UAB
Universitat Autònoma
de Barcelona

Treball final del grau de matemàtiques

Aplicació d'un mètode espectral a fluxos bidimensionals incompressibles en un domini rectangular

Guillem Masdemont Serra

Supervisat per Prof. Jezabel Curbelo
Tutor acadèmic Prof. Armengol Gasull

Facultat de Ciències

Convocatòria
Bellaterra, Juny, 2025

Resum

Aquest treball se centra en la modelització d'un flux bidimensional entre dues parets en moviment, mitjançant mètodes espectrals aplicats a equacions en derivades parcials. Es parteix del mètode clàssic de Moser et al. (1982) per a parets fixes i l'adaptem per incorporar moviment de les parets, forces externes i un tractament propi del terme no lineal, absent a l'article original. Així, aquest estudi persegueix un doble objectiu: d'una banda, aprofundir en les implicacions matemàtiques associades a la modelització de fluxos bidimensionals a partir de fonts de referència fonamentals; i, de l'altra, aplicar una anàlisi computacional pròpia mitjançant la simulació d'exemples originals, sovint absents dels tractaments teòrics habituals presents en la literatura.

Els resultats mostren que el mètode de Petrov-Galerkin proposat és capaç de capturar fenòmens físics rellevants, com la conservació qualitativa de l'energia i la formació de remolins, tot i les limitacions associades a la sensibilitat caòtica del sistema. Paral·lelament, aquest treball ha permès aprofundir en els coneixements teòrics i computacionals sobre els mètodes espectrals. Com a línies futures, es proposa millorar l'eficiència de la implementació computacional, explorar estratègies de resolució més estables i aprofundir en tècniques de canvi de base espectral.

Resumen

En el presente trabajo estudiamos la modelización de un flujo bidimensional entre dos paredes en movimiento mediante métodos espectrales aplicados a ecuaciones en derivadas parciales. Se parte del método clásico de Moser et al. (1982) para paredes fijas y se adapta para incorporar el movimiento de las paredes, fuerzas externas y un tratamiento propio del término no lineal, ausente en el artículo original. Así, este estudio persigue un doble objetivo: por un lado, profundizar en las implicaciones matemáticas asociadas a la modelización de flujos bidimensionales a partir de fuentes de referencia fundamentales; y por otro, aplicar un análisis computacional propio mediante la simulación de ejemplos originales, a menudo ausentes de los tratamientos teóricos habituales presentes en la literatura.

Los resultados muestran que el método de Petrov-Galerkin propuesto es capaz de capturar fenómenos físicos relevantes, como la conservación cualitativa de la energía y la formación de remolinos, a pesar de las limitaciones asociadas a la sensibilidad caótica del sistema. Paralelamente, el desarrollo ha permitido afianzar los conocimientos teóricos y computacionales sobre los métodos espectrales. Como líneas futuras, se propone mejorar la eficiencia de la implementación computacional, explorar nuevas estrategias de resolución estable y avanzar en técnicas de cambio de base espectral.

Abstract

We focus on the modelling of a two-dimensional flow between two moving walls using spectral methods applied to partial differential equations. Our work builds upon the classical method of Moser et al. (1982) for fixed walls, and we adapt it to incorporate wall motion, external forces, and a custom treatment of the nonlinear term, which is not presented in the original paper. The study has a dual objective: to deepen the understanding of spectral methods in simple domains based on key reference works, and to analyse their computational implementation through an original example. The work adopts an applied and numerical perspective, with specifically designed simulations.

The results show that the proposed Petrov-Galerkin method is successful in describing relevant physical phenomena, such as the qualitative conservation of energy and the formation of vortices, despite the limitations associated with the system's chaotic sensitivity. In parallel, the development process has contributed to acquiring a deeper understanding of spectral methods, both from theoretical and computational perspectives. As future directions, we suggest improving computational efficiency, explore new strategies for stable numerical integration, and advancing spectral basis transformation techniques.

Índex

1 Introducció

El desenvolupament del càlcul numèric ha estat estretament vinculat a les necessitats pràctiques de cada època. En l'era de la supercomputació, destaca l'esforç per simular la dinàmica dels fluids terrestres i estudiar l'evolució climàtica global [1]. Per a problemes d'aquesta escala, és habitual descompondre el domini en subdominis de geometria simple [2], on les equacions es resolen localment amb condicions de contorn adequades. En aquests casos, fins i tot els models més simples presenten solucions complexes que resulten d'interès, com ara els problemes de convecció compressible, Rayleigh-Bénard o Taylor-Couette.

En el present treball ens centrarem en la modelització d'equacions en derivades parcials en dominis simples utilitzant mètodes espectrals amb l'objectiu de modelitzar un problema original: el flux bidimensional confinat entre dues parets en moviment. Partirem del mètode proposat per Moser et al. (1982) [3] per al cas amb parets fixes, i l'adaptarem per incorporar el moviment de les parets i la presència de forces externes. A més, proposem i implementem un tractament propi per al terme no lineal, absent en l'article original. Així, aquest estudi persegueix un doble objectiu: d'una banda, entendre les obres fonamentals [4, 5, 6] que sustenten la teoria dels mètodes espectrals en dominis simples; de l'altra, analitzar els costos computacionals i els desafiaments pràctics associats a la seva implementació. El treball se situa en l'àmbit del càlcul numèric aplicat, amb exemples i simulacions dissenyades originalment. Les simulacions han estat realitzades amb un processador AMD Ryzen 5 3500U, utilitzant Python sota Windows 11.

El treball s'estructura de la manera següent: a la Secció 2 s'introdueixen les equacions de Navier-Stokes i els aspectes clau per a la seva modelització mitjançant mètodes espectrals. A la Secció 3 es presenta la base teòrica d'un mètode espectral, amb exemples que serveixen de referència per abordar correctament el mètode de Moser et al. A la Secció 4 es desenvolupa el problema original basat en aquests fonaments, i finalment es recullen les conclusions a la Secció 5. Els apèndixs A i B inclouen eines numèriques complementàries, introduïdes breument al llarg del grau i accessibles per a consulta si escau.

2 La dinàmica de fluids

Les equacions en derivades parcials de Navier-Stokes són conegudes per governar la dinàmica dels fluids. Van ser proposades a mitjan segle XIX de la mà de Claude-Louis Navier i van ser progressivament desenvolupades per George Gabriel Stokes. A les següent pàgines en proposem una derivació intuïtiva pel cas incompressible que ocupa aquest treball. Posteriorment, introduïm el nombre de Reynolds i les nocions essencials de vorticitat d'un camp, funcions de corrent i teoremes d'existència i unicitat. Aquesta informació permetrà motivar els estudis de la Secció 3.4 i Secció 4.

2.1 Equacions de Navier-Stokes

Considerem una regió de l'espai $\Omega \subset \mathbb{R}^3$ (o bidimensional) saturada per un fluid[1] en la qual s'hi defineixen funcions suficientment suaus, intrínseques del sistema: La densitat $\rho(\mathbf{x}) : \mathbb{R}^3 \to \mathbb{R}$, viscositat $\mu(\mathbf{x}) : \mathbb{R}^3 \to \mathbb{R}$ i força externa[2] $\mathbf{b}(\mathbf{x}, t) : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$. L'existència física d'aquestes magnituds està sotmesa a les hipòtesis de la mecànica estadística en un món macroscòpic.

Per a cada posició $\mathbf{x} = (x, y, z) \in \Omega$ existeix una partícula del fluid que hi passa en un instant $t \in \mathbb{R}$ i segueix una trajectòria ben definida $\mathbf{x}(t)$. Es defineix el camp $\mathbf{u}(\mathbf{x}, t) \in \mathbb{R}^3$ com el camp de velocitats del fluid a l'espai-temps. Les equacions que governen la dinàmica del camp $\mathbf{u}$ es basen en tres principis universals: La conservació de la massa, la conservació moment lineal i la conservació de l'energia.

Principi 2.1 (Conservació de la massa)**.** *La massa no es crea ni es destrueix. Per tant, la variació de densitat de massa dins un domini tancat és deguda a l'entrada o sortida de massa a través del contorn del domini. Matemàticament*

$$\frac{d}{dt}\int_W \rho \, dV = -\int_{\partial W} \rho \mathbf{u} \cdot \mathbf{n} \, dS \tag{2.1}$$

on $W \subset \Omega$ és un domini arbitrari tancat, $\mathbf{n}$ és el vector normal exterior la superfície orientable $S = \partial W$. El signe de (2.1) depèn de la tria d'orientació del vector normal $\mathbf{n}$.

Atès que aquest raonament és vàlid per a qualsevol domini W, s'aplica el *teorema de la divergència de Gauss* per transformar el flux a través de la superfície en una integral volumètrica i es reuneixen la part dreta i esquera de la igualtat sota el mateix signe integral. Així, s'arriba a la forma diferencial de la conservació de la massa[3]:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0. \tag{2.2}$$

Definició 2.1 (Fluid homogeni i incompressible)**.** *Diem que un fluid és homogeni si la seva densitat roman constant en l'espai $\rho(\mathbf{x}, t) = \rho(t)$. Diem que un fluid és incompressible si qualsevol porció del seu volum roman invariant amb el temps. El camp de velocitats en un fluid homogeni i incompressible satisfà*

$$\nabla \cdot \mathbf{u} = 0. \tag{2.3}$$

Principi 2.2 (Conservació del moment lineal)**.** *El moment lineal és conserva, és a dir, l'acceleració que experimenta una partícula del fluid és deguda a la suma de forces que actuen sobre ella.*

Matemàticament, l'acceleració d'una partícula amb velocitat $\mathbf{u}(\mathbf{x}(t), t)$ ve donada per

$$\mathbf{a}(t) = \frac{d^2\mathbf{x}}{dt^2} = \frac{d\mathbf{u}}{dt} = \partial_t \mathbf{u} + (\mathbf{u} \cdot \nabla)\mathbf{u},$$

[1] Cal entendre un fluid com una substància que es deforma contínuament (*Hipòtesi del continu*) sota l'acció d'una força de cisallament, per petita que sigui.

[2] Físicament, pot incloure la gravetat o forces electromagnètiques

[3] Anomenada equació de continuïtat; esdevé fonamental en diversos camps de la física, com l'electrodinàmica (conservació de la càrrega), la termodinàmica (conservació de l'energia) o la mecànica quàntica (conservació de la probabilitat).

on a la última igualtat s'aplica la regla de la cadena.
L'operador anterior s'acostuma a anomenar **derivada material** o **derivada total** seguint la trajectòria d'una partícula del fluid:

$$\frac{D(\cdot)}{Dt} = \partial_t(\cdot) + \mathbf{u} \cdot \nabla(\cdot). \tag{2.4}$$

Per aplicar el principi 2.2 es defineixen les forces que actuen sobre la partícula. La força exercida sobre un fluid es mesura prenent un volum tancat $W \subset \Omega$ i analitzant les forces que actuen sobre la seva superfície $S = \partial W$. En mecànica de fluids l'estudi es divideix en dos casos fonamentals, el d'Euler i el de Navier-Stokes

$$\text{Força per unitat d'àrea sobre la superfície S} = \begin{cases} -p(\mathbf{x},t)\mathbf{n} & \text{Cas d'Euler} \\ -p(\mathbf{x},t)\mathbf{n} + \boldsymbol{\sigma}(\mathbf{x},\boldsymbol{t}) \cdot \mathbf{n} & \text{Cas de Navier-Stokes} \end{cases}.$$

on $p(\mathbf{x},t) \in \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}$ és una funció (incògnita o no segons el problema) anomenada pressió, $\mathbf{n}$ és el vector normal exterior a S i $\boldsymbol{\sigma}(\mathbf{x},\boldsymbol{t})$ és el tensor d'estrès, una aplicació tensorial d'ordre dos representada per una matriu 3×3 en coordenades.

En el cas d'Euler, diem que el fluid és **ideal** ja que no assumim interaccions viscoses o fregament entre les partícules. L'única força que es considera és la força normal deguda a la pressió del reste de fluid sobre el volum W. A més d'això, també hi afegim un terme forces externes per unitat de massa, denotat per $\mathbf{b}(\mathbf{x},t)$. Resumidament, la força total exercida sobre ∂W és la suma de forces sobre tot el domini tancat, que es pot escriure com

$$\mathbf{F}_{\partial W} = -\int_{\partial W} p\mathbf{n}\, dA + \int_W \rho\mathbf{b}\, dV = -\int_W (\nabla p + \rho\mathbf{b})\, dV.$$

En la segona igualtat s'utilitza el *Teorema de la divergència de Gauss.*

Pel cas de Navier-Stokes, assumim que les partícules interactuen entre elles mitjançant un tensor d'estrés $\boldsymbol{\sigma}$[4]. Si se suposa que $\boldsymbol{\sigma}$ depen linealment del gradient del camp de velocitats, que és invariant sota rotacions i és simetric (balanç de moment angular), aleshores el tensor $\boldsymbol{\sigma}$ es pot caracteritzar completament mitjançant dos coefficients reals i constants, μ, λ anomenats primer i segon coeficient de viscositat, respectívament. Desenvolupant un raonament similar al cas anterior s'arriba a

$$\mathbf{F}_{\partial W} = -\int_W (\nabla p + (\lambda+\mu)\nabla(\nabla \cdot \mathbf{u}) + \mu\Delta\mathbf{u} + \mathbf{b})\, dV,$$

on $\Delta\mathbf{u} = (\partial_{xx} + \partial_{yy} + \partial_{zz})\,\mathbf{u}$. Un procediment detallat de les deduccions anteriors es pot trobar a [7].

Definició 2.2. *Aplicant el Principi 2.2 i seguint la notació d'aquesta secció, definim la forma diferencial de les equacions del moment pel cas d'Euler i Navier-Stokes com*

$$\rho\frac{D\mathbf{u}}{Dt} = -\nabla p + \rho\mathbf{b}, \qquad \textit{Cas d'Euler.} \tag{2.5a}$$

$$\rho\frac{D\mathbf{u}}{Dt} = -\nabla p + (\lambda+\mu)\nabla(\nabla \cdot \mathbf{u}) + \mu\Delta\mathbf{u} + \rho\mathbf{b}, \qquad \textit{Cas de Navier-Stokes.} \tag{2.5b}$$

[4]Des del punt de vista físic, això equival a incorporar els efectes de la viscositat al problema.

Principi 2.3 (Conservació de l'energia)**.** *L'energia no es crea ni es destrueix.*

Seguint el primer principi de la termodinàmica, podem assumir que l'energia del sistema es pot descompondre com

$$E_{\text{total}} = E_{\text{cinètica}} + E_{\text{interna}}.$$

on l'energia cinètica ve donada per l'expressió següent, mentre que l'energia interna és funció de la temperatura

$$E_{\text{cinètica}} = \frac{1}{2}\int_W \rho\|\mathbf{u}\|^2 dV, \qquad E_{\text{interna}} \propto \text{Temperatura}.$$

Cal igualar les expressions anteriors a l'energia total del sistema. Nogensmenys, tals consideracions depenen de les hipòtesis físiques que es considerin, com ara si el fluid és homogeni, incompressible, no isentrop[5] o isoterm. Presentem el cas incompressible, que esdevé el de més importància per aquest treball. Assumint densitat ρ constant, l'expressió de l'energia cinètica es desenvolupa com

$$\frac{d}{dt}E_{\text{cinètica}} = \frac{d}{dt}\left[\frac{1}{2}\int_{W_t}\rho\|\mathbf{u}\|^2\, dV\right] = \int_{W_t}\rho\left(\mathbf{u}\cdot\left(\frac{\partial\mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\mathbf{u}\right)\right)dV.$$

I, en aquest escenari, la condició d'incompressibilitat ($\nabla\cdot\mathbf{u} = 0$) condueix a la següent igualtat pel cas d'Euler [7]:

$$\int_{W_t}\rho\left(\mathbf{u}\cdot\left(\frac{\partial\mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\mathbf{u}\right)\right)dV = -\int_{W_t}(\mathbf{u}\cdot\nabla p - \rho\mathbf{u}\cdot\mathbf{b})dV$$

el qual s'observa que és essencialment conseqüència directa de la conservació del moment lineal (2.5a). Un resultat anàleg s'obté per a les equacions de Navier–Stokes (2.5b). Resumidament, s'observa que el Principi 2 i el Principi 3 aporten essencialment la mateixa informació per a fluids homogenis incompressibles[6].

Per acabar, les **condicions de contorn** especifiquen com el fluid interactua amb les vores del domini. Per a les equacions d'Euler, només es considera la component normal de la força, cosa que implica que el fluid no pot travessar la paret però pot relliscar-hi:

$$\mathbf{u}\cdot\mathbf{n} = 0 \quad \text{a } \partial\Omega \quad \text{(Euler)}.$$

En canvi, en el cas de Navier-Stokes, s'introdueix també la component tangencial i, d'acord amb l'evidència experimental i també matemàtica (derivades d'ordre superior a (2.5b) a causa del Laplacià) s'imposa que el fluid resti adherit a la paret:

$$\mathbf{u} = 0 \quad \text{a } \partial\Omega \quad \text{(Navier-Stokes)}.$$

Amb aquestes condicions i els principis establerts, el model de Navier-Stokes descriu completament el comportament d'un fluid homogeni i incompressible.

Definició 2.3 (Equacions incompressibles d'Euler i Navier-Stokes)**.** *Siguin $\Omega \subset \mathbb{R}^n$ (amb $n = 2$ o 3) un domini regular i limitat que representa el fluid, amb frontera $\partial\Omega$, i $\mathbf{u} : \Omega \times [0,T] \to \mathbb{R}^n$ el camp de velocitat del fluid. Les equacions d'Euler i de Navier-Stokes per un fluid incompressible es donen, respectivament, i, seguint la notació definida a l'inici d'aquesta secció, per:*

Euler incompressibles	***Navier-Stokes incompressibles***
$\rho\dfrac{D\mathbf{u}}{Dt} = -\nabla p + \rho\mathbf{b}$	$\rho\dfrac{D\mathbf{u}}{Dt} = -\nabla p + \mu\Delta\mathbf{u} + \rho\mathbf{b}$
$\nabla\cdot\mathbf{u} = 0$	$\nabla\cdot\mathbf{u} = 0$
$\mathbf{u}\cdot\mathbf{n} = 0 \quad a\ \partial\Omega$	$\mathbf{u} = \mathbf{0} \quad a\ \partial\Omega$

[5] Isentrop és un procés on l'entropia del sistema, quantitat de desordre físic, és manté constant, com ara un procés reversible.

[6] Aquest fet es troba en acord amb el nombre de incògnites que el sistema presenta.

Les equacions d'Euler constitueixen una formulació fonamental per a l'estudi del comportament dels fluids, ja que representen una simplificació de les equacions de Navier-Stokes en el cas de fluids no viscosos (coeficient de viscositat nul $\mu = 0$). Cal remarcar, però, que la naturalesa d'ambdues formulacions és fonamentalment diferent, ja que les condicions de contorn associades també ho són. Fins i tot en el límit $\mu = 0$, no s'obtenen les equacions d'Euler com a cas particular de les de Navier-Stokes. Per conveniència, es defineixen tot seguit els següents termes:

$$\Delta \mathbf{u} \quad \text{terme difusiu o dissipatiu} \qquad\qquad (\mathbf{u}\cdot\nabla)\mathbf{u} \quad \text{terme d'inèrcia o convectiu}$$

Introduïm a continuació aspectes claus sobre l'existència i l'evolució energètica de les solucions, fonamentals per a la modelització teòrica en els capítols posteriors.

Teorema 2.4 (Existència i unicitat en 2D per Navier–Stokes incompressibles [8])**.** *Siguin* $\mathbf{u}_0 \in H^s(\Omega)$ *i* $\nabla \cdot \mathbf{u}_0 = 0$ *amb* $s > 2$ *i* $\Omega \subset \mathbb{R}^2$ *un domini acotat amb condicions de contorn regulars (Dirichlet o homogènies). Aleshores, existeix una única solució regular*

$$\mathbf{u} \in C([0,\infty); H^s(\Omega)) \cap C^1((0,\infty); H^{s-2}(\Omega))$$

del sistema de Navier–Stokes incompressible, per a tot temps $t > 0$.
A més, la solució és única dins aquesta classe funcional i depèn contínuament de les dades inicials.

Essencialment, [9] exposa que el teorema anterior es demostra mostrant que la solució es manté regular i no desenvolupa singularitats en temps finit. No obstant això, aquest resultat no es pot estendre al cas tridimensional. Actualment, es desconeix si existeixen solucions regulars globals per a les equacions de Navier–Stokes en tres dimensions, així com la possible aparició de singularitats en temps finit. També és desconeguda la unicitat d'aquestes solucions[7]. En qualsevol cas, la demostració per al cas bidimensional es fonamenta en un resultat energètic clau, de gran utilitat pràctica.

Proposició 2.5 (Igualtat de l'energia [10])**.** *Considerem les equacions descrites a la definició 2.3. A les equacions d'Euler assumim que el fluid és ideal, és a dir, que no hi ha interacció microscòpica entre partícules. Si el fluid és, a més, incompressible, es considera que tota l'energia del sistema és energia cinètica. En absència de forces externes, això implica la conservació de l'energia cinètica:*

$$\frac{dE_{cinètica}}{dt} = 0 \qquad \textit{(Euler + incompressible)}$$

En Navier-Stokes, la presència del terme de viscositat permet la dissipació d'energia cinètica per fricció. Aquest fenomen és anàleg a la dissipació de temperatura en l'equació del calor amb condicions de contorn Dirichlet i condueix a

$$\frac{1}{2}\int_\Omega \rho|\mathbf{u}|^2\, dx = \frac{1}{2}\int_\Omega \rho|\mathbf{u}_0|^2\, dx - \mu\int_0^t\int_\Omega |\nabla\mathbf{u}|^2\, dx\, dt'.$$

L'energia cinètica en un instant de temps positiu és equivalent a l'energia cinètica a temps inicial més les pèrdues per calor donades pel terme dissipatiu al llarg del procés, per tant:

$$\frac{d}{dt}E_{cinètica} \leqslant 0 \qquad \textit{(Navier–Stokes)} \tag{2.7}$$

En absència de forces externes.

Cal destacar que, en general, un problema de fluids presenta dues incògnites, la velocitat $\mathbf{u}$ i la pressió $p(\mathbf{x}, t)$. No obstant, en un fluid homogeni incompressible, la pressió no és una quantitat independent, sinó que es determina completament per la condició de conservació de la massa 2.3. En tals casos, la pressió serveix per imposar la divergència nul·la de la velocitat quan s'evoluciona a través de l'equació dels moments, tal i com es veurà en el mètode de Fourier-Galerkin bidimensional de la secció 3.4.

[7] Aquest problema és un dels *set Problemes del Mil·lenni* formulats per l'Institut Clay de Matemàtiques.

2.2 Número de Reynolds

L'adimensionalització d'un problema permet estudiar el comportament de les solucions de manera independent de l'escenari físic concret [11]. Aquesta tècnica consisteix en un canvi de variables que elimina les unitats físiques del sistema. En un problema on el fluid és homogeni ($\rho = \rho_0$ constant), es redefineixen els paràmetres mitjançant el canvi de variable $\nu = \frac{\mu}{\rho_0}$ i l'equació dels moments deriva a

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\mathbf{u} = -\nabla p + \nu\Delta\mathbf{u} + \mathbf{b}.$$

Quan, a més, es tracta un problema sense dimensions, com en seran els d'aquest treball, es prenen valors $(U, L, T) \in \mathbb{R}^3$ anomenats respectivament, velocitat característica, escala característica i temps característic del problema[8] i es consideren els canvis de variables

$$\mathbf{u}' = \frac{\mathbf{u}}{U}, \qquad \mathbf{x}' = \frac{\mathbf{x}}{L}, \qquad t' = \frac{t}{T}.$$

Així, s'obtenen les equacions de Navier-Stokes sense dimensions[9]

$$\begin{aligned} \frac{\partial \mathbf{u}'}{\partial t'} + (\mathbf{u}'\cdot\nabla)\mathbf{u}' &= -\nabla p + \frac{1}{\mathrm{Re}}\Delta\mathbf{u}' + \mathbf{b}\,, \\ \nabla\cdot\mathbf{u}' &= 0\,, \\ \mathbf{u}' &= 0\,, \qquad \text{a } \partial\Omega \end{aligned} \tag{2.8}$$

on el factor dividint el terme dissipatiu s'anomena nombre de Reynolds i engloba tota la informació física del sistema.

$$\mathrm{Re} = \frac{LU}{\nu} \qquad \text{Nombre de Reynolds.}$$

El comportament del fluid depèn de manera significativa del nombre de Reynolds[10]. Quan el nombre de Reynolds és elevat ($\mathrm{Re} \approx 10^4$), el terme convectiu predomina sobre el terme dissipatiu, cosa que fa que el flux presenti turbulències. En canvi, quan el nombre de Reynolds és baix, el terme dissipatiu domina i el flux tendeix a ser laminar.
És també rellevant estudiar el comportament asimptòtic de les equacions quan $\mathrm{Re} \to 0$. Aquest límit dóna lloc al fenomen conegut com a capa límit, una zona propera a la vora on els efectes viscosos són predominants i les propietats del flux canvien ràpidament. A més, quan $\mathrm{Re} \to 0$, el terme convectiu es pot negligir, i per tant, les equacions de Navier–Stokes es redueixen a un sistema simplificat conegut com a **equacions de Stokes**, que descriu el comportament d'un flux altament viscós en un règim de baix nombre de Reynolds.

2.3 Vorticitat del camp i funcions de corrent

En aquest últim apartat, introduïm nocions que permetran caracteritzar un flux homogeni incompressible. Definim la vorticitat d'un flux i les línies de corrent i ho connectem amb la noció de solució estacionària.

Definició 2.4 (Vorticitat d'un flux)**.** *Es defineix la vorticitat d'un flux* $\mathbf{u} \in \mathbb{R}^3$ *a l'espai tridimensional com*

$$\boldsymbol{\omega} := \nabla \times \mathbf{u}. \tag{2.9}$$

La definició es fàcilment extrapol·lable a dimensió 2.

[8] En un domini rectangular, la longitud característica es correspon normalment a la longitud del costat més petit del domini.

[9] Un cop realitzat el canvi de variables, és comú desempallegar-se dels símbols distintius, (per exemple, comes o barrets).

[10] El nombre de Reynolds pot variar àmpliament: al voltant de 10^{-6} per a bacteris desplaçant-se dins d'un vas de vidre, aproximadament 1 per a un peix en un aquari, 4×10^6 per a una persona nedant en una piscina i fins a 10^{12} en el cas d'un huracà. Vegeu [12].

El vector vorticitat esdevé una eina fonamental per a l'estudi de turbulències atès que permet saber com gira el fluid localment en cada punt. El següent resultat en mostra la importància per a camps incompressibles:

Teorema 2.6 (Teorema de Helmholtz, [7])**.** *Un camp de vectors* $\mathbf{w}$ *a* Ω *es pot descompondre de manera única en un camp* $\mathbf{u}$ *de divergència zero, i un camp irrotacional*

$$\mathbf{w} = \mathbf{u} + \nabla p$$

on, a més, $\mathbf{u} \cdot \mathbf{n} = 0$ *a* $\partial\Omega$*.*

Demostració. L'existència és conseqüència de la solució del problema de Neumann següent. Si suposem que $\mathbf{w} = \mathbf{u} + \nabla p$ amb $\mathbf{u}$ i ∇p tal com es descriuen a l'enunciat, aleshores $\nabla \cdot \mathbf{w} = \Delta p$ i $\mathbf{w} \cdot \mathbf{n} = \mathbf{n} \cdot \nabla p$. D'aquesta manera, donat $\mathbf{w}$ podem considerar el problema de Neumann:

$$\Delta p = \nabla \cdot \mathbf{w} \quad \text{a } \Omega, \qquad \frac{\partial p}{\partial n} = \mathbf{w} \cdot \mathbf{n} \quad \text{a } \partial\Omega.$$

Aquest problema té solució per a p llevat d'una constant. Aleshores es defineix $\mathbf{u} = \mathbf{w} - \nabla p$. Per construcció s'obté que $\nabla \cdot \mathbf{u} = \nabla \cdot \mathbf{w} - \Delta p = 0$ i, de forma similar, $\mathbf{u} \cdot \mathbf{n} = 0$ usant que $\nabla p \cdot \mathbf{n} = \frac{\partial p}{\partial n}$. El camp ∇p és irrotacional ja que és un gradient.

Per provar la unicitat, suposem que existeixen camps $\mathbf{u}_1, p_1$ i $\mathbf{u}_2, p_2$ tals que

$$\mathbf{w} = \mathbf{u}_1 + \nabla p_1,$$
$$\mathbf{w} = \mathbf{u}_2 + \nabla p_2.$$

Cal veure que $\mathbf{u}_1 = \mathbf{u}_2$ i $p_1 = p_2$. Primerament

$$0 = (\mathbf{u}_1 - \mathbf{u}_2) + \nabla(p_1 - p_2).$$

Considerem la propietat ortogonal següent

$$\int_\Omega \mathbf{u} \cdot \nabla p \, dV = 0,$$

que és certa ja que

$$\nabla(p\mathbf{u}) = (\nabla \cdot \mathbf{u})p + \mathbf{u} \cdot \nabla p,$$

i usant que $\nabla \cdot \mathbf{u} = 0$ i que $\mathbf{u} \cdot \mathbf{n} = 0$, tenim que:

$$\int_\Omega \mathbf{u} \cdot \nabla p \, dV = \int_\Omega \nabla \cdot (p\mathbf{u}) \, dV = \int_{\partial\Omega} p\mathbf{u} \cdot \mathbf{n} \, dS = 0.$$

Prenent el producte escalar amb $\mathbf{u}_1 - \mathbf{u}_2$ i aplicant la propietat d'ortogonalitat, obtenim

$$0 = \int_\Omega \{\|\mathbf{u}_1 - \mathbf{u}_2\|^2 + (\mathbf{u}_1 - \mathbf{u}_2) \cdot \nabla(p_1 - p_2)\} \, dV = \int_\Omega \|\mathbf{u}_1 - \mathbf{u}_2\|^2 \, dV.$$

Per tant, $\mathbf{u}_1 = \mathbf{u}_2$ i, aleshores, $\nabla p_1 = \nabla p_2$. □

Corol·lari 2.7. *La velocitat es pot reconstruir a partir del vector vorticitat mitjançant una equació de Poisson. Partim de*

$$\nabla \cdot \mathbf{u} = 0 \quad a\ \Omega,$$
$$\nabla \times \mathbf{u} = \omega \quad a\ \Omega,$$
$$\mathbf{u} = 0 \quad a\ \partial\Omega.$$

Com que $\nabla \cdot \mathbf{u} = 0$*, el teorema de Helmholtz assegura que existeix una funció* $\boldsymbol{\psi}$ *tal que* $\nabla \times \boldsymbol{\psi} = \mathbf{u}$*, pel nostre cas bidimensional,* $\boldsymbol{\psi} = (0, 0, \psi)$*. Fent un abús de notació, escrivim* $\nabla \times \psi = \mathbf{u}$*. Usant la*

identitat $\nabla \times (\nabla \times \boldsymbol{\psi}) = \nabla(\nabla \cdot \boldsymbol{\psi}) - \Delta \boldsymbol{\psi}$ *i assumint que el domini* Ω *és simplement connex i acotat*[11]*, el sistema anterior es redueix a*

$$\begin{aligned} -\Delta \psi &= \omega \quad a\ \Omega, \\ \psi &= const. \quad a\ \partial\Omega. \end{aligned}$$

Per tant primer es determina ψ *i posteriorment es calcula la velocitat usant* $\mathbf{u} = \nabla \times \psi$*.*

De fet, la funció ψ és sovint una eina útil en la resolució de les equacions de Navier-Stokes, i rep el nom de funció de corrent. A continuació, en donem una definició formal:

Teorema 2.8 (Funció de corrent, [7])**.** *Donat un fluid incompressible en un domini simplement connex* Ω*, aleshores existeix una única funció* $\psi(\mathbf{x}, t) \in \mathbb{R}$ *llevat d'una constant, tal que*

$$u = \partial_y \psi \quad i \quad v = -\partial_x \psi. \tag{2.10}$$

La funció ψ *s'anomena funció de corrent.*

Definició 2.5 (Línia de corrent, trajectòria d'una partícula i flux estacionari)**.** *Es defineix la trajectòria que recorre una partícula dins el fluid situada a* $\mathbf{x}_0$ *quan* $t = 0$ *com la solució del sistema dinàmic:*

$$\frac{d\mathbf{x}}{dt} = \mathbf{u}(\mathbf{x}(t), t)\,, \qquad \mathbf{x}(0) = \mathbf{x_0}.$$

Paral·lelament, definim les línies de corrent en un instant $t \geqslant 0$ *com les corbes que satisfan:*

$$\frac{d\mathbf{x}}{ds} = \mathbf{u}(\mathbf{x}(s), t). \tag{2.11}$$

Quan $\partial_t \mathbf{u} = 0$*, aleshores les línies de trajectòria i les línies de corrent coincideixen i diem que el fluid és* ***estacionari****.*

Corol·lari 2.9. *Per a fluids incompressibles les línies de corrent són les corbes de nivell de les funcions de corrent ja que, donada* $\mathbf{x}(s) = (x(s), y(s))$ *una línia de corrent, aleshores*

$$\frac{d}{ds}\psi(x(s), y(s), t) = \partial_x \psi \cdot x' + \partial_y \psi \cdot y' = -vu + uv = 0.$$

Com detallarem al Problema 3.7, algunes de les solucions per a problemes incompressibles de Navier-Stokes es poden trobar a través de l'evolució de les funcions de corrent.

