Journal of Computational Physics 229 (2010) 5879–5895

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

A compatible and conservative spectral element method on unstructured grids

Mark A. Taylor [a,*], Aimé Fournier [b]

[a] Sandia National Laboratories, Albuquerque, NM, USA
[b] National Center for Atmospheric Research, Mesoscale and Microscale Meteorology Division, Boulder, CO, USA

ARTICLE INFO

Article history:
Received 23 November 2009
Received in revised form 5 April 2010
Accepted 6 April 2010
Available online 13 April 2010

Keywords:
Spectral element
Finite element
Mimetic
Compatible
Support operators
Atmospheric modeling
Shallow-water equations
Sphere
Conservation
Local conservation
Unstructured grids
Cubed sphere

ABSTRACT

We derive a formulation of the spectral element method which is *compatible* on very general unstructured three-dimensional grids. Here compatible means that the method retains discrete analogs of several key properties of the divergence, gradient and curl operators: the divergence and gradient are anti-adjoints (the negative transpose) of each other, the curl is self-adjoint and annihilates the gradient operator, and the divergence annihilates the curl. The adjoint relations hold globally, and at the element level with the inclusion of a natural discrete element boundary flux term.

We then discretize the shallow-water equations on the sphere using the cubed–sphere grid and show that compatibility allows us to locally conserve mass, energy and potential vorticity. Conservation is obtained without requiring the equations to be in conservation form. The conservation is exact assuming exact time integration.

1. Introduction

The modern form of the spectral finite-element method (henceforth referred to as SEM) dates to [1], which was based on [2]. It can be formulated as a conventional continuous Galerkin polynomial-based finite-element method. The key difference is that the inner product uses an inexact Gauss–Lobatto quadrature. When combined with a nodal basis that interpolates the quadrature nodes, one obtains a diagonal mass matrix. This is a very efficient way to obtain a high-order explicit method on unstructured grids for time-dependent equations. Because of this, the SEM has been used extensively in geophysical applications including global atmospheric circulation modeling [3–9], ocean modeling [10–12] and planetary scale seismology [13,14].

It has been recently discovered that the continuous Galerkin finite-element method is locally conservative [15]. Here we generalize this result to the inexact-integration case of the SEM, where we also obtain a stronger form of local conservation. Local conservation is a statement about the discrete divergence operator. Here we further generalize these results to show that the SEM is *compatible* (also called *mimetic* or the *support operator* method). Compatible discretizations are those that mimic key vector-calculus properties of the divergence, gradient and curl operators [16–22]. Compatible discretizations

* Corresponding author. Tel.: +1 505 284 1874; fax: +1 505 284 0154.
E-mail addresses: mataylo@sandia.gov (M.A. Taylor), fournier@ucar.edu (A. Fournier).

0021-9991/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2010.04.008

can be formulated for finite-difference, finite-volume and finite-element methods and are considered in a common framework in [23]. They are closely related to discretizations which preserve the properties of the Hamiltonian structure of the continuum equations [24–26].

There is no formal definition of a compatible numerical method. Here we show that the compatible properties of the SEM include:

- A divergence theorem: the discrete divergence and gradient operators are anti-adjoints (adjoints with a negative sign) with respect to the SEM inner product.
- A Stokes theorem: the discrete curl operator is self-adjoint with respect to the SEM inner product.
- The discrete gradient operator is annihilated by the discrete curl operator.
- The discrete curl operator is annihilated by the discrete divergence operator.

The divergence/gradient global adjoint relationship is usually obtained in the SEM by defining a weak-gradient operator as the adjoint of the divergence operator. What is shown here is that merely straightforwardly discretizing gradient and divergence leads to matrices whose adjoint relation implies a discrete statement of the divergence theorem. That these adjoint relations hold globally is shown by first showing that the SEM has discrete divergence and Stokes theorems which hold at the element level. In the continuum, the divergence theorem applied to a volume consisting of a single element includes a boundary term: the integral of a flux term over the surface of the element. The SEM divergence theorem includes a discrete analog of this boundary term. Because the SEM basis functions are globally continuous, the discrete flux will be equal and opposite as computed by adjacent elements, providing local conservation. This element boundary term is similar to the element boundary term that is explicitly included in discontinuous Galerkin (DG) methods. But in a continuous Galerkin method like the SEM, this boundary term is never computed as part of a numerical implementation. A similar concept applies to the SEM curl operator and associated Stokes theorem.

The global divergence/gradient adjoint relationship ensures a symmetric discrete Laplacian, which is of great benefit to iterative solvers. In Cartesian coordinates, the elemental version of this identity can be inferred from Eqs. F.54 and F.55 in [Appendix F [27]]. The fact that this elemental divergence theorem applies to unstructured grids in curvilinear coordinates is not generally known (as with the discrete local version of Stokes theorem). This is evidenced by the fact that the SEM has never been considered a locally conservative method, and local conservation is equivalent to having a discrete divergence theorem. To the best of our knowledge, the annihilator properties of the SEM described here are also previously unknown.

In the case of energy conservation, compatible methods are of interest because they allow conservation without utilizing a total-energy equation [17,28]. In atmospheric modeling, an early use of this property in one dimension dates to [29]. Energy is conserved by the careful mimicking of the energy balance in the original equation: the conversion between kinetic and internal energy terms will exactly balance and the advection operator will not dissipate any kinetic energy. Kinetic-energy dissipation, if needed, must be explicitly added in a compatible method via the introduction of limiters, hyper-viscosity or large-eddy-simulation based approaches.

To verify our results, we use the SEM to discretize the shallow-water equations in curvilinear coordinates on the surface of the sphere and show the method conserves mass, energy and potential vorticity.

2. Spectral-element discretization

We now give a summary of the SEM, using the traditional finite-element inner-product formulation with globally defined continuous basis functions [30,31]. This formulation allows for a clearer illustration of the numerical properties of the method, while the more standard matrix–vector formulation [32,33] is useful for efficient numerical implementations. We present many details which are needed later to show precisely that the method is compatible. We consider only periodic domains, such as the surface of the sphere, so that we may ignore the boundary terms and simplify the exposition.

2.1. Discrete spaces for the SEM

We first define the discrete space used by the SEM. Let x^α and $\vec{x} = \sum_{\alpha=1}^{3} x^\alpha \vec{e}_\alpha$ be the Cartesian coordinates and position vector of a point in the reference cube $[-1,1]^3$ and let r^α and $\vec{r}$ be the (possibly curvilinear) coordinates and position vector of a point in the computational domain, denoted by Ω. We denote the space of polynomials up to degree d in $[-1,1]^3$ by

$$\mathcal{P}_d := \operatorname*{span}_{i,j,k=0}^{d} (x^1)^i (x^2)^j (x^3)^k = \operatorname*{span}_{\vec{\imath}\in\mathbb{I}} \phi_{\vec{\imath}}(\vec{x}),$$

where $\mathbb{I} := \{0,\ldots,d\}^3$ contains all the degrees and $\phi_{\vec{\imath}}(\vec{x}) = \prod_{\alpha=1}^{3} \varphi_{i^\alpha}(x^\alpha), i^\alpha = 0,\ldots,d$, are the cardinal functions, namely polynomials that interpolate the 3D degree-d Gauss–Lobatto-Legendre (GLL) nodes $\vec{\xi}_{\vec{\imath}} := \sum_\alpha \xi_{i^\alpha} \vec{e}_\alpha$. The cardinal-function expansion coefficients of a function g are its GLL nodal values, so we have

$$g(\vec{x}) = \sum_{\vec{\imath}\in\mathbb{I}} g(\vec{\xi}_{\vec{\imath}})\phi_{\vec{\imath}}(\vec{x}) = \boldsymbol{g}^T\boldsymbol{\phi}(\vec{x}) \quad \forall g\in\mathcal{P}_d,$$

$$\text{where the column matrices} \quad \boldsymbol{g} = \sum_{\vec{\imath}\in\mathbb{I}} g(\vec{\xi}_{\vec{\imath}})\boldsymbol{e}_{i^1}\otimes\boldsymbol{e}_{i^2}\otimes\boldsymbol{e}_{i^3} \in \mathbb{R}^{(d+1)^3\times 1} \tag{1}$$

$$\text{and} \quad \boldsymbol{\phi}(\vec{x}) = \sum_{\vec{\imath}\in\mathbb{I}} \phi_{\vec{\imath}}(\vec{x})\boldsymbol{e}_{i^1}\otimes\boldsymbol{e}_{i^2}\otimes\boldsymbol{e}_{i^3} \in \mathbb{R}^{(d+1)^3\times 1}, \tag{2}$$

$()^T$ is the matrix transpose operator, $\boldsymbol{e}_i \in \{0,1\}^{(d+1)\times 1}$ is column $i + 1$ of the identity matrix $\mathbf{I} \in \{0,1\}^{(d+1)\times(d+1)}$ and $\otimes$ denotes the Kronecker product. We then decompose the computational domain Ω using a hexahedral finite-element mesh with a set $\{\Omega_m\}_{m=1}^{M}$ of elements. We assume the mesh is conforming (has no hanging nodes), and that each element can be $\mathcal{C}^1$ mapped to the reference element $[-1,1]^3$. We denote this map and its inverse by

$$\vec{r} = \vec{r}(\vec{x}; m), \quad \vec{x} = \vec{x}(\vec{r}; m). \tag{3}$$

These mapping functions must agree along neighboring element boundaries (shared surfaces $\partial\Omega_m = \partial\Omega_{\acute{m}}$). This implies that the tangential derivatives of $\vec{r}$ will also agree along neighboring element boundaries, but the normal derivatives may not.

We can now define the piecewise-polynomial spectral-element spaces $\mathcal{V}^0$ and $\mathcal{V}^1$ as

$$\mathcal{V}^0 = \{f \in (\mathcal{L}^2\Omega) : f(\vec{r}(\cdot; m)) \in \mathcal{P}_d, \forall m\} = \operatorname*{span}_{m=1}^{M}\{\phi_{\vec{\imath}}(\vec{x}(\cdot; m))\}_{\vec{\imath}\in\mathbb{I}} \tag{4}$$

$$\text{and} \quad \mathcal{V}^1 = \mathcal{C}^0(\Omega) \cap \mathcal{V}^0 = \operatorname*{span}_{\ell=1}^{L} \Phi_\ell.$$

Functions in $\mathcal{V}^0$ are polynomial within each element but may be discontinuous at element boundaries and $\mathcal{V}^1$ is the subspace of continuous function in $\mathcal{V}^0$. The SEM is a Galerkin method with respect to the $\mathcal{V}^1$ subspace and it can be formulated soley in terms of functions in $\mathcal{V}^1$. However, for some intermediate quantities used here it is useful to consider the larger $\mathcal{V}^0$ space. We take $M_d = \dim\mathcal{V}^0 = (d+1)^3 M$, and $L = \dim\mathcal{V}^1 < M_d$. For conforming meshes considered here, a global piecewise cardinal-function basis $\{\Phi_\ell(\vec{r})\}_{\ell=1}^{L}$ for $\mathcal{V}^1$ can be constructed by piecing together appropriate combinations of the M_d possible $\phi_{\vec{\imath}}(\vec{x}(\vec{r}; m))$ in the conventional manner [34,30]. For non-conforming meshes, see [31, eq. A6]. We denote the set of L nodes that these global basis functions interpolate by

$$\{\vec{r}_\ell\}_{\ell=1}^{L} := \bigcup_{m=1}^{M} \vec{r}(\{\vec{\xi}_{\vec{\imath}}\}_{\vec{\imath}\in\mathbb{I}}; m), \quad \text{that is,} \quad \Phi_\ell(\vec{r}_{\acute{\ell}}) = \delta_{\ell,\acute{\ell}}. \tag{5}$$

For every point $\vec{r}_\ell$, there exists at least one element Ω_m and at least one GLL node $\vec{\xi}_{\vec{\imath}} = \vec{x}(\vec{r}_\ell; m)$. In 3D, if $\vec{r}_\ell$ belongs to exactly one Ω_m it is an element-interior node (or global boundary node). If it belongs to exactly two Ω_ms, it is an element-face-interior node. Otherwise it is an edge-interior or vertex node.

