Discrete mechanics and optimal control for constrained systems

S. Leyendecker[1,*,†], S. Ober-Blöbaum[2], J. E. Marsden[2] and M. Ortiz[3]

[1]*Aeronautics and Control and Dynamical Systems, California Institute of Technology, 1200 E. California Boulevard, Mail Code 107-81, Pasadena, CA 91125, U.S.A.*
[2]*Control and Dynamical Systems, California Institute of Technology, 1200 E. California Boulevard, Mail Code 205-45, Pasadena, CA 91125, U.S.A.*
[3]*Aeronautics and Mechanical Engineering, California Institute of Technology, 1200 E. California Boulevard, Mail Code 205-45, Pasadena, CA 91125, U.S.A.*

SUMMARY

The equations of motion of a controlled mechanical system subject to holonomic constraints may be formulated in terms of the states and controls by applying a constrained version of the Lagrange-d'Alembert principle. This paper derives a structure-preserving scheme for the optimal control of such systems using, as one of the key ingredients, a discrete analogue of that principle. This property is inherited when the system is reduced to its minimal dimension by the discrete null space method. Together with initial and final conditions on the configuration and conjugate momentum, the reduced discrete equations serve as nonlinear equality constraints for the minimization of a given objective functional. The algorithm yields a sequence of discrete configurations together with a sequence of actuating forces, optimally guiding the system from the initial to the desired final state. In particular, for the optimal control of multibody systems, a force formulation consistent with the joint constraints is introduced. This enables one to prove the consistency of the evolution of momentum maps. Using a two-link pendulum, the method is compared with existing methods. Further, it is applied to a satellite reorientation maneuver and a biomotion problem. Copyright © 2009 John Wiley & Sons, Ltd.

1. INTRODUCTION

This work combines two recently developed methods, namely the discrete null space method that is suitable for the accurate, robust and efficient time integration of constrained dynamical systems (in particular for multibody dynamics) and an approach to discrete mechanics and optimal control (DMOC) based on a discretization of the Lagrange-d'Alembert principle. The new method's acronym is DMOCC. The idea of this combination has been introduced briefly in [1] and is investigated in detail for three-dimensional multibody systems consisting of rigid bodies interconnected by joints in this work.

From various available methods used to enforce holonomic constraints in the framework of the Hamiltonian or Lagrangian formalism (see, for example, [2, 3]

---

*Correspondence to: S. Leyendecker, Aeronautics and Control and Dynamical Systems, California Institute of Technology, 1200 E. California Boulevard, Mail Code 107-81, Pasadena, CA 91125, U.S.A.
†E-mail: sleye@caltech.edu

and for a computational approach [4]), the focus in this paper is on two methods yielding exact constraint fulfillment, the Lagrange multiplier method and a null space method, described in, for example, [5].

Because of the relatively simple structure of the evolution equations derived from the Lagrange multiplier method, their temporal discrete form can be derived easily using mechanical integrators as demonstrated among others in [6–8]. However, the presence of Lagrange multipliers among the set of unknowns enlarges the number of equations and causes the discrete system to be ill-conditioned for small time steps as reported (among others) by [9, 10]. In contrast to this undesirable situation, the use of a specific null space method, especially in conjunction with a reparameterization in generalized coordinates, has the advantageous property of a small dimensional system of equations. On the other hand, these evolution equations have a highly complicated structure, causing the derivation of their temporal discrete form to be expensive and therefore, in most cases, not recommended [11, 12].

A remedy for these difficulties is found in the discrete null space method introduced in [13], which proposes a reversal of two of the main steps when designing a specific numerical method. In the first step, the discrete form of the simple structured DAEs resulting from the use of the Lagrange multiplier method is derived using a mechanical integrator, for example, an energy-momentum conserving integrator [6, 7] or a variational integrator leading to a symplectic-momentum conserving scheme [8]. For forced systems, both methods correctly compute the change in momentum maps. Then, in the second step, the transition to the reduced scheme and finally the nodal reparameterization are performed in the temporal discrete setting in complete analogy to the procedure in the continuous case according to the discrete null space method. The resulting time-stepping scheme performs excellently in all relevant categories. First of all, it yields the smallest possible dimension for the system of equations, promising lower computational costs than other schemes. Second, it is second-order accurate and inherits the conservation properties from the constrained scheme and third, the condition number of the scheme is independent of the time step.

Summarizing, the discrete null space method is especially suited for the accurate simulation of large dimensional systems subject to a large number of constraints. In particular, the resulting equations lend themselves as dynamic constraints in an optimization algorithm since their dimension is minimal, thus only the exactly required number of unknowns has to be determined.

To find local solutions of nonlinear optimal control problems consisting of a given objective functional and equations describing the underlying dynamics of the system, a numerical method falling into the class of direct methods is used here. Thereby, the state and control variables are discretized directly in order to transform the optimal control problem into a finite-dimensional nonlinear constrained optimization problem that can be solved by standard nonlinear optimization techniques such as sequential quadratic programming (see [14, 15]). In contrast to other methods such as, for example, shooting [16–18], multiple shooting [19, 20] or collocation methods [21, 22], relying on a direct integration of the associated ordinary differential equations parameterized by states and controls or the controls only (see also [23] and [24] for an overview of the current state of the art), a recently developed method DMOC (see [25, 26]) is used here. It is based on the discretization of the variational structure of the mechanical system directly. In the context of variational integrators, as in [27], the discretization of the Lagrange-d'Alembert principle leads to the structure-preserving time-stepping equations that serve as equality constraints for the resulting finite-dimensional nonlinear optimization problem. In [26, 28, 29] DMOC was first applied to low-orbital thrust transfers and the optimal control of formation flying satellites including an algorithm that exploits a hierarchical structure of that problem.

In this work, DMOC is used to find optimal trajectories of state and control variables for systems of rigid bodies subject to joint constraints. Each rigid body is viewed as a constrained continuum, that is, it is described in terms of redundant coordinates subject to holonomic constraints [30, 31]. This is in contrast to rotation-based approaches taken, for example, in [32, 33]. Here, the equations of motion assume the form of DAEs with a constant mass matrix. Their temporal

discrete form can be derived and reduced according to the discrete null space method. This procedure has the advantage of circumventing the difficulties associated with rotational parameters [34] and it can be generalized easily to the modelling of geometrically exact beams and shells and to multibody systems consisting of theses structures as developed in [35–37].

The combination of the two proposed methods involves several specific benefits. First of all, the discrete dynamics equation constraining the optimal control problem in DMOC can be formulated easily. Using the discrete Lagrange-d'Alembert principle, they are derived as the discrete analogue to the simple structured evolution equations, where the configuration constraints are enforced using Lagrange multipliers. Second, the discrete null space method reduces the dynamics constraints to the smallest possible number of equations and variables, which leads to lower computational costs for the optimization algorithm. Third, the benefit of exact constraint fulfillment and correct computation of the change in momentum maps is guaranteed by the optimization algorithm. These benefits are important, especially for high-dimensional rigid-body systems with joint constraints.

An outline of the paper is as follows. Section 2 fixes the formulation of the continuous optimal control problem of constrained dynamics, which is formulated in the discrete setting in Section 3. Techniques for rigid-body systems are set up in Section 4. The main contribution is found in Sections 5 and 6. In particular, actuating forces being consistent with the specific joint constraints are given in Section 5 and structure preservation of the resulting time-stepping scheme is proved. Section 6 applies the theory developed in the paper to the optimal control of multibody systems. Numerical examples from the field of satellite reorientation maneuvers and biomotion as well as a comparison with existing methods are presented in Section 7.

2. CONSTRAINED DYNAMICS AND OPTIMAL CONTROL

This section derives the equations of motion for forced holonomically constrained systems; these equations are to be fulfilled as constraints in the optimization

problem. The transformation of the differential algebraic equations by the null space method with reparameterization, and in particular the equivalence of the resulting equations of motion, is described in detail in [38] for conservative systems.

Consider an n-dimensional mechanical system with the time-dependent configuration vector $\boldsymbol{q}(t) \in Q$ and velocity vector $\dot{\boldsymbol{q}}(t) \in T_{\boldsymbol{q}(t)}Q$, where $t \in [t_0, t_N] \subset \mathbb{R}$ denotes the time and $N \in \mathbb{N}$. Let the configuration be constrained by the function $\boldsymbol{g}(\boldsymbol{q}) = \boldsymbol{0} \in \mathbb{R}^m$ with constraint manifold $C = \{\boldsymbol{q} \in Q \mid \boldsymbol{g}(\boldsymbol{q}) = \boldsymbol{0}\}$ and influenced by the force field $\boldsymbol{f} : \mathbb{R}^{n-m} \times TQ \to T^*Q$.

The optimization problem: The goal is to determine the optimal trajectory and force field, such that the system is moved from the initial state $(\boldsymbol{q}^0, \dot{\boldsymbol{q}}^0) \in TC$ to the final state $(\boldsymbol{q}^N, \dot{\boldsymbol{q}}^N) \in TC$, obeying the equations of motion and at the same time, the objective functional

$$J(\boldsymbol{q}, \dot{\boldsymbol{q}}, \boldsymbol{f}) = \int_{t_0}^{t_N} B(\boldsymbol{q}, \dot{\boldsymbol{q}}, \boldsymbol{f}) \, \mathrm{d}t \tag{1}$$

is to be minimized. Here, $B(\boldsymbol{q}, \dot{\boldsymbol{q}}, \boldsymbol{f}) : TC \times T_{\boldsymbol{q}}^*Q \to \mathbb{R}$ is a given cost function.

The constrained Lagrange-d'Alembert principle: As already mentioned, the motion has to obey the equations of motion, which, in the present case, are based on a constrained version of the Lagrange-d'Alembert principle (see e.g. [39]), which requires that

$$\delta \int_{t_0}^{t_N} [L(\boldsymbol{q}, \dot{\boldsymbol{q}}) - \boldsymbol{g}^{\mathrm{T}}(\boldsymbol{q}) \cdot \boldsymbol{\lambda}] \, \mathrm{d}t + \int_{t_0}^{t_N} \boldsymbol{f} \cdot \delta \boldsymbol{q} \, \mathrm{d}t = 0 \tag{2}$$

for all variations $\delta \boldsymbol{q} \in TQ$ vanishing at the endpoints and $\delta \boldsymbol{\lambda} \in \mathbb{R}^m$. The Lagrangian $L : TQ \to \mathbb{R}$ equals the kinetic energy $\frac{1}{2}\dot{\boldsymbol{q}}^{\mathrm{T}} \cdot \boldsymbol{M} \cdot \dot{\boldsymbol{q}}$ including the consistent mass matrix $\boldsymbol{M} \in \mathbb{R}^{n \times n}$ minus a potential function $V : Q \to \mathbb{R}$. Furthermore, $\boldsymbol{\lambda}(t) \in \mathbb{R}^m$ represents the vector of time-dependent Lagrange multipliers. The last term represents the virtual work resulting from the force field. The constrained Lagrange-d'Alembert principle (2) leads to the differential–algebraic system of equations of motion

$$\frac{\partial L(\boldsymbol{q}, \dot{\boldsymbol{q}})}{\partial \boldsymbol{q}} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L(\boldsymbol{q}, \dot{\boldsymbol{q}})}{\partial \dot{\boldsymbol{q}}}\right) - \boldsymbol{G}^{\mathrm{T}}(\boldsymbol{q}) \cdot \boldsymbol{\lambda} + \boldsymbol{f} = \boldsymbol{0}$$

$$\boldsymbol{g}(\boldsymbol{q}) = \boldsymbol{0} \tag{3}$$

where $G(q) = Dg(q)$ denotes the Jacobian of the constraints. The vector $-G^T(q) \cdot \lambda$ represents the constraint forces that prevent the system from deviations off the constraint manifold.

The null space method: Assuming that the constraints are independent, for every $q \in C$ the basis vectors of $T_q C$ form an $n \times (n-m)$ matrix $P(q)$ with corresponding linear map $P(q): \mathbb{R}^{n-m} \to T_q C$. This matrix is called null space matrix, since $\text{range}(P(q)) = \text{null}(G(q)) = T_q C$. Thus, a premultiplication of the differential equation $(3)_1$ by $P^T(q)$ eliminates the constraint forces including the Lagrange multipliers from the system.

Reparameterization: For many applications, it is possible to find a local parameterization of the constraint manifold $F: U \subseteq \mathbb{R}^{n-m} \to C$ in terms of independent generalized coordinates $u \in U$. Then the Jacobian $DF(u)$ of the coordinate transformation plays the role of a null space matrix. Since the constraints $(3)_2$ are fulfilled automatically by the reparameterized configuration variable $q = F(u)$, the system is reduced to $n-m$ second-order differential equations

$$P^T(q) \cdot \left[\frac{\partial L(q, \dot{q})}{\partial q} - \frac{d}{dt} \left(\frac{\partial L(q, \dot{q})}{\partial \dot{q}} \right) + f \right] = 0 \quad (4)$$

Owing to the presence of constraints, the forces f are not independent. They can be calculated in terms of the time-dependent generalized control forces $\tau(t) \in T^*U$. Consequently, there are $n-m$ independent generalized forces acting on the generalized degrees of freedom. These can be calculated as $\tau = (\partial F / \partial u)^T \cdot f$, see, for example, [40]. On the other hand, a redundant force vector $f \in T^*Q$ can be computed via

$$f = B^T(q) \cdot \tau \quad (5)$$

with the $n \times (n-m)$ configuration-dependent input transformation matrix $B^T: T^*U \to T^*Q$. Therefore, the choice of the transformation matrix must ensure the consistency of momentum maps in the sense that they change only and exactly according to the generalized force.

Thus, the optimal control of constrained dynamics gives rise to the optimization problem in Bolza form consisting of the optimization of the objective function (1) subject to the reduced equations of constrained motion (4).

3. CONSTRAINED DISCRETE DYNAMICS AND OPTIMAL CONTROL

Analogous steps are performed in the temporal discrete variational setting to derive the forced constrained discrete Euler–Lagrange equations and their reduction to minimal dimension. Corresponding to the configuration manifold Q, the discrete phase space is defined by $Q \times Q$, which is locally isomorphic to TQ. For a constant time step $h \in \mathbb{R}$, a path $q: [t_0, t_N] \to Q$ is replaced by a discrete path $q_d: \{t_0, t_0 + h, , \ldots, t_0 + Nh = t_N\} \to Q$, $N \in \mathbb{N}$, where $q_n = q_d(t_n)$ is viewed as an approximation to $q(t_n)$ at $t_n = t_0 + nh$. Similarly, $\lambda_n = \lambda_d(t_n)$ approximates the Lagrange multipliers, while the force field f is approximated by two discrete forces $f_n^-, f_n^+: T^*U \times Q \to T^*Q$ in a way that respects work, as is explained below.