[11] Aquesta hipòtesi és necessària. Sovint cal canviar les condicions de contorn en funció del domini i adaptar-les corresponentment, tal i com es detalla a [13].

3 Mètodes espectrals

Els mètodes espectrals sorgeixen a finals del segle XX com a tècnica numèrica per resoldre equacions en derivades parcials. Van ser proposats inicialment per les matemàtiques russes Blinova i Silberman (1944) en el context d'equacions de vòrtex i van ser desenvolupats posteriorment per autors com Orszag, Eliason, Rasmussen i Machenhauer. L'obra matemàticament fonamental aparegué als anys 80 amb el nom de Spectral Methods in Fluid Dynamics [5], ampliada a [6] el 2010. En aquesta secció, usem la teoria i tècniques descrites a [6] junt amb els llibres de Boyd. [4] i D. Gottlieb [14].

Un mètode espectral consisteix en comprimir la informació d'una funció en un espai de funcions, per exemple, $L^2(a,b)$ a un espai més manejable com l'espai de successions de quadrat integrable $l^2(\mathbb{C})$. Això típicament s'assoleix buscant una forma d'escriure la funció com una combinació lineal d'una base completa $\{\phi_n\}_{n\in\mathbb{N}}$ de funcions de l'espai de manera que la informació de f es col·lapsa als coeficients α_n,

$$f(x) = \sum_n \alpha_n \phi_n(x). \tag{3.1}$$

Els mètodes espectrals són de caràcter global: per calcular les derivades de f en un punt, cal conèixer tots els coeficients α_n. Això contrasta amb mètodes locals, com els d'elements finits, que només utilitzen la informació dels punts veïns. Aquesta globalitat permet obtenir una precisió molt elevada: mentre que en els mètodes d'elements finits l'error decreix algebraicament amb el nombre de punts N, en els mètodes espectrals pot disminuir més ràpidament que exponencialment. Ara bé, els mètodes espectrals resulten menys eficients en dominis complexos, ja que és difícil trobar funcions base ϕ_n que compleixin adequadament les condicions de contorn, o bé quan la funció f presenta discontinuïtats o forts gradients.

3.1 Mètode dels residus ponderats

El mètode de residus ponderats esdevé l'eina per excel·lència per a trobar els coeficients de l'expansió espectral. Es considera un problema formulat de manera general com:

Definició 3.1 (Problema de valors inicials i de contorn (IBVP))**.** *Considerem el problema valors inicials i de contorn*

$$\begin{aligned} &\frac{\partial \mathbf{u}}{\partial t} = \mathcal{M}(\mathbf{u}) \qquad \mathbf{x}\in\Omega, \quad t \geqslant 0, \\ &\mathcal{B}(\mathbf{u}) = 0 \qquad \mathbf{x}\in\partial\Omega, \quad t \geqslant 0, \\ &\mathbf{u}(\mathbf{x},0) = \mathbf{g}(\mathbf{x}) \qquad \mathbf{x}\in\Omega, \end{aligned} \tag{3.2}$$

on $\mathbf{u}(\mathbf{x},t)\in\mathbb{R}^n$ és una funció escalar o vectorial definida sobre el domini espacial $\Omega\subseteq\mathbb{R}^m$ amb frontera $\partial\Omega$, i depenent del temps $t\geqslant 0$; $\mathcal{M}$ és un operador diferencial que defineix la dinàmica del sistema, i $\mathcal{B}$ un operador lineal que imposa les condicions de contorn.

Es vol aproximar una solució del problema 3.1. S'assumeix que per a cada temps $t\geqslant 0$, la solució $\mathbf{u}(\mathbf{x},t)$ pertany a un espai de Hilbert $\mathcal{H}(\Omega)^n$ que es pot expandir sobre una base $\{\Phi_j\}_{j\in\mathbb{N}}$. Aleshores es planteja una aproximació del següent tipus

$$\mathbf{u}(\mathbf{x},t) = \sum_{n\in\mathbb{N}} \alpha_n(t)\Phi_n(\mathbf{x}) \tag{3.3}$$

on els coeficients $\alpha_n(t)$ prenen valors complexos per a cada temps i esdevenen les incògnites del mètode espectral[12]. El mètode consisteix en considerar un truncament per $N\in\mathbb{N}$ fixat

$$\mathbf{u}_N(\mathbf{x},t) = \sum_{n=0}^{N} \alpha_n(t)\Phi_n(\mathbf{x}) \tag{3.4}$$

[12]El desacoblament temporal-espacial mitjançant productes de funcions assumeix, sovint de manera implícita, una regularitat suficient tant en l'espai com en el temps. Tanmateix, aquesta hipòtesi no sempre és satisfeta, i per mantenir la igualtat formal caldria relaxar, en general, la continuïtat temporal de les funcions $\alpha_n(t)$ [14].

i s'assumeix que al ser l'ansatz de (3.3) cert per a cada temps $t \geqslant 0$, l'expressió $\mathbf{u}_N(\mathbf{x}, t)$ esdevindrà una bona aproximació de la funció $\mathbf{u}(\mathbf{x}, t)$. Per tant, estarà prop de ser solució del Problema 3.1 i es podrà definir el residu com:

$$\mathcal{R}(\mathbf{u}_N) = \frac{\partial \mathbf{u}_N}{\partial t} - \mathcal{M}(\mathbf{u}_N). \tag{3.5}$$

Els coeficients $\alpha_n(t)$ s'obtenen minimitzant (3.5) respecte a la norma de l'espai $\mathcal{H}(\Omega)^n$. Això es garanteix imposant que el residu sigui ortogonal als subespais generats per un sistema linealment independent $\{\boldsymbol{\varphi}_j\}_{j=1}^N \subset \mathcal{H}(\Omega)^n$, és a dir,

$$\langle \mathcal{R}(\mathbf{u}_N), \boldsymbol{\varphi}_j \rangle_{\mathcal{H}} = 0, \quad j = 1, \ldots, N.$$

L'elecció i nombre de funcions $\boldsymbol{\varphi}_j(x)$ determinarà el tipus de mètode espectral que s'aplica i ha d'anar en acord a minimitzar el residu a l'espai $\mathcal{H}(\Omega)^n$. En aquest sentit, els mètodes espectrals contenen dos passos essencials:

1. La tria de les funcions de l'espai d'aproximació $\operatorname{span}\langle \Phi_1, \ldots, \Phi_n \rangle$.
2. La tria de les funcions de l'espai de projecció $\operatorname{span}\langle \boldsymbol{\varphi}_1, \ldots, \boldsymbol{\varphi}_n \rangle$ on fer el residu ortogonal.

En general, s'imposen N condicions d'ortogonalitat al residu, obtenint-se un sistema diferencial de N equacions per a les incògnites $\alpha_n(t)$. La selecció de les funcions d'aproximació es pot basar, per exemple, en problemes de Sturm–Liouville, definits a l'Apèndix A.1 i en bases adequades detallades a l'Apèndix A.2. A l'espai de projecció, es busca triar funcions que simplifiquin els càlculs i siguin ortogonals. A continuació, presentem els mètodes més habituals:

1. **Mètodes de Galerkin**: Consisteixen en prendre com a funcions base, funcions Φ_n que satisfan les condicions de contorn, i pendre com a funcions test les funcions base $\boldsymbol{\varphi}_n = \Phi_n$. Una variació del mètode de Galerkin és el de Petrov-Galerkin, on la restricció $\boldsymbol{\varphi}_n = \Phi_n$ deixa d'aplicar-se i, en canvi, s'imposa $\operatorname{span}\langle \boldsymbol{\varphi}_1, \ldots, \boldsymbol{\varphi}_n \rangle = \operatorname{span}\langle \Phi_1, \ldots, \Phi_n \rangle$.
2. **Mètodes de col·locació**: Consisteixen en prendre com a funcions base, funcions Φ_n que satisfan les condicions de contorn, i prendre com a funcions test deltes de Dirac centrades en punts específics, $\boldsymbol{\varphi}_n = \delta(\mathbf{x} - \mathbf{x}_n)$. En un mètode de col·locació, s'imposa la solució exacte a certs punts.
3. **Mètode tau**: Consisteixen en prendre com a funcions base funcions que no necessàriament satisfan les condicions de contorn. Les condicions de contorn s'imposen posteriorment afegint restriccions addicionals sobre els coeficients α_n.

Formalment el mètode de residus ponderats vol solucionar el problema de condicions inicial i de contorn (3.2) en un subespai de dimensió N tal i com es detalla a [14]. Un mètode dels residus ponderats estarà ben definit si la tria de funcions base dota de convergència uniforme els coeficients α_n. En cas contrari, el residu de (3.6) pot perdre el sentit matemàtic atribuït [5].

Amb l'objectiu de modelar el problema formulat a la Secció 4, analitzem l'aplicació del mètode de residus ponderats als diferents termes de les equacions de Navier-Stokes (2.3). Desglossem l'estudi en dues parts: primer, el tractament d'un terme dissipatiu lineal, i després, el d'un terme convectiu no lineal. L'anàlisi aplicada destaca les subtileses de cada terme i estableix les eines i aproximacions per a la modelització espectral considerada a la Secció 4.

3.2 Tractament d'un terme dissipatiu lineal

El tractament lineal en mètodes espectrals és ampli; aquí ens centrem en l'equació de la calor amb condicions de contorn de tipus Dirichlet, que modela el terme difusiu de les equacions de Navier–Stokes. La metodologia es mostra mitjançant tres exemples.

Problema 3.1 (Equació de la calor). *Considerem el problema diferencial*

$$\frac{\partial u}{\partial t}(x,t) = \kappa \frac{\partial^2 u}{\partial x^2}(x,t) \qquad 0 \leqslant x \leqslant L,$$
$$u(0,t) = u(L,t) = 0,$$
$$u(x,0) = f(x),$$

on $\kappa \in \mathbb{R}$ és el coeficient de difusió.

S'estudien els mètodes de Galerkin i col·locació pel cas $\kappa = 1$ amb bases de polinomis de Fourier i Txebishev, detallats a l'Apèndix A.2.

Exemple 3.2 (Fourier-Galerkin). *Es considera l'espai $\mathcal{H} = \mathcal{L}^2(0,\pi)$ i una expansió en funcions senars*

$$u_N(x,t) = \sum_{n=1}^{N} \alpha_n(t) \sin nx.$$

El mètode de Galerkin pels residus ponderats deriva al següent sistema d'equacions:

$$\int_0^\pi \left(\frac{\partial u_N}{\partial t} - \frac{\partial^2 u_N}{\partial x^2} \right) \sin nx = 0, \qquad n \in \{1, \dots, N\}. \tag{3.6}$$

Resolent la integral anterior, s'obté el sistema d'equacions diferencials i condició inicial

$$\frac{d\alpha_n}{dt} = -n^2 \alpha_n, \qquad \alpha_n(0) = \frac{2}{\pi} \int_0^\pi u_0(x) \sin nx \; dx \qquad n = \{1, \dots, N\}.$$

El sistema anterior es resol de forma analítica d'on s'obté que

$$u_N(x,t) = \sum_{n=1}^{N} \alpha_n(t) \sin nx, \qquad \alpha_n(t) = \alpha_n(0) e^{-n^2 t}.$$

L'ordre de convergència del mètode està relacionat amb la suavitat i periodicitat de la condició inicial. En el millor dels casos, la solució pot pertànyer a l'espai de Sobolev de funcions periòdiques amb derivades infinitament periòdiques i regulars $H_p^\infty(0,\pi)$. Aleshores, els coeficients $\alpha_n(0)$ decreixen més ràpid que algebraicament, tal i com es detalla en el teorema A.2 de l'Apèndix A.3.1. En qualsevol cas, l'error d'aproximació per a temps positius decreix com $\epsilon \propto \mathcal{O}(e^{-N^2})$ a diferència d'un mètode d'elements finits on la convergència és algebraica [15]. El mètode de residus ponderats considerat reprodueix correctament la solució estàndard per una equació del calor amb condicions de contorn de Dirichlet.

Exemple 3.3 (Txebishev - Galerkin). *Es considera l'espai $\mathcal{H} = \mathcal{L}^2_\omega([-1,1])$ amb pes de Txebishev $\omega = (1-x^2)^{-1/2}$ i una expansió en polinomis de Txebishev que satisfan les condicions de contorn*

$$u^N(x,t) = \sum_{n=0}^{N} a_k(t)(1-x^2)T_k(x).$$

El mètode de Galerkin pels residus ponderats deriva al següent sistema d'equacions:

$$\int_{-1}^{1} \frac{\partial u^N(x,t)}{\partial t} (1-x^2)\, T_i(x)\, \omega(x) \; dx = \int_{-1}^{1} \frac{\partial^2 u^N(x,t)}{\partial x^2} (1-x^2)\, T_i(x)\, \omega(x)\, dx \qquad i \in \{0, \dots N\}.$$

Es defineixen les matrius $A := (A_{ij})$ i $B := (B_{ij})$ per

$$A_{ij} = \int_{-1}^{1} \omega(x)\, T_i(x)\, T_j(x)\, (1-x^2)^2\, dx \qquad B_{ij} = \int_{-1}^{1} \omega(x)\, (1-x^2)\, T_i(x)\, \frac{\partial^2}{\partial x^2} \left((1-x^2)\, T_j(x) \right) dx \tag{3.7}$$

i s'obté la següent equació diferencial pels coeficients de l'expansió:

$$A\frac{d\mathbf{a}}{dt} = B\mathbf{a} \qquad \mathbf{a}^T = [a_0, \dots, a_N]$$

Els panells 1a i 1b detallen els valors obtinguts per les matrius de (3.7). El sistema anterior no presenta grans avantatges per integració implícita respecte a explicíta atès que les matrius que se'n deriven són de tipus banda i estàtiques amb el temps. D'aquesta manera, hom pot considerar un algoritme d'inversió per A i després resoldre l'equació lineal analíticament.

Exemple 3.4 (Col·locació Txebishev [6])**.** *Es considera l'espai* $\mathcal{H} = \mathcal{L}^2_\omega([-1,1])$ *i una expansió en polinomis de Txebishev*

$$u^N(x,t) = \sum_{n=0}^{N} a_n(t)T_n(x). \tag{3.8}$$

Es trien com a funcions test les deltes de Dirac $\delta(x - x_j)$ *centrades als punts de quadratura Gauss-Lobatto,* $x_j = \cos\left(\frac{\pi j}{N}\right)$, $j \in \{0, \dots, N\}$*. El mètode dels residus imposa*

$$\frac{\partial u^N}{\partial t} - \frac{\partial^2 u^N}{\partial x^2}\bigg|_{x=x_j} = 0 \qquad j = 1, \dots, N-1 \tag{3.9}$$

S'afegeixen els punts $u^N(-1,t) = u^N(1,0) = 0$ *idènticament com a condicions de contorn. L'aproximació descrita a (3.8) convé reescriure-la com a suma de polinomi de Lagrange*

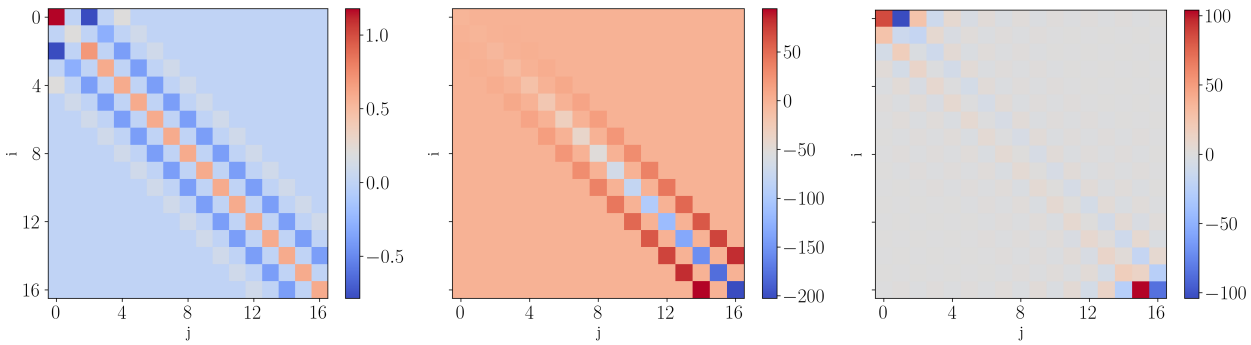
$$u^N(x,t) = \sum_{j=0}^{N} u_j(t)\mathcal{L}_j(x), \tag{3.10}$$

de manera que $u_j(t)$ *són els valors de la solució als punts de Gauss-Lobatto per a cada temps t i* $\mathcal{L}_n(x_j) = \delta_{nj}$ *els polinomis de Lagrange o deltes de Dirac discretes. Aquesta formalisme es descriu amb més detall a l'Apèndix A.4.2 i permet implementar el mètode eficientment a través de la matriu de diferenciació dels polinomis de Txebishev (D). S'observa que*

$$\frac{\partial u^N}{\partial x} = \sum_{j=0}^{N} u_j(t)\mathcal{L}'_j(x) = \sum_{j=0}^{N}\sum_{k=0}^{N} D_{jk}u_k(t)\mathcal{L}_j(x), \tag{3.11}$$

on D_{jk} *són les cel·les de la matriu D detallada a l'Apèndix A.4.2. Usant aquest formalisme, la condició de (3.9) esdevé fàcil d'imposar. El resultat es mostra a (3.12). La part dreta involucra les derivades temporals i repetint l'argument de (3.11) per escriure la segona derivada s'obté el sistema diferencial*

$$\frac{d\mathbf{u}}{dt} = D^2\mathbf{u} \qquad \mathbf{u}^T = [u_0, u_1, u_2, \cdots, u_{N-1}, u_N]. \tag{3.12}$$



(a) Matriu A (Equació 3.7) (b) Matriu B (Equació 3.7) (c) Matriu D (Equació 3.11)

Figura 1: (a) i (b): Estructura de les matrius A i B per $N = 16$ harmònics descrites a (3.7) i emprades per un mètode de Txebishev-Galerkin a l'equació del calor (Exemple 3.3). (c): Estructura de la matriu de diferenciació D de (3.11), pel mètode de col·locació de Txebishev (Exemple 3.4).

Cal imposar la condició de contorn $u_0(t) = u_N(t) = 0$ *idènticament i donar una condició inicial per a definir el sistema*[13]*. A la Figura 1c es mostra l'estructura de la matriu de derivades (D). El sistema anterior es troba resolt analíticament a [6] per una condició inicial* $u_0 = \sin(x)$ *usant funcions de Bessel.*

S'estudien computacionalment les implementacions dels mètodes anteriors amb codis detallats a l'Apèndix **??** per $-1 \leqslant x \leqslant 1$. El mètode de Fourier-Galerkin s'implementa eficientment mitjançant la transformada ràpida de Fourier i tenint en compte un mostreig de nodes suficient seguint la taxa de Nqysit, detallada a A.3.3. El cost algorítmic del mètode és de $\mathcal{O}(N \log N)$ ja que l'evolució temporal dels coeficients es deriva analíticament i només cal usar la transformada discreta de Fourier un parell de cops. El mètode de Txebishev-Galerkin presenta més irregularitats. Cal implementar una modificació de la transformada discreta del cosinus que augmenta el cost algorítmic de $\mathcal{O}(N \log N) \to \mathcal{O}(N^2)$, el major cost computacional sorgeix del pas d'invertir[14] la matriu A, amb $\mathcal{O}(N^3)$ operacions. La integració temporal es resol analíticament. Pel mètode de col·locació, els nodes es prenen en acord per aplicar els mètodes de quadratura detallats a l'Apèndix A.5.2, i simplificar els efectes del fenomen de Runge[15]. En tal cas s'aplica la matriu de diferenciació de Txebishev $\mathcal{O}(N^2)$, tot i que per $N \geqslant 128$, cal començar a considerar aplicar el procediment que minimitza en un ordre de magnitud el cost computacional, $N^2 \to (5 \log_2 N + 12)N$ detallat a l'Apèndix A.4.2. En aquests últims dos casos, l'evolució numèrica del sistema diferencial esdevé més inestable que en el cas de Fourier, atès que les funcions no són funcions pròpies ortogonals del problema de condicions inicials i de contorn.

D'altra banda, s'estudia una integració numèrica d'Euler per l'equació (3.12) i s'observa que falla per la falta de precisió a partir de $N \geqslant 10$ inclús per a passos de temps molt reduïts. En tal cas, un mètode de Crank-Nicolson per un pas suficientment petit esdevé apropiat i suficient[16]. Tal mètode s'estudiarà a la Secció 3.5.

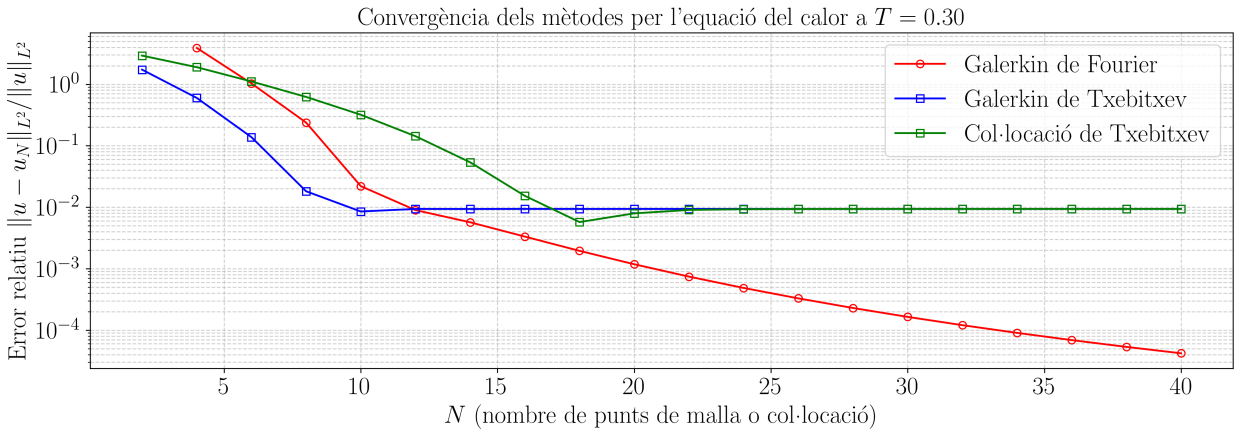


Figura 2: Simulació numèrica de l'error relatiu de la solució de l'equació del calor amb terme difusiu $\kappa = 0.1$ i a temps $T = 0.3$ per una condició inicial donada per una gaussiana amb desviació estàndard de $\sigma = 0.1$. Es simulen els exemples 3.2, 3.3 i 3.4 per a diferent nombre de modes N. S'observa que el mètode de Fourier-Galerkin, amb ordre computacional de $\mathcal{O}(N \log N)$ minimitza més ràpidament l'error respecte dels mètodes amb bases de Txebishev d'odres més elevats $\mathcal{O}(N^3), \mathcal{O}(N^2)$. Consultar codis als Apèndixs **??**

La Figura 2 detalla els errors obtinguts pels diferents mètodes usant com a condició inicial una Gaussiana centrada amb desviació estàndard de $\sigma = 0.1$ o la solució fonamental de Dirichlet per l'equació

[13] S'observa que el sistema està ben definit afegint tals condicions ja que la matriu D^2 té determinant 0.

[14] Aquest valor es pot reduir a $\mathcal{O}(16N)$ usant paquets per invertir matrius de tipus banda k

[15] El fenomen de Runge és un error d'oscil·lació que apareix en interpolacions polinòmiques quan s'usen punts equidistants, especialment quan s'utilitzen polinomis d'alt grau.

[16] Cal notar, que reduir el pas d'un mètode d'Euler no implica acostar-se arbitràriament a l'eficiència un Crank-Nicolson tal i com s'explica a l'Apèndix B.

del calor[17]. Per a $T = 0.3$, s'observa que tots els mètodes presenten un error relatiu baix prenent més de 20 nodes. El mètode de Fourier-Galerkin esdevé computacionalment molt eficient i estable. D'altra banda, els mètodes de Txebishev saturen l'error relatiu de la solució a partir de $N = 20$. Creiem que això és degut a l'ϵ de l'ordinador ja que s'observa un mateix fenomen aproximant la condició inicial. Tots els mètodes són estables amb el temps i mantenen l'error relatiu respectiu mostrat a la figura 2.

3.3 Tractament d'un terme convectiu no lineal

El tractament de problemes no lineals en mètodes espectrals depèn del cas concret i de l'espai funcional considerat. En aquest treball, s'estudia específicament el terme convectiu de les equacions de Navier–Stokes (2.3), usant bases de Fourier i Txebishev, i que, essencialment, pren la forma

$$\mathcal{F}(u) = \frac{1}{2}\frac{\partial u^2}{\partial x} = u\frac{\partial u}{\partial x}.$$

En generalitzem el cas per a dues funcions arbitràries $u(x)$ i $v(x)$. Suposem que es vol aplicar un mètode de Fourier-Galerkin a una equació diferencial on hi apareix el producte $s(x) = u(x)v(x)$. De manera que és d'interès resoldre eficientment

$$\widehat{uv}_k = \int_0^{2\pi} uv\, e^{-ikx} dx \ .$$

Quan hi ha un producte de funcions, la transformada es dona per una convolució discreta dels coeficients de les transformades $\hat{u}_k$ i $\hat{v}_k$:

$$u = \sum \hat{u}_k\, e^{ikx} \qquad v = \sum \hat{v}_k\, e^{ikx} \qquad \hat{s}_k = \widehat{(uv)}_k = \sum_{m+n=k} \hat{u}_m \hat{v}_n \tag{3.13}$$

Donat un truncament finit de N harmònics, aleshores el valor anterior esdevé en l'aproximació

$$\hat{s}_k = \sum_{\substack{m+n=k \\ |m|,|n| \leqslant N/2}} \hat{u}_m \hat{v}_n \qquad |k| \leqslant N/2. \tag{3.14}$$

L'operació analítica de (3.14) comporta un cost de $\mathcal{O}(N^2)$ operacions, en tres dimensions el cost ascendeix a $\mathcal{O}(N^4)$. En tal cas es desenvolupa la tècnica detallada a la Figura 3 seguint el mètode de la transformada i s'assoleixen costos algorítmics de $\frac{15}{2} N \log_2 N$ operacions i de l'ordre de $\mathcal{O}(N^3 \log N)$ pel cas tridimensional[18], on N denota el nombre de punts de discretització. Nogensmenys, a canvi, hi apareixen errors d'aliatge ja que la transformada ràpida de Fourier no computa exactament els coeficients sinó una versió perioditzada tal i com es detalla a l'Apèndix A.3.2. En concret s'obté

$$\tilde{s}_k = \sum_{m+n=k} \hat{u}_m \hat{v}_n + \sum_{m+n=k\pm N} \hat{u}_m \hat{v}_n = \hat{s}_k + \underbrace{\sum_{m+n=k\pm N} \hat{u}_m \hat{v}_n}_{\text{Error d'aliatge}} . \tag{3.15}$$

on $\tilde{s}_k$ és el terme obtingut usant la transformada ràpida de Fourier seguint la Figura 3, i $\hat{s}_k$ és el terme que s'obtindria analíticament. L'error restant s'anomena error d'aliatge i convé minimitzar-lo. Si no es minimitza, el mètode esdevé fortament inestable, ja que coeficients amb magnituds grans de baixes freqüències poden retroalimentar coeficients petits d'altes freqüències.

[17] Malgrat que la condició inicial no satisfà les condicions de contorn, l'error relatiu amb un gaussiana que ho fa és minim ($\approx 10^{-12}$)

[18] Per un mètode d'element finits unidimensional el cost és de $\mathcal{O}(N)$.

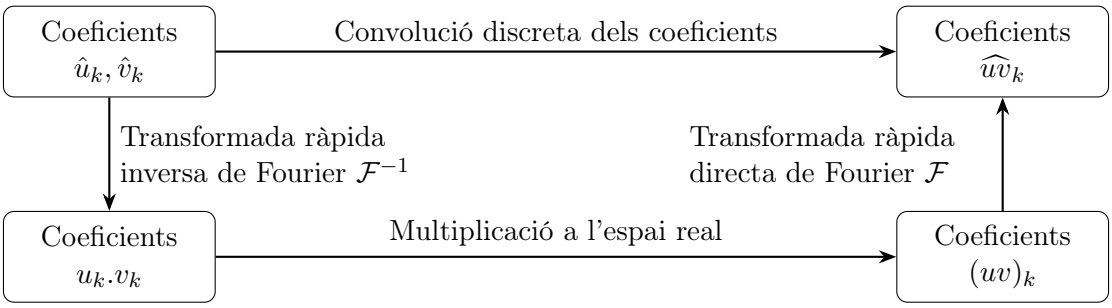


Figura 3: Diagrama del procés per avaluar la transformada d'un producte de funcions $u(x)$, $v(x)$ usant una convolució discreta i la transformada ràpida de Fourier. Usant una convolució discreta donada per (3.14) calen $\mathcal{O}(N^2)$ operacions. Usant la transformada ràpida de Fourier, l'algorisme disminueix a $\mathcal{O}(N \log_2 N)$ però introdueix aliatges.

