

INTERIOR POINT METHODS IN FUNCTION SPACE*

MARTIN WEISER†

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 path-following method is established.

Key words. interior point methods in function space, optimal control, complementarity functions

AMS subject classifications. 49M15, 90C48, 90C51

DOI. 10.1137/S0363012903437277

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, 31] relying on Pontryagin’s maximum principle, and the direct methods [7, 17, 21, 30, 37] based on the Karush–Kuhn–Tucker necessary conditions. Direct methods can be characterized by several features. Among them are the following:

- (i) 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 sequential quadratic programming (SQP) methods where linear-quadratic programs are discretized.

- (ii) Type of optimization algorithm: Among the most popular algorithms employed for solving the optimization problems arising in optimal control are SQP and interior point methods. A recent alternative are semismooth Newton methods [5, 34].

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. This redundancy can be directly exploited by active set-type methods. In contrast, interior point methods are considered to benefit less from this redundancy [20, 40]. Nevertheless, interior point methods are reported to be very efficient for solving optimal control problems—a fact that is not well explained by straightforward application of finite-dimensional interior point convergence theory to

*Received by the editors November 5, 2003; accepted for publication (in revised form) April 14, 2005; published electronically November 23, 2005. This work was supported by the DFG Research Center MATHEON in Berlin. A preprint of this paper appeared as ZIB Report 03-35.

<http://www.siam.org/journals/sicon/44-5/43727.html>

†Zuse Institute Berlin, Takustr. 7, 14195 Berlin, Germany (weiser@zib.de).

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, 22, 23, 27, 32, 33]. Here, Robinson’s theory of generalized equations [29] 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 [35, 36, 24].

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 path-following 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

$$(2.1) \quad \min J(x) \quad \text{subject to} \quad \begin{aligned} c(x) &= 0 \quad \text{a.e.}, \\ r(x) &= 0, \\ g(x) &\geq 0 \quad \text{a.e.} \end{aligned}$$

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

$$(2.2) \quad c(x) = \begin{bmatrix} \bar{c}(x) \\ y(0) - y_0 \end{bmatrix}, \quad \bar{c}(x)(t) = \bar{c}(x(t)) - \dot{y}(t),$$

$$(2.3) \quad r(x) = \tilde{r}(y(1))$$

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 $\tilde{f} : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \rightarrow \mathbb{R}$, $\bar{c} : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_y}$, $\tilde{r} : \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_r}$, $\tilde{g}^u : \mathbb{R}^{n_u} \rightarrow \mathbb{R}^{n_{g^u}}$, and $\tilde{g}^y : \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_{g^y}}$ 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 [38].

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

$$(2.4) \quad |f^{(k)}(x) - f^{(k)}(y)| \leq \kappa|x - y|,$$

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 $\mathbf{f}^{(k)}(u) : (\prod_{j=1}^k L_{pk}^n) \rightarrow L_p^m$ that inherits boundedness and Lipschitz continuity from $f^{(k)}$:

$$(2.5) \quad \left\| \mathbf{f}^{(k)}(u) \right\|_{(\prod_{j=1}^k L_{pk}^n) \rightarrow L_p^m} \leq \sup_{|x| \leq \|u\|_{L_\infty^n}} |f^{(k)}(x)|,$$

$$(2.6) \quad \left\| \mathbf{f}^{(k)}(u + \delta u) - \mathbf{f}^{(k)}(u) \right\|_{(\prod_{j=1}^k L_{pk}^n) \rightarrow L_p^m} \leq \kappa \|\delta u\|_{L_\infty^n}.$$

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

$$(2.7) \quad \left\| \mathbf{f}^{(k)}(u) \right\|_{(\prod_{j=1}^k (W_{pk}^1)^n) \rightarrow (W_p^1)^m} \leq \sup_{|x| \leq \|u\|_{L_\infty^n}} (k + 1)|f^{(k)}(x)| + |f^{(k+1)}(x)|,$$

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

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

With Theorem 2.1 earlier, we conclude that

$$J : L_\infty^{n_u} \times (W_\infty^1)^{n_y} \rightarrow \mathbb{R},$$

$$c : L_\infty^{n_u} \times (W_\infty^1)^{n_y} \rightarrow L_\infty^{n_y}, \quad \text{and}$$

$$g : L_\infty^{n_u} \times (W_\infty^1)^{n_y} \rightarrow L_\infty^{n_u} \times L_\infty^{n_y}$$

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 $\lambda_c \in \mathbb{R}^{n_y} \times (L_\infty^{n_y})^*$, $\lambda_r \in \mathbb{R}^{n_r}$, and $\eta \in (L_\infty^{n_u})^* \times ((W_\infty^1)^{n_y})^*$ such that the following conditions are satisfied:

$$(2.8) \quad J'(x_*) - c'(x_*)^* \lambda_c - r'(x_*)^* \lambda_r - g'(x_*)^* \eta = 0,$$

$$c(x_*) = 0, \quad r(x_*) = 0,$$

$$g(x_*) \geq 0, \quad \eta \geq 0, \quad \langle \eta, g(x_*) \rangle = 0.$$

Under certain assumptions (see, e.g., [26, 28]) these conditions are necessary for x_* to be a local solution of (2.1). Thus, Kuhn–Tucker points are promising candidates for solutions.

