DSCS

Dynamics and Statistics of the Climate System, 2016, 1–31
doi:10.1093/climsys/dzw005
Advance Access Publication Date: 22 November 2016
Review Article

OXFORD

Review Article

Compatible finite element spaces for geophysical fluid dynamics

Andrea Natale, Jemma Shipton and Colin J Cotter*

Department of Mathematics, Imperial College London, London SW7 2AZ, UK

*Correspondence Colin J Cotter, Department of Mathematics, Imperial College London, London SW7 2AZ, UK; E-mail: colin.cotter@imperial.ac.uk

Received 21 April 2016; Accepted 21 November 2016

Abstract

Compatible finite elements provide a framework for preserving important structures in equations of geophysical fluid dynamics and are becoming important in their use for building atmosphere and ocean models. We survey the application of compatible finite element spaces to geophysical fluid dynamics, including the application to the non-linear rotating shallow water equations, and the three-dimensional compressible Euler equations. We summarize analytic results about dispersion relations and conservation properties and present new results on approximation properties in three dimensions on the sphere and on hydrostatic balance properties.

Key words: Mixed finite elements, numerical analysis, numerical weather prediction

1. Introduction

Finite element methods have recently become a popular discretization approach for numerical weather prediction, mostly using spectral elements or discontinuous Galerkin methods (Bao *et al.*, 2015; Brdar *et al.*, 2013; Dennis *et al.*, 2011; Fournier *et al.*, 2004; Giraldo *et al.* 2013; Kelly and Giraldo, 2012; Marras *et al.*, 2013; Thomas and Loft, 2005); the use of these methods is reviewed in the work by Marras *et al.* (2015). Finite element methods provide the opportunity to use more general grids in numerical weather prediction, either to improve load balancing in massively parallel computation or to facilitate adaptive mesh refinement. They also allow the development of higher order discretizations. Compatible finite element methods are built from finite element spaces that have differential operators that map from one space to another. They have a long history in both numerical analysis (see Boffi *et al.*, 2013, for a summary of contributions) and applications, including aerodynamics, structural mechanics, electromagnetism and porous media flows. Until recently, they have not been considered much in the numerical weather prediction context, although the lowest Raviart–Thomas element (denoted as RT_0) has been proposed for ocean modelling (Walters and Casulli, 1998), and the RT_0 and lowest order Brezzi–Douglas–Marini element (denoted BDM_1) were both analysed in Rostand and Le Roux (2008) as part of the quest for a mixed finite element pair with good dispersion properties when applied to geophysical fluid dynamics. The realization in Cotter and Shipton (2012) that compatible finite element methods have steady geostrophic modes when applied to the linear shallow water equations, combined with other properties that make them analogous to the Arakawa C-grid staggered finite difference method, has led to an effort to develop these methods for numerical weather prediction, including the Gung Ho dynamical core project in the UK.

In this paper, we review recent work on the development of compatible finite element methods for numerical weather prediction, concentrating on theoretical results about stability, accuracy and conservation properties. In particular, we restrict ourselves to primal-only formulations, i.e. involving a single mesh, on simplicial or tensor product elements, although the use of primal–dual formulation has also been explored in the context of atmosphere simulations using arbitrarily polygonal meshes (Thuburn and Cotter, 2015). Within this survey, we also include some new results, listed below:

- Inclusion of the consideration of inertial modes in the analysis of numerical solutions of the linear rotating shallow water equations on the plane.
- An extension of recent results on approximation theory to triangular prism elements, including approximation properties when solving problems in an 'atmosphere-shaped' spherical shell domain. These are listed below.
    1. For arbitrary non-affine prismatic meshes, there is a loss of consistency of at least 2 degrees for H(curl) elements, 3 degrees for H(div) elements (which are used for velocity in this framework), and 4 degrees for L^2 Elements (which are used for density/pressure in this framework). The consistency loss increases with the polynomial order.
    2. If the prismatic elements are arranged in columns, but with arbitrary slopes on top and bottom faces of each element, then the loss of consistency only affects the polynomial spaces in the horizontal direction.
    3. If the mesh can be obtained from a global continuously differentiable transformation from a smooth triangulation of the sphere extruded into uniform layers, then the expected consistency order of the chosen finite element spaces is obtained. This would arise in the case of a terrain-following mesh with smooth terrain, for example.
    4. The latter result also applies to a cubed sphere mesh under the same conditions.
- An analysis of hydrostatic balance in the discretization of three-dimensional compressible Euler equations.

The rest of this paper is organized as follows. In Section 2, we introduce the concept of Hilbert complex and establish the notation. In Section 3, we describe compatible finite elements in two dimensions and review some relevant analytical results. In Section 4, we review how these results are applied to the linear rotating shallow water equations (including a proper analysis of inertial modes). In Section 5, we review how compatible finite element methods can be used to build discretizations of the non-linear rotating shallow water equations with conservation properties. In Section 6, we describe how to extend compatible finite element spaces to three dimensions for the columnar meshes used in numerical weather prediction (including new results on approximation properties for compatible finite element spaces on columnar meshes). In Section 7, we describe how to treat the discretization of temperature within this framework (including new results on existence and uniqueness of hydrostatic balance). Finally, we give a summary and outlook in Section 8.

2. Hilbert complexes

In this section, we introduce the concept of Hilbert complex, which is the main ingredient to construct compatible finite element spaces, and we establish the notation for the Sobolev spaces used throughout the paper.

Let Ω be an n-dimensional domain with $n \leq 3$, or in general a bounded Riemmannian manifold with or without boundary. Scalar and vector function spaces on Ω can be collected in a diagram which has the following form:

$$0 \to V^0(\Omega) \xrightarrow{\mathrm{d}^0} V^1(\Omega) \xrightarrow{\mathrm{d}^1} \ldots \xrightarrow{\mathrm{d}^{n-1}} V^n(\Omega) \to 0\,, \tag{1}$$

where $V^k(\Omega)$ is a scalar function space for $k = 0, n$ and a vector function space for $0 < k < n$. The operators d^k for $k = 0,\ldots,n-1$ are derivative operators, i.e. they satisfy the Leibniz rule with respect to an opportunely defined product, and they verify $d^k \circ d^{k-1} = 0$ for all $k = 1,\ldots,n-1$. If the spaces $V^k(\Omega)$ are Hilbert spaces, than the sequence in (1) is a 'Hilbert complex'.

If $n = 1$, Ω is locally isomorphic to an open interval in $\mathbb{R}$. If x denotes a local coordinate on Ω, we choose as Hilbert complex on Ω the sequence:

$$0 \to H^1(\Omega) \xrightarrow{\frac{d}{dx}} L^2(\Omega) \to 0\,. \tag{2}$$

For $n = 2$, we consider the Hilbert complex:

$$0 \to H^1(\Omega) \xrightarrow{\nabla^\perp} H(\mathrm{div};\Omega) \xrightarrow{\nabla\cdot} L^2(\Omega) \to 0\,. \tag{3}$$

Finally, for $n = 3$, we consider the Hilbert complex:

$$0 \to H^1(\Omega) \xrightarrow{\nabla} H(\mathrm{curl};\Omega) \xrightarrow{\nabla\times} \mathrm{H}(\mathrm{div};\Omega) \xrightarrow{\nabla\cdot} \mathrm{L}^2(\Omega) \to 0\,. \tag{4}$$

From a numerical perspective reproducing such complex structures is key in achieving a number of properties that are crucial for geophysical fluid dynamics simulations, as it will be shown in the following sections.

It will be useful to have a unified notation to denote the spaces and the differential operators in the diagrams in (2), (3) and (4). For this purpose, we introduce the following diagram:

$$0 \to HV^0(\Omega) \xrightarrow{\mathrm{d}^0} HV^1(\Omega) \xrightarrow{\mathrm{d}^1} \ldots \xrightarrow{\mathrm{d}^{n-1}} HV^n(\Omega) \to 0\,, \tag{5}$$

where the spaces $HV^k(\Omega)$ is the k^{th} space in any of the sequences in (2), (3) or (4), depending on the dimension n of Ω. We denote by $\|\cdot\|_{L^2(\Omega)}$ (or simply $\|\cdot\|_{L^2}$ when the domain is clear from the context) the L^2 norm on any of these spaces, and we do not change notation depending on whether $V^k(\Omega)$ contains scalar or vector functions.

The spaces $HV^k(\Omega)$ are Hilbert spaces with respect to the norm $\|\cdot\|_{HV^k(\Omega)}$, which is defined by:

$$\|u\|^2_{HV^k(\Omega)} := \|u\|^2_{L^2(\Omega)} + \|\mathrm{d}^k u\|^2_{L^2(\Omega)}\,, \tag{6}$$

for all $u \in HV^k(\Omega)$. For example, we have for $n = 2, 3$:

$$\|u\|^2_{HV^{n-1}(\Omega)} = \|u\|^2_{H(\mathrm{div};\Omega)} := \|u\|^2_{L^2(\Omega)} + \|\nabla\cdot u\|^2_{L^2(\Omega)}\,. \tag{7}$$

Furthermore, we define $H^sV^k(\Omega)$ to be the spaces of scalar or vector functions with components in $H^s(\Omega)$. The norm associated to any of these spaces is defined by summing up the norms of each component in $H^s(\Omega)$. Analogous definitions hold for the spaces $W^s_pV^k(\Omega)$.

3. Compatible finite element spaces in two dimensions

In this section, we provide a brief introduction to compatible finite element spaces in two dimensions and review some of their properties. We start with a definition.

Definition 3.1. *Let* $V^0_h(\Omega) \subset H^1(\Omega)$, $V^1_h(\Omega) \subset H(\mathrm{div};\Omega)$, $V^2_h(\Omega) \subset L^2(\Omega)$ *be a sequence of finite element spaces* ($V^0_h(\Omega)$ *and* $V^2_h(\Omega)$ *contain scalar valued functions, while* $V^1_h(\Omega)$ *contains vector valued functions). These spaces are called compatible if:*

1. $\nabla^\perp \psi \in V^1_h(\Omega), \forall \psi \in V^0_h(\Omega)$,
2. $\nabla\cdot u \in V^2_h(\Omega), \forall u \in V^1_h(\Omega)$, *and*
3. *there exist bounded projections* π^i, $i = 0, 1, 2$, *such that the following diagram commutes.*

$$\begin{tikzcd} H^1(\Omega) \arrow[r, "\nabla^\perp"] \arrow[d, "\pi^0"] & H(\mathrm{div};\Omega) \arrow[r, "\nabla\cdot"] \arrow[d, "\pi^1"] & L^2(\Omega) \arrow[d, "\pi^2"] \\ V^0_h(\Omega) \arrow[r, "\nabla^\perp"] & V^1_h(\Omega) \arrow[r, "\nabla\cdot"] & V^2_h(\Omega) \end{tikzcd} \tag{8}$$

The use of these commutative diagrams to prove stability and convergence of mixed finite element methods for elliptic problems has a long history, as detailed in Boffi *et al.* (2013). More recently, this structure has been translated into the language of differential forms. This provides a unifying framework that relates properties between the various spaces and for different dimensions, called finite element exterior calculus (Arnold et al., 2006, 2010). In computational electromagnetism, the term 'discrete differential forms' is used to denote this choice of finite element spaces (Hiptmair, 2002). Here, since we use the language of vector calculus, we use the term compatible finite element spaces.

If the domain Ω has a boundary $\partial\Omega$, then we define $\mathring{H}^1(\Omega)$ and $\mathring{V}_h^0(\Omega)$ as the subsets of $H^1(\Omega)$ and $V_h^0(\Omega)$, where functions vanish on $\partial\Omega$, respectively. Similarly, we define $\mathring{H}(\mathrm{div};\Omega)$ and $\mathring{V}_h^1(\Omega)$ as the subsets of $H(\mathrm{div};\Omega)$ and $V_h^1(\Omega)$, where the normal components of functions vanish on $\partial\Omega$, respectively. Then, the above commutative diagram still holds with subspaces appropriately substituted. Since these are the required coastal boundary conditions for stream function and velocity, we shall assume that we use these subspaces throughout the rest of the paper, but shall drop the $\mathring{\cdot}$ notation for brevity.

A large number of these sets of spaces can be found in the Periodic Table of Finite Elements (Arnold and Logg, 2014). Examples include:

- (CG_k,RT_{k-1},DG_{k-1}) on triangular meshes for $k > 0$, where CG_k denotes the continuous Lagrange elements of degree k, RT_k denotes the Raviart–Thomas elements of degree k, and DG_k denotes the discontinuous Lagrange elements of degree k. (This corresponds to the $\mathcal{P}_k^-\Lambda^r$ families in the finite element exterior calculus.)
- $(CG_k,BDM_{k-1},DG_{k-2})$ on triangular meshes for $k > 1$, where BDM_k is the Brezzi–Fortin–Marini element of degree k. (This corresponds to the $\mathcal{P}_k\Lambda^r$ families in the finite element exterior calculus.)
- (CG_k,RT_{k-1},DG_{k-1}) on quadrilateral meshes for $k > 0$, where RT_k now denotes the quadrilateral Raviart–Thomas elements. [This corresponds to the $\mathcal{Q}_k^-\Lambda^r$ spaces in the finite element exterior calculus (Arnold *et al.*, 2014)].

While explicit constructions of the triangular mesh examples above exist on affine meshes, the approach that is usually taken computationally is to define these elements on a canonical reference cell (triangle or square) and then to define elements on each physical cell via the pullback mappings listed in Table 1. More precisely, if F_T is a mapping from the reference cell $\widehat{T}$ to the physical cell T, then finite element functions are related via:

- $\psi \in V_h^0(\widehat{T}) \Rightarrow \psi \circ F_T^{-1} \in V_h^0(T)$,
- $u \in V_h^1(\widehat{T}) \Rightarrow Ju/\det(J) \circ F_T^{-1} \in V_h^1(T)$,
- $\rho \in V_h^2(\widehat{T}) \Rightarrow \rho/\det(J) \circ F_T^{-1} \in V_h^2(T)$,

where J is the Jacobian matrix DF_T. More precisely, given a decomposition $\mathcal{T}_h$ of the domain Ω in elements T, with maximum element diameter h, the finite element space $V_h^k(\Omega)$ on $\mathcal{T}_h$ is defined by requiring $V_h^k(T) = F_T^{-1*}(V_h^k(\hat{T}))$ and $V_h^k(\Omega) \subset HV^k(\Omega)$. The last requirement enforces the correct inter-element continuity accordingly to the degree k in the complex.

The transformation for $V_h^1(\Omega)$ is referred to as the contravariant Piola transformation for $H(\mathrm{div})$ elements. Note that for $V_h^1(\Omega)$, scaling of edge basis functions is also necessary to achieve the correct inter-element continuity; see Rognes *et al.* (2009) for details of how to implement this efficiently. This transformation can be extended to surfaces embedded in three dimensions so that the $V_h^1(\Omega)$ vector fields remain tangential to the surface; see Rognes *et al.* (2013), for details of efficient implementation, including examples on the sphere.

It is also useful to define dual operators to the $\nabla^\perp$ and $\nabla\cdot$ operators through integration by parts; these allow us to approximate the curl of functions in V_h^1 and the gradient of functions in V_h^2.

Table 1. $F^{-1*}v$ for $1 \le n \le 3$ and $0 \le k \le 3$. J is the Jacobian matrix DF of the map F.

