CrossMark

Simulation of multibody systems with servo constraints through optimal control

R. Altmann[1] · **J. Heiland**[2]

Abstract We consider mechanical systems where the dynamics are partially constrained to prescribed trajectories. An example for such a system is a building crane with a load and the requirement that the load moves on a certain path.

Enforcing this condition directly in form of a servo constraint leads to differential-algebraic equations (DAEs) of arbitrarily high index. Typically, the model equations are of index 5, which already poses high regularity conditions. If we relax the servo constraints and consider the system from an optimal control point of view, the strong regularity conditions vanish, and the solution can be obtained by standard techniques.

By means of the well-known n-car example and an overhead crane, the theoretical and expected numerical difficulties of the direct DAE and the alternative modeling approach are illustrated. We show how the formulation of the problem in an optimal control context works and address the solvability of the optimal control system. We discuss that the problematic DAE behavior is still inherent in the optimal control system and show how its evidences depend on the regularization parameters of the optimization.

Keywords Servo constraints · Inverse dynamics · High-index DAEs · Optimal control · Underactuated mechanical systems

1 Introduction

We consider mechanical systems with servo constraints; see, e.g., [7, 8, 19], for which a part of the motion is specified. This includes crane models where one seeks for an input

✉ J. Heiland
heiland@mpi-magdeburg.mpg.de

R. Altmann
raltmann@math.tu-berlin.de

[1] Institut für Mathematik MA4-5, Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany

[2] Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106 Magdeburg, Germany

that makes an end effector follow a prescribed trajectory. Often, these configurations are called *inverse dynamics* problems or, since the number of degrees of freedom exceeds the number of controls, *underactuated mechanical systems*. A recent overview of the diversity of given applications can be found, e.g., in [9, 27]. In this paper, we consider a two-car example [8, Sect. II], its generalization to *n* connected cars, and an overhead crane example [5, Ex. 4]. In all of these examples, the model equations are *differentially flat*, which means that the input can be expressed and determined solely by means of the desired output and its derivatives [12].

In the direct modeling approach, the inputs are regarded as variables, whereas the desired output is formulated as a constraint. This constraint makes the model equations a system of differential-algebraic equations (DAE) even though the dynamics of the system may be given in the form of an ODE. Even more, the resulting DAE is typically of high index so that for its application in simulations, we need to employ a suitable *index reduction* [1, 5, 7]. Apart from the numerical difficulties that come with high index DAEs, an immediate drawback of the DAE approach is that a solution to the problem can only exist if the target trajectory is sufficiently smooth.

As a cure to both mentioned shortcomings of the DAE approach, we consider an optimal control approach that relaxes the constraints and balances the approximation to the target with the control effort. The relaxation of the constraint $Cx = y$ softens the strong regularity assumptions in the DAE setting. In theory, the desired trajectory y may then be even discontinuous.

The same approach was discussed in [26], however, without analyzing the optimality system such that, due to otherwise inconsistent boundary conditions, it becomes necessary to introduce an additional regularization.

Apart from inconsistent boundary conditions, also nonsmooth data may cause problems—not in theory but in the practical application of the optimization approach. In other words, the DAE problematic is not simply gone since the weak coupling of the masses and the noncollocated sensors and activators may lead to oscillations in the output and strong peaks in the (unknown) input. A remedy is the penalization of the derivatives of the control force at the expense of a worse performance and of less standard systems of equations for the numerical realization.

In this paper, we analyze the optimal control approach and investigate methods for the solution of the resulting equation systems. In particular, the following issues are addressed:

1. The very weak coupling of input and output that, in the (DAE) limit, may lead to singular actuations. We will investigate the dependency of the penalty or regularization parameters and the behavior in the limit case.
2. Necessary and sufficient optimality conditions with an emphasis on their use for the solution of the optimization problem. Particularly, in the case of holonomic constraints, the formally derived first-order optimality conditions are preferable over alternative formulations, but they may not be solvable due to inconsistent data.

The paper is organized as follows. In Sect. 2, we give the formulation of the servo-constraint problem as a DAE for which additional holonomic constraints are allowed. The counterpart is then presented in Sect. 3 in which the same configuration is modeled as an optimal control problem. Here we formulate the optimality conditions, discuss the consistency conditions of the boundary data, and prove the existence of an optimal solution. The relation of the two approaches is then the topic of Sect. 4, in particular, the optimal control problem in which the input is not penalized is analyzed. Section 5 gives an overview of the different solution strategies. This includes the DAE case and the optimal control approach for which

Fig. 1 Illustration of the mechanical system from Example 1 including two cars connected by a spring with parameters k and d
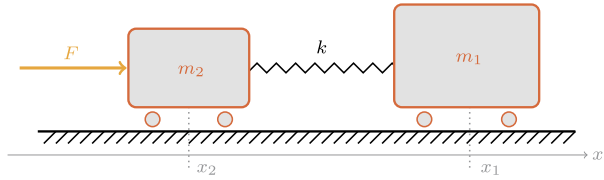
boundary-value problems have to be solved. In Sect. 6, we compare the two approaches by means of two numerical examples. Finally, a conclusion is given in Sect. 7.

2 DAE setting

This section is devoted to the original formulation of the servo-constraint configuration as a high-index DAE. For this, we consider the dynamics of a mechanical system with servo and possibly holonomic constraints. We start with a prototype of a mechanical system with servo constraints.

Example 1 Consider two cars that are connected via a spring (see Fig. 1), as it was described, e.g., in [8, Sect. II]. This mechanical system is modeled with two degrees of freedom, namely the positions x_1, x_2, and one servo constraint. The aim is to make the car on the right follow a desired trajectory given by the function y, i.e., to have that $x_1 = y$. To achieve this, a certain input force F has to be applied. The constrained system has the form

$$m_1 \ddot{x}_1 = -k(x_1 - x_2 - d), \tag{1a}$$

$$m_2 \ddot{x}_2 = k(x_1 - x_2 - d) + F, \tag{1b}$$

$$x_1 = y. \tag{1c}$$

Here, the spring constant k and the spring length d are positive. The given equations of motion, with input F treated as an unknown, form a DAE of (differentiation) index 5. For a definition of the index, we refer to [4, Def. 2.2.2].

Example 1 is a realization of a general second-order system with inputs and servo constraints. In the DAE formulation, where the input is considered as an unknown variable, it reads as follows.

Configuration 1 For a time interval $[0, T]$, initial values $x^0, v^0 \in \mathbb{R}^n$ and a forcing term $f \in \mathcal{C}([0, T]; \mathbb{R}^n)$, for matrices $M \in \mathbb{R}^{n,n}$ symmetric and strictly positive definite and $A \in \mathbb{R}^{n,n}$, for an output operator $C \in \mathbb{R}^{m,n}$ and a desired output y with $y(t) \in \mathbb{R}^m$, and for an input operator $B \in \mathbb{R}^{n,m}$, find an input $u \in \mathcal{C}([0, T]; \mathbb{R}^m)$ and a state trajectory $x \in \mathcal{C}^2([0, T]; \mathbb{R}^n)$ such that

$$M \ddot{x} = Ax + Bu + f, \tag{2a}$$

$$y = Cx, \tag{2b}$$

$$x(0) = x^0, \quad \text{and} \quad \dot{x}(0) = v^0. \tag{2c}$$

Another example of Configuration 1 is given by the generalization of Example 1 to a spring-mass chain of n masses connected by $n-1$ springs.

Example 2 A mass-spring chain, where we want to steer the first mass x_1 along a trajectory y by applying a force u to the last mass x_n, can be modeled as in Configuration 1, where $x = [x_1 \ x_2 \ \cdots \ x_n]^\mathsf{T}$ is the vector of coordinates, $M \in \mathbb{R}^{n,n}$ is the diagonal matrix of the masses, and $A \in \mathbb{R}^{n,n}$, $B \in \mathbb{R}^{n,1}$, $C \in \mathbb{R}^{1,n}$, and $f \in \mathbb{R}^{n,1}$ are given as

$$
A = \begin{bmatrix}
-k_1 & k_1 \\
k_1 & -k_1 - k_2 & k_2 \\
& k_2 & -k_2 - k_3 & k_3 \\
& & \ddots & \ddots & \ddots \\
& & & k_{n-2} & -k_{n-2} - k_{n-1} & k_{n-1} \\
& & & & k_{n-1} & -k_{n-1}
\end{bmatrix}, \qquad
B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix},
$$

$$
C = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}, \quad \text{and} \quad
f = \begin{bmatrix}
k_1 d_1 \\
-k_1 d_1 + k_2 d_2 \\
\vdots \\
-k_{n-2} d_{n-2} + k_{n-1} d_{n-1} \\
-k_{n-1} d_{n-1}
\end{bmatrix}
\tag{3}
$$

for given spring constants $k_1, \dots, k_{n-1} > 0$ and spring lengths $d_1, \dots, d_{n-1} > 0$. This system then defines a DAE of index $2n + 1$.

In order to model general multibody systems, such as the crane model in [5, Ex. 4], we need to include holonomic constraints of type $g(x) = 0$, where g is a suitable possibly nonlinear function. Therefore, we extend Configuration 1 as follows.

