, the nearly active part of the complementarity equation is multiplied by Ψ_g^{-1}:

$$\begin{bmatrix} M_u & M_{uy} & C_u^* & & G^*\chi^A \\ M_{yu} & M_y & C_y^* & R_y^* & \\ C_u & C_y & & & \\ & R_y & & & \\ \chi^A G & & & & -\chi^A \Psi_g^{-1} \Psi_\eta \end{bmatrix} \begin{bmatrix} \Delta u \\ \Delta y \\ \Delta\lambda_c \\ \Delta\lambda_r \\ \chi^A \Delta\eta \end{bmatrix} = \begin{bmatrix} \bar z_a^u \\ z_a^y \\ z_c \\ z_r \\ \chi^A \Psi_g^{-1} z_p \end{bmatrix}$$

with $\bar z_a^u = z_a^u - G^* \chi^I \Psi_\eta^{-1} z_p$. Note that χ^A, Ψ_g^{-1}, and Ψ_η commute. Continuing with the elimination of Δy and λ_c as in the proof of Theorem 3.3 we end up with

$$\begin{bmatrix} T & -(R_y S)^* & G^*\chi^A \\ -R_y S & & \\ \chi^A G & & -\chi^A \Psi_g^{-1} \Psi_\eta \end{bmatrix} \begin{bmatrix} \Delta u \\ \Delta\lambda_r \\ \chi^A \Delta\eta \end{bmatrix} = \begin{bmatrix} \hat z_a^u \\ \hat z_r \\ \chi^A \Psi_g^{-1} z_p \end{bmatrix},$$

where $T = M_u + S^* M_y S - (M_{uy} S + S^* M_{yu})$, $\hat z_a^u = \bar z_a^u - M_{uy} C_y^{-1} z_c - S^*(\bar z_a^y - M_y C_y^{-1} z_c)$, and $\hat z_r = z_r - R_y C_y^{-1} z_c$. Due to assumption 3, $M_u + S^* M_y S - (M_{uy} S + S^* M_{yu})$ is positive definite on the nullspace of $\chi^A G$ and positive semidefinite on the whole space. Assumption 1 provides the inf-sup condition for the combined operator

$$\begin{bmatrix} -R_y S \\ \chi^A G \end{bmatrix},$$

and $\chi^A \Psi_g^{-1} \Psi_\eta$ is positive semidefinite. In this situation, the application of Brezzi's splitting theorem is substituted by a theorem of Braess and Blömer [9] on saddle point problems with penalty term. This guarantees the existence of a solution $(\Delta u, \Delta\lambda_r, \chi^A \Delta\eta) \in L_\infty^{n_u} \times \mathbb{R}^{n_r} \times L_\infty(\Omega^A)$ with

$$\|\Delta u\|_{L_2} + |\Delta\lambda_r| + \|\Delta\eta^A\|_{L_2} \leq \text{const}\, \kappa (\|\hat z_a^u\|_{L_2} + |\hat z_r| + \|\chi^A \Psi_g^{-1} z_p\|_{L_2})\,,$$

where $\kappa = \|T\| + \|R_y S\| + \|G_A\| + \|\chi^A \Psi_g^{-1} \Psi_\eta\| + \alpha + \beta$. Note that due to Lemma 3.7 the operators $\chi^A \Psi_g^{-1} \Psi_\eta$ and $\chi^I \Psi_\eta^{-1} \Psi_g$ are bounded independently

of μ. This property is inherited by κ and $\|\chi^A \Psi_g^{-1} z_p\|$, such that $\|\Delta u\|_{L_2}$, $|\Delta\lambda_r|$, and $\|\chi^A \Delta\eta\|_{L_2}$ are bounded independently of μ.

Subsequently, the L_∞-regularity of Δu and $\chi^A \Delta\eta$ is established. As in the proof of Theorem 3.3, we have

$$(S^* M_y S - M_{uy} S - S^* M_{yu})\Delta u + S^* R_y^* \Delta\lambda_r \in L_\infty^{n_u} ,$$

such that for almost all $t \in \Omega$ the finite dimensional linear equation system

$$\begin{bmatrix} M_u(t) & \tilde{g}'(u(t)^T \chi^A(t) \\ \chi^A(t)\tilde{g}'(u(t)) & -B \end{bmatrix} \begin{bmatrix} \Delta u(t) \\ \chi^A \Delta\eta(t) \end{bmatrix} = \begin{bmatrix} a \\ \chi^A(t) b \end{bmatrix}$$

