

Factorization of the Fourier transform of the pressure-Poisson equation using finite differences in colocated grids

Juan Pedro Mellado* and Cedrick Ansoorge

Max Planck Institute for Meteorology, Bundesstr. 53, 20146 Hamburg, Germany

Received 23 June 2011, revised 15 and 23 November 2011, accepted 27 November 2011

Published online 29 February 2012

Key words Compact finite-difference methods, solenoidal constraint, incompressible flows.

The zero-divergence constraint on the velocity field in the numerical simulation of incompressible flows can be reduced, in certain cases, to a set of one-dimensional linear difference equations for the pressure. These equations involve the second-order derivative $\delta_x \delta_x \mathbf{p}$ expressed in terms of twice the first-order derivative. When implicit finite-difference schemes are used, those equations lead to full linear systems, which are computationally prohibitive. Hence, it is a common practice to substitute $\delta_x \delta_x \mathbf{p}$ by a different discretization $\delta_{xx} \mathbf{p}$. However, it is well known that this step results in a non-zero divergence in the velocity field. This paper presents a factorization of the original equation that allows to satisfy the discrete solenoidal constraint exactly while maintaining a linear relation between the number of operations and the grid size. As an example, the method is particularized to compact schemes often found in the literature.

© 2012 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

1 Introduction

When using finite-difference methods to solve the Navier-Stokes equations in an incompressible flow, the discrete zero-divergence constraint on the velocity field leads, in certain cases, to the following linear difference equation

$$\delta_x \delta_x \mathbf{p}|_j - \lambda^2 \mathbf{p}|_j = \mathbf{f}|_j, \quad j = 2, \dots, n-1, \quad (1)$$

$\lambda \in \mathfrak{R}$. As usual, the components of the n -dimensional vectors $\mathbf{p} = (p_j)$ and $\delta_x \mathbf{p} = (p'_j)$ are, respectively, approximations to the values $\{p(x_j) : j = 1, \dots, n\}$ of a function $p(\cdot)$ and its derivative defined on a finite interval $[x_1, x_n]$. The boundary-value problem is completed with the boundary conditions

$$\left. \begin{aligned} \alpha_1 p'_1 + \alpha_0 p_1 &= \alpha \\ \beta_1 p'_n + \beta_0 p_n &= \beta \end{aligned} \right\}. \quad (2)$$

Dirichlet and/or Neumann conditions are the usual case, but a Robin type will be retained for generality. One common scenario for (1-2) occurs when all the directions except one are periodic and a discrete Fourier transformation is applied to the governing equations, obtaining then (1-2) for each Fourier mode \mathbf{p} of the pressure field. This problem can of course be obtained by discretizing directly the pressure-Poisson equation resulting from the continuous form of the solenoidal constraint [9, 11–13]. However, it is important to realize that the term $\delta_x \delta_x \mathbf{p}$ in (1) appears by the successive application of the first-order derivative operator δ_x that is used to calculate the pressure gradient and the velocity divergence, which is in general different from the discretization of a second-order derivative operator δ_{xx} [8, 15, 24].

The difficulty to solve (1-2) using implicit finite differences of the form

$$A \delta_x \mathbf{p} = (1/h) B \mathbf{p}, \quad (3)$$

where $h = (x_n - x_1)/(n-1)$ is a reference space step and the square matrices $A = (a_{ij})$ and $B = (b_{ij})$ are narrow banded [10, 23], is that the system obtained after substituting $\delta_x \delta_x \mathbf{p}|_j = \lambda^2 \mathbf{p}|_j + \mathbf{f}|_j$ from (1) into $\delta_x \delta_x \mathbf{p} = (1/h^2)(A^{-1}B)^2 \mathbf{p}$ is full and therefore the computational cost for large n is unaffordable. Several approaches are adopted to overcome this difficulty. The term $\delta_x \delta_x \mathbf{p}$ is sometimes substituted by a different discretization $\delta_{xx} \mathbf{p}$ in order to have a banded linear system easier to solve [6, 21, 24]; unfortunately, this simplification leads to a velocity field that does not satisfy the discrete solenoidal constraint, a small truncation error always remaining in the solution. In other cases, appropriate symmetries in the non-periodic direction allow an additional Fourier transformation and then high-order accuracy can be preserved [16];

* Corresponding author E-mail: juan-pedro.mellado@zmaw.de, Phone: +49 40 41173 354, Fax: +049 40 41173 430

these symmetries, however, do not always apply. This paper solves problem (1-2) exactly without resorting to any of these simplifications.

In particular, we derive new algorithms based on the factorization $\delta_x \delta_x \mathbf{p} - \lambda^2 \mathbf{p} = (\delta_x - \lambda)(\delta_x + \lambda) \mathbf{p}$ that exploit the small bandwidth of the system matrices A and B to maintain a linear relation between the number of operations and the size n of the problem. The corresponding well-posedness and convergence rates are studied thoroughly. The resulting algorithms are applicable to non-uniform grids, which is strongly desirable for near-wall regions and sponge layers. The formulation is general, but we particularize it to two compact schemes often found in the literature, one fourth-order and one sixth-order. These compact finite-difference schemes are frequently used due to their spectral-like resolution and high-order accuracy [3, 6, 16, 17, 21, 24]

The paper is organized as follows. Section 2 contains definitions and preliminary results to be used thereafter. We analyze the problem $\delta_x \mathbf{p}|_j + \lambda p_j = f_j$ in Sect. 3 before studying the complete case in Sect. 4. Section 3 includes the discussion of numerical integration ($\lambda = 0$) and is therefore interesting in its own. Section 5 validates the results using a two-dimensional laminar Taylor-Green vortex, and the three-dimensional turbulent Ekman boundary layer.

2 Preliminaries

The linear finite-difference operator $\delta_x : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $\mathbf{p}' = \delta_x \mathbf{p} = (1/h)D\mathbf{p}$, where $D = A^{-1}B$, is not invertible. The reason is that the sum of the columns of B is always zero, by construction of the finite-difference scheme. In other words, $B\mathbf{i} = 0$, where $\mathbf{i} = (1 \ 1 \ \dots \ 1)^T$ is the n -dimensional unit vector. Consequently, given any $\mathbf{q} \in \mathbb{R}^n$, there is no guarantee that a unique \mathbf{p} exists such that $\mathbf{q} = \delta_x \mathbf{p}$. Since $\text{rank}(D) = n - 1$, it is known from basic linear algebra that solution exists if and only if the rank of the extended $n \times (n + 1)$ matrix $[D \ \mathbf{q}]$ is $n - 1$. However, there is an infinite number of solutions, all of them varying in a constant vector. This indeterminacy will be resolved below by the corresponding boundary condition.

Let us introduce the block matrices

$$A = \begin{pmatrix} a_{11} & \mathbf{a}_{12}^T \\ \mathbf{a}_{21} & A_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & \mathbf{b}_{12}^T \\ \mathbf{b}_{21} & B_{22} \end{pmatrix}.$$

Splitting (3) into the first and the last $n - 1$ equations, and eliminating p'_1 , we obtain

$$hA_{22}^R \begin{pmatrix} p'_2 \\ \vdots \\ p'_n \end{pmatrix} = B_{22}^R \begin{pmatrix} p_2 \\ \vdots \\ p_n \end{pmatrix} + p_1 \mathbf{b}_{21}^R, \quad (4)$$

where the $(n - 1) \times (n - 1)$ matrices $\{A_{22}^R, B_{22}^R\}$ and the column vector $\mathbf{b}_{21}^R \in \mathbb{R}^{n-1}$ are

$$A_{22}^R = A_{22} - \frac{1}{a_{11}} \mathbf{a}_{21} \mathbf{a}_{12}^T, \quad B_{22}^R = B_{22} - \frac{1}{a_{11}} \mathbf{a}_{21} \mathbf{b}_{12}^T, \quad \mathbf{b}_{21}^R = \mathbf{b}_{21} - \frac{b_{11}}{a_{11}} \mathbf{a}_{21}. \quad (5)$$

Note that A_{22}^R and B_{22}^R have the same bandwidths as A and B , respectively. The element p'_1 can be calculated by

$$p'_1 = \frac{1}{ha_{11}} \begin{pmatrix} b_{11} & \mathbf{b}_{12}^T \end{pmatrix} \mathbf{p} - \frac{1}{a_{11}} \mathbf{a}_{12}^T \begin{pmatrix} p'_2 \\ \vdots \\ p'_n \end{pmatrix}. \quad (6)$$

A similar decomposition can be introduced by splitting (3) into the first $n - 1$ and the last equations and eliminating p'_n , defining the corresponding matrices $\{A_{11}^R, B_{11}^R\}$ and the vector \mathbf{b}_{1n}^R .

