Available online at www.sciencedirect.com
ScienceDirect
Comput. Methods Appl. Mech. Engrg. 374 (2021) 113475
Computer methods in applied mechanics and engineering
www.elsevier.com/locate/cma

Multisymplectic Galerkin Lie group variational integrators for geometrically exact beam dynamics based on unit dual quaternion interpolation — no shear locking

Thomas Leitz*, Rodrigo T. Sato Martín de Almagro, Sigrid Leyendecker

Friedrich-Alexander-Universität Erlangen-Nürnberg, Lehrstuhl für Technische Dynamik, Immerwahrstrasse 1, D-91058 Erlangen, Germany

Received 4 November 2018; received in revised form 11 September 2020; accepted 24 September 2020
Available online 8 December 2020

Abstract

We present a Galerkin multisymplectic Lie group variational integrator. It is suitable for dynamical systems defined on a two dimensional space–time and the integrator allows arbitrary convergence orders independently for both dimensions. As an example we use geometrically exact beam dynamics where a slender structure is modelled as a centre line with a cross section at every point. The Lie group in question is the special Euclidean group in three-dimensional space, $SE(3)$, which we parametrize using unit dual quaternions. This allows a very simple and efficient interpolation method to be used, which additionally prevents shear locking present in more naive discretizations of geometrically exact beams.

Keywords: Lie group; Variational integrator; Multisymplectic; Geometrically exact beam; Space–time discretization; High–order integration

1. Introduction

Variational integrators were first introduced by Moser and Veselov [1]. They are based on a discrete version of Hamilton's principle leading to corresponding discrete Euler–Lagrange equations. Symmetries in a system lead to the exact preservation of discrete momentum maps, i.e. there is a discrete Noether's theorem. Variational integrators show very good energy behaviour, i.e. there is no artificial gain or dissipation in a conservative system [2]. A great overview of variational integrators is given by Marsden et al. [3,4]. High-order variational integrators on linear spaces [5–7] and on Lie groups [8,9] have been developed.

Lie groups are smooth manifolds endowed with a group structure which can be exploited to derive globally defined continuous and discrete Euler–Lagrange equations. Furthermore, the Lie group exponential map can be used to solve the nonlinear system of equations representing the integrator directly on the manifold without the need for Lagrange multipliers and constraint functions [10].

Physical systems described by partial differential equations (PDE) in the context of multisymplectic geometry are discussed in [11]. Galerkin multisymplectic variational integrators have been developed [12,13] and applied to geometrically exact beam dynamics [14–16].

* Corresponding author.
E-mail addresses: thomas.leitz@fau.de (T. Leitz), rodrigo.t.sato@fau.de (R.T. Sato Martín de Almagro), sigrid.leyendecker@fau.de (S. Leyendecker).

https://doi.org/10.1016/j.cma.2020.113475
0045-7825/

In this paper we present multisymplectic Lie group variational integrators of arbitrary order based on unit dual quaternion interpolation. A multisymplectic quaternion formulation of the problem was already formulated by [17] on $SO(3)\times\mathbb{R}^3$ and in this paper we reformulate it in $SE(3)$. This choice of space has consequences in the discrete setting when considering Galerkin-based integrators. The choice leads to the use of different spaces of interpolation functions, where the former will treat rotations and translations as independently and the latter will combine both. Unless carefully applied, the first choice can lead to the pathological behaviour known as shear locking. Our approach was inspired by the works of Hall et al. [8] but generalized to a two-dimensional space–time. The use of dual-quaternions to parametrize the position and orientation of the cross sections of geometrically exact beams allows a very simple and efficient interpolation method which additionally circumvents the problem of shear locking found in other Galerkin-based approaches. This interpolation method was already used in [18] in the context of the finite element method.

2. Lagrangian field theories

The geometry involved in Lagrangian field theories is certainly more complex than the one in standard mechanics and an in-depth exploration of it is outside of the scope of this paper (see for instance [11]).

A Lagrangian field theoretical system is a pair $(\pi_Y, \mathcal{L})$, where $\pi_Y : Y \to X$ is a fibre bundle with X, Y, the base and configuration spaces, smooth manifolds of dimension r and $r+m$ respectively, and $\mathcal{L} = L\eta$ is a Lagrangian density, where $\eta \in \Omega^k(X)$ is a volume form on X. L is called the Lagrangian of the system, and if the theory is of *first-order*, then it will be defined on the first jet bundle of π, i.e. $L : J^1\pi \to \mathbb{R}$. We will denote the boundary of X by ∂X.

Assuming adapted bundle coordinates $(x^i, u^a, v_i^a) \in J^1\pi$, $i = 1,\dots,r$, $a = 1,\dots,m$, a Lagrangian is said to be regular if the Hessian matrix

$$\text{blockdiag}\left(\left(\frac{\partial^2 L}{\partial v_k^1 \partial v_l^1}\right)_{k,l=1}^{r}, \left(\frac{\partial^2 L}{\partial v_k^1 \partial v_l^2}\right)_{k,l=1}^{r}, \dots, \left(\frac{\partial^2 L}{\partial v_k^{m-1} \partial v_l^m}\right)_{k,l=1}^{r}, \left(\frac{\partial^2 L}{\partial v_k^m \partial v_l^m}\right)_{k,l=1}^{r}\right),$$

is regular everywhere. We will assume that our Lagrangian is regular.

In this setting, instead of working with curves as in standard mechanics, which can be thought of as sections of a line bundle, we will work with sections of π_Y which we call fields, $u : X \to Y$, i.e. $\pi_Y \circ u = \mathrm{Id}_X$. The prolongation of our fields into the first jet bundle, $j^1u : X \to J^1\pi$ will be of the form $(x^i, u^a(x), u^a_{,i}(x))$, where $u^a_{,i} = \frac{\partial u^a}{\partial x^i}$.

We define the mechanical action associated to our fields as

$$\mathcal{S}[u] = \int_X (j^1u)^*\mathcal{L} = \int_X L(j^1u)\eta$$

We can assume an arbitrary family of smooth sections $u_\epsilon(x)$ with fixed values at ∂X, parametrized by $\epsilon \in \mathbb{R}$, and $u_0(x) = u(x)$. Hamilton's principle states that $u(x)$ is a physical trajectory of the system iff such a section is a stationary point of the action, i.e.

$$\left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \mathcal{S}[u_\epsilon] = 0$$

This is equivalent to the field $u(x)$ being a solution of the field Euler–Lagrange equations

$$\frac{\partial L}{\partial u^a} - \frac{\partial}{\partial x^i}\left(\frac{\partial L}{\partial u^a_{,i}}\right) = 0, \quad \text{for } a = 1,\dots,m$$

Assume now that Y is embedded in another bundle $\pi_Z : Z \to X$, with inclusion $\iota : Y \hookrightarrow Z$ such that it is a bundle homomorphism over X, i.e. such that $\pi_Z \circ \iota = \pi_Y$. Y may then be defined by the null-set of a *constraint function* $c : Z \to \mathbb{R}^{n-m}$, i.e. $Y = \{q_x \in Z \mid c(q_x) = 0\}$, and clearly $c \circ \iota = 0$.

Given the parametrized family of sections u_ϵ we can construct another family $\iota(u_\epsilon) = q_\epsilon : X \to Z$, with $q_0(x) = q(x) = \iota(u(x))$. This allows us to define variations of $q(x)$ in terms of $\delta u(x) = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} u_\epsilon(x)$, noting that δu is a vertical vector field on Y,

$$\delta q(x) = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} q_\epsilon(x) = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \iota(u_\epsilon(x)) = T_{u(x)}\iota(\delta u(x)).$$

In coordinates, this is nothing but $\delta q^b(x) = \frac{\partial \iota^b}{\partial u^a}(x,u)\delta u^a(x)$.

Consider a Lagrangian system $\left(\pi_{\mathcal{N}}, \widetilde{\mathcal{L}}\right)$, possibly singular, which we will call *ambient system*, with $\widetilde{\mathcal{L}} = \widetilde{L}\eta$ such that $L = \widetilde{L} \circ j^1\iota$. Then, we will be able to cast the original problem as a constrained one in π_Z.

In that case, it can be shown that the resulting constrained Euler–Lagrange equations can be written as

$$\left\{ \begin{array}{r} \left(T^*\iota\right)_q \left[\dfrac{\partial \widetilde{L}}{\partial q} - \dfrac{\partial}{\partial x^i}\left(\dfrac{\partial \widetilde{L}}{\partial q_{,i}}\right)\right] = 0 \\ c\left(q\right) = 0 \end{array} \right.$$

which must be solved together with appropriate boundary conditions imposed at ∂X.

Here $(T^*\iota) : T^*Z|_Y \to T^*Y$ can be understood as a matrix of dimensions $(r+m)\times(r+n)$. This is acting on a 1-form $D_{EL}\widetilde{L} : J^2\pi_Z \to T^*Z$, $\left[\frac{\partial \widetilde{L}}{\partial q^b} - \frac{\partial}{\partial x^i}\left(\frac{\partial \widetilde{L}}{\partial q^b_i}\right)\right] \mathrm{d}q^b$, so the only part that interests us is the $m \times n$ term. Therefore, when complemented with the $n-m$ constraint equations, the entire system is of dimension n.

We will actually consider a slightly simpler setting, where our bundles Y, Z are trivial bundles of the form $X \times M$ and $X \times N$, respectively. π_Y, π_Z reduce to the canonical projections onto their respective first arguments, pr_1. If $M \subset N$, we may have an inclusion map $\hat{\iota} : M \hookrightarrow N$ and a constraint function $\hat{c} : N \to \mathbb{R}^{n-m}$. The first induces a bundle homomorphism over X, $\iota : Y \hookrightarrow Z$, while the second induces a function $c : Z \to \mathbb{R}^{n-m}$. This time it makes sense to talk about *base independence*, which ι and c are, as should be the constraint function $\hat{c} : N \to \mathbb{R}^{n-m}$. In this case the term $(T^*\iota)_q$ coincides with $T^*\hat{\iota}_u$ for $q = \iota(u)$ and can be written as $\left(\frac{\partial \hat{\iota}^b}{\partial u^a}(u)\right)^T$.

2.1. Field theories on Lie groups

Since our goal is to simulate the dynamics of the geometrically exact beam, we are interested in the particular case where M is a Lie group G.

Recall that a Lie group $(G, \cdot)$ is a smooth manifold endowed with a group structure where '$\cdot$' denotes the group operation. We denote an element of G as g, its inverse as g^{-1} and the identity element as e. Given $h \in G$ we can define:

$$\begin{aligned} L_h &: G \to G, \qquad g \mapsto h \cdot g, \\ R_h &: G \to G, \qquad g \mapsto g \cdot h, \end{aligned}$$

the *left* and *right translation*, respectively. The set of vector fields on G that are invariant under left translation, i.e. $X \in \mathfrak{X}(G)$ such that $T_g L_h X(g) = X(h \cdot g)$, are called *left-invariant* vector fields. The set of all left-invariant vector fields on G, which we will denote as $\mathfrak{g}$, is a finite dimensional vector space which is diffeomorphic to the tangent space at the identity, T_eG, and is closed under the usual commutator of vector fields.

A Lie algebra is a vector space V closed under a given skew-symmetric bilinear operation, the Lie bracket $[\cdot,\cdot] : V \times V \to V$, satisfying the Jacobi identity, i.e. that for all $\xi, \eta, \zeta \in V$, $[\xi,[\eta,\zeta]]+[\eta,[\zeta,\xi]]+[\zeta,[\xi,\eta]] = 0$. Thus, $\mathfrak{g}$ is a Lie algebra whose bracket is the usual commutator. The exponential map, $\exp : \mathfrak{g} \to G$, allows us to move from the algebra to the group, and it is locally invertible in a neighbourhood of the identity, U_e.

Given a section $g(x)$ of $Y = X \times G$, we may construct a family $g_\epsilon(x)$ parametrized by ϵ such that $g_0(x) = g(x)$ from another family of sections $\xi_\epsilon(x)$ of $X \times \mathfrak{g}$ with $\xi_0(x) = 0$ simply by exponentiation as $g_\epsilon(x) = L_{g(x)} \exp(\xi_\epsilon(x))$. Its variations are:

$$\delta g(x) = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} g_\epsilon(x) = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} L_{g(x)} \exp(\xi_\epsilon(x)) = T_e L_{g(x)} \left(\delta\xi(x)\right),$$

where we have used the fact that $T_0 \exp = \mathrm{Id}$, and $\left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \xi_\epsilon(x) = \delta\xi(x)$.

