Discrete conservation properties for shallow water flows using mixed mimetic spectral elements

Lee, D.; Palha, A.; Gerritsma, M.

DOI
10.1016/j.jcp.2017.12.022

Publication date
2018

Document Version
Final published version

Published in
Journal of Computational Physics

Citation (APA)
Lee, D., Palha, A., & Gerritsma, M. (2018). Discrete conservation properties for shallow water flows using mixed mimetic spectral elements. *Journal of Computational Physics*, *357*, 282-304. https://doi.org/10.1016/j.jcp.2017.12.022

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Journal of Computational Physics 357 (2018) 282–304

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Discrete conservation properties for shallow water flows using mixed mimetic spectral elements

D. Lee [a,*], A. Palha [b], M. Gerritsma [c]

[a] Computer, Computational and Statistical Sciences, Los Alamos National Laboratory, Los Alamos, NM 87545, USA
[b] Eindhoven University of Technology, Department of Mechanical Engineering, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
[c] Delft University of Technology, Faculty of Aerospace Engineering, P.O. Box 5058, 2600 GB Delft, The Netherlands

ARTICLE INFO

Article history:
Received 2 August 2017
Received in revised form 23 November 2017
Accepted 17 December 2017
Available online 20 December 2017

Keywords:
Mimetic
Spectral elements
High order
Shallow water
Energy and potential enstrophy conservation

ABSTRACT

A mixed mimetic spectral element method is applied to solve the rotating shallow water equations. The mixed method uses the recently developed spectral element histopolation functions, which exactly satisfy the fundamental theorem of calculus with respect to the standard Lagrange basis functions in one dimension. These are used to construct tensor product solution spaces which satisfy the generalized Stokes theorem, as well as the annihilation of the gradient operator by the curl and the curl by the divergence. This allows for the exact conservation of first order moments (mass, vorticity), as well as higher moments (energy, potential enstrophy), subject to the truncation error of the time stepping scheme. The continuity equation is solved in the strong form, such that mass conservation holds point wise, while the momentum equation is solved in the weak form such that vorticity is globally conserved. While mass, vorticity and energy conservation hold for any quadrature rule, potential enstrophy conservation is dependent on exact spatial integration. The method possesses a weak form statement of geostrophic balance due to the compatible nature of the solution spaces and arbitrarily high order spatial error convergence.

Published by Elsevier Inc.

1. Introduction

In recent years there has been much interest in the use of finite element methods for the development of geophysical fluid solvers. This is in large part due to the recognition of the importance of conservation over long time integrations as a means of mitigating against both numerical instabilities and biases in the solution [1], and the capacity of finite elements to satisfy the conservation of various moments via the use of compatible or mimetic finite element spaces [2]. Various finite element spaces have been explored for their suitability for modeling geophysical flows including Raviart–Thomas, Brezzi–Douglas–Marini and Brezzi–Douglas–Fortin–Marini elements [3–5]. When used in a sequence of element types such as the standard continuous and discontinuous Galerkin elements, these can be shown to exactly satisfy the Kelvin–Stokes and Gauss-divergence theorems when applying the curl and divergence operators respectively in the weak form, as well as the annihilation of the gradient operator by the curl. Satisfying these properties exactly in the discrete form is necessary in order to conserve both first order and higher order moments for the shallow water equations in rotational form, as first presented for a C-grid finite difference scheme [6], and later formalized via the derivation of the finite difference operators from Hamiltonian methods [7,8].

* Corresponding author.
E-mail address: davelee2804@gmail.com (D. Lee).

https://doi.org/10.1016/j.jcp.2017.12.022
0021-9991/Published by Elsevier Inc.

The application of mimetic finite elements for shallow water flows has also been formalized in the language of exterior calculus in order to generalize the expression of their conservation properties [9]. Mimetic properties have also been demonstrated for the standard A-grid spectral element method on cubed sphere geometries via careful use of covariant and contra-variant transformations in order to evaluate the curl and divergence operators respectively so as to exactly satisfy the Kelvin–Stokes and Gauss-divergence theorems [10].

The present study explores the use of spectral elements using mixed basis functions in order to preserve mimetic properties for geophysical flows. These recently developed methods invoke the use of *histopolation* functions [11], which are defined such that they exactly satisfy the fundamental theorem of calculus with respect to the nodal Lagrange basis functions from which they are derived. In the language of exterior calculus these edge functions may then be regarded as defining the space of 1-forms (with the Lagrange polynomials defining the space of 0-forms). By taking tensor product combinations of these basis functions and their associated edge functions higher dimensional differential k-forms may also be constructed for which the Kelvin–Stokes and Gauss-divergence theorems are satisfied [12].

As for other compatible tensor product finite element methods, the mixed mimetic spectral elements provide an effective means of preserving many of the properties of a geophysical fluid over long time integrations. These include:

- *Conservation of mass, vorticity, total energy and potential enstrophy*
  Conservation of mass ensures that the solution does not drift over time and develop large biases. Vorticity is an important moment to conserve since the large scale circulation of the atmosphere and oceans at mid latitudes is dominated by a quasi-balance between rotation and pressure gradients. Conservation of energy ensures that solutions remain bounded and numerical instabilities do not grow exponentially. The importance of potential enstrophy conservation is less immediately obvious meanwhile since for two dimensional turbulent fluids this cascades to small scales where some dissipation mechanism is necessary [13].
- *Stationary geostrophic modes*
  The large scale circulations of the atmosphere and ocean are dominated by the slow evolution of modes balanced between Coriolis forces and pressure gradients. If this balance cannot be exactly replicated in a numerical model then the process of adjustment will result in the radiation of fast gravity waves that can contaminate the solution [14].
- *High order spatial error convergence*
  The spectral element method defines the nodes of the basis functions to cluster towards element boundaries such that spurious oscillations due to spectral ringing are avoided, allowing convergence of errors at arbitrarily high order.

The rest of this paper proceeds as follows: In the following section the shallow water equations will be briefly introduced in the continuous form. Section 3 discusses the mixed mimetic spectral element method, as introduced by previous authors [11,12,15,16], and its application to the rotating shallow water equations. Section 4 explores the conservation properties of the discrete system. In Section 5 some results are presented, demonstrating the error convergence, conservation and balance properties of the method. Section 6 details some conclusions regarding the suitability of the method for geophysical flows, its advantages and limitations, as well as some future work we intend to pursue on the topic.

2. The 2D shallow water equations

The two dimensional shallow water equations present an excellent testing ground for primitive equation atmospheric and oceanic models since they exhibit many of the same features, including slowly evolving Rossby waves and fast gravity waves, nonlinear cascades of higher moments (kinetic energy and potential enstrophy), and the conservation of various moments (mass, total energy, vorticity, and an infinite number of rotational moments, of which potential enstrophy is the first).

Let $\Omega = [x_{\min}, x_{\max}] \times [y_{\min}, y_{\max}] \subset \mathbb{R}^2$ such that $x_{\min} < x_{\max}$ and $y_{\min} < y_{\max}$ be a doubly periodic domain, and let $t_F > 0$. The rotational form of the shallow water problem, [6], consists in finding the prognostic variables velocity, $\vec{u} : \Omega \times (0, t_F] \mapsto \mathbb{R}^2$, and fluid depth, $h : \Omega \times (0, t_F] \mapsto \mathbb{R}$, such that[1]

$$\begin{cases} \dfrac{\partial \vec{u}}{\partial t} + q \times \vec{F} + \nabla(K + gh) = 0\,, & \text{in } \Omega \times (0, t_F]\,, \quad \text{(a)} \\ \dfrac{\partial h}{\partial t} + \nabla \cdot \vec{F} = 0\,, & \text{in } \Omega \times (0, t_F]\,. \quad \text{(b)} \end{cases} \tag{1}$$

The diagnostic variables are potential vorticity $q : \Omega \times (0, t_F] \mapsto \mathbb{R}$, mass flux $\vec{F} : \Omega \times (0, t_F] \mapsto \mathbb{R}^2$, and kinetic energy per unit mass $K : \Omega \times (0, t_F] \mapsto \mathbb{R}$, defined as

[1] The shallow water equations are formulated in 2D, such that $q \times \vec{F} := q\vec{e}_z \times \vec{F} = q\vec{F}^{\perp}$.

$$\begin{cases} q := \dfrac{\nabla \times \vec{u} + f}{h}, & \text{(a)} \\ \vec{F} := h\vec{u}, & \text{(b)} \\ K := \dfrac{1}{2}\vec{u} \cdot \vec{u}, & \text{(c)} \end{cases} \tag{2}$$

where $f : \Omega \mapsto \mathbb{R}$ is the Coriolis term. Note that we assume that this Coriolis term does not explicitly depend on time.

The shallow water equations conserve the following invariants, [6]:

- Volume: integrating (1a) over the domain Ω and assuming periodic boundary conditions gives the result that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega} h \,\mathrm{d}\Omega = 0\,. \tag{3}$$

  If we assume constant density, then this also gives mass conservation.
- Vorticity: taking the curl of (1a) leads to a conservation equation for vorticity $\omega = \nabla \times \vec{u} : \Omega \times (0, t_F] \mapsto \mathbb{R}$, of the form

$$\frac{\partial \omega}{\partial t} + \nabla \cdot (\vec{u}\,(\omega + f)) = 0\,. \tag{4}$$

  Integrating over the domain (and assuming periodic boundary conditions) gives the conservation of vorticity as

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega} \omega \,\mathrm{d}\Omega = 0\,. \tag{5}$$

- Energy: computing the inner product of (1a) with $\vec{F}$ and multiplying (1a) by gh gives

$$\frac{\partial hK}{\partial t} + \nabla \cdot (\vec{F}K) + \vec{F} \cdot \nabla(gh) = 0\,, \tag{6}$$

$$\frac{\partial (\frac{1}{2}gh^2)}{\partial t} + \nabla \cdot (gh\vec{F}) - \vec{F} \cdot \nabla(gh) = 0\,. \tag{7}$$

  Combining these and integrating over Ω gives energy conservation as

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega} \left\{ hK + \frac{1}{2}gh^2 \right\} \mathrm{d}\Omega = 0\,. \tag{8}$$

- Potential enstrophy: expressing the vorticity as $\omega = hq - f$ within the vorticity advection equation and subtracting q times (1a) gives an advection equation for the potential vorticity as

$$\frac{\partial q}{\partial t} + \vec{u} \cdot \nabla q = 0. \tag{9}$$

  Multiplying this by hq and (1a) by $\frac{1}{2}q^2$ and adding gives

$$\frac{\partial \frac{1}{2}hq^2}{\partial t} + \nabla \cdot \left(\frac{1}{2}\vec{F}q^2 \right) = 0. \tag{10}$$

  Integrating over the domain then gives potential enstrophy conservation as

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega} \frac{1}{2}hq^2 \,\mathrm{d}\Omega = 0\,. \tag{11}$$

For a detailed derivation of these conservation laws the reader is referred to previous studies [4,6].

One of the central objectives of this paper is to show that these properties can be satisfied at the discrete level within the mimetic spectral element discretization framework. Ensuring the conservation of invariants in discrete form helps to mitigate against biases and instabilities in the solution of the original system [1]. For this reason, it is considered important to satisfy these conservation properties at the discrete level in geophysical flow solvers, especially when long time simulation is the goal.

3. Spatial discretization

In this work we set out to construct a mimetic spectral finite element discretization for the shallow water equations as given by (1), together with the diagnostic equations (2). In particular we use a mixed finite element formulation, for more details on mixed finite elements see for example [17,18].

The mimetic finite element discretization presented here differs from other finite element methods in the degrees of freedom. Here the degrees of freedom, the expansion coefficients, represent integral values. For nodal basis functions, the degrees of freedom represent the values at points, see (31), the degrees of freedom for vectors will be the (two-dimensional) fluxes over line segments, see (35), (36) in the section. The degrees of freedom for densities will be their integrated values over 2D volumes, see (40). The main motivations for the use of these degrees of freedom are: 1. that integral values are well defined on non-orthogonal grids and 2. that the optimal order of convergence is displayed on both affine and non-affine meshes. Alternative formulations may encounter loss of optimality, [19–21]. An important feature of this sequence of basis functions is that the derivative can be applied directly to the degrees of freedom by multiplying the degrees of freedom by an appropriate incidence matrix [22–26], as shown in (44) and (47) in this section. The incidence matrices do not depend on the polynomial degree or the shape of the grid; they represent the topological derivative. In Section 4 we will prove conservation properties in terms of these topological structures. Since the proofs are based on metric-free concepts, this ensures that these properties will also hold on highly deformed grids.

3.1. Weak formulation

The first step to construct this discretization is the weak form of (1) and (2). In this work, as usual, $\langle\cdot,\cdot\rangle_\Omega$ represents the L^2 inner product

$$\langle f, g\rangle_\Omega := \int_\Omega f \cdot g \, \mathrm{d}\Omega \,. \tag{12}$$

The weak formulation reads: for any time $t \in (0, t_F]$ and for a given Coriolis term $f \in L^2(\Omega)$, find $\vec{u}, \vec{F} \in H(\mathrm{div}, \Omega)$, $h, K \in L^2(\Omega)$, and $q \in H(\mathrm{rot}, \Omega)$[2] such that

$$\begin{cases} \langle \vec{v}, \dfrac{\partial \vec{u}}{\partial t}\rangle_\Omega + \langle \vec{v}, q \times \vec{F}\rangle_\Omega - \langle \nabla\cdot\vec{v}, K + gh\rangle_\Omega = 0\,, & \forall \vec{v} \in H(\mathrm{div}, \Omega) \quad \text{(a)} \\ \langle \sigma, \dfrac{\partial h}{\partial t}\rangle_\Omega + \langle \sigma, \nabla\cdot\vec{F}\rangle_\Omega = 0\,, & \forall \sigma \in L^2(\Omega), \quad \text{(b)} \\ \langle \zeta, hq\rangle_\Omega = -\langle \nabla^\perp\zeta, \vec{u}\rangle_\Omega + \langle \zeta, f\rangle_\Omega\,, & \forall \zeta \in H(\mathrm{rot}, \Omega), \quad \text{(c)} \\ \langle \vec{\varphi}, \vec{F}\rangle_\Omega = \langle \vec{\varphi}, h\vec{u}\rangle_\Omega\,, & \forall \vec{\varphi} \in H(\mathrm{div}, \Omega), \quad \text{(d)} \\ \langle \kappa, K\rangle_\Omega = \dfrac{1}{2}\langle \kappa, \vec{u}\cdot\vec{u}\rangle_\Omega\,, & \forall \kappa \in L^2(\Omega), \quad \text{(e)} \end{cases} \tag{13}$$

where we have used integration by parts and the periodic boundary conditions to obtain the identities [4]

$$\begin{cases} \langle \vec{v}, \nabla(K + gh)\rangle_\Omega = -\langle \nabla\cdot\vec{v}, K + gh\rangle_\Omega, & \text{(a)} \\ \langle \varsigma, \nabla\times\vec{u}\rangle_\Omega = -\langle \nabla^\perp\varsigma, \vec{u}\rangle_\Omega. & \text{(b)} \end{cases} \tag{14}$$

These two relations show that minus the gradient is the Hilbert adjoint of the divergence operator and minus the curl is the Hilbert adjoint of the rot, provided the boundary integrals vanish. The space $L^2(\Omega)$ corresponds to square integrable functions and the spaces $H(\mathrm{div}, \Omega)$ and $H(\mathrm{rot}, \Omega)$ contain square integrable functions whose divergence and rot are also square integrable.