We also define similar spaces for 3D vectors. We introduce two families of spaces, with a subscript of either *con* or *cov*, denoting if the contravariant or covariant components of the vectors are piecewise polynomial, respectively.

$$\mathcal{V}^0_{\rm con} = \{\vec{u} \in (\mathcal{L}^2\Omega)^3 : u^\alpha \in \mathcal{V}^0, \alpha = 1,2,3\}$$
$$\text{and} \quad \mathcal{V}^1_{\rm con} = \mathcal{C}^0(\Omega)^3 \cap \mathcal{V}^0_{\rm con},$$

where u^α, α = 1,2,3 are the contravariant components of $\vec{u}$ defined below. Vectors in $\mathcal{V}^1_{\rm con}$ are globally continuous and their contravariant components are polynomials in each element. Similarly,

$$\mathcal{V}^0_{\rm cov} = \left\{\vec{u} \in \mathcal{L}^2(\Omega)^3 : u_\beta \in \mathcal{V}^0, \beta = 1,2,3\right\}$$
$$\text{and} \quad \mathcal{V}^0_{\rm cov} = \mathcal{C}^0(\Omega)^3 \cap \mathcal{V}^0_{\rm cov}.$$

In this work, for functions $f \in \mathcal{V}^0$, we will rely on the cardinal-function (2) expansion local to each element,

$$f(\vec{r}) = \sum_{\vec{\imath}\in\mathbb{I}} f\left(\vec{r}\left(\vec{\xi}_{\vec{\imath}}; m\right)\right)\phi_{\vec{\imath}}(\vec{x}(\vec{r}; m)) = \boldsymbol{f}^{\,T}_{m}\boldsymbol{\phi}(\vec{x}(\vec{r}; m)) \quad \forall \vec{r} \in \Omega_m, \tag{6}$$

where the expansion coefficients are the function values at the mapped GLL nodes, and $\boldsymbol{f}_m$ is defined as in (1) with $g(\vec{x}) = f(\vec{r}(\vec{x}; m))$. We also define a column vector $\boldsymbol{f} = (\boldsymbol{f}^{\,T}_{1}, \boldsymbol{f}^{\,T}_{2}, \ldots, \boldsymbol{f}^{\,T}_{M})^T \in \mathbb{R}^{M_d\times 1}$. As functions f in $\mathcal{V}^0$ can be multiple-valued at GLL nodes that are *redundant* (i.e., shared by more than one element), so $\boldsymbol{f}$ contains all such values. For $f \in \mathcal{V}^1$, the values at any redundant points must all be the same. Note that since $f(\vec{r}(\cdot; m))$ is a polynomial of degree d and there are d + 1 GLL nodes along each edge, then agreement at these points is equivalent to agreement along the entire edge, as required for $\mathcal{V}^1$. To remove the extra degrees of freedom for $f \in \mathcal{V}^1$ represented by these duplicate values, we rely on the expansion in terms of the global cardinal-function basis

$$f(\vec{r}) = \sum_{\ell=1}^{L} f(\vec{r}_\ell)\Phi_\ell(\vec{r}) = \bar{\boldsymbol{f}}^T\boldsymbol{\Phi}(\vec{r}), \tag{7}$$

where $\bar{\boldsymbol{f}}$ and $\boldsymbol{\Phi}(\vec{r})$ are defined in Appendix A. We note that for $f \in \mathcal{V}^1$, we can relate $\boldsymbol{f}$ and $\bar{\boldsymbol{f}}$ by introducing the "scatter matrix" $\mathbf{Q}$ such that

$$\boldsymbol{f} = \mathbf{Q}\bar{\boldsymbol{f}}, \tag{8}$$

and $\mathbf{Q} = (\mathbf{Q}_1, \ldots, \mathbf{Q}_L) \in \{0, 1\}^{M_d \times L}$ denotes the identity with those rows repeated that correspond to redundant degrees of freedom [e.g., [35] p. 79] i.e., each row of the column $\mathbf{Q}_\ell \in \{0, 1\}^{M_d \times 1}$ equals the corresponding value, 0 or 1, of Φ_ℓ. In this work we restrict ourselves to conforming meshes, but note that for non-conforming meshes $\mathbf{Q}$ can be defined to include interpolation factors [e.g., [35] pp. 64 & 79].

2.2. The SEM differential operators in curvilinear coordinates

We first give the standard curvilinear coordinate formulas for vector operators we will use, following [36]. Given the 3 × 3 Jacobian of the the mapping (3) from $[-1,1]^3$ to Ω_m, we denote its determinant-magnitude by

$$J := |\vec{g}_1 \times \vec{g}_2 \cdot \vec{g}_3|, \tag{9}$$

$$\text{where} \quad \vec{g}_\beta := \frac{\partial \vec{r}}{\partial x^\beta} \tag{10}$$

is a covariant basis vector. A vector $\vec{v}$ may be written in terms of physical or covariant or contravariant components, $v[\gamma]$ or v_β or v^α,

$$\vec{v} = \sum_{\gamma=1}^{3} v[\gamma]\frac{\partial \vec{r}}{\partial r^\gamma} = \sum_{\beta=1}^{3} v_\beta \vec{g}^\beta = \sum_{\alpha=1}^{3} v^\alpha \vec{g}_\alpha, \tag{11}$$

that are related by $v_\beta := \vec{v} \cdot \vec{g}_\beta$ and $v^\alpha := \vec{v} \cdot \vec{g}^\alpha$, where $\vec{g}^\alpha := \nabla x^\alpha$ is a contravariant basis vector. The dot product and contravariant components of the cross product are [e.g., [36] Table 1]

$$\vec{u} \cdot \vec{v} = \sum_{\alpha=1}^{3} u_\alpha v^\alpha \quad \text{and} \quad (\vec{u} \times \vec{v})^\alpha = \frac{1}{J}\sum_{\beta,\gamma=1}^{3} \epsilon^{\alpha\beta\gamma} u_\beta v_\gamma \tag{12}$$

where $\epsilon^{\alpha\beta\gamma} \in \{0,\pm 1\}$ is the Levi–Civita symbol.

The divergence, covariant coordinates of the gradient and contravariant coordinates of the curl are [e.g., [36] eqs. 2.1.1, 2.1.4 & 2.1.6]

$$\nabla \cdot \vec{v} = \frac{1}{J}\sum_\alpha \frac{\partial}{\partial x^\alpha}(J v^\alpha), \quad (\nabla f)_\alpha = \frac{\partial f}{\partial x^\alpha} \quad \text{and} \quad (\nabla \times \vec{v})^\alpha = \frac{1}{J}\sum_{\beta,\gamma} \epsilon^{\alpha\beta\gamma}\frac{\partial v_\gamma}{\partial x^\beta}. \tag{13}$$

In the SEM, these operators are all computed in terms of the derivatives with respect to $\vec{x}$ in the reference element, computed exactly (to machine precision) by differentiating the local element expansion (6). For the gradient, the covariant coordinates of $\nabla f, f \in \mathcal{V}^0$ are thus computed exactly within each element. Note that $\nabla f \in \mathcal{V}^0_{\text{cov}}$, but may not be in $\mathcal{V}^1_{\text{cov}}$ even for $f \in \mathcal{V}^1$ due to the fact that its components will be multi-valued at element boundaries because ∇f computed in adjacent elements will not necessarily agree along their shared surfaces. In the case where J is constant within each element, the SEM curl of $\vec{v} \in \mathcal{V}^0_{\text{cov}}$ and the divergence of $\vec{u} \in \mathcal{V}^0_{\text{con}}$ will also be exact, but as with the gradient, multiple-valued at element boundaries.

For non-constant J, these operators may not be computed exactly by the SEM due to the Jacobian factors in the operators and the Jacobian factors that appear when converting between covariant and contravariant coordinates. In [3], these formulas were expanded via the product rule and all derivatives of metric terms were computed analytically. For the compatible version of SEM, we follow [37] and evaluate these operators in the form shown in (13). The quadratic terms that appear are first projected into $\mathcal{V}^0$ via interpolation at the GLL nodes and then this interpolant is differentiated exactly using (6). For example, to compute the divergence of $\vec{v} \in \mathcal{V}^0_{\text{con}}$, we first compute the interpolant $\mathcal{I}(Jv^\alpha) \in \mathcal{V}^0$ of $J\ v^\alpha$, defined by

$$\mathcal{I}f(\vec{r}) := \boldsymbol{\phi}^T(\vec{x}(\vec{r};m))\boldsymbol{f}_{\ m} \quad \forall \vec{r} \in \Omega_m, \quad m = 1, \ldots M, \tag{14}$$

and GLL interpolant of a product fg derives simply from the product of the GLL nodal values of f and g. This operation is just a reinterpretation of the nodal values and is essentially free in the SEM. The derivatives of this interpolant are then computed exactly from (6). The sum of partial derivatives are then divided by J at the GLL nodal values and thus the SEM divergence operator $\nabla_\text{d}\cdot()$ is given by

$$\nabla \cdot \vec{v} \approx \nabla_\text{d} \cdot \vec{v} := \mathcal{I}\left(\frac{1}{J}\sum_\alpha \frac{\partial \mathcal{I}(J v^\alpha)}{\partial x^\alpha}\right) \in \mathcal{V}^0. \tag{15}$$

Similarly, the gradient and curl are approximated by

$$(\nabla f)_\alpha \approx (\nabla_\text{d} f)_\alpha := \frac{\partial f}{\partial x^\alpha} \tag{16}$$

$$\text{and} \quad (\nabla \times \vec{v})^\alpha \approx (\nabla_\text{d} \times \vec{v})^\alpha := \sum_{\beta,\gamma} \epsilon^{\alpha\beta\gamma}\mathcal{I}\left(\frac{1}{J}\frac{\partial v_\gamma}{\partial x^\beta}\right) \tag{17}$$

with $\nabla_{\mathrm{d}} f \in \mathcal{V}^0_{\mathrm{cov}}$ and $\nabla_{\mathrm{d}} \times \vec{v} \in \mathcal{V}^0_{\mathrm{con}}$. The SEM is well known for being quite efficient in computing these types of operations. The SEM divergence, gradient and curl can all be evaluated at the $(d+1)^3$ GLL nodes within each element in $\mathcal{O}(d)$ operations per node using the tensor-product property of these points [33,32]. Besides (14), we will also have use of the interpolating projection of vectors:

$$\mathcal{I}_{\mathrm{con}}\vec{v} := \sum_\alpha \mathcal{I}(v^\alpha)\vec{g}_\alpha \in \mathcal{V}^0_{\mathrm{con}}, \quad \mathcal{I}_{\mathrm{cov}}\vec{v} := \sum_\alpha \mathcal{I}(v_\alpha)\vec{g}^\alpha \in \mathcal{V}^0_{\mathrm{cov}}. \tag{18}$$

Note that for $\vec{v} \in \mathcal{C}^0\Omega)^3, \mathcal{I}_{\mathrm{con}}(\vec{v}) \in \mathcal{V}^1_{\mathrm{con}}$ and $\mathcal{I}_{\mathrm{cov}}(\vec{v}) \in \mathcal{V}^1_{\mathrm{cov}}$.

2.3. The SEM discrete inner product

Instead of using exact integration of the basis functions as in a traditional finite-element method, the SEM uses a GLL quadrature approximation for the integral over Ω, that we denote by $\langle \cdot \rangle$. We define the unlabeled integral as the usual volume-weighted integral over the entire domain Ω. We can write this integral as a sum of volume-weighted integrals over the set of elements $\{\Omega_m\}_{m=1}^M$ used to decompose the domain,

$$\int fg = \sum_{m=1}^M \int_{\Omega_m} fg.$$

The integral over a single element Ω_m is written as an integral over $[-1,1]^3$ by

$$\int_{\Omega_m} fg = \int\int\int_{[-1,1]^3} f(\vec{r}(\cdot;m))g(\vec{r}(\cdot;m))J_m\, dx^1\, dx^2\, dx^3 \approx \langle fg\rangle_{\Omega_m},$$

where we approximate the integral over $[-1,1]^3$ by GLL quadrature,

$$\langle fg\rangle_{\Omega_m} := \sum_{\vec{i}\in\mathbb{I}} w_{i^1} w_{i^2} w_{i^3} J_m\left(\vec{\xi}_{\vec{i}}\right) f\left(\vec{r}\left(\vec{\xi}_{\vec{i}};m\right)\right) g\left(\vec{r}\left(\vec{\xi}_{\vec{i}};m\right)\right) = \boldsymbol{g}_m^T \mathbf{W}_m \boldsymbol{f}_{\ m}, \tag{19}$$

and the element mass matrix $\mathbf{W}_m$ for Ω_m is defined in Appendix A. The SEM approximation to the global integral is then naturally defined as

$$\int fg \approx \sum_{m=1}^M \langle fg\rangle_{\Omega_m} = \langle fg\rangle := \boldsymbol{f}^{\ T}\mathbf{W}\boldsymbol{g}, \quad \text{and similarly } \langle \vec{u}\cdot\vec{v}\rangle = \sum_\alpha \boldsymbol{u}_\alpha^T \mathbf{W} \boldsymbol{v}^\alpha, \tag{20}$$
$$\text{where} \quad \mathbf{W} := \operatorname*{diag}_m \mathbf{W}_m \in \mathbb{R}^{M_d \times M_d}.$$

The SEM global mass matrix for $\mathcal{V}^1$ is diagonal and given by $\mathbf{Q}^T\mathbf{W}\mathbf{Q} \in \mathbb{R}^{L\times L}$. This can be seen by the fact that for $f,g \in \mathcal{V}^1$, using (8) we can also write

$$\langle fg\rangle = \bar{\boldsymbol{f}}^T \mathbf{Q}^T \mathbf{W} \mathbf{Q} \bar{\boldsymbol{g}}. \tag{21}$$

It will be useful below to write either side of (20) in expressions, depending on the need to emphasize a function abstractly as f or concretely in terms of its set $\boldsymbol{f}$ of mapped GLL nodal values, as would be employed in computer codes.