Discrete constrained Lagrange-d'Alembert principle: According to the derivation of variational integrators for constrained dynamics in [38], the action integral in (2) is approximated in a time interval $[t_n, t_{n+1}]$ using the discrete Lagrangian $L_d: Q \times Q \to \mathbb{R}$ and the discrete constraint function $g_d: Q \to \mathbb{R}$ via

$$L_d(q_n, q_{n+1}) - \frac{1}{2} g_d^T(q_n) \cdot \lambda_n - \frac{1}{2} g_d^T(q_{n+1}) \cdot \lambda_{n+1}$$

$$\approx \int_{t_n}^{t_{n+1}} L(q, \dot{q}) - g^T(q) \cdot \lambda \, dt \quad (6)$$

Among various possible choices to approximate this integral, in this work the midpoint rule is used for the Lagrangian, that is,

$$L_d(q_n, q_{n+1}) = hL \left(\frac{q_{n+1} + q_n}{2}, \frac{q_{n+1} - q_n}{h} \right) \quad (7)$$

and for the constraints $g_d^T(q_n) = hg^T(q_n)$ is used. Similarly, the virtual work is approximated by

$$f_n^- \cdot \delta q_n + f_n^+ \cdot \delta q_{n+1} \approx \int_{t_n}^{t_{n+1}} f \cdot \delta q \, dt \quad (8)$$

where f_n^+, f_n^- are called the left and right discrete forces, respectively. They are specified in (12).

The discrete version of the constrained Lagrange-d'Alembert principle (2) requires the discrete path $\{q_n\}_{n=0}^N$ and multipliers $\{\lambda_n\}_{n=0}^N$ to fulfill

$$\delta \sum_{n=0}^{N-1} L_d(q_n, q_{n+1}) - \frac{1}{2} g_d^{\mathrm{T}}(q_n) \cdot \lambda_n - \frac{1}{2} g_d^{\mathrm{T}}(q_{n+1}) \cdot \lambda_{n+1}$$

$$+ \sum_{n=0}^{N-1} f_n^- \cdot \delta q_n + f_n^+ \cdot \delta q_{n+1} = 0 \qquad (9)$$

for all variations $\{\delta q_n\}_{n=0}^N$ and $\{\delta \lambda_n\}_{n=0}^N$ with $\delta q_0 = \delta q_N = 0$, which is equivalent to the constrained forced discrete Euler–Lagrange equations

$$D_2 L_d(q_{n-1}, q_n) + D_1 L_d(q_n, q_{n+1})$$

$$- G_d^{\mathrm{T}}(q_n) \cdot \lambda_n + f_{n-1}^+ + f_n^- = 0 \qquad (10)$$

$$g(q_{n+1}) = 0$$

for $n = 1, \ldots, N-1$, where $G_d(q_n)$ denotes the Jacobian of $g_d(q_n)$. Note that the time-stepping scheme (10) has not been deduced by discretizing (3), but rather via a discrete variational principle.

The discrete null space method: Analogous to the continuous case, to eliminate the discrete constraint forces from (10), a discrete null space matrix fulfilling $\mathrm{range}(P(q_n)) = \mathrm{null}(G_d(q_n))$ is employed. Pre-multiplying $(10)_1$ by the transposed discrete null space matrix cancels the constraint forces; that is, the Lagrange multipliers are eliminated from the set of unknowns and the system's dimension is reduced to n.

Nodal reparameterization: As in the continuous case, a reduction of the system to the minimal possible dimension can be accomplished by a local reparametrization of the constraint manifold in the neighborhood of the discrete configuration variable. At the time nodes, q_n is expressed in terms of the discrete generalized coordinates $u_n \in U \subseteq \mathbb{R}^{n-m}$ by the map $F: U \subseteq \mathbb{R}^{n-m} \times Q \to C$, such that the constraints are fulfilled.

$$q_n = F(u_n, q_{n-1}) \quad \text{with } g(q_n) = g(F(u_n, q_{n-1})) = 0$$
$$(11)$$

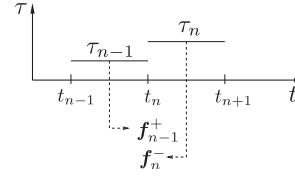The discrete generalized control forces are assumed to be constant in each time interval, see Figure 1.

Figure 1. Relation of redundant forces f_{n-1}^+, f_n^- to discrete generalized forces τ_{n-1}, τ_n.

First of all, the effect of the generalized forces acting in $[t_{n-1}, t_n]$ and in $[t_n, t_{n+1}]$ is transformed to the time node t_n, via $\tau_{n-1}^+ = (h/2) \tau_{n-1}$ and $\tau_n^- = (h/2) \tau_n$. Second, the components of the discrete force vectors $f_{n-1}^+, f_n^- \in T_{q_n}^* Q$ can be calculated similar to (5) as

$$f_{n-1}^+ = B^{\mathrm{T}}(q_n) \cdot \tau_{n-1}^+, \quad f_n^- = B^{\mathrm{T}}(q_n) \cdot \tau_n^-$$
$$f_n = f_n^+ + f_n^-, \quad f_d = \{f_n\}_{n=0}^{N-1}$$
$$(12)$$

Thus, f_{n-1}^+ denotes the effect of the generalized force τ_{n-1} acting in $[t_{n-1}, t_n]$ on q_n, whereas f_n^- denotes the effect on q_n of τ_n acting in $[t_n, t_{n+1}]$.

Insertion of the nodal reparameterization for the configuration (11) into the scheme redundantizes $(10)_2$. The resulting scheme

$$P^{\mathrm{T}}(q_n) \cdot [D_2 L_d(q_{n-1}, q_n) + D_1 L_d(q_n, F(u_{n+1}, q_n))$$

$$+ f_{n-1}^+ + f_n^-] = 0 \qquad (13)$$

has to be solved for u_{n+1}, whereupon q_{n+1} is obtained from (11). Equation (13) is equivalent to the constrained scheme (10), thus it also has the key properties of exact constraint fulfillment, symplecticity and momentum consistency, that is, any change in the value of a momentum map reflects exactly the applied forces as will be shown in Section 7. When no load is present, momentum maps are conserved exactly. While the constrained scheme (10) becomes increasingly ill-conditioned for decreasing time steps, the condition number of (13) is independent of the time step.

Boundary conditions: In the next step, the boundary conditions $q(t_0) = q^0, \dot{q}(t_0) = \dot{q}^0$ and $q(t_N) = q^N, \dot{q}(t_N) = \dot{q}^N$ have to be formulated in the discrete setting. Let $q_{00} \in C$ be a fixed reference configuration, relative to which the initial configuration is computed as $q_0 = F(u_0, q_{00})$. To prescribe an initial configuration

at t_0, one can request $\boldsymbol{u}_0 = \boldsymbol{u}^0$. If an absolute reparameterization is used, that is, Equation (11) is changed to $\boldsymbol{q}_n = \boldsymbol{F}(\boldsymbol{u}_n, \boldsymbol{q}_{00})$, then $\boldsymbol{u}_N = \boldsymbol{u}^N$ defines the final configuration $\boldsymbol{q}_N$ uniquely (see Sections 7.2 and 7.3, for examples). However, for the relative reparameterization (11), $n-m$ independent final configuration conditions have to be identified with the function $\boldsymbol{D} : Q \times Q \to \mathbb{R}^{n-m}$ depending on the specific system under consideration. Since in the present formulation of constrained forced discrete variational dynamics on $Q \times Q$, velocities are not properly defined, the velocity conditions have to be transformed into conditions on the conjugate momenta, which are defined at each and every time node using the discrete Legendre transform. Three different versions have been defined in [38] for the conservative case. Now, the presence of forces at the time nodes has to be incorporated into that transformation leading to the constrained forced discrete Legendre transforms $\mathbb{F}^{\mathrm{cf}^-} L_d : Q \times Q \to T^*Q$ and $\mathbb{F}^{\mathrm{cf}^+} L_d : Q \times Q \to T^*Q$ reading

$$
\begin{aligned}
\mathbb{F}^{\mathrm{cf}^-} L_d \ &: \ (\boldsymbol{q}_n, \boldsymbol{q}_{n+1}) \mapsto (\boldsymbol{q}_n, \boldsymbol{p}_n^-) \\
\boldsymbol{p}_n^- &= -D_1 L_d(\boldsymbol{q}_n, \boldsymbol{q}_{n+1}) + \tfrac{1}{2}\boldsymbol{G}_d^{\mathrm{T}}(\boldsymbol{q}_n) \cdot \boldsymbol{\lambda}_n - \boldsymbol{f}_n^- \\
\mathbb{F}^{\mathrm{cf}^+} L_d \ &: \ (\boldsymbol{q}_{n-1}, \boldsymbol{q}_n) \mapsto (\boldsymbol{q}_n, \boldsymbol{p}_n^+) \\
\boldsymbol{p}_n^+ &= D_2 L_d(\boldsymbol{q}_{n-1}, \boldsymbol{q}_n) - \tfrac{1}{2}\boldsymbol{G}_d^{\mathrm{T}}(\boldsymbol{q}_n) \cdot \boldsymbol{\lambda}_n + \boldsymbol{f}_{n-1}^+
\end{aligned}
\tag{14}
$$

As in the conservative case, the time-stepping scheme $(10)_1$ can be interpreted as matching of momenta $\boldsymbol{p}_n^+ - \boldsymbol{p}_n^- = \boldsymbol{0}$ such that along the discrete trajectory, there is a unique momentum at each time node n, which can be denoted by $\boldsymbol{p}_n$. However, just as the appearance of Lagrange multipliers is avoided in the discrete equations of motion (13), their presence in the initial and final momentum conditions complicates matters unnecessarily. A formula to recover the Lagrange multipliers from the discrete trajectory in a post-processing step can be found in [38]. The following versions of the discrete Legendre transforms do not use Lagrange multipliers. The projected discrete Legendre transforms $^{Q}\mathbb{F}^{\mathrm{cf}^-} L_d : Q \times Q \to \eta(T_{\boldsymbol{q}_n}^* C)$ and $^{Q}\mathbb{F}^{\mathrm{cf}^+} L_d : Q \times Q \to \eta(T_{\boldsymbol{q}_n}^* C)$ read

$$
\begin{aligned}
^{Q}\boldsymbol{p}_n^- &= \boldsymbol{Q}(\boldsymbol{q}_n) \cdot [-D_1 L_d(\boldsymbol{q}_n, \boldsymbol{q}_{n+1}) - \boldsymbol{f}_n^-] \\
^{Q}\boldsymbol{p}_n^+ &= \boldsymbol{Q}(\boldsymbol{q}_n) \cdot [D_2 L_d(\boldsymbol{q}_{n-1}, \boldsymbol{q}_n) + \boldsymbol{f}_{n-1}^+]
\end{aligned}
\tag{15}
$$

where $\boldsymbol{Q}(\boldsymbol{q}_n)$ is given by $\boldsymbol{Q} = \boldsymbol{I}_{n \times n} - \boldsymbol{G}_d^{\mathrm{T}} \cdot [\boldsymbol{G}_d \cdot \boldsymbol{M}^{-1} \cdot \boldsymbol{G}_d^{\mathrm{T}}]^{-1} \boldsymbol{G}_d \cdot \boldsymbol{M}^{-1}$ and fulfills $\boldsymbol{Q}(\boldsymbol{q}_n) \cdot \boldsymbol{G}_d^{\mathrm{T}}(\boldsymbol{q}_n) = \boldsymbol{0}_{n \times m}$. Note that for the constrained discrete Legendre transforms and for the projected discrete Legendre transforms, the output is an n-dimensional momentum vector. In the projected case, it lies in the $(n-m)$-dimensional submanifold $\eta(T_{\boldsymbol{q}_n}^* C)$ being the embedding of $T_{\boldsymbol{q}_n}^* C$ into $T_{\boldsymbol{q}_n}^* Q$. Yet another possibility is to compute an $(n-m)$-dimensional momentum vector by projecting with the discrete null space matrix. The reduced discrete Legendre transforms $^{P}\mathbb{F}^{\mathrm{cf}^-} L_d : Q \times Q \to T^*U$ and $^{P}\mathbb{F}^{\mathrm{cf}^+} L_d : Q \times Q \to T^*U$ are given by

$$
\begin{aligned}
^{P}\boldsymbol{p}_n^- &= \boldsymbol{P}^{\mathrm{T}}(\boldsymbol{q}_n) \cdot [-D_1 L_d(\boldsymbol{q}_n, \boldsymbol{q}_{n+1}) - \boldsymbol{f}_n^-] \\
^{P}\boldsymbol{p}_n^+ &= \boldsymbol{P}^{\mathrm{T}}(\boldsymbol{q}_n) \cdot [D_2 L_d(\boldsymbol{q}_{n-1}, \boldsymbol{q}_n) + \boldsymbol{f}_{n-1}^+]
\end{aligned}
\tag{16}
$$

This version is most appropriate to be used as a constraint in the optimization problem, since it yields the minimal number of independent conditions, while conditions formulated using (15) are redundant and (14) involves the Lagrange multipliers.

Note that according to the range of the projection, $^{Q}\boldsymbol{p}_n$ fulfills the constraints on the momentum level; that is, $\boldsymbol{h}_d(\boldsymbol{q}_n, {}^{Q}\boldsymbol{p}_n) = \boldsymbol{G}(\boldsymbol{q}_n) \cdot \boldsymbol{M}^{-1} \cdot {}^{Q}\boldsymbol{p}_n = \boldsymbol{0}$ while this is not in general the case for $\boldsymbol{p}_n$. This question is superfluous for $^{P}\boldsymbol{p}_n$.