	$k = 0$	$k = 1$	$k = 2$	$k = 3$
$n = 1$	$v \circ F^{-1}$	$v/\det(J) \circ F^{-1}$		
$n = 2$	$v \circ F^{-1}$	$Jv/\det(J) \circ F^{-1}$	$v/\det(J) \circ F^{-1}$	
$n = 3$	$v \circ F^{-1}$	$J^{-T}v \circ F^{-1}$	$Jv/\det(J) \circ F^{-1}$	$v/\det(J) \circ F^{-1}$

Definition 3.2. *Let* $v \in V_h^1(\Omega)$, $\rho \in V_h^2(\Omega)$. *Then define* $\widetilde{\nabla}^{\perp} \cdot v \in V_h^0(\Omega)$ *and* $\widetilde{\nabla}\rho \in V_h^1(\Omega)$ via:

$$\int_\Omega \gamma \widetilde{\nabla}^{\perp} \cdot v dx = -\int_\Omega \nabla^{\perp}\gamma \cdot v dx, \quad \forall \gamma \in V_h^0(\Omega), \tag{9}$$

$$\int_\Omega w \cdot \widetilde{\nabla}\rho dx = -\int_\Omega \nabla \cdot w \rho dx, \quad \forall w \in V_h^1(\Omega). \tag{10}$$

A very important result about compatible finite element spaces is the existence of a Helmholtz decomposition for $V_h^1(\Omega)$.

Theorem 3.1. Helmholtz decomposition. *Let* Ω *be a suitably smooth (see* Arnold *et al.*, 2010 *for details) domain, and let* $V_h^0(\Omega)$, $V_h^1(\Omega)$, $V_h^2(\Omega)$ *be compatible finite element spaces on* Ω. *Define the subspaces*:

$$B_h = \{u \in V_h^1(\Omega) \, : \, u = \nabla^{\perp}\psi, \, \psi \in V_h^0(\Omega)\}, \tag{11}$$

$$B_h^* = \{u \in V_h^1(\Omega) \, : \, u = \widetilde{\nabla}\phi, \, \phi \in V_h^2(\Omega)\}, \tag{12}$$

$$\mathfrak{h}_h = \left\{u \in V_h^1(\Omega) \, : \, \nabla \cdot u = 0, \, \widetilde{\nabla}^{\perp} \cdot u = 0\right\}. \tag{13}$$

Then,

$$V_h^1(\Omega) = B_h \oplus \mathfrak{h}_h \oplus B_h^*, \tag{14}$$

where $\oplus$ *indicates an orthogonal decomposition with respect to the* L^2 *inner product.*

Further, the space $\mathfrak{h}_h$ *of harmonic functions of* $V_h^1(\Omega)$ *has the same dimension as the corresponding space* $\mathfrak{h}$ *of harmonic functions of* $H(\mathrm{div};\Omega)$ *(which is determined purely from the topology of* Ω*).*

Proof. See Arnold *et al.* (2006). □

3.1. Mixed discretization for the Poisson equation

The mixed discretization of the Poisson equation:

$$-\nabla^2 p = f, \quad \int_\Omega p dx = 0, \quad \frac{\partial p}{\partial n} = 0 \text{ on } \partial\Omega, \tag{15}$$

using compatible finite elements is very well studied. Defining

$$\overline{V}_h^2(\Omega) = \left\{p \in V_h^2 : \int_\Omega p dx = 0\right\}, \tag{16}$$

we seek $u^h \in V_h^1(\Omega)$ and $p^h \in \overline{V}_h^2(\Omega)$, such that

$$\int_\Omega w^h \cdot u^h dx + \int_\Omega (\nabla \cdot w^h) p^h dx = 0, \quad \forall w^h \in V_h^1(\Omega), \tag{17}$$

$$\int_\Omega \phi^h \nabla \cdot u^h dx = \int_\Omega \phi^h f dx, \quad \forall \phi^h \in \overline{V}_h^2(\Omega). \tag{18}$$

The compatible finite element structure between $V_h^1(\Omega)$ and $V_h^2(\Omega)$ is behind the proof of the following results, all of which can be found in Boffi *et al.* (2013). The first result is about the stability of the $\widetilde{\nabla}$ operator defined in Definition 3.2.

Theorem 3.2. (inf-sup condition). *Define the bilinear form*:

$$b(w,p)=\int\limits_{\Omega}(\nabla\cdot w)p\,dx. \tag{19}$$

There exists $c > 0$*, independent of mesh resolution, such that:*

$$\inf_{p\in\overline{V}_h^2(\Omega)}\sup_{w\in V_h^1(\Omega)}\frac{b(w,p)}{\|w\|_{H(div)}\|p\|_{L^2}}\geq c, \tag{20}$$

where

$$\|w\|_{H(\mathrm{div})}^2=\|w\|_{L^2}^2+\|\nabla\cdot w\|_{L^2}^2. \tag{21}$$

Proof. The result is a particular case of Proposition 5.4.2 in Boffi *et al.* (2013). ☐

This result is then key to the proof of the following theorem.

Theorem 3.3. *Equations (17) and (18) have a unique solution* (u^h,p^h)*, satisfying*

$$\|u^h\|_{H(\mathrm{div})}\leq a_1\|f\|_{L^2}, \tag{22}$$

$$\|p^h\|_{L^2}\leq a_2\|f\|_{L^2}, \tag{23}$$

$$\|u-u^h\|_{H(\mathrm{div})}+\|p-p^h\|_{L^2}\leq a_3\|f-f^h\|_{L^2}, \tag{24}$$

for positive constants a_1, a_2, a_3 *independent of mesh resolution, where* p *is the exact solution to the Poisson equation,* $u = -\nabla p$*, and* f^h *is the* L^2 *projection of* f *into* $V_h^2(\Omega)$.

Proof. This is a particular case of the abstract results in Chapters 4 and 5 of Boffi *et al.* (2013). See in particular Theorems 4.3.2, 5.2.1 and 5.2.5. ☐

Finally, we have the following convergence result for eigenvalues.

Theorem 3.4. *Let* $\lambda_0 < \lambda_1 < \ldots < \lambda_i < \ldots$ *be the eigenvalues of the Laplacian operator* $-\nabla^2$*, with corresponding eigenspaces* $\{E_i\}_{i=0}^{\infty}$.

Let $\lambda_0^h<\lambda_1^h<\ldots<\lambda_m^h$*, be the complete set of eigenvalues of the discrete Laplacian operator* $-\nabla\cdot\widetilde{\nabla}$*, with corresponding eigenspaces* $\{E_i^h\}_{i=0}^{m}$.

Then $\forall\epsilon > 0$, $\forall\, k\in\mathbb{N}$, $\exists h_0 > 0$*, such that* $\forall h \leq h_0$,

$$\max_{i=0,\ldots,k}|\lambda_i-\lambda_i^h|\leq\epsilon, \tag{25}$$

$$\hat{\delta}\left(\oplus_{i=0}^k E_i,\oplus_{i=0}^k E_i^h\right)\leq\epsilon, \tag{26}$$

where

$$\hat{\delta}(E,F)=\max(\delta(E,F),\delta(F,E)), \tag{27}$$

$$\delta(E,F)=\sup_{p\in E,\|p\|_{L^2}=1}\inf_{\phi\in F}\|p-\phi\|_{L^2}. \tag{28}$$

Proof. The result follows from applying the abstract results of Section 6.5 in Boffi *et al.* (2013), and in particular as a consequence of Theorem 6.5.1. ☐

4. Compatible finite element methods for the linear shallow water equations

In this section, we summarize some properties of compatible finite element methods applied to the linear shallow water equations. The equations take the form:

$$u_t + fu^{\perp} + g\nabla h = 0, \tag{29}$$

$$h_t + H\nabla \cdot u = 0, \tag{30}$$

where the $\perp$ operator maps (u_1,u_2) to $(-u_2,u_1)$, $u_t := \partial u/\partial t$, h is the layer height, and f is the (possibly spatially dependent) Coriolis parameter, g is the acceleration due to gravity and H is a constant representing the reference depth. This system of equations is a toy model of barotropic waves in the atmosphere, and it is used to explore important properties of numerical discretizations.

The compatible finite element spatial discretisation for the shallow water equations takes $u \in V_h^1(\Omega)$, $h \in V_h^2(\Omega)$, and solves:

$$\int_{\Omega} w \cdot u_t dx + \int_{\Omega} fw \cdot u^{\perp} dx - g\int_{\Omega} (\nabla \cdot w)h dx = 0, \quad \forall w \in V_h^1(\Omega), \tag{31}$$

$$\int_{\Omega} \phi(h_t + H\nabla \cdot u)dx = 0, \quad \forall \phi \in V_h^2(\Omega). \tag{32}$$

Staniforth and Thuburn (2012) provided a number of desirable properties for numerical discretizations when applied to the linear shallow water equations, which we consider in the following subsections. We note that all of these results apply on arbitrarily unstructured grids.

4.1. Energy conservation

Combining (31) and (32), with $w = u$ in the first and $\phi = h$ in the second, shows that the energy:

$$E = \frac{1}{2}\int_{\Omega} H|u|^2 + gh^2 dx, \tag{33}$$

is conserved by these solutions. Note that in the fully discrete system, i.e. when time is also discretized, the level of energy conservation will depend upon the chosen scheme and time step.

4.2. Stability of pressure gradient operator

It is crucial that the discretization of the pressure gradient term ∇h is stable; this means that there are no functions h for which h is large amplitude, but the discrete approximation to ∇h is small. In other words, there are no spurious eigenvalues for the discrete Laplacian obtained by combining ∇ with $\widetilde{\nabla}$. For mixed finite element discretizations, this requirement is formalized by the inf-sup condition, which we have already stated in the previous section. In fact, further than this, the eigenvalue convergence result for $\nabla \cdot \widetilde{\nabla}$ shows that the wave equation with $f = 0$ has convergent eigenvalues.

4.3. Geostrophic balance

A crucial property for large-scale modelling is that the linearized equations on the f-plane (*i.e.* Ω is the periodic plane and f is constant) have geostrophic steady states (Staniforth and Thuburn, 2012). This is proved in the following.

Theorem 4.1. *For all divergence-free* $u \in B_h$, *there exists* h *such that* (u,h) *is a steady-state solution (31) and (32)* (Cotter and Shipton, 2012).

Proof. Since $\nabla \cdot u = 0$, (32) immediately implies that $h_t = 0$. Since, $u \in B_h$ and so there exists $\psi \in V_h^0(\Omega)$ such that $u = \nabla^{\perp}\psi$. Take $h \in V_h^2(\Omega)$ according to:

$$\int_{\Omega} \phi gh dx = \int_{\Omega} \phi f\psi dx \quad \forall \phi \in V_h^2(\Omega). \tag{34}$$

Then,

$$\int_\Omega w\cdot u_t dx = \int_\Omega fw\cdot\nabla\psi + (\nabla\cdot w)ghdx, \tag{35}$$

$$= \int_\Omega -(\nabla\cdot w)f\psi + (\nabla\cdot w)ghdx, \tag{36}$$

$$= \int_\Omega -(\nabla\cdot w)gh + (\nabla\cdot w)ghdx = 0, \tag{37}$$

after integrating by parts, which is permitted since the product $\psi w\in H(\mathrm{div})$. □

If this property is not satisfied, then a solution with balanced initial condition oscillates on the fast timescale. This is catastrophic when the non-linear terms are introduced since the solution is polluted by rapid oscillations before the slow balanced state has had time to evolve; this renders a discretization useless for large-scale atmosphere or ocean modelling.

4.4. Absence of spurious inertial oscillations

Inertial oscillations are solutions of the shallow water equations on the f-plane with $h = 0$. For the continuous equations, (30) implies that $\nabla\cdot u = 0$. Then, (29) implies that:

$$u_t + fu^\perp = 0. \tag{38}$$

Taking the divergence then implies that $\nabla^\perp\cdot u = 0$. Therefore, u is a harmonic function. On the periodic plane, $\mathfrak{h}$ is two-dimensional, consisting of all spatially constant velocity fields, and the solution just rotates around with frequency f.

Some discretizations, i.e. the $DG_k - CG_{k-1}$ discretization considered in Cotter and Ham (2011), have subspaces of spurious inertial oscillations, i.e. extra divergence-free oscillatory solutions with $h = 0$, which are not spatially uniform. These are known to cause spurious grid scale oscillations in regions of high shear such as representations of the Gulf Stream in ocean models (S Danilov, personal communication). The staggered C-grid finite difference method is free from such modes. It does, however, have a steady divergence-free solution in the kernel of the Coriolis operator, which is known as a 'Coriolis mode'. We shall see in this section that similar modes can exist for compatible finite element methods.

Inertial oscillations for compatible finite element discretizations were not properly considered by Cotter and Shipton (2012), so we consider them here. We start by observing that for a compatible discretisation $\mathfrak{h}_h\equiv\mathfrak{h}$, i.e. it is the space of constant vector fields v_0. To see this, first note that $\nabla\cdot v_0 = 0$. Secondly,

$$\int_\Omega \nabla^\perp\gamma\cdot v_0 dx = -\int_\Omega \gamma\underbrace{\nabla^\perp\cdot v_0}_{=0}dx = 0, \quad \forall\gamma\in V_h^0(\Omega), \tag{39}$$

after integration by parts (since γ and v_0 are both continuous), and therefore $\tilde{\nabla}^\perp\cdot v_0 = 0$. Therefore, $v_0\in\mathfrak{h}_h$. Since $\dim(\mathfrak{h}_h) = \dim(\mathfrak{h}) = 2$ by Theorem 1, we have therefore characterized all discrete harmonic functions in $V_h^1(\Omega)$.

Theorem 4.2. *The only time-varying solutions of the compatible finite element discretization of the shallow water equations on the f-plane with divergence-free u and h = 0 are the solutions with spatially constant u satisfying $u_t + fu^\perp = 0$. Any time-independent solutions are in the kernel of the discrete Coriolis operator,* i.e.:

$$\int_\Omega w\cdot u^\perp dx = 0, \quad \forall w\in V_h^1(\Omega). \tag{40}$$

Proof. To see that spurious inertial oscillations do not exist for compatible finite element discretizations, we first note that $h = 0$ implies that $\nabla\cdot u = 0$. Then, (31) implies that:

$$\int_\Omega w\cdot u_t dx + f\int_\Omega w\cdot u^\perp dx = 0, \quad \forall w\in V_h^1(\Omega). \tag{41}$$

Taking $w = \nabla^{\perp}\gamma$ for $\gamma \in V_h^0(\Omega)$, we obtain:

$$\int_{\Omega} \nabla^{\perp}\gamma \cdot u_t dx = -f \int_{\Omega} \nabla \gamma \cdot u dx, \tag{42}$$

$$= f \int_{\Omega} \gamma \nabla \cdot u dx = 0, \quad \forall \gamma \in V_h^0(\Omega), \tag{43}$$

where we may integrate by parts for the same reason as in the geostrophic balance calculation. This means that u_t is orthogonal to everything in B_h, i.e. $u_t \in \mathfrak{h}_h$. Hence, we may write $u = k + u_0$, where $k \in \mathfrak{h}_h$, and u_0 is divergence-free and independent of time. Therefore, by the geostrophic balance condition, there exists $h_0 \in V_h^2(\Omega)$ such that:

$$f \int_{\Omega} w \cdot u_0^{\perp} dx - g \int_{\Omega} (\nabla \cdot w) h_0 dx = 0, \quad \forall w \in V_h^1(\Omega), \tag{44}$$

and therefore substitution of $u = k + u_0$ into (41) leads to

$$\int_{\Omega} w \cdot k_t dx = -f \int_{\Omega} w \cdot k^{\perp} dx + g \int_{\Omega} (\nabla \cdot w) h_0 dx, \quad \forall w \in V_h^1(\Omega). \tag{45}$$

Since $\mathfrak{h}_h$ is the space of constant vector fields, if $k \in \mathfrak{h}_h$, then so is $k^{\perp}$. Therefore, both k_t and $k^{\perp}$ are orthogonal to B_h^*, and therefore

$$\int_{\Omega} (\nabla \cdot w) h_0 dx = 0, \quad \forall w \in B_h^*. \tag{46}$$

This means that if we define $v \in V_h^1(\Omega)$ via

$$\int_{\Omega} w \cdot v dx = -\int_{\Omega} (\nabla \cdot w) h_0 dx, \quad \forall w \in V_h^1(\Omega), \tag{47}$$

then $\nabla \cdot v = 0$. This means that $v \in V_h^1(\Omega)$ and $h_0 \in V_h^2(\Omega)$ satisfy the following coupled system of equations,

$$\int_{\Omega} w \cdot v dx + \int_{\Omega} (\nabla \cdot w) g h_0 dx = 0, \quad \forall w \in V_h^1(\Omega), \tag{48}$$

$$-\int_{\Omega} \phi \nabla \cdot v dx = 0, \quad \forall \phi \in V_h^2(\Omega). \tag{49}$$

This is just the mixed finite element discretization of the Poisson equation with 0 on the right-hand side, and therefore $h_0 = 0$ by uniqueness. Therefore, we have

$$\int_{\Omega} w \cdot u_0^{\perp} dx = 0, \quad \forall w \in V_h^1(\Omega). \tag{50}$$

This means that u_0 must be in the kernel of the Coriolis operator.