holds, width $B = \chi^A(t)\partial_g \psi(g(u(t)), \eta(t))^{-1} \partial_\eta \psi(g(u(t)), \eta(t))$. Here, a and b denote generic right hand side vectors the norm of which is bounded by a constant independent of μ. By assumption 2, $M_u(t)$ is positive definite on the nullspace of $\tilde{g}'(u(t))$, such that we can again apply the Lemma by Braess and Blömer. This yields

$$|\Delta u(t)| + |\chi^A \Delta\eta(t)| \leq \text{const}\,(\|M_u(t)\| + \|\tilde{g}'(u(t))\| + \|B\| + \alpha + \beta)(|a| + |b|) \quad (35)$$

for almost all $t \in \Omega$, and hence

$$\|\Delta u\|_{L_\infty} \leq \text{const} \quad (36)$$

$$\|\chi^A \Delta\eta^A\|_{L_\infty} \leq \text{const} , \quad (37)$$

independently of μ. Finally, tracing back the elimination stack as in Theorem 3.3 verifies the claim. □

As in Corollary 3.4, local existence of the central path can be shown. Moreover, the a-priori bound of the solution given by Theorem 3.6 eliminates the possibility of premature termination of the path. Finally, the fact that the inverse of $\partial_v F$ can be bounded independently of μ limits the length of the path and thus ensures convergence.

Theorem 3.9. *Assume Theorem 3.6 holds, providing a bounded set $D_0 \subset V$ containing the central path. Define $D = \bigcup_{v \in D_0} S(v, \epsilon)$ for some $\epsilon > 0$. Suppose the assumptions of Theorem 3.8 hold on D.*

If there are $v_0 \in D_0$ and $\mu_0 > 0$ with $F(v_0; \mu_0) = 0$, then the central path $v(\mu)$ exists for all $0 < \mu \leq \mu_0$ and converges to a Kuhn-Tucker point $v(0)$:

$$\|v(\mu) - v(0)\|_V \leq \text{const}\,\sqrt{\mu}$$

Proof. First we notice that due to Theorem 3.2, there is some $\epsilon > 0$ such that $\partial_v F(v; \mu)^{-1}$ is uniformly bounded on the neighborhood

$$U = \bigcup_{(v;\mu) \text{ with } F(v;\mu)=0} S((v, \mu), \epsilon)$$

of the central path solutions $v(\mu)$. As in the proof of Theorem 3.4, the central path exists on a maximal interval I_μ containing μ_0. Since due to Theorem 3.6 this central path is bounded away from ∂D, we have $\inf I_\mu = 0$. Thus, the central path exists for all $0 < \mu \leq \mu_0$.

Next we estimate $\partial_\mu F(v(\mu);\mu)$. Since only the complementarity function Ψ depends on μ, this is given by $\partial\mu\Psi(g(u),\eta;\mu) = -(g(u)^2+\eta^2+2\mu)^{-1/2}$. On the central path we have $g(u)\cdot\eta = \mu$ a.e. and thus

$$\|\partial_\mu \Psi(g(u),\eta;\mu)\|_{L_\infty} \le (4\mu)^{-1/2}.$$

Now the derivative of the central path is given by

$$v'(\mu) = \partial_v F(v(\mu);\mu)^{-1}\partial_\mu F(v(\mu);\mu).$$

Theorem 3.8 yields

$$\|v'(\mu)\|_V \le \|\partial_v F(v(\mu);\mu)^{-1}\|_{Z\to V}\|\partial_\mu F(v(\mu);\mu)\|_Z \le \text{const}\,\mu^{-1/2}. \tag{38}$$

Therefore the central path is uniformly continuous and converges to some limit point $v(0)\in D$ at a rate of

$$\|v(\mu)-v(0)\|_V \le \int_0^\mu \|v'(s)\|_V\,ds \le \text{const}\int_0^\mu s^{-1/2}\,ds = \text{const}\,\sqrt{\mu}.$$

The continuity of F on $D\times[0,\infty[$ implies that $F(v(0);0)=0$, such that $v(0)$ satisfies the first order necessary conditions (8). □

In [38], the assumptions of the preceeding theorems have been verified for a class of simple optimal control problems, and numerical results are given.

