



Discrete adjoint method for variational integration of constrained ODEs and its application to optimal control of geometrically exact beam dynamics

Matthias Schubert¹ · Rodrigo T. Sato Martín de Almagro¹ · Karin Nachbagauer^{2,3} · Sina Ober-Blöbaum⁴ · Sigrid Leyendecker¹

Received: 30 November 2022 / Accepted: 5 September 2023 / Published online: 6 October 2023
© The Author(s) 2023

Abstract

Direct methods for the simulation of optimal control problems apply a specific discretization to the dynamics of the problem, and the discrete adjoint method is suitable to calculate corresponding conditions to approximate an optimal solution. While the benefits of structure preserving or geometric methods have been known for decades, their exploration in the context of optimal control problems is a relatively recent field of research. In this work, the discrete adjoint method is derived for variational integrators yielding structure preserving approximations of the dynamics firstly in the ODE case and secondly for the case in which the dynamics is subject to holonomic constraints. The convergence rates are illustrated by numerical examples. Thirdly, the discrete adjoint method is applied to geometrically exact beam dynamics, represented by a holonomically constrained PDE.

Keywords Optimal control · Discrete adjoint method · Variational integrators · Geometrically exact beam · Holonomically constrained system

✉ M. Schubert
matthias.schubert@fau.de

R.T. Sato Martín de Almagro
rodrigo.t.sato@fau.de

K. Nachbagauer
karin.nachbagauer@fh-wels.at

S. Ober-Blöbaum
sinaober@math.uni-paderborn.de

S. Leyendecker
sigrid.leyendecker@fau.de

¹ Friedrich-Alexander-Universität Erlangen-Nürnberg, Institute of Applied Dynamics, Immerwahrstrasse 1, 91058 Erlangen, Germany

² Faculty of Engineering and Environmental Sciences, University of Applied Sciences Upper Austria, Stelzhamerstraße 23, 4600 Wels, Austria

³ Institute for Advanced Study, Technical University of Munich, Lichtenbergstraße 2a, 85748 Garching, Germany

⁴ Universität Paderborn, Warburger Straße 100, 33098 Paderborn, Germany

1 Introduction

There are two alternative ways to handle an optimal control problem numerically. The so-called *indirect methods* first derive the necessary conditions for optimality in the continuous-time setting by applying PONTYAGIN's maximum principle and then discretizing the resulting equations. In contrast, *direct methods* first discretize the continuous problem, turning it into a finite dimensional one, and then apply a discrete version of PONTYAGIN's maximum principle. In both cases, one is led to the augmentation of the original objective with the different constraints enforced by LAGRANGE multipliers. The LAGRANGE multipliers enforcing the plant (the dynamic equations) of the problem are commonly called adjoint or co-state variables. In the multibody systems literature it is common to refer to this as the adjoint method, and in particular, the discrete adjoint method when considered as a direct method. In this contribution, we apply the discrete adjoint method to optimal control problems with variational integrators approximating the dynamics.

In general, for direct approaches, the discretization of the ODE governing the dynamics results in a specific discretization of the adjoint variables especially for symplectic methods as e.g. variational integrators [1–3]. Variational and thus symplectic numerical methods are worthy of consideration as they can benefit the solution of boundary value problems [4]. For the optimal control of constrained ODEs, discretizations with conservation properties are of interest as well [5, 6].

The optimal control of mechanical PDEs, such as string and beam dynamics, is an active field of research [7–9]. The discrete adjoint method has been used for the optimization of flexible multibody systems [10] as well as for parameter identification in rigid body dynamics [11, 12]. The discrete adjoint method for variational integrators with holonomic constraints is discussed in [13]. The discrete adjoint method is derived for a specific discretization of dynamics, and this matches the chosen integrator. Therefore, it suggests itself to be applied to integrators that are structure preserving [14].

In this work we briefly summarize variational integrators and then show how to derive the discrete adjoint equations for this class of integrators. The basic principles, the derivation of boundary conditions, and the discretization of forces are explained. The discrete adjoint method is then extended to variational integrators for holonomically constrained ODEs. The convergence behavior of both methods is investigated with the example of a mathematical pendulum. Finally, the method is applied to the constrained PDE case of geometrically exact beam dynamics.

2 Discrete adjoint method for variational integrators

2.1 Variational integrators

This section illustrates the derivation of the equations of motion for forced systems via variational principles in the continuous and discrete setting [2, 15]. These equations have to be fulfilled as constraints for the optimal control problem.

Consider a Lagrangian mechanical system whose configuration space is the n -dimensional smooth manifold Q . The motion of our system is represented by a curve $q : [0, T] \rightarrow Q$, $t \mapsto q(t)$. We denote the velocity of the configuration at time t by $\dot{q}(t) \in T_{q(t)}Q$. The Lagrangian is a function defined on the tangent bundle of Q , $\mathcal{T}Q$, $\mathcal{L} : \mathcal{T}Q \rightarrow \mathbb{R}$. It usually represents the difference of kinetic and potential energy. An external Lagrangian control force is a map $f_{\mathcal{L}} : \mathcal{T}Q \times \mathcal{U} \rightarrow T^*Q$ where $\mathcal{U} \subseteq \mathbb{R}^l$, $l \leq n$, is the

space of admissible controls. A control is thus a curve $u : [0, T] \rightarrow \mathcal{U}$. The total virtual work of such a system vanishes

$$\delta \int_0^T \mathcal{L}(q(t), \dot{q}(t)) dt + \int_0^T f_{\mathcal{L}}(q(t), \dot{q}(t), u(t)) \delta q(t) dt = 0, \quad \forall \delta q(t). \tag{1}$$

This is the LAGRANGE–D’ALEMBERT principle (with controls), which states that the total virtual work evaluated over a physical trajectory of the system q (and a control u) vanishes for all variations $\delta q(t)$ with fixed end-points $\delta q(0) = \delta q(T) = 0$. This leads to the equations of motion, the forced EULER–LAGRANGE equations:

$$-\frac{d}{dt} \frac{\partial \mathcal{L}(q, \dot{q})}{\partial \dot{q}} + \frac{\partial \mathcal{L}(q, \dot{q})}{\partial q} + f_{\mathcal{L}}(q, \dot{q}, u) = 0. \tag{2}$$

This principle is an extension of HAMILTON’S principle to include nonconservative forces such as control or dissipative forces. A forced variational integrator is derived via the approximation of the action and the virtual work of nonconservative forces and subsequent variation in the discrete setting [15–17]. The time interval $[0, T]$ is discretized by N time nodes, we consider a discrete configuration path $\{q_n\}_{n=0}^N$ with $q_n \approx q(t_n)$ with linear approximation of $q(t)$ in $[t_n, t_{n+1}]$. The approximation of the action integral via the discrete Lagrangian L_d and the approximation of the virtual work of nonconservative forces via the left and right side discrete forces f_d^- and f_d^+ is considered. The input variable is approximated as $u_n \approx u(t_n)$. In each time interval $[t_n, t_{n+1}]$, the control path $u_d = \{u_n\}_{n=0}^{N-1}$ is approximated constant.

$$\int_{t_n}^{t_{n+1}} \mathcal{L}(q(t), \dot{q}(t)) dt \approx L_d(q_n, q_{n+1}), \tag{3}$$

$$\begin{aligned} \int_{t_n}^{t_{n+1}} f_{\mathcal{L}}(q(t), \dot{q}(t), u(t)) \delta q(t) dt &\approx f_d^-(q_n, q_{n+1}, u_n) \delta q_n \\ &+ f_d^+(q_n, q_{n+1}, u_n) \delta q_{n+1}. \end{aligned} \tag{4}$$

The discrete total virtual work vanishes:

$$\sum_{n=0}^{N-1} [\delta L_d(q_n, q_{n+1}) + f_d^-(q_n, q_{n+1}, u_n) \delta q_n + f_d^+(q_n, q_{n+1}, u_n) \delta q_{n+1}] = 0, \quad \forall \delta q_n \tag{5}$$

with $\delta q_0 = \delta q_N = 0$. The discrete LAGRANGE–D’ALEMBERT principle leads to the discrete, forced EULER–LAGRANGE equations, which are derived via discrete variation and subsequent rearrangement of terms for fixed boundary conditions. The slot derivatives D_k denote derivatives with respect to the k th argument.

$$D_1 L_d(q_n, q_{n+1}) + D_2 L_d(q_{n-1}, q_n) + f_d^-(q_n, q_{n+1}, u_n) + f_d^+(q_{n-1}, q_n, u_{n-1}) = 0 \tag{6}$$

for $n = 1, \dots, N - 1$. This equation takes two positions at the current and the previous time node and defines the relation with the next one. Given q_{n-1}, q_n, u_{n-1} , and u_n , this equation determines a unique q_{n+1} provided the discrete Lagrangian is regular, i.e., the matrix $D_1 D_2 L_d = D_2 D_1 L_d$ is regular.

The initial conditions are usually defined on \mathcal{TQ} as position and velocity or on \mathcal{T}^*Q as position and momentum, but not on $Q \times Q$ as two positions at different points in time.

To initialize this time stepping scheme, both a continuous and a discrete version of the LEGENDRE transformation are needed.

The continuous LEGENDRE transformation $\mathbb{F}\mathcal{L} : \mathcal{T}\mathcal{Q} \rightarrow \mathcal{T}^*\mathcal{Q}$, $(q, \dot{q}) \mapsto (q, p = D_2L(q, \dot{q}))$ connects the Lagrangian and the Hamiltonian formulations of dynamics. It allows us to compute an initial momentum p^0 from an initial configuration and velocity (q^0, \dot{q}^0) . In the discrete setting, the (forced) discrete LEGENDRE transformation defines two distinct maps from the discrete state space to the cotangent bundle, $\mathbb{F}^\pm L_d : \mathcal{Q} \times \mathcal{Q} \times \mathcal{U} \rightarrow \mathcal{T}^*\mathcal{Q}$, defined by

$$\begin{aligned} \mathbb{F}^- L_d : (q_n, q_{n+1}, u_n) &\mapsto (q_n, p_n^-) & (7a) \\ &= \left(q_n, -D_1 L_d(q_n, q_{n+1}) - f_d^-(q_n, q_{n+1}, u_n) \right), \end{aligned}$$

$$\begin{aligned} \mathbb{F}^+ L_d : (q_n, q_{n+1}, u_n) &\mapsto (q_{n+1}, p_{n+1}^+) & (7b) \\ &= \left(q_{n+1}, D_2 L_d(q_n, q_{n+1}) + f_d^+(q_n, q_{n+1}, u_n) \right), \end{aligned}$$

with the left and right side discrete momenta p_n^- and p_n^+ . These allow us to interpret the discrete EULER-LAGRANGE equations (6) as a matching of momenta $p_n^- = p_n^+$ for $n = 1, \dots, N - 1$.

To initialize the algorithm, given a configuration q^0 , a velocity \dot{q}^0 , and an initial control u_0 , the relation

$$D_2 L(q^0, \dot{q}^0) = p^0 = -D_1 L_d(q^0, q_1) - f_d^-(q^0, q_1, u_0) \tag{8}$$

determines q_1 .

2.2 Derivation of the discrete adjoint method for variational integrators

Similar to the discrete variational principle in Sect. 2.1, now the discrete adjoint method for variational integrators in (6) is derived via a discrete variational principle, and the structure and the resulting numerical method for the adjoint equations are illustrated.

Here, we concentrate on a discrete objective J_d containing a quadratic MAYER term:

$$J_M(q_N, p_N) = \frac{1}{2}(q_N - q^N)^T S_q(q_N - q^N) + \frac{1}{2}(p_N - p^N)^T S_p(p_N - p^N), \tag{9}$$

where S_q and S_p are positive semidefinite matrices. The MAYER term is used to relax the enforcement of the end state conditions (q^N, p^N) , introducing weights for the reaching of the configuration and the momentum at the last time step N .

The discrete adjoint method is derived by augmenting the objective with the variational integrator (6) and (8) as constraints and by taking variations of the augmented objective [2].

