INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING
Int. J. Numer. Meth. Engng (2016)
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/nme.5330

Variational integrators – A continuous time approach

M. Muehlebach[1,*], T. Heimsch[2] and Ch. Glocker[3]

[1]*Department of Mechanical and Process Engineering, Institute for Dynamic Systems and Control, ETH Zurich, Sonneggstrasse 3, 8092 Zurich, Switzerland*
[2]*Department of Mechanical and Process Engineering, Institute for Mechanical Systems, ETH Zurich, Tannenstrasse 3, 8092 Zurich, Switzerland*
[3]*Segantinistrasse 47, 8049 Zurich, Switzerland*

SUMMARY

This article presents a family of variational integrators from a continuous time point of view. A general procedure for deriving symplectic integration schemes preserving an energy-like quantity is shown, which is based on the principle of virtual work. The framework is extended to incorporate holonomic constraints without using additional regularization. In addition, it is related to well-known partitioned Runge–Kutta methods and to other variational integration schemes. As an example, a concrete integration scheme is derived for the planar pendulum using both polar and Cartesian coordinates. Copyright © 2016 John Wiley & Sons, Ltd.

Received 15 March 2015; Revised 21 February 2016; Accepted 8 July 2016

KEY WORDS: multibody dynamics; numerical integration methods; variational integrators; symplectic integrators; discontinuous Galerkin methods

1. INTRODUCTION

For most mechanical systems, we must rely on numerical methods to investigate their properties. For obtaining at least a qualitative picture of the dynamics, we would like to ensure that even with a large step size and a long time horizon, the approximate solutions represent qualitatively the true motion. To that extent, integration methods based on Hamilton's principle or on a related variational principle, characterizing the motion as a whole, seem particularly suitable. Thus, variational integrators approach the numerical integration from a variational principle rather than a discretization of the corresponding ordinary differential equations. By doing so, underlying geometric properties of the motion are often preserved.

The recent literature on variational integrators, for example [1–3], relies typically on Hamilton's principle, which relates the dynamic equilibrium to stationary trajectories of a certain action functional. In the discrete mechanics framework, the action functional is discretized in the first place. In the spirit of Hamilton's principle, the resulting discrete action is required to be stationary subject to fixed boundary conditions, which leads to a discrete analogue of the Euler–Lagrange equations. These equations describe the evolution of the system according to the numerical integration scheme. By discretizing the action in the first place, a predominantly discrete-time point of view is adopted.

This article suggests an alternative approach: The dynamic equilibrium is approximated via discontinuous ansatz functions. Therefore, the principle of virtual action is formulated such that piecewise continuous trajectories are allowed and the virtual action is required to vanish for all variations spanned by the piecewise continuous basis functions. This leads in a consistent way to causal, iterative integration algorithms, which are shown to be symplectic. In contrast to the

*Correspondence to: Michael Muehlebach, Department of Mechanical and Process Engineering, Institute of Dynamical Systems and Control, ETH Zurich, Sonneggstrasse 3, 8092 Zurich, Switzerland.
†E-mail: michaemu@ethz.ch

discrete mechanics framework, the 'discretization' (in this context meant as an approximation of the underlying infinite dimensional problem by a finite dimensional one) happens implicitly by restricting the variations to be spanned by the basis functions, thereby offering a clear connection to the continuous-time dynamics of the system. In particular, according to the principle of d'Alembert–Lagrange, the proposed method has the interpretation of adding artificial constraint forces, constraining the trajectories to lie in the span of the basis functions. These constraint forces can be calculated a posteriori and can be used to quantify the approximation error. In addition, the presented approach can be extended to incorporate holonomic constraints by simply accounting for the virtual work generated by these constraints.

The presented approach offers also a connection to finite element methods, which apply to boundary value problems encountered in elastomechanics. In the standard finite element formulation, the displacements are approximated by basis functions, which are defined over sub-domains, thereby dividing the domain into finite elements. The displacements are typically required to be continuous, which results in a coupling of the finite elements. The approach presented herein is similar in the sense that the trajectories are approximated by basis functions defined over sub-intervals (in time). However, the trajectories are not required to be continuous, which decouples the time intervals in-between the discontinuities and results in a causal integration scheme applicable to initial value problems, as shown in the remainder.

1.1. Existing literature and classification of the presented work

One of the first publications about finite elements in space and time can be found by Fried [4] and Argyris [5]. Since then, several procedures under different names, for example, Petroff–Galerkin, weighted residuals, Hamilton's law, and time finite elements, have been proposed, which, although based on the same variational principles, differ greatly in their numerical application. A major source of disagreement represents the variations of the boundary values, which are required to vanish in the usual continuous version of Hamilton's principle. In [6], for example, the boundary values are retained, whereas in [5], the variations of the boundary values are required to vanish. For the algorithm presented in [7], a convergence proof is provided. The work is followed up by Aharoni and Bar-Yoseph [8], where discontinuous and continuous Galerkin approaches are presented and compared. Similarly, discontinuous ansatz functions are used in [9], and a link to Runge–Kutta methods is established. It happens that the good performance of well-known numerical integration schemes such as the Newmark method, the midpoint rule, or the Strömer-Verlet algorithm can be explained by tracing the integration schemes back to variational principles [1]. In [10], a discontinuous Galerkin approach is proposed for the simulation of nonsmooth dynamical systems, which, in particular, contains the classical Moreau–Jean timestepping schemes. The problem of simulating multibody systems subject to constraints (using a finite element method) is, for example, addressed in [11]. A similar approach is applied to the simulation of helicopter rotor dynamics in [12]. A spatial and temporal discretization based on a finite element approach is presented in [13] to solve problems arising in thermo-elastodynamics. The energy-momentum consistent timestepping schemes are shown to retain the first and second laws of thermodynamics in the discrete setting, and a convergence criterion is derived.

In the discrete mechanics framework, the action integral of Hamilton's principle is directly discretized (e.g., [1–3], and [14, p. 204]), which provides a discrete-time point of view. Both Marsden and West [1] and Lacoursière [3] interpret the classical results of Lagrangian and Hamiltonian mechanics, for example, Noether's theorem and symplectic reductions, in a discrete framework, with a special emphasis on symplectic integration. The approach is extended to incorporate holonomic constraints and is used to solve optimal control problems in [15–17], respectively. The holonomic constraints are enforced by augmenting the discrete action sum with Lagrange multipliers [16]. These multipliers, which have the physical interpretation of constraint forces, can be eliminated using suitable projections, leading to the so-called discrete null space method. The discrete mechanics framework, as presented in [1, 3], and [14, p. 204], expresses the discrete action sum using Lagrangian formalism. The Hamiltonian formalism is introduced in [18] by performing

DOI: 10.1002/nme

so-called left or right discrete Legendre transforms. A more direct approach is proposed in [2], where the action, expressed using Hamiltonian formalism, is discretized without recourse to Lagrangian formalism.

Several other contributions exist, for example, [19] and [20], where integration methods based on generating functions and approximative solutions to the Hamilton–Jacobi equation are presented.

In contrast to the discrete mechanics framework, for example, the work in [1–3], [14, p. 204], [15–18], we do not start from a discretized action integral. We use discontinuous basis functions to parametrize the solutions over the whole time interval of interest. Instead of Hamilton's principle, we use the principle of virtual action to approximate the dynamic equilibrium and restrict variations to be spanned by the basis functions. Consequently, our formulation is close to the continuous time dynamics and provides a clear interpretation of the discretization in terms of the principle of d'Alembert–Lagrange. Similar to [6], the variations of the boundary values at the discontinuous time instants will be retained and play a central role. The approach presented herein is based on Hamiltonian formalism, and therefore, the kinematic link between generalized coordinates and generalized momenta is not necessarily fulfilled exactly. The parametrization using basis functions allows to evaluate time derivatives in a straightforward and unambiguous manner, without referring to discrete Legendre transformations, as carried out in [1]. The proposed approach is also flexible enough to encompass, for instance, the Galerkin variational integration approach from [1] (in case of a constant mass matrix) or the approach from [2] as special cases.

In addition, the developed framework is used to incorporate holonomic constraints in a simple and consistent way. Unlike Lacoursière [3], no additional damping and regularization are needed for a numerically stable integration. Similar to the unconstrained case and in contrast to [1] and [16], the principle of virtual action is used as a starting point, and the discretization is carried out implicitly by restricting the variations to be spanned by the basis functions. The constraint is incorporated on velocity level. Compared with the constrained Galerkin variational integration approach presented in [1, p. 452], the Lagrange multiplier imposing the constraint, which has the physical interpretation of a generalized constraint force, is treated on equal footing with generalized coordinates and generalized momenta and therefore parametrized using (discontinuous) basis functions. As a result, the constraint is not enforced exactly at all time instants but is fulfilled only at certain predefined time instants (provided that the basis functions fulfill the basic assumption of containing at least a constant element).

1.2. Outline

The paper is organized as follows: In Section 2, we introduce the principle of virtual action, that is, the virtual work integrated over time. By inserting discontinuous ansatz functions into the principle of virtual action, a set of possibly nonlinear equations is obtained approximating the dynamic equilibrium. The properties of the resulting numerical integrator are analyzed in Section 3. In particular, conditions are established for which an energy-like quantity is conserved and symplectic integration is demonstrated. This provides a geometric characterization of the flow. For example, it follows that the area enclosed by an arbitrary closed contour in the phase space remains constant as the contour evolves in time. Moreover, the approximate solutions can be regarded as the exact solutions of a slightly perturbed Hamiltonian system. In Section 4, the integration framework is extended to incorporate holonomic constraints by simply accounting for the virtual work exerted by the constraint. The resulting integration scheme uses no additional regularization and will be shown to conserve the gap function from one timestep to the next (under some weak assumptions). In Section 5, the relation to partitioned Runge–Kutta methods is shortly discussed. The presented integration methods are illustrated on the example of the planar pendulum in Section 6. In a first step, polar coordinates are used to model the pendulum, and the presented method is compared with the Galerkin variational integrator approach from [1, p. 415]. In a second step, the addition of holonomic constraints is demonstrated by using a parametrization based on Cartesian coordinates. The obtained simulation results are discussed subsequently. Finally, the article concludes with a summary and outlook in Section 7.

DOI: 10.1002/nme

2. DERIVATION OF VARIATIONAL INTEGRATORS

2.1. Notation and definitions

The class of continuously differentiable functions mapping an interval $I \subset \mathbb{R}$ to the Euclidean space $\mathbb{R}^n$ is denoted by $C^1(I, \mathbb{R}^n)$, a vector-valued quantity is written in boldface, and ansatz functions are marked with a tilde, that is, $\tilde{\boldsymbol{q}}$ denotes an ansatz for $\boldsymbol{q}$.

Additionally, left and right limits of a discontinuous function $\boldsymbol{f} : I \to \mathbb{R}^n$ at the discontinuity point $t = t_d$ are denoted by

$$\boldsymbol{f}(t_d)^+ = \lim_{t \downarrow t_d} \boldsymbol{f}(t), \quad \text{respectively} \quad \boldsymbol{f}(t_d)^- = \lim_{t \uparrow t_d} \boldsymbol{f}(t). \tag{1}$$

The set of all piecewise continuously differentiable functions mapping an interval $I = [t_0, t_N] \subset \mathbb{R}$ to the Euclidean space $\mathbb{R}^n$ is denoted by $C^1_{\text{pc}}(I, \mathbb{R}^n)$. Hence, given a function $\boldsymbol{f} \in C^1_{\text{pc}}([t_0, t_N], \mathbb{R}^n)$, there exists a finite number of discontinuous time instants $t_0 \leqslant t_1 < t_2 < \cdots \leqslant t_N$ such that $\boldsymbol{f}$ is continuously differentiable on the intervals (t_i, t_{i+1}), $i = 0, 1, \ldots, N-1$. A typical function $f \in C^1_{\text{pc}}([t_0, t_N], \mathbb{R})$ is depicted in Figure 1. At the discontinuous time instant $t = t_d$, left and right limits exist, which, in general, do not agree with the function value at $t = t_d$ nor with each other, that is, $f(t_d)^- \neq f(t_d) \neq f(t_d)^+$.

The configuration of the mechanical system is expressed in local coordinates by $\boldsymbol{q}(t)$ with $\boldsymbol{q} \in C^1_{\text{pc}}(I, \mathbb{R}^n)$ and $I = [t_0, t_N] \subset \mathbb{R}$. The generalized momenta are expressed in local coordinates by $\boldsymbol{p}(t)$ with $\boldsymbol{p} \in C^1_{\text{pc}}(I, (\mathbb{R}^n)^*)$, where $(\mathbb{R}^n)^*$ is the dual of $\mathbb{R}^n$. Note that $\boldsymbol{q}$ and $\boldsymbol{p}$ are almost everywhere continuously differentiable. In the following, the dual space $(\mathbb{R}^n)^*$ is identified with $\mathbb{R}^n$ such that both elements $\boldsymbol{q}(t) \in \mathbb{R}^n$ and $\boldsymbol{p}(t) \in (\mathbb{R}^n)^*$ are written as column vectors. The generalized momenta are defined almost everywhere by

$$\boldsymbol{p} := \boldsymbol{M}(\boldsymbol{q})\dot{\boldsymbol{q}}, \tag{2}$$

where $\boldsymbol{M}(\boldsymbol{q})$ denotes the symmetric, positive-definite mass matrix. Moreover, the associated Hamiltonian is given by

$$H(\boldsymbol{p}, \boldsymbol{q}, t) = \frac{1}{2}\boldsymbol{p}^\top \boldsymbol{M}(\boldsymbol{q})^{-1}\boldsymbol{p} + V(\boldsymbol{q}, t), \tag{3}$$

where $V(\boldsymbol{q}, t)$ represents the potential energy. A derivation of these expressions for multibody systems can be found in [21].

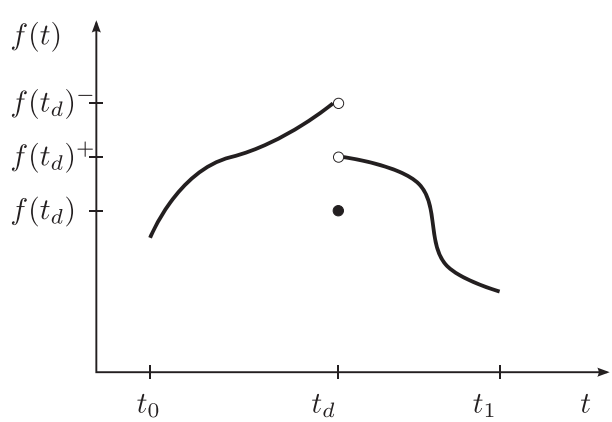

Figure 1. A possible example of a function $f \in C^1_{\text{pc}}([t_0, t_1], \mathbb{R})$, which is discontinuous at time $t = t_d$.

DOI: 10.1002/nme

Weak variations of a function $\boldsymbol{f} \in C^1_{\mathrm{pc}}(I, \mathbb{R}^n)$, denoted by $\delta \boldsymbol{f}$, are defined as

$$\delta \boldsymbol{f}(t) := \left.\frac{\mathrm{d}\hat{\boldsymbol{f}}}{\mathrm{d}\varepsilon}\right|_{\varepsilon_0} (\varepsilon - \varepsilon_0) = \boldsymbol{v}(t)(\varepsilon - \varepsilon_0), \tag{4}$$

where $\hat{\boldsymbol{f}} : \mathbb{R} \times I \to \mathbb{R}^n$, $\hat{\boldsymbol{f}}(\varepsilon, t) = \boldsymbol{f}(t) + (\varepsilon - \varepsilon_0)\boldsymbol{v}(t)$ with $\boldsymbol{v} \in C^1_{\mathrm{pc}}(I, \mathbb{R}^n)$. These variations are referred to as weak because they converge in a weak sense, that is, $\lim_{\varepsilon \downarrow 0} ||\hat{\boldsymbol{f}} - \boldsymbol{f}||_1 = 0$ (e.g., [22]), where $||\cdot||_1$ is defined as

$$||\boldsymbol{f}||_1 = \sup_{t \in I} |\boldsymbol{f}(t)| + \sup_{t \in I} |\dot{\boldsymbol{f}}(t)|. \tag{5}$$

By restricting ourselves to weak variations, we avoid explicitly the use of variations induced by comparison functions of the form

$$\hat{\boldsymbol{f}}(\varepsilon, t) = \boldsymbol{f}\left(t - \hat{t}(\varepsilon, t)\right) + \boldsymbol{v}\left(t - \hat{t}(\varepsilon, t)\right)(\varepsilon - \varepsilon_0), \quad \text{with} \quad \hat{\boldsymbol{f}}(\varepsilon_0, t) = \boldsymbol{f}(t), \quad \forall t \in I. \tag{6}$$

These variations do not converge in the weak norm because they induce a time and value shift [21]. In particular, the time shift would introduce couplings between two neighboring time intervals separated by a discontinuity.[‡] These couplings need to be avoided in order to obtain a causal integration algorithm.

2.2. *Virtual action*

The starting point of the following derivation represents the principle of virtual action, where the virtual action is defined as the virtual work integrated over time. As piecewise continuously differentiable ansatz functions are used, similar to [9], the principle of virtual action is introduced in a form where the generalized coordinates and generalized momenta are not required to be continuous. It is therefore postulated that

Postulate 2.1
(Principle of virtual action) Let $\boldsymbol{q} \in C^1_{\mathrm{pc}}(I, \mathbb{R}^n)$, $\boldsymbol{p} \in C^1_{\mathrm{pc}}(I, (\mathbb{R}^n)^*)$, with $I = [t_0, t_N] \subset \mathbb{R}$. The non-potential forces $\boldsymbol{f}_{NP}(t) \in \mathbb{R}^n$ are assumed to be absolutely continuous for all $t \in I$. If the virtual action expressed by

$$\delta A := \int_I \delta \boldsymbol{q}^\top \mathrm{d}\boldsymbol{p} - \delta \boldsymbol{p}^\top \mathrm{d}\boldsymbol{q} + \left(\delta H(\boldsymbol{p}, \boldsymbol{q}, t) - \delta \boldsymbol{q}^\top \boldsymbol{f}_{NP}(t)\right) \mathrm{d}t, \tag{7}$$

where the Hamiltonian $H(\boldsymbol{p}, \boldsymbol{q}, t)$ is defined almost everywhere by Equation (3) vanishes for all variations $\delta \boldsymbol{q}$ and $\delta \boldsymbol{p}$, then the system is almost everywhere in dynamic equilibrium for times $t \in I$.

Note that $\mathrm{d}\boldsymbol{q}$ and $\mathrm{d}\boldsymbol{p}$ refer to the differential measure of $\boldsymbol{q} \in C^1_{\mathrm{pc}}(I, \mathbb{R}^n)$ and $\boldsymbol{p} \in C^1_{\mathrm{pc}}(I, (\mathbb{R}^n)^*)$. The differential measure contains a density with respect to the Lebesgue measure and with respect to an atomic measure [21]. The Lebesgue measure is denoted by $\mathrm{d}t$.

This postulate can be derived from the principle of the virtual work in the following way: (1) introduce local generalized coordinates $\boldsymbol{q}(t) \in \mathbb{R}^n$, which describe the motion of each material point of the mechanical system; (2) define the virtual action to be the integral of the virtual work over time. Owing to the fundamental lemma of the calculus of variations, requiring the virtual action to vanish for all weak virtual variation $\delta \boldsymbol{q}$ is equivalent to the principle of virtual work (almost everywhere). Note that we do not require $\delta \boldsymbol{q}$ to vanish at the time interval boundaries, that is, $\delta \boldsymbol{q}(t_0)$

[‡]Variations with respect to time would impose the second Weierstrass–Erdmann condition at time instants where $\boldsymbol{q}$ or $\boldsymbol{p}$ is not continuous [21].

DOI: 10.1002/nme

and $\delta \boldsymbol{q}(t_N)$ are not required to vanish[§]; (3) decompose the external and internal forces acting on each material point in potential and non-potential forces. Under the assumption that variation and time differentiation commute, the virtual work can be split into the variation of the Lagrangian (denoted by L), the virtual work exerted by the non-potential forces, and a remaining term, leading to $\delta W \, \mathrm{d}t = \mathrm{d}\left(\delta \boldsymbol{q}^\top \boldsymbol{p}\right) - \delta L \, \mathrm{d}t - \delta \boldsymbol{q}^\top \boldsymbol{f}_{NP} \, \mathrm{d}t$, for example, [21]. The obtained expression is commonly referred to as Lagrange's central equation [23, 24]. Note that $\boldsymbol{p}$ is not yet identified as independent quantity but is considered as placeholder for $\boldsymbol{M}(\boldsymbol{q})\dot{\boldsymbol{q}}$ according to Equation (2); (4) perform the Legendre transformation of the Lagrangian with respect to $\dot{\boldsymbol{q}}$ to obtain the Hamiltonian and introduce the generalized momenta $\boldsymbol{p}$. The kinematic link between the generalized momenta $\boldsymbol{p}$ and the generalized velocities $\dot{\boldsymbol{q}}$ can be enforced by requiring the virtual action to vanish for variations of the generalized momenta. Thus, the change from Lagrangian to Hamiltonian formalism is performed by regarding generalized coordinates and momenta as a priori independent and introducing independent variations $\delta \boldsymbol{p}$.

A proof of Postulate 2.1 for smooth systems can be found in Appendix A.

Note that the use of the differential measures $\mathrm{d}\boldsymbol{q}$ and $\mathrm{d}\boldsymbol{p}$ has the effect of forcing left and right limits of $\boldsymbol{q}$ and $\boldsymbol{p}$ to agree at discontinuous time instants. This fact is shortly illustrated by considering the particular variations $\delta \boldsymbol{q}, \delta \boldsymbol{p}$, which are everywhere zero, except at the time instant $t = t_d$. The virtual action then simplifies to the integral over the time singleton $\{t_d\}$. Using linearity of the integral and recalling that the differential measure has a density with respect to an atomic measure, we obtain

$$\begin{aligned} \delta A &= \int_{\{t_d\}} \delta \boldsymbol{q}(t_d)^\top \mathrm{d}\boldsymbol{p} - \delta \boldsymbol{p}(t_d)^\top \mathrm{d}\boldsymbol{q} = \delta \boldsymbol{q}(t_d)^\top \int_{\{t_d\}} \mathrm{d}\boldsymbol{p} - \delta \boldsymbol{p}(t_d)^\top \int_{\{t_d\}} \mathrm{d}\boldsymbol{q} \\ &= \delta \boldsymbol{q}(t_d)^\top \left(\boldsymbol{p}(t_d)^+ - \boldsymbol{p}(t_d)^-\right) - \delta \boldsymbol{p}(t_d)^\top \left(\boldsymbol{q}(t_d)^+ - \boldsymbol{q}(t_d)^-\right). \end{aligned} \tag{8}$$

Requiring the virtual action to vanish for all variations $\delta \boldsymbol{q}(t_d)$ and $\delta \boldsymbol{p}(t_d)$ implies therefore the agreement of the left and right limits of $\boldsymbol{q}$ and $\boldsymbol{p}$ at a possible discontinuity.

Nonetheless, at first sight, the derivation of Postulate 2.1 from the principle of virtual work seems rather lengthy and involved. The motivation is twofold: By relaxing continuity requirements and allowing discontinuous trajectories, we are eventually able to decouple the time intervals in-between discontinuities leading to a causal and iterative integration algorithm. Moreover, the change to Hamiltonian formalism offers additional flexibility in the approximation of the equations of motion for the purpose of numerical integration. This comes at the cost of only approximately fulfilling the kinematic link between generalized velocities and generalized momenta given by Equation (2).

To evaluate the principle of virtual action given by Postulate 2.1, infinitely many test functions need to be compared, and in order to find the trajectories compatible with the dynamic equilibrium, the virtual action has to vanish for each one of them. The approach to variational integration presented herein is based on evaluating the principle of virtual action only for test functions parametrized by a given ansatz; that is, the virtual action will be required to vanish only for variations of the ansatz parameters. Equivalently, we introduce constraint forces, forcing the trajectories to lie within the class of functions spanned by the ansatz. By the principle of d'Alembert–Lagrange, the virtual work is required to vanish for admissible variations, which are in this case given by variations of the ansatz parameters only. It will be shown that this approach leads for a well-chosen ansatz to convergence and even to symplectic integration.