4 A short-step pathfollowing method

With the refined estimates from Section 3.2, we can show linear convergence of a short step path-following method. Note that this is a purely theoretical algorithm, since it relies on the exact solution of operator equations in function space and on knowledge of global Lipschitz constants. For an implementable approximation via inexact Newton corrector and inexact tangential predictor we refer to [38].

We consider the following simple algorithm.

Algorithm 4.1.
initialize v_0, μ_0 *such that* $F(v_0;\mu_0)=0$
choose $\sigma<1$ *sufficiently large*
while $\mu_k>0$
  advance $\mu_{k+1}\leftarrow\sigma\mu_k$
  compute one corrector step $\partial_v F(v_k;\mu_{k+1})\delta v_k = -\partial_\mu F(v_k;\mu_{k+1})$
  advance $v_{k+1}\leftarrow v_k+\delta v_k$, $k\leftarrow k+1$

The sequence v_k of iterates converges to the Kuhn-Tucker point $v(0)$.

First we recall the essentials of an affine covariant Newton-Mysovskikh theorem from [16].

Theorem 4.2. *Assume $F: X\to Y$ is a differentiable mapping with $F(x^*)=0$. Assume the derivative $F'(x)$ is invertible on $D=S(x^*,\delta)$ and satisfies*

$$\|F'(x)^{-1}(F'(y)-F'(x))\| \le \omega\|y-x\| \tag{39}$$

for $x, y \in D$. Let the ordinary Newton sequence x^k starting at $x^0 \in D$ be defined by $x^{k+1} = x^k - F'(x^k)^{-1}F(x^k)$. Then x^k converges to x^* at a rate of

$$\|x^{k+1} - x^*\| \le \frac{\omega}{2}\|x^k - x^*\|^2.$$

Theorem 4.3. *Suppose that F satisfies the assumptions of Theorem 3.9, providing a bounded set D. Let $v_0 \in D$ and $\mu_0 > 0$ be given such that $F(v_0; \mu_0) = 0$. Then there is a constant $\sigma < 1$ such that the sequence v_k of iterates generated by Algorithm 4.1 converges linearly to the limit point $v(0)$ of the central path.*

Proof. To begin with, we verify the assumptions of Theorem 4.2. By Theorems 3.2 and 3.8 there are constants γ_1 and γ_2 independent of $\mu \le \mu_0$, such that $\|\partial_v F(v;\mu) - \partial_v F(v(\mu);\mu)\|_{V\to Z} \le \gamma_1 \mu^{-1/2}$ and $\|\partial_v F(v(\mu);\mu)^{-1}\|_{Z\to V} \le \gamma_2$. Omitting the argument μ from F we use the Banach perturbation lemma to derive

$$\begin{aligned}
&\|\partial_v F(v)^{-1}\|_{Z\to V} \\
&\quad \le \|\partial_v F(v(\mu))^{-1}\|_{Z\to V}\|(I - (\partial_v F(v(\mu)) - \partial_v F(v))\partial_v F(v(\mu))^{-1})^{-1}\|_{Z\to Z} \\
&\quad \le \frac{\gamma_2}{1 - \gamma_1\mu^{-1/2}\|v - v(\mu)\|_V\gamma_2} \le 2\gamma_2
\end{aligned}$$

for $v \in D = S\big(v(\mu), \sqrt{\mu}/(2\gamma_2\gamma_1)\big)$. For $v_1, v_2 \in D$ we thus obtain

$$\begin{aligned}
\|\partial_v F(v_1)^{-1}(\partial_v F(v_2) - \partial_v F(v_2))\|_{V\to V} \\
&\le \|\partial_v F(v_1)^{-1}\|_{Z\to V}\|(\partial_v F(v_2) - \partial_v F(v_1))\|_{V\to Z} \\
&\le 2\gamma_2\gamma_1\mu^{-1/2}\|v_2 - v_1\|_V,
\end{aligned}$$

which establishes the Lipschitz condition (39) with

$$\omega(\mu) \le \frac{2\gamma_2\gamma_1}{\sqrt{\mu}}.$$

As in (38) in the proof of Theorem 3.9, we obtain a bound on the derivative of the central path in the form of

$$\|v'(\mu)\|_V \le \frac{\beta}{\sqrt{\mu}}$$

with $\beta < \infty$ independent of μ. Define

$$\delta = (2\gamma_2\gamma_1)^{-1} \quad \text{and} \quad \sigma \ge \left(1 - \frac{\delta}{2(\delta+\beta)}\right)^2. \tag{40}$$

