ERRATUM: GLOBAL CONVERGENCE OF A NONSMOOTH NEWTON’S METHOD FOR CONTROL-STATE CONSTRAINED OPTIMAL CONTROL PROBLEMS

MATTHIAS GERDTS∗ AND BJÖRN HÜPPING†

ERRATUM to SIAM Journal on Optimization, Vol. 19 (1), pp. 326–350. Lemma 2.5 (i) in the reference is not correct. In order to correct this, additional assumptions and a smoothing step are needed. However, the correction is conceptual only as the existence and construction of the smoothing step are still open questions.

Abstract. We investigate a nonsmooth Newton’s method for the numerical solution of optimal control problems subject to mixed control-state constraints. The necessary conditions are stated in terms of a local minimum principle. By use of the Fischer-Burmeister function the local minimum principle is transformed into an equivalent nonlinear and nonsmooth equation in appropriate Banach spaces. This nonlinear and nonsmooth equation is solved by a nonsmooth Newton’s method. We prove the global convergence and the locally superlinear convergence under certain regularity conditions. The globalized method is based on the minimization of the squared residual norm. Numerical examples for the Rayleigh problem conclude the article.

Key words. optimal control, nonsmooth Newton’s method, control-state constraints, global convergence

AMS subject classifications. 49J15, 49J52, 49M15

1. Introduction. We consider the following optimal control problem subject to mixed control-state constraints:

\[
\begin{align*}
\text{Minimize} & \quad \int_0^1 f_0(x(t), u(t)) dt \\
\text{w.r.t.} & \quad x \in W^{1,\infty}([0,1],\mathbb{R}^{n_x}), u \in L^\infty([0,1],\mathbb{R}^{n_u}), \\
\text{s.t.} & \quad x'(t) = f(x(t), u(t)) \text{ a.e. in } [0,1], \\
& \quad \psi(x(0),x(1)) = 0, \\
& \quad c(x(t),u(t)) \leq 0 \text{ a.e. in } [0,1].
\end{align*}
\]

(\text{OCP})

Without loss of generality the discussion is restricted to autonomous problems on the fixed time interval [0,1]. The functions \( f_0 : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}, f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x}, \psi : \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}, c : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_c}, \) are supposed to be at least twice continuously differentiable w.r.t. to all arguments. As usual, the Banach space \( L^\infty([0,1],\mathbb{R}^n) \) consists of all measurable functions \( h : [0,1] \to \mathbb{R}^n \) with

\[
\|h\|_\infty := \text{ess sup}_{0 \leq t \leq 1} \|h(t)\| < \infty,
\]

where \( \| \cdot \| \) denotes the Euclidian norm on \( \mathbb{R}^n \). For \( 1 \leq r < \infty \) the Banach spaces \( L^r([0,1],\mathbb{R}^n) \) consist of all measurable functions \( h : [0,1] \to \mathbb{R}^n \) with

\[
\|h\|_r := \left( \int_0^1 \|h(t)\|^r dt \right)^{1/r} < \infty.
\]

∗Institut für Mathematik, Universität Würzburg, Am Hubland, 97074 Würzburg, Germany, (gerdts@mathematik.uni-wuerzburg.de)
†School of Mathematics, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom, (huppingb@maths.bham.ac.uk)
For $1 \leq r \leq \infty$ the Banach spaces $W^{1,r}([0,1],\mathbb{R}^n)$ consist of all absolutely continuous functions $h: [0,1] \rightarrow \mathbb{R}^n$ with
\[
\|h\|_{1,r} := \max\{\|h\|_r, \|h'\|_r\} < \infty.
\]

Several approaches towards the numerical solution of OCP have been investigated in the literature. The so-called direct discretization method is based on a discretization of the infinite dimensional optimal control problem and leads to a finite dimensional nonlinear program, cf., e.g., Gerdts [10]. The latter can be solved numerically by suitable programming methods such as, e.g., sequential quadratic programming. The direct discretization method turns out to be very robust in practice. Nevertheless, the computational effort grows at a nonlinear rate with the number of grid points used for discretization. Convergence results for discretized optimal control problems can be found in Dontchev et al. [7, 6], Hager [16], and Malanowski et al. [24].

The so-called indirect method for optimal control problems attempts to satisfy the necessary conditions that are provided by the well-known minimum principle numerically, cf. Hartl et al. [17] for an overview on minimum principles. The exploitation of the minimum principle leads to a nonlinear multi-point boundary value problem that has to be solved numerically, cf. Oberle and Grimm [29] for an implementation of a multiple shooting algorithm. Although the indirect method usually leads to the most accurate solutions, it suffers from the drawback that it requires a good initial guess in order to convergence. One crucial task is to estimate the sequence of active and inactive intervals of the control-state constraint.

We refer to Büskens [3], Gerdts [12], chapter one of Grötschel et al. [14], Ioffe and Tihomirov [18] and the literature cited therein for an overview on direct discretization methods and indirect methods.

Our intention is to analyze the local and global convergence properties of an alternative method – the nonsmooth Newton’s method. The method is based on a nonsmooth reformulation of the necessary optimality conditions and it was introduced for the problem class OCP in Gerdts [11]. A brief outline of the essential ideas of the algorithm is as follows. The reformulation of the necessary conditions leads to the nonsmooth equation
\[
F(z) = 0, \quad F: Z \rightarrow Y,
\]

where $Z$ and $Y$ are appropriate Banach spaces. Application of the globalized nonsmooth Newton’s method generates sequences $\{z^k\}$, $\{d^k\}$ and $\{\alpha_k\}$ related by the iteration
\[
z^{k+1} = S_k(z^k + \alpha_k d^k), \quad k = 0, 1, 2, \ldots.
\]

Herein, the search direction $d^k$ is the solution of the linear operator equation $V_k(d^k) = -F(z^k)$ and the step length $\alpha_k > 0$ is determined by a line-search procedure of Armijo’s type for a suitably defined merit function. The linear operator $V_k$ is chosen from an appropriately defined generalized Jacobian $\partial^* F(z^k)$. The smoothing step $S_k$ is needed to map the search direction into the correct space.

The nonsmooth Newton’s method was investigated in finite dimensions amongst others by Qi [30] and Qi and Sun [31]. Extensions to infinite spaces can be found in Kummer [20, 21], Chen et al. [4], and Ulbrich [32, 33]. Our approach follows the general framework of Ulbrich [32, 33] which was used to solve certain optimal control problems subject to partial differential equations. The novelty of this paper is the
application to the problem class OCP. The application of the nonsmooth Newton’s method to this problem class has not been investigated in detail by now. The structure of the problem is exploited and leads to a new global convergence result in Section 4. Moreover, sufficient conditions for the non-singularity of the operator $V_k$ are derived in Section 3.

The paper is organized as follows. Section 2 introduces the nonsmooth Newton’s method and establishes the locally superlinear convergence under comparatively mild assumptions. Herein, we also correct Lemma 2.5 (i) in Gerds [13], which fails to hold in the present form. In order to guarantee the semi-smoothness of the mapping $F$, it is necessary to consider $F$ as a mapping from $Z = Z_∞$ (using the supremum norm) to $Y = Y_r$ (using an $L_r$-norm) with $r < ∞$ rather than to $Y_∞$. This in turn requires the introduction of a smoothing operator $S_k$ that maps the search direction $d_k$ from $Z_r$ to $Z_∞$. Note, that a smoothing operator is not needed in the finite dimensional case. Appropriate conditions will be formulated that guarantee the local and global convergence of the method. However, the existence and construction of the smoothing step are open questions. In this sense, the resulting algorithm has to be considered a concept. In Section 3 details of the computation of the search direction are shown. It turns out that the search direction solves a linear boundary value problem with a differential-algebraic equation (DAE). If a certain operator is invertible, the so-called index of the DAE is one and the DAE can be transformed easily into an ordinary differential equation. A sufficient condition for the existence of the inverse operator is provided. Section 4 analyzes the global convergence properties of the nonsmooth Newton’s method. Finally, numerical illustrations are presented in Section 5.

2. Local Convergence of the Nonsmooth Newton’s Method. The (augmented) Hamilton function $H : \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_c} \rightarrow \mathbb{R}$ is defined by

$$
H(x, u, \lambda, \eta) := f_0(x, u) + \lambda^\top f(x, u) + \eta^\top c(x, u).
$$

We summarize the well-known minimum principle for OCP. Throughout the rest of the paper we will use the abbreviation $f[t]$ for $f(x(t), u(t))$ and likewise for other functions with time dependent arguments. Moreover, for an index set $I$ and a vector $c$ with components $c_i$ we define $c_I := (c_i)_{i\in I}$.

Let $(x_*, u_*)$ be a (weak) local minimum of OCP and, in addition to the smoothness assumptions made above, let the following assumptions be satisfied at $(x_*, u_*)$:

(i) Linear independence: There exist $\alpha > 0$ and $\beta > 0$ such that

$$
\|c_{I_\alpha(t)}^\top t \| \geq \beta \| t \|
$$

for all $\zeta$ of appropriate dimension. Herein, the index set $I_\alpha$ is defined by

$I_\alpha(t) := \{i \in \{1, \ldots, n_c\} \mid c_i[t] \geq -\alpha\}.$

(ii) Controllability: For every $q \in \mathbb{R}^{n_c}$ there exists a solution of the linear system

$$
x'(t) - f'_0(t)[x(t) - f'_0(t)[u(t)] = 0,
\psi_\sigma x(0) + \psi_x x(1) = q,
\psi_{u}^c(t)[x(t) + c_{I_\alpha(t)}^\top t u(t)] + S_{\alpha}(t) \sigma(t) = 0,
$$

where $S_{\alpha}(t) := \text{diag}(c_{I_\alpha(t)})$ and $c_{I_\alpha(t)} := \min\{c_i[t] + \alpha, 0\}$. Under these assumptions, Malanowski [23] shows in Theorem 4.3 on page 86 the regularity of the Lagrange multipliers associated with OCP. In particular, the multiplier $l_0$ associated with the objective function can be normalized to one and the linear
operator defined by the linear system in (ii) is surjective under the assumptions (i) and (ii), cf. Lemma 4.1 in Malanowski [23]. Under assumptions (i) and (ii) there exist Lagrange multipliers \( \lambda_* \in W^{1,\infty}([0,1], \mathbb{R}^{n_\lambda}), \eta_* \in L^\infty([0,1], \mathbb{R}^{n_\eta}), \) and \( \sigma_* \in \mathbb{R}^{n_\sigma} \) with

\[
x_*'(t) - f(x_*(t), u_*(t)) = 0, \\
\lambda_*'(t) + H'_x(x_*(t), u_*(t), \lambda_*(t), \eta_*(t)) = 0, \\
\psi(x_*(0), x_*(1)) = 0, \\
\lambda_*(0) + \psi_x(x_*(0), x_*(1)) \sigma_* = 0, \\
\lambda_*(1) - \psi_x(x_*(0), x_*(1)) \sigma_* = 0, \\
H'_u(x_*(t), u_*(t), \lambda_*(t), \eta_*(t)) = 0.
\]

Furthermore, the complementarity conditions hold a.e. in \([0,1]\\):

\[
eta_*(t) \geq 0, \quad c(x_*(t), u_*(t)) \leq 0, \quad \eta_*(t)^T c(x_*(t), u_*(t)) = 0.
\]

**Remark 2.1.** Similar necessary conditions can be found in Neustadt [28], Ch. VI.3, and Zeidan [35], Th. 3.1. A regularity condition based on a controllability condition can be found in Zeidan [35], Proposition 4.2. The regularity assumptions (i) and (ii) with suitable extensions occur in the context of sufficient conditions, cf. Malanowski et al. [26], convergence of discretization methods, cf. Dontchev et al. [6], and sensitivity analysis, cf. Maurer and Augustin [27].