When applied to the product of functions $f,g \in \mathcal{V}^0$, the quadrature approximation $\langle fg\rangle$ defines a discrete inner-product in the usual manner. The quadrature approximation can be applied to any function in $\mathcal{C}^0$. In particular, triple products often occur in the weak formulation of nonlinear equations, so we note for later use that since the quadrature and interpolation nodes coincide,

$$\langle \mathcal{I}(fgh)\rangle = \langle \mathcal{I}(fg)h\rangle = \langle \mathcal{I}(f)gh\rangle = \langle fgh\rangle \quad \forall f,g,h \in \mathcal{C}^0. \tag{22}$$

2.4. The SEM discrete surface integral

For an arbitrary element Ω_m, we will use $\langle\cdot\rangle_{\partial\Omega_m}$ to denote the GLL quadrature approximation to the surface area integral over the boundary of Ω_m,

$$\oint_{\partial\Omega_m} \vec{v}\cdot\hat{n}\, dA \approx \langle \vec{v}\cdot\hat{n}\rangle_{\partial\Omega_m},$$

with dA being the area measure and $\hat{n}$ the outward unit normal. We now give the GLL approximation to this integral in curvilinear coordinates. For

$$(\alpha,\beta,\gamma) \in \mathbb{K} = \{(1,2,3),(2,3,1),(3,1,2)\}, \tag{23}$$

consider the partition $\partial\Omega_m = \cup_{(\alpha,\beta,\gamma)\in\mathbb{K}}(\partial^\gamma_+\Omega_m \cup \partial^\gamma_-\Omega_m)$ into pairs of surfaces

$$\partial^\gamma_\pm\Omega_m := \vec{r}(\square^\gamma_\pm; m), \tag{24}$$

where $\square^\gamma_\pm := \{\vec{x} \in [-1,1]^3 : x^\gamma = \pm 1\}$ denotes the corresponding end faces of the reference cube and $\vec{r}$ (See Eq. (3)) is the map from the reference cube into Ω. Two tangent vectors in $\partial^\gamma_\pm \Omega_m$ are $\vec{g}_\alpha$ and $\vec{g}_\beta$ (10), and thus the outward normal vector and area measure can be written

$$\vec{n} = \pm \vec{g}_\alpha \times \vec{g}_\beta \quad \text{and} \quad dA = |\vec{n}|\, dx^\alpha dx^\beta \quad (\vec{x} \in \square^\gamma_\pm), \tag{25}$$

so that $\hat{n}\, dA = \vec{n}\, dx^\alpha\, dx^\beta$ and we have

$$\int_{\partial^\gamma_\pm \Omega_m} \vec{v} \cdot \hat{n}\, dA = \int\int_{\square^\gamma_\pm} \vec{v} \cdot \vec{n}\, dx^\alpha\, dx^\beta \approx \sum_{i,j=0}^{d} w_i w_j (\vec{v} \cdot \vec{n})|_{x^\alpha = \xi_i, x^\beta = \xi_j, x^\gamma = \pm 1} \tag{26}$$

Since

$$\oint_{\partial \Omega_m} \vec{v} \cdot \hat{n}\, dA = \sum_{(\alpha,\beta,\gamma) \in \mathbb{K}} \left(\int_{\square^\gamma_+} \vec{v} \cdot \vec{n}\, dx^\alpha\, dx^\beta + \int_{\square^\gamma_-} \vec{v} \cdot \vec{n}\, dx^\alpha\, dx^\beta \right),$$

it is natural to define

$$\langle \vec{v} \cdot \hat{n} \rangle_{\partial \Omega_m} = \sum_{(\alpha,\beta,\gamma) \in \mathbb{K}} \sum_{i,j} w_i w_j (\vec{v} \cdot \vec{n}) \Big|^{x^\alpha = \xi_i, x^\beta = \xi_j, x^\gamma = 1}_{x^\alpha = \xi_i, x^\beta = \xi_j, x^\gamma = -1}. \tag{27}$$

Eq. (27) is expressed in Appendix A in the form using matrices that would be applied to a local value-column $\boldsymbol{v}^\tau_m$.

2.5. The projection/Assembly operator

Let us define $\wp : \mathcal{V}^0 \to \mathcal{V}^1$ to be the unique orthogonal (self-adjoint) projection operator from $\mathcal{V}^0$ onto $\mathcal{V}^1$ w.r.t. the SEM discrete inner product (20). The operation $\wp$ is essentially the same as the common procedure in the SEM described as *assembly* [e.g., [32] p. 7], or *direct stiffness summation* [e.g., [33] eq. 4.5.8]. Thus the SEM assembly procedure is not an ad hoc way to remove the redundant degrees of freedom in $\mathcal{V}^0$, but is in fact the natural projection operator $\wp$. At element interior points, it leaves the nodal values unchanged, while at element boundary points shared by multiple elements it is a Jacobian-weighted average over all redundant values. To apply the projection $\wp : \mathcal{V}^0_{\text{cov}} \to \mathcal{V}^1_{\text{cov}}$ to vectors $\vec{u}$, one cannot project the covariant components since the corresponding basis vectors $\vec{g}_\beta$ and $\vec{g}^\alpha$ do not necessarily agree along element faces. Instead we must define the projection as acting on the components using a globally continuous basis such as $\partial \vec{r} / \partial r^\gamma$,

$$\wp(\vec{u}) = \mathcal{I}_{\text{cov}} \left(\sum_\alpha \wp(u[\alpha]) \frac{\partial \vec{r}}{\partial x^\alpha} \right) \quad \vec{u} \in \mathcal{V}^0_{\text{cov}}$$

with a similar definition for $\vec{u} \in \mathcal{V}^0_{\text{con}}$.

To write $g = \wp f$ in matrix form $\mathbf{g} = \mathbf{P}\boldsymbol{f}$, for column matrices $\mathbf{g}, \boldsymbol{f} \in \mathbb{R}^{M_d \times 1}$, we start with the fact that $\langle \Phi_\ell\, g \rangle = \langle \Phi_\ell f \rangle \quad \forall \ell$. Using (5) and (21) to write this in terms of $\bar{\mathbf{g}} \in \mathbb{R}^{L \times 1}$ we have $\mathbf{Q}^T \mathbf{W} \mathbf{Q} \bar{\mathbf{g}} = \mathbf{Q}^T \mathbf{W} \boldsymbol{f}$. Multiplying by the inverse mass matrix and then $\mathbf{Q}$ (8), we obtain $\mathbf{g} = \mathbf{Q}(\mathbf{Q}^T \mathbf{W} \mathbf{Q})^{-1} \mathbf{Q}^T \mathbf{W} \boldsymbol{f}$ and thus

$$\mathbf{P} := \mathbf{Q}(\mathbf{Q}^T \mathbf{W} \mathbf{Q})^{-1} \mathbf{Q}^T \mathbf{W} \in \mathbb{R}^{M_d \times M_d}.$$

The projection corresponds to a weighted sum of redundant, possibly disagreeing values, $\mathbf{Q}^T \mathbf{W}$, followed by the diagonal mass matrix inversion, $(\mathbf{Q}^T \mathbf{W} \mathbf{Q})^{-1}$, followed by a *scatter* to redundant agreeing values by $\mathbf{Q}$ [cf. [6] Eq. (14)]. One can easily verify $\mathbf{P}$ is a projection (idempotent) and self-adjoint ($\mathbf{WP}$ is symmetric).

3. Compatibility in the SEM

3.1. The discrete divergence theorem for an element

In the continuum case, for a single element Ω_m, we always have

$$\int_{\Omega_m} \vec{v} \cdot \nabla f + \int_{\Omega_m} f \nabla \cdot \vec{v} = \oint_{\partial \Omega_m} f \vec{v} \cdot \hat{n}\, dA.$$

We now show the discrete analog of this relation,

$$\langle \vec{v} \cdot \nabla_{\text{d}} f \rangle_{\Omega_m} + \langle f \nabla_{\text{d}} \cdot \vec{v} \rangle_{\Omega_m} = \langle f \vec{v} \cdot \hat{n} \rangle_{\partial \Omega_m} \quad \forall f \in \mathcal{V}^0, \vec{v} \in \mathcal{V}^0_{\text{con}}. \tag{28}$$

Eq. (28) is expressed in Appendix A in the form of matrices that would be applied to local value-columns $\boldsymbol{v}^\tau_m$ and $\boldsymbol{f}_m$. We start by expanding the differential operators on the (28) l.h.s. using the SEM formulation as prescribed in (15). For convenience $u^\gamma := \mathcal{I}(J v^\gamma) \in \mathcal{V}^0$. Then

$$\sum_{\gamma=1}^{3}\left\langle v^{\gamma}\frac{\partial f}{\partial x^{\gamma}}+\frac{f}{J}\frac{\partial\mathcal{I}(Jv^{\gamma})}{\partial x^{\gamma}}\right\rangle_{\Omega_m}=\sum_{\gamma=1}^{3}\sum_{\vec{i}}w_{i^1}w_{i^2}w_{i^3}J\left(v^{\gamma}\frac{\partial f}{\partial x^{\gamma}}+\frac{f}{J}\frac{\partial\mathcal{I}(Jv^{\gamma})}{\partial x^{\gamma}}\right)\Big|_{\vec{x}=\vec{\xi}_{\vec{i}}}$$

$$=\sum_{\gamma=1}^{3}\sum_{\vec{i}}w_{i^1}w_{i^2}w_{i^3}\left(u^{\gamma}\frac{\partial f}{\partial x^{\gamma}}+f\frac{\partial u^{\gamma}}{\partial x^{\gamma}}\right)\Big|_{\vec{x}=\vec{\xi}_{\vec{i}}}$$

$$=\sum_{\gamma=1}^{3}\sum_{\vec{i}}w_{i^1}w_{i^2}w_{i^3}\frac{\partial fu^{\gamma}}{\partial x^{\gamma}}\Big|_{\vec{x}=\vec{\xi}_{\vec{i}}}\tag{29}$$

$$=\sum_{(\alpha,\beta,\gamma)\in\mathbb{K}}\sum_{i,j}w_iw_j\int_{-1}^{1}\frac{\partial fu^{\gamma}}{\partial x^{\gamma}}\Big|_{x^{\alpha}=\xi_i,x^{\beta}=\xi_j}dx^{\gamma}\tag{30}$$

$$=\sum_{(\alpha,\beta,\gamma)\in\mathbb{K}}\sum_{i,j}w_iw_j(fu^{\gamma})\Big|_{x^{\alpha}=\xi_i,x^{\beta}=\xi_j,x^{\gamma}=-1}^{x^{\alpha}=\xi_i,x^{\beta}=\xi_j,x^{\gamma}=1}\tag{31}$$

$$=\sum_{(\alpha,\beta,\gamma)\in\mathbb{K}}\sum_{i,j}w_iw_j(Jfv^{\gamma})\Big|_{\alpha=\xi_i,x^{\beta}=\xi_j,x^{\gamma}=-1}^{x^{\alpha}=\xi_i,x^{\beta}=\xi_j,x^{\gamma}=1}.\tag{32}$$

$$=\sum_{(\alpha,\beta,\gamma)\in\mathbb{K}}\sum_{i,j}w_iw_j(f\vec{v}\cdot\vec{n})\Big|_{x^{\alpha}=\xi_i,x^{\beta}=\xi_j,x^{\gamma}=-1}^{x^{\alpha}=\xi_i,x^{\beta}=\xi_j,x^{\gamma}=1}.\tag{33}$$

We note that the term $\partial f u^{\gamma}/\partial x^{\gamma}$ above is the derivative of a polynomial of degree $2d$ in x^{α} ($\alpha = 1,2,3$), which should not be confused with $\partial\mathcal{I}(fu^{\gamma})/\partial x^{\gamma}$, the derivative of the degree d interpolant that usually appears in the SEM. Hence the (29) summand is a polynomial in x^{γ} of degree at most $2d-1$, evaluated at its GLL nodal values. Because GLL quadrature in the x^{γ} direction is *exact* for such polynomials, we can replace this sum by the x^{γ}-integral (30) which is then evaluated in (31). To get from (31) to (32), we substitute back in $u^{\gamma} := \mathcal{I}(Jv^{\gamma})$ and drop the interpolation operator since the sumand is only evaluated at the interpolation nodes. For the last step, we use (25) and the fact that on the surface $x^{\gamma}=\pm1$, $\vec{v}\cdot\vec{n}=v^{\gamma}n_{\gamma}=\pm Jv^{\gamma}$ by (9). Combining (33) with (27) proves (28).

