Radial collocation methods for the onset of convection in rotating spheres.

J. Sánchez[a,*], F. Garcia[a], M. Net[a]

[a] *Dept. de Física Aplicada, Universitat Politècnica de Catalunya, Jordi Girona Salgado s/n. Campus Nord. Mòdul B4, 08034 Barcelona, Spain*

Abstract

The viability of using collocation methods in radius and spherical harmonics in the angular variables to calculate convective flows in full spherical geometry is examined. As a test problem the stability of the conductive state of a self-gravitating fluid sphere subject to rotation and internal heating is considered. A study of the behavior of different radial meshes previously used by several authors in polar coordinates, including or not the origin, is first performed. The presence of spurious modes due to the treatment of the singularity at the origin, to the spherical harmonics truncation, and to the initialization of the eigenvalue solver is shown, and ways to eliminate them are presented. Finally, to show the usefulness of the method, the neutral stability curves at very high Taylor and moderate and small Prandtl numbers are calculated and shown.

Keywords: Collocation methods, spherical harmonics, thermal convection, linear stability analysis.

1. Introduction

Spectral methods have been used during the last three decades for the solution of partial differential equations, and particularly in Fluid Mechanics. The books by Gottlieb and Orszag, Canuto et al. and Boyd [1, 2, 3] were specially influential in the field. They were followed by other authors ([4, 5, 6, 7, 8, 9, 10] for instance), including the new edition in two volumes of [2] ([11, 12]). Although the theory is well developed, many practical implementation details have to be studied when particular problems are undertaken.

In many applications, the domain in which the equations have to be solved is a product of intervals in some particular coordinate system, and when Galerkin methods are used, the tensorial product of bases of functions in each of the intervals provides a base for the full domain. In the particular case of polar, cylindrical polar, or spherical coordinates, the singularity at the origin, if it is

*Corresponding author's e-mail: juan.j.sanchez@upc.edu

inside the domain, introduces an additional complication, and regularity conditions might have to be imposed. The case of polar and cylindrical polar systems has been largely studied in the past, both for Galerkin and collocation methods. Several ways of achieving spectral accuracy have been developed. A recent article [13] compares seven different radial bases for the approximation of functions and for solving the Poisson equation on the unit disk. When collocation methods are used, some authors add regularity conditions at the origin [14, 15, 16], while others avoid including the origin in the collocation mesh [17, 18, 19, 20].

The case of a full sphere has been considered more recently. In the angular coordinates (colatitude and longitude) the expansion in spherical harmonics is analogous to the expansion in Fourier series in the angle of polar coordinates, or in the two angles in toroidal coordinates (see [3]). Since they are the eigenfunctions of the Laplace operator on the unit sphere, all diffusion terms of functions expanded in this base become block-diagonal, with blocks of the size of the radial discretization. Moreover, the spherical harmonics do not suffer the pole problem, i.e. instability due to accumulation of points near the poles when uniform meshes in the two angles are taken, but the problem at the origin is still present. If the domain is a spherical shell, the combination of spherical harmonics with Chebyshev expansions or collocation in the radius ensures spectral accuracy. Fourier expansions in the colatitude were also considered very early for problems on spherical surfaces, to allow the use of fast Fourier transforms in both angular coordinates [21, 22, 23].

Two main options are available for the treatment of the radial coordinate in full spherical domains; expansions in radial basis functions followed by Galerkin, Petrov-Galerkin, tau or collocation projections, to obtain a system of equations for the expansion coefficients, or a pseudospectral collocation method in which the derivatives are substituted by difference operators on a radial mesh of points.

A list of possible radial functions can be found in [24], and in some recent works [25, 26]. In [25], how much a expansion in even or odd degree Chebyshev polynomials violates the regularity conditions at the origin is quantified, and two basis of one-sided Jacobi polynomials (Verkley and Worland) are compared. The main conclusion is that unless there is a need of a faithful representation of the solution, preserving the regularity near the origin, Chebyshev methods would be the choice, in general, due to the availability of fast transforms. In [26] a methodology for constructing orthogonal Galerkin basis satisfying quite general boundary conditions is described. These type of radial functions, satisfying also the full regularity conditions have been recently used to study the onset of kinematic dynamos in [27].

A similar method is based on Petrov-Galerkin projections. The idea goes back to [28], and consists in trying to preserve, as much as possible, the banded structure of the derivative matrices of Chebyshev expansions. The inclusion of tau implementations of the boundary conditions, and the use of a different basis of test functions lead to almost banded matrices which can be solved efficiently. The method can be applied not only to linear ordinary differential equations of constant coefficients but also to the case of non-constant smooth coefficients [29, 30]. Moreover the linear systems are well-conditioned. These

methods have been recently extended to two-dimensional partial differential equations in rectangular domains in [31], and used in fluid flow applications [32].

In the present article the possibility of using collocation methods in the radius is studied. The authors have developed several time integration and stability analysis codes, using this approach, for the convection in spherical shells of pure and binary fluids with very good results [33, 34, 35, 36, 37, 38]. The main reasons for using collocation methods were simplicity and that changing boundary conditions was easy. Therefore it seems natural trying to extend the algorithms for the full sphere. Before starting to develop a full 3D time integrator, it is worth studying the stability analysis of the conduction state. Most sources of instability of the time stepping algorithms can be detected and corrected by looking at the linearized problem, which separates into one for each azimuthal wave number. It is known that collocation methods usually give rise to worst conditioned systems than Galerkin or Petrov-Galerkin methods [11] but, if one has in mind fully tridimensional time evolutions or to apply continuation techniques like in [39, 40], it is not expected that more than a few hundreds of radial collocation points will be used. Otherwise the total size of the discretization would be prohibitive. Moreover, the problems solved in this article imply the solution of discretized systems of linear differential equations with a block tridiagonal structure. Therefore the nice banded shape of Galerkin methods, which is the main reason for their efficiency, is destroyed when a block LU decomposition is used to solve the linear systems. In addition, when the non-linear terms are included in a time-evolution code, their evaluation requires passing to physical space by some kind of transformation. If the radial discretization is by a pseudospectral method there is no need to make forward and backward transformations in this coordinate, which represents some CPU time saving.

The rest of the article is organized as follows. In Section 2, the formulation of the problem and the angular discretization of the equations are introduced. In Section 3, the radial discretization is discussed together with some details of the implementation. It includes a comparison of the results obtained by using different radial meshes on several tests. In Section 4 the results for moderate and small Prandtl numbers at high Taylor numbers are presented. Finally, the paper closes in Section 5 with a brief summary of the main conclusions.

2. Mathematical model and angular discretization

Consider the thermal convection of a spherical fluid sphere of radius r_o, internally and uniformly heated, rotating about an axis of symmetry of direction $\mathbf{k}$ with constant angular velocity $\mathbf{\Omega} = \Omega\mathbf{k}$, and subject to radial gravity $\mathbf{g} = -\gamma\mathbf{r}$, where γ is a constant and $\mathbf{r}$ the position vector. This is the gravity field inside a spherical mass of uniform density. The mass, momentum and energy equations are written, using the Boussinesq approximation, in a rotating frame of reference with angular velocity $\mathbf{\Omega}$, and in spherical coordinates, (r, θ, φ), with θ measuring the colatitude, and φ the longitude. In addition, the centrifugal force is neglected, since $\Omega^2/\gamma \ll 1$ in most celestial bodies, for instance in the

major planets, and the density in the Coriolis term is taken constant. The units to write the dimensionless equations are the radius of the sphere, r_o, for the distance, $\nu^2/\gamma\alpha r_o^4$ for the temperature, and r_o^2/ν for the time, ν and α being the kinematic viscosity and the thermal expansion coefficient, respectively.

The divergence-free velocity field is expressed in terms of toroidal, Ψ, and poloidal, Φ, potentials

$$\mathbf{v} = \nabla \times (\Psi \mathbf{r}) + \nabla \times \nabla \times (\Phi \mathbf{r})\,. \tag{1}$$

The non-dimensional equations for both potentials, and the temperature perturbation, $\Theta = T - T_c$, from the conduction state, $\mathbf{v} = \mathbf{0}$ and $T = T_c(r)$, are obtained by taking the radial component of the curl and double curl of the momentum equation and substituting $T = T_c + \Theta$ into the temperature equation. They are

$$\left[(\partial_t - \nabla^2)L_2 - 2Ta^{1/2}\,\partial_\varphi\right]\Psi \;=\; -2Ta^{1/2}\mathcal{Q}\Phi - \mathbf{r}\cdot\nabla\times(\boldsymbol{\omega}\times\mathbf{v}), \tag{2}$$

$$\left[(\partial_t - \nabla^2)L_2 - 2Ta^{1/2}\,\partial_\varphi\right]\nabla^2\Phi + L_2\Theta \;= 2Ta^{1/2}\mathcal{Q}\Psi + \mathbf{r}\cdot\nabla\times\nabla\times(\boldsymbol{\omega}\times\mathbf{v}), \tag{3}$$

$$\left(\sigma\partial_t - \nabla^2\right)\Theta - Ra\,L_2\Phi \;=\; -(\mathbf{v}\cdot\nabla)\Theta, \tag{4}$$

where $\boldsymbol{\omega} = \nabla \times \mathbf{v}$ is the vorticity.

The non-dimensional parameters of the problem are the Rayleigh, Ra, Prandtl, σ, and Taylor, Ta, numbers, defined as

$$Ra = \frac{\beta\gamma\alpha\Delta T r_o^4}{\kappa\nu}, \quad Ta^{1/2} = \frac{\Omega r_o^2}{\nu}, \quad \sigma = \frac{\nu}{\kappa}, \tag{5}$$

where $\beta = q/3\kappa c_p$, q is the (uniform) rate of heat generation per unit mass, κ is the thermal diffusivity, and c_p the specific heat at constant pressure. In these non-dimensional units the conduction state is $T_c(r) = T_0 - Ra\,r^2/2\sigma$.