Unfortunately, these necessary conditions are not directly solvable for the variable \((x_*, u_*, \lambda_*, \eta_*, \sigma_*)\\) owing to the complementarity conditions. Therefore, the subsequent considerations aim at the reformulation of this set of equalities and inequalities as an equivalent system of equations, which will be solved by a generalized version of Newton’s method. Notice, that if the mixed-control state constraints were not present in the optimal control problem, then the generalized version of Newton’s method will coincide with the (classical) Lagrange-Newton method. The Lagrange-Newton method for control constrained optimal control problems was analyzed in Alt and Malanowski [2]. Under suitable conditions, the authors obtain a locally quadratic convergence rate for problems with control constraints in the first paper (Theorem 4) and a superlinear convergence rate for problems with control and state constraints in the second (Theorem 5.2). This is more than we are able to show for our approach so far, as only a locally superlinear convergence rate will be established in this paper. The results of Alt and Malanowski [1, 2] suggest, that it might be possible to improve the local convergence rate of our method.

The convex and locally Lipschitz continuous Fischer-Burmeister function \( \varphi : \mathbb{R}^2 \to \mathbb{R} \) is defined by

\[
\varphi(a, b) := \sqrt{a^2 + b^2} - a - b,
\]

cf. Fischer [8]. The Fischer-Burmeister function has the nice property that \( \varphi(a, b) = 0 \) holds if and only if \( a, b \geq 0 \) and \( ab = 0 \). Hence, the complementarity conditions (2.7) are equivalent with the equality

\[
\varphi(-c_i(x_*(t), u_*(t)), \eta_*(t)) = 0, \quad i = 1, \ldots, n_c,
\]

that has to hold almost everywhere in \([0,1]\\). Rather than working with the derivative of \( \varphi \), which does not exist at the origin, we will work with Clarke’s generalized
Jacobian of $\varphi$:

$$\partial \varphi(a, b) = \begin{cases} \left( \frac{a}{\sqrt{a^2 + b^2}} - 1, \frac{b}{\sqrt{a^2 + b^2}} - 1 \right), & \text{if } (a, b) \neq (0, 0), \\ \{(s, r) \in \mathbb{R}^2 \mid (s + 1)^2 + (r + 1)^2 \leq 1\}, & \text{if } (a, b) = (0, 0). \end{cases}$$

Notice, that $\partial \varphi(a, b)$ is a nonempty, convex and compact set. For $1 \leq r \leq \infty$ let the Banach spaces

$$Z_r = W^{1,r}([0, 1], \mathbb{R}^{n_x}) \times L^r([0, 1], \mathbb{R}^{n_x}) \times W^{1,r}([0, 1], \mathbb{R}^{n_u}) \times L^r([0, 1], \mathbb{R}^{n_u}) \times \mathbb{R}^{n_v},$$

$$Y_{1,r} = L^r([0, 1], \mathbb{R}^{n_x}) \times L^r([0, 1], \mathbb{R}^{n_u}) \times \mathbb{R}^{n_v} \times \mathbb{R}^{n_v} \times L^r([0, 1], \mathbb{R}^{n_u}),$$

$$Y_{2,r} = L^r([0, 1], \mathbb{R}^{n_v})$$

be equipped with the maximum norm for product spaces and $z_* = (x_*, u_*, \lambda_*, \eta_*, \sigma_*)$. Then, the necessary conditions (2.1)-(2.7) are equivalent with the nonlinear equation

$$F(z_*) = \begin{pmatrix} F_1(z_*) \\ F_2(z_*) \end{pmatrix} = 0,$$

where $F_1 : Z_\infty \to Y_{1,r}$ and $F_2 : Z_\infty \to Y_{2,r}$ denote the smooth and the nonsmooth part of $F : Z_\infty \to Y_r := Y_{1,r} \times Y_{2,r}$ with $1 \leq r \leq \infty$, respectively:

$$F_1(z)(\cdot) := \begin{pmatrix} x'(\cdot) - f(x(\cdot), u(\cdot)) \\ \lambda'(\cdot) + H'_x(x(\cdot), u(\cdot), \lambda(\cdot), \eta(\cdot)) \top \\ \psi(x(0), x(1)) \\ \lambda(0) + \psi'_{x_u}(x(0), x(1)) \top \sigma \\ \lambda(1) - \psi'_{x_u}(x(0), x(1)) \top \sigma \\ H'_u(x(\cdot), u(\cdot), \lambda(\cdot), \eta(\cdot)) \top \end{pmatrix}, \quad F_2(z)(\cdot) := \omega(z(\cdot)),$$

(2.10)

where $\omega = (\omega_1, \ldots, \omega_{n_v}) \top : \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_v} \times \mathbb{R}^{n_v} \to \mathbb{R}^{n_v}$ and

$$\omega_i(x, \bar{u}, \bar{\lambda}, \bar{\eta}, \bar{\sigma}) := \varphi(-c_i(x, \bar{u}, \bar{\eta}), \bar{\sigma}), \quad i = 1, \ldots, n_c.$$

(2.11)

For technical reasons, which become apparent later, we consider $F$ as a mapping from $Z_\infty$ into $Y_r$. However, we note that

$$\text{im}(F) \subseteq Y_\infty \subset Y_r \quad \text{for every } 1 \leq r < \infty.$$

The standard approach to solve (2.9) numerically would be to apply the classical Newton’s method. Unfortunately, the derivative $F'(z^k)$ does not exist since the component $F_2$ is not differentiable. Hence, we have to find a substitute for the derivative $F'$ in the classical Newton’s method. In finite dimensional spaces, such a substitute for locally Lipschitz continuous functions may be chosen from the generalized Jacobian of $F$ defined by

$$\partial F(z) := \text{co} \left\{ V \mid V = \lim_{z_i \to z} D_F F'(z_i) \right\},$$

where $D_F$ denotes the set of points where $F$ is differentiable, cf. Clarke [5]. However, in infinite dimensional spaces it is more difficult to define an appropriate generalized Jacobian since locally Lipschitz continuous functions in general are not differentiable.
almost everywhere. Motivated by the chain rule in finite dimensions we define the point to set mapping $\partial F : Z_\infty \Rightarrow \mathcal{L}(Y_r, Y_r)$ for some $1 \leq r \leq \infty$ according to

$$\partial F(z^k)(z) := \begin{cases} F'(z^k)(z) & S = \text{diag}(s_1, \ldots, s_n), \\ -S(\epsilon_z)'x + \epsilon_z'u + R \eta & R = \text{diag}(r_1, \ldots, r_m), \\ (s_i, r_i) \in \partial \varphi_i[^z] \text{ a.e.,} & s_i(\cdot), r_i(\cdot) \text{ measurable} \end{cases}$$

and use this set as a generalized Jacobian. The same idea was introduced earlier in Ulbrich [32], Def. 3.35, p. 47. Notice that the first component $F_1$ of $F$ in (2.10) as a mapping from $Z_\infty$ to $Y_\infty$ is continuously Fréchet-differentiable with

$$F_1'(z^k)(z) = \begin{pmatrix} x'(\cdot) - f_z(\cdot)x(\cdot) - f_z'(\cdot)u(\cdot) & \lambda'(\cdot) + H'_z[u]x(\cdot) + H''_{zu}[\cdot]u(\cdot) + H'''_{z\lambda}[\cdot]\lambda(\cdot) + H''_{z\eta}[\cdot]\eta(\cdot) \\ \psi'_z x(0) + \psi'_z x(1) & \lambda(0) + \psi'_{z0x0}(\sigma, x(0)) + \psi'_{z0x1}(\sigma, x(1)) + \left(\psi'_{z0}\right)'^\top \sigma \\ \lambda(1) - \psi'_{z1x0}(\sigma, x(0)) - \psi'_{z1x1}(\sigma, x(1)) - \left(\psi'_{z1}\right)'^\top \sigma & H'_z[u]x(\cdot) + H''_{uu}[\cdot]u(\cdot) + H''_{u\lambda}[\cdot]\lambda(\cdot) + H''_{u\eta}[\cdot]\eta(\cdot) \end{pmatrix},$$

provided that the functions $f_0, f, c, \psi$ are twice continuously differentiable w.r.t. all arguments. Herein, all functions are evaluated at $z^k = (z^k, u^k, \lambda^k, \eta^k, \sigma^k) \in Z_\infty$. The Fréchet differentiability from $Z_\infty$ to $Y_\infty$ implies that $F_1$ is continuously Fréchet-differentiable as a mapping from $Z_\infty$ to $Y_1, r$ for every $1 \leq r \leq \infty$, because

$$\lim_{\|h\|_{Z_\infty} \to 0} \frac{\|F_1(z + h) - F_1(z) - F'_1(z)(h)\|_{Y_r}}{\|h\|_{Z_\infty}} \leq \lim_{\|h\|_{Z_\infty} \to 0} \frac{C\|F_1(z + h) - F_1(z) - F'_1(z)(h)\|_{Y_r}}{\|h\|_{Z_\infty}} = 0,$$

where $C > 0$ is a constant.

It is straightforward to show that every $V \in \partial F(z^k)$ defines a linear and bounded operator from $Z_r$ into $Y_r$ for every $1 \leq r \leq \infty$.

Replacing the non-existing Jacobian $F'$ in the classical Newton’s method by the generalized Jacobian $\partial F(z^k)$ leads to the following algorithm. The algorithm makes use of a smoothing operator $S_k : Z_r \to Z_\infty$, see Ulbrich [32], which maps $z^k + d^k \in Z_r$ back to $Z_\infty$.

**Algorithm 2.2 (Local Nonsmooth Newton’s Method).**

1. Choose $z^0 \in Z_\infty$.
2. If some stopping criterion is satisfied, stop.
3. Choose an arbitrary $V_k \in \partial F(z^k)$ and compute the search direction $d^k$ from the linear equation

$$V_k(d^k) = -F(z^k).$$

4. Set $z^{k+1} = S_k(z^k + d^k)$, $k = k + 1$, and goto (1).

The assumptions needed to prove local convergence of the method are similar to those in Qi [30], Qi and Sun [31], Jiang [19], and Ulbrich [32]. $\partial F(z)$ is called nonsingular if for every $V \in \partial F(z)$ the inverse operator $V^{-1}$ exists and if it is linear and bounded, i.e. $V^{-1} \in \mathcal{L}(Y_r, Z_r)$.

