Necessary Conditions, Sufficient Conditions, and Convergence Analysis for Optimal Control Problems with Differential-Algebraic Equations

Björn Martens

Vollständiger Abdruck der von der Fakultät für Luft- und Raumfahrttechnik der Universität der Bundeswehr München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Gutachter/Gutachterin:

1. Prof. Dr. rer. nat. Matthias Gerdts
2. Prof. Dr. rer. nat. Sabine Pickenhain

Die Dissertation wurde am 11.03.2019 bei der Universität der Bundeswehr München eingereicht und durch die Fakultät für Luft- und Raumfahrttechnik am 17.06.2019 angenommen. Die mündliche Prüfung fand am 12.07.2019 statt.

Abstract

In this thesis, we study optimal control problems subject to differential-algebraic equations in Hessenberg form and mixed control-state constraints. Specifically, we derive necessary and sufficient conditions for problems with Hessenberg differential-algebraic equations of arbitrary order, and then examine convergence properties of an approximated optimal control problem for the index two case.

The first part of this thesis is dedicated to proving vital lemmas and theorems for later chapters. Herein, we consider linear operators, bilinear forms, generalized equations, parametric optimization problems, and linear differential-algebraic equations.

In the second part, we derive a local minimum principle for optimal control problems with index one differential-algebraic equations and mixed control-state constraints by writing the problem as an infinite optimization problem, for which non-trivial Lagrange multipliers exist. Then, an explicit representation is derived for these multipliers, which yields a local minimum principle. The results are then applied to an optimal control problem with Hessenberg differential-algebraic equations of arbitrary order by reducing the index to one.

The third part of this thesis examines second-order sufficient conditions for optimal control problems subject to index one differential-algebraic equations and mixed control-state constraints. Herein, a Riccati equation is used to construct a quadratic function, which satisfies a Hamilton Jacobi inequality. The main task of the verification is to prove second-order sufficient conditions for a parametric optimization problems with the assumptions at hand, which is done in the first part. Analog to the second part, the results are applied to problems with Hessenberg differential-algebraic equations of arbitrary order by reducing the index.

In the last part of this thesis, we consider the implicit Euler discretization for an optimal control problem subject to an index two differential-algebraic equation in semi-explicit form and mixed control-state constraints. Typically, convergence is proven by comparing the respective Karush-Kuhn-Tucker conditions. However, there is a discrepancy between the continuous and discrete necessary conditions of optimal control problems with differential-algebraic equations of index two or higher. Hence, standard techniques fail. This was overcome by equivalently reformulating the discrete optimization problem, which has suitable Karush-Kuhn-Tucker conditions. The respective necessary conditions are then rewritten as generalized equations and a fitting convergence theorem is applied, which results in a linear convergence rate of the solution and multipliers in the essential supremum norm.

Kurzzusammenfassung

In dieser Dissertation studieren wir Optimalsteuerungsprobleme mit differential-algebraischen Gleichungen in Hessenberg-Form und gemischten Steuer- und Zustandsbeschränkungen. Wir leiten notwendige und hinreichende Bedingungen für Probleme mit Hessenberg differential-algebraischen Gleichungen mit beliebiger Ordnung her. Des weiteren untersuchen wir Konvergenzeigenschaften von einem approximierten Optimalsteuerungsproblem mit einer Index zwei differential-algebraischen Gleichung und gemischten Steuer- und Zustandsbeschränkungen.

Der erste Teil der Dissertation ist den Beweisen von grundlegenden Lemmas und Theoremen gewidmet, welche wir in den darauffolgenden Kapiteln benötigen. Hierbei betrachten wir lineare Operatoren, Bilinearformen, verallgemeinerte Gleichungen, parametrische Optimierungsprobleme und lineare differential-algebraischen Gleichungen.

Im zweiten Teil leiten wir ein lokales Minimumprinzip für Optimalsteuerungsprobleme mit einer differential-algebraischen Gleichung vom Index eins und gemischten Steuer- und Zustandsbeschränkungen her, indem wir das Problem als Infinites Optimierungsproblem betrachten, für welches nicht-triviale Lagrange-Multiplikatoren existieren. Danach leiten wir eine explizite Darstellung für diese Multiplikatoren her, wodurch wir ein lokales Minimumprinzip erhalten. Diese Resultate wenden wir dann auf ein Optimalsteuerungsprobleme mit einer differential-algebraischen Gleichung in Hessenberg-Form mit beliebiger Ordnung an, indem wir den Index der differential-algebraischen Gleichung auf eins reduzieren.

Im dritten Teil der Dissertation untersuchen wir hinreichende Bedingungen zweiter Ordnung für Optimalsteuerungsprobleme mit einer differential-algebraischen Gleichung vom Index eins und gemischten Steuer- und Zustandsbeschränkungen. Dabei verwenden wir eine Riccati-Gleichung, um eine quadratische Funktion zu konstruieren, welche eine Hamilton-Jacobi Ungleichung erfüllt. Die Hauptaufgabe der Verifikation besteht darin, hinreichende Bedingungen für ein parametrische Optimierungsproblem mit den verfügbaren Annahmen zu beweisen. Die Resultate werden dann analog zum zweiten Teil auf Probleme mit differential-algebraischen Gleichungen in Hessenberg-Form mit beliebiger Ordnung angewandt, indem der Index auf eins reduziert wird.

Im letzten Teil betrachten wir die implizite Eulerdiskretisierung für ein Optimalsteuerungsproblem mit einer Index zwei differential-algebraischen Gleichung und gemischten Steuer- und Zustandsbeschränkungen. Üblicherweise wird Konvergenz durch den Vergleich der Karush-Kuhn-Tucker Bedingungen des diskretisierten Problems mit denen des Ausgangsproblems bewiesen. Bei Probleme mit differential-algebraischen Gleichungen mit einem Index größer als eins liegt jedoch eine strukturelle Diskrepanz zwischen den jeweiligen notwendigen Bedingungen vor. Aus diesem Grund sind versuchte Konvergenzbeweise mit Standardtechniken aus der Theorie gescheitert. Die Diskrepanz wurde durch eine Umformulierung des diskretisierten Optimierungsproblems überwunden, wodurch wir geeignete Karush-Kuhn-Tucker Bedingungen erhalten haben. Die jeweiligen notwendigen Bedingungen werden dann als verallgemeinerte Gleichungen betrachtet und ein entsprechendes Konvergenz-Theorem wird angewandt. Das Resultat ist die Existenz von einer diskreten Lösung und Multiplikatoren, welche linear gegen die Lösung und Multiplikatoren des Ausgangsproblems in der essentiellen Supremumsnorm konvergieren.

Acknowledgments

First, I would like to sincerely thank my advisor Professor Dr. Matthias Gerdts for giving me the opportunity to work on the project that resulted in this thesis, for his guidance, support, and patience, and for always taking the time to discuss my studies.

I would also like to thank Professor Dr. Sabine Pickenhain for accepting to review my thesis, and for her valuable suggestions, which helped to improve this thesis.

I am very grateful to the Air Force Office of Scientific Research and the Deutsche Forschungsgemeinschaft for funding my research.

Furthermore, I would like to thank my colleagues who accompanied me through my time at the Bundeswehr University. Special thanks goes to Dr. Axel Dreves whose constructive input helped greatly in the finalization of this thesis.

Finally, I would like to express my gratitude to my family for their continuous support throughout the years, and I would like to dedicate this thesis to my late mother.

Contents

Notations

Set-related Symbols

$V \cup W$	the union of the sets V and W
$V \cap W$	the intersection of the sets V and W
$V \setminus W$	the set of elements in V, but not in W
$V \times W$	the Cartesian product of the sets V and W
V^n	the n-fold Cartesian product of the set V
$\{v\}$	the set consisting of the point v
$\mathrm{card}\,(V)$	the cardinality of the set V

Sets and Spaces

$\mathbb{N}$	set of natural numbers
$J \subset \mathbb{N}$	finite set of indexes, ordered from smallest to largest
$\mathbb{R}$	set of real numbers
$\mathbb{R}_+$	set of non-negative real numbers
$\lvert a \rvert$	absolute value of $a \in \mathbb{R}$
$\mathbb{R}^n$	n-dimensional Euclidean space with norm $\lVert \cdot \rVert$
$[a, b]$	compact time interval in $\mathbb{R}$ with fixed $a < b$
$\mathbb{G}_N$	grid with $N \in \mathbb{N}$ subintervals
X, Y, Z	Banach spaces
$\mathbf{0}_X$	zero of a Banach space X
$\mathbf{0}$	generic zero element of some space
$\lVert \cdot \rVert_X$	norm on a Banach space X
$\langle \cdot, \cdot \rangle_X$	inner product on a pre-Hilbert space X
$\mathcal{B}_\rho\,(x)$	the closed ball with radius ρ and center x
$\mathfrak{L}(X, Y)$	set of all linear, continuous operators from X to Y
$\lVert T \rVert_{\mathfrak{L}(X,Y)} = \sup\limits_{x \in X \setminus \{\mathbf{0}_X\}} \frac{\lVert Tx \rVert_Y}{\lVert x \rVert_X}$	norm on $\mathfrak{L}(X, Y)$
$X^\star = \mathfrak{L}(X, \mathbb{R})$	topological dual space of X
K^*	positive dual cone of K
$\lVert x^\star \rVert_{X^\star} = \lVert x^\star \rVert_{\mathfrak{L}(X,\mathbb{R})}$	norm on $X^\star$ defining the strong topology
$L_p^n\,([a, b])$	space of equivalence classes, which consist of measurable functions $f : [a, b] \to \mathbb{R}^n$ that are bounded in the norm $\lVert \cdot \rVert_p$
$W_{q,p}^n\,([a, b])$	Sobolev space of all absolutely continuous functions $f : [a, b] \to \mathbb{R}^n$ that are bounded in the norm $\lVert \cdot \rVert_{q,p}$
$\mathcal{C}_0^n\,([a, b])$	space of continuous functions $f : [a, b] \to \mathbb{R}^n$
$\mathcal{C}_1^n\,([a, b])$	space of continuously differentiable functions $f : [a, b] \to \mathbb{R}^n$

Mappings

$T : X \to Y$	a mapping from X to Y
$\mathrm{im}(T) := \{T(x) \mid x \in X\}$	image of X under the map $T : X \to Y$
$\ker(T) := \{x \in X \mid T(x) = \mathbf{0}_Y\}$	kernel or null space of a linear map $T : X \to Y$
T^{-1}	inverse mapping of $T : X \to Y$
$T^{\star}$	adjoint operator of the linear map $T : X \to Y$
$T'(x)$	Fréchet derivative of F at x
$T'_x(x,y) = \frac{\partial T(x,y)}{\partial x}$	partial Fréchet derivative of T at (x,y)
$\nabla T(x) = T'(x)^{\top} \in \mathbb{R}^n$	gradient of $T : \mathbb{R}^n \to \mathbb{R}$
$\nabla_x T(x,y) = T'_x(x,y)^{\top}$	partial gradient of $T : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$
$\nabla^2_{xy} T(x,y) = \frac{\partial^2 T(x,y)}{\partial x \partial y} \in \mathbb{R}^{n\times m}$	second partial Fréchet derivative of $T : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$
$F : X \rightrightarrows Y$	a set-valued mapping from X to Y
$\mathrm{graph}\,(F) := \{(x,y) \in X \times Y \mid y \in F(x)\}$	the graph of a set-valued mapping $F : X \rightrightarrows Y$
$\mathcal{H}$	Hamilton function of an optimal control problem
$\mathcal{L}$	Lagrange function
$\exp(x) = e^x$	the real exponential function

Vectors and Matrices

$v \in \mathbb{R}^n$	a column vector in $\mathbb{R}^n$
$v^{\top}$	transpose of a vector v
$\lVert v \rVert$	Euclidean norm of a vector v
$A \in \mathbb{R}^{n\times m}$	a matrix with n rows and m columns
$A^{\top} \in \mathbb{R}^{m\times n}$	transpose of a matrix $A \in \mathbb{R}^{n\times m}$
A^{-1}	inverse of a matrix $A \in \mathbb{R}^{n\times n}$
$A^{\curlywedge} = A^{\top}\left(AA^{\top}\right)^{-1}$	right inverse of a matrix A with full row rank
$\mathrm{rank}\,(A)$	rank of a matrix A
$\mathbf{I}_n$	unit matrix in $\mathbb{R}^{n\times n}$
$\mathrm{diag}\,[a_j]_{j\in J}$	diagonal matrix with entries $a_j \in \mathbb{R}$, ordered from smallest to largest index
$\mathbf{0}_{n\times m}$	matrix with only zero entries of dimension $n \times m$
$\lVert A \rVert$	spectral norm of a matrix $A \in \mathbb{R}^{n\times m}$

Chapter 1

Introduction

The study of optimal control has its origins in the theory of variational problems. Research for variational problems started to increase in 1696, when Johann Bernoulli (1667–1748) first proposed the Brachistochrone problem, and has since then been broadened by several renowned mathematicians, including Leonhard Euler (1707–1793), Ludovico Lagrange (1736–1813), Andrien Legendre (1752–1833), Carl Jacobi (1804–1851), and William Hamilton (1805–1865).
By separating state and control variables and permitting control constraints variational problems were generalized to optimal control problems. Herein, the behavior of the state is described by dynamic equations, typically as a system of ordinary or partial differential equations (ODEs or PDEs). Usually, the dynamic behavior is influenced through the choice of control variables, which in turn are often subject to constraints, such as physical limitations. The objective of optimal control is to find a control such that all constraints are satisfied, and a particular optimality criterion is obtained.
Since the 20^{th} century optimal control theory has been applied to numerous fields, for instance, aerospace, biological engineering, economics, and robotics. An example of an optimal control problem would be an autonomously driven car, where the dynamic behavior is given by some equations of motion, which can be controlled by, e.g., steering, accelerating, and braking. Such a system is typically restricted by the boundaries of the road, as well as the maximum steering angle and acceleration. The objective could be to minimize the travel time from start to destination or the fuel consumption.
Many dynamic behaviors in, e.g., electronics, mechanics, and process engineering are described by *differential-algebraic equations* (DAEs), which have been investigated more thoroughly since the 1970's. DAEs in their most general form are implicit differential equations

$$\mathbf{0}_{\mathbb{R}^{n_z}} = F\left(t, z(t), \dot{z}(t)\right), \quad t \in [t_0, t_f], \tag{1.1}$$

where z is called state variable. This type of equation consists of differential equations and algebraic equations, if the partial derivative $F'_{\dot{z}}(\cdot)$ is singular. Otherwise, the implicit function theorem can be applied to solve for $\dot{z}$, resulting in an explicit ODE. A special class of DAEs are the so-called *semi-explicit* DAEs, where the state variable z is decomposed into the *differential variable* x and *algebraic variable* y, and (1.1) into differential and algebraic equations:

$$\begin{aligned} \dot{x}(t) &= f\left(t, x(t), y(t)\right), \quad t \in [t_0, t_f], \\ \mathbf{0}_{\mathbb{R}^{n_y}} &= g\left(t, x(t), y(t)\right), \quad t \in [t_0, t_f]. \end{aligned} \tag{1.2}$$

The concept of an *index* was introduced for DAEs in order to measure its degree of regularity. There are numerous index definitions, such as the differentiation index, the perturbation index, the strangeness index, and the tractability index (cf. [20, 38, 42, 43, 66, 74–76]). In (1.2), let f be uniformly Lipschitz continuous with respect to x and y, and g be continuously differentiable in sufficiently large convex compact sets. The DAE (1.2) has index one, if the partial derivative $g_y'(\cdot)$ is non-singular, and $g_y'(\cdot)$, $g_y'(\cdot)^{-1}$ are bounded. Then, the algebraic equation in (1.2) is implicitly solvable for y. DAEs are called higher index DAEs, if the index is greater than one. In that case, the algebraic equation does not depend on all components of the algebraic variable y. For the DAE

$$\begin{array}{rcll} \dot{x}(t) &=& f\left(t, x(t), y(t)\right), & t \in [t_0, t_f], \\ \mathbf{0}_{\mathbb{R}^{n_y}} &=& g\left(t, x(t)\right), & t \in [t_0, t_f], \end{array} \tag{1.3}$$

let f be continuously differentiable, and let g be twice continuously differentiable in sufficiently large convex compact sets. Then, (1.3) has index two, if $g_x'(\cdot)\, f_y'(\cdot)$ is non-singular, and $g_x'(\cdot)\, f_y'(\cdot)$, $\left(g_x'(\cdot)\, f_y'(\cdot)\right)^{-1}$ are bounded. In Chapter 3 and Chapter 4, we consider specially structured *Hessenberg* DAEs of order $k > 2$

$$\begin{array}{rcll} \dot{x}_1(t) &=& f_1\left(t, x_1(t), x_2(t), \dots, x_{k-1}(t), y(t)\right), & t \in [t_0, t_f], \\ \dot{x}_2(t) &=& f_2\left(t, x_1(t), x_2(t), \dots, x_{k-1}(t)\right), & t \in [t_0, t_f], \\ \dot{x}_3(t) &=& f_3\left(t, x_2(t), x_3(t), \dots, x_{k-1}(t)\right), & t \in [t_0, t_f], \\ & \vdots & & \\ \dot{x}_{k-1}(t) &=& f_{k-1}\left(t, x_{k-2}(t), x_{k-1}(t)\right), & t \in [t_0, t_f], \\ \mathbf{0}_{\mathbb{R}^{n_y}} &=& g\left(t, x_{k-1}(t)\right), & t \in [t_0, t_f]. \end{array} \tag{1.4}$$

Herein, $(x_1, x_2, \dots, x_{k-1})$ is the differential variable and y is the algebraic variable. Let the functions f_i be i-times continuously differentiable for $i = 1, \dots, k-1$, and g be k-times continuously differentiable in sufficiently large convex compact sets. Then, (1.4) has index k, if

$$E(\cdot) := g_{x_{k-1}}'(\cdot)\, f_{k-1,x_{k-2}}'(\cdot)\, f_{k-2,x_{k-3}}'(\cdot) \cdots f_{2,x_1}'(\cdot)\, f_{1,y}'(\cdot)$$

is non-singular, and $E(\cdot)$, $E(\cdot)^{-1}$ are bounded.
In this thesis, we analyze optimal control problems of the following type:

Problem 1.1 (DAE Optimal Control Problem)
Minimize the objective function

$$\varphi\left(x(0), x(1)\right)$$

with respect to

$$x : [0,1] \to \mathbb{R}^{n_x}, \quad y : [0,1] \to \mathbb{R}^{n_y}, \quad u : [0,1] \to \mathbb{R}^{n_u},$$

subject to the semi-explicit DAE

$$\begin{array}{rcl} \dot{x}(t) &=& f\left(x(t), y(t), u(t)\right), \\ \mathbf{0}_{\mathbb{R}^{n_y}} &=& g\left(x(t), y(t), u(t)\right), \end{array}$$

the boundary condition

$$\mathbf{0}_{\mathbb{R}^{n_\psi}} = \psi\left(x(0), x(1)\right),$$

and the mixed control-state constraint

$$\mathbf{0}_{\mathbb{R}^{n_c}} \geq c\left(x(t), y(t), u(t)\right).$$

Herein, the inequality is considered componentwise, i.e., $0 \geq c_j\left(x(t), y(t), u(t)\right)$ for $j = 1, \ldots, n_c$. Since we also consider inequality constraints in Problem 1.1, we will later combine the index notions above for the algebraic equation $\mathbf{0}_{\mathbb{R}^{n_y}} = g\left(x(t), y(t), u(t)\right)$ with the regularity of the mixed control-state constraints $\mathbf{0}_{\mathbb{R}^{n_c}} \geq c\left(x(t), y(t), u(t)\right)$. Moreover, through several transformation techniques it is possible to transform more general optimal control problems to the form of Problem 1.1. For instance, problems with free final time, non-autonomous problems, and problems with integral objective functional can be transformed. In order to properly define Problem 1.1, we introduce Lebesgue and Sobolev spaces.

Definition 1.2 (Lebesgue and Sobolev Spaces)
Let $a, b \in \mathbb{R}$ with $a < b$.

- *For $1 \leq p < \infty$ the Lebesgue space $L_p\left([a,b]\right)$ is the Banach space of all equivalence classes, which consist of measurable functions $v : [a,b] \to \mathbb{R}$ (compare Definition A.8) with*

$$\int_a^b |v(t)|^p \, dt < \infty,$$

  i.e., the p-th power of the absolute value of $v(\cdot)$ is Lebesgue integrable on $[a,b]$. The Lebesgue space $L_\infty\left([a,b]\right)$ is the Banach space of all equivalence classes, which consist of measurable functions $v : [a,b] \to \mathbb{R}$ (compare Definition A.8) with

$$\underset{t \in [a,b]}{\text{ess sup}} |v(t)| < \infty,$$

  i.e., $v(\cdot)$ is essentially bounded on $[a,b]$. Two functions are in the same equivalence class, if they are equal almost everywhere on $[a,b]$ in terms of the Lebesgue measure.

- *For $k \in \mathbb{N}$ and $1 \leq p \leq \infty$ the Sobolev space $W_{k,p}\left([a,b]\right)$ consists of absolutely continuous functions $v : [a,b] \to \mathbb{R}$ with absolutely continuous derivatives up to order $k-1$ and*

$$\frac{d^k}{dt^k} v \in L_p\left([a,b]\right),$$

  where $\frac{d^k}{dt^k} v$ is the weak derivative of order k.

- *For $1 \leq p \leq \infty$ and $k, n \in \mathbb{N}$ the spaces $L_p^n\left([a,b]\right)$ and $W_{k,p}^n\left([a,b]\right)$ are the product spaces*

$$\begin{aligned} L_p^n\left([a,b]\right) &:= L_p\left([a,b]\right) \times \cdots \times L_p\left([a,b]\right) \\ W_{k,p}^n\left([a,b]\right) &:= W_{k,p}\left([a,b]\right) \times \cdots \times W_{k,p}\left([a,b]\right) \end{aligned}$$

  equipped with the norms

$$\|v\|_p := \left(\int_a^b \|v(t)\|^p \, dt \right)^{\frac{1}{p}}, \quad 1 \leq p < \infty, \qquad \|v\|_\infty := \underset{t \in [a,b]}{\text{ess sup}} \|v(t)\|,$$

$$\|v\|_{k,p} := \max\left\{ \|v\|_p, \left\| \frac{dv}{dt} \right\|_p, \ldots, \left\| \frac{d^k v}{dt^k} \right\|_p \right\}, \quad 1 \leq p \leq \infty.$$

For Problem 1.1, we consider the differential state x to be in the Sobolev space $W^{n_x}_{1,\infty}([0,1])$, and the algebraic state y and control u to be in the Lebesgue spaces $L^{n_y}_{\infty}([0,1])$ and $L^{n_u}_{\infty}([0,1])$, respectively. Occasionally, we consider matrix valued functions $A \in L^{n\times m}_{\infty}([a,b])$ bounded in the norm $\|A\|_{\infty} := \operatorname*{ess\,sup}_{t\in[a,b]} \|A(t)\|$, in which case $\|\cdot\|$ is the spectral norm. It will be clear from the context whether the Euclidean or the spectral norm is used.

In optimal control one can distinguish between *weak local minimizer* and *strong local minimizer*:

Definition 1.3 (Weak and Strong Local Minimizer)
Suppose $(\hat{x},\hat{y},\hat{u}) \in W^{n_x}_{1,\infty}([0,1]) \times L^{n_y}_{\infty}([0,1]) \times L^{n_u}_{\infty}([0,1])$ is feasible for Problem 1.1. Then, $(\hat{x},\hat{y},\hat{u})$ is called a

- *(weak) local minimizer of Problem 1.1, if*

$$\varphi(x(0),x(1)) \geq \varphi(\hat{x}(0),\hat{x}(1))$$

  for all admissible (x,y,u) with

$$\|x-\hat{x}\|_{1,\infty} < \rho, \quad \|y-\hat{y}\|_{\infty} < \rho, \quad \|u-\hat{u}\|_{\infty} < \rho$$

  for some $\rho > 0$.

- *strong local minimizer of Problem 1.1, if*

$$\varphi(x(0),x(1)) \geq \varphi(\hat{x}(0),\hat{x}(1))$$

  for all admissible (x,y,u) with

$$\|x-\hat{x}\|_{\infty} < \rho$$

  for some $\rho > 0$.

Note that a strong local minimizer is also a weak local minimizer, since strong local minimizer are optimal on a larger set of algebraic states and controls. In this thesis, we only consider *weak local minimizer* and aim to derive the following for optimal control problems subject to a semi-explicit DAE, a boundary condition, and a mixed control-state constraint:

(a) Necessary conditions in form of a local minimum principle for problems with index one DAEs and Hessenberg DAEs of arbitrary order. (Theorem 3.1.15 / Theorem 3.2.5)

(b) Second-order sufficient conditions for problems with index one DAEs and Hessenberg DAEs of arbitrary order. (Theorem 4.1.12 / Theorem 4.2.2)

(c) Conditions such that there exist a solution of an approximation of the optimal control problem with an index two DAE that converges to a solution of the continuous optimal control problem. (Theorem 5.5.6)

Necessary conditions, also called *maximum principles* or *minimum principles*, for optimal control problems have been investigated since the 1950's. Early proofs of the maximum principle can be found in Pontryagin et al. [104] and Hestenes [57]. In [96, 124], optimal control problems subject to ordinary differential equations with mixed control-state constraints have been analyzed. Problems with pure state constraints are discussed in, e.g., [59–61, 63, 86, 87, 89, 90]. In more recent years, the research has been expanded to optimal control problems subject to DAEs. In [8, 65, 94], linear quadratic DAE optimal control problems are discussed. Herein, [94] consider descriptor systems with constant coefficient matrices, whereas time-variant systems are considered in [65]. In [8], nonlinear quasi-linear DAEs are examined. Optimal control problems with nonlinear index one DAEs in semi-explicit form are inspected in [29, 47, 102], where [102] consider set constraints on the controls, and [29, 47] consider pure state and mixed control-state constraints. Necessary conditions for problems with higher index DAEs were derived in [45, 47, 83, 111]. In [45, 47], Index two DAEs with pure state constraints, mixed control-state constraints, and set constraints on the controls are considered. Problems with Hessenberg DAEs up to index three are analyzed in [111], and in [83], problems with Hessenberg DAEs of arbitrary order are investigated. By reducing the optimal control problem to an equivalent nonsmooth variational problem, a maximum principle for problems with implicit control systems is derived in [28]. In [67], general unstructured DAE optimal control problems are studied. [69, 88] establish necessary conditions for infinite optimization problems, which are closely related to optimal control problems.

In Chapter 3 we expand the research on necessary conditions for optimal control problems. Specifically, Theorem 3.1.15 generalizes [47, Theorem 3.4.4], and Theorem 3.2.5 generalizes [47, Theorem 3.3.8] by weakening the assumptions and considering Hessenberg DAEs of arbitrary order. Moreover, [83, Theorem 3.1] is generalized by including boundary conditions and weakening the assumptions.

Sufficient conditions can be utilized in order to verify, if a solution of local minimum principle is a weak local minimizer. Results on sufficient conditions have primarily been established for optimal control problems subject to explicit ODEs. In [21, 31, 98, 123], problems with control constraints are analyzed. Mixed control-state constraints are considered in [92, 93]. In [13, 79, 82, 99, 100], optimal control problems subject to mixed control-state constraints and pure state constraints are investigated, where [13] considers multiple pure state constraints of arbitrary order. Problems with free final time have been discussed in [21, 58, 93]. In [14, 15], sufficient conditions for strong local minimizer were derived.
By adding boundary conditions to the optimal control problem we expand the results of [83, Theorem 4.1] in Theorem 4.1.12/ Theorem 4.2.2. The results in [92] are also generalized, since DAEs are included in Problem 3.1.1/ Problem 3.2.1.

In general, optimal control problems are not analytically solvable. Therefore, numerically methods are used in order to obtain an approximated solution. From a theoretical point of view, it is of interest to find conditions that guarantee convergence of the approximated solution. Herein, the type and rate of convergence depends on the discretization method and the problem itself.

The most commonly used approximation method is the (explicit or implicit) *Euler discretization*. It was used for optimal control problems with mixed control-state constraints in [48,80]. Linear convergence in the L_∞-norm is obtained in [80], whereas convergence with a rate of $\frac{1}{p}$ in the L_p-norm is achieved in [48] for optimal controls of bounded variation. In [16,32,34], problems with pure state constraints of order one are investigated. Herein, [32,34] prove linear convergence in the L_2-norm and convergence of order $\frac{2}{3}$ in the L_∞-norm, while [16] obtain linear convergence in the L_∞-norm. Problems with DAEs are considered in [84,85]. A Linear convergence rate in the L_∞-norm for problems with index one DAEs is achieved in [84]. Problems with index two DAEs and mixed control-state constraints are examined in [85], and linear convergence in the L_∞-norm is proven. In order to obtain a higher rate of convergence Runge-Kutta methods are utilized. In [33,53,117], they are applied to problems with set constraints on the control, where [33,117] achieve a quadratic convergence rate for a second order Runge-Kutta approximation, and [53] obtain convergence of arbitrary rate with a suitable Runge-Kutta scheme. In [73], convergence for the value of the objective function is proven. In [3–7,101,113,118], problems with discontinuous (bang-bang type) controls are inspected. Linear problems are discussed in [4,101,118], where [4] obtain linear convergence in the L_1-norm, and of rate $\frac{1}{2}$ in the L_2-norm for the control. In [3,5,6,113], linear quadratic systems are investigated. Herein, [3] achieve corresponding results to [4], whereas [6] obtain a linear convergence rate. In [7], nonlinear optimal control problems, where the control appears linearly, are considered.
Theorem 5.5.6 generalizes the results in [80] by including index two DAEs. We expand the research on convergence analysis for optimal control problems, and establish a technique that deals with the discrepancy between the necessary conditions of the problem and its approximation. This method could also be used for problems with DAEs of higher index. Additionally, the techniques in Chapter 5 are also applicable to problems subject to index one DAEs and mixed control-state constraints.

The thesis is structured as follows:
In Chapter 2, we gather fundamental results and definitions. The aim is to shorten the technical and repetitive proofs in later chapters by considering more general settings such that the results can be applied to various problems. The analysis includes properties of linear operators and bilinear forms, regularity of generalized equations, sufficient conditions and sensitivity for finite-dimensional parametric optimization problems, and characteristics of linear time-variant DAEs. In particular, approximation conditions are derived for bilinear forms, generalized equations, and linear time-variant DAEs.
Necessary conditions for optimal control problems subject to semi-explicit DAEs and mixed control-state constraints are derived in Chapter 3. Therein, we first consider index one DAEs and obtain a local minimum principle under regularity and controllability assumptions. Then, we apply the results to optimal control problems with Hessenberg DAEs of arbitrary order via index-reduction.

In Chapter 4, we provide second-order sufficient conditions using a *Hamilton Jacobi inequality*. Similar to Chapter 3, we first discuss the index one case and then DAEs of higher index. With aid of appropriate *Riccati equations*, suitable quadratic functions are constructed, which satisfy the respective Hamilton Jacobi inequalities.

Chapter 5 is dedicated to proving convergence of approximations of optimal control problems subject to semi-explicit index two DAEs and mixed control-state constraints. The standard scheme for proving convergence is to compare the continuous and discrete necessary conditions, and apply a suitable approximation result. However, this method fails for higher index DAEs, since there is a *structural discrepancy* between the respective necessary conditions. This was overcome by finding an equivalent reformulation of the discretized optimal control problem. Convergence is then proven for states and control, as well as the associated Lagrange multipliers of the transformed discrete problem. Additionally, a relationship between the multipliers of the modified discrete problem and the directly discretized problem is established.

In Chapter 6, we summarize the main results, and give a perspective of open questions and future research topics that arose from this thesis.

In the Appendix, we collect some definitions and auxiliary statements for the other chapters, which would have disturbed the reading flow.

In the following chapters, we frequently use the abbreviation $F[t]$ for functions of type $F(z(t))$ in order to simplify notation. Usually, these functions will be evaluated at a minimizer (or Karush-Kuhn-Tucker point), e.g.,

$$f[t] := f(\hat{x}(t), \hat{y}(t), \hat{u}(t)), \quad g'_x[t] := g'_x(\hat{x}(t), \hat{y}(t), \hat{u}(t)).$$

Chapter 2

Fundamental Results

The intention of this chapter is to collect fundamental definitions and statements, which are significant for later chapters, in one place, thus improving the reading flow. In Section 2.1, we gather some properties of linear operators and bilinear forms on Hilbert spaces. The main result of this section is Theorem 2.1.10, which states that approximations of bilinear forms inherit a particular coercivity property under appropriate assumptions. In Section 2.2, we introduce the concept of generalized equations and prove a convergence result, which is essential for Chapter 5. Section 2.3 deals with finite-dimensional parametric optimization problems. Herein, we provide, among other things, second-order sufficient conditions and a sensitivity result, which prove to be crucial for Chapter 4 and Section 5.5, respectively. Finally, in Section 2.4 we examine linear time-variant DAEs. In particular, we investigate the connection between time-continuous and time-discrete systems, using the results of Section 2.1.

2.1 Linear Operators and Bilinear Forms

Definitions and statements on Banach spaces, Hilbert spaces, and linear operators can be found in, e.g., [1, 51, 72, 107, 120].

Let X be a Hilbert space with the inner product $\langle \cdot, \cdot \rangle_X$, and the induced norm $\|\cdot\|_X$ given by $\|x\|_X = \sqrt{\langle x, x \rangle_X}$.

Definition 2.1.1 (Dual Space)

For a normed vector space X the dual space $X^\star$ is the set of all linear, continuous functionals from X to $\mathbb{R}$, i.e., $X^\star = \mathfrak{L}(X, \mathbb{R})$.

Remark 2.1.2

For $p, q \in \mathbb{R}$ with $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$ the dual space of $L_p([a,b])$ can be (isometric isomorphically) identified with $L_q([a,b])$, i.e., for every $f \in (L_p([a,b]))^\star$ exists a unique equivalence class in $L_q([a,b])$ such that for all elements v_f of the equivalence class it holds $\|f\|_{(L_p([a,b]))^\star} = \|v_f\|_q$ and

$$f(u) = \int_a^b v_f(t)\, u(t)\, dt \quad \text{for all} \;\; u \in L_p([a,b]).$$

Moreover, the dual space of $L_1([a,b])$ can be identified with $L_\infty([a,b])$. However, the dual space of $L_\infty([a,b])$ cannot be identified with $L_1([a,b])$, since it also contains so called finitely additive measures (cf. [41, Proposition 7.16], [112, Theorem 6.19], [121]).

The following well-known theorem describes the relation between a Hilbert space and its dual:

Theorem 2.1.3 (Fréchet-Riesz Theorem, [120, Theorem V.3.6])
Let X be a Hilbert space. Then, the linear map $\mathcal{I}_X : X \to X^\star$ defined by $\mathcal{I}_X x := \langle \cdot, x \rangle_X$ has the following properties:

(i) $\mathcal{I}_X$ is bijective.

(ii) $\mathcal{I}_X$ is a linear isometry, i.e., $\|\mathcal{I}_X x\|_{X^\star} = \|x\|_X$ for all $x \in X$.

The mapping $\mathcal{I}_X : X \to X^\star$ is called *canonical isomorphism* between X and $X^\star$. It follows straightforwardly from condition *(ii)* that $\|\mathcal{I}_X\|_{\mathfrak{L}(X,X^\star)} = \left\|\mathcal{I}_X^{-1}\right\|_{\mathfrak{L}(X^\star,X)} = 1$.
In the sequel, we examine linear, continuous operators $T : X \to Y$ between Hilbert spaces and their *adjoint operators* $T^\star : Y^\star \to X^\star$. In particular, we aim to derive conditions for the existence of a *generalized right inverse*. To that end, we require the following notion of a strengthened surjectivity property:

Definition 2.1.4 (Uniform Surjectivity)
Let X, Y be Banach spaces. A mapping $T \in \mathfrak{L}(X, Y)$ is called uniformly surjective, if the following holds:
There exists a constant $\kappa > 0$ such that for all $y \in Y$ there exists $x \in X$ with

$$\begin{aligned} Tx &= y\,, \\ \kappa \|x\|_X &\leq \|y\|_Y\,. \end{aligned}$$

Note that in case of finite dimensional spaces this definition is equivalent to standard surjectivity.
We recall that the adjoint operator $T^\star : Y^\star \to X^\star$ of the linear map $T : X \to Y$ is defined by

$$(T^\star y^\star)(x) := y^\star(Tx) \quad \text{for } y^\star \in Y^\star, x \in X, \tag{2.1.1}$$

which allows us to prove the following using Theorem 2.1.3:

Lemma 2.1.5
Let X, Y be Hilbert spaces. Suppose $T \in \mathfrak{L}(X, Y)$ is uniformly surjective with constant κ. Then, the adjoint operator $T^\star$ is injective and satisfies

$$\|T^\star y^\star\|_{X^\star} \geq \kappa \|y^\star\|_{Y^\star} \quad \textit{for all } y^\star \in Y^\star\,.$$

Proof. Let $y^\star \in Y^\star \setminus \{\mathbf{0}_{Y^\star}\}$ be arbitrary. According to Theorem 2.1.3, for the element $v_y := \mathcal{I}_Y^{-1} y^\star \in Y \setminus \{\mathbf{0}_Y\}$ it holds

$$y^\star(\cdot) = \langle \cdot, v_y \rangle_Y\,, \tag{2.1.2}$$

$$\|y^\star\|_{Y^\star} = \|v_y\|_Y\,. \tag{2.1.3}$$

Then, the uniform surjectivity of T assures the existence of $x_v \in X \setminus \{\mathbf{0}_X\}$ with

$$Tx_v = v_y\,, \tag{2.1.4}$$

$$\|x_v\|_X \leq \frac{1}{\kappa} \|v_y\|_Y\,. \tag{2.1.5}$$

It follows

$$\begin{aligned}
\|T^\star y^\star\|_{X^\star} &= \sup_{x\in X\setminus\{\mathbf{0}_X\}} \frac{|(T^\star y^\star)(x)|}{\|x\|_X} \overset{(2.1.1)}{=} \sup_{x\in X\setminus\{\mathbf{0}_X\}} \frac{|y^\star(Tx)|}{\|x\|_X} \\
&\overset{(2.1.2)}{=} \sup_{x\in X\setminus\{\mathbf{0}_X\}} \frac{|\langle Tx, v_y\rangle_Y|}{\|x\|_X} \geq \frac{|\langle Tx_v, v_y\rangle_Y|}{\|x_v\|_X} \\
&\overset{(2.1.4)}{=} \frac{|\langle v_y, v_y\rangle_Y|}{\|x_v\|_X} \overset{(2.1.5)}{\geq} \kappa \frac{|\langle v_y, v_y\rangle_Y|}{\|v_y\|_Y} \\
&= \kappa \|v_y\|_Y \overset{(2.1.3)}{=} \kappa \|y^\star\|_{Y^\star},
\end{aligned}$$

which implies $T^\star y^\star \neq 0$ for all $y^\star \in Y^\star \setminus \{\mathbf{0}_{Y^\star}\}$. Thus, $\ker(T^\star) = \{\mathbf{0}_{Y^\star}\}$, which is equivalent to $T^\star$ being injective. □

Consequently, if Y is finite dimensional, the following holds:

Lemma 2.1.6

Let X, Y be Hilbert spaces with $\dim(Y) \leq \dim(X)$, and $\dim(Y) < \infty$. Suppose $T \in \mathfrak{L}(X, Y)$ is uniformly surjective with constant κ. Then, the mapping $S := T \circ \mathcal{I}_X^{-1} \circ T^\star \in \mathfrak{L}(Y^\star, Y)$ is bijective and the inverse is bounded by

$$\left\|S^{-1}\right\|_{\mathfrak{L}(Y,Y^\star)} \leq \frac{1}{\kappa^2}.$$

Proof. Let $y^\star \in Y^\star \setminus \{\mathbf{0}_{Y^\star}\}$ be arbitrary. Then, due to Lemma 2.1.5, it holds

$$\begin{aligned}
&T^\star y^\star \neq \mathbf{0}_{X^\star}, \\
&\|T^\star y^\star\|_{X^\star} \geq \kappa \|y^\star\|_{Y^\star}. \qquad (2.1.6)
\end{aligned}$$

In addition, Theorem 2.1.3 implies $u_{T_y} := \mathcal{I}_X^{-1}(T^\star y^\star) \in X \setminus \{\mathbf{0}_X\}$ satisfies

$$(T^\star y^\star)(\cdot) = \langle \cdot, u_{T_y}\rangle_X, \qquad (2.1.7)$$

$$\|T^\star y^\star\|_{X^\star} = \|u_{T_y}\|_X, \qquad (2.1.8)$$

and therefore

$$Sy^\star = \left(T \circ \mathcal{I}_X^{-1} \circ T^\star\right)(y^\star) = T u_{T_y}. \qquad (2.1.9)$$

Exploiting the inequality $y^\star(y) \leq \|y^\star\|_{Y^\star}\|y\|_Y$ leads to

$$\begin{aligned}
\|y^\star\|_{Y^\star}\|Sy^\star\|_Y &\geq y^\star(Sy^\star) \overset{(2.1.9)}{=} y^\star(Tu_{T_y}) \overset{(2.1.1)}{=} (T^\star y^\star)(u_{T_y}) \\
&\overset{(2.1.7)}{=} \langle u_{T_y}, u_{T_y}\rangle_X = \|u_{T_y}\|_X^2 \overset{(2.1.8)}{=} \|T^\star y^\star\|_{X^\star}^2 \\
&\overset{(2.1.6)}{\geq} \kappa^2 \|y^\star\|_{Y^\star}^2.
\end{aligned}$$

Dividing with respect to $\|y^\star\|_{Y^\star} > 0$ yields

$$\|Sy^\star\|_Y \geq \kappa^2 \|y^\star\|_{Y^\star} > 0, \qquad (2.1.10)$$

which implies $Sy^\star \neq \mathbf{0}_Y$ for all $y^\star \in Y^\star \setminus \{\mathbf{0}_{Y^\star}\}$. Hence, $\ker(S) = \{\mathbf{0}_{Y^\star}\}$, which proves the injectivity of S. Utilizing $\dim(Y^\star) = \dim(Y) < \infty$ and the rank-nullity theorem results in

$$\dim(Y^\star) = \dim(\operatorname{im}(S)) + \dim(\ker(S)) = \dim(\operatorname{im}(S)) = \dim(Y)\ ,$$

and therefore $\operatorname{im}(S) = Y$. Thus, S is surjective and S^{-1} exists. Finally, for an arbitrary $y \in Y \setminus \{\mathbf{0}_Y\}$ we conclude

$$\|y\|_Y = \left\|S\left(S^{-1}y\right)\right\|_Y \overset{(2.1.10)}{\geq} \kappa^2 \left\|S^{-1}y\right\|_{Y^\star}\ .$$

Dividing by $\kappa^2 \|y\|_Y$ and taking the supremum with respect to $y \in Y \setminus \{\mathbf{0}_Y\}$ yields the bound

$$\left\|S^{-1}\right\|_{\mathfrak{L}(Y,Y^\star)} \leq \frac{1}{\kappa^2},$$

which completes the proof. □

Remark 2.1.7 (Generalized Right Inverse)
In Lemma 2.1.6, we have shown that $\left(T \circ \mathcal{I}_X^{-1} \circ T^\star\right)^{-1}$ *exists and is bounded by* $\frac{1}{\kappa^2}$*, if* T *is uniformly surjective and* Y *is finite dimensional. Therefore, the generalized right inverse*

$$R := \mathcal{I}_X^{-1} \circ T^\star \circ \left(T \circ \mathcal{I}_X^{-1} \circ T^\star\right)^{-1} \in \mathfrak{L}(Y,X)\,, \tag{2.1.11}$$
$$\text{with} \quad (T \circ R)(y) = y \;\; \text{for all} \;\; y \in Y$$

exists and is bounded by $\frac{\|T\|_{\mathfrak{L}(X,Y)}}{\kappa^2}$*, since* $\|T\|_{\mathfrak{L}(X,Y)} = \|T^\star\|_{\mathfrak{L}(Y^\star,X^\star)}$ *(cf. [120, Satz III.4.2]) and* $\left\|\mathcal{I}_X^{-1}\right\|_{\mathfrak{L}(X^\star,X)} = 1$.

Next, we study *bilinear forms* $\mathcal{P} : X \times X \to \mathbb{R}$ on Hilbert spaces, for which we introduce the following notions:

Definition 2.1.8 (Continuity, Symmetry, Coercivity)
Let X *be a Hilbert space,* $\mathcal{P} : X \times X \to \mathbb{R}$ *a bilinear form, and* $U \subseteq X$ *a subset.*

(i) $\mathcal{P}$ *is called continuous, if there exists a constant* $\Gamma_{\mathcal{P}} \geq 0$ *such that for all* $x_1, x_2 \in X$ *it holds*

$$|\mathcal{P}(x_1,x_2)| \leq \Gamma_{\mathcal{P}} \|x_1\|_X \|x_2\|_X\ .$$

(ii) $\mathcal{P}$ *is called symmetric, if for all* $x_1, x_2 \in X$ *it holds* $\mathcal{P}(x_1,x_2) = \mathcal{P}(x_2,x_1)$.

(iii) $\mathcal{P}$ *is called (uniformly) coercive on* U*, if there exists a constant* $\gamma > 0$ *such that for all* $u \in U$ *it holds*

$$\mathcal{P}(u,u) \geq \gamma \|u\|_X^2\ .$$

If a bilinear form is continuous, symmetric, and (uniformly) coercive on U, then it was shown in [88, Lemma 5.5] that the coercivity can be expanded to an even larger set than U:

Theorem 2.1.9 (Maurer, Zowe)
Let X be a Hilbert space and $\mathcal{P} : X \times X \to \mathbb{R}$ be a continuous, symmetric bilinear form. Furthermore, suppose $\mathcal{P}$ is coercive on a subset $U \subseteq X$ with constant γ. Then, there exist $\gamma_0, \Gamma_U > 0$ such that

$$\mathcal{P}(u+x, u+x) \geq \gamma_0 \|u+x\|_X^2$$

is satisfied for all $u \in U, x \in X$ with $\|x\|_X \leq \Gamma_U \|u\|_X$.

For a parameter $h \in (0, \infty)$ we consider the finite dimensional, closed subspaces $X_h \subseteq X$ and $Y_h \subseteq Y$, which equipped with the inner products $\langle \cdot, \cdot \rangle_X$ and $\langle \cdot, \cdot \rangle_Y$, respectively, are complete, and therefore Hilbert spaces. Additionally, let $T \in \mathfrak{L}(X, Y), T_h \in \mathfrak{L}(X_h, Y_h)$ be linear, continuous operators, and $\mathcal{P} : X \times X \to \mathbb{R}, \mathcal{P}_h : X_h \times X_h \to \mathbb{R}$ be bilinear forms. Our aim is to prove that under certain conditions, the coercivity of the bilinear form $\mathcal{P}$ on $\ker(T)$ is inherited (for sufficiently small h) by the parametric bilinear form $\mathcal{P}_h$ on $\ker(T_h)$ with a constant independent of h.

Theorem 2.1.10 (Parametric Coercivity)
Let X, Y be Hilbert spaces and $X_h \subseteq X$, $Y_h \subseteq Y$ finite dimensional closed subspaces. Suppose $T \in \mathfrak{L}(X, Y), T_h \in \mathfrak{L}(X_h, Y_h)$ and $\mathcal{P} : X \times X \to \mathbb{R}, \mathcal{P}_h : X_h \times X_h \to \mathbb{R}$ are linear, bounded operators and continuous, symmetric bilinear forms, respectively, where the bounds of T_h and $\mathcal{P}_h$ are independent of h. Furthermore, the following properties hold:

(i) *T is uniformly surjective with constant κ.*

(ii) *$\mathcal{P}$ is coercive on $\ker(T) \subseteq X$ with constant γ.*

(iii) *There exist $\tilde{\kappa}, h_1 > 0$ independent of h such that for all $0 < h \leq h_1$ the mapping T_h is uniformly surjective with constant $\tilde{\kappa}$.*

(iv) *There exists a constant $\mathbf{L}_T \geq 0$ independent of h such that*

$$\|Tx_h - T_h x_h\|_Y \leq \mathbf{L}_T h \|x_h\|_X$$

for all $x_h \in X_h$.

(v) *There exists a constant $\mathbf{L}_\mathcal{P} \geq 0$ independent of h such that*

$$\mathcal{P}(x_h, x_h) - \mathcal{P}_h(x_h, x_h) \leq \mathbf{L}_\mathcal{P} h \|x_h\|_X^2$$

for all $x_h \in X_h \cap \ker(T)$.

Then, there exist $\tilde{\gamma}, \tilde{h} > 0$ independent of h such that for every $0 < h \leq \tilde{h}$ the bilinear form $\mathcal{P}_h$ is coercive on $\ker(T_h)$ with constant $\tilde{\gamma}$.

Proof. Define the restriction $\tilde{T}_h := T|_{X_h}$ with

$$\tilde{T}_h \in \mathfrak{L}\left(X_h, \tilde{Y}_h\right) , \quad \tilde{Y}_h := T\left(X_h\right) \subseteq Y ,$$

which is uniformly surjective with constant κ, since $\operatorname{im}\left(\tilde{T}_h\right) = \tilde{Y}_h$ and (i) hold. Moreover, $\dim\left(\tilde{Y}_h\right) \leq \dim\left(X_h\right)$, $\dim\left(\tilde{Y}_h\right) < \infty$, and $\tilde{Y}_h$ is closed, due to the surjectivity of $\tilde{T}_h$. Equipped with the inner product of Y, the subspace $\tilde{Y}_h$ is complete, and therefore a Hilbert space. This allows us to apply Lemma 2.1.6 for $\tilde{T}_h$, thus the mapping $\tilde{T}_h \circ \mathcal{I}_{X_h}^{-1} \circ \tilde{T}_h^\star \in \mathfrak{L}\left(\tilde{Y}_h^\star, \tilde{Y}_h\right)$ is bijective and the inverse operator is bounded by $\left\|\left(\tilde{T}_h \circ \mathcal{I}_{X_h}^{-1} \circ \tilde{T}_h^\star\right)^{-1}\right\|_{\mathfrak{L}\left(\tilde{Y}_h, \tilde{Y}_h^\star\right)} \leq \frac{1}{\kappa^2}$. In addition, according to [120, Satz III.4.2], it holds $\left\|\tilde{T}_h^\star\right\|_{\mathfrak{L}\left(\tilde{Y}_h^\star, X_h^\star\right)} = \left\|\tilde{T}_h\right\|_{\mathfrak{L}\left(X_h, \tilde{Y}_h\right)}$. Hence, by Remark 2.1.7, the right inverse $\tilde{R}_h := \mathcal{I}_{X_h}^{-1} \circ \tilde{T}_h^\star \circ \left(\tilde{T}_h \circ \mathcal{I}_{X_h}^{-1} \circ \tilde{T}_h^\star\right)^{-1}$ satisfies

$$\left\|\tilde{R}_h\right\|_{\mathfrak{L}\left(\tilde{Y}_h, X_h\right)} \leq \frac{\left\|\mathcal{I}_{X_h}^{-1}\right\|_{\mathfrak{L}\left(X_h^\star, X_h\right)} \left\|\tilde{T}_h\right\|_{\mathfrak{L}\left(X_h, \tilde{Y}_h\right)}}{\kappa^2} \leq \Gamma_R, \tag{2.1.12}$$

where $\Gamma_R := \frac{\|T\|_{\mathfrak{L}(X,Y)}}{\kappa^2}$ is independent of h. For an arbitrary $u_h \in \ker\left(\tilde{T}_h\right)$ it holds

$$\mathbf{L}_{\mathcal{P}} h \left\|u_h\right\|_X^2 \overset{(v)}{\geq} \mathcal{P}\left(u_h, u_h\right) - \mathcal{P}_h\left(u_h, u_h\right) \overset{(ii)}{\geq} \gamma \left\|u_h\right\|_X^2 - \mathcal{P}_h\left(u_h, u_h\right) .$$

Reordering and choosing $h \leq \frac{\gamma}{2\mathbf{L}_{\mathcal{P}}}$ yields

$$\mathcal{P}_h\left(u_h, u_h\right) \geq \left(\gamma - \mathbf{L}_{\mathcal{P}} h\right) \left\|u_h\right\|_X^2 \geq \frac{\gamma}{2} \left\|u_h\right\|_X^2 .$$

Then, according to Theorem 2.1.9, there exist $\gamma_0, \Gamma_U > 0$ independent of h such that

$$\mathcal{P}_h\left(u_h + z_h, u_h + z_h\right) \geq \gamma_0 \left\|u_h + z_h\right\|_X^2 \tag{2.1.13}$$

for all $u_h \in \ker\left(\tilde{T}_h\right)$ and $z_h \in X_h$ with $\|z_h\|_X \leq \Gamma_U \|u_h\|_X$. Set $\tilde{\gamma} := \gamma_0$ and

$$\tilde{h} := \min\left\{h_1, \frac{\gamma}{2\mathbf{L}_{\mathcal{P}}}, \frac{1}{2\Gamma_R \mathbf{L}_T}, \frac{\Gamma_U}{2\Gamma_R \mathbf{L}_T}\right\} . \tag{2.1.14}$$

Let $0 < h \leq \tilde{h}$ and $x_h \in \ker\left(T_h\right)$ be arbitrary. Then, it holds

$$\left\|\tilde{T}_h x_h\right\|_Y = \left\|\tilde{T}_h x_h - T_h x_h\right\|_Y \overset{(iv)}{\leq} \mathbf{L}_T h \left\|x_h\right\|_X . \tag{2.1.15}$$

Define

$$z_h := \tilde{R}_h\left(\tilde{T}_h x_h\right) , \qquad u_h := x_h - z_h ,$$

which satisfy $\tilde{T}_h u_h = \tilde{T}_h x_h - \left(\tilde{T}_h \circ \tilde{R}_h\right)\left(\tilde{T}_h x_h\right) = \tilde{T}_h x_h - \tilde{T}_h x_h = \mathbf{0}_Y$, hence $u_h \in \ker\left(\tilde{T}_h\right)$. Furthermore, the choice of $\tilde{h}$ implies

$$\begin{aligned} \|u_h\|_X &= \|x_h - z_h\|_X \geq \|x_h\|_X - \|z_h\|_X = \|x_h\|_X - \left\|\tilde{R}_h\left(\tilde{T}_h x_h\right)\right\|_X \\ &\overset{(2.1.12)}{\geq} \|x_h\|_X - \Gamma_R \left\|\tilde{T}_h x_h\right\|_Y \overset{(2.1.15)}{\geq} \left(1 - \Gamma_R \mathbf{L}_T h\right) \|x_h\|_X \overset{(2.1.14)}{\geq} \frac{1}{2} \|x_h\|_X , \end{aligned} \tag{2.1.16}$$

and consequently

$$\|z_h\|_X \overset{(2.1.12)}{\leq} \Gamma_R \left\|\tilde{T}_h x_h\right\|_Y \overset{(2.1.15)}{\leq} \Gamma_R \mathbf{L}_T h \|x_h\|_X \overset{(2.1.16)}{\leq} 2\Gamma_R \mathbf{L}_T h \|u_h\|_X \overset{(2.1.14)}{\leq} \Gamma_U \|u_h\|_X .$$

Finally, (2.1.13) yields

$$\mathcal{P}_h(x_h, x_h) = \mathcal{P}_h(u_h + z_h, u_h + z_h) \overset{(2.1.13)}{\geq} \gamma_0 \|u_h + z_h\|_X^2 = \tilde{\gamma} \|x_h\|_X^2 ,$$

which completes the proof. □

2.2 Generalized Equations

In this section, we examine properties of *generalized equations*, i.e., inclusions of the form

$$\mathbf{0}_Y \in T(x) + F(x). \tag{2.2.1}$$

Herein, X, Y are Banach spaces, $T : X \to Y$ is a single-valued function, and $F : X \rightrightarrows Y$ is a *set-valued mapping*. Problems of this type have been extensively investigated in [36]. The inverse of the set-valued mapping F is defined as

$$F^{-1}(y) = \{x \in X \mid y \in F(x)\}.$$

For a subset $P \subseteq Y$ we consider the parametric generalized equation

$$\mathbf{0}_Y \in T(x) + F(x) + p, \quad p \in P. \tag{2.2.2}$$

First, we derive an implicit function theorem for (2.2.2) similar to [30, 35], that is finding conditions under which (2.2.2) has a solution x depending on p near a reference solution $(\hat{x}, \hat{p})$. In case of a single-valued equation $\mathbf{0}_Y = T(x) + p$ a sufficient assumption would amount to $T^{-1}(\hat{x})$ being a linear, continuous operator. However, for (2.2.2) we require a different property introduced by Robinson [109]:

Definition 2.2.1 (Strong Regularity)
Let $\Omega \subseteq X$ be open and $\hat{x} \in \Omega$ such that $\hat{x}$ solves (2.2.1) and T is Fréchet differentiable at $\hat{x}$. Furthermore, let the set-valued mapping $S : X \rightrightarrows Y$ be defined by

$$S(x) := T(\hat{x}) + T'(\hat{x})(x - \hat{x}) + F(x). \tag{2.2.3}$$

We call (2.2.1) strongly regular at $\hat{x}$ with associated Lipschitz constant $\mathbf{L} > 0$, if there exist neighborhoods V of $\hat{x}$ and U of $\mathbf{0}_Y$ such that for every $u \in U$ the set

$$S^{-1}(u) \cap V = \{v \in V \mid u \in S(v)\} \tag{2.2.4}$$

contains a single element $v(u)$, and the mapping $G^S_{U,V} : U \to V$ defined as

$$G^S_{U,V}(u) := v(u) \tag{2.2.5}$$

is Lipschitz continuous with constant $\mathbf{L}$.

The property of *strong regularity* allows us to prove the following *implicit function theorem* for (2.2.2):

Theorem 2.2.2 (Implicit Function Theorem)
Let T be Fréchet differentiable on Ω. Suppose for $\hat{x} \in \Omega$ and $\hat{p} \in P$ the operators T, T' are continuous at $\hat{x}$, and

$$\mathbf{0}_Y \in T(x) + \hat{p} + F(x) \tag{2.2.6}$$

is strongly regular at $\hat{x}$ with Lipschitz constant $\mathbf{L}$. Then, for any $\varepsilon > 0$ there exist neighborhoods W_ε of $\hat{x}$ and N_ε of $\hat{p}$, and a single-valued mapping $x : N_\varepsilon \to W_\varepsilon$ such that for each $p \in N_\varepsilon$, $x(p)$ is the unique solution of the inclusion (2.2.2) *in W_ε. Moreover, for every $p_1, p_2 \in N_\varepsilon$ it holds*

$$\|x(p_1) - x(p_2)\|_X \le (\mathbf{L} + \varepsilon) \|p_1 - p_2\|_Y .$$

Proof. Let $\varepsilon > 0$ be arbitrary and let us define

$$S(x) := T(\hat{x}) + \hat{p} + T'(\hat{x})(x - \hat{x}) + F(x).$$

Then, by strong regularity of (2.2.6) there exist neighborhoods V of $\hat{x}$ and U of $\mathbf{0}_Y$ such that for every $u \in U$ the set

$$S^{-1}(u) \cap V = \{v \in V \mid u \in S(v)\}$$

contains a single element $v(u)$, and the mapping $G^S_{U,V} : U \to V$ defined as

$$G^S_{U,V}(u) := v(u)$$

is Lipschitz continuous with constant $\mathbf{L}$. Choose $\delta > 0$ such that $\mathbf{L}\delta < \frac{\varepsilon}{\mathbf{L}+\varepsilon} < 1$, and define the parametric mapping $r : \Omega \times P \to Y$ by

$$r(x,p) := T(\hat{x}) + \hat{p} + T'(\hat{x})(x - \hat{x}) - T(x) - p.$$

Since T' is continuous in $\hat{x}$, there exists $\rho_1 > 0$ such that

$$\|T'(x) - T'(\hat{x})\|_{\mathfrak{L}(X,Y)} \le \delta \tag{2.2.7}$$

for every $x \in \mathcal{B}_{\rho_1}(\hat{x}) \subseteq \Omega$. Now, choose $\rho, \varrho > 0$, $\rho \le \rho_1$ such that $\mathcal{B}_{\delta\rho+\varrho}(\mathbf{0}_Y) \subseteq U$ and

$$\mathbf{L}\varrho \le (1 - \mathbf{L}\delta)\rho. \tag{2.2.8}$$

Hence, (2.2.7) is also satisfied for every $x \in \mathcal{B}_\rho(\hat{x})$. In addition, using the mean-value theorem in [59, p. 40], it follows that for every $x \in \mathcal{B}_\rho(\hat{x})$ and $p \in \mathcal{B}_\varrho(\hat{p})$ it holds

$$\begin{aligned}
\|r(x,p)\|_Y &\le \|T(\hat{x}) - T(x) + T'(\hat{x})(x - \hat{x})\|_Y + \|\hat{p} - p\|_Y \\
&\le \sup_{\theta \in (0,1)} \|T'((1-\theta)x + \theta\hat{x}) - T'(\hat{x})\|_{\mathfrak{L}(X,Y)} \|x - \hat{x}\|_X + \|\hat{p} - p\|_Y \\
&\overset{(2.2.7)}{\le} \delta \|x - \hat{x}\|_X + \|\hat{p} - p\|_Y \le \delta\rho + \varrho,
\end{aligned}$$

since $\|(1-\theta)x+\theta\hat{x}-\hat{x}\|_X = (1-\theta)\|x-\hat{x}\|_X \le \rho$. Thus, $r(x,p) \in \mathcal{B}_{\delta\rho+\varrho}(\mathbf{0}_Y) \subseteq U$. Set $W_\varepsilon := \mathcal{B}_\rho(\hat{x})$ and $N_\varepsilon := \mathcal{B}_\varrho(\hat{p})$. For arbitrary $p \in N_\varepsilon$ we define the function $\Pi_p : W_\varepsilon \to V$ by

$$\Pi_p(x) := G^S_{U,V}(r(x,p)),$$

which satisfies

$$\begin{aligned} & \Pi_p(x) = x \\ \Leftrightarrow\quad & r(x,p) \in S(x) \\ \Leftrightarrow\quad & T(\hat{x}) + \hat{p} + T'(\hat{x})(x-\hat{x}) - T(x) - p \in T(\hat{x}) + \hat{p} + T'(\hat{x})(x-\hat{x}) + F(x) \\ \Leftrightarrow\quad & \mathbf{0}_Y \in T(x) + p + F(x) \end{aligned} \tag{2.2.9}$$

for every $x \in W_\varepsilon$. Utilizing the Lipschitz continuity of $G^S_{U,V}$ and the mean-value theorem in [59, p. 40] implies

$$\begin{aligned} \|\Pi_p(x_1) - \Pi_p(x_2)\|_X &\le \mathbf{L}\|r(x_1,p) - r(x_2,p)\|_Y \\ &= \mathbf{L}\|T'(\hat{x})(x_1-x_2) - (T(x_1) - T(x_2))\|_Y \\ &\le \mathbf{L}\sup_{\theta\in(0,1)}\|T'((1-\theta)x_1+\theta x_2) - T'(\hat{x})\|_{\mathfrak{L}(X,Y)}\|x_1-x_2\|_X \\ &\overset{(2.2.7)}{\le} \mathbf{L}\delta\|x_1-x_2\|_X \end{aligned} \tag{2.2.10}$$

for all $x_1, x_2 \in W_\varepsilon$, since $\|(1-\theta)x_1+\theta x_2-\hat{x}\|_X \le (1-\theta)\rho + \theta\rho = \rho$. It follows that Π_p is a contraction mapping on W_ε, because $\mathbf{L}\delta < \frac{\varepsilon}{\mathbf{L}+\varepsilon} < 1$. Moreover, it holds

$$\mathbf{0}_Y \in T(\hat{x}) + \hat{p} + F(\hat{x}) = S(\hat{x}),$$

and therefore $G^S_{U,V}(\mathbf{0}_Y) = \hat{x}$. Consequently, since $r(\hat{x},p) = \hat{p} - p$, we obtain

$$\begin{aligned} \|\Pi_p(\hat{x}) - \hat{x}\|_X &= \left\|G^S_{U,V}(r(\hat{x},p)) - G^S_{U,V}(\mathbf{0}_Y)\right\|_X \\ &\le \mathbf{L}\|r(\hat{x},p)\|_Y = \mathbf{L}\|p-\hat{p}\|_Y \\ &\le \mathbf{L}\varrho \overset{(2.2.8)}{\le} (1-\mathbf{L}\delta)\rho. \end{aligned} \tag{2.2.11}$$

Then, for every $x \in W_\varepsilon$ we have $\Pi_p(x) \in W_\varepsilon$, since by (2.2.10) and (2.2.11) it holds

$$\begin{aligned} \|\Pi_p(x) - \hat{x}\|_X &\le \|\Pi_p(x) - \Pi_p(\hat{x})\|_X + \|\Pi_p(\hat{x}) - \hat{x}\|_X \\ &\le \mathbf{L}\delta\|x-\hat{x}\|_X + (1-\mathbf{L}\delta)\rho \\ &\le \mathbf{L}\delta\rho + (1-\mathbf{L}\delta)\rho = \rho. \end{aligned}$$

According to the Banach contraction principle (cf. [40, Theorem 12.3]), the self-map Π_p has a unique fixed point $x(p)$ with

$$\|x(p) - x\|_X \le \frac{1}{1-\mathbf{L}\delta}\|\Pi_p(x) - x\|_X \quad \text{for every } x \in W_\varepsilon. \tag{2.2.12}$$

It follows from (2.2.9) that for each $p \in N_\varepsilon$ the relation

$$\Pi_p(x(p)) = x(p) \Leftrightarrow \mathbf{0}_Y \in T(x(p)) + p + F(x(p))$$

is satisfied. In addition, by uniqueness of the fixed point, $x(p)$ is the unique solution of the inclusion (2.2.2). Furthermore, for arbitrary $p_1, p_2 \in N_\varepsilon$ it holds

$$\begin{aligned} \|x(p_1) - x(p_2)\|_X &\overset{(2.2.12)}{\leq} \frac{1}{1-\mathbf{L}\delta} \|\Pi_{p_1}(x(p_2)) - x(p_2)\|_X \\ &= \frac{1}{1-\mathbf{L}\delta} \|\Pi_{p_1}(x(p_2)) - \Pi_{p_2}(x(p_2))\|_X, \end{aligned}$$

and, since $r(x, p_1) - r(x, p_2) = p_2 - p_1$, we obtain

$$\begin{aligned} \|\Pi_{p_1}(x(p_2)) - \Pi_{p_2}(x(p_2))\|_X &= \left\| G^S_{U,V}(r(x(p_2), p_1)) - G^S_{U,V}(r(x(p_2), p_2)) \right\|_X \\ &\leq \mathbf{L} \|r(x(p_2), p_1) - r(x(p_2), p_2)\|_Y \\ &= \mathbf{L} \|p_1 - p_2\|_Y. \end{aligned}$$

Finally, we conclude

$$\|x(p_1) - x(p_2)\|_X \leq \frac{\mathbf{L}}{1-\mathbf{L}\delta} \|p_1 - p_2\|_Y \leq (\mathbf{L}+\varepsilon) \|p_1 - p_2\|_Y,$$

since $1 - \mathbf{L}\delta > 1 - \frac{\varepsilon}{\mathbf{L}+\varepsilon} = \frac{\mathbf{L}}{\mathbf{L}+\varepsilon}$, which completes the proof. □

Let H be a Banach space of parameters and $\tilde{H} \subseteq H$ a neighborhood of the origin $\mathbf{0}_H$. For $h \in \tilde{H} \setminus \{\mathbf{0}_H\}$ we denote the subspaces $X_h \subseteq X$, $Y_h \subseteq Y$ supplied with the same respective norms. For a function $T_h : X_h \to Y_h$ and a set-valued mapping $F_h : X_h \rightrightarrows Y_h$ we consider the generalized equation

$$\mathbf{0}_Y \in T_h(x_h) + F_h(x_h). \tag{2.2.13}$$

Let $\hat{x}$ denote a solution of (2.2.1). Our goal is to find conditions such that for sufficiently small $\|h\|_H$ the inclusion (2.2.13) has a unique solution $\hat{x}_h$ with

$$\|\hat{x}_h - \hat{x}\|_X \to 0 \quad \text{for} \quad \|h\|_H \to 0.$$

To that end, we consider the parametric generalized equation

$$\mathbf{0}_Y \in T_h(x_h) + p_h + F_h(x_h). \tag{2.2.14}$$

Remark 2.2.3

Suppose for a parameter $\hat{p}_h$ the inclusion (2.2.14) is strongly regular at z_h with Lipschitz constant $\mathbf{L}_h$, and $\|z_h - \hat{x}\|_X \to 0$, $\|\hat{p}_h\|_Y \to 0$ for $\|h\|_H \to 0$. Then, there are $\rho_h, \varrho_h > 0$ such that for all $p_h \in \mathcal{B}_{\varrho_h}(\hat{p}_h)$ the inclusion (2.2.14) has a unique solution $x_h(p_h)$ in $\mathcal{B}_{\rho_h}(z_h)$, which is Lipschitz continuous with constant $\mathbf{L}_h$. If the constants $\mathbf{L}_h = \mathbf{L}$, $\rho_h = \rho$, and $\varrho_h = \varrho$ were independent of h, then we could choose $\|h\|_H$ sufficiently small such that $\|\hat{p}_h\|_Y < \varrho$, thus $\mathbf{0}_Y \in \mathcal{B}_\varrho(\hat{p}_h)$. Therefore, (2.2.13) would have a unique solution $\hat{x}_h$ in $\mathcal{B}_{\rho_h}(z_h)$ with

$$\|\hat{x}_h - \hat{x}\|_X \leq \|z_h - \hat{x}\|_X + \|\hat{x}_h - z_h\|_X \leq \|z_h - \hat{x}\|_X + (\mathbf{L}+\varepsilon)\|\hat{p}_h\|_Y \to 0$$

for $\|h\|_H \to 0$.

This means we require a reinforced strong regularity condition such that the constants $\mathbf{L}_h$, ρ_h, and ϱ_h are independent of h.

Definition 2.2.4 (Uniform Strong Regularity)
Let $z_h \in X_h$ exist such that z_h solves (2.2.13) *and T_h is Fréchet differentiable at z_h. We call* (2.2.13) *uniformly strongly regular at z_h, if there exist $\varsigma, \varrho, \rho, \mathbf{L} > 0$ independent of h such that for every $h \in \mathcal{B}_\varsigma(\mathbf{0}_H) \setminus \{\mathbf{0}_H\}$ the inclusion* (2.2.13) *is strongly regular at z_h with neighborhoods $\mathcal{B}_\rho(z_h)$, $\mathcal{B}_\varrho(\mathbf{0}_Y)$ and Lipschitz constant $\mathbf{L}$.*

Remark 2.2.5
In order to verify uniform strong regularity at z_h, one has to prove the existence of $\mathbf{L}, \varsigma, \varrho, \rho > 0$ independent of h such that for every $h \in \mathcal{B}_\varsigma(\mathbf{0}_H) \setminus \{\mathbf{0}_H\}$ and every $y_h \in \mathcal{B}_\varrho(\mathbf{0}_Y)$ the inclusion

$$y_h \in T_h(z_h) + T_h'(z_h)(x_h - z_h) + F_h(x_h)$$

has a unique solution $x_h(y_h)$ in $\mathcal{B}_\rho(z_h)$, which is Lipschitz continuous with respect to y_h and Lipschitz constant $\mathbf{L}$.

The notion of uniform strong regularity allows us to prove that the generalized equation (2.2.13) has a unique solution $\hat{x}_h$ that converges to the solution $\hat{x}$ of inclusion (2.2.1) for $\|h\|_H \to 0$, which is fundamental for the main result of Chapter 5 (Theorem 5.5.6):

Theorem 2.2.6 (Convergence)
Let $\hat{x}$ be a solution of (2.2.1) *and let T_h be Fréchet differentiable. Furthermore, let the following conditions hold:*

(i) Let $T_h'(\cdot)$ be Lipschitz continuous with constant $\mathbf{L}_T > 0$.

(ii) Let there exist $z_h \in X_h$ and $\hat{p}_h \in Y_h$ such that

$$\mathbf{0}_Y \in T_h(z_h) + \hat{p}_h + F_h(z_h), \tag{2.2.15}$$

$$\textit{and} \qquad \|z_h - \hat{x}\|_X \to 0, \quad \|\hat{p}_h\|_Y \to 0, \quad \textit{for } \|h\|_H \to 0. \tag{2.2.16}$$

(iii) Let (2.2.15) *be uniformly strongly regular at z_h with associated Lipschitz constant $\mathbf{L} > 0$.*

Then, there exists $\boldsymbol{l}, \tilde{\varsigma} > 0$ such that for each $h \in \mathcal{B}_{\tilde{\varsigma}}(\mathbf{0}_H) \setminus \{\mathbf{0}_H\}$ the inclusion (2.2.13) *has a locally unique solution $\hat{x}_h$ with*

$$\|\hat{x}_h - \hat{x}\|_X \leq \boldsymbol{l}(\|z_h - \hat{x}\|_X + \|\hat{p}_h\|_Y)$$

Proof. According to (iii), there exist $\mathbf{L}, \varsigma, \varrho, \rho > 0$, which satisfy the conditions in Definition 2.2.4. For arbitrary $h \in \mathcal{B}_\varsigma(\mathbf{0}_H) \setminus \{\mathbf{0}_H\}$ we apply Theorem 2.2.2 to (2.2.15) for some $\varepsilon > 0$. By condition (i), the inequality (2.2.7) holds for T_h' with a constant independent of h. Additionally, the neighborhoods W_ε and $N_\varepsilon = \mathcal{B}_\varrho(\hat{p}_h)$ can be chosen independent of h, since uniform strong regularity holds for neighborhoods independent of h. Select $\tilde{\varsigma} > 0$ such that $\|\hat{p}_h\|_Y \leq \frac{\varrho}{2}$ for each $h \in \mathcal{B}_{\tilde{\varsigma}}(\mathbf{0}_H) \setminus \{\mathbf{0}_H\}$, which is possible according to (2.2.16). Then, $\mathbf{0}_Y \in N_\varepsilon$, hence there exists a solution $\hat{x}_h$ of (2.2.13) with $\|\hat{x}_h - z_h\|_X \leq (\mathbf{L} + \varepsilon) \|\hat{p}_h\|_Y$. Exploiting the triangle inequality yields

$$\|\hat{x}_h - \hat{x}\|_X \leq \|z_h - \hat{x}\|_X + \|\hat{x}_h - z_h\|_X \leq \boldsymbol{l} \left(\|z_h - \hat{x}\|_X + \|\hat{p}_h\|_Y\right)$$

for $\boldsymbol{l} := \max\{1, (\mathbf{L} + \varepsilon)\}$, which proves the assertion. $\square$

Remark 2.2.7

Condition (ii) *in Theorem 2.2.6 is often referred to as consistency. If* (2.2.13) *is an approximation of* (2.2.1)*, then one usually chooses the projection of* $\hat{x}$ *into the subspace* X_h *as the solution* z_h *of the inclusion* (2.2.15)*, and* $\hat{p}_h$ *is chosen accordingly. Therefore,* $\|z_h - \hat{x}\|_X$ *is called interpolation error, and* $\|\hat{p}_h\|_Y$ *is called consistency error. The rate of convergence depends on these errors, e.g., if there exist* $\mathbf{L}_1, \Gamma_1 \geq 0$ *independent of* h *such that* $\|z_h - \hat{x}\|_X \leq \mathbf{L}_1 \|h\|$ *and* $\|\hat{p}_h\|_Y \leq \Gamma_1 \|h\|$*, then the solution* $\hat{x}_h$ *of the approximated generalized equation* (2.2.13) *would converge linearly to the solution* $\hat{x}$ *of* (2.2.1).

2.3 Parametric Nonlinear Optimization Problems

In this section, we aim to derive a *sensitivity result* and *second-order sufficient conditions* for special cases of the following problem:

Problem 2.3.1 (Parametric Nonlinear Optimization Problem)
For a parameter $p \in \mathcal{P} \subseteq \mathbb{R}^{n_p}$, $\mathcal{P}$ closed, and $n_z, n_H, n_G \in \mathbb{N}$ with $n_H + n_G \leq n_z$ let the functions $\mathcal{J} : \mathbb{R}^{n_z} \times \mathcal{P} \to \mathbb{R}$, $G : \mathbb{R}^{n_z} \times \mathcal{P} \to \mathbb{R}^{n_G}$, $H : \mathbb{R}^{n_z} \times \mathcal{P} \to \mathbb{R}^{n_H}$ be sufficiently smooth.

$$\begin{aligned} \textit{Minimize} \quad & \mathcal{J}(z,p) \\ \textit{with respect to} \quad & z \in \mathbb{R}^{n_z} \\ \textit{subject to} \quad & H(z,p) = \mathbf{0}_{\mathbb{R}^{n_H}}, \\ & G(z,p) \leq \mathbf{0}_{\mathbb{R}^{n_G}}. \end{aligned}$$

Problems of this type have been extensively treated in the literature, e.g., [2, 10, 44, 97]. Let the Lagrange function of Problem 2.3.1 be denoted by

$$\mathcal{L}(\ell_0, z, \lambda, \eta, p) := \ell_0 \mathcal{J}(z,p) + \lambda^\top H(z,p) + \eta^\top G(z,p) \ ,$$

and the set of feasible vectors by

$$\Sigma(p) := \{ z \in \mathbb{R}^{n_z} \mid H(z,p) = \mathbf{0}_{\mathbb{R}^{n_H}}, G(z,p) \leq \mathbf{0}_{\mathbb{R}^{n_G}} \}.$$

According to [10, Theorem 4.3.2], necessary conditions for a fixed parameter can be expressed as:

Theorem 2.3.2 (Fritz John Necessary Conditions)
For a fixed parameter $\hat{p} \in \mathcal{P}$ let $\hat{z}$ be a local minimum of Problem 2.3.1, and let $\mathcal{J}(\cdot, \hat{p})$, $G(\cdot, \hat{p})$, $H(\cdot, \hat{p})$ be continuously differentiable. Then, there exist multipliers $\ell_0 \geq 0$, $\lambda \in \mathbb{R}^{n_H}$, and $\eta \in \mathbb{R}^{n_G}$ not all zero such that

$$\begin{aligned} \nabla_z \mathcal{L}(\ell_0, \hat{z}, \lambda, \eta, \hat{p}) &= \mathbf{0}_{\mathbb{R}^{n_z}}, \\ \eta &\geq \mathbf{0}_{\mathbb{R}^{n_G}}, \\ \eta^\top G(\hat{z}, \hat{p}) &= 0. \end{aligned}$$

If so called *constraint qualifications* hold (cf. [10, Chapter 5]), then there exist multipliers that satisfy Theorem 2.3.2 with $\ell_0 = 1$. Among the most commonly used constraint qualifications are the *Mangasarian-Fromovitz constraint qualification* and the *linear independence constraint qualification* (LICQ). For our purposes we choose the latter:

Definition 2.3.3 (Linear Independence Constraint Qualification)
For a fixed parameter $\hat{p} \in \mathcal{P}$ the linear independence constraint qualification is satisfied at $\hat{z}$ for Problem 2.3.1, if the vectors

$$\begin{aligned} &\nabla_z H_j(\hat{z}, \hat{p}), \quad j = 1, \ldots, n_H, \\ &\nabla_z G_j(\hat{z}, \hat{p}), \quad j = 1, \ldots, n_G \textit{ with } G_j(\hat{z}, \hat{p}) = 0 \end{aligned}$$

are linear independent.

Remark 2.3.4 (Uniqueness of Multipliers)
If the assumptions of Theorem 2.3.2 are satisfied and the linear independence constraint qualification hold at $\hat{z}$, then the necessary conditions hold for $\ell_0 = 1$, and the associated multipliers λ, η are unique, cf. [48, Corollary 2.3.39].