If $G \subset N$, and $q_\epsilon(x) = \iota(L_{g(x)} \exp(\xi_\epsilon(t)))$, the final form of our Euler–Lagrange equations are

$$\left.\begin{array}{r} P^T(q)\left[\dfrac{\partial \widetilde{L}}{\partial q} - \dfrac{\partial}{\partial x^i}\left(\dfrac{\partial \widetilde{L}}{\partial q_{,i}}\right)\right] = 0 \\ c\left(q\right) = 0 \end{array}\right\}$$

where now $P(q)$ is the nullspace matrix formed by $T_{L_g e}\iota T_e L_g$ defined on $X \times N|_{X\times G}$. Whenever $X \times N \cong \mathbb{R}^{r+n}$ these equations will be valid on a globally defined coordinate frame.

One could choose to work intrinsically on the Lie group and further exploit the structure of the group by applying trivialization. In the mechanical setting, this is the process of representing points in TG by elements in the product space $G \times \mathfrak{g}$, where $\mathfrak{g} \cong T_eG$ denotes the Lie algebra associated to the Lie group. It can be accomplished by applying left (or right) translation to the fibres at a point g, leading us to represent (g, v) as $(g, g^{-1} \cdot v = \omega)$. In that setting it leads to trivialized Lagrangians $\ell : G \times \mathfrak{g} \to \mathbb{R}$ and trivialized equations of motion, and the same can be applied in the field theoretical setting. This is not the main focus of this paper and describing how to obtain said trivialized equations lies outside of its scope (see for instance [10,19,20]).

3. Unit dual quaternions and interpolation in $SE(3)$

For the particular model we intend to study, the geometrically exact beam, which will be introduced in Section 4, we will be focusing on the particular Lie group $SE(3)$. It is well-known that unit quaternions provide a double covering of $SO(3)$. Something similar can be said about unit dual quaternions and $SE(3)$, and it is this that we will apply.

Due to the vector space structure of the space of quaternions, quaternion interpolation can be performed with ease. Furthermore, one can obtain an interpolation on the space of unit quaternions via normalization. Exactly the same applies on the space of dual quaternions and unit dual quaternion interpolation, which we use in Section 3.4. The identification with $SE(3)$ then allows us to perform high-order interpolation on this space without resourcing to more involved approaches.

First, let us introduce quaternions, dual numbers and finally dual quaternions.

3.1. Quaternions and unit quaternions

Quaternions can be thought of as a four-dimensional generalization of the complex numbers defined as

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

where multiplication, represented by juxtaposition, is non-commutative.

From its definition we can obtain the multiplication rule for any pair of quaternions, $q, p \in \mathbb{H}$, which we will denote by simple juxtaposition, $qp \in \mathbb{H}$. A left and right translation on quaternions can then be defined as $L_q : p \mapsto qp$ and $R_p : q \mapsto qp$. In analogy with complex numbers, if $q = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k} = q_0 + \mathbf{q}$ we define its conjugate to be $\bar{q} = q_0 - \mathbf{q}$, so that $\overline{qp} = \bar{p}\bar{q}$, and its norm as $\|q\| = \sqrt{\bar{q}q}$. The inverse of $q \neq 0$ is $q^{-1} = \bar{q} / \|q\|^2$. Exponential and logarithm operations can be defined in terms of power series leading to expressions

$$\exp(q) = e^{q_0}\left(\cos(\|\mathbf{q}\|) + \sin(\|\mathbf{q}\|)\frac{\mathbf{q}}{\|\mathbf{q}\|}\right),$$

$$\ln(q) = \ln(\|q\|) + \cos^{-1}\left(\frac{q_0}{\|q\|}\right)\frac{\mathbf{q}}{\|\mathbf{q}\|}.$$

The set of unit quaternions, defined as $\mathbb{H}_1 := \{q \in \mathbb{H} \,|\, \|q\| = 1\} \cong \mathbb{R}^3$, is diffeomorphic to the 3-sphere, S^3. It forms a Lie group under quaternion multiplication with unit $e = 1$ (the real unit) and the inverse reduces to conjugation, being homomorphic to the Lie group $SU(2)$. The Euler–Rodrigues formula provides us with a surjective map $\chi : \mathbb{H}_1 \to SO(3)$, $q \mapsto \chi(q) = I_3 + 2q_0\widehat{\mathbf{q}} + 2\widehat{\mathbf{q}}^2$, where $\widehat{\cdot} : \mathbb{R}^3 \to \mathfrak{so}(3)$ is the standard map of vectors to skew-symmetric 3×3 matrices.

The Lie algebra of $\mathbb{H}_1$ is the set of purely imaginary quaternions, i.e. $\mathfrak{h}_1 := \{x \in \mathbb{H} \,|\, \bar{x} + x = 0\}$. The restriction $\exp|_{\mathfrak{h}_1} : \mathfrak{h}_1 \to \mathbb{H}$ can be thought of as the composition of the group exponential $\exp_{\mathbb{H}_1} : \mathfrak{h}_1 \to \mathbb{H}_1$ with an inclusion map $\iota : \mathbb{H}_1 \hookrightarrow \mathbb{H}$, which remarkably is such that

$$\iota(g \cdot h) = \iota(g)\iota(h), \tag{1}$$

where $\cdot$ represents the group multiplication in $\mathbb{H}_1$, and $\iota(e) = 1$. This leads to the popular parametrization

$$\exp(x) = \cos(\|\mathbf{x}\|) + \sin(\|\mathbf{x}\|)\frac{\mathbf{x}}{\|\mathbf{x}\|} = \cos\left(\frac{\theta}{2}\right) + \sin\left(\frac{\theta}{2}\right)\mathbf{u},$$

where $\mathbf{u}$ is a unit vector in $\mathbb{R}^3$ representing the axis of rotation and θ is the actual angle of rotation.

Taking the differential of the Euler–Rodrigues map at $q = 1$ we get the algebra isomorphism $T_1\chi : \mathfrak{h}_1 \to \mathfrak{so}(3)$, $T_1\chi x = 2\hat{\mathbf{x}}$.

Given a curve $q(t) \in \iota(\mathbb{H}_1)$ and a family of curves $x_\varepsilon(t) \in \mathfrak{h}_1$ such that $x_0(t) = 0$ and $\delta x(t) = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\big|_{\varepsilon=0} x_\varepsilon(t)$, we can define the family $q_\varepsilon(t) = q(t)\exp(x_\varepsilon(t))$. This, together with Eq. (1), leads to variations of the form

$$\delta q(t) = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0} q_\varepsilon(t) = T_1 L_{q(t)} \left(T_e\iota \left(T_0 \exp_{\mathbb{H}_1} \delta x(t)\right)\right) = T_1 L_{q(t)} \left(T_e\iota\, \delta x(t)\right)$$

Therefore, the null space matrix is:

$$P(q) = T_1 L_q T_e\iota = \begin{pmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & -q_3 & q_2 \\ q_2 & q_3 & q_0 & -q_1 \\ q_3 & -q_2 & q_1 & q_0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -q_1 & -q_2 & -q_3 \\ q_0 & -q_3 & q_2 \\ q_3 & q_0 & -q_1 \\ -q_2 & q_1 & q_0 \end{pmatrix}$$

If instead of $x_\varepsilon(t) \in \mathfrak{h}_1$ we work with $\xi_\varepsilon(t) = (T_1\chi)^{-1}(x_\varepsilon(t)) \in \mathfrak{so}(3)$, we obtain the same null space matrix but multiplied by a factor $1/2$.

3.2. Dual numbers

Dual numbers are a generalization of real numbers [21,22]. They are defined as

$$\mathbb{D} := \left\{y = y_r + y_\varepsilon\varepsilon \,|\, y_r, y_\varepsilon \in \mathbb{R}, \varepsilon^2 = 0\right\}$$

Algebraically dual numbers form a 2-dimensional unital commutative associative algebra over $\mathbb{R}$. If $y, z \in \mathbb{D}$, its product is $yz = zy = y_r z_r + (y_r z_\varepsilon + y_\varepsilon z_r)\varepsilon$. Any element $y \in \mathbb{D}$ whose real part is non-zero, i.e. $y_r \neq 0$, has an inverse $y^{-1} = \frac{1}{y_r} - \frac{y_\varepsilon}{y_r^2}\varepsilon$. Pure dual numbers have no inverse, which makes $\mathbb{D}$ a commutative ring but not a field. Given any real analytic function $f(x)$, we may extend it into the dual numbers, and it can be shown that

$$f(y) = f(y_r + y_\varepsilon\varepsilon) = f(y_r) + f'(y_r)y_\varepsilon\varepsilon$$

This fact has led to their recent application in numerics, in the field of automatic differentiation.

3.3. Dual quaternions and unit dual quaternions

Dual quaternions [22–24] can be defined as

$$\widetilde{\mathbb{H}} := \left\{\tilde{q} = \tilde{q}_0 + \tilde{q}_1\mathbf{i} + \tilde{q}_2\mathbf{j} + \tilde{q}_3\mathbf{k} \,|\, \tilde{q}_i \in \mathbb{D}, \mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1\right\} = \{\tilde{q} = q_r + q_\varepsilon\varepsilon \,|\, q_r, q_\varepsilon \in \mathbb{H}\}$$

Conjugation is naturally extended for any dual quaternion $\tilde{q} = q_r + q_\varepsilon\varepsilon$ as $\bar{\tilde{q}} = \bar{q}_r + \bar{q}_\varepsilon\varepsilon$. So is the norm, $\|\tilde{q}\| = \sqrt{\bar{\tilde{q}}\tilde{q}} = \|q_r\| + \frac{\langle q_r, q_\varepsilon\rangle}{\|q_r\|}\varepsilon$, where $\langle\cdot,\cdot\rangle$ is the inner product in $\mathbb{R}^4$. Unit dual quaternions are defined as $\widetilde{\mathbb{H}}_1 = \left\{\tilde{q} \in \widetilde{\mathbb{H}} \,|\, \|\tilde{q}\| = 1\right\}$, which implies that if $\tilde{q} \in \widetilde{\mathbb{H}}_1$, $q_r \in \mathbb{H}_1$ and $\langle q_r, q_\varepsilon\rangle = 0$, i.e. $\widetilde{\mathbb{H}}_1 \cong S^3 \times \mathbb{R}^3$. Every unit dual quaternion has an inverse, making $\widetilde{\mathbb{H}}_1$ a Lie group, and it follows that, as in the unit quaternion case, the inverse of a unit dual quaternion is its conjugate.

A unit dual quaternion can be parametrized by a pair $(q, x) \in \mathbb{H}_1 \times \mathfrak{h}_1$ representing a rotation and a subsequent translation respectively,

$$\tilde{q} = q_r(q) + q_\varepsilon(q, x)\varepsilon = q + \frac{1}{2}xq\varepsilon.$$

The Lie algebra of $\widetilde{\mathbb{H}}_1$ is the set of purely imaginary dual quaternions, i.e. $\tilde{\mathfrak{h}}_1 := \left\{\tilde{x} \in \widetilde{\mathbb{H}} \,|\, \bar{\tilde{x}} + \tilde{x} = 0\right\} \cong \mathbb{R}^3 \times \mathbb{R}^3$. The exponential map $\exp|_{\tilde{\mathfrak{h}}_1} : \tilde{\mathfrak{h}}_1 \to \tilde{\mathbb{H}}$ can be computed from the one obtained for quaternions by extending it into $\mathbb{D}$. The real part is exactly $\exp(x_r)$, while the dual part then takes the form

$$\frac{1}{\|\mathbf{x_r}\|^2}\left[(2s\mathbf{x}_\varepsilon\cos(s\,\|\mathbf{x}_r\|)\sin(s\,\|\mathbf{x}_r\|) + s\mathbf{x}_r d)s\,\|\mathbf{x}_r\| - \cos(s\,\|\mathbf{x}_r\|)\sin(s\,\|\mathbf{x}_r\|)s\mathbf{x}_r d + 2s\mathbf{x}_r \times \mathbf{x}_\varepsilon \sin(s\,\|\mathbf{x}_r\|)^2\right],$$

where $d = 2\langle\mathbf{x_r}, \mathbf{x_\varepsilon}\rangle/\|\mathbf{x_r}\|$.