**Theorem 2.3.** Let $z_* \in Z_\infty$ be a zero of $F$. Suppose that there exist constants $\Delta > 0$ and $C > 0$ such that for every $\|z - z_*\|_{Z_\infty} < \Delta$ the generalized Jacobian
\( \partial_z F(z) \) of the mapping \( F : Z_\infty \to Y \) is nonsingular and \( \|V^{-1}\|_{\mathcal{L}(Y, Z_\infty)} \leq C \) for every \( V \in \partial_z F(z) \). Moreover, let there exist a constant \( C_S > 0 \) such that

\[
\|S_k(z^k + d^k) - z_\ast\|_{Z_\infty} \leq C_S \|z^k + d^k - z_\ast\|_{Z_\infty}
\]

for all \( k \).

(i) Let

\[
\|F(z) - F(z_\ast) - V(z - z_\ast)\|_{Y_\ast} = o(\|z - z_\ast\|_{Z_\infty}) \quad \forall V \in \partial_z F(z)
\]

as \( \|z - z_\ast\|_{Z_\infty} \to 0 \). Then, for \( z^0 \) sufficiently close to \( z_\ast \) the nonsmooth Newton’s method converges superlinearly to \( z_\ast \) in the norm \( \| \cdot \|_{Z_\infty} \).

(ii) Let

\[
\|F(z) - F(z_\ast) - V(z - z_\ast)\|_{Y_\ast} = O(\|z - z_\ast\|_{Z_\infty}^{1+p}) \quad \forall V \in \partial_z F(z)
\]

as \( \|z - z_\ast\|_{Z_\infty} \to 0 \). Then, for \( z^0 \) sufficiently close to \( z_\ast \) the nonsmooth Newton’s method converges at order \( 1 + p \) to \( z_\ast \) in the norm \( \| \cdot \|_{Z_\infty} \).

Furthermore, if \( F(z^k) \neq 0 \) for all \( k \) and if there is a constant \( \bar{C}_S \) with

\[
\|S_k(z^k + d^k) - z^k\|_{Z_\infty} \leq \bar{C}_S \cdot \|d^k\|_{Z_\ast}
\]

for every \( k \), then the residual values converge superlinearly:

\[
\lim_{k \to \infty} \frac{\|F(z^k+1)\|_{Y_\ast}}{\|F(z^k)\|_{Y_\ast}} = 0.
\]

Proof. Due to the first assumption, the algorithm is well-defined in some neighborhood of \( z_\ast \). It holds

\[
V_k(z^k + d^k - z_\ast) = V_k(z^k - z_\ast) + V_kd^k = V_k(z^k - z_\ast) - F(z^k) + F(z_\ast).
\]

The assertions in (i) and (ii) follow from

\[
\|z^{k+1} - z_\ast\|_{Z_\infty} = \|S_k(z^k + d^k) - z_\ast\|_{Z_\infty}
\]

\[
\leq C_S \|z^k + d^k - z_\ast\|_{Z_\ast}
\]

\[
= C_S \|V_k^{-1}(V_k(z^k - z_\ast) - F(z^k) + F(z_\ast))\|_{Z_\ast}
\]

\[
\leq C_S \|V_k^{-1}\|_{\mathcal{L}(Y_\ast, Z_\ast)} \cdot \|F(z^k) - F(z_\ast) - V_k(z^k - z_\ast)\|_{Y_\ast}
\]

\[
\leq C_S \cdot C \cdot \|F(z^k) - F(z_\ast) - V_k(z^k - z_\ast)\|_{Y_\ast}
\]

\[
= \begin{cases} 
  o(\|z^k - z_\ast\|_{Z_\infty}), & \text{in case (i)}, \\
  O(\|z^k - z_\ast\|_{Z_\infty}^{1+p}), & \text{in case (ii)}. 
\end{cases}
\]

Let \( \varepsilon > 0 \) be arbitrary. According to Equation (2.16) there exists \( \delta > 0 \) with

\[
\|z^{k+1} - z_\ast\|_{Z_\infty} \leq \varepsilon \|z^k - z_\ast\|_{Z_\infty} \quad \text{whenever} \quad \|z^k - z_\ast\|_{Z_\infty} \leq \delta.
\]

Notice, that for any \( \delta > 0 \) there exists some \( k_0(\delta) \) such that \( \|z^k - z_\ast\|_{Z_\infty} \leq \delta \) for every \( k \geq k_0(\delta) \) since \( z^k \) converges to \( z_\ast \). By the local Lipschitz continuity of \( F \) we get

\[
\|F(z^{k+1})\|_{Y_\ast} = \|F(z^{k+1}) - F(z_\ast)\|_{Y_\ast} \leq L \|z^{k+1} - z_\ast\|_{Z_\infty} \leq L \varepsilon \|z^k - z_\ast\|_{Z_\infty}
\]

locally around \( z_\ast \) and the Newton iteration implies

\[
\|z^{k+1} - z^k\|_{Z_\infty} \leq \bar{C}_S \cdot \|V_k^{-1}\|_{\mathcal{L}(Y_\ast, Z_\ast)} \cdot \|F(z^k)\|_{Y_\ast} \leq \bar{C}_S \cdot C \cdot \|F(z^k)\|_{Y_\ast}.
\]
Thus,
\[ \|z^k - z_*\|_{Z_\infty} \leq \|z^{k+1} - z_*\|_{Z_\infty} + \|z^{k+1} - z_*\|_{Z_\infty} \]
\[ \leq \tilde{C}_S \cdot C \cdot \|F(z^k)\|_{Y_r} + \|z^{k+1} - z_*\|_{Z_\infty} \]
and
\[ \|z^k - z_*\|_{Z_\infty} \leq \frac{\tilde{C}_S \cdot C}{1 - \varepsilon} \|F(z^k)\|_{Y_r}. \]
Finally,
\[ \|F(z^{k+1})\|_{Y_r} \leq L \|z^k - z_*\|_{Z_\infty} \leq \frac{L \varepsilon \tilde{C}_S C}{1 - \varepsilon} \|F(z^k)\|_{Y_r}. \]

Since \( F(z^k) \neq 0 \) and \( \varepsilon \) may be arbitrarily small this shows the last assertion. \( \Box \)

**Remark 2.4.**
- The properties (2.12) and (2.13) can be written as
  \[ \sup_{v \in \partial F(z)} \|F(z) - F(z_*) - V(z - z_*)\|_{Y_r} = \mathcal{O}(\|z - z_*\|_{Z_\infty}), \]
  \[ \sup_{v \in \partial F(z)} \|F(z) - F(z_*) - V(z - z_*)\|_{Y_r} = \mathcal{O}(\|z - z_*\|_{Z_\infty}^{1+p}) \]
as \( \|z - z_*\|_{Z_\infty} \to 0 \) and are referred to as semismoothness and \( p \)-order semismoothness of \( F \) at \( z_* \), cf. Ulbrich [32], Def. 3.1, p. 34.
- It suffices if the assumptions are satisfied for certain elements of \( \partial F \) provided that only these elements are used in the algorithm. For the upcoming computations we used the element corresponding to the choices
  \[ s_i(t) = \begin{cases} -1, & \text{if } c_i[t] = 0, \eta_i(t) = 0, \\ \frac{-c_i(t)}{\sqrt{c_i(t)^2 + \eta_i(t)^2}} - 1, & \text{otherwise}, \end{cases} \]
  \[ r_i(t) = \begin{cases} 0, & \text{if } c_i[t] = 0, \eta_i(t) = 0, \\ \frac{-\eta_i(t)}{\sqrt{c_i(t)^2 + \eta_i(t)^2}} - 1, & \text{otherwise}. \end{cases} \]

We now show that the operator \( F : Z_\infty \to Y_r \) is semismooth for every \( 1 \leq r < \infty \).
As mentioned before, the first component \( F_1 \) of \( F \) is continuously Fréchet-differentiable as a mapping from \( Z_\infty \) to \( Y_r \) for every \( 1 \leq r \leq \infty \), if \( f_0, f, c, \psi \) are twice continuously differentiable. The Fréchet-differentiability immediately yields (2.12) for the component \( F_1 \) for \( Y = Y_r, 1 \leq r \leq \infty \). If the second derivatives of \( f_0, f, c, \psi \) are even locally Lipschitz continuous, then \( F_1 \) also satisfies a local Lipschitz condition of type
\[ \|F_1'(z + d) - F_1'(z)\|_{\mathcal{L}(Z_\infty, Y_{1,r})} \leq L \|d\|_{Z_\infty}. \]
Using this property and the mean-value theorem we find
\[ \|F_1(z + d) - F_1(z) - F_1'(z + d)(d)\|_{Y_{1,r}} \leq \int_0^1 \|F_1'(z + td) - F_1'(z + d)(d)\|_{Y_{1,r}} dt \]
\[ \leq \int_0^1 \|F_1'(z + td) - F_1'(z + d(\mathcal{L}(Z_\infty, Y_{1,r}) dt \cdot \|d\|_{Z_\infty} \]
\[ \leq \frac{L}{2} \|d\|_{Z_\infty}^2. \]
and thus (2.13) with \( p = 1 \) holds for \( F_1 \).

The second component \( F_2(z)(t) = \omega(z(t)) \) of \( F \) in (2.10) is a superposition operator as in Ulbrich [32], Sec. 3.3, which maps \( L^\infty \) into \( L' \). It was shown in Ulbrich [32], Theorems 3.44 and 3.48, that the superposition operator \( F_2 \) is semismooth as a mapping from \( Z_\infty \) to \( Y_{2,r} \) for every \( 1 \leq r < \infty \), if the following assumptions are satisfied:

- The operator \( G : Z_\infty \rightarrow Y_{2,r}, 1 \leq r < \infty \), defined by \( G(z)(\cdot) = \langle e(x(\cdot)), u(\cdot), \eta(\cdot) \rangle \) is continuously Fréchet differentiable.
- The mapping \( z \in Z_\infty \mapsto G(z) \in Y_{2,\infty} \) is locally Lipschitz continuous.
- \( \varphi \) is Lipschitz continuous and semismooth.

Please note that \( r = \infty \) is excluded.

The Fischer-Burmeister function \( \varphi : \mathbb{R}^2 \rightarrow \mathbb{R} \) is Lipschitz continuous and 1-order semismooth (and particularly semismooth) according to Fischer [9], Lemma 20. The mapping \( z \in Z_\infty \mapsto G(z) \in Y_{2,\infty} \) is continuously Fréchet differentiable (and thus locally Lipschitz continuous), if \( c \) is continuously differentiable. Using the same reasoning as before, this implies that the operator \( G \) as a mapping from \( Z_\infty \) to \( Y_{2,r} \) for every \( 1 \leq r < \infty \) is continuously Fréchet differentiable. Hence, the operator \( F_2 \) is semismooth as an operator from \( Z_\infty \) to \( Y_{2,r} \) with \( 1 \leq r < \infty \). Summarizing, we obtain the following local convergence result.

**Theorem 2.5.** Let \( z_0 \in Z_\infty \) be a zero of \( F \). Let \( 1 \leq r < \infty \). Suppose that there exist constants \( \Delta > 0 \) and \( C > 0 \) such that for every \( \|z - z_0\|_Z < \Delta \) the generalized Jacobian \( F(z) \) of the mapping \( F : Z_\infty \rightarrow Y_r \) is nonsingular and \( \|V^{-1}\|_{c(Y_r, Z_r)} \leq C \) for every \( V \in \partial F(z) \). Moreover, let there exist a constant \( C_S > 0 \) such that

\[
\|S_k(z^k + d^k) - z_0\|_Z \leq C_S \|d^k\|_Y, \quad \text{for all } k.
\]

Then, the nonsmooth Newton’s method converges locally at a superlinear rate, if \( f_0, f, c, \psi \) are twice continuously differentiable.

