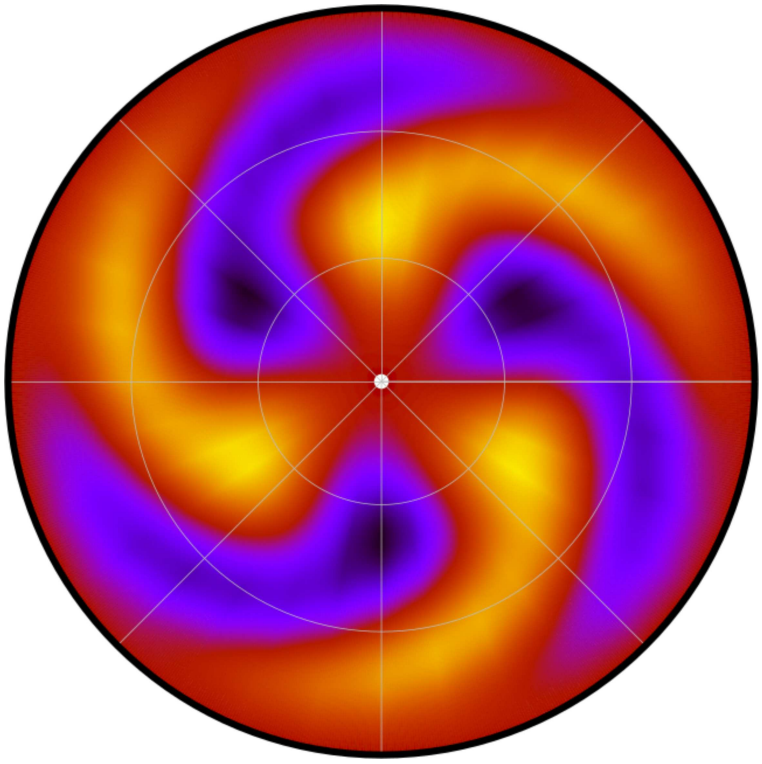
CyFlow 1

Spectral solver
for incompressible flows in cylinders

January 30, 2025



Developer: Cédric Beaume (University of Leeds)
User: Ellen Bartle (PhD student, University of Leeds, 2024–now)

Contents

1 History

This project started in 2021 motivated by several questions that involved cylindrical geometries. I initially tried mainstream spectral methods but was not satisfied with the way they handled the singularity at the cylinder axis. I later discovered the disk discretization of Matsushima and Marcus [6], which handles the intricacies of the geometry elegantly. Unfortunately, its implementation is not the easiest but, after all, YODO[1]. The first Helmholtz solvers were validated in 2022. As I reflected about the method to solve the Navier–Stokes equation together with the incompressibility constraint, I remembered my many discussions with Laurette Tuckerman. Her passion for divergence-free methods swayed me and I opted for a potential approach, inspired by the impressive work of Piotr Boroński. CyFlow 1 was completed in 2023.

[1] You Only Discretize Once

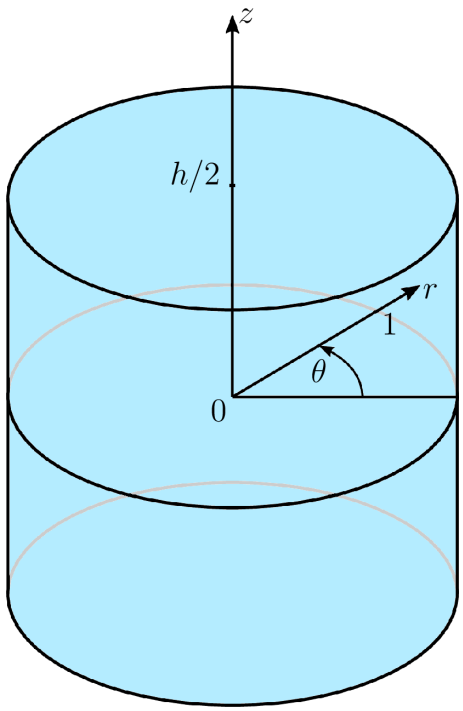


Figure 1: Sketch of the cylindrical domain of interest. The domain is centered at the origin. The cylinder side-walls are located at $r = 1$ and its end-walls are located at $z = \pm h/2$. The azimuthal direction is denoted θ.

2 Incompressible flow equations

We consider an incompressible flow inside a cylinder of radius 1 and height h, as sketched in Figure 1. The cylinder is centered at the origin of the coordinate system, r denotes the radial coordinate, θ is the azimuthal coordinate and z denotes the axial one.

The fluid flow is governed by the Navier–Stokes equation and the incompressibility condition:

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\,\mathbf{u} = -\nabla p + \frac{1}{Re}\nabla^2\mathbf{u}, \tag{1}$$

$$\nabla\cdot\mathbf{u} = \mathbf{0}, \tag{2}$$

where $\mathbf{u} = u_r\hat{\mathbf{r}} + u_\theta\hat{\theta} + u_z\hat{\mathbf{z}}$ is the velocity field with its projections onto the unit vectors in the radial direction, $\hat{\mathbf{r}}$, the azimuthal direction, $\hat{\theta}$, and the axial direction, $\hat{\mathbf{z}}$, p is the pressure, t is the time and Re is the Reynolds number.

Ensuring the incompressibility of the flow numerically is not a straightforward task due to the fact that the incompressibility constraint (2) is not an evolution equation and due to the absence of such an equation for the pressure. However, it can be done exactly by reformulating the equation using poloidal and toroidal streamfunctions. The derivation of this formulation is not trivial and relies heavily on vector calculus. I have provided necessary identities in Appendix A & B. If something appears non-trivial, it is likely that these will help.

2.1 Poloidal-toroidal formulation

We introduce the poloidal streamfunction, ϕ, and the toroidal streamfunction, ψ, such that:

$$\mathbf{u} = \nabla \times (\psi \hat{\mathbf{z}}) + \nabla \times \nabla \times (\phi \hat{\mathbf{z}}). \tag{3}$$

This decomposition automatically ensures that the velocity is divergence-free.

Under this decomposition, each velocity component derives from the streamfunction as follows:

$$\mathbf{u} = \begin{pmatrix} u_r \\ u_\theta \\ u_z \end{pmatrix} = \begin{pmatrix} \frac{1}{r}\partial_\theta \psi + \partial_{rz}\phi \\ \frac{1}{r}\partial_{\theta z}\phi - \partial_r \psi \\ -\nabla_h^2 \phi \end{pmatrix}, \tag{4}$$

where $\nabla_h^2 = r^{-1}\partial_r(r\partial_r) + r^{-2}\partial_\theta^2$ is the Laplacian restricted to the disk components. Note that the subscript h is used to match literature notation, where its stands for *horizontal*, the axis of the cylinder often being represented in the vertical direction. The vorticity follows directly:

$$\omega = \nabla \times \mathbf{u} = \begin{pmatrix} \omega_r \\ \omega_\theta \\ \omega_z \end{pmatrix} = \begin{pmatrix} \partial_{rz}\psi - \frac{1}{r}\partial_\theta \nabla^2 \phi \\ \partial_r \nabla^2 \phi + \frac{1}{r}\partial_{\theta z}\psi \\ -\nabla_h^2 \psi \end{pmatrix}, \tag{5}$$

where $\nabla^2 = \nabla_h^2 + \partial_z^2$ is the full Laplacian. Lastly, the double curl of the velocity will also be useful:

$$\nabla \times \nabla \times \mathbf{u} = \begin{pmatrix} -\partial_{rz}\nabla^2 \phi - \frac{1}{r}\partial_\theta \nabla^2 \psi \\ \partial_r \nabla^2 \psi - \frac{1}{r}\partial_{\theta z}\nabla^2 \phi \\ \nabla^2 \nabla_h^2 \phi \end{pmatrix}. \tag{6}$$

The Navier–Stokes equation is transformed by taking the projection on $\hat{\mathbf{z}}$ of its curl and its double curl:

$$\left(\partial_t - \frac{1}{Re}\nabla^2\right)\nabla_h^2 \psi = \hat{\mathbf{z}} \cdot \nabla \times [(\mathbf{u}\cdot\nabla)\mathbf{u}], \tag{7}$$

$$\left(\partial_t - \frac{1}{Re}\nabla^2\right)\nabla^2 \nabla_h^2 \phi = -\hat{\mathbf{z}} \cdot \nabla \times \nabla \times [(\mathbf{u}\cdot\nabla)\mathbf{u}]. \tag{8}$$

The resulting equations are periodic in the azimuthal direction, so they will not require any boundary conditions in θ. However, they require 5 boundary conditions in the radial direction (noting that $r = 1$ covers "both sides" of the origin) as well as 3 boundary conditions per end-wall in z.

2.2 Gauge condition

By introducing the poloidal-toroidal decomposition (3), we have introduced a gauge freedom for both ϕ and ψ. To understand the structure of the gauge, we need to solve:

$$\begin{aligned}
\mathbf{u}^g = \mathbf{0} \quad &\Rightarrow \quad \nabla \times (\psi^g \hat{\mathbf{z}}) + \nabla \times \nabla \times (\phi^g \hat{\mathbf{z}}) = \mathbf{0} && (9)\\
&\Rightarrow \quad \nabla_h \times (\psi^g \hat{\mathbf{z}}) + \nabla \left[\nabla \cdot (\phi^g \hat{\mathbf{z}})\right] - \nabla^2 (\phi^g \hat{\mathbf{z}}) = \mathbf{0} && (10)\\
&\Rightarrow \quad \nabla_h \times (\psi^g \hat{\mathbf{z}}) + \nabla \left[\partial_z \phi^g\right] - (\nabla^2 \phi^g) \hat{\mathbf{z}} = \mathbf{0} && (11)\\
&\Rightarrow \quad \nabla_h \times (\psi^g \hat{\mathbf{z}}) + \nabla_h \left[\partial_z \phi^g\right] - (\nabla_h^2 \phi^g) \hat{\mathbf{z}} = \mathbf{0} && (12)\\
&\Rightarrow \quad \begin{cases} \nabla_h \times (\psi^g \hat{\mathbf{z}}) = -\nabla_h \left(\partial_z \phi^g\right) \\ \nabla_h^2 \phi^g = 0 \end{cases}, && (13)
\end{aligned}$$

where $\nabla_h = \hat{\mathbf{r}} \partial_r + r^{-1} \hat{\theta} \partial_\theta$ is the gradient restricted to the disk components.

The poloidal gauge ϕ^g exists and its uniqueness can be imposed by providing a suitable boundary condition on the disk. To guarantee the existence of a toroidal gauge, we note that:

$$\begin{aligned}
\nabla_h^2 \phi^g = 0 \quad &\Rightarrow \quad \nabla_h \cdot \nabla_h \phi^g = 0 && (14)\\
&\Rightarrow \quad \nabla_h \cdot \left[\nabla_h \left(\partial_z \phi^g\right)\right] = 0. && (15)
\end{aligned}$$

Observing that this calculation yields a divergence-free field in the (r, θ)-plane, we use the identity: $\nabla_h \cdot [\nabla_h \times (f \hat{\mathbf{z}})] = 0$ for any suitably regular function f, to deduce:

$$\begin{aligned}
\nabla_h \times (f \hat{\mathbf{z}}) &= \nabla_h \left(\partial_z \phi^g\right) && (16)\\
&= -\nabla_h \times (\psi^g \hat{\mathbf{z}}), && (17)
\end{aligned}$$

obtained by using equation (13) and thereby proving the existence of a toroidal gauge ψ^g.