2.3. The parametrization of ansatz functions

Let $\tilde{C}^1_{\mathrm{pc}}(I \times \mathbb{R}^m, \mathbb{R}^n)$, with $I \subset \mathbb{R}$, be the space of piecewise continuously differentiable ansatz functions defined by

[§]The fundamental lemma of the calculus of variations is commonly formulated for vanishing variations at the end points. As this represents a special case of variations with non-vanishing end points, the result holds likewise.

DOI: 10.1002/nme

$$\tilde{C}^1_{\mathrm{pc}}(I \times \mathbb{R}^m, \mathbb{R}^n) := \Big\{ \boldsymbol{y} : I \times \mathbb{R}^m \to \mathbb{R}^n \,|\, \boldsymbol{y}(\cdot, \boldsymbol{\eta}_y) \in C^1_{\mathrm{pc}}(I, \mathbb{R}^n),\ \forall \boldsymbol{\eta}_y \in \mathbb{R}^m, \\ \boldsymbol{y}(t, \cdot) : \mathbb{R}^m \to \mathbb{R}^n \text{ smooth } \forall t \in I \Big\}.$$

Moreover, an ansatz for $\boldsymbol{q}$ on the interval I will be denoted by $\tilde{\boldsymbol{q}}(t, \boldsymbol{\eta}_q)$, where $\boldsymbol{\eta}_q$ is the parameter vector. Similarly, an ansatz for $\boldsymbol{p}$ will be denoted by $\tilde{\boldsymbol{p}}(t, \boldsymbol{\eta}_p)$, with $\boldsymbol{\eta}_p$ the ansatz parameters. In general, the number of parameters, that is, the dimension of $\boldsymbol{\eta}_q$ and $\boldsymbol{\eta}_p$, does not need to agree. To simplify notation, we suppress occasionally the dependence on the parameter vector; that is, $\tilde{\boldsymbol{q}}(t, \boldsymbol{\eta}_q)$ will be denoted by $\tilde{\boldsymbol{q}}(t)$ and likewise $\tilde{\boldsymbol{p}}(t, \boldsymbol{\eta}_p)$ by $\tilde{\boldsymbol{p}}(t)$.

As an example, consider the interval $[0, 2]$ and the case $n = 1$: A piecewise polynomial ansatz of order m for q is given by

$$\tilde{q}(t, \boldsymbol{\eta}_q) = \begin{cases} \eta_{q_0} + \eta_{q_1}\, t + \eta_{q_2}\, t^2 + \ldots + \eta_{q_m}\, t^m, & t \in [0, 1] \\ \eta_{q_{m+1}} + \eta_{q_{m+2}}\, t + \eta_{q_{m+3}}\, t^2 + \ldots + \eta_{q_{2m+1}}\, t^m, & t \in (1, 2] \end{cases} \tag{9}$$

where $\boldsymbol{\eta}_q = \left(\eta_{q_0}, \eta_{q_1}, \ldots, \eta_{q_{2m+1}}\right)^{\mathsf{T}} \in \mathbb{R}^{2m+2}$ is the parameter vector.

Next, variations restricted to a class of ansatz functions are introduced. Let $\tilde{\boldsymbol{f}} \in \tilde{C}^1_{\mathrm{pc}}(I \times \mathbb{R}^m, \mathbb{R}^n)$ be an ansatz and $\boldsymbol{\eta}_f : \mathbb{R} \to \mathbb{R}^m$ be a continuously differentiable function. Variations restricted to this ansatz will be denoted by

$$\delta \tilde{\boldsymbol{f}}(t) := \left.\frac{\mathrm{d}\tilde{\boldsymbol{f}}(t, \boldsymbol{\eta}_f(\varepsilon))}{\mathrm{d}\varepsilon}\right|_{\varepsilon_0} (\varepsilon - \varepsilon_0) = \left.\frac{\partial \tilde{\boldsymbol{f}}}{\partial \boldsymbol{\eta}_f} \frac{\partial \boldsymbol{\eta}_f}{\partial \varepsilon}\right|_{\varepsilon_0} (\varepsilon - \varepsilon_0) = \left.\frac{\partial \tilde{\boldsymbol{f}}}{\partial \boldsymbol{\eta}_f}\right|_{\boldsymbol{\eta}_f(\varepsilon_0)} \delta \boldsymbol{\eta}_f. \tag{10}$$

Note that the variations $\delta \boldsymbol{\eta}_f : \mathbb{R} \to \mathbb{R}^m$ are therefore defined as

$$\delta \boldsymbol{\eta}_f(\varepsilon) := \left.\frac{\partial \boldsymbol{\eta}_f}{\partial \varepsilon}\right|_{\varepsilon_0} (\varepsilon - \varepsilon_0) \tag{11}$$

and to simplify notation, the dependence on ε is omitted. The variations $\delta \boldsymbol{\eta}_f$ are in general unrelated to the parameter values $\boldsymbol{\eta}_f(\varepsilon_0)$. For ease of notation, $\boldsymbol{\eta}_f(\varepsilon_0)$ is simply written as $\boldsymbol{\eta}_f$ and can be thought of the parameter vector.

For the previous example of $\tilde{q}(t, \boldsymbol{\eta}_q)$, (restricted) variations are given by

$$\delta \tilde{q} = \begin{cases} \delta\eta_{q_0} + \delta\eta_{q_1}\, t + \delta\eta_{q_2}\, t^2 + \ldots + \delta\eta_{q_m}\, t^m, & t \in [0, 1] \\ \delta\eta_{q_{m+1}} + \delta\eta_{q_{m+2}}\, t + \delta\eta_{q_{m+3}}\, t^2 + \ldots + \delta\eta_{q_{2m+1}}\, t^m, & t \in (1, 2]\,. \end{cases} \tag{12}$$

2.4. The variational integration approach

The presented approach is similar to the derivation of the finite elements method in elastomechanics. By restricting variations to be nonzero in an interval where the ansatz is continuously differentiable, a time segment of the virtual action is cut out. Requiring the virtual action to vanish for these variations leads to a set of nonlinear equations describing the evolution of the approximate trajectories from one time interval to the next.

Depending on the continuity assumptions of the ansatz functions, different integration schemes can be derived. To fix the ideas, we will make the restriction of $\tilde{\boldsymbol{q}}(t, \boldsymbol{\eta}_q)$ being left continuous and $\tilde{\boldsymbol{p}}(t, \boldsymbol{\eta}_p)$ being right continuous (for all parameter vectors $\boldsymbol{\eta}_q$ and $\boldsymbol{\eta}_p$). However, the presented framework extends naturally to various other cases, including the configuration where $\tilde{\boldsymbol{q}}$ and $\tilde{\boldsymbol{p}}$ are assumed to be both left continuous or both right continuous. In Appendix B, the derivation is presented for the case $\tilde{\boldsymbol{q}}$ continuous and $\tilde{\boldsymbol{p}}$ neither left nor right continuous.

The time interval of interest, $I = [t_0, t_N]$ with $t_N > t_0$, is divided into sub-intervals $I_i := [t_i, t_{i+1}] \subset \mathbb{R}$ $(i = 0, 1, \ldots, N-1)$, where each sub-interval $I_i := [t_i, t_{i+1}] \subset \mathbb{R}$ is chosen such

DOI: 10.1002/nme

that $\bigcup_{i=0}^{N-1}(I_i \setminus \{t_{i+1}\}) = I \setminus \{t_N\}$. The two ansatz functions $\tilde{\boldsymbol{q}}(t)$ and $\tilde{\boldsymbol{p}}(t)$ for $\boldsymbol{q}$ and $\boldsymbol{p}$ are required to be everywhere continuously differentiable, except at the time instants $t_0, t_1, \ldots, t_N$. Hence, they are piecewise continuously differentiable, with

$$\tilde{\boldsymbol{q}} \in \tilde{C}^1_{\mathrm{pc}}\left(I \times \mathbb{R}^{n \cdot m_q \cdot N}, \mathbb{R}^n\right), \quad \text{and} \quad \tilde{\boldsymbol{p}} \in \tilde{C}^1_{\mathrm{pc}}\left(I \times \mathbb{R}^{n \cdot m_p \cdot N}, (\mathbb{R}^n)^*\right). \tag{13}$$

An example of the configuration is depicted in Figure 2.

Inserting the ansatz into the expression of the virtual action leads to

$$\delta A = \int_I \delta\tilde{\boldsymbol{q}}^\top \mathrm{d}\tilde{\boldsymbol{p}} - \delta\tilde{\boldsymbol{p}}^\top \mathrm{d}\tilde{\boldsymbol{q}} + \left(\delta H(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{q}}, t) - \delta\tilde{\boldsymbol{q}}^\top \boldsymbol{f}_{NP}\right) \mathrm{d}t. \tag{14}$$

We require the virtual action to vanish for all variations $\delta\tilde{\boldsymbol{q}}$ and $\delta\tilde{\boldsymbol{p}}$ to obtain an approximation of the dynamic equilibrium. Because of the discontinuous ansatz, this is equivalent to requiring the virtual action to vanish for all variations $\delta\tilde{\boldsymbol{q}}$ and $\delta\tilde{\boldsymbol{p}}$, which are everywhere zero except on the interval I_i, and then expanding the requirement to all intervals I_i, with $i = 0, 1, \ldots, N-1$. Requiring the virtual action to vanish for variations that are zero for $t \notin I_i$ simplifies Equation (14) to

$$\begin{aligned}
\delta A &= \int_{I_i} \delta\tilde{\boldsymbol{q}}^\top \mathrm{d}\tilde{\boldsymbol{p}} - \delta\tilde{\boldsymbol{p}}^\top \mathrm{d}\tilde{\boldsymbol{q}} + \left(\delta H(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{q}}, t) - \delta\tilde{\boldsymbol{q}}^\top \boldsymbol{f}_{NP}\right) \mathrm{d}t \\
&= \int_{\{t_i\}} \delta\tilde{\boldsymbol{q}}^\top \mathrm{d}\tilde{\boldsymbol{p}} - \delta\tilde{\boldsymbol{p}}^\top \mathrm{d}\tilde{\boldsymbol{q}} + \int_{\{t_{i+1}\}} \delta\tilde{\boldsymbol{q}}^\top \mathrm{d}\tilde{\boldsymbol{p}} - \delta\tilde{\boldsymbol{p}}^\top \mathrm{d}\tilde{\boldsymbol{q}} \\
&\quad + \int_{(t_i, t_{i+1})} \left(\delta\tilde{\boldsymbol{q}}^\top \dot{\tilde{\boldsymbol{p}}} - \delta\tilde{\boldsymbol{p}}^\top \dot{\tilde{\boldsymbol{q}}} + \delta H(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{q}}, t) - \delta\tilde{\boldsymbol{q}}^\top \boldsymbol{f}_{NP}\right) \mathrm{d}t.
\end{aligned}$$

The differential measures $\mathrm{d}\tilde{\boldsymbol{q}}$ and $\mathrm{d}\tilde{\boldsymbol{p}}$ are expressed by $\dot{\tilde{\boldsymbol{q}}}\,\mathrm{d}t$ and $\dot{\tilde{\boldsymbol{p}}}\,\mathrm{d}t$ in the interval (t_i, t_{i+1}), where $\tilde{\boldsymbol{q}}$ and $\tilde{\boldsymbol{p}}$ are continuously differentiable. Because $\tilde{\boldsymbol{q}}$ is left continuous (for any parameter vector $\boldsymbol{\eta}_q$), the variation $\delta\tilde{\boldsymbol{q}}$ is left continuous as well. The variation $\delta\tilde{\boldsymbol{q}}(t)$ was chosen to vanish for $t \notin I_i$ and hence, $\delta\tilde{\boldsymbol{q}}(t_i)^- = 0$. By left continuity of $\delta\tilde{\boldsymbol{q}}$, this implies $\delta\tilde{\boldsymbol{q}}(t_i)^- = \delta\tilde{\boldsymbol{q}}(t_i) = 0$. Similarly, $\tilde{\boldsymbol{p}}$ is right continuous, which leads to $\delta\tilde{\boldsymbol{p}}(t_{i+1}) = 0$. Thus, the expression of the virtual action simplifies further to

$$\delta A = -\int_{\{t_i\}} \delta\tilde{\boldsymbol{p}}^\top \mathrm{d}\tilde{\boldsymbol{q}} + \int_{\{t_{i+1}\}} \delta\tilde{\boldsymbol{q}}^\top \mathrm{d}\tilde{\boldsymbol{p}} + \int_{(t_i, t_{i+1})} \left(\delta\tilde{\boldsymbol{q}}^\top \dot{\tilde{\boldsymbol{p}}} - \delta\tilde{\boldsymbol{p}}^\top \dot{\tilde{\boldsymbol{q}}} + \delta H(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{q}}, t) - \delta\tilde{\boldsymbol{q}}^\top \boldsymbol{f}_{NP}\right) \mathrm{d}t.$$

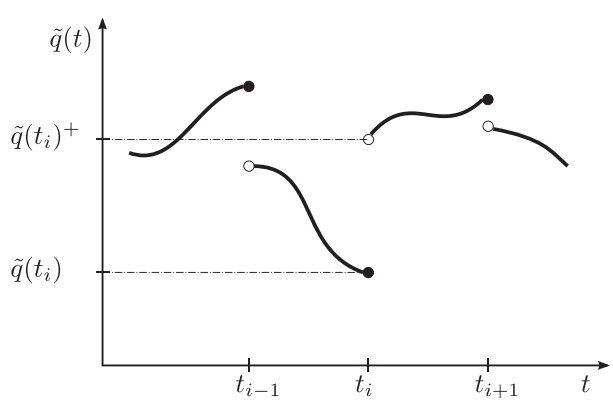

Figure 2. A schematic example of a possible ansatz function $\tilde{q}$ for q. The ansatz is left continuous, and within the intervals $(t_{i-1}, t_i]$ and $(t_i, t_{i+1}]$, $\tilde{q}$ is continuously differentiable.

DOI: 10.1002/nme

Linearity of the integral can be used to reformulate the previous expression:

$$\delta A = -\delta\tilde{\boldsymbol{p}}(t_i)^{\mathsf{T}} \int\limits_{\{t_i\}} \mathrm{d}\tilde{\boldsymbol{q}} + \delta\tilde{\boldsymbol{q}}(t_{i+1})^{\mathsf{T}} \int\limits_{\{t_{i+1}\}} \mathrm{d}\tilde{\boldsymbol{p}} + \int\limits_{(t_i,t_{i+1})} \left(\delta\tilde{\boldsymbol{q}}^{\mathsf{T}}\dot{\tilde{\boldsymbol{p}}} - \delta\tilde{\boldsymbol{p}}^{\mathsf{T}}\dot{\tilde{\boldsymbol{q}}} + \delta H\left(\tilde{\boldsymbol{p}},\tilde{\boldsymbol{q}},t\right) - \delta\tilde{\boldsymbol{q}}^{\mathsf{T}}\boldsymbol{f}_{NP}\right)\mathrm{d}t,$$

leading to the strong form of the virtual action given by

$$\begin{aligned}\delta A = &-\delta\tilde{\boldsymbol{p}}(t_i)^{\mathsf{T}}\left(\tilde{\boldsymbol{q}}(t_i)^{+} - \tilde{\boldsymbol{q}}(t_i)^{-}\right) + \delta\tilde{\boldsymbol{q}}(t_{i+1})^{\mathsf{T}}\left(\tilde{\boldsymbol{p}}(t_{i+1})^{+} - \tilde{\boldsymbol{p}}(t_{i+1})^{-}\right)\\ &+ \int\limits_{(t_i,t_{i+1})} \left(\delta\tilde{\boldsymbol{q}}^{\mathsf{T}}\dot{\tilde{\boldsymbol{p}}} - \delta\tilde{\boldsymbol{p}}^{\mathsf{T}}\dot{\tilde{\boldsymbol{q}}} + \delta H(\tilde{\boldsymbol{p}},\tilde{\boldsymbol{q}},t) - \delta\tilde{\boldsymbol{q}}^{\mathsf{T}}\boldsymbol{f}_{NP}\right)\mathrm{d}t.\end{aligned} \tag{15}$$

Applying integration by parts on the integrand $\delta\tilde{\boldsymbol{q}}^{\mathsf{T}}\dot{\tilde{\boldsymbol{p}}}$ yields the weak form:

$$\begin{aligned}\delta A = &-\delta\tilde{\boldsymbol{p}}(t_i)^{\mathsf{T}}\left(\tilde{\boldsymbol{q}}(t_i)^{+} - \tilde{\boldsymbol{q}}(t_i)^{-}\right) - (\delta\tilde{\boldsymbol{q}}(t_i)^{+})^{\mathsf{T}}\tilde{\boldsymbol{p}}(t_i)^{+} + \delta\tilde{\boldsymbol{q}}(t_{i+1})^{\mathsf{T}}\tilde{\boldsymbol{p}}(t_{i+1})^{+}\\ &- \int\limits_{(t_i,t_{i+1})} \left(\delta(\dot{\tilde{\boldsymbol{q}}}^{\mathsf{T}}\tilde{\boldsymbol{p}} - H(\tilde{\boldsymbol{p}},\tilde{\boldsymbol{q}},t)) + \delta\tilde{\boldsymbol{q}}^{\mathsf{T}}\boldsymbol{f}_{NP}\right)\mathrm{d}t.\end{aligned} \tag{16}$$

The term $-\delta\tilde{\boldsymbol{q}}(t_{i+1})^{\mathsf{T}}\tilde{\boldsymbol{p}}(t_{i+1})^{-} + (\delta\tilde{\boldsymbol{q}}(t_{i+1})^{-})^{\mathsf{T}}\tilde{\boldsymbol{p}}(t_{i+1})^{-}$ vanishes because of the left continuity of $\tilde{\boldsymbol{q}}$, asserting that $\delta\tilde{\boldsymbol{q}}(t_{i+1})^{-} = \delta\tilde{\boldsymbol{q}}(t_{i+1})$. Note that the differentiability requirement of $\tilde{\boldsymbol{p}}$ could potentially be alleviated because only the time derivative $\dot{\tilde{\boldsymbol{q}}}$ appears in the weak form of the virtual action. If both $\tilde{\boldsymbol{q}}$ and $\tilde{\boldsymbol{p}}$ are piecewise continuously differentiable, the weak and the strong forms are equivalent.

Recall that the variations were chosen everywhere zero except for $t \in [t_i, t_{i+1}]$. In the following, it is assumed that within the interval $t \in (t_i, t_{i+1}]$, the ansatz $\tilde{\boldsymbol{q}}$ depends on time but only on a subset of the parameters $\boldsymbol{\eta}_q$, denoted by $\boldsymbol{\eta}_{q_i} \in \mathbb{R}^{n\cdot m_q}$. Similarly, within the interval $t \in [t_i, t_{i+1})$, $\tilde{\boldsymbol{p}}$ is assumed to depend on time but only on a subset of the parameters $\boldsymbol{\eta}_p$, denoted by $\boldsymbol{\eta}_{p_i} \in \mathbb{R}^{n\cdot m_p}$. This simplifies the variations $\delta\tilde{\boldsymbol{q}}$ and $\delta\tilde{\boldsymbol{p}}$, which were chosen to vanish for $t \notin [t_i, t_{i+1}]$ to

$$\delta\tilde{\boldsymbol{q}}(t) = \begin{cases} \frac{\partial\tilde{\boldsymbol{q}}}{\partial\boldsymbol{\eta}_{q_i}}\delta\boldsymbol{\eta}_{q_i}, & t \in (t_i, t_{i+1}] \\ 0, & \text{else} \end{cases} \quad \text{and} \quad \delta\tilde{\boldsymbol{p}}(t) = \begin{cases} \frac{\partial\tilde{\boldsymbol{p}}}{\partial\boldsymbol{\eta}_{p_i}}\delta\boldsymbol{\eta}_{p_i}, & t \in [t_i, t_{i+1}) \\ 0, & \text{else} \end{cases}. \tag{17}$$

Therefore, by factorizing the time-independent parameter variations $\delta\boldsymbol{\eta}_{q_i}$ and $\delta\boldsymbol{\eta}_{p_i}$ out, the virtual action can be rewritten as

$$\delta A = \delta\boldsymbol{\eta}_{q_i}^{\mathsf{T}}\boldsymbol{A}_q\left(\boldsymbol{\eta}_{q_i}, \boldsymbol{\eta}_{p_i}, \tilde{\boldsymbol{p}}(t_{i+1})^{+}\right) + \delta\boldsymbol{\eta}_{p_i}^{\mathsf{T}}\boldsymbol{A}_p\left(\boldsymbol{\eta}_{q_i}, \boldsymbol{\eta}_{p_i}, \tilde{\boldsymbol{q}}(t_i)\right), \tag{18}$$

with

$$\begin{aligned}\boldsymbol{A}_q\left(\boldsymbol{\eta}_{q_i}, \boldsymbol{\eta}_{p_i}, \tilde{\boldsymbol{p}}(t_{i+1})^{+}\right) := &-\left.\frac{\partial\tilde{\boldsymbol{q}}}{\partial\boldsymbol{\eta}_{q_i}}\right|_{t\downarrow t_i}^{\mathsf{T}}\tilde{\boldsymbol{p}}(t_i)^{+} + \left.\frac{\partial\tilde{\boldsymbol{q}}}{\partial\boldsymbol{\eta}_{q_i}}\right|_{t=t_{i+1}}^{\mathsf{T}}\tilde{\boldsymbol{p}}(t_{i+1})^{+}\\ &- \int\limits_{(t_i,t_{i+1})}\left[\frac{\partial\dot{\tilde{\boldsymbol{q}}}}{\partial\boldsymbol{\eta}_{q_i}}^{\mathsf{T}}\tilde{\boldsymbol{p}} - \frac{\partial\tilde{\boldsymbol{q}}}{\partial\boldsymbol{\eta}_{q_i}}^{\mathsf{T}}\left(\frac{\partial H}{\partial\boldsymbol{q}}^{\mathsf{T}} - \boldsymbol{f}_{NP}\right)\right]\mathrm{d}t \in \mathbb{R}^{n\cdot m_q}\end{aligned} \tag{19}$$

and

DOI: 10.1002/nme

$$\begin{aligned} \boldsymbol{A}_p\left(\boldsymbol{\eta}_{q_i}, \boldsymbol{\eta}_{p_i}, \tilde{\boldsymbol{q}}(t_i)^-\right) := &- \left.\frac{\partial \tilde{\boldsymbol{p}}}{\partial \boldsymbol{\eta}_{p_i}}\right|_{t=t_i}^{\mathsf{T}} \left(\tilde{\boldsymbol{q}}(t_i)^+ - \tilde{\boldsymbol{q}}(t_i)^-\right) \\ &- \int\limits_{(t_i, t_{i+1})} \frac{\partial \tilde{\boldsymbol{p}}}{\partial \boldsymbol{\eta}_{p_i}}^{\mathsf{T}} \left(\dot{\tilde{\boldsymbol{q}}} - \frac{\partial H}{\partial \boldsymbol{p}}^{\mathsf{T}}\right) \mathrm{d}t \in \mathbb{R}^{n \cdot m_p}. \end{aligned} \tag{20}$$

Note that $\tilde{\boldsymbol{q}}(t_i)^+$ and $\tilde{\boldsymbol{p}}(t_i)^+$ are uniquely determined by the parameter vectors $\boldsymbol{\eta}_{q_i}$, $\boldsymbol{\eta}_{p_i}$, that is, $\tilde{\boldsymbol{q}}(t_i)^+ = \lim_{t \downarrow t_i} \tilde{\boldsymbol{q}}(t, \boldsymbol{\eta}_{q_i})$ and $\tilde{\boldsymbol{p}}(t_i)^+ = \lim_{t \downarrow t_i} \tilde{\boldsymbol{p}}(t, \boldsymbol{\eta}_{p_i})$ (c.f. Equation (17)). As a consequence, $\boldsymbol{A}_q$ and $\boldsymbol{A}_p$ depend solely on the parameters $\boldsymbol{\eta}_{q_i}$, $\boldsymbol{\eta}_{p_i}$ and on the values $\tilde{\boldsymbol{q}}(t_i)^-$ and $\tilde{\boldsymbol{p}}(t_{i+1})^+$, describing the coupling to the neighboring time intervals. Although the states $\tilde{\boldsymbol{q}}(t_i)^-$ and $\tilde{\boldsymbol{p}}(t_{i+1})^+$ depend on the parameters $\boldsymbol{\eta}_{q_{i-1}}$ and $\boldsymbol{\eta}_{p_{i+1}}$, they are regarded as vectors in $\mathbb{R}^n$ and $(\mathbb{R}^n)^*$, respectively, when solving for the trajectories in the interval $[t_i, t_{i+1}]$ in order to obtain a causal integration scheme. More precisely, from Equation (18), it follows that δA vanishes for all $\delta\boldsymbol{\eta}_{q_i}$ and $\delta\boldsymbol{\eta}_{p_i}$ if and only if the set of equations

$$\begin{aligned} \boldsymbol{A}_q\left(\boldsymbol{\eta}_{q_i}, \boldsymbol{\eta}_{p_i}, \tilde{\boldsymbol{p}}(t_{i+1})^+\right) &= \boldsymbol{0}, \\ \boldsymbol{A}_p\left(\boldsymbol{\eta}_{q_i}, \boldsymbol{\eta}_{p_i}, \tilde{\boldsymbol{q}}(t_i)^-\right) &= \boldsymbol{0} \end{aligned} \tag{21}$$

is fulfilled. For a given boundary condition $\{\tilde{\boldsymbol{q}}(t_i), \tilde{\boldsymbol{p}}(t_i)\}$, which yields the two additional equations $\tilde{\boldsymbol{p}}(t_i) = \tilde{\boldsymbol{p}}(t_i, \boldsymbol{\eta}_{p_i}) = \tilde{\boldsymbol{p}}(t_i, \boldsymbol{\eta}_{p_i})^+$ by right continuity of $\tilde{\boldsymbol{p}}$ and $\tilde{\boldsymbol{q}}(t_i) = \tilde{\boldsymbol{q}}(t_i)^-$ by left continuity of $\tilde{\boldsymbol{q}}$, the set of equations (21) is solved for the parameters $\boldsymbol{\eta}_{q_i}, \boldsymbol{\eta}_{p_i}$, and the vectors $\tilde{\boldsymbol{q}}(t_i)^-$ and $\tilde{\boldsymbol{p}}(t_{i+1})^+$.[¶] Thereby, the boundary values $\tilde{\boldsymbol{q}}(t_i)^-$ and $\tilde{\boldsymbol{p}}(t_{i+1})^+$ are treated as unknown vectors in $\mathbb{R}^n$ and $(\mathbb{R}^n)^*$, respectively. The obtained solution can be used to construct the approximate trajectories of the system in the interval $[t_i, t_{i+1}]$. The values at $\{t_{i+1}\}$, that is, $\tilde{\boldsymbol{q}}(t_{i+1})$ and $\tilde{\boldsymbol{p}}(t_{i+1}) = \tilde{\boldsymbol{p}}(t_{i+1})^+$, yield the boundary conditions for the next time interval. By repeating this procedure for all time intervals I_i, $i = 0, 1, \ldots, N-1$, the virtual action vanishes for all variations $\delta\tilde{\boldsymbol{q}}$ and $\delta\tilde{\boldsymbol{p}}$, and trajectories approximating the dynamic equilibrium are obtained.