Let us assume by induction that $\|v_k - v(\mu_k)\|_V \le \delta\sqrt{\mu_k}/2$. Then we have

$$\begin{aligned}
\|v_k - v(\sigma\mu_k)\|_V &\le \|v_k - v(\mu_k)\|_V + (1-\sigma)\mu_k \sup_{\mu\in[\sigma\mu_k,\mu_k]}\|v'(\mu)\|_V \\
&\le \frac{\delta\sqrt{\mu_k}}{2} + (1-\sigma)\mu_k\beta(\sigma\mu_k)^{-1/2} \\
&= \sqrt{\mu_k}\left(\frac{\delta}{2} + \frac{\beta}{\sqrt{\sigma}} + \beta\sqrt{\sigma}\right).
\end{aligned}$$

With σ given by (40), some tedious calculation verifies

$$\frac{\delta}{2} + \frac{\beta}{\sqrt{\sigma}} + \beta\sqrt{\sigma} \le \delta\sqrt{\sigma}$$

and hence

$$\|v_k - v(\sigma\mu_k)\|_V \le \delta\sqrt{\mu_k\sigma}.$$

Now the corrector step, which is a Newton step for the problem $F(v;\sigma\mu_k) = 0$, leads to

$$\begin{aligned}\|v_{k+1} - v(\mu_{k+1})\|_V &\le \frac{\omega(\mu)}{2}\|v_k - v(\mu_{k+1})\|_V^2 \le \frac{\omega(\mu)}{2}\delta^2\mu_{k+1} \\ &\le \frac{\delta}{2}\sqrt{\mu_{k+1}},\end{aligned}$$

which completes the induction. As for the convergence of the iterates, we observe that by Theorem 3.9

$$\begin{aligned}\|v_k - v(0)\|_V &\le \|v_k - v(\mu_k)\|_V + \|v(\mu_k) - v(0)\|_V \\ &\le \frac{\delta}{2}\sqrt{\mu_k} + \text{const}\,\sqrt{\mu_k} \\ &\le \text{const}\,\sigma^{k/2}\sqrt{\mu_0},\end{aligned}$$

which proves linear convergence of $v_k \to v(0)$. □

Acknowledgement The author gratefully acknowledges careful reading of the manuscript by A. Schiela.

References

[1] W. Alt and K. Malanowski. The Lagrange-Newton method for state constrained optimal control problems. *Comput. Optim. Appl.*, 4(3):217–239, 1995.

[2] U. Ascher, J. Christiansen, and R.D. Russell. Collocation software for boundary-value odes. *ACM Trans. Math. Softw.*, 7:209–222, 1981.

[3] U. Ascher, R.M.M. Mattheij, and R.D. Russell. *Numerical solution of boundary value problems for ordinary differential equations*. Prentice Hall, 1988.

[4] G. Bader and U. Ascher. A new basis implementation for a mixed order boundary value ode solver. *SIAM J. Sci. Stat. Comput.*, 8:483–500, 1987.

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

[6] H.G. Bock. Numerische Behandlung von zustandsbeschränkten und Chebychef-Steuerungs-Problemen. , Carl-Cranz-Gesellschaft, Oberpfaffenhofen, 1981.

[7] H.G. Bock and K.-J. Plitt. A multiple shooting algorithm for direct solution of optimal control problems. In *Proceedings of the 9th IFAC World Congress, Budapest*. Pergamon Press, 1984.

[8] D. Braess. *Finite Elements*. Cambridge University Press, second edition, 2001.

[9] D. Braess and C. Blömer. A multigrid method for a parameter dependent problem in solid mechanics. *Numer. Math.*, 57, 1990.

[10] F. Brezzi. On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. *Rev. Franc. Automat. Inform. Rech. Operat.*, 8:129–151, 1974.

[11] J.V. Burke and S. Xu. The global linear convergence of a noninterior path-following algorithm for linear complementarity problems. *Math. Oper. Res.*, 23(3):719–734, 1998.

[12] B. Chen and N. Xiu. A global linear and local quadratic noninterior continuation method for nonlinear complementarity problems based on Chen-Mangasarian smoothing functions. *SIAM J. Optim.*, 9(3):605–623, 1999.

