TECHNISCHE UNIVERSITÄT
ILMENAU

Master thesis

Optimal control of linear differential-algebraic equations

Jonas Witschel

Advisor:
Prof. Dr. Achim Ilchmann

Ilmenau, 9th May 2017 (revised edition)

Abstract

The topic of the master thesis is linear-quadratic optimal control of time-varying and time-invariant differential-algebraic equations (DAEs). The thesis is divided into two main parts: in the first part, we investigate linear time-varying DAEs. We recall the solution theory of DAEs and introduce the optimal control problem. We then proceed to prove that the optimal value is a quadratic function and fulfils Bellman's principle of optimality. Using these results, we can characterize the optimal value as an extremal solution of the Kalman-Yakubovich-Popov inequality. In the second part, we turn our attention towards time-invariant, regular DAEs. We first derive a differentiability condition that the control input of the system needs to fulfil. Using these results, we introduce an augmented system that includes certain derivatives of the input as system states. For this augmented system, an optimal control problem equivalent to the one of the nominal system is defined that can be solved easily using results from the theory of optimal control for ordinary differential equations. This enables us to explicitly calculate the optimal control of the nominal DAE, which can be implemented as a state feedback as well.

Zusammenfassung

Das Thema der Masterarbeit ist die linear-quadratische Optimalsteuerung zeitvarianter und zeitinvarianter differentiell-algebraischer Gleichungen (DAEs). Die Arbeit besteht aus zwei Hauptteilen: Im ersten Teil betrachten wir zeitvariante DAEs. Wir wiederholen die Lösungstheorie von DAEs und definieren das von uns betrachtete Optimalsteuerungsproblem. Anschließend zeigen wir, dass der Optimalwert eine quadratische Funktion ist und das Bellmansche Optimalitätsprinzip erfüllt. Mithilfe dieser Ergebnisse können wir den Optimalwert als extremale Lösung der Kalman-Yakubovich-Popov-Ungleichung charakterisieren. Im zweiten Teil widmen wir uns zeitinvarianten regulären DAEs. Wir leiten zunächst eine Differenzierbarkeitsbedingung her, die der Steuereingang des Systems erfüllen muss. Mithilfe dieser Resultate führen wir ein erweitertes System ein, das gewisse Ableitungen des Eingangs als Systemzustände enthält. Für das so erweiterte System kann ein zum Optimalsteuerungsproblem des originalen Systems äquivalentes Optimalsteuerungsproblem definiert werden. Dieses lässt sich mithilfe der Theorie der Optimalsteuerung gewöhnlicher Differentialgleichungen leicht lösen. Das ermöglicht uns, die optimale Steuerung der ursprünglichen DAE explizit anzugeben. Weiterhin lässt sich diese auch als Zustandsrückführung implementieren.

Acknowledgements

First of all, I want to sincerely thank my advisor, Prof. Achim Ilchmann, for his advice and support, which enabled writing this thesis. I further extend my gratitude to Prof. Karl Worthmann and Dr. Leslie Leben for all their fruitful discussions and suggestions.

Contents

Nomenclature

$\mathbb{N}, \mathbb{N}_0$	the natural numbers excluding and including 0, respectively
$\mathbb{R}$	the real numbers
$\mathbb{C}$	the complex numbers
$\overline{\mathbb{C}}_+$	$:= \{z \in \mathbb{C} \mid \operatorname{Re} z \geq 0\}$, the closed right half complex plane
$\mathbb{R}[s]$	ring of polynomials with real coefficients and the indeterminate s
$\mathbb{R}(s)$	field of rational functions with real coefficients (the quotient field of $\mathbb{R}[s]$)
$\mathbb{R}^n$	real column vectors of dimension n
$\mathbb{R}^{m\times n}$	real matrices with m rows and n columns
$M^\top$	transpose of the matrix $M \in \mathbb{R}^{m\times n}$
I_n	the $n \times n$ identity matrix
$0_{m\times n}$	the $m \times n$ zero matrix
$GL_n(\mathbb{R})$	invertible real matrices of size $n \times n$
$\mathbb{R}^{n\times n}_{\text{sym}}$	symmetric matrices
$\text{rk}_{\mathbb{F}}$	rank of a matrix over the field $\mathbb{F}$
$\text{nil ind}\, N$	$:= \min\{i \in \mathbb{N} \mid N^i = 0\}$, i. e. the index of nilpotency of a nilpotent matrix $N \in \mathbb{R}^{n\times n}$
$M \geq 0, M > 0$	$M \in \mathbb{R}^{n\times n}$ is positive semidefinite or positive definite, respectively
$M \geq N, M > N$	$M - N \in \mathbb{R}^{n\times n}$ is positive semidefinite or positive definite, respectively
$\mathcal{C}(X \to Y)$	space of continous functions with domain X and codomain Y
$\mathcal{C}^1$	space of continously differentiable functions
$\mathcal{L}^1_{\text{loc}}$	space of locally integrable functions
$\mathcal{L}^2_{\text{loc}}$	space of locally square integrable functions
$\mathcal{L}^\infty_{\text{loc}}$	space of locally essentially bounded functions
$\mathcal{W}^{k,p}_{\text{loc}}$	Sobolev space of functions whose weak derivatives up to the order $k \in \mathbb{N}_0$ exist and are in $\mathcal{L}^p_{\text{loc}}$
ess sup	essential supremum
$\lvert\cdot\rvert$	absolute value of a real or complex number
$\lVert\cdot\rVert$	norm of a vector or function
dom	domain of a function
$f\vert_I$	restriction of the function f to the set I
$\overset{\text{ae}}{=}$	equality almost everywhere
$[f(t)]^b_{t=a}$	evaluation of the function f defined by $[f(t)]^b_{t=a} := f(b) - f(a)$
$\frac{\mathrm{d}f}{\mathrm{d}t}, \dot{f}$	(weak) derivative of f

1 Introduction

In this master thesis, we will consider linear time-varying differential-algebraic equation systems (DAEs) of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}(E(t)x(t)) = A(t)x(t) + B(t)u(t),$$

where $E(\cdot)$, $A(\cdot)$ and $B(\cdot)$ are matrix-valued functions of time and $x(\cdot)$ and $u(\cdot)$ are vector-valued. $x(\cdot)$ can be seen as the state of the dynamic system and $u(\cdot)$ is some form of external control input. Such systems are a generalization of linear time-varying ordinary differential equations (ODEs)

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = A(t)x(t) + B(t)u(t)$$

that are able to include additional algebraic constraints, expressed by the, generally singular, matrix function $E(\cdot)$. Such systems arise for example in circuit theory or when modelling mechanical systems, when dynamic elements such as capacitors or inductors need to be coupled. This coupling introduces algebraic constraints, as for example described by Kirchhoff's laws. In recent years, such systems have come into focus both from an engineering point of view and also as a research area in applied mathematics: large-scale electric circuits like microprocessors can be modelled automatically as differential-algebraic equation systems, which can then be analysed algebraically and numerically.

The focus of the thesis is the optimal control of differential-algebraic systems: we want to achieve a fast convergence of the state to a desired setpoint (usually assumed to be 0), using a reasonable amount of control effort. To accomplish this goal, we define a so-called performance index that "penalizes" high deviations of the state from the setpoint as well as high control inputs and try do calculate an optimal trajectory that minimizes this performance index. Choosing different performance indices suited to the real-world process we want to control allows us to trade off fast convergence to the setpoint with high control inputs against slower convergence with more feasible inputs.

1.1 Historical overview

The problem we want to approach has two sides: differential-algebraic equation systems and optimal control. Differential-algebraic equations have been in the focus of mathematical studies since the 1970ies. There are numerous textbooks and research monographs available, e. g. [Cam80; Cam82; KM06; LMT13]. Optimal control for *ordinary* differential equations has been covered extraordinarily comprehensively both from an engineering as well as from a mathematical point of view, see e. g. [Bro70; CA78; AM90; LR95; ZDG96] to name but a few. An important class of optimal control problems (OCPs) are so-called non-singular linear-quadratic OPCs, which is what we will consider in this thesis. They can be solved using differential and algebraic *Riccati equations*, which are very well-studied for ordinary differential equations [LR95; AFIJ03].

The consolidation of both topics, the optimal control of *differential-algebraic* equations, has however received much less attention: there are a few attempts to generalize Riccati equations to differential-algebraic systems [Lew85; KM90], however as [BL87] and more recently [LMT13] have shown, one needs to be very careful in pursuing this approach, as some of these generalizations may not have a solution even if the optimal control problem is indeed feasible. In the time-invariant case, one may consider so-called Lur'e equations [RRV15] instead, which are closely related to the Kalman-Yakubovich-Popov inequality we will study in Section 2.4. So far no dedicated textbook focusing on linear-quadratic optimal control problems for differential-algebraic equations exists.

1.2 Structure of the thesis

The thesis consists of two main parts. In Chapter 2 we will recall the solution theory of time-varying linear differential-algebraic equations. Afterwards we will introduce a linear-quadratic optimal control problem for time-varying DAEs and prove basic properties of the optimal value of this optimal control problem, namely the quadratic representation of the optimal value and Bellman's famous principle of optimality. Using these results, we will see that the optimal value fulfils the so-called Kalman-Yakubovich-Popov inequality. Finally, we will be able to characterize the optimal value as an extremal solution of this inequality. The results in this chapter closely follow the proofs given in [CA78, Chapter II], tailored to differential-algebraic equations.

Chapter 3 is the main chapter of the thesis. We will investigate time-invariant, regular differential-algebraic equations to explicitly calculate the optimal control. To this end we first derive differentiability conditions that the control input of the differential-algebraic equation must fulfil to admit a solution. Using these results, we introduce an augmented system whose input, unlike the nominal system, may be prescribed freely. For the augmented system, an optimal control problem equivalent to the one of the nominal system can be defined. The new optimal control problem is that of an ordinary differential equation and can be solved using well-established methodology. Exploiting this fact, we can state the optimal control of the nominal differential-algebraic system as well as derive optimal state feedbacks. To the best of our knowledge, both the differentiability condition as well as the augmented system and the optimal controls building on it cannot be found in the current literature.

2 Optimal control of time-varying linear differential-algebraic equations

2.1 Time-varying linear differential-algebraic equations

We consider time-varying linear differential-algebraic equations

$$\frac{\mathrm{d}}{\mathrm{d}t}(E(t)x(t)) = A(t)x(t) + B(t)u(t) \tag{2.1}$$

where $E, A \in \mathcal{C}(I \to \mathbb{R}^{n\times n})$ and $B \in \mathcal{C}(I \to \mathbb{R}^{n\times m})$ for some perfect interval $I \subseteq \mathbb{R}$. We will denote the system (2.1) by $[E, A, B]$ for short.

With this class of systems, we face two main problems:

(a) Unlike ordinary differential equations, not every input $u \in \mathcal{L}^2_{\text{loc}}(\mathbb{R} \to \mathbb{R}^m)$ and every initial value $x(t_0) = x^0 \in \mathbb{R}^n$ admits a solution of the DAE (2.1). The input needs to fulfil certain differentiability conditions. This is already true for time-invariant DAEs and will be investigated for these in more detail in Section 3.4. Hence we always need to consider pairs of state and input trajectories (x, u).

(b) As the DAE (2.1) is time-varying, it can have a *finite escape time*, i. e. a solution that exists on an interval $J \subseteq I$ might not be extendable to the whole interval I.

These considerations lead to the following concept of a solution.

Definition 2.1.1 (Solutions and behaviour)**.** A tuple $(x, u) \in \mathcal{L}^2_{\text{loc}}(J \to \mathbb{R}^n \times \mathbb{R}^m)$ is called a *solution* of $[E, A, B]$ on some *interval of existence* $J \subseteq I$ if, and only if, $Ex \in \mathcal{W}^{1,2}_{\text{loc}}(J \to \mathbb{R}^n)$ and (2.1) is fulfilled almost everywhere on J.

The set of all (local) solutions of $[E, A, B]$ is denoted by the *behaviour*

$$\mathfrak{B}_{[E,A,B]} := \{(x, u) \in \mathcal{L}^2_{\text{loc}}(J \to \mathbb{R}^n \times \mathbb{R}^m) \mid J \subseteq \mathbb{R} \text{ perfect interval, } (x, u) \text{ solves (2.1)}\}.$$

Furthermore, we are interested in solutions defined on a fixed perfect interval $J \subseteq I$

$$\mathfrak{B}^J_{[E,A,B]} := \{(x, u) \in \mathcal{L}^2_{\text{loc}}(J \to \mathbb{R}^n \times \mathbb{R}^m) \mid (x, u) \text{ solves (2.1)}\}.$$

Note that throughout the whole thesis, we will always use the unique (absolutely) continuous representative of functions from the Sobolev space $\mathcal{W}^{1,2}_{\text{loc}}$ so that pointwise evaluation of these functions is well-defined.

Some care must be exercised when specifying initial values: unlike ordinary differential equations, not every initial value allows for a solution. Instead we regard the *consistent initial values*,

i. e. the initial values we can prescribe at a given time t_0 so that a solution of the DAE (2.1) exists. For this definition, we also need to require x to be continuous, as we have only assumed Ex to be continuous in our definition of a solution and E will generally be singular.
An alternative is to only prescribe the initial value for Ex, such that $(Ex)(t_0) = Ex^0$ for a given *differential initial value* $x^0 \in \mathbb{R}^n$. As we have required Ex to be continuous, this does not need any extra assumptions, unlike the consistent initial values. Both concepts are formalized in the following definition.

Definition 2.1.2 (Consistent initial values)**.** The vector space of *consistent initial values* for a given time instant $t_0 \in I$ is given by

$$\mathcal{V}_{[E,A,B]}(t_0) := \{x^0 \in \mathbb{R}^n \mid \exists x \in \mathfrak{B}_{[E,A,B]} : x \in \mathcal{C}(\operatorname{dom} x \to \mathbb{R}^n),\ t_0 \in \operatorname{dom} x,\ x(t_0) = x^0\}.$$

Likewise, the space of *consistent initial differential values* is defined by

$$\mathcal{V}^{\text{diff}}_{[E,A,B]}(t_0) := \{x^0 \in \mathbb{R}^n \mid \exists x \in \mathfrak{B}_{[E,A,B]} : t_0 \in \operatorname{dom} x,\ (Ex)(t_0) = E(t_0)x^0\}.$$

Obviously, $\mathcal{V}_{[E,A,B]}(t_0) \subseteq \mathcal{V}^{\text{diff}}_{[E,A,B]}(t_0)$.

In this thesis, we will mainly consider systems that do not exhibit a finite escape time, so-called *full systems* [INS84].

Definition 2.1.3 (Fullness)**.** A system (2.1) is called *full* if, and only if, $I = \mathbb{R}$ and every solution on an interval $J \subseteq \mathbb{R}$ can be extended to the whole of $\mathbb{R}$, i. e.

$$\forall J \subseteq \mathbb{R} \text{ perfect interval } \forall (x,u) \in \mathfrak{B}^J_{[E,A,B]}\ \exists (\hat{x},\hat{u}) \in \mathfrak{B}^{\mathbb{R}}_{[E,A,B]} : (\hat{x},\hat{u})|_J = (x,u).$$

Example 2.1.4. Consider for $I = \mathbb{R}$ the DAE

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & t \end{bmatrix}\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(t). \tag{2.2}$$

Choose any perfect interval $J \subseteq \mathbb{R}$ with $0 \notin J$. Evaluating the second row of (2.2), we see that for almost all $t \in J$

$$x_2(t) = -\frac{u(t)}{t}.$$

By Definition 2.1.1, we have that $x_2 \in \mathcal{W}^{1,2}_{\text{loc}}(J \to \mathbb{R})$, so $t \mapsto -\frac{u(t)}{t}$ and therefore u must be differentiable almost everywhere. Hence evaluating the first row of (2.2) yields

$$x_1(t) = \dot{x}_2(t) = \frac{u(t)}{t^2} - \frac{\dot{u}(t)}{t} \quad \text{for almost all } t \in J.$$

The behaviour is given by

$$\mathfrak{B}^J_{[E,A,B]} = \left\{ (x,u) \in \mathcal{L}^2_{\text{loc}}(J \to \mathbb{R}^n \times \mathbb{R}^m) \;\middle|\; u \in \mathcal{W}^{1,2}_{\text{loc}}(J \to \mathbb{R}),\ x(t) \overset{\text{ae}}{=} \begin{pmatrix} \frac{u(t)}{t^2} - \frac{\dot{u}(t)}{t} \\ -\frac{u(t)}{t} \end{pmatrix} \right\}.$$

For $t_0 \neq 0$, choose any perfect interval $J \subseteq \mathbb{R}$ with $t_0 \in J$, $0 \notin J$. Let $x^0 \in \mathbb{R}^2$ be an arbitrary and define $u \in \mathcal{W}^{1,2}_{\mathrm{loc}}(J \to \mathbb{R})$ by

$$u(t) := -(t-t_0)(t_0 x_1^0 - x_2^0) - t_0 x_2^0.$$

Hence for $x \in C(J \to \mathbb{R}^2)$ with

$$x(t) = \begin{pmatrix} \frac{u(t)}{t^2} - \frac{\dot{u}(t)}{t} \\ -\frac{u(t)}{t} \end{pmatrix} = \begin{pmatrix} \frac{2t-t_0}{t^2}(t_0 x_1^0 - x_2^0) + \frac{1}{t^2} t_0 x_2^0 \\ \frac{t-t_0}{t}(t_0 x_1^0 - x_2^0) + \frac{1}{t} t_0 x_2^0 \end{pmatrix},$$

we have $(x,u) \in \mathfrak{B}^J_{[E,A,B]}$ and $x(t_0) = x^0$, therefore

$$\forall t_0 \neq 0 : \mathcal{V}_{[E,A,B]}(t_0) = \mathcal{V}^{\mathrm{diff}}_{[E,A,B]}(t_0) = \mathbb{R}^2.$$

The system (2.2) is not full: choose any perfect interval $J \subseteq \mathbb{R}$ with $0 \notin J$ and set $u \equiv 1$, then we see that the solution

$$\left(t \mapsto -\frac{1}{t}, 1\right) \in \mathfrak{B}^J_{[E,A,B]}$$

cannot be extended to the whole of $\mathbb{R}$ as it has a singularity at $t = 0$.