Configuration 2 For a time interval $[0, T]$, initial values $x^0, v^0 \in \mathbb{R}^n$ and a forcing term $f \in \mathcal{C}([0, T]; \mathbb{R}^n)$, for $g \colon \mathbb{R}^n \to \mathbb{R}^r$, for $M \in \mathbb{R}^{n,n}$ symmetric and strictly positive definite and $A \in \mathbb{R}^{n,n}$, for an output operator $C \in \mathbb{R}^{m,n}$ and a desired output y with $y(t) \in \mathbb{R}^m$, and for an input operator $B \in \mathbb{R}^{n,m}$, find an input $u \in \mathcal{C}([0, T]; \mathbb{R}^m)$, a state trajectory $x \in \mathcal{C}^2([0, T]; \mathbb{R}^n)$, and a Lagrange multiplier $p \in \mathcal{C}([0, T]; \mathbb{R}^r)$ such that

$$
M\ddot{x} = Ax + G^\mathsf{T}(x)p + Bu + f, \tag{4a}
$$

$$
0 = g(x), \tag{4b}
$$

$$
y = Cx, \tag{4c}
$$

$$
x(0) = x^0, \quad \text{and} \quad \dot{x}(0) = v^0. \tag{4d}
$$

Here, the Lagrange multiplier p couples the holonomic constraint (4b) to the dynamical equations (4a), and G is the Jacobian of the holonomic constraint, i.e., $G(x) = \frac{\partial g}{\partial x}(x) \in \mathbb{R}^{r,n}$.

The following observations and assumptions will be formulated for Configuration 2, but they directly apply to Configuration 1 if we omit g and $G^\mathsf{T}(x)p$. From DAE theory it is known that the derivatives of the right-hand side of system (4a)–(4d) appear in the solution [4, Ch. 2]. Because of the assumed semiexplicit structure of the system, only the derivatives of g and y are part of the solution and, in particular, of the desired input u. We can show that

in the index-5 case of Example 1, the input depends on the 4th derivative of y, whereas for the n-spring-mass chain of Example 2, it depends on $y^{(2n)}$. Thus, the following assumptions are made to guarantee a continuous solution u of Configuration 2 or 1.

Assumption 1 (*DAE setting*) *In the formulation of Configurations 2 and 1, it is assumed that*:

1. *Smoothness of the data*: $f \in \mathcal{C}([0, T]; \mathbb{R}^n)$ *and* $y \in \mathcal{C}^{\nu_d - 1}([0, T]; \mathbb{R}^m)$, *where* ν_d *is the (differentiation) index of the system equations.*
2. *Consistency of the initial values with respect to the holonomic constraints*: $g(x^0) = 0$ *and* $G(x^0)v^0 = 0$, *if applicable.*
3. *Consistency of the initial values with respect to the target output*: $Cx^0 = y(0)$ *and* $Cv^0 = \dot{y}(0)$, *and also the conditions that result from the insertion of the differential equation into the servo constraint,*

$$\ddot{y} = C\ddot{x} = CM^{-1}\left(Ax + G^{\mathsf{T}}(x)p + Bu + f\right).$$

Remark 1 In the n-car example from Example 2, the consistency conditions, which directly follow from equation (2b), are $y(0) = x_1^0$ and $\dot{y}(0) = v_1^0$. Furthermore, through the combination of Eqs. (2a) and (2b), we obtain the two conditions

$$m_1\ddot{y}(0) = k_1\left(-x_1^0 + x_2^0 + d_1\right), \qquad m_1 y^{(3)}(0) = k_1\left(-v_1^0 + v_2^0\right).$$

We emphasize that the numerical integration of high-index DAEs involves many difficulties; see, e.g., [21, Ch. II]. Furthermore, the high-index property and the resulting assumptions hypothesize that the DAE formulation (4a)–(4d) does not provide an appropriate model. Thus, we propose a remodeling process that leads to an optimal control problem. This involves a modeling error that is adjustable by a parameter as discussed in the following sections.

Remark 2 From the observation that the solution of an index-ν_d system depends on the ν_dth derivative of the right-hand side it follows that the index is closely related to the so-called *relative degree*; see [18] for a definition. Indeed, for the examples considered in this paper, the rule of thumb that *the relative degree is equal to the index minus 1* applies. However, since the same input-to-output behavior can be realized through DAEs of different indices, this rule does not hold in general.

3 Formulation as optimal control problem

Instead of prescribing the servo constraint (4c) in a rigid way, we can formulate it as the target of an optimization problem. Accordingly, the solution does not follow the trajectory exactly, which allows for a less regular target, and which typically leads to smaller input forces.

Configuration 3 For a $\nu \in \mathbb{N}$, find $u \in \mathcal{C}^\nu([0, T]; \mathbb{R}^m)$ that minimizes the cost functional

$$\mathcal{J}(x, u) := \mathcal{S}\left(x(T)\right) + \int_0^T \mathcal{Q}(x) + \mathcal{R}(u)\, \mathrm{d}t \tag{5}$$

with the quadratic performance criteria

$$\mathcal{Q}(x) := \frac{1}{2}(Cx - y)^\mathsf{T} Q(Cx - y), \qquad \mathcal{R}(u) := \frac{1}{2}\sum_{i=0}^{\nu} u^{(i)\mathsf{T}} R_i u^{(i)}, \tag{6a}$$

$$\text{and} \quad \mathcal{S}\big(x(T)\big) := \frac{1}{2}\big(Cx(T) - y(T)\big)^\mathsf{T} S\big(Cx(T) - y(T)\big) \tag{6b}$$

for given $Q \in \mathbb{R}^{m,m}$, $S \in \mathbb{R}^{m,m}$, and $R_0, \ldots, R_\nu \in \mathbb{R}^{m,m}$ symmetric and positive semidefinite. The state $x = x(u) \in \mathcal{C}^2([0, T]; \mathbb{R}^n)$ is related to u through the dynamics (4a), the holonomic constraint (4b), and the initial conditions (4d).

The parameters Q, S, and $R_0, \ldots, R_\nu$ can be chosen to meet certain requirements for the minimization. With this, we may install different kinds of penalizations of the derivatives of this input variable u. Note that $R_0, \ldots, R_\nu$ describe the modeling error compared to the DAE formulation in Configuration 2; see also the discussion in Sect. 4.

Note that in Configuration 3 the fulfillment of the constraint (4c) is balanced with the cost of the input u, including its derivatives up to order ν. Also, the necessary smoothness conditions on y for a continuous solution u and the consistency condition of the initial values with respect to the target output (see Assumption 1) can be relaxed.

Assumption 2 (*Optimal control setting*) *In the formulation of Configuration 3 it is assumed that*:

1. *Smoothness of the data*: *f and y are continuous on $[0, T]$.*
2. *Consistency of the initial values with respect to the holonomic constraints*: *$g(x^0) = 0$ and $G(x^0)v^0 = 0$, if applicable.*

3.1 Optimality conditions

We derive *formal* optimality conditions for the optimization problem in Configuration 3; cf. [25, Ch. 6] and [22] for the DAE case with holonomic constraints. In order to apply the standard variational approach, we have to make the following assumption.

Assumption 3 *For any input u, the state equations (4a)–(4b) with (4d) have a unique solution $x = x(u)$, and the map $u \mapsto x(u)$ that assigns an input u to the corresponding solution is differentiable.*

For the considered setups and in particular for the linear problem from Example 2, this assumption is readily confirmed. To derive the formal optimality conditions, we consider the Lagrange functional

$$\mathcal{L}(u; \lambda, \mu) = \mathcal{J}\big(x(u), u\big) + \int_0^T \lambda^\mathsf{T}\big(M\ddot{x} - Ax - G^\mathsf{T}(x)p - Bu - f\big) + \mu^\mathsf{T} g(x)\, dt \tag{7}$$

for formally introduced multipliers $\lambda \in \mathcal{C}^2([0, T]; \mathbb{R}^n)$ and $\mu \in \mathcal{C}([0, T]; \mathbb{R}^r)$.

The formal optimality conditions are derived from the requirement that, for suitable (λ, μ), an optimal u^* marks a stationary point of $\mathcal{L}(u; \lambda, \mu)$, i.e.,

$$\frac{\partial}{\partial u}\mathcal{L}\big(u^*; \lambda, \mu\big)\delta_u = 0 \tag{8}$$

for every variation δ_u; cf. [29, Ch. 14]. This is the case if (x, p, u, λ, μ) solves the following formal optimality system.

Problem 1 Consider the functions and coefficients defined in Configurations 2 and 3. Find $(x, p), (\lambda, \mu) \in \mathcal{C}^2([0, T]; \mathbb{R}^n) \times \mathcal{C}([0, T]; \mathbb{R}^r)$, and $u \in \mathcal{C}([0, T]; \mathbb{R}^m)$ such that

$$M\ddot{x} = Ax + G^{\mathsf{T}}(x)p + Bu + f, \tag{9a}$$

$$0 = g(x), \tag{9b}$$

$$M^{\mathsf{T}}\ddot{\lambda} = A^{\mathsf{T}}\lambda + \frac{\partial}{\partial x}\big(G(x)^{\mathsf{T}}p\big)\lambda - G^{\mathsf{T}}(x)\mu - C^{\mathsf{T}}QCx + C^{\mathsf{T}}Qy, \tag{9c}$$

$$0 = G(x)\lambda, \tag{9d}$$

$$0 = \sum_{i=0}^{\nu}(-1)^i R_i u^{(2i)} - B^{\mathsf{T}}\lambda \tag{9e}$$

with the initial conditions for x as in Eq. (4d) and the terminal conditions for the dual variable λ given by

$$M^{\mathsf{T}}\lambda(T) = 0, \quad M^{\mathsf{T}}\dot{\lambda}(T) = C^{\mathsf{T}}S\big(Cx(T) - y(T)\big). \tag{10}$$

Depending on the parameter ν, we obtain the boundary conditions

$$\sum_{i=1}^{\nu}\sum_{k=1}^{i}(-1)^k\,\delta_u^{(i-k)\mathsf{T}}\,R_i u^{(i+k-1)}|_0^T = 0 \tag{11}$$

for all variations δ_u from a suitable subset of $\mathcal{C}([0, T]; \mathbb{R}^m)$.