The constraints of Problem 2.3.1 together with the necessary conditions for $\ell_0 = 1$, i.e.,

$$\begin{aligned} \nabla_z \mathcal{L}(1, z, \lambda, \eta, p) &= \mathbf{0}_{\mathbb{R}^{n_z}}, \\ H(z,p) &= \mathbf{0}_{\mathbb{R}^{n_H}}, \\ G(z,p) &\leq \mathbf{0}_{\mathbb{R}^{n_G}}, \\ \eta &\geq \mathbf{0}_{\mathbb{R}^{n_G}}, \\ \eta^\top G(z,p) &= 0. \end{aligned} \tag{2.3.1}$$

are called *Karush-Kuhn-Tucker conditions* (KKT conditions). For a non-empty, closed, and convex set $C \subseteq \mathbb{R}^n$ we denote the *normal cone operator* by

$$\mathcal{N}_C : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$$
$$\mathcal{N}_C(x) := \begin{cases} \left\{ y \in \mathbb{R}^n \mid y^\top (c - x) \leq 0 \text{ for all } c \in C \right\}, & \text{if } x \in C \\ \emptyset, & \text{if } x \notin C \end{cases}.$$

Note that $\mathcal{N}_C$ satisfies the following conditions

$$\begin{aligned} \mathcal{N}_{\mathbb{R}^n}(x) &= \{\mathbf{0}_{\mathbb{R}^n}\}, \\ \mathcal{N}_{\mathbb{R}^n_+}(x) &= \begin{cases} \left\{ y \in \mathbb{R}^n \mid y^\top x = 0,\ y \leq \mathbf{0}_{\mathbb{R}^n} \right\}, & \text{if } x \in \mathbb{R}^n_+ \\ \emptyset, & \text{if } x \notin \mathbb{R}^n_+ \end{cases}, \\ \mathcal{N}_{C_1 \times C_2}(x_1, x_2) &= \begin{cases} \mathcal{N}_{C_1}(x_1) \times \mathcal{N}_{C_2}(x_2), & \text{if } (x_1, x_2) \in C_1 \times C_2 \\ \emptyset, & \text{if } (x_1, x_2) \notin C_1 \times C_2 \end{cases}. \end{aligned}$$

This allows us to write the KKT-conditions (2.3.1) as a generalized equation of the form

$$\mathbf{0}_{\mathbb{R}^{n_z} \times \mathbb{R}^{n_H} \times \mathbb{R}^{n_G}} \in - \begin{pmatrix} \nabla_z \mathcal{L}(1, z, \lambda, \eta, p) \\ H(z,p) \\ G(z,p) \end{pmatrix} + \mathcal{N}_{\mathbb{R}^{n_z} \times \mathbb{R}^{n_H} \times \mathbb{R}^{n_G}_+} \begin{pmatrix} z \\ \lambda \\ \eta \end{pmatrix}. \tag{2.3.2}$$

For a fixed parameter $p \in \mathcal{P}$ let $\left(\hat{z}(p), \hat{\lambda}(p), \hat{\eta}(p)\right)$ be a *Karush-Kuhn-Tucker point* (KKT-point) of Problem 2.3.1, i.e., $\left(\hat{z}(p), \hat{\lambda}(p), \hat{\eta}(p)\right)$ satisfies the KKT-conditions (2.3.1). Furthermore, for the KKT-point and $p \in \mathcal{P}$, we decompose the index set $J := \{1, \ldots, n_G\}$ into the subsets

$$\begin{aligned} \hat{J}^+(p) &:= \{ j \in J \mid G_j(\hat{z}(p), p) = 0,\ \hat{\eta}_j(p) > 0 \}, \quad \hat{j}^+(p) := \operatorname{card}\left(\hat{J}^+(p)\right), \\ \hat{J}^0(p) &:= \{ j \in J \mid G_j(\hat{z}(p), p) = 0,\ \hat{\eta}_j(p) = 0 \}, \quad \hat{j}^0(p) := \operatorname{card}\left(\hat{J}^0(p)\right), \\ \hat{J}^-(p) &:= \{ j \in J \mid G_j(\hat{z}(p), p) < 0,\ \hat{\eta}_j(p) = 0 \}, \quad \hat{j}^-(p) := \operatorname{card}\left(\hat{J}^-(p)\right), \end{aligned}$$

and use the abbreviations

$$\begin{aligned}
\hat{A}_H(p) &:= H'(\hat{z}(p), p), \\
\hat{B}_G^+(p) &:= \left[G_j'(\hat{z}(p), p)\right]_{j\in\hat{J}^+(p)}, \\
\hat{B}_G^0(p) &:= \left[G_j'(\hat{z}(p), p)\right]_{j\in\hat{J}^0(p)}, \\
\hat{B}_G^-(p) &:= \left[G_j'(\hat{z}(p), p)\right]_{j\in\hat{J}^-(p)}.
\end{aligned}$$

If $\hat{J}^0(p) = \emptyset$, then the strict complementarity condition $\hat{\eta}_j(p) - G_j(\hat{z}(p), p) > 0$ holds for all $j \in J$. For a fixed parameter $\hat{p} \in \mathcal{P}$ second-order sufficient conditions for Problem 2.3.1 were proven in, e.g., [10, Theorem 4.4.2], [47, Theorem 6.1.3]:

Theorem 2.3.5 (Second-Order Sufficient Conditions)
For a fixed parameter $\hat{p} \in \mathcal{P}$ let $\left(\hat{z}(\hat{p}), \hat{\lambda}(\hat{p}), \hat{\eta}(\hat{p})\right)$ be a KKT-point of Problem 2.3.1 and let the functions $\mathcal{J}(\cdot, \hat{p}), H(\cdot, \hat{p})$, and $G(\cdot, \hat{p})$ be twice continuously differentiable. Suppose for every $v \in \mathbb{R}^{n_z} \setminus \{\mathbf{0}_{\mathbb{R}^{n_z}}\}$ with

$$\begin{aligned}
\hat{A}_H(\hat{p})\, v &= \mathbf{0}_{\mathbb{R}^{n_H}} \\
\hat{B}_G^+(\hat{p})\, v &= \mathbf{0}_{\mathbb{R}^{n_G}} \\
\hat{B}_G^0(\hat{p})\, v &\le \mathbf{0}_{\mathbb{R}^{n_G}}
\end{aligned} \tag{2.3.3}$$

the inequality

$$v^\top \nabla_{zz}^2 \mathcal{L}\left(1, \hat{z}(\hat{p}), \hat{\lambda}(\hat{p}), \hat{\eta}(\hat{p}), \hat{p}\right) v > 0$$

is satisfied. Then, there exist $\alpha, \rho > 0$ such that for every $z \in \Sigma(\hat{p}) \cap \mathcal{B}_\rho(\hat{z}(\hat{p}))$ it holds

$$\mathcal{J}(z, \hat{p}) \ge \mathcal{J}(\hat{z}(\hat{p}), \hat{p}) + \alpha \left\| z - \hat{z}(\hat{p}) \right\|^2.$$

Remark 2.3.6
The set of all vectors $v \in \mathbb{R}^{n_z}$ that satisfy (2.3.3) *is called critical cone. For our purposes we require a stronger condition than the one in Theorem 2.3.5. Particularly, by removing the inequality $\hat{B}_G^0(\hat{p})\, v \le \mathbf{0}_{\mathbb{R}^{n_G}}$ in* (2.3.3) *we obtain the condition*

$$v^\top \nabla_{zz}^2 \mathcal{L}\left(1, \hat{z}(\hat{p}), \hat{\lambda}(\hat{p}), \hat{\eta}(\hat{p}), \hat{p}\right) v > 0 \quad \textit{for all } v \in \ker\left(\begin{bmatrix} \hat{A}_H(\hat{p}) \\ \hat{B}_G^+(\hat{p}) \end{bmatrix}\right) \setminus \{\mathbf{0}_{\mathbb{R}^{n_z}}\}. \tag{2.3.4}$$

Of course, if $\hat{J}^0(p)$ is empty, then $\hat{B}_G^0(\hat{p})$ is vacuous. Thus, the condition in Theorem 2.3.5 and (2.3.4) *would be equivalent. However, in general*

$$\ker\left(\begin{bmatrix} \hat{A}_H(\hat{p}) \\ \hat{B}_G^+(\hat{p}) \end{bmatrix}\right)$$

is a superset of the critical cone.

Condition (2.3.4) together with the linear independence constraint qualification are sufficient for strong regularity of the generalized equation (2.3.2), as shown in [109, Theorem 4.1]:

Theorem 2.3.7 (Strongly Regular KKT-Conditions, Robinson)
For a fixed parameter $\hat{p} \in \mathcal{P}$ let $\left(\hat{z}(\hat{p}), \hat{\lambda}(\hat{p}), \hat{\eta}(\hat{p})\right)$ be a KKT-point of Problem 2.3.1 and let the functions $\mathcal{J}(\cdot, \hat{p}), H(\cdot, \hat{p})$, and $G(\cdot, \hat{p})$ be twice differentiable at $\hat{z}(\hat{p})$. Furthermore, let the following conditions hold:

(i) The matrix

$$\begin{bmatrix} \hat{A}_H(\hat{p}) \\ \hat{B}_G^{+}(\hat{p}) \\ \hat{B}_G^{0}(\hat{p}) \end{bmatrix}$$

has full row rank.

(ii) For all $v \in \ker\left(\begin{bmatrix} \hat{A}_H(\hat{p}) \\ \hat{B}_G^{+}(\hat{p}) \end{bmatrix}\right) \setminus \{\mathbf{0}_{\mathbb{R}^{n_z}}\}$ one has $v^\top \nabla_{zz}^2 \mathcal{L}\left(1, \hat{z}(\hat{p}), \hat{\lambda}(\hat{p}), \hat{\eta}(\hat{p}), \hat{p}\right) v > 0$.

Then, (2.3.2) *is strongly regular at $\left(\hat{z}(\hat{p}), \hat{\lambda}(\hat{p}), \hat{\eta}(\hat{p})\right)$.*

This statement allows us to derive a crucial sensitivity result for the following special case of Problem 2.3.1:

Problem 2.3.8
For a parameter $p \in \mathcal{P} \subseteq \mathbb{R}^{n_p}$, $\mathcal{P}$ closed, $n_z, n_H, n_G \in \mathbb{N}$ with $n_H + n_G \leq n_z$, and matrices $R \in \mathbb{R}^{n_z \times n_p}$, $C \in \mathbb{R}^{n_H \times n_p}$, $D \in \mathbb{R}^{n_G \times n_p}$ let the functions $\mathcal{J} : \mathbb{R}^{n_z} \times \mathcal{P} \to \mathbb{R}$, $G : \mathbb{R}^{n_z} \times \mathcal{P} \to \mathbb{R}^{n_G}$, $H : \mathbb{R}^{n_z} \times \mathcal{P} \to \mathbb{R}^{n_H}$ defined by

$$\mathcal{J}(z,p) := \tilde{\mathcal{J}}(z) + z^\top Rp, \quad H(z,p) := \tilde{H}(z) + Cp, \quad G(z,p) := \tilde{G}(z) + Dp$$

be sufficiently smooth.

$$\begin{aligned} \textit{Minimize} \quad & \tilde{\mathcal{J}}(z) + z^\top Rp \\ \textit{with respect to} \quad & z \in \mathbb{R}^{n_z} \\ \textit{subject to} \quad & \tilde{H}(z) + Cp = \mathbf{0}_{\mathbb{R}^{n_H}}, \\ & \tilde{G}(z) + Dp \leq \mathbf{0}_{\mathbb{R}^{n_G}}. \end{aligned} \tag{2.3.5}$$

Analog to (2.3.2), we write the KKT-conditions of Problem 2.3.8 as the generalized equation

$$\mathbf{0}_{\mathbb{R}^{n_z} \times \mathbb{R}^{n_H} \times \mathbb{R}^{n_G}} \in -\begin{pmatrix} \nabla_z \mathcal{L}(1, z, \lambda, \eta, p) \\ \tilde{H}(z) + Cp \\ \tilde{G}(z) + Dp \end{pmatrix} + \mathcal{N}_{\mathbb{R}^{n_z} \times \mathbb{R}^{n_H} \times \mathbb{R}_+^{n_G}} \begin{pmatrix} z \\ \lambda \\ \eta \end{pmatrix} \tag{2.3.6}$$

$$\Leftrightarrow \quad \mathbf{0}_{\mathbb{R}^{n_z} \times \mathbb{R}^{n_H} \times \mathbb{R}^{n_G}} \in -\begin{pmatrix} \nabla_z \mathcal{L}(1, z, \lambda, \eta, \mathbf{0}_{\mathbb{R}^{n_p}}) \\ \tilde{H}(z) \\ \tilde{G}(z) \end{pmatrix} - \begin{pmatrix} R \\ C \\ D \end{pmatrix} p + \mathcal{N}_{\mathbb{R}^{n_z} \times \mathbb{R}^{n_H} \times \mathbb{R}_+^{n_G}} \begin{pmatrix} z \\ \lambda \\ \eta \end{pmatrix}$$

with the Lagrange function

$$\mathcal{L}(\ell_0, z, \lambda, \eta, p) := \ell_0 \left(\tilde{\mathcal{J}}(z) + z^\top Rp\right) + \lambda^\top \left(\tilde{H}(z) + Cp\right) + \eta^\top \left(\tilde{G}(z) + Dp\right).$$

For the new parameter $q(p) := -\begin{pmatrix} R \\ C \\ D \end{pmatrix} p \in \mathbb{R}^m$ with $m = n_z + n_H + n_G$ we consider the parametric generalized equation

$$\mathbf{0}_{\mathbb{R}^m} \in -\begin{pmatrix} \nabla_z \mathcal{L}\left(1, z, \lambda, \eta, \mathbf{0}_{\mathbb{R}^{n_p}}\right) \\ \tilde{H}(z) \\ \tilde{G}(z) \end{pmatrix} + q(p) + \mathcal{N}_{\mathbb{R}^{n_z} \times \mathbb{R}^{n_H} \times \mathbb{R}_+^{n_G}} \begin{pmatrix} z \\ \lambda \\ \eta \end{pmatrix}, \tag{2.3.7}$$

and prove the following:

Theorem 2.3.9 (Sensitivity of the KKT-Conditions)
Let the assumptions of Theorem 2.3.7 hold for a fixed parameter $\hat{p} \in \mathcal{P}$ and a local minimum $\hat{z}(\hat{p})$ with the associated multipliers $\hat{\lambda}(\hat{p}), \hat{\eta}(\hat{p})$ of Problem 2.3.8. Furthermore, let the functions $\tilde{\mathcal{J}}, \tilde{H}$, and $\tilde{G}$ be twice continuously differentiable. Then, there exist $\varrho > 0$, a neighborhood W of $\left(\hat{z}(\hat{p}), \hat{\lambda}(\hat{p}), \hat{\eta}(\hat{p})\right)$, and Lipschitz continuous functions $z : \mathcal{B}_\varrho(\hat{p}) \to \mathbb{R}^{n_z}$, $\lambda : \mathcal{B}_\varrho(\hat{p}) \to \mathbb{R}^{n_H}$, $\eta : \mathcal{B}_\varrho(\hat{p}) \to \mathbb{R}^{n_G}$ such that for each $p \in \mathcal{B}_\varrho(\hat{p})$, $(z(p), \lambda(p), \eta(p))$ is the unique solution of the inclusion (2.3.6) *in W, and $z(p)$ is a local minimum of Problem 2.3.8.*

Proof. $\left(\hat{z}(\hat{p}), \hat{\lambda}(\hat{p}), \hat{\eta}(\hat{p})\right)$ satisfies the linear independence constraint qualification in Definition 2.3.3, hence it solves the inclusion (2.3.7). According to Theorem 2.3.7, for $q(\hat{p})$ the inclusion (2.3.7) is strongly regular at $\left(\hat{z}(\hat{p}), \hat{\lambda}(\hat{p}), \hat{\eta}(\hat{p})\right)$ with some Lipschitz constant $\mathbf{L} > 0$. Then, by Theorem 2.2.2 for a fixed $\varepsilon > 0$ there exist neighborhoods W_ε of $\left(\hat{z}(\hat{p}), \hat{\lambda}(\hat{p}), \hat{\eta}(\hat{p})\right)$ and N_ε of $q(\hat{p})$, a single-valued mapping $\left(\tilde{z}, \tilde{\lambda}, \tilde{\eta}\right) : N_\varepsilon \to W_\varepsilon$, which is Lipschitz continuous with constant $\mathbf{L} + \varepsilon$, and for each $q \in N_\varepsilon$, $\left(\tilde{z}, \tilde{\lambda}, \tilde{\eta}\right)(q)$ is the unique solution of (2.3.7) in W_ε. In addition, for $p_1, p_2 \in \mathbb{R}^{n_p}$ with $q(p_1), q(p_2) \in N_\varepsilon$ it holds

$$\left\| \left(\tilde{z}, \tilde{\lambda}, \tilde{\eta}\right)(q(p_1)) - \left(\tilde{z}, \tilde{\lambda}, \tilde{\eta}\right)(q(p_2)) \right\| \leq (\mathbf{L} + \varepsilon) \|q(p_1) - q(p_2)\| \leq \boldsymbol{l} \|p_1 - p_2\|$$

for $\boldsymbol{l} := (\mathbf{L} + \varepsilon)(\|R\| + \|C\| + \|D\|)$. Set $\alpha := \min_{j \in \hat{J}^+(\hat{p})} \hat{\eta}_j(\hat{p}) > 0$, $W := W_\varepsilon$, and select $\varrho > 0$ such that

$$q\left(\mathcal{B}_\varrho(\hat{p})\right) \subseteq N_\varepsilon \text{ and } \varrho \leq \frac{\alpha}{2\boldsymbol{l}}. \tag{2.3.8}$$

Then, the functions $z : \mathcal{B}_\varrho(\hat{p}) \to \mathbb{R}^{n_z}$, $\lambda : \mathcal{B}_\varrho(\hat{p}) \to \mathbb{R}^{n_H}$, $\eta : \mathcal{B}_\varrho(\hat{p}) \to \mathbb{R}^{n_G}$ defined as

$$z(p) := \tilde{z}(q(p)), \quad \lambda(p) := \tilde{\lambda}(q(p)), \quad \eta(p) := \tilde{\eta}(q(p))$$

are Lipschitz continuous with constant $\boldsymbol{l}$. Additionally, $(z(p), \lambda(p), \eta(p))$ is the unique solution of (2.3.6) in W for each $p \in \mathcal{B}_\varrho(\hat{p})$, and it holds $(z(\hat{p}), \lambda(\hat{p}), \eta(\hat{p})) = \left(\hat{z}(\hat{p}), \hat{\lambda}(\hat{p}), \hat{\eta}(\hat{p})\right)$. We denote

$$\begin{aligned} J^+(p) &:= \{j \in J \mid G_j(z(p), p) = 0,\ \eta_j(p) > 0\}, \\ A_H(p) &:= H'(z(p), p), \\ B_G^+(p) &:= \left[G_j'(z(p), p)\right]_{j \in J^+(p)}, \end{aligned}$$

which satisfy $J^+(\hat{p}) = \hat{J}^+(\hat{p})$, $A_H(\hat{p}) = \hat{A}_H(\hat{p})$, and $B_G^+(\hat{p}) = \hat{B}_G^+(\hat{p})$. In order to verify the optimality of $z(p)$ for Problem 2.3.8, we show that for every $p \in \mathcal{B}_\varrho(\hat{p})$ there exists $\gamma(p) > 0$ such that

$$v^\top \nabla_{zz}^2 \mathcal{L}(1, z(p), \lambda(p), \eta(p), p)\, v \geq \gamma(p) \|v\|^2 \quad \text{for every } v \in \ker\left(\begin{bmatrix} A_H(p) \\ B_G^+(p) \end{bmatrix}\right),$$

which is sufficient for the assumption in Theorem 2.3.5. To that end, let $j \in J^+(\hat{p}) = \hat{J}^+(\hat{p})$ be arbitrary. Then, by choice of ϱ

$$\eta_j(\hat{p}) - \eta_j(p) \leq \|\eta(\hat{p}) - \eta(p)\| \leq \boldsymbol{l}\, \|\hat{p} - p\| \overset{(2.3.8)}{\leq} \frac{\alpha}{2},$$

is satisfied for every $p \in \mathcal{B}_\varrho(\hat{p})$. This and $\eta_j(\hat{p}) \geq \alpha$ imply

$$0 < \frac{\alpha}{2} \leq \hat{\eta}_j(\hat{p}) - \frac{\alpha}{2} \leq \eta_j(p),$$

hence $j \in J^+(p)$ for each $p \in \mathcal{B}_\varrho(\hat{p})$. We conclude $J^+(\hat{p}) \subseteq J^+(p)$ and therefore

$$\ker\left(\begin{bmatrix} A_H(p) \\ B_G^+(p) \end{bmatrix}\right) \subseteq \ker\left(\begin{bmatrix} A_H(p) \\ \left[\tilde{G}_j'(z(p))\right]_{j \in J^+(\hat{p})} \end{bmatrix}\right) =: \mathcal{K}(p).$$

We will prove positive definiteness of $\nabla_{zz}^2 \mathcal{L}(1, z(p), \lambda(p), \eta(p), p)$ on the larger set $\mathcal{K}(p)$ by assuming the contrary. Thus, for each $i \in \mathbb{N}$ there exists some $p_i \in \mathcal{B}_{\frac{\varrho}{i}}(\hat{p})$, and for every $k \in \mathbb{N}$ there exists some $v_{i_k} \in \mathcal{K}(p_i)$ with

$$v_{i_k}^\top \nabla_{zz}^2 \mathcal{L}(1, z(p_{i_k}), \lambda(p_{i_k}), \eta(p_{i_k}), p_{i_k})\, v_{i_k} < \frac{1}{i_k} \|v_{i_k}\|^2. \tag{2.3.9}$$

Note that the parameter p only appears linearly in

$$\mathcal{L}(\ell_0, z, \lambda, \eta, p) = \ell_0 \left(\tilde{\mathcal{J}}(z) + z^\top R p\right) + \lambda^\top \left(\tilde{H}(z) + Cp\right) + \eta^\top \left(\tilde{G}(z) + Dp\right),$$

hence $\nabla_{zz}^2 \mathcal{L}(1, \cdot, \cdot, \cdot, p) = \nabla_{zz}^2 \mathcal{L}(1, \cdot, \cdot, \cdot, \hat{p})$. Since the balls $\mathcal{B}_1(\mathbf{0}_{\mathbb{R}^{n_z}})$ and $\mathcal{B}_\varrho(\hat{p})$ are compact with respect to Euclidean norm $\|\cdot\|$, there exist convergent subsequences of $(p_{i_k})_{i_k \in \mathbb{N}} \subseteq \mathcal{B}_\varrho(\hat{p})$ and $\left(\frac{v_{i_k}}{\|v_{i_k}\|}\right)_{i_k \in \mathbb{N}} \subseteq \mathcal{B}_1(\mathbf{0}_{\mathbb{R}^{n_z}})$ with limits $\hat{p}$ and $\hat{v} \in \mathcal{K}(\hat{p})$, respectively. In order to minimize the use of indexes, we assume without loss of generality that $(p_{i_k})_{i_k \in \mathbb{N}}$ and $\left(\frac{v_{i_k}}{\|v_{i_k}\|}\right)_{i_k \in \mathbb{N}}$ are convergent with limits

$$\lim_{k \to \infty} p_{i_k} = \hat{p}, \quad \lim_{k \to \infty} \frac{v_{i_k}}{\|v_{i_k}\|} = \hat{v} \in \mathcal{K}(\hat{p}) = \ker\left(\begin{bmatrix} A_H(\hat{p}) \\ B_G^+(\hat{p}) \end{bmatrix}\right) = \ker\left(\begin{bmatrix} \hat{A}_H(\hat{p}) \\ \hat{B}_G^+(\hat{p}) \end{bmatrix}\right), \quad \|\hat{v}\| = 1.$$

Then, dividing (2.3.9) by $\|v_{i_k}\|^2$ and taking the limit yields $\hat{v}^\top \nabla_{zz}^2 \mathcal{L}\left(1, \hat{z}, \hat{\lambda}, \hat{\eta}, \hat{p}\right) \hat{v} \leq \mathbf{0}_{\mathbb{R}^{n_z}}$, which contradicts assumption (ii) in Theorem 2.3.7. Thus, $\nabla_{zz}^2 \mathcal{L}(1, z(p), \lambda(p), \eta(p), p)$ is positive definite on $\mathcal{K}(p)$, which is a superset of $\ker\left(\begin{bmatrix} A_H(p) \\ B_G^+(p) \end{bmatrix}\right)$. $\square$

Linear Quadratic Case:
Let us consider a special case of Problem 2.3.8, where

$$\tilde{\mathcal{J}}(z) := \frac{1}{2} z^\top Q z, \quad \tilde{H}(z) := Az, \quad \tilde{G}(z) := Bz.$$

with a symmetric matrix $Q \in \mathbb{R}^{n_z \times n_z}$, $A \in \mathbb{R}^{n_H \times n_z}$, and $B \in \mathbb{R}^{n_G \times n_z}$. Thus, Problem 2.3.8 becomes a linear quadratic optimization problem:

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2} z^\top Q z + z^\top R p \\ \text{with respect to} \quad & z \in \mathbb{R}^{n_z} \\ \text{subject to} \quad & Az + Cp = \mathbf{0}_{\mathbb{R}^{n_H}}, \\ & Bz + Dp \leq \mathbf{0}_{\mathbb{R}^{n_G}}. \end{aligned} \tag{2.3.10}$$

If the matrix $\begin{bmatrix} A \\ B \end{bmatrix}$ has full row rank, then the set of admissible vectors is not empty, and the linear independence constraint qualification in Definition 2.3.3 is satisfied for every $p \in \mathcal{P}$. Furthermore, if $v^\top Q v > 0$ for all $v \in \ker(A) \setminus \{\mathbf{0}_{\mathbb{R}^{n_z}}\}$, then the sufficient conditions in Theorem 2.3.5 hold for any KKT-point. Hence, the KKT-conditions of (2.3.10), expressed as the linear generalized equation

$$\mathbf{0}_{\mathbb{R}^m} \in - \begin{bmatrix} Q & A^\top & B^\top \\ A & \mathbf{0}_{n_H \times n_H} & \mathbf{0}_{n_H \times n_G} \\ B & \mathbf{0}_{n_G \times n_H} & \mathbf{0}_{n_G \times n_G} \end{bmatrix} \begin{pmatrix} z \\ \lambda \\ \eta \end{pmatrix} - \begin{pmatrix} R \\ C \\ D \end{pmatrix} \hat{p} + \mathcal{N}_{\mathbb{R}^{n_z} \times \mathbb{R}^{n_H} \times \mathbb{R}_+^{n_G}} \begin{pmatrix} z \\ \lambda \\ \eta \end{pmatrix}, \tag{2.3.11}$$

have a unique (global) solution $\left(\hat{z}(\hat{p}), \hat{\lambda}(\hat{p}), \hat{\eta}(\hat{p})\right)$ for every $\hat{p} \in \mathbb{R}^{n_p}$. Moreover, for every $q \in \mathbb{R}^m$ the perturbed inclusion

$$q \in - \begin{bmatrix} Q & A^\top & B^\top \\ A & \mathbf{0}_{n_H \times n_H} & \mathbf{0}_{n_H \times n_G} \\ B & \mathbf{0}_{n_G \times n_H} & \mathbf{0}_{n_G \times n_G} \end{bmatrix} \begin{pmatrix} z \\ \lambda \\ \eta \end{pmatrix} - \begin{pmatrix} R \\ C \\ D \end{pmatrix} \hat{p} + \mathcal{N}_{\mathbb{R}^{n_z} \times \mathbb{R}^{n_H} \times \mathbb{R}_+^{n_G}} \begin{pmatrix} z \\ \lambda \\ \eta \end{pmatrix},$$

also has a unique solution. Since (2.3.11) is already linear, it follows from Theorem 2.3.7 and Definition 2.2.1, that (2.3.11) is strongly regular at the unique KKT-point $\left(\hat{z}(\hat{p}), \hat{\lambda}(\hat{p}), \hat{\eta}(\hat{p})\right)$ for every $\hat{p} \in \mathbb{R}^{n_p}$. According to Theorem 2.3.9, for every $\hat{p} \in \mathbb{R}^{n_p}$ there exist $\varrho(\hat{p}), \boldsymbol{l}(\hat{p}) > 0$, a neighborhood $W(\hat{p})$ of $\left(\hat{z}(\hat{p}), \hat{\lambda}(\hat{p}), \hat{\eta}(\hat{p})\right)$, and Lipschitz continuous functions

$$\tilde{z}_{\hat{p}} : \mathcal{B}_{\varrho(\hat{p})}(\hat{p}) \to \mathbb{R}^{n_z}, \quad \tilde{\lambda}_{\hat{p}} : \mathcal{B}_{\varrho(\hat{p})}(\hat{p}) \to \mathbb{R}^{n_H}, \quad \tilde{\eta}_{\hat{p}} : \mathcal{B}_{\varrho(\hat{p})}(\hat{p}) \to \mathbb{R}^{n_G}$$

with constant $\boldsymbol{l}(\hat{p})$ such that for each $p \in \mathcal{B}_{\varrho(\hat{p})}(\hat{p})$, $\left(\tilde{z}_{\hat{p}}(p), \tilde{\lambda}_{\hat{p}}(p), \tilde{\eta}_{\hat{p}}(p)\right)$ is the unique solution of

$$\mathbf{0}_{\mathbb{R}^m} \in - \begin{bmatrix} Q & A^\top & B^\top \\ A & \mathbf{0}_{n_H \times n_H} & \mathbf{0}_{n_H \times n_G} \\ B & \mathbf{0}_{n_G \times n_H} & \mathbf{0}_{n_G \times n_G} \end{bmatrix} \begin{pmatrix} z \\ \lambda \\ \eta \end{pmatrix} - \begin{pmatrix} R \\ C \\ D \end{pmatrix} (\hat{p} + p) + \mathcal{N}_{\mathbb{R}^{n_z} \times \mathbb{R}^{n_H} \times \mathbb{R}_+^{n_G}} \begin{pmatrix} z \\ \lambda \\ \eta \end{pmatrix}.$$

Let $\hat{\varrho} > 0$ be arbitrary and let $\mathcal{U}_{\varrho(\hat{p})}(\hat{p})$ denote the open ball around $\hat{p}$ with radius $\varrho(\hat{p})$. Then,

$$\bigcup_{\hat{p}\in\mathcal{B}_{\hat{\varrho}}(\mathbf{0}_{\mathbb{R}^{n_p}})} \mathcal{U}_{\varrho(\hat{p})}(\hat{p}) \tag{2.3.12}$$

is an open cover of $\mathcal{B}_{\hat{\varrho}}(\mathbf{0}_{\mathbb{R}^{n_p}})$, since the union of an arbitrary number of open sets is open, and $\varrho(\hat{p}) > 0$ for every $\hat{p} \in \mathcal{B}_{\hat{\varrho}}(\mathbf{0}_{\mathbb{R}^{n_p}})$. In addition, by compactness of $\mathcal{B}_{\hat{\varrho}}(\mathbf{0}_{\mathbb{R}^{n_p}})$, there exists a finite subcover of (2.3.12). Hence, there exist $i \in \mathbb{N}$ and $\hat{p}_1, \ldots, \hat{p}_i$ such that

$$\bigcup_{k=1}^{i} \mathcal{U}_{\varrho(\hat{p}_k)}(\hat{p}_k) \supseteq \mathcal{B}_{\hat{\varrho}}(\mathbf{0}_{\mathbb{R}^{n_p}}).$$

With the functions $\tilde{z}_{\hat{p}_k}, \tilde{\lambda}_{\hat{p}_k}, \tilde{\eta}_{\hat{p}_k}$ for $k = 1, \ldots, i$ we are able to construct Lipschitz continuous functions $z : \mathcal{B}_{\hat{\varrho}}(\mathbf{0}_{\mathbb{R}^{n_p}}) \to \mathbb{R}^{n_z}$, $\lambda : \mathcal{B}_{\hat{\varrho}}(\mathbf{0}_{\mathbb{R}^{n_p}}) \to \mathbb{R}^{n_H}$, $\eta : \mathcal{B}_{\hat{\varrho}}(\mathbf{0}_{\mathbb{R}^{n_p}}) \to \mathbb{R}^{n_G}$ such that for every $\hat{p} \in \mathcal{B}_{\hat{\varrho}}(\mathbf{0}_{\mathbb{R}^{n_p}})$, $(z(\hat{p}), \lambda(\hat{p}), \eta(\hat{p}))$ is the unique solution of the inclusion (2.3.11). Since $\left(\hat{z}(\hat{p}), \hat{\lambda}(\hat{p}), \hat{\eta}(\hat{p})\right)$ is the unique solution of (2.3.11) for every $\hat{p} \in \mathbb{R}^{n_p}$, it holds

$$\left(\hat{z}(\hat{p}), \hat{\lambda}(\hat{p}), \hat{\eta}(\hat{p})\right) = (z(\hat{p}), \lambda(\hat{p}), \eta(\hat{p})),$$

for each $\hat{p} \in \mathcal{B}_{\hat{\varrho}}(\mathbf{0}_{\mathbb{R}^{n_p}})$, and $\left(\hat{z}(\hat{p}), \hat{\lambda}(\hat{p}), \hat{\eta}(\hat{p})\right)$ is Lipschitz continuous with respect to $\hat{p}$ and constant $\boldsymbol{l} := \max_{k=1,\ldots,i} \boldsymbol{l}(\hat{p}_k)$.

Let us summarize these statements for the special case (2.3.10) of Problem 2.3.8 in the following corollary:

Corollary 2.3.10
Let $Q \in \mathbb{R}^{n_z \times n_z}$ be symmetric, $A \in \mathbb{R}^{n_H \times n_z}$, $B \in \mathbb{R}^{n_G \times n_z}$, $R \in \mathbb{R}^{n_z \times n_p}$, $C \in \mathbb{R}^{n_H \times n_p}$, and $D \in \mathbb{R}^{n_G \times n_p}$. Furthermore, let the following conditions hold:

(i) The matrix $\begin{bmatrix} A \\ B \end{bmatrix}$ has full row rank.

(ii) For all $v \in \ker(A) \setminus \{\mathbf{0}_{\mathbb{R}^{n_z}}\}$ one has $v^\top Q v > 0$.

Then, for an arbitrary $\varrho > 0$ and each $p \in \mathcal{B}_{\varrho}(\mathbf{0}_{\mathbb{R}^{n_p}})$ the linear quadratic optimization problem

$$\begin{aligned} \textit{Minimize} \quad & \frac{1}{2} z^\top Q z + z^\top R p \\ \textit{with respect to} \quad & z \in \mathbb{R}^{n_z} \\ \textit{subject to} \quad & Az + Cp = \mathbf{0}_{\mathbb{R}^{n_H}}, \\ & Bz + Dp \leq \mathbf{0}_{\mathbb{R}^{n_G}}, \end{aligned}$$

has a unique solution $\hat{z}(p)$ together with unique Lagrange multipliers $\hat{\lambda}(p)$, $\hat{\eta}(p)$. Moreover, $\left(\hat{z}(p), \hat{\lambda}(p), \hat{\eta}(p)\right)$ is Lipschitz continuous with respect to $p \in \mathcal{B}_{\varrho}(\mathbf{0}_{\mathbb{R}^{n_p}})$.

This result is crucial for the proof of Lemma 5.5.4, where a parametric optimization problem of type (2.3.10) occurs.

Our next goal is to derive (uniform) second-order sufficient conditions for the following optimization problem with parametric objective function:

Problem 2.3.11

Let $p \in \mathcal{P} \subset \mathbb{R}$ be a given parameter, $\mathcal{P}$ compact, and for $n_z, n_H, n_G \in \mathbb{N}$ with $n_H + n_G \leq n_z$ let the functions $\mathcal{J} : \mathbb{R}^{n_z} \times \mathcal{P} \to \mathbb{R}$, $G : \mathbb{R}^{n_z} \to \mathbb{R}^{n_G}$, $H : \mathbb{R}^{n_z} \to \mathbb{R}^{n_H}$ be sufficiently smooth.

$$\begin{aligned} \textit{Minimize} \quad & \mathcal{J}(z, p) \\ \textit{with respect to} \quad & z \in \mathbb{R}^{n_z} \\ \textit{subject to} \quad & H(z) = \mathbf{0}_{\mathbb{R}^{n_H}}, \\ & G(z) \leq \mathbf{0}_{\mathbb{R}^{n_G}}. \end{aligned}$$

With the Lagrange function

$$\mathcal{L}(\ell_0, z, \lambda, \eta, p) := \ell_0 \mathcal{J}(z, p) + \lambda^\top H(z) + \eta^\top G(z) \ ,$$

the KKT conditions of Problem 2.3.11 can be expressed as

$$\begin{aligned} \nabla_z \mathcal{L}(1, z, \lambda, \eta, p) &= \mathbf{0}_{\mathbb{R}^{n_z}}, \\ H(z) &= \mathbf{0}_{\mathbb{R}^{n_H}}, \\ G(z) &\leq \mathbf{0}_{\mathbb{R}^{n_G}}, \\ \eta &\geq \mathbf{0}_{\mathbb{R}^{n_G}}, \\ \eta^\top G(z) &= 0. \end{aligned} \tag{2.3.13}$$

For every $p \in \mathcal{P}$ let $\left(\hat{z}(p), \hat{\lambda}(p), \hat{\eta}(p)\right)$ be a KKT-point of Problem 2.3.11. Furthermore, we define the *linearizing cone* $K(p)$ and the cone $K^+(p)$ by

$$\begin{aligned} K(p) &:= \left\{ d \in \ker\left(H'(\hat{z}(p))\right) \mid G'_j(\hat{z}(p))\, d \leq 0,\ j \in \hat{J}^0(p) \cup \hat{J}^+(p) \right\}, \\ K^+(p) &:= \left\{ d \in \ker\left(H'(\hat{z}(p))\right) \mid G'_j(\hat{z}(p))\, d = 0,\ j \in \hat{J}^+(p) \right\}. \end{aligned}$$

Clearly, $K^+(p) \subseteq K(p)$ is satisfied, if the *strict complementarity condition* $\hat{J}^0(p) = \emptyset$ holds. According to the KKT-conditions (2.3.13), for an arbitrary $d \in \ker\left(H'(\hat{z}(p))\right)$ it holds

$$\begin{aligned} 0 &= \nabla_z \mathcal{J}(\hat{z}(p), p)^\top d + \hat{\lambda}(p)^\top H'(\hat{z}(p))\, d + \hat{\eta}(p)^\top G'(\hat{z}(p))\, d \\ &= \nabla_z \mathcal{J}(\hat{z}(p), p)^\top d + \sum_{j \in \hat{J}^+(p)} \hat{\eta}_j(p)\, G'_j(\hat{z}(p))\, d, \end{aligned}$$

which implies the relation

$$d \in K^+(p) \quad \Leftrightarrow \quad \nabla_z \mathcal{J}(\hat{z}(p), p)^\top d = 0 \text{ and } d \in \ker\left(H'(\hat{z}(p))\right). \tag{2.3.14}$$

Moreover, if $\hat{J}^0(p) = \emptyset$, then it holds

$$d \in K(p) \quad \Leftrightarrow \quad \nabla_z \mathcal{J}(\hat{z}(p), p)^\top d \geq 0 \text{ and } d \in \ker\left(H'(\hat{z}(p))\right). \tag{2.3.15}$$

For a constant $\nu \geq 0$ we define the cone

$$\tilde{K}^{\nu}(p) := \left\{ d \in \ker\left(H'\left(\hat{z}(p)\right)\right) \mid 0 \leq \nabla_z \mathcal{J}\left(\hat{z}(p), p\right)^\top d \leq \nu \|d\| \right\},$$

which satisfies $K^+(p) \subset \tilde{K}^{\nu}(p)$ and $K^+(p) = \tilde{K}^0(p)$. Additionally, if the strict complementarity condition $\hat{J}^0(p) = \emptyset$ holds, then the relation (2.3.15) implies $\tilde{K}^{\nu}(p) \subseteq K(p)$. Similar to [88, Lemma 5.7] we prove that, if the matrix $\nabla^2_{zz}\mathcal{L}\left(1, \hat{z}(p), \hat{\lambda}(p), \hat{\eta}(p), p\right)$ is uniformly positive definite on the cone $K^+(p)$, then it is also uniformly positive definite on $\tilde{K}^{\nu}(p)$ for sufficiently small $\nu > 0$.

Lemma 2.3.12
Let the functions H and G be twice continuously differentiable, let $\mathcal{J}(\cdot, p)$ be twice continuously differentiable for every $p \in \mathcal{P}$, and let $\mathcal{J}(z, \cdot)$, $\nabla_z \mathcal{J}(z, \cdot)$, and $\nabla^2_{zz}\mathcal{J}(z, \cdot)$ be continuous for every $z \in \mathbb{R}^{n_z}$. Let there exist continuous functions $\hat{z}(\cdot) : \mathcal{P} \to \mathbb{R}^{n_z}$, $\hat{\lambda}(\cdot) : \mathcal{P} \to \mathbb{R}^{n_H}$, $\hat{\eta}(\cdot) : \mathcal{P} \to \mathbb{R}^{n_G}$ such that $\left(\hat{z}(p), \hat{\lambda}(p), \hat{\eta}(p)\right)$ is a KKT-point of Problem 2.3.11 for every $p \in \mathcal{P}$. Moreover, let the following be satisfied:

(i) There exists a constant $\beta > 0$ such that for all $p \in \mathcal{P}$ and every $d \in \mathbb{R}^{n_H} \times \mathbb{R}^{\hat{j}^+(p)} \times \mathbb{R}^{\hat{j}^0(p)}$ it holds

$$\left\| \begin{bmatrix} \hat{A}_H(p) \\ \hat{B}^+_G(p) \\ \hat{B}^0_G(p) \end{bmatrix}^\top d \right\| \geq \beta \|d\|.$$

(ii) There exists a constant $\gamma > 0$ such that for all $p \in \mathcal{P}$ and every $v \in K^+(p)$ it holds

$$v^\top \nabla^2_{zz}\mathcal{L}\left(1, \hat{z}(p), \hat{\lambda}(p), \hat{\eta}(p), p\right) v \geq \gamma \|v\|^2.$$

Then, there exist $\tilde{\gamma}, \nu > 0$ such that for all $p \in \mathcal{P}$ and every $v \in \tilde{K}^{\nu}(p)$ it holds

$$v^\top \nabla^2_{zz}\mathcal{L}\left(1, \hat{z}(p), \hat{\lambda}(p), \hat{\eta}(p), p\right) v \geq \tilde{\gamma} \|v\|^2.$$

Proof. Assume the opposite is true. Thus, for every $i \in \mathbb{N}$ there exist $p_i \in \mathcal{P}$ and a vector $v_i \in \ker\left(H'\left(\hat{z}(p_i)\right)\right)$ such that

$$\begin{aligned} 0 \leq \nabla_z \mathcal{J}\left(\hat{z}(p_i), p_i\right)^\top v_i &\leq \frac{1}{i}\|v_i\|, \\ v_i^\top \nabla^2_{zz}\mathcal{L}\left(1, \hat{z}(p_i), \hat{\lambda}(p_i), \hat{\eta}(p_i), p_i\right) v_i &< \frac{1}{i}\|v_i\|^2. \end{aligned}$$

The second inequality implies $v_i \neq \mathbf{0}_{\mathbb{R}^{n_z}}$. Define the set-valued function $M : \mathcal{P} \rightrightarrows \mathbb{R}^{n_z}$ as $M(p) := \ker\left(H'\left(\hat{z}(p)\right)\right) \cap \mathcal{B}_1\left(\mathbf{0}_{\mathbb{R}^{n_z}}\right)$. According to Lemma A.6, $\operatorname{graph}(M)$ is compact and furthermore it holds $\left(p_i, \frac{v_i}{\|v_i\|}\right) \in \operatorname{graph}(M)$. The compactness implies that there exists a convergent sub-sequence $\left(\left(p_{i_k}, \frac{v_{i_k}}{\|v_{i_k}\|}\right)\right)_{i_k \in \mathbb{N}} \subseteq \operatorname{graph}(M)$ with some limit $(\tilde{p}, \tilde{v})$ in $\operatorname{graph}(M)$, hence $\tilde{v} \in \ker\left(H'\left(\hat{z}(\tilde{p})\right)\right)$ and $\|\tilde{v}\| = 1$. Additionally, it holds

$$\begin{aligned} 0 \leq \nabla_z \mathcal{J}\left(\hat{z}(p_{i_k}), p_{i_k}\right)^\top \frac{v_{i_k}}{\|v_{i_k}\|} &\leq \frac{1}{i_k}, \\ \frac{v_{i_k}}{\|v_{i_k}\|}^\top \nabla^2_{zz}\mathcal{L}\left(1, \hat{z}(p_{i_k}), \hat{\lambda}(p_{i_k}), \hat{\eta}(p_{i_k}), p_{i_k}\right) \frac{v_{i_k}}{\|v_{i_k}\|} &< \frac{1}{i_k}. \end{aligned}$$

Taking the limit yields

$$\begin{aligned} \nabla_z \mathcal{J}\left(\hat{z}\left(\tilde{p}\right),\tilde{p}\right)^\top \tilde{v} &= 0, \\ \tilde{v}^\top \nabla^2_{zz}\mathcal{L}\left(1,\hat{z}\left(\tilde{p}\right),\hat{\lambda}\left(\tilde{p}\right),\hat{\eta}\left(\tilde{p}\right),\tilde{p}\right)\tilde{v} &\leq 0, \end{aligned}$$

which by (2.3.14) implies $\tilde{v} \in K^+\left(\tilde{p}\right)$, and $\tilde{v}^\top \nabla^2_{zz}\mathcal{L}\left(1,\hat{z}\left(\tilde{p}\right),\hat{\lambda}\left(\tilde{p}\right),\hat{\eta}\left(\tilde{p}\right),\tilde{p}\right)\tilde{v} \leq 0$. This contradicts condition (*ii*), which proves the assertion. □

Remark 2.3.13
If the strict complementarity condition $\hat{J}^0\left(p\right) = \emptyset$ holds, then the relation $\tilde{K}^\nu\left(p\right) \subseteq K\left(p\right)$ is satisfied, which allows us to distinguish between two cases:

$$v \in \tilde{K}^\nu\left(p\right) \quad \textit{and} \quad v \in K\left(p\right) \setminus \tilde{K}^\nu\left(p\right).$$

This is essential for the proof of Theorem 2.3.14, since we obtain the coercivity condition $v^\top \nabla^2_{zz}\mathcal{L}\left(1,\hat{z}\left(p\right),\hat{\lambda}\left(p\right),\hat{\eta}\left(p\right),p\right)v \geq \tilde{\gamma}\left\|v\right\|^2$, if $v \in \tilde{K}^\nu\left(p\right)$, and for the other case we get the lower bound $\nabla_z\mathcal{J}\left(\hat{z}\left(p\right),p\right)^\top v > \nu\left\|v\right\|$, which holds by definition. These properties are exploited in Taylor expansions of the functions $\mathcal{L}$ and $\mathcal{J}$, respectively, which permits us to prove optimality of $\hat{z}\left(p\right)$.

Since the constraints in Problem 2.3.11 are independent of the parameter we denote the set of admissible vectors by Σ, and prove the following uniform second-order sufficient conditions, which are fundamental for Chapter 4:

Theorem 2.3.14 (Uniform Second-Order Sufficient Conditions)
Let the functions H and G be twice continuously differentiable, let $\mathcal{J}\left(\cdot,p\right)$ be twice continuously differentiable for every $p \in \mathcal{P}$, and let $\mathcal{J}\left(z,\cdot\right)$, $\nabla_z\mathcal{J}\left(z,\cdot\right)$, and $\nabla^2_{zz}\mathcal{J}\left(z,\cdot\right)$ be continuous for every $z \in \mathbb{R}^{n_z}$. Let there exist continuous functions $\hat{z}\left(\cdot\right) : \mathcal{P} \to \mathbb{R}^{n_z}$, $\hat{\lambda}\left(\cdot\right) : \mathcal{P} \to \mathbb{R}^{n_H}$, $\hat{\eta}\left(\cdot\right) : \mathcal{P} \to \mathbb{R}^{n_G}$ such that $\left(\hat{z}\left(p\right),\hat{\lambda}\left(p\right),\hat{\eta}\left(p\right)\right)$ is a KKT-point of Problem 2.3.11 for every $p \in \mathcal{P}$. Moreover, let the following be satisfied:

(i) There exists a constant $\beta > 0$ such that for all $p \in \mathcal{P}$ and every $d \in \mathbb{R}^{n_H} \times \mathbb{R}^{\hat{j}^+(p)} \times \mathbb{R}^{\hat{j}^0(p)}$ it holds

$$\left\| \begin{bmatrix} \hat{A}_H\left(p\right) \\ \hat{B}^+_G\left(p\right) \\ \hat{B}^0_G\left(p\right) \end{bmatrix}^\top d \right\| \geq \beta\left\|d\right\|.$$

(ii) There exists a constant $\gamma > 0$ such that for all $p \in \mathcal{P}$ and every $v \in K^+\left(p\right)$ it holds

$$v^\top \nabla^2_{zz}\mathcal{L}\left(1,\hat{z}\left(p\right),\hat{\lambda}\left(p\right),\hat{\eta}\left(p\right),p\right)v \geq \gamma\left\|v\right\|^2.$$

(iii) The strict complementarity condition $\hat{J}^0\left(p\right) = \emptyset$ is only violated by finitely many $p \in \mathcal{P}$.

Then, there exist $\alpha, \rho > 0$ such that for all $p \in \mathcal{P}$ and every $z \in \Sigma \cap \mathcal{B}_\rho\left(\hat{z}\left(p\right)\right)$ it holds

$$\mathcal{J}\left(z,p\right) \geq \mathcal{J}\left(\hat{z}\left(p\right),p\right) + \alpha\left\|z - \hat{z}\left(p\right)\right\|^2.$$

Proof. According to the assumptions, the function $\nabla^2_{zz}\mathcal{L}\left(1,\hat{z}\left(p\right),\hat{\lambda}\left(p\right),\hat{\eta}\left(p\right),p\right)$ is continuous with respect to $p\in\mathcal{P}$ and $\mathcal{P}$ is compact. Thus, there exists a constant $\Gamma_{\mathcal{L}}>0$ such that

$$\max_{p\in\mathcal{P}}\left\|\nabla^2_{zz}\mathcal{L}\left(1,\hat{z}\left(p\right),\hat{\lambda}\left(p\right),\hat{\eta}\left(p\right),p\right)\right\|\leq\Gamma_{\mathcal{L}}.$$

By Lemma 2.3.12, there exist $\tilde{\gamma},\nu>0$ such that for all $p\in\mathcal{P}$ and every $v\in\tilde{K}^{\nu}\left(p\right)$ it holds

$$v^{\top}\nabla^2_{zz}\mathcal{L}\left(1,\hat{z}\left(p\right),\hat{\lambda}\left(p\right),\hat{\eta}\left(p\right),p\right)v\geq\tilde{\gamma}\left\|v\right\|^2.$$

It follows from Theorem 2.1.9 that there exist $\gamma_0,\Gamma_r>0$ depending only on $\tilde{\gamma},\Gamma_{\mathcal{L}}$ such that for all $p\in\mathcal{P}$, every $v\in\tilde{K}^{\nu}\left(p\right)$, and every $w\in\mathbb{R}^{n_z}$ with $\left\|w\right\|\leq\Gamma_r\left\|v\right\|$ it holds

$$\left(v+w\right)^{\top}\nabla^2_{zz}\mathcal{L}\left(1,\hat{z}\left(p\right),\hat{\lambda}\left(p\right),\hat{\eta}\left(p\right),p\right)\left(v+w\right)\geq\gamma_0\left\|v+w\right\|^2. \tag{2.3.16}$$

According to [88, Theorem 4.2], for every $p\in\mathcal{P}$ the set Σ is approximated at $\hat{z}\left(p\right)$ by the linearizing cone $K\left(p\right)$, i.e., there exists mappings $s:\Sigma\times\mathcal{P}\to\mathbb{R}^{n_z}$, $r:\Sigma\times\mathcal{P}\to\mathbb{R}^{n_z}$ such that for every $\left(z,p\right)\in\Sigma\times\mathcal{P}$ it holds

$$z-\hat{z}\left(p\right)=s\left(z,p\right)+r\left(z,p\right),\quad s\left(z,p\right)\in K\left(p\right),\ \lim_{z\to\hat{z}(p)}\frac{r\left(z,p\right)}{\left\|z-\hat{z}\left(p\right)\right\|}=0.$$

Let $p\in\mathcal{P}$ be arbitrary such that the strict complementarity condition $\hat{\mathcal{J}}^0\left(p\right)=\emptyset$ holds, hence $\tilde{K}^{\nu}\left(p\right)\subseteq K\left(p\right)$ is satisfied. As described in Remark 2.3.13 we distinguish between two cases:

$$s\left(z,p\right)\in\tilde{K}^{\nu}\left(p\right)\quad\text{and}\quad s\left(z,p\right)\in K\left(p\right)\setminus\tilde{K}^{\nu}\left(p\right).$$

Firstly, we assume $s\left(z,p\right)\in\tilde{K}^{\nu}\left(p\right)$. Then, by [88, Lemma 4.2], we can choose $\tilde{\rho}>0$ such that we obtain

$$\left\|r\left(z,p\right)\right\|=\left\|s\left(z,p\right)-\left(z-\hat{z}\left(p\right)\right)\right\|\leq\Gamma_r\left\|s\left(z,p\right)\right\|\quad\text{for all } z\in\mathcal{B}_{\tilde{\rho}}\left(\hat{z}\left(p\right)\right).$$

Consequently, by (2.3.16), it holds

$$\begin{aligned}\left(s\left(z,p\right)+r\left(z,p\right)\right)^{\top}\nabla^2_{zz}\mathcal{L}\left(1,\hat{z}\left(p\right),\hat{\lambda}\left(p\right),\hat{\eta}\left(p\right),p\right)\left(s\left(z,p\right)+r\left(z,p\right)\right)&\\ \geq\gamma_0\left\|s\left(z,p\right)+r\left(z,p\right)\right\|^2.&\end{aligned} \tag{2.3.17}$$

Since $\nabla^2_{zz}\mathcal{L}\left(1,z,\hat{\lambda}\left(p\right),\hat{\eta}\left(p\right),p\right)$ is continuous with respect to $z\in\mathbb{R}^{n_z}$ and $p\in\mathcal{P}$, there exists a $0<\rho\leq\tilde{\rho}$ such that for $p\in\mathcal{P}$ and every $z\in\mathcal{B}_{\rho}\left(\hat{z}\left(p\right)\right)$ it holds

$$\left\|\nabla^2_{zz}\mathcal{L}\left(1,z,\hat{\lambda}\left(p\right),\hat{\eta}\left(p\right),p\right)-\nabla^2_{zz}\mathcal{L}\left(1,\hat{z}\left(p\right),\hat{\lambda}\left(p\right),\hat{\eta}\left(p\right),p\right)\right\|\leq\frac{\gamma_0}{2}. \tag{2.3.18}$$

Exploiting the Taylor expansion for $\mathcal{L}\left(1,z,\hat{\lambda}\left(p\right),\hat{\eta}\left(p\right),p\right)$ at $\hat{z}\left(p\right)$ yields

$$\begin{aligned}\mathcal{L}\left(1,z,\hat{\lambda}\left(p\right),\hat{\eta}\left(p\right),p\right)&=\mathcal{L}\left(1,\hat{z}\left(p\right),\hat{\lambda}\left(p\right),\hat{\eta}\left(p\right),p\right)\\&\quad+\nabla_z\mathcal{L}\left(1,\hat{z}\left(p\right),\hat{\lambda}\left(p\right),\hat{\eta}\left(p\right),p\right)\left(z-\hat{z}\left(p\right)\right)\qquad(2.3.19)\\&\quad+\frac{1}{2}\left(z-\hat{z}\left(p\right)\right)^{\top}\nabla^2_{zz}\mathcal{L}\left(1,\xi\left(z\right),\hat{\lambda}\left(p\right),\hat{\eta}\left(p\right),p\right)\left(z-\hat{z}\left(p\right)\right)\\&\overset{(2.3.13)}{=}\mathcal{J}\left(\hat{z}\left(p\right),p\right)\\&\quad+\frac{1}{2}\left(z-\hat{z}\left(p\right)\right)^{\top}\nabla^2_{zz}\mathcal{L}\left(1,\xi\left(z\right),\hat{\lambda}\left(p\right),\hat{\eta}\left(p\right),p\right)\left(z-\hat{z}\left(p\right)\right),\end{aligned}$$

for every $z \in \Sigma \cap \mathcal{B}_\rho(\hat{z}(p))$ and a certain $\xi(z) \in \mathcal{B}_\rho(\hat{z}(p))$. Inspecting the quadratic term and utilizing (2.3.17) and (2.3.18) results in

$$\begin{aligned}
&(z-\hat{z}(p))^\top \nabla^2_{zz}\mathcal{L}\left(1,\xi(z),\hat{\lambda}(p),\hat{\eta}(p),p\right)(z-\hat{z}(p))\\
&\quad= (s(z,p)+r(z,p))^\top \nabla^2_{zz}\mathcal{L}\left(1,\xi(z),\hat{\lambda}(p),\hat{\eta}(p),p\right)(s(z,p)+r(z,p))\\
&\quad= (s(z,p)+r(z,p))^\top \nabla^2_{zz}\mathcal{L}\left(1,\hat{z}(p),\hat{\lambda}(p),\hat{\eta}(p),p\right)(s(z,p)+r(z,p))\\
&\qquad- (s(z,p)+r(z,p))^\top \left[\nabla^2_{zz}\mathcal{L}\left(1,\hat{z}(p),\hat{\lambda}(p),\hat{\eta}(p),p\right)\right.\\
&\qquad\left.- \nabla^2_{zz}\mathcal{L}\left(1,\xi(z),\hat{\lambda}(p),\hat{\eta}(p),p\right)\right](s(z,p)+r(z,p))\\
&\quad\geq \gamma_0\|s(z,p)+r(z,p)\|^2 - \frac{\gamma_0}{2}\|s(z,p)+r(z,p)\|^2 = \frac{\gamma_0}{2}\|z-\hat{z}(p)\|^2.
\end{aligned} \tag{2.3.20}$$

Furthermore, for every $z \in \Sigma \cap \mathcal{B}_\rho(\hat{z}(p))$ it holds

$$\mathcal{J}(z,p) \geq \mathcal{L}\left(1,z,\hat{\lambda}(p),\hat{\eta}(p),p\right),$$

since $H(z) = \mathbf{0}_{\mathbb{R}^{n_H}}$ and $\hat{\eta}(p)^\top G(z) \leq 0$. Consequently, using (2.3.19) and (2.3.20) yields

$$\mathcal{J}(z,p) \geq \mathcal{J}(\hat{z}(p),p) + \frac{\gamma_0}{4}\|z-\hat{z}(p)\|^2,$$

for every $z \in \Sigma \cap \mathcal{B}_\rho(\hat{z}(p))$.

Now, suppose $s(z,p) \in K(p) \setminus \tilde{K}^\nu(p)$, thus $\nabla_z\mathcal{J}(\hat{z}(p),p)^\top s(z,p) > \nu\|s(z,p)\|$. Since the function $\nabla_z\mathcal{J}(\hat{z}(\cdot),\cdot)$ is continuous with respect to p, there exists a constant $\Gamma_\mathcal{J} > 0$ such that $\Gamma_\mathcal{J} = \sup_{p\in\mathcal{P}}\|\nabla_z\mathcal{J}(\hat{z}(p),p)\|$. Select $0 < \rho < 1$ satisfying

$$\left\|\nabla_z\mathcal{J}(z,p)^\top - \nabla_z\mathcal{J}(\hat{z}(p),p)^\top\right\| \leq \frac{\nu}{2} \text{ and } \|r(z,p)\| \leq \frac{\nu}{4(\Gamma_\mathcal{J}+\nu)}\|z-\hat{z}(p)\|$$

for every $z \in \Sigma \cap \mathcal{B}_\rho(\hat{z}(p))$. Exploiting $\|s(z,p)\| \geq \|z-\hat{z}(p)\| - \|r(z,p)\|$ and the mean-value theorem for a certain $\xi(z) \in \mathcal{B}_\rho(\hat{z}(p))$ yields

$$\begin{aligned}
\mathcal{J}(z,p) - \mathcal{J}(\hat{z}(p),p) &= \nabla_z\mathcal{J}(\xi(z),p)^\top(z-\hat{z}(p))\\
&= \nabla_z\mathcal{J}(\hat{z}(p),p)^\top s(z,p) + \nabla_z\mathcal{J}(\hat{z}(p),p)^\top r(z,p)\\
&\quad- \left(\nabla_z\mathcal{J}(\hat{z}(p),p)^\top - \nabla_z\mathcal{J}(\xi(z),p)^\top\right)(z-\hat{z}(p))\\
&\geq \nu\|s(z,p)\| - \Gamma_\mathcal{J}\|r(z,p)\| - \frac{\nu}{2}\|z-\hat{z}(p)\|\\
&\geq \frac{\nu}{2}\|z-\hat{z}(p)\| - (\Gamma_\mathcal{J}+\nu)\|r(z,p)\|\\
&\geq \frac{\nu}{2}\|z-\hat{z}(p)\| - \frac{\nu(\Gamma_\mathcal{J}+\nu)}{4(\Gamma_\mathcal{J}+\nu)}\|z-\hat{z}(p)\|\\
&= \frac{\nu}{4}\|z-\hat{z}(p)\| \geq \frac{\nu}{4}\|z-\hat{z}(p)\|^2.
\end{aligned}$$

We take the minimum radius ρ of both cases and set $\alpha := \min\left\{\frac{\gamma_0}{4},\frac{\nu}{4}\right\}$, which yields

$$\mathcal{J}(z,p) \geq \mathcal{J}(\hat{z}(p),p) + \alpha\|z-\hat{z}(p)\|^2 \quad \text{for every } z \in \Sigma \cap \mathcal{B}_\rho(\hat{z}(p)).$$

According to (iii), this relation is only violated by finitely many (isolated) points $p \in \mathcal{P}$. Suppose there exists $\tilde{p} \in \mathcal{P}$ such that there exists a $\tilde{z} \in \Sigma \cap \mathcal{B}_\rho\left(\hat{z}\left(\tilde{p}\right)\right)$ with

$$\mathcal{J}\left(\tilde{z}, \tilde{p}\right) - \mathcal{J}\left(\hat{z}\left(\tilde{p}\right), \tilde{p}\right) - \alpha \left\|\tilde{z} - \hat{z}\left(\tilde{p}\right)\right\|^2 < 0. \tag{2.3.21}$$

Then, there exists a convergent sequence $\left(\left(p_i, z_i\right)\right)_{i\in\mathbb{N}} \subseteq \operatorname{graph}\left(\Sigma \cap \mathcal{B}_\rho\left(\hat{z}\left(\cdot\right)\right)\right)$ with limit $\left(\tilde{p}, \tilde{z}\right)$, and

$$\mathcal{J}\left(z_i, p_i\right) - \mathcal{J}\left(\hat{z}\left(p_i\right), p_i\right) - \alpha \left\|z_i - \hat{z}\left(p_i\right)\right\|^2 \geq 0.$$

Since left-hand side is continuous with respect to (p, z), taking the limits yields

$$\mathcal{J}\left(\tilde{z}, \tilde{p}\right) - \mathcal{J}\left(\hat{z}\left(\tilde{p}\right), \tilde{p}\right) - \alpha \left\|\tilde{z} - \hat{z}\left(\tilde{p}\right)\right\|^2 \geq 0,$$

which contradicts (2.3.21). This completes the proof. □

2.4 Linear Time-Variant Differential-Algebraic Equations

An in-depth analysis of linear, nonlinear, and other types of DAEs was covered in the textbooks [66, 71] and the references therein. We limit our investigations to linear, time-variant DAEs in semi-explicit form

$$\begin{aligned}
\dot{z}\left(t\right) &= A\left(t\right)z\left(t\right) + B\left(t\right)v\left(t\right), && \text{a.e. in } [0,1],\\
\mathbf{0}_{\mathbb{R}^m} &= C\left(t\right)z\left(t\right) + D\left(t\right)v\left(t\right), && \text{a.e. in } [0,1],\\
\mathbf{0}_{\mathbb{R}^{n_E}} &= E_0 z\left(0\right) + E_1 z\left(1\right),
\end{aligned} \tag{2.4.1}$$

where $z \in W_{1,p}^{n_z}\left([0,1]\right)$ is the differential state and $v \in L_p^{n_v}\left([0,1]\right)$ is the control $(p = 2, \infty)$, which is partly determined by the algebraic equation. For $n_z, n_v, m, n_E \in \mathbb{N}$ with $m \leq n_v, n_E \leq 2n_z$ the matrix functions have the dimensions

$$\begin{aligned}
&A \in L_\infty^{n_z\times n_z}\left([0,1]\right), && B \in L_\infty^{n_z\times n_v}\left([0,1]\right),\\
&C \in L_\infty^{m\times n_z}\left([0,1]\right), && D \in L_\infty^{m\times n_v}\left([0,1]\right),\\
&E_0, E_1 \in \mathbb{R}^{n_E\times n_z}.
\end{aligned}$$

We introduce the following terminology for system (2.4.1):

Definition 2.4.1 (Uniform Linear Independence, Controllability)

(i) The matrix function $D\left(\cdot\right)$ in system (2.4.1) is uniformly linear independent, if there exists a constant $\beta > 0$ such that for almost every $t \in [0,1]$ and for all $\varpi \in \mathbb{R}^m$ it holds

$$\left\|D\left(t\right)^\top \varpi\right\| \geq \beta \left\|\varpi\right\|.$$

(ii) System (2.4.1) is completely controllable, if for every $e \in \mathbb{R}^{n_E}$ there exist $(z, v) \in W_{1,p}^{n_z}\left([0,1]\right) \times L_p^{n_v}\left([0,1]\right)$ satisfying

$$\begin{aligned}
\dot{z}\left(t\right) &= A\left(t\right)z\left(t\right) + B\left(t\right)v\left(t\right), && \textit{a.e. in } [0,1],\\
\mathbf{0}_{\mathbb{R}^m} &= C\left(t\right)z\left(t\right) + D\left(t\right)v\left(t\right), && \textit{a.e. in } [0,1],\\
e &= E_0 z\left(0\right) + E_1 z\left(1\right).
\end{aligned}$$

Remark 2.4.2 (Reduced System)
If the matrix function $D(\cdot)$ in system (2.4.1) is uniformly linear independent, then for almost every $t \in [0,1]$ and all $\varpi \in \mathbb{R}^m$ it holds

$$\varpi^\top D(t) D(t)^\top \varpi = \left\| D(t)^\top \varpi \right\|^2 \geq \beta^2 \|\varpi\|^2 .$$

Consequently, for almost every $t \in [0,1]$ the matrix $D(t) D(t)^\top \in \mathbb{R}^{m \times m}$ is uniformly positive definite, hence the inverse exists and is uniformly bounded by $\frac{1}{\beta^2}$ (see Lemma 2.1.6). Therefore, the right inverse $D(\cdot)^\curlywedge := D(\cdot)^\top \left(D(\cdot) D(\cdot)^\top \right)^{-1}$ is essentially bounded. For almost every $t \in [0,1]$ consider the inhomogeneous linear equation

$$D(t) v = b, \quad v \in \mathbb{R}^{n_v}, b \in \mathbb{R}^m, \tag{2.4.2}$$

where $D(t)$ has full row rank. The general solution of this system is the sum of a particular solution and the general solution of the homogeneous system. For almost every $t \in [0,1]$ the linear mapping

$$\left(\mathbf{I}_{n_v} - D(t)^\curlywedge D(t) \right) : \mathbb{R}^{n_v} \to \ker(D(t))$$

is surjective, since for every $b \in \ker(D(t))$ it holds

$$\left(\mathbf{I}_{n_v} - D(t)^\curlywedge D(t) \right) b = b - D(t)^\curlywedge D(t) b = b.$$

Thus, for almost every $t \in [0,1]$ the general solution of the homogeneous system can be expressed by $\left(\mathbf{I}_{n_v} - D(t)^\curlywedge D(t) \right) w$ for $w \in \mathbb{R}^{n_v}$. Moreover, $D(t)^\curlywedge b$ is a particular solution of the inhomogeneous system for almost every $t \in [0,1]$. We conclude that the general solution of (2.4.2) is determined by

$$v = D(t)^\curlywedge b + \left(\mathbf{I}_{n_v} - D(t)^\curlywedge D(t) \right) w$$

for $w \in \mathbb{R}^{n_v}$ and almost every $t \in [0,1]$. Hence, if $D(\cdot)$ is uniformly linear independent, then we are able to write the control $v(\cdot)$ as

$$v(\cdot) = -D(\cdot)^\curlywedge C(\cdot) z(\cdot) + \left(\mathbf{I}_{n_v} - D(\cdot)^\curlywedge D(\cdot) \right) w(\cdot)$$

for an arbitrary $w \in L_p^{n_v}([0,1])$. Inserting this expression into (2.4.1) yields the reduced system

$$\begin{aligned} \dot{z}(t) &= \tilde{A}(t) z(t) + \tilde{B}(t) w(t), \qquad \textit{a.e. in } [0,1], \\ \mathbf{0}_{\mathbb{R}^{n_E}} &= E_0 z(0) + E_1 z(1), \end{aligned} \tag{2.4.3}$$

with the notation

$$\begin{aligned} \tilde{A}(\cdot) &:= A(\cdot) - B(\cdot) D(\cdot)^\curlywedge C(\cdot) \in L_\infty^{n_z \times n_z}([0,1]), \\ \tilde{B}(\cdot) &:= B(\cdot) \left(\mathbf{I}_{n_v} - D(\cdot)^\curlywedge D(\cdot) \right) \in L_\infty^{n_z \times n_v}([0,1]). \end{aligned} \tag{2.4.4}$$

We denote the solution of the matrix differential equation

$$\dot{\Phi}(t) = A(t) \Phi(t), \qquad \text{a.e. in } [0,1], \qquad \Phi(0) = \mathbf{I}_{n_z},$$

by $\Phi_A \in W_{1,\infty}^{n_z \times n_z}([0,1])$, and use the abbreviations

$$\begin{aligned} R &:= E_0 + E_1 \Phi_{\tilde{A}}(1) \in \mathbb{R}^{n_E \times n_z}, \\ S(\cdot) &:= E_1 \Phi_{\tilde{A}}(1) \Phi_{\tilde{A}}(\cdot)^{-1} \tilde{B}(\cdot) \in L_\infty^{n_E \times n_v}([0,1]), \\ G &:= RR^\top + \int_0^1 S(t) S(t)^\top dt \in \mathbb{R}^{n_E \times n_E}, \end{aligned} \tag{2.4.5}$$

to prove the following relation between *controllability* and the *Gramian matrix* G:

Lemma 2.4.3 (Controllability, Gramian Matrix)
If the matrix function $D(\cdot)$ in system (2.4.1) is uniformly linear independent, then the following holds: System (2.4.1) is completely controllable, if and only if $\operatorname{rank}(G) = n_E$.

Proof. For $e \in \mathbb{R}^{n_E}$ we consider the inhomogeneous system

$$\begin{aligned} \dot{z}(t) &= A(t) z(t) + B(t) v(t), && \text{a.e. in } [0,1], \\ \mathbf{0}_{\mathbb{R}^m} &= C(t) z(t) + D(t) v(t), && \text{a.e. in } [0,1], \\ e &= E_0 z(0) + E_1 z(1). \end{aligned}$$

Analog to Remark 2.4.2, we obtain the reduced system

$$\begin{aligned} \dot{z}(t) &= \tilde{A}(t) z(t) + \tilde{B}(t) w(t), && \text{a.e. in } [0,1], \\ e &= E_0 z(0) + E_1 z(1), \end{aligned}$$

for a $w \in L_p^{n_v}([0,1])$ with

$$v(\cdot) = -D(\cdot)^{\wedge} C(\cdot) z(\cdot) + \left(\mathbf{I}_{n_v} - D(\cdot)^{\wedge} D(\cdot)\right) w(\cdot).$$

This differential equation has the solution

$$z(\cdot) = \Phi_{\tilde{A}}(\cdot) z(0) + \Phi_{\tilde{A}}(\cdot) \int_0^{\cdot} \Phi_{\tilde{A}}(\tau)^{-1} \tilde{B}(\tau) w(\tau) d\tau,$$

which allows us to write the boundary conditions as

$$\begin{aligned} e &= \left(E_0 + E_1 \Phi_{\tilde{A}}(1)\right) z(0) + E_1 \Phi_{\tilde{A}}(1) \int_0^1 \Phi_{\tilde{A}}(\tau)^{-1} \tilde{B}(\tau) w(\tau) d\tau \\ &= Rz(0) + \int_0^1 S(\tau) w(\tau) d\tau. \end{aligned} \tag{2.4.6}$$

First, suppose (2.4.1) is completely controllable and $\operatorname{rank}(G) < n_E$. Then, there exists a $e \in \mathbb{R}^{n_E} \setminus \{\mathbf{0}_{\mathbb{R}^{n_E}}\}$ with $Ge = \mathbf{0}_{\mathbb{R}^{n_E}}$. Additionally, it holds

$$0 = e^\top G e = e^\top R R^\top e + \int_0^1 e^\top S(t) S(t)^\top e \, dt = \left\| R^\top e \right\|^2 + \left\| S(\cdot)^\top e \right\|_2^2,$$

which implies $0 = \left\|R^\top e\right\| = \left\|S(\cdot)^\top e\right\|_2$. Using Hölder's inequality and the rewritten boundary conditions in (2.4.6) yields

$$\begin{aligned} 0 &\leq \|e\|^2 = e^\top e \overset{(2.4.6)}{=} e^\top R z(0) + \int_0^1 e^\top S(\tau)\, w(\tau)\, d\tau \\ &\overset{\text{Hölder}}{\leq} \left\|R^\top e\right\| \|z(0)\| + \left\|S(\cdot)^\top e\right\|_2 \|w\|_2 = 0. \end{aligned}$$

Thus, $\|e\| = 0$, which contradicts $e \in \mathbb{R}^{n_E} \setminus \{\mathbf{0}_{\mathbb{R}^{n_E}}\}$.
Now, we assume $\operatorname{rank}(G) = n_E$, hence the inverse exists. For an arbitrary $e \in \mathbb{R}^{n_E}$ set

$$\begin{aligned} z(0) &= R^\top G^{-1} e, \\ w(\cdot) &= S(\cdot)^\top G^{-1} e \in L_p^{n_v}([0,1]). \end{aligned}$$

These satisfy the boundary conditions in (2.4.6), since

$$Rz(0) + \int_0^1 S(\tau)\, w(\tau)\, d\tau = \left(RR^\top + \int_0^1 S(\tau)\, S(\tau)^\top d\tau\right) G^{-1} e = GG^{-1} e = e.$$

Then, $(z, v) \in W_{1,p}^{n_z}([0,1]) \times L_p^{n_v}([0,1])$ with

$$\begin{aligned} z(\cdot) &= \Phi_{\tilde{A}}(\cdot)\, z(0) + \Phi_{\tilde{A}}(\cdot) \int_0^{\cdot} \Phi_{\tilde{A}}(\tau)^{-1} \tilde{B}(\tau)\, w(\tau)\, d\tau, \\ v(\cdot) &= -D(\cdot)^{\curlywedge} C(\cdot)\, z(\cdot) + \left(\mathbf{I}_{n_v} - D(\cdot)^{\curlywedge} D(\cdot)\right) w(\cdot), \end{aligned}$$

satisfy

$$\begin{aligned} \dot{z}(t) &= A(t)\, z(t) + B(t)\, v(t), && \text{a.e. in } [0,1], \\ \mathbf{0}_{\mathbb{R}^m} &= C(t)\, z(t) + D(t)\, v(t), && \text{a.e. in } [0,1], \\ e &= E_0 z(0) + E_1 z(1), \end{aligned}$$

which completes the proof. □

In the sequel, we consider an approximation of the linear system (2.4.1). To that end, let $\mathbb{G}_N := \{0 = t_0 < t_1 < t_2 < \ldots < t_{N-1} < t_N = 1\}$ be a grid of $[0,1]$ with $t_i := ih$, $i = 0, 1, \ldots, N$, $N \in \mathbb{N}$, and the mesh size $h := \frac{1}{N}$. For $i = 1, \ldots, N$ let us denote the *discrete derivative* (backwards difference approximation) at t_i by $u'(t_i) := \frac{u(t_i) - u(t_{i-1})}{h}$. Furthermore, for $p = 2, \infty$ we define the finite dimensional subspaces

$$\begin{aligned} L_{p,h}^n([0,1]) &:= \left\{u \in L_p^n([0,1]) \mid u(t) = u(t_i),\, t \in (t_{i-1}, t_i],\, i = 1, \ldots, N\right\}, \\ W_{1,p,h}^n([0,1]) &:= \Big\{u \in W_{1,p}^n([0,1]) \mid u(t) = u'(t_i)(t - t_{i-1}) + u(t_{i-1}), \\ &\qquad t \in (t_{i-1}, t_i],\, i = 1, \ldots, N\Big\}, \end{aligned}$$

where $L_{p,h}^n([0,1])$ consist of piece-wise constant functions, and $W_{1,p,h}^n([0,1])$ consists of piece-wise linear, continuous functions (compare Figure 2.1).

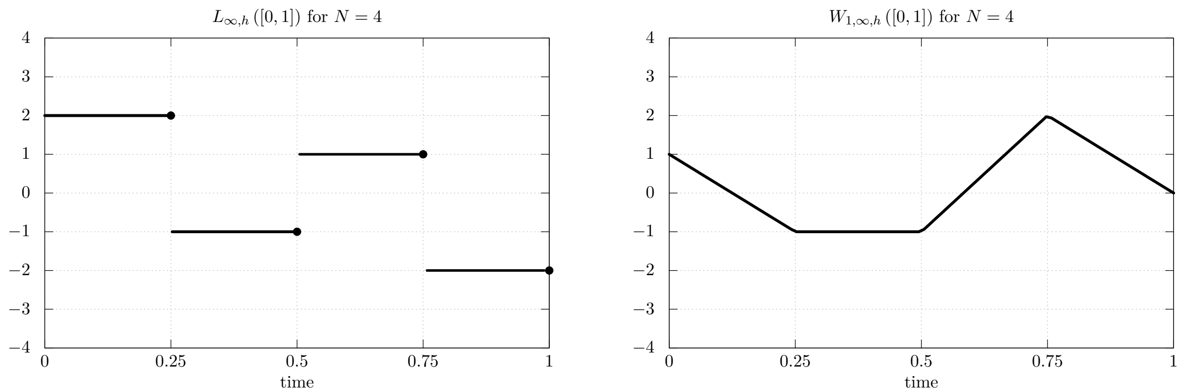

Figure 2.1: Illustration of elements in $L^n_{\infty,h}([0,1])$ and $W^n_{1,\infty,h}([0,1])$.

Consider the following time-discrete system

$$\begin{aligned} z_h'(t_i) &= A_h(t_i) z_h(t_i) + B_h(t_i) v_h(t_i), \qquad i = 1,\dots,N, \\ \mathbf{0}_{\mathbb{R}^m} &= C_h(t_i) z_h(t_i) + D_h(t_i) v_h(t_i), \qquad i = 1,\dots,N, \\ \mathbf{0}_{\mathbb{R}^{n_E}} &= E_{0,h} z_h(t_0) + E_{1,h} z_h(t_N), \end{aligned} \tag{2.4.7}$$

where

$$\begin{aligned} &z_h \in W^{n_z}_{1,p,h}([0,1]), && v_h \in L^{n_v}_{p,h}([0,1]), \\ &A_h \in L^{n_z \times n_z}_{\infty,h}([0,1]), && B_h \in L^{n_z \times n_v}_{\infty,h}([0,1]), \\ &C_h \in L^{m \times n_z}_{\infty,h}([0,1]), && D_h \in L^{m \times n_v}_{\infty,h}([0,1]), \\ &E_{0,h}, E_{1,h} \in \mathbb{R}^{n_E \times n_z}. \end{aligned}$$

We will refer to (2.4.1) and (2.4.7) as the continuous system and the discrete system, respectively. Approximations like (2.4.7) occur in Chapter 5, where we apply the implicit Euler discretization to an optimal control problem subject to a DAE. Our aim is to show that, under appropriate assumptions, the properties in Definition 2.4.1 and Lemma 2.4.3 for the continuous system are inherited by the discrete system for sufficiently small mesh size h. To that end, we assume the following:

Assumption 2.4.4 (Data Approximation)
There exists a constant $\mathbf{L} > 0$ *such that for all* $h > 0$ *is holds*

$$\begin{aligned} &\|A(\cdot) - A_h(\cdot)\|_\infty \le \mathbf{L}h, && \|B(\cdot) - B_h(\cdot)\|_\infty \le \mathbf{L}h, \\ &\|C(\cdot) - C_h(\cdot)\|_\infty \le \mathbf{L}h, && \|D(\cdot) - D_h(\cdot)\|_\infty \le \mathbf{L}h, \\ &\|E_0 - E_{0,h}\| \le \mathbf{L}h, && \|E_1 - E_{1,h}\| \le \mathbf{L}h. \end{aligned}$$

This condition allows us to prove, that $D_h(\cdot)$ is uniformly linear independent for sufficiently small h.

Lemma 2.4.5 (Discrete Linear Independence)
Let $D(\cdot)$ be uniformly linear independent with constant $\beta > 0$ and let Assumption 2.4.4 hold. Then there exists $h_1 > 0$ such that:

(i) For all $0 < h \leq h_1$: $D_h(\cdot)$ is uniformly linear independent with constant $\frac{\beta}{2}$.

(ii) There exists a constant $\mathbf{L}_D > 0$ such that for all $0 < h \leq h_1$ it holds

$$\left\| D(\cdot)^{\curlywedge} - D_h(\cdot)^{\curlywedge} \right\|_{\infty} \leq \mathbf{L}_D h$$

Proof.

Set $h_1 := \frac{\beta}{2\mathbf{L}}$ and let $0 < h \leq h_1$ be arbitrary.

(i) For almost every $t \in [0,1]$ and for all $\varpi \in \mathbb{R}^m$ it holds

$$\begin{aligned} \left\| D_h(t)^{\top} \varpi \right\| &= \left\| D(t)^{\top} \varpi - \left(D(t)^{\top} - D_h(t)^{\top} \right) \varpi \right\| \\ &\geq \left\| D(t)^{\top} \varpi \right\| - \left\| D(t)^{\top} - D_h(t)^{\top} \right\| \|\varpi\| \\ &\geq \beta \|\varpi\| - \mathbf{L}h \|\varpi\| \geq \frac{\beta}{2} \|\varpi\|, \end{aligned}$$

which proves the assertion.

(ii) Analog to Remark 2.4.2, we show that for almost every $t \in [0,1]$ the matrix $D_h(t) D_h(t)^{\top}$ is non-singular and the inverse is essentially bounded by $\frac{4}{\beta^2}$. Then, for almost every $t \in [0,1]$ we obtain

$$\begin{aligned} &\left\| D(t) D(t)^{\top} - D_h(t) D_h(t)^{\top} \right\| \\ &\leq \|D(t) - D_h(t)\| \left\| D(t)^{\top} \right\| + \|D_h(t)\| \left\| D(t)^{\top} - D_h(t)^{\top} \right\| \\ &\leq \left(\|D\|_{\infty} + \|D_h\|_{\infty} \right) \mathbf{L}h. \end{aligned}$$

Using Lemma A.2 we conclude

$$\begin{aligned} &\left\| \left(D(t) D(t)^{\top} \right)^{-1} - \left(D_h(t) D_h(t)^{\top} \right)^{-1} \right\| \\ &\leq \left\| \left(D(t) D(t)^{\top} \right)^{-1} \right\| \left\| \left(D_h(t) D_h(t)^{\top} \right)^{-1} \right\| \left\| D(t) D(t)^{\top} - D_h(t) D_h(t)^{\top} \right\| \\ &\leq \frac{1}{\beta^2} \frac{4}{\beta^2} \left(\|D\|_{\infty} + \|D_h\|_{\infty} \right) \mathbf{L}h \end{aligned}$$

for almost every $t \in [0,1]$. Therefore, it holds

$$
\begin{aligned}
&\left\| D(t)^{\curlywedge} - D_h(t)^{\curlywedge} \right\| \\
&= \left\| D(t)^\top \left(D(t)\, D(t)^\top \right)^{-1} - D_h(t)^\top \left(D_h(t)\, D_h(t)^\top \right)^{-1} \right\| \\
&\leq \left\| D(t)^\top - D_h(t)^\top \right\| \left\| \left(D(t)\, D(t)^\top \right)^{-1} \right\| \\
&\quad + \left\| D_h(t)^\top \right\| \left\| \left(D(t)\, D(t)^\top \right)^{-1} - \left(D_h(t)\, D_h(t)^\top \right)^{-1} \right\| \\
&\leq \mathbf{L}h \frac{1}{\beta^2} + \|D_h\|_\infty \frac{4}{\beta^4} \left(\|D\|_\infty + \|D_h\|_\infty \right) \mathbf{L}h
\end{aligned}
$$

for almost every $t \in [0,1]$, hence (ii) is satisfied for

$$
\mathbf{L}_D := \frac{\mathbf{L}}{\beta^4} \left[\beta^2 + 4 \|D_h\|_\infty \left(\|D\|_\infty + \|D_h\|_\infty \right) \right].
$$

□

In order to prove controllability for the discrete system (2.4.7), we show that the associated Gramian matrix has full rank. To that end, we prove that the matrices contained in the definition of the Gramian matrix satisfy a condition as in Assumption 2.4.4 with respect to their continuous counterparts in (2.4.4) and (2.4.5).

If the conditions of Lemma 2.4.5 hold, then we are able to reduce the discrete system (2.4.7) analog to the continuous case in Remark 2.4.2 to

$$
\begin{aligned}
z_h'(t_i) &= \tilde{A}_h(t_i)\, z_h(t_i) + \tilde{B}_h(t_i)\, w_h(t_i), \qquad i = 1, \ldots, N, \\
\mathbf{0}_{\mathbb{R}^{n_E}} &= E_{0,h} z_h(t_0) + E_{1,h} z_h(t_N),
\end{aligned}
\tag{2.4.8}
$$

where $w_h \in L_{p,h}^{n_v}([0,1])$ and

$$
\begin{aligned}
\tilde{A}_h(\cdot) &:= A_h(\cdot) - B_h(\cdot)\, D_h(\cdot)^{\curlywedge} C_h(\cdot) \in L_{\infty,h}^{n_z \times n_z}([0,1]), \\
\tilde{B}_h(\cdot) &:= B_h(\cdot) \left(\mathbf{I}_{n_v} - D_h(\cdot)^{\curlywedge} D_h(\cdot) \right) \in L_{\infty,h}^{n_z \times n_v}([0,1]),
\end{aligned}
$$

which retain the property of Assumption 2.4.4.