Finally, we have

$$\int_{\Omega} w \cdot k_t dx + \int_{\Omega} f w \cdot k^{\perp} dx = 0, \quad \forall w \in \mathfrak{h}_h \subset V_h^1(\Omega), \tag{51}$$

which implies the pointwise equation $k_t + fk^{\perp} = 0$, since the L^2-projection onto $\mathfrak{h}_h$ is trivial. □

4.5. Inertia–gravity waves

Continuing the analysis of solutions in the periodic f-plane, Cotter and Shipton (2012) considered the time-dependent divergent solutions of (31) and (32), i.e. the inertia–gravity wave solutions. One way to calculate the equation for these modes is to make a discrete Helmholtz decomposition of both the solution u and the test function w,

$$u = \nabla^{\perp} \psi + \widetilde{\nabla} \phi + k, \quad \psi \in \overline{V}_h^0(\Omega),\ \phi \in \overline{V}_h^2(\Omega),\ k \in \mathfrak{h}_h, \tag{52}$$

$$w = \nabla^{\perp} \gamma + \widetilde{\nabla} \alpha + l, \quad \gamma \in \overline{V}_h^0(\Omega),\ \alpha \in \overline{V}_h^2(\Omega),\ l \in \mathfrak{h}_h, \tag{53}$$

where

$$\overline{V}_h^0(\Omega) = \left\{ \psi \in V_h^0 : \int_{\Omega} \psi dx = 0 \right\}, \tag{54}$$

and substitute into the equations, using orthogonality to obtain

$$\int_{\Omega} \nabla \gamma \cdot \nabla \psi_t - f \nabla \gamma \cdot \widetilde{\nabla} \phi dx = 0, \quad \forall \gamma \in \overline{V}_h^0(\Omega), \tag{55}$$

$$\int_{\Omega} \widetilde{\nabla} \alpha \cdot \widetilde{\nabla} \phi_t + f \widetilde{\nabla} \alpha \cdot (\widetilde{\nabla} \phi)^{\perp} - \widetilde{\nabla} \alpha \cdot \nabla \psi - g(\nabla \cdot \widetilde{\alpha}) h dx = 0, \quad \forall \alpha \in \overline{V}_h^2(\Omega), \tag{56}$$

where we have made use of the fact that the discrete harmonic functions consists of constant vector fields. Taking the time derivative of the second equation and substituting $h_t = -H\nabla\cdot u$, which holds pointwise since $\nabla \cdot \boldsymbol{u} \in V_h^2(\Omega)$, we obtain

$$\int_{\Omega} \widetilde{\nabla} \alpha \cdot \widetilde{\nabla} \phi_{tt} + f \widetilde{\nabla} \alpha \cdot (\widetilde{\nabla} \phi)_t^{\perp} - \widetilde{\nabla} \alpha \cdot \nabla \psi_t + gH \nabla \cdot \widetilde{\nabla} \alpha \nabla \cdot \widetilde{\nabla} \phi dx = 0,\ \forall \alpha \in \overline{V}_h^2(\Omega).$$

Finally, we define a projection $P : \overline{V}_h^2(\Omega) \rightarrow \overline{V}_h^0(\Omega)$. by

$$\int_{\Omega} \nabla \gamma \cdot \nabla P\phi = \int_{\Omega} \nabla \gamma \cdot \widetilde{\nabla} \phi dx, \quad \forall \gamma \in \overline{V}_h^0(\Omega),$$

which is solvable since the left-hand side is the standard continuous finite element discretization of the Laplace operator. Then, the equation becomes

$$\int_{\Omega} \widetilde{\nabla} \alpha \cdot \widetilde{\nabla} \phi_{tt} + f^2 \nabla P\alpha \cdot \nabla P\phi + gH \nabla \cdot \widetilde{\nabla} \alpha \nabla \cdot \widetilde{\nabla} \phi \mathrm{d}x = -\int_{\Omega} f \widetilde{\nabla} \alpha \cdot (\widetilde{\nabla} \phi)_t^{\perp} \mathrm{d}x, \quad \forall \alpha \in \overline{V}_h^2(\Omega). \tag{59}$$

Upon inspection, when $f = 0$ this is just the mixed compatible finite element discretization of the wave equation, which we know to have convergent eigenvalues from the previous section.

The continuous form of this equation is

$$\nabla^2\left(\frac{\partial^2}{\partial tt}+f^2\right)\phi + gH\,\nabla^2\,\nabla^2\,\phi = 0. \tag{60}$$

For $f \neq 0$, we might hope that the right-hand side of the discretized equation is small for resolved modes. However, the projection is more problematic, since $V_h^0(\Omega)$ and $V_h^2(\Omega)$ might have different dimensions. If $\dim\,(V_h^0(\Omega)) < \dim\,(V_h^2(\Omega))$, then some modes are projected out and the f^2 contribution (which originates from the Coriolis term) is invisible to those modes. Hence, Cotter and Shipton (2012) concluded that $\dim\,(V_h^0(\Omega)) \geq \dim\,(V_h^2(\Omega))$ is a necessary (but not sufficient) condition for the absence of spurious inertia–gravity modes. This necessary condition is not satisfied for the RT_k family on triangles. For example, when $k = 0$, we have $\dim\,(V_h^2(\Omega)) = 2\dim\,(V_h^0(\Omega))$. This leads to the spurious inertia–gravity mode behaviour that was observed by Danilov (2010) for $RT_0 - DG_0$ and the related C-grid staggered finite difference method (obtained by lumping the mass in the $RT_0 - DG_0$ discretization). On the other hand, the BDM_k family does satisfy this necessary condition. Cotter and Shipton (2012) noted that the $BDFM_1$ spaces exactly satisfy $\dim\,(V_h^2(\Omega)) = \dim\,(V_h^0(\Omega))$,which might be additionally desirable, particularly since they also analysed the Rossby wave equation (not reproduced here) and concluded that $\dim\,(V_h^2(\Omega)) \leq \dim\,(V_h^0(\Omega))$ is a necessary condition for the absence of spurious Rossby waves. However, it is not clear that spurious Rossby waves are damaging for numerical simulations, since high spatial frequency Rossby waves do not propagate anyway.

To explore sufficient conditions for absence of spurious inertia–gravity modes, it might only be possible to make progress by assuming symmetric grids and performing von Neumann analysis. On triangular grids this is very challenging, but the analysis has been successfully carried out for RT_0. and BDM_1 spaces in Rostand and Le Roux (2008); it is still unclear how to dealias the high mode branch identified for BDM_1, but we believe that they are physical modes. On rectangular elements, it is much easier since one can exploit the tensor product structure. Staniforth *et al.* (2013) showed that (under the assumption of y-independent solutions) the $RT_0 - DG_0$ space is free from spurious modes and has a similar, but more accurate, dispersion relation to the finite difference C-grid staggering. They investigated, and found two mode branches, a high branch and a low branch. Although both of these branches correspond to physical modes, there is a small jump between the branches, where the group velocity goes to 0. This might be problematic in a non-linear model if numerical noise, data assimilation or physics parameterizations were to excite this mode, since it would not propagate away. The authors were able to remove the jump, and the zero group velocity region, by applying a perturbation to the mass matrix (which does not affect consistency). Melvin *et al.* (2014) considered two-dimensional solutions and showed the same results; no spurious inertia gravity modes but jumps do appear in the dispersion relation that can be fixed by consistent modification of the mass matrix.

5. Compatible finite element methods for the non-linear shallow water equations

In this section, we survey the application of compatible finite element methods to the non-linear shallow water equations.

5.1. Energy–enstrophy conservation

The compatible finite element structure facilitates the design of spatial discretizations that conserve mass, energy and enstrophy. An energy–enstrophy conserving scheme was presented in McRae and Cotter (2014). The starting point is the vector invariant form of the shallow water equations,

$$u_t + qF^{\perp} + \nabla\left(gD + \frac{1}{2}|u|^2\right) = 0, \tag{61}$$

$$D_t + \nabla\cdot F = 0, \tag{62}$$

where D is the layer depth, F is the mass flux and q is the potential vorticity, defined by

$$F = uD, \quad q = \frac{\nabla^{\perp}\cdot u + f}{D}, \tag{63}$$

respectively. It is well known that these equations have conserved energy E and mass M, where

$$E = \frac{1}{2}\int_\Omega D|u|^2 + gD^2 dx, \quad M = \int_\Omega D dx, \tag{64}$$

plus an infinity hierarchy of conserved Casimirs, given by

$$C_i = \int_\Omega D q^i dx, \quad i = 1, 2, \ldots. \tag{65}$$

The energy–enstrophy conserving scheme, which can be thought of as an extension of the scheme of Arakawa and Lamb (1981) to compatible finite element methods, conserves E, M, the total vorticity C_1. and the enstrophy C_2. The scheme takes $q \in V_h^0(\Omega)$, $u, F \in V_h^1(\Omega)$, and $D \in V_h^2(\Omega)$, with

$$\int_\Omega w \cdot u_t dx + \int_\Omega w \cdot F^\perp q dx = \int_\Omega \nabla \cdot w \left(gD + \frac{1}{2}|u|^2 \right) dx, \quad \forall w \in V_h^1(\Omega), \tag{66}$$

$$\int_\Omega \phi (D_t + \nabla \cdot F) dx = 0, \quad \forall \phi \in V_h^2(\Omega), \tag{67}$$

$$\int_\Omega w \cdot F dx = \int_\Omega w \cdot u D dx, \quad \forall w \in V_h^1(\Omega), \tag{68}$$

$$\int_\Omega \gamma q D dx = -\int_\Omega \nabla^\perp \gamma \cdot u + \gamma f dx, \quad \forall \gamma \in V_h^0(\Omega). \tag{69}$$

In this scheme, u and D are the prognostic variables and F and q can be diagnosed from them using their definitions in integral form. Despite u and D being prognostic, one can derive a potential vorticity equation by taking $w = \nabla^\perp \gamma$ in (66), for $\gamma \in V_h^0(\Omega)$, which is permitted due to the compatible structure. Then, combining this with the time derivative of (69), we obtain

$$\int_\Omega \gamma (qD)_t dx - \int_\Omega \nabla \gamma \cdot F q dx = 0, \quad \forall \gamma \in V_h^0(\Omega), \tag{70}$$

which is the standard finite element discretization of the potential vorticity equation. Direct computation plus integration by parts then shows that this scheme conserves M, E, C_0 and C_1. Enstrophy conservation is particularly important because it specifies control of q in the L^2 norm (weighted by D), which implies control of gradients of the divergence-free component of u *via* the discrete Helmholtz decomposition. Without this control, even with energy conservation implying control of u in L^2, we would observe unbounded growth in the derivatives of u as the mesh is refined.

In two dimensions, although energy cascades to large scales, enstrophy cascades to small scales. Since (70) is a regular Galerkin discretization of the law of conservation of potential vorticity, this means that enstrophy will accumulate at the grid scale in general, leading to oscillations. It was noted by McRae and Cotter (2014) that it is possible to modify the potential vorticity flux qF in such a way that enstrophy is dissipated at the grid scale while energy is still conserved, by introducing a finite element version of the anticipated potential vorticity scheme (Ringler *et al.*, 2010). This was demonstrated to produce less oscillatory solutions in vortex merger experiments.

All of the properties described in this section can be expressed in the language of finite element exterior calculus; this was presented in Cotter and Thuburn (2014). Moreover, note that conservation of energy and enstrophy may not be exact when time is also discretized. Then, the level of conservation will depend upon the chosen scheme and time step.

5.2. Stable transport schemes

When these schemes are extended to three dimensions, we encounter the problem that energy is now also cascading to small scales. Historically, the solution to this was to add eddy viscosity terms to collect energy at small scales. However, this is unnecessarily dissipative and so does not make the best use of available resolution. The modern approach to developing numerical weather prediction models is to use high-order upwind transport schemes to control energy and enstrophy in the numerical solution. The shallow water equations provide a useful vehicle for exploring these issues in a two-dimensional model that can be run quickly; we review recent results here.

Since we are using different finite element spaces for different fields, this calls for different approaches to the design of transport schemes for them. For $D \in V_h^2(\Omega)$, a standard upwind discontinuous Galerkin discretization can be applied to the mass conservation equation:

$$\int_\Omega \phi D_t dx - \int_\Omega \nabla_h \phi \cdot u D dx + \int_\Gamma [\![u\phi]\!] \widetilde{D} dx = 0, \quad \forall \phi \in V_h^2(\Omega), \tag{71}$$

where Γ is the set of interior facets in the mesh, the 'jump' operator is defined for vectors w by

$$[\![w]\!] = w^+ \cdot n^+ + w^- \cdot n^-, \tag{72}$$

with ± indicating the values of functions on different sides of each edge and $n^\pm$ is the unit normal vector pointing from the ± side to the ∓ side, $\widetilde{D}$ indicates the value of D taken from the upwind side, i.e. the side where $u^\pm \cdot n^\pm$ is negative, and where ∇h is the usual 'broken' gradient defined in L^2 as

$$\nabla_h \big|_T \phi = \nabla \big|_T \phi. \tag{73}$$

Here $|_T$ indicates the restriction of a function to a single cell T, for all cells T in the mesh decomposition of Ω. It was shown in Cotter and Thuburn (2014) and Shipton and Cotter (submitted for publication) that it is possible to find a mass flux F such that

$$D_t + \nabla \cdot F = 0, \tag{74}$$

in a (cheap) post-processing step after the transport scheme step has been completed.

There are several different approaches to velocity transport. One possibility is to maintain (69) and to modify the potential vorticity flux qF so that (70) is replaced by an upwind discretization of the potential vorticity conservation law. For example, choosing the streamline upwind Petrov–Galerkin (Brooks and Hughes, 1982) discretization

$$\int_\Omega \left(\gamma + \frac{\eta h F}{|F|} \cdot \nabla \gamma \right) (Dq)_t dx - \int_\Omega \nabla \gamma \cdot qF dx + \int_\Omega \frac{\eta h F}{|F|} \cdot \nabla \gamma \nabla \cdot (Fq) dx = 0, \quad \forall \gamma \in V_h^0(\Omega), \tag{75}$$

where $0 < \eta < 1$ is a constant upwind parameter and $h(x)$ is an indicator function returning the size of the cell containing the point x. This is a high-order consistent upwind discretization. We obtain a modified potential vorticity flux:

$$qF - \frac{\eta h F}{|F|} ((qD)_t + (Fq)), \tag{76}$$

which can then replace qF in (66). Since this flux is still proportional to F, energy is still conserved (neglecting any energy changes from the D transport scheme). In Shipton and Cotter (submitted for publication), this idea was extended to higher order Taylor–Galerkin time stepping schemes. A related approach is to introduce dual grids, so that the dual grid potential vorticity is represented in a discontinuous finite element space. At lowest order, these are $\mathcal{P}_0$ spaces (piecewise constant functions) that can be treated with finite volume advection schemes that achieve higher spatial order by increasing the stencil size. A discretization for the shallow water equations using finite element spaces for arbitrary polygonal grids, and mappings between primal and dual grids, was developed in Thuburn and Cotter (2015). Results demonstrating this approach on a standard set of test cases on the sphere and comparing cubed sphere and dual icosahedral grids were also presented.