After some computations, one obtains the following null space matrix

$$P(\tilde{q}) = \begin{bmatrix} P(q_r) & 0^{4\times 3} \\ P(q_\varepsilon) & P(q_r) \end{bmatrix}$$

3.4. Unit dual quaternion interpolation

Dual quaternion interpolation can be performed exactly as standard quaternion interpolation. Assume we are interested in interpolating sections of the trivial bundle $\mathrm{pr}_1 : X \times \widetilde{\mathbb{H}} \to X$, where $X = [0, 1]^2$ is the unit square with coordinates (τ, σ). We can apply a tensor product interpolation using a set of nodes $\left\{\left(\tau^l, \sigma_k, \tilde{q}_k^l\right) \in X \times \widetilde{\mathbb{H}}\right\}_{k=0,\dots,K}^{l=0,\dots,L}$, for $K, L \in \mathbb{N}$ with

$$0 = \tau^0 < \tau^1 < \cdots < \tau^{L-1} < \tau^L = 1\,,$$
$$0 = \sigma_0 < \sigma_1 < \cdots < \sigma_{K-1} < \sigma_K = 1\,.$$

This way our interpolant in $\widetilde{\mathbb{H}}$ can be written as

$$\tilde{q}\left(\left\{\tilde{q}_k^l\right\}_{k=0,\dots,K}^{l=0,\dots,L}, \tau, \sigma\right) = \sum_{l=0}^{L}\sum_{k=0}^{K} w_k^l(\tau,\sigma)\,\tilde{q}_k^l$$

If $\tilde{q}_k^l \in \widetilde{\mathbb{H}}_1$, unit dual quaternion interpolation can be obtained by normalization

$$\tilde{q}\left(\left\{\tilde{q}_k^l\right\}_{k=0,\dots,K}^{l=0,\dots,L}, \tau, \sigma\right) = \frac{\sum_{l=0}^{L}\sum_{k=0}^{K} w_k^l(\tau,\sigma)\,\tilde{q}_k^l}{\left\|\sum_{l=0}^{L}\sum_{k=0}^{K} w_k^l(\tau,\sigma)\,\tilde{q}_k^l\right\|}. \tag{2}$$

The weights are Lagrange polynomials of the form

$$w_k^l(\tau,\sigma) = \left(\prod_{\substack{n=0\\ n\neq l}}^{L} \frac{\tau-\tau^n}{\tau^l-\tau^n}\right)\left(\prod_{\substack{m=0\\ m\neq k}}^{K} \frac{\sigma-\sigma_m}{\sigma_k-\sigma_m}\right).$$

4. Geometrically exact beam dynamics

Geometrically exact beams, originally developed by Simo [25], model the behaviour of slender structures taking into consideration longitudinal, shear, flexural and torsional deformations. They are described by the position of a line of centroids in three-dimensional space, $u \in \mathbb{R}^3$, and the orientation of rigid cross sections along the line, $\Lambda \in SO(3)$. It is assumed that the local frame of each section is such that its first basis vector is perpendicular to the section while its second and third lie on the section along its principal axes (when applicable). Although its natural configuration space is $SO(3) \times \mathbb{R}^3$, as we will discuss in Section 4.2, it will be advantageous to substitute this space by the group $SE(3)$.

Our configuration bundle will then be the trivial bundle $\mathrm{pr}_1 : X \times SE(3) \to X$, where $X \equiv \mathbb{R} \times [0, \ell]$. Given a section of this bundle $g = (\Lambda, u)$, the following magnitudes can be defined:

$$(\omega, v) = \dot{g}\cdot g^{-1} = (\dot{\Lambda}\Lambda^{-1}, \dot{u} + \dot{\Lambda}\Lambda^{-1}u) \qquad \text{Spatial angular and linear velocities}$$
$$(\kappa, w) = g'\cdot g^{-1} = (\Lambda'\Lambda^{-1}, u' + \Lambda'\Lambda^{-1}u) \qquad \text{Spatial angular and linear strains}$$
$$(\Omega, V) = g^{-1}\cdot \dot{g} = (\Lambda^{-1}\dot{\Lambda}, \Lambda^{-1}\dot{u}) \qquad \text{Body angular and linear velocities}$$
$$(K, W) = g^{-1}\cdot g' = (\Lambda^{-1}\Lambda', \Lambda^{-1}\cdot u') \qquad \text{Body angular and linear strains}$$

We assume a hyperelastic material behaviour, implying the existence of a bilinear quadratic elastic energy function of the angular and linear strains of the form

$$\Psi(K, W) = \frac{1}{2}\left[\langle W - e_1, \mathbb{C}_l(W - e_1)\rangle + \langle K, \mathbb{C}_a K\rangle\right].$$

Here $\mathbb{C}_l = \mathrm{diag}([EA_1, GA_2, GA_3])$ and $\mathbb{C}_a = \mathrm{diag}([GI_1, EI_2, EI_3])$ represent the linear and angular stiffness in body frame. E and G are the elastic (Young) and shear modulus respectively, and A_i and I_i are effective areas and area moments of inertia respectively. e_1 is the perpendicular basis vector as seen in the body frame.

The Lagrangian of the model [14,16,17] can be defined on $J^1\mathrm{pr}_1$, which itself can be left-trivialized into the bundle $\check{\mathrm{pr}}_1 : X \times SE(3) \times \mathfrak{se}(3) \to X$. Using coordinates $(t, s, \Lambda, u, \Omega, V, K, W)$, we can write the trivialized Lagrangian $\ell : J^1\check{\mathrm{pr}}_1 \to \mathbb{R}$ as:

$$\ell(t, s, \Lambda, u, \Omega, V, K, W) = \frac{1}{2}\left[\langle V, \rho V\rangle + \langle \Omega, \mathbb{J}\Omega\rangle - \langle W - e_1, \mathbb{C}_l(W - e_1)\rangle - \langle K, \mathbb{C}_a K\rangle\right] - U(\Lambda, u) \tag{3}$$

where $\rho > 0$ is the linear density of the beam, $U : SE(3) \to \mathbb{R}$ is a potential function and $\mathbb{J} = \mathrm{diag}([J_1, J_2, J_3])$ is the inertia matrix of the sections in the body frame. We will assume that $\mathbb{J}$, $\mathbb{C}_a$ and $\mathbb{C}_l$ are non-singular. Any of these may also be considered functions of X, but for the sake of simplicity we will assume they are constant, rendering our Lagrangian autonomous.

Using the notation $N := \mathbb{C}_l(W - e_1)$ and $M := \mathbb{C}_a K$ for the body linear and angular stresses, the trivialized Euler–Lagrange equations of the model can be shown to be

$$\begin{cases} \mathbb{J}\dot{\Omega} - (\mathbb{J}\Omega)\times\Omega &= M' - M\times K - N\times W + \mathrm{T}, \\ \rho(\dot{V} + \Omega\times V) &= N' - N\times K + F. \end{cases}$$

where F and T are terms due to the potential.

In our case, the boundary conditions will be of the form

$$\begin{aligned} g(0, s) &= g_0(s), & \dot{g}(0, s) &= v_0(s), & &\text{for } s \in [0, \ell], \\ & & g'(t, 0) &= w_{\text{left}}(t), & &\text{for } t \in [0, T], \\ & & g'(t, \ell) &= w_{\text{right}}(t), & &\text{for } t \in [0, T], \end{aligned} \tag{4}$$

with $(g_0(s), v_0(s))$, $(g(t, 0), w_{\text{left}}(t))$ and $(g(t, \ell), w_{\text{right}}(t)) \in TG$, which translate to conditions,

$$\begin{aligned} &(\Lambda(0, s), u(0, s)) = (\Lambda_0(s), u_0(s)), && (\Omega(0, s), V(0, s)) = (\Omega_0(s), V_0(s)), && \text{for } s \in [0, \ell], \\ & && (K(t, 0), W(t, 0)) = (K_{\text{left}}(t), W_{\text{left}}(t)), && \text{for } t \in [0, T], \\ & && (K(t, \ell), W(t, \ell)) = (K_{\text{right}}(t), W_{\text{right}}(t)), && \text{for } t \in [0, T]. \end{aligned}$$

4.1. Geometrically exact beams in terms of dual quaternions

Given a smooth curve $\tilde{q}(t, s) \in \widetilde{\mathbb{H}}_1$, the left-trivialized velocity and strains are defined as

$$\widetilde{\Omega} := 2\bar{\tilde{q}}\dot{\tilde{q}} = \Omega + \varepsilon V, \quad \widetilde{K} := 2\bar{\tilde{q}}\tilde{q}' = K + \varepsilon W,$$

where $\Omega = 2\bar{q}\dot{q}$ and $V = \bar{q}\dot{x}q$, $K = 2\bar{q}q'$ and $W = \bar{q}x'q$.

Given the bundle $\tilde{\pi} : \widetilde{\mathbb{H}} \to X \equiv \mathbb{R}\times[0, \ell]$, we may define the following ambient Lagrangian $\tilde{L} : J^1\tilde{\pi} \to \mathbb{R}$ (cf. [17]):

$$\tilde{L}\left(t, s, \tilde{q}, \dot{\tilde{q}}, \tilde{q}'\right) = 2\left[\widetilde{M}\left(\bar{\tilde{q}}\dot{\tilde{q}}, \bar{\tilde{q}}\dot{\tilde{q}}\right) - \widetilde{C}\left(\bar{\tilde{q}}\tilde{q}' - \varepsilon\mathbf{i}, \bar{\tilde{q}}\tilde{q}' - \varepsilon\mathbf{i}\right)\right] - \widetilde{U}(\tilde{q})$$

Here $\widetilde{M}, \widetilde{C} : \widetilde{\mathbb{H}}\times\widetilde{\mathbb{H}} \to \mathbb{R}$, $\widetilde{M}(\tilde{q}, \tilde{p}) = \langle q_r, \widetilde{\mathbb{J}}p_r\rangle + \langle q_\epsilon, \tilde{\rho}p_\epsilon\rangle$, $\widetilde{C}(\tilde{q}, \tilde{p}) = \langle q_r, \widetilde{\mathbb{C}}_a p_r\rangle + \langle q_\epsilon, \widetilde{\mathbb{C}}_l p_\epsilon\rangle$ are symmetric bilinear forms with associated matrices $\widetilde{\mathbb{J}} = \mathrm{diag}([\alpha_1, \mathbb{J}])$, $\tilde{\rho} = \mathrm{diag}([\alpha_2, \rho I_3])$, $\widetilde{\mathbb{C}}_a = \mathrm{diag}([\alpha_3, \mathbb{C}_a])$ and $\widetilde{\mathbb{C}}_l = \mathrm{diag}([\alpha_4, \mathbb{C}_l])$, and $\alpha_i \in \mathbb{R}$. This Lagrangian is regular if $\alpha_i \neq 0$ for all $i = 1, \ldots, 4$ and it can be seen that if $\tilde{\imath} : \widetilde{\mathbb{H}}_1 \hookrightarrow \widetilde{\mathbb{H}}$, $L = \tilde{L}\circ T\tilde{\imath}$ is regular and independent of α_i.

4.2. Shear locking and unit dual quaternion interpolation

As suggested in [18], the type of interpolation presented in Section 3.4 prevents the numerical phenomenon known as "shear locking" that occurs in standard Galerkin-type interpolations. To understand why this is so, first we need to understand what causes shear locking in the first place.