Lemma 2.4.6 (Reduced Data Approximation)
Let $D(\cdot)$ be uniformly linear independent with constant $\beta > 0$ and let Assumption 2.4.4 hold. Then, there exists $h_1 > 0$ and $\tilde{\mathbf{L}} \geq 0$ such that for all $0 < h \leq h_1$ it holds

$$
\left\| \tilde{A}(\cdot) - \tilde{A}_h(\cdot) \right\|_\infty \leq \tilde{\mathbf{L}} h, \qquad \left\| \tilde{B}(\cdot) - \tilde{B}_h(\cdot) \right\|_\infty \leq \tilde{\mathbf{L}} h.
$$

Proof. Set $h_1 := \frac{\beta}{2\mathbf{L}}$ and let $0 < h \leq h_1$ be arbitrary. Then, for almost every $t \in [0,1]$

$$\begin{aligned}
\left\|\tilde{A}(t) - \tilde{A}_h(t)\right\| &= \left\|A(t) - B(t)\,D(t)^{\curlywedge}\,C(t) - A_h(t) + B_h(t)\,D_h(t)^{\curlywedge}\,C_h(t)\right\| \\
&\leq \|A(t) - A_h(t)\| + \left\|B(t)\,D(t)^{\curlywedge}\,C(t) - B_h(t)\,D_h(t)^{\curlywedge}\,C_h(t)\right\| \\
&\overset{\text{Lemma}}{\underset{A.1}{\leq}} \mathbf{L}h + \|B(t) - B_h(t)\|\left\|D(t)^{\curlywedge}\right\|\|C(t)\| \\
&\quad + \left\|D(t)^{\curlywedge} - D_h(t)^{\curlywedge}\right\|\|B_h(t)\|\|C(t)\| \\
&\quad + \|C(t) - C_h(t)\|\|B_h(t)\|\left\|D_h(t)^{\curlywedge}\right\| \\
&\overset{\text{Lemma}}{\underset{2.4.5}{\leq}} \mathbf{L}h + \mathbf{L}h\left\|D(\cdot)^{\curlywedge}\right\|_\infty\|C(\cdot)\|_\infty + \mathbf{L}_D h\|B_h(\cdot)\|_\infty\|C(\cdot)\|_\infty \\
&\quad + \mathbf{L}h\|B_h(\cdot)\|_\infty\left\|D_h(\cdot)^{\curlywedge}\right\|_\infty,
\end{aligned}$$

is satisfied, and furthermore for almost every $t \in [0,1]$

$$\begin{aligned}
\left\|\tilde{B}(t) - \tilde{B}_h(t)\right\| &= \left\|B(t) - B(t)\,D(t)^{\curlywedge}\,D(t) - B_h(t) + B_h(t)\,D_h(t)^{\curlywedge}\,D_h(t)\right\| \\
&\leq \|B(t) - B_h(t)\| + \left\|B(t)\,D(t)^{\curlywedge}\,D(t) - B_h(t)\,D_h(t)^{\curlywedge}\,D_h(t)\right\| \\
&\overset{\text{Lemma}}{\underset{A.1}{\leq}} \mathbf{L}h + \|B(t) - B_h(t)\|\left\|D(t)^{\curlywedge}\right\|\|D(t)\| \\
&\quad + \left\|D(t)^{\curlywedge} - D_h(t)^{\curlywedge}\right\|\|B_h(t)\|\|D(t)\| \\
&\quad + \|D(t) - D_h(t)\|\|B_h(t)\|\left\|D_h(t)^{\curlywedge}\right\| \\
&\overset{\text{Lemma}}{\underset{2.4.5}{\leq}} \mathbf{L}h + \mathbf{L}h\left\|D(\cdot)^{\curlywedge}\right\|_\infty\|D(\cdot)\|_\infty + \mathbf{L}_D h\|B_h(\cdot)\|_\infty\|D(\cdot)\|_\infty \\
&\quad + \mathbf{L}h\|B_h(\cdot)\|_\infty\left\|D_h(\cdot)^{\curlywedge}\right\|_\infty.
\end{aligned}$$

Thus, the assertion holds for

$$\begin{aligned}
\tilde{\mathbf{L}} := \max\Big\{&\mathbf{L} + \mathbf{L}\left\|D(\cdot)^{\curlywedge}\right\|_\infty\|C(\cdot)\|_\infty + \mathbf{L}_D\|B_h(\cdot)\|_\infty\|C(\cdot)\|_\infty + \mathbf{L}\|B_h(\cdot)\|_\infty\left\|D_h(\cdot)^{\curlywedge}\right\|_\infty, \\
&\mathbf{L} + \mathbf{L}\left\|D(\cdot)^{\curlywedge}\right\|_\infty\|D(\cdot)\|_\infty + \mathbf{L}_D\|B_h(\cdot)\|_\infty\|D(\cdot)\|_\infty + \mathbf{L}\|B_h(\cdot)\|_\infty\left\|D_h(\cdot)^{\curlywedge}\right\|_\infty\Big\}.
\end{aligned}$$

□

We denote the function $\Phi : [0,1] \to \mathbb{R}^{n_z \times n_z}$ with

$$\Phi(t) = \Phi'(t_i)(t - t_{i-1}) + \Phi(t_{i-1}) \quad \text{for } \; t \in (t_{i-1}, t_i]\,,\; i = 1, \ldots, N,$$

which satisfies the matrix difference equation

$$\begin{aligned}
\Phi'(t_i) &= \tilde{A}_h(t_i)\,\Phi(t_i)\,, \qquad i = 1, \ldots, N, \\
\Phi(t_0) &= \mathbf{I}_{n_z},
\end{aligned}$$

by $\Phi_{\tilde{A}_h}(\cdot)$. Then, for $t \in (t_{i-1}, t_i]\,,\; i = 1, \ldots, N$ it holds

$$\begin{aligned}
\Phi_{\tilde{A}_h}(t) &= \Phi'_{\tilde{A}_h}(t_i)(t - t_{i-1}) + \Phi_{\tilde{A}_h}(t_{i-1}) \\
&= \tilde{A}_h(t_i)\,\Phi_{\tilde{A}_h}(t_i)(t - t_{i-1}) + \left(\mathbf{I}_{n_z} - h\tilde{A}_h(t_i)\right)\Phi_{\tilde{A}_h}(t_i) \qquad (2.4.9) \\
&= \left(\mathbf{I}_{n_z} - (t_i - t)\,\tilde{A}_h(t_i)\right)\Phi_{\tilde{A}_h}(t_i)\,.
\end{aligned}$$

Applying Lemma A.5 for $h \leq \frac{1}{2\left\|\tilde{A}_h\right\|_\infty}$ yields

$$
\begin{aligned}
\left\|\mathit{\Phi}_{\tilde{A}_h}(t)\right\| &\leq \left\|\left(\mathbf{I}_{n_z} - (t_i - t)\,\tilde{A}_h(t_i)\right)\right\| \left\|\mathit{\Phi}_{\tilde{A}_h}(t_i)\right\| \\
&\leq \left(1 + h\left\|\tilde{A}_h\right\|_\infty\right)\exp\left(2\left\|\tilde{A}_h\right\|_\infty\right) \leq \frac{3}{2}\exp\left(2\left\|\tilde{A}_h\right\|_\infty\right), \\
\left\|\mathit{\Phi}_{\tilde{A}_h}(t)^{-1}\right\| &\leq \left\|\left(\mathbf{I}_{n_z} - (t_i - t)\,\tilde{A}_h(t_i)\right)^{-1}\right\| \left\|\mathit{\Phi}_{\tilde{A}_h}(t_i)^{-1}\right\| \qquad (2.4.10)\\
&\leq \frac{1}{1 - h\left\|\tilde{A}_h\right\|_\infty}\exp\left(\left\|\tilde{A}_h\right\|_\infty\right) \leq 2\exp\left(\left\|\tilde{A}_h\right\|_\infty\right), \\
\left\|\dot{\mathit{\Phi}}_{\tilde{A}_h}(t)\right\| &\leq \left\|\tilde{A}_h(t_i)\right\| \left\|\mathit{\Phi}_{\tilde{A}_h}(t_i)\right\| \leq \left\|\tilde{A}_h\right\|_\infty \exp\left(2\left\|\tilde{A}_h\right\|_\infty\right)
\end{aligned}
$$

for $t \in (t_{i-1}, t_i]$, $i = 1, \ldots, N$, hence $\mathit{\Phi}_{\tilde{A}_h}(\cdot) \in W^{n_z \times n_z}_{1,\infty,h}([0,1])$.

Lemma 2.4.7 (Solution Matrix Difference)

Let $D(\cdot)$ be uniformly linear independent with constant $\beta > 0$ and let Assumption 2.4.4 hold. Then, there exists $h_1 > 0$ and $\mathbf{L}_\Phi \geq 0$ such that for all $0 < h \leq h_1$ it holds

$$
\left\|\Phi_{\tilde{A}}(\cdot) - \mathit{\Phi}_{\tilde{A}_h}(\cdot)\right\|_\infty \leq \mathbf{L}_\Phi h, \qquad \left\|\Phi_{\tilde{A}}(\cdot)^{-1} - \mathit{\Phi}_{\tilde{A}_h}(\cdot)^{-1}\right\|_\infty \leq \mathbf{L}_\Phi h.
$$

Proof. Set $h_1 := \min\left\{\frac{\beta}{2\mathbf{L}}, \frac{1}{2\left\|\tilde{A}_h\right\|_\infty}\right\}$ and let $0 < h \leq h_1$ be arbitrary. Then, for all $i = 1, \ldots, N$ and $t \in (t_{i-1}, t_i]$ we obtain

$$
\begin{aligned}
\Phi_{\tilde{A}}(t) - \mathit{\Phi}_{\tilde{A}_h}(t) &= \Phi_{\tilde{A}}(t_0) - \mathit{\Phi}_{\tilde{A}_h}(t_0) + \int_{t_0}^{t} \frac{d}{d\tau}\left(\Phi_{\tilde{A}}(\tau) - \mathit{\Phi}_{\tilde{A}_h}(\tau)\right) d\tau \\
&= \sum_{k=1}^{i-1} \int_{t_{k-1}}^{t_k} \frac{d}{d\tau}\left(\Phi_{\tilde{A}}(\tau) - \mathit{\Phi}_{\tilde{A}_h}(\tau)\right) d\tau + \int_{t_{i-1}}^{t} \frac{d}{d\tau}\left(\Phi_{\tilde{A}}(\tau) - \mathit{\Phi}_{\tilde{A}_h}(\tau)\right) d\tau \\
&= \sum_{k=1}^{i-1} \int_{t_{k-1}}^{t_k} \tilde{A}(\tau)\,\Phi_{\tilde{A}}(\tau) - \tilde{A}_h(t_k)\,\mathit{\Phi}_{\tilde{A}_h}(t_k)\, d\tau \\
&\quad + \int_{t_{i-1}}^{t} \tilde{A}(\tau)\,\Phi_{\tilde{A}}(\tau) - \tilde{A}_h(t_i)\,\mathit{\Phi}_{\tilde{A}_h}(t_i)\, d\tau \\
&= \sum_{k=1}^{i-1} \int_{t_{k-1}}^{t_k} \tilde{A}(\tau)\,\Phi_{\tilde{A}}(\tau) - \tilde{A}_h(t_k)\,\mathit{\Phi}_{\tilde{A}_h}(\tau)\, d\tau \\
&\quad + \int_{t_{i-1}}^{t} \tilde{A}(\tau)\,\Phi_{\tilde{A}}(\tau) - \tilde{A}_h(t_i)\,\mathit{\Phi}_{\tilde{A}_h}(\tau)\, d\tau \\
&\quad - \sum_{k=1}^{i-1} \int_{t_{k-1}}^{t_k} \tilde{A}_h(t_k)\left(\mathit{\Phi}_{\tilde{A}_h}(t_k) - \mathit{\Phi}_{\tilde{A}_h}(\tau)\right) d\tau \\
&\quad - \int_{t_{i-1}}^{t} \tilde{A}_h(t_i)\left(\mathit{\Phi}_{\tilde{A}_h}(t_i) - \mathit{\Phi}_{\tilde{A}_h}(\tau)\right) d\tau.
\end{aligned}
$$

Furthermore, (2.4.9) yields

$$\mathit{\Phi}_{\tilde{A}_h}(t_i) - \mathit{\Phi}_{\tilde{A}_h}(t) = (t_i - t)\,\tilde{A}_h(t_i)\,\mathit{\Phi}_{\tilde{A}_h}(t_i) \tag{2.4.11}$$

for every $i = 1, \ldots, N$ and $t \in (t_{i-1}, t_i]$. Recall, $\tilde{A}_h(t) = \tilde{A}_h(t_i)$ for every $i = 1, \ldots, N$ and $t \in (t_{i-1}, t_i]$. Consequently, it holds

$$\begin{aligned}
\left\|\Phi_{\tilde{A}}(t) - \mathit{\Phi}_{\tilde{A}_h}(t)\right\| \quad &\leq \quad \sum_{k=1}^{i-1} \int_{t_{k-1}}^{t_k} \left\|\tilde{A}(\tau)\,\Phi_{\tilde{A}}(\tau) - \tilde{A}_h(t_k)\,\mathit{\Phi}_{\tilde{A}_h}(\tau)\right\| d\tau \\
&\quad + \int_{t_{i-1}}^{t} \left\|\tilde{A}(\tau)\,\Phi_{\tilde{A}}(\tau) - \tilde{A}_h(t_i)\,\mathit{\Phi}_{\tilde{A}_h}(\tau)\right\| d\tau \\
&\quad + \sum_{k=1}^{i-1} \int_{t_{k-1}}^{t_k} \left\|\tilde{A}_h(t_k)\left(\mathit{\Phi}_{\tilde{A}_h}(t_k) - \mathit{\Phi}_{\tilde{A}_h}(\tau)\right)\right\| d\tau \\
&\quad + \int_{t_{i-1}}^{t} \left\|\tilde{A}_h(t_i)\left(\mathit{\Phi}_{\tilde{A}_h}(t_i) - \mathit{\Phi}_{\tilde{A}_h}(\tau)\right)\right\| d\tau \\
&\overset{(2.4.11)}{\leq} \sum_{k=1}^{i-1} \int_{t_{k-1}}^{t_k} \left\|\tilde{A}(\tau) - \tilde{A}_h(t_k)\right\| \left\|\mathit{\Phi}_{\tilde{A}_h}(\tau)\right\| + \left\|\tilde{A}(\tau)\right\| \left\|\Phi_{\tilde{A}}(\tau) - \mathit{\Phi}_{\tilde{A}_h}(\tau)\right\| d\tau \\
&\quad + \int_{t_{i-1}}^{t} \left\|\tilde{A}(\tau) - \tilde{A}_h(t_i)\right\| \left\|\mathit{\Phi}_{\tilde{A}_h}(\tau)\right\| + \left\|\tilde{A}(\tau)\right\| \left\|\Phi_{\tilde{A}}(\tau) - \mathit{\Phi}_{\tilde{A}_h}(\tau)\right\| d\tau \\
&\quad + \sum_{k=1}^{i-1} \int_{t_{k-1}}^{t_k} \left\|\tilde{A}_h(t_k)^2\,\mathit{\Phi}_{\tilde{A}_h}(t_k)\,(t_k - \tau)\right\| d\tau \\
&\quad + \int_{t_{i-1}}^{t} \left\|\tilde{A}_h(t_i)^2\,\mathit{\Phi}_{\tilde{A}_h}(t_i)\,(t_i - \tau)\right\| d\tau \\
&\underset{2.4.6}{\overset{\text{Lemma}}{\leq}} \sum_{k=1}^{i-1} \int_{t_{k-1}}^{t_k} \tilde{\mathbf{L}}h \left\|\mathit{\Phi}_{\tilde{A}_h}\right\|_\infty + \left\|\tilde{A}\right\|_\infty \left\|\Phi_{\tilde{A}}(\tau) - \mathit{\Phi}_{\tilde{A}_h}(\tau)\right\| d\tau \\
&\quad + \int_{t_{i-1}}^{t} \tilde{\mathbf{L}}h \left\|\mathit{\Phi}_{\tilde{A}_h}\right\|_\infty + \left\|\tilde{A}\right\|_\infty \left\|\Phi_{\tilde{A}}(\tau) - \mathit{\Phi}_{\tilde{A}_h}(\tau)\right\| d\tau \\
&\quad + \sum_{k=1}^{i-1} \int_{t_{k-1}}^{t_k} \left\|\tilde{A}_h\right\|_\infty^2 \left\|\mathit{\Phi}_{\tilde{A}_h}\right\|_\infty h\, d\tau + \int_{t_{i-1}}^{t} \left\|\tilde{A}_h\right\|_\infty^2 \left\|\mathit{\Phi}_{\tilde{A}_h}\right\|_\infty h\, d\tau \\
&= \quad \int_{t_0}^{t} \tilde{\mathbf{L}}h \left\|\mathit{\Phi}_{\tilde{A}_h}\right\|_\infty + \left\|\tilde{A}\right\|_\infty \left\|\Phi_{\tilde{A}}(\tau) - \mathit{\Phi}_{\tilde{A}_h}(\tau)\right\| + \left\|\tilde{A}_h\right\|_\infty^2 \left\|\mathit{\Phi}_{\tilde{A}_h}\right\|_\infty h\, d\tau \\
&\leq \quad \left(\tilde{\mathbf{L}} + \left\|\tilde{A}_h\right\|_\infty^2\right) \left\|\mathit{\Phi}_{\tilde{A}_h}\right\|_\infty h + \left\|\tilde{A}\right\|_\infty \int_{t_0}^{t} \left\|\Phi_{\tilde{A}}(\tau) - \mathit{\Phi}_{\tilde{A}_h}(\tau)\right\| d\tau,
\end{aligned}$$

for every $i = 1, \ldots, N$ and $t \in (t_{i-1}, t_i]$. Applying Lemma A.9 (Gronwall) for $\left\|\Phi_{\tilde{A}}(t) - \mathit{\Phi}_{\tilde{A}_h}(t)\right\|$ yields

$$\begin{aligned}\left\|\Phi_{\tilde{A}}(t) - \mathit{\Phi}_{\tilde{A}_h}(t)\right\| &\leq \left(\tilde{\mathbf{L}} + \left\|\tilde{A}_h\right\|_\infty^2\right)\left\|\mathit{\Phi}_{\tilde{A}_h}\right\|_\infty h \exp\left(\left\|\tilde{A}\right\|_\infty t\right) \\ &\overset{(2.4.10)}{\leq} \frac{3}{2}\left(\tilde{\mathbf{L}} + \left\|\tilde{A}_h\right\|_\infty^2\right)\exp\left(2\left\|\tilde{A}_h\right\|_\infty\right)\exp\left(\left\|\tilde{A}\right\|_\infty\right) h\end{aligned}$$

for all $t \in [0, 1]$. In addition, by Lemma A.2, it holds

$$\begin{aligned}\left\|\Phi_{\tilde{A}}(t)^{-1} - \mathit{\Phi}_{\tilde{A}_h}(t)^{-1}\right\| &\leq \left\|\Phi_{\tilde{A}}(\cdot)^{-1}\right\|_\infty \left\|\mathit{\Phi}_{\tilde{A}_h}(\cdot)^{-1}\right\|_\infty \left\|\Phi_{\tilde{A}}(t) - \mathit{\Phi}_{\tilde{A}_h}(t)\right\| \\ &\overset{(2.4.10)}{\leq} 2\exp\left(\left\|\tilde{A}\right\|_\infty\right)\exp\left(\left\|\tilde{A}_h\right\|_\infty\right)\left\|\Phi_{\tilde{A}}(t) - \mathit{\Phi}_{\tilde{A}_h}(t)\right\|\end{aligned}$$

for every $t \in [0, 1]$, where we exploited $\left\|\Phi_{\tilde{A}}(\cdot)^{-1}\right\|_\infty \leq \exp\left(\left\|\tilde{A}\right\|_\infty\right)$. Hence, the assertion is satisfied for $\mathbf{L}_\Phi := 3\exp\left(2\left\|\tilde{A}\right\|_\infty + 3\left\|\tilde{A}_h\right\|_\infty\right)\left(\tilde{\mathbf{L}} + \left\|\tilde{A}_h\right\|_\infty^2\right)$. □

We define the abbreviations $R_h \in \mathbb{R}^{n_E \times n_z}$, $S_h \in L_{\infty,h}^{n_E \times n_v}([0, 1])$ by

$$\begin{aligned}R_h &:= E_{0,h} + E_{1,h}\mathit{\Phi}_{\tilde{A}_h}(t_N), \\ S_h(t_i) &:= E_{1,h}\mathit{\Phi}_{\tilde{A}_h}(t_N)\mathit{\Phi}_{\tilde{A}_h}(t_{i-1})^{-1}\tilde{B}_h(t_i), \quad i = 1, \ldots, N,\end{aligned}$$

which satisfy the following:

Lemma 2.4.8

Let $D(\cdot)$ be uniformly linear independent with constant $\beta > 0$ and let Assumption 2.4.4 hold. Then, there exists $h_1 > 0$ and $\mathbf{L}_R, \mathbf{L}_S \geq 0$ such that for all $0 < h \leq h_1$ it holds

$$\|R - R_h\| \leq \mathbf{L}_R h, \qquad \|S(\cdot) - S_h(\cdot)\|_\infty \leq \mathbf{L}_S h.$$

Proof. Set $h_1 := \min\left\{\frac{\beta}{2\mathbf{L}}, \frac{1}{2\left\|\tilde{A}_h\right\|_\infty}\right\}$, $\mathbf{L}_R := \left(\mathbf{L} + \mathbf{L}\exp\left(\left\|\tilde{A}\right\|_\infty\right) + \|E_{1,h}\|\,\mathbf{L}_\Phi\right)$, and let $0 < h \leq h_1$ be arbitrary. Then, it holds

$$\begin{aligned}\|R - R_h\| &\leq \|E_0 - E_{0,h}\| + \|E_1 - E_{1,h}\|\left\|\Phi_{\tilde{A}}(1)\right\| + \|E_{1,h}\|\left\|\Phi_{\tilde{A}}(1) - \mathit{\Phi}_{\tilde{A}_h}(1)\right\| \\ &\underset{2.4.7}{\overset{\text{Lemma}}{\leq}} \mathbf{L}h + \mathbf{L}h\exp\left(\left\|\tilde{A}\right\|_\infty\right) + \|E_{1,h}\|\,\mathbf{L}_\Phi h \\ &= \mathbf{L}_R h.\end{aligned}$$

Moreover, utilizing Lemma A.2, Lemma 2.4.7, (2.4.10), and (2.4.11) yields

$$\begin{aligned}\left\|\Phi_{\tilde{A}}(t)^{-1} - \mathit{\Phi}_{\tilde{A}_h}(t_{i-1})^{-1}\right\| &\leq \left\|\Phi_{\tilde{A}}(t)^{-1} - \mathit{\Phi}_{\tilde{A}_h}(t)^{-1}\right\| + \left\|\mathit{\Phi}_{\tilde{A}_h}(t)^{-1} - \mathit{\Phi}_{\tilde{A}_h}(t_{i-1})^{-1}\right\| \\ &\leq \mathbf{L}_\Phi h + \left\|\mathit{\Phi}_{\tilde{A}_h}(t)^{-1}\right\|\left\|\mathit{\Phi}_{\tilde{A}_h}(t_{i-1})^{-1}\right\|\left\|\mathit{\Phi}_{\tilde{A}_h}(t) - \mathit{\Phi}_{\tilde{A}_h}(t_{i-1})\right\| \\ &\leq \mathbf{L}_\Phi h + \left(2\exp\left(\left\|\tilde{A}_h\right\|_\infty\right)\right)^2\left\|\mathit{\Phi}'_{\tilde{A}_h}(t_i)(t - t_{i-1})\right\| \\ &\leq \left(\mathbf{L}_\Phi + 4\left\|\tilde{A}_h\right\|_\infty\exp\left(4\left\|\tilde{A}_h\right\|_\infty\right)\right) h \qquad (2.4.12)\end{aligned}$$

for $i = 1, \ldots, N$ and $t \in (t_{i-1}, t_i]$. By using (2.4.12) and applying Lemma A.1, Lemma A.5, Lemma 2.4.6, and Lemma 2.4.7 we conclude

$$\begin{aligned}
\|S(t) - S_h(t)\| &= \left\| E_1 \Phi_{\tilde{A}}(1) \Phi_{\tilde{A}}(t)^{-1} \tilde{B}(t) - E_{1,h} \varPhi_{\tilde{A}_h}(t_N) \varPhi_{\tilde{A}_h}(t_{i-1})^{-1} \tilde{B}_h(t_i) \right\| \\
&\leq \|E_1 - E_{1,h}\| \|\Phi_{\tilde{A}}(1)\| \left\|\Phi_{\tilde{A}}(t)^{-1}\right\| \left\|\tilde{B}(t)\right\| \\
&\quad + \|E_{1,h}\| \left\|\Phi_{\tilde{A}}(1) - \varPhi_{\tilde{A}_h}(t_N)\right\| \left\|\Phi_{\tilde{A}}(t)^{-1}\right\| \left\|\tilde{B}(t)\right\| \\
&\quad + \|E_{1,h}\| \left\|\varPhi_{\tilde{A}_h}(t_N)\right\| \left\|\Phi_{\tilde{A}}(t)^{-1} - \varPhi_{\tilde{A}_h}(t_{i-1})^{-1}\right\| \left\|\tilde{B}(t)\right\| \\
&\quad + \|E_{1,h}\| \left\|\varPhi_{\tilde{A}_h}(t_N)\right\| \left\|\varPhi_{\tilde{A}_h}(t_{i-1})^{-1}\right\| \left\|\tilde{B}(t) - \tilde{B}_h(t_i)\right\| \\
&\leq \mathbf{L} h \exp\left(\left\|\tilde{A}\right\|_\infty\right)^2 \left\|\tilde{B}\right\|_\infty + \|E_{1,h}\| \mathbf{L}_\Phi h \exp\left(\left\|\tilde{A}\right\|_\infty\right) \left\|\tilde{B}\right\|_\infty \\
&\quad + \|E_{1,h}\| \exp\left(2\left\|\tilde{A}_h\right\|_\infty\right) \left(\mathbf{L}_\Phi + 4\left\|\tilde{A}_h\right\|_\infty \exp\left(4\left\|\tilde{A}_h\right\|_\infty\right)\right) h \left\|\tilde{B}\right\|_\infty \\
&\quad + \|E_{1,h}\| \exp\left(2\left\|\tilde{A}_h\right\|_\infty\right) \exp\left(\left\|\tilde{A}_h\right\|_\infty\right) \tilde{\mathbf{L}} h
\end{aligned}$$

for $i = 1, \ldots, N$ and almost every $t \in (t_{i-1}, t_i]$. Hence, the assertion holds for

$$\begin{aligned}
\mathbf{L}_S := \mathbf{L} \exp\left(\left\|\tilde{A}\right\|_\infty\right)^2 \left\|\tilde{B}\right\|_\infty + \|E_{1,h}\| \mathbf{L}_\Phi \exp\left(\left\|\tilde{A}\right\|_\infty\right) \left\|\tilde{B}\right\|_\infty \\
+ \|E_{1,h}\| \exp\left(2\left\|\tilde{A}_h\right\|_\infty\right) \left(\mathbf{L}_\Phi + 4\left\|\tilde{A}_h\right\|_\infty \exp\left(4\left\|\tilde{A}_h\right\|_\infty\right)\right) \left\|\tilde{B}\right\|_\infty \\
+ \|E_{1,h}\| \exp\left(3\left\|\tilde{A}_h\right\|_\infty\right) \tilde{\mathbf{L}}
\end{aligned}$$

□

Finally, for the discrete Gramian $G_h := R_h R_h^\top + \sum_{k=1}^{N} \int_{t_{k-1}}^{t_k} S_h(t_k) S_h(t_k)^\top d\tau$ we can prove the following:

Lemma 2.4.9 (Discrete Gramian Matrix)
Let $D(\cdot)$ be uniformly linear independent with constant $\beta > 0$ and let Assumption 2.4.4 hold. Suppose (2.4.1) is completely controllable. Then, there exists $h_1 > 0$ such that for all $0 < h \leq h_1$ the discrete Gramian G_h is non-singular.

Proof. According to Lemma 2.4.3, the continuous Gramian G is non-singular. In addition, G is symmetric and positive definite, since for arbitrary $\varpi \in \mathbb{R}^{n_E} \setminus \{\mathbf{0}_{\mathbb{R}^{n_E}}\}$ it holds

$$0 \neq \varpi^\top G \varpi = \left\|R^\top \varpi\right\|^2 + \left\|S(\cdot)^\top \varpi\right\|_2^2 \geq 0,$$

thus $\varpi^\top G \varpi > 0$. Furthermore, since the eigenvalues of a symmetric positive definite matrix are positive, for $\gamma_G := \min\{\lambda \mid \lambda \text{ is an eigenvalue of } G\} > 0$ we obtain $\varpi^\top G \varpi \geq \gamma_G \|\varpi\|^2$ for every $\varpi \in \mathbb{R}^{n_E}$. Set

$$\mathbf{L}_G := \mathbf{L}_R(\|R\| + \|R_h\|) + \mathbf{L}_S(\|S\|_\infty + \|S_h\|_\infty) \text{ and } h_1 := \min\left\{\frac{\beta}{2\mathbf{L}}, \frac{1}{2\left\|\tilde{A}_h\right\|_\infty}, \frac{\gamma_G}{2\mathbf{L}_G}\right\}.$$

Then, by Lemma 2.4.8, for arbitrary $0 < h \leq h_1$ and $\varpi \in \mathbb{R}^{n_E}$ it holds

$$\begin{aligned}
\varpi^\top G\varpi - \varpi^\top G_h\varpi &= \left\|R^\top\varpi\right\|^2 + \left\|S\left(\cdot\right)^\top\varpi\right\|_2^2 - \left\|R_h^\top\varpi\right\|^2 - \left\|S_h\left(\cdot\right)^\top\varpi\right\|_2^2 \\
&= \left(\left\|R^\top\varpi\right\| - \left\|R_h^\top\varpi\right\|\right)\left(\left\|R^\top\varpi\right\| + \left\|R_h^\top\varpi\right\|\right) \\
&\quad + \left(\left\|S\left(\cdot\right)^\top\varpi\right\|_2 - \left\|S_h\left(\cdot\right)^\top\varpi\right\|_2\right)\left(\left\|S\left(\cdot\right)^\top\varpi\right\|_2 + \left\|S_h\left(\cdot\right)^\top\varpi\right\|_2\right) \\
&\leq \left\|R^\top\varpi - R_h^\top\varpi\right\|\left(\left\|R^\top\right\| + \left\|R_h^\top\right\|\right)\left\|\varpi\right\| \\
&\quad + \left\|S\left(\cdot\right)^\top\varpi - S_h\left(\cdot\right)^\top\varpi\right\|_2\left(\left\|S\left(\cdot\right)^\top\right\|_\infty + \left\|S_h\left(\cdot\right)^\top\right\|_\infty\right)\left\|\varpi\right\| \\
&\leq \mathbf{L}_R h\left(\left\|R\right\| + \left\|R_h\right\|\right)\left\|\varpi\right\|^2 + \mathbf{L}_S h\left(\left\|S\right\|_\infty + \left\|S_h\right\|_\infty\right)\left\|\varpi\right\|^2 \\
&= \mathbf{L}_G h\left\|\varpi\right\|^2.
\end{aligned}$$

Consequently,

$$0 < \gamma_G\left\|\varpi\right\|^2 \leq \varpi^\top G\varpi \leq \mathbf{L}_G h\left\|\varpi\right\|^2 + \varpi^\top G_h\varpi \leq \frac{\gamma_G}{2}\left\|\varpi\right\|^2 + \varpi^\top G_h\varpi,$$

hence G_h is positive definite, and therefore non-singular. □

In the same way as in the second part of the proof of Lemma 2.4.3, it follows straightforwardly from Lemma 2.4.9 that the discrete system 2.4.7 is completely controllable, if the discrete Gramian is non-singular. Summarizing, we showed that, if the continuous system 2.4.1 satisfies the linear independence conditions, is completely controllable, and Assumption 2.4.4 holds, then the linear independence and complete controllability are retained by the discrete system 2.4.7.

Occasionally, it is more convenient to view DAEs of type (2.4.1) as an operator equation $F(z,v) = \mathbf{0}$ for a linear operator F. To that end, for $p = 2, \infty$ and spaces

$$\begin{aligned}
Z_p &:= W_{1,p}^{n_z}\left([0,1]\right) \times L_p^{n_v}\left([0,1]\right), & Y_p &:= L_p^{n_z}\left([0,1]\right) \times L_p^{m}\left([0,1]\right) \times \mathbb{R}^{n_E}, \\
Z_{p,h} &:= W_{1,p,h}^{n_z}\left([0,1]\right) \times L_{p,h}^{n_v}\left([0,1]\right), & Y_{p,h} &:= L_{p,h}^{n_z}\left([0,1]\right) \times L_{p,h}^{m}\left([0,1]\right) \times \mathbb{R}^{n_E},
\end{aligned}$$

with the norms $\left\|(z,v)\right\|_{Z_p} := \max\left\{\left\|z\right\|_{1,p}, \left\|v\right\|_p\right\}$, $\left\|(a,b,e)\right\|_{Y_p} := \max\left\{\left\|a\right\|_p, \left\|b\right\|_p, \left\|e\right\|\right\}$ we define the linear operators $F : Z_p \to Y_p$, $F_h : Z_{p,h} \to Y_{p,h}$ as

$$F\left(z,v\right) := \begin{pmatrix} \dot{z}\left(\cdot\right) - A\left(\cdot\right)z\left(\cdot\right) - B\left(\cdot\right)v\left(\cdot\right) \\ C\left(\cdot\right)z\left(\cdot\right) + D\left(\cdot\right)v\left(\cdot\right) \\ E_0 z\left(0\right) + E_1 z\left(1\right) \end{pmatrix},$$

$$F_h\left(z_h,v_h\right)\left(t\right) := \begin{pmatrix} z_h'\left(t_i\right) - A_h\left(t_i\right)z_h\left(t_i\right) - B_h\left(t_i\right)v_h\left(t_i\right) \\ C_h\left(t_i\right)z_h\left(t_i\right) + D_h\left(t_i\right)v_h\left(t_i\right) \\ E_{0,h}z_h\left(t_0\right) + E_{1,h}z_h\left(t_N\right) \end{pmatrix}, \quad t \in \left(t_{i-1}, t_i\right], \, i = 1, \ldots, N.$$

Since all the matrix functions are essentially bounded, the linear operators are also bounded. The following lemma gives conditions under which F and F_h are uniformly surjective in virtue of Definition 2.1.4:

Lemma 2.4.10 (Uniform Surjective Operators)
Let system (2.4.1) *and* (2.4.7) *satisfy the conditions in Definition 2.4.1 and Assumption 2.4.4, respectively. Then, it holds:*

(i) The operator F is uniformly surjective.

(ii) There exists $h_1 > 0$ such that for every $0 < h \leq h_1$ the operator F_h is uniformly surjective with a constant independent of h.

Proof.

(i) For arbitrary $(a(\cdot), b(\cdot), e) \in Y_p$ we show that the inhomogeneous linear system

$$\begin{aligned} \dot{z}(t) &= A(t) z(t) + B(t) v(t) - a(t), && \text{a.e. in } [0,1], \\ \mathbf{0}_{\mathbb{R}^m} &= C(t) z(t) + D(t) v(t) - b(t), && \text{a.e. in } [0,1], \\ e &= E_0 z(0) + E_1 z(1). \end{aligned} \tag{2.4.13}$$

has a solution $(z, v) \in Z_p$ satisfying $\kappa \|(z,v)\|_{Z_p} \leq \|(a,b,e)\|_{Y_p}$ for some constant $\kappa > 0$. Analog to Remark 2.4.2, we reduce the system to

$$\begin{aligned} \dot{z}(t) &= \tilde{A}(t) z(t) + \tilde{B}(t) w(t) + \tilde{a}(t), && \text{a.e. in } [0,1], \\ e &= E_0 z(0) + E_1 z(1), \end{aligned}$$

where $w \in L_p^{n_v}([0,1])$ and $\tilde{a}(\cdot) := -a(\cdot) + B(\cdot) D(\cdot)^{\curlywedge} b(\cdot)$, which has the solution

$$z(\cdot) = \Phi_{\tilde{A}}(\cdot) z(0) + \Phi_{\tilde{A}}(\cdot) \int_0^{\cdot} \Phi_{\tilde{A}}(\tau)^{-1} \left(\tilde{B}(\tau) w(\tau) + \tilde{a}(\tau)\right) d\tau.$$

Inserting this into the boundary conditions yields

$$\begin{aligned} e &= \left(E_0 + E_1 \Phi_{\tilde{A}}(1)\right) z(0) + E_1 \Phi_{\tilde{A}}(1) \int_0^1 \Phi_{\tilde{A}}(\tau)^{-1} \tilde{B}(\tau) w(\tau) d\tau \\ &\quad + E_1 \Phi_{\tilde{A}}(1) \int_0^1 \Phi_{\tilde{A}}(\tau)^{-1} \tilde{a}(\tau) d\tau \\ &= R z(0) + \int_0^1 S(\tau) w(\tau) d\tau + E_1 \Phi_{\tilde{A}}(1) \int_0^1 \Phi_{\tilde{A}}(\tau)^{-1} \tilde{a}(\tau) d\tau. \end{aligned} \tag{2.4.14}$$

For $\tilde{e} := e - E_1 \Phi_{\tilde{A}}(1) \int_0^1 \Phi_{\tilde{A}}(\tau)^{-1} \tilde{a}(\tau) d\tau$ we choose

$$z(0) = R^\top G^{-1} \tilde{e}, \quad w(\cdot) = S(\cdot)^\top G^{-1} \tilde{e},$$

which satisfy the boundary conditions (2.4.14). We conclude that (2.4.13) holds for

$$z(\cdot) = \Phi_{\tilde{A}}(\cdot) R^\top G^{-1} \tilde{e} + \Phi_{\tilde{A}}(\cdot) \int_0^{\cdot} \Phi_{\tilde{A}}(\tau)^{-1} \left(\tilde{B}(\tau) S(\tau)^\top G^{-1} \tilde{e} + \tilde{a}(\tau)\right) d\tau, \tag{2.4.15}$$

$$v(\cdot) = -D(\cdot)^{\curlywedge} C(\cdot) z(\cdot) + \left(\mathbf{I}_{n_v} - D(\cdot)^{\curlywedge} D(\cdot)\right) S(\cdot)^\top G^{-1} \tilde{e} + D(\cdot)^{\curlywedge} b(\cdot),$$

hence $F : Z_p \to Y_p$ is surjective. Additionally, we obtain the bounds

$$\begin{aligned}
\|\tilde{a}\|_p &\leq \|a\|_p + \|B\|_\infty \left\|D^{\curlywedge}\right\|_\infty \|b\|_p \\
&\leq \left(1 + \|B\|_\infty \left\|D^{\curlywedge}\right\|_\infty\right) \|(a,b,e)\|_{Y_p}, \\
\|\tilde{e}\| &\leq \|e\| + \|E_1\| \exp\left(\left\|\tilde{A}\right\|_\infty\right) \|\tilde{a}\|_p \\
&\leq \left[1 + \|E_1\| \exp\left(\left\|\tilde{A}\right\|_\infty\right) \left(1 + \|B\|_\infty \left\|D^{\curlywedge}\right\|_\infty\right)\right] \|(a,b,e)\|_{Y_p}.
\end{aligned}$$

Set $\kappa_1 := \left[1 + \|E_1\| \exp\left(\left\|\tilde{A}\right\|_\infty\right) \left(1 + \|B\|_\infty \left\|D^{\curlywedge}\right\|_\infty\right)\right]$. Using (2.4.15) and the differential equation in (2.4.13) we find a constant $\kappa_2 > 0$ such that

$$\|z\|_p \leq \kappa_2 \|(a,b,e)\|_{Y_p}, \quad \|\dot{z}\|_p \leq \kappa_2 \|(a,b,e)\|_{Y_p}, \quad \|v\|_p \leq \kappa_2 \|(a,b,e)\|_{Y_p},$$

and therefore $\kappa \|(z,v)\|_{Z_p} \leq \|(a,b,e)\|_{Y_p}$ for $\kappa = \frac{1}{\kappa_2}$.

(ii) Set $h_1 := \min\left\{\frac{\beta}{2\mathbf{L}}, \frac{1}{2\left\|\tilde{A}_h\right\|_\infty}, \frac{\gamma_G}{2\mathbf{L}_G}\right\}$ and let $0 < h \leq h_1$ be arbitrary. Then, by Lemma 2.4.9, the discrete Gramian G_h is non-singular. In the same way as in (i), we consider the inhomogeneous linear system

$$\begin{aligned}
z_h'(t_i) &= A_h(t_i) z_h(t_i) + B_h(t_i) v_h(t_i) - a_h(t_i), \qquad i = 1, \ldots, N, \\
\mathbf{0}_{\mathbb{R}^m} &= C_h(t_i) z_h(t_i) + D_h(t_i) v_h(t_i) - b_h(t_i), \qquad i = 1, \ldots, N, \\
e_h &= E_{0,h} z_h(t_0) + E_{1,h} z_h(t_N).
\end{aligned} \tag{2.4.16}$$

for arbitrary $(a_h(\cdot), b_h(\cdot), e_h) \in Y_{p,h}$. Analog to the continuous case, we reduce the system to

$$\begin{aligned}
z_h'(t_i) &= \tilde{A}_h(t_i) z_h(t_i) + \tilde{B}_h(t_i) w_h(t_i) + \tilde{a}_h(t_i), \qquad i = 1, \ldots, N, \\
e_h &= E_{0,h} z(t_0) + E_{1,h} z(t_N),
\end{aligned}$$

where $w_h \in L_{p,h}^{n_v}([0,1])$ and $\tilde{a}_h(\cdot) := -a_h(\cdot) + B_h(\cdot) D_h(\cdot)^{\curlywedge} b_h(\cdot)$. This system has the solution

$$z_h(t_i) = \Phi_{\tilde{A}_h}(t_i) z_h(t_0) + \Phi_{\tilde{A}_h}(t_i) h \sum_{k=1}^{i} \Phi_{\tilde{A}_h}(t_{k-1})^{-1} \left(\tilde{B}_h(t_k) w_h(t_k) + \tilde{a}_h(t_k)\right),$$

which inserted into boundary conditions yields

$$e_h = R_h z_h(t_0) + h \sum_{k=1}^{N} S_h(t_k) w_h(t_k) + E_{1,h} \Phi_{\tilde{A}_h}(t_N) h \sum_{k=1}^{N} \Phi_{\tilde{A}_h}(t_{k-1})^{-1} \tilde{a}_h(t_k).$$

Choose

$$z_h(t_0) = R_h^\top G_h^{-1} \tilde{e}_h, \quad w_h(\cdot) = S_h(\cdot)^\top G_h^{-1} \tilde{e}_h,$$

where $\tilde{e}_h = e_h - E_{1,h}\Phi_{\tilde{A}_h}(t_N)\, h \sum_{k=1}^{N} \Phi_{\tilde{A}_h}(t_{k-1})^{-1} \tilde{a}_h(t_k)$, which results in the solution

$$
\begin{aligned}
z_h(t_i) &= \Phi_{\tilde{A}_h}(t_i)\, R_h^\top G_h^{-1} \tilde{e}_h \\
&\quad + \Phi_{\tilde{A}_h}(t_i)\, h \sum_{k=1}^{i} \Phi_{\tilde{A}_h}(t_{k-1})^{-1} \left(\tilde{B}_h(t_k)\, S_h(t_k)^\top G_h^{-1} \tilde{e}_h + \tilde{a}_h(t_k) \right), \\
v_h(t_i) &= -D_h(t_i)^{\curlywedge} C_h(t_i)\, z_h(t_i) \\
&\quad + \left(\mathbf{I}_{n_v} - D_h(t_i)^{\curlywedge} D_h(t_i) \right) S_h(t_i)^\top G_h^{-1} \tilde{e_h} + D_h(t_i)^{\curlywedge} b_h(t_i),
\end{aligned}
$$

for $i = 1, \ldots, N$, thus F_h is surjective. Analog to the continuous case, we find a constant $\tilde{\kappa} > 0$ independent of h such that $\tilde{\kappa} \left\| (z_h, v_h) \right\|_{Z_p} \leq \left\| (a_h, b_h, e_h) \right\|_{Y_p}$, which completes the proof.

□

Consider the following bilinear forms

$$
\mathcal{P} : Z_2 \times Z_2 \to \mathbb{R}, \quad \mathcal{P}_h : Z_{2,h} \times Z_{2,h} \to \mathbb{R},
$$

defined by

$$
\begin{aligned}
\mathcal{P}\left(\left(z^1, v^1 \right), \left(z^2, v^2 \right) \right) &:= \begin{pmatrix} z^1(0) \\ z^1(1) \end{pmatrix}^\top \Lambda \begin{pmatrix} z^2(0) \\ z^2(1) \end{pmatrix} \\
&\quad + \int_0^1 \begin{pmatrix} z^1(t) \\ v^1(t) \end{pmatrix}^\top Q(t) \begin{pmatrix} z^2(t) \\ v^2(t) \end{pmatrix} dt, \\
\mathcal{P}_h\left(\left(z_h^1, v_h^1 \right), \left(z_h^2, v_h^2 \right) \right) &:= \begin{pmatrix} z_h^1(t_0) \\ z_h^1(t_N) \end{pmatrix}^\top \Lambda_h \begin{pmatrix} z_h^2(t_0) \\ z_h^2(t_N) \end{pmatrix} \\
&\quad + \sum_{k=1}^{N} \int_{t_{k-1}}^{t_k} \begin{pmatrix} z_h^1(t_k) \\ v_h^1(t_k) \end{pmatrix}^\top Q_h(t_k) \begin{pmatrix} z_h^2(t_k) \\ v_h^2(t_k) \end{pmatrix} dt,
\end{aligned}
$$

where $\Lambda, \Lambda_h \in \mathbb{R}^{2n_z \times 2n_z}$, and $Q \in L_\infty^{(n_z+n_v)\times(n_z+n_v)}([0,1])$, $Q_h \in L_{\infty,h}^{(n_z+n_v)\times(n_z+n_v)}([0,1])$ are symmetric matrices and symmetric matrix functions, respectively. These bilinear forms are symmetric and also continuous, since for every $\left(z^1, v^1\right), \left(z^2, v^2\right) \in Z_2$ it holds

$$
\begin{aligned}
\left| \mathcal{P}\left(\left(z^1, v^1 \right), \left(z^2, v^2 \right) \right) \right| &\leq \left\| \left(z^1(0), z^1(1) \right) \right\| \left\| \Lambda \right\| \left\| \left(z^2(0), z^2(1) \right) \right\| \\
&\quad + \int_0^1 \left\| \left(z^1(t), v^1(t) \right) \right\| \left\| Q \right\|_\infty \left\| \left(z^2(t), v^2(t) \right) \right\| dt \\
&\leq \left(\left\| z^1(0) \right\| + \left\| z^1(1) \right\| \right) \left\| \Lambda \right\| \left(\left\| z^2(0) \right\| + \left\| z^2(1) \right\| \right) \\
&\quad + \left\| Q \right\|_\infty \left\| \left(z^1, v^1 \right) \right\|_2 \left\| \left(z^2, v^2 \right) \right\|_2 \\
&\leq 4 \left\| z^1 \right\|_{1,2} \left\| \Lambda \right\| 4 \left\| z^2 \right\|_{1,2} + \left\| Q \right\|_\infty \left(\left\| z^1 \right\|_2 + \left\| v^1 \right\|_2 \right) \left(\left\| z^2 \right\|_2 + \left\| v^2 \right\|_2 \right) \\
&\leq \left(16 \left\| \Lambda \right\| + 4 \left\| Q \right\|_\infty \right) \left\| \left(z^1, v^1 \right) \right\|_{Z_2} \left\| \left(z^2, v^2 \right) \right\|_{Z_2}.
\end{aligned}
$$

Herein, we exploited the Cauchy-Schwartz and the Sobolev (Lemma A.7) inequalities. By the same token, we can prove continuity for $\mathcal{P}_h$. Using Theorem 2.1.10 we want to show that coercivity of the bilinear form $\mathcal{P}$ on $\ker(F)$ is inherited by the discrete bilinear form $\mathcal{P}_h$ on $\ker(F_h)$ for sufficiently small h. To that end, we assume the following:

Assumption 2.4.11 (Bilinear Form Approximation)
There exist constants $\mathbf{L}_\Lambda, \mathbf{L}_Q \geq 0$ such that for all $h > 0$ it holds

$$\|\Lambda - \Lambda_h\| \leq \mathbf{L}_\Lambda h, \quad \|Q(\cdot) - Q_h(\cdot)\|_\infty \leq \mathbf{L}_Q h.$$

This assumption allows us to prove coercivity for the discrete bilinear form. Additionally, if we decompose Q, Q_h in a certain way, we are able to show that coercivity also implies that Legendre-Clebsch conditions are satisfied (compare Figure 2.2). Using the same techniques

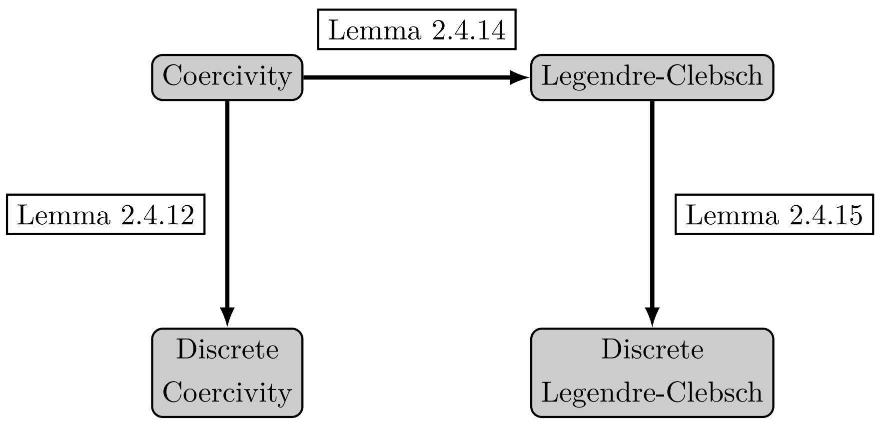

Figure 2.2: Relation between coercivity and Legendre-Clebsch conditions.

as in [85, Lemma 9], one can show that the discrete Legendre-Clebsch condition follows from discrete coercivity directly.

Lemma 2.4.12 (Discrete Coercivity)
Let system (2.4.1) *and* (2.4.7) *satisfy the conditions in Definition 2.4.1 and Assumption 2.4.4, respectively. Furthermore, suppose Assumption 2.4.11 holds and $\mathcal{P}$ is coercive on* $\ker(F)$ *with constant $\gamma > 0$. Then, there exist $\tilde{\gamma}, \tilde{h} > 0$ such that for every $0 < h \leq \tilde{h}$ the bilinear form $\mathcal{P}_h$ is coercive on* $\ker(F_h)$ *with constant $\tilde{\gamma}$.*

Proof. Our goal is to apply Theorem 2.1.10, i.e., we have to verify conditions (i) - (v).

(i) F is uniformly surjective by Lemma 2.4.10.

(ii) The coercivity of $\mathcal{P}$ on $\ker(F)$ is explicitly assumed.

(iii) By Lemma 2.4.10, there exists a $h_1 > 0$ such that F_h is uniformly surjective for every $0 < h \leq h_1$.

(iv) We need to show that there exists a constant $\mathbf{L}_F \geq 0$ independent of h such that

$$\|F(z_h, v_h) - F_h(z_h, v_h)\|_{Y_2} \leq \mathbf{L}_F h \|(z_h, v_h)\|_{Z_2} \quad \text{for all } (z_h, v_h) \in Z_{2,h}.$$

To that end, let $(z_h, v_h) \in Z_{2,h}$ be arbitrary. Then, for $i = 1, \ldots, N$ and almost every $t \in (t_{i-1}, t_i]$ we obtain

$$
\begin{aligned}
&F\left(z_h, v_h\right)(t) - F_h\left(z_h, v_h\right)(t) \\
&= \begin{pmatrix} \dot{z}_h(t) - A(t) z_h(t) - B(t) v_h(t) \\ C(t) z_h(t) + D(t) v_h(t) \\ E_0 z_h(t_0) + E_1 z_h(t_N) \end{pmatrix} - \begin{pmatrix} z_h'(t_i) - A_h(t_i) z_h(t_i) - B_h(t_i) v_h(t_i) \\ C_h(t_i) z_h(t_i) + D_h(t_i) v_h(t_i) \\ E_{0,h} z_h(t_0) + E_{1,h} z_h(t_N) \end{pmatrix} \\
&= \begin{pmatrix} -A(t) z_h(t) + A_h(t_i) z_h(t_i) - (B(t) - B_h(t_i)) v_h(t_i) \\ C(t) z_h(t) - C_h(t_i) z_h(t_i) + (D(t) - D_h(t_i)) v_h(t_i) \\ (E_0 - E_{0,h}) z_h(t_0) + (E_1 - E_{1,h}) z_h(t_N) \end{pmatrix} =: \begin{pmatrix} a_h(t) \\ b_h(t) \\ e_h \end{pmatrix}.
\end{aligned}
$$

Since

$$
z_h(t_i) - z_h(t) = z_h'(t_i)(t_i - t) \tag{2.4.17}
$$

for every $t \in (t_{i-1}, t_i]$ and $i = 1, \ldots, N$, it holds

$$
\begin{aligned}
\|a_h(t)\| &\le \|A(t) - A_h(t_i)\| \|z_h(t_i)\| + \|A(t)\| \|z_h(t) - z_h(t_i)\| \\
&\quad + \|B(t) - B_h(t_i)\| \|v_h(t_i)\| \\
&\le \mathbf{L} h \|z_h(t_i)\| + \|A\|_\infty h \|z_h'(t_i)\| + \mathbf{L} h \|v_h(t_i)\| \\
&\le (\mathbf{L} + \|A\|_\infty) h \left(\|z_h(t_i)\| + \|z_h'(t_i)\| + \|v_h(t_i)\|\right)
\end{aligned}
$$

for $i = 1, \ldots, N$ and almost every $t \in (t_{i-1}, t_i]$. Thus, using the Sobolev inequality in Lemma A.7 yields

$$
\begin{aligned}
\|a_h\|_2^2 &= \int_0^1 \|a_h(t)\|^2 dt = \sum_{k=1}^N \int_{t_{k-1}}^{t_k} \|a_h(t)\|^2 dt \\
&\le (\mathbf{L} + \|A\|_\infty)^2 h^2 \sum_{k=1}^N \int_{t_{k-1}}^{t_k} \left(\|z_h(t_k)\| + \|z_h'(t_k)\| + \|v_h(t_k)\|\right)^2 dt \\
&\overset{\text{Sobolev}}{\le} (\mathbf{L} + \|A\|_\infty)^2 h^2 \sum_{k=1}^N \int_{t_{k-1}}^{t_k} \left(2\|z_h\|_{1,2} + \|z_h'(t_k)\| + \|v_h(t_k)\|\right)^2 dt \\
&\le (\mathbf{L} + \|A\|_\infty)^2 h^2 \sum_{k=1}^N \int_{t_{k-1}}^{t_k} 2\left(\left(2\|z_h\|_{1,2}\right)^2 + \left(\|z_h'(t_k)\| + \|v_h(t_k)\|\right)^2\right) dt \\
&\le (\mathbf{L} + \|A\|_\infty)^2 h^2 \sum_{k=1}^N \int_{t_{k-1}}^{t_k} 2\left(4\|z_h\|_{1,2}^2 + 2\left(\|z_h'(t_k)\|^2 + \|v_h(t_k)\|^2\right)\right) dt \\
&\le (\mathbf{L} + \|A\|_\infty)^2 h^2 \sum_{k=1}^N \int_{t_{k-1}}^{t_k} 4\left(2\|z_h\|_{1,2}^2 + \|z_h'(t_k)\|^2 + \|v_h(t_k)\|^2\right) dt \\
&= 4(\mathbf{L} + \|A\|_\infty)^2 h^2 \left(2\|z_h\|_{1,2}^2 + \|z_h'\|_2^2 + \|v_h\|_2^2\right) \\
&\le 4(\mathbf{L} + \|A\|_\infty)^2 h^2 \left(2\|(z_h, v_h)\|_{Z_2}^2 + \|(z_h, v_h)\|_{Z_2}^2 + \|(z_h, v_h)\|_{Z_2}^2\right) \\
&= 16(\mathbf{L} + \|A\|_\infty)^2 h^2 \|(z_h, v_h)\|_{Z_2}^2 .
\end{aligned}
$$

Analog, we obtain $\|b_h\|_2 \leq 4\left(\mathbf{L} + \|C\|_\infty\right) h \left\|(z_h, v_h)\right\|_{Z_2}$ and additionally

$$\begin{aligned}\|e_h\| &\leq \|E_0 - E_{0,h}\| \left\|z_h(t_0)\right\| + \|E_1 - E_{1,h}\| \left\|z_h(t_N)\right\| \leq \mathbf{L}h\left(\|z_h(t_0)\| + \|z_h(t_N)\|\right)\\ &\overset{\text{Sobolev}}{\leq} 4\mathbf{L}h\|z_h\|_{1,2} \leq 4\mathbf{L}h\left\|(z_h, v_h)\right\|_{Z_2}.\end{aligned}$$

Hence, $\|F(z_h, v_h) - F_h(z_h, v_h)\|_{Y_2} \leq \mathbf{L}_F h \left\|(z_h, v_h)\right\|_{Z_2}$ for $\mathbf{L}_F := 4\left(\mathbf{L} + \|A\|_\infty + \|C\|_\infty\right)$.

(v) It remains to show that there exists a constant $\mathbf{L}_\mathcal{P} \geq 0$ independent of h such that

$$\mathcal{P}\left((z_h, v_h), (z_h, v_h)\right) - \mathcal{P}_h\left((z_h, v_h), (z_h, v_h)\right) \leq \mathbf{L}_\mathcal{P} h \left\|(z_h, v_h)\right\|_{Z_2}^2 \quad \text{for all } (z_h, v_h) \in Z_{2,h}.$$

For arbitrary $(z_h, v_h) \in Z_{2,h}$ it holds

$$\begin{aligned}\begin{pmatrix} z_h(t_0) \\ z_h(t_N)\end{pmatrix}^\top (\Lambda - \Lambda_h) \begin{pmatrix} z_h(t_0) \\ z_h(t_N)\end{pmatrix} &\underset{2.4.11}{\overset{\text{Assumption}}{\leq}} \mathbf{L}_\Lambda h\left(\|z_h(t_0)\|^2 + \|z_h(t_N)\|^2\right)\\ &\overset{\text{Sobolev}}{\leq} 8\mathbf{L}_\Lambda h \|z_h\|_{1,2}^2,\end{aligned}$$

and utilizing (2.4.17) yields

$$\begin{aligned}&\int_0^1 \begin{pmatrix} z_h(t) \\ v_h(t)\end{pmatrix}^\top Q(t) \begin{pmatrix} z_h(t) \\ v_h(t)\end{pmatrix} dt - \sum_{k=1}^N \int_{t_{k-1}}^{t_k} \begin{pmatrix} z_h(t_k) \\ v_h(t_k)\end{pmatrix}^\top Q_h(t_k) \begin{pmatrix} z_h(t_k) \\ v_h(t_k)\end{pmatrix} dt\\
=\;& \sum_{k=1}^N \int_{t_{k-1}}^{t_k} \begin{pmatrix} z_h(t_k) \\ v_h(t_k)\end{pmatrix}^\top \left(Q(t) - Q_h(t_k)\right) \begin{pmatrix} z_h(t_k) \\ v_h(t_k)\end{pmatrix} dt\\
&+ \sum_{k=1}^N \int_{t_{k-1}}^{t_k} \begin{pmatrix} z_h(t) - z_h(t_k) \\ \mathbf{0}_{\mathbb{R}^{n_v}}\end{pmatrix}^\top Q(t) \begin{pmatrix} z_h(t) - z_h(t_k) \\ \mathbf{0}_{\mathbb{R}^{n_v}}\end{pmatrix} dt\\
&+2 \sum_{k=1}^N \int_{t_{k-1}}^{t_k} \begin{pmatrix} z_h(t) - z_h(t_k) \\ \mathbf{0}_{\mathbb{R}^{n_v}}\end{pmatrix}^\top Q(t) \begin{pmatrix} z_h(t_k) \\ v_h(t_k)\end{pmatrix} dt\\
\underset{2.4.11}{\overset{\text{Assumption}}{\leq}}\;& \sum_{k=1}^N \int_{t_{k-1}}^{t_k} \mathbf{L}_Q h\left(\|z_h(t_k)\|^2 + \|v_h(t_k)\|^2\right) dt\\
&+ \sum_{k=1}^N \int_{t_{k-1}}^{t_k} \|Q\|_\infty \left(h^2 \|z_h'(t_k)\|^2 + 2h\|z_h'(t_k)\| \left\|(z_h(t_k), v_h(t_k))\right\|\right) dt\\
\overset{\text{Sobolev}}{\leq}\;& \sum_{k=1}^N \int_{t_{k-1}}^{t_k} \mathbf{L}_Q h\left(\left(2\|z_h\|_{1,2}\right)^2 + \|v_h(t_k)\|^2\right) dt\\
&+ \|Q\|_\infty \left(h^2 \|z_h'\|_2^2 + \sum_{k=1}^N \int_{t_{k-1}}^{t_k} 2h \|z_h'(t_k)\| \left\|(z_h(t_k), v_h(t_k))\right\| dt\right)\\
\underset{\text{Schwarz}}{\overset{\text{Cauchy}}{\leq}}\;& \mathbf{L}_Q h\left(4\|z_h\|_{1,2}^2 + \|v_h\|_2^2\right) + \|Q\|_\infty\left(h^2\|z_h'\|_2^2 + 2h\|z_h'\|_2 \left\|(z_h, v_h)\right\|_2\right)\\
\leq\;& 5\mathbf{L}_Q h \left\|(z_h, v_h)\right\|_{Z_2}^2 + \|Q\|_\infty \left(h^2 + 4h\right) \left\|(z_h, v_h)\right\|_{Z_2}^2.\end{aligned}$$

Thus, since $h + 4 \leq 5$, for $\mathbf{L}_\mathcal{P} := 8\mathbf{L}_\Lambda + 5\mathbf{L}_Q + 5\|Q\|_\infty$ it holds

$$\mathcal{P}\left((z_h, v_h), (z_h, v_h)\right) - \mathcal{P}_h\left((z_h, v_h), (z_h, v_h)\right) \leq \mathbf{L}_\mathcal{P} h \left\|(z_h, v_h)\right\|_{Z_2}^2.$$

All conditions in Theorem 2.1.10 are satisfied, which completes the proof. □

Let us decompose the matrix function $Q(\cdot)$ into $\begin{bmatrix} M(\cdot) & K(\cdot)^\top \\ K(\cdot) & \Pi(\cdot) \end{bmatrix} \in L_\infty^{(n_z+n_v)\times(n_z+n_v)}([0,1])$. If $\mathcal{P}$ is coercive, then $\Pi(\cdot)$ satisfies a *Legendre-Clebsch condition.* To prove this we require the following notion, cf. [50, 2.18 Definition]:

Definition 2.4.13 (Lebesgue Point)
For a locally Lebesgue integrable function $f : [0,1] \to \mathbb{R}^n$ *a point* $s \in (0,1)$ *is a Lebesgue point, if it holds*

$$\lim_{\varepsilon\to 0+} \frac{1}{2\varepsilon} \int_{s-\varepsilon}^{s+\varepsilon} \|f(t) - f(s)\| \, dt = 0.$$

According to [50, 2.19 Theorem], for a function $f \in L_1^n([0,1])$ almost every point $s \in [0,1]$ is a Lebesgue point, which allows us to show the following:

Lemma 2.4.14 (Legendre-Clebsch Condition)
Let system (2.4.1) *satisfy the conditions in Definition 2.4.1. Furthermore, suppose the bilinear form* $\mathcal{P}$ *with*

$$Q(\cdot) = \begin{bmatrix} M(\cdot) & K(\cdot)^\top \\ K(\cdot) & \Pi(\cdot) \end{bmatrix} \in L_\infty^{(n_z+n_v)\times(n_z+n_v)}([0,1])$$

is coercive on $\ker(F)$ *with constant* $\gamma > 0$. *Then, for almost every* $t \in [0,1]$ *and* $\varpi \in \ker(D(t))$ *the Legendre-Clebsch condition*

$$\varpi^\top \Pi(t)\, \varpi \geq \gamma \|\varpi\|^2$$

is satisfied.

Proof. Let $s \in (0,1)$ be an arbitrary Lebesgue point of D and Π, i.e.,

$$\lim_{\varepsilon\to 0+} \frac{1}{2\varepsilon} \int_{s-\varepsilon}^{s+\varepsilon} \|D(t) - D(s)\| \, dt = 0, \tag{2.4.18}$$

$$\lim_{\varepsilon\to 0+} \frac{1}{2\varepsilon} \int_{s-\varepsilon}^{s+\varepsilon} \|\Pi(t) - \Pi(s)\| \, dt = 0. \tag{2.4.19}$$

Additionally, from $0 \leq \|D(t) - D(s)\|^2 \leq 2\|D\|_\infty \|D(t) - D(s)\|$ for almost every $t \in [0,1]$, it follows

$$\lim_{\varepsilon\to 0+} \frac{1}{2\varepsilon} \int_{s-\varepsilon}^{s+\varepsilon} \|D(t) - D(s)\|^2 \, dt = 0. \tag{2.4.20}$$

Let $\varpi \in \ker(D(s))$ be arbitrary. Then, it holds

$$\begin{aligned} \left\|D(t)^\curlywedge D(t)\,\varpi\right\| &\leq \left\|D(\cdot)^\curlywedge\right\|_\infty \|D(t)\,\varpi\| = \left\|D(\cdot)^\curlywedge\right\|_\infty \|D(t)\,\varpi - D(s)\,\varpi\| \\ &\leq \left\|D(\cdot)^\curlywedge\right\|_\infty \|D(t) - D(s)\| \, \|\varpi\| \end{aligned}$$

for almost every $t \in [0,1]$. Thus, (2.4.18) and (2.4.20) yield

$$\lim_{\varepsilon\to 0+} \frac{1}{2\varepsilon} \int_{s-\varepsilon}^{s+\varepsilon} \left\|D(t)^\curlywedge D(t)\,\varpi\right\| dt = 0, \quad \lim_{\varepsilon\to 0+} \frac{1}{2\varepsilon} \int_{s-\varepsilon}^{s+\varepsilon} \left\|D(t)^\curlywedge D(t)\,\varpi\right\|^2 dt = 0. \tag{2.4.21}$$

The control and state

$$w_\varepsilon(\cdot) := \begin{cases} \varpi & , \; t \in [s-\varepsilon, s+\varepsilon] \\ \mathbf{0}_{\mathbb{R}^{n_v}} & , \; \text{otherwise} \end{cases} \quad , \quad z_\varepsilon(\cdot) := \Phi_{\tilde{A}}(\cdot) \int_0^{\cdot} \Phi_{\tilde{A}}(\tau)^{-1} \tilde{B}(\tau) w_\varepsilon(\tau) \, d\tau,$$

satisfy the reduced differential equation

$$\begin{aligned} \dot{z}(t) &= \tilde{A}(t) z(t) + \tilde{B}(t) w(t), \qquad \text{a.e. in } [0,1], \\ z(0) &= \mathbf{0}_{\mathbb{R}^{n_z}}, \end{aligned}$$

and in addition for $t \in [0,1]$ we obtain

$$\|z_\varepsilon(t)\| \le \exp\left(\left\|\tilde{A}\right\|_\infty\right) \left\|\tilde{B}\right\|_\infty \int_{s-\varepsilon}^{s+\varepsilon} \|w_\varepsilon(\tau)\| \, dt \le 2\varepsilon \exp\left(\left\|\tilde{A}\right\|_\infty\right) \left\|\tilde{B}\right\|_\infty \|\varpi\|. \tag{2.4.22}$$

Then, for the control $v_\varepsilon(\cdot) := -D(\cdot)^{\curlywedge} C(\cdot) z_\varepsilon(\cdot) + \left(\mathbf{I}_{n_v} - D(\cdot)^{\curlywedge} D(\cdot)\right) w_\varepsilon(\cdot)$ it holds

$$F(z_\varepsilon, v_\varepsilon) = \begin{pmatrix} \mathbf{0}_{L_\infty^{n_z}([0,1])} \\ \mathbf{0}_{L_\infty^{m}([0,1])} \\ E_1 z_\varepsilon(1) \end{pmatrix}.$$

According to Lemma 2.4.10, there exist $\kappa > 0$ and $(\tilde{z}_\varepsilon, \tilde{v}_\varepsilon) \in Z_\infty$ such that

$$F(\tilde{z}_\varepsilon, \tilde{v}_\varepsilon) = \begin{pmatrix} \mathbf{0}_{L_\infty^{n_z}([0,1])} \\ \mathbf{0}_{L_\infty^{m}([0,1])} \\ -E_1 z_\varepsilon(1) \end{pmatrix}$$

and

$$\kappa \left\|(\tilde{z}_\varepsilon, \tilde{v}_\varepsilon)\right\|_{Z_\infty} \le \|E_1 z_\varepsilon(1)\| \overset{(2.4.22)}{\le} 2\varepsilon \|E_1\| \exp\left(\left\|\tilde{A}\right\|_\infty\right) \left\|\tilde{B}\right\|_\infty \|\varpi\|. \tag{2.4.23}$$

Introduce the new state and control

$$x := z_\varepsilon + \tilde{z}_\varepsilon, \quad u := v_\varepsilon + \tilde{v}_\varepsilon,$$

which due to the linearity of F satisfy $F(x,u) = \mathbf{0}_{Y_\infty}$. Moreover, we denote

$$\tilde{u}_\varepsilon(\cdot) := -D(\cdot)^{\curlywedge} C(\cdot) z_\varepsilon(\cdot) + \tilde{v}_\varepsilon(\cdot),$$

hence $u(\cdot) = \tilde{u}_\varepsilon(\cdot) + \left(\mathbf{I}_{n_v} - D(\cdot)^{\curlywedge} D(\cdot)\right) w_\varepsilon(\cdot)$.

Then, (2.4.22), (2.4.23), and

$$\Gamma := 2\exp\left(\left\|\tilde{A}\right\|_\infty\right)\left\|\tilde{B}\right\|_\infty\left(1+\frac{\|E_1\|}{\kappa}+\left\|D(\cdot)^{\curlywedge}\right\|_\infty\|C\|_\infty\right)$$

yield

$$\|x\|_\infty \leq \Gamma\|\varpi\|\varepsilon, \quad \|\tilde{u}_\varepsilon\|_\infty \leq \Gamma\|\varpi\|\varepsilon.$$

Furthermore, it holds

$$\begin{aligned}
\|u\|_2^2 &= \int_0^1 \left\|\tilde{u}_\varepsilon(t) - D(t)^{\curlywedge} D(t) w_\varepsilon(t) + w_\varepsilon(t)\right\|^2 dt \\
&\geq \int_0^1 \left(\left\|\tilde{u}_\varepsilon(t) - D(t)^{\curlywedge} D(t) w_\varepsilon(t)\right\| - \|w_\varepsilon(t)\|\right)^2 dt \\
&\geq \int_0^1 -2\left\|\tilde{u}_\varepsilon(t) - D(t)^{\curlywedge} D(t) w_\varepsilon(t)\right\|\|w_\varepsilon(t)\| + \|w_\varepsilon(t)\|^2 dt \\
&\geq \int_0^1 -2\|\tilde{u}_\varepsilon(t)\|\|w_\varepsilon(t)\| - 2\left\|D(t)^{\curlywedge} D(t) w_\varepsilon(t)\right\|\|w_\varepsilon(t)\| + \|w_\varepsilon(t)\|^2 dt \\
&\geq \int_{s-\varepsilon}^{s+\varepsilon} -2\Gamma\|\varpi\|\varepsilon\|\varpi\| - 2\left\|D(t)^{\curlywedge} D(t)\varpi\right\|\|\varpi\| + \|\varpi\|^2 dt \\
&= -4\Gamma\|\varpi\|^2\varepsilon^2 - 2\int_{s-\varepsilon}^{s+\varepsilon}\left\|D(t)^{\curlywedge} D(t)\varpi\right\|\|\varpi\| dt + 2\|\varpi\|^2\varepsilon,
\end{aligned}$$

which exploiting

$$\mathcal{P}((x,u),(x,u)) \geq \gamma\|(x,u)\|_{Z_2} \geq \gamma\|u\|_2^2$$

implies

$$\mathcal{P}((x,u),(x,u)) \geq \gamma\left(-4\Gamma\|\varpi\|^2\varepsilon^2 - 2\int_{s-\varepsilon}^{s+\varepsilon}\left\|D(t)^{\curlywedge} D(t)\varpi\right\|\|\varpi\| dt + 2\|\varpi\|^2\varepsilon\right). \tag{2.4.24}$$

For an upper bound we examine

$$
\begin{aligned}
\mathcal{P}\left(\left(x,u\right),\left(x,u\right)\right) &= \begin{pmatrix} x\left(0\right) \\ x\left(1\right) \end{pmatrix}^{\top} \Lambda \begin{pmatrix} x\left(0\right) \\ x\left(1\right) \end{pmatrix} + \int\limits_{0}^{1} \begin{pmatrix} x\left(t\right) \\ u\left(t\right) \end{pmatrix}^{\top} \begin{pmatrix} M\left(t\right) & K\left(t\right)^{\top} \\ K\left(t\right) & \Pi\left(t\right) \end{pmatrix} \begin{pmatrix} x\left(t\right) \\ u\left(t\right) \end{pmatrix} dt \\
&\leq 2\left\|\Lambda\right\|\Gamma^{2}\left\|\varpi\right\|^{2}\varepsilon^{2} + \left\|M\right\|_{\infty}\Gamma^{2}\left\|\varpi\right\|^{2}\varepsilon^{2} + 2\int\limits_{0}^{1} x\left(t\right)^{\top}K\left(t\right)^{\top}\tilde{u}_{\varepsilon}\left(t\right)dt \\
&\quad + 2\int\limits_{0}^{1} x\left(t\right)^{\top}K\left(t\right)^{\top}\left(\mathbf{I}_{n_v} - D\left(t\right)^{\curlywedge}D\left(t\right)\right)w_{\varepsilon}\left(t\right)dt + \int\limits_{0}^{1}\tilde{u}_{\varepsilon}\left(t\right)^{\top}\Pi\left(t\right)\tilde{u}_{\varepsilon}\left(t\right)dt \\
&\quad + 2\int\limits_{0}^{1}\tilde{u}_{\varepsilon}\left(t\right)^{\top}\Pi\left(t\right)\left(\mathbf{I}_{n_v} - D\left(t\right)^{\curlywedge}D\left(t\right)\right)w_{\varepsilon}\left(t\right)dt \\
&\quad + \int\limits_{0}^{1} w_{\varepsilon}\left(t\right)^{\top}\left(\mathbf{I}_{n_v} - D\left(t\right)^{\curlywedge}D\left(t\right)\right)\Pi\left(t\right)\left(\mathbf{I}_{n_v} - D\left(t\right)^{\curlywedge}D\left(t\right)\right)w_{\varepsilon}\left(t\right)dt \\
&\leq \left[\left(2\left\|\Lambda\right\| + \left\|M\right\|_{\infty} + 2\left\|K\right\|_{\infty}\right)\Gamma^{2} + 4\left\|K\right\|_{\infty}\Gamma\left(1 + \left\|D\left(\cdot\right)^{\curlywedge}D\left(\cdot\right)\right\|_{\infty}\right)\right]\left\|\varpi\right\|^{2}\varepsilon^{2} \\
&\quad + \left\|\Pi\right\|_{\infty}\Gamma^{2}\left\|\varpi\right\|^{2}\varepsilon^{2} + 4\left\|\Pi\right\|_{\infty}\Gamma\left\|\varpi\right\|^{2}\left(1 + \left\|D\left(\cdot\right)^{\curlywedge}D\left(\cdot\right)\right\|_{\infty}\right)\varepsilon^{2} \\
&\quad + \left\|\Pi\right\|_{\infty}\int\limits_{s-\varepsilon}^{s+\varepsilon} 2\left\|\varpi\right\|\left\|D\left(t\right)^{\curlywedge}D\left(t\right)\varpi\right\| + \left\|D\left(t\right)^{\curlywedge}D\left(t\right)\varpi\right\|^{2}dt \\
&\quad + \int\limits_{s-\varepsilon}^{s+\varepsilon}\left\|\Pi\left(t\right) - \Pi\left(s\right)\right\|\left\|\varpi\right\|^{2}dt + 2\varepsilon\varpi^{\top}\Pi\left(s\right)\varpi.
\end{aligned}
$$

Dividing the upper bound and lower bound (2.4.24) by 2ε, taking the limit $\varepsilon \to 0+$, and exploiting (2.4.19), (2.4.21) yields

$$
\varpi^{\top}\Pi\left(s\right)\varpi \geq \gamma\left\|\varpi\right\|^{2}, \quad \text{a.e. in } \left(0,1\right),
$$

which proves the assertion. □

Similar to Lemma 2.4.12, we can prove that the Legendre-Clebsch condition holds for the discrete case, if $Q_h\left(\cdot\right)$ is decomposed into $\begin{bmatrix} M_h\left(\cdot\right) & K_h\left(\cdot\right)^{\top} \\ K_h\left(\cdot\right) & \Pi_h\left(\cdot\right) \end{bmatrix} \in L_{\infty,h}^{\left(n_z+n_v\right)\times\left(n_z+n_v\right)}\left(\left[0,1\right]\right)$.

Lemma 2.4.15 (Discrete Legendre-Clebsch Condition)

Let system (2.4.1) *and* (2.4.7) *satisfy the conditions in Definition 2.4.1 and Assumption 2.4.4, respectively. Furthermore, suppose there exists* $\mathbf{L}_{\Pi} \geq 0$ *with* $\left\|\Pi\left(\cdot\right) - \Pi_h\left(\cdot\right)\right\|_{\infty} \leq \mathbf{L}_{\Pi}h$*, and the bilinear form* $\mathcal{P}$ *with* $Q\left(\cdot\right) = \begin{bmatrix} M\left(\cdot\right) & K\left(\cdot\right)^{\top} \\ K\left(\cdot\right) & \Pi\left(\cdot\right) \end{bmatrix} \in L_{\infty}^{\left(n_z+n_v\right)\times\left(n_z+n_v\right)}\left(\left[0,1\right]\right)$ *is coercive on* $\ker\left(F\right)$ *with constant* $\gamma > 0$*. Then, there exist* $\tilde{\gamma}, \tilde{h} > 0$ *such that for every* $0 < h \leq \tilde{h}$*,* $i = 1, \ldots, N$*, and* $\varpi \in \ker\left(D_h\left(t_i\right)\right)$ *the discrete Legendre-Clebsch condition*

$$
\varpi^{\top}\Pi_h\left(t_i\right)\varpi \geq \tilde{\gamma}\left\|\varpi\right\|^{2}
$$

is satisfied.

Proof. Let $i \in \{1, \ldots, N\}$ be arbitrary. Then, by Lemma 2.4.14, for almost every $t \in (t_{i-1}, t_i]$ the symmetric and continuous bilinear form $\varpi^\top \Pi(t) \varpi$ is coercive on $\ker(D(t))$ with constant $\gamma > 0$. Consequently, the conditions (i) - (v) of Theorem 2.1.10 hold for the symmetric and continuous bilinear form $\varpi^\top \Pi_h(t_i) \varpi$, because:

(i) For almost every $t \in (t_{i-1}, t_i]$ let us consider the linear equation

$$D(t) v = b, \quad v \in \mathbb{R}^{n_v}, b \in \mathbb{R}^m,$$

which has the solution $v(t) := D(t)^\curlywedge b$. By Definition 2.4.1 and Remark 2.4.2, the right inverse $D(\cdot)^\curlywedge$ is uniformly bounded. Therefore, for almost every $t \in (t_{i-1}, t_i]$ the linear operator $D(t)$ is uniformly surjective with constant $\frac{1}{\left\|D(\cdot)^\curlywedge\right\|_\infty}$.

(ii) The bilinear form $\varpi^\top \Pi(t) \varpi$ is coercive on $\ker(D(t))$ for almost every $t \in (t_{i-1}, t_i]$.

(iii) According to Lemma 2.4.5 there exists a $h_1 > 0$ such that for all $0 < h \leq h_1$ the matrix function $D_h(\cdot)$ is uniformly linear independent. Using the same arguments as in (i) yields the uniform surjectivity of $D_h(t_i)$.

(iv) By Assumption 2.4.4, for every $v \in \mathbb{R}^{n_v}$ and almost every $t \in (t_{i-1}, t_i]$ it holds

$$\|D(t) v - D_h(t_i) v\| \leq \mathbf{L} h \|v\|.$$

(v) Since $\Pi_h(t) = \Pi_h(t_i)$ for $t \in (t_{i-1}, t_i]$, it holds $\|\Pi(t) - \Pi_h(t_i)\|_\infty \leq \mathbf{L}_\Pi h$ for almost every $t \in (t_{i-1}, t_i]$, due to the assumptions. Thus,

$$\varpi^\top \Pi(t) \varpi - \varpi^\top \Pi_h(t_i) \varpi = \varpi^\top (\Pi(t) - \Pi_h(t_i)) \varpi \leq \mathbf{L}_\Pi \|\varpi\|^2$$

for all $\varpi \in \mathbb{R}^{n_v}$ and almost every $t \in (t_{i-1}, t_i]$.

All conditions in Theorem 2.1.10 are satisfied, which completes the proof. □

Chapter 3

Necessary Conditions

Since the 1950's necessary conditions, also called *maximum principles* or *minimum principles*, for optimal control problems have been studied. Pontryagin et al. [104] and Hestenes [57] contributed early proofs of the maximum principle. Optimal control problems subject to ordinary differential equations with mixed control-state constraints have been studied in [96, 124], where [124] also provides second-order necessary (and sufficient) conditions. Problems with pure state constraints are investigated in, e.g., [59–61, 63, 87].

The order of an inequality constraint is given by the number of time-derivatives it takes for the control to appear in the constraint. Mixed control-state constraints can be considered as the order zero case. In order to derive necessary conditions, it is usually assumed that the constraints satisfy some *regularity condition*. For mixed control-state constraints one could assume that the derivative of the active constraints with respect to the control are linear independent. In case of a pure state constraint of order k, a regularity condition would be that the derivative of the k-th time-derivative with respect to the control is unequal to zero on boundary intervals. For multiple constraints it is assumed that the respective derivatives are linear independent. Additionally, if boundary conditions are present, usually some rank or *controllability condition* is imposed.

There are two main approaches for deriving maximum/minimum principles for optimal control problems with pure state constraints of order k: The *direct adjoining* and the *indirect adjoining* approach. In the former case, the inequality constraints are directly adjoined to the Hamilton function, whereas in the latter case, the k-th time-derivative is adjoined to the Hamilton function. The direct approach is used in, e.g., [59, 60, 89, 90]. For pure state constraints of order one the indirect approach was used in [104], and for constraints of higher order it was used in [86]. More references can be found in the survey paper on maximum principles [56].

The maximum/minimum principles are generally expressed in form of *adjoint differential equations*, *transversality conditions*, *complementarity conditions*, and *jump conditions*. The latter conditions appear, if pure state constraints are imposed. In that case, the adjoint multiplier may have discontinuities at junction or contact points. For higher order inequality constraints the multipliers are usually only piecewise absolutely continuous, piecewise continuous, or of bounded variation, as opposed to the case with mixed control-state constraints, where the adjoint multiplier is absolutely continuous.

Linear quadratic DAE optimal control problems are discussed in, e.g., [8, 65, 94]. Descriptor systems with constant coefficient matrices are considered in [94], whereas time-variant systems are considered in [65]. In [8], nonlinear quasi-linear DAEs are considered as well. Optimal control problems with index one DAEs in semi-explicit form with set constraints on the controls and with pure state and mixed control-state constraints are investigated in [102] and [29,47], respectively. In the index one case, the algebraic constraint can be directly adjoined to the Hamilton function. Thus, the results coincide with results for problems with explicit ODEs. This changes for higher index DAEs. In [8, Example 3.16], it was shown that the usual necessary conditions do not hold for the Hamilton function, where the algebraic constraint was directly adjoined. Instead, the time-derivative of the algebraic constraint, where the algebraic variable first appears, is adjoined, similar to indirect adjoining approach for pure state constraints. In the index k case, one would adjoin the $(k-1)$-st derivative to the Hamilton function. Necessary conditions for problems with higher index DAEs were derived in [45, 47, 83, 111]. Index two DAEs with pure state constraints, mixed control-state constraints, and set constraints on the controls are considered in [45,47]. In [111], problems with Hessenberg DAEs up to index three are examined, and in [83] problems with Hessenberg DAEs of arbitrary order are analyzed. In [28], a maximum principle for problems with implicit control systems is derived, by reducing the optimal control problem to an equivalent nonsmooth variational problem. General unstructured DAE optimal control problems are investigated in [67]. [69, 88] provide necessary conditions for infinite optimization problems, which are closely related to optimal control problems.