The resulting nonlinear constrained optimization problem reads

$$\min_{q_d, u_d} J_d(q_d, u_d) = J_M(q_N, p_N) + \sum_{n=0}^{N-1} \frac{1}{2} u_n^T R u_n \tag{10a}$$

subject to:

$$q_0 = q^0, \tag{10b}$$

$$p^0 = -D_1 L_d(q^0, q_1) - f_d^-(q^0, q_1, u_0), \tag{10c}$$

$$0 = D_1 L_d(q_n, q_{n+1}) + f_d^-(q_n, q_{n+1}, u_n) \tag{10d}$$

$$+ D_2 L_d(q_{n-1}, q_n) + f_d^+(q_{n-1}, q_n, u_{n-1}), \quad \text{for } n = 1, \dots, N - 1,$$

$$p_N = D_2 L_d(q_{N-1}, q_N) + f_d^+(q_{N-1}, q_N, u_{N-1}). \tag{10e}$$

The quantities p^0 and q^0 are prescribed initial conditions at the initial time node. The objective also includes a LAGRANGE term, which is quadratic in the control and R is a positive-definite weight matrix. Equation (10e) defining p_N corresponds to the discrete LEGENDRE transformation $\mathbb{F}^+ L_d(q_{N-1}, q_N, u_{N-1})$.

Remark 1 The dependence on q_{N-1} and q_N of the momentum term (10e) of the MAYER term makes it more prone to producing larger contributions than the configuration term. This can make the optimization process unstable and possibly not convergent. To improve this, an iterative approach may be used where the end momentum of the (i)th iteration $p_N^{(i)}$ is used to inform the choice of a modified desired end momentum \tilde{p}^N such that

$$\|p_N^{(i)} - \tilde{p}^N(p_N^{(i)}, p^N)\| \leq \|p_N^{(i)} - p^N\|$$

with $\tilde{p}^N(p^N, p^N) = p^N$. The procedure can be initialized by considering a first iteration with $S_p = 0$ and ended once $\|p_N^{(i)} - p^N\|$ is sufficiently small to allow us to substitute \tilde{p}^N by p^N in a final iteration.

The objective J_d is augmented to \tilde{J}_d by the initial conditions and the discrete EULER-LAGRANGE equations via adjoint variables $\lambda_n \approx \lambda(t_n)$ with the discrete adjoint path $\lambda_d = \{\lambda_n\}_{n=0}^{N-1}$. The indices are chosen such that λ_n pairs with the corresponding momenta p_n^\pm .

$$\begin{aligned} \tilde{J}_d(q_d, u_d, \lambda_d) &= J_M(q_N, p_N(q_{N-1}, q_N, u_{N-1})) + \sum_{n=0}^{N-1} \frac{1}{2} u_n^T R u_n \tag{11} \\ &+ \lambda_0^T \left[p^0 + D_1 L_d(q^0, q_1) + f_d^-(q^0, q_1, u_0) \right] \\ &+ \sum_{n=1}^{N-1} \lambda_n^T \left[D_1 L_d(q_n, q_{n+1}) + D_2 L_d(q_{n-1}, q_n) \right. \\ &\left. + f_d^-(q_n, q_{n+1}, u_n) + f_d^+(q_{n-1}, q_n, u_{n-1}) \right]. \end{aligned}$$

The discrete variation of the augmented objective $\delta \tilde{J}_d = 0$ has to vanish for variations δu_n , $\delta \lambda_n$, and δq_n with boundary conditions $\delta q_0 = 0$, that is, directly enforced as $q_0 = q^0$ at the initial time node is specified in problem (10a)–(10e). The variation of the three types of variables leads to three sets of equations. The variation w.r.t the adjoint variables leads to the discrete EULER-LAGRANGE equations, the constraints in (10a)–(10e). The variation with respect to the configuration variable yields the adjoint equations, reading with rearrangement of terms as follows:

$$\begin{aligned} &\lambda_{N-1}^T [D_2 D_1 L_d(q_{N-1}, q_N) + D_2 f_d^-(q_{N-1}, q_N, u_{N-1})] \tag{12a} \\ &= -S_q (q_N - q^N) - S_p [p_N(q_{N-1}, q_N, u_{N-1}) - p^N] \end{aligned}$$

$$\begin{aligned} &\times [D_2 D_2 L_d(q_{N-1}, q_N) + D_2 f_d^+(q_{N-1}, q_N, u_{N-1})], \\ \lambda_{N-2}^T &[D_2 D_1 L_d(q_{N-2}, q_{N-1}) + D_2 f_d^-(q_{N-2}, q_{N-1}, u_{N-2})] \end{aligned} \tag{12b}$$

$$\begin{aligned} &+ \lambda_{N-1}^T [D_2 D_2 L_d(q_{N-2}, q_{N-1}) + D_2 f_d^+(q_{N-2}, q_{N-1}, u_{N-2}) \\ &+ D_1 D_1 L_d(q_{N-1}, q_N) + D_1 f_d^-(q_{N-1}, q_N, u_{N-1})] \\ &= -S_p [p_N(q_{N-1}, q_N, u_{N-1}) - p^N] \\ &\times [D_1 D_2 L_d(q_{N-1}, q_N) + D_1 f_d^+(q_{N-1}, q_N, u_{N-1})], \end{aligned}$$

$$0 = \lambda_{n-1}^T [D_2 D_1 L_d(q_{n-1}, q_n) + D_2 f_d^-(q_{n-1}, q_n, u_{n-1})] \tag{12c}$$

$$\begin{aligned} &+ \lambda_n^T [D_2 D_2 L_d(q_{n-1}, q_n) + D_2 f_d^+(q_{n-1}, q_n, u_{n-1}) \\ &+ D_1 D_1 L_d(q_n, q_{n+1}) + D_1 f_d^-(q_n, q_{n+1}, u_n)] \\ &+ \lambda_{n+1}^T [D_1 D_2 L_d(q_n, q_{n+1}) + D_1 f_d^+(q_n, q_{n+1}, u_n)] \quad \text{for } n = N - 2, \dots, 1. \end{aligned}$$

The discrete variational principle directly provides the boundary conditions (12a) and (12b) for the two last adjoint variables, as no boundary conditions for the state variables are prescribed at these time nodes. The variation w.r.t. the input u_n yields the optimality conditions. Note that the last equation is different:

$$\begin{aligned} 0 &= Ru_n + \lambda_n^T D_3 f_d^-(q_n, q_{n+1}, u_n) \\ &+ \lambda_{n+1}^T D_3 f_d^+(q_n, q_{n+1}, u_n), \quad \text{for } n = 0, \dots, N - 2, \end{aligned} \tag{13a}$$

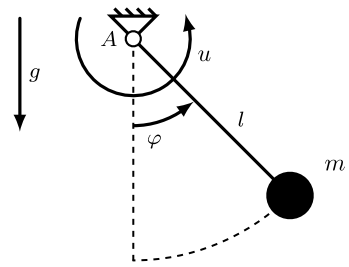
$$\begin{aligned} 0 &= Ru_{N-1} + \lambda_{N-1}^T D_3 f_d^-(q_{N-1}, q_N, u_{N-1}) \\ &+ S_p [p_N(q_{N-1}, q_N, u_{N-1}) - p^N] D_3 f_d^+(q_{N-1}, q_N, u_{N-1}). \end{aligned} \tag{13b}$$

The discrete EULER-LAGRANGE equations (6) can be solved forward in time and the adjoint equations (12a)–(12c) backward in time sequentially given the configuration path to determine the discrete adjoint variables as a shooting method while using the input equations (13a)–(13b) to update the input. Such a direct shooting algorithm directly uses the equations derived above and thus is simple to implement. However, an appropriately small time step h is necessary for stable integration in both directions in time. The discrete optimization problem with respect to q_d , u_d , and λ_d can also be solved by applying an interior point algorithm [18] or sequential quadratic programming [19]. In those, the variational integrator is used as equality constraints for the optimization as in (10a)–(10e).

2.3 Application of the discrete adjoint method to a mathematical pendulum

Let us consider a mathematical pendulum as depicted in Fig. 1, in minimal coordinates $q = \varphi$ with the Lagrangian $\mathcal{L}(\varphi, \dot{\varphi}) = \frac{1}{2}ml^2\dot{\varphi}^2 - mgl \cos(\varphi)$ that is actuated by a torque $f = u$. The discrete Lagrangian approximated with the midpoint rule is $L_d(\varphi_n, \varphi_{n+1}) = \frac{1}{2}hml^2(\frac{\varphi_{n+1}-\varphi_n}{h})^2 - hmg \cos(\frac{\varphi_{n+1}+\varphi_n}{2})$ with the time step h . The discrete forces are

Fig. 1 Torque-controlled mathematical pendulum



$f_d^\pm(\varphi_n, \varphi_{n+1}, u_n) = \frac{1}{2}hu_n$. The desired end configuration in this problem is $q^N = \varphi^N = \pi$. The end momentum has to vanish $p^N = 0$. The first slot derivatives of the discrete Lagrangian used for the discrete EULER-LAGRANGE equations are as follows:

$$D_1 L_d(\varphi_n, \varphi_{n+1}) = -ml^2 \frac{\varphi_{n+1} - \varphi_n}{h} + \frac{h}{2} mgl \sin\left(\frac{\varphi_{n+1} + \varphi_n}{2}\right), \tag{14}$$

$$D_2 L_d(\varphi_{n-1}, \varphi_n) = ml^2 \frac{\varphi_n - \varphi_{n-1}}{h} + \frac{h}{2} mgl \sin\left(\frac{\varphi_n + \varphi_{n-1}}{2}\right). \tag{15}$$

The time stepping scheme (6) for the configuration is

$$0 = \frac{\varphi_{n+1} - 2\varphi_n + \varphi_{n-1}}{h} - \frac{h}{2l} g \sin\left(\frac{\varphi_{n+1} + \varphi_n}{2}\right) - \frac{h}{2l} g \sin\left(\frac{\varphi_n + \varphi_{n-1}}{2}\right) - h \frac{u_n + u_{n-1}}{2}. \tag{16}$$

It is initialized with

$$0 = p^0 - ml^2 \frac{\varphi_1 - \varphi_0}{h} + \frac{h}{2} mgl \sin\left(\frac{\varphi_1 + \varphi_0}{2}\right) + h \frac{u_0}{2}. \tag{17}$$

The second derivatives of the discrete Lagrangian inserted in the adjoint equations (12c) leads to

$$0 = \frac{\lambda_{n-1}^T - 2\lambda_n^T + \lambda_{n+1}^T}{h} - \frac{\lambda_n^T + \lambda_{n-1}^T}{2} \frac{h}{2l} g \cos\left(\frac{\varphi_n + \varphi_{n-1}}{2}\right) - \frac{\lambda_{n+1}^T + \lambda_n^T}{2} \frac{h}{2l} g \cos\left(\frac{\varphi_{n+1} + \varphi_n}{2}\right). \tag{18}$$

Two equations according to (12a) and (12b) are necessary to initialize the backward integration (18) in time:

$$0 = \lambda_{N-1}^T \left[-\frac{ml^2}{h} + \frac{h}{4} mg \cos\left(\frac{\varphi_N + \varphi_{N-1}}{2}\right) \right] + S_q(\varphi_N - \pi) + S_p \left[ml^2 \frac{\varphi_N - \varphi_{N-1}}{h} + \frac{h}{2} mgl \sin\left(\frac{\varphi_N + \varphi_{N-1}}{2}\right) \right] \tag{19a}$$

$$\begin{aligned}
 & \times \left[\frac{ml^2}{h} + \frac{h}{4} mg \cos \left(\frac{\varphi_N + \varphi_{N-1}}{2} \right) \right], \\
 0 = & \frac{2\lambda_{N-1}^T - \lambda_{N-2}^T}{h} \tag{19b} \\
 & + \frac{\lambda_{N-1}^T + \lambda_{N-2}^T}{2} \frac{h}{2} \frac{g}{l} \cos \left(\frac{\varphi_{N-1} + \varphi_{N-2}}{2} \right) + \lambda_{N-1}^T \frac{h}{4} \frac{g}{l} \cos \left(\frac{\varphi_N + \varphi_{N-1}}{2} \right) \\
 & + S_p \left[ml^2 \frac{\varphi_N - \varphi_{N-1}}{h} + \frac{h}{2} mgl \sin \left(\frac{\varphi_N + \varphi_{N-1}}{2} \right) \right] \\
 & \times \left[\frac{ml^2}{h} + \frac{h}{4} mg \cos \left(\frac{\varphi_N + \varphi_{N-1}}{2} \right) \right].
 \end{aligned}$$

