On the Interpolation Problem for the Poisson Equation on Collocated Meshes

Daniel Santos, Jordi Muela, Nicolas Valle, F. Xavier Trias

ON THE INTERPOLATION PROBLEM FOR THE POISSON EQUATION ON COLLOCATED MESHES

D. Santos[1], J. Muela[1], N. Valle[1], F. X. Trias[1]

[1] Heat and Mass Transfer Technological Center,
Technical University of Catalonia, C/Colom 11, 08222 Terrassa (Barcelona)
daniel.santos.serrano@upc.edu
jordi.muela@upc.edu
nicolas.valle@upc.edu
francesc.xavier.trias@upc.edu

Key words: Symmetry-preserving discretization, collocated formulation, structured grid, Poisson equation, stability, interpolator, FSM

Abstract. The appearence of unphysical velocities in highly distorted meshes is a common problem in many simulations. In collocated meshes, this problem arises from the interpolation of the pressure gradient from faces to cells. Using an algebraic form for the classical incompressible Navier-Stokes equations, this problem is adressed. Starting from the work of F. X. Trias et. al. [FX.Trias et al. JCP 258: 246-267, 2014], a new approach for studying the Poisson equation obtained using the Fractional Step Method is found, such as a new interpolator is proposed in order to found a stable solution, which avoid the appearence of these unpleasant velocities. The stability provided by the interpolator is formally proved for cartesian meshes and its rotations, using fully-explicit time discretizations. The construction of the Poisson equation is supported on mimicking the symmetry properties of the differential operators and the Fractional Step Method. Then it is reinterpreted using a recursive application of the Fractional Step Method in order to study the system as an stationary iterative solver. Furthermore, a numerical analysis for unstructured mesh is also provided.

1 INTRODUCTION

The continuity and Navier-Stokes equations are the equations which describe the mass and momentum conservation. Even though they were formulated long time ago, they still do not have an analytical solution for general cases. This is why numerical methods are so important, because they give us numerical approximations of these solutions. One family of numerical methods that can be applied to the resolution of this equations are the Finite Volume Methods. These methods are characterised by their apparent mathematical simplicity and intuitive description of the fundamental physics of the equations. Despite the fact that a staggered configuration would be the natural way in which the FVM should be applied to discretise the equations [1, 2, 3], a collocated grid arrangement is often preferred due to its simplicity in order to deal with complex geometries and its computational cost [4, 5]. Indeed, this approach is found in the most popular codes, such as ANSYS-FLUENT or OpenFOAM.

One method used to solve the velocity-pressure coupling is the classical fractional step method [6, 7, 8].

However, working with collocated meshes and a wide stencil Poisson equation in such a method lead us to the well-known checkerboard problem. This problem can be solved by means of working with a compact stencil Poisson equation [9, 10] or by regularizing of the convective term such as [4].

Furthermore, other problems or unstabilities can be found in the numerical solution of these equations. One of these problems is the appearance of spurious velocities in highly distorted meshes, coming from the interpolation from faces to cells of the discrete pressure gradient, which is strongly related to the resolution of the Poisson equation itself. Many reconstructions of the discrete gradient have been proposed [11, 12, 13, 14]. Some of them were proposed in terms of improving accuracy, but they can also fix the interpolation problem. However, they do not usually respect the underlying symmetries of the continuos gradient and divergence operator. This work is focused on this problem from the point of view of respecting these symmetries.

2 Symmetry-preserving finite volume discretization on collocated grids

The well-known dimensionless Navier-Stokes equations for the simulation of laminar or turbulent, Newtonian and incompressible flows in primitive variables are

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

$$\nabla \cdot u = 0, \tag{1b}$$

where Re is the dimensionless Reynolds number. A fully-conservative finite-volume discretization respecting the symmetries of the differential operators for structured and unstructured meshes was presented in [4]. Using this discretization approach, the properties of the continuous equations are preserved, as far as the symmetries of the convective and diffusive operators.

As a collocated formulation is preferred over a staggered one due to its simplicity, pressure and velocity are computed at the control volume center. Following the notation used in [4], let us define $p_c = (p_1, p_2, \ldots, p_n)^* \in \mathbb{R}^n$ and $u_c \in \mathbb{R}^{3n}$ as the cell-centered pressure and velocity fields, where n is the number of control volumes. The subindices c and s refer to whether the variables are cell-centered or staggered at the faces. For simplicity u_c is defined as a column vector and arranged as $u_c = (u_1, u_2, u_3)^*$, where $u_i = \left((u_i)_1, (u_i)_2, \ldots, (u_i)_n \right)^*$ are the vectors containing the velocity components corresponding to the x_i-spatial direction. In order to assure mass conservation in the control volumes, a secondary velocity field is defined at the faces $u_s = \left((u_s)_1, (u_s)_2, (u_s)_3, \ldots, (u_s)_m \right)^* \in \mathbb{R}^m$ where m is the number of faces on the computational domain. Defining a linear shift transformation from cell to faces $\Gamma_{c \to s} \in \mathbb{R}^{m \times 3n}$, both velocities can be related as follows

$$u_s \equiv \Gamma_{c \to s} u_c. \tag{2}$$

The matrices $\Omega \in \mathbb{R}^{3n \times 3n}$, $\mathsf{C}(u_s) \in \mathbb{R}^{3n \times 3n}$ and $\mathsf{D} \in \mathbb{R}^{3n \times 3n}$ are block diagonal matrices given by