Prescribed initial and final velocities of course should be consistent with the constraints on velocity level. Using the standard continuous Legendre transform $\mathbb{F}L : TC \to T^*C$

$$
\mathbb{F}L : (\boldsymbol{q}, \dot{\boldsymbol{q}}) \mapsto (\boldsymbol{q}, \boldsymbol{p}) = (\boldsymbol{q}, D_2 L(\boldsymbol{q}, \dot{\boldsymbol{q}}))
\tag{17}
$$

yields momenta that are consistent with the constraints on momentum level as well. With these preliminaries, the velocity boundary conditions are transformed to the following conditions on momentum level: $\boldsymbol{p}(t_0) = D_2 L(\boldsymbol{q}(t_0), \dot{\boldsymbol{q}}(t_0)) = \boldsymbol{p}^0$ and $\boldsymbol{p}(t_N) = D_2 L(\boldsymbol{q}(t_N), \dot{\boldsymbol{q}}(t_N)) = \boldsymbol{p}^N$, respectively. Then, $\boldsymbol{p}^0 = \boldsymbol{p}_0^-$ and $\boldsymbol{p}^N = \boldsymbol{p}_N^+$ are the corresponding conditions on the discrete level, which read in detail as

$$
\begin{aligned}
\boldsymbol{P}^{\mathrm{T}}(\boldsymbol{q}_0) \cdot [&D_2 L(\boldsymbol{q}_0, \dot{\boldsymbol{q}}_0) \\
&+ D_1 L_d(\boldsymbol{q}_0, \boldsymbol{q}_1) + \boldsymbol{f}_0^-] = \boldsymbol{0} \\
\boldsymbol{P}^{\mathrm{T}}(\boldsymbol{q}_N) \cdot [&D_2 L(\boldsymbol{q}_N, \dot{\boldsymbol{q}}_N) \\
&- D_2 L_d(\boldsymbol{q}_{N-1}, \boldsymbol{q}_N) - \boldsymbol{f}_{N-1}^+] = \boldsymbol{0}
\end{aligned}
\tag{18}
$$

The discrete constrained optimization problem: Now the optimal control problem for the constrained discrete motion can be formulated. To begin with, an approximation

$$B_d(\boldsymbol{q}_n, \boldsymbol{q}_{n+1}, \boldsymbol{f}_n) \approx \int_{t_n}^{t_{n+1}} B(\boldsymbol{q}, \dot{\boldsymbol{q}}, \boldsymbol{f}) \, \mathrm{d}t \qquad (19)$$

of the continuous objective functional (1) has to be defined. As with the approximations (7), the midpoint rule is used:

$$B_d(\boldsymbol{q}_n, \boldsymbol{q}_{n+1}, \boldsymbol{f}_n) = h B\left(\frac{\boldsymbol{q}_{n+1} + \boldsymbol{q}_n}{2}, \frac{\boldsymbol{q}_{n+1} - \boldsymbol{q}_n}{h}, \boldsymbol{f}_n\right) \quad (20)$$

This yields the discrete objective function

$$J_d(\boldsymbol{q}_d, \boldsymbol{f}_d) = \sum_{n=0}^{N-1} B_d(\boldsymbol{q}_n, \boldsymbol{q}_{n+1}, \boldsymbol{f}_n) \qquad (21)$$

where the discrete configurations and forces are expressed in terms of their corresponding independent generalized quantities $\boldsymbol{u}_d = \{\boldsymbol{u}_n\}_{n=0}^{N}$ and $\boldsymbol{\tau}_d = \{\boldsymbol{\tau}_n\}_{n=0}^{N-1}$, respectively. Alternatively, a new objective function can be formulated directly in the generalized quantities

$$\bar{J}_d(\boldsymbol{u}_d, \boldsymbol{\tau}_d) = \sum_{n=0}^{N-1} \bar{B}_d(\boldsymbol{u}_n, \boldsymbol{u}_{n+1}, \boldsymbol{\tau}_n) \qquad (22)$$

depending on the desired interpretation of the optimization problem. In any case, (21) or (22) has to be minimized with respect to $\boldsymbol{u}_d, \boldsymbol{\tau}_d$ subject to the initial and final configuration constraints $\boldsymbol{u}_0 - \boldsymbol{u}^0 = \boldsymbol{0}$, $\boldsymbol{D}(\boldsymbol{q}_N, \boldsymbol{q}^N) = \boldsymbol{0}$, the initial and final momenta constraints (18), and the discretized dynamics (13).

Accuracy and efficiency: The order of approximation of the discrete Lagrangian and the discrete virtual work given in (7) and (8), respectively, determines the accuracy and the order of convergence of the optimal control method. In general, one uses polynomial approximations to the trajectories and numerical quadrature to approximate the integrals. Then, the order of the discrete Lagrangian and the discrete virtual work is given by the order of the quadrature rule in use. In [25], (Ober-Blöbaum *et al.*, submitted) it is shown for DMOC (unconstrained dynamics) that a discrete Lagrangian and discrete virtual work of order κ lead to

an optimal control scheme of order κ.[‡] That means, the state and control trajectories as well as the Lagrange multipliers resulting from the Pontryagin maximum principle are approximated with an accuracy of $\mathcal{O}(h^\kappa)$. This is in contrast to other schemes, for example, standard Runge–Kutta or collocation methods, where the approximation of the Lagrange multipliers may be of one or more orders less and is a result of the symplectic nature of the underlying discretization.

When accuracy is improved by increasing the number of discretization points, an implementation on configuration level $Q \times Q$ only, rather than on configuration-momentum or configuration-velocity level, leads to a smaller number of optimization parameters and therefore to a smaller number of SQP iterations[§] compared with collocation methods. Therefore, the optimal trajectory for the configuration and the control forces is determined while the corresponding momenta and velocities are reconstructed. Note that the purely configuration-based formulation of DMOC leads to equivalent discrete solution as the configuration-momentum-based collocation resulting from symplectic Runge–Kutta methods, see (Ober-Blöbaum *et al.*, submitted).

In the case of constrained dynamics, the use of the discrete null space method with nodal reparametrization in DMOCC yields a constrained optimization problem of minimal possible dimension: the optimization of (21) or (22) subject to the discretized equations includes $(2N+1)(n-m)$ variables and $(N+3)(n-m)$ constraints. In contrast to that, the optimization problem resulting from the Lagrange multiplier formulation involves $N(2n+m)+n$ unknowns and $(N-1)(n+m)+4n$ constraints (i.e. $(3N+1)m$ more variables and $(N+1)2m$ more constraints). Certainly, this influences the computational costs of the SQP solver substantially. Both formulations, DMOC in generalized coordinates and DMOCC, converge to

---

[‡]Here, smoothness and coercivity of the solution as well as bounded variation of the controls are assumed.

[§]By exploiting the sparse structure of the optimzation problem, the number of SQP iterations grows approximately linearly w.r.t. the number of optimization variables, see (Ober-Blöbaum *et al.*, submitted).

the same solution as shown for a simple example in Section 7.1.

Of course, the SQP solver provides only local solutions that are strongly dependent on the initial guess. Since the focus of this work is on the structure-preserving approximation of the optimal control of constrained motion, the problem of finding global solutions is not pursued here.

4. OPTIMAL CONTROL FOR RIGID-BODY DYNAMICS

A constrained formulation of rigid-body dynamics (see [30, 41]) is used in this work. The time-dependent configuration variable of a rigid body

$$q(t) = [\varphi(t) \ \ d_1(t) \ \ d_2(t) \ \ d_3(t)]^{\mathrm{T}} \in \mathbb{R}^{12} \quad (23)$$

consists of the placement of the center of mass $\varphi \in \mathbb{R}^3$ and the directors $d_I \in \mathbb{R}^3$, $I = 1, 2, 3$, which are constrained to stay orthonormal during the motion, see Figure 2. Of course this is equivalent to specify that the configuration manifold is the Euclidean group, SE(3), which is common in other treatments, such as [39]. These constraints on the directors are called internal constraints, since they represent the underlying kinematic assumptions. Then the body's Euler tensor with respect to the center of mass can be related to the inertia tensor J via $E = \frac{1}{2}(\mathrm{tr}J)I - J$, where I denotes the 3×3 identity matrix. The principal values of the Euler tensor, E_i, together with the body's total mass M_φ are ingredients in the rigid-body's constant symmetric positive-definite mass matrix

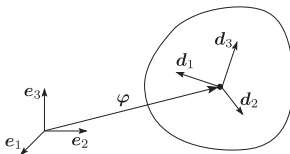$$M = \mathrm{diag}(M_\varphi I \ \ E_1 I \ \ E_2 I \ \ E_3 I) \quad (24)$$

Figure 2. Configuration of a rigid body with respect to an orthonormal frame $\{e_I\}$ fixed in space.

The angular momentum of the rigid body can be computed as

$$L = \varphi \times p_\varphi + d_I \times p_I \quad (25)$$

where summation convention is used to sum over the repeated index I.

Null space matrix: An account of rigid-body dynamics is given in [36, 42] where also the null space matrix

$$P_{\mathrm{int}}(q) = \begin{bmatrix} I & 0 \\ 0 & -\widehat{d}_1 \\ 0 & -\widehat{d}_2 \\ 0 & -\widehat{d}_3 \end{bmatrix} \quad (26)$$

has been derived. Here, $\widehat{a}$ denotes the skew-symmetric 3×3 matrix with corresponding axial vector $a \in \mathbb{R}^3$. For a single rigid-body moving free in space, no external constraints are present; therefore, $P(q) = P_{\mathrm{int}}(q)$.

Nodal reparameterization: When the nodal reparameterization of unknowns is applied, the configuration of the free-rigid body is specified by six unknowns $u_{n+1} = (u_{\varphi_{n+1}}, \theta_{n+1}) \in U \subset \mathbb{R}^3 \times \mathbb{R}^3$, characterizing the incremental displacement and incremental rotation, respectively. Accordingly, in the present case the nodal reparameterization $F : U \times Q \to C$ introduced in (11) assumes the form

$$\begin{aligned} q_{n+1} &= F_d(u_{n+1}, q_n) \\ &= [\varphi_n + u_{\varphi_{n+1}} \\ &\quad \exp(\widehat{\theta_{n+1}}) \cdot (d_1)_n \\ &\quad \exp(\widehat{\theta_{n+1}}) \cdot (d_2)_n \\ &\quad \exp(\widehat{\theta_{n+1}}) \cdot (d_3)_n]^{\mathrm{T}} \end{aligned} \quad (27)$$

where Rodrigues' formula is used to obtain a closed-form expression of the exponential map, see, for example, [39].

Actuation of the rigid body: Consider a single rigid body that is actuated by generalized forces $\tau_{\mathrm{rb}} = [\tau_\varphi \ \ \tau_\theta]^{\mathrm{T}} \in \mathbb{R}^6$ consisting of a translational force $\tau_\varphi \in \mathbb{R}^3$ and a torque $\tau_\theta \in \mathbb{R}^3$. Assume that the force is not applied in the center of mass, but in material points of the rigid body located at $\varrho^{\mathrm{rb}} = \varrho_I^{\mathrm{rb}} d_I$ away from the center of mass. This results in a force τ_φ applied at the center of mass and a torque $\varrho^{\mathrm{rb}} \times \tau_\varphi + \tau_\theta$ that are

given by

$$\begin{bmatrix} \boldsymbol{\tau}_\varphi \\ \boldsymbol{\varrho}^{\mathrm{rb}} \times \boldsymbol{\tau}_\varphi + \boldsymbol{\tau}_\theta \end{bmatrix} = \boldsymbol{C}_{\mathrm{rb}}(\boldsymbol{q}) \cdot \boldsymbol{\tau}_{\mathrm{rb}}$$

$$\boldsymbol{C}_{\mathrm{rb}}(\boldsymbol{q}) = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \widehat{\boldsymbol{\varrho}}^{\mathrm{rb}} & \boldsymbol{I} \end{bmatrix} \in \mathbb{R}^{6 \times 6} \qquad (28)$$

As with (5), the redundant forces can be computed according to

$$\boldsymbol{f} = [\boldsymbol{f}_\varphi \ \boldsymbol{f}_1 \ \boldsymbol{f}_2 \ \boldsymbol{f}_3]^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}(\boldsymbol{q}) \cdot \boldsymbol{\tau}_{\mathrm{rb}} \in \mathbb{R}^{12} \qquad (29)$$

with

$$\boldsymbol{B}^{\mathrm{T}}(\boldsymbol{q}) = \boldsymbol{P}_{\mathrm{int}}(\boldsymbol{q}) \cdot \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \frac{1}{2}\boldsymbol{I} \end{bmatrix} \cdot \boldsymbol{C}_{\mathrm{rb}}(\boldsymbol{q}) \in \mathbb{R}^{12 \times 6} \qquad (30)$$

A straightforward calculation shows

$$\boldsymbol{P}^{\mathrm{T}}(\boldsymbol{q}) \cdot \boldsymbol{f} = \boldsymbol{C}_{\mathrm{rb}}(\boldsymbol{q}) \cdot \boldsymbol{\tau}_{\mathrm{rb}} \qquad (31)$$

The resulting reduced forces in (13) represent the effect of the applied forces and torques $\boldsymbol{\tau}_{\mathrm{rb}}$ on the generalized degrees of freedom.

The same holds in the discrete setting, where the resulting reduced forces in (13)

$$\boldsymbol{P}^{\mathrm{T}}(\boldsymbol{q}_n) \cdot (\boldsymbol{f}_{n-1}^+ + \boldsymbol{f}_n^-) = \boldsymbol{C}_{\mathrm{rb}}(\boldsymbol{q}_n) \cdot (\boldsymbol{\tau}_{\mathrm{rb}_{n-1}}^+ + \boldsymbol{\tau}_{\mathrm{rb}_n}^-) \qquad (32)$$

represent the effect of the applied forces and torques at t_n on the generalized degrees of freedom.

Proposition 4.1
The above definition of the redundant left and right discrete forces guarantees that, in the absence of a potential energy, the change in the angular momentum along the solution trajectory $\boldsymbol{q}_d$ of (13) is induced only by the effect of the discrete generalized forces.

Proof
Computation of $\boldsymbol{p}_{n+1}^+$ and via $\boldsymbol{p}_n^-$ the discrete Legendre transforms (14) and insertion into the definition of angular momentum (25) yield

$$\boldsymbol{L}_{n+1} - \boldsymbol{L}_n = \boldsymbol{\varphi}_{n+1} \times \boldsymbol{p}_{\varphi_{n+1}}^+ + \boldsymbol{d}_{In+1} \times \boldsymbol{p}_{In+1}^+$$

$$- \boldsymbol{\varphi}_n \times \boldsymbol{p}_{\varphi_n}^- - \boldsymbol{d}_{In} \times \boldsymbol{p}_{In}^-$$

$$= \boldsymbol{\varphi}_{n+1} \times (\boldsymbol{f}_{\varphi_n}^+) + \boldsymbol{d}_{In+1} \times (\boldsymbol{f}_{In}^+)$$

$$- \boldsymbol{\varphi}_n \times (-\boldsymbol{f}_{\varphi_n}^-) - \boldsymbol{d}_{In} \times (-\boldsymbol{f}_{In}^-)$$

$$= \boldsymbol{\varphi}_{n+1} \times \boldsymbol{\tau}_{\varphi_n}^+$$

$$+ \boldsymbol{d}_{In+1} \times (\tfrac{1}{2}(\boldsymbol{\varrho}_{n+1}^{\mathrm{rb}} \times \boldsymbol{\tau}_{\varphi_n}^+ + \boldsymbol{\tau}_{\theta_n}^+) \times \boldsymbol{d}_{In+1})$$

$$+ \boldsymbol{\varphi}_n \times \boldsymbol{\tau}_{\varphi_n}^-$$

$$+ \boldsymbol{d}_{In} \times (\tfrac{1}{2}(\boldsymbol{\varrho}_n^{\mathrm{rb}} \times \boldsymbol{\tau}_{\varphi_n}^- + \boldsymbol{\tau}_{\theta_n}^-) \times \boldsymbol{d}_{In})$$

$$= \boldsymbol{\varphi}_{n+1} \times \boldsymbol{\tau}_{\varphi_n}^+ + \boldsymbol{\varrho}_{n+1}^{\mathrm{rb}} \times \boldsymbol{\tau}_{\varphi_n}^+ + \boldsymbol{\tau}_{\theta_n}^+$$

$$+ \boldsymbol{\varphi}_n \times \boldsymbol{\tau}_{\varphi_n}^- + \boldsymbol{\varrho}_n^{\mathrm{rb}} \times \boldsymbol{\tau}_{\varphi_n}^- + \boldsymbol{\tau}_{\theta_n}^- \qquad (33)$$

A straightforward calculation shows that all terms stemming from the kinetic energy and the constraint forces cancel. □

Remark 4.2 (The presence of gravity)
With an acceleration $g \in \mathbb{R}$ due to gravity in the negative $\boldsymbol{e}_3$-direction, the corresponding potential reads $V(\boldsymbol{q}) = [0 \ 0 \ -M_\varphi g \ 0 \ \ldots \ 0] \cdot \boldsymbol{q}$. In this case, Equation (33) yields

$$\boldsymbol{L}_{n+1} - \boldsymbol{L}_n = \boldsymbol{\varphi}_{n+1} \times \boldsymbol{\tau}_{\varphi_n}^+ + \boldsymbol{\varrho}_{n+1}^{\mathrm{rb}} \times \boldsymbol{\tau}_{\varphi_n}^+ + \boldsymbol{\tau}_{\theta_n}^+$$

$$+ \boldsymbol{\varphi}_n \times \boldsymbol{\tau}_{\varphi_n}^- + \boldsymbol{\varrho}_n^{\mathrm{rb}} \times \boldsymbol{\tau}_{\varphi_n}^- + \boldsymbol{\tau}_{\theta_n}^-$$

$$- (\boldsymbol{\varphi}_{n+1} + \boldsymbol{\varphi}_n) \times \frac{h}{2}[0 \ 0 \ -M_\varphi g]^{\mathrm{T}} \qquad (34)$$

meaning that the third component of the angular momentum changes only according to the applied forces, while the change in the first and second component is influenced by gravity as well. In particular, in the absence of any external forces, this shows that the third component of the angular momentum is conserved exactly.