3.2. The discrete divergence theorem for the whole domain

In the continuum case in a periodic domain, we have

$$\int\vec{v}\cdot\nabla f+\int f\nabla\cdot\vec{v}=0$$

i.e., the gradient and divergence are each the anti-adjoint of the other. We now show the discrete analog of this relation,

$$\langle\vec{v}\cdot\nabla_{\mathrm{d}}f\rangle+\langle f\nabla_{\mathrm{d}}\cdot\vec{v}\rangle=0\quad\forall f\in\mathcal{V}^1,\vec{v}\in\mathcal{V}^1_{\mathrm{con}}.\tag{34}$$

A matrix notation of (34) is given in Appendix A. To show (34), we sum (28) over all elements and use (26) and (27) to show

$$\sum_{m=1}^{M}\langle f\vec{v}\cdot\hat{n}\rangle_{\partial\Omega_m}=\sum_{m=1}^{M}\sum_{(\alpha,\beta,\gamma)\in\mathbb{K}}\sum_{i,j}w_iw_j(f\vec{v}\cdot\vec{n})\Big|_{x^{\alpha}=\xi_i,x^{\beta}=\xi_j,x^{\gamma}=-1}^{x^{\alpha}=\xi_i,x^{\beta}=\xi_j,x^{\gamma}=1}=0.\tag{35}$$

This is a sum over only GLL nodes that lie on element boundaries. For periodic domains, each of the six surfaces in the sum over $(\alpha,\beta,\gamma)\in\mathbb{K}$ from an element m will be shared by an adjacent element $\acute{m}$. Without loss of generality, assume a coordinate labeling so that this surface is given by $x^{\gamma}=1$ for Ω_m and $x^{\gamma}=-1$ for $\Omega_{\acute{m}}$. Since the surfaces coincide, $\vec{r}(\vec{x};m)|_{x^{\gamma}=1}=\vec{r}(\vec{x};\acute{m})|_{x^{\gamma}=-1}$, the tangent vectors $\vec{g}_{\alpha}$ and $\vec{g}_{\beta}$ (10) computed in elements Ω_m and $\Omega_{\acute{m}}$ must agree on this surface and thus the outward surface normal $\vec{n}$ (25) will be equal and opposite in elements Ω_m and $\Omega_{\acute{m}}$,

$$\vec{n}(\vec{r}(\vec{x};m))|_{x^{\gamma}=1}=-\vec{n}(\vec{r}(\vec{x};\acute{m}))|_{x^{\gamma}=-1}.$$

Finally, since $\vec{v}\in\mathcal{V}^1_{\mathrm{con}}$, $\vec{v}(\vec{r}(\vec{x};m))=\vec{v}(\vec{r}(\vec{x};\acute{m}))$ and we see that the discrete flux out of a surface of element Ω_m is identical to the flux into the adjacent element which shares that surface. The net flux when summing over all surfaces of all elements will thus exactly cancel, establishing (35). Note that we require only that the $\vec{r}(\cdot;m)$ agree along element boundaries for neighboring Ω_m. The derivatives of these maps must be well defined within each element but their normal derivatives are not required to agree along element boundaries.

3.3. The discrete Stokes theorem for an element

In the continuum case, we have that the curl operator is self-adjoint. Applied to a single element Ω_m, this takes the form of the identity

$$\int_{\Omega_m}\vec{u}\cdot\nabla\times\vec{v}-\int_{\Omega_m}\vec{v}\cdot\nabla\times\vec{u}=\oint_{\partial\Omega_m}(\vec{v}\times\vec{u})\cdot\hat{n}\,dA,$$

with the SEM discrete analog

$$\langle \vec{u} \cdot \nabla_{\mathrm{d}} \times \vec{v} \rangle_{\Omega_m} - \langle \vec{v} \cdot \nabla_{\mathrm{d}} \times \vec{u} \rangle_{\Omega_m} = \langle (\vec{v} \times \vec{u}) \cdot \hat{n} \rangle_{\partial\Omega_m} \quad \forall \vec{u}, \vec{v} \in \mathcal{V}^0_{\mathrm{cov}}. \tag{36}$$

A matrix notation of (36) is given in Appendix A. Using the form of the curl operator from (17), the (36) l.h.s. is

$$\sum_{\gamma,\sigma,\tau=1}^{3} \left\langle \frac{\epsilon^{\sigma\gamma\tau}}{J} \left(u_\sigma \frac{\partial v_\tau}{\partial x^\gamma} - v_\sigma \frac{\partial u_\tau}{\partial x^\gamma} \right) \right\rangle = \sum_{\gamma,\sigma,\tau=1}^{3} \sum_{\vec{i}} w_{i^1} w_{i^2} w_{i^3} \epsilon^{\sigma\gamma\tau} \left(u_\sigma \frac{\partial v_\tau}{\partial x^\gamma} + v_\tau \frac{\partial u_\sigma}{\partial x^\gamma} \right) \Bigg|_{\vec{x}=\vec{\xi}_{\vec{i}}} \tag{37}$$

$$= \sum_{\gamma,\sigma,\tau=1}^{3} \sum_{\vec{i}} w_{i^1} w_{i^2} w_{i^3} \epsilon^{\sigma\gamma\tau} \frac{\partial u_\sigma v_\tau}{\partial x^\gamma} \Bigg|_{\vec{x}=\vec{\xi}_{\vec{i}}}, \tag{38}$$

where we changed the sign of the second (37) term by swapping σ and τ and using $\epsilon^{\tau\gamma\sigma} = -\epsilon^{\sigma\gamma\tau}$. Following the same procedure as for the discrete divergence theorem, stepping from (29) to (30), we see that the GLL quadrature will be exact with respect to $\mathrm{d}x^\gamma$ for the (38) terms involving $\partial/\partial x^\gamma$. As in the step from (30) to (31), replacing these sums with the integrals and changing the sum over γ to a sum over the triplets (23), (38) becomes

$$\sum_{(\alpha,\beta,\gamma)\in\mathbb{K}} \sum_{\sigma,\tau} \sum_{i,j} w_i w_j \epsilon^{\sigma\gamma\tau} (u_\sigma v_\tau) \Bigg|_{x^\alpha=\xi_i, x^\beta=\xi_j, x^\gamma=-1}^{x^\alpha=\xi_i, x^\beta=\xi_j, x^\gamma=1} \tag{39}$$

We note that on the surface $x^\gamma = \pm 1$, for $(\alpha, \beta, \gamma) \in \mathbb{K}$ we have $(\vec{v} \times \vec{u}) \cdot \vec{n} = \pm \sum_{\sigma,\tau} \epsilon^{\sigma\gamma\tau} u_\sigma v_\tau$ which can be derived from (12) and the fact that $\vec{g}_\alpha \cdot \vec{n} = \vec{g}_\beta \cdot \vec{n} = 0$ and $\vec{g}_\gamma \cdot \vec{n} = \pm J$ (9), (25). Thus the surface integral, (36) r.h.s., expanded using (12), (27) is

$$\langle (\vec{v} \times \vec{u}) \cdot \hat{n} \rangle_{\partial\Omega_m} = \sum_{(\alpha,\beta,\gamma)\in\mathbb{K}} \sum_{i,j} w_i w_j \sum_{\sigma,\tau} \epsilon^{\sigma\gamma\tau} (v_\tau u_\sigma) \Bigg|_{x^\alpha=\xi_i, x^\beta=\xi_j, x^\gamma=-1}^{x^\alpha=\xi_i, x^\beta=\xi_j, x^\gamma=1},$$

which equals (39), and thus we have shown (36).

3.4. The discrete Stokes theorem for the whole domain

The extension of this result to the global identity

$$\langle \vec{u} \cdot \nabla_{\mathrm{d}} \times \vec{v} \rangle - \langle \vec{v} \cdot \nabla_{\mathrm{d}} \times \vec{u} \rangle = 0 \quad \forall \vec{u}, \vec{v} \in \mathcal{V}^1_{\mathrm{cov}}, \tag{40}$$

proceeds exactly as in Section 3.2, by showing the boundary flux, (36) r.h.s., will sum to zero. In particular, we have that since $\vec{u}$ is continuous at element boundaries, then on a surface $\partial^\gamma_\pm \Omega_m$ (24), with $\alpha \neq \beta \neq \gamma, u_\alpha$ and u_β will also be continuous by the same argument used to show $\vec{n}$ is continuous (because the maps agree along boundaries) in Section 3.2. This also shows that the result can be generalized to

$$\langle \nabla_{\mathrm{d}} f \cdot \nabla_{\mathrm{d}} \times \vec{v} \rangle - \langle \vec{v} \cdot \nabla_{\mathrm{d}} \times \nabla_{\mathrm{d}} f \rangle = 0 \quad \forall \vec{v} \in \mathcal{V}^1_{\mathrm{con}}, f \in \mathcal{V}^1, \tag{41}$$

since in this formulation, in the flux terms u_α is replaced by $\partial f/\partial x^\alpha$ and u_β is replaced by $\partial f/\partial x^\beta$, both of which also agree along element boundaries $\partial^\gamma_\pm \Omega_m$ if $f \in \mathcal{V}^1$.

3.5. Annihilator properties: $\nabla_{\mathrm{d}} \times \nabla_{\mathrm{d}} f = \vec{0}, \nabla_{\mathrm{d}} \cdot \nabla_{\mathrm{d}} \times \vec{\psi} = 0$

Let $f \in \mathcal{V}^0$ and $\vec{\psi} \in \mathcal{V}^0_{\mathrm{cov}}$. If one simply applies the differential operators within an element using the formulations (15)–(17), we trivially obtain $\nabla_{\mathrm{d}} \times \nabla_{\mathrm{d}} f = \vec{0}$ and $\nabla_{\mathrm{d}} \cdot \nabla_{\mathrm{d}} \times \vec{\psi} = 0$ pointwise because of symmetry: $\sum_{\alpha\beta} \epsilon^{\alpha\beta\gamma} \mathbf{D}_\alpha \mathbf{D}_\beta = \mathbf{0} \quad \forall \gamma$ (that $\mathbf{D}_\alpha \mathbf{D}_\beta = \mathbf{D}_\beta \mathbf{D}_\alpha$ is evident from the tensor-product formula in Appendix A). Restricting the spaces to $\mathcal{V}^1$ and $\mathcal{V}^1_{\mathrm{cov}}$ so we can apply (41), these relations are equivalent to

$$\langle \nabla_{\mathrm{d}} \times \vec{\psi} \cdot \nabla_{\mathrm{d}} f \rangle = 0 \quad \forall \vec{\psi} \in \mathcal{V}^1_{\mathrm{cor}}, f \in \mathcal{V}^1. \tag{42}$$

Closely related and often important identities are

$$\wp \nabla_{\mathrm{d}} \times \wp \nabla_{\mathrm{d}} f = \vec{0} \quad \text{and} \quad \wp \nabla_{\mathrm{d}} \cdot \wp \nabla_{\mathrm{d}} \times \vec{\psi} = 0 \quad \forall \vec{\psi} \in \mathcal{V}^1_{\mathrm{cov}}, f \in \mathcal{V}^1. \tag{43}$$

The procedure to establish each of these identities is identical and thus we discuss only $\wp \nabla_{\mathrm{d}} \times \wp \nabla_{\mathrm{d}} f = \vec{0}$. We first note that this identity is equivalent to

$$0 = \langle \vec{\psi} \cdot \wp \nabla_{\mathrm{d}} \times \wp \nabla_{\mathrm{d}} f \rangle = \langle \vec{\psi} \cdot \nabla_{\mathrm{d}} \times \wp \nabla_{\mathrm{d}} f \rangle = \langle \nabla_{\mathrm{d}} \times \vec{\psi} \cdot \wp \nabla_{\mathrm{d}} f \rangle \quad \forall \vec{\psi} \in \mathcal{V}^1_{\mathrm{cov}}, f \in \mathcal{V}^1.$$

where we used that $\wp$ is self-adjoint and $\wp \vec{\psi} = \vec{\psi}$. Combining this with (42), to show (43) it is sufficient to show

$$\langle \nabla_{\mathrm{d}} \times \vec{\psi} \cdot (\wp - 1) \nabla_{\mathrm{d}} f \rangle = 0. \tag{44}$$

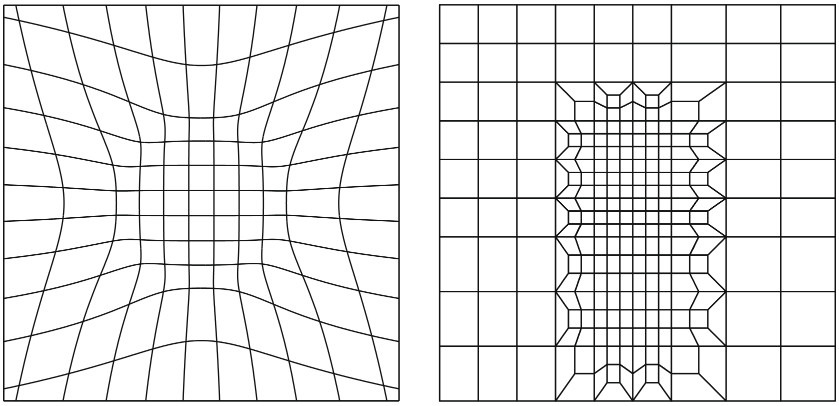

Fig. 1. Example two-dimensional grids. On both grids the SEM has a discrete divergence and Stokes theorem and $\nabla_{\rm d} \times \nabla_{\rm d} f = \vec{0}$ and $\nabla_{\rm d} \cdot \nabla_{\rm d} \times \vec{\psi} = 0$. For the unstructured grid on the left, the SEM also satisfies $\wp\nabla_{\rm d} \times \wp\nabla_{\rm d} f = \vec{0}$ and $\wp\nabla_{\rm d} \cdot \wp\nabla_{\rm d} \times \vec{\psi} = 0$.

