CONVERGENCE ANALYSIS OF EULER DISCRETIZATION OF
CONTROL-STATE CONSTRAINED OPTIMAL CONTROL
PROBLEMS WITH CONTROLS OF BOUNDED VARIATION

MATTHIAS GERDTS
Universität der Bundeswehr München
Institut für Mathematik und Rechneranwendung
Werner-Heisenberg-Weg 39, 85577 Neubiberg, Germany

MARTIN KUNKEL
Elektrobit Automotive GmbH
Am Wolsmantel 46, 91058 Erlangen, Germany

Abstract. We study convergence properties of Euler discretization of optimal control problems with ordinary differential equations and mixed control-state constraints. Under suitable consistency and stability assumptions a convergence rate of order $1/p$ of the discretized control to the continuous control is established in the $L^p$-norm. Throughout it is assumed that the optimal control is of bounded variation. The convergence proof exploits the reformulation of first order necessary optimality conditions as nonsmooth equations.

1. Introduction. Direct discretization methods are widely used to solve challenging optimal control problems numerically and many different implementations are available. The theoretical convergence properties of direct discretization methods, however, are not yet fully understood, e.g. for problems with discontinuous or singular controls, and only few papers address this issue. Please refer to [23] for a survey on discretization of optimal control problems.

Malanowski et al. [22] show the convergence of an Euler discretization for optimal control problems subject to mixed control-state constraints. The convergence rates for state, control, adjoints and multipliers turn out to be linear with respect to the supremum norm under the assumption (amongst others) that the optimal control is continuous. Likewise, assuming – amongst others – that the optimal control is Lipschitz continuous, Dontchev et al. [9] investigate an Euler discretization of a state and mixed control-state constrained optimal control problem and establish a linear convergence rate for the state in the $W^{1,2}$-norm and for the control in the $L_2$-norm. These papers rewrite necessary conditions as generalized equations and exploit stability results based on Robinson’s implicit function theorem for generalized equations to derive the convergence results. Higher order Runge-Kutta discretizations are considered in Hager [15] and in Dontchev et al. [10], where the discussion is restricted to second order Runge-Kutta methods. It is shown that

2010 Mathematics Subject Classification. Primary: 49J15, 49J52, 49M25; Secondary: 49K40.

Key words and phrases. Optimal control, discretization, convergence, control-state constraints, bounded variation, nonsmooth equation.

The reviewing process of the paper was handled by Ryan Loxton and Qun Lin as Guest Editors.
a higher order convergence rate is possible, if the optimal control is sufficiently smooth.

Discontinuous controls are permitted in Alt et al. [3, 4] and Veliov [28], wherein the discussion is restricted to linear and linear-quadratic optimal control problems, respectively. These papers assume the optimal control to be of bang-bang type and exclude singular subarcs.

Mesh independence results for Newton’s method and the Lagrange-Newton method can be found in Alt [1, 2]. Herein, consistent and stable discretizations for generalized equations play a central role and, as we shall see, are related to our approach.

Recently, Loxton et al. [25] investigate convergence properties for the control parametrization enhancing technique subject to control-state constraints and show convergence of the objective function values. Herein, the objective contains a term measuring the total variation of the control.

The purpose of this paper is to obtain convergence results for discretized optimal control problems subject to mixed control-state constraints without assuming continuity of the optimal control. Instead we will merely assume the optimal control to be of bounded variation, thus discontinuities in the control are permitted. The proof technique follows the general discretization theory for operator equations of Stetter [26], extends it towards nonsmooth equations, and aims at showing that consistency and stability implies convergence. Instead of considering generalized equations we will make use of nonsmooth reformulations of first order necessary optimality conditions. The assumptions in the convergence proof, however, exclude problems with pure state constraints and are not automatically satisfied for problems with the control appearing linearly.

Note that this paper is an abridged version of Chapter 5 in Kunkel [18].

In this article we discuss the following optimal control problem which we call the continuous optimal control problem in the sequel:

**Problem 1 (ODE-OCP).** Let \( J : \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \to \mathbb{R}, f : T \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x}, \psi : \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_y}, c : T \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_c} \) be Lipschitz-continuous w.r.t. \( t \) and twice continuously differentiable w.r.t. all other arguments. Find a state variable \( x \in W^{1,\infty}(T, \mathbb{R}^{n_x}) \), and a control variable \( u \in L^\infty(T, \mathbb{R}^{n_u}) \) such that the objective function

\[
J(x(t_0), x(t_f))
\]

is minimized subject to the ODE

\[
x'(t) = f(t, x(t), u(t)), \quad a.e. \text{ in } T,
\]

the boundary conditions

\[
\psi(x(t_0), x(t_f)) = 0,
\]

and the mixed control-state constraints

\[
c(t, x(t), u(t)) \leq 0, \quad a.e. \text{ in } T.
\]

Without loss of generality the discussion is restricted to problems on the fixed and bounded time interval \( T := [t_0, t_f] \subset \mathbb{R} \). As usual, the Banach space \( L^\infty(T, \mathbb{R}^{n}) \) consists of all essentially bounded functions \( v : T \to \mathbb{R}^{n} \) and the Banach spaces \( W^{1,p}(T, \mathbb{R}^{n}) \) with \( p \in \mathbb{N} \) or \( p = \infty \) consist of all absolutely continuous functions \( v : T \to \mathbb{R}^{n} \) with bounded derivative \( v' \) in the \( L^p \)-norm. The numbers \( n_x, n_u \in \mathbb{N} \), and \( n_c, n_\psi \in \mathbb{N}_0 \) denote the numbers of states, controls, and constraints.
$BV([a, b], \mathbb{R})$ denotes the space of all functions $v : [a, b] \to \mathbb{R}$ of bounded variation. The norm $\| \cdot \|_{BV}$ is defined by

$$
\|v\|_{BV} := |v(a)| + \text{Var}^b_a(v) = |v(a)| + \sup_{\text{all grids}}\sum_{a=t_0}^{b} |v(t_{k+1}) - v(t_k)|,
$$

and with this norm, the space $BV([a, b], \mathbb{R})$ is a Banach space, see Heuser [16]. The space $NBV([a, b], \mathbb{R})$ is the space of all normalized functions of bounded variation, i.e. all functions $v \in BV([a, b], \mathbb{R})$ which are continuous from the right and satisfy $v(a) = 0$. Furthermore we define the space $BV^1([a, b], \mathbb{R})$ as the space of all $v \in W^{1,\infty}([a, b], \mathbb{R})$ with $v' \in BV([a, b], \mathbb{R})$. The space $NBV^1([a, b], \mathbb{R})$ is the space of all $v \in W^{1,\infty}([a, b], \mathbb{R})$ with $v' \in NBV([a, b], \mathbb{R})$.

Our aim is to show convergence of the Euler discretization of Problem 1 for controls of bounded variation. The main tool in our discussion is the reformulation of necessary optimality conditions as system of equations. The optimality conditions consists of equations and complementarity conditions. The complementarity conditions will be reformulated as equations by applying an NCP (nonlinear complementarity problem) function $\varphi : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m$ with the property

$$
\varphi(a, b) = 0 \iff a \geq 0, \; b \geq 0, \; a^\top b = 0,
$$

for $a, b \in \mathbb{R}^m$. One popular choice for $\varphi$ in the special case $m = 1$ is the convex and Lipschitz continuous Fischer-Burmeister function $\varphi(a, b) := \sqrt{a^2 + b^2} - a - b$, compare Geiger and Kanzow [12]. Componentwise application of the Fischer-Burmeister function yields a vector-valued NCP function as in (1). Amongst others, Sun and Qi [27], Chen et al. [7], and Geiger and Kanzow [12] define and investigate alternative NCP functions.

The Fischer-Burmeister function $\varphi$ is not differentiable at $(a, b) = (0, 0)$. As a substitute for the derivative of $\varphi$ we use Clarke’s generalized Jacobian, compare [8]:

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

This paper is organized as follows. In Section 2 we summarize necessary conditions for local minimizers of Problem 1. These conditions are rewritten as an equivalent nonsmooth system of equations. In Section 3 we discretize Problem 1 using Euler’s method for the differential equation and obtain a finite dimensional optimization problem. The first order necessary conditions for minimizers of this finite problem are transformed into an equivalent nonsmooth system of equations using the same techniques as in Section 2. The convergence of the discretization method for controls of bounded variation is established in Theorem 4.6 in Section 4, which, together with Theorem 4.7, contains the main contribution of the paper. Finally a numerical illustration is presented in Section 5.

2. The continuous problem. The analysis of the discretization method which we define in Section 3 is based on necessary optimality conditions for the continuous and discretized optimal control problems. In this section we discuss the continuous problem.

Let $(x_+, u_+)$ be a weak local minimizer of Problem 1. The Jacobian matrices of $f$ and $c$ with respect to $x$ and $u$ are denoted by $f'_x$, $f'_u$, $c'_x$, and $c'_u$, respectively. Throughout, we use the notion $f'_x(t)$ as an abbreviation for $f'_x(t, x(t), u(t))$ and similarly for functions with the same structure, e.g. the Hamilton function.
defined below. We define the time dependent index set $I_\alpha(t)$ by $I_\alpha(t) := \{i \in \{1, \ldots, n_\alpha\} \mid c_i[t] \geq -\alpha\}$, the functions $c_i,\alpha(t)$ by $c_i,\alpha(t) := \min\{c_i[t] + \alpha, 0\}$, and the diagonal matrix $S_\alpha(t)$ by $S_\alpha(t) := \text{diag}(c_i,\alpha(t), i = 1, \ldots, n_\alpha)$. Let $c_{I_\alpha(t),\alpha}[t] := (c_{i,\alpha})_{i \in I_\alpha(t)}$ denote the submatrix of the Jacobian matrix $c_{\alpha}$ that consists of the rows $i \in I_\alpha(t)$.

Moreover, let the following two regularity conditions hold:

(A1) Linear independence: There exist $\alpha > 0$ and $\beta > 0$ such that

$$\|c_{I_\alpha(t),\alpha}[t]^{\top}\| \geq \beta\|\zeta\|$$

almost everywhere for all $\zeta$ of appropriate dimension.

(A2) Controllability: For every $q \in \mathbb{R}^{n_\psi}$ there exists $(y, v, w) \in W^{1,\infty}(T, \mathbb{R}^{n_\alpha}) \times L^\infty(T, \mathbb{R}^{n_\psi}) \times L^\infty(T, \mathbb{R}^{n_\nu})$, a solution of the linear system

$$y'(t) - f'_x[t]y(t) - f'_u[t]v(t) = 0,$$

$$\psi'_x(t_0) + \psi'_x(t_f) = q,$$

$$c'_x[t]y(t) + c'_u[t]v(t) + S_\alpha(t)w(t) = 0.$$

Under assumptions (A1) and (A2), Malanowski [21] shows the regularity of the Lagrange multipliers. For the special case of box constrained controls, Banishashemi and Kaya [5, Prop. 2.1] provide a linear independence condition, which is easier to verify than the one in (A1).