3.2. Finite dimensional mimetic function spaces

The second step to construct this discretization is the definition of the spatial conforming function spaces, where we will seek the discrete solutions for velocity $\vec{u}_h$, fluid depth h_h, mass flux $\vec{F}_h$, kinetic energy K_h and potential vorticity q_h:

$$q_h \in W_h \subset H(\mathrm{rot}, \Omega), \quad \vec{u}_h, \vec{F}_h \in U_h \subset H(\mathrm{div}, \Omega), \quad \text{and} \quad h_h, K_h \in Q_h \subset L^2(\Omega)\,. \tag{15}$$

[2] For scalar variable ψ the rot operator $\nabla^\perp := \vec{e}_z \times \nabla$ is defined as

$$\nabla^\perp\psi = \begin{pmatrix} -\partial\psi/\partial y \\ \partial\psi/\partial x \end{pmatrix}.$$

The choice of finite dimensional function spaces determines the properties of the discretization, see for example [27,28] for a more general discussion, and [2] for a recent discussion focused on the 2D Navier–Stokes equations.

Therefore, the finite dimensional function spaces used in this work are such that when combined form a Hilbert subcomplex

$$\mathbb{R} \longrightarrow W_h \xrightarrow{\nabla^\perp} U_h \xrightarrow{\nabla\cdot} Q_h \longrightarrow 0\,. \tag{16}$$

The meaning of this Hilbert subcomplex is that

$$\{\nabla^\perp \omega_h \,|\, \omega_h \in W_h\} \subset U_h \quad \text{and} \quad \{\nabla\cdot \vec{u}_h \,|\, \vec{u}_h \in U_h\} \subseteq Q_h\,. \tag{17}$$

In other words, the rot operator must map W_h into U_h and the div operator must map U_h onto Q_h.

This discrete subcomplex mimics the 2D Hilbert complex associated to the continuous functional spaces

$$\mathbb{R} \longrightarrow H(\mathrm{rot},\Omega) \xrightarrow{\nabla^\perp} H(\mathrm{div},\Omega) \xrightarrow{\nabla\cdot} L^2(\Omega) \longrightarrow 0\,. \tag{18}$$

The Hilbert complex is an important structure that is intimately connected to the de Rham complex of differential forms. Therefore, the construction of a discrete subcomplex is an important requirement to obtain a stable and accurate finite element discretization, see for example [22–26,28,29] for a detailed discussion.

Each of these finite dimensional function spaces W_h, U_h, and Q_h has an associated finite set of basis functions ${\epsilon_i}^W, \vec{\epsilon}_i^{\,U}, {\epsilon_i}^Q$, such that

$$W_h = \mathrm{span}\{{\epsilon_1}^W, \ldots, {\epsilon_{d_W}}^W\}, \quad U_h = \mathrm{span}\{\vec{\epsilon}_1^{\,U}, \ldots, \vec{\epsilon}_{d_U}^{\,U}\}, \quad \text{and} \quad Q_h = \mathrm{span}\{\epsilon_1^Q, \ldots, \epsilon_{d_Q}^Q\}\,, \tag{19}$$

where d_W, d_U, and d_Q denote the dimension of the discrete function spaces and therefore correspond to the number of degrees of freedom associated to each of the unknowns. As a consequence, the approximate solutions for vorticity, potential vorticity, Coriolis, velocity, mass flux, fluid depth, and kinetic energy can be expressed as a linear combination of these basis functions

$$\begin{cases} q_h := \displaystyle\sum_{i=1}^{d_W} q_i\, {\epsilon_i}^W, \quad f_h := \sum_{i=1}^{d_W} f_i {\epsilon_i}^W, & \text{(a)} \\ \vec{u}_h := \displaystyle\sum_{i=1}^{d_U} u_i\, \vec{\epsilon}_i^{\,U}, \quad \vec{F}_h := \sum_{i=1}^{d_U} F_i\, \vec{\epsilon}_i^{\,U}, & \text{(b)} \\ h_h := \displaystyle\sum_{i=1}^{d_Q} h_i\, \epsilon_i^Q, \quad K_h := \sum_{i=1}^{d_Q} K_i\, \epsilon_i^Q, & \text{(c)} \end{cases} \tag{20}$$

with q_i, f_i, u_i, F_i, h_i, K_i, the degrees of freedom for vorticity, potential vorticity, Coriolis, velocity, mass flux, fluid depth, and kinetic energy, respectively. Since the shallow water equations form a time dependent set of equations, these coefficients will be time dependent.

3.2.1. One dimensional nodal and histopolant polynomials

To define the two-dimensional basis functions ${\epsilon_i}^W$, $\vec{\epsilon}_i^{\,U}$, and ϵ_i^Q, we first introduce two types of one-dimensional polynomials: one associated with nodal interpolation, and the other with integral interpolation (histopolation).[3] Subsequently, these two types of polynomials will be combined to generate the two-dimensional polynomial basis functions used to discretize the physical quantities that appear in this problem.

Consider the canonical interval $I = [-1, 1] \subset \mathbb{R}$ and the Legendre polynomials, $L_p(\xi)$ of degree p with $\xi \in I$. The $p+1$ roots, ξ_i, of the polynomial $(1-\xi^2)\frac{\mathrm{d}L_p}{\mathrm{d}\xi}$ are called Gauss–Lobatto–Legendre (GLL) nodes and satisfy $-1 = \xi_0 < \xi_1 < \cdots < \xi_{p-1} < \xi_p = 1$. Let $l_i^p(\xi)$ be the Lagrange polynomial of degree p through the GLL nodes, such that

$$l_i^p(\xi_j) := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}, \quad i, j = 0, \ldots, p\,. \tag{21}$$

The explicit form of these Lagrange polynomials is given by

$$l_i^p(\xi) = \prod_{\substack{k=0 \\ k\neq i}}^{p} \frac{\xi - \xi_k}{\xi_i - \xi_k}\,. \tag{22}$$

[3] For an extensive discussion of integral interpolation (histopolation) see [11,30].

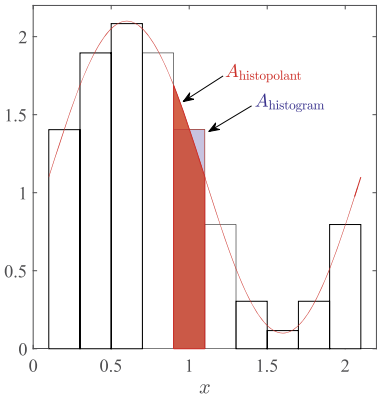

Fig. 1. Histogram and an example of a histopolant (red curve). By definition, the integral of the histopolant over each cell (or bin) $A_{\text{histopolant}}$ is equal to the area of the corresponding bar of the histogram $A_{\text{histogram}}$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Let $q_h(\xi)$ be a polynomial of degree p defined on $I = [-1, 1]$ and $q_i = q_h(\xi_i)$, then the expansion of $q_h(\xi)$ in terms of Lagrange polynomials is given by

$$q_h(\xi) := \sum_{i=0}^{p} q_i l_i^p(\xi)\,. \tag{23}$$

Because the expansion coefficients in (23) are given by the value of q_h in the nodes ξ_i, we refer to this interpolation as a *nodal interpolation* and we will denote the Lagrange polynomials in (22) by *nodal polynomials*.

Before introducing the second set of basis polynomials that will be used in this work it is important to introduce the reader to the concept of *histopolant*. Given a histogram, i.e. a piece-wise constant function, a histopolant is a smooth function whose integrals over the cells (or bins) of the histogram are equal to the area of the corresponding bars of the histogram, see Fig. 1. If the histopolant is a polynomial we say that it is a *polynomial histopolant*. In the same way as a polynomial interpolant that passes exactly through $p+1$ points has degree p, a polynomial that exactly histopolates a histogram with $p+1$ bins has polynomial degree p. Consider now a function $g(x)$ and its associated integrals over a set of cells $[a_{j-1}, a_j]$, $g_j = \int_{a_{j-1}}^{a_j} g(x)\mathrm{d}x$, with $a_0, < \cdots < a_j < \cdots < a_p$. The set of integral values g_j and cells $[a_{j-1}, a_j]$ can be seen as a histogram. As mentioned before, it is possible to construct a histopolant of this histogram. This histopolant will be an approximating function of g that has the particular property of having the same integral over the cells $[a_{j-1}, a_j]$ as the original g. Just like a nodal interpolation exactly reconstructs the original function at the interpolating points, a histopolant exactly reconstructs the integral of the original function over the cells.

Using the nodal polynomials we can define another set of basis polynomials, $e_i^p(\xi)$, as

$$e_i^p(\xi) := -\sum_{k=0}^{i-1} \frac{\mathrm{d}l_k^p(\xi)}{\mathrm{d}\xi}\,, \qquad i = 1, \ldots, p\,. \tag{24}$$

These polynomials $e_i^p(\xi)$ have polynomial degree $p-1$ and satisfy,

$$\int_{\xi_{j-1}}^{\xi_j} e_i^p(\xi)\,\mathrm{d}\xi = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}\,, \quad i, j = 1, \ldots, p\,. \tag{25}$$

The proof that the polynomials $e_i^p(\xi)$ have degree $p-1$ follows directly from the fact that their definition (24) involves a linear combination of the derivative of polynomials of degree p. The proof of (25) results from the properties of $l_k^p(\xi)$. Using (24) the integral of $e_i^p(\xi)$ becomes

$$\int_{\xi_{j-1}}^{\xi_j} e_i^p(\xi)\,\mathrm{d}\xi = -\int_{\xi_{j-1}}^{\xi_j} \sum_{k=0}^{i-1} \frac{\mathrm{d}l_k^p(\xi)}{\mathrm{d}\xi} = -\sum_{k=0}^{i-1} \int_{\xi_{j-1}}^{\xi_j} \frac{\mathrm{d}l_k^p(\xi)}{\mathrm{d}\xi} = -\sum_{k=0}^{i-1} \left(l_k^p(\xi_j) - l_k^p(\xi_{j-1})\right) = -\sum_{k=0}^{i-1} \left(\delta_{k,j} - \delta_{k,j-1}\right)\,,$$

where $\delta_{i,j}$ is the Kronecker delta. It is straightforward to see that

$$-\sum_{k=0}^{i-1} \left(\delta_{k,j} - \delta_{k,j-1}\right) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}\,, \quad i, j = 1, \ldots, p\,.$$

For more details see [11,30].

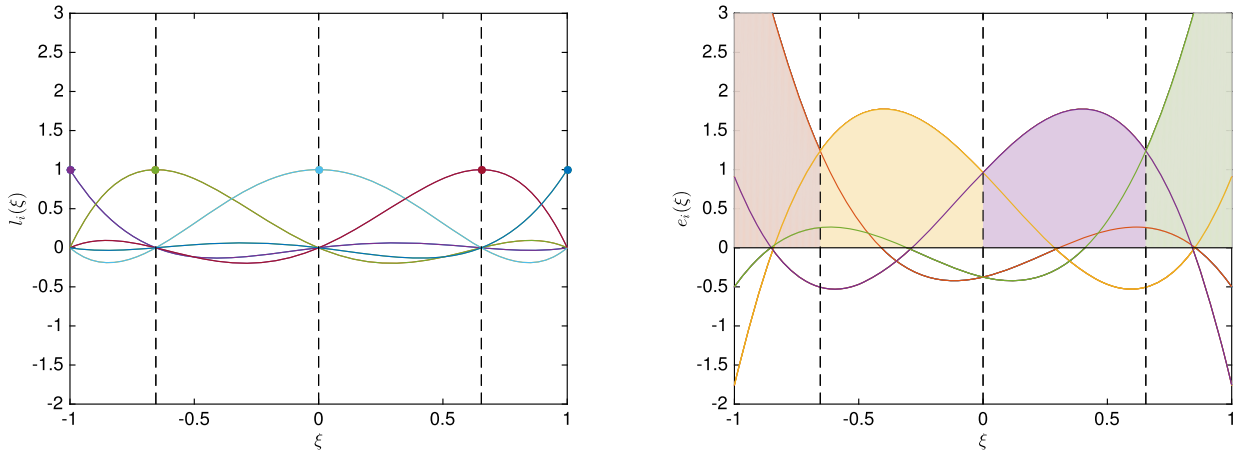

Fig. 2. Basis polynomials associated to $p = 4$. Left: nodal polynomials, the value of the basis polynomial at the corresponding node is one and on the other nodes is zero. Right: histopolant polynomials, the integral of the basis polynomials over the corresponding shaded area evaluates to one and to zero on the others.

Let $g_h(\xi)$ be a polynomial of degree $(p-1)$ defined on $I = [-1, 1]$ and $g_i = \int_{\xi_{i-1}}^{\xi_i} g_h(\xi)\,\mathrm{d}\xi$, then its expansion in terms of the polynomials $e_i^p(\xi)$ is given by

$$g_h(\xi) = \sum_{i=1}^{p} g_i e_i^p(\xi)\,. \tag{26}$$

Because the expansion coefficients in (26) are the integral values of $g_h(\xi)$, we denote the polynomials in (24) by *histopolant polynomials*[4] and refer to (26) as *histopolation*.

It can be shown, [11,30], that if $q_h(\xi)$ is expanded in terms of nodal polynomials, as in (23), then the expansion of its derivative $\frac{\mathrm{d}q_h(\xi)}{\mathrm{d}\xi}$ in terms of histopolant polynomials is

$$\begin{aligned}\left(\frac{\mathrm{d}q_h(\xi)}{\mathrm{d}\xi}\right)_h &= \sum_{i=1}^{p}\left(\int_{\xi_{i-1}}^{\xi_i}\frac{\mathrm{d}q_h(\xi)}{\mathrm{d}\xi}\mathrm{d}\xi\right)e_i^p(\xi) = \sum_{i=1}^{p}\left(q_h(\xi_i) - q_h(\xi_{i-1})\right)e_i^p(\xi)\\ &= \sum_{i=1}^{p}\left(q_i - q_{i-1}\right)e_i^p(\xi) = \sum_{i=1,j=0}^{p}\mathsf{E}_{i,j}^{1,0}q_j e_i^p(\xi)\,,\end{aligned} \tag{27}$$

where $\mathsf{E}_{i,j}^{1,0}$ are the coefficients of the $p \times (p+1)$ matrix $\mathbf{E}^{1,0}$

$$\mathbf{E}^{1,0} := \begin{pmatrix} -1 & 1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 1 & 0 & \ddots & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & \ddots & 0 & -1 & 1 & 0 \\ 0 & \dots & 0 & 0 & -1 & 1 \end{pmatrix}. \tag{28}$$

The following identity holds (Commuting property)

$$\left(\frac{\mathrm{d}q(\xi)}{\mathrm{d}\xi}\right)_h = \frac{\mathrm{d}q_h(\xi)}{\mathrm{d}\xi}\,. \tag{29}$$

For an example of the one-dimensional basis polynomials corresponding to $p = 4$, see Fig. 2.

[4] In earlier work we referred to these functions as *edge functions*.