Before discussing the quality of this approximation in the next section, it is interesting to note that the boundary conditions are imposed a posteriori, together with the set of equations (21). This amounts to solve $nm_q + nm_p + 2n$ equations for the nm_q parameters $\boldsymbol{\eta}_{q_i}$, the nm_p parameters $\boldsymbol{\eta}_{p_i}$, the vector $\tilde{\boldsymbol{p}}(t_{i+1})^+ \in (\mathbb{R}^n)^*$, and the vector $\tilde{\boldsymbol{q}}(t_i)^- \in \mathbb{R}^n$. By doing so, we do not restrict the variations $\delta\tilde{\boldsymbol{q}}$ and $\delta\tilde{\boldsymbol{p}}$ to vanish at the time interval boundaries, that is, at the time instants t_i, $i = 0, 1, \ldots, N$. This turns out to be essential for guaranteeing symplectic integration as shown in Section 3.2. Moreover, Equation (16) can be interpreted as a stationarity condition with respect to the ansatz parameters $\boldsymbol{\eta}_{q_i}$ and $\boldsymbol{\eta}_{p_i}$ leading to a discrete analogue of Hamilton's principle in the absence of non-potential forces.

3. PROPERTIES

In the next section, some properties of the integration algorithm based on the stepping equations (21) are analyzed. Given a family of ansatz functions, for example, polynomials, that are dense in $C([t_i, t_{i+1}], \mathbb{R}^n)$, respectively $C\left([t_i, t_{i+1}], (\mathbb{R}^n)^*\right)$, the approximate solution converges to the exact trajectory in the limit of $m_q \to \infty$ and $m_p \to \infty$.[‖] This is because $\delta\tilde{\boldsymbol{q}}$ and $\delta\tilde{\boldsymbol{p}}$ can approximate any continuous variation arbitrarily well (uniform convergence) by the Stone–Weierstrass theorem. Solving therefore Equation (21) amounts in requiring the virtual action to vanish for all piecewise

[¶]From the boundary condition and the left continuity of $\tilde{\boldsymbol{q}}$, it follows that $\tilde{\boldsymbol{q}}(t_i)^- = \tilde{\boldsymbol{q}}(t_i)$, and therefore, solving for $\tilde{\boldsymbol{q}}(t_i)^-$ is trivial in the sense that the known boundary condition $\tilde{\boldsymbol{q}}(t_i)$ can be directly inserted into the expression for $\boldsymbol{A}_p$ in Equation (21).

[‖]The time intervals $[t_i, t_{i+1}]$, $i = 0, 1, \ldots, N-1$ are assumed to be small enough such that the trajectories of the mechanical system are unique. Local uniqueness is guaranteed by the theorem of Piccard–Lindelöf, for example, [25, p. 156].

DOI: 10.1002/nme

continuous variations $\delta \boldsymbol{q}$ and $\delta \boldsymbol{p}$ and guarantees, according to Postulate 2.1, dynamic equilibrium for all times $t \in I$ (almost everywhere).

In addition, expression (16) is invariant to the choice of ansatz parameters. Assume two parametrizations related by $\hat{\boldsymbol{\eta}}_{q_i} = \boldsymbol{f}(\boldsymbol{\eta}_{q_i})$, where $\boldsymbol{f} : \mathbb{R}^{n \cdot m_q} \to \mathbb{R}^{n \cdot m_q}$ is a diffeomorphism. Given that the expression of the virtual action vanishes for all variations $\delta \boldsymbol{\eta}_{q_i}$, it vanishes also for all variations $\delta \hat{\boldsymbol{\eta}}_{q_i}$, which is because of the relationship

$$\delta \hat{\boldsymbol{\eta}}_{q_i} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{\eta}_q} \delta \boldsymbol{\eta}_{q_i}, \qquad \text{with} \quad \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{\eta}_q} \quad \text{invertible.} \tag{22}$$

Therefore, the trajectories obtained by solving Equation (21) are only dependent on the choice of ansatz functions (e.g., polynomials) but independent of their actual parametrization (e.g., linear combination of monomials and linear combination of Legendre polynomials). In practice, however, the choice of the parametrization influences the conditioning of the (implicit) equations.

Moreover, the presented algorithms are shown to have some additional beneficial properties. To simplify the analysis, the non-potential forces $\boldsymbol{f}_{NP}$ are assumed to vanish in the following.

3.1. Conservation of energy

As introduced previously, the ansatz $\tilde{\boldsymbol{q}}$ is assumed to be left continuous and $\tilde{\boldsymbol{p}}$ right continuous. It is again assumed that the ansatz $\tilde{\boldsymbol{q}}(t)$ depends on time and only on the parameters $\boldsymbol{\eta}_{q_i} \in \mathbb{R}^{n \cdot m_q}$ within the interval $t \in (t_i, t_{i+1}]$. Similarly, $\tilde{\boldsymbol{p}}(t)$ is assumed to depend on time and only on the parameters $\boldsymbol{\eta}_{p_i} \in \mathbb{R}^{n \cdot m_p}$ within the interval $t \in [t_i, t_{i+1})$. Thus, as in the previous section, variations given by Equation (17) are introduced. The following additional assumption on the two ansatz functions is made.

Assumption 3.1
Let

$$\tilde{\boldsymbol{q}} \in \tilde{C}^1_{\text{pc}}\left(I \times \mathbb{R}^{n \cdot m_q \cdot N}, \mathbb{R}^n\right) \qquad \text{and} \qquad \tilde{\boldsymbol{p}} \in \tilde{C}^1_{\text{pc}}\left(I \times \mathbb{R}^{n \cdot m_p \cdot N}, (\mathbb{R}^n)^*\right),$$

be such that there exists for every $\delta t \in \mathbb{R}$ and every $\boldsymbol{\eta}_{q_i} \in \mathbb{R}^{n \cdot m_q}$, $\boldsymbol{\eta}_{p_i} \in \mathbb{R}^{n \cdot m_p}$ two vectors $\delta \boldsymbol{\eta}_{q_i} \in \mathbb{R}^{n \cdot m_q}$ and $\delta \boldsymbol{\eta}_{p_i} \in \mathbb{R}^{n \cdot m_p}$ that fulfill

$$\begin{aligned} \frac{\partial \tilde{\boldsymbol{q}}(\boldsymbol{\eta}_{q_i}, t)}{\partial \boldsymbol{\eta}_{q_i}} \delta \boldsymbol{\eta}_{q_i} &= \dot{\tilde{\boldsymbol{q}}}(\boldsymbol{\eta}_{q_i}, t)\, \delta t, \quad \forall t \in [t_i, t_{i+1}) \qquad \text{and} \\ \frac{\partial \tilde{\boldsymbol{p}}(\boldsymbol{\eta}_{p_i}, t)}{\partial \boldsymbol{\eta}_{p_i}} \delta \boldsymbol{\eta}_{p_i} &= \dot{\tilde{\boldsymbol{p}}}(\boldsymbol{\eta}_{p_i}, t)\, \delta t, \quad \forall t \in (t_i, t_{i+1}], \qquad i = 0, 1, 2, \ldots, N-1. \end{aligned}$$

Example 1
As an example, consider $q(t) \in \mathbb{R}$ for $t \in [0, 1]$. Choosing a polynomial ansatz function of second order leads to

$$\tilde{q}(t, \boldsymbol{\eta}_q) = \eta_0 + \eta_1 t + \eta_2 t^2$$

with $\boldsymbol{\eta}_q = (\eta_0, \eta_1, \eta_2)^{\mathrm{T}}$, $t \in [0, 1]$. Because $\dot{\tilde{q}} = \eta_1 + 2\eta_2 t$ lies in the subspace spanned by $\delta \tilde{q} = \delta\eta_0 + \delta\eta_1 t + \delta\eta_2 t^2$, Assumption 3.1 is fulfilled.

It is clear that the variations introduced in Assumption 3.1 are not variations with respect to time as described in [21] or [26], as for example, the time instants t_i, $i = 0, 1, \ldots, N$ remain fixed. Variations of the form $\delta \tilde{\boldsymbol{q}} = \dot{\tilde{\boldsymbol{q}}}\, \delta t$ and $\delta \tilde{\boldsymbol{p}} = \dot{\tilde{\boldsymbol{p}}}\, \delta t$ are merely a trick to extract the conservation of an energy-like quantity from the virtual action. In the case of polynomials, Assumption 3.1 requires that the ansatz functions should be able to capture an arbitrary time shift, such that the same trajectory

DOI: 10.1002/nme

is being obtained regardless of choosing $\tilde{\boldsymbol{q}}(t-\hat{t},\boldsymbol{\eta}_{q_i})$ or $\tilde{\boldsymbol{q}}(t,\boldsymbol{\eta}_{q_i})$ as ansatz (for a fixed $\hat{t}\in\mathbb{R}$). This condition is well known in the finite elements literature where ansatz functions are required to capture rigid body modes (for example [27, p. 31]). Ansatz functions fulfilling Assumption 3.1 are sometimes referred to as complete.

For this particular choice of variations, that is, $\delta\tilde{\boldsymbol{q}}=\dot{\tilde{\boldsymbol{q}}}\,\delta t$, $\delta\tilde{\boldsymbol{p}}=\dot{\tilde{\boldsymbol{p}}}\,\delta t$, the strong form of the virtual action (15) simplifies to

$$\begin{aligned}\delta A &= -\delta t\;\dot{\tilde{\boldsymbol{p}}}(t_i)^{\mathsf{T}}\left(\tilde{\boldsymbol{q}}(t_i)^{+}-\tilde{\boldsymbol{q}}(t_i)^{-}\right)+\delta t\;\dot{\tilde{\boldsymbol{q}}}(t_{i+1})^{\mathsf{T}}\left(\tilde{\boldsymbol{p}}(t_{i+1})^{+}-\tilde{\boldsymbol{p}}(t_{i+1})^{-}\right)\\ &\quad+\delta t\int\limits_{(t_i,t_{i+1})}\left(\frac{\partial H}{\partial\boldsymbol{p}}\dot{\tilde{\boldsymbol{p}}}+\frac{\partial H}{\partial\boldsymbol{q}}\dot{\tilde{\boldsymbol{q}}}\right)\mathrm{d}t\\ &= -\delta t\;\dot{\tilde{\boldsymbol{p}}}(t_i)^{\mathsf{T}}\left(\tilde{\boldsymbol{q}}(t_i)^{+}-\tilde{\boldsymbol{q}}(t_i)^{-}\right)+\delta t\;\dot{\tilde{\boldsymbol{q}}}(t_{i+1})^{\mathsf{T}}\left(\tilde{\boldsymbol{p}}(t_{i+1})^{+}-\tilde{\boldsymbol{p}}(t_{i+1})^{-}\right)\\ &\quad+\delta t\int\limits_{(t_i,t_{i+1})}\left(\frac{dH}{dt}-\frac{\partial H}{\partial t}\right)\mathrm{d}t\\ &= -\delta t\;\dot{\tilde{\boldsymbol{p}}}(t_i)^{\mathsf{T}}\left(\tilde{\boldsymbol{q}}(t_i)^{+}-\tilde{\boldsymbol{q}}(t_i)^{-}\right)+\delta t\;\dot{\tilde{\boldsymbol{q}}}(t_{i+1})^{\mathsf{T}}\left(\tilde{\boldsymbol{p}}(t_{i+1})^{+}-\tilde{\boldsymbol{p}}(t_{i+1})^{-}\right)\\ &\quad+\delta t\left(H\left(\tilde{\boldsymbol{p}}(t_{i+1})^{-},\tilde{\boldsymbol{q}}(t_{i+1})^{-},t_{i+1}\right)-H\left(\tilde{\boldsymbol{p}}(t_i)^{+},\tilde{\boldsymbol{q}}(t_i)^{+},t_i\right)-\int\limits_{(t_i,t_{i+1})}\frac{\partial H}{\partial t}\mathrm{d}t\right).\end{aligned}$$

Because the virtual work is required to vanish for all $\delta\tilde{\boldsymbol{q}}$ and all $\delta\tilde{\boldsymbol{p}}$, it must also vanish for this particular choice of variations (hence for all δt), which leads to

$$\begin{aligned}&\dot{\tilde{\boldsymbol{q}}}(t_{i+1})^{\mathsf{T}}\left(\tilde{\boldsymbol{p}}(t_{i+1})^{+}-\tilde{\boldsymbol{p}}(t_{i+1})^{-}\right)+H\left(\tilde{\boldsymbol{p}}(t_{i+1})^{-},\tilde{\boldsymbol{q}}(t_{i+1})^{-},t_{i+1}\right)=\\ &\qquad\dot{\tilde{\boldsymbol{p}}}(t_i)^{\mathsf{T}}\left(\tilde{\boldsymbol{q}}(t_i)^{+}-\tilde{\boldsymbol{q}}(t_i)^{-}\right)+H\left(\tilde{\boldsymbol{p}}(t_i)^{+},\tilde{\boldsymbol{q}}(t_i)^{+},t_i\right)+\int\limits_{(t_i,t_{i+1})}\frac{\partial H}{\partial t}\mathrm{d}t.\end{aligned}\tag{23}$$

If the system is conservative, that is, the Hamiltonian is not explicitly dependent on time, this simplifies further to

$$\begin{aligned}&\dot{\tilde{\boldsymbol{q}}}(t_{i+1})^{\mathsf{T}}\left(\tilde{\boldsymbol{p}}(t_{i+1})^{+}-\tilde{\boldsymbol{p}}(t_{i+1})^{-}\right)+H\left(\tilde{\boldsymbol{p}}(t_{i+1})^{-},\tilde{\boldsymbol{q}}(t_{i+1})^{-}\right)=\\ &\qquad\dot{\tilde{\boldsymbol{p}}}(t_i)^{\mathsf{T}}\left(\tilde{\boldsymbol{q}}(t_i)^{+}-\tilde{\boldsymbol{q}}(t_i)^{-}\right)+H\left(\tilde{\boldsymbol{p}}(t_i)^{+},\tilde{\boldsymbol{q}}(t_i)^{+}\right).\end{aligned}\tag{24}$$

The aforementioned equations can be interpreted as the conservation of an energy-like quantity in the following sense: All the terms appearing in Equation (24) can be expressed in terms of the values $\tilde{\boldsymbol{q}}(t_i)$, $\tilde{\boldsymbol{p}}(t_i)$ using the stepping equation (21). By moving the terms to the left-hand side the expression

$$\Delta H\left(\tilde{\boldsymbol{p}}(t_i),\tilde{\boldsymbol{q}}(t_i)\right)=0,\tag{25}$$

is obtained, which is valid for all time instants t_i, $i=0,1,\ldots,N-1$. This can be interpreted as an energy balance, augmented with the terms $\dot{\tilde{\boldsymbol{q}}}(t_{i+1})^{\mathsf{T}}\left(\tilde{\boldsymbol{p}}(t_{i+1})^{+}-\tilde{\boldsymbol{p}}(t_{i+1})^{-}\right)$ and $\dot{\tilde{\boldsymbol{p}}}(t_i)^{\mathsf{T}}\left(\tilde{\boldsymbol{q}}(t_i)^{+}-\tilde{\boldsymbol{q}}(t_i)^{-}\right)$ accounting for the discontinuities of $\tilde{\boldsymbol{q}}$ and $\tilde{\boldsymbol{p}}$.

3.2. Symplectic integration

To show symplectic integration, we choose an oriented smooth two-dimensional manifold with boundary, denoted by Γ, which is embedded in the phase space, $\Gamma\subset\mathbb{R}^n\times(\mathbb{R}^n)^*$ (e.g., [28, p. 411]). The boundary $\partial\Gamma$ defines a closed contour along which we choose initial conditions $\tilde{\boldsymbol{q}}(t_i,s)$ and $\tilde{\boldsymbol{p}}(t_i,s)$ (parametrized by $s\in[0,1)$). Solving Equation (21) for initial conditions along $\partial\Gamma$

DOI: 10.1002/nme

determines the parameters $\boldsymbol{\eta}_{q_i}(s)$ and $\boldsymbol{\eta}_{p_i}(s)$ as well as the final state $\tilde{\boldsymbol{q}}(t_{i+1}, \boldsymbol{\eta}_{q_i}(s))$, $\tilde{\boldsymbol{p}}(t_{i+1}, \boldsymbol{\eta}_{p_i}(s))$ for each $s \in [0, 1)$. The invariance of the two-form $\mathrm{d}\boldsymbol{q} \wedge \mathrm{d}\boldsymbol{p}$ follows directly by integrating Equation (21) along $\partial\Gamma$:

$$
\begin{aligned}
0 &= \int_{\partial\Gamma} \left(\boldsymbol{A}_q(s)^\top \frac{\partial \boldsymbol{\eta}_{q_i}}{\partial s} + \boldsymbol{A}_p(s)^\top \frac{\partial \boldsymbol{\eta}_{p_i}}{\partial s} \right) \mathrm{d}s \\
&= \int_{\partial\Gamma} \Bigg(- \mathrm{d}\left((\tilde{\boldsymbol{p}}(t_i)^+)^\top \tilde{\boldsymbol{q}}(t_i)^+ \right) + (\tilde{\boldsymbol{q}}(t_i)^-)^\top \mathrm{d}\tilde{\boldsymbol{p}}(t_i) + (\tilde{\boldsymbol{p}}(t_{i+1})^+)^\top \mathrm{d}\tilde{\boldsymbol{q}}_i(t_{i+1}) \\
&\qquad - \mathrm{d} \int\limits_{(t_i, t_{i+1})} \left(\tilde{\boldsymbol{p}}^\top \dot{\tilde{\boldsymbol{q}}} - H(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{q}}, t) \right) \mathrm{d}t \Bigg) \\
&= \int_{\partial\Gamma} \tilde{\boldsymbol{q}}(t_i)^\top \mathrm{d}\tilde{\boldsymbol{p}}(t_i) + \tilde{\boldsymbol{p}}(t_{i+1})^\top \mathrm{d}\tilde{\boldsymbol{q}}(t_{i+1}).
\end{aligned}
$$

Applying Stoke's theorem (e.g., [28, p. 411]) to the previous equation leads to

$$
\int_\Gamma \mathrm{d}\tilde{\boldsymbol{q}}(t_i) \wedge \mathrm{d}\tilde{\boldsymbol{p}}(t_i) = \int_\Gamma \mathrm{d}\tilde{\boldsymbol{q}}(t_{i+1}) \wedge \mathrm{d}\tilde{\boldsymbol{p}}(t_{i+1}). \tag{26}
$$

In other words, the integration preserves the two-form $\mathrm{d}\boldsymbol{q} \wedge \mathrm{d}\boldsymbol{p}$ and is therefore symplectic.

In practice, the integrals occurring in Equation (21) are typically approximated using quadrature. Nonetheless, the resulting integration scheme is guaranteed to be symplectic (under mild assumptions on the quadrature rule)**, as the previous argument can be carried over.

4. HOLONOMIC CONSTRAINTS

Next, the addition of holonomic constraints is discussed. The constraints are parametrized by the gap function $\boldsymbol{g} \in C^2(\mathbb{R}^n \times I, \mathbb{R}^{n_c})$ such that the system lies on the constraint manifold at time $t \in I$ if

$$
\boldsymbol{g}(\boldsymbol{q}(t), t) = \boldsymbol{0}.
$$

The integer n_c denotes the number of constraints, that is, the dimension of $\boldsymbol{g}(\boldsymbol{q}(t), t)$. Similar to Section 2, we postulate the principle of virtual action in the constrained case.

Postulate 4.1
Let $\boldsymbol{q} \in C^1_{\mathrm{pc}}(I, \mathbb{R}^n)$, $\boldsymbol{p} \in C^1_{\mathrm{pc}}(I, (\mathbb{R}^n)^*)$, $\boldsymbol{p}_g \in C^1_{\mathrm{pc}}(I, (\mathbb{R}^{n_c})^*)$, and $I = [t_0, t_N] \subset \mathbb{R}$. The Hamiltonian $H(\boldsymbol{p}, \boldsymbol{q}, t)$ corresponding to the unconstrained system is defined almost everywhere by Equation (3). It is assumed that the non-potential forces $\boldsymbol{f}_{NP}(t)$ are absolutely continuous for all $t \in I$ and that compliant initial conditions with $\boldsymbol{g}(\boldsymbol{q}(t_0), t_0) = 0$ are provided. If the virtual action expressed by

$$
\delta A := \int_I \delta\boldsymbol{q}^\top \mathrm{d}\boldsymbol{p} - \delta\boldsymbol{p}^\top \mathrm{d}\boldsymbol{q} - \delta\boldsymbol{g}^\top \mathrm{d}\boldsymbol{p}_g + \delta\boldsymbol{p}_g^\top \mathrm{d}\boldsymbol{g} + \left(\delta H(\boldsymbol{p}, \boldsymbol{q}, t) - \delta\boldsymbol{q}^\top \boldsymbol{f}_{NP}(t) \right) \mathrm{d}t, \tag{27}
$$

vanishes for all variations $\delta\boldsymbol{q}, \delta\boldsymbol{p}$, and $\delta\boldsymbol{p}_g$, then the system is almost everywhere in dynamic equilibrium for times $t \in I$.