The local truncation error of (3) is defined by

$$\mathbf{t} = \frac{1}{h} B \begin{pmatrix} p(x_1) \\ \vdots \\ p(x_n) \end{pmatrix} - A \begin{pmatrix} \frac{dp}{dx}(x_1) \\ \vdots \\ \frac{dp}{dx}(x_n) \end{pmatrix}, \quad (7)$$

which yields $\boldsymbol{\epsilon} = -A^{-1}\mathbf{t}$ as the discretization errors $\{\epsilon_j = dp/dx(x_j) - p'_j : j = 1, \dots, n\}$. With these definitions, the matrix decomposition introduced before leads to the relation

$$B_{22}^R \begin{pmatrix} p(x_2) \\ \vdots \\ p(x_n) \end{pmatrix} + p(x_1) \mathbf{b}_{21}^R = hA_{22}^R \begin{pmatrix} \frac{dp}{dx}(x_2) \\ \vdots \\ \frac{dp}{dx}(x_n) \end{pmatrix} + h \left[\begin{pmatrix} t_2 \\ \vdots \\ t_n \end{pmatrix} - \frac{t_1}{a_{11}} \mathbf{a}_{21} \right]. \quad (8)$$

Table 1 Coefficients of the finite-difference formulae (9). The first three rows are centered differences, the last three are biased differences. The matrix A in (3) is constructed in terms of the coefficients $\{r_i\}$, the matrix B in terms of $\{s_i\}$. The last column contains the leading order term of the local truncation error defined by (7).

	r_{-1}	r_0	r_{+1}	s_{-2}	s_{-1}	s_0	s_{+1}	s_{+2}	t
C2	0	1	0	0	$-1/2$	0	$1/2$	0	$1/3! h^2 p^{(3)}$
C4	$1/4$	1	$1/4$	0	$-3/4$	0	$3/4$	0	$-1/5! h^4 p^{(5)}$
C6	$1/3$	1	$1/3$	$-1/36$	$-7/9$	0	$7/9$	$1/36$	$4/7! h^6 p^{(7)}$
B1	0	1	0	0	0	-1	1	0	$1/2! h p^{(2)}$
B3	0	1	2	0	0	$-5/2$	2	$1/2$	$-2/4! h^3 p^{(4)}$
B5	$1/6$	1	$1/2$	0	$-10/18$	$-1/2$	1	$1/18$	$-2/6! h^5 p^{(6)}$

As explained in Sect. 1, we particularize the method to compact schemes. We restrict ourselves to the 5-point stencils

$$r_{-1}p'_{j-1} + r_0p'_j + r_{+1}p'_{j+1} = \frac{1}{h}(s_{-2}p_{j-2} + s_{-1}p_{j-1} + s_0p_j + s_{+1}p_{j+1} + s_{+2}p_{j+2}) \tag{9}$$

given in Table 1 and discussed by Lele [17]. Global schemes (3) are constructed as a combination of n of these formulae. We define the global algorithm (35653) by using the centered scheme C6 at the $(n - 4)$ interior points, and the biased schemes B5 at $j = 2$ and B3 at $j = 1$ with the corresponding symmetric counterpart at $j = n - 1$ and $j = n$, respectively [3]. We compare it with the classical Padé scheme (343). For these two schemes, matrices A are tridiagonal, matrices B are pentadiagonal. As a reference, we also include some of the results for the explicit second-order scheme (121), for which A is the identity matrix.

A uniform grid $\{x_j = x_1 + (j - 1)h : j = 1, \dots, n\}$ has been considered in (9) and Table 1. If a non-uniform grid $\{x_j : j = 1, \dots, n\}$ is employed instead, we can define $\mathbf{x}' = (1/h)A^{-1}B\mathbf{x}$ from the mapping between the computational and the physical domains and A should be replaced by $A \text{diag}(\mathbf{x}')$, where $\text{diag}(\mathbf{x}')$ is a diagonal matrix with $\{x'_j\}$ as diagonal elements. Then, Eq. (3) and the algorithms discussed in the following sections also apply to these non-uniform grids, though specific matrix parameters like the norms and condition numbers discussed below should be appropriately recalculated.

3 First-order problem

We seek the n -tuple $\{p_j\}$ such that

$$\left. \begin{aligned} p'_j + \lambda p_j &= f_j \quad j = 2, \dots, n \\ p_1 &= \alpha \end{aligned} \right\}, \quad \lambda \geq 0, \tag{10}$$

consistently with $\mathbf{p}' = (1/h)D\mathbf{p}$; thus, there are $2n$ equations for the $2n$ unknowns $\{p_j, p'_j\}$. The case $\lambda = 0$ corresponds to the problem of numerical integration given the discretized integrand $\{f_j : j = 2, \dots, n\}$. The restriction to $\lambda > 0$ will be explained later.

3.1 Solution

The problem is solved substituting (10) into (4) to get the $n - 1$ equations

$$C_{22}^R \begin{pmatrix} p_2 \\ \vdots \\ p_n \end{pmatrix} = hA_{22}^R \begin{pmatrix} f_2 \\ \vdots \\ f_n \end{pmatrix} - \alpha \mathbf{b}_{21}^R \tag{11}$$

for the $n - 1$ unknowns $\{p_2, \dots, p_n\}$, where the system matrix is

$$C_{22}^R = B_{22}^R + \lambda hA_{22}^R. \tag{12}$$

The value of p'_1 that makes the system solvable is obtained according to (6), if needed.

The linearity of the problem allows us to express the solution as

$$\mathbf{p} = \mathbf{p}^{(0)} + \alpha \mathbf{e}^{(-)}, \quad (13)$$

where the first term is determined by the non-homogeneous part of the problem

$$\left. \begin{array}{l} \delta_x \mathbf{p}^{(0)}|_j + \lambda \mathbf{p}^{(0)}|_j = f_j \quad j = 2, \dots, n \\ \mathbf{p}^{(0)}|_1 = 0 \end{array} \right\}, \quad (14)$$

and the second term by the homogeneous part

$$\left. \begin{array}{l} \delta_x \mathbf{e}^{(-)}|_j + \lambda \mathbf{e}^{(-)}|_j = 0 \quad j = 2, \dots, n \\ \mathbf{e}^{(-)}|_1 = 1 \end{array} \right\}. \quad (15)$$

For the case $\lambda = 0$, $C_{22}^R = B_{22}^R$, and $\mathbf{e}^{(-)}$ is equal to \mathbf{i} , the unit vector. This can be easily seen from the relation $B\mathbf{i} = 0$ and the matrix decomposition introduced in Sect. 2, which leads to $-(B_{22}^R)^{-1}\mathbf{b}_{21}^R = (1 \ 1 \ \dots \ 1)^T$. For $\lambda \neq 0$, the notation $\mathbf{e}^{(-)}$ has been chosen because this n -dimensional vector is the finite-difference approximation to $\exp[-\lambda(\mathbf{x} - x_1\mathbf{i})]$, the exponential function applied on the vector of relative grid positions $\mathbf{x} - x_1\mathbf{i}$.

3.2 Round-off errors

The condition numbers $\mu(C_{22}^R) = \|C_{22}^R\| \| (C_{22}^R)^{-1} \|$ have been calculated numerically as a function of the two parameters n and $\lambda(x_n - x_1)$ for the compact schemes introduced in Sect. 2. Matrix norms, unless otherwise stated, are induced L_2 -norms derived from vector L_2 -norms

$$\|\mathbf{p}\| = \left(h \sum_{j=1}^n |p_j|^2 \right)^{1/2},$$

where the reference grid step h has been defined before [18]. The condition numbers quantify the sensitivity of the solution of a linear system to perturbations of the system matrix and the right-hand side term. Thus, they are relevant because they provide an estimate for the loss of accuracy associated with the floating-point arithmetic when solving (11) (see for instance Golub and Loan [10]). For the case $\lambda = 0$ (numerical integration), $C_{22}^R = B_{22}^R$ and μ augments proportional to the grid size n ; linear fits are summarized in Table 2 and show a variation between $\simeq 2n$ and $\simeq 35.2n$, depending on the scheme. (For comparison, condition numbers $\mu(A) = \mu(A_{22}^R)$ and norms $\|A_{22}^R\| = \|A\|$ for matrices A are also included in that table.)

Table 2 System parameters for different schemes and for different grid sizes n . See text.