3.2.2. Two dimensional basis functions

Consider the nodal polynomials (22), $l_i^p(\xi)$ of degree p, the histopolant polynomials (24), $e_i^p(\xi)$ of degree $p-1$, the canonical square $\Omega = I \times I \subset \mathbb{R}^2$, and take $\xi, \eta \in I = [-1, 1]$.

Basis functions for W_h Combining nodal polynomials we can construct the polynomial basis functions for W_h on a reference quadrilateral. Consider the canonical interval $I = [-1, 1]$, the canonical square $\Omega = I \times I \subset \mathbb{R}^2$, the nodal polynomials (22), $l_i^p(\xi)$ of degree p, and take $\xi, \eta \in I$. Then a set of two-dimensional basis polynomials, ${\epsilon_k}^W(\xi, \eta; p)$, on Ω can be constructed as the tensor product of the one-dimensional ones

$$ {\epsilon_k}^W(\xi, \eta; p) := l_i^p(\xi)\, l_j^p(\eta)\,, \qquad i, j = 0, \ldots, p, \quad k = j + 1 + i(p+1)\,. \tag{30} $$

These polynomials, ${\epsilon_k}^W(\xi, \eta; p)$, have degree p in each direction and from (21) it follows that they satisfy, see [11,29],

$$ {\epsilon_k}^W(\xi_i, \eta_j; p) = \begin{cases} 1 & \text{if } k = i(p+1) + j + 1 \\ \\ 0 & \text{if } k \neq i(p+1) + j + 1 \end{cases}, \quad i, j = 0, \ldots, p, \quad k = 1, \ldots, (p+1)^2\,. \tag{31} $$

Here, as before, ξ_i and η_i with $i = 0, \ldots, p$ are the Gauss–Lobatto–Legendre (GLL) nodes. Let $\omega_h(\xi, \eta)$ be a polynomial function of degree p in ξ and η, defined on Ω and

$$ \omega_k^p = \omega_h(\xi_i, \eta_j), \qquad \text{with } k = i(p+1) + j + 1, \tag{32} $$

then its expansion in terms of these polynomials, $\omega_h(\xi, \eta; p)$, is given by

$$ \omega_h(\xi, \eta; p) = \sum_{k=1}^{(p+1)^2} \omega_k^p\, {\epsilon_k}^W(\xi, \eta; p)\,. \tag{33} $$

For this relation between the expansion coefficients and nodal interpolation we denote the polynomials in (30) by *nodal polynomials*. Therefore we set $W_h^p := \text{span}\{{\epsilon_1}^W(\xi, \eta; p), \ldots, {\epsilon_{(p+1)^2}}^W(\xi, \eta; p)\}$. To simplify the notation, the explicit reference to the polynomial degree p will be dropped from the function space, the basis functions, and the coefficient expansion, therefore from here on we will simply use W_h, ${\epsilon_k}^W(\xi, \eta)$, and ω_k.

Basis functions for U_h In a similar way, but combining nodal polynomials with histopolant polynomials, we can construct the polynomial basis functions for U_h on quadrilaterals. Consider the nodal polynomials (22), $l_i^p(\xi)$ of degree p, the histopolant polynomials (24), $e_i^p(\xi)$ of degree $p-1$, the canonical square $\Omega = I \times I \subset \mathbb{R}^2$, and take $\xi, \eta \in I = [-1, 1]$. A set of two-dimensional basis polynomials, $\vec{\epsilon}_k^{\,U}(\xi, \eta; p)$, can be constructed as the tensor product of the one-dimensional basis functions

$$ \begin{aligned} &\vec{\epsilon}_k^{\,U}(\xi, \eta; p) \\ &\quad := \begin{cases} l_i^p(\xi) e_j^p(\eta)\, \vec{e}_\xi & \text{if } k \leq p(p+1), \quad \text{with } i = 0, \ldots, p, \quad j = 1, \ldots, p, \quad k = ip + j\,, \\ e_i^p(\xi) l_j^p(\eta)\, \vec{e}_\eta & \text{if } k > p(p+1), \quad \text{with } i = 1, \ldots, p, \quad j = 0, \ldots, p, \quad k = (p+i-1)(p+1) + j + 1\,. \end{cases} \end{aligned} \tag{34} $$

These polynomials, $\vec{\epsilon}_k^{\,U}(\xi, \eta; p)$, have degree p in ξ and $p-1$ in η if $k \leq p(p+1)$. If $k > p(p+1)$, then the degree in ξ is $p-1$ and the degree in η is p. Using (21) and (25) it follows that these polynomials satisfy, [11,29]

$$ \int_{\eta_{j-1}}^{\eta_j} \vec{\epsilon}_k^{\,U}(\xi_i, \eta; p) \cdot \vec{e}_\xi \, \mathrm{d}\xi = \begin{cases} 1 & \text{if } k = ip + j\,, \\ 0 & \text{if } k \neq ip + j\,, \end{cases} \quad \text{with} \quad \begin{cases} i = 0, \ldots, p, \\ j = 1, \ldots, p, \\ k = 1, \ldots, 2p(p+1)\,, \end{cases} \tag{35} $$

and

$$ \int_{\xi_{i-1}}^{\xi_i} \vec{\epsilon}_k^{\,U}(\xi, \eta_j) \cdot \vec{e}_\eta \, \mathrm{d}\eta = \begin{cases} 1 & \text{if } k = (p+i-1)(p+1) + j + 1\,, \\ 0 & \text{if } k \neq (p+i-1)(p+1) + j + 1\,, \end{cases} \quad \text{with} \quad \begin{cases} i = 1, \ldots, p, \\ j = 0, \ldots, p, \\ k = 1, \ldots, 2p(p+1)\,. \end{cases} \tag{36} $$

Let $\vec{u}_h(\xi, \eta; p)$ be a vector valued polynomial function defined on Ω and

$$u_k^p = \begin{cases} \displaystyle\int_{\eta_{j-1}}^{\eta_j} \vec{u}_h(\xi_i,\eta)\cdot\vec{e}_\xi \,\mathrm{d}\eta & \text{if } k \le p(p+1), \quad \text{with} \quad \begin{cases} i=0,\dots,p, \\ j=1,\dots,p, \\ k=ip+j\,, \end{cases} \\ \displaystyle\int_{\xi_{i-1}}^{\xi_i} \vec{u}_h(\xi,\eta_j)\cdot\vec{e}_\eta \,\mathrm{d}\xi & \text{if } k > p(p+1), \quad \text{with} \quad \begin{cases} i=1,\dots,p, \\ j=0,\dots,p, \\ k=(p+i-1)(p+1)+j+1\,. \end{cases} \end{cases} \tag{37}$$

Then its expansion in terms of these polynomials, $\vec{u}_h(\xi,\eta;p)$, is given by

$$\vec{u}_h(\xi,\eta;p) = \sum_{k=1}^{2p(p+1)} u_k^p \vec{\epsilon}_k^{\,Q}(\xi,\eta;p)\,. \tag{38}$$

The expansion $\vec{u}_h(\xi,\eta;p)$ is a two-dimensional polynomial edge histopolant (interpolates integral values along lines). Since the coefficients of this expansion are edge (or flux) integrals, we denote the polynomials in (34) by *edge polynomials*. We set $U_h := \mathrm{span}\{\epsilon_1^U(\xi,\eta;p),\dots,\epsilon_{2p(p+1)}^U(\xi,\eta;p)\}$. To simplify the notation, the explicit reference to the polynomial degree p will be dropped from the function space, the basis functions, and the expansion coefficients, therefore from here on we will simply use U_h, $\vec{\epsilon}_k^{\,U}(\xi,\eta)$, and u_k.

Basis functions for Q_h Combining histopolant polynomials we can construct the polynomial basis functions for Q_h on a quadrilateral. Consider the canonical interval $I=[-1,1]$, the canonical square $\Omega = I\times I \subset \mathbb{R}^2$, the histopolant polynomials (24), $e_i^p(\xi)$ of degree $p-1$, and take $\xi,\eta\in I$. Then a set of two-dimensional basis polynomials, $\epsilon_k^Q(\xi,\eta;p)$, can be constructed as the tensor product of the one-dimensional ones

$$\epsilon_k^Q(\xi,\eta;p) := e_i^p(\xi)\, e_j^p(\eta), \qquad i,j=1,\dots,p, \quad k=j+(i-1)p\,. \tag{39}$$

These polynomials, $\epsilon_k^Q(\xi,\eta;p)$, have degree $p-1$ in each variable and satisfy, see [11,29],

$$\int_{\xi_{i-1}}^{\xi_i}\int_{\eta_{j-1}}^{\eta_j} \epsilon_k^Q(\xi,\eta;p)\,\mathrm{d}\xi\,\mathrm{d}\eta = \begin{cases} 1 & \text{if } k=(i-1)p+j \\ 0 & \text{if } k\neq(i-1)p+j \end{cases}, \quad i,j=1,\dots,p, \quad k=1,\dots,p^2\,. \tag{40}$$

Here, as before, ξ_i and η_i with $i=0,\dots,p$ are the Gauss–Lobatto–Legendre (GLL) nodes. Let $K_h(\xi,\eta)$ be a polynomial function defined on Ω and $K_k^p = \int_{\xi_{i-1}}^{\xi_i}\int_{\eta_{j-1}}^{\eta_j} K_h(\xi,\eta)\,\mathrm{d}\xi\,\mathrm{d}\eta$ with $k=j+(i-1)p$, then its expansion in terms of these polynomials, $K_h(\xi,\eta;p)$, is given by

$$K_h(\xi,\eta;p) = \sum_{k=1}^{p^2} K_k^p\, \epsilon_k^Q(\xi,\eta;p)\,. \tag{41}$$

For this relation between the expansion coefficients and surface integration we denote the polynomials in (39) by *surface polynomials*. Moreover, these basis polynomials satisfy $\epsilon_k^Q(\xi,\eta;p)\in L^2(\Omega)$. Therefore we set $Q_h^p := \mathrm{span}\{\epsilon_1^Q(\xi,\eta;p),\dots,\epsilon_{p^2}^Q(\xi,\eta;p)\}$. To simplify the notation, the explicit reference to the polynomial degree p will be dropped from the function space, the basis functions, and the coefficient expansion, therefore from here on we will simply use Q_h, $\epsilon_k^Q(\xi,\eta)$, and K_k.

3.2.3. Properties of the basis functions

The first property that can be shown, [11,29], is that if $\omega_h(\xi,\eta)\in W_h$, then $\nabla^\perp\omega_h(\xi,\eta)\in U_h$, where $\omega_h(\xi,\eta)$ is expanded as (33).

$$\begin{aligned} \left(\nabla^\perp\omega_h(\xi,\eta)\right) \;=\; & \sum_{i=0,j=1}^{p} \left(\int_{\eta_{j-1}}^{\eta_j} \nabla^\perp\omega_h(\xi_i,\eta)\cdot\vec{e}_\xi\,\mathrm{d}\eta \right) \vec{\epsilon}_{ip+j}^{\,U}(\xi,\eta) \\ & + \sum_{i=1,j=0}^{p} \left(\int_{\xi_{i-1}}^{\xi_i} \nabla^\perp\omega_h(\xi,\eta_j)\cdot\vec{e}_\eta\,\mathrm{d}\xi \right) \vec{\epsilon}_{(p+i-1)(p+1)+j+1}^{\,U}(\xi,\eta) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=0,\,j=1}^{p} \big(\omega_h(\xi_i,\eta_{j-1}) - \omega_h(\xi_i,\eta_j)\big)\,\vec{\epsilon}^{\,U}_{ip+j}(\xi,\eta) \\
&\qquad + \sum_{i=1,\,j=0}^{p} \big(\omega_h(\xi_i,\eta_j) - \omega_h(\xi_{i-1},\eta_j)\big)\,\vec{\epsilon}^{\,U}_{(p+i-1)(p+1)+j+1}(\xi,\eta) \\
&\overset{(32)}{=} \sum_{i=0,\,j=1}^{p} \big(\omega_{ip+j} - \omega_{ip+j+1}\big)\,\vec{\epsilon}^{\,U}_{ip+j}(\xi,\eta) \\
&\qquad + \sum_{i=1,\,j=0}^{p} \big(\omega_{i(p+1)+j+1} - \omega_{(i-1)(p+1)+j+1}\big)\,\vec{\epsilon}^{\,U}_{(p+i-1)(p+1)+j+1}(\xi,\eta) \\
&= \sum_{k=1}^{2p(p+1)} \sum_{j=1}^{(p+1)^2} \mathsf{E}^{1,0}_{k,j}\omega_j \vec{\epsilon}^{\,U}_k(\xi,\eta)\,, \qquad (42)
\end{aligned}
$$

where $\mathsf{E}^{1,0}_{k,j}$ are the coefficients of the $2p(p+1)\times(p+1)^2$ matrix $\mathbf{E}^{1,0}$ and are defined as

$$
\mathsf{E}^{1,0}_{k,j} := \begin{cases}
-1 & \text{if } j = k+1+(k\,\mathrm{div}(p+1)) \quad \text{and} \quad 1 \le k \le p(p+1)\,, \\
1 & \text{if } j = k+(k\,\mathrm{div}(p+1)) \quad \text{and} \quad 0 \le k \le p(p+1)\,, \\
1 & \text{if } j = k-(p-1)(p+1) \quad \text{and} \quad p(p+1) < k \le 2p(p+1)\,, \\
-1 & \text{if } j = k-p(p+1) \quad \text{and} \quad p(p+1) < k \le 2p(p+1)\,, \\
0 & \text{otherwise}\,.
\end{cases} \qquad (43)
$$

Here $(k\,\mathrm{div}p)$ denotes integer division in which the remainder is discarded.

From (42) it follows that

$$
\omega_h(\xi,\eta) \in W_h \quad \Longrightarrow \quad \nabla^\perp \omega_h(\xi,\eta) \in U_h\,,
$$

which is the finite-dimensional analogue of

$$
\omega(\xi,\eta) \in H(\mathrm{rot};\Omega) \quad \Longrightarrow \quad \nabla^\perp \omega_(\xi,\eta) \in H(\mathrm{div};\Omega)\,.
$$

Or fully discrete, if ω_j are the expansion coefficients of $\omega_h \in W_h$ with respect to the basis $e^W_i(\xi,\eta)$, then $\mathsf{E}^{1,0}_{k,j}\omega_j$ are the expansion coefficients of $\nabla^\perp\omega_h$ in U_h with respect to the basis $\vec{u}^{\,U}_j(\xi,\eta)$.

As a special case we have that

$$
\nabla^\perp \epsilon^W_j = \sum_{k=1}^{2p(p+1)} \mathsf{E}^{1,0}_{k,j}\vec{\epsilon}^{\,U}_k\,, \qquad (44)
$$

and therefore $\nabla^\perp \vec{\epsilon}^{\,W}_j \in U_h$, with $j = 1,\ldots,(p+1)^2$, and these basis functions satisfy (16).