2.2 The performance index and the optimal value

Now we turn our attention towards optimal control problems, where we want to minimize a given *performance index* defined for solutions $(x,u) \in \mathfrak{B}^{[t_0,t_\mathrm{f})}_{[E,A,B]}$ on the *time horizon* $[t_0, t_\mathrm{f}) \subseteq I$, $t_\mathrm{f} \in (t_0, \infty]$. In this thesis, we will study the quadratic performance index

$$J(t_0, t_\mathrm{f}, x, u) := \int_{t_0}^{t_\mathrm{f}} \ell(x(t), u(t), t)\,\mathrm{d}t, \tag{2.3}$$

where

$$\ell : \mathbb{R}^n \times \mathbb{R}^m \times [t_0, t_\mathrm{f}),\ \ell(x,u,t) := \begin{pmatrix} x \\ u \end{pmatrix}^\top \begin{bmatrix} Q(t) & H(t) \\ H(t)^\top & R(t) \end{bmatrix} \begin{pmatrix} x \\ u \end{pmatrix},$$
$$Q \in \mathcal{L}^\infty_{\mathrm{loc}}(I \to \mathbb{R}^{n\times n}_{\mathrm{sym}}),\ H \in \mathcal{L}^\infty_{\mathrm{loc}}(I \to \mathbb{R}^{n\times m}),\ R \in \mathcal{L}^\infty_{\mathrm{loc}}(I \to \mathbb{R}^{m\times m}_{\mathrm{sym}}).$$

The performance index "penalizes" high values of the state x by the scalar product $x^\top Q x$ as well as high control inputs u by the scalar product $u^\top R u$. We can also include an evaluation of coupled states and inputs as given by $2x^\top H u$.

For a given consistent initial differential value $x^0 \in \mathcal{V}^{\mathrm{diff}}_{[E,A,B]}(t_0)$, we are interested in the infimum of the performance index (2.3) (if it exists), denoted by the *optimal value*

$$J^*(t_0, t_\mathrm{f}, E(t_0)x^0) := \inf\{J(t_0, t_\mathrm{f}, x, u) \mid (x,u) \in \mathfrak{B}^{[t_0,t_\mathrm{f})}_{[E,A,B]},\ (Ex)(t_0) = E(t_0)x^0\} \in [-\infty, \infty]. \tag{2.4}$$

Note that the optimal value might not exist or be infinite for a number of reasons:

(a) For a given consistent initial differential value $x^0 \in \mathcal{V}^{\mathrm{diff}}_{[E,A,B]}(t_0)$, a solution is only guaranteed to exist on a possibly small interval that includes t_0. The solution might not be extendable to the whole interval $[t_0, t_\mathrm{f})$ that the performance index is defined on, thus rendering the set over which to take the infimum in (2.4) empty.

(b) The integral in (2.3) might be $+\infty$ for all solutions $\mathfrak{B}^{[t_0,t_\mathrm{f})}_{[E,A,B]}$ with $(Ex)(t_0) = E(t_0)x^0$. This is only a concern on an infinite time horizon, as for $t_\mathrm{f} < \infty$, the Hölder inequality [AE08, Theorem X.4.2] and our assumptions guarantee

$$|J(t_0, t_\mathrm{f}, x, u)| \leq \|Q\|_{L^\infty}\|x\|^2_{L^2} + 2\|x\|_{L^2}\|H\|_{L^\infty}\|u\|_{L^2} + \|R\|_{L^\infty}\|u\|^2_{L^2} < \infty.$$

(c) We might have $J^*(t_0, t_\mathrm{f}, E(t_0)x^0) = -\infty$, i. e. the performance index can be made arbitrarily small by a series of solutions $(x_n, u_n) \in \left(\mathfrak{B}^{[t_0,t_\mathrm{f})}_{[E,A,B]}\right)^{\mathbb{N}}$. This can be prevented by requiring

$$\begin{bmatrix} Q(\cdot) & H(\cdot) \\ H(\cdot)^\top & R(\cdot) \end{bmatrix} \geq 0 \text{ almost everywhere on } [t_0, t_\mathrm{f}), \tag{2.5}$$

as this gives a lower bound of 0 for the performance index (2.3). A somewhat milder assumption is to postulate $J(t_0, t_\mathrm{f}, x, u) \geq 0$ for all $(x, u) \in \mathfrak{B}^{[t_0,t_\mathrm{f})}_{[E,A,B]}$, without directly requiring the matrix in (2.5) to be positive semidefinite. This assumption is what we will use in this chapter.

Assumption 2.2.1. Throughout this chapter, we will require the performance index (2.3) to fulfil

$$\forall (x, u) \in \mathfrak{B}^{[t_0,t_\mathrm{f})}_{[E,A,B]} : J(t_0, t_\mathrm{f}, x, u) \geq 0.$$

This assumption is notably fulfilled if $\left[\begin{smallmatrix} Q(\cdot) & H(\cdot) \\ H(\cdot)^\top & R(\cdot) \end{smallmatrix}\right]$ is pointwise positive semidefinite.

2.3 Properties of the optimal value

We want to solve the optimal control problem, i. e. for a given consistent inital differential value $x^0 \in \mathcal{V}^{\mathrm{diff}}_{[E,A,B]}(t_0)$, we want to find an *optimal control* $(x^*, u^*) \in \mathfrak{B}^{[t_0,t_\mathrm{f})}_{[E,A,B]}$ such that $(Ex^*)(t_0) = E(t_0)x^0$ and

$$J(t_0, t_\mathrm{f}, x^*, u^*) = J^*(t_0, t_\mathrm{f}, Ex^0),$$

if such a solution (x^*, u^*) exists. An important step in this process is to be able to characterize the optimal value J^*. As a first result, we prove that the optimal value is a quadratic function in Ex^0.

Proposition 2.3.1 (Quadratic optimal value)**.** *Let $[E, A, B]$ be given as in* (2.1) *with the performance index* (2.3) *given on $[t_0, t_f) \subseteq I$ and the optimal value J^* as in* (2.4)*. If $J^*(t_0, t_f, E(t_0)x^0)$ exists and is finite for all $x^0 \in \mathcal{V}^{diff}_{[E,A,B]}(t_0)$, then it has a quadratic representation*

$$\exists P(t_0, t_f) \in \mathbb{R}^{n\times n}_{sym}\ \forall x^0 \in \mathcal{V}^{diff}_{[E,A,B]}(t_0) : J^*(t_0, t_f, E(t_0)x^0) = (E(t_0)x^0)^\top P(t_0, t_f)\,(E(t_0)x^0).$$

Proof. To show that $J^*(t_0, t_\mathrm{f}, \cdot)$ is a quadratic form, it is according to [CA78, Lemma II.2.2] enough to prove the following three equalities for all $x^0, x^1 \in \mathcal{V}^{\mathrm{diff}}_{[E,A,B]}(t_0)$, $\lambda \in \mathbb{R}$:

$$J^*(t_0, t_\mathrm{f}, \lambda E(t_0)x^0) = \lambda^2 J^*(t_0, t_\mathrm{f}, E(t_0)x^0), \tag{2.6}$$

$$\begin{aligned} &J^*(t_0, t_\mathrm{f}, E(t_0)(x^0 + x^1)) + J^*(t_0, t_\mathrm{f}, E(t_0)(x^0 - x^1)) \\ &\quad = 2J^*(t_0, t_\mathrm{f}, E(t_0)x^0) + 2J^*(t_0, t_\mathrm{f}, E(t_0)x^1), \end{aligned} \tag{2.7}$$

$$\begin{aligned} &J^*(t_0, t_\mathrm{f}, E(t_0)(x^0 + \lambda x^1)) - J^*(t_0, t_\mathrm{f}, E(t_0)(x^0 - \lambda x^1)) \\ &\quad = \lambda J^*(t_0, t_\mathrm{f}, E(t_0)(x^0 + x^1)) - \lambda J^*(t_0, t_\mathrm{f}, E(t_0)(x^0 - x^1)). \end{aligned} \tag{2.8}$$

We show (2.6): Let $x^0 \in \mathcal{V}^{\mathrm{diff}}_{[E,A,B]}(t_0)$ be arbitrary and choose $(x, u) \in \mathfrak{B}^{[t_0,t_\mathrm{f})}_{[E,A,B]}$ such that $(Ex)(t_0) = E(t_0)x^0$. Then by linearity of the DAE (2.1), $(\lambda x, \lambda u) \in \mathfrak{B}^{[t_0,t_\mathrm{f})}_{[E,A,B]}$ fulfils $(E\lambda x)(t_0) = \lambda E(t_0)x^0$. So by definition of the performance index,

$$J(t_0, t_\mathrm{f}, \lambda x, \lambda u) = \lambda^2 J(t_0, t_\mathrm{f}, x, u).$$

For arbitrary $\varepsilon > 0$, choose $(x, u) \in \mathfrak{B}^{[t_0,t_\mathrm{f})}_{[E,A,B]}$ with $(Ex)(t_0) = E(t_0)x^0$ such that

$$J(t_0, t_\mathrm{f}, x, u) \le J^*(t_0, t_\mathrm{f}, E(t_0)x^0) + \varepsilon.$$

So we have

$$J^*(t_0, t_\mathrm{f}, \lambda E(t_0)x^0) \le J(t_0, t_\mathrm{f}, \lambda x, \lambda u) = \lambda^2 J(t_0, t_\mathrm{f}, x, u) \le \lambda^2 J^*(t_0, t_\mathrm{f}, E(t_0)x^0) + \lambda^2 \varepsilon. \tag{2.9}$$

Similarly, choosing a solution $(\lambda x, \lambda u) \in \mathfrak{B}^{[t_0,t_\mathrm{f})}_{[E,A,B]}$ with $(E\lambda x)(t_0) = \lambda E(t_0)x^0$ such that

$$J(t_0, t_\mathrm{f}, \lambda x, \lambda u) \le J^*(t_0, t_\mathrm{f}, \lambda E(t_0)x^0) + \varepsilon$$

leads to

$$\lambda^2 J^*(t_0, t_\mathrm{f}, E(t_0)x^0) \le \lambda^2 J(t_0, t_\mathrm{f}, x, u) = J(t_0, t_\mathrm{f}, \lambda x, \lambda u) \le J^*(t_0, t_\mathrm{f}, \lambda E(t_0)x^0) + \varepsilon. \tag{2.10}$$

Since ε was arbitrary, we obtain (2.6) from (2.9) and (2.10).

The other two equations are proved similarly. □

Note that $P(t_0, t_\mathrm{f})$ given in the previous proposition is not unique: as $E(t_0)$ generally is singular, we can choose any projector $T \in \mathbb{R}^{n\times n}$ with $TE(t_0) = E(t_0)$ and obtain

$$\begin{aligned} \forall x^0 \in \mathcal{V}^{\mathrm{diff}}_{[E,A,B]}(t_0) : J^*(t_0, t_\mathrm{f}, E(t_0)x^0) &= (E(t_0)x^0)^\top P(t_0, t_\mathrm{f})\,(E(t_0)x^0) \\ &= (E(t_0)x^0)^\top\, [T^\top P(t_0, t_\mathrm{f})\, T]\,(E(t_0)x^0). \end{aligned}$$

In the previous proposition, we have studied the optimal value "pointwise" for a fixed t_0. To find a characterization of the optimal value, we need to study the optimal value for *all* $t_1 \in (t_0, t_\mathrm{f}]$ instead. In order to do this, and fulfil the existence condition for J^* from the previous proposition, we will introduce the concept of *reachability*.

Definition 2.3.2 (Strong reachability)**.** The system $[E, A, B]$ is called *strongly reachable on the interval* $[t_0, t_f]$ if, and only if,

$$\forall x^{\mathrm{f}} \in \mathbb{R}^n \; \exists (x, u) \in \mathfrak{B}_{[E,A,B]}^{[t_0,t_{\mathrm{f}}]} : (Ex)(t_0) = 0 \wedge (Ex)(t_{\mathrm{f}}) = E(t_{\mathrm{f}})x^{\mathrm{f}}.$$

Note that the definition implies $\mathcal{V}_{[E,A,B]}^{\mathrm{diff}}(t_{\mathrm{f}}) = \mathbb{R}^n$. As we will see in the next proposition, reachability guarantees the existence of the optimal value on a finite horizon. Additionally, we need the system to be full in order to be able to extend every solution to the whole interval $[t_0, t_{\mathrm{f}})$ so that we can use Assumption 2.2.1.

Proposition 2.3.3. *Let* $t_f < \infty$ *and* $[E, A, B]$ *given by* (2.1) *be full and strongly reachable on the interval* $[t_0, t_1]$ *for all* $t_1 \in (t_0, t_f]$*. Assume further that the performance index* (2.3) *fulfils Assumption 2.2.1. Then*

$$\exists P(\cdot, t_f) : (t_0, t_f] \to \mathbb{R}_{sym}^{n\times n} \; \forall x^1 \in \mathbb{R}^n \; \forall t_1 \in (t_0, t_f] : J^*(t_1, t_f, E(t_1)x^1) = (E(t_1)x^1)^\top P(t_1, t_f)\, E(t_1)x^1.$$

Proof. Let $x^1 \in \mathbb{R}^n$ be arbitrary. As the system is strongly reachable on $[t_0, t_1]$, there exists a solution $(x_{\mathrm{r}}, u_{\mathrm{r}}) \in \mathfrak{B}_{[E,A,B]}^{[t_0,t_1]}$ with $(Ex_{\mathrm{r}})(t_1) = E(t_1)x^1$. By assumption that the system is full, this solution extends to a solution $(\hat{x}, \hat{u}) \in \mathfrak{B}_{[E,A,B]}^{[t_0,t_f)}$, therefore

$$J^*(t_1, t_{\mathrm{f}}, E(t_1)x^1) \leq J(t_1, t_{\mathrm{f}}, \hat{x}|_{[t_1,t_{\mathrm{f}})}, \hat{u}|_{[t_1,t_{\mathrm{f}})}) \leq J(t_0, t_{\mathrm{f}}, \hat{x}, \hat{u}) < \infty.$$

To prove the inequality $J^*(t_1, t_{\mathrm{f}}, E(t_1)x^1) > -\infty$, let $(x, u) \in \mathfrak{B}_{[E,A,B]}^{[t_1,t_{\mathrm{f}})}$ be any solution with $(Ex)(t_1) = x^1$. Then for

$$\tilde{x}(t) := \begin{cases} x_{\mathrm{r}}(t), & t \in [t_0, t_1) \\ x(t), & t \in [t_1, t_{\mathrm{f}}) \end{cases}, \quad \tilde{u}(t) := \begin{cases} u_{\mathrm{r}}(t), & t \in [t_0, t_1) \\ u(t), & t \in [t_1, t_{\mathrm{f}}) \end{cases},$$

we have $(\tilde{x}, \tilde{u}) \in \mathfrak{B}_{[E,A,B]}^{[t_0,t_{\mathrm{f}})}$. Furthermore, by Assumption 2.2.1

$$J(t_0, t_{\mathrm{f}}, \tilde{x}, \tilde{u}) = \int_{t_0}^{t_1} \ell(x_{\mathrm{r}}(t), u_{\mathrm{r}}(t), t)\,\mathrm{d}t + \int_{t_1}^{t_{\mathrm{f}}} \ell(x(t), u(t), t)\,\mathrm{d}t \geq 0.$$

Therefore

$$J(t_1, t_{\mathrm{f}}, x, u) = \int_{t_1}^{t_{\mathrm{f}}} \ell(x(t), u(t), t)\,\mathrm{d}t \geq -\int_{t_0}^{t_1} \ell(x_{\mathrm{r}}(t), u_{\mathrm{r}}(t), t)\,\mathrm{d}t,$$

and since $(x, u) \in \mathfrak{B}_{[E,A,B]}^{[t_1,t_{\mathrm{f}})}$ is arbitrary, this gives a constant lower bound for $J^*(t_1, t_{\mathrm{f}}, E(t_1)x^1)$.

Therefore, $J^*(t_1, t_{\mathrm{f}}, E(t_1)x^1)$ exists for all $x^1 \in \mathbb{R}^n$, $t_1 \in (t_0, t_{\mathrm{f}}]$. According to Proposition 2.3.1, it can therefore be expressed as

$$J^*(t_1, t_{\mathrm{f}}, E(t_1)x^1) = (E(t_1)x^1)^\top P(t_1, t_{\mathrm{f}})E(t_1)x^1.$$

This concludes the proof. □

As in Proposition 2.3.1, the function $P(\cdot, t_f)$ generally is not uniquely determined. It is currently not known how the DAE (2.1) and the performance index (2.3) relate to the smoothness of the function $P(\cdot, t_f)$. To proceed further, we will require the latter to be continuously differentiable in the next section, however it is not clear how smooth the coefficients in (2.1) and (2.3) need to be chosen to warrant this assumption.