Unfortunately, the unwieldy complementarity condition (2.8) is difficult to handle numerically. The idea of primal-dual interior point methods is to relax the complementarity condition by

$$(2.9) \quad \eta \cdot g(x) = \mu, \quad \eta \geq 0, \quad g(x) \geq 0$$

and to consider the homotopy $\mu \rightarrow 0$. Alternatively, complementarity functions $\psi(a, b; \mu) : \mathbb{R}^2 \times \mathbb{R} \rightarrow \mathbb{R}$ 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 (2.9).

These relaxations, however, are only well defined if $\eta \in L_1$, and are continuously differentiable only in case $\eta \in L_\infty$. Note that this is required to hold only 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 $\eta \in L_\infty^{n_u} \times L_\infty^{n_\eta} \subset (L_\infty^{n_u})^* \times ((W_\infty^1)^{n_\eta})^*$ 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

$$(2.10) \quad F(x, \lambda_c, \lambda_r, \eta; \mu) = \begin{bmatrix} \partial_x L(x, \lambda_c, \lambda_r, \eta) \\ -c(x) \\ -r(x) \\ \Psi(\eta, g(x); \mu) \end{bmatrix}.$$

As will be shown in Theorem 3.2 later, F maps

$$(2.11) \quad V \times \mathbb{R}_+ = (L_\infty^{n_u} \times (W_\infty^1)^{n_\eta}) \times (\mathbb{R}^{n_y} \times L_\infty^{n_y}) \times \mathbb{R}^{n_r} \times (L_\infty^{n_u} \times L_\infty^{n_\eta}) \times \mathbb{R}_+$$

into

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

3. The central path. The main object of analytical interest is the *central path* defined by the homotopy (2.9) in μ . First we consider its actual existence in the regular setting given by (2.11) before discussing convergence.

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

$$(3.1) \quad \psi(a, b; \mu) = a + b - \sqrt{a^2 + b^2 + 2\mu}$$

as an example from a large class of different complementarity functions (see [11, 12, 13, 25]).

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 (3.1) 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*

$$(3.2) \quad \|\partial_g \Psi\|_{L_\infty \rightarrow L_\infty} \leq 2,$$

$$(3.3) \quad \|(\partial_g \Psi)^{-1}\|_{L_\infty \rightarrow L_\infty} \leq \max\left(3, \frac{2}{\mu} \|g\|_{L_\infty}^2\right),$$

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 operator 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

$$(3.4) \quad \begin{aligned} \min\left(\frac{\mu}{2a^2}, \frac{1}{3}\right) &= 1 - \max\left(1 - \frac{\mu}{2a^2}, \frac{2}{3}\right) \leq 1 - \frac{1}{\sqrt{1 + \frac{2\mu}{a^2}}} \\ &\leq 1 - \frac{1}{\sqrt{1 + \frac{b^2}{a^2} + \frac{2\mu}{a^2}}} = 1 - \frac{|a|}{\sqrt{a^2 + b^2 + 2\mu}} \\ &\leq \partial_a \psi(a, b; \mu) \\ &\leq 1 + \frac{|a|}{\sqrt{a^2 + b^2 + 2\mu}} \leq 2. \end{aligned}$$

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 (3.2) and has an inverse that is bounded by (3.3).

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 $\partial_g \Psi$. \square

THEOREM 3.2. *The complementarity formulation (2.10) 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*

$$(3.5) \quad \|\partial_v F(v + \delta v; \mu) - \partial_v F(v; \mu)\|_{V \rightarrow Z} \leq c(1 + \mu^{-1/2}) \|\delta v\|_V$$

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 $J(x) = \langle \mathbf{1}, \tilde{\mathbf{f}}(x) \rangle$ 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

$$(3.6) \quad J'(x) = \tilde{\mathbf{f}}'(x)^* \mathbf{1} \in (L_1^{n_u} \times (W_1^1)^{n_y})^* .$$

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

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

again by Theorem 2.1 such that

$$(3.7) \quad c'(x)^* \lambda_c \in (L_1^{n_u} \times (W_1^1)^{n_y})^* .$$

Similarly, we obtain

$$(3.8) \quad r'(x)^* \lambda_r \in (L_1^{n_u} \times (W_1^1)^{n_y})^* \quad \text{and} \quad g'(x)^* \eta \in (L_1^{n_u} \times (W_1^1)^{n_y})^* .$$

Collecting (3.6)–(3.8), $F(v; \mu) \in Z$ is verified. Continuous differentiability is inherited from J, c, g , and ψ .

As for the Lipschitz continuity of the derivative, we have to estimate the 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 blocks separately. First we see that

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

Since x_1 and x_2 are bounded in terms of D , the derivative of the Nemyckii operator \bar{c} inherits the Lipschitz constant $\kappa_c(D)$ of \bar{c}' due to (2.6) of Theorem 2.1 with $p = \infty$. Thus, we conclude

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

Analogously, we obtain

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

Concerning the dual operators $c'(x)^*$ and $g'(x)^*$, we apply Theorem 2.1 with $p = 1$ in (2.6) and obtain

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

and

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

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

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

as before. In view of

$$c''(x_1)^* \lambda_{c_1} - c''(x_2)^* \lambda_{c_2} = c''(x_1)^* (\lambda_{c_1} - \lambda_{c_2}) + (c''(x_1)^* - c''(x_2)^*) \lambda_{c_2}$$

and the boundedness of $c''(x_1)^*$ due to (2.5) of Theorem 2.1, we derive a constant $\kappa(D)$ for

$$\|c''(x_1)^* \lambda_{c_1} - c''(x_2)^* \lambda_{c_2}\|_{X \rightarrow L^\infty \times ((W_1^1)^{n_y})^*} \leq \bar{\kappa}_c(D) \|v_1 - v_2\|_X.$$