$$\Omega = \mathsf{I}_3 \otimes \Omega_c, \quad \mathsf{C}(u_s) = \mathsf{I}_3 \otimes \mathsf{C}_c(u_s), \quad \mathsf{D} = \mathsf{I}_3 \otimes \mathsf{D}_c, \tag{3}$$

where $\mathsf{I}_3 \in \mathbb{R}^{3 \times 3}$ is the identity matrix and $\Omega_c \in \mathbb{R}^{n \times n}$ is a diagonal matrix with the cell-centered control volumes. $\mathsf{C}_c(u_s) \in \mathbb{R}^{n \times n}$ and $\mathsf{D}_c \in \mathbb{R}^{n \times n}$ are the cell-centered convective and diffusive operators for a

discrete scalar field, respectively. Finally, $\mathsf{G}_c \in \mathbb{R}^{3n \times n}$ represents the discrete gradient operator and the matrix $\mathsf{M} \in \mathbb{R}^{n \times m}$ is the face-to-center discrete divergence operator.

In this way, the equations on an arbitrary collocated mesh can be written as follows

$$\Omega \frac{du_c}{dt} + \mathsf{C}(u_s) u_c + \mathsf{D}u_c + \Omega \mathsf{G}_c p_c = 0_c, \tag{4a}$$

$$\mathsf{M}u_s = 0_c. \tag{4b}$$

The construction of the discrete operators is adressed in section 3 of [4] and it is not the object of this work.

2.1 Solving the pressure-velocity coupling. Checkerboard problem

In order to simplify the notation, semidiscretized momentum equation (4a) can be rewritten as follows

$$\frac{du_c}{dt} = \mathsf{R}(u_c) - \mathsf{G}_c p_c, \tag{5}$$

where $\mathsf{R}(u_c) \equiv -\Omega^{-1}(\mathsf{C}(u_s) u_c + \mathsf{D}u_c)$ represents the convective and diffusive terms. For the temporal discretization, a fully explicit scheme is assumed whereas the incompressibility constraint is treated implicitly. Introducing the functions f and $\mathbf{R}$, which represent a general fully explicit discretization, the fully-discretized NS equations read

$$\mathsf{M}u_s^{n+1} = 0_c, \tag{6a}$$

$$\frac{f(u_c^{n+1}, ..., u_c^0)}{\Delta t} = \mathbf{R}(u_c^n, ..., u_c^0) - \mathsf{G}_c p_c^{n+1},$$

In the case presented in [4], $f(u_c^{n+1}, ..., u_c^0) = (\kappa + 1/2) u_c^{n+1} - 2\kappa u_c^n + (\kappa - 1/2) u_c^{n-1}$, and $\mathbf{R}(u_c^n, ..., u_c^0) = \mathsf{R}((1 + \kappa) u_c^n - \kappa u_c^{n-1})$, where the parameter κ is computed each time-step to adapt the linear stability domain of the time-integration scheme to the instantaneous flow conditions in order to maximize the time-step, Δt [15]. The simplest case would be a first-order forward Euler: $f(u_c^{n+1}, ..., u_c^0) = u_c^{n+1} - u_c^n$.

Observe that, in general, we can split f into two parts. The first one will contain the new velocity u_c^{n+1} and the second one will contain all the previous ones. This split can be represented as $f(u_c^{n+1}, ..., u_c^0) = \alpha u_c^{n+1} + h(u_c^n, ..., u_c^0)$, where α is a general parameter and g is a function containing the contribution of all previous velocites.

To solve the velocity-pressure coupling, a classical fractional step method [6, 7, 8] is used. For the staggered velocity field, u_s, this projection is 'naturally' derived from the Helmholtz-Hodge vector decomposition theorem [16], whereby a velocity u_s^p can be uniquely decomposed into a solenoidal vector, u_s^{n+1}, and a curl-free vector, expressed as the gradient of a scalar field, $\mathsf{G}\tilde{p}_c'$. This decomposition is written as

$$u_s^p = u_s^{n+1} + \mathsf{G}\tilde{p}_c', \tag{7}$$

where $\mathsf{G} \in \mathbb{R}^{m \times n}$ is the center-to-face staggered gradient operator which is related with the divergence operator via [4]

$$\mathsf{G} \equiv -\Omega_s^{-1} \mathsf{M}^*, \tag{8}$$

where $\Omega_s \in \mathbb{R}^{m \times m}$ is a diagonal matrix with the staggered control volumes. Then, taking the divergence of Eq.(7) yields a discrete Poisson equation for $\tilde{p}_c{}'$

$$\mathsf{M} u_s^p = \mathsf{M} u_s^{n+1} + \mathsf{M} \mathsf{G} \tilde{p}_c{}' \quad \longrightarrow \quad \mathsf{M} \mathsf{G} \tilde{p}_c{}' = \mathsf{M} u_s^p. \tag{9}$$