Furthermore, if \( F(z^k) \neq 0 \) for all \( k \) and if there is a constant \( C_S \) with \( \|S_k(z^k + d^k) - z^k\|_Z \leq C_S \|d^k\|_Y \) for every \( k \), then the residual values converge superlinearly:

\[
\lim_{k \rightarrow \infty} \frac{\|F(z^{k+1})\|_Y}{\|F(z^k)\|_Y} = 0.
\]

3. Computation of the Search Direction. For brevity we neglect the arguments whenever possible. The linear operator equation \( V_k(d^k) = -F(z^k) \) in step (2) of Algorithm 2.2 reads as

\[
\begin{pmatrix}
x'(t) \\
\lambda'(t)
\end{pmatrix} =
\begin{pmatrix}
f'_x & 0 \\
-H''_{xx} & -H''_{x}\lambda
\end{pmatrix}
\begin{pmatrix}
x(t) \\
\lambda(t)
\end{pmatrix} -
\begin{pmatrix}
f'_u & 0 \\
-H''_{ux} & -H''_{x}\eta
\end{pmatrix}
\begin{pmatrix}
u(t) \\
\eta(t)
\end{pmatrix} -
\begin{pmatrix}
\psi'_{x_0} & 0 \\
0 & \psi'_{x_0}^T
\end{pmatrix}
\begin{pmatrix}
x(0) \\
\lambda(0)
\end{pmatrix} +
\begin{pmatrix}
\psi'_{x_1} & 0 \\
0 & \psi'_{x_1}^T
\end{pmatrix}
\begin{pmatrix}
x(1) \\
\lambda(1)
\end{pmatrix}
\]

and

\[
\begin{pmatrix}
\psi'(x^k(0), x^k(1)) \\
\lambda^k(0) + \psi_{x_0}^T(\sigma^k)
\end{pmatrix}
\begin{pmatrix}
x(0) \\
\lambda(0)
\end{pmatrix} +
\begin{pmatrix}
\psi'(x^k(0), x^k(1)) \\
\lambda(1) - \psi_{x_1}^T(\sigma^k)
\end{pmatrix}
\begin{pmatrix}
x(1) \\
\lambda(1)
\end{pmatrix}
\]

\[
= - \begin{pmatrix}
\psi'(x^k(0), x^k(1)) \\
\lambda^k(0) + \psi_{x_0}^T(\sigma^k)
\end{pmatrix},
\]

\[
= - \begin{pmatrix}
\psi'(x^k(0), x^k(1)) \\
\lambda(1) - \psi_{x_1}^T(\sigma^k)
\end{pmatrix},
\]

(3.2)
and

\[
\mathcal{A} \begin{pmatrix} u \\ \eta \end{pmatrix} + \begin{pmatrix} H''_{u\xi} & H''_{u\lambda} \\ -S\ell_x & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = -\begin{pmatrix} H'_u \\ \omega(z^k(\cdot)) \end{pmatrix},
\]

where

\[
\mathcal{A} := \begin{pmatrix} H''_{u\mu} & (c'_u)^\top \\ -S\ell'_u & R \end{pmatrix}.
\]

Herein, every function is evaluated at the current iterate \( z^k \). If the inverse operator \( \mathcal{A}^{-1} \) exists, equation (3.3) can be solved for \( u \) and \( \eta \) according to

\[
\begin{pmatrix} u \\ \eta \end{pmatrix} = -\mathcal{A}^{-1} \left[ \begin{pmatrix} H''_{u\xi} & H''_{u\lambda} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} + \begin{pmatrix} H'_u \\ \omega(z^k(\cdot)) \end{pmatrix} \right],
\]

A sufficient condition for the non-singularity of \( \mathcal{A} \) is given below in Theorem 3.2. The constant \( \sigma \) in (3.2) can be viewed as a solution of the differential equation \( \sigma' = 0 \).

Introducing (3.5) into the differential equation (3.1), augmenting this system by \( \sigma' = 0 \), and taking into account the boundary conditions (3.2), yields the linear boundary value problem for \( \xi = (x, \lambda, \sigma)^\top \):

\[
\xi' = B\xi + b, \quad E_0\xi(0) + E_1\xi(1) = q,
\]

where

\[
B = \begin{pmatrix} f'_{x} & 0 & 0 \\ -H''_{u\xi} & -H''_{u\lambda} & 0 \\ 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} f'_{u} & 0 \\ -H''_{u\xi} & 0 & 0 \\ 0 & -H''_{u\lambda} & 0 \end{pmatrix} \mathcal{A}^{-1} \begin{pmatrix} H''_{u\xi} & H''_{u\lambda} & 0 \\ 0 & 0 & 0 \end{pmatrix},
\]

\[
b = -\begin{pmatrix} (x^k)' \quad f \\ (\lambda^k)' + H'_{x} \quad 0 \end{pmatrix} + \begin{pmatrix} f'_{u} & 0 \\ -H''_{u\xi} & 0 & 0 \end{pmatrix} \mathcal{A}^{-1} \begin{pmatrix} H'_u \\ \omega(z^k(\cdot)) \end{pmatrix},
\]

\[
E_0 = \begin{pmatrix} \psi'_{x_{1,0}} \quad \sigma^k_{x_{1,0}} & 0 & 0 \\ -\psi'_{x_{0,1} x_{1,0}} & I & \psi'_{x_{0,1}} \quad \sigma^k_{x_{0,1}} \\ 0 & -\psi'_{x_{1,1} x_{1,0}} & I \end{pmatrix},
\]

\[
E_1 = \begin{pmatrix} \psi'_{x_{0,1}} \quad \sigma^k_{x_{0,1}} & 0 & 0 \\ -\psi'_{x_{1,1} x_{0,1}} & I \end{pmatrix},
\]

\[
q = -\begin{pmatrix} \lambda^k(0) + \psi'_{x_{0,1}} \quad \sigma^k \\ \lambda^k(1) - \psi'_{x_{1,1}} \quad \sigma^k \end{pmatrix}.
\]

Hence, in each iteration of Algorithm 2.2 we have to solve the linear boundary value problem (3.6).

If the operator \( \mathcal{A} \) is not invertible, the situation becomes more involved. In this case, equation (3.3) imposes algebraic constraints and the equations (3.1) and (3.3) form a differential algebraic equation (DAE) with an index of at least two. Actually, the case when \( \mathcal{A} \) is invertible corresponds to the index one case. We will not go into detail here and remain this problem open for future research.
We state a sufficient condition for the existence and boundedness of the inverse operator of $\mathcal{A}$ in (3.4). The proof of this condition uses the Banach lemma, cf. Ljusternik and Sobolew [22], Th. 3, p. 108.

**Lemma 3.1 (Banach Lemma).** Let $X_1$ and $X_2$ be Banach spaces and $M, \Delta : X_1 \to X_2$ linear and continuous operators. Let $M^{-1}$ exist and let $\|M^{-1}\Delta\| < 1$. Then, the operator $M + \Delta$ possesses an inverse $(M + \Delta)^{-1}$ and

$$\|(M + \Delta)^{-1}\| \leq \frac{1}{1 - \|M^{-1}\Delta\|}\|M^{-1}\|.$$ 

The following sufficient conditions for the boundedness of $\mathcal{A}^{-1}$ aim at the formulation of conditions that do not assume that the underlying process $z$ satisfies the first-order necessary optimality conditions. This is important in view of globalization of the method as the iterate $z^k$ may be arbitrary.

**Theorem 3.2.** Let $z = (x, u, \lambda, \eta, \sigma) \in Z$ be given. Define the index sets

$$I_z(t) := \{i \in \{1, \ldots, n_e\} \mid c_i(t) = 0, \eta_i(t) > 0\},$$

$$J_z(t) := \{i \in \{1, \ldots, n_e\} \mid |c_i(t)| \leq \gamma \eta_i(t), \eta_i(t) \geq 0\}, \quad \gamma > 0.$$ 

Let the following assumptions hold at $z$:

(i) Let there exist constants $C_1, C_2, C_3$ such that a.e. in $[0, 1]$ it holds

$$\|H_{uu}''[t]\| \leq C_1, \quad \|c''_u[t]\|^2 \leq C_2, \quad \|c'_u[t]\| \leq C_3.$$

(ii) (Coercivity) Let there exist a constant $\alpha > 0$ such that a.e. in $[0, 1]$ it holds

$$d^\top H_{uu}''[t]d \geq \alpha\|d\|^2 \quad \text{for all} \quad d \in \mathbb{R}^{n_u} : c'_{f_z(t), u}[t]d = 0.$$

(iii) (Linear independence) Let there exist constants $\gamma > 0$ and $\beta > 0$ such that a.e. in $[0, 1]$ it holds

$$\|c'_{f_z(t), u}[t] \zeta\| \geq \beta \|\zeta\| \quad \text{for all} \quad \zeta \text{ of appropriate dimension.}$$

Then, a.e. in $[0, 1]$ the inverse operator $\mathcal{A}^{-1}(t)$ exists and it holds $\|\mathcal{A}^{-1}(t)\| \leq C$ for some constant $C$.

**Proof.** In the sequel we will make use of the following notation. For an index set $I \subseteq \{1, \ldots, n_e\}$ let $S_I := \text{diag}(s_i \mid i \in I)$, $R_I := \text{diag}(r_i \mid i \in I)$, and $A_I := (c'_{s_i, u} \mid i \in I)$. Moreover, $I^c := \{1, \ldots, n_e\} \setminus I$ denotes the complementary index set of the index set $I$ and $Q := H_{uu}''$.

Without loss of generality, using row and column permutations the operator $\mathcal{A}$ in (3.4) can be partitioned as

$$\mathcal{A}(t) = \begin{pmatrix}
Q(t) & A_{I_z(t)}(t) & A_{I_z(t)}(t) \\
-S_{I_z(t)}(t)A_{I_z(t)}(t) & R_{I_z(t)}(t) & 0 \\
-S_{I_z(t)}(t)A_{I_z(t)}(t) & 0 & R_{I_z(t)}(t)
\end{pmatrix},$$

where $I_z(t)$ is a suitable index set depending on a constant $0 < \varepsilon < 1$. The idea behind this partition is to collect all indices $i$ with $-\varepsilon \leq r_i(t) \leq 0$ in the set $I_z(t)$, i.e.

$$I_z(t) := \{i \in \{1, \ldots, n_e\} \mid -\varepsilon \leq r_i(t) \leq 0\}.$$ 

Consequently, the index set $I_z^c(t)$ is given by

$$I_z^c(t) := \{i \in \{1, \ldots, n_e\} \mid 1 \geq r_i(t) \leq -\varepsilon\}.$$
Recall, that a.e. in $[0,1]$ we have $(s_i(t), r_i(t)) \in \{(s, r) \in \mathbb{R}^2 \mid (s + 1)^2 + (r + 1)^2 \leq 1\}$ and hence a.e. in $[0,1]$ it holds $-2 \leq s_i(t) \leq 0$ and $-2 \leq r_i(t) \leq 0$ for all $i \in \{1, \ldots, n_c\}$.