The gauge problem (13) can equivalently be presented in the following form [5]:

$$\begin{cases} \nabla_h^2 \phi^g = 0 \\ \nabla_h \psi^g = \nabla_h \left(\partial_z \phi^g\right) \times \hat{\mathbf{z}} \end{cases}, \tag{18}$$

and implies that the poloidal gauge ϕ^g is defined up to a harmonic function in (r, θ) and arbitrary in z. Its uniqueness can be imposed by simply setting the value of the gauge at the cylinder side-wall. The toroidal gauge ψ^g is determined by its poloidal counterpart up to a constant that may vary with z. A simple way to impose the uniqueness of the toroidal gauge is to set its value on the centerline of the cylinder. A simple choice of gauge condition is [2]:

$$\phi^g = 0 \text{ at } r = 1, \tag{19}$$

$$\psi^g = 0 \text{ at } r = 0. \tag{20}$$

2.3 Radial boundary condition

The side-wall of the cylinder is impenetrable, so we impose $u_r(r = 1) = u_r^w = 0$, where the superscript w indicates a value at the wall. This boundary condition becomes:

$$\frac{1}{r} \partial_\theta \psi + \partial_{rz} \phi = 0 \text{ at } r = 1. \tag{21}$$

We also impose no-slip boundary condition on a static side-wall:

$$\left.\begin{array}{l} u_\theta(r=1) = u_\theta^w \\ u_z(r=1) = u_z^w \end{array}\right\} \Rightarrow \left\{\begin{array}{l} \dfrac{1}{r}\partial_{\theta z}\phi - \partial_r\psi = u_\theta^w \\ -\nabla_h^2\phi = u_z^w \end{array}\right. \quad \text{at } r=1. \tag{22}$$

The poloidal gauge condition (19) is set at the same boundary and implies that $\partial_{\theta z}\phi(r=1) = 0$. Consequently, the radial no-slip boundary condition reads:

$$\left\{\begin{array}{l} \partial_r\psi = -u_\theta^w \\ \nabla_h^2\phi = -u_z^w \end{array}\right. \quad \text{at } r=1. \tag{23}$$

2.4 Compatibility condition

We need to ensure that the poloidal-toroidal problem (7) & (8) is equivalent to the primitive problem we wish to solve (equations (1) & (2)). Let us define:

$$\mathbf{f} = \partial_t\mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} - \frac{1}{Re}\nabla^2\mathbf{u} \tag{24}$$

$$= -\nabla p, \tag{25}$$

in such a way that the primitive problem can be written as $\mathbf{g} = \nabla\times\mathbf{f} = \mathbf{0}$ [5]. The following implication is natural:

$$\mathbf{g} = \mathbf{0} \Rightarrow \left\{\begin{array}{l} \mathbf{g}\cdot\hat{\mathbf{z}} = 0 \\ (\nabla\times\mathbf{g})\cdot\hat{\mathbf{z}} = 0 \end{array}\right. , \tag{26}$$

where the right-hand-side is, in fact, the poloidal-toroidal problem (7) & (8). For the equivalence to stand, we need to prove the leftward implication. Unfortunately, conditions are missing to establish it as the right-hand-side problem is of higher order [1]:

$$\left.\begin{array}{l} \mathbf{g}\cdot\hat{\mathbf{z}} = 0 \\ (\nabla\times\mathbf{g})\cdot\hat{\mathbf{z}} = 0 \end{array}\right\} \Rightarrow \mathbf{g} = \nabla_h\kappa, \tag{27}$$

where κ does not automatically yield $\mathbf{g} = \mathbf{0}$. For this equality to stand, we need $\kappa = \kappa(z)$. We can easily get κ of this form by making it a solution of a Laplace equation on the disk with homogeneous Neumann boundary conditions at its edge:

$$\left.\begin{array}{l} \mathbf{g} = \nabla_h\kappa \\ \nabla\cdot\mathbf{g} = 0 \\ \mathbf{g}\cdot\hat{\mathbf{r}} = 0 \text{ at } r=1 \end{array}\right\} \Rightarrow \left.\begin{array}{l} \nabla_h^2\kappa = 0 \\ \partial_r\kappa = 0 \text{ at } r=1 \end{array}\right\} \Rightarrow \kappa = \kappa(z) \Rightarrow \mathbf{g} = \mathbf{0}. \tag{28}$$

Consequently, the exact equivalence between the primitive problem and the potential problem is as follows:

$$\mathbf{g} = \mathbf{0} \iff \left\{\begin{array}{l} \mathbf{g}\cdot\hat{\mathbf{z}} = 0 \\ (\nabla\times\mathbf{g})\cdot\hat{\mathbf{z}} = 0 \\ \nabla\cdot\mathbf{g} = 0 \\ \mathbf{g}\cdot\hat{\mathbf{r}} = 0 \text{ at } r=1 \end{array}\right. . \tag{29}$$

The first additional condition, $\nabla \cdot \mathbf{g} = 0$, is automatically satisfied in the case of incompressible flows. Furthermore, the additional boundary condition is trivial for axisymmetric modes since:

$$\begin{aligned} \nabla_h^2 \kappa = 0 \quad &\Rightarrow \quad \frac{1}{r}\partial_r(r\partial_r\kappa) = 0 \text{ for } \kappa \text{ axisymmetric} && (30)\\ &\Rightarrow \quad \kappa = k_1(z)\ln(r) + k_2(z) && (31)\\ &\Rightarrow \quad \kappa = k_2(z), && (32) \end{aligned}$$

without the need for any additional condition.

As a result of the above equivalence and remarks, we need to impose a compatibility condition between the primitive and potential formulations, which takes the form of a radial boundary condition for non-axisymmetric modes:

$$\nabla \times \left[\partial_t \mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} - \frac{1}{Re}\nabla^2\mathbf{u}\right] \cdot \hat{\mathbf{r}} = 0 \text{ at } r = 1. \tag{33}$$

The first term of equation (33) yields:

$$\begin{aligned} \nabla \times \partial_t \mathbf{u} \cdot \hat{\mathbf{r}} &= \partial_t \left[(\nabla \times \mathbf{u}) \cdot \hat{\mathbf{r}}\right] && (34)\\ &= \partial_t \omega_r && (35)\\ &= 0 \text{ at } r = 1, && (36) \end{aligned}$$

at a non-deformable side-wall ($\omega_r = 0$ at $r = 1$). The second term of equation (33) requires more care:

$$\begin{aligned} \nabla \times \left[(\mathbf{u}\cdot\nabla)\,\mathbf{u}\right] \cdot \hat{\mathbf{r}} &= -\nabla \times (\mathbf{u} \times \omega) \cdot \hat{\mathbf{r}} && (37)\\ &= \partial_z \left(u_z\omega_r - u_r\omega_z\right) - \frac{1}{r}\partial_\theta \left(u_r\omega_\theta - u_\theta\omega_r\right) && (38)\\ &= \partial_z(u_z\omega_r) + \frac{1}{r}\partial_\theta(u_\theta\omega_r) \quad \text{as } u_r^w = 0 && (39)\\ &= 0 \text{ at } r = 1, && (40) \end{aligned}$$

for the same reason as before. Equation (33) thus reduces to:

$$\begin{aligned} \nabla \times \left(\nabla^2\mathbf{u}\right) \cdot \hat{\mathbf{r}} = 0 \quad &\Rightarrow \quad \nabla \times \left[\nabla \times (\nabla \times \mathbf{u})\right] \cdot \hat{\mathbf{r}} = 0 && (41)\\ &\Rightarrow \quad \nabla \times \begin{pmatrix} -\partial_{rz}(\nabla^2\phi) - \frac{1}{r}\partial_\theta(\nabla^2\psi) \\ \partial_r(\nabla^2\psi) - \frac{1}{r}\partial_{\theta z}(\nabla^2\phi) \\ \nabla^2\nabla_h^2\phi \end{pmatrix} \cdot \hat{\mathbf{r}} = 0 && (42)\\ &\Rightarrow \quad \frac{1}{r}\partial_\theta\nabla^2\nabla_h^2\phi - \partial_z\left[\partial_r(\nabla^2\psi) - \frac{1}{r}\partial_{\theta z}(\nabla^2\phi)\right] = 0 && (43)\\ &\Rightarrow \quad \frac{1}{r}\partial_\theta\nabla^2\nabla_h^2\phi - \partial_{rz}(\nabla^2\psi) + \frac{1}{r}\partial_{\theta zz}(\nabla^2\phi) = 0 \text{ at } r = 1. && (44) \end{aligned}$$

This expression can be further simplified by noticing that the velocity definition and the poloidal gauge imply:

$$\left.\begin{aligned} u_z^w = -\nabla_h^2\phi \text{ at } r = 1 \\ \phi(r = 1) = 0 \end{aligned}\right\} \quad \Rightarrow \quad \nabla^2\phi = -u_z^w \text{ at } r = 1, \tag{45}$$

$$\left.\begin{aligned} u_\theta^w = \frac{1}{r}\partial_{\theta z}\phi - \partial_r\psi \text{ at } r = 1 \\ \phi(r = 1) = 0 \end{aligned}\right\} \quad \Rightarrow \quad \partial_r\psi = -u_\theta^w \text{ at } r = 1, \tag{46}$$

in such a way that the last term and the axial contribution of the Laplacian of the second term of equation (44) vanish for a non-deformable side-wall ($\partial_\theta u_z^w = 0$ and $\partial_z u_\theta^w = 0$). Eventually, the compatibility condition reads:

$$\frac{1}{r}\partial_\theta \nabla^2 \nabla_h^2 \phi - \partial_{rz} \nabla_h^2 \psi = 0 \text{ at } r = 1 \text{ for non-axisymmetric modes.} \tag{47}$$

Note here that the derivation of this compatibility condition relies solely on three conditions: the impenetrability condition $u_r = 0$ at $r = 1$ as well as the non-deformability of the side-wall $\partial_\theta u_z^w = \partial_z u_\theta^w = 0$. If these conditions are not met, additional terms would enter the compatibility condition.

2.5 Axial boundary condition

The impenetrability and no-slip boundary conditions in the axial direction take the form $\mathbf{u} = \mathbf{u}^\pm$ at $z = \pm h/2$. Imposing this boundary condition by simply expressing the velocity components in terms of the potentials is not straightforward as the resulting expressions couple the potentials together. Instead, we can make use of equivalence (29) for $\mathbf{g} = \mathbf{u} - \mathbf{u}^\pm$ and on the top and bottom disks, rather than on the cylinder. We can thus impose on the surfaces $z = \pm h/2$:

$$\left.\begin{array}{l} u_z = u_z^\pm \\ (\nabla_h \times \mathbf{u}) \cdot \hat{\mathbf{z}} = (\nabla_h \times \mathbf{u}^\pm) \cdot \hat{\mathbf{z}} \\ \nabla_h \cdot \mathbf{u} = \nabla_h \cdot \mathbf{u}^\pm \\ u_r = u_r^\pm \text{ at } r = 1 \end{array}\right\} \Rightarrow \left\{\begin{array}{l} u_z = 0 \text{ (impenetrability condition)} \\ (\nabla_h \times \mathbf{u}) \cdot \hat{\mathbf{z}} = (\nabla_h \times \mathbf{u}^\pm) \cdot \hat{\mathbf{z}} \\ -\partial_z u_z = \nabla_h \cdot \mathbf{u}^\pm \\ u_r = 0 \text{ at } r = 1 \quad \text{as } u_r^w = 0 \end{array}\right. \tag{48}$$

$$\Rightarrow \left\{\begin{array}{l} \nabla_h^2 \phi = 0 \\ -\nabla_h^2 \psi = \dfrac{1}{r}\left[\partial_r (r u_\theta^\pm) - \partial_\theta u_r^\pm\right] \\ -\partial_z u_z = \dfrac{1}{r}\left[\partial_r (r u_r^\pm) + \partial_\theta u_\theta^\pm\right] \end{array}\right. , \tag{49}$$

where we dropped the fourth condition as it is already imposed by the radial no-slip boundary condition. The non-deformability of the top and bottom walls require $u_r^\pm = \partial_\theta u_\theta^\pm = 0$ and the axial boundary condition reduces to:

$$\left\{\begin{array}{l} \nabla_h^2 \phi = 0 \\ \nabla_h^2 \psi = -\dfrac{1}{r}\partial_r (r u_\theta^\pm) \quad \text{on } z = \pm h/2. \\ \partial_z \nabla_h^2 \phi = 0 \end{array}\right. \tag{50}$$

2.6 Summary

The poloidal-toroidal formulation of the Navier–Stokes equation for an incompressible flow is:

$$\left(\partial_t - \frac{1}{Re}\nabla^2\right) \nabla_h^2 \psi = \hat{\mathbf{z}} \cdot \nabla \times \left[(\mathbf{u} \cdot \nabla)\mathbf{u}\right], \tag{51}$$

$$\left(\partial_t - \frac{1}{Re}\nabla^2\right) \nabla^2 \nabla_h^2 \phi = -\hat{\mathbf{z}} \cdot \nabla \times \nabla \times \left[(\mathbf{u} \cdot \nabla)\mathbf{u}\right]. \tag{52}$$

These equations are to be solved alongisde the following radial boundary conditions:

$$\begin{cases} \dfrac{1}{r}\partial_\theta\psi + \partial_{rz}\phi = 0 \\ \partial_r\psi = -u_\theta^w \\ \nabla_h^2\phi = -u_z^w \\ \phi = 0 \end{cases} \quad \text{at } r = 1, \tag{53}$$

the following additional conditions:

$$\begin{cases} \psi(r=0) = 0 \text{ for axisymmetric modes} \\ \dfrac{1}{r}\partial_\theta\nabla^2\nabla_h^2\phi - \partial_{rz}\nabla_h^2\psi = 0 \text{ for non axisymmetric modes at } r = 1 \end{cases} \tag{54}$$

and the following axial boundary conditions:

$$\begin{cases} \nabla_h^2\phi = 0 \\ \nabla_h^2\psi = -\dfrac{1}{r}\partial_r(ru_\theta^\pm) \\ \partial_z\nabla_h^2\phi = 0 \end{cases} \quad \text{on } z = \pm h/2. \tag{55}$$

This derivation relies on the fact that the side-wall is impenetrable ($u_r^w = 0$) and non-deformable ($\partial_\theta u_z^w = \partial_z u_\theta^w = 0$) and that the end-walls are impenetrable ($u_z^\pm = 0$) and non-deformable ($u_r^\pm = \partial_\theta u_\theta^\pm = 0$).