We aim to derive a local minimum principle for optimal control problems subject to Hessenberg DAEs of arbitrary order and mixed control-state constraints, similar to [83]. By including boundary conditions and weakening the regularity assumptions on the mixed control-state constraints we expand the results of [83, Theorem 3.1]. Furthermore, Theorem 3.2.5 generalizes the results of [47, Theorem 3.3.8]. In order to derive necessary conditions for problems with higher index DAEs, we first consider DAEs with index one, for which we derive necessary conditions using the techniques developed in [81]. By reducing the index of a Hessenberg DAEs with higher index to an equivalent system with index one we obtain a local minimum principle for optimal control problems subject to Hessenberg DAEs with arbitrary index.

For the derivation we use the following scheme (see Figure 3.1):

(a) Modify the optimal control problem by introducing a slack variable to the mixed control-state constraints such that a weak local minimizer of the initial problem is also a solution of the modified problem, if the slack variable is set to zero.

(b) Interpret the modified problem as an infinite optimization problem and use the results in [69], which yields non trivial Lagrange multipliers.

(c) Deduce an explicit representation of these Lagrange multipliers and apply variation lemmas to derive a local minimum principle for the modified problem.

(d) Show that the Lagrange multipliers for the initial problem associated with the weak local minimizer are equal to the Lagrange multipliers for the modified problem associated with the same weak local minimizer and the slack variable equal to zero.

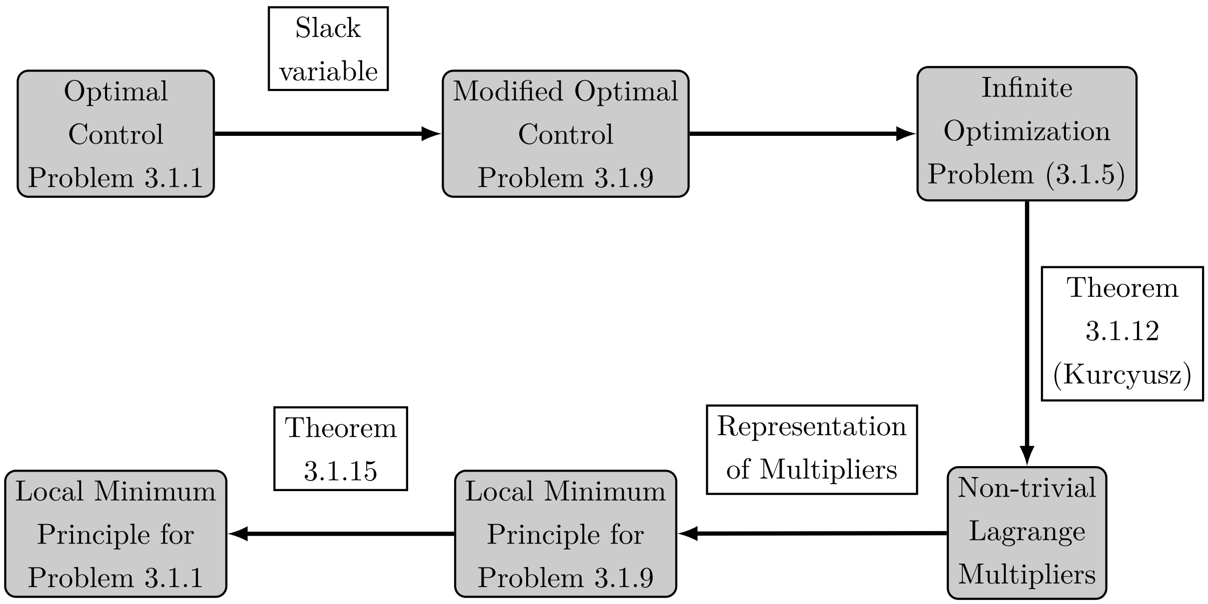

Figure 3.1: Scheme to derive necessary conditions for Problem 3.1.1.

3.1 Necessary Conditions for Index One Problems

Consider the following optimal control problem:

Problem 3.1.1 (Optimal Control Problem with Index One DAE)

Let $n_x, n_y, n_u, n_\psi, n_c \in \mathbb{N}$ *with* $n_\psi \leq 2n_x$, $n_c \leq n_u$. *Let*

$$\begin{aligned} &\varphi : \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \to \mathbb{R}, && \psi : \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_\psi}, \\ &f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x}, && g : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_y}, \quad c : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_c}, \end{aligned}$$

be functions.

$$\begin{array}{lrcll} \textit{Minimize} & \multicolumn{4}{l}{\varphi(x(0), x(1)),} \\ \textit{with respect to} & \multicolumn{4}{l}{x \in W_{1,\infty}^{n_x}([0,1]), y \in L_\infty^{n_y}([0,1]), u \in L_\infty^{n_u}([0,1]),} \\ \textit{subject to} & \dot{x}(t) & = & f(x(t), y(t), u(t)), & \textit{a.e. in } [0,1], \\ & \mathbf{0}_{\mathbb{R}^{n_y}} & = & g(x(t), y(t), u(t)), & \textit{a.e. in } [0,1], \\ & \mathbf{0}_{\mathbb{R}^{n_\psi}} & = & \psi(x(0), x(1)), & \\ & \mathbf{0}_{\mathbb{R}^{n_c}} & \geq & c(x(t), y(t), u(t)), & \textit{a.e. in } [0,1]. \end{array}$$

In order to derive necessary conditions, we assume Problem 3.1.1 has a weak local minimizer and the functions are sufficiently smooth.

Assumption 3.1.2

(3.1.A1) *(Existence of a Minimizer)*
Let $(\hat{x}, \hat{y}, \hat{u}) \in W^{n_x}_{1,\infty}([0,1]) \times L^{n_y}_{\infty}([0,1]) \times L^{n_u}_{\infty}([0,1])$ *be a weak local minimizer of Problem 3.1.1.*

(3.1.A2) *(Smoothness of the System)*

(a) φ *and* ψ *are continuously differentiable with respect to all arguments.*

(b) For a sufficiently large convex compact neighborhood $\mathcal{M}$ *of*

$$\{(\hat{x}(t), \hat{y}(t), \hat{u}(t)) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \mid t \in [0,1]\},$$

let the mappings

$$(x,y,u) \mapsto f(x,y,u), \quad (x,y,u) \mapsto g(x,y,u), \quad (x,y,u) \mapsto c(x,y,u),$$

be continuously differentiable, and the derivatives

$$f'_{(x,y,u)}, \quad g'_{(x,y,u)}, \quad c'_{(x,y,u)},$$

be bounded in $\mathcal{M}$. *Set*

$$\Gamma := \sup_{(x,y,u)\in\mathcal{M}} \max\Big\{\left\|f'_{(x,y,u)}(x,y,u)\right\|, \left\|g'_{(x,y,u)}(x,y,u)\right\|, \left\|c'_{(x,y,u)}(x,y,u)\right\|, \\ \|f(x,y,u)\|, \|g(x,y,u)\|, \|c(x,y,u)\|\Big\}$$

We abbreviate the derivatives at the minimizer by

$$\begin{array}{lll} A_f(\cdot) := f'_x[\cdot], & B_f(\cdot) := f'_y[\cdot], & C_f(\cdot) := f'_u[\cdot], \\ A_g(\cdot) := g'_x[\cdot], & B_g(\cdot) := g'_y[\cdot], & C_g(\cdot) := g'_u[\cdot], \\ \Psi_0 := \psi'_{x_0}(\hat{x}(0), \hat{x}(1)), & \Psi_1 := \psi'_{x_1}(\hat{x}(0), \hat{x}(1)), & \\ A_c(\cdot) := c'_x[\cdot], & B_c(\cdot) := c'_y[\cdot], & C_c(\cdot) := c'_u[\cdot]. \end{array}$$

In [83], a rank condition was introduced that included full row rank of $[B_c(\cdot), C_c(\cdot)]$. This assumption is often too strong, since it does not even hold for simple box constraints. To weaken that assumption we define

$$\begin{aligned} c_j^{\alpha}(\cdot) &:= \min\{c_j[\cdot] + \alpha, 0\}, \quad j \in J := \{1, \ldots, n_c\}, \\ D^{\alpha}(\cdot) &:= \operatorname{diag}\left[c_j^{\alpha}(\cdot)\right]_{j\in J}, \\ E^{\alpha}(\cdot) &:= \begin{bmatrix} B_g(\cdot) & C_g(\cdot) & \mathbf{0}_{n_y\times n_c} \\ B_c(\cdot) & C_c(\cdot) & D^{\alpha}(\cdot) \end{bmatrix}. \end{aligned}$$

for a constant $\alpha \geq 0$, and assume the following linear independence and controllability conditions:

Assumption 3.1.3 (Linear Independence, Controllability)

(3.1.A3) *(Index One / Regularity Condition)*
There exist constants $\alpha > 0$ and $\beta > 0$ such that for all $\varpi \in \mathbb{R}^{n_y+n_c}$ it holds

$$\left\| E^{\alpha}(t)^{\top} \varpi \right\| \geq \beta \|\varpi\|, \quad a.e. \ in \ [0,1].$$

(3.1.A4) *(Controllability)*
For any $e \in \mathbb{R}^{n_\psi}$ there exist $(x,y,u,v) \in W_{1,\infty}^{n_x}([0,1]) \times L_{\infty}^{n_y}([0,1]) \times L_{\infty}^{n_u}([0,1]) \times L_{\infty}^{n_c}([0,1])$ such that the DAE

$$\begin{aligned}
\dot{x}(t) &= A_f(t)\,x(t) + B_f(t)\,y(t) + C_f(t)\,u(t), && a.e. \ in \ [0,1], \\
\mathbf{0}_{\mathbb{R}^{n_y}} &= A_g(t)\,x(t) + B_g(t)\,y(t) + C_g(t)\,u(t), && a.e. \ in \ [0,1], \\
e &= \Psi_0 x(0) + \Psi_1 x(1), && \\
\mathbf{0}_{\mathbb{R}^{n_c}} &= A_c(t)\,x(t) + B_c(t)\,y(t) + C_c(t)\,u(t) + D^{\alpha}(t)\,v(t), && a.e. \ in \ [0,1],
\end{aligned}$$

is satisfied.

Remark 3.1.4
Note that (3.1.A3) combines the index one property of the DAE and the regularity of the inequality constraint. We require this condition in order to reduce the linearized system as in Remark 2.4.2. One could also assume that it is possible to first solve the algebraic equation $\mathbf{0}_{\mathbb{R}^{n_y}} = A_g(\cdot)\,x(\cdot) + B_g(\cdot)\,y(\cdot) + C_g(\cdot)\,u(\cdot)$ for $y(\cdot)$, and then solve the remaining equation $\mathbf{0}_{\mathbb{R}^{n_c}} = A_c(\cdot)\,x(\cdot) + B_c(\cdot)\,y(\cdot) + C_c(\cdot)\,u(\cdot) + D^{\alpha}(\cdot)\,v(\cdot)$ with inserted $y(\cdot)$ for $u(\cdot)$ and $v(\cdot)$. However, condition (3.1.A3) is more general. Furthermore, assumption (3.1.A3) excludes active pure state constraints. To see this, let there exist $\hat{j} \in J$ with $c_{\hat{j}}(x,y,u) := s(x)$, where $s(\cdot)$ is continuously differentiable. Then, we can choose $\varpi \in \mathbb{R}^{n_y+n_c}$ with

$$\varpi_j := \begin{cases} 1, & \text{if } j = n_y + \hat{j} \\ 0, & \text{otherwise} \end{cases}$$

in (3.1.A3), which implies

$$\left\| E^{\alpha}(t)^{\top} \varpi \right\| = \left|\min\{s(\hat{x}(t)) + \alpha, 0\}\right| \geq \beta \|\varpi\| = \beta, \quad a.e. \ in \ [0,1].$$

Since $s(\hat{x}(\cdot))$ is continuous and non-positive on $[0,1]$ we obtain

$$s(\hat{x}(t)) \leq -\alpha - \beta < 0, \quad in \ [0,1].$$

Thus, in our case the pure state constraint does not have active boundary arcs, contact points, or touch points. Therefore, without loss of generality we can assume that for all $j \in J$ and $(x,y,u) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u}$ it holds $\nabla_{(y,u)} c_j(x,y,u) \neq \mathbf{0}_{\mathbb{R}^{n_y+n_u}}$, i.e., pure state constraints do not occur.

For a constant $\tilde{\alpha} \geq 0$ we denote by $J^{\tilde{\alpha}}(\cdot) := \{j \in J \mid c_j[\cdot] \geq -\tilde{\alpha}\}$ the set of indexes for $\tilde{\alpha}-$active constraints, and by $j^{\tilde{\alpha}}(\cdot) := \operatorname{card}\left(J^{\tilde{\alpha}}(\cdot)\right)$ the number of $\tilde{\alpha}-$active constraints. Moreover, we define the matrix functions

$$A_c^{\tilde{\alpha}}(\cdot) := \left[c'_{j,x}[\cdot]\right]_{j \in J^{\tilde{\alpha}}(\cdot)}, \quad B_c^{\tilde{\alpha}}(\cdot) := \left[c'_{j,y}[\cdot]\right]_{j \in J^{\tilde{\alpha}}(\cdot)}, \quad C_c^{\tilde{\alpha}}(\cdot) := \left[c'_{j,u}[\cdot]\right]_{j \in J^{\tilde{\alpha}}(\cdot)},$$

which we consider to be vacuous, if $J^{\tilde{\alpha}}(t)$ is empty.

Lemma 3.1.5

Suppose (3.1.A1) and (3.1.A2) hold. Then, (3.1.A3) is satisfied, if and only if there exist $\tilde{\alpha} > 0$ and $\tilde{\beta} > 0$ such that for almost every $t \in [0,1]$ and every $b^{\tilde{\alpha}} \in \mathbb{R}^{n_y + j^{\tilde{\alpha}}(t)}$ it holds

$$\left\| \begin{bmatrix} B_g(t) & C_g(t) \\ B_c^{\tilde{\alpha}}(t) & C_c^{\tilde{\alpha}}(t) \end{bmatrix}^\top b^{\tilde{\alpha}} \right\| \geq \tilde{\beta} \left\| b^{\tilde{\alpha}} \right\|. \tag{3.1.1}$$

Proof. First, let (3.1.A3) hold for $\alpha, \beta > 0$ and set $\tilde{\alpha} = \alpha$, $\tilde{\beta} = \beta$. For almost every $t \in [0,1]$ and an arbitrary $b^{\tilde{\alpha}} = \begin{pmatrix} b_g \\ b_c^{\tilde{\alpha}} \end{pmatrix} \in \mathbb{R}^{n_y} \times \mathbb{R}^{j^{\tilde{\alpha}}(t)}$ we define $\varpi_c \in \mathbb{R}^{n_c}$ as

$$\varpi_{c,j} = \begin{cases} b_{c\ ,j}^{\tilde{\alpha}}, & \text{if } j \in J^{\tilde{\alpha}}(t) \\ 0, & \text{otherwise} \end{cases} \qquad j \in \{1, \ldots, n_c\},$$

and set $\varpi := \begin{pmatrix} b_g \\ \varpi_c \end{pmatrix}$. Then, it holds for almost every $t \in [0,1]$

$$\begin{aligned} \left\| \begin{bmatrix} B_g(t)^\top & B_c^{\tilde{\alpha}}(t)^\top \\ C_g(t)^\top & C_c^{\tilde{\alpha}}(t)^\top \end{bmatrix} b^{\tilde{\alpha}} \right\| &= \left\| \begin{bmatrix} B_g(t)^\top & B_c^{\tilde{\alpha}}(t)^\top \\ C_g(t)^\top & C_c^{\tilde{\alpha}}(t)^\top \\ \mathbf{0}_{n_c \times n_y} & \mathbf{0}_{n_c \times j^{\tilde{\alpha}}(t)} \end{bmatrix} b^{\tilde{\alpha}} \right\| = \left\| \begin{bmatrix} B_g(t)^\top & B_c(t)^\top \\ C_g(t)^\top & C_c(t)^\top \\ \mathbf{0}_{n_c \times n_y} & D^\alpha(t)^\top \end{bmatrix} \varpi \right\| \\ &= \left\| E^\alpha(t)^\top \varpi \right\| \overset{(3.1.A3)}{\geq} \beta \left\| \varpi \right\| = \tilde{\beta} \left\| b^{\tilde{\alpha}} \right\|, \end{aligned}$$

which proves the assertion.

Now, suppose (3.1.1) holds for $\tilde{\alpha}, \tilde{\beta} > 0$ and set $\alpha := \frac{\tilde{\alpha}}{2}$, $\beta := \frac{\tilde{\alpha}\tilde{\beta}}{\tilde{\alpha} + 2\tilde{\beta} + 2\|B_c\|_\infty + 2\|C_c\|_\infty}$. For almost every $t \in [0,1]$ and an arbitrary $\varpi := \begin{pmatrix} \varpi_g \\ \varpi_c \end{pmatrix} \in \mathbb{R}^{n_y} \times \mathbb{R}^{n_c}$ we define

$$\begin{pmatrix} a_B \\ a_C \\ a_{D^\alpha} \end{pmatrix} := \begin{bmatrix} B_g(t)^\top & B_c(t)^\top \\ C_g(t)^\top & C_c(t)^\top \\ \mathbf{0}_{n_c \times n_y} & D^\alpha(t)^\top \end{bmatrix} \varpi,$$

and set $b^{\tilde{\alpha}} := \begin{pmatrix} \varpi_g \\ b_c^{\tilde{\alpha}} \end{pmatrix} \in \mathbb{R}^{n_y} \times \mathbb{R}^{j^{\tilde{\alpha}}(t)}$ with $b_{c\ ,j}^{\tilde{\alpha}} = \varpi_{c,j}$ for $j \in J^{\tilde{\alpha}}(t)$. For all $j \in J \setminus J^{\tilde{\alpha}}(t)$ it holds $c_j^\alpha(t) = \min\{c_j[t] + \alpha, 0\} < \min\{-\tilde{\alpha} + \alpha, 0\} = -\alpha < 0$ for almost every $t \in [0,1]$. We denote the diagonal matrices

$$D_{\tilde{\alpha},+}^\alpha(\cdot) := \operatorname{diag}\left[c_j^\alpha(\cdot)\right]_{j \in J^{\tilde{\alpha}}(\cdot)}, \quad D_{\tilde{\alpha},-}^\alpha(\cdot) := \operatorname{diag}\left[c_j^\alpha(\cdot)\right]_{j \in J \setminus J^{\tilde{\alpha}}(\cdot)},$$

where $D^{\alpha}_{\tilde{\alpha},-}(t)$ is non-singular and satisfies $\left\| D^{\alpha}_{\tilde{\alpha},-}(t)^{-1} \right\| \leq \max_{j \in J \setminus J^{\tilde{\alpha}}(t)} \frac{1}{|c_j^{\alpha}(t)|} \leq \frac{1}{\alpha}$ for almost every $t \in [0,1]$. Additionally, we denote $\varpi_c^{\tilde{\alpha},-} := \left(\varpi_{c,j}\right)_{j \in J \setminus J^{\tilde{\alpha}}(t)}$ and decompose $a_{D^{\alpha}} = D^{\alpha}(t)^{\top} \varpi_c$ into

$$a_{D^{\alpha}}^{\tilde{\alpha},+} := D^{\alpha}_{\tilde{\alpha},+}(t)\, b_c^{\tilde{\alpha}}, \quad a_{D^{\alpha}}^{\tilde{\alpha},-} := D^{\alpha}_{\tilde{\alpha},-}(t)\, \varpi_c^{\tilde{\alpha},-},$$

which yields

$$\left\| \varpi_c^{\tilde{\alpha},-} \right\| = \left\| D^{\alpha}_{\tilde{\alpha},-}(t)^{-1} a_{D^{\alpha}}^{\tilde{\alpha},-} \right\| \leq \frac{1}{\alpha} \left\| a_{D^{\alpha}}^{\tilde{\alpha},-} \right\| \leq \frac{1}{\alpha} \left\| \begin{pmatrix} a_B \\ a_C \\ a_{D^{\alpha}} \end{pmatrix} \right\|$$

for almost every $t \in [0,1]$. With $B_c^{\tilde{\alpha},-}(\cdot) := \left[c'_{j,y}[\cdot]\right]_{j \in J \setminus J^{\tilde{\alpha}}(\cdot)}, C_c^{\tilde{\alpha},-}(\cdot) := \left[c'_{j,u}[\cdot]\right]_{j \in J \setminus J^{\tilde{\alpha}}(\cdot)}$ we get

$$\begin{bmatrix} B_g(t)^{\top} & B_c^{\tilde{\alpha}}(t)^{\top} & B_c^{\tilde{\alpha},-}(t) \\ C_g(t)^{\top} & C_c^{\tilde{\alpha}}(t)^{\top} & C_c^{\tilde{\alpha},-}(t) \end{bmatrix} \begin{pmatrix} \varpi_g \\ b_c^{\tilde{\alpha}} \\ \varpi_c^{\tilde{\alpha},-} \end{pmatrix} = \begin{pmatrix} a_B \\ a_C \end{pmatrix}$$

for almost every $t \in [0,1]$. By reordering this equation we obtain

$$\begin{bmatrix} B_g(t)^{\top} & B_c^{\tilde{\alpha}}(t)^{\top} \\ C_g(t)^{\top} & C_c^{\tilde{\alpha}}(t)^{\top} \end{bmatrix} \begin{pmatrix} \varpi_g \\ b_c^{\tilde{\alpha}} \end{pmatrix} = \begin{pmatrix} a_B \\ a_C \end{pmatrix} - \begin{bmatrix} B_c^{\tilde{\alpha},-}(t) \\ C_c^{\tilde{\alpha},-}(t) \end{bmatrix} \varpi_c^{\tilde{\alpha},-}$$

for almost every $t \in [0,1]$. Exploiting (3.1.1) yields

$$\begin{aligned} \tilde{\beta} \left\| b^{\tilde{\alpha}} \right\| &\leq \left\| \begin{bmatrix} B_g(t)^{\top} & B_c^{\tilde{\alpha}}(t)^{\top} \\ C_g(t)^{\top} & C_c^{\tilde{\alpha}}(t)^{\top} \end{bmatrix} b^{\tilde{\alpha}} \right\| = \left\| \begin{bmatrix} B_g(t)^{\top} & B_c^{\tilde{\alpha}}(t)^{\top} \\ C_g(t)^{\top} & C_c^{\tilde{\alpha}}(t)^{\top} \end{bmatrix} \begin{pmatrix} \varpi_g \\ b_c^{\tilde{\alpha}} \end{pmatrix} \right\| \\ &\leq \left\| \begin{pmatrix} a_B \\ a_C \end{pmatrix} \right\| + \left\| \begin{bmatrix} B_c^{\tilde{\alpha},-}(t) \\ C_c^{\tilde{\alpha},-}(t) \end{bmatrix} \varpi_c^{\tilde{\alpha},-} \right\| \leq \left(1 + \frac{\|B_c\|_{\infty} + \|C_c\|_{\infty}}{\alpha}\right) \left\| \begin{pmatrix} a_B \\ a_C \\ a_{D^{\alpha}} \end{pmatrix} \right\| \end{aligned}$$

for almost every $t \in [0,1]$. Finally, using $\alpha = \frac{\tilde{\alpha}}{2}$ we conclude

$$\begin{aligned} \|\varpi\| &\leq \left\| b^{\tilde{\alpha}} \right\| + \left\| \varpi_c^{\tilde{\alpha},-} \right\| \leq \left(\frac{1}{\tilde{\beta}} + \frac{\|B_c\|_{\infty} + \|C_c\|_{\infty}}{\alpha \tilde{\beta}} + \frac{1}{\alpha} \right) \left\| \begin{pmatrix} a_B \\ a_C \\ a_{D^{\alpha}} \end{pmatrix} \right\| \\ &\leq \left(\frac{\tilde{\alpha} + 2\tilde{\beta} + 2\|B_c\|_{\infty} + 2\|C_c\|_{\infty}}{\tilde{\alpha}\tilde{\beta}} \right) \left\| \begin{pmatrix} a_B \\ a_C \\ a_{D^{\alpha}} \end{pmatrix} \right\| \\ &= \frac{1}{\beta} \left\| \begin{bmatrix} B_g(t)^{\top} & B_c(t)^{\top} \\ C_g(t)^{\top} & C_c(t)^{\top} \\ \mathbf{0}_{n_c \times n_y} & D^{\alpha}(t)^{\top} \end{bmatrix} \varpi \right\| \end{aligned}$$

for almost every $t \in [0,1]$, which completes the proof. □

Remark 3.1.6

Lemma 3.1.5 allows us to substitute assumption (3.1.A3) for condition (3.1.1) *in the local minimum principle in Theorem 3.1.15. In the same way, the requirements of Theorem 3.2.5 can be adjusted. For convenience, we use a condition similar to* (3.1.1) *instead of assumption (3.2.A3) in Chapter 5 for the index two case.*

Let us consider the linearized DAE

$$\begin{aligned}
\dot{x}(t) &= A_f(t)\,x(t) + B_f(t)\,y(t) + C_f(t)\,u(t), && \text{a.e. in } [0,1], \\
\mathbf{0}_{\mathbb{R}^{n_y}} &= A_g(t)\,x(t) + B_g(t)\,y(t) + C_g(t)\,u(t), && \text{a.e. in } [0,1], \qquad (3.1.2)\\
\mathbf{0}_{\mathbb{R}^{n_\psi}} &= \Psi_0 x(0) + \Psi_1 x(1), \\
\mathbf{0}_{\mathbb{R}^{n_c}} &= A_c(t)\,x(t) + B_c(t)\,y(t) + C_c(t)\,u(t) + D^\alpha(t)\,v(t), && \text{a.e. in } [0,1],
\end{aligned}$$

which we studied in Section 2.4. If (3.1.A1) - (3.1.A3) hold, then we are able to reduce this system analog to Remark 2.4.2, i.e., for $w \in L_\infty^{n_y+n_u+n_c}([0,1])$ we get the system

$$\begin{aligned}
\dot{x}(t) &= \tilde{A}^\alpha(t)\,x(t) + \tilde{B}^\alpha(t)\,w(t), && \text{a.e. in } [0,1], \\
\mathbf{0}_{\mathbb{R}^{n_\psi}} &= \Psi_0 x(0) + \Psi_1 x(1),
\end{aligned}$$

where

$$\begin{aligned}
\tilde{A}^\alpha(\cdot) &:= A_f(\cdot) - \left(B_f(\cdot), C_f(\cdot), \mathbf{0}_{n_x\times n_c}\right) E^\alpha(\cdot)^\curlywedge \begin{pmatrix} A_g(\cdot) \\ A_c(\cdot) \end{pmatrix}, \\
\tilde{B}^\alpha(\cdot) &:= \left(B_f(\cdot), C_f(\cdot), \mathbf{0}_{n_x\times n_c}\right) \Big(\mathbf{I}_{n_y+n_u+n_c} - E^\alpha(\cdot)^\curlywedge E^\alpha(\cdot)\Big).
\end{aligned}$$

Moreover, we define the Gramian matrix G as

$$\begin{aligned}
R &:= \Psi_0 + \Psi_1 \Phi_{\tilde{A}^\alpha}(1), \\
S(\cdot) &:= \Psi_1 \Phi_{\tilde{A}^\alpha}(1)\,\Phi_{\tilde{A}^\alpha}(\cdot)^{-1}\,\tilde{B}^\alpha(\cdot), \\
G &:= RR^\top + \int_0^1 S(t)\,S(t)^\top\,dt.
\end{aligned}$$

Then, according to Lemma 2.4.3, the following holds:

Lemma 3.1.7

Let (3.1.A1) - (3.1.A3) be satisfied. Then, (3.1.A4) holds, if and only if $\operatorname{rank}(G) = n_\psi$.

Remark 3.1.8

Similar to Lemma 3.1.5 we can replace

$$\mathbf{0}_{\mathbb{R}^{n_c}} = A_c(t)\,x(t) + B_c(t)\,y(t) + C_c(t)\,u(t) + D^\alpha(t)\,v(t), \qquad \textit{a.e. in } [0,1] \tag{3.1.3}$$

in condition (3.1.A4) with

$$\mathbf{0}_{\mathbb{R}^{j^{\tilde{\alpha}}(t)}} = A_c^{\tilde{\alpha}}(t)\,x(t) + B_c^{\tilde{\alpha}}(t)\,y(t) + C_c^{\tilde{\alpha}}(t)\,u(t), \qquad \textit{a.e. in } [0,1], \tag{3.1.4}$$

which yields reduced matrices different from $\tilde{A}^{\alpha}(\cdot)$, $\tilde{B}^{\alpha}(\cdot)$, and therefore a different Gramian matrix $\bar{G}$. Since (3.1.1) *and (3.1.A3) are equivalent, the Gramian matrix $\bar{G}$ exists, if and only if G exists, which follows from the reduction of the respective linear DAEs (compare Remark 2.4.2). Using Lemma 2.4.3 and the structure of the Gramian matrices we conclude that the linear DAE* (3.1.2) *with* (3.1.4) *instead of* (3.1.3) *is completely controllable, if and only if (3.1.A4) holds. Thus, the result in Theorem 3.1.15 remains valid, if we use this new controllability condition and* (3.1.1) *instead of (3.1.A4) and (3.1.A3), respectively. Furthermore, the requirements of Theorem 3.2.5 can be adjusted in the same way.*

In order to derive a local minimum principle for Problem 3.1.1, we consider the following modification of Problem 3.1.1:

Problem 3.1.9 (Modified Optimal Control Problem)
Let (3.1.A1) - (3.1.A3) hold for constants $\alpha, \beta > 0$.

$$\begin{array}{lllll}
\textit{Minimize} & \multicolumn{4}{l}{\varphi(x(0), x(1)) + \frac{1}{2}\int_0^1 \|v(t)\|^2\, dt,} \\[2ex]
\textit{with respect to} & \multicolumn{4}{l}{x \in W_{1,\infty}^{n_x}([0,1]),\, y \in L_\infty^{n_y}([0,1]),\, u \in L_\infty^{n_u}([0,1]),\, v \in L_\infty^{n_c}([0,1]),} \\[2ex]
\textit{subject to} & \dot{x}(t) & = & f(x(t), y(t), u(t)), & \textit{a.e. in } [0,1], \\
 & \mathbf{0}_{\mathbb{R}^{n_y}} & = & g(x(t), y(t), u(t)), & \textit{a.e. in } [0,1], \\
 & \mathbf{0}_{\mathbb{R}^{n_\psi}} & = & \psi(x(0), x(1)), & \\
 & \mathbf{0}_{\mathbb{R}^{n_c}} & \geq & c(x(t), y(t), u(t)) + D^{\alpha}(t)\, v(t), & \textit{a.e. in } [0,1].
\end{array}$$

The obvious choice for a weak local minimizer of Problem 3.1.9 would be $\left(\hat{x}, \hat{y}, \hat{u}, \mathbf{0}_{L_\infty^{n_c}([0,1])}\right)$, which we show in the following lemma:

Lemma 3.1.10
Let (3.1.A1) - (3.1.A3) hold for constants $\alpha, \beta > 0$. Then, $\left(\hat{x}, \hat{y}, \hat{u}, \mathbf{0}_{L_\infty^{n_c}([0,1])}\right)$ is a weak local minimizer of Problem 3.1.9.

Proof. Since $\left(\hat{x}, \hat{y}, \hat{u}, \mathbf{0}_{L_\infty^{n_c}([0,1])}\right)$ is feasible for Problem 3.1.9, it remains to show (local) optimality. Set $\rho := \frac{\alpha}{2\Gamma}$ with Γ defined in (3.1.A2), and choose an arbitrary

$$(x, y, u, v) \in \mathcal{B}_\rho\left(\left(\hat{x}, \hat{y}, \hat{u}, \mathbf{0}_{L_\infty^{n_c}([0,1])}\right)\right),$$

which is admissible for Problem 3.1.9. Note that (x, y, u) satisfies the constraints of Problem 3.1.1, except for the inequality constraints. For almost every $t \in [0,1]$ and each $j \in J^{\alpha}(t)$ it holds $c_j^{\alpha}(t) = 0$, hence

$$c_j(x(t), y(t), u(t)) = c_j(x(t), y(t), u(t)) + c_j^{\alpha}(t)\, v_j(t) \leq 0.$$

In addition, using the mean-value theorem in [59, p. 40] yields

$$\begin{aligned}
|c_j(x(t), y(t), u(t)) - c_j(\hat{x}(t), \hat{y}(t), \hat{u}(t))| \\
&\leq \sup_{\theta \in (0,1)} \left| c_{j,(x,y,u)}((1-\theta)(x(t), y(t), u(t)) + \theta(\hat{x}(t), \hat{y}(t), \hat{u}(t))) \right| \\
&\qquad \|(x(t), y(t), u(t)) - (\hat{x}(t), \hat{y}(t), \hat{u}(t))\| \\
&\leq \Gamma \|(x(t), y(t), u(t)) - (\hat{x}(t), \hat{y}(t), \hat{u}(t))\| \\
&\leq \Gamma \rho \leq \frac{\alpha}{2}
\end{aligned}$$

for all $j \in J \setminus J^{\alpha}(t)$, thus $c_j(x(t), y(t), u(t)) \leq \frac{\alpha}{2} + c_j(\hat{x}(t), \hat{y}(t), \hat{u}(t)) < -\frac{\alpha}{2} < 0$ is satisfied for almost every $t \in [0,1]$. Consequently, it holds $c(x(t), y(t), u(t)) \leq \mathbf{0}_{\mathbb{R}^{n_c}}$ for almost every $t \in [0,1]$, and therefore (x, y, u) is feasible for Problem 3.1.1.

Now, assume that $\left(\hat{x}, \hat{y}, \hat{u}, \mathbf{0}_{L_{\infty}^{n_c}([0,1])}\right)$ is not a weak local minimizer of Problem 3.1.9. Then, for any $\tilde{\rho} \leq \rho$ there exists $(\tilde{x}, \tilde{y}, \tilde{u}, \tilde{v}) \in \mathcal{B}_{\tilde{\rho}}\left(\left(\hat{x}, \hat{y}, \hat{u}, \mathbf{0}_{L_{\infty}^{n_c}([0,1])}\right)\right)$, which is admissible for Problem 3.1.9 and satisfies

$$\varphi(\tilde{x}(0), \tilde{x}(1)) + \frac{1}{2} \int_0^1 \|\tilde{v}(t)\|^2 \, dt < \varphi(\hat{x}(0), \hat{x}(1)).$$

It follows that $(\tilde{x}, \tilde{y}, \tilde{u})$ is feasible for Problem 3.1.1 and $\varphi(\tilde{x}(0), \tilde{x}(1)) < \varphi(\hat{x}(0), \hat{x}(1))$, which contradicts the (local) optimality of $(\hat{x}, \hat{y}, \hat{u})$. □

Next, we rewrite Problem 3.1.9 as an *infinite optimization problem* of the form

Problem 3.1.11 (Infinite Optimization Problem)

Let Z, V be spaces, $K \subseteq V$ a cone, and $\mathcal{J} : Z \to \mathbb{R}$, $F : Z \to V$ a functional and operator, respectively.

$$\begin{array}{ll}
\textit{Minimize} & \mathcal{J}(z), \\
\textit{with respect to} & z \in Z, \\
\textit{subject to} & F(z) \in K,
\end{array}$$

by defining the following:

$$\begin{aligned}
Z &:= W_{1,\infty}^{n_x}([0,1]) \times L_{\infty}^{n_y}([0,1]) \times L_{\infty}^{n_u}([0,1]) \times L_{\infty}^{n_c}([0,1]), \\
V &:= L_{\infty}^{n_x}([0,1]) \times L_{\infty}^{n_y}([0,1]) \times \mathbb{R}^{n_\psi} \times L_{\infty}^{n_c}([0,1]), \\
K &:= \left\{\mathbf{0}_{L_{\infty}^{n_x}([0,1])}\right\} \times \left\{\mathbf{0}_{L_{\infty}^{n_y}}\right\} \times \{\mathbf{0}_{\mathbb{R}^{n_\psi}}\} \times K_c, \\
K_c &:= \{k_c \in L_{\infty}^{n_c}([0,1]) \mid k_c(t) \geq \mathbf{0}_{\mathbb{R}^{n_c}} \text{ a.e. in } [0,1]\}, \\
z &:= (x, y, u, v), \qquad (3.1.5) \\
\mathcal{J}(z) &:= \varphi(x(0), x(1)) + \frac{1}{2} \int_0^1 \|v(t)\|^2 \, dt, \\
F^{\alpha}(z) &:= \begin{pmatrix} f(x(\cdot), y(\cdot), u(\cdot)) - \dot{x}(\cdot) \\ g(x(\cdot), y(\cdot), u(\cdot)) \\ \psi(x(0), x(1)) \\ -c(x(\cdot), y(\cdot), u(\cdot)) - D^{\alpha}(\cdot) v(\cdot) \end{pmatrix}.
\end{aligned}$$

Kurcyusz [69] derived necessary conditions for Problem 3.1.11, specifically, the existence of *nontrivial Lagrange multipliers.*

Theorem 3.1.12 (Kurcyusz)
Suppose for Problem 3.1.11 the following is satisfied: Z, V are Banach spaces, $K \subseteq V$ is a closed convex cone with vertex at zero, $\mathcal{J} : Z \to \mathbb{R}$ is Fréchet differentiable, and $F : Z \to V$ is continuously Fréchet differentiable. Furthermore, there exists a local minimizer $\hat{z} \in Z$ and $im(F'(\hat{z})) = V$. Then, there exist nontrivial Lagrange multipliers $(\ell_0, \ell^\star) \in \mathbb{R} \times V^\star$ satisfying

$$\begin{aligned}
&(\ell_0, \ell^\star) \neq (0, \mathbf{0}_{V^\star}), \quad \ell_0 \geq 0, \quad \ell^\star \in K^*, \\
&\ell^\star(F(\hat{z})) = 0, \\
&\ell_0 \mathcal{J}'(\hat{z})(z) - \ell^\star(F'(\hat{z})(z)) = 0 \quad \textit{for all } z \in Z.
\end{aligned}$$

In order to apply Theorem 3.1.12 for the spaces and functions defined in (3.1.5), and the weak local minimizer $\left(\hat{x}, \hat{y}, \hat{u}, \mathbf{0}_{L_\infty^{n_c}([0,1])}\right)$, it remains to show the surjectivity of the linear operator $F^{\alpha\prime}\left(\hat{x}, \hat{y}, \hat{u}, \mathbf{0}_{L_\infty^{n_c}([0,1])}\right) : Z \to V$. To that end, we verify that the requirements of Lemma 2.4.10 are satisfied for the linear operator $F^{\alpha\prime}\left(\hat{x}, \hat{y}, \hat{u}, \mathbf{0}_{L_\infty^{n_c}([0,1])}\right)$:

Lemma 3.1.13
Let (3.1.A1) - (3.1.A4) hold for constants $\alpha, \beta > 0$. Then, $F^{\alpha\prime}\left(\hat{x}, \hat{y}, \hat{u}, \mathbf{0}_{L_\infty^{n_c}([0,1])}\right) : Z \to V$ is surjective.

Proof. For an arbitrary $z = (x, y, u, v) \in Z$ it holds

$$F^{\alpha\prime}\left(\hat{x}, \hat{y}, \hat{u}, \mathbf{0}_{L_\infty^{n_c}([0,1])}\right)(z) = \begin{pmatrix} A_f(\cdot)x(\cdot) + B_f(\cdot)y(\cdot) + C_f(\cdot)u(\cdot) - \dot{x}(\cdot) \\ A_g(\cdot)x(\cdot) + B_g(\cdot)y(\cdot) + C_g(\cdot)u(\cdot) \\ \Psi_0 x(0) + \Psi_1 x(1) \\ -A_c(\cdot)x(\cdot) - B_c(\cdot)y(\cdot) - C_c(\cdot)u(\cdot) - D^\alpha(\cdot)v(\cdot) \end{pmatrix}.$$

Then, the assumptions (3.1.A3), (3.1.A4) for $E^\alpha(\cdot) = \begin{bmatrix} B_g(\cdot) & C_g(\cdot) & \mathbf{0}_{n_y \times n_c} \\ B_c(\cdot) & C_c(\cdot) & D^\alpha(\cdot) \end{bmatrix}$ and this linearized system correspond to the uniform linear independence condition (i) and the controllability condition (ii) in Definition 2.4.1. Thus, by Lemma 2.4.10 the linear operator $F^{\alpha\prime}\left(\hat{x}, \hat{y}, \hat{u}, \mathbf{0}_{L_\infty^{n_c}([0,1])}\right)$ is surjective. □

This allows us to apply Theorem 3.1.12. Hence, assumptions (3.1.A1) - (3.1.A4) are sufficient for the existence of nontrivial Lagrange multipliers $(\ell_0, \ell^\star) \in \mathbb{R} \times V^\star$ for Problem 3.1.11 with the functions defined in (3.1.5), and the weak local minimizer $\left(\hat{x}, \hat{y}, \hat{u}, \mathbf{0}_{L_\infty^{n_c}([0,1])}\right)$. Unfortunately, these necessary conditions are impractical, since they involve the functional $\ell^\star \in K^* \subset V^\star$. Our objective is to find an explicit representation of this multiplier. To that end, we decompose the functional $\ell^\star$ into $\left(\lambda_f^\star, \lambda_g^\star, -\sigma^\top, \eta^\star\right)$ and consider the variational equation

$$\begin{aligned}
0 &= \ell_0 \mathcal{J}'\left(\hat{x}, \hat{y}, \hat{u}, \mathbf{0}_{L_\infty^{n_c}([0,1])}\right)(z) - \ell^\star\left(F^{\alpha\prime}\left(\hat{x}, \hat{y}, \hat{u}, \mathbf{0}_{L_\infty^{n_c}([0,1])}\right)(z)\right) \\
&= \ell_0\left(\varphi'_{x_0}(\hat{x}(0), \hat{x}(1))\, x(0) + \varphi'_{x_1}(\hat{x}(0), \hat{x}(1))\, x(1)\right) \qquad (3.1.6)\\
&\quad - \lambda_f^\star\left(A_f(\cdot)\, x(\cdot) + B_f(\cdot)\, y(\cdot) + C_f(\cdot)\, u(\cdot) - \dot{x}(\cdot)\right) \\
&\quad - \lambda_g^\star\left(A_g(\cdot)\, x(\cdot) + B_g(\cdot)\, y(\cdot) + C_g(\cdot)\, u(\cdot)\right) \\
&\quad + \sigma^\top\left(\Psi_0 x(0) + \Psi_1 x(1)\right) \\
&\quad + \eta^\star\left(A_c(\cdot)\, x(\cdot) + B_c(\cdot)\, y(\cdot) + C_c(\cdot)\, u(\cdot) + D^\alpha(\cdot)\, v(\cdot)\right)
\end{aligned}$$

for all $z = (x, y, u, v) \in Z$. Consequently, the following variational equations hold for every $x \in W_{1,\infty}^{n_x}([0,1]), y \in L_\infty^{n_y}([0,1]), u \in L_\infty^{n_u}([0,1])$, and $v \in L_\infty^{n_c}([0,1])$, respectively:

$$\begin{aligned}
0 &= \vartheta'_{x_0}(\hat{x}(0), \hat{x}(1), \ell_0, \sigma)\, x(0) + \vartheta'_{x_1}(\hat{x}(0), \hat{x}(1), \ell_0, \sigma)\, x(1) \\
&\quad + \lambda_f^\star\left(\dot{x}(\cdot) - A_f(\cdot)\, x(\cdot)\right) - \lambda_g^\star\left(A_g(\cdot)\, x(\cdot)\right) + \eta^\star\left(A_c(\cdot)\, x(\cdot)\right), \qquad (3.1.7)\\
0 &= -\lambda_f^\star\left(B_f(\cdot)\, y(\cdot)\right) - \lambda_g^\star\left(B_g(\cdot)\, y(\cdot)\right) + \eta^\star\left(B_c(\cdot)\, y(\cdot)\right), \qquad (3.1.8)\\
0 &= -\lambda_f^\star\left(C_f(\cdot)\, u(\cdot)\right) - \lambda_g^\star\left(C_g(\cdot)\, u(\cdot)\right) + \eta^\star\left(C_c(\cdot)\, u(\cdot)\right), \qquad (3.1.9)\\
0 &= \eta^\star\left(D^\alpha(\cdot)\, v(\cdot)\right), \qquad (3.1.10)
\end{aligned}$$

where $\vartheta(x_0, x_1, \ell_0, \sigma) := \ell_0 \varphi(x_0, x_1) + \sigma^\top \psi(x_0, x_1)$. For arbitrary

$$a_f \in L_\infty^{n_x}([0,1]), \quad a_g \in L_\infty^{n_y}([0,1]), \quad a_c \in L_\infty^{n_c}([0,1])$$

we consider the inhomogeneous, linear initial value problem

$$\begin{aligned}
\dot{x}(t) &= A_f(t)\, x(t) + B_f(t)\, y(t) + C_f(t)\, u(t) + a_f(t), && \text{a.e. in } [0,1], \qquad (3.1.11)\\
\mathbf{0}_{\mathbb{R}^{n_y}} &= A_g(t)\, x(t) + B_g(t)\, y(t) + C_g(t)\, u(t) + a_g(t), && \text{a.e. in } [0,1],\\
\mathbf{0}_{\mathbb{R}^{n_x}} &= x(0), \\
a_c(t) &= A_c(t)\, x(t) + B_c(t)\, y(t) + C_c(t)\, u(t) + D^\alpha(t)\, v(t), && \text{a.e. in } [0,1].
\end{aligned}$$

This system has the solution

$$\begin{aligned}
x(\cdot) &= \Phi_{\tilde{A}^\alpha}(\cdot) \int_0^{\cdot} \Phi_{\tilde{A}^\alpha}(\tau)^{-1}\, \tilde{a}^\alpha(\tau)\, d\tau, \\
\begin{pmatrix} y(\cdot) \\ u(\cdot) \\ v(\cdot) \end{pmatrix} &= -E^\alpha(\cdot)^\curlywedge \left[\begin{pmatrix} A_g(\cdot) \\ A_c(\cdot) \end{pmatrix} \Phi_{\tilde{A}^\alpha}(\cdot) \int_0^{\cdot} \Phi_{\tilde{A}^\alpha}(\tau)^{-1}\, \tilde{a}^\alpha(\tau)\, d\tau - \begin{pmatrix} a_g(\cdot) \\ -a_c(\cdot) \end{pmatrix}\right],
\end{aligned}$$

where

$$\begin{aligned}
\tilde{A}^\alpha(\cdot) &= A_f(\cdot) - \left(B_f(\cdot), C_f(\cdot), \mathbf{0}_{n_x \times n_c}\right) E^\alpha(\cdot)^\curlywedge \begin{pmatrix} A_g(\cdot) \\ A_c(\cdot) \end{pmatrix}, \\
\tilde{a}^\alpha(\cdot) &= a_f(\cdot) - \left(B_f(\cdot), C_f(\cdot), \mathbf{0}_{n_x \times n_c}\right) E^\alpha(\cdot)^\curlywedge \begin{pmatrix} a_g(\cdot) \\ -a_c(\cdot) \end{pmatrix}.
\end{aligned}$$

Inserting the relations (3.1.11) for (x, y, u, v) into the variational equation (3.1.6) yields

$$
\begin{aligned}
0 &= \left(\vartheta'_{x_0}\left(\hat{x}(0), \hat{x}(1), \ell_0, \sigma\right) x(0) + \vartheta'_{x_1}\left(\hat{x}(0), \hat{x}(1), \ell_0, \sigma\right) x(1)\right) \\
&\quad + \lambda_f^\star\left(a_f(\cdot)\right) + \lambda_g^\star\left(a_g(\cdot)\right) + \eta^\star\left(a_c(\cdot)\right) \\
&= \vartheta'_{x_1}\left(\hat{x}(0), \hat{x}(1), \ell_0, \sigma\right) \Phi_{\tilde{A}^\alpha}(1) \int_0^1 \Phi_{\tilde{A}^\alpha}(t)^{-1} \tilde{a}^\alpha(t)\, dt \\
&\quad + \lambda_f^\star\left(a_f(\cdot)\right) + \lambda_g^\star\left(a_g(\cdot)\right) + \eta^\star\left(a_c(\cdot)\right) \\
&= \vartheta'_{x_1}\left(\hat{x}(0), \hat{x}(1), \ell_0, \sigma\right) \Phi_{\tilde{A}^\alpha}(1) \int_0^1 \Phi_{\tilde{A}^\alpha}(t)^{-1} a_f(t)\, dt \\
&\quad - \vartheta'_{x_1}\left(\hat{x}(0), \hat{x}(1), \ell_0, \sigma\right) \Phi_{\tilde{A}^\alpha}(1) \int_0^1 \Phi_{\tilde{A}^\alpha}(t)^{-1} \left(B_f(t), C_f(t), \mathbf{0}_{n_x \times n_c}\right) E^\alpha(t)^\curlywedge \begin{pmatrix} a_g(t) \\ -a_c(t) \end{pmatrix} dt \\
&\quad + \lambda_f^\star\left(a_f(\cdot)\right) + \lambda_g^\star\left(a_g(\cdot)\right) + \eta^\star\left(a_c(\cdot)\right).
\end{aligned}
$$

Defining the multipliers

$$
\begin{aligned}
\lambda_f(\cdot)^\top &:= \vartheta'_{x_1}\left(\hat{x}(0), \hat{x}(1), \ell_0, \sigma\right) \Phi_{\tilde{A}^\alpha}(1) \Phi_{\tilde{A}^\alpha}(\cdot)^{-1}, \\
\lambda_g(\cdot)^\top &:= -\lambda_f(\cdot)^\top \left(B_f(t), C_f(\cdot), \mathbf{0}_{n_x \times n_c}\right) E^\alpha(\cdot)^\curlywedge \begin{pmatrix} \mathbf{I}_{n_y} \\ \mathbf{0}_{n_c \times n_y} \end{pmatrix}, \\
\eta(\cdot)^\top &:= -\lambda_f(\cdot)^\top \left(B_f(t), C_f(\cdot), \mathbf{0}_{n_x \times n_c}\right) E^\alpha(\cdot)^\curlywedge \begin{pmatrix} \mathbf{0}_{n_y \times n_c} \\ \mathbf{I}_{n_c} \end{pmatrix},
\end{aligned}
$$

results in

$$
\begin{aligned}
&\lambda_f^\star\left(a_f(\cdot)\right) + \lambda_g^\star\left(a_g(\cdot)\right) + \eta^\star\left(a_c(\cdot)\right) = \\
&\qquad - \int_0^1 \lambda_f(t)^\top a_f(t)\, dt - \int_0^1 \lambda_g(t)^\top a_g(t)\, dt + \int_0^1 \eta(t)^\top a_c(t)\, dt.
\end{aligned}
$$

Hence, we obtain the explicit representations

$$
\begin{aligned}
\lambda_f^\star\left(a_f(\cdot)\right) &= -\int_0^1 \lambda_f(t)^\top a_f(t)\, dt, \\
\lambda_g^\star\left(a_g(\cdot)\right) &= -\int_0^1 \lambda_g(t)^\top a_g(t)\, dt, \\
\eta^\star\left(a_c(\cdot)\right) &= \int_0^1 \eta(t)^\top a_c(t)\, dt,
\end{aligned}
$$

with $\lambda_f \in W_{1,\infty}^{n_x}([0,1])$, $\lambda_g \in L_\infty^{n_y}([0,1])$, and $\eta \in L_\infty^{n_c}([0,1])$. Let us denote the (augmented) *Hamilton function* by

$$
\begin{aligned}
&\mathcal{H} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_c} \to \mathbb{R}, \\
&\mathcal{H}\left(x, y, u, \lambda_f, \lambda_g, \eta\right) := \lambda_f^\top f(x, y, u) + \lambda_g^\top g(x, y, u) + \eta^\top c(x, y, u).
\end{aligned} \tag{3.1.12}
$$

Investigating (3.1.7) for an arbitrary $x \in W_{1,\infty}^{n_x}([0,1])$ yields

$$\begin{aligned}
0 &= \vartheta'_{x_0}(\hat{x}(0), \hat{x}(1), \ell_0, \sigma)\, x(0) + \vartheta'_{x_1}(\hat{x}(0), \hat{x}(1), \ell_0, \sigma)\, x(1) \\
&\quad + \lambda_f^\star(\dot{x}(\cdot) - A_f(\cdot)\, x(\cdot)) - \lambda_g^\star(A_g(\cdot)\, x(\cdot)) + \eta^\star(A_c(\cdot)\, x(\cdot)) \\
&= \vartheta'_{x_0}(\hat{x}(0), \hat{x}(1), \ell_0, \sigma)\, x(0) + \vartheta'_{x_1}(\hat{x}(0), \hat{x}(1), \ell_0, \sigma)\, x(1) \\
&\quad - \int_0^1 \lambda_f(t)^\top \dot{x}(t)\, dt + \int_0^1 \nabla_x \mathcal{H}[t]^\top x(t)\, dt \\
&= \vartheta'_{x_0}(\hat{x}(0), \hat{x}(1), \ell_0, \sigma)\, x(0) + \vartheta'_{x_1}(\hat{x}(0), \hat{x}(1), \ell_0, \sigma)\, x(1) \\
&\quad - \left[\lambda_f(t)^\top x(t)\right]_0^1 + \int_0^1 \left(\dot{\lambda}_f(t) + \nabla_x \mathcal{H}[t]\right)^\top x(t)\, dt \\
&= \left(\lambda_f(0)^\top + \vartheta'_{x_0}(\hat{x}(0), \hat{x}(1), \ell_0, \sigma)\right) x(0) \\
&\quad + \left(-\lambda_f(1)^\top + \vartheta'_{x_1}(\hat{x}(0), \hat{x}(1), \ell_0, \sigma)\right) x(1) \\
&\quad + \int_0^1 \left(\dot{\lambda}_f(t) + \nabla_x \mathcal{H}[t]\right)^\top x(t)\, dt,
\end{aligned} \tag{3.1.13}$$

where $\mathcal{H}[\cdot] = \mathcal{H}(\hat{x}(\cdot), \hat{y}(\cdot), \hat{u}(\cdot), \lambda_f(\cdot), \lambda_g(\cdot), \eta(\cdot))$. By the same token, using (3.1.8), (3.1.9), and (3.1.10) leads to

$$0 = \int_0^1 \nabla_y \mathcal{H}[t]^\top y(t)\, dt, \quad 0 = \int_0^1 \nabla_u \mathcal{H}[t]^\top u(t)\, dt, \quad 0 = \int_0^1 \eta(t)^\top D^\alpha(t)\, v(t)\, dt, \tag{3.1.14}$$

for all $y \in L_\infty^{n_y}([0,1])$, $u \in L_\infty^{n_u}([0,1])$, and $v \in L_\infty^{n_c}([0,1])$, respectively. Applying Lemma A.11 for $x \in W_{1,\infty}^{n_x}([0,1])$ with $x(0) = x(1) = \mathbf{0}_{\mathbb{R}^{n_x}}$ to (3.1.13) yields

$$\mathbf{0}_{\mathbb{R}^{n_x}} = \dot{\lambda}_f(t) + \nabla_x \mathcal{H}[t], \quad \text{a.e. in } [0,1].$$

Then, choosing variations of (3.1.13) with $x(0) = \mathbf{0}_{\mathbb{R}^{n_x}}, x(1) \neq \mathbf{0}_{\mathbb{R}^{n_x}}$ and vice versa results in

$$\begin{aligned}
\mathbf{0}_{\mathbb{R}^{n_x}} &= \lambda_f(0) + \nabla_{x_0} \vartheta(\hat{x}(0), \hat{x}(1), \ell_0, \sigma), \\
\mathbf{0}_{\mathbb{R}^{n_x}} &= -\lambda_f(1) + \nabla_{x_1} \vartheta(\hat{x}(0), \hat{x}(1), \ell_0, \sigma).
\end{aligned}$$

Since $W_{1,\infty}^n([0,1]) \subset L_\infty^n([0,1])$, we can apply Lemma A.11 to (3.1.14) and obtain

$$\mathbf{0}_{\mathbb{R}^{n_y}} = \nabla_y \mathcal{H}[t], \quad \mathbf{0}_{\mathbb{R}^{n_u}} = \nabla_u \mathcal{H}[t], \quad \mathbf{0}_{\mathbb{R}^{n_c}} = D^\alpha(t)\, \eta(t), \quad \text{a.e. in } [0,1].$$

Finally, we investigate the complementarity condition

$$\ell^\star\left(F^\alpha\left(\left(\hat{x}, \hat{y}, \hat{u}, \mathbf{0}_{L_\infty^{n_c}([0,1])}\right)\right)\right) = 0, \quad \ell^\star \in K^*,$$

which according to the explicit representation of the multipliers satisfies

$$\begin{aligned} 0 &= \ell^\star \left(F^\alpha \left(\left(\hat{x}, \hat{y}, \hat{u}, \mathbf{0}_{L_\infty^{n_c}([0,1])} \right) \right) \right) = -\eta^\star \left(c \left(\hat{x}\left(\cdot\right), \hat{y}\left(\cdot\right), \hat{u}\left(\cdot\right) \right) \right) \\ &= -\int_0^1 \eta\left(t\right)^\top c\left[t\right] dt, \qquad (3.1.15) \\ 0 &\le \eta^\star\left(k_c\right) = \int_0^1 \eta\left(t\right)^\top k_c\left(t\right) dt \end{aligned}$$

for all $k_c \in K_c$, i.e., $k_c \in L_\infty^{n_c}([0,1])$ with $k_c(t) \geq \mathbf{0}_{\mathbb{R}^{n_c}}$ almost everywhere in $[0,1]$. Applying Lemma A.12 yields $\eta(t) \geq \mathbf{0}_{\mathbb{R}^{n_c}}$ for almost every $t \in [0,1]$, and therefore $-\eta(t)^\top c[t] \geq \mathbf{0}_{\mathbb{R}^{n_c}}$ almost everywhere in $[0,1]$. Thus, (3.1.15) implies $\eta(t)^\top c[t] = \mathbf{0}_{\mathbb{R}^{n_c}}$ for almost every $t \in [0,1]$. We summarize the results in the following theorem:

Theorem 3.1.14

Let (3.1.A1) - (3.1.A4) hold for constants $\alpha, \beta > 0$. Then, there exist multipliers

$$\ell_0 \in \mathbb{R}, \lambda_f \in W_{1,\infty}^{n_x}([0,1]), \lambda_g \in L_\infty^{n_y}([0,1]), \sigma \in \mathbb{R}^{n_\psi}, \eta \in L_\infty^{n_c}([0,1])$$

associated with the weak local minimizer $\left(\hat{x}, \hat{y}, \hat{u}, \mathbf{0}_{L_\infty^{n_c}([0,1])}\right)$ of Problem 3.1.9 such that the following holds with the Hamilton function defined in (3.1.12)*:*

(i) $\ell_0 \geq 0$, $(\ell_0, \lambda_f, \lambda_g, \sigma, \eta) \neq \mathbf{0}$.

(ii) Adjoint DAE: Almost everywhere in $[0,1]$ *it holds*

$$\begin{aligned} \dot{\lambda}_f\left(t\right) &= -\nabla_x \mathcal{H}\left(\hat{x}\left(t\right), \hat{y}\left(t\right), \hat{u}\left(t\right), \lambda_f\left(t\right), \lambda_g\left(t\right), \eta\left(t\right)\right), \\ \mathbf{0}_{\mathbb{R}^{n_y}} &= \nabla_y \mathcal{H}\left(\hat{x}\left(t\right), \hat{y}\left(t\right), \hat{u}\left(t\right), \lambda_f\left(t\right), \lambda_g\left(t\right), \eta\left(t\right)\right). \end{aligned}$$

(iii) Transversality conditions:

$$\begin{aligned} \lambda_f\left(0\right) &= -\ell_0 \nabla_{x_0} \varphi\left(\hat{x}\left(0\right), \hat{x}\left(1\right)\right) - \psi'_{x_0}\left(\hat{x}\left(0\right), \hat{x}\left(1\right)\right)^\top \sigma, \\ \lambda_f\left(1\right) &= \ell_0 \nabla_{x_1} \varphi\left(\hat{x}\left(0\right), \hat{x}\left(1\right)\right) + \psi'_{x_1}\left(\hat{x}\left(0\right), \hat{x}\left(1\right)\right)^\top \sigma. \end{aligned}$$

(iv) Stationarity of Hamilton function: Almost everywhere in $[0,1]$ *it holds*

$$\mathbf{0}_{\mathbb{R}^{n_u}} = \nabla_u \mathcal{H}\left(\hat{x}\left(t\right), \hat{y}\left(t\right), \hat{u}\left(t\right), \lambda_f\left(t\right), \lambda_g\left(t\right), \eta\left(t\right)\right).$$

(v) Complementarity condition: Almost everywhere in $[0,1]$ *it holds*

$$\eta\left(t\right) \geq \mathbf{0}_{\mathbb{R}^{n_c}}, \quad 0 = \eta\left(t\right)^\top c\left(\hat{x}\left(t\right), \hat{y}\left(t\right), \hat{u}\left(t\right)\right), \quad \mathbf{0}_{\mathbb{R}^{n_c}} = D^\alpha\left(t\right) \eta\left(t\right).$$

Now, it remains to show that the multipliers for the modified problem in Theorem 3.1.14 associated with the weak local minimizer $\left(\hat{x}, \hat{y}, \hat{u}, \mathbf{0}_{L_\infty^{n_c}([0,1])}\right)$ coincide with Lagrange multipliers for Problem 3.1.1 associated with $(\hat{x}, \hat{y}, \hat{u})$, which yields the main result of this section:

Theorem 3.1.15 (Local Minimum Principle for Problem 3.1.1)
Let (3.1.A1) - (3.1.A4) hold for constants $\alpha, \beta > 0$. Then, there exist multipliers

$$\hat{\ell}_0 \in \mathbb{R}, \hat{\lambda}_f \in W_{1,\infty}^{n_x}([0,1]), \hat{\lambda}_g \in L_{\infty}^{n_y}([0,1]), \hat{\sigma} \in \mathbb{R}^{n_\psi}, \hat{\eta} \in L_{\infty}^{n_c}([0,1])$$

associated with the weak local minimizer $(\hat{x}, \hat{y}, \hat{u})$ of Problem 3.1.1 such that the following holds with the Hamilton function defined in (3.1.12)*:*

(i) $\hat{\ell}_0 \geq 0$, $\left(\hat{\ell}_0, \hat{\lambda}_f, \hat{\lambda}_g, \hat{\sigma}, \hat{\eta}\right) \neq \mathbf{0}$.

(ii) Adjoint DAE: Almost everywhere in $[0,1]$ *it holds*

$$\begin{aligned} \dot{\hat{\lambda}}_f(t) &= -\nabla_x \mathcal{H}\left(\hat{x}(t), \hat{y}(t), \hat{u}(t), \hat{\lambda}_f(t), \hat{\lambda}_g(t), \hat{\eta}(t)\right), \\ \mathbf{0}_{\mathbb{R}^{n_y}} &= \nabla_y \mathcal{H}\left(\hat{x}(t), \hat{y}(t), \hat{u}(t), \hat{\lambda}_f(t), \hat{\lambda}_g(t), \hat{\eta}(t)\right). \end{aligned}$$

(iii) Transversality conditions:

$$\begin{aligned} \hat{\lambda}_f(0) &= -\ell_0 \nabla_{x_0} \varphi(\hat{x}(0), \hat{x}(1)) - \psi'_{x_0}(\hat{x}(0), \hat{x}(1))^\top \hat{\sigma}, \\ \hat{\lambda}_f(1) &= \ell_0 \nabla_{x_1} \varphi(\hat{x}(0), \hat{x}(1)) + \psi'_{x_1}(\hat{x}(0), \hat{x}(1))^\top \hat{\sigma}. \end{aligned}$$

(iv) Stationarity of Hamilton function: Almost everywhere in $[0,1]$ *it holds*

$$\mathbf{0}_{\mathbb{R}^{n_u}} = \nabla_u \mathcal{H}\left(\hat{x}(t), \hat{y}(t), \hat{u}(t), \hat{\lambda}_f(t), \hat{\lambda}_g(t), \hat{\eta}(t)\right).$$

(v) Complementarity condition: Almost everywhere in $[0,1]$ *it holds*

$$\hat{\eta}(t) \geq \mathbf{0}_{\mathbb{R}^{n_c}}, \quad 0 = \hat{\eta}(t)^\top c(\hat{x}(t), \hat{y}(t), \hat{u}(t)).$$

Proof. We show that $(\ell_0, \lambda_f, \lambda_g, \sigma, \eta)$ is a Lagrange multiplier associated with the weak local minimizer $\left(\hat{x}, \hat{y}, \hat{u}, \mathbf{0}_{L_{\infty}^{n_c}([0,1])}\right)$ of Problem 3.1.9, if and only if it is a Lagrange multiplier associated with the weak local minimizer $(\hat{x}, \hat{y}, \hat{u})$ of Problem 3.1.1.
Clearly, if $(\ell_0, \lambda_f, \lambda_g, \sigma, \eta)$ satisfies $(i)-(v)$ in Theorem 3.1.14, it also satisfies $(i)-(v)$ in Theorem 3.1.15. For the opposite direction we only need to verify that, if $(\ell_0, \lambda_f, \lambda_g, \sigma, \eta)$ is a Lagrange multiplier associated with $(\hat{x}, \hat{y}, \hat{u})$, then it holds

$$\mathbf{0}_{\mathbb{R}^{n_c}} = D^\alpha(t)\,\eta(t), \quad \text{a.e. in } [0,1].$$

We recall $D^\alpha(\cdot) = \operatorname{diag}\left[c_j^\alpha(\cdot)\right]_{j \in J}$, $c_j^\alpha(\cdot) = \min\{c_j[\cdot] + \alpha, 0\}$. Then, for almost every $t \in [0,1]$ and each $j \in J^\alpha(t)$ it holds $c_j^\alpha(t) = 0$, hence $c_j^\alpha(t)\,\eta_j(t) = 0$. On the other hand, for each $j \in J \setminus J^\alpha(t)$ we have $c_j[t] < -\alpha$, and therefore, by (v), we obtain $0 = \eta_j(t) = c_j^\alpha(t)\,\eta_j(t)$ for almost every $t \in [0,1]$, which proves the assertion. □

As described in Remark 3.1.6 and Remark 3.1.8 one could substitute assumptions (3.1.A3) and (3.1.A4) in Theorem 3.1.15 for (3.1.1) in Lemma 3.1.5 and the controllability condition depicted in Remark 3.1.8, respectively.
The assumptions in Theorem 3.1.15 are actually sufficient for a stronger result, in particular, the existence of Lagrange multipliers with $\hat{\ell}_0 > 0$.

Corollary 3.1.16 (Constraint Qualification)
Let (3.1.A1) - (3.1.A4) hold for constants $\alpha, \beta > 0$. Then, the conditions in Theorem 3.1.15 (and Theorem 3.1.14, respectively) hold for $\hat{\ell}_0 > 0$.

Proof. Suppose the contrary is true, i.e., $(i) - (v)$ hold for $\hat{\ell}_0 = 0$. Then, $(ii), (iv)$, and $\mathbf{0}_{\mathbb{R}^{n_c}} = D^{\alpha}(t)\,\eta(t)$ for almost every $t \in [0,1]$ imply

$$\mathbf{0}_{\mathbb{R}^{n_y+n_u+n_c}} = \begin{bmatrix} B_f(t)^{\top} \\ C_f(t)^{\top} \\ \mathbf{0}_{n_c \times n_x} \end{bmatrix} \hat{\lambda}_f(t) + E^{\alpha}(t)^{\top} \begin{pmatrix} \hat{\lambda}_g(t) \\ \hat{\eta}(t) \end{pmatrix}, \qquad \text{a.e. in } [0,1].$$

Consequently, it holds

$$\begin{pmatrix} \hat{\lambda}_g(t) \\ \hat{\eta}(t) \end{pmatrix} = -\left(E^{\alpha}(t)^{\curlywedge}\right)^{\top} \begin{bmatrix} B_f(t)^{\top} \\ C_f(t)^{\top} \\ \mathbf{0}_{n_c \times n_x} \end{bmatrix} \hat{\lambda}_f(t), \tag{3.1.16}$$

$$\begin{aligned} \mathbf{0}_{\mathbb{R}^{n_y+n_u+n_c}} &= \left(\mathbf{I}_{n_y+n_u+n_c} - E^{\alpha}(t)^{\curlywedge} E^{\alpha}(t)\right)^{\top} \begin{bmatrix} B_f(t)^{\top} \\ C_f(t)^{\top} \\ \mathbf{0}_{n_c \times n_x} \end{bmatrix} \hat{\lambda}_f(t) \\ &= \tilde{B}^{\alpha}(t)^{\top} \hat{\lambda}_f(t) \end{aligned} \tag{3.1.17}$$

for almost every $t \in [0,1]$. Inserting $\begin{pmatrix} \hat{\lambda}_g(\cdot) \\ \hat{\eta}(\cdot) \end{pmatrix}$ into (ii) yields the differential equation

$$\dot{\hat{\lambda}}_f(t) = -\tilde{A}^{\alpha}(t)^{\top} \hat{\lambda}_f(t), \qquad \text{a.e. in } [0,1]$$

with the solution

$$\hat{\lambda}_f(\cdot) = \left(\Phi_{\tilde{A}^{\alpha}}(\cdot)^{\top}\right)^{-1} \hat{\lambda}_f(0). \tag{3.1.18}$$

Since $\hat{\ell}_0 = 0$ the transversality conditions

$$\begin{aligned} \hat{\lambda}_f(0) &= -\Psi_0^{\top}\hat{\sigma}, \\ \hat{\lambda}_f(1) &= \Psi_1^{\top}\hat{\sigma}, \end{aligned} \tag{3.1.19}$$

are satisfied. Together with (3.1.18) this result in

$$\Psi_1^{\top}\hat{\sigma} = \left(\Phi_{\tilde{A}^{\alpha}}(1)^{\top}\right)^{-1} \hat{\lambda}_f(0) = -\left(\Phi_{\tilde{A}^{\alpha}}(1)^{\top}\right)^{-1} \Psi_0^{\top}\hat{\sigma},$$

and therefore $\mathbf{0}_{\mathbb{R}^{n_\psi}} = \left(\Psi_0^{\top}\hat{\sigma} + \Phi_{\tilde{A}^{\alpha}}(1)^{\top}\Psi_1^{\top}\hat{\sigma}\right) = R^{\top}\hat{\sigma}$. Moreover, for almost every $t \in [0,1]$ it holds

$$\begin{aligned} S(t)^{\top}\hat{\sigma} &= \tilde{B}^{\alpha}(t)^{\top}\left(\Phi_{\tilde{A}^{\alpha}}(t)^{-1}\right)^{\top}\Phi_{\tilde{A}^{\alpha}}(1)^{\top}\Psi_1^{\top}\hat{\sigma} \\ &= \tilde{B}^{\alpha}(t)^{\top}\left(\Phi_{\tilde{A}^{\alpha}}(t)^{-1}\right)^{\top}\Phi_{\tilde{A}^{\alpha}}(1)^{\top}\hat{\lambda}_f(1) \\ &= \tilde{B}^{\alpha}(t)^{\top}\left(\Phi_{\tilde{A}^{\alpha}}(t)^{-1}\right)^{\top}\hat{\lambda}_f(0) \\ &= \tilde{B}^{\alpha}(t)^{\top}\hat{\lambda}_f(t) = \mathbf{0}_{\mathbb{R}^{n_y+n_u+n_c}}. \end{aligned}$$

We conclude $G\hat{\sigma} = RR^\top \hat{\sigma} + \int_0^1 S(t) S(t)^\top \hat{\sigma} dt = \mathbf{0}_{\mathbb{R}^{n_\psi}}$. According to Lemma 3.1.7, the matrix G is non-singular, thus $\hat{\sigma} = \mathbf{0}_{\mathbb{R}^{n_\psi}}$. Then, (3.1.18), (3.1.19) imply $\hat{\lambda}_f(t) = \mathbf{0}_{\mathbb{R}^{n_x}}$ for all $t \in [0,1]$, and therefore, by (3.1.16), we also have $\begin{pmatrix} \hat{\lambda}_g(t) \\ \hat{\eta}(t) \end{pmatrix} = \mathbf{0}_{\mathbb{R}^{n_y+n_c}}$ for almost every $t \in [0,1]$. This contradicts $\left(\hat{\ell}_0, \hat{\lambda}_f, \hat{\lambda}_g, \hat{\sigma}, \hat{\eta}\right) \neq \mathbf{0}$, which proves the assertion. □

3.2 Necessary Conditions for Higher Index Problems

Consider the following optimal control problem:

Problem 3.2.1 (Optimal Control Problem with Higher Index Hessenberg DAE) *Let $k \in \mathbb{N}, k \geq 2$, $n_{x_1}, \ldots, n_{x_{k-1}}, n_y, n_u, n_\psi, n_c \in \mathbb{N}$, and $n_{\boldsymbol{x}} := \sum_{i=1}^{k-1} n_{x_i}$ with $(k-1) n_y \leq n_{\boldsymbol{x}}$, $n_\psi + (k-1) n_y \leq 2n_{\boldsymbol{x}}$, $n_c \leq n_u$. Let*

$$\begin{aligned}
&\varphi : \mathbb{R}^{n_{\boldsymbol{x}}} \times \mathbb{R}^{n_{\boldsymbol{x}}} \to \mathbb{R}, && \psi : \mathbb{R}^{n_{\boldsymbol{x}}} \times \mathbb{R}^{n_{\boldsymbol{x}}} \to \mathbb{R}^{n_\psi}, \\
&f_1 : \bigtimes_{l=1}^{k-1} \mathbb{R}^{n_{x_l}} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x}, && f_i : \bigtimes_{l=i-1}^{k-1} \mathbb{R}^{n_{x_l}} \to \mathbb{R}^{x_i}, \quad i = 2, \ldots, k-1, \\
&g : \mathbb{R}^{n_{x_{k-1}}} \to \mathbb{R}^{n_y}, && c : \bigtimes_{l=1}^{k-1} \mathbb{R}^{n_{x_l}} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_c},
\end{aligned}$$

be functions.

$$\begin{array}{lrcll}
\textit{Minimize} & \multicolumn{4}{l}{\varphi(\boldsymbol{x}(0), \boldsymbol{x}(1)),} \\[2ex]
\textit{with respect to} & \multicolumn{4}{l}{\boldsymbol{x} = (x_1, \ldots, x_{k-1}) \in \bigtimes_{i=1}^{k-1} W_{i,\infty}^{n_{x_i}}([0,1]), y \in L_\infty^{n_y}([0,1]), u \in L_\infty^{n_u}([0,1]),} \\[2ex]
\textit{subject to} & \dot{x}_1(t) & = & f_1(x_1(t), \ldots, x_{k-1}(t), y(t), u(t)), & \textit{a.e. in } [0,1], \\
& \dot{x}_2(t) & = & f_2(x_1(t), \ldots, x_{k-1}(t)), & \textit{in } [0,1], \\
& \dot{x}_3(t) & = & f_3(x_2(t), \ldots, x_{k-1}(t)), & \textit{in } [0,1], \\
& & \vdots & & \\
& \dot{x}_{k-1}(t) & = & f_{k-1}(x_{k-2}(t), x_{k-1}(t)), & \textit{in } [0,1], \\
& \mathbf{0}_{\mathbb{R}^{n_y}} & = & g(x_{k-1}(t)), & \textit{in } [0,1], \\
& \mathbf{0}_{\mathbb{R}^{n_\psi}} & = & \psi(\boldsymbol{x}(0), \boldsymbol{x}(1)), & \\
& \mathbf{0}_{\mathbb{R}^{n_c}} & \geq & c(x_1(t), \ldots, x_{k-1}(t), y(t), u(t)), & \textit{a.e. in } [0,1].
\end{array}$$

Note that the smoothness assumption $x_i \in W_{i,\infty}^{n_{x_i}}([0,1])$ for $i = 2, \ldots, k-1$ is not restrictive, since the smoothness follows automatically from the structure of the Hessenberg DAE and $x_1 \in W_{1,\infty}^{n_{x_1}}([0,1])$. Usually, the semi-explicit DAE in Problem 3.2.1 is called Hessenberg DAE of order k, if the $(k-1)$-st derivative of the algebraic equation $\mathbf{0}_{\mathbb{R}^{n_y}} = g(x_{k-1}(t))$ with respect to t is implicitly solvable for the algebraic variable y. To that end, the functions in Problem 3.2.1 need to be sufficiently smooth, therefore we assume the following:

Assumption 3.2.2

(3.2.A1) *(Existence of a Minimizer)*

Let $(\hat{\boldsymbol{x}}, \hat{y}, \hat{u}) \in \bigtimes_{i=1}^{k-1} W_{i,\infty}^{n_{x_i}}([0,1]) \times L_\infty^{n_y}([0,1]) \times L_\infty^{n_u}([0,1])$ *be a weak local minimizer of Problem 3.2.1.*

(3.2.A2) *(Smoothness of the System)*

(a) φ *and* ψ *are continuously differentiable with respect to all arguments.*

(b) For a sufficiently large convex compact neighborhood $\mathcal{M}_1$ *of*

$$\left\{ (\hat{x}_1(t), \ldots, \hat{x}_{k-1}(t), \hat{y}(t), \hat{u}(t)) \in \bigtimes_{i=1}^{k-1} \mathbb{R}^{n_{x_i}} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \mid t \in [0,1] \right\},$$

let the mappings

$$(x_1, \ldots, x_{k-1}, y, u) \mapsto f_1(x_1, \ldots, x_{k-1}, y, u),$$
$$(x_1, \ldots, x_{k-1}, y, u) \mapsto c(x_1, \ldots, x_{k-1}, y, u),$$

be continuously differentiable, and the derivatives

$$f'_{1,(x_1,\ldots,x_{k-1},y,u)}, \quad c'_{(x_1,\ldots,x_{k-1},y,u)},$$

be bounded in $\mathcal{M}_1$*. Furthermore, for* $i = 2, \ldots, k-1$ *and sufficiently large convex compact neighborhoods* $\mathcal{M}_i$ *of*

$$\left\{ (\hat{x}_{i-1}(t), \ldots, \hat{x}_{k-1}(t)) \in \bigtimes_{l=i-1}^{k-1} \mathbb{R}^{n_{x_l}} \mid t \in [0,1] \right\},$$

let the mappings

$$(x_{i-1}, \ldots, x_{k-1}) \mapsto f_i(x_{i-1}, \ldots, x_{k-1}),$$

be i*-times continuously differentiable, and the derivatives*

$$\frac{\partial f_i(x_{i-1}, \ldots, x_{k-1})}{\partial (x_{i-1}, \ldots, x_{k-1})}, \frac{\partial^2 f_i(x_{i-1}, \ldots, x_{k-1})}{(\partial (x_{i-1}, \ldots, x_{k-1}))^2}, \ldots, \frac{\partial^i f_i(x_{i-1}, \ldots, x_{k-1})}{(\partial (x_{i-1}, \ldots, x_{k-1}))^i},$$

be bounded in $\mathcal{M}_i$*. For a sufficiently large convex compact neighborhood* $\mathcal{M}_k$ *of*

$$\{\hat{x}_{k-1}(t) \in \mathbb{R}^{n_{k-1}} \mid t \in [0,1]\},$$

let the mapping

$$x_{k-1} \mapsto g(x_{k-1}),$$

be k*-times continuously differentiable, and the derivatives*

$$\frac{\partial g_i(x_{k-1})}{\partial x_{k-1}}, \frac{\partial^2 g_i(x_{k-1})}{(\partial x_{k-1})^2}, \ldots, \frac{\partial^k g_i(x_{k-1})}{(\partial x_{k-1})^k},$$

be bounded in $\mathcal{M}_k$*.*

Something that occurs in higher index DAEs are the so called *hidden constraints*, i.e., not only does the algebraic constraint $\mathbf{0}_{\mathbb{R}^{n_y}} = g(x_{k-1}(t))$ for $t \in [0,1]$ hold, but also derivatives of $g(x_{k-1}(t))$ with respect to t up to the $(k-1)$-st derivative vanish. For $t \in [0,1]$ we denote

$$\begin{aligned}
g_{k-1}(x_{k-1}(t)) &:= g(x_{k-1}(t)), \\
g_{k-2}(x_{k-2}(t), x_{k-1}(t)) &:= \frac{d}{dt} g(x_{k-1}(t)) = g'(x_{k-1}(t)) f_{k-1}(x_{k-2}(t), x_{k-1}(t)), \\
g_{k-3}(x_{k-3}(t), x_{k-2}(t), x_{k-1}(t)) &:= \frac{d}{dt} g_{k-2}(x_{k-2}(t), x_{k-1}(t)) \\
&= g'_{k-2,x_{k-2}}(x_{k-2}(t), x_{k-1}(t)) \\
&\qquad f_{k-2}(x_{k-3}(t), x_{k-2}(t), x_{k-1}(t)) \\
&\quad + g'_{k-2,x_{k-1}}(x_{k-2}(t), x_{k-1}(t)) f_{k-1}(x_{k-2}(t), x_{k-1}(t)), \\
&\vdots \\
g_1(x_1(t), x_2(t), \ldots, x_{k-1}(t)) &:= \frac{d}{dt} g_2(x_2(t), x_3(t), \ldots, x_{k-1}(t)) \\
&= \sum_{i=2}^{k-1} g'_{2,x_i}(x_2(t), x_3(t), \ldots, x_{k-1}(t)) \\
&\qquad f_i(x_{i-1}(t), x_i(t), \ldots, x_{k-1}(t)),
\end{aligned}$$

and for almost every $t \in [0,1]$ we define

$$\begin{aligned}
g_0(x_1(t), \ldots, x_{k-1}(t), y(t), u(t)) &:= \frac{d}{dt} g_1(x_1(t), x_2(t), \ldots, x_{k-1}(t)) \\
&= \sum_{i=2}^{k-1} g'_{1,x_i}(x_1(t), x_2(t), \ldots, x_{k-1}(t)) \\
&\qquad f_i(x_{i-1}(t), x_i(t), \ldots, x_{k-1}(t)) \\
&\quad + g'_{1,x_1}(x_1(t), x_2(t), \ldots, x_{k-1}(t)) \\
&\qquad f_1(x_1(t), \ldots, x_{k-1}(t), y(t), u(t)).
\end{aligned}$$

Then, the Hessenberg DAE has index k, if the matrix function

$$g'_{0,y}(\cdot) = g_{k-1,x_{k-1}}(\cdot) f_{k-1,x_{k-2}}(\cdot) f_{k-2,x_{k-3}}(\cdot) \cdots f_{2,x_1}(\cdot) f_{1,y}(\cdot)$$

is non-singular and the inverse is essentially bounded along a trajectory. Analog to the index one case in Section 3.1, we will later combine the index condition of the algebraic equation with the regularity of the mixed control-state constraints. Instead of considering the Hessenberg DAE of order k, it is possible to reduce the index of the DAE to one. To that end, we denote

$$\boldsymbol{x}_{[i]} := \begin{pmatrix} x_i \\ x_{i+1} \\ \vdots \\ x_{k-1} \end{pmatrix}, \quad i = 1, \ldots, k-1,$$

$$\boldsymbol{f}(\boldsymbol{x}, y, u) := \begin{pmatrix} f_1\left(\boldsymbol{x}_{[1]}, y, u\right) \\ f_2\left(\boldsymbol{x}_{[1]}\right) \\ f_3\left(\boldsymbol{x}_{[2]}\right) \\ \vdots \\ f_{k-1}\left(\boldsymbol{x}_{[k-2]}\right) \end{pmatrix}, \quad \boldsymbol{g}(\boldsymbol{x}) := \begin{pmatrix} g_1\left(\boldsymbol{x}_{[1]}\right) \\ g_2\left(\boldsymbol{x}_{[2]}\right) \\ g_3\left(\boldsymbol{x}_{[3]}\right) \\ \vdots \\ g_{k-1}\left(\boldsymbol{x}_{[k-1]}\right) \end{pmatrix},$$

which yields the index reduced system defined by

$$\begin{aligned} \dot{\boldsymbol{x}}(t) &= \boldsymbol{f}(\boldsymbol{x}(t), y(t), u(t)), && \text{a.e. in } [0,1], \\ \mathbf{0}_{\mathbb{R}^{n_y}} &= g_0(\boldsymbol{x}(t), y(t), u(t)), && \text{a.e. in } [0,1], \\ \mathbf{0}_{\mathbb{R}^{(k-1)n_y}} &= \boldsymbol{g}(\boldsymbol{x}(0)), \\ \mathbf{0}_{\mathbb{R}^{n_\psi}} &= \psi(\boldsymbol{x}(0), \boldsymbol{x}(1)), \\ \mathbf{0}_{\mathbb{R}^{n_c}} &\geq c(\boldsymbol{x}(t), y(t), u(t)), && \text{a.e. in } [0,1]. \end{aligned} \tag{3.2.1}$$

In the following, we will prove that (3.2.1) is indeed equivalent to the system in Problem 3.2.1:

Lemma 3.2.3
$(\boldsymbol{x}, y, u) \in \bigtimes_{i=1}^{k-1} W_{i,\infty}^{n_{x_i}}([0,1]) \times L_\infty^{n_y}([0,1]) \times L_\infty^{n_u}([0,1])$ *is a solution of the system in Problem 3.2.1, if and only if it is a solution of the reduced system* (3.2.1).