[13] C. Chen and O L. Mangasarian. A class of smoothing functions for nonlinear and mixed complementarity problems. *Comput. Optim. Appl.*, 5(2):97–138, 1996.

[14] P. Deuflhard. A modified Newton method for the solution of ill-conditioned systems of nonlinear equations with application to multiple shooting. *Numer. Math.*, 22:289–315, 1974.

[15] P. Deuflhard. A stepsize control for continuation methods and its special application to multiple shooting techniques. *Numer. Math.*, 33:115–146, 1979.

[16] P. Deuflhard and F.A. Potra. Asymptotic mesh independence of Newton-Galerkin methods via a refined Mysovskii theorem. *SIAM J. Numer. Anal.*, 29(5):1395–1412, 1992.

[17] A.L. Dontchev, W.W. Hager, and V.M. Veliov. Uniform convergence and mesh independence of Newton's method for discretized variational problems. *SIAM J. Control Optim.*, 39(3):961–980, 2000.

[18] A. Fischer. A special Newton-type optimization method. *Optimization*, 24:269–284, 1992.

[19] N.I.M. Gould and Ph.L. Toint. Numerical methods for large-scale non-convex quadratic programming. In A.H. Siddiqi and M. Kocvara, editors, *Trends in Industrial and Applied Mathematics*, pages 149–179. Kluwer, 2002.

[20] W. Hager. Runge-kutta methods in optimal control and the transformed adjoint system. *Numer. Math.*, 87:247–282, 2000.

[21] M. Heinkenschloss and M. Tröltzsch. Analysis of the lagrange-sqp-newton method for the control of a phase field equation. *Control and Cybernetics*, 28:177–211, 1999.

[22] M. Hintermüller and M. Hinze. A SQP-semi-smooth Newton-type algorithm applied to control of the instationary Navier-Stokes system subject to control constraints. TR 03-11, Department of Computational and Applied Mathematics, Rice University, 2003.

[23] S. Ito, C.T. Kelley, and E.W. Sachs. Inexact primal-dual interior point iteration for linear programs in function spaces. *Comput. Optim. Appl.*, 4:189–201, 1995.

[24] C. Kanzow. Some noninterior continuation methods for linear complementarity problems. *SIAM J. Matrix Anal. Appl.*, 17(4):851–868, 1996.

[25] S. Kurcyusz. On the Existence and Nonexistence of Lagrange Multipliers in Bachach Spaces. *J. Optim. Theory Appl.*, 20(1):81–110, 1976.

[26] K.C.P. Machielsen. *Numerical solution of optimal control problems with state constraints by sequential quadratic programming in function space*, volume 53 of *CWI Tract*. Centrum voor Wiskunde en Informatica, Amsterdam, 1988.

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

[28] S.M. Robinson. Generalized equations. In A. Bachem, M. Grötschel, and B. Korte, editors, *Mathematical programming. The state of the art*, pages 346–367. Springer, 1983.

[29] V. Schulz. Solving discretized optimization problems by partially reduced SQP methods. *Comput. Vis. Sci.*, 1:83–96, 1998.

[30] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*. Springer, 1993.

[31] F. Tröltzsch. An sqp method for the optimal control of a nonlinear heat equation. *Control Cybern.*, 23:267–288, 1994.

[32] F. Tröltzsch. On the lagrange-newton-sqp method for the optimal control of semilinear parabolic equations. *SIAM J. Control Optim.*, 38(1):294–312, 1999.

[33] M. Ulbrich. Semismooth Newton methods for operator equations in function spaces. *SIAM J. Optim.*, 13:805–842, 2003.

[34] 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(6):1938–1984, 2000.

[35] 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 Optim.*, 37(3):731–764, 1999.

[36] O. von Stryk. Numerical solution of optimal control problems by direct collocation. In R. Bulirsch, editor, *Optimal control. Calculus of variations, optimal control theory and numerical methods*, volume 111 of *ISNM*, pages 129–143. Birkhäuser, 1993.

[37] M. Weiser. *Function Space Complementarity Methods for Optimal Control Problems*. PhD thesis, Free University of Berlin, Dept. Math. and Comp. Sci., 2001.

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

[39] E.A. Yildirim and S.J. Wright. Warm-start strategies in interior-point methods for linear programming. *SIAM J. Optim.*, 12:782–810, 2002.

[40] E. Zeidler. *Nonlinear Functional Analysis and its Applications*, volume I. Springer, New York, 1986.