The operators L_2 and $\mathcal{Q}$ are defined by $L_2 = -r^2\nabla^2 + \partial_r(r^2\partial_r)$, $\mathcal{Q} = r\cos\theta\nabla^2 - (L_2 + r\partial_r)(\cos\theta\partial_r - r^{-1}\sin\theta\partial_\theta)$.

Stress-free and perfect thermally conducting boundary conditions are assumed at $r = 1$, which become

$$\partial_r(\Psi/r) = 0, \quad \Phi = 0, \quad \partial_{rr}^2\Phi = 0, \quad \text{and} \quad \Theta = 0, \tag{6}$$

in terms of the velocity potentials and the temperature perturbation.

To discretize the equations the functions $X = (\Psi, \Phi, \Theta)$ are first expanded in spherical harmonic series with a triangular truncation of maximal degree L, namely

$$X(t, r, \theta, \varphi) = \sum_{l=0}^{L}\sum_{m=-l}^{l} X_l^m(r, t)Y_l^m(\theta, \varphi), \tag{7}$$

with $\Psi_l^{-m} = \overline{\Psi_l^m}$, $\Phi_l^{-m} = \overline{\Phi_l^m}$, $\Theta_l^{-m} = \overline{\Theta_l^m}$, and imposing $\Psi_0^0 = \Phi_0^0 = 0$ to uniquely determine the two scalar potentials. The spherical harmonics are normalized as

$$Y_l^m(\theta,\varphi) = \sqrt{\frac{2l+1}{2}\frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta) e^{im\varphi} \quad l \geq 0, \quad 0 \leq m \leq l,$$

P_l^m being the associated Legendre functions of degree l and order m. The equations (2)-(4) written for the complex coefficients become

$$\partial_t \Psi_l^m = \mathcal{D}_l \Psi_l^m + \frac{1}{l(l+1)} \left[2Ta^{1/2} \left(im\Psi_l^m - [\mathcal{Q}\Phi]_l^m \right) - [\mathbf{r}\cdot\nabla\times(\boldsymbol{\omega}\times\mathbf{v})]_l^m \right], \tag{8}$$

$$\partial_t \mathcal{D}_l \Phi_l^m = \mathcal{D}_l^2 \Phi_l^m - \Theta_l^m + \frac{1}{l(l+1)} \Big[2Ta^{1/2} \left(im\mathcal{D}_l\Phi_l^m + [\mathcal{Q}\Psi]_l^m \right) + [\mathbf{r}\cdot\nabla\times\nabla\times(\boldsymbol{\omega}\times\mathbf{v})]_l^m \Big], \tag{9}$$

$$\partial_t \Theta_l^m = \sigma^{-1}\mathcal{D}_l\Theta_l^m + \sigma^{-1} l(l+1) Ra\, \Phi_l^m - [(\mathbf{v}\cdot\nabla)\Theta]_l^m, \tag{10}$$

with

$$\mathcal{D}_l = \partial_{rr}^2 + \frac{2}{r}\partial_r - \frac{l(l+1)}{r^2}, \tag{11}$$

and the boundary conditions decouple for each degree and order, i.e.

$$\partial_r(\Psi_l^m/r) = 0, \quad \Phi_l^m = 0, \quad \partial_{rr}^2\Phi_l^m = 0, \quad \text{and} \quad \Theta_l^m = 0 \tag{12}$$

at $r = 1$. The square bracket $[\cdot]_l^m$ indicates extracting the spherical harmonic coefficient of degree l and order m. The operator $\mathcal{Q}$ is

$$[\mathcal{Q}f]_l^m = -(l-1)(l+1)c_l^m D_{1-l}^+ f_{l-1}^m - l(l+2)c_{l+1}^m D_{l+2}^+ f_{l+1}^m, \tag{13}$$

$$\text{with} \quad D_l^+ = \partial_r + \frac{l}{r}, \quad \text{and} \quad c_l^m = \left(\frac{l^2-m^2}{4l^2-1} \right)^{1/2}. \tag{14}$$

To study the stability of the conduction state, which is $(\Psi,\Phi,\Theta) = 0$, the eigenvalue problems

$$\lambda\Psi_l^m = \mathcal{D}_l\Psi_l^m + \frac{1}{l(l+1)} \left[2Ta^{1/2} \left(im\Psi_l^m - [\mathcal{Q}\Phi]_l^m \right) \right], \tag{15}$$

$$\lambda\mathcal{D}_l\Phi_l^m = \mathcal{D}_l^2\Phi_l^m - \Theta_l^m + \frac{1}{l(l+1)} \left[2Ta^{1/2} \left(im\mathcal{D}_l\Phi_l^m + [\mathcal{Q}\Psi]_l^m \right) \right], \tag{16}$$

$$\lambda\Theta_l^m = \sigma^{-1}\mathcal{D}_l\Theta_l^m + \sigma^{-1} l(l+1) Ra\, \Phi_l^m, \tag{17}$$

for $m = 0, 1, \ldots$ have to be solved, λ being the eigenvalue.

Since the operator $\mathcal{Q}$ only couples the spherical harmonics coefficients by their degree, l, and not by their order, m, (see Eq. (13)), a sequence of uncoupled eigenproblems are obtained parameterized by the integer wave number $m \geq 0$.

Due to the triangular truncation used (see Eq. (7)), the number of coefficients involved in the eigenvalue problem corresponding to the wave number m is $3(L-m+1)$ if $m \geq 1$, and $3L$ if $m = 0$ because $\Psi_0^0 = \Phi_0^0 = 0$ and Eq. (17) for Θ_0^0 is uncoupled from the rest and purely diffusive, and consequently never gives rise to instability.

3. Radial discretization

For the discretization of the radial operators collocation methods have been employed. These methods have been used successfully in the case of spherical shells ([33, 34, 35, 37, 38]). They are very efficient due to the possibility of using optimized matrix-matrix products to evaluate the action by the operators in order to compute the nonlinear terms in the time evolution of the full non-linear equations. Moreover the simplicity of collocation methods has also allowed to change easily the boundary conditions and developing different high order time integrators. With collocation methods each radial operator in Eqs. (15)-(17) must be substituted by an approximate derivative matrix associated to a given radial mesh of points. This mesh might include or not the origin at which the equations become singular. This section is devoted to study different radial meshes.

3.1. Derivation matrices and implementation of boundary conditions

When spectral collocation methods are used in the radial coordinate, the derivatives are approximated at the points of a collocation mesh of points r_i, $i = 0, \ldots, n_r$ by formulas of the form

$$f^{(k)}(r_i) \approx \sum_{j=0}^{n_r} d_{i,j}^{(k)} f(r_j), \quad i = 0, \ldots, n_r, \quad k = 1, 2, \ldots \tag{18}$$

It is assumed here that $r_0 = 0$ and $r_{n_r} = 1$. If the mesh does not include the origin the sum in (18) starts with $j = 1$, and the derivatives at $r = 0$ are not evaluated. When boundary value problems for f are solved the values $f(r_i)$ have to be found. The symbol f_i is used for the approximation obtained to $f(r_i)$. The method used to obtain the coefficients of formulas (18) is that described in [41]. The codes provided in [6] were used, written in the extended precision provided by the gfortran compiler of the Gnu compiler collection to minimize the error. It must be said that for particular meshes (Chebyshev-Gauss-Lobatto points, for instance) closed form expressions for the coefficients are available for instance in [2], at least for $k = 1, 2$.

Implementing homogeneous Dirichlet boundary conditions is trivial. If, for instance, $f(1) = 0$ then f_{n_r} is no longer an unknown and the sums in (18) end at $j = n_r - 1$. For the implementation of Neumann and Robin boundary conditions, or for fourth–order problems we follow the methods described, among others, in [2, 42, 7]. In all cases considered the values of f at r_0 and r_{n_r} can be written, from the boundary conditions, as linear combinations of those at the inner

Symbol	Points ($i = 1, \ldots, n_r$)	Description
CGL	$r_i = (1 - \cos(\pi i/n_r))/2$	Chebyshev–Gauss–Lobatto
SCG	$r_i = (2 - \cos(\pi(2i-1)/(2n_r))/2$ $+ \cos(\pi(2n_r - 1)/(2n_r)))$	Shifted Chebyshev–Gauss
CGR	$r_i = (1 + \cos(2\pi(n_r - i)/(2n_r - 1)))/2$	Chebyshev–Gauss–Radau
HCGL	$r_i = -\cos(\pi(i + n_r - 1)/(2n_r - 1))$	Half Chebyshev–Gauss–Lobatto
ES	$r_i = i/n_r$	Equally spaced points

Table 1: Radial meshes used in this study.

points. After substituting these expressions in Eqs. (18), derivative matrices are obtained which include the boundary conditions

$$f^{(k)}(r_i) \approx \sum_{j=1}^{n_r-1} \tilde{d}_{i,j}^{(k)} f(r_j), \quad i = 1, \ldots, n_r - 1, \quad k = 0, 1, 2, \ldots, \tag{19}$$

and only involve the values at the inner points. The Appendix includes the details for a case with two conditions at $r = 1$.

The radial differential operators in Eqs. (15)-(17) are substituted by the corresponding approximating matrices, giving rise to a complex eigenvalue problem. The matrix at the right side is block-tridiagonal due to the coupling of the spherical harmonics of adjacent degrees by the operator $\mathcal{Q}$ (see Eq. (13)). It must be noticed that with the boundary conditions used here the operator $\mathcal{D}_l$ appearing at the left side of Eq. (16) is invertible and therefore the eigenvalue problem (15)-(17) can be written as

$$A_m x = \lambda x, \tag{20}$$

x being the vector containing the values at the collocation points of the amplitudes in spherical harmonics of the potentials and the temperature perturbation, of order m and degrees $m \leq l \leq L$, $(\Psi_l^m(r_i), \Phi_l^m(r_i), \Theta_l^m(r_i))$. The matrix A_m keeps the block-tridiagonal structure.