Exemple 3.5 (Equació de Burgers. Mètode de Fourier - Galerkin)**.** *Es considera un problema de condicions inicials i de contorn donada per l'equació no viscosa de Burgers amb condicions de contorn periòdiques*

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0, \qquad u(0,t) = u(2\pi,t), \qquad u_0(x) = \sin(x).$$

Es considera l'espai $\mathcal{H} = \mathcal{L}^2(0, 2\pi)$ *i una expansió en sèries de Fourier:*

$$u^N(x,t) = \sum_{k=-N/2}^{N/2-1} \hat{u}_k(t) e^{ikx}. \tag{3.16}$$

El mètode dels residus defineix el següent sistema diferencial pels coeficients:

$$\frac{d\hat{u}_k}{dt} = -\left(\widehat{u^N \frac{\partial u^N}{\partial x}}\right)_k \qquad k = -\frac{N}{2}, \ldots, \frac{N}{2} - 1,$$

on el terme no-lineal acobla tots els modes de Fourier anteriors i representa:

$$\left(\widehat{u^N \frac{\partial u^N}{\partial x}}\right)_k = \frac{1}{2\pi}\int_0^{2\pi} u^N(x,t)\frac{\partial u^N(x,t)}{\partial x} e^{-ikx}\, dx.$$

El sistema anterior és no lineal i la solució s'integra numèricament amb condicions inicials donades per la transformada del sinus: $\hat{u}_{\pm 1}(x) = \mp i$ *i* $\hat{u}_k = 0$ *per a* $|k| \neq 1$.

Computacionalment, caldria seguir el formalisme d'una transformada discreta de Fourier, detallat a la Figura 3, seguida d'una regla de 3/2 o una matriu de filtratge, detallades en les properes Seccions 3.3.1, 3.3.2. El següent anàlisi es presenta calculant les integrals explícitament i usant (3.14) amb la finalitat de només tenir error procedent de l'integrador numèric. Aquesta pràctica és ineficient i només té sentit des d'un punt de vista teòric. A la figura 4 s'observa el comportament del mètode Fourier-Galerkin per l'equació de Burgers de l'exemple 3.5, prop del temps de xoc per a diferent nombre de nodes (N). L'anàlisi considera un mètode d'Euler d'odre 1 amb l'objectiu d'estudiar l'error numèric. Per a temps propers al temps de xoc $T \approx 1$, s'observa el fenomen de Gibbs per aproximacions de nodes baixos. D'altra banda, l'integrador numèric esdevé inestable per un nombre de nodes gran $N \approx 125$. Existeix, doncs, un balanç entre nombre d'harmònics i precisió de l'integrador. Una explicació detallada es donarà a la Secció 3.5. En mètodes espectrals, hom busca el balanç entre l'integrador numèric d'odre més baix possible i que, simultàniament, pugui capturar la finesa del nombre de nodes de Fourier que calgui representar

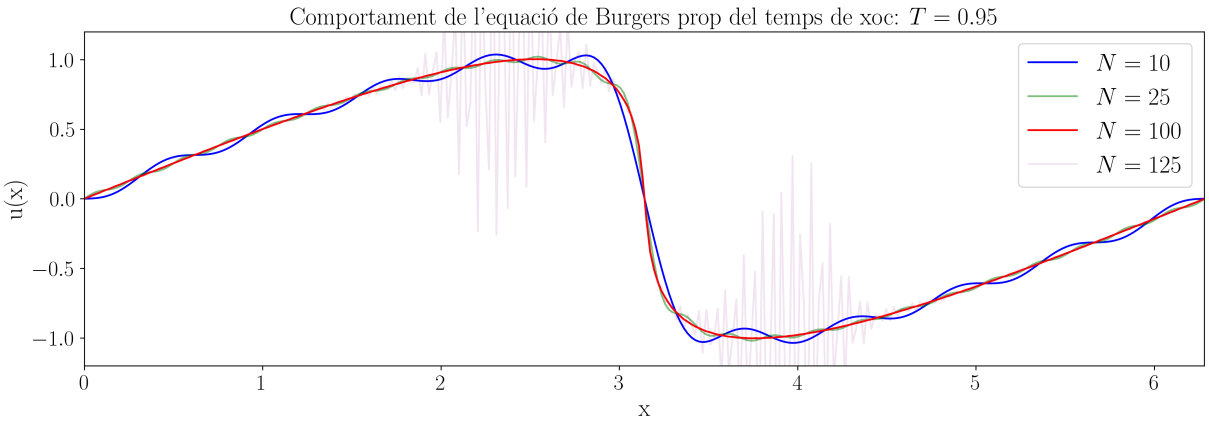


Figura 4: Simulació numèrica d'un mètode de Fourier-Galerkin per a l'equació de Burgers per a diferent nombre de nodes de Fourier prop del temps de xoc. Especificacions: Mètode d'integració d'Euler amb pas de temps de $t = 0.005$ i temps final $T = 0.95$. S'observa l'aparició del fenomen de Gibbs per a $N \leqslant 50$. Per a $N = 125$ el mètode d'integració té pas de temps massa baix. Consultar codi a l'Apèndix **??**.

3.3.1 Regla de 3/2

El tractament del terme no lineal comporta grans dificultats en la resolució d'una equació diferencial quan no se n'extreu el terme d'aliatge presentat a (3.15). En aquest treball, descrivim i implementem la regla del 3/2, tot i que també existeixen altres tècniques com les regles de desfasament, les quals no apliquem pel fet que són lleugerament més exigents computacionalment [16].

Donades dues funcions $u(x)$ i $v(x)$ i el seu producte $s(x) = u(x)v(x)$ amb N valors coneguts equiespaiats, volem calcular la transformada de Fourier de $s(x)$ a partir del valor de les transformades de $u(x)$ i $v(x)$, evitant els errors d'aliatge que s'observen fent-ho directament amb (3.15). La regla de 3/2 consisteix en:

1. S'obtenen els coeficients de Fourier $\hat{u}_k, \hat{v}_k$ de dues funcions $u(x), v(x)$, amb valors u_j, v_j coneguts en una malla uniforme $x_j = \frac{2\pi j}{N}$, mitjançant:

$$\hat{u}_k = \frac{1}{N}\sum_{j=0}^{N-1} u_j e^{-ikx_j}, \qquad \hat{v}_k = \frac{1}{N}\sum_{j=0}^{N-1} v_j e^{-ikx_j}, \qquad \text{per } k = -N/2, \dots, N/2-1.$$

2. S'amplia la informació de les coeficients de Fourier fins a $M > N$ amb valors nuls fora del rang original

$$\breve{u}_k = \begin{cases} \hat{u}_k, & |k| \leqslant N/2 \\ 0, & \text{altrament} \end{cases}, \qquad \breve{v}_k = \begin{cases} \hat{v}_k, & |k| \leqslant N/2 \\ 0, & \text{altrament} \end{cases}, \qquad |k| \leqslant M/2.$$

3. Es fan les transformades inverses per conèixer $\bar{u}_j, \bar{v}_j$ en una malla més fina $y_j = \frac{2\pi j}{M}$, i es calcula el producte punt a punt:

$$\bar{u}_j = \sum_{k=-M/2}^{M/2-1} \breve{u}_k e^{iky_j}, \qquad \bar{v}_j = \sum_{k=-M/2}^{M/2-1} \breve{v}_k e^{iky_j}, \qquad \bar{s}_j = \bar{u}_j \bar{v}_j$$

4. Es calcula la transformada directa del producte interpolat

$$\breve{s}_k = \frac{1}{M}\sum_{j=0}^{M-1} \bar{s}_j \, e^{-iky_j}.$$

5. El valor d'aquesta transformada esdevé

$$\breve{s}_k = \sum_{m+n=k} \breve{u}_m \breve{v}_n + \sum_{m+n=k\pm M} \breve{u}_m \breve{v}_n. \tag{3.17}$$

Interessa aconseguir que el segon sumand sigui idènticament nul. Per això, n'hi ha prou amb triar M de manera que, en cada cas, un dels dos coeficients del producte sigui nul i, així, la contaminació per aliatge desaparegui. Com que $\breve{u}_m = \breve{v}_m = 0$ per a $|m| > N/2$, el pitjor cas que cal considerar és

$$-\frac{N}{2} - \frac{N}{2} \leqslant \frac{N}{2} - 1 - M \quad \Rightarrow \quad M \geqslant \frac{3}{2}N - 1.$$

Eliminem el segon sumand de l'equació (3.17) i, per tant, s'elimina el terme d'aliatge de l'equació (3.15).

Aquesta tècnica permet recuperar correctament la convolució de les transformades (3.14), a canvi d'augmentar el cost computacional de $\frac{15}{2}N\log N$ a $\frac{45}{4}N\log_2\left(\frac{3}{2}N\right)$. Idealment, per a mantenir l'eficiència de les transformades de Fourier, cal prendre $M = 2^n$, conseqüentment, aquest algoritme és aproximadament un 50% més costos que l'algoritme tradicional presentat en la Figura 3 però, a canvi, s'evita completament l'error d'aliatge en el rang espectral d'interès.

3.3.2 Tècniques d'Estabilització

En l'exemple 3.5 s'ha observat que sovint cal discernir entre sacrificar resolució prenent un nombre de nodes més baix N o sacrificar l'exactitud d'algorismes d'integració numèrics. Presentem tècniques comuns que permeten prolongar l'estabilitat d'un mètode numèric espectral. El teorema de Nyquist, Apèndix. A.3.3, mostra que per a capturar les freqüències dels nodes k cal fer un mostreig de més de $N \geqslant 2k$. En fenòmens de turbulències, esdevé comú que la transformada de Fourier no capturi tots els modes de la simulació i modes de freqüències altes es barregin en freqüències baixes. Quan això passa, el terme no-lineal tendeix a amplificar aquest errors i esdevé comú aplicar una matriu de filtratge en la integració numèrica que suavitzi els modes alts de menys importància:

Definició 3.2 (Matriu de filtratge)**.** *Sigui σ una aplicació*

$$\sigma : \{-N, -N+1, \cdots, N\} \to [0, 1].$$

Diem que σ és una matriu de filtratge si: (a) satisfà que $\sigma(0) = 1$ (el mode fonamental es conserva), (b) $\sigma(k) \approx 1$ per modes propers al mode fonamental, (c) $\sigma(k) = \sigma(-k)$ (per simetria lateral) i (d) $\sigma(k) \to 0$ quan $|k| \to N/2$.

Per millorar l'estabilitat en la integració numèrica de certs tipus d'equacions diferencials, es pot aplicar la tècnica del factor d'integració. Aquesta tècnica permet reduir l'error en sistemes que combinen termes lineals i no lineals, i resulta especialment útil per a la resolució de les equacions de Navier-Stokes.

Factor d'integració Considerem una equació diferencial de la forma

$$\frac{d\hat{\mathbf{u}}}{dt} = \mathcal{N}(\hat{\mathbf{u}}) + \mathcal{L}(\hat{\mathbf{u}})$$

on $\mathcal{N}$ és un operador no lineal i $\mathcal{L}$ és un operador lineal. Sigui L la matriu que indueix l'operador $\mathcal{L}$, aleshores, s'observa la següent cadena d'igualtats:

$$\frac{d}{dt}\left(e^{-Lt}\hat{\mathbf{u}}\right) = -Le^{-Lt}\hat{\mathbf{u}} + e^{-Lt}\frac{d\hat{\mathbf{u}}}{dt} = e^{-Lt}(-L\hat{\mathbf{u}} + \frac{d\hat{\mathbf{u}}}{dt}) = e^{-Lt}N(u). \tag{3.18}$$

Això motiva a considerar un canvi de variables $\hat{\mathbf{v}} = e^{-Lt}\hat{\mathbf{u}}$ i s'arriba al sistema diferencial:

$$\frac{d\hat{\mathbf{v}}}{dt} = e^{-Lt}N(e^{Lt}\hat{\mathbf{v}}).$$

El terme lineal queda absorbit per un operador exponencial i, en conseqüència, el mètode esdevé més estable. Computacionalment, si resolem l'equació diferencial amb integradors numèric d'un pas, detallats a l'Apèndix B. Manipulació algebraica senzilla porta a l'algorisme:

$$\hat{\mathbf{u}}(t+\Delta t) \approx e^{L\Delta t}[\,\hat{\mathbf{u}}(t) + \Phi(\hat{\mathbf{u}}(t), t, \Delta t)\Delta t\,], \qquad \hat{\mathbf{u}}(0) = \hat{\mathbf{u}}_0.$$

On $\Phi(\hat{\mathbf{u}}(t), t, \Delta t)$ és la funció que resulta de l'elecció de l'integrador numèric d'un pas i que depèn de $\mathcal{N}(\hat{\mathbf{u}})$. Aquest tractament fa que el terme lineal sigui totalment estable i exacte, de manera que l'error només prové del terme no lineal. Veurem que aquesta tècnica permet ampliar el pas d'integració considerablement a través dels arguments exposats al a Secció 3.5.

Exemple 3.6 (Galerkin - Fourier per l'equació de Korteg-de-Vries)**.** *Es considera un problema de condicions inicials i de contorn donada per l'equació de Korteg-de-Vries amb condicions de contorn periòdiques:*

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + 6\nu\frac{\partial^3 u}{\partial x^3} = 0, \qquad u(0,t) = u(2\pi, t), \qquad u(x,0) = u_0(x).$$

Es considera l'espai $\mathcal{H} = \mathcal{L}^2(0, 2\pi)$ i una expansió en sèries de Fourier similar a l'exemple 3.5 per l'equació de Burgers. S'obté el següent sistema d'equacions:

$$\frac{d\hat{u}_k}{dt} = -\left(\widehat{u^N \frac{\partial u^N}{\partial x}}\right)_k + 6i\nu k^3 \hat{u}_k, \qquad k = -\frac{N}{2}, \ldots, \frac{N}{2} - 1. \tag{3.19}$$

El sistema anterior es soluciona aplicant la tècnica de factor d'integració, Secció 3.3.2 i escollint un integrador numèric adequat. S'obté el següent algorisme amb passa de temps Δt fixada[19]*.*

$$\hat{\mathbf{u}}_{i+1} = e^{-i\mathbf{k}^3\Delta t}\left(\hat{\mathbf{u}}_i + \Phi(\hat{\mathbf{u}}_i, t, \Delta t)\,\Delta t\right), \qquad \hat{\mathbf{u}}_0 = \hat{\mathbf{u}}(0). \tag{3.20}$$

Es presenta una simulació per un domini $L = 20$ i amb condició inicial donada per un solitó[20]

$$u_0 = \frac{1}{2}\left(\frac{1}{\cosh\left(\frac{1}{2}x\right)}\right)^2. \tag{3.21}$$

Atès que l'energia o massa d'un solitó roman invariant amb el temps, s'estudia l'evolució de la massa per diferents integradors numèrics. Es prenen $N = 128$ nodes de Fourier i es comparen 6 versions diferents. A totes s'usa el factor d'integració (3.20) seguit d'un Runge-Kutta que varia en ordre $1, 2, 4$. També s'afegeix la implementació de la regla de 3/2 a l'estudi. La figura 5 en detalla el resultats. S'observa l'error intrínsec que té associat cada integrador numèric mitjançant el perfil de cada recta aproximada. Els Runge-Kutta d'odres 2 i 4 mantenen satisfactòriament la massa de la solució afitada per a 2000 passos de temps amb interval $\Delta t = 0.01$. D'altra banda, s'observa que quan no s'aplica la regla de 3/2 el terme no lineal presenta errors que s'acumulen i divergeixen generalment després d'aproximadament 1000 iteracions. Per a $N = 128$ s'observa que un mètode d'Euler i regla de 3/2 no és suficient per a simular aquesta equació lineal, verificant les conclusions de l'exemple 3.5. Un integrador d'odre 2 i 4 amb la regla de 3/2 es suficient per modelitzar el problema per a temps $T \leqslant 20$.

[19] Es veurà a la Secció 3.5 que, mitjançant el factor d'integració, la condició $\Delta t \leqslant \mathcal{O}(\Delta x^3)$ es pot relaxar fins a $\Delta t < \mathcal{O}(\Delta x)$.

[20] Els solitons són solucions analítiques de l'equació de Korteg-de-Vries. Es poden descriure com petites onades o campanes que es desplacen amb el temps mantenint la forma [17].

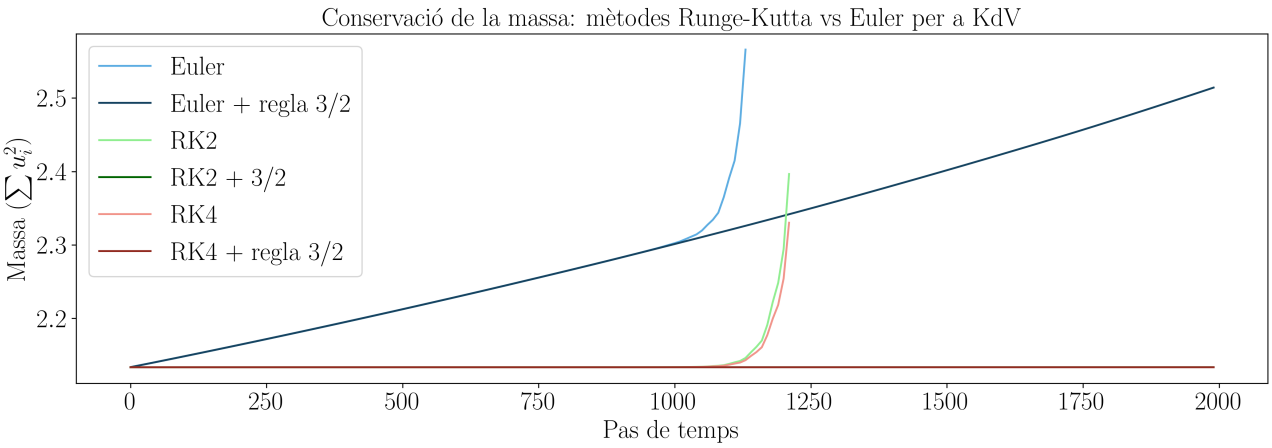


Figura 5: Simulació de l'evolució de la massa de la solució en funció del temps per un equació de Korteg-de-Vries mitjançant un mètode numèric de Fourier-Galerkin. Com a condició inicial s'usa el solitó descrit a (3.21), el pas de temps és de $t = 0.01$. S'observa com la regla de 3/2 esdevé essencial per remoure l'error d'aliatge en la integració del terme no lineal. D'altra banda, s'observa com el mètode de Runge-Kutta 2 i 4 són suficients per a simular el comportament de l'equació per a temps mínims de $T \leqslant 20$. Consultar codi **??**

.

3.4 Aplicació d'un mètode de Galerkin–Fourier a les equacions de Navier-Stokes en 2D

Modelitzem les equacions de Navier–Stokes presentades a la Secció 2 en una geometria bidimensional amb condicions de contorn periòdiques, integrant la teoria del terme dissipatiu lineal i del terme convectiu no lineal. Apliquem les tècniques desenvolupades a les Seccions 3.2 i 3.3, generalitzant-les al cas bidimensional.

Problema 3.7. *Sigui la regió $\Omega = (0, 2\pi)\times(0, 2\pi)$, i considerem el problema de contorn amb condicions inicials donat per la següent equació en derivades parcials:*

$$\begin{aligned} \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\mathbf{u} &= -\nabla p + \frac{1}{Re}\Delta\mathbf{u}, \\ \nabla\cdot\mathbf{u} &= 0, \end{aligned} \tag{3.22}$$

$$\mathbf{u}(t,0,y) = \mathbf{u}(t,2\pi,y), \quad \mathbf{u}(t,x,0) = \mathbf{u}(t,x,2\pi), \qquad \forall t \geqslant 0$$

amb condició inicial suficientment suau $\mathbf{u}(0,\mathbf{x}) = \mathbf{u}_0 \in \Omega$ i on el significat de les variables segueix la notació definida a la secció 2 pel cas d'un fluid incompressible.

Es considera l'espai $\mathcal{H} = \mathcal{L}^2(\Omega)$ i s'usa una expansió en harmònics de Fourier atès que les condicions de contorn són totalment periòdiques:

$$\mathbf{u}(\mathbf{x},t) = \sum_{\mathbf{k}} \hat{\mathbf{u}}_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{x}} = \sum_{k_x}\sum_{k_x} \begin{pmatrix} \hat{a}_{\mathbf{k}} \\ \hat{b}_{\mathbf{k}} \end{pmatrix} e^{ik_x x} e^{ik_y y}$$

on els termes $\hat{a}_{\mathbf{k}}, \hat{b}_{\mathbf{k}}$ són funcions complexes. Es desenvolupa la expressió per la pressió similarment

$$p(\mathbf{x},t) = \sum_{\mathbf{k}} \hat{p}_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{x}}.$$

S'aplica el mètode de Galerkin de residus ponderats a les equacions anteriors de forma similar als Exemples 3.5 i 3.6. L'equació diferencial que en deriva combina els estudis previs lineal i no lineal

$$\begin{aligned} \left(\frac{d}{dt} + \frac{1}{\mathrm{Re}}|\mathbf{k}|^2\right)\hat{\mathbf{u}}_{\mathbf{k}} &= -i\mathbf{k}\hat{p}_{\mathbf{k}} - \widehat{(\mathbf{u}\cdot\nabla\mathbf{u})}_{\mathbf{k}}, \\ i\mathbf{k}\cdot\hat{\mathbf{u}}_{\mathbf{k}} &= 0. \end{aligned} \tag{3.23}$$

El sistema anterior és funció de dues variables $\hat{\mathbf{u}}_k$ i $\hat{p}_k$. No obstant, la pressió en un flux incompressible està totalment determinada[21], i la podem obtenir a través de l'equació de la conservació de la massa prenent el producte escalar de $i\mathbf{k}$ amb l'equació dels moments de (3.23):

$$\hat{p}_{\mathbf{k}} = -\frac{1}{|\mathbf{k}|^2} i\mathbf{k}\cdot\hat{\mathbf{f}}_{\mathbf{k}}, \qquad \hat{\mathbf{f}}_{\mathbf{k}} = -\widehat{(\mathbf{u}\cdot\nabla\mathbf{u})}_{\mathbf{k}}.$$

Aleshores, el mètode espectral queda reduït a una equació diferencial per $\hat{\mathbf{u}}_k$[22]:

$$\left(\frac{d}{dt} + \frac{1}{\mathrm{Re}}|\mathbf{k}|^2\right)\hat{\mathbf{u}}_{\mathbf{k}} = \hat{\mathbf{f}}_{\mathbf{k}} - \mathbf{k}\frac{\mathbf{k}\cdot\hat{\mathbf{f}}_{\mathbf{k}}}{|\mathbf{k}|^2}, \qquad \hat{\mathbf{f}}_{\mathbf{k}} = -\widehat{(\mathbf{u}\cdot\nabla\mathbf{u})}_{\mathbf{k}}. \tag{3.24}$$

La integració temporal d'aquesta equació permet usar el factor d'integració detallat a la secció 3.3.2 i es complementa amb un Runge-Kutta d'odre 4. L'algorisme esdevé:

$$\begin{cases} \hat{\mathbf{a}}_{n+1} = e^{-\frac{1}{Re}|\mathbf{k}|^2\Delta t}\ (\hat{\mathbf{a}}_n + \Phi_x(\hat{\mathbf{u}}_n, t, \Delta t)\,\Delta t\ ) \\ \hat{\mathbf{b}}_{n+1} = e^{-\frac{1}{Re}|\mathbf{k}|^2\Delta t}\ \left(\hat{\mathbf{b}}_n + \Phi_y(\hat{\mathbf{u}}_n, t, \Delta t)\,\Delta t\ \right) \end{cases}$$

on Φ_x, Φ_y són les funcions de Runge-Kutta 4 obtingudes a partir de (3.24) per les dues components del camp.

Es segueixen els formalismes i els algorismes d'optimització i d'eliminació de l'alitatge comentats a les seccions anteriors. En dues dimensions la transformada de Fourier ascendeix de $15/2N\log N$ a $5N^2\log N$ operacions[23] i el terme no lineal involucra 4 operacions addicionals. Sumant-hi la regla de 3/2, el còmput total del terme no lineal requereix de $30N^2\log N$ operacions. D'altra banda, atès que l'integrador Runge-Kutta 4 programat usa 4 avaluacions de la funció, l'avanç d'un pas de temps ascendeix a aproximadament $120N^2\log N$ operacions. Finalment, tenint en compte la condició CFL d'estabilitat, $\Delta t \leqslant \frac{2.88\,\mathrm{Re}}{N}$ que detallarem a la propera Secció 3.5, s'obté que el cost algorítmic total del mètode ascendeix com $\mathcal{O}(N^3)$. El codi per simular aquest problema es troba a **??**.

Proposem el següent estudi: Considerem 3 funcions de corrent periòdiques per obtenir condicions inicials que permetin estudiar solucions analítiques de turbulència i estabilitat

$$\begin{aligned} \psi_1(x,y) &= \sin(x)\sin(y) && \text{(Taylor-Green)}, \\ \psi_2(x,y) &= \frac{\sin(2y)}{2} && \text{(Corrent de cisalla)}, \\ \psi_3(x,y) &= \sin(2x)\cos(x)\sin^2(2y). \end{aligned} \tag{3.25}$$

La definció de funció de corrent es troba a la Secció 2.3. Per una condició inicial donada per ψ_1, s'obtenen les equacions de Taylor-Green vòrtex amb solució analítica $\mathbf{u}(t) = (u(t), v(t))$ per tot temps $t \geqslant 0$ com:

$$u(x,y,t) = e^{-2\nu t}\sin x\cos y, \quad v(x,y,t) = -e^{-2\nu t}\cos x\sin y,$$

Similarment, l'equació de corrent ψ_2 porta fluxos de cisallament (Shear flows)

$$u(x,y,t) = \cos(2y)\,e^{-4\nu t}, \quad v(x,y,t) = 0 \tag{3.26}$$

S'obtenen les solucions aproximades pel mètode de Galerkin a temps $t = 0.5$, per una malla de 128 nodes i $\mathrm{Re} = 100$, s'observen errors-relatius a $L^2(\Omega)$ amb la solució analítica de $\mathcal{O}(10^{-5})$ unitats, validant el correcte funcionament del mètode.
S'estudia l'aparició de turbulències sota una petita pertubació de magnitud $\epsilon \geqslant 0$. En tal cas, s'usa un flux de cisallament lleugerament pertorbat donat per

$$\psi_{2,\epsilon} = \frac{1}{2}\left(\sin(2y) - \epsilon\sin(2x)\cos(2y)\right). \tag{3.27}$$

[21] En un flux incompressible la pressió perd el significat físic, tal i com es comenta a la Secció 2

[22] S'observa com la pressió s'usa per inserir l'equació de la massa a la l'equació de la conservació dels moments.

[23] El factor N addicional, apareix de fer una transformada per a cada columna de la matriu de mostreig.

La vorticitat del camp inicial donat per (3.27) es detalla al panell esquerra de la Figura 6 prenent $\epsilon = 0.1$ i $N = 256$. Al panell central i dret es mostren les dues evolucions de (3.27) per a diferent nombre de Reynolds, Re = 50 a Re = 5000, a temps $T = 5$. Per un flux laminar Re $\leqslant 100$, s'observa que la pertubació s'estabilitza i recupera la solució de l'equació (3.26). Per a fluxos amb Reynolds Re $\geqslant 1000$, s'observen turbulències. Les transicions a turbulències per aquest nombre de Reynolds es troben en acord al valors descrits a la literatura [18].

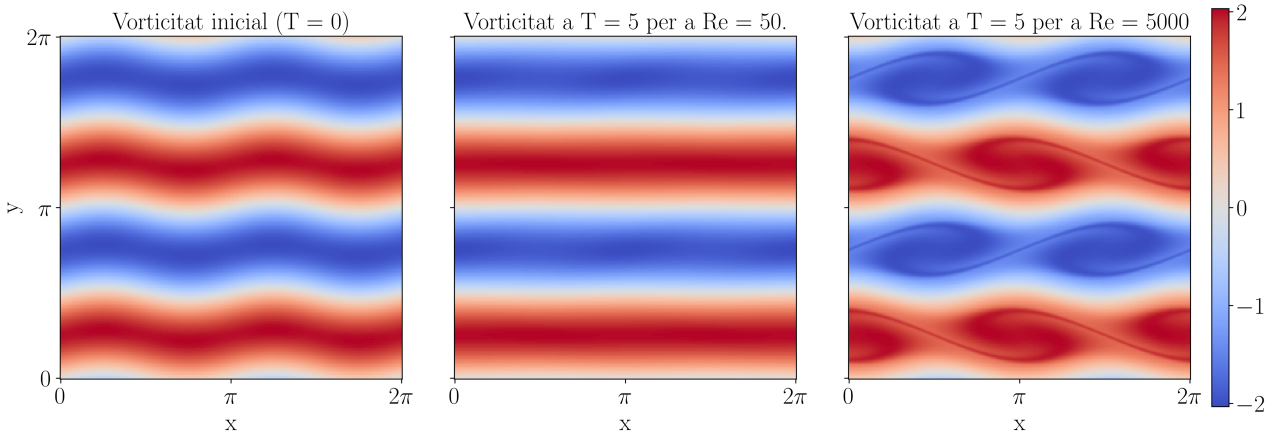


Figura 6: Evolució del vector vorticitat $\omega = \nabla \times \mathbf{u}$, per a dos fluxos amb mateixa condició inicial (panell esquerra) i diferent nombre de Reynolds fins a $T = 5$. A la imatge de l'esquerra es mostra la vorticitat inicial. A la imatge central, s'evoluciona el flux per Re = 50. A la imatge de la dreta s'evoluciona el flux amb Re = 5000. S'observen fluxos laminars i turbulents. Especificacions: Nodes de Fourier $N = 256$, pas d'integració seguint la condició de CFL, $\Delta t = 0.001$. Temps de simulació aproximat de ≈ 500 segons.

S'elaboren funcions que verifiquin numèricament la correcta implementació del mètode. En detallem:

Conservació de la massa: La divergència del camp es calcula mitjançant la transformada discreta de Fourier i, alternativament, amb extrapolació de Richardson d'ordre 2 (Apèndix A.5.1). Els errors en norma L^2 són de l'ordre de 10^{-12} amb Fourier i 10^{-5} amb Richardson, gràcies a la suavitat i periodicitat de la solució.

Conservació dels moments: Es verifica l'equació dels moments a temps $T = 5$ a partir de tres captures consecutives $\mathbf{u}_{i-1}, \mathbf{u}_i, \mathbf{u}_{i+1}$, amb aproximació temporal $\partial_t \mathbf{u}_i \approx (\mathbf{u}_{i+1} - \mathbf{u}_{i-1})/(2\Delta t)$. Les derivades espacials s'obtenen també amb Richardson. El residu resultant és de l'ordre de 10^{-5} en norma L^2.

Pas de temps: El pas de temps Δt es determina empíricament comparant l'energia respecte una referència $\Delta t = 0.001$, usant com a condició inicial la funció ψ_3 a (3.25). Per $\Delta t \leqslant 0.01$ (un ordre de magnitud superior) l'error relatiu respecte $\Delta t = 0.001$ és de $\mathcal{O}(10^{-4})$ a les primeres 10 unitats de temps; amb $\Delta t = 0.05$, es detecten desbordaments de memòria, estudiarem la raó d'aquest fet a la següent Secció 3.5 i implementarem un Runge-Kutta Felhberg a la Secció 4 per ajustar el pas de temps automàticament. Per a $\Delta t \leqslant 0.01$, l'energia total decreix monòtonament, en línia amb (2.7).