Remark 3 For the case $\nu = 0$, by (9e) we obtain the often used algebraic relation

$$0 = -B^{\mathsf{T}}\lambda + R_0 u$$

and no boundary conditions for the input u.

In general, for $\nu > 0$, the space of suitable variations δ_u depends on the incorporation of the inputs in the optimality problem. Either, we may add the derivatives of the input to the cost functional as in (5) without any specifications of initial conditions or incorporate the inputs $\dot{u}, \ldots, u^{(\nu-1)}$ as part of the state vector as in [25, Rem. 3.8]. In the latter case, initial conditions have to be stated for $u(0), \ldots, u^{(\nu-1)}(0)$, which may be unphysical. Furthermore, the corresponding variations $\delta_u, \dot{\delta}_u, \ddot{\delta}_u, \ldots, \delta_u^{(\nu-1)}$ have to vanish at $t = 0$.

Remark 4 If we do not impose restrictions on admissible inputs u and, thus, on δ_u, then for fixed ν and positive definite R_ν, the following conditions need to hold:

$$\nu = 1: \quad \dot{u}(0) = \dot{u}(T) = 0,$$

$$\nu = 2: \quad \ddot{u}(0) = \ddot{u}(T) = 0,$$

$$R_1\dot{u}(0) - R_2 u^{(3)}(0) = R_1\dot{u}(T) - R_2 u^{(3)}(T) = 0.$$

3.2 First-order formulation

In order to apply standard theory and standard numerical routines, we reformulate the optimality system as a first-order system. For this, we restrict ourselves to the unconstrained case $r = 0$ with $\nu = 1$. We introduce the variables

$$z = \begin{bmatrix} x \\ \dot{x} \end{bmatrix}, \qquad \zeta := \begin{bmatrix} -\lambda \\ \lambda \end{bmatrix}, \qquad v := \begin{bmatrix} u \\ -\dot{u} \end{bmatrix}.$$

The DAE (2a)–(2c) in first-order form is given by

$$\tilde{M}\dot{z} = \tilde{A}z + \tilde{B}u + \tilde{f}, \tag{12a}$$

$$\tilde{C}z = g, \tag{12b}$$

$$z(0) = \begin{bmatrix} x^0 \\ v^0 \end{bmatrix} \tag{12c}$$

with

$$\tilde{M} := \begin{bmatrix} I_n & 0 \\ 0 & M \end{bmatrix}, \qquad \tilde{A} := \begin{bmatrix} 0 & I_n \\ A & 0 \end{bmatrix}, \qquad \tilde{B} := \begin{bmatrix} 0 \\ B \end{bmatrix},$$

$$\tilde{C} := \begin{bmatrix} C & 0 \end{bmatrix}, \qquad \tilde{f} := \begin{bmatrix} 0 \\ f \end{bmatrix},$$

where I_n denotes the identity matrix in $\mathbb{R}^n$. In order to write the resulting optimality system (9a)–(9e) as a first-order system, we further introduce the matrices

$$\hat{M}^\mathsf{T} := \begin{bmatrix} M^\mathsf{T} & 0 \\ 0 & I_n \end{bmatrix}, \qquad \tilde{J} := \begin{bmatrix} 0 & I_m \\ -I_m & 0 \end{bmatrix},$$

$$\tilde{I}_\beta := \begin{bmatrix} -\beta_0 I_m & 0 \\ 0 & \beta_1 I_m \end{bmatrix}, \quad \text{and} \quad \hat{B} := \begin{bmatrix} \tilde{B} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ B & 0 \end{bmatrix}.$$

The optimality system with the Lagrange multiplier μ and control v is given via

$$\begin{bmatrix} 0 & \tilde{M} & \\ -\hat{M}^\mathsf{T} & 0 & \\ & & \beta_1 \tilde{J} \end{bmatrix} \begin{bmatrix} \dot{\zeta} \\ \dot{z} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} 0 & \tilde{A} & \hat{B} \\ \tilde{A}^\mathsf{T} & -\tilde{C}^\mathsf{T}\tilde{C} & \\ \hat{B}^\mathsf{T} & & \tilde{I}_\beta \end{bmatrix} \begin{bmatrix} \zeta \\ z \\ v \end{bmatrix} + \begin{bmatrix} \tilde{f} \\ \tilde{C}^\mathsf{T}g \\ 0 \end{bmatrix} \tag{13}$$

with the initial and terminal conditions

$$z(0) = \begin{bmatrix} x^0 \\ v^0 \end{bmatrix}, \qquad -\hat{M}^\mathsf{T}\zeta(T) = \gamma \tilde{C}^\mathsf{T}\big(\tilde{C}x(T) - g(T)\big), \qquad \dot{u}(0) = 0, \qquad \dot{u}(T) = 0. \tag{14}$$

In the case where an initial condition for the input was prescribed, i.e., $u(0) = u^0$ is given, the conditions for u in (14) reduce to $\dot{u}(T) = 0$. With appropriate matrices $\mathcal{B}_0$ and $\mathcal{B}_T$ and a vector $\rho \in \mathbb{R}^{4n+2m}$, these boundary conditions can also be written in the form

$$\mathcal{B}_0 \begin{bmatrix} \zeta(0) \\ z(0) \\ v(0) \end{bmatrix} + \mathcal{B}_T \begin{bmatrix} \zeta(T) \\ z(T) \\ v(T) \end{bmatrix} = \rho.$$

This formulation is used in Sect. 5.2.2.

3.3 Necessary conditions for the existence of an optimal solution

In this subsection, the particular aspect of *consistency* as it arises in the context of optimal control of DAEs is discussed. Also, a mathematical result that can be used two overcome this *consistency* issue in the case of optimal control of multibody systems with possibly nonlinear holonomic constraints is presented.

If the optimality system (9a)–(11) has a solution, then it provides necessary optimality conditions for $(x(u), u)$. However, in the considered DAE context, i.e., when holonomic constraints are applied, it may happen that the optimization problem has a solution whereas the *formal* optimality system is not solvable [23]. Apart from the general case that the boundary values do not permit a solution [2], for a DAE, a solution may not exist because of insufficient smoothness of the data or because of inconsistent initial or terminal values.

Thus, it is an important task to establish necessary conditions for solvability of the formal optimality system. By Assumption 2 the initial conditions for x are consistent. The adjoint equations (9b) and (9d) have the same differential-algebraic structure, so that from (9d) we can read off the consistency conditions for the terminal values for λ, namely

$$G\big(x(T)\big)\lambda(T) = 0 \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}t}\big(G\big(x(T)\big)\big)\lambda(T) + G\big(x(T)\big)\dot{\lambda}(T) = 0; \tag{15}$$

cf. Assumption 2(2). Comparing the prescribed terminal conditions (10) for λ to (15), we obtain the necessary and sufficient condition for consistency as

$$0 = G\big(x(T)\big)M^{-\mathsf{T}}C^{\mathsf{T}}S\big(Cx(T) - y(T)\big). \tag{16}$$

Similar conditions in a slightly different formulation have been reported in [26], where the authors proposed the variants to remove the end point penalization from the cost functional or to consider a regularization of the dynamical equation. Within this regularization, the constraint (4b) is replaced by its derivative.

The following theorem shows that, instead of the state equations, we can modify the cost functional. This modification ensures consistency while affecting neither the performance criterion nor the necessity of the formal optimal conditions.

Theorem 1 (*Ensuring consistency*) *Let $P_{x^*(T)}$ be a projector onto the kernel of $G(x^*(T))$ that satisfies $M^{-\mathsf{T}}P_{x^*(T)}^{\mathsf{T}} = P_{x^*(T)}M^{-\mathsf{T}}$. Then, replacing the terminal conditions (10) for λ by the conditions*

$$M^{\mathsf{T}}\lambda(T) = 0, \qquad M^{\mathsf{T}}\dot{\lambda}(T) = P_{x^*(T)}C^{\mathsf{T}}S\big(Cx^*(T) - y(T)\big) \tag{17}$$

ensures consistency of the terminal conditions for λ. Moreover, if $(x^*, p^*, u^*, \lambda, \mu)$ solve the optimality system with (17), then u^* is a stationary point of (7).