The equations for the input are as follows:

$$0 = Rh u_n + h \frac{\lambda_n^T + \lambda_{n+1}^T}{2}, \quad \text{for } n = 0, \dots, N - 2, \tag{20a}$$

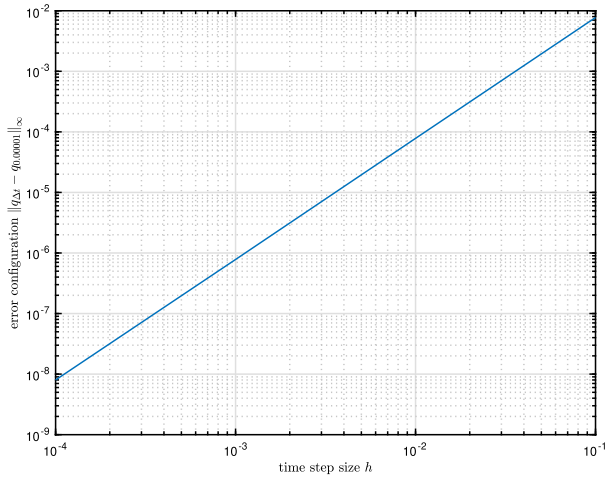
$$\begin{aligned}
 0 = & Rh u_{N-1} + \frac{h}{2} \lambda_{N-1}^T \tag{20b} \\
 & + \frac{h}{2} S_p \left[ml^2 \frac{\varphi_N - \varphi_{N-1}}{h} + \frac{h}{2} mgl \sin \left(\frac{\varphi_N + \varphi_{N-1}}{2} \right) \right].
 \end{aligned}$$

The convergence of the configuration q_d and the adjoint variables λ_d is illustrated in Figs. 2(a) and 2(b), respectively. A simulation time of $T = 2$ and a constant input of $u^n = 1$ for $n = 0, \dots, N - 1$ is used; the pendulum has a length of $L = 1$ with a gravitational constant of $g = 9.81$. The mass of the pendulum is $m = 1$. For the input weight $R = 10^{-5}h$ is used. The weights in the MAYER term are $S_q = 10^3$ and $S_p = 10^{-2}$. These values were chosen to obtain solutions that achieve the upswing of the pendulum to the upper equilibrium point, with minimal effort. Larger values for the input weighting lead to solutions with end configuration at the lower equilibrium of the pendulum. The absolute error in these plots is computed using the infinity norm of the difference of the variables and a reference solution $(q_{\text{ref}}, \lambda_{\text{ref}})$, which is a simulation with a fine discretization of $h = 10^{-5}$, $\|q_d - q_{\text{ref}}\|_\infty$, and $\|\lambda_d - \lambda_{\text{ref}}\|_\infty$, respectively. The convergence rate for the configuration and adjoint variables is equal, we observe second order convergence. This is in accordance with the theoretical results in [2]. These convergence results are derived for the forward integration of the time stepping scheme (16) and the subsequent backwards solution of (18) using the configuration variables calculated with the same time step width.

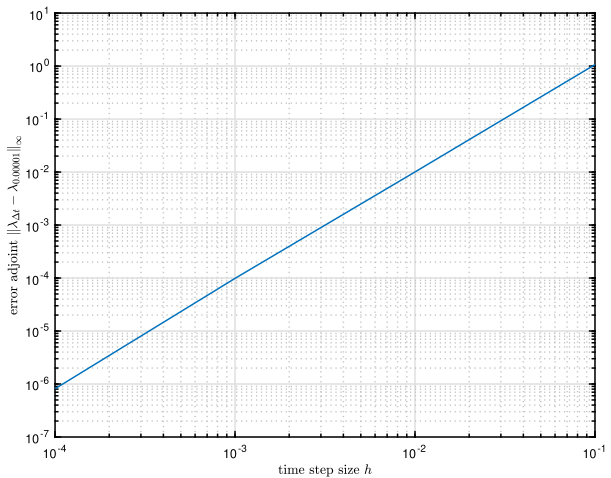
The optimized motion of the pendulum is depicted in Figs. 3(a), 3(b), 3(c), and 3(d). The momentum p and the kinetic energy T are close to zero at the end of the simulation with the optimized input acting on the pendulum.

Remark 2 PONTYAGIN’s maximum principle leads to necessary conditions for optimality in the continuous-time setting. The resulting adjoint equations are $\ddot{\lambda}^T - g/l \cos \varphi = 0$ and the control equations are $Ru + \lambda = 0$. It can be checked that the discrete equations (18) and (20a) are the corresponding discrete versions of these equations when discretized using a midpoint rule. The discrete boundary conditions (19a)–(19b) and (20b), however, are not so easy to relate to their continuous counterparts. We plan to address this very issue in a future publication.

Fig. 2 Error of the configuration q_d and adjoint variable λ_d



(a) Error of configuration q_d versus time step for the mathematical pendulum in minimal coordinates.



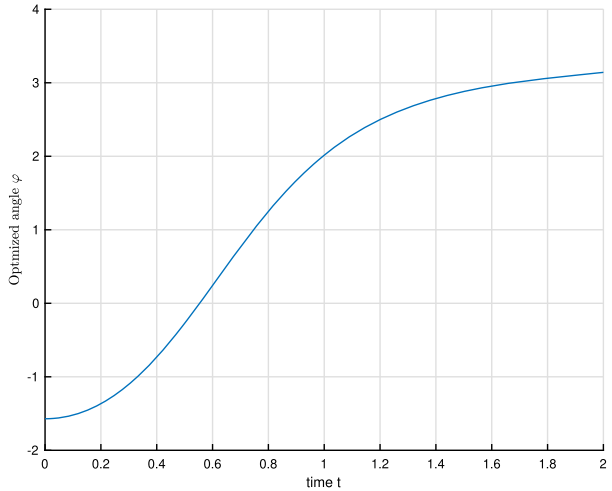
(b) Error of adjoint variables λ_d versus time step for the mathematical pendulum in minimal coordinates.

3 Discrete adjoint method for variational integration of constrained dynamics

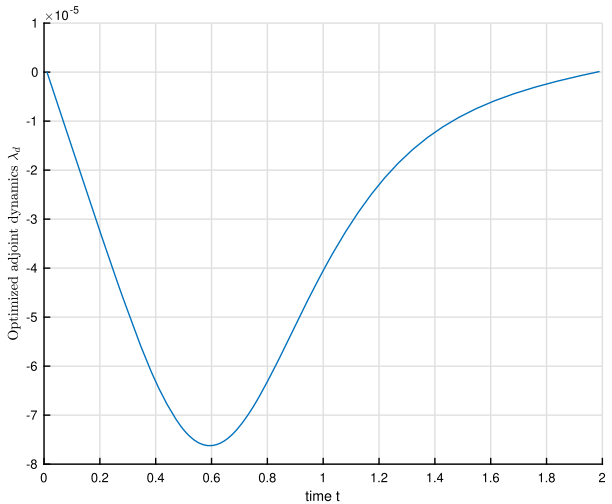
3.1 Variational integration of constrained dynamics

The derivation of variational integrators for constrained systems that use null space projection and nodal reparametrization [20] is shortly summarized in the following section, using similar steps as in Sect. 2. The discrete adjoint method for such systems is derived thereafter similar to Sect. 2.2.

Fig. 3 Optimization results for the pendulum, using the discrete adjoint method and single shooting



(a) Optimized angle φ .



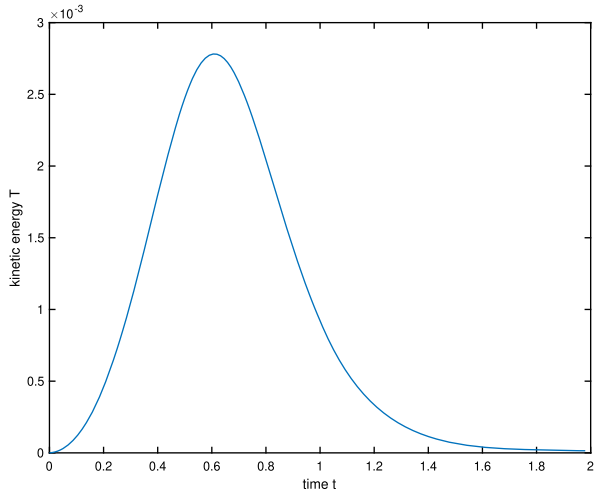
(b) Optimized adjoint variable λ .

Up until now, we have worked in local coordinates directly on the configuration manifold \mathcal{Q} . However, it can be advantageous to consider \mathcal{Q} an ambient (vector) space parametrized by redundant coordinates and constrain the motion by constraints. Given a scleromic, holonomic constraint function $g : \mathcal{Q} \rightarrow \mathbb{R}^m$, the constraint submanifold is then

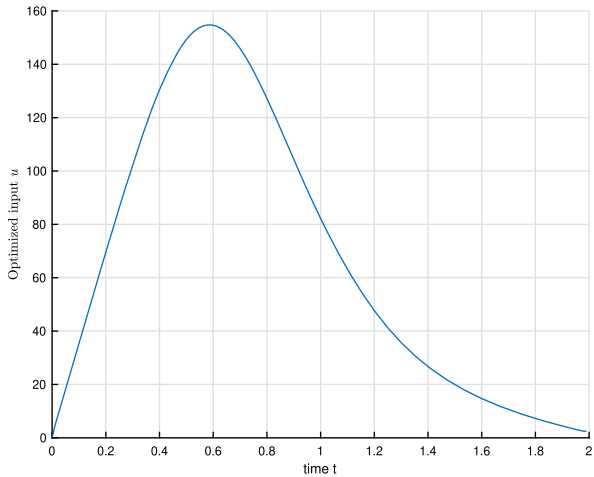
$$\mathcal{M} := \{q \in \mathcal{Q} \mid g(q) = 0\}. \tag{21}$$

We assume that the Jacobian $\frac{\partial g}{\partial q}$ has full rank m , so the dimension of the constraint manifold is $n - m$, the number of degrees of freedom of the mechanical system. We also assume consistent initial conditions (q^0, \dot{q}^0) that fulfill the constraints on configuration level $g(q^0) = 0$ as well as on velocity level $\frac{d}{dt}g(q^0) = \frac{\partial g(q^0)}{\partial q} \dot{q}^0 = 0$.

Fig. 3 (Continued)



(c) Kinetic energy T of the optimized upswing maneuver.



(d) Optimized input u .

