Spectral discretization of a model for organic pollution in waters

Christine Bernardi, Driss Yakoubi

Spectral discretization
of a model for organic pollution in waters

by Christine Bernardi* and Driss Yakoubi*

Abstract: We are interested with a mixed reaction diffusion system describing the organic pollution in stream-waters. In this work, we propose a mixed variational formulation and recall its well-posedness. Next, we consider a spectral discretization of this problem and establish nearly optimal error estimates. Numerical experiments confirm the interest of this approach.

Résumé: Nous nous intéressons à un système d'équations aux dérivées partielles qui modélise la pollution organique dans des rivières. Nous en écrivons une formulation variationnelle mixte et rappelons qu'il est bien posé. Puis nous considérons sa discrétisation par méhode spectrale et démontrons des estimations presque optimales de l'erreur. Des expériences numériques confirment l'intérêt de cette méthode.

---

*   CNRS, UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France.
    Sorbonne Universits, UPMC Univ Paris 06, UMR 7598, LJLI, F-75005, Paris, France.
    e-mail addresses: bernardi@ann.jussieu.fr, yakoubi@giref.ulaval.ca

1. Introduction.

Let Ω be a bounded connected two- or three-dimensional domain with a Lipschitz-continuous boundary $\partial\Omega$. We are interested in the following system, where the unknowns are the density b of the Biochemical Oxygen Demand (BOD) and the concentration c of the Dissolved Oxygen (DO):

$$
\begin{cases}
-\mathrm{div}\,(d\,\nabla b) + r\,b = f & \text{in } \Omega, \\[2mm]
-\mathrm{div}\,(d\,\nabla c) + r_*\,c + r\,b = g & \text{in } \Omega, \\[2mm]
c = \alpha & \text{on } \partial\Omega, \\[2mm]
d\,\partial_n c = \beta & \text{on } \partial\Omega.
\end{cases}
\tag{1.1}
$$

The dispersion coefficient d and the reaction parameters r and r_* are positive but depend on the space variable. The right-hand side f represents the source of the pollution, while the datum g can for instance describe the uptake oxygen from the atmosphere to reduce the deficit of oxygen caused by the biodegradation of the organic pollutants.

Such a model was derived from an early formulation written by Streeter and Phelps in 1925 to study the Ohio river (see [13]), from which sophisticated modeling of the organic pollution in stream-waters has been elaborated. We refer to [11], [5] and [12] for the dispersion-reaction models. The central element of such problems is the oxygen. The main tracers currently used are the BOD which is the amount of oxygen per unit volume, necessary for the micro-organisms and aerobic bacteria to break down the organic matter contained in the water and the DO which is the oxygen concentration housed in a unit volume of water. The coupling term represented by $r\,b$ in the second equation is the depletion of oxygen due to elevated BOD.

One of the mathematical difficulties in system (1.1) is the existence of two boundary conditions on c and none on b, which is similar to the stream-function/vorticity formulation of the Stokes problem. Indeed, from a practical point of view, measurements on c are easy and instantaneously obtained while those on b require a strict chemical protocol and last five days. So, one of the stakes of working with problem (1.1) is to determine, for given values of the data f, g and also α, β, the values of $d\,\partial_n b$ on the boundary.

Despite all of this, the theoretical results for the model given in (1.1) have already been derived in [1], and we only recall them. Next, we propose a spectral discretization of this model. It must be noted that, in contrast with the finite element discretization proposed in [1], no stabilization is needed for this method. Thus, we construct a discrete problem by the Galerkin method with numerical integration, we prove its well-posedness and establish nearly optimal error estimates. Some numerical experiments confirm the good properies of this discretization.

The outline of the paper is as follows:

- In Section 2, we write the variational formulation of problem (1.1) and recall the arguments for its well-posedness.
- In Section 3, we describe the discrete problem and prove its well-posedness.
- Quasi-optimal a priori error estimates are derived in Section 4.
- Numerical experiments are presented in Section 5.

2. The mixed variational framework.

From now on, we assume that the reaction and diffusion coefficients r, r_* and d are piecewise continuous on $\overline{\Omega}$ and also that there exist positive real constants $r_\flat$, $r_\sharp$, $d_\flat$ and $d_\sharp$ such that

$$\forall \boldsymbol{x} \in \overline{\Omega}, \qquad r_\flat \leq r(\boldsymbol{x}), r_*(\boldsymbol{x}) \leq r_\sharp \quad \text{and} \quad d_\flat \leq d(\boldsymbol{x}) \leq d_\sharp. \tag{2.1}$$

We also suppose the diffusion coefficient d to be piecewise continuously differentiable on $\overline{\Omega}$. To develop our analysis we make a simplification assumption about the boundary data, that is $(\alpha, \beta) = (0, 0)$, and discuss the case of non-homogeneous boundary data only at the end of this section.

2.1. The variational formulation.

In what follows, we use the whole scale of Sobolev spaces $H^s(\Omega)$, $s \geq 0$. We also need the space $H_0^1(\Omega)$ of functions in $H^1(\Omega)$ which vanish on $\partial\Omega$ and its dual space $H^{-1}(\Omega)$, the duality pairing being represented by $\langle \cdot, \cdot \rangle_{H^{-1}, H_0^1}$. The functional framework involves the following space, introduced in [3] for the stream-function/vorticity formulation of the Stokes problem,

$$\mathbb{V} = \left\{ \chi \in L^2(\Omega); \ \text{div}\,(d\,\nabla\chi) \in H^{-1}(\Omega) \right\},$$

which is a Hilbert space when endowed with the graph norm

$$\|\chi\|_{\mathbb{V}} = \left(\|\text{div}\,(d\,\nabla\chi)\|_{H^{-1}(\Omega)}^2 + \|\chi\|_{L^2(\Omega)}^2 \right)^{1/2}.$$

To write the variational formulation of the system, let us introduce three bilinear forms,

$$\forall (\chi, \varphi) \in \mathbb{V} \times \mathbb{V}, \qquad a(\chi, \varphi) = \int_\Omega r(\boldsymbol{x})\,\chi(\boldsymbol{x})\varphi(\boldsymbol{x})\,d\boldsymbol{x},$$

$$\forall (\psi, \varphi) \in H_0^1(\Omega) \times \mathbb{V}, \qquad m(\psi, \varphi) = \langle -\text{div}\,(d\,\nabla\varphi) + r\,\varphi, \psi \,\rangle_{H^{-1}, H_0^1},$$

$$\forall (\psi, \varphi) \in H_0^1(\Omega) \times \mathbb{V}, \qquad m_*(\psi, \varphi) = \langle -\text{div}\,(d\,\nabla\varphi) + r_*\,\varphi, \psi \,\rangle_{H^{-1}, H_0^1}.$$

The mixed variational problem may then be expressed in terms of these bilinear forms as follows: For any data f in $H^{-1}(\Omega)$ and g in $L^2(\Omega)$, it reads

Find (b, c) in $\mathbb{V} \times H_0^1(\Omega)$ such that

$$\forall \psi \in H_0^1(\Omega), \qquad m\,(\psi, b) = \langle f, \psi \rangle_{H^{-1}, H_0^1},$$

$$\forall \varphi \in \mathbb{V}, \qquad m_*(c, \varphi) + a(b, \varphi) = \int_\Omega g(\boldsymbol{x})\varphi(\boldsymbol{x})\,d\boldsymbol{x}. \tag{2.2}$$

3

The following equivalence property is readily checked: A pair (b, c) in $\mathbb{V} \times H_0^1(\Omega)$ is a solution of the mixed variational problem (2.2) if and only if it satisfies the boundary value system (1.1) with $\alpha = \beta = 0$.

2.2. Well-posedness.

We recall that, according to the analysis of the abstract problem achieved in [10] (see also [2]), necessary and sufficient conditions on the bilinear forms are required to ensure existence, uniqueness, and stability for the mixed problem (2.2). The inf-sup conditions on the forms $m(\cdot, \cdot)$ and $m_*(\cdot, \cdot)$ easily follow from the imbedding of $H_0^1(\Omega)$ into $\mathbb{V}$, see [1, Lemmas 2.2 and 2.3].