Proof. It remains to show that, if $(\boldsymbol{x}, y, u)$ satisfies

$$\begin{aligned} \mathbf{0}_{\mathbb{R}^{n_y}} &= g_0(\boldsymbol{x}(t), y(t), u(t)), && \text{a.e. in } [0,1], \\ \mathbf{0}_{\mathbb{R}^{(k-1)n_y}} &= \boldsymbol{g}(\boldsymbol{x}(0)), \end{aligned}$$

then $\mathbf{0}_{\mathbb{R}^{n_y}} = g(x_{k-1}(t))$ holds in $[0,1]$. Defining the function $\tilde{g}(\cdot) := g(x_{k-1}(\cdot)) \in W_{k-1,\infty}^{n_y}([0,1])$ and exploiting the Taylor expansion at $t = 0$ yields

$$\begin{aligned} \tilde{g}(t) = \tilde{g}(0) + \tilde{g}'(0)\,t + \frac{1}{2}\tilde{g}''(0)\,t^2 + \ldots + \frac{1}{(k-2)!}\tilde{g}^{(k-2)}(0)\,t^{k-2} \\ + \int_0^t \frac{(t-\tau)^{k-2}}{(k-2)!}\tilde{g}^{(k-1)}(\tau)\,d\tau \end{aligned}$$

for all $t \in [0,1]$. Moreover, since $\tilde{g}^{(i)}(\cdot) = g_{k-1-i}\left(\boldsymbol{x}_{[k-1-i]}\right)$ holds for $i = 0, 1, \ldots, k-2$ and $\tilde{g}^{(k-1)}(\cdot) = g_0(\boldsymbol{x}(\cdot), y(\cdot), u(\cdot))$, we obtain

$$\begin{aligned}
\mathbf{0}_{\mathbb{R}^{n_y}} &= g_0(\boldsymbol{x}(t), y(t), u(t)) = \tilde{g}^{(k-1)}(t), \qquad \text{a.e. in } [0,1], \\
\mathbf{0}_{\mathbb{R}^{(k-1)n_y}} &= \boldsymbol{g}(\boldsymbol{x}(0)) = \begin{pmatrix} \tilde{g}^{(k-2)}(0) \\ \tilde{g}^{(k-3)}(0) \\ \vdots \\ \tilde{g}'(0) \\ \tilde{g}(0) \end{pmatrix},
\end{aligned}$$

and therefore $\mathbf{0}_{\mathbb{R}^{n_y}} = \tilde{g}(t) = g(x_{k-1}(t))$ in $[0,1]$, which proves the assertion. □

This allows us to consider Problem 3.2.1 subject to the index reduced system (3.2.1). Thus, we can deduce a local minimum principle for Problem 3.2.1 by applying the results in Theorem 3.1.15. Therefore, the same way as in Section 3.1, we abbreviate the derivatives at the minimizer by

$$\begin{array}{lll}
A_{\boldsymbol{f}}(\cdot) := \boldsymbol{f}'_{\boldsymbol{x}}[\cdot], & B_{\boldsymbol{f}}(\cdot) := \boldsymbol{f}'_{y}[\cdot], & C_{\boldsymbol{f}}(\cdot) := \boldsymbol{f}'_{u}[\cdot], \\
A_{g_0}(\cdot) := g'_{0,\boldsymbol{x}}[\cdot], & B_{g_0}(\cdot) := g'_{0,y}[\cdot], & C_{g_0}(\cdot) := g'_{0,u}[\cdot], \\
\Psi_0^{\boldsymbol{g}} := \left[\begin{array}{c} \boldsymbol{g}'[0] \\ \psi'_{\boldsymbol{x}_0}(\hat{\boldsymbol{x}}(0), \hat{\boldsymbol{x}}(1)) \end{array}\right], & \Psi_1^{\boldsymbol{g}} := \left[\begin{array}{c} \mathbf{0}_{(k-1)n_y \times n_{\boldsymbol{x}}} \\ \psi'_{\boldsymbol{x}_1}(\hat{\boldsymbol{x}}(0), \hat{\boldsymbol{x}}(1)) \end{array}\right], & \\
A_c(\cdot) := c'_{\boldsymbol{x}}[\cdot], & B_c(\cdot) := c'_{y}[\cdot], & C_c(\cdot) := c'_{u}[\cdot],
\end{array}$$

and define

$$\begin{aligned}
c_j^{\alpha}(\cdot) &:= \min\{c_j[\cdot] + \alpha, 0\}, \quad j \in J, \\
D^{\alpha}(\cdot) &:= \operatorname{diag}\left[c_j^{\alpha}(\cdot)\right]_{j \in J}, \\
E_0^{\alpha}(\cdot) &:= \left[\begin{array}{ccc} B_{g_0}(\cdot) & C_{g_0}(\cdot) & \mathbf{0}_{n_y \times n_c} \\ B_c(\cdot) & C_c(\cdot) & D^{\alpha}(\cdot) \end{array}\right]
\end{aligned}$$

for a constant $\alpha \geq 0$. Note that in application it might be necessary to eliminate redundant boundary conditions in (3.2.1). Otherwise, the associated Gramian matrix is singular, and therefore, by Lemma 2.4.3, the following controllability assumption will not hold:

Assumption 3.2.4 (Linear Independence, Controllability)

(3.2.A3) *(Index k / Regularity Condition)*

There exist constants $\alpha > 0$ and $\beta > 0$ such that for all $\varpi \in \mathbb{R}^{n_y+n_c}$ it holds

$$\left\| E_0^{\alpha}(t)^{\top} \varpi \right\| \geq \beta \|\varpi\|, \quad \textit{a.e. in } [0,1].$$

(3.2.A4) *(Controllability)*

For any $e \in \mathbb{R}^{(k-1)n_y + n_\psi}$ *there exists*

$$(\boldsymbol{x}, y, u, v) \in \bigtimes_{i=1}^{k-1} W_{i,\infty}^{n_{x_i}}([0,1]) \times L_\infty^{n_y}([0,1]) \times L_\infty^{n_u}([0,1]) \times L_\infty^{n_c}([0,1])$$

such that the DAE

$$\begin{aligned}
\dot{\boldsymbol{x}}(t) &= A_{\boldsymbol{f}}(t)\,\boldsymbol{x}(t) + B_{\boldsymbol{f}}(t)\,y(t) + C_{\boldsymbol{f}}(t)\,u(t), && a.e.\ in\ [0,1],\\
\mathbf{0}_{\mathbb{R}^{n_y}} &= A_{g_0}(t)\,\boldsymbol{x}(t) + B_{g_0}(t)\,y(t) + C_{g_0}(t)\,u(t), && a.e.\ in\ [0,1],\\
e &= \Psi_0^{\boldsymbol{g}}\boldsymbol{x}(0) + \Psi_1^{\boldsymbol{g}}\boldsymbol{x}(1),\\
\mathbf{0}_{\mathbb{R}^{n_c}} &= A_c(t)\,\boldsymbol{x}(t) + B_c(t)\,y(t) + C_c(t)\,u(t) + D^\alpha(t)\,v(t), && a.e.\ in\ [0,1],
\end{aligned}$$

is satisfied.

For the same reasons described in Remark 3.1.4 we combined the index property of the DAE with the regularity of the inequality constraint in (3.2.A3). With the (augmented) Hamilton function defined by

$$\begin{aligned}
&\mathcal{H} : \bigtimes_{l=1}^{k-1} \mathbb{R}^{n_{x_l}} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \times \bigtimes_{l=1}^{k-1} \mathbb{R}^{n_{x_l}} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_c} \to \mathbb{R},\\
&\mathcal{H}(\boldsymbol{x}, y, u, \boldsymbol{\lambda_f}, \lambda_{g_0}, \eta) := \boldsymbol{\lambda_f}^\top \boldsymbol{f}(\boldsymbol{x}, y, u) + \lambda_{g_0}^\top g_0(\boldsymbol{x}, y, u) + \eta^\top c(\boldsymbol{x}, y, u),
\end{aligned}$$

where $\boldsymbol{\lambda_f} = \left(\lambda_{f_1}, \ldots, \lambda_{f_{k-1}}\right)$, we apply Theorem 3.1.15 and Lemma 3.1.16 to Problem 3.2.1 subject to the reduced system (3.2.1), which yields the main result of this section:

Theorem 3.2.5 (Local Minimum Principle for Problem 3.2.1)
Let (3.2.A1) - (3.2.A4) hold for constants $\alpha, \beta > 0$. *Then, there exist multipliers*

$$\begin{aligned}
&\hat{\ell}_0 \in \mathbb{R}, \hat{\boldsymbol{\lambda}}_{\boldsymbol{f}} = \left(\hat{\lambda}_{f_1}, \ldots, \hat{\lambda}_{f_{k-1}}\right) \in \bigtimes_{l=1}^{k-1} W_{1,\infty}^{n_{x_l}}([0,1]), \hat{\lambda}_{g_0} \in L_\infty^{n_y}([0,1]),\\
&\hat{\sigma}_\psi \in \mathbb{R}^{n_\psi}, \hat{\boldsymbol{\sigma}}_{\boldsymbol{g}} \in \mathbb{R}^{(k-1)n_y}, \hat{\eta} \in L_\infty^{n_c}([0,1])
\end{aligned}$$

associated with the weak local minimizer $(\hat{\boldsymbol{x}}, \hat{y}, \hat{u})$ *of Problem 3.2.1 such that:*

(i) $\hat{\ell}_0 = 1$.

(ii) Adjoint DAE: Almost everywhere in $[0,1]$ *it holds*

$$\begin{aligned}
\dot{\hat{\lambda}}_{f_1}(t) &= -\nabla_{x_1}\mathcal{H}\left(\hat{\boldsymbol{x}}(t), \hat{y}(t), \hat{u}(t), \hat{\boldsymbol{\lambda}}_{\boldsymbol{f}}(t), \hat{\lambda}_{g_0}(t), \hat{\eta}(t)\right),\\
\dot{\hat{\lambda}}_{f_2}(t) &= -\nabla_{x_2}\mathcal{H}\left(\hat{\boldsymbol{x}}(t), \hat{y}(t), \hat{u}(t), \hat{\boldsymbol{\lambda}}_{\boldsymbol{f}}(t), \hat{\lambda}_{g_0}(t), \hat{\eta}(t)\right),\\
&\ \vdots\\
\dot{\hat{\lambda}}_{f_{k-1}}(t) &= -\nabla_{x_{k-1}}\mathcal{H}\left(\hat{\boldsymbol{x}}(t), \hat{y}(t), \hat{u}(t), \hat{\boldsymbol{\lambda}}_{\boldsymbol{f}}(t), \hat{\lambda}_{g_0}(t), \hat{\eta}(t)\right),\\
\mathbf{0}_{\mathbb{R}^{n_y}} &= \nabla_y\mathcal{H}\left(\hat{\boldsymbol{x}}(t), \hat{y}(t), \hat{u}(t), \hat{\boldsymbol{\lambda}}_{\boldsymbol{f}}(t), \hat{\lambda}_{g_0}(t), \hat{\eta}(t)\right).
\end{aligned}$$

(iii) *Transversality conditions:*

$$\begin{aligned}\hat{\boldsymbol{\lambda}}_{\boldsymbol{f}}(0) &= -\nabla_{\boldsymbol{x}_0}\varphi\left(\hat{\boldsymbol{x}}(0),\hat{\boldsymbol{x}}(1)\right) - \psi'_{\boldsymbol{x}_0}\left(\hat{\boldsymbol{x}}(0),\hat{\boldsymbol{x}}(1)\right)^\top \hat{\sigma}_\psi - \boldsymbol{g}'\left(\hat{\boldsymbol{x}}(0)\right)^\top \hat{\sigma}_{\boldsymbol{g}},\\ \hat{\boldsymbol{\lambda}}_{\boldsymbol{f}}(1) &= \nabla_{\boldsymbol{x}_1}\varphi\left(\hat{\boldsymbol{x}}(0),\hat{\boldsymbol{x}}(1)\right) + \psi'_{\boldsymbol{x}_1}\left(\hat{\boldsymbol{x}}(0),\hat{\boldsymbol{x}}(1)\right)^\top \hat{\sigma}_\psi.\end{aligned}$$

(iv) *Stationarity of Hamilton function: Almost everywhere in* $[0,1]$ *it holds*

$$\mathbf{0}_{\mathbb{R}^{n_u}} = \nabla_u \mathcal{H}\left(\hat{\boldsymbol{x}}(t),\hat{y}(t),\hat{u}(t),\hat{\boldsymbol{\lambda}}_{\boldsymbol{f}}(t),\hat{\lambda}_{g_0}(t),\hat{\eta}(t)\right).$$

(v) *Complementarity condition: Almost everywhere in* $[0,1]$ *it holds*

$$\hat{\eta}(t) \geq \mathbf{0}_{\mathbb{R}^{n_c}}, \quad 0 = \hat{\eta}(t)^\top c\left(\hat{\boldsymbol{x}}(t),\hat{y}(t),\hat{u}(t)\right).$$

Note that $\hat{\boldsymbol{\lambda}}_{\boldsymbol{f}}$ is only Lipschitz continuous and not as smooth as the differential state $\hat{\boldsymbol{x}}$. Moreover, the assumptions (3.2.A3) and (3.2.A4) in Theorem 3.2.5 could be substituted for conditions analog to the ones described in Lemma 3.1.5 / Remark 3.1.6 and Remark 3.1.8, respectively.

3.3 Example

In order to illustrate that a potential solution of an optimal control problem can be obtained from the local minimum principles in Theorem 3.1.15 / Theorem 3.2.5, we consider the following variation of the minimum energy problem:

Example 3.3.1

$$\begin{array}{llcll} \textit{Minimize} & x_4(1), & & & \\ \textit{subject to} & \dot{x}_1(t) &=& u(t)-y(t), & 0 = x_1(0), \quad 0 = x_1(1),\\ & \dot{x}_2(t) &=& u(t), & 0 = x_2(0)-1, \quad 0 = x_2(1)+1,\\ & \dot{x}_3(t) &=& -x_2(t), & \\ & \dot{x}_4(t) &=& \frac{1}{2}u(t)^2, & 0 = x_4(0),\\ & 0 &=& x_1(t)+x_3(t). & \end{array}$$

Herein, the DAE has index two, since we can (explicitly) solve the hidden constraint

$$0 = \dot{x}_1(t) + \dot{x}_3(t) = u(t) - y(t) - x_2(t)$$

for the algebraic variable y. Replacing the algebraic constraint with the hidden constraint and adding the initial condition $0 = x_1(0) + x_3(0)$ yields the reduced system

$$\begin{array}{llcll} \text{Minimize} & x_4(1), & & & \\ \text{subject to} & \dot{x}_1(t) &=& u(t)-y(t), & 0 = x_1(0), \quad 0 = x_1(1),\\ & \dot{x}_2(t) &=& u(t), & 0 = x_2(0)-1, \quad 0 = x_2(1)+1,\\ & \dot{x}_3(t) &=& -x_2(t), & \\ & \dot{x}_4(t) &=& \frac{1}{2}u(t)^2, & 0 = x_4(0),\\ & 0 &=& u(t)-y(t)-x_2(t), & \\ & 0 &=& x_1(0)+x_3(0). & \end{array}$$

Then, by Theorem 3.2.5, the necessary conditions (for $\ell_0 = 1$) can be expressed as

$$\begin{array}{rcll}
\dot{\lambda}_{f_1}(t) & = & 0, & \lambda_{f_1}(0) = -\sigma_1, \\
\dot{\lambda}_{f_2}(t) & = & \lambda_{f_3}(t) + \lambda_g(t), & \lambda_{f_2}(0) = -\sigma_2, \\
\dot{\lambda}_{f_3}(t) & = & 0, & \lambda_{f_3}(1) = 0, \\
\dot{\lambda}_{f_4}(t) & = & 0, & \lambda_{f_4}(1) = 1, \\
0 & = & -\lambda_{f_1}(t) - \lambda_g(t), & \\
0 & = & \lambda_{f_1}(t) + \lambda_{f_2}(t) + \lambda_{f_4}(t)\, u(t) + \lambda_g(t), &
\end{array}$$

where we neglected some of the redundant transversality conditions. These KKT-conditions have the solution

$$\begin{array}{rclrcl}
x_1(t) & = & -t^2 + t, & \lambda_{f_1}(t) & = & 0, \\
x_2(t) & = & -2t + 1, & \lambda_{f_2}(t) & = & 2, \\
x_3(t) & = & t^2 - t, & \lambda_{f_3}(t) & = & 0, \\
x_4(t) & = & 2t, & \lambda_{f_4}(t) & = & 1, \\
y(t) & = & 2t - 3, & \lambda_g(t) & = & 0, \\
u(t) & = & -2. & & &
\end{array}$$

Furthermore, we check, if the conditions (3.2.A3), (3.2.A4) hold. To that end, we compute the (constant) matrices

$$A_f = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad B_f = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad C_f = \begin{pmatrix} 1 \\ 1 \\ 0 \\ -2 \end{pmatrix},$$

$$A_{g_0} = (0, -1, 0, 0), \quad B_{g_0} = -1, \quad C_{g_0} = 1,$$

$$\Psi_0^{\boldsymbol{g}} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \Psi_1^{\boldsymbol{g}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

(3.2.A3) is satisfied, since

$$\left\| \left[B_{g_0}, C_{g_0}\right]^\top \varpi \right\| = \left\| \begin{pmatrix} -1 \\ 1 \end{pmatrix} \varpi \right\| = \sqrt{2}\,|\varpi| \quad \text{for all} \quad \varpi \in \mathbb{R}.$$

The linear system in (3.2.A4) has the associated Gramian matrix

$$G = \begin{pmatrix} 2 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 2\exp\left(\frac{1}{2}\right) - 2 & \exp\left(\frac{1}{2}\right) \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 2\exp\left(\frac{1}{2}\right) - 2 & 0 & 6\exp(1) - 16\exp\left(\frac{1}{2}\right) + 13 & 3\exp(1) - 4\exp\left(\frac{1}{2}\right) + 1 \\ 0 & 0 & \exp\left(\frac{1}{2}\right) & 0 & 3\exp(1) - 4\exp\left(\frac{1}{2}\right) + 1 & \frac{3\exp(1)-1}{2} \end{pmatrix},$$

which is non-singular, and therefore, by Lemma 2.4.3, condition (3.2.A4) holds.

In this chapter, we derived necessary conditions for optimal control problems subject to Hessenberg DAEs of arbitrary order and mixed control-state constraints. The results of [83] were extended by including boundary conditions and weakening the assumptions, which also hold for, e.g., box-constraints. We first derived a local minimum principle for the index one case by modifying the problem, and then showed that higher index DAEs can be reduced to index one, thus obtaining a local minimum principle for Hessenberg DAEs of arbitrary order.

Chapter 4

Sufficient Conditions

In contrast to nonlinear problems subject to DAEs, sufficient conditions for problems with explicit ODEs have been well studied. Problems with control constraints are investigated in [21, 31, 98, 123]. In [92, 93], mixed control-state constraints are considered. Optimal control problems subject to mixed control-state constraints and pure state constraints are analyzed in [13, 79, 82, 99, 100], where [13] considers multiple pure state constraints of arbitrary order. Problems with free final time have been examined in [21, 58, 93]. Sufficient conditions for strong local minimizer were discussed in [14, 15]. Second-order sufficient conditions usually include a *continuous control* assumption, *linear independence of active constraints*, and some type of *Legendre-Clebsch condition*. Herein, the Legendre-Clebsch condition might include the active part of the constraints (compare assumption (4.1.A6)), or a strengthened Legendre-Clebsch condition is used as in [13, 91]. If boundary conditions are present, then it is generally assumed that a rank or *controllability condition* holds. Furthermore, junction point and *complementarity conditions* can occur.

One approach to derive sufficient conditions is to view the optimal control problem as an *infinite optimization problem* as in Section 3.1, and apply corresponding results. In the finite dimensional case (compare Problem 2.3.1), a second-order sufficient condition is the positive definiteness of the Hessian of the Lagrange function at a KKT-point on a certain cone as in Theorem 2.3.5. The standard proof techniques rely in a decisive way on the compactness of the unit sphere in $\mathbb{R}^n$. Unfortunately, this property does not carry over to infinite dimensional spaces. Thus, one requires stronger conditions for infinite optimization problems, for instance, (uniform) coercivity of the Hessian of the Lagrange function as in Definition 2.1.8 (cf. [88]). However, something that occurs in sufficient conditions for optimal control problems is the so called *two-norm discrepancy*. In particular, the optimal control problem viewed as an infinite optimization problem is well defined and differentiable in the L_∞-norm, but the (uniform) coercivity condition only holds in a weaker L_2-norm, in which the Lagrange function is not differentiable. In [91], the two-norm discrepancy is overcome by introducing approximation conditions for the functions, which are satisfied in the weaker L_2-norm.

A different approach to derive second-order sufficient conditions is to use a *Hamilton-Jacobi inequality* (cf. [82, 92, 98–100, 123]). Herein, the task is to construct a quadratic function corresponding to the optimal control problem, which satisfies the mentioned inequality. One key assumption in deriving such a function is the existence of a (bounded) solution of a *Riccati equation*.

Using the techniques in [92] we aim to construct a quadratic function, which satisfies a Hamilton-Jacobi inequality. With this approach, we aim to derive second-order sufficient optimality conditions for Problem 3.1.1 and Problem 3.2.1. Thus, we extend the results in [83] by including boundary conditions, and we generalize the results in [92] by considering problems with DAEs.

4.1 Sufficient Conditions for Index One Problems

Our goal in this section is to construct a quadratic function for Problem 3.1.1

$$\begin{array}{lrcll}
\text{Minimize} & \multicolumn{3}{l}{\varphi\left(x\left(0\right), x\left(1\right)\right),} & \\[2ex]
\text{with respect to} & \multicolumn{3}{l}{x \in W_{1,\infty}^{n_x}\left(\left[0,1\right]\right), y \in L_{\infty}^{n_y}\left(\left[0,1\right]\right), u \in L_{\infty}^{n_u}\left(\left[0,1\right]\right),} & \\[2ex]
\text{subject to} & \dot{x}\left(t\right) & = & f\left(x\left(t\right), y\left(t\right), u\left(t\right)\right), & \text{a.e. in } \left[0,1\right], \\
 & \mathbf{0}_{\mathbb{R}^{n_y}} & = & g\left(x\left(t\right), y\left(t\right), u\left(t\right)\right), & \text{a.e. in } \left[0,1\right], \\
 & \mathbf{0}_{\mathbb{R}^{n_\psi}} & = & \psi\left(x\left(0\right), x\left(1\right)\right), & \\
 & \mathbf{0}_{\mathbb{R}^{n_c}} & \geq & c\left(x\left(t\right), y\left(t\right), u\left(t\right)\right), & \text{a.e. in } \left[0,1\right],
\end{array}$$

which satisfies a Hamilton-Jacobi inequality. To that end, we define the following sets for an element $(x, y, u) \in W_{1,\infty}^{n_x}\left(\left[0,1\right]\right) \times L_{\infty}^{n_y}\left(\left[0,1\right]\right) \times L_{\infty}^{n_u}\left(\left[0,1\right]\right)$ and $\rho > 0$:

$$\begin{aligned}
\mathcal{B}_{\rho}^{\infty}\left(x, y, u\right) &:= \left\{(z, w, v) \in L_{\infty}^{n_x}\left(\left[0,1\right]\right) \times L_{\infty}^{n_y}\left(\left[0,1\right]\right) \times L_{\infty}^{n_u}\left(\left[0,1\right]\right)\right. \\
&\qquad \left. \mid \left\|(z, w, v) - (x, y, u)\right\|_{\infty} \leq \rho\right\}, \\
\Sigma_{\rho}\left(x\left(t\right), y\left(t\right), u\left(t\right)\right) &:= \mathcal{B}_{\rho}\left(x\left(t\right), y\left(t\right), u\left(t\right)\right) \\
&\qquad \cap \left\{(z, w, v) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \,\middle|\, \begin{array}{l} g\left(z, w, v\right) = \mathbf{0}_{\mathbb{R}^{n_y}} \\ c\left(z, w, v\right) \leq \mathbf{0}_{\mathbb{R}^{n_c}} \end{array}\right\}, \\
\operatorname{graph}\left(\Sigma_{\rho}\left(x, y, u\right)\right) &:= \left\{(t, z, w, v) \in \left[0,1\right] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u}\right. \\
&\qquad \left. \mid (z, w, v) \in \Sigma_{\rho}\left(x\left(t\right), y\left(t\right), u\left(t\right)\right)\right\}, \\
\Upsilon_{\rho}\left(x_0, x_1\right) &:= \mathcal{B}_{\rho}\left(x_0, x_1\right) \cap \left\{(z_0, z_1) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \mid \psi\left(z_0, z_1\right) = \mathbf{0}_{\mathbb{R}^{n_\psi}}\right\}.
\end{aligned}$$

For problems without inequality constraints $\mathbf{0}_{\mathbb{R}^{n_c}} \geq c\left(x\left(t\right), y\left(t\right), u\left(t\right)\right)$ the Hamilton-Jacobi equation (cf. [25, 27]) is applicable, i.e., one has to verify the existence of a function $V\left(t, x\right)$ satisfying the partial differential equation

$$V_t'\left(t, x\right) + V_x'\left(t, x\right) f\left(x, y, u\right) = 0, \tag{4.1.1}$$

and suitable boundary conditions in order to obtain optimality. In the same way as in [92, Theorem 3.1], we consider a generalized version of this approach, specifically, a Hamilton-Jacobi inequality with a quadratic deviation term, where $V\left(t, x\right)$ satisfies the equality (4.1.1) along the (potential) optimal solution $\left(\hat{x}\left(t\right), \hat{y}\left(t\right), \hat{u}\left(t\right)\right)$. The deviation term yields a quadratic growth condition for the objective function in the L_2-norm.

Theorem 4.1.1 (Hamilton-Jacobi Inequality)
Let $(\hat{x}, \hat{y}, \hat{u})$ be admissible for Problem 3.1.1 and $\hat{y}, \hat{u}$ be continuous on $[0,1]$. Suppose there exists a function $V : [0,1] \times \mathbb{R}^{n_x} \to \mathbb{R}$, which is piece-wise continuously differentiable with respect to t and continuously differentiable with respect to x, and there exist constants $\rho, \gamma > 0$ such that:

(i) For every $(t, x, y, u) \in graph\,(\Sigma_\rho (\hat{x}, \hat{y}, \hat{u}))$ it holds

$$V_t'(t,x) + V_x'(t,x)\, f(x,y,u) \geq \gamma \left(\|x - \hat{x}(t)\|^2 + \|y - \hat{y}(t)\|^2 + \|u - \hat{u}(t)\|^2 \right),$$

and for all $t \in [0,1]$ it holds

$$V_t'(t, \hat{x}(t)) + V_x'(t, \hat{x}(t))\, f(\hat{x}(t), \hat{y}(t), \hat{u}(t)) = 0.$$

(ii) For all $(x_0, x_1) \in \Upsilon_\rho (\hat{x}(0), \hat{x}(1))$ it holds

$$\begin{aligned} & V(1, \hat{x}(1)) - V(0, \hat{x}(0)) - [V(1, x_1) - V(0, x_0)] + \varphi(x_0, x_1) - \varphi(\hat{x}(0), \hat{x}(1)) \\ & \quad \geq \gamma \left(\|x_0 - \hat{x}(0)\|^2 + \|x_1 - \hat{x}(1)\|^2 \right). \end{aligned}$$

Then, the optimality condition

$$\varphi(x(0), x(1)) \geq \varphi(\hat{x}(0), \hat{x}(1)) + \gamma \left[\|(x,y,u) - (\hat{x}, \hat{y}, \hat{u})\|_2^2 + \|(x(0), x(1)) - (\hat{x}(0), \hat{x}(1))\|^2 \right]$$

is satisfied for all admissible $(x, y, u) \in \mathcal{B}_\rho^\infty (\hat{x}, \hat{y}, \hat{u})$.

Proof. Let $(x, y, u) \in \mathcal{B}_\rho^\infty (\hat{x}, \hat{y}, \hat{u})$ be feasible for Problem 3.1.1. Then, it holds

$$\begin{aligned} 0 &= \int_0^1 \frac{d}{dt} [V(t, x(t)) - V(t, \hat{x}(t))] - \frac{d}{dt} [V(t, x(t)) - V(t, \hat{x}(t))]\, dt \\ &= \int_0^1 V_t'(t, x(t)) + V_x'(t, x(t))\, f(x(t), y(t), u(t)) \\ &\quad - V_t'(t, \hat{x}(t)) - V_x'(t, \hat{x}(t))\, f(\hat{x}(t), \hat{y}(t), \hat{u}(t))\, dt \\ &\quad - [V(t, x(t)) - V(t, \hat{x}(t))]_{t=0}^{t=1} \\ &\overset{(i)}{\geq} \gamma \|(x,y,u) - (\hat{x}, \hat{y}, \hat{u})\|_2^2 + V(1, \hat{x}(1)) - V(0, \hat{x}(0)) - [V(1, x(1)) - V(0, x(0))] \\ &\overset{(ii)}{\geq} -\varphi(x(0), x(1)) + \varphi(\hat{x}(0), \hat{x}(1)) \\ &\quad + \gamma \left[\|(x,y,u) - (\hat{x}, \hat{y}, \hat{u})\|_2^2 + \|(x(0), x(1)) - (\hat{x}(0), \hat{x}(1))\|^2 \right], \end{aligned}$$

which proves the assertion. □

In order to derive second-order sufficient conditions for Problem 3.1.1, we use the following scheme (see Figure 4.1):

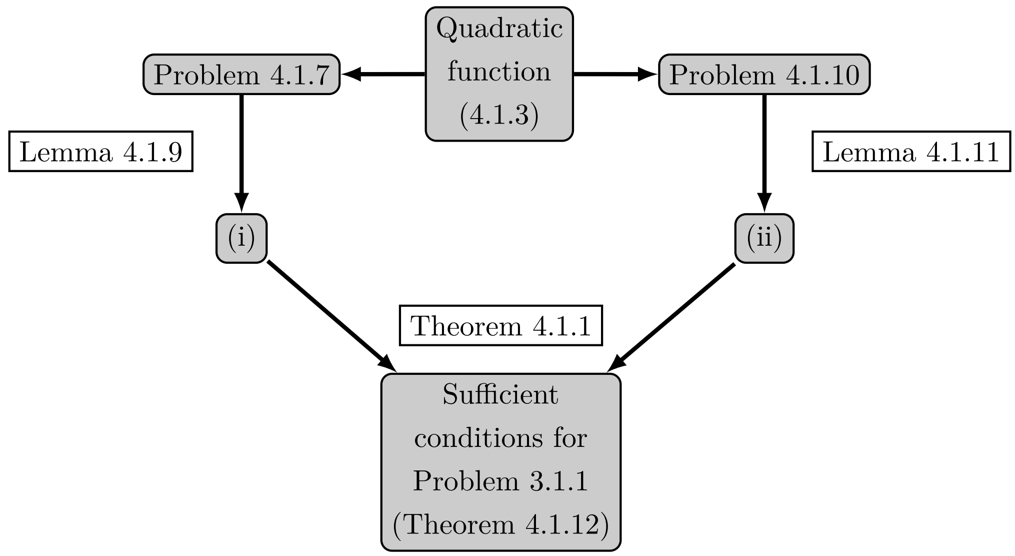

Figure 4.1: Scheme to derive second-order sufficient conditions for Problem 3.1.1.

(a) Construct a quadratic function using the solution of a Riccati equation.

(b) Verify conditions (i) and (ii) of Theorem 4.1.1 for the quadratic function:

  (i) Consider a parametric optimization problem depending on $t \in [0, 1]$, which is in the class of Problem 2.3.11. Show that the problem satisfies the sufficient conditions in Theorem 2.3.14 at a KKT-point.

  (ii) Verify the sufficient conditions in Theorem 2.3.5 for a finite dimensional optimization problem at a KKT-point.

(c) Apply Theorem 4.1.1.

Since we want to derive second-order sufficient conditions, we consider second derivatives of the functions in Problem 3.1.1. Hence, we need stronger smoothness assumptions than in Section 3.1. Furthermore, we require $\hat{y}$ and $\hat{u}$ to be continuous.

Assumption 4.1.2

(4.1.A1) *(Existence / Smoothness of a KKT-Point)*
Let

$$\left(\hat{x}, \hat{y}, \hat{u}, \hat{\lambda}_f, \hat{\lambda}_g, \hat{\sigma}, \hat{\eta}\right) \in W_{1,\infty}^{n_x}([0,1]) \times \mathcal{C}_0^{n_y}([0,1]) \times \mathcal{C}_0^{n_u}([0,1]) \\ \times W_{1,\infty}^{n_x}([0,1]) \times L_\infty^{n_y}([0,1]) \times \mathbb{R}^{n_\psi} \times L_\infty^{n_c}([0,1])$$

be a KKT-point of Problem 3.1.1.

(4.1.A2) *(Smoothness of the System)*

(a) *φ and ψ are twice continuously differentiable and all the derivatives are Lipschitz continuous with respect to all arguments.*

(b) *For a sufficiently large convex compact neighborhood $\mathcal{M}$ of*

$$\{(\hat{x}(t), \hat{y}(t), \hat{u}(t)) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \mid t \in [0,1]\},$$

let the mappings

$$(x,y,u) \mapsto f(x,y,u), \quad (x,y,u) \mapsto g(x,y,u), \quad (x,y,u) \mapsto c(x,y,u),$$

be twice continuously differentiable, and let all the derivatives be Lipschitz continuous in $\mathcal{M}$ with respect to all arguments.

Let $\mathbf{L}$ denote maximum over all Lipschitz constants of

$$f,\ g,\ c,\ f'_{(x,y,u)},\ g'_{(x,y,u)},\ c'_{(x,y,u)},\ f''_{(x,y,u)(x,y,u)},\ g''_{(x,y,u)(x,y,u)},\ c''_{(x,y,u)(x,y,u)}.$$

We call $\left(\hat{x}, \hat{y}, \hat{u}, \hat{\lambda}_f, \hat{\lambda}_g, \hat{\sigma}, \hat{\eta}\right)$ a KKT-point of Problem 3.1.1, if it solves the local minimum principle in Theorem 3.1.15 with $\ell_0 = 1$. Similar to Assumption 3.1.3, we require that the functions in Problem 3.1.1 satisfy certain regularity conditions at the KKT-point. Thus, for $t \in [0,1]$ we decompose the index set $J := \{1, \ldots, n_c\}$ into the index set of *active constraints* and the index set, where *strict complementarity* holds

$$\begin{aligned} J^0(t) &:= \{j \in J \mid c_j[t] = 0\}, & j^0(t) &:= \operatorname{card}\left(J^0(t)\right), \\ J^+(t) &:= \left\{j \in J^0(t) \mid \hat{\eta}_j(t) > 0\right\}, & j^+(t) &:= \operatorname{card}\left(J^+(t)\right). \end{aligned}$$

In Lemma 4.1.4 we show that $\hat{\eta}(\cdot)$ is continuous on $[0,1]$, if certain conditions are satisfied. Thus, $J^+(t)$ is defined for all $t \in [0,1]$. Moreover, we use the abbreviations

$$\begin{aligned} &A_f(\cdot) := f'_x[\cdot], && B_f(\cdot) := f'_y[\cdot], && C_f(\cdot) := f'_u[\cdot], \\ &A_g(\cdot) := g'_x[\cdot], && B_g(\cdot) := g'_y[\cdot], && C_g(\cdot) := g'_u[\cdot], \\ &\Psi_0 := \psi'_{x_0}(\hat{x}(0), \hat{x}(1)), && \Psi_1 := \psi'_{x_1}(\hat{x}(0), \hat{x}(1)), && \end{aligned}$$

and define the functions

$$A_c^0(\cdot) := \left[c'_{j,x}[\cdot]\right]_{j \in J^0(\cdot)}, \quad B_c^0(\cdot) := \left[c'_{j,y}[\cdot]\right]_{j \in J^0(\cdot)}, \quad C_c^0(\cdot) := \left[c'_{j,u}[\cdot]\right]_{j \in J^0(\cdot)},$$

$$A_c^+(\cdot) := \left[c'_{j,x}[\cdot]\right]_{j \in J^+(\cdot)}, \quad B_c^+(\cdot) := \left[c'_{j,y}[\cdot]\right]_{j \in J^+(\cdot)}, \quad C_c^+(\cdot) := \left[c'_{j,u}[\cdot]\right]_{j \in J^+(\cdot)},$$

$$\hat{\eta}^0(\cdot) := [\hat{\eta}_j(\cdot)]_{j \in J^0(\cdot)}, \qquad \hat{\eta}^+(\cdot) := [\hat{\eta}_j(\cdot)]_{j \in J^+(\cdot)},$$

which we consider to be vacuous, if $J^0(t)$ or $J^+(t)$ are empty, respectively. With these notations we assume the following:

Assumption 4.1.3 (Linear Independence, Controllability, Complementarity)

(4.1.A3) *(Index one / Regularity Condition)*
There exists a constant $\beta > 0$ such that for all $t \in [0,1]$ and $\varpi \in \mathbb{R}^{n_y} \times \mathbb{R}^{j^0(t)}$ it holds

$$\left\| \begin{bmatrix} B_g(t) & C_g(t) \\ B_c^0(t) & C_c^0(t) \end{bmatrix}^\top \varpi \right\| \geq \beta \|\varpi\| .$$

(4.1.A4) *(Controllability)*
For any $e \in \mathbb{R}^{n_\psi}$ there exists $(x,y,u) \in W_{1,\infty}^{n_x}([0,1]) \times L_\infty^{n_y}([0,1]) \times L_\infty^{n_u}([0,1])$ such that the DAE

$$\begin{aligned}
\dot{x}(t) &= A_f(t)x(t) + B_f(t)y(t) + C_f(t)u(t), && a.e.\ in\ [0,1],\\
\mathbf{0}_{\mathbb{R}^{n_y}} &= A_g(t)x(t) + B_g(t)y(t) + C_g(t)u(t), && a.e.\ in\ [0,1],\\
e &= \Psi_0 x(0) + \Psi_1 x(1),\\
\mathbf{0}_{\mathbb{R}^{j^0(t)}} &= A_c^0(t)x(t) + B_c^0(t)y(t) + C_c^0(t)u(t), && a.e.\ in\ [0,1],
\end{aligned}$$

is satisfied.

(4.1.A5) *(Strict Complementarity)*
The set $\{t \in [0,1] \mid J^0(t) \neq J^+(t)\}$ consists of finitely many junction points.

Note that (4.1.A3) combines the index property of the DAE and the regularity of the active inequality constraints (compare Remark 3.1.4). Condition (4.1.A5) implies that the strict complementarity condition is only violated by finitely many junction points in $[0,1]$, which corresponds to condition *(iii)* in Theorem 2.3.14. We recall the abbreviation (compare Chapter 1)

$$\mathcal{H}[\cdot] := \mathcal{H}\left(\hat{x}(\cdot), \hat{y}(\cdot), \hat{u}(\cdot), \hat{\lambda}_f(\cdot), \hat{\lambda}_g(\cdot), \hat{\eta}(\cdot)\right).$$

for the Hamilton function and its derivatives at the KKT-point. Condition (4.1.A3) assures that the multipliers are actually smoother than assumed, as the following lemma shows:

Lemma 4.1.4
If (4.1.A1) - (4.1.A3) hold, then we have

$$\left(\hat{\lambda}_g, \hat{\eta}\right) \in \mathcal{C}_0^{n_y}([0,1]) \times \mathcal{C}_0^{n_c}([0,1]) \quad and \quad \left(\hat{x}, \hat{\lambda}_f\right) \in \mathcal{C}_1^{n_x}([0,1]) \times \mathcal{C}_1^{n_x}([0,1]).$$

Proof. According to Theorem 3.1.15, for almost every $t \in [0,1]$ the KKT-point satisfies

$$\begin{aligned}
\mathbf{0}_{\mathbb{R}^{n_y}} &= \nabla_y \mathcal{H}[t] = B_f(t)^\top \hat{\lambda}_f(t) + B_g(t)^\top \hat{\lambda}_g(t) + B_c(t)^\top \hat{\eta}(t)\\
&= B_f(t)^\top \hat{\lambda}_f(t) + B_g(t)^\top \hat{\lambda}_g(t) + B_c^0(t)^\top \hat{\eta}^0(t),\\
\mathbf{0}_{\mathbb{R}^{n_u}} &= \nabla_u \mathcal{H}[t] = C_f(t)^\top \hat{\lambda}_f(t) + C_g(t)^\top \hat{\lambda}_g(t) + C_c(t)^\top \hat{\eta}(t)\\
&= C_f(t)^\top \hat{\lambda}_f(t) + C_g(t)^\top \hat{\lambda}_g(t) + C_c^0(t)^\top \hat{\eta}^0(t).
\end{aligned}$$

By (4.1.A3), the right inverse $\begin{bmatrix} B_g(\cdot) & C_g(\cdot) \\ B_c^0(\cdot) & C_c^0(\cdot) \end{bmatrix}^{\curlywedge}$ exists, thus we obtain

$$\begin{pmatrix} \hat{\lambda}_g(\cdot) \\ \hat{\eta}^0(\cdot) \end{pmatrix} = -\left(\begin{bmatrix} B_g(\cdot) & C_g(\cdot) \\ B_c^0(\cdot) & C_c^0(\cdot) \end{bmatrix}^{\curlywedge}\right)^\top \begin{bmatrix} B_f(\cdot)^\top \\ C_f(\cdot)^\top \end{bmatrix} \hat{\lambda}_f(\cdot).$$

Therefore, it holds $\left(\hat{\lambda}_g, \hat{\eta}\right) \in \mathcal{C}_0^{n_y}([0,1]) \times \mathcal{C}_0^{n_c}([0,1])$, since the right-hand side is continuous with respect to $t \in [0,1]$. Consequently, $\nabla_x \mathcal{H}[\cdot]$ is continuous and, by (4.1.A1), $f[\cdot]$ is continuous as well. Exploiting the differential equations

$$\begin{aligned} \dot{x}(t) &= f[t], && \text{in } [0,1], \\ \dot{\hat{\lambda}}_f(t) &= -\nabla_x \mathcal{H}[t], && \text{in } [0,1], \end{aligned}$$

yields $\left(\hat{x}, \hat{\lambda}_f\right) \in \mathcal{C}_1^{n_x}([0,1]) \times \mathcal{C}_1^{n_x}([0,1])$. □

For the construction of the function in Theorem 4.1.1 we require the existence of a bounded solution of a Riccati equation. To that end, for $\vartheta(x_0, x_1, \ell_0, \sigma) := \ell_0 \varphi(x_0, x_1) + \sigma^\top \psi(x_0, x_1)$ we define

$$\begin{aligned} &\Lambda_{00} := \nabla^2_{x_0 x_0} \vartheta(\hat{x}(0), \hat{x}(1), 1, \hat{\sigma}), \quad \Lambda_{01} := \nabla^2_{x_0 x_1} \vartheta(\hat{x}(0), \hat{x}(1), 1, \hat{\sigma}), \\ &\Lambda_{11} := \nabla^2_{x_1 x_1} \vartheta(\hat{x}(0), \hat{x}(1), 1, \hat{\sigma}), \end{aligned}$$

and assume the following:

Assumption 4.1.5 (Legendre-Clebsch, Riccati)

(4.1.A6) *(Legendre-Clebsch Condition)*

There exists a constant $\delta > 0$ such that for all $t \in [0,1]$ and $\varpi \in \ker\left(\begin{bmatrix} B_g(t) & C_g(t) \\ B_c^+(t) & C_c^+(t) \end{bmatrix}\right)$ the uniform Legendre-Clebsch condition

$$\varpi^\top \nabla^2_{(y,u)(y,u)} \mathcal{H}[t]\, \varpi \geq \delta \|\varpi\|^2$$

is satisfied.

(4.1.A7) *(Riccati Condition)*

For the matrix functions

$$Q(\cdot) := \nabla^2_{xx} \mathcal{H}[\cdot], \quad R^+(\cdot) := \begin{bmatrix} \nabla^2_{yy}\mathcal{H}[\cdot] & \nabla^2_{yu}\mathcal{H}[\cdot] & B_g(\cdot)^\top & B_c^+(\cdot)^\top \\ \nabla^2_{uy}\mathcal{H}[\cdot] & \nabla^2_{uu}\mathcal{H}[\cdot] & C_g(\cdot)^\top & C_c^+(\cdot)^\top \\ B_g(\cdot) & C_g(\cdot) & \mathbf{0}_{n_y \times n_y} & \mathbf{0}_{n_y \times j^+(\cdot)} \\ B_c^+(\cdot) & C_c^+(\cdot) & \mathbf{0}_{j^+(\cdot) \times n_y} & \mathbf{0}_{j^+(\cdot) \times j^+(\cdot)} \end{bmatrix},$$

$$S^+(\cdot) := \begin{bmatrix} \nabla^2_{yx}\mathcal{H}[\cdot] \\ \nabla^2_{ux}\mathcal{H}[\cdot] \\ A_g(\cdot) \\ A_c^+(\cdot) \end{bmatrix}, \quad K^+(\cdot) := \begin{bmatrix} B_f(\cdot)^\top \\ C_f(\cdot)^\top \\ \mathbf{0}_{n_y \times n_x} \\ \mathbf{0}_{j^+(\cdot) \times n_x} \end{bmatrix},$$

the Riccati equation

$$\begin{aligned}\dot{P}(t) &= -P(t)\,A_f(t) - A_f(t)^\top P(t) - Q(t)\\ &\quad + \left(K^+(t)\,P(t) + S^+(t)\right)^\top R^+(t)^{-1}\left(K^+(t)\,P(t) + S^+(t)\right) \quad \text{in } [0,1]\end{aligned}$$

subject to the boundary condition

$$\varpi^\top \begin{bmatrix} P(0)+\Lambda_{00} & \Lambda_{01} \\ \Lambda_{01}^\top & \Lambda_{11} - P(1) \end{bmatrix} \varpi > 0 \quad \text{for all } \varpi \in \ker([\Psi_0, \Psi_1]) \setminus \{\mathbf{0}_{\mathbb{R}^{2n_x}}\}$$

has a bounded solution.

(4.1.A6) implies that $R^+(\cdot)$ is non-singular on $[0,1]$ and the inverse is uniformly bounded, which assures that assumption (4.1.A7) is sensible. Furthermore, since we assumed that the Riccati equation possesses a bounded solution, it follows from a stability result for differential equations that this property is preserved, if we add the perturbation $\varepsilon\mathbf{I}_{n_x}$ to the right hand side for sufficiently small $\varepsilon > 0$. Additionally, the perturbed solution also satisfies the boundary conditions in (4.1.A7) for sufficiently small $\varepsilon > 0$.

Lemma 4.1.6
If (4.1.A1) - (4.1.A7) hold, then there exists a constant $\tilde{\varepsilon} > 0$ such that for all $0 \leq \varepsilon \leq \tilde{\varepsilon}$ the perturbed Riccati equation

$$\begin{aligned}\dot{P}(t) &= -P(t)\,A_f(t) - A_f(t)^\top P(t) - Q(t) + \varepsilon\mathbf{I}_{n_x}\\ &\quad + \left(K^+(t)\,P(t) + S^+(t)\right)^\top R^+(t)^{-1}\left(K^+(t)\,P(t) + S^+(t)\right) \quad \text{in } [0,1]\end{aligned} \tag{4.1.2}$$

subject to the boundary condition

$$\varpi^\top \begin{bmatrix} P(0)+\Lambda_{00} & \Lambda_{01} \\ \Lambda_{01}^\top & \Lambda_{11} - P(1) \end{bmatrix} \varpi > 0 \quad \text{for all } \varpi \in \ker([\Psi_0, \Psi_1]) \setminus \{\mathbf{0}_{\mathbb{R}^{2n_x}}\}$$

has a bounded solution P_ε.

Proof. According to [119, p. 103], there exists $\varepsilon_1 > 0$ such that for all $0 \leq \varepsilon \leq \varepsilon_1$ the perturbed Riccati equation has a bounded solution P_ε with $\lim_{\varepsilon\to 0} \|P_0 - P_\varepsilon\|_\infty = 0$, where P_0 is the reference solution, which satisfies the boundary conditions. Since

$$\begin{bmatrix} P_0(0)+\Lambda_{00} & \Lambda_{01} \\ \Lambda_{01}^\top & \Lambda_{11} - P_0(1) \end{bmatrix}$$

is positive definite on $\ker([\Psi_0, \Psi_1])$, there exists $\nu > 0$ such that

$$\varpi^\top \begin{bmatrix} P_0(0)+\Lambda_{00} & \Lambda_{01} \\ \Lambda_{01}^\top & \Lambda_{11} - P_0(1) \end{bmatrix} \varpi \geq \nu \|\varpi\|^2 \quad \text{for all } \ \varpi \in \ker([\Psi_0, \Psi_1]).$$

Choose $\tilde{\varepsilon} > 0$ such that for all $0 \leq \varepsilon \leq \tilde{\varepsilon}$ it holds $\|P_0 - P_\varepsilon\|_\infty \leq \frac{\nu}{4}$. Then, we obtain

$$\begin{aligned}
\varpi^\top \begin{bmatrix} P_\varepsilon(0) + \Lambda_{00} & \Lambda_{01} \\ \Lambda_{01}^\top & \Lambda_{11} - P_\varepsilon(1) \end{bmatrix} \varpi &= \varpi^\top \begin{bmatrix} P_0(0) + \Lambda_{00} & \Lambda_{01} \\ \Lambda_{01}^\top & \Lambda_{11} - P_0(1) \end{bmatrix} \varpi \\
&\quad - \varpi^\top \begin{bmatrix} P_0(0) - P_\varepsilon(0) & \mathbf{0}_{n_x \times n_x} \\ \mathbf{0}_{n_x \times n_x} & P_\varepsilon(1) - P_0(1) \end{bmatrix} \varpi \\
&\geq \nu \|\varpi\|^2 - \frac{\nu}{2} \|\varpi\|^2 = \frac{\nu}{2} \|\varpi\|^2
\end{aligned}$$

for every $\varpi \in \ker([\Psi_0, \Psi_1])$, which proves the assertion. □

For a fixed $\hat{\varepsilon} > 0$ such that $P_{\hat{\varepsilon}}(\cdot)$ is a bounded solution of the perturbed Riccati equation (4.1.2) subject to the boundary condition we define the quadratic function by

$$\begin{aligned}
&V(\cdot,\cdot) : [0,1] \times \mathbb{R}^{n_x} \to \mathbb{R}, \\
&V(t,x) := \hat{\lambda}_f(t)^\top (x - \hat{x}(t)) + \frac{1}{2} (x - \hat{x}(t))^\top P_{\hat{\varepsilon}}(t) (x - \hat{x}(t)) .
\end{aligned} \tag{4.1.3}$$

For this quadratic function we will first verify condition (i) in Theorem 4.1.1. To that end, for $(t,x) \in [0,1] \times \mathbb{R}^{n_x}$ we consider the derivatives

$$\begin{aligned}
V_t'(t,x) &= \dot{\hat{\lambda}}_f(t)^\top (x - \hat{x}(t)) - \hat{\lambda}_f(t)^\top \dot{\hat{x}}(t) - \dot{\hat{x}}(t)^\top P_{\hat{\varepsilon}}(t) (x - \hat{x}(t)) \\
&\quad + \frac{1}{2} (x - \hat{x}(t))^\top \dot{P}_{\hat{\varepsilon}}(t) (x - \hat{x}(t)) , \\
V_x'(t,x) f(x,y,u) &= \hat{\lambda}_f(t)^\top f(x,y,u) + (x - \hat{x}(t))^\top P_{\hat{\varepsilon}}(t) f(x,y,u) .
\end{aligned}$$

Adding these equations and evaluating at $(\hat{x}(t), \hat{y}(t), \hat{u}(t))$ yields

$$V_t'(t, \hat{x}(t)) + V_x'(t, \hat{x}(t)) f(\hat{x}(t), \hat{y}(t), \hat{u}(t)) = 0 \tag{4.1.4}$$

for all $t \in [0,1]$. Consider the parametric optimization problem depending on $t \in [0,1]$

Problem 4.1.7 (Parametric Optimization Problem)

$$\begin{aligned}
\textit{Minimize} \quad & V_t'(t,x) + V_x'(t,x) f(x,y,u) \\
\textit{with respect to} \quad & (x,y,u) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \\
\textit{subject to} \quad & g(x,y,u) = \mathbf{0}_{\mathbb{R}^{n_y}} , \\
& c(x,y,u) \leq \mathbf{0}_{\mathbb{R}^{n_c}} .
\end{aligned}$$

For fixed $t \in [0,1]$ the inequality of (i) represents an optimality condition of this problem, in particular, the quadratic growth condition

$$\begin{aligned}
&V_t'(t,x) + V_x'(t,x) f(x,y,u) - V_t'(t,\hat{x}(t)) + V_x'(t,\hat{x}(t)) f(\hat{x}(t), \hat{y}(t), \hat{u}(t)) \\
&\quad \geq \gamma \left(\|x - \hat{x}(t)\|^2 + \|y - \hat{y}(t)\|^2 + \|u - \hat{u}(t)\|^2 \right) ,
\end{aligned}$$

for all admissible (x,y,u) in a neighborhood of $(\hat{x}(t), \hat{y}(t), \hat{u}(t))$. In order to obtain this optimality condition, we aim to apply the sufficient conditions for problems with parametric objective function in Theorem 2.3.14. First, we show that $\left(\hat{x}(t), \hat{y}(t), \hat{u}(t), \hat{\lambda}_g(t), \hat{\eta}(t)\right)$ is a KKT-point of Problem 4.1.7 for all $t \in [0,1]$. To that end, we denote the associated Lagrange function by

$$\mathcal{L}(\ell_0, x, y, u, \lambda_g, \eta, t) := \ell_0 \left(V_t'(t,x) + V_x'(t,x) f(x,y,u)\right) + \lambda_g^\top g(x,y,u) + \eta^\top c(x,y,u) \quad (4.1.5)$$

for $t \in [0,1]$, which allows us to express the first order necessary conditions of Problem 4.1.7 as

$$\begin{aligned}
\mathbf{0}_{\mathbb{R}^{n_x}} &= \nabla_x \mathcal{L}(\ell_0, x, y, u, \lambda_g, \eta, t) \\
&= \ell_0 \Big[\dot{\hat{\lambda}}_f(t) - P_{\hat{\varepsilon}}(t)\dot{\hat{x}}(t) + \dot{P}_{\hat{\varepsilon}}(t)(x - \hat{x}(t)) \\
&\quad + f_x'(x,y,u)^\top \hat{\lambda}_f(t) + P_{\hat{\varepsilon}}(t) f(x,y,u) + f_x'(x,y,u)^\top P_{\hat{\varepsilon}}(t)(x - \hat{x}(t))\Big] \\
&\quad + g_x'(x,y,u)^\top \lambda_g + c_x'(x,y,u)^\top \eta, \\
\mathbf{0}_{\mathbb{R}^{n_y}} &= \nabla_y \mathcal{L}(\ell_0, x, y, u, \lambda_g, \eta, t) \\
&= \ell_0 \Big[f_y'(x,y,u)^\top \hat{\lambda}_f(t) + f_y'(x,y,u)^\top P_{\hat{\varepsilon}}(t)(x - \hat{x}(t))\Big] \\
&\quad + g_y'(x,y,u)^\top \lambda_g + c_y'(x,y,u)^\top \eta, \\
\mathbf{0}_{\mathbb{R}^{n_u}} &= \nabla_u \mathcal{L}(\ell_0, x, u, u, \lambda_g, \eta, t) \\
&= \ell_0 \Big[f_u'(x,u,u)^\top \hat{\lambda}_f(t) + f_u'(x,u,u)^\top P_{\hat{\varepsilon}}(t)(x - \hat{x}(t))\Big] \\
&\quad + g_u'(x,u,u)^\top \lambda_g + c_u'(x,u,u)^\top \eta, \\
0 &= \eta^\top c(x,y,u), \quad \eta \geq \mathbf{0}_{\mathbb{R}^{n_c}}
\end{aligned}$$

according to Theorem 2.3.2. If (4.1.A1) - (4.1.A3) hold, then for all $t \in [0,1]$ these conditions are satisfied by $\left(1, \hat{x}(t), \hat{y}(t), \hat{u}(t), \hat{\lambda}_g(t), \hat{\eta}(t)\right)$, since

$$\begin{aligned}
\nabla_x \mathcal{L}\left(1, \hat{x}(t), \hat{y}(t), \hat{u}(t), \hat{\lambda}_g(t), \hat{\eta}(t), t\right) &= \dot{\hat{\lambda}}_f(t) + \nabla_x \mathcal{H}[t] = \mathbf{0}_{\mathbb{R}^{n_x}}, \\
\nabla_y \mathcal{L}\left(1, \hat{x}(t), \hat{y}(t), \hat{u}(t), \hat{\lambda}_g(t), \hat{\eta}(t), t\right) &= \nabla_y \mathcal{H}[t] = \mathbf{0}_{\mathbb{R}^{n_y}}, \\
\nabla_u \mathcal{L}\left(1, \hat{x}(t), \hat{y}(t), \hat{u}(t), \hat{\lambda}_g(t), \hat{\eta}(t), t\right) &= \nabla_u \mathcal{H}[t] = \mathbf{0}_{\mathbb{R}^{n_u}}, \\
0 &= \hat{\eta}(t)^\top c(\hat{x}(t), \hat{y}(t), \hat{u}(t)), \quad \hat{\eta}(t) \geq \mathbf{0}_{\mathbb{R}^{n_c}},
\end{aligned}$$

correspond to the necessary conditions in Theorem 3.1.15. In order to apply Theorem 2.3.14, we have to verify the following:

(a) There exists a constant $\beta > 0$ such that for all $t \in [0,1]$ and $\varpi \in \mathbb{R}^{n_y} \times \mathbb{R}^{j^0(t)}$ it holds

$$\left\| \begin{bmatrix} B_g(t) & C_g(t) \\ B_c^0(t) & C_c^0(t) \end{bmatrix}^\top \varpi \right\| \geq \beta \|\varpi\|.$$

(b) There exists a constant $\kappa > 0$ such that for all $t \in [0,1]$ and for every

$$d \in \ker\left(\begin{bmatrix} A_g(t) & B_g(t) & C_g(t) \\ A_c^+(t) & B_c^+(t) & C_c^+(t) \end{bmatrix}\right)$$

it holds

$$d^\top \nabla^2_{(x,y,u)(x,y,u)} \mathcal{L}\left(1, \hat{x}(t), \hat{y}(t), \hat{u}(t), \hat{\lambda}_g(t), \hat{\eta}(t), t\right) d \geq \kappa \|d\|^2.$$

(c) The strict complementarity condition $J^0(t) = J^+(t)$ is only violated by finitely many $t \in [0,1]$.

The conditions (a) and (c) hold due to assumptions (4.1.A3) and (4.1.A5), respectively. Thus, it remains to show, that the matrix

$$\begin{aligned}
&\nabla^2_{(x,y,u)(x,y,u)}\mathcal{L}\left(1,\hat{x}(t),\hat{y}(t),\hat{u}(t),\hat{\lambda}_g(t),\hat{\eta}(t),t\right)\\
&= \begin{bmatrix} \dot{P}_{\hat{\varepsilon}}(t) + P_{\hat{\varepsilon}}(t)A_f(t) + A_f(t)^\top P_{\hat{\varepsilon}}(t) + Q(t) & \left([B_f(t),C_f(t)]^\top P_{\hat{\varepsilon}}(t) + \nabla^2_{(y,u)(x)}\mathcal{H}[t]\right)^\top \\ [B_f(t),C_f(t)]^\top P_{\hat{\varepsilon}}(t) + \nabla^2_{(y,u)(x)}\mathcal{H}[t] & \nabla^2_{(y,u)(y,u)}\mathcal{H}[t] \end{bmatrix}
\end{aligned}$$

is uniformly positive definite on $\ker\left(\begin{bmatrix} A_g(t) & B_g(t) & C_g(t) \\ A_c^+(t) & B_c^+(t) & C_c^+(t)\end{bmatrix}\right)$ for all $t \in [0,1]$. Therefore, we define the matrix functions

$$\begin{aligned}
M^+_{\hat{\varepsilon}}(\cdot) &:= \begin{bmatrix} \dot{P}_{\hat{\varepsilon}}(\cdot) + P_{\hat{\varepsilon}}(\cdot)A_f(\cdot) + A_f(\cdot)^\top P_{\hat{\varepsilon}}(\cdot) + Q(\cdot) & \left(K^+(\cdot)P_{\hat{\varepsilon}}(\cdot) + S^+(\cdot)\right)^\top \\ K^+(\cdot)P_{\hat{\varepsilon}}(\cdot) + S^+(\cdot) & R^+(\cdot)\end{bmatrix},\\
N^+(\cdot) &:= \begin{bmatrix} A_g(\cdot) & B_g(\cdot) & C_g(\cdot) \\ A_c^+(\cdot) & B_c^+(\cdot) & C_c^+(\cdot)\end{bmatrix},
\end{aligned}$$

which satisfy the following:

Lemma 4.1.8

If (4.1.A1) - (4.1.A7) hold, then there exists a constant $\kappa > 0$ such that for all $t \in [0,1]$ and for every $d \in \ker\left(N^+(t)\right)$ it holds

$$\begin{aligned}
&d^\top \nabla^2_{(x,y,u)(x,y,u)}\mathcal{L}\left(1,\hat{x}(t),\hat{y}(t),\hat{u}(t),\hat{\lambda}_g(t),\hat{\eta}(t),t\right)d\\
&= \begin{pmatrix} d \\ \mathbf{0}_{\mathbb{R}^{n_y+j^+(t)}}\end{pmatrix}^\top M^+_{\hat{\varepsilon}}(t)\begin{pmatrix} d \\ \mathbf{0}_{\mathbb{R}^{n_y+j^+(t)}}\end{pmatrix} \geq \kappa\|d\|^2.
\end{aligned}$$

Proof. We recall

$$R^+(\cdot) = \begin{bmatrix} \nabla^2_{yy}\mathcal{H}[\cdot] & \nabla^2_{yu}\mathcal{H}[\cdot] & B_g(\cdot)^\top & B_c^+(\cdot)^\top \\ \nabla^2_{uy}\mathcal{H}[\cdot] & \nabla^2_{uu}\mathcal{H}[\cdot] & C_g(\cdot)^\top & C_c^+(\cdot)^\top \\ B_g(\cdot) & C_g(\cdot) & \mathbf{0}_{n_y\times n_y} & \mathbf{0}_{n_y\times j^+(\cdot)} \\ B_c^+(\cdot) & C_c^+(\cdot) & \mathbf{0}_{j^+(\cdot)\times n_y} & \mathbf{0}_{j^+(\cdot)\times j^+(\cdot)}\end{bmatrix},$$

where $\begin{bmatrix} B_g(t) & C_g(t) \\ B_c^+(t) & C_c^+(t)\end{bmatrix}$ has full rank for all $t \in [0,1]$ by (4.1.A3), and $\begin{bmatrix} \nabla^2_{yy}\mathcal{H}[t] & \nabla^2_{yu}\mathcal{H}[t] \\ \nabla^2_{uy}\mathcal{H}[t] & \nabla^2_{uu}\mathcal{H}[t]\end{bmatrix}$ is uniformly positive definite on $\ker\left(\begin{bmatrix} B_g(t) & C_g(t) \\ B_c^+(t) & C_c^+(t)\end{bmatrix}\right)$ with constant $\delta > 0$ for all $t \in [0,1]$ according to (4.1.A6) . Additionally, it holds

$$\begin{aligned}
&\dot{P}_{\hat{\varepsilon}}(t) + P_{\hat{\varepsilon}}(t)A_f(t) + A_f(t)^\top P_{\hat{\varepsilon}}(t) + Q(t)\\
&\quad = \hat{\varepsilon}\mathbf{I}_{n_x} + \left(K^+(t)P_{\hat{\varepsilon}}(t) + S^+(t)\right)^\top R^+(t)^{-1}\left(K^+(t)P_{\hat{\varepsilon}}(t) + S^+(t)\right) \quad \text{in } [0,1].
\end{aligned}$$

Thus, for all $t \in [0,1]$ we have

$$M_{\hat{\varepsilon}}^{+}(t) = \begin{bmatrix} \hat{\varepsilon}\mathbf{I}_{n_x} + \left(K^{+}(t) P_{\hat{\varepsilon}}(t) + S^{+}(t)\right)^{\top} R^{+}(t)^{-1} \left(K^{+}(t) P_{\hat{\varepsilon}}(t) + S^{+}(t)\right) & \left(K^{+}(t) P_{\hat{\varepsilon}}(t) + S^{+}(t)\right)^{\top} \\ K^{+}(t) P_{\hat{\varepsilon}}(t) + S^{+}(t) & R^{+}(t) \end{bmatrix},$$

$$K^{+}(t) P_{\hat{\varepsilon}}(t) + S^{+}(t) = \begin{bmatrix} \left[B_f(t), C_f(t)\right]^{\top} P_{\hat{\varepsilon}}(t) + \nabla^2_{(y,u)(x)} \mathcal{H}[t] \\ \begin{bmatrix} A_g(\cdot) \\ A_c^{+}(\cdot) \end{bmatrix} \end{bmatrix},$$

and therefore all conditions of Lemma A.4 are satisfied. Since all the matrix functions are continuous on $[0,1]$ and $\delta, \hat{\varepsilon}$ are independent of t, we find a uniform constant $\kappa > 0$ such that the assertion holds. □

Consequently, we are able to prove that the inequality of (i) in Theorem 4.1.1 holds by applying Theorem 2.3.14:

Lemma 4.1.9

If (4.1.A1) - (4.1.A7) hold, then there exist constants $\tilde{\rho}, \tilde{\gamma} > 0$ such that for all $t \in [0,1]$ and every $(x,y,u) \in \mathcal{B}_{\tilde{\rho}}\left((\hat{x}(t), \hat{y}(t), \hat{u}(t))\right)$, which is feasible for Problem 4.1.7, it holds

$$\begin{aligned} & V_t'(t,x) + V_x'(t,x) f(x,y,u) \\ & \quad \geq V_t'(t,\hat{x}(t)) + V_x'(t,\hat{x}(t)) f(\hat{x}(t), \hat{y}(t), \hat{u}(t)) + \tilde{\gamma} \left\|(x,y,u) - (\hat{x}(t), \hat{y}(t), \hat{u}(t))\right\|^2 . \end{aligned}$$

Proof. According to (4.1.A3), (4.1.A5), and Lemma 4.1.8, all the requirements of Theorem 2.3.14 are satisfied for Problem 4.1.7, which proves the assertion. □

In (4.1.4) and Lemma 4.1.9 we have shown that condition (i) in Theorem 4.1.1 holds for the quadratic function defined in (4.1.3). In order to verify condition (ii), we define the objective function

$$\mathcal{J}(x_0, x_1) := \varphi(x_0, x_1) - \varphi(\hat{x}(0), \hat{x}(1)) - \left(V(1, x_1) - V(0, x_0)\right)$$

and consider the following optimization problem:

Problem 4.1.10

$$\begin{aligned} \textit{Minimize} \quad & \mathcal{J}(x_0, x_1) \\ \textit{with respect to} \quad & (x_0, x_1) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \\ \textit{subject to} \quad & \psi(x_0, x_1) = \mathbf{0}_{\mathbb{R}^{n_\psi}} . \end{aligned}$$

According to Theorem 2.3.2, the associated necessary conditions can be expressed as

$$\begin{aligned} \mathbf{0}_{\mathbb{R}^{n_x}} &= \nabla_{x_0} \mathcal{L}_{0,1}(\ell_0, x_0, x_1, \sigma) \\ &= \ell_0 \nabla_{x_0} \varphi(x_0, x_1) + \psi_{x_0}'(x_0, x_1)^{\top} \sigma + \ell_0 \left(\hat{\lambda}_f(0) + P_{\hat{\varepsilon}}(0)(x_0 - \hat{x}(0))\right), \\ \mathbf{0}_{\mathbb{R}^{n_x}} &= \nabla_{x_1} \mathcal{L}_{0,1}(\ell_0, x_0, x_1, \sigma) \\ &= \ell_0 \nabla_{x_1} \varphi(x_0, x_1) + \psi_{x_1}'(x_0, x_1)^{\top} \sigma - \ell_0 \left(\hat{\lambda}_f(1) + P_{\hat{\varepsilon}}(1)(x_1 - \hat{x}(1))\right), \end{aligned}$$

where the Lagrange function is defined by

$$\mathcal{L}_{0,1}\left(\ell_0, x_0, x_1, \sigma\right) := \ell_0 \mathcal{J}\left(x_0, x_1\right) + \sigma^\top \psi\left(x_0, x_1\right).$$

These conditions hold for the KKT-point $\left(\ell_0, x_0, x_1, \sigma\right) = \left(1, \hat{x}\left(0\right), \hat{x}\left(1\right), \hat{\sigma}\right)$ according to the transversality conditions in Theorem 3.1.15. Using the same techniques as in Lemma 3.1.7 we can show that the Gramian matrix associated with the linear system in (4.1.A4) has full rank, if and only if (4.1.A4) holds. The Gramian matrix is of the form

$$G = \left[\Psi_0, \Psi_1\right] \tilde{G} \begin{bmatrix} \Psi_0^\top \\ \Psi_1^\top \end{bmatrix},$$

which implies that $\left[\Psi_0, \Psi_1\right] = \psi'_{(x_0,x_1)}\left(\hat{x}\left(0\right), \hat{x}\left(1\right)\right)$ must have full rank. Furthermore, the matrix

$$\nabla^2_{(x_0,x_1)(x_0,x_1)} \mathcal{L}_{0,1}\left(1, \hat{x}\left(0\right), \hat{x}\left(1\right), \hat{\sigma}\right) = \begin{bmatrix} P_{\hat{\varepsilon}}\left(0\right) + \Lambda_{00} & \Lambda_{01} \\ \Lambda_{01}^\top & \Lambda_{11} - P_{\hat{\varepsilon}}\left(1\right) \end{bmatrix}$$

is positive definite on $\ker\left(\left[\Psi_0, \Psi_1\right]\right)$ by Lemma 4.1.6, which allows us to prove the following:

Lemma 4.1.11
If (4.1.A1) - (4.1.A7) hold, then there exist $\tilde{\rho}, \tilde{\gamma} > 0$ such that for every $\left(x_0, x_1\right) \in \mathcal{B}_{\tilde{\rho}}\left(\hat{x}\left(0\right), \hat{x}\left(1\right)\right)$ with $\psi\left(x_0, x_1\right) = \mathbf{0}_{\mathbb{R}^{n_\psi}}$ it holds

$$\begin{aligned} &\varphi\left(x_0, x_1\right) - \varphi\left(\hat{x}\left(0\right), \hat{x}\left(1\right)\right) - \left(V\left(1, x_1\right) - V\left(0, x_0\right)\right) \\ &\quad \geq -\left(V\left(1, \hat{x}\left(1\right)\right) - V\left(0, \hat{x}\left(0\right)\right)\right) + \tilde{\gamma}\left\|\left(x_0, x_1\right) - \left(\hat{x}\left(0\right), \hat{x}\left(1\right)\right)\right\|^2. \end{aligned}$$

Proof. All the requirements of Theorem 2.3.5 are satisfied for Problem 4.1.10 at $\left(\hat{x}\left(0\right), \hat{x}\left(1\right), \hat{\sigma}\right)$, which proves the assertion. □

We summarize the main result of this section in the following theorem:

Theorem 4.1.12 (Second-Order Sufficient Conditions for Problem 3.1.1)
If (4.1.A1) - (4.1.A7) hold, then there exist constants $\rho, \gamma > 0$ such that the optimality condition

$$\varphi\left(x\left(0\right), x\left(1\right)\right) \geq \varphi\left(\hat{x}\left(0\right), \hat{x}\left(1\right)\right) + \gamma\left[\left\|\left(x, y, u\right) - \left(\hat{x}, \hat{y}, \hat{u}\right)\right\|_2^2 + \left\|\left(x\left(0\right), x\left(1\right)\right) - \left(\hat{x}\left(0\right), \hat{x}\left(1\right)\right)\right\|^2\right]$$

is satisfied for all admissible $\left(x, y, u\right) \in \mathcal{B}^\infty_\rho\left(\hat{x}, \hat{y}, \hat{u}\right)$ of Problem 3.1.1.

4.2 Sufficient Conditions for Higher Index Problems

Similar to Section 3.2, we also derive sufficient conditions for Problem 3.2.1

$$\begin{array}{ll} \text{Minimize} & \varphi\left(\boldsymbol{x}\left(0\right),\boldsymbol{x}\left(1\right)\right), \\ \\ \text{with respect to} & \boldsymbol{x}=\left(x_1,\ldots,x_{k-1}\right)\in \bigtimes_{i=1}^{k-1} W_{i,\infty}^{n_{x_i}}\left(\left[0,1\right]\right), y\in L_{\infty}^{n_y}\left(\left[0,1\right]\right), u\in L_{\infty}^{n_u}\left(\left[0,1\right]\right), \end{array}$$

$$\text{subject to}\quad \begin{array}{rcll} \dot{x}_1\left(t\right) &=& f_1\left(x_1\left(t\right),\ldots,x_{k-1}\left(t\right),y\left(t\right),u\left(t\right)\right), & \text{a.e. in } \left[0,1\right], \\ \dot{x}_2\left(t\right) &=& f_2\left(x_1\left(t\right),\ldots,x_{k-1}\left(t\right)\right), & \text{in } \left[0,1\right], \\ \dot{x}_3\left(t\right) &=& f_3\left(x_2\left(t\right),\ldots,x_{k-1}\left(t\right)\right), & \text{in } \left[0,1\right], \\ &\vdots&& \\ \dot{x}_{k-1}\left(t\right) &=& f_{k-1}\left(x_{k-2}\left(t\right),x_{k-1}\left(t\right)\right), & \text{in } \left[0,1\right], \\ \mathbf{0}_{\mathbb{R}^{n_y}} &=& g\left(x_{k-1}\left(t\right)\right), & \text{in } \left[0,1\right], \\ \mathbf{0}_{\mathbb{R}^{n_\psi}} &=& \psi\left(\boldsymbol{x}\left(0\right),\boldsymbol{x}\left(1\right)\right), & \\ \mathbf{0}_{\mathbb{R}^{n_c}} &\geq& c\left(x_1\left(t\right),\ldots,x_{k-1}\left(t\right),y\left(t\right),u\left(t\right)\right), & \text{a.e. in } \left[0,1\right]. \end{array}$$

For the index reduced system (3.2.1) we use the abbreviations at the KKT-point

$$\begin{array}{lll} A_{\boldsymbol{f}}\left(\cdot\right):=\boldsymbol{f}'_{\boldsymbol{x}}\left[\cdot\right], & B_{\boldsymbol{f}}\left(\cdot\right):=\boldsymbol{f}'_{y}\left[\cdot\right], & C_{\boldsymbol{f}}\left(\cdot\right):=\boldsymbol{f}'_{u}\left[\cdot\right], \\ A_{g_0}\left(\cdot\right):=g'_{0,\boldsymbol{x}}\left[\cdot\right], & B_{g_0}\left(\cdot\right):=g'_{0,y}\left[\cdot\right], & C_{g_0}\left(\cdot\right):=g'_{0,u}\left[\cdot\right], \\ \Psi_0^{\boldsymbol{g}}:=\begin{bmatrix} \boldsymbol{g}'\left[0\right] \\ \psi'_{\boldsymbol{x}_0}\left(\hat{\boldsymbol{x}}\left(0\right),\hat{\boldsymbol{x}}\left(1\right)\right)\end{bmatrix}, & \Psi_1^{\boldsymbol{g}}:=\begin{bmatrix} \mathbf{0}_{(k-1)n_y\times n_{\boldsymbol{x}}} \\ \psi'_{\boldsymbol{x}_1}\left(\hat{\boldsymbol{x}}\left(0\right),\hat{\boldsymbol{x}}\left(1\right)\right)\end{bmatrix}, & \\ \\ A_c^0\left(\cdot\right):=\left[c'_{j,\boldsymbol{x}}\left[\cdot\right]\right]_{j\in J^0(\cdot)}, & B_c^0\left(\cdot\right):=\left[c'_{j,y}\left[\cdot\right]\right]_{j\in J^0(\cdot)}, & C_c^0\left(\cdot\right):=\left[c'_{j,u}\left[\cdot\right]\right]_{j\in J^0(\cdot)}, \\ \\ A_c^+\left(\cdot\right):=\left[c'_{j,\boldsymbol{x}}\left[\cdot\right]\right]_{j\in J^+(\cdot)}, & B_c^+\left(\cdot\right):=\left[c'_{j,y}\left[\cdot\right]\right]_{j\in J^+(\cdot)}, & C_c^+\left(\cdot\right):=\left[c'_{j,u}\left[\cdot\right]\right]_{j\in J^+(\cdot)}, \end{array}$$

and define

$$\begin{aligned} &\vartheta\left(\boldsymbol{x}_0,\boldsymbol{x}_1,\ell_0,\sigma_\psi,\sigma_{\boldsymbol{g}}\right):=\ell_0\varphi\left(\boldsymbol{x}_0,\boldsymbol{x}_1\right)+\sigma_\psi^\top\psi\left(\boldsymbol{x}_0,\boldsymbol{x}_1\right)+\sigma_{\boldsymbol{g}}^\top\boldsymbol{g}\left(\boldsymbol{x}_0\right), \\ &\quad\Lambda_{00}:=\nabla^2_{\boldsymbol{x}_0\boldsymbol{x}_0}\vartheta\left(\hat{\boldsymbol{x}}\left(0\right),\hat{\boldsymbol{x}}\left(1\right),1,\hat{\sigma}_\psi,\hat{\sigma}_{\boldsymbol{g}}\right), \quad \Lambda_{01}:=\nabla^2_{\boldsymbol{x}_0\boldsymbol{x}_1}\vartheta\left(\hat{\boldsymbol{x}}\left(0\right),\hat{\boldsymbol{x}}\left(1\right),1,\hat{\sigma}_\psi,\hat{\sigma}_{\boldsymbol{g}}\right), \\ &\quad\Lambda_{11}:=\nabla^2_{\boldsymbol{x}_1\boldsymbol{x}_1}\vartheta\left(\hat{\boldsymbol{x}}\left(0\right),\hat{\boldsymbol{x}}\left(1\right),1,\hat{\sigma}_\psi,\hat{\sigma}_{\boldsymbol{g}}\right). \end{aligned}$$

Then, with the (augmented) Hamilton function

$$\mathcal{H}\left(\boldsymbol{x},y,u,\boldsymbol{\lambda_f},\lambda_{g_0},\eta\right):=\boldsymbol{\lambda}_{\boldsymbol{f}}^\top\boldsymbol{f}\left(\boldsymbol{x},y,u\right)+\lambda_{g_0}^\top g_0\left(\boldsymbol{x},y,u\right)+\eta^\top c\left(\boldsymbol{x},y,u\right)$$

we assume the following:

Assumption 4.2.1

(4.2.A1) *(Existence / Smoothness of a KKT-Point)*
Let

$$\left(\hat{\boldsymbol{x}}, \hat{y}, \hat{u}, \hat{\boldsymbol{\lambda}}_{\boldsymbol{f}}, \hat{\lambda}_{g_0}, \hat{\sigma}_{\psi}, \hat{\sigma}_{\boldsymbol{g}}, \hat{\eta}\right) \in \bigtimes_{i=1}^{k-1} W_{i,\infty}^{n_{x_i}}\left([0,1]\right) \times \mathcal{C}_0^{n_y}\left([0,1]\right) \times \mathcal{C}_0^{n_u}\left([0,1]\right) \times \bigtimes_{i=1}^{k-1} W_{1,\infty}^{n_{x_i}}\left([0,1]\right) \\ \times L_{\infty}^{n_y}\left([0,1]\right) \times \mathbb{R}^{n_\psi} \times \mathbb{R}^{(k-1)n_y} \times L_{\infty}^{n_c}\left([0,1]\right)$$

be a KKT-point of of Problem 3.2.1.

(4.2.A2) *(Smoothness of the System)*

(a) φ and ψ are twice continuously differentiable and all the derivatives are Lipschitz continuous with respect to all arguments.

(b) For a sufficiently large convex compact neighborhood $\mathcal{M}_1$ of

$$\left\{\left(\hat{x}_1\left(t\right), \ldots, \hat{x}_{k-1}\left(t\right), \hat{y}\left(t\right), \hat{u}\left(t\right)\right) \in \bigtimes_{i=1}^{k-1} \mathbb{R}^{n_{x_i}} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \mid t \in [0,1]\right\},$$

let the mappings

$$\left(x_1, \ldots, x_{k-1}, y, u\right) \mapsto f_1\left(x_1, \ldots, x_{k-1}, y, u\right),$$
$$\left(x_1, \ldots, x_{k-1}, y, u\right) \mapsto c\left(x_1, \ldots, x_{k-1}, y, u\right),$$

be twice continuously differentiable, and all the derivatives be Lipschitz continuous in $\mathcal{M}_1$ with respect to all arguments. Furthermore, for $i = 2, \ldots, k-1$ and sufficiently large convex compact neighborhoods $\mathcal{M}_i$ of

$$\left\{\left(\hat{x}_{i-1}\left(t\right), \ldots, \hat{x}_{k-1}\left(t\right)\right) \in \bigtimes_{l=i-1}^{k-1} \mathbb{R}^{n_{x_l}} \mid t \in [0,1]\right\},$$

let the mappings

$$\left(x_{i-1}, \ldots, x_{k-1}\right) \mapsto f_i\left(x_{i-1}, \ldots, x_{k-1}\right),$$

be $i+1$-times continuously differentiable, and all the derivatives be Lipschitz continuous in $\mathcal{M}_i$ with respect to all arguments. For a sufficiently large convex compact neighborhood $\mathcal{M}_k$ of

$$\left\{\hat{x}_{k-1}\left(t\right) \in \mathbb{R}^{n_{k-1}} \mid t \in [0,1]\right\},$$

let the mapping

$$x_{k-1} \mapsto g\left(x_{k-1}\right),$$

be $k+1$-times continuously differentiable, and all the derivatives be Lipschitz continuous in $\mathcal{M}_k$ with respect to all arguments.