Proof Let u^* be a solution to the optimality system and consider the first variation $\frac{\partial}{\partial u}\mathcal{L}(u^*; \lambda, \mu)$. The relation that defines the terminal condition for $\dot{\lambda}^{\mathsf{T}}$ is given by

$$0 = \frac{\partial}{\partial u}\big(\mathcal{S}\big(x^*(T)\big)\big)\delta_u(T) - \dot{\lambda}^{\mathsf{T}}(T)\delta_{x(\delta_u)}(T)$$

$$= \frac{\partial}{\partial x}\big(S\big(x^*(T)\big)\big)\delta_{x(\delta_u)}(T) - \dot{\lambda}^{\mathsf{T}}(T)\delta_{x(\delta_u)}(T), \tag{18}$$

where $\delta_{x(\delta_u)} = \frac{\partial}{\partial u}(x^*(u))\delta_u$ is the variation in x^* induced by the variation of u. Since $\delta_{x(\delta_u)}$ solves the state equations (4a)–(4b) linearized about x^* with input δ_u (cf. [28, Ch. 2]), we have that $\delta_{x(\delta_u)}(T)$ fulfills the linearized constraint (4b), i.e., $G(x^*(T))\delta_{x(\delta_u)}(T) = 0$ or, equivalently, $\delta_{x(\delta_u)}(T) = P_{x^*(T)}\delta_{x(\delta_u)}(T)$. Accordingly, relation (18) does not change if we replace $\frac{\partial}{\partial x}(S(x^*(T)))$ by $\frac{\partial}{\partial x}(S(x^*(T)))P_{x^*(T)}$. For the considered quadratic cost functional (5), this means that the formal conditions (16) are equivalent (in the sense that the first variation of $\mathcal{L}$ is not affected) to

$$\dot{\lambda}(T) = P_{x^*(T)}\frac{\partial}{\partial x}\big(\mathcal{S}\big(x^*(T)\big)\big)^{\mathsf{T}} = P_{x^*(T)}C^{\mathsf{T}}S\big(Cx^*(T) - y(T)\big),$$

which concludes the proof.   □

Remark 5 In the general case, $P_{x^*(T)}$ is defined implicitly since it depends on the unknown solution x^*. In the case of linear holonomic constraints, $P_{x^*(T)}$ is readily computed; see [15, Rem. 8.20]. As in the example presented further, in order to ensure consistency of the terminal conditions, we may also use a projection onto a subspace of $G(x(T))$ that is possibly independent of x. This, however, will effectively alter the performance criterion $\mathcal{S}$.

Remark 6 If M is symmetric, then the condition $M^{-\mathsf{T}}P_{x^*(T)}^{\mathsf{T}} = P_{x^*(T)}M^{-\mathsf{T}}$ is the orthogonality condition in the inner product induced by M^{-1}, which is the natural inner product in PDE applications.

3.4 Existence of optimal solutions

For Configuration 3 constrained by linear equations without holonomic constraints as in (2a), the existence of solutions is provided by well-known results.

Lemma 1 (*Existence of an optimal solution*) *For $\nu \geq 0$, consider the optimal control problem with cost functional (5) constrained by (2a) and let Assumption 2 hold. If $R_\nu > 0$ and if $u(0), \dot{u}(0), \ldots, u^{(\nu-1)}(0)$ are given, then system (9a)–(9e) and the optimal control problem have a unique solution for any $T < \infty$ and initial data x^0 and v^0.*

Proof Recall that, by the standard order reduction approach, the second-order system (9a)–(9e) can be reformulated as an equivalent first-order system; see Sect. 3.2.

Then, for $\nu = 0$, the result is given in [25, Rem. 3.6]. For $\nu = 1$ with $R_1 > 0$, we may introduce a new variable for the derivative of the control u. Interpreting u as a part of the state variable whereas its derivative $v := \dot{u}$ is the new control variable, the same arguments apply; cf. [25, Rem. 3.8]. Note that this ansatz requires an initial value for u. This procedure may be successively repeated for $\nu > 1$.   □

Remark 7 Note that the existence result in Lemma 1 is true for all initial values x^0 and v^0 in contrast to the DAE (4a)–(4d), which requires consistent initial data. The case $\nu = 0$ with $R_0 = 0$ yields again the DAE formulation of the problem and thus, needs consistent boundary conditions. This case is discussed in Sect. 4.

For the nonlinear optimality system (9a)–(9e) with holonomic constraints, we use the strong but reasonable assumption that the state equations (4a)–(4d) have a solution for any input u under the consideration that the solutions of the state equations depend smoothly on the input (Assumption 3) to state that existence of a solution to the formal optimality system is indeed a necessary condition for optimality.

We first show that the adjoint equations have a solution for every state trajectory and, thus, also at the optimal solution. Then we confer that the smoothness of the input to state map implies that, at an optimal solution, the gradient condition (9e) must also be fulfilled.

Lemma 2 *Consider a solution (x, p) of (4a)–(4d). If g is sufficiently smooth and $G(x(t))$ has full row rank for all $t \in [0, T]$ and if the end condition*

$$M^\mathsf{T}\lambda(T) = 0, \qquad M^\mathsf{T}\dot{\lambda}(T) = C^\mathsf{T}S\big(Cx(T) - y(T)\big) \tag{19}$$

is consistent, then the adjoint equations (9c)–(9d) with end condition (19) have a unique solution.

Proof We rewrite the adjoint equations as the first-order system (cf. the second equation of (13))

$$\begin{bmatrix} M^\mathsf{T} & 0 \\ 0 & M^\mathsf{T} \end{bmatrix} \frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} \lambda \\ \dot{\lambda} \end{bmatrix} = \begin{bmatrix} 0 & M^\mathsf{T} \\ \tilde{A}^\mathsf{T} & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ \dot{\lambda} \end{bmatrix} + \begin{bmatrix} 0 \\ G^\mathsf{T} \end{bmatrix} \mu + \begin{bmatrix} 0 \\ \tilde{f} \end{bmatrix}, \tag{20a}$$

$$G\lambda = 0, \tag{20b}$$

where μ is the multiplier that accounts for the holonomic constraint in (7). Therein, the dependencies on x and t were omitted, and all linear coefficients and inhomogeneities in $\tilde{A}$ and $\tilde{f}$ were clustered. Then, via another multiplier η, the differentiated constraint $\dot{G}\lambda + G\dot{\lambda} = 0$ is added to the system and yields

$$\begin{bmatrix} M^\mathsf{T} & 0 \\ 0 & M^\mathsf{T} \end{bmatrix} \frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} \lambda \\ \dot{\lambda} \end{bmatrix} = \begin{bmatrix} 0 & M^\mathsf{T} \\ \tilde{A}^\mathsf{T} & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ \dot{\lambda} \end{bmatrix} + \begin{bmatrix} G^\mathsf{T} & \dot{G}^\mathsf{T} \\ 0 & G^\mathsf{T} \end{bmatrix} \begin{bmatrix} \eta \\ \mu \end{bmatrix} + \begin{bmatrix} 0 \\ \tilde{f} \end{bmatrix}, \tag{21a}$$

$$\begin{bmatrix} G & 0 \\ \dot{G} & G \end{bmatrix} \begin{bmatrix} \lambda \\ \dot{\lambda} \end{bmatrix} = 0. \tag{21b}$$

Since G has pointwise full row rank and since the terminal conditions are assumed to be consistent, by [15, Thm. 8.6] we can state that system (21a)–(21b) has a unique solution. We can show that the parts (λ, μ) of a solution (λ, η, μ) to (21a)–(21b) also solve (20a)–(20b). The other way round, by construction and by the smoothness assumption on G, a solution (λ, μ) to (20a)–(20b) partially defines a solution to (21a)–(21b) and, thus, is unique. $\qquad \square$

Theorem 2 *Assume that $u \mapsto x$ is Lipschitz continuous. If for $(x(u_0), u_0)$, the constraints and the cost functional are Gâteaux differentiable with respect to x at $x(u_0)$ and if the terminal conditions (10) are consistent, then the optimality system (9a)–(9e) is a necessary condition for optimality of $(x(u_0), u_0)$.*

Proof At every candidate solution $x(u_0)$, the adjoint equations (9c) and (9d) with (10) are solvable by Lemma 2. Then, the claim follows from the result given in [15, Thm. 5.5]. □

Concerning sufficiency for the existence of unique global or local solutions, general results for constrained optimization extended to optimal control problems can be consulted; see, e.g., [15, Ch. 5.3].

3.5 Various optimality systems

In the remaining part of this section, we consider particular cases of the optimality system (9a)–(9e) for constrained and nonconstrained systems and different values of v. For this, consider again the cost functional (5) and set

$$Q = I_m, \qquad R_i = \beta_i I_m, \qquad S = \gamma I_m.$$

Thus, the remaining parameters for the optimization problem are $\gamma, \beta_0, \ldots, \beta_v \geq 0$. For different values of v, the various structures of the optimality system are analyzed. Note that this subsection is restricted to $\beta_v > 0$. The particular case $v = 0$ with $\beta_0 = 0$, i.e., with no constraints on the input at all, is then discussed in Sect. 4.

3.5.1 Case $r = 0$, $v = 0$

Consider the case of Configuration 3 with $r = 0$, i.e., the optimization is constrained by an ODE instead of a DAE. As mentioned before, it is assumed here that $\beta_0 > 0$. This then leads to the optimality system

$$\begin{bmatrix} 0 & M & \\ M^\mathsf{T} & 0 & \\ & & 0 \end{bmatrix} \begin{bmatrix} \ddot{\lambda} \\ \ddot{x} \\ \ddot{u} \end{bmatrix} = \begin{bmatrix} 0 & A & B \\ A^\mathsf{T} & -C^\mathsf{T}C & \\ -B^\mathsf{T} & & \beta_0 I_m \end{bmatrix} \begin{bmatrix} \lambda \\ x \\ u \end{bmatrix} + \begin{bmatrix} f \\ C^\mathsf{T}y \\ 0 \end{bmatrix}.$$

Thus, we obtain a DAE of index 1 but with initial and terminal conditions of the form

$$x(0) = x^0, \qquad \dot{x}(0) = v^0, \qquad \lambda(T) = 0, \qquad M^\mathsf{T}\dot{\lambda}(T) = \gamma C^\mathsf{T}\big(Cx(T) - y(T)\big).$$