The second property that can be shown, [11,29], is that if $\vec{u}_h(\xi,\eta) \in U_h$ is expanded in terms of edge polynomials, as in (38), then the expansion of $\nabla\cdot\vec{u}_h$ in terms of the surface polynomials, (39), is

$$
\begin{aligned}
\nabla\cdot\vec{u}_h(\xi,\eta) &= \sum_{i,j=1}^{p} \left(\int_{\xi_{i-1}}^{\xi_i}\int_{\eta_{j-1}}^{\eta_j} \nabla\cdot\vec{u}_h(\xi,\eta)\,\mathrm{d}\xi\,\mathrm{d}\eta \right) \epsilon^{\mathrm{Q}}_{j+(i-1)p}(\xi,\eta) \\
&= \sum_{i,j=1}^{p} \left(\int_{\eta_{j-1}}^{\eta_i} \vec{u}_h(\xi_i,\eta)\cdot\vec{e}_\xi\,\mathrm{d}\eta + \int_{\xi_{i-1}}^{\xi_i} \vec{u}_h(\xi,\eta_j)\cdot\vec{e}_\eta\,\mathrm{d}\xi - \int_{\eta_{j-1}}^{\eta_i} \vec{u}_h(\xi_{i-1},\eta)\cdot\vec{e}_\xi\,\mathrm{d}\eta \right. \\
&\qquad \left. - \int_{\xi_{i-1}}^{\xi_i} \vec{u}_h(\xi,\eta_{j-1})\cdot\vec{e}_\eta\,\mathrm{d}\xi \right) \epsilon^{\mathrm{Q}}_{j+(i-1)p}(\xi,\eta)
\end{aligned}
$$

$$\overset{(37)}{=} \sum_{k=1}^{p^2} \left(u_{k+p} + u_{k+p(p+1)+(k\mathrm{div}(p+1))+1} - u_k - u_{k+p(p+1)+(k\mathrm{div}(p+1))} \right) \epsilon_k^Q(\xi,\eta)$$

$$= \sum_{k=1}^{p^2} \sum_{j=1}^{2p(p+1)} \mathsf{E}^{2,1}_{k,j} u_j \epsilon_k^Q(\xi,\eta)\,, \tag{45}$$

where $\mathsf{E}^{2,1}_{k,j}$ are the coefficients of the $p^2 \times 2p(p+1)$ incidence matrix $\mathbf{E}^{2,1}$ defined as

$$\mathsf{E}^{2,1}_{k,j} := \begin{cases} 1 & \text{if } j = k+p\,, \\ 1 & \text{if } j = k+p(p+1)+(k\,\mathrm{div}(p+1))+1\,, \\ -1 & \text{if } j = k\,, \\ -1 & \text{if } j = k+p(p+1)+(k\,\mathrm{div}(p+1))\,, \\ 0 & \text{otherwise}\,. \end{cases} \tag{46}$$

Equation (45) confirms that we have a finite dimensional Hilbert sequence as in (16), because

$$\vec{u}_h(\xi,\eta) \in U_h \quad \Longrightarrow \quad \nabla\cdot\vec{u}_h(\xi,\eta) \in Q_h\,,$$

which is the finite dimensional analogue of

$$\vec{u} \in H(\mathrm{div};\Omega) \quad \Longrightarrow \quad \nabla\cdot\vec{u} \in L^2(\Omega)\,.$$

In terms of the expansion coefficients we have: If u_j are the expansion coefficients of $\vec{u}_h \in U_h$ with respect to the basis $\vec{\epsilon}_j^U$, then the expansion coefficients of $\nabla\cdot\vec{u}_h \in Q_h$ with respect to the basis ϵ_k^Q are given by $\sum_{j=1}^{2p(p+1)} \mathsf{E}^{2,1}_{k,j} u_j$. As a special case we have that

$$\nabla\cdot\vec{\epsilon}_j^{\,U} = \sum_{k=1}^{p^2} \mathsf{E}^{2,1}_{k,j} \epsilon_k^Q\,. \tag{47}$$

If ω_j are the expansion coefficients of $\omega_h \in W_h$, then $\mathsf{E}^{1,0}_{k,j}\omega_j$ are the expansion coefficients of $\nabla^\perp\omega_h \in U_h$. Then $\mathsf{E}^{2,1}_{i,k}\mathsf{E}^{1,0}_{k,j}\omega_j$ are the expansion coefficients of $\nabla\cdot\nabla^\perp\omega_h \in Q_h$. Since $\nabla\cdot\nabla^\perp\omega_h = 0$ for all ω_h (and therefore ω_j) and because ϵ_k^Q forms a basis for Q_h, we need to have that

$$\mathsf{E}^{2,1}_{i,k} \circ \mathsf{E}^{1,0}_{k,j} \equiv 0\,.$$

This is the fully discrete representation of the vector identity $\nabla\cdot\nabla^\perp \equiv 0$.

3.2.4. Some remarks

Note that the degrees of freedom for W_h are associated with the points in the GLL-grid. In a multi-element setting, neighboring elements share the GLL-points on the boundary of the element, thus imposing C^0-continuity for functions in W_h. The degrees of freedom for U_h are the integrals of the normal components (the flux) of $\vec{u}_h$ over the edge in the GLL-grid. In a multi-element setting, neighboring elements share an edge and therefore also the degree of freedom associated with that edge. So only the normal component of $\vec{u}_h \in U_h$ is continuous between elements, thus making U_h a proper subspace of $H(\mathrm{div};\Omega)$. The degrees of freedom for Q_h are the integral values over the two-dimensional surfaces in the GLL-grid. A surface in one spectral element is entirely disjoint from a surface in neighboring elements and therefore, the approximation in Q_h is discontinuous between elements. In that sense Q_h is a proper subspace of $L^2(\Omega)$.

There is a close relation between this mimetic finite element discretization and the traditional C-grid finite difference discretization [6], where rotational moments are located at vertices, normal velocities on edges and pressure and mass variables at cell centers. It is possible to construct an analogue to the D-grid finite difference discretization, for which the tangential and not the normal velocities reside on the edges. Two advantages of the mimetic finite element discretization is the immediate generalization to arbitrary order and the possibility to treat deformed meshes, which we have not addressed in the current work.

There exists for the Hilbert subcomplex of discrete function spaces described above a discrete Helmholtz decomposition [3] of vector fields in U_h into a unique partition of rotational components in W_h and divergent components in Q_h which are orthogonal to one another. This result has been shown previously for the mixed mimetic spectral element function spaces [15].

3.3. Discrete weak formulation

Consider again the domain $\Omega \subset \mathbb{R}^2$ and its tessellation $\mathcal{T}(\Omega)$ consisting of N arbitrary quadrilaterals (possibly curved), Ω_m, with $m = 1, \dots, N$. We assume that all quadrilateral elements Ω_m can be obtained from a map $\Phi_m : (\xi, \eta) \in I^2 \mapsto (x, y) \in \Omega_m$. Then the pushforward $\Phi_{m,*}$ maps functions in the reference element I^2 to functions in the physical element Ω_m, see for example [31,32]. For this reason it suffices to explore the analysis on the reference domain I^2. Additionally, the multi-element case follows the standard approach in finite elements.

Remark 1. If a differential geometry formulation was used, the physical quantities would be represented by differential k-forms and the map $\Phi_m : (\xi, \eta) \in I^2 \mapsto (x, y) \in \Omega_m$ would generate a pullback, Φ_m^*, mapping k-forms in physical space, Ω_m, to k-forms in the reference element, I^2, [15].

The discrete weak formulation can be stated as: given $\Omega = I^2$, the polynomial degree p and a Coriolis term $f_h \in W_h$, for any time $t \in (0, t_F]$ find $\vec{u}_h, \vec{F}_h \in U_h$, $h_h, K_h \in Q_h$, and $q_h \in W_h$ such that

$$
\begin{cases}
\langle \vec{v}_h, \dfrac{\partial \vec{u}_h}{\partial t} \rangle_\Omega + \langle \vec{v}_h, q_h \times \vec{F}_h \rangle_\Omega - \langle \nabla \cdot \vec{v}_h, K_h + g h_h \rangle_\Omega = 0\,, & \forall \vec{v}_h \in U_h, \quad \text{(a)} \\
\langle \sigma_h, \dfrac{\partial h_h}{\partial t} \rangle_\Omega + \langle \sigma_h, \nabla \cdot \vec{F}_h \rangle_\Omega = 0\,, & \forall \sigma_h \in Q_h, \quad \text{(b)} \\
\langle \zeta_h, h_h q_h \rangle_\Omega = -\langle \nabla^\perp \zeta_h, \vec{u}_h \rangle_\Omega + \langle \zeta_h, f_h \rangle_\Omega\,, & \forall \zeta_h \in W_h, \quad \text{(c)} \\
\langle \vec{\varphi}_h, \vec{F}_h \rangle_\Omega = \langle \vec{\varphi}_h, h_h \vec{u}_h \rangle_\Omega\,, & \forall \vec{\varphi}_h \in U_h, \quad \text{(d)} \\
\langle \kappa_h, K_h \rangle_\Omega = \dfrac{1}{2} \langle \kappa_h, \vec{u}_h \cdot \vec{u}_h \rangle_\Omega\,, & \forall \kappa_h \in Q_h. \quad \text{(e)}
\end{cases}
\tag{48}
$$

Using the expansions for $\vec{u}_h$, $\vec{F}_h$, h_h, K_h and q_h in (20), (48) can be written as: find $\boldsymbol{u}, \boldsymbol{F} \in \mathbb{R}^{d_U}$, $\boldsymbol{h}, \boldsymbol{K} \in \mathbb{R}^{d_Q}$, and $\boldsymbol{q} \in \mathbb{R}^{d_W}$ such that

$$
\begin{cases}
\displaystyle\sum_{i=1}^{d_U} \langle \vec{\epsilon}_j^{\,U}, \vec{\epsilon}_i^{\,U} \rangle_\Omega \frac{\mathrm{d}u_i}{\mathrm{d}t} + \sum_{i=1}^{d_U} \langle \vec{\epsilon}_j^{\,U}, q_h \times \vec{\epsilon}_i^{\,U} \rangle_\Omega F_i - \sum_{i,k=1}^{d_Q} \mathsf{E}^{2,1}_{k,j} \langle \epsilon_k^Q, \epsilon_i^Q \rangle_\Omega (K_i + g h_i) = 0\,, & j = 1, \dots, d_U\,, \quad \text{(a)} \\
\displaystyle\sum_{i=1}^{d_Q} \langle \epsilon_j^Q, \epsilon_i^Q \rangle_\Omega \frac{\mathrm{d}h_i}{\mathrm{d}t} + \sum_{i=1}^{d_U} \sum_{i,k=1}^{d_Q} \mathsf{E}^{2,1}_{k,i} \langle \epsilon_j^Q, \epsilon_k^Q \rangle_\Omega F_i = 0\,, & j = 1, \dots, d_Q\,, \quad \text{(b)} \\
\displaystyle\sum_{i=1}^{d_W} \langle \epsilon_j^{\,W}, h_h \epsilon_i^{\,W} \rangle_\Omega q_i = -\sum_{i,k=1}^{d_U} \mathsf{E}^{1,0}_{k,j} \langle \vec{\epsilon}_k^{\,U}, \vec{\epsilon}_i^{\,U} \rangle_\Omega u_i + \sum_{i=1}^{d_W} \langle \epsilon_j^{\,W}, \epsilon_i^{\,W} \rangle_\Omega f_i\,, & j = 1, \dots, d_W\,, \quad \text{(c)} \\
\displaystyle\sum_{i=1}^{d_U} \langle \vec{\epsilon}_j^{\,U}, \vec{\epsilon}_i^{\,U} \rangle_\Omega F_i = \sum_{i=1}^{d_U} \langle \vec{\epsilon}_j^{\,U}, h_h \vec{\epsilon}_i^{\,U} \rangle_\Omega u_i\,, & j = 1, \dots, d_U\,, \quad \text{(d)} \\
\displaystyle\sum_{i=1}^{d_Q} \langle \epsilon_j^Q, \epsilon_i^Q \rangle_\Omega K_i = \frac{1}{2} \sum_{i=1}^{d_U} \langle \epsilon_j^Q, \vec{u}_h \cdot \vec{\epsilon}_i^{\,U} \rangle_\Omega u_i\,, & j = 1, \dots, d_Q\,, \quad \text{(e)}
\end{cases}
\tag{49}
$$

with $\boldsymbol{u} := [u_1, \dots, u_{d_U}]^\top$, $\boldsymbol{F} := [F_1, \dots, F_{d_U}]^\top$, $\boldsymbol{h} := [h_1, \dots, h_{d_Q}]^\top$, $\boldsymbol{K} := [K_1, \dots, K_{d_Q}]^\top$ and $\boldsymbol{q} := [q_1, \dots, q_{d_W}]^\top$.

Using matrix notation, (49) can be expressed more compactly as

$$
\begin{cases}
\mathsf{U} \dfrac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} + \mathsf{U}^q \boldsymbol{F} - \left(\mathsf{E}^{2,1} \right)^\top \mathsf{Q} \left(\boldsymbol{K} + g \boldsymbol{h} \right) = 0\,, & \text{(a)} \\
\mathsf{Q} \dfrac{\mathrm{d}\boldsymbol{h}}{\mathrm{d}t} + \mathsf{Q} \mathsf{E}^{2,1} \boldsymbol{F} = 0\,, & \text{(b)} \\
\mathsf{W}^h \boldsymbol{q} = -\left(\mathsf{E}^{1,0} \right)^\top \mathsf{U} \boldsymbol{u} + \mathsf{W} \boldsymbol{f}\,, & \text{(c)} \\
\mathsf{U} \boldsymbol{F} = \mathsf{U}^h \boldsymbol{u}\,, & \text{(d)} \\
\mathsf{Q} \boldsymbol{K} = \dfrac{1}{2} \mathsf{U}^u \boldsymbol{u}\,. & \text{(e)}
\end{cases}
\tag{50}
$$

The coefficients of the matrices U, Q, and W are given by

$$
\mathsf{U}_{ij} := \langle \vec{\epsilon}_j^{\,U}, \vec{\epsilon}_i^{\,U} \rangle_\Omega, \quad \mathsf{Q}_{ij} := \langle \epsilon_j^Q, \epsilon_i^Q \rangle_\Omega, \quad \text{and} \quad \mathsf{W}_{ij} := \langle \epsilon_j^{\,W}, \epsilon_i^{\,W} \rangle_\Omega\,. \tag{51}
$$

Similarly, the coefficients of the matrices $\mathbf{U}^q$, $\mathbf{U}^h$, $\mathbf{U}^u$, and $\mathbf{W}^h$ are given by

$$\mathsf{U}^q_{ij} := \langle \vec{\epsilon}_i^{\,U}, q_h \times \vec{\epsilon}_j^{\,U} \rangle_\Omega, \quad \mathsf{U}^h_{ij} := \langle \vec{\epsilon}_i^{\,U}, h_h \vec{\epsilon}_j^{\,U} \rangle_\Omega, \quad \mathsf{U}^u_{ij} := \langle \epsilon_i^{\,Q}, \vec{u}_h \cdot \vec{\epsilon}_j^{\,U} \rangle_\Omega, \quad \text{and} \quad \mathsf{W}^h_{ij} := \langle \epsilon_i^{\,W}, h_h \epsilon_j^{\,W} \rangle_\Omega . \tag{52}$$

The mass matrix $\mathbf{Q}$ may be canceled entirely from (50b), such that the continuity equation holds in the strong form.

$$\frac{\mathrm{d}\boldsymbol{h}}{\mathrm{d}t} + \mathbf{E}^{2,1}\boldsymbol{F} = 0 .$$

This reflects the fact that the divergence theorem is satisfied point wise as given by (47). We further note that (48e), (49e) and (50e) are redundant since $\vec{u}_h \cdot \vec{u}_h$ may be directly projected onto $\vec{v}_h$ in (48a), (49a) and (50a) respectively. We have preserved this expression however as it makes explicit that $K_h \in Q_h$ is a scalar quantity in the same discrete function space as h_h, and is efficient to compute since it may be done so individually for each element since Q_h is discontinuous across element boundaries.