We normalize the multiplier $l_0$ associated with the objective function to $l_0 = 1$ as usual. Then there exist multipliers $\lambda_\ast \in W^{1,\infty}(T, \mathbb{R}^{n_\psi})$, $\eta_\ast \in L^\infty(T, \mathbb{R}^{n_\nu})$, and $\sigma_\ast \in \mathbb{R}^{n_\psi}$ such that the following conditions are satisfied, compare [21]:

$$x'_\ast(t) - f(t, x_\ast(t), u_\ast(t)) = 0, \quad (2)$$

$$\lambda'_\ast(t) + H'_x(t, x_\ast(t), u_\ast(t), \lambda_\ast(t), \eta_\ast(t))^{\top} = 0, \quad (3)$$

$$\psi(x_\ast(t_0), x_\ast(t_f)) = 0, \quad (4)$$

$$\lambda_\ast(t_0) + \gamma'_x(x_\ast(t_0), x_\ast(t_f), \sigma_\ast)^{\top} = 0, \quad (5)$$

$$\lambda_\ast(t_f) - \gamma'_x(x_\ast(t_0), x_\ast(t_f), \sigma_\ast)^{\top} = 0, \quad (6)$$

$$H'_u(t, x_\ast(t), u_\ast(t), \lambda_\ast(t), \eta_\ast(t))^{\top} = 0. \quad (7)$$

with the Hamilton function $H : \mathbb{R} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\psi} \times \mathbb{R}^{n_\nu} \to \mathbb{R}$,

$$H(t, x, u, \lambda, \eta) = \lambda^{\top} f(t, x, u) + \eta^{\top} c(t, x, u).$$

and $\gamma : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\psi} \to \mathbb{R}$,

$$\gamma(x_0, x_f, \sigma) := J(x_0, x_f) + \sigma^{\top} \psi(x_0, x_f).$$

Furthermore, the complementarity conditions hold a.e. in $T$:

$$\eta_\ast(t) \geq 0, \quad c(t, x_\ast(t), u_\ast(t)) \leq 0, \quad \eta_\ast(t)^{\top} c(t, x_\ast(t), u_\ast(t)) = 0. \quad (8)$$

Using the Fischer-Burmeister function defined in Section 1 and due to (1), the necessary conditions (2)–(8) can be equivalently written as a nonlinear equation

$$F(z) = 0, \quad (9)$$

with $F : Z_\infty \to Y_\infty$,

$$Z_\infty = W^{1,\infty}(T, \mathbb{R}^{n_\psi}) \times L^\infty(T, \mathbb{R}^{n_\nu}) \times W^{1,\infty}(T, \mathbb{R}^{n_\psi}) \times L^\infty(T, \mathbb{R}^{n_\nu}) \times \mathbb{R}^{n_\psi},$$

$$Y_\infty = L^\infty(T, \mathbb{R}^{n_\psi}) \times L^\infty(T, \mathbb{R}^{n_\nu}) \times \mathbb{R}^{n_\psi} \times \mathbb{R}^{n_x} \times L^\infty(T, \mathbb{R}^{n_\nu}) \times L^\infty(T, \mathbb{R}^{n_\nu}),$$

$$F(z) = 0.$$
and

\[
F(z) := \begin{pmatrix}
x' - f(\cdot, x(t), u(t)) \\
\lambda' + H_u'(:, x(t), u(t), \lambda(\cdot), \eta(\cdot)) \\
\psi(x(t_0), x(t_f)) \\
\lambda(t_0) + \gamma'_{x_0}(x(t_0), x(t_f), \sigma) \\
\lambda(t_f) - \gamma'_{x_f}(x(t_0), x(t_f), \sigma) \\
H'_{u}(\cdot, x(\cdot), u(\cdot), \lambda(\cdot), \eta(\cdot)) \\
\varphi(-c(\cdot, x(\cdot), u(\cdot)), \eta(\cdot))
\end{pmatrix}.
\] (10)

3. The discretized problem. In this section we introduce the Euler discretization of Problem 1. Let \( \{G_N\}_{N \in \mathbb{N}}, G_N := \{t_0 < t_1 < \ldots < t_N = t_f\} \), be a sequence of time meshes on the interval \( T = [t_0, t_f] \) with maximal step width

\[
h := \max_{i \in \{0, \ldots, N-1\}} h_i = \max_{i \in \{0, \ldots, N-1\}} (t_{i+1} - t_i),
\]

and let there exist constants \( c_1 > 0 \) and \( c_2 > 0 \) (independent of \( N \)) such that

\[
c_1 \frac{1}{N} \leq h_i \leq c_2 \frac{1}{N}, \quad \forall i \in \{0, \ldots, N-1\}.
\]

The space \( W^{1,p}_h(T, \mathbb{R}^n) \) is defined to be the space of continuous functions which are affine on each interval \([t_i, t_{i+1}]\) equipped with the norm \( \| \cdot \|_{1,p} \). The space \( L^p(T, \mathbb{R}^n) \) is defined to be the space of functions \( g \) which are constant and continuous from the right on each interval \([t_i, t_{i+1}]\) with \( g(t_N) := g(t_{N-1}) \) equipped with the norm \( \| \cdot \|_p \).

We discretize the control \( u \) by a piecewise constant function \( u_h \), the state \( x \) by a piecewise affine function \( x_h \), and focus on explicit Euler’s method for the differential equation, i.e.

\[
x'_h(t_i) := \frac{x_h(t_{i+1}) - x_h(t_i)}{h_i} = f(t_i, x_h(t_i), u_h(t_i)), \quad \forall i \in \{0, \ldots, N-1\}.
\]

Note that \( x'_h(t_i) \) is the derivative of \( x_h \) in the interval \((t_i, t_{i+1})\). The discretized problem reads as follows and we call it discrete problem throughout:

**Problem 2** (Discretized ODE-OCP). Find the discretized state variable \( x_h \in W^{1,\infty}_h(T, \mathbb{R}^n) \), and the discretized control variable \( u_h \in L^\infty(T, \mathbb{R}^n) \) such that the objective function

\[
J(x_h(t_0), x_h(t_N))
\]

is minimized subject to

\[
f(t_i, x_h(t_i), u_h(t_i)) - x'_h(t_i) = 0, \quad \forall i \in \{0, \ldots, N-1\},
\]

the boundary conditions

\[
\psi(x_h(t_0), x_h(t_N)) = 0,
\]

and the mixed control-state constraints

\[
c(t_i, x_h(t_i), u_h(t_i)) \leq 0, \quad \forall i \in \{0, \ldots, N-1\}.
\]

Problem 2 is a finite dimensional nonlinear program. Note that the continuous mixed control-state constraints \( c(t, x(t), u(t)) \leq 0 \) and the differential equations are satisfied at the grid points only. Again, we assume that some constraint qualification holds, so that we can normalize the Lagrange multiplier \( l_0 \) associated with the objective function to \( l_0 = 1 \).
Theorem 3.1. Let \((x_h, u_h)\) be a local minimum of Problem 2 and let some constraint qualification be satisfied, e.g. the Mangasarian-Fromovitz constraint qualification, compare \([12]\). Then there exist discrete Lagrange multipliers \(\lambda_h \in W_h^{1,\infty}(T, \mathbb{R}^n), \eta_h \in L_h^\pi(T, \mathbb{R}^n), \) and \(\sigma_h \in \mathbb{R}^{n_v}\) which satisfy:

\[
\begin{align*}
    \lambda'_h(t_i) + H'_x(t_i, x_h(t_i), u_h(t_i), \lambda_h(t_{i+1}), \eta_h(t_{i+1})) & = 0, \quad (11) \\
    H'_u(t_i, x_h(t_i), u_h(t_i), \lambda_h(t_{i+1}), \eta_h(t_{i+1})) & = 0, \quad (12) \\
    \eta_h(t_i) & \geq 0, \quad \eta_h(t_i) \top c(t_i, x_h(t_i), u_h(t_i)) = 0, \quad (13)
\end{align*}
\]

for all \(i \in \{0, \ldots, N - 1\}\) and

\[
\begin{align*}
    \lambda_h(t_0) + \gamma'_{x_0}(x_h(t_0), x_h(t_N), \sigma_h) & = 0, \quad (14) \\
    \lambda_h(t_N) - \gamma'_{x_0}(x_h(t_0), x_h(t_N), \sigma_h) & = 0. \quad (15)
\end{align*}
\]

Proof. Evaluate the KKT conditions for Problem 2, see Kunkel \([18]\). \(\square\)

Remark 1. Note that \(h_i \eta_h(t_i)\) is the Lagrange multiplier associated with the constraint

\[c(t_i, x_h(t_i), u_h(t_i)) \leq 0\]

in Problem 2 and \(h_i \lambda_h(t_{i+1})\) is the multiplier associated with equation

\[f(t_i, x_h(t_i), u_h(t_i)) - x'_h(t_i) = 0\]

in Problem 2.

Theorem 3.1 leads to the discrete counter-part of Equation (9). Find \(z_h \in Z_{h,p}\) with

\[F_h(z_h) = 0, \quad F_h : Z_{h,p} \to Y_{h,p},\]

where \(z_h = (x_h, u_h, \lambda_h, \eta_h, \sigma)\),

\[
\begin{align*}
    Z_{h,p} & = W_{h,p}^1(T, \mathbb{R}^{n_x}) \times L_h^p(T, \mathbb{R}^{n_u}) \times W_{h,p}^1(T, \mathbb{R}^{n_x}) \times L_h^p(T, \mathbb{R}^{n_u}) \times \mathbb{R}^{n_v}, \\
    Y_{h,p} & = L_h^p(T, \mathbb{R}^{n_x}) \times L_h^p(T, \mathbb{R}^{n_u}) \times \mathbb{R}^{n_v} \times \mathbb{R}^{n_x} \times L_h^p(T, \mathbb{R}^{n_u}) \times L_h^p(T, \mathbb{R}^{n_u})
\end{align*}
\]

and

\[
F_h(z_h)(t_i) = \begin{pmatrix}
    x'_h(t_i) - f(t_i, x_h(t_i), u_h(t_i)) \\
    \lambda'_h(t_i) + H'_x(t_i, x_h(t_i), u_h(t_i), \lambda_h(t_{i+1}), \eta_h(t_{i+1})) \\
    \lambda_h(t_0) + \gamma'_{x_0}(x_h(t_0), x_h(t_N), \sigma_h) \\
    \lambda_h(t_N) - \gamma'_{x_0}(x_h(t_0), x_h(t_N), \sigma_h) \\
    H'_u(t_i, x_h(t_i), u_h(t_i), \lambda_h(t_{i+1}), \eta_h(t_{i+1})) \\
    \varphi(-c(t_i, x_h(t_i), u_h(t_i)), \eta_h(t_i))
\end{pmatrix}
\]

for \(i = 0, \ldots, N - 1\).

By \(\partial F_h\) we denote Clarke’s generalized Jacobian of the locally Lipschitz continuous map \(F_h\), see Clarke \([8]\). According to Clarke \([8\), Thm 2.6.6, Prop. 2.6.2\] the generalized Jacobian \(\partial F_h\) is a subset of the larger set \(\partial \tilde{F}_h\), which is defined to be the point to set mapping \(\partial \tilde{F}_h : Z_{h,p} \rightrightarrows L(Z_{h,p}, Y_{h,p})\) given by

\[
\partial \tilde{F}_h(z_h)(z_h) := \begin{cases}
    G_h(z_h)(z_h), & \text{if } S = \text{diag}(s_1, \ldots, s_{n_v}), \\
    -S(c'_u[\cdot]x_h + c'_u[\cdot]u_h + R\eta_h), & \text{if } (s_i, r_i) \in \partial \varphi[\cdot] \text{ on } G_N
\end{cases}
\]

where \(S = \text{diag}(s_1, \ldots, s_{n_v}), \quad R = \text{diag}(r_1, \ldots, r_{n_v}), \quad (s_i, r_i) \in \partial \varphi[\cdot] \text{ on } G_N\).
As elements \((s_i, r_i) \in \partial \varphi[\cdot]\) we might choose for \(t_j \in \mathbb{G}_N\) and \(i \in \{1, \ldots, n_c\}\),

\[
s_i(t_j) = \begin{cases} 
-1, & \text{if } c_i[t_j] = 0, \eta_h, i(t_j) = 0, \\
\frac{-c_i[t_j]}{\sqrt{c_i[t_j]^2 + \eta_h, i(t_j)^2}} - 1, & \text{otherwise,}
\end{cases}
\]

\[
r_i(t_j) = \begin{cases} 
0, & \text{if } c_i[t_j] = 0, \eta_h, i(t_j) = 0, \\
\frac{\eta_h, i(t_j)}{\sqrt{c_i[t_j]^2 + \eta_h, i(t_j)^2}} - 1, & \text{otherwise,}
\end{cases}
\]

where the index \(i\) denotes the \(i\)-th component of the respective vectors.