3.2. Test for the radial meshes

Several radial meshes have been considered. Some of them have been used in polar coordinates by different authors [14, 19, 15, 20]. They are listed in Table 1. Stretched versions have been also considered to separate the nodes accumulated at the origin. Some researchers claim that this is needed to satisfy Courant stability conditions in time integrators. Although this is not our case, because we are interested in solving eigenvalue problems, we have included this possibility just for comparison purposes. The transformation used is $r_i^{(t)} = r_i^s$ with a power $s < 1$, where r_i are the points of the meshes in the table, and $r_i^{(t)}$ refers to the stretched mesh. The meshes considered, not including the origin, are (see [2]):

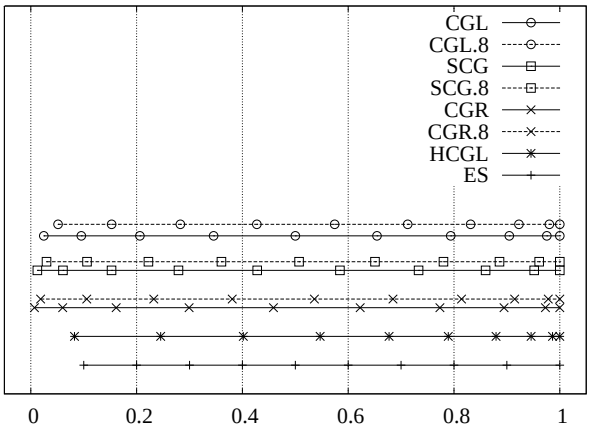

Figure 1: Radial meshes used in this study for $n_r = 10$, when the origin $r = 0$ is not included.

- Chebyshev-Gauss-Lobatto points (CGL) given by

$$r_i = (1 - \cos(\pi i/n_r)/2, \quad i = 1, \cdots, n_r,$$

which are the zeros of the polynomial $q_{n_r+1}(x) = T_{n_r+1}(x) + aT_{n_r}(x) + bT_{n_r-1}(x)$ in the interval $[-1, 1]$ transformed into the interval $[0, 1]$ by means of $r = (x + 1)/2$. Here $T_k(x)$ is the Chebyshev polynomial of the first kind of degree k and the coefficients a and b are selected such that $q(-1) = q(1) = 0$. This is the mesh most usually employed in pseudo-spectral collocation methods.

- Shifted Chebyshev-Gauss points (SCG) given by

$$r_i = (2 - \cos(\pi(2i - 1)/2n_r)/2 + \cos(\pi(2n_r - 1)/2n_r), \quad i = 1, \cdots, n_r,$$

which are the zeros of the polynomial $T_{n_r}(x)$ transformed, as before, into the interval $[0, 1]$ and shifted to the right to make $r_{n_r} = 1$. The original Chebyshev-Gauss points in the interval $[-1, 1]$ minimize the maximum norm, $\|f\|_\infty = \sup_{r\in[-1,1]} |f(r)|$, of the polynomial $(x - x_1) \cdots (x - x_{n_r})$ among all possible meshes of n_r points in the interval. This minimizes the interpolation error in this norm.

- Chebyshev-Gauss-Radau points (CGR) given by

$$r_i = (1 + \cos(2\pi(n_r - i)/(2n_r - 1)))/2, \quad i = 1, \cdots, n_r$$

which are the zeros of the polynomial $q_{n_r}(x) = T_{n_r}(x) + aT_{n_r-1}(x)$ with $a = -T_{n_r}(-1)/T_{n_r-1}(-1)$ such that $q_{n_r}(-1) = 0$. The zeros are reversed in order and transformed to the interval $[0, 1]$. These points are the nodes of the Chebyshev-Gauss-Radau quadrature formulas used when one of the two extreme points of the interval is not included.

- Half Chebyshev-Gauss-Lobatto points (HCGL) given by

$$r_i = -\cos(\pi(n_r + i - 1)/(2n_r - 1)))/2, \quad i = 1, \cdots, n_r$$

  which are the right half of the $2n_r$ Chebyshev-Gauss-Lobatto points in the interval $[-1, 1]$ transformed to $[0, 1]$. It has been included just for comparison purposes, to add another example of the effect of transforming the previous meshes to separate the points close the origin from it.

- Equally spaced mesh (ES), given by

$$r_i = i/n_r, \quad i = 1, \cdots, n_r.$$

  It has also been included for comparison purposes, to see the growth of the Lebesgue constant with n_r.

All the sets of points have been defined such that the index $i = n_r$ corresponds to the point $r = 1$. Figure 1 displays the main five meshes and three stretched ones for $s = 0.8$, not including the origin. The suffixes .8 or .9 after the name of a mesh indicate the value of s used in the transformation.

In the case of polar cylindrical coordinates a way of obtaining spectral accuracy for the approximation of boundary-value problems in the unit disk is to take the radius $r \in [-1, 1]$ instead of $r \in [0, 1]$ and the angle $\theta \in [-\pi/2, \pi/2]$ instead of $\theta \in [0, 2\pi]$, and to use an even number of Chebyshev-Gauss-Lobatto points in $[-1, 1]$ to avoid the origin (see [17, 6] for details). This method could also be used in spherical coordinates but in this case it involves both the longitude θ and the colatitude φ, since the point $(-r, \theta, \varphi)$ is the same as $(r, \pi - \theta, \varphi + \pi)$. This makes things much more complicated than for polar coordinates. Since we are interested in simple extensions of the methods already used for a spherical shell, this possible method has not been considered.

Before using these meshes in the stability problem for the convection in spheres they have been compared by computing the dependence with n_r of their Lebesgue constant, the condition number of the discretized Bessel operator and its first and fifth eigenvalues.

The Lebesgue constant of the mesh r_i, $i = 1, \ldots, n_r$ is defined as

$$\Lambda_{n_r} = \max_{r \in [0,1]} \sum_{i=1}^{n_r} |L_i(r)|, \tag{21}$$

where

$$L_i(r) = \prod_{\substack{j=1 \\ j \neq i}}^{n_r} \frac{r - r_j}{r_i - r_j}, \quad i = 1, \ldots, n_r, \tag{22}$$

are the Lagrange polynomials satisfying $L_i(r_j) = \delta_{ij}$. When the origin is included in the collocation mesh the indices in the last two formulas start at $i = 0$ and $j = 0$, respectively. We recall that the Lebesgue constant satisfies

$$\|f - p_{n_r}\|_\infty \leq (1 + \Lambda_{n_r})\|f - p_{n_r}^{opt}\|_\infty,$$

where f is an arbitrary function, p_{n_r} is the interpolation polynomial of degree n_r based on the collocation mesh, and $p_{n_r}^{opt}$ is the best polynomial approximation in the maximum norm of degree n_r. Roughly speaking this inequality says that the interpolation polynomial is at most a factor $1+\Lambda_{n_r}$ worse than the best possible approximation in the maximum norm. It is known that Λ_{n_r} grows at least as $O(\ln n_r)$ for any distribution of interpolation nodes. For the Chebyshev-Gauss points it is $\Lambda_{n_r} \sim O(\ln n_r)$, and for equally spaced points $\Lambda_{n_r} \sim O(2^{n_r}/n_r \ln n_r)$ (see [43, 6] for instance).

The spherical Bessel equation is

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{df}{dr}\right) - \frac{l(l+1)}{r^2}f = -\lambda f, \quad 0 < r < 1, \quad l = 0, 1, \ldots, \tag{23}$$

to which we add the boundary condition $f(1) = 0$. Its eigenvalues are $\lambda_{lp} = r_{lp}^2$, $p = 1, 2, \ldots$, where r_{lp} are the zeros of the spherical Bessel function

$$j_l(r) = \sqrt{\frac{\pi}{2r}}\, J_{l+1/2}(r), \tag{24}$$

$J_{l+1/2}(r)$ being the cylindrical Bessel function of order $l + 1/2$ [44]. The corresponding eigenfunctions are $f_{lp}(r) = j_l(r_{lp}r)$. We will refer to the operator defined by the left hand side of Eq. (23) as the Bessel operator. It is the operator $\mathcal{D}_l$ already defined in Eq. (11). It was discretized by substituting the derivatives by the approximation derivative matrices on the test meshes. When the origin is included in the mesh a regularity condition has to be imposed there. Since the spherical Bessel functions satisfy

$$j_l(r) \approx r^l/(2l+1)!! \tag{25}$$

for $r \ll 1$ [44], and Eq. (23) is of second order, we have just added the condition $f(0) = 0$ for $l > 0$. The case $l = 0$ has not been considered in these preliminary tests, but the corresponding boundary condition would be $df/dr(0) = 0$. See below for the conditions used in the case of the stability problem of the fluid sphere.

The maximum and spectral norms of the discretized spherical Bessel operator, and its maximum-norm condition number were computed. Their definitions, for a generic matrix A of dimension n, are $\|A\|_\infty = \max_{1\le i\le n}\sum_{j=1}^n |a_{ij}|$, $\|A\|_2 = \sigma_{max}(A)$, $\sigma_{max}(A)$ being the largest singular value of A, and $\kappa_\infty(A) = \|A\|_\infty\|A^{-1}\|_\infty$, respectively.

The approximate eigenvalues, norms and condition numbers were obtained by using the double precision subroutines DGEEV, DLANGE, DGETRF, DGECON and DGESVD of the LAPACK library [45].