Using this proposition, we can prove an important result for optimal control problems, the so-called *Bellman's principle of optimality*. Roughly speaking, the principle asserts that for every optimal trajectory $(x^*, u^*) \in \mathfrak{B}^{[t_0,t_f)}_{[E,A,B]}$, the trajectory $(x^*|_{[t_1,t_f)}, u^*|_{[t_1,t_f)}) \in \mathfrak{B}^{[t_1,t_f)}_{[E,A,B]}$ is also an optimal trajectory on $[t_1, t_f)$.

Theorem 2.3.4 (Bellman's principle of optimality)**.** *Let $t_f < \infty$ and $[E, A, B]$ given by* (2.1) *be full and strongly reachable on the interval $[t_0, t_1]$ for all $t_1 \in (t_0, t_f]$. Assume further that the performance index* (2.3) *fulfils Assumption 2.2.1. Then*

$$\forall x^1 \in \mathbb{R}^n \, \forall [t_1, t_2] \subseteq (t_0, t_f] \, \forall (x, u) \in \mathfrak{B}^{[t_1,t_2]}_{[E,A,B]} \text{ with } (Ex)(t_1) = E(t_1)x^1 :$$
$$J^*(t_1, t_f, E(t_1)x^1) \leq \int_{t_1}^{t_2} \ell(x(t), u(t), t)\,\mathrm{d}t + J^*(t_2, t_f, (Ex)(t_2)).$$

Proof. Let $x^1 \in \mathbb{R}^n$, $[t_1, t_2] \subseteq (t_0, t_f]$ and $(x, u) \in \mathfrak{B}^{[t_1,t_2]}_{[E,A,B]}$ with $(Ex)(t_1) = E(t_1)x^1$ be arbitrary. According to Proposition 2.3.3, the optimal values $J^*(t_1, t_f, E(t_1)x^1)$ and $J^*(t_2, t_f, (Ex)(t_2))$ exist and are finite. Choose $(\hat{x}, \hat{u}) \in \mathfrak{B}^{[t_2,t_f)}_{[E,A,B]}$ for $\varepsilon > 0$ with $(E\hat{x})(t_2) = (Ex)(t_2)$ and $J(t_2, t_f, \hat{x}, \hat{u}) \leq J^*(t_2, t_f, (Ex)(t_2)) + \varepsilon$. Then for

$$\tilde{x}(t) := \begin{cases} x(t), & t \in [t_1, t_2) \\ \hat{x}(t), & t \in [t_2, t_f) \end{cases}, \quad \tilde{u}(t) := \begin{cases} u(t), & t \in [t_1, t_2) \\ \hat{u}(t), & t \in [t_2, t_f) \end{cases},$$

we have $(\tilde{x}, \tilde{u}) \in \mathfrak{B}^{[t_1,t_f)}_{[E,A,B]}$ with $(E\tilde{x}_1)(t_1) = (Ex)(t_1) = E(t_1)x^1$ and therefore

$$\begin{aligned} J^*(t_1, t_f, E(t_1)x^1) \leq J(t_1, t_f, \tilde{x}, \tilde{u}) &= \int_{t_1}^{t_2} \ell(x(t), u(t), t)\,\mathrm{d}t + J(t_2, t_f, \hat{x}, \hat{u}) \\ &\leq \int_{t_1}^{t_2} \ell(x(t), u(t), t)\,\mathrm{d}t + J^*(t_2, t_f, (Ex)(t_2)) + \varepsilon. \end{aligned}$$

Since ε is arbitrary, the inequality is proved. □

2.4 The Kalman-Yakubovich-Popov inequality

After the previous considerations regarding the optimal value, in this section we will characterize the optimal value as the maximal solution of the so-called *Kalman-Yakubovich-Popov (KYP) inequality*. To do so, we need the additional assumption that the quadratic representation $P(t_1, t_f)$ found in Proposition 2.3.3 is differentiable with respect to t_1.

Assumption 2.4.1. We require the function

$$P_*(\cdot) := P(\cdot, t_f) : (t_0, t_f] \to \mathbb{R}^{n\times n}_{\mathrm{sym}} \tag{2.11}$$

as given in Proposition 2.3.3 to be continuously differentiable: $P_* \in C^1((t_0, t_f] \to \mathbb{R}^{n\times n}_{\mathrm{sym}})$.

In this thesis we will not further investigate under which conditions this assumption holds.

Theorem 2.4.2 (Kalman-Yakubovich-Popov inequality). *Let $t_f < \infty$, $[E, A, B]$ be given by* (2.1) *and $[t_1, t_2] \subseteq (t_0, t_f]$. Then for any $P \in C^1([t_1, t_2] \to \mathbb{R}^{n\times n}_{sym})$ and $(x, u) \in \mathfrak{B}^{[t_1,t_2]}_{[E,A,B]}$ we have*

$$\int_{t_1}^{t_2} \begin{pmatrix} x(t) \\ u(t) \end{pmatrix}^\top \begin{bmatrix} \mathcal{L}(P)(t) & E(t)^\top P(t) B(t) \\ B(t)^\top P(t) E(t) & 0 \end{bmatrix} \begin{pmatrix} x(t) \\ u(t) \end{pmatrix} \mathrm{d}t = \left[(Ex)(t)^\top P(t)(Ex)(t)\right]_{t=t_1}^{t_2}, \tag{2.12}$$

where $\mathcal{L}(P)$ is given by

$$\mathcal{L}(P)(\cdot) := E(\cdot)^\top \dot{P}(\cdot)E(\cdot) + A(\cdot)^\top P(\cdot)E(\cdot) + E(\cdot)^\top P(\cdot)A(\cdot) + Q(\cdot).$$

Additionally, let $[E, A, B]$ be full, strongly reachable on the interval $[t_0, t_1]$ for all $t_1 \in (t_0, t_f]$, and assume that the performance index (2.3) *fulfils Assumption 2.2.1. If P_* as defined in* (2.11) *fulfils Assumption 2.4.1, then we have the* Kalman-Yakubovich-Popov inequality

$$\begin{pmatrix} x(t) \\ u(t) \end{pmatrix}^\top \begin{bmatrix} \mathcal{L}(P_*)(t) & E(t)^\top P_*(t)B(t) + H(t) \\ B(t)^\top P_*(t)E(t) + H(t)^\top & R(t) \end{bmatrix} \begin{pmatrix} x(t) \\ u(t) \end{pmatrix} \geq 0 \text{ for almost all } t \in (t_0, t_f]. \tag{2.13}$$

Proof. Evaluating the integral in (2.12) yields

$$\begin{aligned}
&\int_{t_1}^{t_2} x^\top E^\top \dot{P} E x + x^\top E^\top P(Ax + Bu) + (Ax + Bu)^\top P E x \,\mathrm{d}t \\
&= \int_{t_1}^{t_2} (Ex)^\top \dot{P} E x + (Ex)^\top P \frac{\mathrm{d}}{\mathrm{d}t}(Ex) + \left(\frac{\mathrm{d}}{\mathrm{d}t} Ex\right)^\top P E x \,\mathrm{d}t \\
&= \int_{t_1}^{t_2} \frac{\mathrm{d}}{\mathrm{d}t}\left((Ex)^\top P (Ex)\right) \mathrm{d}t = \left[(Ex)(t)^\top P(t)(Ex)(t)\right]_{t=t_1}^{t_2}.
\end{aligned}$$

For the second assertion (2.13), note that Bellman's principle of optimality and Proposition 2.3.3 give

$$\left[(Ex)(t)^\top P_*(t)(Ex)(t)\right]_{t=t_1}^{t_2} = J^*(t_2, t_f, (Ex)(t_2)) - J^*(t_1, t_f, (Ex)(t_1)) \geq -\int_{t_1}^{t_2} \ell(x(t), u(t), t)\,\mathrm{d}t. \tag{2.14}$$

So we have

$$\begin{aligned}
&\int_{t_1}^{t_2} \begin{pmatrix} x \\ u \end{pmatrix}^\top \begin{bmatrix} E^\top \dot{P}_* E + E^\top P_* A + A^\top P_* E + Q & E^\top P_* B + H \\ B^\top P_* E + H^\top & R \end{bmatrix} \begin{pmatrix} x \\ u \end{pmatrix} \mathrm{d}t \\
&= \int_{t_1}^{t_2} \begin{pmatrix} x \\ u \end{pmatrix}^\top \begin{bmatrix} E^\top \dot{P}_* E + E^\top P_* A + A^\top P_* E & E^\top P_* B \\ B^\top P_* E & 0 \end{bmatrix} \begin{pmatrix} x \\ u \end{pmatrix} + \ell(x, u, t)\,\mathrm{d}t \\
&= \left[(Ex)(t)^\top P_*(t)(Ex)(t)\right]_{t=t_1}^{t_2} + \int_{t_1}^{t_2} \ell(x(t), u(t), t)\,\mathrm{d}t \overset{(2.14)}{\geq} 0.
\end{aligned}$$

Since this equation holds for all $t_1 \in (t_0, t_f)$, $t_2 \in (t_1, t_f]$, the assertion (2.13) follows. □

We can now characterize the optimal value as one of the solutions of the Kalman-Yakubovich-Popov inequality which are maximal on the image of $E(\cdot)$, i. e. $E(\cdot)^\top P(\cdot)E(\cdot) \leq E(\cdot)^\top P_*(\cdot)E(\cdot)$ pointwise for all solutions P of the KYP inequality.

Theorem 2.4.3. *Let $t_f < \infty$ and $[E, A, B]$ given by* (2.1) *be full and strongly reachable on the interval $[t_0, t_1]$ for all $t_1 \in (t_0, t_f]$. Assume further that the performance index* (2.3) *fulfils Assumption 2.2.1, and let Assumption 2.4.1 hold for P_* as defined in* (2.11)*. Then for all $P \in C^1((t_0, t_f] \to \mathbb{R}^{n\times n}_{sym})$ that fulfil*

$$\begin{aligned}
&\forall (x,u) \in \mathfrak{B}^{[t_0,t_f)}_{[E,A,B]} : \begin{pmatrix} x \\ u \end{pmatrix}^\top \begin{bmatrix} \mathcal{L}(P) & E^\top PB + H \\ B^\top PE + H^\top & R \end{bmatrix} \begin{pmatrix} x \\ u \end{pmatrix} \overset{ae}{\geq} 0, \\
&E(t_f)^\top P(t_f)E(t_f) = 0,
\end{aligned}$$

we have

$$\forall t \in (t_0, t_f] : E(t)^\top P(t)E(t) \leq E(t)^\top P_*(t)E(t). \tag{2.15}$$

Proof. Let $t \in (t_0, t_\mathrm{f}]$, $x^1 \in \mathbb{R}^n$ be arbitrary and choose $(x,u) \in \mathfrak{B}^{[t,t_\mathrm{f}]}_{[E,A,B]}$ with $(Ex)(t) = E(t)x^1$. By Theorem 2.4.2, we know that for all $t \in (t_0, t_\mathrm{f}]$,

$$\int_t^{t_\mathrm{f}} \begin{pmatrix} x \\ u \end{pmatrix}^\top \begin{bmatrix} \mathcal{L}(P) & E^\top PB + H \\ B^\top PE + H & R \end{bmatrix} \begin{pmatrix} x \\ u \end{pmatrix} \mathrm{d}\tau = \left[(Ex)^\top P(Ex)\right]_{\tau=t}^{t_\mathrm{f}} + \int_t^{t_\mathrm{f}} \ell(x(\tau), u(\tau), \tau)\,\mathrm{d}\tau.$$

Therefore by assumption,

$$\begin{aligned}
J(t, t_\mathrm{f}, x, u) &= (Ex)(t)^\top P(t)(Ex)(t) - (Ex)(t_\mathrm{f})^\top P(t_\mathrm{f})(Ex)(t_\mathrm{f}) \\
&\quad + \int_t^{t_\mathrm{f}} \begin{pmatrix} x \\ u \end{pmatrix}^\top \begin{bmatrix} \mathcal{L}(P) & E^\top PB + H \\ B^\top PE + H & R \end{bmatrix} \begin{pmatrix} x \\ u \end{pmatrix} \mathrm{d}\tau \\
&\geq (Ex)(t)^\top P(t)(Ex)(t) = (x^1)^\top E(t)^\top P(t)E(t)x^1.
\end{aligned}$$

It follows that

$$\begin{aligned}
(x^1)^\top E(t)^\top P_*(t)\, E(t)x^1 &= J^*(t, t_\mathrm{f}, E(t)x^1) \\
&= \inf\{J(t, t_\mathrm{f}, x, u) \mid (x,u) \in \mathfrak{B}^{[t,t_\mathrm{f})}_{[E,A,B]}, (Ex)(t) = E(t)x^1\} \\
&\geq (x^1)^\top E(t)^\top P(t)E(t)x^1.
\end{aligned}$$

As $x^1 \in \mathbb{R}^n$ is arbitrary, this shows the assertion (2.15). □

As for Proposition 2.3.3, there will generally be multiple maximal solutions, any of which describes the optimal value.

This concludes the treatment of time-varying DAEs in this chapter. We have seen that the optimal value J^* is a quadratic function that fulfils Bellman's principle of optimality and the Kalman-Yakubovich-Popov inequality. Furthermore, we can find the optimal value by searching for the maximal solutions of the KYP inequality. However this is not a viable way to obtain the optimal value, as it involves finding the extremal solution of a matrix-valued function. In the next chapter, we will therefore turn our attention to time-invariant DAEs which allow for a much easier and numerically more feasible treatment of the optimal control problem.

3 Optimal control of time-invariant regular differential-algebraic equations

3.1 Time-invariant linear differential-algebraic equations

In this chapter, we will consider time-invariant, single-input differential-algebraic systems

$$\frac{\mathrm{d}}{\mathrm{d}t}Ex(t) = Ax(t) + bu(t), \tag{3.1}$$

where $E, A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$; we write $[E, A, b] \in \mathbb{R}^{n\times(2n+1)}$. The notion of a solution stays the same as in Section 2.1. Note however that for time-invariant linear DAEs, no finite escape time is possible, so we can define the behaviour on the whole of $\mathbb{R}$:

$$\mathfrak{B}_{[E,A,b]} := \left\{(x,u) \in \mathcal{L}^2_{\mathrm{loc}}(\mathbb{R} \to \mathbb{R}^n \times \mathbb{R}) \;\middle|\; \begin{array}{l} Ex \in \mathcal{W}^{1,2}_{\mathrm{loc}}(\mathbb{R} \to \mathbb{R}^n), \\ \text{(3.1) holds for almost all } t \in \mathbb{R} \end{array}\right\}. \tag{3.2}$$

Likewise, the consistent (differential) initial values are not dependent on t_0 any more, so without loss of generality, we fix $t_0 = 0$.

$$\mathcal{V}_{[E,A,b]} := \{x^0 \in \mathbb{R}^n \mid \exists (x,u) \in \mathfrak{B}_{[E,A,b]} : x \in \mathcal{C}(\mathbb{R} \to \mathbb{R}^n) \wedge x(0) = x^0\},$$
$$\mathcal{V}^{\mathrm{diff}}_{[E,A,b]} := \{x^0 \in \mathbb{R}^n \mid \exists (x,u) \in \mathfrak{B}_{[E,A,b]} : (Ex)(0) = Ex^0\}.$$

For the optimal control, we also start at $t_0 = 0$ without loss of generality. Using constant matrices for $Q(\cdot)$, $H(\cdot)$ and $R(\cdot)$ in (2.3), the performance index of the time-invariant system $[E, A, b] \in \mathbb{R}^{n\times(2n+1)}$ becomes

$$J(t_\mathrm{f}, x, u) = \int_0^{t_\mathrm{f}} \begin{pmatrix} x(t) \\ u(t) \end{pmatrix}^\top \begin{bmatrix} Q & h \\ h^\top & r \end{bmatrix} \begin{pmatrix} x(t) \\ u(t) \end{pmatrix} \mathrm{d}t,$$

where $t_\mathrm{f} \in (0, \infty]$, $(x, u) \in \mathfrak{B}_{[E,A,b]}$ and $Q \in \mathbb{R}^{n\times n}_{\mathrm{sym}}$, $h \in \mathbb{R}^n$ and $r \in \mathbb{R}$.

To ensure that J is bounded from below, we can assume that

$$\begin{bmatrix} Q & h \\ h^\top & r \end{bmatrix} \geq 0, \tag{3.3}$$

as explained in Section 2.2. However, we will see in Section 3.6 that this assumption is not necessary for the approach chosen in this chapter, as we will reduce the DAE optimal control problem to an optimal control problem for an ordinary differential equation. To obtain a finite optimal value, we must ensure that the optimal value of the ODE control problem is finite. This is guaranteed for a finite time horizon $t_\mathrm{f} < \infty$ by requiring (2.5), but can also be achieved under weaker assumptions.

The definition of the optimal value (2.4) is left unchanged:

$$J^*(t_\mathrm{f}, Ex^0) := \inf\{J(t_\mathrm{f}, x, u) \mid (x,u) \in \mathfrak{B}_{[E,A,b]}, (Ex)(0) = Ex^0\} \in [-\infty, \infty],\; t_\mathrm{f} \in (0, \infty].$$

3.2 Controllability and stabilizability

For the following considerations of time-invariant optimal control problems, we will use some further notions of controllability and stabilizability. The first is a weaker version of the strong reachability we introduced in Definition 2.3.2: instead of requiring that it is possible to strongly reach every $x^{\mathrm{f}} \in \mathbb{R}^n$, we only need to be able to steer from each solution trajectory $(x_1, u_1) \in \mathfrak{B}_{[E,A,b]}$ to each other trajectory $(x_2, u_2) \in \mathfrak{B}_{[E,A,b]}$ in a finite time t_f.