3.5.2 Case $r = 0$, $v = 1$

For the same case but with the inclusion of a penalization of the first derivative of the input u with parameter $\beta_1 \neq 0$, the optimality system (9a)–(9e) reads

$$\begin{bmatrix} 0 & M & \\ M^\mathsf{T} & 0 & \\ & & \beta_1 I_m \end{bmatrix} \begin{bmatrix} \ddot{\lambda} \\ \ddot{x} \\ \ddot{u} \end{bmatrix} = \begin{bmatrix} 0 & A & B \\ A^\mathsf{T} & -C^\mathsf{T}C & \\ -B^\mathsf{T} & & \beta_0 I_m \end{bmatrix} \begin{bmatrix} \lambda \\ x \\ u \end{bmatrix} + \begin{bmatrix} f \\ C^\mathsf{T}y \\ 0 \end{bmatrix}. \qquad (22)$$

In contrast to the case with $v = 0$, the leading matrix on the left-hand side is invertible. Thus, the optimality system (22) is an ODE. The corresponding boundary conditions are given by

$$x(0) = x^0, \qquad \dot{x}(0) = v^0, \qquad \dot{u}(0) = 0, \qquad \dot{u}(T) = 0,$$
$$\lambda(T) = 0, \qquad M^\mathsf{T}\dot{\lambda}(T) = \gamma C^\mathsf{T}\big(Cx(T) - y(T)\big).$$

3.5.3 Case $r = 0$, $v = 2$

In the case $v = 2$, the second derivatives of the inputs are also penalized, i.e., $\beta_2 \neq 0$. As seen in Eq. (9e), the fourth derivative of u appears in the optimality system. In order to write the system in a second-order form, we introduce a new variable $v := \ddot{u}$. The optimality system then has the form

$$
\begin{bmatrix}
0 & M & & \\
M^\mathsf{T} & 0 & & \\
& & 0 & \beta_2 I_m \\
& & I_m & 0
\end{bmatrix}
\begin{bmatrix}
\ddot{\lambda} \\
\ddot{x} \\
\ddot{u} \\
\ddot{v}
\end{bmatrix}
=
\begin{bmatrix}
0 & A & B & \\
A^\mathsf{T} & -C^\mathsf{T}C & & \\
B^\mathsf{T} & & -\beta_0 I_m & \beta_1 I_m \\
& & 0 & I_m
\end{bmatrix}
\begin{bmatrix}
\lambda \\
x \\
u \\
v
\end{bmatrix}
+
\begin{bmatrix}
f \\
C^\mathsf{T}y \\
0 \\
0
\end{bmatrix}. \quad (23)
$$

This is again an ODE, and the corresponding boundary values read:

$$
\begin{aligned}
& x(0) = x^0, \qquad \dot{x}(0) = v^0, \qquad \ddot{u}(0) = 0, \qquad \ddot{u}(T) = 0, \\
& \beta_1 \dot{u}(0) = \beta_2 \dot{v}(0), \qquad \beta_1 \dot{u}(T) = \beta_2 \dot{v}(T), \\
& \lambda(T) = 0, \qquad M^\mathsf{T}\dot{\lambda}(T) = \gamma C^\mathsf{T}\big(Cx(T) - y(T)\big).
\end{aligned}
$$

3.5.4 Case $r > 0$, $v = 0$

Finally, we give an example of the optimality system if the cost functional is constrained by a DAE. Here, no derivatives of the inputs are penalized, i.e., $v = 0$ and $\beta_0 \neq 0$. Assuming that the Jacobian G is constant, i.e., Eq. (9b) is of the form $0 = g(x) = Gx$, we obtain the optimality system

$$
\begin{bmatrix}
0 & M & & & \\
M^\mathsf{T} & 0 & & & \\
& & 0 & & \\
& & & 0 & \\
& & & & 0
\end{bmatrix}
\begin{bmatrix}
\ddot{\lambda} \\
\ddot{x} \\
\ddot{p} \\
-\ddot{\mu} \\
-\ddot{u}
\end{bmatrix}
$$

$$
=
\begin{bmatrix}
0 & A & G^\mathsf{T} & 0 & -B \\
A^\mathsf{T} & -C^\mathsf{T}C & 0 & G^\mathsf{T} & 0 \\
G & 0 & 0 & & \\
0 & G & & 0 & \\
-B^\mathsf{T} & 0 & & & -\beta_0 I_m
\end{bmatrix}
\begin{bmatrix}
\lambda \\
x \\
p \\
-\mu \\
-u
\end{bmatrix}
+
\begin{bmatrix}
f \\
C^\mathsf{T}y \\
0 \\
0 \\
0
\end{bmatrix}.
$$

In contrast to the previous cases, this is a DAE of index 3, but again with a mixture of initial and terminal conditions. This can be seen by the particular structure of the system, which is the same as that for constrained multibody systems; cf. [17, Ch. VII.1]. This is no surprise since only the "index-5 constraint" was removed by the cost functional.

4 Comparison of DAE and optimal control solutions

To discuss the qualitative behavior of the solutions of the optimal control problem, we consider the linear case without holonomic constraints (Configuration 1) and, in particular, discuss the n-element mass-spring chains as in Example 2.

In the optimal control setting of Sect. 3, it is reasonable to assume that R_v is positive definite. In the sequel, we analyze the limit case with $v = 0$ and $R_0 = 0$, i.e., the case in which the control is not constrained at all.

4.1 Equivalence for $R_0 = 0$

We show that for Example 2 with $\nu = 0$ and $R_0 = 0$, the DAE approach of Configuration 1 is equivalent to the optimal control formulation in Configuration 3, provided that $Q > 0$. Note that this implies that the corresponding optimality system is only solvable for $y \in \mathcal{C}^{2n}([0, T]; \mathbb{R})$. Recall that n denotes the number of coupled cars.

It is easy to see that a solution (x, u) of the original DAE (2a)–(2c) minimizes the cost functional for $R_0 = 0$. For a solution x of the DAE, we have $Cx = y$ such that the cost functional $\mathcal{J}$ from (5) is minimized since

$$\mathcal{J}(x, u) = \mathcal{S}\big(x(T)\big) + \frac{1}{2} \int_0^T (Cx - y)^\mathsf{T} Q(Cx - y) \, \mathrm{d}t = 0.$$

Let us consider the optimality system for the case $R_0 = 0$. Equation (9e) reduces to $0 = B^\mathsf{T} \lambda$, which directly implies that the last component of λ vanishes, i.e., $\lambda_n = 0$. As a result, Eq. (9c) has the form

$$\begin{bmatrix} \ddot{\lambda}_1 \\ \ddot{\lambda}_2 \\ \vdots \\ \ddot{\lambda}_{n-1} \\ 0 \end{bmatrix} = \begin{bmatrix} -k_1 & k_1 & & & \\ k_1 & -k_1 - k_2 & k_2 & & \\ & \ddots & \ddots & \ddots & \\ & & k_{n-2} & -k_{n-2} - k_{n-1} & k_{n-1} \\ & & & k_{n-1} & -k_{n-1} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_{n-1} \\ 0 \end{bmatrix} - \begin{bmatrix} Q(x_1 - y) \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}.$$

In agreement with the boundary conditions of λ, we obtain successively $\lambda_{n-1} = \cdots = \lambda_1 = 0$. If Q is invertible, then this implies that $x_1 = Cx = y$, which then resembles Configuration 1. In this case, also condition (10) is satisfied, and, thus, Configurations 1 and 3 are equivalent. If Q is not invertible, then the system is not uniquely solvable.

Remark 8 The preceding observation is an instance of the general fact that if the linear system without holonomic constraints is controllable and observable and if Q is invertible, then, provided that the data is sufficiently smooth, a solution to the optimal control problem (Configuration 3) resembles the solution of the DAE of Configuration 1. To see this, recall that in the considered situation, the system is observable if and only if $Cx - y = 0$ implies $x_1 = y$, and, by duality, that the system is controllable if and only if $B^\mathsf{T} \lambda = 0$ implies that $\lambda = 0$ for all time.

Remark 9 The equivalence of the DAE and optimal control approach for $R_0 = 0$ can also be shown for the overhead crane from the example in Sect. 6.3, which includes a holonomic constraint. In this case, we can show in a similar manner that the dual variables λ and μ vanish, so that the servo constraint $Cx = y$ has to be satisfied.

4.2 Convergence barriers

By Lemma 1, if $y \in \mathcal{C}([0, T], \mathbb{R})$ and $R_0 > 0$, then Configuration 3 with $\nu = 0$, subject to linear constraints, has a unique solution. By the results of the previous Sect. 4.1, for $R_0 = 0$, a solution only exists if y sufficiently smooth.