5. OPTIMAL CONTROL FOR KINEMATIC PAIRS

In the sequel, actuating forces being consistent with the specific joint constraints are given and structure preservation of the resulting time-stepping scheme is proved.

The coupling of two neighboring links (body 1 and body 2) by a specific joint J yields $m_{\text{ext}}^{(J)}$ external constraints on the configuration variable

$$q = [q^1 \ q^2]^{\mathrm{T}} \in \mathbb{R}^{24} \qquad (35)$$

consisting of $q^{\alpha}, \alpha = 1, 2$ of the form (23). The degrees of freedom of the relative motion of one body with respect to the other are decreased from 6 to $r^{(J)} = 6 - m_{\text{ext}}^{(J)}$. The location of a specific joint in the αth body is characterized by the coordinates ϱ_i^{α} in the body frame $\{d_I^{\alpha}\}$ as $\varrho^{\alpha} = \varrho_i^{\alpha} d_i^{\alpha}$, for $\alpha = 1, 2$.

Null space matrix: The total null space matrix associated with a kinematic pair can be calculated from

$$P^{(J)}(q) = \begin{bmatrix} P_{\text{int}}(q^1) & 0_{12 \times r^{(J)}} \\ P_{\text{int}}(q^2) \cdot P_{\text{ext}}^{2,(J)}(q) \end{bmatrix} \qquad (36)$$

where the internal null space matrix of each body is given in (26) and the $6 \times (6 + r^{(J)})$ matrix $P_{\text{ext}}^{2,(J)}(q)$ accounts for the coupling induced by a specific joint.

Nodal reparameterization: The redundant coordinates $q \in \mathbb{R}^{24}$ of each kinematic pair may be expressed in terms of $6 + r^{(J)}$ independent generalized coordinates. When using a reparameterization of unknowns in the discrete null space method, relationships of the form

$$q_{n+1} = F^{(J)}(\mu_{n+1}^{(J)}, q_n) \qquad (37)$$

are required, where

$$\mu_{n+1}^{(J)} = (u_{\varphi_{n+1}}^1, \theta_{n+1}^1, \vartheta_{n+1}^{(J)}) \in \mathbb{R}^{6 + r^{(J)}} \qquad (38)$$

consists of a minimal number of incremental unknowns in $[t_n, t_{n+1}]$ for a specific kinematic pair. In (38), $(u_{\varphi_{n+1}}^1, \theta_{n+1}^1) \in \mathbb{R}^3 \times \mathbb{R}^3$ are incremental displacements and rotations, respectively, associated with the first body (see Section 4). Furthermore, $\vartheta_{n+1}^{(J)} \in \mathbb{R}^{r^{(J)}}$ denotes incremental unknowns, which characterize the configuration of the second body relative to the axis (or plane in case of the E pair) of relative motion fixed in the first body. In view of (35), the mapping in (37) may be partitioned according to

$$\begin{aligned} q_{n+1}^1 &= F^1(u_{\varphi_{n+1}}^1, \theta_{n+1}^1, q_n^1) \\ q_{n+1}^2 &= F^{2,(J)}(\mu_{n+1}^{(J)}, q_n) \end{aligned} \qquad (39)$$

where $F^1(u_{\varphi_{n+1}}^1, \theta_{n+1}^1, q_n^1)$ is given by (27) and $F^{2,(J)}(\mu_{n+1}^{(J)}, q_n)$ remains to be specified for each kinematic pair.

Actuation of a kinematic pair: The actuation of kinematic pairs is twofold. First of all, the overall motion of the pair can be influenced by applying translational forces and torques $\tau_{\text{rb}} \in \mathbb{R}^6$ to one of the bodies, say body 1. Any resulting change in the first bodies velocities will be transferred to the second body via the constrained equations of motion. Second, the relative motion of the pair can be influenced. Actuation of the joint connection itself by joint forces $\tau^{(J)} \in \mathbb{R}^{r^{(J)}}$ effects both bodies, where according to 'action equals reaction', the resulting generalized forces on the bodies are equal, but opposite in sign, see, for example, [43]. The dimension of the joint force $\tau^{(J)}$ is determined by the number of relative degrees of freedom $r^{(J)}$ permitted by the specific joint.

Altogether, the generalized forces

$$[\tau_{\text{rb}} \ \tau^{(J)}]^{\mathrm{T}} \in \mathbb{R}^{6 + r^{(J)}} \qquad (40)$$

act on the kinematic pair. The redundant forces can then be computed similar to (5) as

$$f = [f^1 \ f^2]^{\mathrm{T}} = B^{\mathrm{T}}(q) \cdot [\tau_{\text{rb}} \ \tau^{(J)}]^{\mathrm{T}} \in \mathbb{R}^{24} \qquad (41)$$

with the $24 \times (6 + r^{(J)})$ matrix

$$\begin{aligned} B^{\mathrm{T}}(q) &= \begin{bmatrix} P_{\text{int}}(q^1) & 0 \\ 0 & P_{\text{int}}(q^2) \end{bmatrix} \\ &\quad \times \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & \frac{1}{2}I & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & \frac{1}{2}I \end{bmatrix} \\ &\quad \times \begin{bmatrix} C_{\text{rb}}(q^1) & C^{1,(J)}(q) \\ 0 & C^{2,(J)}(q) \end{bmatrix} \end{aligned} \qquad (42)$$

and the $6 \times r^{(J)}$ matrices $C^{\alpha,(J)}(q), \alpha = 1, 2$ being specified according to the specific joint in use.

As with Equation (31), the product of the transposed null space matrix and the redundant forces yields
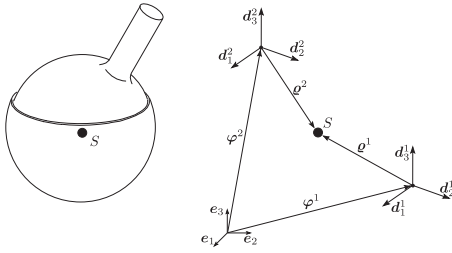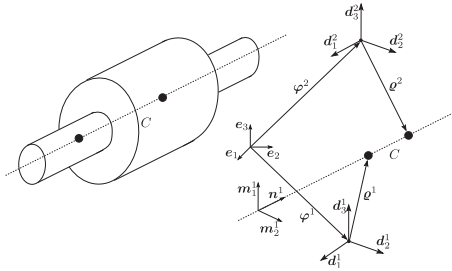
Figure 3. Spherical pair.

Figure 4. Cylindrical pair.

the effect of the generalized forces on the generalized degrees of freedom of the kinematic pair.

5.1. Spherical pair

The S pair (Figure 3) prevents all relative translation between the two bodies. The null space matrix and nodal reparameterization for the S pair are equal to those given in [36, 42].

Actuation of the spherical pair: A torque $\boldsymbol{\tau}^{(S)} \in \mathbb{R}^3$ can be applied at the spherical joint. Then the forces on each body are computed according to (41) with

$$\boldsymbol{C}^{1,(S)}(\boldsymbol{q}) = [\boldsymbol{0} \ -\boldsymbol{I}]^{\mathrm{T}}, \quad \boldsymbol{C}^{2,(S)}(\boldsymbol{q}) = [\boldsymbol{0} \ \boldsymbol{I}]^{\mathrm{T}} \quad (43)$$

The generalized forces affect the following actuation of the generalized degrees of freedom of the spherical pair. In addition to the rigid-body actuation, the first body's rotation is influenced by the joint torque, which also actuates the relative rotation of the second body

$$(\boldsymbol{P}^{(S)}(\boldsymbol{q}))^{\mathrm{T}} \cdot \boldsymbol{f} = [\tau_\varphi \ \ \varrho^{\mathrm{rb}} \times \tau_\varphi + \tau_\theta - \boldsymbol{\tau}^{(S)} \ \ \boldsymbol{\tau}^{(S)}]^{\mathrm{T}} \quad (44)$$

Proposition 5.1
This definition of the redundant left and right discrete forces guarantees that the change of angular momentum

along the solution trajectory $\boldsymbol{q}_d$ of (13) is induced only by the effect of the discrete generalized forces $\boldsymbol{\tau}_{\mathrm{rb}}$. In particular, it is conserved exactly, if the motion of the pair is induced by shape changes only.

Proof
In proving the second statement, it is assumed that only the joint is actuated, that is, $\boldsymbol{\tau}_{\mathrm{rb}} = \boldsymbol{0}$. Computation of $\boldsymbol{p}_{n+1}^+$ and via $\boldsymbol{p}_n^-$ the discrete Legendre transform (14) and insertion into the definition of angular momentum (25) yield

$$\begin{aligned}
\boldsymbol{L}_{n+1} - \boldsymbol{L}_n &= \boldsymbol{\varphi}_{n+1}^1 \times \boldsymbol{p}_{\varphi_{n+1}}^{1\,+} + \boldsymbol{d}_{In+1}^1 \times \boldsymbol{p}_{In+1}^{1+} \\
&\quad + \boldsymbol{\varphi}_{n+1}^2 \times \boldsymbol{p}_{\varphi_{n+1}}^{2\,+} + \boldsymbol{d}_{In+1}^2 \times \boldsymbol{p}_{In+1}^{2+} \\
&\quad - \boldsymbol{\varphi}_n^1 \times \boldsymbol{p}_{\varphi_n}^{1\,-} - \boldsymbol{d}_{In}^1 \times \boldsymbol{p}_{In}^{1\,-} \\
&\quad - \boldsymbol{\varphi}_n^2 \times \boldsymbol{p}_{\varphi_n}^{2\,-} - \boldsymbol{d}_{In}^2 \times \boldsymbol{p}_{In}^{2\,-} \\
&= \boldsymbol{d}_{In+1}^1 \times (-\tfrac{1}{2}\boldsymbol{\tau}_n^{(S)+} \times \boldsymbol{d}_{In+1}^1) \\
&\quad + \boldsymbol{d}_{In+1}^2 \times (\tfrac{1}{2}\boldsymbol{\tau}_n^{(S)+} \times \boldsymbol{d}_{In+1}^2) \\
&\quad + \boldsymbol{d}_{In}^1 \times (-\tfrac{1}{2}\boldsymbol{\tau}_n^{(S)-} \times \boldsymbol{d}_{In}^1) \\
&\quad + \boldsymbol{d}_{In}^2 \times (\tfrac{1}{2}\boldsymbol{\tau}_n^{(S)-} \times \boldsymbol{d}_{In}^2) \\
&= -\boldsymbol{\tau}_n^{(S)+} + \boldsymbol{\tau}_n^{(S)+} - \boldsymbol{\tau}_n^{(S)-} + \boldsymbol{\tau}_n^{(S)-} = \boldsymbol{0} \quad (45)
\end{aligned}$$

The first statement follows by combining (45) and (33). □

5.2. Cylindrical pair

Let $\{\boldsymbol{m}_1^1, \boldsymbol{m}_2^1, \boldsymbol{n}^1\}$ constitute a right-handed orthonormal frame, which is fixed in the first body and specified by $\boldsymbol{n}^1 = n_I^1 \boldsymbol{d}_I^1$ and $\boldsymbol{m}_\kappa^1 = (m_\kappa^1)_I \boldsymbol{d}_I^1$ for $\kappa = 1, 2$. The motion of the second body relative to the axis $\boldsymbol{n}^1$ can be described by $r^{(C)} = 2$ degrees of freedom: translation u^2 along $\boldsymbol{n}^1$ and rotation θ^2 about $\boldsymbol{n}^1$.

Remark 5.2 (Comparison with Betsch and Leyendecker [36] and Leyendecker [42])
The assumption

$$\boldsymbol{\omega}^2 = \boldsymbol{\omega}^1 + \dot{\theta}^2 \boldsymbol{n}^1 \quad (46)$$

used in [36, 42] induces the second body to perform the same rotational motion as the first one and to additionally rotate relative to it about the axis $\boldsymbol{n}^1$. In particular, the second body follows the first bodies rotation about $\boldsymbol{n}^1$ if the relative velocity is zero. For example, a pure rotation of the first body about $\boldsymbol{n}^1$ would affect

The rotation of the second body's directors consists of this rotation and that part of the rotation of body one which is not about the axis $\boldsymbol{n}^1$. Using the notation

$$\boldsymbol{R}^{1,2} = \exp(\widehat{\boldsymbol{\theta}^1_{n+1}}) \cdot \exp(-(\widehat{\boldsymbol{n}^1_n \otimes \boldsymbol{n}^1_n}) \cdot \boldsymbol{\theta}^1_{n+1}) \qquad (49)$$

it may be expressed via the product of exponentials formula

$$\boldsymbol{q}^2_{n+1} = \boldsymbol{F}^{2,(C)}(\boldsymbol{\mu}^{(C)}_{n+1}, \boldsymbol{q}_n) = \begin{bmatrix} \boldsymbol{\varphi}^1_n + \boldsymbol{I}^{11}_n \cdot \boldsymbol{u}^1_{\varphi_{n+1}} + \exp(\widehat{\boldsymbol{\theta}^1_{n+1}}) \cdot [\boldsymbol{\varrho}^1_n + (\bar{u}^2_n + u^2_{n+1})\boldsymbol{n}^1_n] - \boldsymbol{R}^{1,2} \cdot \exp(\theta^2_{n+1}\widehat{\boldsymbol{n}^1_n}) \cdot \boldsymbol{\varrho}^2_n \\ \boldsymbol{R}^{1,2} \cdot \exp(\theta^2_{n+1}\widehat{\boldsymbol{n}^1_n}) \cdot (\boldsymbol{d}^2_1)_n \\ \boldsymbol{R}^{1,2} \cdot \exp(\theta^2_{n+1}\widehat{\boldsymbol{n}^1_n}) \cdot (\boldsymbol{d}^2_2)_n \\ \boldsymbol{R}^{1,2} \cdot \exp(\theta^2_{n+1}\widehat{\boldsymbol{n}^1_n}) \cdot (\boldsymbol{d}^2_3)_n \end{bmatrix} \qquad (50)$$

the second body, but not vice versa according to (46). The new kinematic assumptions, to be introduced in (47), in combination with the new update formula (50) completely decouple the bodies with respect to rotations about and translations along $\boldsymbol{n}^1$. Therefore, it is easier to apply joint actuations that lead to momentum-consistent dynamics.