(4.2.A3) *(Index k / Regularity Condition)*

There exists a constant $\beta > 0$ *such that for all* $t \in [0,1]$ *and* $\varpi \in \mathbb{R}^{n_y + j^0(t)}$ *it holds*

$$\left\| \begin{bmatrix} B_{g_0}(t) & C_{g_0}(t) \\ B_c^0(t) & C_c^0(t) \end{bmatrix}^\top \varpi \right\| \geq \beta \left\|\varpi\right\|.$$

(4.2.A4) *(Controllability)*

For any $e \in \mathbb{R}^{(k-1)n_y + n_\psi}$ *there exists* $(\boldsymbol{x}, y, u) \in \bigtimes_{i=1}^{k-1} W_{i,\infty}^{n_{x_i}}([0,1]) \times L_\infty^{n_y}([0,1]) \times L_\infty^{n_u}([0,1])$ *such that the DAE*

$$\begin{aligned}
\dot{\boldsymbol{x}}(t) &= A_{\boldsymbol{f}}(t)\,\boldsymbol{x}(t) + B_{\boldsymbol{f}}(t)\,y(t) + C_{\boldsymbol{f}}(t)\,u(t), && \textit{a.e. in } [0,1], \\
\mathbf{0}_{\mathbb{R}^{n_y}} &= A_{g_0}(t)\,\boldsymbol{x}(t) + B_{g_0}(t)\,y(t) + C_{g_0}(t)\,u(t), && \textit{a.e. in } [0,1], \\
e &= \Psi_0^{\boldsymbol{g}}\boldsymbol{x}(0) + \Psi_1^{\boldsymbol{g}}\boldsymbol{x}(1), \\
\mathbf{0}_{\mathbb{R}^{j^0(t)}} &= A_c^0(t)\,\boldsymbol{x}(t) + B_c^0(t)\,y(t) + C_c^0(t)\,u(t), && \textit{a.e. in } [0,1],
\end{aligned}$$

is satisfied.

(4.2.A5) *(Strict Complementarity)*

The set $\left\{t \in [0,1] \mid J^0(t) \neq J^+(t)\right\}$ *consists of finitely many junction points.*

(4.2.A6) *(Legendre-Clebsch Condition)*

There exists a constant $\delta > 0$ *such that for all* $t \in [0,1]$ *and* $\varpi \in \ker\left(\begin{bmatrix} B_{g_0}(t) & C_{g_0}(t) \\ B_c^+(t) & C_c^+(t) \end{bmatrix}\right)$ *the uniform Legendre-Clebsch condition*

$$\varpi^\top \nabla^2_{(y,u)(y,u)}\mathcal{H}[t]\,\varpi \geq \delta \left\|\varpi\right\|^2$$

is satisfied.

(4.2.A7) *(Riccati Condition)*

For the matrix functions

$$Q(\cdot) := \nabla^2_{\boldsymbol{xx}}\mathcal{H}[\cdot], \quad R^+(\cdot) := \begin{bmatrix} \nabla^2_{yy}\mathcal{H}[\cdot] & \nabla^2_{yu}\mathcal{H}[\cdot] & B_{g_0}(\cdot)^\top & B_c^+(\cdot)^\top \\ \nabla^2_{uy}\mathcal{H}[\cdot] & \nabla^2_{uu}\mathcal{H}[\cdot] & C_{g_0}(\cdot)^\top & C_c^+(\cdot)^\top \\ B_{g_0}(\cdot) & C_{g_0}(\cdot) & \mathbf{0}_{n_y \times n_y} & \mathbf{0}_{n_y \times j^+(\cdot)} \\ B_c^+(\cdot) & C_c^+(\cdot) & \mathbf{0}_{j^+(\cdot) \times n_y} & \mathbf{0}_{j^+(\cdot) \times j^+(\cdot)} \end{bmatrix},$$

$$S^+(\cdot) := \begin{bmatrix} \nabla^2_{y\boldsymbol{x}}\mathcal{H}[\cdot] \\ \nabla^2_{u\boldsymbol{x}}\mathcal{H}[\cdot] \\ A_{g_0}(\cdot) \\ A_c^+(\cdot) \end{bmatrix}, \quad K^+(\cdot) := \begin{bmatrix} B_{\boldsymbol{f}}(\cdot)^\top \\ C_{\boldsymbol{f}}(\cdot)^\top \\ \mathbf{0}_{n_y \times n_{\boldsymbol{x}}} \\ \mathbf{0}_{j^+(\cdot) \times n_{\boldsymbol{x}}} \end{bmatrix},$$

the Riccati equation

$$\begin{aligned}\dot{P}(t) &= -P(t)\,A_{\boldsymbol{f}}(t) - A_{\boldsymbol{f}}(t)^{\top} P(t) - Q(t) \\ &\quad + \left(K^{+}(t)\,P(t) + S^{+}(t)\right)^{\top} R^{+}(t)^{-1}\left(K^{+}(t)\,P(t) + S^{+}(t)\right) \quad \textit{in } [0,1]\end{aligned}$$

subject to the boundary condition

$$\varpi^{\top}\left[\begin{array}{cc} P(0)+\Lambda_{00} & \Lambda_{01} \\ \Lambda_{01}^{\top} & \Lambda_{11}-P(1)\end{array}\right]\varpi > 0 \quad \textit{for all } \varpi \in \ker\left(\left[\Psi_0^{\boldsymbol{g}}, \Psi_1^{\boldsymbol{g}}\right]\right) \setminus \left\{\mathbf{0}_{\mathbb{R}^{2n_{\boldsymbol{x}}}}\right\}$$

has a bounded solution.

Analog to Lemma 4.1.4, we can show that $\hat{\lambda}_g$, $\hat{\eta}$ are continuous and $\hat{x}_1$, $\hat{\lambda}_{f_1}$ are continuously differentiable in $[0,1]$. For a sufficiently small $\hat{\varepsilon} > 0$ we denote the bounded solution of the perturbed Riccati equation

$$\begin{aligned}\dot{P}(t) &= -P(t)\,A_{\boldsymbol{f}}(t) - A_{\boldsymbol{f}}(t)^{\top} P(t) - Q(t) + \hat{\varepsilon}\mathbf{I}_{n_{\boldsymbol{x}}} \\ &\quad + \left(K^{+}(t)\,P(t) + S^{+}(t)\right)^{\top} R^{+}(t)^{-1}\left(K^{+}(t)\,P(t) + S^{+}(t)\right) \quad \text{in } [0,1]\end{aligned}$$

subject to the boundary condition in $(4.2.A7)$ by $P_{\hat{\varepsilon}}(\cdot)$. We define the quadratic function

$$\begin{aligned}&V(\cdot,\cdot) : [0,1]\times\mathbb{R}^{n_{\boldsymbol{x}}} \to \mathbb{R}, \\ &V(t,\boldsymbol{x}) := \hat{\boldsymbol{\lambda}}_{\boldsymbol{f}}(t)^{\top}(\boldsymbol{x}-\hat{\boldsymbol{x}}(t)) + \frac{1}{2}(\boldsymbol{x}-\hat{\boldsymbol{x}}(t))^{\top} P_{\hat{\varepsilon}}(t)\,(\boldsymbol{x}-\hat{\boldsymbol{x}}(t)),\end{aligned}$$

which satisfies the condition of Theorem 4.1.1 for Problem 3.2.1 subject to the equivalent reduced (index one) system (3.2.1). Thus, we obtain the main result of this section:

Theorem 4.2.2 (Second-Order Sufficient Conditions for Problem 3.2.1)
If (4.2.A1) - (4.2.A7) hold, then there exist constants $\rho, \gamma > 0$ such that the optimality condition

$$\varphi(\boldsymbol{x}(0),\boldsymbol{x}(1)) \geq \varphi(\hat{\boldsymbol{x}}(0),\hat{\boldsymbol{x}}(1)) + \gamma\left[\|(\boldsymbol{x},y,u)-(\hat{\boldsymbol{x}},\hat{y},\hat{u})\|_2^2 + \|(\boldsymbol{x}(0),\boldsymbol{x}(1))-(\hat{\boldsymbol{x}}(0),\hat{\boldsymbol{x}}(1))\|^2\right]$$

is satisfied for all admissible $(\boldsymbol{x},y,u) \in \mathcal{B}_{\rho}^{\infty}(\hat{\boldsymbol{x}},\hat{y},\hat{u})$ of Problem 3.2.1.

4.3 Example

For Example 3.3.1 we want to verify, if the calculated solution is actually a weak local minimizer by applying Theorem 4.2.2. To that end, it remains to show that the conditions (4.2.A6) and (4.2.A7) hold. The associated Hamilton function is defined by

$$\begin{aligned}&\mathcal{H}(x_1,x_2,x_3,x_4,y,u,\lambda_{f_1},\lambda_{f_2},\lambda_{f_3},\lambda_{f_4},\lambda_g) \\ &\quad := \lambda_{f_1}(u-y) + \lambda_{f_2}u - \lambda_{f_3}x_2 + \frac{1}{2}\lambda_{f_4}u^2 + \lambda_g(u-y-x_2).\end{aligned}$$

Then, (4.2.A6) is satisfied, since $\ker([B_{g_0}(t)\,C_{g_0}(t)]) = \ker((-1,1))$, and therefore the Legendre-Clebsch condition

$$\left(\begin{array}{c}\varpi\\ \varpi\end{array}\right)^{\top}\nabla^2_{(y,u)(y,u)}\mathcal{H}[t]\left(\begin{array}{c}\varpi\\ \varpi\end{array}\right) = \left(\begin{array}{c}\varpi\\ \varpi\end{array}\right)^{\top}\left(\begin{array}{cc}0&0\\0&1\end{array}\right)\left(\begin{array}{c}\varpi\\ \varpi\end{array}\right) = \varpi^2 = \frac{1}{2}\left\|\left(\begin{array}{c}\varpi\\ \varpi\end{array}\right)\right\|^2$$

holds for all $\begin{pmatrix} \varpi \\ \varpi \end{pmatrix} \in \ker\left((-1,1)\right)$. For (4.2.A6) we compute the (constant) matrices

$$\begin{gathered}
R^{+} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & 0 \end{pmatrix}, \quad \left(R^{+}\right)^{-1} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \\
S^{+} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad K^{+} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 1 & 1 & 0 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\
Q = \mathbf{0}_{4\times 4}, \quad \Lambda_{00} = \mathbf{0}_{4\times 4}, \quad \Lambda_{01} = \mathbf{0}_{4\times 4}, \quad \Lambda_{11} = \mathbf{0}_{4\times 4}, \\
\left[\Psi_0^{\boldsymbol{g}}, \Psi_1^{\boldsymbol{g}}\right] = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix},
\end{gathered}$$

which yields the Riccati equation

$$\begin{aligned}
\dot{p}_{11}(t) &= \left(p_{12}(t) - 2p_{14}(t)\right)^2 \\
\dot{p}_{12}(t) &= \left(p_{12}(t) - 2p_{14}(t)\right)\left(p_{22}(t) - 2p_{24}(t)\right) + \left(p_{13}(t) - p_{11}(t)\right) \\
\dot{p}_{13}(t) &= \left(p_{12}(t) - 2p_{14}(t)\right)\left(p_{23}(t) - 2p_{34}(t)\right) \\
\dot{p}_{14}(t) &= \left(p_{12}(t) - 2p_{14}(t)\right)\left(p_{24}(t) - 2p_{44}(t)\right) \\
\dot{p}_{22}(t) &= \left(p_{22}(t) - 2p_{24}(t)\right)^2 + 2\left(p_{23}(t) - p_{12}(t)\right) \\
\dot{p}_{23}(t) &= \left(p_{22}(t) - 2p_{24}(t)\right)\left(p_{23}(t) - 2p_{34}(t)\right) + \left(p_{33}(t) - p_{13}(t)\right) \\
\dot{p}_{24}(t) &= \left(p_{22}(t) - 2p_{24}(t)\right)\left(p_{24}(t) - 2p_{44}(t)\right) + \left(p_{34}(t) - p_{14}(t)\right) \\
\dot{p}_{33}(t) &= \left(p_{23}(t) - 2p_{34}(t)\right)^2 \\
\dot{p}_{34}(t) &= \left(p_{23}(t) - 2p_{34}(t)\right)\left(p_{24}(t) - 2p_{44}(t)\right) \\
\dot{p}_{44}(t) &= \left(p_{24}(t) - 2p_{44}(t)\right)^2
\end{aligned}$$

for the symmetric matrix function

$$P(\cdot) := \begin{pmatrix} p_{11}(\cdot) & p_{12}(\cdot) & p_{13}(\cdot) & p_{14}(\cdot) \\ p_{12}(\cdot) & p_{22}(\cdot) & p_{23}(\cdot) & p_{24}(\cdot) \\ p_{13}(\cdot) & p_{23}(\cdot) & p_{33}(\cdot) & p_{34}(\cdot) \\ p_{14}(\cdot) & p_{24}(\cdot) & p_{34}(\cdot) & p_{44}(\cdot) \end{pmatrix}.$$

This Riccati equation has the bounded (constant) solution

$$P(\cdot) = \begin{pmatrix} -1 & 0 & -1 & 0 \\ 0 & -4 & 0 & -2 \\ -1 & 0 & -1 & 0 \\ 0 & -2 & 0 & -1 \end{pmatrix}.$$

Since $\ker\left(\left[\Psi_0^{\boldsymbol{g}}, \Psi_1^{\boldsymbol{g}}\right]\right) = \left\{(0,0,0,0,0,0,\varpi_1,\varpi_2) \mid \varpi_1,\varpi_2 \in \mathbb{R}\right\}$, only the definiteness of

$$\begin{pmatrix} p_{33}(1) & p_{34}(1) \\ p_{34}(1) & p_{44}(1) \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

is relevant for the boundary conditions to be satisfied. We obtain

$$\varpi^\top \begin{bmatrix} P(0) + \Lambda_{00} & \Lambda_{01} \\ \Lambda_{01}^\top & \Lambda_{11} - P(1) \end{bmatrix} \varpi = \varpi_1^2 + \varpi_2^2 > 0$$

for all $\varpi \in \ker\left(\left[\Psi_0^{\boldsymbol{g}}, \Psi_1^{\boldsymbol{g}}\right]\right) \setminus \{\mathbf{0}_{\mathbb{R}^8}\}$, and therefore the boundary conditions are satisfied. For illustration purposes, we show that for $\varepsilon = \frac{1}{16}$ the perturbed Riccati has a bounded solution, which satisfies the boundary conditions (see Figure 4.2), since the matrix

$$\begin{pmatrix} p_{33,\varepsilon}(1) & p_{34,\varepsilon}(1) \\ p_{34,\varepsilon}(1) & p_{44,\varepsilon}(1) \end{pmatrix} \approx \begin{pmatrix} -0.937278 & -0.001125 \\ -0.001125 & -0.931479 \end{pmatrix}$$

is negative definite. Moreover, the error with respect to the reference solution $p_{33}(1) = -1$, $p_{34}(1) = 0$, $p_{44}(1) = -1$ is declining with decreasing perturbation (see Table 4.1).

ε	$p_{33,\varepsilon}(1)$	$p_{34,\varepsilon}(1)$	$p_{44,\varepsilon}(1)$
0.25	−0.7435632620790397	−0.3365590016931926	−0.5681414623634595
0.125	−0.8739585476272056	−0.0053191628693259	−0.8464515869344885
0.0625	−0.9372772356430186	−0.0011254386360057	−0.9314790146543683
0.02	−0.9799791865196865	−0.0001044079420916	−0.9794426109138125
0.01	−0.9899949005508796	−0.0000255387721687	−0.9898637265415218
0.005	−0.9949987377039996	−0.0000063166079344	−0.9949663031908956
0.002	−0.9979997992181540	−0.0000010042350419	−0.9979946435566353
0.001	−0.9989999499024794	−0.0000002505282293	−0.9989986637842081

Table 4.1: Illustration of decreasing error with respect to the reference solution.

In this chapter, we derived second-order sufficient conditions for optimal control problems subject to Hessenberg DAEs of arbitrary order and mixed control-state constraints by using a Hamilton Jacobi inequality. In contrast to [83] and [92], we also included boundary conditions and algebraic equations, respectively. The main task was to prove Theorem 2.3.14 with the assumptions at hand, which was essential in order to show that second-order sufficient conditions hold for Problem 4.1.7.

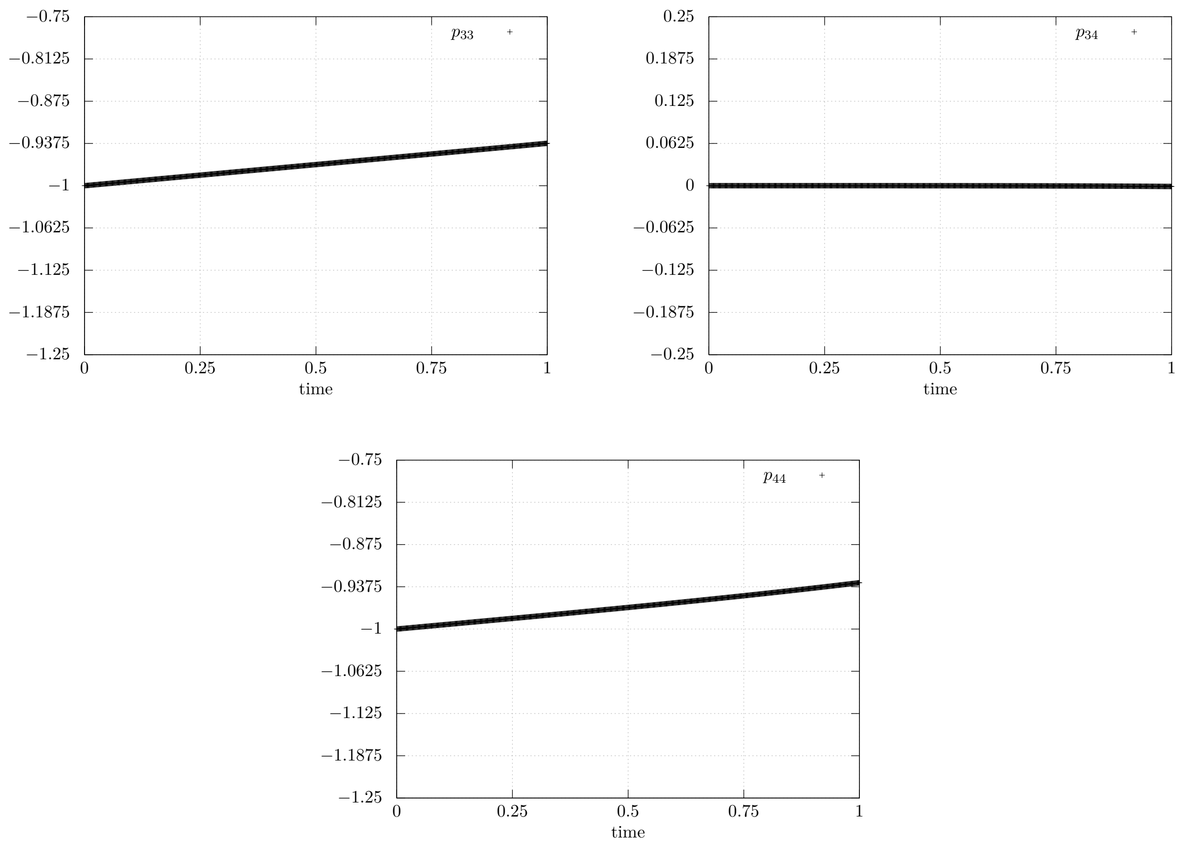

Figure 4.2: Solution of the perturbed Riccati equation for $\varepsilon = \frac{1}{16}$.

Chapter 5

Convergence Analysis

There are numerous methods for numerically solving optimal control problems that split into three classes: direct methods, indirect methods, and function space methods. In this chapter, we focus on the direct approach, which is often used for solving problems in practical applications, since direct discretization methods are robust, user friendly, and are able to deal with difficult problems with control and state constraints (cf. [11]).

In [48, 80], Euler discretizations for problems with mixed control-state constraints are discussed. Herein, [80] achieve convergence of order one in the L_∞-norm with assumptions sufficient for a Lipschitz continuous optimal control, whereas [48] consider controls of bounded variation to obtain a convergence rate of $\frac{1}{p}$ in the L_p-norm. Optimal control problems with pure state constraints of order one are analyzed in [16, 32, 34]. In [32, 34], linear convergence is achieved in the L_2-norm and convergence of order $\frac{2}{3}$ in the L_∞-norm. Via a strengthened Legendre-Clebsch condition, [16] obtain linear convergence in the L_∞-norm. Runge-Kutta methods for problems with set constraints on the control are studied in [33, 53, 117]. Herein, [33, 117] use a second order Runge-Kutta approximation in order to obtain convergence of order two. In [53], convergence of arbitrary order is achieved with a Runge-Kutta scheme of appropriate order and a sufficiently smooth optimal control. Convergence for the value of the objective function is obtained through a control parametrization enhancing technique in [73].

In order to prove convergence, one usually requires similar conditions as in Chapter 4, in particular, regularity of the constraints, controllability, and second-order conditions, e.g., a Legendre-Clebsch condition or a coercivity property of the Hessian of the Lagrange function. Additionally, the optimal control is assumed to be continuous or Lipschitz continuous. Using the techniques developed in [37, 52, 81] and the above conditions, it might be possible to show that a continuous optimal control is actually Lipschitz continuous.

A common strategy to prove convergence for nonlinear problems with smooth optimal control is to compare the KKT-conditions of the continuous problem with the KKT-conditions of the discretized problem. The respective conditions are expressed as generalized equations and an approximation result as in Theorem 2.2.6 is applied.

Convergence for problems with discontinuous controls is discussed in [3–7, 101, 113, 118]. Linear problems are considered in [4, 101, 118]. In [4], convergence of order one in the L_1-norm, and of order $\frac{1}{2}$ in the L_2-norm is shown for the control, if the switching function satisfies a suitable growth condition around its zeros, and the optimal control is of bang-bang type. In [101, 118], a controllability assumption is used to prove convergence of an order depending on the controllability index. Linear quadratic systems are examined in [3, 5, 6, 113]. [3] obtain results similar

to [4]. A L_1 control cost depending on a parameter is augmented in [6]. The result is an optimal control of bang-zero-bang type and linear convergence is obtained, if the switching function has a stable structure. For nonlinear optimal control problems with linearly appearing control and bang-bang solutions convergence is analyzed in [7].
A general convergence theory applicable to approximations of optimal control problems is provided in [116], which was used in [84] to prove convergence of order one for optimal control problems subject to index one DAEs without inequality constraints.
In this chapter, we prove convergence for approximations of optimal control problems subject to index two DAEs with mixed control-state constraints and boundary conditions, therefore generalizing the results in [80, 85]. We use the following scheme:

(a) First, we gather assumptions for the continuous problem in Section 5.1, which are sufficient for the KKT-conditions in Theorem 3.2.5. Furthermore, we introduce a coercivity condition for the Hessian of the Lagrange function.

(b) In Section 5.2, we approximate the optimal control problem with the implicit Euler discretization. We modify the discrete problem such that the associated KKT-conditions are consistent with the KKT-conditions of the continuous problem.

(c) In Section 5.3, we gather properties, which follow from the results in Section 2.4.

(d) We express the respective KKT-conditions as generalized equations in Section 5.4.

(e) In Section 5.5, we apply Theorem 2.2.6 to the generalized equations, which yields a solution of the discrete KKT-conditions that converges linearly to the continuous KKT-point in the L_∞-norm (Theorem 5.5.6).

(f) Finally, in Section 5.7 we establish a relationship between the multipliers of the necessary conditions for the modified discrete problem and the multipliers associated with the directly discretized problem.

These techniques can also be applied to problems with index one DAEs by skipping the modification step in (b).

5.1 Continuous Problem

Consider the problem:

Problem 5.1.1 (Optimal Control Problem with Index Two DAE)
Let $n_x, n_y, n_u, n_\psi, n_c \in \mathbb{N}$ with $n_\psi + n_y \leq 2n_x$, $n_c \leq n_u$. Let

$$\begin{array}{lll} \varphi : \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \to \mathbb{R}, & \psi : \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_\psi}, & \\ f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x}, & g : \mathbb{R}^{n_x} \to \mathbb{R}^{n_y}, & c : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_c}, \end{array}$$

be functions.

$$\begin{array}{lrcll} \textit{Minimize} & \multicolumn{4}{l}{\varphi\left(x\left(0\right), x\left(1\right)\right),} \\ & & & & \\ \textit{with respect to} & \multicolumn{4}{l}{x \in W_{1,\infty}^{n_x}\left(\left[0,1\right]\right), y \in L_\infty^{n_y}\left(\left[0,1\right]\right), u \in L_\infty^{n_u}\left(\left[0,1\right]\right),} \\ & & & & \\ \textit{subject to} & \dot{x}\left(t\right) & = & f\left(x\left(t\right), y\left(t\right), u\left(t\right)\right), & \textit{a.e. in } \left[0,1\right], \\ & \mathbf{0}_{\mathbb{R}^{n_y}} & = & g\left(x\left(t\right)\right), & \textit{in } \left[0,1\right], \\ & \mathbf{0}_{\mathbb{R}^{n_\psi}} & = & \psi\left(x\left(0\right), x\left(1\right)\right), & \\ & \mathbf{0}_{\mathbb{R}^{n_c}} & \geq & c\left(x\left(t\right), y\left(t\right), u\left(t\right)\right), & \textit{a.e. in } \left[0,1\right]. \end{array}$$

Throughout this chapter, we assume the following similar to Section 4.2:

Assumption 5.1.2

(5.A1) *(Existence / Smoothness of a Minimizer)*
Let $(\hat{x}, \hat{y}, \hat{u}) \in W_{2,\infty}^{n_x}([0,1]) \times W_{1,\infty}^{n_y}([0,1]) \times W_{1,\infty}^{n_u}([0,1])$ be a weak local minimizer of Problem 5.1.1.

(5.A2) *(Smoothness of the System)*

(a) φ and ψ are twice continuously differentiable with respect to all arguments and the derivatives are Lipschitz continuous

(b) For a sufficiently large convex compact neighborhood $\mathcal{M}_1$ of

$$\left\{\left(\hat{x}\left(t\right), \hat{y}\left(t\right), \hat{u}\left(t\right)\right) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \mid t \in [0,1]\right\},$$

let the mappings

$$\begin{aligned} (x, y, u) &\mapsto f\left(x, y, u\right), \\ (x, y, u) &\mapsto c\left(x, y, u\right), \end{aligned}$$

be twice continuously differentiable, and the derivatives be Lipschitz continuous in $\mathcal{M}_1$. Furthermore, for a sufficiently large convex compact neighborhood $\mathcal{M}_2$ of

$$\left\{\hat{x}\left(t\right) \in \mathbb{R}^{n_x} \mid t \in [0,1]\right\},$$

let the mapping

$$x \mapsto g(x),$$

be three-times continuously differentiable, and the derivatives be Lipschitz continuous in $\mathcal{M}_2$.

Let $\mathbf{L}$ *denote the maximum over all Lipschitz constants of*

$$f, g, c, \frac{\partial f}{\partial(x,y,u)}, \frac{\partial g}{\partial x}, \frac{\partial c}{\partial(x,y,u)}, \frac{\partial^2 f}{(\partial(x,y,u))^2}, \frac{\partial^2 g}{(\partial x)^2}, \frac{\partial^2 c}{(\partial(x,y,u))^2}, \frac{\partial^3 g}{(\partial x)^3}.$$

Remark 5.1.3

Note that we assume the minimizer $(\hat{x}, \hat{y}, \hat{u})$ *to be in* $W^{n_x}_{2,\infty}([0,1]) \times W^{n_y}_{1,\infty}([0,1]) \times W^{n_u}_{1,\infty}([0,1])$ *instead of* $W^{n_x}_{1,\infty}([0,1]) \times L^{n_y}_{\infty}([0,1]) \times L^{n_u}_{\infty}([0,1])$. *This condition is crucial for the convergence proof in Section 5.5, since we require the derivatives of the systems functions at the minimizer to be Lipschitz continuous. It might be possible to weaken this assumption by using the techniques developed in [37, 52, 81].*

According to Lemma 3.2.3, the constraints of Problem 5.1.1 are equivalent to the reduced system

$$\begin{array}{rcll} \dot{x}(t) & = & f(x(t), y(t), u(t)), & \text{a.e. in } [0,1], \\ \mathbf{0}_{\mathbb{R}^{n_y}} & = & g'(x(t)) f(x(t), y(t), u(t)), & \text{a.e. in } [0,1], \\ \mathbf{0}_{\mathbb{R}^{n_y}} & = & g(x(0)), & \\ \mathbf{0}_{\mathbb{R}^{n_\psi}} & = & \psi(x(0), x(1)), & \\ \mathbf{0}_{\mathbb{R}^{n_c}} & \geq & c(x(t), y(t), u(t)), & \text{a.e. in } [0,1]. \end{array} \tag{5.1.1}$$

For a constant $\alpha \geq 0$ and $t \in [0,1]$ we define the following sets

$$\begin{aligned} J &:= \{1, \ldots, n_c\}, \\ J^\alpha(t) &:= \{j \in J \mid c[t] \geq -\alpha\}, \quad j^\alpha(t) := \operatorname{card}(J^\alpha(t)), \\ \Theta^\alpha_j &:= \{t \in [0,1] \mid j \in J^\alpha(t)\}, \end{aligned}$$

and we abbreviate the derivatives at the minimizer by

$$\begin{array}{lll} A_f(\cdot) := f'_x[\cdot], & B_f(\cdot) := f'_y[\cdot], & C_f(\cdot) := f'_u[\cdot], \\ A_g(\cdot) := g'[\cdot], & & \\ A^g_f(\cdot) := \dot{A}_g(\cdot) + A_g(\cdot) A_f(\cdot), & B^g_f(\cdot) := A_g(\cdot) B_f(\cdot), & C^g_f(\cdot) := A_g(\cdot) C_f(\cdot), \\ E_0 := A_g(0), & \Psi_0 := \psi'_{x_0}(\hat{x}(0), \hat{x}(1)), & \Psi_1 := \psi'_{x_1}(\hat{x}(0), \hat{x}(1)), \\ A_c(\cdot) := c'_x[\cdot], & B_c(\cdot) := c'_y[\cdot], & C_c(\cdot) := c'_u[\cdot], \\ A^\alpha_c(\cdot) := \left[c'_{j,x}[\cdot]\right]_{j \in J^\alpha(\cdot)}, & B^\alpha_c(\cdot) := \left[c'_{j,y}[\cdot]\right]_{j \in J^\alpha(\cdot)}, & C^\alpha_c(\cdot) := \left[c'_{j,u}[\cdot]\right]_{j \in J^\alpha(\cdot)}, \end{array}$$

where we consider $A^\alpha_c(t), B^\alpha_c(t), C^\alpha_c(t)$ to be vacuous, if $J^\alpha(t)$ is empty. With this notation we assume the following:

Assumption 5.1.4 (Linear Independence, Controllability)

(5.A3) *(Index two / Regularity Condition)*
There exist constants $\alpha > 0$ *and* $\beta > 0$ *such that for all* $t \in [0,1]$ *and every* $\varpi \in \mathbb{R}^{n_y + j^{\alpha}(t)}$ *it holds*

$$\left\| \begin{bmatrix} B_f^g(t) & C_f^g(t) \\ B_c^{\alpha}(t) & C_c^{\alpha}(t) \end{bmatrix}^{\top} \varpi \right\| \geq \beta \|\varpi\| .$$

(5.A4) *(Controllability)*
For any $e \in \mathbb{R}^{n_\psi}$ *there exists* $(x, y, u) \in W_{1,\infty}^{n_x}([0,1]) \times L_\infty^{n_y}([0,1]) \times L_\infty^{n_u}([0,1])$ *such that the DAE*

$$\begin{aligned}
\dot{x}(t) &= A_f(t)\,x(t) + B_f(t)\,y(t) + C_f(t)\,u(t), && \text{a.e. in } [0,1], \\
\mathbf{0}_{\mathbb{R}^{n_y}} &= A_f^g(t)\,x(t) + B_f^g(t)\,y(t) + C_f^g(t)\,u(t), && \text{a.e. in } [0,1], \\
e &= \begin{bmatrix} E_0 \\ \Psi_0 \end{bmatrix} x(0) + \begin{bmatrix} \mathbf{0}_{n_y \times n_x} \\ \Psi_1 \end{bmatrix} x(1), && \\
\mathbf{0}_{\mathbb{R}^{j^{\alpha}(t)}} &= A_c^{\alpha}(t)\,x(t) + B_c^{\alpha}(t)\,y(t) + C_c^{\alpha}(t)\,u(t), && \text{a.e. in } [0,1]
\end{aligned}$$

is satisfied.

Please note that we coupled the index property of the DAE with the regularity of the inequality constraint (compare Remark 3.1.4). If (5.A1) - (5.A4) hold, then, by Theorem 3.2.5, there exist multipliers

$$\hat{\ell}_0 \in \mathbb{R}, \hat{\lambda}_f \in W_{1,\infty}^{n_x}([0,1]), \hat{\lambda}_g \in L_\infty^{n_y}([0,1]), \hat{\sigma} \in \mathbb{R}^{n_\psi}, \hat{\varsigma} \in \mathbb{R}^{n_y}, \hat{\eta} \in L_\infty^{n_c}([0,1])$$

associated with the weak local minimizer $(\hat{x}, \hat{y}, \hat{u})$ of Problem 5.1.1 such that $\hat{\ell}_0 = 1$ and

$$\begin{aligned}
\dot{\hat{\lambda}}_f(t) &= -\nabla_x \mathcal{H}\left(\hat{x}(t), \hat{y}(t), \hat{u}(t), \hat{\lambda}_f(t), \hat{\lambda}_g(t), \hat{\eta}(t)\right), && \text{a.e. in } [0,1], \\
\mathbf{0}_{\mathbb{R}^{n_y}} &= \nabla_y \mathcal{H}\left(\hat{x}(t), \hat{y}(t), \hat{u}(t), \hat{\lambda}_f(t), \hat{\lambda}_g(t), \hat{\eta}(t)\right), && \text{a.e. in } [0,1], \\
\mathbf{0}_{\mathbb{R}^{n_u}} &= \nabla_u \mathcal{H}\left(\hat{x}(t), \hat{y}(t), \hat{u}(t), \hat{\lambda}_f(t), \hat{\lambda}_g(t), \hat{\eta}(t)\right), && \text{a.e. in } [0,1], \\
\hat{\lambda}_f(0) &= -\nabla_{x_0} \varphi(\hat{x}(0), \hat{x}(1)) - \psi_{x_0}'(\hat{x}(0), \hat{x}(1))^{\top} \hat{\sigma} - g'(\hat{x}(0))^{\top} \hat{\varsigma}, && \\
\hat{\lambda}_f(1) &= \nabla_{x_1} \varphi(\hat{x}(0), \hat{x}(1)) + \psi_{x_1}'(\hat{x}(0), \hat{x}(1))^{\top} \hat{\sigma}, && \\
0 &= \hat{\eta}(t)^{\top} c(\hat{x}(t), \hat{y}(t), \hat{u}(t)), \quad \hat{\eta}(t) \geq \mathbf{0}_{\mathbb{R}^{n_c}}, && \text{a.e. in } [0,1]
\end{aligned} \tag{5.1.2}$$

is satisfied. Herein, the (augmented) Hamilton function is defined by

$$\begin{aligned}
&\mathcal{H} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_c} \to \mathbb{R}, \\
&\mathcal{H}(x, y, u, \lambda_f, \lambda_g, \eta) := \lambda_f^{\top} f(x, y, u) + \lambda_g^{\top} g'(x) f(x, y, u) + \eta^{\top} c(x, y, u).
\end{aligned}$$

Analog to Lemma 4.1.4, we can show that

$$\hat{\lambda}_f \in W^{n_x}_{2,\infty}([0,1]), \quad \hat{\lambda}_g \in W^{n_y}_{1,\infty}([0,1]), \quad \hat{\eta} \in W^{n_c}_{1,\infty}([0,1]),$$

if (5.A1) - (5.A3) hold. Moreover, for a constant $\nu \geq 0$ and $t \in [0,1]$ we define

$$\begin{aligned}
&J^\nu_+(t) := \left\{ j \in J^0(t) \mid \hat{\eta}(t) > \nu \right\}, \quad j^\nu_+(t) := \operatorname{card}\left(J^\nu_+(t)\right), \quad \Upsilon^\nu_j := \left\{ t \in [0,1] \mid j \in J^\nu_+(t) \right\},\\
&\check{A}^\nu_c(t) := \left[c'_{j,x}[t] \right]_{j \in J^\nu_+(t)}, \quad \check{B}^\nu_c(t) := \left[c'_{j,y}[t] \right]_{j \in J^\nu_+(t)}, \quad \check{C}^\nu_c(t) := \left[c'_{j,u}[t] \right]_{j \in J^\nu_+(t)},
\end{aligned}$$

where we consider $\check{A}^\nu_c(t), \check{B}^\nu_c(t), \check{C}^\nu_c(t)$ to be vacuous, if $J^\nu_+(t)$ is empty, and for the functional $\vartheta(x_0, x_1, \sigma, \varsigma) := \varphi(x_0, x_1) + \sigma^\top \psi(x_0, x_1) + \varsigma^\top g(x_0)$ we denote

$$\begin{aligned}
&\Lambda_{00} := \nabla^2_{x_0 x_0} \vartheta(\hat{x}(0), \hat{x}(1), \hat{\sigma}, \hat{\varsigma}), \quad \Lambda_{01} := \nabla^2_{x_0 x_1} \vartheta(\hat{x}(0), \hat{x}(1), \hat{\sigma}, \hat{\varsigma}),\\
&\Lambda_{11} := \nabla^2_{x_1 x_1} \vartheta(\hat{x}(0), \hat{x}(1), \hat{\sigma}, \hat{\varsigma}).
\end{aligned}$$

For the space

$$X_2 := W^{n_x}_{1,2}([0,1]) \times L^{n_y}_2([0,1]) \times L^{n_u}_2([0,1])$$

equipped with the norm $\|(x,y,u)\|_{X_2} := \max\left\{ \|x\|_{1,2}, \|y\|_2, \|u\|_2 \right\}$ we define the symmetric bilinear form

$$\begin{aligned}
&\mathcal{P} : X_2 \times X_2 \to \mathbb{R},\\
&\mathcal{P}\left(\left(x^1, y^1, u^1\right), \left(x^2, y^2, u^2\right)\right) := \begin{pmatrix} x^1(0) \\ x^1(1) \end{pmatrix}^\top \begin{bmatrix} \Lambda_{00} & \Lambda_{01} \\ \Lambda_{01}^\top & \Lambda_{11} \end{bmatrix} \begin{pmatrix} x^2(0) \\ x^2(1) \end{pmatrix} \qquad (5.1.3)\\
&+ \int_0^1 \begin{pmatrix} x^1(t) \\ y^1(t) \\ u^1(t) \end{pmatrix}^\top \begin{bmatrix} \nabla^2_{xx}\mathcal{H}[t] & \nabla^2_{xy}\mathcal{H}[t] & \nabla^2_{xu}\mathcal{H}[t] \\ \nabla^2_{yx}\mathcal{H}[t] & \nabla^2_{yy}\mathcal{H}[t] & \nabla^2_{yu}\mathcal{H}[t] \\ \nabla^2_{ux}\mathcal{H}[t] & \nabla^2_{uy}\mathcal{H}[t] & \nabla^2_{uu}\mathcal{H}[t] \end{bmatrix} \begin{pmatrix} x^2(t) \\ y^2(t) \\ u^2(t) \end{pmatrix} dt,
\end{aligned}$$

which is continuous, since $\nabla^2_{(x,y,u)(x,y,u)}\mathcal{H}[t]$ is bounded. We assume the following uniform coercivity condition:

Assumption 5.1.5 (Coercivity)

(5.A5) *(Coercivity)*

There exist constants $\nu > 0$ and $\gamma > 0$ such that for every $(x,y,u) \in X_2$, which satisfies

$$\begin{aligned}
\dot{x}(t) &= A_f(t)x(t) + B_f(t)y(t) + C_f(t)u(t), && a.e.\ in\ [0,1],\\
\mathbf{0}_{\mathbb{R}^{n_y}} &= A^g_f(t)x(t) + B^g_f(t)y(t) + C^g_f(t)u(t), && a.e.\ in\ [0,1],\\
\mathbf{0}_{\mathbb{R}^{n_y}} &= E_0 x(0),\\
\mathbf{0}_{\mathbb{R}^{n_\psi}} &= \Psi_0 x(0) + \Psi_1 x(1),\\
\mathbf{0}_{\mathbb{R}^{j^\nu_+(t)}} &= \check{A}^\nu_c(t)x(t) + \check{B}^\nu_c(t)y(t) + \check{C}^\nu_c(t)u(t), && a.e.\ in\ [0,1],
\end{aligned}$$

it holds

$$\mathcal{P}((x,y,u),(x,y,u)) \geq \gamma \|(x,y,u)\|^2_{X_2}.$$

5.2 Discrete Problem

For $N \in \mathbb{N}$ let $\mathbb{G}_N := \{0 = t_0 < t_1 < t_2 < \ldots < t_{N-1} < t_N = 1\}$ be a grid of $[0,1]$ with $t_i := ih$, $i = 0, 1, \ldots, N$, and the mesh size $h := \frac{1}{N}$. We consider the following direct discretization of Problem 5.1.1:

Problem 5.2.1 (Discrete Optimal Control Problem with Index Two DAE)

$$\begin{array}{lllll}
\textit{Minimize} & \varphi\left(x_0, x_N\right), & & & \\
& & & & \\
\textit{with respect to} & \multicolumn{4}{l}{x_0 \in \mathbb{R}^{n_x}, \left(x_i, y_i, u_i\right) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u}, i = 1, \ldots, N,} \\
& & & & \\
\textit{subject to} & \frac{x_i - x_{i-1}}{h} & = & f\left(x_i, y_i, u_i\right), & i = 1, \ldots, N, \\
& \mathbf{0}_{\mathbb{R}^{n_y}} & = & g\left(x_i\right), & i = 0, 1, \ldots, N, \\
& \mathbf{0}_{\mathbb{R}^{n_\psi}} & = & \psi\left(x_0, x_N\right), & \\
& \mathbf{0}_{\mathbb{R}^{n_c}} & \geq & c\left(x_i, y_i, u_i\right), & i = 1, \ldots, N.
\end{array}$$

For the (augmented) Hamilton function

$$\begin{aligned}
&\breve{\mathcal{H}} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_c} \to \mathbb{R} \\
&\breve{\mathcal{H}}\left(x, y, u, \breve{\lambda}_f, \breve{\lambda}_g, \breve{\eta}\right) := \breve{\lambda}_f^\top f\left(x, y, u\right) + \breve{\lambda}_g^\top g\left(x\right) + \breve{\eta}^\top c\left(x, y, u\right),
\end{aligned} \tag{5.2.1}$$

and multipliers

$$\begin{array}{c}
\breve{\ell}_0 \in \mathbb{R}, \breve{\sigma} \in \mathbb{R}^{n_\psi}, \breve{\varsigma} \in \mathbb{R}^{n_y}, \\
\begin{array}{rcll}
\breve{\lambda}_{f,i} & \in & \mathbb{R}^{n_x}, & i = 0, 1, \ldots, N, \\
\breve{\lambda}_{g,i} & \in & \mathbb{R}^{n_y}, & i = 1, \ldots, N, \\
\breve{\eta}_i & \in & \mathbb{R}^{n_c}, & i = 1, \ldots, N,
\end{array}
\end{array}$$

we get the following necessary conditions

$$\begin{array}{rcll}
\frac{\breve{\lambda}_{f,i} - \breve{\lambda}_{f,i-1}}{h} & = & -\nabla_x \breve{\mathcal{H}}\left(x_i, y_i, u_i, \breve{\lambda}_{f,i-1}, \breve{\lambda}_{g,i-1}, \breve{\eta}_i\right), & i = 1, \ldots, N, \\
\mathbf{0}_{\mathbb{R}^{n_y}} & = & \nabla_y \breve{\mathcal{H}}\left(x_i, y_i, u_i, \breve{\lambda}_{f,i-1}, \breve{\lambda}_{g,i-1}, \breve{\eta}_i\right), & i = 1, \ldots, N, \\
\breve{\lambda}_{f,0} & = & -\breve{\ell}_0 \nabla_{x_0} \varphi\left(x_0, x_N\right) - \psi'_{x_0}\left(x_0, x_N\right)^\top \breve{\sigma} - g'\left(x_0\right)^\top \breve{\varsigma}, & \\
\breve{\lambda}_{f,N} & = & \breve{\ell}_0 \nabla_{x_1} \varphi\left(x_0, x_N\right) + \psi'_{x_1}\left(x_0, x_N\right)^\top \breve{\sigma}, & \\
\mathbf{0}_{\mathbb{R}^{n_u}} & = & \nabla_u \breve{\mathcal{H}}\left(x_i, y_i, u_i, \breve{\lambda}_{f,i-1}, \breve{\lambda}_{g,i-1}, \breve{\eta}_i\right), & i = 1, \ldots, N, \\
0 & = & c\left(x_i, y_i, u_i\right)^\top \breve{\eta}_i, \quad \breve{\eta}_i \geq 0, & i = 1, \ldots, N,
\end{array} \tag{5.2.2}$$

cf. [47, Theorem 5.4.4].

Remark 5.2.2

Please note that the discrete necessary conditions hold for the Hamilton function $\breve{\mathcal{H}}$ *defined in* (5.2.1) *(with an index shift) instead of*

$$\mathcal{H}\left(x, y, u, \lambda_f, \lambda_g, \eta\right) = \lambda_f^\top f\left(x, y, u\right) + \lambda_g^\top g'\left(x\right) f\left(x, y, u\right) + \eta^\top c\left(x, y, u\right),$$

as in the continuous case. This leads to a discrepancy between the respective necessary conditions, since the continuous necessary conditions with the Hamilton function $\breve{\mathcal{H}}$ *do not have a solution* *in general (cf. [8, Example 3.16]). Therefore, the respective KKT-conditions are not consistent with each other (compare (ii) in Theorem 2.2.6), i.e., the error that arises from inserting the continuous KKT-point into the discrete KKT-conditions does not converge to zero, if the mesh size h tends to zero. This discrepancy can be overcome by exhibiting extra assumptions such that the continuous KKT-conditions with Hamilton function $\breve{\mathcal{H}}$ defined in (5.2.1) have a solution. This permitted us to show that the respective KKT-conditions with the Hamilton function $\breve{\mathcal{H}}$ are consistent with each other. However, this approach failed in the attempt to verify uniform strong regularity (compare Definition 2.2.4). Specifically, we were unable to prove the step in Lemma 5.5.4. Thus, a different strategy was required.*

As illustrated in Section 3.2, the necessary conditions of optimal control problems with higher index DAEs actually coincide with the necessary conditions of the index reduced problem. In order to obtain suitable discrete necessary conditions, we emulate the index reduction idea of the continuous case, i.e., replacing the algebraic constraint $\mathbf{0}_{\mathbb{R}^{n_y}} = g(x(t))$ in $[0,1]$ with the hidden constraint $\mathbf{0}_{\mathbb{R}^{n_y}} = g'(x(t)) f(x(t), y(t), u(t))$ almost everywhere in $[0,1]$, and the extra initial condition $\mathbf{0}_{\mathbb{R}^{n_y}} = g(x(0))$. To that end, we consider the discrete algebraic constraint

$$\mathbf{0}_{\mathbb{R}^{n_y}} = g(x_i), \quad i = 0, 1, \ldots, N,$$

and replace it with a discrete derivative, in particular, the backwards difference approximation

$$\mathbf{0}_{\mathbb{R}^{n_y}} = \frac{g(x_i) - g(x_{i-1})}{h}, \quad i = 1, \ldots, N,$$

together with the initial condition $\mathbf{0}_{\mathbb{R}^{n_y}} = g(x_0)$. In addition, we solve the difference equation

$$\frac{x_i - x_{i-1}}{h} = f(x_i, y_i, u_i)$$

for x_{i-1}, which yields $x_{i-1} = x_i - hf(x_i, y_i, u_i)$, and insert it into the backwards difference approximation, which results in

$$\mathbf{0}_{\mathbb{R}^{n_y}} = \frac{g(x_i) - g(x_i - hf(x_i, y_i, u_i))}{h}, \quad i = 1, \ldots, N.$$

With the notation $\tilde{g}_h(x, y, u) := \frac{g(x) - g(x - hf(x,y,u))}{h}$ we get the reduced problem:

Problem 5.2.3 (Discrete Optimal Control Problem with Reduced DAE)

$$\begin{array}{lrcll}
\textit{Minimize} & \multicolumn{4}{l}{\varphi(x_0, x_N),} \\
\textit{with respect to} & \multicolumn{4}{l}{x_0 \in \mathbb{R}^{n_x}, (x_i, y_i, u_i) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u}, i = 1, \ldots, N,} \\
\textit{subject to} & \frac{x_i - x_{i-1}}{h} & = & f(x_i, y_i, u_i), & i = 1, \ldots, N, \\
 & \mathbf{0}_{\mathbb{R}^{n_y}} & = & \tilde{g}_h(x_i, y_i, u_i), & i = 1, \ldots, N, \\
 & \mathbf{0}_{\mathbb{R}^{n_y}} & = & g(x_0), & \\
 & \mathbf{0}_{\mathbb{R}^{n_\psi}} & = & \psi(x_0, x_N), & \\
 & \mathbf{0}_{\mathbb{R}^{n_c}} & \geq & c(x_i, y_i, u_i), & i = 1, \ldots, N.
\end{array}$$

Note that Problem 5.2.1 and Problem 5.2.3 are equivalent, since

$$\begin{array}{rcll} \mathbf{0}_{\mathbb{R}^{n_y}} &=& g\left(x_i\right), & \text{for } i=0,1,\ldots,N, \\ \Leftrightarrow\quad \mathbf{0}_{\mathbb{R}^{n_y}} &=& \frac{g(x_i)-g(x_{i-1})}{h}, & \text{for } i=1,\ldots,N, \quad \text{and} \quad \mathbf{0}_{\mathbb{R}^{n_y}} = g\left(x_0\right), \end{array}$$

and

$$\begin{array}{rcll} \frac{x_i-x_{i-1}}{h} &=& f\left(x_i,y_i,u_i\right), & \text{for } i=1,\ldots,N, \\ \Leftrightarrow\quad x_{i-1} &=& x_i - hf\left(x_i,y_i,u_i\right), & \text{for } i=1,\ldots,N. \end{array}$$

Therefore, the difference-algebraic equations of the respective problems are equivalent. We define the discrete Hamilton function by

$$\begin{aligned} &\widetilde{\mathcal{H}}_h : \mathbb{R}^{n_x}\times\mathbb{R}^{n_y}\times\mathbb{R}^{n_u}\times\mathbb{R}^{n_x}\times\mathbb{R}^{n_y}\times\mathbb{R}^{n_c}\to\mathbb{R}, \\ &\widetilde{\mathcal{H}}_h\left(x,y,u,\lambda_f,\lambda_g,\eta\right) := \lambda_f^\top f\left(x,y,u\right) + \lambda_g^\top \tilde{g}_h\left(x,y,u\right) + \eta^\top c\left(x,y,u\right). \end{aligned} \tag{5.2.3}$$

Then, with multipliers

$$\ell_0\in\mathbb{R}, \sigma\in\mathbb{R}^{n_\psi}, \varsigma\in\mathbb{R}^{n_y},$$

$$\begin{array}{rcll} \lambda_{f,i} &\in& \mathbb{R}^{n_x} & i=0,1,\ldots,N, \\ \lambda_{g,i} &\in& \mathbb{R}^{n_y} & i=1,\ldots,N, \\ \eta_i &\in& \mathbb{R}^{n_c} & i=1,\ldots,N, \end{array}$$

the necessary conditions of Problem 5.2.3 can be expressed as follows (cf. [47, Theorem 5.4.4]):

$$\begin{array}{rcll} \frac{\lambda_{f,i}-\lambda_{f,i-1}}{h} &=& -\nabla_x\widetilde{\mathcal{H}}_h\left(x_i,y_i,u_i,\lambda_{f,i-1},\lambda_{g,i-1},\eta_i\right), & i=1,\ldots,N, \\ \mathbf{0}_{\mathbb{R}^{n_y}} &=& \nabla_y\widetilde{\mathcal{H}}_h\left(x_i,y_i,u_i,\lambda_{f,i-1},\lambda_{g,i-1},\eta_i\right), & i=1,\ldots,N, \\ \lambda_{f,0} &=& -\ell_0\nabla_{x_0}\varphi\left(x_0,x_N\right) - \psi'_{x_0}\left(x_0,x_N\right)^\top\sigma - g'\left(x_0\right)^\top\varsigma, & \\ \lambda_{f,N} &=& \ell_0\nabla_{x_1}\varphi\left(x_0,x_N\right) + \psi'_{x_1}\left(x_0,x_N\right)^\top\sigma, & \\ \mathbf{0}_{\mathbb{R}^{n_u}} &=& \nabla_u\widetilde{\mathcal{H}}_h\left(x_i,y_i,u_i,\lambda_{f,i-1},\lambda_{g,i-1},\eta_i\right), & i=1,\ldots,N, \\ 0 &=& c\left(x_i,y_i,u_i\right)^\top\eta_i, \quad \eta_i\geq 0, & i=1,\ldots,N. \end{array} \tag{5.2.4}$$

The multipliers, which satisfy (5.2.2) and (5.2.4), respectively, are not equal in general. In Section 5.7 we derive a relationship between the respective multipliers.

Remark 5.2.4
Another strategy for deriving suitable discrete KKT-conditions would be to first reduce the continuous DAE in Problem 5.1.1 to an index one DAE and then apply a discretization scheme. However, this approach has several drawbacks from a practical point of view, where the dynamics are often automatically produced by software packages. Thus, an index reduction can only be done numerically. Furthermore, since there is no process to enforce that the discretized reduced system satisfies the algebraic constraint $\mathbf{0}_{\mathbb{R}^{n_y}} = g\left(x\left(t\right)\right)$*, the so-called drift-off effect might occur (cf. [19, 54]).*

The main goal in this chapter is to prove that Problem 5.2.3 has a solution with associated multipliers satisfying (5.2.4), which converge towards the weak local minimizer $(\hat{x},\hat{y},\hat{u})$ and its associated Lagrange multipliers, respectively. To that end, we first write the respective KKT-conditions as generalized equations in Section 5.4. Then, we verify the conditions of Theorem 2.2.6 for these equations in Section 5.5. Thus, we obtain a solution of the discrete KKT-conditions, which converges linearly to the solution of the continuous KKT-conditions.

5.3 Preparations

In the same way as in Section 2.4, we consider finite dimensional subspaces of $L_p^n([0,1])$ and $W_{1,p}^n([0,1])$ defined as

$$\begin{aligned} L_{p,h}^n([0,1]) &:= \Big\{v \in L_p^n([0,1]) \mid v(t) = v(t_i),\, t \in (t_{i-1}, t_i],\, i = 1, \ldots, N\Big\}, \\ W_{1,p,h}^n([0,1]) &:= \Big\{v \in W_{1,p}^n([0,1]) \mid v(t) = v'(t_i)(t - t_{i-1}) + v(t_{i-1}), \\ &\qquad t \in (t_{i-1}, t_i],\, i = 1, \ldots, N\}, \end{aligned}$$

for $p = 2, \infty$, where $v'(t_i) := \frac{v(t_i) - v(t_{i-1})}{h}$ denotes the backwards difference approximation. This allows us to write the constraints of Problem 5.2.3 as

$$\begin{aligned} x_h'(t_i) &= f(x_h(t_i), y_h(t_i), u_h(t_i)), && i = 1, \ldots, N, \\ \mathbf{0}_{\mathbb{R}^{n_y}} &= \tilde{g}_h(x_h(t_i), y_h(t_i), u_h(t_i)), && i = 1, \ldots, N, \\ \mathbf{0}_{\mathbb{R}^{n_y}} &= g(x_h(t_0)), \\ \mathbf{0}_{\mathbb{R}^{n_\psi}} &= \psi(x_h(t_0), x_h(t_N)), \\ \mathbf{0}_{\mathbb{R}^{n_c}} &\geq c(x_h(t_i), y_h(t_i), u_h(t_i)), && i = 1, \ldots, N, \end{aligned}$$

for $(x_h, y_h, u_h) \in W_{1,\infty,h}^{n_x}([0,1]) \times L_{\infty,h}^{n_y}([0,1]) \times L_{\infty,h}^{n_u}([0,1])$. Let us abbreviate the derivatives of $\tilde{g}_h$ at $(\hat{x}, \hat{y}, \hat{u})$ by

$$\tilde{A}_{f,h}^g(\cdot) := \tilde{g}_{h,x}'[\cdot], \quad \tilde{B}_{f,h}^g(\cdot) := \tilde{g}_{h,y}'[\cdot], \quad \tilde{C}_{f,h}^g(\cdot) := \tilde{g}_{h,u}'[\cdot].$$

We introduce the spaces and subspaces

$$\begin{aligned} X_p &:= W_{1,p}^{n_x}([0,1]) \times L_p^{n_y}([0,1]) \times L_p^{n_u}([0,1]), \\ X_{p,h} &:= W_{1,p,h}^{n_x}([0,1]) \times L_{p,h}^{n_y}([0,1]) \times L_{p,h}^{n_u}([0,1]), \\ Y_p^\alpha &:= L_p^{n_x}([0,1]) \times L_p^{n_y}([0,1]) \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_\psi} \times L_p^{n_c}(\Theta^\alpha), \\ Y_{p,h}^\alpha &:= L_{p,h}^{n_x}([0,1]) \times L_{p,h}^{n_y}([0,1]) \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_\psi} \times L_{p,h}^{n_c}(\Theta^\alpha), \\ \check{Y}_p^\nu &:= L_p^{n_x}([0,1]) \times L_p^{n_y}([0,1]) \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_\psi} \times L_p^{n_c}(\Upsilon^\nu), \\ \check{Y}_{p,h}^\nu &:= L_{p,h}^{n_x}([0,1]) \times L_{p,h}^{n_y}([0,1]) \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_\psi} \times L_{p,h}^{n_c}(\Upsilon^\nu), \\ L_p^{n_c}(\Theta^\alpha) &:= \bigtimes_{j \in J} L_p\left(\Theta_j^\alpha\right), \quad L_p^{n_c}(\Upsilon^\nu) := \bigtimes_{j \in J} L_p\left(\Upsilon_j^\nu\right), \\ L_{p,h}^{n_c}(\Theta^\alpha) &:= \bigtimes_{j \in J} L_{p,h}\left(\Theta_j^\alpha\right), \quad L_{p,h}^{n_c}(\Upsilon^\nu) := \bigtimes_{j \in J} L_{p,h}\left(\Upsilon_j^\nu\right) \end{aligned}$$

equipped with the norms

$$\begin{aligned} \|(x,y,u)\|_{X_p} &:= \max\left\{\|x\|_{1,p}, \|y\|_p, \|u\|_p\right\}, \\ \left\|\left(a_f, a_f^g, a_E, a_\Psi, a_c^\alpha\right)\right\|_{Y_p^\alpha} &:= \max\left\{\|a_f\|_p, \left\|a_f^g\right\|_p, \|a_E\|, \|a_\Psi\|, \|a_c^\alpha\|_p\right\}, \\ \left\|\left(a_f, a_f^g, a_E, a_\Psi, \check{a}_c^\nu\right)\right\|_{\check{Y}_p^\nu} &:= \max\left\{\|a_f\|_p, \left\|a_f^g\right\|_p, \|a_E\|, \|a_\Psi\|, \|\check{a}_c^\nu\|_p\right\}. \end{aligned}$$

For these spaces we define the linear operators

$$\begin{array}{rclrcl} F^{\alpha} & \in & \mathfrak{L}\left(X_{\infty}, Y_{\infty}^{\alpha}\right) \cap \mathfrak{L}\left(X_2, Y_2^{\alpha}\right), & \check{F}^{\nu} & \in & \mathfrak{L}\left(X_{\infty}, \check{Y}_{\infty}^{\nu}\right) \cap \mathfrak{L}\left(X_2, \check{Y}_2^{\nu}\right), \\ F_h^{\alpha} & \in & \mathfrak{L}\left(X_{\infty,h}, Y_{\infty,h}^{\alpha}\right) \cap \mathfrak{L}\left(X_{2,h}, Y_{2,h}^{\alpha}\right), & \check{F}_h^{\nu} & \in & \mathfrak{L}\left(X_{\infty,h}, \check{Y}_{\infty,h}^{\nu}\right) \cap \mathfrak{L}\left(X_{2,h}, \check{Y}_{2,h}^{\nu}\right) \end{array}$$

by

$$F^{\alpha}(x,y,u) := \begin{pmatrix} \dot{x}(\cdot) - A_f(\cdot)x(\cdot) - B_f(\cdot)y(\cdot) - C_f(\cdot)u(\cdot) \\ A_f^g(\cdot)x(\cdot) + B_f^g(\cdot)y(\cdot) + C_f^g(\cdot)u(\cdot) \\ E_0 x(0) \\ \Psi_0 x(0) + \Psi_1 x(1) \\ A_c^{\alpha}(\cdot)x(\cdot) + B_c^{\alpha}(\cdot)y(\cdot) + C_c^{\alpha}(\cdot)u(\cdot) \end{pmatrix},$$

$$\check{F}^{\nu}(x,y,u) := \begin{pmatrix} \dot{x}(\cdot) - A_f(\cdot)x(\cdot) - B_f(\cdot)y(\cdot) - C_f(\cdot)u(\cdot) \\ A_f^g(\cdot)x(\cdot) + B_f^g(\cdot)y(\cdot) + C_f^g(\cdot)u(\cdot) \\ E_0 x(0) \\ \Psi_0 x(0) + \Psi_1 x(1) \\ \check{A}_c^{\nu}(\cdot)x(\cdot) + \check{B}_c^{\nu}(\cdot)y(\cdot) + \check{C}_c^{\nu}(\cdot)u(\cdot) \end{pmatrix},$$

$$F_h^{\alpha}(x_h,y_h,u_h)(t) := \begin{pmatrix} x_h'(t_i) - A_f(t_i)x_h(t_i) - B_f(t_i)y_h(t_i) - C_f(t_i)u_h(t_i) \\ \tilde{A}_{f,h}^g(t_i)x_h(t_i) + \tilde{B}_{f,h}^g(t_i)y_h(t_i) + \tilde{C}_{f,h}^g(t_i)u_h(t_i) \\ E_0 x_h(t_0) \\ \Psi_0 x_h(t_0) + \Psi_1 x_h(t_N) \\ A_c^{\alpha}(t_i)x_h(t_i) + B_c^{\alpha}(t_i)y_h(t_i) + C_c^{\alpha}(t_i)u_h(t_i) \end{pmatrix},$$

$$\check{F}_h^{\nu}(x_h,y_h,u_h)(t) := \begin{pmatrix} x_h'(t_i) - A_f(t_i)x_h(t_i) - B_f(t_i)y_h(t_i) - C_f(t_i)u_h(t_i) \\ \tilde{A}_{f,h}^g(t_i)x_h(t_i) + \tilde{B}_{f,h}^g(t_i)y_h(t_i) + \tilde{C}_{f,h}^g(t_i)u_h(t_i) \\ E_0 x_h(t_0) \\ \Psi_0 x_h(t_0) + \Psi_1 x_h(t_N) \\ \check{A}_c^{\nu}(t_i)x_h(t_i) + \check{B}_c^{\nu}(t_i)y_h(t_i) + \check{C}_c^{\nu}(t_i)u_h(t_i) \end{pmatrix},$$

for $t \in (t_{i-1}, t_i]$, $i = 1, \ldots, N$, which satisfy the following according to Lemma 2.4.10:

Lemma 5.3.1 (Surjectivity Conditions)
Let (5.A1) - (5.A4) hold. Then, F^{α} and $\check{F}^{\nu}$ are uniformly surjective with some constant $\kappa > 0$. Furthermore, there exists a $h_1 > 0$ such that for all $0 < h \leq h_1$ the operators F_h^{α} and $\check{F}_h^{\nu}$ are uniformly surjective with some constant $\tilde{\kappa} > 0$ independent of h.

Proof. Since $J_+^{\nu}(t) \subseteq J^{\alpha}(t)$ for all $t \in [0,1]$, assumption (5.A3) implies

$$\left\| \begin{bmatrix} B_f^g(t) & C_f^g(t) \\ \check{B}_c^{\nu}(t) & \check{C}_c^{\nu}(t) \end{bmatrix}^{\top} \varpi \right\| \geq \beta \left\|\varpi\right\|$$

for all $t \in [0,1]$ and every $\varpi \in \mathbb{R}^{n_y + j_+^\nu(t)}$. Thus, the linear system

$$\begin{aligned}
\dot{x}(t) &= A_f(t)\,x(t) + B_f(t)\,y(t) + C_f(t)\,u(t), && \text{a.e. in } [0,1], \\
\mathbf{0}_{\mathbb{R}^{n_y}} &= A_f^g(t)\,x(t) + B_f^g(t)\,y(t) + C_f^g(t)\,u(t), && \text{a.e. in } [0,1], \\
\mathbf{0}_{\mathbb{R}^{n_y}\times\mathbb{R}^{n_\psi}} &= \begin{bmatrix} E_0 \\ \Psi_0 \end{bmatrix} x(0) + \begin{bmatrix} \mathbf{0}_{n_y \times n_x} \\ \Psi_1 \end{bmatrix} x(1), && \\
\mathbf{0}_{\mathbb{R}^{j_+^\nu(t)}} &= \check{A}_c^\nu(t)\,x(t) + \check{B}_c^\nu(t)\,y(t) + \check{C}_c^\nu(t)\,u(t), && \text{a.e. in } [0,1]
\end{aligned}$$

is completely controllable. In order to apply Lemma 2.4.10, it remains to show that the conditions in Assumption 2.4.4 hold. To that end, for an arbitrary $i \in \{1,\ldots,N\}$ and $t \in (t_{i-1}, t_i]$ using the mean-value theorem in [59, p. 40] yields

$$\begin{aligned}
\left\|A_f^g(t) - \tilde{A}_{f,h}^g(t_i)\right\| &= \left\|\frac{d}{dt}\left(g'[t]\right) + g'[t]\,A_f(t) - \frac{g'[t_i] - g'\left(\hat{x}(t_i) - hf[t_i]\right)}{h}\right. \\
&\qquad \left. - g'\left(\hat{x}(t_i) - hf[t_i]\right)A_f(t_i)\right\| \\
&\le \left\| g''\left(\hat{x}(t)\right) f[t] - \int_0^1 g''\left(\hat{x}(t_i) - \theta h f[t_i]\right) f[t_i]\, d\theta \right\| \\
&\qquad + \left\|g'[t] - g'\left(\hat{x}(t_i) - hf[t_i]\right)\right\| \left\|A_f\right\|_\infty + \left\|g'[\cdot]\right\|_\infty \left\|A_f(t) - A_f(t_i)\right\| \\
&\le \left\| \int_0^1 g''\left(\hat{x}(t)\right) - g''\left(\hat{x}(t_i) - \theta h f[t_i]\right) d\theta \right\| \left\|f[\cdot]\right\|_\infty + \left\|g''[\cdot]\right\|_\infty \left\|f[t] - f[t_i]\right\| \\
&\qquad + \mathbf{L}\left(\left\|f[\cdot]\right\|_\infty \left\|A_f\right\|_\infty + \left\|g'[\cdot]\right\|_\infty\right) h \\
&\le \mathbf{L}\left(\left\|f[\cdot]\right\|_\infty^2 + \left\|g''[\cdot]\right\|_\infty + \left\|f[\cdot]\right\|_\infty \left\|A_f\right\|_\infty + \left\|g'[\cdot]\right\|_\infty\right) h
\end{aligned}$$

Moreover, for all $t \in (t_{i-1}, t_i]$ it holds

$$\begin{aligned}
\left\|B_f^g(t) - \tilde{B}_{f,h}^g(t_i)\right\| &= \left\|g'[t]\,B_f(t) - g'\left(\hat{x}(t_i) - hf[t_i]\right) B_f(t_i)\right\| \\
&\le \left\|g'[t] - g'\left(\hat{x}(t_i) - hf[t_i]\right)\right\| \left\|B_f\right\|_\infty + \left\|g'[\cdot]\right\|_\infty \left\|B_f(t) - B_f(t_i)\right\| \\
&\le \mathbf{L}\left(\left\|f[\cdot]\right\|_\infty \left\|B_f\right\|_\infty + \left\|g'[\cdot]\right\|_\infty\right) h,
\end{aligned}$$

and by the same token $\left\|C_f^g(t) - \tilde{C}_{f,h}^g(t_i)\right\| \le \mathbf{L}\left(\left\|f[\cdot]\right\|_\infty \left\|C_f\right\|_\infty + \left\|g'[\cdot]\right\|_\infty\right) h$ for all $t \in (t_{i-1}, t_i]$, which proves the assertion. □

We abbreviate the discrete Hamilton function (5.2.3) at the continuous KKT-point by

$$\widetilde{\mathcal{H}}_h[t_i, t_{i-1}] := \widetilde{\mathcal{H}}_h\left(\hat{x}(t_i), \hat{y}(t_i), \hat{u}(t_i), \hat{\lambda}_f(t_{i-1}), \hat{\lambda}_g(t_{i-1}), \hat{\eta}(t_i)\right), \quad i = 1,\ldots,N,$$

and analog for its derivatives. Additionally, we define the discrete bilinear form as

$$\mathcal{P}_h : X_{2,h} \times X_{2,h} \to \mathbb{R},$$

$$\mathcal{P}_h\left(\left(x_h^1, y_h^1, u_h^1\right), \left(x_h^2, y_h^2, u_h^2\right)\right) := \begin{pmatrix} x_h^1(t_0) \\ x_h^1(t_N) \end{pmatrix}^\top \begin{bmatrix} \Lambda_{00} & \Lambda_{01} \\ \Lambda_{01}^\top & \Lambda_{11} \end{bmatrix} \begin{pmatrix} x_h^2(t_0) \\ x_h^2(t_N) \end{pmatrix} \tag{5.3.1}$$
$$+ \sum_{i=1}^{N} h \begin{pmatrix} x_h^1(t_i) \\ y_h^1(t_i) \\ u_h^1(t_i) \end{pmatrix}^\top \begin{bmatrix} \nabla_{xx}^2 \widetilde{\mathcal{H}}_h[t_i, t_{i-1}] & \nabla_{xy}^2 \widetilde{\mathcal{H}}_h[t_i, t_{i-1}] & \nabla_{xu}^2 \widetilde{\mathcal{H}}_h[t_i, t_{i-1}] \\ \nabla_{yx}^2 \widetilde{\mathcal{H}}_h[t_i, t_{i-1}] & \nabla_{yy}^2 \widetilde{\mathcal{H}}_h[t_i, t_{i-1}] & \nabla_{yu}^2 \widetilde{\mathcal{H}}_h[t_i, t_{i-1}] \\ \nabla_{ux}^2 \widetilde{\mathcal{H}}_h[t_i, t_{i-1}] & \nabla_{uy}^2 \widetilde{\mathcal{H}}_h[t_i, t_{i-1}] & \nabla_{uu}^2 \widetilde{\mathcal{H}}_h[t_i, t_{i-1}] \end{bmatrix} \begin{pmatrix} x_h^2(t_i) \\ y_h^2(t_i) \\ u_h^2(t_i) \end{pmatrix},$$

which is symmetric and continuous, since

$$\nabla^2_{(x,y,u)(x,y,u)} \widetilde{\mathcal{H}}_h[t_i, t_{i-1}], \quad i = 1, \ldots, N$$

is bounded according to (5.A1) and (5.A2). In addition, by Lemma 2.4.12, the bilinear form $\mathcal{P}_h$ satisfies the following:

Lemma 5.3.2 (Discrete Coercivity)
Let (5.A1) - (5.A5) hold. Then, there exist $h_1, \tilde{\gamma} > 0$ such that for all $0 < h \leq h_1$ the bilinear form $\mathcal{P}_h$ is coercive on $\ker\left(\check{F}_h^\nu\right)$ with constant $\tilde{\gamma} > 0$.

Proof. In order to apply Lemma 2.4.12, it remains to show that the conditions in Assumption 2.4.11 hold. Since the KKT-point $\left(\hat{x}, \hat{y}, \hat{u}, \hat{\lambda}_f, \hat{\lambda}_g, \hat{\varsigma}, \hat{\sigma}, \hat{\eta}\right)$ is Lipschitz continuous and (5.A2) holds, we find a constant $\mathbf{L}_{\nabla^2 \mathcal{H}} \geq 0$ using similar techniques as in Lemma 5.3.1 such that for all $i \in \{1, \ldots, N\}$ and $t \in (t_{i-1}, t_i]$ we obtain

$$\left\| \begin{bmatrix} \nabla_{xx}^2 \mathcal{H}[t] & \nabla_{xy}^2 \mathcal{H}[t] & \nabla_{xu}^2 \mathcal{H}[t] \\ \nabla_{yx}^2 \mathcal{H}[t] & \nabla_{yy}^2 \mathcal{H}[t] & \nabla_{yu}^2 \mathcal{H}[t] \\ \nabla_{ux}^2 \mathcal{H}[t] & \nabla_{uy}^2 \mathcal{H}[t] & \nabla_{uu}^2 \mathcal{H}[t] \end{bmatrix} - \begin{bmatrix} \nabla_{xx}^2 \widetilde{\mathcal{H}}_h[t_i, t_{i-1}] & \nabla_{xy}^2 \widetilde{\mathcal{H}}_h[t_i, t_{i-1}] & \nabla_{xu}^2 \widetilde{\mathcal{H}}_h[t_i, t_{i-1}] \\ \nabla_{yx}^2 \widetilde{\mathcal{H}}_h[t_i, t_{i-1}] & \nabla_{yy}^2 \widetilde{\mathcal{H}}_h[t_i, t_{i-1}] & \nabla_{yu}^2 \widetilde{\mathcal{H}}_h[t_i, t_{i-1}] \\ \nabla_{ux}^2 \widetilde{\mathcal{H}}_h[t_i, t_{i-1}] & \nabla_{uy}^2 \widetilde{\mathcal{H}}_h[t_i, t_{i-1}] & \nabla_{uu}^2 \widetilde{\mathcal{H}}_h[t_i, t_{i-1}] \end{bmatrix} \right\| \leq \mathbf{L}_{\nabla^2 \mathcal{H}} h,$$

which proves the assertion. □

From this, we can immediately conclude with Lemma 2.4.14 and Lemma 2.4.15 that continuous and discrete Legendre-Clebsch conditions are satisfied:

Lemma 5.3.3 (Legendre-Clebsch Conditions)
Let (5.A1) - (5.A5) hold with coercivity constant $\gamma > 0$. Then, for every $t \in [0, 1]$ and each $v \in \ker\left(\begin{bmatrix} B_f^g(t) & C_f^g(t) \\ \check{B}_c^\nu(t) & \check{C}_c^\nu(t) \end{bmatrix}\right)$ the continuous Legendre-Clebsch condition

$$v^\top \begin{bmatrix} \nabla_{yy}^2 \mathcal{H}[t] & \nabla_{yu}^2 \mathcal{H}[t] \\ \nabla_{uy}^2 \mathcal{H}[t] & \nabla_{uu}^2 \mathcal{H}[t] \end{bmatrix} v \geq \gamma \|v\|^2$$

holds.

Furthermore, there exist $h_1, \tilde{\gamma} > 0$ *such that for all* $0 < h \leq h_1$*, every* $i \in \{1, \ldots, N\}$*, and each* $v \in \ker\left(\begin{bmatrix} \tilde{B}^g_{f,h}(t_i) & \tilde{C}^g_{f,h}(t_i) \\ \check{B}^\nu_c(t_i) & \check{C}^\nu_c(t_i) \end{bmatrix}\right)$ *the discrete Legendre-Clebsch condition*

$$v^\top \begin{bmatrix} \nabla^2_{yy}\widetilde{\mathcal{H}}_h[t_i, t_{i-1}] & \nabla^2_{yu}\widetilde{\mathcal{H}}_h[t_i, t_{i-1}] \\ \nabla^2_{uy}\widetilde{\mathcal{H}}_h[t_i, t_{i-1}] & \nabla^2_{uu}\widetilde{\mathcal{H}}_h[t_i, t_{i-1}] \end{bmatrix} v \geq \tilde{\gamma}\,\|v\|^2$$

holds.

□

5.4 Abstract Setting

Our aim is to apply Theorem 2.2.6 to the continuous necessary conditions (5.1.2) of Problem 5.1.1 and the discrete necessary conditions (5.2.4) of Problem 5.2.3. Thus, we write the KKT-conditions of the respective systems as generalized equations of the form (2.2.1). To that end, for $p = 2, \infty$ we define the following spaces

$$\begin{aligned}
\Xi_p &:= W^{n_x}_{1,p}([0,1]) \times L^{n_y}_p([0,1]) \times L^{n_u}_p([0,1]) \\
&\quad \times W^{n_x}_{1,p}([0,1]) \times L^{n_y}_p([0,1]) \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_\psi} \times L^{n_c}_p([0,1]), \\
\xi &:= (x, y, u, \lambda_f, \lambda_g, \varsigma, \sigma, \eta) \\
\Xi_{p,h} &:= W^{n_x}_{1,p,h}([0,1]) \times L^{n_y}_{p,h}([0,1]) \times L^{n_u}_{p,h}([0,1]) \\
&\quad \times W^{n_x}_{1,p,h}([0,1]) \times L^{n_y}_{p,h}([0,1]) \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_\psi} \times L^{n_c}_{p,h}([0,1]), \\
\xi_h &:= (x_h, y_h, u_h, \lambda_{f,h}, \lambda_{g,h}, \varsigma_h, \sigma_h, \eta_h) \\
\Omega_p &:= L^{n_x}_p([0,1]) \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \times L^{n_y}_p([0,1]) \times L^{n_u}_p([0,1]) \\
&\quad \times L^{n_x}_p([0,1]) \times L^{n_y}_p([0,1]) \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_\psi} \times L^{n_c}_p([0,1]), \\
\omega &:= \left(\omega_{\mathcal{H}_x}, \omega_{\vartheta_0}, \omega_{\vartheta_1}, \omega_{\mathcal{H}_y}, \omega_{\mathcal{H}_u}, \omega_f, \omega^g_f, \omega_{g_0}, \omega_\psi, \omega_c\right) \\
\Omega_{p,h} &:= L^{n_x}_{p,h}([0,1]) \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \times L^{n_y}_{p,h}([0,1]) \times L^{n_u}_{p,h}([0,1]) \\
&\quad \times L^{n_x}_{p,h}([0,1]) \times L^{n_y}_{p,h}([0,1]) \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_\psi} \times L^{n_c}_{p,h}([0,1]), \\
\omega_h &:= \left(\omega_{\mathcal{H}_x,h}, \omega_{\vartheta_0,h}, \omega_{\vartheta_1,h}, \omega_{\mathcal{H}_y,h}, \omega_{\mathcal{H}_u,h}, \omega_{f,h}, \omega^g_{f,h}, \omega_{g_0,h}, \omega_{\psi,h}, \omega_{c,h}\right),
\end{aligned}$$

equipped with the norms

$$\begin{aligned}
\|\xi\|_{\Xi_p} &:= \max\left\{\|x\|_{1,p}, \|y\|_p, \|u\|_p, \|\lambda_f\|_{1,p}, \|\lambda_g\|_p, \|\varsigma\|, \|\sigma\|, \|\eta\|_p\right\}, \\
\|\omega\|_{\Omega_p} &:= \max\Big\{\|\omega_{\mathcal{H}_x}\|_p, \|\omega_{\vartheta_0}\|, \|\omega_{\vartheta_1}\|, \|\omega_{\mathcal{H}_y}\|_p, \|\omega_{\mathcal{H}_u}\|_p, \\
&\qquad \|\omega_f\|_p, \left\|\omega^g_f\right\|_p, \|\omega_{g_0}\|, \|\omega_\psi\|, \|\omega_c\|_p\Big\}.
\end{aligned}$$

Moreover, we define the functions $\mathcal{T} : \Xi_\infty \to \Omega_\infty$ and $\mathcal{T}_h : \Xi_{\infty,h} \to \Omega_{\infty,h}$ as

$$\mathcal{T}(\xi) := \begin{pmatrix} \dot{\lambda}_f(\cdot) + \nabla_x \mathcal{H}(x(\cdot), y(\cdot), u(\cdot), \lambda_f(\cdot), \lambda_g(\cdot), \eta(\cdot)) \\ \lambda_f(0) + \nabla_{x_0}\varphi(x(0), x(1)) + \psi'_{x_0}(x(0), x(1))^\top \sigma + g'(x(0))^\top \varsigma \\ \lambda_f(1) - \nabla_{x_1}\varphi(x(0), x(1)) - \psi'_{x_1}(x(0), x(1))^\top \sigma \\ \nabla_y \mathcal{H}(x(\cdot), y(\cdot), u(\cdot), \lambda_f(\cdot), \lambda_g(\cdot), \eta(\cdot)) \\ \nabla_u \mathcal{H}(x(\cdot), y(\cdot), u(\cdot), \lambda_f(\cdot), \lambda_g(\cdot), \eta(\cdot)) \\ \dot{x}(\cdot) - f(x(\cdot), y(\cdot), u(\cdot)) \\ g'(x(\cdot)) f(x(\cdot), y(\cdot), u(\cdot)) \\ g(x(0)) \\ \psi(x(0), x(1)) \\ c(x(\cdot), y(\cdot), u(\cdot)) \end{pmatrix},$$

and for $i = 1, \ldots, N, \ t \in (t_{i-1}, t_i]$

$$\mathcal{T}_h(\xi_h)(t) := \begin{pmatrix} \lambda'_{f,h}(t_i) + \nabla_x \widetilde{\mathcal{H}}_h(x_h(t_i), y_h(t_i), u_h(t_i), \lambda_{f,h}(t_{i-1}), \lambda_{g,h}(t_{i-1}), \eta_h(t_i)) \\ \lambda_{f,h}(t_0) + \nabla_{x_0}\varphi(x_h(t_0), x_h(t_N)) + \psi'_{x_0}(x_h(t_0), x_h(t_N))^\top \sigma_h \\ \hfill + g'(x_h(t_0))^\top \varsigma_h \\ \lambda_{f,h}(t_N) - \nabla_{x_1}\varphi(x_h(t_0), x_h(t_N)) - \psi'_{x_1}(x_h(t_0), x_h(t_N))^\top \sigma_h \\ \nabla_y \widetilde{\mathcal{H}}_h(x_h(t_i), y_h(t_i), u_h(t_i), \lambda_{f,h}(t_{i-1}), \lambda_{g,h}(t_{i-1}), \eta_h(t_i)) \\ \nabla_u \widetilde{\mathcal{H}}_h(x_h(t_i), y_h(t_i), u_h(t_i), \lambda_{f,h}(t_{i-1}), \lambda_{g,h}(t_{i-1}), \eta_h(t_i)) \\ x'_h(t_i) - f(x_h(t_i), y_h(t_i), u_h(t_i)) \\ \tilde{g}_h(x_h(t_i), y_h(t_i), u_h(t_i)) \\ g(x_h(t_0)) \\ \psi(x_h(t_0), x_h(t_N)) \\ c(x_h(t_i), y_h(t_i), u_h(t_i)) \end{pmatrix}.$$

Note that $\mathcal{T}$ is defined with the (continuous) Hamilton function $\mathcal{H}$ and $\mathcal{T}_h$ with the (discrete) Hamilton function $\widetilde{\mathcal{H}}_h$ in (5.2.3), which occurs in the necessary conditions (5.2.4) for Problem 5.2.3. Additionally, the set valued mappings $\mathcal{F} : \Xi_\infty \rightrightarrows \Omega_\infty$ and $\mathcal{F}_h : \Xi_{\infty,h} \rightrightarrows \Omega_{\infty,h}$ are defined by

$$\mathcal{F}(\xi) := \begin{pmatrix} \left\{\mathbf{0}_{L_\infty^{n_x}([0,1])}\right\} \\ \{\mathbf{0}_{\mathbb{R}^{n_x}}\} \\ \{\mathbf{0}_{\mathbb{R}^{n_x}}\} \\ \left\{\mathbf{0}_{L_\infty^{n_y}([0,1])}\right\} \\ \left\{\mathbf{0}_{L_\infty^{n_u}([0,1])}\right\} \\ \left\{\mathbf{0}_{L_\infty^{n_x}([0,1])}\right\} \\ \left\{\mathbf{0}_{L_\infty^{n_y}([0,1])}\right\} \\ \{\mathbf{0}_{\mathbb{R}^{n_y}}\} \\ \{\mathbf{0}_{\mathbb{R}^{n_\psi}}\} \\ \mathcal{F}_c(\eta) \end{pmatrix}, \quad \mathcal{F}_h(\xi_h) := \begin{pmatrix} \left\{\mathbf{0}_{L_\infty^{n_x}([0,1])}\right\} \\ \{\mathbf{0}_{\mathbb{R}^{n_x}}\} \\ \{\mathbf{0}_{\mathbb{R}^{n_x}}\} \\ \left\{\mathbf{0}_{L_\infty^{n_y}([0,1])}\right\} \\ \left\{\mathbf{0}_{L_\infty^{n_u}([0,1])}\right\} \\ \left\{\mathbf{0}_{L_\infty^{n_x}([0,1])}\right\} \\ \left\{\mathbf{0}_{L_\infty^{n_y}([0,1])}\right\} \\ \{\mathbf{0}_{\mathbb{R}^{n_y}}\} \\ \{\mathbf{0}_{\mathbb{R}^{n_\psi}}\} \\ \mathcal{F}_{c,h}(\eta_h) \end{pmatrix},$$

$$\mathcal{F}_c : L_\infty^{n_c}([0,1]) \rightrightarrows L_\infty^{n_c}([0,1]), \quad \mathcal{F}_{c,h} : L_{\infty,h}^{n_c}([0,1]) \rightrightarrows L_{\infty,h}^{n_c}([0,1]),$$

$$\mathcal{F}_c(\eta) := \begin{cases} \left\{\omega_c \in L_{\infty,+}^{n_c}([0,1]) \mid \eta(t)^\top \omega_c(t) = 0, \text{ a.e. in } [0,1]\right\}, & \text{if } \eta \in L_{\infty,+}^{n_c}([0,1]) \\ \emptyset, & \text{otherwise} \end{cases},$$

$$\mathcal{F}_{c,h}(\eta_h) := \begin{cases} \left\{\omega_{c,h} \in L_{\infty,h,+}^{n_c}([0,1]) \mid \eta_h(t_i)^\top \omega_{c,h}(t_i) = 0, \; i = 1,\ldots,N\right\}, & \text{if } \eta_h \in L_{\infty,h,+}^{n_c}([0,1]) \\ \emptyset, & \text{otherwise} \end{cases},$$

where $L_{\infty,+}^{n_c}([0,1]) \subseteq L_\infty^{n_c}([0,1])$ and $L_{\infty,h,+}^{n_c}([0,1]) \subseteq L_{\infty,h}^{n_c}([0,1])$ consist of functions, which are non-negative almost everywhere on $[0,1]$. For this notation we consider the generalized equations

$$\mathbf{0}_\Omega \in \mathcal{T}(\xi) + \mathcal{F}(\xi), \tag{5.4.1}$$

$$\mathbf{0}_\Omega \in \mathcal{T}_h(\xi_h) + \mathcal{F}_h(\xi_h). \tag{5.4.2}$$

5.5 Convergence Proof

Throughout this section, we assume that the conditions (5.A1) - (5.A5) hold. Our goal is to apply Theorem 2.2.6 for the KKT-point $\hat{\xi} := \left(\hat{x}, \hat{y}, \hat{u}, \hat{\lambda}_f, \hat{\lambda}_g, \hat{\varsigma}, \hat{\sigma}, \hat{\eta}\right)$ (which solves the continuous inclusion (5.4.1)) and the discrete generalized equation (5.4.2), thus we need to verify the following conditions:

(i) $\mathcal{T}_h'(\cdot)$ is Lipschitz continuous with a constant independent of h.