This is a phenomenon frequently occurring in beam simulations where separate bending and shear energies are considered. Shear locking leads to incorrect strains and overall deformation due to an apparent increase in the shear stiffness of the beam. The cause is easy to identify, especially if we consider a static 2D version of our problem

without gravity: assume $\Lambda \in SO(2)$ and $u \in \mathbb{R}^2$, so that our problem can be described by a curve in $S^1 \times \mathbb{R}^2$. The Lagrangian reduces to a deformation energy of the form:

$$L(\theta, u_1, u_2, W_\theta, W_1, W_2) = \frac{1}{2}\left[EA_1(W_1\cos\theta + W_2\sin\theta - 1)^2 + GA_2(W_1\cos\theta - W_2\sin\theta)^2 + EI_3W_\theta^2\right]$$

If we use linear Galerkin interpolation, this is the same as saying that, given an interval of length Δs, we will have initial and final values for (θ, u_1, u_2), namely $\theta^1, u_1^1, u_2^1, \theta^2, u_1^2, u_2^2$, which irremediably lead to $W_\theta^1 = W_\theta^2 = \frac{\theta^2-\theta^1}{\Delta s}$, $W_1^1 = W_1^2 = \frac{u_1^2-u_1^1}{\Delta s}$, $W_2^1 = W_2^2 = \frac{u_2^2-u_2^1}{\Delta s}$, being constant in the interval. Thus, functional evaluations of any quadrature rule results in these values. If we have an arbitrary linear change in the angle of the sections, this will generally lead to a non-vanishing shear term. This is because what the shear term measures is the tangential discrepancy between the normal to the section and the tangent vector to the centroid line, which in this case is just a straight line. The choice of linear interpolation in space leads to tangent vectors which are entirely different from the analytical solution.

Contrary to that, performing linear interpolation in $\widetilde{\mathbb{H}}$ already leads to quadratic interpolation in the translational variables, and by normalizing, the interpolatory curve stops being a polynomial altogether. This eliminates the issue, as the tangent to the curve can now better match the analytical solution.

It should be noted that this analysis does not apply to other types of discretization, such as collocation, where the effect may not be present.

5. Discrete field theories

Variational integrators are based on a discrete version of Hamilton's principle applied to an approximation of the action. They are characterized by the conservation of the symplectic and multisymplectic form associated to the problem which translate into good long-term energy and momentum behaviour (see [14,16,17] for more information on this particular model).

We propose the following space–time discretization. The time interval $t = \left[t^0, t^J\right] \subset \mathbb{R}$ is discretized into $J+1$ main time nodes t^j, where $j \in \{0, \dots, J\}$ and the space interval $s = [s_0, s_A] \subset \mathbb{R}$ is discretized into $A+1$ main space nodes s_a, where $a \in \{0, \dots, A\}$. This is called the primary space–time grid. The time intervals $\left[t^j, t^{j+1}\right]$ are each further discretized with $L+1$ time nodes $t^{j,l}$, where $l \in \{0, \dots, L\}$ and $t^j = t^{j,0} = t^{j-1,L}$ for every $j \in \{1, \dots, J\}$. The space intervals $[s_a, s_{a+1}]$ are each further discretized with $K+1$ space nodes $s_{a,k}$, where $k \in \{0, \dots, K\}$ and $s_a = s_{a,0} = s_{a-1,A}$ for every $a \in \{1, \dots, A\}$. We call this the secondary space–time grid and the space–time nodes between the main space–time nodes we call 'control points'. Consequently $\tilde{q}_a^j = \tilde{q}\left(t^j, s_a\right)$ are the main variables and $\tilde{q}_{a,k}^{j,l} = \tilde{q}\left(t^{j,l}, s_{a,k}\right)$ with $(l,k) \in [0, \dots, L] \times [0, \dots, K] \setminus \{(0,0), (L,0), (0,K), (L,K)\}$ are the secondary variables. The space–time-grid is visualized in Fig. 1.

With that in mind, the discrete Lagrangian $(\tilde{L}_d)_a^j : \widetilde{\mathbb{H}}^{(L+1)(K+1)} \to \mathbb{R}$ is an approximation of the continuous action on the space–time element $\left[t^j, t^{j+1}\right] \times [s_a, s_{a+1}]$ and reads

$$(\tilde{L}_d)_a^j \approx \int_{t^j}^{t^{j+1}} \int_{s_a}^{s_{a+1}} \tilde{L}\left(t, s, \tilde{q}, \dot{\tilde{q}}, \tilde{q}'\right) \mathrm{d}s\, \mathrm{d}t$$

The approximation is done in two steps. Firstly, the curve $\tilde{q}(t,s)$ is approximated on the space–time element $\left[t^j, t^{j+1}\right] \times [s_a, s_{a+1}]$ using some interpolation method

$$\hat{\tilde{q}}_a^j\left(\left\{\tilde{q}_{a,k}^{j,l}\right\}_{k=0\dots K}^{l=0\dots L}, \tau, \sigma\right) \approx \tilde{q}(t,s) \quad \text{for} \quad (t,s) \in \left[t^j, t^{j+1}\right] \times [s_a, s_{a+1}]$$

where $\tau(t) = \frac{t-t^j}{\Delta t^j}$ with $\Delta t^j = t^{j+1} - t^j$ is the pseudo time and $\tau \in [0,1]$ and $\sigma(s) = \frac{s-s_a}{\Delta s_a}$ with $\Delta s_a = s_{a+1} - s_a$ is the pseudo space and $\sigma \in [0,1]$ and $\left\{g_{a,k}^{j,l}\right\}_{k=0\dots K}^{l=0\dots L}$ are the points of the interpolation. Secondly, a quadrature rule is applied to approximate the integral and the discrete Lagrangian is defined as

$$\begin{aligned}(\tilde{L}_d)_a^j\left(\left\{\tilde{q}_{a,k}^{j,l}\right\}_{k=0\dots K}^{l=0\dots L}\right) := \Delta t^j \Delta s_a \sum_{i_l=0}^{I_l}\sum_{i_k=0}^{I_k} b^{i_l} b_{i_k} (\tilde{L}_d)\left(\hat{\tilde{q}}_a^j\left(\left\{\tilde{q}_{a,k}^{j,l}\right\}_{k=0\dots K}^{l=0\dots L}, \tau^{i_l}, \sigma_{i_k}\right),\right.\\ \left.\dot{\hat{\tilde{q}}}_a^j\left(\left\{\tilde{q}_{a,k}^{j,l}\right\}_{k=0\dots K}^{l=0\dots L}, \tau^{i_l}, \sigma_{i_k}\right), \left(\hat{\tilde{q}}_a^j\right)'\left(\left\{\tilde{q}_{a,k}^{j,l}\right\}_{k=0\dots K}^{l=0\dots L}, \tau^{i_l}, \sigma_{i_k}\right),\ t\left(\tau^{i_l}\right), s\left(\sigma_{i_k}\right)\right)\end{aligned} \tag{5}$$

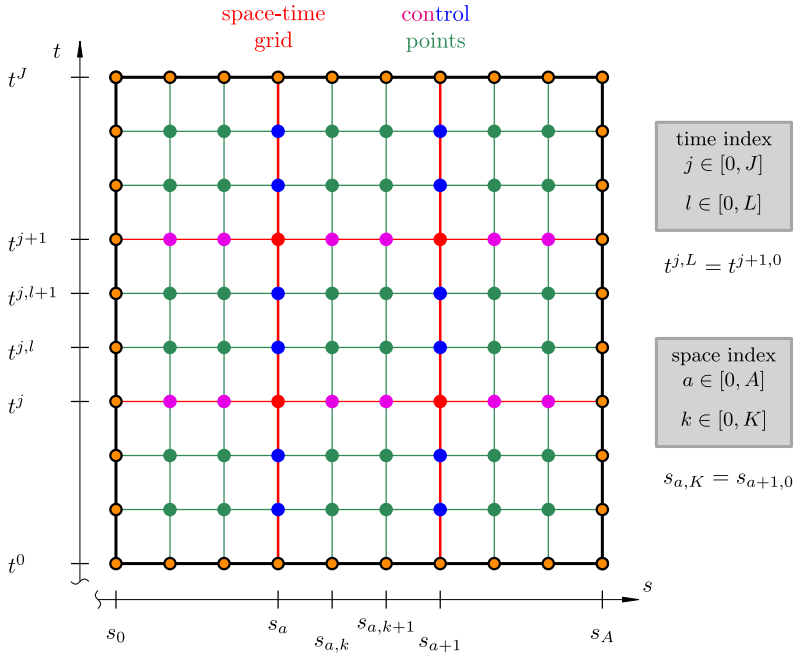

Fig. 1. Visualization of the space–time grid.

where b^{i_l} and b_{i_k} with $i_l \in \{0, \dots, I_l\}$ and $i_k \in \{0, \dots, I_k\}$ are weights with $\sum_{i_l=0}^{I_l} \sum_{i_k=0}^{I_k} b^{i_l} b_{i_k} = 1$ and $\left(\tau^{i_l}, \sigma_{i_k}\right)$ are the corresponding quadrature points.

According to the principle of stationary action, the variation of the discrete action $\mathcal{S}_\mathrm{d} = \sum_{j=0}^{J-1} \sum_{a=0}^{A-1} (\tilde{L}_d)_a^j$ vanishes for all $\bigcup_{j=0}^{J-1} \bigcup_{a=0}^{A-1} \left\{\delta x_{a,k}^{j,l}\right\}_{k=0\dots K}^{l=0\dots L}$ with $\delta x_{a,k}^{j,l} = 0$ on the boundary of space–time [4], this leads to the discrete Euler–Lagrange equations

$$P^T\left(\tilde{q}_{a,0}^{j,0}\right) \cdot \left[\frac{\partial (\tilde{L}_d)_a^j}{\partial \tilde{q}_{a,0}^{j,0}} + \frac{\partial (\tilde{L}_d)_a^{j-1}}{\partial \tilde{q}_{a,0}^{j-1,L}} + \frac{\partial (\tilde{L}_d)_{a-1}^j}{\partial \tilde{q}_{a-1,K}^{j,0}} + \frac{\partial (\tilde{L}_d)_{a-1}^{j-1}}{\partial \tilde{q}_{a-1,K}^{j-1,L}}\right] = 0 \tag{6a}$$

$$P^T\left(\tilde{q}_{a,0}^{j,l}\right) \cdot \left[\frac{\partial (\tilde{L}_d)_a^j}{\partial \tilde{q}_{a,0}^{j,l}} + \frac{\partial (\tilde{L}_d)_{a-1}^j}{\partial \tilde{q}_{a-1,K}^{j,l}}\right] = 0 \qquad \text{for } 0 < l < L \tag{6b}$$

$$P^T\left(\tilde{q}_{a,k}^{j,0}\right) \cdot \left[\frac{\partial (\tilde{L}_d)_a^j}{\partial \tilde{q}_{a,k}^{j,0}} + \frac{\partial (\tilde{L}_d)_a^{j-1}}{\partial \tilde{q}_{a,k}^{j-1,L}}\right] = 0 \qquad \text{for } 0 < k < K \tag{6c}$$

$$P^T\left(\tilde{q}_{a,k}^{j,l}\right) \cdot \left[\frac{\partial (\tilde{L}_d)_a^j}{\partial \tilde{q}_{a,k}^{j,l}}\right] = 0 \qquad \text{for } \begin{matrix} 0 < l < L \\ 0 < k < K \end{matrix} \tag{6d}$$

for $j = 1, \dots, J-1$ and $a = 1, \dots, A-1$. Again, note that $\tilde{q}_a^j = \tilde{q}_{a,0}^{j,0} = \tilde{q}_{a,0}^{j-1,L} = \tilde{q}_{a-1,K}^{j,0} = \tilde{q}_{a-1,K}^{j-1,L}$. The discrete Euler–Lagrange equations above are a generalization of [13, equations 4:15a through 4.15d]. To give boundary conditions in the discrete setting, we need to utilize the continuous conjugate momenta $\frac{\partial \tilde{L}}{\partial \dot{\tilde{q}}}(t, s, \tilde{q}, \dot{\tilde{q}}, \tilde{q}')$ and $\frac{\partial \tilde{L}}{\partial \tilde{q}'}(t, s, \tilde{q}, \dot{\tilde{q}}, \tilde{q}')$. Inserting the continuous boundary conditions given in Eq. (4) into the conjugate momenta, Eq. (6a) is subject to the boundary conditions