Figure 2 shows the dependence of the Lebesgue constant with the number of mesh points. The top and bottom rows of plots correspond to the meshes with or without the origin, respectively. The two left plots include the ES and HCGL meshes for comparison purposes (see table 1 for their definitions), and the right the stretched meshes for $s = 0.9$. It can be seen the growth of Λ_{n_r}

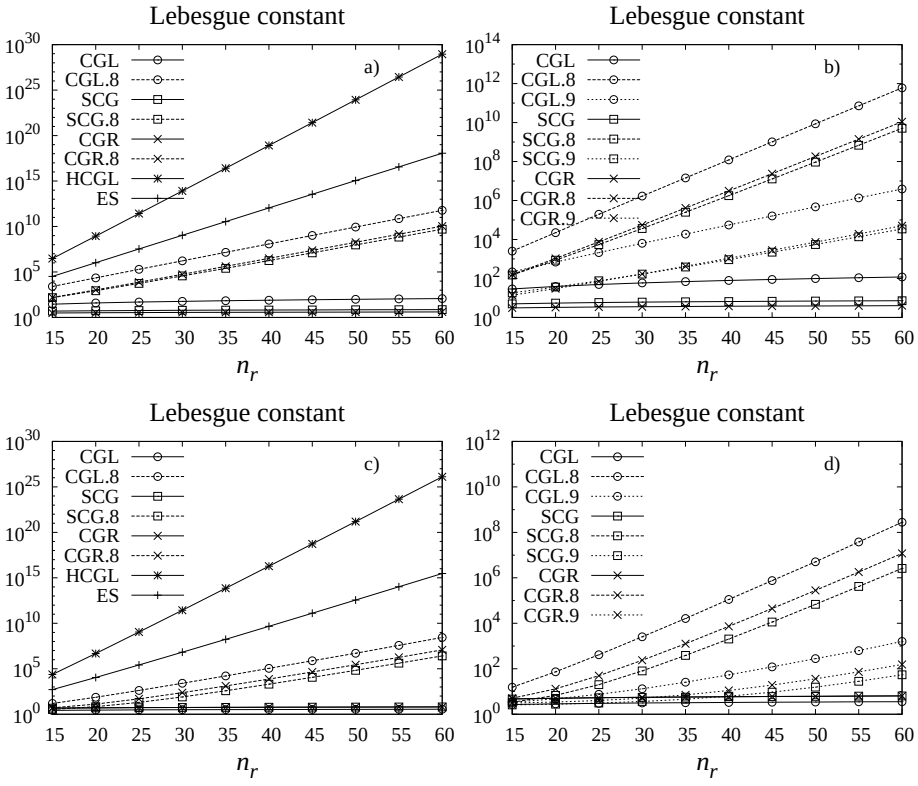

Figure 2: Variation of the Lebesgue constant with the number of mesh points. The top and bottom rows correspond to the meshes without or with the origin, respectively.

for the HCGL mesh is even greater than for the ES, due to the separation of the points close to the origin. The same holds for all the transformed sets of points when s is low enough. The right plots also show the substantial increase in the growth of Λ_{n_r} when s moves from 0.9 to 0.8. The effect of including or not the origin in the mesh can be seen by comparing the two rows of plots. The presence of the origin contributes, as could be expected, to reduce Λ_{n_r} substantially. For instance, in the case of CGL.8 with $n_r = 60$, $\Lambda_{n_r} = O(10^{12})$ if the origin is not included, but $\Lambda_{n_r} = O(10^8)$ if it is. In the case of CGL, also for $n_r = 60$, the two values are $\Lambda_{n_r} = 119$ and $\Lambda_{n_r} = 3.6$. For CGR and SCG the difference is much lesser, with all the values of Λ_{n_r} being of order one. The reason for this difference between the CGL, and CGR and SCG meshes is that the first point of the mesh, r_1, is closer to the origin in the two latter cases (see Fig. 1 for $n_r = 10$). From this first study it can be deduced that the CGL, CGR and SCG are the best meshes. Moreover CGL points have the advantage of the easy closed form for the coefficients of the derivative matrices for orders one and two [2].

The same conclusion can be obtained from the dependence with n_r of the maximum norm, the condition number or the difference between the exact and the approximate eigenvalues of the spherical Bessel operator for $l = 1$ shown in Figs. 3 and 4. The method we use to compute eigenvalues in the fluid sphere problem involves, among other things, the computation of matrix products and the solution of linear systems with the discretized version of $\mathcal{D}_l$, which is the spherical Bessel operator. The latter implies computing its LU decomposition. It is known that the propagation of rounding errors in the evaluation of matrix products or in linear systems solvings is directly related to the norm and the condition number of the matrix [46]. As for the Lebesgue constant the presence of the origin in the collocation mesh helps reducing both the norm and the condition number, and they grow when the stretching power s decreases.

Finally Fig. 4 shows the dependence with n_r of the first and fifth eigenvalues of the Bessel operator for $l = 1$. In the right plots the case $s = 0.9$ has not been included because, as happens for the first eigenvalue, the results are very similar to those of the untransformed meshes. The curves for the stretched meshes show an initial decay of the error, followed by an steady increase, in contrast to those for the original sets of points, for which once the error reaches a threshold, it remains close to constant. From these calculations it might seem that the inclusion of the origin is not as important as before. In next section we will see that in the problem of thermal convection in spheres, spurious eigenvalues of the eigenproblem (15)-(17) are found, for the lower values of the azimuthal wave numbers ($m = 0, 1$), when the origin is not included.

3.3. Numerical techniques for the eigenvalue problem in spherical geometry

After the results of the previous tests, the CGL mesh has always been used in the rest of calculations. The only variation considered was the inclusion or not of the origin.

It is known that if a smooth function f is expanded in spherical harmonics as

$$f(r, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} f_l^m(r) Y_l^m(\theta, \varphi) \tag{26}$$

then $f_l^m(r) = r^l g_l^m(r)$, with $g_l^m(r)$ smooth and even in r. This means that

$$\frac{d^k f_l^m}{dr^k}(0) = 0, \quad k = 0, \ldots, l-1 \tag{27}$$

as for the spherical Bessel functions (see Eq. (25)). Several previous studies support the idea of imposing just a minimum number of these regularity conditions in the numerical approximations. In the case of expansions of smooth functions in Fourier series in the two angular variables on a spherical surface as

$$F(\theta, \varphi) = \sum_{m=-\infty}^{\infty} F_m(\theta) \exp(im\varphi), \tag{28}$$

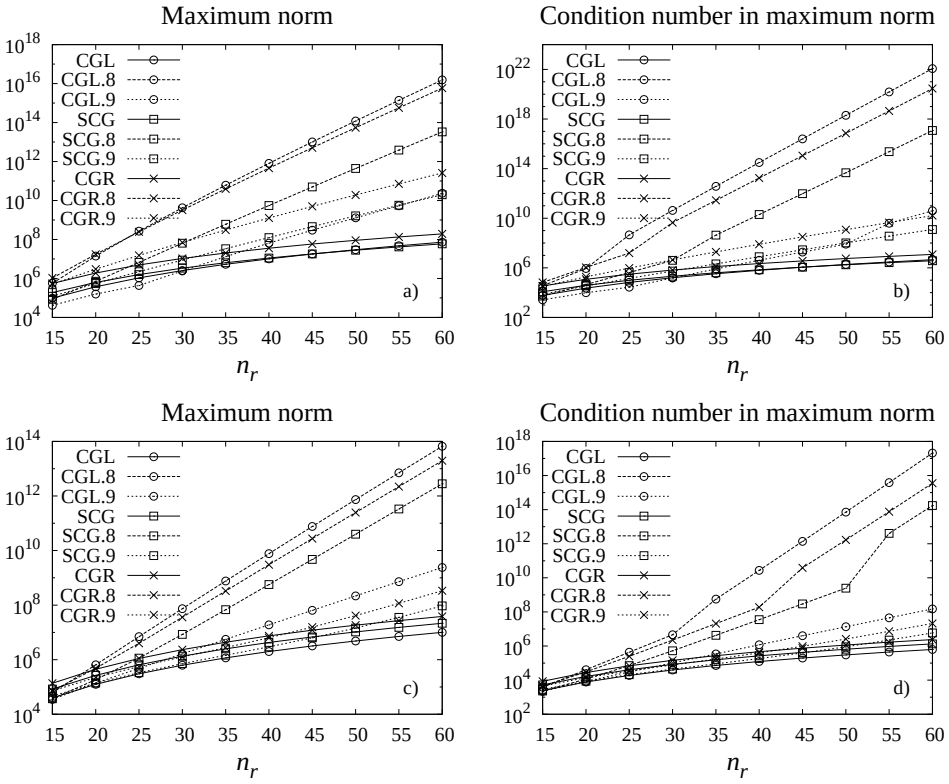

Figure 3: Variation of the maximum norm and the maximum norm condition number of the discretized version of the spherical Bessel operator for $l = 1$. The top and bottom rows correspond to the meshes without or with the origin, respectively.

the coefficient $F_m(\theta)$ has zeros at the poles of order m [22]. It was shown in [21] that it is possible to apply all the above regularity conditions. The modified Robert functions were considered there, finding that they form a very ill-conditioned basis. However, by using parity modified Fourier series, which have at most a first order zero at the poles, very good approximations to the eigenvalues of the Laplacian were found in [23]. In the case of polar coordinates the regularity conditions at the origin are exactly the same as for the spheres at the poles, as can be seen by just changing θ by r in Eq. (28). A Chebyshev-tau solution of the eigenvalue problem for the cylindrical Bessel's equation was computed in [1], which was highly improved by enforcing a zero-derivative of the function at the origin. The same holds in the collocation case [14] (see [24] for more details and additional remarks on the bibliography). Therefore, this has also been done in our calculations when the mesh includes the origin. Just the boundary condition $f(0) = 0$ is imposed as in the previous tests, independently of the order of the equations. For the fourth order equation one could add two regularity conditions, but we have found it unnecessary. Since the coefficients