Lemma 2.1. *The bilinear form $m(\cdot, \cdot)$ satisfies the inf-sup condition for a positive constant η,*

$$\forall \psi \in H_0^1(\Omega), \quad \sup_{\varphi \in \mathbb{V}} \frac{m(\psi, \varphi)}{\|\varphi\|_{\mathbb{V}}} \geq \eta \, \|\psi\|_{H^1(\Omega)}. \tag{2.3}$$

The bilinear form $m_*(\cdot, \cdot)$ satisfies the inf-sup condition for a positive constant η_*,

$$\forall \psi \in H_0^1(\Omega), \quad \sup_{\varphi \in \mathbb{V}} \frac{m_*(\psi, \varphi)}{\|\varphi\|_{\mathbb{V}}} \geq \eta_* \, \|\psi\|_{H^1(\Omega)}. \tag{2.4}$$

The inf-sup conditions on $a(\cdot, \cdot)$ use the kernels of the bilinear forms $m(\cdot, \cdot)$ and $m_*(\cdot, \cdot)$, defined to be

$$\mathcal{K} = \left\{ \varphi \in \mathbb{V}; \ \forall \psi \in H_0^1(\Omega), \ m(\psi, \varphi) = 0 \right\},$$

$$\mathcal{K}_* = \left\{ \varphi \in \mathbb{V}; \ \forall \psi \in H_0^1(\Omega), \ m_*(\psi, \varphi) = 0 \right\}.$$

Note that $\mathcal{K}$ and $\mathcal{K}_*$ are closed subspaces in $\mathbb{V}$, and are then Hilbert spaces, when endowed with $\| \cdot \|_{\mathbb{V}}$. We now recall the inf-sup conditions on the form $a(\cdot, \cdot)$ from [1, Lemma 2.7], their proof relies on the construction of an isomorphism between the spaces $\mathcal{K}$ and $\mathcal{K}_*$.

Lemma 2.2. *Assume that*

$$osc \, \sqrt{\frac{r_*}{r}} = \max_{\boldsymbol{x} \in \overline{\Omega}} \sqrt{\frac{r_*}{r}(\boldsymbol{x})} - \min_{\boldsymbol{x} \in \overline{\Omega}} \sqrt{\frac{r_*}{r}(\boldsymbol{x})} < 2. \tag{2.5}$$

Then, the bilinear form $a(\cdot, \cdot)$ satisfies the two inf-sup conditions for a positive constant τ

$$\forall \chi \in \mathcal{K}, \quad \sup_{\varphi \in \mathcal{K}_*} \frac{a(\chi, \varphi)}{\|\varphi\|_{\mathbb{V}}} \geq \tau \, \|\chi\|_{\mathbb{V}},$$

$$\forall \varphi \in \mathcal{K}_*, \quad \sup_{\chi \in \mathcal{K}} \frac{a(\chi, \varphi)}{\|\chi\|_{\mathbb{V}}} \geq \tau \, \|\varphi\|_{\mathbb{V}}. \tag{2.6}$$

We are hence in a position to state the main result of this section, which is straightforwardly derived from [10, Thm 3.1] and the generalized inf-sup conditions stated in Lemmas 2.1 and 2.2.

Theorem 2.3. *Assume that (2.5) holds true. Then, for any data (f, g) in $H^{-1}(\Omega) \times L^2(\Omega)$, the mixed problem (2.2) has a unique solution (b, c) in $\mathbb{V} \times H_0^1(\Omega)$. Moreover this solution satisfies*

$$\|b\|_{\mathbb{V}} + \|c\|_{H^1(\Omega)} \leq C(\|f\|_{H^{-1}(\Omega)} + \|g\|_{L^2(\Omega)}). \tag{2.7}$$

Remark 2.4. Some further regularity of the solution can easily be derived from [9, Cor. 3.5.22]. For instance, if the domain Ω is a convex polygon or polyhedron, for any data (f, g) in $H^{-1}(\Omega) \times H^1(\Omega)$, the solution (b, c) of problem (2.2) belongs to $H^1(\Omega) \times H^3(\Omega)$.

2.3. Nonhomogeneous boundary conditions.

Enforcing the nonhomogeneous boundary conditions on the boundary $\partial\Omega$ is worth some comments. In view of the functional framework used in our analysis, prescribing the Dirichlet condition $c = \alpha$ seems natural in $H^{1/2}(\partial\Omega)$ while imposing the Neumann one $d\,\partial_n c = \beta$ arises some trouble. We start from the following variational formulation

Find (b, c) in $\mathbb{V} \times H^1(\Omega)$ satisfying $c_{|\partial\Omega} = \alpha$ and such that

$$\begin{aligned}
\forall \psi \in H_0^1(\Omega), \qquad & m\,(\psi, b) = \langle f, \psi \rangle_{H^{-1}, H_0^1}, \\
\forall \varphi \in \mathbb{V}, \qquad & m_*(c, \varphi) + a(b, \varphi) = \int_\Omega g(\boldsymbol{x})\varphi(\boldsymbol{x})\,d\boldsymbol{x} - \langle \beta, \varphi \rangle_{\partial\Omega}.
\end{aligned} \tag{2.8}$$

The key point is to give a sense to the duality product $\langle \beta, \varphi \rangle_{\partial\Omega}$, which seems meaningless for all φ in $\mathbb{V}$. We recall from [9, Section 1.5] that, if the boundary $\partial\Omega$ is smooth, any φ in $\mathbb{V}$ has a trace $\varphi_{|\partial\Omega}$ that belongs to $H^{-1/2}(\partial\Omega)$. As a result for $\langle \beta, \varphi \rangle_{\partial\Omega}$ to make sense, it is necessary to assume that β lies in $H^{1/2}(\partial\Omega)$. This regularity seems too strong to consider on a Neumann data. This difficulty can be overcome by using an appropriate lifting, see [1, Thm 2.13], and we have therefore the following result.

Theorem 2.5. *Assume that (2.5) holds true. Then, for any data (f, g) in $H^{-1}(\Omega) \times L^2(\Omega)$ and (α, β) in $H^{1/2}(\partial\Omega) \times H^{-1/2}(\partial\Omega)$, the boundary value system (2.8) has a unique solution (b, c) in $\mathbb{V} \times H^1(\Omega)$.*

3. The discrete problem and its well-posedness.

We first describe the construction of the discrete problem. Next, we prove the analogous lemmas as in Section 2 and we derive its well-posedness.

3.1. The discrete problem.

From now on, we work with the domain Ω equal to the square $]-1,1[^2$ or to the cube $]-1,1[^3$. For each nonnegative integer n, we denote by $\mathbb{P}_n(\Omega)$ the space of polynomials with 2 or 3 variables and degree $\leq n$ with respect to each variable. Next, we fix an integer $N \geq 2$ and define the variational spaces

$$\mathbb{V}_N = \mathbb{P}_N(\Omega), \qquad \mathbb{X}_N = \mathbb{P}_N(\Omega) \cap H_0^1(\Omega). \tag{3.1}$$

Despite the complex definition of $\mathbb{V}$, they lead to a conforming discretization.