En conclusió, el mètode espectral descrit permet una aproximació acurada de les solucions de les equacions de Navier–Stokes en una geometria bidimensional amb condicions de contorn periòdiques. El cost computacional és de l'ordre $\mathcal{O}(N^3)$. Amb un pas de temps òptim $\Delta t = 0.01$ usant un RK4, la relació entre temps de simulació i temps real és d'uns 10 s per unitat temporal de simulació, per un mallat de 256×256 harmònics en un processador AMD Ryzen 5 3500U usant Python sota Windows 11.

3.5 Condició de Courant–Friedrichs–Lewy

En aquesta secció es detalla la selecció dels passos d'integració numèrica en els exemples previs. Si el mètode espectral és consistent, s'obté un sistema diferencial

$$\frac{d\mathbf{u}}{dt} = f(\mathbf{u}) \tag{3.28}$$

on f pot ser lineal o no lineal. Per a mètodes consistents i fortament estables, com els de Runge-Kutta, detallats a l'Apèndix B.1.1 i B.1.3, la qüestió principal és l'elecció del pas de temps Δt.
Considerem l'equació diferencial lineal

$$\frac{du}{dt} = \lambda u, \tag{3.29}$$

on $\lambda \in \mathbb{C}$ és el valor propi dominant d'un operador lineal o el major valor propi de la Jacobiana $(\partial f_i/\partial u_j)_{ij}$. En general, un mètode d'integració numèric es pot expressar com un sistema iteratiu, el comportament del qual depèn del valor del producte complex

$$\alpha = \lambda \Delta t,$$

i és estable si α pertany a la regió d'estabilitat $\mathcal{A}$ pròpia del mètode [16].
A la Figura 7 mostrem les regions d'estabilitat $\mathcal{A}$ dels integradors usats. Crank-Nicolson amplia notablement la regió d'estabilitat (la meitat real negativa del pla complex), permetent passos de temps més grans però a un cost computacional més elevat i implicit. Aquest canvi resol la inestabilitat observada amb Euler per $N \geqslant 10$ a l'Exemple 3.4, deguda als valors propis creixents de la matriu de diferenciació D^2. Aquest comportament es repeteix en molts problemes de difussió [5]. Així Crank-Nicolson és ideal per a l'integració d'equacions de Navier–Stokes. Els mètodes de Runge-Kutta ofereixen un bon compromís entre precisió i estabilitat sense augmentar excessivament la complexitat computacional. La seva regió d'estabilitat creix amb l'ordre, millorant la convergència respecte a l'Euler endavant[24] (FE).

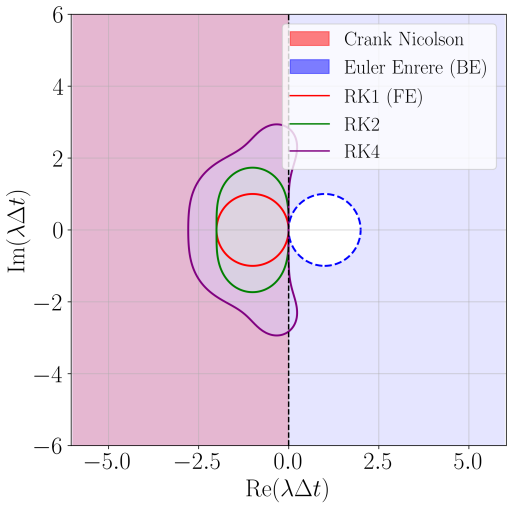


Figura 7: Regions d'estabilitat $\mathcal{A}$ al pla complex per mètodes θ d'integració numèrica. Per $\theta = 0$ sorgeix el mètode d'Euler endavant (corba vermella) amb estabilitat donada per un cercle de radi 1 centrat a $z = -1$. Per a $\theta = 1/2$ sorgeix el mètode de Crank-Nicolson amb estabilitat donada per $\mathrm{Re}(\lambda\Delta t) < 0$. Per a $\theta = 1$ sorgeix un mètode d'Euler enrere amb regió d'estabilitat fora del cercle puntejat blau de radi 1 a $z = 1$. Les zones d'estabilitat pels mètodes de Runge-Kutta d'ordres d'odres 1,2,4 es mostren mitjançant els colors vermells, verd i lila respectivament. S'observa com els Runge-Kutta's presenten un ordre de convergència més elevat que els mètodes d'Euler i Crank-Nicolson, però tenen una regió d'estabilitat més petita.

[24] Per mètodes iteratius lineals de passos múltiples, cal analitzar les arrels del polinomi característic i triar un pas de temps que les mantingui dins del cercle unitari.

Pel cas inestable, com la col·locació amb Txebishev per l'equació de la calor, Exemple 3.4, passar d'Euler explícit a Crank-Nicolson (mètode θ amb $\theta = 1/2$) millora significativament l'estabilitat. El mètode θ es defineix com

$$\mathbf{u}_{n+1} = \mathbf{u}_n + \Delta t[\theta \mathbf{f}^{n+1} + (1-\theta)\mathbf{f}^n], \qquad \mathbf{f}^n = \mathbf{f}(\mathbf{u}^n, t^n)$$

on $\theta = 0$ correspon a Euler endavant (Runge-Kutta 1) i $\theta = 1$ a Euler.

A continuació, detallem el procediment emprat per escollir els passos de temps per al Problema 3.7 de Navier-Stokes i l'Exemple 3.6 de l'equació de Korteweg-de Vries. Partim de les regions d'estabilitat mostrades a la Figura 7. Quan el sistema diferencial no s'ajusta directament a (3.29), cal ajustar i obtenir una aproximació adequada.

En el cas del mètode de Fourier-Galerkin aplicat a Navier-Stokes en 2D, Problema 3.7, es tracta d'integrar el sistema:

$$\frac{d\mathbf{u_k}}{dt} = \hat{\mathbf{f}}_\mathbf{k} - \mathbf{k}\frac{\mathbf{k}\cdot\hat{\mathbf{f}}_\mathbf{k}}{|\mathbf{k}|^2} - \frac{1}{\mathrm{Re}}|\mathbf{k}|^2\mathbf{u_k}, \qquad \hat{\mathbf{f}}_\mathbf{k} = -\widehat{(\mathbf{u}\cdot\nabla\mathbf{u})}_\mathbf{k}.$$

Observem que el terme no lineal creix com $\|\mathbf{k}\|$, mentre que el terme lineal creix com $\|\mathbf{k}\|^2$. Per tant, una cota habitual és aproximar λ pel terme dominant, és a dir, $\lambda = -\frac{N^2}{\mathrm{Re}}$, on N és el valor màxim del vector d'harmònics $\mathbf{k}$. Prenent un mètode de Runge-Kutta d'ordre 4, la condició d'estabilitat $\alpha = \lambda\Delta t \in \mathcal{A}$ implica:

$$-\frac{N^2}{\mathrm{Re}}\Delta t \in \mathcal{A} \iff \frac{N^2}{\mathrm{Re}}\Delta t \leqslant 2.88 \iff \Delta t \leqslant \frac{2.88\ \mathrm{Re}}{N^2},$$

i considerant que la transformada de Fourier pren nodes equidistants amb separació Δx en el domini $\Omega = (0, 2\pi) \times (0, 2\pi)$, obtenim:

$$\Delta t \leqslant 2.88\frac{\mathrm{Re}}{N^2} = \frac{2.88}{4\pi^2}\mathrm{Re}\,(\Delta x)^2.$$

De manera similar, per a l'equació de Korteweg-de Vries, Exemple 3.6, la fita teòrica és:

$$\Delta t \leqslant \frac{2.88}{N^3} = \frac{2.88}{8\pi^3}(\Delta x)^3.$$

Definició 3.3 (Condició CFL)**.** *Les condicions anteriors es poden generalitzar a una gran varietat de mètodes iteratius i es coneixen com a* ***Condició de Courant–Friedrichs–Lewy (CFL)*** *quan s'obtenen per a mètodes d'elements finits.*

Cal destacar que, en els nostres exemples, aquesta cota s'ha millorat mitjançant l'ús de factors d'integració que estabilitzen el terme lineal, com es detalla a la Secció 3.3.2. Això ens ha permès centrar l'estabilitat en el terme no lineal, de manera que $\lambda = \mathcal{O}(N)$. Per tant, els mètodes d'integració considerats en el Problema 3.7 i en l'Exemple 3.6 són notablement més estables i eficients amb el factor d'integració. Concretament, per a Navier-Stokes es passa de $\Delta t = \mathcal{O}(\Delta x^3)$ a $\Delta t = \mathcal{O}(\Delta x)$ i s'han pogut usar passos de temps fins a $\Delta t = 0.02$, valor proper a la cota teòrica $\Delta t \approx 0.01$ obtinguda de la CFL. De manera anàloga, per a Korteweg-de Vries s'ha escollit $\Delta t = 0.01$ per a $N = 128$, obtenint bons resultats. D'altra banda, el fet que s'observin oscil·lacions en l'Exemple 3.5 per a l'equació de Burgers amb mètode d'Euler quan $\Delta t = 0.005$, és essencialment perquè supera la cota $\frac{1}{128^2} = 6.4\times10^{-5}$ de CFL.

Tots els resultats obtinguts en els exemples validen les cotes obtingudes per les condicions de CFL.

4 Aplicació d'un mètode de Petrov–Galerkin al flux periòdic bidimensional confinat entre dues parets en moviment

En aquesta secció es desenvolupa un mètode numèric per a resoldre les equacions de Navier-Stokes per a un flux confinat entre dues parets paral·leles en moviment. Els fonaments matemàtics es basen en l'article [3], on es presenta un mètode per a tractar un flux entre parets fixes mitjançant les equacions de Navier-Stokes. En aquest treball fem ús d'aquest enfocament, però desenvolupem una estratègia pròpia per al càlcul del terme no lineal, que no es tracta a l'article, i modifiquem l'ansatz de la solució per tal de permetre la simulació de parets en moviment. El codi de la implementació i els detalls numèrics són propis i és troben a l'Apèndix ??. Utilitzem la majoria de tècniques descrites a la Secció 3.1, amb petits incisos extrets de [6].

L'element central d'estudi és un problema de contorn i condicions inicials de fluxos motivat per les equacions incompressibles de Navier-Stokes en dues dimensions. La versió que es presenta és equivalent a les equacions obtingudes a la Secció 2, però s'hi introdueix el vector vorticitat $\boldsymbol{\omega}$, que en serà una quantitat d'interès. Una derivació de l'equació (4.1) a partir de (2.6) es mostra a l'Apèndix C.1.

Problema 4.1. *Sigui un regió d'estudi bidimensional $\Omega = (0, 2\pi) \times (-1, 1)$ i considerem el problema de contorn i condicions inicials proposat per la següent equació en derivades parcials:*

$$\begin{aligned} \frac{\partial \mathbf{u}}{\partial t} &= -\nabla p - \frac{1}{Re}\nabla \times (\nabla \times \mathbf{u}) + \mathbf{u} \times \boldsymbol{\omega} + \mathbf{F}(\mathbf{x}, t), \\ \nabla \cdot \mathbf{u} &= 0, \end{aligned} \tag{4.1}$$

$$\mathbf{u}(t, 0, y) = \mathbf{u}(t, 2\pi, y)\,, \quad \mathbf{u}(t, x, \pm 1) = (\pm 1, 0), \qquad \forall t \geqslant 0,$$

on $\mathbf{u}\colon \mathbb{R}_+ \times \Omega \to \mathbb{R}^2$ és una funció vectorial amb condició inicial suficientment suau

$$\mathbf{u}(0, \mathbf{x}) = \mathbf{u}_0 \in \Omega.$$

El significat de les variables segueix la notació definida a la secció 2 pel cas d'un fluid incompressible. Aquest problema pot modelar un flux sotmès a la força de la gravetat, periòdic en l'eix de les x i confinat entre dues parets en moviments oposats a l'eix de les y, tal i com es descriu a la Figura 8.

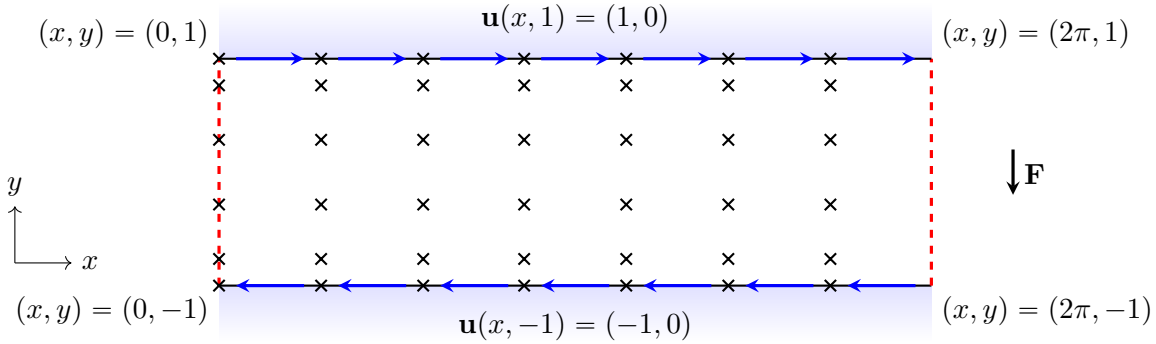


Figura 8: Representació esquemàtica del domini bidimensional $\Omega = (0, 2\pi) \times (-1, 1)$. Les línies vermelles discontínues als extrems verticals ($x = 0$ i $x = 2\pi$) indiquen condicions de contorn periòdiques $\mathbf{u}(t, 0, y) = \mathbf{u}(t, 2\pi, y)$. Les parets superior i inferior, representades amb un gradient de color, es desplacen horitzontalment a una velocitat $\mathbf{u}(x, \pm 1) = (\pm 1, 0)$ il·lustrada per les fletxes blaves. S'afegeix una força externa en direcció vertical $\mathbf{F}$. Els punts en forma de creu corresponen a un mallat espectral equiespaiat de tipus Fourier en la direcció x i Gauss-Lobatto en la direcció y, adaptat per a la resolució del terme no-lineal.

Es poden considerar mètodes espectrals de diferent naturalesa per aproximar (4.1), en podríem considerar un usant una base completa de funcions a l'espai bidimensional composta per Fourier a l'eix

de les x i polinomis complets a l'eix de les y. En tal cas, caldria calcular el valor del gradient de pressió, que es trobaria a través de l'equació de conservació de la massa i ajudaria a imposar que l'evolució temporal també tingués divergència zero, de forma similar al mètode de Fourier-Galerkin per a Navier-Stokes en dues dimensions, Secció 3.4. Aquests mètodes poden esdevenir inestables sota integració temporal i cal afegir petites correccions en cada pas de temps, tal i com descriu Canuto et al. [6].

En aquest treball, es planteja un mètode de Galerkin prenent funcions en un espai de divergència zero que permet estalviar els entrebancs comentats anteriorment. Es busca, aleshores, una combinació lineal de funcions que puguin aproximar una solució definida a l'espai

$$\mathcal{V}(\Omega) = \left\{\mathbf{u} \in L^2\big(\Omega,\, w(y)\big) \;\middle|\; \nabla\cdot\mathbf{u} = 0, \quad \mathbf{u}(0,y) = \mathbf{u}(2\pi,y), \quad \mathbf{u}(x,\pm 1) = (\pm 1, 0)\right\},$$

on

$$L^2(\Omega, w(y)) := \left\{ f : \Omega \to \mathbb{R} \;\middle|\; \int_\Omega |f(x,y)|^2\, w(y)\, dy\, dx < \infty \right\}$$

i $w(y) = \frac{1}{\sqrt{1-y^2}}$ és el pes de Txebishev.
Malauradament, $\mathcal{V}(\Omega)$ no és un espai vectorial, ja que les condicions de contorn que el defineixen no són homogènies. En aquests cas, escrivim la funció com la suma d'una funció particular que satisfà les condicions de contorn no homogènies i una combinació lineal de funcions de l'espai amb condicions de contorn homogènies.

En primer lloc, escollim una funció particular de $\mathcal{V}(\Omega)$ que compleixi les condicions no homogènies. En aquest cas, una tria natural és

$$\mathbf{u} = \begin{pmatrix} y \\ 0 \end{pmatrix}.$$

Un cop fixada la funció particular, busquem la solució general com la suma d'aquesta i una funció addicional que pertanyi a l'espai

$$\mathcal{H}(\Omega) \in \left\{\mathbf{u} \in L^2(\Omega, w(y)) : \nabla\cdot\mathbf{u} = 0, \quad \mathbf{u}(x=0,y) = \mathbf{u}(x=2\pi,y), \quad \mathbf{u}(x,y=\pm 1) = (0,0)\right\},$$

que presenta condicions de contorn homogènies i té estructura d'espai vectorial.

Donada una funció $\mathbf{u} \in \mathcal{H}(\Omega)$, aquesta es pot expressar com una sèrie en un sistema complet $\{\phi_{jk}(x,y) \mid j,k \in \mathbb{N}\}$ de $\mathcal{H}(\Omega)$, amb convergència de la sèrie en la norma $L^2(\Omega, w(y))$:

$$\mathbf{u} \overset{L^2}{=} \sum_{j,k} \begin{bmatrix} a_{jk} \\ b_{jk} \end{bmatrix} \phi_{jk}(x,y), \quad \mathbf{u} \in \mathcal{H}(\Omega), \tag{4.2}$$

on els coefficients $a_{jk}, b_{jk} \in \mathbb{C}$ són coeficients a determinar.

És d'interès escollir adequadament els valors de les funcions $\phi_{jk}(x,y)$ per tal de minimitzar els costos algorítmics del mètode de residus ponderats. En tal geometria, *Moser et al.* [3] presenta polinomis de Fourier a l'eix x i deixa una modificació de Txebishev a l'eix y com a conseqüència de les condicions de contorn. A continuació, presentem una derivació formal de les següents funcions. Una breu descripció i propietats d'aquestes funcions es troben a les Seccions A.3 i A.4.

Per a cada parell d'harmònics (j,k), prenem $\phi_{jk}(x,y) = f_j(y)e^{ikx}$, on $\{f_j(y)\}_{j\in\mathbb{N}}$ és un sistema polinòmic dens a l'espai de les funcions admissibles en la variable y. La condició de divergència zero de l'espai imposa la relació

$$\nabla\cdot\left(\begin{bmatrix} a_{jk} \\ b_{jk} \end{bmatrix} f_j(y) e^{ikx}\right) = 0$$

entre els coeficients de l'expansió, d'on s'obté la restricció

$$a_{jk} = b_{jk}\frac{i f_j'}{k f_j} \tag{4.3}$$

per $k \neq 0$. Per tant, l'expansió descrita a l'equació (4.2) esdevé

$$\mathbf{u} = \sum_{jk} b_{jk} \begin{bmatrix} \frac{i}{k} f_j' \\ f_j \end{bmatrix} e^{ikx} = \sum_{jk} \alpha_{jk} \begin{bmatrix} i f_j' \\ k f_j \end{bmatrix} e^{ikx}, \tag{4.4}$$

on $\alpha_{jk} k = b_{jk}$ i l'estudi de trobar els coeficients $(a, b)_{jk}$ es redueix a trobar els valors α_{jk}. El cas per a $k = 0$ el tractarem com un cas apart posteriorment.

Deduïm una expressió per les funcions $f_j(y)$. Les condicions de contorn imposen que $f_j(\pm 1) = 0$ per a tot $j \in \mathbb{N}$. D'altra banda, si escrivim $\mathbf{u} = (u, v)$, i recordant que el camp de velocitat és incompressible, és a dir, $\nabla \cdot \mathbf{u} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0$, es dedueix que

$$\frac{\partial u}{\partial x}(x, \pm 1) = 0 \quad \Rightarrow \quad \frac{\partial v}{\partial y}(x, \pm 1) = 0. \tag{4.5}$$

En conseqüència, les funcions $f_j(y)$ tenen una arrel doble a tots els punts del contorn. Pel teorema fonamental de descomposició de polinomis, s'obté que:

$$f_j(y) = (1 - y^2)^2 Q(y),$$

on $Q(y)$ és un polinomi de grau $\deg(Q) = j - 4$. Atès que la família de polinomis de Txebishev és completa a l'espai de funcions $L^2([-1, 1], w(y))$, el polinomi $Q(y)$ pot ser escrit com a combinació lineal de polinomis de Txebishev T_j. Així que, per simplicitat, escollim f_j com:

$$f_j = (1 - y^2)^2 T_j(y)$$

Pel cas $k = 0$, s'observa a (4.4) que la segona component del camp és idènticament nul·la. En aquest escenari, l'única condició que resta és $u(x, \pm 1) = 0$, la qual afecta exclusivament la primera component. Per satisfer aquesta condició, es proposa emprar una combinació lineal de les funcions

$$h_j = (1 - y^2) T_j(y) \tag{4.6}$$

per a la primera component[25].
En conclusió, una base de l'espai homogeni $\mathcal{H}(\Omega)$ és:

$$\Phi_{jk}(x, y) = \begin{cases} \begin{bmatrix} i f_j' \\ k f_j \end{bmatrix} e^{ikx} & \text{si} \quad k \neq 0 \\ \begin{bmatrix} h_j \\ 0 \end{bmatrix} & \text{si} \quad k = 0 \end{cases} \qquad \text{on} \quad \begin{cases} f_j = (1 - y^2)^2 T_j(y) \\ h_j = (1 - y^2) T_j(y) \end{cases} . \tag{4.7}$$

Combinant-ho amb la solució particular, es conclou que una funció $\mathbf{u}$ pertanyent a l'espai no homogeni $\mathcal{V}$ es pot escriure com a combinació lineal de funcions base ϕ_{jk} de l'espai homogeni $\mathcal{H}(\Omega)$, multiplicades per coeficients $\alpha_{jk} \in \mathbb{C}$, més una solució particular de l'espai no homogeni:

$$\mathbf{u} = \sum_{jk} \alpha_{jk} \Phi_{jk}(x, y) + \begin{pmatrix} y \\ 0 \end{pmatrix}. \tag{4.8}$$

El terme no homogeni es pot adaptar en funció de les condicions de contorn. Aquí en mostrem alguns exemples més:

[25] Una altra manera d'entendre aquesta construcció és notar que el procediment previ per $k \neq 0$ seria equivalent si a la funció u hi afegíssim un polinomi en la variable y:

$$u = \sum_{j,k} \alpha_{jk} \; i f_j'(y) \; e^{i k_x x} + \beta P(y), \quad \beta \in \mathbb{C}$$

on $P(y)$ és un polinomi arbitrari que compleix $P(\pm 1) = 0$. Així, qualsevol terme addicional ha de respectar les condicions de contorn, i l'ús de les funcions h_j garanteix aquest comportament de manera sistemàtica.

Condició de contorn	Terme no homogeni
$u(t,x,\pm 1) = \pm 1$	$\begin{pmatrix} y \\ 0 \end{pmatrix}$
$u(t,x,1) = 1 \quad u(t,x,-1) = 0$	$\begin{pmatrix} \dfrac{y+1}{2} \\ 0 \end{pmatrix}$
$u(t,x,\pm 1) = f(t)$	$\begin{pmatrix} f(t)\, y^2 \\ 0 \end{pmatrix}$

Taula 1: Condicions de contorn utilitzades per al flux confinat entre dues parets i els corresponents termes no homogenis a considerar. Cada fila mostra una configuració diferent de velocitat a les parets.

Cal destacar que, pel cas que correspon a una paret fixa i l'altra en moviment, el problema resultant correspon al cas de Taylor-Couette en el pla [19]. A la següent secció desenvolupem el mètode espectral per a trobar l'evolució dels coeficients α_{jk}.

4.1 Mètode de Petrov-Galerkin

El cas que resulta més avantatjós en un mètode espectral és aquell que les funcions base són també funcions pròpies del problema de condicions de contorn i valors inicials. No obstant, resoldre el possible problema linealitzat de Sturm-Liouville imposa condicions addicionals sobre les derivades de les funcions, cosa que fa que la convergència del mètode decreixi ràpidament [3]. Escollim el mètode de Petrov-Galerkin, tal com es presenta a l'article [3], el qual consisteix a emprar com a funcions de prova unes funcions lleugerament modificades de la base de funcions. Més endavant, es compararà l'avantatge obtingut respecte d'un mètode de Galerkin. Així, considerem les funcions de prova

$$\Psi_{jk}(x,y) = \begin{cases} \begin{bmatrix} ig_j' \\ kg_j \end{bmatrix} e^{ikx} & \text{si} \quad k \neq 0 \\ \begin{bmatrix} P_j \\ 0 \end{bmatrix} & \text{si} \quad k = 0 \end{cases} \quad \text{on} \quad \begin{cases} g_j(y) = \left(\frac{T_{l+2}(y)}{l(l+1)} - \frac{2T_l(y)}{(l+1)(l-1)} + \frac{T_{l-2}(y)}{l(l-1)} \right)/4 & l = j+2 \\ P_j = (T_{l-1} - T_{l+1})/(2l) & l = j+1 \end{cases}.$$

Aquesta base de funcions $g_j(y)$ ve motivada de resultes del mètode de Txebishev-Galerkin per una equació ordinària lineal amb condicions de contorn Dirichlet que es pot trobar a la Secció 4 de [16]. Després de manipulació algebraica, s'obtenen els polinomis anteriors $g_j(y)$. D'altra banda, al ser Ψ_{jk} una base de $\mathcal{H}(\Omega)$, les funcions anteriors presenten divergència nul·la i satisfan les condicions de contorn homogènies En tal cas, el mètode esdevé un mètode de Petrov-Galerkin seguint la terminologia de la Secció 3.1.

Es considera l'aproximació donada per un truncament fins harmònics N_x i N_y per a Fourier i Txebishev.

$$\mathbf{u}_n = \sum_{j=0}^{N_y} \sum_{k=-N_x}^{N_x} \alpha_{jk}(t) \Phi_{jk}(x,y) + \begin{pmatrix} y \\ 0 \end{pmatrix}. \tag{4.9}$$

Fem notar que aquí usem un abús de notació per descriure $k \in \{-N_x, \ldots, N_x\}$. A la realitat, el mètode s'implementa prenent exactament $2^N, N \in \mathbb{N}$ harmònics per maximitzar l'eficiència de la transformada ràpida de Fourier, per tant, el desenvolupament de Fourier no és simètric tal i com es descriu a (4.9).

Apliquem el mètode de residus ponderats a $\mathbf{u}_n$ usant cada funció test Ψ_{jk} per a cada harmònic $(j,k) \in \{0, \ldots, N_y\} \times \{-N_x, \ldots, N_x\}$. Per ara, negligim el terme de força a (4.1), el qual tractarem més endavant a la secció 4.1.2. Obtenim les condicions

$$\left(\frac{\partial \mathbf{u}_n}{\partial t}, \Psi_{jk} \right) = -\left(\nabla P, \Psi_{jk} \right) - \frac{1}{\mathrm{Re}} \left(\nabla \times (\nabla \times \mathbf{u}_n), \Psi_{jk} \right) + \left(\mathbf{u}_n \times \boldsymbol{\omega}_n, \Psi_{jk} \right). \tag{4.10}$$

Com s'esmenta a la Secció 2, el significat físic de la pressió per un problema de Navier-Stokes incompressible es perd i s'usa com a eina per imposar divergència zero a l'equació dels moments. Com que

les funcions base ja són incompressibles, la contribució per pressió és nul·la. Matemàticament, això s'observa ja que el gradient és l'operador adjunt de la divergència en l'espai de funcions $\mathcal{L}^2(\Omega, w(y))$. Per tant

$$(\nabla P, \Psi_{jk}) = -(P, \nabla \cdot \Psi_{jk}) = 0.$$

Per reescriure (4.10) de forma convenient, es defineixen les funcions

$$\boldsymbol{\xi}_i = \begin{pmatrix} ig_j' \\ kg_j \end{pmatrix} \qquad \boldsymbol{\Gamma}_j = \begin{pmatrix} if_j' \\ kf_j \end{pmatrix}$$

i s'usa l'ortogonalitat de la base de Fourier. S'obté el següent sistema d'equacions diferencials per cada mode $k \in \{-N_x, \dots, N_x\}$:

$$\sum_{j=0}^{N_y} \frac{d\alpha_{jk}}{dt} \int_{-1}^{1} \bar{\boldsymbol{\xi}}_{\boldsymbol{i}} \cdot \boldsymbol{\Gamma}_j \, w(y)\, dy = -\frac{1}{\mathrm{Re}} \sum_{j=0}^{N_y} \alpha_{jk} \int_{-1}^{1} \bar{\boldsymbol{\xi}}_i \cdot \widehat{\nabla \times \nabla \times} \boldsymbol{\Gamma}_j \, w(y)\, dy + \int_{-1}^{1} \bar{\boldsymbol{\xi}}_i \cdot \widehat{\mathbf{u}_n \times \boldsymbol{\omega}}_n \, w(y)\, dy$$

on $i \in \{0, \dots, N_y\}$. La transformada de Fourier del doble rotacional es detalla a l'Apèndix C.1 . El sistema anterior s'escriu de forma convenient per a cada harmònic $k \in \{-N_x, \dots, N_x\}$ com

$$A_{(k)} \frac{d\boldsymbol{\alpha}_k}{dt} = \frac{1}{Re} B_{(k)} \, \boldsymbol{\alpha}_k + \mathbf{F}_{(k)}(\boldsymbol{\alpha}), \qquad \boldsymbol{\alpha}_k = \begin{pmatrix} \alpha_{0k} \\ \alpha_{1k} \\ \vdots \\ \alpha_{N_y k} \end{pmatrix}, \quad \boldsymbol{F} = \begin{pmatrix} F_{0k} \\ F_{1k} \\ \vdots \\ F_{N_y k} \end{pmatrix}.$$

Aquí, les matrius $A_{(k)}$ i $B_{(k)}$ i el vector $\mathbf{F}_{(k)}$ del terme no lineal depenen de l'harmònic k. Per $k \neq 0$ fixat, cada cel·la (i, j) està definida com

$$A_{ij} = (\boldsymbol{\xi}_i, \boldsymbol{\Gamma}_j) \qquad B_{ij} = (\boldsymbol{\xi}_i, \mathcal{L}(\boldsymbol{\Gamma}_j)) \qquad F_{ik} = (\boldsymbol{\xi}_i, \widehat{\mathbf{u}_n \times \boldsymbol{\omega}}_n) \tag{4.11}$$

on definim l'operador de Laplace com

$$\mathcal{L} = \frac{d^2}{dy^2} - k^2.$$

Integrant per parts i usant les condicions de contorn, s'obtenen les expressions

$$A_{ij} = -\int_{-1}^{1} \left[\mathcal{L}(f_j)\, w(y) - f_j'\, w'(y)\right] g_i \, dy \qquad B_{ij} = \int_{-1}^{1} \mathcal{L}(f_j) \, \left[\mathcal{L}(g_i)\, w(y) \, + \, g_i'\, w'(y)\right] dy. \tag{4.12}$$

Finalment, falta tractar el cas $k = 0$, que segueix

$$A_{ij} = \int_{-1}^{1} h_j P_i \, w(y)\, dy \qquad B_{ij} = \int_{-1}^{1} \mathcal{L}(h_j)\, P_i \, w(y)\, dy. \tag{4.13}$$

La implementació numèrica d'aquestes matrius l'hem verificat de dues maneres diferents usant (4.12) i (4.11) directament. Una derivació de les expressions de (4.12) es pot consular a l'Apèndix C.2.