3 Temporal discretization

Fill out this section

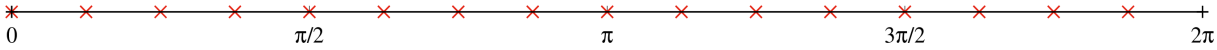


Figure 2: Equidistributed mesh for one-dimensional, 2π-periodic functions. The location of the $N = 16$ meshpoints is shown in red.

4 Spatial discretization

4.1 Fourier–Galerkin projection for the azimuthal direction

4.1.1 Principles

As we seek periodic solutions in the azimuthal direction, it is natural to turn to a Fourier–Galerkin discretization based on the equidistributed mesh:

$$\theta_j = \frac{2\pi j}{N}, \qquad j = 0, \dots, N-1, \tag{56}$$

where N is the number of points and where the last point ($j = N$) is omitted as it is unnecessary: $f(x_N) = f(x_0)$ due to the periodicity of the function f. An example of such a mesh is shown in figure 2. Note that to produce good quality plots of any function represented using this type of mesh, one needs to include the value of the function at x_N by copying its value at x_0.

We can express $f(\theta)$ as a discrete Fourier series:

$$f(\theta) = \sum_{n=-\infty}^{\infty} \hat{f}_n \exp(\imath n\theta), \tag{57}$$

where the $\hat{f}_n$ are the complex coefficients associated with the Fourier expansion and $\imath$ is the imaginary unit number. The above choice of basis implicitly takes care of the periodic boundary condition so that expansion (57) is, in fact, a Galerkin projection. By using the finite mesh (56), we can write:

$$f(\theta_j) = \sum_{n=-N/2+1}^{N/2} \left[\hat{f}_n \exp\left(\imath n j \frac{2\pi}{N} \right) \right] \tag{58}$$

$$\Rightarrow f(\theta_j) = \hat{f}_0 + \sum_{n=1}^{N/2-1} \left[\hat{f}_n \exp\left(\imath n j \frac{2\pi}{N} \right) + \hat{f}_n^* \exp\left(-\imath n j \frac{2\pi}{N} \right) \right] + \hat{f}_{N/2}(-1)^j, \tag{59}$$

where $j = 0, \dots, N-1$ and where the asterisk denotes complex conjugation. Some simplifications were carried out: (i) the $n = 0$ and $n = N/2$ modes have been taken out of the sum; and (ii) the replacement of the "negative" frequency coefficients by the complex conjugate of the corresponding positive frequency is a consequence of the fact that $f(\theta)$ is a real-valued function ($\hat{f}_{-n} = \hat{f}_n^*$). The reason for the limited range of n is explained in Section 4.1.2. The operation that takes the function values in physical space, $f(\theta_j)$, and returns its values in wavenumber space, $\hat{f}_n$, is called forward Fourier transform:

$$\hat{f}_n = \frac{1}{N} \sum_{j=0}^{N-1} f(\theta_j) \exp\left(-\imath n j \frac{2\pi}{N} \right), \qquad n = 0, \dots, N/2, \tag{60}$$

while the reverse operation, expressed in equation (59), is called backward Fourier transform. Note that the normalization by N, here in expression (60), can either be carried out in the forward, in the backward Fourier transform or split between both. It is, obviously, advantageous to do it all at once. Our Fourier transforms are carried out efficiently by an algorithm called fast Fourier transform (FFT) and implemented via the use of the library FFTW [4]. This algorithm takes care of the normalization during the forward transform.

4.1.2 Danger: Aliasing

The range of n in equation (59) is better understood by examining mode representation. First of all, modes with wavenumber $n \geq N$ do not carry any information as our choice of meshgrid makes them project exactly onto modes with wavenumber $m < N$, where $n \equiv m (\mathrm{mod}\, N)$:

$$\exp\left(\imath(m+N)j\frac{2\pi}{N}\right) = \exp\left(\imath mj\frac{2\pi}{N}\right). \tag{61}$$

As a result, all the information is contained in wavenumbers $0 \leq n < N$.

Let us now have a look at a discretized mode of wavenumber $N/2 + m$, with $m < N/2$:

$$\begin{aligned}
\exp\left[\imath\left(\frac{N}{2}+m\right)j\frac{2\pi}{N}\right] &= (-1)^j \exp\left(\imath mj\frac{2\pi}{N}\right) && (62)\\
&= \left[(-1)^j \exp\left(-\imath mj\frac{2\pi}{N}\right)\right]^* && (63)\\
&= \left\{\exp\left[\imath\left(\frac{N}{2}-m\right)j\frac{2\pi}{N}\right]\right\}^*. && (64)
\end{aligned}$$

This introduces a special wavenumber called *Nyqvist frequency*: $n = N/2$. The symmetry around this wavenumber is responsible for a phenomenon called *frequency folding*, whereby modes beyond the Nyqvist frequency fold back onto modes of lower wavenumber: their coefficients appear complex conjugated and with wavenumber reflected with respect to the Nyqvist frequency.

As a result from the above remarks, modes with wavenumber higher than the Nyqvist frequency are incorrectly represented. This phenomenon is called *aliasing* and is exemplified in Figure 3 for two functions. While the continuous representations of $\cos(9\theta)$ and $\cos(7\theta)$ look drastically different, their values coincide on the 16 equidistributed meshpoints between 0 and 2π. This results from the fact that, given a 16 point mesh, the Nyqvist frequency is $N/2 = 8$ and, thus, that wavenumber $n = 9$ folds back onto $n = 7$. A similar observation can be made with $\sin(9\theta)$, with the exception that a change of sign is applied during frequency folding, owing to the complex conjugation of the coefficients, as shown above.

4.1.3 Information storage

The discretized function $f(\theta_j)$ contains N real values. This is the amount of information available at the collocation points. The Fourier coefficients are complex, so the storage of the same amount of information necessitates $N/2$ coefficients, corresponding to wavenumbers 0 to $N/2-1$. However, the first coefficient ($n = 0$) corresponds to the constant mode: $\cos(0\theta) + \imath \sin(0\theta) = 1$. Since f is real-valued, the imaginary part of the associated coefficient contains no information, so one real-valued piece of information has to be contained elsewhere. The first mode above

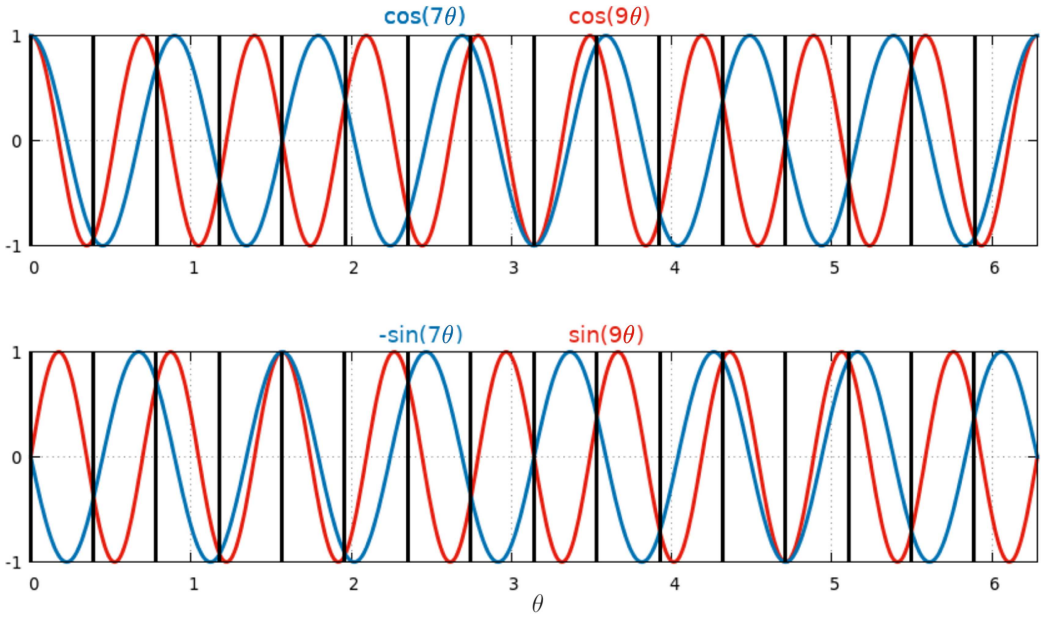


Figure 3: Illustration of the phenomenon of aliasing on a $\Gamma = 2\pi$-periodic domain in θ meshed with $N = 16$ equidistributed points given by equation (56) and whose location is shown by the thick vertical lines. The Nyqvist frequency associated with this mesh is $n_{Nyq} = N/2 = 8$. The top (resp. bottom) figure shows that $\cos(9\theta)$ (resp. $\sin(9\theta)$) coincides with $\cos(7\theta)$ (resp. $-\sin(7\theta)$) onto the meshpoints.

wavenumber $N/2 - 1$ is the Nyqvist mode and takes the form:

$$\exp\left(\imath\frac{N}{2}j\frac{2\pi}{N}\right) = \exp\left(\imath j\pi\right) \tag{65}$$

$$= (-1)^j. \tag{66}$$

Since it is real-valued, it completes the representation of the function in wavenumber space.

In practice, both the coefficients of the constant mode and of the Nyqvist frequency modes are stored as complex numbers but their imaginary part does not enter any calculation. Fourier coefficient information storage is illustrated in Figure 4 for a function discretized over $N = 16$ meshpoints. The physical representation of the function contains 16 real-valued pieces of information When the Fourier transform of the function is taken, a set of 9 complex coefficients are returned, corresponding to wavenumber 0 to 8. However, the imaginary part of the coefficient of the mode of wavenumber 0 and that of the coefficient of the mode of wavenumber 8 do not contain any information. The Fourier coefficients therefore also contain 16 real-valued pieces of information, as many as the collocation representation of the function.

4.1.4 Prevention: De-aliasing

The product of two Fourier modes of wavenumbers n and m redistributes energy to a mode whose wavenumber is $m + n$:

$$\exp\left(\imath n\theta\frac{2\pi}{\Gamma}\right)\exp\left(\imath m\theta\frac{2\pi}{\Gamma}\right) = \exp\left(\imath(n+m)\theta\frac{2\pi}{\Gamma}\right). \tag{67}$$

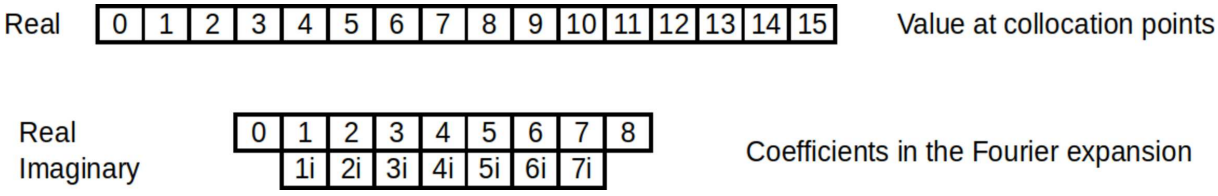


Figure 4: Top: Representation of the information contained in a real-valued, discretized function over $N = 16$ collocation points, labeled 0 to 15. The information contained in physical space is real-valued. Bottom: Representation of the information of the same function via its Fourier coefficients. All coefficients are complex but only the real and imaginary parts that contain information are represented. The numerical label indicates the wavenumber and the suffix "i" indicates the imaginary part of the coefficient.

Figure 5: Graphical representation of the frequency folding phenomenon. In order of ascending wavenumber, modes project onto themselves until the Nyqvist frequency, $N/2$, then spuriously down until wavenumber 0 before going back up again in a periodic process. The idea behind dealiasing via Fourier filtering is to filter out modes with wavenumber larger than k in such a way that when the nonlinearity is applied, frequency folding occurs only onto modes with wavenumber greater than k, as shown by the dark red arrow.

By extension, we can easily see that a nonlinearity of order η applied onto a mode of wavenumber k injects energy at wavenumber ηk. If our mesh does not possess a sufficiently large number of points to accommodate such a function, we will end up with aliasing: the energy contained in that function will spuriously be redistributed to other, lower wavenumber modes, as previously discussed.