Owing to these considerations, a.e. in $[0,1]$ the matrices $R_i$ and $S_i$ are nonsingular and the following estimates hold (w.r.t. the spectral norm):

$$\|R_i\| \leq \varepsilon, \quad \varepsilon < \|R_i^\dagger\| \leq 2, \quad \frac{1}{2} \leq \|R_i^\dagger\| < \frac{1}{\varepsilon},$$

$$0 \leq \|S_i\| \leq 2, \quad 1 - \sqrt{\varepsilon(2 - \varepsilon)} \leq \|S_i\| \leq 2, \quad \frac{1}{2} \leq \|S_i^\dagger\| \leq \frac{1}{1 - \sqrt{\varepsilon(2 - \varepsilon)}}.$$

Herein, and in the sequel as well, we omitted the explicit dependence on $t$ for brevity.

In order to show the nonsingularity of $A$ we investigate the linear equation

$$\begin{pmatrix}
    Q & A_i^\top & A_i^\top \\
    -S_i & A_i & 0 \\
    -S_i & A_i & 0
  \end{pmatrix}
  \begin{pmatrix}
    w_1 \\
    w_2 \\
    w_3
  \end{pmatrix}
  =
  \begin{pmatrix}
    e_1 \\
    e_2 \\
    e_3
  \end{pmatrix}
$$

and we will show that $\|A_ww\| \geq C\|w\|$ holds for all $w = (w_1, w_2, w_3)^\top$ and some $C > 0$.

Since $S_i$ and $R_i$ are nonsingular, we obtain

$$\begin{pmatrix}
    Q + A_i^\top R_i^{-1} S_i^\top A_i & A_i^\top \\
    -S_i & A_i & 0
  \end{pmatrix}
  \begin{pmatrix}
    w_1 \\
    w_2 \\
    w_3
  \end{pmatrix}
  =
  \begin{pmatrix}
    e_1 - A_i^\top R_i^{-1} e_3 \\
    -S_i e_2 \\
    e_3
  \end{pmatrix},$$

$$w_3 = R_i^{-1} e_3 + S_i^\top A_i^\top w_1. \quad (3.7)$$

We will now show that the operator

$$M_\varepsilon := \begin{pmatrix}
    Q + T & A_i^\top \\
    A_i & 0
  \end{pmatrix}, \quad T = A_i^\top R_i^{-1} S_i^\top A_i^\top,$$

is nonsingular for $\varepsilon > 0$ sufficiently small and that there exists a constant $K > 0$ independent of $\varepsilon$ with $\|M_\varepsilon^{-1}\| \leq K$. Notice, that $T$ is symmetric and positive semidefinite as $R_i^{-1} S_i^\top$ is a diagonal matrix with non-negative entries.

We need to specify the index set $I_\varepsilon$ in more detail. It holds

$$I_\varepsilon = \{i \in \{1, \ldots, n_c\} \mid |c_i| \leq \delta \eta_i, \eta_i > 0\}$$

$$\cup \{i \in \{1, \ldots, n_c\} \mid c_i = 0, \eta_i = 0, r_i \geq -\varepsilon\}, \quad (3.8)$$

where $\delta = \frac{\sqrt{\varepsilon(2 - \varepsilon)}}{1 - \varepsilon}$. This can be seen as follows: If $|c_i| \leq \delta \eta_i$ and $\eta_i > 0$ then

$$r_i = \frac{\eta_i}{\sqrt{c_i^2 + \eta_i^2}} - 1 \geq \frac{\eta_i}{\sqrt{\delta^2 \eta_i^2 + \eta_i^2}} - 1 = \frac{1}{\sqrt{1 + \delta^2}} - 1 = -\left(1 - \frac{1}{\sqrt{1 + \delta^2}}\right) = -\varepsilon.$$

Notice, that for those indices with $c_i = 0 = \eta_i$ the corresponding values $(s_i, r_i)$ can be chosen arbitrarily from the set $\{(s, r) \in \mathbb{R}^2 \mid (s + 1)^2 + (r + 1)^2 \leq 1\}$. This explains the second set on the right hand side of (3.8). On the other hand, if $\eta_i < 0$ then $r_i < -1$. If $\eta_i = 0$ and $c_i \neq 0$, then $r_i = -1$. If $|c_i| > \delta \eta_i$ and $\eta_i > 0$, then as above $r_i < -\varepsilon$. Finally, if $c_i = 0 = \eta_i$ and $r_i < -\varepsilon$, then evidently $r_i \notin I_\varepsilon$.

Notice that $I_\varepsilon \subseteq I_\varepsilon$ for every $\varepsilon > 0$ as $c_i = 0$ and $\eta_i > 0$ implies $r_i = 0$. Hence,

$$\{d \in \mathbb{R}^{n_c} \mid A_i d = 0\} \subseteq \{d \in \mathbb{R}^{n_c} \mid A_i d = 0\} \quad \forall \varepsilon > 0$$
and (ii) implies
\[ d^T (Q + T)d \geq d^T Qd \geq \alpha \|d\|^2 \quad \text{for all} \quad d \in \mathbb{R}^{n_u} : A_L d = 0. \tag{3.9} \]

Now, choose \( \varepsilon > 0 \) such that \( \sqrt{\frac{\varepsilon (2-\varepsilon)}{1-\varepsilon}} \leq \gamma \). Then, \( I_\varepsilon \subseteq J_\gamma \) and assumption (iii) implies
\[ \|A_L^T \tilde{\zeta}\| = \|A_L^T \tilde{\zeta} + A_L^T J_\gamma \cap I_\varepsilon \cdot 0\| \geq \beta \|\tilde{\zeta}, 0\| = \beta \|\tilde{\zeta}\| \tag{3.10} \]

for every \( \tilde{\zeta} \) of appropriate dimension.

Using (3.9), (3.10), assumption (i) and the same arguments as Hager [15] in the proof of Lemma 3.2, it can be shown that the matrix \( M_\varepsilon \) is nonsingular and there exists a constant \( K \) with \( \|M_\varepsilon^{-1}\| \leq K \) for every \( \varepsilon > 0 \) satisfying \( \sqrt{\frac{\varepsilon (2-\varepsilon)}{1-\varepsilon}} \leq \gamma \). It is important to point out that the constant \( K \) only depends on the constants \( C_1, C_2, C_3, \alpha, \beta \) but not on \( \varepsilon \).

The operator
\[ \Gamma = \begin{pmatrix} Q + T & A_L^T \\ A_L & -S_{I_\varepsilon}^{-1}R_{I_\varepsilon} \end{pmatrix} = M_\varepsilon + \begin{pmatrix} 0 & 0 \\ 0 & -S_{I_\varepsilon}^{-1}R_{I_\varepsilon} \end{pmatrix} =: M_\varepsilon + \Delta_\varepsilon \]
can be viewed as a perturbation of \( M_\varepsilon \) with
\[ \|\Delta_\varepsilon\| = \|S_{I_\varepsilon}^{-1}R_{I_\varepsilon}\| \leq \|S_{I_\varepsilon}^{-1}\| \cdot \|R_{I_\varepsilon}\| \leq \frac{\varepsilon}{1 - \sqrt{\varepsilon (2-\varepsilon)}}. \]

Let \( \varepsilon > 0 \) be such that
\[ \frac{\varepsilon}{1 - \sqrt{\varepsilon (2-\varepsilon)}} < \frac{1}{K}, \quad \frac{\varepsilon (2-\varepsilon)}{1 - \varepsilon} \leq \gamma. \]

Then, \( \|\Delta_\varepsilon\| < \frac{1}{\|M_\varepsilon^{-1}\|} \) and according to the Banach lemma 3.1, \( \Gamma^{-1} \) exists and there is a constant \( \tilde{K} \) with
\[ \|\Gamma^{-1}\| \leq \tilde{K}. \]

Equation (3.7) yields the estimates
\[ \| (w_1, w_2) \| \leq \| \Gamma^{-1} \| \left( \| e_1 \| + \| A_L^T \| \cdot \| R_{I_\varepsilon}^{-1} \| \cdot \| e_3 \| + \| S_{I_\varepsilon}^{-1} \| \cdot \| e_2 \| \right) \]
\[ \leq \tilde{K} \left( 1 + \frac{C_2}{\varepsilon} + \frac{1}{1 - \sqrt{\varepsilon (2-\varepsilon)}} \right) \| e \| \]
\[ =: \tilde{C} \| e \| \]
and
\[ \| w_3 \| \leq \| R_{I_\varepsilon}^{-1} \| \left( \| e_3 \| + \| S_{I_\varepsilon} \| \cdot \| A_L \| \cdot \| w_1 \| \right) \leq \frac{1}{\varepsilon} \left( 1 + 2C_3 \tilde{C} \right) \| e \|. \]

The triangle inequality yields \( \| w \| \leq \| (w_1, w_2) \| + \| w_3 \| \leq C \| e \| \) where \( C = \tilde{C} + \frac{1}{\varepsilon} \left( 1 + 2C_3 \tilde{C} \right) \) and the assertion follows with Ljusternik and Sobolew [22], Th. 1, p. 106. \( \square \)
Remark 3.3. The assumptions (ii) and (iii) of Theorem 3.2 are related to the linear independence condition and the Legendre-Clebsch condition which were imposed in Malanowski et al. [26], assumptions (A1) and (B), and in Malanowski and Maurer [25], assumptions (A3) and (A4). However, they differ in some details. In particular, as the proof indicates, the region of validity of the uniform linear independence condition in (iii) has to be coupled to the value of the multiplier \( \eta \).

It remains to establish the non-singularity and the boundedness of the inverse of the linear operator defining the boundary value problem (3.6). For \( 1 \leq r \leq \infty \) this operator \( G : W^{1,r}([0,1],\mathbb{R}^{2n_x+n_y}) \rightarrow L^r([0,1],\mathbb{R}^{2n_x+n_y}) \times \mathbb{R}^{2n_x+n_y} =: \Omega \) is defined by

\[
G(\xi)(t) = \begin{pmatrix}
\xi'(t) - B(t)\xi(t) \\
E_0\xi(0) + E_1\xi(1)
\end{pmatrix},
\]

with \( \| (\omega_1,\omega_2) \|_\Omega = \max\{ \| \omega_1 \|_r, \| \omega_2 \|_r \} \).

Theorem 3.4. Let the following assumptions be satisfied.

(i) Let there exist a constant \( C \) such that a.e. in \([0,1]\) it holds \( \| B(t) \| \leq C \).

(ii) Let there exist \( \kappa > 0 \) such that for all \( \zeta \in \mathbb{R}^{2n_x+n_y} \) it holds

\[
\| (E_0\Phi(0) + E_1\Phi(1))\zeta \| \geq \kappa \| \zeta \|,
\]

where \( \Phi \) is a fundamental solution with \( \Phi'(t) = B(t)\Phi(t) \), \( \Phi(0) = I \).

Then, for \( 1 \leq r \leq \infty \) the inverse operator \( G^{-1} \) exists and it holds \( \| G^{-1} \| \leq K \) for some constant \( K \).