	$\mu(B_{22}^R)$	$\ B_{22}^R\ $	$\mu(A)$	$\ A\ $	γ_1	γ_2
(35653)	$35.2(n-1)$	3.29	15.2	2.49	11.3	$0.231(n-1)$
(343)	$20.2(n-1)$	3.35	14.6	2.48	8.00	$0.333(n-1)$
(121)	$1.93(n-1)$	1.51	1	1	2.00	$1.000(n-1)$

The variation of the condition number of the system matrix C_{22}^R with $\lambda(x_n - x_1)$ and n is depicted in Fig. 1 in terms of the ratio $\mu(C_{22}^R)/\mu(B_{22}^R)$. This ratio is always less than 1 and therefore the condition number of C_{22}^R is bounded by that of B_{22}^R , shown in Table 2. It is seen in Fig. 1 that all of the curves coincide for relatively small values of λ , and the dependence on the grid size n appears only beyond $\lambda(x_n - x_1) \simeq (n-1)$, as the asymptotic values $\mu(A)/\mu(B_{22}^R) < 1$ are approached for $\lambda(x_n - x_1) \rightarrow \infty$. Figure 1 shows only the scheme (35653), but the rest of them exhibit the same qualitative behavior and the only quantitative differences occur for $\lambda(x_n - x_1) \gg (n-1)$, as in the case of the dependence with n . However, this region of the parameter space $(n, \lambda(x_n - x_1))$ is arguably of limited interest because a well resolved solution of (10) requires a number of points n large enough for the previous inequality not to hold. This can be understood by interpreting λ^{-1} as a length scale, so that small values of λh correspond to a well resolved homogeneous solution and reasonable maximum values are $\lambda h \simeq 1$. From the definition of the grid spacing $h = (x_n - x_1)/(n-1)$, this upper limit corresponds to the line $\lambda(x_n - x_1) \simeq (n-1)$ in the parameter space $(n, \lambda(x_n - x_1))$, indicating that for $\lambda(x_n - x_1) \gg (n-1)$ the solution is under-resolved and n should be increased.

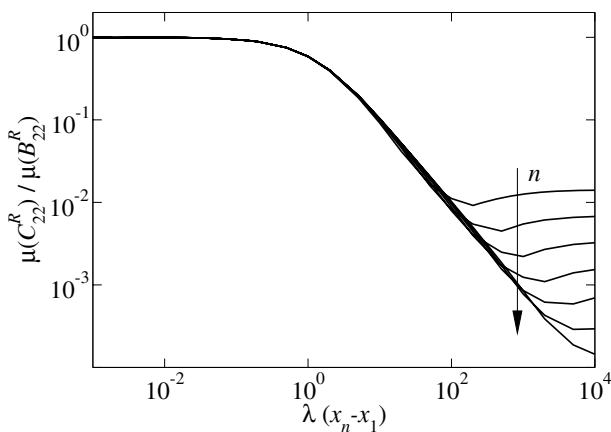


Fig. 1 Condition numbers μ of the system matrix $C_{22}^R = B_{22}^R + \lambda h A_{22}^R$ for scheme (35653), normalized with $\mu(B_{22}^R)$. Each curve corresponds to a system size $n \in \{32, 64, \dots, 2048\}$.

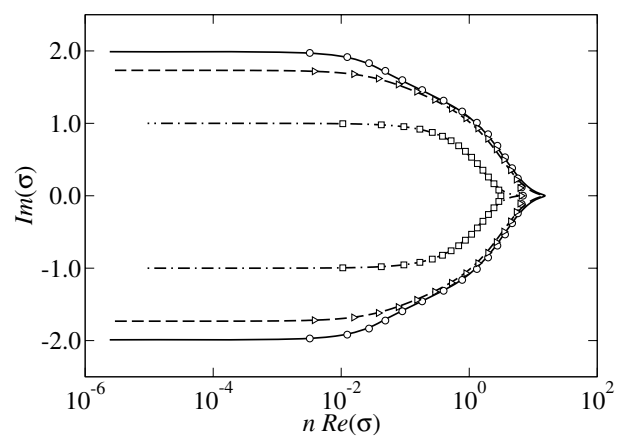


Fig. 2 Spectra of the matrix $(A_{22}^R)^{-1} B_{22}^R$: solid line, scheme (35653); dashed line, scheme (343); dot-dashed line, scheme (121); lines, grid size $n = 1024$; symbols, $n = 32$.

3.3 Discretization errors

Problem (10) is of course a discretization of the ordinary differential equation $dp/dx + \lambda p = f$ in $[x_1, x_2]$ with the boundary condition $p(x_1) = \alpha$. The corresponding discretization errors $\{\epsilon_j = p(x_j) - p_j : j = 1, \dots, n\}$ are obtained if the identities $dp/dx \equiv f - \lambda p$ and $f(x_j) \equiv f_j$ are substituted into (8) and relation (11) is used to eliminate $\{f_j\}$, which leads to $\epsilon = [p(x_1) - \alpha]e^{(-)} + \epsilon^Q$ if we define

$$\begin{pmatrix} \epsilon_2^Q \\ \vdots \\ \epsilon_n^Q \end{pmatrix} = h(C_{22}^R)^{-1} \left[\begin{pmatrix} t_2 \\ \vdots \\ t_n \end{pmatrix} - \frac{t_1}{a_{11}} \mathbf{a}_{21} \right],$$

$\epsilon_1^Q = 0$. In this expression, the truncation errors $\{t_j\}$ are given in terms of the function $p(\cdot)$ and not $f(\cdot)$, but this suffices for the purpose here of obtaining a convergence rate in the limit $h \rightarrow 0$. Using the identity $C_{22}^R = B_{22}^R [I + \lambda h (B_{22}^R)^{-1} A_{22}^R]$, an upper bound of $\|\epsilon^Q\|$ can be obtained by

$$\|\epsilon^Q\| \leq \gamma_0 \|h(B_{22}^R)^{-1} \left[\begin{pmatrix} t_2 \\ \vdots \\ t_n \end{pmatrix} - \frac{t_1}{a_{11}} \mathbf{a}_{21} \right]\|.$$

The first factor $\gamma_0 = \|[I + \lambda h (B_{22}^R)^{-1} A_{22}^R]^{-1}\|$ can be studied as a function of the grid size n and $\lambda(x_n - x_1)$ to observe that it is of order one and becomes independent of n as n increases. Hence, the order of the error $\|\epsilon^Q\|$ is determined by the second factor in the inequality above.

A bound of this second term can be obtained by introducing the set of column vectors $\{\mathbf{c}_j\}$ such that $(B_{22}^R)^{-1} = (\mathbf{c}_2 \dots \mathbf{c}_n)$, which allows us to write

$$\|\epsilon^Q\| \leq \gamma_0 h \sum_{j=2}^n |t_j - t_1 \frac{a_{j1}}{a_{11}}| \|\mathbf{c}_j\| \leq \gamma_0 \max\{\|\mathbf{c}_j\|\} h \sum_{j=2}^n |t_j - t_1 \frac{a_{j1}}{a_{11}}|. \tag{16}$$

If $\max\{\|\mathbf{c}_j\|\}$ is also bounded independently of n , then the convergence rate is determined by the last factor consisting of a linear combination of the local truncation errors $\{t_j\}$ multiplied by the space step h . This is indeed the case. Defining $\gamma_1 = \max\{\|\mathbf{c}_j\|\} / (x_n - x_1)^{1/2}$ and studying $\gamma_1(n)$ as a function of the grid size n , results indicate that γ_1 rapidly approaches a constant as n increases; asymptotic values are given in Table 2. With respect to the last summation in (16), the matrix A is narrow banded and therefore $a_{j1} = 0$ for j greater than some particular value independent of n (e.g., for the compact schemes here considered, $a_{j1} = 0$ for $j > 2$). Hence, we can group together the low-order terms from the boundary points and the rest to write

$$\|\epsilon^Q\| \leq \theta_1 h^{\gamma_1+1} + \theta_2 (n-1) h^{\gamma_2+1} / (x_n - x_1) = \theta_1 h^{\gamma_1+1} + \theta_2 h^{\gamma_2},$$

where γ_1 and γ_2 are the order of the truncation errors at the boundary and interior points, respectively. The coefficients θ_1 and θ_2 may depend on $x_n - x_1$ and $\lambda(x_n - x_1)$, but not on n . For the compact schemes (35653) and (343) considered in this study we have $4 = \gamma_1 + 1 \leq \gamma_2$, and therefore the convergence rate in the limit $h \rightarrow 0$ is 4, as validated later in Sect. 5.

For $\lambda = 0$, we have simply numerical integration [14]. Then, (14) constitutes a quadrature to calculate approximations $\mathbf{p}^{(0)}$ to the integral of $f(x)$ on the set of intervals $\{[x_1, x_j] : j = 2, \dots, n\}$, and $\{\epsilon_j^Q\}$ is the set of quadrature errors

$$\epsilon_j^Q = \int_{x_1}^{x_j} f(\xi) d\xi - \mathbf{p}^{(0)}|_j, \quad j = 2, \dots, n.$$

It is noted that this result allows to build easily well balanced schemes, i. e. schemes balancing the dominant source and divergence terms down to machine accuracy (e. g. for conditions near hydrostatic equilibrium [2]).