A LAGRANGE multiplier ν is used to enforce the constraint by appending the term $-g(q)^T \nu$ to the Lagrangian in the action integral. Thus, the LAGRANGE–D’ALEMBERT principle in this setting reads

$$0 = \delta \int_0^T [\mathcal{L}(q(t), \dot{q}(t)) - g(q)^T \nu] dt + \int_0^T f_{\mathcal{L}}(q(t), \dot{q}(t), u(t)) \delta q dt, \quad \forall \delta q, \delta \nu \quad (22)$$

with $\delta q(0) = \delta q(T) = 0$. The constraint part of the action integral is approximated with the trapezoidal rule:

$$\int_{t_n}^{t_{n+1}} g(q(t))^T \nu(t) dt \approx \frac{1}{2} [g_d(q_n) \nu_n + g_d(q_{n+1}) \nu_{n+1}] \quad (23)$$

with $g_d(q_n) = hg(q_n)$. Including this in the discrete variational principle in (5), in the constrained case, the discrete variational principle, the variation of the discrete action sum with the variations δq_n and δv_n and $\delta q_0 = \delta q_N = 0$ with subsequent rearrangement of terms leads to the discrete, constrained EULER-LAGRANGE equations

$$0 = D_1 L_d(q_n, q_{n+1}) + D_2 L_d(q_{n-1}, q_n) + \frac{\partial g_d(q_n)^T}{\partial q_n} v_n \tag{24a}$$

$$+ f_d^-(q_n, q_{n+1}, u_n) + f_d^-(q_{n-1}, q_n, u_{n-1}),$$

$$0 = g(q_{n+1}) \tag{24b}$$

of dimension $n + m$. To reduce the dimension of (24a) from n to $n - m$ and eliminate the LAGRANGE multipliers, thus avoiding conditioning problems related to these, a discrete null space matrix $P(q_n) \in \mathbb{R}^{n \times (n-m)}$, with columns spanning the tangent space $T_{q_n} \mathcal{M}$, that only depends on quantities at the current step can be applied such that the constraint forces are eliminated. Further, a nodal reparametrization $q_{n+1} = F_d(q_n, v_{n+1})$ with $v_{n+1} \in \mathcal{V} \subseteq \mathbb{R}^{n-m}$ is then used to eliminate the constraints as $g(F_d(q_n, v_{n+1})) = 0$, $\forall v_{n+1} \in \mathcal{V}$, for $n = 0, \dots, N - 1$. Together with the null space matrix, the reparametrization $F_d : \mathcal{V} \times \mathcal{Q} \rightarrow \mathcal{M}$ leads to the integration scheme

$$P^T(q_n) [D_1 L_d(q_n, F_d(q_n, v_{n+1})) + D_2 L_d(q_{n-1}, q_n) + f_d^-(q_n, F_d(q_n, v_{n+1}), u_n) + f_d^+(q_{n-1}, q_n, u_{n-1})] = 0, \text{ for } n = 1, \dots, N - 1 \tag{25}$$

that has to be iteratively solved for v_{n+1} in each time step, given q_{n-1}, q_n, u_{n-1} and u_n .

The redundant control forces $f(q, u) = B^T(q)\tau(u) \in \mathbb{R}^n$ depend on the generalized control forces $\tau(u) \in \mathbb{R}^{n-m}$ and the input transformation matrix $B^T(q) \in \mathbb{R}^{n \times (n-m)}$ that must be chosen such that the consistency with the constraints and consistency of momentum maps are ensured [6]. The discrete approximations of the redundant forces

$$f_d^-(q_n, q_{n+1}, u_n) = \frac{h}{2} B^T(q_n)\tau(u_n),$$

$$f_d^+(q_n, q_{n+1}, u_n) = \frac{h}{2} B^T(q_{n+1})\tau(u_n)$$

capture the effect of the generalized forces acting on the time $[t_n, t_{n+1}]$. We have assumed that u is approximated constant in each time interval.

3.2 Derivation of the discrete adjoint method for variational integration of constrained dynamics

The constrained setting with null space projection and reparametrization for a mechanical system leads to implicit equations of minimal dimension. The discrete adjoint method applied to such a system leads to adjoint variables of minimal dimension $n - m$. It also involves the null space projection for the adjoint equations.

The starting point is a problem such as in equation (10a)–(10e), but now constrained by the discrete EULER-LAGRANGE equations for the constrained system with null space projection and nodal reparametrization (25) as in [6]. Similar to the procedure outlined in Sect. 2.2, the objective is augmented with the discrete EULER-LAGRANGE equations. As

these equations are defined on \mathcal{M} using the nodal reparametrization $F_d(q_n, v_{n+1})$, adjoint variables of the same dimension as v_{n+1} are necessary.

An objective J_d consisting of a MAYER term and an integral term quadratic in the control, similar to the discrete adjoint method for systems without constraints in equation (10a)–(10e) is considered:

$$J_d = \frac{1}{2}(q_N - q^N)^T S_q(q_N - q^N) + \sum_{n=0}^{N-1} \frac{1}{2} u_n^T R u_n. \tag{26}$$

However, to simplify matters, the MAYER term of the momentum has been omitted since it can be handled similarly as in the unconstrained case. The variation of the objective $\delta J_d = 0$ with respect to all variables δq_n , $\delta \lambda_n$, δu_n , and δv_{n+1} at all time steps has to vanish. The variation of the redundant configuration δq_n with respect to the minimal coordinate δv_n reads

$$\delta q_n = D_2 F_d(q_{n-1}, v_n) \delta v_n. \tag{27}$$

The Jacobian matrix $\frac{\partial F_d}{\partial v_n}$ is a null space matrix [21]. After applying this relation, the adjoint equations become

$$\begin{aligned} \lambda_{N-1}^T P^T(q_{N-1}) [D_2 D_1 L_d(q_{N-1}, F_d(q_{N-1}, v_N))] \\ = -S_q(q_N - q^N) D_2 F_d(q_{N-1}, v_N), \end{aligned} \tag{28a}$$

$$\begin{aligned} \lambda_{N-2}^T P^T(q_{N-2}) [D_2 D_1 L_d(q_{N-2}, F_d(q_{N-2}, v_{N-1}))] \\ = - \left\{ \lambda_{N-1}^T D_1 P^T(q_{N-1}) [D_1 L_d(q_{N-1}, F_d(q_{N-1}, v_N)) + D_2 L_d(q_{N-2}, q_{N-1}) \right. \\ \left. + f_d^-(q_{N-1}, F_d(q_{N-1}, v_N), u_{N-1}) + f_d^+(q_{N-2}, q_{N-1}, u_{N-2}) \right] \\ \left. + \lambda_{N-1}^T P^T(q_{N-1}) [D_1 D_1 L_d(q_{N-1}, F_d(q_{N-1}, v_N)) + D_2 D_2 L_d(q_{N-2}, q_{N-1}) \right] \right\} \\ \times D_2 F_d(q_{N-2}, v_{N-1}), \end{aligned} \tag{28b}$$

$$\begin{aligned} \lambda_{n-1}^T P^T(q_{n-1}) [D_2 D_1 L_d(q_{n-1}, F_d(q_{n-1}, v_n))] \\ = - \left\{ \lambda_n^T D_1 P^T(q_n) [D_1 L_d(q_n, F_d(q_n, v_{n+1})) + D_2 L_d(q_{n-1}, q_n) \right. \\ \left. + f_d^-(q_n, F_d(q_n, v_{n+1}), u_n) + f_d^+(q_{n-1}, q_n, u_{n-1}) \right] \\ \left. + \lambda_n^T P^T(q_n) [D_1 D_1 L_d(q_n, F_d(q_n, v_{n+1})) + D_2 D_2 L_d(q_{n-1}, q_n) \right] \\ \left. + \lambda_{n+1}^T P^T(q_{n+1}) D_1 D_2 L_d(q_n, q_{n+1}) \right\} D_2 F_d(q_{n-1}, v_n) \text{ for } n = N - 2, \dots, 1. \end{aligned} \tag{28c}$$

The variations with respect to the input variables vanish if

$$0 = R u_n + \lambda_n^T P^T(q_n) D_3 f_d^-(q_n, F_d(q_n, v_{n+1}), u_n) \tag{29a}$$

$$+ \lambda_{n+1}^T P^T(q_{n+1}) D_3 f_d^+(q_n, F_d(q_n, v_{n+1}), u_n) \text{ for } n = 1, \dots, N - 2,$$

$$0 = Ru_{N-1} + \lambda_{N-1}^T P^T(q_{N-1}) D_3 f_d^-(q_{N-1}, F_d(q_{N-1}, v_N), u_{N-1}) \tag{29b}$$

hold. The evaluation of these equations can be used to update the input variables in a shooting method.

3.3 Discrete adjoint method for a mathematical pendulum described as constrained system

The mathematical pendulum is described as a constrained system in the ambient space $\mathcal{Q} = \mathbb{R}^2$ with redundant coordinates $q = [x \ y]^T$ and the constraint equation $g(q) = 1/2(x^2 + y^2 - l^2)$. The null space matrix is $P(q_n)^T = [-y_n \ x_n]$, the input transformation matrix is $B(q_n)^T = [\frac{-y_n}{2l^2} \ \frac{x_n}{2l^2}]$, the generalized force is $\tau(u) = u$, and the nodal reparametrization reads

$$q_{n+1} = F_d(q_n, v_{n+1}) = \begin{bmatrix} \cos(v_{n+1}) & -\sin(v_{n+1}) \\ \sin(v_{n+1}) & \cos(v_{n+1}) \end{bmatrix} q_n. \tag{30}$$

The input variable can be interpreted as the physical torque and the variable v as the incremental angle. The Figs. 4(a) and 4(b) show the convergence results for the pendulum in the constrained case. The adjoint variables are of minimum dimension $(n - m)$ just as the configuration variables. The error is calculated in the same way as for the unconstrained case in Sect. 2.3 as infinity norm of the difference to the reference trajectory using the same parameters. These errors are determined with solutions obtained via forward time stepping for the configuration and backward time stepping for the adjoint variables with fixed input. It can be observed in the figures that also in the constrained case the convergence rate is quadratic. However, note that the theoretical results in [2] only consider the case in minimal coordinates and not the constrained case.

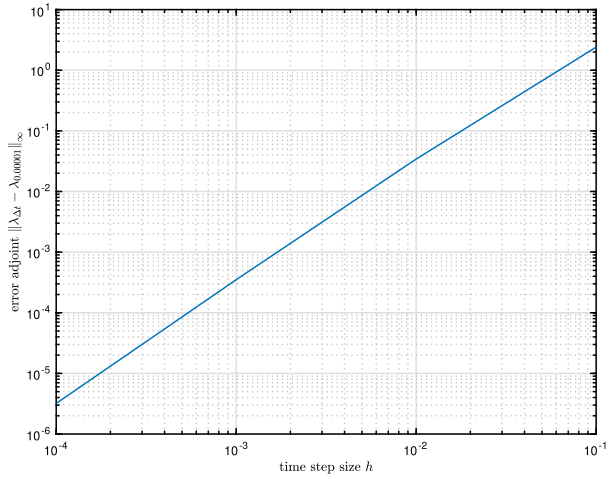
The optimized motion of the pendulum is depicted in Figs. 5(a), 5(b), 5(c), and 5(d). The input u and the kinetic energy T are close to zero at the end of the simulation with the optimized input acting on the pendulum. The end configuration is weighted with $S_q = 10^3$, the end momentum weight is $S_p = 10^{-2}$. The weight for the input is $R = 10^{-5}h$. This low weight for the input is chosen to reach the upper equilibrium position of the pendulum. It reduces the input from a constant initial guess of 1 as well as regularizing the optimization problem.

The results are similar to those obtained previously by the pendulum in minimal coordinates. Small differences in the solution are visible but show a similar optimized result.

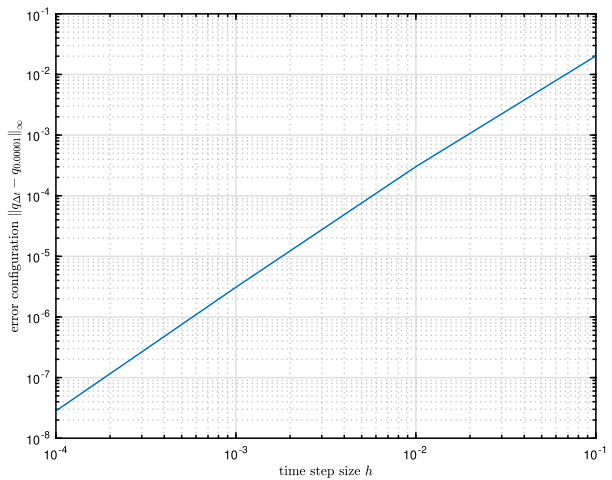
4 Discrete adjoint method for geometrically exact beam dynamics

In this section, the discrete adjoint method is applied to an optimal control problem involving dynamics of a geometrically exact beam being approximated via the multisymplectic integrator found in [22].