Definition 3.2.1 (Behavioural controllability [PW98, Definition 5.2.2])**.** The system $[E, A, b] \in \mathbb{R}^{n\times(2n+1)}$ is called *behaviourally controllable* if, and only if,

$$\forall (x_1, u_1), (x_2, u_2) \in \mathfrak{B}_{[E,A,b]}\, \exists t_{\mathrm{f}} > 0\, \exists (x, u) \in \mathfrak{B}_{[E,A,b]} : ((x(t), u(t)) = \begin{cases} (x_1(t), u_1(t)), & t < 0 \\ (x_2(t), u_2(t)), & t > t_{\mathrm{f}}. \end{cases}$$

The following concept of stabilizability is an even weaker condition, where we only require to be able to reach 0 asymptotically, starting from a given solution trajectory $(x, u) \in \mathfrak{B}_{[E,A,b]}$.

Definition 3.2.2 (Behavioural stabilizability [PW98, Definition 5.2.29])**.** The system $[E, A, b] \in \mathbb{R}^{n\times(2n+1)}$ is called *behaviourally stabilizable* if, and only if,

$$\forall (x, u) \in \mathfrak{B}_{[E,A,b]}\, \exists (\tilde{x}, \tilde{u}) \in \mathfrak{B}_{[E,A,b]} : (x, u)|_{(-\infty,0)} \overset{\text{ae}}{=} (\tilde{x}, \tilde{u})|_{(-\infty,0)} \wedge \lim_{t\to\infty} \operatorname*{ess\,sup}_{[t,\infty)} \|(\tilde{x}, \tilde{u})\| = 0.$$

Note that we have [BR13, Corollary 4.3]

$$\begin{aligned} & [E, A, b] \text{ is strongly reachable} \\ \Longrightarrow\ & [E, A, b] \text{ is behaviourally controllable} \\ \Longrightarrow\ & [E, A, b] \text{ is behaviourally stabilizable.} \end{aligned}$$

3.3 The quasi Weierstraß Form

A very important subclass of time-invariant linear DAEs are so-called *regular* DAEs. This type of DAE has the useful property that the solution for a given input u and initial value $x^0 \in \mathcal{V}_{[E,A,b]}$, provided that it exists, is unique. [BIT12, Proposition 2.16]

Definition 3.3.1. The pencil $(sE - A) \in \mathbb{R}[s]^{n\times n}$ is said to be *regular* if, and only if,

$$\det(sE - A) \in \mathbb{R}[s] \setminus \{0\}.$$

Otherwise, it is called *singular*. We call $[E, A, b] \in \mathbb{R}^{n\times(2n+1)}$ regular (singular) if, and only if, the pencil $sE - A$ is regular (singular), respectively.

It is well-known, see e. g. [BIT12, Theorem 2.4], that any regular pencil $(sE - A) \in \mathbb{R}[s]^{n\times n}$ can be transformed into quasi Weierstraß form, i. e. there are $S, T \in GL_n(\mathbb{R})$ such that

$$S(sE - A)T = s\overline{E} - \overline{A} := s \begin{bmatrix} I_{n_J} & 0 \\ 0 & N \end{bmatrix} - \begin{bmatrix} J & 0 \\ 0 & I_{n_N} \end{bmatrix}, \; J \in \mathbb{R}^{n_J\times n_J},\; N \in \mathbb{R}^{n_N\times n_N} \text{ nilpotent},$$

where n_J and n_N are uniquely determined, and J and N are unique up to similarity. If N and J are in Jordan canonical form, the pencil $s\overline{E} - \overline{A}$ is said to be in *Weierstraß canonical form*. Unlike the quasi Weierstraß form, this is a canonical form for regular pencils, however the transformation matrices S and T are harder to compute numerically and the additional structure is not advantageous for our purposes.

For any regular system $[E, A, b] \in \mathbb{R}^{n\times(2n+1)}$, a corresponding system in quasi Weierstraß form is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} I_{n_J} & 0 \\ 0 & N \end{bmatrix}\begin{pmatrix} x_J(t) \\ x_N(t) \end{pmatrix} = \begin{bmatrix} J & 0 \\ 0 & I_{n_N} \end{bmatrix}\begin{pmatrix} x_J(t) \\ x_N(t) \end{pmatrix} + \begin{pmatrix} b_J \\ b_N \end{pmatrix} u(t), \tag{3.4}$$

where for $(z, u) \in \mathfrak{B}_{[\overline{E},\overline{A},\overline{b}]}$, we have partitioned the state and input vector in two components

$$\begin{pmatrix} x_J \\ x_N \end{pmatrix} := z,\ x_J \in \mathcal{L}^2_{\mathrm{loc}}(\mathbb{R} \to \mathbb{R}^{n_J}),\ x_N \in \mathcal{L}^2_{\mathrm{loc}}(\mathbb{R} \to \mathbb{R}^{n_N}),$$
$$\begin{pmatrix} b_J \\ b_N \end{pmatrix} := \overline{b} := Sb,\ b_J \in \mathbb{R}^{n_J},\ b_N \in \mathbb{R}^{n_N}.$$

We can immediately see that the dynamic behaviour of x_J and x_N is "decoupled" and, in particular, that x_J fulfils the ordinary differential equation

$$\dot{x}_J(t) = Jx_J(t) + b_J u(t).$$

The solution of the *nilpotent part* x_N, i. e. $\frac{\mathrm{d}}{\mathrm{d}t}Nx_N(t) = x_N(t) + b_N u(t)$, warrants some further investigation, which will be done in Proposition 3.4.2.

The solutions of the original and the transformed DAE are related as follows

$$(x, u) \in \mathfrak{B}_{[E,A,b]} \iff (z, u) = (T^{-1}x, u) \in \mathfrak{B}_{[\overline{E},\overline{A},\overline{b}]}.$$

The performance index of the transformed system (3.4) equals that of the nominal system (3.1) if the performance index of the transformed system is determined by the matrix

$$\begin{bmatrix} T^\top QT & T^\top h \\ (T^\top h)^\top & r \end{bmatrix}.$$

This is easily seen by evaluating the performance index

$$\begin{aligned}\overline{J}(t_{\mathrm{f}}, z, u) &= \int_0^{t_{\mathrm{f}}} \begin{pmatrix} z(t) \\ u(t) \end{pmatrix}^\top \begin{bmatrix} T^\top QT & T^\top h \\ (T^\top h)^\top & r \end{bmatrix}\begin{pmatrix} z(t) \\ u(t) \end{pmatrix} \mathrm{d}t \\ &= \int_0^{t_{\mathrm{f}}} \begin{pmatrix} Tz(t) \\ u(t) \end{pmatrix}^\top \begin{bmatrix} Q & h \\ h^\top & r \end{bmatrix}\begin{pmatrix} Tz(t) \\ u(t) \end{pmatrix} \mathrm{d}t \\ &= J(t_{\mathrm{f}}, x, u). \end{aligned} \tag{3.5}$$

It directly follows that

$$\begin{bmatrix} Q & h \\ h^\top & r \end{bmatrix} \geq 0 \iff \begin{bmatrix} T^\top QT & T^\top h \\ (T^\top h)^\top & r \end{bmatrix} = \begin{bmatrix} T & 0 \\ 0 & 1 \end{bmatrix}^\top \begin{bmatrix} Q & h \\ h^\top & r \end{bmatrix}\begin{bmatrix} T & 0 \\ 0 & 1 \end{bmatrix} \geq 0.$$

The optimal value for $z^0 \in \mathcal{V}^{\mathrm{diff}}_{[\overline{E},\overline{A},\overline{b}]} = T^{-1}\mathcal{V}^{\mathrm{diff}}_{[E,A,b]}$ is given by

$$\begin{aligned}\overline{J}^*(t_\mathrm{f}, \overline{E}z^0) &= \inf\{\overline{J}(t_\mathrm{f}, z, u) \mid (z,u) \in \mathfrak{B}_{[\overline{E},\overline{A},\overline{b}]}, (\overline{E}z)(0) = \overline{E}z^0\} \\ &= \inf\{J(t_\mathrm{f}, x, u) \mid (x,u) \in \mathfrak{B}_{[E,A,b]}, (SETT^{-1}x)(0) = SETz^0\} \\ &= J^*(t_\mathrm{f}, E(Tz^0)).\end{aligned}$$

For the system in quasi Weierstraß form (3.4), we will partition the matrix $T^\top QT$ and the vector $T^\top h$ into submatrices of the appropriate shape:

$$\begin{bmatrix} Q_J & Q_{JN} & h_J \\ Q_{JN}^\top & Q_N & h_N \\ h_J^\top & h_N^\top & r \end{bmatrix} := \begin{bmatrix} T^\top QT & T^\top h \\ (T^\top h)^\top & r \end{bmatrix},$$
$$Q_J \in \mathbb{R}^{n_J \times n_J},\ Q_{JN} \in \mathbb{R}^{n_J \times n_N},\ Q_N \in \mathbb{R}^{n_N \times n_N},\ h_J \in \mathbb{R}^{n_J},\ h_N \in \mathbb{R}^{n_N}.$$

To sum up, the equivalent performance index for the system in quasi Weierstraß form (3.4) is given by

$$\overline{J}\big(t_\mathrm{f}, \big(\begin{smallmatrix} x_J \\ x_N \end{smallmatrix}\big), u\big) = \int_0^{t_\mathrm{f}} \begin{pmatrix} x_J(t) \\ x_N(t) \\ u(t) \end{pmatrix}^\top \begin{bmatrix} Q_J & Q_{JN} & h_J \\ Q_{JN}^\top & Q_N & h_N \\ h_J^\top & h_N^\top & r \end{bmatrix} \begin{pmatrix} x_J(t) \\ x_N(t) \\ u(t) \end{pmatrix} \mathrm{d}t. \tag{3.6}$$

3.4 Differentiability of the input

When dealing with control of differential-algebraic equations, we face one main difficulty: the "control input" u cannot be chosen freely, as it needs to fulfil certain differentiability conditions. This is a problem when trying to find the optimal control input, as we need to know which inputs we can implement. In this section, we will determine a necessary differentiability condition for u. This will be useful in the following sections, where we will derive an augmented system from the original system, whose input can be chosen freely from $\mathcal{L}^2_{\mathrm{loc}}$. In a way, this new input can be seen as the "true" input of the differential-algebraic system.

Definition 3.4.1 (Input index)**.** Let $[E, A, b] \in \mathbb{R}^{n\times(2n+1)}$ be given in quasi Weierstraß form (3.4). The *input index* of the system is given by

$$\omega_u := \min\{i \in \mathbb{N}_0 \mid N^i b_N = 0\}.$$

The index does not depend on the chosen quasi Weierstraß form, as N is unique up to similarity, therefore

$$\begin{aligned}\forall S \in GL_{n_N}(\mathbb{R}) : \min\{i \in \mathbb{N}_0 \mid (S^{-1}NS)^i S^{-1} b_N = 0\} &= \min\{i \in \mathbb{N}_0 \mid S^{-1}N b_N = 0\} \\ &= \min\{i \in \mathbb{N}_0 \mid N^i b_N = 0\}.\end{aligned}$$

Note that

$$\omega_u = 0 \iff b_N = 0.$$

Proposition 3.4.2. *Consider the system* $[E, A, b] \in \mathbb{R}^{n\times(2n+1)}$ *in quasi Weierstraß form* (3.4). *Then for every* $(x, u) \in \mathfrak{B}_{[E,A,b]}$ *we have that*

$$x_N \stackrel{ae}{=} -\sum_{i=0}^{\omega_u-1} N^i b_N u^{(i)}, \tag{3.7}$$

where $u \in \mathcal{W}^{\omega_u-1,2}_{loc}(\mathbb{R} \to \mathbb{R})$ *for* $\omega_u \geq 2$.

Proof. For the purpose of this proposition, it is enough to consider the algebraic part of the DAE, i. e.

$$\frac{\mathrm{d}}{\mathrm{d}t} N x_N(t) = x_N(t) + b_N u(t). \tag{3.8}$$

In passing, we note that $\varphi \in \mathcal{W}^{1,2}_{\mathrm{loc}}(\mathbb{R} \to \mathbb{R}^n)$ if, and only if, $\varphi_1, \ldots, \varphi_n \in \mathcal{W}^{1,2}_{\mathrm{loc}}(\mathbb{R} \to \mathbb{R})$ and therefore $M\varphi \in \mathcal{W}^{1,2}_{\mathrm{loc}}(\mathbb{R} \to \mathbb{R}^l)$ for any $M \in \mathbb{R}^{l\times n}$.

Set $\left(\left(\begin{smallmatrix} x_J \\ x_N \end{smallmatrix}\right), u\right) \in \mathfrak{B}_{[E,A,b]}$, hence $(x_N, u) \in \mathfrak{B}_{[N,I_{n_N},b_N]}$. We proceed in several steps.

Step 1: We show that

$$\forall i \geq \omega_u : N^i x_N = 0. \tag{3.9}$$

For $i \geq \operatorname{nil\,ind} N \geq \omega_u$, we have $N^i = 0$ and therefore (3.9) follows.

Assume that (3.9) holds for $i > \omega_u$. Then we have

$$N^{i-1} x_N = N^{i-1}(x_N + b_N u) \underset{(3.8)}{\stackrel{\mathrm{ae}}{=}} N^{i-1} \frac{\mathrm{d}}{\mathrm{d}t} N x_N = \frac{\mathrm{d}}{\mathrm{d}t} N^i x_N \stackrel{(3.9)}{=} 0.$$

Therefore (3.9) is shown for $i - 1$.

Step 2: For $\omega_u = 0$, we have $x_N = N^0 x_N = 0$ and (3.7) follows.

For $\omega_u = 1$, we have $x_N = \frac{\mathrm{d}}{\mathrm{d}t}(N x_N) - b_N u = -b_N u$ and (3.7) is shown.

For $\omega_u \geq 2$, we show the following statement by induction:

$$\forall i \in \{0, \ldots, \omega_u - 2\} : u \in \mathcal{W}^{i+1,2}_{\mathrm{loc}}(\mathbb{R} \to \mathbb{R}) \text{ and} \tag{3.10}$$

$$N^{\omega_u-1} b_N u^{(i)} \stackrel{\mathrm{ae}}{=} -N^{\omega_u-1-i} x_N - \sum_{k=0}^{i-1} N^{\omega_u-1-i+k} b_N u^{(k)}. \tag{3.11}$$

For $i = 0$, we have

$$0 \stackrel{(3.9)}{=} \frac{\mathrm{d}}{\mathrm{d}t} N^{\omega_u} x_N = N^{\omega_u-1} \frac{\mathrm{d}}{\mathrm{d}t} N x_N \stackrel{\mathrm{ae}}{=} N^{\omega_u-1} x_N + N^{\omega_u-1} b_N u$$
$$\implies N^{\omega_u-1} b_N u \stackrel{\mathrm{ae}}{=} -N^{\omega_u-1} x_N \stackrel{(3.2)}{\in} \mathcal{W}^{1,2}_{\mathrm{loc}}(\mathbb{R} \to \mathbb{R}^n)$$

and so (3.11) follows. Furthermore, as $N^{\omega_u-1} b_N \neq 0$ and u is scalar, we have (3.10).

Assume that (3.10) and (3.11) hold for $i \in \{0, 1, \dots, \omega_u - 3\}$. Then we have that

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} N^{\omega_u - 1} b_N u^{(i)} &\underset{(3.11)}{\overset{\text{ae}}{=}} \frac{\mathrm{d}}{\mathrm{d}t}\Big(-N^{\omega_u - 1 - i} x_N - \sum_{k=0}^{i-1} N^{\omega_u - 1 - i + k} b_N u^{(k)} \Big) \\
&\underset{(3.8)}{\overset{\text{ae}}{=}} -N^{\omega_u - 1 - (i+1)} (x_N + b_N u) - \sum_{k=0}^{i-1} N^{\omega_u - 1 - i + k} b_N u^{(k+1)} \\
&= -\underbrace{N^{\omega_u - 1 - (i+1)} x_N}_{\overset{(3.2)}{\in} \mathcal{W}^{1,2}_{\mathrm{loc}}(\mathbb{R} \to \mathbb{R}^n)} - \sum_{k=0}^{(i+1)-1} \underbrace{N^{\omega_u - 1 - (i+1) + k} b_N u^{(k)}}_{\overset{(3.10)}{\in} \mathcal{W}^{1,2}_{\mathrm{loc}}(\mathbb{R} \to \mathbb{R}^n)}.
\end{aligned} \tag{3.12}$$

This shows (3.11) for $i + 1$.

Furthermore, as the right-hand side of (3.12) is in $\mathcal{W}^{1,2}_{\mathrm{loc}}(\mathbb{R} \to \mathbb{R}^n)$, $N^{\omega_u - 1} b_N \neq 0$ and u is scalar, it follows that $u^{(i+1)} \in \mathcal{W}^{1,2}_{\mathrm{loc}}(\mathbb{R} \to \mathbb{R})$. Therefore (3.10) holds for $i + 1$.