3.4 Further discussion

The reason for the requirement $\lambda > 0$ in (10) is now explained. We can write $C_{22}^R = A_{22}^R[(A_{22}^R)^{-1}B_{22}^R + \lambda hI]$ and the condition $\det(C_{22}^R) \neq 0$ necessary for existence and uniqueness of the solution is equivalent to the condition that $-\lambda h$ does not belong to the spectrum of the matrix $(A_{22}^R)^{-1}B_{22}^R$. Figure 2 plots the spectra of this matrix for two problem sizes n and shows that the eigenvalues have always positive real part. Hence, $\lambda > 0$ guarantees the solution of (10). If $\lambda < 0$, let us redefine the problem to

$$\left. \begin{aligned} p'_j - \lambda p_j &= f_j \quad j = 1, \dots, n-1 \\ p_n &= \beta \end{aligned} \right\}, \quad \lambda \geq 0. \quad (17)$$

Proceeding similarly to (10), we obtain

$$C_{11}^R \begin{pmatrix} p_1 \\ \vdots \\ p_{n-1} \end{pmatrix} = hA_{11}^R \begin{pmatrix} f_1 \\ \vdots \\ f_{n-1} \end{pmatrix} - \beta \mathbf{b}_{1n}^R,$$

where the system matrix is now $C_{11}^R = B_{11}^R - \lambda hA_{11}^R$. Analogously, we can express that solution as $\mathbf{p} = \mathbf{p}^{(0)} + \beta \mathbf{e}^{(+)}$, a linear combination of the non-homogeneous part, $\mathbf{p}^{(0)}$, and the contribution from the homogeneous problem

$$\left. \begin{aligned} \delta_x \mathbf{e}^{(+)}|_j - \lambda \mathbf{e}^{(+)}|_j &= 0 \quad j = 1, \dots, n-1 \\ \mathbf{e}^{(+)}|_n &= 1 \end{aligned} \right\}. \quad (18)$$

This n -dimensional vector is the finite-difference approximation to $\exp[\lambda(\mathbf{x} - x_n \mathbf{i})]$, the exponential function applied to the vector of relative grid positions. As before, if $\lambda = 0$ then $\mathbf{e}^{(+)} = \mathbf{i}$.

Last, it is observed that the definitions of $\mathbf{e}^{(-)}$ and $\mathbf{e}^{(+)}$ by (15) and (18), respectively, imply $\delta_x \mathbf{e}^{(-)}|_1 + \lambda \mathbf{e}^{(-)}|_1 \neq 0$ and $\delta_x \mathbf{e}^{(+)}|_n - \lambda \mathbf{e}^{(+)}|_n \neq 0$. It can be shown, however, that these two scalars are proportional to the truncation error, i. e. $\lambda + \delta_x \mathbf{e}^{(-)}|_1 \sim O(\|\mathbf{t}\|)$ and $\lambda - \delta_x \mathbf{e}^{(+)}|_n \sim O(\|\mathbf{t}\|)$. Therefore, they tend to zero as the space step h is reduced. This result will be used later.

4 Second-order problem

We consider now the complete problem (1-2). Because of linearity, $(\delta_x + \lambda)(\delta_x - \lambda)\mathbf{p} = \delta_x \delta_x \mathbf{p} - \lambda^2 \mathbf{p}$ for any \mathbf{p} , and the problem can be formulated in terms of the operators $\delta_x + \lambda$ and $\delta_x - \lambda$. We develop a two-step algorithm based on this factorization.

4.1 Solution

First, let us define $\mathbf{q} = (\delta_x - \lambda)\mathbf{p}$ and solve

$$\delta_x \mathbf{q}|_j + \lambda \mathbf{q}|_j = \mathbf{f}|_j, \quad j = 2, \dots, n-1$$

as a function of two free parameters. Without loss of generality we can assume $\lambda > 0$, and then $f_n = q'_n + \lambda q_n$ (note that f_n is not given) and a boundary value q_1 at $j = 1$ are a reasonable choice as free parameters because Sect. 3 has shown that such a problem is well posed. From (11) and (13) the solution can be written as

$$\mathbf{q} = \mathbf{q}^{(0)} + f_n \mathbf{q}^{(1)} + q_1 \mathbf{e}^{(-)},$$

where $\mathbf{e}^{(-)}$ is given by (15) and

$$\left. \begin{aligned} \delta_x \mathbf{q}^{(k)}|_j + \lambda \mathbf{q}^{(k)}|_j &= \mathbf{g}^{(k)}|_j \quad j = 2, \dots, n \\ \mathbf{q}^{(k)}|_1 &= 0 \end{aligned} \right\}$$

for $k \in \{0, 1\}$, having defined the auxiliary $(n-1)$ -dimensional vectors $\mathbf{g}^{(0)} = (f_2 \dots f_{n-1} 0)^T$ and $\mathbf{g}^{(1)} = (0 \dots 0 1)^T$. Each of these two problems is similar to the first-order cases discussed before in Sect. 3 and can be solved accordingly.

Second, let us solve

$$\left. \begin{aligned} \delta_x \mathbf{p}^{(k)}|_j - \lambda \mathbf{p}^{(k)}|_j &= \mathbf{q}^{(k)}|_j \quad j = 1, \dots, n-1 \\ \mathbf{p}^{(k)}|_n &= 0 \end{aligned} \right\} \quad (19)$$

for $k \in \{0, 1\}$ and define $\mathbf{s}^{(+)}$ by

$$\left. \begin{aligned} \delta_x \mathbf{s}^{(+)}|_j - \lambda \mathbf{s}^{(+)}|_j &= \mathbf{e}^{(-)}|_j \quad j = 1, \dots, n-1 \\ \mathbf{s}^{(+)}|_n &= 0 \end{aligned} \right\}. \quad (20)$$

Then, the solution to the original problem (1-2) is

$$\mathbf{p} = \mathbf{p}^{(0)} + f_n \mathbf{p}^{(1)} + q_1 \mathbf{s}^{(+)} + p_n \mathbf{e}^{(+)} \quad (21)$$

if the compatibility condition $\delta_x \mathbf{p}|_n - \lambda p_n = q_n$ is enforced by

$$f_n = - \frac{\mathbf{q}^{(0)}|_n - \delta_x \mathbf{p}^{(0)}|_n + q_1 (\mathbf{e}^{(-)}|_n - \delta_x \mathbf{s}^{(+)}|_n) + p_n (\lambda - \delta_x \mathbf{e}^{(+)}|_n)}{\mathbf{q}^{(1)}|_n - \delta_x \mathbf{p}^{(1)}|_n}. \quad (22)$$

Note that case $k = 1$ above does not depend on the forcing $\{f_j : j = 2, \dots, n-1\}$ and can be precomputed and employed for different forcings. The denominator in (22) can be written as $\mathbf{q}^{(1)}|_n - \delta_x \mathbf{p}^{(1)}|_n = h\gamma_2$ and γ_2 can be studied as a function of $(n, \lambda(x_n - x_1))$ in the same way as the condition numbers in Sect. 3. Calculations show that $0 < \gamma_2(n, \lambda(x_n - x_1)) < \gamma_2(n, 0)$ and therefore $\gamma_2 \neq 0$ and f_n is well defined. In particular, $\gamma_2(n, 0)$ is well approximated by a linear fit (shown in Table 2), and $\gamma_2(n, \lambda(x_n - x_1))$ decreases monotonically with $\lambda(x_n - x_1)$ for a fixed n towards the asymptotic behavior $[\lambda(x_n - x_1)]^{-1}$.

Finally, the boundary conditions provide the two additional equations

$$\left. \begin{aligned} \alpha_1 q_1 + (\alpha_0 + \alpha_1 \lambda) p_1 &= \alpha \\ \beta_1 q_n + (\beta_0 + \beta_1 \lambda) p_n &= \beta \end{aligned} \right\},$$

which written in terms of the unknowns $\{q_1, p_n\}$ yield

$$\begin{aligned} \begin{pmatrix} \alpha_1 + (\alpha_0 + \alpha_1 \lambda) \mathbf{s}^{(+)}|_1 & (\alpha_0 + \alpha_1 \lambda) \mathbf{e}^{(+)}|_1 \\ \beta_1 \mathbf{e}^{(-)}|_n & \beta_0 + \beta_1 \lambda \end{pmatrix} \begin{pmatrix} q_1 \\ p_n \end{pmatrix} \\ = \begin{pmatrix} \alpha - (\alpha_0 + \alpha_1 \lambda) [\mathbf{p}^{(0)}|_1 + f_n \mathbf{p}^{(1)}|_1] \\ \beta - \beta_1 [\mathbf{q}^{(0)}|_n + f_n \mathbf{q}^{(1)}|_n] \end{pmatrix}. \end{aligned} \quad (23)$$