Fig. 4 Error of the configuration q_d and adjoint variable λ_d



(a) Error of adjoint variables λ_d versus time step for the mathematical pendulum as constrained system.

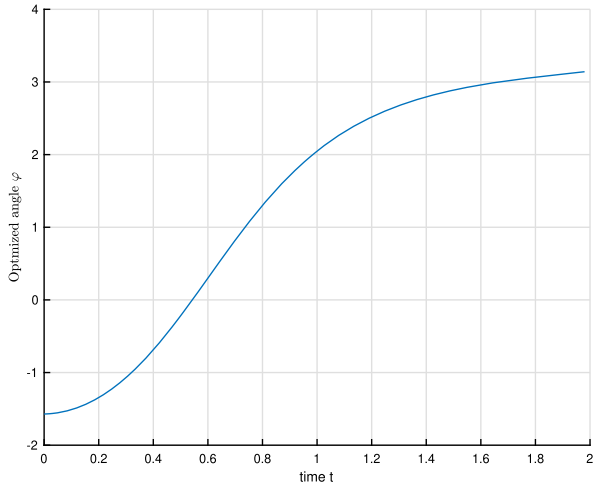


(b) Error of configuration q_d versus time step for the mathematical pendulum as constrained system.

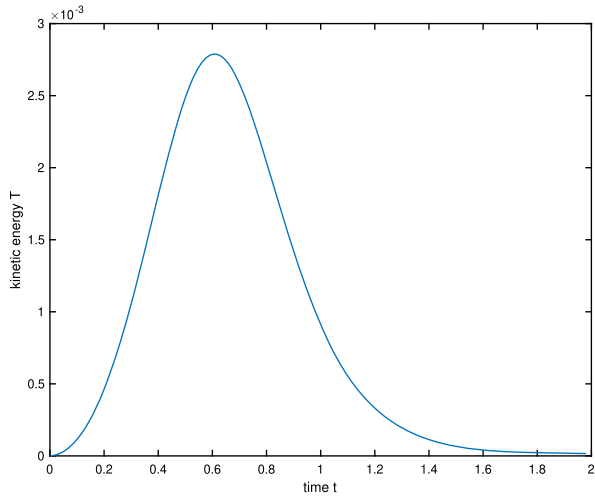
4.1 Geometrically exact beam model

The geometrically exact beam [26] models a rod-like deformable body as a curve $x(t, s) \in \mathbb{R}^3$, with a rigid cross section attached to each of its points. Here, $t \in [0, T]$ is used again to parametrize time, while $s \in [0, \ell]$ parametrizes the longitudinal position along the curve. The orientation of the cross section at s is described by a rotation $R(t, s) \in SO(3)$. When considered as a collection of columns $R(t, s) = [d_1(t, s), d_2(t, s), d_3(t, s)]$, the triad of vectors are known as the directors of the cross section (see Fig. 6).

Fig. 5 Optimization results for the pendulum in the constraint setting, using the discrete adjoint method and single shooting



(a) Optimized angle φ , calculated as sum of v .



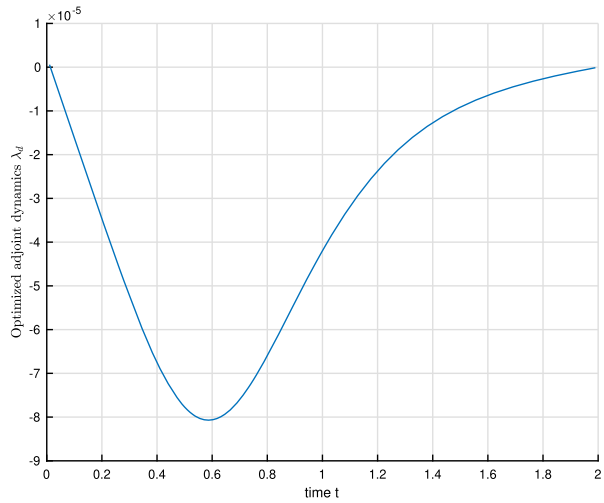
(c) Kinetic energy T of the optimized upswing maneuver.

This can be considered as a Lagrangian field theory with configuration space $\mathcal{Q} = \mathbb{R}^3 \times SO(3)$. This space is diffeomorphic to the group of special Euclidean transformations in 3D, $SE(3)$, to which it differs only in terms of group structure. In [22, 27], the authors claim it to be numerically more advantageous to consider this latter space.

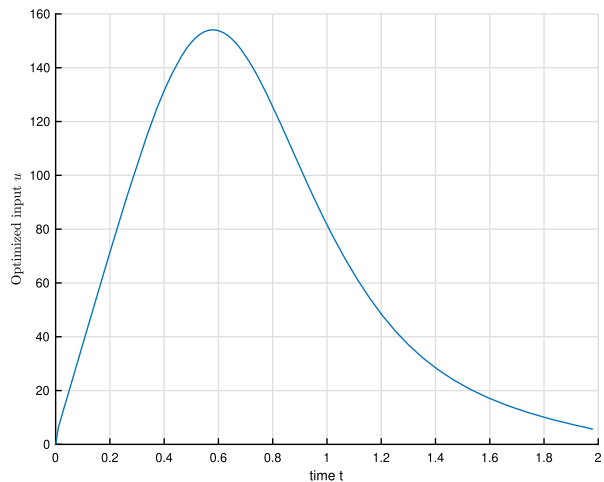
If $g(t, s) = (R(t, s), x(t, s)) \in SE(3)$ denotes the configuration of a cross section, its derivatives with respect to t and s are related to velocities and strains respectively. More specifically,

$$\begin{aligned}
 3(\Omega, V) &= g^{-1}\dot{g} = (R^{-1}\dot{R}, R^{-1}\dot{x}) && \text{body angular and linear time derivatives,} \\
 (K, W) &= g^{-1}g' = (R^{-1}R', R^{-1}x') && \text{body angular and linear space derivatives,}
 \end{aligned}$$

Fig. 5 (Continued)



(b) Optimized adjoint variable λ .



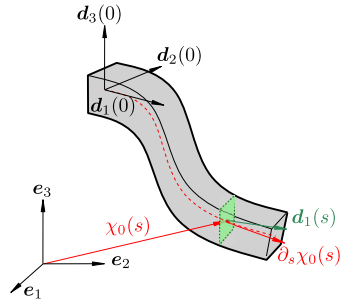
(d) Optimized input u .

where we have used $\dot{X} = \frac{\partial X}{\partial t}$ and $X' = \frac{\partial X}{\partial s}$ and “body” is meant to signify “in the reference frame of the section itself”. Considering a reference configuration $g_{\text{ref}}(s) \in SE(3)$, we also define the strains

$$(\Lambda, \Gamma) = (K - K_{\text{ref}}, W - W_{\text{ref}}).$$

The simple case of a straight initial configuration along the e_1 axis, $g_{\text{ref}}(s) = (I, se_1)$, where I is the identity matrix, leads to $\Lambda = K$ and $\Gamma = W - e_1$. One can see that Λ measures the curvature (bending and torsion) and Γ measures the difference between d_1 and x' (elongation and shear).

Fig. 6 Configuration of a geometrically exact beam



Considering a hyperelastic material model with moderate strains, the Lagrangian density of the system can be written as

$$\mathcal{L}(g, \dot{g}, g') = \frac{1}{2} (\Omega^T \mathbb{J} \Omega + \rho A V^T V - \Lambda^T \mathbb{C}_1 \Lambda - \Gamma^T \mathbb{C}_2 \Gamma) - U_{\text{ext}}(R, x),$$

where $\rho > 0$ is the linear density of the beam, $U_{\text{ext}} : SE(3) \rightarrow \mathbb{R}$ is an external potential function, and $\mathbb{J} = \rho \text{diag}([J_1, J_2, J_3])$ is the matrix of moments of inertia of the sections in the body frame. Assuming uniform cross sections and directors d_2 and d_3 coincident with the principal moments of area I_2 and I_3 , one gets that $J_1 = \rho (I_2 + I_3)$, $J_2 = \rho I_2$ and $J_3 = \rho I_3$, and $\mathbb{C}_1 = \text{diag}([G(I_2 + I_3), EI_2, EI_3])$, $\mathbb{C}_2 = \text{diag}([EA, \kappa_2 GA, \kappa_3 GA])$, which are the matrices representing the corresponding stiffness parameters of the sections. κ_2 and κ_3 are possible shear correction factors.

4.2 Unit dual quaternion formulation

Working on $SE(3)$ is difficult. In [22] the authors propose the use of a constrained approach where the space of dual quaternions $\tilde{\mathbb{H}}$, which is a vector space, is considered as ambient manifold and the unit dual quaternions $\tilde{\mathbb{H}}_1$ as constraint submanifold since it is well known that this latter space provides a double covering of $SE(3)$.

The space of dual quaternions is defined by

$$\tilde{\mathbb{H}} := \{ \tilde{q} = q_r + q_d \epsilon \mid q_r, q_d \in \mathbb{H}, \epsilon^2 = 0 \},$$

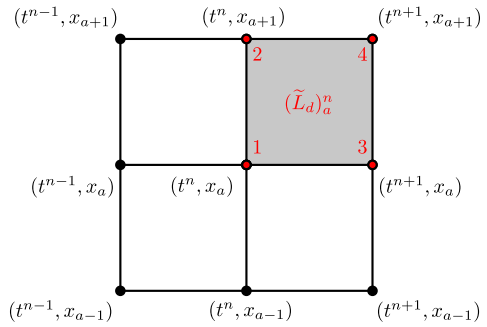
where ϵ is the so-called dual unit and

$$\mathbb{H} := \{ q = q_0 + q_1 i + q_2 j + q_3 k \mid q_i \in \mathbb{R}, i^2 = j^2 = k^2 = ijk = -1 \}$$

is the space of quaternions. Both of these are vector spaces, so working with them is quite simple.

Similar to complex numbers, a conjugation operation is defined on the space of quaternions, namely, if $p = p_0 + p_1 i + p_2 j + p_3 k$, then $\bar{p} = p_0 - q_1 i - q_2 j - q_3 k$, and this operation is inherited by dual quaternions. This defines a norm on \mathbb{H} , $\|p\| = \sqrt{\bar{p}p}$, and lets us write the inverse of p as $p^{-1} = \bar{p}/\|p\|^2$. This also defines a seminorm on $\tilde{\mathbb{H}}$ by $\|\tilde{p}\| = \sqrt{\tilde{p}\tilde{p}} = \|p_r\| + \frac{p_r^T p_\epsilon}{\|p_r\|} = \sqrt{p_r^T p_r} + \frac{p_r^T p_\epsilon}{\sqrt{p_r^T p_r}}$, where in the last equality we consider the quaternions q_r, q_ϵ as vectors in \mathbb{R}^4 . The set of unit quaternions and unit dual quaternions are

Fig. 7 Spacetime grid for the multisymplectic variational integrator with relative indices marked in red



thus $\mathbb{H}_1 := \{q \in \mathbb{H} \mid \|q\| = 1\}$ and $\tilde{\mathbb{H}}_1 := \{\tilde{q} \in \tilde{\mathbb{H}} \mid \|\tilde{q}\| = 1\}$ respectively. More explicitly, the latter implies

$$q_{0,r}^2 + q_{1,r}^2 + q_{2,r}^2 + q_{3,r}^2 = 1, \tag{31a}$$

$$q_{0,r}q_{0,\epsilon} + q_{1,r}q_{1,\epsilon} + q_{2,r}q_{2,\epsilon} + q_{3,r}q_{3,\epsilon} = 0. \tag{31b}$$