4. Discrete conservation properties

4.1. Conservation of mass

Conservation of mass is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega h \,\mathrm{d}\Omega = 0 . \tag{53}$$

At the discrete level we have that $h_h \in Q_h$, therefore $h_h = \sum_j h_j \epsilon_j^Q$. Recalling that the basis functions ϵ_j^Q satisfy (40), we have that

$$\int_\Omega h_h \,\mathrm{d}\Omega = \sum_j h_j = \mathbf{1}^\mathsf{T}\boldsymbol{h} , \tag{54}$$

with $\mathbf{1}^\mathsf{T} = [1, \ldots, 1] \in \mathbb{R}^{dQ}$. Eliminating $\mathbf{Q}$ in (50b) results in

$$\frac{\mathrm{d}\boldsymbol{h}}{\mathrm{d}t} = -\mathbf{E}^{2,1}\boldsymbol{F} . \tag{55}$$

Combining (53) with (54) and (55), results in conservation of mass at the algebraic level

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega h_h \,\mathrm{d}\Omega \overset{(54)}{=} \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{1}^\mathsf{T}\boldsymbol{h}\right) = \mathbf{1}^\mathsf{T}\frac{\mathrm{d}\boldsymbol{h}}{\mathrm{d}t} \overset{(55)}{=} -\mathbf{1}^\mathsf{T}\mathbf{E}^{2,1}\boldsymbol{F} = 0 . \tag{56}$$

The last identity results from the telescoping property of the incidence matrix on periodic domains, $\mathbf{1}^\mathsf{T}\mathbf{E}^{2,1} = 0$.

4.2. Conservation of vorticity

Conservation of total vorticity is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \omega \,\mathrm{d}\Omega = 0 . \tag{57}$$

At the discrete level, vorticity is defined as: for any time $t \in (0, t_F]$, given $\vec{u}_h \in U_h$, find $\omega_h \in W_h$ such that

$$\langle \varsigma_h, \omega_h \rangle_\Omega = -\langle \nabla^\perp \varsigma_h, \vec{u}_h \rangle_\Omega , \qquad \forall \varsigma_h \in W_h . \tag{58}$$

At the algebraic level (58) becomes

$$\mathbf{W}\boldsymbol{\omega} = -\left(\mathbf{E}^{1,0}\right)^\mathsf{T}\mathbf{U}\boldsymbol{u} , \tag{59}$$

and its time derivative

$$\mathbf{W}\frac{\mathrm{d}\boldsymbol{\omega}}{\mathrm{d}t} = -\left(\mathbf{E}^{1,0}\right)^\mathsf{T}\mathbf{U}\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} . \tag{60}$$

Since $\omega_h \in W_h$, we have that $\omega_h = \sum_j \omega_k \epsilon_j^W$. Recalling that $\sum_j \epsilon_j^W = 1$, we have

$$\int_\Omega \omega_h \,\mathrm{d}\Omega = \int_\Omega \sum_j \omega_j \epsilon_j^W \,\mathrm{d}\Omega = \int_\Omega \sum_j \left(\sum_k \epsilon_k^W\right) \omega_j \epsilon_j^W \,\mathrm{d}\Omega = \sum_{j,k} \omega_j \int_\Omega \epsilon_k^W \epsilon_j^W \,\mathrm{d}\Omega = \mathbf{1}^\mathsf{T}\mathbf{W}\boldsymbol{\omega}\,, \tag{61}$$

with $\mathbf{1}^\mathsf{T} = [1, \ldots, 1] \in \mathbb{R}^{d_W}$. Therefore (57) becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \omega_h \,\mathrm{d}\Omega \overset{(61)}{=} \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{1}^\mathsf{T}\mathbf{W}\boldsymbol{\omega}\right) = \mathbf{1}^\mathsf{T}\mathbf{W}\frac{\mathrm{d}\boldsymbol{\omega}}{\mathrm{d}t}\,, \tag{62}$$

Multiplying both sides of (60) by $\mathbf{1}^\mathsf{T}$ and combining with (62) gives the time conservation of total vorticity at the algebraic level

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \omega_h \,\mathrm{d}\Omega = \mathbf{1}^\mathsf{T}\mathbf{W}\frac{\mathrm{d}\boldsymbol{\omega}}{\mathrm{d}t} = -\mathbf{1}^\mathsf{T}\left(\mathbf{E}^{1,0}\right)^\mathsf{T}\mathbf{U}\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} = 0\,. \tag{63}$$

The last identity again results from the telescoping property of the incidence matrix on periodic domains, $\mathbf{1}^\mathsf{T}\left(\mathbf{E}^{1,0}\right)^\mathsf{T} = \left(\mathbf{E}^{1,0}\mathbf{1}\right)^\mathsf{T} = 0$.

Note that the conservation of vorticity is satisfied irrespective of how q_h is constructed in (49a). As such vorticity conservation is preserved in the event that the anticipated potential vorticity method [33] is used to truncate the potential enstrophy cascade, since this involves removing some downstream *anticipated* potential vorticity from the right hand side in order to introduce some dispersion into the vorticity advection equation. We construct this anticipated potential vorticity $\hat{q}_h$ in the weak form as

$$\langle \varsigma_h, \hat{q}_h\rangle_\Omega = \langle \varsigma_h, q_h\rangle_\Omega - \Delta\tau\langle \varsigma_h, \vec{u}_h \times \nabla^\perp q_h\rangle_\Omega \tag{64}$$

where $\Delta\tau$ is some time scale associated with the evaluation of the downstream potential vorticity.

4.3. Conservation of energy

For physical problems governed by the shallow water equations, the total energy $\mathcal{E}$, is given by the sum of kinetic and potential energies

$$\mathcal{E} := \int_\Omega \left(Kh + \frac{1}{2}gh^2\right)\mathrm{d}\Omega = \langle h, K\rangle_\Omega + \frac{1}{2}\langle h, gh\rangle_\Omega\,. \tag{65}$$

Conservation of total energy is then expressed as

$$\frac{\mathrm{d}\mathcal{E}}{\mathrm{d}t} = \langle h, \frac{\partial K}{\partial t}\rangle_\Omega + \langle \frac{\partial h}{\partial t}, K\rangle_\Omega + \langle \frac{\partial h}{\partial t}, gh\rangle_\Omega\,. \tag{66}$$

For the discrete variables, the time variation of energy takes a similar form

$$\frac{\mathrm{d}\mathcal{E}_h}{\mathrm{d}t} = \langle h_h, \frac{\partial K_h}{\partial t}\rangle_\Omega + \langle \frac{\partial h_h}{\partial t}, K_h\rangle_\Omega + \langle \frac{\partial h_h}{\partial t}, gh_h\rangle_\Omega\,, \tag{67}$$

which can be written at the algebraic level as

$$\frac{\mathrm{d}\mathcal{E}_h}{\mathrm{d}t} = \boldsymbol{h}^\mathsf{T}\mathbf{Q}\frac{\mathrm{d}\boldsymbol{K}}{\mathrm{d}t} + \frac{\mathrm{d}\boldsymbol{h}^\mathsf{T}}{\mathrm{d}t}\mathbf{Q}\boldsymbol{K} + g\frac{\mathrm{d}\boldsymbol{h}^\mathsf{T}}{\mathrm{d}t}\mathbf{Q}\boldsymbol{h}\,. \tag{68}$$

From (50e) we have that

$$\mathbf{Q}\boldsymbol{K} = \frac{1}{2}\mathbf{U}^u\boldsymbol{u}\,, \tag{69}$$

and therefore its time derivative is given by

$$\mathbf{Q}\frac{\mathrm{d}\boldsymbol{K}}{\mathrm{d}t} = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{U}^u\boldsymbol{u}\right)\,. \tag{70}$$

If we now recall the definition of $\mathbf{U}^u$, (52), we can rewrite the right hand side of (70) as

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{U}^u\boldsymbol{u}\right) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{j=1}^{d_U}\mathsf{U}_{ij}^u u_j\right) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{k,j=1}^{d_U} u_k\langle \epsilon_i^Q, \vec{\epsilon}_k^{\,U}\cdot\vec{\epsilon}_j^{\,U}\rangle_\Omega u_j\right) = \sum_{j=1}^{d_U}\mathsf{U}_{ij}^u\frac{\mathrm{d}u_j}{\mathrm{d}t} = \mathbf{U}^u\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t}\,. \tag{71}$$

Substituting (71) into (68) yields

$$\frac{\mathrm{d}\mathcal{E}_h}{\mathrm{d}t} = \boldsymbol{h}^\mathsf{T}\mathbf{U}^u\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} + \frac{\mathrm{d}\boldsymbol{h}^\mathsf{T}}{\mathrm{d}t}\mathbf{Q}\boldsymbol{K} + g\frac{\mathrm{d}\boldsymbol{h}^\mathsf{T}}{\mathrm{d}t}\mathbf{Q}\boldsymbol{h}\,. \tag{72}$$

Noting that

$$\boldsymbol{h}^\mathsf{T}\mathbf{U}^u = \sum_{i=1}^{d_Q} h_i \mathsf{U}^u_{ij} = \sum_{i,k=1}^{d_Q,d_U} h_i \langle \epsilon_i^Q, \vec{\epsilon}_k^{\,U}\cdot\vec{\epsilon}_j^{\,U}\rangle_\Omega u_k = \sum_{i,k=1}^{d_Q,d_U} h_i \langle \vec{\epsilon}_k^{\,U}, \epsilon_i^Q\vec{\epsilon}_j^{\,U}\rangle_\Omega u_k = \sum_{k=1}^{d_U} u_k \mathsf{U}^h_{kj} = \boldsymbol{u}^\mathsf{T}\mathbf{U}^h\,, \tag{73}$$

we can rewrite (72) as

$$\frac{\mathrm{d}\mathcal{E}_h}{\mathrm{d}t} = \boldsymbol{u}^\mathsf{T}\left(\mathbf{U}^h\right)^\mathsf{T}\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} + \frac{\mathrm{d}\boldsymbol{h}^\mathsf{T}}{\mathrm{d}t}\mathbf{Q}\boldsymbol{K} + g\frac{\mathrm{d}\boldsymbol{h}^\mathsf{T}}{\mathrm{d}t}\mathbf{Q}\boldsymbol{h}\,, \tag{74}$$

since $\mathbf{U}^h$ is a symmetric matrix. If we now use (50d) on the first term on the right hand side of (74) and (50b) on the other two terms, we obtain

$$\frac{\mathrm{d}\mathcal{E}_h}{\mathrm{d}t} = \boldsymbol{F}^\mathsf{T}\mathbf{U}\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} - \boldsymbol{F}^\mathsf{T}\left(\mathbf{E}^{2,1}\right)^\mathsf{T}\mathbf{Q}(\boldsymbol{K}+g\boldsymbol{h})\,, \tag{75}$$

again because $\mathbf{U}$ and $\mathbf{Q}$ are symmetric matrices.

Finally, substituting (50a) into (75) yields

$$\frac{\mathrm{d}\mathcal{E}_h}{\mathrm{d}t} = -\boldsymbol{F}^\mathsf{T}\mathbf{U}^q\boldsymbol{F} + \boldsymbol{F}^\mathsf{T}\left(\mathbf{E}^{2,1}\right)^\mathsf{T}\mathbf{Q}(\boldsymbol{K}+g\boldsymbol{h}) - \boldsymbol{F}^\mathsf{T}\left(\mathbf{E}^{2,1}\right)^\mathsf{T}\mathbf{Q}(\boldsymbol{K}+g\boldsymbol{h}) = 0\,, \tag{76}$$

because the last two terms directly cancel each other, and the first one is zero since $\mathbf{U}^q$ is a skew-symmetric matrix. The necessary conditions for energy conservation in space, namely that $\mathbf{U}^q$ is skew-symmetric and that the gradient and divergence operators are anti-adjoints, as given in (14a) and applied in (48a) are more fundamentally derived from the structure of the Poisson bracket used to construct the rotating shallow water equations in Hamiltonian form [7,34]. Note that (71) requires that the chain rule holds in the discrete form for $\vec{u}_h$ with respect to t. The conservation of energy therefore is limited to the truncation error in the time stepping scheme.

4.4. Conservation of potential enstrophy

Potential enstrophy, $\mathcal{Q}$, is defined as, [4],

$$\mathcal{Q} := \int_\Omega hq^2\,\mathrm{d}\Omega = \langle hq, q\rangle_\Omega\,. \tag{77}$$

Conservation of potential enstrophy states that

$$\frac{\mathrm{d}\mathcal{Q}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\langle hq, q\rangle_\Omega = \langle\frac{\partial h}{\partial t}q, q\rangle_\Omega + 2\langle hq, \frac{\partial q}{\partial t}\rangle_\Omega = 0\,. \tag{78}$$

When considering the discrete variables, the time variation of potential enstrophy is expressed in an analogous way

$$\frac{\mathrm{d}\mathcal{Q}_h}{\mathrm{d}t} = \langle\frac{\partial h_h}{\partial t}q_h, q_h\rangle_\Omega + 2\langle h_h q_h, \frac{\partial q_h}{\partial t}\rangle_\Omega\,, \tag{79}$$

which at the algebraic level becomes

$$\frac{\mathrm{d}\mathcal{Q}_h}{\mathrm{d}t} = \frac{\mathrm{d}\boldsymbol{h}^\mathsf{T}}{\mathrm{d}t}\mathbf{W}^q\boldsymbol{q} + 2\boldsymbol{h}^\mathsf{T}\mathbf{W}^q\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}t}\,, \tag{80}$$

with

$$\mathbf{W}^q_{ij} := \langle \epsilon_i^W, q_h\epsilon_j^Q\rangle_\Omega\,. \tag{81}$$

Noting that

$$\left(\mathbf{W}^q\boldsymbol{h}\right)_i = \sum_{k=1}^{d_Q}\mathbf{W}^q_{ik}h_k = \sum_{k=1}^{d_Q}\langle\epsilon_i^W, q_h\epsilon_k^Q\rangle_\Omega h_k = \sum_{k=1}^{d_W}\langle\epsilon_i^W, h_h\epsilon_k^W\rangle_\Omega q_k = \sum_{k=1}^{d_W}\mathbf{W}^h_{ik}q_k = \left(\mathbf{W}^h\boldsymbol{q}\right)_i\,, \tag{82}$$

it is possible to rewrite (80) as

$$\frac{\mathrm{d}\mathcal{Q}_h}{\mathrm{d}t} = \frac{\mathrm{d}\boldsymbol{h}^\mathsf{T}}{\mathrm{d}t}\left(\mathbf{W}^q\right)^\mathsf{T}\boldsymbol{q} + 2\boldsymbol{q}^\mathsf{T}\mathbf{W}^h\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}t}\,. \tag{83}$$