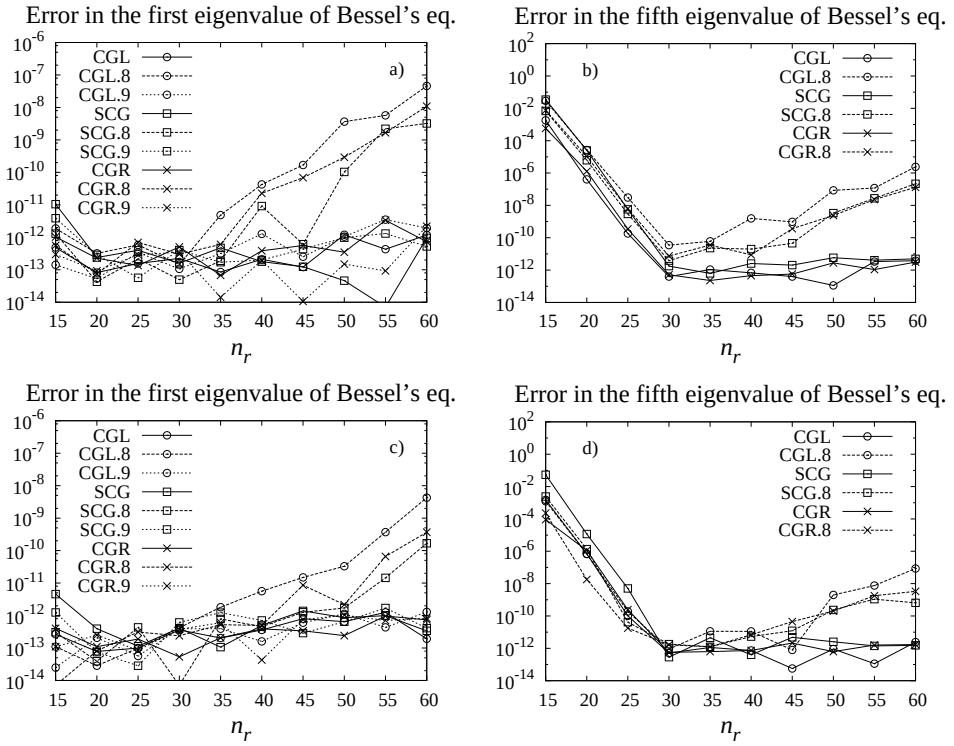

Figure 4: Variation of the error in the approximation of the first and fifth eigenvalue of the spherical Bessel operator for $l = 1$. The top and bottom rows correspond to the meshes without or with the origin, respectively.

of the potentials and temperature perturbation with $m = 0$ and $l = 0$ do not appear in the stability analysis (see Sect. 2), the condition $df/dr(0) = 0$ need not be used. It should be implemented just for Θ_0^0 in a time evolution code for Eqs. (8)-(10).

Another way of imposing regularity conditions is to add a prefactor r^l, substituting the prognostic functions f_l by $r^l f_l$. It is known that large values of l lead to ill-conditioning [24], similarly to what happens with the modified Robert functions. However, lower values of l can be used safely. A recent implementation of this procedure can be found in [47], where a multi-layer code for the simulation of stellar core convection was presented.

In order to compute the leading eigenvalues (maximal real part) of the problem (15)-(17), we use the following technique. Assume that the problem is written as $A_m x = \lambda x$ (see comments before Eq. (20)). Given a complex shift ξ, the eigenvalues of

$$(A_m - \xi I)^{-1} \tag{29}$$

of maximal modulus (a fixed number of them) are found by using Arnoldi's method implemented in subroutine DNAUPD of the ARPACK library [48]. This

shift-invert transformation provides the eigenvalues of A_m closest to ξ. If (x, μ) is an eigenpair of $(A_m - \xi I)^{-1}$, then $(x, \xi + 1/\mu)$ is an eigenpair of A_m. Instead of using a single shift ξ we repeat the process with a set ξ_k, $k = 1, \ldots, k_{max}$, all of them close to the imaginary axis, eliminating repeated eigenvalues found for different values of ξ. In order to select the ξ_k, information from previous cases with nearby parameters is used. Some experiments were done using Cayley transformations of the form $(A_m - \zeta I)(A_m - \xi I)^{-1}$, but the performance was poor. It is known that in the presence of eigenvalues of large negative real part or large imaginary part, the convergence to those of maximal real part can be very slow [49] with this transformation.

The subroutine DNAUPD needs a user code computing the action by the operator $(A_m - \xi I)^{-1}$ or, equivalently, which solves systems with matrix $A_m - \xi I$. The matrix A_m of the eigenvalue problem (15)-(17) has block tridiagonal structure due to the coupling between the spherical harmonic coefficients of the same order m and degrees $l - 1$, l and $l + 1$ coming from the operator $\mathcal{Q}$ (see Eq. (13)). The subtraction of the term ξI does not alter this structure. To solve these linear systems an adapted block LU decomposition is performed.

The initial condition for Arnoldi's method can be a random vector, but we have observed that for some values of the parameters of the problem and the azimuthal wave number m, spurious eigenvalues appear. This is due to the presence of components in the initial condition along the eigenvectors associated to eigenvalues of large negative real part, that Arnoldi's method combined with the shift-invert strategy seems to be unable to filter completely. To avoid them and to improve the convergence we use a previous filtering process based on time integration. Suppose that x^0 is a random initial condition. Then we apply several steps of the implicit Euler's method

$$(I - \Delta t A_m) x^{k+1} = x^k, \quad k = 0, 1, \ldots. \tag{30}$$

It is known that this integration damps very quickly the components of x^0 along the eigenvectors corresponding to eigenvalues of A_m of large negative real part. Since implicit Euler's method is unconditionally stable and we are only interested in its filtering properties, Δt can be much larger than that required to have an accurate integration. This helps to minimize the CPU time involved in this process. Moreover the linear systems to be solved during the time integration have the same block-tridiagonal structure than A_m. In fact, Eq. (30) can be written as $(A_m - (1/\Delta t) I) x^{k+1} = -(1/\Delta t) x^k$, where the matrix has the form of a real shift-invert transformation (see Eq. (29)).

In order to find the critical values for the onset of convection the zeros of $\mathrm{Re}(\lambda_{max})$ are found by means of Brent's method [50], λ_{max} being the eigenvalue of maximal real part obtained by the swept of shifts just described. It is very important to get rid of all possible spurious eigenvalues for this method to work consistently when curves of bifurcation on two-dimensional parameter spaces are traced.

In the axisymmetric case, $m = 0$, with the stress-free boundary conditions used in this study, there is always a zero eigenvalue of Eqs. (15)-(17). The

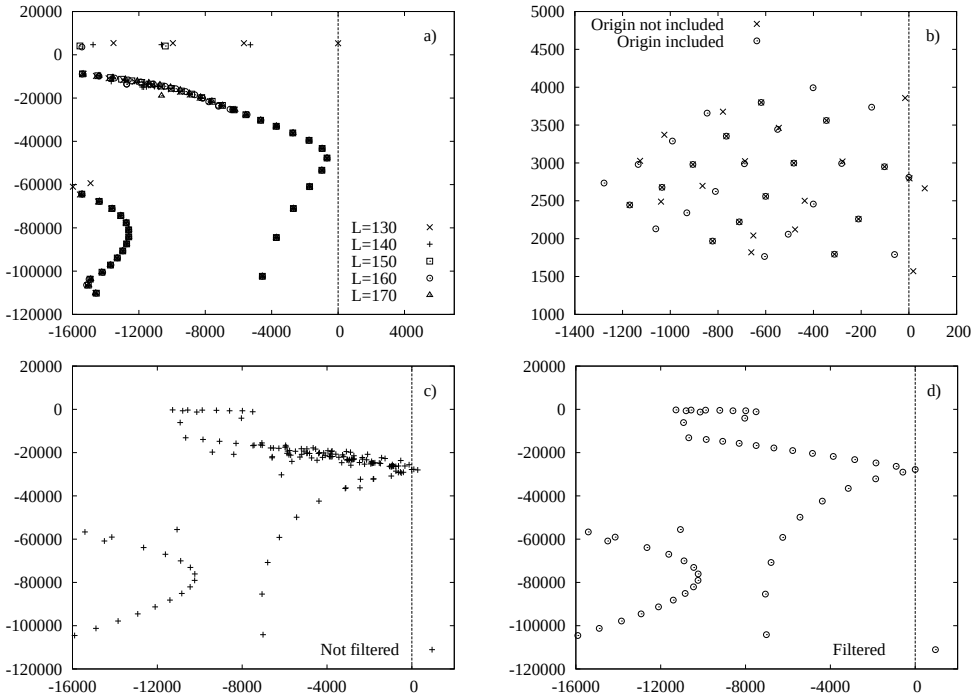

Figure 5: Examples of spurious eigenvalues.

corresponding eigenfunction is given by $\psi_1^0 = cr$ with c a non-zero constant, and the rest of coefficients of ψ and those of ϕ and Θ equal to zero. Taking into account that $Y_1^0(\theta, \varphi) = \sqrt{3/2}\cos\theta$, and using Eq. (1), it turns out that the corresponding velocity field is $(v_r, v_\theta, v_\varphi) = (0, 0, cr\sqrt{3/2}\sin\theta)$ which is a rigid rotation about the $\mathbf{k}$ axis. This zero eigenvalue has to be removed from the spectra to be able to find the critical value of the parameters as described in the previous paragraph. For this purpose we use Wieland's deflation on A_0 [51]. If a matrix A of dimension n has a spectrum $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ and λ_1 is simple with eigenvector u_1, then the matrix $A - \xi u_1 u_1^\top$ has spectrum $\{\lambda_1 - \xi, \lambda_2, \ldots, \lambda_n\}$, i.e. λ_1 is shifted to $\lambda_1 - \xi$. In our case $\lambda_1 = 0$ and we take $\xi = -1000$. Since our method finds the eigenvalues of maximal real part, the shifted one will not interfere in the computation of the zeros of $\mathrm{Re}(\lambda_{max})$. Moreover, since the corresponding eigenvector has a single non-vanishing spherical harmonic coefficient, ψ_1^0, the matrix $u_1 u_1^\top$ is zero except for a single diagonal block of dimension $n_r - 1$, letting the block-tridiagonal structure of A_0 unaltered.