As stated before, an element $\tilde{q} \in \tilde{\mathbb{H}}_1$ can be put into correspondence with an element of $SE(3)$. In particular, we can parametrize \tilde{q} by a rotation angle θ and two purely imaginary quaternions n, x , i.e., $n_0 = x_0 = 0$, with $\|n\| = 1$, representing a rotation axis and a three-dimensional translation respectively. This way $q = \cos(\theta/2) + n \sin(\theta/2)$ and $\tilde{q} = q + \frac{1}{2}xq\epsilon$. If $\tilde{q}(t, s) \in \tilde{\mathbb{H}}_1$, then

$$\tilde{\Omega} := 2\tilde{q}\dot{\tilde{q}} = \Omega + V\epsilon, \quad \tilde{K} := 2\tilde{q}\dot{\tilde{q}}' = K + W\epsilon.$$

One can thus define an ambient Lagrangian in the dual quaternions

$$\tilde{\mathcal{L}}(\tilde{q}, \dot{\tilde{q}}, \tilde{q}') = 2\tilde{M}(\tilde{q}\dot{\tilde{q}}, \tilde{q}\dot{\tilde{q}}) - 2\tilde{C}(\tilde{q}\tilde{q}' - \tilde{q}_{\text{ref}}\tilde{q}'_{\text{ref}}, \tilde{q}\tilde{q}' - \tilde{q}_{\text{ref}}\tilde{q}'_{\text{ref}}) - \tilde{U}(\tilde{q}), \tag{32}$$

where $\tilde{M}(\tilde{q}, \tilde{p}) = q_r^T \tilde{\mathbb{J}} p_r + q_\epsilon^T \tilde{\rho} p_\epsilon$, $\tilde{C}(\tilde{q}, \tilde{p}) = q_r^T \tilde{\mathbb{C}}_1 p_r + q_\epsilon^T \tilde{\mathbb{C}}_2 p_\epsilon$, with $\tilde{\mathbb{J}} = \text{diag}([\alpha_1, \mathbb{J}])$, $\tilde{\rho} = \text{diag}([\alpha_2, \rho AI])$, $\tilde{\mathbb{C}}_1 = \text{diag}([\alpha_3, \mathbb{C}_1])$, and $\tilde{\mathbb{C}}_2 = \text{diag}([\alpha_4, \mathbb{C}_2])$, and $\alpha_i \in \mathbb{R}$. These α can be chosen arbitrarily as they play no role in the dynamics once the unity constraints (31a)–(31b) are enforced.

4.3 Discrete Lagrangian

To discretize the beam, the spacetime $[0, T] \times [0, \ell]$ is discretized into a regular grid (see Fig. 7) with constant space and time steps Δs and Δt respectively.

We discretize the ambient Lagrangian density (32) applying the trapezoidal rule in both space and time

$$\begin{aligned} \tilde{\mathcal{L}}_d(\tilde{q}_a^n, \tilde{q}_{a+1}^n, \tilde{q}_a^{n+1}, \tilde{q}_{a+1}^{n+1}) &= \frac{1}{4\Delta t \Delta s} \\ &\times \left[\tilde{\mathcal{L}}\left(\tilde{q}_a^n, \frac{\tilde{q}_a^{n+1} - \tilde{q}_a^n}{\Delta t}, \frac{\tilde{q}_{a+1}^{n+1} - \tilde{q}_{a+1}^n}{\Delta s}\right) + \tilde{\mathcal{L}}\left(\tilde{q}_{a+1}^n, \frac{\tilde{q}_{a+1}^{n+1} - \tilde{q}_{a+1}^n}{\Delta t}, \frac{\tilde{q}_{a+1}^{n+1} - \tilde{q}_a^n}{\Delta s}\right) \right. \\ &\left. + \tilde{\mathcal{L}}\left(\tilde{q}_a^{n+1}, \frac{\tilde{q}_a^{n+1} - \tilde{q}_a^n}{\Delta t}, \frac{\tilde{q}_{a+1}^{n+1} - \tilde{q}_{a+1}^n}{\Delta s}\right) + \tilde{\mathcal{L}}\left(\tilde{q}_{a+1}^{n+1}, \frac{\tilde{q}_{a+1}^{n+1} - \tilde{q}_{a+1}^n}{\Delta t}, \frac{\tilde{q}_{a+1}^{n+1} - \tilde{q}_a^{n+1}}{\Delta s}\right) \right] \end{aligned}$$

and introduce the notation $(\tilde{L}_d)_a^n := \tilde{L}_d(\tilde{q}_a^n, \tilde{q}_{a+1}^n, \tilde{q}_a^{n+1}, \tilde{q}_{a+1}^{n+1})$ to simplify the formulas.

As derived in [22], the discrete constrained EULER-LAGRANGE field equations are derived via a discrete variational principle in space and time and subsequent rearrangement of terms in space index a and time index n . As shown there, a natural choice of null space matrix is

$$\tilde{P}(\tilde{q}) = \begin{bmatrix} P(q_r) & 0 \\ P(q_\epsilon) & P(q_r) \end{bmatrix}, \quad P(q) = \frac{1}{2} \begin{bmatrix} -q_1 & -q_2 & -q_3 \\ q_0 & -q_3 & q_2 \\ q_3 & q_0 & -q_1 \\ -q_2 & q_1 & q_0 \end{bmatrix}.$$

The forced version of these equations results from the application of the discrete LA-GRANGE-D’ALEMBERT principle, similar to (5),

$$\sum_a \sum_n [\delta(\tilde{L}_d)_a^n + (f_d^1)_a^n \delta \tilde{q}_a^n + (f_d^2)_a^n \delta \tilde{q}_{a+1}^n + (f_d^3)_a^n \delta \tilde{q}_a^{n+1} + (f_d^4)_a^n \delta \tilde{q}_{a+1}^{n+1}] = 0$$

with $(f_d^i)_a^n := f_d^i(\tilde{q}_a^n, \tilde{q}_{a+1}^n, \tilde{q}_a^{n+1}, \tilde{q}_{a+1}^{n+1}, u_a^n)$ denoting all external and control forces, and $i = 1, \dots, 4$ coinciding with the corresponding relative node on which they are applied, as in Fig. 7. This leads to DEL with a force contribution from each adjacent spacetime rectangle sharing the node under consideration:

$$\begin{aligned} \tilde{P}(\tilde{q}_a^n)^T \left[D_1(\tilde{L}_d)_a^n + D_2(\tilde{L}_d)_{a-1}^n + D_3(\tilde{L}_d)_{a-1}^{n-1} + D_4(\tilde{L}_d)_{a-1}^{n-1} \right. \\ \left. + (f_d^1)_a^n + (f_d^2)_{a-1}^n + (f_d^3)_{a-1}^{n-1} + (f_d^4)_{a-1}^{n-1} \right] = 0. \end{aligned} \tag{33}$$

Suitable boundary conditions in space and time as well at the spacetime corners are directly derived via the discrete variational principle.

KELVIN-VOIGT type viscous damping is included as external forces that are proportional to the discrete approximation of the strain rate [23] with bulk viscosity ζ and shear viscosity η . In the moderate strain regime these result in a damping matrix $\tilde{\mathbb{D}} = \text{diag}([0, \eta(I_2 + I_3), \chi I_2, \chi I_3, 0, \chi A, \eta A, \eta A])$, where $\chi = \zeta(3 - E/G)^2 + \eta(E/G)^2/3$ is the extensional viscosity. The corresponding discrete force is

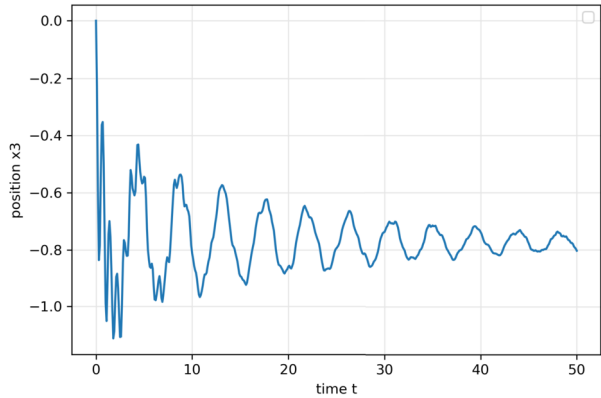
$$\begin{aligned} L_{(\tilde{q}_a^n)}^T (f_d^{\text{KV},1})_a^n &= L_{(\tilde{q}_{a+1}^n)}^T (f_d^{\text{KV},3})_a^n = \frac{\Delta t \Delta s}{4} \tilde{\mathbb{D}} \left(\frac{\tilde{K}_a^{n+1} - \tilde{K}_a^n}{\Delta t} \right), \\ L_{(\tilde{q}_{a+1}^n)}^T (f_d^{\text{KV},2})_a^n &= L_{(\tilde{q}_{a+1}^n)}^T (f_d^{\text{KV},4})_{a+1}^n = \frac{\Delta t \Delta s}{4} \tilde{\mathbb{D}} \left(\frac{\tilde{K}_{a+1}^{n+1} - \tilde{K}_{a+1}^n}{\Delta t} \right), \end{aligned}$$

where by $L_{\tilde{q}}^T$ we denote the transposed of the dual quaternion left multiplication operation by \tilde{q} , and $\tilde{K}_a^n = \tilde{K}(\tilde{q}_a^n, (\tilde{q}')_a^n)$ with $(\tilde{q}')_a^n = (\tilde{q}')_{a+1}^n = (\tilde{q}_{a+1}^n - \tilde{q}_a^n)/\Delta s$ and $(\tilde{q}')_{a+1}^{n+1} = (\tilde{q}')_{a+1}^{n+1} = (\tilde{q}_{a+1}^{n+1} - \tilde{q}_{a+1}^n)/\Delta s$. Figure 8 shows the position of the tip of a cantilever beam with fixed-free boundary conditions that is initially straight under gravity. The strain-rate proportional damping leads to reduced high frequency oscillations.

4.4 The discrete adjoint method in spacetime

The discrete adjoint method for the geometrically exact beam considers the configuration variables as well as the adjoint variables in space and time to derive the discrete adjoint equations in space and time. Single shooting in time while simultaneously solving the equations

Fig. 8 Viscous damping of a cantilever beam



in space is used for the solution of the optimal control problem. The BARZILEI–BORWEIN gradient method [24, 25] is used for the update. Here, a pendulum-like beam subject to gravity and fixed-free translation and free-free rotation boundary conditions is considered with a torque u applied at the fixed end as discrete redundant control forces

$$L_{(\tilde{q}_0^n)}^T (f_d^{C,1})_0^n = L_{(\tilde{q}_0^{n+1})}^T (f_d^{C,3})_0^n = 2\Delta t \Delta s u_0^n \mathbf{k}.$$

Since our control is only applied at the boundary, this is a boundary control problem for a PDE.

The desired configuration is the upright rotated position of the beam, specified for each node in space. The final position considered is undeformed. The desired maneuver is from the lower position to the upright position in such a way that the inertial terms cancel the strains in the end configuration. As the system is heavily underactuated, the chosen input does not allow us to control the motion in the axial direction and does not lead to a stationary upright position. Hence, no end momentum is imposed. Nonetheless, the control task should demonstrate the presented method in an academic example that resembles the previous pendulum examples sufficiently.