Furthermore, if we now observe that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{W}^h\boldsymbol{q}\right)=\frac{\mathrm{d}\mathbf{W}^h}{\mathrm{d}t}\boldsymbol{q}+\mathbf{W}^h\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}t}=\mathbf{W}^q\frac{\mathrm{d}\boldsymbol{h}}{\mathrm{d}t}+\mathbf{W}^h\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}t}\,, \tag{84}$$

equation (83) becomes

$$\frac{\mathrm{d}\mathcal{Q}_h}{\mathrm{d}t}=\frac{\mathrm{d}\boldsymbol{h}^\mathsf{T}}{\mathrm{d}t}\left(\mathbf{W}^q\right)^\mathsf{T}\boldsymbol{q}-2\frac{\mathrm{d}\boldsymbol{h}^\mathsf{T}}{\mathrm{d}t}\left(\mathbf{W}^q\right)^\mathsf{T}\boldsymbol{q}+2\boldsymbol{q}^\mathsf{T}\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{W}^h\boldsymbol{q}\right)=-\frac{\mathrm{d}\boldsymbol{h}^\mathsf{T}}{\mathrm{d}t}\left(\mathbf{W}^q\right)^\mathsf{T}\boldsymbol{q}+2\boldsymbol{q}^\mathsf{T}\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{W}^h\boldsymbol{q}\right)\,. \tag{85}$$

If we now assume that $\frac{\mathrm{d}\boldsymbol{f}}{\mathrm{d}t}=0$, using (50c) we can rewrite (85) as

$$\frac{\mathrm{d}\mathcal{Q}_h}{\mathrm{d}t}=-\frac{\mathrm{d}\boldsymbol{h}^\mathsf{T}}{\mathrm{d}t}\left(\mathbf{W}^q\right)^\mathsf{T}\boldsymbol{q}-2\boldsymbol{q}^\mathsf{T}\left(\mathbf{E}^{1,0}\right)^\mathsf{T}\mathbf{U}\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t}\,. \tag{86}$$

Substituting (50a) into the last term on the right hand side yields

$$\frac{\mathrm{d}\mathcal{Q}_h}{\mathrm{d}t}=-\frac{\mathrm{d}\boldsymbol{h}^\mathsf{T}}{\mathrm{d}t}\left(\mathbf{W}^q\right)^\mathsf{T}\boldsymbol{q}+2\boldsymbol{q}^\mathsf{T}\left(\mathbf{E}^{1,0}\right)^\mathsf{T}\mathbf{U}^q\boldsymbol{F}-2\boldsymbol{q}^\mathsf{T}\left(\mathbf{E}^{1,0}\right)^\mathsf{T}\left(\mathbf{E}^{2,1}\right)^\mathsf{T}\mathbf{Q}\left(\boldsymbol{K}+g\boldsymbol{h}\right)\,. \tag{87}$$

The last term on the right hand side cancels out because $\mathbf{E}^{2,1}\mathbf{E}^{1,0}=\mathbf{0}$, therefore

$$\frac{\mathrm{d}\mathcal{Q}_h}{\mathrm{d}t}=-\frac{\mathrm{d}\boldsymbol{h}^\mathsf{T}}{\mathrm{d}t}\left(\mathbf{W}^q\right)^\mathsf{T}\boldsymbol{q}+2\boldsymbol{q}^\mathsf{T}\left(\mathbf{E}^{1,0}\right)^\mathsf{T}\mathbf{U}^q\boldsymbol{F}\,. \tag{88}$$

Since $\frac{\mathrm{d}\mathcal{Q}_h}{\mathrm{d}t}$ is a scalar, we have that

$$\frac{\mathrm{d}\boldsymbol{h}^\mathsf{T}}{\mathrm{d}t}\left(\mathbf{W}^q\right)^\mathsf{T}\boldsymbol{q}=\boldsymbol{q}^\mathsf{T}\mathbf{W}^q\frac{\mathrm{d}\boldsymbol{h}}{\mathrm{d}t}\,, \tag{89}$$

and substituting into (88), results in

$$\frac{\mathrm{d}\mathcal{Q}_h}{\mathrm{d}t}=-\boldsymbol{q}^\mathsf{T}\mathbf{W}^q\frac{\mathrm{d}\boldsymbol{h}}{\mathrm{d}t}+2\boldsymbol{q}^\mathsf{T}\left(\mathbf{E}^{1,0}\right)^\mathsf{T}\mathbf{U}^q\boldsymbol{F}\,. \tag{90}$$

Moreover, using (50b) we can write

$$\frac{\mathrm{d}\mathcal{Q}_h}{\mathrm{d}t}=-\boldsymbol{q}^\mathsf{T}\mathbf{W}^q\mathbf{E}^{2,1}\boldsymbol{F}+2\boldsymbol{q}^\mathsf{T}\left(\mathbf{E}^{1,0}\right)^\mathsf{T}\mathbf{U}^q\boldsymbol{F}\,. \tag{91}$$

Now, noting that

$$\left[\left(\mathbf{U}^q\right)^\mathsf{T}\mathbf{E}^{1,0}\boldsymbol{q}\right]_j=\sum_{i,k=1}^{d_W,d_U}\langle\vec{\epsilon}_k^{\,U},q_h\times\vec{\epsilon}_j^{\,U}\rangle_\Omega\mathbf{E}^{1,0}_{k,i}q_i=-\sum_{i,k=1}^{d_W,d_U}\langle\vec{\epsilon}_j^{\,U},q_h\times\vec{\epsilon}_k^{\,U}\rangle_\Omega\mathbf{E}^{1,0}_{k,i}q_i=-\langle\vec{\epsilon}_j^{\,U},q_h\times\left(\sum_{i,k=1}^{d_W,d_U}\mathbf{E}^{1,0}_{k,i}q_i\vec{\epsilon}_k^{\,U}\right)\rangle_\Omega\,. \tag{92}$$

Recalling (42), we can rewrite the last term of (92) as

$$\sum_{i,k=1}^{d_W,d_U}\mathbf{E}^{1,0}_{k,i}q_i\vec{\epsilon}_k^{\,U}=\nabla^\perp q_h\,, \tag{93}$$

and consequently (92) becomes

$$\left[\left(\mathbf{U}^q\right)^\mathsf{T}\mathbf{E}^{1,0}\boldsymbol{q}\right]_j=-\langle\vec{\epsilon}_j^{\,U},q_h\times\nabla^\perp q_h\rangle_\Omega\,. \tag{94}$$

The right hand side of this equation can be transformed in the following way

$$-\langle\vec{\epsilon}_j^{\,U},q_h\times\nabla^\perp q_h\rangle_\Omega\overset{*}{=}\frac{1}{2}\langle\vec{\epsilon}_j^{\,U},\nabla(q_h\cdot q_h)\rangle_\Omega\overset{*}{=}\frac{1}{2}\langle\nabla\cdot\vec{\epsilon}_j^{\,U},q_h\cdot q_h\rangle_\Omega\,. \tag{95}$$

Therefore we have

$$\left[\left(\mathbf{U}^q\right)^\mathsf{T}\mathbf{E}^{1,0}\boldsymbol{q}\right]_j=\frac{1}{2}\langle\nabla\cdot\vec{\epsilon}_j^{\,U},q_h\cdot q_h\rangle_\Omega\,. \tag{96}$$

It is important to note that the two middle identities, with $\overset{*}{=}$, in (95) are only valid if exact integration is used in the inner products. If approximate integration is used, these identities are only guaranteed to be approximately valid. Now, the right hand term in (96) can be written as

$$\langle \nabla \cdot \vec{\epsilon}_j^{\,U}, q_h \cdot q_h \rangle_\Omega = \sum_{k=1}^{d_Q} \mathbf{E}^{2,1}_{k,j} \langle \epsilon_k^Q, q_h \cdot q_h \rangle_\Omega = \sum_{k=1}^{d_Q} \mathbf{E}^{2,1}_{k,j} \langle \epsilon_k^Q q_h, \epsilon_i^W \rangle_\Omega q_i = \left[\left(\mathbf{E}^{2,1} \right)^\intercal \left(\mathbf{W}^q \right)^\intercal \boldsymbol{q} \right]_j . \tag{97}$$

Therefore, (91) becomes

$$\frac{\mathrm{d}\mathcal{Q}_h}{\mathrm{d}t} = -\boldsymbol{q}^\intercal \mathbf{W}^q \mathbf{E}^{2,1} \boldsymbol{F} + \boldsymbol{q}^\intercal \mathbf{W}^q \mathbf{E}^{2,1} \boldsymbol{F} = 0 , \tag{98}$$

thus proving conservation of potential enstrophy at the discrete level.

As for energy conservation, potential enstrophy conservation requires that the chain rule holds for differentiation in time, as applied in (78). However unlike the energy conservation argument, potential enstrophy conservation also requires that this holds for spatial differentiation, as given in (95). Potential enstrophy conservation is therefore also subject to exact spatial integration in the weak form.

4.5. Geostrophic balance

The existence of stationary geostrophic modes derives from the linearization of (1) as [3]

$$\begin{cases} \dfrac{\partial \vec{u}}{\partial t} + f \times \vec{u} + g \nabla h = 0 , & \text{in } \Omega \times (0, t_F] , \quad \text{(a)} \\ \dfrac{\partial h}{\partial t} + H \nabla \cdot \vec{u} = 0 , & \text{in } \Omega \times (0, t_F] , \quad \text{(b)} \end{cases} \tag{99}$$

where H is the mean depth of the fluid layer. If we assume a stationary solution then we have the balanced system

$$f \times \vec{u} + g \nabla h = 0 \qquad \nabla \cdot \vec{u} = 0 \tag{100}$$

In order to show that this balance is satisfied for our formulation we begin by introducing the stream function as $\vec{u} = \nabla^\perp \psi$. Note that this identity is satisfied in the strong form for our discrete formulation via (44). Substituting this into (100) gives

$$-f \nabla \psi + g \nabla h = 0 \qquad \nabla \cdot \vec{u} = 0 \tag{101}$$

To show that this relation is satisfied at the algebraic level we begin by taking the linearized form of (50a), (50b) as

$$\mathbf{U} \frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} + \mathbf{U}^f \boldsymbol{u} - g \left(\mathbf{E}^{2,1} \right)^\top \mathbf{Q} \boldsymbol{h} = 0 \tag{102}$$

$$\frac{\mathrm{d}\boldsymbol{h}}{\mathrm{d}t} + \mathbf{E}^{2,1} \boldsymbol{u} = 0 \tag{103}$$

where $\mathsf{U}^f_{ij} := \langle \vec{\epsilon}_i^{\,U}, f_h \times \vec{\epsilon}_j^{\,U} \rangle_\Omega$. Note that we have eliminated the $\mathbf{Q}$ matrix in the linearized continuity equation (103) as done in (55). Applying the stream function identity to (102) and then the adjoint relation (14a) gives

$$\mathbf{U} \frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} + f \left(\mathbf{E}^{2,1} \right)^\top \mathbf{Q}^w \boldsymbol{\psi} - g \left(\mathbf{E}^{2,1} \right)^\top \mathbf{Q} \boldsymbol{h} = 0 \tag{104}$$

where $\mathsf{Q}^w_{ij} := \langle \epsilon_i^Q, \epsilon_j^W \rangle_\Omega$. The last two terms of (104) constitute the weak form of (101) and so balance up to truncation error of the interpolating polynomials, ϵ_i^W, ϵ_i^Q is achieved such that $\boldsymbol{u}$ is approximately stationary. Since (103) holds point wise, the absence of divergence leads to a stationary fluid depth $\boldsymbol{h}$, thus demonstrating geostrophic balance.

In the following section we will illustrate for a simple divergence free linear test case that the errors in geostrophic balance remain bounded and convergent with temporal and spatial resolution.

5. Results

5.1. Convergence

In order to first validate our mixed mimetic spectral element formulation we inspect the L_2 norm error convergence of the various operators. For this we compare the three diagnostic equations (50c)–(50e) to the analytic solutions for a specified velocity and depth field, as derived from a stream function solution of

$$\psi = 0.1 \cos(x - \pi) \cos(y - \pi) \qquad \Omega = (0, 2\pi] \times (0, 2\pi], \tag{105}$$

where $\vec{u} = \nabla^\perp \psi$, $h = (f/g)\psi + H$ via geostrophic balance, $f = g = 8.0$ and $H = 0.2$ is the constant of integration in the geostrophic balance relation. In each case exact spatial integration is applied. By demonstrating both the algebraic convergence of the errors for constant polynomial degree with decreasing mesh size and the spectral convergence with

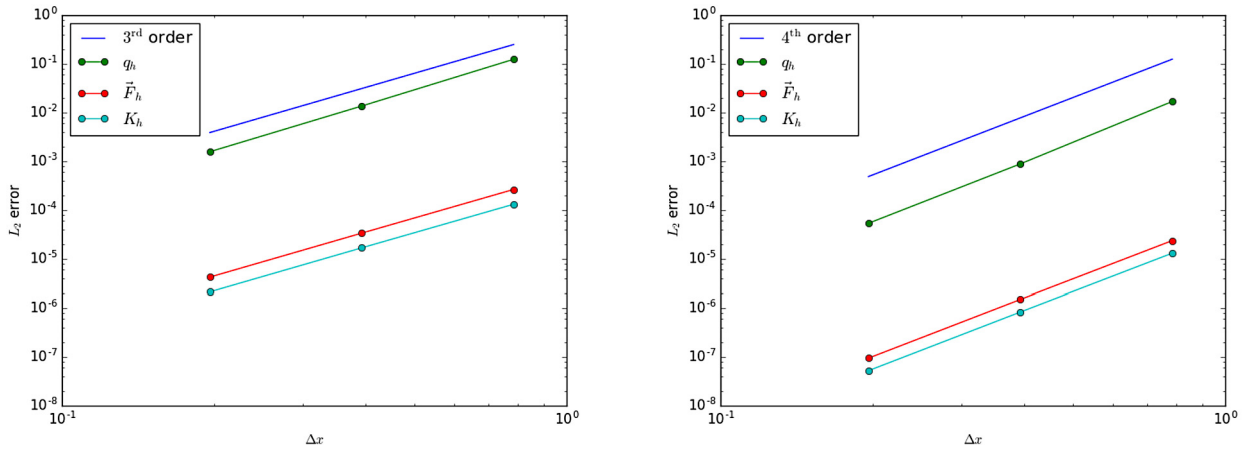

Fig. 3. Algebraic convergence for the 3rd order (left) and 4th order (right) basis functions for q_h, $\vec{F}_h$ and K_h. Blue slopes show the theoretical convergence rates for 3rd and 4th order respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

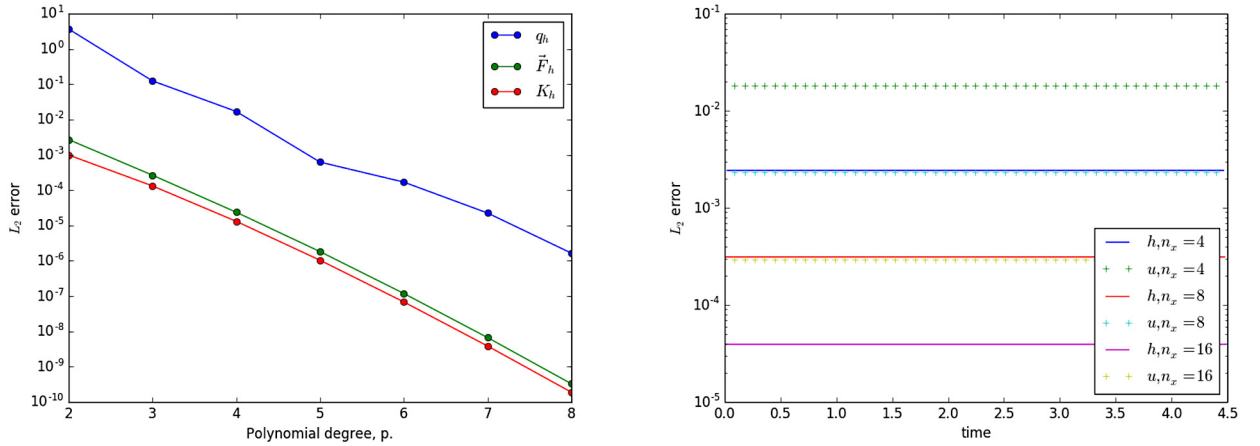

Fig. 4. Spectral convergence of the diagnostic equations with polynomial degree (left) and convergence of errors for the linear geostrophic balance test (right).

increasing polynomial degree and constant mesh size of the solution for each of the diagnostic equations we aim to validate the basis functions for each of the W_h, U_h and Q_h function spaces.