4. **Convergence analysis for controls of bounded variation.** In this section we show convergence of sequences of solutions \(\{z_h\}\) of the discrete problem \(F_h(z_h) = 0\) against the solution \(z\) of the continuous problem \(F(z) = 0\). Note that we have to select the range and image space of \(F\) carefully to obtain convergence. Here, we only consider controls of bounded variation and measure the error in \(L^p\)-norms with \(1 \leq p < \infty\). We cannot expect convergence in \(L^\infty\) since \(u\) may possess jumps.

**Assumption 4.1.** Throughout this section we assume that the control \(u_*\) and the multiplier \(\eta_*\) of the mixed control-state constraints satisfy

\[
\begin{align*}
    u_* &\in BV(T, \mathbb{R}^{n_u}), & \eta_* &\in BV(T, \mathbb{R}^{n_e}).
\end{align*}
\]

The following lemma can be found in \cite[Thm 4]{17}:

**Lemma 4.2.** Let \(v : \mathbb{R}^n \to \mathbb{R}\) be a Lipschitz continuous function with Lipschitz constant \(L\) and let \(z : [a,b] \to \mathbb{R}^n\) be a function of bounded variation.

Then, the composition \(v \circ z\) is of bounded variation.

The absolutely continuous functions \(x_*\) and \(\lambda_*\) are functions of bounded variation. Under Assumption 4.1, Lemma 4.2 implies that the compositions \(f(\cdot, x_*(\cdot), u_*(\cdot))\) and \(c(\cdot, x_*(\cdot), u_*(\cdot))\) as well as the first and second derivatives \(f'_*[\cdot], f''_*[\cdot], c'_*[\cdot], \ldots\), etc., are functions of bounded variation. Furthermore, since the multiplication and summation of two functions of bounded variation is of bounded variation, see Natanson \cite[Chapter VIII]{24}, the Hamilton function and its first and second derivatives are also of bounded variation. This shows that

\[
\begin{align*}
    x'_*(\cdot) &\in BV(T, \mathbb{R}^{n_x})), \\
    \lambda'_*(\cdot) &\in H'_*([\cdot])^T BV(T, \mathbb{R}^{n_x}),
\end{align*}
\]

and thus, \(x_* \in BV^1(T, \mathbb{R}^{n_x})\) and \(\lambda_* \in BV^1(T, \mathbb{R}^{n_e})\).

We define the spaces

\[
\begin{align*}
    (\tilde{Z}_p, \|\cdot\|_{Z_p}) &= (BV^1(T, \mathbb{R}^{n_x}), \|\cdot\|_{1,p}) \times (BV(T, \mathbb{R}^{n_x}), \|\cdot\|_p) \\
    &\times (BV^1(T, \mathbb{R}^{n_e}), \|\cdot\|_{1,p}) \times (BV(T, \mathbb{R}^{n_e}), \|\cdot\|_p) \times (\mathbb{R}^{n_e}, \|\cdot\|),
\end{align*}
\]

\[
\begin{align*}
    (\tilde{Y}_p, \|\cdot\|_{Y_p}) &= (BV(T, \mathbb{R}^{n_x}), \|\cdot\|_p) \times (BV(T, \mathbb{R}^{n_x}), \|\cdot\|_p) \\
    &\times (\mathbb{R}^{n_e+2n_e}, \|\cdot\|_p) \times (BV(T, \mathbb{R}^{n_e}), \|\cdot\|_p) \times (BV(T, \mathbb{R}^{n_e}), \|\cdot\|_p),
\end{align*}
\]

and restriction operators \(\Delta^0_h : BV(T, \mathbb{R}) \to L^p_h(T, \mathbb{R})\) with

\[
\Delta^0_h(v)(t_i) = v(t_i), \quad \forall i \in \{0, 1, \ldots, N - 1\},
\]

and \(\Delta^1_h : BV^1(T, \mathbb{R}) \to W_h^{1,p}(T, \mathbb{R})\) with

\[
\Delta^1_h(v)(t_i) = v(t_i), \quad \forall i \in \{0, 1, \ldots, N\}.
\]
The discretization operators are combined in an obvious way, so that we obtain restriction operators for the spaces $Z_p$ and $Y_p$, i.e.

$$\Delta_h^Z : \tilde{Z}_p \to Z_{h,p}, \quad \Delta_h^Z (z) := (\Delta_h^1 (x), \Delta_h^0 (u), \Delta_h^1 (\lambda), \Delta_h^0 (\sigma))^\top,$$

$$\Delta_h^Y : \tilde{Y}_p \to Y_{h,p}, \quad \Delta_h^Y (y) := (\Delta_h^0 (y_u), \Delta_h^0 (y_\lambda), y_\psi, y_{\lambda_0}, y_{\lambda_f}, \Delta_h^0 (y_u), \Delta_h^0 (y_\psi))^\top,$$

for $z = (x, u, \lambda, \eta, \sigma)^\top \in \tilde{Z}_p$ and $y = (y_x, y_\lambda, y_\psi, y_{\lambda_0}, y_{\lambda_f}, y_u, y_\psi)^\top \in \tilde{Y}_p$. Along with an operator $\Delta_h : \tilde{Z}_p \to Y_{h,p}$ which maps $F$ to $F_h$ the defined discretization method is denoted by $\mathcal{M} := (Z_{h,p}, Y_{h,p}, \Delta_h^Z, \Delta_h^Y, \Delta_h)$.

Note that by $\tilde{Z}_p$ and $\tilde{Y}_p$ we denote the restriction of $Z_p$ and $Y_p$ to functions of bounded variation.

We exploit the following basic approximation properties for the discretization operators.

**Lemma 4.3.** (i) **Norm equivalence in $\mathbb{R}^N$:** There exists $c_3 > 0$, independent of $h$, such that for all $a \in \mathbb{R}^N$ it holds

$$\max_{i \in \{1, \ldots, N\}} |a_i| \leq \left( \sum_{i=1}^N |a_i|^p \right)^{1/p} \leq \sum_{i=1}^N |a_i| \leq \frac{c_3}{h^{1/p}} \left( \sum_{i=1}^N |a_i|^p \right)^{1/p}.$$

(ii) **For all $v \in W^{1,\infty}(T, \mathbb{R})$ and $i = 0, \ldots, N - 1$ it holds**

$$v(t_{i+1}) - v(t_i) = \int_{t_i}^{t_{i+1}} v'(t) \, dt.$$

(iii) **For all $v_h \in L_h^p(T, \mathbb{R})$, e.g. piecewise constant functions, it holds**

$$\|v_h\|_p \leq \sum_{i=0}^{N-1} h_i^{1/p} |v_h(t_i)|.$$  

(iv) **For all $v \in BV(T, \mathbb{R})$ it holds**

$$|v(t_i) - v(t)| \leq \text{Var}_{t_i}^{t_{i+1}} (v),$$

for $t \in [t_i, t_{i+1}]$ and

$$\|\Delta_h^0 (v) - v\|_p \leq h^{1/p} \|v\|_{BV}.$$  

(v) **For all $v \in BV^1(T, \mathbb{R})$ it holds**

$$\|\Delta_h^1 (v) - v\|_p \leq 2 h^{1/p} \|v\|_{BV}.$$  

(vi) **For all $v \in BV^1(T, \mathbb{R})$ it holds**

$$\|\Delta_h^1 (v) - v\|_p \leq 2 h^{1/p} \|v\|_{BV}.$$

Note that (vi) is not implied by (iv), since $\Delta_h^1 (v)'(t_i) \neq \Delta_h^0 (v')(t_i)$ in general.

**Proof.** See Appendix A. $\square$

**Corollary 1.** For $v \in BV(T, \mathbb{R})$ the restriction operator $\Delta_h^0$ satisfies

$$\lim_{h \to 0} \|\Delta_h^0 (v)\|_p = \|v\|_p.$$

For $v \in BV^1(T, \mathbb{R})$ the restriction operator $\Delta_h^1$ satisfies

$$\lim_{h \to 0} \|\Delta_h^1 (v)\|_{1,p} = \|v\|_{1,p}.$$
Theorem 4.4 (Consistency)

Proof. Lemma 4.3 (iv) implies
\[
\lim_{h \to 0} \|\Delta_h^0(v)\|_{L^p} - \|v\|_{L^p} \leq \lim_{h \to 0} \|\Delta_h^1(v) - v\|_{L^p} \leq \lim_{h \to 0} h^{1/p} \|v\|_{BV} = 0,
\]
and Lemma 4.3 (v) and (vi) imply
\[
\lim_{h \to 0} \|\Delta_h^1(v)\|_{L^1} - \|v\|_{L^1} \leq \lim_{h \to 0} \max\{\|\Delta_h^1(v) - v\|_{L^1}, \|\Delta_h^1(v) - v'\|_{L^1}\}
\]
\[
\leq \lim_{h \to 0} \max\{2h^{1/p} \|v\|_{BV}, 2h^{1/p} \|v'\|_{BV}\} = 0.
\]
\]

We prove consistency and stability of the discretization method:

Theorem 4.4 (Consistency). Let \(z_* \in \tilde{Z}_p\) be a locally optimal solution satisfying \(F(z_*) = 0\). Then the discretization method is consistent of order \(1/p\) at \(z_*\), i.e.
\[
\|F_h(z_*) - \Delta_h^p(F(z_*))\|_{\mathcal{Y}_{h,p}} = \mathcal{O}(h^{1/p}) \quad \text{as} \quad h \to 0.
\]

Proof. Component-wise application of Lemma 4.3 (iii) yields
\[
\|F_h(z_*) - \Delta_h^p(F(z_*))\|_{\mathcal{Y}_{h,p}} \leq \max \left\{ h^{1/p} \sum_{i=0}^{N-1} \left| \frac{\Delta^1_h(x_*(t_{i+1})) - \Delta^1_h(x_*(t_i))}{h_i} \right| + f(t_i, \Delta^1_h(x_*(t_i)), \Delta^0_h(u_*(t_i)) \right|, \right.
\]
\[
\left. h^{1/p} \sum_{i=0}^{N-1} \left| \frac{\Delta^1_h(\lambda_*(t_{i+1})) - \Delta^1_h(\lambda_*(t_i))}{h_i} \right| + H^\prime_x(t_i, \Delta^1_h(x_*(t_i)), \Delta^0_h(u_*(t_i)), \Delta^1_h(\lambda_*(t_{i+1})), \Delta^0_h(\eta_*(t_i))) \right|, \right.
\]
\[
\left. \left| \psi(\Delta^1_h(x_*(t_0)), \Delta^1_h(x_*(t_N))) \right|, \right.
\]
\[
\left. \left| \Delta^1_h(\lambda_*(t_0)) + \gamma^\lambda_{x}(\Delta^1_h(x_*(t_0)), \Delta^1_h(x_*(t_N)), \sigma_x) \right|, \right.
\]
\[
\left. \left| \Delta^1_h(\lambda_*(t_N)) - \gamma^\lambda_{x} (\Delta^1_h(x_*(t_0)), \Delta^1_h(x_*(t_N)), \sigma_x) \right|, \right.
\]
\[
\left. h^{1/p} \sum_{i=0}^{N-1} \left| \Delta^0_h(u_*(t_i)) \right|, \right.
\]
\[
\left. h^{1/p} \sum_{i=0}^{N-1} \left| \varphi(-c(t_i, \Delta^1_h(x_*(t_i)), \Delta^0_h(u_*(t_i)), \Delta^0_h(\eta_*(t_i)) \right), \right.
\]
\[
\left. \left| \psi(x_*(t_0)), x_*(t_N) \right|, \right.
\]
\[
\left. \left| \lambda_*(t_0) + \gamma_{z}(x_*(t_0), x_*(t_N), \sigma_x) \right|, \right. \]
\]

\(\square\)
\[ \left\| \lambda_*(t_N) - \gamma'_{xf}(x_*(t_0), x_*(t_N), \sigma_*)^T \right\|, \]
\[ h^{1/p} \sum_{i=0}^{N-1} \left\| H'_x(t_i, x_*(t_i), u_*(t_i), \lambda_*(t_{i+1}), \eta_*(t_i))^T \right\|, \]
\[ h^{1/p} \sum_{i=0}^{N-1} \left\| \varphi(-c(t_i, x_*(t_i), u_*(t_i)), \eta_*(t_i)) \right\|. \]

The expressions \( \left\| \psi(x_*(t_0), x_*(t_N)) \right\|, \left\| \lambda_*(t_0) + \gamma'_{xf}(x_*(t_0), x_*(t_N), \sigma_*)^T \right\|, \left\| \lambda_*(t_N) - \gamma'_{xf}(x_*(t_0), x_*(t_N), \sigma_*)^T \right\|, \left\| \varphi(-c(t_i, x_*(t_i), u_*(t_i)), \eta_*(t_i)) \right\| \) vanish, since \( F(z_* = 0) \).