For the SEM, an analysis of (44) shows that it will be zero only for grids with certain geometric properties. We have not developed necessary conditions for this identity. But we will note (without proof) for later use that this property will hold for all 2D grids which satisfy:

1. All corner nodes must have redundancy of either 3 or 4.
2. For corner nodes with redundancy of 4: the 4 possible values of the Jacobian determinant (as computed within the 4 elements containing the node) must agree.

3.6. Example grids

We now consider the two-dimensional grids in Fig. 1. The grid on the left is logically Cartesian, generated by the radially stretched, non-conformal 2D map

$$\begin{pmatrix} r^1 \\ r^2 \end{pmatrix} \leftarrow \frac{\varrho\left(\sqrt{(r^1)^2 + (r^2)^2}\right)}{\varrho(1)\sqrt{(r^1)^2 + (r^2)^2}} \begin{pmatrix} r^1 \\ r^2 \end{pmatrix},$$

$$\text{where} \quad \varrho(r) := \frac{2}{5} r W(a-r) + \left(r - \frac{3}{5}a\right) W(r-a)$$

is a ramp that tilts smoothly at $r = a = \frac{7}{10}$ between slopes $\frac{2}{5}$ and 1, and $W(x) = (1 + \tanh 8x)/2$ is a semi-infinite window function. The grid on the right is unstructured and has all straight sided elements that can be mapped to the reference quadrilateral via the conventional bi-linear map. For both grids the map to the reference element is $\mathcal{C}^1$ within each element and globally $\mathcal{C}^0$, and thus the tangential-derivative $\vec{g}_\alpha$ and $\vec{g}_\beta$ (10) agree along every boundary $\partial^{\gamma}_{\pm}\Omega_m$ and so the SEM on these grids will obey the discrete divergence theorem (28), (34), Stokes theorem 36, 40, 41 and $\nabla_{\rm d} \times \nabla_{\rm d} f = \vec{0}$ and $\nabla_{\rm d} \cdot \nabla_{\rm d} \times \vec{\psi} = 0$ identities. For the left grid, since the map is globally $\mathcal{C}^1$ it satisfies items 1 and 2 in Section 3.5 and thus the SEM for this grid will also obey the $\wp\nabla_{\rm d} \times \wp\nabla_{\rm d} f = \vec{0}$ and $\wp\nabla_{\rm d} \cdot \wp\nabla_{\rm d} \times \vec{\psi} = 0$ identity. This identity will not hold for the grid on the right.

4. Shallow-water equations on the surface of the sphere

We will now apply the compatible spectral-element formulation to the shallow-water equations on the surface of the unit sphere. We use spherical coordinates $\lambda = r^1$ for longitude, $\theta = r^2$ for latitude and $r = r^3$ for radius, with associated unit vectors $\hat{\lambda}, \hat{\theta}$ and $\hat{k}$, respectively. We restrict our functions to the surface of the unit sphere ($r = 1$) and assume $\partial/\partial r = 0$. To discretize, we use the cubed-sphere grid (Fig. 2) first used in [38]. Each cube face is mapped to the surface of the sphere with the equal-angle gnomonic projection [39]. The map from the reference element $[-1,1]^2$ to the cube face is a translation and scaling. The composition of these two maps we denote by $\lambda = \lambda(x^1,x^2)$ and $\theta = \theta(x^1,x^2)$. For the cubed sphere, all vertex nodes have redundancy 3 or 4. Within each cube face, the Jacobian of the mapping is globally $\mathcal{C}^1$. For the redundancy-4 nodes on the cube-face edges, the Jacobian is also continuous by symmetry and thus this grid also satisfies items 1 and 2 in Section 3.5 and thus the SEM on this grid will obey (43).

We solve the shallow-water equations in rotational form,

$$\frac{\partial \vec{u}}{\partial t} = -\omega \hat{k} \times \vec{u} - \nabla\left(\frac{1}{2}\vec{u}^2 + gH\right),$$

$$\frac{\partial h}{\partial t} = -\nabla \cdot h\vec{u},$$

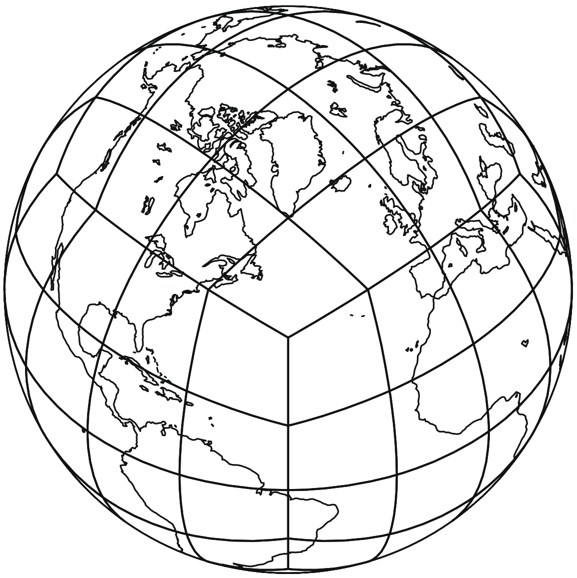

Fig. 2. Tiling the surface of the sphere with quadrilaterals. An inscribed cube is projected to the surface of the sphere. The faces of the cubed sphere are further subdivided to form a quadrilateral grid of the desired resolution. Coordinate lines from the gnomonic equal-angle projection are shown.

for velocity $\vec{u}$ and fluid thickness h, with the absolute vorticity $\omega = \hat{k} \cdot (\nabla \times \vec{u} + 2\vec{\Omega})$, gravity g, rotation $\vec{\Omega}$, and bottom-surface elevation h_s and thus fluid-surface height $H = h + h_s$. The SEM discretization of the shallow-water system is to find $\vec{u}(\cdot,t) \in \mathcal{V}^1_{\text{cov}}$ and $h(\cdot,t) \in \mathcal{V}^1$ such that for all $\vec{\psi} \in \mathcal{V}^1_{\text{con}}$ and $\psi \in \mathcal{V}^1$,

$$\left\langle \vec{\psi} \cdot \frac{\partial \vec{u}}{\partial t} \right\rangle = -\langle \vec{\psi} \cdot \omega \hat{k} \times \vec{u} \rangle - \left\langle \vec{\psi} \cdot \nabla_{\text{d}} \mathcal{I}\left(\frac{1}{2}\vec{u}^2 + gH\right) \right\rangle, \tag{45}$$

$$\left\langle \psi \frac{\partial h}{\partial t} \right\rangle = -\langle \psi \nabla_{\text{d}} \cdot h\vec{u} \rangle. \tag{46}$$

The SEM (assuming exact time integration) solves (45), (46) *exactly*. Note that the argument to ∇_{d} must be in $\mathcal{V}^0$, and thus we have inserted an $\mathcal{I}$ operator since in general $\vec{u}^2 \notin \mathcal{V}^0$ is a polynomial of degree $2d$ within each element. No $\mathcal{I}$ operator is needed in the remaining terms due to (22) and the the definition of $\nabla_{\text{d}} \cdot ()$ in (15).

The choices $\vec{\psi} = \mathcal{I}_{\text{con}}(\Phi_\ell \hat{\lambda})$ or $\mathcal{I}_{\text{con}}(\Phi_\ell \hat{\theta})$ take (45), (46) to assembled equations for the physical-component value-arrays $\frac{d}{dt}\bar{\boldsymbol{u}}[\lambda]$ or $\frac{d}{dt}\bar{\boldsymbol{u}}[\theta]$ and $\frac{d}{dt}\bar{\boldsymbol{h}}$, all members of $\mathbb{R}^{L\times 1}$. There are several equivalent numerical approaches to solving the resulting system of ODEs for these three $\mathbb{R}^{L\times 1}$ vectors, with different efficiencies depending on the polynomial degree, implementation issues and machine architectures [40]. The global matrix approach works directly with the $\mathbb{R}^{L\times 1}$ solution vectors and is a commen view of the SEM. Here we use an elemental decomposition, where the numerical implementation works with functions in the larger $\mathcal{V}^0, \mathcal{V}^0_{\text{cov}}$ spaces (vectors in $\mathbb{R}^{M_d\times 1}$) with projections back to $\mathcal{V}^1, \mathcal{V}^1_{\text{cov}}$ (vectors in $\mathbb{R}^{L\times 1}$) when necessary. Note that in all cases, the boundary integral terms that appear in the discrete divergence and Stokes theorems do not appear in the numerical implementation. They represent flux terms that are implicit in the SEM, ensuring local conservation, but the SEM does not make use of them directly.

With the elemental decomposition, the solution of (45), (46) is solved with the following two-step procedure. For simplicity, assume a forward-Euler discretization in time:

1. From a given state $h(t) \in \mathcal{V}^1, \vec{u}(t) \in \mathcal{V}^1_{\text{cov}}$ at time t, advance the solution within each element by Δt.

$$\frac{\vec{u}(*) - \vec{u}(t)}{\Delta t} = -\omega(t)\hat{k} \times \vec{u}(t) - \nabla_{\text{d}} \mathcal{I}\left(\frac{1}{2}\vec{u}(t)^2 + gH(t)\right) \tag{47}$$

$$\frac{h(*) - h(t)}{\Delta t} = -\nabla_{\text{d}} \cdot h(t)\vec{u}(t). \tag{48}$$

where $h(*) \in \mathcal{V}^0, \vec{u}(*) \in \mathcal{V}^0_{\text{cov}}$. This step is completely local to the element, making it extremely efficient on parallel computers if each processor has one or more elements in memory. We write this step in terms of solutions in $\mathbb{R}^{M_d\times 1}$:

$$\frac{\boldsymbol{u}[\alpha](*)-\boldsymbol{u}[\alpha](t)}{\Delta t}=\boldsymbol{\omega}(t)\boxtimes\sum_{\beta}\epsilon^{\alpha\beta3}\boldsymbol{u}[\beta](t)-\mathsf{D}[\alpha]\left(\frac{1}{2}\sum_{\beta}\boldsymbol{u}[\beta](t)\boxtimes\boldsymbol{u}[\beta](t)+g\boldsymbol{H}(t)\right),$$

$$\frac{\boldsymbol{h}(*)-\boldsymbol{h}(t)}{\Delta t}=\sum_{\beta}\boldsymbol{\Delta}_{\beta}\boldsymbol{h}(t)\boxtimes\boldsymbol{u}^{\beta}(t),$$

$$\text{where}\quad \mathsf{D}[\alpha]=\operatorname*{diag}_{m}\operatorname*{diag}_{\vec{\imath}}\sum_{\gamma}\frac{\partial x^{\gamma}}{\partial r^{\alpha}}(\vec{r}(\vec{\xi}_{\vec{\imath}};m);m)\mathsf{D}_{\gamma}$$

is the physical-component gradient matrix, $\boldsymbol{\omega}=2\boldsymbol{\Omega}_3-\sum_{\gamma}\mathsf{C}^{\gamma3}\boldsymbol{u}_{\gamma}$ contains absolute-vorticity values, $\boldsymbol{u}_{\beta}=\sum_{\gamma}\boldsymbol{G}_{\beta,\gamma}\boxtimes\boldsymbol{u}^{\gamma}$ contains covariant velocity values, $G_{\beta,\gamma;m,\vec{\imath}}=\vec{g}_{\beta}\cdot\vec{g}_{\gamma}(\vec{\xi}_{\vec{\imath}};m)$ contains metric-tensor values, entry i of the Hadamard-Schur product $\boldsymbol{a}\boxtimes\boldsymbol{b}$ is just a_ib_i, and other matrices are defined in Appendix A.