Proof. The proof uses a similar reasoning as Malanowski and Maurer [25] in Section 4. Consider the boundary value problem

\[
\begin{align*}
\xi'(t) - B(t)\xi(t) &= \omega_1(t) \\
E_0\xi(0) + E_1\xi(1) &= \omega_2.
\end{align*}
\]

Since \( \| G^{-1} \| = \frac{1}{\inf\{ \| G(\xi) \| \mid \| \xi \|_1 = 1 \}} \), we must show that \( \| (\omega_1,\omega_2) \|_\Omega \geq \| \xi \|_1/r \) for all \( (\omega_1,\omega_2) \in \Omega \) and \( \xi \) solving the above linear equation.

Consider the initial value problem

\[
\tilde{\xi}'(t) = B(t)\tilde{\xi}(t) + \omega_1(t), \quad \tilde{\xi}(0) = 0.
\]

The solution is given implicitly by

\[
\tilde{\xi}(t) = \int_0^t B(\tau)\tilde{\xi}(\tau) + \omega_1(\tau)d\tau.
\]

Gronwall’s lemma yields

\[
\| \tilde{\xi}(t) \| \leq \| \omega_1 \| \exp(\| B \|_\infty) \leq \| \omega_1 \|_1 \exp(C).
\]

Hölder’s inequality implies

\[
\| \tilde{\xi}(t) \| \leq C_1\| \omega_1 \|_r \exp(C)
\]

for \( 1 \leq r \leq \infty \) and some constant \( C_1 \).

Similarly, we find

\[
\| \xi(t) \| \leq (\| \xi(0) \| + C_1\| \omega_1 \|_r \exp(C).
\]
For the fundamental system $\Phi$ we obtain

$$
\|\Phi(t)\| \leq 1 + \|B\|_\infty \int_0^t \|\Phi(\tau)\| d\tau \leq \exp(\|B\|_\infty) \leq \exp(C).
$$

Using the solution formula for linear differential equations we find

$$
\xi(t) = \Phi(t) \left( \xi(0) + \int_0^t \Phi(\tau)^{-1} \omega_1(\tau) d\tau \right) = \Phi(t)\xi(0) + \tilde{\xi}(t).
$$

Moreover,

$$
(E_0\Phi(0) + E_1\Phi(1))\xi(0) = \omega_2 - E_1\Phi(1) \int_0^1 \Phi(\tau)^{-1} \omega_1(\tau) d\tau = \omega_2 - E_1\tilde{\xi}(1).
$$

It follows

$$
\kappa \|\xi(0)\| \leq \|\omega_2\| + \|E_1\|\|\tilde{\xi}(1)\| \leq \|\omega_2\| + C_1\|E_1\|\|\omega_1\|_r \exp(C)
$$

and thus

$$
\|\xi(0)\| \leq \frac{1}{\kappa} (\|\omega_2\| + C_1\|E_1\|\|\omega_1\|_r \exp(C))
$$

$$
\leq \frac{1}{\kappa} (1 + C_1\|E_1\| \exp(C)) \max\{\|\omega_2\|, \|\omega_1\|_r\}
$$

$$
= : \kappa_1\|\omega_1, \omega_2\|_\Omega.
$$

Hence,

$$
\|\xi(t)\| \leq (\|\xi(0)\| + C_1\|\omega_1\|_r \exp(C)) \leq (\kappa_1 + C_1 \exp(C))\|\omega_1, \omega_2\|_\Omega.
$$

With $\kappa_2 := (\kappa_1 + C_1 \exp(C))$ we proved $\|\xi\|_r \leq \kappa_2\|\omega_1, \omega_2\|_\Omega$.

Minkowski's inequality yields

$$
\|\xi'(r)\|_r = \|B(\cdot)\xi(\cdot) + \omega_1(\cdot)\|_r \leq \|B(\cdot)\xi(\cdot)\|_r + \|\omega_1\|_r \leq C_2\|\xi\|_r + \|\omega_1\|_r
$$

and thus

$$
\|\xi'(r)\|_r \leq (1 + C_2\kappa_2)\|\omega_1, \omega_2\|_\Omega.
$$

With $K := \max\{\kappa_2, 1 + C_2\kappa_2\}$ we obtain $\|\xi\|_{1,r} \leq K\|\omega_1, \omega_2\|_\Omega$, which shows the assertion.

**Remark 3.5.** An alternative way to show the (unique) solvability of the boundary value problem (3.6) can be found in sections 3 and 4 of Malanowski and Maurer [25]. The idea is to interpret the boundary value problem (3.6) as the first order necessary optimality conditions for a linear-quadratic accessory problem. Then, the unique solvability of the accessory problem is shown under a complete controllability condition and a coercivity condition for the objective function. The latter will be satisfied if – in addition to other assumptions – a suitably defined Riccati equation has a bounded solution.

A combination of Theorems 3.2 and 3.4 leads to the following result, which establishes the uniform non-singularity of the generalized Jacobian $\partial_z F(z)$ required in Theorems 2.3 and 2.5.
Theorem 3.6. Let $z_*$ be a zero of $F$. Suppose that there exists a constant $\Delta > 0$ such that for every $\|z - z_*\|_{Z_{\infty}} < \Delta$ the assumptions of Theorems 3.2 and 3.4 hold with uniform constants. Then, for every $1 \leq r \leq \infty$ the generalized Jacobian $\partial_r F(z) \in \mathcal{L}(Y_r, Z_r)$ is nonsingular and there exists a constant $C > 0$ such that $\|V^{-1}\|_{\mathcal{L}(Y_r, Z_r)} \leq C$ for every $V \in \partial_r F(z)$.

Remark 3.7. Theorem 3.6 holds for every $1 \leq r \leq \infty$. In particular, this implies that every element $V \in \partial_r F(z)$ maps a function in $Z_r$ to a function in $Y_r$. In particular, if $F(z^k) \in Y_\infty$ then $d^k = -V_k^{-1}F(z^k) \in Z_\infty$. $F(z^k) \in Y_\infty$ holds, if $z^k \in Z_\infty$. Hence, the smoothing operator $S_k$ in step (3) of Algorithm 2.2 can be chosen to be the identity if the initial $z^0$ is chosen to be in $Z_\infty$. But then the condition

$$\|S_k(z^k + d^k) - z_*\|_{Z_{\infty}} \leq C_S\|z^k + d^k - z_*\|_{Z_{r}}$$

reduces to

$$\|z^{k+1} - z_*\|_{Z_{\infty}} \leq C_S\|z^{k+1} - z_*\|_{Z_r}.$$  
Note that these conditions do not hold in general, but may hold for a specific sequence.

4. Globalization. One reason that makes the Fischer-Burmeister function appealing is the fact that its square

$$\phi(a, b) := \varphi(a, b)^2 = \left(\sqrt{a^2 + b^2} - a - b\right)^2$$

is continuously differentiable with $\phi'(a, b) = 2\varphi(a, b)v$, where $v \in \partial\varphi(a, b)$ is arbitrary. Hence, the mappings

$$(\bar{x}, \bar{u}, \bar{\eta}) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\eta} \mapsto \phi(-c_i(\bar{x}, \bar{u}, \bar{\eta}), \quad i = 1, \ldots, n_c)$$

are continuously differentiable by the chain rule. This allows to globalize the local nonsmooth Newton’s method using the squared $L^2$-norm of $F$ as a merit function:

$$\Theta(z) := \frac{1}{2}\|F(z)\|_{V_2}^2$$

$$= \frac{1}{2} \int_0^1 \|x'(t) - f(x(t), u(t))\|^2 dt$$

$$+ \frac{1}{2} \int_0^1 \|\lambda'(t) + H'_s(x(t), u(t), \lambda(t), \eta(t))\|^2 dt$$

$$+ \frac{1}{2} \int_0^1 \|H'_c(x(t), u(t), \lambda(t), \eta(t))\|^2 dt + \frac{1}{2} \sum_{i=1}^{n_c} \int_0^1 \phi(-c_i(x(t), u(t)), \eta_i(t))dt$$

$$+ \frac{1}{2}\|\psi(x(0), x(1))\|^2 + \frac{1}{2}\|\lambda(0) + \psi_{\lambda}(x(0), x(1))\|^2$$

$$+ \frac{1}{2}\|\lambda(1) - \psi_{\lambda}(x(0), x(1))\|^2.$$ 

\(\Theta\) is Fréchet-differentiable in $Z_{\infty}$ if $f_0, f, c, \psi$ are twice continuously differentiable. An analysis of the derivative of $\Theta$ reveals that for $d^k$ with $V_k(d^k) = -F(z^k)$ it holds

$$\Theta'(z^k)(d^k) = -2\Theta(z^k) = -\|F(z^k)\|_{V_2}^2. \quad (4.1)$$

As a consequence, $d^k$ is a direction of descent of $\Theta$ at $z^k$ and for some $\bar{\sigma} \in (0, 1)$ there exists $\alpha > 0$ such that

$$\Theta(z^k + \alpha d^k) \leq \Theta(z^k) + \bar{\sigma}\alpha\Theta'(z^k)(d^k). \quad (4.2)$$
Instead for \( z^k + \alpha d^k \) we intend to perform a line-search using the smoothing operator \( S_k(z^k + \alpha d^k) \). The following growth condition is sufficient to prove the well-posedness of the line-search procedure.

**Assumption 4.1.** For \( z^k \) and \( d^k = -V_k^{-1} F(z^k) \) let there be a smoothing operator \( S_k \) and constants \( 0 \leq L < 1, \gamma > 0 \) with

\[
|\Theta(S_k(z^k + \alpha d^k)) - \Theta(z^k + \alpha d^k)| \leq 2L\alpha^{1+\gamma}\Theta(z^k) \tag{4.3}
\]

for all \( 0 \leq \alpha \leq 1 \) and all \( k \).

**Lemma 4.2.** Let Assumption 4.1 be satisfied. Then there exists \( \alpha > 0 \) such that

\[
\Theta(S_k(z^k + \alpha d^k)) \leq \Theta(z^k) + \sigma\alpha\Theta'(z^k)(d^k) = \Theta(z^k)(1 - 2\sigma\alpha)
\]

for some \( \sigma \in (0, 1 - L) \). Moreover, it holds \( 0 < 1 - 2\sigma\alpha < 1 \) whenever \( \sigma \in (0, \min\{1/2, 1 - L\}) \).

**Proof.** Inequality (4.2) together with Assumption 4.1 and exploiting \( \alpha^{1+\gamma} \leq \alpha \) for \( 0 \leq \alpha \leq 1 \) implies

\[
\Theta(S_k(z^k + \alpha d^k)) \leq \Theta(z^k)(1 - 2(\sigma - L)\alpha).
\]

Define \( \hat{\sigma} := \sigma + L \in (0, 1) \), i.e. \( \sigma \in (-L, 1 - L) \). Then

\[
\Theta(S_k(z^k + \alpha d^k)) \leq \Theta(z^k)(1 - 2\sigma\alpha) = \Theta(z^k) + \sigma\alpha\Theta'(z^k)(d^k).
\]

Armijo’s rule requires \( 0 < \sigma < 1 \). Since \( 0 \leq L < 1 \) and together with \( \sigma \in (-L, 1 - L) \) this implies \( \sigma \in (0, 1 - L) \). Because \( 0 < \alpha \leq 1 \) it holds \( 0 < 1 - 2\sigma\alpha < 1 \) whenever \( \sigma \in (0, \min\{1/2, 1 - L\}) \). \( \blacksquare \)

Lemma 4.2 guarantees that the line-search in the following global version of the semi-smooth Newton method is well-defined unless \( z^k \) is a zero of \( F \).