**In particular, symplectic integration is guaranteed if the integrals in the expressions $\boldsymbol{A}_q$ and $\boldsymbol{A}_p$ are approximated using the same quadrature rule.

DOI: 10.1002/nme

Compared with Section 2, the expression of the virtual action is augmented by the two terms $-\delta \boldsymbol{g}^{\mathsf{T}} \mathrm{d}\boldsymbol{p}_g$ and $\delta \boldsymbol{p}_g^{\mathsf{T}} \mathrm{d}\boldsymbol{g}$ to account for the constraint. The addition of the second term, $\delta \boldsymbol{p}_g^{\mathsf{T}} \mathrm{d}\boldsymbol{g}$, enforces the constraint on the velocity level. More precisely, it implies $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{g}(\boldsymbol{q}(t),t) = 0$, whenever $\boldsymbol{q}$ is continuously differentiable and $\boldsymbol{g}(\boldsymbol{q}(t_d)^-, t_d) = \boldsymbol{g}(\boldsymbol{q}(t_d)^+, t_d)$ at time instants t_d, where $\boldsymbol{q}$ is discontinuous.[††] Hence, provided that a compliant initial condition $\boldsymbol{g}(\boldsymbol{q}(t_0), t_0) = 0$ is given, the term $\delta \boldsymbol{p}_g^{\mathsf{T}} \mathrm{d}\boldsymbol{g}$ ensures that the constraint will be fulfilled for all $t \in I$ (almost everywhere).

The addition of the first term, $-\delta \boldsymbol{g}^{\mathsf{T}} \mathrm{d}\boldsymbol{p}_g$, which can be expanded to

$$-\delta \boldsymbol{g}^{\mathsf{T}} \mathrm{d}\boldsymbol{p}_g = -\delta \boldsymbol{q}^{\mathsf{T}} \boldsymbol{W}(\boldsymbol{q}(t),t) \mathrm{d}\boldsymbol{p}_g, \qquad \text{with} \qquad \boldsymbol{W}(\boldsymbol{q}(t),t) := \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{q}}^{\mathsf{T}}, \tag{28}$$

describes the impulse $\mathrm{d}\boldsymbol{p}_g$ imposing the constraint. The matrix $\boldsymbol{W}(\boldsymbol{q}(t),t)$ assigns to each component of the impulse $\mathrm{d}\boldsymbol{p}_g$ a vector orthogonal to the tangent space of $\boldsymbol{g}(\boldsymbol{q}(t),t)$. Whenever $\boldsymbol{p}_g$ is continuously differentiable, $\mathrm{d}\boldsymbol{p}_g$ can be expressed as $\dot{\boldsymbol{p}}_g \mathrm{d}t$, where the components of $\dot{\boldsymbol{p}}_g$ represent the constraint forces and the columns of $\boldsymbol{W}(\boldsymbol{q}(t),t)$ the corresponding directions (e.g., [26, p. 141], [29, p. 51]). Together they form the generalized force $\boldsymbol{W}(\boldsymbol{q}(t),t)\dot{\boldsymbol{p}}_g$ imposing the constraint. Clearly for admissible variations,

$$\delta \boldsymbol{q}_{\mathrm{ad}}(t) \in \left\{ \boldsymbol{\xi} \in \mathbb{R}^n \mid \boldsymbol{W}(\boldsymbol{\xi},t)^{\mathsf{T}} \boldsymbol{\xi} = 0 \right\}, \tag{29}$$

$\delta \boldsymbol{q}_{\mathrm{ad}}(t)^{\mathsf{T}} \boldsymbol{W}(\boldsymbol{q}(t),t) \mathrm{d}\boldsymbol{p}_g$ vanishes. Provided that the constraint $\boldsymbol{g}(\boldsymbol{q}(t),t) = 0$ is fulfilled for all $t \in I$, the virtual action reduces therefore for admissible variations to the expression given by Equation (7) in Section 2, which is in accordance with the principle of d'Alembert–Lagrange (e.g., [29, p. 48]).

Using the expression of the virtual action given by Equation (27), the procedure of Section 2 is applied next to derive a variational integration algorithm for the constrained case.

4.1. Variational integration – the constrained case

Three ansatz functions $\tilde{\boldsymbol{q}}$, $\tilde{\boldsymbol{p}}$, and $\tilde{\boldsymbol{p}}_g$ for $\boldsymbol{q}$, $\boldsymbol{p}$, and $\boldsymbol{p}_g$, belonging to $\tilde{C}^1_{\mathrm{pc}}$, are introduced. We will make the additional assumption that $\tilde{\boldsymbol{q}}$ is left continuous, whereas $\tilde{\boldsymbol{p}}$ and $\tilde{\boldsymbol{p}}_g$ are both right continuous, which is analogous to Section 2. The time interval of interest, $I = [t_0, t_N]$, with $t_N > t_0$, is again divided into sub-intervals $I_i := [t_i, t_{i+1}] \subset \mathbb{R}$. The intervals $I_i \subset \mathbb{R}$ are chosen ($i = 0, 1, 2, \ldots, N-1$) such that $\bigcup_{i=0}^{N-1} (I_i \setminus \{t_{i+1}\}) = I \setminus \{t_N\}$. The ansatz functions

$$\begin{aligned} &\tilde{\boldsymbol{q}} \in \tilde{C}^1_{\mathrm{pc}}\left(I \times \mathbb{R}^{n \cdot m_q \cdot N}, \mathbb{R}^n\right), \qquad \tilde{\boldsymbol{p}} \in \tilde{C}^1_{\mathrm{pc}}\left(I \times \mathbb{R}^{n \cdot m_p \cdot N}, (\mathbb{R}^n)^*\right), \\ &\tilde{\boldsymbol{p}}_g \in \tilde{C}^1_{\mathrm{pc}}\left(I \times \mathbb{R}^{n_c \cdot m_g \cdot N}, (\mathbb{R}^{n_c})^*\right), \end{aligned} \tag{30}$$

are chosen to be continuous in the interior of the intervals I_i, $i = 0, 1, \ldots, N-1$, and are only allowed to be discontinuous at the time instants t_i, $i = 0, 1, 2, \ldots, N$.

The ansatz is inserted into the virtual action given by Equation (27). By requiring the virtual action to vanish for all variations $\delta\tilde{\boldsymbol{q}}$, $\delta\tilde{\boldsymbol{p}}$, and $\delta\tilde{\boldsymbol{p}}_g$, an approximation to the dynamic equilibrium is obtained. We choose variations $\delta\tilde{\boldsymbol{q}}$, $\delta\tilde{\boldsymbol{p}}$, and $\delta\tilde{\boldsymbol{p}}_g$, which are zero everywhere except for $t \in [t_i, t_{i+1}]$ to cut the interval I_i out of the virtual action (for a fixed integer $i > 0$). We then require the virtual action to vanish for all variations $\delta\tilde{\boldsymbol{q}}$, $\delta\tilde{\boldsymbol{p}}$, and $\delta\tilde{\boldsymbol{p}}_g$, vanishing everywhere except for $t \in I_i$. Repeating this procedure for all intervals I_i, $i = 0, 1, \ldots, N-1$, is equivalent to requiring the virtual action to vanish for arbitrary variations $\delta\tilde{\boldsymbol{q}}$, $\delta\tilde{\boldsymbol{p}}$, and $\delta\tilde{\boldsymbol{p}}_g$. This leads naturally to an iterative algorithm yielding approximate solutions to the dynamic equilibrium. We evaluate the virtual action for variations that vanish everywhere except on I_i. This results in

[††]The continuity of the gap function implies that $\lim_{t \uparrow t_d} \boldsymbol{g}(\boldsymbol{q}(t),t) = \boldsymbol{g}(\boldsymbol{q}(t_d)^-, t_d)$ and $\lim_{t \downarrow t_d} \boldsymbol{g}(\boldsymbol{q}(t),t) = \boldsymbol{g}(\boldsymbol{q}(t_d)^+, t_d)$ hold.

DOI: 10.1002/nme

$$\delta A = \int\limits_{\{t_i\}} \delta\tilde{\boldsymbol{q}}^\top \mathrm{d}\tilde{\boldsymbol{p}} - \delta\tilde{\boldsymbol{p}}^\top \mathrm{d}\tilde{\boldsymbol{q}} + \delta\tilde{\boldsymbol{p}}_g{}^\top \mathrm{d}\tilde{\boldsymbol{g}} - \delta\tilde{\boldsymbol{g}}^\top \mathrm{d}\tilde{\boldsymbol{p}}_g + \int\limits_{\{t_{i+1}\}} \delta\tilde{\boldsymbol{q}}^\top \mathrm{d}\tilde{\boldsymbol{p}} - \delta\tilde{\boldsymbol{p}}^\top \mathrm{d}\tilde{\boldsymbol{q}} + \delta\tilde{\boldsymbol{p}}_g{}^\top \mathrm{d}\tilde{\boldsymbol{g}} - \delta\tilde{\boldsymbol{g}}^\top \mathrm{d}\tilde{\boldsymbol{p}}_g$$
$$+ \int\limits_{(t_i,t_{i+1})} \left(\delta\tilde{\boldsymbol{q}}^\top \dot{\tilde{\boldsymbol{p}}} - \delta\tilde{\boldsymbol{p}}^\top \dot{\tilde{\boldsymbol{q}}} + \delta\tilde{\boldsymbol{p}}_g^\top \dot{\boldsymbol{g}} - \delta\tilde{\boldsymbol{g}}^\top \dot{\tilde{\boldsymbol{p}}}_g + \delta H\left(\tilde{\boldsymbol{p}},\tilde{\boldsymbol{q}},t\right) - \delta\tilde{\boldsymbol{q}}^\top \boldsymbol{f}_{NP}\right) \mathrm{d}t.$$

The differential measures $\mathrm{d}\tilde{\boldsymbol{q}}$, $\mathrm{d}\tilde{\boldsymbol{p}}$, and $\mathrm{d}\tilde{\boldsymbol{p}}_g$ are again expressed by $\dot{\tilde{\boldsymbol{q}}}\ \mathrm{d}t$, $\dot{\tilde{\boldsymbol{p}}}\ \mathrm{d}t$, and $\dot{\tilde{\boldsymbol{p}}}_g\ \mathrm{d}t$ in the interval (t_i, t_{i+1}). Note that $\tilde{\boldsymbol{g}}$ stands for $\tilde{\boldsymbol{g}} := \boldsymbol{g}(\tilde{\boldsymbol{q}}(t), t)$, so for example, $\tilde{\boldsymbol{g}}(t_i)^+$ is a shorthand notation for $\lim_{t\downarrow t_i} \boldsymbol{g}(\tilde{\boldsymbol{q}}(t), t)$. In the same way, the variation $\delta\tilde{\boldsymbol{g}}$ is an abbreviation of $\delta\boldsymbol{g}(\tilde{\boldsymbol{q}}(t), t) = \boldsymbol{W}(\tilde{\boldsymbol{q}}(t), t)^\top \delta\tilde{\boldsymbol{q}}(t)$.

The left and right continuities of $\tilde{\boldsymbol{q}}$, $\tilde{\boldsymbol{p}}$, and $\tilde{\boldsymbol{p}}_g$, implying that $\delta\tilde{\boldsymbol{q}}(t_i) = 0$, $\delta\tilde{\boldsymbol{p}}(t_{i+1}) = 0$, $\delta\tilde{\boldsymbol{p}}_g(t_{i+1}) = 0$ due to the assumption that the variations vanish outside the interval $[t_i, t_{i+1}]$, simplify the previous expression to

$$\begin{aligned}\delta A = &-\delta\tilde{\boldsymbol{p}}(t_i)^\top \left(\tilde{\boldsymbol{q}}(t_i)^+ - \tilde{\boldsymbol{q}}(t_i)^-\right) + \delta\tilde{\boldsymbol{q}}(t_{i+1})^\top \left(\tilde{\boldsymbol{p}}(t_{i+1})^+ - \tilde{\boldsymbol{p}}(t_{i+1})^-\right)\\ &+ \delta\tilde{\boldsymbol{p}}_g(t_i)^\top \left(\tilde{\boldsymbol{g}}(t_i)^+ - \tilde{\boldsymbol{g}}(t_i)^-\right) - \delta\tilde{\boldsymbol{g}}(t_{i+1})^\top \left(\tilde{\boldsymbol{p}}_g(t_{i+1})^+ - \tilde{\boldsymbol{p}}_g(t_{i+1})^-\right)\\ &+ \int\limits_{(t_i,t_{i+1})} \left(\delta\tilde{\boldsymbol{q}}^\top \dot{\tilde{\boldsymbol{p}}} - \delta\tilde{\boldsymbol{p}}^\top \dot{\tilde{\boldsymbol{q}}} + \delta\tilde{\boldsymbol{p}}_g{}^\top \tilde{\boldsymbol{g}} - \delta\tilde{\boldsymbol{g}}^\top \dot{\tilde{\boldsymbol{p}}}_{\boldsymbol{g}} + \delta H\left(\tilde{\boldsymbol{p}},\tilde{\boldsymbol{q}},t\right) - \delta\tilde{\boldsymbol{q}}^\top \boldsymbol{f}_{NP}\right) \mathrm{d}t,\end{aligned} \tag{31}$$

which is nothing but the strong form of the virtual action.

Applying integration by parts to the terms $\delta\tilde{\boldsymbol{q}}^\top \dot{\tilde{\boldsymbol{p}}}$ and $\delta\tilde{\boldsymbol{g}}^\top \dot{\tilde{\boldsymbol{p}}}_{\boldsymbol{g}}$ results in[‡‡]

$$\begin{aligned}\delta A = &-\delta\tilde{\boldsymbol{p}}(t_i)^\top \left(\tilde{\boldsymbol{q}}(t_i)^+ - \tilde{\boldsymbol{q}}(t_i)^-\right) - \left(\delta\tilde{\boldsymbol{q}}(t_i)^+\right)^\top \tilde{\boldsymbol{p}}(t_i)^+ + \delta\tilde{\boldsymbol{q}}(t_{i+1})^\top \tilde{\boldsymbol{p}}(t_{i+1})^+\\ &+ \delta\tilde{\boldsymbol{p}}_g(t_i)^\top \left(\tilde{\boldsymbol{g}}(t_i)^+ - \tilde{\boldsymbol{g}}(t_i)^-\right) + \left(\delta\tilde{\boldsymbol{g}}(t_i)^+\right)^\top \tilde{\boldsymbol{p}}_g(t_i)^+ - \delta\tilde{\boldsymbol{g}}(t_{i+1})^\top \tilde{\boldsymbol{p}}_g(t_{i+1})^+\\ &- \int\limits_{(t_i,t_{i+1})} \left[\delta\left(\tilde{\boldsymbol{p}}^\top \dot{\tilde{\boldsymbol{q}}} - \tilde{\boldsymbol{p}}_g^\top \dot{\tilde{\boldsymbol{g}}} - H\left(\tilde{\boldsymbol{p}},\tilde{\boldsymbol{q}},t\right)\right) + \delta\tilde{\boldsymbol{q}}^\top \boldsymbol{f}_{NP}\right] \mathrm{d}t,\end{aligned} \tag{32}$$

that is, the weak form of the virtual action.

As in Section 2, it is assumed that the ansatz $\tilde{\boldsymbol{q}}$ depends on time and only on the parameters $\boldsymbol{\eta}_{q_i} \in \mathbb{R}^{n\cdot m_q}$ within the interval $t \in (t_i, t_{i+1}]$. Likewise, $\tilde{\boldsymbol{p}}$ and $\tilde{\boldsymbol{p}}_g$ are assumed to depend on time and only on the parameters $\boldsymbol{\eta}_{p_i} \in \mathbb{R}^{n\cdot m_p}$ and $\boldsymbol{\eta}_{p_{gi}} \in \mathbb{R}^{n_c\cdot m_g}$ within the interval $t \in [t_i, t_{i+1})$. This simplifies the variations $\delta\tilde{\boldsymbol{q}}$, $\delta\tilde{\boldsymbol{p}}$, and $\delta\tilde{\boldsymbol{p}}_g$, which were chosen to vanish for $t \notin [t_i, t_{i+1}]$ to

$$\begin{aligned}\delta\tilde{\boldsymbol{q}} &= \begin{cases} \frac{\partial\tilde{\boldsymbol{q}}}{\partial\boldsymbol{\eta}_{q_i}}\delta\boldsymbol{\eta}_{q_i}, & t \in (t_i, t_{i+1}]\\ 0, & \text{else}\end{cases}, \qquad \delta\tilde{\boldsymbol{p}} = \begin{cases} \frac{\partial\tilde{\boldsymbol{p}}}{\partial\boldsymbol{\eta}_{p_i}}\delta\boldsymbol{\eta}_{p_i}, & t \in [t_i, t_{i+1})\\ 0, & \text{else}\end{cases} \qquad \text{and}\\ \delta\tilde{\boldsymbol{p}}_g &= \begin{cases} \frac{\partial\tilde{\boldsymbol{p}}}{\partial\boldsymbol{\eta}_{p_{gi}}}\delta\boldsymbol{\eta}_{p_{gi}}, & t \in [t_i, t_{i+1})\\ 0, & \text{else}\end{cases}.\end{aligned} \tag{33}$$

Factorizing the time-independent variations $\boldsymbol{\eta}_{q_i}$, $\boldsymbol{\eta}_{p_i}$, and $\boldsymbol{\eta}_{p_{g_i}}$ out leads to

$$\delta A = \delta\boldsymbol{\eta}_{q_i}^\top \boldsymbol{A}_q + \delta\boldsymbol{\eta}_{p_i}^\top \boldsymbol{A}_p + \delta\boldsymbol{\eta}_{p_{g_i}}^\top \boldsymbol{A}_{p_g}, \tag{34}$$

with

[‡‡]For a holonomic constraint, it holds that $\frac{\mathrm{d}}{\mathrm{d}t}\left(\delta\boldsymbol{g}(\tilde{\boldsymbol{q}}(t), t)\right) = \delta\left(\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{g}(\tilde{\boldsymbol{q}}(t), t)\right)$.

DOI: 10.1002/nme

$$
\begin{aligned}
\boldsymbol{A}_q := &- \left.\frac{\partial \tilde{\boldsymbol{q}}}{\partial \boldsymbol{\eta}_{q_i}}\right|_{t\downarrow t_i}^{\mathsf{T}} \left(\tilde{\boldsymbol{p}}(t_i)^+ - \boldsymbol{W}\left(\tilde{\boldsymbol{q}}(t_i)^+, t_i\right)\, \tilde{\boldsymbol{p}}_g(t_i)^+\right) \\
&+ \left.\frac{\partial \tilde{\boldsymbol{q}}}{\partial \boldsymbol{\eta}_{q_i}}\right|_{t=t_{i+1}}^{\mathsf{T}} \left(\tilde{\boldsymbol{p}}(t_{i+1})^+ - \boldsymbol{W}(\tilde{\boldsymbol{q}}(t_{i+1}), t_{i+1})\, \tilde{\boldsymbol{p}}_g(t_{i+1})^+\right) \\
&- \int\limits_{(t_i, t_{i+1})} \left(\frac{\partial \dot{\tilde{\boldsymbol{q}}}}{\partial \boldsymbol{\eta}_{q_i}}^{\mathsf{T}} \tilde{\boldsymbol{p}} - \frac{\partial \dot{\tilde{\boldsymbol{g}}}}{\partial \boldsymbol{\eta}_{q_i}}^{\mathsf{T}} \tilde{\boldsymbol{p}}_g - \frac{\partial \tilde{\boldsymbol{q}}}{\partial \boldsymbol{\eta}_{q_i}}^{\mathsf{T}} \left(\frac{\partial H}{\partial \boldsymbol{q}}^{\mathsf{T}} - \boldsymbol{f}_{NP}\right)\right) \mathrm{d}t,
\end{aligned}
$$

$$
\boldsymbol{A}_p := - \left.\frac{\partial \tilde{\boldsymbol{p}}}{\partial \boldsymbol{\eta}_{p_i}}\right|_{t=t_i}^{\mathsf{T}} \left(\tilde{\boldsymbol{q}}(t_i)^+ - \tilde{\boldsymbol{q}}(t_i)^-\right) - \int\limits_{(t_i, t_{i+1})} \left(\frac{\partial \tilde{\boldsymbol{p}}}{\partial \boldsymbol{\eta}_{p_i}}^{\mathsf{T}} \left(\dot{\tilde{\boldsymbol{q}}} - \frac{\partial H}{\partial \boldsymbol{p}}^{\mathsf{T}}\right)\right) \mathrm{d}t,
$$

and

$$
\boldsymbol{A}_{p_g} := \left.\frac{\partial \tilde{\boldsymbol{p}}_g}{\partial \boldsymbol{\eta}_{p_{gi}}}\right|_{t=t_i}^{\mathsf{T}} \left(\boldsymbol{g}\left(\tilde{\boldsymbol{q}}(t_i)^+, t_i\right) - \boldsymbol{g}(\tilde{\boldsymbol{q}}(t_i)^-, t_i)\right) + \int\limits_{(t_i, t_{i+1})} \frac{\partial \tilde{\boldsymbol{p}}_g}{\partial \boldsymbol{\eta}_{p_{gi}}}^{\mathsf{T}} \dot{\boldsymbol{g}}(\tilde{\boldsymbol{q}}(t), t)\, \mathrm{d}t.
$$

According to Equation (33), the generalized coordinates $\tilde{\boldsymbol{q}}(t_i)$ at time t_i are only dependent on the parameters $\boldsymbol{\eta}_{q_{i-1}}$, and owing to the left continuity of $\tilde{\boldsymbol{q}}$, it holds that $\tilde{\boldsymbol{q}}(t_i)^- = \tilde{\boldsymbol{q}}(t_i) \neq \tilde{\boldsymbol{q}}(t_i)^+$. As in the unconstrained case, $\tilde{\boldsymbol{q}}(t_i)^-$ is considered to be a fixed vector in $\mathbb{R}^n$ when approximating the trajectories in the interval $[t_i, t_{i+1}]$.

The ansatz for $\tilde{\boldsymbol{p}}$ and $\tilde{\boldsymbol{p}}_g$ is by definition right continuous, which implies that $\tilde{\boldsymbol{p}}(t_{i+1})^- \neq \tilde{\boldsymbol{p}}(t_{i+1}) = \tilde{\boldsymbol{p}}(t_{i+1})^+$ and $\tilde{\boldsymbol{p}}_g(t_{i+1})^- \neq \tilde{\boldsymbol{p}}_g(t_{i+1}) = \tilde{\boldsymbol{p}}_g(t_{i+1})^+$. According to Equation (33), the variables $\tilde{\boldsymbol{p}}(t_{i+1})$, $\tilde{\boldsymbol{p}}(t_{i+1})^+$ and $\tilde{\boldsymbol{p}}_g(t_{i+1})$, $\tilde{\boldsymbol{p}}_g(t_{i+1})^+$ are only dependent on the parameters $\boldsymbol{\eta}_{p_{i+1}}$ and $\boldsymbol{\eta}_{p_{g(i+1)}}$, respectively. By examination of $\boldsymbol{A}_q$ in Equation (34), it is apparent that the term $\tilde{\boldsymbol{p}}(t_{i+1})^+$ and $\tilde{\boldsymbol{p}}_g(t_{i+1})^+$, describing the coupling to the neighboring time interval, appears only in the linear combination

$$
\tilde{\boldsymbol{p}}(t_{i+1})^+ - \boldsymbol{W}\left(\tilde{\boldsymbol{q}}(t_{i+1}), t_{i+1}\right)\, \tilde{\boldsymbol{p}}_g(t_{i+1})^+ = \left.\left[\tilde{\boldsymbol{p}}(t) - \boldsymbol{W}(\tilde{\boldsymbol{q}}(t), t)\, \tilde{\boldsymbol{p}}_g(t)\right]\right|_{t=t_{i+1}}, \tag{35}
$$

which is therefore regarded as a fixed vector in $(\mathbb{R}^n)^*$ when solving for the approximate trajectories in the interval $[t_i, t_{i+1}]$.