We also make use of the Gauss-Lobatto formula on the interval $]-1,1[$. Let $\mathbb{P}_n(-1,1)$ denote the space of restrictions to $]-1,1[$ of polynomials with degree $\leq n$; setting: $\xi_0 = -1$ and $\xi_N = 1$, we recall that there exist $N-1$ nodes ξ_j, $1 \leq j \leq N-1$, in $]-1,1[$, with $\xi_0 < \xi_1 < \ldots < \xi_N$, and $N+1$ positive weights ρ_j, $0 \leq j \leq N$, such that

$$\forall \Phi \in \mathbb{P}_{2N-1}(-1,1), \quad \int_{-1}^1 \Phi(\zeta)\, d\zeta = \sum_{j=0}^N \Phi(\xi_j)\, \rho_j. \tag{3.2}$$

Moreover the following property holds [4, form. (13.20)]

$$\forall \varphi \in \mathbb{P}_N(-1,1), \quad \|\varphi\|_{L^2(-1,1)}^2 \leq \sum_{j=0}^N \varphi^2(\xi_j)\, \rho_j \leq 3\, \|\varphi\|_{L^2(-1,1)}^2. \tag{3.3}$$

This leads to define a discrete product on Ω: For any continuous functions u and v on $\overline{\Omega}$,

$$(u,v)_N = \begin{cases} \sum_{i=0}^N \sum_{j=0}^N u(\xi_i,\xi_j)v(\xi_i,\xi_j)\, \rho_i\rho_j & \text{in dimension 2,} \\ \sum_{i=0}^N \sum_{j=0}^N \sum_{k=0}^N u(\xi_i,\xi_j,\xi_k)v(\xi_i,\xi_j,\xi_k)\, \rho_i\rho_j\rho_k & \text{in dimension 3.} \end{cases} \tag{3.4}$$

It follows from (3.3) that this product is a scalar product on $\mathbb{P}_N(\Omega)$. Let also $\mathcal{I}_N$ be the Lagrange interpolation operator on the nodes (ξ_i,ξ_j) or (ξ_i,ξ_j,ξ_k) with values in $\mathbb{P}_N(\Omega)$.

We can now define the bilinear forms needed for the discrete problem:

$$\forall(\chi_N,\varphi_N) \in \mathbb{V}_N \times \mathbb{V}_N, \qquad a_N(\chi_N,\varphi_N) = (r\,\chi_N,\varphi_N)_N,$$
$$\forall(\psi_N,\varphi_N) \in \mathbb{X}_N \times \mathbb{V}_N, \qquad m_N(\psi_N,\varphi_N) = (d\,\nabla\varphi_N,\nabla\psi_N)_N + (r\,\varphi_N,\psi_N)_N,$$
$$\forall(\psi_N,\varphi_N) \in \mathbb{X}_N \times \mathbb{V}_N, \qquad m_{*N}(\psi_N,\varphi_N) = (d\,\nabla\varphi_N,\nabla\psi_N)_N + (r_*\,\varphi_N,\psi_N)_N.$$

Note from (3.3) that all these forms are continuous when the spaces $\mathbb{V}_N$ and $\mathbb{X}_N$ are both provided with the norm of $H^1(\Omega)$ and also that replacing the coefficients d, r and r_* by their Lagrange interpolates does not change the previous definition.

The discrete problem is now constructed from problem (2.2) by the Gakerkin method with numerical integration. In the case $\alpha = \beta = 0$ of homogeneous boundary conditions and for continuous data f and g on $\overline{\Omega}$, it reads

Find (b_N, c_N) in $\mathbb{V}_N \times \mathbb{X}_N$ such that

$$
\begin{aligned}
\forall \psi_N \in \mathbb{X}_N, \qquad & m_N(\psi_N, b_N) = (f, \psi_N)_N, \\
\forall \varphi \in \mathbb{V}_N, \qquad & m_{*N}(c_N, \varphi_N) + a_N(b_N, \varphi_N) = (g, \varphi_N)_N.
\end{aligned}
\tag{3.5}
$$

Obviously this problem results into a square linear system.

3.2. Wellposedness.

Even if the well-posedness of problem (3.5) can be derived in a simpler way, we have decided to use the same arguments as for the continuous problem for that, namely to prove inf-sup conditions on the forms $a_N(\cdot, \cdot)$, $m_N(\cdot, \cdot)$ and $m_{*N}(\cdot, \cdot)$.

Lemma 3.1. *The bilinear form $m_N(\cdot, \cdot)$ satisfies the inf-sup condition for a positive constant η' independent of N*

$$
\forall \psi_N \in \mathbb{X}_N, \quad \sup_{\varphi_N \in \mathbb{V}_N} \frac{m_N(\psi_N, \varphi_N)}{\|\varphi_N\|_{H^1(\Omega)}} \geq \eta' \, \|\psi_N\|_{H^1(\Omega)}.
\tag{3.6}
$$

The bilinear form $m_*(\cdot, \cdot)$ satisfies the inf-sup condition for a positive constant η'_* independent of N

$$
\forall \psi_N \in \mathbb{X}_N, \quad \sup_{\varphi_N \in \mathbb{V}_N} \frac{m_{*N}(\psi_N, \varphi_N)}{\|\varphi_N\|_{H^1(\Omega)}} \geq \eta'_* \, \|\psi_N\|_{H^1(\Omega)}.
\tag{3.7}
$$

Proof. Since the proofs of the two conditions are exactly the same, we only check the first one. Using the embedding of $\mathbb{X}_N$ into $\mathbb{V}_N$, we take φ_N equal to ψ_N, which gives

$$
m_N(\psi_N, \varphi_N) = (d\, \nabla \psi_N, \nabla \psi_N)_N + (r\, \psi_N, \psi_N)_N.
$$

It thus follows from the positivity of d and r combined with (3.3) that

$$
\sup_{\varphi_N \in \mathbb{V}_N} m_N(\psi_N, \varphi_N) \geq \min\{d_\flat, r_\flat\} \, \|\psi_N\|^2_{H^1(\Omega)} = \min\{d_\flat, r_\flat\} \, \|\psi_N\|_{H^1(\Omega)} \|\varphi_N\|_{H^1(\Omega)}.
$$

This gives the desired inf-sup condition.

Unfortunately, the conditions on $a_N(\cdot, \cdot)$ turn out to be a little more complex. They require the discrete kernels

$$\mathcal{K}_N = \Big\{ \varphi_N \in \mathbb{V}_N; \ \forall \psi_N \in \mathbb{X}_N, \ m_N(\psi_N, \varphi_N) = 0 \Big\},$$

$$\mathcal{K}_{*N} = \Big\{ \varphi_N \in \mathbb{V}_N; \ \forall \psi_N \in \mathbb{X}_N, \ m_{*N}(\psi_N, \varphi_N) = 0 \Big\}.$$

There is no reason for them to be imbedded in the continuous kernels $\mathcal{K}$ and $\mathcal{K}_*$.

Next, with any χ_N in $\mathcal{K}_N$, we associate the solution θ_N in $\mathbb{X}_N$ of the following problem

$$\forall \psi_N \in \mathbb{X}_N, \quad (d \nabla \theta_N, \nabla \psi_N)_N + (r_* \theta_N, \psi_N)_N = \big((r - r_*)\chi_N, \psi_N\big)_N. \tag{3.8}$$

Obviously, thanks to (3.3) and the Lax–Milgram lemma, this problem has a unique solution. Then, we define an operator $\mathcal{M}_N$ on $\mathcal{K}_N$ by the formula

$$\mathcal{M}_N \chi_N = \chi_N + \theta_N.$$

We now give some properties of this operator.

Lemma 3.2. *The operator $\mathcal{M}_N$ is one-to-one from $\mathcal{K}_N$ onto $\mathcal{K}_{*N}$. Moreover, the following stability properties hold for positive constants $\sigma_\flat$ and $\sigma_\sharp$ independent of N*

$$\forall \chi_N \in \mathcal{K}_N, \quad \sigma_\flat \, \|\chi_N\|_{L^2(\Omega)} \leq \|\mathcal{M}_N \chi_N\|_{L^2(\Omega)} \leq \sigma_\sharp \, \|\chi_N\|_{L^2(\Omega)}. \tag{3.9}$$

Proof. The fact that $\mathcal{M}_N$ maps $\mathcal{K}_N$ into $\mathcal{K}_{*N}$ follows from its definition and the fact that it is one-to-one follows from estimate (3.9) that we now prove. By taking ψ_N equal to θ_N in (3.8), we derive from (3.3)

$$\|\theta_N\|_{L^2(\Omega)} \leq c \, \|\chi_N\|_{L^2(\Omega)},$$

whence the second estimate in (3.9). The proof of the first estimate follows from the same lines by using the formula $\chi_N = \mathcal{M}_N \chi_N - \theta_N$ and the fact that problem (3.8) can equivalently be written

$$\forall \psi_N \in \mathbb{X}_N, \quad (d \nabla \theta_N, \nabla \psi_N)_N + (r \, \theta_N, \psi_N)_N = \big((r - r_*)\mathcal{M}_N \chi_N, \psi_N\big)_N.$$

We are now in a position to state and prove the next lemma. Its rather technical proof relies on the same arguments as for [1, Lemma 2.7].