3.4. Spurious eigenvalues

Several sources of spurious eigenvalues have been found. Boyd [3] distinguishes between *physically spurious eigenvalues*, which are in error because of misapplication of boundary conditions or some other misrepresentation of the

physics, and *numerically spurious eigenvalues*, which are poor approximations to exact eigenvalues because the mode is oscillating too rapidly to be resolved by the degrees of freedom of the discretization. In principle the latter could always be eliminated by increasing the resolution of the discretization. In our particular problem, the way the operator

$$[\mathcal{Q}f]_l^m = -(l-1)(l+1)c_l^m D_{1-l}^+ f_{l-1}^m - l(l+2)c_{l+1}^m D_{l+2}^+ f_{l+1}^m,$$

is truncated is also relevant. If we truncate at a degree $l = L$ several options are available. One is just forgetting the term $-L(L+2)c_{L+1}^m D_{L+2}^+ f_{L+1}^m$, but this leads in many cases to the appearance of spurious modes. Similar problems occur in time evolution codes in spheres or spherical shells when the nonlinear terms are truncated. This was noticed in [52] and by the authors in their implementations of codes in the case of spherical shells [34, 35, 37, 38]. In some cases the numerical instability shows up only after long time integrations. This makes the diagnostic of the problem very difficult unless the stability problem of the conduction state is first studied in some detail, including azimuthal wave numbers, m, which are not necessarily those which give the critical parameters for the onset of convection. A second option is taking a higher L when spurious modes appear, but this is not a systematic approach and the required L can be much larger. A third method consists in eliminating completely the term $[\mathcal{Q}f]_L^m$ in the equation for the order m and degree L. If the solutions are symmetric with respect to the equatorial plane, only the even or odd coefficients of the potentials and the temperature perturbation ($\Psi_l^m(r_i), \Phi_l^m(r_i)$, and $\Theta_l^m(r_i)$) are nonzero. Therefore enforcing $[\mathcal{Q}f]_L^m = 0$ can be done for some equations only, depending on the parity of m. This has been the option adopted in our codes to minimize the truncation parameter L and thus the size of the systems to be solved.

As mentioned before we have found another source of numerical spurious modes; those found by the Arnoldi's eigenvalue solver which cannot completely filter the eigenvalues of large negative real part. We have found that they do not disappear just by increasing the resolution of the discretization, on the contrary more spurious modes might appear.

Figure 5 shows three examples of spurious eigenvalues, corresponding to the three mentioned possible sources. In Fig. 5a the values of the parameters are $\sigma = 0.037$, $Ta = 5 \times 10^{12}$ and $Ra = 4.44 \times 10^8$, the azimuthal wave number is $m = 21$, the mesh is CGL including the origin, and the truncation parameters are $n_r = 70$ and L from 130 to 170 by increments of 10 units. The shifts used have real part 0.01 and imaginary parts ranging from 0 to -10^5. There is a set of spurious eigenvalues of positive imaginary part at the top of the plot. That on the imaginary axis disappears when L is incremented from 130 to 140. The second, with real part close to -5700, disappears when L changes from 140 to 150, and the third, with real part close to -10000, when L changes from 150 to 160. It is clear that if $L < 130$ the critical Rayleigh number for the onset of convection, and the corresponding frequency and eigenmode are wrong. To find the right ones it is necessary to set $L > 130$. Only an increase in a few

units is enough to move the first spurious eigenvalue to the left the required distance to allow the codes to detect correctly the critical point, which has an imaginary part close to -48000. This is an example of spurious modes due to the lack of resolution of the discretization. They are recognized, apart from being strongly dependent on L, because the eigenfunctions present oscillations with the frequency of the radial mesh.

Figure 5b shows two spectra for $\sigma = 0.01$, $Ta = 10^7$ and $Ra = 1.84 \times 10^4$, and $m = 0$. A single shift $0.01 + 2800i$ close to the right critical eigenvalue was used to compute them. The two meshes are CGL with or without the origin. When it is included (circles) the eigenfunctions are smooth. When not, they oscillate as in the previous example, indicating spurious modes. We have found this behavior only for low values of m. When it is large enough, the fast decay of the amplitudes of the spherical harmonics together with their derivatives for $r \to 0$ makes the presence of the origin less important, allowing using meshes with only internal points and the left limit $r = 1$. The differences between the results with the two types of meshes are then not significant.

Finally, Figs. 5c and 5d show two leading spectra for $\sigma = 0.1$, $Ta = 5 \times 10^{12}$ and $Ra = 9.09 \times 10^8$, and $m = 20$, the mesh is CGL including the origin, and the truncation parameters are $n_r = 70$ and $L = 130$. The shifts used have real part 0.01 and imaginary parts form 0 to -10^5. If the initial condition for Arnoldi's method is not filtered (Figs. 5c) a change in the shift produces eigenvalues slightly different which are difficult to recognize as being the same. It can be observed that this happens for the eigenvalues of imaginary part close to -20000, where there is a cloud of eigenvalues. We think that the reason for this spreading is the bad conditioning of the matrix $A_m - \xi I$ for ξ close to the critical eigenvalue, which makes the eigenvalues strongly dependent on ξ. The filtering process makes that the action of this matrix be restricted to the subspace generated by the eigenvector of eigenvalues with real part closer to the imaginary axis. This would enhance the conditioning of the action by $(A_m - \xi I)^{-1}$, reducing the noise in the computations of the spectra as can be seen in the filtered case in Fig. 5d.

4. Results on the stability analysis

To show the reliability of the method described, the Taylor number dependence of the critical Rayleigh number, Ra_c, precession frequency, ω_c, and wave number, m_c, for the onset of convection have been computed. A moderate Prandtl number ($\sigma = 0.78$) is first considered, and the most difficult and expensive case of large Ta is undertaken. The resolutions used are in the ranges $50 \leq n_r \leq 70$ and $80 \leq L \leq 180$, depending on the azimuthal wave number, to obtain Ra_c with errors below 1% and much lesser for ω_c. The results are summarized in Fig. 6. In the first plot, the left axis and the solid curve correspond to critical Rayleigh number, while the discontinuous line and the right axis indicate the drifting frequency. The neutral curves are the envelopes of others sixty, one for each of the wave numbers taken into account and shown in

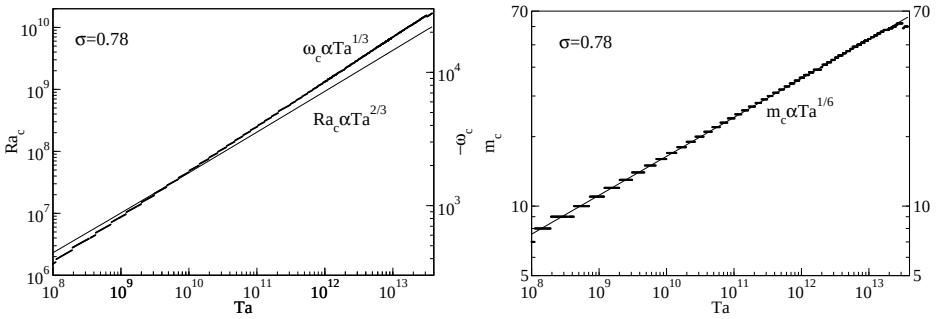

Figure 6: At left, critical Rayleigh number (solid line) and minus the precession frequency (discontinuous curve), and at right longitudinal wave number versus the Taylor number.

the second figure with a discontinuous line. The curves fit very well to the theoretical power laws predicted in the asymptotic theories of [53, 54, 55, 56, 57, 58]. The numerical laws found are $Ra_c = 12.744 \times Ta^{0.656}$, $\omega_c = -0.771 \times Ta^{0.337}$ and $m_c = 0.339 \times Ta^{0.169}$ (solid line of the right plot of Fig. 6).

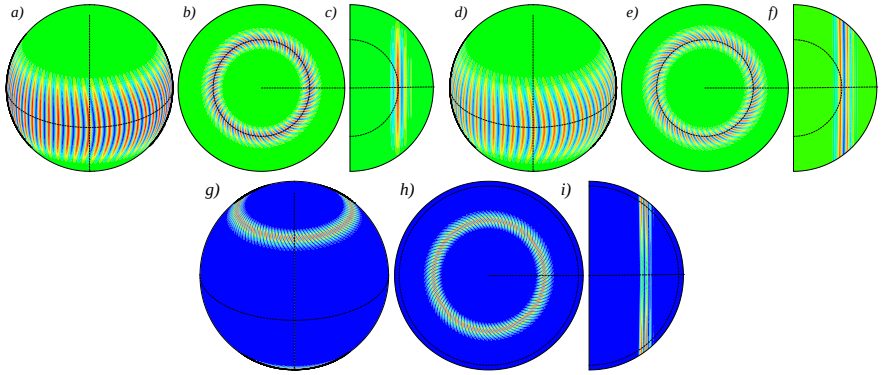

Figure 7: Contour plots of (a-c) Θ, (d-f) v_φ, and (g-i) E_K, of the eigenfunction with $Ta = 7.03 \times 10^{12}$, $Ra_c = 3.38 \times 10^9$, $\omega_c = -1.64 \times 10^4$ and $m_c = 50$. Light grey (green) means zero for Θ and v_φ, and black (blue) means zero for E_K. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

A sample of the eigenfunctions that can be found along the neutral curves of Fig. 6 are drawn in Fig. 7. Each set of three figures contains projections of Θ, the longitudinal velocity v_φ and the kinetic energy density E_K on the sphere, and on the equatorial and a meridional plane. The lines superposed in each surface indicate the place where the other two projections are taken. The patterns of convection are spiraling columnar (SP modes), very localized near the mid radius. The velocity field fulfills quite well the Taylor-Proudman theorem, i.e. it is quite independent of the axial coordinate, and E_K reach the maximum in the outer surface at mid latitudes, in agreement with early experimental

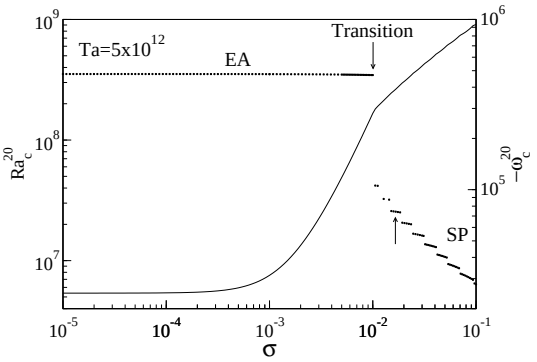

Figure 8: Critical Rayleigh number Ra_c^{20} (solid line) and minus critical precession frequency ω_c^{20} (discontinuous curve) of the solutions of longitudinal wave number $m = 20$ depicted versus σ.

observations [59]. The higher Ta the narrower the spiral. According [58] this solutions can be described as a superposition of quasi-geostrophic inertial-wave modes.