As the expression on the right-hand side of Equation (35) is extensively used in the following, the function

$$
\boldsymbol{p}_z\left(\boldsymbol{q}(t), \boldsymbol{p}(t), \boldsymbol{p}_g(t), t\right) := \boldsymbol{p}(t) - \boldsymbol{W}(\boldsymbol{q}(t), t)\, \boldsymbol{p}_g(t) \tag{36}
$$

is introduced. To simplify notation, its arguments $\boldsymbol{q}(t)$, $\boldsymbol{p}(t)$, $\boldsymbol{p}_g(t)$ will be omitted and $\tilde{\boldsymbol{p}}_z(t)$ will be used to denote $\tilde{\boldsymbol{p}}_z(t) = \boldsymbol{p}_z(\tilde{\boldsymbol{q}}(t), \tilde{\boldsymbol{p}}(t), \tilde{\boldsymbol{p}}_g(t), t)$. It follows from the discontinuity of $\tilde{\boldsymbol{q}}$, $\tilde{\boldsymbol{p}}$, and $\tilde{\boldsymbol{p}}_g$ that $\tilde{\boldsymbol{p}}_z$ is neither left nor right continuous.

As pointed out in Section 2, making the virtual action vanish for all variations $\delta\tilde{\boldsymbol{q}}$, $\delta\tilde{\boldsymbol{p}}$, and $\delta\tilde{\boldsymbol{p}}_g$ is equivalent to requiring the virtual action to vanish for all variations $\delta\boldsymbol{\eta}_{q_i}$, $\delta\boldsymbol{\eta}_{p_i}$, and $\delta\boldsymbol{\eta}_{p_{gi}}$, with $i = 0, 1, \ldots, N-1$. The latter amounts in solving the set of equations

$$
\begin{aligned}
&\boldsymbol{A}_q\left(\boldsymbol{\eta}_{q_i}, \boldsymbol{\eta}_{p_i}, \boldsymbol{\eta}_{p_{g_i}}, \tilde{\boldsymbol{p}}_z(t_{i+1})\right) = \boldsymbol{0}, \\
&\boldsymbol{A}_p\left(\boldsymbol{\eta}_{q_i}, \boldsymbol{\eta}_{p_i}, \tilde{\boldsymbol{q}}(t_i)^-\right) = \boldsymbol{0}, \\
&\boldsymbol{A}_{p_g}\left(\boldsymbol{\eta}_{q_i}, \boldsymbol{\eta}_{p_{g_i}}, \tilde{\boldsymbol{q}}(t_i)^-\right) = \boldsymbol{0},
\end{aligned} \tag{37}
$$

DOI: 10.1002/nme

for the parameters $\boldsymbol{\eta}_{q_i}$, $\boldsymbol{\eta}_{p_i}$, $\boldsymbol{\eta}_{p_{g_i}}$, and the vectors $\tilde{\boldsymbol{q}}(t_i)^-$ and $\tilde{\boldsymbol{p}}_z(t_{i+1})$. Note that the quantities $\tilde{\boldsymbol{p}}(t_i)^+$, $\tilde{\boldsymbol{p}}_g(t_i)^+$, $\tilde{\boldsymbol{q}}(t_i)^+$ are all dependent on the parameters $\boldsymbol{\eta}_{q_i}$, $\boldsymbol{\eta}_{p_i}$, and $\boldsymbol{\eta}_{p_{g_i}}$ and are therefore not listed as arguments of $\boldsymbol{A}_q$, $\boldsymbol{A}_p$, $\boldsymbol{A}_{p_g}$ in Equation (37). Given initial conditions $\tilde{\boldsymbol{q}}(t_i)$ and $\tilde{\boldsymbol{p}}_z(t_i)$, this allows to calculate the approximate trajectories $\tilde{\boldsymbol{q}}$, $\tilde{\boldsymbol{p}}$, and $\tilde{\boldsymbol{p}}_g$ in the interval (t_i, t_{i+1}). Furthermore, the values $\tilde{\boldsymbol{q}}(t_{i+1})$, given by evaluating $\tilde{\boldsymbol{q}}(t, \boldsymbol{\eta}_{q_i})$ at $t = t_{i+1}$, and $\tilde{\boldsymbol{p}}_z(t_{i+1})$ are obtained and are used as boundary conditions for the next time interval. More precisely, given the boundary conditions $\left\{\tilde{\boldsymbol{q}}(t_i), \tilde{\boldsymbol{p}}_z(t_i)\right\}$, we solve Equation (37) together with

$$\tilde{\boldsymbol{q}}(t_i) = \tilde{\boldsymbol{q}}(t_i)^-, \qquad \tilde{\boldsymbol{p}}_z(t_i) = \tilde{\boldsymbol{p}}\left(t_i, \boldsymbol{\eta}_{p_i}\right) - \boldsymbol{W}\left(\tilde{\boldsymbol{q}}(t_i), t_i\right) \tilde{\boldsymbol{p}}_g(t_i, \boldsymbol{\eta}_{p_{g_i}}) \tag{38}$$

for $\boldsymbol{\eta}_{q_i}, \boldsymbol{\eta}_{p_i}, \boldsymbol{\eta}_{p_{g_i}}$, $\tilde{\boldsymbol{q}}(t_i)^-$, and $\tilde{\boldsymbol{p}}_z(t_{i+1})$. This amounts in solving $nm_q + nm_p + n_c m_g + 2n$ equations for the nm_q parameters $\boldsymbol{\eta}_{q_i}$, the nm_p parameters $\boldsymbol{\eta}_{p_i}$, the $n_c m_g$ parameters $\boldsymbol{\eta}_{p_{g_i}}$, the vector $\tilde{\boldsymbol{p}}_z(t_{i+1}) \in (\mathbb{R}^n)^*$, and the vector $\tilde{\boldsymbol{q}}(t_i)^- \in \mathbb{R}^n$.[§§] Note that the values $\tilde{\boldsymbol{q}}(t_i)^-$ and $\tilde{\boldsymbol{p}}_z(t_{i+1})$ are regarded as unknown vectors in $\mathbb{R}^n$ and $(\mathbb{R}^n)^*$ when solving the stepping equations (37). To initialize the algorithm, compliant initial conditions $\boldsymbol{q}_0$ and $\boldsymbol{p}_0$ must be provided, that is, fulfilling $\boldsymbol{g}(\boldsymbol{q}_0, t_0) = 0$, and the boundary condition on the generalized momentum is set by $\tilde{\boldsymbol{p}}_z(t_0) = \boldsymbol{p}_0$.

4.2. Properties

The properties of the integration are analyzed next. To simplify the analysis, non-potential forces $\boldsymbol{f}_{NP}$ are again assumed to be absent.

A similar argument as in Section 3 can be made to ensure convergence of the approximate trajectories for $m_q \to \infty, m_p \to \infty$, and $m_g \to \infty$.

4.2.1. Conservation of the gap function $\boldsymbol{g}$. It will be shown that the gap function $\boldsymbol{g}(\boldsymbol{q}(t), t)$ is conserved if the ansatz fulfills the following condition.

Assumption 4.1
Let $\tilde{\boldsymbol{p}}_g \in \tilde{C}^1_{\mathrm{pc}}(I \times \mathbb{R}^{n_c \cdot m_g \cdot N}, (\mathbb{R}^{n_c})^*)$ be such that for each $k \in \{1, 2, \ldots, n_c\}$ and each $i \in \{0, 1, \ldots, N-1\}$, a $j \in \{1, 2, \ldots, n_c \cdot m_g \cdot N\}$ exists, which fulfills

$$\tilde{p}_{g_k}(t, \boldsymbol{\eta}_{p_{g_i}}) = \eta_{p_{g_j}}, \quad \forall t \in [t_i, t_{i+1}), \tag{39}$$

where $\tilde{p}_{g_k}$ denotes the kth component of $\tilde{\boldsymbol{p}}_g$ and $\eta_{p_{g_j}}$ the jth component of $\boldsymbol{\eta}_{p_g}$.

This assumption implies that the ansatz has at least a constant element. It is assumed to hold throughout the next section.

Evaluating the kth component of $\boldsymbol{A}_{p_g}$ in Equation (37) with respect to the variation $\delta\eta_{p_{g_j}}$, such that $\delta\tilde{p}_{g_k}(t, \boldsymbol{\eta}_{p_{g_i}}) = \delta\eta_{p_{g_j}}$, leads to

$$\begin{aligned} A_{p_{g_k}} &= g_k\left(\tilde{\boldsymbol{q}}(t_i)^+, t_i\right) - g_k\left(\tilde{\boldsymbol{q}}(t_i)^-, t_i\right) + \int\limits_{(t_i, t_{i+1})} \dot{g}_k(\tilde{\boldsymbol{q}}(t), t)\, \mathrm{d}t \\ &= g_k\left(\tilde{\boldsymbol{q}}(t_{i+1})^-, t_{i+1}\right) - g_k\left(\tilde{\boldsymbol{q}}(t_i)^-, t_i\right) = 0. \end{aligned}$$

By Assumption 4.1, this holds likewise for all $k \in \{1, 2, \ldots, n_c\}$ and concludes that

$$\boldsymbol{g}\left(\tilde{\boldsymbol{q}}(t_{i+1})^-, t_{i+1}\right) = \boldsymbol{g}\left(\tilde{\boldsymbol{q}}(t_i)^-, t_i\right).$$

In other words, if the algorithm is initialized with an initial configuration $\boldsymbol{q}_0$ compliant with the constraint, the approximate trajectory will always stay in accordance with the constraint at the

[§§] Like in the unconstrained case, solving for $\tilde{\boldsymbol{q}}(t_i)^-$ is trivial in the sense that the known boundary condition $\tilde{\boldsymbol{q}}(t_i)$ can be directly inserted into the expressions $\boldsymbol{A}_p$ and $\boldsymbol{A}_{p_g}$ in Equation (37).

DOI: 10.1002/nme

discontinuous time instants t_i, $i = 0, 1, 2, \ldots, N$. This ensures that no additional regularization is needed.

4.2.2. Conservation of energy. Given that Assumption 3.1 holds for $\tilde{\boldsymbol{q}}$, $\tilde{\boldsymbol{p}}$, and $\tilde{\boldsymbol{p}}_g$, an energy-like quantity is shown to be conserved, similar to Section 3.1. Again, applying the specific variations $\delta\tilde{\boldsymbol{q}} = \dot{\tilde{\boldsymbol{q}}}\,\delta t$, $\delta\tilde{\boldsymbol{p}} = \dot{\tilde{\boldsymbol{p}}}\,\delta t$, and $\delta\tilde{\boldsymbol{p}}_g = \dot{\tilde{\boldsymbol{p}}}_g\,\delta t$ to the strong form of the virtual work leads to

$$\begin{aligned}\dot{\tilde{\boldsymbol{q}}}(t_{i+1})^{\mathsf{T}}\left(\tilde{\boldsymbol{p}}_z(t_{i+1}) - \tilde{\boldsymbol{p}}_z(t_{i+1})^-\right) + H\left(\tilde{\boldsymbol{p}}(t_{i+1})^-, \tilde{\boldsymbol{q}}(t_{i+1})^-, t_{i+1}\right) &= \dot{\tilde{\boldsymbol{p}}}(t_i)^{\mathsf{T}}\left(\tilde{\boldsymbol{q}}(t_i)^+ - \tilde{\boldsymbol{q}}(t_i)^-\right)\\ - \dot{\tilde{\boldsymbol{p}}}_g(t_i)^{\mathsf{T}}\left(\tilde{\boldsymbol{g}}(t_i)^+ - \tilde{\boldsymbol{g}}(t_i)^-\right) + H\left(\tilde{\boldsymbol{p}}(t_i)^+, \tilde{\boldsymbol{q}}(t_i)^+, t_i\right) &+ \int\limits_{(t_i, t_{i+1})} \left(\frac{\partial H}{\partial t} - \frac{\partial \tilde{\boldsymbol{g}}^{\mathsf{T}}}{\partial t}\dot{\tilde{\boldsymbol{p}}}_{\boldsymbol{g}}\right)\mathrm{d}t.\end{aligned} \tag{40}$$

Similar to the discussion in Section 3.1, the aforementioned equation can be expressed using the boundary values $\tilde{\boldsymbol{q}}(t_i)$, $\tilde{\boldsymbol{p}}_z(t_i)$, yielding the discrete energy balance

$$\Delta H_c\left(\tilde{\boldsymbol{q}}(t_0), \tilde{\boldsymbol{p}}_z(t_0)\right) = \Delta H_c\left(\tilde{\boldsymbol{q}}(t_1), \tilde{\boldsymbol{p}}_z(t_1)\right) = \cdots = \Delta H_c\left(\tilde{\boldsymbol{q}}(t_{N-1}), \tilde{\boldsymbol{p}}_z(t_{N-1})\right) = 0, \tag{41}$$

in case the Hamiltonian and the gap function are not explicitly dependent on time.

4.2.3. Symplectic integration. Next, it will be shown that the transformation from $\{\tilde{\boldsymbol{q}}(t_i), \tilde{\boldsymbol{p}}(t_i)\}$ to $\{\tilde{\boldsymbol{q}}(t_{i+1}), \tilde{\boldsymbol{p}}(t_{i+1})\}$ is symplectic. We choose an oriented smooth two-dimensional manifold with boundary [28, p. 411], denoted by Γ, which is embedded in the phase space, $\Gamma \subset \mathbb{R}^n \times (\mathbb{R}^n)^*$, and is compliant with the constraint, that is, $\boldsymbol{g}(\boldsymbol{q}, t_i) = 0$, $\forall(\boldsymbol{q}, \boldsymbol{p}) \in \Gamma$. The boundary $\partial\Gamma$ defines a closed contour, along which initial conditions $\tilde{\boldsymbol{q}}(t_i, s)$, $\tilde{\boldsymbol{p}}(t_i, s)$, parametrized by $s \in [0, 1)$, are chosen. Thus, the generalized coordinates $\tilde{\boldsymbol{q}}(t_i, s)$ are in agreement with the constraint, that is,

$$\boldsymbol{g}(\tilde{\boldsymbol{q}}(t_i, s), t_i) = 0, \quad \forall s \in [0, 1). \tag{42}$$

Solving Equation (37) yields the parameters $\boldsymbol{\eta}_{q_i}(s)$, $\boldsymbol{\eta}_{p_i}(s)$, and $\boldsymbol{\eta}_{p_{g_i}}(s)$ for every $s \in [0, 1)$, which parametrize the approximate trajectories $\tilde{\boldsymbol{p}}(t, \boldsymbol{\eta}_{p_i}(s))$, $\tilde{\boldsymbol{q}}(t, \boldsymbol{\eta}_{q_i}(s))$, and $\tilde{\boldsymbol{p}}_g(t, \boldsymbol{\eta}_{p_{g_i}}(s))$ for $t \in (t_i, t_{i+1})$.[¶¶]

Integration of Equation (34) along $\partial\Gamma$ leads to

$$\begin{aligned}0 &= \int_{\partial\Gamma} \boldsymbol{A}_q(s)^{\mathsf{T}}\mathrm{d}\boldsymbol{\eta}_{q_i} + \boldsymbol{A}_p^{\mathsf{T}}\mathrm{d}\boldsymbol{\eta}_{p_i} + \boldsymbol{A}_{pq}^{\mathsf{T}}\mathrm{d}\boldsymbol{\eta}_{p_{g_i}}\\ &= \int_{\partial\Gamma}\Bigg(-\mathrm{d}\int\limits_{(t_i, t_{i+1})}\left(\tilde{\boldsymbol{p}}^{\mathsf{T}}\dot{\tilde{\boldsymbol{q}}} - \tilde{\boldsymbol{p}}_g{}^{\mathsf{T}}\dot{\tilde{\boldsymbol{g}}} - H\left(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{q}}, t\right)\right)\mathrm{d}t - \mathrm{d}\left(\left(\tilde{\boldsymbol{q}}(t_i)^+\right)^{\mathsf{T}}\tilde{\boldsymbol{p}}(t_i)^+\right)\\ &\quad + \mathrm{d}\left(\left(\tilde{\boldsymbol{p}}_g(t_i)^+\right)^{\mathsf{T}}\tilde{\boldsymbol{g}}(t_i)^+\right) + \tilde{\boldsymbol{p}}_z(t_{i+1})^{\mathsf{T}}\mathrm{d}\tilde{\boldsymbol{q}}(t_{i+1}) + \tilde{\boldsymbol{q}}(t_i)^{\mathsf{T}}\mathrm{d}\tilde{\boldsymbol{p}}(t_i) - \tilde{\boldsymbol{g}}(t_i)^{\mathsf{T}}\mathrm{d}\tilde{\boldsymbol{p}}_g(t_i)\Bigg)\\ &= \int_{\partial\Gamma}\tilde{\boldsymbol{p}}_z(t_{i+1})^{\mathsf{T}}\mathrm{d}\tilde{\boldsymbol{q}}(t_{i+1}) + \tilde{\boldsymbol{q}}(t_i)^{\mathsf{T}}\mathrm{d}\tilde{\boldsymbol{p}}(t_i) - \tilde{\boldsymbol{g}}(t_i)^{\mathsf{T}}\mathrm{d}\tilde{\boldsymbol{p}}_g(t_i).\end{aligned} \tag{43}$$

Because $\tilde{\boldsymbol{q}}(t_i, s)$ is chosen such that $\boldsymbol{g}\left(\tilde{\boldsymbol{q}}(t_i, s), t_i\right) = 0$, $\forall s \in [0, 1)$, it holds according to Section 4.2.1 that $\boldsymbol{g}\left(\tilde{\boldsymbol{q}}(t_{i+1}, s), t_{i+1}\right) = 0$, $\forall s \in [0, 1)$. Hence, it follows that

[¶¶]The dependence on s is occasionally omitted in the following to simplify notation.

DOI: 10.1002/nme

$$\begin{aligned}\tilde{\boldsymbol{p}}_z(t_{i+1})^{\mathsf{T}}\mathrm{d}\tilde{\boldsymbol{q}}(t_{i+1}) &= \left(\tilde{\boldsymbol{p}}(t_{i+1}) - \boldsymbol{W}\left(\tilde{\boldsymbol{q}}(t_{i+1})\right)\ \tilde{\boldsymbol{p}}_g(t_{i+1})\right)^{\mathsf{T}}\mathrm{d}\tilde{\boldsymbol{q}}(t_{i+1})\\ &= \tilde{\boldsymbol{p}}(t_{i+1})^{\mathsf{T}}\mathrm{d}\tilde{\boldsymbol{q}}(t_{i+1}) - \tilde{\boldsymbol{p}}_g(t_{i+1})^{\mathsf{T}}\mathrm{d}\boldsymbol{g}(\tilde{\boldsymbol{q}}(t_{i+1}))\\ &= \tilde{\boldsymbol{p}}(t_{i+1})^{\mathsf{T}}\mathrm{d}\tilde{\boldsymbol{q}}(t_{i+1}),\end{aligned}$$

which simplifies Equation (43) to

$$0 = \int_{\partial\Gamma} \tilde{\boldsymbol{p}}(t_{i+1})^{\mathsf{T}}\mathrm{d}\tilde{\boldsymbol{q}}(t_{i+1}) + \tilde{\boldsymbol{q}}(t_i)^{\mathsf{T}}\mathrm{d}\tilde{\boldsymbol{p}}(t_i).$$

Applying Stoke's theorem [28, p. 411] to the previous equation yields

$$\int_{\Gamma} \mathrm{d}\tilde{\boldsymbol{q}}(t_i) \wedge \mathrm{d}\tilde{\boldsymbol{p}}(t_i) = \int_{\Gamma} \mathrm{d}\tilde{\boldsymbol{q}}(t_{i+1}) \wedge \mathrm{d}\tilde{\boldsymbol{p}}(t_{i+1}), \tag{44}$$

which shows the invariance of the two-form $\mathrm{d}\boldsymbol{q} \wedge \mathrm{d}\boldsymbol{p}$ and implies symplectic integration.

Note that the integration scheme remains symplectic even if the integrals occurring in (37) are approximated using quadrature (under mild assumptions on the quadrature rule); the argument showing symplectic integration is analogous.

5. RELATION TO RUNGE–KUTTA METHODS

In the next section, the presented integration methods are related to well-known Runge–Kutta methods. This is carried out by approximating the action integral with quadrature. Note that links between previously proposed variational integration schemes and Runge–Kutta methods are, for example, given in [1, 9, 14], and references therein. However, the integration scheme proposed herein is different, and therefore, the connection to the Runge–Kutta methods is shortly discussed.

To simplify notation, the derivations are presented on the example of a one-dimensional system ($n = 1$), where linear basis functions are chosen. Note that it is straightforward to extend the derivation to systems with more degrees of freedom or more complex basis functions.

Consider the interval [0, 1] with the following ansatz for $\tilde{q}$ and $\tilde{p}$:

$$\tilde{q}\left(t, \boldsymbol{\eta}_q\right) := \begin{cases} q_0, & t = 0 \\ \frac{1+\sqrt{3}}{2}\eta_{q0} + \frac{1-\sqrt{3}}{2}\eta_{q1} + \sqrt{3}\left(\eta_{q1} - \eta_{q0}\right)t, & t \in (0, 1], \end{cases} \tag{45}$$

$$\tilde{p}(t, \boldsymbol{\eta}_p) := \begin{matrix} \frac{1+\sqrt{3}}{2}\eta_{p0} + \frac{1-\sqrt{3}}{2}\eta_{p1} + \sqrt{3}\left(\eta_{p1} - \eta_{p0}\right)t, & t \in [0, 1) \\ p_1, & t = 1, \end{matrix} \tag{46}$$

where $\boldsymbol{\eta}_q = (\eta_{q0}, \eta_{q1})^{\mathsf{T}}$ and $\boldsymbol{\eta}_p = (\eta_{p0}, \eta_{p1})^{\mathsf{T}}$. The dependence of $\tilde{q}$ on q_0 and $\tilde{p}$ on p_1 is omitted to simplify notation. Note that the ansatz is parametrized in such a way that the evaluation at the Gauss quadrature points $t_{g1} = \frac{1}{2} - \frac{\sqrt{3}}{6}$ and $t_{g2} = \frac{1}{2} + \frac{\sqrt{3}}{6}$ corresponds exactly to the ansatz parameters, that is, $\tilde{q}\left(t = t_{g0}, \boldsymbol{\eta}_q\right) = \eta_{q0}$ and $\tilde{q}\left(t = t_{g1}, \boldsymbol{\eta}_q\right) = \eta_{q1}$. The same applies for $\tilde{p}$.