4.1.1 Case $\lambda = 0$

For $\lambda = 0$, it is $\mathbf{e}^{(+)} = \mathbf{e}^{(-)} = \mathbf{i}$. In addition, the solution to (20) can be calculated explicitly, obtaining $\mathbf{s}^{(+)} = \mathbf{x} - x_n \mathbf{i}$. Hence, (21) simplifies to

$$\mathbf{p} = \mathbf{p}^{(0)} + f_n \mathbf{p}^{(1)} + q_1 (\mathbf{x} - x_n \mathbf{i}) + p_n \mathbf{i} \quad (24)$$

and the compatibility condition (22), which is now simply $\delta_x \mathbf{p}|_n = q_n$, reduces to

$$f_n = - \frac{\mathbf{q}^{(0)}|_n - \delta_x \mathbf{p}^{(0)}|_n}{\mathbf{q}^{(1)}|_n - \delta_x \mathbf{p}^{(1)}|_n}. \quad (25)$$

This result implies that f_n depends on $\{f_j : j = 2, \dots, n-1\}$ through the numerator but it is independent of the boundary conditions. Once f_n has been calculated, the two unknowns $\{q_1, p_n\}$ are determined by (23). For a unique solution of this system to exist, the boundary conditions need to satisfy

$$\begin{vmatrix} \alpha_1 - \alpha_0(x_n - x_1) & \alpha_0 \\ \beta_1 & \beta_0 \end{vmatrix} \neq 0.$$

This result is precisely the generalized Wronskian condition required to guarantee the existence and uniqueness of the solution of the problem in the continuous form [1] (the two fundamental solutions are $p_1(x) = 1$ and $p_2(x) = x$). Hence, if the continuous problem is well posed, so is the discrete form (1-2) for $\lambda = 0$.

In the case of the application of this methodology to obtain the pressure for the Fourier modes $\lambda = 0$, we are normally interested in the case $\alpha_0 = \beta_1 = 0$, that is, Neumann and Dirichlet boundary conditions at $j = 1$ and $j = n$, respectively. Then, p_n and q_1 are explicitly known and problem (1-2) for $\lambda = 0$ is solved by (24) and (25). In the symmetric case, if q_n and p_1 are given, note that we could have proceed differently; we could have started solving for \mathbf{q} taking q_n and f_1 as free parameters, and proceed consistently to end up with the three unknowns $\{f_1, q_n, p_1\}$ and obtain explicitly the solution.

4.1.2 Case $\lambda \neq 0$

In this case, there is no explicit expression for $\mathbf{s}^{(+)}$ and it needs to be calculated numerically from (20) for each λ . However, for the purpose of analysis, the following expression derived in Appendix A proves useful:

$$2\lambda\mathbf{s}^{(+)} = \mathbf{e}^{(-)}|_n \mathbf{e}^{(+)} - \mathbf{e}^{(-)} + (\lambda + \delta_x \mathbf{e}^{(-)}|_1) \begin{pmatrix} h(C_{11}^R)^{-1} & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}. \quad (26)$$

From Sect. 3 we know that $\lambda + \delta_x \mathbf{e}^{(-)}|_1 \sim O(\|\mathbf{t}\|)$ and that $h(C_{11}^R)^{-1}$ is uniformly bounded (independently of n). Thus, the last term in the right-hand side of (26) is of the order of the truncation error, that is, negligibly small as the space step h is reduced.

A second difference compared to the previous case $\lambda = 0$ is that condition (22), which makes the system $(\delta_x - \lambda)\mathbf{p} = \mathbf{q}$ solvable, depends both on the forcing term $\{f_j\}$ and on the boundary conditions through p_n and q_1 . However, the determinant Δ of the linear system (23),

$$2\lambda\Delta = (\alpha_1\lambda - \alpha_0)(\beta_1\lambda + \beta_0) - (\alpha_1\lambda + \alpha_0)(\beta_1\lambda - \beta_0)\mathbf{e}^{(+)}|_1 \mathbf{e}^{(-)}|_n + (\alpha_1\lambda + \alpha_0)(\beta_1\lambda + \beta_0)O(\|\mathbf{t}\|),$$

where (26) has been used, can also be related to the generalized Wronskian W of the continuous problem,

$$2\lambda W = (\alpha_1\lambda - \alpha_0)(\beta_1\lambda + \beta_0) - (\alpha_1\lambda + \alpha_0)(\beta_1\lambda - \beta_0) \exp[-2\lambda(x_n - x_1)],$$

to prove that $\Delta = W + O(\|\mathbf{t}\|)$ (the two fundamental solutions are $p_1(x) = \exp[\lambda(x - x_n)]$ and $p_2(x) = \exp[-\lambda(x - x_1)]$). Therefore, if the continuous problem is well posed, there always exists a small enough h for which (23) can be solved to obtain

$$q_1 = a_0 + f_n a_1, \quad p_n = b_0 + f_n b_1,$$

for some particular values $\{a_0, a_1, b_0, b_1\}$. Hence, f_n is given explicitly by

$$f_n = -\frac{\mathbf{q}^{(0)}|_n - \delta_x \mathbf{p}^{(0)}|_n + a_0(\mathbf{e}^{(-)}|_n - \delta_x \mathbf{s}^{(+)}|_n) + b_0(\lambda - \delta_x \mathbf{e}^{(+)}|_n)}{\mathbf{q}^{(1)}|_n - \delta_x \mathbf{p}^{(1)}|_n + a_1(\mathbf{e}^{(-)}|_n - \delta_x \mathbf{s}^{(+)}|_n) + b_1(\lambda - \delta_x \mathbf{e}^{(+)}|_n)}. \quad (27)$$

In the usual case of Neumann boundary conditions for the pressure (resulting from the no-penetration constraint on the velocity field at the boundary), it is $\alpha_1 = \beta_1 = 1$ and $\alpha_0 = \beta_0 = 0$, and the coefficients $\{a_0, a_1, b_0, b_1\}$ are

$$a_0\Delta = [\alpha - \lambda \mathbf{p}^{(0)}|_1 + \mathbf{e}^{(+)}|_1(\mathbf{q}^{(0)}|_n - \beta)]\lambda,$$

$$a_1\Delta = (\mathbf{q}^{(1)}|_n \mathbf{e}^{(+)}|_1 - \lambda \mathbf{p}^{(1)}|_1)\lambda,$$

$$b_0\Delta = (\lambda \mathbf{p}^{(0)}|_1 - \alpha)\mathbf{e}^{(-)}|_n + (\beta - \mathbf{q}^{(0)}|_n)(1 + \lambda \mathbf{s}^{(+)}|_1),$$

$$b_1\Delta = \lambda \mathbf{p}^{(1)}|_1 \mathbf{e}^{(-)}|_n - \mathbf{q}^{(1)}|_n(1 + \lambda \mathbf{s}^{(+)}|_1).$$

where the determinant satisfies $2\Delta/\lambda = 2(1 - \mathbf{e}^{(+)}|_1 \mathbf{e}^{(-)}|_n) + 2\lambda \mathbf{s}^{(+)}|_1 = 1 - \mathbf{e}^{(+)}|_1 \mathbf{e}^{(-)}|_n + O(\|\mathbf{t}\|) \neq 0$, the last equality following from (26).

4.2 Error analysis

Regarding the discretization error, the two-step algorithm discussed above is a linear combination of first-order problems, whose error was studied in Sect. 3.3. Hence, the order of the approximation to the solution of the continuous problem associated to (1-2) is expected to be the same as that obtained previously for the first-order problem. In particular, a convergence rate $O(h^4)$ in the limit $h \rightarrow 0$ is predicted for the compact schemes (35653) and (343) considered in this study. This result is confirmed later in Sect. 5.

On the contrary, the loss of accuracy associated with floating-point arithmetic increases with respect to the first-order problem. An estimate of this round-off error ϵ^R when solving a linear system like (11) is provided by $\epsilon^R \simeq \mu(C_{22}^R)(\epsilon_1^R + \epsilon_2^R)$ in terms of the condition number $\mu(C_{22}^R)$ and relative errors of the system matrix itself, ϵ_1^R , and the forcing term, ϵ_2^R [10]. In Sect. 3.2, the condition number was observed to vary with the grid size as $\mu(C_{22}^R) \simeq n$ (see Fig. 1 and Table 2). Therefore, for first-order problems, a loss of precision of the order of $n\epsilon$ is introduced, where ϵ is machine accuracy (of the order of 10^{-16} using an eight-byte representation). However, the second-order problem accumulates an error of the order of $n^2\epsilon$ because the right-hand side term in the second step of the factorization, Eq. (19), contains already a round-off error of order $\epsilon_2^R \simeq n\epsilon$.