Null space matrix: Specifically, the new relation between the angular velocities reads

$$\boldsymbol{\omega}^2 = \boldsymbol{I}^{11} \cdot \boldsymbol{\omega}^1 + \dot{\theta}^2 \boldsymbol{n}^1 \qquad (47)$$

It ensures that the translation along and rotation about the axis $\boldsymbol{n}^1$ of one body leaves the other body motionless. With $\boldsymbol{I}^{11} = \boldsymbol{I} - \boldsymbol{n}^1 \otimes \boldsymbol{n}^1$, the null space matrix for the C pair can be inferred from (36) with

$$\boldsymbol{P}_{\text{int}}(\boldsymbol{q}^2) \cdot \boldsymbol{P}^{2,(C)}_{\text{ext}}(\boldsymbol{q})$$

$$= \begin{bmatrix} \boldsymbol{I}^{11} & \widehat{\boldsymbol{\varrho}^2} \cdot \boldsymbol{I}^{11} - \widehat{\boldsymbol{\varrho}^1} - \widehat{u^2 \boldsymbol{n}^1} & \boldsymbol{n}^1 & \boldsymbol{\varrho}^2 \times \boldsymbol{n}^1 \\ \boldsymbol{0} & -\widehat{\boldsymbol{d}^2_1} \cdot \boldsymbol{I}^{11} & \boldsymbol{0} & \boldsymbol{n}^1 \times \boldsymbol{d}^2_1 \\ \boldsymbol{0} & -\widehat{\boldsymbol{d}^2_2} \cdot \boldsymbol{I}^{11} & \boldsymbol{0} & \boldsymbol{n}^1 \times \boldsymbol{d}^2_2 \\ \boldsymbol{0} & -\widehat{\boldsymbol{d}^2_3} \cdot \boldsymbol{I}^{11} & \boldsymbol{0} & \boldsymbol{n}^1 \times \boldsymbol{d}^2_3 \end{bmatrix} \qquad (48)$$

Nodal Reparameterization: For the C pair, the configuration of the second body with respect to the axis $\boldsymbol{n}^1$ can be characterized by $\boldsymbol{\vartheta}^{(C)}_{n+1} = (u^2_{n+1}, \theta^2_{n+1}) \in \mathbb{R}^2$. Here, θ^2_{n+1} accounts for the incremental rotation.

where $\bar{u}^2_n = \boldsymbol{n}^1_n \cdot (\boldsymbol{\varphi}^2_n + \boldsymbol{\varrho}^2_n - \boldsymbol{\varphi}^1_n - \boldsymbol{\varrho}^1_n)$ denotes the translation of the second body relative to the first one in the direction of the axis $\boldsymbol{n}^1_n$ at time t_n.

Actuation of the cylindrical pair: The two relative degrees of freedom allowed by the cylindrical joint can be actuated by a translational force $\tau^{(C)}_\varphi \in \mathbb{R}$ that acts in the direction of the axis $\boldsymbol{n}^1$ and a torque $\tau^{(C)}_\theta \in \mathbb{R}$ about $\boldsymbol{n}^1$. Even if the joint is located away from the centers of mass, translational force along $\boldsymbol{n}^1$ cannot cause a relative rotation of the second body for this pair. However, it causes the pair to rotate according to a torque about $(\boldsymbol{\varphi}^1 - \boldsymbol{\varphi}^2) \times \boldsymbol{n}^1$, which is assigned to the first body. Using the matrices

$$\boldsymbol{C}^{1,(C)}(\boldsymbol{q}) = \begin{bmatrix} -\boldsymbol{n}^1 & \boldsymbol{0} \\ (\boldsymbol{\varphi}^1 - \boldsymbol{\varphi}^2) \times \boldsymbol{n}^1 & -\boldsymbol{n}^1 \end{bmatrix}$$

$$\boldsymbol{C}^{2,(C)}(\boldsymbol{q}) = \begin{bmatrix} \boldsymbol{n}^1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{n}^1 \end{bmatrix} \qquad (51)$$

consistent forces can be computed from (41). The actuation of the generalized degrees of freedom reads

$$(\boldsymbol{P}^{(C)}(\boldsymbol{q}))^{\mathrm{T}} \cdot \boldsymbol{f} = [\boldsymbol{\tau}_\varphi - \tau^{(C)}_\varphi \boldsymbol{n}^1 \quad \boldsymbol{\varrho}^{\mathrm{rb}} \times \boldsymbol{\tau}_\varphi + \boldsymbol{\tau}_\theta - \tau^{(C)}_\theta \boldsymbol{n}^1$$

$$\tau^{(C)}_\varphi \quad \tau^{(C)}_\theta]^{\mathrm{T}} \qquad (52)$$

Proposition 5.1 stating the consistency of angular momentum holds for the C pair and can be proved in exactly the same steps as for the S pair.

5.3. Revolute and prismatic pair

The revolute and the prismatic pair are both special cases of the cylindrical pair, having one of the two

characterized by the incremental variables $\vartheta_{n+1}^{(E)} = (u_{1_{n+1}}^2, u_{2_{n+1}}^2, \theta_{n+1}^2) \in \mathbb{R}^3$. Accordingly, the mapping $q_{n+1}^2 = F_{q_n}^{2,(E)}(\mu_{n+1}^{(E)})$ can be written in the form

$$
F_{q_n}^{2,(E)}(\mu_{n+1}^{(E)}) = \begin{bmatrix} \varphi_n^1 + (I^{12})_n \cdot u_{\varphi_{n+1}}^1 + R^{1,2} \cdot [\varrho_n^1 + ((\bar{u}_\kappa^2)_n + u_{\kappa_{n+1}}^2)(m_\kappa^1)_n] - R^{1,2} \cdot \exp\left(\theta_{n+1}^2 \widehat{n_n^1}\right) \cdot \varrho_n^2 \\ R^{1,2} \cdot \exp\left(\theta_{n+1}^2 \widehat{n_n^1}\right) \cdot (d_1^2)_n \\ R^{1,2} \cdot \exp\left(\theta_{n+1}^2 \widehat{n_n^1}\right) \cdot (d_2^2)_n \\ R^{1,2} \cdot \exp\left(\theta_{n+1}^2 \widehat{n_n^1}\right) \cdot (d_3^2)_n \end{bmatrix} \tag{54}
$$

relative degrees of freedom present in the C pair, respectively. Their null space matrices, nodal reparameterization and actuation, can be inferred from the previous treatment of the C pair by eliminating the translational degree of freedom in case of the R pair and the rotational degree of freedom in case of the P pair.

5.4. Planar pair

For the E pair (Figure 5), the second body may rotate about the axis specified by n^1 and translate in the plane spanned by m_1^1 and m_2^1.

Null space matrix: Using $I^{12} = I - m_1^1 \otimes m_1^1 - m_2^1 \otimes m_2^1$, the relation between the angular velocities reads $\omega^2 = I^{11} \cdot \omega^1 + \dot{\theta}^2 n^1$. With regard to (36), the null space matrix for the E pair is given by

$$
P_{\text{int}}(q^2) \cdot P_{\text{ext}}^{2,(E)}(q)
$$

$$
= \begin{bmatrix} I^{12} & \widehat{\varrho^2} \cdot I^{11} - \widehat{\varrho^1} - \widehat{u_\kappa^2 m_\kappa^1} & m_1^1 & m_2^1 & \varrho^2 \times n^1 \\ 0 & -\widehat{d_1^2} \cdot I^{11} & 0 & 0 & n^1 \times d_1^2 \\ 0 & -\widehat{d_2^2} \cdot I^{11} & 0 & 0 & n^1 \times d_2^2 \\ 0 & -\widehat{d_3^2} \cdot I^{11} & 0 & 0 & n^1 \times d_3^2 \end{bmatrix} \tag{53}
$$

Nodal reparameterization: In the present case the configuration of the second body can be

Here, $(\bar{u}_\kappa^2)_n = (m_\kappa^1)_n \cdot (\varphi 2_n + \varrho 2_n - \varphi_n^1 - \varrho_n^1)$ denotes the translation of the second body relative to the first one in the direction of the axis $(m_\kappa^1)_n$ at time t_n.

Actuation of the planar pair: The three relative degrees of freedom allowed by the planar joint can be actuated by translational forces $\tau_{\varphi_1}^{(E)}, \tau_{\varphi_2}^{(E)} \in \mathbb{R}$ acting in the directions of m_1^1, m_2^1 and a torque $\tau_\theta^{(E)} \in \mathbb{R}$ about n^1. In (41), they are accounted for using

$$
C^{1,(E)}(q) = \begin{bmatrix} -m_1^1 & -m_2^1 & 0 \\ (\varphi_1 - \varphi_2) \times m^1 & (\varphi_1 - \varphi_2) \times m^2 & -n^1 \end{bmatrix}
$$

$$
C^{2,(E)}(q) = \begin{bmatrix} m_1^1 & m_2^1 & 0 \\ 0 & 0 & n^1 \end{bmatrix}
$$

$$
\tag{55}
$$

Similar to (52), the torque induced by a translational joint force away from the center of mass affects the generalized rotational degrees of freedom only with
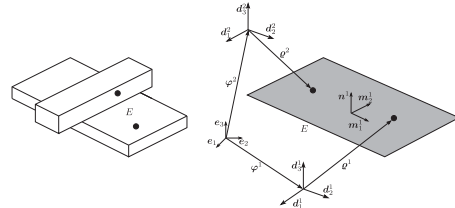

Figure 5. Planar pair.

respect to the allowed rotation around $\boldsymbol{n}^1$

$$(\boldsymbol{P}^{(E)}(\boldsymbol{q}))^{\mathrm{T}}\cdot\boldsymbol{f}=\begin{bmatrix} \boldsymbol{\tau}_\varphi - \tau^{(E)}_{\varphi_1}\boldsymbol{m}^1_1 - \tau^{(E)}_{\varphi_2}\boldsymbol{m}^1_2 \\ \varrho^{\mathrm{rb}}\times\boldsymbol{\tau}_\varphi + \boldsymbol{\tau}_\theta + (\boldsymbol{n}^1)^{\mathrm{T}}\cdot(\varrho^2\times(\tau^{(E)}_{\varphi_1}\boldsymbol{m}^1_1 + \tau^{(E)}_{\varphi_2}\boldsymbol{m}^1_2))\cdot\boldsymbol{n}^1 - \tau^{(E)}_\theta\boldsymbol{n}^1 \\ \tau^{(E)}_{\varphi_1} \\ \tau^{(E)}_{\varphi_2} \\ -(\boldsymbol{n}^1)^{\mathrm{T}}\cdot(\varrho^2\times(\tau^{(E)}_{\varphi_1}\boldsymbol{m}^1_1 + \tau^{(E)}_{\varphi_2}\boldsymbol{m}^1_2)) + \tau^{(E)}_\theta \end{bmatrix} \qquad (56)$$

Again, Proposition 5.1 stating the consistency of angular momentum holds and the proof is straightforward.

6. OPTIMAL CONTROL FOR MULTIBODY SYSTEMS

In a kinematic chain or tree-structured system, where N bodies are interconnected by $N-1$ joints, the multibody system consists of $N-1$ pairs. The configuration variable

$$\boldsymbol{q}=[\boldsymbol{q}^1 \ \ldots \ \boldsymbol{q}^N]^{\mathrm{T}}\in\mathbb{R}^{12N} \qquad (57)$$

is a generalization of (35).

Actuation of a multibody system: As a generalization of (40), the forces and torques acting on the multibody system can be collected in

$$[\boldsymbol{\tau}_{\mathrm{rb}} \ \boldsymbol{\tau}^{(J_1)} \ \ldots \ \boldsymbol{\tau}^{(J_{N-1})}]^{\mathrm{T}}\in\mathbb{R}^{6+\sum_{i=1}^{N-1} r^{(J_i)}} \qquad (58)$$

The redundant forces for each body can be computed as

$$\begin{aligned} \boldsymbol{f}&=[\boldsymbol{f}^1 \ \ldots \ \boldsymbol{f}^N]^{\mathrm{T}} \\ &=\boldsymbol{B}^{\mathrm{T}}(\boldsymbol{q})\cdot[\boldsymbol{\tau}_{\mathrm{rb}}\boldsymbol{\tau}^{(J_1)} \ \ldots \ \boldsymbol{\tau}^{(J_{N-1})}]^{\mathrm{T}}\in\mathbb{R}^{12N} \quad (59) \end{aligned}$$

with the $12N\times(6+\sum_{i=1}^{N-1} r^{(J_i)})$ matrix $\boldsymbol{B}^{\mathrm{T}}(\boldsymbol{q})$ being the product of three matrices as in (42). The first matrix corresponds to the internal constraints of each body and consists of N blocks $\boldsymbol{P}_{\mathrm{int}}(\boldsymbol{q}^\alpha)$, $\alpha=1,\ldots,N$. The second $6N\times 6N$ diagonal matrix is an obvious extension of the one given in (42) consisting of multiples of the identity matrix. The third matrix is given by concatenating

columnwise the matrices $\boldsymbol{C}_{\mathrm{rb}}(\boldsymbol{q}^1)$ and $\boldsymbol{C}^{1,(J_i)},\boldsymbol{C}^{2,(J_i)}$ for each joint $J_i, i=1,\ldots,N-1$

into a $6N\times(6+\sum_{i=1}^{N-1} r^{(J_i)})$ matrix. The first $6N\times 6$ column consists of $\boldsymbol{C}_{\mathrm{rb}}(\boldsymbol{q}^1)$ and a zero matrix below. In the following, $6N\times r^{(J_i)}$ columns, $\boldsymbol{C}^{1,(J_i)}(\boldsymbol{q})$ and $\boldsymbol{C}^{2,(J_i)}(\boldsymbol{q})$ occur in the rows corresponding to the forces $\boldsymbol{f}^\alpha$ and $\boldsymbol{f}^\beta$ (if the αth and βth body are connected by the joint J_i), respectively. See Sections 7.2 and 7.3 for examples of this matrix in the context of a tree-structured multibody system and a kinematic chain.