Next, the ansatz is inserted in Equation (21). However, instead of solving the integrals exactly, they are approximated by Gauss quadrature, which leads to

$$\begin{aligned}\mathbf{0} = -\left.\frac{\partial\tilde{q}}{\partial\boldsymbol{\eta}_q}\right|_{t\downarrow 0}^{\mathsf{T}} \tilde{p}(0)^+ + \left.\frac{\partial\tilde{q}}{\partial\boldsymbol{\eta}_q}\right|_{t=1}^{\mathsf{T}} \tilde{p}(1) - \frac{1}{2}\sum_{j=0}^{1}\Bigg[&\left.\frac{\partial\dot{\tilde{q}}}{\partial\boldsymbol{\eta}_q}\right|_{t=t_{gj}}^{\mathsf{T}} \tilde{p}(t_{gj})\\ &+ \left.\frac{\partial\tilde{q}}{\partial\boldsymbol{\eta}_q}\right|_{t=t_{gj}}^{\mathsf{T}} \left(-\left.\frac{\partial H}{\partial q}\right|_{t=t_{gj}} + f_{NP}(t_{gj})\right)\Bigg],\end{aligned} \tag{47}$$

DOI: 10.1002/nme

$$\mathbf{0} = -\left.\frac{\partial \tilde{p}}{\partial \boldsymbol{\eta}_p}\right|_{t=0}^{\mathsf{T}} (\tilde{q}(0)^+ - \tilde{q}(0)) - \frac{1}{2}\sum_{j=0}^{1} \left.\frac{\partial \tilde{p}}{\partial \boldsymbol{\eta}_p}\right|_{t=t_{gj}}^{\mathsf{T}} \left(\dot{\tilde{q}}(t_{gj}) - \left.\frac{\partial H}{\partial p}\right|_{t=t_{gj}}\right). \tag{48}$$

Given initial conditions q_0 and p_0 such that $\tilde{q}(0) = q_0$ and $\tilde{p}(0) = p_0$, the previous set of equations can be solved for $\boldsymbol{\eta}_q$, $\boldsymbol{\eta}_p$, and p_1. This results in

$$\begin{aligned}
\eta_{q0} &= q_0 + \frac{1}{3}\left.\frac{\partial H}{\partial p}\right|_{t=t_{g0}} + \frac{1-\sqrt{3}}{6}\left.\frac{\partial H}{\partial p}\right|_{t=t_{g1}},\\
\eta_{q1} &= q_0 + \frac{\sqrt{3}+1}{6}\left.\frac{\partial H}{\partial p}\right|_{t=t_{g0}} + \frac{1}{3}\left.\frac{\partial H}{\partial p}\right|_{t=t_{g1}},\\
\tilde{q}(1) &= q_0 + \frac{1}{2}\left.\frac{\partial H}{\partial p}\right|_{t=t_{g0}} + \frac{1}{2}\left.\frac{\partial H}{\partial p}\right|_{t=t_{g1}},
\end{aligned}$$

and

$$\begin{aligned}
\eta_{p0} &= p_0 + \frac{1}{6}\left(-\left.\frac{\partial H}{\partial q}\right|_{t=t_{g0}} + f_{NP}(t_{g0})\right) + \left(\frac{1}{3} - \frac{\sqrt{3}}{6}\right)\left(-\left.\frac{\partial H}{\partial q}\right|_{t=t_{g1}} + f_{NP}(t_{g1})\right),\\
\eta_{p1} &= p_0 + \left(\frac{1}{3} + \frac{\sqrt{3}}{6}\right)\left(-\left.\frac{\partial H}{\partial q}\right|_{t=t_{g0}} + f_{NP}(t_{g0})\right) + \frac{1}{6}\left(-\left.\frac{\partial H}{\partial q}\right|_{t=t_{g1}} + f_{NP}(t_{g1})\right),\\
\tilde{p}(1) &= p_0 + \frac{1}{2}\left(-\left.\frac{\partial H}{\partial q}\right|_{t=t_{g0}} + f_{NP}(t_{g0})\right) + \frac{1}{2}\left(-\left.\frac{\partial H}{\partial q}\right|_{t=t_{g1}} + f_{NP}(t_{g1})\right).
\end{aligned}$$

Thus, the expression for the states at time $t = 1$, $\{\tilde{q}(1), \tilde{p}(1)\}$, corresponds exactly to a partitioned Runge–Kutta procedure with the Butcher tableau depicted in Figure 3.

However, in the Runge–Kutta setting, the trajectories are only evaluated at discrete time instants, that is, given initial conditions q_0 and p_0, the Runge–Kutta method provides the trajectories evaluated at time $t = 1$, $\tilde{q}(1)$ and $\tilde{p}(1)$. In contrast, the approach presented herein that provides approximate trajectories in the time interval $[0, 1]$ (c.f. Equations (45) and (46)) is guaranteed to be symplectic and conserves an energy-like quantity. Furthermore, the approach is flexible as the choice of basis functions or the quadrature rule is not restricted.

6. NUMERICAL EXAMPLES

Next, the variational integration methods are illustrated in two examples. In a first step, the numerical integration of the mathematical pendulum, parametrized by polar coordinates, is discussed. In a second step, the pendulum is described by Cartesian coordinates, providing the need to include holonomic constraints.

$$\begin{array}{c|cc}
\frac{1}{2} - \frac{\sqrt{3}}{6} & \frac{1}{3} & \frac{1-\sqrt{3}}{6}\\
\frac{1}{2} + \frac{\sqrt{3}}{6} & \frac{1+\sqrt{3}}{6} & \frac{1}{3}\\
\hline
 & \frac{1}{2} & \frac{1}{2}
\end{array}
\qquad
\begin{array}{c|cc}
\frac{1}{2} - \frac{\sqrt{3}}{6} & \frac{1}{6} & \frac{1}{3} - \frac{\sqrt{3}}{6}\\
\frac{1}{2} + \frac{\sqrt{3}}{6} & \frac{1}{3} + \frac{\sqrt{3}}{6} & \frac{1}{6}\\
\hline
 & \frac{1}{2} & \frac{1}{2}
\end{array}$$

Figure 3. Butcher tableau for $m_q = m_p = 1$. Left: update equations for q and right: update equations for p.

DOI: 10.1002/nme

6.1. The pendulum in polar coordinates

The pendulum is modeled as point mass (with mass m) subjected to gravity and constrained to the circle of radius l. In polar coordinates, the Hamiltonian of the system is given by

$$H(p,q) = \frac{1}{2ml^2}p^2 - mlg\cos q, \tag{49}$$

where $q \in [-\pi,\pi)$ refers to the angle, $p \in \mathbb{R}$ denotes the corresponding generalized momentum, and $g = 9.81\mathrm{m/s}^2$ denotes the gravitational acceleration. The angle is introduced such that $q = 0$ corresponds to the hanging equilibrium.

For simplicity, we choose polynomial ansatz functions over equidistant time intervals of length T for both $\tilde{q}$ and $\tilde{p}$, that is,

$$\tilde{q}\left(t,\boldsymbol{\eta}_q\right) = \begin{cases} q_0 & t = 0 \\ \eta_{q0} + \eta_{q1}t + \eta_{q2}t^2 & t \in (0,T] \\ \eta_{q3} + \eta_{q4}(t-T) + \eta_{q5}(t-T)^2 & t \in (T,2T] \\ \vdots \end{cases} \tag{50}$$

$$\tilde{p}(t,\boldsymbol{\eta}_p) = \begin{cases} \eta_{p0} + \eta_{p1}t + \eta_{p2}t^2 & t \in [0,T) \\ \eta_{p3} + \eta_{p4}(t-T) + \eta_{p5}(t-T)^2 & t \in [T,2T) \\ \vdots \end{cases}. \tag{51}$$

Next, we will consider the time interval $t \in [0,T]$ in more details. Because the Hamiltonian is not explicitly time-dependent, we can treat the subsequent time intervals analogously. For the time interval $t \in [0,T]$, the ansatz functions can be rewritten as

$$\tilde{q}\left(t,\boldsymbol{\eta}_{q0}\right) = \begin{cases} q_0 & t = 0 \\ \boldsymbol{\tau}(t)^\mathsf{T}\boldsymbol{\eta}_{q0} & t \in (0,T], \end{cases} \qquad \tilde{p}\left(t,\boldsymbol{\eta}_{p0}\right) = \begin{cases} \boldsymbol{\tau}(t)^\mathsf{T}\boldsymbol{\eta}_{p0} & t \in [0,T) \\ \eta_{p3} & t = T, \end{cases} \tag{52}$$

where $\boldsymbol{\tau}(t) := (1,\ t,\ t^2)^\mathsf{T}$, $\boldsymbol{\eta}_{q0} := (\eta_{q0},\eta_{q1},\eta_{q2})^\mathsf{T}$, and $\boldsymbol{\eta}_{p0} := (\eta_{p0},\eta_{p1},\eta_{p2})^\mathsf{T}$. The dependence on q_0 and η_{p3} is omitted to simplify notation. In Section 2, it was shown that requiring the virtual action to vanish for all variations $\delta\boldsymbol{\eta}_q$, $\delta\boldsymbol{\eta}_p$ is equivalent to impose that the set of equations (21) has to be fulfilled. By inserting the ansatz functions, we obtain

$$\begin{aligned} \boldsymbol{A}_q\left(\boldsymbol{\eta}_{q0},\boldsymbol{\eta}_{p0},\tilde{p}\left(T,\boldsymbol{\eta}_p\right)^+\right) &= -\boldsymbol{\tau}(0)\boldsymbol{\tau}(0)^\mathsf{T}\boldsymbol{\eta}_{p0} + \boldsymbol{\tau}(T)\tilde{p}\left(T,\boldsymbol{\eta}_p\right)^+ \\ &\quad - \int_{(0,T)} \dot{\boldsymbol{\tau}}\boldsymbol{\tau}^\mathsf{T}\mathrm{d}t\ \boldsymbol{\eta}_{p0} + mlg\int_{(0,T)} \boldsymbol{\tau}\sin\boldsymbol{\tau}^\mathsf{T}\boldsymbol{\eta}_{q0}\mathrm{d}t = \mathbf{0}, \end{aligned} \tag{53}$$

$$\begin{aligned} \boldsymbol{A}_p\left(\boldsymbol{\eta}_{q0},\boldsymbol{\eta}_{p0},\tilde{q}\left(0,\boldsymbol{\eta}_q\right)^-\right) &= -\boldsymbol{\tau}(0)\boldsymbol{\tau}(0)^\mathsf{T}\boldsymbol{\eta}_{q0} + \boldsymbol{\tau}(0)\tilde{q}\left(0,\boldsymbol{\eta}_q\right) \\ &\quad - \int_{(0,T)} \boldsymbol{\tau}\dot{\boldsymbol{\tau}}^\mathsf{T}\mathrm{d}t\ \boldsymbol{\eta}_{q0} + \frac{1}{ml^2}\int_{(0,T)} \boldsymbol{\tau}\boldsymbol{\tau}^\mathsf{T}\mathrm{d}t\ \boldsymbol{\eta}_{p0} = \mathbf{0}. \end{aligned} \tag{54}$$

The initial conditions for q and p, denoted by q_0^G and p_0^G, determine the value $\tilde{q}(0,\boldsymbol{\eta}_q)$ and $\tilde{p}(0,\boldsymbol{\eta}_p)$. They can be enforced by adding the two equations $\tilde{q}(0,\boldsymbol{\eta}_q) = q_0^G$ and $\tilde{p}(0,\boldsymbol{\eta}_p) = p_0^G$, or equivalently

$$q_0 - q_0^G = 0, \qquad \boldsymbol{\tau}(0)^\mathsf{T}\boldsymbol{\eta}_{p0} - p_0^G = 0, \tag{55}$$

DOI: 10.1002/nme

to the set of equations (53) and (54). By solving Equations (53), (54), and (55) for the unknowns $\boldsymbol{\eta}_{q0}$, $\boldsymbol{\eta}_{p0}$, q_0, and $\tilde{p}\left(T, \boldsymbol{\eta}_p\right)^+$, the approximate trajectories for the time interval $t \in [0, T]$ are obtained. The values $\tilde{q}\left(T, \boldsymbol{\eta}_q\right)$ and $\tilde{p}\left(T, \boldsymbol{\eta}_p\right)$, which is by right continuity equal to $\tilde{p}\left(T, \boldsymbol{\eta}_p\right)^+$, determine the initial conditions for the next time interval, that is, $t \in [T, 2T]$.

The numerical values of the different parameters are listed in Table I. To solve the nonlinear stepping equations, the Newton method with a relative tolerance of 10^{-13} has been used. The integrals in Equations (53) and (54) are calculated analytically, except for

$$\int_{(0,T)} \boldsymbol{\tau} \sin \boldsymbol{\tau}^{\mathsf{T}} \boldsymbol{\eta}_{q0} \mathrm{d}t,$$

which is approximated by the rectangle method using a width of 10^{-4}s. Figure 4 shows the phase portrait. The total energy is depicted in Figure 5, where it can be confirmed that the energy oscillates around the correct value. Both plots indicate a physically consistent integration. The root mean squared error is approximately $1.24 \cdot 10^{-4}$rad for the angle and $1.40 \cdot 10^{-2}$kg m^2/s for the angular

Table I. Parameters and initial conditions.

Mass	$m = 1$kg
Pendulum length	$l = 2$m
Time interval	$T = 0.3$s
Initial coordinates	$q_0^G = 120$deg
Initial momenta	$p_0^G = 0$kg m^2/s

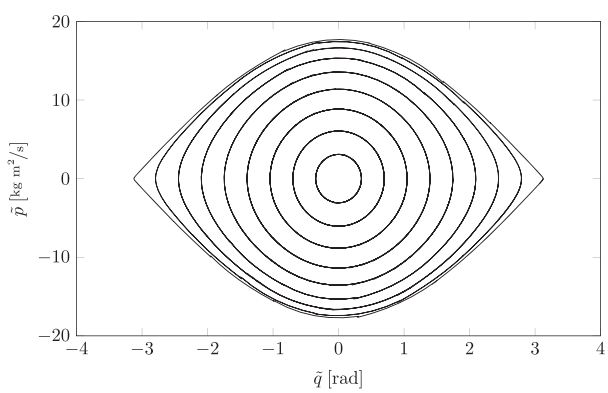

Figure 4. Phase diagram of the pendulum.

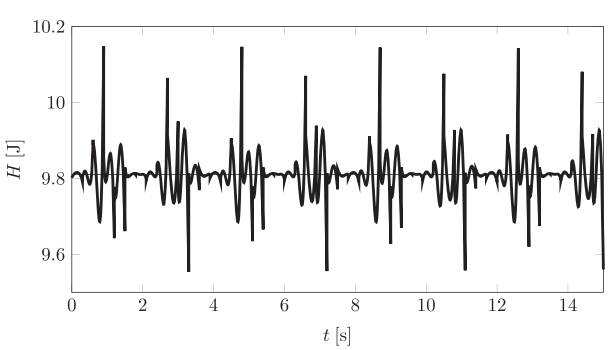

Figure 5. Total energy of the pendulum. The thin black line indicates the exact value of 9.81 J.

DOI: 10.1002/nme

momentum, which is small compared with the integration timesteps.[||||] Moreover, owing to the symplectic integration, an excellent long-time behavior (no energy drift) can be observed.

6.2. *Galerkin variational integration approach from* [1, p. 415]

We compare the integration scheme from the previous section with the Galerkin variational integrator approach presented in [1, p. 415].

Similarly, polynomial basis functions of second order over equidistant time intervals of length T are used to approximate the generalized coordinates $\tilde{q}$, that is,

$$\tilde{q}\left(t, \boldsymbol{\eta}_q\right) = \begin{cases} q_0 & t = 0 \\ \eta_{q0} + \eta_{q1} t + \eta_{q2} t^2 & t \in (0, T] \\ \eta_{q3} + \eta_{q4}(t - T) + \eta_{q5}(t - T)^2 & t \in (T, 2T]. \\ \vdots \end{cases} \tag{56}$$

In contrast to the approach presented in the previous section, the generalized coordinates are required to be continuous, which yields the constraints $q_0 = \eta_{q0}$, $\eta_{q3} = \eta_{q0} + \eta_{q1} T + \eta_{q2} T^2$, and so on. The integration scheme is derived by requiring the action

$$A = \int_{[0,NT]} L\left(\tilde{q}, \dot{\tilde{q}}\right) \mathrm{d}t = \sum_{k=0}^{N-1} \int_{(kT,(k+1)T]} L\left(\tilde{q}, \dot{\tilde{q}}\right) \mathrm{d}t \tag{57}$$

to be stationary with respect to the ansatz parameters $\boldsymbol{\eta}_q$ and with fixed boundaries, $\delta\tilde{q}(0) = 0$, $\delta\tilde{q}(NT) = 0$, where TN refers to the length of the prediction horizon. Note that the time singleton $\{0\}$ has zero Lebesgue measure, and therefore, the integral over $\{0\}$ vanishes in the right-hand side of Equation (57). The Lagrangian is given by the difference of kinetic and potential energy, which leads for the pendulum in polar coordinates to

$$L(q, \dot{q}) = \frac{1}{2} m l^2 \dot{q}^2 + m l g \cos q. \tag{58}$$

We impose the continuity requirements on $\tilde{q}$ using the Lagrange multipliers $p_k, k = 0, 1, \ldots, N-1$ and augment the action (57) to

$$A = \int_0^{NT} L\left(\tilde{q}, \dot{\tilde{q}}\right) \mathrm{d}t + \sum_{k=0}^{N-1} p_k \left(\tilde{q}\left(kT, \boldsymbol{\eta}_q\right)^+ - \tilde{q}\left(kT, \boldsymbol{\eta}_q\right)\right). \tag{59}$$

Consequently, the action can be made stationary by requiring its variation with respect to the ansatz parameters $\boldsymbol{\eta}_q$ to vanish (disregarding the continuity requirements on $\tilde{q}$). Variations with respect to the multipliers p_k impose the continuity requirements on $\tilde{q}$.

We consider the time interval $t \in [0, T]$ in more details and rewrite the ansatz function as

$$\tilde{q}\left(t, \boldsymbol{\eta}_{q0}\right) = \begin{cases} q_0 & t = 0 \\ \boldsymbol{\tau}(t)^{\mathsf{T}} \boldsymbol{\eta}_{q0} & t \in (0, T], \end{cases} \tag{60}$$

with $\boldsymbol{\tau}(t) = (1, t, t^2)^{\mathsf{T}}$, $\boldsymbol{\eta}_{q0} = \left(\eta_{q0}, \eta_{q1}, \eta_{q2}\right)^{\mathsf{T}}$. The dependence of $\tilde{q}$ on q_0 is omitted to simplify notation. Because the Lagrangian is not explicitly dependent on time, we can treat the subsequent time intervals analogously. Requiring the action to be stationary with respect to the ansatz parameters $\boldsymbol{\eta}_{q0}$ yields

[||||]The solutions from the variational integrator were compared with the trajectories obtained by MATLAB's ode45 (MathWorks, Natick, MA, USA) with relative and absolute tolerances of 10^{-12}.

DOI: 10.1002/nme

$$\delta A = \delta \boldsymbol{\eta}_{q0}^{\mathsf{T}} \left(\int_0^T \boldsymbol{\tau}(t) \frac{\partial L}{\partial q} + \dot{\boldsymbol{\tau}}(t) \frac{\partial L}{\partial \dot{q}} \mathrm{d}t + \boldsymbol{\tau}(0) p_0 - \boldsymbol{\tau}(T) p_1 \right) = 0, \tag{61}$$

for all variations $\delta \boldsymbol{\eta}_{q0}$. The variation with respect to p_0 and p_1 yields $\tilde{q}(0, \boldsymbol{\eta}_{q0})^+ = q_0$, $\tilde{q}(T, \boldsymbol{\eta}_{q0}) = \tilde{q}(T, \boldsymbol{\eta}_q)^+$, respectively, and imposes continuity on $\tilde{q}$.

Therefore, the update step is given by the implicit equation

$$\boldsymbol{A}_q \left(\boldsymbol{\eta}_{q0}, p_0, p_1 \right) = \boldsymbol{\tau}(0) p_0 - \boldsymbol{\tau}(T) p_1 + m l^2 \int_0^T \dot{\boldsymbol{\tau}} \dot{\boldsymbol{\tau}}^{\mathsf{T}} \mathrm{d}t \; \boldsymbol{\eta}_{q0} - m l g \int_0^T \boldsymbol{\tau} \sin \boldsymbol{\tau}^{\mathsf{T}} \boldsymbol{\eta}_{q0} \mathrm{d}t = \mathbf{0} \tag{62}$$

together with the boundary condition $q_0 = q_0^G$ and continuity requirement $\tilde{q}\left(0, \boldsymbol{\eta}_{q0}\right) = \tilde{q}\left(0, \boldsymbol{\eta}_{q0}\right)^+$. According to Marsden [1], the Lagrange multipliers p_0 and p_1 correspond to the generalized momentum at the time instants $t = 0$ and $t = T$. Imposing the boundary conditions $q_0 = q_0^G$ and $p_0 = p_0^G$ and the continuity requirement $\tilde{q}(0, \boldsymbol{\eta}_{q0}) = \tilde{q}(0, \boldsymbol{\eta}_{q0})^+$, Equation (62) can be solved for $\boldsymbol{\eta}_{q0}$ and p_1, yielding the approximate trajectory $\tilde{q}(t)$ in the time interval $t \in [0, T]$, as well as the generalized momentum p_1, which is used as boundary condition for the next time interval.

By comparing (62) with (53), it can be seen that the terms accounting for the potential energy agree. By defining the generalized momentum as $\tilde{p} := m l^2 \dot{\tilde{q}}$, the terms resulting from the kinetic energy can be matched as well. In the previous section, the generalized momentum was assumed to be right continuous, which implies $\tilde{p}(T, \boldsymbol{\eta}_p)^+ = \tilde{p}(T, \boldsymbol{\eta}_p)$. Hence, the boundary terms $\boldsymbol{\tau}(T) p_1$ and $\boldsymbol{\tau}(T) \tilde{p}(T, \boldsymbol{\eta}_p)^+$ can be identified with each other. The difference between Equations (62) and (53) lies therefore in the way the boundary condition at $t = 0$ is incorporated. Indeed, it turns out that the Galerkin variational integration approach as proposed in [1] can be derived as a special case from the framework presented herein. It suffices to assume the generalized coordinates to be continuous and the generalized momentum to be neither left nor right continuous. The corresponding derivations are presented in Appendix B.

The values for the pendulum mass, the gravitational acceleration, and the pendulum length are chosen as in the previous section (Table I). The nonlinear Equation (62), where the second integral is again approximated by the rectangle rule (with a width of 10^{-4}s), is solved using Newton's method with the same tolerance as in Section 6.1. The resulting phase portrait is depicted in Figure 6, where the discontinuities of the generalized momentum are clearly visible. Note that the generalized momentum is defined as $\tilde{p} := m l^2 \dot{\tilde{q}}$. For the initial conditions given in Table I, the root mean squared error of the integration scheme is compared with the approach presented in Section 6.1. The root mean squared error is approximated by averaging over a time period of 15 s, which corresponds to approximately five pendulum swings. The method from Section 6.1 has a smaller error in the angle variable q as well as in the momentum p (c.f. Table II).

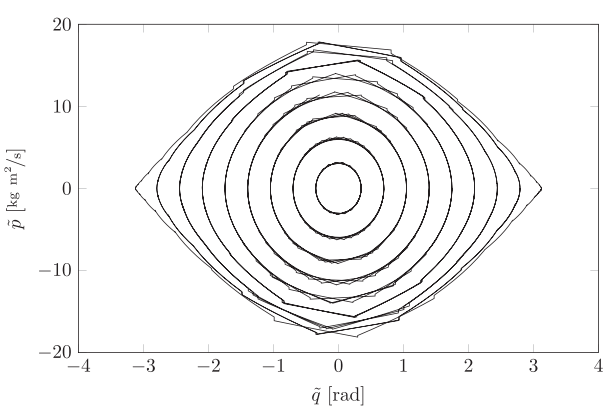

Figure 6. Phase diagram of the pendulum, simulated according to the Galerkin variational integration approach presented in [1].

DOI: 10.1002/nme

Table II. Root mean squared error of the methods presented in Section 6.1 and 6.2.

	rms error q	rms error p
Method from Sec. 6.1	$1.24 \cdot 10^{-4}$rad	$1.40 \cdot 10^{-2}$kg m^{2}s^{-1}
Method from Sec. 6.2	$1.48 \cdot 10^{-3}$rad	$1.14 \cdot 10^{-1}$kg m^{2}s^{-1}

Note that the root mean squared error of the method presented in 6.2 is roughly one order of magnitude larger.