Finally, using the definition of G given in Eq.(8) the previous equation becomes

$$\mathsf{L} \tilde{p}_c{}' = \mathsf{M} u_s^p \qquad \text{with} \qquad \mathsf{L} \equiv -\mathsf{M} \Omega_s^{-1} \mathsf{M}^*, \tag{10}$$

where the discrete Laplacian operator $\mathsf{L} \in \mathbb{R}^{n \times n}$ is, by construction, a symmetric negative-definite matrix. Once the solution is obtained, u_s^{n+1} results from the correction (7)

$$u_s^{n+1} = u_s^p - \mathsf{G} \tilde{p}_c{}'. \tag{11}$$

2.1.1 Constraints on the shift operators

Linear shift operators are needed to relate the cell-centered velocity fields to the staggered ones and vice-versa. Namely, the linear shift operator, $\Gamma_{c \to s} \in \mathbb{R}^{m \times 3n}$, transforms a cell-centered velocity field into a staggered one

$$u_s = \Gamma_{c \to s} u_c, \tag{12}$$

whereas the cell-centered fields are related to the staggered ones via the linear shift transformation $\Gamma_{s \to c} \in \mathbb{R}^{3n \times m}$,

$$u_c = \Gamma_{s \to c} u_s. \tag{13}$$

Notice, that in general $\Gamma_{s \to c} \Gamma_{c \to s} = \mathsf{I}$ holds only approximately *i.e.* $u_c \approx \Gamma_{s \to c} \Gamma_{c \to s} u_c$. More importantly, recalling the definition for the staggered gradient operator given in Eq.(8), the cell-centered discrete gradient operator results

$$\mathsf{G}_c = -\Gamma_{s \to c} \Omega_s^{-1} \mathsf{M}^*, \tag{14}$$

and, therefore, the face-to-cell shift operator $\Gamma_{s \to c}$ is restricted as follows

$$\left(\Omega \Gamma_{s \to c} \Omega_s^{-1} \mathsf{M}^* \right)^* = \mathsf{M} \Gamma_{c \to s} \longrightarrow \Gamma_{s \to c} = \Omega^{-1} \Gamma_{c \to s}^* \Omega_s, \tag{15}$$

to force that the pressure gradient contribution to the global kinetic energy exactly vanishes (see section 2.2 of [4]).

3 Analysis on the Poisson equation obtained using the FSM

3.1 Computing the cell-center predictor velocity u_c^p

Let us introduce the following notation: $h^i = h(u_c^i, ..., u_c^0)$ and $\mathbf{R}^i = \frac{1}{\alpha} \mathbf{R}(u_c^i, ..., u_c^0)$. In the projection method, assuming we are at the $n+1$ time step, the predicted velocity, u_c^p, is computed as follows

$$u_c^{p1} = -\frac{1}{\alpha} h^n + \Delta t^{n+1} \mathbf{R}^n, \tag{16}$$

$$u_c^p = u_c^{p1} - \mathsf{G}_c \tilde{p}_c{}^p, \tag{17}$$

where $\tilde{p}_c{}^p$ is a 'prediction' of the pressure at the next time step, $\tilde{p}_c{}^{n+1} = \tilde{p}_c{}^p + \tilde{p}_c{}'$. The definition of $\tilde{p}_c{}'$ introduce an additional error to the momentum equation in a collocated formulation. Furthermore, as we are working with collocated meshes, we will face the checkerboard problem. This problem is widely adressed in sections 2.3.3 and 2.4 of [4].

So equation (17) can be rewritten as:

$$u_c^p = -\frac{1}{\alpha}h^n + \Delta t^{n+1}\mathbf{R}^n - \mathsf{G}_c\tilde{p}_c{}^p. \tag{18}$$

Introducing the last result in equation (10), we obtain:

$$\mathsf{L}\tilde{p}_c{}' = \mathsf{M}u_s^p = \mathsf{M}_c u_c^p = -\frac{1}{\alpha}\mathsf{M}_c h^n + \Delta t^{n+1}\mathsf{M}_c\mathbf{R}^n - \mathsf{L}_c\tilde{p}_c{}^p, \tag{19}$$

where $\mathsf{L}_c = \mathsf{M}_c\mathsf{G}_c \in \mathbb{R}^{n \times n}$ is the collocated Laplacian operator. Finally, adding $\mathsf{L}\tilde{p}_c{}^p$ at both sides, we arrive at:

$$\mathsf{L}\tilde{p}_c{}^{n+1} = -\frac{1}{\alpha}\mathsf{M}_c h^n + \Delta t^{n+1}\mathsf{M}_c\mathbf{R}^n - (\mathsf{L} - \mathsf{L}_c)\tilde{p}_c{}^p. \tag{20}$$

3.2 Reinterpretation of the recursive application of the FSM as a stationary iterative solver of the wide-stencil Poisson equation

3.2.1 Particular case: a first-order forward Euler

For the first-order forward Euler scheme, we can specify the previous functions and constants as $\alpha = 1$, $h^n = -u_c^n$ and $\mathbf{R}^n = \mathsf{R}(u_c^n)$. Assuming we are in $n+1$ time step, the velocity centered field at time n can be computed as follows:

$$\begin{aligned} u_c^n &= u_c^p - \mathsf{G}_c\tilde{p}_c{}' = u_c^{n-1} + \Delta t^n\mathsf{R}\left(u_c^{n-1}\right) - \mathsf{G}_c(\tilde{p}_c{}^p + \tilde{p}_c{}') = \\ &= \left(u_c^{n-2} + \Delta t^{n-1}\mathsf{R}\left(u_c^{n-2}\right) - \mathsf{G}_c\tilde{p}_c{}^{n-1}\right) + \Delta t^n\mathsf{R}\left(u_c^{n-1}\right) - \mathsf{G}_c\tilde{p}_c{}^n = ... = \\ &= u_c^0 + \sum_{i=1}^{n}\Delta t^i\mathsf{R}\left(u_c^{i-1}\right) - \mathsf{G}_c\sum_{i=0}^{n}\tilde{p}_c{}^i \end{aligned} \tag{21}$$