In this subsection, we examine how the optimal control u behaves when $R_0 \to 0$ in dependence of the smoothness of y. Consider the n-car example from Example 2 and the associated adjoint equations

$$M^\mathsf{T} \ddot{\lambda} = A^\mathsf{T} \lambda - C^\mathsf{T}(Cx - y) \tag{24}$$

with A and C from (3). With

$$\Delta_j := k_j \left(\frac{\lambda_j}{m_j} - \frac{\lambda_{j-1}}{m_{j-1}} \right), \quad j = 2, 3, \ldots, n, \tag{25}$$

we can rewrite Eq. (24) as

$$\ddot{\lambda}_1 = -\Delta_2 - (x_1 - y), \tag{26a}$$

$$\ddot{\lambda}_2 = \Delta_2 - \Delta_3, \tag{26b}$$

$$\vdots$$

$$\ddot{\lambda}_{n-1} = \Delta_n - \Delta_{n-1}, \tag{26c}$$

$$\ddot{\lambda}_n = \Delta_n. \tag{26d}$$

The gradient condition (9e) gives $\lambda_n = R_0 u$ and $\Delta_n = \ddot{\lambda}_n = R_0 \ddot{u}$. By a combination of (25) and (26a)–(26d) we recursively compute $\lambda_{n-1}, \lambda_{n-2}, \ldots, \lambda_1$ via the formulas

$$\lambda_{j-1} = -\frac{1}{k_{j-1}} \Delta_j + \frac{m_{j-1}}{m_j} \lambda_j \quad \text{and} \quad \Delta_{j-1} = \Delta_j - \ddot{\lambda}_{j-1}.$$

Finally, via (26a), we can directly relate the difference in the target $x_1 - y$ to the computed input u. Assuming uniform masses m and uniform spring constants k, for $n = 2$, we find

$$x_1 - y = R_0 \left(2\ddot{u} - \frac{1}{k} u^{(4)} \right), \tag{27}$$

whereas for $n = 3$ it must hold that

$$x_1 - y = R_0 \left(\left(\frac{1}{k} + 1 \right) u^{(4)} - \frac{1}{k} \left(\frac{1}{k} + 1 \right) u^{(6)} \right). \tag{28}$$

We observe that

1. For nonsmooth y, where we cannot expect the convergence of x_1 to y, the control u has strong peaks in its derivatives in order to fulfill (27) or (28) as $R_0 \to 0$.
2. For moderate values of R_0, the tracking error $x_1 - y$ is affected by the oscillations in the derivatives of u multiplied by multiples of $\frac{1}{k}$ depending on the length of the considered chain.
3. As $k \to \infty$, i.e., when the connections between the cars become more rigid, the higher derivatives of u are damped out from the tracking error. In fact, if one connection is rigid, then the two connected cars can be considered as one, and the index of the system reduces.

5 Solution strategies

Within this section, we review several concepts how to solve numerically mechanical systems with servo constraints. First, we comment on the classical approach where the model is given by the DAE (4a)–(4d). In this case, index reduction methods are applied, which

then allow us to integrate the resulting equations. Second, using the optimal control ansatz (5), we consider the two cases of either solving directly the optimality system, which is a boundary value problem (BVP), or the resulting Riccati equations. The latter approach may then be used to define a feedback control.

5.1 Solving high-index DAEs

As mentioned already in the introduction, the computation of the inverse dynamics of a discrete mechanical system given by a specification of a trajectory is a highly challenging problem [7, 8]. The reason is the high-index structure of the resulting DAEs. In the case of underactuated mechanical systems considered here, the systems are often of (differentiation) index 5 but may be arbitrarily high as shown in Example 2.

In order to realize the so-called *feedforward control* of the system, we have to solve this high-index DAE. Here, it is advisable to apply index reduction methods instead of solving the equations directly [4, Ch. 5.4]. A well-known approach based on a projection of the dynamics was introduced in [8]; see also [6]. For this, we have to compute time-dependent projection matrices in order to split the dynamics of the underactuated system into constrained and unconstrained parts.

Instead, we may also use the index reduction technique called *minimal extension* [20]. This technique profits from the given semiexplicit structure of the dynamical system and can be easily applied. The application to a wide range of crane models can be found in the recent papers [1, 3]. Therein, it is shown that the method of minimal extension may even be applied the second time, which then leads to a DAE of index 1 for which the numerical integration works essentially as for stiff ODEs [17, Ch. VI.1].

We remark that index reduction techniques are inevitable for numerical simulations of high-index problems. However, for applications like the n-car example given in the introduction, which is of index $2n + 1$, the DAE approach does not seem to be applicable. The modeling presented here as an optimal control problem still works properly for the general case. For a numerical example including a 3-car model, we refer to Sect. 6.

5.2 Direct solution of the optimality BVP

In this subsection, we discuss the application of the finite difference method and shooting approach in order to solve the optimality system (9a)–(9e).

5.2.1 Finite differences

The optimality system includes both initial and terminal conditions, so that the application of standard time-stepping methods is not possible. A straight-forward approach is to introduce a grid of the time domain and to apply the method of finite differences to the differential coupled equations, which leads to a (large but block-sparse) algebraic system. Alternatively, we can apply finite elements or more general collocation methods.

5.2.2 Shooting method

For the application of the shooting method, we consider the first-order system (13), i.e., we consider again the case with $r = 0$ (no additional holonomic constraint) and $\nu = 1$.

For notational reasons, we write system (13) shortly as $\mathcal{M}\dot{y} = \mathcal{K}y + h$, i.e., the vector y includes all state, input, and dual variables. Recall that the boundary conditions can be written in the form $\mathcal{B}_0 y(0) + \mathcal{B}_T y(T) = \rho$. Because of the special structure of the given boundary

conditions (initial and terminal conditions are not mixed), we may assume a reordering of the variables in order to get a system of the form

$$\mathcal{M}\dot{y} = \mathcal{K}y + h, \qquad \begin{bmatrix} \mathcal{B}_{01} \\ 0 \end{bmatrix} y(0) + \begin{bmatrix} 0 \\ \mathcal{B}_{T2} \end{bmatrix} y(T) = \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix}.$$

The aim of the shooting method is to restore the initial conditions for the entire vector y such that methods for initial value problems are applicable again. Since the initial values of the state variables and $\dot{u}$ (respectively, u if initial data for the input were prescribed) are already given, we can apply the so-called *reduced superposition* [2, Ch. 4.2.4]. This reduces the computational effort of the method. Within the following algorithm, we denote by s the size of the original system and, thus, by $2s$ the size of the first-order system we want to solve.

Step 1: Search for the fundamental solution of the corresponding homogeneous system. However, using the reduced superposition, it is sufficient to compute $Y \in \mathbb{R}^{2s,s}$ solving

$$\mathcal{M}\dot{Y} = \mathcal{K}Y, \qquad Y(0) = \begin{bmatrix} I & & 0 & \\ & I & & 0 \\ & & & I & 0 \end{bmatrix}^\mathsf{T}.$$

Step 2: Find a solution $w \in \mathbb{R}^{2s}$ of the initial value problem

$$\mathcal{M}\dot{w} = \mathcal{K}w + h, \qquad w(0) = \begin{bmatrix} 0 & 0 & x^0 & v^0 & 0 & 0 \end{bmatrix}^\mathsf{T}.$$

Step 3: Find the coefficients $c \in \mathbb{R}^s$ given by the linear system

$$\mathcal{B}_{T2}Y(T)c = \rho_2 - \mathcal{B}_{T2}w(T),$$

which then gives the solution of the BVP as $y = Yc + w$. Thus, an approximation of y can be either given directly if the matrices Y and w were stored on the entire time grid or by solving the IVP

$$\mathcal{M}\dot{y} = \mathcal{K}y + h \quad \text{with } y(0) = Y(0)c + w(0).$$

Remark 10 (Comparison of computational effort) Assume that we always use the same time step size with N grid points. Then, the finite difference method leads to a system of size $2sN$ such that the computational effort is quadratic in N. For the shooting method, we have to solve several initial value problems (each using N time steps). Note that the size of the systems is bounded by the size of the original BVP such that the overall costs are only linear in N (but with a large constant depending on s^2).

Remark 11 A more stable extension of the (single) shooting method is called the *multiple shooting method* [2, Ch. 4.4.3]. For this, the time interval $[0, T]$ is partitioned by shooting points $0 = t_1 < t_2 < \cdots < t_{N+1} = T$. On each subinterval $[t_i, t_{i+1}]$, we may compute a solution $y_i(t) = Y_i(t)c_i + w_i(t)$ similarly as before. The coefficient vectors $c_i \in \mathbb{R}^s$ are given by a linear system that contains the boundary and the continuity conditions in-between the time steps.

Remark 12 For the other cases, i.e., for $r > 0$ (with holonomic constraints) or different values of ν, we may need to use different techniques, depending in the structure of the BVP.

In the case $r = 0$, $\nu = 0$ (cf. Sect. 3.5.1, where we obtain as an optimality system an index-1 DAE) and need to consider shooting methods. For this, we refer to [24]. In the case $r > 0$, for which we obtain index-3 systems, we refer to [11] or, after an index reduction to index 2, also to [13].

5.3 Riccati approach

In the linear case and if $\nu = 0$, i.e., if no derivatives of u appear in the optimality system (9a)–(9e), the BVP can be solved via a Riccati decoupling. This requires the formulation as a first-order system as in (13), which already is in the standard form considered; see, e.g., in [25, Ch. 5]. In the case of holonomic constraints, we can use the results on constrained Riccati equations given in [15], which readily apply to constrained multibody equations in the *Gear–Gupta–Leimkuhler* formulation [14].

6 Numerical examples

In this section, we provide several numerical experiments. First, we consider the two-car system from Example 1, i.e., an example without holonomic constraints. Second, we add a third car, which then gives an index-7 DAE in the original formulation. Finally, we consider an overhead crane as an example with $r > 0$, i.e., with a holonomic constraint. The code is written in *Python* and can be obtained from the author's public *Github* repository [16].