Lemma 3.3. *Assume that condition (2.5) holds. Then, the bilinear form $a(\cdot, \cdot)$ satisfies the two inf-sup conditions for a positive constant τ'*

$$\forall \chi_N \in \mathcal{K}_N, \quad \sup_{\varphi_N \in \mathcal{K}_{*N}} \frac{a_N(\chi_N, \varphi_N)}{\|\varphi_N\|_{L^2(\Omega)}} \geq \tau' \, \|\chi_N\|_{L^2(\Omega)},$$

$$\forall \varphi_N \in \mathcal{K}_{*N}, \quad \sup_{\chi_N \in \mathcal{K}_N} \frac{a_N(\chi_N, \varphi_N)}{\|\chi_N\|_{L^2(\Omega)}} \geq \tau' \, \|\varphi_N\|_{L^2(\Omega)}. \tag{3.10}$$

Proof. Owing to Lemma 3.2, with each χ_N in $\mathcal{K}_N$, we associate the function $\varphi_N = \mathcal{M}_N \chi_N$ in $\mathcal{K}_{*N}$ and with each φ_N in $\mathcal{K}_{*N}$, we associate the function $\chi_N = \mathcal{M}_N^{-1} \varphi_N$ in $\mathcal{K}_N$, so that both inf-sup conditions rely on the evaluation of the quantity $a_N(\chi_N, \varphi_N)$. Since the function $\theta_N = \varphi_N - \chi_N$ belongs to $\mathbb{X}_N$ and satisfies

$$\forall \psi_N \in \mathbb{X}_N, \quad (d\,\nabla\theta_N, \nabla\psi_N)_N + (r_* \,\varphi_N, \psi_N)_N - (r\chi_N, \psi_N)_N = 0,$$

we obtain

$$(d\,\nabla\theta_N, \nabla\theta_N)_N + (r_* \,\varphi_N, \varphi_N)_N + (r\,\chi_N, \chi_N)_N = (r\chi_N, \varphi_N)_N + (r_*\chi_N, \varphi_N)_N,$$

whence, for any positive real number ξ,

$$(d\,\nabla\theta_N, \nabla\theta_N)_N + (r_* \,\varphi_N, \varphi_N)_N + (r\chi_N, \chi_N)_N = (1+\xi)a_N(\chi_N, \varphi_N) + \big((r_* - \xi r)\chi_N, \varphi_N\big)_N.$$

Using the fact that $(\cdot, \cdot)_N$ is a product together with the formula $ab \le a^2 + \frac{b^2}{4}$, we derive

$$(r_* \,\varphi_N, \varphi_N)_N + (r\,\chi_N, \chi_N)_N \le (1+\xi)\,a_N(\chi_N, \varphi_N) + (r_* \,\varphi_N, \varphi_N)_N + \Big(\frac{(r_* - \xi r)^2}{4r_*}\,\chi_N, \chi_N\Big)_N,$$

whence

$$(1+\xi)\,a_N(\chi_N, \varphi_N) \ge \Big(\big(1 - \frac{(r_* - \xi r)^2}{4rr_*}\big)r\,\chi_N, \chi_N\Big)_N.$$

Now, assumption(2.5) is equivalent, see [1, Lemma 2.6], to the existence of a $\zeta > 0$ such that, for an appropriate value of ξ,

$$\min_{x \in \overline{\Omega}}\big(1 - \frac{(r_* - \xi r)^2}{4rr_*}\big) = \zeta.$$

Combining all this with (3.3) and Lemma 3.2 leads to

$$(1+\xi)\,a_N(\chi_N, \varphi_N) \ge r_\flat \zeta\,(\chi_N, \chi_N)_N \ge \frac{r_\flat \zeta}{\sigma_\sharp}\,\|\chi_N\|_{L^2(\Omega)}\|\varphi_N\|_{L^2(\Omega)},$$

whence the desired inf-sup conditions.

Even if the norms in Lemma 3.1 and 3.3 are not coherent, using the equivalence of norms on the finite-dimensional spaces $\mathbb{V}_N$ and $\mathbb{X}_N$ allows us to apply [10, Thm 3.1].

Theorem 3.4. *Assume that (2.5) holds true. Then, for any data (f, g) continuous on $\overline{\Omega}$, the discrete problem (3.5) has a unique solution (b_N, c_N) in $\mathbb{V}_N \times \mathbb{X}_N$.*

3.3. Nonhomogeneous boundary conditions.

We now define a discrete scalar product on $\partial\Omega$ as follows: On each edge (in dimension 2) or face (in dimension 3) of Ω, a discrete product is defined as in (3.4); then, the discrete product $(\cdot,\cdot)_N^{\partial\Omega}$ is defined by summing these local products on the 4 edges or 6 faces of Ω. We also introduce an interpolation operator $\mathcal{I}_N^{\partial\Omega}$: On each edge or face Γ of Ω, it coincides with the interpolation operator on the nodes (ξ_i, ξ_j) or (ξ_i, ξ_j, ξ_k) which belongs to $\overline{\Gamma}$ with values in $\mathbb{P}_N(\Gamma)$. The discrete problem is constructed in an obvious way:

Find (b_N, c_N) in $\mathbb{V}_N \times \mathbb{V}_N$ satisfying $c_{N\,|\partial\Omega} = \mathcal{I}_N^{\partial\Omega}\alpha$ and such that

$$
\begin{aligned}
&\forall \psi_N \in \mathbb{X}_N, \Gamma m_N\left(\psi_N, b_N\right) = (f, \psi_N)_N, \\
&forall\varphi \in \mathbb{V}_N, \qquad m_{*N}(c_N, \varphi_N) + a(b_N, \varphi_N) = (g, \varphi_N)_N - (\beta, \varphi_N)_N^{\partial\Omega}.
\end{aligned}
\tag{3.11}
$$

When all assumptions of Theorem 3.4 are satisfied and if moreover the data (α, β) are continuous on $\partial\Omega$, this problem has a unique solution.

4. A priori error analysis.

We first intend to bound $\|b - b_N\|_{L^2(\Omega)}$. For this, we introduce the affine subspace

$$\mathcal{K}_N(f) = \left\{ \varphi_N \in \mathbb{V}_N; \ \forall \psi_N \in \mathbb{X}_N, \ m_N\left(\psi_N, \varphi_N\right) = (f, \psi_N)_N \right\}. \tag{4.1}$$

When taking β_N in $\mathcal{K}_N(f)$, we observe that $b_N - \beta_N$ belongs to $\mathcal{K}_N$. So the idea is to apply the first inf-sup condition (3.10).

In all this section, we assume that the coefficients d, r and r_* are constant, only for simplicity of the estimates (otherwise, further terms of type $\|d - \mathcal{I}_N d\|_{L^\infty(\Omega)}$ are involved and when the coefficients are regular, which is most often the case, they can be considered as neglectable).

The next proof requires the orthogonal projection operator Π_{N-1} from $L^2(\Omega)$ onto $\mathbb{P}_{N-1}(\Omega)$.