$$P^T\left(\tilde{q}_a^0\right) \cdot \left[\frac{\partial (\tilde{L}_d)_a^0}{\partial \tilde{q}_{a,0}^{0,0}} + \frac{\partial (\tilde{L}_d)_{a-1}^0}{\partial \tilde{q}_{a-1,K}^{0,0}}\right] = -P^T\left(\tilde{q}_a^0\right) \cdot \left(\int_{s_{a-1}}^{s_{a+1}} \left[\frac{\partial \tilde{L}}{\partial \dot{\tilde{q}}} \cdot \frac{\partial \hat{\tilde{q}}}{\partial \tilde{q}_a^0}\right]_{t^0} \mathrm{d}s\right) \qquad \text{for } 0 < a < A$$

$$P^T\left(\tilde{q}_0^j\right)\cdot\left[\frac{\partial(\tilde{L}_d)_0^j}{\partial\tilde{q}_{0,0}^{j,0}}+\frac{\partial(\tilde{L}_d)_0^{j-1}}{\partial\tilde{q}_{0,0}^{j-1,L}}\right]=-P^T\left(\tilde{q}_0^j\right)\cdot\left(\int_{t^{j-1}}^{t^{j+1}}\left[\frac{\partial\tilde{L}}{\partial\tilde{q}'}\cdot\frac{\partial\hat{\tilde{q}}}{\partial\tilde{q}_0^j}\right]_{s_0}\mathrm{d}t\right)\qquad\text{for}\quad 0<j<J$$

$$P^T\left(\tilde{q}_A^j\right)\cdot\left[\frac{\partial(\tilde{L}_d)_{A-1}^j}{\partial\tilde{q}_{A-1,K}^{j,0}}+\frac{\partial(\tilde{L}_d)_{A-1}^{j-1}}{\partial\tilde{q}_{A-1,K}^{j-1,L}}\right]=P^T\left(\tilde{q}_A^j\right)\cdot\left(\int_{t^{j-1}}^{t^{j+1}}\left[\frac{\partial\tilde{L}}{\partial\tilde{q}'}\cdot\frac{\partial\hat{\tilde{q}}}{\partial\tilde{q}_A^j}\right]_{s_A}\mathrm{d}t\right)\qquad\text{for}\quad 0<j<J$$

and

$$P^T\left(\tilde{q}_0^0\right)\cdot\left[\frac{\partial(\tilde{L}_d)_0^0}{\partial\tilde{q}_{0,0}^{0,0}}\right]=P^T\left(\tilde{q}_0^0\right)\cdot\left(-\int_{s_0}^{s_1}\left[\frac{\partial\tilde{L}}{\partial\dot{\tilde{q}}}\cdot\frac{\partial\hat{\tilde{q}}}{\partial\tilde{q}_0^0}\right]_{t^0}\mathrm{d}s-\int_{t^0}^{t^1}\left[\frac{\partial\tilde{L}}{\partial\tilde{q}'}\cdot\frac{\partial\hat{\tilde{q}}}{\partial\tilde{q}_0^0}\right]_{s_0}\mathrm{d}t\right)$$

$$P^T\left(\tilde{q}_A^0\right)\cdot\left[\frac{\partial(\tilde{L}_d)_{A-1}^0}{\partial\tilde{q}_{A-1,K}^{0,0}}\right]=P^T\left(\tilde{q}_A^0\right)\cdot\left(-\int_{s_{A-1}}^{s_A}\left[\frac{\partial\tilde{L}}{\partial\dot{\tilde{q}}}\cdot\frac{\partial\hat{\tilde{q}}}{\partial\tilde{q}_A^0}\right]_{t^0}\mathrm{d}s+\int_{t^0}^{t^1}\left[\frac{\partial\tilde{L}}{\partial\tilde{q}'}\cdot\frac{\partial\hat{\tilde{q}}}{\partial\tilde{q}_A^0}\right]_{s_A}\mathrm{d}t\right)$$

These are visualized in Fig. 2 as $\mathcal{A}$, $\mathcal{B}_{\text{left}}$, $\mathcal{B}_{\text{right}}$, $\mathcal{C}_{\text{left}}$ and $\mathcal{C}_{\text{right}}$ from top to bottom. Eq. (6b) is subject to the boundary conditions

$$P^T\left(\tilde{q}_{0,0}^{j,l}\right)\cdot\left[\frac{\partial(\tilde{L}_d)_0^j}{\partial\tilde{q}_{0,0}^{j,l}}\right]=-P^T\left(\tilde{q}_{0,0}^{j,l}\right)\cdot\left(\int_{t^j}^{t^{j+1}}\left[\frac{\partial\tilde{L}}{\partial\tilde{q}'}\cdot\frac{\partial\hat{\tilde{q}}}{\partial\tilde{q}_{0,0}^{j,l}}\right]_{s_0}\mathrm{d}t\right)\qquad\text{for}\quad\begin{matrix}0\le j<J\\0<l<L\end{matrix}$$

$$P^T\left(\tilde{q}_{A,0}^{j,l}\right)\cdot\left[\frac{\partial(\tilde{L}_d)_{A-1}^j}{\partial\tilde{q}_{A-1,K}^{j,l}}\right]=P^T\left(\tilde{q}_{A,0}^{j,l}\right)\cdot\left(\int_{t^j}^{t^{j+1}}\left[\frac{\partial\tilde{L}}{\partial\tilde{q}'}\cdot\frac{\partial\hat{\tilde{q}}}{\partial\tilde{q}_{A,0}^{j,l}}\right]_{s_A}\mathrm{d}t\right)\qquad\text{for}\quad\begin{matrix}0\le j<J\\0<l<L\end{matrix}$$

These are visualized in Fig. 2 as $\mathcal{D}_{\text{left}}$ and $\mathcal{D}_{\text{right}}$ from top to bottom. Eq. (6c) is subject to the boundary conditions

$$P^T\left(\tilde{q}_{a,k}^{0,0}\right)\cdot\left[\frac{\partial(\tilde{L}_d)_a^0}{\partial\tilde{q}_{a,k}^{0,0}}\right]=-P^T\left(\tilde{q}_{a,k}^{0,0}\right)\cdot\left(\int_{s_a}^{s_{a+1}}\left[\frac{\partial\tilde{L}}{\partial\dot{\tilde{q}}}\cdot\frac{\partial\hat{\tilde{q}}}{\partial\tilde{q}_{a,k}^{0,0}}\right]_{t^0}\mathrm{d}s\right)\qquad\text{for}\quad\begin{matrix}0\le a<A\\0<k<K\end{matrix}$$

and is visualized in Fig. 2 as $\mathcal{E}$. The boundary conditions given here are equivalent to the continuous boundary conditions given in Eq. (4) and allow forward integration in time. Other use cases, such as optimal control, might require different sets of boundary conditions which are deducible from the above.

5.1. Convergence

In standard discrete mechanics we have the theory of variational error which relates the order of convergence of a constant-step method with how good our discrete Lagrangian approximates the corresponding exact discrete Lagrangian of the problem. As far as we know, no equivalent result has been obtained for field theories as of yet. Still, we work under the assumption that a similar result must be true. Moreover, in a problem with a geometry on the base space X that allows one to separate into distinct dimensions such as the spatial and the temporal in this one, we may *integrate out* either of these dimensions producing discrete Lagrangians akin to those of standard discrete mechanics, albeit in higher dimensions:

$$(\tilde{L}_d^{\text{temporal}})^j:\widetilde{\mathbb{H}}^{(L+1)(K+1)(A+1)}\to\mathbb{R},\quad(\tilde{L}_d^{\text{temporal}})^j=\sum_{a=0}^{A}(\tilde{L}_d)_a^j;$$

$$(\tilde{L}_d^{\text{spatial}})^j:\widetilde{\mathbb{H}}^{(L+1)(K+1)(J+1)}\to\mathbb{R},\quad(\tilde{L}_d^{\text{spatial}})_a=\sum_{j=0}^{J}(\tilde{L}_d)_a^j.$$

Thus, it makes sense to study the convergence in one dimension with the other one fixed, though we do not claim that these necessarily converge to an exact discrete Lagrangian. In fact, the solutions of the discrete Euler-Lagrange

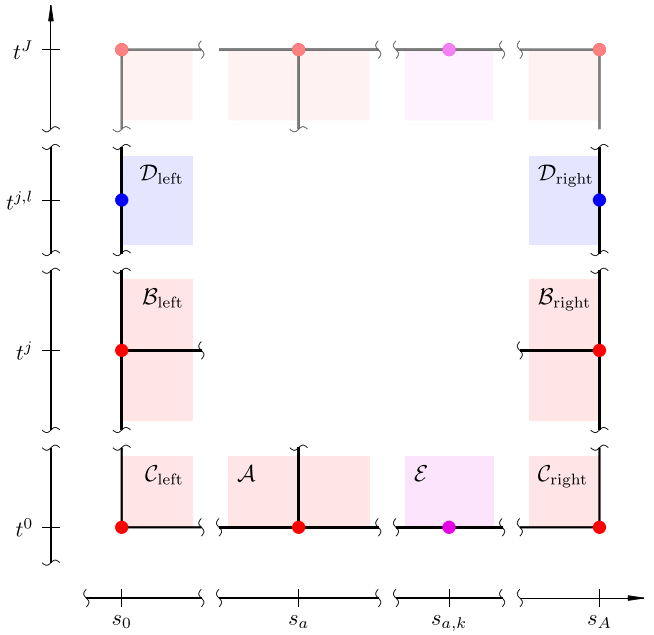

Fig. 2. Visualization of the boundary conditions for the discrete Euler–Lagrange equations on a two-dimensional space–time.

equations of these *integrated* discrete Lagrangians will not converge to the exact solution unless the step in the integrated dimension also goes to zero.

In [6,7] the order of convergence of high-order Galerkin variational methods is studied for free and constrained systems respectively in vector spaces. This is standard nodal value convergence order (i.e. at each step), as the step-size h is modified. We assume that the theory applies to our case in each dimension. There, the global order is shown to be $p = \min(2s, u)$, s is the degree of the underlying polynomial interpolant and u is the order of the quadrature.

The error inside each step, between the Galerkin interpolant curve and the exact solution, for Galerkin variational integrators was studied in [8], Theorem 3.9. There it was shown that if an integrator is of order p (local error h^{p+1}), an estimate of this error is $h^{(p+1)/2}$. We have another estimate from Sobolev interpolation theory (see [26], Theorem 3.1.6 and [27], Theorem 4.4.4). In that case, assuming the exact solution is in $H^{s+1}([0,h]) = W^{s+1,2}([0,h])$, the deviation between exact solution and interpolant measured in $W^{0,2}([0,h])$, with s the order of the underlying polynomial basis of the approximation, is of order h^{s+1}.

The order of convergence as the degree of the underlying functional space for the interpolant and the quadrature grow, is what we call polynomial refinement order. This was studied also in [8], Theorem 3.4. The result states that polynomial refinement provides a bound of the form C^s, with C constant, for the difference between the exact discrete Lagrangian and the resulting approximation via interpolation. No information on the magnitude of these constants is provided.

6. Numerical tests and convergence analysis

For the numerical convergence analysis, two batches of simulations were run which involved the same physical system: A beam with the parameters shown in Table 1 is let go without gravity and oscillates due to an initial deformation. The chosen initial configuration is $(p_0(s) = p_{\text{bend}}(s)p_{\text{twist}}(s), x_0(s)) \in \mathbb{H}_1 \times \mathfrak{h}_1$, where $p_{\text{bend}}(s) = \cos(\alpha s) + \sin(\alpha s)\mathbf{k}$, $p_{\text{twist}}(s) = \cos(\beta s) + \sin(\beta s)\mathbf{i}$ and $x_0(s) = \sin(\gamma s)\mathbf{i} + (\cos(\gamma s) - 1)\mathbf{j}$, with $\alpha = \beta = \gamma/4 = \pi/(400\ell)$ and $\ell = 1$.

For all cases, the nodes of the interpolation polynomials (τ^l, σ_k), for $l = 0, \ldots, L$ and $k = 0, \ldots, K$, are Chebyshev–Lobatto points of the form

$$\tau^l = \frac{1}{2}\left[1 + \cos\left(\frac{(L-l)}{L}\pi\right)\right]$$

Table 1
Beam parameters for the simulation of the numerical convergence analysis of MGLGVI.