(ii) There exist $\tilde{\xi}_h \in \Xi_{\infty,h}$ and $\hat{\omega}_h \in \Omega_{\infty,h}$ such that

$$\mathbf{0}_\Omega \in \mathcal{T}_h\left(\tilde{\xi}_h\right) + \hat{\omega}_h + \mathcal{F}_h\left(\tilde{\xi}_h\right)$$

$$\text{and} \qquad \left\|\tilde{\xi}_h - \hat{\xi}\right\|_{\Xi_\infty} \to 0, \quad \|\hat{\omega}_h\|_{\Omega_\infty} \to 0 \quad \text{for } h \to 0.$$

(iii) There exist $h_1, \varrho, \rho > 0$ such that for all $0 < h \le h_1$ and every $\zeta \in \mathcal{B}_\varrho(\mathbf{0}_{\Omega_\infty})$ the inclusion

$$\zeta \in \mathcal{T}_h\left(\tilde{\xi}_h\right) + \hat{\omega}_h + \mathcal{T}_h'\left(\tilde{\xi}_h\right)\left(\xi_h - \tilde{\xi}_h\right) + \mathcal{F}_h(\xi_h)$$

has a unique solution $\xi_h(\zeta) \in \mathcal{B}_\rho\left(\tilde{\xi}_h\right)$, which is Lipschitz continuous with respect to ζ and a Lipschitz constant independent of h.

According to the smoothness assumption (5.A2), the function $\mathcal{T}_h'(\cdot)$ is Lipschitz continuous, hence condition (i) of Theorem 2.2.6 is satisfied.
For (ii) we define the projections $\Delta_h : \Xi_\infty \to \Xi_{\infty,h}$, $\Delta_h^{1,n} : W_{1,\infty}^n([0,1]) \to W_{1,\infty,h}^n([0,1])$, and $\Delta_h^{0,n} : L_\infty^n([0,1]) \cap \mathcal{C}_0^n([0,1]) \to L_{\infty,h}^n([0,1])$ as

$$\Delta_h(\xi) := \left(\Delta_h^{1,n_x}(x), \Delta_h^{0,n_y}(y), \Delta_h^{0,n_u}(u), \Delta_h^{1,n_x}(\lambda_f), \Delta_h^{0,n_y}(\lambda_g), \varsigma, \sigma, \Delta_h^{0,n_c}(\eta)\right),$$

$$\begin{aligned} \Delta_h^{1,n}(v)(t_0) &:= v(t_0), & \\ \Delta_h^{1,n}(v)(t) &:= v'(t_i)(t - t_{i-1}) + v(t_{i-1}), & t \in (t_{i-1}, t_i], \; i = 1,\ldots,N, \\ \Delta_h^{0,n}(v)(t) &:= v(t_i), & t \in (t_{i-1}, t_i], \; i = 1,\ldots,N, \end{aligned}$$

(compare Figure 2.1). Then, $\Delta_h\left(\hat{\xi}\right)$ and $\hat{\omega}_h \in \Omega_{\infty,h}$ defined by

$$\hat{\omega}_h(t) := \begin{pmatrix} -\frac{\hat{\lambda}_f(t_i)-\hat{\lambda}_f(t_{i-1})}{h} - \nabla_x \widetilde{\mathcal{H}}_h[t_i, t_{i-1}] \\ \mathbf{0}_{\mathbb{R}^{n_x}} \\ \mathbf{0}_{\mathbb{R}^{n_x}} \\ -\nabla_y \widetilde{\mathcal{H}}_h[t_i, t_{i-1}] \\ -\nabla_u \widetilde{\mathcal{H}}_h[t_i, t_{i-1}] \\ -\frac{\hat{x}(t_i)-\hat{x}(t_{i-1})}{h} + f[t_i] \\ -\tilde{g}_h[t_i] \\ \mathbf{0}_{\mathbb{R}^{n_y}} \\ \mathbf{0}_{\mathbb{R}^{n_\psi}} \\ \mathbf{0}_{\mathbb{R}^{n_c}} \end{pmatrix},$$

for $t \in (t_{i-1}, t_i]$, $i = 1, \ldots, N$ satisfy the inclusion

$$\mathbf{0}_\Omega \in \mathcal{T}_h\left(\Delta_h\left(\hat{\xi}\right)\right) + \hat{\omega}_h + \mathcal{F}_h\left(\Delta_h\left(\hat{\xi}\right)\right).$$

Furthermore, since

$$\begin{aligned}\left(\hat{x}, \hat{y}, \hat{u}, \hat{\lambda}_f, \hat{\lambda}_g, \hat{\eta}\right) \in\; & W_{2,\infty}^{n_x}([0,1]) \times W_{1,\infty}^{n_y}([0,1]) \times W_{1,\infty}^{n_u}([0,1]) \\ & \times W_{2,\infty}^{n_x}([0,1]) \times W_{1,\infty}^{n_y}([0,1]) \times W_{1,\infty}^{n_c}([0,1]),\end{aligned} \tag{5.5.1}$$

it holds

$$\begin{aligned}\left\|\hat{x}(t) - \Delta_h^{1,n_x}(\hat{x})(t)\right\| &= \left\|\hat{x}(t) - \hat{x}(t_{i-1}) - \hat{x}'(t_i)(t - t_{i-1})\right\| \leq 2\left\|\dot{\hat{x}}\right\|_\infty h, \\ \left\|\frac{d}{dt}\left(\hat{x}(t) - \Delta_h^{1,n_x}(\hat{x})(t)\right)\right\| &= \left\|\dot{\hat{x}}(t) - \hat{x}'(t_i)\right\| \leq \left\|\ddot{\hat{x}}\right\|_\infty h, \\ \left\|\hat{y}(t) - \Delta_h^{0,n_y}(\hat{y})(t)\right\| &= \left\|\hat{y}(t) - \hat{y}(t_i)\right\| \leq \left\|\dot{\hat{y}}\right\|_\infty h\end{aligned}$$

for $t \in (t_{i-1}, t_i]$, $i = 1, \ldots, N$. Analog, we obtain bounds for $\left(\hat{u}, \hat{\lambda}_f, \hat{\lambda}_g, \hat{\eta}\right)$, thus

$$\left\|\hat{\xi} - \Delta_h\left(\hat{\xi}\right)\right\|_\Xi \to 0 \quad \text{for } h \to 0,$$

with a linear convergence rate. By (5.A2) and (5.5.1), we find a constant $\mathbf{L}_{\nabla\mathcal{H}} \geq 0$ such that

$$\begin{aligned}\left\|\nabla_x \mathcal{H}[t] - \nabla_x \widetilde{\mathcal{H}}_h[t_i, t_{i-1}]\right\| &\leq \mathbf{L}_{\nabla\mathcal{H}} h, \\ \left\|-\nabla_y \widetilde{\mathcal{H}}_h[t_i, t_{i-1}]\right\| = \left\|\nabla_y \mathcal{H}[t] - \nabla_y \widetilde{\mathcal{H}}_h[t_i, t_{i-1}]\right\| &\leq \mathbf{L}_{\nabla\mathcal{H}} h, \\ \left\|-\nabla_u \widetilde{\mathcal{H}}_h[t_i, t_{i-1}]\right\| = \left\|\nabla_u \mathcal{H}[t] - \nabla_u \widetilde{\mathcal{H}}_h[t_i, t_{i-1}]\right\| &\leq \mathbf{L}_{\nabla\mathcal{H}} h\end{aligned}$$

for $t \in (t_{i-1}, t_i]$, $i = 1, \ldots, N$, and therefore we have

$$\begin{aligned}\left\|-\frac{\hat{\lambda}_f(t_i) - \hat{\lambda}_f(t_{i-1})}{h} - \nabla_x \widetilde{\mathcal{H}}_h[t_i, t_{i-1}]\right\| &= \left\|\frac{1}{h}\int_{t_{i-1}}^{t_i} -\dot{\hat{\lambda}}_f(t) - \nabla_x \widetilde{\mathcal{H}}_h[t_i, t_{i-1}]\, dt\right\| \\ &= \left\|\frac{1}{h}\int_{t_{i-1}}^{t_i} \nabla_x \mathcal{H}[t] - \nabla_x \widetilde{\mathcal{H}}_h[t_i, t_{i-1}]\, dt\right\| \\ &\leq \mathbf{L}_{\nabla\mathcal{H}} h\end{aligned}$$

for $i = 1, \dots, N$. By the same token, we get $\left\| -\frac{\hat{x}(t_i) - \hat{x}(t_{i-1})}{h} + f[t_i] \right\| \leq \mathbf{L}h$ for $i = 1, \dots, N$. Finally, using the Taylor expansion and exploiting $g'[t_i] f[t_i] = \mathbf{0}_{\mathbb{R}^{n_y}}$ yields for $i = 1, \dots, N$

$$\begin{aligned}
&\left\| \frac{g(\hat{x}(t_i)) - g(\hat{x}(t_i) - hf[t_i])}{h} \right\| \\
&\quad = \left\| -g'[t_i] f[t_i] + h \int_0^1 (1-\theta) g''(\hat{x}(t_i) - \theta h f[t_i]) (f[t_i], f[t_i]) \, d\theta \right\| \\
&\quad \leq h \int_0^1 (1-\theta) \left(\| g''(\hat{x}(t_i)) \| + \| g''(\hat{x}(t_i)) - g''(\hat{x}(t_i) - \theta h f[t_i]) \| \right) \| f[\cdot] \|_\infty^2 \, d\theta \\
&\quad \leq h \int_0^1 (1-\theta) \left(\| g''[\cdot] \|_\infty + \theta h \mathbf{L} \| f[\cdot] \|_\infty \right) \| f[\cdot] \|_\infty^2 \, d\theta \leq \left(\| g''[\cdot] \|_\infty + \mathbf{L} \| f[\cdot] \|_\infty \right) \| f[\cdot] \|_\infty^2 \, h.
\end{aligned}$$

Overall, we conclude $\|\hat{\omega}_h\|_\Omega \to 0$ for $h \to 0$ with a linear convergence rate, hence condition (ii) holds.

Using the techniques in [80] we take the following steps in order to verify condition (iii) (see Figure 5.1):

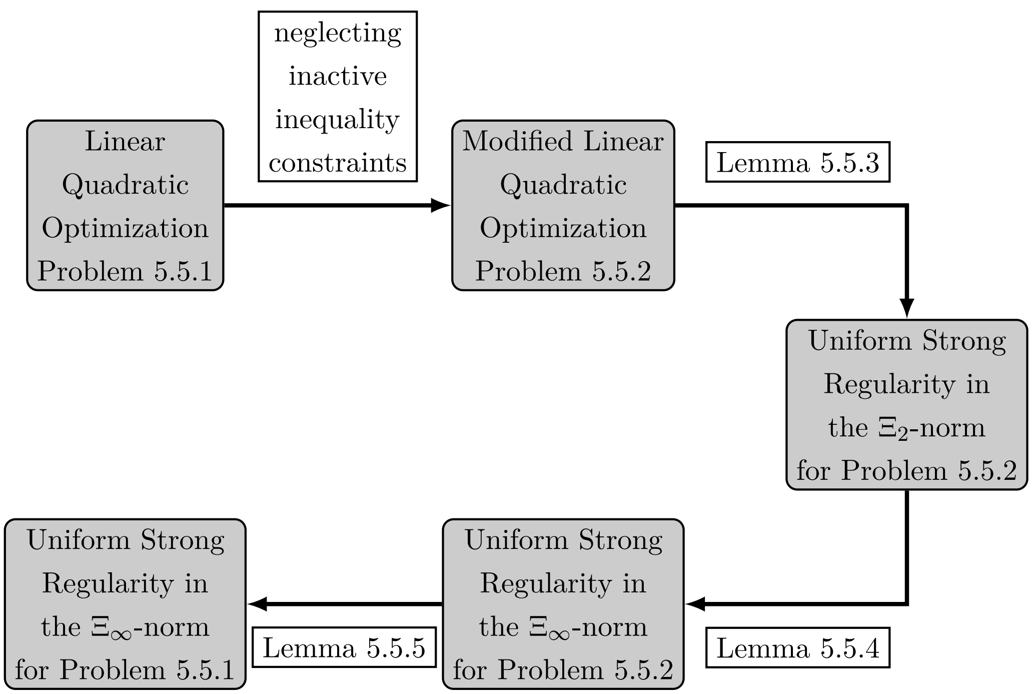

Figure 5.1: Scheme to verify condition (iii) (uniform strong regularity) for (5.4.2).

(a) First, we see that the (perturbed) linearized inclusion

$$\zeta \in \mathcal{T}_h\left(\Delta_h\left(\hat{\xi}\right)\right) + \hat{\omega}_h + \mathcal{T}_h'\left(\Delta_h\left(\hat{\xi}\right)\right)\left(\xi_h - \Delta_h\left(\hat{\xi}\right)\right) + \mathcal{F}_h\left(\xi_h\right) \tag{5.5.2}$$

represents the KKT-conditions of a linear quadratic optimization problem (Problem 5.5.1).

(b) Then, we modify the inequality constraints of that problem by neglecting inactive constraints and obtain a system, where the gradients of all constraints are linear independent (Problem 5.5.2).

(c) Using the discrete coercivity condition in Lemma 5.3.2 we show that the modified problem has a unique solution for every perturbation, which satisfies condition (*iii*) in the weaker Ξ_2-norm (Lemma 5.5.3).

(d) Exploiting the discrete Legendre-Clebsch condition in Lemma 5.3.3 and the sensitivity result in Corollary 2.3.10 yields uniform strong regularity in the Ξ_∞-norm (Lemma 5.5.4).

(e) Finally, we show that for sufficiently small perturbations the unique solution of the modified problem is also the unique solution of the original linear quadratic problem, thus (5.5.2) satisfies (*iii*) in the Ξ_∞-norm (Lemma 5.5.5).

For an arbitrary perturbation $\zeta \in \Omega_{\infty,h}$ we denote

$$\pi_h\left(\zeta\right) = \begin{pmatrix} \pi_{\mathcal{H}_x,h}\left(\zeta\right) \\ \pi_{\vartheta_0,h}\left(\zeta\right) \\ \pi_{\vartheta_1,h}\left(\zeta\right) \\ \pi_{\mathcal{H}_y,h}\left(\zeta\right) \\ \pi_{\mathcal{H}_u,h}\left(\zeta\right) \\ \pi_{f,h}\left(\zeta\right) \\ \pi_{f,h}^{g}\left(\zeta\right) \\ \pi_{g_0,h}\left(\zeta\right) \\ \pi_{\psi,h}\left(\zeta\right) \\ \pi_{c,h}\left(\zeta\right) \end{pmatrix} := \mathcal{T}_h\left(\Delta_h\left(\hat{\xi}\right)\right) + \hat{\omega}_h - \mathcal{T}_h'\left(\Delta_h\left(\hat{\xi}\right)\right)\Delta_h\left(\hat{\xi}\right) - \zeta.$$

Then, the (perturbed) linearized inclusion becomes

$$\mathbf{0}_\Omega \in \mathcal{T}_h'\left(\Delta_h\left(\hat{\xi}\right)\right)\xi_h + \mathcal{F}_h\left(\xi_h\right) + \pi_h\left(\zeta\right),$$

which we can write as

$$
\begin{array}{rcrcrcl}
\lambda_{f,h}'(t_i) & + & \nabla_{xx}^2 \widetilde{\mathcal{H}}_h [t_i, t_{i-1}] \, x_h(t_i) & + & \nabla_{xy}^2 \widetilde{\mathcal{H}}_h [t_i, t_{i-1}] \, y_h(t_i) & & \\
 & + & \nabla_{xu}^2 \widetilde{\mathcal{H}}_h [t_i, t_{i-1}] \, u_h(t_i) & + & A_f(t_i)^\top \lambda_{f,h}(t_{i-1}) & & \\
 & + & \tilde{A}_{f,h}^g(t_i)^\top \lambda_{g,h}(t_{i-1}) & + & A_c(t_i)^\top \eta_h(t_i) & & \\
 & & & + & \pi_{\mathcal{H}_x,h}(\zeta)(t_i) & = & \mathbf{0}_{\mathbb{R}^{n_x}}, \\[2ex]
\lambda_{f,h}(t_0) & + & \Lambda_{00} x_h(t_0) & + & \Lambda_{01} x_h(t_N) & & \\
 & + & \Psi_0^\top \sigma_h & + & E_0^\top \varsigma_h & & \\
 & & & + & \pi_{\vartheta_0,h}(\zeta) & = & \mathbf{0}_{\mathbb{R}^{n_x}}, \\[2ex]
\lambda_{f,h}(t_N) & + & \Lambda_{01}^\top x_h(t_0) & + & \Lambda_{11} x_h(t_N) & & \\
 & + & \Psi_1^\top \sigma_h & + & \pi_{\vartheta_1,h}(\zeta) & = & \mathbf{0}_{\mathbb{R}^{n_x}}, \\[2ex]
 & & \nabla_{yx}^2 \widetilde{\mathcal{H}}_h [t_i, t_{i-1}] \, x_h(t_i) & + & \nabla_{yy}^2 \widetilde{\mathcal{H}}_h [t_i, t_{i-1}] \, y_h(t_i) & & \\
 & + & \nabla_{yu}^2 \widetilde{\mathcal{H}}_h [t_i, t_{i-1}] \, u_h(t_i) & + & B_f(t_i)^\top \lambda_{f,h}(t_{i-1}) & & \\
 & + & \tilde{B}_{f,h}^g(t_i)^\top \lambda_{g,h}(t_{i-1}) & + & B_c(t_i)^\top \eta_h(t_i) & & \\
 & & & + & \pi_{\mathcal{H}_y,h}(\zeta)(t_i) & = & \mathbf{0}_{\mathbb{R}^{n_y}}, \\[2ex]
 & & \nabla_{ux}^2 \widetilde{\mathcal{H}}_h [t_i, t_{i-1}] \, x_h(t_i) & + & \nabla_{uy}^2 \widetilde{\mathcal{H}}_h [t_i, t_{i-1}] \, y_h(t_i) & & \\
 & + & \nabla_{uu}^2 \widetilde{\mathcal{H}}_h [t_i, t_{i-1}] \, u_h(t_i) & + & C_f(t_i)^\top \lambda_{f,h}(t_{i-1}) & & \\
 & + & \tilde{C}_{f,h}^g(t_i)^\top \lambda_{g,h}(t_{i-1}) & + & C_c(t_i)^\top \eta_h(t_i) & & \\
 & & & + & \pi_{\mathcal{H}_u,h}(\zeta)(t_i) & = & \mathbf{0}_{\mathbb{R}^{n_u}},
\end{array}
$$

$$
\begin{array}{rcrcrcrcrcl}
x_h'(t_i) & - & A_f(t_i) \, x_h(t_i) & - & B_f(t_i) \, y_h(t_i) & - & C_f(t_i) \, u_h(t_i) & + & \pi_{f,h}(\zeta)(t_i) & = & \mathbf{0}_{\mathbb{R}^{n_x}}, \\[2ex]
 & & \tilde{A}_{f,h}^g(t_i) \, x_h(t_i) & + & \tilde{B}_{f,h}^g(t_i) \, y_h(t_i) & + & \tilde{C}_{f,h}^g(t_i) \, u_h(t_i) & + & \pi_{f,h}^g(\zeta)(t_i) & = & \mathbf{0}_{\mathbb{R}^{n_y}}, \\[2ex]
 & & & & E_0 x_h(t_0) & & & + & \pi_{g_0,h}(\zeta) & = & \mathbf{0}_{\mathbb{R}^{n_y}}, \\
 & & & & \Psi_0 x_h(t_0) & + & \Psi_1 x_h(t_N) & + & \pi_{\psi,h}(\zeta) & = & \mathbf{0}_{\mathbb{R}^{n_\psi}}, \\[2ex]
 & & A_c(t_i) \, x_h(t_i) & + & B_c(t_i) \, y_h(t_i) & + & C_c(t_i) \, u_h(t_i) & + & \pi_{c,h}(\zeta)(t_i) & \leq & \mathbf{0}_{\mathbb{R}^{n_c}}, \\[2ex]
 & & & & & & & & \eta_h(t_i) & \geq & \mathbf{0}_{\mathbb{R}^{n_c}}, \\[2ex]
 & & \eta_h(t_i)^\top (A_c(t_i) \, x_h(t_i) & + & B_c(t_i) \, y_h(t_i) & + & C_c(t_i) \, u_h(t_i) & + & \pi_{c,h}(\zeta)(t_i)) & = & 0,
\end{array}
$$

for $i = 1, \ldots, N$. These are the KKT-conditions of the following Problem:

Problem 5.5.1 (Discrete Linear Quadratic Optimization Problem)

$$
\begin{array}{ll}
\textit{Minimize} & \frac{1}{2}\mathcal{P}_h\left(\left(x_h,y_h,u_h\right),\left(x_h,y_h,u_h\right)\right)+\sum\limits_{i=1}^{N} h \begin{pmatrix} x_h\left(t_i\right) \\ y_h\left(t_i\right) \\ u_h\left(t_i\right) \end{pmatrix}^{\top} \begin{pmatrix} \pi_{\mathcal{H}_x,h}\left(\zeta\right)\left(t_i\right) \\ \pi_{\mathcal{H}_y,h}\left(\zeta\right)\left(t_i\right) \\ \pi_{\mathcal{H}_u,h}\left(\zeta\right)\left(t_i\right) \end{pmatrix} \\
& +x_h\left(t_0\right)^{\top}\pi_{\vartheta_0,h}\left(\zeta\right)+x_h\left(t_N\right)^{\top}\pi_{\vartheta_1,h}\left(\zeta\right),
\end{array}
$$

$$
\textit{with respect to} \quad x_h\in W_{1,\infty,h}^{n_x}\left(\left[0,1\right]\right), y_h\in L_{\infty,h}^{n_y}\left(\left[0,1\right]\right), u_h\in L_{\infty,h}^{n_u}\left(\left[0,1\right]\right),
$$

$$
\begin{array}{lrcl}
\textit{subject to} & x_h'\left(t_i\right) & = & A_f\left(t_i\right)x_h\left(t_i\right)+B_f\left(t_i\right)y_h\left(t_i\right)+C_f\left(t_i\right)u_h\left(t_i\right)-\pi_{f,h}\left(\zeta\right)\left(t_i\right), \\
& \mathbf{0}_{\mathbb{R}^{n_y}} & = & \tilde{A}_{f,h}^{g}\left(t_i\right)x_h\left(t_i\right)+\tilde{B}_{f,h}^{g}\left(t_i\right)y_h\left(t_i\right)+\tilde{C}_{f,h}^{g}\left(t_i\right)u_h\left(t_i\right)+\pi_{f,h}^{g}\left(\zeta\right)\left(t_i\right), \\
& \mathbf{0}_{\mathbb{R}^{n_y}} & = & E_0x_h\left(t_0\right)+\pi_{g_0,h}\left(\zeta\right), \\
& \mathbf{0}_{\mathbb{R}^{n_\psi}} & = & \Psi_0x_h\left(t_0\right)+\Psi_1x_h\left(t_N\right)+\pi_{\psi,h}\left(\zeta\right), \\
& \mathbf{0}_{\mathbb{R}^{n_c}} & \geq & A_c\left(t_i\right)x_h\left(t_i\right)+B_c\left(t_i\right)y_h\left(t_i\right)+C_c\left(t_i\right)u_h\left(t_i\right)+\pi_{c,h}\left(\zeta\right)\left(t_i\right),
\end{array}
$$

for $i=1,\ldots,N$.

The index set of the active constraints of Problem 5.5.1 stay the same as in the unperturbed case for sufficiently small perturbations, as we will prove later in Lemma 5.5.5. Thus, we modify Problem 5.5.1 by neglecting inactive inequality constraints:

Problem 5.5.2 (Modified Discrete Linear Quadratic Optimization Problem)

$$
\begin{array}{ll}
\textit{Minimize} & \frac{1}{2}\mathcal{P}_h\left(\left(x_h,y_h,u_h\right),\left(x_h,y_h,u_h\right)\right)+\sum\limits_{i=1}^{N} h \begin{pmatrix} x_h\left(t_i\right) \\ y_h\left(t_i\right) \\ u_h\left(t_i\right) \end{pmatrix}^{\top} \begin{pmatrix} \pi_{\mathcal{H}_x,h}\left(\zeta\right)\left(t_i\right) \\ \pi_{\mathcal{H}_y,h}\left(\zeta\right)\left(t_i\right) \\ \pi_{\mathcal{H}_u,h}\left(\zeta\right)\left(t_i\right) \end{pmatrix} \\
& +x_h\left(t_0\right)^{\top}\pi_{\vartheta_0,h}\left(\zeta\right)+x_h\left(t_N\right)^{\top}\pi_{\vartheta_1,h}\left(\zeta\right),
\end{array}
$$

$$
\textit{with respect to} \quad x_h\in W_{1,\infty,h}^{n_x}\left(\left[0,1\right]\right), y_h\in L_{\infty,h}^{n_y}\left(\left[0,1\right]\right), u_h\in L_{\infty,h}^{n_u}\left(\left[0,1\right]\right),
$$

$$
\begin{array}{lrcl}
\textit{subject to} & x_h'\left(t_i\right) & = & A_f\left(t_i\right)x_h\left(t_i\right)+B_f\left(t_i\right)y_h\left(t_i\right)+C_f\left(t_i\right)u_h\left(t_i\right)-\pi_{f,h}\left(\zeta\right)\left(t_i\right), \\
& \mathbf{0}_{\mathbb{R}^{n_y}} & = & \tilde{A}_{f,h}^{g}\left(t_i\right)x_h\left(t_i\right)+\tilde{B}_{f,h}^{g}\left(t_i\right)y_h\left(t_i\right)+\tilde{C}_{f,h}^{g}\left(t_i\right)u_h\left(t_i\right)+\pi_{f,h}^{g}\left(\zeta\right)\left(t_i\right), \\
& \mathbf{0}_{\mathbb{R}^{n_y}} & = & E_0x_h\left(t_0\right)+\pi_{g_0,h}\left(\zeta\right), \\
& \mathbf{0}_{\mathbb{R}^{n_\psi}} & = & \Psi_0x_h\left(t_0\right)+\Psi_1x_h\left(t_N\right)+\pi_{\psi,h}\left(\zeta\right),
\end{array}
$$

$$
\begin{array}{ll}
\textit{and} & c_{j,x}'\left[t_i\right]x_h\left(t_i\right)+c_{j,y}'\left[t_i\right]y_h\left(t_i\right)+c_{j,u}'\left[t_i\right]u_h\left(t_i\right) \\
& +\pi_{c,h,j}\left(\zeta\right)\left(t_i\right)\begin{cases} =0, & \textit{if } t_i\in\Upsilon_j^{\nu} \\ \leq 0, & \textit{if } t_i\in\Theta_j^{\alpha}\setminus\Upsilon_j^{\nu} \end{cases}, \quad j\in J^{\alpha}\left(t_i\right),
\end{array}
$$

for $i=1,\ldots,N$.

If the sufficient conditions in Theorem 2.3.5 hold, then this linear quadratic optimization problem has a global minimum. According to Lemma 5.3.2, the bilinear form $\mathcal{P}_h$ is coercive on $\ker\left(\check{F}_h^{\nu}\right)$, hence Problem 5.5.2 has a global minimum. Moreover, by Remark 2.3.4, the associated Lagrange multipliers are unique, if the linear independence constraint qualification in Definition 2.3.3 is satisfied. Consider the system of linear equations

$$\begin{array}{ccccccccc}
 & & & & & & E_0 x_h(t_0) & = & a_{E_0} \\
 & & & & \Psi_0 x_h(t_0) & + & \Psi_1 x_h(t_N) & = & a_\Psi \\
x_h'(t_1) & - & A_f(t_1) x_h(t_1) & - & B_f(t_1) y_h(t_1) & - & C_f(t_1) u_h(t_1) & = & a_f(t_1) \\
 & & \tilde{A}^g_{f,h}(t_1) x_h(t_1) & + & \tilde{B}^g_{f,h}(t_1) y_h(t_1) & + & \tilde{C}^g_{f,h}(t_1) u_h(t_1) & = & \tilde{a}^g_{f,h}(t_1) \\
 & & A^\alpha_c(t_1) x_h(t_1) & + & B^\alpha_c(t_1) y_h(t_1) & + & C^\alpha_c(t_1) u_h(t_1) & = & a^\alpha_c(t_1) \\
x_h'(t_2) & - & A_f(t_2) x_h(t_2) & - & B_f(t_2) y_h(t_2) & - & C_f(t_2) u_h(t_2) & = & a_f(t_2) \\
 & & \tilde{A}^g_{f,h}(t_2) x_h(t_2) & + & \tilde{B}^g_{f,h}(t_2) y_h(t_2) & + & \tilde{C}^g_{f,h}(t_2) u_h(t_2) & = & \tilde{a}^g_{f,h}(t_2) \\
 & & A^\alpha_c(t_2) x_h(t_2) & + & B^\alpha_c(t_2) y_h(t_2) & + & C^\alpha_c(t_2) u_h(t_2) & = & a^\alpha_c(t_2) \\
 & & & & \vdots & & & \vdots & \vdots \\
x_h'(t_N) & - & A_f(t_N) x_h(t_N) & - & B_f(t_N) y_h(t_N) & - & C_f(t_N) u_h(t_N) & = & a_f(t_N) \\
 & & \tilde{A}^g_{f,h}(t_N) x_h(t_N) & + & \tilde{B}^g_{f,h}(t_N) y_h(t_N) & + & \tilde{C}^g_{f,h}(t_N) u_h(t_N) & = & \tilde{a}^g_{f,h}(t_N) \\
 & & A^\alpha_c(t_N) x_h(t_N) & + & B^\alpha_c(t_N) y_h(t_N) & + & C^\alpha_c(t_N) u_h(t_N) & = & a^\alpha_c(t_N)
\end{array},$$

where $a_{E_0} \in \mathbb{R}^{n_y}, a_\Psi \in \mathbb{R}^{n_\psi}$, and $a_f(t_i) \in \mathbb{R}^{n_x}, \tilde{a}^g_{f,h}(t_i) \in \mathbb{R}^{n_y}, a^\alpha_c(t_i) \in \mathbb{R}^{j^\alpha(t_i)}$ for $i = 1, \ldots, N$. According to Lemma 5.3.1, this system has a solution for arbitrary

$$\left(a_{E_0}, a_\Psi, a_f(t_1), \tilde{a}^g_{f,h}(t_1), a^\alpha_c(t_1), \ldots, a_f(t_N), \tilde{a}^g_{f,h}(t_N), a^\alpha_c(t_N)\right).$$

Thus, for every $\zeta \in \Omega_{\infty,h}$ the set of admissible vectors for Problem 5.5.2 is not empty and the matrix

$$\begin{bmatrix}
\begin{bmatrix} E_0 \\ \Psi_0 \end{bmatrix} & & & & \begin{bmatrix} \mathbf{0}_{n_y \times n_x} & \mathbf{0}_{n_y \times n_y} & \mathbf{0}_{n_y \times n_u} \\ \Psi_1 & \mathbf{0}_{n_\psi \times n_y} & \mathbf{0}_{n_\psi \times n_u} \end{bmatrix} \\
R_h(t_1) & S_h(t_1) & & & \\
 & R_h(t_2) & S_h(t_2) & & \\
 & & \ddots & & \\
 & & & R_h(t_N) & S_h(t_N)
\end{bmatrix},$$

where

$$R_h(t_1) := \begin{bmatrix} \frac{1}{h}\mathbf{I}_{n_x} \\ \mathbf{0}_{n_y \times n_x} \\ \mathbf{0}_{j^\alpha(t_1) \times n_x} \end{bmatrix}, \quad R_h(t_i) := \begin{bmatrix} \frac{1}{h}\mathbf{I}_{n_x} & \mathbf{0}_{n_x \times n_y} & \mathbf{0}_{n_x \times n_u} \\ \mathbf{0}_{n_y \times n_x} & \mathbf{0}_{n_y \times n_y} & \mathbf{0}_{n_y \times n_u} \\ \mathbf{0}_{j^\alpha(t_i) \times n_x} & \mathbf{0}_{j^\alpha(t_i) \times n_y} & \mathbf{0}_{j^\alpha(t_i) \times n_u} \end{bmatrix}, \quad i = 2, \ldots, N,$$

$$S_h(t_i) := \begin{bmatrix} -\frac{1}{h}\mathbf{I}_{n_x} + A_f(t_i) & B_f(t_i) & C_f(t_i) \\ \tilde{A}^g_{f,h}(t_i) & \tilde{B}^g_{f,h}(t_i) & \tilde{C}^g_{f,h}(t_i) \\ A^\alpha_c(t_i) & B^\alpha_c(t_i) & C^\alpha_c(t_i) \end{bmatrix}, \quad i = 1, \ldots, N,$$

has full row rank. Therefore, the linear independence constraint qualification is satisfied for Problem 5.5.2. Consequently, for every $\zeta \in \Omega_{\infty,h}$ Problem 5.5.2 has a unique (global) minimizer together with unique multipliers, which we denote by

$$\bar{\xi}_h(\zeta) := \left(\bar{x}_h(\zeta), \bar{y}_h(\zeta), \bar{u}_h(\zeta), \bar{\lambda}_{f,h}(\zeta), \bar{\lambda}_{g,h}(\zeta), \bar{\varsigma}_h(\zeta), \bar{\sigma}_h(\zeta), \bar{\eta}_h(\zeta)\right),$$

where we set $\bar{\eta}_{h,j}(\zeta)(t_i) = 0$, if $j \in J \setminus J^\alpha(t_i)$ for $i = 1, \ldots, N$. Moreover, for all $i = 1, \ldots, N$ we denote $\bar{\eta}_h^\alpha(\zeta)(t_i) := [\bar{\eta}_{h,j}(\zeta)(t_i)]_{j \in J^\alpha(t_i)}$, and prove the following (compare step (c) above):

Lemma 5.5.3
Let (5.A1) - (5.A5) hold. Then, there exist $h_1, \boldsymbol{l}_{L_2} > 0$ such that for all $0 < h \leq h_1$ and every $\zeta_1, \zeta_2 \in \Omega_{\infty,h}$ it holds

$$\left\|\bar{\xi}_h(\zeta_1) - \bar{\xi}_h(\zeta_2)\right\|_{\Xi_2} \leq \boldsymbol{l}_{L_2} \|\zeta_1 - \zeta_2\|_{\Omega_\infty}. \tag{5.5.3}$$

Proof. Since we aim to obtain a bound in the Ξ_2-norm, we consider Problem 5.5.2 in $X_{2,h}$ instead of $X_{\infty,h}$. This does not change the optimization problem, because these spaces are finite dimensional, and therefore isomorphic. Choose $h_1 > 0$ such that Lemma 5.3.1 and Lemma 5.3.2 hold for constants $\tilde{\kappa} > 0$ and $\tilde{\gamma} > 0$, respectively. Then, by Lemma 2.1.6, the operator $F_h^\alpha \circ \mathcal{I}_{X_{2,h}}^{-1} \circ F_h^{\alpha\star} \in \mathfrak{L}\left(Y_{2,h}^{\alpha\;\star}, Y_{2,h}^\alpha\right)$, where $\mathcal{I}_{X_{2,h}}$ is the canonical isomorphism between $X_{2,h}$ and $X_{2,h}^\star$ (compare Theorem 2.1.3), is bijective and the inverse is uniformly bounded. Additionally, there exist constants $\Gamma_{F^\alpha}, \Gamma_{\mathcal{P}} \geq 0$ satisfying

$$\begin{aligned} &\left\|\mathcal{I}_{X_{2,h}}^{-1} \circ F_h^{\alpha\star} \circ \left(F_h^\alpha \circ \mathcal{I}_{X_{2,h}}^{-1} \circ F_h^{\alpha\star}\right)^{-1}\right\|_{\mathfrak{L}\left(Y_{2,h}^\alpha, X_{2,h}\right)} \leq \Gamma_{F^\alpha}, \\ &\left|\mathcal{P}_h\left(\left(x_h^1, y_h^1, u_h^1\right), \left(x_h^2, y_h^2, u_h^2\right)\right)\right| \leq \Gamma_{\mathcal{P}} \left\|\left(x_h^1, y_h^1, u_h^1\right)\right\|_{X_2} \left\|\left(x_h^2, y_h^2, u_h^2\right)\right\|_{X_2} \end{aligned} \tag{5.5.4}$$

for all $\left(x_h^1, y_h^1, u_h^1\right), \left(x_h^2, y_h^2, u_h^2\right) \in X_{2,h}$. We abbreviate the perturbation that appears in the constraints of Problem 5.5.2 by

$$s(\zeta)(t_i) := \begin{pmatrix} \pi_{f,h}(\zeta)(t_i) \\ \pi_{f,h}^g(\zeta)(t_i) \\ \pi_{g_0,h}(\zeta) \\ \pi_{\psi,h}(\zeta) \\ [\pi_{c,h,j}(\zeta)(t_i)]_{j \in J^\alpha(t_i)} \end{pmatrix}$$

for $\zeta \in \Omega_{\infty,h}$ and $i = 1, \ldots, N$. Now, we can transform Problem 5.5.2 such that the perturbation does not appear in the constraints, in particular, we introduce new variables

$$\begin{pmatrix} z_h \\ w_h \\ v_h \end{pmatrix} := \begin{pmatrix} x_h \\ y_h \\ u_h \end{pmatrix} + \mathcal{I}_{X_{2,h}}^{-1} \circ F_h^{\alpha\star} \circ \left(F_h^\alpha \circ \mathcal{I}_{X_{2,h}}^{-1} \circ F_h^{\alpha\star}\right)^{-1} s(\zeta), \tag{5.5.5}$$

which yields the optimization problem

$$
\begin{aligned}
&\mathbf{OP}(\zeta)\\
&\quad \text{Minimize} \quad \mathcal{J}_h(z_h, w_h, v_h, \zeta),\\
&\quad \text{with respect to} \quad z_h \in W^{n_x}_{1,\infty,h}([0,1]), w_h \in L^{n_y}_{\infty,h}([0,1]), v_h \in L^{n_u}_{\infty,h}([0,1]),\\
&\quad \text{subject to} \quad
\begin{array}{rcl}
z_h'(t_i) &=& A_f(t_i) z_h(t_i) + B_f(t_i) w_h(t_i) + C_f(t_i) v_h(t_i),\\
\mathbf{0}_{\mathbb{R}^{n_y}} &=& \tilde{A}^g_{f,h}(t_i) z_h(t_i) + \tilde{B}^g_{f,h}(t_i) w_h(t_i) + \tilde{C}^g_{f,h}(t_i) v_h(t_i),\\
\mathbf{0}_{\mathbb{R}^{n_y}} &=& E_0 z_h(t_0),\\
\mathbf{0}_{\mathbb{R}^{n_\psi}} &=& \Psi_0 z_h(t_0) + \Psi_1 z_h(t_N),
\end{array}\\
&\quad \text{and} \quad
\begin{array}{l}
c'_{j,x}[t_i] z_h(t_i) + c'_{j,y}[t_i] w_h(t_i)\\
\quad + c'_{j,u}[t_i] v_h(t_i) \begin{cases} = 0, & \text{if } t_i \in \Upsilon^\nu_j \\ \leq 0, & \text{if } t_i \in \Theta^\alpha_j \setminus \Upsilon^\nu_j \end{cases}, \quad j \in J^\alpha(t_i),
\end{array}\\
&\quad \text{for } i = 1, \ldots, N,
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{J}_h(z_h, w_h, v_h, \zeta) &:= \frac{1}{2}\mathcal{P}_h((z_h, w_h, v_h), (z_h, w_h, v_h)) + \mathcal{Q}_h(z_h, w_h, v_h, \zeta),\\
\mathcal{Q}_h(z_h, w_h, v_h, \zeta) &:= \sum_{i=1}^N h \begin{pmatrix} z_h(t_i) \\ w_h(t_i) \\ v_h(t_i) \end{pmatrix}^\top \begin{pmatrix} \pi_{\mathcal{H}_x,h}(\zeta)(t_i) \\ \pi_{\mathcal{H}_y,h}(\zeta)(t_i) \\ \pi_{\mathcal{H}_u,h}(\zeta)(t_i) \end{pmatrix}\\
&\quad + z_h(t_0)^\top \pi_{\vartheta_0,h}(\zeta) + z_h(t_N)^\top \pi_{\vartheta_1,h}(\zeta)\\
&\quad - \mathcal{P}_h\left(\mathcal{I}^{-1}_{X_{2,h}} \circ F^{\alpha\star}_h \circ \left(F^\alpha_h \circ \mathcal{I}^{-1}_{X_{2,h}} \circ F^{\alpha\star}_h\right)^{-1} s(\zeta), (z_h, w_h, v_h)\right).
\end{aligned}
$$

Utilizing the Cauchy-Schwarz inequality yields the existence of a constant $\Gamma_1 \geq 0$ such that

$$
|\mathcal{Q}_h(z_h, w_h, v_h, \zeta)| \leq \Gamma_1 \|(z_h, w_h, v_h)\|_{X_{2,h}} \|\zeta\|_{\Omega_{\infty,h}}
$$

for all $(z_h, w_h, v_h) \in X_{2,h}, \zeta \in \Omega_{\infty,h}$. Since the perturbation appears linearly in the objective function $\mathcal{J}_h(z_h, w_h, v_h, \zeta)$, the properties of Lemma 5.3.1 and Lemma 5.3.2 remain valid. Hence, $\mathbf{OP}(\zeta)$ has a unique minimizer for every $\zeta \in \Omega_{\infty,h}$, which we denote by $(\bar{z}_h(\zeta), \bar{w}_h(\zeta), \bar{v}_h(\zeta))$. Consequently, for every $\zeta \in \Omega_{\infty,h}$ the objective function $\mathcal{J}_h$ satisfies the optimality condition at $(\bar{z}_h(\zeta), \bar{w}_h(\zeta), \bar{v}_h(\zeta))$, i.e.,

$$
\begin{aligned}
0 &\leq \nabla_{(z_h,w_h,v_h)} \mathcal{J}_h(\bar{z}_h(\zeta), \bar{w}_h(\zeta), \bar{v}_h(\zeta), \zeta)^\top ((z_h, w_h, v_h) - (\bar{z}_h(\zeta), \bar{w}_h(\zeta), \bar{v}_h(\zeta)))\\
&= \mathcal{P}_h((\bar{z}_h(\zeta), \bar{w}_h(\zeta), \bar{v}_h(\zeta)), (z_h, w_h, v_h) - (\bar{z}_h(\zeta), \bar{w}_h(\zeta), \bar{v}_h(\zeta)))\\
&\quad + \mathcal{Q}_h(z_h - \bar{z}_h(\zeta), w_h - \bar{w}_h(\zeta), v_h - \bar{v}_h(\zeta), \zeta)
\end{aligned}
$$

for all admissible (z_h, w_h, v_h). Since the constraints in $\mathbf{OP}(\zeta)$ are independent of the perturbation, $(\bar{z}_h(\zeta_1), \bar{w}_h(\zeta_1), \bar{v}_h(\zeta_1))$ is feasible for $\mathbf{OP}(\zeta_2)$ and $(\bar{z}_h(\zeta_2), \bar{w}_h(\zeta_2), \bar{v}_h(\zeta_2))$ is feasible for

$\mathbf{OP}(\zeta_1)$ for all $\zeta_1, \zeta_2 \in \Omega_{\infty,h}$. This in turn implies

$$\begin{aligned}
0 \leq\ & \mathcal{P}_h\left(\left(\bar{z}_h(\zeta_1), \bar{w}_h(\zeta_1), \bar{v}_h(\zeta_1)\right), \left(\bar{z}_h(\zeta_2), \bar{w}_h(\zeta_2), \bar{v}_h(\zeta_2)\right) - \left(\bar{z}_h(\zeta_1), \bar{w}_h(\zeta_1), \bar{v}_h(\zeta_1)\right)\right) \\
& + \mathcal{Q}_h\left(\bar{z}_h(\zeta_2) - \bar{z}_h(\zeta_1), \bar{w}_h(\zeta_2) - \bar{w}_h(\zeta_1), \bar{v}_h(\zeta_2) - \bar{v}_h(\zeta_1), \zeta_1\right) \\
0 \leq\ & \mathcal{P}_h\left(\left(\bar{z}_h(\zeta_2), \bar{w}_h(\zeta_2), \bar{v}_h(\zeta_2)\right), \left(\bar{z}_h(\zeta_1), \bar{w}_h(\zeta_1), \bar{v}_h(\zeta_1)\right) - \left(\bar{z}_h(\zeta_2), \bar{w}_h(\zeta_2), \bar{v}_h(\zeta_2)\right)\right), \\
& + \mathcal{Q}_h\left(\bar{z}_h(\zeta_1) - \bar{z}_h(\zeta_2), \bar{w}_h(\zeta_1) - \bar{w}_h(\zeta_2), \bar{v}_h(\zeta_1) - \bar{v}_h(\zeta_2), \zeta_2\right).
\end{aligned}$$

Adding these inequalities and utilizing the coercivity of $\mathcal{P}_h$ yields

$$\begin{aligned}
& \mathcal{Q}_h\left(\bar{z}_h(\zeta_2) - \bar{z}_h(\zeta_1), \bar{w}_h(\zeta_2) - \bar{w}_h(\zeta_1), \bar{v}_h(\zeta_2) - \bar{v}_h(\zeta_1), \zeta_1 - \zeta_2\right) \\
& \geq \mathcal{P}_h\left(\left(\bar{z}_h(\zeta_2), \bar{w}_h(\zeta_2), \bar{v}_h(\zeta_2)\right) - \left(\bar{z}_h(\zeta_1), \bar{w}_h(\zeta_1), \bar{v}_h(\zeta_1)\right),\right. \\
& \quad \left.\left(\bar{z}_h(\zeta_2), \bar{w}_h(\zeta_2), \bar{v}_h(\zeta_2)\right) - \left(\bar{z}_h(\zeta_1), \bar{w}_h(\zeta_1), \bar{v}_h(\zeta_1)\right)\right) \\
& \geq \tilde{\gamma} \left\|\left(\bar{z}_h(\zeta_2), \bar{w}_h(\zeta_2), \bar{v}_h(\zeta_2)\right) - \left(\bar{z}_h(\zeta_1), \bar{w}_h(\zeta_1), \bar{v}_h(\zeta_1)\right)\right\|_{X_2}^2.
\end{aligned}$$

Exploiting the boundedness of $\mathcal{Q}_h$ results in

$$\left\|\left(\bar{z}_h(\zeta_2), \bar{w}_h(\zeta_2), \bar{v}_h(\zeta_2)\right) - \left(\bar{z}_h(\zeta_1), \bar{w}_h(\zeta_1), \bar{v}_h(\zeta_1)\right)\right\|_{X_2} \leq \frac{\Gamma_1}{\tilde{\gamma}} \left\|\zeta_2 - \zeta_1\right\|_{\Omega_\infty}.$$

With (5.5.4) and (5.5.5) we conclude that there exists a constant $\boldsymbol{l}_1 \geq 0$ such that

$$\left\|\left(\bar{x}_h(\zeta_2), \bar{y}_h(\zeta_2), \bar{u}_h(\zeta_2)\right) - \left(\bar{x}_h(\zeta_1), \bar{y}_h(\zeta_1), \bar{u}_h(\zeta_1)\right)\right\|_{X_2} \leq \boldsymbol{l}_1 \left\|\zeta_2 - \zeta_1\right\|_{\Omega_\infty}$$

for all $\zeta_1, \zeta_2 \in \Omega_{\infty,h}$. Since $\left(\bar{\lambda}_{f,h}(\zeta), \bar{\lambda}_{g,h}(\zeta), \bar{\varsigma}_h(\zeta), \bar{\sigma}_h(\zeta), \bar{\eta}_h^\alpha(\zeta)\right)$ is an element of the Hilbert space $Y_{2,h}^\alpha$, there exists a unique operator $\left(\bar{\lambda}_{f,h}(\zeta), \bar{\lambda}_{g,h}(\zeta), \bar{\varsigma}_h(\zeta), \bar{\sigma}_h(\zeta), \bar{\eta}_h^\alpha(\zeta)\right)^\star$ in the dual space $Y_{2,h}^{\alpha\ \star}$ with equal norm value according to Theorem 2.1.3. This allows us to express the stationarity of the Lagrange function for Problem 5.5.2 as

$$\mathbf{0}_{X_{2,h}^\star} = \mathcal{P}_h\left(\left(\bar{x}_h(\zeta), \bar{y}_h(\zeta), \bar{u}_h(\zeta)\right), \cdot\right) + F_h^{\alpha\star}\left(\left(\bar{\lambda}_{f,h}(\zeta), \bar{\lambda}_{g,h}(\zeta), \bar{\varsigma}_h(\zeta), \bar{\sigma}_h(\zeta), \bar{\eta}_h^\alpha(\zeta)\right)^\star\right)(\cdot).$$

Solving for $\left(\bar{\lambda}_{f,h}(\zeta), \bar{\lambda}_{g,h}(\zeta), \bar{\varsigma}_h(\zeta), \bar{\sigma}_h(\zeta), \bar{\eta}_h^\alpha(\zeta)\right)^\star$ yields

$$\begin{aligned}
& \left(\bar{\lambda}_{f,h}(\zeta), \bar{\lambda}_{g,h}(\zeta), \bar{\varsigma}_h(\zeta), \bar{\sigma}_h(\zeta), \bar{\eta}_h^\alpha(\zeta)\right)^\star(\cdot) \\
& \quad = -\left(F_h^\alpha \circ \mathcal{I}_{X_{2,h}}^{-1} \circ F_h^{\alpha\star}\right)^{-1} \circ F_h^\alpha \circ \mathcal{I}_{X_{2,h}}^{-1} \circ \mathcal{P}_h\left(\left(\bar{x}_h(\zeta), \bar{y}_h(\zeta), \bar{u}_h(\zeta)\right), \cdot\right).
\end{aligned}$$

Thus, by exploiting (5.5.4) we find a constant $\Gamma_2 \geq 0$ such that for all $\zeta_1, \zeta_2 \in \Omega_{\infty,h}$ it holds

$$\begin{aligned}
& \Big\|\left(\bar{\lambda}_{f,h}(\zeta_2), \bar{\lambda}_{g,h}(\zeta_2), \bar{\varsigma}_h(\zeta_2), \bar{\sigma}_h(\zeta_2), \bar{\eta}_h^\alpha(\zeta_2)\right) \\
& \qquad - \left(\bar{\lambda}_{f,h}(\zeta_1), \bar{\lambda}_{g,h}(\zeta_1), \bar{\varsigma}_h(\zeta_1), \bar{\sigma}_h(\zeta_1), \bar{\eta}_h^\alpha(\zeta_1)\right)\Big\|_{Y_2^\alpha} \\
& \quad = \Big\|\left(\bar{\lambda}_{f,h}(\zeta_2), \bar{\lambda}_{g,h}(\zeta_2), \bar{\varsigma}_h(\zeta_2), \bar{\sigma}_h(\zeta_2), \bar{\eta}_h^\alpha(\zeta_2)\right)^\star \\
& \qquad - \left(\bar{\lambda}_{f,h}(\zeta_1), \bar{\lambda}_{g,h}(\zeta_1), \bar{\varsigma}_h(\zeta_1), \bar{\sigma}_h(\zeta_1), \bar{\eta}_h^\alpha(\zeta_1)\right)^\star\Big\|_{Y_{2,h}^{\alpha\ \star}} \\
& \quad \leq \Gamma_2 \left\|\left(\bar{x}_h(\zeta_2), \bar{y}_h(\zeta_2), \bar{u}_h(\zeta_2)\right) - \left(\bar{x}_h(\zeta_1), \bar{y}_h(\zeta_1), \bar{u}_h(\zeta_1)\right)\right\|_{X_2} \leq \Gamma_2 \boldsymbol{l}_1 \left\|\zeta_2 - \zeta_1\right\|_{\Omega_\infty}.
\end{aligned}$$

Finally, by using the difference equation for $\bar{\lambda}_{f,h}$ we obtain a constant $\boldsymbol{l}_2 \geq 0$ with

$$\left\|\bar{\lambda}'_{f,h}(\zeta_2) - \bar{\lambda}'_{f,h}(\zeta_1)\right\|_2 \leq \boldsymbol{l}_2 \|\zeta_2 - \zeta_1\|_{\Omega_\infty},$$

hence the assertion (5.5.3) holds for $\boldsymbol{l}_{L_2} := \max\{\boldsymbol{l}_1, \Gamma_2 \boldsymbol{l}_1, \boldsymbol{l}_2\}$. □

Thus far, we have shown uniform strong regularity in the weaker Ξ_2-norm. In order to obtain uniform strong regularity in the Ξ_∞-norm, we consider a parametric optimization problem depending on $\left(\bar{x}_h(\zeta), \bar{\lambda}_{f,h}(\zeta), \pi(\zeta)\right)$ and apply the sensitivity result in Corollary 2.3.10, which yields the following (compare step (d) above):

Lemma 5.5.4

Let (5.A1) - (5.A5) hold. Then, there exist $h_1, \boldsymbol{l}_{L_\infty}, \varrho > 0$ *such that for all* $0 < h \leq h_1$ *and every* $\zeta_1, \zeta_2 \in \mathcal{B}_\varrho(\mathbf{0}_{\Omega_\infty})$ *it holds*

$$\left\|\bar{\xi}_h(\zeta_1) - \bar{\xi}_h(\zeta_2)\right\|_{\Xi_\infty} \leq \boldsymbol{l}_{L_\infty} \|\zeta_1 - \zeta_2\|_{\Omega_\infty}.$$

Proof. Choose $h_1 > 0$ such that Lemma 5.3.1 and the discrete Legendre-Clebsch condition in Lemma 5.3.3 are satisfied, and (compare Lemma 2.4.5) for all $0 < h \leq h_1$, $i = 1, \ldots, N$, and every $\varpi \in \mathbb{R}^{n_y} \times \mathbb{R}^{j^\alpha(t_i)}$ it holds

$$\left\| \begin{bmatrix} \tilde{B}^g_{f,h}(t_i) & \tilde{C}^g_{f,h}(t_i) \\ B^\alpha_c(t_i) & C^\alpha_c(t_i) \end{bmatrix}^\top \varpi \right\| \geq \frac{\beta}{2} \|\varpi\|. \tag{5.5.6}$$

For arbitrary $0 < h \leq h_1$ and $i \in \{1, \ldots, N\}$ we consider the parametric optimization problem depending on $\chi(\zeta)(t_i) := \left(\bar{x}_h(\zeta)(t_i), \bar{\lambda}_{f,h}(\zeta)(t_{i-1}), \pi(\zeta)(t_i)\right)$:

LQP $(\chi(\zeta)(t_i))$

Minimize
$$\frac{1}{2}\begin{pmatrix} w \\ v \end{pmatrix}^\top \begin{bmatrix} \nabla^2_{yy}\widetilde{\mathcal{H}}_h[t_i, t_{i-1}] & \nabla^2_{yu}\widetilde{\mathcal{H}}_h[t_i, t_{i-1}] \\ \nabla^2_{uy}\widetilde{\mathcal{H}}_h[t_i, t_{i-1}] & \nabla^2_{uu}\widetilde{\mathcal{H}}_h[t_i, t_{i-1}] \end{bmatrix} \begin{pmatrix} w \\ v \end{pmatrix}$$
$$+ \begin{pmatrix} w \\ v \end{pmatrix}^\top \begin{bmatrix} \nabla^2_{yx}\widetilde{\mathcal{H}}_h[t_i, t_{i-1}] & B_f(t_i)^\top \\ \nabla^2_{ux}\widetilde{\mathcal{H}}_h[t_i, t_{i-1}] & C_f(t_i)^\top \end{bmatrix} \begin{pmatrix} \bar{x}_h(\zeta)(t_i) \\ \bar{\lambda}_{f,h}(\zeta)(t_{i-1}) \end{pmatrix}$$
$$+ \begin{pmatrix} w \\ v \end{pmatrix}^\top \begin{pmatrix} \pi_{\mathcal{H}_y,h}(\zeta)(t_i) \\ \pi_{\mathcal{H}_u,h}(\zeta)(t_i) \end{pmatrix},$$

with respect to $w \in \mathbb{R}^{n_y}, v \in \mathbb{R}^{n_u}$,

subject to
$$\mathbf{0}_{\mathbb{R}^{n_y}} = \tilde{A}^g_{f,h}(t_i)\,\bar{x}_h(\zeta)(t_i) + \tilde{B}^g_{f,h}(t_i)\,w + \tilde{C}^g_{f,h}(t_i)\,v + \pi^g_{f,h}(\zeta)(t_i),$$

and
$$c'_{j,x}[t_i]\,\bar{x}_h(\zeta)(t_i) + c'_{j,y}[t_i]\,w + c'_{j,u}[t_i]\,v + \pi_{c,h,j}(\zeta)(t_i) \begin{cases} = 0, & \text{if } t_i \in \Upsilon^\nu_j \\ \leq 0, & \text{if } t_i \in \Theta^\alpha_j \setminus \Upsilon^\nu_j \end{cases}, \quad j \in J^\alpha(t_i).$$

(5.5.6) and the discrete Legendre-Clebsch condition correspond to the linear independence constraint qualification in Definition 2.3.3 and the second-order sufficient conditions in Theorem 2.3.5, respectively. Hence, $(\bar{y}_h(\zeta)(t_i), \bar{u}_h(\zeta)(t_i))$ is the unique minimizer of $\mathbf{LQP}(\chi(\zeta)(t_i))$ together with the unique multiplier $\left(\bar{\lambda}_{g,h}(\zeta)(t_{i-1}), \bar{\eta}_h^{\alpha}(\zeta)(t_i)\right)$. Moreover, (5.5.6) and the discrete Legendre-Clebsch condition coincide with (i) and (ii) in Corollary 2.3.10, respectively. Thus, for an arbitrary $\tilde{\varrho} > 0$, there exist Lipschitz continuous functions

$$
\begin{aligned}
w &: \mathcal{B}_{\tilde{\varrho}}(\chi(\mathbf{0}_{\Omega_\infty})(t_i)) \to \mathbb{R}^{n_y}, & v &: \mathcal{B}_{\tilde{\varrho}}(\chi(\mathbf{0}_{\Omega_\infty})(t_i)) \to \mathbb{R}^{n_u},\\
\lambda_g &: \mathcal{B}_{\tilde{\varrho}}(\chi(\mathbf{0}_{\Omega_\infty})(t_i)) \to \mathbb{R}^{n_y}, & \eta^{\alpha} &: \mathcal{B}_{\tilde{\varrho}}(\chi(\mathbf{0}_{\Omega_\infty})(t_i)) \to \mathbb{R}^{j^{\alpha}(t_i)}
\end{aligned}
$$

such that $(w(\chi(\zeta)(t_i)), v(\chi(\zeta)(t_i)))$ is the unique solution of $\mathbf{LQP}(\chi(\zeta)(t_i))$ together with the unique Lagrange multiplier $(\lambda_g(\chi(\zeta)(t_i)), \eta^{\alpha}(\chi(\zeta)(t_i)))$ for each $\chi(\zeta)(t_i) \in \mathcal{B}_{\tilde{\varrho}}(\chi(\mathbf{0}_{\Omega_\infty})(t_i))$. Note that $\pi(\zeta_1) - \pi(\zeta_2) = \zeta_1 - \zeta_2$ is satisfied for all $\zeta_1, \zeta_2 \in \Omega_{\infty,h}$. Utilizing the Sobolev inequality in Lemma A.7 and the result of Lemma 5.5.3 yield

$$
\begin{aligned}
\|\bar{x}_h(\zeta_1)(t_i) - \bar{x}_h(\zeta_2)(t_i)\| &\le 2\|\bar{x}_h(\zeta_1) - \bar{x}_h(\zeta_2)\|_{1,2} \le 2\boldsymbol{l}_{L_2}\|\zeta_1 - \zeta_2\|_{\Omega_\infty}, \qquad (5.5.7)\\
\left\|\bar{\lambda}_{f,h}(\zeta_1)(t_{i-1}) - \bar{\lambda}_{f,h}(\zeta_2)(t_{i-1})\right\| &\le 2\left\|\bar{\lambda}_{f,h}(\zeta_1) - \bar{\lambda}_{f,h}(\zeta_2)\right\|_{1,2} \le 2\boldsymbol{l}_{L_2}\|\zeta_1 - \zeta_2\|_{\Omega_\infty}.
\end{aligned}
$$

Therefore, we find a $\varrho > 0$ such that

$$
\chi(\zeta)(t_i) = \left(\bar{x}_h(\zeta)(t_i), \bar{\lambda}_{f,h}(\zeta)(t_{i-1}), \pi(\zeta)(t_i)\right) \in \mathcal{B}_{\tilde{\varrho}}(\chi(\mathbf{0}_{\Omega_\infty})(t_i))
$$

for every $\zeta \in \mathcal{B}_{\varrho}(\mathbf{0}_{\Omega_\infty})$. Since $(\bar{y}_h(\zeta)(t_i), \bar{u}_h(\zeta)(t_i))$ is the unique minimizer of $\mathbf{LQP}(\chi(\zeta)(t_i))$ and $\left(\bar{\lambda}_{g,h}(\zeta)(t_{i-1}), \bar{\eta}_h^{\alpha}(\zeta)(t_i)\right)$ is the associated, unique Lagrange multiplier, for every perturbation $\zeta \in \mathcal{B}_{\varrho}(\mathbf{0}_{\Omega_\infty})$ it holds

$$
\begin{aligned}
&(w(\chi(\zeta)(t_i)), v(\chi(\zeta)(t_i)), \lambda_g(\chi(\zeta)(t_i)), \eta^{\alpha}(\chi(\zeta)(t_i)))\\
&\quad = \left(\bar{y}_h(\zeta)(t_i), \bar{u}_h(\zeta)(t_i), \bar{\lambda}_{g,h}(\zeta)(t_{i-1}), \bar{\eta}_h^{\alpha}(\zeta)(t_i)\right).
\end{aligned}
$$

Exploiting the Lipschitz continuity of $(w(\chi(\zeta)(t_i)), v(\chi(\zeta)(t_i)), \lambda_g(\chi(\zeta)(t_i)), \eta^{\alpha}(\chi(\zeta)(t_i)))$ we find a constant $\boldsymbol{l}_1 \ge 0$ satisfying

$$
\begin{aligned}
&\Big\|\left(\bar{y}_h(\zeta_1)(t_i), \bar{u}_h(\zeta_1)(t_i), \bar{\lambda}_{g,h}(\zeta_1)(t_{i-1}), \bar{\eta}_h^{\alpha}(\zeta_1)(t_i)\right)\\
&\qquad - \left(\bar{y}_h(\zeta_2)(t_i), \bar{u}_h(\zeta_2)(t_i), \bar{\lambda}_{g,h}(\zeta_2)(t_{i-1}), \bar{\eta}_h^{\alpha}(\zeta_2)(t_i)\right)\Big\|\\
&\quad \le \boldsymbol{l}_1 \max\Big\{\|\bar{x}_h(\zeta_1)(t_i) - \bar{x}_h(\zeta_2)(t_i)\|, \left\|\bar{\lambda}_{f,h}(\zeta_1)(t_{i-1}) - \bar{\lambda}_{f,h}(\zeta_2)(t_{i-1})\right\|,\\
&\qquad\qquad \|\pi(\zeta_1)(t_i) - \pi(\zeta_2)(t_i)\|\Big\}
\end{aligned}
$$

for every $\zeta_1, \zeta_2 \in \mathcal{B}_{\varrho}(\mathbf{0}_{\Omega_\infty})$. Using (5.5.7) we conclude there exists a $\boldsymbol{l}_2 \ge 0$ such that

$$
\begin{aligned}
&\Big\|\left(\bar{y}_h(\zeta_1)(t_i), \bar{u}_h(\zeta_1)(t_i), \bar{\lambda}_{g,h}(\zeta_1)(t_{i-1}), \bar{\eta}_h^{\alpha}(\zeta_1)(t_i)\right)\\
&\qquad - \left(\bar{y}_h(\zeta_2)(t_i), \bar{u}_h(\zeta_2)(t_i), \bar{\lambda}_{g,h}(\zeta_2)(t_{i-1}), \bar{\eta}_h^{\alpha}(\zeta_2)(t_i)\right)\Big\| \le \boldsymbol{l}_2\|\zeta_1 - \zeta_2\|_{\Omega_\infty}.
\end{aligned}
$$

Finally, utilizing the difference equation for $\bar{x}_h$ and $\bar{\lambda}_{f,h}$ we find $\boldsymbol{l}_{L_\infty} \ge 0$ satisfying

$$
\left\|\bar{\xi}_h(\zeta_1) - \bar{\xi}_h(\zeta_2)\right\|_{\Xi_\infty} \le \boldsymbol{l}_{L_\infty}\|\zeta_1 - \zeta_2\|_{\Omega_\infty}
$$

for every $\zeta_1, \zeta_2 \in \mathcal{B}_{\varrho}(\mathbf{0}_{\Omega_\infty})$, which completes the proof. □

We now have uniform strong regularity for the solution and multipliers $\bar{\xi}_h(\zeta)$ of the modified system in Problem 5.5.2. It remains to show that for sufficiently small h and ζ, $\bar{\xi}_h(\zeta)$ is also the (unique) solution of Problem 5.5.1 (compare step (e) above).

Lemma 5.5.5
Let (5.A1) - (5.A5) hold. Then, there exist $h_1, \varrho, \rho > 0$ such that for all $0 < h \leq h_1$ and every $\zeta \in \mathcal{B}_\varrho(\mathbf{0}_{\Omega_\infty})$, $(\bar{x}_h(\zeta), \bar{y}_h(\zeta), \bar{u}_h(\zeta))$ is the unique minimizer of Problem 5.5.1 together with the unique Lagrange multipliers $\left(\bar{\lambda}_{f,h}(\zeta), \bar{\lambda}_{g,h}(\zeta), \bar{\varsigma}_h(\zeta), \bar{\sigma}_h(\zeta), \bar{\eta}_h(\zeta)\right)$, and it holds

$$\left(\bar{x}_h(\zeta), \bar{y}_h(\zeta), \bar{u}_h(\zeta), \bar{\lambda}_{f,h}(\zeta), \bar{\lambda}_{g,h}(\zeta), \bar{\varsigma}_h(\zeta), \bar{\sigma}_h(\zeta), \bar{\eta}_h(\zeta)\right) \in \mathcal{B}_\rho\left(\Delta_h\left(\hat{\xi}\right)\right).$$

Proof. Choose $h_1, \varrho, \rho > 0$ such that Lemma 5.5.4 holds with constant $\boldsymbol{l}_{L_\infty}$ and

$$\begin{aligned}
&\varrho \leq \min\left\{\frac{\alpha}{2\left(3\boldsymbol{l}_{L_\infty}\left(\|A_c\|_\infty + \|B_c\|_\infty + \|C_c\|_\infty\right) + 1\right)}, \frac{\nu}{2\boldsymbol{l}_{L_\infty}}\right\}, \\
&\rho := \boldsymbol{l}_{L_\infty}\varrho.
\end{aligned}$$

Let $\zeta \in \mathcal{B}_\varrho(\mathbf{0}_{\Omega_\infty})$ be arbitrary. We recall

$$\pi_{c,h}(\zeta)(t_i) = c[t_i] - c'_x[t_i]\hat{x}(t_i) - c'_y[t_i]\hat{y}(t_i) - c'_u[t_i]\hat{u}(t_i) - \zeta_c(t_i), \quad i = 1, \ldots, N,$$

which for every $j \in J$, $i \in \{1, \ldots, N\}$ gives us

$$\begin{aligned}
&c'_{j,x}[t_i]\bar{x}_h(\zeta)(t_i) + c'_{j,y}[t_i]\bar{y}_h(\zeta)(t_i) + c'_{j,u}[t_i]\bar{u}_h(\zeta)(t_i) + \pi_{c,h,j}(\zeta)(t_i) \\
&\quad = c'_{j,x}[t_i]\left(\bar{x}_h(\zeta)(t_i) - \hat{x}(t_i)\right) + c'_{j,y}[t_i]\left(\bar{y}_h(\zeta)(t_i) - \hat{y}(t_i)\right) \\
&\qquad + c'_{j,u}[t_i]\left(\bar{u}_h(\zeta)(t_i) - \hat{u}(t_i)\right) + c_j[t_i] - \zeta_{c,j}(t_i) \\
&\quad \leq \left(\|A_c\|_\infty + \|B_c\|_\infty + \|C_c\|_\infty\right)\left(\|\bar{x}_h(\zeta)(t_i) - \hat{x}(t_i)\| + \|\bar{y}_h(\zeta)(t_i) - \hat{y}(t_i)\|\right. \\
&\qquad \left. + \|\bar{u}_h(\zeta)(t_i) - \hat{u}(t_i)\|\right) + c_j[t_i] + \|\zeta\|_{\Omega_\infty} \\
&\quad \leq \left(\|A_c\|_\infty + \|B_c\|_\infty + \|C_c\|_\infty\right)3\boldsymbol{l}_{L_\infty}\|\zeta\|_{\Omega_\infty} + c_j[t_i] + \|\zeta\|_{\Omega_\infty} \\
&\quad \leq \frac{\alpha}{2} + c_j[t_i]
\end{aligned}$$

according to Lemma 5.5.4 and the choice of ϱ. Then, for any $j \in J$, $i \in \{1, \ldots, N\}$ with $t_i \notin \Theta_j^\alpha$ it holds

$$\begin{aligned}
&c'_{j,x}[t_i]\bar{x}_h(\zeta)(t_i) + c'_{j,y}[t_i]\bar{y}_h(\zeta)(t_i) + c'_{j,u}[t_i]\bar{u}_h(\zeta)(t_i) + \pi_{c,h,j}(\zeta)(t_i) \\
&\quad \leq \frac{\alpha}{2} + c_j[t_i] < \frac{\alpha}{2} - \alpha = -\frac{\alpha}{2} < 0.
\end{aligned}$$

Moreover, for any $j \in J$, $i \in \{1, \ldots, N\}$ with $t_i \in \Upsilon_j^\nu$ we have

$$\nu - \bar{\eta}_{h,j}(\zeta)(t_i) < \hat{\eta}_j(t_i) - \bar{\eta}_{h,j}(\zeta)(t_i) \leq |\hat{\eta}_j(t_i) - \bar{\eta}_{h,j}(\zeta)(t_i)| \leq \boldsymbol{l}_{L_\infty}\|\zeta\|_{\Omega_\infty} \leq \frac{\nu}{2},$$

hence $\bar{\eta}_{h,j}(\zeta)(t_i) > \frac{\nu}{2} > 0$. This implies that $(\bar{x}_h(\zeta), \bar{y}_h(\zeta), \bar{u}_h(\zeta))$ is admissible for Problem 5.5.1. Furthermore, by Lemma 5.3.2 and Lemma 5.3.1, the sufficient conditions of Theorem 2.3.5 and the linear constraint qualification of Definition 2.3.3 are satisfied for $(\bar{x}_h(\zeta), \bar{y}_h(\zeta), \bar{u}_h(\zeta))$.

Thus, it is the unique minimizer of Problem 5.5.1 and the associated Lagrange multipliers $\left(\bar{\lambda}_{f,h}(\zeta), \bar{\lambda}_{g,h}(\zeta), \bar{\varsigma}_h(\zeta), \bar{\sigma}_h(\zeta), \bar{\eta}_h(\zeta)\right)$ are unique for every $\zeta \in \mathcal{B}_\varrho(\mathbf{0}_{\Omega_\infty})$ according to Remark 2.3.4. Finally, since $\bar{\xi}_h(\zeta) = \left(\bar{x}_h(\zeta), \bar{y}_h(\zeta), \bar{u}_h(\zeta), \bar{\lambda}_{f,h}(\zeta), \bar{\lambda}_{g,h}(\zeta), \bar{\varsigma}_h(\zeta), \bar{\sigma}_h(\zeta), \bar{\eta}_h(\zeta)\right)$ and $\bar{\xi}_h(\mathbf{0}_{\Omega_\infty}) = \Delta_h\left(\hat{\xi}\right)$ it holds

$$\left\|\bar{\xi}_h(\zeta) - \Delta_h\left(\hat{\xi}\right)\right\|_{\Xi_\infty} \leq \boldsymbol{l}_{L_\infty} \|\zeta\|_{\Omega_\infty} \leq \boldsymbol{l}_{L_\infty}\varrho = \rho,$$

which completes the proof. □

We summarize the main convergence result of this chapter in a final theorem, which establishes the convergence of the solution of Problem 5.2.3 to the solution of Problem 5.1.1 by applying Theorem 2.2.6:

Theorem 5.5.6 (Convergence)
Let (5.A1) - (5.A5) hold. Then, there exist $\hat{h}, \boldsymbol{l} > 0$ such that for every $0 < h \leq \hat{h}$, Problem 5.2.3 has a unique solution and associated Lagrange multipliers that converge linearly to the weak local minimizer of Problem 5.1.1 and the associated Lagrange multipliers with respect to the L_∞-norm.

Remark 5.5.7
Please note that convergence was shown for the multipliers associated with Problem 5.2.3. In Section 5.7, a relationship between the respective multipliers of Problem 5.2.1 and Problem 5.2.3 is derived. However, it is unclear, if the multipliers associated with Problem 5.2.1 are convergent, since a jump condition occurs for the adjoint multipliers. Therefore, convergence with respect to the L_∞-norm can not be expected.

5.6 Example

Consider the implicit Euler discretization of Example 3.3.1

$$\begin{array}{lrcll}
\text{Minimize} & x_{4,h}(t_N), & & & \\
\text{subject to} & x'_{1,h}(t_i) & = & u_h(t_i) - y_h(t_i), & 0 = x_{1,h}(t_0), \quad 0 = x_{1,h}(t_N), \\
 & x'_{2,h}(t_i) & = & u_h(t_i), & 1 = x_{2,h}(t_0), \quad -1 = x_{2,h}(t_N), \\
 & x'_{3,h}(t_i) & = & -x_{2,h}(t_i), & \\
 & x'_{4,h}(t_i) & = & \frac{1}{2}u_h(t_i)^2, & 0 = x_{4,h}(t_0), \\
 & 0 & = & x_{1,h}(t_i) + x_{3,h}(t_i), &
\end{array}$$

and the system with a discrete index reduction as described in Problem 5.2.3

(**RDOCP – 1**)

$$\begin{array}{lrcll}
\text{Minimize} & x_{4,h}(t_N), & & & \\
\text{subject to} & x'_{1,h}(t_i) & = & u_h(t_i) - y_h(t_i), & 0 = x_{1,h}(t_0), \quad 0 = x_{1,h}(t_N), \\
 & x'_{2,h}(t_i) & = & u_h(t_i), & 1 = x_{2,h}(t_0), \quad -1 = x_{2,h}(t_N), \\
 & x'_{3,h}(t_i) & = & -x_{2,h}(t_i), & \\
 & x'_{4,h}(t_i) & = & \frac{1}{2}u_h(t_i)^2, & 0 = x_{4,h}(t_0), \\
 & 0 & = & u_h(t_i) - y_h(t_i) - x_{2,h}(t_i), & \\
 & 0 & = & x_{1,h}(t_0) + x_{3,h}(t_0). &
\end{array}$$

For problem (**RDOCP − 1**) we want to compare the solution of the (discrete) KKT-conditions with the continuous KKT-point of Example 3.3.1, in order to illustrate that the discrete Lagrange multipliers converge to the continuous multipliers. The necessary conditions for (**RDOCP − 1**) read as

$$\begin{array}{rcll} \lambda'_{f_1,h}(t_i) &=& 0, & \lambda_{f_1,h}(t_0) = -\sigma_{1,h}, \\ \lambda'_{f_2,h}(t_i) &=& \lambda_{f_3,h}(t_{i-1}) + \lambda_{g,h}(t_{i-1}), & \lambda_{f_2,h}(t_0) = -\sigma_{2,h}, \\ \lambda'_{f_3,h}(t_i) &=& 0, & \lambda_{f_3,h}(t_N) = 0, \\ \lambda'_{f_4,h}(t_i) &=& 0, & \lambda_{f_4,h}(t_N) = 1, \\ 0 &=& -\lambda_{f_1,h}(t_{i-1}) - \lambda_{g,h}(t_{i-1}), & \\ 0 &=& \lambda_{f_1,h}(t_{i-1}) + \lambda_{f_2,h}(t_{i-1}) + \lambda_{f_4,h}(t_{i-1})\, u_h(t_i) + \lambda_{g,h}(t_{i-1}) & \end{array}$$

for $i = 1, \ldots, N$. The KKT-conditions have the solution

$$\begin{array}{rcl} x_{1,h}(t_i) &=& -\frac{2h}{1-h^2} t_i^3 - \frac{(1-h)(1-2h)}{1-h^2} t_i^2 + \frac{1-h+2h^2}{1-h^2} t_i, \\ x_{2,h}(t_i) &=& -\frac{6h}{1-h^2} t_i^2 - 2\frac{1-3h-h^2}{1-h^2} t_i + 1, \\ x_{3,h}(t_i) &=& \frac{2h}{1-h^2} t_i^3 + \frac{(1-h)(1-2h)}{1-h^2} t_i^2 - \frac{1-h+2h^2}{1-h^2} t_i, \\ x_{4,h}(t_i) &=& \frac{24h^2}{(1-h^2)^2} t_i^3 + \frac{12h\left(1-3h-h^2\right)}{(1-h^2)^2} t_i^2 + 2\frac{1-6h+7h^2+6h^3-2h^4}{(1-h^2)^2} t_i, \\ y_h(t_i) &=& \frac{6h}{1-h^2} t_i^2 + 2\frac{1-9h-h^2}{1-h^2} t_i - 3\frac{(1+h)(1-3h)}{1-h^2}, \\ u_h(t_i) &=& -\frac{12h}{1-h^2} t_i - 2\frac{1-3h-4h^2}{1-h^2}, \\ && \\ \lambda_{f_1,h}(t_i) &=& -\frac{12h}{1-h^2}, \\ \lambda_{f_2,h}(t_i) &=& \frac{12h}{1-h^2} t_i + 2\frac{(1-h)(1-2h)}{1-h^2}, \\ \lambda_{f_3,h}(t_i) &=& 0, \\ \lambda_{f_4,h}(t_i) &=& 1, \\ \lambda_{g,h}(t_i) &=& \frac{12h}{1-h^2}, \end{array}$$

which compared to the continuous solution we calculated in Section 3.3 yields

$$\begin{array}{rcl} |x_1(t_i) - x_{1,h}(t_i)| &\leq& 6h, \\ |x_2(t_i) - x_{2,h}(t_i)| &\leq& 13h, \\ |x_3(t_i) - x_{3,h}(t_i)| &\leq& 6h, \\ |x_4(t_i) - x_{4,h}(t_i)| &\leq& 28h, \\ |y(t_i) - y_h(t_i)| &\leq& 34h, \\ |u(t_i) - u_h(t_i)| &\leq& 21h, \\ && \\ |\lambda_{f_1}(t_i) - \lambda_{f_1,h}(t_i)| &\leq& 13h, \\ |\lambda_{f_2}(t_i) - \lambda_{f_2,h}(t_i)| &\leq& 20h, \\ |\lambda_{f_3}(t_i) - \lambda_{f_3,h}(t_i)| &=& 0, \\ |\lambda_{f_4}(t_i) - \lambda_{f_4,h}(t_i)| &=& 0, \\ |\lambda_g(t_i) - \lambda_{g,h}(t_i)| &\leq& 13h, \end{array}$$

for $i = 1, \ldots, N$, $h \leq \frac{1}{4}$. Therefore, the solution of (**RDOCP − 1**), as well as the associated Lagrange multipliers converge linearly to the continuous solution and its associated multipliers (compare Figure 5.2, Figure 5.3, and Figure 5.4).