Lemma 4.1. *Assume that the data (f, g) belong to $H^\sigma(\Omega) \times H^\sigma(\Omega)$, $\sigma > \frac{d}{2}$. Then, the following error estimate holds between the solutions (b, c) of problem (2.2) and (b_N, c_N) of problem (3.5)*

$$\|b - b_N\|_{L^2(\Omega)} \leq C \left(N^2 \inf_{\gamma_{N-1} \in \mathbb{X}_{N-1}} \|c - \gamma_{N-1}\|_{H^1(\Omega)} \right.$$
$$\left. + \inf_{\beta_N \in \mathcal{K}_N(f)} \|b - \beta_N\|_{L^2(\Omega)} + \|b - \Pi_{N-1} b\|_{L^2(\Omega)} + \varepsilon_N^{(D)} \right), \tag{4.2}$$

where the quantity $\varepsilon_N^{(D)}$ is given by

$$\varepsilon_N^{(D)} = N^{-\sigma} \left(\|f\|_{H^\sigma(\Omega)} + \|g\|_{H^\sigma(\Omega)} \right). \tag{4.3}$$

Proof. Let β_N be any polynomial in $\mathcal{K}_N(f)$. It follows from (3.10) that

$$\|b_N - \beta_N\|_{L^2(\Omega)} \leq \tau'^{-1} \sup_{\varphi_N \in \mathcal{K}_{*N}} \frac{a_N\left(b_N - \beta_N, \varphi_N\right)}{\|\varphi_N\|_{L^2(\Omega)}}.$$

By applying problem (3.5), then problem (2.2), we derive that, for any φ_N in $\mathcal{K}_{*N}$,

$$a_N(b_N - \beta_N, \varphi_N) = -a_N(\beta_N, \varphi_N) + (g, \varphi_N)_N$$
$$= m_*(c, \varphi_N) + a(b, \varphi_N) - a_N(\beta_N, \varphi_N) - \int_\Omega g(\boldsymbol{x}) \varphi_N(\boldsymbol{x}) \, d\boldsymbol{x} + (g, \varphi_N)_N.$$

Since φ_N belongs to $\mathcal{K}_{*N}$, we have for any γ_{N-1} in $\mathbb{X}_{N-1}$,

$$m_*(c, \varphi_N) = m_*(c - \gamma_{N-1}, \varphi_N),$$

11

whence, by applying the standard inverse inequality [4, Thm 5.1] to φ_N,

$$m_*(c, \varphi_N) \leq CN^2 \|c - \gamma_{N-1}\|_{H^1(\Omega)} \|\varphi_N\|_{L^2(\Omega)}.$$

On the other hand, we derive by using first (3.2), next (3.3),

$$a(b, \varphi_N) - a_N(\beta_N, \varphi_N) \leq C \left(\|b - \Pi_{N-1}b\|_{L^2(\Omega)} + \|b - \beta_N\|_{L^2(\Omega)} \right) \|\varphi_N\|_{L^2(\Omega)}.$$

Finally, the same arguments yield

$$\left| (g, -\varphi_N)_N - \int_\Omega g(\boldsymbol{x})\varphi_N)(\boldsymbol{x}) \, d\boldsymbol{x} \right|$$
$$= \left| (\mathcal{I}_N g - \Pi_{N-1}g, \varphi_N)_N - \int_\Omega (g - \Pi_{N-1}g)(\boldsymbol{x})\varphi_N(\boldsymbol{x}) \, d\boldsymbol{x} \right|$$
$$\leq \left(27 \|g - \mathcal{I}_N g\|_{L^2(\Omega)} + 28 \|g - \Pi_{N-1}g\|_{L^2(\Omega)} \right) \|\varphi_N\|_{L^2(\Omega)}.$$

Using the standard approximation properties of the operators $\mathcal{I}_N$ and Π_{N-1}, see [4, Thms 7.1 and 14.2], implies that this last term is bounded by $\varepsilon_N^{(D)} \|\varphi_N\|_{L^2(\Omega)}$. Combining all this gives the desired estimate.

We now prove an estimate for $\|c - c_N\|_{H^1(\Omega)}$.

Lemma 4.2. *If the assumptions of Lemma 4.1 are satisfied, the following error estimate holds between the solutions (b, c) of problem (2.2) and (b_N, c_N) of problem (3.5)*

$$\|c - c_N\|_{H^1(\Omega)} \leq C \left(N^2 \inf_{\gamma_{N-1} \in \mathbb{X}_{N-1}} \|c - \gamma_{N-1}\|_{H^1(\Omega)} \right.$$
$$\left. + \inf_{\beta_N \in \mathcal{K}_N(f)} \|b - \beta_N\|_{L^2(\Omega)} + \|b - \Pi_{N-1}b\|_{L^2(\Omega)} + \varepsilon_N^{(D)} \right), \tag{4.4}$$

where $\varepsilon_N^{(D)}$ is defined in (4.3).

Proof. It follows from problems (2.2) and (3.5) that, for any γ_{N-1} in $\mathbb{X}_{N-1}$;

$$m_{*N}(c_N - \gamma_{N-1}, c_N - \gamma_{N-1}) = m_*(c, c_N - \gamma_{N-1}) - m_{*N}(\gamma_{N-1}, c_N - \gamma_{N-1})$$
$$- \int_\Omega g(\boldsymbol{x})(c_N - \gamma_{N-1})(\boldsymbol{x}) \, dx + (g, c_N - \gamma_{N-1})_N + a(b, c_N - \gamma_{N-1}) - a_N(b_N, c_N - \gamma_{N-1}).$$

Thus, exactly the same arguments as in the proof of Lemma 4.1 lead to

$$\|c_N - \gamma_{N-1}\|_{H^1(\Omega)}^2 \leq C \|c - \gamma_{N-1}\|_{H^1(\Omega)} \|c_N - \gamma_{N-1}\|_{H^1(\Omega)}$$
$$+ C' \left(\|b - \Pi_{N-1}b\|_{L^2(\Omega)} + \varepsilon_N^{(D)} + \|b - b_N\|_{L^2(\Omega)} \right) \|c_N - \gamma_{N-1}\|_{L^2(\Omega)}.$$

12

Using estimate (4.2) gives the desired bound.

To make these estimates complete, we must only prove an upper bound for the quantity $\inf_{\beta_N \in \mathcal{K}_N(f)} \|b - \beta_N\|_{L^2(\Omega)}$. The next lemma is just an extension of [8, Chap. II, eq. (1.16)].

Lemma 4.3. *Assume that the part b of the solution (b, c) of problem (2.2) belongs to $H^1(\Omega)$. If the assumptions of Lemma 4.1 are satisfied, the following approximation estimate holds*

$$\inf_{\beta_N \in \mathcal{K}_N(f)} \|b - \beta_N\|_{H^1(\Omega)} \leq C \left(\inf_{\delta_{N-1} \in \mathbb{V}_{N-1}} \|b - \delta_{N-1}\|_{H^1(\Omega)} + \varepsilon_N^{(D)} \right), \tag{4.5}$$

where $\varepsilon_N^{(D)}$ is defined in (4.3).

Proof. Let δ_{N-1} be any function in $\mathbb{V}_{N-1}$. Owing to the inf-sup condition (3.6), there exists a polynomial $\delta_N^\diamond$ in $\mathbb{V}_N$ such that

$$\forall \psi_N \in \mathbb{X}_N, \quad m_N(\psi_N, \delta_N^\diamond) = (f, \psi_N)_N - m_N(\psi_N, \delta_{N-1}),$$

and which satisfies [8, Chap. I, Lemma 4.1]

$$\|\delta_N^\diamond\|_{H^1(\Omega)} \leq \eta'^{-1} \sup_{\psi_N \in \mathbb{X}_N} \frac{(f, \psi_N)_N - m_N(\psi_N, \delta_{N-1}),}{\|\psi_N\|_{H^1(\Omega)}}. \tag{4.6}$$

Obviously, the polynomial $\beta_N = \delta_{N-1} + \delta_N^\diamond$ belongs to $\mathcal{K}_N(f)$. Moreover, we have

$$\|b - \beta_N\|_{H^1(\Omega)} \leq \|b - \delta_{N-1}\|_{H^1(\Omega)} + \|\delta_N^\diamond\|_{H^1(\Omega)}.$$

To evaluate this last quantity, we observe from the first line of problem (2.2) that

$$(f, \psi_N)_N - m_N(\psi_N, \delta_{N-1}) = (f, \psi_N)_N - \int_\Omega f(\boldsymbol{x}) \psi_N(\boldsymbol{x}) \, d\boldsymbol{x} + m(\psi_N, b - \delta_{N-1}).$$

Inserting this into (4.6) gives the desired estimate.