Introducing (21) in (20) leads to:

$$\begin{aligned} \mathsf{L}\tilde{p}_c{}^{n+1} &= \mathsf{M}_c u_c^n + \Delta t^{n+1}\mathsf{M}_c\mathbf{R}^n - (\mathsf{L} - \mathsf{L}_c)\tilde{p}_c{}^p = \\ &= \mathsf{M}_c u_c^0 + \mathsf{M}_c\sum_{i=1}^{n+1}\Delta t^i\mathsf{R}\left(u_c^{i-1}\right) - \mathsf{L}_c\sum_{i=0}^{n}\tilde{p}_c{}^i + (\mathsf{L} - \mathsf{L}_c)\tilde{p}_c{}^p. \end{aligned} \tag{22}$$

Now, we can define $\bar{p}_c{}^n = \sum_{i=0}^{n}\tilde{p}_c{}^i$. Observe that $\tilde{p}_c{}^{n+1} = \bar{p}_c{}^{n+1} - \bar{p}_c{}^n$. This allows us to rearrange the previous equation as:

$$\mathsf{L}\bar{p}_c{}^{n+1} = \mathsf{M}_c u_c^0 + \mathsf{M}_c \sum_{i=1}^{n+1} \Delta t^i \mathsf{R}\left(u_c^{i-1}\right) + (\mathsf{L} - \mathsf{L}_c)\bar{p}_c{}^n + (\mathsf{L} - \mathsf{L}_c)\tilde{p}_c{}^p. \tag{23}$$

This can be viewed as a stationary iterative solver. The stability of this process will depend on the eigenvalues of $\mathsf{L}^{-1}\mathsf{L}_c$ and $(\mathsf{L} - \mathsf{L}_c)$. This is going to be analysed deeply for the general case in the next section.

3.2.2 The general case: assuming a general fully-explicit scheme

Recalling we are working with collocated meshes and assuming we are in $n+1$ time step, the velocity centered field at time n can be computed as follows:

$$
\begin{aligned}
u_c^n &= u_c^p - \mathsf{G}_c \tilde{p}_c{}' = -\frac{1}{\alpha} h^{n-1} + \Delta t^n \mathbf{R}^{n-1} - \mathsf{G}_c(\tilde{p}_c{}^p + \tilde{p}_c{}') = \\
&= -\frac{1}{\alpha} h^{n-1} + \Delta t^n \mathbf{R}^{n-1} - \mathsf{G}_c \tilde{p}_c{}^n.
\end{aligned}
\tag{24}
$$

Recalling that h is a function which contains a linear combination of the previous velocities, we can apply this result recursively up to u_c^0 and rearrange the results in new functions:

$$u_c^n = ... = \mathbf{K}u_c^0 + \mathsf{G}_c F(\tilde{p}_c{}^n, \tilde{p}_c{}^{n-1}, ..., \tilde{p}_c{}^0) + H(\Delta t^n \mathbf{R}^{n-1}, \Delta t^{n-1}\mathbf{R}^{n-2}, ..., \Delta t^1 \mathbf{R}^0) \tag{25}$$

where $\mathbf{K}$ is a constant, and F and H are functions that depend on $\tilde{p}_c{}^i$ and $\mathbf{R}^i$ respectively, and also in the election of the time step Δt at each iteration.

Observe that, in the same way we procedeed for (25), we can rewrite equation (20) in terms of all the previous time steps using the last result as follows:

$$\mathsf{L}\tilde{p}_c{}^{n+1} = \tilde{\mathbf{K}} + \mathsf{L}_c \tilde{F}(\tilde{p}_c{}^n, \tilde{p}_c{}^{n-1}, ..., \tilde{p}_c{}^0) + \tilde{H}(\Delta t^{n+1}\mathbf{R}^n, \Delta t^n \mathbf{R}^{n-1}, ..., \Delta t^1 \mathbf{R}^0) - (\mathsf{L} - \mathsf{L}_c)\tilde{p}_c{}^p. \tag{26}$$

Here, we have done an arrangement in terms of functions as we did in (25). This can be viewed as a stationary iterative solver. Note that inside $\mathbf{R}^i$ we also have the velocities at previous time steps and, consequently, the pressures at previous time steps, but, if the time step Δt is chosen wisely, the stability of this process will depend on the eigenvalues of $\mathsf{L}^{-1}\mathsf{L}_c$ and $(\mathsf{L} - \mathsf{L}_c)$, which subsequently depend on the interpolator operators. This will be carefully analysed in the next sections.