The behavior of the problem at low Prandtl numbers is very different. The main characteristic is that the patterns of convection depend on the value of σ, and the preferred eigenfunctions are even axisymmetric in the $\sigma \to 0$ limit [60]. Therefore, it does not exist in this case a single theory describing the onset of convection. To see what kind of solutions are preferred, the neutral curve of a selected longitudinal wave number $m = 20$ has been computed and represented in Fig. 8 for $\sigma < 0.1$ and $Ta = 5 \times 10^{12}$. The axis and types of lines mean the same as in Fig. 6. Up to $\sigma \approx 0.01$ the preferred modes of convection are equatorially attached (EA). The cells of convection and the vortices of the velocity field are trapped close to the boundary of the sphere near the equator, forming a wave guide (see the first row of Fig. 9). This solutions can be represented essentially by a single quasi-geostrophic inertial wave mode [58]. For these solutions for $\sigma < 0.001$, Ra_c is very low and ω_c very high and constant in the full interval. At $\sigma \approx 0.01$ there is an interchange of neutral curves signaled by the change of slope of the curve of Ra_c and the jump of that of ω_c, of almost and order of magnitude. From this point to 0.018, approximately, there is a small range of σ in which the cells of Θ spread and spiral to the interior of the fluid forming several humps and finally detach from the equator, but the vortices still remain attached to the outer boundary, despite they also spiral to the interior. Contour plots of these solutions are depicted in the second and third rows of Fig. 9. At higher σ the patterns of convection are SP, more or less localized depending on σ (see fourth row of Fig. 9). The arrows in Fig. 8 mark the region between the EA and the SP modes.

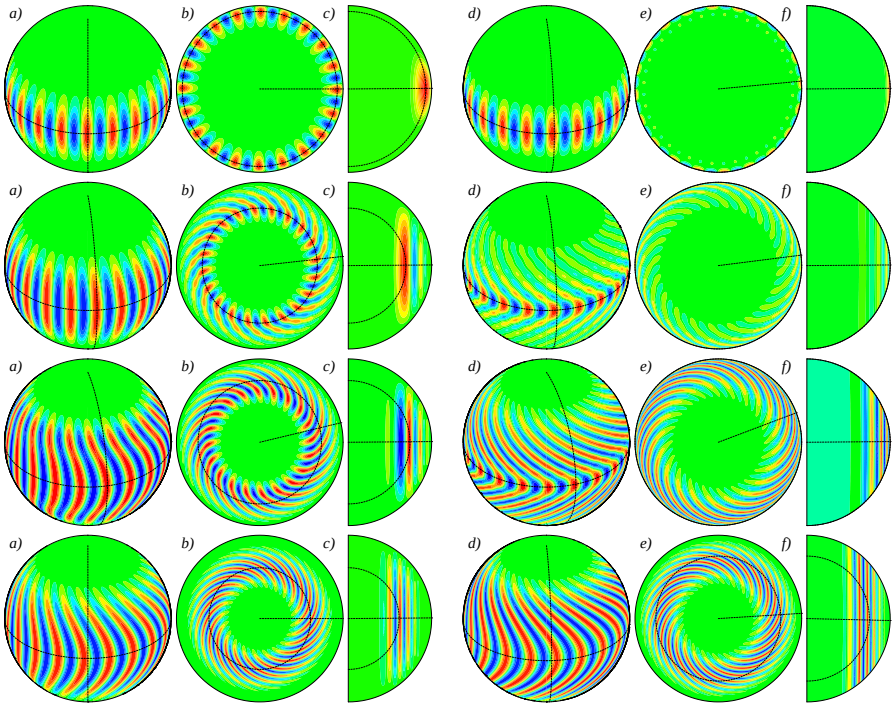

Figure 9: Contour plots of (a-c) Θ and (d-f) v_φ for: First row $\sigma = 1. \times 10^{-4}$, $Ra_c^{20} = 5.40 \times 10^6$, and $\omega_c^{20} = -4.78 \times 10^5$. Second row $\sigma = 1.11 \times 10^{-2}$, $Ra_c^{20} = 1.90 \times 10^8$, $\omega_c^{20} = -1.05 \times 10^5$. Third row $\sigma = 1.75 \times 10^{-2}$, $Ra_c^{20} = 2.74 \times 10^8$ and $\omega_c^{20} = -7.39 \times 10^4$. Fourth row $\sigma = 0.1$, $Ra_c^{20} = 9.09 \times 10^8$ and $\omega_c^{20} = -2.78 \times 10^4$.

5. Conclusions

The study of section 3.2 shows that collocation on a Gauss-Lobatto mesh (or similar) together with the inclusion of simple regularity conditions at the origin provides an accurate way of discretizing the radial part of equations in spherical domains. This has been checked having in mind a spherical-harmonic Galerkin projection in the angular variables. Stretching the meshes, separating the nodes from the origin, or avoiding the inclusion of the origin lead to larger interpolation errors, and worst conditioning of the derivative matrices and approximation of the spectra of eigenvalue problems.

The study was motivated by a problem of astrophysical interest; the convection of a self-gravitating spherical fluid mass, subject to rotation and internal heating. The stability of the basic conduction state has been used as a non-trivial test of the collocation methods. The appearance of spurious modes, not detected in the simple initial study, shows the importance of testing the algorithms in realistic situations. Three kinds of spurious modes have been found

and ways to eliminate them proposed. An example is their appearance due to the lack of regularity condition at the origin. The SP and EA modes with high wave numbers have spatial structures away from the origin, consequently the critical parameters and eigenfunctions are well approximated by using Gauss-Lobatto meshes without the origin. However, for very low σ, the instability affects the center of the fluid sphere. Then, including the origin in the mesh together with regularity conditions is essential to avoid the spurious modes described above.

The results represented in Figs. 6 and 8 required the computation of thousands of spectra. Without a robust method to obtain them, not contaminated by spurious solutions, it would have been impossible to explore all the wave numbers in the plots, which range from $m = 5$ to $m = 65$. The different spatial structures of the critical eigenmodes for high Taylor and low Prandtl numbers have been shown, and the dependence on Ta of the critical parameters previously found by approximate asymptotic theories has been confirmed.

6. Acknowledgments

This research has been supported by MEC-DGICT/FEDER project FIS2013-40674-P and AGAUR-GENCAT project 2014-SGR-1145.

Appendix

An example of the construction of derivative matrices implementing the boundary conditions is presented here. It corresponds to $f(1) = 0$ and $f''(1) = 0$, without any condition at $r = 0$, i.e., in the case of a mesh which does not include the origin, or with condition $f(0) = 0$ if it is included.

Assume first that $r_0 = 0$ is not included in the mesh and that $f(r)$ is approximated by the Lagrange interpolation polynomial

$$f(r) \approx \sum_{k=1}^{n_r} L_k(r) f(r_k), \tag{31}$$

with $L_k(r)$ the Lagrange polynomial (22) of degree $n_r - 1$ satisfying $L_k(r_j) = \delta_{kj}$ for $i = 1, \ldots, n_r$. If $d_{j,k}^{(l)} = L_k^{(l)}(r_j)$ for $j = 1, \ldots, n_r$, the following derivative approximation formula is obtained

$$f^{(l)}(r_j) \approx \sum_{k=1}^{n_r} d_{j,k}^{(l)} f(r_k), \quad j = 1, \ldots, n_r, \tag{32}$$

at the nodes of the mesh. From now on we will write $f_i = f(r_i)$ for any of the functions which will appear.

In order to implement $f(1) = 0$ and $f''(1) = 0$ let us write $f(r) = (1-r)g(r)$. The condition $f(1) = 0$ is then satisfied. By taking derivatives it follows that

$$f^{(l)}(r) = (1 - r)g^{(l)}(r) - l g^{(l-1)}(r). \tag{33}$$

The condition $f''(1) = 0$ implies $g'(1) = 0$. This condition is approximated by using Eq. (32) for $l = 1$, and then g_{n_r} is written as a linear combination of the values of g at the inner nodes

$$g_{n_r} = -\sum_{j=1}^{n_r-1} d^{(1)}_{n_r,j}/d^{(1)}_{n_r,n_r} g_j. \tag{34}$$

By substituting the derivatives in Eq. (33) by their approximations, and using (34) and that $g_i = f_i/(1 - r_i)$ one finally obtains

$$f^{(l)}(r_j) \approx \sum_{k=1}^{n_r-1} \tilde{d}^{(l)}_{j,k} f(r_k), \quad j = 1, \ldots, n_r - 1, \tag{35}$$

with

$$\tilde{d}^{(l)}_{j,k} = \left[(1 - r_j)d^{(l)}_{j,k} - l d^{(l-1)}_{j,k} - (d^{(1)}_{n_r,k}/d_{n_r,n_r})((1 - r_j)d^{(l)}_{j,k} - l d^{(l-1)}_{j,k})\right]/(1 - r_k),$$

where only the inner points of the mesh are involved in the calculations.

If the origin is included in the mesh, with $f(0) = 0$, then the index k in Eqs. (31) and (32) starts in principle at $k = 0$, and the Lagrange polynomials are different and have degree n_r. Since $f_0 = f(0) = 0$ the coefficients $d^{(l)}_{j,0}$, $j = 1, \ldots, n_r$ will never appear and the index in Eq. (32) can be started again at $k = 1$. Since $f_0 = 0$ implies $g_0 = 0$, the rest remains the same. Therefore the only difference between the two cases is the value of the coefficients $d^{(l)}_{j,k}$ which are based on different meshes. The rest remains unchanged.

The trick of writing $f(r) = (1 - r)g(r)$ has to be used when there are two boundary conditions at $r = 1$. If there would be just one, the value of f_{n_r} can be obtained from those at the inner points directly from the approximate boundary condition.

References

[1] D. Gottlieb, S. Orszag, Numerical Analysis of Spectral Methods: Theory and Applications, SIAM, 1977.