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

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

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 (3.5). \square

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 exist 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,

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

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

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

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

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

$$(3.9) \quad \|\partial_v F(v; \mu)^{-1}\|_{Z \rightarrow V} \leq \text{const}(1 + \mu^{-3})$$

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 $\Psi_\eta = \partial_\eta \Psi(g(x), \eta)$, $\Psi_g = \partial_g \Psi(g(x), \eta)$,

$\Psi_\eta^u = \partial_{\eta^u} \Psi^u(g^u(u), \eta^u)$, $\Psi_g^u = \partial_{g^u} \Psi^u(g^u(u), \eta^u)$, 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 (2.2) and has a bounded solution for any right-hand side. Thus, C_y has a bounded inverse. More precisely, for any $p \geq 1$,

$$(3.10) \quad C_y^{-1} : L_p^{n_y} \rightarrow (W_p^1)^{n_y} \text{ is bounded uniformly for } v \in D.$$

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

$$(3.11) \quad \begin{bmatrix} M_u + S^* M_y S - (M_{uy} S + S^* M_{yu}) & -S^* R_y^* \\ -R_y S & \end{bmatrix} \begin{bmatrix} \Delta u \\ \Delta \lambda_r \end{bmatrix} = \begin{bmatrix} \hat{z}_a^u \\ \hat{z}_r \end{bmatrix}.$$

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 (3.11), 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 (3.10), M_u , $S^* M_y S$, $M_{uy} S$, and $S^* M_{yu}$ can all be continuously extended to L_2 . Then, $M_u + S^* M_y S - (M_{uy} S + S^* M_{yu}) : L_2^{n_u} \rightarrow L_2^{n_u}$ 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 (3.11) with

$$(3.12) \quad \begin{aligned} \|\Delta u\|_{L_2} &\leq \text{const} (\|\hat{z}_a^u\|_{L_2} + \kappa |\hat{z}_r|) \quad \text{and} \\ |\Delta \lambda_r| &\leq \text{const} (\kappa \|\hat{z}_a^u\|_{L_2} + \kappa^2 |\hat{z}_r|), \end{aligned}$$

where

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

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 :

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

$$\begin{aligned}
 \|M_y\|_{W_2^1 \rightarrow (W_2^1)^*} &= \|\partial_y^2 L\|_{W_2^1 \rightarrow (W_2^1)^*} + \|G_y^* (\Psi_\eta^y)^{-1} \Psi_y^y G_y\|_{W_2^1 \rightarrow (W_2^1)^*} \\
 &\leq \text{const} + \|G_y^*\|_{L_2 \rightarrow (W_2^1)^*} \|\Psi_\eta^y\|_{L_2 \rightarrow L_2}^{-1} \|\Psi_y^y\|_{L_2 \rightarrow L_2} \|G_y\|_{W_2^1 \rightarrow L_2} \\
 &\leq \text{const} \left(1 + \|(\Psi_\eta^y)^{-1} \Psi_y^y\|_{L_2 \rightarrow L_2} \right) \\
 (3.13) \quad &\leq \text{const}(1 + \mu^{-1}), \\
 \kappa &\leq 1 + \|M_u\|_{L_2 \rightarrow L_2} + \text{const} \|M_y\|_{W_2^1 \rightarrow (W_2^1)^*} + \text{const} \\
 &\leq \text{const}(1 + \mu^{-1}).
 \end{aligned}$$

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

$$\begin{aligned}
 \|\bar{z}_a^u\|_{L_2} &\leq \|z_a\|_{L_2} + \|G_u^* (\Psi_\eta^u)^{-1} z_p^u\|_{L_2} \leq \text{const}(1 + \mu^{-1}) \|z\|_Z, \\
 \|S^* M_y C_y^{-1} z_c\|_{L_2} &\leq \|S^*\|_{(W_2^1)^* \rightarrow L_2} \|M_y\|_{W_2^1 \rightarrow (W_2^1)^*} \|C_y^{-1} z_c\|_{W_2^1} \\
 &\leq \text{const}(1 + \mu^{-1}) \|z_c\|_{L_2} \leq \text{const}(1 + \mu^{-1}) \|z\|_Z,
 \end{aligned}$$

and hence

$$(3.14) \quad \|\hat{z}_a^u\|_{L_2} \leq \text{const}(1 + \mu^{-1}) \|z\|_Z.$$

From this we conclude that

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

Moreover, $|\hat{z}_r| \leq \text{const} \|z\|_Z$ is evident from (3.10). Observing that $S : L_2^{n_u} \rightarrow (W_2^1)^{n_y}$ and $S^* : (W_1^1)^{n_y^*} \rightarrow L_\infty^{n_u}$ due to (3.10), and additionally $R_y^* : \mathbb{R}^{n_r} \rightarrow (W_1^1)^{n_y^*}$, we infer

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

such that (3.11) implies

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

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

$$(3.15) \quad \|\Delta u\|_{L_\infty} \leq \text{const} \|\hat{z}_a^u - (S^* M_y S - M_{uy} S - S^* M_{yu}) \Delta u + S^* R_y^* \Delta \lambda_r\|_{L_\infty}.$$