As shown in Fig. 3, $\vec{F}_h$ and K_h satisfy their expected rates of convergence for both 3rd and 4th order accurate basis representations. However $q_h \in W_h$ should in principle converge at one order higher, since its polynomial expansion is one degree higher. This higher rate of convergence is not observed from Fig. 3. The fact that q_h converges at the same rate as $\vec{F}_k$ and K_h may be possibly be attributed to the fact that h_h as used in the left hand side of diagnostic equation (50c) is discontinuous across element boundaries and so may be projecting spurious gradients onto q_h. Convergence studies of q_h using the $||q_h||^2_{H(\mathrm{rot})}$ norm with an analytic (continuous) depth field (not shown), which exhibit the correct rate of convergence, would seem to suggest this.

Fig. 4 shows the spectral convergence of errors for constant mesh size and increasing polynomial degree for each of the W_h, U_h and Q_h tensor product element diagnostic equations.

5.2. Linearized geostrophic balance

As a second test the linearized system (99a), (99b) is solved in the discrete form for a divergence free velocity field and a depth field derived from geostrophic balance via the stream function solution (105). The discrete problem for this system is equivalent to (50) with $\boldsymbol{q} = \boldsymbol{f}$, $\boldsymbol{K} = 0$ and $\boldsymbol{F} = H\boldsymbol{u}$. For this and all subsequent tests we use an explicit second order Runge–Kutta scheme, given for the continuous form (1) as

$$y' = y^n - 0.5\Delta t G(y^n), \qquad y^{n+1} = y^n - \Delta t G(y')$$

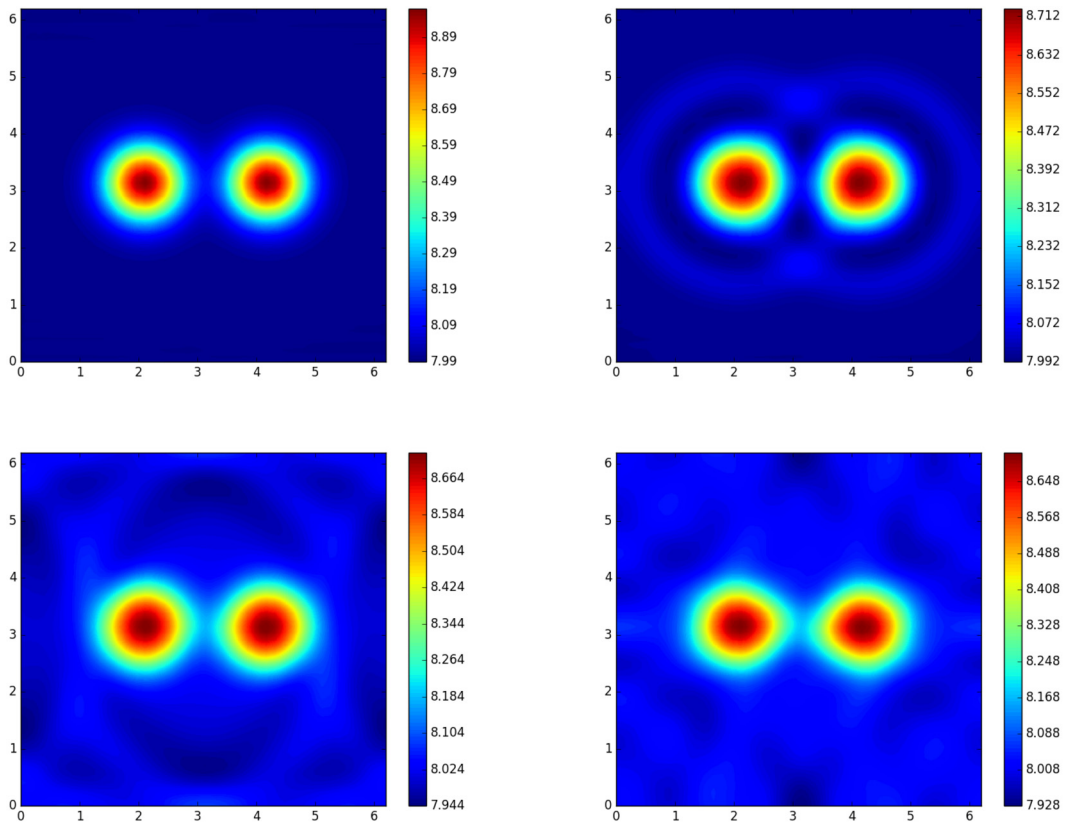

Fig. 5. Fluid depth h at times $t = 0.0$, 0.5, 1.0 and 2.0 dimensionless units. Approximate geostrophic balance is preserved while fast gravity waves radiate outwards from the disturbances.

for $y = [\vec{u}, h]^\top$, $G = [q \times \vec{F} + \nabla(K + gh), \nabla \cdot \vec{F}]^\top$. Fig. 4 shows the convergence of errors for the linearized geostrophic balance problem with $n_x = 4, 8$ and 16 3rd order elements in each dimension (and a time step of $\Delta t = 0.02/n_x$). As can be seen the errors for a given resolution remain constant, showing that once the weak form geostrophic balance relation has been approximated for the first time step there is no subsequent accumulation of errors.

5.3. Conservation

Having verified the construction of our mixed mimetic spectral element operators and solver via various convergence and geostrophic balance tests, we proceed to investigate the conservation properties of the full system (50), as derived in Section 4. The model is tested for a pair of vortices starting in approximate geostrophic balance, as given from the stream function solution

$$\psi = e^{-2.5((x-\pi)^2+(y-2\pi/3)^2)} + e^{-2.5((x-\pi)^2+(y-4\pi/3)^2)} \qquad \Omega = (0, 2\pi] \times (0, 2\pi], \tag{106}$$

where the velocity is given as $\vec{u} = \nabla^\perp \psi$, and the depth is derived from geostrophic balance as $f \times \vec{u} + g\nabla h = 0$, with $f = g = H = 8.0$, using 20×20 3rd order elements and a maximum time step of $\Delta t = 0.0052$. The solution behaves as expected, with fast gravity waves radiating out from the initial disturbance, which is preserved for long times due to the approximate geostrophic balance of the initial condition, as shown in Fig. 5. The ratio of the deformation radius $L_d = \sqrt{gH}/f$ to the nodal grid spacing is approximately 9.55.

As discussed in Section 4, mass and vorticity conservation hold independent of time step, as shown in Fig. 6. This is due to the point wise satisfaction of the divergence theorem in the case of mass, and the elimination of the gradient operator by the curl in the weak form in the case of vorticity. Total energy and potential enstrophy are conserved to truncation error in time, as shown in Fig. 7, with the second order Runge–Kutta time stepping scheme applied with varying time step size.

Gauss–Lobatto–Legendre (GLL) quadrature is known to be exact for polynomials of degree $p = 2n - 3$, where n is the number of quadrature points [35]. Therefore in order to exactly integrate all nonlinear matrices in (50) we use $n = (3p + 3)/2$ quadrature points (where $3p$ is the maximum degree of the test function, trial function and nonlinear function basis expansion product). A second set of tests was also run using inexact quadrature with $n = p + 1$ in order to derive diagonal mass matrices for test functions in W_h, since this significantly increases the computational efficiency of the scheme,

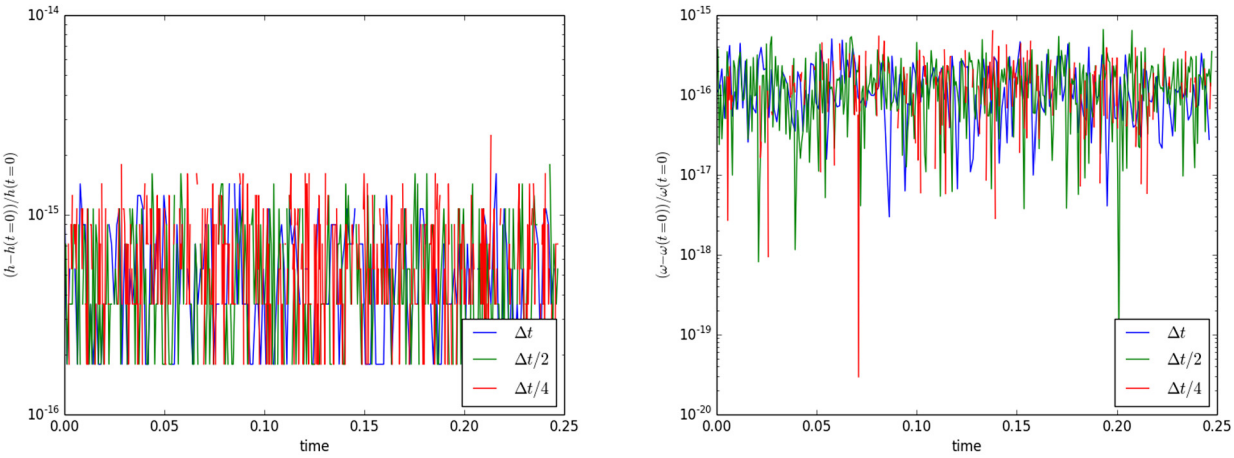

Fig. 6. Exact conservation for the volume, h (left) and vorticity, ω (right) with time.

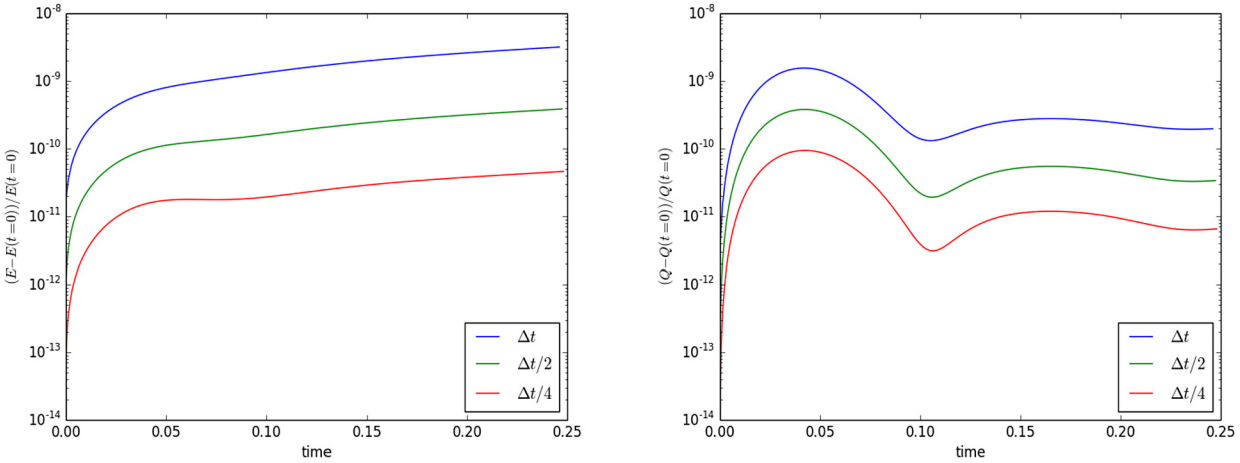

Fig. 7. Convergence of conservation errors for the energy, $\mathcal{E}$ (left) and potential enstrophy, $\mathcal{Q}$ (right) with time step Δt (exact spatial integration).

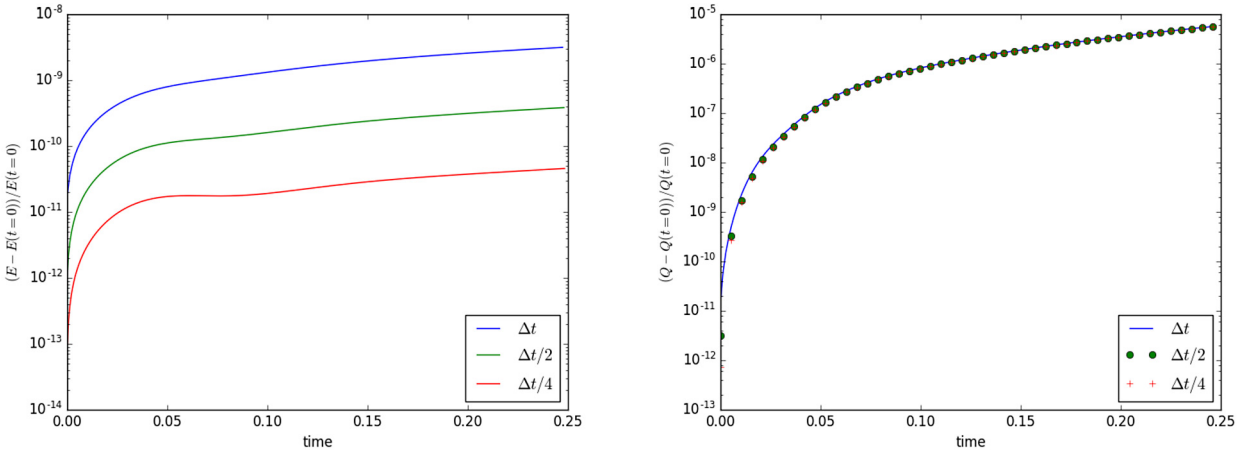

Fig. 8. Convergence of conservation errors for the energy, $\mathcal{E}$ (left) and potential enstrophy, $\mathcal{Q}$ (right) with time step Δt (inexact spatial integration).

and is customary for spectral element models [10]. While energy conservation holds for inexact spatial integration, potential enstrophy conservation fails for inexact integration, as shown in Fig. 8.