[2] C. Canuto, M. Y. Hussaini, A. Quarteroni, T. A. Zang, Spectral Methods in Fluid Dynamics, Springer, 1988.

[3] J. P. Boyd, Chebyshev and Fourier Spectral Methods, Springer, New York, 1989.

[4] C. Bernardi, Y. Maday, Approximations spectrales de problemes aux limites elliptiques, Vol. 10 of Mathématiques et Applications, Springer, Paris, 1992.

[5] D. Funaro, Polynomial Approximation of Differential Equations, Springer-Verlag, New York, 1992.

[6] B. Fornberg, A Practical Guide to Pseudospectral Methods, Cambridge University Press, New York, 1996.

[7] R. Peyret, Spectral methods for incompressible viscous flow, Vol. 148 of Applied Mathematical Sciences, Springer, 2002.

[8] M. O. Deville, P. F. Fischer, M. E. H., High-order methods for incompressible fluid flow, Cambridge University Press, 2002.

[9] G. M. Karniadakis, S. J. Sherwin, Spectral/hp element methods for CFD, Numerical mathematics and scientific computations, Oxford, 1999.

[10] J. Hesthaven, S. Gottlieb, D. Gottlieb, Spectral Methods for Time-Dependent Problems, Vol. 21 of Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, Cambridge, 2007.

[11] C. Canuto, M. Y. Hussaini, A. Quarteroni, T. A. Zang, Spectral Methods. Fundamentals in single domains., Springer, 2006.

[12] C. Canuto, M. Y. Hussaini, A. Quarteroni, T. A. Zang, Spectral Methods. Evolution to complex geometries and applications to Fluid Dynamics, Springer, 2007.

[13] J. P. Boyd, F. Yu, Comparing seven spectral methods for interpolation and for solving the Poisson equation in a disk: Zernike polynomials, Logan–Shepp ridge polynomials, Chebyshev–Fourier series, cylindrical Robert functions, Bessel–Fourier expansions, square-to-disk conformal mapping and radial basis functions, J. Comput. Phys. 230 (4) (2011) 1408 – 1438.

[14] W. Huang, D. M. Sloan, Pole condition for singular problems: The pseudospectral approximation, J. Comput. Phys. 107 (2) (1993) 254–261.

[15] W. Huang, H. Ma, W. Sun, Convergence analysis of spectral collocation methods for a singular differential equation, SIAM J. Numer. Anal. 41 (6) (2003) 2333–2349.

[16] D. J. Torres, E. A. Coutsias, Pseudospectral solution of the two-dimensional navier–stokes equations in a disk, SIAM J. Sci. Comput. 21 (1) (1999) 378–403.

[17] B. Fornberg, A pseudospectral approach for polar ans spherical geometries, SIAM J. Sci. Comput. 16 (5) (1995) 1071–1081.

[18] K. Mohseni, T. Colonius, Numerical treatment of polar coordinate singularities, J. Comput. Phys. 157 (2) (2000) 787 – 795.

[19] H. Chen, Y. Su, B. D. Shizgal, A direct spectral collocation Poisson solver in polar and cylindrical coordinates, J. Comput. Phys. 160 (2) (2000) 453–469.

[20] W. Heinrichs, Spectral collocation schemes on the unit disc, J. Comput. Phys. 199 (1) (2004) 66 – 86.

[21] P. E. Merilees, An alternative for the summation of a series of spherical harmonics, J. Appl. Meteor. 12 (1973) 224–227.

[22] S. A. Orszag, Fourier series on spheres, Month. Weather. Rev. 102 (1974) 56–75.

[23] J. P. Boyd, The choice of spectral functions on a sphere for boundary and eigenvalue problems: A comparison of Chebyshev, Fourier and associated Legendre expansions, Month. Weather. Rev. 106 (1978) 1184–1191.

[24] J. P. Boyd, Chebyshev and Fourier Spectral Methods, Dover, New York, 1999.

[25] P. W. Livermore, C. A. Jones, S. J. Worland, Spectral radial basis functions for full sphere computations, J. Comput. Phys. 227 (2) (2007) 1209 – 1224.

[26] P. W. Livermore, Galerkin orthogonal polynomials, J. Comput. Phys. 229 (6) (2010) 2046 – 2060.

[27] K. Li, P. W. Livermore, A. Jackson, An optimal Galerkin scheme to solve the kinematic dynamo eigenvalue problem in a full sphere, J. Comput. Phys. 229 (23) (2010) 8666 – 8683.

[28] C. W. Clenshaw, The numerical solution of linear differential equations in Chebyshev series, Math. Proc. Camb. Phil. Soc. 53 (1957) 134–149.

[29] E. A. Coutsias, T. Hagstrom, D. Torres, An efficient spectral method for ordinary differential equations with rational function coefficients, Math. Comp. 65 (1996) 611–635.

[30] S. Olver, A. Townsend, A fast and well-conditioned spectral method, SIAM Rev. 55 (3) (2013) 462–489.

[31] A. Townsend, S. Olver, The automatic solution of partial differential equations using a global spectral method, J. Comput. Phys. 299 (15) (2015) 106–123.

[32] M. Berioz, T. Hagstrom, S. R. Lau, R. H. Price, Multidomain, sparse, spectral-tau method for helically symmetric flow, Computers and Fluids 102 (2014) 250–265.

[33] F. Garcia, J. Sánchez, M. Net, Antisymmetric polar modes of thermal convection in rotating spherical fluid shells at high Taylor numbers, Phys. Rev. Lett. 101 (19) (2008) 194501.

[34] M. Net, F. Garcia, J. Sánchez, On the onset of low-Prandtl-number convection in rotating spherical shells: non-slip boundary conditions, J. Fluid Mech. 601 (2008) 317–337.

[35] F. Garcia, M. Net, B. García-Archilla, J. Sánchez, A comparison of high-order time integrators for the Boussinesq Navier-Stokes equations in rotating spherical shells, J. Comput. Phys. 229 (2010) 7997–8010.

[36] M. Net, F. Garcia, J. Sánchez, Numerical study of the onset of thermosolutal convection in rotating spherical shells, Phys. Fluids 24 (2012) 064101–1–064101–21.

[37] F. Garcia, J. Sánchez, M. Net, Numerical simulations of high-Rayleigh-number convection in rotating spherical shells under laboratory conditions, Phys. Earth Planet. Inter. 230 (2014) 28–44.

[38] F. Garcia, L. Bonaventura, M. Net, J. Sánchez, Exponential versus IMEX high-order time integrators for thermal convection in rotating spherical shells, J. Comput. Phys. 264 (2014) 41–54.

[39] J. Sánchez, F. Garcia, M. Net, Computation of azimuthal waves and their stability in thermal convection in rotating spherical shells with application to the study of a double-Hopf bifurcation, Phys. Rev. E 87 (2013) 033014.

[40] F. Garcia, M. Net, J. Sánchez, Continuation and stability of convective modulated rotating waves in spherical shells, Accepted in Phys. Rev. E.

[41] B. Fornberg, Generation of finite difference formulas on arbitrary spaced grids, Math. Comput. 51 (1988) 699–706.

[42] D. Funaro, W. Heinrichs, Some results about the pseudospectral approximation of one-dimensional fourth-order problems, Numer. Math. 58 (1) (1990) 399–418.

[43] T. J. Rivlin, An introduction to the approximation of functions, Dover, 1969.

[44] G. B. Arfken, H. J. Weber, Mathematical methods for physicists, fifth edition, Academic Press, 2001.

[45] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, D. Sorensen, LAPACK Users' Guide, Third Edition, SIAM, 1999.

[46] G. H. Golub, C. F. Van Loan, Matrix Computations (3rd Ed.), Johns Hopkins University Press, Baltimore, MD, USA, 1996.

[47] T. Cai, K. L. Chan, L. Deng, Numerical simulation of core convection by a multi-layer semi-implicit spherical spectral method, Journal of Computational Physics 230 (24) (2011) 8698 – 8712.

[48] R. B. Lehoucq, D. C. Sorensen, C. Yang, ARPACK User's Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods, Software, Environments, Tools, SIAM, 1998.

[49] K. Meerbergen, D. Roose, Matrix transformations for computing rightmost eigenvalues of large sparse non-symmetric eigenvalue problems, IMA J. Numer. Anal. 16 (3) (1996) 297–346.

[50] R. P. Brent, Algorithms for Minimization without Derivatives, Prentice-Hall, New Jersey, 1973.

[51] Y. Saad, Numerical Methods for Large Eigenvalue Problems, Manchester University Press, Manchester, 1992.

[52] A. Tilgner, Spectral methods for the simulation of incompressible flow in spherical shells, Int. J. Numer. Meth. Fluids 30 (1999) 713–724.

[53] P. H. Roberts, On the thermal instability of a rotating fluid sphere containing heat sources, Phil. Trans. R. Soc. Lond. A 263 (1136) (1968) 93–117.

[54] F. H. Busse, Thermal instabilities in rapidly rotating systems, J. Fluid Mech. 44 (1970) 441–460.

[55] K. Zhang, Spiralling columnar convection in rapidly rotating spherical fluid shells, J. Fluid Mech. 236 (1992) 535–556.

[56] C. A. Jones, A. M. Soward, A. I. Mussa, The onset of thermal convection in a rapidly rotating sphere, J. Fluid Mech. 405 (2000) 157–179.

[57] E. Dormy, A. M. Soward, C. A. Jones, D. Jault, P. Cardin, The onset of thermal convection in rotating spherical shells, J. Fluid Mech. 501 (2004) 43–70.

[58] K. Zhang, X. Liao, A new asymptotic method for the analysis of convection in a rapidly rotating sphere, J. Fluid Mech. 518 (2004) 319–346.

[59] C. R. Carrigan, F. H. Busse, An experimental and theoretical investigation of the onset of convection in rotating spherical shells, J. Fluid Mech. 126 (1983) 287–305.

[60] J. Sánchez, F. Garcia, M. Net, Critical torsional modes of convection in rotating fluid spheres at high Taylor numbers, Submitted to J. Fluid Mech.