One way to prevent aliasing is to use a Fourier filter, which is a simple low-pass filter that replaces the Fourier coefficients of modes whose wavenumber is larger than a cutoff value by 0. To calculate this cutoff value, we needs to make use of our understanding of frequency folding. We want the application of the nonlinearity onto our largest unfiltered wavenumber mode to fold back as close as possible but not onto itself, as illustrated in Figure 5. This can be put in equation as follows:

$$\eta k < \frac{N}{2} + \left(\frac{N}{2} - k \right), \tag{68}$$

where the left-hand-side is the wavenumber resulting from the application of the nonlinearity of order η and where the right-hand-side is the "distance" traveled on the line of wavenumbers from 0 to the folding point $(N/2)$ and back to the original wavenumber $N/2 - k$. It follows that

$$k < \frac{N}{\eta + 1} \tag{69}$$

identifies all the wavenumbers k that can remain unfiltered.

In our case, the nonlinearity is of second order, so we have $k < N/3$, filtering out about $1/3$ of the modes. To take a concrete example, let us consider the case shown in Figure 4, whereby a function is evaluated on $N = 16$ meshpoints. This function is characterized numerically by 16 pieces of real-valued information, whether it is considered in its collocation or its Fourier coefficient form. Using the Fourier filter associated with the quadratic nonlinearity, we should

only retain modes associated with wavenumbers $k < N/3 \approx 5.33$. We then filter out the modes associated with wavenumbers 6, 7 and 8, which corresponds to 5 real-valued pieces of information (remember that wavenumber 8 is the Nyqvist frequency and only contain one piece of real-valued information), thereby losing 31.25% of the information initially contained by the function.

4.2 Zernike tau method for the radial direction

4.2.1 The pole condition

Assuming a Fourier expansion in the azimuthal direction θ, it is natural to think of a discretized function in terms of a linear combination of non-singular monomials in the radial direction. The resulting basis functions take the form:

$$h_{m,n}(r,\theta) = r^n \exp(\imath m\theta), \tag{70}$$

where $n = 0, 1, 2, \ldots$ and where m is the azimuthal wavenumber.

Focusing on the real line, i.e. $x = r$ for $\theta = 0$ and $x = -r$ for $\theta = \pi$, this monomial can be expressed as:

$$\begin{aligned} h_{m,n}(r,0) &= x^n, & (71)\\ h_{m,n}(r,\pi) &= (-x)^n \exp(\imath m\pi) \\ &= (-1)^{m+n} x^n. & (72) \end{aligned}$$

From these expression, one can see that if $m + n$ is not even, the n^{th} derivative of the basis function with respect to r is not continuous at $r = 0$. This imposes a parity condition: if n is odd (resp. even), then m has to be odd (resp. even).

A second condition stems from the application of the polar Laplacian onto our basis functions:

$$\begin{aligned} \nabla^2_{\text{polar}}\, h_{m,n}(r,\theta) &= \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}\right) r^n \exp(\imath m\theta) \\ &= \left(n^2 - m^2\right) r^{n-2} \exp(\imath m\theta). & (73) \end{aligned}$$

This expression contains valuable information: (i) applying the Laplacian onto our basis function drops the r-polynomial degree by 2 and (ii) basis functions with $m = n$ are in the kernel of the Laplacian. Together with the parity condition, we can see that basis functions for which $n = m + 2p$ with $p = 0, 1, 2, \ldots$ are well-behaved when under the action of the Laplacian. Basis functions for which $n = m - 2 - 2p$ are, however, affected by clustering around $r = 0$ as we shall see. For m and n even, successive applications of the Laplacian yield a function that does no longer vary with r. Azimuthal variations of such functions:

$$\frac{1}{r}\frac{\partial}{\partial \theta} \exp(\imath m\theta) = \frac{\imath m \exp(\imath m\theta)}{r} \tag{74}$$

diverge when r tends to zero (the case $m = 0$ yields $n < 0$ and, thus, is not considered). For m and n odd, successive applications of the Laplacian yield a function proportional to r. Applying the Laplacian to this function once more, we get:

$$\nabla^2_{\text{polar}}\, [r \exp(\imath m\theta)] = \frac{1 - m^2}{r} \exp(\imath m\theta). \tag{75}$$

Azimuthal basis function	Allowed radial monomials
1	$1, r^2, r^4, r^6, r^8, r^{10}, r^{12}, \ldots$
$\exp(\imath\theta)$	$r, r^3, r^5, r^7, r^9, r^{11}, r^{13}, \ldots$
$\exp(2\imath\theta)$	$r^2, r^4, r^6, r^8, r^{10}, r^{12}, r^{14}, \ldots$
$\exp(3\imath\theta)$	$r^3, r^5, r^7, r^9, r^{11}, r^{13}, r^{15}, \ldots$
$\exp(4\imath\theta)$	$r^4, r^6, r^8, r^{10}, r^{12}, r^{14}, r^{16}, \ldots$
$\exp(5\imath\theta)$	$r^5, r^7, r^9, r^{11}, r^{13}, r^{15}, r^{17}, \ldots$
$\exp(6\imath\theta)$	$r^6, r^8, r^{10}, r^{12}, r^{14}, r^{16}, r^{18}, \ldots$
$\vdots$	$\vdots$

Table 1: List of the radial monomials satisfying the pole condition for each azimuthal basis function.

This function is also singular at $r = 0$ (the case $m = 1$ also yields $n < 0$ and, thus, is not considered). Examples of troublesome basis functions are provided for illustrative purposes in [3].

As a result, the pole condition yields $n = m + 2p$, where n is the monomial order in the radial direction, m is the azimuthal wavenumber and $p = 0, 1, 2, \ldots$. This condition provides a sparse structure for the function basis, which is shown in Table 1.

4.2.2 The Zernike polynomials

Unfortunately, the radial monomials do not possess good numerical properties: the transfer matrix associated with the transform from physical space to monomial space is ill-conditioned. A natural choice of basis function in polar coordinates is that of the Bessel functions. This, too, is a poor choice: expressing solutions in a basis of Bessel functions displays unsuitably slow convergence due the Gibbs phenomenon at the edge of the disk. A tempting, flexible choice is that of the Chebyshev polynomials but it is far from optimal: any basis of Chebyshev polynomials contains all the monomials and, thus, would waste memory and computing time (half of the modes have the wrong parity, and half of the remaining ones have $n < m$).

The Zernike polynomials provide a natural polynomial basis that satisfies the pole condition. They were developed in optics where they serve, among other things, as basis functions to describe features on circular lenses. These polynomials were used for the first time in fluid dynamics by Matsushima & Marcus [6] and can be defined in several ways. They are solutions of the following Sturm–Liouville problem:

$$\left[\frac{1}{r}\frac{d}{dr}\left((1-r^2)r\frac{d}{dr}\right) - \frac{m^2}{r^2} + n(n+2)\right] Q_n^m(r) = 0, \tag{76}$$

for $0 \leq r \leq 1$ and $0 \leq |m| \leq n$, where m and n are integers. Equation (76) is singular at $r = 1$, thereby preventing the Gibbs phenomenon from developing at the edge of the disk. The indices m and n respectively refer to the azimuthal wavenumber and the radial polynomial order, and are related by the condition: $n = m + 2p$, with $p = 0, 1, 2, \ldots$. The Zernike polynomials can be

obtained using the Frobenius method:

$$Q_n^m(r) = \sum_{q=0}^{\frac{n-|m|}{2}} \frac{(-1)^{q+\frac{n-|m|}{2}}\,\Gamma\left(\frac{n+|m|}{2}+q+1\right)}{q!\,\left(\frac{n-|m|}{2}-q\right)!\,\Gamma\left(|m|+q+1\right)} r^{|m|+2q}, \tag{77}$$

where the Gamma function $\Gamma(x) = \int_0^\infty y^{x-1}\exp(-y)dy$, for a complex number x with positive real part. This special function is often encountered as an extension of the factorial due to the property: $\Gamma(n) = (n-1)!$, when n is a strictly positive integer.

The Zernike polynomials are orthogonal with respect to the weight function r:

$$\int_0^1 Q_n^m(r)\,Q_{n'}^m\,r\,dr = I_n^m\,\delta_{nn'}, \tag{78}$$

where

$$I_n^m = \frac{1}{2n+2} \tag{79}$$

is the normalization constant and $\delta_{nn'}$ is the Kronecker delta.

The first Zernike polynomials are:

$$\begin{aligned}
Q_0^0(r) &= 1, && (80)\\
Q_2^0(r) &= 2r^2-1, && (81)\\
Q_4^0(r) &= 6r^4-6r^2+1, && (82)\\
Q_6^0(r) &= 20r^6-30r^4+12r^2-1, && (83)\\
Q_1^1(r) &= r, && (84)\\
Q_3^1(r) &= 3r^3-2r, && (85)\\
Q_5^1(r) &= 10r^5-12r^3+3r, && (86)\\
Q_2^2(r) &= r^2, && (87)\\
Q_4^2(r) &= 4r^4-3r^2, && (88)\\
Q_6^2(r) &= 15r^6-20r^4+6r^2, && (89)\\
Q_3^3(r) &= r^3, && (90)\\
Q_5^3(r) &= 5r^5-4r^3, && (91)\\
Q_4^4(r) &= r^4, && (92)\\
Q_6^4(r) &= 6r^6-5r^4. && (93)
\end{aligned}$$

The fully two-dimensional basis functions that they generate are represented in Figure 6 for an arbitrary phase. Other phases can be obtained by rotation of the basis function around the disk center. The way these modes generate solutions can be understood from the classification provided by the figure. The higher the azimuthal wavenumber m, the more wavelengths in the azimuthal direction. The higher the radial polynomial order n, the more sign changes from $r = 0$ until $r = 1$. Note that the sign change density increases toward the edge of the disk to prevent clustering at $r = 0$.

4.2.3 Discretizing the disk

We use a standard equidistributed meshgrid in the azimuthal direction:

$$\theta_j = \frac{2\pi j}{M}, \qquad j = 0,\ldots,M-1, \tag{94}$$

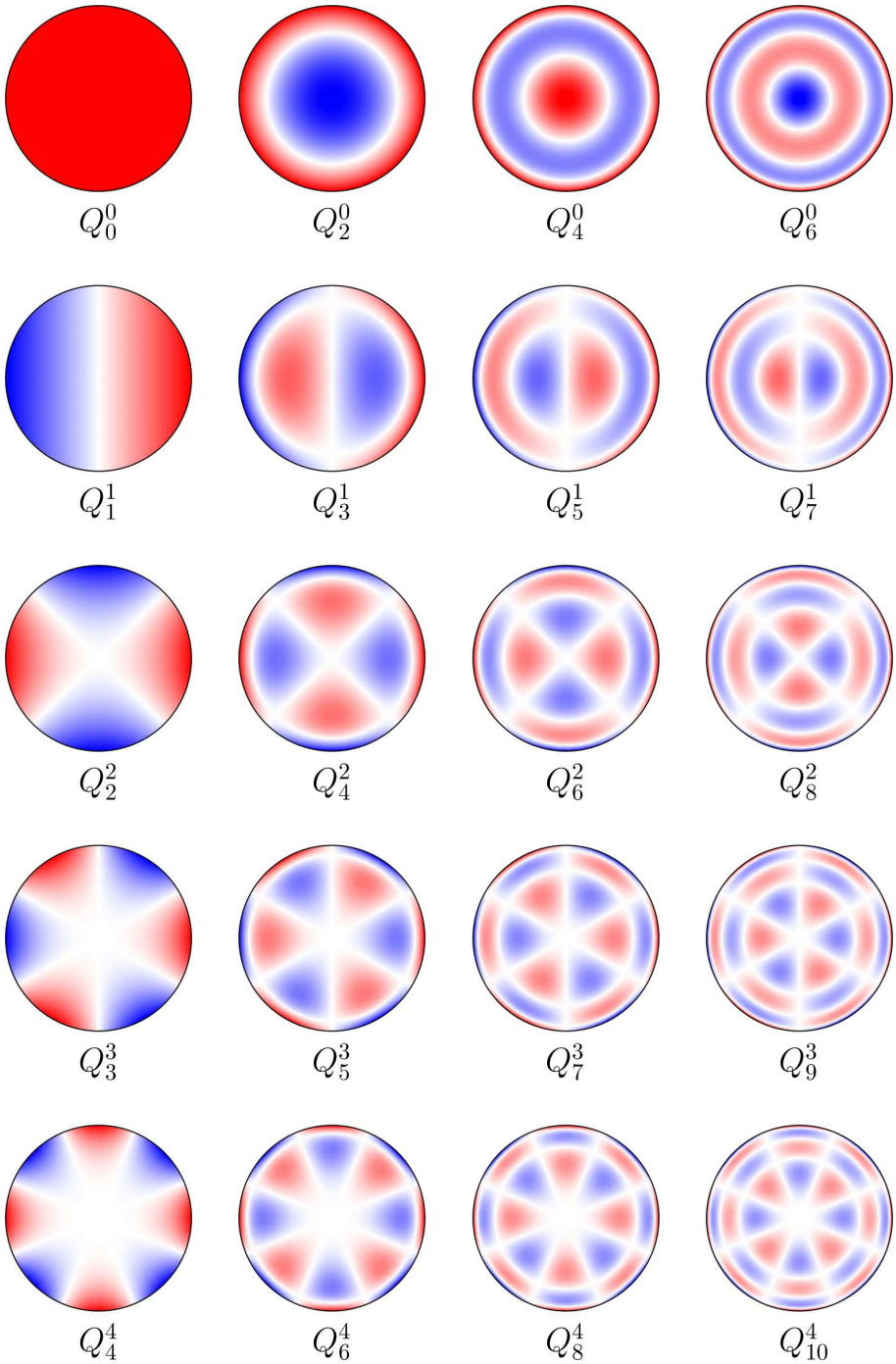


Figure 6: First Zernike modes Q_n^m, where $m \leq 4$ is the azimuthal wavenumber and $n \leq m + 6$ is the radial polynomial order. Red (resp. blue) areas represent positive (resp. negative) values of the mode. Only one phase of the mode is shown here (phase 0, corresponding to the real part of $\exp(\imath m\theta)$). All other phases can be obtained by circular rotation of the modes.