Remark 4.4. Since the domain Ω is convex, the assumptions on the data imply that b belongs to $H^1(\Omega)$, see Remark 2.4.

Even if the estimate for $b - \beta_N$ is not in the right norm, we are now in a position to state the final error estimate which follows from Lemmas 4.1 to 4.3 together with standard approximation properties.

Theorem 4.5. *Assume that the data (f, g) belong to $H^\sigma(\Omega) \times H^\sigma(\Omega)$, $\sigma > \frac{d}{2}$, and that the solution (b, c) of problem (2.2) belongs to $H^{s-1}(\Omega) \times H^{s+1}(\Omega)$, $s \geq 2$. Then, the following error estimate holds between this solution and the solution (b_N, c_N) of problem (3.5)*

$$\|b - b_N\|_{L^2(\Omega)} + \|c - c_N\|_{H^1(\Omega)}$$
$$\leq C \left(N^{2-s} \left(\|b\|_{H^{s-1}(\Omega)} + \|c\|_{H^{s+1}(\Omega)} \right) + N^{-\sigma} \left(\|f\|_{H^\sigma(\Omega)} + \|g\|_{H^\sigma(\Omega)} \right) \right). \tag{4.7}$$

Even if this estimate is not fully optimal, it proves the good properties of our discretization.

5. Numerical experiments.

We report in this section several numerical tests, the aimi being to evaluate the performance of the spectral discretization in two and three space dimensions. All the computations have been performed on the code FreeFEM3D, spectral version, developped during the thesis of D. Yakoubi [14], in collaboration with S. Del Pino [6]. Note that, in all simulations, the polynomial degree in the directions x, y and z are the same.

5.1. Accuracy test.

In this test, the rate of convergence with respect to polynomial degree N for the couple (b, c) in the L^2-norm and H^1-norm have been tested numerically on the square

$$\Omega =] - \frac{1}{2\pi}, -\frac{1}{2\pi} + 1[^2.$$

A two-dimensional analytic and very smooth solution of problem (1.1) is given by, with the appropriate source terms f and g,

$$b(x, y) = \cos(\frac{3\pi}{2}x) \cos(\frac{3\pi}{2}y), \qquad c(x, y) = \sin(\frac{3\pi}{2}x) \sin(\frac{3\pi}{2}y). \qquad (5.1)$$

The dispersion and reaction parameters are variable and depend on N:

$$d(x, y) = 0.1 \cos\left(\pi(x^2 + y^2)\right) + 0.2,$$
$$r(x, y) = x^N + 3y^N + 4 \qquad \text{and} \qquad r_*(x, y) = 2x^N - y^N + 10.cr$$

Numerically, these parameters are replaced by their Lagrange interpolates.

We recall that this situation is already investigated in [1, Section 5] when the physical parameters are constant but seem close to real-life values.

We illustrate the behavior of the error between the exact solution (b, c) and the discrete solution (b_N, c_N) versus the polynomial degree N which is varying between $N = 5$ and $N = 30$. Figure 1 presents this error in the norm of $L^2(\Omega)$, in plain red line for b and dashed blue line for c, in semi-logarithmic scales and as a function of N. Figure 2 is similar but presents the error in the norm of $H^1(\Omega)$. In this simple geometry and for a smooth solution, the error is of course of spectral type.
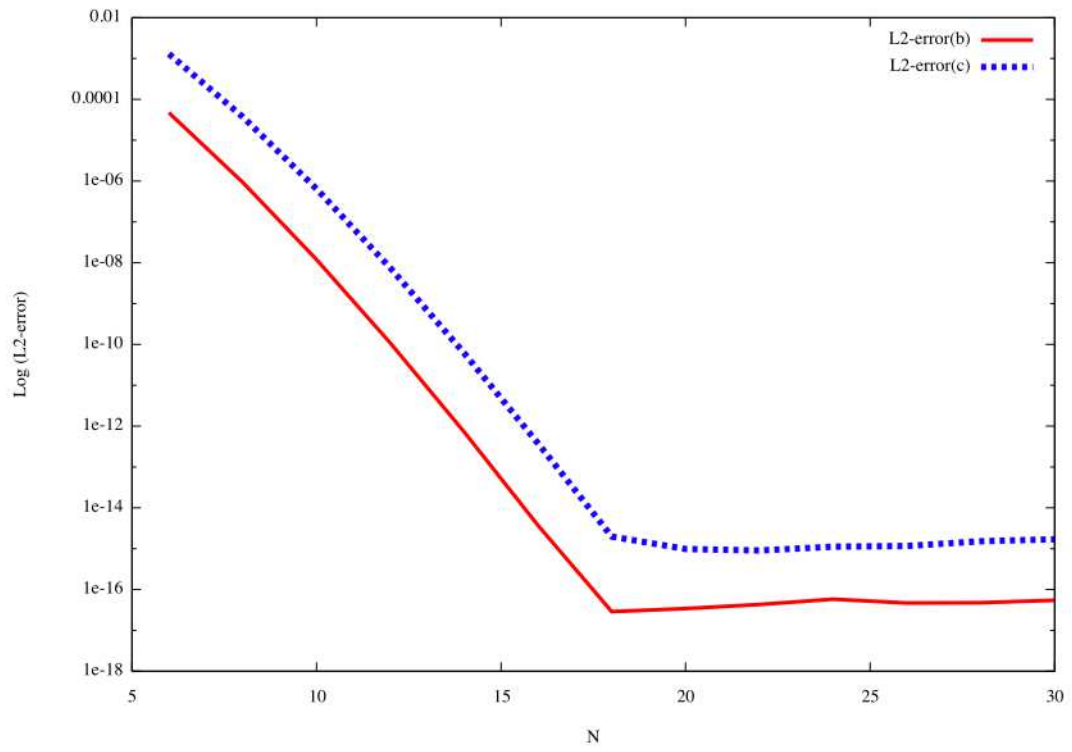
**Figure 1.** The L^2-errors on b and c as a function of N
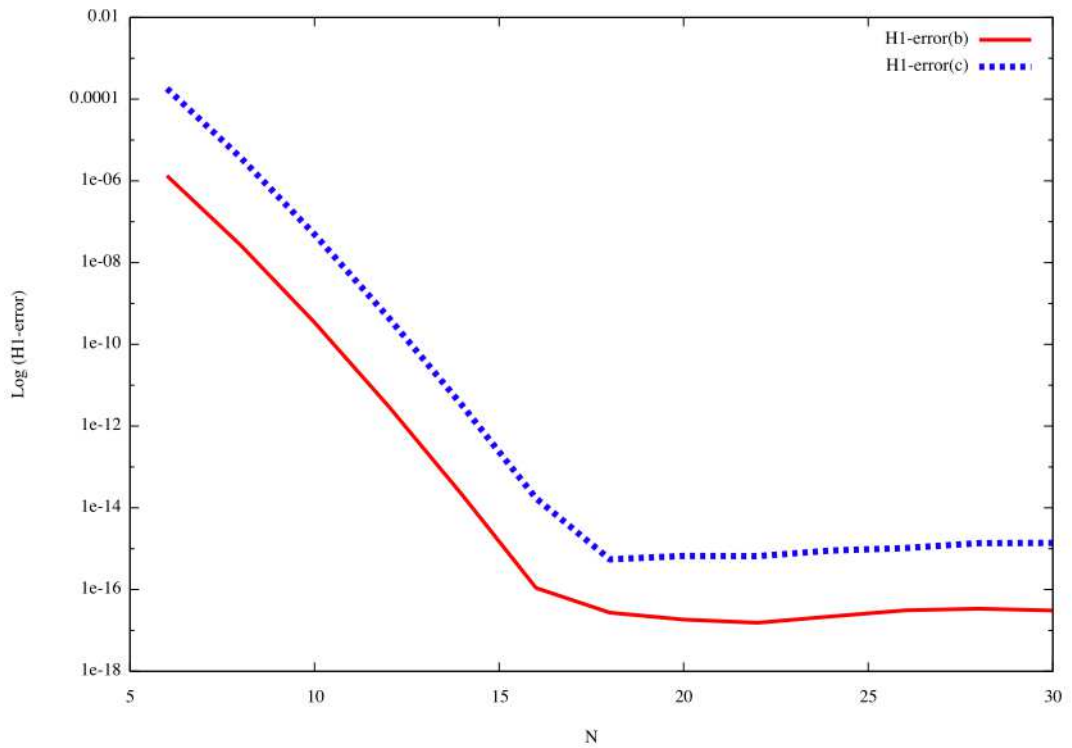
Figure 2. The H^1-errors on b and c as a function of N