We exploit Lemma 4.3 (ii) and obtain
\[ \sum_{i=0}^{N-1} \frac{x_*(t_{i+1}) - x_*(t_i)}{h_i} - f(t_i, x_*(t_i), u_*(t_i)) \leq \sum_{i=0}^{N-1} \frac{1}{h_i} \int_{t_i}^{t_{i+1}} \left\| x'_*(t) - f[t_i] \right\| dt \]
\[ = \sum_{i=0}^{N-1} \frac{1}{h_i} \int_{t_i}^{t_{i+1}} x'_*(t) - f[t_i] dt \leq \sum_{i=0}^{N-1} \frac{1}{h_i} \int_{t_i}^{t_{i+1}} \| f[t] - f[t_i] \| dt \]
\[ \hfill (26) \]
\[ \text{Appendix A} \]
\[ \sum_{i=0}^{N-1} \text{Var}_{t_i}^{t_{i+1}} (f[\cdot]) \leq \| f[\cdot] \|_{BV}. \]

Similarly it holds
\[ \sum_{i=0}^{N-1} \left\| \frac{\lambda_*(t_{i+1}) - \lambda_*(t_i)}{h_i} + H'_x(t_i, x_*(t_i), u_*(t_i), \lambda_*(t_{i+1}), \eta_*(t_i))^T \right\| \]
\[ \leq \sum_{i=0}^{N-1} \frac{1}{h_i} \int_{t_i}^{t_{i+1}} \left\| \lambda'_x(t) + H'_x(t_i, x_*(t_i), u_*(t_i), \lambda_*(t_{i+1}), \eta_*(t_i))^T \right\| dt \]
\[ = \sum_{i=0}^{N-1} \frac{1}{h_i} \int_{t_i}^{t_{i+1}} \left\| -H'_x[t_i]^T + H'_x(t_i, x_*(t_i), u_*(t_i), \lambda_*(t_{i+1}), \eta_*(t_i))^T \right\| dt \]
\[ \leq \sum_{i=0}^{N-1} \frac{1}{h_i} \int_{t_i}^{t_{i+1}} \left\| -c'_x[t_i]^T \eta(t) + c'_x[t_i]^T \eta_*(t_i) \right\| dt \]
\[ + \sum_{i=0}^{N-1} \frac{1}{h_i} \int_{t_i}^{t_{i+1}} \left\| f'_x[t_i]^T \lambda_*(t_{i+1}) - f'_x[t_i]^T \lambda_*(t) \right\| dt \]
\[ \leq \sum_{i=0}^{N-1} \text{Var}_{t_i}^{t_{i+1}} (\| c'_x[\cdot]^T \eta_*(\cdot) \|) \]
\[ + \sum_{i=0}^{N-1} \frac{1}{h_i} \int_{t_i}^{t_{i+1}} \left\| f'_x[t_i]^T \lambda_*(t_{i+1}) - f'_x[t_i]^T \lambda_*(t) \right\| \]
\[ + \left\| f'_x[t_i]^T \lambda_*(t) \right\| - f'_x[t_i]^T \lambda_*(t) \right\| dt \]
\[ \leq \| c'_x[\cdot]^T \eta_*(\cdot) \|_{BV} + \sum_{i=0}^{N-1} \left( \left\| f'_x[t_i]^T \lambda_*(t_{i+1}) - \lambda_*(t_i) \right\| + \text{Var}_{t_i}^{t_{i+1}} (f'_x[\cdot]^T \lambda_*(\cdot)) \right) \]
\[ \leq \| c'_x[\cdot]^T \eta_*(\cdot) \|_{BV} + \| f'_x[\cdot]^T \|_{\infty} \left\| \lambda_*(\cdot) \right\|_{BV} + \| f'_x[\cdot]^T \lambda_*(\cdot) \|_{BV}. \]
It remains to discuss the equation which arises from the optimality condition:

\[
\sum_{i=0}^{N-1} \left\| H_u'(t_i, x_*(t_i), u_*(t_i), \lambda_*(t_{i+1}), \eta_*(t_i))^T \right\| \\
= \sum_{i=0}^{N-1} \left\| H_u'(t_i, x_*(t_i), u_*(t_i), \lambda_*(t_{i+1}), \eta_*(t_i))^T - \frac{H_u'[t_i]}{\delta} \right\| \\
= \sum_{i=0}^{N-1} \left\| f_u'[t_i]^T (\lambda_*(t_{i+1}) - \lambda_*(t_i)) \right\| \\
\leq \| f_u'[\cdot]^T \|_{\infty} \| \lambda_*(\cdot) \|_{BV}.
\]

Finally we obtain

\[
\| F_h(\Delta_h^B(z_*)) - \Delta_h^V(F(z_*)) \|_{Y_{h,p}} \leq h^{1/p} \max \left\{ \| f[\cdot] \|_{BV}, \| c_z[1]^T \eta_*(\cdot) \|_{BV} + \| f_x'[\cdot]^T \|_{\infty} \| \lambda_*(\cdot) \|_{BV} + \| f_x'[\cdot]^T \lambda_*(\cdot) \|_{BV}, \| f_u'[\cdot]^T \|_{\infty} \| \lambda_*(\cdot) \|_{BV} \right\} \\
= O(h^{1/p}).
\]

\[\square\]

In the subsequent considerations let \( F_{h,j} \) denote the \( j \)-th functional component of \( F_h \) and let \( M := 4n_c + n_\psi + n_u + n_e \) denote the number of components of \( F_h \), e.g.

\[ F_{h,1} : Z_{h,p} \to L^p_h(T, \mathbb{R}), \quad F_{h,1}(z_h)(t_i) = x_{h,1}(t_i) - f_1(t_i, x_h(t_i), u_h(t_i)). \]

\textbf{Theorem 4.5 (Stability).} Let \( z_* \in \tilde{Z}_p \) be a locally optimal solution of Problem 1 with \( F(z_*) = 0 \). Let the inverse operators \( W_{h}^{-1} \) exist and let there exist constants \( C > 0 \) and \( \delta > 0 \) (independent of \( h \)) such that \( \| W_{h}^{-1} \|_{L(Y_{h,p}, Z_{h,p})} \leq C \) for every

\[
W_h = \begin{pmatrix}
\partial F_{h,1}(z_h^1) \\
\vdots \\
\partial F_{h,M}(z_h^M)
\end{pmatrix},
\]

every \( z_h^j \in Z_{h,p} \) with \( \| z_h^j - \Delta_h^B(z_*) \|_{Z_{h,p}} < \delta, j = 1, \ldots, M \), and every \( h > 0 \) sufficiently small.

Then the discretization method is stable at \( \Delta_h^B(z_*) \) and there exist solutions \( z_h \in Z_{h,p} \) of \( F_h(z_h) = 0 \) with \( \| z_h - \Delta_h^B(z_*) \|_{Z_{h,p}} < \delta \) for almost all \( h > 0 \) sufficiently small.

\textbf{Proof.} Let \( \epsilon > 0 \) be some fixed constant. Let \( h > 0 \) and \( z_h^1, z_h^2 \in Z_{h,p} \) with \( \| z_h^1 - \Delta_h^B(z_*) \|_{Z_{h,p}} < \delta \) and \( \| F_h(z_h^1) - F_h(\Delta_h^B(z_*)) \|_{Y_{h,p}} < \epsilon, i = 1, 2 \), be given.

The functions \( F_{h,j} \) are Lipschitz continuous. We apply Lebourg’s mean-value theorem, see Clarke [8, Theorem 2.3.7], component-wise on \( F_h \). The theorem yields the existence of \( \zeta_h^i \in [z_h^1, z_h^2] \subset Z_{h,p} \), such that

\[
F_{h,j}(z_h^1)(t_i) - F_{h,j}(z_h^2)(t_i) = \partial F_{h,j}(\zeta_h^i)(t_i)(z_h^1(t_i) - z_h^2(t_i)), \quad j = 1, \ldots, M.
\]
Since
\[ F_h(z_h^i)(t_i) - F_h(z_h^j)(t_i) \in \begin{pmatrix} \partial F_{h,1}(\zeta^1_h)(t_i) \\ \vdots \\ \partial F_{h,M}(\zeta^M_h)(t_i) \end{pmatrix} (z_h^i(t_i) - z_h^j(t_i)), \]
there exists
\[ W_h \in \begin{pmatrix} \partial F_{h,1}(\zeta^1_h) \\ \vdots \\ \partial F_{h,M}(\zeta^M_h) \end{pmatrix} \]
with
\[ F_h(z_h^1) - F_h(z_h^2) = W_h(z_h^1 - z_h^2). \]
The assumptions yield
\[ \| z_h^1 - z_h^2 \|_{Z_h,p} = \| W_h^{-1}(F_h(z_h^1) - F_h(z_h^2)) \|_{Z_h,p} \leq C \| F_h(z_h^1) - F_h(z_h^2) \|_{Y_h,p}. \]
Under these assertions, Lempio [19, Corollary 6.1.20] shows the stability at \( \Delta_z^{h,p} \) with stability bound \( C \) and stability threshold \( \min\{\epsilon, \delta/C\} \) and the existence of solutions of \( F(z_h) = 0 \) if
\[ \lim_{h \to 0} F_h(\Delta_z^{h,p}(z_*)) = 0, \]
which follows from consistency, see Theorem 4.4.

The convergence of \( M \) follows instantly:

**Theorem 4.6.** Let all the assumptions from Theorem 4.5 be satisfied. Then the discretization method is convergent of order \( 1/p \) at \( z_* \).