A difficulty with the approach described above is the loss of consistency near the boundary in (66), due to the fact $V_h^0(\Omega)$ consists of functions that vanish on the boundary. It can be shown that this results in a loss of consistency of one order, i.e. from second order to first order consistency. There is a similar problem with the dual grid approach. An alternative approach to this is to treat the curl operator directly in the velocity equation (still in vector invariant form) without introducing an intermediate potential vorticity variable. This can be developed by dotting with a test function w and integrating over a single-cell T, then integrating by parts to obtain:

$$\int_T w\cdot\frac{F^\perp}{D}(\nabla^\perp\cdot u+f)dx=-\int_T \nabla^\perp\left(w\cdot\frac{F^\perp}{D}\right)\cdot u dx+\int_{\partial T} n^\perp\cdot\frac{F^\perp}{D}\tilde{u} dS, \tag{77}$$

where n is the unit outward pointing normal to ∂T, the boundary of e and $\tilde{u}$ is either the averaged or the upwind value of u on ∂T. Even in the presence of upwinding, this discretization is still energy conserving, since the energy equation sets $w = F$ and $F{\cdot}F^\perp = 0$. In Natale and Cotter (submitted for publication), it was shown that the application of this discretization can be derived from a variational principle; the paper also proved convergence (suboptimal by 1 degree) of solutions for the centred flux variant of the discretization in two and three dimensions.

Finally, we remark that all of these velocity transport schemes require solution of global mass matrix systems, one for each stage of an explicit Runge–Kutta method, for example. If one is willing to give up energy conservation, then an embedded discontinuous Galerkin transport scheme (Cotter and Kuzmin, 2016) can be used. In these schemes, the advected field $u\in V_h^1(\Omega)$ is injected into the broken space $\widehat{V}_h^1(\Omega)$ (the same finite element space but with no continuity constraints between cells) at the start of the time step. Then, one or more explicit upwind discontinuous Galerkin time steps are taken using a stable Runge–Kutta scheme. Finally, the solution is projected back into $V_h^1(\Omega)$. This final step does involve a global mass solve, but this can be incorporated into a semi-implicit update at no extra cost. We are currently investigating the stability and accuracy of these schemes.

5.3. Semi-implicit implementation

In semi-implicit discretizations, a crucial aspect is the efficient solution of the resulting coupled linear system that must be solved during each non-linear iteration of the semi-implicit scheme. For example, consider the following implicit time stepping scheme.

$$\begin{aligned}\int_\Omega w\cdot u^d dx &= \int_\Omega w\cdot u^n dx+\frac{\Delta t}{2}\int_\Omega w\cdot f(u^n)^\perp dx\\ &\quad+\frac{\Delta t}{2}\int_\Omega \nabla\cdot wgD^n dx, \quad \forall w\in V_h^1(\Omega),\end{aligned} \tag{78}$$

$$u^a=\Phi^u(u^d;u^{n+1/2}), \tag{79}$$

$$D^{n+1}=\Phi^D(D^d;u^{n+1/2}), \tag{80}$$

$$\begin{aligned}\int_\Omega w\cdot u^{n+1} dx &= \int_\Omega w\cdot u^a dx+\frac{\Delta t}{2}\int_\Omega w\cdot f(u^{n+1})^\perp dx\\ &\quad+\frac{\Delta t}{2}\int_\Omega (\nabla\cdot w)gD^{n+1} dx, \quad \forall w\in V_h^1(\Omega),\end{aligned} \tag{81}$$

where $\Phi^u(v; u^{n+1/2})$ is the application of a chosen velocity advection scheme over one time step, with initial condition v and using advecting velocity $u^{n+1/2}$, while Φ^D is the same thing but for layer depth. A Picard iteration to obtain u^{n+1}, D^{n+1} requires the solution of the following linear system for the iterative corrections to u^{n+1} and D^{n+1},

$$\begin{aligned}\int_\Omega w \cdot \Delta u dx + \frac{\Delta t}{2}\int_\Omega w \cdot f \Delta u^\perp dx \\ -\frac{\Delta t}{2}\int_\Omega (\nabla \cdot w) g \Delta D dx = -R_u[w], \quad \forall w \in V_h^1(\Omega),\end{aligned} \tag{82}$$

$$\int_\Omega \phi \Delta D dx + \frac{\Delta t}{2}\int_\Omega \phi(\nabla \cdot H)\Delta u dx = -R_D[\phi], \quad \forall \phi \in V_h^2(\Omega), \tag{83}$$

where $R_{\mathrm{u}}[w]$ and $R_{\mathrm{D}}[\phi]$ are residuals for the implicit system, and H is the value of D in a state of rest. After obtaining Δu and ΔD, we replace $u^{n+1} \mapsto u^{n+1} + \Delta u$, $D^{n+1} \mapsto D^{n+1} + \Delta D$ and repeat the iterative procedure again for a fixed number of times. In staggered finite difference models, the standard approach to solving (82) and (83) is to neglect the Coriolis term and then to eliminate u to obtain a discrete Helmholtz equation for which smoothers such as SOR are convergent, and therefore the system is amenable to multigrid methods or Krylov subspace methods (or a combination of the two). This approach is problematic in the compatible finite element setting because the velocity mass matrix is globally coupled and thus has a dense inverse.

An alternative approach, which also allows us to keep the Coriolis term in the linear system, is called hybridization; this technique dates back to the 1960s in engineering use, and to the late 1970s and 1980s in terms of numerical analysis (Boffi *et al.*, 2013). In this approach, we consider solutions u in the broken discontinuous space $\widehat{V}_h^1(\Omega)$. Then, Lagrange multipliers are introduced to enforce continuity of the normal component of u, to force it back into $V_h^1(\Omega) \subset \widehat{V}_h^1(\Omega)$. We define a trace space for the Lagrange multipliers, W, consisting of piecewise polynomials on each edge $e \in \Gamma$, with no continuity requirements between edges, and so the degree of the polynomials is the same as the degree of the normal component of functions from $V_h^1(\Omega)$. The resulting system is:

$$\begin{aligned}\int_\Omega w \cdot \Delta u dx + \frac{\Delta t}{2}\int_\Omega w \cdot f \Delta u^\perp dx \\ -\frac{\Delta t}{2}\int_\Omega (\nabla \cdot w) g \Delta D dx + \int_\Gamma \lambda [\![w]\!] dS = -R_u[w], \quad \forall w \in \widehat{V}_h^1(\Omega),\end{aligned} \tag{84}$$

$$\int_\Omega \phi \Delta D dx + \frac{\Delta t}{2}\int_\Omega \phi(\nabla \cdot H)\Delta u dx = -R_D[\phi], \quad \forall \phi \in V_h^2(\Omega), \tag{85}$$

$$\int_\Gamma \mu [\![u]\!] dS = 0, \quad \forall \mu \in W. \tag{86}$$

We see that if u, D, λ solve this system of equations, then $u \in V_h^1(\Omega)$, through (86). Further, if we take $w \in V_h^1(\Omega) \subset \widehat{V}_h^1(\Omega)$, then the λ term vanishes and so u, D also solves (82) and (83). Further, since $\widehat{V}_h^1(\Omega)$. and $V_h^2(\Omega)$ are *both* discontinuous spaces, they can be sparsely eliminated to obtain a positive-definite system for λ. The resulting system is symmetric and positive definite, so can be solved using the conjugate gradient method preconditioned by Gauss–Seidel iteration, for example. In addition, it can be shown that λ is in fact an approximation to $g\,\Delta D\,\Delta t/2$; this fact can then be used to show that a multigrid method for the system for λ will converge (Gopalakrishnan and Tan, 2009). Having solved for λ, u and D can be obtained by back-substitution separately in each element. If the system for λ has only been solved approximately, then some form of projection (i.e. averaging values on each side of element boundaries) is required to obtain $\Delta u \in V_h^1(\Omega)$. This approach to solving the linear system was explored in the context of shallow water equations on the sphere in Shipton and Cotter (submitted for publication). We are currently extending the implementation and analysis to hybridization of the linear system arising in three-dimensional compressible dynamical cores.

6. Three-dimensional compatible finite element spaces

The shallow water equations provide a useful stepping stone for development of numerical methods for three-dimensional dynamical cores. We now discuss the additional aspects that occur when building discretizations in three dimensions. The extension of vertical grid staggering to finite element methods, including high-order finite elements, was introduced by Guerra and Ullrich (2016). Here, we consider the use of three-dimensional compatible finite elements that have the same structure in the vertical.

In analogy to the discussion of Section 3, we start with the following definition.

Definition 6.1. *Let* $V_h^0(\Omega) \subset H^1(\Omega)$, $V_h^1(\Omega) \subset H(\text{curl};\Omega)$, $V_h^2(\Omega) \subset H(\text{div};\Omega)$ and $V_h^3(\Omega) \subset L^2(\Omega)$ *be a sequence of finite element spaces (*$V_h^0(\Omega)$ *and* $V_h^3(\Omega)$ *contain scalar valued functions, while* $V_h^1(\Omega)$ *and* $V_h^2(\Omega)$ *contain vector valued functions). These spaces are called compatible if*:

1. $\nabla\psi \in V_h^1(\Omega), \forall \psi \in V_h^0(\Omega)$,
2. $\nabla \times v \in V_h^2(\Omega), \forall v \in V_h^1(\Omega)$,
3. $\nabla \cdot \boldsymbol{u} \in V_h^3(\Omega), \forall \boldsymbol{u} \in V_h^2(\Omega)$, *and*
4. *there exist bounded projections* π^i, $i = 0, 1, 2, 3$, *such that the following diagram commutes.*

$$\begin{tikzcd}
H^1(\Omega) \arrow[r, "\nabla"] \arrow[d, "\pi^0"] & H(\text{curl};\Omega) \arrow[r, "\nabla\cdot"] \arrow[d, "\pi^1"] & H(\text{div};\Omega) \arrow[r, "\nabla\cdot"] \arrow[d, "\pi^2"] & L^2(\Omega) \arrow[d, "\pi^3"] \\
V_h^0(\Omega) \arrow[r, "\nabla"] & V_h^1(\Omega) \arrow[r, "\nabla\times"] & V_h^2(\Omega) \arrow[r, "\nabla\cdot"] & V_h^3(\Omega)
\end{tikzcd} \tag{87}$$

In a compatible finite element discretization of the three-dimensional compressible Euler equations, we choose the velocity $\boldsymbol{u} \in V_h^2(\Omega)$ and the density $\rho \in V_h^3(\Omega)$. We will consider how to represent temperature in a later section. There are good reasons why all successful numerical weather prediction models have mesh cells aligned in vertical columns. The dynamics is very close to hydrostatic balance, which is obtained by solving

$$\frac{\partial p}{\partial z} = -g\rho. \tag{88}$$

In large-scale models, the domain is very thin, so cells have much greater extent in the horizontal than the vertical direction. If cells are not aligned in columns, then the large horizontal errors in p will be coupled into the hydrostatic equation, and the discretization generates spurious waves that rapidly degrade the state of hydrostatic balance. We will see later that using a mesh that is aligned in columns leads to well-defined hydrostatic pressures when compatible finite element spaces are used. Hence, we consider meshes constructed from hexahedra or triangular prisms, where all cells have side walls aligned in the vertical direction. We do allow for sloping cell tops and bottoms, which facilitates terrain-following meshes.

Under these constraints, we build cells in the following way. We construct a reference cell as the tensor product of a horizontal cell (a triangle or quadrilateral) and an interval. Then, each cell in the physical mesh is obtained by a mapping from the reference cell into $\mathbb{R}^3$, such that the above constraints are satisfied.

To build finite element spaces on this mesh, we first construct a finite element on the reference cell. The finite element is formed from tensor products of finite elements defined on the base triangle or square and finite elements defined on the vertical interval. Of course, the square can also then be defined as the product of two intervals. We then define finite element spaces on the mesh *via* the pullback mappings listed in Table 1. More precisely, if F_K is a mapping from the reference cell $\widehat{K}$ to the physical cell K, then finite element functions are related *via*

- $\psi \in V_h^0(\widehat{K}) \Rightarrow \psi \circ F_K^{-1} \in V_h^0(K)$,
- $v \in V_h^1(\widehat{K}) \Rightarrow J^{-T}\, v \circ F_K^{-1} \in V_h^1(K)$,
- $\boldsymbol{u} \in V_h^2(\widehat{K}) \Rightarrow J\boldsymbol{u}/\det(J) \circ F_K^{-1} \in V_h^2(K)$,
- $\rho \in V_h^3(\widehat{K}) \Rightarrow \rho/\det(J) \circ F_K^{-1} \in V_h^3(K)$,

where J is the Jacobian matrix DF_K. Then, as in the two-dimensional case, given a decomposition $\mathcal{T}_h$ of the domain Ω in elements K, the finite element space $V_h^k(\Omega)$ on $\mathcal{T}_h$ is defined by requiring $V_h^k(K) = F_K^{-1*}(V_h^k(\widehat{K}))$ and $V_h^k(\Omega) \subset HV^k(\Omega)$.

A key problem is that for the velocity space $V_p^2(\Omega)$, this transformation results in non-polynomial functions when the transformation is not affine (a linear transformation composed with translation). We obtain non-affine transformations whenever we use spherical shell (i.e. atmosphere-shaped) domains, or terrain-following coordinates, whether we have triangular prismatic or hexahedral cells. This problem was discussed and the general loss of error quantified using FEEC in the case of compatible finite element spaces on cubical elements in Arnold *et al.* (2014). Here, we extend this discussion to cover the triangular prismatic cells that we wish to use in horizontally unstructured mesh numerical weather prediction. We also show that this general loss of error is still possible for columnar meshes. For the case of manifolds embedded in $\mathbb{R}^n$, this problem was addressed by Holst and Stern (2012), where it was shown that if the computational mesh is a piecewise polynomial approximation of a smooth manifold, then functions can be approximated at the optimal rate and further that numerical solutions of mixed Hodge Laplacian problems converge optimally. Here, we extend these results to spherical shell domains in the case of extruded meshes where the base mesh is either a cubed sphere mesh of quadrilaterals or an arbitrarily structured mesh of triangles. This extension makes use of a transformation of the sphere domain to a three-dimensional submanifold of $\mathbb{R}^4$, for which an affine mesh is possible.

The rest of this section is structured as follows. In Section 6.1, we describe the construction of tensor product finite element spaces, including the extension to triangular prism elements. In Section 6.2, we prove some approximation properties of these spaces. In Section 6.3, we discuss global transformations of meshes and analyse how these affect the approximation properties of finite element spaces. This analysis is then applied in Section 6.4 to the case of a spherical annulus.

6.1. Tensor product finite element spaces

The construction of tensor product compatible spaces is addressed in Arnold (2013) and Arnold *et al.* (2014) with application to cubic meshes. However, the procedure for the definition of such spaces is rather general and applies to any mesh whose elements are generated as Cartesian product of simplices. This covers both cell types, triangular prisms and hexahedra, which we are interested in. The construction relies on having two simplices $T \subset \mathbb{R}^n$ and $S \subset \mathbb{R}^m$ and a polynomial complex of compatible finite element spaces on each of these. Then, one can produce a polynomial complex of compatible finite element spaces on $T \times S$ using the ones defined on T and S. Here, we only show the result of this procedure for our case of interest, i.e. when $n = 2$, $m = 1$ and $T \times S$ is a prism.