3.2.3 Conditions to solve the system

Inverting L in (26), and taking into account that $\tilde{\mathbf{K}}$ and the function H can be bounded by means of a wisely chosen time step, then, let us define $\mathsf{T}_1 = \mathsf{L}^{-1}(\mathsf{L} - \mathsf{L}_c)$ and $\mathsf{T}_2 = \mathsf{L}^{-1}\mathsf{L}_c$. If $\mu_i^{\mathsf{T}_1}$ and $\mu_i^{\mathsf{T}_2}$ are their eigenvalues, the process will converge if $|\mu_i^{\mathsf{T}_1}| < 1$ and $|\mu_i^{\mathsf{T}_2}| < 1$.
Notice that

$$\mathsf{L}^{-1}(\mathsf{L} - \mathsf{L}_c) = \mathsf{I} - \mathsf{L}^{-1}\mathsf{L}_c \implies \mu_i^{\mathsf{T}_1} = 1 - \mu_i^{\mathsf{T}_2}. \tag{27}$$

One the one hand, $L^{-1}L_c$ is the product of two negative definite matrices (see Corollary 1), then, by Corollary 2, $L^{-1}L_c$ is a positive definite matrix. So $\mu_i^{\mathsf{T}_1} = 1 - \mu_i^{\mathsf{T}_2} < 1$ (it is a well-known result in solving iterative systems).

On the other hand, if $L - L_c$ is negative definite, then $L^{-1}(L - L_c)$ is positive definite, so $\mu_i^{\mathsf{T}_1} > 0$. We can conclude that $\mu_i^{\mathsf{T}_1} \in (0, 1)$ and it implies that $\mu_i^{\mathsf{T}_2} \in (0, 1)$ and the process converge.

In conclusion, we have to study the definiteness of $L - L_c$.

3.2.4 Relation of $L - L_c$ and the interpolator operators

We have to proof when $L - L_c$ is negative definite:

$$
\begin{aligned}
L - L_c &= -M\Omega_s^{-1}M^* + M\Gamma_{c \to s}\Omega^{-1}\Gamma_{c \to s}^* M^* \\
&= -M(\Omega_s^{-1} - \Gamma_{c \to s}\Omega^{-1}\Gamma_{c \to s}^*)M^*.
\end{aligned}
\tag{28}
$$

According to Lemma 1, the definiteness of $-M(\Omega_s^{-1} - \Gamma_{c \to s}\Omega^{-1}\Gamma_{c \to s}^*)M^*$ is given by the definiteness of $\Omega_s^{-1} - \Gamma_{c \to s}\Omega^{-1}\Gamma_{c \to s}^*$, so, as we want $L - L_c$ to be negative definite, we should impose $\Omega_s^{-1} - \Gamma_{c \to s}\Omega^{-1}\Gamma_{c \to s}^*$ to be positive definite. Using the shift operators relation (15):

$$
\Omega_s^{-1} - \Gamma_{c \to s}\Omega^{-1}\Gamma_{c \to s}^* = \Omega_s^{-1} - \Omega_s^{-1}\Gamma_{s \to c}^*\Omega\Gamma_{s \to c}\Omega_s^{-1}
\tag{29}
$$

Multiplying each side of the expression by Ω_s and applying again Lemma 1, we conclude that

$$
\Omega_s(\Omega_s^{-1} - \Omega_s^{-1}\Gamma_{s \to c}^*\Omega\Gamma_{s \to c}\Omega_s^{-1})\Omega_s = \Omega_s - \Gamma_{s \to c}^*\Omega\Gamma_{s \to c}
\tag{30}
$$

should be positive definite. As we can see here, the definiteness of $L - L_c$ are directly related to the interpolation from faces to cells matrix. Ideally, as it is easier to define the interpolation from cells to faces via $\Gamma_{c \to s}$, we would like some expression involving this matrix. In order to do that, using (15) and multiplying each side of the expression by $\Omega_s^{-\frac{1}{2}}$:

$$
I - \Omega_s^{-\frac{1}{2}}\Gamma_{c \to s}\Omega^{-1}\Gamma_{c \to s}^T\Omega_s^{-\frac{1}{2}}.
\tag{31}
$$

If $B = \Omega_s^{-\frac{1}{2}}\Gamma_{c \to s}\Omega^{-1}\Gamma_{c \to s}^T\Omega_s^{-\frac{1}{2}}$, we will need $\rho(B) \leq 1$, being $\rho(\cdot)$ the spectral radius (observe that B is positive definite).

Defining $D = \Omega^{-\frac{1}{2}}\Gamma_{c \to s}^T\Omega_s^{-\frac{1}{2}}$, we can construct the previous matrix $B = D^T D$ and another matrix $C = DD^T$. A general theorem states that $eigen(D^T D) = eigen(DD^T)$ except for $\lambda = 0$ (and its multiplicity) eigenvalue. So,

$$
\rho(C) = \rho(B) \Rightarrow (\rho(B) \leq 1 \Leftrightarrow \rho(C) \leq 1).
\tag{32}
$$

So, we need $I - \Omega^{-\frac{1}{2}}\Gamma_{c \to s}^T\Omega_s\Gamma_{c \to s}\Omega^{-\frac{1}{2}}$ to be positive definite. Finally, we can multiply the previous expression by $\Omega^{\frac{1}{2}}$ and we will obtain that

$$
\Omega - \Gamma_{c \to s}^T\Omega_s\Gamma_{c \to s}.
\tag{33}
$$

Finally, we have proven that $L - L_c$ being negative definite is equivalent to $\Omega - \Gamma_{c \to s}^T\Omega_s\Gamma_{c \to s}$ being positive definite. An interpretation of this in terms of a Cartesian mesh can be found in the next section.