5.2. In an ellipse.

Now, we show the numerical simulation in non-parallelipedic domains obtained following the spectral method in complex geometries proposed and analyzed in the thesis of D. Yakoubi [14], see also [6]. In fact, this method uses a fictitious domain method combined with a Nitsche penalization approach to enforce the boundary condition, see for instance Girault and Glowinski [7], in order to allow the use of global high order polynomials over the domain of interest. The numerical analysis of the method relies on the stability of an extension operator from the domain of interest to a larger Cartesian domain, as well as on the estimate on the penalization error. Finally, in order to compute integrals of polynomials over general domains, an approximate internal covering of the domain by simple parallelipedic domains is considered, where to use the standard quadrature formulas. So, the computational domain we consider in this section is the interior of the ellipse defined by:

$$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = 1, \qquad \alpha = 0.2 \quad \text{and} \quad \beta = 0.1.$$
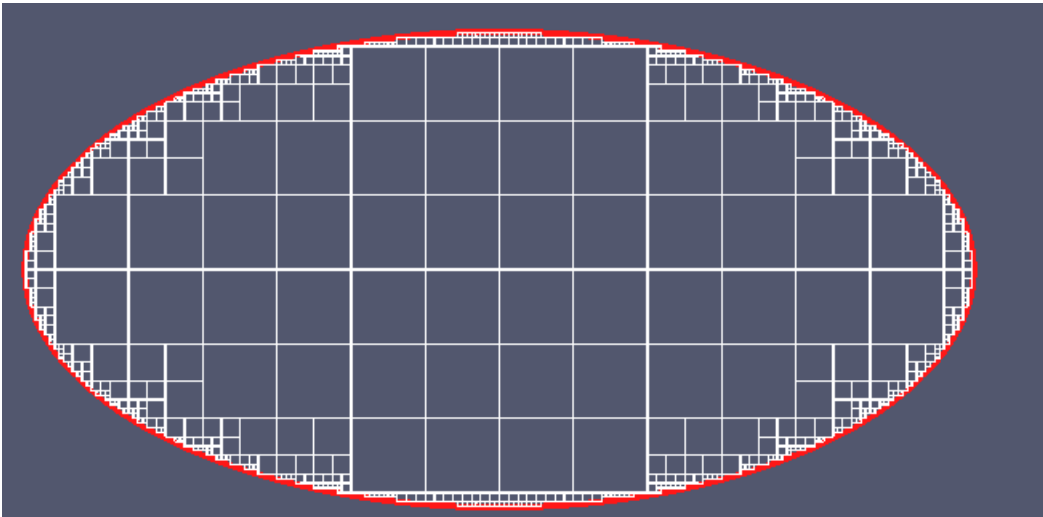
**Figure 3.** Internal covering of the ellipse

Figure 3. shows the approximation of the ellipse by the union of the parallelipedic domains. As proposed in [1, Section 5], the dispersion and reaction parameters are constant and fixed to

$$d = 0.151, \qquad r = 0.2, \qquad r_* = 0.4. \tag{5.2}$$

We conserve the exact solution cited above.

We display in the left panel of Figure 4 the exact solution b and in the right its approximation, while in Figure 5 the exact solution c and its approximation are presented.

This spectral solution is issued from the discretization for $N = 15$. Clearly, as standard in spectral methods, both the exact solution and discrete solution are very similar.
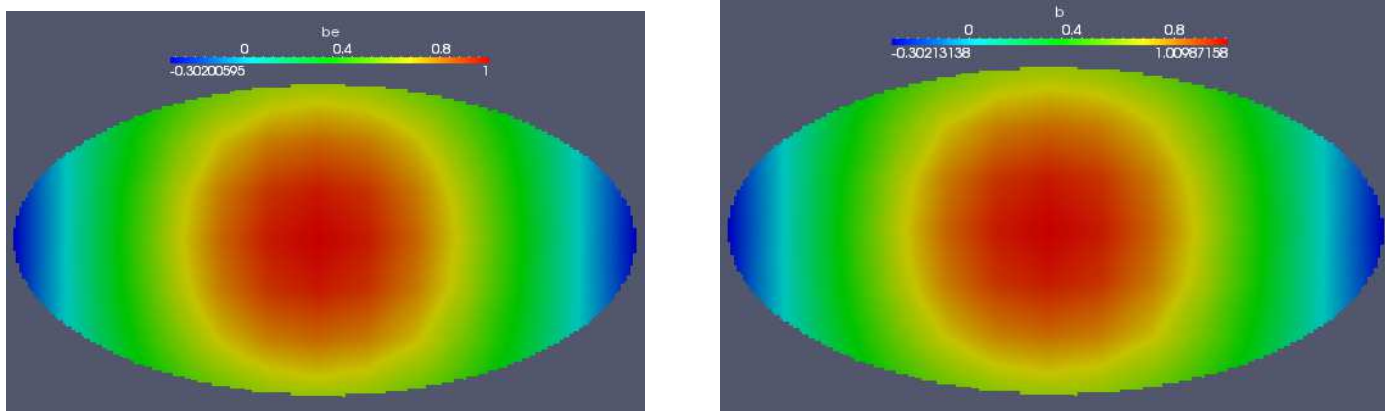
Figure 4. The exact solution b (left part) and the discrete solution b_N (right part)
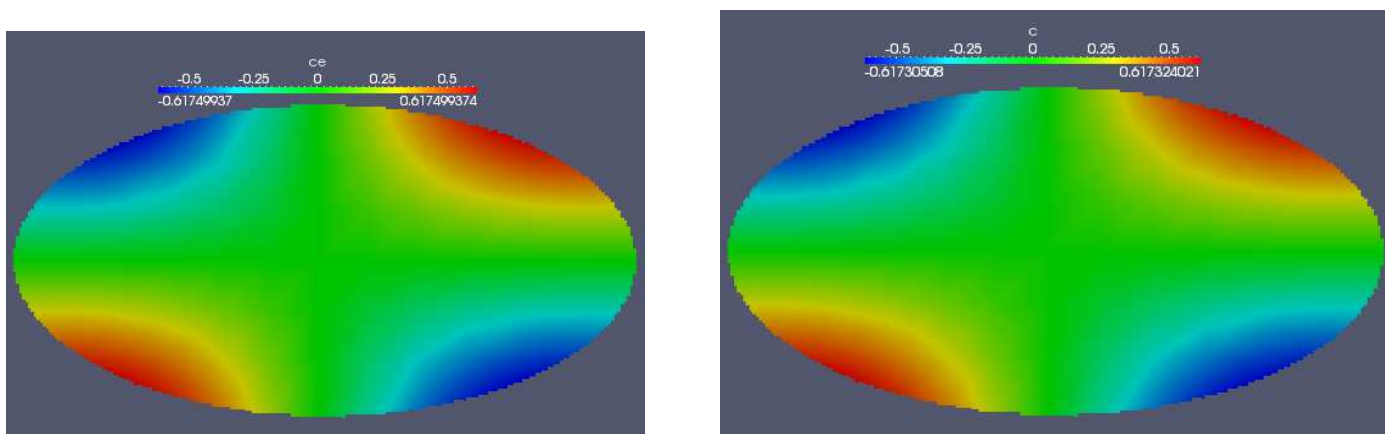
Figure 5. The exact solution c (left part) and the discrete solution c_N (right part)