D'altra banda, un podria considerar un mètode de Galerkin i prendre de funcions test les funcions base. El procediment anterior desacoblaria correctament els termes en k però, malauradament, no desacoblaria totalment els harmònics j. En tal cas les matrius $A_{(k)}$ i $B_{(k)}$ en funció de cada valor k fixat esdevindrien

$$A_{ij} = (\boldsymbol{\xi}_i, \boldsymbol{\xi}_j) \ , \qquad B_{ij} = (\boldsymbol{\xi}_i, \mathcal{L}(\boldsymbol{\xi}_j)) \, .$$

La millora computacional del mètode de Petrov-Galerkin respecte al mètode de Galerkin convencional la detallem a la Figura 9, on comparem l'estructura de les matrius $A_{(1)}$ i $B_{(1)}$ per a $N_y = 16$ en ambdós mètodes. Observem que el mètode de Petrov-Galerkin genera dues matrius de tipus banda més estretes on, a més, els coeficients romanen constants o creixen molt més lentament que en el cas de Galerkin. Aquest comportament es manté per a tots els modes k. Això es deu al denominador de les funcions de prova de Petrov-Galerkin, que modera els valors alts de les derivades dels polinomis de

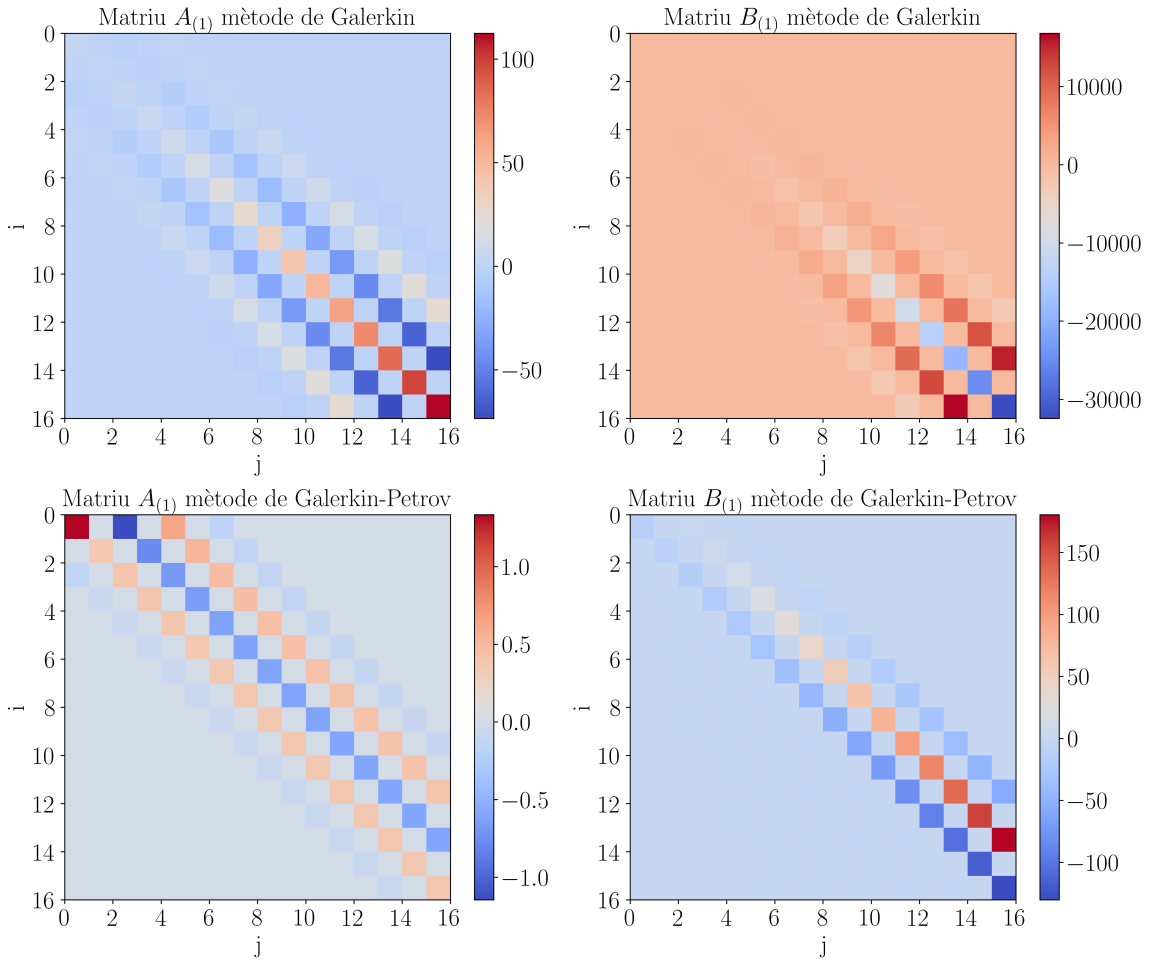


Figura 9: Matrius obtingudes mitjançant el mètode de Galerkin i el mètode de Petrov-Galerkin. En el cas del mètode de Galerkin (panells superiors), s'observa una matriu tipus banda amb vuit diagonals no nul·les i valors creixents. En canvi, amb el mètode de Petrov-Galerkin (panells inferiors) s'obtenen matrius també tipus banda, però amb cinc diagonals no nul·les i amb valors que creixen de manera més suau.

Txebishev , detallats a l'Apèndix A.4. Per tant, concloem que aquest mètode és més eficient i estable, ja que les matrius resultants són més properes a la diagonal i amb coeficients acotats.

El càlcul d'aquestes matrius cal fer-les un cop i són productes de polinomis. Per a calcular els valors, es segueix una quadratura amb nodes de Gauss-Txebishev detallats a l'Apèndix A.5.2. El nombre de nodes (N) que prenem per a cada integral ve donat per una cota superior vulgar $N = (j+4)+(j+2)$ (grau màxim del polinomi a integrar) atès que només cal calcular-les un cop. El còmput de les integrals per a $N_x = 32$, $N_y = 32$ tarda aproximadament 300 segons en un processador AMD Ryzen 5 3500U, utilitzant Python sota Windows 11.

Inversió de matrius Fem un breu incís a l'estudi de la inversió de les matrius $A_{(k)}$ el qual esdevindrà d'utilitat a les properes seccions.

En general, invertir matrius és un problema mal condicionat i potencialment inestable, especialment quan la matriu és gairebé singular. Per aquest motiu, gairebé mai es recomana aplicar la inversió directa. No obstant això, el cas que ens ocupa és de naturalesa diferent, ja que les matrius involucrades són tipus banda, com es mostra a la Figura 9. En aquests casos, existeixen mètodes d'inversió robustos, com el descrit per Emrah Kılıç et al. a [20], basats en una modificació del mètode d'inversió LU per a matrius tipus banda. En aquest treball, hem optat per realitzar un estudi senzill i eficient

que sigui factible dins del marc pràctic i temporal d'aquesta memòria.

Primer estudiem la descomposició en valors singulars, tal com es detalla a l'Apèndix C.3. Cada matriu tipus banda $A := A_{(k)}$ s'escriu mitjançant la descomposició SVD com $A = U\Sigma V^T$, on U i V són matrius unitàries i Σ és una matriu diagonal amb els valors singulars de A. Com que les matrius real unitàries tenen com a inversa la seva transposada, la inversa (o pseudo-inversa si cal) de la matriu A es pot calcular com:

$$A^{-1} = V\Sigma^{-1}U^T. \tag{4.14}$$

D'altra banda, presentem una descomposició LU, que es caracteritza per descompondre la matriu A com a producte de dues matrius L i U triangulars, una amb la diagonal i els elements sobre la diagonal diferents de zero, i l'altra amb la diagonal i els elements sota la diagonal diferents de zero. Posteriorment, es resolen n sistemes lineals, els quals són senzills de resoldre, usant, per exemple, triangulació de Gauss, tal i com es detalla als apunts de Mondelo [21].

Essencialment, el mètode de descomposició en valors singulars (SVD) és especialment útil quan la matriu està mal condicionada, ja que permet calcular-ne la pseudo-inversa de manera estable. En el nostre cas, considerem la matriu $A_{(1)}$ de dimensions 32×32 i n'estudiem primer el nombre de condició. Per a una matriu quadrada invertible A, el nombre de condició respecte a la norma 2 és:

$$\kappa(A) = \|A\|_2 \cdot \|A^{-1}\|_2 = \frac{\sigma_{\max}}{\sigma_{\min}},$$

on $\sigma_{\max}$ i $\sigma_{\min}$ són, respectivament, el valor singular màxim i mínim de la matriu A, obtinguts a partir de la matriu diagonal Σ en la seva descomposició SVD.

El nombre de condició obtingut per $A_{(1)}$ de dimensió 32 és $\kappa(A) = 281{,}369$, valor significativament superior a 1. Idealment, es vol que $\kappa(A)$ sigui proper a 1, fet que indicaria una matriu ben condicionada. Tanmateix, per $\kappa(A) \in (0, 10^8)$, el condicionament es considera acceptable dins del marge de precisió d'un ordinador estàndard. Quan $\kappa(A) \geqslant 10^8$, la inversió esdevé numèricament problemàtica i pot excedir la precisió màquina. En el nostre cas, podem concloure que la inversió de la matriu $A_{(1)}$ és acceptable. Les següents estimacions mostren el comportament dels mètodes d'inversió:

$$\text{LU:} \quad \left\| A_{(1)}^{-1}A_{(1)} - I \right\|_2 = \mathcal{O}(10^{-12}) \qquad \text{SVD:} \quad \left\| A_{(1)}^{-1}A_{(1)} - I \right\|_2 = \mathcal{O}(10^{-8}).$$

Concloem que optem per una inversió LU per les matrius ja que presenta menys error.

4.1.1 Tractament del terme no lineal

Complementem l'estudi de l'article [3] afegint el terme no-lineal convectiu que converteixen les equacions de Stokes en les equacions de Navier-Stokes. El terme no-lineal està acoblat per a diferents nombres d'ona k d'una manera semblant a com es descriu a l'Exemple 3.5 de l'equació de Burgers. Els termes d'interès són els coeficients descrits (4.11) com:

$$F_{ik} = \int_{-1}^{1} \bar{\boldsymbol{\xi}}_i \cdot \widehat{(\mathbf{u}_n \times \boldsymbol{\omega}_n)}_k \, w(y)\, dy. \tag{4.15}$$

Proposem el següent tractament per a calcular eficientment el terme no-lineal que acobla tots els harmònics. Primer es desenvolupa el producte vectorial, el procediment es detalla a l'Apèndix C.2 i s'obté

$$\widehat{\mathbf{u}_n \times \boldsymbol{\omega}_n} = \left(\widehat{v_n \frac{\partial v_n}{\partial x}} - \widehat{v_n \frac{\partial u_n}{\partial y}}, -\widehat{u_n \frac{\partial v_n}{\partial x}} + \widehat{u_n \frac{\partial u_n}{\partial y}} \right) \tag{4.16}$$

on $\mathbf{u}_n = (u_n, v_n)$. Cada terme de l'expressió (4.16) s'ha avaluat seguint el formalisme per avaluar ràpidament derivades usant nodes de Fourier i Txebishev. Primer es pren un mallat de punts del domini Ω on s'evalua la funció $\mathbf{u}_n$. Els punts es prenen equidistants en l'eix x per la FFT i seguint els nodes de Gauss-Lobatto en la direcció y,

$$(x,y) = \left(\frac{\pi j}{N_x}, \cos\left(\frac{\pi i}{N_y}\right)\right) \quad (i,j) \in \{0,\dots,2N_x-1\} \times \{0,\dots,N_y\}, \tag{4.17}$$

tal i com es descriu a la Figura 8. Posteriorment, per les derivades a l'eix de les x s'usa la transformada ràpida de Fourier per multiplicar pel vector d'harmònics $\mathbf{k}$. El procediment es detalla a (4.18). D'altra banda, per efectuar les derivades a l'eix de les y, es multiplica per la matriu de diferenciació, tal i com es detalla a (4.19). Aquests mètodes es troben explicats en detall als Apèndixs A.3.4 i A.4.2 respectivament.

$$v\frac{\partial v}{\partial x} = v\mathcal{F}^{-1}(i\mathbf{k}\mathcal{F}(v)) \ , \qquad u\frac{\partial u}{\partial x} = u\mathcal{F}^{-1}(i\ \mathbf{k}\ \mathcal{F}(v)). \tag{4.18}$$

$$v\frac{\partial u}{\partial y} = v\mathcal{D}u \qquad u\frac{\partial u}{\partial y} = u\mathcal{D}u. \tag{4.19}$$

El resultat que s'obté cal aplicar-hi la transformada de Fourier tal i com marca (4.16), en tal cas, s'implementa la regla de 3/2 per remoure l'error d'aliatge, detallada a la Secció 3.3.1. Finalment, al mètode també s'hi aplica una matriu de filtre que arrodoneix valors petits a zero, Secció 3.3.2, i millorar l'estabilitat en els valors de α_{jk} per a fluxos turbulents.

El resultat de tot aquest procés és una matriu amb els valors de (4.16) seguint el mallat de punts (4.17), que esdevé perfecte per aplicar un mètode de quadratura de Gauss-Lobatto

$$F_{jk} = \int_{-1}^{1} \bar{\boldsymbol{\xi}}_j \cdot \widehat{(\mathbf{u}_n \times \boldsymbol{\omega}_n)}_k \, w(y)\, dy = \sum_i \bar{\boldsymbol{\xi}}_j(y_i) \cdot \widehat{(\mathbf{u}_n \times \boldsymbol{\omega}_n)}_k(y_i)\, w(y_i)\, dy \tag{4.20}$$

detallat a l'Apèndix A.5.2. Atès que es vol capturar la màxima subtilesa en tals integracions, es pren un nombre de nodes elevat, que es detallarà a les properes seccions.

Observació 4.2. *Una bona comprovació és veure que la solució* $\mathbf{u} = (y, 0)$ *és estacionària. Si descomponem* $\mathbf{u}_n = \begin{pmatrix} u_n \\ v_n \end{pmatrix} + \begin{pmatrix} y \\ 0 \end{pmatrix}$ *on* $\begin{pmatrix} u_n \\ v_n \end{pmatrix}$ *és la part homogènia i* $\begin{pmatrix} y \\ 0 \end{pmatrix}$ *és la part inhomogènia obtenim*

$$\widehat{\mathbf{u}_n \times \boldsymbol{\omega}_n} = \left(\widehat{v_n\frac{\partial v_n}{\partial x}} - \widehat{v_n\frac{\partial u_n}{\partial y}}, -\widehat{u_n\frac{\partial v_n}{\partial x}} + \widehat{u_n\frac{\partial u_n}{\partial y}}\right) + \left(-\widehat{v_n}, -\widehat{y\frac{\partial v_n}{\partial x}} + \widehat{y\frac{\partial u_n}{\partial y}} + \widehat{u_n + y}\right). \tag{4.21}$$

Quan $u_n = v_n = 0$, veiem que (4.21) esdevé només una constant a la segona component del segon sumant, conseqüència de la transformada de Fourier a l'eix x, i desapareix posteriorment sota el signe integral a (4.15) per l'ortogonalitat dels polinomis de Txebishev.

4.1.2 Tractament per una força externa

Mostrem com es pot implementar un vector de força extern, el qual pot variar en el temps[26]. Per tal d'il·lustrar la seva introducció dins el mètode espectral, considerem un vector força constant

$$\mathbf{F} = \rho \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$$

on $\theta \in (0, 2\pi)$ és l'angle que determina la direcció del vector respecte a l'eix horitzontal, i ρ representa una densitat de força a cada punt. Aquest vector pot modelar, per exemple, l'efecte de la gravetat inclinada sobre un canal bidimensional.

[26] Tot i que aquest plantejament és relativament senzill des d'un punt de vista matemàtic, esdevé rellevant en simulacions físiques, ja que permet incorporar efectes com la gravetat, acceleracions externes o camps aplicats.

Quan s'introdueix aquest terme dins el mètode espectral, observem que només contribueix a l'harmònic fonamental de Fourier $k = 0$. En aquest cas, la força es projecta sobre els polinomis base mitjançant

$$F_{j0} = \int_{-1}^{1} P_j(y)\, \widehat{\rho \cos\theta}_0 \; dy.$$

Això implica que el mètode només és sensible a la component horitzontal de la força, que és proporcional a $\cos\theta$. En particular, si s'intenta simular una força purament vertical, el mètode espectral no la detecta, cosa coherent amb el fet que el fluid és incompressible i es troba confinat en una regió plana i tancada. En aquest context, la gravetat no produeix cap efecte net sobre el sistema.

4.1.3 Esquema d'integració temporal

S'arriba a un sistema de $N = (2N_x + 1)(N_y + 1)$ equacions reunides en paquets a través del nombre d'ona $k \in \{-N_x, \ldots, N_x\}$ acoplades a través del terme no-lineal que s'escriu dins $\mathbf{F}_{(k)}(\boldsymbol{\alpha})$ junt a forces externes:

$$A_{(k)} \frac{d\boldsymbol{\alpha}_k}{dt} = \frac{1}{Re} B_{(k)} \; \boldsymbol{\alpha}_k + \mathbf{F}_{(k)}(\boldsymbol{\alpha}). \tag{4.22}$$

El fet que $\mathbf{F}_{(k)}(\boldsymbol{\alpha})$ acopli tots els modes fa que la l'equació anterior no es pugui paral·lelitzar a diferència del cas mostrat a [3] per les equacions de Stokes. En tal cas, esdevé essencial calcular el terme no-lineal amb els valors del pas anterior.

Es poden considerar mètodes implícits o explícits per resoldre numèricament el sistema donat a (4.22). El fet que la família de matrius $A_{(k)}$ siguin de tipus banda, permet invertir bé el sistema i fer-lo explícit, tal i com s'ha comentat a 4.1. S'obté

$$\begin{cases} \dfrac{d\boldsymbol{\alpha}_{-N_x}}{dt} & = \dfrac{1}{Re} \bar{A}_{(-N_x)} \boldsymbol{\alpha}_{-N_x} & + \bar{\mathbf{F}}_{(-N_x)}(\boldsymbol{\alpha}) \\ \vdots & & \\ \dfrac{d\boldsymbol{\alpha}_k}{dt} & = \dfrac{1}{Re} \bar{A}_{(k)} \boldsymbol{\alpha}_k & + \bar{\mathbf{F}}_{(k)}(\boldsymbol{\alpha}) \\ \vdots & & \\ \dfrac{d\boldsymbol{\alpha}_{N_x}}{dt} & = \dfrac{1}{Re} \bar{A}_{(N_x)} \boldsymbol{\alpha}_{N_x} & + \bar{\mathbf{F}}_{(N_x)}(\boldsymbol{\alpha}) \end{cases} \tag{4.23}$$

on $\bar{A}_{(k)} = A_{(k)}^{-1} B_{(k)}$ i $\bar{\mathbf{F}}_{(k)} = A_{(k)}^{-1} \mathbf{F}_{(k)}$.

Motivat pels exemples estudiats anteriorment, proposem una millora de l'integrador numèric emprant el mètode de Runge-Kutta-Fehlberg d'ordres 4 i 5. A diferència de les implementacions prèvies, aquest integrador adapta el pas de temps en funció de l'error associat a la diferència entre les solucions obtingudes pels integradors d'ordres diferents.

Presentem breument la forma com s'ajusta el pas de temps; una descripció detallada del funcionament del mètode es pot consultar a l'Apèndix B.1.2. Considerem dos esquemes iteratius corresponents a integradors de Runge-Kutta d'ordres 4 i 5, respectivament,

$$\begin{aligned} \boldsymbol{\alpha}_{i+1} &= \boldsymbol{\alpha}_i + h\, \Phi(t_i, \boldsymbol{\alpha}_i, h), \\ \tilde{\boldsymbol{\alpha}}_{i+1} &= \tilde{\boldsymbol{\alpha}}_i + h\, \tilde{\Phi}(t_i, \tilde{\boldsymbol{\alpha}}_i, h), \end{aligned}$$

on $i \geqslant 0$.

Els arguments d'ordre de convergència condueixen a la definició del següent factor $q \in \mathbb{R}$, emprat per ajustar el pas de temps de h a qh, en funció d'una tolerància d'error prescrita $\varepsilon > 0$ entre els dos mètodes,

$$q \leqslant 0.84 \left(\frac{\varepsilon h}{|\tilde{\boldsymbol{\alpha}}_{i+1} - \boldsymbol{\alpha}_{i+1}|} \right)^{1/4}. \tag{4.24}$$

La implementació s'efectua utilitzant un paquet d'integradors numèrics, amb una tolerància escollida de $\varepsilon \approx 10^{-9}$ i es fan proves del mètode descrit anteriorment per a diferents valors del nombre de nodes N_x i N_y.

4.2 Simulació numèrica

Fem un breu incís a la implementació numèrica del mètode. El lector interessat pot consultar el codi a l'Apèndix **??**. *Algunes observacions i detalls interessants sobre convenis de programació i detalls de precisió de l'ordinador s'han recollit i comentat a l'Apèndix C.4.*

Resumim el mètode numèric de Petrov-Galerkin per a resoldre el flux bidimensional confinat entre dues parets en moviment a la Figura 10. Computacionalment es guarden els valors de α_{jk} en una matriu de manera que la majoria de càlculs es poden fer mitjançant multiplicació de matrius. Hem verificat el mètode pels diferents passos. Pel terme lineal, les matrius de tipus banda $A_{(k)}$ i $B_{(k)}$ s'han calculat de dues formes diferents, usant (4.11) i (4.12). Similarment, pel que fa el terme no lineal, la transformació $\boldsymbol{\alpha} \to \mathbf{u}$ s'ha obtingut a través de multiplicar per matrius $\mathcal{A}_{(k)}$ en la variable u i $\mathcal{B}_{(k)}$ en la variable v i usant la transformada de Fourier. El resultat s'ha verificat comparant-ho punt a punt usant directament de l'expressió de (4.8). Per les matrius de transformació $\mathcal{A}_{(k)}, \mathcal{B}_{(k)}$, hem considerat el doble de nodes en y per a maximitzar la precisió del càlcul de terme no-lineal en la integració de Gauss-Lobatto. D'altra banda, el plantejament de les derivades a (4.19) s'ha verificat per a diverses funcions i la integració numèrica de (4.20) ha estat verificada analíticament pels casos més simples. Pel terme de força, hem decidit fer servir una integració seguint els nodes de Txebishev-Gauss, a diferència dels de Gauss-Lobatto, ja que en milloren la precisió lleugerament i, en tal cas, no cal lligar-ho amb la matriu de diferenciació, la qual es calcula en els nodes de Gauss-Lobatto. El sistema es integrat usant l'integrador de Runge-Kutta Felhberg 45 amb tolerància $\varepsilon = 10^{-9}$, i els paràmetres de sortida es transformen de tornada a l'espai físic.

Computacionalment, s'observa que el terme no lineal és el principal factor que limita l'eficiència del mètode, especialment, la transformació dels coeficients espectrals α_{jk} als valors de l'espai físic $\mathbf{u}(x,t)$ al llarg de l'eix y. En aquest procés, cal resoldre un sistema lineal per a cada harmònic k, cosa que implica un total de N matrius per realitzar el canvi de base complet, amb un cost computacional de l'ordre de $\mathcal{O}(N^3)$ operacions. D'altra banda, els passos intermedis impliquen càlculs de termes no lineals en dues dimensions, similars als del Problema 3.7 d'un flux bidimensional periòdic, amb un cost aproximat de $120N^2 \log N$ operacions, considerant la regla de 3/2 detallada a la Secció 3.3.1. A més, a cada pas de temps cal aplicar la matriu de diferenciació de Txebishev a l'eix y a l'equació (4.19) la qual té un cost de $2N^2$ operacions. Finalment, el pas d'integració no lineal, detallat a l'equació (4.20), també requereix un cost de l'ordre de de $\mathcal{O}(N^2)$ operacions. En conjunt, el cost total per calcular el terme no lineal ascendeix, aleshores, a aproximadament $N^3 + 120N^2 \log N$ operacions.

A cada iteració són necessàries 6 avaluacions de la funció no lineal. Considerant la condició de CFL, detalla a la Secció 3.5, el pas de temps hauria de ser de l'ordre de $\Delta t = \mathcal{O}(N^{-2})$. Això condueix a un cost algorítmic global per simulació de l'ordre de $\mathcal{O}(N^5)$ operacions[27].

[27] L'algorisme ha estat optimitzat pre-computant i emmagatzemant totes les matrius de canvi de base i del terme lineal en subfitxers recarregables.

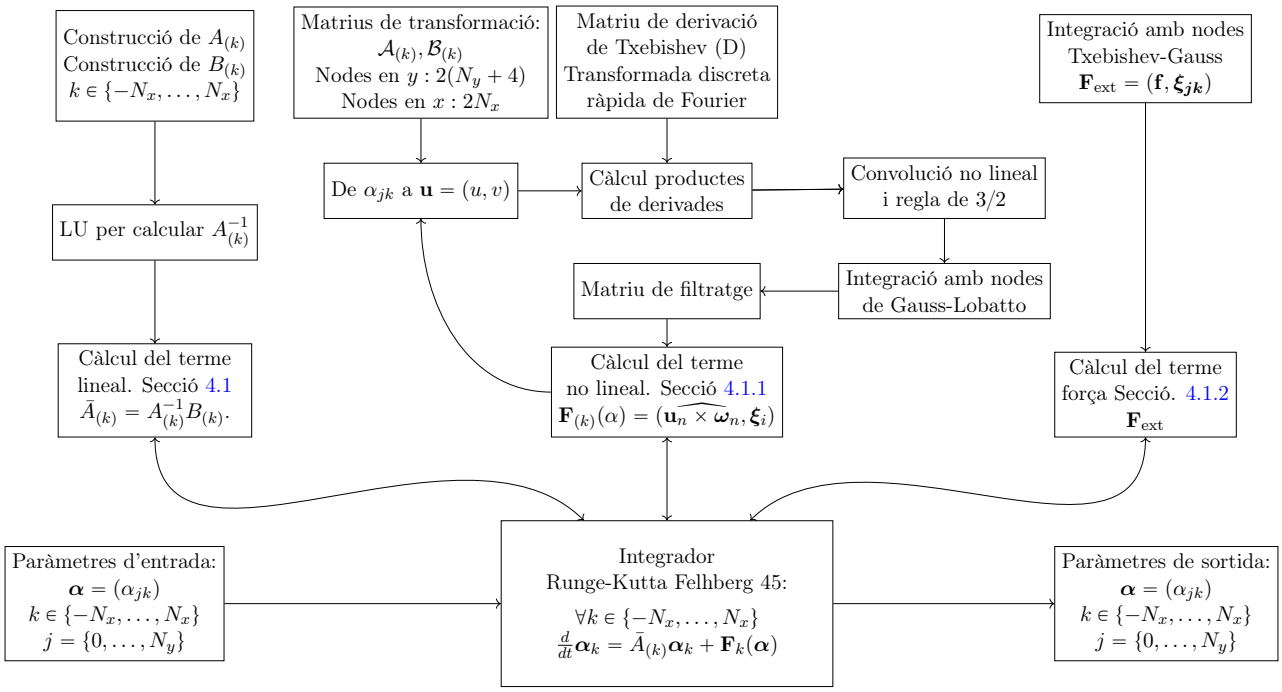


Figura 10: Diagrama del mètode numèric de Petrov-Galerkin implementat al flux periòdic bidimensional confinat entre dues parets en moviment. Es treballa amb els coeficients α_{jk} organitzats en forma matricial per optimitzar els càlculs mitjançant multiplicació de matrius. S'han verificat els components del mètode: els termes lineals amb dues expressions diferents per les matrius $A_{(k)}$ i $B_{(k)}$, els termes no lineals mitjançant transformacions espectrals i derivades calculades amb matrius i FFT, i el terme forçant amb integració sobre nodes de Gauss-Txebishev. La integració temporal es realitza amb l'integrador de Runge-Kutta-Fehlberg 45 amb tolerància $\varepsilon = 10^{-9}$ i els resultats es transformen novament a l'espai físic. El cost total del mètode ascendeix com $\mathcal{O}(N^5)$.

4.3 Anàlisi de resultats

Estudiem l'estabilitat i la coherència del mètode numèric. En primer lloc, verifiquem-ne la qualitat en absència del terme convectiu no lineal ($\mathbf{u} \times \boldsymbol{\omega}$) present a les equacions (4.1) que defineixen el problema. En aquest cas, es simulen les equacions de Stokes, que es comporten de manera similar a una equació del calor amb condicions de contorn fixades i evolució amb divergència nul·la. Es consideren diverses condicions inicials consistents a activar un harmònic particular, assignant $\alpha_{jk} = 1$ i fixant la resta de coeficients a zero. Observem com la solució decau progressivament fins a estabilitzar-se al voltant de la solució particular de l'expansió espectral. Aquest comportament és esperable, ja que el terme difusiu extreu energia del sistema d'acord amb la condició (2.7), fins a assolir un estat estacionari al voltant de la solució particular.

D'altra banda, es verifica que, en afegir el terme no lineal, els fluxos perfectament laminars, que són aquells formats únicament pels harmònics $k = 0$, romanen laminars en el temps i evolucionen cap a la solució estacionària corresponent al terme no homogeni. Aquest comportament és coherent, ja que en fluxos laminars el terme convectiu no lineal és idènticament nul, i, per tant, la dinàmica es desacobla en els harmònics de Fourier, comportant-se de manera anàloga a les equacions de Stokes. En conseqüència, l'evolució del flux queda governada exclusivament pel terme dissipatiu parabòlic, el qual tendeix a estabilitzar la solució al voltant de $\mathbf{u} = (y, 0)$, que correspon a la solució particular imposada pel terme no homogeni. La Figura 11 mostra l'evolució de l'energia cinètica d'aquests fluxos laminars al llarg del temps per a diferents modes inicials. S'hi observa com les parets en moviment tendeixen a estabilitzar el flux cap a la solució estacionària, cedint o absorbint energia segons la condició inicial.

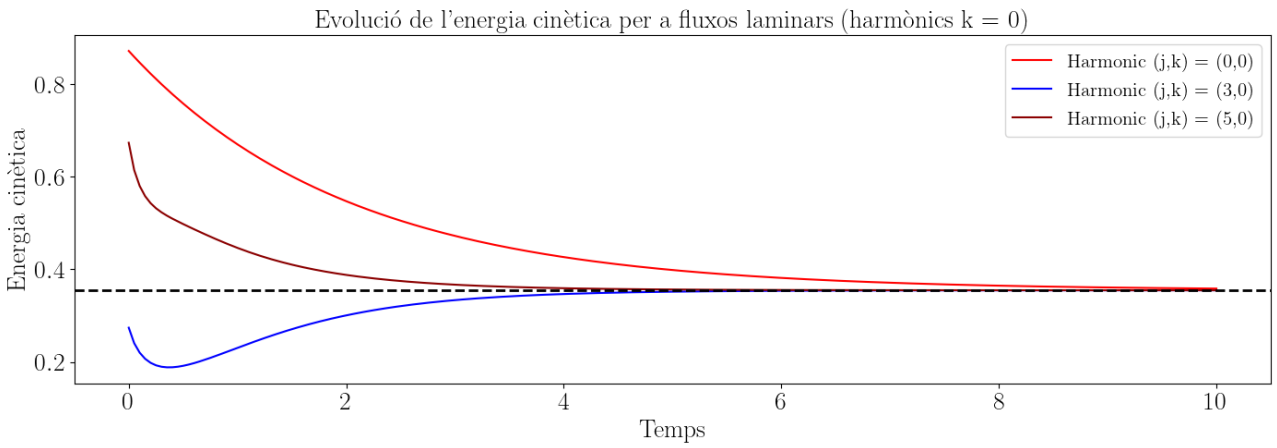


Figura 11: Evolució de l'energia cinètica per a fluxos laminars en funció del temps per a $Re = 100$. Línia negra correspon a l'energia cinètica de la solució $\mathbf{u} = (y, 0)$. S'observa com l'evolució de diferents condicions inicials sempre s'estabilitza al voltant de l'energia. Aquest fenòmen és independent del nombre de Reynolds considerat.