Our optimal control problem is of the form

$$\min_{\tilde{q}_d, u_d} J_d(\tilde{q}_d, u_d) = \sum_{a=0}^A [\tilde{q}_a^N - (\tilde{q}_a^N)_*]^T S_q [\tilde{q}_a^N - (\tilde{q}_a^N)_*] + \sum_{n=0}^{N-1} \frac{1}{2} (u_0^n)^T R (u_0^n) \tag{34a}$$

subject to:

$$\tilde{q}_a^0 = (\tilde{q}_a^0)_*, \tag{34b} \quad \text{for } a = 0, \dots, N,$$

$$u_0^0 = (u_0^0)_*, \tag{34c}$$

$$0 = \tilde{P}^T (\tilde{q}_0^n) [D_1(\tilde{L}_d)_0^n + D_3(\tilde{L}_d)_0^{n-1} + (f_d^1)_0^n + (f_d^3)_0^{n-1}], \tag{34d}$$

for $n = 1, \dots, N - 1,$

$$0 = \tilde{P}^T (\tilde{q}_a^n) [D_1(\tilde{L}_d)_a^n + D_2(\tilde{L}_d)_{a-1}^n + D_3(\tilde{L}_d)_a^{n-1} + D_4(\tilde{L}_d)_{a-1}^{n-1} \tag{34e}$$

$$+ (f_d^1)_a^n + (f_d^2)_{a-1}^n + (f_d^3)_a^{n-1} + (f_d^4)_{a-1}^{n-1}],$$

for $a = 1, \dots, A - 1,$ for $n = 1, \dots, N - 1,$

where $(\tilde{q}_a^0)_*$, $(u_0^0)_*$ are given initial discrete values and $(\tilde{q}_a^N)_*$ denotes the discretized desired end configuration.

The adjoint equations are obtained similar to the constrained temporal case by applying discrete variational calculus and nodal reparametrization, but now in space and time. However, the resulting equations are quite long, and so they will not be reproduced here in their entirety. For instance, the equations obtained by taking variations of the inputs at the fixed boundary $a = 0$ are

$$0 = (\lambda_0^n)^T \tilde{P}^T(\tilde{q}_0^n)(D_5 f_d^1)_0^n + (\lambda_0^{n+1})^T \tilde{P}^T(\tilde{q}_0^{n+1})(D_5 f_d^2)_0^n, \tag{35a}$$

for $n \in [1, N - 2]$,

$$0 = (\lambda_0^{N-1})^T \tilde{P}^T(\tilde{q}_0^{N-1})(D_5 f_d^1)_0^{N-1}. \tag{35b}$$

These are used to update the torque. If instead of boundary control we had controls over the bulk, then these equations would generalize to all nodes as follows:

$$0 = (\lambda_a^n)^T \tilde{P}^T(\tilde{q}_a^n)(D_5 f_d^1)_a^n + (\lambda_{a+1}^n)^T \tilde{P}^T(\tilde{q}_{a+1}^n)(D_5 f_d^2)_a^n + (\lambda_a^{n+1})^T \tilde{P}^T(\tilde{q}_a^{n+1})(D_5 f_d^3)_a^n + (\lambda_{a+1}^{n+1})^T \tilde{P}^T(\tilde{q}_{a+1}^{n+1})(D_5 f_d^4)_a^n,$$

for $a = 1, \dots, A - 2, n = 1, \dots, N - 2,$

$$0 = (\lambda_{A-1}^n)^T \tilde{P}^T(\tilde{q}_{A-1}^n)(D_5 f_d^1)_{A-1}^n + (\lambda_{A-1}^{n+1})^T \tilde{P}^T(\tilde{q}_{A-1}^{n+1})(D_5 f_d^3)_{A-1}^n,$$

for $n \in [1, N - 2]$,

$$0 = (\lambda_a^{N-1})^T \tilde{P}^T(\tilde{q}_a^{N-1})(D_5 f_d^1)_a^{N-1} + (\lambda_{a+1}^{N-1})^T \tilde{P}^T(\tilde{q}_{a+1}^{N-1})(D_5 f_d^2)_a^{N-1},$$

for $a \in [1, A - 2]$,

$$0 = (\lambda_{A-1}^{N-1})^T \tilde{P}^T(\tilde{q}_{A-1}^{N-1})(D_5 f_d^1)_{A-1}^{N-1}.$$

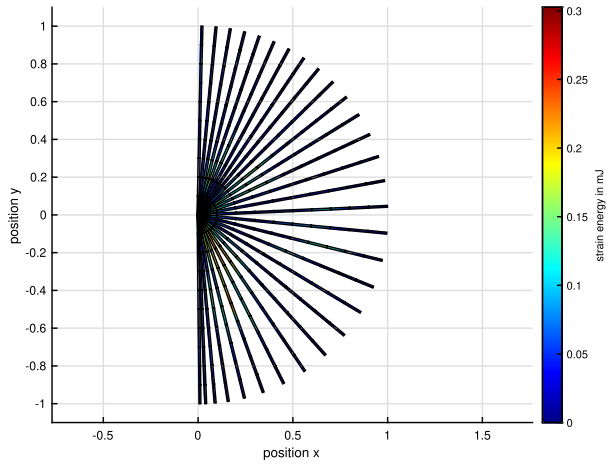
4.5 Fairly rigid beam

The fairly rigid beam demonstrates the sequential optimization of the beam dynamics with objective minimization of the control effort. The simulation of the beam dynamics uses $A = 10$ nodes in space and $N = 3000$ nodes in time. The beam has a length of $L = 1$. The simulation duration is $T = 1$. The resulting time step is $h = \frac{1}{3000}$ in time and the step size in space is $\Delta s = \frac{1}{10}$. A constant initial guess of $u^0 = 1500$ is used. The beam has a square cross-section of $A_{\text{cross}} = 0.01$ with a side length of $l_s = 0.1$. The chosen weighting for the end term is $S_q = 10^8$, and $R = 10^{-2}$ for the input.¹ The material of the beam is fairly rigid with a YOUNG'S modulus of $E = 210,000$ and a POISSON ratio of $\nu = 0.3$. The mass density is $\rho = 7.85$. The beam is damped with $\eta = 1 \cdot 10^{-1}$ and $\zeta = 1 \cdot 10^{-2}$.

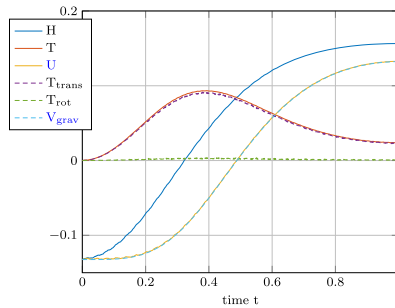
Figure 9(a) shows snapshot of the motion of the beam. Figure 9(b) shows the total energy H as well as all its contributions over time. The deformation energy is the difference between the total potential energy of the system U and the gravitational potential energy U_{grav} . The main contribution to the kinetic energy T is due to translation. At the end of the simulation, the kinetic energy reduces due to the input weight. The optimized input is

¹These values have been chosen to provide similar magnitudes to the different terms of the discrete objective. Notice that S_q affects only a single time step, multiplying terms with values around π and 0. R appears in a sum containing the 3000 time steps with input values between 2500 and 0.

Fig. 9 Motion of a fairly rigid beam and its energy



(a) Deformation of a fairly rigid beam during the upward motion.

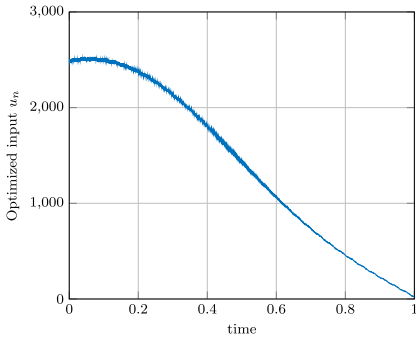


(b) Energy of the beam for the optimized motion over time t .

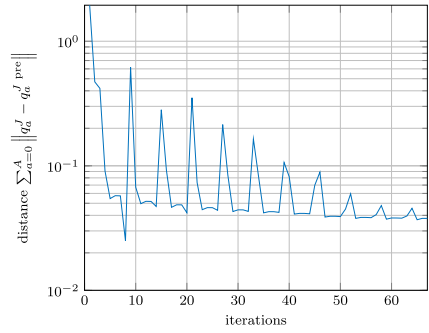
depicted in Fig. 10(a), it decreases to zero at the end of the simulation time. The optimized quantities, the distance of the beam to the desired end configuration as well as the control effort are depicted in Figs. 10(b) and 10(c), respectively. The gradient depicted in Fig. 10(d) shows heavy oscillations.

4.6 Very flexible beam

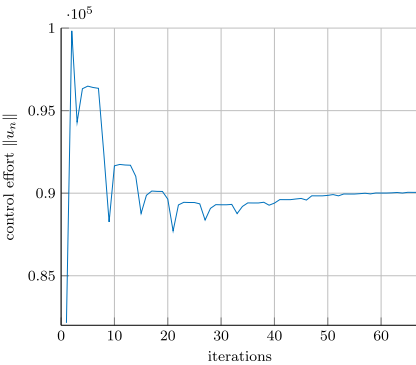
A very flexible beam demonstrates the sequential optimization for more flexible beams that show larger deformations and are therefore harder to control. The simulation of the beam and adjoint dynamics uses $A = 5$ nodes in space for a length of $L = 1$. This results in a space step width of $\Delta s = \frac{1}{5}$. The simulation time is $T = 0.5$ using $N = 600$ node in time and a time step of $h = \frac{1}{1200}$. The initial guess for the input is $u^0 = 50$ for all time intervals. The beam has a square cross-section of $A_{\text{cross}} = 0.0025$ with a side length of $l_s = 0.05$. The YOUNG'S modulus is $E = 50,000$ and the mass density $\rho = 1000$. The POISSON ratio is $\nu = 0.35$. KELVIN-VOIGT type damping is applied with $\eta = 1 \cdot 10^{-1}$ and $\zeta = 1 \cdot 10^{-2}$.



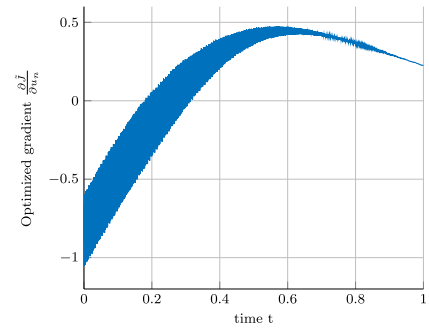
(a) Input u_n over time t .



(b) Distance to end configuration over optimization iterations.



(c) Control effort $\|u_n\|$ over optimization iterations.



(d) Gradient of the augmented objective with respect to the input $\frac{\partial J}{\partial u_n}$ over time t .

Fig. 10 Optimization results of beam dynamics using single shooting for a fairly rigid beam

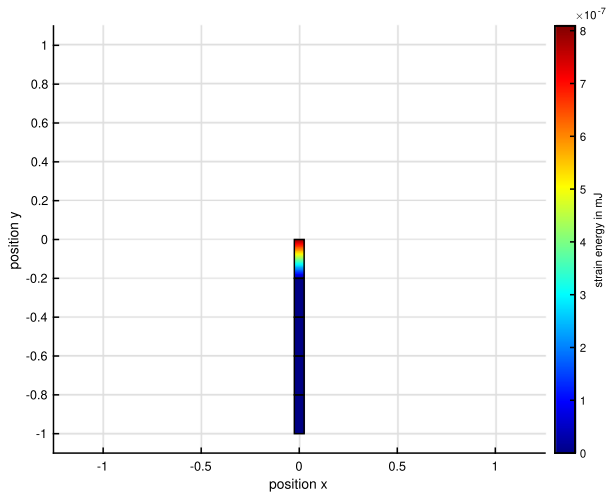
The weighting for the end configuration is $S_q = 10^2$. For this numerical experiment, the input weight was set to $R = 0$ since the chosen end configuration gets increasingly harder to reach for more flexible beams.

The optimization results are depicted in Fig. 11. The input in Fig. 12(a) is increased compared to the initial guess. In addition, oscillations are present. The gradient depicted in Fig. 12(b) shows oscillations with high frequency that are likely caused by the dynamics of the beam in normal direction as these deformations are of much higher frequency than bending deformations due to the difference in stiffness. The objective is depicted in 12(c). The largest decrease happens at the start of the optimization. Figure 12(d) depicts the total energy and its parts. During the optimization, mainly the translational part of the kinetic energy increases as well as the potential energy due to the gravitation.