For large grid sizes n , this result sets an upper limit to the advantage of the two-step algorithm presented here compared to the substitution of the operator $\delta_x\delta_x$ by an approximation δ_{xx} , one of the simplifications mentioned in the introduction. The reason is that the latter only includes a round-off error of order $n\epsilon$ because it only needs to solve one single linear system instead of two of them sequentially. However, for typical values of n of the order of 10^3 or 10^4 and using an eight-byte representation of floating-point numbers, the factorization procedure is still better than the approximation δ_{xx} . For instance, Mellado et al. [20] investigates the errors associated with the simplification $\delta_x\delta_x \simeq \delta_{xx}$ for typical resolutions in direct numerical simulations of free convection using the compact schemes considered in this work. Results show that the discretization errors in the divergence constraint, measured relative to other non-zero quantities of the velocity-gradient tensor, like the enstrophy, are in the range $10^{-8} - 10^{-4}$. The round-off errors obtained with the factorization approach discussed above reach the lower bound of this interval, 10^{-8} , for grid sizes of the order of $n \simeq 10^4$. Below this cross-over value, the two-step algorithm presented in this paper is more accurate. This is a relatively large number of points for three-dimensional simulations of turbulent flows, where the computational domain in the other two periodic directions tends to be significantly larger due to requirements in statistical convergence and de-correlation lengths, and show the advantages of the factorization technique for the current and near-future computational resources.

5 Validation

The governing equations in a frame of reference rotating around Ox_3 are

$$\frac{\partial u_i}{\partial t} = -u_k \frac{\partial u_i}{\partial x_k} + \text{Ro}^{-1} \epsilon_{ik3} u_k + \text{Re}^{-1} \frac{\partial^2 u_i}{\partial x_k \partial x_k} - \frac{\partial p}{\partial x_i}, \quad (28)$$

$$\frac{\partial u_k}{\partial x_k} = 0, \quad (29)$$

where (u_1, u_2, u_3) is the velocity vector, p is the kinematic pressure, Re is the Reynolds number and Ro is the Rossby number. The Einstein summation convention is used and ϵ_{ijk} is the permutation symbol. Periodicity is assumed in the directions Ox_1 and Ox_2 , and then the numerical solution of these equations leads to problem (1-2) for the variable p . We briefly review this result for completeness before discussing the simulations in Sect. 5.1 and Sect. 5.2.

The spatial derivatives are discretized using the sixth-order compact schemes discussed in Sect. 2; the periodic form is applied along Ox_1 and Ox_2 and the biased scheme (35653) along Ox_3 . The advancement in time is performed with a low-storage five-stage, fourth-order Runge-Kutta scheme, so that stage M can formally be written as

$$(U_i^{M+1} - U_i^M) / (\zeta^{M+1} \Delta t) = H_i^M - G_i P^M. \quad (30)$$

The vector U_i contains the velocity component along Ox_i at all the interior grid points, and the vector P contains the pressure at all the grid points, including those at the boundary. The coefficients ζ^{M+1} are not important for the current discussion and can be found elsewhere [4, 24]. The matrix G_i is the discretization of the partial derivative along the direction Ox_i , and H_i contains the rest of the right-hand side of (28) and includes additional Runge-Kutta coefficients. The discrete form of the solenoidal constraint (29) is

$$G_1 U_1^{M+1} + G_2 U_2^{M+1} + G_3 U_3^{M+1} = 0. \quad (31)$$

These equations are imposed at the interior points and need to be complemented by boundary conditions at the boundary planes $j = 1$ and $j = n$.

Substituting (30) into (31), a system of equations for P is obtained. A discrete Fourier transformation \mathcal{F} along Ox_1 and Ox_2 applied to that equation yields

$$\delta_3 \delta_3 \hat{\mathbf{p}}^M|_j - (\lambda_1^2 + \lambda_2^2) \hat{\mathbf{p}}^M|_j = \hat{\mathbf{f}}|_j, \quad j = 2, \dots, n-1, \quad (32)$$

for each of the horizontal Fourier modes (the corresponding sub-indexes have been dropped for notational convenience). The n -dimensional vector in the right-hand side is $\hat{\mathbf{f}} = (\hat{f}_1 \dots \hat{f}_n)^T = \mathcal{F}(G_3 H_3^M + G_2 H_2^M + G_1 H_1^M)$, where n is the number of grid points along Ox_3 (the residual dilatation at stage M can be retained in this term, if desired). The coefficients λ_1 and λ_2 are the modified wavenumbers associated to the finite-difference discretization of the first-order derivative along the directions Ox_1 and Ox_2 , that is $\mathcal{F}(G_1 P) = i\lambda_1 \hat{\mathbf{p}}$ and $\mathcal{F}(G_2 P) = i\lambda_2 \hat{\mathbf{p}}$, where $i = \sqrt{-1}$. These modified wavenumbers characterize the spectral accuracy of the finite-difference scheme and are known explicitly as a function of the wavenumber itself [17, 21].

Splitting (32) into the real and imaginary parts, Eq. (1) is recovered. With respect to the boundary conditions for $\hat{\mathbf{p}}$, different approaches can be found in the literature (a comprehensive discussion is presented in Gresho and Sani [12] and Guermond et al. [13]). We assume $U_3 = 0$ at the boundary planes (no penetration), which implies $\{\hat{\mathbf{u}}_3|_j = 0 : j = 1, n\}$, and impose the normal momentum equation at those boundary nodes, which yields $\{\delta_3 \hat{\mathbf{p}}|_j = \hat{\mathbf{H}}_3|_j : j = 1, n\}$. We obtain thereby the problem (1-2) studied in this paper. In particular, if $\lambda_1 \neq 0$ or $\lambda_2 \neq 0$, then (32) with this set of Neumann boundary conditions corresponds to the case studied in Sect. 4.1.2.

On the other hand, if the modified wavenumbers are both zero (that is, $\lambda_1 = \lambda_2 = 0$, which is the case for 2 modes in two dimensions, 4 modes in three dimensions), then the Fourier transform of the zero-divergence constraint (31)

$$i\lambda_1 \hat{\mathbf{u}}_1^{M+1} + i\lambda_2 \hat{\mathbf{u}}_2^{M+1} + \delta_3 \hat{\mathbf{u}}_3^{M+1} = 0 \quad (33)$$

implies $\delta_3 \hat{\mathbf{u}}_3|_j = 0$ at the interior points, which together with the no-penetration boundary conditions leads to $\hat{\mathbf{u}}_3 = 0$ for these Fourier modes. Hence, the vertical momentum equation provides the first-order equation $\delta_3 \hat{\mathbf{p}}|_j = \hat{\mathbf{H}}_3|_j$ for the corresponding pressure mode $\hat{\mathbf{p}}$ at the interior points $j = 2, \dots, n-1$. With respect to the two boundary conditions $\{\delta_3 \hat{\mathbf{p}}|_j = \hat{\mathbf{H}}_3|_j : j = 1, n\}$, one of them simply extends this first-order equation to the corresponding boundary, whereas the other boundary condition needs to be substituted by a Dirichlet one. One of these Dirichlet boundary conditions defines the mean pressure, and is irrelevant because only the gradient of p is important for the evolution of the flow. The rest of the cases $\lambda_1 = \lambda_2 = 0$ represent part of the $2\Delta x_1$ and $2\Delta x_2$ plane waves, the smallest wavelengths supported in the grid. Although present in the pressure field, these Fourier modes do not enter either in the evolution of the flow because the discrete form of the gradient $G_i P$ sets them precisely equal to zero (the modified wavenumbers λ_1 and λ_2 corresponding to those Fourier modes are zero). Thus, we use the boundary condition $\hat{\mathbf{p}}|_1 = 0$ without loss of generality, and cases $\lambda_1 = \lambda_2 = 0$ correspond to the class of problems studied in Sect. 3.

5.1 Laminar Taylor-Green vortex

We consider first a Taylor-Green vortex ($\text{Ro}^{-1} = 0$) in a two-dimensional domain $(0, 1) \times (0, 1/2)$,

$$u_1 = \sin(2\pi x_1) \cos(2\pi x_2) f(t), \quad u_2 = -\cos(2\pi x_1) \sin(2\pi x_2) f(t), \\ p = 1/4[\cos(4\pi x_1) + \cos(4\pi x_2)][f(t)]^2,$$

where the temporal variation is given by $f(t) = \exp(-8\pi^2 \text{Re}^{-1} t)$. The boundary conditions are periodic in Ox_1 and, along the direction Ox_2 , homogeneous Dirichlet in u_2 (no penetration) and homogeneous Neumann in u_1 (free slip). Two cases are studied: (1) $\text{Re} = 8\pi^2 \times 100$ integrated over the turnover time $t = 1.0$, dominated by the advection terms; (2) $\text{Re} = 8\pi^2/10$ integrated over the viscous e -folding time $t = 0.1$, dominated by the viscous terms.