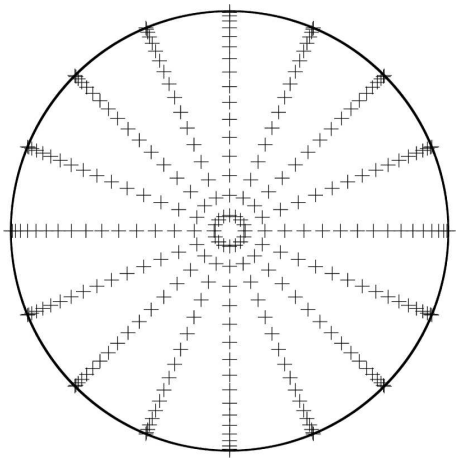


Figure 7: Example of meshgrid generated on the disk for 17 points in the radial direction and 16 points in the azimuthal direction.

in preparation for the use of the Fourier transform. In the radial direction, we wish to use a meshgrid including a point at the boundary to apply boundary conditions. The first N radial points are roots of a polynomial:

$$r_i = \text{roots } \left(Q^0_{2N+2}(r) - Q^0_{2N}(r)\right), \qquad i = 0, \dots, N-1, \tag{95}$$

while the last point $r_N = 1$ is located at the boundary. An example of meshgrid resulting from the above choices is shown in figure 7.

The interplay between the azimuthal wavenumber and the radial polynomial order in the Zernike basis enforces a condition on the meshing parameters. With M points in the azimuthal direction, the maximum azimuthal wavenumber modeled, once the Fourier transform is taken, is $M/2$. Owing to equation (95), an $N+1$-point discretization in the radial direction yields a maximum radial polynomial order of $2N$. To avoid generating a suboptimal meshgrid, we need to ensure that there is at least one radial polynomial associated with the highest azimuthal wavenumbers. This condition implies:

$$M \leq 4N. \tag{96}$$

4.2.4 The Zernike transform

To send a function expressed in the space of the disk into Zernike space, we make first use of the Fourier transform associated with the azimuthal direction:

$$f^m(r) = \frac{1}{M} \sum_{j=0}^{M-1} f(r, \theta_j) \exp\left(-\imath m j \frac{2\pi}{M}\right), \qquad m = 0, \dots, M/2, \tag{97}$$

where the superscript m indicates the azimuthal wavenumber and $\imath$ is the unit imaginary number. The radial direction is then treated using the Zernike transform:

$$f^m_n = \sum_{i=0}^{N} w_i \, f^m(r_i) \, \frac{Q^m_n(r_i)}{I^m_n}, \qquad n = 0, \dots, 2N, \text{with } n = m + 2p \text{ and } p = 0, 1, \dots, \tag{98}$$

where the quadrature weights are:

$$w_i = \frac{(2N+1)\, r_i^2\, I_{2N}^0}{(N+1)^2 \left[Q_{2N}^0(r_i)\right]^2}, \qquad i = 0, \ldots, N-1 \tag{99}$$

$$w_N = I_0^0 - \sum_{i=0}^{N-1} w_i. \tag{100}$$

To send a function back into physical space, we first take the reverse Zernike transform:

$$f^m(r) = \sum_{\substack{n=|m| \\ n+|m| \text{ even}}}^{2N} f_n^m\, Q_n^m(r), \tag{101}$$

followed by the inverse Fourier transform:

$$f(r,\theta_j) = f^0(r) + \sum_{m=1}^{M/2-1} \left[f^m(r) \exp\left(\imath m j \frac{2\pi}{M}\right) + f^{m*}(r) \exp\left(-\imath m j \frac{2\pi}{M}\right)\right] + f^{M/2}(r)(-1)^j, \tag{102}$$

where $j = 0, \ldots, N-1$ and where the asterisk denotes complex conjugation.

4.2.5 Numerical procedure

The Frobenius definition of the Zernike polynomials is numerically unstable: the ratios involved remain of a reasonable magnitude but each of their components becomes large for moderate n and m, which leads to a loss of numerical accuracy. For this reason, we use the recurrence formulae from [6], which consist in generating the polynomial values for each wavenumber m from starting values:

$$Q_{|m|}^m(r) = r^{|m|}, \tag{103}$$

$$Q_{|m|+2}^m(r) = (|m|+1)\left[\frac{|m|+2}{|m|+1}\, r^2 - 1\right] Q_{|m|}^m(r), \tag{104}$$

and successively determining the other polynomial values using the recurrence relation:

$$Q_n^m(r) = \frac{(n-1)\left[4n(n-2)r^2 - 2n(n-1) - 2m^2\right]}{(n-|m|)(n+|m|)(n-2)}\, Q_{n-2}^m(r) \ldots - \frac{n(n+|m|-2)(n-|m|-2)}{(n-|m|)(n+|m|)(n-2)}\, Q_{n-4}^m(r). \tag{105}$$

The numerical procedure to generate the meshpoints and the basis functions is as follows. First of all, we generate the radial mesh by using successive bisection down to a machine-precision gap to find r_i, $i = 0, \ldots, N-1$ from expression (95) and where the successive evaluations of the polynomials are calculated using equations (103)–(105) with $m = 0$. The last point is $r_N = 1$. The value of the basis functions is then computed at the location of the meshpoints using equations (103)–(105). The normalization constants are subsequently computed using equation (79), followed by the quadrature weights, computed using equations (99) and (100). At this point, I noticed that taking the Zernike transform of standard functions forward then backward generated a small error, typically $O(10^{-12})$. This error is due to the accumulation

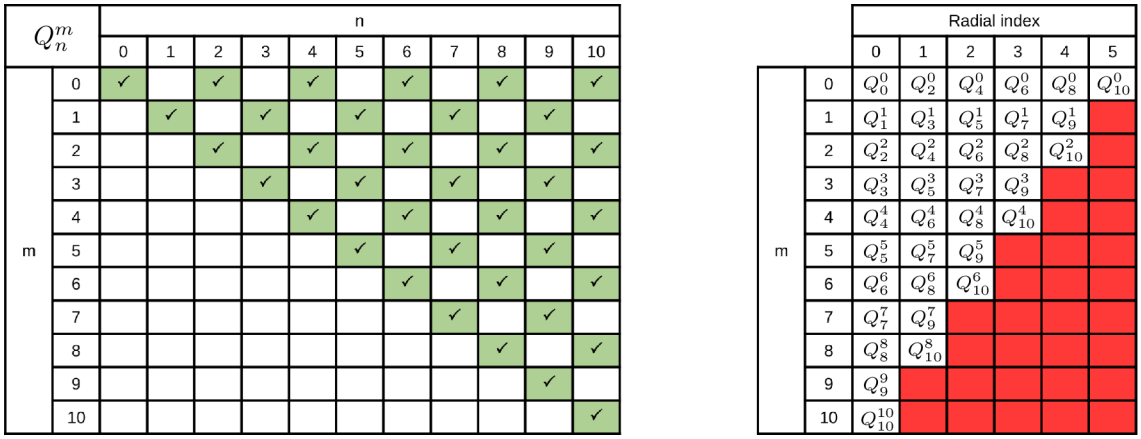


Figure 8: Left: The green boxes with the check mark represent the non-zero Q_n^m modes, where m represents the Fourier index of the azimuthal wavenumber and n the radial Zernike polynomial number. The table represents the modes obtained for a disk discretized using $N+1=6$ radial points and $M=20$ azimuthal points. The modes that are not marked violate the pole condition and, thus, cannot contribute to the representation of the solution. Right: A simple reorganization of the non-zero Q_n^m modes makes information storage optimal.

of operations during the recurrence procedure. To improve the polynomial values, I used a modified Gram–Schmidt projection of the polynomials using the quadrature in expression (98). Using the Gram–Schmidt orthogonalized polynomials, I did not observe any error above 10^{-13} when Zernike transforming forward then backward the same functions.

Question: since the polynomials have been corrected and are used to generate the quadrature weights, should we not recompute the latter? I tried: the process is unstable and the optimal stopping point was after the first Gram–Schmidt projection. Why?

4.2.6 Information storage

The structure of the modes introduced in the previous section is peculiar as $Q_n^m(r) \neq 0$ only when n is greater than or equal to m and shares the same parity. This creates sparse matrices, leading to a waste of memory and of computing time, as shown in the left panel of Figure 8. For $N+1=6$ radial points and $M=20$ azimuthal points, we notice that only 36% of the the Q_n^m satisfy the pole condition. Increasing the number of azimuthal points would not allow the storage of more information as all the newly added modes violate the pole condition, while increasing the number of radial points would increase the proportion of useful modes (to a limit at 50% due to the parity condition).

To take advantage of this structure, we reorganize the modes by replacing the radial number n by a radial index equal to $(n-m)/2$. The modes are now stored according to the radial index and m, as shown on the right panel of Figure 8. To avoid calculating over inexisting modes, we define the azimuthal Zernike dimension:

$$N_{zer}(m) = \text{floor}\left(\frac{2N-m}{2}\right), \tag{106}$$

where m is the azimuthal wavenumber, $N+1$ is the number of radial points (implying that $2N$ is the maximum degree of the radial polynomial). Matrices in Zernike spaces (as the one

represented on the right panel of Figure 8) only contain value from 0 to $N_{zer}(m)$ and loops involving them will simply stop at $N_{zer}(m)$.

Information storage for each mode in the azimuthal direction is carried out as outlined in Section 4.1.3 independently for each m.

4.2.7 Generating operator matrices

Matsushima & Marcus [6] provided a recurrence method to generate the matrix formulation of the Helmholtz operator. First of all, let us assume the following expansion for the component of a function u corresponding to azimuthal wavenumber m:

$$u^m(r) = \sum_{\substack{n=|m| \\ n+|m| even}}^{2N} u_n^m \, Q_n^m(r), \tag{107}$$

which follows from Section 4.2.4 and where we have assumed that the radial direction is discretized using $N+1$ Zernike points. Since differential operations in the azimuthal directions can be dealt with as scalar operations in Fourier space, differential operators are diagonal in that direction and can be written for any given azimuthal wavenumber m independently. For a differential operator acting on azimuthal wavenumber m, $\mathcal{L}^m$, we have:

$$\begin{aligned} f^m(r) &= \mathcal{L}^m u^m(r) \\ &= \sum_{\substack{n=|m| \\ n+|m| even}}^{2N} f_n^m \, Q_n^m(r). \end{aligned} \tag{108}$$

For $\mathcal{L}^m = r^2 \mathcal{H}^m(\kappa)$, where:

$$\mathcal{H}^m(\kappa) = \frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr} - \frac{m^2}{r^2} - \kappa, \tag{109}$$

is the Helmholtz operator parameterized by coefficient κ and written for azimuthal wavenumber m, the following recurrence relation can be written:

$$\begin{aligned} 4\left[(n-|m|)(n+|m|)(n+4)(n+5) - \kappa\frac{n+5}{n}\left(n+m^2\right)\right] u_n^m + \dots \\ 2\left[4(n+1)(n+5)\left(n^2+6n+m^2+4\right) + \kappa\left(n^2+6n+3m^2+2\right)\right] u_{n+2}^m + \dots \\ 4\left[(n-|m|+6)(n+|m|+6)(n+1)(n+2) + \kappa\frac{n+1}{n+6}\left(n-m^2\right)\right] u_{n+4}^m - \dots \\ \kappa\frac{(n-|m|+6)(n+|m|+6)(n+1)(n+2)}{(n+6)(n+7)} u_{n+6}^m - \dots \\ \kappa\frac{(n-|m|)(n+|m|)(n+4)(n+5))}{(n-1)n} u_{n-2}^m = \dots \\ (2n+8)(2n+10) f_n^m - 2(2n+2)(2n+10) f_{n+2}^m + (2n+2)(2n+4) f_{n+4}^m. \end{aligned} \tag{110}$$

Following the notation of [6], this relationship can be written in matrix form:

$$L^m(\kappa) u_n^m = S^m f_n^m, \tag{111}$$

with $L^m(\kappa)$ and S^m suitably defined by equation (110).