In order to estimate the right-hand side of (3.15), we first note that since \dot{y} appears linearly in c , M_y is a Nemyckii operator. We thus infer

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

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 (3.15) as follows:

$$\begin{aligned}
 \|S^* M_y S - M_{uy} S - S^* M_{yu}\|_{L_2 \rightarrow L_\infty} \|\Delta u\|_{L_2} &\leq \|S^*\|_{L_\infty \rightarrow L_\infty} \|M_y\|_{L_\infty \rightarrow L_\infty} \|S\|_{L_2 \rightarrow L_\infty} \text{const}(1 + \mu^{-1}) \|z\|_Z \\
 &\leq \text{const}(1 + \mu^{-2}) \|z\|_Z, \\
 \|S^* R_y^*\|_{\mathbb{R}^{n_r} \rightarrow L_\infty} |\Delta \lambda_r| &\leq \text{const}(1 + \mu^{-2}) \|z\|,
 \end{aligned}$$

and $\|\hat{z}_a^u\|_{L_\infty} \leq \text{const}(1 + \mu^{-1})$ analogously to (3.14). Thus, we conclude

$$(3.16) \quad \|\Delta u\|_{L_\infty} \leq \text{const}(1 + \mu^{-2})\|z\|_Z.$$

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

$$(3.17) \quad \begin{aligned} \|\Delta \lambda_c\|_{\mathbb{R}^{n_r} \times L_\infty} &= \|C_y^{-*} \left(\bar{z}_a^y - M_y C_y^{-1} z_c - (M_{yu} - M_y S) \Delta u - R_y^* \Delta \lambda_r \right)\|_{\mathbb{R}^{n_r} \times L_\infty} \\ &\leq \text{const} \left\| \bar{z}_a^y - M_y C_y^{-1} z_c - (M_{yu} - M_y S) \Delta u - R_y^* \Delta \lambda_r \right\|_{(W_1^1)^*} \\ &\leq \text{const} \left(\|\bar{z}_a^y\|_{(W_1^1)^*} + \|M_y\|_{L_\infty \rightarrow L_\infty} \|C_y^{-1} z_c\|_{W_1^1} \right. \\ &\quad \left. + \|M_{yu} - M_y S\|_{L_\infty \rightarrow L_\infty} \|\Delta u\|_{L_\infty} \right. \\ &\quad \left. + \|R_y^*\|_{\mathbb{R}^{n_r} \rightarrow (W_1^1)^*} |\Delta \lambda_r| \right) \\ &\leq \text{const} \left(\|z_a^y - G_y^* (\Psi_\eta^y)^{-1} (z_p^y - \Psi_w^y z_s^y)\|_{(W_1^1)^*} + (1 + \mu^{-1}) \|z\|_Z \right. \\ &\quad \left. + (1 + \mu^{-1}) \|\Delta u\|_{L_\infty} + |\Delta \lambda_r| \right) \\ &\leq \text{const} \left(\|z\|_Z + \|G_y^*\|_{L_\infty \rightarrow (W_1^1)^*} \|(\Psi_\eta^y)^{-1}\|_{L_\infty \rightarrow L_\infty} \|z_p^y - \Psi_w^y z_s^y\|_{L_\infty} \right. \\ &\quad \left. + (1 + \mu^{-3}) \|z\|_Z \right) \\ &\leq \text{const}(1 + \mu^{-3}) \|z\|_Z. \end{aligned}$$

The state Δy is bounded by

$$(3.18) \quad \|\Delta y\|_{W_1^1} \leq \|C_y^{-1} z_c\|_{W_1^1} + \|S\|_{L_\infty \rightarrow W_1^1} \|\Delta u\|_{L_\infty} \leq \text{const}(1 + \mu^{-2}) \|z\|_Z.$$

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

$$(3.19) \quad \begin{aligned} \|\Delta \eta\|_{L_\infty} &\leq \|\Psi_\eta^{-1}\|_{L_\infty \rightarrow L_\infty} (\|z_p\|_{L_\infty} + \|\Psi_g G \Delta x\|_{L_\infty}) \\ &\leq \text{const}(1 + \mu^{-3}) \|z\|_Z. \end{aligned}$$

Collecting (3.12) and (3.16)–(3.19) we obtain the claim (3.9). \square

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 $I_\mu \subset \mathbb{R}_+$ around μ_0 and a continuously differentiable central path $v : I_\mu \rightarrow D$ with the following properties:*

1. $v(\mu_0) = v_0$.
2. $F(v(\mu); \mu) = 0$ for all $\mu \in I_\mu$.
3. Either $\text{dist}(v(I_\mu), \partial D) = 0$ 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. [41, Thm. 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. \square

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 occurrence 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 $\eta \in L_\infty^{n_\eta}(\Omega)$, define the characteristic function $\chi^A = \chi^A(t; u, \eta, \mu)$ of the nearly active set vector Ω^A componentwise as

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

The corresponding characteristic function χ^I of the nearly inactive set vector Ω^I is defined as $\mathbf{1} - \chi^A$, where $\mathbf{1} \in L_\infty^{n_u}$ 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:

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

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

$$(3.20) \quad \inf_{h \in \mathbb{R}^{n_r}, \xi \in L_\infty^{n_\xi}} \sup_{\delta u \in L_1^{n_u}} \frac{h^T \partial_y r(x) \partial_y c(x)^{-1} \partial_u c(x) \delta u + \langle \chi^A \xi, g'(u) \delta u \rangle}{(|h| + \|\chi^A \xi\|_{L_\infty^{n_\xi}}) \|\delta u\|_{L_1^{n_u}}} \geq \beta$$