6.1 Two-car example

We consider the two-car example from the introduction; see Fig. 1. Recall that the equations of motion (1a)–(1c) form a DAE of index 5. As in [5, Ex. 3], the following parameters were used within the computations:

$$m_1 = 2 \text{ kg}, \qquad m_2 = 1 \text{ kg}, \qquad k = 1 \tfrac{\text{N}}{\text{m}}, \qquad d = 0.5 \text{ m}.$$

6.1.1 Comparison of DAE and optimal control solution

The initial values within the computations are given by

$$x_1^0 = 0.5 \text{ m}, \qquad v_1^0 = 0 \tfrac{\text{m}}{\text{s}}, \qquad x_2^0 = 0 \text{ m}, \qquad v_2^0 = 0 \tfrac{\text{m}}{\text{s}}.$$

For the definition of a rest-to-rest maneuver, we introduce the polynomial

$$p(s) = 1716s^7 - 9009s^8 + 20020s^9 - 24024s^{10} + 16380s^{11} - 6006s^{12} + 924s^{13}.$$

With this and $y_0 = 0.5$ m, $y_f = 2.5$ m, we define on the time interval $[0, 4s]$ the target trajectory

$$y(t) = \begin{cases} y_0 & \text{if } 0 \leq t < 1, \\ y_0 + p(\tfrac{t-1}{2})(y_f - y_0) & \text{if } 1 \leq t \leq 3, \\ y_f & \text{if } 3 < t \leq 4. \end{cases} \tag{29}$$
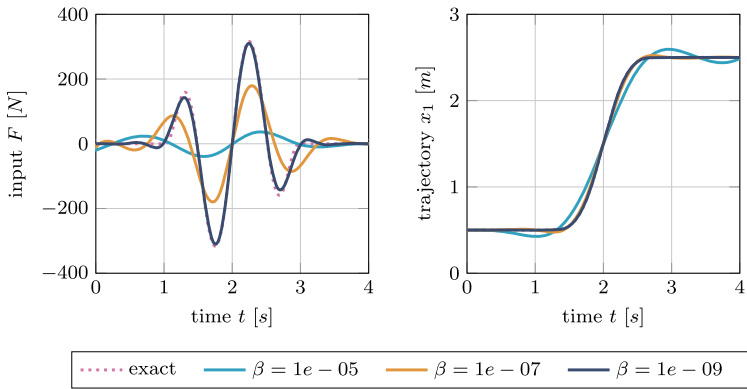
Fig. 2 The exact input force F (DAE solution) and the forces obtained through the optimal control formulation for different values of the penalization parameter β (*left*) and the corresponding trajectories (*right*) for the two-car example

Note that y is smooth enough, so that the DAE solution (which we refer to as an exact solution) exists. Recall that this requires consistent initial positions and initial velocities of the cars as mentioned in Remark 1, namely

$$y(0) = x_1^0, \qquad \dot{y}(0) = v_1^0, \qquad m_1\ddot{y}(0) = -\left(x_1^0 - x_2^0 - d\right), \qquad m_1 y^{(3)}(0) = -\left(v_1^0 - v_2^0\right).$$

We compare the exact solutions, which are readily computable from the systems equation, to the trajectories and input forces obtained from solving the associated optimal control problem of Configuration 3 with parameters $Q = S = 1$ and for varying $R_0 = \beta \in \mathbb{R}$. The occurring linear boundary value problem is solved by finite differences on a regular grid of size $\tau = 0.01$ s.

As expected, depending on the penalization parameter β, the optimal control approach leads to input forces that are smaller than the exact force F; see Fig. 2 (left). The reduction of the amplitude is best seen for large values of β. The optimal control problem is a compromise of costs and accuracy, as can be seen from the deviations from the target trajectory, which decrease for smaller values of the penalization parameter β; see Fig. 2 (right).

6.1.2 Feedback representations of the optimization solutions

Another advantage of the optimization approach is that the optimal control can be realized as a feedback. In fact, the first-order optimality conditions (9a)–(9e) suggest that u depends linearly on λ, which depends, possibly nonlinearly, on the state x. For the considered linear case of Example 1, the optimality system can be solved via a differential Riccati equation [25, Ch. 5.1], which directly leads to a feedback representation of the optimal control.

We stay with the example of the 2-car setup to illustrate the benefits of the feedback representation. If we simply apply the known exact control solution to the considered system, then a perturbation, e.g., in the initial position or initial velocity, necessarily leads to a drift off the desired trajectory; 0 see Fig. 3 (left). In contrast, the feedback solution of the optimal control problem with $\beta_0 = 10^{-9}$ detects and damps possible perturbations; see Fig. 3 (right).
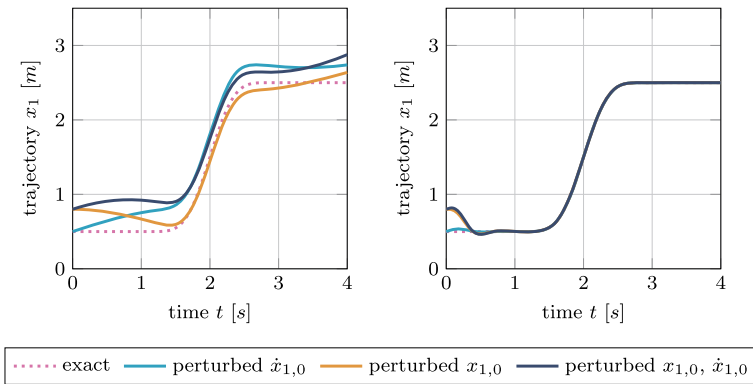
Fig. 3 Benefits of the feedback representation: The output obtained from applying the exact control solution directly (*left*) and the output obtained via a feedback representation of the control solution of the optimal control problem (*right*) in the case of perturbed initial values

6.2 Three-car example

In this subsection, we add an additional car, i.e., we consider Example 2 with $n = 3$. This means that the positions of the bodies are given by x_1, x_2, and x_3, where the trajectory of x_1 is prescribed by y. Recall that this gives a DAE of index 7 rather than of index 5 as in the previous example. As parameters, we set

$$m_1 = m_2 = 1 \text{ kg}, \qquad m_3 = 2 \text{ kg}, \qquad k = 1 \frac{\text{N}}{\text{m}}, \qquad d = 0.5 \text{ m}.$$

As initial conditions we have

$$x_1^0 = 0.5 \text{ m}, \ v_1^0 = 0 \frac{\text{m}}{\text{s}}, \qquad x_2^0 = 0 \text{ m}, \ v_2^0 = 0 \frac{\text{m}}{\text{s}}, \qquad x_3^0 = -0.5 \text{ m}, \ v_3^0 = 0 \frac{\text{m}}{\text{s}}.$$

For the simulation, we consider the same rest-to-rest maneuver as in (29). As required, the prescribed trajectory is six times continuously differentiable, and the initial conditions satisfy all consistency conditions; cf. Assumption 1 and Remark 1.

The exact solution and the results from the optimal control problem for different values of the penalization parameter are shown in Fig. 4. The weak coupling from the input on the third car to the output measured on the first car is apparent in the large peaks in the exact input force. The optimization approach leads to significantly reduced amplitudes in the input at the expense of a certain deviation from the prescribed target trajectory.

6.3 Overhead crane

The servo-constraint problem of this section was originally formulated in terms of minimal coordinates in [5] and was recast in redundant coordinates in [10]; see Fig. 5. We follow here the latter approach, which then fits into the framework of Sect. 2.

For this, we consider the state and input variables

$$x = [s, \ \alpha, \ x_d, \ z_d]^\mathsf{T}, \qquad u = [F, \ M_n]^\mathsf{T}.$$

With these redundant coordinates (one additional variable and one additional equation), we need to add one holonomic constraint and the corresponding Lagrange multiplier $p \in \mathbb{R}$.
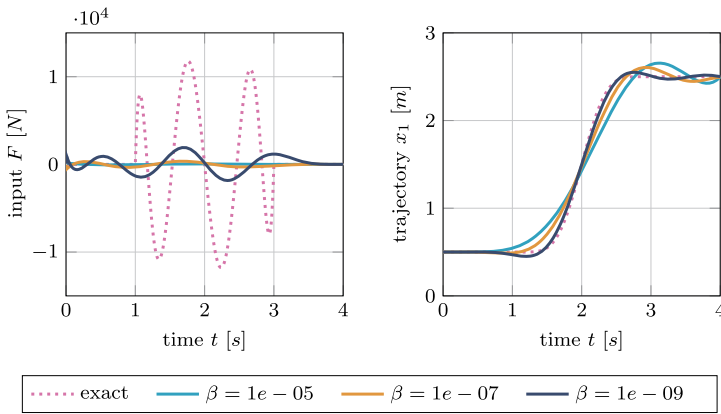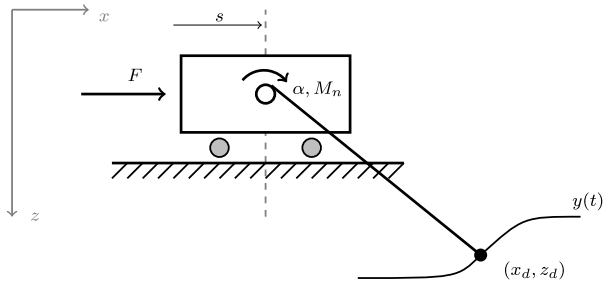
Fig. 4 The exact input force F (DAE solution) and the forces obtained through the optimal control formulation for different values of the penalization parameter β (*left*) and the corresponding trajectories (*right*) for the three-car example

Fig. 5 Overhead trolley crane with the notation of the rotationless formulation introduced in [10]

The overall system reads

$$
\begin{bmatrix} m_t & & & \\ & J & & \\ & & m & \\ & & & m \end{bmatrix} \begin{bmatrix} \ddot{s} \\ \ddot{\alpha} \\ \ddot{x}_d \\ \ddot{z}_d \end{bmatrix} - G^{\mathsf{T}}(x)p - \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} u = \begin{bmatrix} 0 \\ -rmg \\ 0 \\ mg \end{bmatrix}, \tag{30a}
$$

$$
(x_d - s)^2 + z_d^2 - (r\alpha)^2 = 0, \tag{30b}
$$

$$
\begin{bmatrix} x_d \\ z_d \end{bmatrix} - y = 0 \tag{30c}
$$

with

$$
G(x) = 2 \begin{bmatrix} s - x_d & -r^2\alpha & x_d - s & z_d \end{bmatrix}, \qquad \left(G(x)^{\mathsf{T}} p \right)_x = 2p \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & -r^2 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.
$$