Another recurrence relationship can be obtained for operator $\mathcal{L}^m = r^2$:

$$\frac{n^2+m^2+2}{2n(n+2)}u_n^m + \frac{(n-|m|+2)(n+|m|+2)}{4(n+3)(n+2)}u_{n+2}^m + \dots \\ \frac{(n-|m|)(n+|m|)}{4n(n-1)}u_{n-2}^m = f_n^m, \tag{112}$$

which can be similarly written in matrix form:

$$T^m u_n^m = f_n^m. \tag{113}$$

By combining the above operators, we can write the Helmholtz operator:

$$\mathcal{H}^m(\kappa) = (S^m T^m)^{-1} L^m(\kappa). \tag{114}$$

In preparing these matrices, it is important to construct S^m and T^m as non-square matrices. The former has one more column than rows and the latter one more row than columns. Keeping these extra terms by allowing these matrices not to be square takes advantage of the form of the recurrence relations and improves accuracy by 2 orders of magnitude on typical Helmholtz operator inversions.

Another useful operator is $\mathcal{L}^m = r\partial_r$, also defined by a recurrence relationship:

$$nu_n^m + \frac{(n+1)(n+4)}{n+3}u_{n+2}^m = f_n^m - \frac{n+1}{n+3}f_{n+2}^m, \tag{115}$$

where the f_n^m can be stably solved backward by starting from the value $f_{N_{zer}(m)}^m = 0$.

Are there any other operator that need defining?

4.2.8 Boundary conditions

Boundary conditions on the circle can be turned into equations using the Zernike discretization.

For Dirichlet boundary conditions at $r = 1$:

$$u(r=1,\theta) = u_{bc}(\theta) \Longrightarrow \sum_{i=1}^{N_{zer}(m)} Q_{|m|+2(i-1)}^m(r=1)\, u_{|m|+2(i-1)}^m = u_{bc}^m, \qquad m = 0,\dots,M/2, \tag{116}$$

where the boundary condition value on the circle $u_{bc}(\theta)$ is Fourier transformed to yield coefficients u_{bc}^m, $m = 0,\dots,M/2$.

For Neumann boundary conditions at $r = 1$, we use the relation for the operator $r\partial_r$ provided by Matsushima & Marcus [6] to obtain:

$$\frac{du}{dr}(r=1) = u_{bc}(\theta) \Longrightarrow \sum_{i=1}^{N_{zer}(m)} \Bigg[|m| + 2(i-1) Q_{|m|+2(i-1)}^m(r=1)\, u_{|m|+2(i-1)}^m \cdots \\ +2(|m|+2i) Q_{|m|+2(i-1)}^m(r=1) \sum_{p=|m|+2i}^{\infty} u_p^m \Bigg] = u_{bc}^m. \tag{117}$$

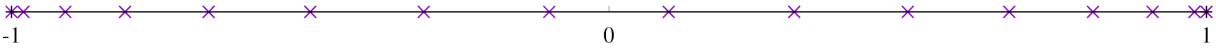


Figure 9: Gauss–Lobatto–Chebyshev mesh for one-dimensional functions between $z = -1$ and $z = 1$. The location of the $N_z + 1 = 16$ meshpoints is shown in purple.

These boundary conditions can easily be programmed as a matrix of coefficients c_i^m and a vector of values u_{bc}^m such that:

$$\sum_{i=1}^{N_{zer}(m)} c_i^m u_i^m = u_{bc}^m, \qquad m = 0, \dots, M/2. \tag{118}$$

4.3 Chebyshev Schur method in the axial direction

4.3.1 Discretization

We use a Chebyshev Schur method in the axial direction based on the Gauss–Lobatto–Chebyshev meshgrid:

$$z_k = z_{min} + \frac{z_{max} - z_{min}}{2}\left[1 + \cos\left(\frac{k\pi}{N_z}\right)\right], \qquad k = 0, 1, \dots, N_z, \tag{119}$$

where $N_z + 1$ is the number of points used to mesh the interval $[z_{min}; z_{max}]$. An example of such a meshgrid is shown in Figure 9. Bear in mind that, due to the (standard) way in which we defined the meshgrid, the meshpoints are organized in reverse order, from the right to the left.

The Gauss–Lobatto–Chebyshev meshpoints allow us to work in the space of the Chebyshev polynomials of the first kind, $T_k(z)$, $k = 0, 1, \dots$. These are typically defined on $[-1; 1]$ where they are solutions of the Sturm–Liouville problem:

$$\left[\frac{d}{dz}\left(\sqrt{1-z^2}\,\frac{d}{dz}\right) + \frac{k^2}{\sqrt{1-z^2}}\right] T_k(z) = 0, \tag{120}$$

augmented with the boundary condition $T_k(z = 1) = 1$. Note that these polynomials are symmetric: T_k is even (resp. odd) if k is even (resp. odd), such that a full set of boundary conditions would be redundant. The Chebyshev polynomials of the first kind are:

$$T_k(z) = \cos\left[k \cos^{-1}(z)\right], \tag{121}$$

which can then be suitably expressed for any interval $[z_{min}; z_{max}]$ by a simple change of variable.

Functions of z can be expressed in the basis of Chebyshev polynomials of the first kind:

$$f(z) = \sum_{k=0}^{\infty} f_k T_k(z). \tag{122}$$

The coefficients f_k can be obtained using the following matrix transform based on the Gauss–Lobatto–Chebyshev meshgrid:

$$f_k = T_{kj}^{fwd} f(z_j), \tag{123}$$

where

$$T_{kj}^{fwd} = \frac{2}{c_j c_k N_z} \cos\left(\frac{\pi j k}{N_z}\right), \tag{124}$$

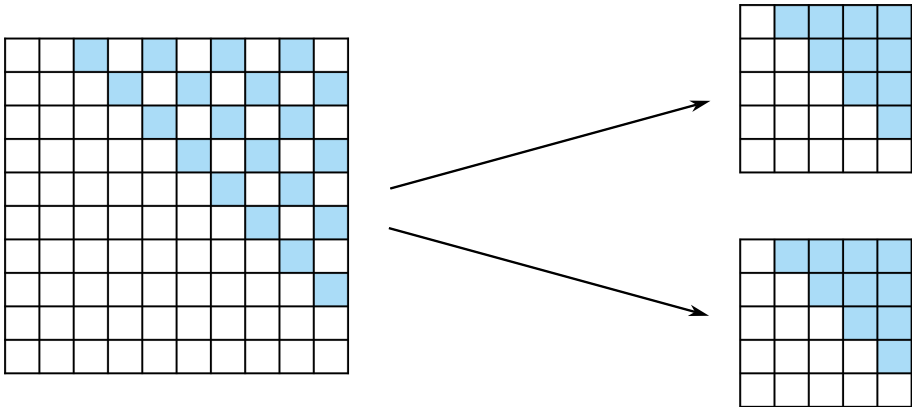


Figure 10: Left: Matrix structure for the full Laplacian operator in Chebyshev space, where the non-zero entries are shown in light blue. Right: Matrix structure for the parity decoupled Laplacian operator in Chebyshev space, where each matrix corresponds to a parity.

where $c_j = 2$ if $j = 0$ or $j = N_z$ and $c_j = 1$ otherwise. The inverse transform is given by:

$$f(z_j) = T^{bwd}_{jk} f_k, \tag{125}$$

where

$$T^{bwd}_{jk} = \cos\left(\frac{\pi jk}{N_z}\right). \tag{126}$$

In practice, due to the structure of these transforms, we can simply obtain the coefficients f_k by the type-I discrete cosine transform. These transforms are then performed using FFTW [4], which scales like $N_z \ln(N_z)$ rather than the above matrix multiplication, which scales like N_z^2.

4.3.2 The Laplacian and its boundary conditions

In the space of the Chebyshev polynomials of the first kind over $[z_{min}; z_{max}]$, the Laplacian matrix is expressed as follows:

$$f = \frac{d^2u}{dz^2} \qquad \Rightarrow \qquad f_k = D_{kj} u_j, \tag{127}$$

where

$$D_{kj} = \begin{cases} \dfrac{4\,j\,(j^2 - k^2)}{c_k (z_{max} - z_{min})^2}, & \text{if } j \geq k+2 \text{ and } j+k \text{ even} \\ = 0, \text{ otherwise} \end{cases}. \tag{128}$$

The Laplacian matrix, D, is less than 1/4 full due to the conditions $j+k$ even (the parity of the Laplacian of a function is the same as the parity of that function) and $j \geq k+2$ (the Laplacian of a polynomial decreases its degree). The resulting matrix structure is shown in the left panel of Figure 10. We can exploit this structure by decoupling the odd and even Chebyshev modes. To do so, we construct the matrix acting on the odd Chebyshev modes, D^{od}, by only keeping the odd rows and columns of the full Laplacian matrix. We also construct the matrix acting on the even Chebyshev modes, D^{ev}, by keeping the even rows and columns of the full Laplacian matrix. The result of this decoupling is two upper-triangular parity-Laplacian matrices that are each of size $(N_z + 1)/2 \times (N_z + 1)/2$, i.e., 4 times smaller than the full Laplacian matrix. In

CyFlow, we define $N_{zpp} = (N_z + 1)/2$ in order to control the number of mode per parity, rather than the number of points N_z.

The last line of each parity-Laplacian matrix is empty (the Laplacian of a polynomial decreases its degree!). We take advantage of this structure to impose the boundary conditions. The Dirichlet and Neumann boundary conditions in Chebyshev space take the form:

$$b_k u_k = \beta, \tag{129}$$

where b_k is the boundary condition vector and β is the boundary condition value in physical space. The boundary conditions are then given by:

- $b_k = 1$ and $\beta = u(z_{max})$ for Dirichlet boundary conditions at z_{max}
- $b_k = (-1)^k$ and $\beta = u(z_{min})$ for Dirichlet boundary conditions at z_{min}
- $b_k = k^2$ and $\beta = u'(z_{max})$ for Neumann boundary conditions at z_{max}
- $b_k = (-1)^k k^2$ and $\beta = u'(z_{max})$ for Neumann boundary conditions at z_{min}

These boundary conditions need to be transformed to be compatible with the parity-Laplacian matrices. They are thus replaced by:

- $b_k = 1$ and $\beta = [u(z_{max}) + u(z_{min})]/2$ for the Dirichlet boundary condition on even modes
- $b_k = 1$ and $\beta = [u(z_{max}) - u(z_{min})]/2$ for the Dirichlet boundary condition on odd modes
- $b_k = k^2$ and $\beta = [u'(z_{max}) + u'(z_{min})]/2$ for the Neumann boundary condition on even modes
- $b_k = k^2$ and $\beta = [u'(z_{max}) - u'(z_{min})]/2$ for the Neumann boundary condition on odd modes

4.3.3 Inverting the Laplacian

To invert the parity Laplacian matrices, D^{od} and D^{ev}, we use a Schur method where the equations corresponding to the two highest frequencies are replaced by the boundary condition equations (129). Since the treatment of the odd and even parity matrices is the same, I will only focus on the odd one here. The system is originally written by separating the high and the low frequency matrix components:

$$\begin{pmatrix} D^{od}_{lolo} & D^{od}_{lohi} \\ D^{od}_{hilo} & D^{od}_{hihi} \end{pmatrix} \begin{pmatrix} u^{od}_{lo} \\ u^{od}_{hi} \end{pmatrix} = \begin{pmatrix} f^{od}_{lo} \\ f^{od}_{hi} \end{pmatrix}, \tag{130}$$

where the odd component of the solution, u^{od}, is of dimension N_{zpp} and decomposed into a vector of coefficients u^{od}_{lo} corresponding to the $N_{zpp}-1$ lowest Chebyshev frequencies and a scalar corresponding to the largest Chebyshev frequencies u^{od}_{hi}. A similar decomposition is applied to the right-hand-side f. The operator matrix is then consequently split into four parts:

- D^{od}_{lolo} is of dimension $(N_{zpp} - 1) \times (N_{zpp} - 1)$ and corresponds to the low frequency components of the operator acting on the low frequency part of the solution
- D^{od}_{lohi} is of dimension $(N_{zpp} - 1) \times 1$ and corresponds to the low frequency components of the operator acting on the high frequency part of the solution

- D^{od}_{hilo} is of dimension $1 \times (N_{zpp} - 1)$ and corresponds to the high frequency components of the operator acting on the low frequency part of the solution
- D^{od}_{hihi} is of dimension 1×1 and corresponds to the high frequency components of the operator acting on the high frequency part of the solution