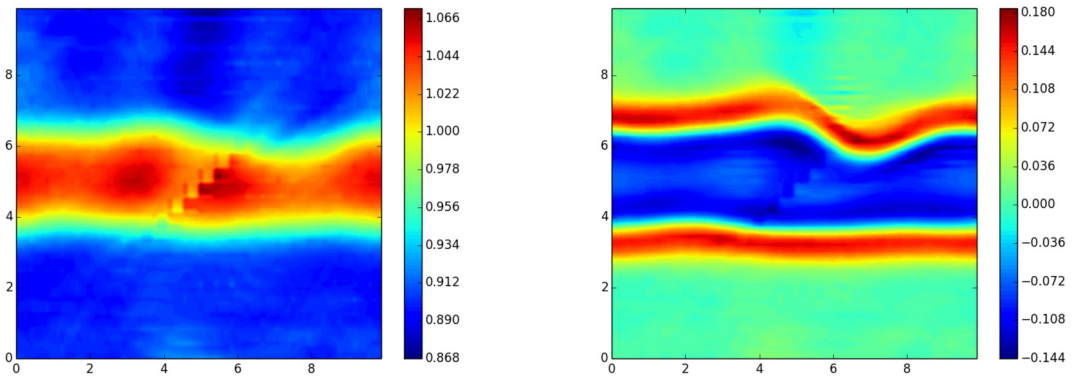

Fig. 9. Fluid depth h (left) and vorticity ω (right) fields for shear flow over orography at $t = 44$, $\Delta\tau = 0.02$.

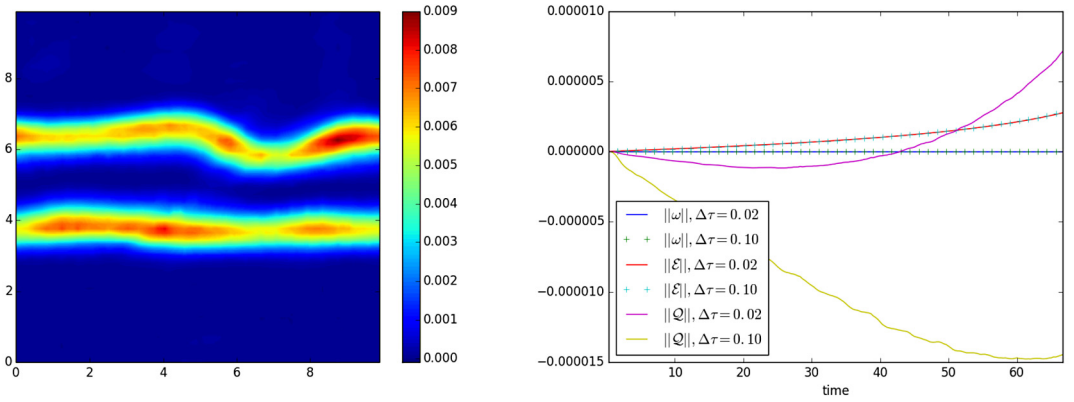

Fig. 10. Kinetic energy per unit volume K for shear flow over orography at $t = 44$, $\Delta\tau = 0.02$ (left), and growth of normalized vorticity ω, energy $\mathcal{E}$ and potential enstrophy $\mathcal{Q}$ with time for varying values of the anticipated potential vorticity coefficient $\Delta\tau$, where $||A|| = (A(t) - A(t=0))/A(t=0)$.

5.4. Stabilization of shear flow via removal of anticipated potential vorticity

Our final numerical experiment involves an investigation of the anticipated potential vorticity method [33] for a shear flow initial condition over a stationary orographic feature. The orography is implemented as a $b_h \in Q_h$, such that the momentum equation (50a) becomes

$$\mathbf{U}\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} + \mathbf{U}^q \boldsymbol{F} - \left(\mathbf{E}^{2,1}\right)^\top \mathbf{Q}\,(\boldsymbol{K} + g\boldsymbol{h} + g\boldsymbol{b}) = 0 \tag{107}$$

and the energy is correspondingly defined as $\mathcal{E} = \langle h_h, K_h + 0.5gh_h + gb_h\rangle_\Omega$. The initial conditions are given as

$$h = H + 0.1\tanh\left(\frac{1-y^2}{2}\right) \qquad \vec{u} = \left(-\frac{\partial h}{\partial y}, 0\right) \tag{108}$$

and the orography as

$$b = \begin{cases} 0.0125(\cos(4\pi x/L)+1)(\cos(4\pi y/L)+1) & \text{if } |x| \le L/4, |y| \le L/4 \\ 0 & \text{otherwise} \end{cases} \tag{109}$$

with $\Omega = (-L/2, +L/2] \times (-L/2, L/2]$, $L = 10$ and $f = g = H = 1.0$. In order to stabilize the potential enstrophy cascade, the anticipated potential vorticity $\hat{q}_h$ (64) substitutes for the potential vorticity q_h in the rotational term of the momentum equation. This results in a loss of potential enstrophy such that the cascade to sub grid scales is arrested [33]. The test is run on 24×24 3rd order elements, with a ratio of L_d to the average nodal grid spacing of 7.2.

Fig. 9 and Fig. 10 show the depth, vorticity and kinetic energy fields after 44 units of dimensionless time have elapsed. While fast gravity waves radiate out from the topography as the solution adjusts, the vorticity field evolves on a slow time scale. The outline of the discontinuous Q_h elements of the orography field is visible in the fluid depth h_h. Fig. 10 also shows the growth in vorticity ω, energy $\mathcal{E}$, and potential enstrophy $\mathcal{Q}$ with time for varying values of the anticipated potential vorticity time scale $\Delta\tau$. As can be seen, vorticity conservation is preserved to machine precision independent of $\Delta\tau$, and

the rate of energy growth is also similar for both values of $\Delta\tau$. The rates of potential enstrophy growth differ however, with $\Delta\tau = 0.02$ resulting in the faster growth of potential enstrophy (following an initial loss) than $\Delta\tau = 0.1$. Neither values of $\Delta\tau$ is able to fully suppress the growth of $\mathcal{Q}$, which over long time integrations will cascade to sub grid scales and contaminate the solution. This failure to fully stabilize the solution via the removal of anticipated potential vorticity is perhaps a result of the projection of the discontinuous h_h field onto q_h within the potential vorticity diagnostic equation. Some method of averaging of h_h at the element boundary quadrature points within (50c) may help to ameliorate this.

6. Conclusion

In this paper we have built upon the work of previous authors in the development of compatible finite element methods for geophysical fluid dynamics [3–5] in order to present a shallow water solver which exactly conserves first order and higher order moments using the mixed mimetic spectral element method [11,12,16]. The conservation of second order moments (energy and potential enstrophy) is subject to the truncation error in the time integration scheme, and the conservation of potential enstrophy also requires exact spatial integration, as shown by the conservation arguments and demonstrated for the test cases given above.

We note the performance constraints of the different diagnostic and prognostic equations as follows:

- Fluid depth, h_h: The continuity equation (50b) is satisfied point wise in the strong form, and may be evaluated purely from the topology, with the divergence theorem satisfied exactly such that the change in fluid depth is simply the sum of the momentum fluxes across the adjacent U_h basis functions of $\vec{F}_h$, and so is extremely fast to compute.
- Kinetic energy per unit volume, K_h: Since K_h also exists on the discontinuous spaces of Q_h, the weak form diagnostic equation (50e) may be solved as a discontinuous Galerkin problem without the need for a global matrix solve.
- Potential vorticity, q_h: If we are prepared to sacrifice potential enstrophy conservation for an inexact quadrature rule, then the left hand side of (50c) is diagonal due to the orthogonality of the W_h basis functions. This again avoids the need for a global matrix solve. The use of inexact GLL quadrature also reduces the computational cost of assembling all other matrices and vectors. This saving may prove significant and override the desirability of potential enstrophy conservation for production codes.
- Velocity, $\vec{u}_h$ and momentum, $\vec{F}_h$: Equations (50a) and (50d) require a global matrix solve, since the function space U_h is continuous across element boundaries. The solution of these equations therefore represents the major computational bottleneck of the scheme.

While we have derived the potential vorticity from a diagnostic equation (50c) in our current formulation, this could alternatively be derived from a prognostic equation by taking the curl of the momentum equation (i.e. by substituting (50a) into (60)). Doing so may have the added advantage of allowing for conservation of energy and potential enstrophy independent of time step via a time staggering of the variables, as has been previously shown for the 2D Navier–Stokes equations [2].

In order to compare the properties of the mixed mimetic spectral element method to the standard A-grid spectral element method [10], the gravity wave dispersion relation for the two method will also be compared, in order to determine if the mixed mimetic method improves upon the spurious discontinuities present in the standard spectral method dispersion relation [36]. The power spectra for nonlinear problems will also be compared to determine if the mixed method can be run with lower amounts of diffusion due to the absence of collocated velocity and pressure degrees of freedom.

Acknowledgements

David Lee would like to thank Drs. Chris Eldred and René Hiemstra for their helpful discussions and insights. This research was supported as part of the Launching an Exascale ACME Prototype (LEAP) project, funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, under contract DE-AC52-06NA25396. Los Alamos Report LA-UR-17-24044.

References

[1] J. Thuburn, Some conservation issues for the dynamical cores of NWP and climate models, J. Comput. Phys. 227 (2008) 3715–3730.
[2] A. Palha, M. Gerritsma, A mass, energy, enstrophy and vorticity conserving (MEEVC) mimetic spectral element discretization for the 2D incompressible Navier–Stokes equations, J. Comput. Phys. 328 (2017) 200–220.
[3] C. Cotter, J. Shipton, Mixed finite elements for numerical weather prediction, J. Comput. Phys. 231 (2012) 7076–7091.
[4] A. McRae, C. Cotter, Energy- and enstrophy-conserving schemes for the shallow-water equations, based on mimetic finite elements, Q. J. R. Meteorol. Soc. 140 (2014) 2223–2234.
[5] A. Natale, J. Shipton, C. Cotter, Compatible finite element spaces for geophysical fluid dynamics, Dyn. Stat. Climate Syst. 1 (2016) 1–31.
[6] A. Arakawa, V. Lamb, A potential enstrophy and energy conserving scheme for the shallow water equations, Mon. Weather Rev. 109 (1981) 18–36.
[7] R. Salmon, Poisson-bracket approach to the construction of energy- and potential-enstrophy-conserving algorithms for the shallow-water equations, J. Atmos. Sci. 61 (2004) 2016–2036.
[8] R. Salmon, A general method for conserving energy and potential enstrophy in shallow-water models, J. Atmos. Sci. 64 (2007) 505–531.
[9] C. Cotter, J. Thuburn, A finite element exterior calculus framework for the rotating shallow-water equations, J. Comput. Phys. 257 (2014) 1506–1526.

[10] M. Taylor, A. Fournier, A compatible and conservative spectral element method on unstructured grids, J. Comput. Phys. 229 (2010) 5879–5895.
[11] M. Gerritsma, Edge functions for spectral element methods, in: Spectral and High Order Methods for Partial Differential Equations, in: Lect. Notes Comput. Sci. Eng., vol. 76, Springer, 2011, pp. 199–207.
[12] J. Kreeft, M. Gerritsma, Mixed mimetic spectral element method for stokes flow: a pointwise divergence-free solution, J. Comput. Phys. 240 (2013) 284–309.
[13] G. Vallis, Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-Scale Circulation, Cambridge University Press, Cambridge, 2006.
[14] J. Thuburn, T. Ringler, W. Skamarock, J. Klemp, Numerical representation of geostrophic modes on arbitrarily structured C-grids, J. Comput. Phys. 228 (2009) 8321–8335.
[15] J. Kreeft, A. Palha, M. Gerritsma, Mimetic framework on curvilinear quadrilaterals of arbitrary order, arXiv:1111.4304, 2011.
[16] R. Hiemstra, D. Toshniwal, R. Huijsmans, M. Gerritsma, High order geometric methods with exact conservation properties, J. Comput. Phys. 257 (2014) 1444–1471.
[17] F. Brezzi, M. Fortin, Mixed and Hybrid Finite Element Methods, Springer Ser. Comput. Math., vol. 15, Springer, 1991.
[18] D. Boffi, F. Brezzi, M. Fortin, Mixed Finite Element Methods and Applications, Springer Ser. Comput. Math., vol. 1, Springer, 2013.
[19] D. Arnold, D. Boffi, F. Bonizzoni, Finite element differential forms on curvilinear cubic meshes and their approximation properties, Numer. Math. 129 (2015) 1–20.
[20] D. Arnold, G. Awanou, Finite element differential forms on cubical meshes, Math. Comput. 83 (288) (2013) 1551–1570.
[21] D. Arnold, D. Boffi, R. Falk, Quadrilateral h(div) finite elements, SIAM J. Numer. Anal. 43 (6) (2005) 2429–2451.
[22] A. Bossavit, Computational electromagnetism and geometry: (1) Network equations, J. Jpn. Soc. Appl. Electromagn. 7 (2) (1999) 150–159.
[23] A. Bossavit, Computational electromagnetism and geometry: (2) Network constitutive laws, J. Jpn. Soc. Appl. Electromagn. 7 (3) (1999) 294–301.
[24] A. Bossavit, Computational electromagnetism and geometry: (3) Convergence, J. Jpn. Soc. Appl. Electromagn. 7 (4) (1999) 401–408.
[25] A. Bossavit, Computational electromagnetism and geometry: (4) From degrees of freedom to fields, J. Jpn. Soc. Appl. Electromagn. 8 (1) (2000) 102–109.
[26] A. Bossavit, Computational electromagnetism and geometry: (5) The “Galerkin Hodge”, J. Jpn. Soc. Appl. Electromagn. 8 (2) (2000) 203–209.
[27] D.N. Arnold, R.S. Falk, R. Winther, Finite element exterior calculus, homological techniques, and applications, Acta Numer. 15 (2006) 1–155.
[28] D.N. Arnold, R.S. Falk, R. Winther, Finite element exterior calculus: from Hodge theory to numerical stability, Bull. Am. Math. Soc. 47 (2) (2010) 281–354.
[29] A. Palha, P. Rebelo, R. Hiemstra, J. Kreeft, M. Gerritsma, Physics-compatible discretization techniques on single and dual grids, with application to the Poisson equation of volume forms, J. Comput. Phys. 257 (2014) 1394–1422.
[30] N. Robidoux, Polynomial histopolation, superconvergent degrees of freedom, and pseudospectral discrete Hodge operators, Unpublished, http://people.math.sfu.ca/~nrobidou/public_html/prints/histogram/histogram.pdf.
[31] R. Abraham, J.E. Marsden, T. Ratiu, Manifolds, Tensor Analysis, and Applications, Appl. Math. Sci., vol. 75, Springer, 2001.
[32] T. Frankel, The Geometry of Physics, 2nd edition, Cambridge University Press, 2004.
[33] R. Sadourny, C. Basdevant, Parameterization of subgrid scale barotropic and baroclinic eddies in quasi-geostrophic models: anticipated potential vorticity method, J. Atmos. Sci. 42 (1985) 1353–1363.
[34] C. Eldred, D. Randall, Total energy and potential enstrophy conserving schemes for the shallow water equations using hamiltonian methods. Part 1. Derivation and properties, Geosci. Model Dev. 10 (2017) 791–810.
[35] G. Karniadakis, S. Sherwin, Spectral/hp Element Methods for Computational Fluid Dynamics, second edition, Oxford University Press, 2005.
[36] T. Melvin, A. Staniforth, J. Thuburn, Dispersion analysis of the spectral element method, Q. J. R. Meteorol. Soc. 138 (2012) 1934–1947.