L_{ength}	1 m
Θ	$\text{diag}\,(1, 1, 2) \times 10^{-1}$ N m
ρ	10 kg m^{-1}
C_1	$\text{diag}\,(2.1277, 2.1277, 5) \times 10^5$ N
C_2	$\text{diag}\,(416.6667, 416.6667, 354.6099)$ N m^2

Table 2
Parameters of the simulation batches for the numerical convergence analysis of MGLGVI.

Para.	$\mathcal{B}_t$	$\mathcal{B}_s$
T	0.001 s	0.005 s
J	2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 19, 21, 23, 25, 28, 31, 34, 37, 40, 44, 49, 54, 59, 65, 71, 78, 86, 94, 104, 114, 125, 137, 151, 166, 182, 200	2
A	12	10, 11, 12, 14, 15, 17, 19, 21, 23, 25, 28, 31, 35, 38, 42, 47, 52, 58, 64, 71, 79, 88, 97, 108, 119, 132, 147, 163, 180, 200
L	1, 2, 3, 4, 5	1
K	1	1, 2, 3, 4

and analogous for σ_k. The quadrature rules used on the spatial and the temporal directions are L and K-stage Gaussian rules, respectively.

The computations are carried out in a similar way as described in Section 6 in [9]. In particular, the Jacobian was pre-generated in pieces and automatically assembled and a standard Newton–Raphson scheme implementation was used. Details about the used Cayley map and the structure of the Jacobian can be found in [9].

Batch $\mathcal{B}_t$ is used to investigate convergence rates with temporal refinements and batch $\mathcal{B}_s$ for spatial refinements. In $\mathcal{B}_t$, both the number of space steps, as well as the spatial polynomial degree are kept constant, while the number of time steps and the temporal polynomial degree are varied. Consequently, in $\mathcal{B}_s$ the number of time steps, as well as the temporal polynomial degree are kept constant, while the number of space steps and the spatial polynomial degree are varied. The batch parameters are given in Table 2, where T is the simulation time, J is the number of time steps, A is the number of space steps and L and K are the temporal and spatial polynomial degrees respectively.

In the following sections, we discuss the convergence orders which are all compiled in Table 3.

The tolerance for the Newton–Raphson scheme for solving the discrete Euler–Lagrange equations is set to 10^{-9}. All the plots in the figures we discuss in the following sections exhibit a convergence cutoff at around 10^{-9}. This is caused both by the tolerance in the Newton–Raphson iteration and the accuracy of the reference solution, which is produced using the same integrator with more time and space steps and higher polynomial degree. In order to reduce the cutoff, one would have to decrease the tolerance of the Newton–Raphson scheme and also produce a sufficiently accurate reference solution.

6.1. Time and space step refinement

We have two types of errors: nodal or main grid error, e^{vi}, and Galerkin interpolation error, e^{gc}, respectively defined in Eq. (A.3) and Eq. (A.2).

The nodal errors over time and space step refinement, $e^{\text{vi},t}$ and $e^{\text{vi},s}$ respectively, are presented in Figs. 7 and 8. Here the errors are of order $\mathcal{O}\left(\Delta t^{\Sigma^{\text{vi},t}}\right)$ and $\mathcal{O}\left(\Delta s^{\Sigma^{\text{vi},s}}\right)$ where $\Sigma^{\text{vi},t} = 2L$ and $\Sigma^{\text{vi},s} = 2K$ respectively. These results were expected, as they coincide with the error corresponding to the standard Gaussian quadrature rule used. This is consistent with the findings reported in [6,7].

Table 3
Convergence orders over space, time and polynomial refinement.

K	1	2	3	4	–	Figure
$\Sigma^{\mathrm{gc},s}$	2	3	4	5	–	5
$\Sigma^{\mathrm{vi},s}$	2	4	6	8	–	7
L	1	2	3	4	5	Figure
$\Sigma^{\mathrm{gc},t}$	2	4*	5*	6*	7*	6
$\Sigma^{\mathrm{vi},t}$	2	4	6	8	10	8
A	15	25	8	71	–	Figure
$\Gamma^{\mathrm{gc},s}$	0.085	0.05	0.035	0.02	–	9
$\Gamma^{\mathrm{vi},s}$	0.01	0.003	0.002	0.001	–	11
J	6	11	25	59		Figure
$\Gamma^{\mathrm{gc},t}$	0.1	0.055	0.03	0.005	–	10
$\Gamma^{\mathrm{vi},t}$	0.06	0.044	0.03	0.003	–	12

Numbers marked with '$*$' are higher than expected.

The Galerkin error over space-step refinement, $e^{\mathrm{gc},s}$, is presented in the Fig. 5. The error is of order $\mathcal{O}\left(\Delta s^{\Sigma^{\mathrm{gc},s}}\right)$ where $\Sigma^{\mathrm{gc},s} = K + 1$. This is consistent with standard Sobolev interpolation theory and smaller than the analytical upper bound given in [8].

The Galerkin error over time-step refinement, $e^{\mathrm{gc},t}$, is presented in the Fig. 6. The error is of order $\mathcal{O}\left(\Delta t^{\Sigma^{\mathrm{gc},t}}\right)$ where $\Sigma^{\mathrm{gc},t} = L + 1$ for $L = 1$ and $\Sigma^{\mathrm{gc},t} = L + 2$ for $L \geq 2$. It is not clear why there is a jump in the order between $L = 1$ and $L \geq 2$ but we can only speculate that it might be something particular to this problem.

6.2. Polynomial refinement

The errors e^{gc} and e^{vi} over temporal and spatial polynomial refinement are presented in Figs. 9–12. The errors are of order $\mathcal{O}\left(\left(\Gamma^{\mathrm{vi},t}\right)^L\right)$ and $\mathcal{O}\left(\left(\Gamma^{\mathrm{vi},s}\right)^L\right)$ for the main grid and $\mathcal{O}\left(\left(\Gamma^{\mathrm{gc},t}\right)^L\right)$ and $\mathcal{O}\left(\left(\Gamma^{\mathrm{gc},s}\right)^L\right)$ for the Galerkin curve where $\Gamma^{\mathrm{vi},t} < \Gamma^{\mathrm{gc},t} < 1$ and $\Gamma^{\mathrm{vi},s} < \Gamma^{\mathrm{gc},s} < 1$. The solid lines in the figures provide information about the constants $\Gamma^{\mathrm{vi},t}$, $\Gamma^{\mathrm{vi},s}$, $\Gamma^{\mathrm{gc},t}$ and $\Gamma^{\mathrm{gc},s}$. Since no information on the constants is given in [8], we can conclude only for the Galerkin curve, that due to $\Gamma^{\mathrm{gc},t} < \sqrt{\Gamma^{\mathrm{vi},t}}$ and $\Gamma^{\mathrm{gc},s} < \sqrt{\Gamma^{\mathrm{vi},s}}$, our numerically obtained errors are smaller than the analytically given upper bound.

6.3. Shear locking tests

As previously discussed in Section 3.4, by interpolating in $\widetilde{\mathbb{H}}$ the method sidesteps the phenomenon known as shear locking. Fig. 3 shows the static problem of a homogeneous beam with a torque acting on both ends which should bend it into a circumference, the exact solution of which can be given analytically. In both cases, $K = 1$ was used, i.e. linear interpolation in space and trapezoidal rule for the approximation of the action integral, yet the red beam uses an interpolation in $\mathbb{H} \times \mathbb{R}^3$, treating positions and orientations independently, while the green beam uses the interpolation method proposed. As a result, the red beam appears to be too stiff and so suffers from shear locking. On the bottom, Fig. 3 shows the influence of the orientation of the interpolated cross sections on the positions on the right and the lack thereof on the left.

Fig. 4 shows the interpolation between the cross sections of the four corners of a space–time element, highlighting the difference in treating positions and orientations independently at the top and using unit dual quaternion interpolation at the bottom. At the bottom we see that the orientation of the cross sections has effects on the positions, which is why shear locking is prevented.

7. Conclusion and outlook

The multisymplectic Galerkin Lie group variational integrator presented in this paper shows the expected convergence rate on the main grid [6] while outperforming the expected convergence rate for the entire Galerkin

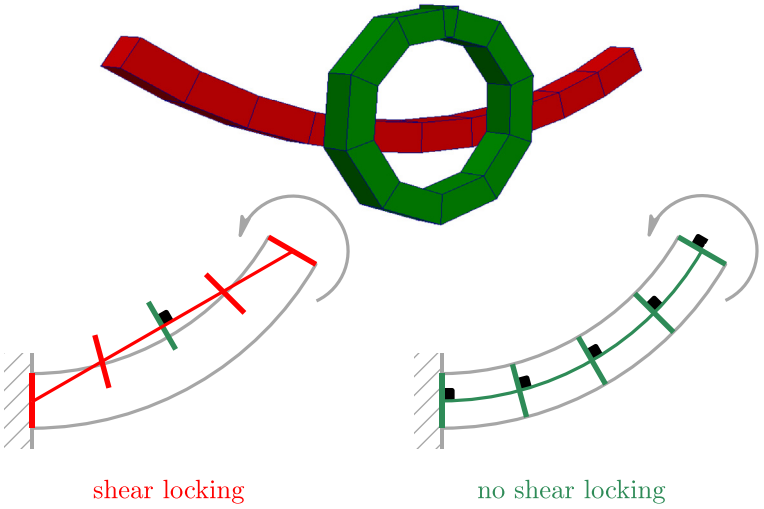

Fig. 3. Visualization of shear locking.

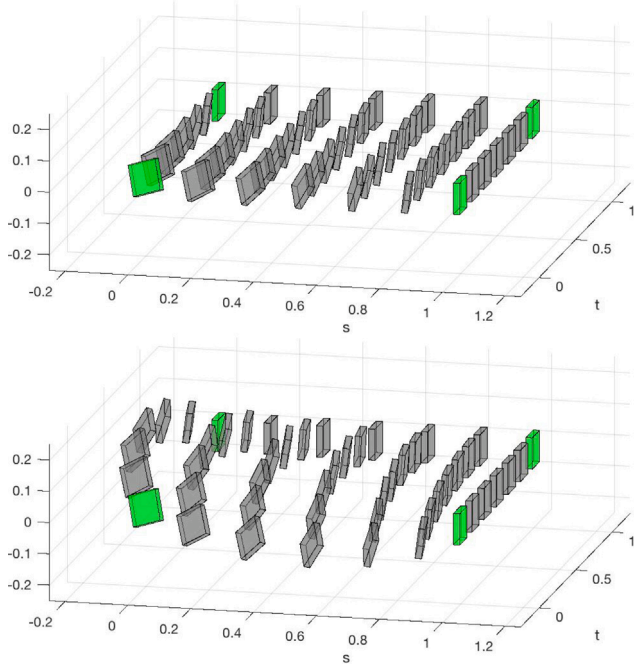

Fig. 4. Comparison of two interpolation methods. Treating positions and orientations independently (top) and unit dual quaternion interpolation (bottom).

curve [8] as we have shown for the symplectic Galerkin Lie group variational integrator based on unit quaternion interpolation in [9].