**Algorithm 4.3 (Global Nonsmooth Newton’s Method).**

(0) Choose \( z^0 \in Z_\infty, \beta \in (0, 1), \sigma \in (0, \min\{1/2, 1 - L\}) \).

(1) If some stopping criterion is satisfied, stop.

(2) Chose an arbitrary \( V_k \in \partial F(z^k) \) and compute the search direction \( d^k \) from

\[
V_k(d^k) = -F(z^k).
\]

(3) Find smallest \( i_k \in N_0 \) with

\[
\Theta(S_k(z^k + \beta^i d^k)) \leq \Theta(z^k) + \sigma\beta^i \Theta'(z^k)(d^k)
\]

and set \( \alpha_k = \beta^i \).

(4) Set \( z^{k+1} = S_k(z^k + \alpha_k d^k), k = k + 1, \) and goto (1).

The upcoming global convergence proof extends the proof presented in Jiang [19] for finite dimensions into infinite dimensions.

**Theorem 4.4.** Let the inverse operators \( V_k^{-1} \) exist for all \( k \) and let \( C > 0 \) be a constant such that \( ||V_k^{-1}||_{C(Y_\infty, Z_\infty)} \leq C \) holds for all \( k \). Let Assumption 4.1 hold. Let \( z_* \) be an accumulation point of the sequence \( \{z^k\} \) generated by the global nonsmooth Newton’s method.

Then, \( z_* \) is a zero of \( F \).

**Proof.** Let \( \{z^k\}_{k\in\mathbb{N}} \) be a subsequence with \( z^k \to z_* \) and \( F(z^k) \neq 0 \). Then,

\[
\Theta'(z^k)(d^k) = -2\Theta(z^k) = -||F(z^k)||^2_{z^k} < 0.
\]

The line-search is well-defined by the differentiability of \( \Theta \).
(i) Case 1: Assume

$$\alpha := \liminf_{j \to \infty} \alpha_{k,j} > 0.$$  

Then

$$0 \leq \Theta(z^{k,j+1}) \leq \Theta(z^{k,j}) + \sigma \alpha_{k,j} \Theta'(z^{k,j})(d^{k,j}) = \Theta(z^{k,j})(1 - 2\sigma \alpha_{k,j}).$$  

With $\sigma \in (0, \min\{1/2, 1 - L\})$ and $\alpha \leq \alpha_{k,j} \leq 1$ it follows $0 < 1 - 2\sigma \alpha_{k,j} 

$$0 \leq \Theta(z^{k,j}) \leq \Theta(z^{k_0})(1 - 2\sigma \alpha)^j \to 0.$$  

By the continuity of $F$, $z_*$ is a zero of $F$.

(ii) Case 2: Assume that there is a subsequence $\{z^k\}_{k \in J}$, $J \subseteq \{k_j \mid j \in \mathbb{N}\}$, with $\alpha_{k_j} \to 0$, $k_j \in J$.

The sequence $\{d^k\}$ is bounded since $\{V_k^{-1}\}$ is bounded and

$$0 \leq \|d^k\|_{Z_\infty} = \|V_k^{-1}(F(z^k))\|_{Z_\infty} \leq C\|F(z^k)\|_{V_\infty} \leq C\|F(z^0)\|_{V_\infty}.$$  

Unfortunately, the boundedness of $\{d^k\}$ in an infinite dimensional space does not imply that there exists a convergent subsequence. However, since $d^k$ is bounded in $Z_\infty$ it is also bounded in the space $Z_2$, which is a Hilbert space and thus reflexive. According to Theorem III.3.7 in Werner [34] there exists a weakly convergent subsequence $\{d^k\}$, $k \in J \subseteq J$. Hence, there exists some $d_* \in Z_2$ such that for every element $g \in Z_2^*$ it holds

$$g(d^k) \to g(d_*). \tag{4.4}$$

Herein, $Z_2^*$ denotes the topological dual space of $Z_2$. The derivative $\Theta'(z_*)(\cdot)$ is an element of $Z_2^*$ and an investigation turns out that it is essentially made up of linear functionals of type

$$g_1(z) = \int_0^1 h_1(z_*(t))z(t)dt, \quad g_2(z) = \int_0^1 h_2(z_*(t))z'(t)dt$$

with essentially bounded functions $h_1(z_*(\cdot))$ and $h_2(z_*(\cdot))$. Thus, by application of the Cauchy-Schwartz inequality, the functionals $g_1$ and $g_2$ are also linear continuous functionals on $Z_2$ and thus $g_1, g_2$, and in particular $\Theta'(z_*)(\cdot)$ can be viewed as elements of $Z_2^*$. Hence, (4.4) holds for $g(\cdot) = \Theta'(z_*)(\cdot)$:

$$\Theta'(z_*)(d^k) \to \Theta'(z_*)(d_*).$$

Furthermore, due to the continuity of $\Theta'(\cdot)$ (in $Z_\infty$) for every $\varepsilon > 0$ there exists $\delta > 0$ such that for every $\|z_k - z_*\|_{Z_\infty} \leq \delta$ it holds

$$|\Theta'(z^k)(d^k) - \Theta'(z_*)(d_k)| = \|d^k\|_{Z_\infty} \left| \Theta'(z^k) \left( \frac{d^k}{\|d^k\|_{Z_\infty}} \right) - \Theta'(z_*) \left( \frac{d_k}{\|d_k\|_{Z_\infty}} \right) \right|$$

$$\leq \|d^k\|_{Z_\infty} \cdot \sup_{\|d\|_{Z_\infty} = 1} |\Theta'(z^k)(d) - \Theta'(z_*)(d)|$$

$$= \|d^k\|_{Z_\infty} \cdot \|\Theta'(z^k) - \Theta'(z_*)\|_{\mathcal{L}(Z_\infty, \mathbb{R})} \leq \varepsilon \|d^k\|_{Z_\infty}.$$
For arbitrary \( \varepsilon > 0 \) we find
\[
|\Theta'(z_k^k)(d^k) - \Theta'(z_*)| \leq |\Theta'(z_k^k)(d^k) - \Theta'(z_*)| + |\Theta'(z_*)(d^k)|
\]
\[
\leq \varepsilon \|d^k\|_{L^\infty} + |\Theta'(z_*)(d^k) - \Theta'(z_*)(d^k)|.
\]
Since \( \varepsilon > 0 \) was arbitrary and since \( d^k \) is weakly convergent it holds
\[
\Theta'(z_k^k)(d^k) \to \Theta'(z_*)\text{ as } k \to \infty, \quad k \in \hat{J}.
\]
In a similar way the Fréchet differentiability of \( \Theta \) yields
\[
\frac{1}{\alpha_k} \left( \frac{\Theta(S_k(z^k + \alpha_k d^k)) - \Theta(z^k)}{\alpha_k} \right) - \Theta'(z_*)|d_*| \leq \frac{1}{\alpha_k} \left( \frac{\Theta(S_k(z^k + \alpha_k d^k)) - \Theta(z^k)}{\alpha_k} \right) - \Theta'(z_*)|d_*|
\]
\[
\leq 2L \alpha_k \Theta(z^k) + \frac{1}{\alpha_k} \phi(\|\alpha_k d^k\|_{L^\infty}) + |\Theta'(z^k)(d^k) - \Theta'(z_*)(d^k)|
\]
\[
\leq 2L \alpha_k \Theta(z^k) + \|d^k\|_{L^\infty} \frac{\phi(\|\alpha_k d^k\|_{L^\infty})}{\alpha_k} + \varepsilon \|d^k\|_{L^\infty} + |\Theta'(z^k)(d^k) - \Theta'(z_*)(d^k)|.
\]
Since \( d^k \) is weakly convergent and \( \alpha_k^k \to 0 \) it holds
\[
\frac{1}{\alpha_k} \left( \frac{\Theta(S_k(z^k + \alpha_k d^k)) - \Theta(z^k)}{\alpha_k} \right) - \Theta'(z_*)|d_*| \to \Theta'(z_*)|d_*|\text{ as } k \to \infty, \quad k \in \hat{J}.
\]
The line search in step (3) of the algorithm yields
\[
\frac{\Theta(S_k(z^k + \alpha_k d^k)) - \Theta(z^k)}{\alpha_k} \leq \sigma \Theta'(z^k)(d^k),
\]
\[
\frac{\Theta(S_k(z^k + \alpha_k d^k)) - \Theta(z^k)}{\alpha_k} \geq \sigma \Theta'(z^k)(d^k).
\]
Passing to the limit and exploiting the previous considerations yields
\[
\sigma \Theta'(z_*)|d_*| = \Theta'(z_*)|d_*|.
\]
Since \( \sigma \in (0, \min\{1/2, 1 - L\}) \) this only holds for \( \Theta'(z_*)|d_*| = 0 \). Thus, we have shown
\[
-\|F(z^k)\|_{L^2}^2 = \Theta'(z^k)(d^k) \to \Theta'(z_*)|d_*| = 0.
\]
By the continuity of \( F, z_* \) is a zero of \( F \).
The previous result only shows that each accumulation point is a zero of $F$. It would be nice to have also the fast local convergence properties of the local method. The locally superlinear convergence would follow from the local convergence theorem 2.3 if we were able to show that $\alpha_k = 1$ satisfies Armijo’s rule for all sufficiently large $k$.

The proof of the local convergence theorem 2.3 showed the superlinear convergence of the values $\|F(z^k)\|_Y$, i.e. for any $\varepsilon > 0$ there exists $\delta > 0$ such that for all $\|z^k - z_*\|_{Z_\infty} \leq \delta$ it holds

$$\|S_k(z^k + d^k) - z_*\|_{Z_\infty} \leq \varepsilon \|z^k - z_*\|_{Z_\infty},$$

where $d^k = -V^{-1}F(z^k)$, $V \in \partial F(z^k)$. In particular, there exists $\delta > 0$ such that for all $\|z^k - z_*\|_{Z_\infty} \leq \delta$ it holds

$$\|S_k(z^k + d^k) - z_*\|_{Z_\infty} \leq \frac{1}{2} \|z^k - z_*\|_{Z_\infty},$$

where

$$\|F(S_k(z^k + d^k))\|_Y \leq \sqrt{1 - 2\sigma\|F(z^k)\|_Y}.$$ 

With $r = 2$ this implies

$$\Theta(S_k(z^k + d^k)) = \frac{1}{2}\|F(S_k(z^k + d^k))\|^2_{Y_2} \leq \frac{1 - 2\sigma}{2}\|F(z^k)\|^2_{Y_2} = (1 - 2\sigma)\Theta(z^k)$$

resp.

$$\Theta(S_k(z^k + d^k)) \leq \Theta(z^k) - 2\sigma\Theta(z^k) = \Theta(z^k) + \sigma\Theta'(z^k)(d^k),$$

i.e. Armijo’s line-search accepts $\alpha_k = 1$ and $z^{k+1} = S_k(z^k + d^k)$. Furthermore, $\|z^{k+1} - z_*\|_{Z_\infty} \leq \frac{1}{2}\|z^k - z_*\|_{Z_\infty} \leq \delta$ and we are in the same situation as above and the argument could be repeated.