We start by selecting polynomial spaces on S and T on which we base the tensor product construction. Since S is one-dimensional, we can only take on S the sequence of spaces (CG_s, DG_{s-1}), for $s \geq 1$, as this is the only valid choice of compatible polynomial spaces. As for the complex on T we have more freedom. Here, we restrict ourself to the examples discussed in Section 3. In particular, if we consider on T the sequence of compatible spaces $(CG_r, RT_{r-1}, DG_{r-1})$, we obtain the following compatible spaces on $T \times S$,

$$\begin{aligned}
\mathcal{E}^-_{r,s}V^0(T\times S) &= CG_r(T)\otimes CG_s(S),\\
\mathcal{E}^-_{r,s}V^1(T\times S) &= [RT_{r-1}(T)\otimes CG_s(S)]\oplus[CG_r(T)\otimes DG_{s-1}(S)\hat{z}],\\
\mathcal{E}^-_{r,s}V^2(T\times S) &= [DG_r(T)\hat{z}\otimes CG_s(S)]\oplus[RT_{r-1}(T)\times DG_{s-1}(S)\,\hat{z}],\\
\mathcal{E}^-_{r,s}V^3(T\times S) &= DG_{r-1}(T)\otimes DG_{s-1}(S),
\end{aligned}$$

where $\hat{z}$ is a unit vector field oriented in the extrusion direction. Alternatively, if we use the sequence $(CG_r, BDM_{r-1}, DG_{r-2})$ complex on T we obtain the following compatible spaces on $T \times S$:

$$\begin{aligned}
\mathcal{E}_{r,s}V^0(T\times S) &= CG_r(T)\otimes CG_s(S),\\
\mathcal{E}_{r,s}V^1(T\times S) &= [BDM_{r-1}(T)\otimes CG_s(S)]\oplus[CG_r(T)\otimes DG_{s-1}(S)\,\hat{z}],\\
\mathcal{E}_{r,s}V^2(T\times S) &= [DG_{r-2}(T)\,\hat{z}\otimes CG_s(S)]\oplus[BDM_{r-1}(T)\times DG_{s-1}(S)\hat{z}],\\
\mathcal{E}_{r,s}V^3(T\times S) &= DG_{r-2}(T)\otimes DG_{s-1}(S)\,.
\end{aligned}$$

Note that in both cases, these spaces have been defined in the way they would appear in the complex, i.e $d^k : \mathcal{E}_{r,s}V^k \to \mathcal{E}_{r,s}V^{k+1}$ and the same for the $\mathcal{E}^-_{r,s}V^k$ spaces. The fact that such spaces can be called compatible follows from the possibility to define bounded projections in accordance to Definition 6.1. This can indeed be done by exploiting the tensor product structure as shown in Arnold *et al.* (2014). More details on the definition and the implementation of the tensor product spaces discussed in this section can also be found in (McRae *et al.* (2014).

6.2. Approximation properties and curvilinear elements

In this section, we analyse the convergence properties of the finite element spaces defined in Section 1 for specific classes of maps F_K. In particular, we assume that the mesh is generated by either affine or multilinear transformations and, for the resulting finite dimensional space $V_h^k(\Omega)$, we are interested in providing bounds for the quantity:

$$\inf_{p \in V_h^k(\Omega)} \|u - p\|_{L^2(\Omega)}. \tag{89}$$

As for non-affine meshes, we will mostly use the results in Arnold *et al.* (2014), which focus on tensor product spaces on curvilinear cubic elements and extend them to general tensor product elements.

We start by recalling that a finite element space is constructed using a set of functions on a reference element $\hat{K}$ and by a set of maps $F_K : \hat{K} \to K$, where K is an element in the domain decomposition $\mathcal{T}_h$. The maps F_K can be arbitrary apart from some regularity conditions. These are determined by the way the pullback of F_K transforms under Sobolev norms, which is described in the following theorem.

Theorem 6.1. [Arnold *et al.*, 2014, *Theorem 2.1*]: *Let M be a positive constant and r a non-negative integer. There exists a constant C depending only on r, M and the dimension of the domain n such that*

$$\|F_K^* v\|_{H^r V^k(\hat{K})} \leq C\|v\|_{H^r V^k(K)}, \quad v \in H^r V^k(K), \tag{90}$$

whenever $\hat{K}$ *is a domain in* $\mathbb{R}^n$ *and* $F_K : \hat{K} \to K$ *is a* C^{r+1} *diffeomorphism satisfying*

$$\max_{1 \leq s \leq r+1} |F_K|_{W^s_\infty(\hat{T})} \leq M, \quad |F_K^{-1}|_{W^1_\infty(\hat{T})} \leq M. \tag{91}$$

In order to get an estimate for the quantity in (89), the Clément interpolant construction is generally invoked. First, for any q-dimensional domain Q define $\mathcal{P}_r(Q)$ to be the space of polynomial scalar functions on Q with degree up to r. Moreover, let $\mathcal{P}_r V^k(Q)$ be equal to $\mathcal{P}_r(Q)$ if $k = 0, q$, or to the space of vector fields with coefficients in $\mathcal{P}_r(Q)$ if $0 < k < q$. If we let $V_h^k(K)$ be the restriction of the finite element space $V_h^k(\Omega)$ to the element K and assume $\mathcal{P}_r V^k(K) \subset V_h^k(K)$, then one can design a projection $\pi : H^s V^k(\Omega) \to V_h^k(\Omega)$, called the Clément interpolant, for any $(n-k)/2 < s \leq r+1$, such that $\pi\big|_K u = (\pi u)\big|_K$ for any $u \in H^s V^k(\Omega)$ and

$$\|u - \pi u\|_{L^2(K)} \leq C h_K^s |u|_{H^s V^k(K)}, \quad u \in H^s V^k(K), \tag{92}$$

where h_K denotes the diameter of K. Summing up over all elements on both sides of (92) yields the global estimate

$$\|u - \pi u\|_{L^2(\Omega)} \leq C h^s \sum_{T \in \mathcal{T}_h} |u|_{H^s V^k(T)} = C h^s |u|_{H^s V^k(\Omega)}, \quad u \in H^s V^k(\Omega), \tag{93}$$

where h is the maximum element diameter in the decomposition $\mathcal{T}_h$ of Ω. Clearly, then, (93) implies the same estimate for the best approximation error in (89).

Therefore, for a given choice of the function space $V_h^k(\hat{K})$ on the reference element $\hat{K}$, one is led to determine the conditions under which $\mathcal{P}_r V^k(K) \subset V_h^k(K) = F_K^{-1*}(V_h^k(\hat{K}))$ in order for the estimates in (93) to hold. We carry out this analysis for the polynomial spaces defined in Section 1, restricting the choice of F_K to affine and multilinear transformations as in Arnold *et al.* (2014). The results are expressed in the following theorems which essentially are a restatement of Theorem 6.1 in Arnold *et al.* (2014), to the specific spaces considered here.

We assume that the mesh is constituted by tensor product elements, so that we look for a polynomial space on the reference element $\hat{K}$ belonging to either one of the families $\mathcal{E}_{r,s}$ or $\mathcal{E}^-_{r,s}$, such that the estimates in Equations (92) and (93) hold. The resulting sufficient conditions for optimal convergence rates are stated in the following theorem.

Theorem 6.2. *Consider a three-dimensional element* $\hat{K} = \hat{T} \times \hat{S}$ *obtained as tensor product of a two-dimensional simplex* $\hat{T}$ *and a one-dimensional simplex* $\hat{S}$*. Let* F_K *be an affine map and* $\mathcal{E}_{r,r}V^k(\hat{K}) \subset V_h^k(\hat{K})$*; or let be a multilinear map and* $\mathcal{E}_{2(r+k),r+k}V^k(\hat{K}) \subset V_h^k(\hat{K})$*. Then* $\mathcal{P}_r V^k(K) \subset V_h^k(K)$ *and (92) and (93) hold.*

Finally, we consider a specific class of transformations that leave the coordinates on one of the two simplices $\hat{T}$ or $\hat{S}$ invariant up to an affine transformation. This corresponds to the case of local area models with terrain-following meshes. Specifically, we assume $F_K = A_K \circ D_K$, where $D_K : \widehat{K} \to \widetilde{K}$ is a special multilinear deformation of the reference element and $A_K : \widetilde{K} \to K$ is an arbitrary affine map. The map D_K is defined in coordinates by $D_K = (I_T, M_S)$, with I_T being the identity map on $\widehat{T}$ and $M_S : \widehat{K} \to \mathbb{R}^m$ being a sufficiently smooth multilinear map. We call the maps F_K constructed in this way as multilinear maps invariant on $\widehat{T}$. Analogously, if $D_K = (M_T, I_S)$, with I_S being the identity map on $\hat{S}$ and $M_T : \widehat{K} \to \mathbb{R}^n$ being a sufficiently smooth multilinear map, we call the relative F_K as a multilinear map invariant on $\hat{S}$. A graphical representation of such constructions is given in Figure 1. The sufficient conditions for optimal convergence relative to these classes of transformations is given by the following theorem.

Theorem 6.3. *Consider a three-dimensional element* $\widehat{K} = \widehat{T} \times \widehat{S}$ *obtained as tensor product of a two-dimensional simplex* $\widehat{T}$ *and a one-dimensional simplex* $\widehat{S}$*. Let* F_K *be a multilinear map invariant on* $\widehat{T}$, $\mu = \min\,(k, 1)$ *and* $\mathcal{E}_{2(r+k),r+\propto}V^k(\widehat{K}) \subset V_h^k(\widehat{K})$*; or let* F_K *be a multilinear map invariant on* $\hat{S}$, $\nu = \min\,(k, 2)$ *and* $\mathcal{E}_{2(r+\nu),r+k}V^k(\widehat{K}) \subset V_h^k(\widehat{K})$*. Then,* $\mathcal{P}_r V^k(K) \subset V_h^k(K)$ *and (92) and (93) hold.*

As one would expect, when F_K is a multilinear map invariant on $\widehat{T}$ (or $\hat{S}$) we obtain sufficient conditions that are considerably less strict than the ones relative to general multlinear maps. However, this is true only when $k > 1$ (or $k > 2$ respectively) as the conditions that apply for lower values of k are the same in both cases.

To summarize the results of this section, we can state that the spaces generated using the function spaces $V_h^k(K) = F_K^{-1*}(\mathcal{E}_{r,s}V^k(\widehat{K}))$ yield the convergence rates shown in Table 2. Note that multilinear transformations invariant on the one-dimensional simplex $\widehat{T}$ are precisely the type of transformations that occur in atmosphere simulations when dealing with meshes following a varying topography.

6.3. Approximation properties for global transformations

The last subsection was devoted to the study of the approximation properties of finite element spaces defined on meshes generated by arbitrary local transformations of a reference element. In several applications, however, the mesh is generated by means of a global transformation from a reference mesh with optimal convergence properties. In these cases, if the global map is sufficiently regular, one can transfer the convergence properties of the original mesh to the deformed one.

Numerical weather prediction models require the capability to solve in a spherical annulus domain; we are also interested in domains with topography and terrain-following coordinates. These are not polygonal domains, and we must model them using piecewise polynomial approximations $\widetilde{\Omega}$. With this in mind, let $\widetilde{\Omega}$ be also an n-dimensional Riemmanian manifold, such that there exists a C^1 diffeomorphism $F : \Omega \to \widetilde{\Omega}$. We want to show how elements of the spaces $HV^k(\Omega)$ can be transformed by F into elements of $HV^k(\widetilde{\Omega})$. It is understood that simple composition with F^{-1} is not enough for this purpose. A well-known example for this is the case of the space $H(\mathrm{div}, \Omega)$, where the correct transformation is given by the contravariant Piola transform which defines an isomorphism between $H(\mathrm{div}, \Omega)$ and $H(\mathrm{div}; \widetilde{\Omega})$.

It is easy to check that one needs to define appropriate transformations, analogous to the Piola transform, for each space $HV^k(\Omega)$. We denote such transformations with a unique symbol, $F^* : HV^k(\widetilde{\Omega}) \to HV^k(\Omega)$, namely the pullback of F, with inverse $F^{-1*} : HV^k(\Omega) \to HV^k(\widetilde{\Omega})$, which defines an isomorphism between $HV^k(\Omega)$ and $HV^k(\widetilde{\Omega})$. F^{-1*} coincides with the contravariant Piola transform when applied to $HV^{n-1}(\Omega)$ and takes a different form depending on the degree

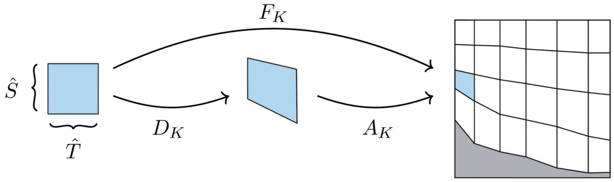

Figure 1. A two-dimensional mesh generated by multilinear transformations invariant on $\hat{T}$ of a reference element (on the left).

Table 2. Convergence rates for finite element spaces obtained by pullback of $\mathcal{E}_{r,s}V^k(\hat{T} \times \hat{S})$ with $\dim(\hat{T}) = 2$ and $\dim(\hat{S}) = 1$ (triangular prism elements), using different classes of transformations F_K. The symbol $\lfloor x \rfloor$ denotes the largest integer smaller or equal to x.

$\mathcal{E}_{r,s}\Lambda^k(\mathcal{T}_h)$	**Affine**	**Multilinear**	**Multilinear invariant on $\hat{T}$**
$k = 0$	$\min(r,s) + 1$	$\min(\lfloor r / 2 \rfloor, s) + 1$	$\min(\lfloor r / 2 \rfloor, s) + 1$
$k = 1$	$\min(r,s) + 1$	$\min(\lfloor r / 2 \rfloor, s)$	$\min(\lfloor r / 2 \rfloor, s)$
$k = 2$	$\min(r,s) + 1$	$\min(\lfloor r / 2 \rfloor, s) - 1$	$\min(\lfloor r / 2 \rfloor - 1, s)$
$k = 3$	$\min(r,s) + 1$	$\min(\lvert r / 2 \rvert, s) - 2$	$\min(\lvert r / 2 \rvert - 2, s)$

k in the complex and the dimension of the domain n. The action of F^{-1*} on an element $v \in HV^k(\Omega)$ is defined in Table 1. Such definitions allow to preserve the Hilbert complex structure, i.e. the following diagram commutes

$$\begin{tikzcd}
0 \arrow[r] & HV^0(\Omega) \arrow[r, "\mathrm{d}^0"] \arrow[d, "F^{-1*}"] & HV^1(\Omega) \arrow[r, "\mathrm{d}^1"] \arrow[d, "F^{-1*}"] & \dots \arrow[r, "\mathrm{d}^{n-1}"] & HV^n(\Omega) \arrow[r] \arrow[d, "F^{-1*}"] & 0 \\
0 \arrow[r] & HV^0(\tilde{\Omega}) \arrow[r, "\mathrm{d}^0"] & HV^1(\tilde{\Omega}) \arrow[r, "\mathrm{d}^1"] & \dots \arrow[r, "\mathrm{d}^{n-1}"] & HV^n(\tilde{\Omega}) \arrow[r] & 0
\end{tikzcd} \tag{94}$$

Note that in the diagram above the differential operators are expressed accordingly to the metric defined on Ω or $\tilde{\Omega}$.

Consider two n-dimensional manifolds $\tilde{\Omega}$ and Ω and a decomposition $\mathcal{T}_{\tilde{h}}$ of $\tilde{\Omega}$ by either simplices or tensor product elements. In other words, each element $\tilde{K}$ in $\mathcal{T}_{\tilde{h}}$ is generated via a sufficiently regular transformation $F_{\tilde{K}} : \hat{K} \to \tilde{K}$ as in Section 2, where $\hat{K}$is a reference element. Let G be a diffeomorphism from $\tilde{\Omega}$ to Ω, and let $G_K : \tilde{K} \to K$ be the restriction of G to $\tilde{K}$. This defines a decomposition T_h of Ω in which each element K is given by a transformation of the reference element $\hat{K}$ defined by $F_K := G_K \circ F_{\tilde{K}}$. Then one can define a compatible function spaces on each K as pullback of the ones defined on $\hat{K}$. In particular, given a space $V^k(\hat{K})$ on the reference element $\hat{K}$, the corresponding space on K is given by $V^k(K) := G_K^{-1*} F_{\tilde{K}}^{-1*} V^k(\hat{K})$.