Figure 3 shows the numerical error in the pressure field as a function of the space step $h = 1/(n-1)$ for these two cases. At the initial time, shown in panel (a), a fourth-order accuracy is exhibited, as obtained theoretically in Sect. 3 and Sect. 4. After the solution is advanced in time, panel (b), the convergence rate for the case (1) remains the same, but that of the viscous-dominated case (2) decreases to three. This result is also to be expected because the order of accuracy of the derivative terms in the right-hand side of the transport equations is three at the boundaries (see Sect. 2). The normal velocity at the boundary is zero and then $u_3 \partial u_i / \partial x_3 = 0$, reducing the local boundary error in the advection term below the remaining fourth-order contributions in the advection-dominated case. On the other hand, if the viscous term controls the error, the third-order convergence rate is maintained, as observed in Fig. 3 (b).

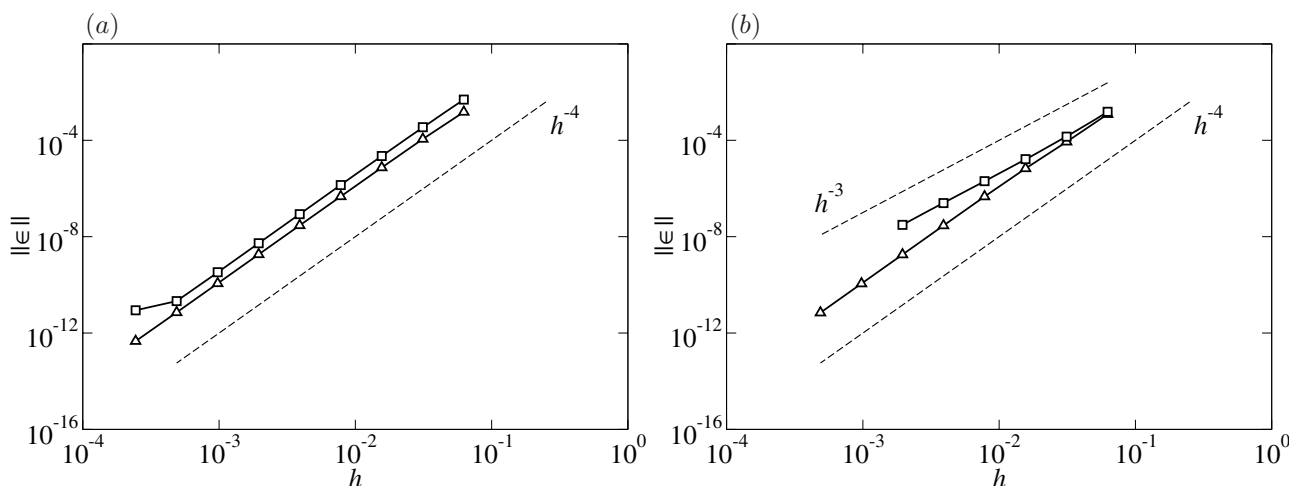


Fig. 3 Simulations of the laminar Taylor-Green vortex: triangles, advection-dominated case (1); squares, viscous-dominated case (2). Panel (a) corresponds to $t = 0$ and panel (b) to times $t = 1.0$ (advection-dominated) and $t = 0.1$ (viscous-dominated).

5.2 Turbulent Ekman layer

The Ekman boundary layer over a smooth surface is considered to demonstrate the applicability of the algorithm to a three-dimensional turbulent flow using grid stretching. This particular flow is commonly employed to study different aspects of the atmospheric boundary layer [22]. The direction Ox_1 is chosen along the mean horizontal geostrophic wind so that far above the surface, where the Coriolis force and the mean pressure gradient are in balance, $\mathbf{u} \rightarrow (1, 0, 0)$. At the top of the domain, no-penetration and free-slip boundary conditions are used. At the bottom $x_3 = 0$, $\mathbf{u} = \mathbf{0}$ applies. Periodicity is imposed along the two horizontal directions Ox_1 and Ox_2 . Equations are non-dimensionalized such that $\text{Ro}^{-1} = 1$ and the Reynolds number is set to $\text{Re} = 1.25 \times 10^5$ in order to compare with the results from Coleman et al. [5].

A broadband random initial condition is defined with a Gaussian energy spectrum peaking at the most unstable mode of the laminar Ekman layer [19]. After an initial transient of the order of the inertial period $2\pi/\text{Ro}^{-1}$, a turbulent boundary layer with a constant mean thickness of about $0.7 u_*/\text{Ro}^{-1}$ is established, where u_* is the friction velocity normalized by the geostrophic wind velocity. The system is statistically homogeneous along Ox_1 , Ox_2 and time.

The grid is $256 \times 256 \times 192$, uniform in the horizontal directions with a space step approximately equal to 4 viscous lengths Re^{-1}/u_* . The grid is stretched along the vertical direction Ox_3 using a hyperbolic tangent to map the computational and the physical domains. Following the literature, the first 10 points are placed within $10 \text{Re}^{-1}/u_*$ from the surface. The vertical space step Δx_3 is about $0.55 \text{Re}^{-1}/u_*$ at the wall, where the stretching factor $(\Delta x_3)_{j+1}/(\Delta x_3)_j$ reaches its maximum of about 1.06. The vertical step augments then continuously and equals the horizontal one at a height of the order of $z \simeq 0.4 u_*/\text{Ro}^{-1}$, where the stretching factor has decreased to about 1.01.

Statistics have been calculated averaging within the horizontal planes and then along 7 inertial periods. The non-dimensional friction velocity u_* and the angle between the shear stress at the surface and the geostrophic wind are 0.062 and 25.7° , respectively, in very good agreement with the values 0.0627 and 25.4° reported by Coleman et al. [5] using a spectral algorithm. The same result is observed in the mean velocities $\langle u_1 \rangle$ and $\langle u_2 \rangle$, presented in Fig. 4 (a) in terms of the hodograph, and confirms further the accuracy and validity of the methodology discussed in this paper.

The pressure field has also been investigated in detail in order to verify that the results are smooth and there is no even-odd decoupling. This is a known deficiency of explicit second-order centered finite differences using collocated variables, the so-called checkerboard problem, and it is normally solved with staggered grids [8, 9]. However, compact schemes like the one used here do not suffer from this problem [7]. Figure 4 (b) plots the root-mean-square profiles of the pressure field and a smooth variation is observed over two orders of magnitude (note the logarithmic scale). Instantaneous profiles are included at particular times in intervals of approximately 2 inertial periods to show that a smooth curve is obtained independently of the temporal average. These results demonstrate that the pressure is indeed well-behaved and there is no unphysical small-scale oscillation; visual inspection of instantaneous fields (not shown) also confirms this behavior.

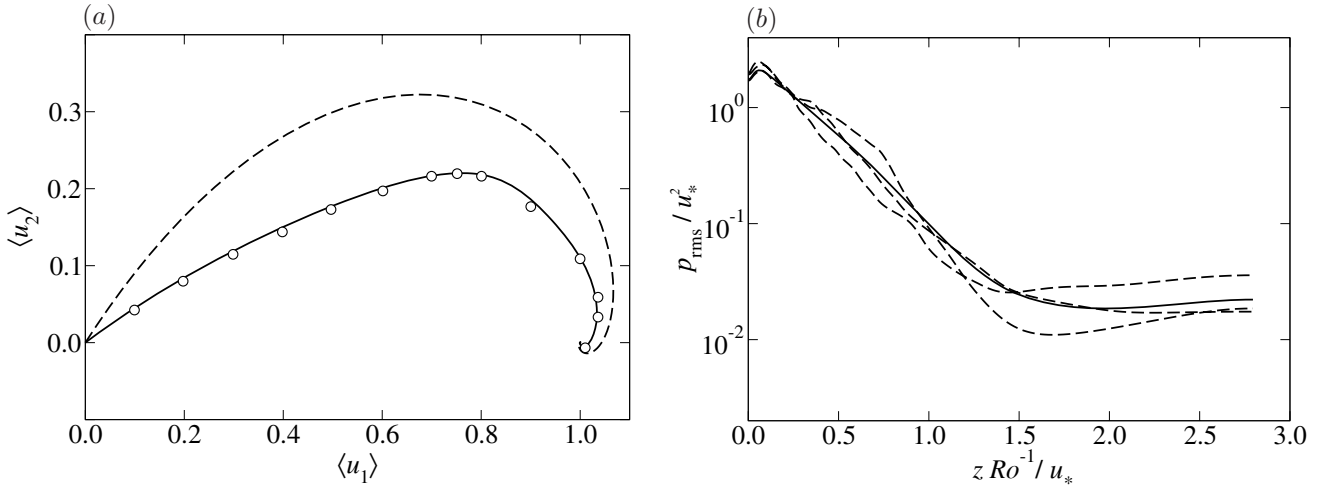


Fig. 4 Simulations of the turbulent Ekman layer. Panel (a) shows the velocity hodograph: solid, turbulent case; dashed, laminar case; symbols, from Coleman et al. [5]. Panel (b) shows the pressure root-mean-square as a function of the distance to the wall: solid, total average; dashed, plane average at different times.