In the optimal control approach, with a cost functional as in Configuration 3, the solution is obtained through the solution of the additional adjoint equations (9c) and (9d), the gradient condition (9e), and the boundary conditions (10) and (11).
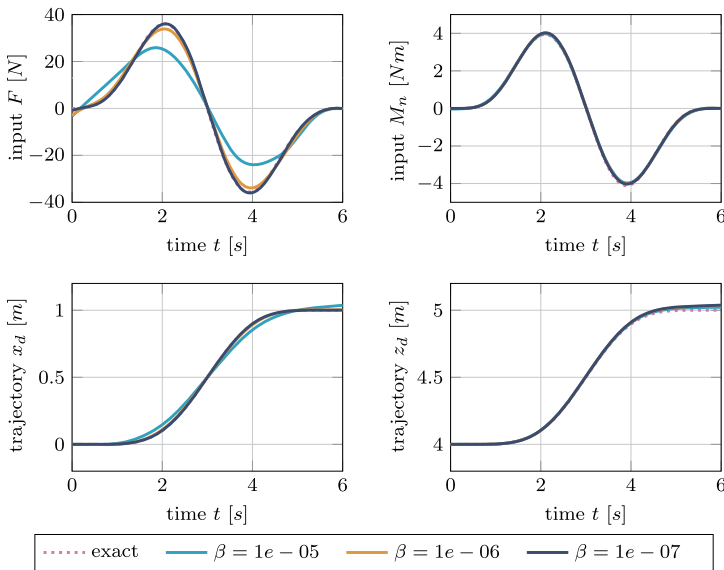
Fig. 6 The forces F and M_n obtained through linearization and optimization for varying penalization parameter β (*top*) and the resulting trajectories (*bottom*) for the example of the overhead crane

We consider the system parameters

$$m_t = 10 \text{ kg}, \qquad J = 0.1 \text{ Nm}, \qquad m = 1 \text{ kg}, \qquad r = 0.1 \text{ m}, \qquad g = 9.81 \, \tfrac{\text{m}}{\text{s}^2}$$

with initial values

$$x_0 = [0 \text{ m}, \ 40 \text{ m}, \ 0 \text{ m}, \ 4 \text{ m}]^\mathsf{T} \quad \text{and} \quad \dot{x}_0 = [0, \ 0, \ 0, \ 0]^\mathsf{T}.$$

Furthermore, we consider a target trajectory as defined in (29) but on the time interval $[0, 6 \text{ s}]$ and with the vector-valued starting and terminal points

$$y_0 = [0, \ 4]^\mathsf{T} \quad \text{and} \quad y_f = [1, \ 5]^\mathsf{T}.$$

We linearize the resulting nonlinear boundary value problems with holonomic constraints around the constant solution obtained with $u \equiv [0, \ 0]^\mathsf{T}$ and solve it via finite differences. The computed approximation to the optimal control is then evaluated in the actual nonlinear model (30a), (30b) and compared to the analytical solution of (30a)–(30c).

As the plots in Fig. 6 show, this combined linearization and optimization approach leads to a decent approximation of the actual control.

7 Conclusion

Within this paper, we have considered mechanical systems with a partly specified motion, which are usually modeled by DAEs of index ≥ 5. Such models require strong regularity assumptions, and their numerical treatment is extremely challenging because of the sensitivity to perturbations. Because of this, an alternative modeling approach was introduced, which

relaxes the prescribed servo constraint and considers a minimization problem instead. By this we decrease the possible errors that occur in the simulation of a high-index DAE but include an additional error since we do not satisfy the constraint exactly. However, this modeling error is controllable by the penalization parameters.

By numerical examples we have shown the advantages of the optimal control approach. First, the resulting control effort is much smaller at the price of only a small error in the constraint and, thus, more realistic since this corresponds to a reduction of costs in real-world applications. Second, the approach is less sensitive to perturbations such as inconsistent initial data.

Finally, we point out that the usefulness of the proposed approach needs to be determined in experiments as well. Thus, the next steps in this development direction will be the inclusion of sensor and actuator dynamics and the implementation in an experimental setup.

References

1. Altmann, R., Betsch, P., Yang, Y.: Index reduction by minimal extension for the inverse dynamics simulation of cranes. Multibody Syst. Dyn. **36**(3), 295–321 (2016)
2. Ascher, U.M., Mattheij, R.M.M., Russell, R.D.: Numerical Solution of Boundary Value Problems for Ordinary Differential Equations. SIAM, Philadelphia (1995)
3. Betsch, P., Altmann, R., Yang, Y.: Numerical integration of underactuated mechanical systems subjected to mixed holonomic and servo constraints. In: Font-Llagunes, J.M. (ed.) Multibody Dynamics: Computational Methods and Applications, vol. 42, pp. 1–18. Springer, Berlin (2016)
4. Brenan, K.E., Campbell, S.L., Petzold, L.R.: Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations. SIAM, Philadelphia (1996)
5. Blajer, W., Kołodziejczyk, K.: A geometric approach to solving problems of control constraints: Theory and a DAE framework. Multibody Syst. Dyn. **11**(4), 343–364 (2004)
6. Blajer, W., Kołodziejczyk, K.: Improved DAE formulation for inverse dynamics simulation of cranes. Multibody Syst. Dyn. **25**(2), 131–143 (2011)
7. Blajer, W.: Index of differential-algebraic equations governing the dynamics of constrained mechanical systems. Appl. Math. Model. **16**(2), 70–77 (1992)
8. Blajer, W.: Dynamics and control of mechanical systems in partly specified motion. J. Franklin Inst. B **334**(3), 407–426 (1997)
9. Blajer, W.: The use of servo-constraints in the inverse dynamics analysis of underactuated multibody systems. J. Comput. Nonlinear Dyn. **9**(4), 1–11 (2014)
10. Betsch, P., Uhlar, S., Quasem, M.: On the incorporation of servo constraints into a rotationless formulation of flexible multibody dynamics. In: Proceedings of Multibody Dynamics 2007—ECCOMAS Thematic Conference, 25–28 June Milano, Italy (2007)
11. Clark, K.D., Petzold, L.R.: Numerical solution of boundary value problems in differential-algebraic systems. SIAM J. Sci. Stat. Comput. **10**(5), 915–936 (1989)
12. Fliess, M., Lévine, J., Martin, P., Rouchon, P.: Flatness and defect of non-linear systems: Introductory theory and examples. Int. J. Control **61**(6), 1327–1361 (1995)
13. Gerdts, M.: Direct shooting method for the numerical solution of higher-index DAE optimal control problems. J. Optim. Theory Appl. **117**(2), 267–294 (2003)
14. Gear, C.W., Gupta, G.K., Leimkuhler, B.: Automatic integration of Euler–Lagrange equations with constraints. J. Comput. Appl. Math. **12–13**, 77–90 (1985)

15. Heiland, J.: Decoupling and optimization of differential-algebraic equations with application in flow control. PhD thesis, TU, Berlin (2014)
16. Heiland, J.: holo-servo-opt—A Python module for the solution of multi-body systems with holonomic and servo constraints via optimal control (2015). https://github.com/highlando/holo-servo-opt
17. Hairer, E., Wanner, G.: Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems, 2nd edn. Springer, Berlin (1996)
18. Isidori, A.: Nonlinear Control Systems, 2nd edn. Springer, Berlin (1989)
19. Kirgetov, V.I.: The motion of controlled mechanical systems with prescribed constraints (servoconstraints). J. Appl. Math. Mech. **31**, 465–477 (1967)
20. Kunkel, P., Mehrmann, V.: Index reduction for differential-algebraic equations by minimal extension. Z. Angew. Math. Mech. **84**(9), 579–597 (2004)
21. Kunkel, P., Mehrmann, V.: Differential-Algebraic Equations. Analysis and Numerical Solution. European Mathematical Society Publishing House, Zürich (2006)
22. Kunkel, P., Mehrmann, V.: Optimal control for unstructured nonlinear differential-algebraic equations of arbitrary index. Math. Control Signals Syst. **20**(3), 227–269 (2008)
23. Kunkel, P., Mehrmann, V.: Formal adjoints of linear DAE operators and their role in optimal control. Electron. J. Linear Algebra **22**, 672–693 (2011)
24. Lamour, R.: A shooting method for fully implicit index-2 differential algebraic equations. SIAM J. Sci. Comput. **18**(1), 94–114 (1997)
25. Locatelli, A.: Optimal Control. An Introduction. Birkhäuser, Basel (2001)
26. Nachbagauer, K., Oberpeilsteiner, S., Sherif, K., Steiner, W.: The use of the adjoint method for solving typical optimization problems in multibody dynamics. J. Comput. Nonlinear Dyn. **10**(6), 061011 (2015)
27. Seifried, R., Blajer, W.: Analysis of servo-constraint problems for underactuated multibody systems. Mech. Sci. **4**(1), 113–129 (2013)
28. Sontag, E.D.: Mathematical Control Theory. Deterministic Finite Dimensional Systems. Springer, New York (1998)
29. Wan, F.Y.M.: Introduction to the Calculus of Variations and Its Applications. Chapman & Hall, New York (1995)