Posteriorment, desenvolupem l'anàlisi sobre un condició inicial lleugerament pertorbada per observar la sensibilitat del mètode al nombre de Reynolds. Atès que les parets es mouen en sentit contrari, esperem observar remolins per a fluxos suficientment turbulents. Tal com s'observa a la Figura 12, hem vist que aquest n'és el cas a temps $T = 10$ per condicions inicials amb petites pertubacions. Donada una condició inicial lleugerament pertorbada, el flux esdevé laminar per un nombre de Reynolds baix de Re = 100, d'altra banda, quan s'incrementa el nombre de Reynolds a prop de Re = 1000 es comencen a observar turbulències i per a Re = 10000 el flux es molt turbulent. Matemàticament, aquest fet s'atribueix a l'activació d'harmònics d'odre superior.

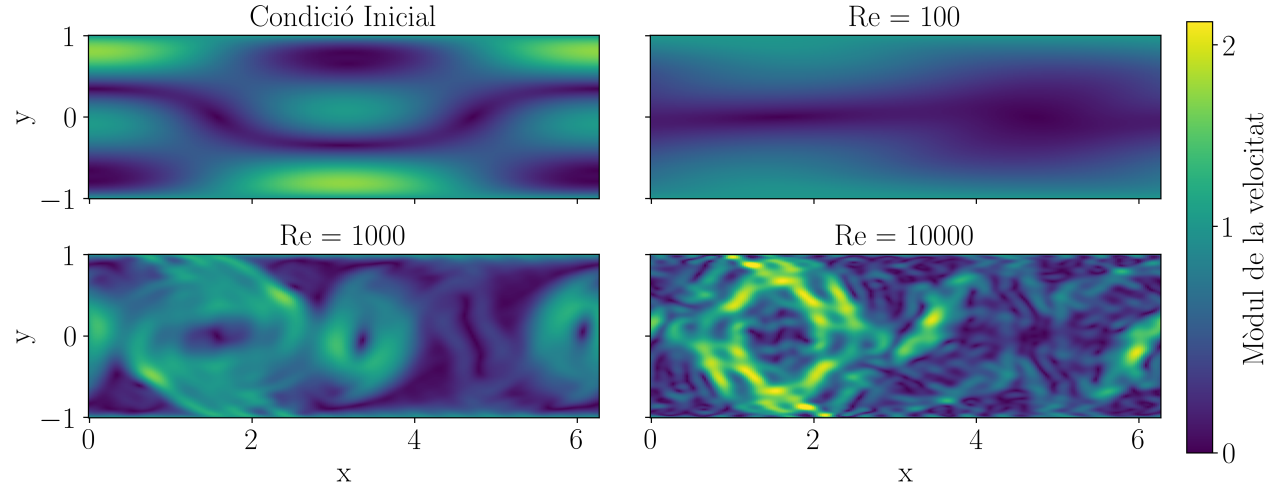


Figura 12: Evolució del mòdul de la velocitat donada una condició inicial (imatge superior esquerra) dotada d'un petita pertubació per a diferents nombre de Reynolds a temps $T = 10$. Per a Re = 100 (imatge superior dreta) s'observa com el flux esdevé poc turbulent i s'estabilitza, amb valors propers al terme no homogeni. Quan el nombre de Reynolds augmenta a Re = 1000 (imatge inferior esquerra) s'observa un flux lleugerament turbulent. El resultat s'amplifica considerablement al considerar Re = 10000 (imatge inferior dreta). En tots casos, l'energia del sistema roman afitada i estable. Simulació numèrica prenent $N_x = N_y = 32$.

El fenomen de turbulències i diagrames de transició obtinguts estan en acord amb altres estudis de caràcter semblant, on es manifesten turbulències per a nombres de Reynolds superiors a Re $\geqslant$ 500 [22].

Observem, a més, que les turbulències es mantenen en el temps per a valors elevats del nombre de Rey-

nolds, sense decaure, tal com s'espera en situacions reals. Aquest comportament contrasta amb el que s'obté a la simulació de les equacions de Navier–Stokes en un domini bidimensional periòdic, detallada al problema 3.7. En aquell cas, l'absència de parets en moviment feia que l'energia cinètica del sistema disminuís progressivament, a diferència del present mètode, on les parets injecten energia al sistema de manera sostinguda. S'observa, a més, que el nombre de remolins formats depèn directament de la condició inicial, i està estretament relacionat amb els harmònics de Fourier i de Txebishev activats en la simulació. D'altra banda, cal destacar que el mètode esdevé inestable i pot divergir si es pren una condició inicial que excita els harmònics més alts. En aquest cas, s'ha constatat que l'energia pot créixer sense control, i que seria necessari un mallat més fi amb més harmònics per estabilitzar la simulació.

Finalment, proposem estudiar la estabilitat i sensibilitat del mètode numèric a condicions inicial i nombre de harmònics considerats. A la figura 13, panell (a), s'observa el comportament de l'energia cinètica per a diferents nombre de Reynolds fins a 10 unitats de temps donada la condició inicial pertorbada de la figura 12. S'observa que el mètode numèric manté l'energia afitada en el rang de temps considerats. Podem discernir 3 comportaments. En verd, fluxos poc turbulents, on el terme difusiu porta la solució a estabilitzar-se ràpidament. En blau, fluids que comencen a ser turbulents però el nombre de Reynolds es troba encara sota el llindar de turbulència. En tal cas, observem l'aparició de turbulències que es perdran eventualment en el temps i, finalment, en vermell, fluxos que superen el llindar i són turbulents, en tal cas, s'han simulat intervals de fins a 100 segons i no s'ha observat que l'energia caigués. També hem analitzat els límits computacionals del mètode, i observem que simulacions extremadament turbulents de Re = 100.000 poden divergir i són inestables. Aquestes observacions són coherents amb l'anàlisi teòrica general del comportament de fluids en règim turbulent recollida a [6].

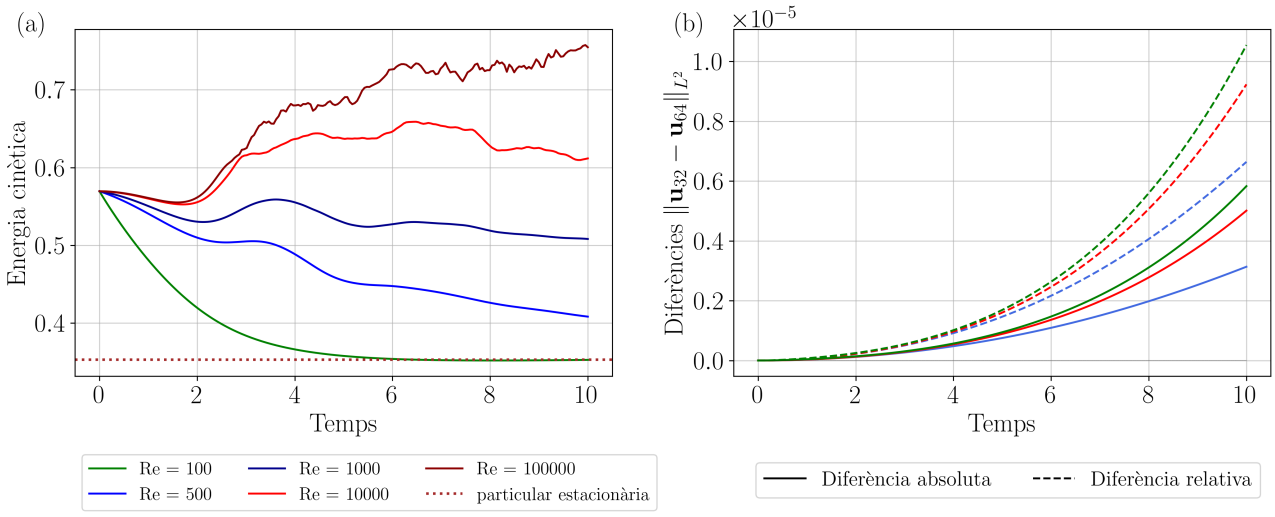


Figura 13: Anàlisi de l'energia cinètica i les diferències de les solucions per diferents harmònics segons Reynolds. (a) Evolució de l'energia cinètica per diferents valors de Reynolds sota una condició inicial pertorbada. Es distingeixen tres règims: laminar (verd), lleugerament turbulent (blau) i turbulent (vermell). La línia puntejada correspon a l'energia de la solució particular no homogènia estacionària $\mathbf{u} = (y, 0)$. El mètode manté l'energia afitada en règims laminars i transitoris. En règims turbulents l'energia no decau. Per $\mathrm{Re} \geqslant 10^5$, les simulacions poden ser numèricament inestables. (b) Sensibilitat a la resolució espectral comparant mallats de 32×32 i 64×64. L'error entre solucions és similar per a tots els règims i dominat pel càlcul del terme no lineal, amb una precisió limitada a 10^{-6}–10^{-5} per la mala condició de les matrius. El sistema mostra sensibilitat caòtica a les condicions inicials i limita la predictibilitat quantitativa a llarg termini.

A la Figura 13, panell (b), s'analitza la sensibilitat de la solució al nombre de nodes, comparant l'evolució d'una mateixa condició inicial (de la Figura 12) amb mallats de 32×32 i 64×64. Es calcula la diferència absoluta i relativa per a tres règims de flux: laminar, poc turbulent i turbulent. S'observa que l'error en la solució és pràcticament independent del règim del flux i que, fins i tot amb toleràncies

estrictes en l'integrador de Runge-Kutta-Fehlberg, no es pot reduir l'error per sota de 10^{-6} per unitat de temps. Aquest comportament revela dues qüestions fonamentals: (1) la naturalesa caòtica del sistema, que fa que sigui altament sensible a petites variacions de la condició inicial i del discretitzat, i (2) la precisió limitada en el càlcul del terme no lineal, que presenta discrepàncies al cinquè decimal entre els dos mallats considerats. Un anàlisi més detallat mostra que l'error relatiu associat al terme lineal ($\nabla \times (\nabla \times \mathbf{u})$) és de l'ordre de 10^{-8} per unitat de temps, mentre que el del terme no lineal ($\boldsymbol{\omega} \times \mathbf{u}$) arriba a 10^{-6}. Aquest últim error es vincula a la inestabilitat numèrica derivada d'invertir matrius amb nombres de condició $\sim 10^5$, fins i tot si són de tipus banda, fet que condiciona la precisió màxima assolible. Això implica que el model és útil per simular fluxos qualitativament, però no amb precisió determinista. A continuació proposem possibles millores per abordar aquestes limitacions en futurs treballs.

4.4 Propostes de millora

Finalment, proposem algunes millores potencials per augmentar l'estabilitat del mètode numèric. Primer, considerem l'aplicació del factor d'integració (Secció 3.3.2) per estabilitzar el terme lineal. Seguint aquest enfocament, la matriu L associada a l'operador lineal del problema seria:

$$L(\boldsymbol{\alpha}) = \begin{bmatrix} \bar{A}_{(-N_x)} & 0 & \cdots & 0 \\ 0 & \bar{A}_{(-N_x+1)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \bar{A}_{(N_x)} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_{(-N_x)} \\ \boldsymbol{\alpha}_{(-N_x+1)} \\ \vdots \\ \boldsymbol{\alpha}_{(N_x)} \end{bmatrix}, \tag{4.25}$$

que té dimensió $(2N_x + 1) \times N_y$. Seguint les pautes del mètode, caldria calcular e^{Lt}, cosa que només seria factible si es diagonalitzés la matriu L mitjançant la diagonalització de tots els blocs $\bar{A}_{(k)}$. Un cop fet això, es podria tractar el terme lineal de manera analítica, de forma similar al que s'ha fet en el Problema 3.7 amb condicions de contorn periòdiques. Cal destacar que aquesta proposta milloraria l'estabilitat de l'algorisme pel que fa al terme lineal, però no solucionaria la inestabilitat associada al terme no lineal.

Per tal d'afrontar aquest problema, es podria proposar un canvi de base mitjançant la descomposició en valors singulars. En particular, es descompondrien les matrius quadrades $A_{(k)} = U_{(k)}\Sigma_{(k)}V_{(k)}^T$, i per cada mode $k \in \{-N_x, \ldots, N_x\}$, es reescriuria el sistema com:

$$U_{(k)}\Sigma_{(k)}\frac{d\ (V_{(k)}^T\boldsymbol{\alpha}_k)}{dt} = \frac{1}{Re}B_{(k)}\ \boldsymbol{\alpha}_k + \mathbf{F}_{(k)}(\boldsymbol{\alpha}).$$

Ara considerem el canvi de variable ben definit $\boldsymbol{\beta}_k = V_{(k)}^T\boldsymbol{\alpha}_{(k)}$ i obtenim un sistema alternatiu escrit com

$$\frac{d\boldsymbol{\beta}_{(k)}}{dt} = \frac{1}{Re}\Sigma^{-1}U_{(k)}^T B_{(k)}\ V_{(k)}^T\ \boldsymbol{\beta}_k + \Sigma^{-1}U_{(k)}^T\mathbf{F}_{(k)}(V^T\boldsymbol{\beta}).$$

A diferència dels casos anteriors, ara només cal invertir trivialment les matriu $\Sigma_{(k)}$, que són diagonals. Per tant, aquest canvi de variables podria mitigar la inestabilitat d'invertir les matriu $A_{(k)}$ en els mètodes anteriors i reduir l'error a la cinquena xifra decimal.

Tot i això, hem observat que la naturalesa caòtica dels sistemes de fluids, especialment en règims turbulents, introdueix una imprevisibilitat intrínseca que compromet la fiabilitat de les solucions numèriques a llarg termini. Aquest comportament es manifesta en diverses dificultats matemàtiques: una alta sensibilitat a les condicions inicials i a les pertorbacions numèriques, nombres de condició molt elevats en les matrius del sistema, que es disparen amb el refinament de la malla, i la necessitat d'incloure harmònics d'ordre molt alt per representar correctament els efectes turbulents. En conjunt, aquests factors posen de manifest les limitacions inherents dels mètodes numèrics davant la complexitat de la dinàmica de fluids.

5 Conclusions

Aquest treball ha tractat la modelització i simulació numèrica d'un flux bidimensional entre dues parets paral·leles en moviment, a partir d'una simulació i resolució numèrica original. Per fer-ho, s'ha desenvolupat un mètode numèric basat en els fonaments matemàtics proposats per Moser et al. [3], que ha estat adaptat i ampliat amb tècniques extretes de la literatura especialitzada. Concretament, s'han incorporat modificacions en l'ansatz de la solució per considerar parets en moviment, així com un tractament específic del terme no lineal i un estudi del mètode de Petrov-Galerkin. El mètode s'ha completat amb l'elecció d'una estratègia temporal adequada per a la integració numèrica, així com amb una anàlisi de les equacions de Navier-Stokes i una secció dedicada a l'estudi per separat dels termes lineals i no lineals.

Els resultats obtinguts mostren que el mètode numèric proposat és capaç de simular amb coherència física les transicions del flux a règims turbulents, mantenint una representació raonable de l'energia i la formació de remolins. Això ens permet descriure qualitativament l'evolució del sistema a llarg termini, tot i les limitacions numèriques associades a la sensibilitat caòtica i al tractament del terme no lineal. En particular, s'ha observat com la naturalesa inherentment inestable del problema té associada vàries implicacions matemàtiques, com nombres de condició elevats, dependència crítica de les condicions inicials i presència d'harmònics d'alt ordre. Paral·lelament, aquest treball ha permès aprofundir en els mètodes espectrals i la seva implementació, tot integrant conceptes de càlcul, anàlisi i mètodes numèrics desenvolupats al llarg del grau. La diversitat d'exemples tractats ha estat clau per captar les subtileses i els límits pràctics de cada algorisme.

De cara al futur, assenyalem diverses línies de millora i investigació: el desenvolupament d'algorismes eficients per fer el canvi de base espectral, la reducció del cost computacional mitjançant tècniques inspirades en la transformada de Fourier, i l'estudi de nous esquemes per a la resolució estable del sistema diferencial. Aquestes propostes obren la porta a una modelització interessant, robusta i eficient, en la qual l'autor expressa interès per continuar aprofundint.

Referències

[1] S. Zhang, S. Xu, H. Fu, L. Wu, Z. Liu, Y. Gao, C. Zhao, W. Wan, L. Wan, H. Lu, C. Li, Y. Liu, X. Lv, J. Xie, Y. Yu, J. Gu, X. Wang, Y. Zhang, C. Ning, Y. Fei, and X. Shen, "Toward earth system modeling with resolved clouds and ocean submesoscales on heterogeneous many-core hpcs," *National Science Review*, vol. 10, no. 6, p. nwad069, 2023.

[2] W. D. Gropp and D. E. Keyes, "Domain decomposition methods in computational fluid dynamics," *International Journal for Numerical Methods in Fluids*, vol. 14, no. 2, pp. 147–165, 1992.

[3] R. Moser, P. Moin, and A. Leonard, "A spectral numerical method for the Navier-Stokes equations with applications to Taylor-Couette flow," *Journal of Computational Physics*, vol. 52, no. 3, pp. 524–544, 1983.

[4] J. P. Boyd, *Chebyshev and Fourier Spectral Methods*. Mineola, NY: Dover Publications, 2nd ed., 2001. Originally published by Dover; electronic version available from the author.

[5] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang, *Spectral Methods in Fluid Dynamics*. Scientific Computation, Berlin, Heidelberg: Springer-Verlag, 1988.

[6] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang, *Spectral Methods: Fundamentals in Single Domains*. Scientific Computation, Springer Berlin, Heidelberg, 1 ed., 2006.

[7] A. J. Chorin and J. E. Marsden, *A Mathematical Introduction to Fluid Mechanics*, vol. 4 of *Texts in Applied Mathematics*. Springer New York, NY, 3 ed., 1993.

[8] O. A. Ladyzhenskaya, *The Mathematical Theory of Viscous Incompressible Flow*, vol. 2 of *Mathematics and Its Applications*. Camberwell, Australia: Gordon and Breach Science Publishers, 2nd ed., 1969. Print.

[9] X. Mora, "Les equacions de Navier-Stokes: Impredictibilitat fins i tot sense papallones?," *Mètode Science Studies Journal*, no. 8, 2017. Article rebut: 19/12/2016, acceptat: 10/03/2017.

[10] H. Abidi, G. Gui, and P. Zhang, "On the global existence and uniqueness of solution to 2-d inhomogeneous incompressible Navier-Stokes equations in critical spaces," 2023.

[11] N. J. Mauser, H. P. Stimming, D. Bäumer, M. Dörfler, and M. Ehler, "Applied analysis." https://www.univie.ac.at, 2024. Lecture notes, University of Vienna, Winter semester 2024/2025, Applied Analysis course.

[12] Wikipedia contributors, "Navier–Stokes equations — Wikipedia, the free encyclopedia," 2025. [Online; accessed 27-April-2025].

[13] J. Gabbard, T. Gillis, P. Chatelain, and W. M. van Rees, "An immersed interface method for the 2d vorticity-velocity Navier-Stokes equations with multiple bodies," *Journal of Computational Physics*, vol. 464, p. 111339, 2022.

[14] D. Gottlieb and S. A. Orszag, *Numerical Analysis of Spectral Methods: Theory and Applications*, vol. 26 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Philadelphia: Society for Industrial and Applied Mathematics (SIAM), 1977.

[15] A. Ern and J. Guermond, *Theory and Practice of Finite Elements*. Springer, 2004.

[16] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang, *Spectral Methods: Evolution to Complex Geometries and Applications to Fluid Dynamics*. Scientific Computation, Berlin, Heidelberg: Springer, 2007.

[17] K. Brauer, "The korteweg-de vries equation: History, exact solutions, and graphical representation." https://www.mathematik.uni-osnabrueck.de, 2014. University of Osnabrück, Germany. Last revision: February 2014.

[18] B. Fornberg, "A numerical study of 2-d turbulence," *Journal of Computational Physics*, vol. 25, no. 1, 1977.

[19] R. J. Donnelly, "Taylor-couette flow: The early days," *Physics Today*, vol. 44, pp. 32–39, Nov. 1991.

[20] E. Kılıç and P. Stanica, "The inverse of banded matrices," *Journal of Computational and Applied Mathematics*, vol. 237, no. 1, pp. 126–135, 2013.

[21] J. M. Mondelo, "Apunts de mètodes numèrics," 2009. Grau de Matemàtiques UAB, Departament de Matemàtiques, Universitat Autònoma de Barcelona, 5 de juny de 2009.

[22] J. Tao, S. Chen, and W. Su, "Local reynolds number and thresholds of transition in shear flows," *Science China Physics, Mechanics and Astronomy*, vol. 56, 02 2013.

[23] I. Peral, *Primer curso de ecuaciones en derivadas parciales*. Madrid: Universidad Autónoma de Madrid, 1995. / Ireneo Peral Alonso.

[24] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier Series," *Mathematics of Computation*, vol. 19, no. 90, pp. 297–301, 1965.

[25] R. Peyret, *Spectral Methods for Incompressible Viscous Flow*. Heidelberg: Springer, 2002.

[26] M. Deville, P. F. Fischer, and E. H. Mund, *High-Order Methods for Incompressible Fluid Flow*. Cambridge University Press, 2002.

[27] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*. Texts in Applied Mathematics, New York, NY: Springer, 3 ed., 2002.

[28] C. Bonet, Jorba, M. T. Martínez, J. Masdemont, M. Ollé, A. Susín, and M. València, *Notes on Numerical Calculus*. Edicions UPC, 1994. Access restricted to the UPC community.

[29] R. Burden, J. Faires, and A. Burden, *Numerical Analysis*. Cengage Learning, 2015.

A Preliminars als mètodes espectrals

A.1 Problemes de Sturm-Liouville

En aquesta primera secció, desenvolupem els problemes de Sturm-Liouville com a eina per trobar funcions base (vegeu la Secció 3.1), que esdevenen el fonament per construir espais de funcions amb base ortonormal [23].

Un mètode espectral, Secció 3, és sovint un problema de Sturm-Liouville generalitzat. Normalment, la solució espectral més assequible d'un problema de condicions inicials i de contorn consisteix a resoldre el problema de valors propis de Sturm-Liouville associat. Considerem un operador diferencial $\tilde{\mathcal{L}}$ definit per a cada $y \in \mathcal{C}^2([a,b])$ com

$$\tilde{\mathcal{L}}(y)(x) = a_0(x)y(x) + a_1(x)y'(x) + a_2(x)y''(x),$$

on $a_0 \in \mathcal{C}^1([a,b])$, $a_1, a_2 \in \mathcal{C}([a,b])$ i $a_0(x) \neq 0$ per tot $x \in [a,b]$.

Aleshores, per una funció $g \in \mathcal{C}^1([a,b])$ tal que $g(x) \neq 0$ per tot $x \in [a,b]$, es pot demostrar que el problema $\tilde{\mathcal{L}}(y)(x) = f$ i el problema $g\tilde{\mathcal{L}}(y) = gf$ són equivalents en el sentit que tenen les mateixes solucions. D'aquesta manera, s'escull una funció g que permeti reformular el problema anterior en el problema

$$\mathcal{L}(y) = g\tilde{\mathcal{L}}(y) = (p(x)y'(x))' + q(x)y(x)$$

que se'n diu forma auto-adjunta del problema de Sturm-Liouville. D'aquí es defineix el problema de contorn de Sturm-Liouville:

Definició A.1. *Un problema de Sturm-Liouville presenta la forma*

$$\begin{cases} \mathcal{L}(y) = f \\ \mathcal{U}(y) = h \end{cases},$$

on les condicions de contorn anteriors $\mathcal{U}(y) = h$ és comú escriure-les com:

$$\underbrace{\begin{pmatrix} m_1 & n_1 & 0 & 0 \\ 0 & 0 & p_1 & q_2 \end{pmatrix}}_{\mathcal{U}} \cdot \underbrace{\begin{pmatrix} y(a) \\ y(b) \\ y'(a) \\ y'(b) \end{pmatrix}}_{y} = \underbrace{\begin{pmatrix} h_1 \\ h_2 \end{pmatrix}}_{h}.$$

Es defineix el problema de valors propis de Sturm-Liouville com:

$$\begin{cases} \mathcal{L}(y) = \lambda w(x)y(x) \\ \mathcal{U}(y) = 0 \end{cases},$$

on $w(x)$ és anomenada funció pes i es suposa real, contínua i estrictament positiva a $[a,b]$. Es diu que $\lambda \in \mathbb{C}$ és un valor propi del problema anterior si existeix una solució no trivial per tal λ.

Proposició A.1. *Considerem el problema de valors propis de Sturm-Liouville anterior. Aleshores:*

1. *Si λ és un valor propi, aleshores $\lambda \in \mathbb{R}$.*
2. *Si $\lambda_1 \neq \lambda_2$ són valors propis i ϕ_1 i ϕ_2 les funcions pròpies corresponents, aleshores es verifica que*

$$\langle \phi_1, \phi_2 \rangle_w = \int_a^b \phi_1(x)\,\phi_2(x)\,w(x)\,dx = 0,$$

on $w(x) > 0$ és la funció pes del problema.

Aquesta estructura és d'interès en el context dels mètodes espectrals, ja que les funcions pròpies ϕ_n que se'n deriven formen una base ortogonal en un espai de funcions adequat com $L^2_w[a,b]$.

A.2 Bases de polinomis

En aquesta secció s'exposen els principals resultats en la teoria d'aproximació per a polinomis trigonomètrics, tot i que farem especial èmfasi en els polinomis trigonomètrics de Fourier i Txebishev usats en la Secció 3.2 i 3.3.

Donades dues funcions $f(x)$, $g(x)$ i un pes $\omega(x) \geqslant 0$ a $[a,b]$, la forma

$$\langle f, g\rangle := \int_a^b \omega(x)\, f(x) g(x)\, dx$$

defineix un producte interior semidefinit positiu, que introdueix una noció d'ortogonalitat. El teorema d'existència dels polinomis ortogonals garanteix que, a partir d'un pes $\omega(x)$, es poden construir recursivament polinomis ortogonals començant per $p_0(x) = 1$, tal i com es descriu als apunts de Mondelo [21]. Aquests polinomis satisfan problemes de Sturm–Liouville, que en determinen les equacions diferencials més adequades per a mètodes espectrals. Presentem a continuació els més habituals:

Nom	**Definició**	**Domini**	**Relació d'ortogonalitat**	**Pes**
Fourier Series	$e_n(x) = e^{inx}$	$[-\pi, \pi]$	$\langle e_m, e_n\rangle = \int_{-\pi}^{\pi} e^{imx}\overline{e^{inx}}dx = 2\pi\delta_{mn}$	$w(x) = 1$
Txebishev	$T_0(x) = 1$ $T_1(x) = x$ $T_2(x) = 2x^2 - 1$	$[-1, 1]$	$\int_{-1}^{1} \frac{T_m(x)T_n(x)}{\sqrt{1-x^2}}dx = \begin{cases} 0 & m \neq n \\ \pi & m = n = 0 \\ \pi/2 & m = n \neq 0 \end{cases}$	$w(x) = \frac{1}{\sqrt{1-x^2}}$
Legendre	$P_0(x) = 1$ $P_1(x) = x$ $P_2(x) = \frac{1}{2}(3x^2 - 1)$	$[-1, 1]$	$\int_{-1}^{1} P_m(x)P_n(x)dx = \frac{2}{2n+1}\delta_{mn}$	$w(x) = 1$
Hermite	$H_0(x) = 1$ $H_1(x) = 2x$ $H_2(x) = 4x^2 - 2$	$(-\infty, \infty)$	$\int_{-\infty}^{\infty} H_m(x)H_n(x)e^{-x^2}dx = \sqrt{\pi}2^n n!\delta_{mn}$	$w(x) = e^{-x^2}$
Laguerre	$L_0(x) = 1$ $L_1(x) = 1 - x$ $L_2(x) = \frac{1}{2}(z^2 - 4x + 2)$	$[0, \infty)$	$\int_0^{\infty} L_m(x)L_n(x)e^{-x}dx = \delta_{mn}$	$w(x) = e^{-x}$

Taula 2: Comparació dels diferents polinomis ortogonals, s'hi inclouen les definicions dels primers termes de cada família (Fourier, Txebishev, Legendre, Hermite i Laguerre), els seus dominis naturals, les relacions d'ortogonalitat amb les respectives funcions pes, i les formes integrals que caracteritzen la seva ortogonalitat ($\delta_{mn} = 1 \iff m = n$ és la delta discreta de Dirac).

Donada una funció $f(x) \in \mathcal{L}^2_\omega([a,b])$ es busca aleshores la combinació lineal de polinomis que satisfà la següent igualtat, on els coeficients $a_n \in \mathbb{R}$ anteriors s'obtenen a partir de les relacions d'ortogonalitat

$$f(x) = \sum_n a_n\, p_n(x), \qquad a_n \propto \int_a^b f(x)\, p_n(x)\, dx.$$

Presentarem els principals resultats per a sèries de Fourier i de Txebishev ja que seran els polinomis que s'usaran per a modelar el problema de la Secció 4.

A.3 Sèries de Fourier: Resultats d'aproximació, FFT, Taxa de Nyquist i derivació

Fem un breu incís a les sèries de Fourier presentant els principals resultats. S'usarà aquesta teoria al llarg del treball per entendre la convergència.

Les sèries de Fourier es motiven demostrant que el següent conjunt de funcions a l'interval $(0, 2\pi)$ indueix una base ortogonal respecte el pes $\omega(x) = 1$:

$$\phi_k(x) = e^{ikx} \quad k \in \mathbb{N} \qquad \langle \phi_k, \phi_l \rangle = \int_0^{2\pi} \phi_k(x)\overline{\phi_l(x)}dx = 2\pi\delta_{kl} \ .$$

Denotem per S_N el subespai donat pels polinomis trigonomètrics de fins a dimensió N:

$$S_N = \text{span}\{e^{ikx} \mid k \in \{-N_x, \dots, N_x\}\}$$

I per $P_N u$ la sèrie de Fourier, que és calcula com:

$$P_N u(x) = \sum_{k=-N}^{N} \hat{u}_k e^{ikx}, \qquad \hat{u}_k = \frac{1}{2\pi}\int_0^{2\pi} u(x)e^{-ikx}dx \quad k = 0, \pm 1, \dots \tag{A.1}$$

Al llarg del grau, s'ha estudiat que la sèrie de Fourier és L^2-convergent, és a dir:

$$\int_0^{2\pi} |u(x) - P_N u(x)|^2 dx \to 0 \qquad \text{quan} \qquad N \to \infty.$$

També es pot demostrar, mitjançant una convolució amb una aproximació a la identitat, que per a funcions contínues, periòdiques i de variació afitada, la sèrie de Fourier convergeix uniformement. Resultats més generals mostren que, per a funcions no contínues, la sèrie de Fourier convergeix puntualment al valor

$$\frac{u(x^+) + u(x^-)}{2} \quad \text{per a tot } x \in (0, 2\pi). \tag{A.2}$$

A.3.1 Resultats d'aproximació usant sèries de Fourier:

Detallem alguns resultat usats en els Exemples 3.2 i 3.5. La identitat de Parseval permet obtenir una quota per l'error que és fa aproximant la solució u per la sèrie de Fourier $P_N u$, a L^2:

$$||u - P_N u||_{L^2} = \left(2\pi \sum_{|k| \geqslant N} |\hat{u}_k|^2\right)^{1/2}.$$

Si, a més, la funció u és continua, l'error puntual be donat per

$$\max_{0 \leqslant x \leqslant 2\pi} |u(x) - P_N u(x)| \leqslant \sum_{|k| \geqslant N} |\hat{u}_k|.$$

És d'interès, aleshores, conèixer com decreixen els coeficients $\hat{u}_k$ en funció de k. El següent resultat en dona un criteri que s'aplicara a l'Exemple 3.2.