The possibility to tune orders of convergence for temporal and spatial discretization arbitrarily is one benefit of the Galerkin approach. While of course also other choices of configuration space, interpolation and discretization may lead to approximate solutions that respect the deformation (thus showing no shear locking), increasing their order of convergence might be challenging. Nevertheless, for Galerkin variational integrators, avoiding shear locking

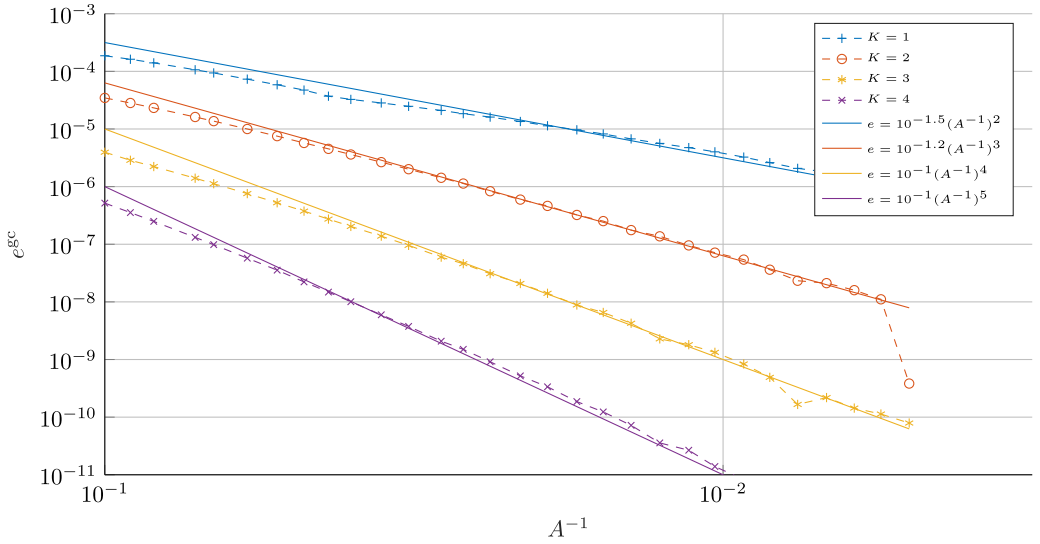

Fig. 5. Error in $\tilde{p}$ on Galerkin curve (dashed line) versus space step refinement for different spatial polynomial degrees, solid line for comparison.

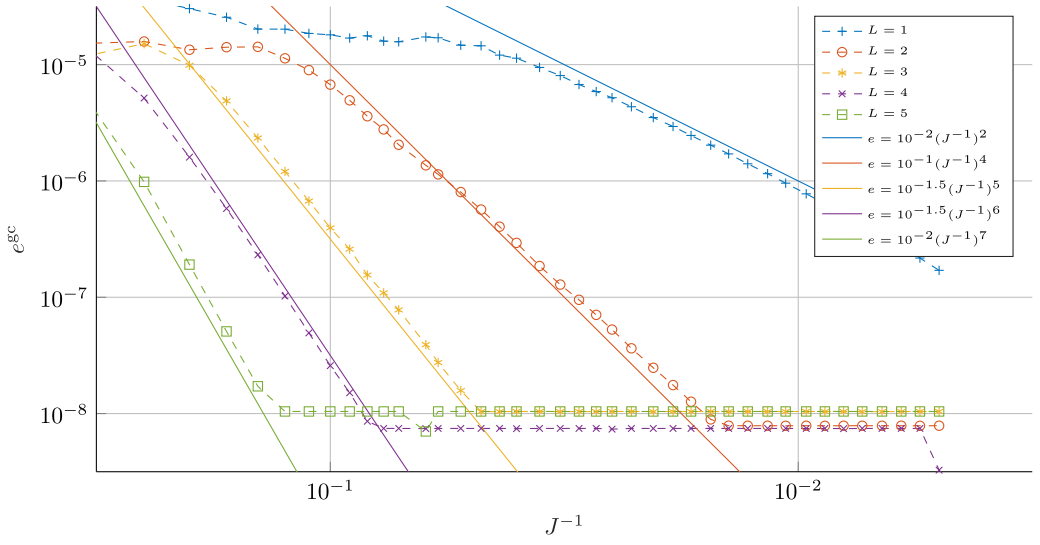

Fig. 6. Error in $\tilde{p}$ on Galerkin curve (dashed line) versus time step refinement for different temporal polynomial degrees, solid line for comparison.

may not always be possible or easily achievable by other means, while the unit dual quaternion approach provides a nice alternative that is not difficult to implement. Furthermore, the vector space nature of dual quaternions and compatibility of operations between these and $SE(3)$ make it a prime candidate to use them as ambient space to treat the problem as a constrained theory.

In the future an analysis of the computational efficiency and a comparison with other methods should be investigated. The order for each space–time element could be chosen independently e.g. based on the stiffness of the beam in that area.

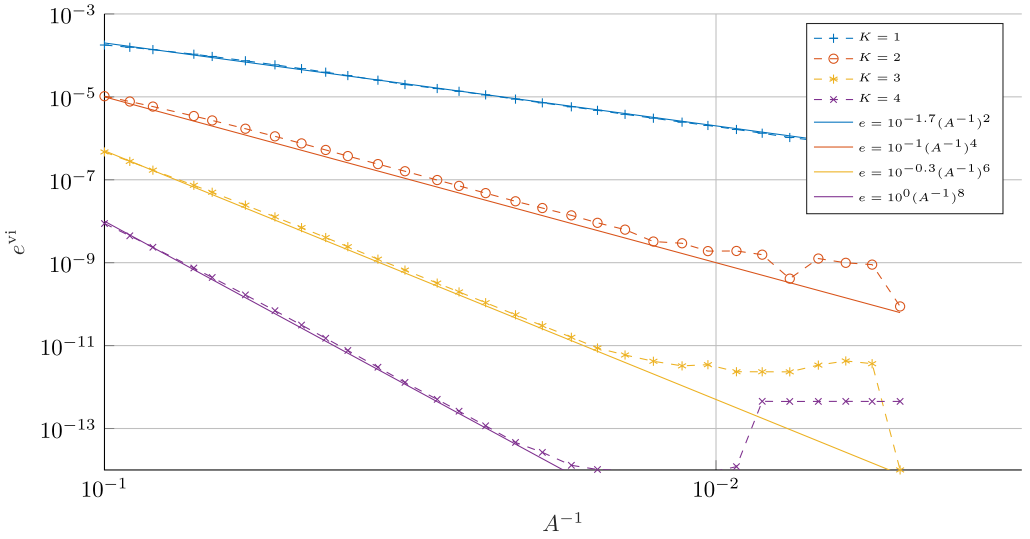

Fig. 7. Error in $\tilde{p}$ on the main space–time nodes (dashed line) versus space step refinement for different spatial polynomial degrees, solid line for comparison.

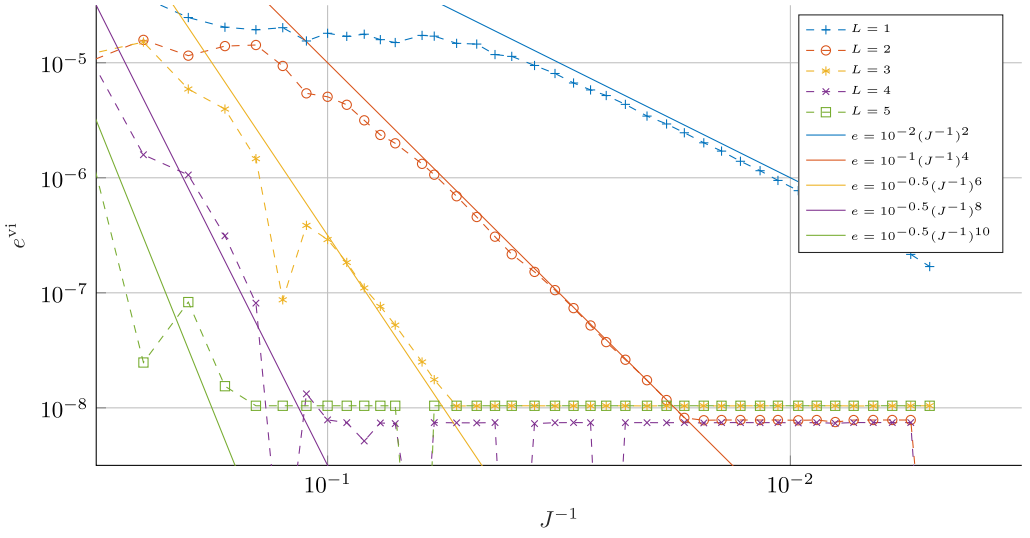

Fig. 8. Error in $\tilde{p}$ on the main space–time nodes (dashed line) versus time step refinement for different temporal polynomial degrees, solid line for comparison.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix. Errors

The function $\tilde{p}\,(t,s) : X \to \tilde{\mathbb{H}}^1$ is the Galerkin curve of the solution of the discrete Euler–Lagrange equations given in Eqs. (6a)–(6d). In absence of an analytical solution of the continuous Euler–Lagrange equations, a reference solution $\tilde{p}_{\mathrm{ref}}\,(t,s)$ is produced using the same integrator with more time and space steps and higher polynomial

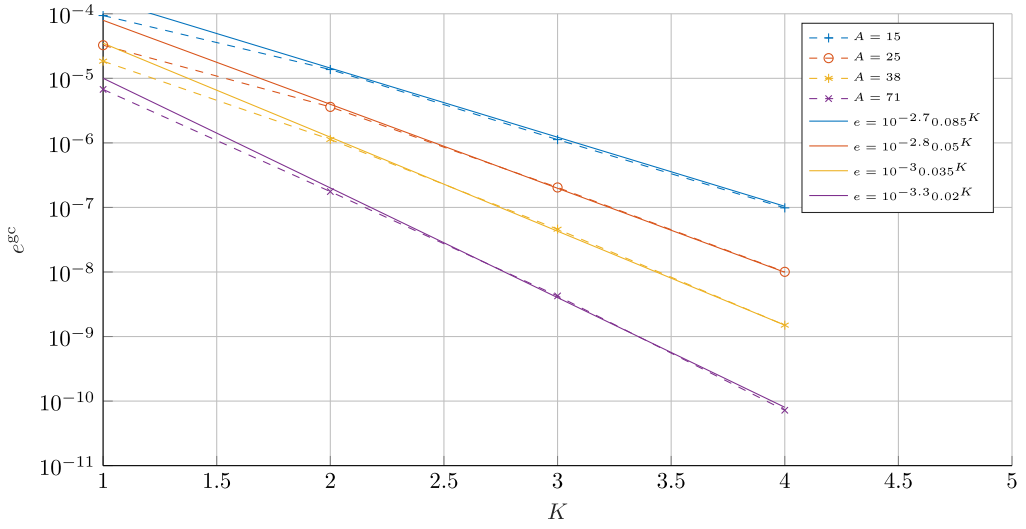

Fig. 9. Error in $\tilde{p}$ on Galerkin curve (dashed line) versus spatial polynomial refinement for different numbers of space steps, solid line for comparison.

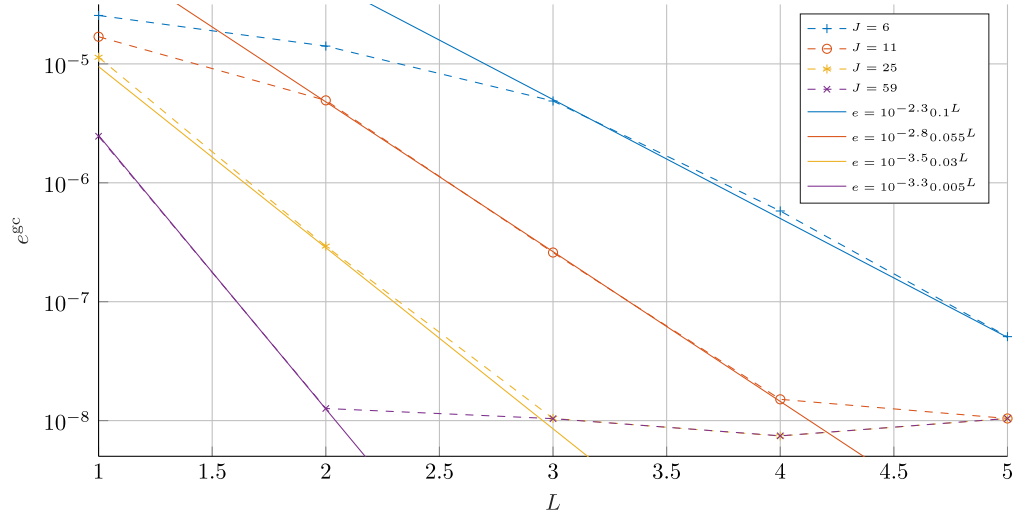

Fig. 10. Error in $\tilde{p}$ on Galerkin curve (dashed line) versus temporal polynomial refinement for different numbers of time steps, solid line for comparison.

degree. $\tilde{p}(t,s)$ and $\tilde{p}_{\text{ref}}(t,s)$ are shorthands for

$$\tilde{p}(t,s) = \tilde{p}\left(\left\{\tilde{p}^l\right\}_{k=0\ldots K}^{l=0\ldots L}, \tau(t), \sigma(s)\right)$$
$$\tilde{p}_{\text{ref}}(t,s) = \tilde{p}_{\text{ref}}\left(\left\{p_{\text{ref}}^l\right\}_{k=0\ldots K}^{l=0\ldots L}, \tau_{\text{ref}}(t), \sigma_{\text{ref}}(s)\right)$$

as defined in Eq. (2).