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 $\partial_y c(x) : (W_1^1) \rightarrow L_1 \times \mathbb{R}^{n_y}$ is uniformly bounded on $D_u \times D_y$. Thus, we can conclude that

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

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

$$(3.21) \quad \|\lambda_c\|_{L_\infty \times \mathbb{R}^{n_y}} \leq \text{const}(1 + |\lambda_r|).$$

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 (3.20) 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 earlier. Similarly, $\|\partial_u J(x)\|_{L_\infty}$ is bounded. $\|g'(u)\|_{L_\infty \rightarrow L_\infty}$ and $\|\partial_u c(x)\|_{L_1 \rightarrow L_1 \times \mathbb{R}^{n_y}}$ 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 $\mu \rightarrow 0$. 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:*

$$(3.22) \quad \|\chi^A \partial_g \Psi(g(u), \eta)^{-1}\|_{L_\infty \rightarrow L_\infty} \leq \left(1 - \frac{\rho}{\sqrt{1 + \rho^2}}\right)^{-1},$$

$$(3.23) \quad \|\chi^I \partial_\eta \Psi(g(u), \eta)^{-1}\|_{L_\infty \rightarrow L_\infty} \leq \left(1 - \frac{1}{\sqrt{1 + \rho^2}}\right)^{-1}.$$

In particular, both bounds are independent of μ .

Proof. In the relevant inequality (3.4) 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, (3.22) is verified. By symmetry, (3.23) is verified using the complementary assumption $a > \rho b$. \square

THEOREM 3.8. *Assume $n_\eta^y = 0$; i.e., only control constraints are present. Suppose there exist 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*

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

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

2. *The augmented second derivative of the Lagrangian*

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

is positive semidefinite on the nullspace of the linearized state equation:

$$(3.24) \quad \langle \xi, M\xi \rangle \geq 0 \quad \text{for } \xi \in \ker c'(x),$$

$$(3.25) \quad \langle \xi, M\xi \rangle \geq \alpha \|\xi\|_{L_2^{n_u} \times (W_2^1)^{n_y}}^2 \quad \text{for } \xi \in \ker c'(x) \cap \ker \chi^A g'(u).$$

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

Before delving into the proof, let us briefly discuss the assumptions of Theorem 3.8. Mostly, they have counterparts in well-known optimality conditions, but they need to be extended a priori to a neighborhood of the central path in order to be able to show convergence.

Assumption 1 is a direct generalization of the linear independence constraint qualification (LICQ; see, e.g., [19, Def. 2.9]) from nonlinear programming to the infinite-dimensional setting. It is also a reinterpretation of regular points (cf. [28, (2.1)]) in the setting of interior points. It provides uniqueness of the Lagrange multipliers and is therefore necessary for proving invertibility of $\partial_v F$.

Convexity of the Lagrangian on the nullspace of the linearized state equation is generally required for sufficient second order optimality conditions. In particular, requirement (3.25) can be interpreted as an adaptation of the convexity condition given by Maurer [28, Thm. 3.5], whereas (3.24) is only technically necessary for invoking a certain saddle point lemma in the proof. In the control constrained setting, the Legendre–Clebsch condition that has been assumed explicitly in Theorem 3.3 is implied by the earlier convexity assumption.

LEMMA 3.9. *Assumption 2 of Theorem 3.8 implies a strengthened Legendre–Clebsch-type condition for almost all $t \in \Omega$:*

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

satisfies

$$(3.26) \quad \xi^T M_u(t) \xi \geq 0 \quad \text{for } \xi \in \mathbb{R}^{n_u},$$

$$(3.27) \quad \xi^T M_u(t) \xi \geq \alpha |\xi|^2 \quad \text{for } \xi \in \ker \chi^A(t) \tilde{g}'(u(t)).$$

Proof. Let $\xi \in \mathbb{R}^{n_u}$ be arbitrary and define $\delta u = \xi \chi_{[t-\epsilon, t+\epsilon]}$ for arbitrary $t \in \text{int}(\Omega)$ and sufficiently small $\epsilon > 0$. Defining M_{yu} , M_{uy} , and S as in Theorem 3.3, we introduce $\delta y = S \delta u$ such that $(\delta u, \delta y) \in \ker c'(x)$. From standard ODE theory we know that $\|\delta y\|_{L_\infty} \leq \text{const} \|\delta u\|_{L_1} \leq \text{const} \epsilon$. Let $M_u = \partial_u^2 L(v) + g'(u)^* \chi^I \Psi_\eta (g(u), \eta)^{-1} \Psi_g (g(u), \eta) g'(u)$, $M_y = \partial_y^2 L(v)$, and

$$M = \begin{bmatrix} M_u & M_{uy} \\ M_{yu} & M_y \end{bmatrix}.$$

Since M_y , M_{yu} , and M_{uy} are uniformly bounded Nemyckii operators, we have by (3.24)

$$\begin{aligned} \langle \delta u, M_u \delta u \rangle &= \langle (\delta u, \delta y), M(\delta u, \delta y) \rangle - \langle \delta y, M_y \delta y \rangle - \langle \delta y, M_{uy} \delta u \rangle - \langle \delta u, M_{yu} \delta y \rangle \\ &\geq 0 - \text{const} \|\delta y\|_{L_\infty}^2 - 2 \text{const} \|\delta y\|_{L_\infty} \|\delta u\|_{L_1} \\ &\geq -\text{const} \epsilon^2 \end{aligned}$$

for all t and $\epsilon > 0$, and hence $\xi^T M_u(t) \xi \geq 0$ for all ξ and almost all $t \in \Omega$, which verifies (3.26). Restricting ξ to $\ker \chi^A(t) \tilde{g}'(u(t))$ and using (3.25) instead of (3.24) finally proves (3.27). \square

Proof of Theorem 3.8. 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, define M_u and M_y as in Lemma 3.9.