Teorema A.2 (Ordre de decreixement dels coeficients de Fourier)**.** *Sigui $u \in \mathcal{L}^2(0, 2\pi)$ tal que $u \in \mathcal{C}^m(0, 2\pi)$ i la derivada j-èssima de u és 2π-periòdica per a tot $0 \leqslant j \leqslant m - 1$, aleshores*

$$\hat{u}_k = \mathcal{O}\left(|k|^{-m}\right)$$

Demostració. Només cal desenvolupar la integral usant la regla de la cadena

$$2\pi\hat{u}_k = \int_0^{2\pi} u(x)e^{-ikx}dx = \frac{-1}{ik}(u(2\pi^-) - u(0^+)) + \frac{1}{ik}\int_0^{2\pi} u'(x)e^{-ikx}dx,$$

on el primer terme de la igualtat es cancel·la atès que $u(2\pi^-) = u(0^+)$ i el resultat es prova per inducció sobre el nombre de derivades. □

Corol·lari A.3. *Sigui $u \in \mathcal{L}^2([0,2\pi])$ infinitament diferenciable i amb derivades infinitament periòdiques i regulars. Aleshores els coeficients $\hat{u}_k$ decreixen més ràpid que qualsevol potència negativa de k.*

Quan no es compleixen les hipòtesis anteriors, la convergència acostuma a ser lenta i apareix el fenomen de Gibbs, que consisteix en petites oscil·lacions al voltant dels punts on la funció té discontinuïtats o canvis bruscos.

Una manera de formalitzar els resultats anteriors és introduint els espais de Sobolev per a funcions periòdiques. Es recorda la definició d'espai de Sobolev:

Definició A.2 (Espai de Sobolev $H^m(a,b)$)**.** *Per a un interval $(a,b) \subset \mathbb{R}$ i un enter $m \geqslant 0$, l'espai de Sobolev $H^m(a,b)$ és el conjunt de totes les funcions $u : [a,b] \to \mathbb{R}$ que són derivables fins a l'ordre m i tal que cadascuna d'aquestes derivades $u^{(k)}$, per $k = 0, 1, \ldots, m$, és quadrat integrable:*

$$H^m(a,b) = \{u \in L^2(a,b) : u, u', \ldots, u^{(m)} \in L^2(a,b)\}.$$

La norma associada és

$$\|u\|_{H^m(a,b)} = \left(\sum_{k=0}^{m} \int_a^b |u^{(k)}(x)|^2 \, dx \right)^{1/2}.$$

Considerem també el subespai periòdic

$$H_p^m(0,2\pi) = \left\{ u \in H^m(0,2\pi) \mid u^{(i)}(0) = u^{(i)}(2\pi), \quad i = 0, \ldots, m-1 \right\}.$$

Heurísticament, els espais $H_p^m(0,2\pi)$ contenen les funcions periòdiques que es poden derivar m vegades i per a les quals la seva sèrie de Fourier convergeix uniformement. Per exemple, per a $u \in H_p^1(0,2\pi)$, tenim que

$$(P_N u)' = P_N u', \quad \forall N \in \mathbb{N},$$

on $P_N u$ és la millor aproximació de u en la norma $L^2(0,2\pi)$ sobre les funcions de grau menor o igual a N.

Teorema A.4. *Per tot $u \in H_p^m(0,2\pi)$ i $m \geqslant 0$, existeix una constant $C > 0$ tal que*

$$\|u - P_N u\|_{L^2(0,2\pi)} \leqslant C N^{-m} \|u^{(m)}\|_{L^2(0,2\pi)}.$$

Demostració. Apliquem les definicions i la representació de Fourier:

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \|u - P_N u\|_{L^2(0,2\pi)} &= \left(\sum_{|k| \geqslant N} |\hat{u}_k|^2 \right)^{1/2} = \left(\sum_{|k| \geqslant N} \frac{1}{|k|^{2m}} |k|^{2m} |\hat{u}_k|^2 \right)^{1/2} \\ &\leqslant N^{-m} \left(\sum_{|k| \geqslant N} |k|^{2m} |\hat{u}_k|^2 \right)^{1/2}, \end{aligned}$$

on l'última suma està afitada per $\|u^{(m)}\|_{L^2(0,2\pi)}$. □

Fórmulacions més àmplies de les convergències anteriors es donen a través dels teoremes de Jackson, detallades a [6].

A.3.2 Transformada discreta de Fourier

Quan apliquem la transformada de Fourier a la pràctica, sovint no es pot definir exactament una funció a tot arreu i convé aproximar la integral per una suma discreta de Riemann. En aquest context, la manera de computar la transformada de Fourier es realitza mitjançant un dels algorismes més rellevants del segle XX: La transformada ràpida de Fourier (Fast Fourier Transform, FFT) [24].

Per a qualsevol nombre enter $N > 0$, considerem el conjunt de punts:

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

aleshores, donada una funció $u \in L^2(0, 2\pi)$, es defineix la transformada directa i inversa discreta de Fourier respecte a aquests punts com:

$$\tilde{u}_k = \frac{1}{N} \sum_{j=0}^{N-1} u(x_j)\, e^{-ikx_j}, \qquad k = -\frac{N}{2} + 1, \dots, \frac{N}{2}, \tag{A.3a}$$

$$u(x_j) = \sum_{k=-N/2+1}^{N/2} \tilde{u}_k\, e^{ikx_j}, \qquad j = 0, \dots, N-1. \tag{A.3b}$$

La matriu associada a la transformada discreta de Fourier és essencialment unitària. Això implica, segons el *teorema espectral de l'àlgebra lineal*, que la seva inversa coincideix amb la seva transposada conjugada i això permet obtenir la transformada inversa discreta de Fourier prenent signes oposats. L'algorisme de la FFT permet assolir una rapidesa computacional de $\mathcal{O}(N \log(N))$.

Teorema A.5. *Siguin $\hat{u}_k$ els coeficients obtinguts a través de (A.1) i $\tilde{u}_k$ els obtinguts a través de (A.3). Aleshores*

$$\tilde{u}_k = \hat{u}_k + \sum_{m \neq 0} \hat{u}_{k+Nm}.$$

És a dir, la transformada discreta de Fourier amb N punts no reprodueix directament la transformada teòrica contínua, sinó que en captura una versió discretitzada i periòdica de l'espectre, la qual limita la resolució.

A.3.3 Taxa de Nyquist i ample de banda

La Taxa de Nyquist s'ha usat als Exemples 3.2, 3.5. La pregunta fonamental a l'hora de mostrejar una funció és quants punts cal prendre i, quin error s'introdueix si se'n prenen menys.

Definició A.3. *Per a $t > 0$, definim l'espai de Paley-Wiener $PW(t)$ com*

$$PW(t) = \left\{ f \in L^2(\mathbb{R}^2) \quad : \quad supp(\hat{f}) \subset [-t, t] \right\}.$$

Teorema A.6 (Teorema de mostreig de Shannon II [11])**.** *Sigui $f \in PW(t)$ aleshores f pot ser perfectament reconstruïda només amb els valors a nT seguint*

$$f(t) = \sum_{n \in \mathbb{Z}} f(nT) \sin\left(\frac{t - nT}{T}\right).$$

Corol·lari A.7. *Per tal de reconstruir una funció en un interval donat sense pèrdua d'informació, la taxa de mostreig (número de mostrar per segon f_s) ha de satisfer la condició de Nyquist:*

$$f_s \geqslant 2 \cdot f_{max}$$

Per reconstruir perfectament una senyal amb freqüències de fins a 5 kHz, cal mostrejar a 10 kHz. Si el mostreig és insuficient, apareixen errors d'aproximació segons la filosofia de la transformada ràpida de Fourier: les altes freqüències es confonen amb les més baixes. Un exemple comú és la funció $\sin(mx)$, i veure que per a freqüències múltiples $m \in \mathbb{N}$, es produeixen els mateixos valors de mostreig.

A.3.4 La derivada per a nodes de Fourier

Introduïm la filosofia general per derivar funcions periòdiques que hem usat a (4.18). Donada una sèrie de Fourier uniformement convergent, la relació entre la sèrie $P_N u$ i la seva derivada $P_N u'$ ve donada per

$$P_N u = \sum_{k=-N}^{N} \hat{u}_k e^{ikx}, \qquad P_N u' = \sum_{k=-N}^{N} ik\hat{u}_k e^{ikx}.$$

Per fer-ho de manera eficient, es calcula la transformada discreta de Fourier de la funció u, es multiplica cada coeficient pel factor ik corresponent, i després es fa la transformada inversa per tornar a l'espai físic. Així, l'aproximació de la derivada als punts de la graella $x_j = \frac{2\pi j}{N}$, amb $j = 0, \ldots, N-1$, $(D_N u)_j$ és

$$(\mathcal{D}_N u)_j = \sum_{k=-N/2}^{N/2-1} \tilde{u}_k e^{2\pi ikj/N}, \qquad \tilde{u}_k = \frac{ik}{N} \sum_{l=0}^{N-1} u(x_l) e^{-2ikl\pi}/N, \qquad j = 0, \ldots, N-1$$

El procediment es representa pel diagrama següent:

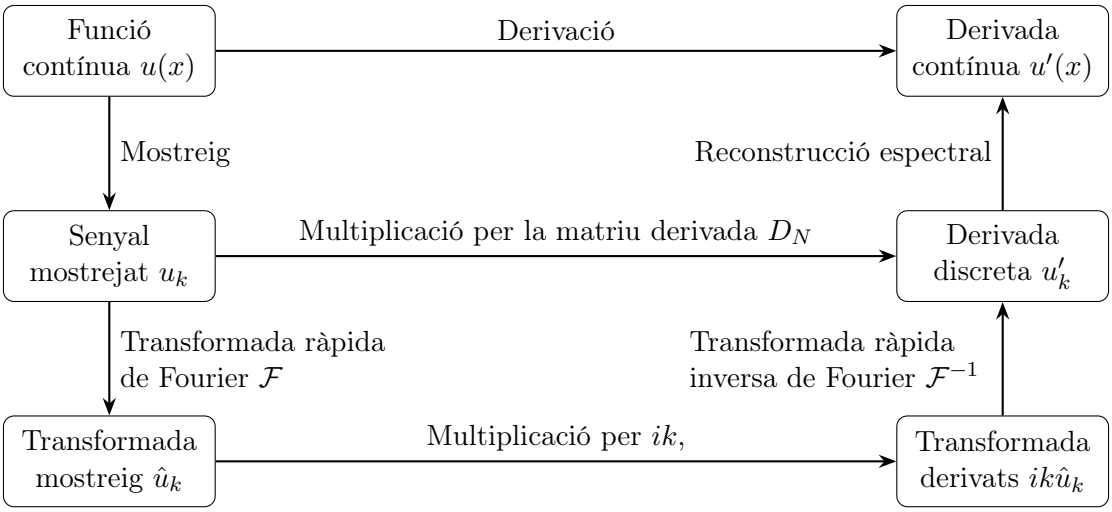


Figura 14: Comparació entre la derivació directa per matriu i la derivació espectral basada en la transformada de Fourier. Multiplicar per la matriu de derivada D_N comporta $2N^2$ operacions. Mentre que el camí inferior amb l'ús de l'algorisme FFT, comporta $5(\log_2 N - 5)N$ operacions.

Tot el procés es pot representar per una matriu derivada D_N que s'extreu de l'analisi anterior:

$$(D_n u)_j = \sum_{l=0}^{N-1} (D_N)_{jl} u_l, \qquad \text{on} \qquad (D_N)_{jl} = \frac{1}{N} \sum_{k=-N/2}^{N/2-1} ike^{2ik(j-l)\pi/N}.$$

Aquestes matrius són vàlides per a un nombre parell de punts, específicament de 2^k; per a nombres imparells es poden trobar a [25].

Tot i la simplicitat de la multiplicació per la matriu derivada D_N, el procés amb la matriu anterior requereix $2N^2$ operacions, mentre que el segon només $(5\log_2 N - 5)N$. Per tant, el primer és més eficient només quan $N \leqslant 8$.

A.4 Polinomis de Txebishev: Transformada discreta del cosinus i matriu de derivació

Els polinomis de Txebishev s'obtenen del problema de Sturm-Liouville

$$\left(\sqrt{1-x^2}T_k'(x)\right)' + \frac{k^2}{\sqrt{1-x^2}}T_k(x) = 0, \qquad |x| \leqslant 1.$$

aleshores $\omega(x) = 1/\sqrt{1-x^2}$. Una expressió compacte per aquest polinomis és:

$$T_k(x) = \cos k\theta, \qquad \theta = \arccos x, \qquad |x| \leqslant 1$$

Fent servir que $\cos(k+1)\theta + \cos(k-1)\theta = 2\cos\theta\cos k\theta$, s'obté que

$$T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x) \qquad T_0(x) = 1 \qquad T_1(x) = x$$

Algunes propietats dels polinomis de Txebishev són

$$|T_k(x)| \leqslant 1, \qquad T_k(\pm 1) = (\pm 1)^k.$$
$$|T_k'(x) \leqslant k^2|, \qquad T_k'(\pm 1) = (\pm 1)^{k+1}k^2.$$

Per una funció $u \in L^2_\omega(-1,1)$ una expansió de Txebishev ve donada per

$$u(x) = \sum_{k=0}^{\infty} \hat{u}_k T_k(x), \quad \hat{u}_k = \frac{2}{\pi c_k}\int_{-1}^{1} u(x)T_k(x)\omega(x)\,dx.$$

El desenvolupament anterior en polinomis de Txebitxev està relacionat amb la sèries de Fourier atès que si es defineix $\bar{u}(\theta) = u(\cos\theta)$, aleshores

$$\bar{u}(\theta) = \sum_{k=0}^{\infty} \hat{u}_k \cos k\theta$$

Així, si $u(x) \in \mathcal{C}^\infty$ a l'espai diferenciable $[-1,1]$, aleshores $\hat{u}(\theta)$ és $\mathcal{H}_p^\infty$ i usant el teorema A.2, s'obté que els coeficients decreixen més ràpid que algebraicament.

Teorema A.8 (Estimació d'aproximació espectral [16])**.** *Sigui $u \in H^m_\omega(-1,1)$, i sigui $P_N u$ la projecció ortogonal sobre l'espai de polinomis de grau $\leqslant N$ respecte al producte escalar ponderat per $\omega(x)$. Aleshores, de forma similar a la teoria de Fourier, existeix una constant $C > 0$, independent de N i u, tal que*

$$\|u - P_N u\|_{L^2_\omega(-1,1)} \leqslant CN^{-m}\,|u|_{H^m_\omega(-1,1)}, \qquad \forall u \in H^m_\omega(-1,1).$$

A.4.1 Transformada discreta del cosinus

En general, per a una família de polinomis ortogonals $\{p_k(x)\}_{k=0}^N$ respecte d'un pes $w(x)$ sobre un interval $[a,b]$, es pot construir una transformada discreta anàloga a la de Fourier.
Suposem que els punts $\{x_j\}_{j=0}^N$ són els punts de quadratura i que es defineix el producte escalar discret associat com:

$$(u,v)_N := \sum_{j=0}^{N} w_j\, u(x_j)v(x_j)$$

on w_j són els pesos de quadratura corresponents.

Els polinomis $p_k(x)$ es consideren ortogonals respecte aquest producte escalar discret, és a dir

$$(p_k, p_m)_N = \gamma_k\,\delta_{km},$$

on γ_k són constants de normalització.

Sigui $u(x)$ una funció discreta donada en els punts $\{x_j\}$. La seva interpolació en la base ortogonal $\{p_k\}$ pren la forma

$$I_N u(x) = \sum_{k=0}^{N} \tilde{u}_k \, p_k(x)$$

on els coeficients espectrals $\tilde{u}_k$ es poden obtenir projectant u sobre cada base

$$(u, p_k)_N = (I_N u, p_k)_N = \sum_{m=0}^{N} \tilde{u}_m \, (p_m, p_k)_N = \gamma_k \, \tilde{u}_k,$$

i, per tant,

$$\tilde{u}_k = \frac{1}{\gamma_k}(u, p_k)_N.$$

En el cas particular en què la base $\{p_k\}$ consisteix en cosinus (com en la transformada discreta del cosinus), i els punts x_j són els punts de Txebishev

$$x_j = \cos\left(\frac{\pi j}{N}\right), \qquad j = 0, \ldots, N$$

llavors els coeficients espectrals prenen la forma explícita

$$\tilde{u}_k = \frac{2 - \delta_{k0}}{N} \sum_{j=0}^{N} u_j \cos\left(\frac{\pi k j}{N}\right).$$

La matriu associada a aquesta transformada discreta del cosinus és:

$$C_{kj} = \frac{2}{N \, \hat{c}_j \, \hat{c}_k} \cos\left(\frac{\pi j k}{N}\right) \qquad \text{on} \qquad \hat{c}_j = \begin{cases} 2, & j = 0 \text{ o } j = N \\ 1, & j = 1, \ldots, N-1 \end{cases}.$$

Els punts de quadratura més habituals en aquest context són els punts de Gauss–Lobatto. Per a més detalls, vegeu la Secció A.5.2.

A.4.2 Matriu de derivada de Txebishev

De manera anàloga al cas de Fourier, volem derivar eficientment funcions expressades en la base de Txebitxev, tal i com hem usat a (3.11) i (4.19). Tenim

$$u(x) = \sum_{k=0}^{\infty} \hat{u}_k T_k(x).$$

Llavors, la seva derivada formal es pot expressar com:

$$u'(x) = \sum_{k=0}^{\infty} \hat{u}_k \frac{d}{dx} T_k(x) = \sum_{k=0}^{\infty} \tilde{u}_k T_k(x), \qquad \tilde{u}_k(x) = \frac{2}{c_k} + \sum_{\substack{p>k \\ p+k \text{ senar}}} p \hat{u}_p T_{p-1}(x). \tag{A.4}$$

amb $c_k = 2$ si $k = 0$ i $c_k = 1$ si $k \geqslant 1$.

Existeix una fórmula de recurrència lineal que permet calcular els valors de la derivada $u'(x)$ a partir dels valors de $u(x)$ coneguts en certs punts. En aquest treball presentem el formalisme de la matriu de derivació per a punts de Gauss-Lobatto. Donats els valors de $u(x_j)$ als punts de Gauss-Lobatto $x_j = \cos\left(\frac{\pi j}{N}\right), j = 0, 1, \ldots, N$, es pot construir una matriu real D_N de dimensions $(N+1) \times (N+1)$ tal que

$$u'(x_j) \approx (D_N u)(x_j) := \sum_{l=0}^{N} (D_N)_{jl} u(x_l), \qquad j = 0, \ldots, N. \tag{A.5}$$

Les entrades de la matriu D_N depenen dels punts que s'escullin per interpolar els polinomis de Txebishev, pels punts de Gauss-Lobatto es té

$$(D_N)_{jl} = \begin{cases} \dfrac{\bar{c}_j}{\bar{c}_l}\dfrac{(-1)^{j+l}}{x_j - x_l}, & j \neq l, \\ -\dfrac{x_l}{2(1-x_l^2)}, & 1 \leqslant j = l \leqslant N-1, \\ \dfrac{2N^2+1}{6}, & j = l = 0, \\ -\dfrac{2N^2+1}{6}, & j = l = N. \end{cases}$$

L'aplicació d'aquestes matrius segueix un esquema anàleg al de la Figura 14, adaptat al cas de Txebixev. El cost algorítmic de transformar, derivar per recurrència i transformar de tornada seguint (A.4) és de $(5\log_2 N + 8 + 2q)N$, on q és l'ordre de la derivada. En canvi, aplicar directament la matriu de derivació seguint (A.5) requereix $2N^2$ operacions. Conseqüentment, per $N \leqslant 12$, multiplicar per matrius és més eficient; per $12 \leqslant N \leqslant 128$, els costos són comparables; i per $N \geqslant 128$, cal utilitzar transformades. Tècniques per optimitzar la velocitat i el càlcul amb matrius es poden trobar a [26].

A.5 Mètodes numèrics: Extrapol·lació de Richardson i mètodes de quadratura

A.5.1 Extrapolació de Richardson

S'usa l'extrapolació de Richardson com a mètode per aproximar derivades de funcions. Tal mètode es fa servir per calcular l'error en les simulacions de les equacions de Navier-Stokes.

Sigui $f \in \mathcal{C}^\infty$ i l'expansió de Taylor de f al voltant d'un punt a convergent en un radi r_a:

$$f(a+h) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} h^n \qquad |h| \leqslant r_a$$

d'on s'obté una primera aproximació per la derivada que ve donada per

$$f'(a) = \frac{f(a+h) - f(a)}{h} - \sum_{n=1} \frac{f^{n+1}(a) h^n}{(n+1)!}. \tag{A.6}$$

Per tant

$$f'(a) = \frac{f(a+h) - f(a)}{h} + \mathcal{O}(h).$$

L'extrapolació de Richardson esdevé util per a calcular el valor de la derivada de forma més precisa. Expandim la sèrie a l'esquerra i a la dreta

$$f(a-h) = \sum_{n=0}^{\infty} \frac{f^{(n)(a)}}{n!}(-h)^n, \qquad |h| \leqslant r_a,$$

$$f(a+h) = \sum_{n=0}^{\infty} \frac{f^{(n)(a)}}{n!} h^n, \qquad |h| \leqslant r_a.$$

Combinant les dues expressions anteriors s'obté

$$f'(a) = \frac{f(a+h) - f(a-h)}{2h} - \sum_{k=1}^{\infty} \frac{f^{(2k+1)}(a)}{(2k+1)!} h^{2k}.$$

I per tant

$$f'(a) = \frac{f(a+h) - f(a-h)}{2h} + \mathcal{O}(h^2).$$

Recursivament, la fita es pot millorar considerant punts més propers amb distància $h \to \frac{h}{2}$:

$$f'(a) = \frac{4D_2\left(\frac{h}{2}\right) - D_2(h)}{3} + \mathcal{O}(h^4) \qquad \text{on} \qquad D_2(h) = \frac{f(a+h) - f(a-h)}{2h}.$$

Iterant el procés anterior, un pot aproximar una derivada d'ordre $2k$ seguint:

$$f'(a) = D_{2k}(h) + \mathcal{O}(h^{2k}), \qquad D_{2(k+1)} = \frac{4^k D_{2k}\left(\frac{h}{2}\right) - D_{2k}(h)}{4^k - 1}.$$

No obstant, cal pensar que $D^{2k}(h)$ requereix de conèixer els punts de la funció a $a \pm h, a \pm \frac{h}{2}, \ldots, a \pm \frac{h}{2(k-1)}$[28].

Per aproximar la segona derivada, el procediment és similar, en tal cas s'obte que:

$$f''(a) = \frac{f(a+h) - 2f(a) + f(a-h)}{h^2} - \frac{f^{(4)}(\xi)}{12}h^2, \qquad \xi \in (a, a+h).$$

L'extrapolació de Richardson, es pot repetir per la segona derivada recursivament, definint ara:

$$D_2(h) = \frac{f(a+h) - 2f(a) + f(a-h)}{h^2}, \qquad D_{2(k+1)} = \frac{4^k D_{2k}\left(\frac{h}{2}\right) - D_{2k}(h)}{4^k - 1}.$$

A.5.2 Mètodes de quadratura

Presentem les fórmules de quadratura per l'Exemple 3.4 i Problema 4.1 a l'equació (4.20) que permet calcular integrals amb un pes $\omega(x)$ no negatiu definit a l'interval $[a, b]$ mitjançant una combinació ponderada

$$\int_a^b \omega(x) f(x) dx \approx \sum_{i=1}^{n} \omega_i f(x_i). \tag{A.7}$$

on s'escullen els valors d'interpolació x_i i també els valors dels pesos ω_i per tal que la fórmula sigui exacte per qualsevol polinomi d'odre més petit que n. En particular, es defineix la fórmula de quadratura per:

$$w_i = \int_a^b \omega(x) L_{i,\{x_j\}}(x)\, dx \quad \text{amb} \quad L_{i,\{x_j\}}(x) = \prod_{\substack{j=1 \\ j\neq 1}}^{n} \frac{x - x_j}{x_i - x_j}. \tag{A.8}$$

Tota quadratura amb pes de n nodes és exacte per polinomis de fins a grau $n-1$. No obstant, si s'escullen els nodes sàviament, s'aconsegueix una aproximació millor:

Teorema A.9 ([21]). *Sigui x_i els punts d'interpolació de les arrels dels polinomi ortogonal φ_n construïts respecte el pes $\omega(x)$ a l'interval $[a, b]$. Aleshores la fórmula de quadratura descrita (A.7) amb pesos descrits a (A.8), té una exactitud de fins a ordre $2n-1$.*

En aquest treball considerem diverses fórmules de quadratura basades en els polinomis de Txebixev, totes amb pes $\omega(x) = 1/\sqrt{1-x^2}$. La Taula 3 recull els nodes x_j i els pesos ω_j per als tres tipus principals: Gauss-Txebixev, Gauss-Radau i Gauss-Lobatto. Cada cas incorpora el paràmetre δ, que en determina l'exactitud segons la fórmula (A.9). En particular les formules d'integració de Gauss impliquen que:

$$(u, v)_N = \sum_{j=0}^{N} u(x_j) v(x_j) w_j = (u, v)_w \quad \text{per } uv \in \mathbb{P}_{2N+\delta}. \tag{A.9}$$

on $\mathbb{P}_k$ denota l'espai de polinomis de grau com a màxim k.

[28] La fórmula de Richardson està estrètament relacionada amb la triangle de tartària.

Tipus	**Nodes** x_j	**Peses** ω_j	**Valor:** δ
Gauss-Txebishev	$x_j = \cos\left(\frac{(2j+1)\pi}{2N+2}\right)$	$\omega_j = \frac{\pi}{N+1}, \quad j = 0, \dots, N$	1
Gauss-Radau	$x_j = \cos\left(\frac{2\pi j}{2N+1}\right)$	$\omega_j = \begin{cases} \frac{\pi}{2N+1}, & j = 0 \\ \frac{2\pi}{2N+1}, & j = 1, \dots, N \end{cases}$	0
Gauss-Lobatto	$x_j = \cos\left(\frac{\pi j}{N}\right)$	$\omega_j = \begin{cases} \frac{\pi}{2N}, & j = 0, N \\ \frac{\pi}{N}, & j = 1, \dots, N-1 \end{cases}$	-1

Taula 3: Nodes i pesos per a diferents fórmules de quadratura associades als polinomis de Txebixev amb pes $\omega(x) = 1/\sqrt{1-x^2}$, el valor δ mostra l'exactitud de cada tipus a (A.9)

B Integradors numèrics

En aquesta secció fem un breu incís a la teoria d'integració numèrica per una equació diferencial ordinària usada per integrar les equacions diferencials dels Problemes 3.7 i 4.1. La informació s'ha extret de [27] i de [28].

Es proposa resoldre un problema de Cauchy

$$\frac{\partial y}{\partial t} = f(y,t) \qquad y(0) = y_0 \qquad t \geqslant 0$$

Considerem un interval temporal d'estudi donat per $t \in [0,T]$. Dividim aquest interval en N parts i definim el pas de temps com $h = \frac{T}{N}$. Definim els punts temporals[29] com $t_n = nh$. Partim de l'aproximació de la derivada mitjançant la recta tangent:

$$y(t) \approx y_0 + f(t_0, y_0)(t - t_0), \quad t \in [t_0, t_1].$$

Prenent $t = t_1$, obtenim el mètode d'Euler:

$$y_1 \approx y_0 + hf(t_0, y_0).$$

Per un instant (t_n, y_n) arbitrari, s'obté l'algorisme d'Euler:

$$\begin{cases} y_0 = y(0), \\ y_{n+1} = y_n + hf(t_n, y_n). \end{cases}$$

B.1 Mètodes d'un pas: Integradors de Runge-Kutta-Fehlberg i estabilitat i convergència

Introduïm els mètodes d'un pas a través de tres definicions essencials i un teorema.

Definició B.1 (Mètodes d'un pas)**.** *Anomenarem mètodes d'un pas aquells mètodes donats per*

$$\begin{cases} y(0) = y_0, \\ y_{n+1} = y_n + h\,\Phi(t_n, y_n, h), \end{cases}$$

on Φ *és una funció contínua sobre* $[0,T] \times \mathbb{R} \times [0, h_0]$ *i Lipschitz en* y*.*

Definició B.2 (Error de truncament)**.** *Anomenem error de truncament en el punt* t_n *al valor* $\epsilon_n = |y(t_n) - y_n|$*.*

[29] Cal tenir en compte que es busquen algorismes que permetin simular el valor explícit de $y(t)$ només a instants $t \in \{t_0, \dots, t_n\}$. En el nostre treball, aquesta condició és suficient.

Definició B.3 (Ordre d'un mètode)**.** *Direm que un mètode d'integració té ordre p, amb $p \in \mathbb{N}$, si existeixen constants $h_0 > 0$ i $k > 0$ tals que $\epsilon_n \leqslant kh^p$ per a tot $h \in [0, h_0]$ i per a tot $n \in \{0, \dots, N\}$. Escriurem que l'error és $\mathcal{O}(h^p)$.*

Observem que els valors h_0 i k poden dependre del nombre de passos N. Tanmateix, aquest fet no és essencial, ja que la definició pretén posar èmfasi en el comportament de l'error ϵ_n en funció de la mida del pas h. El següent teorema permet desenvolupar integradors numèrics de diferents ordres:

Teorema B.1. *Si es compleix la següent condició:*

$$\frac{y(t+h) - y(t)}{h} - \Phi(t, y(t), h) = \mathcal{O}(h^p), \qquad \forall t \in [a, b], \quad p \in \mathbb{N},$$

aleshores el mètode d'integració té ordre p.