**Proof.** Theorem 4.5 guarantees the existence of solutions \( z_h \) of \( F_h(z_h) = 0 \) for sufficiently small \( h > 0 \). It holds
\[ \| F_h(z_h) - F_h(\Delta_z^{h,p}(z_*)) \|_{Y_h,p} = \| F_h(\Delta_z^{h,p}(z_*)) \|_{Y_h,p}. \]
Stability at \( \Delta_z^{h,p}(z_*) \) implies
\[ \| z_h - \Delta_z^{h,p}(z_*) \|_{Z_h,p} \leq C \| F_h(z_h) - F_h(\Delta_z^{h,p}(z_*)) \|_{Y_h,p} = C \| F_h(\Delta_z^{h,p}(z_*)) \|_{Y_h,p} \]
Since \( \| F_h(\Delta_z^{h,p}(z_*)) \|_{Y_h,p} = O(h^{1/p}) \) holds by Theorem 4.4, it follows
\[ \| z_h - \Delta_z^{h,p}(z_*) \|_{Z_h,p} = O(h^{1/p}). \]

It remains to establish sufficient conditions for the regularity of
\[ W_h \in \begin{pmatrix} \partial F_{h,1}(\zeta^1_h) \\ \vdots \\ \partial F_{h,M}(\zeta^M_h) \end{pmatrix} \]
and the uniform boundedness of \( W_h^{-1} \) in \( \mathcal{L}(Y_h,p, Z_h,p) \). The chain rule for generalized derivatives in finite dimension, see Clarke [8, p. 72ff], yields \( \partial F_{h,j}(\zeta^j_h) \subseteq \partial, F_{h,j}(\zeta^j_h) \) with \( \partial F_h \) from Section 3.

We will show the assertions for all \( W_h \in \mathcal{L}(Z_h,p, Y_h,p) \) in the larger set
\[ \begin{pmatrix} \partial F_{h,1}(\zeta^1_h) \\ \vdots \\ \partial F_{h,M}(\zeta^M_h) \end{pmatrix}. \]
The linear operator equation \( W_h(\delta z_h) = w_h \) reads:

\[
\begin{align*}
\left( \frac{\delta x_i(t_i)}{\delta \lambda_h(t_i)} \right) + A_h(t_i) \left( \frac{\delta x_h(t_i)}{\delta \lambda_h(t_{i+1})} \right) + B_h(t_i) \left( \frac{\delta u_h(t_i)}{\delta \eta_h(t_i)} \right) &= w_{h,d}(t_i), \\
C_h(t_i) \left( \frac{\delta x_h(t_i)}{\delta \lambda_h(t_{i+1})} \right) + D_h(t_i) \left( \frac{\delta u_h(t_i)}{\delta \eta_h(t_i)} \right) &= w_{h,a}(t_i), \\
E_{h,0} \left( \frac{\delta x_h(t_0)}{\delta \lambda_h(t_0)} \right) + E_{h,f} \left( \frac{\delta x_h(t_N)}{\delta \lambda_h(t_N)} \right) + E_{h,r} \delta \sigma_h &= w_{h,r},
\end{align*}
\]

for \( i = 0, \ldots, N - 1 \), where \( \delta z_h := (\delta x_h, \delta u_h, \delta \lambda_h, \delta \eta_h, \delta \sigma_h) \) and \( w_h := (w_{h,d}, w_{h,a}, w_{h,r}) \). We reordered the components of \( F_h \) for the sake of presentation. The matrices \( A_h, B_h, C_h, D_h, E_{h,0}, E_{h,f}, \) and \( E_{h,r} \) depend on \( \zeta_1^h, \ldots, \zeta_M^h \).

If \( D_h(t_i) \) is nonsingular we can solve Equation (17) with respect to \( \delta u_h(t_i) \) and \( \delta \eta_h(t_i) \) and insert the resulting expression for \( (\delta u_h(t_i), \delta \eta_h(t_i)) \) into (16). This yields the reduced system

\[
\begin{align*}
\left( \frac{\delta x_i(t_i)}{\delta \lambda_h(t_i)} \right) + A_h(t_i) \left( \frac{\delta x_h(t_i)}{\delta \lambda_h(t_{i+1})} \right) &= b_h(t_i), \\
E_{h,0} \left( \frac{\delta x_h(t_0)}{\delta \lambda_h(t_0)} \right) + E_{h,f} \left( \frac{\delta x_h(t_N)}{\delta \lambda_h(t_N)} \right) + E_{h,r} \delta \sigma_h &= w_{h,r},
\end{align*}
\]

for \( i = 0, \ldots, N - 1 \), where

\[
\begin{align*}
A_h(t_i) &:= A_h(t_i) - B_h(t_i)D_h(t_i)^{-1}C_h(t_i), \\
b_h(t_i) &:= w_{h,d}(t_i) - B_h(t_i)D_h(t_i)^{-1}w_{h,a}(t_i).
\end{align*}
\]

Note that the system (19)–(20) resembles the system which is discussed in Gerds [13, Eq. (3.6)]. Also the subsequent theorem can be considered as a discrete version of Gerds [13, Theorem 3.4] and the techniques used in the proof are similar.

**Theorem 4.7.** Let \( z_* \in Z_p \) be a locally optimal solution satisfying \( F(z_*) = 0 \) and let the following assumptions hold with uniform constants for all \( \zeta_1^h \subseteq Z_{h,p} \) with \( \|\zeta_1^h - \Delta_h^p(z_*)\|_{Z_{h,p}} < \delta \), \( j = 1, \ldots, M \) and all \( h > 0 \) sufficiently small.

(i) Let \( D_h(t_i) \) be nonsingular for all \( i \in \{0, \ldots, N - 1\} \) and let there exist a constant \( C_D > 0 \) such that

\[
\|D_h^{-1}(t_i)\| \leq C_D,
\]

for all \( i \in \{0, \ldots, N - 1\} \).

(ii) Let there exist \( C_B > 0 \) such that

\[
\|B_h(t_i)\| \leq C_B,
\]

for all \( i \in \{0, \ldots, N - 1\} \).

(iii) Let there exist \( C_C > 0 \) such that

\[
\|C_h\|_p := \left( \sum_{i=0}^{N-1} h_i \|C_h(t_i)\|^p \right)^{1/p} \leq C_C.
\]

(iv) Let there exist constants \( C_1 \subseteq (0,1), C_2 > 0, \) and \( C_3 > 0 \) with

\[
h_i \|\hat{A}_h(t_i)\| \leq C_1
\]
By Assumption (iv) it holds
\[ \prod_{i=0}^{N-1} (1 + h_i \| \hat{A}_h(t_i) \|) \leq C_2, \quad \prod_{i=0}^{N-1} (1 - h_i \| \hat{A}_h(t_i) \|) \geq C_3. \]

Furthermore, let there exist \( C_4 > 0 \) such that
\[ \| \hat{A}_h \|_p := \left( \sum_{i=0}^{N-1} h_i \| \hat{A}_h(t_i) \|^p \right)^{1/p} \leq C_4. \]

Note that all assumptions in (iv) are satisfied, if \( \| \hat{A}_h(t_i) \| \) is uniformly bounded.

(v) Let the matrix
\[ R_h := (E_{h,0} \Phi_h(t_0) + E_{h,f} \Phi_h(t_N) \quad E_{h,r}) \]
be nonsingular and let there exist constants \( C_5 > 0 \) and \( C_6 > 0 \) s.t.
\[ \| E_{h,f} \| \leq C_5, \quad \| R_h^{-1} \| \leq C_6. \]

Herein, \( \Phi_h(t_i) = (X_h(t_i), \Lambda_h(t_i))^\top \) is the solution of the homogeneous problem
\[ \begin{cases} X'_h(t_i) + \hat{A}_h(t_i) X_h(t_i) = 0, & i = 0, \ldots, N-1, \quad \Phi_h(t_0) = I. \end{cases} \]

Then, \( W_h \) is nonsingular and there exists \( K > 0 \) with \( \| W_h^{-1} \| \leq K \) for all \( \zeta_h^i \in Z_{h,p} \) with \( \zeta_h^i - \Delta_h^p (z^*_s) \|_{Z_{h,p}} < \delta \) and all \( h > 0 \) sufficiently small.

Proof. Let \( w_h \in Y_{h,p}, \: D_h(t_i) \) is nonsingular by Assumption (i) and we first consider the reduced system (19)–(20). For this system we show the uniform nonsingularity. Thus we have to show that there exists a constant \( K > 0 \) s.t.
\[ \max \{ \| \delta x_h \|_{1,p}, \| \delta \lambda_h \|_{1,p}, \| \delta \sigma_h \| \} \leq K \max \{ \| b_h \|_p, \| w_{h,r} \| \}. \]

Reordering of Equation (19) with respect to \( \xi_h(t) := (\delta x_h(t_i), \delta \lambda_h(t_i)) \)
\[ \begin{pmatrix} I & h_i \hat{A}_h^{12}(t_i) \\ 0 & I + h_i \hat{A}_h^{22}(t_i) \end{pmatrix} \xi_h(t_{i+1}) = \begin{pmatrix} I & -h_i \hat{A}_h^{11}(t_i) \\ -h_i \hat{A}_h^{21}(t_i) & I \end{pmatrix} \xi_h(t_i) + h_i b_h(t_i). \] (21)

The matrices \( M_h(t_i) \) and \( P_h(t_i) \) are perturbed identity matrices with perturbations
\[ \Delta M_h(t_i) := \begin{pmatrix} 0 & h_i \hat{A}_h^{12}(t_i) \\ 0 & h_i \hat{A}_h^{22}(t_i) \end{pmatrix}, \quad \Delta P_h(t_i) := \begin{pmatrix} -h_i \hat{A}_h^{11}(t_i) & 0 \\ -h_i \hat{A}_h^{21}(t_i) & 0 \end{pmatrix}. \]

By Assumption (iv) it holds
\[ \| \Delta M_h(t_i) \| \leq h_i \| \hat{A}_h(t_i) \| \leq C_1 < 1 \quad \text{and} \quad \| \Delta P_h(t_i) \| \leq h_i \| \hat{A}_h(t_i) \| \leq C_1 < 1. \]

This yields \( \| M_h(t_i) \| = \| I + \Delta M_h(t_i) \| \leq 1 + C_1 \) and \( \| P_h(t_i) \| \leq 1 + C_1. \) The nonsingularity of the matrix \( M_h(t_i) \) is provided by the perturbation lemma, see [20, Thm. 3]. Furthermore the perturbation lemma yields
\[ \| M_h(t_i)^{-1} \| = \|(I + \Delta M_h(t_i))^{-1} \| = \|(I - (-\Delta M_h(t_i)))^{-1} \| \leq \frac{1}{1 - h_i \| \hat{A}_h(t_i) \|} \leq \frac{1}{1 - C_1}. \]