We summarize:

**Theorem 4.5.** Let the assumptions of Theorem 4.4 and Theorem 2.5 be valid with $r = 2$.

Then, for sufficiently large $k$ the step length $\alpha_k = 1$ is accepted and the global method turns into the local one.

**5. Numerical Results.** All computations were performed on a PC with 3 GHz processing speed. We used $\|F(z^k)\|^2 \leq 10^{-15}$ as a stopping criterion in the nonsmooth Newton’s method and did not use a smoothing step.

**5.1. Rayleigh Problem, Version 1.** We illustrate the method for the Rayleigh problem, cf. Maurer and Augustin [27], p. 39: Minimize

$$\int_0^{4.5} u(t)^2 + x_1(t)^2 \, dt$$

subject to

\begin{align*}
x_1' &= x_2, \\
x_2' &= -x_1 + x_2(1.4 - 0.14x_2^2) + 4u, \\
x_1(0) &= -5, \\
x_2(0) &= -5 
\end{align*}

and

$$u + \frac{1}{6}x_1 \leq 0.$$
With \( x = (x_1, x_2)^\top, \lambda = (\lambda_1, \lambda_2)^\top, \sigma = (\sigma_1, \sigma_2)^\top \) the Hamilton function reads as
\[
H(x, u, \lambda, \eta) = u^2 + x_1^2 + \lambda_1 x_2 + \lambda_2 \left( -x_1 + x_2 \left( 1.4 - 0.14x_2^2 \right) + 4u \right) + \eta \left( u + \frac{1}{6} x_1 \right).
\]

With \( z = (x, u, \lambda, \eta, \sigma) \) the function \( F \) in (2.9) is given by
\[
F(z) = \begin{pmatrix}
x_1' - x_2 \\
x_2' - \left( -x_1 + x_2 \left( 1.4 - 0.14x_2^2 \right) + 4u \right) \\
\lambda_1' + 2x_1 - \lambda_2 + \frac{1}{2}\eta \\
\lambda_2 + \lambda_1 + \lambda_2 \left( 1.4 - 0.42x_2^2 \right) \\
x_1(0) + 5 \\
x_2(0) + 5 \\
\lambda_1(0) + \sigma_1 \\
\lambda_2(0) + \sigma_2 \\
\lambda_1(4.5) \\
\lambda_2(4.5) \\
2u + 4\lambda_2 + \eta \\
\phi \left( -\left( u + \frac{1}{6} x_1 \right), \eta \right)
\end{pmatrix}
\]

In each iteration of the nonsmooth Newton’s method we have to solve the linear boundary value problem (3.2), (3.6) for \( x, \lambda, \sigma \). We leave the details of the boundary value problem (3.2),(3.6) and equation (3.5) to the reader. We note, that for all \((s+1)^2 + (r+1)^2 \leq 1\) it holds
\[
\det \mathcal{A} = \det \begin{pmatrix} 2 & 1 \\ -s & r \end{pmatrix} = 2r + s \neq 0
\]
and thus the operator \( \mathcal{A} \) in (3.4) is invertible. The differential equations are discretized on \([0, 4.5]\) using forward Euler’s method with \( N \) equidistant subintervals. The occurring derivatives \((x^k)'\) and \((\lambda^k)'\) are approximated by finite forward differences. Moreover, it turned out that it is advisable to scale the boundary conditions and the transversality conditions in the merit function by the step size \( h = 1/N \). The boundary value problem was solved by the single shooting method. Table 5.1 shows the output of the globalized nonsmooth Newton’s method, i.e. step size \( \alpha \), residual norm \( ||F||^2 \), and \( ||d^k|| \) during iteration. The iterations show the rapid quadratic convergence at the end of the iteration sequence. Recall, that only a locally superlinear convergence rate was established in Theorem 2.5.

| ITER | ALPHA | 1/2 | ||F||^2 | ||d^k|| |
|------|-------|-----|--------|--------|
| 0    | 0.000000E+00 | 0.245000E+04 | 0.173257E+04 |
| 1    | 0.531441E+00 | 0.173372E+04 | 0.316003E+04 |
| 2    | 0.717898E-01 | 0.170185E+04 | 0.897810E+03 |
| 3    | 0.185302E+00 | 0.155477E+04 | 0.653211E+03 |
| ...  |       |      |        |        |
| 10   | 0.100000E+01 | 0.147905E-05 | 0.263768E-06 |
| 11   | 0.100000E+01 | 0.152582E-25 | 0.598877E-12 |

Table 5.1

Output of globalized nonsmooth Newton’s method for the first version of Rayleigh’s problem for \( N = 100 \) subintervals and Euler discretization: local quadratic convergence.
The following table summarizes results for different step sizes. The number of iterations differs only by one, which indicates – at least numerically – the mesh independence of the method. Furthermore, the CPU time grows at a linear rate with $N$.

<table>
<thead>
<tr>
<th>$N$</th>
<th>CPU time [s]</th>
<th>Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0.027</td>
<td>13</td>
</tr>
<tr>
<td>500</td>
<td>0.136</td>
<td>14</td>
</tr>
<tr>
<td>1000</td>
<td>0.271</td>
<td>14</td>
</tr>
<tr>
<td>2000</td>
<td>0.505</td>
<td>14</td>
</tr>
<tr>
<td>4000</td>
<td>1.083</td>
<td>14</td>
</tr>
<tr>
<td>8000</td>
<td>2.065</td>
<td>14</td>
</tr>
</tbody>
</table>

Figure 5.1 illustrates the iterates of the nonsmooth Newton’s method. Notice the small inactive arc of the control-state constraint at the end of the time interval.

For comparison reasons the same optimal control problem was solved alternatively by a direct discretization method as in Gerdts [10] with Euler discretization and $N = 100$ subintervals. For this method the overall CPU time was 3.81 CPU seconds on the same processor. Furthermore, for the direct discretization method the CPU time grows nonlinearly with $N$. Hence, if all regularity assumptions are fulfilled the nonsmooth Newton’s method is an extremely efficient method.
5.2. Rayleigh Problem, Version 2. We consider a slight variation of the Rayleigh problem where boundary conditions are added and the control-state constraint is replaced by box constraints for the control, cf. Maurer and Augustin [27], p. 39: Minimize \((5.1)\) subject to \((5.2)\) and \(x_1(4.5) = 0, x_2(4.5) = 0\) and

\[-1 \leq u \leq 1.\]
With $x = (x_1, x_2)^\top$, $\lambda = (\lambda_1, \lambda_2)^\top$, $\sigma = (\sigma_1, \ldots, \sigma_4)^\top$, $\eta = (\eta_1, \eta_2)^\top$ the Hamilton function reads as

$$H(x, u, \lambda, \eta) = u^2 + x_1^2 + \lambda_1 x_2 + \lambda_2 (-x_1 + x_2 (1.4 - 0.14 x_2^2) + 4u) + \eta_1 (u - 1) + \eta_2 (-u - 1).$$

With $z = (x, u, \lambda, \eta, \sigma)$ the function $F$ in (2.9) is given by

$$F(z) = \begin{pmatrix}
  x_1' - x_2 \\
  x_2' - (-x_1 + x_2 (1.4 - 0.14 x_2^2) + 4u) \\
  \lambda_1' + 2 x_1 - \lambda_2 \\
  \lambda_2' + \lambda_1 + \lambda_2 (1.4 - 0.42 x_2^2) \\
  x_1(0) + 5 \\
  x_2(0) + 5 \\
  x_1(4.5) \\
  x_2(4.5) \\
  \lambda_1(0) + \sigma_1 \\
  \lambda_2(0) + \sigma_2 \\
  \lambda_1(4.5) - \sigma_3 \\
  \lambda_2(4.5) - \sigma_4 \\
  2u + 4 \lambda_2 + \eta_1 - \eta_2 \\
  \varphi (- (u - 1), \eta_1) \\
  \varphi (- (-u - 1), \eta_2)
\end{pmatrix}.$$ 

Again, we leave the details of the linear boundary value problem (3.2), (3.6) and equation (3.5) to the reader. An investigation of the generalized differential of $\varphi$ yields

$$\det A = \det \begin{pmatrix}
  2 & 1 & -1 \\
  -s_1 & r_1 & 0 \\
  s_2 & 0 & r_2
\end{pmatrix} = 2r_1 r_2 + r_1 s_2 + r_2 s_1 \neq 0$$

for any $(s_1, r_1) \in \partial \varphi(-(u - 1), \eta_1)$ and $(s_2, r_2) \in \partial \varphi(-(u - 1), \eta_2)$. Figure 5.2 illustrates the iterates of the nonsmooth Newton’s method for $N = 100$. Table 5.2 shows more detailed information about the iterations, i.e. step size $\alpha$, residual norm $\|F\|^2$, and $\|d^k\|$. Again, the boundary conditions and the transversality conditions in the merit function were scaled by the step size $h = 1/N$. The iterations show the rapid quadratic convergence at the end of the iteration sequence. Recall, that only a locally superlinear convergence rate was established in Theorem 2.5.
Figure 5.2. Numerical solution of the second version of Rayleigh’s problem for \( N = 100 \) Euler steps: Intermediate iterates (thin lines) and converged solution (thick lines).
Output of globalized nonsmooth Newton’s method for the second version of Rayleigh’s problem for \( N = 100 \) subintervals and Euler discretization: local quadratic convergence.

The number of iterations remains nearly constant, which indicates—a least numerically—the mesh independence of the method. Furthermore, the CPU time grows at a linear rate with \( N \).

<table>
<thead>
<tr>
<th>( N )</th>
<th>CPU time [s]</th>
<th>Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0.049</td>
<td>17</td>
</tr>
<tr>
<td>500</td>
<td>0.204</td>
<td>15</td>
</tr>
<tr>
<td>1000</td>
<td>0.502</td>
<td>18</td>
</tr>
<tr>
<td>2000</td>
<td>0.848</td>
<td>16</td>
</tr>
<tr>
<td>4000</td>
<td>1.785</td>
<td>17</td>
</tr>
<tr>
<td>8000</td>
<td>3.713</td>
<td>17</td>
</tr>
</tbody>
</table>

Again, the same optimal control problem was solved alternatively by a direct discretization method as in Gerdts [10] with Euler discretization and \( N = 100 \) subintervals. Herein, for better comparability the control constraints \(-1 \leq u \leq 1\) are not viewed as simple box constraints but are treated algorithmically as two ‘nonlinear’ mixed control-state constraints. For the direct method the overall CPU time was 2.41 CPU seconds. As mentioned before, the CPU time grows at a nonlinear rate with \( N \). Again, if all regularity assumptions are fulfilled the nonsmooth Newton’s method turns out to be extremely efficient.

Acknowledgments. The authors are supported by DFG grant GE 1163/5-1.

REFERENCES