2. Project the solution back to $\mathcal{V}^1,\mathcal{V}^1_{\text{cov}}$:

$$\vec{u}(t+\Delta t)=\wp(\vec{u}(*)),\quad h(t+\Delta t)=\wp(h(*)). \tag{49}$$

4.1. Global mass conservation

Define the discrete mass to be $M:=\langle h\rangle$. Taking ψ = 1 in (46), applying (34) and noting that $\nabla_{\text{d}}1=\vec{0}$ will be computed exactly, we see that $\frac{d}{dt}M=\langle\frac{\partial h}{\partial t}\rangle=0$. Mass conservation will be exact for any reasonable time stepping scheme.

If we also advect a tracer,

$$\frac{\partial q}{\partial t}+\vec{u}\cdot\nabla q=0,$$

by solving

$$\left\langle\psi\frac{\partial q}{\partial t}\right\rangle+\langle\psi\vec{u}\cdot\nabla_{\text{h}}q\rangle=0, \tag{50}$$

then its mass $\langle h\,q\rangle$ is also exactly conserved with exact time-stepping. This can be seen by testing (46) with ψ = q and testing (50) with ψ = h, summing and applying (34) to derive

$$\frac{d}{dt}\langle qh\rangle=\left\langle q\frac{\partial h}{\partial t}\right\rangle+\left\langle h\frac{\partial q}{\partial t}\right\rangle=-\langle h\vec{u}\cdot\nabla_{\text{h}}q\rangle-\langle q\nabla_{\text{d}}\cdot h\vec{u}\rangle=0.$$

4.2. Local mass conservation

To show local conservation within a single element Ω_m, one would like to choose a test function for (46) that takes the value 1 in Ω_m and 0 elsewhere. But such a test function would belong to $\mathcal{V}^0$, not $\mathcal{V}^1$ and is thus not an allowable choice in (46). Instead, we examine steps 1 and 2 in Section 4 separately and show each is locally conservative. Considering the mass in Ω_m,

$$M_m(t):=\langle h(t)\rangle_{\Omega_m},\quad M_m(*):=\langle h(*)\rangle_{\Omega_m},$$

and applying (28)–(48), we have that

$$\frac{M_m(*)-M_m(t)}{\Delta t}=-\langle h(t)\vec{u}(t)\cdot\hat{n}\rangle_{\partial\Omega_m}$$

and thus we have a strong form of local conservation for step 1. The change in mass within an element is exactly equal to the flux of mass through the boundaries. The flux is continuous across elements and thus the flux of mass out of Ω_m is exactly offset by the gain in mass of the adjacent elements.

We then apply step 2, which is mass conserving since $\wp$ is self-adjoint (i.e. $\langle\wp(h)\rangle$ = $\langle 1\wp(h)\rangle$ = $\langle\wp(1)h\rangle$ = $\langle h\rangle$). In the case of the SEM, $\wp$ is a local operation when expressed in terms of the point values at the GLL nodes, as it is the Jacobian weighted average over redundant nodal values at element boundaries, and thus step 2 is also locally conservative. The projection operator would also be locally conservative in a mass-lumped finite-element method. But in general, for a finite element method with a non-diagonal mass matrix, this strong form of local conservation would be *lost* and instead one would have local conservation in the sense of [15].

4.3. Energy conservation

For the momentum Eq. (45), take the test function $\vec{\psi}=\mathcal{I}_{\text{con}}(h\vec{u})$.

By the t-derivative chain rule and (22) we have

$$\left\langle\frac{1}{2}h\frac{\partial\vec{u}^2}{\partial t}\right\rangle=-\left\langle h\vec{u}\cdot\nabla_{\text{d}}\mathcal{I}\left(\frac{1}{2}\vec{u}^2+gH\right)\right\rangle.$$

For the continuity Eq. (46), we test with
$\psi = \mathcal{I}(\frac{1}{2}\vec{u}^2)$ and $\psi = g\,H$, applying (34) to the r.h.s. in both cases, to obtain

$$\left\langle \frac{1}{2}\vec{u}^2 \frac{\partial h}{\partial t} \right\rangle = \left\langle h\vec{u} \cdot \nabla_{\mathrm{d}} \mathcal{I}\left(\frac{1}{2}\vec{u}^2\right) \right\rangle$$

$$\text{and} \quad \left\langle \frac{1}{2} \frac{\partial gH^2}{\partial t} \right\rangle = \langle h\vec{u} \cdot \nabla_{\mathrm{d}} gH \rangle.$$

Summing these three equations, we obtain a discrete analog of total-energy conservation:

$$\frac{dE}{dt} = 0, \quad \text{where} \quad E = \frac{1}{2}\langle h\vec{u}^2 + gH^2 \rangle = \frac{1}{2}\boldsymbol{h}^T \boxtimes \vec{\boldsymbol{u}}^T \cdot \mathbf{W}\vec{\boldsymbol{u}} + \frac{1}{2} g\boldsymbol{H}^T \mathbf{W}\boldsymbol{H}. \tag{51}$$

Local conservation of energy, in the sense shown in Section 4.2, can be obtained if one retains all the element boundary terms in the above manipulations.

4.4. Potential vorticity conservation

The potential vorticity in the shallow-water equations is given by $q = \omega/h$. The equation for potential vorticity,

$$\frac{\partial q}{\partial t} = -\vec{u} \cdot \nabla q,$$

written in conservation form is

$$\frac{\partial \omega}{\partial t} = -\nabla \cdot \omega\vec{u}. \tag{52}$$

The SEM discretization of (45), (46) will locally conserve potential vorticity as a consequence of the compatibility of the divergence operator and that $\nabla \times \nabla f = \vec{0}$. To show this, we use $\wp(\nabla_{\mathrm{d}} \times \psi\hat{k})$ as a test function in (45). Using $\vec{u} = \wp\vec{u}$ and that $\wp$ is self-adjoint, we have

$$\left\langle (\nabla_{\mathrm{d}} \times \psi\hat{k}) \cdot \frac{\partial \vec{u}}{\partial t} \right\rangle = -\left\langle (\nabla_{\mathrm{d}} \times \psi\hat{k}) \cdot \wp(\omega\hat{k} \times \vec{u}) \right\rangle - \left\langle (\nabla_{\mathrm{d}} \times \psi\hat{k}) \cdot \wp\nabla_{\mathrm{d}} \mathcal{I}\left(\frac{1}{2}\vec{u}^2 + gH\right) \right\rangle.$$

On the cubed-sphere grid, the last term vanishes by (43). Applying (40) to the remaining terms, we have

$$\left\langle \psi \frac{\partial \omega}{\partial t} \right\rangle = -\left\langle \hat{k}\psi \cdot \nabla_{\mathrm{d}} \times \wp(\omega\hat{k} \times \vec{u}) \right\rangle.$$

We further reduce the equation to the desired form by using the identities $\hat{k} \cdot \nabla_{\mathrm{d}} \times (\hat{k} \times \vec{v}) = \nabla_{\mathrm{d}} \cdot \vec{v}$, $\langle \psi \frac{\partial \omega}{\partial t} \rangle = \langle \psi \frac{\partial \wp(\omega)}{\partial t} \rangle$ and $\wp(\omega\hat{k} \times \vec{u}) = \mathcal{I}(\hat{k} \times \wp(\omega)\vec{u})$. The latter two identities can be shown using the identities $\psi = \wp\psi, \vec{u} = \wp\vec{u}$, the self-adjointness of $\wp$ and (22). Combining these results, we have that

$$\left\langle \psi \frac{\partial \wp(\omega)}{\partial t} \right\rangle = -\langle \psi \nabla_{\mathrm{d}} \cdot (\wp(\omega)\vec{u}) \rangle \quad \forall \psi \in \mathcal{V}^1.$$

Thus the ω diagnosed from the solution to (45) and then projected into $\mathcal{V}^1$ satisfies the SEM discretization of (52) and is then locally conserved as in Section 4.2.

4.5. Test Case 5: Conservation

The compatible version of the SEM described here has been implemented in HOMME, the High-Order Method Modeling Environment [41]. HOMME contains both spectral element and discontinuous Galerkin methods for solving two-dimensional and three-dimensional equations on the sphere. Only minor modifications were needed to HOMME's original spectral-element derivative operators and inner product from [37] to make it compatible and thus conservative. We use test case 5 from [42], a suite of well established shallow-water test cases for the sphere, to verify these new conservation properties. Test case 5 was designed to study the effectiveness of a scheme in conserving several integral invariants of the flow. It consists of zonal flow impinging on a mountain. No analytic solution is known but a high-resolution reference solutions is provided by the authors of [42], computed from a T213 spherical-harmonic spectral transform model [43].

We use HOMME to solve Eqs. (45), (46) exactly as written. In order to illustrate best the conservation properties of the method, no additional filters or diffusion terms of any kind are used. HOMME normally uses an explicit leapfrog or semi-implicit time stepping method, both of which require the use of the Robert filter. To remove this filter we replaced the leapfrog method with the leapfrog-trapezoidal method (a trapezoidal method with leapfrog predictor [44]), which has a large stability region that contains the imaginary axis and is easy to implement within a leapfrog based code. We use a time step of 320 s and a grid of 1350 elements with degree-3 polynomial representation (4 × 4 GLL grid) within each element. This grid has an average GLL grid spacing of 2 degrees at the Equator. We purposely choose degree-3 in order to demonstrate that the conservation results hold even when the SEM is run in a low order (by SEM standards) configuration and thus do not rely on low truncation error.

A contour plot of the solution from this configuration after 15 days, is shown in Fig. 3. The l_1,l_2 and l_∞ errors at this final time, normalized as in [42], are 0.00043, 0.00061 and 0.0087 respectively. The l_2 error is close to the uncertainty level in the reference solution (estimated in [3] as 0.00072) and matches the error level obtained by the non-compatible SEM results at 1.4 degree grid spacing in [3].

For this simulation, the discrete total mass is conserved to 15 digits after 15 days. The integrals of divergence and ω (mass weighted potential vorticity) remain below 10^{-20} for the 15 day simulation. We do not expect exact energy conservation, since the compatible SEM conserves energy exactly only with exact time stepping. But any energy conservation errors must be due entirely to the time discretization. On a fixed grid with a second-order accurate time stepping method this error should decrease to machine precision as $\mathcal{O}((\Delta t)^2)$. We show this result in Fig. 4, plotting $(E - E_0)/E_0$, where E is the energy (defined in (51)) after 15 days and E_0 is its initial value. At the smallest time step, $\Delta t = 1$ s, $(E - E_0)/E_0 = -0.95 \times 10^{-13}$. Test case 5 also considers the total potential enstrophy, $P = \langle h\omega^2\rangle$. In HOMME, this quantity is only conserved to truncation error levels. After 15 days, $(P - P_0)/P_0 = 0.00011$, for all time steps used above.

4.6. Test Case 2 and 6: Grid Convergence

For completeness, we present grid convergence results to show that the compatible SEM does not change the formal order of accuracy of the method and remains competitive with global spectral models. The T213 reference solution provided for test case 5 has a large uncertainty. At low resolutions most models, including HOMME, have already converged to within this uncertainty level making the reference solution unusable for grid convergence studies. We will instead use test case 2 and 6. Test case 2 is an analytically known steady state solution of the full nonlinear shallow-water equations. It consists of solid body rotation for the velocity with the corresponding balanced height field. The height field consists of lows over each pole. Test case 6 is an $R = 4$ Rossby-Haurwitz wave which moves from west to east without change of shape in the

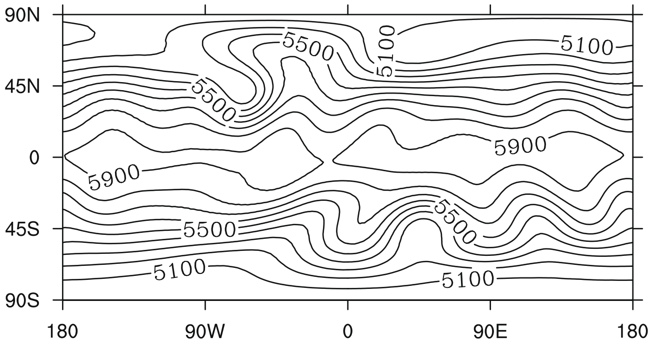

Fig. 3. Contour plot of the test case 5 height field h in meters at 15 days from a 2 degree HOMME simulation. The contour interval is 100 m.