3.2.5 Relation between the interpolator matrix and the defineteness

Taking into account that $\Gamma_{c\rightarrow s} = N_s\Pi = \begin{pmatrix} N_{s,1} & N_{s,2} & N_{s,3} \end{pmatrix}(I_3 \otimes \Pi_{c\rightarrow s})$, where the matrices $N_{s,i} \in \mathbb{R}^{m \times m}$ are diagonal matries containing the x^i-spatial components of the face normal vectors, and $\Pi_{c\rightarrow s} \in \mathbb{R}^{m \times n}$ is the operator that interpolates a cell-centered scalar field to the faces. The calculation of the matrix and the coefficients are straightforward:

$$\Gamma_{c\rightarrow s}^T \Omega_s \Gamma_{c\rightarrow s} = \begin{pmatrix} \Pi_{c\rightarrow s}^T N_{s,1}\Omega_s N_{s,1}\Pi_{c\rightarrow s} & \Pi_{c\rightarrow s}^T N_{s,1}\Omega_s N_{s,2}\Pi_{c\rightarrow s} & \Pi_{c\rightarrow s}^T N_{s,1}\Omega_s N_{s,3}\Pi_{c\rightarrow s} \\ \Pi_{c\rightarrow s}^T N_{s,2}\Omega_s N_{s,1}\Pi_{c\rightarrow s} & \Pi_{c\rightarrow s}^T N_{s,2}\Omega_s N_{s,2}\Pi_{c\rightarrow s} & \Pi_{c\rightarrow s}^T N_{s,2}\Omega_s N_{s,3}\Pi_{c\rightarrow s} \\ \Pi_{c\rightarrow s}^T N_{s,3}\Omega_s N_{s,1}\Pi_{c\rightarrow s} & \Pi_{c\rightarrow s}^T N_{s,3}\Omega_s N_{s,2}\Pi_{c\rightarrow s} & \Pi_{c\rightarrow s}^T N_{s,3}\Omega_s N_{s,3}\Pi_{c\rightarrow s} \end{pmatrix}, \qquad (34)$$

and

$$(\Pi_{c\rightarrow s}^T N_{s,i}\Omega_s N_{s,j}\Pi_{c\rightarrow s})_{r,s} = \sum_{l=1}^{m} a_{lr}n_l^i V_l^{stagg} n_l^j a_{ls}, \qquad (35)$$

where a_{ij} are the coefficients of $\Pi_{c\rightarrow s}$. For more details in the notation, the reader is referred again to [4].

Now, in order to relate all this to the grid, let us recall the Greshgorin circle theorem. It tells us that, if $A \in \mathbb{R}^{n \times n}$, $R_i = \sum_{j \neq i}|a_{ij}|$, and $D(a_{ii},R)$ is the closed disc of radius R and center a_{ii}, then every eigenvalue of A lies within at least one $D(a_{ii},R)$. Given a general scalar interpolation matrix from cells to faces,

$$\Pi_{c\rightarrow s} = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ . & . & . & . \\ . & . & . & . \\ a_{m1} & a_{m2} & ... & a_{mn} \end{pmatrix}, \qquad (36)$$

and applying the Greshgorin circle theorem to our previous matrix $\Omega - \Gamma_{c\rightarrow s}^T \Omega_s \Gamma_{c\rightarrow s}$, we will find that, if $V_r - \sum_{s=1}^{n}\sum_{j=x,y,z}|\sum_{l=1}^{m} a_{lr}V_l^{stagg}a_{ls}n_l^i n_l^j| \geq 0 \quad \forall r \in \{1,...,n\}, i \in \{x,y,z\}$, then $\Omega - \Gamma_{c\rightarrow s}^T \Omega_s \Gamma_{c\rightarrow s}$ is positive definite and consequently, $\mathsf{L} - \mathsf{L}_c$ is negative definite.

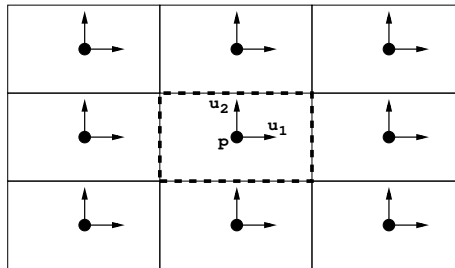### 3.3 A stable interpolation for Cartesian grids

Figure 1: Representation of a collocated-Cartesian grid.

We will understand a Cartesian mesh as a mesh whose faces are parallel to Cartesian axis. For our purpose, we can have elements with different volumes. Let us assume we apply the previous result to this kind of mesh, in particular to the highlighted control volume. If we take $i = x$:

$$
\begin{aligned}
V_P - \sum_{s=0}^{4} \sum_{j=x,y} |\sum_{l=1}^{4} a_{lP} V_l^{stagg} a_{ls} n_l^x n_l^j| &= V_P - \sum_{s=0}^{4} \sum_{j=x,y} |a_{eP} V_e^{stagg} a_{es} n_e^x n_e^j + a_{wP} V_w^{stagg} a_{ws} n_w^x n_w^j| = \\
&= V_P - \sum_{s=0}^{4} \left(a_{eP} V_e^{stagg} a_{es} + a_{wP} V_w^{stagg} a_{ws} \right) = \\
&= V_P - a_{eP} V_e^{stagg} (a_{eP} + a_{eE}) - a_{wP} V_w^{stagg} (a_{wP} + a_{wW}) \quad (37)
\end{aligned}
$$

Here, we are changing the numbers for the cardinal points N, S, W, E for simplicity, s is the subindex referred to control volumes and l is the subindex referred to faces. For example, a_{eP} is the interpolator coefficient from cell P to face e. The faces n, s, e, w are referred to the control volume P.