Step 3: For $\omega_u \geq 2$, reconsider (3.11) for $i = \omega_u - 2$: by (3.10), we know that $u \in \mathcal{W}^{i+1,2}(\mathbb{R} \to \mathbb{R})$, so (3.11) is differentiable. Carrying out the differentiation in the same was as in (3.12), we arrive at

$$N^{\omega_u - 1} b_N u^{(\omega_u - 1)} \overset{\text{ae}}{=} -x_N - \sum_{k=0}^{\omega_u - 2} N^k b_N u^{(k)}. \tag{3.13}$$

Rearranging (3.13) for x_N immediately gives (3.11).

This completes the proof of the proposition. □

Remark 3.4.3. In this thesis, we will not consider systems in quasi Weierstraß form (3.4) with $\omega_u = 0$: by Proposition 3.4.2, we know that for these systems $x_N \overset{\text{ae}}{=} 0$, so the performance index (3.6) becomes

$$\begin{aligned}
J(t_{\mathrm{f}}, \left(\begin{smallmatrix} x_J \\ x_N \end{smallmatrix}\right), u) &= \int_0^{t_{\mathrm{f}}} \begin{pmatrix} x_J(t) \\ x_N(t) \\ u(t) \end{pmatrix}^{\top} \begin{bmatrix} Q_J & Q_{JN} & h_J \\ Q_{JN}^{\top} & Q_N & h_N \\ h_J^{\top} & h_N^{\top} & r \end{bmatrix} \begin{pmatrix} x_J(t) \\ x_N(t) \\ u(t) \end{pmatrix} \mathrm{d}t \\
&= \int_0^{t_{\mathrm{f}}} \begin{pmatrix} x_J(t) \\ u(t) \end{pmatrix}^{\top} \begin{bmatrix} Q_J & h_J \\ h_J^{\top} & r \end{bmatrix} \begin{pmatrix} x_J(t) \\ u(t) \end{pmatrix} \mathrm{d}t.
\end{aligned}$$

This is a mere ordinary differential equation optimal control problem (with the ODE being $\dot{x}_J = J x_J + b_J u$), which can be solved using the standard repertoire, see e. g. [LR95].

3.5 An augmented system

Now that we know how many times the input u must be differentiable, we are in a position to reformulate the differential-algebraic system into a form which allows for a much easier solution of the optimal control problem, as its input can be chosen freely from $\mathcal{L}^2_{\mathrm{loc}}$. The basic idea is the following: instead of considering u as the input, we use its highest derivative we know to exist, which is $u^{(\omega_u-1)}$. The lower order derivatives $u, \dots, u^{(\omega_u-2)}$ are in turn introduced as new states of the augmented system. The relation between the solutions of the nominal and the augmented system is bijective: in order to obtain u from our new input $u^{(\omega_u-1)}$, we only need the "integration constants" $u(0), \dots, u^{(\omega_u-2)}(0)$, which correspond to the initial values of the augmented system.

Definition 3.5.1 (Augmented system)**.** For the system $[E, A, b] \in \mathbb{R}^{n\times(2n+1)}$ in quasi Weierstraß form (3.4) with $\omega_u > 0$, the *augmented system* $[\widehat{E}, \widehat{A}, \widehat{b}] \in \mathbb{R}^{(n+\omega_u-1)\times(2(n+\omega_u-1)+1)}$ is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\underbrace{\begin{bmatrix} I_{n_J} & 0 & 0 \\ 0 & I_{\omega_u-1} & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{=:\widehat{E}}\widehat{x} = \underbrace{\begin{bmatrix} J & b_J & 0 & 0 \\ 0 & 0 & I_{\omega_u-2} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & b_N & Nb_N, \dots, N^{\omega_u-2}b_N & I_{n_N} \end{bmatrix}}_{=:\widehat{A}}\widehat{x} + \underbrace{\begin{pmatrix} 0_{n_J} \\ 0_{\omega_u-2} \\ 1 \\ N^{\omega_u-1}b_N \end{pmatrix}}_{=:\widehat{b}}\widehat{u}, \tag{3.14}$$

where

$$\widehat{E} := \begin{bmatrix} I_{n_J+\omega_u-1} & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{(n+\omega_u-1)\times(n+\omega_u-1)},$$

$$\widehat{A} := \begin{bmatrix} A_1 & 0 \\ A_2 & I_{n_N} \end{bmatrix} \in \mathbb{R}^{(n+\omega_u-1)\times(n+\omega_u-1)},$$

$$A_1 := \begin{cases} J, & \omega_u = 1 \\ \begin{bmatrix} J & b_J & 0 \\ 0 & 0 & I_{\omega_u-2} \\ 0 & 0 & 0 \end{bmatrix}, & \omega_u > 1, \end{cases} \quad A_1 \in \mathbb{R}^{(n_J+\omega_u-1)\times(n_J+\omega_u-1)},$$

$$A_2 := \begin{cases} 0_{n_N\times n_J}, & \omega_u = 1 \\ \begin{bmatrix} 0_{n_N\times n_J} & b_N & Nb_N & \dots & N^{\omega_u-2}b_N \end{bmatrix}, & \omega_u > 1, \end{cases} \quad A_2 \in \mathbb{R}^{n_N\times(n_J+\omega_u-1)},$$

$$\widehat{b} := \begin{pmatrix} b_1 \\ N^{\omega_u-1}b_N \end{pmatrix} \in \mathbb{R}^{n+\omega_u-1},$$

$$b_1 := \begin{cases} b_J, & \omega_u = 1 \\ \begin{pmatrix} 0 & 0 & \dots & 0 & 1 \end{pmatrix}^\top, & \omega_u > 1, \end{cases} \quad b_1 \in \mathbb{R}^{n_J+\omega_u-1}.$$

Remark 3.5.2.

(a) For $\omega_u > 1$, the states of the augmented system $\widehat{x}$ consist of the states of the nominal system as well as the new states $u, \dots, u^{(\omega_u-2)}$, while the input of the augmented system $\widehat{u}$ is the $(\omega_u - 1)$-th derivative of the original input:

$$\widehat{x} = \begin{pmatrix} x_J, & u, & \dot{u}, & \dots, & u^{(\omega_u-2)}, & x_N \end{pmatrix}^\top,$$
$$\widehat{u} = u^{(\omega_u-1)}.$$

For $\omega_u = 1$, the states and input of the augmented system remain unaltered compared to the nominal system.

(b) Note that the differential equation for x_J is unchanged from the original system: by reading the first n_J columns of the system (3.14), we see that

$$\dot{x}_J = Jx_J + b_J u. \tag{3.15}$$

The only difference is that for $\omega_u > 1$, u is now a state of the system and not an external input any more.

(c) For $\omega_u > 1$, the dynamics of the new states $u, \dots, u^{(\omega_u-2)}$ are given by a simple "integrator chain" in the rows $n_J + 1$ to $n_J + \omega_u - 1$ of (3.14):

$$\begin{aligned} &\forall i = 0, \dots, \omega_u - 3 : \frac{\mathrm{d}}{\mathrm{d}t} u^{(i)} = \frac{\mathrm{d}}{\mathrm{d}t} \widehat{x}_{n_J+i+1} = \widehat{x}_{n_J+i+2} = u^{(i+1)} \\ &\frac{\mathrm{d}}{\mathrm{d}t} u^{(\omega_u-2)} = \frac{\mathrm{d}}{\mathrm{d}t} \widehat{x}_{n_J+\omega_u-1} = \widehat{u} = u^{(\omega_u-1)}. \end{aligned} \tag{3.16}$$

For $\omega_u = 1$, this part is not present.

(d) Finally, in the last n_N rows of (3.14), we find

$$0 = b_N u + N b_N \dot{u} + \dots + N^{\omega_u-2} b_N u^{(\omega_u-2)} + I_{n_N} x_N + N^{\omega_u-1} b_N u^{(\omega_u-1)}, \tag{3.17}$$

which rearranged yields the solution formula (3.7).

(e) The augmented system depends on the chosen quasi Weierstraß form of $[E, A, b]$ and is therefore not uniquely determined. However the dimensions of the extended system are uniquely determined as the input index is invariant for all quasi Weierstraß forms.

Example 3.5.3. Consider the nilpotent DAE in quasi Weierstraß form

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} x_{N,1}(t) \\ x_{N,2}(t) \end{pmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} x_{N,1}(t) \\ x_{N,2}(t) \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(t). \tag{3.18}$$

The DAE has input index $\omega_u = 2$, and the augmented system is given by

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} u(t) \\ x_{N,1}(t) \\ x_{N,2}(t) \end{pmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{pmatrix} u(t) \\ x_{N,1}(t) \\ x_{N,2}(t) \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \widehat{u}(t). \tag{3.19}$$

The relation between the behaviour of the nominal system and the behaviour of the augmented system is given in the following proposition.

Proposition 3.5.4. *Let* $[E, A, b] \in \mathbb{R}^{n \times (2n+1)}$ *be in quasi Weierstraß form* (3.4) *with* $\omega_u > 0$. *With the augmented system* $[\widehat{E}, \widehat{A}, \widehat{b}]$ *as in Definition 3.5.1, it holds that*

$$\left(\begin{pmatrix} x_J \\ x_N \end{pmatrix}, u \right) \in \mathfrak{B}_{[E,A,b]} \iff \left(\begin{pmatrix} x_J \\ u \\ \dot{u} \\ \vdots \\ u^{(\omega_u-2)} \\ x_N \end{pmatrix}, u^{(\omega_u-1)} \right) \in \mathfrak{B}_{[\widehat{E},\widehat{A},\widehat{b}]}. \tag{3.20}$$

Furthermore, $x_1 := (x_J^\top, \quad u, \quad \dots, \quad u^{(\omega_u-2)})^\top$ *fulfils the ordinary differential equation initial value problem*

$$\boxed{\begin{aligned} &\dot{x}_1(t) = A_1 x_1(t) + b_1 u^{(\omega_u-1)}(t), \\ &x_1(0) = \big(x_J(0), \quad u(0), \quad \dot{u}(0), \quad \dots, \quad u^{(\omega_u-2)}(0)\big)^\top \in \mathbb{R}^{n_J+\omega_u-1}. \end{aligned}} \tag{3.21}$$

Proof. This is a direct consequence of interpreting the DAE (3.14) row-wise as in eqs. (3.15) to (3.17), using the fact that $u \in \mathcal{W}^{\omega_u-1,2}_{\mathrm{loc}}(\mathbb{R} \to \mathbb{R})$ as stated in Proposition 3.4.2. □

The differential equation (3.21) will be of particular importance for the optimal control problem, as seen in the next section. It retains some controllability and stabilizability properties of the differential-algebraic system.

Lemma 3.5.5. *Let* $[E, A, b] \in \mathbb{R}^{n\times(2n+1)}$ *be in quasi Weierstraß form* (3.4) *with* $\omega_u > 0$*. Then* $[E, A, b]$ *is behaviourally controllable if, and only if, the ordinary differential equation* (3.21) *is controllable.*

Proof. By [BR13, Corollary 4.3], we have that

$$[E, A, b] \text{ is behaviourally controllable } \iff \forall \lambda \in \mathbb{C} : \mathrm{rk}_{\mathbb{C}}[\lambda E - A, b] = \mathrm{rk}_{\mathbb{R}(s)}[sE - A, b].$$

As $\mathrm{rk}_{\mathbb{C}}(\lambda N - I_{n_N}) = \mathrm{rk}_{\mathbb{R}(s)}(\lambda N - I_{n_N}) = n_N$ for all $\lambda \in \mathbb{C}$, we have

$$\forall \lambda \in \mathbb{C} : \mathrm{rk}_{\mathbb{C}}[\lambda E - A, b] = \mathrm{rk}_{\mathbb{C}} \begin{bmatrix} \lambda I_{n_J} - J & 0 & b_J \\ 0 & \lambda N - I_{n_N} & b_N \end{bmatrix} = n_N + \mathrm{rk}_{\mathbb{C}}[\lambda I_{n_J} - J, b_J]. \tag{3.22}$$

Together with $\mathrm{rk}_{\mathbb{R}(s)}[sE - A, b] = \mathrm{rk}_{\mathbb{R}(s)}(sE - A) = n$ due to regularity of $sE - A$, we obtain

$$[E, A, b] \text{ is behaviourally controllable } \iff \forall \lambda \in \mathbb{C} : \mathrm{rk}_{\mathbb{C}}[\lambda I_{n_J} - J, b_J] = n_J. \tag{3.23}$$

For $\omega_u = 1$, the assertion now directly follows by the Hautus criterion [TSH01, Theorem 3.13(i)]. For $\omega_u > 1$, we have, again by the Hautus criterion,

$$\begin{aligned} &\text{(3.21) is controllable} \\ \iff &\forall \lambda \in \mathbb{C} : \mathrm{rk}_{\mathbb{C}}[\lambda I - A_1, b_1] = \mathrm{rk}_{\mathbb{C}} \begin{bmatrix} \lambda I_{n_J} - J & b_J & 0 & 0 \\ 0 & 0 & I_{\omega_u-2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = n_J + \omega_u - 1 \\ \iff &\forall \lambda \in \mathbb{C} : \mathrm{rk}_{\mathbb{C}}[\lambda I_{n_J} - J, b_J] = n_J \\ \overset{(3.23)}{\iff} &[E, A, b] \text{ is behaviourally controllable.} \end{aligned}$$

Thus the lemma is proven. □

Lemma 3.5.6. *Let* $[E, A, b] \in \mathbb{R}^{n\times(2n+1)}$ *be in quasi Weierstraß form* (3.4) *with* $\omega_u > 0$*. Then* $[E, A, b]$ *is behaviourally stabilizable if, and only if, the ordinary differential equation* (3.21) *is stabilizable.*

Proof. By [BR13, Corollary 4.3] we have that

$$\begin{aligned}
& [E, A, b] \text{ is behaviourally stabilizable} \\
\iff & \forall \lambda \in \overline{\mathbb{C}}_+ : \mathrm{rk}_{\mathbb{C}}[\lambda E - A, b] = \mathrm{rk}_{\mathbb{R}(s)}[sE - A, b] = n \\
\overset{(3.22)}{\iff} & \forall \lambda \in \overline{\mathbb{C}}_+ : \mathrm{rk}_{\mathbb{C}}[\lambda I_{n_J} - J, b_J] = n_J.
\end{aligned}$$

The rest of the proof is completely analogous to Lemma 3.5.5. □

3.6 The optimal control problem for the augmented system

In this section, we will solve the optimal control problem for the nominal system $[E, A, b]$ by solving an equivalent optimal control problem for the augmented system $[\widehat{E}, \widehat{A}, \widehat{b}]$ as in Definition 3.5.1. As we will see, this actually turns out to be an *ordinary differential equation* optimal control problem, with the ODE given as in (3.21).

Lemma 3.6.1. *Let* $[E, A, b] \in \mathbb{R}^{n \times (2n+1)}$ *be in quasi Weierstraß form* (3.4) *with* $\omega_u > 0$. *The performance index for the system* (3.4) *is given as in* (3.6) *by*

$$J(t_f, x, u) = \int_0^{t_f} \begin{pmatrix} x_J(t) \\ x_N(t) \\ u(t) \end{pmatrix}^\top \begin{bmatrix} Q_J & Q_{JN} & h_J \\ Q_{JN}^\top & Q_N & h_N \\ h_J^\top & h_N^\top & r \end{bmatrix} \begin{pmatrix} x_J(t) \\ x_N(t) \\ u(t) \end{pmatrix} \mathrm{d}t.$$

Then for any $\left(\left(\begin{smallmatrix} x_J \\ x_N \end{smallmatrix} \right), u \right) \in \mathfrak{B}_{[E,A,b]}$ *we have that*

$$\begin{aligned}
J\left(t_f, \left(\begin{smallmatrix} x_J \\ x_N \end{smallmatrix} \right), u\right) &= \widehat{J}(t_f, x_1, u^{(\omega_u - 1)}) \\
&:= \int_0^{t_f} \begin{pmatrix} x_1(t) \\ u^{(\omega_u-1)}(t) \end{pmatrix}^\top \begin{bmatrix} \widehat{Q} & \widehat{h} \\ \widehat{h}^\top & \widehat{r} \end{bmatrix} \begin{pmatrix} x_1(t) \\ u^{(\omega_u-1)}(t) \end{pmatrix} \mathrm{d}t,
\end{aligned} \tag{3.24}$$

where x_1 *is the solution of the ordinary initial value problem* (3.21)

$$\begin{aligned}
\dot{x}_1(t) &= A_1 x_1(t) + b_1 u^{(\omega_u-1)}(t), \\
x_1(0) &= \left(x_J(0), \quad u(0), \quad \dot{u}(0), \quad \dots, \quad u^{(\omega_u-2)}(0) \right)^\top \in \mathbb{R}^{n_J + \omega_u - 1}
\end{aligned}$$

and, for $\omega_u = 1$,

$$\begin{aligned}
\widehat{Q} &:= Q_J \in \mathbb{R}^{n_J \times n_J}_{sym}, \\
\widehat{h} &:= h_J - Q_{JN} b_N \in \mathbb{R}^{n_J}, \\
\widehat{r} &:= r + b_N^\top Q_N b_N - b_N^\top h_N - h_N^\top b_N \in \mathbb{R},
\end{aligned}$$

while, for $\omega_u > 1$,

$$
\begin{aligned}
\widehat{Q} &:= Q_1 + A_2^\top Q_N A_2 - Q_2 A_2 - A_2^\top Q_2^\top \in \mathbb{R}_{sym}^{(n_J+\omega_u-1)\times(n_J+\omega_u-1)},\\
\widehat{h} &:= (A_2^\top Q_N - Q_2) N^{\omega_u-1} b_N \in \mathbb{R}^{n_J+\omega_u-1},\\
\widehat{r} &:= b_N^\top (N^{\omega_u-1})^\top Q_N N^{\omega_u-1} b_N \in \mathbb{R},\\
Q_1 &:= \begin{bmatrix} Q_J & h_J & 0\\ h_J^\top & r & 0\\ 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}_{sym}^{(n_J+\omega_u-1)\times(n_J+\omega_u-1)},\\
Q_2 &:= \begin{bmatrix} Q_{JN}\\ h_N^\top\\ 0 \end{bmatrix} \in \mathbb{R}^{(n_J+\omega_u-1)\times n_N}.
\end{aligned}
$$