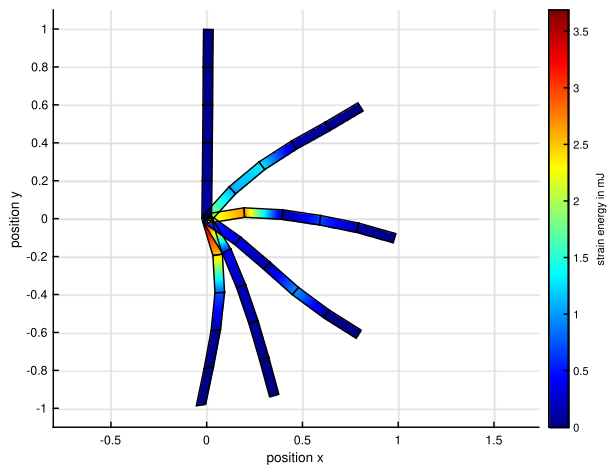
5 Summary

The discrete adjoint method for variational integration of (constrained) ODEs is derived, and its convergence properties are demonstrated with the help of numerical examples. Quadratic

Fig. 11 Snapshot of a very flexible beam at $n = 0$ as well as during the upward movement of a very flexible beam



(a) Initial configuration of the beam by example of a very flexible beam.



(b) Deformation of a very flexible beam during its motion.

convergence results of the configuration variables as well as for the adjoint variables based on simulations of a mathematical pendulum are observed. The discrete adjoint method is also applied to the multisymplectic Galerkin Lie group integrator for geometrically exact beam dynamics, in particular to the optimal control of the upward motion of a pendulum-like beam. The discrete adjoint method directly derives fitting equations at the boundary based on the discretization chosen for the variational integrator. The discrete adjoint method for constrained systems with null space projection and nodal reparametrization also directly results in the null space projection of the discrete adjoint equations. The properties of the discrete adjoint method applied to structure preserving integrators have to be analyzed further as to understand the connection in a more general setting.

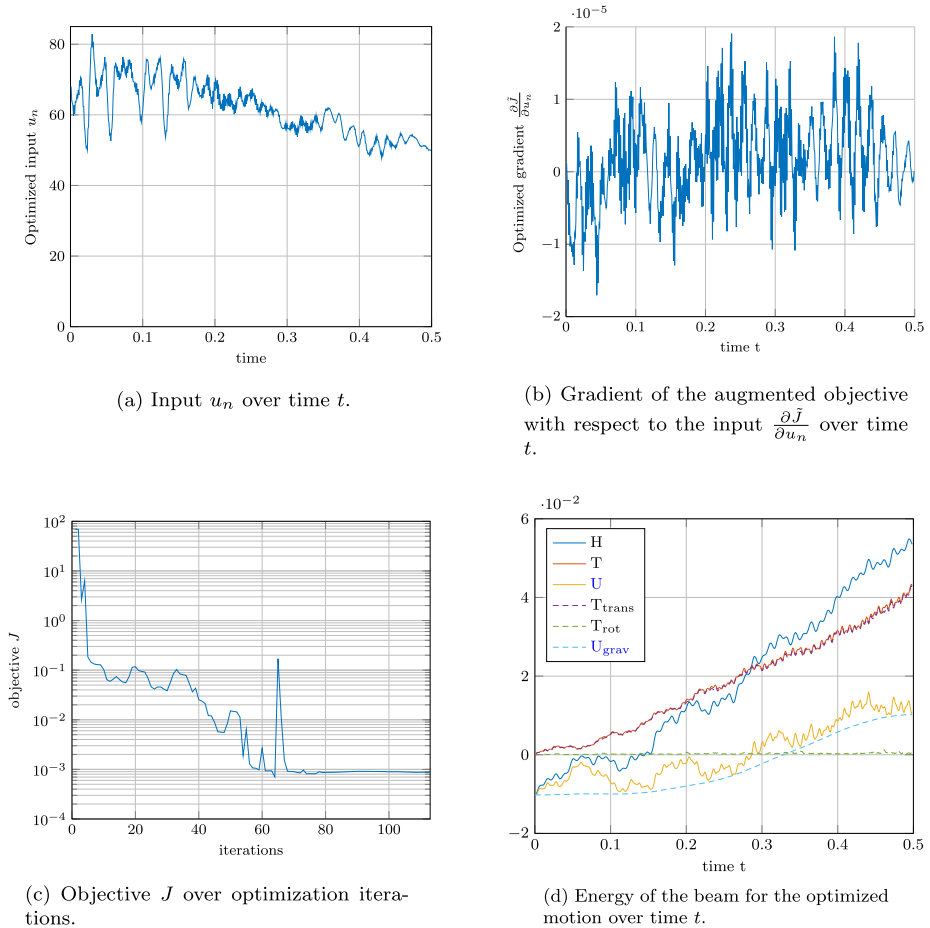


Fig. 12 Optimization results of beam dynamics of a very flexible beam

Author contributions M.S. wrote the initial version of the manuscript. R.S.T.M.A. contributed to the discussions, wrote much of the theoretical part of section 4 and provided additional help with figures and rewrites in other sections. All authors reviewed the manuscript. K.N., S.O., S.L. posed the research question and conducted the first research on the topic of this paper. SL continuously supervised M.S.’s work. M.S. wrote all code.

Funding Open Access funding enabled and organized by Projekt DEAL. This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860124. This publication reflects only the authors’ view, and the Research Executive Agency is not responsible for any use that may be made of the information it contains.

This work was partly supported by the German Research Foundation (DFG, German Research Foundation) under Grant SFB 1483 – Project-ID 442419336.

Karin Nachbagauer acknowledges support from the Technical University of Munich - Institute for Advanced Study.

Data Availability Not applicable.

Materials Availability Not applicable.

Declarations

Ethical approval Not applicable.

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Campos, C.M., Ober-Blöbaum, S., Trélat, E.: High order variational integrators in the optimal control of mechanical systems. *Discrete Contin. Dyn. Syst.* **35**(9), 4193–4223 (2015). <https://doi.org/10.3934/dcds.2015.35.4193>
2. Ober-Blöbaum, S., Junge, O., Marsden, J.E.: Discrete mechanics and optimal control: an analysis. *ESAIM Control Optim. Calc. Var.* **17**, 322–352 (2011). <https://doi.org/10.1051/cocv/2010012>
3. Bonnans, J.F., Laurent-Varin, J.: Computation of order conditions for symplectic partitioned Runge-Kutta schemes with application to optimal control. *Numer. Math.* **103**, 1–10 (2006). <https://doi.org/10.1007/s00211-005-0661-y>
4. McLachlan, R.I., Offen, C.: Bifurcation of solutions to Hamiltonian boundary value problems. *Nonlinearity* **31**, 2895 (2018). <https://doi.org/10.1088/1361-6544/aab630>
5. Betsch, P., Schneider, S.: Conservation of generalized momentum maps in the optimal control of constrained mechanical systems. *IFAC-PapersOnLine* **54**, 615–619 (2021). <https://doi.org/10.1016/j.ifacol.2021.06.123>
6. Leyendecker, S., Ober-Blöbaum, S., Marsden, J.E., Ortiz, M.: Discrete mechanics and optimal control for constrained systems. *Optim. Control Appl. Methods* **31**, 505–528 (2010). <https://doi.org/10.1002/oca.912>
7. Ströhle, T., Betsch, P.: A simultaneous space-time discretization approach to the inverse dynamics of geometrically exact strings. *Int. J. Numer. Methods Eng.* **123**, 2573–2609 (2022). <https://doi.org/10.1002/nme.6951>
8. Lismonde, A., Sonnevile, V., Brüls, O.: A geometric optimization method for the trajectory planning of flexible manipulators. *Multibody Syst. Dyn.* **47**, 347–362 (2019). <https://doi.org/10.1007/s11044-019-09695-z>
9. Brüls, O., Bastos, G. Jr, Seifried, R.: A stable inversion method for feedforward control of constrained flexible multibody systems. *J. Comput. Nonlinear Dyn.* **9**, 011014 (2014). <https://doi.org/10.1115/1.4025476>
10. Callejo, A., Sonnevile, V., Bauchau, O.A.: Discrete adjoint method for the sensitivity analysis of flexible multibody systems. *J. Comput. Nonlinear Dyn.* **14** (2019). <https://doi.org/10.1115/1.4041237>
11. Lauß, T., Oberpeilsteiner, S., Steiner, W., Nachbagauer, K.: The discrete adjoint method for parameter identification in multibody system dynamics. *Multibody Syst. Dyn.* **42**, 397–410 (2018). <https://doi.org/10.1007/s11044-017-9600-9>
12. Lauß, T., Oberpeilsteiner, S., Steiner, W., Nachbagauer, K.: The discrete adjoint gradient computation for optimization problems in multibody dynamics. *J. Comput. Nonlinear Dyn.* **12**, 031016 (2017). <https://doi.org/10.1115/1.4035197>
13. Ebrahimi, M., Butscher, A., Cheong, H., Iorio, F.: Design optimization of dynamic flexible multibody systems using the discrete adjoint variable method. *Comput. Struct.* **213**, 82–99 (2019). <https://doi.org/10.1016/j.compstruc.2018.12.007>
14. Sanz-Serna, J.M.: Symplectic Runge–Kutta schemes for adjoint equations, automatic differentiation, optimal control, and more. *SIAM Rev.* **58**, 3–33 (2016). <https://doi.org/10.1137/151002769>
15. Marsden, J.E., West, M.: Discrete mechanics and variational integrators. *Acta Numer.* **10**, 357–514 (2001). <https://doi.org/10.1017/S096249290100006X>
16. Hairer, E.: *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*, vol. 31. Springer, Berlin (2010)

17. Jordan, B.W., Polak, E.: Theory of a class of discrete optimal control systems. *J. Electron. Control* **17**(6), 697–711 (1964). <https://doi.org/10.1080/00207216408937740>
18. Wächter, A., Biegler, L.: On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Program.* **106**, 25–57 (2006). <https://doi.org/10.1007/s10107-004-0559-y>
19. Gill, P.E., Murray, W., Saunders, M.A.: SNOPT: an SQP algorithm for large-scale constrained optimization. *SIAM Rev.* **47**(1), 99–131 (2005). <https://doi.org/10.1137/S0036144504446096>
20. Leyendecker, S., Marsden, J.E., Ortiz, M.: Variational integrators for constrained dynamical systems. *Z. Angew. Math. Mech.* **88**(9), 677–708 (2008). https://doi.org/10.1007/0-387-24255-4_10
21. Betsch, P., Leyendecker, S.: The discrete null space method for the energy consistent integration of constrained mechanical systems. Part II: multibody dynamics. *Int. J. Numer. Methods Eng.* **67**, 499–552 (2006). <https://doi.org/10.1002/nme.1639>
22. Leitz, T., Sato Martín de Almagro, R.T., Leyendecker, S.: Multisymplectic Galerkin Lie group variational integrators for geometrically exact beam dynamics based on unit dual quaternion interpolation – no shear locking. *Comput. Methods Appl. Mech. Eng.* **374**, 113475 (2021). <https://doi.org/10.1016/j.cma.2020.113475>
23. Linn, J., Lang, H., Tuganov, A.: Geometrically exact Cosserat rods with Kelvin–Voigt type viscous damping. *Mech. Sci.* **4**, 79–96 (2013). <https://doi.org/10.5194/ms-4-79-2013>
24. Barzilai, J., Borwein, J.M.: Two-point step size gradient methods. *IMA J. Numer. Anal.* **8**, 141–148 (1988). <https://doi.org/10.1093/imanum/8.1.141>
25. Fletcher, R.: On the Barzilai-Borwein method. *Optim. Control Appl.* **96** (2001). https://doi.org/10.1007/0-387-24255-4_10
26. Simo, J.: A finite strain beam formulation. The three-dimensional dynamic problem. Part I. *Comput. Methods Appl. Mech. Eng.* **49**(1), 55–70 (1985)
27. Sonnevile, V., Brüls, O., Bauchau, O.A.: Interpolation schemes for geometrically exact beams: a motion approach. *Int. J. Numer. Methods Eng.* (2017). <https://doi.org/10.1002/nme.5548>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.