Demostració. Volem acotar l'error de truncament $\epsilon_n = |y(t_n) - y_n|$ i demostrar que satisfà la definició B.3. Aleshores

$$\begin{aligned}
\epsilon_n &= |y(t_n) - y_n| \\
&\leqslant |y(t_n) - y(t_{n-1}) - h\Phi(t_{n-1}, y(t_{n-1}), h)| \\
&\quad + |y(t_{n-1}) + h\Phi(t_{n-1}, y(t_{n-1}), h) - y_n| \\
&= |y(t_{n-1} + h) - y(t_n) - h\Phi(t_{n-1}, y(t_{n-1}), h)| + |y(t_{n-1}) - y_{n-1}| \\
&\quad + h|\Phi(t_{n-1}, y(t_{n-1}), h) - \Phi(t_{n-1}, y_{n-1}, h)| \\
&\leqslant kh^{p+1} + \epsilon_{n-1} + hL|y(t_{n-1}) - y_{n-1}| \\
&\leqslant kh^{p+1} + (1 + hL)\epsilon_{n-1},
\end{aligned}$$

on hem suposat que Φ és Lipschitz respecte al segon argument, amb constant $L > 0$. Aplicant recursivament aquesta desigualtat

$$\begin{aligned}
\epsilon_n &\leqslant kh^{p+1} + (1 + hL)\epsilon_{n-1} \\
&\leqslant kh^{p+1} + (1 + hL)kh^{p+1} + (1 + hL)^2\epsilon_{n-2} \\
&\leqslant \cdots \leqslant kh^{p+1} \sum_{i=0}^{n-1} (1 + hL)^i.
\end{aligned}$$

Com que $(1 + hL)^n \leqslant e^{Lnh} = e^{LT}$ per $t_n \leqslant T$, es té

$$\epsilon_n \leqslant kh^{p+1} \sum_{i=0}^{n-1} (1 + hL)^i \leqslant kh^{p+1} \cdot \frac{(1 + hL)^n - 1}{hL} \leqslant \frac{k}{L}(1 + hL)^n h^p \leqslant \tilde{k}h^p,$$

per una constant $\tilde{k} > 0$, i per tant, el mètode d'integració té ordre p. □

Aplicant el resultat anterior a $\Phi(t, y, h) = f(t, y)$ es demostra que el mètode d'Euler té ordre 1.

El primer que hom pensa és en dissenyar un mètode d'odre superior usant el Teorema B.1 i prendre per $\Phi(t, y(t), h)$ una aproximació de f per Taylor. Aquest mètodes es coneixen com a mètodes de Taylor. A la pràctica no solen implementar-se ja que habitualment és costós obtenir les derivades parcials de f. Per a contrarestar això, s'introdueixen els mètodes de Runge-Kutta, que consisteixen en emprar el teorema anterior i construir una funció Φ que permeti obtenir una convergència d'ordre superior usant només evaluacions de la funció f.

B.1.1 Mètodes de Runge-Kutta

Els mètodes de Runge-Kutta generalitzen el mètode d'Euler tot combinant diverses estimacions del pendent. Es defineixen per l'algorisme següent:

$$y(0) = y_0,$$
$$y_{n+1} = y_n + h\sum_{i=1}^{k} c_i k_i^n,$$

on $k \in \mathbb{N}$ és el nombre d'etapes i els valors intermedis k_i^n es defineixen recursivament com

$$\begin{cases} k_1^n = f(t_n, y_n), \\ k_i^n = f\left(t_n + a_i h,\ y_n + h\sum_{j=1}^{i-1} b_{ij} k_j^n\right), \quad \text{per } i = 2, \ldots, k. \end{cases}$$

L'objectiu d'aquests mètodes és aproximar derivades d'ordre superior mitjançant mitjanes ponderades dels valors de f en diversos punts. El mètode d'Euler és un cas particular amb $k = 1$ i $c_1 = 1$, i és d'ordre 1.

En el cas $k = 2$, la funció Φ que en resulta és

$$\Phi(t, y, h) = c_1 k_1 + c_2 k_2 = c_1 f(t, y) + c_2 f(t + a_2 h, y + h b_{21} f(t, y)).$$

Per tal que el mètode tingui ordre p, es demana que:

$$\left| \frac{y(t+h) - y(t)}{h} - \Phi(t, y, h) \right| = \mathcal{O}(h^p).$$

D'aquesta manera, es buscen els coeficients que satisfan la condició anterior:

$$\begin{aligned} \frac{y(t+h) - y(t)}{h} &= y'(t) + \frac{h}{2} y''(t) + \frac{h^2}{6} y'''(t) + \mathcal{O}(h^3) \\ &= f(t, y) + \frac{h}{2}\left[f_t(t, y) + f_y(t, y) f(t, y)\right] \\ &\quad + \frac{h^2}{6}\left[f_{tt}(t, y) + 2 f_{ty}(t, y) f(t, y) + f_t(t, y) f_y(t, y)\right. \\ &\quad \left. + f_y^2(t, y) f(t, y) + f_{yy}(t, y) f^2(t, y)\right] + \mathcal{O}(h^3). \end{aligned}$$

D'altra banda, desenvolupem la funció $\Phi(t, y, h)$ per a un mètode de Runge-Kutta amb dues etapes:

$$\begin{aligned} \Phi(t, y, h) &= c_1 f(t, y) + c_2 f\left(t + a_2 h, y + h b_{21} f(t, y)\right) \\ &= c_1 f(t, y) + c_2\left[f(t, y) + a_2 h f_t(t, y) + h b_{21} f_y(t, y) f(t, y)\right. \\ &\quad + \frac{(a_2 h)^2}{2} f_{tt}(t, y) + a_2 b_{21} h^2 f_{ty}(t, y) f(t, y) \\ &\quad \left. + \frac{(b_{21} h)^2}{2} f_y^2(t, y) f(t, y) + \frac{(b_{21} h)^2}{2} f_{yy}(t, y) f^2(t, y)\right] + \mathcal{O}(h^3). \end{aligned}$$

Igualant els dos desenvolupaments per tal de maximitzar l'ordre, es compara terme a terme. Per garantir que Φ coincideixi amb el desenvolupament de Taylor fins a ordre 2, s'ha de satisfer el següent sistema d'equacions:

$$\begin{cases} c_1 + c_2 = 1 \\ a_2 c_2 = \dfrac{1}{2} \\ b_{21} c_2 = \dfrac{1}{2} \end{cases}$$

Aquest és un sistema compatible indeterminat i qualsevol tria que satisfaci els coeficients serà d'odre 2. El més conegut que pren el nom de Runge-Kutta d'ordre 2 s'aconsegueix prenent $a_2 = b_{21} = 1$ i $c_1 = c_2 = 1/2$. L'algorisme esdevé aleshores

$$\begin{cases} y(0) = y_0 \\ y_{n+1} = y_n + h\left[\frac{1}{2}f(x_n, y_n) + \frac{1}{2}f(x_n + h, y_n + hf(x_n, y_n))\right] \end{cases}$$

De forma similar al cas anterior, es pot obtenir el Runge-Kutta d'odre 4, el qual ve definit per l'algorisme

$$\begin{cases} y(0) = y_0 \\ y_{n+1} = y_n + \frac{h}{6}[k_1^n + 2k_2^n + 2k_3^2 + k_4^n] \end{cases},$$

on

$$\begin{cases} k_1^n = f(x_n, y_n) \\ k_2^n = f(x_n + \frac{h}{2}, y_n + \frac{h}{2}k_1^n) \\ k_3^n = f(x_n + \frac{h}{2}, y_n + \frac{h}{2}k_2^n) \\ k_4^n = f(x_n + h, y_n + hk_3^n) \end{cases}$$

B.1.2 Integrador de Runge-Kutta-Fehlberg

Descrivim en detall l'integrador numèric presentat a la secció 4.1.3 del problema 4. Essencialment l'integrador de Runge-Kutta-Fehlberg considera dos integradors Runge-Kutta d'ordres n i $n+1$ i una expressió que permeti definir un pas h de manera que l'error entre integradors estigui acotat [29]. Així, es defineixen dos mètodes que satisfan l'aproximació següent en ordres n i $n+1$:

$$\alpha(t_{i+1}) = \alpha(t_i) + h\Phi(t_i, y(t_i), h) + \mathcal{O}(h^{n+1}), \tag{B.1a}$$

$$\alpha(t_{i+1}) = \alpha(t_i) + h\tilde{\Phi}(t_i, y(t_i), h) + \mathcal{O}(h^{n+2}) \tag{B.1b}$$

on Φ i $\tilde{\Phi}$ són les funcions d'un pas dels mètodes d'odres n i $n+1$, i partir d'això, definim els mètodes iteratius respectivament per

$$\alpha_{i+1} = \alpha_i + h\Phi(t_i, \alpha_i, h) \qquad \alpha_0 = \alpha(0) \qquad i > 0,$$

$$\tilde{\alpha}_{i+1} = \tilde{\alpha}_i + h\tilde{\Phi}(t_i, \tilde{\alpha}_i, h) \qquad \alpha_0 = \alpha(0) \qquad i > 0.$$

Suposem que $\alpha(t_i) = \alpha_i = \tilde{\alpha}_i$. D'aquesta manera l'error de l'integrador d'odre 4 ve donat per:

$$\begin{aligned} \tau_{i+1}(h) &= \frac{\alpha(t_{i+1}) - \alpha(t_i)}{h} - \Phi(t_i, \alpha(t_i), h) \\ &= \frac{\alpha(t_{i+1}) - \alpha_i}{h} - \Phi(t_i, \alpha_i, h) \\ &= \frac{\alpha(t_{i+1}) - [\alpha_i + h\Phi(t_i, \alpha_i, h)]}{h} = \frac{1}{h}\left(\alpha(t_{i+1}) - \alpha_{i+1}\right). \end{aligned}$$

De forma similar s'obté que

$$\tilde{\tau}_{i+1}(h) = \frac{1}{h}(\alpha(t_{i+1}) - \tilde{\alpha}_{i+1}).$$

Conseqüentment, s'obté

$$\begin{aligned} \tau_{i+1}(h) &= \frac{1}{h}\left(\alpha(t_{i+1}) - \alpha_{i+1}\right) \\ &= \frac{1}{h}\left[\alpha(t_{i+1}) - \tilde{\alpha}_{i+1} + (\tilde{\alpha}_{i+1} - \alpha_{i+1})\right] \\ &= \tilde{\tau}_{i+1}(h) + \frac{1}{h}\left(\tilde{\alpha}_{i+1} - \alpha_{i+1}\right). \end{aligned}$$

Com que $\tau_{i+1}(h)$ té ordre $\mathcal{O}(h^n)$ a diferència de $\tilde{\tau}_{i+1}(h)$ que té ordre $\mathcal{O}(h^{n+1})$, tal i com es veu a (B.1), aleshores l'aproximació següent esdevé raonable

$$\tau_{i+1}(h) \approx \frac{1}{h}(\tilde{\alpha}_{i+1} - \alpha_{i+1}) \approx Kh^n.$$

Per un nombre K independent de h. L'expressió anterior dona una aproximació de l'error entre els dos mètodes. Si modifiquem la passa per un factor q, aleshores:

$$\tau_{i+1}(qh) \approx K(qh)^n = q^n(Kh^n) \approx \frac{q^n}{h}(\tilde{\alpha}_{i+1} - \alpha_{i+1}).$$

Si volem que l'error entre els dos mètodes esdevingui més petit que ϵ, l'equació anterior es rescriu com

$$q \leqslant \left(\frac{\epsilon h}{|\tilde{\alpha}_{i+1} - \alpha_{i+1}|}\right)^{1/n}.$$

En el cas que ens ocupa, considerem integradors de Runge-Kutta d'ordre 4 i 5, utilitzant una taula d'avaluacions eficient per a f, de manera que, per a cada pas, només calen un total de 6 avaluacions de f per definir ambdós mètodes. Això suposa un avantatge d'un 40% respecte a la utilització separada de dos mètodes RK4 i RK5. La integració numèrica s'efectua escollint una passa inicial h i ajustant q a cada iteració. A causa dels errors d'arrodoniment, sovint es fixa una fita més conservadora per a q. Una elecció habitual en el mètode de Runge-Kutta 4(5) és [29]

$$q = 0.84\left(\frac{\varepsilon h}{|\widetilde{\alpha}_{i+1} - \alpha_{i+1}|}\right)^{1/4}.$$

Aleshores, pel mètode de Runge-Kutta Felhberg, es consideren els mètodes iteratius:

$$\tilde{\alpha}_{i+1} = \alpha_i + \frac{16}{135}k_1 + \frac{6656}{12825}k_3 + \frac{28561}{56430}k_4 - \frac{9}{50}k_5 + \frac{2}{55}k_6$$
$$\alpha_{i+1} = \alpha_i + \frac{25}{216}k_1 + \frac{1408}{2565}k_3 + \frac{2197}{4104}k_4 - \frac{1}{5}k_5$$

on les equacions dels coeficients són:

$$\begin{aligned}
k_1 &= hf(t_i, \alpha_i),\\
k_2 &= hf\left(t_i + \tfrac{1}{4}h, \alpha_i + \tfrac{1}{4}k_1\right),\\
k_3 &= hf\left(t_i + \tfrac{3}{8}h, \alpha_i + \tfrac{3}{32}k_1 + \tfrac{9}{32}k_2\right),\\
k_4 &= hf\left(t_i + \tfrac{12}{13}h, \alpha_i + \tfrac{1932}{2197}k_1 - \tfrac{7200}{2197}k_2 + \tfrac{7296}{2197}k_3\right),\\
k_5 &= hf\left(t_i + h, \alpha_i + \tfrac{439}{216}k_1 - 8k_2 + \tfrac{3680}{513}k_3 - \tfrac{845}{4104}k_4\right),\\
k_6 &= hf\left(t_i + \tfrac{1}{2}h, \alpha_i - \tfrac{8}{27}k_1 + 2k_2 - \tfrac{3544}{2565}k_3 + \tfrac{1859}{4104}k_4 - \tfrac{11}{40}k_5\right).
\end{aligned}$$

Per implementar aquest mètode de forma eficient es discerneixen dos casos. Quan $q \leqslant 1$ es rebutja la tria inicial de h al pas i i es repeteixen els càlculs usant $q_i h$. Quan $q \geqslant 1$, s'accepta el valor computat al pas i amb la mida del pas h, però es canvia la mida del pas a $q_i h$ per el pas $(i+1)$.

B.1.3 Estabilitat i convergència

Introduïm les nocions de convergència i estabilitat dels mètodes de Runge-Kutta. Cal distingir entre l'estabilitat del mètode numèric i la de l'equació diferencial. Si aquesta és mal condicionada, l'integrador serà inestable independentment del pas temporal. Assumirem, per ara, que l'equació diferencial és estable.

Perquè un mètode estigui ben definit, ha de tenir ordre més gran que 1.

Definició B.4 (Consistència)**.** *Direm que un mètode és consitent si té ordre més gran que* 1*.*

Els mètodes de Runge-Kutta es poden englobar dins dels anomenats mètodes multipàs, que es poden consultar a [28]. En tot cas, generalitzem la teoria de convergència a:

Definició B.5 (Estabilitat)**.** *Donat un mètode definit per l'algorisme*

$$y_{n+1} = \sum_{i=1}^{k} \alpha_i y_{n+1-i} + h\,\Phi(x_{n+1}, x_n, \ldots, x_{n+1-m}, y_n, \ldots, y_{n+1-k}),$$

es diu que és estable si totes les arrels α_i *del polinomi característic*

$$p(\alpha) = \alpha^k - \alpha_1\alpha^{k-1} - \cdots - \alpha_k,$$

satisfan $|\alpha_i| \leqslant 1$ *per a tot* $i = 1, \ldots, k$*, i les que compleixen* $|\alpha_i| = 1$ *són arrels simples.*

D'altra banda, el mètode es diu fortament estable si exactament $k-1$ *arrels satisfan* $|\alpha_i| < 1$ *i una (simple) té mòdul* 1*.*

Atès que un mètode d'un pas té polinomi característic $p(\alpha) = 1 - \alpha$, es dedueix que tots els mètodes d'un pas són fortament estables.

Quan un mètode és fortament estable i consistent, la solució analítica i l'aproximada tenen el mateix límit quan $n \to \infty$. Això es deu al fet que el polinomi característic té totes les arrels amb mòdul menor que 1, excepte una d'elles, que té mòdul 1 i és simple. La teoria de funcions discretes mostra que, en aquest cas, les potències de les arrels amb mòdul menor que 1 tendeixen a zero, i la contribució dominant prové de l'arrel de mòdul 1. Per tant, el mètode és convergent a l'infinit.

Definició B.6. *Diem que un mètode és convergent quan* $\epsilon_n \to 0$ *quan* $n \to \infty$*.*

C Equivalències analítiques i càlculs complementaris de la Secció 4

C.1 Equivalències en les equacions de Navier-Stokes incompressibles

Sovint s'utilitza una formulació alternativa de les equacions de Navier-Stokes per a fluids incompressibles en termes de la vorticitat i la pressió dinàmica. Segons el teorema de Helmholtz 2.6, la vorticitat conté pràcticament tota la informació sobre el flux. Així, és convenient reescriure el sistema original en funció d'aquesta magnitud. Comencem amb les equacions de Navier-Stokes per a un fluid incompressible:

$$\begin{aligned} \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\,\mathbf{u} &= -\nabla \mathrm{P} + \frac{1}{\mathrm{Re}}\Delta\mathbf{u} + \mathbf{f}, \\ \nabla\cdot\mathbf{u} &= 0, \\ \mathbf{u} &= 0 \quad \text{a} \quad \partial\Omega. \end{aligned}$$

Primer, destaquem que el terme de dissipació es pot escriure com un doble rotacional mitjançant la identitat següent

$$\nabla\times\nabla\times\mathbf{u} = \nabla(\nabla\cdot\mathbf{u}) - (\nabla\cdot\nabla)\mathbf{u} = -\Delta\mathbf{u} \tag{C.1}$$

atès que $\nabla\cdot\mathbf{u} = 0$.

D'altra banda, es pot fer ús de la identitat de Helmholtz per a la vorticitat ($\mathbf{w} = \nabla\times\mathbf{u}$). Així

$$\mathbf{u}\times(\nabla\times\mathbf{u}) + (\mathbf{u}\cdot\nabla)\mathbf{u} = \nabla\left(\frac{|\mathbf{u}|^2}{2}\right) \implies (\mathbf{u}\cdot\nabla)\mathbf{u} = \nabla\left(\frac{|\mathbf{u}|^2}{2}\right) - \mathbf{u}\times\mathbf{w}.$$

El terme $\left(\frac{|\mathbf{u}|^2}{2}\right)$ representa habitualment l'energia cinètica per unitat de volum, i es pot absorbir dins de la pressió. En tal cas, es formula una versió equivalent de l'equació de Navier-Stokes, on P rep el nom de pressió dinàmica, en el nostre problema aquest canvi esdevé útil ja que la pressió desapareixerà sota el mètode espectral.
Això condueix al sistema d'equacions presentat a la Secció 4, on P és la pressió dinàmica i s'introdueix el doble rotacional:

$$\begin{aligned} \frac{\partial \mathbf{u}}{\partial t} &= -\nabla \mathrm{P} - \frac{1}{\mathrm{Re}}\nabla\times\nabla\times\mathbf{u} + \mathbf{u}\times\mathbf{w}, \\ \nabla\cdot\mathbf{u} &= 0. \end{aligned}$$

A continuació motivem l'aparició de l'operador $\mathcal{L}$ a (4.11) mitjançant les següent igualtats. Es coneix la següent identitat vectorial

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\cdot\mathbf{c})\,\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\,\mathbf{c}.$$

Utilitzant aquesta identitat vectorial, primer escrivim les transformades de Fourier dels termes derivats com

$$\widehat{\nabla\times\mathbf{u}} = i\boldsymbol{k}\times\hat{\mathbf{u}}_{\mathbf{k}}, \qquad \widehat{\nabla\cdot\mathbf{u}} = i\mathbf{k}\cdot\hat{\mathbf{u}}_{\mathbf{k}} \qquad .$$

Fent servir la identitat anterior s'observa que

$$\widehat{\nabla\times\nabla\times}\,\mathbf{u}_{\mathbf{k}} = i\mathbf{k}\times(i\mathbf{k}\times\hat{\mathbf{u}}_{\mathbf{k}}) = -k^2\hat{\mathbf{u}}_{\mathbf{k}} + (\mathbf{k}\cdot\hat{\mathbf{u}}_{\mathbf{k}})\,\mathbf{k} = -k^2\hat{\mathbf{u}}_{\mathbf{k}},$$

i això ens porta a la forma següent per una transformada només en l'eix de les x,

$$\widehat{\nabla\times\nabla\times}\,\mathbf{u} = -\widehat{\nabla^2\mathbf{u}} = -\left(k^2 + \frac{\partial^2}{\partial y^2}\right)\hat{\mathbf{u}}.$$

C.2 Càlculs de derivades, matrius i termes no lineal:

Les derivades de f necessàries per calcular (4.12) es donen per:

$$\begin{aligned}
f_j(y) &= (1-y^2)^2 T_j(y),\\
f_j'(y) &= -4y\,(1-y^2)T_j(y) + (1-y^2)^2 T_j'(y),\\
f_j''(y) &= -4(1-3y^2)T_j(y) - 8y(1-y^2)T_j'(y) + (1-y^2)^2 T_j''(y),\\
f_j'''(y) &= 24yT_j(y) - 12(1-y^2)T_j'(y) - 12y(1-y^2)T_j''(y) + (1-y^2)^2 T_j'''(y).
\end{aligned}$$

Similarment, per les derivades de les funcions g amb pes es de Txebishev tenim:

$$g_j(y) = \left(\frac{T_{j+2}(y)}{j(j+1)} - \frac{2T_j(y)}{(j+1)(j-1)} + \frac{T_{j+2}}{j(j-1)}\right)\frac{1}{4\sqrt{1-y^2}} = \frac{A(y)}{\sqrt{1-y^2}},$$

$$g_j'(y) = \frac{A'(y)}{\sqrt{1-y^2}} + \frac{yA(y)}{(1-y^2)^{3/2}},$$

$$g_j''(y) = \frac{A''(y)}{\sqrt{1-y^2}} + \frac{yA'(y)}{(1-y^2)^{3/2}} + \frac{A(y)+yA'(y)}{(1-y^2)^{3/2}} + \frac{3y^2A(y)}{(1-y^2)^{5/2}}.$$

Es presenten les derivacions del valor de les matrius presentades a (4.12). Els valors obtinguts es comproven mitjançant dues integracions que són equivalents

$$\begin{aligned}
A_{ij} &= \int_{-1}^{1} \mathbf{u}_j \cdot \boldsymbol{\xi}_i\, w(y)\, dy\\
&= \int_{-1}^{1} (g_i' f_j' + k^2\, g_i\, f_j)\, w(y)\,,
\end{aligned}$$

$$\begin{aligned}
A_{ij} &= \int_{-1}^{1} \mathbf{u}_j \cdot \boldsymbol{\xi}_i\, w(y)\, dy\\
&= \int_{-1}^{1} (g_i' f_j' + k^2\, g_i\, f_j)\, w(y)\, dy\\
&= \int_{-1}^{1} \left(-\left(f_j'' + f_j' \frac{y}{1+y^2}\right) g_i + k^2\, g_i\, f_j\right)\, w(y)\, dy\\
&= -\int_{-1}^{1} \left(\mathcal{L}(f_j) + \frac{y}{1+y^2} f_j'\right)\, g_i\, w(y)\, dy.
\end{aligned} \tag{C.2}$$

Similarment es fa el mateix per la matriu B, d'on s'obté

$$\begin{aligned}
B_{ij} &= \int_{-1}^{1} \mathcal{L}(\mathbf{u}_j) \cdot \boldsymbol{\xi}_i\, w(y)\, dy\\
&= \int_{-1}^{1} (f_j'''\, g_i' - k^2\, f_j'\, g_i' + k^2\, f_j''\, g_i - k^4\, f_j\, g_i)\, w(y)\, dy.
\end{aligned}$$

$$\begin{aligned}
B_{ij} &= \int_{-1}^{1} \mathcal{L}(\mathbf{u}_j) \cdot \boldsymbol{\xi}_i\, w(y)\, dy\\
&= \int_{-1}^{1} (\mathcal{L}(f_j') g_i' + \mathcal{L}(f_j)\, k^2\, g_i)\, w(y)\, dy\\
&= \int_{-1}^{1} \left(-\mathcal{L}(f_j)\left(g_i'' + g_i' \frac{y}{1-y^2}\right) + \mathcal{L}(f_j)\, k^2\, g_i\right)\, w(y)\, dy\\
&= -\int_{-1}^{1} \mathcal{L}(f_j)\left(g_i' \frac{y}{1-y^2} + \mathcal{L}(g_i)\right)\, w(y)\, dy.
\end{aligned} \tag{C.3}$$

En ambdós casos s'usa la regla de la cadena.

D'altra banda el terme no lineal presentat a (4.15) es desenvolupa de la següent manera. El terme de vorticitat és

$$\boldsymbol{\omega} = \nabla \times \mathbf{u} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial_x & \partial_y & \partial_z \\ u & v & 0 \end{vmatrix} = \left(0, 0, \ \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right).$$

Ja que els vector velocitats no tenen component z. Aleshores el terme no lineal esdevé:

$$\begin{aligned} \mathbf{u} \times \boldsymbol{\omega} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u & v & 0 \\ 0 & 0 & \omega \end{vmatrix} = (v\omega, -u\omega, 0) \\ &= \left[v\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right), -u\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)\right] \\ &= \left(v\frac{\partial v}{\partial x} - v\frac{\partial u}{\partial y}, -u\frac{\partial v}{\partial x} + u\frac{\partial u}{\partial y}\right). \end{aligned}$$

D'altra banda si es reescriu $\mathbf{u}$ com una suma del terme homogeni i no homogeni, $\mathbf{u} \to \mathbf{u} + y$. Aleshores el terme de (4.16) es desenvolupa com:

$$\begin{aligned} \widehat{\mathbf{u} \times \boldsymbol{\omega}} &= \left(\widehat{v\frac{\partial v}{\partial x}} - \widehat{v\frac{\partial (u+y)}{\partial y}}, -\widehat{(u+y)\frac{\partial v}{\partial x}} + \widehat{(u+y)\frac{\partial (u+y)}{\partial y}}\right) \\ &= \left(\widehat{v\frac{\partial v}{\partial x}} - \widehat{v\frac{\partial u}{\partial y}} - \widehat{v}, -\widehat{u\frac{\partial v}{\partial x}} - \widehat{y\frac{\partial v}{\partial x}} + \widehat{u\frac{\partial u}{\partial y}} + \widehat{y\frac{\partial u}{\partial y}} + \widehat{u+y}\right) \\ &= \left(\widehat{v\frac{\partial v}{\partial x}} - \widehat{v\frac{\partial u}{\partial y}}, -\widehat{u\frac{\partial v}{\partial x}} + \widehat{u\frac{\partial u}{\partial y}}\right) + \left(-\widehat{v}, -\widehat{y\frac{\partial v}{\partial x}} + \widehat{y\frac{\partial u}{\partial y}} + \widehat{u+y}\right). \end{aligned}$$

C.3 Descomposició en valors singulars (SVD)

Presentem la descomposició en valors singulars de (4.14). Invertir matrius té un cost algorítmic de $\mathcal{O}(N^3)$ i no resulta estable si el problema està mal condicionat. Per a resoldre aquest problema s'utilitza el mètode de descomposició en valors singulars per invertir matrius de manera robusta.

Teorema C.1 (Descomposició en valors singulars [11])**.** *Sigui A una aplicació lineal representada per una matriu $A \in \mathbb{R}^{k\times n}$. Aleshores, existeixen matrius U, Σ i V tal que:*

$$A = U\Sigma V^T,$$

on $U \in \mathbb{R}^{k\times k}$ és una matriu ortogonal (unitària si treballem amb nombres complexos), $\Sigma \in \mathbb{R}^{k\times n}$ és una matriu diagonal amb valors no negatius anomenats valors singulars, $V \in \mathbb{R}^{n\times n}$ és també una matriu ortogonal. El símbol T denota conjugada.

La descomposició SVD permet aproximar la inversa d'una matriu, fins i tot quan A no és invertible o està mal condicionada. Tal i com s'ha explicat anteriorment, per una matriu unitària:

$$U^{-1} = U^T \qquad \text{i} \qquad (V^T)^{-1} = V.$$

Així, si Σ és invertible (o pseudo-invertible), s'obté una expressió per la inversa de A com:

$$A^{-1} = V\Sigma^{-1}U^T.$$

C.4 Recomanacions d'implementació i observacions

En aquesta secció es presenten algunes recomanacions pràctiques i observacions rellevants per a la implementació del mètode numèric descrit a la Secció 4.

1. **Normalització de la transformada de Fourier:** Quan s'implementa la transformada ràpida de Fourier mitjançant paquets de funcions, és habitual que no s'inclogui el factor de normalització N descrit a (A.3) en realitzar les transformades. Cal tenir-ho en compte i aplicar-lo manualment si és necessari.

2. **Ordre dels punts de quadratura i de les transformades:** La convenció per a definir els punts de quadratura en els polinomis de Txebixev és de dreta a esquerra, tal com es detalla a l'Apèndix A.5.2. De manera anàloga, computacionalment la transformada de Fourier s'ordena començant per l'harmònic fonamental en lloc de per l'harmònic menor. Aquestes convencions són estàndard i ben establertes en l'àmbit de la simulació numèrica.

3. **Filtrat per rigidesa de matrius:** En el càlcul de les entrades de la matriu, quan s'obtenen valors petits de l'ordre de 10^{-7}, és convenient ajustar-los automàticament a zero. Aquesta pràctica millora l'esparsitat i la condició numèrica del sistema sense comprometre la convergència.

4. **Representació del terme de força:** La representació de termes externs, com ara la gravetat, pot resultar inadequada quan es considera el seu desenvolupament espectral i s'insereix dins el mètode. En aquests casos, cal afinar la formulació numèrica per assegurar una traducció correcta.

5. **Errors d'aproximació en la integració numèrica i el pas de temps:** Tot i que els mètodes d'integració d'equacions diferencials estudiats són fortament convergents, no és cert que la precisió millori indefinidament en reduir el pas de temps h.

   Quan s'implementa numèricament un mètode d'integració apareixen dos tipus d'error: (i) l'error de discretització, degut a l'aproximació de l'equació diferencial per passos discrets, i (ii) l'error d'aritmètica i d'arrodoniment propi de l'ordinador. A mesura que h es fa més petit, el nombre d'operacions augmenta i l'error d'arrodoniment esdevé dominant. Per tant, la corba que descriu l'error global de l'integrador en funció de h presenta un mínim per a un valor positiu de h, a diferència de l'error teòric, que decreix monòtonament quan $h \to 0$.

D Codis

El codi complet emprat per a la generació de les gràfiques i la realització de les simulacions es troba disponible en un repositori de GitHub associat a l'autor, accessible a l'enllaç següent: https://github.com/GuillemMasdemont/Spectral-Methods.

Aquest repositori es mantindrà disponible mentre duri el procés de revisió. Les simulacions s'han dut a terme en `Python 3.12.3` (versió del 15 d'abril de 2024), executades en un processador `AMD Ryzen 5 3500U` sota `Windows 11`.

Requeriments:

```
Python 3.12.3 | packaged by conda-forge | (main, Apr 15 2024, 18:20:11) [MSC
    v.1938 64 bit (AMD64)]

aiohttp==3.12.11
beautifulsoup4==4.13.3
ipykernel==6.29.5
ipython==8.30.0
jupyterlab==4.3.5
jsonschema==4.23.0
matplotlib==3.9.2
nbconvert==7.16.6
nbformat==5.10.4
numpy==2.1.1
pandas==2.2.3
pillow==10.4.0
plotly==6.0.1
pycryptodome==3.23.0
python-dateutil==2.9.0.post0
pyyaml==6.0.2
requests==2.32.3
scipy==1.15.2
tqdm==4.66.5
```