Proof. Let $\omega_u > 1$, and $\left(\left(\begin{smallmatrix} x_J\\ x_N\end{smallmatrix}\right), u\right) \in \mathfrak{B}_{[E,A,b]}$ be arbitrary. It follows by (3.20) that

$$
J(t_\mathrm{f}, x, u) = \int_0^{t_\mathrm{f}} \begin{pmatrix} x_1(t)\\ x_N(t)\end{pmatrix}^\top \begin{bmatrix} Q_1 & Q_2\\ Q_2^\top & Q_N\end{bmatrix} \begin{pmatrix} x_1(t)\\ x_N(t)\end{pmatrix} \mathrm{d}t, \tag{3.25}
$$

with x_1 as defined in Proposition 3.5.4. By substituting

$$
x_N = -A_2 x_1 - N^{\omega_u-1} b_N u^{(\omega_u-1)}
$$

and rewriting the integrand of (3.25) as

$$
\begin{aligned}
\begin{pmatrix} x_1\\ x_N\end{pmatrix}^\top \begin{bmatrix} Q_1 & Q_2\\ Q_2^\top & Q_N\end{bmatrix} \begin{pmatrix} x_1\\ x_N\end{pmatrix} &= x_1^\top Q_1 x_1 + x_1^\top Q_2 x_N + x_N^\top Q_2^\top x_1 + x_N^\top Q_N x_N\\
&= x_1^\top Q_1 x_1 + x_1^\top Q_2 [-A_2 x_1 - N^{\omega_u-1} b_N u^{(\omega_u-1)}]\\
&\quad + [-A_2 x_1 - N^{\omega_u-1} b_N u^{(\omega_u-1)}]^\top Q_2^\top x_1\\
&\quad + [-A_2 x_1 - N^{\omega_u-1} b_N u^{(\omega_u-1)}]^\top Q_N [-A_2 x_1 - N^{\omega_u-1} b_N u^{(\omega_u-1)}]\\
&= x_1^\top \widehat{Q} x_1 + x_1^\top \widehat{h} u^{(\omega_u-1)} + u^{(\omega_u-1)} \widehat{h}^\top x_1 + u^{(\omega_u-1)} \widehat{r} u^{(\omega_u-1)},
\end{aligned}
$$

we arrive at

$$
J(t_\mathrm{f}, x, u) = \int_0^{t_\mathrm{f}} \begin{pmatrix} x_1(t)\\ u^{(\omega_u-1)}(t)\end{pmatrix}^\top \begin{bmatrix} \widehat{Q} & \widehat{h}\\ \widehat{h}^\top & \widehat{r}\end{bmatrix} \begin{pmatrix} x_1(t)\\ u^{(\omega_u-1)}(t)\end{pmatrix} \mathrm{d}t.
$$

From Proposition 3.5.4 we derive that x_1 fulfils the initial value problem (3.21). This shows the assertion for $\omega_u > 1$. For $\omega_u = 1$, completely analogous considerations can be made. □

Example 3.6.2. Consider the DAE

$$
\frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} 1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{bmatrix} \begin{pmatrix} x_J(t)\\ x_{N,1}(t)\\ x_{N,2}(t)\end{pmatrix} = \begin{bmatrix} -1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} \begin{pmatrix} x_J(t)\\ x_{N,1}(t)\\ x_{N,2}(t)\end{pmatrix} + \begin{pmatrix} 2\\ 0\\ 3\end{pmatrix} u(t)
$$

with the performance index

$$
J(t_\mathrm{f}, x, u) = \int_0^{t_\mathrm{f}} \left\| \left(x_J(t),\quad x_{N,1}(t),\quad x_{N,2}(t),\quad u(t)\right)^\top \right\|^2 \mathrm{d}t.
$$

The DAE has input index $\omega_u = 2$, and the augmented system is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix}\begin{pmatrix}x_J(t)\\u(t)\\x_{N,1}(t)\\x_{N,2}(t)\end{pmatrix}=\begin{bmatrix}-1&2&0&0\\0&0&0&0\\0&0&1&0\\0&3&0&1\end{bmatrix}\begin{pmatrix}x_J(t)\\u(t)\\x_{N,1}(t)\\x_{N,2}(t)\end{pmatrix}+\begin{pmatrix}0\\1\\3\\0\end{pmatrix}\widehat{u}(t).$$

The differential equation for the augmented system is

$$\dot{x}_1(t)=\begin{pmatrix}\dot{x}_J(t)\\\dot{u}(t)\end{pmatrix}=\begin{bmatrix}-1&2\\0&0\end{bmatrix}\begin{pmatrix}x_J(t)\\u(t)\end{pmatrix}+\begin{pmatrix}0\\1\end{pmatrix}\widehat{u}(t),$$

and the performance index for the augmented system is

$$\widehat{J}(t_\mathrm{f},x_1,\widehat{u})=\int_0^{t_\mathrm{f}}\begin{pmatrix}x_J(t)\\u(t)\\\widehat{u}(t)\end{pmatrix}^\top\begin{bmatrix}1&0&0\\0&10&0\\0&0&9\end{bmatrix}\begin{pmatrix}x_J(t)\\u(t)\\\widehat{u}(t)\end{pmatrix}\mathrm{d}t.$$

The previous lemma shows a relation between the performance indices of the nominal and the augmented system, where the latter can be calculated using only an ordinary differential equation. Our goal now is to be able to solve the DAE optimal control problem by completely reducing it to an ODE control problem, with the differential equation given by (3.21). In order to do so, we still need to determine a relation between the initial differential values of the nominal and the augmented system. The following lemma will help us in this process, as it effectively gives a bijection between the initial values of both systems, as further studied in Theorem 3.6.6.

Lemma 3.6.3. *Let* $[E,A,b]\in\mathbb{R}^{n\times(2n+1)}$ *be in quasi Weierstraß form* (3.4) *with* $\omega_u>0$. *Then the matrix*

$$\begin{bmatrix}b_N, & Nb_N, & \dots, & N^{\omega_u-1}b_N\end{bmatrix}\in\mathbb{R}^{n_N\times\omega_u}$$

has full (column) rank $\omega_u\le n_N$.

Proof. We show by induction for $0\le i\le\omega_u-1$ that

$$\operatorname{rk}\begin{bmatrix}N^ib_N, & N^{i+1}b_N, & \dots, & N^{\omega_u-1}b_N\end{bmatrix}=\omega_u-i. \tag{3.26}$$

For $i=\omega_u-1$, (3.26) follows from $N^{\omega_u-1}b_N\neq 0$.

Assume that (3.26) holds for some $i\in\{1,\dots,\omega_u-1\}$. Consider the linear combination

$$\alpha_{i-1}N^{i-1}b_N+\cdots+\alpha_{\omega_u-1}N^{\omega_u-1}b_N=0 \tag{3.27}$$

for $\alpha_{i-1},\dots,\alpha_{\omega_u-1}\in\mathbb{R}$. Multiplying (3.27) from the left by N, we arrive at

$$0=\alpha_{i-1}N^ib_N+\cdots+\alpha_{\omega_u-2}N^{\omega_u-1}b_N+\alpha_{\omega_u-1}N^{\omega_u}b_N=\alpha_{i-1}N^ib_N+\cdots+\alpha_{\omega_u-2}N^{\omega_u-1}b_N.$$

By (3.26), we see that

$$\alpha_{i-1}=\cdots=\alpha_{\omega_u-2}=0.$$

Therefore by (3.27) and $N^{\omega_u-1}b_N\neq 0$, we obtain $\alpha_{\omega_u-1}=0$ and hence (3.26) for $i-1$.

Now (3.26) for $i=0$ shows the assertion. □

Before we state the main result, we stress that a sufficient condition for the optimal value of the augmented system to be finite is the assumption $\begin{bmatrix} Q_J & Q_{JN} & h_J \\ Q_{JN}^\top & Q_N & h_N \\ h_J^\top & h_N^\top & r \end{bmatrix} \geq 0$ for the nominal system (3.4) with the performance index (3.6), which can be transferred to the augmented system.

Lemma 3.6.4. *Let* $[E, A, b] \in \mathbb{R}^{n\times(2n+1)}$ *be in quasi Weierstraß form* (3.4) *with* $\omega_u > 0$ *and the performance index as defined in Lemma 3.6.1. Then we have*

$$\begin{bmatrix} Q_J & Q_{JN} & h_J \\ Q_{JN}^\top & Q_N & h_N \\ h_J^\top & h_N^\top & r \end{bmatrix} \geq 0 \implies \begin{bmatrix} \widehat{Q} & \widehat{h} \\ \widehat{h}^\top & \widehat{r} \end{bmatrix} \geq 0.$$

In particular, this means $\widehat{Q} \geq 0, \widehat{r} \geq 0$.

Proof. Seeking a contradiction, suppose that there exists $\left(\begin{smallmatrix} \tilde{x} \\ \tilde{u} \end{smallmatrix}\right) \in \mathbb{R}^{n_J+\omega_u}$ such that

$$\ell(\tilde{x}) := \begin{pmatrix} \tilde{x} \\ \tilde{u} \end{pmatrix}^\top \begin{bmatrix} \widehat{Q} & \widehat{h} \\ \widehat{h}^\top & \widehat{r} \end{bmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{u} \end{pmatrix} < 0.$$

Let x_1 be the solution of

$$\dot{x}_1(t) = A_1 x_1(t) + b_1 \tilde{u}, \quad x_1(0) = \tilde{x}.$$

By continuity of $\ell(x_1(\cdot))$ and $\ell(x_1(0)) = \ell(\tilde{x}) < 0$, there exists a $t_\mathrm{f} > 0$ such that $\ell(t) < 0$ for all $t \in [0, t_\mathrm{f})$. By Proposition 3.5.4, we have that

$$\begin{aligned} x_J &= [I_{n_J}, 0_{n_J\times(\omega_u-1)}]\, x_1, \\ x_N &= -A_2 x_1 - N^{\omega_u-1} b_N \tilde{u}, \\ u &= \begin{cases} \tilde{u}, & \omega_u = 1 \\ [0_{1\times n_J}, 1, 0_{1\times(\omega_u-2)}]\, x_1, & \omega_u > 1 \end{cases} \end{aligned} \tag{3.28}$$

fulfil

$$\left(\begin{pmatrix} x_J \\ x_N \end{pmatrix}, u \right) \in \mathfrak{B}_{[E,A,b]}.$$

Therefore by Lemma 3.6.1, we obtain

$$\begin{aligned} 0 > \int_0^{t_\mathrm{f}} \ell(x_1(t))\,\mathrm{d}t &= \widehat{J}(t_\mathrm{f}, x_1, \tilde{u}) = J\big(t_\mathrm{f}, \left(\begin{smallmatrix} x_J \\ x_N \end{smallmatrix}\right), u\big) \\ &= \int_0^{t_\mathrm{f}} \begin{pmatrix} x_J(t) \\ x_N(t) \\ u(t) \end{pmatrix}^\top \begin{bmatrix} Q_J & Q_{JN} & h_J \\ Q_{JN}^\top & Q_N & h_N \\ h_J^\top & h_N^\top & r \end{bmatrix} \begin{pmatrix} x_J(t) \\ x_N(t) \\ u(t) \end{pmatrix} \mathrm{d}t, \end{aligned}$$

in contrast to the assumption. □

Remark 3.6.5. Note that the converse equality generally does not hold: as an example, for the DAE (3.18) with the performance index (3.6) given by

$$\begin{bmatrix} Q_N & h_N \\ h_N^\top & r \end{bmatrix} := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

which is not positive semidefinite, we still have

$$\begin{bmatrix} \widehat{Q} & \widehat{h} \\ \widehat{h}^\top & \widehat{r} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} > 0.$$

So the positive semidefiniteness of the DAE performance index (3.6) is a sufficient condition for the positive semidefiniteness of the performance index of the augmented system (3.24), however we only require the latter in the following considerations.

Now we are able to state the main result of this section: to solve the optimal control problem for the differential-algebraic system, we can instead solve the corresponding problem for the augmented system. However, for the augmented system this merely equates to solving an optimal control problem for an ordinary differential equation, which is a well understood area [CA78; LR95].

Theorem 3.6.6. *Let* $[E, A, b] \in \mathbb{R}^{n\times(2n+1)}$ *be in quasi Weierstraß form* (3.4) *with* $\omega_u > 0$, *equipped with the performance index* (3.6). *Then the optimal value is given by*

$$\begin{aligned} \forall x^0 \in \mathcal{V}^{diff}_{[E,A,b]}: \quad & J^*(t_f, Ex^0) = \widehat{J}^*(t_f, Ex^0) \\ & := \inf\left\{ \widehat{J}(t_f, x_1, \widehat{u}) \;\middle|\; \begin{array}{l} \widehat{u} \in \mathcal{L}^2_{loc}(\mathbb{R} \to \mathbb{R}), \\ \dot{x}_1(t) = A_1 x_1(t) + b_1 \widehat{u}(t),\ x_1(0) = F_1^{-1} E x^0 \end{array} \right\}, \end{aligned} \tag{3.29}$$

with $\widehat{J}$ *as defined in* (3.24) *and*

$$F_1 := \begin{bmatrix} [I_{n_J}, 0_{n_J\times(\omega_u-1)}] \\ -NA_2 \end{bmatrix} \in \mathbb{R}^{n\times(n_J+\omega_u-1)} \text{ left invertible.}$$

Proof. First, note that

$$F_1 = \begin{bmatrix} I_{n_J} & 0 \\ 0 & -[Nb_N, N^2 b_N, \dots, N^{\omega_u-1} b_N] \end{bmatrix}$$

has full rank according to Lemma 3.6.3 and therefore has a left inverse F_1^{-1}.

Let $x^0 \in \mathbb{R}^n$ and $\varepsilon > 0$ be arbitrary. We proceed in two steps:

Step 1: We show

$$J^*(t_f, Ex^0) \le \widehat{J}^*(t_f, Ex^0). \tag{3.30}$$

For $\widehat{J}^*(t_f, Ex^0) = +\infty$, there is nothing to show. Otherwise, choose $\widehat{u} \in \mathcal{L}^2_{\text{loc}}(\mathbb{R} \to \mathbb{R})$ such that

$$\widehat{J}(t_f, x_1, \widehat{u}) \le \begin{cases} \widehat{J}^*(t_f, Ex^0) + \varepsilon, & \widehat{J}^*(t_f, Ex^0) > -\infty \\ -\varepsilon, & \widehat{J}^*(t_f, Ex^0) = -\infty \end{cases},$$

where x_1 is the solution of the initial value problem

$$\dot{x}_1(t) = A_1 x_1(t) + b_1 \widehat{u}(t), \; x_1(0) = F_1^{-1} E x^0.$$

Define x_J, x_N and u as in (3.28), then we see that

$$\left(\begin{pmatrix} x_1 \\ x_N \end{pmatrix}, \widehat{u} \right) \in \mathfrak{B}_{[\widehat{E},\widehat{A},\widehat{B}]} \overset{\text{Prop. 3.5.4}}{\Longrightarrow} (x, u) := \left(\begin{pmatrix} x_J \\ x_N \end{pmatrix}, u \right) \in \mathfrak{B}_{[E,A,b]}.$$

Furthermore, we have

$$(Ex)(0) = \begin{pmatrix} x_J(0) \\ (Nx_N)(0) \end{pmatrix} = \begin{bmatrix} [I_{n_J}, 0_{n_J \times m\, n_N}] \\ -NA_2 \end{bmatrix} x_1(0) = F_1 x_1(0) = Ex^0.$$

So by Lemma 3.6.1 we get

$$J^*(t_f, Ex^0) \le J(t_f, x, u) = \widehat{J}(t_f, x_1, \widehat{u}) \le \begin{cases} \widehat{J}^*(t_f, Ex^0) + \varepsilon, & \widehat{J}^*(t_f, Ex^0) > -\infty \\ -\varepsilon, & \widehat{J}^*(t_f, Ex^0) = -\infty \end{cases}.$$

As $\varepsilon > 0$ is arbitrary, we obtain (3.30).