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} :

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

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 2, T 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 $(\Delta u, \Delta \lambda_r, \chi^A \Delta \eta) \in L_2^{n_u} \times \mathbb{R}^{n_r} \times L_2(\Omega^A)$ 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

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

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

$$(3.28) \quad \begin{bmatrix} M_u(t) & \tilde{g}'(u(t))^T \chi^A(t) \\ \chi^A(t) \tilde{g}'(u(t)) & -B \end{bmatrix} \begin{bmatrix} \Delta u(t) \\ \chi^A \Delta \eta(t) \end{bmatrix} = \begin{bmatrix} a \\ \chi^A(t) b \end{bmatrix}$$

holds, with $B = \chi^A(t) \partial_g \psi(g(u(t)), \eta(t))^{-1} \partial_\eta \psi(g(u(t)), \eta(t))$. Here, a and b denote generic right-hand side vectors the norm of which is bounded by a constant independent of μ . By Lemma 3.9, $M_u(t)$ is positive definite on the nullspace of $\tilde{g}'(u(t))$, such that we can again apply the lemma by Braess and Blömer, now for the finite-dimensional equation (3.28). This yields

$$(3.29) \quad |\Delta u(t)| + |\chi^A \Delta \eta(t)| \leq \text{const} (\|M_u(t)\| + \|\tilde{g}'(u(t))\| + \|B\| + \alpha + \beta) (|a| + |b|)$$

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

$$(3.30) \quad \|\Delta u\|_{L_\infty} \leq \text{const},$$

$$(3.31) \quad \|\chi^A \Delta \eta^A\|_{L_\infty} \leq \text{const}$$

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

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 3.10. *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

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

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

$$(3.32) \quad \|v'(\mu)\|_V \leq \|\partial_v F(v(\mu); \mu)^{-1}\|_{Z \rightarrow V} \|\partial_\mu F(v(\mu); \mu)\|_Z \leq \text{const } \mu^{-1/2}.$$

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

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

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 (2.8). \square

In the remainder of the section, we will apply the preceding theorems to a class of prototypical optimal control problems. We consider

$$\begin{aligned} & \min \int_0^1 \left(\tilde{f}^y(y(t)) + \frac{\alpha}{2} |u(t)|^2 \right) dt \\ & \text{subject to } \dot{y}(t) = Ay(t) + Bu(t), \\ & \quad y(0) = y_0, \\ & \quad a \leq u(t) \leq b. \end{aligned}$$

THEOREM 3.11. *Suppose that \tilde{f}^y is convex and twice Lipschitz-continuously differentiable, $\alpha > 0$, $a < b$, $A \in \mathbb{R}^{n_y \times n_y}$, and $B \in \mathbb{R}^{n_y \times n_u}$. Assume there are v_0 and $\mu_0 > 0$ such that $F(v_0; \mu_0) = 0$. Then the central path $v(\mu)$ converges to a Kuhn-Tucker point $v(0) \in D$ at a rate of*

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

Proof. We restrict the discussion to a scalar control, i.e., $n_u = 1$. The extension to vector valued controls is straightforward but notationally more involved. We start with Theorem 3.6, choosing

$$(3.33) \quad \rho < \frac{1}{\mu_0} \left(\frac{b-a}{2} \right)^2$$

for separating nearly active and nearly inactive constraints. Due to the box constraints and the linearity of the state equation, conditions (i) and (ii) are satisfied. Since no terminal boundary conditions are given, the inf-sup condition (3.20) simplifies to

$$\inf_{\xi \in L^2_p} \sup_{\delta u \in L^1_q} \frac{\langle \chi^A \xi, g'(u) \delta u \rangle}{\|\chi^A \xi\|_{L^2_p} \|\delta u\|_{L^1_q}} \geq \beta \quad \text{with } g'(u) = \begin{pmatrix} I \\ -I \end{pmatrix}.$$

Assume that for a central path solution (v, μ) with $\mu \leq \mu_0$, the lower constraint $u \geq a$ is nearly active at t , i.e., $\rho\eta^a(t) \geq u(t) - a$. For simplicity, we will omit the argument t in the following. Together with (3.33) and the interior point condition $\eta^a(u - a) = \mu = \eta^b(b - u)$ holding for all central path solutions, this implies

$$\begin{aligned} b - u &= b - a - (u - a) \geq b - a - \sqrt{\rho\eta^a(u - a)} = b - a - \sqrt{\rho\mu} \\ &\geq b - a - \frac{b - a}{2} = \frac{b - a}{2} > \sqrt{\rho\mu} = \sqrt{\rho\eta^b(b - u)}. \end{aligned}$$