The Schur method simply consists in replacing the equations on the largest frequency by the boundary conditions. The resulting system reads:

$$\begin{pmatrix} D^{od}_{lolo} & D^{od}_{lohi} \\ b^{od}_{lo} & b^{od}_{hi} \end{pmatrix} \begin{pmatrix} u^{od}_{lo} \\ u^{od}_{hi} \end{pmatrix} = \begin{pmatrix} f^{od}_{lo} \\ \beta^{od} \end{pmatrix}, \tag{131}$$

where b^{od}_{lo} is the row constituted of the first $N_{zpp} - 1$ boundary condition coefficients b_k associated with the desired boundary condition, b^{od}_{hi} is comprised of the corresponding high-frequency coefficient and β^{od} is the right-hand-side β of the boundary condition. Solving the resulting system is done in two steps. First, we solve the low frequencies via:

$$\left[D^{od}_{lolo} - D^{od}_{lohi}(b^{od}_{hi})^{-1} b^{od}_{lo}\right] u^{od}_{lo} = f^{od}_{lo} - D^{od}_{lohi}(b^{od}_{hi})^{-1}\beta^{od}, \tag{132}$$

where we diagonalize the left-hand-side as a preliminary step to the calculation:

$$\left[D^{od}_{lolo} - D^{od}_{lohi}(b^{od}_{hi})^{-1} b^{od}_{lo}\right] = (P^{od})^{-1}\Lambda^{od}P^{od}, \tag{133}$$

where P^{od} is the transfer matrix comprised of the eigenvectors of the left-hand-side matrix and Λ^{od} is the diagonal matrix of its eigenvalues. The inversion is a three-step process easily understandable when equation (132) is cast in the following form:

$$\Lambda^{od}P^{od}u^{od}_{lo} = P^{od}\left[f^{od}_{lo} - D^{od}_{lohi}(b^{od}_{hi})^{-1}\beta^{od}\right]. \tag{134}$$

After evaluating the right-hand-side of equation (132), it is transformed into the eigenspace of the Laplacian by multiplying it by P^{od}. Then, each component is divided by the corresponding eigenvalue of Λ^{od}. The solution u^{od}_{lo} is then recovered by multiplying the result by $(P^{od})^{-1}$.

The final step is to recover the high frequencies through:

$$u^{od}_{hi} = (b^{od}_{hi})^{-1}\left(\beta^{od} - b^{od}_{lo}u^{od}_{lo}\right). \tag{135}$$

5 Solving the Helmholtz problem

Let us now consider the Helmholtz problem:

$$\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}+\frac{\partial^2}{\partial z^2}\right)u=f, \tag{136}$$

where u is the solution and f is a known right-hand-side. For this explanation, we will consider the following Dirichlet boundary conditions:

$$u=u_{rbc} \quad \text{at } r=1, \tag{137}$$

$$u=u_{zbc} \quad \text{at } z=\pm\frac{h}{2}, \tag{138}$$

where $r=1$ is the side-wall of the cylinder and $z=\pm h/2$ are the end-walls of the cylinder.

5.1 Diagonalizations

The first step to solve the Helmholtz equation (136) is to Fourier transform it in the azimuthal direction:

$$\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}-\frac{m^2}{r^2}+\frac{\partial^2}{\partial z^2}\right)u^m(r,z)=f^m(r,z), \tag{139}$$

where the superscript m indicates the azimuthal wavenumber m component. This results in $M/2+1$ decoupled complex equations, where M is the number of meshpoints in the azimuthal direction.

The equations are then transformed in r using the Zernike transform (98) and in z using the Chebyshev transform, taking care of separating the even and odd Chebyshev contributions (see Section 4.3). The resulting equation for the odd modes is:

$$\left(\mathcal{L}^m\otimes I+I\otimes D^{od}\right)U^{od,m}=(I\otimes I)\,F^{od,m}, \tag{140}$$

where $\otimes$ indicates the tensor product between the r-tensor on the left and the z-tensor on the right, $\mathcal{L}^m$ is the discrete version of $\partial^2/\partial r^2+1/r\partial/\partial r-m/r^2$ in Zernike space (it is not needed as such, so read on!), D^{od} is the odd mode contribution to the discretized Laplacian in Chebyshev space (see Section 4.3), I is the identity matrix, $U^{od,m}$ represents the odd modes in z at azimuthal wavenumber m of the solution and where $F^{od,m}$ is the projection of the right-hand-side onto azimuthal wavenumber m and the odd Chebyshev basis in z. A similar equation is found for the even modes and results in the same treatment.

Equation (140) is tensorial, so it is useful to clarify the dimension of its terms. First of all, there are $M/2+1$ such complex equations, due to the azimuthal Fourier transform but these equations are decoupled. For $N+1$ radial points and N_z+1 axial points, matrix $\mathcal{L}^m$ is of dimension $(N+1)\times(N+1)$, matrix D^{od} is of dimension $N_{zpp}\times N_{zpp}$ (cf Section 4.3) and $U^{od,m}$ and $F^{od,m}$ are tensors of dimensions $(N+1)\times N_{zpp}$.

We use a Schur method to apply the boundary condition on the disks: the highest Chebyshev frequency mode equation is discarded and replaced by the boundary conditions in z. This method is described in Section 4.3 and yields:

$$\left[\mathcal{L}^m\otimes I+I\otimes(D^{od}_{lolo}-D^{od}_{lohi}(b^{od}_{hi})^{-1}b^{od}_{lo})\right]U^{od,m}_{lo}=(I\otimes I)\,\hat{F}^{od,m}_{lo}, \tag{141}$$

where $(I \otimes I)\,\hat{F}_{lo}^{od,m} = (I \otimes I)\,F_{lo}^{od,m} - \left(I \otimes D_{lohi}^{od}(b_{hi}^{od})^{-1}\beta^{od}\right)$ is the right-hand-side corrected by the boundary condition contribution and where all the new terms have already been introduced in Section 4.3. Note here that the z-dimension of the tensors in this equation has been reduced to $N_{zpp}-1$ due to the Schur method. The left-hand-side z-operator can be diagonalized like in Section 4.3 to yield:

$$\left[\mathcal{L}^m \otimes I + I \otimes (P^{od})^{-1}\Lambda^{od}P^{od}\right] U_{lo}^{od,m} = (I \otimes I)\,\hat{F}_{lo}^{od,m}, \tag{142}$$

which can be multiplied by $I \otimes P^{od}$ to yield:

$$\left[\mathcal{L}^m \otimes P^{od} + I \otimes \Lambda^{od}P^{od}\right] U_{lo}^{od,m} = \left(I \otimes P^{od}\right)\hat{F}_{lo}^{od,m}. \tag{143}$$

By introducing $\tilde{U}_{lo}^{od,m} = \left(I \otimes P^{od}\right)U_{lo}^{od,m}$ and $\tilde{F}_{lo}^{od,m} = \left(I \otimes P^{od}\right)\hat{F}_{lo}^{od,m}$, we can see the simple numerical structure of the resulting equation:

$$\left[\mathcal{L}^m \otimes I + I \otimes \Lambda^{od}\right]\tilde{U}_{lo}^{od,m} = (I \otimes I)\,\tilde{F}_{lo}^{od,m}. \tag{144}$$

Since the problem is now diagonalized in z, we can decouple all its components and write:

$$\left(\mathcal{L}^m + \Lambda_k^{od}\right)\tilde{U}_{lo,k}^{od,m} = \tilde{F}_{lo,k}^{od,m}, \tag{145}$$

for $k = 1, \ldots, N_{zpp}-1$ and $m = 0, \ldots, M/2$. We note that the left-hand-side operator is nothing else but $\mathcal{H}^m(\Lambda_k^{od})$, defined in Section 4.2.7, so that the equation we need to solve is:

$$\mathcal{H}^m(\Lambda_k^{od})\tilde{U}_{lo,k}^{od,m} = \tilde{F}_{lo,k}^{od,m}. \tag{146}$$

For each azimuthal wavenumber m, we have $N_{zpp}-1$ such equations to solve, each corresponding to an eigendirection of the odd parity Laplace operator in Chebyshev space and as many for the even parity Laplace operator. In total, we have $(M/2+1)(2N_{zpp}-2)$ such problems to solve. Each of these equations needs to be complemented with its corresponding radial boundary condition. Equation (137) is Fourier transformed in θ and Chebyshev transformed in z. The resulting contribution for the odd modes is:

$$u_k^{od,m} = u_{rbc,k}^{od,m}, \qquad \text{at } r = 1, \tag{147}$$

where the superscript m indicates the m^{th} azimuthal wavenumber and the subscript k indicates the k^{th} Chebyshev mode of the parity identified in the first superscript. A similar equation is also obtained for the even modes.

5.2 Tau method on the disks

For simplicity, we recast equation (146) into the following form:

$$\mathcal{H}U = F, \tag{148}$$

with boundary condition (147) expressed in matrix form using (118):

$$c \cdot U = s, \tag{149}$$

where the notation has been simplified for clarity. There are $(M/2+1)(2N_{zpp}-2)$ such problems, i.e., $N_{zpp}-1$ per axial parity per azimuthal wavenumber and all of these problems are solved

in the same way. We continue the calculation in index notation, which makes the algebra more intuitive.

We implement the boundary condition using a tau method:

$$R_{ij}^{-1} L_{jk} U_k = F_i + \tau_1 \delta_{i\bar{N}}, \tag{150}$$

where $\mathcal{H} = R^{-1}L = (ST)^{-1}L$ (see equation (114)), where R and L are of dimension $\bar{N} \times \bar{N}$, where $\bar{N}$ is a function of N_{zer} (see Section 4.2.7), τ_1 is the tau correction to satisfy the boundary condition and δ_{ij} is the Kronecker delta ($\delta_{ij} = 1$ is $i = j$ and 0 otherwise). Sending R to the right-hand-side, we get:

$$L_{jk} U_k = R_{ji} F_i + \tau_1 R_{j\bar{N}}. \tag{151}$$

We want to take advantage of the structure of the matrix L to use fast solvers. By moving all the rows down one row and bringing the last row into first row position, we obtain a penta-diagonal matrix, with the exception of the first row. We apply such a transformation to L and R, and name the resulting matrices $\tilde{L}$ and $\tilde{R}$:

$$\tilde{L}_{jk} U_k = \tilde{R}_{ji} F_i + \tau_1 \tilde{R}_{j\bar{N}}. \tag{152}$$

Since the first row of $\tilde{L}$ prevents it from being penta-diagonal, we replace it with δ_{1j}, leaving 1 on its first component and 0 elsewhere and making the resulting matrix, L', at last penta-diagonal. By doing this, we have modified one row of system (152), so we need to include a second tau correction, τ_2, to allow the solution to be solution of the initial problem:

$$L'_{jk} U_k = \tilde{R}_{ji} F_i + \tau_1 \tilde{R}_{j\bar{N}} + \tau_2 \delta_{j1}. \tag{153}$$

By sending the left-hand-operator to the right-hand-side, we get:

$$U_k = L'^{-1}_{kj} \tilde{R}_{ji} F_i + \tau_1 L'^{-1}_{kj} \tilde{R}_{j\bar{N}} + \tau_2 L'^{-1}_{k1}, \tag{154}$$

which can be written using the notation in [6]:

$$U_k = \bar{A}_k + \tau_1 G_{1,k} + \tau_2 G_{2,k}, \tag{155}$$

where new vectors are introduced: $\bar{A}_k = L'^{-1}_{kj} Y_j = L'^{-1}_{kj} \tilde{R}_{ji} F_i$, $G_{1,k} = L'^{-1}_{kj} G_{3,j} = L'^{-1}_{kj} \tilde{R}_{j\bar{N}}$ and $G_{2,k} = L'^{-1}_{k1}$.

Equation (155) needs to be complemented by the boundary condition

$$c_k U_k = s, \tag{156}$$

as well as the high-frequency equation that we discarded when we replaced the first line of $\tilde{L}$ by δ_{1j}:

$$L_{\bar{N}k} U_k = R_{\bar{N}i} F_i + \tau_1 R_{\bar{N}\bar{N}}. \tag{157}$$

By replacing U_k using equation (155) in these two relationships, we obtain the following system:

$$\begin{cases} c_k G_{1,k} \tau_1 + c_k G_{2,k} \tau_2 = s - c_k \bar{A}_k \\ (f_k G_{1,k} - R_{\bar{N}\bar{N}}) \tau_1 + f_f G_{2,k} \tau_2 = Y_1 - f_k \bar{A}_k \end{cases}, \tag{158}$$

where $f_k = L_{\bar{N}k}$. The solution to this system is:

$$\begin{cases} \tau_1 = \dfrac{c_l G_{2,l} (Y_1 - f_k \bar{A}_k) - f_k G_{2,k} (s - c_l \bar{A}_l)}{\xi} \\ \tau_2 = \dfrac{(f_k G_{1,k} - R_{\bar{N}\bar{N}})(s - c_l \bar{A}_l) - c_l G_{1,l} (Y_1 - f_k \bar{A}_k)}{\xi} \end{cases}, \tag{159}$$

where $\xi = (f_k G_{1,k} - R_{\bar{N}\bar{N}}) c_l G_{2,l} - c_l G_{1,l} f_k G_{2,k}$.