Thus, Equation (21) can be solved for \( \xi_h(t_{i+1}) \)
\[ \xi_h(t_{i+1}) = M_h(t_i)^{-1} P_h(t_i) \xi_h(t_i) + h_i M_h(t_i)^{-1} b_h(t_i), \]
and $\xi_h(t_N)$ is uniquely defined by $\xi_h(t_0)$. It holds
\[
\|\xi_h(t_{i+1})\| \leq \|M_h(t_i)^{-1}\|\|P_h(t_i)\|\|\xi_h(t_i)\| + h_i\|M_h(t_i)^{-1}\|\|b_h(t_i)\|,
\]
and repeated application yields
\[
\|\xi_h(t_i)\| \leq \left( \prod_{j=0}^{i-1} \|M_h(t_j)^{-1}\|\|P_h(t_j)\| \right) \|\xi_h(t_0)\|
\]
\[
+ \sum_{j=0}^{i-1} \left( \prod_{k=j+1}^{i-1} \|M_h(t_k)^{-1}\|\|P_h(t_k)\| \right) h_j\|M_h(t_j)^{-1}\|\|b_h(t_j)\|. \tag{22}
\]
With the same argumentation we show that the homogeneous problem from Assumption (v) has a unique solution $\Phi_h$. We obtain an estimate for $\Phi_h(t_i)$ if we set $b_h = 0$ and exploit $\Phi(t_0) = I$ in (22):
\[
\|\Phi_h(t_i)\| \leq \prod_{j=0}^{i-1} \|M_h(t_j)^{-1}\|\|P_h(t_j)\| \leq \prod_{j=0}^{N-1} \frac{(1 + \hat{h}_j \|A_h(t_j)\|)}{(1 - \hat{h}_j \|A_h(t_j)\|)} \leq \frac{C_2}{C_3},
\]
by Assumption (iv).
One can easily verify that $\xi_h$ may be represented as
\[
\xi_h(t_i) = \Phi_h(t_i)\xi_h(t_0) + \hat{\xi}_h(t_i), \tag{23}
\]
where $\hat{\xi}_h$ is the solution of the inhomogeneous initial value problem
\[
\begin{pmatrix}
\delta x_h'(t_i) \\
\delta \lambda_h'(t_i)
\end{pmatrix}
+ \hat{A}_h(t_i)
\begin{pmatrix}
\delta x_h(t_i) \\
\delta \lambda_h(t_{i+1})
\end{pmatrix}
= b_h(t_i), \quad \hat{\xi}_h(t_0) = 0.
\]
Equation (22) with $\hat{\xi}_h(t_0) = 0$ yields
\[
\|\hat{\xi}_h(t_i)\| \leq \sum_{j=0}^{i-1} \left( \prod_{k=j+1}^{i-1} \|M_h(t_k)^{-1}\|\|P_h(t_k)\| \right) h_j\|M_h(t_j)^{-1}\|\|b_h(t_j)\|
\]
\[
\leq \sum_{j=0}^{i-1} \frac{\prod_{k=j+1}^{i-1} (1 + h_i\|A_h(t_i)\|) h_j\|M_h(t_j)^{-1}\|\|b_h(t_j)\|}{\prod_{k=j+1}^{N-1} (1 - h_i\|A_h(t_i)\|)}
\]
\[
\leq \sum_{j=0}^{i-1} \frac{\prod_{k=0}^{N-1} (1 + h_i\|A_h(t_i)\|) h_j\|M_h(t_j)^{-1}\|\|b_h(t_j)\|}{\prod_{k=0}^{N-1} (1 - h_i\|A_h(t_i)\|)}
\]
\[
\leq \frac{C_2}{C_3} \sum_{j=0}^{i-1} h_j\|M_h(t_j)^{-1}\|\|b_h(t_j)\|. \tag{24}
\]
for \( i = 0, \ldots, N - 1 \) with a constant \( \kappa_1 := \frac{C_2}{C_3(1 - C_1)}. \)

We introduce \( \xi_h(t_N) \) from Equation (23) into the boundary conditions (20) and obtain

\[
\begin{align*}
(E_{h,0} \Phi_h(t_0) &+ E_{h,f} \Phi_h(t_N) - E_{h,r}) \left( \xi_h(t_0) \right) \frac{\delta \sigma_h}{\delta} + E_{h,f} \tilde{\xi}_h(t_N) \\
&= R_h \left( \xi_h(t_0) \right) \frac{\delta \sigma_h}{\delta} + E_{h,f} \tilde{\xi}_h(t_N) = w_{h,r}.
\end{align*}
\]

The matrix \( R_h \) is nonsingular and \( \tilde{\xi}_h(t_N) \) does not depend on the boundary conditions. This shows that the reduced system has a unique solution. Furthermore it holds

\[
\max(\|\xi_h(t_0)\|, \|\sigma_h\|) \leq \|R_h^{-1}(\|w_{h,r}\| + \|E_{h,f}\|\|\tilde{\xi}_h(t_N)\|)) \\
\leq \|R_h^{-1}\| (\|w_{h,r}\| + \kappa_1\|E_{h,f}\|\|b_h\|_1) \\
\leq C_6\|w_{h,r}\| + C_5C_6\kappa_1\|b_h\|_1,
\]

where we exploit (25) and Assumption (v). We conclude that

\[
\|\xi_h(t_i)\| = \|\Phi_h(t_i)\xi_h(t_0) + \tilde{\xi}_h(t_i)\| \\
\leq \|\Phi_h(t_i)\|(C_6\|w_{h,r}\| + C_5C_6\kappa_1\|b_h\|_1) + \kappa_1\|b_h\|_1 \\
\leq \frac{C_2C_6}{C_3}\|w_{h,r}\| + \left(1 + \frac{C_2C_6}{C_3}\right) \kappa_1\|b_h\|_1 \\
\leq \kappa_2\|w_{h,r}\| + \kappa_3\|b_h\|_1,
\]

with constants \( \kappa_2 = \frac{C_2C_6}{C_3} \) and \( \kappa_3 = 1 + \frac{C_2C_6}{C_3} \). This yields

\[
\max(\|\xi_h\|_p, \|\sigma_h\|) \leq \max(\|\xi_h\|_\infty, \|\sigma_h\|) \\
\leq \max(\kappa_2\|w_{h,r}\| + \kappa_3\|b_h\|_1, C_6\|w_{h,r}\| + C_5C_6\kappa_1\|b_h\|_1) \\
\leq \kappa_4 \max(\|b_h\|_p, \|w_{h,r}\|),
\]

with a constant \( \kappa_4 > 0. \)

It remains to analyze the \( p \)-norm of the derivative of \( \xi_h \). We apply Minkowski’s inequality and conclude

\[
\|\xi_h'\|_p = \left( \sum_{i=0}^{N-1} h_i \left\| b_h(t_i) - \hat{A}_h(t_i) \left( \frac{\delta x_h(t_i)}{\delta \lambda_h(t_{i+1})} \right) \right\|^p \right)^{1/p}
\]
\[
\left( \sum_{i=0}^{N-1} h_i \| b_h(t_i) \|^p \right)^{1/p} + \left( \sum_{i=0}^{N-1} h_i \left\| \hat{A}_h(t_i) \left( \frac{\delta x_h(t_i)}{\delta \lambda_h(t_{i+1})} \right) \right\|^p \right)^{1/p} \\
\leq \| b_h \|_p + \left( \sum_{i=0}^{N-1} h_i \| \hat{A}_h(t_i) \|^p \| \xi_h(t_i) \| + \| \xi_h(t_{i+1}) \| \right)^{1/p} \\
\leq \| b_h \|_p + \left( \sum_{i=0}^{N-1} h_i \| \hat{A}_h(t_i) \|^p (2\kappa_2 \| w_{h,r} \| + 2\kappa_3 \| b_h \|_1 \right)^{1/p} \\
\leq \| b_h \|_p + \left( \sum_{i=0}^{N-1} h_i \| \hat{A}_h(t_i) \|^p \kappa_7 \max \{ \| b_h \|_1, \| w_{h,r} \| \} \right)^{1/p} \\
\leq \| b_h \|_p + \kappa_5 \max \{ \| b_h \|_1, \| w_{h,r} \| \} \left( \sum_{i=0}^{N-1} h_i \| \hat{A}_h(t_i) \|^p \right)^{1/p} \\
\leq \| b_h \|_p + C_4 \kappa_5 \max \{ \| b_h \|_1, \| w_{h,r} \| \} \\
\leq \kappa_6 \max \{ \| b_h \|_p, \| w_{h,r} \| \},
\]

with constants \( \kappa_5 > 0, \kappa_6 > 0, \) and \( \kappa_7 > 0. \) This yields the uniform nonsingularity of the reduced system:

\[
\max \{ \| \xi_h \|_1, \| \delta \sigma_h \| \} \leq \max \{ \| \xi_h \|_p, \| \xi_h \|_p, \| \delta \sigma_h \| \} \\
\leq \max \{ \kappa_4, \kappa_6 \} \max \{ \| b_h \|_p, \| w_{h,r} \| \}.
\]

The full system \((16)-(18)\) has a unique solution: the solution \( x_h, \lambda_h \) and \( \sigma_h \) of the reduced system and

\[
\left( \begin{array}{c}
\delta u_h(t_i) \\
\delta \eta_h(t_i)
\end{array} \right) = D_h(t_i)^{-1} w_{h,a}(t_i) - D_h(t_i)^{-1} C_h(t_i) \left( \begin{array}{c}
\delta x_h(t_i) \\
\delta \lambda_h(t_{i+1})
\end{array} \right).
\]

By Assumptions (i) and (iii) we show in a similar way that

\[
\left\| \left( \begin{array}{c}
\delta u_h \\
\delta \eta_h
\end{array} \right) \right\|_p \leq \left( \sum_{i=0}^{N-1} h_i \| D_h(t_i)^{-1} w_{h,a}(t_i) \|^p \right)^{1/p} + \left( \sum_{i=0}^{N-1} h_i \left\| D_h(t_i)^{-1} C_h(t_i) \left( \begin{array}{c}
\delta x_h(t_i) \\
\delta \lambda_h(t_{i+1})
\end{array} \right) \right\|^p \right)^{1/p} \\
\leq C_D \left( \sum_{i=0}^{N-1} h_i \| w_{h,a}(t_i) \|^p \right)^{1/p} + C_D \left( \sum_{i=0}^{N-1} h_i \| C_h(t_i) \|^p \| \xi_h(t_i) \| + \| \xi_h(t_{i+1}) \| \right)^{1/p} \\
\leq C_D \| w_{h,a} \|_p + C_D \kappa_5 \max \{ \| b_h \|_1, \| w_{h,r} \| \} \left( \sum_{i=0}^{N-1} h_i \| C_h(t_i) \|^p \right)^{1/p} \\
\leq C_D \| w_{h,a} \|_p + C_D C_4 \kappa_5 \max \{ \| b_h \|_p, \| w_{h,r} \| \}, \\
\leq \kappa_8 \max \{ \| w_{h,a} \|_p, \| b_h \|_p, \| w_{h,r} \| \}
\]

with a constant \( \kappa_8 > 0. \)
The right hand side of the reduced system
\[ b_h(t_i) = w_{h,a}(t_i) - B_h(t_i) D_h(t_i)^{-1} w_{h,a}(t_i) \]
satisfies
\[
\|b_h\|_p \leq \|w_{h,a}\|_p + \left( \sum_{i=0}^{N-1} h_i \|B_h(t_i) D_h(t_i)^{-1} w_{h,a}(t_i)\|_p \right)^{1/p}
\]
\[
\leq \|w_{h,a}\|_p + C_D C_B \left( \sum_{i=0}^{N-1} h_i \|w_{h,a}(t_i)\|_p \right)^{1/p}
\]
\[
\leq \|w_{h,a}\|_p + C_D C_B \|w_{h,a}\|_p,
\]
by assumptions (i) and (ii).

Summarizing, we have shown that there exists \( K > 0 \) such that
\[
\max\{\|\delta x_h\|_{1,p}, \|\delta \lambda_h\|_{1,p}, \|\delta u_h\|_p, \|\delta \eta_h\|_p, \|\delta \sigma_h\|\} \leq K \max\{\|w_{h,a}\|_p, \|w_{h,a}\|_p, \|w_{h,a}\|\}.
\]

\[ \square \]

**Remark 2.** The discrete error \( \|z_h - \Delta^p_h(z_*)\|_{Z_{h,p}} \) lacks the problem that it does not necessarily see jumps in \( z_* \in \tilde{Z}_p \). One would rather like to work with the continuous error
\[
\|z_h - z_*\|_{Z_p}.
\]
This error is well defined, since \( Z_{h,p} \subset \tilde{Z}_p \). We also get convergence in \( \tilde{Z}_p \), since
\[
\|z_h - z_*\|_{\tilde{Z}_p} \leq \|z_h - \Delta^p_h(z_*)\|_{\tilde{Z}_p} + \|\Delta^p_h(z_*) - z_*\|_{\tilde{Z}_p}
\]
\[
= \|z_h - \Delta^p_h(z_*)\|_{Z_{h,p}} + \|\Delta^p_h(z_*) - z_*\|_{\tilde{Z}_p}
\]
\[
\leq O(h^{1/p}) + \|\Delta^p_h(z_*) - z_*\|_{\tilde{Z}_p}.
\]
The quantity \( \|\Delta^p_h(z_*) - z_*\|_{\tilde{Z}_p} \) is again of order \( h^{1/p} \) which can be seen from the proofs of Lemma 4.3 and Corollary 1.