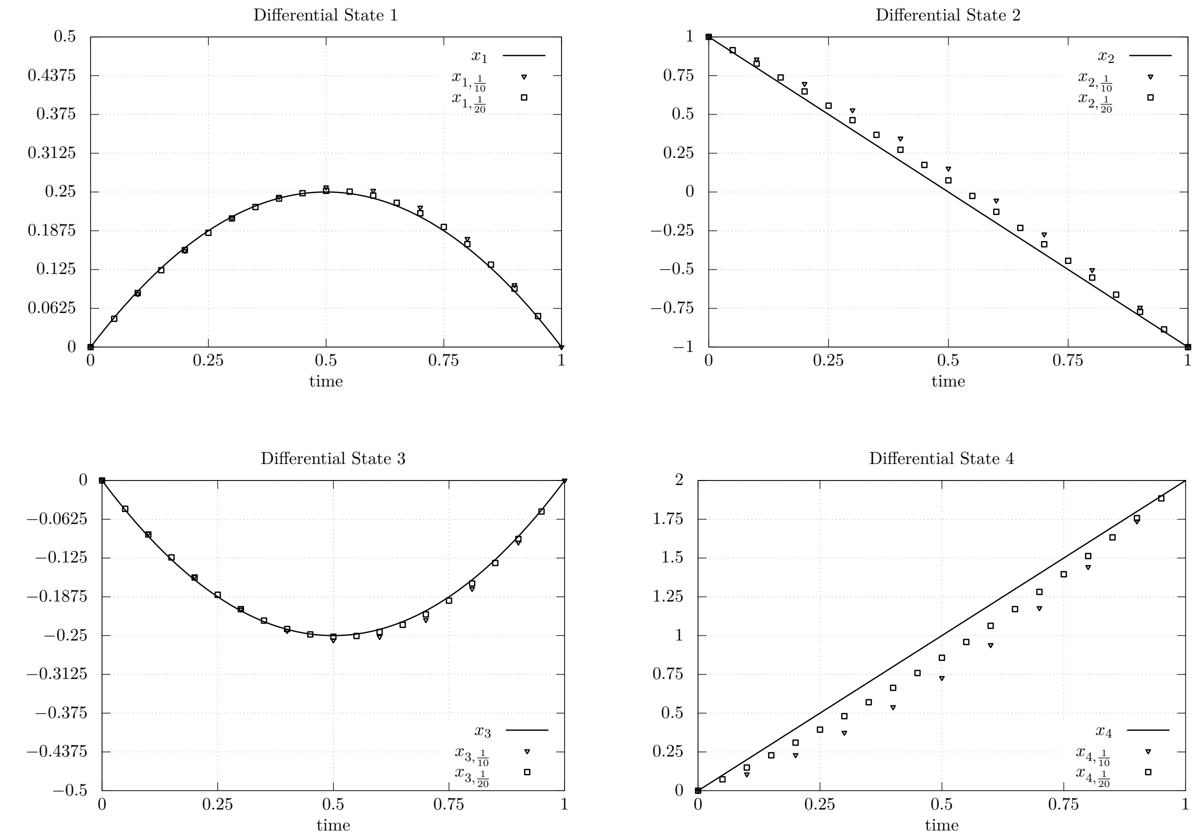

Figure 5.2: Comparison of differential states for $N = 10$ and $N = 20$.

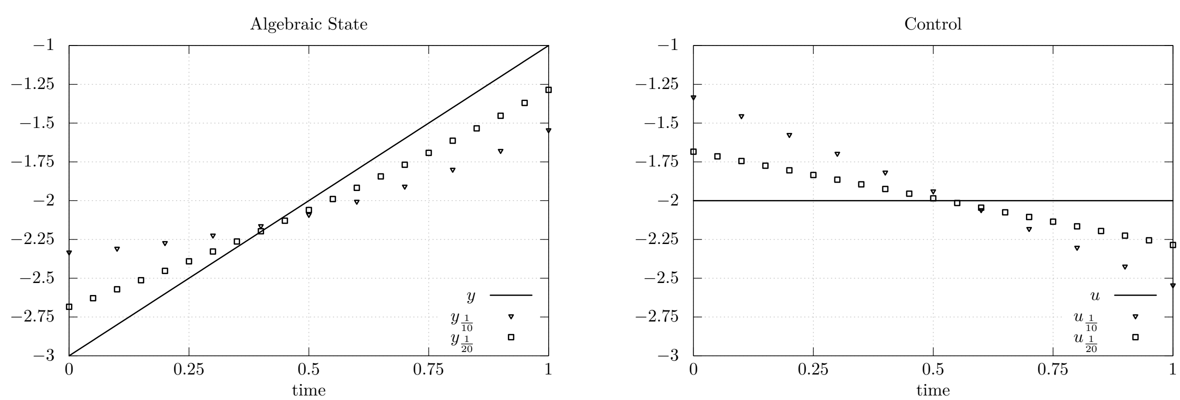

Figure 5.3: Comparison of algebraic states and controls for $N = 10$ and $N = 20$.

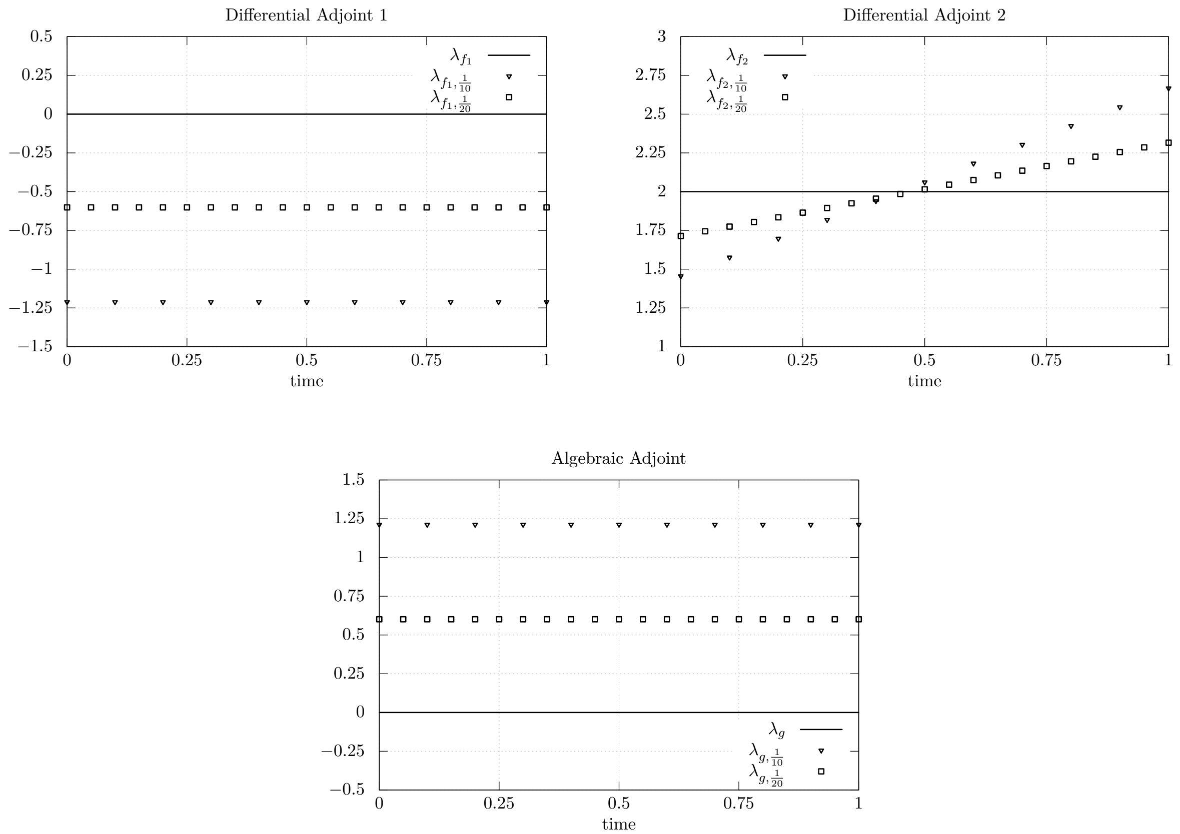

Figure 5.4: Comparison of multipliers for $N = 10$ and $N = 20$.

5.7 Relationship between Discrete Multipliers

In this section, we aim to derive a relationship between the multipliers, which satisfy the respective necessary conditions of

$$
\begin{array}{lllll}
\text{Minimize} & \multicolumn{4}{l}{\varphi\left(x_0, x_N\right),} \\
\text{with respect to} & \multicolumn{4}{l}{x_0 \in \mathbb{R}^{n_x}, \left(x_i, y_i, u_i\right) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u}, i = 1, \ldots, N,} \\
\text{subject to} & \frac{x_i - x_{i-1}}{h} & = & f\left(x_i, y_i, u_i\right), & i = 1, \ldots, N, \\
 & \mathbf{0}_{\mathbb{R}^{n_y}} & = & g\left(x_i\right), & i = 0, 1, \ldots, N, \\
 & \mathbf{0}_{\mathbb{R}^{n_\psi}} & = & \psi\left(x_0, x_N\right), & \\
 & \mathbf{0}_{\mathbb{R}^{n_c}} & \geq & c\left(x_i, y_i, u_i\right), & i = 1, \ldots, N,
\end{array}
\tag{5.7.1}
$$

and the reduced problem

$$\begin{array}{lll} \text{Minimize} & \varphi\left(x_0, x_N\right), & \\[1em] \text{with respect to} & x_0 \in \mathbb{R}^{n_x}, \left(x_i, y_i, u_i\right) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u}, i = 1, \dots, N, & \\[1em] \text{subject to} & \begin{array}{rcll} \frac{x_i - x_{i-1}}{h} & = & f\left(x_i, y_i, u_i\right), & i = 1, \dots, N, \\ \mathbf{0}_{\mathbb{R}^{n_y}} & = & \tilde{g}_h\left(x_i, y_i, u_i\right), & i = 1, \dots, N, \\ \mathbf{0}_{\mathbb{R}^{n_y}} & = & g\left(x_0\right), & \\ \mathbf{0}_{\mathbb{R}^{n_\psi}} & = & \psi\left(x_0, x_N\right), & \\ \mathbf{0}_{\mathbb{R}^{n_c}} & \geq & c\left(x_i, y_i, u_i\right), & i = 1, \dots, N, \end{array} & \end{array} \tag{5.7.2}$$

where $\tilde{g}_h\left(x, y, u\right) := \frac{g(x) - g(x - hf(x,y,u))}{h}$. We recall the necessary conditions of (5.7.1)

$$\begin{array}{rcll} \frac{\breve{\lambda}_{f,i} - \breve{\lambda}_{f,i-1}}{h} & = & -\nabla_x \breve{\mathcal{H}}\left(x_i, y_i, u_i, \breve{\lambda}_{f,i-1}, \breve{\lambda}_{g,i-1}, \breve{\eta}_i\right), & i = 1, \dots, N, \\ \mathbf{0}_{\mathbb{R}^{n_y}} & = & \nabla_y \breve{\mathcal{H}}\left(x_i, y_i, u_i, \breve{\lambda}_{f,i-1}, \breve{\lambda}_{g,i-1}, \breve{\eta}_i\right), & i = 1, \dots, N, \\ \breve{\lambda}_{f,0} & = & -\breve{\ell}_0 \nabla_{x_0} \varphi\left(x_0, x_N\right) - \psi'_{x_0}\left(x_0, x_N\right)^\top \breve{\sigma} - g'\left(x_0\right)^\top \breve{\varsigma}, & \\ \breve{\lambda}_{f,N} & = & \breve{\ell}_0 \nabla_{x_1} \varphi\left(x_0, x_N\right) + \psi'_{x_1}\left(x_0, x_N\right)^\top \breve{\sigma}, & \\ \mathbf{0}_{\mathbb{R}^{n_u}} & = & \nabla_u \breve{\mathcal{H}}\left(x_i, y_i, u_i, \breve{\lambda}_{f,i-1}, \breve{\lambda}_{g,i-1}, \breve{\eta}_i\right), & i = 1, \dots, N, \\ 0 & = & c\left(x_i, y_i, u_i\right)^\top \breve{\eta}_i, \quad \breve{\eta}_i \geq 0, & i = 1, \dots, N, \end{array} \tag{5.7.3}$$

where the (augmented) Hamilton function is defined by

$$\breve{\mathcal{H}}\left(x, y, u, \breve{\lambda}_f, \breve{\lambda}_g, \breve{\eta}\right) := \breve{\lambda}_f^\top f\left(x, y, u\right) + \breve{\lambda}_g^\top g\left(x\right) + \breve{\eta}^\top c\left(x, y, u\right), \tag{5.7.4}$$

and the necessary conditions of (5.7.2)

$$\begin{array}{rcll} \frac{\lambda_{f,i} - \lambda_{f,i-1}}{h} & = & -\nabla_x \widetilde{\mathcal{H}}_h\left(x_i, y_i, u_i, \lambda_{f,i-1}, \lambda_{g,i-1}, \eta_i\right), & i = 1, \dots, N, \\ \mathbf{0}_{\mathbb{R}^{n_y}} & = & \nabla_y \widetilde{\mathcal{H}}_h\left(x_i, y_i, u_i, \lambda_{f,i-1}, \lambda_{g,i-1}, \eta_i\right), & i = 1, \dots, N, \\ \lambda_{f,0} & = & -\ell_0 \nabla_{x_0} \varphi\left(x_0, x_N\right) - \psi'_{x_0}\left(x_0, x_N\right)^\top \sigma - g'\left(x_0\right)^\top \varsigma, & \\ \lambda_{f,N} & = & \ell_0 \nabla_{x_1} \varphi\left(x_0, x_N\right) + \psi'_{x_1}\left(x_0, x_N\right)^\top \sigma, & \\ \mathbf{0}_{\mathbb{R}^{n_u}} & = & \nabla_u \widetilde{\mathcal{H}}_h\left(x_i, y_i, u_i, \lambda_{f,i-1}, \lambda_{g,i-1}, \eta_i\right), & i = 1, \dots, N, \\ 0 & = & c\left(x_i, y_i, u_i\right)^\top \eta_i, \quad \eta_i \geq 0, & i = 1, \dots, N, \end{array} \tag{5.7.5}$$

where the (augmented) Hamilton function is defined by

$$\widetilde{\mathcal{H}}_h\left(x, y, u, \lambda_f, \lambda_g, \eta\right) := \lambda_f^\top f\left(x, y, u\right) + \lambda_g^\top \tilde{g}_h\left(x, y, u\right) + \eta^\top c\left(x, y, u\right). \tag{5.7.6}$$

In order to derive a relationship, we first consider the adjoint algebraic equation and stationarity condition in (5.7.3)

$$\begin{array}{rcl} \mathbf{0}_{\mathbb{R}^{n_y}} & = & \nabla_y \breve{\mathcal{H}}\left(x_i, y_i, u_i, \breve{\lambda}_{f,i-1}, \breve{\lambda}_{g,i-1}, \breve{\eta}_i\right) \\ & = & f'_y\left(x_i, y_i, u_i\right)^\top \breve{\lambda}_{f,i-1} + c'_y\left(x_i, y_i, u_i\right)^\top \breve{\eta}_i, \\ \mathbf{0}_{\mathbb{R}^{n_u}} & = & \nabla_u \breve{\mathcal{H}}\left(x_i, y_i, u_i, \breve{\lambda}_{f,i-1}, \breve{\lambda}_{g,i-1}, \breve{\eta}_i\right) \\ & = & f'_u\left(x_i, y_i, u_i\right)^\top \breve{\lambda}_{f,i-1} + c'_u\left(x_i, y_i, u_i\right)^\top \breve{\eta}_i \end{array}$$

for $i = 1, \ldots, N$, and in (5.7.5)

$$\begin{aligned}
\mathbf{0}_{\mathbb{R}^{n_y}} &= \nabla_y \widetilde{\mathcal{H}}_h \left(x_i, y_i, u_i, \lambda_{f,i-1}, \lambda_{g,i-1}, \eta_i\right) \\
&= f'_y \left(x_i, y_i, u_i\right)^\top \lambda_{f,i-1} + f'_y \left(x_i, y_i, u_i\right)^\top g' \left(x_i - hf\left(x_i, y_i, u_i\right)\right)^\top \lambda_{g,i-1} + c'_y \left(x_i, y_i, u_i\right)^\top \eta_i \\
&= f'_y \left(x_i, y_i, u_i\right)^\top \left(\lambda_{f,i-1} + g' \left(x_i - hf\left(x_i, y_i, u_i\right)\right)^\top \lambda_{g,i-1}\right) + c'_y \left(x_i, y_i, u_i\right)^\top \eta_i, \\
\mathbf{0}_{\mathbb{R}^{n_u}} &= \nabla_u \widetilde{\mathcal{H}}_h \left(x_i, y_i, u_i, \lambda_{f,i-1}, \lambda_{g,i-1}, \eta_i\right) \\
&= f'_u \left(x_i, y_i, u_i\right)^\top \lambda_{f,i-1} + f'_u \left(x_i, y_i, u_i\right)^\top g' \left(x_i - hf\left(x_i, y_i, u_i\right)\right)^\top \lambda_{g,i-1} + c'_u \left(x_i, y_i, u_i\right)^\top \eta_i \\
&= f'_u \left(x_i, y_i, u_i\right)^\top \left(\lambda_{f,i-1} + g' \left(x_i - hf\left(x_i, y_i, u_i\right)\right)^\top \lambda_{g,i-1}\right) + c'_u \left(x_i, y_i, u_i\right)^\top \eta_i
\end{aligned}$$

for $i = 1, \ldots, N$. It follows from the difference equation that $x_{i-1} = x_i - hf\left(x_i, y_i, u_i\right)$ for $i = 1, \ldots, N$. Thus, the conditions are equal, if

$$\begin{aligned}
\breve{\lambda}_{f,i-1} &= \lambda_{f,i-1} + g'\left(x_{i-1}\right)^\top \lambda_{g,i-1}, & i &= 1, \ldots, N, \\
\breve{\eta}_i &= \eta_i, & i &= 1, \ldots, N.
\end{aligned} \tag{5.7.7}$$

With this relation and the transversality conditions we obtain

$$\begin{aligned}
\breve{\lambda}_{f,0} &= -\breve{\ell}_0 \nabla_{x_0} \varphi\left(x_0, x_N\right) - \psi'_{x_0}\left(x_0, x_N\right)^\top \breve{\sigma} - g'\left(x_0\right)^\top \breve{\varsigma} \\
&= \lambda_{f,0} + g'\left(x_0\right)^\top \lambda_{g,0} \\
&= -\ell_0 \nabla_{x_0} \varphi\left(x_0, x_N\right) - \psi'_{x_0}\left(x_0, x_N\right)^\top \sigma - g'\left(x_0\right)^\top \varsigma + g'\left(x_0\right)^\top \lambda_{g,0}.
\end{aligned}$$

This yields the natural choice

$$\begin{aligned}
\breve{\ell}_0 &= \ell_0, \\
\breve{\sigma} &= \sigma, \\
\breve{\varsigma} &= \varsigma - \lambda_{g,0}.
\end{aligned} \tag{5.7.8}$$

For $i = 1, \ldots, N-1$ we consider the difference equation in (5.7.3)

$$\begin{aligned}
\frac{\breve{\lambda}_{f,i} - \breve{\lambda}_{f,i-1}}{h} &= -\nabla_x \breve{\mathcal{H}}\left(x_i, y_i, u_i, \breve{\lambda}_{f,i-1}, \breve{\lambda}_{g,i-1}, \breve{\eta}_i\right) \\
&= -f'_x\left(x_i, y_i, u_i\right)^\top \breve{\lambda}_{f,i-1} - g'\left(x_i\right)^\top \breve{\lambda}_{g,i-1} - c'_x\left(x_i, y_i, u_i\right)^\top \breve{\eta}_i \\
&\overset{(5.7.7)}{=} -f'_x\left(x_i, y_i, u_i\right)^\top \left(\lambda_{f,i-1} + g'\left(x_{i-1}\right)^\top \lambda_{g,i-1}\right) - g'\left(x_i\right)^\top \breve{\lambda}_{g,i-1} \\
&\quad -c'_x\left(x_i, y_i, u_i\right)^\top \eta_i.
\end{aligned} \tag{5.7.9}$$

Additionally, (5.7.7) and the difference equation in (5.7.5) imply

$$\begin{aligned}
\frac{\breve{\lambda}_{f,i}-\breve{\lambda}_{f,i-1}}{h} &= \frac{\lambda_{f,i}-\lambda_{f,i-1}}{h}+\frac{g'(x_i)^\top\lambda_{g,i}-g'(x_{i-1})^\top\lambda_{g,i-1}}{h} \\
&= -\nabla_x\widetilde{\mathcal{H}}_h(x_i,y_i,u_i,\lambda_{f,i-1},\lambda_{g,i-1},\eta_i)+\frac{g'(x_i)^\top\lambda_{g,i}-g'(x_{i-1})^\top\lambda_{g,i-1}}{h} \\
&= -f'_x(x_i,y_i,u_i)^\top\lambda_{f,i-1}-\left(\frac{g'(x_i)-g'(x_i-hf(x_i,y_i,u_i))}{h}\right)^\top\lambda_{g,i-1} \\
&\quad -\left(g'(x_i-hf(x_i,y_i,u_i))f'_x(x_i,y_i,u_i)\right)^\top\lambda_{g,i-1}-c'_x(x_i,y_i,u_i)^\top\eta_i \\
&\quad +\frac{g'(x_i)^\top\lambda_{g,i}-g'(x_{i-1})^\top\lambda_{g,i-1}}{h} \\
&= -f'_x(x_i,y_i,u_i)^\top\lambda_{f,i-1}-\left(\frac{g'(x_i)-g'(x_{i-1})}{h}\right)^\top\lambda_{g,i-1} \\
&\quad -\left(g'(x_{i-1})f'_x(x_i,y_i,u_i)\right)^\top\lambda_{g,i-1}-c'_x(x_i,y_i,u_i)^\top\eta_i \\
&\quad +\frac{g'(x_i)^\top\lambda_{g,i}-g'(x_{i-1})^\top\lambda_{g,i-1}}{h} \\
&= -f'_x(x_i,y_i,u_i)^\top\left(\lambda_{f,i-1}+g'(x_{i-1})^\top\lambda_{g,i-1}\right) \\
&\quad +g'(x_i)^\top\frac{\lambda_{g,i}-\lambda_{g,i-1}}{h}-c'_x(x_i,y_i,u_i)^\top\eta_i
\end{aligned}$$

for $i=1,\ldots,N-1$. Subtracting this equation from (5.7.9) yields

$$\mathbf{0}_{\mathbb{R}^{n_x}} = -g'(x_i)^\top\breve{\lambda}_{g,i-1}+g'(x_i)^\top\frac{\lambda_{g,i}-\lambda_{g,i-1}}{h},$$

hence the conditions are equal, if

$$\breve{\lambda}_{g,i-1} = -\frac{\lambda_{g,i}-\lambda_{g,i-1}}{h}, \quad i=1,\ldots,N-1. \tag{5.7.10}$$

Furthermore, the transversality conditions and (5.7.8) imply

$$\begin{aligned}
\breve{\lambda}_{f,N} &= \breve{\ell}_0\nabla_{x_1}\varphi(x_0,x_N)+\psi'_{x_1}(x_0,x_N)^\top\breve{\sigma} \\
&= \ell_0\nabla_{x_1}\varphi(x_0,x_N)+\psi'_{x_1}(x_0,x_N)^\top\sigma \\
&= \lambda_{f,N}.
\end{aligned}$$

Consequently, $\breve{\lambda}_{f,N}=\lambda_{f,N}+g'(x_N)^\top\lambda_{g,N}$ holds, if we set $\lambda_{g,N}=\mathbf{0}_{\mathbb{R}^{n_x}}$. Therefore, (5.7.10) implies $\breve{\lambda}_{g,N-1}=\frac{\lambda_{g,N-1}}{h}$. Finally, we summarize the relationships between the respective multipliers

$$\begin{aligned}
\breve{\lambda}_{f,i-1} &= \lambda_{f,i-1}+g'(x_{i-1})^\top\lambda_{g,i-1}, && i=1,\ldots,N, \\
\breve{\lambda}_{g,i-1} &= -\tfrac{\lambda_{g,i}-\lambda_{g,i-1}}{h}, && i=1,\ldots,N-1, \\
\breve{\lambda}_{g,N-1} &= \tfrac{\lambda_{g,N-1}}{h}, && \\
\breve{\eta}_i &= \eta_i, && i=1,\ldots,N, \\
\breve{\ell}_0 &= \ell_0, && \\
\breve{\sigma} &= \sigma, && \\
\breve{\varsigma} &= \varsigma-\lambda_{g,0}.
\end{aligned} \tag{5.7.11}$$

Example 5.7.1

In order to illustrate the relationship of the discrete multipliers in (5.7.11), *we consider the following index two problem from [8, Example 3.16]:*

$$\begin{array}{llcll} \textit{Minimize} & \frac{1}{2}x_1(1)^2 + x_3(1), & & & \\ \textit{subject to} & \dot{x}_1(t) & = & u(t), & 1 = x_1(0), \\ & \dot{x}_2(t) & = & -y(t) + u(t), & \\ & \dot{x}_3(t) & = & \frac{1}{2}\left(y(t)^2 + u(t)^2\right), & 0 = x_3(0), \\ & 0 & = & x_2(t). & \end{array} \tag{5.7.12}$$

The necessary conditions for this problem with the Hamilton function in (5.7.4)

$$\breve{\mathcal{H}}\left(x_1, x_2, x_3, y, u, \breve{\lambda}_{f_1}, \breve{\lambda}_{f_2}, \breve{\lambda}_{f_3}, \breve{\lambda}_g\right) = \breve{\lambda}_{f_1} u + \breve{\lambda}_{f_2}(-y+u) + \frac{1}{2}\breve{\lambda}_{f_3}\left(y^2+u^2\right) + \breve{\lambda}_g x_2$$

have no solution. However, the local minimum principle in Theorem 3.2.5 for $\ell_0 = 1$ with the Hamilton function

$$\mathcal{H}\left(x_1, x_2, x_3, y, u, \lambda_{f_1}, \lambda_{f_2}, \lambda_{f_3}, \lambda_g\right) = \lambda_{f_1} u + \lambda_{f_2}(-y+u) + \frac{1}{2}\lambda_{f_3}\left(y^2+u^2\right) + \lambda_g(-y+u)$$

has the solution

$$\begin{array}{rclrcl} x_1(t) & = & -\frac{1}{3}t + 1, & \lambda_{f_1}(t) & = & \frac{2}{3}, \\ x_2(t) & = & 0, & \lambda_{f_2}(t) & = & 0, \\ x_3(t) & = & \frac{1}{9}t, & \lambda_{f_3}(t) & = & 1, \\ y(t) & = & -\frac{1}{3}, & \lambda_g(t) & = & -\frac{1}{3}, \\ u(t) & = & -\frac{1}{3}. & & & \end{array}$$

For (5.7.12) *we consider the discretization in* (5.7.1)

$(\mathbf{DOCP-2})$

$$\begin{array}{llcll} \textit{Minimize} & \frac{1}{2}x_{1,N}^2 + x_{3,N}, & & & \\ \textit{subject to} & \frac{x_{1,i}-x_{1,i-1}}{h} & = & u_i, & 1 = x_{1,0}, \\ & \frac{x_{2,i}-x_{2,i-1}}{h} & = & -y_i + u_i, & \\ & \frac{x_{3,i}-x_{3,i-1}}{h} & = & \frac{1}{2}\left(y_i^2 + u_i^2\right), & 0 = x_{3,0}, \\ & 0 & = & x_{2,i}, & \end{array}$$

and the reduced discretization in (5.7.2)

$(\mathbf{RDOCP-2})$

$$\begin{array}{llcll} \textit{Minimize} & \frac{1}{2}x_{1,N}^2 + x_{3,N}, & & & \\ \textit{subject to} & \frac{x_{1,i}-x_{1,i-1}}{h} & = & u_i, & 1 = x_{1,0}, \\ & \frac{x_{2,i}-x_{2,i-1}}{h} & = & -y_i + u_i, & \\ & \frac{x_{3,i}-x_{3,i-1}}{h} & = & \frac{1}{2}\left(y_i^2 + u_i^2\right), & 0 = x_{3,0}, \\ & 0 & = & \frac{x_{2,i}-(x_{2,i}-h(-y_i+u_i))}{h} & \\ & & = & -y_i + u_i, & \\ & 0 & = & x_{2,0}. & \end{array}$$

The necessary conditions for $(\mathbf{DOCP-2})$ *stated in* (5.7.3) *with the Hamilton function* $\breve{\mathcal{H}}$ *yield the solution*

$$\begin{array}{rclll rclll}
x_{1,i} & = & -\frac{1}{3}ih+1, & i=0,1,\ldots,N, & & \breve{\lambda}_{f_1,i} & = & \frac{2}{3}, & i=0,1,\ldots,N, \\
x_{2,i} & = & 0, & i=0,1,\ldots,N, & & \breve{\lambda}_{f_2,i} & = & -\frac{1}{3}, & i=0,1,\ldots,N-1, \\
 & & & & & \breve{\lambda}_{f_2,N} & = & 0, & \\
x_{3,i} & = & \frac{1}{9}ih, & i=0,1,\ldots,N, & & \breve{\lambda}_{f_3,i} & = & 1, & i=0,1,\ldots,N, \\
y_i & = & -\frac{1}{3}, & i=1,\ldots,N, & & \breve{\lambda}_{g,i} & = & 0, & i=0,1,\ldots,N-2, \\
 & & & & & \breve{\lambda}_{g,N-1} & = & -\frac{1}{3h}, & \\
u_i & = & -\frac{1}{3}, & i=1,\ldots,N, & & \breve{\ell}_0 & = & 1. &
\end{array} \tag{5.7.13}$$

Note that $\breve{\lambda}_{f_2,i}$ *and* $\breve{\lambda}_{g,i}$ *satisfy a jump condition at their respective final index. Since we have linear algebraic constraints the discrete Hamilton function* $\widetilde{\mathcal{H}}_h$ *for* $(\mathbf{RDOCP-2})$ *coincides with the continuous Hamilton function, thus*

$$\widetilde{\mathcal{H}}_h\left(x_1,x_2,x_3,y,u,\lambda_{f_1},\lambda_{f_2},\lambda_{f_3},\lambda_g\right) = \lambda_{f_1}u + \lambda_{f_2}\left(-y+u\right) + \frac{1}{2}\lambda_{f_3}\left(y^2+u^2\right) + \lambda_g\left(-y+u\right).$$

With the necessary conditions in (5.7.5) *we obtain the solution*

$$\begin{array}{rcll rcll}
x_{1,i} & = & -\frac{1}{3}ih+1, & i=0,1,\ldots,N, & \lambda_{f_1,i} & = & \frac{2}{3}, & i=0,1,\ldots,N, \\
x_{2,i} & = & 0, & i=0,1,\ldots,N, & \lambda_{f_2,i} & = & 0, & i=0,1,\ldots,N, \\
x_{3,i} & = & \frac{1}{9}ih, & i=0,1,\ldots,N, & \lambda_{f_3,i} & = & 1, & i=0,1,\ldots,N, \\
y_i & = & -\frac{1}{3}, & i=1,\ldots,N, & \lambda_{g,i} & = & -\frac{1}{3}, & i=0,1,\ldots,N-1, \\
u_i & = & -\frac{1}{3}, & i=1,\ldots,N, & \ell_0 & = & 1, &
\end{array} \tag{5.7.14}$$

for $(\mathbf{RDOCP-2})$. *The multipliers in* (5.7.13) *and* (5.7.14) *satisfy the relations in* (5.7.11), *since* $g'\left(x_{1,i},x_{2,i},x_{3,i}\right) = (0,1,0)$ *and therefore*

$$\begin{array}{rclcll}
\breve{\lambda}_{f_1,i} & = & \lambda_{f_1,i} & = & \frac{2}{3}, & i=0,1,\ldots,N, \\
\breve{\lambda}_{f_2,i} & = & \lambda_{f_2,i}+\lambda_{g,i} & = & -\frac{1}{3}, & i=0,1,\ldots,N-1, \\
\breve{\lambda}_{f_2,N} & = & \lambda_{f_2,N} & = & 0, & \\
\breve{\lambda}_{f_3,i} & = & \lambda_{f_3,i} & = & 1, & i=0,1,\ldots,N, \\
\breve{\lambda}_{g,i-1} & = & -\frac{\lambda_{g,i}-\lambda_{g,i-1}}{h} & = & 0, & i=1,\ldots,N-1, \\
\breve{\lambda}_{g,N-1} & = & \frac{\lambda_{g,N-1}}{h} & = & -\frac{1}{3h}. &
\end{array}$$

Please note that $\breve{\lambda}_{g,N-1} \to -\infty$ *as* $h \to 0$.

In this chapter, we examined the convergence property of the implicit Euler discretization of an optimal control problem subject to an index two DAE and a mixed control-state constraint by writing the respective KKT-conditions as generalized equations and applying the approximation result in Theorem 2.2.6. A suitable reformulation of the discretized optimization problem was used in order to obtain consistent KKT-conditions. This transformation allowed us to prove convergence not only for the states and control, but also for the associated Lagrange multipliers in the L_∞-norm with linear convergence rate. The results of [80] and [85] were generalized by including algebraic equations and boundary conditions, respectively.

Chapter 6

Conclusion and Outlook

In this thesis, optimal control problems subject to Hessenberg DAEs and mixed control-state constraints are studied, extending results in the literature on necessary conditions, sufficient conditions and convergence analysis.

In Chapter 3, a local minimum principle is derived for the index one case with weakened regularity assumptions on the mixed control-state constraints. The results are then applied to optimal control problems with higher index DAEs by reducing the index to one. One drawback of this approach could be that, in the higher index case, the multipliers of the adjoint differential equations are only Lipschitz continuous, thus less smooth than the differential state. In the future, the results could be improved with a different approach, and further generalized by including pure state constraints.

A Hamilton Jacobi inequality is exploited in Chapter 4 in order to derive second-order sufficient conditions for our problem class. Herein, a quadratic function is constructed by using the solution of an appropriate Riccati equation. The conditions of the Hamilton Jacobi inequality are verified by considering two finite dimensional optimization problems, and applying suitable sufficient conditions to these problems. The main effort was to prove sufficient conditions for a special type of a parametric optimization problem with the available assumptions. Problems with DAEs and pure state constraints, as well as problems with more general DAEs could be investigated in the future.

The investigation in Chapter 5 was limited to Hessenberg DAEs of order two. Convergence was proven for the states and control, as well as the associated Lagrange multipliers by comparing the continuous and discrete KKT-conditions written as generalized equations and applying a suitable approximation theorem. In contrast to the index one case, there is a structural discrepancy between the continuous and discrete KKT-conditions, caused by an implicit index reduction in the continuous local minimum principle. Therefore, standard techniques for proving convergence fail. The first approach to overcome the discrepancy was to consider systems that satisfy the continuous necessary conditions with the Hamilton function associated with the discrete necessary conditions. However, for this class of problem it seemed impossible to verify the uniform strong regularity condition, since the adjoint DAE has index two and therefore certain regularity conditions were not satisfied. The main task then became to find a suitable reformulation of the discretized optimization problem, which yields discrete KKT-conditions that are consistent with continuous KKT-conditions. This was achieved by emulating the continuous index reduction. It is conceivable that this method of discrete index reduction may be generalized and applied to problems with Hessenberg DAEs of arbitrary order in future research. The groundwork in terms of necessary and sufficient conditions has already been established in Chapter 3 and Chapter 4, respectively. Problems with higher index DAEs might require suitable Runge-Kutta schemes, which could also be used to obtain convergence of higher order. Furthermore, problems with DAEs and pure state constraints, and linear problems with bang-bang optimal controls could be investigated in the future. Additionally, it might be possible to weaken the smoothness assumption of the minimizer by using the techniques developed in [37, 52, 81].

Chapter A

Appendix

In order to improve the reading flow, we collected auxiliary lemmas in this appendix that are unsuitable for the other chapters.

First, we prove some properties for matrices, including bounds for solutions of matrix difference equations. Then, we show that the graphs of certain types of set-valued mappings are closed / compact. Lastly, we provide the definition of measurable functions, Gronwall lemmas, a Sobolev inequality, and variational lemmas.

The first two proofs examine the norms of specific matrix differences:

Lemma A.1

Let there exist a sequence $(n_i)_{i\in\mathbb{N}} \subset \mathbb{N}$ and matrices $A_i, B_i \in \mathbb{R}^{n_i \times n_{i+1}}$ for $i \in \mathbb{N}$. Then, for every $N \in \mathbb{N}$ it holds

$$\left\| \bigtimes_{i=1}^{N} A_i - \bigtimes_{i=1}^{N} B_i \right\| \le \sum_{k=1}^{N} \|A_k - B_k\| \left(\bigtimes_{i=1}^{k-1} \|B_i\| \right) \left(\bigtimes_{i=k+1}^{N} \|A_i\| \right).$$

Proof. We use the induction principle to proof the assertion.

For $N = 2$ the assertion is true, since

$$\|A_1A_2 - B_1B_2\| = \|A_1A_2 - B_1A_2 + B_1A_2 - B_1B_2\| \le \|A_1 - B_1\| \|A_2\| + \|A_2 - B_2\| \|B_1\|.$$

Suppose the assertion holds for $N \in \mathbb{N}$. Then, it follows

$$\begin{aligned}
\left\| \bigtimes_{i=1}^{N+1} A_i - \bigtimes_{i=1}^{N+1} B_i \right\| &= \left\| \left(\bigtimes_{i=1}^{N} A_i \right) A_{N+1} - \left(\bigtimes_{i=1}^{N} B_i \right) B_{N+1} \right\| \\
&\le \left\| \bigtimes_{i=1}^{N} A_i - \bigtimes_{i=1}^{N} B_i \right\| \|A_{N+1}\| + \|A_{N+1} - B_{N+1}\| \left\| \bigtimes_{i=1}^{N} B_i \right\| \\
&\le \sum_{k=1}^{N} \|A_k - B_k\| \left(\bigtimes_{i=1}^{k-1} \|B_i\| \right) \left(\bigtimes_{i=k+1}^{N} \|A_i\| \right) \|A_{N+1}\| \\
&\quad + \|A_{N+1} - B_{N+1}\| \bigtimes_{i=1}^{N} \|B_i\| \\
&= \sum_{k=1}^{N} \|A_k - B_k\| \left(\bigtimes_{i=1}^{k-1} \|B_i\| \right) \left(\bigtimes_{i=k+1}^{N+1} \|A_i\| \right) \\
&\quad + \|A_{N+1} - B_{N+1}\| \bigtimes_{i=1}^{(N+1)-1} \|B_i\| \\
&= \sum_{k=1}^{N+1} \|A_k - B_k\| \left(\bigtimes_{i=1}^{k-1} \|B_i\| \right) \left(\bigtimes_{i=k+1}^{N+1} \|A_i\| \right),
\end{aligned}$$

which completes the proof. □

Lemma A.2
Suppose the matrices $A, B \in \mathbb{R}^{n\times n}$ are non-singular. Then it holds

$$\left\|A^{-1} - B^{-1}\right\| \leq \left\|A^{-1}\right\| \left\|B^{-1}\right\| \|A - B\| .$$

Proof. Exploiting the inequality

$$\left\|I_n - A^{-1}B\right\| = \left\|A^{-1}(A-B)\right\| \leq \left\|A^{-1}\right\| \|A - B\|$$

yields

$$\begin{aligned}\left\|A^{-1} - B^{-1}\right\| &= \left\|-\left(I_n - A^{-1}B\right)B^{-1}\right\| \\ &\leq \left\|B^{-1}\right\| \left\|I_n - A^{-1}B\right\| \\ &\leq \left\|A^{-1}\right\| \left\|B^{-1}\right\| \|A - B\| ,\end{aligned}$$

which proves the assertion. □

In Riccati equations, the following well-known type of matrix occurs:

Lemma A.3
For $n, m \in \mathbb{N}$ with $n \geq m$ suppose $Q \in \mathbb{R}^{n\times n}$ is symmetric, and $C \in \mathbb{R}^{m\times n}$ has full rank m. Moreover, it holds

$$d^\top Q d > 0 \quad \textit{for all } d \in \ker(C) \setminus \{\mathbf{0}_{\mathbb{R}^n}\} . \tag{A.1}$$

Then, the matrix

$$T := \begin{bmatrix} Q & C^\top \\ C & \mathbf{0}_{m\times m} \end{bmatrix} \in \mathbb{R}^{(n+m)\times(n+m)}$$

is non-singular.

Proof. Assume $\operatorname{rank}(T) < n + m$. Then, there exist $\begin{pmatrix} d \\ e \end{pmatrix} \in \mathbb{R}^n \times \mathbb{R}^m \setminus \left\{\begin{pmatrix} \mathbf{0}_{\mathbb{R}^n} \\ \mathbf{0}_{\mathbb{R}^m} \end{pmatrix}\right\}$ with

$$\begin{pmatrix} \mathbf{0}_{\mathbb{R}^n} \\ \mathbf{0}_{\mathbb{R}^m} \end{pmatrix} = T \begin{pmatrix} d \\ e \end{pmatrix} = \begin{pmatrix} Qd + C^\top e \\ Cd \end{pmatrix},$$

hence $d \in \ker(C)$. We conclude

$$0 = d^\top \left(Qd + C^\top e\right) = d^\top Q d,$$

which, by (A.1), implies $d = \mathbf{0}_{\mathbb{R}^n}$. Thus, $C^\top e = \mathbf{0}_{\mathbb{R}^n}$, and since C has full row rank it holds $e = \mathbf{0}_{\mathbb{R}^m}$, which contradicts $\begin{pmatrix} d \\ e \end{pmatrix} \neq \begin{pmatrix} \mathbf{0}_{\mathbb{R}^n} \\ \mathbf{0}_{\mathbb{R}^m} \end{pmatrix}$. □

Lemma A.3 allows us to prove the following:

Lemma A.4

Let $E \in \mathbb{R}^{n\times n}, Q \in \mathbb{R}^{m\times m}$ *be symmetric,* $C \in \mathbb{R}^{l\times m}$ *with* $\mathrm{rank}(C) = l \le m$*, and* $A \in \mathbb{R}^{m\times n}$*,* $B \in \mathbb{R}^{l\times n}$*. Furthermore, there exist* $\varepsilon, \gamma > 0$ *such that*

$$\begin{aligned} d^\top E d &\ge \varepsilon \|d\|^2 \quad \textit{for all } d \in \mathbb{R}^n, \\ e^\top Q e &\ge \gamma \|e\|^2 \quad \textit{for all } e \in \ker(C). \end{aligned} \tag{A.2}$$

Then, there exists a constant $\kappa > 0$ *such that the matrices*

$$M := \begin{bmatrix} E + D^\top T^{-1} D & D^\top \\ D & T \end{bmatrix}, \quad T := \begin{bmatrix} Q & C^\top \\ C & \mathbf{0}_{l\times l} \end{bmatrix}, \quad D := \begin{bmatrix} A \\ B \end{bmatrix},$$

satisfy

$$\begin{pmatrix} d \\ e \end{pmatrix}^\top \begin{bmatrix} E + D^\top T^{-1} D & A^\top \\ A & Q \end{bmatrix} \begin{pmatrix} d \\ e \end{pmatrix} = \begin{pmatrix} d \\ e \\ \mathbf{0}_{\mathbb{R}^l} \end{pmatrix}^\top M \begin{pmatrix} d \\ e \\ \mathbf{0}_{\mathbb{R}^l} \end{pmatrix} \ge \kappa \left\| \begin{pmatrix} d \\ e \end{pmatrix} \right\|^2$$

for all $\begin{pmatrix} d \\ e \end{pmatrix} \in \ker([B, C])$*.*

Proof. Define the non-singular matrices

$$S := \begin{bmatrix} \mathbf{I}_n & \mathbf{0}_{n\times(m+l)} \\ -T^{-1}D & \mathbf{I}_{m+l} \end{bmatrix}, \quad S^{-1} = \begin{bmatrix} \mathbf{I}_n & \mathbf{0}_{n\times(m+l)} \\ T^{-1}D & \mathbf{I}_{m+l} \end{bmatrix},$$

which satisfy $S^\top M S = \begin{bmatrix} E & \mathbf{0}_{n\times(m+l)} \\ \mathbf{0}_{(m+l)\times n} & T \end{bmatrix}$. For an arbitrary $\begin{pmatrix} d \\ e \end{pmatrix} \in \ker([B, C])$ we denote

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} := S^{-1} \begin{pmatrix} d \\ e \\ \mathbf{0}_{\mathbb{R}^l} \end{pmatrix} = \begin{pmatrix} d \\ T^{-1}Dd + \begin{pmatrix} e \\ \mathbf{0}_{\mathbb{R}^l} \end{pmatrix} \end{pmatrix},$$

which with $D = \begin{bmatrix} A \\ B \end{bmatrix}$ implies

$$\begin{aligned} Cb = [C, \mathbf{0}_{l\times l}] \begin{pmatrix} b \\ c \end{pmatrix} &= [C, \mathbf{0}_{l\times l}] \left(T^{-1}Dd + \begin{pmatrix} e \\ \mathbf{0}_{\mathbb{R}^l} \end{pmatrix} \right) \\ &= [\mathbf{0}_{l\times m}, \mathbf{I}_l]\, T T^{-1} D d + Ce = Bd + Ce = \mathbf{0}_{\mathbb{R}^l}, \end{aligned}$$

thus $b \in \ker(C)$. In addition, $a = d$ and $c = [\mathbf{0}_{l\times m}, \mathbf{I}_l]\, T^{-1}Dd = [\mathbf{0}_{l\times m}, \mathbf{I}_l]\, T^{-1}Da$, hence for $\alpha := \|T^{-1}\| \|D\| + 1 > 0$ it holds $\|c\| \le \alpha \|a\|$. Since

$$\left\| \begin{pmatrix} d \\ e \end{pmatrix} \right\| = \left\| \begin{pmatrix} d \\ e \\ \mathbf{0}_{\mathbb{R}^l} \end{pmatrix} \right\| = \left\| S \begin{pmatrix} a \\ b \\ c \end{pmatrix} \right\| \le \|S\| \left\| \begin{pmatrix} a \\ b \\ c \end{pmatrix} \right\|,$$

we conclude

$$
\begin{aligned}
\left(d^\top, e^\top, \mathbf{0}_{1\times l}\right) M \begin{pmatrix} d \\ e \\ \mathbf{0}_{\mathbb{R}^l} \end{pmatrix} &= \left(a^\top, b^\top, c^\top\right) S^\top M S \begin{pmatrix} a \\ b \\ c \end{pmatrix} \\
&= \left(a^\top, b^\top, c^\top\right) \begin{bmatrix} E & \mathbf{0}_{n\times(m+l)} \\ \mathbf{0}_{(m+l)\times n} & T \end{bmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} \\
&= a^\top E a + \left(b^\top, c^\top\right) \begin{bmatrix} Q & C^\top \\ C & \mathbf{0}_{l\times l} \end{bmatrix} \begin{pmatrix} b \\ c \end{pmatrix} \\
&\overset{Cb=\mathbf{0}}{=} a^\top E a + b^\top Q b \\
&\overset{(A.2)}{\geq} \varepsilon \left\|a\right\|^2 + \gamma \left\|b\right\|^2 \\
&\geq \frac{\varepsilon}{2} \left\|a\right\|^2 + \gamma \left\|b\right\|^2 + \frac{\varepsilon}{2\alpha} \left\|c\right\|^2 \\
&\geq \min\left\{\frac{\varepsilon}{2}, \gamma, \frac{\varepsilon}{2\alpha}\right\} \left\| \begin{pmatrix} a \\ b \\ c \end{pmatrix} \right\|^2 \\
&\geq \kappa \left\| \begin{pmatrix} d \\ e \end{pmatrix} \right\|^2
\end{aligned}
$$

for $\kappa := \frac{\min\left\{\frac{\varepsilon}{2}, \gamma, \frac{\varepsilon}{2\alpha}\right\}}{\|S\|^2}$, which completes the proof. □

For $N \in \mathbb{N}$, $h := \frac{1}{N}$, and $t_i := ih$, $i = 0, 1, \ldots, N$ we consider the matrix difference equation

$$
\begin{aligned}
&\Phi_h'(t_i) = A_h(t_i)\Phi_h(t_i), \quad i = 1, \ldots, N, \\
&\Phi_h(t_0) = I_n,
\end{aligned}
$$

where $A_h \in L_{\infty,h}^{n\times n}([0,1])$. The solution of this equation satisfies the following inequalities:

Lemma A.5

(i) If $h \leq \frac{1}{2\|A_h\|_\infty}$, *then for all* $i = 1, \ldots, N$ *it holds*

$$
\begin{aligned}
\|\Phi_h(t_i)\| &\leq \exp\left(2 \|A_h\|_\infty t_i\right), \\
\|\Phi_h'(t_i)\| &\leq \|A_h\|_\infty \exp\left(2 \|A_h\|_\infty t_i\right).
\end{aligned}
$$

(ii) If $h < \frac{1}{\|A_h\|_\infty}$, *then for all* $i = 1, \ldots, N$ *the matrix* $\Phi_h(t_i)$ *is non-singular and it holds*

$$
\left\|\Phi_h(t_i)^{-1}\right\| \leq \exp\left(\|A_h\|_\infty t_i\right).
$$

Proof.

(i) For $i = 1, \dots, N$ the inequality $\|hA_h(t_i)\| \le h \|A_h\|_\infty \le \frac{1}{2}$ is satisfied, thus the Neumann series of $(I_n - hA_h(t_i))$ converges and the inverse exists. This implies

$$\Phi_h(t_i) = (I_n - hA_h(t_i))^{-1} \Phi_h(t_{i-1}), \quad i = 1, \dots, N.$$

Additionally, it holds

$$\begin{aligned}
\|\Phi_h(t_i)\| &\le \left\|(I_n - hA_h(t_i))^{-1}\right\| \|\Phi_h(t_{i-1})\| \le \frac{1}{1 - h\|A_h\|_\infty} \|\Phi_h(t_{i-1})\| \\
&= \frac{1 - h\|A_h\|_\infty + h\|A_h\|_\infty}{1 - h\|A_h\|_\infty} \|\Phi_h(t_{i-1})\| \\
&= \left(1 + \frac{h\|A_h\|_\infty}{1 - h\|A_h\|_\infty}\right) \|\Phi_h(t_{i-1})\| \\
&\le \left(1 + \frac{h\|A_h\|_\infty}{1 - \frac{1}{2}}\right) \|\Phi_h(t_{i-1})\| \\
&= (1 + 2h\|A_h\|_\infty) \|\Phi_h(t_{i-1})\|,
\end{aligned}$$

for $i = 1, \dots, N$. Define $a_i := \|\Phi_h(t_i)\| \ge 0$, $i = 0, 1, \dots, N$ and $\beta := 2\|A_h\|_\infty > 0$, which satisfy

$$a_i \le (1 + \beta h)\, a_{i-1}, \quad i = 1, \dots, N.$$

Applying Lemma A.10 yields

$$\|\Phi_h(t_i)\| = a_i \le \exp(\beta i h)\, a_0 = \exp(2\|A_h\|_\infty t_i) \|\Phi_h(t_0)\| = \exp(2\|A_h\|_\infty t_i)$$

for $i = 0, 1, \dots, N$. This immediately implies

$$\|\Phi_h'(t_i)\| \le \|A_h(t_i)\| \|\Phi_h(t_i)\| \le \|A_h(t_i)\| \exp(2\|A_h\|_\infty t_i) \le \|A_h\|_\infty \exp(2\|A_h\|_\infty t_i)$$

for all $i = 1, \dots, N$.

(ii) Analog to (i), we prove that for $h < \frac{1}{\|A_h\|_\infty}$ and $i = 1, \dots, N$ the matrix $(I_n - hA_h(t_i))$ is non-singular. This implies $\Phi_h(t_i) = (I_n - hA_h(t_i))^{-1} \Phi_h(t_{i-1}) = \bigtimes_{k=0}^{i-1} (I_n - hA_h(t_{i-k}))^{-1}$, hence $\Phi_h(t_i)$ is non-singular for $i = 0, 1, \dots, N$. Moreover, for $i = 1, \dots, N$ it holds $\Phi_h(t_{i-1})^{-1} (I_n - hA_h(t_i)) = \Phi_h(t_i)^{-1}$ and therefore

$$\left\|\Phi_h(t_i)^{-1}\right\| \le \left\|\Phi_h(t_{i-1})^{-1}\right\| \|(I_n - hA_h(t_i))\| \le (1 + h\|A_h\|_\infty) \left\|\Phi_h(t_{i-1})^{-1}\right\|.$$

Applying Lemma A.10 for $a_i := \|\Phi_h(t_i)^{-1}\| \ge 0$, $i = 0, 1, \dots, N$ and $\beta := \|A_h\|_\infty$ yields

$$\left\|\Phi_h(t_i)^{-1}\right\| = a_i \le \exp(\beta i h)\, a_0 = \exp(\|A_h\|_\infty t_i), \quad i = 0, 1, \dots, N.$$

$\square$

For continuous set-valued functions defined on a compact interval we prove the following:

Lemma A.6

Let $n, m, k \in \mathbb{N}$ with $k + m < n$, $a, b \in \mathbb{R}$ with $a < b$, $\rho > 0$, and the matrix functions $A : [a, b] \to \mathbb{R}^{k\times n}$, $B : [a, b] \to \mathbb{R}^{m\times n}$ be continuous. Furthermore, let the set valued mappings $K : [a, b] \rightrightarrows \mathbb{R}^n$ and $M : [a, b] \rightrightarrows \mathbb{R}^n$ be defined by

$$K(t) := \{d \in \mathbb{R}^n \mid A(t)d = \mathbf{0}_{\mathbb{R}^k},\, B(t)d \leq \mathbf{0}_{\mathbb{R}^m}\}, \quad M(t) := K(t) \cap \mathcal{B}_\rho\left(\mathbf{0}_{\mathbb{R}^n}\right).$$

Then, graph (K) is closed and graph (M) is compact.

Proof. Since for every $t \in [a, b]$ the matrices $A(t) \in \mathbb{R}^{k\times n}$ and $B(t) \in \mathbb{R}^{m\times n}$ are linear, continuous operators and $k + m < n$, the set $K(t)$ is non-empty and closed, and the set $M(t)$ is non-empty and compact. Let $[(t_i, d_i)]_{i\in\mathbb{N}} \subseteq \operatorname{graph}(K)$ be a convergent sequence with

$$\lim_{i\to\infty} (t_i, d_i) = \left(\hat{t}, \hat{d}\right).$$

Then, it holds $A(t_i)d_i = \mathbf{0}_{\mathbb{R}^k}$ and $B(t_i)d_i \leq \mathbf{0}_{\mathbb{R}^m}$ for every $i \in \mathbb{N}$. By continuity of $A(\cdot)$, for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $i \geq N$ it holds $\left\|A(t_i) - A(\hat{t})\right\| < \frac{\varepsilon}{2\|\hat{d}\|}$ and $\left\|d_i - \hat{d}\right\| < \frac{\varepsilon}{2\|A\|_\infty}$. Consequently,

$$\left\|A(t_i)d_i - A(\hat{t})\hat{d}\right\| \leq \left\|A(t_i) - A(\hat{t})\right\| \left\|\hat{d}\right\| + \left\|d_i - \hat{d}\right\| \|A\|_\infty < \varepsilon, \quad i \geq N,$$

and therefore $\mathbf{0}_{\mathbb{R}^k} = \lim\limits_{i\to\infty} A(t_i)d_i = A(\hat{t})\hat{d}$. By the same token, it follows $B(\hat{t})\hat{d} \leq \mathbf{0}_{\mathbb{R}^m}$, hence $\left(\hat{t}, \hat{d}\right) \in \operatorname{graph}(K)$, which proves the closeness of graph (K). Finally, since

$$\operatorname{graph}(M) = \operatorname{graph}(K) \cap \left([a, b] \times \mathcal{B}_\rho\left(\mathbf{0}_{\mathbb{R}^n}\right)\right)$$

and $[a, b] \times \mathcal{B}_\rho\left(\mathbf{0}_{\mathbb{R}^n}\right)$ is compact, we conclude that graph (M) is also compact. □

For functions in $W^n_{1,2}([0, 1])$ the following inequality holds:

Lemma A.7 (Sobolev Inequality)

Let $u \in W^n_{1,2}([0, 1])$ with norm $\|u\|_{1,2} := \max\{\|u\|_2, \|\dot{u}\|_2\}$ be given. Then, it holds

$$\|u\|_\infty \leq 2\,\|u\|_{1,2}.$$

Proof. Define $v : [0, 1] \to \mathbb{R}^n$ by $v(\cdot) := \int_0^{\cdot} u(\tau)\, d\tau$. Then, by the mean-value theorem, there exists $s \in (0, 1)$ such that

$$\int_0^1 u(\tau)\, d\tau = v(1) - v(0) = \dot{v}(s) = u(s).$$

Thus, by exploiting the Cauchy-Schwartz inequality we get for every $t \in [0,1]$

$$\begin{aligned}
\|u(t)\| &= \left\| u(s) + \int_s^t \dot{u}(\tau)\, d\tau \right\| \\
&\leq \|u(s)\| + \int_s^t \|\dot{u}(\tau)\|\, d\tau \\
&\leq \int_0^1 \|u(\tau)\|\, d\tau + \int_0^1 \|\dot{u}(\tau)\|\, d\tau \\
&\leq \|u\|_2 + \|\dot{u}\|_2 \\
&\leq 2\, \|u\|_{1,2}\,.
\end{aligned}$$

□

In order to properly define the Lebesgue space $L_p^n([a,b])$ we require the notion of a measurable function (cf. [9]):

Definition A.8 (Measurable Function)
Let $(\Omega, \mathcal{A})$ and $\left(\tilde{\Omega}, \tilde{\mathcal{A}}\right)$ be measurable spaces, i.e., Ω and $\tilde{\Omega}$ are sets and $\mathcal{A}$ and $\tilde{\mathcal{A}}$ are σ-algebras on Ω and $\tilde{\Omega}$, respectively. The mapping $T : \Omega \to \tilde{\Omega}$ is called ($\mathcal{A} - \tilde{\mathcal{A}}-$)measurable, if it holds

$$T^{-1}\left(\tilde{A}\right) \in \mathcal{A} \quad \textit{for all} \quad \tilde{A} \in \tilde{\mathcal{A}}.$$

For $d \in \mathbb{N}$ and $\Omega \subseteq \mathbb{R}^d$ we consider the measurable space $\left(\Omega, \mathscr{B}^d\right)$ endowed with the Lebesgue measure defined on $\mathbb{R}^d$, where $\mathscr{B}^d$ is the σ-algebra generated by the set of all half open intervals $[a_1, b_1) \times \cdots \times [a_d, b_d) \subset \mathbb{R}^d$. Thus, the equivalence classes in the Lebesgue space $L_p^n([a,b])$ contain ($\mathscr{B} - \mathscr{B}^n$-)measurable functions $v : [a,b] \to \mathbb{R}^n$ (compare Definition 1.2).
Finally, we provide some Gronwall and variational lemmas, which proofs can be found in, e.g., [47]:

Lemma A.9 (Continuous Gronwall Lemma, [47, Lemma 1.1.14])
Let $a : [0,1] \to \mathbb{R}$ be a integrable function, $\alpha \in L_\infty([0,1])$, and $\beta \geq 0$ with

$$a(t) \leq \alpha(t) + \beta \int_0^t a(\tau) d\tau$$

for almost every $t \in [0,1]$. Then, it holds

$$a(t) \leq \|\alpha(\cdot)\|_\infty \exp(\beta t)$$

for almost every $t \in [0,1]$.

Lemma A.10 (Discrete Gronwall Lemma, [47, Lemma 4.1.21])
Let $h > 0$, $\beta > 0$, $a_n \geq 0$, $n = 0, 1, \ldots, N$ be related by

$$a_n \leq (1 + h\beta)\, a_{n-1}, \quad n = 1, \ldots, N.$$

Then, it holds

$$a_n \leq \exp(\beta n h)\, a_0, \quad n = 0, 1, \ldots, N.$$

Lemma A.11 (Variational Lemma, [47, Lemma 3.1.9])
Let $v \in L_\infty([a,b])$. *If for every* $w \in W_{1,\infty}([a,b])$ *with* $w(a) = w(b) = 0$

$$\int_a^b v(t)w(t)dt = 0$$

is satisfied, then $v(t) = 0$ *almost everywhere in* $[a,b]$.

Lemma A.12 (Variational Lemma, [47, Lemma 3.3.7])
Let $v \in L_1([a,b])$. *If for every* $w \in L_\infty([a,b])$ *with* $w(t) \geq 0$ *almost everywhere in* $[a,b]$ *the inequality*

$$\int_a^b v(t)w(t)dt \geq 0$$

is satisfied, then $v(t) \geq 0$ *almost everywhere in* $[a,b]$.

Bibliography

[1] Alt, H. W.: *Lineare Funktionalanalysis*, Springer, Berlin-Heidelberg-New York, 4th edition, (2002)

[2] Alt, W.: *Nichtlineare Optimierung: Eine Einführung in Theorie, Verfahren und Anwendungen*, Vieweg, Braunschweig/Wiesbaden, (2002)

[3] Alt, W., Baier, R., Gerdts, M., Lempio, F.: *Error Bounds for Euler Approximation of Linear-Quadratic Control Problems with Bang-Bang Solutions*, Numerical Algebra, Control and Optimization, 2(3), 547–570 (2012)

[4] Alt, W., Baier, R., Gerdts, M., Lempio, F.: *Approximations of linear control problems with bang-bang solutions*, Optimization, 62 (1), 9–32 (2013)

[5] Alt, W., Seydenschwanz, M.: *An implicit discretization scheme for linear-quadratic control problems with bang-bang solutions*, Optimization Methods and Software, 29(3), 535–560 (2014)

[6] Alt, W., Schneider, C.: *Linear-quadratic control problems with L^1-control costs*, Optimal Control Applications and Methods, 36(4), 512–534 (2015)

[7] Alt, W., Felgenhauer, U., Seydenschwanz, M.: *Euler discretization for a class of nonlinear optimal control problems with control appearing linearly*, Computational Optimization and Applications, 69(3), 825–856 (2018)

[8] Backes, A.: *Extremalbedingungen für Optimierungs-Probleme mit Algebro-Differentialgleichungen*, PhD thesis, Mathematisch-Naturwissenschaftliche Fakultät, Humboldt Universität Berlin, Berlin, Germany, (2006)

[9] Bauer, H.: *Maß- und Integrationstheorie.* Walter de Gruyter, Berlin/New York, (1990)

[10] Bazaraa, M. S., Sherali, H. D., Shetty, C. M.: *Nonlinear Programming. Theory and Algorithms*, 872p. John Wiley & Sons (2006)

[11] Betts, J. T.: *Practical methods for optimal control using nonlinear programming*, Volume 3 of Advances in Design and Control, SIAM, Philadelphia, (2001)

[12] Bock, H. G., Plitt, K. J.: *A Multiple Shooting Algorithm for Direct Solution of Optimal Control Problems*, Proceedings of the 9th IFAC Worldcongress, Budapest, Hungary, (1984)

[13] Bonnans, J. F., Hermant, A.: *Second-Order Analysis for Optimal Control Problems with Pure State Constraints and Mixed Control-State Constraints*, ELSEVIER, Annales de l'Institut Henri Poincaré, 26(2), 561–598 (2009)

[14] Bonnans, J. F., Osmolovskiĭ, N. P.: *Second-Order Analysis for Optimal Control Problems with Control and Initial-Final State Constraints*, Journal of Convex Analysis, 17(3&4), 885–913 (2010)

[15] Bonnans, J. F., Dupuis, X., Pfeiffer, L.: *Second-order sufficient conditions for strong solutions to optimal control problems*, ESAIM, Control, Optimisation and Calculus of Variations, EDP Sciences, 20(3), 704–724 (2014)

[16] Bonnans, F., Festa, A.: *Error estimates for the Euler discretization of an optimal control problem with first-order state constraints*, SIAM Journal on Numerical Analysis, 55(2), 445–471 (2017)

[17] Bryson, A. E., Jr., Denham, W. F., Dreyfus, S. E.: *Optimal Programming Problems with Inequality Constraints, I: Necessary Conditions for Extremal Solutions*, AIAA Journal, 1, 2544–2550 (1963)

[18] Büskens, C.: *Optimierungsmethoden und Sensitivitätsanalyse für optimale Steuerprozesse mit Steuer- und Zustandsbeschränkungen*, PhD thesis, Fachbereich Mathematik, Westfälische Wilhems-Universität Münster, (1998)

[19] Burger, M., Gerdts M.: *A Survey on Numerical Methods for the Simulation of Initial Value Problems with sDAEs*, In: Ilchmann A., Reis T. (eds) Surveys in Differential-Algebraic Equations IV, Differential-Algebraic Equations Forum, Springer, Cham, (2017)

[20] Campbell, S. L., Gear, C. W.: *The index of general nonlinear DAEs*, Numerische Mathematik, 72, 173–196 (1995)

[21] Chamberland, M., Zeidan, V.: *Second order necessity and sufficiency theory for the free final time problem*, in Proceedings of the 31st IEEE Conference on Decision and Control, Tucson, AZ, 1518–1525 (1992)

[22] Chen, J., Gerdts, M.: *Numerical solution of control-state constrained optimal control problems with an inexact smoothing Newton method*, IMA Journal of Numerical Analysis, 31(4), 1598–1624 (2011)

[23] Chen, J., Gerdts, M.: *Smoothing technique of nonsmooth Newton methods for control-state constrained optimal control problems*, SIAM Journal on Numerical Analysis, 50(4), 1982-2011 (2012)

[24] Chernousko, F. L., Lyubushin, A. A.: *Method of successive approximations for solution of optimal control problems*, Optimal Control Applications and Methods, 3, 101–114 (1982)

[25] Clarke, F. H.: *Methods of Dynamic and Nonsmooth Optimization*, CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, Vol. 57, (1989)

[26] Clarke, F. H., Ledyaev, Y., de Pinho, M. d. R.: *An extension of the Schwarzkopf multiplier rule in optimal control*, SIAM Journal on Control and Optimization, 49, 599 – 610 (2011)

[27] Clarke, F.: *Functional Analysis, Calculus of Variations and Optimal Control* Springer, London, (2013)

[28] Devdariani, E. N., Ledyaev, Yu. S.: *Maximum Principle for Implicit Control Systems*, Applied Mathematics and Optimization, 40, 79–103 (1999)

[29] Dmitruk, A. V.: *Maximum principle for the general optimal control problem with phase and regular mixed constraints.* Computational Mathematics and Modeling, 4, 364–377 (1993)

[30] Dontchev, A. L.: *Implicit function theorems for generalized equations*, Mathematical Programming, 70, 91–106 (1995)

[31] Dontchev, A. L., Hager, W. W., Poore, A. B., Yang, B.: *Optimality, stability and convergence in nonlinear control*, Applied Mathematics and Optimization, 31, 297–326 (1995)

[32] Dontchev, A. L., Hager, W. W., Malanowski, K.: *Error Bounds for Euler Approximation of a State and Control Constrained Optimal Control Problem*, Numerical Functional Analysis and Optimization, 21(5 & 6), 653–682 (2000)

[33] Dontchev, A. L., Hager, W. W., Veliov, V. M.: *Second-Order Runge-Kutta Approximations in Control Constrained Optimal Control*, SIAM Journal on Numerical Analysis, 38(1), 202–226 (2000)

[34] Dontchev, A. L., Hager, W. W.: *The Euler approximation in state constrained optimal control*, Mathematics of Computation, 70, 173–203 (2001)

[35] Dontchev, A. L.: *Robinson's implicit function theorem*, Control and Cybernetics, 32(3), 529–541 (2003)

[36] Dontchev, A. L., Rockafellar, R. T.: *Implicit Functions and Solution Mappings. A View from Variational Analysis*, Springer, New York, (2009)

[37] Dontchev, A. L., Krastanov, M. I., Veliov, V. M.: *On the existence of Lipschitz continuous optimal feedback control*, Research Report 2018-04, ISSN 2521-313X, Technische Universität Wien, (2018)

[38] Duff, I. S., Gear, C. W.: *Computing the structural index*, SIAM Journal on Algebraic Discrete Methods, 7(4), 594–603 (1986)

[39] Evans, L. C.: *Partial Differential Equations, Second Edition*, Graduate Studies in Mathematics, volume 19, American Mathematical Society, (2010)

[40] Fabian, M., Habala, P., Hájek, P., Montesinos, V., Zizler, V.: *Banach Space Theory: The Basis for Linear and Nonlinear Analysis*, CMS Books in Mathematics, Springer, New York, (2011)

[41] Folland, G.: *Real Analysis*, John Wiley & Sons, New York, (1984)

[42] Gear, C. W.: *Differential-algebraic equation index transformations*, SIAM Journal on Scientific and Statistical Computing, 9, 39–47 (1988)

[43] Gear, C. W.: *Differential algebraic equations, indices, and integral algebraic equations*, SIAM Journal on Numerical Analysis, 27(6), 1527–1534 (1990)

[44] Geiger, C., Kanzow, C.: *Theorie und Numerik restringierter Optimierungsaufgaben*, Springer, Berlin–Heidelberg–New York, (2002)

[45] Gerdts, M.: *Local Minimum principle for optimal control problems subject to index-two differential-algebraic equations*, Journal of Optimization Theory and Applications, 130(3), 443–462 (2006)

[46] Gerdts, M., Kunkel, M.: *A globally convergent semi-smooth Newton method for control-state constrained DAE optimal control problems*, Computational Optimization and Applications, 48(3), 601–633 (2011)

[47] Gerdts, M.: *Optimal control of ODEs and DAEs.* Walter de Gruyter, Berlin/Boston, (2012)

[48] Gerdts, M., Kunkel, M.: *Convergence Analysis of Euler Discretization of Control-State Constrained Optimal Control Problems with Controls of Bounded Variation*, Journal of Industrial and Management Optimization, 10(1), 311–336 (2014)

[49] Gerdts, M.: *A survey on optimal control problems with differential-algebraic equations*, Differential-Algebraic Equations Forum, Editors: Ilchmann, A., Reis, T., 103–161, (2015)

[50] Giaquinta, M., Modica, G.: *Mathematical Analysis. An Introduction to Functions of Several Variables*, Springer, New York, (2009)

[51] Göpfert, A., Riedrich, T.: *Funktionalanalysis*, Teubner, Leipzig, (1980)

[52] Hager, W. W.: *Lipschitz Continuity for Constrained Processes*, SIAM Journal on Control and Optimization, 17(3), 321–338 (1979)

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

[54] Hairer, E., Wanner, G.: *Solving ordinary differential equations II: Stiff and differential-algebraic problems*, Springer Series in Computational Mathematics, volume 14, 2nd edition, Springer, Berlin/Heidelberg/New York, (1996)

[55] Hargraves, C. R., Paris, S. W.: *Direct trajectory optimization using nonlinear programming and collocation*, Journal of Guidance, Control and Dynamics, 10, 338–342 (1987)

[56] Hartl, R. F., Sethi, S. P., Vickson, G.: *A Survey of the Maximum Principles for Optimal Control Problems with State Constraints*, SIAM Review, 37(2), 181–218 (1995)

[57] Hestenes, M. R.: *Calculus of variations and optimal control theory,* John Wiley & Sons, New York, (1966)

[58] Hull, D. G.: *Sufficient Conditions for a Minimum of the Free-Final-Time Optimal Control Problem*, Journal of Optimization Theory and Applications, 68(2), 275–287 (1991)

[59] Ioffe, A. D., Tihomirov, V. M.: *Theory of extremal problems,* volume 6 of Studies in Mathematics and its Applications, Amsterdam, New York, Oxford, North-Holland Publishing Company, (1979)

[60] Jacobson, D. H., Lele, M. M., Speyer, J. L.: *New Necessary Conditions of Optimality for Constrained Problems with State-Variable Inequality Constraints*, Journal of Mathematical Analysis and Applications, 35, 255–284 (1971)

[61] Knobloch, H. W.: *Das Pontryaginsche Maximumprinzip für Probleme mit Zustandsbeschränkungen i und ii*, Zeitschrift für Angewandte Mathematik und Mechanik, 55, 545–556, 621–634, (1975)

[62] Kraft, D.: *On Converting Optimal Control Problems into Nonlinear Programming Problems*, In Schittkowski, K., editor, Computational Mathematical Programming, volume F15 of NATO ASI Series, Berlin-Heidelberg-New York, Springer, 261–280 (1985)

[63] Kreindler, E.: *Additional Necessary Conditions for Optimal Control with State-Variable Inequality Constraints*, Journal of Optimization Theory and Applications, 38(2), 241–250 (1982)

[64] Kunkel, M.: *Nonsmooth Newton Methods and Convergence of Discretized Optimal Control Problems subject to DAEs*, Dissertation, Universität der Bunderwehr München, Neubiberg, Germany, (2011)

[65] Kunkel, P., Mehrmann, V.: *The linear quadratic optimal control problem for linear descriptor systems with variable coefficients*, Mathematics of Control, Signals, and Systems, 10, 247–264 (1997)

[66] Kunkel, P., Mehrmann, V.: *Differential-algebraic equations. Analysis and numerical solution*, Zürich, European Mathematical Society Publishing House, viii, (2006)

[67] Kunkel, P., Mehrmann, P.: *Optimal control for unstructured nonlinear differential-algebraic equations of arbitrary index*, Mathematics of Control, Signals, and Systems, 20(3), 227–269 (2008)

[68] Kunkel, P., Mehrmann, P.: *Formal adjoints of linear DAE operators and their role in optimal control*, Electronic Journal of Linear Algebra, 22(1), 672–693 (2011)

[69] Kurcyusz, S.: *On the Existence and Nonexistence of Lagrange Multipliers in Banach Spaces*, Journal of Optimization Theory and Applications, 20(1), 81–110 (1976)

[70] Kurina, G. A., März, R.: *On linear-quadratic optimal control problems for time-varying descriptor systems*, SIAM Journal on Control and Optimization, 42, 2062–2077 (2004)

[71] Lamour, R., März, R., Tischendorf, C.: *Differential-Algebraic Equations: A Projector Based Analysis*, Differential-Algebraic Equations Forum, Berlin, Springer, (2013)

[72] Ljusternik, L. A., Sobolew, W. I.: *Elemente der Funktionalanalysis*, Verlag Harri Deutsch, Zürich-Frankfurt/Main-Thun, (1976)

[73] Loxton, R., Lin, Q., Rehbock, V., Teo, K. L.: *Control parametrization for optimal control problems with continuous inequality constraints: new convergence results*, Numerical Algebra, Control and Optimization, 2, 571–599 (2012)

[74] März, R.: *On linear differential-algebraic equations and linerizations*, Applied Numerical Mathematics, 18(1), 267–292 (1995)

[75] März, R.: *Criteria of the trivial solution of differential algebraic equations with small nonlinearities to be asymptotically stable*, Journal of Mathematical Analysis and Applications, 225(2), 587–607 (1998)

[76] März, R.: *EXTRA-ordinary differential equations: Attempts to an analysis of differential-algebraic systems*, In A. Balog, editor, European congress of mathematics, volume 1 of Progress in Mathematics 168, Basel, Birkhäuser, 313–334 (1998)

[77] Malanowski, K.: *Second-Order Conditions and Constraint Qualifications in Stability and Sensitivity Analysis of Solutions to Optimization Problems in Hilbert Spaces*, Applied Mathematics and Optimization, 25, 51–79 (1992)

[78] Malanowski, K., Maurer, H.: *Sensitivity Analysis for Parametric Control Problems with Control-State Constraints*, Computational Optimization and Applications, 5, 253–283 (1996)

[79] Malanowski, K.: *Sufficient Optimality Conditions for Optimal Control subject to State Constraints*, SIAM Journal on Control and Optimization, 35(1), 205–227 (1997)

[80] Malanowski, K., Büskens, C., Maurer, H.: *Convergence of Approximations to Nonlinear Optimal Control Problems*, In: Anthony Fiacco, editor, Mathematical programming with data perturbations, volume 195, Dekker, Lecture Notes in Pure and Applied Mathematics, 253–284 (1997)

[81] Malanowski, K.: *On normality of Lagrange multipliers for state constrained optimal control problems*, Optimization, 52(1), 75–91 (2003)

[82] Malanowski, K., Maurer, H., Pickenhain, S.: *Second-order Sufficient Conditions for State-constrained Optimal Control Problems*, Journal of Optimization Theory and Applications, 123(3), 595–617 (2004)

[83] Martens, B., Gerdts, M.: *Necessary and Sufficient Conditions for Optimal Control Problems Subject to Hessenberg Differential-Algebraic Equations of Arbitrary Index and Mixed Control-State Constraints*, to appear in: Pure and Applied Functional Analysis, (2019)

[84] Martens, B., Gerdts, M.: *Convergence Analysis of the Implicit Euler-discretization and Sufficient Conditions for Optimal Control Problems Subject to Index-one Differential-algebraic Equations*, Set-Valued and Variational Analysis, DOI 10.1007/s11228-018-0471-x, (2018)

[85] Martens, B., Gerdts, M.: *Convergence Analysis for Approximations of Optimal Control Problems Subject to Higher Index Differential-Algebraic Equations and Mixed Control-State Constraints*, Preprint, Bundeswehr University Munich, Institute of Mathematics and Applied Computing, submitted, (2018)

[86] Maurer, H.: *Optimale Steuerungsprozesse Mit Zustandsbeschränkungen*, Mathematisches Institut, Universität Würzburg, Germany, (1976)

[87] Maurer, H.: *On Optimal Control Problems with Boundary State Variables and Control Appearing Linearly*, SIAM Journal on Control and Optimization, 15(3), 345–362 (1977)

[88] Maurer, H., Zowe, J.: *First- and second-order necessary and sufficient optimality conditions in infinite dimensional programming Problems*, Mathematical Programming, 16, 98–110 (1979)

[89] Maurer, H.: *Differential stability in optimal control problems*, Applied Mathematics and Optimization, 5, 283–295 (1979)

[90] Maurer, H.: *On the minimum principle for optimal control problems with state constraints*, Schriftenreihe des Rechenzentrums der Universität Münster Nr. 41, Münster, (1979)

[91] Maurer, H.: *First and Second Order Sufficient Optimality Conditions in Mathematical Programming and Optimal Control*, Mathematical Programming Study, 14, 163–177 (1981)

[92] Maurer, H., Pickenhain, S.: *Second-order Sufficient Conditions for Control Problems with Mixed Control-state Constraints*, Journal of Optimization Theory and Applications, 86(3), 649–667 (1995)

[93] Maurer, H., Oberle, H. J.: *Second Order Sufficient Conditions for Optimal Control Problems with free Final Time: The Riccati Approach*, SIAM Journal on Control and Optimization, 41(2), 380–403 (2002)

[94] Mehrmann, V.: *Existence, uniqueness, and stability of solutions to singular linear quadratic optimal control problems*, Linear Algebra and Applications, 121, 291–331 (1989)

[95] Mehrmann, V.: *The autonomous linear quadratic control problem. Theory and numerical solution*, Lecture Notes in Control and Information Sciences, volume 163, Springer, Berlin, (1991)

[96] Neustadt, L. W.: *Optimization: A Theory of Necessary Conditions*, Princeton, New Jersey, (1976)

[97] Nocedal, J., Wright, S. J.: *Numerical Optimization 2nd ed.*, 664p., Springer Series in Operations Research and Financial Engineering, New York, (2006)

[98] Orrell, D., Zeidan, V.: *Another Jacobi Sufficiency Criterion for Optimal Control with Smooth Constraints*, Journal of Optimization Theory and Applications, 58, 283–300 (1988)

[99] Pickenhain, S., Tammer, K.: *Sufficient Conditions for Local Optimality in Multidimensional Control Problems with State Restrictions*, Zeitschrift für Analysis und ihre Anwendungen, 10, 397–405 (1991)

[100] Pickenhain, S.: *Sufficiency Conditions for Weak Local Minima in Multidimensional Optimal Control Problems with Mixed Control-State Restrictions*, Zeitschrift für Analysis und ihre Anwendungen, 11, 559–568 (1992)

[101] Pietrus, A., Scarinci, T., Veliov, V. M.: *High order discrete approximations to Mayer's problems for linear systems*, SIAM Journal on Control and Optimization, 56(1), 102–119 (2018)

[102] de Pinho, M., Vinter, R. B.: *Necessary Conditions for Optimal Control Problems Involving Nonlinear Differential Algebraic Equations*, Journal of Mathematical Analysis and Applications, 212, 493–516 (1997)

[103] de Pinho, M., Ilchmann, A.: *Weak maximum principle for optimal control problems with mixed constraints*, Nonlinear Analysis, 48(8), 1179–1196 (2002)

[104] Pontryagin, L. S., Boltyanskij, V. G., Gamkrelidze, R. V., Mishchenko, E. F.: *Mathematische Theorie optimaler Prozesse*, Oldenbourg, München, (1964)

[105] Pytlak, R.: *Optimal control of differential-algebraic equations of higher index. II: Necessary optimality conditions*, Journal of Optimization Theory and Applications, 134(1), 77–90 (2007)

[106] Reid, W. T.: *Riccati Differential Equations*, Mathematics in Science and Engineering, volume 86, Academic Press, New York, (1972)

[107] Riesz, F., Sz.-Nagy, B.: *Functional Analysis*, Dover Publications Inc., New York, (1990), originally published by Frederick Ungar Publishing Co., (1955)

[108] Robinson, S. M.: *Stability Theory for Systems of Inequalities, Part II: Differentiable Nonlinear Systems*, SIAM Journal on Numerical Analysis, 13(4), 487–513 (1976)

[109] Robinson, S. M.: *Strongly regular generalized equations*, Mathematics of Operations Research, 5, 43–62 (1980)

[110] Robinson, S. M.: *Local structure of feasible sets in nonlinear programming. Part III: Stability and sensitivity*, Mathematical Programming Study, volume 30, 45–66 (1987)

[111] Roubicek, T., Valásek, M.: *Optimal control of causal differential-algebraic systems*, Journal of Mathematical Analysis and Applications, 269(2), 616–641 (2002)

[112] Rudin, W.: *Real and Complex Analysis*, McGraw–Hill, (1987)

[113] Scarinci, T., Veliov, V. M.: *Higher-order numerical scheme for linear-quadratic problems with bang-bang controls*, Computational Optimization and Applications, 69(2), 403–422 (2018)

[114] Schulz, V. H., Bock, H. G., Steinbach, M. C.: *Exploiting invariants in the numerical solution of multipoint boundary value problems for DAE*, SIAM Journal on Scientific Computing, 19(2), 440–467 (1998)

[115] Steinbach, M. C.: *Fast Recursive SQP Methods for Large-Scale Optimal Control Problems*, PhD thesis, Interdisziplinäres Zentrum für Wissenschaftliches Rechnen, Universität Heidelberg, Heidelberg, Germany, (1995)

[116] Stetter, H. J.: *Analysis of discretization methods for ordinary differential equations*, Springer Tracts in Natural Philosophy, volume 23, Berlin-Heiderlberg-New York, Springer Verlag, (1973)

[117] Veliov, V. M.: *On the Time-Discretization of Control Systems*, SIAM Journal on Control Optimization, 35(5), 1470–1486 (1997)

[118] Veliov, V. M.: *Error analysis of discrete approximations to bang-bang optimal control problems: the linear case*, Control and cybernetics, 34, 967–982 (2005)

[119] Walter, W.: *Gewöhnliche Differentialgleichungen*, Heidelberger Taschenbücher, Springer, Heidelberg, (1976)

[120] Werner, D.: *Funktionalanalysis*, Springer, Berlin-Heidelberg-New York, (1995)

[121] Yosida, K., Hewitt, E.: *Finitely additive measures*, Transactions of the American Mathematical Society, 72, 46–66 (1952)

[122] Zeidan, V.: *First and Second-Order Sufficient for Optimal Control and the Calculus of Variations*, Applied Mathematics and Optimization, 11(1), 209–226 (1984)

[123] Zeidan, V.: *Sufficiency Conditions with Minimal Regularity Assumption*, Applied Mathematics and Optimization, 20, 19–31 (1989)

[124] Zeidan, V.: *The Riccati Equation for Optimal Control Problems with Mixed State-Control Constraints: Necessity and Sufficiency*, SIAM Journal on Control and Optimization, 32(5), 1297–1321 (1994)