7. NUMERICAL EXAMPLES

For all numerical examples, we apply the midpoint rule to approximate the relevant integrals leading to second-order optimal control schemes. The solution of the resulting restricted optimization problem is solved using a sparse SQP optimization algorithm based on SNOPT (see [14] for details) that is implemented in the routine nag_opt_nlp_sparse of the NAG library.[¶]

7.1. Two-link pendulum

As a first numerical example, the optimal control of a two-link pendulum is considered that was already investigated in [25], (Ober-Blöbaum *et al.*, submitted). In the mentioned work, the system was directly formulated in generalized coordinates, which is easily possible for this two-dimensional problem. For this system, a comparison between DMOC and a collocation method
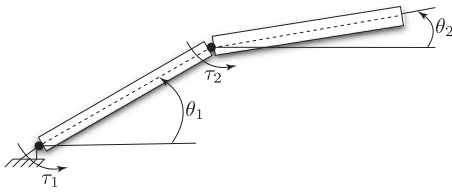

[¶]www.nag.com.

Figure 6. Model of a two-link pendulum.

was performed with respect to convergence rates, the consistency of momentum maps, and the number of iterations executed by the SQP solver. In this contribution, the numerical example is completed by a comparison with DMOCC.

Setup and problem statement: The two-link pendulum in Figure 6 consists of two-coupled planar rigid bodies, of which one is fixed in space. θ_i, $i=1,2$, denotes the orientation of the ith link measured counterclockwise from the positive horizontal axis. The configuration of the system is specified by $q=(\theta_1, \theta_2)$.

The control torques τ_1, τ_2 are applied at the base of the first link and at the joint between the two links. The two-link pendulum is to be steered from the stable equilibrium point $q^0=(-\pi/2, -\pi/2)$ with zero angular velocity $\dot{q}^0=(0,0)$ to the unstable equilibrium point $q^T=(\pi/2, \pi/2)$ with velocity $\dot{q}^T=(0,0)$. For the motion of the pendulum, the control effort

$$J(\tau_1, \tau_2) = \int_0^T \frac{1}{2}(\tau_1^2(t) + \tau_2^2(t))\, \mathrm{d}t \qquad (60)$$

is to be minimized, where the final time $T=1$ is fixed.

Problem solution: Three different methodologies are compared. First, a collocation method of order 2 is applied to two different models: the differential equation systems resulting from the Hamiltonian system with generalized coordinates (q, p) as well as the system formulated on the tangent space with generalized coordinates $(q, \dot{q})=(q, v)$. Second, DMOC in generalized coordinates, and third, DMOCC is performed. To insure that the same local minimum is found by the SQP solver for different numbers of discretization points, the reference solution with $N=512$ is computed first, while initial guesses for

coarser discretizations are extracted from the reference trajectory. Note, that, as already described in (Ober-Blöbaum *et al.*; submitted), identical solutions are obtained for DMOC and the equivalent collocation method for the Hamiltonian formulation.

In Figure 7 (left), the convergence rates for the configurations u_d and the control torques τ_d are depicted. Here, for each method a reference trajectory has been created with $N=512$ discretizations points and time step $h=1.9\times10^{-3}$, and the maximum norm is used to compute the error. For all methods the convergence rate for the configuration and control trajectory is quadratic. Note that the convergence rate for DMOCC is slightly better than for DMOC in generalized coordinates. For all methods also the objective function value converges quadratically as shown in Figure 7 (right). In Figure 8 (left) the convergence of the DMOCC solution to the DMOC solution in generalized coordinates is depicted. Indeed, both methods converge to the same discrete solution.

In Figure 8 (right), the consistency in the momentum map for all methods is shown. For the solution resulting from DMOC, the collocation approach applied to the Hamiltonian system and DMOCC, the change in angular momentum exactly equals the sum of the applied control forces (to numerical accuracy). These results are consistent with the well-known conservation properties of variational integrators, which provide discretizations that preserve continuous properties such as momentum maps in the discrete setting in a natural way. On the other hand, the collocation method applied to the tangent space system described in velocities fails to capture the change in angular momentum accurately because the discrete tangent space formulation destroys the discrete Hamiltonian structure and the resulting scheme is not momentum-preserving anymore.

The methods based on a formulation in generalized coordinates behave equally well compared with DMOCC for the two-dimensional example under consideration. However, for the optimal control of more complex three-dimensional systems as the examples described next, a discrete formulation in generalized coordinates becomes cumbersome and the structure-preserving discretizate equations of motion are far more complicated.
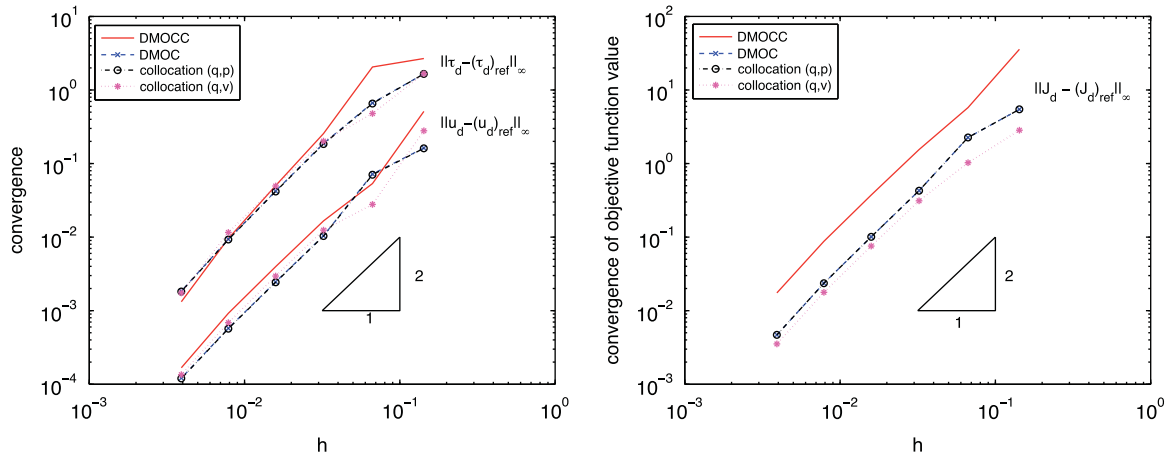
Figure 7. Two-link pendulum. Comparison of the convergence rates for DMOCC, DMOC, and a collocation approach. Left: Configuration and control. Right: Objective function value.
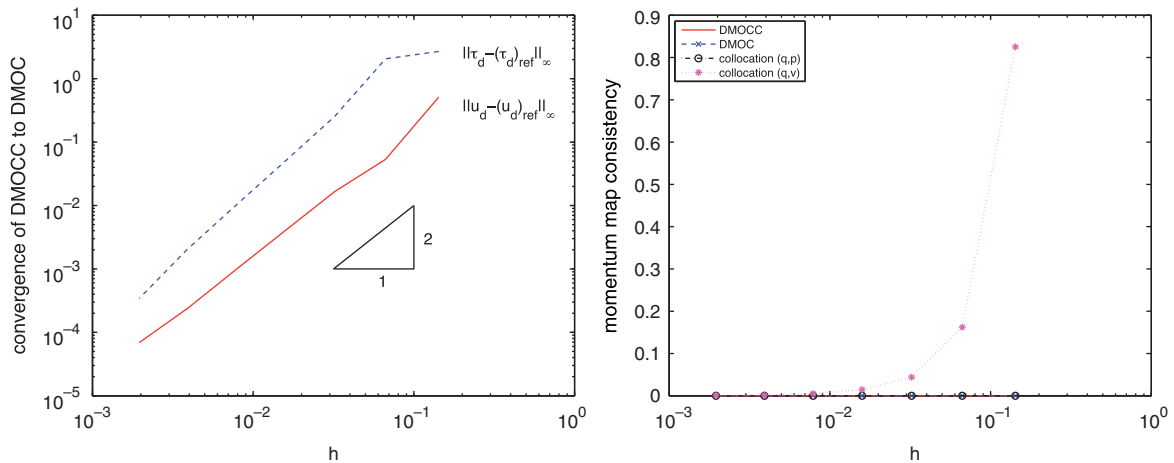

Figure 8. Two-link pendulum. Left: Convergence of DMOCC solution to DMOC solution in generalized coordinates. Right: Momentum map consistency.

7.2. *Optimal control of a rigid body with rotors*

Fully actuated case—setup and problem statement. Inspired by space telescopes such as the Hubble telescope, whose change in orientation is induced by spinning rotors, a multibody system consisting of a main body to which rotors are connected by revolute joints has been analyzed. The revolute joints allow each rotor to rotate relative to the main body around an axis through its center, which is fixed in the main body and

are actuated by torques $\tau^{(R_1)}$, $\tau^{(R_2)}$, $\tau^{(R_3)} \in \mathbb{R}$. No other force and torque are applied to this tree-structured system; therefore, in (42), the last matrix reduces to

$$
\begin{bmatrix}
C^{1,(R_1)}(q) & C^{1,(R_2)}(q) & C^{1,(R_3)}(q) \\
C^{2,(R_1)}(q) & 0 & 0 \\
0 & C^{2,(R_2)}(q) & 0 \\
0 & 0 & C^{2,(R_3)}(q)
\end{bmatrix}
\tag{61}
$$

which is then used in (59) to compute the redundant forces on each body.

The goal is to determine optimal torques to guide the main body from the initial orientation $\boldsymbol{u}_\theta^0 = [0, 0, 0]$ into the final position $\boldsymbol{u}_\theta^N = (\pi/14)[1, 2, 3]$, where the absolute reparametrization $\boldsymbol{q}_n = \boldsymbol{F}(\boldsymbol{u}_n, \boldsymbol{q}_{00})$ is used instead of (11) here. The motion starts and ends at rest. The maneuver time is $T = 5$ and the time step is $h = 0.1$, thus $N = 50$. As in the first example, the objective function represents the control effort that has to be minimized. Owing to the presence of three rotors with nonplanar axes of rotation, this problem is fully actuated.

Fully actuated case—problem solution: The employed initial guess does not fulfill the discrete equations of motion, the main body rotates uniformly into the final orientation, while the rotors do not move and the control torques are set to zero. Figure 9 shows the configuration of the system at $t = 0, 1, \ldots, 5$. The static frame represents the required final orientation where the axes must coincide with the centers of the rotors as the motion ends (see last picture). The optimal torques, which are constant in each time interval, are depicted in Figure 10. They yield a control effort of $\bar{J}_d = 2.8242 \times 10^6$. Finally, Figure 10 illustrates the evolution of the kinetic energy and a special attribute of the system under consideration. Owing to a geometric phase, the motion occurs although the total angular momentum remains zero at all times. As shown in Proposition 5.1, the algorithm is able to represent this correctly.

Remark 7.1 (Dimension of the constrained optimization problem)
Using the advocated method, the problem consists of 909 unknowns and 477 constraints, rather than 6798 variables and 4555 constraints using the Lagrange multiplier formulation.

Underactuated case: The same rest to rest maneuver is investigated for the underactuated system where one momentum wheel has been removed. The fully actuated maneuver serves as an initial guess. The reorientation maneuver depicted in Figure 11 requires only slightly more control effort $\bar{J}_d = 2.9168 \times 10^6$ than the fully actuated case.

Consistency of angular momentum is observable from Figure 12. It also shows that the energy does not evolve as symmetrically as for the fully actuated problem. That means that acceleration phase and braking phase are not exactly inverse to each other. This becomes also obvious from Figure 12 showing the evolution of the optimal generalized forces.

7.3. Optimal control of a pitcher's motion

As an example of biomotion in sports, the optimal pitch of an athlete is investigated in this section. For simplicity, a kinematic chain representing the pitcher's arm is considered including the collarbone, the upper, and the forearm (see Figure 13). Here, the control torques in the joints represent the muscle activation.

Setup and problem statement: The first rigid body, representing the collarbone, is assumed to be fixed in the inertial frame via a revolute joint modelling the rotation of the torso around the $\boldsymbol{e}_3$-axis, thus the axis of the first revolute joint is $\boldsymbol{n}^1 = \boldsymbol{e}_3$. Collarbone and upper arm are connected via a spherical joint, representing the three-dimensional rotation of the shoulder. A revolute joint serves as the elbow between the upper and the forearm allowing the forearm to rotate around a prescribed axis $\boldsymbol{n}^2$ fixed in the upper arm.

It is assumed that all degrees of freedom, that is the rotations of the collarbone $\theta^1 \in \mathbb{R}$, the shoulder θ^2, θ^3, $\theta^4 \in \mathbb{R}$, and the elbow $\theta^5 \in \mathbb{R}$, are directly steerable. There is a rotational torque $\boldsymbol{\tau}^{(S)} \in \mathbb{R}^3$ acting in the shoulder joint and two scalar torques $\tau^{(R_1)}$, $\tau^{(R_2)} \in \mathbb{R}$ acting in the first revolute joint and the elbow joint, respectively. Since the first body is fixed in space, the last matrix in (42) reduces to

$$\begin{bmatrix} \boldsymbol{C}^{2,(R_1)}(\boldsymbol{q}) & \boldsymbol{C}^{1,(S)}(\boldsymbol{q}) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C}^{2,(S)}(\boldsymbol{q}) & \boldsymbol{C}^{1,(R_2)}(\boldsymbol{q}) \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{C}^{2,(R_2)}(\boldsymbol{q}) \end{bmatrix} \quad (62)$$

Remark 7.2
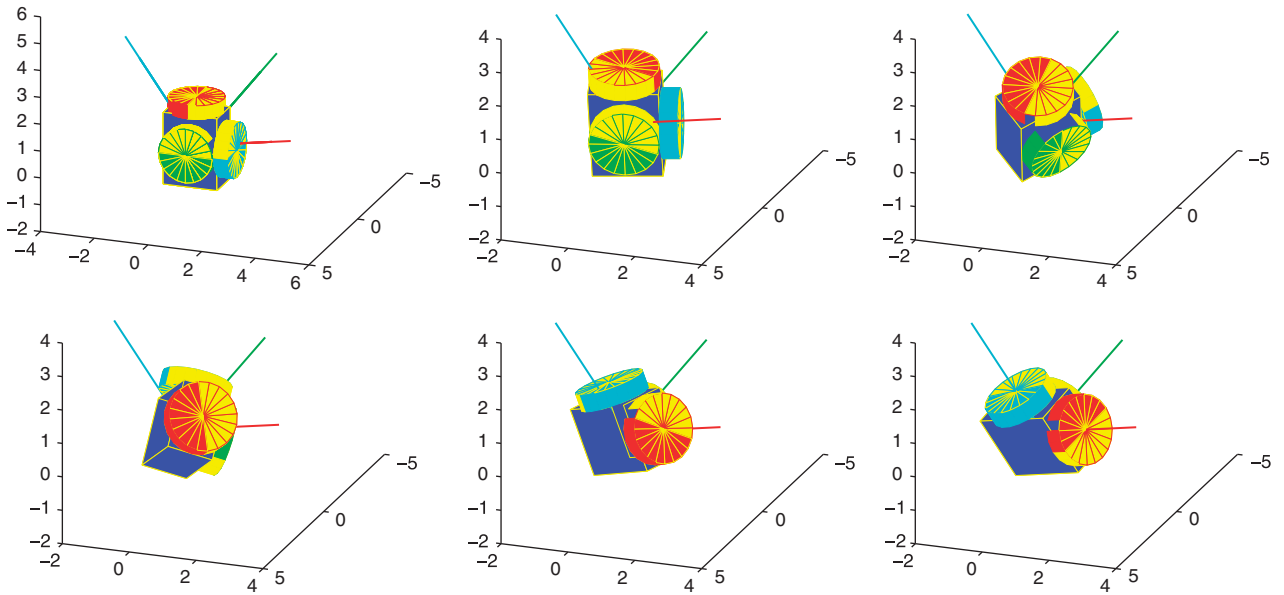For the pitcher, the effect of the actuating torques in the joints on the generalized degrees of freedom takes