**Remark 3.** The assumptions in Theorem 4.7 and especially the assumption on the nonsingularity of the matrices \( D_h \) exclude problems with pure state constraints and are not automatically satisfied for problems with the control appearing linearly. To this end a more detailed investigation of subclasses that satisfy the assumptions of Theorem 4.7 would be desirable and is the subject of future research.

5. **Numerical illustrations.** Problem 2 is a finite optimization problem and various established software packages for its solution are available, see \url{http://plato.asu.edu/guide.html} for an overview and classification. Note that although the convergence proof is motivated by a nonsmooth reformulation of optimality conditions, it is not necessary to solve the discrete problems with Newton’s method. The subsequent results are obtained with the solver WORHP, see Büskens et al. [6]. WORHP is an implementation of the sequential quadratic programming (SQP) method designed for smooth large-scale and sparse nonlinear optimization problems. The solver implements two strategies to tackle the quadratic subproblems, a nonsmooth Newton’s method and an interior point method. Both methods lead to large-scale and sparse linear equation systems which are solved by the direct decomposition solver MA57, see Duff [11]. WORHP implements several fall-back strategies for
difficulties (like ill-conditioning, breakdown of linesearch strategies, etc) that might occur while solving the optimization problem.

We consider the optimal control problem

**Problem 3.** Let $T := [t_0, t_f]$, $a \in BV(T, \mathbb{R})$ continuous with only countably many roots and $b \in BV(T, \mathbb{R}_+^1)$ continuous.

Find $x \in W^{1, \infty}(T, \mathbb{R})$ and $u \in L^{\infty}(T, \mathbb{R})$ such that

$$x(t_f)$$
is minimized subject to

$$x'(t) = a(t)u(t), \quad a.e. \text{ in } [t_0, t_f],$$

$$x(t_0) = 0,$$

and

$$0 \leq u(t) \leq b(t), \quad a.e. \text{ in } [t_0, t_f].$$

We set parameters $t_0 := -2$, $t_f := 2$,

$$a(t) := \begin{cases} 0, & \text{if } t = 0, \\ t^2 \sin \left(\frac{5\pi}{T}\right), & \text{if } t \neq 0, \end{cases}$$

and $b(t) := t^2$.

The globally optimal control $u$ of Problem 3 is given by

$$u(t) = \begin{cases} b(t), & \text{if } a(t) < 0, \\ 0, & \text{if } a(t) > 0. \end{cases}$$

$$= \begin{cases} t^2, & \text{if } t \in [-2, -\frac{5}{3}), \\ t^2, & \text{if } t \in \left[-\frac{5}{k}, -\frac{5}{k+1}\right), \quad k = 4, 6, 8, \ldots, \\ t^2, & \text{if } t \in \left[\frac{5}{k}, \frac{5}{k-1}\right), \quad k = 4, 6, 8, \ldots, \\ 0, & \text{else.} \end{cases}$$

The state $x$ associated with the optimal control is obtained by integrating the state equation

$$x(t) = x(t_0) + \int_{t_0}^{t} a(\tau)u(\tau) \, d\tau = \int_{t_0}^{t} a(\tau)u(\tau) \, d\tau.$$  

The optimal objective value is approximately $x(2) \approx -4.3349$. From the local minimum principle we derive $\lambda(t) = 1$, $\sigma = -1$, and

$$\eta_1(t) = \begin{cases} 0, & \text{if } a(t) < 0, \\ a(t), & \text{if } a(t) > 0, \end{cases}$$

$$\eta_2(t) = \begin{cases} -a(t), & \text{if } a(t) < 0, \\ 0, & \text{if } a(t) > 0, \end{cases}$$

where $\eta_1$ and $\eta_2$ are the multipliers associated with $-u(t) \leq 0$ and $u(t) - b(t) \leq 0$ respectively.

$u$ is piecewise monotone and has jumps at $s_k := \frac{5}{k}, \quad k = 3, 4, 5, \ldots$ and $-s_k$. Let $s_2 = 2$. The total variation of $u$ is (compare Figure 1)

$$\text{Var}_{-2}^2(u) = \sum_{k=3}^{\infty} |u(-s_k^-) - u(-s_k^+)| + \sum_{k=3}^{\infty} |u(s_k^-) - u(s_k^+)|$$

jumps in $(-2, 0)$

jumps in $(0, 2)$
Note that all single series converge since they are dominated by all $t_i$ assumption, $F(z)$ admits a unique solution in $Z_{\infty}$. The discrete problems $F_h(z_i)$ only posses unique solutions, if $a(t_i) \neq 0$ for $i = 0, \ldots, N - 1$. Thus, we assume

$$
\{ t \in I : a(t) = 0 \} \cap \mathbb{G}_h = \emptyset.
$$

Similarly we prove that the multipliers $\eta_1$ and $\eta_2$ are of bounded variation with $\text{Var}_2^2(\eta_1) = \text{Var}_2^2(\eta_2) = -\frac{731}{36} + \frac{25\pi^2}{6}$. The multipliers are plotted in Figure 2.

**Remark 4.** Note that $u$, $\eta_1$, and $\eta_2$ are uniquely determined by $F(z)(t) = 0$ at all $t$ with $a(t) > 0$ or $a(t) < 0$. Since the set $\{ t \in I : a(t) = 0 \}$ is countable by assumption, $F(z)$ admits a unique solution in $Z_{\infty}$. The discrete problems $F_h(z_i)$ only posses unique solutions, if $a(t_i) \neq 0$ for $i = 0, \ldots, N - 1$. Thus, we assume

$$
\{ t \in I : a(t) = 0 \} \cap \mathbb{G}_h = \emptyset.
$$

---

**Figure 1.** The optimal control $u$ has countable infinitely many jumps.
Let us now verify the assumptions of Theorem 4.7 for solutions $z_h$ of the discrete problems $F_h(z_h) = 0$. The matrices in (16)–(17) read

$$A_h(t_i) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad B_h(t_i) = \begin{pmatrix} -a(t_i) & 0 \\ 0 & 0 \end{pmatrix},$$

$$C_h(t_i) = \begin{pmatrix} 0 & a(t_i) \\ 0 & 0 \end{pmatrix}, \quad D_h(t_i) = \begin{pmatrix} 0 & 1 & -1 \\ s_1^2(t_i) & r_1^2(t_i) & 0 \\ -s_2^2(t_i) & 0 & r_2^2(t_i) \end{pmatrix},$$

the matrix $\hat{A}_h(t)$ in the reduced system (19) reads

$$\hat{A}_h(t_i) = \frac{a(t_i)^2 r_1^2(t_i) r_2^2(t_i)}{\det D_h(t_i)} \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

if $D_h(t_i)$ is nonsingular, and the matrices in (18) are given by

$$E_{h,0} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_{h,f} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_{h,r} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. $$

It holds

$$(s_1^1(t_i), r_1^1(t_i)) \in \partial \varphi(u_h(t_i), \eta_{h,1}(t_i)) = \begin{cases} (0,-1), & a(t_i) < 0, \\ (-1,0), & a(t_i) > 0, \end{cases}$$

$$(s_2^2(t_i), r_2^2(t_i)) \in \partial \varphi(b(t_i) - u_h(t_i), \eta_{h,2}(t_i)) = \begin{cases} (-1,0), & a(t_i) < 0, \\ (0,-1), & a(t_i) > 0, \end{cases}$$

and

$$\det D_h(t_i) = -s_1^2(t_i) r_2^2(t_i) - s_2^2(t_i) r_1^2(t_i) = -1,$$

for all $i = 0, \ldots, N - 1$. Thus, we obtain $\|D_h(t_i)^{-1}\| = 2$, $\hat{A}_h(t_i) = 0$, $\|B_h(t_i)\| \leq \abs{a(t_i)}$, $\|C_h(t_i)\| \leq \abs{a(t_i)}$, and assumptions (i)–(iv) of Theorem 4.7 are fulfilled since $a$ is bounded by definition. Assumption (v) is satisfied since from $\hat{A}_h(t_i) = 0$ follows $\Phi_h(t_N) = I$ and

$$R_h := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}. $$
Table 1. Order of convergence for discontinuous control

| p  | q_p | C_p   |
|----|-----|-------|
| 1  | 1.006241 | 30.47329 |
| 2  | 0.5381130 | 4.381088  |
| 3  | 0.3569428 | 3.429845  |
| 4  | 0.2661071 | 3.137542  |

The order of convergence may be estimated by determining the optimal $C_p$ and $q_p$ such that $\|z_h - z_\ast\|_{Z_p} \approx C_p h^{q_p}$ in least-squares sense: we solve the discretized problem for various $h_i$, $i = 1, \ldots, M$, and then solve

$$\minimize \sum_{i=1}^{M} \left( \frac{\|z_h - z_\ast\|_{Z_p} - C_p h_i^{q_p}}{\|z_h_i - z_\ast\|_{Z_p}} \right)^2$$

with respect to $C_p$ and $q_p$. The continuous $L^p$ norms are integrated by an adaptive integration scheme. Table 1 summarizes the orders $q_p$, which are in good agreement with the theoretical orders $1/p$.

6. Conclusion. We studied discretized control-state constrained optimal control problems using explicit Euler’s method and prove its convergence of order $1/p$ in the $L^p$-norm. In the proof we assumed the optimal control to be merely of bounded variation. The proof technique follows and extends the general convergence theory in [26] towards nonsmooth equations since the classic approach cannot be used for discontinuous controls. The main convergence result in Theorem 4.6 can be summarized as follows: consistency and stability implies convergence.

There are several directions for future research. The convergence proof was conducted for state discretization by explicit Euler’s method. Further research towards higher order discretization methods has to show whether higher order approximations are possible under suitable smoothness assumptions. Another topic for future research is the technical assumptions in Theorem 4.5 regarding the uniform nonsingularity of the operators $W_h$. It would be desirable to have assumptions that can be verified in the optimal solution of the continuous problem rather than in the optimal solutions of the discretized problems. Finally, the approach used in this paper cannot be extended in a straightforward way to problems with pure state constraints. A potential approach for pure state constraints could be to use regularization techniques similar to [14] in order to embed the problem with pure state constraints into a family of problems with mixed control-state constraints. Then, the uniform convergence with regard to the regularization parameter has to be shown. This will be the subject of future research.

Appendix A. Proof of Lemma 4.3.

Proof. (i) Norm equivalence in the finite dimensional space $\mathbb{R}^N$ yields

$$\max_{i \in \{1, \ldots, N\}} |a_i| \leq \left( \sum_{i=1}^{N} |a_i|^p \right)^{1/p} \leq \sum_{i=1}^{N} |a_i| \leq N^{1-1/p} \left( \sum_{i=1}^{N} |a_i|^p \right)^{1/p} \leq (c_2 h^{-1})^{1-1/p} \left( \sum_{i=1}^{N} |a_i|^p \right)^{1/p}$$
Application of the fundamental theorem of calculus yields the assertion, see Natanson [24, Chapter IX, 4].