Now, let us define the following interpolator, being i the index used for faces and j the index used for control volumes:

$$
a_{ij} = \frac{\tilde{V}_{ij}}{V_i^{stagg}}, \quad (38)
$$

where $\tilde{V}_{ij}$ is the quantity of volume of V_i^{stagg} which is inside the control volume j. This quantity can be computed as $\tilde{V}_{ij} = d_{ij} * S_i$, where d_{ij} is the distance from control volume center j to the face i, and S_i is the surface of i. The interpolator is called volume weighted. Observe that, if j and k are two adjacent control volumes which share face i, then $a_{ij} + a_{ik} = 1$. Applying this interpolator to equation (37) we obtain:

$$
\begin{aligned}
V_P - a_{eP} V_e^{stagg} (a_{eP} + a_{eE}) - a_{wP} V_w^{stagg} (a_{wP} + a_{wW}) &= V_P - a_{eP} V_e^{stagg} - a_{wP} V_w^{stagg} = \\
&= V_P - \tilde{V}_{eP} - \tilde{V}_{wP} = 0. \quad (39)
\end{aligned}
$$

Last equality holds because $\tilde{V}_{eP} - \tilde{V}_{wP} = d_{eP} * S_e + d_{wP} * S_w = V_P$, independently of the cell center location (at least if we define d_{ij} as the distance from the cell node to the plane which contains the face). However, this suggests putting the cell center right in the center of the rectangle, which is the common way. Furthermore, if we have a non-periodic case, and we are in the adjacent cell of the (for example) left boundary, observe that we would get $V_P - \tilde{V}_{eP} \geq 0$. So, the interpolator gives a stable solution for the Poisson equation in Cartesian grids even in non-periodic cases.

Observe that the results for the y-coordinate are analogous to the previous ones. Furthermore, due to the same symmetries, this result can be extended to a 3-dimensional case (or to a general n-dimensional case). The extension is straightforward. In addition, the stability can be also proved for rotations of the Cartesian meshes, but the proof is not presented here due to its extension.

3.4 Comments on unstructured grids

As no theoretical result could be found for triangular unstructured meshes, a numerical test has been carried out:

- Checking the sign of the eigenvalues of $L - L_c$.

- Computing the global error and the rate of convergence of the new interpolator.

When placing the cell center in the orthocenter, some eigenvalues of $L - L_c$ lie on the positive axis, even using the weighted volume scheme. This problem seems to be avoided if the cell center is placed in the circumcenter and it is placed inside the triangle. However, when the circumcenter lies outside the triangle the previous problem appears again. The possibility of working with negative staggered volumes has been tested, but it introduces other problems such as L not being negative definite. A possible solution would be to work with meshes constructed with acute-angled triangles, because they have the circumcenter inside the triangle.

For computing the global error and the rate of convergence,

- 60 random meshes have been used.

- Average step-size is defined as $trace(\Omega_c)/n$, where n is the number of control volumes.

- Velocity was assumed to be $u = cos(2\pi x)sin(2\pi y)$ and $v = -sin(2\pi x)cos(2\pi y)$.

- The next mapping function has been used to stretch the meshes:

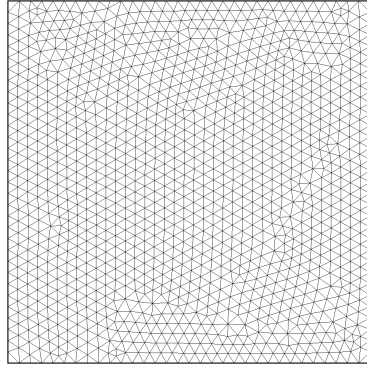$$x_i^{new} = \frac{1}{2}\left(1 + \frac{tanh(\gamma(2x_i - 1))}{tanh(\gamma)}\right)$$
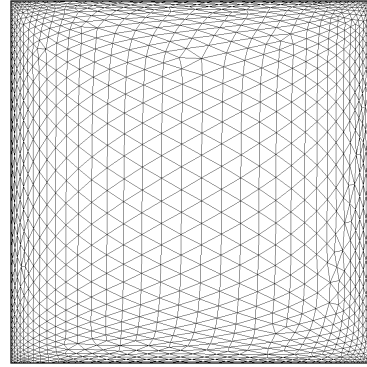

Figure 2: Example of mesh used with $\gamma = 0$.

Figure 3: Example of mesh used with $\gamma = 2$.

The global error is understood as $||p_c^s - p_c^c||_2$, where:

- p_c^s is obtained solving $Lp_c^s = Mu_s$.

- p_c^c is obtained solving $Lp_c^c = M\Gamma_{c\to s}u_c$.

So, we understand the global error as the mean difference between the result of an iteration using the staggered velocity and the collocated velocity.