Squaring and dividing by $b - u$ finally yields $b - u > \rho\eta^b$, which implies that the upper constraint $u \leq b$ is nearly inactive whenever the lower constraint is nearly active. Analogously, the converse can be shown, such that at most one of the two constraints is active. Since in $\chi^A\xi$ at least one component vanishes, we see that

$$(3.34) \quad \inf_{\xi \in L_p^2} \sup_{\delta u \in L_q^1} \frac{\langle \chi^A\xi, g'(u)\delta u \rangle}{\|\chi^A\xi\|_{L_p^2} \|\delta u\|_{L_q^1}} \geq \inf_{\xi \in L_p^1} \sup_{\delta u \in L_q^1} \frac{\langle \xi, \delta u \rangle}{\|\xi\|_{L_p^1} \|\delta u\|_{L_q^1}} \geq 1$$

for both $(p, q) = (\infty, 1)$ and $(p, q) = (2, 2)$, which confirms the inf-sup condition.

Now we verify the assumptions of Theorem 3.8 on the whole space $D = V$. Assumption 1 is again the inf-sup condition (3.34). The Legendre–Clebsch condition 2 is satisfied due to $\alpha > 0$ and the linearity of the constraints, as is the positive definiteness condition 3 for $\partial_x^2 L(v)$. Since Theorem 3.8 thus holds on V , we can apply Theorem 3.10, which yields the claim. \square

Remark 3.12. The main conditions to verify are the inf-sup constraint qualification and the convexity. While the latter has been explicitly assumed, the former is a direct consequence of the box constraints. More complex optimization problems require more work to verify the assumptions of Theorem 3.10. Nonlinearity of the state equation needs to be compensated by convexity and an a priori bound on λ as given by Theorem 3.6 in order to obtain convexity of the Lagrangian with respect to x . The inf-sup constraint qualification can be shown for more general constraints, e.g., pointwise convex polyhedral admissible sets for the control. It needs to be verified that at most n_u constraints are nearly active.

Numerical results for a specific problem of this class are given in [39].

4. A short-step path-following 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 [39].

We consider the following simple algorithm.

ALGORITHM 4.1.

- 1 initialize v_0, μ_0 such that $F(v_0; \mu_0) = 0$
- 2 choose $\sigma < 1$ sufficiently large
- 3 while $\mu_k > 0$
- 4 advance $\mu_{k+1} \leftarrow \sigma\mu_k$
- 5 compute one corrector step $\partial_v F(v_k; \mu_{k+1})\delta v_k = -\partial_\mu F(v_k; \mu_{k+1})$
- 6 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 $F : X \rightarrow Y$ is a differentiable mapping with $F(x^*) = 0$. Assume the derivative $F'(x)$ is invertible on $D = S(x^*, \delta)$ and satisfies

$$(4.1) \quad \|F'(x)^{-1}(F'(y) - F'(x))\| \leq \omega \|y - x\|$$

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

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

THEOREM 4.3. Suppose that F satisfies the assumptions of Theorem 3.10, 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 $\mu \leq \mu_0$, such that $\|\partial_v F(v; \mu) - \partial_v F(v(\mu); \mu)\|_{V \rightarrow Z} \leq \gamma_1 \mu^{-1/2}$ and $\|\partial_v F(v(\mu); \mu)^{-1}\|_{Z \rightarrow V} \leq \gamma_2$. Omitting the argument μ from F , we use the Banach perturbation lemma to derive

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

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

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

which establishes the Lipschitz condition (4.1) with

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

As in (3.32) in the proof of Theorem 3.10, we obtain a bound on the derivative of the central path in the form of

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

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

$$(4.2) \quad \delta = (2\gamma_2\gamma_1)^{-1} \quad \text{and} \quad \sigma \geq \left(1 - \frac{\delta}{2(\delta + \beta)}\right)^2.$$

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

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

With σ given by (4.2), some tedious calculation verifies

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

and hence

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

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

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

which completes the induction. As for the convergence of the iterates, we observe that by Theorem 3.10

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

which proves linear convergence of $v_k \rightarrow v(0)$. \square

Acknowledgment. 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 (1995), pp. 217–239.
- [2] U. ASCHER, J. CHRISTIANSEN, AND R. RUSSELL, *Collocation software for boundary-value ODEs*, *ACM Trans. Math. Software*, 7 (1981), pp. 209–222.
- [3] U. ASCHER, R. MATTHEIJ, AND R. RUSSELL, *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*, Prentice-Hall, Englewood Cliffs, NJ, 1988.
- [4] G. BADER AND U. ASCHER, *A new basis implementation for a mixed order boundary value ODE solver*, *SIAM J. Sci. Statist. Comput.*, 8 (1987), pp. 483–500.
- [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 (2000), pp. 495–521.
- [6] H. BOCK, *Numerische Behandlung von zustandsbeschränkten und Chebychef-Steuerungs-Problemen*, Technical report, Carl-Cranz-Gesellschaft, Oberpfaffenhofen, 1981.
- [7] H. 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, Elmsford, NY, 1984.
- [8] D. BRAESS, *Finite Elements*, 2nd ed., Cambridge University Press, Cambridge, UK, 2001.
- [9] D. BRAESS AND C. BLÖMER, *A multigrid method for a parameter dependent problem in solid mechanics*, *Numer. Math.*, 57 (1990), pp. 747–761.
- [10] F. BREZZI, *On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers*, *Rev. Franc. Automat. Inform. Rech. Oper.*, 8 (1974), pp. 129–151.
- [11] J. BURKE AND S. XU, *The global linear convergence of a noninterior path-following algorithm for linear complementarity problems*, *Math. Oper. Res.*, 23 (1998), pp. 719–734.