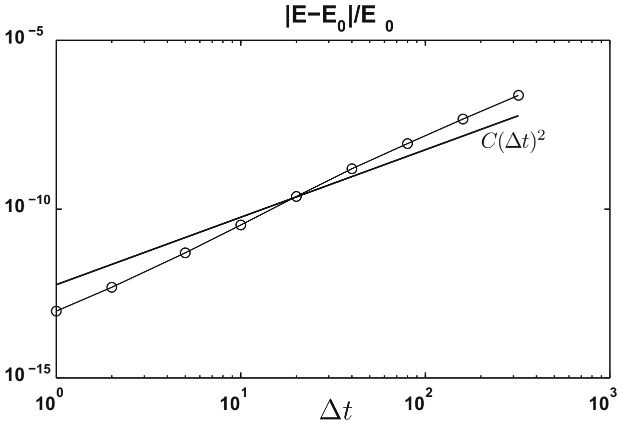

Fig. 4. Relative error in energy conservation after 15 days in shallow water test case 5, using a 2 degree cubed-sphere grid as a function of Δt (seconds). The error converges to near machine precision at better than a second-order rate.

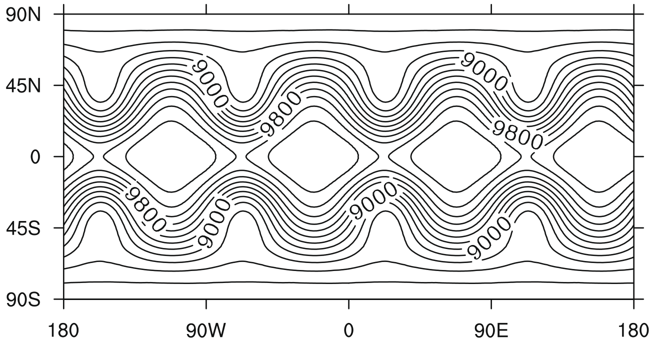

Fig. 5. Contour plot of the test case 6 height field h in meters at 14 days from HOMME (1-degree resolution). The contour interval is 200 m.

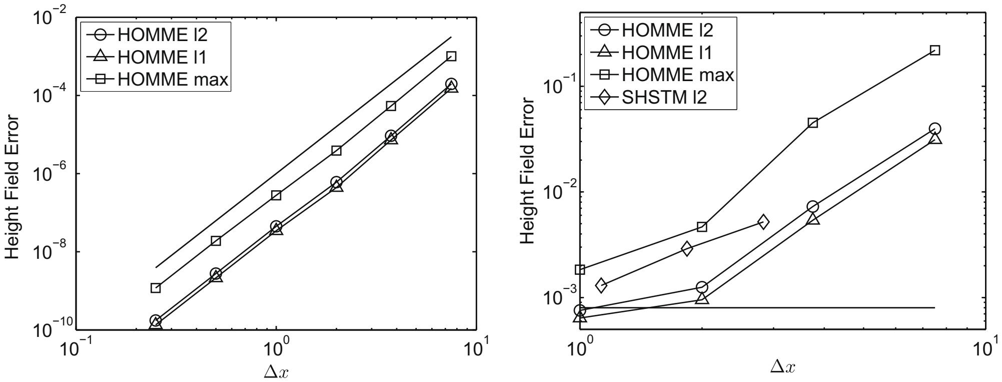

Fig. 6. Mesh convergence in shallow-water test case 2 (left) and case 6 (right). HOMME's l_1,l_2 and l_∞ relative errors are plotted as a function of Δx (degrees), the average grid spacing at the Equator. For test case 2, the convergence rate exactly matches the formal accuracy of the method (4th order) shown by the fainter line. For test case 6, we also show the l_2 errors from a spherical-harmonic spectral transform model (labeled SHSTM) and the l_2 error uncertainty in the reference solution (horizontal line).

non-divergent barotropic equations. This motion is only approximated in the shallow-water equations, so again results from the NCAR T213 spectral model are used as the reference solution.

The simulations are run with the same configuration used with test case 5: leapfrog-trapezoidal time stepping, degree-3 polynomials and no additional filters or diffusion terms. The number of elements used ranges from 96 for the lowest resolution (representing an average grid spacing of the GLL points at the Equator of 7.5 degrees) to 86,400 for the highest resolution (0.25 degree equatorial spacing). The timestep at the lowest resolution is 640 s, decreasing linearly with increasing resolution. A contour plot of the solution from HOMME at 1 degree resolution after 14 days is shown in Fig. 5.

To show grid convergence, in Fig. 6 we plot the l_∞,l_1 and l_2 errors after 5 days (case 2) and 14 days (case 6), normalized as specified in [42]. For case 2, which has an analytical solution, HOMME obtains a convergence rate which matching the formal order of accuracy of the method (4'th order in the case of this configuration of HOMME). As is typical for the realistic shallow-water test cases without analytic solutions, the l_2 error for case 6 converges at a rate less than the formal order of accuracy of the method. Here the convergence rate is approximately $O((\Delta\ x)^{2.6})$. The convergence stops as the l_2 error level approaches the T213 reference solution l_2 error uncertainty (0.0008, from [45,3]) plotted as a horizontal line in the figure. For comparison, the figure also contains data from T42, T63 and T106 simulations [45,43] from the same model used to generate the T213 reference solution. This spectral model converges at a rate close to $O((\Delta x)^{1.5})$, where Δx is taken to be the grid spacing of the transform grid at the equator.

5. The inviscid, incompressible Navier–Stokes equations

The shallow-water energy E includes kinetic energy only due to horizontal 2-velocity $\vec{u}$, but vertical motion $u_r \sim \frac{\partial h}{\partial t}$ is included implicitly by the variable-height factors that make the E integrand trilinear in $\vec{u}$ and h. This trilinearity in turn means (34) is required to obtain (51). In contrast, the incompressible Navier–Stokes equations for 3-velocity $\vec{w} := \vec{u} + u_r\hat{k}$ conserve a simpler quadratic kinetic energy $\frac{1}{2}\vec{w}^2$. However, the replacement of 2D mass conservation (46) by the 3D incompressibility constraint (54) avoids the (34) requirement but complicates numerical energy conservation.

Whereas h uniquely specifies a physical, hydrostatic pressure $p_H = g(H - r)$, the Navier–Stokes pressure p is a Lagrange multiplier for enforcing incompressibility. The discrete weak-form problem is to find $\vec{w}(\cdot, t) \in \mathcal{V}^1_{\text{cov}}$ and $p \in \widetilde{\mathcal{V}}^0$ such that

$$\left\langle \vec{\psi} \cdot \frac{\partial \vec{w}}{\partial t} \right\rangle = \left\langle p \nabla \cdot \vec{\psi} \right\rangle_{\text{G}} - \left\langle \vec{\psi} \cdot \vec{a}(\vec{w}) \right\rangle \quad \forall \vec{\psi} \in \mathcal{V}^1_{\text{con}} \tag{53}$$

$$\text{and,} \quad \langle \psi \nabla \cdot \vec{w} \rangle_{\text{G}} = 0 \quad \forall \psi \in \widetilde{\mathcal{V}}^0, \tag{54}$$

where $\vec{a}(\vec{w}) = (\nabla_{\text{d}} \times \vec{w}) \times \vec{w}$ is nonlinear advection. Because a naïve elimination of $\vec{w}$ between (53) and (54) with $\widetilde{\mathcal{V}}^0 = \mathcal{V}^0$ would lead to spurious modes, typically $\widetilde{\mathcal{V}}^0$ is taken to be (4) but with degree $\tilde{d} = d - 2$, and the $\langle\rangle_{\text{G}}$ terms are approximated by Gauss quadrature [e.g.,[30] ch. 7, [33] ch. 6, [32] ch. 8].To show conservation, we simply take $\vec{\psi} = \mathcal{I}_{\text{con}} \vec{w}$ (18) in (53):

$$\frac{d}{dt} \left\langle \frac{1}{2} \vec{w}^2 \right\rangle = \langle p \nabla \cdot w \rangle_{\text{G}} = 0 \quad \text{if} \quad \psi \to p \quad \text{in (54)}. \tag{55}$$

Unpublished simulations using GASpAR [cf. [31]] confirm (55) holds better than if $\vec{a}(\vec{w}) \to (\vec{w} \cdot \nabla_{\text{d}})\vec{w}$.

6. Conclusions

The spectral-element method is an arbitrarily high-order finite-element method that is very efficient due to its diagonal mass matrix. Here we showed that with a careful treatment of the divergence, gradient and curl operators, the method is *compatible* on nearly arbitrary unstructured grids in three dimensions. It preserves the local (at the element level) adjoint relationships of the discrete divergence, gradient and curl operators with respect to the natural spectral-element inner product and element boundary integral. It also preserves the annihilator properties of these operators. With the primitive-variable rotational form of the shallow-water equations, these compatible properties result in a discretization that locally conserves mass and potential vorticity to machine precision and energy to within the time-truncation error. These results were verified numerically in HOMME, on the cubed-sphere grid, for the standard flow over a mountain and Rossby–Haurwitz shallow-water test cases for the sphere.

Acknowledgments

MT thanks P. Bochev and M. Shashkov for helpful discussions about mimetic methods and J. Drake for suggesting we look at local conservation. AF thanks the U Reading Department of Meteorology and NCAR Turbulence Numerics Team for supporting contributions to this work. This work supported in part by DOE/BER Grant 06-13194.

Appendix A. Matrix expressions

Here we give matrix notation of selected expressions. The set of GLL weights is $\{w_i\}_{i=0}^{d}$. The $\boldsymbol{J}_m$ that multiplies $\boldsymbol{\Gamma}_\tau$ arises from computing $\hat{n} \cdot \vec{g}_\tau$ using (9), (25).

Equation	Expression	Definitions	Descriptions
(7)	$\bar{\boldsymbol{f}}^T \boldsymbol{\Phi}(\vec{r})$	$\bar{\boldsymbol{f}} := (f(\vec{r}_1), \ldots, f(\vec{r}_L))^T$	Unique values of f
		$\boldsymbol{\Phi}(\vec{r}) := (\Phi_1(\vec{r}), \ldots, \Phi_L(\vec{r}))^T \in (\mathcal{V}^0)^L$	Interpolating functions
(15)	$\nabla_{\text{d}} \cdot \vec{v}\vert_{\Omega_m} \boldsymbol{\phi}^T(\vec{x}(\vec{r}; m)) \sum_\alpha \boldsymbol{\Delta}_{\alpha,m} v^\alpha_m$	$\boldsymbol{\Delta}_{\alpha,m} := \mathbf{J}_m^{-1} \mathbf{D}_\alpha \mathbf{J}_m \quad \in \mathbb{R}^{(d+1)^3 \times (d+1)^3}$	Divergence matrix for Ω_m
		$\mathbf{J}_m := \text{diag} \boldsymbol{J}_m \in \mathbb{R}^{(d+1)^3 \times (d+1)^3}$	Jacobian matrix for Ω_m
(16)	$\vec{g}_\alpha \cdot \nabla_{\text{d}} f\vert_{\Omega_m} = \boldsymbol{\phi}^T(\vec{x}(\vec{r}; m)) \mathbf{D}_\alpha \boldsymbol{f}_m$	$\mathbf{D}_\alpha := \mathbf{D} \otimes \mathbf{I} \otimes \mathbf{I} \delta^1_\alpha + \mathbf{I} \otimes \mathbf{D} \otimes \mathbf{I} \delta^2_\alpha$ $+\mathbf{I} \otimes \mathbf{I} \otimes \mathbf{D} \delta^3_\alpha \in \mathbb{R}^{(d+1)^3 \times (d+1)^3}$	Gradient matrix
		$D_{i,j} := \frac{d\varphi_j}{dx} \xi_i) \quad (i,j = 0, \ldots d)$	Derivative matrix
(17)	$\vec{g}^\alpha \cdot \nabla_{\text{d}} \times \vec{v}\vert_{\Omega_m} = \boldsymbol{\phi}^T(\vec{x}(\vec{r}; m)) \sum_\gamma \mathbf{C}^{\alpha\gamma}_m \boldsymbol{v}_{\gamma,m}$	$\mathbf{C}^{\alpha\gamma}_m := \sum_\beta \epsilon^{\alpha\beta\gamma} \mathbf{J}_m^{-1} \mathbf{D}_\beta$	Curl matrix
(19)	$\langle fg \rangle_{\Omega_m} = \boldsymbol{g}^T_m \mathbf{W}_m \boldsymbol{f}_m$	$\mathbf{W}_m := (\mathbf{w} \otimes \mathbf{w} \otimes \mathbf{w}) \mathbf{J}_m \in \mathbb{R}^{(d+1)^3 \times (d+1)^3}$	Mass matrix for Ω_m
		$\mathbf{w} := \text{diag} \boldsymbol{w} \in \mathbb{R}^{(d+1) \times (d+1)}$	Mass matrix for [−1,1]