The rate of convergence, p_{conv}, is found fitting the points assuming a relation $y = kx^{p_{conv}}$, where k is a constant. As we can see in the next table, the rate of convergence value is around 1 for both schemes. Furthermore, the order of the error found with both schemes is similar, and it becomes practically the same for stretched meshes.
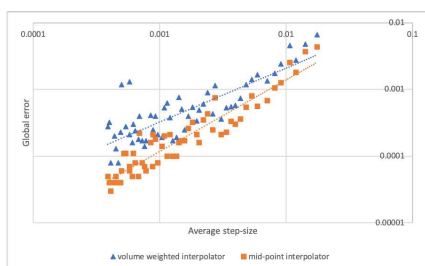
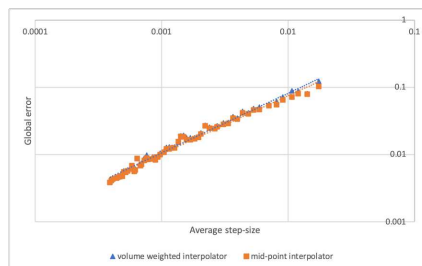Figure 4: Mean global error results for the mesh with $\gamma = 0$.

Figure 5: Mean global error results for the mesh with $\gamma = 2$.

γ	0	0.5	1	1.5	2
p (volume weighted interpolator)	0.81	0.96	1.00	1.05	0.88
p (mid-point interpolator)	1.07	0.97	0.98	1.02	0.87

4 Conclusions

Apart from the well-known checkerboard problem, the appearance of unphysical velocities is a common problem found in highly distorted meshes. The volume weighted scheme solves this problem, at least for Cartesian meshes. A theoretical result could not be found for unstructured triangular meshes, but the accuracy for the resolution of one iteration of the Poisson equation has been tested. The results show that we have a first-order error solving the Poisson equation when approaching the staggered velocity using the volume weighted scheme. Furthermore, problems with extending the methodology are found when placing the cell-center outside the control volume, such as breaking the negative definiteness of L. Finally, it is interesting to remark that an interpolator that makes (33) positive definite (thus, guaranteeing stability) could not be unique, so other possibilites (also for unstructured meshes) may be found.

Acknowledgments

A Some mathematical results about matrices and eigenvalues

All the results presented in this appendix have been proved by the authors. However, the authors do not claim the authority of them, because they probably have been previously demonstrated.

Lemma 1 *Let $A \in \mathbb{R}^{m \times m}$ be a positive (or negative) definite matrix. If $M \in \mathbb{R}^{n \times m}$, $m \geq n$, is another matrix with full rank, then, MAM^T is a positive (or negative) definite matrix.*

Corollary 1 *The matrices L and L_c are negative defined.*

Lemma 2 *If A is a symmetric positive definite matrix and B is a positive (or negative) definite matrix, then AB is a positive (or negative) definite matrix.*

11

Corollary 2 *If A is a symmetric negative definite matrix and B is a positive (or negative) definite matrix, then AB is a negative (or positive) definite matrix.*

REFERENCES

[1] F. Harlow and J. Welch. Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface. *The Physics of Fluids*, 8(12):2182–2189, 1965.

[2] R. Verstappen and A. Veldman. Symmetry-preserving discretization of turbulent flow. *Journal of Computational Physics*, 187:343–368, 2003.

[3] F. Ham, F. Lien, A. Strong. A fully conservative second-order finitie difference scheme for incompressible flow on nonuniform grids. *Journal of Computational Physics*, 177(1):117–133, 2002.

[4] F. X. Trias, O. Lehmkuhl, A. Oliva, C.D. Pérez-Segarra, and R.W.C.P. Verstappen. Symmetry-preserving discretization of Navier-Stokes equations on collocated unstructured meshes. *Journal of Computational Physics*, 258:246–267, 2014.

[5] A. Pascau. Cell face velocity alternatives in a structured colocated grid for the unsteady Navier-Stokes equations. *International Journal for Numerical Methods in Fluids*, 65:812–833, 2011.

[6] A. J. Chorin. Numerical Solution of the Navier-Stokes Equations. *Mathematics of Computation*, 22:745–762, 1968.

[7] N. N. Yanenko. *The Method of Fractional Steps*. Springer-Verlag, 1971.

[8] J. B. Perot. An analysis of the fractional step method. *Journal of Computational Physics*, 108:51–58, 1993.

[9] C. M. Rhie and W. L. Chow. Numerical study of the turbulent flow past an airfoil with trailing edge separation. *AIAA Journal*, 21(11):1525–1532, 1983.

[10] C. M. Klaij. On the stabilization of finite volume methods with co-located variables for incompressible flow. *Journal of Computational Physics*, 297:84–89, 2015.

[11] B. Diskin and J. Thomas. Accuracy of gradient reconstruction on grids with high aspect ratio. *NIA Report*, (2008-12), 2009.

[12] A. Syrakos, S. Varchanis, Y. Dimakopoulos, A. Goulas and J. Tsamopoulous. A critical analysis of some popular methods for the discretisation of the gradient operator in finite volume methods. *Physics of Fluids*, 29, 2017.

[13] M. Deka, S. Brahmachary, R. Thirumalaisamy, A. Dalal and G. Natarajan. A new Green–Gauss reconstruction on unstructured meshes. Part I: Gradient reconstruction. *Journal of Computational Physics*, 422, 2020.

[14] A. Katz and V. Sankaran. High aspect ratio grid effects on the accuracy of navier–stokes solutions on unstructured meshes. *Computers Fluids*, (65):66–79, 2012.

[15] F. X. Trias and O. Lehmkuhl. A self-adaptive strategy for the time-integration of Navier-Stokes equations. *Numerical Heat Transfer, part B*, 60(2):116–134, 2011.

[16] A. J. Chorin and J. E. Marsden. *A Mathematical Introduction to Fluid Mechanics*. Springer, third edition, 2000.