- [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 (1999), pp. 605–623.
- [13] C. CHEN AND O. L. MANGASARIAN, *A class of smoothing functions for nonlinear and mixed complementarity problems*, Comput. Optim. Appl., 5 (1996), pp. 97–138.
- [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 (1974), pp. 289–315.
- [15] P. DEUFLHARD, *A stepsize control for continuation methods and its special application to multiple shooting techniques*, Numer. Math., 33 (1979), pp. 115–146.
- [16] P. DEUFLHARD AND F. POTRA, *Asymptotic mesh independence of Newton–Galerkin methods via a refined Mysovskii theorem*, SIAM J. Numer. Anal., 29 (1992), pp. 1395–1412.
- [17] A. DONTCHEV, W. HAGER, AND V. VELIOV, *Uniform convergence and mesh independence of Newton’s method for discretized variational problems*, SIAM J. Control Optim., 39 (2000), pp. 961–980.
- [18] A. FISCHER, *A special Newton-type optimization method*, Optimization, 24 (1992), pp. 269–284.
- [19] A. FORSGREN, P. GILL, AND M. WRIGHT, *Interior methods for nonlinear optimization*, SIAM Rev., 44 (2002), pp. 525–597.
- [20] N. GOULD AND P. TOINT, *Numerical methods for large-scale non-convex quadratic programming*, in Trends in Industrial and Applied Mathematics, A. Siddiqi and M. Kocvara, eds., Kluwer Academic, Norwell, MA, 2002, pp. 149–179.
- [21] W. HAGER, *Runge–Kutta methods in optimal control and the transformed adjoint system*, Numer. Math., 87 (2000), pp. 247–282.
- [22] M. HEINKENSCHLOSS AND M. TRÖLTZSCH, *Analysis of the Lagrange–SQP–Newton method for the control of a phase field equation*, Control Cybernet., 28 (1999), pp. 177–211.
- [23] M. HINTERMÜLLER AND M. HINZE, *A SQP-Semi-Smooth Newton-Type Algorithm Applied to Control of the Stationary Navier–Stokes System Subject to Control Constraints*, Technical report TR 03-11, Department of Computational and Applied Mathematics, Rice University, Houston, TX, 2003.
- [24] S. ITO, C. KELLEY, AND E. SACHS, *Inexact primal-dual interior point iteration for linear programs in function spaces*, Comput. Optim. Appl., 4 (1995), pp. 189–201.
- [25] C. KANZOW, *Some noninterior continuation methods for linear complementarity problems*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 851–868.
- [26] S. KURCYUSZ, *On the existence and nonexistence of Lagrange multipliers in Banach spaces*, J. Optim. Theory Appl., 20 (1976), pp. 81–110.
- [27] K. MACHIELSEN, *Numerical Solution of Optimal Control Problems with State Constraints by Sequential Quadratic Programming in Function Space*, CWI Tract 53, Centrum voor Wiskunde en Informatica, Amsterdam, 1988.
- [28] H. MAURER, *First and second order sufficient optimality conditions in mathematical programming and optimal control*, Math. Program. Study, 14 (1981), pp. 163–177.
- [29] S. ROBINSON, *Generalized equations*, in Mathematical Programming. The State of the Art, A. Bachem, M. Grötschel, and B. Korte, eds., Springer, Berlin, 1983, pp. 346–367.
- [30] V. SCHULZ, *Solving discretized optimization problems by partially reduced SQP methods*, Comput. Vis. Sci., 1 (1998), pp. 83–96.
- [31] J. STOER AND R. BULIRSCH, *Introduction to Numerical Analysis*, Springer, Berlin, 1993.
- [32] F. TRÖLTZSCH, *An SQP method for the optimal control of a nonlinear heat equation*, Control Cybernet., 23 (1994), pp. 267–288.
- [33] F. TRÖLTZSCH, *On the Lagrange–Newton–SQP method for the optimal control of semilinear parabolic equations*, SIAM J. Control Optim., 38 (1999), pp. 294–312.
- [34] M. ULBRICH, *Semismooth Newton methods for operator equations in function spaces*, SIAM J. Optim., 13 (2003), pp. 805–842.
- [35] 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 (2000), pp. 1938–1984.
- [36] 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 (1999), pp. 731–764.
- [37] O. VON STRYK, *Numerical solution of optimal control problems by direct collocation*, in Optimal Control. Calculus of Variations, Optimal Control Theory and Numerical Methods, Internat. Ser. Numer. Math. 111, R. Bulirsch, ed., Birkhäuser, Boston, 1993, pp. 129–143.
- [38] M. WEISER, *Function Space Complementarity Methods for Optimal Control Problems*, Ph.D. thesis, Free University of Berlin, Berlin, 2001.

- [39] M. WEISER AND P. DEUFLHARD, *The Central Path towards the Numerical Solution of Optimal Control Problems*, ZIB report 01-12, Zuse Institute Berlin, 2001.
- [40] E. YILDIRIM AND S. WRIGHT, *Warm-start strategies in interior-point methods for linear programming*, SIAM J. Optim., 12 (2002), pp. 782–810.
- [41] E. ZEIDLER, *Nonlinear Functional Analysis and Its Applications*, Vol. I, Springer, New York, 1986.