Step 2: We prove

$$\widehat{J}^*(t_f, Ex^0) \le J^*(t_f, Ex^0). \tag{3.31}$$

For $J^*(t_f, Ex^0) = +\infty$, there is nothing to show. Otherwise, choose $(x, u) \in \mathfrak{B}_{[E,A,b]}$ such that $(Ex)(0) = Ex^0$ and

$$J(t_f, x, u) \le \begin{cases} J^*(t_f, Ex^0) + \varepsilon, & J^*(t_f, Ex^0) > -\infty \\ -\varepsilon, & J^*(t_f, Ex^0) = -\infty \end{cases}.$$

By Proposition 3.5.4, we have that x_1 solves the initial value problem

$$\dot{x}_1(t) = A_1 x_1(t) + b_1 u^{(\omega_u - 1)}(t),$$
$$x_1(0) = \left(x_J(0), \quad u(0), \quad \dot{u}(0), \quad \dots, \quad u^{(\omega_u - 2)}(0) \right).$$

Furthermore,

$$(Nx_N)(0) = -NA_2 x_1(0) - NN^{\omega_u - 1} b_N u(0) = -NA_2 x_1(0),$$

and therefore

$$F_1 x_1(0) = \begin{bmatrix} [I_{n_J}, 0_{n_J \times (\omega_u - 1)}] \\ -NA_2 \end{bmatrix} x_1(0) = \begin{pmatrix} x_J(0) \\ (Nx_N)(0) \end{pmatrix} = (Ex)(0).$$

So we have that

$$\widehat{J}^*(t_f, Ex^0) \le \widehat{J}(t_f, x_1, u^{(\omega_u - 1)}) = J(t_f, x, u) \le \begin{cases} J^*(t_f, Ex^0) + \varepsilon, & J^*(t_f, Ex^0) > -\infty \\ -\varepsilon, & J^*(t_f, Ex^0) = -\infty \end{cases}.$$

For $\varepsilon \to 0$ or $\varepsilon \to \infty$, respectively, we obtain (3.31).

Now (3.30) and (3.31) prove the assertion. □

Remark 3.6.7. Theorem 3.6.6 shows for $t_\mathrm{f} < \infty$ that

$$\begin{bmatrix} \widehat{Q} & \widehat{h} \\ \widehat{h}^\top & \widehat{r} \end{bmatrix} \geq 0 \tag{3.32}$$

is a sufficient condition for the DAE optimal value (2.4) to be finite, as it ensures that $\widehat{J}$ as defined in (3.24) is bounded from below by 0. This is the condition we will use for the following results, instead of the more restrictive $\begin{bmatrix} Q_J & Q_{JN} & h_J \\ Q_{JN}^\top & Q_N & h_N \\ h_J^\top & h_N^\top & r \end{bmatrix} \geq 0$.

3.7 Finite horizon optimal control

Now that we have transformed the differential-algebraic optimal control problem into an equivalent ordinary differential equation one, we are also able to state the solution explicitly, provided that the ODE optimal control problem is non-singular.

Proposition 3.7.1. *Let $[E, A, b] \in \mathbb{R}^{n\times(2n+1)}$ be in quasi Weierstraß form* (3.4)*, with $\omega_u > 0$ and the performance index defined as in* (3.6) *for $t_f < \infty$. If* (3.32) *holds with $\widehat{r} > 0$, then the unique optimal control u^* is given by the output of the time-varying ordinary linear differential equation system*

$$\begin{aligned} \dot{x}_1^*(t) &= \Big[A_1 - b_1 \frac{1}{\widehat{r}}(b_1^\top P(t) + \widehat{h}^\top)\Big] x_1^*(t), \quad x_1^*(0) = F_1^{-1} E x^0, \\ u^*(t) &= \begin{cases} -\frac{1}{\widehat{r}}(b_1^\top P(t) + \widehat{h}^\top)\, x_1^*(t), & \omega_u = 1 \\ [0_{1\times n_J}, 1, 0_{1\times(\omega_u-2)}]\, x_1^*(t), & \omega_u > 1, \end{cases} \end{aligned}$$

where $P \in \mathcal{W}^{1,1}_{loc}(\mathbb{R} \to \mathbb{R}^{(n_J+\omega_u-1)\times(n_J+\omega_u-1)}_{sym})$ is the solution of the Riccati equation

$$\dot{P}(t) = -A_1^\top P(t) - P(t) A_1 - \widehat{Q} + (P(t) b_1 + \widehat{h}) \frac{1}{\widehat{r}} (b_1^\top P(t) + \widehat{h}^\top), \quad P(t_f) = 0.$$

Proof. It is well-known, see e. g. [LR95, Lemma 16.4.2, Theorem 16.4.3], that

$$\widehat{u}^*(\cdot) = -\frac{1}{\widehat{r}}(b_1^\top P(\cdot) + \widehat{h}^\top) x_1(\cdot)$$

is the unique optimal control for the ordinary differential equation control problem, i. e.

$$\widehat{J}^*(t_\mathrm{f}, Ex^0) = \widehat{J}(t_\mathrm{f}, x_1^*, \widehat{u}^*),$$

where x_1^* is the solution of the ordinary differential equation $\dot{x}_1^*(t) = \big[A_1 - \frac{1}{\widehat{r}}(b_1^\top P(t) + h^\top)\big] x_1^*(t)$, $x_1^*(0) = F_1^{-1} E x^0$. By defining x_J^*, x_N^* and u^* as in (3.28), we therefore get $\left(\begin{pmatrix} x_J^* \\ x_N^* \end{pmatrix}, u^* \right) \in \mathfrak{B}_{[E,A,b]}$, and by Lemma 3.6.1 and Theorem 3.6.6 it follows that

$$J^*(t_\mathrm{f}, Ex^0) = \widehat{J}^*(t_\mathrm{f}, Ex^0) = \widehat{J}(t_\mathrm{f}, x_1^*, \widehat{u}^*) = J\Big(t_\mathrm{f}, \begin{pmatrix} x_J^* \\ x_N^* \end{pmatrix}, u^*\Big).$$

Therefore u^* is an optimal control.

To prove the uniqueness, assume that we have $(\tilde{x}, \tilde{u}) \in \mathfrak{B}_{[E,A,b]}$ with $(E\tilde{x})(0) = Ex^0$ and $J(t_\mathrm{f}, \tilde{x}, \tilde{u}) = J^*(t_\mathrm{f}, Ex^0)$. By Lemma 3.6.1 and Theorem 3.6.6 we have

$$\widehat{J}(t_\mathrm{f}, \tilde{x}_1, \widehat{\tilde{u}}) = J(t_\mathrm{f}, \tilde{x}, \tilde{u}) = J^*(t_\mathrm{f}, Ex^0) = \widehat{J}^*(t_\mathrm{f}, Ex^0),$$

where $\tilde{x}_1$ and $\widehat{\tilde{u}}$ are defined as in (3.28). As u^* is the unique optimal control of the ODE (3.21), we have that $\widehat{\tilde{u}} = \widehat{u}^*$ and hence $\tilde{x}_1 = x_1^*$. For $\omega_u = 1$, this directly shows the uniqueness of u^* as $\tilde{u} = \widehat{\tilde{u}} = \widehat{u}^* = u^*$. For $\omega_u > 1$, we see by Proposition 3.5.4 that

$$\tilde{u} = \tilde{x}_{1,n_J+1} = x^*_{1,n_J+1} = u^*,$$

which shows uniqueness of u^* and concludes the proof. □

Remark 3.7.2. For $\omega_u > 1$, the assumption $\widehat{r} > 0$ in the previous proposition is notably fulfilled if $Q_N > 0$. Furthermore, the former assumption does not depend on the chosen quasi Weierstraß form (3.4): let

$$\left[\begin{bmatrix} I_{n_J} & 0 \\ 0 & N_1 \end{bmatrix}, \begin{bmatrix} J_1 & 0 \\ 0 & I_{n_N} \end{bmatrix}, \begin{pmatrix} b_{J,1} \\ b_{N,1} \end{pmatrix}\right], \left[\begin{bmatrix} I_{n_J} & 0 \\ 0 & N_2 \end{bmatrix}, \begin{bmatrix} J_2 & 0 \\ 0 & I_{n_N} \end{bmatrix}, \begin{pmatrix} b_{J,2} \\ b_{N,2} \end{pmatrix}\right]$$

be two quasi Weierstraß form representations of the regular system $[E, A, b] \in \mathbb{R}^{n\times(2n+1)}$. As N_1 and N_2 are similar, we obtain for the system matrices and performance indices of the systems in quasi Weierstraß form that

$$\exists T_N \in GL_{n_N}(\mathbb{R}) : N_2 = T_N^{-1} N_1 T_N,\ b_{N,2} = T_N^{-1} b_{N,1},\ Q_{N,2} = T_N^\top Q_{N,1} T_N,\ h_{N,2} = T_N^\top h_{N,1}.$$

For $\omega_u \geq 2$, it follows that

$$\begin{aligned} \widehat{r}_2 &= b_{N,2}^\top (N_2^{\omega_u-1})^\top Q_{N,2} N_2^{\omega_u-1} b_{N,2} \\ &= b_{N,1}^\top T_N^{-\top} (T_N^{-1} N_1^{\omega_u-1} T_N)^\top T_N^\top Q_{N,1} T_N T_N^{-1} N_1^{\omega_u-1} T_N T_N^{-1} b_{N,1} \\ &= b_{N,1}^\top (N_1^{\omega_u-1})^\top Q_{N,1} N_1^{\omega_u-1} b_{N,1} \\ &= \widehat{r}_1. \end{aligned}$$

For $\omega_u = 1$ an analogue derivation shows $\widehat{r}_1 = \widehat{r}_2$.

3.8 Infinite horizon optimal control

In the previous section, we have solved the DAE optimal control problem for a finite time horizon $t_\mathrm{f} < \infty$. In this section, we turn our attention to the infinite horizon case $t_\mathrm{f} = \infty$. This poses some additional challenges, as it is not guaranteed any more that J^* is finite. On the other hand, we will see that the infinite horizon allows us to implement the optimal control as a static state feedback of the nominal DAE system.

Proposition 3.8.1. *Let* $[E, A, b] \in \mathbb{R}^{n\times(2n+1)}$ *be in quasi Weierstraß form* (3.4)*, with* $\omega_u > 0$ *and the performance index defined as in* (3.6)*. Assume further that* (3.32) *holds with* $\operatorname{rk}\left[\begin{smallmatrix} \widehat{Q} & \widehat{h} \\ \widehat{h}^\top & \widehat{r} \end{smallmatrix}\right] = \operatorname{rk}\widehat{Q} + 1$ *and* $\widehat{r} > 0$*,* $[E, A, b]$ *is behaviourally stabilizable, and* $(\widehat{Q}, A_1)$ *is observable. Then the optimal control* u^* *is given by the output of the ordinary linear differential equation system*

$$\dot{x}_1^*(t) = \left[A_1 - b_1\frac{1}{\widehat{r}}(b_1^\top P + \widehat{h}^\top)\right]x_1^*(t), \quad x_1^*(0) = F_1^{-1}Ex^0,$$

$$u^*(t) = \begin{cases} -\frac{1}{\widehat{r}}(b_1^\top P + \widehat{h}^\top)\, x_1^*(t), & \omega_u = 1 \\ [0_{1\times n_J}, 1, 0_{1\times(\omega_u-2)}]\, x_1^*(t), & \omega_u > 1, \end{cases}$$

where $P \in \mathbb{R}_{sym}^{(n_J+\omega_u-1)\times(n_J+\omega_u-1)}$ *is given by the unique symmetric, positive definite solution of the algebraic Riccati equation (ARE)*

$$A_1^\top P + PA_1 + \widehat{Q} - (Pb_1 + \widehat{h})\frac{1}{\widehat{r}}(b_1^\top P + \widehat{h}^\top) = 0. \tag{3.33}$$

Proof. Note that by Lemma 3.5.6 the pair (A_1, b_1) is stabilizable. Together with the assumption that $(\widehat{Q}, A_1)$ is observable, it follows that the solution of the ODE optimal control problem (3.29) satisfies [LR95, Theorem 16.3.3]

$$\widehat{J}^*(\infty, Ex^0) = \widehat{J}(\infty, x_1^*, -\frac{1}{\widehat{r}}(b_1^\top P + \widehat{h}^\top)x_1^*(\cdot)) < \infty.$$

The rest of the proof is analoguous to Proposition 3.7.1. □

Under some further assumptions, the optimal control can also be implemented as a state feedback of the nominal DAE system (3.4). Note that we need to make sure that the closed-loop system is still regular, otherwise it might have multiple solutions, as shown in the following example.

Example 3.8.2. Consider the nilpotent DAE (3.18). By implementing the feedback

$$u(\cdot) = -\begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}x_{N,1}(\cdot) \\ x_{N,2}(\cdot)\end{pmatrix} = -x_{N,2}(\cdot), \tag{3.34}$$

the closed-loop system becomes singular:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}0 & 1 \\ 0 & 0\end{bmatrix}\begin{pmatrix}x_{N,1}(t) \\ x_{N,2}(t)\end{pmatrix} = \begin{bmatrix}1 & 0 \\ 0 & 0\end{bmatrix}\begin{pmatrix}x_{N,1}(t) \\ x_{N,2}(t)\end{pmatrix}, \text{ where } \det\begin{bmatrix}-1 & s \\ 0 & 0\end{bmatrix} = 0.$$

$x_{N,2}$ is a "free variable" which is not restricted by the system dynamics: the behaviour is given by

$$\left\{\begin{pmatrix}x_{N,1} \\ x_{N,2}\end{pmatrix} \in \mathcal{L}^2_{\text{loc}}(\mathbb{R} \to \mathbb{R}^2) \,\middle|\, x_{N,2} \in \mathcal{W}^{1,2}_{\text{loc}}(\mathbb{R} \to \mathbb{R}) \wedge x_{N,1} = \dot{x}_{N,2}\right\}.$$

Note that this happens because we are giving the system "redundant information" by implementing the feedback (3.34): the behaviour of the open-loop system (3.18) is

$$\left\{\left(\begin{pmatrix}x_{N,1} \\ x_{N,2}\end{pmatrix}, u\right) \in \mathcal{L}^2_{\text{loc}}(\mathbb{R} \to \mathbb{R}^2 \times \mathbb{R}) \,\middle|\, u \in \mathcal{W}^{1,2}_{\text{loc}}(\mathbb{R} \to \mathbb{R}) \wedge x_{N,1} = -\dot{u} \wedge x_{N,2} = -u\right\},$$

and we see that the feedback (3.34) only implements the condition $u = -x_{N,2}$, which was already present in the open-loop system.

Lemma 3.8.3. *Let* $[E, A, b] \in \mathbb{R}^{n\times(2n+1)}$ *be in quasi Weierstraß form* (3.4) *and* $k = (k_J, k_N) \in \mathbb{R}^{1\times n_J} \times \mathbb{R}^{1\times n_N}$*. Then*

$$\begin{aligned} &[E, A - bk, 0] \text{ is singular} \\ \iff\ & k_J(sI_{n_J} - J)^{-1}b_J = 0 \ \wedge\ k_N\,[b_N, Nb_N, \ldots, N^{\omega_u-1}b_N] = (1, 0, \ldots, 0). \end{aligned} \tag{3.35}$$

Proof. Since $[E, A, b]$ is in quasi Weierstraß form, it follows that $(sE - A)$ is regular. Therefore, the Sherman-Morrison-Woodbury formula [Ber09, Fact 2.16.3]

$$\det((sE - A) + bk) = (1 + k(sE - A)^{-1}b)\det(sE - A)$$

yields

$$sE - (A - bk) \text{ is singular} \iff 1 + k(sE - A)^{-1}b = 0. \tag{3.36}$$

Since $[E, A, b]$ is in quasi Weierstraß form (3.4) we have

$$\begin{aligned} k(sE - A)^{-1}b &= \begin{pmatrix} k_J & k_N \end{pmatrix} \begin{bmatrix} (sI_{n_J} - J)^{-1} & 0 \\ 0 & (sN - I_{n_N})^{-1} \end{bmatrix} \begin{pmatrix} b_J \\ b_N \end{pmatrix} \\ &= k_J(sI_{n_J} - J)^{-1}b_J + k_N(sN - I_{n_N})^{-1}b_N \in \mathbb{R}(s). \end{aligned} \tag{3.37}$$

By [Ber09, Equation (4.4.23)], we have

$$(sI_{n_J} - J)^{-1} = \frac{\sum_{i=0}^{n_J-1} J_i s^i}{\det(sI_{n_J} - J)} \ \text{ for } J_0, \ldots, J_{n_J-2} \in \mathbb{R}^{n_J\times n_J} \text{ and } J_{n_J-1} = I_{n_J} \tag{3.38}$$

and the nilpotency of N gives

$$(sN - I_{n_N})^{-1} = -\sum_{i=0}^{\operatorname{nil\,ind} N-1} N^i s^i. \tag{3.39}$$

"⇒": Suppose that $(sE - (A - bk))$ is singular. Then (3.37) applied to (3.36), using the two identities (3.38) and (3.39) yields

$$-\det(sI_{n_J} - J) = \sum_{i=0}^{n_J-1} k_J J_i b_J s^i - \sum_{i=0}^{\operatorname{nil\,ind} N-1} k_N N^i b_N s^i \cdot \det(sI_{n_J} - J). \tag{3.40}$$

Since

$$\det(sI_{n_J} - J) = s^{n_J} + \alpha_{n_J-1}s^{n_J-1} + \cdots + \alpha_0 s^0 \in \mathbb{R}[s]$$

we conclude from (3.40) that

$$\sum_{i=0}^{\operatorname{nil\,ind} N-1} k_N N^i b_N s^i = k_N b_N$$

and therefore

$$-1 = -k_N b_N, \quad k_N[Nb_N, N^2b_N, \ldots, N^{\omega_u-1}b_N] = 0_{n_N\times\omega_u}. \tag{3.41}$$

Substituting (3.41) in (3.40) yields

$$\sum_{i=0}^{n_J-1} k_J J_i b_J s^i = 0.$$

This applied to (3.38) shows the first of the necessary conditions in (3.35). The second necessary condition was already shown in (3.41).

"$\Leftarrow$": Substituting the necessary conditions in (3.35) into (3.37) together with (3.39) directly shows

$$k(sE - A)^{-1} b = -1$$

and singularity of $[E, A - bk, 0]$ follows from (3.36). This concludes the proof. □

Now we are able to state the optimal feedback in terms of the nominal DAE in quasi Weierstraß form.