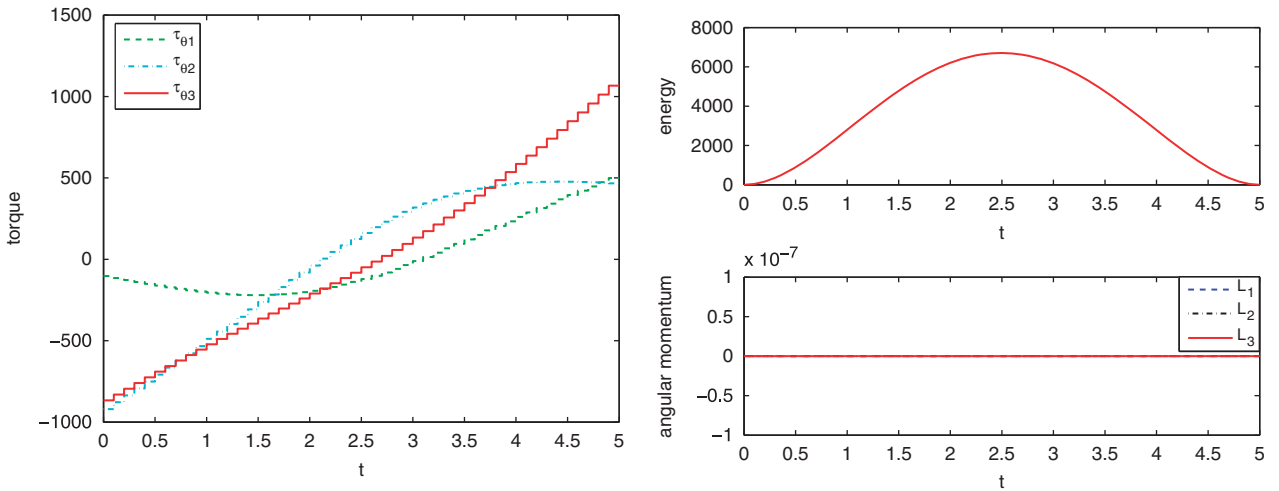Figure 9. Rigid body with three rotors. Configuration at $t = 10nh, n = 0, \ldots, 5$ ($h = 0.1$).

Figure 10. Rigid body with three rotors. Left: Torque (N m). Right: Energy (J) and components of angular momentum vector $\boldsymbol{L} = L_I \boldsymbol{e}_I$ (N ms) ($h = 0.1$).

the form

$$\boldsymbol{P}^{\mathrm{T}}(\boldsymbol{q}) \cdot \boldsymbol{f}$$
$$= [\tau^{(R_1)} - (\boldsymbol{n}^1)^{\mathrm{T}} \cdot \boldsymbol{\tau}^{(S)} \quad \boldsymbol{\tau}^{(S)} - \tau^{(R_2)} \boldsymbol{n}^2 \quad \tau^{(R_2)}]^{\mathrm{T}} \quad (63)$$

The control torque $\boldsymbol{\tau}^{(S)}$ in the spherical joint acts with different signs on the collarbone and on the upper arm. Only $(\boldsymbol{n}^1)^{\mathrm{T}} \cdot \boldsymbol{\tau}^{(S)}$, the part of $\boldsymbol{\tau}^{(S)}$ in the direction of $\boldsymbol{n}^1$, influences the collarbone's rotation, since the collarbone is constrained to perform rotational motion
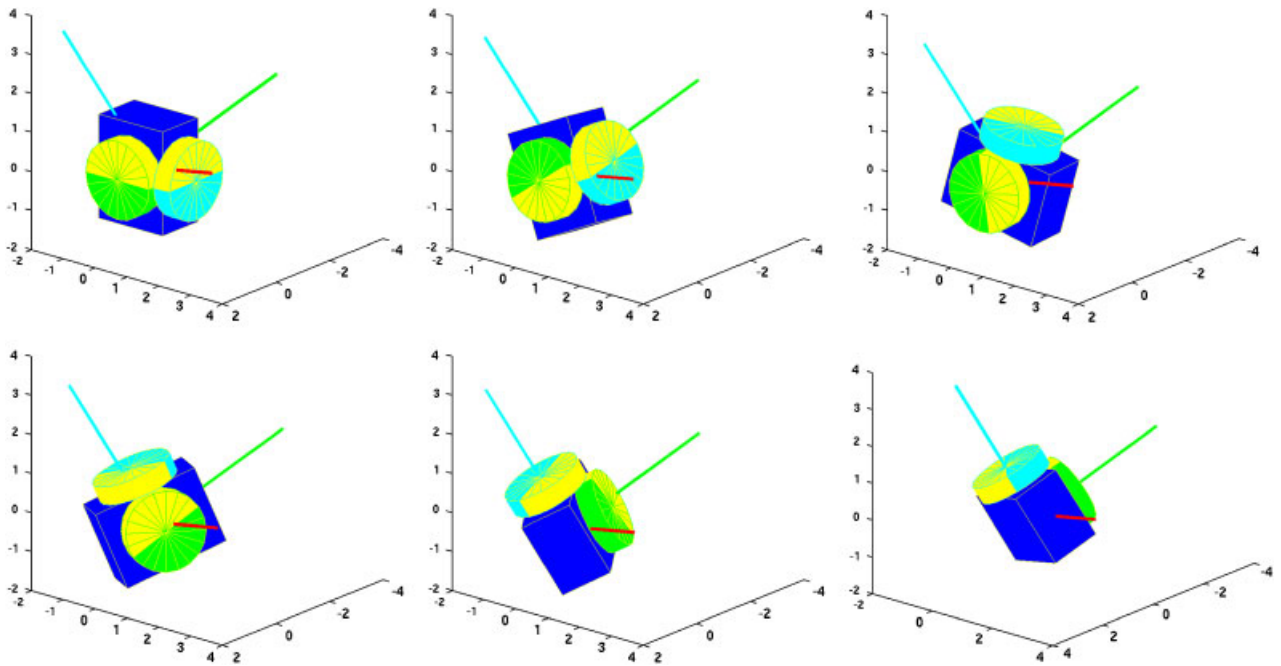
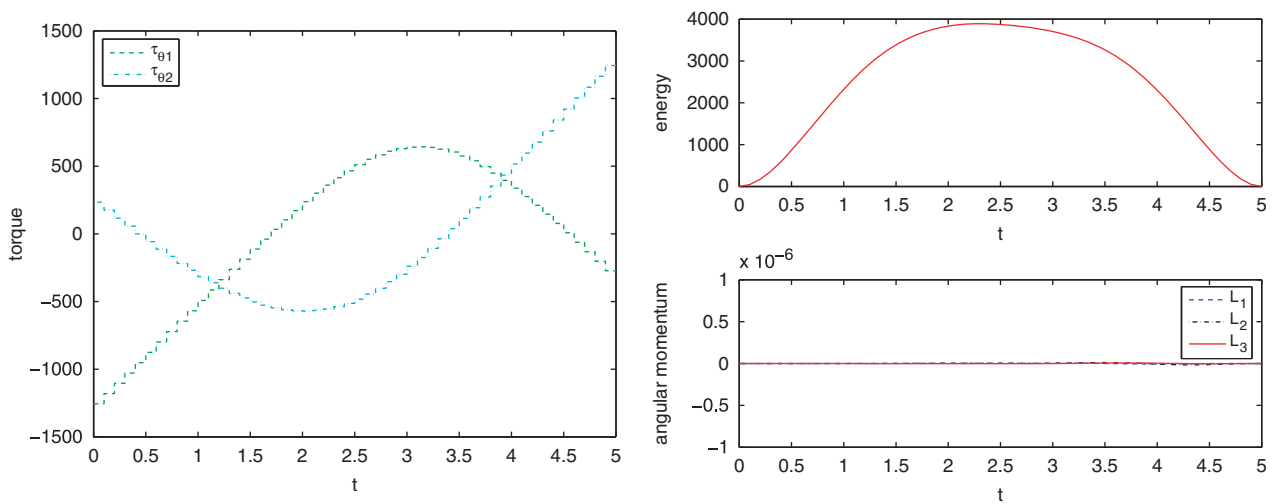Figure 11. Rigid body with two rotors. Configuration at $t = 10nh, n = 0, \ldots, 5$ ($h = 0.1$).

Figure 12. Rigid body with two rotors. Left: Torque (N m). Right: Energy (J) and components of angular momentum vector $\boldsymbol{L} = L_I \boldsymbol{e}_I$ (N ms) ($h = 0.1$).
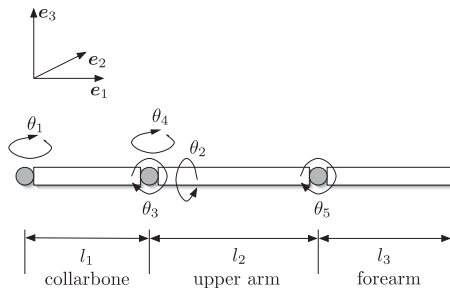
Figure 13. The optimal pitch. Model for the arm consisting
of collarbone, upper arm, and forearm.

around $\boldsymbol{n}^1$ only. Similarly, $\tau^{(R_2)}$ acts on the upper and forearm with different signs; therefore, it influences the three generalized degrees of freedom in the shoulder by $\tau^{(R_2)}\boldsymbol{n}^2$.

The pitcher is assumed to begin the motion with prescribed initial configuration and zero velocity. Rather than prescribing final configurations for all present bodies, a limited but not fixed final configuration is defined for the hand position‖, for example, positive $\boldsymbol{e}_2$- and $\boldsymbol{e}_3$-position. Owing to the human body's anatomy, the relative motion in each joint is limited. To obtain a realistic motion, each generalized configuration variable is bounded, for example, the forearm is assumed to bend in only one direction. In addition, the incorporation of bounds on the control torques is needed, since the muscles are not able to create an arbitrary amount of strength.

The goal is to maximize the final momentum of the hand in $\boldsymbol{e}_2$-direction. More specifically, the projected discrete Legendre transform (15) is used to compute the discrete objective function $J_d(\boldsymbol{q}_d,\boldsymbol{f}_d)= -\boldsymbol{e}_2^{\mathrm{T}}\cdot({}^Q\boldsymbol{p}_N^+)_{\mathrm{hand}}$. During the optimization the final time is free, which also means that the optimal duration of the pitch is determined as a variable.

Problem solution: The number of time nodes has been set to $N=35$. Starting from an initial position of the joints as $\theta_0^1=\theta_0^2=\theta_0^3=\theta_0^4=0$, $\theta_0^5=-\pi/4$, different local solutions for the optimal motion are obtained,

depending on the initial guess in use, where initial guesses were constructed by interpolating coarse pitch sequences and setting the control parameters to zero. In Figure 14 snapshots of a particular locally optimal motion are depicted with the optimal final time $T=0.41545$, that is, $h=0.01187$. Starting from the initial configuration shown in the first picture, the pitcher strikes his arm out, moves it rearwards, pulls it above his head, before he finally moves his arm like a whip to obtain the necessary swing to maximize the final momentum $J_d=24.502$. The evolution of discrete generalized coordinates and torques can be observed from Figure 15. Figure 16 illustrates the consistency of angular momentum. Owing to the presence of gravity and the fixing of the chain in space by a revolute joint with rotation axis $\boldsymbol{e}_3$, the only symmetry of the augmented discrete Lagrangian (6) is rotation about $\boldsymbol{e}_3$. Therefore, the corresponding component of the angular momentum L_3 changes exactly according to the torque $\tau^{(R_1)}$, applied in the supporting joint. The kinetic energy, which is increasing substantially towards the end of the movement, is shown in Figure 16.

Remark 7.3 (*Dimension of the constrained optimization problem*)
Including the free final time, the number of variables is 356 and the initial conditions and dynamic constraints sum up to 190. In the Lagrange multiplier formulation, one is faced with 3641 variables and 2422 constraints.

The next step is to consider more complex models that behave more realistically. For example, the interaction of the muscles and the resulting muscle force can be included, see [44]. Owing to the constrained formulation of multibody dynamics, model extensions can easily be incorporated by coupling new bodies to the system via constraints.

8. CONCLUSION

This paper proposes a new approach, DMOCC, to the solution of optimal control problems for constrained mechanical systems via the combination of two recently
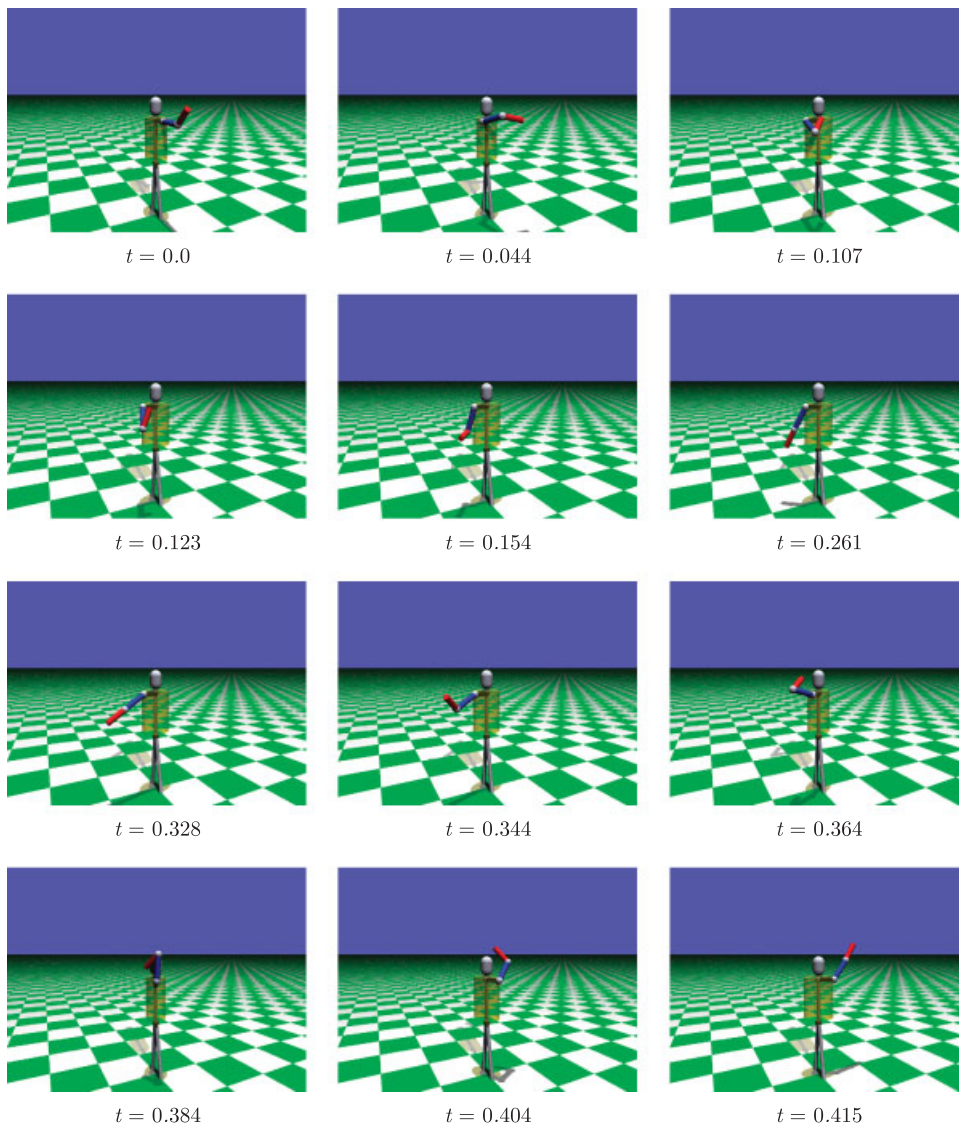

‖Since the hand is not modelled as a separate rigid body within the system, it is assumed to be located at the endpoint of the forearm.