Let $V_{\tilde{h}}^k(\tilde{\Omega})$ and $V_h^k(\Omega)$ be the resulting finite element spaces on $\tilde{\Omega}$ and Ω, respectively. The definition of the space $V_h^k(\Omega)$ makes sense only if $V_h^k(\Omega) \subset HV^k(\Omega)$. For this to hold, it is sufficient that G is C^0 at the element interfaces and G_K^{-1*} is bounded inside each element. These conditions, however, are not sufficient for optimal convergence. The following theorem establishes the level of regularity of G required to guarantee optimal convergence rates for the relative finite element spaces on Ω.

Theorem 6.4. *Assume that* $\mathcal{P}_r V^k(\tilde{K}) \subset V_h^k(\tilde{K})$ *and let* $G : \tilde{\Omega} \to \Omega$ *be a* C^0 *diffeomorphism. Moreover, assume that for any* $\tilde{K} \in \mathcal{T}_{\tilde{h}}$ *the restriction* $G_K : \tilde{K} \to K$ *is a* C^{r+2} *diffeomorphism, such that*

$$\max_{1 \leq s \leq r+2} |G_K|_{W_\infty^s(\tilde{\Omega})} \leq M, \quad |G_K^{-1}|_{W_\infty^1(\tilde{\Omega})} \leq M, \tag{95}$$

for $M > 0$. *Then, for* $(n - k)/2 < s \leq r + 1$,

$$\inf_{p \in V_h^k(\Omega)} \|u - p\|_{L^2(\Omega)} \leq C\tilde{h}^s \|u\|_{H^s V^k(\Omega)}, \quad u \in H^s V^k(\Omega), \tag{96}$$

where $C > 0$ *and* $\tilde{h}$ *is the largest diameter of the elements in the decomposition* $\mathcal{T}_{\tilde{h}}$ *of* $\tilde{\Omega}$.

Proof. By definition any function $p \in V_h^k(\Omega)$ can be expressed formally as $p = G^{-1*}\tilde{p}$ with $\tilde{p} \in V_{\tilde{h}}^k(\tilde{\Omega})$, with the pullback computed elementwise. Therefore,

$$\inf_{p \in V_h^k(\Omega)} \|u - p\|_{L^2(\Omega)} = \inf_{\tilde{p} \in V_{\tilde{h}}^k(\tilde{\Omega})} \|G^{-1*}(G^* u - \tilde{p})\|_{L^2(\Omega)}. \tag{97}$$

Using Theorem 6.1 and (93), we obtain the following bound,

$$\inf_{\tilde{p} \in V_{\tilde{h}}^k(\tilde{\Omega})} \|G^{-1*}(G^* u - \tilde{p})\|_{L^2(\Omega)} \leq C \inf_{\tilde{p} \in V_{\tilde{h}}^k(\tilde{\Omega})} \|G^* u - \tilde{p}\|_{L^2(\tilde{\Omega})} \leq C\tilde{h}^s \sum_{\tilde{K} \in \mathcal{T}_{\tilde{h}}} |G_K^* u|_{H^s V^k(\tilde{K})}, \tag{98}$$

for $(n-k)/2 < s \leq r+1$. Applying again Theorem 1 to the right hand side of (98), we obtain

$$\left| G_K^* u \right|_{H^s V^k(\widetilde{K})} \leq \| G_K^* u \|_{H^s V^k(\widetilde{K})} \leq C \| u \|_{H^s V^k(K)}. \tag{99}$$

The proof is completed by combining (97), (98) and (99). ☐

The estimates derived in this section might seem to have little utility, given that functions in $V_h^k(K)$ cannot be computed accurately for arbitrary diffeomorphisms G. Nonetheless, Theorem 6.1 finds an important application in the context of problems defined on surfaces. To see this, let Ω be a n-dimensional manifold embedded in $\mathbb{R}^{n+1}$. Consider a family of decompositions of Ω by piecewise linear elements parameterized by the maximum element diameter $\tilde{h}$.Such meshes implicitly define a family of approximate domains $\widetilde{\Omega}_{\tilde{h}}$, on which it is possible to define as finite element space $V_{\tilde{h}}^k(\widetilde{\Omega}_{\tilde{h}})$ any of the polynomial spaces considered above. Under some regularity assumption on Ω (Holst and Stern, 2012) one can also construct a family of diffeomorphisms $G_{\tilde{h}} : \widetilde{\Omega}_{\tilde{h}} \to \Omega$, which by construction can only be C^0 at the element interfaces and that approach the identity map as $\tilde{h} \to 0$. The pullbacks $G_{\tilde{h}}^{-1*}$ can be used to construct finite element spaces $V_h^k(\Omega)$. Hence, a natural measure of the error between $G_{\tilde{h}}$ and the identity map on Ω is given by

$$\| \operatorname{adj}(G_{\tilde{h}}^{-1*}) \circ G_{\tilde{h}}^{-1*} - \mathrm{Id} \| , \tag{100}$$

where adj denotes the adjoint with respect to the L_2 inner product, Id is the identity map and $\|\cdot\|$ is the standard operator norm on the space of L^2 functions or vector fields, accordingly to the degree k.

The exact same setting for this problem and its connection with the FEEC are explored in Holst and Stern (2012) for the solution of elliptic problems. Their main conclusion is that even though one uses only functions in $V_{\tilde{h}}^k(\widetilde{\Omega}_{\tilde{h}})$ as computational basis functions, the final error estimates depend on the quantity in (100) and on the best approximation error on the space $V_h^k(\Omega)$, as measured by

$$\inf_{p \in V_h^k(\Omega)} \| u - p \|_{L^2(\Omega)}. \tag{101}$$

Therefore, if Theorem 6.4 holds, with the constant M in (95) being independent of $\tilde{h}$, one can apply the results in Holst and Stern (2012) and obtain optimal convergence rates for the problem defined on the original surface Ω.

6.4. Finite element spaces in 'atmosphere-shaped' domains

In this section, we consider as an example finite element spaces defined on a spherical shell in $\mathbb{R}^3$, which is a typical scenario for atmosphere simulations, and we compute their approximation properties. The spherical shell C is the submanifold of $\mathbb{R}^3$ defined by

$$C := \left\{ x \in \mathbb{R}^3 \,\middle|\, a \leq \|x\| \leq b \right\}, \tag{102}$$

for any $b > a > 0$. The shell C is topologically the same as the generalised cylinder $\widetilde{C} = S \times [a, b] \subset \mathbb{R}^4$, where S is the sphere of radius a,

$$S := \left\{ x \in \mathbb{R}^3 \,\middle|\, \|x\| = a \right\}. \tag{103}$$

In fact, we can construct a continuous transformation $F : \mathbb{R}^4 \to \mathbb{R}^3$ such that the restriction $F|_{\widetilde{C}} : \widetilde{C} \to C$ is a C^∞ diffeomorphism. A particular choice for F is the multilinear map defined by

$$F(x_1, x_2, x_3, x_4) = (x_1, x_2, x_3)\left(1 + \frac{x_4 - a}{a}\right), \tag{104}$$

where $\{x_i\}_{i=1}^3$ and $\{x_i\}_{i=1}^4$ are Cartesian coordinates in $\mathbb{R}^3$ and $\mathbb{R}^4$ respectively.

Given a decomposition of $[a,b]$ and one for S, one can construct a decomposition $\mathcal{T}_{\tilde{h}}$ for $\widetilde{C}$ by tensor product, meaning that each element in $\mathcal{T}_{\tilde{h}}$ is constructed as tensor product element of a simplex in each mesh. The decomposition $\mathcal{T}_{\tilde{h}}$ implicitly defines the approximate cylinder $\widetilde{C}_{\tilde{h}}$. The map F restricted to $\widetilde{C}_{\tilde{h}}$ defines a decomposition $\mathcal{T}_h$ for C and the approximate shell C_h as image of F restricted to $\widetilde{C}_{\tilde{h}}$. Any element in $K \in \mathcal{T}_h$ is constructed using a multilinear map $F|_{\widetilde{K}} : \widetilde{K} \to K$ with $\widetilde{K} \in \mathcal{T}_{\tilde{h}}$.

Since the decomposition $\mathcal{T}_{\tilde{h}}$ is composed by tensor product elements we can define a finite element space $V_{\tilde{h}}^k(\widetilde{C}_{\tilde{h}})$ on $\widetilde{C}_{\tilde{h}}$ to belong to either one of the families $\mathcal{E}_{r,s}V^k$ or $\mathcal{E}_{r,s}^- V^k$. The map F induces via pullback a definition for a finite element space $V_h^k(C_h)$ on C_h. Since the derivatives of the map $F|_{\widetilde{K}}$ are bounded away from 0 by a constant independent of the size of $\widetilde{K}$, we can apply Theorem 6.2 and the classical Clément interpolant construction to obtain a first estimate for the approximation properties of the space $V_h^k(C)$. In this case, however, such estimate is not sharp. As a matter of fact, since $F|_{\widetilde{C}_{\tilde{h}}}$ is certainly C^0 at the element interfaces Theorem 6.4 also applies providing optimal convergence rates for the space $V_h^k(C_h)$ as inherited by the ones of $V_{\tilde{h}}^k(\widetilde{C}_{\tilde{h}})$.

As explained in the last section, the error estimates for $V_h^k(C_h)$ can be transferred to a finite element space defined on C, if we are able to define a diffeomorphism from C_h to C that satisfies the requirements of Theorem 6.4. We now show how this can be done. Using the techniques described in Holst and Stern (2012), one can construct a diffeomorphism $G_{\tilde{h}} : \widetilde{C}_{\tilde{h}} \to \widetilde{C}$ that converges to the identity map on $\widetilde{C}$ as $\tilde{h} \to 0$. The map $G_{\tilde{h}}$ can be lifted via the map F to define the diffeomorphism $G_h : C_h \to C$, i.e.

$$G_h := F|_{\widetilde{C}} \circ G_{\tilde{h}} \circ (F|_{\widetilde{C}_{\tilde{h}}})^{-1} \,. \tag{105}$$

In Figure 2, a graphical representation of the maps is presented in the two-dimensional case. Because of the smoothness of F, if $G_{\tilde{h}}$ satisfies the hypothesis of Theorem 6.4 so does G_h. Therefore, the optimal convergence rates can be extended to the finite element spaces defined on C via G_h. As explained in Section 6.3, in order for these estimates to be useful G_h needs to approach the identity as $\tilde{h} \to 0$, in the sense of (100). Formally, this can be verified directly by letting $\tilde{h} \to 0$ in (105).

It is worth observing that given a differential problem on C, high-order convergence can be achieved just as long as G_h approaches the identity sufficiently fast with mesh refinement. In other words, if the error term in (100) is rewritten for G_h scales with h^{s+1}, then s is the highest convergence rate achievable regardless of the polynomial order of the spaces on $\widetilde{C}_{\tilde{h}}$. In Holst and Stern (2012), it is proven that if one adopts an isoparametric approach s is the order of the polynomial approximation $\widetilde{C}_{\tilde{h}}$ of the surface $\widetilde{C}$. Therefore, if higher convergence rates are required, one just needs to adopt a higher order approximation of $\widetilde{C}$ instead of a piecewise linear one and repeat the considerations above.

Numerical experiments were performed to verify the expected convergence rates for both direct L^2 projections and the Helmholtz problem (Arnold *et al.*, 2006) on the shell as shown in Figures 3 and 4, respectively.

It should be stressed that the final estimates on the curved domain rely essentially on the possibility of defining the global map F. Given this, proving optimal convergence rates in other cases of interest boils down to constructing such maps. For example, since in the meteorological community, the use of cubed sphere meshes with quadrilateral cells is widespread, it is reasonable to ask whether such meshes enjoy optimal approximation properties when using the FEEC. These meshes are constructed using the C^0 map from the surface of a cube to the sphere. This map possesses enough smoothness for Theorem 4 to apply and can be easily extended to the shell case by means of an extrusion in the fourth dimension as described above.

A similar discussion to the one of the present section can be applied to other domains which are relevant for geophysical applications. For example, one is generally interested in performing simulations in a domain that approximates the actual shape of the Earth, including terrain irregularities such as continents and mountains. In this case, instead of modifying the map F one can think of defining a domain $\widetilde{C} \subset \mathbb{R}^4$ such that its image C under F coincides with the desired Earthatmosphere approximation. If C is sufficiently smooth then $\widetilde{C}$ is a regular surface in $\mathbb{R}^4$ and one can repeat the considerations of this section to generate a finite element space on C with optimal convergence properties.

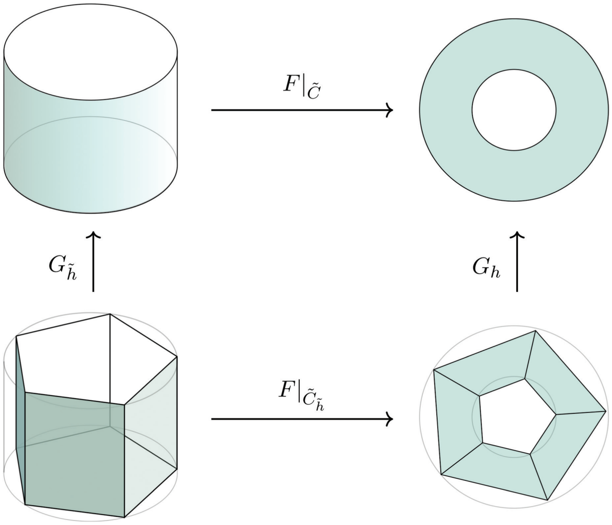

Figure 2. Graphical representation of the maps involved in the construction of G_h in the two-dimensional case.

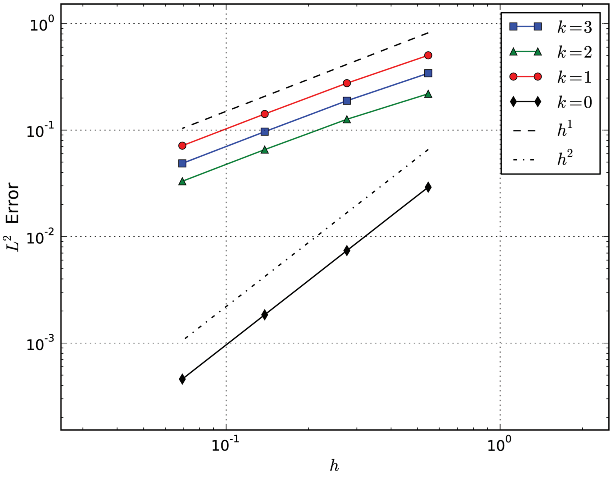

Figure 3. Convergence test for the finite element spaces on C defined via pullback of $\mathcal{E}^{-}_{r,s}V^k(\widetilde{C}_{\tilde{h}})$ with $r = s = 1$. The map G_h is computed using a piecewise linear approximation, which is sufficient to show convergence up to second order.

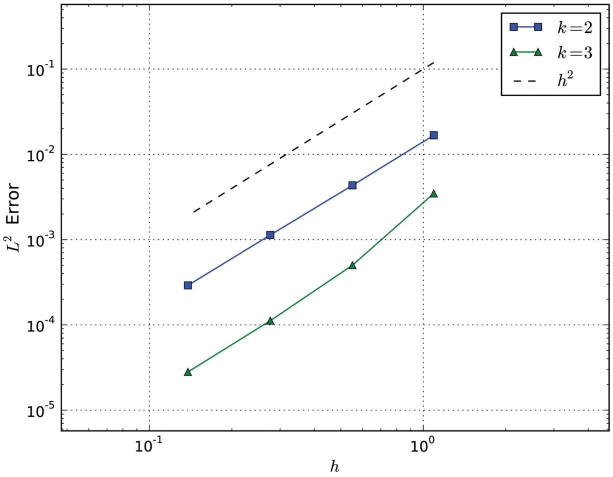

Figure 4. Convergence test for the mixed formulation of the Helmholtz equation for three forms on C, with finite element spaces defined via pullback of $\mathcal{E}^{-}_{r,s}V^k(\widetilde{C}_{\tilde{h}})$ with $r = s = 2$ and $k = 2,3$. The map G_h is computed using a piecewise linear approximation, which is sufficient to show convergence up to second order.