Error $e_{\tilde{p}}$. With the collection of space–time nodes $\boldsymbol{X} = \left[\left(t^0, s_0\right), \ldots, \left(t^J, s_A\right)\right]$, the function $\xi\left[\tilde{p}(t)\right](\boldsymbol{X})$ is defined as

$$\xi\left[\tilde{p}(t,s)\right](\boldsymbol{X}) = \left[\tilde{p}\left(t^0, s_0\right); \ldots; \tilde{p}\left(t^J, s_A\right)\right]$$

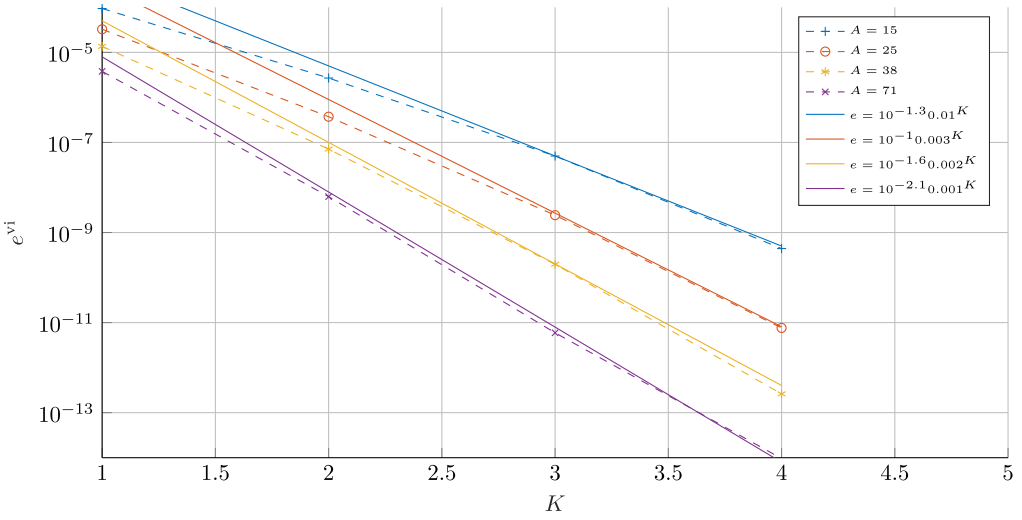

Fig. 11. Error in $\tilde{p}$ on the main space–time nodes (dashed line) versus spatial polynomial refinement for different numbers of space steps, solid line for comparison.

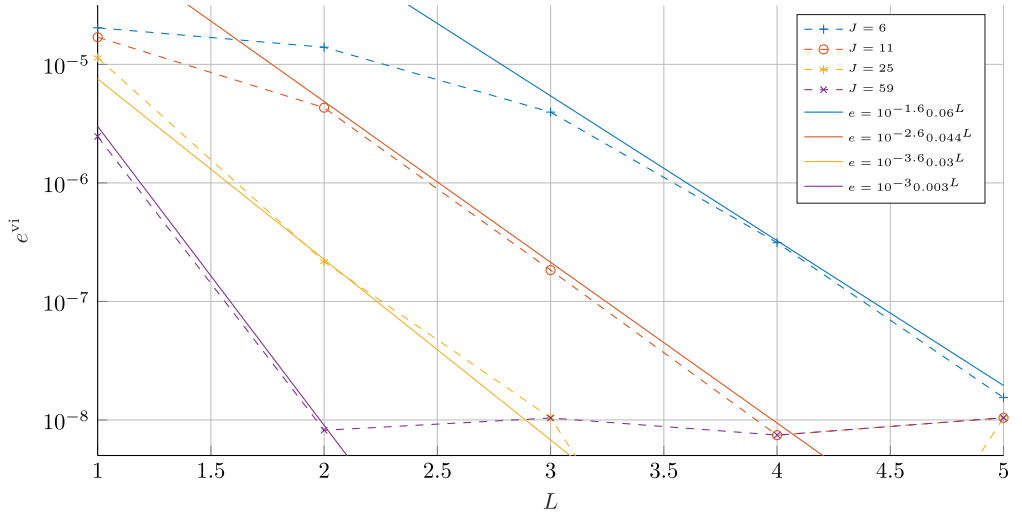

Fig. 12. Error in $\tilde{p}$ on the main space–time nodes (dashed line) versus temporal polynomial refinement for different numbers of time steps, solid line for comparison.

It creates a column vector containing all dual quaternions as 8-dimensional vectors stacked on top of each other. The error $e_{\tilde{p}}$ is defined as

$$e_{\tilde{p}}\left[\tilde{p}_{\mathrm{ref}}\left(t,s\right),\tilde{p}\left(t,s\right)\right]\left(\boldsymbol{X}\right)=\left\|\xi\left[\tilde{p}_{\mathrm{ref}}\left(t,s\right)\right]\left(\boldsymbol{X}\right)-\xi\left[\tilde{p}\left(t,s\right)\right]\left(\boldsymbol{X}\right)\right\|_{\infty} \tag{A.1}$$

where $\|\cdot\|_{\infty}$ is the maximum norm.

Error $e_{\tilde{p}}^{\mathrm{gc}}$. Using the collection of space–time nodes $\boldsymbol{X}^{\mathrm{gc}}$ with a sufficiently large number of space–time nodes, such that the maximum norm

$$\left\| \xi \left(\tilde{p}\left(t,s\right), \boldsymbol{X}^{\mathrm{gc}} \right) \right\|_{\infty} \approx \sup \left\{ \left| \tilde{p}\left(t,s\right) \right| : (t,s) \in \left[t^0, t^J \right] \times \left[s_a, s_A \right] \right\}$$

is a good approximation of the supremum norm, and using Eq. (A.1), the error of the Galerkin curve e_{p}^{gc} is defined as

$$e_{\tilde{p}}^{\mathrm{gc}} \left[\tilde{p}_{\mathrm{ref}}\left(t,s\right), \tilde{p}\left(t,s\right) \right] \left(\boldsymbol{X}^{\mathrm{gc}} \right) = e_{\tilde{p}} \left[\tilde{p}_{\mathrm{ref}}\left(t,s\right), \tilde{p}\left(t,s\right) \right] \left(\boldsymbol{X}^{\mathrm{gc}} \right) \tag{A.2}$$

Error $e_{\tilde{p}}^{\mathrm{vi}}$. Using the collection of space–time nodes $\boldsymbol{X}^{\mathrm{vi}}$ whose elements coincide with the space–time nodes of the main grid of $\tilde{p}\left(t,s\right)$, and using Eq. (A.1), the error $e_{\tilde{p}}^{\mathrm{vi}}$ is defined as

$$e_{\tilde{p}}^{\mathrm{vi}} \left[\tilde{p}_{\mathrm{ref}}\left(t,s\right), \tilde{p}\left(t,s\right) \right] \left(\boldsymbol{X}^{\mathrm{vi}} \right) = e_{\tilde{p}} \left[\tilde{p}_{\mathrm{ref}}\left(t,s\right), \tilde{p}\left(t,s\right) \right] \left(\boldsymbol{X}^{\mathrm{vi}} \right) \tag{A.3}$$

References

[1] Jürgen Moser, Alexander P. Veselov, Discrete versions of some classical integrable systems and factorization of matrix polynomials, Comm. Math. Phys. 139 (2) (1991) 217–243.
[2] Sebastian Reich, Backward error analysis for numerical integrators, SIAM J. Numer. Anal. 36 (5) (1999) 1549–1570.
[3] A. Lew, J.E. Marsden, M. Ortiz, M. West, An overview of variational integrators, in: Finite Element Methods: 1970's and Beyond, 2004, pp. 98–115.
[4] J.E. Marsden, M. West, Discrete mechanics and variational integrators, Acta Numer. 10 (2001) 357–514.
[5] James Hall, Melvin Leok, Spectral variational integrators, Numer. Math. 130 (4) (2015) 681–740.
[6] Sina Ober-Blöbaum, Nils Saake, Construction and analysis of higher order Galerkin variational integrators, Adv. Comput. Math. 41 (6) (2015) 955–986.
[7] Theresa Wenger, Sina Ober-Blöbaum, Sigrid Leyendecker, Construction and analysis of higher order variational integrators for dynamical systems with holonomic constraints, Adv. Comput. Math. (2017) 1–33.
[8] James Hall, Melvin Leok, Lie group spectral variational integrators, Found. Comput. Math. (2015) 1–59.
[9] Thomas Leitz, Sigrid Leyendecker, Galerkin Lie-group variational integrators based on unit quaternion interpolation, Comput. Methods Appl. Mech. Engrg. 338 (2018) 333–361.
[10] T. Lee, M. Leok, N.H. McClamroch, Lie group variational integrators for the full body problem, Comput. Methods Appl. Mech. Engrg. 196 (29–30) (2007) 2907–2924.
[11] Jerrold E. Marsden, George W. Patrick, Steve Shkoller, Multisymplectic geometry, variational integrators, and nonlinear PDEs, Comm. Math. Phys. 199 (1998) 351–395.
[12] Melvin Leok, Generalized Galerkin variational integrators, 2005, arXiv:math/0508360.
[13] Jing-Bo Chen, Variational integrators and the finite element method, Appl. Math. Comput. 196 (2) (2008) 941–958.
[14] F. Demoures, F. Gay-Balmaz, S. Leyendecker, S. Ober-Blöbaum, T.S. Ratiu, Y. Weinand, Discrete variational Lie group formulation of geometrically exact beam dynamics, Numer. Math. 130 (1) (2014) 73–123.
[15] François Demoures, François Gay-Balmaz, Marin Kobilarov, Tudor S. Ratiu, Multisymplectic Lie group variational integrator for a geometrically exact beam in R^3, Commun. Nonlinear Sci. Numer. Simul. 19 (10) (2014) 3492–3512.
[16] François Demoures, François Gay-Balmaz, Tudor S. Ratiu, Multisymplectic variational integrators and space/time symplecticity, 2013, arXiv:1310.4772.
[17] E. Celledoni, N. Säfström, A Hamiltonian and multi-Hamiltonian formulation of a rod model using quaternions, Comput. Methods Appl. Mech. Engrg. 199 (45–48) (2010) 2813–2819.
[18] Valentin Sonneville, Olivier Brüls, Olivier A. Bauchau, Interpolation schemes for geometrically exact beams: A motion approach, Internat. J. Numer. Methods Engrg. (2017).
[19] Jerrold E. Marsden, Tudor S. Ratiu, Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems, second ed., Springer, 1999.
[20] Taeyoung Lee, Computational Geometric Mechanics and Control of Rigid Bodies (Ph.D. thesis), University of Michigan, 2008.
[21] W.K. Clifford, Preliminary sketch of biquaternions, Proc. Lond. Math. Soc. s1-4 (1) (1871) 381–395.
[22] Yan-Bi Jia, Dual Quaternions, in: Lecture Notes, Iowa State University, 2018.
[23] Konstantinos Daniilidis, Hand-eye calibration using dual quaternions, Int. J. Robot. Res. 18 (3) (1999) 286–298.
[24] J.M. Selig, Exponential and cayley maps for dual quaternions, Adv. Appl. Clifford Algebr. 20 (3–4) (2010) 923–936.
[25] J. Simo, A finite strain beam formulation. The three-dimensional dynamic problem. Part I, Comput. Methods Appl. Mech. Engrg. 49 (1) (1985) 55–70.
[26] Philippe G. Ciarlet, The Finite Element Method for Elliptic Problems, in: Classics in Applied Mathematics, vol. 40, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002, p. xxviii+530.
[27] Susanne C. Brenner, L. Ridgway Scott, The Mathematical Theory of Finite Element Methods, third ed., in: Texts in Applied Mathematics, vol. 15, Springer, New York, 2008, p. xviii+397.