In the next table, we give matrix expressions of the compatible identities:

Eq.	Expression	Definitions	Descriptions
(27)	$\langle \vec{v} \cdot \hat{n} \rangle_{\partial\Omega_m} = \boldsymbol{J}_m^T \sum_\tau \boldsymbol{\Gamma}_\tau \boldsymbol{v}_m^\tau$	$\boldsymbol{\Gamma}_\tau := \mathbf{B} \otimes \mathbf{w} \otimes \mathbf{w}\delta_\tau^1 + \mathbf{w} \otimes \mathbf{B} \otimes \mathbf{w}\delta_\tau^2$ $+\mathbf{w} \otimes \mathbf{w} \otimes \mathbf{B}\delta_\tau^3 \in \mathbb{R}^{(d+1)^3 \times (d+1)^3}$	Boundary-flux matrix for Ω_m
		$\mathbf{B} := \boldsymbol{e}_d \boldsymbol{e}_d^T - \boldsymbol{e}_0 \boldsymbol{e}_0^T \in \{0, \pm 1\}^{(d+1)\times(d+1)}$	Difference matrix: $\square_-^\gamma$ to $\square_+^\gamma$
(28)	$\mathbf{W}_m \mathbf{D}_\tau + \boldsymbol{\Delta}_{\tau,m}^T \mathbf{W}_m = \mathbf{J}_m \boldsymbol{\Gamma}_\tau \in \mathbb{R}^{(d+1)^3 \times (d+1)^3}$		Divergence theorem for Ω_m
(34)	$\mathbf{Q}^T(\mathbf{W}(\mathbf{I} \otimes \mathbf{D}_\tau) + \boldsymbol{\Delta}_\tau^T \mathbf{W})\mathbf{Q} = \mathbf{Q}^T \mathbf{J}(\mathbf{I} \otimes \boldsymbol{\Gamma}_\tau)\mathbf{Q} = \mathbf{0}$	$\boldsymbol{\Delta}_\tau := \mathrm{diag}_m \boldsymbol{\Delta}_{\tau,m} \in \mathbb{R}^{M_d \times M_d}$	Divergence theorem for Ω
		$\mathbf{J} := \underset{m}{\mathrm{diag}}\, \mathbf{J}_m \in \mathbb{R}^{M_d \times M_d}$	Jacobian matrix for Ω
(36)	$\mathbf{W}_m \mathbf{C}_m^{\alpha\beta} - \mathbf{C}_m^{\beta\alpha T} \mathbf{W}_m = \sum_\tau \epsilon^{\alpha\beta\tau} \boldsymbol{\Gamma}_\tau$		Curl identity for Ω_m
(44)	$\sum_\alpha \mathbf{C}^{\alpha\beta T} \mathbf{W}\mathbf{P}(\mathbf{I} \otimes \mathbf{D}_\alpha) = \mathbf{0}$	$\mathbf{C}^{\alpha\beta} := \underset{m}{\mathrm{diag}}\, \mathbf{C}_m^{\alpha\beta} \in \mathbb{R}^{M_d \times M_d}$	Curl-grad compatibility

References

[1] Y. Maday, A.T. Patera, Spectral element methods for the incompressible Navier–Stokes equations, in: A.K. Noor, J.T. Oden (Eds.), State of the Art Surveys on Computational Mechanics, ASME, New York, 1987, pp. 71–143.

[2] A. Patera, A spectral element method for fluid dynamics: laminar flow in a channel expansion, J. Comput. Phys. 54 (1984) 468–488.

[3] M. Taylor, J. Tribbia, M. Iskandarani, The spectral element method for the shallow water equations on the sphere, J. Comput. Phys. 130 (1997) 92–108.

[4] F.X. Giraldo, A spectral element shallow water model on spherical geodesic grids, Int. J. Numer. Methods Fluids 35 (2001) 869–901.

[5] F.X. Giraldo, T.E. Rosmond, A scalable spectral element Eulerian atmospheric model (SEE-AM) for NWP: dynamical core tests, Mon. Wea. Rev. 132 (2004) 133–153.

[6] A. Fournier, M. Taylor, J. Tribbia, The spectral element atmosphere model (SEAM): high-resolution parallel computation and localized resolution of regional dynamics, Mon. Wea. Rev. 132 (2004) 726–748.

[7] S. Thomas, R. Loft, The NCAR spectral element climate dynamical core: semi-implicit Eulerian formulation, J. Sci. Comput. 25 (2005) 307–322.

[8] H. Wang, J.J. Tribbia, F. Baer, A. Fournier, M.A. Taylor, A spectral element version of CAM, Mon. Wea. Rev. 135 (2007) 3825G3840, doi:10.1175/2007MWR2058.1.

[9] A. St-Cyr, C. Jablonowski, J.M. Dennis, H.M. Tufo, S.J. Thomas, A comparison of two shallow water models with non-conforming adaptive grids, Mon. Wea. Rev. 136 (2008) 1898–1922.

[10] D. Haidvogel, E.N. Curchitser, M. Iskandarani, R. Hughes, M.A. Taylor, Global modeling of the ocean and atmosphere using the spectral element method, Atmos. Ocean Spec. 35 (1997) 505–531.

[11] M. Iskandarani, D. Haidvogel, J. Levin, E.N. Curchitser, C.A. Edwards, Multiscale geophysical modeling using the spectral element method, Comput. Sci. Eng. 4 (2002) 42–48.

[12] A. Molcard, N. Pinardi, M. Iskandarani, D. Haidvogel, Wind driven circulation of the mediterranean sea simulated with a spectral element ocean model, Dyn. Atmos. Oceans 35 (2002) 97–130.

[13] D. Komatitsch, J. Tromp, Spectral-element simulations of global seismic wave propagation - I. Validation, Geophys. J. Int. 149 (2002) 390–412.

[14] J. Tromp, D. Komatitsch, Q. Liu, Spectral-element and adjoint methods in seismology, Commun. Comput. Phys. 3 (1) (2008) 1–32.

[15] T.J.R. Hughes, G. Engel, L. Mazzei, M.G. Larson, The continuous Galerkin method is locally conservative, J. Comput. Phys. 163 (2000) 467–488, doi:10.1006/jcph.2000.6577.

[16] A.A. Samarskiĭ, V.F. Tishkin, A.P. Favorskiĭ, M.Y. Shashkov, Operator-difference schemes, Differentsial'nye Uravneniya 17 (7) (1981) 1317–1327. p. 1344.

[17] L.G. Margolin, A.E. Tarwater, A diffusion operator for lagrangian codes, in: R. Lewis, K. Morgan, W. Habashi (Eds.), Numerical Methods for Heat Transfer, Pineridge Press, Swansea, 1987, pp. 1252–1260.

[18] R. Nicolaides, Direct discretization of planar div-curl problems, SIAM J. Numer. Anal. (1992) 32–56.

[19] M. Shashkov, S. Steinberg, Support-operator finite-difference algorithms for general elliptic problems, J. Comput. Phys. 118 (1995) 131–151.

[20] M. Shashkov, Conservative Finite Difference Methods on General Grids, CRC-Press, Boca Raton, FL, 1996. p. 384.

[21] J.M. Hyman, M. Shashkov, Adjoint operators for the natural discretizations of the divergence, gradient, and curl on logically rectangular grids, Appl. Numer. Math 25 (1997) 413–442.

[22] J.M. Hyman, M. Shashkov, Natural discretizations for the divergence, gradient and curl on logically rectangular grids, Int. J. Appl. Numer. Math. 33 (1997) 81–104.

[23] P. Bochev, M. Hyman, Principles of compatible discretizations, in: D.N. Arnold, P. Bochev, R. Lehoucq, R. Nicolaides, M. Shashkov (Eds.), Compatible Discretizations, Proceedings of IMA Hot Topics Workshop on Compatible Discretizations, vol. IMA 142, Springer Verlag, 2006, pp. 89–120.

[24] A. Gassmann, H.J. Herzog, Towards a consistent numerical compressible non-hydrostatic model using generalized hamiltonian tools, Q.J.R. Meteorol. Soc. 134 (635) (2008) 1597–1613, doi:10.1002/qj.297.

[25] R. Salmon, Poisson-bracket approach to the construction of energy- and potential-enstrophy-conserving algorithms for the shallow-water equations, J. Atmo. Sci. 61 (2004) 2016–2036.

[26] R. Salmon, A general method for conserving energy and potential enstrophy in shallow-water models, J. Atmo. Sci. 64 (2007) 515–531.

[27] J.P. Boyd, Chebyshev and Fourier Spectral Methods: Second Revised Edition, second ed., Dover Publications, 2001. <http://amazon.com/o/ASIN/0486411834/>.

[28] L.G. Margolin, M. Shashkov, Finite volume methods and the equations of finite scale: a mimetic approach, Int. J. Numer. Meth. Fluids 56 (8) (2008) 991–1002, doi:10.1002/fld.1592.

[29] A.J. Simmons, D.M. Burridge, An energy and angular momentum conserving vertical finite-difference scheme and hybrid vertical coordinates, Mon. Wea. Rev. 109 (1981) 758–766.

[30] C. Canuto, M.Y. Hussaini, A. Quarteroni, T. Zang, Spectral Methods: Evolution to Complex Geometries and Applications to Fluid Dynamics, first ed., Springer, 2007.

[31] A. Fournier, D. Rosenberg, A. Pouquet, Dynamically adaptive spectral-element simulations of 2D incompressible Navier–Stokes vortex decays, Geophys. Astrophys. Fluid Dyn. 103 (2) (2009) 245–268.

[32] G.E. Karniadakis, S.J. Sherwin, Spectral/*hp* Element Methods for Computational Fluid Dynamics (Numerical Mathematics and Scientific Computation), second ed., Oxford University Press, USA, 2005.

[33] M.O. Deville, P.F. Fischer, E.H. Mund, High Order Methods for Incompressible Fluid Flow, first ed., Cambridge University Press, 2002.
[34] P. Solin, K. Segeth, I. Dolezel, Higher-Order Finite Element Methods, Chapman & Hall/CRC Press, 2004.
[35] D. Rosenberg, A. Fournier, P. Fischer, A. Pouquet, Geophysical-astrophysical spectral-element adaptive refinement (GASpAR): object-oriented *h*-adaptive fluid dynamics simulation, J. Comp. Phys. 215 (2006) 59–80, doi:10.1016/j.jcp.2005.10.031.
[36] J.H. Heinbockel, Introduction to Tensor Calculus and Continuum Mechanics, Trafford Publishing, Victoria, B.C., 2001.
[37] S. Thomas, R. Loft, Parallel semi-implicit spectral element methods for atmospheric general circulation models, J. Sci. Comput. 15 (2000) 499–518.
[38] R. Sadourny, Conservative finite-difference approximations of the primitive equations on quasi-uniform spherical grids, Mon. Wea. Rev. 100 (2) (1972) 136–144.
[39] M. Rančić, R. Purser, F. Mesinger, A global shallow-water model using an expanded spherical cube: gnomonic versus conformal coordinates, Q.J.R. Meteorol. Soc. 122 (1996) 959–982.
[40] P.E.J. Vos, S.J. Sherwin, R.M. Kirby, From h to p efficiently: Implementing finite and spectral/*hp* element methods to achieve optimal performance for low and high order discretisations, J. Comput. Phys. 229 (2010) 5161–5181.
[41] J. Dennis, A. Fournier, W.F. Spotz, A. St. -Cyr, M.A. Taylor, S.J. Thomas, H. Tufo, High resolution mesh convergence properties and parallel efficiency of a spectral element atmospheric dynamical core, Int. J. High Perf. Comput. Appl. 19 (2005) 225–235.
[42] D.L. Williamson, J.B. Drake, J.J. Hack, R. Jakob, P.N. Swarztrauber, A standard test set for numerical approximations to the shallow water equations in spherical geometry, J. Comput. Phys. 102 (1992) 211–224.
[43] R. Jakob-Chien, J.J. Hack, D.L. Williamson, Spectral transform solutions to the shallow water test set, J. Comput. Phys. 119 (1995) 164–187.
[44] D.R. Durran, The third-order adams-bashforth method: an attractive alternative to leapfrog time differencing, Mon. Wea. Rev. 119 (1991) 702–720.
[45] R. Jakob, J.J. Hack, D.L. Williamson, Solutions to the shallow water test set using the spectral transform method, Technical Report, NCAR/TN-388+STR, National Center for Atmospheric Research, 1993.