6.3. *The pendulum in Cartesian coordinates*

Next, a description based on Cartesian coordinates is presented. Hence, the Hamiltonian is given by

$$H(\boldsymbol{p}, \boldsymbol{q}) = \frac{1}{2m}\left(p_1^2 + p_2^2\right) + mg\, q_2, \tag{63}$$

where $\boldsymbol{q} = (q_1, q_2)^\mathsf{T}$ denotes the position of the point mass and $\boldsymbol{p} = (p_1, p_2)^\mathsf{T}$ the corresponding generalized momentum. Additionally, the point mass is constrained to a circle of radius l, which leads to the gap function

$$g(\boldsymbol{q}) = q_1^2 + q_2^2 - l^2. \tag{64}$$

Again, polynomials up to second order are taken as ansatz functions. Similar to Equations (50) and (51), a left continuous ansatz $\tilde{\boldsymbol{q}}(t, \boldsymbol{\eta}_q)$ is chosen for $\boldsymbol{q}$ and a right continuous ansatz $\tilde{\boldsymbol{p}}(t, \boldsymbol{\eta}_p)$ and $\tilde{p}_g(t, \boldsymbol{\eta}_{p_g})$ for $\boldsymbol{p}$ and p_g. We recall that the generalized momentum p_g is associated with the constraint $g(\boldsymbol{q}) = 0$ and that $\dot{p}_g$ represents the constraint force. The Hamiltonian is not explicitly dependent on time, and therefore, it is enough to consider the interval $t \in [0, T]$ in more details, as the remaining time intervals are treated in a similar manner. In the interval $t \in [0, T]$, we choose the following parametrization of the generalized coordinates

$$\tilde{\boldsymbol{q}}\left(t, \boldsymbol{\eta}_q\right) = \begin{pmatrix} 1 & t & t^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & t & t^2 \end{pmatrix} \begin{pmatrix} \eta_{q0} \\ \vdots \\ \eta_{q5} \end{pmatrix} = \boldsymbol{\tau}(t)^\mathsf{T} \boldsymbol{\eta}_{q0}, \qquad t \in (0, T], \tag{65}$$

and similarly for the generalized momenta

$$\tilde{\boldsymbol{p}}\left(t, \boldsymbol{\eta}_p\right) = \boldsymbol{\tau}(t)^\mathsf{T} \boldsymbol{\eta}_{p0}, \qquad \tilde{p}_g\left(t, \boldsymbol{\eta}_{p_g}\right) = (1,\ t,\ t^2) \begin{pmatrix} \eta_{p_{g0}} \\ \eta_{p_{g1}} \\ \eta_{p_{g2}} \end{pmatrix} = \boldsymbol{\tau}_g(t)^\mathsf{T} \boldsymbol{\eta}_{p_{g0}}, \qquad t \in [0, T). \tag{66}$$

From Section 4, we infer that the virtual action vanishes for all variations $\delta\tilde{\boldsymbol{q}}$, $\delta\tilde{\boldsymbol{p}}$, and $\delta\tilde{p}_g$ if and only if the set of equations (37) is fulfilled. This yields

$$\begin{aligned} \boldsymbol{A}_q\left(\boldsymbol{\eta}_{q0}, \boldsymbol{\eta}_{p0}, \boldsymbol{\eta}_{p_{g0}}, \tilde{\boldsymbol{p}}_z(T)\right) &= -\boldsymbol{\tau}(0)\left(\boldsymbol{\tau}(0)^\mathsf{T}\boldsymbol{\eta}_{p0} - \boldsymbol{W}\left(\boldsymbol{\tau}(0)^\mathsf{T}\boldsymbol{\eta}_{q0}\right)\ \boldsymbol{\tau}_g(0)^\mathsf{T}\boldsymbol{\eta}_{p_{g0}}\right) + \boldsymbol{\tau}(0)\ \tilde{\boldsymbol{p}}_z(T) \\ &\quad - \int_{(0,T)} \dot{\boldsymbol{\tau}}\boldsymbol{\tau}^\mathsf{T}\mathrm{d}t\ \boldsymbol{\eta}_{p0} + \int_{(0,T)} \dot{\boldsymbol{\tau}}\ \boldsymbol{W}\left(\boldsymbol{\tau}^\mathsf{T}\boldsymbol{\eta}_{q0}\right)\ \boldsymbol{\tau}_g^\mathsf{T}\ \mathrm{d}t\ \boldsymbol{\eta}_{p_{g0}} + \int_{(0,T)} \boldsymbol{\tau}\begin{pmatrix} 0 \\ mg \end{pmatrix}\mathrm{d}t = \boldsymbol{0}, \end{aligned} \tag{67}$$

$$\begin{aligned} \boldsymbol{A}_p\left(\boldsymbol{\eta}_{q0}, \boldsymbol{\eta}_{p0},\ \tilde{\boldsymbol{q}}\left(0, \boldsymbol{\eta}_q\right)^-\right) &= -\boldsymbol{\tau}(0)\boldsymbol{\tau}(0)^\mathsf{T}\boldsymbol{\eta}_{q0} + \boldsymbol{\tau}(0)\tilde{\boldsymbol{q}}(0, \boldsymbol{\eta}_q)^- \\ &\quad - \int_{(0,T)} \boldsymbol{\tau}\dot{\boldsymbol{\tau}}^\mathsf{T}\mathrm{d}t\ \boldsymbol{\eta}_{q0} + \frac{1}{m}\int_{(0,T)} \boldsymbol{\tau}\boldsymbol{\tau}^\mathsf{T}\mathrm{d}t\ \boldsymbol{\eta}_{p0} = \boldsymbol{0}, \end{aligned} \tag{68}$$

DOI: 10.1002/nme

$$\boldsymbol{A}_g\left(\boldsymbol{\eta}_{q0}, \boldsymbol{\eta}_{p0}, \tilde{\boldsymbol{q}}(0, \boldsymbol{\eta}_q)^-\right) = \boldsymbol{\tau}_g(0)\left[\, g(\boldsymbol{\tau}(0)^\top \boldsymbol{\eta}_{q0}) - g\left(\tilde{\boldsymbol{q}}(0, \boldsymbol{\eta}_q)^-\right) \,\right] + \int\limits_{(0,T)} \boldsymbol{\tau}_g \boldsymbol{W}\left(\boldsymbol{\tau}^\top \boldsymbol{\eta}_{q0}\right)^\top \dot{\boldsymbol{\tau}}^\top \,\mathrm{d}t\; \boldsymbol{\eta}_{q0} = \boldsymbol{0}, \tag{69}$$

where the matrix $\boldsymbol{W}(\boldsymbol{q})$ is given by

$$\boldsymbol{W}(\boldsymbol{q}) = \frac{\partial g}{\partial \boldsymbol{q}}^\top = \begin{pmatrix} 2q_1 \\ 2q_2 \end{pmatrix}. \tag{70}$$

The initial conditions $\boldsymbol{q}_0^G$, $\boldsymbol{p}_0^G$ are set by imposing $\tilde{\boldsymbol{q}}(0, \boldsymbol{\eta}_q) = \boldsymbol{q}_0^G$ and $\tilde{\boldsymbol{p}}_z(0) = \boldsymbol{p}_0^G$, or equivalently

$$\tilde{\boldsymbol{q}}\left(0, \boldsymbol{\eta}_q\right)^- - \boldsymbol{q}_0^G = \boldsymbol{0}, \qquad \boldsymbol{\tau}(0)^\top \boldsymbol{\eta}_{p0} - \boldsymbol{W}\left(\boldsymbol{\tau}(0)^\top \boldsymbol{\eta}_q\right)\, \boldsymbol{\tau}_g(0)^\top \boldsymbol{\eta}_{p_{g0}} - \boldsymbol{p}_0^G = \boldsymbol{0}. \tag{71}$$

Solving Equations (67), (68), (69), and (71) yields the parameters $\boldsymbol{\eta}_{q0}$, $\boldsymbol{\eta}_{p0}$, and $\boldsymbol{\eta}_{p_{g0}}$ as well as the boundary term $\tilde{\boldsymbol{p}}_z(T)$. This determines the approximate trajectories $\tilde{\boldsymbol{q}}(t, \boldsymbol{\eta}_q)$, $\tilde{\boldsymbol{p}}(t, \boldsymbol{\eta}_p)$, and $\tilde{p}_g(t, \boldsymbol{\eta}_{p_g})$, for $t \in [0, T)$, as well as the boundary conditions $\tilde{\boldsymbol{q}}(T, \boldsymbol{\eta}_q)$ and $\tilde{\boldsymbol{p}}_z(T)$ for the next time interval. Note that no initial (or boundary) condition is prescribed on the momentum p_g. This allows the integration algorithm to 'absorb' initial or boundary conditions that are not necessarily in agreement with $\frac{\mathrm{d}}{\mathrm{d}t} g(\boldsymbol{q}(0)) = 0$. As the constraint force is given by the time derivative of p_g, the initial value of $p_g(0)$ is not defined in a mechanical sense. For a physically consistent integration, the simulation should be started with initial conditions $\boldsymbol{q}_0^G$ and $\boldsymbol{p}_0^G$ compatible with the constraint, such that $g\left(\boldsymbol{q}_0^G\right) = 0$ and $\frac{\mathrm{d}}{\mathrm{d}t} g\left(\boldsymbol{q}_0^G\right) = \frac{1}{m} \boldsymbol{W}\left(\boldsymbol{q}_0^G\right)^\top \boldsymbol{p}_0^G = 0$.

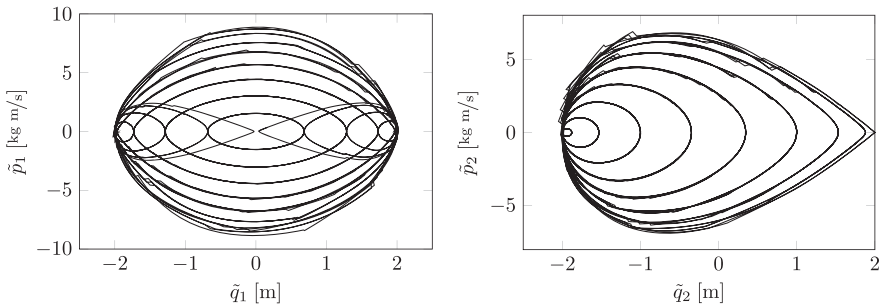

Figure 7. Phase diagram of the x-coordinate q_1 and p_1 (left) and the y-coordinate q_2 and p_2 (right). At higher angular velocities, the discontinuities in position and momentum are clearly visible.

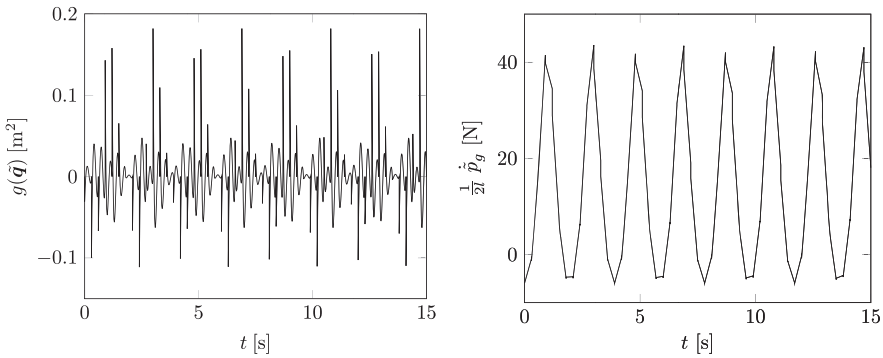

Figure 8. Depicted is the gap function g (left) and the constraint force $\dot{p}_g$ (right). The constraint force $\dot{p}_g$ is scaled by the magnitude of the vector $\boldsymbol{W}(\boldsymbol{q})$ in order to match the mechanical force experienced by the rod. Note that the gap function vanishes only at the time interval boundaries, that is, for $t = iT, i = 0, 1, \ldots, N$ (up to numerical accuracy), which is not visible owing to the constraint violations for $t \in (iT, (i+1)T)$. Furthermore, the ansatz for p_g is of second order, and therefore, the constraint force $\dot{\tilde{p}}_g$ is piecewise linear.

DOI: 10.1002/nme

The parameters given in Table I are used in the numerical experiment. The nonlinear stepping equations are solved by the Newton procedure with a relative tolerance of 10^{-13}. The resulting phase diagrams are shown in Figure 7. Figure 8 shows the time history of the gap function, as well as the constraint force. Again, a physically sensible integration is obtained. Owing to the symplectic integration, a consistent long-time behavior is observed. Despite the comparably large step size of $T = 0.3$s, the root mean squared position error is below $2.26 \cdot 10^{-2}$m and the root mean squared momentum error below $1.51 \cdot 10^{-1}$kg m/s. Note that no additional constraint stabilization was used. Moreover, the Newton procedure used for solving the stepping equations converged in less than 15 iterations for each timestep, indicating a numerically stable algorithm.

7. CONCLUSION

This article discussed a particular approach to variational integration. After deriving integration schemes from the principle of virtual work, some of their properties were analyzed and provided an a posteriori justification. Next, the addition of holonomic constraints was elaborated. Provided that sensible ansatz functions are chosen, the presented framework guarantees symplectic integration and the conservation of an energy-like quantity without additional regularization. Application of the method to a simple mechanical system including a holonomic constraint yielded promising results and demonstrated physically consistent integration. In particular, a numerically stable long-time behavior was observed even for large timesteps.

Simplifying the integration algorithm by approximating the action integral by quadrature showed a connection to the well-known Runge–Kutta methods.

We believe that the close relation to continuous Hamiltonian mechanics, which enables to transfer classical results, combined with a physically consistent long-time behavior (due to the fact that the transformation from one time interval to the next is symplectic) indicates the potential of the proposed integration methods.

APPENDIX A: PROOF OF POSTULATE 2.1

We prove Postulate 2.1 in two steps. Starting from the principle of virtual work, we conclude that the system is in dynamic equilibrium if the corresponding equations of motion are fulfilled.[***] Using this fact, we then show that Postulate 2.1 holds.

According to [21], the principle of virtual work states that a mechanical system is in dynamic equilibrium for times $t \in I := [t_0, t_N]$ if the virtual work

$$\delta W := \frac{\mathrm{d}}{\mathrm{d}t}\left(\delta \boldsymbol{q}^{\top} \boldsymbol{p}\right) - \delta L - \delta \boldsymbol{q}^{\top} \boldsymbol{f}_{NP} \tag{A.1}$$

vanishes for all variations $\delta \boldsymbol{q}$ and all times $t \in I$. Note that the Lagrangian L is defined as

$$L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) := \frac{1}{2}\dot{\boldsymbol{q}}^{\top} \boldsymbol{M}(\boldsymbol{q})\dot{\boldsymbol{q}} - V(\boldsymbol{q}, t) \tag{A.2}$$

and the generalized momenta $\boldsymbol{p}$ are given by Equation (2).

The transition from Lagrangian to Hamiltonian formalism is carried out in two steps. First the Lagrangian is rewritten as

$$L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = \dot{\boldsymbol{q}}^{\top} \boldsymbol{p} - H(\boldsymbol{q}, \boldsymbol{p}, t), \tag{A.3}$$

where the generalized momenta $\boldsymbol{p}$ are considered to be a placeholder for $\boldsymbol{M}(\boldsymbol{q})\dot{\boldsymbol{q}}$ according to Equation (2) and the Hamiltonian is defined in Equation (3). Then, the kinematic link between

[***] We assume that the generalized coordinates $\boldsymbol{q}$ are at least twice continuously differentiable for times $t \in I$, such that the equations of motion are well defined for all $t \in I$.

DOI: 10.1002/nme

the generalized coordinates and the generalized momenta is enforced by (independent) variations of the generalized momenta. Therefore, the variation of the Lagrangian in Equation (A.1) can be replaced by

$$\delta\left(\dot{\boldsymbol{q}}^{\top}\boldsymbol{p} - H(\boldsymbol{q},\boldsymbol{p},t)\right), \tag{A.4}$$

where the variations of $\delta\boldsymbol{p}$ are independent of $\delta\boldsymbol{q}$.

As a result, it follows that the system is in dynamic equilibrium for times $t \in I$ if the virtual work

$$\delta W = \frac{\mathrm{d}}{\mathrm{d}t}\left(\delta\boldsymbol{q}^{\top}\boldsymbol{p}\right) - \delta\left(\dot{\boldsymbol{q}}^{\top}\boldsymbol{p}\right) + \delta H - \delta\boldsymbol{q}^{\top}\boldsymbol{f}_{NP} \tag{A.5}$$

vanishes for all variations $\delta\boldsymbol{q}$ and all variations $\delta\boldsymbol{p}$, and for all $t \in I$.

Using the fact that time differentiation and variation commute, the virtual work can be rewritten as

$$\delta W = \delta\boldsymbol{q}^{\top}\left(\dot{\boldsymbol{p}} + \frac{\partial H}{\partial \boldsymbol{q}}^{\top} - \boldsymbol{f}_{NP}\right) + \delta\boldsymbol{p}^{\top}\left(-\dot{\boldsymbol{q}} + \frac{\partial H}{\partial \boldsymbol{p}}^{\top}\right). \tag{A.6}$$

Therefore, the principle of the virtual work implies that the system is in dynamic equilibrium for times $t \in I$ if

$$\dot{\boldsymbol{p}} = -\frac{\partial H}{\partial \boldsymbol{q}}^{\top} + \boldsymbol{f}_{NP}, \quad \dot{\boldsymbol{q}} = \frac{\partial H}{\partial \boldsymbol{p}}^{\top} \tag{A.7}$$

is fulfilled for times $t \in I$.

We now show that Postulate 2.1 holds by proving that if the virtual action in (7) vanishes for all (discontinuous) variations $\delta\boldsymbol{q}$, $\delta\boldsymbol{p}$, the equations of motion are fulfilled almost everywhere and the trajectories are continuous, except at finitely many discontinuous time instants where left and right limits agree. The generalized coordinates $\boldsymbol{q}$ and the generalized momenta $\boldsymbol{p}$ are assumed to be piecewise continuously differentiable. Therefore, we can choose variations $\delta\boldsymbol{q}$ and $\delta\boldsymbol{p}$, which are everywhere zero except on an interval $(t_0, t_1) \subset I$, where $\boldsymbol{q}$ and $\boldsymbol{p}$ are continuously differentiable. For these specific variations, Equation (7) reduces to

$$\begin{aligned}\int_{(t_0,t_1)} \delta\boldsymbol{q}^{\top}\dot{\boldsymbol{p}} - \dot{\boldsymbol{q}}^{\top}\delta\boldsymbol{p} + \delta H - \delta\boldsymbol{q}^{\top}\boldsymbol{f}_{NP}\,\mathrm{d}t \\ = \int_{(t_0,t_1)} \delta\boldsymbol{q}^{\top}\left(\dot{\boldsymbol{p}} + \frac{\partial H}{\partial \boldsymbol{q}}^{\top} - \boldsymbol{f}_{NP}\right) + \delta\boldsymbol{p}^{\top}\left(-\dot{\boldsymbol{q}} + \frac{\partial H}{\partial \boldsymbol{p}}^{\top}\right)\mathrm{d}t,\end{aligned} \tag{A.8}$$

and has to vanish for all $\delta\boldsymbol{q}$ and $\delta\boldsymbol{p}$ (which are zero except on the interval (t_0, t_1)). The fundamental lemma of the calculus of variations [22] implies that (A.8) vanishes if and only if

$$\dot{\boldsymbol{q}} = \frac{\partial H}{\partial \boldsymbol{p}}^{\top}, \quad \dot{\boldsymbol{p}} = -\frac{\partial H}{\partial \boldsymbol{q}}^{\top} + \boldsymbol{f}_{NP}, \tag{A.9}$$

for all $t \in (t_0, t_1)$. Hence, the system is in dynamic equilibrium for all $t \in (t_0, t_1)$. Repeating the same argument for all time intervals where $\boldsymbol{q}$ and $\boldsymbol{p}$ are continuously differentiable leads to the conclusion that the system fulfills the equations of motion (A.7) for all $t \in I$ (almost everywhere).

It remains to show that the left and right limits of the generalized coordinates and momenta at discontinuous time instants agree in order to conclude that the system is in dynamic equilibrium for all $t \in I$ (almost everywhere). Note that by assumption, there are finitely many time instants at which $\boldsymbol{q}$ and $\boldsymbol{p}$ are allowed to be discontinuous. We choose particular variations $\delta\boldsymbol{q}$, $\delta\boldsymbol{p}$, which are everywhere zero, except at the discontinuous time instant t_d. The virtual action simplifies in that

DOI: 10.1002/nme

case to the integral over the time singleton $\{t_d\}$. By linearity of the integral and the fact that the differential measure has a density with respect to an atomic measure, we obtain

$$\begin{aligned}\delta A &= \int_{\{t_d\}} \delta \boldsymbol{q}(t_d)^\top \mathrm{d}\boldsymbol{p} - \delta \boldsymbol{p}(t_d)^\top \mathrm{d}\boldsymbol{q} = \delta \boldsymbol{q}(t_d)^\top \int_{\{t_d\}} \mathrm{d}\boldsymbol{p} - \delta \boldsymbol{p}(t_d)^\top \int_{\{t_d\}} \mathrm{d}\boldsymbol{q} \\ &= \delta \boldsymbol{q}(t_d)^\top (\boldsymbol{p}(t_d)^+ - \boldsymbol{p}(t_d)^-) - \delta \boldsymbol{p}(t_d)^\top \left(\boldsymbol{q}(t_d)^+ - \boldsymbol{q}(t_d)^-\right).\end{aligned} \tag{A.10}$$

Requiring the virtual action to vanish for all variations $\delta \boldsymbol{q}(t_d)$ and $\delta \boldsymbol{p}(t_d)$ yields therefore

$$\boldsymbol{p}(t_d)^+ = \boldsymbol{p}(t_d)^-, \quad \boldsymbol{q}(t_d)^+ = \boldsymbol{q}(t_d)^-. \tag{A.11}$$

Repeating the same argument for all discontinuous time instants leads to the conclusion that the left and right limits agree at every discontinuous time instant. Therefore, the system is in dynamic equilibrium for all $t \in I$ (almost everywhere) if the virtual action, as defined in Postulate 2.1, vanishes for all variations $\delta \boldsymbol{q}$ and $\delta \boldsymbol{p}$.

APPENDIX B: ADDENDUM TO SECTION 6.2

We derive the integration scheme in case the generalized coordinates are assumed to be continuous and the generalized momenta are neither left nor right continuous. Thereby, we will show that the boundary terms match exactly the Galerkin variational integration approach proposed in [1, p. 415] (c.f. Section 6.2). By approximation of the resulting integrals using quadrature, the integration scheme can be related to the Galerkin variational integration approach in [2] and, in case of polynomial basis functions and a constant mass matrix, to the approach from [1, p. 415].