6 Influence matrix method

To be written

7 Nonlinear terms

7.1 Direct expression

The poloidal-toroidal formulation of the governing equations is:

$$\left(\partial_t - \frac{1}{Re}\nabla^2\right)\nabla_h^2\psi = S_\psi, \tag{160}$$

$$\left(\partial_t - \frac{1}{Re}\nabla^2\right)\nabla^2\nabla_h^2\phi = S_\phi, \tag{161}$$

where we have introduced the following nonlinear terms:

$$S_\psi = \hat{\mathbf{z}}\cdot\nabla\times\left[(\mathbf{u}\cdot\nabla)\mathbf{u}\right], \tag{162}$$

$$S_\phi = -\hat{\mathbf{z}}\cdot\nabla\times\nabla\times\left[(\mathbf{u}\cdot\nabla)\mathbf{u}\right]. \tag{163}$$

Using the identities $(\mathbf{u}\cdot\nabla)\mathbf{u} = 0.5\nabla\mathbf{u}^2 - \mathbf{u}\times(\nabla\times\mathbf{u})$ and $\nabla\times(\nabla f) = \mathbf{0}$ for any suitably regular function f, as well as the definition of the vorticity, $\omega = \nabla\times\mathbf{u}$, these nonlinearities can be simplified into:

$$S_\psi = -\hat{\mathbf{z}}\cdot\nabla\times(\mathbf{u}\times\omega), \tag{164}$$

$$S_\phi = \hat{\mathbf{z}}\cdot\nabla\times\nabla\times(\mathbf{u}\times\omega). \tag{165}$$

For simplicity, we introduce:

$$\mathbf{S} = -\mathbf{u}\times\omega \Rightarrow \begin{cases} S_r = u_z\omega_\theta - u_\theta\omega_z \\ S_\theta = u_r\omega_z - u_z\omega_r \\ S_z = u_\theta\omega_r - u_r\omega_\theta \end{cases}. \tag{166}$$

The nonlinear terms therefore become:

$$\left.\begin{aligned} S_\psi &= -\hat{\mathbf{z}}\cdot\nabla\times\mathbf{S} \\ S_\phi &= \hat{\mathbf{z}}\cdot\nabla\times\nabla\times\mathbf{S} \end{aligned}\right\} \Rightarrow \begin{cases} S_\psi = \dfrac{1}{r}\partial_r(rS_\theta) - \dfrac{1}{r}\partial_\theta S_r \\ S_\phi = \nabla_h^2 S_z - \partial_z\left[\dfrac{1}{r}\partial_r(rS_r) + \dfrac{1}{r}\partial_\theta S_\theta\right] \end{cases}. \tag{167}$$

7.2 A good choice of differential operators

Unfortunately, not all the differential operators involved in the numerical expression of the nonlinear terms (167) are compatible with our choice of disctretization of the disk. For example, $f(r,\theta) = r^n\exp(\imath m\theta)$ does not satisfy the pole condition for $n < m$ (see Section 1). This implies that differential operators are only allowed if they do not generate polynomials in r of lower degree when applied to $r^m\exp(\imath m\theta)$.

We note that expressing $\mathbf{S}$ involves a number of differential operators and, while ∂_z and ∇_h^2 do not pose any issue with the pole condition, $r^{-1}\partial_\theta$ and ∂_r do. We replace these operators by ∂_θ and $r\partial_r$ and define the surrogate velocities and vorticities:

$$\mathbf{u}^* = \begin{cases} u_r^* = \partial_\theta\psi + r\partial_{rz}\phi \\ u_\theta^* = \partial_{\theta z}\phi - r\partial_r\psi \\ u_z^* = -\nabla_h^2\phi \end{cases}, \tag{168}$$

$$\omega^* = \begin{cases} \omega_r^* = \partial_\theta u_z^* - \partial_z u_\theta^* \\ \omega_\theta^* = \partial_z u_r^* - r\partial_r u_z^* \\ \omega_z^* = -\nabla_h^2\psi \end{cases}, \tag{169}$$

noting that $u_r^* = ru_r$, $u_\theta^* = ru_\theta$, $u_z^* = u_z$, $\omega_r^* = r\omega_r$, $\omega_\theta^* = r\omega_\theta$ and $\omega_z^* = \omega_z$. The surrogate velocities and vorticities can be direclty computed in spectral space.

The nonlinearity $\mathbf{S}$ has to be computed in physical space and respect the pole condition (see Section 1). This cannot be achieved directly, so we define a surrogate nonlinearity:

$$\mathbf{S}^* = \begin{cases} S_r^* = u_z^*\omega_\theta^* - u_\theta^*\omega_z^* \\ S_\theta^* = u_r^*\omega_z^* - u_z^*\omega_r^* \\ S_z^* = \dfrac{1}{r^2}\left(u_\theta^*\omega_r^* - u_r^*\omega_\theta^*\right) \end{cases}, \tag{170}$$

where $S_r^* = rS_r$, $S_\theta^* = rS_\theta$ and $S_z^* = S_z$. Note that dividing S_r^* or S_θ^* by r would violate the pole condition and would create a large error when transforming these quantities into spectral space.

The nonlinear terms (167) can be expressed in terms of the surrogate quantities:

$$\begin{cases} S_\psi = \dfrac{1}{r}\partial_r S_\theta^* - \dfrac{1}{r^2}\partial_\theta S_r^* \\ S_\phi = \nabla_h^2 S_z^* - \partial_z\left[\dfrac{1}{r}\partial_r S_r^* + \dfrac{1}{r^2}\partial_\theta S_\theta^*\right] \end{cases}. \tag{171}$$

Here, another problem arises: the operator $r^{-1}\partial_r$ is not compatible with the pole condition. When applying this operator to $r^m\exp(\imath m\theta)$, we notice that the additional operator $\imath r^{-2}\partial_\theta$ is required to satisfy the pole condition. We thus write:

$$\text{System (171)} \quad \Rightarrow \quad \begin{cases} S_\psi = \dfrac{1}{r}\partial_r S_\theta^* + \dfrac{\imath}{r^2}\partial_\theta S_\theta^* - \dfrac{\imath}{r^2}\partial_\theta S_\theta^* - \dfrac{1}{r^2}\partial_\theta S_r^* \\ S_\phi = \nabla_h^2 S_z^* - \partial_z\left[\dfrac{1}{r}\partial_r S_r^* + \dfrac{\imath}{r^2}\partial_\theta S_r^* - \dfrac{\imath}{r^2}\partial_\theta S_r^* + \dfrac{1}{r^2}\partial_\theta S_\theta^*\right] \end{cases} \tag{172}$$

$$\Rightarrow \quad \begin{cases} S_\psi = \left(\dfrac{1}{r}\partial_r + \dfrac{\imath}{r^2}\partial_\theta\right) S_\theta^* + \imath\left(\dfrac{\imath}{r^2}\partial_\theta S_r^* - \dfrac{1}{r^2}\partial_\theta S_\theta^*\right) \\ S_\phi = \nabla_h^2 S_z^* - \partial_z\left[\left(\dfrac{1}{r}\partial_r + \dfrac{\imath}{r^2}\partial_\theta\right) S_r^* - \imath\left(\dfrac{1}{r^2}\partial_\theta S_r^* + \dfrac{\imath}{r^2}\partial_\theta S_\theta^*\right)\right] \end{cases} \tag{173}$$

$$\Rightarrow \quad \begin{cases} S_\psi = \left(\dfrac{1}{r}\partial_r + \dfrac{\imath}{r^2}\partial_\theta\right) S_\theta^* + \imath S_c^* \\ S_\phi = \nabla_h^2 S_z^* - \partial_z\left[\left(\dfrac{1}{r}\partial_r + \dfrac{\imath}{r^2}\partial_\theta\right) S_r^* - S_c^*\right] \end{cases}, \tag{174}$$

where

$$S_c^* = \frac{\imath}{r^2}\partial_\theta\left(S_r^* + \imath S_\theta^*\right). \tag{175}$$

With this algebra, we notice that all the operations carried out on $\mathbf{S}^*$ can be performed in spectral space. The only dangerous evaluations are those of S_c^*. They will be carried out in quasi-physical space (physical in r and z but spectral in θ to take advantage of the differentiation in this direction).

A Vector calculus

Let ϕ be a continuously twice-differentiable scalar field:

$$\nabla^2\phi = \nabla\cdot(\nabla\phi)\,. \tag{176}$$

Let $\mathbf{u}$ be a continuously twice-differentiable vector field:

$$\nabla\cdot(\nabla\times\mathbf{u}) = 0, \tag{177}$$

$$(\mathbf{u}\cdot\nabla)\,\mathbf{u} = \frac{1}{2}\nabla\mathbf{u}^2 - \mathbf{u}\times(\nabla\times\mathbf{u})\,, \tag{178}$$

$$\nabla^2\mathbf{u} = \nabla(\nabla\cdot\mathbf{u}) - \nabla\times(\nabla\times\mathbf{u}) \tag{179}$$

B Cylindrical calculus

Let us consider a suitably regular scalar field ϕ:

$$\nabla^2\phi = \frac{1}{r}\partial_r(r\partial_r\phi) + \frac{1}{r^2}\partial_\theta^2\phi + \partial_z^2\phi. \tag{180}$$

Let us consider a suitably regular vector field $\mathbf{A} = A_r\hat{\mathbf{r}} + A_\theta\hat{\theta} + A_z\hat{\mathbf{z}}$ with its projections along the radial, azimuthal and axial directions of respective unit vectors $\hat{\mathbf{r}}$, $\hat{\theta}$ and $\hat{\mathbf{z}}$:

$$\nabla\times\mathbf{A} = \begin{pmatrix} \dfrac{1}{r}\partial_\theta A_z - \partial_z A_\theta \\ \partial_z A_r - \partial_r A_z \\ \dfrac{1}{r}\left[\partial_r(rA_\theta) - \partial_\theta A_r\right] \end{pmatrix}. \tag{181}$$

Similarly:

$$\nabla\times\nabla\times\mathbf{A} = \begin{pmatrix} \partial_r\left(\dfrac{1}{r}\partial_\theta A_\theta + \partial_z A_z\right) + \dfrac{2}{r^2}\partial_\theta A_\theta - \left(\dfrac{1}{r^2}\partial_\theta^2 + \partial_z^2\right)A_r \\ \dfrac{1}{r}\partial_\theta\left(\partial_r A_r + \partial_z A_z\right) + \dfrac{1}{r^2}\left(A_\theta - \partial_\theta A_r\right) - \left[\dfrac{1}{r}\partial_r(r\partial_r) + \partial_z^2\right]A_\theta \\ \partial_z\left[\dfrac{1}{r}\partial_r(rA_r) + \dfrac{1}{r}\partial_\theta A_\theta\right] - \left[\dfrac{1}{r}\partial_r(r\partial_r) + \dfrac{1}{r^2}\partial_\theta^2\right]A_z \end{pmatrix}. \tag{182}$$

The Laplacian is not straightforward:

$$\nabla^2\mathbf{A} = \begin{pmatrix} \nabla^2 A_r - \dfrac{A_r}{r^2} - \dfrac{2\partial_\theta A_\theta}{r^2} \\ \nabla^2 A_\theta - \dfrac{A_\theta}{r^2} + \dfrac{2\partial_\theta A_r}{r^2} \\ \nabla^2 A_z \end{pmatrix}, \tag{183}$$

where the expression of the scalar field Laplacian is given in equation (180).

References

[1] P. Boroński. A method based on poloidal-toroidal potentials applied to the von kármán flow in a finite cylinder geometry. *PhD*, 2007.

[2] P. Boroński and L. S. Tuckerman. Poloidal-toroidal decomposition in a finite cylinder i. influence matrices for the magnetohydrodynamic equations. *J. Comput. Phys.*, 227, 2007.

[3] P. Boroński and L. S. Tuckerman. Poloidal-toroidal decomposition in a finite cylinder ii. discretization, regularization and validation. *J. Comput. Phys.*, 227, 2007.

[4] M. Frigo and S. G. Johnson. The design and implementation of FFTW3. *Proceedings of the IEEE*, 93, 2005.

[5] F. Marques. On boundary conditions for velocity potentials in confined flows: Application to couette flow. *Phys. Fluids*, 2:729–737, 1990.

[6] T. Matsushima and P. S. Marcus. A spectral method for polar coordinates. *J. Comput. Phys.*, 120:365–374, 1995.