Figure 14. The optimal pitch. Snapshots of the motion sequence ($h = 0.01187$).

developed methods: the discrete null space method and the optimal control method DMOC.

DMOCC is used to compute trajectories for a mechanical system that is optimally guided from an initial to a final configuration via external forces. The given objective function is extremized subject to the reduced discrete dynamic equations of the constrained mechanical system.

The proposed method benefits from an easy derivation and implementation of the constraint equation for the optimization algorithm and ensures exact constraint fulfillment and structure-preserving properties of the
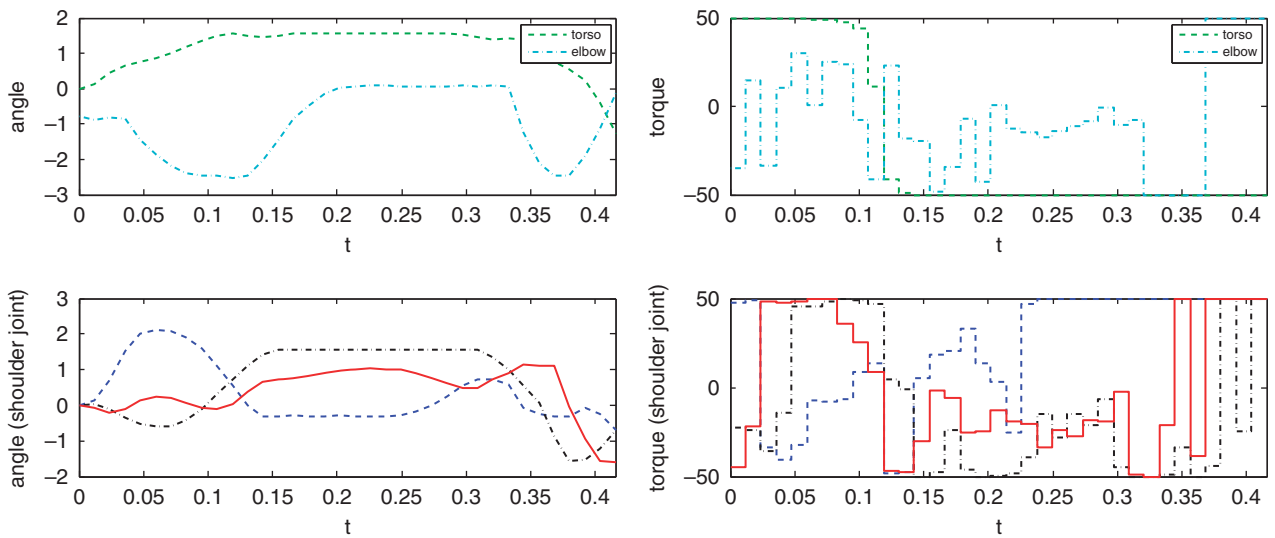
Figure 15. The optimal pitch. Left: Evolution of discrete generalized coordinates (rad). Right: Torque (N m) ($h = 0.01187$).
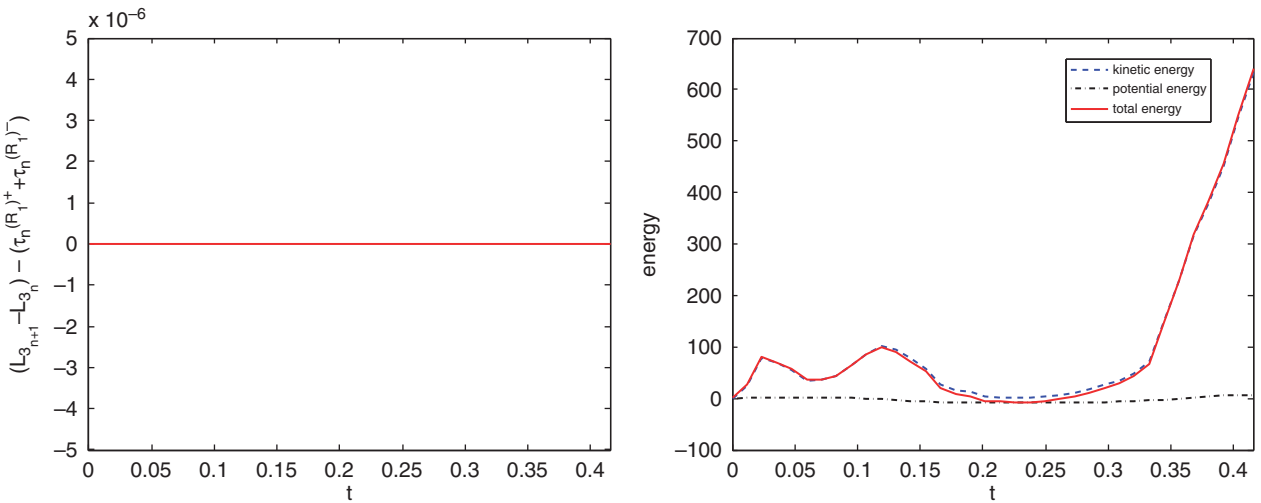

Figure 16. The optimal pitch. Left: Consistency of angular momentum (N m). Right: Evolution of kinetic, potential, and total energy (J) ($h = 0.01187$).

computed solutions. In particular, actuating forces being consistent with the specific joint constraints are given and angular momentum consistency of the resulting time-stepping scheme is proved analytically and verified numerically with a satellite reorientation problem and the optimization of a pitcher's motion.

REFERENCES

1. Leyendecker S, Ober-Blöbaum S, Marsden JE, Ortiz M. Discrete mechanics and optimal control for constrained multibody dynamics. *Proceedings of the 6th International Conference on Multibody Systems*, *Nonlinear Dynamics*, *and Control*, *ASME International Design Engineering Technical Conferences*, Las Vegas, Nevada, 2007.

2. Bertsekas D. *Nonlinear Programming*. Athena Scientific: Nashua, 1995.

3. Luenberger D. *Linear and Nonlinear Programming*. Addison-Wesley: Reading, MA, 1984.

4. Leyendecker S, Betsch P, Steinmann P. Energy-conserving integration of constrained Hamiltonian systems—a comparison of approaches. *Computational Mechanics* 2003; **33**:174–185.

5. Benzi M, Golub G, Liesen J. Numerical solution of saddle point problems. *Acta Numerica* 2005; **1**:1–137.

6. Betsch P, Steinmann P. Conserving properties of a time FE method—Part III: mechanical systems with holonomic constraints. *International Journal for Numerical Methods in Engineering* 2002; **53**:2271–2304.

7. Gonzalez O. Mechanical systems subject to holonomic constraints: differential-algebraic formulations and conservative integration. *Physica D* 1999; **132**:165–174.

8. Wendlandt J, Marsden JE. Mechanical integrators derived from a discrete variational principle. *Physica D* 1997; **106**:223–246.

9. Petzold L, Lostedt P. Numerical solution of nonlinear differential equations with algebraic constraints II: practical implications. *SIAM Journal on Scientific Computing* 1986; **7**:720–733.

10. Hairer E, Lubich C, Roche M. *The Numerical Solution of Differential Algebraic Equations by Runge–Kutta Methods*. Springer: Berlin, 1989.

11. Leimkuhler B, Reich S. *Simulating Hamiltonian Dynamics*. Cambridge University Press: Cambridge, 2004.

12. Rheinboldt W. Solving algebraically explicit DAEs with the MANPAK-manifold-algorithms. *Computers and Mathematics with Applications* 1997; **33**:31–43.

13. Betsch P. The discrete null space method for the energy consistent integration of constrained mechanical systems. Part I: holonomic constraints. *Computer Methods in Applied Mechanics and Engineering* 2005; **194**:5159–5190.

14. Gill P, Murray W, Saunders M. SNOPT: an SQP algorithm for large-scale constrained optimization. *Numerical Analysis Report 97-2*, Department of Mathematics, University of California, San Diego, La Jolla, CA, 1997.

15. Schittkowski K. *Nonlinear Programming Codes*. Lecture Notes in Economics and Mathematical Systems, vol. 183. Springer: Berlin, 1980; viii+242.

16. Stoer J, Bulirsch R. *Introduction to Numerical Analysis*, vol. 12. Springer: Berlin, 2002; xvi+744.

17. Kraft D. On converting optimal control problems into nonlinear programming problems. *Computational Mathematical Programming* 1985; **F15**:261–280.

18. Hicks G, Ray W. Approximation methods for optimal control systems. *Canadian Journal of Chemical Engineering* 1971; **49**:522–528.

19. Deuflhard P. A modified Newton method for the solution of ill-conditioned systems of nonlinear equations with application to multiple shooting. *Numerische Mathematik* 1974; **22**: 289–315.

20. Bock H, Plitt K. A multiple shooting algorithm for direct solution of optimal control problems. *Proceedings of the 9th IFAC World Congress*, Budapest, 1984; 243–247.

21. von Stryk O. Numerical solution of optimal control problems by direct collocation. *Optimal Control* (Freiburg, 1991). International Series in Numerical Mathematics, vol. 111. Birkhauser: Basel, 1993; 129–143.

22. Biegler L. Solution of dynamic optimization problems by successive quadratic programming and orthogonal collocation. *Computers and Chemical Engineering* 1984; **8**:243–248.

23. Betts J. Survey of numerical methods for trajectory optimization. *Journal of Guidance*, *Control*, *and Dynamics* 1998; **21**:193–207.

24. Binder T, Blank L, Bock HG, Bulirsch R, Dahmen W, Diehl M, Kronseder T, Marquardt W, Schlöder JP, von Stryk O. Introduction to model based optimization of chemical processes on moving horizons. In *Online Optimization of Large Scale Systems*: *State of the Art*, Grötschel M, Krumke SO, Rambau J (eds). Springer: Berlin, 2001; 295–340. Available at: http://www.zib.de/dfg-echtzeit/Publikationen/Preprints/Preprint-01-15.html.

25. Ober-Blöbaum S. Discrete mechanics and optimal control. *Ph.D. Thesis*, University of Paderborn, 2008.

26. Junge O, Marsden JE, Ober-Blöbaum S. Discrete mechanics and optimal control. *Proceedings of the 16th IFAC World Congress*, Prague, 2005.

27. Marsden J, West M. Discrete mechanics and variational integrators. *Acta Numerica* 2001; **10**:357–514.

28. Junge O, Ober-Blöbaum S. Optimal reconfiguration of formation flying satellites. *IEEE Conference on Decision and Control and European Control Conference* (*ECC*), Seville, Spain, 2005.

29. Junge O, Ober-Blöbaum S. Optimal reconfiguration of formation flying spacecraft—a decentralized approach. *Proceedings of the IEEE Conference on Decision and Control and European Control Conference* (*ECC*), San Diego, U.S.A., 2006.

30. Betsch P, Steinmann P. Constrained integration of rigid body dynamics. *Computer Methods in Applied Mechanics and Engineering* 2001; **191**:467–488.

31. Reich S. Symplectic integrators for systems of rigid bodies. *Fields Institute Communications* 1996; **10**:181–191.

32. Krysl P. Direct time integration of rigid body motion with discrete-impulse midpoint approximation: explicit Newmark algorithms. *Communications in Numerical Methods in Engineering* 2006; **22**:441–451.

33. Bou-Rabee NM, Marsden JE. Hamilton–Pontryagin integrators on Lie groups: introduction and structure-preserving

properties. *Foundations of Computational Mathematics* 2008; **9**:197–219.

34. Bauchau O, Trainelli L. The vectorial parameterization of rotation. *Nonlinear Dynamics* 2003; **1**:71–92.

35. Leyendecker S, Betsch P, Steinmann P. Objective energy-momentum conserving integration for the constrained dynamics of geometrically exact beams. *Computer Methods in Applied Mechanics and Engineering* 2006; **195**: 2313–2333.

36. Betsch P, Leyendecker S. The discrete null space method for the energy consistent integration of constrained mechanical systems. Part II: multibody dynamics. *International Journal for Numerical Methods in Engineering* 2006; **67**: 499–552.

37. Leyendecker S, Betsch P, Steinmann P. The discrete null space method for the energy consistent integration of constrained mechanical systems. Part III: flexible multibody dynamics. *Multibody System Dynamics* 2008; **19**:45–72.

38. Leyendecker S, Marsden JE, Ortiz M. Variational integrators for constrained dynamical systems. *ZAMM* 2008; **88**:677–708.

39. Marsden JE, Ratiu TS. *Introduction to Mechanics and Symmetry* (2nd edn). Texts in Applied Mathematics, vol. 1994. Springer: Berlin, 1999.

40. Goldstein H, Poole C, Safko J. *Classical Mechanics*. Addison-Wesley: Reading, MA, 2002.

41. Antmann S. *Nonlinear Problems in Elasticity*. Springer: Berlin, 1995.

42. Leyendecker S. Mechanical integrators for constrained dynamical systems in flexible multibody dynamics. *Ph.D. Thesis*, University of Kaiserslautern, 2006.

43. Bullo F, Lewis A. *Geometric Control of Mechanical Systems*. Springer: Berlin, 2004.

44. Timmermann J. Die Nullraum-Methode in Kombination mit DMOC zur optimalen Steuerung mechanischer Systeme mit holonomen Zwangsbedingungen. *Diplomarbeit*, University of Paderborn, 2008.