As in Section 2.4, the time interval of interest, $I = [t_0, t_N]$ with $t_N > t_0$, is divided into subintervals $I_i := [t_i, t_{i+1}] \subset \mathbb{R}$ $(i = 0, 1, \ldots, N-1)$. The two ansatz functions $\tilde{\boldsymbol{q}}(t)$ and $\tilde{\boldsymbol{p}}(t)$ for $\boldsymbol{q}$ and $\boldsymbol{p}$ are required to be everywhere continuously differentiable, except at the time instants $t_0, t_1, \ldots, t_N$. Hence, they are piecewise continuously differentiable, with

$$\tilde{\boldsymbol{q}} \in \tilde{C}^1_{\mathrm{pc}}\left(I \times \mathbb{R}^{n \cdot m_q \cdot N}, \mathbb{R}^n\right), \quad \text{and} \quad \tilde{\boldsymbol{p}} \in \tilde{C}^1_{\mathrm{pc}}\left(I \times \mathbb{R}^{n \cdot m_p \cdot N}, (\mathbb{R}^n)^*\right). \tag{B.1}$$

Inserting the ansatz into the expression of the virtual action leads to

$$\delta A = \int_I \delta \tilde{\boldsymbol{q}}^\top \mathrm{d}\tilde{\boldsymbol{p}} - \delta \tilde{\boldsymbol{p}}^\top \mathrm{d}\tilde{\boldsymbol{q}} + \left(\delta H\left(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{q}}, t\right) - \delta \tilde{\boldsymbol{q}}^\top \boldsymbol{f}_{NP}\right) \mathrm{d}t. \tag{B.2}$$

We require the virtual action to vanish for all variations $\delta \tilde{\boldsymbol{q}}$ and $\delta \tilde{\boldsymbol{p}}$ to obtain an approximation of the dynamic equilibrium. However, $\tilde{\boldsymbol{q}}$ is required to be continuous, and therefore, the variations $\delta \tilde{\boldsymbol{q}}$ must be continuous as well. It is assumed that within the intervals $t \in (t_i, t_{i+1}]$, the ansatz $\tilde{\boldsymbol{q}}$ depends on time but only on a subset of the parameters $\boldsymbol{\eta}_q$, denoted by $\boldsymbol{\eta}_{q_i} \in \mathbb{R}^{n \cdot m_q}$. The continuity of $\tilde{\boldsymbol{q}}$ imposes restrictions on the ansatz parameters $\boldsymbol{\eta}_q$, that is

$$\tilde{\boldsymbol{q}}\left(t_i, \boldsymbol{\eta}_{q_{i-1}}\right)^- = \tilde{\boldsymbol{q}}\left(t_i, \boldsymbol{\eta}_{q_i}\right),$$

for all $i = 0, 1, \ldots, N-1$. As a result, the variations of $\boldsymbol{\eta}_{q_i}$ are not allowed to be arbitrary because it must hold that

$$\delta \tilde{\boldsymbol{q}}(t_i)^- = \left.\frac{\partial \tilde{\boldsymbol{q}}}{\partial \boldsymbol{\eta}_{q_{i-1}}}\right|_{t \uparrow t_i} \delta \boldsymbol{\eta}_{q_{i-1}} = \left.\frac{\partial \tilde{\boldsymbol{q}}}{\partial \boldsymbol{\eta}_{q_i}}\right|_{t = t_i} \delta \boldsymbol{\eta}_{q_i} = \delta \tilde{\boldsymbol{q}}(t_i).$$

DOI: 10.1002/nme

In the following, we impose the continuity requirements on $\tilde{\boldsymbol{q}}$ using the Lagrange multipliers $\boldsymbol{\lambda}_i \in (\mathbb{R}^n)^*$, $i = 1, 2, \ldots, N-1$, such that the parameters $\boldsymbol{\eta}_{q_i}$ and corresponding variations $\delta\boldsymbol{\eta}_{q_i}$ can be treated as if they were unconstrained.[†††] Hence, the virtual action is augmented with the terms

$$\sum_{i=0}^{N-1} \delta\left(\boldsymbol{\lambda}_i^\top\left(\tilde{\boldsymbol{q}}(t_i)^+ - \tilde{\boldsymbol{q}}(t_i)\right)\right) = \sum_{i=0}^{N-1} \delta\boldsymbol{\lambda}_i^\top\left(\tilde{\boldsymbol{q}}(t_i)^+ - \tilde{\boldsymbol{q}}(t_i)\right) + \left(\delta\tilde{\boldsymbol{q}}(t_i)^+\right)^\top\boldsymbol{\lambda}_i - \delta\tilde{\boldsymbol{q}}(t_i)^\top\boldsymbol{\lambda}_i, \quad \text{(B.3)}$$

resulting in

$$\begin{aligned}\delta A = &\int_I \delta\tilde{\boldsymbol{q}}^\top \mathrm{d}\tilde{\boldsymbol{p}} - \delta\tilde{\boldsymbol{p}}^\top \mathrm{d}\tilde{\boldsymbol{q}} + \left(\delta H(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{q}}, t) - \delta\tilde{\boldsymbol{q}}^\top \boldsymbol{f}_{NP}\right)\mathrm{d}t \\ &+ \sum_{i=0}^{N-1} \delta\boldsymbol{\lambda}_i^\top\left(\tilde{\boldsymbol{q}}(t_i)^+ - \tilde{\boldsymbol{q}}(t_i)\right) + \left(\delta\tilde{\boldsymbol{q}}(t_i)^+\right)^\top\boldsymbol{\lambda}_i - \delta\tilde{\boldsymbol{q}}(t_i)^\top\boldsymbol{\lambda}_i,\end{aligned} \quad \text{(B.4)}$$

where the variations $\delta\boldsymbol{\eta}_{q_i} \in \mathbb{R}^{nm_q}$ are unconstrained, $i = 0, 1, \ldots, N-1$. Moreover, if the variations are restricted to be such that $\delta\tilde{\boldsymbol{q}}(t_i)^- = \delta\tilde{\boldsymbol{q}}(t_i)$, $i = 0, 1, \ldots, N-1$, the expression of the virtual action given in Equation (B.4) reduces to the one given in (B.2).

For $t \in (t_i, t_{i+1})$, the ansatz $\tilde{\boldsymbol{p}}$, which is neither left nor right continuous, is assumed to depend on time and only on a subset of the parameters $\boldsymbol{\eta}_p$, denoted by $\boldsymbol{\eta}_{p_i} \in \mathbb{R}^{n \cdot m_p}$, $i = 0, 1, \ldots, N-1$. We fix $0 \leqslant i \leqslant N-1$ and consider variations $\delta\boldsymbol{\eta}_{q_i}$, $\delta\boldsymbol{\eta}_{p_i}$, and $\delta\tilde{\boldsymbol{p}}(t_i)$. In that way, we can 'cut' the time segment I_i out of the virtual action. By repeating the procedure for all time intervals I_i, we require the virtual action to vanish for all variations $\delta\tilde{\boldsymbol{q}}$ and $\delta\tilde{\boldsymbol{p}}$, thereby approximating the dynamic equilibrium in the entire interval $t \in I$. This yields

$$\begin{aligned}\delta A &= \int_{I_i} \delta\tilde{\boldsymbol{q}}^\top \mathrm{d}\tilde{\boldsymbol{p}} - \delta\tilde{\boldsymbol{p}}^\top \mathrm{d}\tilde{\boldsymbol{q}} + \left(\delta H\left(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{q}}, t\right) - \delta\tilde{\boldsymbol{q}}^\top \boldsymbol{f}_{NP}\right)\mathrm{d}t + \left(\delta\tilde{\boldsymbol{q}}(t_i)^+\right)^\top\boldsymbol{\lambda}_i - \delta\tilde{\boldsymbol{q}}(t_{i+1})^\top\boldsymbol{\lambda}_{i+1} \\ &= \int_{\{t_i\}} \delta\tilde{\boldsymbol{q}}^\top \mathrm{d}\tilde{\boldsymbol{p}} - \delta\tilde{\boldsymbol{p}}^\top \mathrm{d}\tilde{\boldsymbol{q}} + \int_{\{t_{i+1}\}} \delta\tilde{\boldsymbol{q}}^\top \mathrm{d}\tilde{\boldsymbol{p}} - \delta\tilde{\boldsymbol{p}}^\top \mathrm{d}\tilde{\boldsymbol{q}} + (\delta\tilde{\boldsymbol{q}}(t_i)^+)^\top\boldsymbol{\lambda}_i - \delta\tilde{\boldsymbol{q}}(t_{i+1})^\top\boldsymbol{\lambda}_{i+1} \\ &\quad + \int_{(t_i, t_{i+1})} \left(\delta\tilde{\boldsymbol{q}}^\top\dot{\tilde{\boldsymbol{p}}} - \delta\tilde{\boldsymbol{p}}^\top\dot{\tilde{\boldsymbol{q}}} + \delta H\left(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{q}}, t\right) - \delta\tilde{\boldsymbol{q}}^\top \boldsymbol{f}_{NP}\right)\mathrm{d}t.\end{aligned}$$

The integral over the time singleton $\{t_i\}$ reduces to

$$\begin{aligned}\int_{\{t_i\}} \delta\tilde{\boldsymbol{q}}^\top \mathrm{d}\tilde{\boldsymbol{p}} - \delta\tilde{\boldsymbol{p}}^\top \mathrm{d}\tilde{\boldsymbol{q}} &= \delta\tilde{\boldsymbol{q}}(t_i)^\top\left(\tilde{\boldsymbol{p}}(t_i)^+ - \tilde{\boldsymbol{p}}(t_i)^-\right) - \delta\tilde{\boldsymbol{p}}(t_i)^\top\left(\tilde{\boldsymbol{q}}(t_i)^+ - \tilde{\boldsymbol{q}}(t_i)^-\right) \\ &= -\delta\tilde{\boldsymbol{p}}(t_i)^\top\left(\tilde{\boldsymbol{q}}(t_i)^+ - \tilde{\boldsymbol{q}}(t_i)^-\right),\end{aligned}$$

because $\delta\tilde{\boldsymbol{q}}(t_i) = 0$ (only the parameters $\boldsymbol{\eta}_{q_i}$ and $\boldsymbol{\eta}_{p_i}$ are varied for now). Note that requiring the virtual action to vanish for variations of the Lagrange multiplier $\boldsymbol{\lambda}_i$ results in the requirement $\tilde{\boldsymbol{q}}(t_i) = \tilde{\boldsymbol{q}}(t_i)^+$. The ansatz $\tilde{\boldsymbol{q}}$ is by construction left continuous, and therefore, $\tilde{\boldsymbol{q}}(t_i)^- = \tilde{\boldsymbol{q}}(t_i)$. As a result, the virtual action evaluated for variations induced by $\delta\boldsymbol{\eta}_{q_i}$, $\delta\boldsymbol{\eta}_{p_i}$, and $\delta\tilde{\boldsymbol{p}}(t_i)$ simplifies to

[†††]This is without loss of generality because one could also work with constrained variations. The resulting integration scheme would yield the same trajectories (almost everywhere).

DOI: 10.1002/nme

$$
\begin{aligned}
\delta A = & \int\limits_{\{t_{i+1}\}} \delta\tilde{\boldsymbol{q}}^{\mathsf{T}} \mathrm{d}\tilde{\boldsymbol{p}} + \left(\delta\tilde{\boldsymbol{q}}(t_i)^{+}\right)^{\mathsf{T}} \boldsymbol{\lambda}_i - \delta\tilde{\boldsymbol{q}}(t_{i+1})^{\mathsf{T}} \boldsymbol{\lambda}_{i+1} \\
& + \int\limits_{(t_i,t_{i+1})} \left(\delta\tilde{\boldsymbol{q}}^{\mathsf{T}} \dot{\tilde{\boldsymbol{p}}} - \delta\tilde{\boldsymbol{p}}^{\mathsf{T}} \dot{\tilde{\boldsymbol{q}}} + \delta H\left(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{q}}, t\right) - \delta\tilde{\boldsymbol{q}}^{\mathsf{T}} \boldsymbol{f}_{NP}\right) \mathrm{d}t \\
= & \, \delta\tilde{\boldsymbol{q}}(t_{i+1})^{\mathsf{T}} \left(\tilde{\boldsymbol{p}}(t_{i+1})^{+} - \boldsymbol{\lambda}_{i+1} - \tilde{\boldsymbol{p}}(t_{i+1})^{-}\right) + \left(\delta\tilde{\boldsymbol{q}}(t_i)^{+}\right)^{\mathsf{T}} \boldsymbol{\lambda}_i \\
& + \int\limits_{(t_i,t_{i+1})} \left(\delta\tilde{\boldsymbol{q}}^{\mathsf{T}} \dot{\tilde{\boldsymbol{p}}} - \delta\tilde{\boldsymbol{p}}^{\mathsf{T}} \dot{\tilde{\boldsymbol{q}}} + \delta H\left(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{q}}, t\right) - \delta\tilde{\boldsymbol{q}}^{\mathsf{T}} \boldsymbol{f}_{NP}\right) \mathrm{d}t.
\end{aligned}
\tag{B.5}
$$

The terms $\boldsymbol{\lambda}_{i+1}$ and $\tilde{\boldsymbol{p}}(t_{i+1})^{+}$, describing the coupling to the neighboring interval, appear only as linear combination. As in Section 4, where the quantity $\tilde{\boldsymbol{p}}_z$ was introduced, we define the generalized momentum at time t_i to be

$$
\tilde{\boldsymbol{p}}(t_i) := \tilde{\boldsymbol{p}}(t_i)^{+} - \boldsymbol{\lambda}_i, \tag{B.6}
$$

which will result in a causal integration scheme in accordance with the requirement that the virtual action has to vanish for all variations $\delta\tilde{\boldsymbol{q}}$ and $\delta\tilde{\boldsymbol{p}}$. This leads to the strong form of the virtual action,

$$
\begin{aligned}
\delta A = & \left(\delta\tilde{\boldsymbol{q}}(t_i)^{+}\right)^{\mathsf{T}} \left(\tilde{\boldsymbol{p}}(t_i)^{+} - \tilde{\boldsymbol{p}}(t_i)\right) + \delta\tilde{\boldsymbol{q}}(t_{i+1})^{\mathsf{T}} \left(\tilde{\boldsymbol{p}}(t_{i+1}) - \tilde{\boldsymbol{p}}(t_{i+1})^{-}\right) \\
& + \int\limits_{(t_i,t_{i+1})} \left(\delta\tilde{\boldsymbol{q}}^{\mathsf{T}} \dot{\tilde{\boldsymbol{p}}} - \delta\tilde{\boldsymbol{p}}^{\mathsf{T}} \dot{\tilde{\boldsymbol{q}}} + \delta H\left(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{q}}, t\right) - \delta\tilde{\boldsymbol{q}}^{\mathsf{T}} \boldsymbol{f}_{NP}\right) \mathrm{d}t.
\end{aligned}
\tag{B.7}
$$

Applying integration by parts on the integrand $\delta\tilde{\boldsymbol{q}}^{\mathsf{T}} \dot{\tilde{\boldsymbol{p}}}$ results in the weak form,

$$
\delta A = -\left(\delta\tilde{\boldsymbol{q}}(t_i)^{+}\right)^{\mathsf{T}} \tilde{\boldsymbol{p}}(t_i) + \delta\tilde{\boldsymbol{q}}(t_{i+1})^{\mathsf{T}} \tilde{\boldsymbol{p}}(t_{i+1}) - \int\limits_{(t_i,t_{i+1})} \left(\delta\left(\dot{\tilde{\boldsymbol{q}}}^{\mathsf{T}} \tilde{\boldsymbol{p}} - H\left(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{q}}, t\right)\right) - \delta\tilde{\boldsymbol{q}}^{\mathsf{T}} \boldsymbol{f}_{NP}\right) \mathrm{d}t. \tag{B.8}
$$

By factorizing the time-independent parameter variations $\delta\boldsymbol{\eta}_{q_i}$ and $\delta\boldsymbol{\eta}_{p_i}$ out, the virtual action can be rewritten as

$$
\delta A = \delta\boldsymbol{\eta}_{q_i}^{\mathsf{T}} \boldsymbol{A}_q\left(\boldsymbol{\eta}_{q_i}, \boldsymbol{\eta}_{p_i}, \tilde{\boldsymbol{p}}(t_i), \tilde{\boldsymbol{q}}(t_i)\right) + \delta\boldsymbol{\eta}_{p_i}^{\mathsf{T}} \boldsymbol{A}_p\left(\boldsymbol{\eta}_{q_i}, \boldsymbol{\eta}_{p_i}\right), \tag{B.9}
$$

with

$$
\begin{aligned}
\boldsymbol{A}_q\left(\boldsymbol{\eta}_{q_i}, \boldsymbol{\eta}_{p_i}, \, \tilde{\boldsymbol{p}}(t_i), \tilde{\boldsymbol{p}}(t_{i+1})\right) := & -\left.\frac{\partial\tilde{\boldsymbol{q}}}{\partial\boldsymbol{\eta}_{q_i}}\right|_{t\downarrow t_i}^{\mathsf{T}} \tilde{\boldsymbol{p}}(t_i) + \left.\frac{\partial\tilde{\boldsymbol{q}}}{\partial\boldsymbol{\eta}_{q_i}}\right|_{t=t_{i+1}}^{\mathsf{T}} \tilde{\boldsymbol{p}}(t_{i+1}) \\
& - \int\limits_{(t_i,t_{i+1})} \left[\frac{\partial\dot{\tilde{\boldsymbol{q}}}}{\partial\boldsymbol{\eta}_{q_i}}^{\mathsf{T}} \tilde{\boldsymbol{p}} - \frac{\partial\tilde{\boldsymbol{q}}}{\partial\boldsymbol{\eta}_{q_i}}^{\mathsf{T}} \left(\frac{\partial H}{\partial\boldsymbol{q}}^{\mathsf{T}} - \boldsymbol{f}_{NP}\right)\right] \mathrm{d}t \in \mathbb{R}^{n\cdot m_q}
\end{aligned}
\tag{B.10}
$$

and

$$
\boldsymbol{A}_p\left(\boldsymbol{\eta}_{q_i}, \boldsymbol{\eta}_{p_i}\right) := -\int\limits_{(t_i,t_{i+1})} \frac{\partial\tilde{\boldsymbol{p}}}{\partial\boldsymbol{\eta}_{p_i}}^{\mathsf{T}} \left(\dot{\tilde{\boldsymbol{q}}} - \frac{\partial H}{\partial\boldsymbol{p}}^{\mathsf{T}}\right) \mathrm{d}t \in \mathbb{R}^{n\cdot m_p}. \tag{B.11}
$$

DOI: 10.1002/nme

It follows that δA vanishes for all $\delta \boldsymbol{\eta}_{q_i}$ and $\delta \boldsymbol{\eta}_{p_i}$ if and only if the set of equations

$$\begin{aligned} \boldsymbol{A}_q\left(\boldsymbol{\eta}_{q_i}, \boldsymbol{\eta}_{p_i}, \tilde{\boldsymbol{p}}(t_i), \tilde{\boldsymbol{p}}(t_{i+1})\right) &= \boldsymbol{0}, \\ \boldsymbol{A}_p\left(\boldsymbol{\eta}_{q_i}, \boldsymbol{\eta}_{p_i}\right) &= \boldsymbol{0} \end{aligned} \tag{B.12}$$

is fulfilled. For a given boundary condition $\{\tilde{\boldsymbol{q}}(t_i), \tilde{\boldsymbol{p}}(t_i)\}$, implying $\tilde{\boldsymbol{q}}(t_i) = \tilde{\boldsymbol{q}}(t_i)^+$ by continuity of $\tilde{\boldsymbol{q}}$, the set of equation is solved for $\boldsymbol{\eta}_{q_i}, \boldsymbol{\eta}_{p_i}$, and the value $\tilde{\boldsymbol{p}}(t_{i+1})$. This allows to generate approximate trajectories of the system in the interval $[t_i, t_{i+1}]$. The values at $\{t_{i+1}\}$, that is, $\tilde{\boldsymbol{q}}(t_{i+1})$ and $\tilde{\boldsymbol{p}}(t_{i+1})$, yield the boundary conditions for the next time interval. By repeating this procedure for all time intervals $I_i, i = 0, 1, \ldots, N-1$, the virtual action vanishes for all variations $\delta\tilde{\boldsymbol{q}}$ and $\delta\tilde{\boldsymbol{p}}$, thereby obtaining trajectories approximating the dynamic equilibrium.

By approximating the integrals using quadrature, the integration scheme can be traced back to the algorithm presented in [2]. Thus, the framework presented herein encompasses the scheme from [2] as a special case by assuming the generalized momenta to be neither left nor right continuous and the generalized coordinates to be continuous.

In case polynomials are used as basis functions and the mass matrix is constant, such that

$$\frac{\partial H}{\partial \boldsymbol{p}}^{\top} = \boldsymbol{M}^{-1}\boldsymbol{p}, \tag{B.13}$$

it can be shown that Equation (B.11) is equivalent to requiring $\boldsymbol{M}\dot{\tilde{\boldsymbol{q}}}(t) = \tilde{\boldsymbol{p}}(t)$ for all $t \in (t_i, t_{i+1})$. Thus, for the pendulum example given in Section 6, the stepping equations (B.12) are equivalent to the so-called Galerkin variational integration procedure, as presented in [1, p. 415] .

ACKNOWLEDGEMENTS

The authors would like to thank Adrian Schweizer and Gabriel Nuetzi for fruitful discussions and helpful comments.

REFERENCES

1. Marsden JE, West M. Discrete mechanics and variational integrators. *Acta Numerica* 2001; **10**:357–514.
2. Leok M, Zhang J. Discrete Hamiltonian variational integrators. *IMA Journal of Numerical Analysis* 2011; **31**(4):1497–1532.
3. Lacoursière C. Ghosts and machines: regularized variational methods for interactive simulations of multibodies with dry frictional contacts. *PhD Thesis*, Umea University, 2007.
4. Fried I. Finite-element analysis of time-dependent phenomena. *AIAA Journal* 1969; **7**(6):1170–1173.
5. Argyris J, Scharpf D. Finite elements in time and space. *Nuclear Engineering and Design* 1969; **10**(4):456–464.
6. Hitzl D, Levinson D. Application of Hamilton's law of varying action to the restricted three-body problem. *Celestial Mechanics* 1980; **22**(3):255–266.
7. Peters D, Izadpanah A. Hp-version finite elements for the space-time domain. *Computational mechanics* 1988; **3**(2):73–88.
8. Aharoni D, Bar-Yoseph P. Mixed finite element formulations in the time domain for solution of dynamic problems. *Computational Mechanics* 1992; **9**(5):359–374.
9. Bottasso C. A new look at finite elements in time: a variational interpretation of Runge–Kutta methods. *Applied Numerical Mathematics* 1997; **25**(4):355–368.
10. Schindler T, Acary V. Timestepping schemes for nonsmooth dynamics based on discontinuous Galerkin methods: definition and outlook. *Mathematics and Computers in Simulation* 2014; **95**:180–199.
11. Borri M, Bottasso C, Mantegazza P. Basic features of the time finite element approach for dynamics. *Meccanica* 1992; **27**(2):119–130.
12. Borri M. Helicopter rotor dynamics by finite element time approximation. *Computers & Mathematics with Applications* 1986; **12**(1):149–160.
13. Groß M, Betsch P. Galerkin-based energy–momentum consistent time-stepping algorithms for classical nonlinear thermo-elastodynamics. *Mathematics and Computers in Simulation* 2011; **82**(4):718–770.
14. Hairer E, Lubich C, Wanner G. *Geometric Numerical Integration*. Springer: Berlin, Heidelberg, 2002.
15. Leyendecker S, Ober-Blöbaum S, Marsden JE, Ortiz M. Discrete mechanics and optimal control for constrained systems. *Optimal Control Applications and Methods* 2010; **31**(6):505–528.

DOI: 10.1002/nme

16. Leyendecker S, Marsden JE, Ortiz M. Variational integrators for constrained dynamical systems. *Journal of Applied Mathematics and Mechanics* 2008; **88**(9):677–708.
17. Ober-Blöbaum S. Discrete mechanics and optimal control. *PhD Thesis*, University of Paderborn, 2008.
18. Lall S, M W. Discrete variational Hamiltonian mechanics. *Journal of Physics A: Mathematical and General* 2006; **39**(19):5509–5519.
19. Channell P, Scovel J. Integrators for Lie–Poisson dynamical systems. *Physica D: Nonlinear Phenomena* 1991; **50**(1):80–88.
20. Ruth R. A canonical integration technique. *IEEE Transactions on Nuclear Science* 1983; **30**(4):2669–2671.
21. Leine RI, Aeberhard U, Glocker C. Hamilton's principle as variational inequality for mechanical systems with impact. *Nonlinear Science* 2009; **19**:633–664.
22. Gelfand I, Fomin S. *Calculus of Variations*. Dover Publications, 1991.
23. Hamel G. *Elementare Mechanik. B.G*. Teubner: Stuttgart, 1912.
24. Papastavridis J. *Analytical mechanics: A Comprehensive Treatise on the Dynamics of Constrained Systems: For Engineers, Physicists, and Mathematicians*. Oxford University Press: Oxford, 2002.
25. Scheck F. *Mechanics: From Newton's Laws to Deterministic Chaos*. Springer: Berlin, Heidelberg, 2010.
26. Lanczos C. *The Variational Principles of Mechanics*. University of Toronto Press: Toronto, 1974.
27. Zienkiewicz O, Taylor R. *The Finite Element Method* (Fifth edn.), Vol. 1. Butterworth Heinemann: Oxford, Woburn, 2000.
28. Lee JM. *Introduction to Smooth Manifolds*. Springer: New York, Heidelberg, Dordrecht, London, 2003.
29. Glocker C. *Set-Valued Force Laws, Dynamics of Non-Smooth Systems Lecture Notes in Applied Mechanics*, Vol. 1. Springer-Verlag: Berlin, 2001.

DOI: 10.1002/nme