Let $v \in L^p_h(T, \mathbb{R})$. Then,

$$
\|v_h\|_p = \left( \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} |v_h(t_i)|^p \, dt \right)^{1/p} \leq \sum_{i=0}^{N-1} h_i |v_h(t_i)|^{1/p} \leq \sum_{i=0}^{N-1} h_i^{1/p} |v_h(t_i)|^{1/p}.
$$

Let $v \in BV(T, \mathbb{R})$. Then, $v$ restricted to the interval $[t_i, t_{i+1}]$ is also a function of bounded variation and by Natanson [24, Chapter VIII, 3] the following holds.

$$
|v(t)| \leq |v(t_i)| + \text{Var}_{t_i}^{t_{i+1}}(v), \quad \forall t \in [t_i, t_{i+1}]. \tag{26}
$$

The function $v(t_i) - v(\cdot)$ is of bounded variation and thus

$$
|v(t_i) - v(t)| \leq |v(t_i) - v(t_i)| + \text{Var}_{t_i}^{t_{i+1}}(v) \leq \text{Var}_{t_i}^{t_{i+1}}(v).
$$

We exploit this assertion piecewise and obtain

$$
\|\Delta_h^0(v) - v\|_p = \left( \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} |v(t_i) - v(t)|^p \, dt \right)^{1/p} \leq \left( \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \left( \text{Var}_{t_i}^{t_{i+1}}(v) \right)^p \, dt \right)^{1/p} \leq h^{1/p} \left( \sum_{i=0}^{N-1} \left( \text{Var}_{t_i}^{t_{i+1}}(v) \right)^p \right)^{1/p} \leq h^{1/p} \sum_{i=0}^{N-1} \text{Var}_{t_i}^{t_{i+1}}(v) \leq h^{1/p} \|v\|_{BV}.
$$

Let $v \in BV^1(T, \mathbb{R})$. Then for $t \in [t_i, t_{i+1}]$ the following holds.

$$
|\Delta_h^1(v)(t) - v(t)| = \left| v(t_i) + \frac{t - t_i}{h_i} (v(t_{i+1}) - v(t_i)) - v(t) \right| \leq |v(t_i) + 0 - v(t_i)| + \text{Var}_{t_i}^{t_{i+1}} \left( \frac{v(t_{i+1}) - v(t_i)}{h_i} t - v(t) \right) \leq \text{Var}_{t_i}^{t_{i+1}} \left( \frac{v(t_{i+1}) - v(t_i)}{h_i} t \right) + \text{Var}_{t_i}^{t_{i+1}}(v) \leq \frac{|v(t_{i+1}) - v(t_i)|}{h_i} (t_{i+1} - t_i) + \text{Var}_{t_i}^{t_{i+1}}(v) \leq 2 \text{Var}_{t_i}^{t_{i+1}}(v).
$$
Thus,

\[ \| \Delta_h^1(v) - v \|_p = \left( \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \left| v(t_{i+1}) - v(t_i) + \frac{t-t_i}{h_i} (v(t_{i+1}) - v(t_i)) - v(t) \right|^p dt \right)^{1/p} \]

\[ \leq \left( \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \left( 2 \text{Var}_{t_i}^{t_{i+1}}(v) \right)^p dt \right)^{1/p} \]

\[ \leq h^{1/p} \left( \sum_{i=0}^{N-1} \left( 2 \text{Var}_{t_i}^{t_{i+1}}(v) \right)^p \right)^{1/p} \]

(i) \[ \leq h^{1/p} \sum_{i=0}^{N-1} 2 \text{Var}_{t_i}^{t_{i+1}}(v) \]

\[ = 2h^{1/p} \text{Var}_{t_0}^{t_N}(v) \leq 2h^{1/p} \| v \|_{BV}. \]

(vi) For all \( v \in BV^1(T, \mathbb{R}) \) the following holds.

\[ \| \Delta_h^1(v)' - v' \|_p = \left( \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \left| \frac{v(t_{i+1}) - v(t_i)}{h_i} - v'(t_i) \right|^p dt \right)^{1/p} \]

\[ \leq \left( \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \left( \left| \frac{v(t_{i+1}) - v(t_i)}{h_i} - v'(t_i) \right| + \text{Var}_{t_i}^{t_{i+1}}(v') \right)^p dt \right)^{1/p} \]

\[ \leq h^{1/p} \left( \sum_{i=0}^{N-1} \left( \left| \frac{v(t_{i+1}) - v(t_i)}{h_i} - v'(t_i) \right| + \text{Var}_{t_i}^{t_{i+1}}(v') \right)^p \right)^{1/p} \]

(i) \[ \leq h^{1/p} \sum_{i=0}^{N-1} \left( \left| \frac{v(t_{i+1}) - v(t_i)}{h_i} - v'(t_i) \right| + \text{Var}_{t_i}^{t_{i+1}}(v') \right) \]

\[ \leq h^{1/p} \sum_{i=0}^{N-1} \frac{1}{h_i} \int_{t_i}^{t_{i+1}} v'(t) dt - v'(t_i) + h^{1/p} \text{Var}_{t_0}^{t_N}(v') \]

\[ = h^{1/p} \sum_{i=0}^{N-1} \frac{1}{h_i} \int_{t_i}^{t_{i+1}} v'(t) - v'(t_i) dt + h^{1/p} \| v' \|_{BV} \]

\[ \leq h^{1/p} \sum_{i=0}^{N-1} \frac{1}{h_i} \int_{t_i}^{t_{i+1}} |v'(t) - v'(t_i)| dt + h^{1/p} \| v' \|_{BV} \]

\[ \leq h^{1/p} \sum_{i=0}^{N-1} \frac{1}{h_i} \int_{t_i}^{t_{i+1}} \text{Var}_{t_i}^{t_{i+1}}(v') dt + h^{1/p} \| v' \|_{BV} \]

\[ \leq h^{1/p} \sum_{i=0}^{N-1} \text{Var}_{t_i}^{t_{i+1}}(v') + h^{1/p} \| v' \|_{BV} \]

\[ \leq h^{1/p} \| v' \|_{BV} + h^{1/p} \| v' \|_{BV} = 2h^{1/p} \| v' \|_{BV}. \]
REFERENCES

[1] W. Alt, Discretization and mesh-independence of Newton’s method for generalized equations, in “Mathematical Programming with Data Perturbations,” (ed. A. Fiacco), 17th symposium, George Washington University, Washington, DC, USA, May 1995. New York, NY: Marcel Dekker. Lect. Notes Pure Appl. Math., 195 (1997), 1–30.

[2] W. Alt, Mesh-independence of the Lagrange-Newton method for nonlinear optimal control problems and their discretizations, Optimization with data perturbations, II. Annals of Operations Research, 101 (2001), 101–117.

[3] W. Alt, R. Baier, M. Gerdts and F. Lempio, Approximation of linear control problems with bang-bang solutions, Optimization, 62 (2013), 9–32.

[4] W. Alt, R. Baier, M. Gerdts and F. Lempio, Error bounds for Euler approximation of linear-quadratic control problems with bang-bang solutions, Numerical Algebra, Control and Optimization, 2 (2012), 547–570.

[5] N. Banihashemi and C. Y. Kaya, Inexact restoration for Euler discretization of box-constrained optimal control problems, Journal of Optimization Theory and Applications, 156 (2013), 726–760.

[6] C. B"uskens, M. Gerdts, T. Nikolayzik, P. Kalmbach, M. Kunkel and D. Wassel, Homepage of the WORHP solver, http://www.worhp.de, 2010.

[7] B. Chen, X. Chen and C. Kanzow, A penalized Fischer-Burmeister NCP-function, Mathematical Programming, Series A, 88 (2000), 211–216.

[8] F. H. Clarke, “Optimization and Nonsmooth Analysis,” Canadian Mathematical Society Series of Monographs and Advanced Texts, New York: John Wiley & Sons, Inc., 1983.

[9] A. L. Dontchev, W. W. Hager and K. Malanowski, Error bounds for Euler approximation of a state and control constrained optimal control problem, Numerical Functional Analysis and Optimization, 21 (2000), 653–682.

[10] A. L. Dontchev, W. W. Hager and V. M. Veliov, Second-order runge-kutta approximations in control constrained optimal control, SIAM Journal on Numerical Analysis, 38 (2000), 202–226.

[11] I. S. Duff, MA57 — A code for the solution of sparse symmetric definite and indefinite systems, ACM Transactions on Mathematical Software, 30 (2004), 118–144.

[12] C. Geiger and C. Kanzow, “Theorie und Numerik Restringierter Optimierungsaufgaben,” Springer, Berlin-Heidelberg-New York, 2002.

[13] M. Gerdts, Global convergence of a nonsmooth Newton method for control-state constrained optimal control problems, SIAM Journal on Optimization, 19 (2008), 326–350. Erratum: M. Gerdts, SIAM Journal on Optimization, 21 (2011), 615–616, http://goo.gl/u4CJe.

[14] M. Gerdts and B. H"upping, Virtual control regularization of state constrained linear quadratic optimal control problems, Comput. Optim. Appl., 51 (2012), 867–882.

[15] W. W. Hager, Runge-Kutta methods in optimal control and the transformed adjoint system, Numerische Mathematik, 87 (2000), 247–282.

[16] H. Heuser, “Funktionalanalysis: Theorie und Anwendung,” B. G. Teubner, Stuttgart, 2006.

[17] M. Josefthy, Compounding functions of bounded variation, Proceedings of the American Mathematical Society, 83 (1981), 354–356.

[18] M. Kunkel, “Nonsmooth Newton Methods and Convergence of Discretized Optimal Control Problems Subject to DAEs,” PhD thesis, Universität der Bundeswehr München, Fakultät für Luft- und Raumfahrttechnik, 2012. urn:nbn:de:bvb:706–2790.

[19] F. Lempio, Numerische mathematik II — methoden der analysis, Bayreuther Mathematische Schriften, 55 (1998).

[20] L. A. Ljusternik and W. I. Sobolew, “Elemente Der Funktionalanalysis,” Fünfte Auflage. Übersetzung der zweiten russischen Auflage von Klaus Fiedler und herausgegeben von Konrad Gröger. Mathematische Lehrbcher und Monographien, I. Abteilung: Mathematische Lehrbcher, Band VIII. Akademie-Verlag, Berlin, 1976.

[21] K. Malanowski, On normality of Lagrange multipliers for state constrained optimal control problems, Optimization, 52 (2003), 75–91.

[22] K. Malanowski, Ch. Biskens and H. Maurer, Convergence of approximations to nonlinear optimal control problems, in “Mathematical Programming with Data Perturbations” (ed. A. Fiacco), volume 195, “Lecture Notes in Pure and Appl. Math.,” Dekker, New York, (1998), 253–284.
[23] M. McAsey, L. Mou and W. Han, *Convergence of the forward-backward sweep method in optimal control*, Computational Optimization and Applications, **53** (2012), 207–226.
[24] I. P. Natanson, “Theorie der Funktionen Einer Reellen Veränderlichen,” Verlag Harri Deutsch, Thun-Frankfurt/M., 1981.
[25] R. Loxton, Q. Lin, V. Rehbock and K. L. Teo, *Control parametrization for optimal control problems with continuous inequality constraints: New convergence results*, Numerical Algebra, Control and Optimization, **2** (2012), 571–599.
[26] H. J. Stetter, *Analysis of discretization methods for ordinary differential equations*, In “Springer Tracts in Natural Philosophy,” Springer, Berlin-Heidelberg-New York, **23** (1973).
[27] D. Sun and L. Qi, *On NCP-functions*, Computational optimizationa tribute to Olvi Mangasarian, Part II. Comput. Optim. Appl., **13** (1999), 201–220.
[28] V. M. Veliov, *Error analysis of discrete approximations to bang-bang optimal control problems: The linear case*, Control Cybern., **34** (2005), 967–982.

Received October 2012; 1st revision February 2013; final revision July 2013.

E-mail address: matthias.gerdts@unibw.de
E-mail address: ich@martinkunkel.de