Proposition 3.8.4. *Let* $[E, A, b] \in \mathbb{R}^{n\times(2n+1)}$ *be in quasi Weierstraß form* (3.4) *with* $\omega_u > 0$ *and the performance index defined as in* (3.6)*. Assume further that* (3.32) *holds with* $\operatorname{rk}\begin{bmatrix} \widehat{Q} & \widehat{h} \\ \widehat{h}^\top & \widehat{r} \end{bmatrix} = \operatorname{rk}\widehat{Q} + 1$ *and* $\widehat{r} > 0$*,* $[E, A, b]$ *is behaviourally stabilizable, and* $(\widehat{Q}, A_1)$ *is observable. Let* P *be defined as in Proposition 3.8.1 and set*

$$p := \frac{1}{\widehat{r}}(b_1^\top P + \widehat{h}^\top) \in \mathbb{R}^{1\times(n_J+\omega_u-1)}.$$

If $\omega_u = 1$*, then the optimal control for the system* $[E, A, b]$ *is given by the feedback*

$$u(\cdot) = -\underbrace{p\begin{bmatrix} I_{n_J} & 0_{n_J\times n_N}\end{bmatrix}}_{=:k} x(\cdot). \tag{3.42}$$

If $\omega_u > 1$ *and* $p_{n_J+1} \neq 0$*, then the optimal control is given by the feedback*

$$u(\cdot) = -\underbrace{\frac{1}{p_{n_J+1}}\begin{pmatrix} p_1, & \dots, & p_{n_J}, & 0, & p_{n_J+2}, & \dots, & p_{n_J+\omega_u-1}, & 1\end{pmatrix} G_1^{-1}}_{=:k} x(\cdot), \tag{3.43}$$

where

$$G_1 := \begin{bmatrix} I_{n_J} & 0 \\ 0 & -[b_N, Nb_N, \dots, N^{\omega_u-1}b_N]\end{bmatrix} \in \mathbb{R}^{n\times(n_J+\omega_u)} \text{ is left invertible.}$$

Proof. Let $(x^*, u^*) \in \mathfrak{B}_{[E,A,b]}$ be the optimal trajectory according to Proposition 3.8.1.

Case 1: For $\omega_u = 1$, we have by Proposition 3.8.1 that

$$u^*(\cdot) = -\frac{1}{\widehat{r}}(b_1^\top P + \widehat{h}^\top)x_1^*(\cdot) = -p x_J^*(\cdot) = -kx^*(\cdot).$$

It remains to show that the closed-loop system $[E, A - bk, 0]$ is regular. By Lemma 3.8.3, it suffices to show that for $k = (k_J, k_N) \in \mathbb{R}^{1\times n_J} \times \mathbb{R}^{1\times n_N}$ the condition

$$k_N b_N \neq 1 \tag{3.44}$$

is fulfilled. In our case, by (3.42) we see that

$$k_J = p, \; k_N = 0_{1\times n_N},$$

so (3.44) is fulfilled as $k_N b_N = 0$.

Case 2: For $\omega_u > 1$, we have by Proposition 3.5.4 and Proposition 3.8.1 that

$$\begin{aligned} u^{*(\omega_u-1)}(\cdot) &= x^{*(\omega_u-1)}_{1,n_J+1}(\cdot) = \dot{x}^*_{1,n_J+\omega_u-1}(\cdot) = -\frac{1}{\widehat{r}}(b_1^\top P + \widehat{h}^\top)x_1^*(\cdot) = -p x_1^*(\cdot) \\ &= -\begin{pmatrix} p, & 0 \end{pmatrix} \begin{pmatrix} x_J^*(\cdot) \\ u^*(\cdot) \\ \vdots \\ u^{*(\omega_u-1)}(\cdot) \end{pmatrix} \\ \iff u^*(\cdot) &= -\frac{1}{p_{n_J+1}} \begin{pmatrix} p_1, \dots, p_{n_J}, & 0, & p_{n_J+2}, \dots, p_{n_J+\omega_u-1}, & 1 \end{pmatrix} \begin{pmatrix} x_1^*(\cdot) \\ u^{*(\omega_u-1)}(\cdot) \end{pmatrix} \end{aligned} \tag{3.45}$$

By (3.7) and Proposition 3.5.4, we have

$$x^*(\cdot) = \begin{pmatrix} x_J^*(\cdot) \\ x_N^*(\cdot) \end{pmatrix} \overset{\text{ae}}{=} \begin{bmatrix} I_{n_J} & 0 \\ 0 & -[b_N, \dots, N^{\omega_u-1}b_N] \end{bmatrix} \begin{pmatrix} x_J^*(\cdot) \\ u^*(\cdot) \\ \vdots \\ u^{*(\omega_u-1)}(\cdot) \end{pmatrix} = G_1 \begin{pmatrix} x_1^*(\cdot) \\ u^{*(\omega_u-1)}(\cdot) \end{pmatrix}.$$

As $[b_N, \dots, N^{\omega_u-1}b_N]$ is left invertible according to Lemma 3.6.3, it follows that G_1 is left invertible and

$$\begin{pmatrix} x_1^*(\cdot) \\ u^{*(\omega_u-1)}(\cdot) \end{pmatrix} = G_1^{-1} x^*(\cdot).$$

Substituting this into (3.45) yields the assertion (3.43).

It remains to show that $[E, A - bk, 0]$ is regular. Note that

$$\begin{aligned} k_N b_N &= (k_J, k_N) \begin{pmatrix} 0_{n_J} \\ b_N \end{pmatrix} = k\, G_1 e_{n_J+1} \\ &= -\frac{1}{p_{n_J+1}} \begin{pmatrix} p_1, & \dots, & p_{n_J}, & 0, & p_{n_J+2}, \dots, p_{n_J+\omega_u-1}, & 1 \end{pmatrix} G_1^{-1} G_1 e_{n_J+1} \\ &= -\frac{1}{p_{n_J+1}} \begin{pmatrix} p_1, & \dots, & p_{n_J}, & 0, & p_{n_J+2}, \dots, p_{n_J+\omega_u-1}, & 1 \end{pmatrix} e_{n_J+1} \\ &= 0, \end{aligned}$$

and regularity of $[E, A - bk, 0]$ follows by Lemma 3.8.3. This concludes the proof. □

Remark 3.8.5. While the optimal *control* for the DAE system (3.4) in quasi Weierstraß form is unique if, and only if, the optimal control for the ODE system (3.21) is unique, the optimal *feedback* is generally not unique: for $2 \leq \omega_u < n_N$, any left inverse of G_1 can be chosen to implement the feedback, giving rise to a possible multitude of optimal feedbacks. Roughly speaking, this is due to the fact that the states $u, \dots, u^{(\omega_u-2)}$ of the augmented system necessary to implement the optimal control can be derived from the nominal system in multiple equivalent ways. However if the actual optimal control is unique, all feedbacks will lead to the same control input and closed-loop system. This will be further investigated in Example 3.8.8.

Remark 3.8.6. It is currently not known whether the condition $p_{n_J+1} \neq 0$ in Proposition 3.8.4 is actually restrictive: so far no example that violates this assumption could be found. It is conjectured that the algebraic Riccati equation, given the structure of the system and the performance index, does not allow for solutions with $p_{n_J+1} = 0$, however no formal proof has been established yet.

Example 3.8.7. Consider the nilpotent DAE in quasi Weierstraß form (3.18) with the performance index

$$J(\infty, x, u) = \int_0^\infty \left\| \begin{pmatrix} x_{N,1}(t) \\ x_{N,2}(t) \end{pmatrix} \right\|^2 \mathrm{d}t.$$

The augmented system is given by (3.19), with x_1 fulfilling the differential equation

$$\dot{x}_1(t) = a_1 x_1(t) + \widehat{u}(t), \quad a_1 = 0. \tag{3.46}$$

The performance index of the augmented system is given by

$$\widehat{J}(\infty, x_1, \widehat{u}) = \int_0^\infty \begin{pmatrix} x_1(t) \\ \widehat{u}(t) \end{pmatrix}^\top \begin{bmatrix} \widehat{q} & 0 \\ 0 & \widehat{r} \end{bmatrix} \begin{pmatrix} x_1(t) \\ \widehat{u}(t) \end{pmatrix} \mathrm{d}t, \quad \widehat{q} = 2,\ \widehat{r} = 1.$$

As (3.46) is stabilizable and $\widehat{q} > 0$, we can apply Proposition 3.8.4. We must find the positive solution of the Riccati equation

$$a_1 p_* + p_* a_1 + \widehat{q} - \frac{1}{\widehat{r}} p_*^2 = 0,$$

which is given by $p_* = \sqrt{2}$. Consequently, the optimal control is given by Proposition 3.8.1 as $u^* = x_1^*$, where x_1^* is the solution of the differential equation

$$\dot{x}_1^*(t) = -p_* x_1^*(t) = -\sqrt{2} x_1^*(t), \quad x_1^*(0) = -x_2^0.$$

To implement an optimal state feedback for the nominal DAE (3.4), we need the inverse of

$$G_1 = -\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = G_1^{-1}.$$

The optimal feedback is then given by

$$u(\cdot) = -\frac{1}{p_*} \begin{pmatrix} 0 & 1 \end{pmatrix} G_1^{-1} \begin{pmatrix} x_{N,1}(\cdot) \\ x_{N,2}(\cdot) \end{pmatrix} = \frac{1}{\sqrt{2}} x_{N,1}(\cdot).$$

Example 3.8.8. Consider the DAE

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}1&0&0&0\\0&0&1&0\\0&0&0&1\\0&0&0&0\end{bmatrix}\begin{pmatrix}x_J(t)\\x_{N,1}(t)\\x_{N,2}(t)\\x_{N,3}(t)\end{pmatrix}=\begin{bmatrix}-1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}\begin{pmatrix}x_J(t)\\x_{N,1}(t)\\x_{N,2}(t)\\x_{N,3}(t)\end{pmatrix}+\begin{pmatrix}0\\0\\1\\0\end{pmatrix}u(t) \tag{3.47}$$

with $\omega_u = 2 < n_N = 3$, equipped with the performance index

$$J(\infty, x, u) = \int_0^\infty \|x(t)\|^2 + u(t)^2 \,\mathrm{d}t.$$

Implementing the procedure established in Proposition 3.8.4, we arrive at

$$P = \begin{bmatrix}\frac{1}{2}&0\\0&\sqrt{2}\end{bmatrix}$$

as the solution of the Riccati equation. Using the family of pseudo inverses

$$G_1^{-1} = \begin{bmatrix}1&0&0&0\\0&0&-1&0\\0&-1&0&\alpha\end{bmatrix}, \quad \alpha \in \mathbb{R},$$

we arrive at the familiy of distinct optimal feedbacks

$$k = \frac{1}{\sqrt{2}}\begin{pmatrix}0&-1&0&\alpha\end{pmatrix}, \quad \alpha \in \mathbb{R}.$$

Note that this is to be expected, as review of the DAE (3.47) reveals that $x_{N,3} \overset{\text{ae}}{=} 0$, so all these feedbacks eventually lead to the same optimal control input u^*.

4 Conclusions and outlook

In this thesis, we have investigated the linear-quadratic optimal control problem for time-varying and time-invariant differential-algebraic equations. For the time-varying case, we have established a characterization of the optimal value, provided that the latter is smooth enough. It needs to be investigated further under which conditions this smoothness assumptions holds. Unfortunately, this characterization does not provide a numerically feasible way to obtain the optimal value. Furthermore the actual structure of the optimal control is not yet known.

For the time-invariant, regular case, we have determined differentiability conditions for the control input. Using these results, we could establish an augmented system for which the optimal control problem can be solved very easily, using the methodology known from ordinary differential equations. However, there are still open research questions for this approach: mainly, we are interested in obtaining the optimal control or feedback without invoking the transformation to the quasi Weierstraß form, as this would make numerical calculations far more efficient.

Furthermore, it would be interesting to further generalize the approach presented in Chapter 3. In particular, it should be possible to derive differentiability conditions for multiple input systems, i. e. $m > 1$. Such results are also be expected to be helpful with regards to the definition of the augmented system for singular DAEs, where we expect more "free" inputs due to underdetermined system states on the one hand, as well as further restrictions to the system input resulting from overdetermined parts of the system. Finally, we pose the question whether it is possible to generalize all the previous results to time-varying systems, at least for particular subclasses such as the standard canonical form [CP83].

Bibliography

[AE08] Herbert Amann and Joachim Escher. *Analysis II*. 1st English. Basel Boston Berlin: Birkhäuser, 2008 (cit. on p. 14).

[AFIJ03] H. Abou-Kandil, G. Freiling, Vlad Ionescu and G. Jank. *Matrix Ricati Equations in Control and Systems Theory*. Basel: Birkhäuser, 2003 (cit. on p. 10).

[AM90] Brian D. O. Anderson and John B. Moore. *Optimal Control – Linear Quadratic Methods*. Englewood Cliffs, NJ: Prentice-Hall, 1990 (cit. on p. 10).

[Ber09] Dennis S. Bernstein. *Matrix Mathematics: Theory, Facts, and Formulas*. 2nd. Princeton and Oxford: Princeton University Press, 2009 (cit. on p. 38).

[BIT12] Thomas Berger, Achim Ilchmann and Stephan Trenn. "The quasi-Weierstraß form for regular matrix pencils". *Linear Algebra Appl.* 436.10 (2012), pp. 4052–4069 (cit. on p. 21).

[BL87] D. Bender and A. Laub. "The linear quadratic optimal regulator problem for descriptor systems". *IEEE Trans. Autom. Control* 32 (1987), pp. 672–688 (cit. on p. 10).

[BR13] Thomas Berger and Timo Reis. "Controllability of linear differential-algebraic systems - a survey". In: *Surveys in Differential-Algebraic Equations I*. Ed. by Achim Ilchmann and Timo Reis. Differential-Algebraic Equations Forum. Berlin-Heidelberg: Springer-Verlag, 2013, pp. 1–61 (cit. on pp. 21, 28, 29).

[Bro70] Roger W. Brockett. *Finite Dimensional Linear Systems*. New York: John Wiley and Sons Inc., 1970 (cit. on p. 10).

[CA78] D.J. Clements and B.D.O. Anderson. *Singular Optimal Control - The Linear-Quadratic Problem*. Lecture Notes in Control and Information Sciences 5. Berlin: Springer-Verlag, 1978 (cit. on pp. 10, 15, 33).

[Cam80] Stephen L. Campbell. *Singular Systems of Differential Equations I*. New York: Pitman, 1980 (cit. on p. 9).

[Cam82] Stephen L. Campbell. *Singular Systems of Differential Equations II*. New York: Pitman, 1982 (cit. on p. 9).

[CP83] Stephen L. Campbell and Linda R. Petzold. "Canonical forms and solvable singular systems of differential equations". *SIAM J. Alg. & Disc. Meth.* 4 (1983), pp. 517–521 (cit. on p. 43).

[INS84] Achim Ilchmann, Ines Nürnberger and Wiland Schmale. "Time-varying polynomial matrix systems". *Int. J. Control* 40.2 (1984), pp. 329–362 (cit. on p. 12).

[KM06] Peter Kunkel and Volker Mehrmann. *Differential-Algebraic Equations. Analysis and Numerical Solution*. Zürich, Switzerland: EMS Publishing House, 2006 (cit. on p. 9).

[KM90] Peter Kunkel and Volker Mehrmann. "Numerical solution of Riccati differential algebraic equations". *Linear Algebra Appl.* 137/138 (1990), pp. 39–66 (cit. on p. 10).

[Lew85] Frank L. Lewis. “Preliminary notes on optimal control for singular systems”. In: *Proc. 24th IEEE Conf. Decis. Control, Ft. Lauderdale, FL*. New York: IEEE Press, 1985, pp. 266–272 (cit. on p. 10).

[LMT13] René Lamour, Roswitha März and Caren Tischendorf. *Differential Algebraic Equations: A Projector Based Analysis*. Vol. 1. Differential-Algebraic Equations Forum. Heidelberg-Berlin: Springer-Verlag, 2013 (cit. on pp. 9, 10).

[LR95] Peter Lancaster and Leiba Rodman. *Algebraic Riccati Equations*. Oxford: Clarendon Press, 1995 (cit. on pp. 10, 25, 33, 35, 37).

[PW98] Jan Willem Polderman and Jan C. Willems. *Introduction to Mathematical Systems Theory. A Behavioral Approach*. New York: Springer-Verlag, 1998 (cit. on p. 21).

[RRV15] Timo Reis, Olaf Rendel and Matthias Voigt. “The Kalman-Yakubovich-Popov Inequality for Differential-Algebraic Systems”. *Linear Algebra Appl.* 86 (2015), pp. 153–193 (cit. on p. 10).

[TSH01] Harry L. Trentelman, Anton A. Stoorvogel and Malo L. J. Hautus. *Control Theory for Linear Systems*. Communications and Control Engineering. London: Springer-Verlag, 2001 (cit. on p. 28).

[ZDG96] Kemin Zhou, John C. Doyle and Keith Glover. *Robust and Optimal Control*. Upper Saddle River, NJ: Prentice Hall, 1996 (cit. on p. 10).

Statement of authorship

I hereby certify that this thesis has been composed by me and is based on my own work, except where explicitly stated otherwise. I confirm that this work has not been submitted elsewhere in any other form for the fulfilment of any other degree or qualification.

Place and date

Jonas Witschel