In general, however, the decomposition of C might be less regular, or in other words, it might not be possible to reinterpret it as image of a smooth map F. In these cases, one can still rely on the 'worst-case' scenario estimates in Theorems 6.2 and 6.3. Note that the results in Theorem 6.3 are particularly relevant for mountain terrain simulation, in which case the computational mesh is usually obtained by modifying in the vertical direction only a decomposition which is otherwise regular in the sense of Theorem 6.4. Note however, that smoothing topography is a standard part of numerical weather prediction.

7. Representation of temperature in compatible finite element spaces

There was one piece of crucial information missing from the last section on three-dimensional finite element spaces: the choice of finite element space for potential temperature θ. This is the temperature a dry fluid parcel would attain if brought adiabatically to a standard reference pressure p_0, typically the pressure at sea level, i.e.

$$\theta = T\left(\frac{p_0}{p}\right)^{\kappa}, \tag{106}$$

where $\kappa = R_d/c_p$, R_d is the gas constant of dry air, and c_p is the specific heat capacity at constant pressure. In this section, we introduce finite element spaces for θ and then describe finite element discretizations for three-dimensional dynamical cores.

In finite difference dynamical cores, there are two main options for vertical staggering: the Lorenz grid (temperature collocated with density) and the Charney–Phillips grid (temperature collocated with vertical velocity). The temperature data are always cell centred in the horizontal.

To mimic the Lorenz grid, we can simply represent temperature in the density space $V_h^3(\Omega)$. However, many dynamical cores prefer the Charney–Phillips grid, since it avoids a spurious mode in the hydrostatic balance equation. To see how to mimic the Charney–Phillips grid in a compatible finite element discretization, we need to examine the structure of $V_h^2(\Omega)$ more closely.

At the level of the reference cell $\hat{K}$, the space $V_h^2(\Omega)$ can be decomposed into a vertical part $V_h^{2,v}(\Omega)$, and a horizontal part $V_h^{2,h}(\Omega)$. For example, for a triangular prism, we could take $V_h^{2,v}(\Omega) = DG_0 \otimes CG_1\hat{z}$, where $\hat{z}$ is the unit vector pointing in the vertical direction. Clearly, all the elements of $V_h^{2,v}(\Omega)$ have vanishing normal components on the side walls of $\hat{K}$. Then, a corresponding choice for $V_h^{2,h}(\Omega)$ is $BDM_1 \otimes DG_0$. Since BDM_1 is defined on the horizontal triangle, all elements of $V_h^{2,h}(\Omega)$ point in the horizontal plane and all the elements of $V_h^{2,h}(\Omega)$ have vanishing normal components on the top and bottom of $\hat{K}$. After Piola transformation to a physical cell K with vertical side walls but possibly sloping top and bottom boundaries, we see that $V_h^{2,v}(K)$ points in the vertical direction, but now $V_h^{2,h}(K)$ has a vertical component as well. It is this non-orthogonality for terrain-following grids that leads to pressure gradient errors just as for the C-grid staggered finite difference method.

We see that $V_h^{2,v}(\widehat{K})$ is vector valued, but always points in the vertical direction. We define $V_h^t(\widehat{K})$ by

$$V_h^t(\widehat{K}) = \{\theta : \theta\hat{z} \in V_h^{2,v}(\widehat{K})\}. \tag{107}$$

Then, given a mapping $F_K : \widehat{K} \to K$, we define

$$V_h^t(K) = \{\theta : \theta \circ F_K \in V_h^t(\widehat{K})\}, \tag{108}$$

i.e. the transformation from the reference element is the usual one for scalar functions, despite $V_h^t(\Omega)$ being constructed from $V_h^{2,v}(\Omega)$. This defines the extension of the Charney–Phillips staggering to compatible finite element discretizations.

The reason for this construction is that it leads to a one-to-one mapping between pressure and potential temperature in the hydrostatic pressure equation. To show this, we will describe a compatible finite element discretization for the three-dimensional compressible Euler equations and then present results about the hydrostatic balance.

7.1. Compatible finite element spatial discretization for the three-dimensional compressible Euler equations

In this section, we present a finite element spatial discretization for the three-dimensional compressible Euler equations. Here, we are mainly addressing the question of how to obtain a consistent discretization of the various terms when compatible finite element spaces with only partial continuity are used. Following the methodology of Wood *et al.* (2014), we consider the system of equations in terms of the density ρ, and the potential temperature θ. These variables provide the opportunity for a better representation of hydrostatic balance in the numerical model. The equations are

$$u_t + (u \cdot \nabla)u + 2\,\Omega \times u = -\theta \nabla \Pi - g\hat{r}, \tag{109}$$

$$\theta_t + u \cdot \nabla \theta = 0, \tag{110}$$

$$\rho_t + \nabla \cdot (u\rho) = 0, \tag{111}$$

where u is the velocity, Ω is the rotation vector in the Coriolis term, g is the acceleration due to gravity, $\hat{r}$ is the unit upward vector (which points away from the origin for a global sphere domain), and Π is a function of ρ and θ given by

$$\Pi = \left(\frac{R_d\rho\theta}{p_0}\right)^{\frac{\kappa}{1-\kappa}}. \tag{112}$$

Note that within the given formulation, the pressure can be determined via the equation of state $p = \rho R_d T$, combined with the definition of Π in (112) and the one of potential temperature in (106).

To develop the discretization for the advection term, we use the vector invariant form

$$(u \cdot \nabla)u = (\nabla \times u) \times u + \nabla \frac{1}{2}|u|^2. \tag{113}$$

Then, extending the development of the advection term from the shallow water equations section, we obtain the weak form

$$\int_\Omega \nabla_h \times (u \times w) \cdot u dx - \int_\Gamma \{\{n \times (u \times w)\}\} \cdot \tilde{u} dS, \quad \forall w \in V_h^1(\Omega), \tag{114}$$

where we adopt the notation that

$$\{\{n \times F\}\} = n^+ \times F^+ + n^- \times F^-, \tag{115}$$

where ∇_h indicates the 'broken' gradient obtained by evaluating the gradient pointwise in each cell, and where $\tilde{u}$ is the value of u on the upwind side of each facet.

To develop the discretization for the pressure gradient term, we also integrate by parts in each cell to obtain:

$$\int_\Omega \nabla_h \cdot (w\theta)\Pi dx - \int_{\Gamma_v} [\![w\theta]\!]\{\Pi\} dS, \quad \forall w \in V_h^1(\Omega), \tag{116}$$

where $\{\cdot\}$ denotes the average for scalars,

$$\{f\} = \frac{1}{2}(f^+ + f^-). \tag{117}$$

where the Γ_v is set of all vertically aligned interior facets (there are no jumps on horizontally aligned facets since $[\![w\theta]\!] = 0$ there). For lowest order RT elements on cuboids, it can be shown that this discretization reduces to the finite difference discretization currently used in the Met Office Unified Model as described in Wood *et al.* (2014).

Finally, we need to provide discretizations of the transport equations for ρ and θ. Since $\rho \in V_h^3(\Omega)$ is discontinuous, we can use the standard upwind discontinuous Galerkin approach. Building a transport scheme for $\theta \in V_h^t(\Omega)$ is more delicate, because θ is continuous in the vertical. This means that there is no upwind stabilization for vertical transport, and so we choose to apply a Streamline Upwind Petrov Galerkin (SUPG) stabilization there. We develop this in two steps. First, we take the θ advection equation, multiply by a test function and integrate by parts, taking the upwind value of θ on the vertical column faces, to get

$$\int_\Omega \gamma\theta_t dx - \int_\Omega \nabla_h \cdot (u\gamma)\, \theta dx + \int_{\Gamma_v} [\![\gamma u]\!]\tilde{\theta} dS = 0, \quad \forall \gamma \in V_h^t(\Omega). \tag{118}$$

Then, to prepare for the SUPG stabilization, we integrate by parts again, to give the equivalent formula,

$$\int_\Omega \gamma\theta_t dx + \int_\Omega \gamma u \cdot \nabla\, \theta dx + \int_{\Gamma_v} ([\![\gamma u]\!]\tilde{\theta} - [\![\gamma u]\!]) dS = 0, \quad \forall \gamma \in V_h^t(\Omega). \tag{119}$$

Since there are now no derivatives applied to γ, we can obtain an SUPG discretization by replacing $\gamma \to \gamma + \eta u \cdot \hat{r}\Delta t\hat{r} \cdot \nabla \gamma$, where $\eta > 0$ is a constant upwinding coefficient, to obtain

$$\int_{\Omega} (\gamma + \eta u \cdot \hat{r}\Delta t\hat{r} \cdot \nabla \gamma)\theta_t dx + \int_{\Omega} (\gamma + \eta u \cdot \hat{r}\Delta t\hat{r} \cdot \nabla \gamma)u \cdot \nabla \theta dx$$
$$+ \int_{\Gamma_v} [\![(\gamma + \eta u \cdot \hat{r}\Delta t\hat{r} \cdot \nabla \gamma)u]\!]\tilde{\theta} - [\![(\gamma + \eta u \cdot \hat{r}\Delta t\hat{r} \cdot \nabla \gamma)u\theta]\!]dS = 0, \quad \forall \gamma \in V_h^t(\Omega). \tag{120}$$

This leads to the following spatial discretization. We seek $u \in V_h^2(\Omega)$, $\rho, \Pi \in V_h^3(\Omega)$, $\theta \in V_h^t(\Omega)$, such that

$$\int_{\Omega} w \cdot u_t + \nabla \times (u \times w) \cdot u - \int_{\Gamma} [\![n \times (u \times w)]\!] \cdot \tilde{u} dS$$
$$+ \int_{\Omega} w \cdot 2\,\Omega \times u dx - \int_{\Omega} \nabla_h \cdot (w\theta)\Pi dx$$
$$+ \int_{\Gamma} [\![w\theta]\!]\{\Pi\} ds + \int_{\Omega} g w \cdot \hat{r} dx = 0, \quad \forall w \in V_h^2(\Omega), \tag{121}$$

$$\int_{\Omega} (\gamma + \eta u \cdot \hat{r}\Delta t\hat{r} \cdot \nabla \gamma)\theta_t dx + \int_{\Omega} (\gamma + \eta u \cdot \hat{r}\Delta t\hat{r} \cdot \nabla \gamma)u \cdot \nabla \theta dx$$
$$+ \int_{\Gamma_v} [\![(\gamma + \eta u \cdot \hat{r}\Delta t\hat{r} \cdot \nabla \gamma)u]\!]\tilde{\theta} - [\![(\gamma + \eta u \cdot \hat{r}\Delta t\hat{r} \cdot \nabla \gamma)u\theta]\!]dS = 0, \quad \forall \gamma \in V_h^t(\Omega), \tag{122}$$

$$\int_{\Omega} \phi \rho_t dx - \int_{\Omega} \nabla_h \phi \cdot \rho dx + \int_{\Gamma} [\![\phi u]\!]\tilde{\rho} dx = 0, \quad \forall \phi \in V_h^3(\Omega), \tag{123}$$

$$\int_{\Omega} \phi \Pi dx - \int_{\Omega} \phi \left(\frac{R_d \rho \theta}{p_0}\right)^{\frac{\kappa}{1-\kappa}} dx = 0, \quad \forall \phi \in V_h^3(\Omega). \tag{124}$$

Some initial numerical results using this spatial discretization are provided in Section 2.

7.2. Vertical slice results

In this section, we present some preliminary results obtained using the discretizations outlined above, in vertical slice (x,z) geometry. The finite element spaces are (CG_2, RT_1, DG_1). The time stepping scheme uses the semi-implicit formulation described in Section 3 along with a third-order, three-stage explicit Strong Stability Preserving Runge Kutta (SSPRK; Gottlieb *et al.*, 2001) scheme for the advection terms. The tests were performed using the Firedrake software suite (see Rathgeber *et al.*, 2016), which allows for symbolic implementation of mixed finite element problems. As the hybridization technique is not currently implemented within Firedrake, we instead solve the linear system by eliminating θ, solving the mixed system for u and p, then reassembling θ.

The first test case is a non-hydrostatic gravity wave, initially described in Skamarock and Klemp (1994) for the Boussinesq model. The test describes the evolution of inertia gravity waves excited by a localized, warm θ perturbation to the background state of constant buoyancy frequency $N = 0.01$ s^{-1}. Figure 5 (top) shows the θ perturbation at $t = 3000$ s from a simulation with $dx = 2000$ m, $dz = 1000$ m, $dt = 6$s. The result compares well with the results produced by other models (e.g. see Melvin *et al.*, 2010).

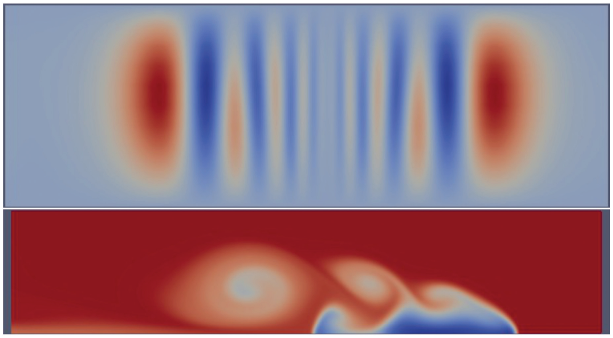

Figure 5. Some initial numerical results using the compatible finite element discretization for the compressible Euler equations in a slice geometry (courtesy: J Shipton). Top: a potential temperature colour plot showing a snapshot from the Skamarock–Klemp gravity wave (Skamarock and Klemp, 1994) test case. Bottom: a potential temperature colour plot showing a snapshot from the Straka *et al.* (1993) dense bubble testcase.

The second test case describes the evolution of a cold bubble as it falls, meets the bottom of the domain, spreads and produces Kelvin–Helmholtz rotors due to the shear instability. The initial conditions are as described in Melvin *et al.* (2010) and the background state is in hydrostatic balance and isentropic with constant potential temperature $\theta = 300\text{K}$. Since the purpose of this test is to study convergence to a well-resolved solution, viscosity and diffusion terms are included, the discretization of which is achieved using the interior penalty method. Figure 5 (bottom) shows the potential temperature at $t = 15$ min of a well-resolved simulation ($dx = dz = 100$ m, $dt = 0.5$ s), showing that we capture the correct position of the front and the details of the instability (Melvin *et al.*, 2010).

Much more comprehensive results in the vertical slice setting will be presented in a forthcoming paper.

7.3. Hydrostatic balance

In this section, we discuss the hydrostatic balance properties of this discretization. It is critical that the basic state of hydrostatic balance is correctly represented in the numerical model. This means that there should be a one-to-one correspondence between potential temperature θ and Π in the state of hydrostatic balance; this precludes spurious hydrostatic modes from emerging in numerical simulations close to a hydrostatic state. To obtain the hydrostatic balance equation, we consider test functions w from the vertical part $V_h^{2,v}(\Omega)$ of the velocity space $V_h^2(\Omega)$. After neglecting the advection terms as well as any non-traditional contribution to the Coriolis force, we obtain

$$0 = -\int_\Omega \nabla_h \cdot (\theta w)\,\Pi dx - \int_\Omega w \cdot \hat{r} g dx, \quad \forall w \in V_h^{2,v}(\Omega). \tag{125}$$

Note that while w is written as a vector, it always points in the $\hat{r}$ direction, since it is in $V_h^{2,v}(\Omega)$. This means that this is really a scalar equation for the vertical derivative of Π. It also explains why the vertical facet integrals have vanished from the expression. In a domain with level topography, vectors in the horizontal space $V_h^{2,h}(\Omega)$ do not have a component in the $\hat{r}$ direction, hence no horizontal motion is generated if this equation is satisfied. However, in terrain-following coordinates, there will be projection of the pressure gradient of Π into the horizontal direction; this is exactly analogous to the pressure gradient errors observed over topography observed in the C-grid staggered finite difference method. We also observe that this equation decomposes into independent equations for each vertical column c. We obtain

$$0 = -\int_c \nabla_h \cdot (\theta w)\Pi dx - \int_c w \cdot \hat{r} g dx, \quad \forall w \in V_h^{2,v}(c), \tag{126}$$

for each column c in the mesh. Note here that we are considering rigid lid boundary conditions; $V_h^{2,v}(\Omega)$ is defined so that $w \cdot n = 0$ for all $w \in V_h^{2,v}(\Omega)$.