6 Summary

A factorization of the linear difference equation $\delta_x \delta_x \mathbf{p}|_j - \lambda^2 \mathbf{p}|_j = \mathbf{f}|_j$ resulting from the discretization of the solenoidal constraint on the velocity field using implicit finite-difference schemes $A \delta_x \mathbf{p} = (1/h) B \mathbf{p}$ has been discussed. This formulation allows to satisfy the zero-divergence condition down to round-off error in collocated meshes, both uniform and non-uniform. It involves only the solution of narrow banded linear systems, which maintains a linear relation between the number of operations and the problem size. We have studied the solvability constraint required by the singular character of the matrix B , using the boundary conditions to resolve the corresponding indeterminacy. It has been proven that the resulting system of equations is well posed. The first-order problem $\delta_x \mathbf{p}|_j + \lambda p_j = f_j$ has been analyzed as an intermediate step. This case includes quadratures for numerical integration ($\lambda = 0$). Global discretization errors are of order $\min\{\gamma_1 + 1, \gamma_2\}$, where γ_1 is the order of the truncation error of the boundary finite-difference formula and γ_2 is that of the interior points. The work has been applied to two compact schemes commonly found in the literature. For these schemes, a fourth-order accuracy is achieved. Round-off errors have been discussed in terms of the condition numbers of the corresponding matrices. It has been shown that, for the first-order problem, this error is proportional to n , the grid size, and for the second-order problem, proportional to n^2 . Using this result and previous work on direct numerical simulations of turbulent flows, we have estimated that the cross-over size n beyond which round-off errors in the factorization technique presented here exceed discretization errors in simplified approaches is about 10^4 .

Acknowledgements Support from the Max Planck Society through its Max Planck Research Groups program is gratefully acknowledged. Computational time was provided by the Jülich Supercomputer Centre. This research was supported in part by the National Science Foundation under Grant No. NSF PHY05-51164.

A Proof of Eq. (26)

Equation (15) can be written as

$$\delta_x \mathbf{e}^{(-)} + \lambda \mathbf{e}^{(-)} = (\lambda + \delta_x \mathbf{e}^{(-)}|_1) \mathbf{g}$$

introducing the n -dimensional auxiliary vector $\mathbf{g} = (1 \ 0 \ \dots \ 0)^T$. Substituting $(1/2)\mathbf{e}^{(-)}$ as obtained from this relation into (20) yields

$$\delta_x \mathbf{s}^{(+)}|_j - \lambda \mathbf{s}^{(+)}|_j = \frac{1}{2} \mathbf{e}^{(-)}|_j + \frac{1}{2\lambda} \left[-\delta_x \mathbf{e}^{(-)}|_j + (\lambda + \delta_x \mathbf{e}^{(-)}|_1) \mathbf{g}|_j \right], \quad j = 1, \dots, n-1,$$

and reorganizing the different terms

$$\delta_x \left(2\lambda \mathbf{s}^{(+)} + \mathbf{e}^{(-)} \right)|_j - \lambda \left(2\lambda \mathbf{s}^{(+)} + \mathbf{e}^{(-)} \right)|_j = (\lambda + \delta_x \mathbf{e}^{(-)}|_1) \mathbf{g}|_j, \quad j = 1, \dots, n-1.$$

We have thereby written an equation for $2\lambda s^{(+)} + e^{(-)}$ of the kind discussed in Sect. 3; the boundary condition is $2\lambda s^{(+)}|_n + e^{(-)}|_n = e^{(-)}|_n$. The solution is given by

$$2\lambda s^{(+)} + e^{(-)} = e^{(-)}|_n e^{(+)} + (\lambda + \delta_x e^{(-)}|_1) \mathbf{w}$$

where \mathbf{w} satisfies

$$\left. \begin{array}{l} \delta_x \mathbf{w}|_j - \lambda \mathbf{w}|_j = \mathbf{g}|_j \quad j = 1, \dots, n-1 \\ \mathbf{w}|_n = 0 \end{array} \right\}.$$

Explicitly, the solution for this last problem is $(w_1 \dots w_{n-1})^T = h(C_{11}^R)^{-1} A_{11}^R (1 \ 0 \dots 0)^T$. Finally, realizing that the first column of matrices A_{11}^R and A_{11} are the same, it is $A_{11}^R (1 \ 0 \dots 0)^T = (a_{11} \dots a_{(n-1)1})^T$ and (26) follows. The notation $s^{(+)}$ is used because this vector is related to the hyperbolic sine function, in particular, it is an approximation to $\sinh[\lambda(\mathbf{x} - x_n \mathbf{i})] / \{\lambda \exp[\lambda(x_n - x_1)]\}$.

References

- [1] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (Springer, Berlin, Heidelberg, New York, 1999).
- [2] N. Botta, R. Klein, S. Langenberg, and S. Lützenkirchen, Well balanced finite volume methods for nearly hydrostatic flows, *J. Comput. Phys.* **196**, 539–565 (1994).
- [3] M. H. Carpenter, D. Gottlieb, and S. Abarbanel, The stability of numerical boundary treatments for compact high-order finite-difference schemes, *J. Comput. Phys.* **108**, 272–295 (1993).
- [4] M. H. Carpenter and C. A. Kennedy, Fourth-Order 2N-storage Runge-Kutta Schemes, Tech. Rep. TM-109112 (NASA Langley Research Center, Langley, 1994).
- [5] G. N. Coleman, J. H. Ferziger, and P. R. Spalart, A numerical study of the turbulent Ekman layer, *J. Fluid Mech.* **213**, 313–348 (1990).
- [6] A. W. Cook and P. E. Dimotakis, Transition stages of Rayleigh-Taylor instability between miscible fluids, *J. Fluid Mech.* **443**, 69–99 (2001).
- [7] E. Dormy, An accurate compact treatment of pressure for colocated variables, *J. Comput. Phys.* **151**, 676–683 (1999).
- [8] J. H. Ferziger and M. Perić, *Computational Methods for Fluid Dynamics* (Springer, Berlin, Heidelberg, New York, 2002).
- [9] C. A. J. Fletcher, *Computational Techniques for Fluid Dynamics* (Springer, Berlin, Heidelberg, New York, 1988).
- [10] G. H. Golub and C. F. V. Loan, *Matrix Computations* (Johns Hopkins University Press, Baltimore, 1983).
- [11] P. M. Gresho, Incompressible fluid dynamics: some fundamental formulation issues, *Annu. Rev. Fluid Mech.* **23**, 413–453 (1991).
- [12] P. M. Gresho and R. L. Sani, On pressure boundary conditions for the incompressible Navier-Stokes equations, *Int. J. Numer. Methods Fluids* **7**, 1111–1145 (1987).
- [13] J. L. Guermond, P. Mineev, and J. Shen, An overview of projection methods for incompressible flows, *Comput. Methods Appl. Mech. Engrg.* **195**, 6011–6045 (2006).
- [14] E. Isaacson and H. B. Keller, *Analysis of Numerical Methods* (John Wiley & Sons, New York, 1966).
- [15] L. Kleiser, C. Härtel, and T. Wintergerste, There is no error in the Kleiser-Schumann influence matrix method, *J. Comput. Phys.* **131**, 85–87 (1998).
- [16] S. Laizet and E. Lamballais, High-order compact schemes for incompressible flows: A simple and efficient method with quasi-spectral accuracy, *J. Comput. Phys.* **228**, 5989–6015 (2009).
- [17] S. K. Lele, Compact finite difference schemes with spectral-like resolution, *J. Comput. Phys.* **103**, 16–42 (1992).
- [18] R. J. LeVeque, *Finite Difference Methods for Ordinary and Partial Differential Equations* (SIAM, PA, 2007).
- [19] D. K. Lilly, On the instability of Ekman boundary flows, *J. Atmos. Sci.* **23**, 481–494 (1966).
- [20] J. P. Mellado, B. Stevens, H. Schmidt, and N. Peters, Buoyancy reversal in cloud-top mixing layers, *Q. J. R. Meteorol. Soc.* **135**, 963–978 (2009).
- [21] J. P. Mellado, B. Stevens, H. Schmidt, and N. Peters, Two-fluid formulation of the cloud-top mixing layer for direct numerical simulation, *Theor. Comput. Fluid Dyn.* **24**, 511–536 (2010).
- [22] R. B. Stull, *An Introduction to Boundary Layer Meteorology* (Kluwer Academic Publishers, Dordrecht, 1988).
- [23] R. Vichnevetsky and J. B. Bowles, *Fourier Analysis of Numerical Approximations of Hyperbolic Equations* (SIAM, PA, 1982).
- [24] R. V. Wilson, A. O. Demuren, and M. Carpenter, Higher-Order Compact Schemes for Numerical Simulation of Incompressible Flows, Tech. Rep. CR-1998-206922 (NASA Langley Research Center, Langley, 1998).