5.3. In a cube.

Finally, we conclude with an three dimensional experiment in the cube

$$\Omega =] - \frac{1}{2\pi}, -\frac{1}{2\pi} + 1[^3.$$

The exact solution is builded from the above solution, as follows

$$b(x, y, z) = \cos(\pi x)\cos(\pi y)\cos(\pi z), \qquad c(x, y, z) = \sin(\pi x)\sin(\pi y)\sin(\pi z). \qquad (5.3)$$

17

The physical parameters d, r and r_* are taken constant and fixed as (5.2). Figure 6 presents the error $\|c - c_N\|_{L^\infty(\Omega)}$ projected on the oblique plan, which is of order 10^{-13}, while the Figure 7 shows the isolines of exact solution in the left panel and its equivalent of discrete solution in the right panel.
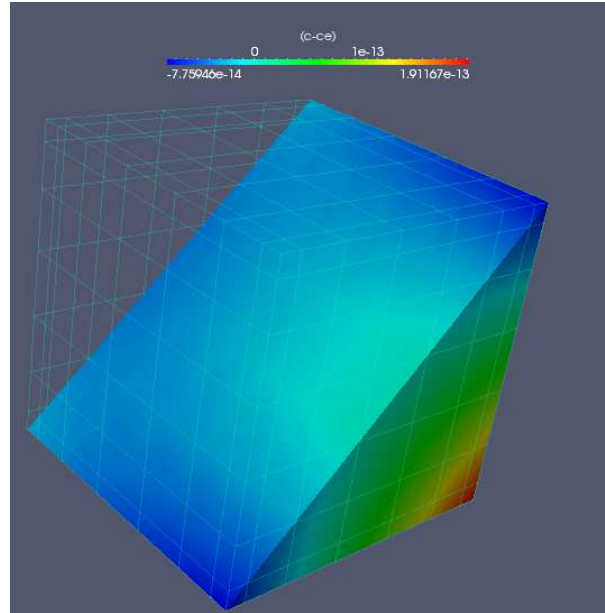
Figure 6. The error between c and c_N
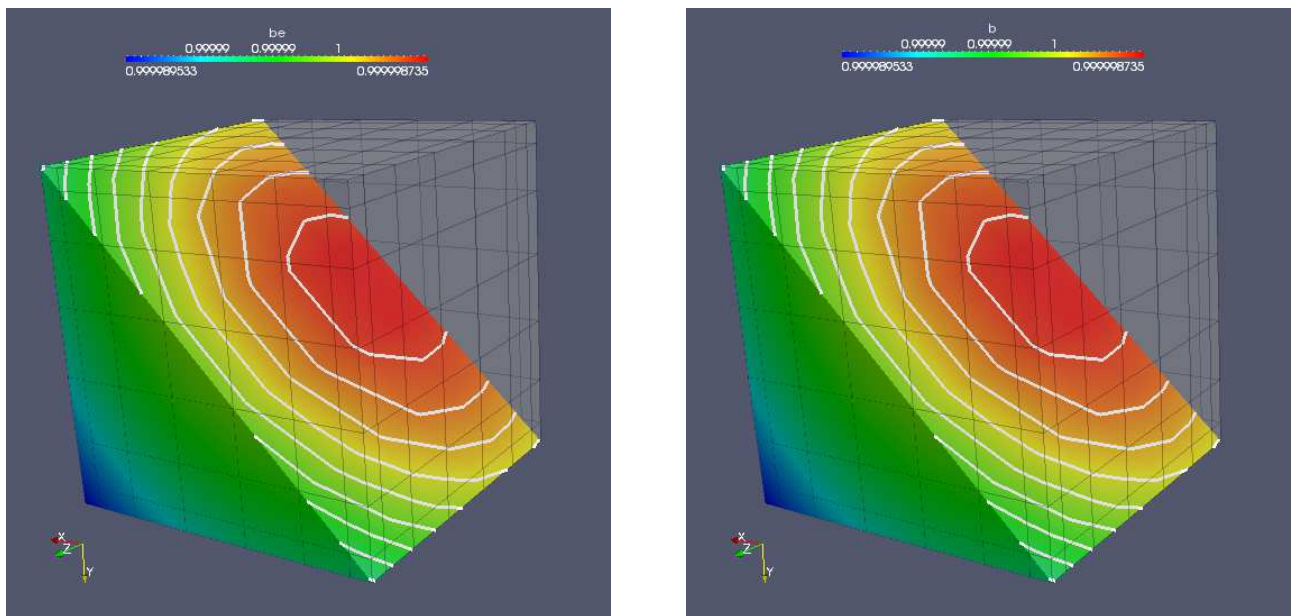
Figure 7. The eaxct b (left part) and the discret b_N (right part)

References

[1] F. Ben Belgacem, C. Bernardi, F. Hecht, S. Salmon — Stabilized finite elements for a reaction-dispersion saddle-point problem with non-constant coefficients, to appear in *SIAM J. Numer. Anal.*

[2] C. Bernardi, C. Canuto, Y. Maday — Generalized inf-sup condition for Chebyshev spectral approximation of the Stokes problem, *SIAM J. Numer. Anal.* **25** (1988), 1237–1271.

[3] C. Bernardi, V. Girault, Y. Maday — Mixed spectral element approximation of the Navier–Stokes equations in the stream-function and vorticity formulation, *IMA J. Numer. Anal.* **12** (1992), 565–608.

[4] C. Bernardi, Y. Maday — *Spectral Methods*, in the *Handbook of Numerical Analysis* **V**, P.G. Ciarlet and J.-L. Lions eds., North-Holland (1997), 209–485.

[5] L.C. Brown, T.O. Barnwell — *The Enhanced Stream Water Quality Models QUAL2E and QUAL2E-UNCAS: Documentation and User Manual*, Environmental Protection Agency (1987).

[6] S. Del Pino, D. Yakoubi — A fictitious-domain like spectral method in complex geometries, in preparation.

[7] V. Girault, R. Glowinski — Error analysis of a fictitious domain method applied to a Dirichlet problem, *Japan J. of Industrial and Applied Math.* **12** (1995), 487–514.

[8] V. Girault, P.-A. Raviart — *Finite Element Methods for Navier–Stokes Equations, Theory and Algorithms*, Springer–Verlag (1986).

[9] P. Grisvard — *Singularities in Boundary Value Problems*, RMA 22, Masson (1982).

[10] R.A. Nicolaides — Existence, uniqueness and approximation for generalized saddle point problems, *SIAM J. Numer. Anal.* **19** (1982), 349–357.

[11] A. Okubo — *Diffusion and Ecological Problems: Mathematical Models*, Springer-Verlag (1980).

[12] C.N. Sawyer, P.L. McCarty, G.F. Parkin — *Chemistry for Environmental Engineering and Science*, McGraw-Hill (2003).

[13] H.W. Streeter, E.B. Phelps — Study of the Pollution and Natural Purification of the Ohio River, *US Public Health Bull.* **146** (1925).

[14] D. Yakoubi — Analyse et mise en œuvre de nouveaux algorithmes en méthodes spectrales, PhD Thesis, Université Pierre et Marie Curie (2007).