We now prove the following two results that show the degree to which there is a one-to-one correspondence between Π and θ through this equation.

Theorem 7.1. *Let $V_h^t(c)$ be the restriction of $V_h^t(\Omega)$ to column c, and let $V_h^{2,v}(c)$, $V_h^3(c)$ be similar restrictions for $V_h^{2,v}(\Omega)$ and $V_h^3(\Omega)$, respectively. Further, define $\overline{V}_h^3(c)$ as*

$$\overline{V}_h^3(c) = \left\{ \rho \in V_h^3(c) \,:\, \int_c \rho dx = 0 \right\}, \tag{127}$$

i.e. the subspace of functions with zero column averages.

For all $\theta \in V_h^t(c)$ satisfying $\theta > 0$, there is a unique solution $\Pi \in \overline{V}_h^3(c)$ satisfying (126).

Proof. First we show that (126) can be reformulated as a coupled system of equations. We first note that if $v \in V_h^{2,v}(c)$, then $\nabla \cdot (\theta v) = \frac{\partial}{\partial r}(\theta v)$, where $\frac{\partial}{\partial r}$ is the vertical derivative). Therefore, if $\nabla \cdot (\theta v) = 0$, then $v=0$, since $v \cdot \hat{r} = 0$ at the top and bottom of the column. This means that we can add a term involving v to (126) without changing the solution. Hence, we obtain the equivalent problem of finding $v \in V_h^{2,v}(c)$, $\Pi \in V_h^3(c)$ such that

$$\int_c w \cdot v dx + \int_c \nabla \cdot (\theta w) \Pi dx = -\int_c g w \cdot \hat{r} dx, \quad \forall\, w \in V_h^{2,v}(c), \tag{128}$$

$$\int_c \phi \nabla \cdot (\theta v) dx = 0, \quad \forall\, \phi \in V_h^3(c). \tag{129}$$

This defines a mixed finite element problem in each vertical column, of the form

$$a(w, v) + b(w, \Pi) = F(w), \quad \forall\, w \in V_h^{2,v}(c), \tag{130}$$

$$b(v, \phi) = 0, \quad \forall\, \phi \in V_h^3(c), \tag{131}$$

with the modified derivative operator $v \mapsto \nabla \cdot (\theta v)$, for which Brezzi's stability conditions can be easily verified. □

Similar results are straightforwardly obtained by considering slip boundary conditions on the bottom surface and constant pressure boundary conditions on the top surface.

Theorem 7.2. *Let $\Pi \in \overline{V}_h^3(c)$, and let $\mathring{V}_h^t(c)$ be the subspace of $V_h^t(c)$ given by*

$$\mathring{V}_h^t(c) = \left\{ \theta \in V_h^t(c) : \theta|_{top} = \theta|_{bottom} = 0 \right\}, \tag{132}$$

where $\theta|_{top}$ and $\theta|_{bottom}$ are the values of θ restricted to the top and bottom surface respectively. Let Π be such that $\hat{r} \cdot [\![\Pi]\!] > 0$ on Γc, where Γc is the set of interior facets in the column c. Further, if $V_h^3(\Omega)$ is degree $p \geq 1$ in the vertical direction, then let $\frac{\partial \Pi}{\partial z} > 0$ in each cell. Then, (126) *has a unique solution for $\theta \in \mathring{V}_h^t(c)$.*

Proof. Consider the reference surface mesh from which the extruded mesh Ω was constructed (either a mesh of a planar surface, or a mesh approximating the sphere). For a given column c, consider the base surface cell on the surface mesh corresponding to c and construct an affine reference volume cell $\hat{e}$ which is of unit height and is extruded orthogonal to the base cell.

Let g_e be the transformation from $\hat{e}$ into each cell e in the column. Then, we note that we may write $w \in V_h^{2,v}(c)$ as

$$w = J \hat{r} \gamma, \quad \gamma \in \mathring{V}_h^t(c), \tag{133}$$

where J is the Jacobian defined by

$$J|_e = \frac{\partial g_e}{\partial \hat{x}} \circ g_e^{-1}, \tag{134}$$

in each cell e. Then, (126) becomes

$$A(\gamma, \theta) = -\int_c \frac{\partial}{\partial z}(\gamma \theta \kappa) \Pi dx = -\int_c \gamma \kappa g dx, \quad \forall\, \gamma \in V_h^t(c), \tag{135}$$

where $\kappa = \hat{r} \cdot J\hat{r}$. We assume a non-degenerate mesh so that $c_0 < \kappa < c_1$ for positive constants c_0, c_1. We then apply integration by parts to obtain

$$A(\gamma, \theta) = \int_c \kappa\gamma\theta \frac{\partial}{\partial z}\bigg|_h \Pi dx + \int_\Gamma \gamma\theta\kappa\hat{r} \cdot [\![\Pi]\!] dS. \tag{136}$$

Under our assumptions on Π, $A(\gamma,\theta)$ defines an inner product on $\mathring{V}_h^t(c)$, and therefore the solution exists and is unique. ☐

This result states that θ is unique up to a choice of values at the boundary. This is the same situation as we have in the C-grid staggered finite difference discretization.

8. Summary and outlook

In this survey article, we reviewed recent work on the application of compatible finite element spaces to geophysical fluid dynamics. In particular, we concentrated on analytic results that provide information about the behaviour of these discretizations when applied to geophysically relevant problems. We also contributed some extra results, including the analysis of inertial modes for the linear shallow water equations, where the conclusion is that there are no spurious inertial modes.

The most substantial contribution is the analysis of the approximation properties of some classes of tensor product finite element spaces that are compatible with the FEEC framework; these spaces have been proposed for the development of global atmosphere models. We applied to such spaces the results of Arnold *et al.* (2014) that state that if the mesh is generated by an arbitrary multilinear transformation one can generally expect a loss in the convergence rate in the L^2 norm, which depends on the way the spaces transform under coordinate transformations. In particular, such convergence loss is at least of $k + 1$ degrees, where k is the order of the space considered as part of the de Rham complex. For example, if the mesh is generated using a prismatic reference element with tensor product polynomial spaces of degree p for both the horizontal and vertical direction, we found that:

- for $H(\text{curl})$ elements ($k = 1$) the expected convergence rate is $\lfloor p/2 \rfloor$;
- for $H(\text{div})$ elements ($k = 2$) the expected convergence rate is $\lfloor p/2 \rfloor - 1$;
- for L^2 elements ($k = 3$) the expected convergence rate is $\lfloor p/2 \rfloor - 2$.

However, if the mesh is obtained via a sufficiently regular global transformation as detailed in the statement of Theorem 6.4, one can retrieve optimal convergence rate $p + 1$ independently of k. In practice, this implies that tensor product finite element spaces on prisms, as derived from classical mixed space such as Raviart–Thomas or Brezzi–Douglas–Marini spaces (Fortin and Brezzi, 1991), can still produce optimal convergence rates even when considered under non-affine transformations. This is the case, e.g. for the spherical shell problem, although generalizations of the approach presented here can be easily adapted to more accurate representations of the Earth surface.

Finally, we provided some new results analysing the hydrostatic balance when a particular choice of temperature space is used, which is compatible with the vertical part of the velocity space. These results showed that, under suitable assumptions of static stability, a hydrostatic pressure can be uniquely determined from a given temperature profile and *vice versa*. All of these theoretical results are underpinning ongoing model development, and detailed numerical studies and discussion of computational performance will be explored in future papers.

References

Arakawa A, Lamb VR. A potential enstrophy and energy conserving scheme for the shallow water equations. *Mon Weather Rev* 1981; 109: 18–36.

Arnold DN. Spaces of finite element differential forms. In: Brezzi F, Franzone PC, Gianazza U, Gilardi G (eds). *Analysis and Numerics of Partial Differential Equations*. Milan, Italy: Springer, 2013, pp. 117–40.

Arnold DN, Boffi D, Bonizzoni F. Finite element differential forms on curvilinear cubic meshes and their approximation properties. *Numer Math* 2014; 129: 1.

Arnold DN, Falk RS, Winther R. Finite element exterior calculus, homological techniques, and applications. *Acta Numer* 2006; 15: 1–155.

Arnold DN, Falk RS, Winther R. Finite element exterior calculus: from Hodge theory to numerical stability. *Bull Amer Math Soc (N.S.)* 2010; 47: 281–354.

Arnold DN, Logg A. Periodic table of the finite elements. *SIAM News* 2014; 47.

Bao L, Klöfkorn R, Nair RD. Horizontally explicit and vertically implicit (HEVI) time discretization scheme for a discontinuous Galerkin nonhydrostatic model. *Mon Weather Rev* 2015; 143: 972–90.

Boffi D, Brezzi F, Fortin M. *Mixed Finite Element Methods and Applications*. Vol. 44. Berlin, Germany: Springer, 2013.

Brdar S, Baldauf M, Dedner A, Klöfkorn R. Comparison of dynamical cores for NWP models: comparison of COSMO and Dune. *Theor Comp Fluid Dynam* 2013; 27: 453–72.

Brooks AN, Hughes TJR. Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. *Comput Methods Appl Mech Eng* 1982; 32: 199–259.

Cotter CJ, Ham DA. Numerical wave propagation for the triangular P1DG-P2 finite element pair. *J Comput Phys* 2011; 230: 2806–20.

Cotter CJ, Kuzmin D. Embedded discontinuous galerkin transport schemes with localised limiters. *J Comput Phys* 2016; 311: 363–73.

Cotter CJ, Shipton J. Mixed finite elements for numerical weather prediction. *J Comput Phys* 2012; 231: 7076–91.

Cotter CJ, Thuburn J. A finite element exterior calculus framework for the rotating shallow-water equations. *J Comput Phys* 2014; 257: 1506–26.

Danilov S. On utility of triangular C-grid type discretization for numerical modeling of large-scale ocean flows. *Ocean Dynam* 2010; 60: 1361–69.

Dennis J, Edwards J, Evans KJ ***et al***. CAM-SE: a scalable spectral element dynamical core for the Community Atmosphere Model. *Int J High Perform Comput Appl* 2011; 26: 74–89.

Fortin M, Brezzi F. *Mixed and Hybrid Finite Element Methods (Springer Series in Computational Mathematics)*. Berlin: Springer and Heidelberg GmbH & Co. K, 1991.

Fournier A, Taylor MA, Tribbia JJ. The spectral element atmosphere model (SEAM): High-resolution parallel computation and localized resolution of regional dynamics. *Mon Weather Rev* 2004; 132: 726–48.

Giraldo FX, Kelly JF, Constantinescu EM. Implicit-explicit formulations of a three-dimensional nonhydrostatic unified model of the atmosphere (NUMA). *SIAM J Sci Comput* 2013; 35: B1162–94.

Gopalakrishnan J, Tan S. A convergent multigrid cycle for the hybridized mixed method. *Numer Linear Algebra Appl* 2009; 16: 689–714.

Guerra JE, Ullrich PA. A high-order staggered finite-element vertical discretization for non-hydrostatic atmospheric models. *Geosci Model Dev* 2016; 2016: 1–53.

Gottlieb S, Shu C-W, Tadmor E. Strong stability-preserving high-order time discretization methods. *SIAM Rev* 2001; 43: 89–112.

Hiptmair R. Finite elements in computational electromagnetism. *Acta Numerica* 2002; 11: 237–339.

Holst M, Stern A. Geometric variational crimes: Hilbert complexes, finite element exterior calculus, and problems on hypersurfaces. *Found Comput Math* 2012; 12: 263–93.

Kelly JF, Giraldo FX. Continuous and discontinuous Galerkin methods for a scalable three-dimensional nonhydrostatic atmospheric model: limited-area mode. *J Comput Phys* 2012; 231: 7988–8008.

Marras S, Kelly JF, Moragues M ***et al***. A review of element-based Galerkin methods for numerical weather prediction: finite elements, spectral elements, and discontinuous Galerkin. *Arch Comput Methods Eng* 2016; 23: 673–722.

Marras S, Moragues M, Vázquez M, Jorba O, Houzeaux G. Simulations of moist convection by a variational multiscale stabilized finite element method. *J Comput Phys* 2013; 252: 195–218.

McRae ATT, Bercea G-T, Mitchell L, Ham DA, Cotter CJ. Automated generation and symbolic manipulation of tensor product finite elements. *SIAM J Sci Comput* 2016; 38: S25–S47.

McRae ATT, Cotter CJ. Energy- and enstrophy-conserving schemes for the shallow-water equations, based on mimetic finite elements. *Quart J R Meteorol Soc*, 2014; 140: 2223–34.

Melvin T, Dubal M, Wood N, Staniforth A, Zerroukat M. An inherently mass-conserving iterative semi-implicit semi-Lagrangian discretization of the non-hydrostatic vertical-slice equations. *Quart J R Meteorol Soc* 2010; 136: 799–814.

Melvin T, Staniforth A, Cotter CJ. A two-dimensional mixed finite-element pair on rectangles. *Quart J R Meteorol Soc* 2014; 140: 930–42.

Rathgeber F, Ham DA, Mitchell L ***et al***. Firedrake: automating the finite element method by composing abstractions. *ACM TOMS*, 2016. doi:10.1145/2998441.

Ringler TD, Thuburn J, Klemp JB, Skamarock WC. A unified approach to energy conservation and potential vorticity dynamics for arbitrarily-structured C-grids. *J Comput Phys* 2010; 229: 3065–90.

Rognes ME, Ham DA, Cotter CJ, McRae ATT. Automating the solution of PDEs on the sphere and other manifolds in FEniCS 1.2. *Geosci Model Dev* 2013; 6: 2099–119.

Rognes ME, Kirby RC, Logg A. Efficient assembly of H(div) and H(curl) conforming finite elements. *SIAM J Sci Comput* 2009; 31: 4130–51.

Rostand V, Le Roux DY. Raviart–Thomas and Brezzi–Douglas–Marini finite-element approximations of the shallow-water equations. *Int J Numer Methods Fluids* 2008; 57: 951–76.

Skamarock WC, Klemp JB. Efficiency and accuracy of the Klemp-Wilhelmson time-splitting technique. *Mon Weather Rev* 1994; 122: 2623–30.

Staniforth A, Melvin T, Cotter CJ. Analysis of a mixed finite-element pair proposed for an atmospheric dynamical core. *Quart J R Meteorol Soc* 2013; 139: 1239–54.

Staniforth A, Thuburn J. Horizontal grids for global weather and climate prediction models: a review. *Q J R Meteorol Soc* 2012; 138: 1–26.

Straka JM, Wilhelmson RB, Wicker LJ, Anderson JR, Droegemeier KK. Numerical solutions of a non-linear density current: a benchmark solution and comparisons. *Int J Numer Methods Fluids* 1993; 17: 1–22.

Thomas SJ, Loft RD. The NCAR spectral element climate dynamical core: Semi-implicit Eulerian formulation. *J Sci Comput* 2005; 25: 307–322.

Thuburn J, Cotter CJ. A primal–dual mimetic finite element scheme for the rotating shallow water equations on polygonal spherical meshes. *J Comput Phys* 2015; 290: 274–97.

Walters RA, Casulli V. A robust, finite element model for hydrostatic surface water flows. *Commun Numer Methods Eng* 1998; 14: 931–40.

Wood N, Staniforth A, White A ***et al***. An inherently mass-conserving semi-implicit semi-Lagrangian discretization of the deep-atmosphere global non-hydrostatic equations. *Quart J R Meteorol Soc* 2014; 140: 1505–20.